

A comparison of bioinformatic workflows for herbivore dietary analysis using a DNA metabarcoding approach

D Botha

 **orcid.org/0000-0002-7321-5382**

Dissertation accepted in fulfilment of the requirements for the
degree *Master of Science in Environmental Sciences* at the North-
West University

Supervisor:	Prof S Barnard
Co-supervisor:	Prof F Siebert
Assistant Supervisor:	Dr MG Du Plessis

Graduation May 2023

27022560

Declaration

This dissertation is being submitted for the degree Master of Science in Environmental Sciences at the North-West University. I declare that the work presented in this MSc dissertation is my own and has not been submitted for any degree or examination at any other university.

D. Botha

24/11/2022

Dedication

To my Heavenly Father, thank you for this great adventure!

“With man this is impossible, but not with God; all things are possible with God” - Mark 10:27

I further dedicate this dissertation to my parents, Shawn and Jolinde Botha, without your love, support, and constant belief in my abilities I would not have been able to achieve this great academic milestone.

Acknowledgements

I would like to single out a few individuals and organisations without whom the completion of this dissertation would not have been possible.

First and foremost, this endeavour would not have been possible without my supervisor, Prof Sandra Barnard. She has been with me since the beginning of my post graduate studies. I have learned so much from her, not just in my role as an academic researcher but her patience, hard-working spirit and calm nature will always inspire me. I am grateful for the opportunity to have learned from someone so knowledgeable and I will always cherish my time as her student. I would also like to thank my co-supervisor, Prof Frances Siebert, for her positivity and motivation during my studies. I immensely appreciate her input in my studies and for providing an environment that has motivated me to focus on my work and be part of an inspiring team of students. I am also extremely grateful to Dr Morné du Plessis for teaching me the ways of technical writing when it comes to bioinformatics and for his assistance with the various analyses. I learned a lot from him, and I am grateful that he could have been part of my academic journey.

Next, I would like to thank the following institutions for their contributions to this study. The South African Limpopo Landscapes Network (SALLNet), an international (Germany-South Africa), interdisciplinary research project, South African Environmental Observation Network (SAEON), Ndlovu Node (Mightymand Mashele) and students/colleagues from the NWU and the University of Venda for assistance with dung sample collection. SALLNet and the National Institute for Communicable Diseases (NICD) for providing financial support for laboratory analyses. A special word of thanks to Mr Kai Behn (Potsdam University, Germany) for his administrative support to the project and the Head of NICD, Dr Arshad Ismail, for providing excellent post-sequencing support. I would also like to acknowledge Mr Riaan Pretorius and Dr Bianca Peterson for their help and patience with learning Bash and R scripting.

Lastly, I would like to thank my family and friends for their encouragement and for always believing in me, especially my parents and my sister, Chane Botha. A special thanks to my fiancé, Ian Oosthuizen, for his unfaltering support, motivation, and love during this stressful time. I could not have done this without him.

Abstract

Dietary analysis based on faecal DNA metabarcoding is dependent on the taxonomic coverage and quality of records available in DNA barcode reference databases, as well as the subsequent choice of bioinformatics pipeline to filter and assign degraded DNA from herbivorous faecal samples to the correct taxonomies. The aim of this study was to (i) create a comprehensive, study area-specific reference database containing the plant barcodes for all possible herbivore food targets in an eastern semi-arid South African savanna, (ii) compare bioinformatics approaches applied to faecal samples of cattle foraging in an eastern semi-arid South African savanna, and (iii) quantify these diets to determine the relative importance of food items. Sequences for the reference sequence databases of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) and transfer RNA leucine (*trnL*) UAA gene intron, which are commonly used barcodes in herbivore diet metabarcoding studies, were obtained from GenBank and BOLDSystems using a study-area-specific plant species list. The taxonomic reliability of these reference libraries was evaluated based on the presence of a barcode gap, selecting an appropriate identification threshold, and evaluating the identification accuracy using primary distance-based criteria. The performance of common bioinformatics pipelines (VSEARCH, OBITOOLS, and DADA2), and *in-silico* filtering parameters were compared using a mock community and evaluated based on its ability to recover these species in their initial abundances (high precision and low Root Mean Square Error (RMSE) values). The chosen method was applied to next-generation sequencing (NGS) datasets obtained from cattle foraging at two sites in the eastern savanna ecosystems of South Africa, namely the Syferkuil experimental farm in Limpopo and the rural Welverdiend village in Mpumalanga. The reference database evaluation revealed that barcode gaps were present for 76% of *rbcL* and 68% of *trnL* species. The identification success rate of these *rbcL* and *trnL* datasets using the *k-nn* criterion was 85,86% and 73,72%, respectively. According to the mock community, the DADA2 pipeline with frequency- and abundance-based filters as well as post-clustering curation performed best with precision values of 0,6 for the *rbcL* and *trnL* datasets and RMSE values of 11,2 and 9,6 for *rbcL* and *trnL*, respectively. Applying this pipeline to the *rbcL* datasets led the recovery of 8 families and 12 genera for Welverdiend, and 9 families and 13 genera for Syferkuil. The *trnL* datasets performed better without the abundance-based filter, identifying 15 families and 9 genera for Welverdiend and 8 families and 10 genera for Syferkuil. Frequency of occurrence and relative read abundance calculations revealed that forbs, non-graminoid herbaceous plants, may be of greater importance than previously recorded. This study provides a study-area-specific *rbcL* and *trnL* reference database that should be used collectively to identify semi-arid eastern savanna flora in South Africa as well as the providing a suitable bioinformatics approach to process DNA amplified from faecal samples with *rbcL* and *trnL* barcodes. This reference database and bioinformatics pipeline will improve our understanding of the composition of the South African savanna flora and therefore improve conservation and management measures targeted at these ecosystems.

Key terms: metabarcoding; bioinformatics; diet analysis; herbivore; reference database; *rbcL*, *trnL*; savanna; forbs

Table of Contents

DECLARATION.....	I
DEDICATION.....	II
ACKNOWLEDGEMENTS.....	III
ABSTRACT	IV
CHAPTER 1 INTRODUCTION	1
1.1 Background and rationale	1
1.2 Research aims and objectives	7
CHAPTER 2 MATERIALS AND METHODS.....	8
2.1 Study area.....	8
2.2 Plant material collection, DNA isolation, PCR amplification and sequencing.....	10
2.3 Faecal sample collection, DNA isolation, PCR amplification and sequencing	11
CHAPTER 3 A RBCL AND TRNL REFERENCE LIBRARY TO AID IN PLANT IDENTIFICATIONS OF A METABARCODING DIETARY ANALYSIS OF HERBIVORES FROM AN EASTERN SEMI-ARID SOUTH AFRICAN SAVANNA.....	13
3.1 Introduction	13
3.2 Materials and methods	14
3.2.1 Study area	14
3.2.2 Barcode reference list assembly	15
3.2.3 Phylogeny	15
3.2.4 Assembly of the DNA barcode reference database	16

3.2.5	Barcode gap analysis	16
3.2.6	Analysis with distance-based criteria	17
3.3	Results.....	17
3.3.1	DNA reference dataset: phylogenetic analysis and summary statistics.....	17
3.3.2	Reference database analysis	20
3.4	Discussion	23
3.5	Conclusion	26
CHAPTER 4 PIPELINE AND FILTERING VALIDATION VIA A MOCK COMMUNITY- BASED ANALYSIS.....		27
4.1	Introduction	27
4.2	Materials and methods	29
4.2.1	Mock community construction.....	30
4.2.2	Selected bioinformatics pipelines and filtering approaches.....	36
4.2.3	Sequence variation analysis.....	38
4.2.4	Final filtering of the mock community	40
4.2.5	Performance comparison criteria	42
4.3	Results.....	42
4.3.1	Simulated plant community	42
4.3.2	Closed-reference OTU picking.....	43
4.3.3	Influence of post-clustering filters.....	44
4.4	Discussion	55
4.5	Conclusion	59

CHAPTER 5	METABARCODING DIETARY ANALYSIS OF CATTLE FORAGING AT TWO DIFFERENT SITES IN AN EASTERN SEMI-ARID SOUTH AFRICAN SAVANNA	60
5.1	Introduction	60
5.2	Materials and methods	62
5.2.1	Study sites and sample collection.....	62
5.2.2	Bioinformatics analysis	62
5.2.2.1	Raw data quality control.....	62
5.2.2.2	Preparation of feature tables via DADA2.....	63
5.2.2.3	<i>In-silico</i> filtering strategies and taxonomic assignments.....	64
5.2.2.4	Dietary richness evaluation.....	65
5.3	Results.....	66
5.3.1	Sequencing results	66
5.3.2	DADA2 bioinformatics workflow	68
5.3.2.1	Filtering and trimming.....	68
5.3.2.2	Analysis of sequence variation	69
5.3.2.3	<i>In-silico</i> filters.....	69
5.3.3	Taxonomic assignment and diversity analysis.....	71
5.4	Discussion	97
5.5	Conclusion	103
CHAPTER 6	OVERALL DISCUSSION	105
CHAPTER 7	OVERALL CONCLUSION	109

BIBLIOGRAPHY	111
ANNEXURES	132
Annexure 1: Supplementary tables supporting Chapter 2.....	132
Annexure 2: Supplementary tables supporting Chapter 3.....	193
Annexure 3: Introducing an <i>rbcL</i> and <i>trnL</i> reference library to aid in the metabarcoding analysis of foraged plants from semi-arid eastern South African savannas	273
Annexure 4: Supplementary tables supporting Chapter 4.....	274
Annexure 5: Supplementary tables and figures supporting Chapter 5.....	278

List of Tables

Table 3.1: Discriminatory power of the <i>rbcL</i> and <i>trnL</i> reference datasets predicted by the application of three distance-based measures: nearest neighbour, Meier's best close match, as well as BOLD identification criterion, for both default and optimized thresholds for the inclusion and exclusion of singletons for the respective datasets. All three analyses were performed to identify queries within the respective identification thresholds to genus level.	22
Table 4.1: Estimated species composition of the floral mock community which represents some dietary items of cattle foraging in an eastern semi-arid South African Savanna ecosystem. The relative abundances of species in the table equal 100% and the accession number (in the case of GenBank), and process identity (in the case of BOLD sequences) are indicated.	31
Table 4.2: A comparison between the next-generation sequencing (NGS) simulators GRINDER and INSILICOSEQ, illustrating the ability of INSILICOSEQ to generate realistic error rates and sequence quality when compared to real data such as the data analysed in the DADA2 tutorial (https://benjjneb.github.io/dada2/tutorial). The quality profiles show a heat map of the frequency of each quality score at each base position. The error rates plots illustrate the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.....	35
Table 4.3: List of approaches used after closed-reference clustering of the VSEARCH bioinformatics pipeline. These values represent sample means retrieved for each method, for the <i>rbcL</i> and <i>trnL</i> plant barcoding marker.....	43
Table 5.1: Summary of the number of unique ASVs detected by each filtering approach as well as the percentage reduction in ASVs between filtering methods.....	70
Supplementary Table 1: Complete species list used in reference database.....	132
Supplementary Table 2: Plants from which leaf material was collected between 26 March and 6 April 2015.....	190

Supplementary Table 3: Taxonomy and accession numbers (GenBank) and/or Process ID's (BOLD) included in the <i>rbcL</i> and <i>trnL</i> reference datasets.	193
Supplementary Table 4: Species sequenced at the NWU with the <i>rbcL</i> and <i>trnL</i> barcode which have been submitted to GenBank.	269
Supplementary Table 5: Comparison criteria of the filtering methods applied to the <i>rbcL</i> mock community dataset. Values of the means for samples and their replicates are used to calculate the final value for the comparison indices. Methods appear in the order of highest precision, lowest RMSE and lastly highest sensitivity.....	274
Supplementary Table 6: Comparison criteria of the filtering methods applied to the <i>trnL</i> mock community dataset. Values of the means for samples and their replicates are used to calculate the final value for the comparison indices. Methods appear in the order of highest precision, lowest RMSE and lastly highest sensitivity.....	276
Supplementary Table 7: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Welverdiend cattle barcoded with <i>rbcL</i>	281
Supplementary Table 8: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Welverdiend cattle barcoded with <i>trnL</i>	283
Supplementary Table 9: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Syferkuil cattle barcoded with <i>rbcL</i>	284
Supplementary Table 10: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of	

ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Syferkuil cattle barcoded with <i>trnL</i>	286

Supplementary Table 11: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Welverdiend <i>rbcL</i> dataset that could not be assigned by the compiled reference database of <i>rbcL</i> and <i>trnL</i> barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.	291

Supplementary Table 12: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Welverdiend <i>trnL</i> dataset that could not be assigned by the compiled reference database of <i>rbcL</i> and <i>trnL</i> barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.	292

Supplementary Table 13: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Syferkuil <i>rbcL</i> dataset that could not be assigned by the compiled reference database of <i>rbcL</i> and <i>trnL</i> barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.	294

Supplementary Table 14: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Syferkuil <i>trnL</i> dataset that could not be assigned by the compiled reference database of <i>rbcL</i> and <i>trnL</i> barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values	

smaller than 0,001. These taxonomic assignments were made with ASVs
resulting from the application of the DADA_1_LULU workflow. 304

Supplementary Table 15: Query sequence identities (“qseqid”) of amplicon sequence variants
(ASVs) from the Syferkuil *rbcL* dataset that could not be assigned by the
compiled reference database of *rbcL* and *trnL* barcodes of potential foraging
targets in an eastern semi-arid South African savanna but could be assigned to
species other than plants by NCBI’s non-redundant nucleotide database with
blast parameters: 80% sequence identity, 80% query coverage and e-values
smaller than 0,001. These taxonomic assignments were made with ASVs
resulting from the application of the DADA_1_RA_LULU workflow. 306

Supplementary Table 16: Query sequence identities (“qseqid”) of amplicon sequence variants
(ASVs) from the Syferkuil *trnL* dataset that could not be assigned by the
compiled reference database of *rbcL* and *trnL* barcodes of potential foraging
targets in an eastern semi-arid South African savanna but could be assigned to
species other than plants by NCBI’s non-redundant nucleotide database with
blast parameters: 80% sequence identity, 80% query coverage and e-values
smaller than 0,001. These taxonomic assignments were made with ASVs
resulting from the application of the DADA_1_RA_LULU workflow. 312

List of Figures

Figure 2.1: Map showing the dietary analysis sites (Syferkuil and Welverdiend), the sites used for collection of plant material (Timbavati, Nkuhlu and Welverdiend) and reference points for the extraction of POSA species lists (all four indicated sites) concerning the Central Bushveld and Lowveld savanna bioregions in South Africa (map courtesy of Wynand Muller, metadata provided by SANBI (2012), Municipal Demarcation Board: https://dataportal-mdb-sa.opendata.arcgis.com/datasets/37d790cf4f3b4b0ebd0ac501eae9b17d_0/explorer?location=29.911228%2C26.005192%2C5.96 and the National Geographic Society: https://services.arcgisonline.com/ArcGIS/rest/services/NatGeo_World_Map/MapServer).	10
Figure 3.1: The composition of the final DNA sequence reference database in terms of number of genera (histogram) and species (line plot) per family.	19
Figure 3.2: Scatterplots of the Barcode gaps for species present A) in the <i>rbcL</i> dataset and (B) <i>trnL</i> dataset. Scatterplots show the relationships between minimum interspecific distance (y-axis) and maximum intraspecific distance (x-axis) where the barcode gap is indicated by points falling above the 1:1 slope. High densities of plots are indicated by the darker colours.	20
Figure 4.1: Overview of the main steps followed in the materials and methods section for the mock community analysis. * An external quality filter was added to the VSEARCH approach since this workflow does not incorporate one in its various steps. ** “OTU” in this diagram may refer to operational taxonomic units as retrieved from the clustering of sequences based on 97% similarity in the VSEARCH approach, amplicon sequence variants retrieved by the inference of an error model in the DADA2 approach or grouping of strictly identical sequences as in the OBITOOLS approach.....	30
Figure 4.2: Abundance compositions of the four mock community samples generated from 23 species.	34
Figure 4.3: A comparison of the performance indices, precision (orange scatterplots), and sensitivity (blue scatterplots) for the OBITOOLS methods OBI_1 (A and B) and OBI_10 (C and D) as applied to the <i>trnL</i> mock community dataset. The	

scatterplots represent the mean index received for each of the 4 samples with its 5 replicates and the boxplots represent the median of the performance index as well as the first and third quartile of the 20 samples.	45
Figure 4.4: Stacked bar plots indicating the relative read abundance (%) by plant per A) plant family and B) genera identified by the OBI_1 method. This filtering method encompasses the removal of features that occur once per sample.....	46
Figure 4.5: Stacked bar plots indicating the relative read abundance (%) by plant per A) plant family and B) genera identified by the OBI_10 method. This filtering method includes the removal of features that occur 10 times or less per sample.....	47
Figure 4.6: The number of sequence proxies identified by the DADA2 denoising methods and the VSEARCH <i>de novo</i> clustering methods for the <i>rbcL</i> mock community dataset. The highest feature richness (69 384 OTUs) was retrieved by the VSEARCH <i>de novo</i> method with no post- or pre- clustering filters applied (VSEARCH_DN_ALL) and the lowest was retrieved by two methods applied for the DADA2 bioinformatics pipeline: DADA_1_RA_LULU identifying 64 ASVs, and DADA_1_LULU also identifying 64 ASVs, where the latter applied both a frequency and abundance filter.....	49
Figure 4.7: The number of sequence proxies identified by the DADA2 denoising methods and the VSEARCH <i>de novo</i> clustering methods for the <i>trnL</i> mock community dataset. The highest feature richness (59 444 OTUs) was retrieved by the VSEARCH <i>de novo</i> method with no post- or pre- clustering filters applied (VSEARCH_DN_ALL) and the lowest was retrieved by two methods applied for the DADA2 bioinformatics pipeline: DADA_1_RA_LULU which identified 132 ASVs and DADA_1_LULU also identifying 132 ASVs, where the latter applied both a frequency and abundance filter.....	49
Figure 4.8: Percentage ASVs, in the case of DADA2 methods, and OTUs, in the case of VSEARCH methods, identified for both the <i>rbcL</i> (red bars) and <i>trnL</i> (blue bars) mock community datasets. These percentages were calculated by dividing the number of assigned reads by the total richness of reads.....	50
Figure 4.9: Comparison indices of precision and RMSE of each method of the a) <i>rbcL</i> and b) <i>trnL</i> mock community datasets. The scatter plot points represent mean values averaged across each method for all samples and their replicates with precision,	

or the positive predictive value, represented by $TP/(TP + FP)$ and RMSE calculated between observed and expected abundances retrieved for each sample replicate, averaged for each method.....	52
Figure 4.10: Tile plot indicating species that could be identified for each method applied for the <i>rbcL</i> mock community. Yellow tiles indicate identified species and purple tiles indicate absent species per method.....	53
Figure 4.11: Tile plot indicating species that could be identified for each method applied for the <i>trnL</i> mock community. Yellow tiles indicate identified species and purple tiles indicate absent species per method.....	54
Figure 5.1: Scatterplots representing the average sequence length per sample plotted against the number of sequenced reads per sample retrieved from the Illumina MiSeq platforms for reads sequenced in the forward direction for the Welverdiend A) <i>rbcL</i> and B) <i>trnL</i> datasets as well as the Syferkuil C) <i>rbcL</i> and D) <i>trnL</i> dataset. The data for these plots were exported from MULTIQC to R to create visualisations with <i>ggplot2</i> for the sake of image resolution and figure uniformity throughout the study.	67
Figure 5.2: Stacked bar plots indicating the number of reads retained (blue bars) relative to the number of reads removed (red bars) as a result of sequence filtering and trimming in DADA2 for the Welverdiend A) <i>rbcL</i> and B) <i>trnL</i> dataset.....	68
Figure 5.3: Stacked bar plots indicating the number of reads retained (blue bars) relative to the number of reads removed (red bars) as a result of sequence filtering and trimming in DADA2 for the Syferkuil A) <i>rbcL</i> and B) <i>trnL</i> dataset.	69
Figure 5.4: Bar plots generated from taxonomic assignments for the Welverdiend <i>rbcL</i> dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).	74
Figure 5.5: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Welverdiend <i>rbcL</i> dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).	75

Figure 5.6: Bar plots generated from taxonomic assignments for the Welverdiend <i>trnL</i> dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).	77
Figure 5.7: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Welverdiend <i>trnL</i> dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).	78
Figure 5.8: Bar plots generated from taxonomic assignments for the Syferkuil <i>rbcL</i> dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).	80
Figure 5.9: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Syferkuil <i>rbcL</i> dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).	81
Figure 5.10: Bar plots generated from taxonomic assignments for the Syferkuil <i>trnL</i> dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).	83
Figure 5.11: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Syferkuil <i>trnL</i> dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).	84
Figure 5.12: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Welverdiend cattle faecal samples by the <i>rbcL</i> barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.	87

Figure 5.13: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Welverdiend cattle faecal samples by the <i>trnL</i> barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.....	90
Figure 5.14: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Syferkuil cattle faecal samples by the <i>rbcL</i> barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.	93
Figure 5.15: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Syferkuil cattle faecal samples by the <i>trnL</i> barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.	96
Figure 5.16: Mirrored bar chart showing the frequency of occurrence (FOO%) of plant families identified as part of the diets of Welverdiend cattle. Orange bars represent families identified barcoded with <i>rbcL</i> and blue bars represent those barcoded with <i>trnL</i>	103
Figure 5.17: Mirrored bar chart showing the frequency of occurrence (FOO%) of plant families identified as part of the diets of Syferkuil cattle. Orange bars represent families identified barcoded with <i>rbcL</i> and blue bars represent those barcoded with <i>trnL</i>	103
Supplementary Figure 1: The quality profiles of Welverdiend sequence data barcoded with <i>rbcL</i> shows a heat map of the frequency of each quality score at each base position.	278
Supplementary Figure 2: The quality profiles of Welverdiend sequence data barcoded with <i>trnL</i> shows a heat map of the frequency of each quality score at each base position.	279
Supplementary Figure 3: The quality profiles of Syferkuil sequence data barcoded with <i>rbcL</i> shows a heat map of the frequency of each quality score at each base position.	280
Supplementary Figure 4: The quality profiles of Syferkuil sequence data barcoded with <i>trnL</i> shows a heat map of the frequency of each quality score at each base position.	281
Supplementary Figure 5: The error rates plot for Welverdiend sequence data barcoded with <i>rbcL</i> illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are	

represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.....288

Supplementary Figure 6: The error rates plot for Welverdiend sequence data barcoded with *trnL* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.....289

Supplementary Figure 7: The error rates plot for Syferkuil sequence data barcoded with *rbcL* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.....290

Supplementary Figure 8: The error rates plot for Syferkuil sequence data barcoded with *trnL* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.....291

Chapter 1 Introduction

Background and rationale

In recent decades, African savannas have developed into a major source for livestock production at the expense of a once diverse suite of native large mammalian herbivore (LMH) populations (Baumgarten *et al.*, 2015), and the restoration of these damaged assemblages has become a key conservation objective (Craigie *et al.*, 2010; Kartzinel *et al.*, 2015; Pansu *et al.*, 2019). Due to the imbalance of woody and grassy vegetation caused by increased livestock production and widespread overgrazing, the growth of small woody plants is encouraged while the amount of herbaceous vegetation is drastically reduced (Wessels *et al.*, 2004; Baumgartner *et al.*, 2015). The quantification of cattle diet composition will therefore highlight key savanna plant species contributing to their diet and have significant implications for field management and conservation efforts (Thomsen & Willerslev, 2015; Pringle *et al.*, 2016; Scasta *et al.*, 2019; Bhattacharyya *et al.*, 2019; Camp *et al.*, 2020). Savannas are usually characterised by the coexistence of only two distinct plant life forms, namely trees and grasses, which compete for the same limited resources. However, there aren't many ecological studies on the function of forbs, or non-graminoid herbaceous vascular plants, in savanna environments, although they make up a sizeable and identifiable portion of the herbaceous layer and contribute to the forage of various herbivore guilds (Siebert & Dreber, 2019; Siebert & Scogings, 2015). Forbs may also play a significant seasonal role in the diets of ungulates (Du Toit *et al.*, 2003; Odadi, 2013; Van Der Merwe & Marshal, 2012) and provide forage stability during stressed conditions, such as drought (Van Coller *et al.*, 2018). The importance of forbs to the ecology of semi-arid South African savannas has recognized the important ecosystem functions they provide (e.g., Van Der Merwe & Marshal, 2012; Siebert & Scogings, 2015; Siebert & Dreber, 2019; Van Coller *et al.*, 2018; Van Coller *et al.*, 2021).

Preceding the determination of ecological roles of various forb species in an African savanna is the establishment of the presence of these species in the diets of herbivores foraging in this ecosystem and to which extent they are utilised. Animal dietary analysis is essential since it provides insight into the dynamics of the food web and how the feeding habits of different herbivore species affect a particular ecosystem (Duffy *et al.*, 2007; Valentini *et al.*, 2008). Therefore, understanding herbivore preferences and browsing intensities among forb species can help determine the resource value of forbs (Siebert & Dreber, 2019). In this regard, the current study aims to determine, firstly, with high accuracy which forb species contribute to the diets of cattle foraging in an eastern semi-arid South African savanna during the dry season and secondly to reliably quantify the diets of these herbivores to determine the relative importance of these species in relation to other dietary objects.

The analysis of animal faeces has become important since ecological information can be gained without the direct observation of the animal (Kohn & Wayne, 1997). Methods such as direct field observations (Owen-Smith, 1979), micro histological techniques aimed at faeces or gut content (Holechek *et al.*, 1982; McInnis *et al.*, 1983), near-infrared reflectance spectroscopy (Foley *et al.*, 1998; Boval *et al.*, 2004), stable isotope analysis (Ben-David *et al.*, 2001) and cafeteria experiments (Johannessen & Samset, 1994; Shuai *et al.*, 2019), have traditionally been used to determine dietary profiles. Although labour-intensive, these methods demand excellent taxonomic knowledge, and they frequently fail to accurately distinguish between related taxa, which could result in mistakes and skewed ecological inferences (Pompanon *et al.*, 2012; Pavan-Kumar *et al.*, 2015; Weiskopf *et al.*, 2016). Molecular species identification, first established by Arnot *et al.* (1993) and advanced by Hebert *et al.* (2003), has enabled diet evaluations through developments in sequencing technology (Vesterinen *et al.*, 2013; Pavan-Kumar *et al.*, 2015). ‘DNA barcoding’ (Hebert *et al.*, 2003) is a straightforward concept of utilising a short, standardised DNA fragment as a barcode for single-species identification via Sanger sequencing. The method has numerous applications in the study of trophic relationships, invasion biology, environmental management, conservation of biodiversity, and food safety (Valentini *et al.*, 2009; Armstrong & Ball, 2005; Pečnikar & Buzan, 2014).

Recent technological developments, based on high-throughput sequencing, have brought about new perspectives for dietary analysis and biodiversity assessments, and have made it possible to evaluate biodiversity more directly than through the screening of individual specimens. The DNA metabarcoding technique makes use of Next Generation Sequencing (NGS) technology, enabling multiple species identification by the enormously parallel sequencing of bulk samples (Taberlet *et al.*, 2012; Pompanon *et al.*, 2012; Clare, 2014). The estimation of animal diets by faecal DNA metabarcoding has been recognized as a non-intrusive, precise, time- and cost-efficient method. The high taxonomic resolution displayed by this technique (Nichols *et al.*, 2016) has enabled the detection of numerous food items that are challenging to identify by traditional methods (Soininen *et al.*, 2009; Bhattacharyya *et al.*, 2018), allowing for the exact mechanisms of niche partitioning in numerous sympatric species to be revealed (Kartzinel *et al.*, 2015; Pansu *et al.*, 2018; Ter Schure *et al.*, 2021; Aylward *et al.*, 2022) and showed several crucial elements of wildlife management such as food selection of endangered species (Iwanowics *et al.*, 2016; Gebremedhin *et al.*, 2016; Goldberg *et al.*, 2020). The accuracy of the faecal DNA metabarcoding method has been thoroughly gouged by studies based on captive feeding trials (Deagle *et al.*, 2010; Scasta *et al.*, 2019; Stapleton *et al.*, 2022).

The reconstruction of herbivore dietary profiles from faeces requires a comprehensive reference database of sequences from potential dietary items which can be matched against sequences recovered from faecal samples (Lee *et al.*, 2018; Gill *et al.*, 2019). The use of DNA barcoding in single species detections has started to satisfy this requirement by the creation of reference sequences libraries of standardised genomic

regions which follows strict inclusion standards such as the provision of metadata and vouchers of the sequenced specimens in museum or herbarium collections (Ferri *et al.*, 2015). The selection of which genomic region, or DNA barcode, to use for species identification in a metabarcoding study is crucial to retrieve and identify target taxa from an environmental sample. These barcodes are chosen for their sequence recovery, annotation, and sequence resolution among taxa (Hajibabaei *et al.*, 2016). While the mitochondrial cytochrome c oxidase subunit 1 (*COI*) gene represents an exemplary barcode for the identification of animals (Coissac *et al.*, 2016), plants possess no such single candidate marker exhibiting the same characteristics (Yu *et al.*, 2011). The Consortium for the Barcode of Life (CBOL) Plant Working Group (2009) has assessed seven plastid DNA (ptDNA) regions and has recommended the use of a two-locus plant identification system that consists of a fragment of the ribulose 1,5 – biphosphate carboxylase gene (*rbcL*) and a region of the maturase gene (*matK*), both of which are located on the chloroplast genome (Lahaye *et al.*, 2008; CBOL Plant Working Group, 2009). It was however emphasised that this barcode standard must be augmented in some studies, such as dietary analysis due to the required ability of the barcoding marker to amplify degraded and low copy number DNA from faecal samples (De Groot *et al.*, 2011; Hollingsworth *et al.*, 2011). Although *matK* may prove useful in single species detection, its application in metabarcoding may be limited since it may lack the conserved sites suitable for PCR primer binding, is notoriously difficult to amplify in some plant groups, and its length (1570 bp) is ill-suited to most metabarcoding studies (Wilkinson *et al.*, 2017).

In the case of herbivore diet analysis, relatively few studies include the use of the core plant barcodes recommended by CBOL Plant Working Group (Garcia-Robledo *et al.*, 2013; Kartzinel *et al.*, 2015; Mallot *et al.*, 2018). The *rbcL* and *trnL*(UAA) barcodes have been proposed as two single-locus gene regions for the analysis of herbivorous- and omnivorous animal diets via DNA metabarcoding (Taberlet *et al.*, 2007; Valentini *et al.*, 2009; Kartzinel *et al.*, 2015; Kress *et al.*, 2015; Mallot *et al.*, 2018), and will constitute the core barcoding regions in this study. Advantageous characteristics of *rbcL* include the universality of its primers (De Groot *et al.*, 2011; Yang *et al.*, 2014), the conservative coding nature of the region that can be used to reconstruct deep evolutionary relationships, high recoverability, and specimen identification up to family level (Taberlet *et al.*, 2007; Ferri *et al.*, 2015; Lee *et al.*, 2018). *TrnL* is also a chloroplast gene and yields small DNA fragments with the whole intron (254-767 bp) representing the forward *c* and reverse *d* regions which are highly conserved among land plants (Taberlet *et al.*, 1991; Taberlet *et al.*, 2007). The *rbcL* and *trnL* barcodes are also well represented on public reference databases with 307 756 *rbcL* and 337 987 *trnL* plant sequences on the non-redundant (nt) nucleotide database of the National Centre for Biotechnology Information (NCBI: Johnson *et al.*, 2008) as of the 7th of November 2022.

Along with the choice of marker to use for the identification of plant species from herbivore diets, the source of the reference database itself also requires some special considerations. Reference sequences used

to assign unknown queries are commonly found on public reference databases (Buchner & Leese, 2020), such as GenBank (Benson *et al.*, 2018) and the Barcode of Life Data System (BOLD: Ratnasingham & Hebert, 2007). The enormous size of these reference databases, however, can create difficulties for molecular taxonomists by containing reference sequences of taxonomies other than plants, reference sequences of poor quality as well as unvouchered or incorrectly vouchered specimens. Studies such as Lamb *et al.* (2016) showed that the use of a reference database specifically curated to contain taxa present in the study area significantly improved the taxonomic identifications which could be related to the actual vegetation available to the animal of interest. Except for the use of a study area-specific reference database, studies such as Gill *et al.* (2019), Sheidai *et al.* (2019) and Lin *et al.* (2021) showed that the efficiency of a reference database to discriminate among species in an environmental sample can be assessed by using the genetic distance approach, i.e., the amount of genetic variation displayed. This approach uses the basis of the “barcoding gap” definition, i.e., the amount of sequence divergence within species should be smaller than that between species for a barcoding locus to discriminate among species. Chapter 3 of this study deals with the compilation and analysis of a comprehensive *rbcL* and *trnL* reference database that can effectively discriminate among plant species in an herbivore faecal sample by containing potential forage from an eastern semi-arid South African savanna.

After the construction of a reference database of standard barcode sequences of potential food items, these same taxonomically informative genes must be recovered and sequenced from herbivorous faecal samples. This process entails DNA extraction, DNA amplification with universal DNA barcode primers and submission of these amplified sequences to high throughput sequencing (HTS) (Deiner *et al.*, 2017; Taberlet *et al.*, 2018). HTS is facilitated by NGS systems that can all simultaneously generate millions of sequencing reads (Pavan-Kumar *et al.*, 2015). An example of such a platform is Roche 545 FLX, used for its ability to generate long read lengths, therefore making it suitable for long amplicon lengths (Pavan-Kumar *et al.*, 2015). However, this platform is unable to reliably read homopolymers and is no longer an option since Roche stopped producing these machines in 2016 (Taberlet *et al.*, 2018). Ion Torrent is another example of an NGS sequencing platform but is used in very few metabarcoding investigations (Deagle *et al.*, 2013; Geml *et al.*, 2014; Kraaijeveld *et al.*, 2015), with a preference for Illumina sequencing in most metabarcoding studies (Taberlet *et al.*, 2018). Amplicon fragment size is an important consideration when selecting an appropriate NGS platform. As demonstrated by the Illumina MiSeq platform which produces 2 x 300 bp paired-end reads, as used in this study, would pose technical limitations for the lengthy amplicons (1570 bp) produced by *matK* primers (Bell *et al.*, 2017).

The downstream analysis of the resulting sequencing data, which typically contains millions of amplicon DNA fragments, would not be possible without the use of bioinformatics (Coissac *et al.*, 2012). Several bioinformatics tools are available for the analysis of NGS datasets and develop at such a phenomenal pace

that new software is made regularly publicly available. Due to the availability of a magnitude of bioinformatics algorithms, programs and scripts, careful consideration is needed when choosing a bioinformatics workflow specific to the needs of the study (Kopylova *et al.*, 2016). Three steps make up a bioinformatics workflow, or pipeline: i) quality control, ii) clustering of DNA sequences based on their similarity, and (iii) taxonomic assignment via a reference library of taxonomically informative barcode regions (Lamb *et al.*, 2016).

Quality screening typically entails trimming raw sequences of Illumina adapters and low-quality bases according to their Phred scores, which is an estimated probability of error for each base introduced during sequencing (Martin, 2011; Caporaso *et al.*, 2010). Forward and reverse reads are also merged during the initial steps of the bioinformatics workflow to create a single consensus sequence (Quail *et al.*, 2012) after which chimeras are detected and removed (Edgar *et al.*, 2011). Chimeras are artifacts of PCR amplification when the extension of an amplicon sequence is aborted in one cycle and subsequently serves as a primer in the next PCR cycle. This faulty “primer” results in the synthesis of a hybrid product by annealing to random DNA templates and a continuous extension (Haas *et al.*, 2011; Fonseca *et al.*, 2012). Sequence clustering typically follows the quality screening, where sequences are grouped into molecular operational taxonomic units (OTUs): read clusters that differ by less than a predetermined threshold of sequence dissimilarity which is often 3% (Westcott & Schloss, 2015; Kopylova *et al.*, 2016; Alberdi *et al.*, 2018). The results of this clustering based on similar sequences results in an OTU table where the observation of an OTU is treated as the observation of a species (Callahan *et al.*, 2017). A key distinction between the definition of OTU’s is whether they are formed by a closed-reference clustering approach, which groups sequences based on their similarity to a reference sequence or whether they are formed by grouping sequences based on a *de novo* method, where clustering occurs as a function of their pairwise sequence similarities (Callahan *et al.*, 2017). Recently, amplicon sequence variant (ASV) methods have been developed that group sequences by resolving ASVs from Illumina amplicon data without the inference of the dissimilarity thresholds, as is needed for forming OTUs (Callahan *et al.*, 2016, Amir *et al.*, 2017). ASV techniques discriminate between sequences by identifying differences of as little as one nucleotide and infer biological sequences in samples before the introduction of PCR or sequencing errors and have further demonstrated to be better than OTUs to discriminate ecological patterns (Callahan *et al.*, 2016; Needham *et al.*, 2017). However, compared to OTU methods, ASV approaches are limited in their capacity to handle short or missing overlap regions and are unable to compute ambiguous bases (Jeske & Gallert, 2022). The method chosen to represent species proxies in metabarcoding studies, therefore, needs to be thoroughly considered to be as representative of the actual vegetation of the studied ecosystem as possible.

The identification of false negatives, false positives, or low taxonomic resolution might result from biases that enter during any step of the experiment, making it difficult to obtain reliable and reproducible

metabarcoding results (Coissac et al., 2016). The feature (OTU/ASV) table resulting from the clustering/denoising of sequences should therefore undergo further revision and filtering before taxonomic assignment. Many rare features, mainly those occurring in low frequencies in certain samples and low abundances across the samples dataset, are the result of PCR and sequencing errors (Huse et al., 2010; Brown et al., 2015). The practice of excluding features based on the suspicion that they represent errors (Huse et al., 2010; Brown et al., 2015; Alberdi et al., 2018) unfortunately leads to the loss of real, but rare biological diversity. Apart from post-clustering filtering strategies of the feature table to remove PCR and sequencing errors according to low abundant/low-frequency reads, post-clustering curation algorithms are also available, such as LULU. LULU identifies and merges often erroneous ‘daughter’ OTUs with consistently abundant co-occurring sequences, or ‘parent’ OTUs (Frøslev et al., 2017).

Due to the magnitude of bioinformatics workflows with stand-alone or built-in quality filters and clustering algorithms, as well as various thresholds of *in-silico* filtering techniques and curation algorithms the choice of an optimal method for species discrimination and estimations of their relative importance to the diets of herbivores is challenging. Dietary metabarcoding also offers a distinctive set of issues impeding quantification. PCR-based metabarcoding data is already difficult to quantify due to problems including PCR primer bias and random sampling during sequencing (Murray et al., 2011). These biases and problems are especially common in dietary metabarcoding and are made worse by the amount of biomass consumed, DNA density in tissues, individual diversity in the digestion of tissues, and variable DNA degradation (Deagle et al., 2010; Thomas et al., 2014). When considering the precautionary principle, which emphasises exercising caution and thoughtful consideration before taking any action when there is insufficient evidence to support it (Van Der Sluijs et al., 2005), the application of a bioinformatics workflow should first be evaluated based on its efficiency to recover species from an herbivore diet in the correct composition and abundance. This is achieved by the creation of a mock community in Chapter 4 (as seen in Matezans et al., 2019; Chua et al., 2021) that mimics the range of species richness and relative abundances expected to be found in cattle diets foraging in an eastern semi-arid South African savanna.

The resulting feature table is a good basis for the downstream analysis where ecological deductions are made and can be performed in R (R Core Team, 2022) with various packages dedicated to such evaluations like PHYLOSEQ (McMurdie & Holmes, 2013) and VEGAN (Oksanen et al., 2017). There is growing proof that DNA metabarcoding is a robust method of assessing biodiversity from bulk environmental samples, however, metabarcoding data has been the subject of much discussion over the possibility of quantifying the contributions of various taxa to a consumer's diet (Deagle et al., 2011; Thomas et al., 2016; Piñol et al., 2018). Some authors limit the use of the metabarcoding technique to strictly determine occurrence data, i.e., presence-absence data, while others promote a quantitative approach (Fonseca et al., 2012; Deiner et al., 2017; Bell et al., 2019). In this study, the comparison between pipelines, specifically

the application of abundance-based filters on data retrieved from faecal samples, as well as the presentation and discussion of both species' lists, and relative abundance data are presented in Chapter 5. Both matrices for dietary quantification are presented to maximise the information level and to facilitate comparisons to other dietary studies. The environmental DNA for this study will be sourced from the faeces of cattle foraging in two systems representative of a semi-arid South African savanna during the dry season. These systems represent different grazing regimes, i.e., non-rotational and rotational grazing, which represent different available forage items. Such variance provides an opportunity to further test the accuracy of the bioinformatics workflow across different plant communities.

Research aims and objectives

This study aims to identify a bioinformatics approach with filtering parameters suited for plant identifications from faeces and dietary quantifications of cattle foraging in an eastern semi-arid South African savanna, with special attention to forb species which have been neglected in previous studies and models concerning savanna ecology. This aim was satisfied by:

- i) compiling a comprehensive, study-area-specific *rbcL* and *trnL* reference database containing all the possible foraging targets of herbivores occurring in an eastern semi-arid South African Savanna,
- ii) determining an optimal bioinformatics approach to recover the range of species and their abundances from a mock community of species which could occur in the diet of cattle foraging in such an ecosystem, and
- iii) applying this approach to cattle faecal samples from study areas situated in an eastern semi-arid South African savanna to identify key dietary items in terms of plant families and genera and inferring the relative importance of forbs as forage to these animals.

Chapter 2 Materials and methods

Study area

The Savanna Biome covers 32,8% (i.e., ~399 600 km²) of the land surface of South Africa and can be divided into six bioregions: Central Bushveld, Mopane, Lowveld, Sub-Escarpment Savanna, Eastern Kalahari Bushveld and Kalahari Duneveld (Rutherford *et al.*, 2006). These bioregions can furthermore be divided into fine-leaf (eutrophic) savannas on nutrient-rich substrates with a higher clay content, or broad-leaf (dystrophic) savannas on nutrient-poor, sandy soil (Huntley, 1982). The Kalahari Duneveld and Kalahari Bushveld bioregions in the western parts of South Africa represent the drier savannas, while the eastern savannas are either moist (>650 mm mean annual rainfall (MAP)) or semi-arid (<650 mm) according to the accepted MAP delineation for semi-arid savannas (Sankaran *et al.*, 2005).

This study was undertaken in two bioregions within the eastern semi-arid savanna of South Africa, namely the Central Bushveld- and the Lowveld Bioregion (Fig. 2.1). These bioregions differ in terms of (i) soil, since the Central Bushveld Bioregion is located on soil with a higher clay content compared to the Lowveld Bioregion which is associated with sandy soils as a result of pedogenesis from granite, and (ii) dominant tree species with the Central Bushveld region being dominated by *Acacia* species (synonym species from the genera *Vachellia* and *Senegalia*) such as *V. tortilis* and *V. nilotica* whereas the Lowveld Bioregion is more favourable to *Combretum* species, such as *C. zeyheri*, *C. apiculatum*. The difference in substrate is also reflected in the broadly different vegetation descriptions with the Central Bushveld on clayey soil representing a typical eutrophic (fine-leaved) savanna, whereas the Lowveld region on sandy soil represents a dystrophic (broad-leaved) savanna. Although these structurally different savanna types affect dominant vegetation types, such differences are much stronger displayed in the tree and shrub layer than in the herbaceous layer. The plant species richness in these semi-arid savanna ecosystems has also been described as similar, and the beta-diversity (i.e., species turnover) of the herbaceous layer in the South African savanna ecosystem much lower than, for instance, in the Grassland Biome (Whittaker *et al.*, 1984).

The different objectives for this study were also achieved at different study sites or areas, although they were all included within the two selected bioregions. The study sites that were selected for faecal sample collection included Welverdiend (Mpumalanga) in the Lowveld Bioregion, and Syferkuil (Limpopo) in the Central Bushveld Bioregion (Fig. 2.1). To optimise the reference database for plant DNA sequences, several dominant plant species of the eastern semi-arid savannas (specifically targeting species that have not been captured in GenBank), were collected for DNA sequencing, of which the voucher specimens were added to the NWU AP Goossens Herbarium. Sites from which such plant material was collected included the Nkuhlu long-term monitoring sites in the Kruger National Park (Mpumalanga), Timbavati (Timbavati

Private Nature Reserve, Limpopo) and Welverdiend communal rangeland (Mpumalanga) (Fig. 2.1). A comprehensive plant species list of all the plants in the areas surrounding the faecal and plant material sampling sites was generated through the implementation of a grid search in the Plants of South Africa (POSA) database. Plant species lists were extracted from a grid of ~10 km radius from the sampling site to capture the representative flora of each sampling site.

Except for the use of POSA lists, species lists were also available for the study sites of Welverdiend and Syferkuil from previous studies at the NWU (Klem, 2019). These records were used as a foundation in building a comprehensive eastern semi-arid South African savanna species list (Supplementary Table 1). This list was also enriched with species from other studies on African savannas (Kartzinel *et al.*, 2015; Gill *et al.*, 2019), based on local expert knowledge of the research team member, Professor Frances Siebert (e.g., Siebert & Dreber, 2019; Silveira *et al.*, 2020; Buisson *et al.*, 2021; Siebert & Eckhardt, 2008; Siebert, 2012, Siebert *et al.*, 2010; Van Coller *et al.*, 2013; Erasmus *et al.*, 2021).

The study sites used for the estimation of dietary profiles of domesticated cattle (Syferkuil and Welverdiend, Fig. 2.1) are also associated with different land-use types. The University of Limpopo's Experimental Farm, Syferkuil, specifically the cattle camp in which the *Drought Act* experiment was conducted (23°50'36.86"S; 29°40'54.99"E), represents a rangeland where rotational grazing is employed. The climate of the area is defined as semi-arid with an average annual rainfall of 350-500 mm, most of which occurs in the summer months (Phefadu & Katu, 2016). The rural village of Welverdiend (24°34'15" S, 31°20'42" E) represents a communal rangeland with a MAP of 600 mm (Shackleton, 2000; Matsika *et al.*, 2013). This village is located close to Bushbuckridge in the Lowveld, in the Bohlabela District, Mpumalanga.

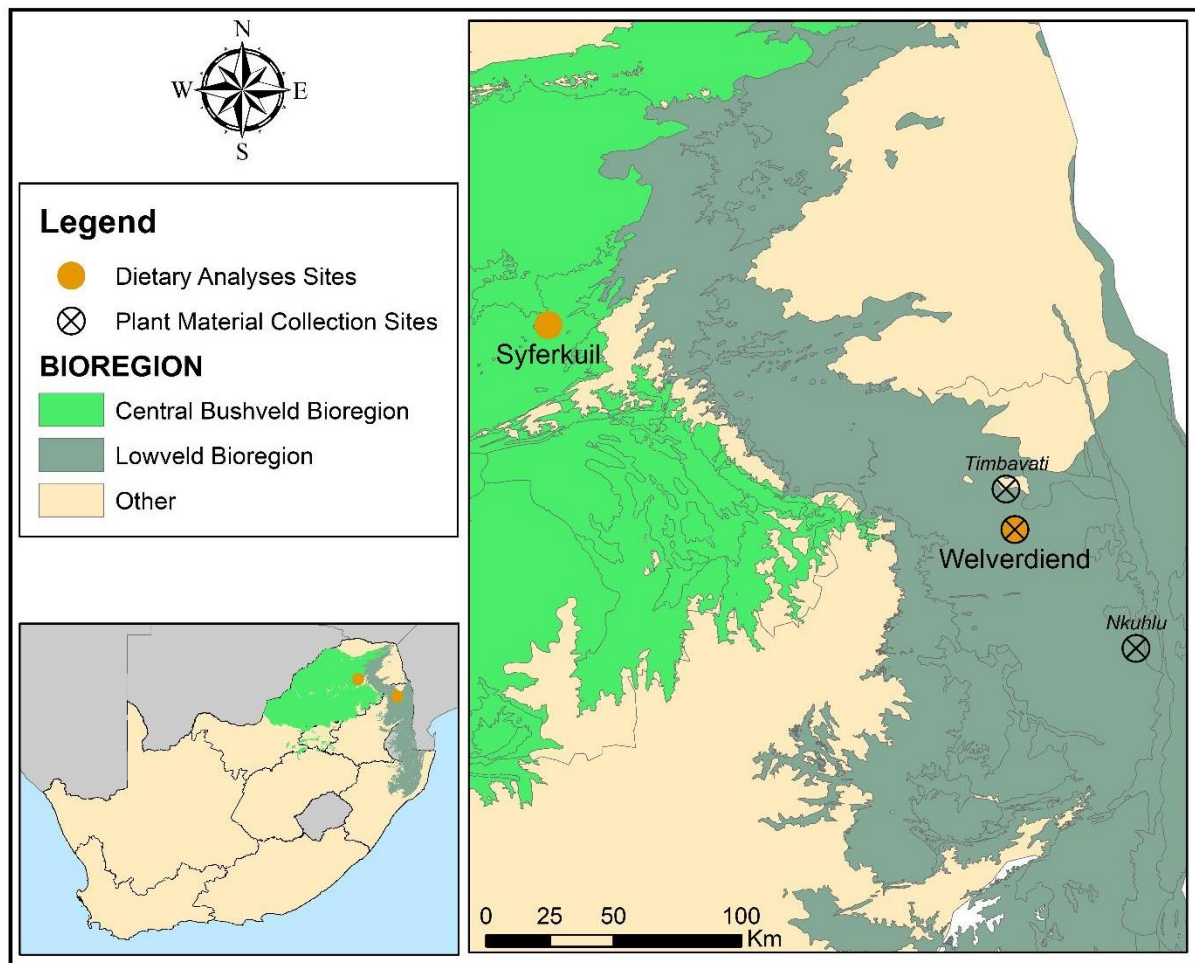


Figure 2.1: Map showing the dietary analysis sites (Syferkuil and Welperdiend), the sites used for collection of plant material (Timbavati, Nkuhlu and Welperdiend) and reference points for the extraction of POSA species lists (all four indicated sites) concerning the Central Bushveld and Lowveld savanna bioregions in South Africa (map courtesy of Wynand Muller, metadata provided by SANBI (2012), Municipal Demarcation Board: https://dataportal-mdb-sa.opendata.arcgis.com/datasets/37d790cf4f3b4b0ebd0ac501eae9b17d_0/explore?location=29.911228%2C26.005192%2C5.96 and the National Geographic Society: https://services.arcgisonline.com/ArcGIS/rest/services/NatGeo_World_Map/MapServer).

Plant material collection, DNA isolation, PCR amplification and sequencing

Leaf material from 48 pre-selected plant species (Supplementary Table 2) was collected between 26 March and 6 April 2015. For DNA extraction, only the top 10 cm of young, fresh leaves were used to eliminate secondary compounds prevalent in older vegetative components that would impede the isolation of a suitable DNA template. Five large leaves or several smaller leaves and herbaceous twigs were sampled and subsequently dried in silica gel in Falcon tubes for storage until DNA extraction. During sampling,

gloves were worn to prevent contamination. Except for the eventual DNA extraction of these dried samples in a laboratory, DNA was also isolated from fresh leaf material using the Qiagen DNeasy Plant Extraction kit following the manufacturer's protocol. In the laboratory facilities at the North-West University (Potchefstroom Campus), 0,10 - 0,15 g of dried plant material was ground into a powder with a mortar and pestle in liquid nitrogen. DNA extraction was again facilitated with the Qiagen DNeasy Plant Extraction kit following the manufacturer's protocol. PCR (polymerase chain reaction) amplification of the *rbcL* and *trnL* barcoding regions was done at the laboratory facilities of the North-West University (Potchefstroom Campus), respectively, with the primers *rbcLa*-F: ATGTCACCACAAACAGAGACTAAAGC (Levin *et al.*, 2013) and *rbcLa*-R: GTAAAATCAAGTCCACCRCG (Kress & Erickson, 2007) as well as with the *trnL* primers (Taberlet *et al.*, 1991), *trnL*-c: CGAAATCGGTAGACGCTACG and *trnL*-d: GGGGATAGAGGGACTTGAAC. Sanger sequencing of the isolated DNA fragments was outsourced to the Central Analytical Facilities (CAF) at the University of Stellenbosch, and the resulting sequences were processed and submitted to GenBank, as described in Chapter 3.1.2.

Faecal sample collection, DNA isolation, PCR amplification and sequencing

The data acquired for the metabarcoding of dietary profiles were sourced from the faeces of 24 individual cattle from Welverdiend, as well as 24 individual cattle from the Syferkuil Experimental farm during the dry season, before the first rains of the new season. The faecal samples from Welverdiend were collected in October 2018 and those from Syferkuil in October 2019. Great care was taken to ensure that these faecal samples were representatives of individuals either by visual observation or by allowing a sufficient gap in time and space in the collection efforts. Approximately 10g of fresh faecal material was collected per individual and ~200mg of this was used for subsequent DNA isolation. Before DNA extraction, faecal samples were stored in a fridge at 10° for no longer than three days. Each extraction series included 24 faecal samples per study site and was facilitated using the Qiagen DNeasy mini-Plant DNA Extraction kit according to the manufacturer's protocol. The isolated DNA was kept frozen at -0 °C until PCR amplification with the *rbcL* and *trnL* plant chloroplast barcode regions, respectively, with the primers *rbcLa*-F: ATGTCACCACAAACAGAGACTAAAGC (Levin *et al.*, 2013) and *rbcLa*-R: GTAAAATCAAGTCCACCRCG (Kress & Erickson, 2007) as well as with the *trnL* primers (Taberlet *et al.*, 1991), *trnL*-c: CGAAATCGGTAGACGCTACG and *trnL*-d: GGGGATAGAGGGACTTGAAC. The expected fragment sizes for 522 bp for *rbcLa*-F/R and 465 bp for *trnL*-c/d.

The *rbcL* and *trnL*(UAA) intron chloroplast gene regions from 48 faecal samples were amplified using the barcode's metagenomic sequencing library protocol (Illumina). Two PCR reactions were completed for *rbcLa* and *trnL*(UAA) library preparations and incorporated the Illumina overhang adapters onto the primers

	listed	above	
			(rbcLa-F: 5'-

TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGTCACCACAAACAGAGACTAAAGC-3'; *rbclL*-R: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTAATAATCAAGTCCACCRG-3'; *trnL*c: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGAAATCGGTAGACGCTACG-3'; and *trnL*d: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGGGATAGAGGGA CTTGAAC-3'). Each PCR reaction had a final volume of ~25 µl which consisted of 12,5 ng DNA template, 5 µl (1 µM) of the forward and reverse primer, respectively, 12,5 µl 2X Kapa HiFi HotStart ReadyMix (KAPA Biosystems Woburn, MA), and 2,5 µl PCR-grade water.

The PCR amplification of these reactions, as well as negative control reaction without any template DNA were carried out simultaneously as follows: heated lid 110 °C, 95 °C for 30 seconds, 25 cycles of 95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds, then 72 °C for 5 minutes and held at 4 °C. Gel electrophoresis (1X TAE buffer, 1.5% agarose, 100 V) was used to visualise the PCR results and successful PCR products were purified using AMPure XP magnetic bead-based purification (Beckman Coulter, IN). Multiplexing was made possible by the Illumina Nextera XT Index kit (Illumina, San Diego, CA) using two 8-base indices. Two indexing primers per sample were utilised and PCR reactions with the purified DNA template (5 µl DNA template, 5 µl x2 index primer, 25 µl 2x KAPA HiFi HotStart ReadyMix and 10 µl PCR-grade water) were performed on a Bio-Rad C1000 Thermal Cycler (Bio-Rad, Hercules, CA). The cycling conditions consisted of an initial denaturation cycle of 95 °C for 3 minutes, followed by eight cycles of denaturation (95 °C for 30 seconds), annealing (55 °C for 30 seconds), and extension (72 °C for 30 seconds). The cycle was completed with a final extension cycle of 72 °C for 5 minutes. The indexed libraries were purified with Ampure XP beads before library pooling, and their quantities were determined using the Qubit dsDNA HS Assay Kit from Thermo Fisher Scientific (Waltham, MA). Prior to sequencing, the quality of the purified amplicons was evaluated using the Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA). Following instructions from Illumina, the sample pool (2 nM) was denatured with 0.2 N NaOH, diluted to 7 pM, and mixed with 10% (v/v) denatured 20 pM PhiX. Samples were sequenced on the Illumina MiSeq platform (2 x 300 bp) by the NICD Sequencing Core Facility service, Pretoria, South Africa.

Chapter 3 A *rbcL* and *trnL* reference library to aid in plant identifications of a metabarcoding dietary analysis of herbivores from an eastern semi-arid South African savanna

Introduction

Dietary analysis is a fundamental part of constructing habitat selection and utilisation models as well as assessing the influence of land use type on the plant community and how this, in turn, can influence foraging strategies (Valentini *et al.*, 2008; Kartzinel *et al.*, 2015). The determination and analysis of food items in each ecosystem will also aid in the identification of key environmental resources for the design of reliable conservation and management strategies (Kowalczyk *et al.*, 2011; Gebremedhin *et al.*, 2016). To determine the composition of animal diets and how this composition reflects the plant community in question, animal faeces can be examined to identify food items (Kartzinel *et al.*, 2015; Lee *et al.*, 2018), as well as the relative abundance thereof ingested per individual sampled (Deagle *et al.*, 2019).

The metabarcoding approach allows for species identification from heterogeneous and degraded environmental samples based on the amplification and subsequent analysis of DNA barcodes (Pegard *et al.*, 2009; Bhargava & Sharma, 2013), which are short gene sequences from a standardised region of the genome that displays low intraspecies and high interspecies variability (Hebert *et al.*, 2003; Galan *et al.*, 2018). DNA metabarcoding provides an attractive alternative to the traditional methods of large-scale species identifications and it can be applied to microscopic (Kang *et al.*, 2021), cryptic (Collins & Cruickshank, 2013), and digested materials (Kartzinel *et al.*, 2015; Mallot *et al.*, 2018).

The *rbcL* and *trnL*(UAA) DNA barcodes have been proposed as two single-locus gene regions that can be used for the analysis of herbivorous- and omnivorous animal diets via DNA metabarcoding (Taberlet *et al.*, 2007; Valentini *et al.*, 2009; Kartzinel *et al.*, 2015; Kress *et al.*, 2015; Mallot *et al.*, 2018). These two barcodes were chosen as the core barcoding regions for this study due to their coverage of different savanna plant species and because of their sequence availability on global databases.

Except for the choice of the barcoding marker to be used, the efficiency and usefulness of metabarcoding as a tool are underpinned by the extent of taxonomic coverage and the quality of records available in the DNA barcode reference database to which the query sequences will be compared and identified (Matesanz *et al.*, 2019; Ando *et al.*, 2018). However, there is a lack of plant taxonomic sequence information for some ecological regions on global DNA sequence databases (Gill *et al.*, 2019). According to Gill *et al.*, (2019), as of 2019 DNA sequences available on the Barcode of Life Data (BOLD) database represent only ~8% of the vascular plants in Africa, while Africa comprises ~20% of the earth's landmass and owns approximately

15% of the global plant diversity (Lindler, 2014). Furthermore, forbs (i.e., non-graminoid herbaceous vascular plants: Scott-Shaw & Morris, 2015; Zaloumis & Bond, 2016), which contribute more than 70% to herbaceous species richness in grasslands, temperate deciduous forests, and savanna ecosystems, have been neglected in vegetation studies and knowledge of their ecological role in savanna ecosystems are limited (Siebert & Dreber, 2019). As a result, forbs are also badly represented on global DNA sequence databases.

Furthermore, it is well known that public DNA reference databases may contain sequencing errors and unverified taxonomies due to a lack of voucher specimen information and links to metadata (Mulcahy *et al.*, 2022). This is especially true for GenBank where the curation of the database is largely left to scientists submitting sequences (Harris, 2003; Hollingsworth *et al.*, 2011). Since the DNA reference database will determine the accuracy of taxonomic assignments, i.e., the success of a barcode in distinguishing between species (Meyer & Paulay, 2005), it is necessary to create a subset of the data available on public DNA databases that are manually curated for sequence quality, taxonomic validity, geographic validity as well as to test the efficiency of this subset reference database to discriminate between and within species with similarity, distance- and tree-based methods.

Accordingly, for our larger study on the dietary analysis of herbivores foraging in a South African Savanna this study aims to: (i) create a comprehensive *rbcL* and *trnL*(UAA) DNA barcode reference database of plant species commonly found in an African Savanna that are potential foraging targets of herbivores, and (ii) evaluate the efficiency of this reference database in discriminating between and within species by testing for the presence of a barcode gap, identifying a data appropriate identification threshold as well as determining the identification accuracy of reference sequences via primary distance-based criteria. The fulfilment of these aims will result in a DNA reference database that can be confidently utilised to delineate between species found in the faeces of herbivores foraging in a South African Savanna as well as to use the successful taxonomic identifications to make certain deductions about the plant community to aid in management and conservation strategies.

Materials and methods

3.1.1 Study area

The study was conducted at two sites that represent different land-use types in semi-arid South African savannas. These study sites, i.e., Syferkuil and Welterdiend, have been described in the overall Materials and Methods section (Chapter 0). As stated, both study sites are representative of eastern semi-arid South African savannas and comprehensive species lists for both these sites were compiled, as described in Chapter 0, and could be used as a reference for sequence mining (see Supplementary Table 1).

3.1.2 Barcode reference list assembly

A subset of *rbcL* and *trnL*(UAA) records available on public databases was created by using study-area specific species lists to manually mine sequences according to specific quality criteria to ensure accurate taxonomic coverage and resolution. The sites used for the compilation of these species' lists have been identified and described in the overall Materials and Methods section (Chapter 0). The criteria for sequence mining included sourcing a maximum of three individuals per species per marker, not including sequences from the same herbarium, and giving preference to sequences with voucher specimens and sequences with the least ambiguous bases. Accession numbers of barcodes that match the quality criteria and have been sequenced with the *rbcL* primers, *rbcLa*-F: ATGTCACCACAAACAGAGACTAAAGC (Levin *et al.*, 2013) and *rbcLa*-R: GTAAAATCAAGTCCACCRCG (Kress & Erickson, 2007) as well as with the *trnL* primers (Taberlet *et al.*, 1991), *trnL*-c: CGAAATCGGTAGACGCTACG and *trnL*-d: GGGGATAGAGGGACTTGAAC, can be found in Supplementary Table 3. The public databases used to source sequences in this study were the GenBank nucleotide database (www.ncbi.nlm.nih.gov) and the BOLD System public database (www.boldsystems.org). Sequences were mined manually according to the species list and quality criteria. Their corresponding accession number (process ID in the case of BOLD) and taxonomic lineage were described in the FASTA header and all entries per marker were contained in separate FASTA files, therefore the reference database consists of two FASTA files, one for *rbcL* and one for *trnL*.

This study also enriched the sourced dataset by supplying twenty-four sequences from plants found in the South African Savanna ecosystem. Most of these sequences are known to not occur at all, or not frequently on public DNA sequence databases. A pipeline for the collection, DNA isolation, and sequencing of *rbcL* and *trnL* barcodes has been established in 2015 (Chapter 0). Leaf material and voucher specimens of woody and herbaceous plants were collected throughout the study sites (Welverdiend, Timbavati and Nkuhlu) as well as neighbouring landscapes. Expert botanists at the NWU examined the voucher specimens to confirm their taxonomy and the voucher specimens were entered into the local AP Goossens Herbarium. The barcode sequences were added to the GenBank nucleotide database and included in the reference database compiled and evaluated in this study. GenBank accession numbers for the *rbcLa* barcode range from MZ461574.1 to MZ461594.1 and the *trnL* barcode from MZ461547.1 to MZ461570.1 (Supplementary Table 4).

3.1.3 Phylogeny

The candidate barcode reference sequences for both markers to be used for phylogenetic analysis were respectively aligned using Geneious (v. 2021.2.2) via MULTiple Sequence Comparison by Log-Expectation (MUSCLE: Edgar, 2004), with a maximum number of eight iterations. The sequences for the *trnL* marker

were not split into families, as they will be for downstream analysis, due to the requirement of robust representations of the alignment and clusters formed for the subsequent “quality” analysis.

The aligned candidate barcode reference sequences were used to reconstruct neighbour-joining (NJ) phylogenetic trees for the *rbcL* and *trnL* dataset, as implemented by Geneious with bootstrap testing of 1000 replicates using the Jukes-Cantor model. Within angiosperms, the genus *Amborella* is reported as the sister to all remaining flowering plants (Drew *et al.*, 2014; Gitzendanner *et al.*, 2018), and accordingly, the *rbcL* and *trnL*(UAA) gene regions of *Amborella trichopoda* were included as an outgroup for the *rbcL* and *trnL* sequences and aligned with the rest of the datasets. The individuals in both the *rbcL* and *trnL* phylogenetic tree were also reviewed manually to validate clusters and sub-clusters that were formed. This evaluation was put in place to improve further analyses recommending the removal of sequences showing an evident deviation from their cluster, be it a cluster of life forms (such as trees, shrubs, grasses, forbs) or plant families. The sequences that were identified to have problematic placements were then replaced with another individual sequence of the species with equal quality if possible, or it was removed from further analysis.

3.1.4 Assembly of the DNA barcode reference database

DNA sequences for both marker datasets were exported to RStudio (RStudio Team, 2022) and aligned to prepare sequences for downstream analysis. Alignment was again facilitated with MUSCLE, but without the outgroup sequences, and subsequent determination of genetic distance and barcode-gap analysis took place. All sequences for *rbcL* were aligned simultaneously using the *AlignSeqs()* function in the DECIPHER package (v2.0: Wright, 2016), and *trnL* sequences were split into families (Gill *et al.*, 2019) and accordingly aligned with the same function due to the often highly divergent sequences displayed for distantly related species. Distance matrices for both datasets were calculated using the Kimura-2 parameter (K2P) model (Kimura, 1980) which is the build-in model of evolution in the APE package’s *dist.dna()* function (v.5.0: Paradis & Schliep, 2019).

3.1.5 Barcode gap analysis

The barcode analysis for each marker dataset used the K2P model to calculate genetic distances for each sequence per marker and was done using the R package SPIDER (v.1.5: Brown *et al.*, 2012). A barcode gap analysis is performed by two approaches, as recommended by Niemann *et al.* (2022), which will allow for barcode gap analyses of the complete datasets, as well as barcode gaps for individual sequences represented in the respective datasets. The first approach deals with comparing the median of inter – and intraspecific distances for each marker dataset as well assessing the significance of the differences between genetic distances with a Wilcoxon rank-sum test with the R function *wilcox.test()* in the R STATS package

(R Core Team, 2022). The second approach follows the determination of the presence of a barcode gap following Meier *et al.* (2008), where the difference between maximum intraspecific distance and the minimum interspecific distance is calculated for each sequence per marker dataset. Both approaches were applied during this study.

3.1.6 Analysis with distance-based criteria

The optimisation of the species identification threshold for each marker dataset is performed to increase identification success (Meyer & Paulay, 2005). During this study, we used the *threshOpt()* function in the SPIDER package of R (v.1.5: Brown *et al.*, 2012) to systematically subject the marker dataset to a range of threshold values (between 0.1 to 0.2% K2P distance). The optimal identification threshold was then calculated based on the number of true positive, false positive, true negative, and false negative identifications for each threshold in the range. The optimal threshold was identified as the value predicting the lowest cumulative error, i.e., the lowest amount of false-positive and negative identifications according to Brown *et al.* (2015).

The predicted accuracy of taxonomic assignment of the reference sequences was determined using three distance-based analyses via the SPIDER package. These primary distance-based criteria included the nearest neighbour (*k*-nn) criterion (Duda *et al.*, 2001), Meier's best close match (BCM) criterion (Meier *et al.*, 2006), as well as BOLD identification criterion (Ratnasingham & Hebert, 2007). The functions by which these measures are executed (*nearNeighbor()*, *bestCloseMatch()*, and *threshID()*), respectively performed an internal taxonomic assignment, where individual sequences were treated as unknown queries and the rest of the marker dataset was treated as the DNA reference database to match corresponding taxonomies within the identified optimal threshold.

Results

3.1.7 DNA reference dataset: phylogenetic analysis and summary statistics

The final species list (Supplementary Table 1) consists of 755 species, 356 genera, 99 families within 3 classes (Magnoliopsida, Liliopsida, and Polypodiopsida) in the phylum Tracheophyta. Of these 755 species, sequences for 108 could not be found on public sequence libraries for either the *rbcL* or *trnL* barcode. Upon inspection of the NJ trees of *rbcL* and *trnL*, it was advised to replace 265 individuals with other sequences from the same species of equal quality since they did not group with other members of their family or life form. Sequences for 59 individuals could be replaced, but 206 sequences had to be removed. This led to a loss of taxonomic coverage in some species, now only represented by < 3 individuals and the complete loss of 57 unique species from the reference dataset. The remaining species and accession numbers, or process

identities (process IDs) in the case of BOLD sequences, can be found in Supplementary Table 3. which now constitutes the complete DNA reference database. The composition of the complete reference database is illustrated in Figure 3.1, where the family Poaceae is the largest representative with 47 genera and 109 species. The second-largest representative family is the Fabaceae with 35 genera and 87 species. The *rbcL* reference dataset consists of 1238 sequences representing 318 genera and 562 species. The *trnL* dataset consists of 921 sequences which represent 270 genera and 461 species. Naturally, there is an overlap between the datasets as sequences were mined for the same species for both markers: sequences for 345 species of 224 genera are found in both marker datasets.

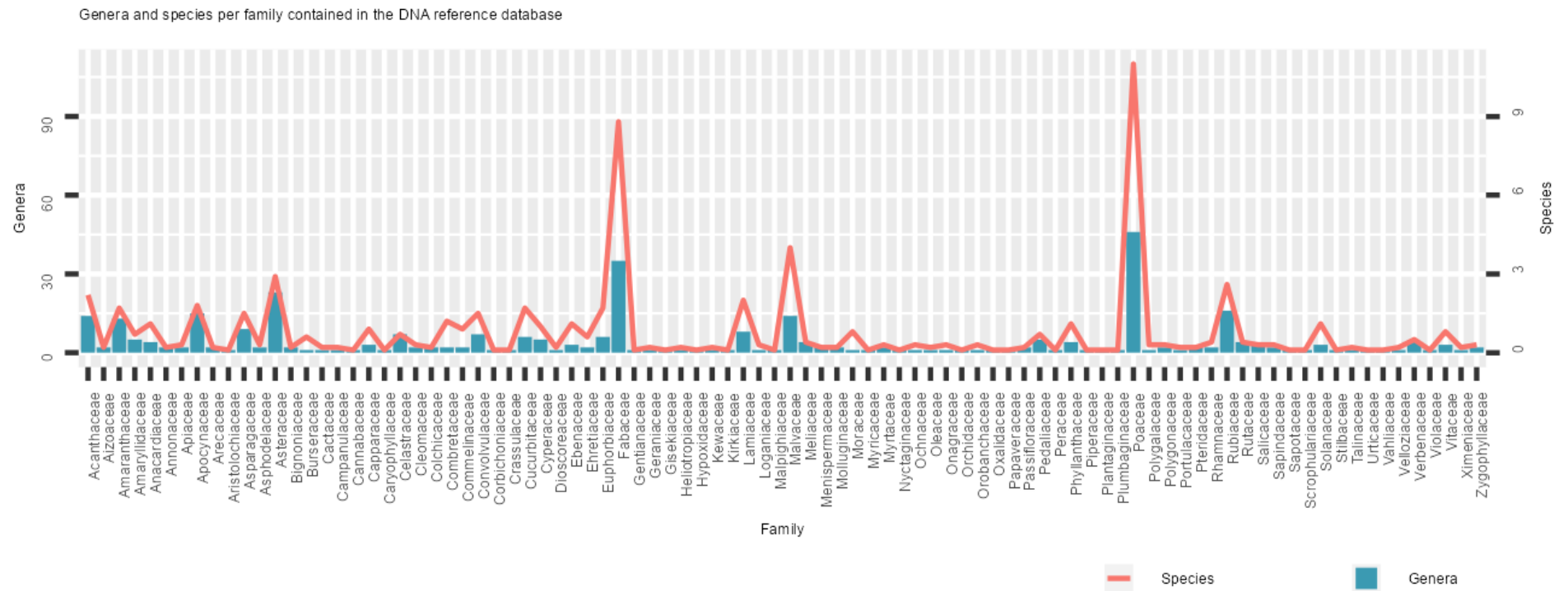


Figure 3.1: The composition of the final DNA sequence reference database in terms of number of genera (histogram) and species (line plot) per family

3.1.8 Reference database analysis

According to the first approach of the barcode gap analysis, by which the presence of a barcode gap was assessed for both marker datasets by the comparisons of medians for inter- and intraspecific genetic distances, it was found that for the *rbcL* dataset, the interspecific K2P distances ranged from 0,00 – 0,94 with a median of 0.01. The interspecific distances were statistically larger than the intraspecific K2P distances (Wilcoxon test: $P < 2.2e^{-16}$) with ranges of 0 – 1.45 and a median of 0.00. The *trnL* dataset followed the same trend with generally larger interspecific distances (ranges from 0,00 – 1,01 and a median of 0,01) than intraspecific distances (ranges: 0 – 1,07; median: 0; Wilcoxon test: $p < 2.2e^{-16}$). The second approach (Meier *et al.*, 2006), revealed a positive difference between maximum intra- and minimum interspecific K2P distances for 76% of the *rbcL* sequences and 68% for the *trnL* sequences. These results are reported as scatterplots (Fig. 3.2) where the sequences that showed a gap appear above the 1:1 slope.

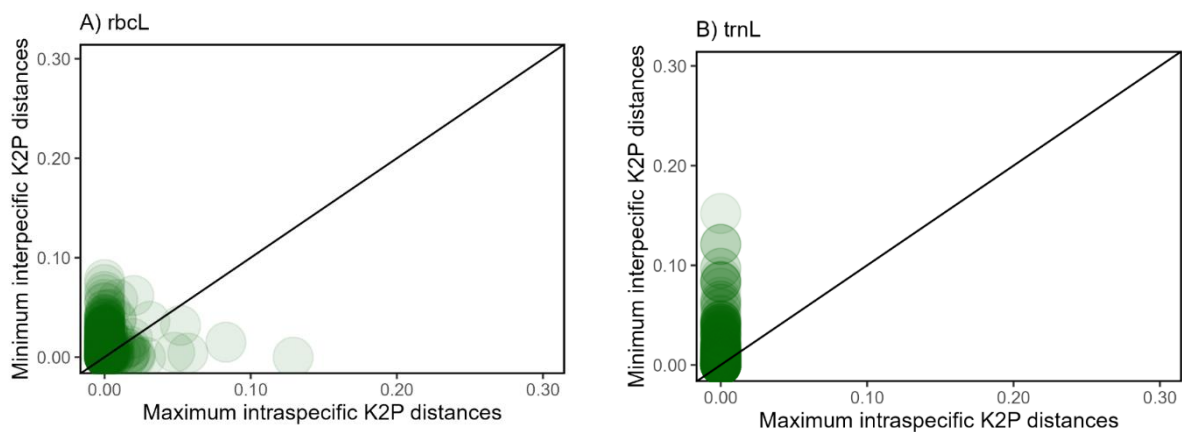


Figure 3.2: Scatterplots of the Barcode gaps for species present A) in the *rbcL* dataset and (B) *trnL* dataset. Scatterplots show the relationships between minimum interspecific distance (y-axis) and maximum intraspecific distance (x-axis) where the barcode gap is indicated by points falling above the 1:1 slope. High densities of plots are indicated by the darker colours.

To improve the accuracy of the DNA reference database, optimal identification thresholds were determined with the *threshOpt* function in the R package SPIDER. The default threshold of species optimization used by identification algorithms such as BOLD is 1%, which may not always be appropriate for all reference datasets (Brown *et al.*, 2012). Accordingly, the determined optimal threshold for the *rbcL* dataset was 0.4% and that of the *trnL* dataset was 0.6%, which clearly indicates that the identification accuracy of the datasets would not have been accurately portrayed by the default threshold value of 1%. These optimal threshold values will be used during the simulated taxonomic identification via the three distance-based analyses which will aid in predicting the accuracy of the reference datasets.

The nearest-neighbour (*k*-NN) criterion performed the best among the distance-based measures during the identification simulations (Table 3.1). The identification success rate, calculated as the number of matches that share the same species index as the query was 85.86% for the complete *rbcL* dataset and 73.72% for the *trnL* dataset. These identification success rates can be increased by excluding singletons, resulting in an identification success rate of 90.78% and therefore an increase of 4.92% for the *rbcL* dataset. Excluding singletons led to an increase of 5.33% resulting in a success rate of 79.05% for the *trnL* dataset. Meier's best close match (BCM) criterion performed similarly to the *k*-NN identification criterion with an identification success rate of 82.66% correct matches with the optimised threshold value for the exclusion of singletons for the *rbcL* dataset and 64.03% for the *trnL* dataset with matches performed in the optimised threshold with the exclusion of singletons. The BOLD identification criterion failed to show the same level of identification success rates with the inclusion of an optimised identification threshold. The BOLD identification criterion performed the worst among the other measures with the highest discriminating power shown for the *trnL* dataset with the exclusion of singletons, reaching 53.20% with the default threshold value of 1%.

Table 3.1: Discriminatory power of the *rbcL* and *trnL* reference datasets predicted by the application of three distance-based measures: nearest neighbour, Meier's best close match, as well as BOLD identification criterion, for both default and optimized thresholds for the inclusion and exclusion of singletons for the respective datasets. All three analyses were performed to identify queries within the respective identification thresholds to genus level.

Dataset	Single-tons	Total number of taxa	Threshold value (%)	Nearest Neighbour *		Best Close Match **				BOLD Identification ***			
				False	True	Ambiguous	Correct	Incorrect	No ID	Ambiguous	Correct	Incorrect	No ID
<i>rbcL</i>	Included	1238	1	175	1063	104	968	72	94	498	604	42	94
			0,4	NA	NA	102	877	38	221	194	794	29	221
	Excluded	1171	1	108	1063	104	968	48	51	498	604	18	51
			0,4	NA	NA	102	877	29	163	194	7943	20	1631
<i>trnL</i>	Included	921	1	242	679	49	550	93	229	183	457	52	229
			0,6	NA	NA	49	508	77	287	129	459	46	287
	Excluded	859	1	180	679	49	550	76	184	183	457	35	184
			0,6	NA	NA	49	508	63	239	129	459	32	239

NA – The identification simulation using the nearest-neighbour criterion does not allow for supplying an alternative identification threshold and only uses the build-in default of 1%.

*Nearest-Neighbour (*k*-NN) criterion result definitions:

False – the nearest species index is not the same as the tested individual; True – the nearest species index is the same as the tested individual. Note that the optimal threshold could not be changed from its default in the `nearestNeighbour()` function.

**Best Close Match (BCM) criterion result definitions:

Ambiguous – Correct and incorrect species are identified as closest matches within the identification threshold; Correct – all close matches within the identification threshold of the query are the same species; Incorrect – all matches within the identification threshold are different species than the query; No ID – no matches could be made within the identification threshold for any individual.

***BOLD identification criterion result definitions:

Ambiguous – Correct and incorrect species are identified as closest matches within the identification threshold; Correct – all close matches within the identification threshold of the query are the same species; Incorrect – all matches within the identification threshold are different species than the query; No ID – no matches could be made within the identification threshold for any individual.

Discussion

Despite several studies published, utilising DNA barcoding in the analyses of herbivore diets (Goldberg *et al.* 2020; Kowalczyk *et al.*, 2019; Iwanowicz *et al.*, 2016), very few focus on herbivore diets in African savannas (Kartzinel *et al.*, 2015; Gill *et al.*, 2019). To the authors' knowledge none focused on South African savannas, and there is not currently a robust DNA reference sequence library available. During this study we have curated a DNA reference library consisting of both the *rbcL* and *trnL* datasets. This is the first step towards facilitating metabarcoding studies aimed at South African savanna ecosystems, specifically for the identification of plant materials obtained in herbivorous faeces. The compilation of this DNA reference database of species found in a South African savanna has been challenging since scientific publications listing all taxa likely to be encountered in the diets of herbivores foraging in these ecosystems were largely unavailable.

The ideal construction of the reference library would be to sample, identify, and sequence the relevant barcodes of all the species on the species list. However, this would be a costly and laborious research campaign. Mining sequences according to comprehensive, geographically constrained species lists (Supplementary Table 1 and Chapter 0) from global databases is a much more cost-effective option and the method illustrated here can be reproduced for any nature of metabarcoding study aimed at representing the composition of an ecosystem without manually sequencing all specimens. The use of GenBank as a taxonomic resource has been questioned since there exists an absence of preserved voucher specimens, non-justified species identifications, and low-quality data (Meier *et al.*, 2006; Meiklejohn *et al.*, 2019). The same is true for BOLD, although this database is better curated due to higher quality submission standards, it may still contain incorrectly identified taxa and low-quality sequence data (Meier *et al.*, 2006; Ratnasingham & Hebert, 2007). Being aware of these limitations enables the user to create certain download

criteria which result in a database of sequences with known localities, known voucher specimens as well as the inclusion of barcodes with the least ambiguous bases. This database, as well as any database mined from global reference databases, is not impervious to imperfection, but this method of using a species list and applying download quality criteria will certainly lift its standard above a taxonomic assignment against whole databases such as GenBank and BOLD. The lowest level of taxonomic resolution achieved in metabarcoding studies is typically genus-level (Goldberg *et al.*, 2020; Turunen *et al.*, 2021) and the assignment of taxonomy at lower taxonomic levels is often infeasible and inaccurate. Therefore, to ensure that this database of barcode sequences is robust enough to be used as an identification tool the barcode gap was analysed together with the barcode phylogeny.

Disparities between inter-and intraspecific distances between and among species are defined as a barcoding gap that, if present at a locus, enables the reliable differentiation of species (Meyer & Paulay, 2005; Niemann *et al.*, 2022). A barcode gap was evident in both approaches used and for both reference sequence datasets. A barcode gap for 76% of the species was obtained for the *rbcL* barcode sequences and for 68% of the species for the *trnL* sequences. This implies that reliable species differentiation would not be possible for 24% of the species in the *rbcL* dataset and 32% in the *trnL* dataset. However, limitations have also been reported in other DNA-barcode studies such as Gill *et al.*, (2019); Mishra *et al.*, (2017), and Kress and Erickson, (2007). Gill *et al.*, (2019) reported a barcode gap for only 73% of the species investigated for the *rbcL* primer and 79% for the *trnL*-F primer in their barcode library of semi-arid East African savanna plants. Mishra *et al.*, (2017) have shown that even within one genus (*Terminalia*) the three different barcodes of *matK*, ITS and *rbcL* contained a barcode gap for <70% of the species. As stated by Gill *et al.*, (2019) savannas are comprised of a diverse range of species that are prone to the absence of a barcoding gap.

The distance-based measures of *k*-nn, BCM, and BOLD identification criteria infer identification success rates by considering the K2P distance matrix of each dataset to simulate taxonomic identifications of one sequence against the rest of the databases to match taxonomies within the identification threshold (Brown *et al.*, 2012). Usually, the distance-based matrices are applied at a species level, but for this metabarcoding study, it was decided to predict the identification success of the datasets based on genus-level identifications. The best success rates across all three methods were seen with the exclusion of singletons, in this case, genera represented by only one individual. The analysis of the identification of success rates, with the exclusion of singletons, is also a prerequisite for the alignment that precedes the barcode gap analysis. Singletons are typically a problem in barcoding studies since the identification simulations will treat the singleton as a query, and it will not have a match available in the reference dataset and will result in either “incorrect” or “no identification” (Brown *et al.*, 2012).

Analyses revealed that the *rbcL* dataset has a predicted identification success rate of 90.78% and *trnL* 79.05% with the *k*-NN method. This would imply that a barcode gap analysis is not always an accurate predictor of the species discrimination success of sequences in a reference database, as concluded in certain studies (Cognato, 2006; Meier *et al.*, 2008; Little *et al.*, 2013). This is also reflected by the findings in this study as illustrated by the inconsistencies between identification success rate simulations and the evaluation of the barcode gap. However, in this case, the barcode gap analysis can be accepted as a more efficient tool to predict database accuracy since it displays the identification success rate achieved by species-level identifications, whereas the distance-based analysis implemented genus-level taxonomic assignments. Furthermore, the discrimination success rates demonstrated by both datasets are above (*rbcL* = 76%) or very close (*trnL* = 68%) to the discrimination success rate of 72% as proposed by the CBOL Plant Working Group (2009). This implies that an adequate discrimination success rate would be possible to the lowest taxonomic level of genus when taking the predicted identification success rates into account.

Among the parameters tested, *k*-NN and BCM performed the best for both the *rbcL* and the *trnL* dataset with success rates for *rbcL* being 90.78% and 82.66%, and for *trnL* being 79.05% and 64.03% for the exclusion of singletons and identifications within the optimal threshold value. *K-nn*'s higher success rate can be attributed to the return of the majority-based identification in cases of a tied closest match, where BCM will return this as an “ambiguous” result. BOLD identification criteria fared the worst with *rbcL* identification success rates reaching 67.81% with the exclusion of singletons at the default threshold and 53.43% for *trnL* with the same parameters. The BOLD identification criteria require that all matches within the identification threshold must be conspecific rather than a single closest match, as in *k*-NN and BCM, for it to return a “correct” result. One misidentification will therefore result in an “ambiguous” result, explaining its poor performance in this study.

The phylogenetic analysis served as a re-identification strategy of sequences included in the *rbcL* and *trnL* reference database. The exclusion of singletons, as recommended by the identification rates of the various distance-based measures (Table 3.1) was not considered in this part of the study, as this would have led to a great loss of taxonomic coverage and identifications as seen by a drop of 5.41% and 7.22% in the total taxonomy included in the *rbcL* and *trnL* databases, respectively. Excluding singletons from the reference datasets would lead to a loss of the families of flowering plants, namely: Aristolochiaceae, Dioscoreaceae, Gentianaceae, Kirkiaceae, Passifloraceae, Piperaceae, Plantaginaceae, Urticaceae in the *rbcL* datasets as well as the loss of Hypoxidaceae, Menispermaceae, Nyctaginaceae, Passifloraceae, Piperaceae, Rutaceae, Scrophulariaceae and Stilbaceae in the *trnL* dataset. We, therefore, sacrificed some identification success for the inclusion of singletons, and therefore an accurate representation, in terms of species present, in a South African savanna. The NJ trees revealed that sequences formed cohesive clusters of orders for the *rbcL* dataset. However, an obvious low taxonomic resolution for some orders in the *trnL* dataset was again obvious, such as for the orders Poales and Commelinales and Fabales. Similar results were also shown by

Gill *et al.*, (2019) for the families Poaceae, Fabaceae, and Malvaceae. Low taxonomic resolutions were expected since many genera which are known to occur in South African savannas are under-sampled in global reference databases or are yet to be barcoded with either *rbcL* or *trnL*, leading to a lack of representatives for certain genera or the inclusion of singletons in the reference database. Species represented in the South African savannas are very diverse as can be seen from the sum of the branch lengths of the NJ trees, namely 108.23 for the 922 *trnL*-barcode taxa and 382.34 for the 1239 *rbcL*-barcode taxa.

Conclusion

During this study a robust DNA barcode reference library was developed for the vegetation found in the eastern semi-arid South African savannas. The library covers a wide diversity of plants that all form an intrinsic part of the ecosystems depending on them. All the different tests used to validate the use and accuracy of this library indicate that it can be used with confidence to assign taxonomies for plants found in the region. We hope that it will add to other similar research conducted not only on the local flora but also to the work done on savannas elsewhere on the African continent. The datasets for *rbcL* and *trnL* are not presented as a complete DNA reference library, but rather as two databases that should be used in unison to identify species foraged on in the eastern semi-arid South African savanna. The work in this chapter has further been submitted to *PLos One* with the title: Introducing an *rbcL* and *trnL* reference library to aid in the metabarcoding analysis of foraged plants from semi-arid eastern South African savannas, and is currently under review (Annexure 3).

Chapter 4 Pipeline and filtering validation via a mock community-based analysis

Introduction

Accurate estimation of trophic interactions enables ecologists to understand ecosystem functions (Duffy, 2007), especially those involved with free-roaming livestock. They contribute to the shape and diversity of terrestrial ecosystems (Hayes & Holl, 2003) and the reconstruction of their diets may give crucial insights into field management practices (Valentini *et al.*, 2008; Shipley *et al.*, 2009; Ter Schure *et al.*, 2021). Many recent studies have employed DNA metabarcoding (Hebert *et al.*, 2003) to reconstruct herbivorous diets (Kartzinel *et al.*, 2015; Scasta *et al.*, 2019; Kowalczyk *et al.*, 2019; Ter Schure *et al.*, 2021; Potter *et al.*, 2022; Meyer *et al.*, 2022). Despite this technique's widespread use, the degraded nature of DNA recovered from faeces (<150 bp) represent non-optimal targets for PCR amplification, resulting in low taxonomic resolution and species abundance data that can often not be quantified (Scasta *et al.*, 2019; Ando *et al.*, 2020). Except for this inherent nature of the DNA to be analysed, certain biases prevent the accurate recovery of the plant community and may be introduced in every step of the experimental workflow.

Both biological and technical biases can be introduced in metabarcoding studies. Biological bias can be introduced before the collection of samples in the form of contamination or differential degradation across ingested species and therefore DNA densities (Deagle *et al.*, 2010). Technical bias may be introduced by varying efficiencies of DNA extraction methods, primer mismatches, and unequal primer binding which may lead to overestimated or underestimated species proportions in the final diet analysis (Alberdi *et al.*, 2018; Deagle *et al.*, 2014). Due to the common occurrence of these biases, many studies rather opt to interpret sequence counts as the frequency of occurrence (FOO), or presence/absence data instead of relative species abundances (Kartzinel *et al.*, 2015; Chua *et al.*, 2021). Although species lists reporting taxa found within samples as well as the frequencies in which they occur can provide quantitative information, it presents some limitations in that all dietary components are equally weighted which may not give an accurate representation of major or minor food items in the diets of animals (Deagle *et al.*, 2019). Reporting the relative read abundance (RRA) would therefore be a superior metric, and, as Deagle *et al.* (2018) stated, RRA can be more accurate than FOO in the quantification of dietary items even with moderate recovery biases.

Several of the above-mentioned biases can be mitigated during the bioinformatics analysis step with the implementation of appropriate filters to ensure firstly, that all species present in the diet may be reflected in the final taxonomic assignments and secondly ensure that the biomass of the ingested food items may be reflected by the number of reads resulting from these filters (Mathon *et al.*, 2021). Many pipelines have been developed that can be used for processing DNA recovered from faeces. These pipelines include,

among others MOTHUR (Schloss *et al.*, 2009), QIIME (Caporaso *et al.*, 2010), DADA2 (Callahan *et al.*, 2016), OBITOOLS (Boyer *et al.*, 2016), USEARCH (Edgar, 2010), its open-source alternative VSEARCH (Rognes *et al.*, 2016), and Claident (Tanabe & Toju, 2013).

However, several aspects render the interpretation of the data difficult, which include: (i) the variable length of the plastid *trnL*(UAA) intron (245 – 767 bp) between taxa and high levels of sequence variability (Taberlet *et al.*, 2007; Jurado-Rivera *et al.*, 2009), (ii) the mentioned challenges associated with different plant barcoding regions (Chapter 1), (iii) the degraded nature of DNA recovered from faeces, and (iii) a lack of consensus regarding a standardised barcoding region for plants (Banchi *et al.*, 2020; Bhargava & Sharma, 2013; Pavan-Kumar *et al.*, 2015).

There is thus a need for bioinformatics pipelines and filtering options specific for the recovery of plant-based animal diets. Some general filtering guidelines for animal dietary analysis by faecal metabarcoding are available, as summarised in the review by Ando *et al.* (2020). These include recommendations such as the removal of sequences of low quality and the creation of Operation Taxonomic Units (OTUs) by defining optimal clustering thresholds, the most common being 97%. During the determination of the diets of European pond turtles (*Emys orbicularis*) via metabarcoding faecal analysis, Ducotterd *et al.* (2020) also noted that a clustering threshold greater than 97.6% to assign mock community members to a sequence cluster will result in fewer false-positive assignments. The clustering approach can approximate species numbers by creating “species equivalents” without taking reference database coverage into account (Cristescu, 2016). This clustering procedure of sequences based on a similarity threshold also acts as a filter that removes sequencing artifacts (Bálint *et al.*, 2016; Alberdi *et al.*, 2018).

A major shift in metabarcoding experiments has been moving from using OTUs to using Amplicon Sequence Variants (ASVs), with the advantage of the latter being able to account for real biological sequence variation by partitioning reads based on error models (Callahan *et al.*, 2017; Joos *et al.*, 2020). After the creation of species equivalents by either an OTU or ASV method, reads with copy numbers or percentages under a certain threshold are removed. A threshold of 1% is recommended (Iwanowics *et al.*, 2016; Kartzinel *et al.*, 2015; Deagle, 2017), which removes rare OTUs and therefore the probability of contamination (Ando *et al.*, 2018). However, Alberdi *et al.* (2018) noted that the removal of rare OTUs should be done with caution since truly rare food items may be excluded, and threshold values may differ among environments and target taxa. After quality filtering and assembling species proxies and their subsequent filtering, the remaining OTUs/ASVs are used as queries in the metabarcoding study which are compared to a reference database to retrieve their taxonomic classifications.

The aim of this study is therefore to identify an optimal bioinformatics pipeline with its succession of filtering software to identify food items, that can be found in the diets of cattle foraging in an eastern semi-arid South African savanna, using the custom *trnL* and *rbcL* reference database developed earlier in the

study (Chapter 3). The final taxonomic assignment of the chosen pipeline should first be able to recover all species consumed and secondly be able to quantify the relative abundance of species within the community of species from the animal diet with the special inclusion of forbs, which are generally neglected in Savanna definitions and management models. In this context, the use of mock communities representing the floristic diversity and their expected proportions found in herbivorous diets can be a useful tool to improve species proportion estimates in DNA metabarcoding studies (see Matezans *et al.*, 2019 and Chua *et al.*, 2021). This is achieved by using the INSILICOSEQ software package (Gourlé *et al.*, 2019) to create an *in-silico* metabarcoding dataset that resembles reads in proportions that are expected from real faecal samples. In investigations of mixed-species samples where the species differ according to relative abundance and degree of relatedness, mock communities serve as an essential positive control (Nguyen *et al.*, 2015). Synthetic faecal samples were created by using *rbcL* and *trnL* sequences of species found in the DNA reference database, created and discussed in Chapter 3 that will represent a sufficiently large taxonomic breadth but also include some closely related genera (Nguyen *et al.*, 2015) and different plant life forms such as trees, shrubs, grasses, sedges, and forbs. The mock community for this study is composed of 23 food items from both the chosen plant barcoding regions which are differentially grazed by domesticated herbivores. A combination of 72 bioinformatics pipelines and filtering options are compared in their efficiency to recover all the species from the *rbcL* and *trnL* mock community in their expected abundances. Except for the removal of rare sequence variants to represent real biological diversity and not PCR or sequencing errors, error-model-based partitioning of reads into ASVs and conventional clustering of reads into OTUs based on similarity were particularly investigated. A new post-clustering curation R (R Core Team, 2022) package was also included to test whether it improved the sensitivity, precision and/or accurate representation of species' proportions in the mock community.

Materials and methods

The flow diagram shown below in Figure 4.1 gives an overview of the workflow followed for each method evaluated in this chapter. Sequences of species chosen to represent the mock community were mined from GenBank, and BOLD. These sequences together with their expected abundances for each sample (see Fig. 4.2) served as input to INSILICOSEQ (Gourlé *et al.*, 2019) where MiSeq sequence reads are generated. These sequences are input into the bioinformatics pipelines of VSEARCH (Rognes *et al.*, 2016), OBITOOLS (Boyer *et al.*, 2016) and DADA2 (Callahan *et al.*, 2016) where quality filtering first takes place before the inference of operational taxonomic units (OTUs) or amplicon sequence variants (ASVs). The VSEARCH approach allows for reference-free clustering of sequences based on their similarity (97%) or clustering against a reference database based on similarities (97%) between query and reference sequences. Final filtering of the six resultant feature tables included abundance-based filtering (*RA*), feature-based filtering (*1* and *10*), combinations of these (*1* and/or *10* and *RA*) as well as curation of the unfiltered and filtered OTU/ASV/unique feature tables with the LULU algorithm (Frøslev *et al.*, 2017). The application

of these steps resulted in a total of 72 OTU/ASV/unique feature tables, representing all the bioinformatics approaches to be compared according to certain criteria described in Section 4.2.5: Performance comparison criteria.

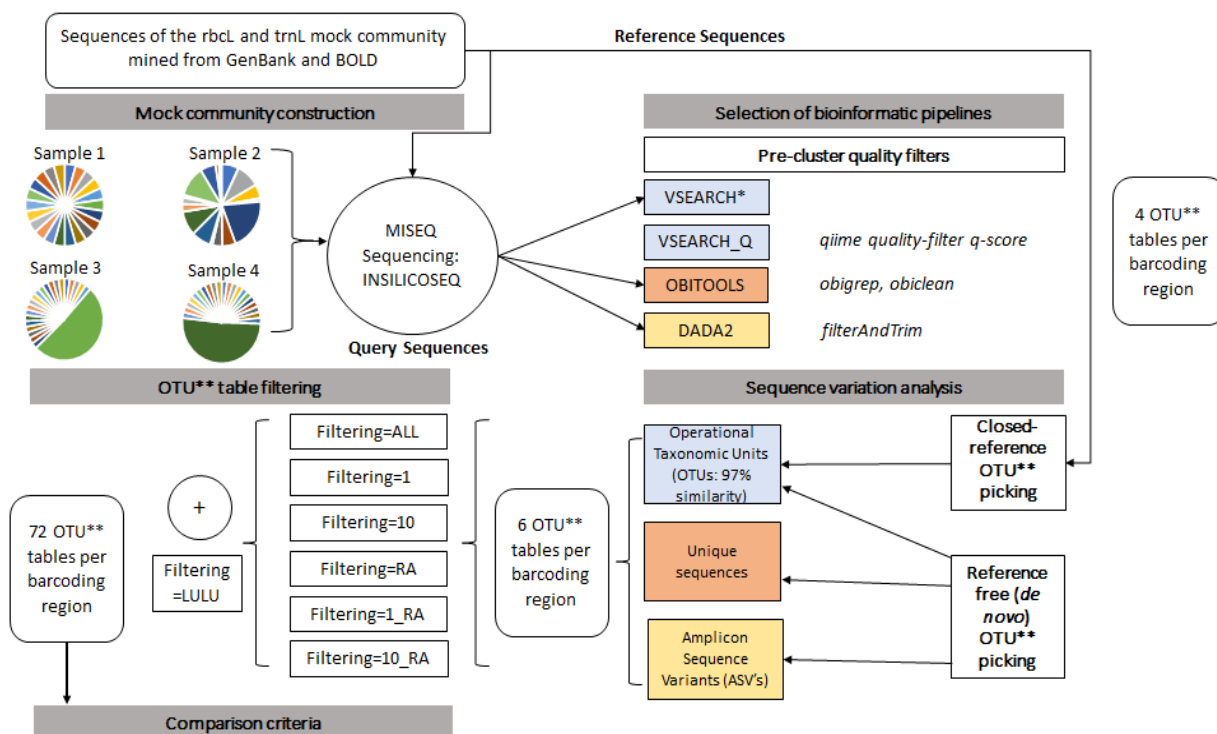


Figure 4.1: Overview of the main steps followed in the materials and methods section for the mock community analysis. * An external quality filter was added to the VSEARCH approach since this workflow does not incorporate one in its various steps. ** “OTU” in this diagram may refer to operational taxonomic units as retrieved from the clustering of sequences based on 97% similarity in the VSEARCH approach, amplicon sequence variants retrieved by the inference of an error model in the DADA2 approach or grouping of strictly identical sequences as in the OBITOOLS approach.

4.1.1 Mock community construction

Species selection

A series of savanna communities were simulated (Table 4.1) to confirm that the Illumina MiSeq platform could recover members from a known community structure. Species were selected based on the insights of a botanist’s knowledge of the floristic diversity of an eastern semi-arid South African savanna as well as the palatability and grazing trends of herbivores found within these ecosystems. The sequences of the selected species were sourced from the compiled reference database in Chapter 3 with their corresponding database sources listed in Table 4.1.

For this study, the chloroplast subunit ‘a’ of *rbcL* (552 bp: Levin *et al.*, 2003; Kress & Erickson, 2007) and whole chloroplast *trnL*(UAA) intron (254 – 767 bp: Taberlet *et al.*, 1991) were chosen as plant barcoding

regions to amplify plant taxa from herbivorous faeces. This choice was advocated by: (i) the standardised use of *rbcL* as a barcoding region (CBOL Plant Working Group, 2009), (ii) its large representation in reference databases, (ii) and its ability to identify query sequences up to family level. The *rbcL* marker formed part of a dual locus identification system with the second the short, and variable *trnL*. It was chosen as the second barcoding marker in this study since it has shown great success in dietary studies due to its ability to resolve taxonomy up to the species level from degraded samples (Lee *et al.*, 2018; Taberlet *et al.*, 2007; Ferri *et al.*, 2015; Yang *et al.*, 2014). It was decided to not merge forward- and reverse reads as generally conducted before data processing and after quality control in these bioinformatics pipelines. Data handling hereafter will therefore only be applied to unmerged, forward reads. The reason for this is that a minimum of 20 base pairs are required to overlap to create a consensus sequence and it would be unsuccessful for larger amplicons expected for the *trnL* intron. This choice is not unusual, especially when large amplicons or amplicons with highly variable lengths or low-quality reverse reads are present (Tan *et al.*, 2019; Pauvert *et al.*, 2019).

Table 4.1: Estimated species composition of the floral mock community which represents some dietary items of cattle foraging in an eastern semi-arid South African Savanna ecosystem. The relative abundances of species in the table equal 100% and the accession number (in the case of GenBank), and process identity (in the case of BOLD sequences) are indicated.

Species	Family	Life form	Abundance (%)	<i>rbcL</i> reference	<i>trnL</i> reference
<i>Sclerocarya birrea</i>	Anacardiaceae	Tree	7	MN216752.1 (GenBank)	KC479214.1 (GenBank)
<i>Tricalysia junodii</i>	Rubiaceae	Tree	0,2	JF265633.1 (GenBank)	AM999496.1 (GenBank)
<i>Combretum apiculatum</i>	Combretaceae	Tree	9,5	KC158503.1 (GenBank)	MK187024.1 (GenBank)
<i>Combretum zeyheri</i>	Combretaceae	Tree	5,5	JF265356.1 (GenBank)	MZ461553.1 (GenBank)
<i>Senegalia nigrescens</i>	Fabaceae	Tree	1	JX572216.1 (GenBank)	MK186955.1 (GenBank)

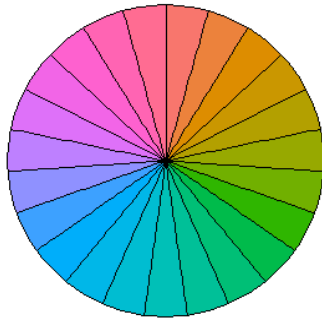
<i>Schotia brachypetala</i>	Fabaceae	Tree	0,3	JQ025087.1 (GenBank)	GQ405142.1 (GenBank)
<i>Dichrostachys cinerea</i>	Fabaceae	Shrub	20,5	JF265387.1 (GenBank)	PNG507-18 (BOLD)
<i>Grewia bicolor</i>	Malvaceae	Shrub	5,5	EU213485.1 (GenBank)	MK261494.1 (GenBank)
<i>Flueggea virosa</i>	Euphorbiaceae	Shrub	4	KP094635.1 (GenBank)	MK187066.1 (GenBank)
<i>Kigelia africana</i>	Bignoniaceae	Shrub	1	JF265490.1 (GenBank)	MK187120.1 (GenBank)
<i>Blepharis integrifolia</i>	Acanthaceae	Forb	8	MZ461576.1 (GenBank)	DQ054860.1 (GenBank)
<i>Abutilon angulatum</i>	Malvaceae	Forb	9,5	JX572177.1 (GenBank)	MK186953.1 (GenBank)
<i>Rhynchosia fleckii</i>	Fabaceae	Forb	0,1	MZ461592.1 (GenBank)	MZ461568.1 (GenBank)
<i>Indigofera tinctoria</i>	Fabaceae	Forb	3	JF265485.1 (GenBank)	MK187108.1 (GenBank)
<i>Evolvulus alsinoides</i>	Convolvulaceae	Forb	2,5	KR737505.1 (GenBank)	KR737964.1 (GenBank)
<i>Ipomoea cairica</i>	Convolvulaceae	Forb	0,8	KR737546.1 (GenBank)	KR738692.1 (GenBank)
<i>Aptosimum lineare</i>	Scrophulariaceae	Forb	0,3	MZ461572.1 (GenBank)	MZ461572.1 (GenBank)
<i>Panicum maximum</i>	Poaceae	Grass	12	KR737299 (GenBank)	KR738467.1 (GenBank)

<i>Urochloa mosambicensis</i>	Poaceae	Grass	6,3	KNPA247-09 (BOLD)	PNG431-18 (BOLD)
<i>Eragrostis superba</i>	Poaceae	Grass	1,5	KR736685.1 (GenBank)	KR737798.1 (GenBank)
<i>Pogonarthria squarrosa</i>	Poaceae	Grass	0,6	KNPA215-09 (BOLD)	DQ655885.1 (GenBank)
<i>Cyperus rotundus</i>	Cyperaceae	Sedge	0,4	JX644667.1 (GenBank)	JX644807.1 (GenBank)
<i>Bulbostylis hispidula</i>	Cyperaceae	Sedge	0,5	AM999781.1 (GenBank)	MG844244.1 (GenBank)

There was a total of 4 synthetic faecal samples with 5 replicates each representing the mock community (Fig. 4.2). The first sample represented species in equal proportions, the second showed species in their real, expected abundances (Table 4.1), and the third represented the most abundant species, in this case, the shrub *Dichrostachys cinerea*, in the largest proportion (50%) and the rest in equal, smaller abundances (2,27%). The fourth sample represented an opposite situation than that of sample three, where the largest abundance is given to the species expected to occur the least in real samples, in this case, the forb *Rhynchosia fleckii*

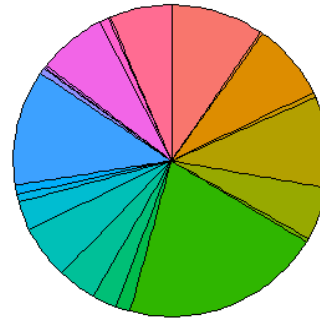
Sample 1

Mock community members in equal abundances



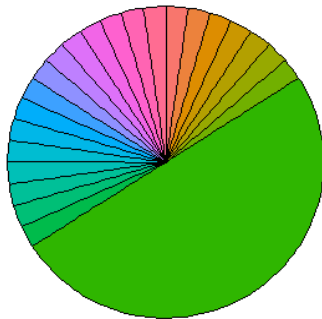
Sample 2

Mock community members in 'real' abundances



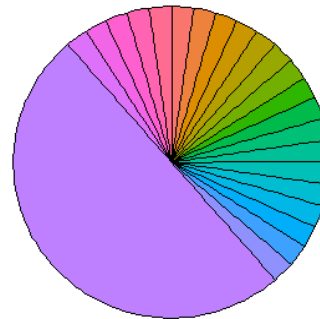
Sample 3

Dichrostachys cinerea representing 50%



Sample 4

Rhynchosia fleckii representing 50%



Species

Abutilon angulatum	Indigofera tinctoria
Aptosimum lineare	Ipomoea cairica
Blepharis integrifolia	Kigelia africana
Bulbostylis hispidula	Megathyrsus maximus
Combretum apiculatum	Pogonarthria squarrosa
Combretum zeyheri	Rhynchosia fleckii
Cyperus rotundus	Schotia brachypetala
Dichrostachys cinerea	Sclerocarya birrea
Eragrostis superba	Senegalia nigrescens
Evolvulus alsinoides	Tricalysia junodii
Flueggea virosa	Urochloa mosambicensis
Grewia bicolor	

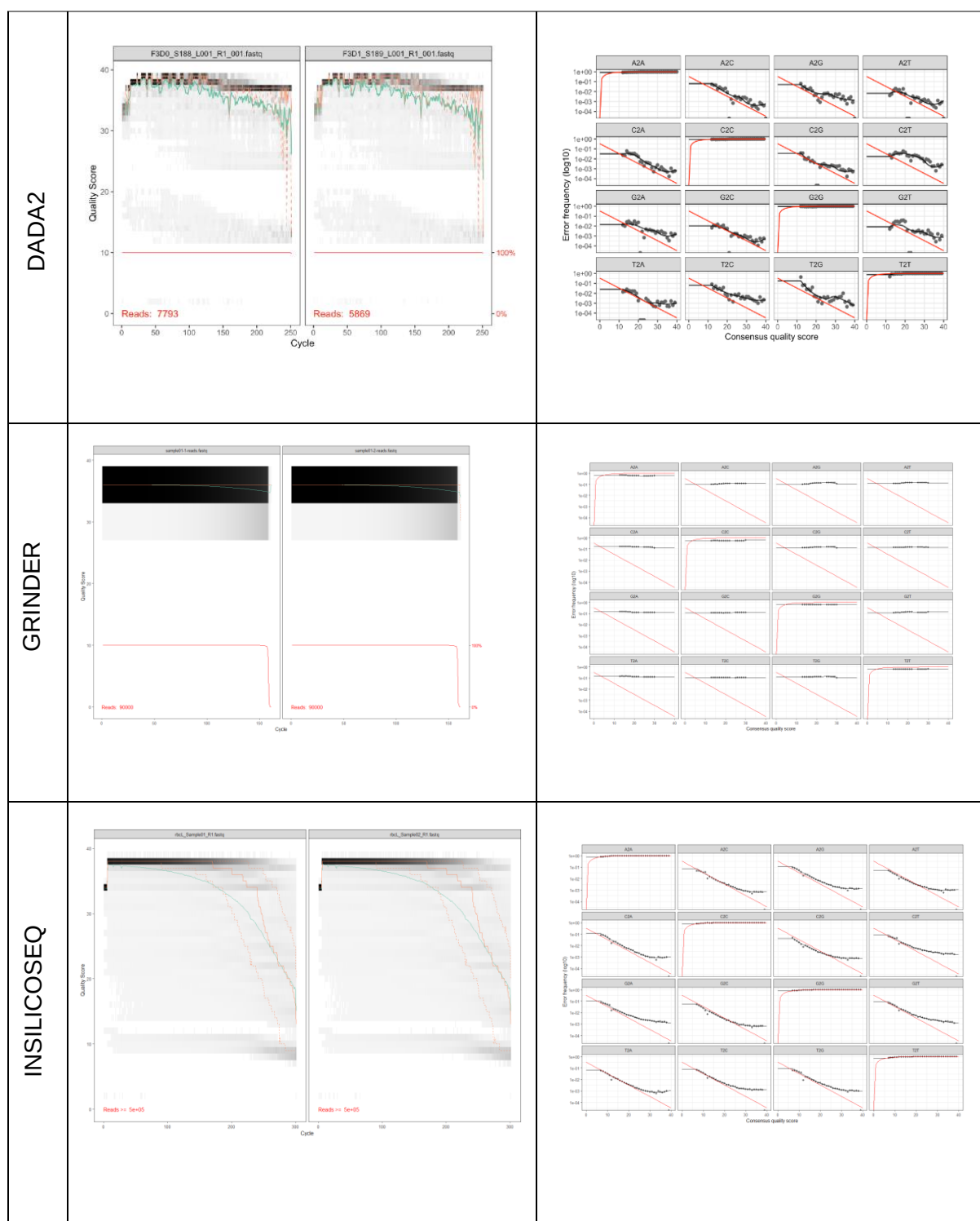
Figure 4.2: Abundance compositions of the four mock community samples generated from 23 species.

Simulation of plant community assemblages contained in faeces

An open-source NGS data simulator, INSILICOSEQ (v. 1.5.4), was chosen to simulate the faecal assemblages, where each sample was replicated 5 times to mimic PCR replicate variability expected when working with real metabarcoding datasets. The pre-built error model of Illumina MiSeq was used to generate 90 000 reads per sample (based on the parameters set by Mathon *et al.* (2021)), where they used 300 bp bidirectional DNA sequencing reads. INSILICOSEQ automatically created outputs for both the forward and reverse FASTQ files as expected after MiSeq sequencing, there was therefore no need to concatenate amplicon reads and quality scores as one would have done with other NGS simulators, such as GRINDER (Angly *et al.*, 2012). INSILICOSEQ also has a unique ability to generate quality scores and error rates that mimics real-life data in such a manner that it is compatible with the quality models implemented in DADA2 (Table 4.2). This NGS simulator would therefore allow the most comparable analysis of bioinformatics pipelines and filtering parameters that can be applied to NGS data retrieved from faecal samples (Table 4.2). Just as sequenced data received from NGS would be subjected to quality control, a quality control step for the simulated dataset was also included and conducted via MULTIQC (v. 1.13: Ewels *et al.*, 2016) which produces summaries of FASTQC (v. 0.11.9: Andrews *et al.*, 2010) reports per sample and sequencing direction regarding the quality and length of sequence reads contained in the NGS dataset.

Table 4.2: A comparison between the next-generation sequencing (NGS) simulators GRINDER and INSILICOSEQ, illustrating the ability of INSILICOSEQ to generate realistic error rates and sequence quality when compared to real data such as the data analysed in the DADA2 tutorial (<https://benjjneb.github.io/dada2/tutorial>). The quality profiles show a heat map of the frequency of each quality score at each base position. The error rates plots illustrate the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.

	Quality Profile	Error Rates



4.1.2 Selected bioinformatics pipelines and filtering approaches

OBITOOLS (v. 1.2.13) (e.g., Kartzinel *et al.*, 2015; Chua *et al.*, 2021), USEARCH (e.g., Edwards *et al.*, 2019; Meyer *et al.*, 2022) and QIIME (e.g., Sickel *et al.*, 2015) were identified as the most common bioinformatics pipelines used to analyse metabarcoding data generated for dietary analysis (Ando *et al.*,

2020). Since USEARCH is not an open-source program and the free version is memory-limited (32-bit), its alternative, VSEARCH (v. 2.7.0), was employed. It offers a 64-bit design, is at least as accurate or more accurate than USEARCH, offers much of the same functions and similar command line inputs and its source code is publicly available (preferred use by, e.g., Matezans *et al.*, 2019).

QIIME1 is no longer maintained, and many of the PYTHON scripts used to analyse NGS data cannot be replicated in new studies. Gardner *et al.* (2019) further stated that QIIME1 is less accurate than QIIME2 in terms of recreating community assemblages as required for this study. Instead of the use of QIIME2 (v. 2022.2) as-is, its plugin DADA2 (v. 1.24.0) (Callahan *et al.*, 2017) was used to assess diet composition due to the success of this method evident across various recent studies (Boggs *et al.*, 2019; Timpano *et al.*, 2020; Johnson *et al.*, 2021; Tsukamoto *et al.*, 2021; Anthoons *et al.*, 2021; Barnes *et al.*, 2022). The bioinformatics pipelines and filtering parameters were applied successively to the raw, simulated data and as standalone methods without any other filtering to assess the true impact of that the pipeline's default parameters as well as the effect that the filtering techniques will have on the observed floristic composition and abundance when compared to the expected composition and abundance. The applied bioinformatics approaches will differ in: (a) the method to analyse sequence variation, which includes pre-clustering quality filters as well as the method to construct species proxies, and (b) filtering of the mock community via contingency-based and abundance filters as well as the implementation of a post-clustering (or post-denoising, in the case of DADA2) curation algorithms. Since NGS simulators, such as INSILICOSEQ and GRINDER cannot simulate chimeras, the chimera removal step of each bioinformatics pipeline could not be evaluated for this mock community.

Some challenges in OBITOOLS installation arose. The original version of OBITOOLS had to be installed in a QIIME1 environment due to the need for the *obiconvert* tool that converts reference sequences' databases to a format compatible with the commands of OBITOOLS. There does not seem to be an equivalent in the new QIIME2-compatible version, OBITOOLS3. The installation instructions on a UBUNTU virtual machine (version 20.04.4 LTS) with its compatible QIIME2 image (2022.2 - 1654113933) installed are, therefore:

```
conda create -n qiime1 python=2.7
```

```
conda activate qiime1
```

```
conda config --add channels conda-forge
```

```
conda install -c bioconda/label/cf201901 obitools
```

```
conda update obitools
```

obiconvert -h # test for successful installation

Another important consideration for OBITOOLS was that the sequence headers in the reference database FASTA files had to be in a specific format to be compatible with the OBITOOLS software. This is a key consideration when using a locally built, custom reference database that contains sequences from various public databases and not just EMBL as is the standard for OBITOOLS tutorials (e.g., <https://pythonhosted.org/OBITools/wolves.html>). Since the reference database used in this study did not contain any barcode primers (not *rbcLaF/R* or *trnLc/d*) or adapters left by the sequencing platform and originated from both GenBank and BOLD, some alternative measures were needed. OBITOOLS' sequence headers must have taxonomic identifiers (taxids) as they are described on NCBI taxonomic repositories. These "taxids" are unique for each taxon across taxonomic ranks. The species' originating from BOLD (Table 4.1) were, therefore, looked up on NCBI's taxonomy browser and the unique "taxid" was manually added to the sequence header. Species for which a GenBank accession number was available (Table 4.1) were downloaded as text files from EMBL and converted to a single FASTA file, per marker, by the *obiconvert* command with input formats being GenBank (*--genbank*) and output formats being FASTA (*--fasta-output*).

The reference database FASTA files that now existed for each marker for GenBank and BOLD were converted to an ECOPCR database by the *obiconvert* command via NCBI's taxonomic repository containing all the taxonomic hierarchies for each "taxid". This repository is represented as "taxdump" files, which need to be downloaded and converted to the appropriate database format by *obitaxonomy*. Once the correct sequence header is retrieved by matching "taxids" described in reference database headers to records in the taxonomic repository, the ECOPCR database could be converted to FASTA again which now represents the reference sequence database used for taxonomic classification in the OBITOOLS workflow. One last step in the reference database construction, specific to OBITOOLS, was to annotate each sequence header with taxonomic ranks from the downloaded database by the *obiannotate* command. No such considerations were necessary for DADA2 and VSEARCH as it only required the reference database to be in FASTA format, however, taxonomic ranks were added to sequence identifiers as a separate metadata file, which was a requirement for import into R and analysis of the performance indices as a PHYLOSEQ object.

4.1.3 Sequence variation analysis

Pre-clustering quality filters

Unlike OBITOOLS and the R package DADA2 which has quality filters built into their workflows with the *obigrep*, *obiclean* and *filterAndTrim* function, respectively, the VSEARCH plugin of QIIME2 does not filter sequences according to quality, length, or ambiguous bases in its clustering parameters or main

workflow steps that precede it. To determine whether quality filtering according to base quality scores or ambiguous bases would influence the performance of the pipeline, a quality filter was applied before reads are clustered into species proxies in the method VSEARCH_Q for both the *rbcL* and *trnL* mock community datasets. The filter was applied to the forward reads of *rbcL* and *trnL* samples that were imported as demultiplexed QIIME2 artifacts via *qiime tools import* with the input format *CasavaOneEightSingleLanePerSampleDirFmt*. The filter was applied by *qiime quality-filter q-score* parameters being 20 for *p-min-quality* and zero for *p-max-ambiguous* (method = VSEARCH_Q). The VSEARCH pipeline was also implemented without quality filtering (method = VSEARCH). For DADA2 (method = DADA2) default parameters were used to filter and trim the reads (*filterAndTrim*) which did not allow for ambiguous bases as required by the DADA2 denoising algorithm, truncated reads according to first instances of low Phred quality scores and allowed for the retainment of reads with a maximum amount of two sequencing errors. No “pre-clustering” quality filters were applied in OBITOOLS since it encompasses removing unaligned sequence records, which will require the forward and reverse reads to be merged which will not be done in this study.

OTU and ASV table construction

DADA2 was used to represent the ASV clustering approach, VSEARCH to represent the OTU clustering approach based on a similarity threshold, and OBITOOLS was used to represent an approach that groups strictly identical sequences. Dereplicated mock community reads passing the quality filter (method = VSEARCH_Q) or those that have not undergone quality filtering (method = VSEARCH) for the *rbcL* and *trnL* dataset that displayed more than 97% similarity (Ducotterd *et al.*, 2020) were clustered into OTUs using the *de novo* clustering approach of a feature table by *qiime vsearch cluster-features-de-novo* (method = VSEARCH_DN) as well as closed-reference clustering of a feature table by *qiime vsearch cluster-features-closed-reference* (VSEARCH_CR). The feature table refers to the output of the *qiime vsearch dereplicate-sequences* command that is a “FeatureTable[Frequency]” QIIME2 artifact which contains frequency information about each sequence.

The reference sequences which are used as a template in closed-reference OTU clustering were the same sequences that were used by INSILICOSEQ to generate mock community assemblages. The closed-reference clustering approach of VSEARCH was employed to act as a positive control for various steps in the workflow and the applied filters. OTUs were created by matching sequences against reference sequences based on a similarity threshold. This clustering method must identify 23 features and sensitivity and precision must equal one. Root mean square error (RMSE) values also must be very low, if not zero. Closed-reference OTU picking will not be considered as the chosen method (method = VSEARCH_CR) to analyse real faecal DNA since all species occurring in the ecosystem had to be well-defined and present in the reference sequence database, which, as stated, is not optimal for environmental DNA metabarcoding

studies (Westcott & Schloss, 2015; Callahan *et al.*, 2017) since much biodiversity is still to be discovered, there is a lack of plant sequences on reference database due to the need for standardisation of barcoding markers and savanna species, especially forbs are poorly presented in DNA reference databases. The number of OTUs resulting from closed-reference picking would therefore not capture the real biodiversity of the animal's diet and its ecosystem.

Sequences left after filtering for the DADA workflow were used in the *learnErrors* command whereby DADA2 uses an algorithm to repeatedly perform sample inference on each sample until an optimal convergence is reached. A clustering threshold was therefore not applicable for DADA2. This error rate and the filtered sequences were then used for sample inference where ASVs were created and could be converted to a feature table by the *makeSequenceTable* command.

Identical sequences left after OBITOOLS filtering were grouped and dereplicated with the *obiuniq* command with its default parameters. The file of representative sequences for OBITOOLS did not need to be converted into a feature table similar to VSEARCH and DADA2 since it would not be analysed by QIIME2's feature table filters since it has its own built-in commands to filter OBITOOLS formatted data. Identical sequences that have been grouped by the *obiuniq* command will hereafter be referred to as "grouped sequences" in contrast to OTUs/ASVs for VSEARCH and DADA2, respectively, and "features" will be used to describe the grouping/clustering output of all three pipelines.

4.1.4 Final filtering of the mock community

Five filtering parameters were applied to all the final representative sequences of each pipeline which in combination is hereafter referred to as "methods" (adapted from Pauvert *et al.*, 2019). Feature-based filtering was represented by Filtering=1 and Filtering=10 which aided in removing features that were only present in one or a few samples, based on suspicions that features with such low frequency represented PCR or sequencing errors and not real biological diversity. Filtering=RA represented abundance-based filtering where 0.005% of the total number of sequences were removed (Bokulich *et al.*, 2013). Filtering=LULU involved clustering curation of amplicon data by the R package, LULU (Frøslev *et al.*, 2017). The purpose of this filter was to reduce the number of erroneous features in feature tables by collapsing problematic features into their parent feature to achieve more realistic biodiversity estimates. Filtering=ALL involved leaving feature data as is after sequence variation analysis to analyse whether further filtering of the feature tables improved the recoverability of the mock community species in the correct proportions.

VSEARCH and DADA2 feature tables for Filtering=1 and Filtering=10 was performed with *qiime feature-table filter-features* with *p-min-samples* 2 for Filtering=1 (method = VSEARCH_1; DADA_1) and *p-min-frequency* 10 for Filtering=10 (method = VSEARCH_10; DADA_10). Abundance filtering for VSEARCH

and DADA2 was performed by first retrieving feature table statistics by the *qiime metadata tabulate* command, unzipping the resultant QZV file, and extracting the metadata TSV file as input for the *qiime feature-table summarize* command, where the output QZV file could be visualised as an HTML file by *qiime tools view* and noting the “total frequency”. This total frequency was then multiplied by 0.00005 and the resulting value was the parameter input, as a whole number, into the *p-min-frequency* option of the *qiime feature-table filter features* command (method = VSEARCH_RA; DADA_RA). Filtering for OBITOOLS was not conducted in QIIME2 since specific commands are built into the workflow with recommended parameters which would not require converting between file formats and possibly compromise the integrity of the OBITOOLS pipeline.

OBITOOLS with the *obigrep* command filters sequences by count and length. For Filtering=1 the specified parameters in *obigrep* was *-p 'count>1'* (method = OBI_1) and for Filtering=10: *-p 'count>10'* (method = OBI_10). The abundance filter, *obiclean* was used to represent Filtering=RA that removed PCR and sequencing errors by only keeping the head of sequences that had no variants with a count of greater than 5% of their own count with the parameter *-r 0.05 -H* (method = OBI_RA). A combination of frequency and abundance filtering was also applied to each pipeline and was defined as METHOD_1_RA and METHOD_10_RA. The post-clustering curation algorithm had to be applied after taxonomic assignment when abundance- and frequency-based filtered feature data could be exported as BIOM files and imported as a PHYLOSEQ object for R analysis.

The taxonomic classification for DADA2 and VSEARCH was performed in QIIME2 by importing the reference database TSV files for each marker containing accession numbers or process identities as well as seven taxonomic ranks of each concerned mock community member. This was done by *qiime tools import* and specifying the input format as *HeaderlessTSVTaxonomyFormat*. Next, the reference database FASTA files were imported as a QIIME2 artifact with the “type” being specified as *FeatureData[Sequence]*. Finally, taxonomy was assigned by *qiime feature-classifier classify-consensus-blast* with a maximum number of 1 hit to keep for each query and a percentage identity match of 0.99 (99% since the mock community reference database only contains 23 species of which the sequences are also query species). Taxonomic assignment of OBITOOLS features was carried out by *ecotag* using the formatted reference database of mock community sequences for the *rbcL* and *trnL* marker as well as the filtered reads. Per the recommendation of Ariza *et al.* (2022), only query sequences with a 100% match to the reference sequences were retained since it was observed that any lower threshold sequences will remain without any taxonomic rank assigned. For the final taxonomic table of OBITOOLS, some tags in sequence headers that will afford no value to interpretation in this study such as “scientific_name_by_db”, “obiclean_samplecount”, “obiclean_count”, “obiclean_singletoncount”, “obiclean_cluster”, “obiclean_internalcount”, “obihead_clean”, “taxid_by_db”, “obiclean_headcount”, “id_status”, “rank_by_db”, “order_name” and “order” were removed with the *delete-tag* option in *obiannotate*. The final assignments of query sequences

to their taxonomy were converted from FASTA files to a tab-delimited file that could be opened by EXCEL or R.

4.1.5 Performance comparison criteria

The ability of the various methods to recover the mock community species and proportions was evaluated based on three comparison criteria: sensitivity (Se), precision (P) and the overall quantitative performance of each bioinformatics approach with its filtering parameters were measured using root mean square error (RMSE) between observed and expected abundances (abundance input files for INSILICOSEQ for each sample).

The sensitivity and precision of each method were respectively defined as the true positive rate $TP/(TP + FN)$ as well as the positive predictive value $TP/(TP + FP)$. TP represents the number of true positive OTUs/ASVs which correspond to mock community species that were present in the mock community and could be identified by the method. FN represented false positive OTUs/ASVs and were species that are present in the mock community but could not be identified by the considered method. Finally, FP was the number of false positive OTUs/ASVs which represented all other OTUs identified by the method as species proxies but were not present in the mock community. These indices of sensitivity and precision represent complementary advantages for the comparison of methods: Sensitivity is useful to identify those methods that miss rare taxa while precision can identify methods that detect false positives.

The quantitative performance of each method was measured by calculating the RMSE for each abundance comparison by implementing the *rmse()* function of the Metrics package (Hamner *et al.*, 2018) in R. The values output by the presence/absence comparison criteria should range between 1 and 0. A value of 1 for sensitivity and precision indicates that the method was effective in identifying all species from the mock community. The lower the RMSE value, the more similar the observed and expected relative abundances were. All three comparison criteria were calculated per sample after summing the species counts per replicate and all statistical analyses were carried out with R v4.2.0.

Results

4.1.6 Simulated plant community

For the 20 replicates simulated with INSILICOSEQ with the MiSeq error model, a total of 1 800 000 paired-end sequencing reads of 301 bp were obtained, where each sample's forward and reverse files contained 45 000 reads respectively. According to MULTIQC quality control files for the forward reads of both the *rbcL* and *trnL* mock community datasets, Phred Scores of less than 20 were observed for all 20 replicates for both markers from 292 bp, meaning they contained very few ambiguous bases which were only discovered from 292 bp and onward, contained overrepresented and duplicated sequences, but no adapter

contamination which is expected since it did not pass through a sequencing platform. Although the datasets displayed unique or very similar reports in terms of quality, overrepresented sequences, duplicate sequences, ambiguous bases, the guanine-cytosine content (GC content) differed which is expected from different barcoding markers.

4.1.7 Closed-reference OTU picking

The closed-reference OTU picking method illustrates that simulated data could be matched to their references with very high confidence (Table 4.3). Significant differences in the case of this approach were observed between the ability of the different methods to recover the relative abundance of mock community members when applying VSEARCH with and without a pre-cluster quality filter. When VSEARCH was applied without a quality filter, the RMSE of the method slightly decreased. Except for the implementation of the pre-clustering filter, there was also no difference in the comparison criteria of the *rbcL* and *trnL* barcoding marker between different methods.

Table 4.3: List of approaches used after closed-reference clustering of the VSEARCH bioinformatics pipeline. These values represent sample means retrieved for each method, for the *rbcL* and *trnL* plant barcoding marker.

Method	Richness		Sensitivity		Precision		RMSE	
	<i>rbcL</i>	<i>trnL</i>	<i>rbcL</i>	<i>trnL</i>	<i>rbcL</i>	<i>trnL</i>	<i>rbcL</i>	<i>trnL</i>
VSEARCH_CR_ALL	23	23	1	1	1	1	0,425959	0,017476
VSEARCH_CR_1	23	23	1	1	1	1	0,425959	0,017476
VSEARCH_CR_10	23	23	1	1	1	1	0,425959	0,017476
VSEARCH_CR_RA	23	23	1	1	1	1	0,425959	0,017476
VSEARCH_CR_1_RA	23	23	1	1	1	1	0,425959	0,017476
VSEARCH_CR_10_RA	23	23	1	1	1	1	0,425959	0,017476
VSEARCH_CR_Q	23	23	1	1	1	1	0,514093	0,029456
VSEARCH_CR_Q_1	23	23	1	1	1	1	0,514093	0,029456

VSEARCH_CR_Q_10	23	23	1	1	1	1	0,514093	0,029456
VSEARCH_CR_Q_RA	23	23	1	1	1	1	0,514093	0,029456
VSEARCH_CR_Q_1_ RA	23	23	1	1	1	1	0,514093	0,029456
VSEARCH_CR_Q_10_ RA	23	23	1	1	1	1	0,514093	0,029456

4.1.8 Influence of post-clustering filters

OBITOOLS methods

OBITOOLS did not perform as expected. For OBI_ALL, where no post-clustering filters were concerned, the command *ecotag* did not result in any taxonomic assignments for both plant barcoding markers and was therefore excluded from further analysis, along with filters that rely on the outputs of this method, including OBI_RA. For the other methods employed to analyse the performance of this bioinformatics pipeline (OBI_1, OBI_1_RA, OBI_10, OBI_10_RA), the *rbcL* mock community dataset was only classified to clade level (Magnoliopsida), meaning that no taxonomic differentiation took place at all. A test to confirm the compatibility of the *rbcL* sequences with the commands implemented in OBITOOLS was conducted where queries are directly used as a reference and the pipeline was then able to differentiate between sequences up to species level. For *trnL*, these filters were able to differentiate between the mock community members, by only filtering according to frequency. The abundance filter, OBI_RA via *obiclean*, did not give any outputs. However, the frequency filters, OBI_1 and OBI_10, were still able to filter the *trnL* mock community dataset (Figure 4.3) and the grouping of features based on similar sequences yielded 18447 features by the OBI_1 method and 805 for the OBI_10 method. OBI_1 proved to be a more sensitive method than OBI_10 based on replicate means, and OBI_10 proved to be the more precise method based on replicate means (Figure 4.3).

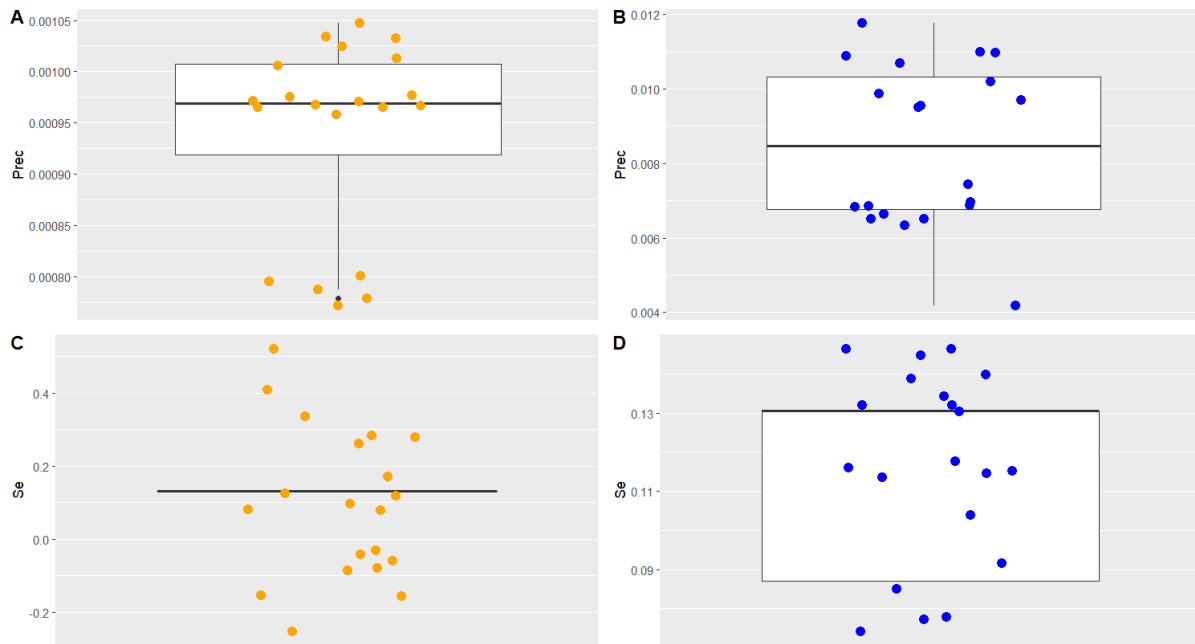


Figure 4.3: A comparison of the performance indices, precision (orange scatterplots), and sensitivity (blue scatterplots) for the OBITOOLS methods OBI_1 (A and B) and OBI_10 (C and D) as applied to the trnL mock community dataset. The scatterplots represent the mean index received for each of the 4 samples with its 5 replicates and the boxplots represent the median of the performance index as well as the first and third quartile of the 20 samples.

The successful OBITOOLS methods were also only able to identify three genera from two families for the entire mock community dataset. The families identified were Fabaceae with members being *Rhynchosia* and *Schotia* as well as Scrophulariaceae with *Aposimum* as a member. Relative read abundance, as indicated by the stacked bar plots (Figures 4.4 and 4.5) showed that the methods retrieved relatively similar proportions for mock community members, with the only difference being a much greater inclusion of unassigned reads for OBI_1. The filtering method OBI_1 could identify Fabaceae in a much greater proportion than just Scrophulariaceae. The relative read abundances of genera identified by the OBI_1 method revealed that *Rhynchosia* was identified in a greater proportion than the other member and makes up the largest proportions of the most abundant family. The OBI_10 method was able to identify these same mock community members with the same relative proportions as described for OBI_1.

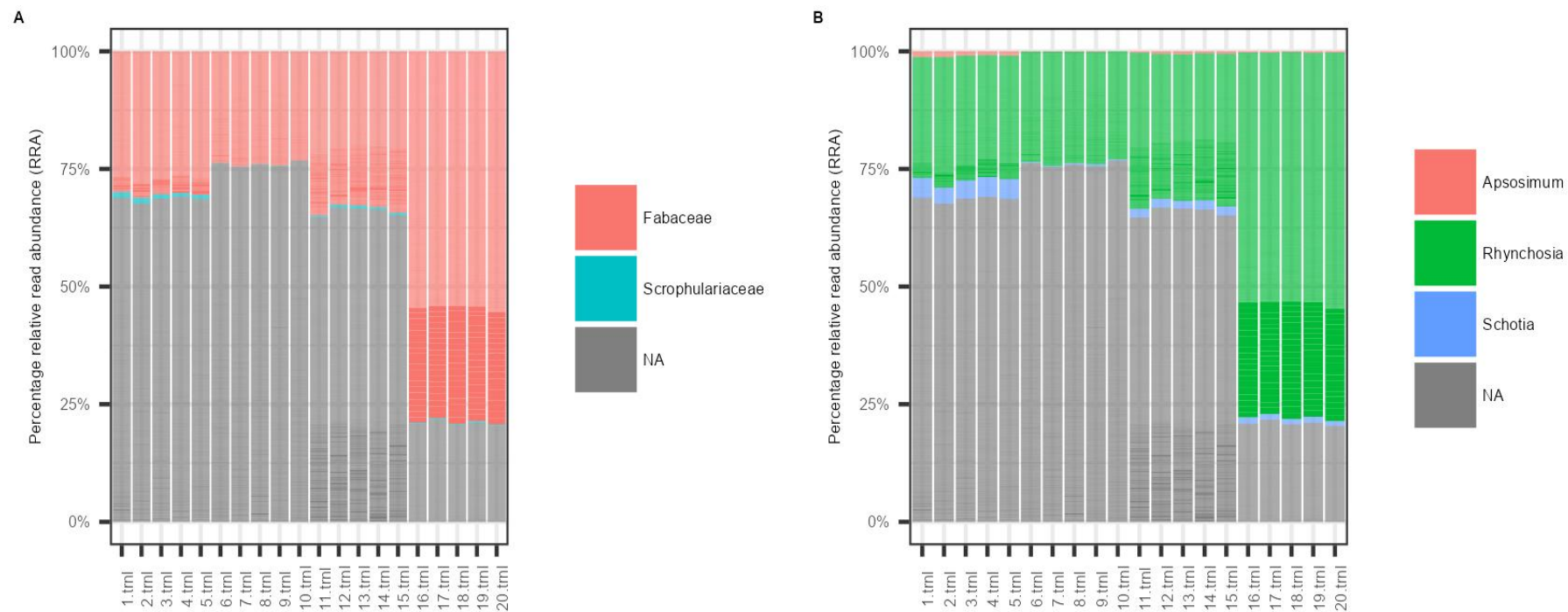


Figure 4.4: Stacked bar plots indicating the relative read abundance (%) by plant per A) plant family and B) genera identified by the OBI_1 method. This filtering method encompasses the removal of features that occur once per sample.

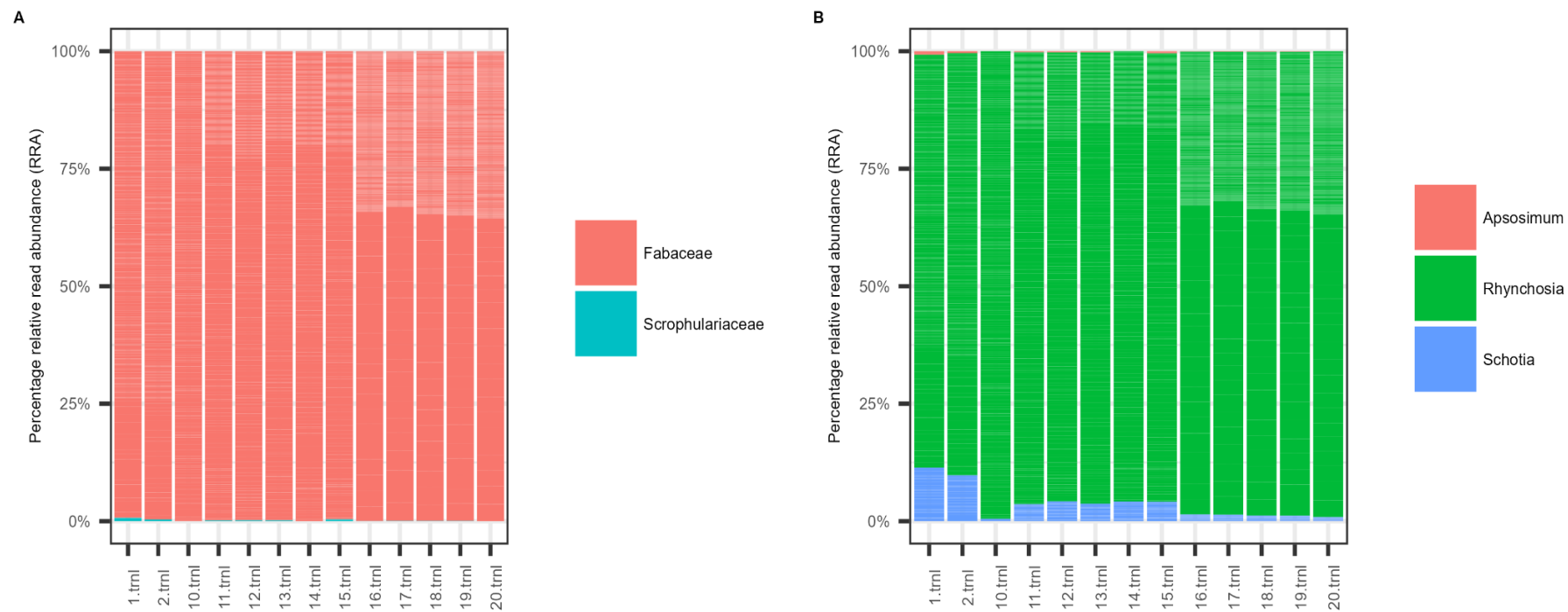


Figure 4.5: Stacked bar plots indicating the relative read abundance (%) by plant per A) plant family and B) genera identified by the OBI_10 method. This filtering method includes the removal of features that occur 10 times or less per sample.

DADA2 and VSEARCH methods

a) Richness

To evaluate and compare the performance of each pipeline, raw values were averaged across samples and replicates for each method tested (Figures 4.6, 4.7, 4.8, and 4.9). For the DADA2 pipeline, applied to the *rbcL* mock community, feature richness retrieved by each method did not differ at all between DADA_ALL, DADA_10, DADA_RA, and DADA_10_RA and remained at 672. The application of these methods to the *trnL* mock community yielded slightly different results with DADA_ALL identifying 1057 OTUs, DADA_10 and DADA_10_RA identifying 1046, and DADA_RA identifying 1 055 OTUs. A significant difference was seen in ASV richness for the method without post-clustering curation when applying a frequency-based filter and a subsequent abundance filter where the richness decreased to 69 for *rbcL* and 135 for the *trnL* dataset (DADA_1 and DADA_1_RA). For the *rbcL* dataset, these trends in feature richness extended to post-clustering curation with LULU with the same richness as seen for DADA_ALL_LULU, DADA_10_LULU, DADA_RA_LULU, and DADA_10_RA_LULU and decrease from 103 to 64 when applying the frequency-based filter with subsequent abundance filtering (DADA_1_LULU and DADA_1_RA_LULU). Therefore, in terms of species richness estimations for the *rbcL* mock community, DADA_1_LULU and DADA_1_RA_LULU outperformed the other methods equally well. The same is true for the *trnL* dataset where the application of the methods with the LULU filter led to a lower feature richness than their non-LULU counterparts and therefore having DADA_1_LULU and DADA_1_RA_LULU retrieving the lowest feature richness of 132. A great overestimation of OTUs took place in the VSEARCH method applied to both mock community datasets. VSEARCH_ALL identified 69 384 sequence proxies for *rbcL* and 59 444 for *trnL*. The application of the post-clustering curation algorithm (VSEARCH_DN_ALL_LULU) decreased the number of OTUs by 72% for the *rbcL* dataset and 64% for the *trnL* dataset. The application of the pre-clustering filter (VSEARCH_DN_ALL_Q) leads to an even larger drop of OTUs retrieved by the VSEARCH_DN_ALL method with 93% for *rbcL* and 95% for *trnL*. The method that performed the best in terms of the retrieval of a feature richness more closely resembling the mock communities for VSEARCH was a combination of pre-and post-clustering filters: VSEARCH_DN_Q_1_RA_LULU, VSEARCH_DN_Q_RA_LULU, VSEARCH_DN_Q_10_RA_LULU identifying 297 OTUs from the *rbcL* database and 317 OTUs from the *trnL* database.

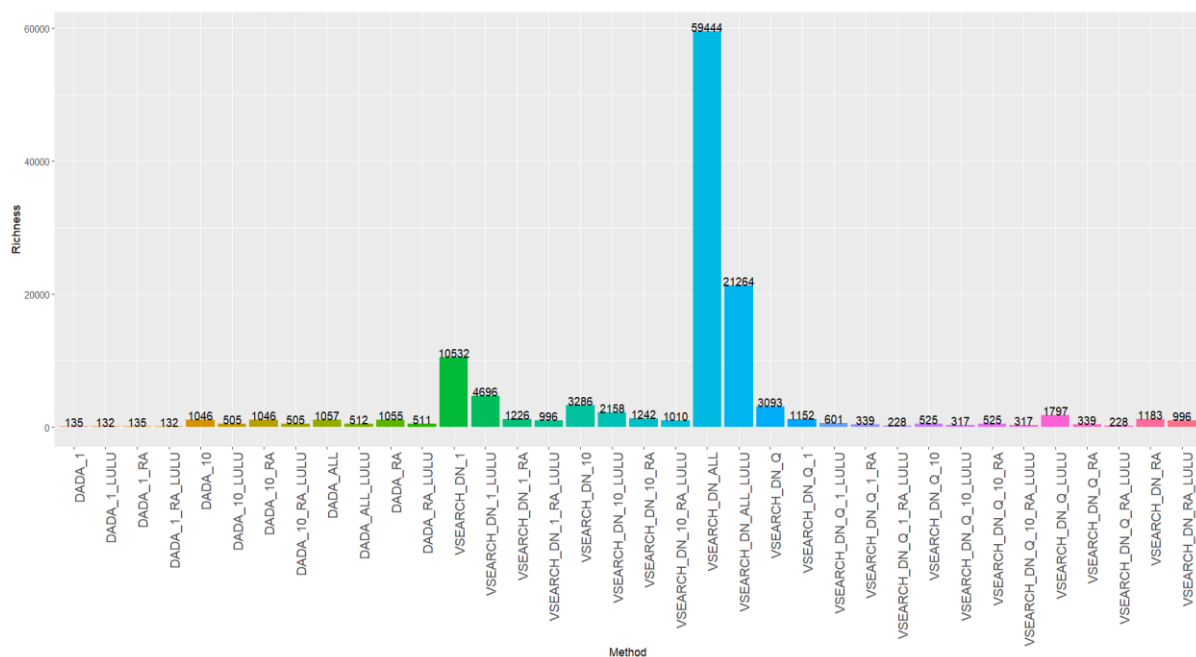


Figure 4.6: The number of sequence proxies identified by the DADA2 denoising methods and the VSEARCH *de novo* clustering methods for the *rbcL* mock community dataset. The highest feature richness (69 384 OTUs) was retrieved by the VSEARCH *de novo* method with no post- or pre- clustering filters applied (VSEARCH_DN_ALL) and the lowest was retrieved by two methods applied for the DADA2 bioinformatics pipeline: DADA_1_RA_LULU identifying 64 ASVs, and DADA_1_LULU also identifying 64 ASVs, where the latter applied both a frequency and abundance filter.

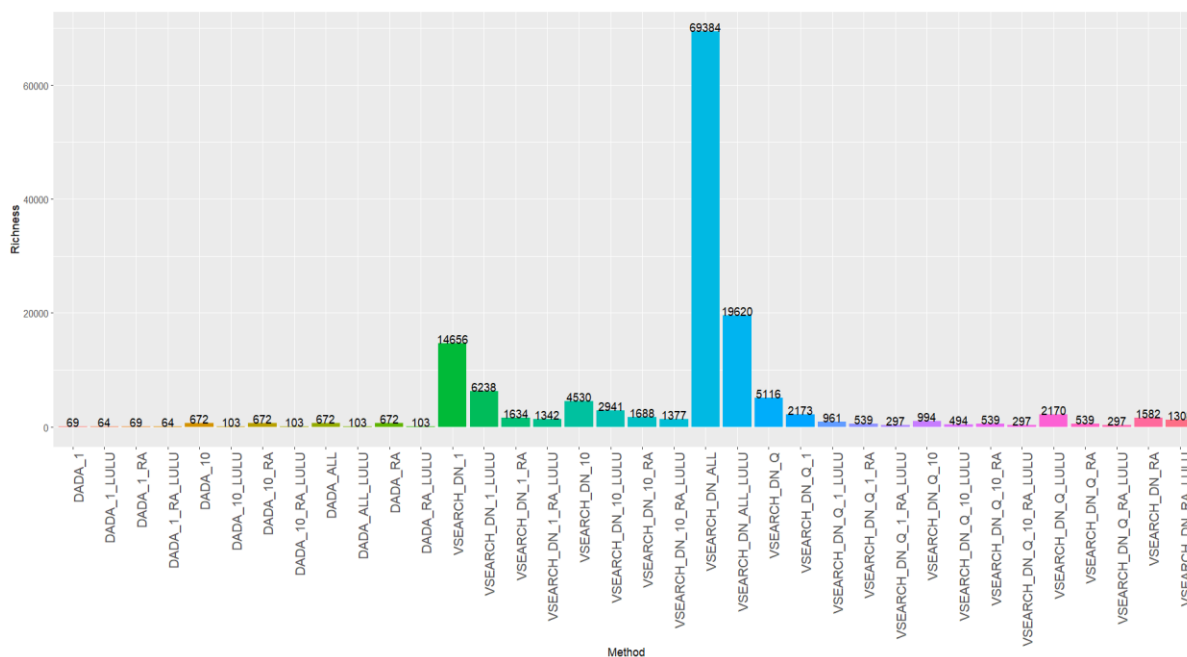


Figure 4.7: The number of sequence proxies identified by the DADA2 denoising methods and the VSEARCH *de novo* clustering methods for the *trnL* mock community dataset. The highest feature

richness (59 444 OTUs) was retrieved by the VSEARCH *de novo* method with no post- or pre- clustering filters applied (VSEARCH_DN_ALL) and the lowest was retrieved by two methods applied for the DADA2 bioinformatics pipeline: DADA_1_RA_LULU which identified 132 ASVs and DADA_1_LULU also identifying 132 ASVs, where the latter applied both a frequency and abundance filter.

b) Taxonomic assignment and comparison criteria

The assignment of these features to species in the mock community reference database with a similarity threshold of 99% significantly differed between DADA2 and VSEARCH (Fig. 4.8). All methods of the DADA2 bioinformatics pipeline were able to match all identified ASVs (100%) to their taxonomies. Methods for VSEARCH that applied the LULU post-curation clustering algorithm performed better than their non-LULU counterparts in terms of percentage assigned reads. The VSEARCH method with the highest percentage of assigned reads for the *rbcL* mock community dataset was VSEARCH_DN_Q_10_RA_LULU with 9.43% and for *trnL* VSEARCH_DN_Q_RA_LULU with 15.79%. The methods that performed the worst in terms of percentage reads identified were VSEARCH_DN_ALL for both *rbcL*, with 0.05% assigned reads, and *trnL* with 0.07%.

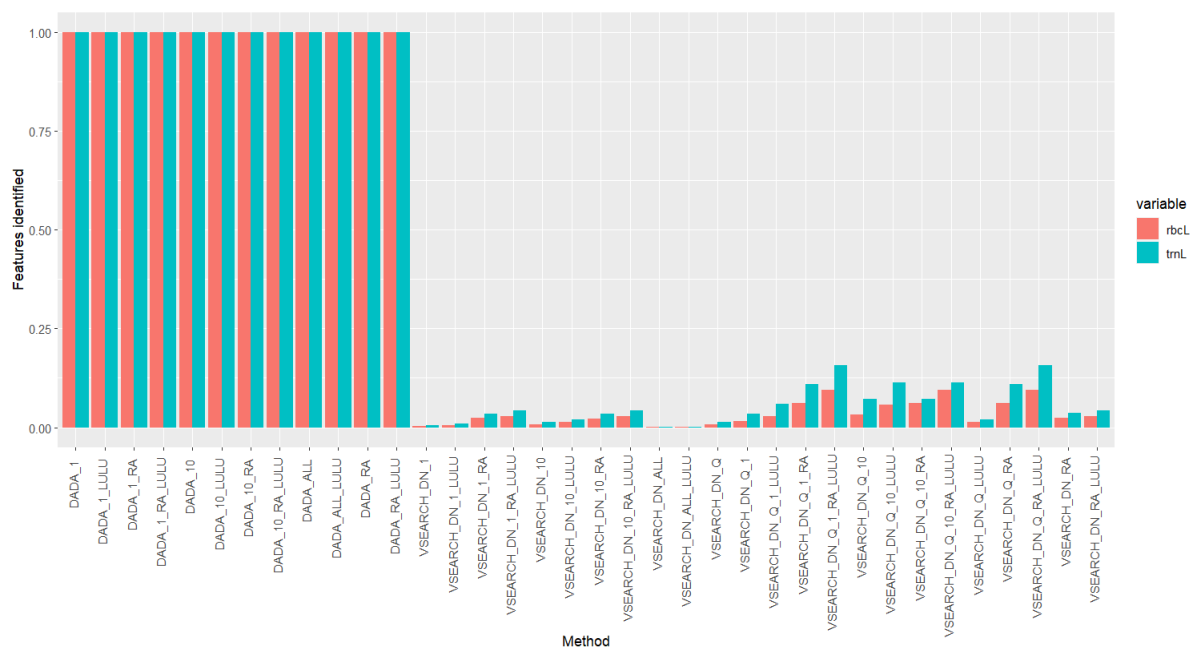


Figure 4.8: Percentage ASVs, in the case of DADA2 methods, and OTUs, in the case of VSEARCH methods, identified for both the *rbcL* (red bars) and *trnL* (blue bars) mock community datasets. These percentages were calculated by dividing the number of assigned reads by the total richness of reads.

The comparison criteria, sensitivity, precision, and RMSE were calculated at the replicate level (Fig. 4.9), ensuring the performance of methods was not biased against certain sample compositions and could recover

the composition of the mock community in any compositional scenario. Bioinformatics methods were therefore ranked in their performance based on the mean value over a total of 4 samples with 5 replicates per mock community assemblage. When sorting the comparison criteria according to precision, RMSE and lastly sensitivity (Fig. 4.9 and Supplementary Tables 5 and 6), DADA_1_RA_LULU and DADA_1_LULU performed equally well for both the *rbcL* and *trnL* datasets with mean precision values of 0.68 for *rbcL* and 0.65 for *trnL*. Although the RMSE values were highest for these two methods (11.214 for *rbcL* and 9.628 for *trnL*), with their non-LULU counterparts coming in a close second, they could retrieve fewer false positives, indicated by high precision values as well as very low feature richness.

The most sensitive bioinformatics pipeline for the *rbcL* and *trnL* mock community datasets was VSEARCH with the methods VSEARCH_DN_RA, VSEARCH_DN_1_RA, VSEARCH_DN_10_RA, VSEARCH_DN_10, VSEARCH_DN_1, VSEARCH_DN_ALL as well as their LULU counterparts proving the most sensitive for the *rbcL* mock community with values of 0.96, identifying all members of the mock community except for the forb *Evolvulus alsinoides* (Fig. 4.10). All VSEARCH methods had a perfect sensitivity value of 1 for the *trnL* mock community dataset, identifying all 23 members of the mock community (Fig. 4.11). The VSEARCH method VSEARCH_DN_ALL_LULU retrieved the lowest RMSE value (0.001937) for the *rbcL* dataset and VSEARCH_DN_Q_10_LULU for the *trnL* dataset with an RMSE value of 0.113565. DADA_1_RA_LULU has been singled out as the method performing equally well in terms of sorting by precision, RMSE, and sensitivity for both the *rbcL* and *trnL* mock community datasets. For the *rbcL* dataset, this method could identify 13 out of 23 members of the mock community: *Blepharis integrifolia*, *Rhynchosia fleckii*, *Dichrostachys cinerea*, *Bulbostylis hispidula*, *Megathyrsus maximus*, *Urochloa mosambicensis*, *Ipomoea cairica*, *Eragrostis superba*, *Sclerocarya birrea*, *Combretum zeyheri*, *Abutilon angulatum*, *Evolvulus alsinoides* and *Combretum apiculatum*, which represents 66% of tree species in the mock community, 25% of shrubs, 71% of forbs, 75% of grasses and 50% of the sedges (Fig. 4.10). The DADA_1_RA_LULU method applied to the *trnL* mock community could identify 21 out of 23 members, except for the trees *Sclerocarya birrea* and *Senegalia nigrescens*, and the sedge *Bulbostylis hispidula*. Therefore, 66.67 % of trees, 100% of shrubs, 100% of forbs, 100% of grasses and 50% of sedges could be identified (Fig. 4.11).

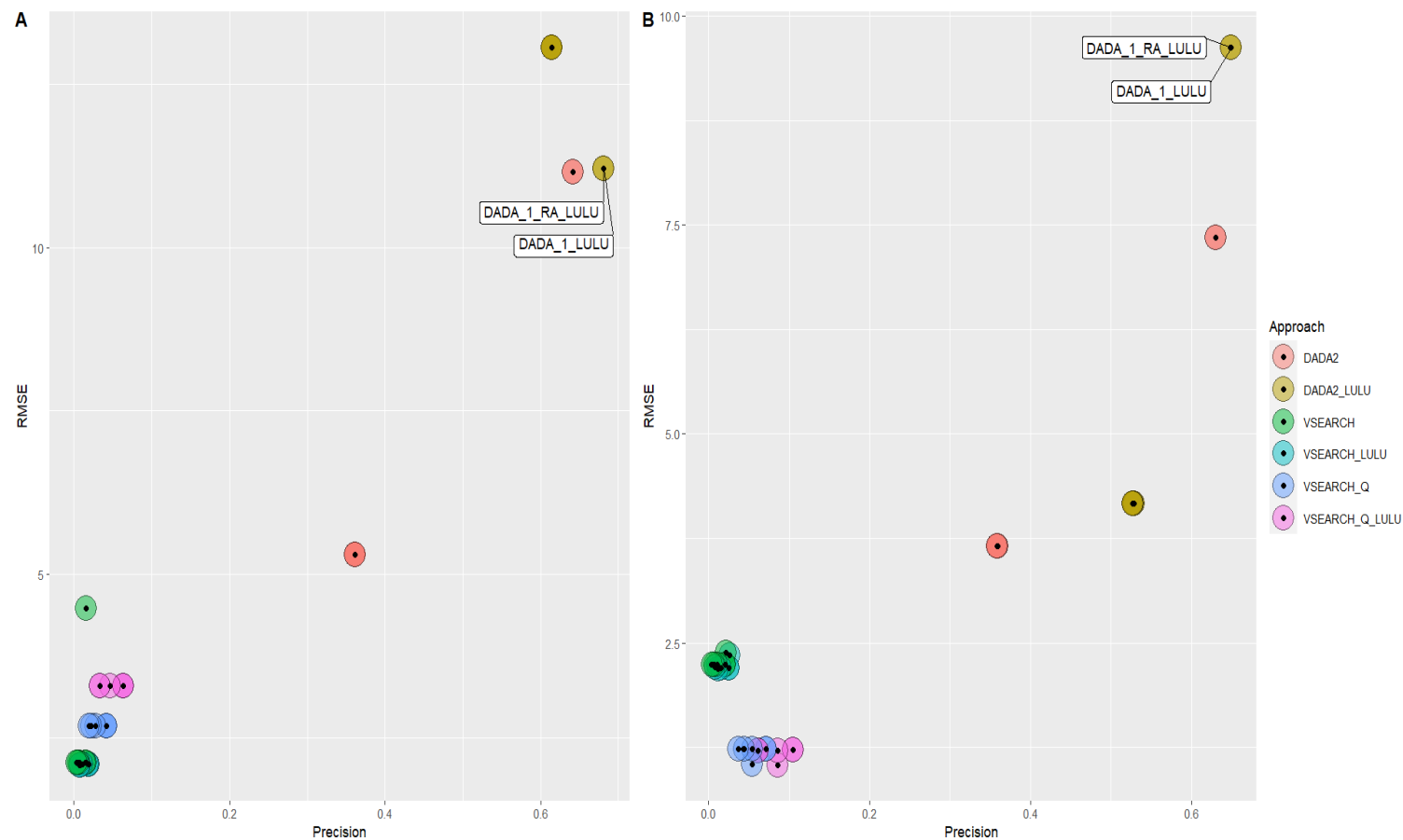


Figure 4.9: Comparison indices of precision and RMSE of each method of the a) *rbcL* and b) *trnL* mock community datasets. The scatter plot points represent mean values averaged across each method for all samples and their replicates with precision, or the positive predictive value, represented by $TP/(TP + FP)$ and RMSE calculated between observed and expected abundances retrieved for each sample replicate, averaged for each method.

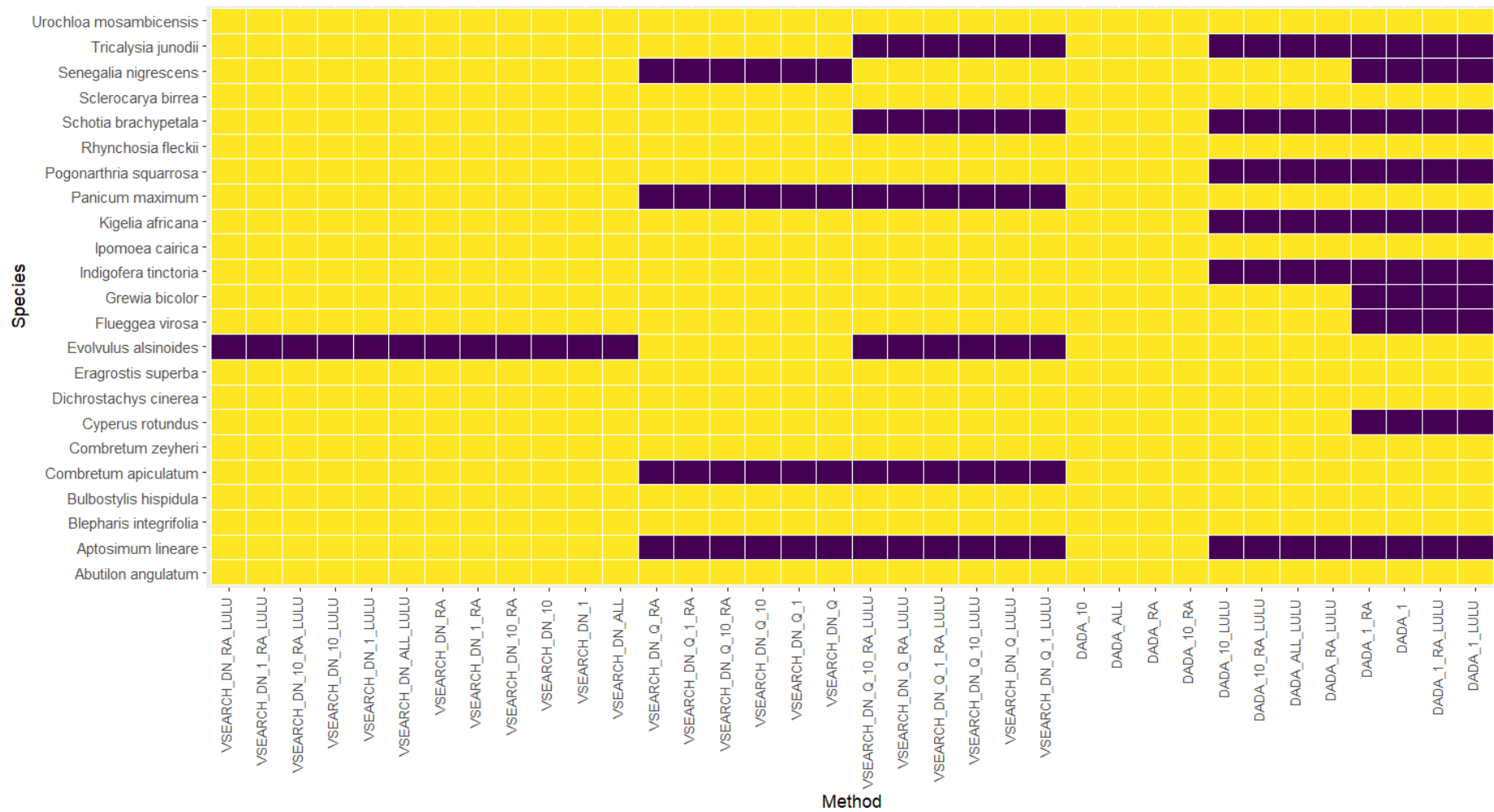


Figure 4.10: Tile plot indicating species that could be identified for each method applied for the *rbcl* mock community. Yellow tiles indicate identified species and purple tiles indicate absent species per method.

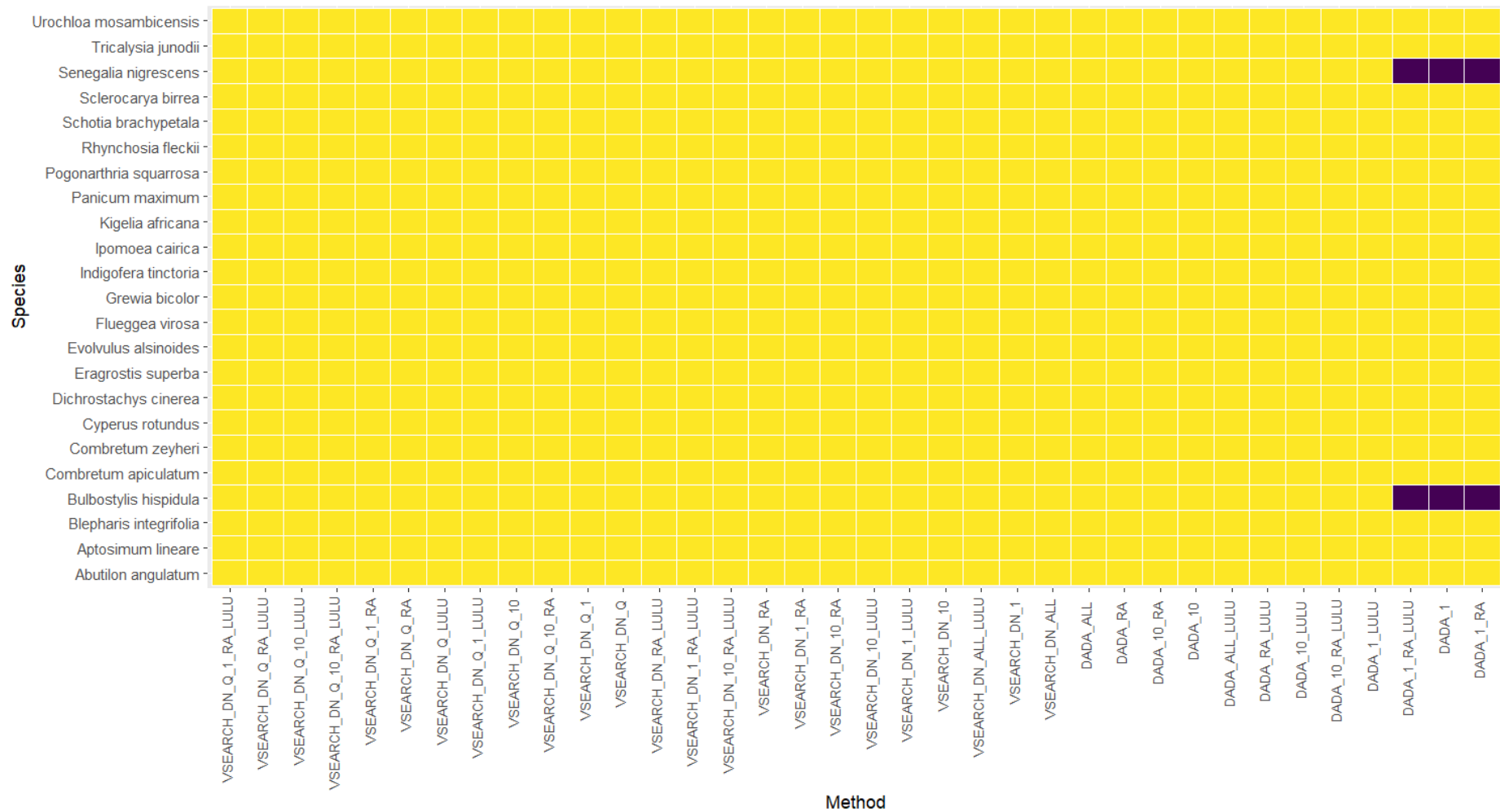


Figure 4.11: Tile plot indicating species that could be identified for each method applied for the *trnL* mock community. Yellow tiles indicate identified species and purple tiles indicate absent species per method.

Discussion

Accurate reconstructions of herbivore diets from terrestrial ecosystems such as savannas enable ecologists to better understand and conserve members contributing to the food chain. The main goal of the overall study was to determine, with high accuracy, which plants contribute to the diets of cattle foraging in an eastern semi-arid South African Savanna. A secondary goal was to determine the degree of contribution of each food item, i.e., the relative abundance of the plant species identified in the faecal samples. These aims were realised specifically by the implementation of a mock community by testing for and discovering which bioinformatics approach and its filtering parameters performed best when evaluated according to (i) the comparison criteria of sensitivity, (ii) precision, (iii) RMSE, (iv) assessments of the feature richness, and (v) the percentage of these features that could be taxonomically assigned. A total of 72 bioinformatics filtering approaches (bioinformatics pipelines with different filtering parameters) were included at the beginning of this study, but after the use of VSEARCH's close-reference clustering as a positive control and the elimination of the OBITOOLS method, 36 methods belonging to DADA2 and VSEARCH's *de novo* clustering method remained to analyse the mock community with. The optimal bioinformatics approach, or method, was chosen based on its ability to recover the richness and abundance of a mock community consisting of 23 species found in different proportions in savanna herbivore diets. The first two comparison criteria, sensitivity, and precision are related to the number of features (OTUs or ASVs) that could be recovered from the NGS dataset and are commonly used in benchmark studies (Chua *et al.*, 2021; Mathon *et al.*, 2021). The third criterion, RMSE, takes relative abundances into account and has been included in studies such as Mathon *et al.* (2021) and Xiong *et al.* (2022).

The data from the positive control method, where a closed-reference OTU picking method was used, indicated that all the members of the mock community were included in the simulated datasets and that they could be referenced back to their taxonomies with all filtering parameters included (Table 4.2). Although OTU richness was perfectly inferred, reflecting the diversity of the mock community, as well as perfect values for precision and sensitivity, RMSE was still not obtained as reflected by a value of zero. That means that true abundances could still not be processed by a method using species' DNA templates for clustering. This has also been observed by other dietary metabarcoding studies focussed on identifying plant food items with chloroplast barcodes such as Bell *et al.* (2021). The analysis further identified VSEARCH *de novo* clustering and DADA2 as approaches that were suitable for species discrimination from forward reads sequenced from both the *rbcL* and *trnL* chloroplast regions. Since OBITOOLS could not process data where forward and reverse reads are not merged, it was eliminated from further comparisons. VSEARCH methods were identified as the most sensitive approach for species recovery from both the *rbcL* and *trnL* datasets as well as the pipeline containing methods that identified the abundances closest to the specified, observed abundances indicated by the lowest RMSE values. DADA2 methods,

specifically DADA_1_RA_LULU has been identified as the method performing equally well for both barcode datasets in terms of precision, RMSE, and sensitivity.

OBITOOLS methods

By using the mock community sequences obtained from GenBank/BOLD as a reference database, *ecotag* was only able to classify reads to clade level for *rbcL*. This means that no sequence differentiation took place since all mock community members stem from the same clade. A possible explanation is that “raw” query reads were 301 bp long, as specified by the MiSeq error model implemented by INSILICSEQ, and reference reads were 500 bp+ (*rbcLa*: 552 bp) since a consensus sequence was created by the publishers of the sequence data on public reference databases by aligning forward and reverse reads. When the query sequences (forward reads) were used as a reference in *ecotag*, they could be successfully identified to species level, demonstrating the influence of the read length discrepancies between the *rbcL* queries and reference reads. However, implementing the OBITOOLS pipeline by using a different database and only forward reads would be misleading in comparison to other bioinformatics pipelines that can use reference sequences directly. The OBITOOLS bioinformatics pipeline was therefore not suitable for the identification of plant species barcoded with the *rbcL* region where only the forward read was used as a query against a reference database of consensus sequences. The identification of two families and three subsequent genera indicated that OBITOOLS methods led to sequence differentiation for the *trnL* mock community dataset, however, this differentiation only took place with the applications of frequency filters since the command for abundance filtering, *obiclean*, did not compute the unmerged sequence headers. A study that also used only the forward reads of the COI barcode for bony fishes to identify species via OBITOOLS, did not report any filtering by abundance and rather employed a frequency filter via *obigrep* that removed sequences containing less than 50 reads (Nobile *et al.*, 2019). Other studies using the OBITOOLS bioinformatics pipeline to analyse the plant components of animal diets had great success. Such studies include Kartzinel *et al.* (2015) who used the *trnL*(UAA) P6-loop as a barcode to illustrate dietary niche partitioning of African large herbivores, Mallot *et al.* (2018) who used this region along with *rbcL* to analyse the plant component in the diets of white-faced capuchins and Kowalczyk *et al.* (2019) that used whole chloroplast intron *trnL*(UAA) for the dietary analysis of the European bison. However, these studies report forming paired-end reads or merging, at the beginning of the OBITOOLS workflow using the *illumina-pairedends* tool, which is not comparable to the use of unmerged, forward reads. Although OBITOOLS could be applied with some level of success to the *trnL* mock community, it was eliminated from further analysis with the LULU post-clustering curation algorithm as well as further inclusion in pipeline and filtering method comparison as it can only be applied to one of the chosen plant barcoding regions that will be used for taxonomic differentiation in this study.

VSEARCH and DADA2 methods

The analysis of feature richness (Figures 4.6 and 4.7) as well as the percentage of these features that could be matched to their taxonomies (Fig. 4.8) for DADA2 and VSEARCH methods revealed a negative correlation: the closer the number of reads was to the original diversity of the mock community consisting of 23 members, the higher the percentage of assigned reads were. This resulted in DADA2 methods assigning 100% of the query sequences to taxonomy in contrast to VSEARCH methods assigning fewer than 9% of queries for the *rbcL* dataset and 16% for the *trnL* dataset. Feature richness is a debatable quantifier in metabarcoding diet analysis (Deagle *et al.*, 2019), but has been employed as such by studies such as Matezans *et al.* (2017) where a strong positive correlation is shown between the percentage of sequences and biomass of sample mixtures of below-ground plant communities. However, the taxonomic assignment of this study showed that although a method displays a feature richness that is very close to the mock community diversity, it does not mean that all members could be identified, and, as the RMSE values showed, does not mean that the mock community members could be identified in the correct proportions.

The results from the comparison criteria per method revealed significant discrepancies between both the bioinformatics pipeline as well as the applied filtering parameters, highlighting the need to choose a bioinformatics pipeline with its subsequent filtering parameters that are suited to fulfil the specific aims set out to answer the research question. VSEARCH methods were able to recover most species from the mock community, apart from the forb *Evolvulus alsinoides* in the *rbcL* mock community dataset, with better reflections of their actual abundances as input into INSILICOSEQ, indicated by RMSE values ranging from 4.48 to 2.07 for the *rbcL* dataset and 1.06 and 0.3 for the *trnL* dataset. However, community richness retrieved with these methods was greatly overestimated and performed much worse than the DADA2 method in terms of precision with values ranging between 0.06 and 0.003 for the *rbcL* dataset, and 0.1 and 0.004 for the *trnL* dataset, where DADA2 was able to score between 0.6 and 0.3 for both the *rbcL* and *trnL* datasets. This means that features in the VSEARCH methods are identified, to a much greater degree than DADA2, that do not exist in the mock community, leading to a high false positive rate. In the case of plant metabarcoding studies, this is problematic since plant reference databases are incomplete due to a lack of a standard barcoding region. Therefore, not all species in the study area may be sequenced with the chosen plant barcodes and will not be identified from the environmental sample. In the case of South African savannas, it is well known that forbs have been neglected in literature and grassland management models (Siebert & Dreber, 2019) and are not well represented on study area species lists and public barcode reference databases. All these factors will have an impact on the interpretation of taxonomic assignments retrieved by bioinformatics pipelines, specifically whether features that have not been assigned to their taxonomy are the result of the clustering or denoising algorithm applied which overestimated the community richness or whether the unassigned features indicate undiscovered plant community richness. The precision criterion is therefore given top priority when ranking the methods used in this study, thereby

not necessarily recovering the entire plant community, but generating high confidence in true positives and not overestimating the richness of the sampled community.

Chosen method

The bioinformatics pipeline's methods which ranked the best according to the precision criterion for the *rbcL* and *trnL* mock community datasets was that of DADA2, the clustering free software package implemented in R. The methods applied for this pipeline had precision values, as mentioned, ranging between 0.6 and 0.3 for both the *rbcL* and *trnL* datasets. The application of the LULU post-clustering curation algorithm, where erroneous ASVs are collapsed into their parent ASVs reduced the community richness by ~80% for the *rbcL* dataset and ~50% for the *trnL* dataset, except for the frequency filter which removes features occurring only once in a sample, that could reduce the richness in the same degree without the application of post-clustering curation. The lowest community richness of 64 for *rbcL* and 132 for *trnL* and highest precision values of 0.68 for *rbcL* and 0.36 for *trnL* were obtained by DADA_1_LULU and DADA_1_RA_LULU, meaning there was no difference in applying the frequency filter before LULU or applying the frequency filter before the abundance filter and then LULU. Based on these results the DADA_1_RA_LULU filter was chosen as the method offering the optimal pipeline and filtering parameters to recover the mock community species with the highest precision and therefore the lowest identification of false positives. The inclusion of the RA filter is also recommended to ensure the removal of PCR or sequencing errors as well as sequences occurring in low frequencies that cannot be simulated. This study aims to determine, firstly and with high precision, the plant species occurring in herbivore diets. The cost of the removal of rare taxa occurring in low frequency is therefore paid rather than including features that may not be assigned by the reference databases and are being interpreted as undiscovered biodiversity. Although the abundance values retrieved for this method are far from perfect, in fact, it resulted in the highest RMSE values, studies have shown that the quantification of the abundances of food items, even when using a dataset derived from faecal samples of cattle that has been fed a known diet can never be perfect (Scasta *et al.*, 2019). Furthermore, the use of simulated mock communities to represent real communities contained in faecal samples has some limitations: the biases associated with the degraded nature of the DNA recovered from faeces cannot be simulated and the real biodiversity contained in a faecal sample, together with contaminations in the field or laboratory as well as the appearances of rare taxa in low frequencies cannot be predicted or simulated. Mallot *et al.* (2018) also noted that controlled feeding studies where the amount of food consumed as well as known gut passage rates are key to determining whether relative read abundances (RRA) can confidently be used to quantify the proportions of food species in animal diets. Therefore, one could use presence/absence data to report the identified components of the diet as a species lists, as frequency of occurrence data of a certain species across all samples, or one could report the species in their relative read abundances per sample (RRA) based on feature richness (e.g.,

Kartzinel *et al.*, 2015), with caution and realisation of all the biases and limitations associated with this quantitative metric.

Conclusion

The selection of an optimal bioinformatics pipeline and filtering parameters to apply to NGS data received from DNA isolated from faeces and barcoded with the *rbcL* and *trnL* primer by first applying the methods to a mock community have highlighted some important considerations. Not all bioinformatics pipelines are suitable for all metabarcoding studies, for example, OBITOOLS is not suitable for the identification of plant species from forward *rbcL* reads. The nature of the research question will be most important when choosing between a clustering method such as VSEARCH which groups sequences based on a similarity threshold, or a denoising method such as DADA2 which groups sequences into ASVs based on small nucleotide differences. The clustering method, VSEARCH, showed that although community richness is greatly overestimated, members of the sampled community are more fully recovered in their expected abundances. The denoising method, DADA2, showed a superior performance when the research question needs to assess community richness and would benefit from precise estimations of features that can be assigned to their taxonomy and not the inclusion of reads that cannot be identified that would be misleading to the diversity contained in the environmental sample. This study also highlights the need for feature table filtering where low-frequency features are removed that may represent contamination, PCR, or sequencing errors and in doing so results in fewer sequences that cannot be taxonomically assigned or the inclusion of features resulting from contamination in the final taxonomic classifications that incorrectly overestimates community richness and do not reflect animal foraging behaviour. The DADA2 method with frequency, abundance, and post-clustering (denoising, in this case) curation is therefore chosen as the bioinformatics method to identify plant species from cattle diets foraging in an eastern semi-arid South African savanna and feature richness will be used to estimate the presence/absence data of the food items in these diets and RRA will cautiously and with a full understanding of the associated biases and limitations, be used to quantify the contribution of each. Future comparisons of bioinformatics approaches for the identification of plant species from herbivorous diets should focus on reducing biological and technical biases associated with metabarcoding studies to gain a better understanding of the proportions of each feature identified and therefore confidently quantify the relative read abundance of food items in dietary analyses. In the future, scientists may also choose to use a reference-based clustering approach, like the VSEARCH method used as a positive control which will deliver very accurate performance criteria values but will exclude the discovery and undermine the importance of unsequenced or novel taxa that is not present on the reference sequence database.

Chapter 5 Metabarcoding dietary analysis of cattle foraging at two different sites in an eastern semi-arid South African savanna

Introduction

This final chapter will investigate what the influence of *in-silico* filtering methods applied will have on taxa detection rate, of two different bioinformatics approaches, specifically DADA_1_LULU and DADA_1_RA_LULU. These approaches performed equally well when applied to a mock community (Chapter 4). This section will therefore culminate both the efforts of the compiled reference database for *rbcL* and *trnL*(UAA) barcoding regions of species commonly found in an eastern semi-arid South African savanna (Chapter 3), and the identification of an optimal bioinformatics approach (Chapter 4) to identify the diets of cattle, which are represented by DNA sequenced from their faeces with the *rbcL* and *trnL*(UAA) plastid regions. The final bioinformatics approach suited for plant identifications from cattle foraging in a semi-arid eastern South African Savanna will be identified per chosen barcoding region based on its ability to recover the highest species diversity in terms of plant families and genera. The chosen bioinformatics approach will be used to identify key species in the diets of cattle foraging in both Welterdiend and Syferkuil as well as to the relative importance of forbs as food items.

Two approaches, namely relative read abundance (RRA) and frequency of occurrence (FOO) are used to derive quantitative estimates of diets from next-generation sequencing data. Diet quantification in terms of FOO is calculated by using read counts per sample as well as presence/occurrence data of this given food item across samples (Biffi *et al.*, 2017; Kartzinel *et al.*, 2015). The contribution of a certain food item to the herbivore diet is then directly proportional to the read count within the population of samples. FOO, therefore, provides quantitative information about animal diet at the population level, i.e., all sampled individuals. This approach, however, requires many sampled individuals, especially in the case of diverse diets (Mata *et al.*, 2019). Another concern regarding the use of this quantification method is that although a food item may occur frequently throughout the population, its ecological importance may be overestimated when individuals in the population each consume it at low frequencies. The second quantification approach, RRA, relies on calculating the relative read abundance, or the relative frequency of sequencing reads per sample/individual, where the number of reads is expected to be proportional to the relative biomass of the associated food items. As stated in Chapter 4, the common occurrences of biases associated with the metabarcoding approach (see Alberdi *et al.*, 2018; Lamb *et al.*, 2019) have resulted in studies rather opting to quantify diets via FOO (Iwanowicz *et al.* 2016; Scasta *et al.*, 2019; Scasta *et al.*, 2020). However, any quantitative results retrieved from metabarcoding are unlikely to be perfect since the process of analysis alone, not considering sample contamination, primer bias, incomplete references databases, etc., consists of many steps and is susceptible to a certain degree of biases and errors. This can

be seen, for example, in a study where cattle are fed known rations, the frequency of occurrence data retrieved from the faecal samples still did not reflect the known dietary composition (Scasta *et al.*, 2019). In terms of dietary studies, RRA is frequently employed to represent the proportion of consumed biomass, and studies specifically focused on LMH diets using the *trnL* approach have supported the use of this metric, at least for taxonomic assignments at the family level (Kartzinel *et al.*, 2015; Willerslev *et al.*, 2014). Moreover, studies such as Gebremedhim *et al* (2016), Kartzinel *et al* (2015), Willerslev *et al* (2014), and Mallot *et al.* (2018) have also shown that quantitative results retrieved by RRA are comparable to those based on FOO but are less sensitive to the inclusion of common biases associated with dietary metabarcoding studies. Examples of such biases are the inclusion of low abundance reads that often do not represent rare taxa, but result from incidental ingestion, sample contamination, or PCR and sequencing errors (Deagle *et al.*, 2018).

The quantification success of data presented in a metabarcoding study is highly dependent on the extent of taxonomic coverage and the quality of records available in the DNA barcode reference database used (Lamb *et al.*, 2016; Weigand *et al.*, 2019). In Chapter 3 a custom barcode reference database was presented for the *rbcL* gene and *trnL*(UAA) intron of plant species that are potential herbivore foraging targets and commonly found in semi-arid savannas of eastern South Africa. A study-area-specific species list of 755 species was compiled and their corresponding sequences were mined from GenBank, BOLD, and NWU repositories according to specific quality criteria to ensure accurate taxonomic coverage and resolution. To ensure that this reference database could be used to discriminate between potential dietary species in a NGS dataset, it was evaluated by testing for the presence of a barcode gap, identifying a data-appropriate identification threshold, and determining the identification accuracy of reference sequences via primary distance-based criteria. Although the use of the reference database was aimed at the optimal scenario of identifying unknown queries to species level, it was decided to evaluate the diets of Welverdiend and Syferkuil cattle only to genus level for the following reasons: i) neither *rbcL* nor *trnL* has particularly high rates of species level discrimination (Valentini *et al.*, 2009; Taberlet *et al.*, 2007; Hollingsworth *et al.*, 2011), ii) operational taxonomic units (OTUs), such as those reported in dietary studies, are difficult to determine at the species level, as previous research have shown (De Barba *et al.*, 2014; Buglione *et al.*, 2018), and iii) it can be difficult to obtain an amplicon length long enough to reliably identify samples down to the species level (Deagle *et al.*, 2010). This problem is mostly brought about by the degradation of food items in faecal samples (Deagle *et al.*, 2006).

Except for the consideration of biases associated with metabarcoding studies when quantifying dietary data and the completeness and discrimination efficiency of the reference database, *in-silico* steps (e.g., choice of filtering parameters and clustering approaches) are known to directly influence taxonomic assignments and diversity quantifications (Frøslev *et al.* 2017). The bioinformatics approach chosen to filter and analyse NGS data is critical to mitigate the impact of biases and result in final taxonomic assignments that are truly

representative of the animal's foraged ecosystem. The comparison of the bioinformatics approaches of VSEARCH, OBITOOLS, and DADA2 with a certain frequency and abundance-based filtering as well as post-clustering curation, identified the DADA2 (DADA_1_RA_LULU) approach as the one that was able to retrieve the mock community composition with the highest precision. This approach included the inference of amplicon sequence variants (ASVs) based on an error model, rather than the traditional clustering of operational taxonomic units (OTUs) based on a similarity threshold. The filtering included in this approach consisted of a frequency filter removing all amplicon sequence variants (ASV) which occurs only once per sample (Filtering=1), an abundance filter which removes ASV occurring less than 0.005% across the sampled community (Filtering=RA), as well as the post clustering curation algorithm, LULU, which collapses erroneous 'daughter' ASVs into their parent ASVs (Filtering=LULU: Frøslev *et al.* 2017).

Despite that DNA metabarcoding pipelines frequently involve abundance-filtering, the relative abundance (RA) cut-off used to decide between the presence or absence of taxa differs significantly between research (Alberdi *et al.*, 2018). A 1% threshold may be appropriate for many dietary studies, according to Deagle *et al.* (2019), but levels between 0% and 5% have been used (Alberdi *et al.*, 2018; Kartzinel *et al.*, 2019; Ter Schure *et al.*, 2021). Littleford-Colquhoun *et al.* (2022) highlight the challenge of filtering NGS datasets of dietary studies in that important food items can occur at low relative abundances and may be filtered out by common bioinformatics procedures. They warn that the continual practice of their exclusion will only reinforce erroneous descriptions of terrestrial ecosystems and animal diet descriptions and further inhibit new understandings of ecological networks. To overcome this challenge and to heed their warning, it was decided to include a less stringent filtering method included in the bioinformatics approach that performed equally well when applied to the mock community (see Chapter 4 Figure 4.9): DADA2 with the same frequency-based filter and post-clustering curation, with the exclusion of the RA filter (DADA_1_LULU).

Materials and methods

5.1.1 Study sites and sample collection

All methodological approaches preceding the bioinformatics analysis of NGS data retrieved from 24 cattle at both the Welverdiend and Syferkuil *Draught Act* study sites were described in Chapter 2. These include (i) a description of the study area, (ii) plant material collection, DNA isolation, PCR amplification and sequencing, and (iii) faecal sample collection, DNA isolation, PCR amplification and sequencing.

5.1.2 Bioinformatics analysis

5.1.2.1 Raw data quality control

In preparation for bioinformatics analysis, raw reads were retrieved from the MiSeq platform for both the *rbcL* and *trnL* markers of Welverdiend and Syferkuil cattle was first visualised by producing FASTQC

reports and a concatenation of these reports per marker and study site with the MULTIQC tool (v. 1.13: Ewels *et al.*, 2016). Hereafter, a dataset refers to either the Welverdiend *rbcL*, Welverdiend *trnL*, Syferkuil *rbcL*, or Syferkuil *trnL* assemblages. The quality reports generated by both FASTQC (per sample quality information) and MULTIQC (per dataset quality information) contain information that should be considered when trimming the sequence reads according to sequence length, the content of adapters added during sequencing, primers left by the barcoding regions, as well as poor read quality towards the beginning or end of sequences. These reports also explain discrepancies between the assigned taxonomy retrieved for the datasets in terms of the quality of raw sequence data.

5.1.2.2 Preparation of feature tables via DADA2

After the quality of the Illumina paired-end reads was inspected, the presence of PCR primers: the chloroplast subunit ‘a’ of *rbcL*, *rbcLaF* (ATGTCACCACAAACAGAGACTAAAGC: Levin *et al.*, 2003), *rbcLaR* (GTAAAATCAAGTCCACCRG: Kress & Erickson, 2007) and the *trnL*(UAA) whole chloroplast intron (Taberlet *et al.*, 1991), *trnLc* (CGAAATCGGTAGACGCTACG) and *trnLd* (GGGGATAGAGGGATTGAAC) in the forward, forward complement, reverse, and reverse complement directions were removed using CUTADAPT (v. 4.1: Martin, 2011). Quality filtering of the raw reads could then commence with the use of the DADA2 package (v. 1.24.0: Callahan *et al.*, 2016) in R (v. 4.2.0: R Core Team, 2022). This included visualising the quality profiles of each dataset, not as part of the filtering and trimming workflow steps, but to serve as a validation of the quality information seen in the MULTIQC reports, to visualise the number of reads, before and after the removal of primers, and sequence adapters as well as the subsequent impact that it had on the quality scores observed for each sample per dataset. The first step of most bioinformatics approaches would be to merge forward and reverse reads, but it will not be applied in this study for reasons listed in Chapter 0. The reads left for each dataset after the removal of barcoding primers were further trimmed by *filterAndTrim* of DADA2 with the following settings: number of ambiguous bases (“N”s) allowed: 0 (maxN=0), minimum length of reads allowed to pass the filter, according to the short nature of reads retrieved from faecal samples: 50 (minLen=50), the truncation of reads at the first instance of a Phred quality score of less than 20 (trunQ=2), and the retainment of reads with a maximum amount of two sequencing errors (maxEE=2). Reads left after the application of *filterAndTrim* were used in a DADA2 error learning algorithm that repeatedly performs sample inference on each sample until an optimal convergence is reached, with the command *learnErrors*. At this point, the filtered sequences were dereplicated with *derepFastq* and sample inference could commence. This learned error rate as well as the dereplicated sequences are then used for sample inference where amplicon sequence variants (ASVs) are created and presented as a sequence table with the *makeSequenceTable* command. An additional filtering step was the removal of chimeric sequences which was not applied in the mock community since NGS simulators like INSILICOSEQ do not have the capability of generating these sequence artifacts. The removal of chimeric sequences from the constructed sequence table was done by

removeBimeraDenovo implemented by DADA2. After the filtering and trimming steps are completed for the DADA2 bioinformatics approach, a feature table (named *Feature_table.tsv* in the following section) and FASTA sequences representing the ASVs in the feature table (named *dada2.fasta* in the following section) were exported for downstream analysis.

5.1.2.3 *In-silico* filtering strategies and taxonomic assignments

The two bioinformatics approaches that was applied to analyse the dataset received from cattle faecal samples from Welperdiend and Syferkuil call for a frequency-based filter combined with the LULU curation algorithm (DADA_1_LULU), as well as the addition of an abundance-based filter (DADA_1_RA_LULU). For post-clustering filtering purposes, as well as taxonomic assignment, certain files needed to be imported as QIIME2 artifacts. It is also important to note that the reference database for both the *rbcL* and *trnL* marker had to be split into FASTA files containing reference sequences as well as process identities or accession numbers as sequence headers, and a tab-delimited file containing the taxonomic descriptions of these reference sequences, with the commonality of the accession numbers or process identities in the first column. This was done to ensure compatibility of the taxonomic assignments to the format of a PHYLOSEQ object which will be analysed in R. The following files were imported as artifacts for further analysis: ASVs resulting from the inference of error rates in DADA2 (*qiime tools import --input-path dada2.fasta --output-path queries_dada2.qza --type 'FeatureData[Sequence]'*), reference database sequences contained in a FASTA file per marker (*qiime tools import --input-path references_[rbcL/trnL].fasta --output-path references_[rbcL/trnL].qza --type 'FeatureData[Sequence]'*), and reference database taxonomic descriptions (*qiime tools import --type 'FeatureData[Taxonomy]' --input-format HeaderlessTSVTaxonomyFormat --input-path reference_taxonomy_[rbcL/trnL].tsv --output-path reference_taxonomy_[rbcL/trnL].qza*). The feature tables extracted from the sequence table created in DADA2 had to be converted to a BIOM file before it could be imported as a QZA file: *biom convert -i Feature_table.tsv -o Feature_table.biom --table-type "Table" --to-json; qiime tools import --input-path Feature_table.biom --type FeatureTable[Frequency] --output-path Feature_table.qza*.

Abundance and frequency-based filtering was applied to the imported feature table artifacts and followed the same steps as given in Chapter 4.1.4 of the materials and methods. For this section, Filtering=1 is applied on its own as well as the combination of Filtering=1 and Filtering=RA to determine the impact that differing filtering parameters will have on the recovered diversity of plant species from herbivorous faecal samples. The two ASV tables resulting from the application of these filters were then piped into post-clustering curation with the LULU algorithm. LULU starts by creating a match list of ASVs contained in a feature table by blasting them against each other using the BLASTN application (v. 2.12.0) in QIIME2. The following recommended BLAST parameters were chosen to create a LULU match list: minimum percentage identity (*-perc_identity*) of 84 and a query coverage high-scoring sequence pair percent (*-*

qcov_hsp_perc) was set to 80. The match list for each ASV table contains the name of the query and reference sequences next to a percentage of identical matches. This match list as well as the ASV files for DADA_1_LULU and DADA_1_RA_LULU were uploaded to R for further analysis. The R package *lulu* (v. 0.1.0: Frøslev *et al.*, 2017) uses the match list and ASV table as input to filter the dataset represented in the ASV table for ASVs that occur in very low abundances and could be PCR errors, sequencing errors or contamination, and either collapses them into their parent ASV or discards them from the dataset.

Finally, these curated ASV tables are imported to QIIME2 to assign their taxonomy. The rest of the taxonomic assignment for DADA_1_LULU and DADA_1_RA_LULU datasets after *in-silico* filtering follows the same protocols as described for the mock community (Chapter 4.1.4) with the *qiime feature-classifier classify-consensus-blast* parameter of percentage identity match specified changed to 80% for unstimulated data rather than 99%, as for the mock community. This command produces a QZA file that, once unzipped, contains a TSV file describing the taxonomic assignments of each ASV with a column with confidence values ranging from zero to one, one being the most confident. BIOM files were created to represent the taxonomic assignments and representative sequences for both approaches by the following command: *biom add-metadata -i Feature_table_LULU.biom -o DADA_1_RA_LULU.biom --observation-metadata-fp BLAST_classification.tsv --observation-header OTUID,taxonomy,confidence*. These BIOM files were then imported to R as PHYLOSEQ objects for analysis with the *phyloseq* package (v. 1.40.0: McMurdie & Holmes, 2013) in R.

Removal of suspected contamination

As part an additional quality control step, the unassigned ASVs resulting from taxonomic assignment by the compiled reference database of both bioinformatics workflows for each dataset were blasted against the contents of NCBI's non-redundant nucleotide (nt) database on the 13th of September 2022. This was done to establish i) the degree of influence of poor-quality profiles retrieved for certain samples especially evident in the Syferkuil dataset in terms of species recovery (i.e., whether the unassigned reads had bacterial/other origins), and (ii) whether the ASVs resulting from these pipelines could not be assigned due to poor taxonomic coverage of the compiled reference database. Reads which were assigned to species other than plants, with a percentage identity of 80%, query coverage of 80% and e-values smaller than 0,001 were removed from the feature (ASV) tables for each dataset since they are not a reflection of the inability of the compiled reference database to identify the composition of herbivores foraging in an eastern semi-arid South African savanna and cannot truly be classified as “unassigned”

5.1.2.4 Dietary richness evaluation

Taxonomic assignments of sequences retrieved from the Syferkuil and Welverdiend cattle faecal samples with both barcoding markers and both bioinformatics approaches were visualised with read abundance bar

plots, bar plots showing the frequency of occurrence of presence/absence data for each dataset as well as stacked bar plots showing the relative read abundance of sequencing reads assigned to family and genera per sample individual. All statistical analyses were carried out in R (v. 4.2.0: R Core Team, 2022).

Read abundance

Read abundances were visualised in a bar plot showing the number of detected ASVs for each marker dataset per study site and bioinformatics approach by the R package *phyloseq*. These bar plots are employed to summarise the number of detected ASVs in relation to those that could not be assigned to their taxonomic level (NA reads), which is helpful since the proportion of NA reads in relation to the rest of the reads can be visualised. For RRA and FOO calculations and visualisations, NA reads were subsequently omitted since the diets of cattle needed to be quantified and visualised with reads that could be taxonomically assigned by the reference database.

Frequency of occurrence (FOO)

FOO was used to visualise the differences in commonalities of frequencies of taxa detected between datasets and bioinformatics approaches across all samples and was visualised as horizontal bar plots with the *ggplot2* package (v. 3.3.6: Wickham, 2016). This summary metric will validate the bioinformatics approach chosen according to the diversity detected in terms of families. The frequency of occurrence percentage for families and genera is calculated by summing all present values per specified family or genus (1 = present; 0 = absent) for samples in the dataset, dividing it by the number of samples and multiplying this total by 100 to retrieve FOO%.

Relative read abundance (RRA)

RRA graphs were prepared in *ggplot2* from the ASV table that was transformed to contain the query sequence identity, sample, and the number of times each ASV appears in a certain sample. RRA was therefore calculated per sample as this ASV count is divided by the sum of the count of all ASVs in a particular sample. This data is visualised as 100% stacked bar plots indicating the proportion of reads assigned to a family or genus level per dataset per bioinformatics approach.

Results

5.1.3 Sequencing results

After the data were subjected to primer removal by CUTADAPT, the quality of reads in each database was inspected. MULTIQC is capable of plotting data contained in the quality report, which was very useful for visualising the number of sequencing reads generated by the Illumina MiSeq platform for both markers per study site which could be plotted against the average sequencing length per sample (~150 bp) to ascertain

whether most sequences matched expected length of reads are present in terms of DNA gathered from faecal samples in general (Fig. 5.1).

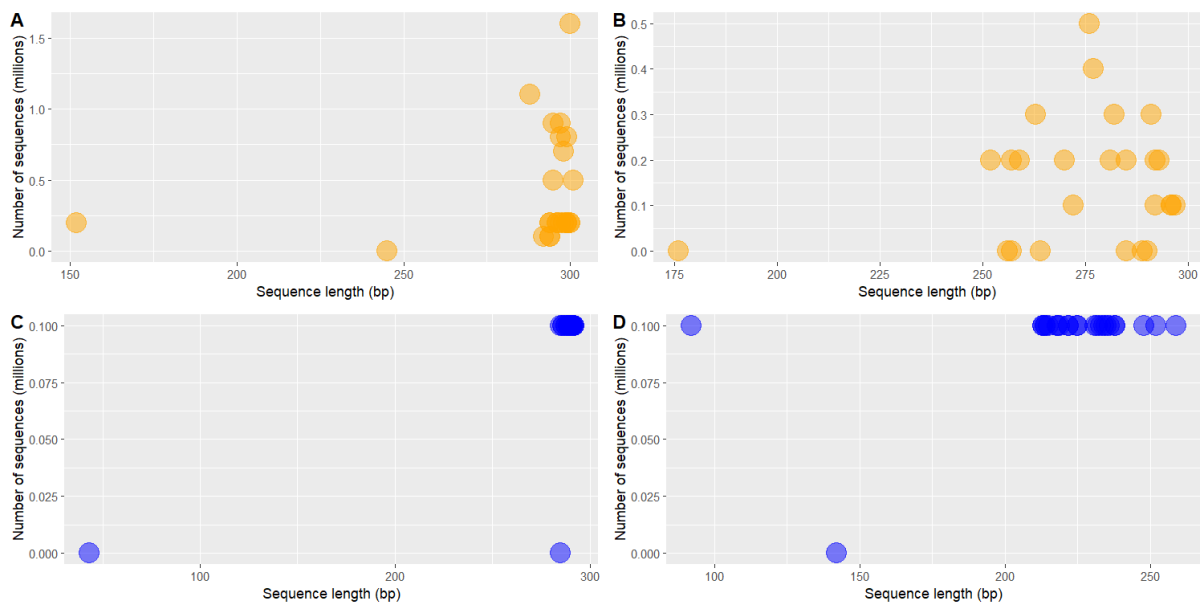


Figure 5.1: Scatterplots representing the average sequence length per sample plotted against the number of sequenced reads per sample retrieved from the Illumina MiSeq platforms for reads sequenced in the forward direction for the Welverdiend A) *rbcL* and B) *trnL* datasets as well as the Syferkuil C) *rbcL* and D) *trnL* dataset. The data for these plots were exported from MULTIQC to R to create visualisations with *ggplot2* for the sake of image resolution and figure uniformity throughout the study.

Syferkuil showed some outlier samples (< 150 bp), where the control sample of *rbcL* contained sequences with an average length of 43 bp. Outliers were also detected in the Syferkuil *trnL* sample where sample 43 contained sequences with an average length of 142 bp and the control contained sequences with an average length of 92 bp. MULTIQC quality reports further revealed typical trends of low-quality sequences near the beginning and end of sequence reads and no dataset contained next-generation sequencing adapters and read lengths. The *rbcL* and *trnL* dataset of Welverdiend was an example of good-quality data and the occurrence of Ns and low-quality Phred scores could easily be filtered out for all samples. MULTIQC flagged sample 43 from Syferkuil's *trnL* dataset as well as the control in the *rbcL* dataset as having very poor-quality sequences, a high inclusion of ambiguous bases and very short sequence lengths. These quality profiles were reflected in plots received by *plotQualityProfile* in DADA2 (Supplementary Figures 1-4).

5.1.4 DADA2 bioinformatics workflow

5.1.4.1 Filtering and trimming

The application of *filterAndTrim* resulted in the removal of reads containing ambiguous bases, reads were truncated at base pair positions at the first instance of quality scores of less than 20 and reads that were shorter than 50 bp were discarded. A total of 10 209 509 reads were used as input for the Welverdiend *rbcL* dataset and 6 225 959 passed the filters. The rate of read removal for this dataset ranged from 33%-57% with sample 4 having the most reads removed and sample 5 the least (Supplementary Table 7). For the *trnL* dataset 3 894 691 reads were used as input and 2 145 297 passed the various filters applied in *filterAndTrim*. The rate of read removal for the *trnL* dataset ranged from 24%-74% with sample 1 representing the maximum and sample 14 representing the minimum removal of reads (Supplementary Table 8). Overall read retention for the Welverdiend dataset after DADA2 filtering was 61% for *rbcL* and 55% for the *trnL* dataset (Fig. 5.2).

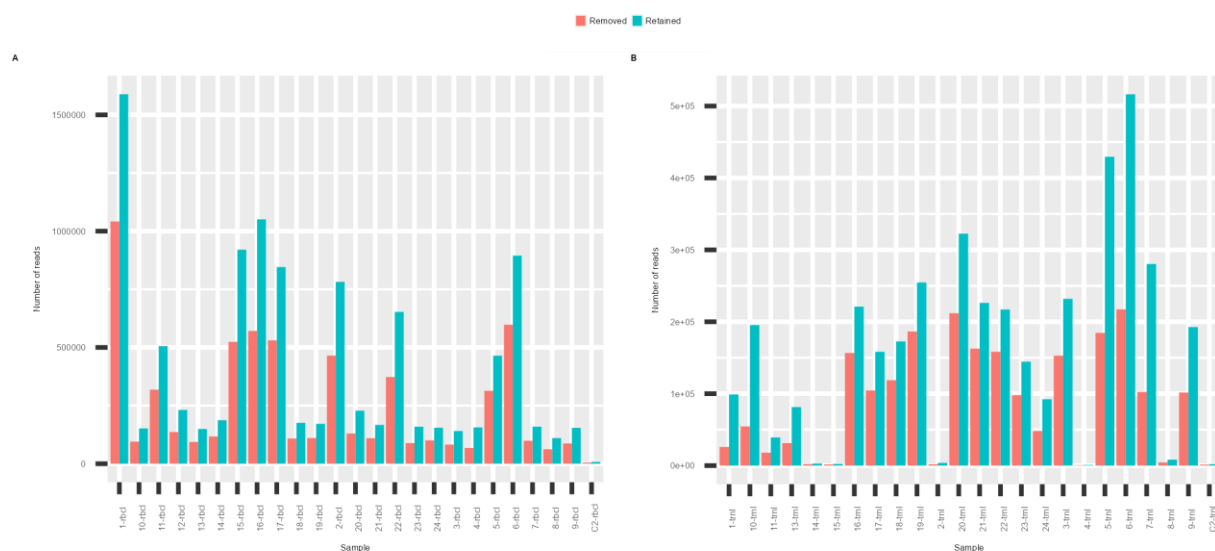


Figure 5.2: Stacked bar plots indicating the number of reads retained (blue bars) relative to the number of reads removed (red bars) as a result of sequence filtering and trimming in DADA2 for the Welverdiend A) *rbcL* and B) *trnL* dataset.

For the Syferkuil dataset, 1 858 096 reads were used as input for the *rbcL* dataset of which 1 566 036 passed the filters implemented in *filterAndTrim*. The rate of read removal for this dataset ranged between 14%-92% with sample 25 representing the least reads removed and the control sample representing the most removed (Supplementary Table 9). A total of 1 985 791 reads were used as input to *filterAndTrim* for the *trnL* dataset of which 1 290 922 passed. The rate of read removal ranged between 26%-58% with sample 27

representing the minimum and sample 43 representing the maximum read removal (Supplementary Table 10). Overall, Syferkuil retained 84% of its *rbcL* reads and 65% of its *trnL* reads (Fig. 5.3)

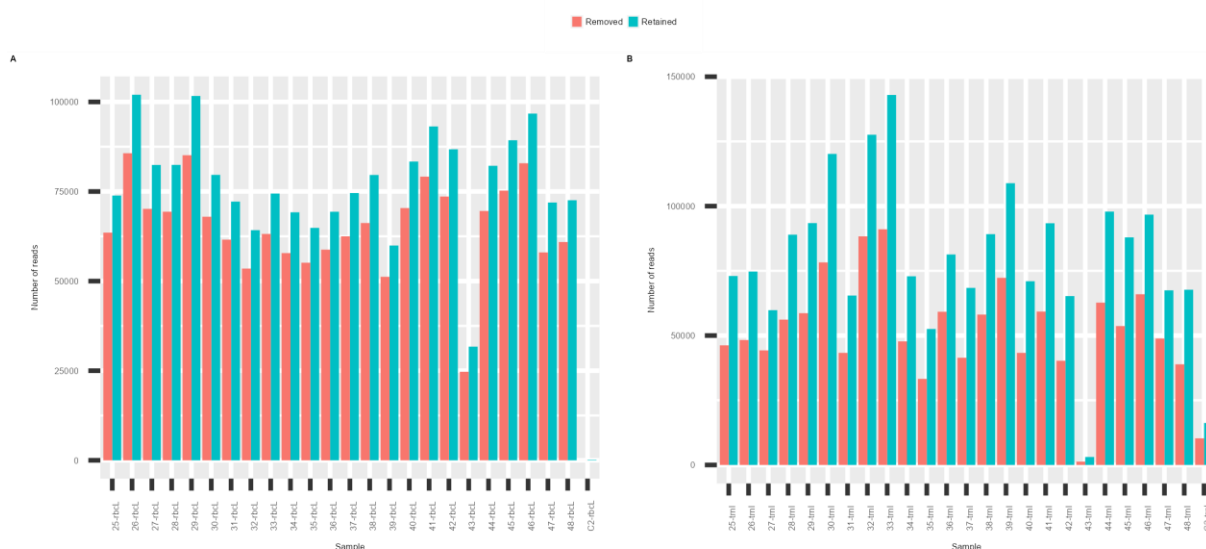


Figure 5.3: Stacked bar plots indicating the number of reads retained (blue bars) relative to the number of reads removed (red bars) as a result of sequence filtering and trimming in DADA2 for the Syferkuil A) *rbcL* and B) *trnL* dataset.

5.1.4.2 Analysis of sequence variation

The core sample inference algorithm was applied to the trimmed and filtered sequence dataset for each dataset after learning the parametric error models. The error models learned for each dataset seemed reasonable according to the DADA2 tutorial (<https://benjjneb.github.io/dada2/tutorial.html> and <https://github.com/benjjneb/dada2/issues/975#issuecomment-602128434>) (Supplementary Figures 5-8) and after sequence dereplication, a DADA II-class object could be formed which contained the amplicon sequence variants. The dada-class objects for the Welverdiend dataset contained 2 523 403 ASVs for *rbcL* and 258 676 for *trnL*. The Syferkuil dataset contained 2 519 703 *rbcL* ASVs and 2 514 133 *trnL* ASVs. Chimera removal ensued on each dataset's ASVs which reduced the number of ASVs in the Welverdiend *rbcL* dataset to 4 989 550 and 1 852 442 for *trnL*. The reads in the Syferkuil dataset were reduced by chimera removal to 1 444 146 for *rbcL* and 1 234 380 for the *trnL* dataset. For sample-specific counts of reads remaining after the denoising process where ASVs are inferred from filtered and trimmed sequences as well as the number of reads remaining after chimera removal see Supplementary Tables 7 to 10.

5.1.4.3 In-silico filters

The application of the QIIME frequency filter where an ASV is removed if it occurs less than once in a sample, based on the suspicion that they may be PCR or sequencing errors, reduced the numbers of ASVs

seen in all the datasets drastically. For Welverdiend this reduction in ASVs was from 4 989 550 to 2 411 for *rbcL* and from 1 852 442 to 1 372 for the *trnL* dataset. The application of the frequency-based filter on the Syferkuil dataset resulted in a reduction from 1 444 146 to 3 365 ASVs for the *rbcL* dataset and 1 234 380 to 3 487 ASVs for the *trnL* dataset. The application of the QIIME abundance-based filter on top of the ASVs received by the frequency-based filter, where ASV occurring less than 0.005% in the dataset is removed, resulted in further reductions in ASV richness. The Welverdiend *rbcL* dataset was reduced to 629 ASVs and 660 ASVs for *trnL*. The number of ASVs in the Syferkuil *rbcL* dataset was reduced to 1 607 for the *rbcL* dataset and 1 527 for the *trnL* dataset. A reduction in the detection of unique ASVs between the two bioinformatics approaches with the application of the abundance filter was therefore more than 50% for all datasets.

These two bioinformatics approaches were then subject to post-clustering with the LULU algorithm. The final number of ASVs present in the Welverdiend *rbcL* dataset filtered only with the frequency-based filter was reduced from 2 411 to 2 058 and 1 372 to 1 217 for *trnL*. The Syferkuil dataset that was also only filtered with the frequency-based filter that contained 3076 ASVs for *rbcL*, previously 3 365, and 3 320 for *trnL*, which was previously 3487 with the frequency filter. The addition of LULU to the approach combining frequency and abundance-based filtering resulted in a reduction of 629 to 494 ASVs for the Welverdiend *rbcL* dataset and 660 to 569 ASVs for the *trnL* dataset. For Syferkuil this reduction was from 1 607 ASVs to 1 403 ASVs for *rbcL* and from 1 527 to 1 367 ASVs for *trnL*. For a summary of the number of ASVs identified by each approach per database see Table 5.1.

Table 5.1: Summary of the number of unique ASVs detected by each filtering approach as well as the percentage reduction in ASVs between filtering methods.

Database	Filtering approach	Number of ASVs	Percentage reduction (%)
Welverdiend <i>rbcL</i>	DADA2	4989550	
	Filter=1	2411	99,95
	Filter=1_LULU	2058	14,64
	Filter=RA_1	629	69,43
	Filter=RA_1_LULU	494	21,46
Welverdiend <i>trnL</i>	DADA2	1852442	

	Filter=1	1372	99,92
	Filter=1_LULU	1217	11,29
	Filter=RA_1	660	45,76
	Filter=RA_1_LULU	569	13,78
Syferkuil <i>rbcL</i>	DADA2	14441465	
	Filter=1	3365	99,76
	Filter=1_LULU	3076	8,58
	Filter=RA_1	1607	47,75
	Filter=RA_1_LULU	1403	12,69
Syferkuil <i>trnL</i>	DADA2	1234380	
	Filter=1	3487	99,71
	Filter=1_LULU	3320	4,78
	Filter=RA_1	1527	54,00
	Filter=RA_1_LULU	1367	10,47

5.1.5 Taxonomic assignment and diversity analysis

Removal of suspected contamination

ASVs that could be matched to non-plant references on NCBI's "nt" database with a query coverage and percentage identity of 80% as well as an e-value smaller than 0.001 were subsequently removed from further analysis (Supplementary Tables 11-16). For the DADA_1_LULU workflow 7 unique ASVs were removed from the Welverdiend *rbcL* dataset and 17 from the *trnL* dataset. For Syferkuil, 15 unique ASVs were removed from the *rbcL* dataset and 18 from the *trnL* dataset. No unassigned sequences obtained from the DADA_1_RA_LULU workflow for Welverdiend were removed for *rbcL* and *trnL*, however for Syferkuil, 493 unique ASVs were removed from the *trnL* dataset and 51 from *rbcL*. The blast results showed that, within the specified parameters, none of the ASVs that could not be assigned by the compiled

reference database could be assigned to a plant species in NCBI's "nt" database, except for one query sequence identity in Syferkuil's *trnL* dataset with the DADA_1_RA_LULU workflow (11d4c6dd6e7eea8d8aef49b1eefb0daa15938a79) where *Helixanthera parasitica* and *Socratina bemarivensis* were assigned.

Taxonomic assignment and diversity analysis of the remaining amplicon sequence variants

Taxonomic assignment was carried out against the compiled reference database (Chapter 3) for both bioinformatics workflows: DADA_1_LULU and DADA_1_RA_LU for each dataset. Bar plots (Figures 5.4, 5.6, 5.8 and 5.10) are used to visualise the number of assigned ASVs to family and genus levels as well as the number of reads that could not be assigned to their taxonomies. This visualisation is important since unassigned reads were removed for RRA and FOO calculations to better represent the diversity and dietary importance of known species. The FOO graphs are presented as results with the read abundance bar plots to highlight the taxonomic diversity and the fact that if a taxonomy is the most abundant it does not necessarily mean it is the most frequent.

The bar plot for *rbcL* sequences of the Welverdiend dataset assigned with ASVs from both bioinformatics approaches (Fig. 5.4) revealed that the number of assigned reads is generally greater than reads that could not be assigned by the reference database (Fig. 5.4) due to either the ASV being a result of PCR or sequencing errors, or the ASV belonging to a species that is not present on the reference database. The read abundance bar plots also revealed that all 25 samples, including the control, remained after the various filtering parameters. Sample 1 had the highest number of reads (ASVs) for both bioinformatics approaches with a total of 770 745 for DADA_1_LULU and 767 641 for DADA_1_RA_LULU, and the control sample had the lowest, with a total of 1 651 for both DADA_1_LULU and DADA_1_RA_LULU. The read abundance bar plots for this dataset further revealed that the plant families identified by the DADA_1_LULU and DADA_1_RA_LULU approach were similar (Fig. 5.4 A and B), except for DADA_1_RA_LULU being able to identify an additional two families: Acanthaceae and Annonaceae. The same trend, where the inclusion of the abundance-based filter results in a higher plant diversity identified for the *rbcL* dataset, extended to identifications on the genus level with three genera identified with the DADA_1_LULU approach (Fig. 5.4 C) and twelve with the DADA_1_RA_LULU approach (Fig. 5.4 D), where in addition to *Albizia*, *Combretum* and *Euclea* as identified by DADA_1_LULU, DADA_1_RA_LULU identified *Rhus*, *Enteropogon*, *Hexalobus*, *Diospyros*, *Dichrostachys*, *Philenoptera*, *Thunbergia*, *Malvastrum* and *Trachypogon*.

The FOO% graphs were constructed by identifying the presence/absence status of each family or genus, summing these values across all samples, and dividing it by the number of samples, in this case, 25 to get a percentage of frequency of occurrence (FOO%) (Fig. 5.5). The Welverdiend *rbcL* dataset that is filtered by the DADA_1_LULU approach showed that the family Malvaceae was least frequently utilised as a food

item with its FOO calculated to 8% and Fabaceae and Combretaceae the most frequently (FOO%=96%) (Fig. 5.5A). The same FOO% for these families are identified by the DADA_1_RA_LULU approach (Fig. 5.5B), with Acanthaceae also identified with the same FOO% as Malvaceae (8%) and Fabaceae and Combretaceae still calculated to FOO of 96%. The most frequently occurring genera in the Welverdiend *rbcL* dataset, according to the DADA_1_LULU and DADA_1_RA_LULU approach were *Albizia* and *Combretum* (FOO%=96%). The genus occurring the least frequent was *Euclea* (88%) identified by the DADA_1_LULU approach (Fig. 5.5C) and the genera *Thunbergia*, *Trachypogon* and *Malvastrum* (8%) by the DADA_1_RA_LULU approach (Fig. 5.5D). The bar and FOO% graphs for the Welverdiend *rbcL* dataset further revealed that the DADA_1_LULU approach was not able to identify the families of Anacardiaceae, Malvaceae and Poaceae to genus level since genera belonging to these families are not present in the read abundance graph and the FOO% graph. DADA_1_RA_LULU, however, was not only able to identify all the present families to genus level but in the case of Fabaceae, Ebenaceae and Poaceae, more than one genus was identified.

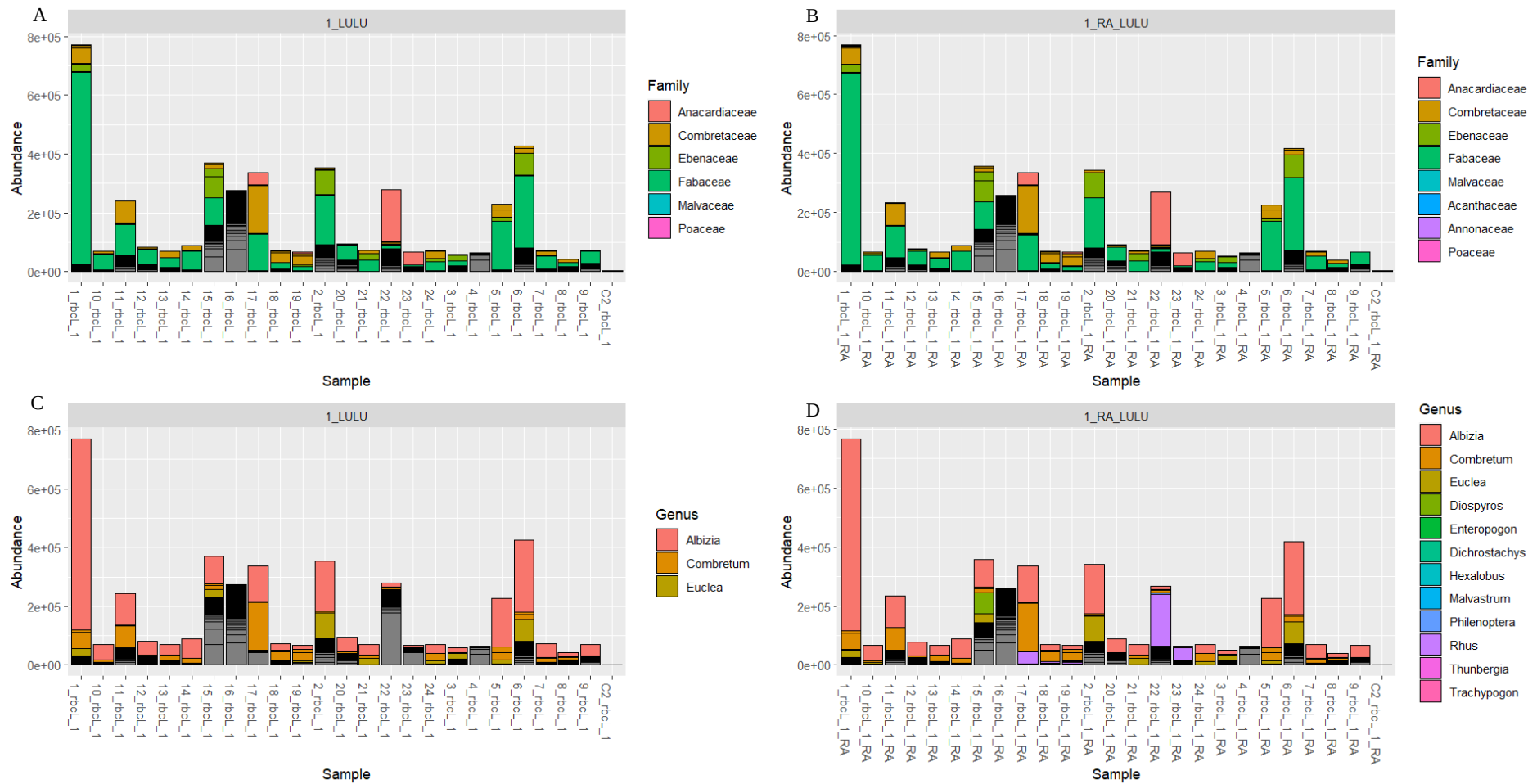


Figure 5.4: Bar plots generated from taxonomic assignments for the Welverdiend *rbcl* dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).

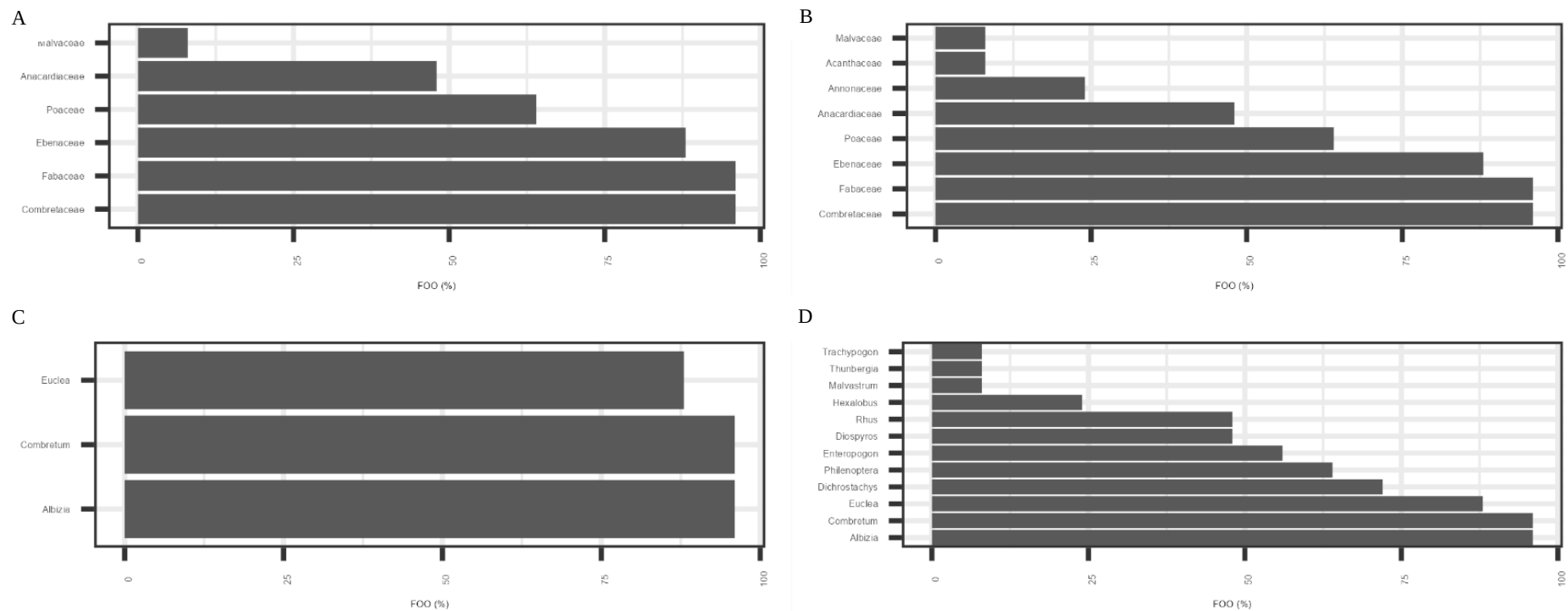


Figure 5.5: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Welverdiend *rbcL* dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).

The number of unassigned reads in the Welverdiend *trnL* dataset was more prevalent than in that of the *rbcL* dataset, but they did still not completely dominate the abundance plots. All 25 samples were present in the final filtered dataset, although, except for the control sample, very low read abundances were seen for samples 2,4,8,14 and 15 (Figure 5.6). The highest recorded read abundance in the DADA_1_LULU dataset was 165 628 ASVs (Fig. 5.6A and C) and 165 438 ASVs in DADA_1_RA_LULU (Fig. 5.6B and D) for sample 19 and the lowest being the control sample with 1 296 reads for both bioinformatics approaches. DADA_1_LULU was able to identify a higher taxonomic diversity in terms of families and genera than the DADA_1_RA_LULU approach. Except for the families Malvaceae, Ebenaceae, Anacardiaceae, Convolvulaceae, Fabaceae, Poaceae, Asparagaceae, Euphorbiaceae, Apocynaceae and Capparaceae identified by DADA_1_RA_LULU, DADA_1_LULU could identify Cyperaceae, Rubiaceae, Ehretiaceae, Asteraceae and Moraceae. For both approaches, there occurred fewer genera than there were families, which means that not all families were identified to the genus level. A total of 8 plant families for DADA_1_LULU and 4 plant families for DADA_1_RA_LULU could not be identified to genus level. The plant family Fabaceae was represented by *Senegalia* and *Vachellia*, Poaceae by *Eragrostis* and *Setaria*, Asparagaceae by *Asparagus*, Ebenaceae by *Euclea*, and Convolvulaceae by *Ipomoea* for both bioinformatics approaches. DADA_1_LULU had additional genera of *Ehretia* from the family Ehretiaceae and *Ficus* from the family Moraceae. The same trend, in terms of family and genus richness, than that observed for the Welverdiend *rbcL* dataset was observed, except that for *trnL* the highest species richness is tied to the use of DADA_1_LULU.

FOO% graphs for the Welverdiend *trnL* dataset revealed that Malvaceae was the most frequently observed plant family in the cattle diets with FOO being 92% for both bioinformatics approaches and the second most frequent families were Ebenaceae and Poaceae with 82% (Fig. 5.7A and B). The least frequent families across the dataset were Asteraceae (8%) for DADA_1_LULU and Capparaceae and Apocynaceae (12%) for DADA_1_RA_LULU. Both bioinformatics approaches agreed that *Euclea* (84%), *Senegalia* (72%), and *Ipomoea* (44%) were the most abundant genera with *Ficus* (12%) and *Ehretia* (12%) representing the least frequent for DADA_1_LULU and *Setaria* (16%) and *Eragrostis* (16%) for DADA_1_RA_LULU (Fig. 5.7C and D).

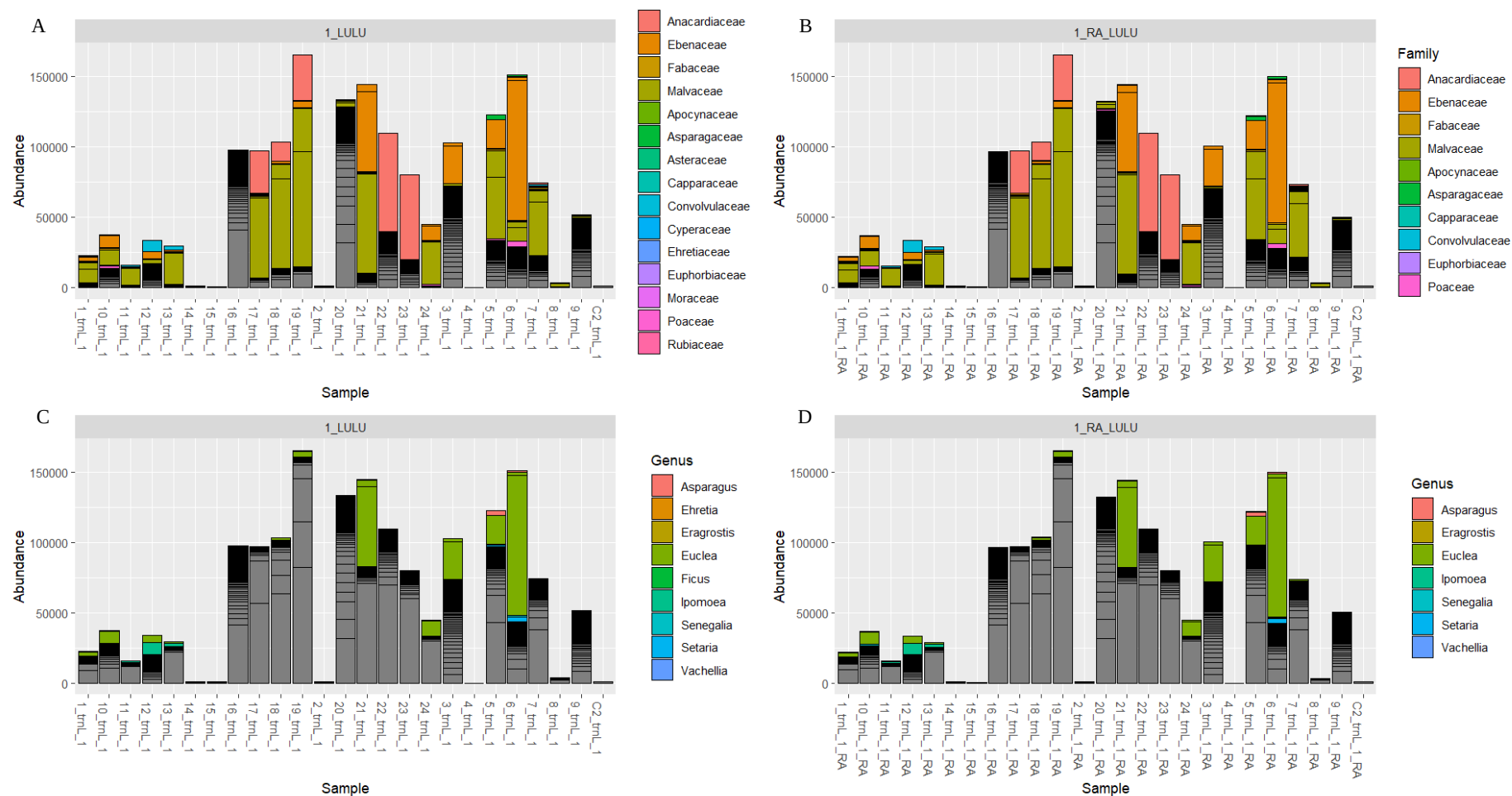


Figure 5.6: Bar plots generated from taxonomic assignments for the Welverdiend *trnL* dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).

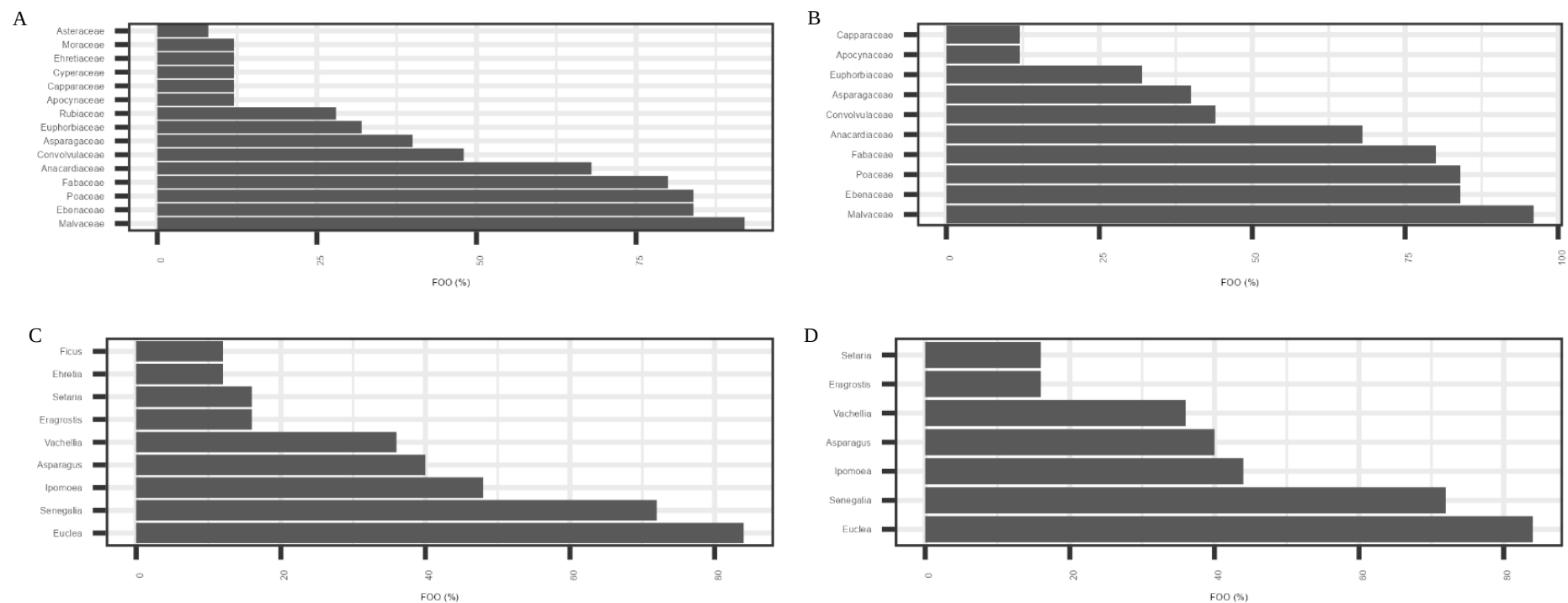


Figure 5.7: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Welverdiend *trnL* dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).

For the Syferkuil *rbcL* dataset, unassigned ASVs were much more prevalent than those that could be linked to a taxonomy. The highest read abundance was recorded for sample 26 with 71 365 reads for DADA_1_LULU (Fig. 5.8A and C) and 68 288 reads for DADA_1_RA_LULU (Fig. 5.8B and D). The lowest read abundance was seen for sample 43 with 4 711 reads for both DADA_1_LULU and DADA_1_RA_LULU. The same trend in species richness as seen for the Welverdiend datasets was evident for Syferkuil: the *rbcL* marker showed a greater success in the recovery of species from cattle faecal samples when filtered with DADA_1_RA_LULU and the *trnL* marker with DADA_1_LULU. This trend in species richness was not observed for families identified with the *rbcL* marker in Syferkuil samples since both approaches identified 9 families. This trend is however evident in the number of genera identified, which almost doubled from the application of DADA_1_LULU (6 genera) to DADA_1_RA_LULU (13 genera). Although the same number of families was identified by both bioinformatics approaches, they differed in composition. Common families identified were Apocynaceae, Ebenaceae, Fabaceae, Phyllanthaceae, Poaceae, Rhamnaceae and Solanaceae and differed in composition in that DADA_1_LULU instead identified Amaranthaceae and Moraceae and DADA_1_RA_LULU identified Caryophyllaceae and Geraniaceae. Of these, only Ebenaceae, Moraceae, Poaceae, Solanaceae and Fabaceae could be identified to genus level for DADA_1_LULU and all nine families could be identified to genus level for DADA_1_RA_LULU.

The FOO% graphs for this dataset's calculation were based on 24 samples since the control sample was lost during filtering. These graphs for the Syferkuil *rbcL* dataset (Fig. 5.9) revealed that DADA_1_LULU identified *Setaria* and *Urochloa* as genera being most frequent (79,2%) and *Ficus* and *Solanum* the least (8,33%). The most frequently occurring family identified by DADA_1_LULU was Poaceae (87,5%) and Moraceae the least (8,33%). DADA_1_RA_LULU identified *Tricholaena* as the most frequently occurring genus (79,2%), and the Fabaceae as the most frequent (75%) family. The genera occurring at lowest frequencies were *Berchemia*, *Monsonia* and *Solanum* with FOO percentages of 8,33% and the least abundant families were Geraniaceae and Solanaceae with FOO percentages of 8,33%.

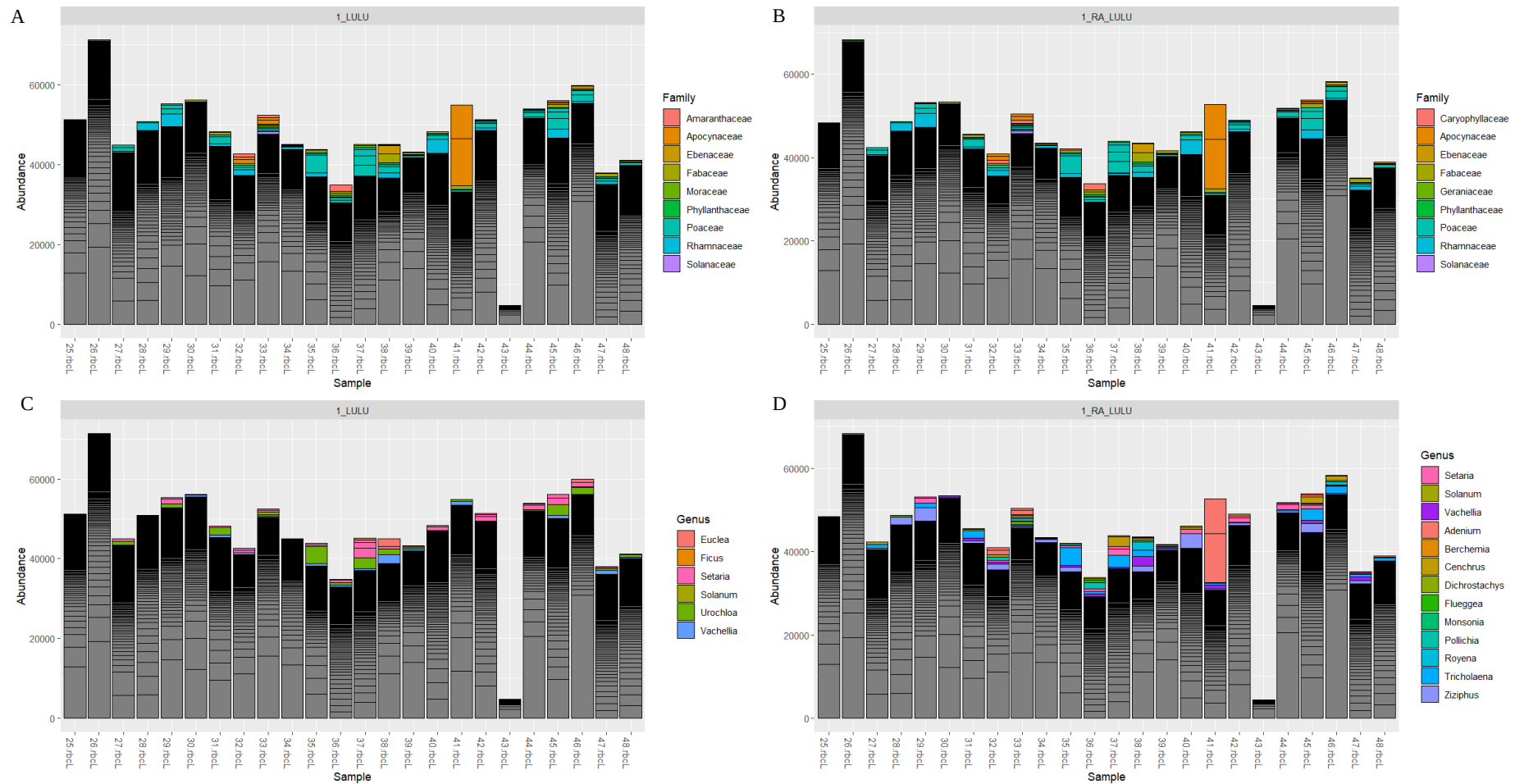


Figure 5.8: Bar plots generated from taxonomic assignments for the Syferkuil *rbcL* dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).

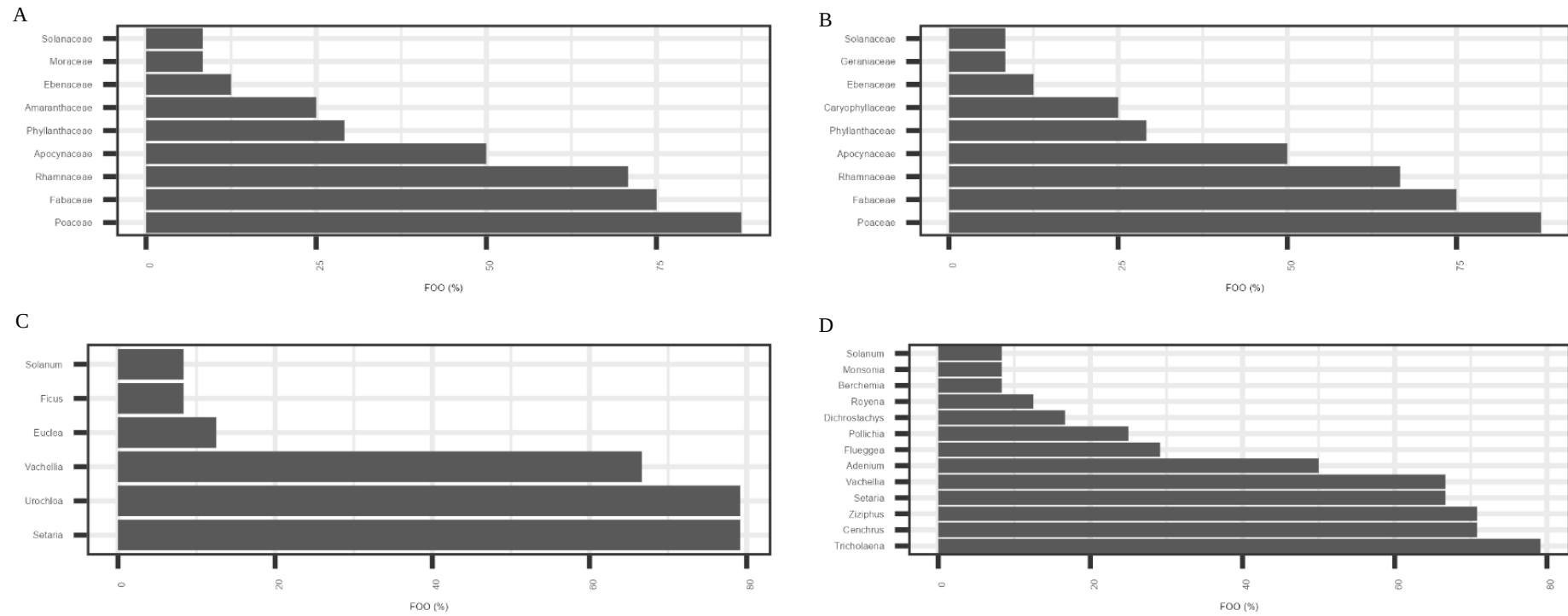


Figure 5.9: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Syferkuil *rbcL* dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).

From the bar plots (Fig. 5.10), it is evident that the Syferkuil *trnL* dataset included the highest abundances of unassigned reads relative to the other three datasets. NA reads were included to such a high degree that the number of ASVs assigned to species could not be visualised with the read abundance bar plots. The results reflect possible contamination since the control sample contained much more reads than seen for any other dataset and was reduced from 9124 seen in DADA_1_LULU (Fig. 5.10A and C) to 78 by DADA_1_RA_LULU (Fig. 5.10B and D). The highest read abundance identified by the DADA_1_LULU approach was 78 226 for sample 33 and the lowest for the control sample. DADA_1_RA_LULU identified sample 36 as having a read abundance of 34 962, and again the control sample had the lowest read abundance. Also, the trend in taxonomic richness, where DADA_1_RA_LULU performs best for *rbcL* datasets and DADA_1_LULU performs best for *trnL* datasets, extends to this dataset where 8 families and 10 genera were identified by DADA_1_LULU and 6 families and 7 genera were identified by DADA_1_RA_LULU. All families that have been identified by the DADA_1_RA_LULU workflow have also been identified by DADA_1_LULU with the addition of Lamiaceae and Euphorbiaceae. In addition to the genera already identified by the DADA_1_RA_LULU approach, DADA_1_LULU was also able to identify *Leucas*, *Pollichia*, *Senegalia* and *Senna*.

Like the *rbcL* dataset for Syferkuil where the control was lost, all reads from sample 43 were lost for the *trnL* dataset during *in-silico* filtering and FOO% was therefore calculated by using 24 samples (Fig. 5.11). The FOO% graphs showed that Poaceae was identified as the most frequently occurring family in the diets of Syferkuil cattle with DADA_1_LULU and DADA_1_RA_LULU indicating a FOO of 95,8% (Fig. 5.11A and B). Both approaches also identified the family Solanaceae as the least or one of the least frequently occurring families (8,33%), with the addition of Lamiaceae and Euphorbiaceae occurring at the same frequency with the DADA_1_LULU approach. The genus *Setaria* was identified as most frequent by both approaches with 87,5% for DADA_1_LULU and 79,2% for DADA_1_RA_LULU. Genera with the lowest FOO, according to DADA_1_LULU were *Senegalia*, *Senna*, *Setaria* and *Solanum* (8,33%) and for DADA_1_RA_LULU, *Solanum* (8,33%).

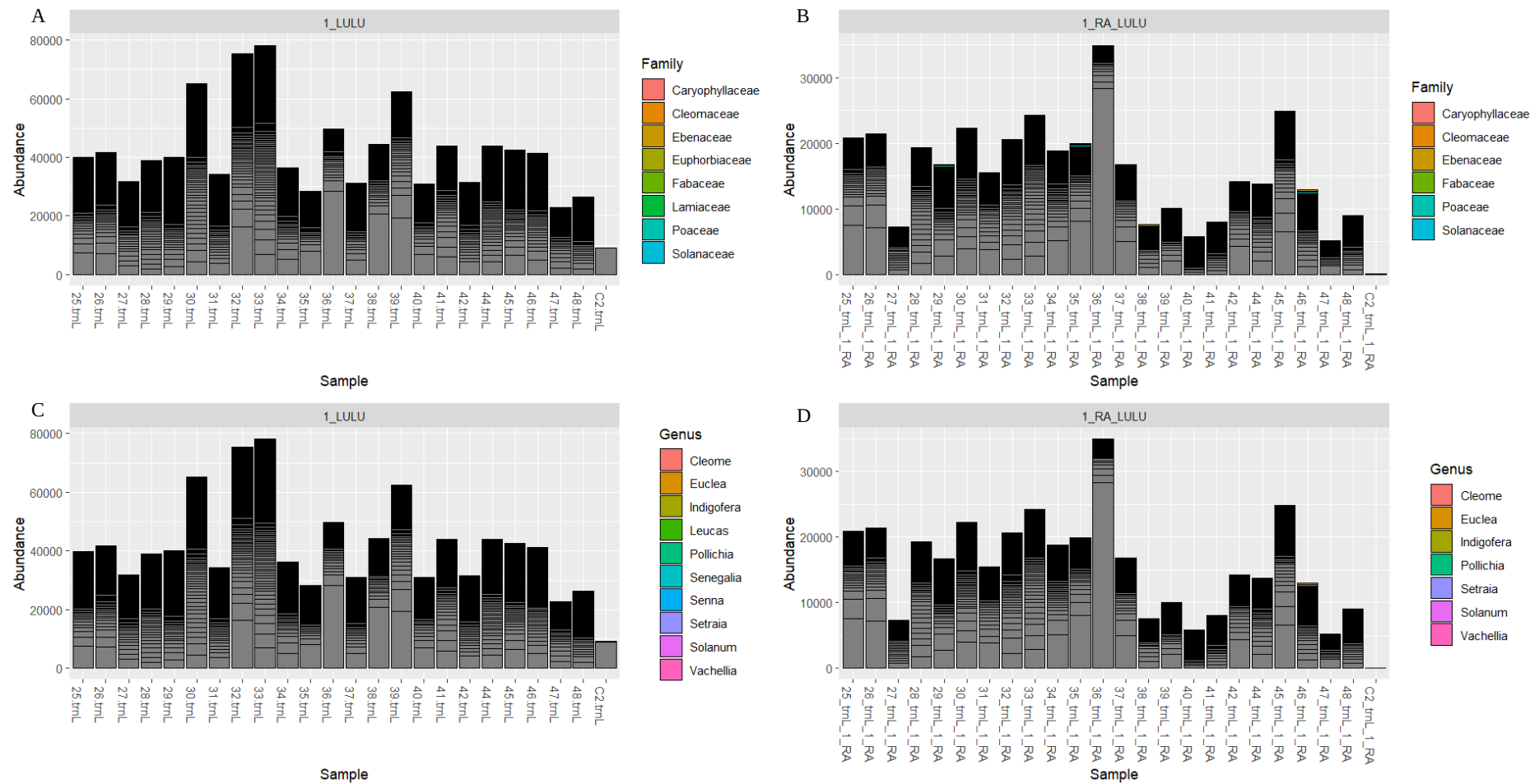


Figure 5.10: Bar plots generated from taxonomic assignments for the Syferkuil *trnL* dataset filtered by bioinformatics approaches identified as most optimal: DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and D) for the abundance of reads of the families (A and B) and genera (C and D) per sample. These plots also indicate the number of unassigned reads per sample (grey).

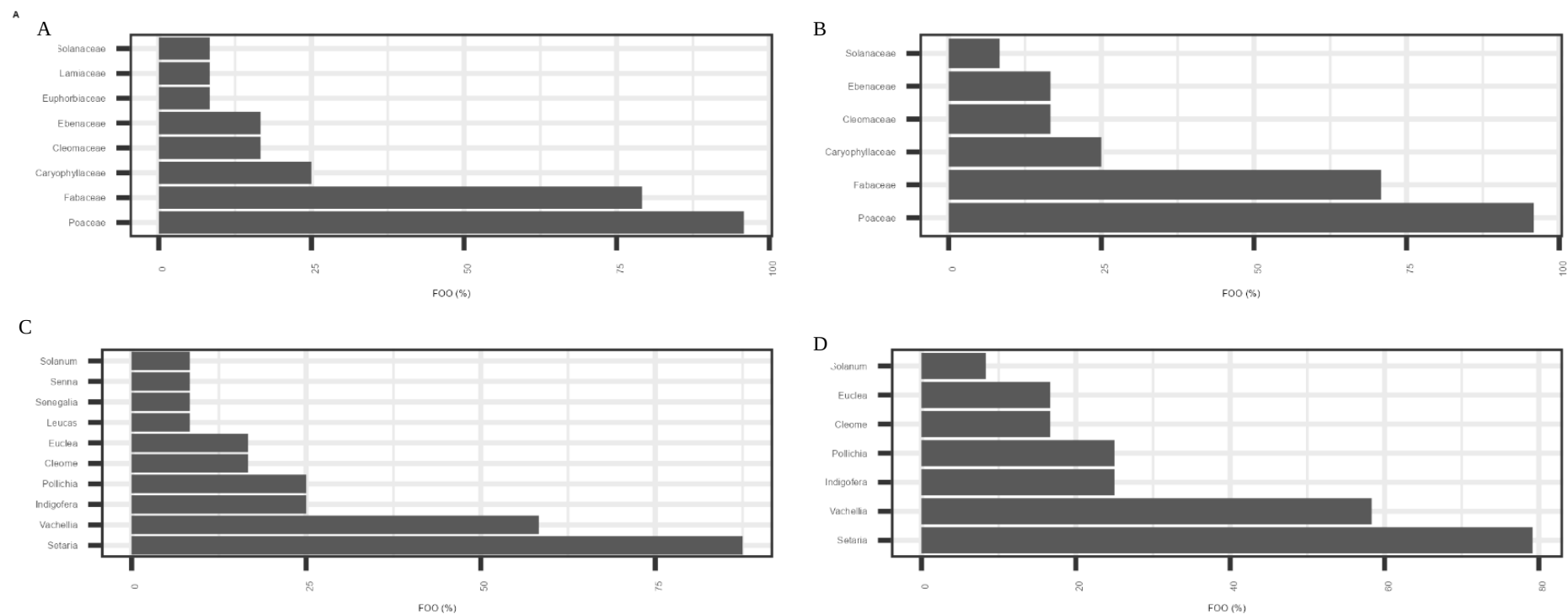
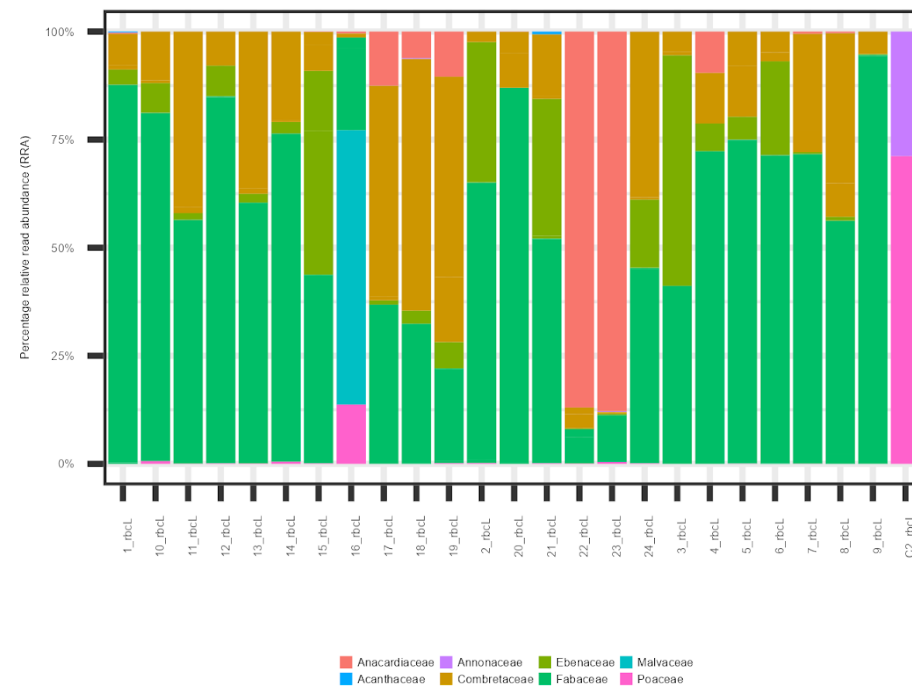


Figure 5.11: Percentage frequency of occurrence (FOO%) graph calculated from presence/absence data of taxonomies per sample for the Syferkuil *trnL* dataset filtered by the DADA_1_LULU (A and C) and DADA_1_RA_LULU (B and C) for the plant families (A and B) and genera (C and D).

The stacked bar plots, reflecting the relative read abundance of identified families per sample for the Welverdiend *rbcL* dataset (Fig. 5.12A), indicated that Fabaceae, Combretaceae and Ebenaceae were the main food items for most sampled individuals, except sample 16 where food items in the family Malvaceae dominated with greater inclusion of Poaceae than in other samples, and sample 22 and 23 where Anacardiaceae made up most of the diet. Although the identified families for DADA_1_LULU and DADA_1_RA_LULU are similar (Fig. 5.12A), DADA_1_RA_LULU could identify more genera (Fig. 5.12B). *Albizia*, *Combretum* and *Euclea*, which are shrubs or trees, were identified as important food items for Welverdiend cattle by the DADA_1_LULU approach. DADA_1_RA_LULU not only showed the relative importance of these three genera but also gave a more complete picture of what were specifically contained in samples 16, 22 and 23. Sample 16 represented the diet of an individual who ingested *Malvastrum* (forb/shrub) in a large amount, followed by *Dichrostachys* (shrub/tree), *Trachypogon* (grass) and lastly *Albizia* (shrub/tree). The genus *Rhus* (shrub with synonym *Searsia*), belonging to the family Anacardiaceae, was also identified as an important food item in samples 17, 18, 19, 4, and specifically samples 22 and 23, where *Rhus* represented more than 75% of the diet.

A



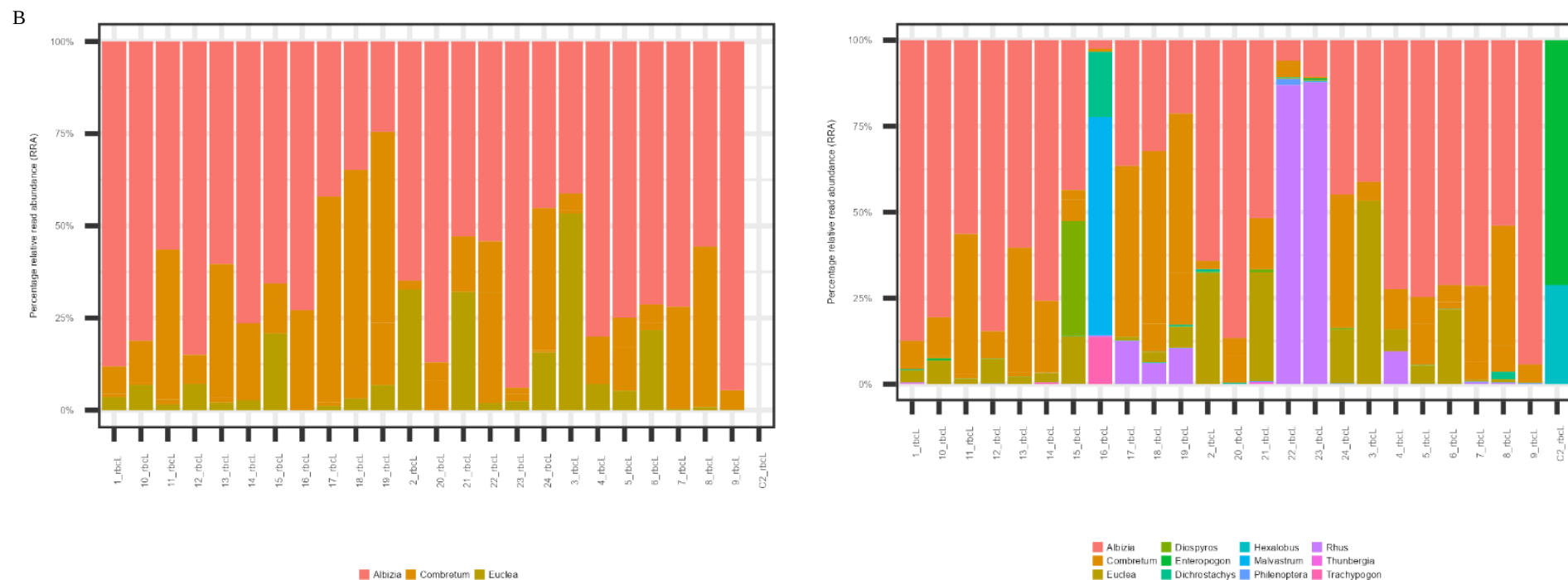
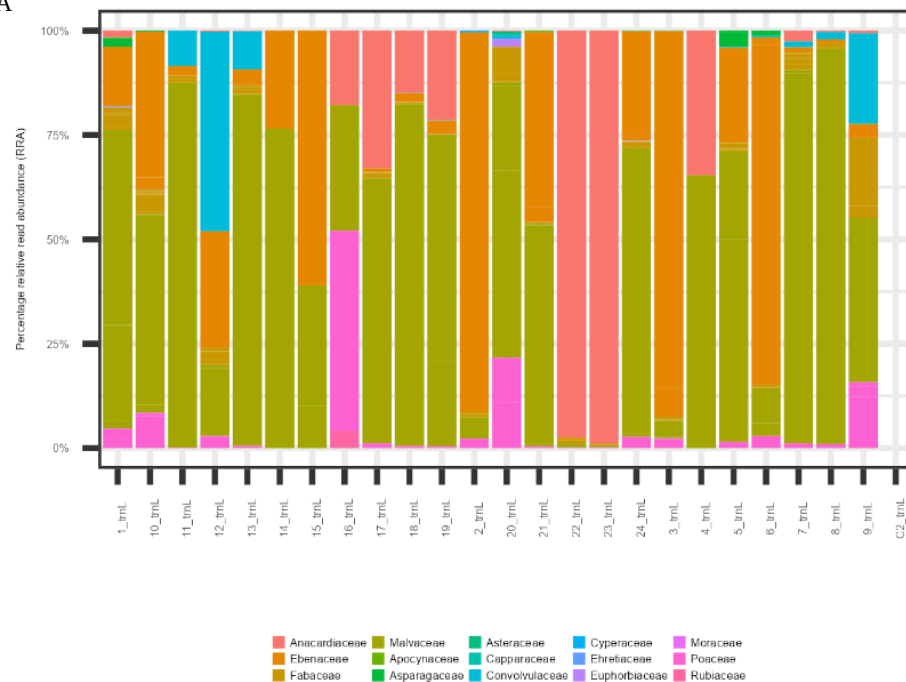


Figure 5.12: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Welverdiend cattle faecal samples by the *rbcL* barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.

The RRA calculations for each sample in the Welverdiend *trnL* dataset filtered with the DADA_1_LULU approach revealed that Malvaceae, Ebenaceae, Convolvulaceae, Poaceae and Acanthaceae are important food items. Asparagaceae, Rubiaceae, and Euphorbiaceae were also identified as important food items in smaller relative amounts (Fig. 5.13A). The same family proportions were seen for DADA_1_RA_LULU except for the exclusion of Rubiaceae in sample 16 which could not be identified by this bioinformatics approach. In the case of genera identified for Welverdiend *trnL* dataset (Fig. 5.13B), samples 4 and 16 did not have any data since the families identified for these samples could not be described to genus level by either bioinformatics approach. Genera contributing to the diets of Welverdiend cattle, identified by DADA_1_LULU and DADA_1_RA_LULU was *Euclea* (shrub/tree) in the greatest amount, followed by *Ipomoea* (forb), *Vachellia* (shrub/tree), *Senegalia* (shrub/tree), *Asparagus* (forb) and *Eragrostis* (grass). DADA_1_LULU also identified *Ficus* (shrub/tree) as a food item in samples 18 and 23.

A



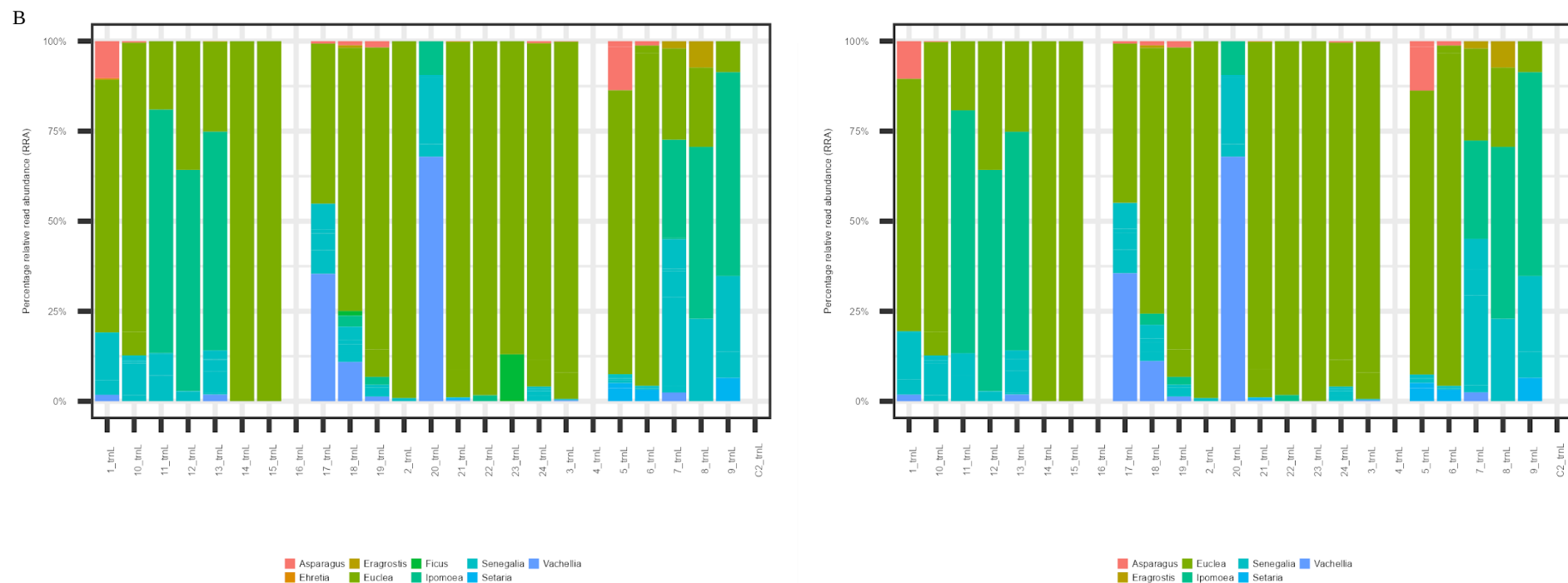
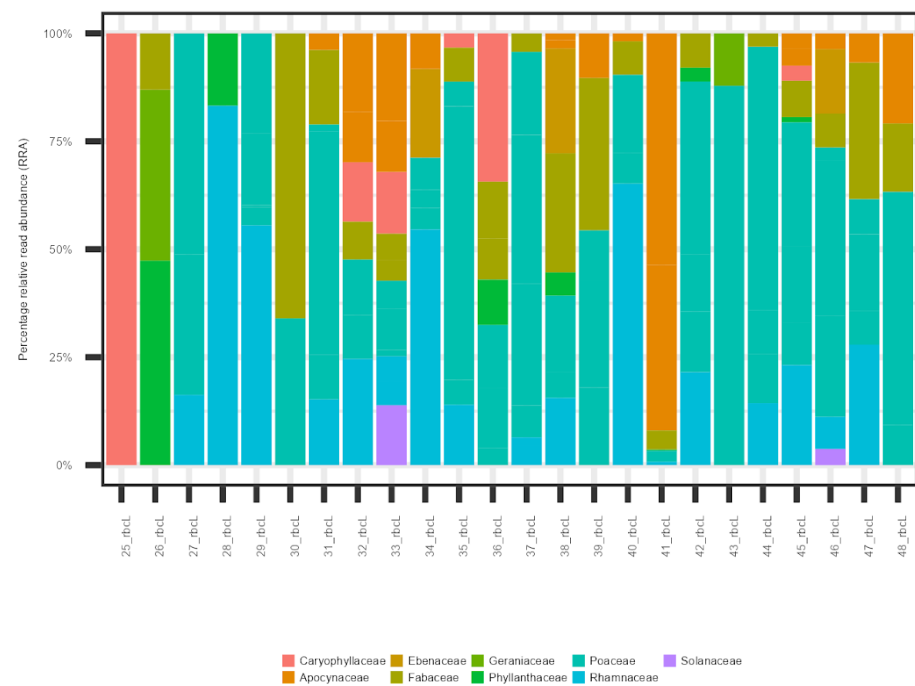
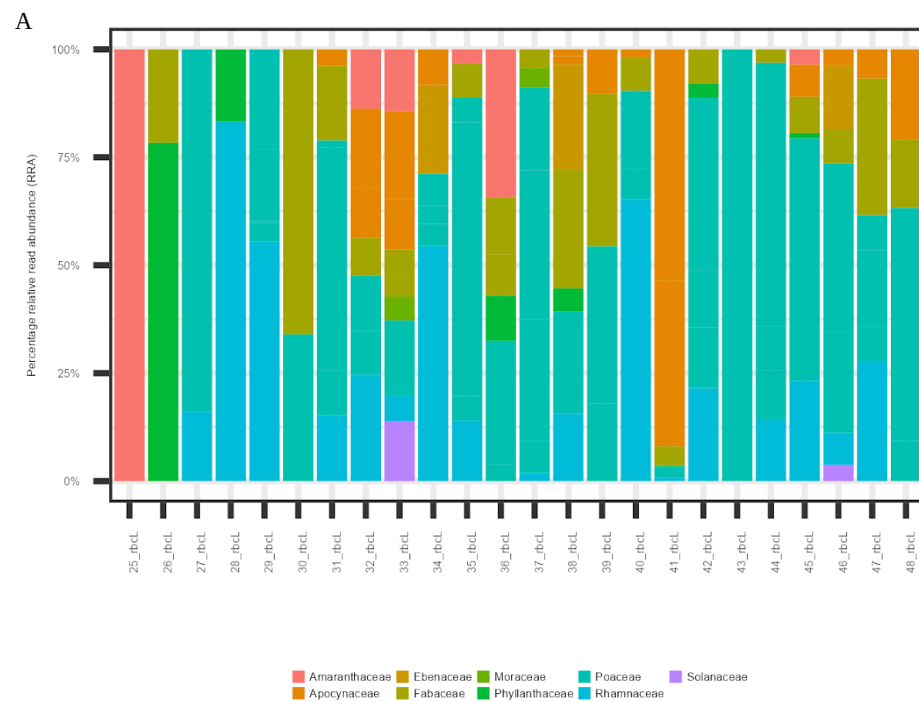


Figure 5.13: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Welverdiend cattle faecal samples by the *trnL* barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.

For cattle foraging at the *Draught Act* section of Syferkuil, the *rbcL* barcoding marker overall identified Poaceae as the most important family, followed by Fabaceae, Amaranthaceae, especially in sample 25 where the latter made up the entire diet, Phyllanthaceae, especially sample 26 where it represents more than 75% of the diet, Apocynaceae, especially in sample 41 where it represented more than 75% of the individuals' diet, Rhamnaceae and Ebenaceae with the addition of Geraniaceae with the DADA_1_LULU approach (Fig. 5.14A). A small amount of the family Solanaceae was also detected in samples 33 and 46- with both bioinformatics approaches. For DADA_1_LULU no genus data was available for samples 25 and 28 since this approach could not identify genera for the families of Amaranthaceae (sample 25) or Rhamnaceae and Phyllanthaceae (sample 28) to genus level. DADA_1_RA_LULU could differentiate between different families on a finer level than DADA_1_LULU, where not only *Setaria* (grass), *Urochloa* (grass), *Euclea* (shrub/tree) and *Vachellia* (shrub/tree) are identified as important contributing food items, but also *Pollichia* (shrub), *Ziziphus* (shrub/tree), *Flueggea* (shrub), *Tricholaena* (grass) (Fig. 5.14B).



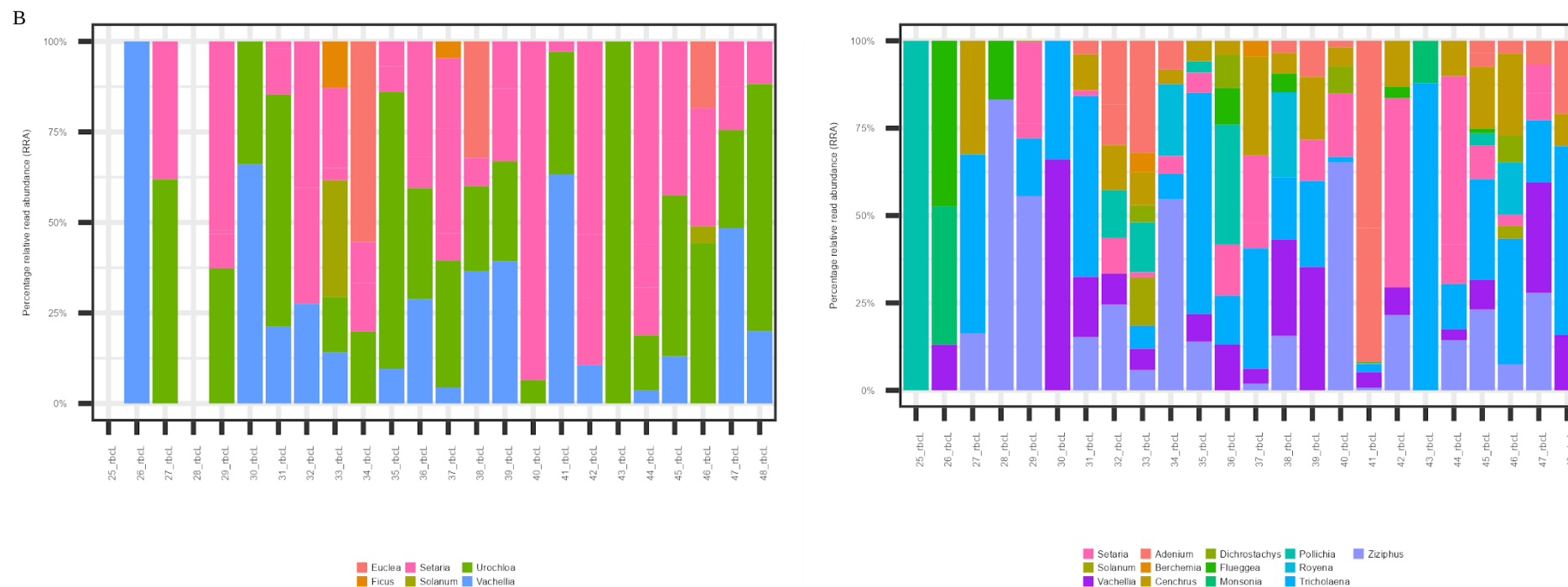
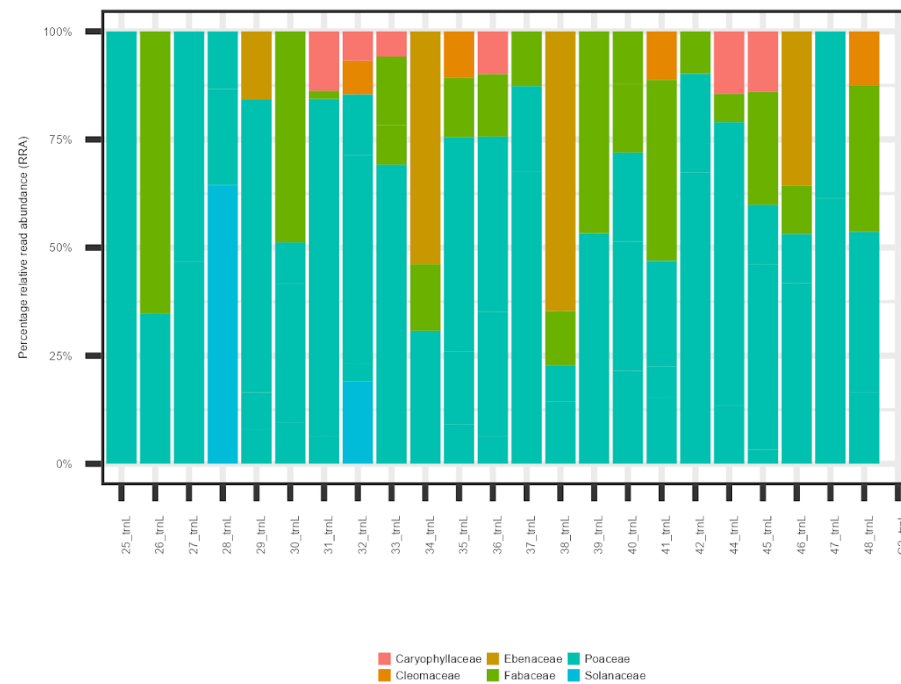
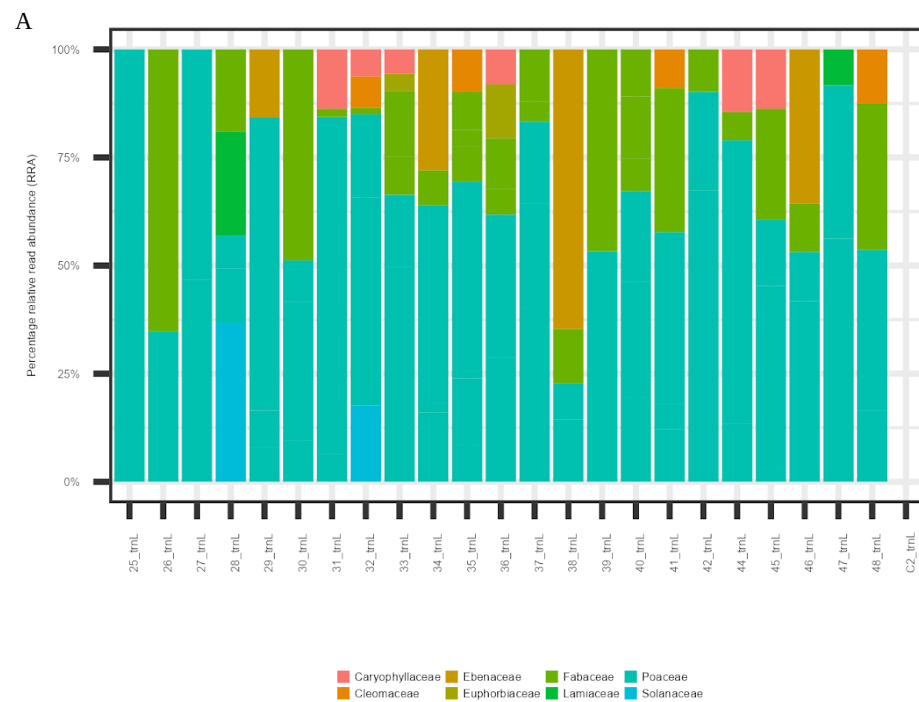


Figure 5.14: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Syferkuil cattle faecal samples by the *rbcl* barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.

Poaceae and Fabaceae were identified as the most important food items for the Syferkuil *trnL* dataset (Fig. 5.15A). Fabaceae, Ebenaceae and Caryophyllaceae were also relatively important food items for samples 31, 32, 33, 36, 44 and 46. Solanaceae was identified as an important food item in samples 28 and 32 and Cleomaceae for samples 32, 35, 41 and 48. *Setaria* (grass) and *Vachellia* (shrub/tree) were identified as dominant genera in Syferkuil cattle diets barcoded with the *trnL* barcoding marker and *Euclea* (shrub/tree), *Pollichia* (forb), *Cleome* (forb) and *Indigofera* (forb/shrub) in lower abundances (Fig. 5.15B). The DADA_1_LULU approach additionally identified a rather large proportion of *Leucas* (forb/shrub) in samples 28 and 47 and *Senna* (shrub/tree) in smaller proportions in samples 28 and 37 (Fig. 5.15B).



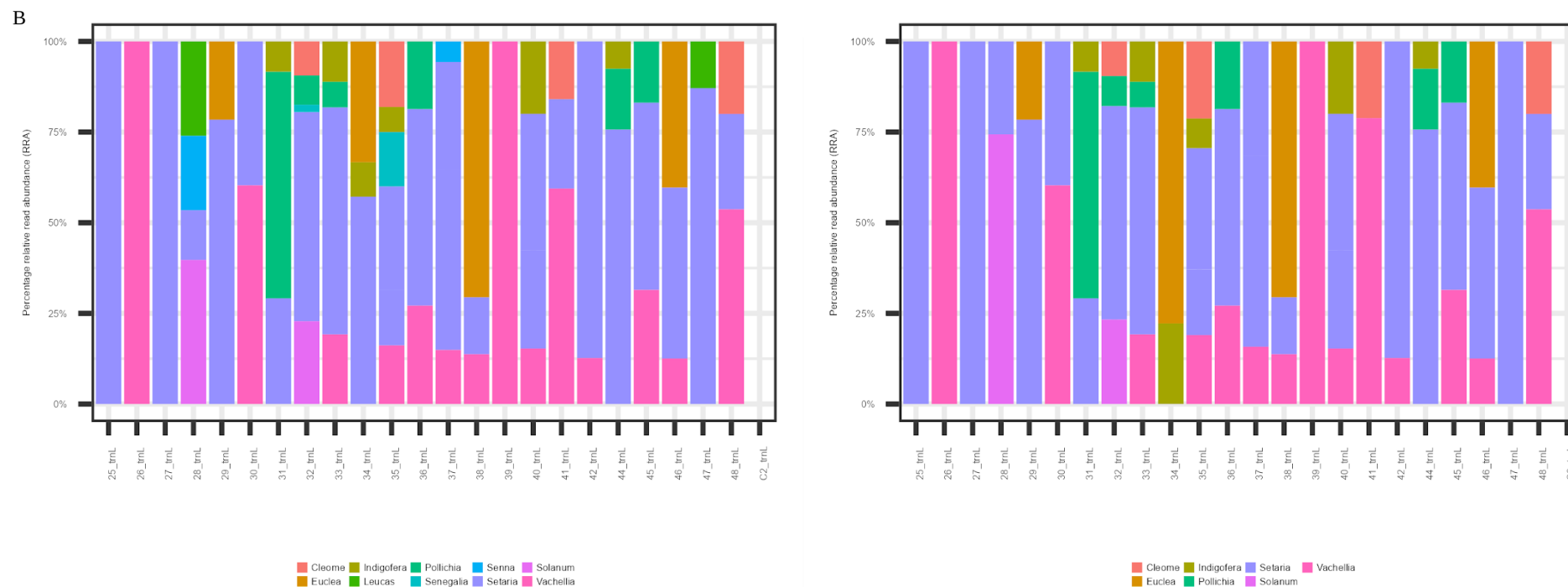


Figure 5.15: Relative read abundance graphs representing the relative read abundance for families (A) and genera (B) identified from Syferkuil cattle faecal samples by the *trnL* barcoding marker by the i) DADA_1_LULU and the ii) DADA_1_RA_LULU approach.

Discussion

Exploring the optimal foraging theory, which holds that the primary objective of a generalist herbivore is to maximise the amount or quality of food when foraging, or to collect food effectively while avoiding predation, requires an in-depth understanding of diet breadth and preferred food items (MacArthur & Pianka, 1983). The quantification of the diets of herbivores foraging in dry savanna ecosystems will therefore highlight which food items are, firstly, available in these ecosystems, and secondly, how important they are as forage to these herbivores. The dietary analysis of cattle foraging in eastern semi-arid South African savannas will therefore allow i) the identification of foraged species in their diets, and ii) the quantification of each of these food items by assessing the diet composition in terms of frequency of occurrence (FOO) and relative read abundance (RRA) per foraged species. These aims were realised by applying an optimal bioinformatics approach that will filter the Illumina sequence data for the *rbcL* and *trnL*(UAA) barcoding markers from cattle faecal samples. This bioinformatics approach has been identified by the application of the mock community (Chapter 4) as the DADA2 bioinformatics pipeline applied with a frequency-based filter and/or an abundance-based filter together with post-clustering curation by the LULU algorithm (DADA_1_LULU and DADA_1_RA_LULU).

Raw sequence data quality

The quality reports generated by MULTIQC from FASTQC reports did not raise any concerns for the Welverdiend dataset. The Syferkuil dataset, however, showed that sample 43 in the *trnL* dataset and the control sample in the *rbcL* dataset could cause some problems in downstream analysis by having very poor quality, in terms of Phred scores, a high inclusion of ambiguous bases and very short average sequence lengths. This low quality resulted in these two samples having much lower reads retained after filtering in the DADA2 workflow, where reads are specifically discarded that contain ambiguous bases as well as the truncation of reads at first instances of low-quality scores. After the completion of the DADA2 workflow, these samples were still present but did not pass the first post-clustering quality filter since none of the ASVs contained in them occurred more than once per sample. The relatively higher quality of the Welverdiend dataset compared to the Syferkuil dataset can also be seen in the number of reads retained after *in-silico* filters and the resulting number of unassigned reads during taxonomic assignment. Read retainment after the application of *in-silico* filters was relatively high for Syferkuil, with the *rbcL* dataset having 33% more reads than Welverdiend after DADA_1_LULU filtering and 69% after DADA_1_RA_LULU filtering. The *trnL* dataset of Syferkuil contained 63% more ASVs than Welverdiend after DADA_1_LULU filtering and 58% after DADA_1_RA_filtering. Despite the high number of ASV retainment after the application of *in-silico* filters, the number of unassigned reads relative to reads that could be assigned to their taxonomy was staggeringly high and were so high for the *trnL* dataset that proportions of assigned versus unassigned could not be visualised. This means that either the unassigned

reads were not present in the reference library, or more likely that these ASVs are present as a result of contamination that could not be filtered out by DADA2 and occurred in such large abundances that they could not be filtered out by *in-silico* filters.

Removal of suspected contamination

By extracting the unassigned ASVs retrieved for each dataset after the comparison of unknown queries to the compiled reference database, it was possible to determine the robustness of this reference database as well as deducing certain quality information, especially for the Syferkuil *trnL* dataset. The compiled reference database of *rbcL* and *trnL* primers for species occurring in an eastern semi-arid South African savanna showed sufficient taxonomic coverage to assign the taxonomies of unknown food items in herbivore diets across the two study sites. The robustness of this database is confirmed by the fact that none of the unassigned reads could be taxonomically assigned to plant species by a large public database of reference sequences. It was evident that non-specific binding to bacterial sequences took place, especially in the case of the *trnL* primers (*trnLc/d*) in the Syferkuil dataset, indicating that it is not as specific to plant taxa as the *rbcL* primers (*rbcLaF/R*). The bacterial assignment of certain sequences could also occur due to the prevalence of chimeric sequences, where the amplicon starts off with the correct PCR primer, followed by random DNA templates. The sequences assigned to bacteria could therefore be artifacts left by the sequencing platform. This does not explain, however, why this same prevalence of bacterial assignments was not seen for the Welverdiend *trnL* dataset, especially after the removal of chimeric sequences by commands built into the DADA2 workflow. Another explanation could be that the PCR amplification of barcoded sequences and its subsequent sequencing did not have the same quality across datasets and that some form of contamination or sequencing error is to blame for the large number of unassigned reads for the Syferkuil *trnL* dataset, as well as the large number of unassigned reads that could be taxonomically assigned to taxa other than plants by the NCBI “nt” database.

Dietary richness quantification

After the processing of raw data and trimming and filtering and the inference of sequence variance by the denoising algorithm of DADA2, as well as the removal of unassigned ASVs suspected to be a result of contamination, taxonomic assignment of the query sequence, or ASVs took place against *rbcL* and *trnL* reference sequences as described in Chapter 3. A comparison of the taxonomy retrieved by each bioinformatics approach per filtering method and barcoding marker revealed that a higher species recovery was evident for the *rbcL* datasets of Welverdiend and Syferkuil with the DADA_1_RA_LULU approach and the DADA_1_LULU approach for the *trnL* datasets. DADA_1_RA_LULU was able to identify 25% more families and 60% more genera from the Welverdiend *rbcL* dataset. The taxonomic resolution was also increased with the use of the DADA_1_RA_LULU approach whereby each family is represented by one or more genera.

This trend continues in the Syferkuil *rbcL* dataset. Each bioinformatics approach was able to identify the same number of families, albeit different families, however, a greater success by DADA_1_RA_LULU was shown to describe these identified families to genus level where 46% more genera were identified by the DADA_1_RA_LULU approach than the DADA_1_LULU approach and all nine identified families could be identified to genus level. It is therefore clear that DADA_1_RA_LULU, where a frequency and abundance filter is applied to ASVs formed by DADA2 and post-clustering curation took place, is applicable for DNA barcoded with the *rbcLa* region to retrieve dietary descriptions with the highest species richness and taxonomic resolution. The opposite trend is observed for the *trnL* datasets of Welverdiend and Syferkuil, where the less stringent filtering approach applied by DADA_1_LULU showed a greater ability to recover the species richness of these cattle diets. The DADA_1_LULU approach was able to identify 33% more families for the Welverdiend dataset than DADA_1_RA_LULU and 22% more genera, however, neither bioinformatics approach was able to identify all identified families to genus level. DADA_1_LULU could identify 25% more families for the Syferkuil *trnL* dataset than DADA_1_RA_LULU and 30% more genera. For this dataset, both bioinformatics approaches were able to describe all the identified families to the genus level. DADA_1_LULU is therefore optimal for sequences of the *trnL* barcoding marker where only a frequency-based filter is applied to ASVs formed by DADA2 and subsequently post-clustering curation. Accordingly, all further discussions of taxonomies retrieved with the *rbcL* and *trnL* barcoding marker to describe the diets of cattle from Welverdiend and Syferkuil will refer to diet quantification by the DADA_1_RA_LULU approach for *rbcL* datasets and the DADA_1_LULU approach for the *trnL* dataset according to the criterion set out for the choice of bioinformatics filtering approach: the optimal bioinformatics approach will be chosen per barcoding marker that retrieved the highest taxonomic richness in terms of plant families and genera.

These taxonomic assignments can be visualised with or without the inclusion of unassigned reads, and the choice of metrics used for the visualisation of taxonomies will depend on the research question, which in this was case establishing the presence of food items and then determining the relative proportions of these food items in the diets of cattle foraging in the ecological systems of Welverdiend and Syferkuil. Diet quantification will therefore be visualised using two distinct metrics and used in the same way to describe a herbivore diet as recommended by Kartzinel *et al.* (2015) and Deagle *et al.* (2019). This entails conservatively using sequence frequency of occurrence (FOO) averaged across all samples to describe species occurrence data and using sequence relative read abundance as a secondary metric to visualise finer differences across datasets and individuals since RRA does not always reflect the relative intake of plant species (Deagle *et al.*, 2019). RRA will also be used to establish the importance of less frequently present food items in the sampled dataset to some sampled individuals. This is an important visualisation since the importance of forbs relative to more frequently occurring species such as grasses will be lost in descriptions based only on the frequency of occurrence; forbs may play an important role in the cattle diets and make

up a large proportion of an individual's diet even though it may not be the most frequently occurring food item.

The taxonomies described using the different barcoding regions highlight the need for the use of a dual-locus identification system for plants (CBOL Plant Working group, 2009; Hollingsworth et al., 2011). The tendency of barcoding loci to fail or faintly amplify specific taxonomic groups (CBOL Plant Working Group, 2009; Zhu *et al.*, 2022) allows for the systematic underrepresentation of inefficiently amplified species and the omission of food items. Using multiple loci to recreate diet profiles allows for mutual verification, which helps reveal such disparities (Camp *et al.*, 2020). The use of the *rbcL* and *trnL* chloroplast genes resulted in complementing taxonomic identifications, facilitating the reconstruction of herbivore diets from an eastern semi-arid South-African savanna (Figures 5.16 and 5.17). Frequently consumed food items in the diets of Welverdiend cattle, according to the *rbcL* and *trnL* marker, were from the Fabaceae, Combretaceae, Malvaceae, Ebenaceae and Anacardiaceae. The family Fabaceae, Poaceae and Ebenaceae also occurred in relatively greater proportions in the diets of these cattle according to the RRA graphs except for Malvaceae in the *rbcL* dataset, where it made up ~63% of the diet of one individual (Fig. 5.12A) and Combretaceae, which could not be identified by the *trnL* barcode. A dissimilarity in the recorded plant diversity identified by barcoding loci could be due to added information due to differing taxonomic breadths (Fahner et al., 2016), which may explain why Combretaceae was missing from identifications with the *trnL* dataset. This family was represented by 12 unique species in the combined reference database, of which only three are barcoded with *trnL* and only consisted of one reference sequence each. The *rbcL* barcode, however, was represented by all 12 species, of which nine were unique to this dataset, with two or more reference sequences each. Interestingly, although Malvaceae was one of the most frequently occurring families in the Welverdiend *trnL* dataset, it could not be identified to genus level, but it was in the *rbcL* dataset in the form of *Malvastrum* where it accounted for ~10% of the FOO and ~63% RRA in sample 16. Importantly Annonaceae has a FOO of 24%, but only occurred in a relative proportion of the control sample of *rbcL* with an RRA of 25% as well as Poaceae with an RRA of 75%. These families were identified to genus level in the same RRA's as *Hexalobus* and *Enteropogon*. The presence of identifiable reads in the control sample should not be possible since it does not stem from a faecal sample but is a mixture of PCR reagents and primers (*rbcLaF/rbcLaR*) added in the same quantities as for the faecal samples. The control sample of Welverdiend, therefore, contained contamination that entered the workflow in a laboratory environment during the preparation of PCR mixtures. Frequently consumed food items for Syferkuil cattle, according to the *rbcL* and *trnL* marker, were relatively similar to those of Welverdiend with Poaceae and Fabaceae showing the highest FOO%. Rhamnaceae and Apocynaceae were also identified as important dietary items by the *rbcL* marker.

Except for the differences in families identified, the *rbcL* and *trnL* barcoding markers have different affinities for amplifying certain genera from a common family identified in the diets of cattle foraging in

each study site. The taxa under consideration, as well as the marker used, have a significant impact on the success of species identification using DNA metabarcoding (De Vere *et al.*, 2012; Amandita *et al.*, 2019; Mallot *et al.*, 2018) and dissimilarities in the identified taxa may increase at finer scales to result in marker disagreement at genus level plant diversity (Yoccoz *et al.*, 2012; Fahner *et al.*, 2016). It is well known that neither *rbcL* or *trnL* have particularly high rates of success in the discrimination of genus or species level taxonomic levels (Valentini *et al.*, 2009; Taberlet *et al.*, 2007; Hollingsworth *et al.*, 2009; Newmaster *et al.*, 2006; Chase *et al.*, 2005; Quéméré *et al.*, 2013), but the observation of an overlapping identification between these markers can provide increased support for the findings (Kress & Erickson, 2007; Fahner *et al.*, 2016). In the Welverdiend study site, the family Fabaceae was represented by *Albizia*, *Dichrostachys* and *Philenoptera* as identified by the *rbcL* barcoding marker and *Senegalia* and *Vachellia* as identified by the *trnL* barcoding marker. For Syferkuil this same family was represented by *Dichrostachys* and *Vachellia* by the *rbcL* marker and *Senegalia*, *Vachellia*, *Indigofera* and *Senna* by the *trnL* marker. The *rbcL* barcoding marker therefore exclusively identified the genera *Dichrostachys* across both datasets and the *trnL* marker exclusively identified *Senegalia*.

According to the comparison of loci as recommended by the CBOL Plant Working Group (2009) or the “overlap rule”, which could be used to reliably determine which genera are present in these study sites according to taxa identified by both markers, no genus level assignments in the Welverdiend dataset could be made and only the identification of *Vachellia* in the Syferkuil dataset. Although this result would have more support, identifications that do not overlap should not be neglected since all species present in the diets of herbivores, especially in an eastern semi-arid South Africa savanna, have not been barcoded with both the chosen plant barcoding markers. The results should therefore be used in a supplementary manner to create a complete picture of the floristic diversity rather than excluding single-locus identifications until the availability of more complete reference sequence databases (Macgillivray & Kaczmarek, 2011). Even if the barcodes are mined to represent equal proportions of the potential dietary items the inherent flaw in sequence mining, i.e., the unknown nature of biases presented in sampling, DNA extraction protocols and sequencing in the originating laboratories (Salter *et al.*, 2014), introduces a source of variation between reference sequences that may lead to the identification of false positives and skewed representations in the final taxonomic assignments. However, the use of a complete, geographically constrained reference database where barcoded species are mined from various sources and represented equally among barcoding markers would already be sufficient for species identification and decrease false positive detection rates based on overlaps between barcoding loci and would certainly lead to more strongly supported deductions on real plant diversity (Fahner *et al.*, 2016).

The importance of forbs

Results indicated that forb genera are frequently identified within the sampled dataset, as well as being foraged by individuals of both Welverdiend and Syferkuil. Genera representing forb species that were identified for the Welverdiend cattle are *Malvastrum* and *Thunbergia* as identified by the *rbcL* marker and *Asparagus*, and *Ipomoea* identified with *trnL*. These genera (*Malvastrum* and *Thunbergia*) are not present in the *trnL* reference dataset and did not appear frequently across the sampled dataset with values of 8%. Forb genera were identified at a much higher frequency in the *trnL* dataset with values of 40% for *Asparagus* and 48% for *Ipomoea*. The *trnL* barcoding marker therefore had the ability to allow for the identification of the genera *Asparagus* and *Ipomoea* even though there were reference sequences available in the *rbcL* barcoding database. For the Syferkuil cattle the forb genera *Pollichia* (FOO = 25%), *Adenium* (FOO = 50%), *Solanum* (FOO = 8,33%), and *Monsonia* (FOO 8,33%) were included in their diets as identified by the *rbcL* marker, as well as *Solanum* (FOO = 8,33), *Cleome* (FOO = 16,7%), *Indigofera* (FOO = 25%), *Leucas* (FOO = 8,33) and *Pollichia* (FOO = 25%) identified by the *trnL* marker. Only the genera of *Pollichia* and *Solanum* could be identified by both the barcoding markers, again indicating the complementary nature of this dual-locus identification system that aids in the determination of dietary profiles when reference sequences are not available for all barcodes (Taberlet *et al.*, 2007) or when a barcode shows preferential amplification of certain plant taxa (Kartzinel *et al.*, 2015; Mallot *et al.*, 2018; Zhu *et al.*, 2022). Forbs are therefore foraged by a noticeable percentage of individuals in the sampled dataset, especially *Asparagus* and *Ipomoea* at Welverdiend and *Adenium*, *Pollichia* and *Indigofera* at Syferkuil. These two ecosystems, therefore, most probably had different compositions of forbs communities and diversity according to the RRA and FOO visualisations. Welverdiend had the highest frequency of occurrence of forbs species across the sampled dataset (i.e., higher forb species richness in the diet), while Syferkuil results illustrated a greater relative abundance of the forbs identified in individual diets (i.e., lower richness, but higher abundance per species). This means that although the forbs occurring at Welverdiend were foraged upon by many members of the sampled dataset, they have not foraged heavily and their diets still included high abundances of other life forms, while at Syferkuil, the forb that an individual managed to come across was heavily grazed, in a higher degree than plants from other life forms.

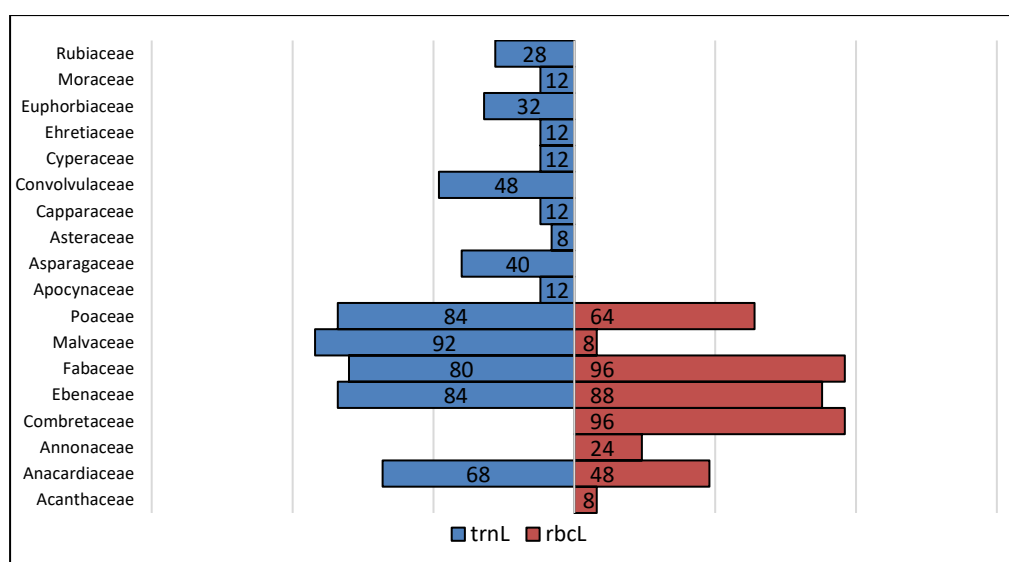


Figure 5.16: Mirrored bar chart showing the frequency of occurrence (FOO%) of plant families identified as part of the diets of Welverdiend cattle. Orange bars represent families identified barcoded with *rbcL* and blue bars represent those barcoded with *trnL*.

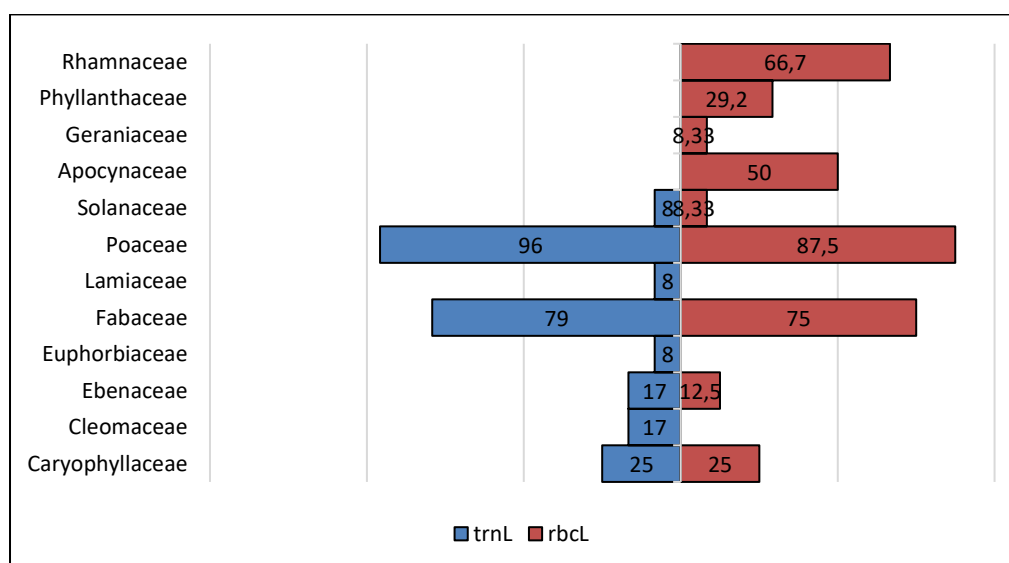


Figure 5.17: Mirrored bar chart showing the frequency of occurrence (FOO%) of plant families identified as part of the diets of Syferkuil cattle. Orange bars represent families identified barcoded with *rbcL* and blue bars represent those barcoded with *trnL*.

Conclusion

The comparison of bioinformatics approaches that performed best in the mock community analysis revealed some important considerations when applying these post-clustering filters and feature table curation to raw Illumina data sequenced from herbivorous faecal samples. The application of an abundance filter to remove

amplicon sequence variants (ASVs) that occur less than 0.005% across the dataset is suited to the bioinformatics analysis of sequences barcoded with the *rbcLa* chloroplast gene region. The application of this filter resulted in a higher taxonomic resolution at both family and genus levels to describe the diets of cattle foraging at Welverdiend and Syferkuil. The application of this abundance filter, however, is not suited to *trnL*(UAA) barcodes since taxonomic resolution greatly decreases with the removal of low abundance reads. Too stringent filtering parameters applied to the *trnL* barcoding marker, therefore, lead to the loss of biodiversity. The decision to use different bioinformatics approaches per barcoding marker will satisfy the research question, in that this overall study first aims to identify all possible food items from the cattle diets. The robustness of the reference sequence database is also confirmed by factually none of the unidentified reads per dataset could be assigned to plant species using a public reference sequence database. The food items occurring in the diets of herbivores foraging in an eastern semi-arid South African savanna could therefore be recovered and taxonomically assigned by the application of the correct bioinformatics workflow as well as the use of a study-area specific reference database. The second aim of the overall study was to determine the contribution of these food items to the cattle diets which was revealed to be the families Fabaceae, Combretaceae, Ebenaceae, and Malvaceae with genera *Albizia*, *Combretum*, and *Euclea* for the Welverdiend study site. Food items that were frequently foraged upon at Syferkuil included the families Poaceae and Fabaceae with genera *Setaria*, *Urochloa* and *Vachellia*. The metrics used for diet quantification, RRA and FOO, also allowed for the determination of the dietary contribution of forbs. Forbs belonging to the genera *Malvastrum*, *Asparagus*, and *Ipomoea* are important food items for cattle grazing in the Welverdiend area, and forbs belonging to the genera *Adenium* and *Pollichia* were important to the nutrition of cattle from Syferkuil. Furthermore, Syferkuil cattle did not seem to visit forbs often, but when an individual encountered a forb the intake of this species was relatively high, while at Welverdiend, more members of the sampled community foraged on specific forbs, but the relative intake thereof wasn't high. The study further highlights the success of a dual-locus barcoding approach for the identification of plants in that the *rbcL* marker supplemented missing taxonomies for *trnL* and vice versa. The use of two plant barcoding regions is, therefore, necessary to retrieve a more accurate taxonomic resolution to reconstruct the diets of herbivores.

Chapter 6 Overall Discussion

This dietary metabarcoding study requires a comprehensive study-area-specific reference database (Lamb *et al.*, 2016) of universal plant barcoding regions (Hollingsworth *et al.*, 2011) that can identify foraging targets present in the diets of a herbivore from an eastern semi-arid South African savanna that is fragmented by digestion and degradation (~150 bp) (Bhargava & Sharma, 2013). This was fulfilled by the compilation and curation of a *rbcL* and *trnL*(UAA) reference library containing both entries of species occurring in the Lowveld and Central Bushveld Bioregions represented by study sites used for faecal sample collection (Wolverdind and Syferkuil) as well as sequences generated from plant material collected from sites Wolverdind, Nkuhlu and Timbavati. This reference database, consisting of the *rbcL* and *trnL* databases exhibits a sufficient barcode gap at species level (76% for the *rbcL* dataset and 68% for the *trnL* dataset) and scored well in other distance-based criteria with high identification success rates for the taxonomic level of genus with 85.86% for *rbcL* and 73.72% for the *trnL*. These values are favourable when compared to Gill *et al.* (2019) where a barcode gap was present in 73% of species barcoded with *rbcL* and 79% for the *trnL* intergenic spacer region, as well as Mishra *et al.* (2017) where the analysis was performed in one genus with the barcodes *matK*, ITS and *rbcL* and where a barcode gap was present for >70% of species.

The discrepancies between the identification success rate and the barcode gap for both datasets may indicate that the determination of a barcode gap is not an efficient way of predicting the species discrimination success rate of a reference database (Cognato, 2006; Meier *et al.*, 2008; Little *et al.*, 2013). In this instance, however, the differences are comparatively small, and the barcode gap analysis can be accepted as a more effective tool to predict database accuracy since it displays the identification success rates at species-level, as opposed to the genus-level taxonomic assignments implemented by distance-based analysis. Furthermore, these values for barcode gaps are not far from the success rate of 72% proposed by the CBOLE Plant Working Group (2009). This suggests that, when considering the expected identification success rates, a sufficient discrimination success rate would be attainable to the lowest taxonomic level of genus. The introduction of a *rbcL* and *trnL* barcode reference library to identify foraging targets in an eastern semi-arid South African savanna is an important output of this study, since metabarcoding studies focusing on African savannas are scarce (e.g., Kartzinell *et al.*, 2015; Gill *et al.*, 2020), and factually non-existing for South African savannas.

Dietary metabarcoding studies further require suitable processing of NGS data by bioinformatics, which include, (i) filtering of sequences according to quality profiles, (ii) grouping of sequencing into similar units by dissimilarity thresholds or denoising algorithms, (iii) post-clustering filtering parameters, and (iv) the application of post-clustering curation. Due to the magnitude of publicly accessible bioinformatics software and filtering threshold recommendations, a mock community was constructed from *rbcL* and *trnL*

barcoding regions of common foraging targets in a South African savanna to evaluate commonly used pipelines and filters based on precision, sensitivity and RMSE.

The mock community analysis determined DADA2 to be the optimal bioinformatics workflow which employs denoising algorithms to infer amplicon sequence variants, with the addition of post-clustering filtering targeting the removal of low frequency and less abundant reads, as well as the application of post-clustering curation (LULU) to further reduce erroneous reads resulting from PCR and sequencing errors (DADA_1_RA_LULU). The QIIME2 plugin, DADA2 has gained popularity in dietary metabarcoding studies (Bell *et al.*, 2021; Stapleton *et al.*, 2022; Díaz-Abad *et al.*, 2022) for its proven efficiency to best represent the sampled community while reducing false positives in relation to the traditional clustering of OTUs based on their sequence similarity (O'Rourke *et al.*, 2020). This pipeline was chosen due to the predetermined goal of providing species abundances that most closely represent reality (indicated by the lowest RMSE values), but to control more strictly for the identification of false positives (indicated by precision values closer or equal to 1). This goal aligns with the overall aims of this study, which firstly aims to identify, with high certainty, which food items constitute the diets of cattle foraging in the chosen ecosystem, and secondly to be able to deduce the relative importance of each food item to the diets of these herbivores.

However, for the metabarcoding dietary analysis of sequences retrieved from faecal samples, it was decided to apply a less stringent bioinformatics approach in parallel to the approach identified by mock community analysis due to concerns raised by authors such as Littleford-Colquhoun *et al.* (2022), regarding the loss of less-frequent, important food items due to stringent filtering parameters. Therefore, a further bioinformatics workflow comparison took place during the analysis of dietary profiles. The results showed that DADA_1_LULU, the less stringent approach, is suitable to describe the composition and enabled the inference of the relative importance of food items in cattle diets of plants barcoded with the *trnL* gene region. The opposite is true for the *rbcL* barcoding marker, where the number of families and genera identified per study site validated the use of the bioinformatics workflow with the addition of an abundance-based filter, as identified by the mock community analysis. After taxonomic assignment by the different bioinformatics workflows per barcoding region, some important features of this herbivore dietary study were highlighted. Firstly, was the importance of employing a dual-locus identification system for plant identifications in a metabarcoding study of plants (CBOL Plant Working group, 2009; Hollingsworth *et al.*, 2011). The full scope of the herbivorous diets would not have been reconstructed without the complementing taxonomic identifications made by both the *rbcL* and *trnL* chloroplast gene regions (Figures 5.16 and 5.17).

Secondly was the observation that *rbcL* and *trnL* have different amplification affinities for different plant families and genera. This observation is certainly in part causative of the first observation in that the

barcodes would not have been complementary if they did not possess different amplification affinities for certain plant families and genera. This is not a new observation and the affinities of certain barcoding primers to certain plant families have been exploited in studies such as Kartzinel *et al.* (2015), where the *trnL*(UAA) P6-loop was used as the core barcoding region and ITS only to ensure the resolution of certain families. Nevertheless, without the use of both plant barcoding regions the full scope of the ingested food items would not be realised as is seen in the Welverdiend dataset where Fabaceae was represented by *Albizia*, *Dichrostachys* and *Philenoptera* with the *rbcL* dataset and *Senegalia* and *Vachellia* with the *trnL* dataset. For Syferkuil this same family was represented by *Dichrostachys* and *Vachellia* with the *rbcL* marker and *Senegalia*, *Vachellia*, *Indigofera* and *Senna* with the *trnL* marker. The *rbcL* barcoding marker therefore exclusively identified the genus *Dichrostachys* across both datasets and the *trnL* marker exclusively identified *Senegalia* across both datasets. Consequently, the recommended use of multiple gene regions (CBOL Plant Working Group, 2009) is crucial for studies using DNA metabarcoding techniques to document the diverse diet of a generalist herbivore.

During taxonomic assignment the compiled reference database proved its efficiency when applied to identify unknown query sequences from Welverdiend and Syferkuil Illumina sequence data. A total of 18 families were identified in the diets of Welverdiend cattle, two of these exclusively belonging to the *rbcL* dataset and ten exclusively belonging to the *trnL* dataset. Twenty genera were identified in this dataset with both barcoding markers with only one family detected by both barcoding datasets. A total of 12 families were identified in the diets of Syferkuil cattle, with four exclusively belonging to the *rbcL* dataset and three to the *trnL* dataset. A total of 19 genera were detected for Syferkuil with nine exclusively belonging to the *rbcL* and six exclusively to the *trnL* dataset. Although the number of identifications seems low, one must remember that these sequences are only identified to genus level, with much more diversity possibly contained in species level assignments. It is also important to consider that the faecal samples originated from cattle foraging in the dry season which would naturally lead to a less diverse flora in their diets. The reference database also reported many unidentified reads for both barcoding markers, especially for the Syferkuil *trnL* dataset. This was concerning, but a blast search only validated the robustness of the compiled reference database since these unidentified reads could also not be assigned to plant taxa by GenBank with BLAST parameters: > 80% query coverage, > 80% percentage identity and e-values smaller than 0,001. These unidentified reads could be a result of nonspecific primer binding, especially exhibited by the *trnL* primers, the inability of the *rbcLa* and *trnL*(UAA) intron primer pairs to amplify certain taxa, chimera formation, contamination entering the workflow at any point from DNA isolation to sequencing (declaring the high number of unidentified reads assigned to bacterial/other species in the Syferkuil *trnL* dataset) as well as simply being undescribed biodiversity that is not in the scope of the compiled reference database or any public reference databases. As stated, forbs are neglected in savanna management models and

definitions, and this further translates to the lack of their representation on public reference databases used for sequence mining.

The application of quantifying diets to illustrate relative read abundance, i.e., an indication of the proportions of food items to individual cattle and not just reporting on the observed frequencies of reads as a measure of describing dietary profiles, have been heavily debated and the use of this metric is cautioned due to the prevalence of biases (Alberdi *et al.*, 2018; Lamb *et al.*, 2019). However, the interpretation of these results, with consideration of the associated biases, may prove helpful in study site comparisons. The application FOO and RRA were useful for inferring the importance of forbs to herbivorous diets at both Welverdiend and Syferkuil. Forbs were foraged by a noticeable percentage of individuals in the sampled dataset, especially *Asparagus* and *Ipomoea* at Welverdiend and *Adenium*, *Pollichia* and *Indigofera* at Syferkuil. Furthermore, Welverdiend had the highest FOO of forb species across the sampled dataset (i.e., more forb species in diets), while Syferkuil results illustrated a greater RRA of forbs in individual diets, i.e., lower richness, but higher abundance of single species in individual diets. This implies that although forbs at Welverdiend were utilised by many members of the sampled dataset, they have not been heavily grazed, and their diets still included high abundances of other plant life forms. At Syferkuil, however, forb frequencies were lower across the sampled dataset but very abundant in the diets of some individuals, meaning that forb species that an individual came across was heavily grazed, in a higher degree than plants from other life forms.

Chapter 7 Overall Conclusion

The first aim of this study was to be able to identify the full scope of plant species present in the diets of cattle foraging in two different bioregions of the South African savanna. This aim was firstly fulfilled by the creation of a reliable DNA barcode reference library, that according to the presence of barcode gaps for species barcoded with both *rbcL* and *trnL* at suitable species discrimination success rates, is safe to use to assign taxonomies to the local flora. This aim was further fulfilled by the identification of an optimal bioinformatics workflow by the application of a mock community. The denoising method, DADA2, showed a superior performance to the other commonly applied workflows (VSAERCH and OBITOOLS) when the research question needs to assess community richness and would benefit from precise estimations of features that can be assigned to their taxonomy and not the inclusion of reads that cannot be identified that would be misleading to the diversity contained in the environmental sample. The mock community analysis results also highlighted the need for feature table filtering to remove possible contamination, PCR, or sequencing errors and in doing so results in fewer sequences that cannot be taxonomically assigned; or the inclusion of features resulting from contamination that would overestimate community richness and not reflect animal foraging behaviour.

The usefulness of evaluating post-clustering, or denoising, filters to retrieve the highest taxonomic richness is illustrated by the application of an abundance filter to the *rbcL* and *trnL* datasets. Too stringent filtering parameters applied to the *trnL* datasets led to the loss of biodiversity. The decision to use different bioinformatics approaches per barcoding marker facilitates the identification of all foraging targets in the diets of these cattle. The fact that the *rbcL* marker filled in for the *trnL* marker's missing taxonomy and vice versa further demonstrated the efficacy of a dual-locus barcoding technique for the identification of plants in this study. The robustness of the reference sequence database was also confirmed by the fact that none of the unidentified reads per dataset could be assigned to plant species using a public reference sequence database using lower stringency parameters. The food items occurring in the diets of herbivores foraging in an eastern semi-arid South African savanna could therefore be recovered and taxonomically assigned by the application of the correct bioinformatics workflow as well as the use of a study-area-specific reference database.

The second aim deals with the quantification of herbivore diets to deduce the importance of plant life forms to the diets of generalist herbivores in a savanna ecosystem. The DADA2 method with frequency, abundance (in the case of *rbcL* datasets), and feature (ASV) curation was chosen as the bioinformatics method to identify plant species in the faecal samples. Feature occurrences (FOO) could be used to estimate the presence/absence data of the food items in these diets and RRA was cautiously, and with a full understanding of the associated biases and limitations, used to quantify the contribution of forbs. Dietary quantification allowed comparisons between study sites used for faecal sample collection: In contrast to

Welverdiend, where more members of the sampled community foraged on certain forbs, the relative intake of these forbs was not high, Syferkuil cattle did not appear to visit forbs very frequently, but when they did, the intake of these species was relatively high. The identification of which forbs are important to the diets of herbivores foraging in savannas as well as the quantification of the importance of these species will allow for the creation of better savanna management models and aid in further research better describing these ecosystems with the inclusion of this previously neglected life form.

Recommendations

- The collection, vouchering and sequencing of more species, especially forb species from South African savannas will aid in more comprehensive public reference databases that can then be used as a sequence mining source to better the resolution of smaller, study-area-specific databases by the addition of new species and the minimization of singletons of other species.
- The decision to not form a consensus sequence for the overlapping region of *rbcLa*-F and *rbcLa*-R represents a missed opportunity for comparisons between the success rate of taxonomic assignment of merged versus unmerged sequence data as well as evaluating the performance of the OBITOOLS bioinformatics pipeline in comparison to DADA2 and VSEARCH.
- The dissimilarity thresholds to cluster sequences into operational taxonomic units, which in this study is 3%, may have been too stringent and led to the removal of several sequences that could have been correctly identified with a higher dissimilarity value. Therefore, considering adaptive thresholds would be crucial to ensure that the full scope of dietary items is retrieved.
- Many more bioinformatics pipelines and filtering thresholds exist, and a comprehensive comparison between these to recover dietary profiles may be helpful to further distinguish and decide upon a standard workflow or best practice guidelines to use in herbivore dietary studies to ensure comparability between studies.
- Although closed-reference OTU picking would not have been able to represent undiscovered biodiversity and therefore not be able to capture the full scope of animal diets, it may be useful for comparisons of known diversity (present on the reference sequence database) between study sites or between different animals foraging in the same ecosystem.

Bibliography

- Alberdi, A., Aizpurua, O., Gilbert, M.T.P. & Bohmann, K. 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, 9(1):134-147.
- Amandita, F.Y., Rembold, K., Vornam, B., Rahayu, S., Siregar, I.Z., Kreft, H. & Finkeldey, R. 2019. DNA barcoding of flowering plants in Sumatra, Indonesia. *Ecology and evolution*, 9(4):1858-1868.
- Amir, A., McDonald, D., Navas-Molina, J.A., Kopylova, E., Morton, J.T., Zech Xu, Z., Kightley, E.P., Thompson, L.R., Hyde, E.R., Gonzalez, A. & Knight, R. 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. *MSystems*, 2(2). <https://doi.org/10.1128/mSystems.00191-16> Date of access: 1 Nov. 2022.
- Ando, H., Fujii, C., Kawanabe, M., Ao, Y., Inoue, T. & Takenaka, A. 2018. Evaluation of plant contamination in metabarcoding diet analysis of a herbivore. *Scientific reports*, 8(1):1-10.
- Ando, H., Mukai, H., Komura, T., Dewi, T., Ando, M. & Isagi, Y. 2020. Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. *Environmental DNA*, 2(4):391-406
- Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> Date of access 17 Mar. 2021.
- Angly, F.E., Willner, D., Rohwer, F., Hugenholtz, P. & Tyson, G.W. 2012. Grinder: a versatile amplicon and shotgun sequence simulator. *Nucleic acids research*, 40(12). <https://doi.org/10.1093/nar/gks251> Date of access: 8 Jul. 2022.
- Anthoons, B., Karamichali, I., Schröder-Nielsen, A., Drouzas, A.D., de Boer, H. & Madesis, P. 2021. Metabarcoding reveals low fidelity and presence of toxic species in short chain-of-commercialization of herbal products. *Journal of Food Composition and Analysis*, 97. <https://doi.org/10.1016/j.jfca.2020.103767> Date of access: 13 May 2022.
- Armstrong, K.F. & Ball, S.L. 2005. DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462):1813-1823.
- Arnot, D.E., Roper, C. & Bayoumi, R.A. 1993. Digital codes from hypervariable tandemly repeated DNA sequences in the *Plasmodium falciparum* circumsporozoite gene can genetically barcode isolates. *Molecular and Biochemical Parasitology*, 61(1):15-24.

- Aylward, C.M., Statham, M.J., Barthman-Thompson, L., Kelt, D.A. & Sacks, B.N. 2022. Dietary characterization of the endangered salt marsh harvest mouse and sympatric rodents using DNA metabarcoding. *Ecology and evolution*, 12(7). <https://doi.org/10.1002/ece3.9121> Date of access: 23 Aug. 2022.
- Bálint, M., Bahram, M., Eren, A.M., Faust, K., Fuhrman, J.A., Lindahl, B., O'Hara, R.B., Öpik, M., Sogin, M.L., Unterseher, M. & Tedersoo, L. 2016. Millions of reads, thousands of taxa: microbial community structure and associations analyzed via marker genes. *FEMS microbiology reviews*, 40(5):686-700.
- Banchi, E., Ametrano, C.G., Greco, S., Stanković, D., Muggia, L. & Pallavicini, A. 2020. PLANiTS: a curated sequence reference dataset for plant ITS DNA metabarcoding. *Database*, 2020. <https://doi.org/10.1093/database/baz155> Date of access: 8 Jul. 2022.
- Barnes, C.J., Nielsen, I.B., Aagaard, A., Ejrnæs, R., Hansen, A.J. & Frøslev, T.G. 2022. Metabarcoding of soil environmental DNA replicates plant community variation but not specificity. *Environmental DNA*, 4(4):732-746.
- Baumgartner, S.A., Treydte, A.C., Grant, C.C. & van Rooyen, J. 2015. Can diverse herbivore communities increase landscape heterogeneity? Comparing wild and domestic herbivore assemblages in a South African savanna. *Perspectives in Plant Ecology, Evolution and Systematics*, 17(1):34-43.
- Bell, K.L., Batchelor, K.L., Bradford, M., McKeown, A., Macdonald, S.L. & Westcott, D. 2021. Optimisation of a pollen DNA metabarcoding method for diet analysis of flying-foxes (*Pteropus* spp.). *Australian Journal of Zoology*, 68:273-284.
- Bell, K.L., Burgess, K.S., Botsch, J.C., Dobbs, E.K., Read, T.D. & Brosi, B.J. 2019. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular Ecology*, 28(2):431-455.
- Bell, K.L., Loeffler, V.M. & Brosi, B.J. 2017. An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Applications in plant sciences*, 5(3). <https://doi.org/10.3732/apps.1600110> Date of access: 15 Feb. 2022.
- Ben-David, M., Shochat, E. & Adams, L.G. 2001. Utility of stable isotope analysis in studying foraging ecology of herbivores: examples from moose and caribou. *Alces: A Journal Devoted to the Biology and Management of Moose*, 37(2):421-434.

- Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Ostell, J., Pruitt, K.D. & Sayers, E.W. 2018. GenBank. *Nucleic acids research*, 46(D1). <https://doi.org/10.1093/nar/gkx1094> Date of access: 1 Nov. 2022.
- Bhargava, M. & Sharma, A. 2013. DNA barcoding in plants: evolution and applications of in silico approaches and resources. *Molecular phylogenetics and evolution*, 67(3):631-641.
- Bhattacharyya, S., Dawson, D.A., Hipperson, H. & Ishtiaq, F. 2019. A diet rich in C3 plants reveals the sensitivity of an alpine mammal to climate change. *Molecular Ecology*, 28(2):250-265.
- Biffi, M., Gillet, F., Laffaille, P., Colas, F., Aulagnier, S., Blanc, F., Galan, M., Tiouchichine, M.L., Némoy, M., Buisson, L. & Michaux, J.R. 2017. Novel insights into the diet of the Pyrenean desman (*Galemys pyrenaicus*) using next-generation sequencing molecular analyses. *Journal of Mammalogy*, 98(5):1497-1507.
- Boggs, L.M., Scheible, M.K., Machado, G. & Meiklejohn, K.A. 2019. Single fragment or bulk soil DNA metabarcoding: which is better for characterizing biological taxa found in surface soils for sample separation? *Genes*, 10(6):431.
- Bokulich, N.A., Subramanian, S., Faith, J.J., Gevers, D., Gordon, J.I., Knight, R., Mills, D.A. & Caporaso, J.G. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature methods*, 10(1):57-59.
- Boval, M., Coates, D.B., Lecomte, P., Decruyenaere, V. & Archimede, H. 2004. Faecal near infrared reflectance spectroscopy (NIRS) to assess chemical composition, in vivo digestibility and intake of tropical grass by Creole cattle. *Animal Feed Science and Technology*, 114(1-4):19-29.
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P. & Coissac, E. 2016. Obitools: A unix-inspired software package for DNA metabarcoding. *Molecular ecology resources*, 16(1):176-182.
- Brown, S.D., Collins, R.A., Boyer, S., Lefort, M.C., Malumbres-Olarte, J.A.G.O.B.A., Vink, C.J. & Cruickshank, R.H. 2012. Spider: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Molecular ecology resources*, 12(3):562-565.
- Buchner, D. & Leese, F. 2020. BOLDigger—a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding and Metagenomics*, 4. <https://doi.org/10.3897/mbmg.4.53535> Date of access: 15 Feb. 2022.

- Buglione, M., Maselli, V., Rippa, D., de Filippo, G., Trapanese, M. & Fulgione, D. 2018. A pilot study on the application of DNA metabarcoding for non-invasive diet analysis in the Italian hare. *Mammalian Biology*, 88:31-42.
- Buisson, E., Fidelis, A., Overbeck, G.E., Schmidt, I.B., Durigan, G., Young, T.P., Alvarado, S.T., Arruda, A.J., Boisson, S., Bond, W. & Coutinho, A. 2021. A research agenda for the restoration of tropical and subtropical grasslands and savannas. *Restoration Ecology*, 29. <https://doi.org/10.1111/rec.13292> Date of access: 20 April 2022.
- Callahan, B.J., McMurdie, P.J. & Holmes, S.P. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal*, 11(12):2639-2643.
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. & Holmes, S.P. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods*, 13(7):581-583.
- Camp, A., Croxford, A.E., Ford, C.S., Baumann, U., Clements, P.R., Hiendleder, S., Woolford, L., Netzel, G., Boardman, W.S., Fletcher, M.T. & Wilkinson, M.J. 2020. Dual-locus DNA metabarcoding reveals southern hairy-nosed wombats (*Lasiorchinus latifrons* Owen) have a summer diet dominated by toxic invasive plants. *PLoS One*, 15(3). <https://doi.org/10.1371/journal.pone.0229390> Date of access: 26 Aug. 2022.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Gonzalez Pena, A., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, C., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J. & Knight, R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, 7(5):335-336.
- CBOL Plant Working Group. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, 106(31):12794-12797.
- Chase, M.W., Salamin, N., Wilkinson, M., Dunwell, J.M., Kesanakurthi, R.P., Haidar, N. & Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1462):1889-1895.
- Chua, P.Y., Crampton-Platt, A., Lammers, Y., Alsos, I.G., Boessenkool, S. & Bohmann, K. 2021. Metagenomics: A viable tool for reconstructing herbivore diet. *Molecular ecology resources*, 21(7):2249-2263.

- Clare, E.L. 2014. Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. *Evolutionary applications*, 7(9):1144-1157.
- Cognato, A.I. 2006. Standard percent DNA sequence difference for insects does not predict species boundaries. *Journal of Economic Entomology*, 99(4):1037-1045.
- Coissac, E., Hollingsworth, P.M., Lavergne, S. & Taberlet, P. 2016. From barcodes to genomes: extending the concept of DNA barcoding. *Molecular Ecology*, 25(7):1423-1428.
- Coissac, E., Riaz, T. & Puillandre, N. 2012. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular ecology*, 21(8):1834-1847.
- Collins, R.A. & Cruickshank, R.H. 2013. The seven deadly sins of DNA barcoding. *Molecular ecology resources*, 13(6):969-975.
- Craigie, I.D., Baillie, J.E., Balmford, A., Carbone, C., Collen, B., Green, R.E. & Hutton, J.M. 2010. Large mammal population declines in Africa's protected areas. *Biological conservation*, 143(9):2221-2228.
- Cristescu, M.E. 2016. Genetic reconstructions of invasion history. In: Barrett, S.C.H., Colautti, R.I., Dlugosch, K.M. & Rieseberg, L.H., eds. *Invasion Genetics: The Baker and Stebbins Legacy*. Wiley Online Library. pp. 267-282.
- Davis, A.L., Scholtz, C.H. & Swemmer, A.M. 2012. Effects of land usage on dung beetle assemblage structure: Kruger National Park versus adjacent farmland in South Africa. *Journal of Insect Conservation*, 16(3):399-411.
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E. & Taberlet, P. 2014. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular ecology resources*, 14(2):306-323.
- De Groot, G.A., During, H.J., Maas, J.W., Schneider, H., Vogel, J.C. & Erkens, R.H. 2011. Use of rbcL and trnL-F as a two-locus DNA barcode for identification of NW-European ferns: an ecological perspective. *PLos One*, 6(1). <https://doi.org/10.1371/journal.pone.0016371> Date of access: 15 Feb. 2022.
- De Vere, N., Rich, T.C., Ford, C.R., Trinder, S.A., Long, C., Moore, C.W., Satterthwaite, D., Davies, H., Allainguillaume, J., Ronca, S. & Tatarinova, T. 2012. DNA barcoding the native flowering plants and conifers of Wales. *PLos One*, 7(6). <https://doi.org/10.1371/journal.pone.0037945> Date of access: 18 Nov. 2022.

- Deagle, B.E., Chiaradia, A., McInnes, J. & Jarman, S.N. 2010. Pyrosequencing faecal DNA to determine diet of little penguins: is what goes in what comes out? *Conservation Genetics*, 11(5):2039-2048.
- Deagle, B.E., Clarke, L.J., Kitchener, J.A., Polanowski, A.M. & Davidson, A.T. 2018. Genetic monitoring of open ocean biodiversity: An evaluation of DNA metabarcoding for processing continuous plankton recorder samples. *Molecular ecology resources*, 18(3):391-406.
- Deagle, B.E., Eveson, J.P. & Jarman, S.N. 2006. Quantification of damage in DNA recovered from highly degraded samples—a case study on DNA in faeces. *Frontiers in zoology*, 3(1):1-10.
- Deagle, B.E., Jarman, S.N., Coissac, E., Pompanon, F. & Taberlet, P. 2014. DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology letters*, 10(9). <https://doi.org/10.1098/rsbl.2014.0562> Date of access: 29 Jul. 2022.
- Deagle, B.E., Kirkwood, R. & Jarman, S.N. 2009. Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular ecology*, 18(9):2022-2038.
- Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L., Kartzinel, T.R. & Eveson, J.P. 2019. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular ecology*, 28(2):391-406.
- Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D.M., de Vere, N. & Pfrender, M.E. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular ecology*, 26(21):5872-5895.
- Díaz-Abad, L., Bacco-Mannina, N., Madeira, F.M., Neiva, J., Aires, T., Serrao, E.A., Regalla, A., Patrício, A.R. & Frade, P.R. 2022. eDNA metabarcoding for diet analyses of green sea turtles (*Chelonia mydas*). *Marine Biology*, 169(1):1-12.
- Drew, B.T., Ruhfel, B.R., Smith, S.A., Moore, M.J., Briggs, B.G., Gitzendanner, M.A., Soltis, P.S. & Soltis, D.E. 2014. Another look at the root of the angiosperms reveals a familiar tale. *Systematic Biology*, 63(3):368-382.
- Du Toit, J.T., Rogers, K.H. & Biggs, H.C. 2003. *The Kruger experience: Ecology and management of savanna heterogeneity*. Washington, DC: Island Press.
- Ducotterd, C., Crovadore, J., Lefort, F., Guisan, A., Ursenbacher, S. & Rubin, J.F. 2020. The feeding behaviour of the European pond turtle (*Emys orbicularis*, L. 1758) is not a threat for other endangered species. *Global Ecology and Conservation*, 23. <https://doi.org/10.1111/1755-0998.13277> Date of access: 4 Aug. 2022.

- Duda, R.O., Hart, P.E. & Stork, D.G. 1973. *Pattern classification and scene analysis*. 2nd ed. New York: Wiley.
- Duffy, J.E., Cardinale, B.J., France, K.E., McIntyre, P.B., Thébault, E. & Loreau, M. 2007. The functional role of biodiversity in ecosystems: incorporating trophic complexity. *Ecology letters*, 10(6):522-538.
- Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5):1792-1797.
- Edgar, R.C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19):2460-2461.
- Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C. & Knight, R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16):2194-2200.
- Edwards, C.E., Swift, J.F., Lance, R.F., Minckley, T.A. & Lindsay, D.L. 2019. Evaluating the efficacy of sample collection approaches and DNA metabarcoding for identifying the diversity of plants utilized by nectivorous bats. *Genome*, 62(1):19-29.
- Erasmus, L.L., Van Coller, H. & Siebert, F. 2021. Teatime in Kruger: Tailoring the application of the Tea Bag Index approach to an African savanna. *South African Journal of Science*, 117(1-2):1-6.
- Ewels, P., Magnusson, M., Lundin, S. & Käller, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19):3047-3048.
- Fahner, N.A., Shokralla, S., Baird, D.J. & Hajibabaei, M. 2016. Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. *PLoS One*, 11(6). <https://doi.org/10.1371/journal.pone.0157505> Date of access: 18 Nov. 2022.
- Ferri, G., Corradini, B., Ferrari, F., Santunione, A.L., Palazzoli, F. & Alu, M. 2015. Forensic botany II, DNA barcode for land plants: Which markers after the international agreement? *Forensic Science International: Genetics*, 15:131-136.
- Foley, W.J., McIlwee, A., Lawler, I., Aragonés, L., Woolnough, A.P. & Berding, N. 1998. Ecological applications of near infrared reflectance spectroscopy—a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia*, 116(3):293-305.
- Fonseca, M.L., Ramírez-Pinzón, M.A., McNeil, K.N., Guevara, M., Gómez-Gutiérrez, L.M., Harter, K., Mongui, A. & Stevenson, P.R. 2022. Dietary preferences and feeding strategies of Colombian highland woolly monkeys. *Scientific reports*, 12(1):1-15.

- Frøslev, T.G., Kjølner, R., Bruun, H.H., Ejrnæs, R., Brunbjerg, A.K., Pietroni, C. & Hansen, A.J. 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature communications*, 8(1):1-11.
- Galan, M., Pons, J.B., Tournayre, O., Pierre, E., Leuchtmann, M., Pontier, D. & Charbonnel, N. 2018. Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis. *Molecular ecology resources*, 18(3):474-489.
- García-Robledo, C., Erickson, D.L., Staines, C.L., Erwin, T.L. & Kress, W.J. 2013. Tropical plant–herbivore networks: reconstructing species interactions using DNA barcodes. *PLos One*, 8(1). <https://doi.org/10.1371/journal.pone.0052967> Date of access: 23 Aug. 2022.
- Gardner, P.P., Watson, R.J., Morgan, X.C., Draper, J.L., Finn, R.D., Morales, S.E. & Stott, M.B. 2019. Identifying accurate metagenome and amplicon software via a meta-analysis of sequence to taxonomy benchmarking studies. *PeerJ*, 7. <https://doi.org/10.7717/peerj.6160> Date of access: 8 Jul. 2022.
- Gebremedhin, B., Flagstad, Ø., Bekele, A., Chala, D., Bakkestuen, V., Boessenkool, S., Popp, M., Gussarova, G., Schrøder-Nielsen, A., Nemomissa, S. & Brochmann, C. 2016. DNA metabarcoding reveals diet overlap between the endangered walia ibex and domestic goats-implications for conservation. *PLos One*, 11(7). <https://doi.org/10.1371/journal.pone.0159133> Date of access: 14 Apr. 2022.
- Geml, J., Pastor, N., Fernandez, L., Pacheco, S., Semenova, T.A., Becerra, A.G., Wicaksono, C.Y. & Nouhra, E.R. 2014. Large-scale fungal diversity assessment in the Andean Yungas forests reveals strong community turnover among forest types along an altitudinal gradient. *Molecular Ecology*, 23(10):2452-2472.
- Gill, B.A., Musili, P.M., Kurukura, S., Hassan, A.A., Goheen, J.R., Kress, W.J., Kuzmina, M., Pringle, R.M. & Kartzinell, T.R. 2019. Plant DNA-barcode library and community phylogeny for a semi-arid East African savanna. *Molecular ecology resources*, 19(4): 838-846.
- Gitzendanner, M.A., Soltis, P.S., Wong, G.K.S., Ruhfel, B.R. & Soltis, D.E. 2018. Plastid phylogenomic analysis of green plants: a billion years of evolutionary history. *American Journal of Botany*, 105(3):291-301.
- Goldberg, A.R., Conway, C.J., Tank, D.C., Andrews, K.R., Gour, D.S. & Waits, L.P. 2020. Diet of a rare herbivore based on DNA metabarcoding of feces: Selection, seasonality, and survival. *Ecology and evolution*, 10(14):7627-7643.

- Gourlé, H., Karlsson-Lindsjö, O., Hayer, J. & Bongcam-Rudloff, E. 2019. Simulating Illumina metagenomic data with InSilicoSeq. *Bioinformatics*, 35(3):521-522.
- Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Sodergren, E. & Methé, B. 2011. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome research*, 21(3):494-504.
- Hajibabaei, M., Baird, D.J., Fahner, N.A., Beiko, R. & Golding, G.B. 2016. A new way to contemplate Darwin's tangled bank: how DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1702). <https://doi.org/10.1098/rstb.2015.0330> Date of access: 26 Aug. 2022.
- Hamner, B., Frasco, M. & LeDell, E. 2018. Metrics: Evaluation metrics for machine learning. *R package version 0.1*, 4.
- Harris, D.J. 2003. Can you bank on GenBank?. *Trends in ecology & evolution*, 18(7): 317-319.
- Hayes, G.F. & Holl K.D. 2003. Cattle Grazing Impacts on Annual Forbs and Vegetation Composition of Mesic Grasslands in California. *Conservation Biology*, 17(6):1697-1702.
- Hebert, P.D., Ratnasingham, S. & De Waard, J.R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270. <https://doi.org/10.1098/rsbl.2003.0025> Date of access: 14 Apr. 2022.
- Holecheck, J.L., Vavra, M. & Pieper, R.D. 1982. Methods for determining the nutritive quality of range ruminant diets: a review. *Journal of Animal Science*, 54(2):363-376.
- Hollingsworth, P.M., Graham, S.W. & Little, D.P. 2011. Choosing and using a plant DNA barcode. *PLoS One*, 6(5). <https://doi.org/10.1371/journal.pone.0019254>. Date of access: 5 May 2022.
- Huntley, B.J. 1982. Southern African savannas. In: Huntley, B.J. & Walker, B.H., eds. *Ecology of tropical savannas*. Vol. 42. Heidelberg: Springer. pp. 101-119.
- Huse, S.M., Welch, D.M., Morrison, H.G. & Sogin, M.L. 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental microbiology*, 12(7):1889-1898.
- Iwanowicz, D.D., Vandergast, A.G., Cornman, R.S., Adams, C.R., Kohn, J.R., Fisher, R.N. & Brehme, C.S. 2016. Metabarcoding of fecal samples to determine herbivore diets: A case study of the endangered Pacific pocket mouse. *PLoS One*, 11(11). <https://doi.org/10.1371/journal.pone.0165366> Date of access: 20 Apr 2022.

- Jeske, J.T. & Gallert, C. 2022. Microbiome Analysis via OTU and ASV-Based Pipelines—A Comparative Interpretation of Ecological Data in WWTP Systems. *Bioengineering*, 9(4):146.
- Johannessen, V. & Samset, E. 1994. Summer diet of the mountain hare (*Lepus timidus* L.) in a low-alpine area in southern Norway. *Canadian Journal of Zoology*, 72(4):652-657.
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezukh, Y., McGinnis, S. & Madden, T.L. 2008. NCBI BLAST: a better web interface. *Nucleic acid research*, 36. <https://doi.org/10.1093/nar/gkn201> Date of access: 17 Mar. 2021.
- Johnson, M.D., Fokar, M., Cox, R.D. & Barnes, M.A. 2021. Airborne environmental DNA metabarcoding detects more diversity, with less sampling effort, than a traditional plant community survey. *BMC Ecology and Evolution*, 21(1):1-15.
- Joos, L., Beirinckx, S., Haegeman, A., Debode, J., Vandecasteele, B., Baeyen, S., Goormachtig, S., Clement, L. & De Tender, C. 2020. Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. *BMC genomics*, 21(1):1-17.
- Jurado-Rivera, J.A., Vogler, A.P., Reid, C.A., Petitpierre, E. and Gómez-Zurita, J. 2009. DNA barcoding insect–host plant associations. *Proceedings of the Royal Society B: Biological Sciences*, 276(1657):639-648.
- Kang, W., Anslan, S., Börner, N., Schwarz, A., Schmidt, R., Künzel, S., Rioual, P., Echeverría-Galindo, P., Vences, M., Wang, J. & Schwalb, A. 2021. Diatom metabarcoding and microscopic analyses from sediment samples at Lake Nam Co, Tibet: The effect of sample-size and bioinformatics on the identified communities. *Ecological Indicators*, 121(2021). <https://doi.org/10.1016/j.ecolind.2020.107070>. Date of access: 1 Mei 2022.
- Kartzinel, T.R. & Pringle, R.M. 2020. Multiple dimensions of dietary diversity in large mammalian herbivores. *Journal of Animal Ecology*, 89(6):1482-1496.
- Kartzinel, T.R., Chen, P.A., Coverdale, T.C., Erickson, D.L., Kress, W.J., Kuzmina, M.L., Rubenstein, D.I., Wang, W. & Pringle, R.M. 2015. DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proceedings of the National Academy of Sciences*, 112(26):8019-8024.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*, 16(2):111-120.

- Klem, J. 2019. *Drought responses of forb and grass communities in communal and protected rangelands*. Potchefstroom: NWU. (Dissertation – MSc).
- Kohn, M.H. & Wayne, R.K. 1997. Facts from feces revisited. *Trends in ecology & evolution*, 12(6):223-227.
- Kopylova, E., Navas-Molina, J.A., Mercier, C., Xu, Z.Z., Mahé, F., He, Y., Zhou, H.W., Rognes, T., Caporaso, J.G. & Knight, R. 2016. Open-source sequence clustering methods improve the state of the art. *MSystems*, 1(1). <https://doi.org/10.1128/mSystems.00003-15> Date of access: 16 May 2022.
- Kowalczyk, R., Wójcik, J.M., Taberlet, P., Kamiński, T., Miquel, C., Valentini, A., Craine, J.M. & Coissac, E. 2019. Foraging plasticity allows a large herbivore to persist in a sheltering forest habitat: DNA metabarcoding diet analysis of the European bison. *Forest Ecology and Management*, 449. <https://doi.org/10.1016/j.foreco.2019.117474> Date of access: 4 Apr. 2022.
- Kraaijeveld, K., De Weger, L.A., Ventayol García, M., Buermans, H., Frank, J., Hiemstra, P.S. & Den Dunnen, J.T. 2015. Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Molecular ecology resources*, 15(1):8-16.
- Kress, W.J. & Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One*, 2(6). <https://doi.org/10.1371/journal.pone.0000508> Date of access: 7 Apr. 2022.
- Kress, W.J., García-Robledo, C., Uriarte, M. & Erickson, D.L. 2015. DNA barcodes for ecology, evolution, and conservation. *Trends in ecology & evolution*, 30(1):25-35.
- Lahaye, R., Van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T.G. & Savolainen, V. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences*, 105(8):2923-2928.
- Lamb, E.G., Winsley, T., Piper, C.L., Freidrich, S.A. & Siciliano, S.D. 2016. A high-throughput belowground plant diversity assay using next-generation sequencing of the *trnL* intron. *Plant and Soil*, 404(1):361-372.
- Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. & Taylor, M.I. 2019. How quantitative is metabarcoding: A meta-analytical approach. *Molecular ecology*, 28(2):420-430.
- Lee, T., Alemseged, Y. & Mitchell, A. 2018. Dropping Hints: Estimating the diets of livestock in rangelands using DNA metabarcoding of faeces. *Metabarcoding and Metagenomics*, 2:1-17.

- Levin, R.A., Wagner, W.L., Hoch, P.C., Nepokroeff, M., Pires, J.C., Zimmer, E.A. & Sytsma, K.J. 2003. Family-level relationships of Onagraceae based on chloroplast *rbcL* and *ndhF* data. *American Journal of Botany*, 90:107-115.
- Lin, X.L., Mo, L., Bu, W.J. & Wang, X.H. 2021. The first comprehensive DNA barcode reference library of Chinese *Tanytarsus* (Diptera: Chironomidae) for environmental DNA metabarcoding. *Diversity and Distributions*, 27(10):1932-1941.
- Linder, H.P. 2014. The evolution of African plant diversity. *Frontiers in Ecology and Evolution*, 2:38.
- Little, D.P., Knopf, P. & Schulz, C. 2013. DNA barcode identification of Podocarpaceae—the second largest conifer family. *PLoS One*, 8(11). <https://doi.org/10.1371/journal.pone.0081008> Date of access: 14 Apr. 2022.
- Littleford-Colquhoun, B.L., Freeman, P.T., Sackett, V.I., Tulloss, C.V., McGarvey, L.M., Geremia, C. & Kartzinell, T.R. 2022. The precautionary principle and dietary DNA metabarcoding: commonly used abundance thresholds change ecological interpretation. *Molecular Ecology*, 31(6):1615-1626.
- MacArthur, R.H. & Pianka, E.R. 1966. On optimal use of a patchy environment. *The American Naturalist*, 100(916):603-609.
- MacGillivray, M.L. & Kaczmarek, I. 2011. Survey of the efficacy of a short fragment of the *rbc L* gene as a supplemental DNA barcode for diatoms. *Journal of Eukaryotic Microbiology*, 58(6):529-536.
- Mallott, E.K., Garber, P.A. & Malhi, R.S. 2018. *trnL* outperforms *rbcL* as a DNA metabarcoding marker when compared with the observed plant component of the diet of wild white-faced capuchins (*Cebus capucinus*, Primates). *PLoS One*, 13(6). <https://doi.org/10.1371/journal.pone.0199556> Date of access: 1 Sep. 2020.
- Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*, 17(1):10-12.
- Mata, V.A., Rebelo, H., Amorim, F., McCracken, G.F., Jarman, S. & Beja, P. 2019. How much is enough? Effects of technical and biological replication on metabarcoding dietary analysis. *Molecular Ecology*, 28(2):165-175.
- Matesanz, S., Pescador, D.S., Pías, B., Sánchez, A.M., Chacón-Labela, J., Illuminati, A., de la Cruz, M., López-Angulo, J., Marí-Mena, N., Vizcaíno, A. & Escudero, A. 2019. Estimating belowground plant abundance with DNA metabarcoding. *Molecular Ecology Resources*, 19(5): 1265-1277.

- Mathon, L., Valentini, A., Guérin, P.E., Normandeau, E., Noel, C., Lionnet, C., Boulanger, E., Thuiller, W., Bernatchez, L., Mouillot, D. & Dejean, T. 2021. Benchmarking bioinformatic tools for fast and accurate eDNA metabarcoding species identification. *Molecular Ecology Resources*, 21(7):2565-2579.
- Matsika, R., Erasmus, B.F. & Twine, W.C. 2013. A tale of two villages: assessing the dynamics of fuelwood supply in communal landscapes in South Africa. *Environmental Conservation*, 40(1):71-83.
- McInnis, M.L., Vavra, M. & Krueger, W.C. 1983. A comparison of four methods used to determine the diets of large herbivores Sheep. *Rangeland Ecology & Management/Journal of Range Management Archives*, 36(3):302-306.
- McMurdie, P.J. & Holmes, S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, 8(4). <https://doi.org/10.1371/journal.pone.0061217> Date of access: 15 Jun. 2022.
- Meier, R., Shiyang, K., Vaidya, G. & Ng, P.K. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic biology*, 55(5):715-728.
- Meier, R., Zhang, G. & Ali, F. 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the “barcoding gap” and leads to misidentification. *Systematic biology*, 57(5):809-813.
- Meiklejohn, K.A., Damaso, N. & Robertson, J.M. 2019. Assessment of BOLD and GenBank –Their accuracy and reliability for the identification of biological materials. *PLoS One*, 14(6). <https://doi.org/10.1371/journal.pone.0217084> Date of access: 10 Apr. 2022.
- Meyer, C.P. & Paulay, G. 2005. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biology*, 3(12). <https://doi.org/10.1371/journal.pbio.0030422>. Date of access: 5 May 2022.
- Meyer, J.M., Honig, N. & Hadly, E.A. 2022. Diet DNA reveals novel African Forest elephant ecology on the grasslands of the Congo Basin. *Environmental DNA*, 4(4):846-867.
- Mishra, P., Kumar, A., Sivaraman, G., Shukla, A.K., Kaliamoorthy, R., Slater, A. & Velusamy, S. 2017. Character-based DNA barcoding for authentication and conservation of IUCN Red listed threatened species of genus *Decalepis* (Apocynaceae). *Scientific reports*, 7(1):1-12.
- Mulcahy, D.G., Ibáñez, R., Jaramillo, C.A., Crawford, A.J., Ray, J.M., Gotte, S.W., Jacobs, J.F., Wynn, A.H., Gonzalez-Porter, G.P., McDiarmid, R.W. & Crombie, R.I. 2022. DNA barcoding of the National Museum of Natural History reptile tissue holdings raises concerns about the use of natural history

collections and the responsibilities of scientists in the molecular age. *PLos One*, 17(3). <https://doi.org/10.1371/journal.pone.0264930> Date of access: 26 Sept. 2022.

Murray, D.C., Coghlan, M.L. & Bunce, M. 2015. From benchtop to desktop: important considerations when designing amplicon sequencing workflows. *PLos One*, 10(4). <https://doi.org/10.1371/journal.pone.0124671> Date of access: 26 Sept. 2022.

Needham, D.M., Sachdeva, R. & Fuhrman, J.A. 2017. Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *The ISME journal*, 11(7):1614-1629.

Newmaster, S.G., Fazekas, A.J. & Ragupathy, S. 2006. DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. *Botany*, 84(3):335-341.

Nguyen, N.H., Smith, D., Peay, K. & Kennedy, P. 2015. Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytologist*, 205(4):1389-1393.

Nichols, R.V., Åkesson, M. & Kjellander, P. 2016. Diet assessment based on rumen contents: a comparison between DNA metabarcoding and macroscopy. *PLos One*, 11(6). <https://doi.org/10.1371/journal.pone.0157977> Date of access: 23 Jan. 2022.

Niemann, H.J., Bezeng, B.S., Orton, R.D., Kabongo, R.M., Pilusa, M. & Van der Bank, M. 2022. Using a DNA barcoding approach to facilitate biosecurity: Identifying invasive alien macrophytes traded within the South African aquarium and pond plant industry. *South African Journal of Botany*, 144:364-376.

Nobile, A.B., Freitas-Souza, D., Ruiz-Ruano, F.J., Nobile, M.L.M., Costa, G.O., De Lima, F.P., Camacho, J.P.M., Foresti, F. & Oliveira, C. 2019. DNA metabarcoding of Neotropical ichthyoplankton: Enabling high accuracy with lower cost. *Metabarcoding and Metagenomics*, 3. <https://doi.org/10.3897/mbmg.3.35060> Date of access: 23 Jan. 2022.

Odadi, W.O., Karachi, M.K., Abdulrazak, S.A. & Young, T.P. 2013. Protein supplementation reduces non-grass foraging by a primary grazer. *Ecological Applications*, 23(2):455-463.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H. & Oksanen, M.J. 2013. Package 'vegan'. *Community ecology package, version*, 2(9):1-295.

O'Rourke, D.R., Bokulich, N.A., Jusino, M.A., MacManes, M.D. & Foster, J.T. 2020. A total crapshoot? Evaluating bioinformatic decisions in animal diet metabarcoding analyses. *Ecology and Evolution*, 10(18):9721-9739.

- Owen-Smith, N. 1979. Assessing the foraging efficiency of a large herbivore, the kudu. *South African Journal of Wildlife Research*, 9(3):102-110.
- Pansu, J., Guyton, J.A., Potter, A.B., Atkins, J.L., Daskin, J.H., Wursten, B., Kartzinell, T.R. & Pringle, R.M. 2019. Trophic ecology of large herbivores in a reassembling African ecosystem. *Journal of Ecology*, 107(3):1355-1376.
- Paradis, E. & Schliep, K. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3):526-528.
- Pauvert, C., Buee, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., Lesur, I., Vallance, J. & Vacher, C. 2019. Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecology*, 41:23-33.
- Pavan-Kumar, A., Gireesh-Babu, P. & Lakra, W.S. 2015. DNA metabarcoding: a new approach for rapid biodiversity assessment. *Journal of Cell Science and Molecular Biology*, 2(1):111.
- Pečnikar, F.Ž. & Buzan, E.V. 2014. 20 years since the introduction of DNA barcoding: from theory to application. *Journal of applied genetics*, 55(1):43-52.
- Pegard, A., Miquel, C., Valentini, A., Coissac, E., Bouvier, F., Francois, D., Taberlet, P., Engel, E. & Pompanon, F. 2009. Universal DNA-based methods for assessing the diet of grazing livestock and wildlife from feces. *Journal of Agricultural and Food Chemistry*, 57(13):5700-5706. <https://doi.org/10.1021/jf803680c> Date of access: 1 Sep. 2020.
- Phefadu, K.C. & Kutu, F.R. 2016. Evaluation of spatial variability of soil physico-chemical characteristics on rhodic ferralsol at the Syferkuil Experimental Farm of University of Limpopo, South Africa. *Journal of Agricultural Science*, 8(10): 92-106.
- Piñol, J., Senar, M.A. & Symondson, W.O. 2019. The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Molecular ecology*, 28(2):407-419.
- Pompanon, F., Deagle, B.E., Symondson, W.O., Brown, D.S., Jarman, S.N. & Taberlet, P. 2012. Who is eating what: diet assessment using next generation sequencing. *Molecular ecology*, 21(8):1931-1950.
- Potter, A.B., Hutchinson, M.C., Pansu, J., Wursten, B., Long, R.A., Levine, J.M. & Pringle, R.M. 2022. Mechanisms of dietary resource partitioning in large-herbivore assemblages: A plant-trait-based approach. *Journal of Ecology*, 110(4):817-832.

- Pringle, R.M., Prior, K.M., Palmer, T.M., Young, T.P. & Goheen, J.R. 2016. Large herbivores promote habitat specialization and beta diversity of African savanna trees. *Ecology*, 97(10):2640-2657.
- Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P. & Gu, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics*, 13(1):1-13.
- Quéméré, E., Hibert, F., Miquel, C., Lhuillier, E., Rasolondraibe, E., Champeau, J., Rabarivola, C., Nusbaumer, L., Chatelain, C., Gautier, L. & Ranirison, P. 2013. A DNA metabarcoding study of a primate dietary diversity and plasticity across its entire fragmented range. *PLoS One*, 8(3).
<https://doi.org/10.1371/journal.pone.0058971> Date of access: 18 Nov. 2022.
- R Core Team. 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/> Date of access: 17 Mar. 2022.
- R.H. Whittaker, J.W. Morris & D. Goodman. 1984. Pattern analysis in savanna woodlands at Nylsvley, South Africa. *Memoirs of the Botanical Survey of South Africa*, 49:1-51
- Ratnasingham, S. & Hebert, P.D. 2007. BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular ecology notes*, 7(3):355-364.
- Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4. <https://doi.org/10.7717/peerj.2584> Date of access: 3 Jul. 2022.
- RStudio Team. 2022. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. <http://www.rstudio.com/> Date of access: 17 Mar. 2022.
- Rutherford, M.C., Mucina, L. & Powrie, L.W. 2006. Biomes and bioregions of southern Africa. In: Mucina, L. & Rutherford, M.C., eds. *The vegetation of South Africa, Lesotho and Swaziland*. Pretoria: South African National Biodiversity Institute. pp. 30-51.
- Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J. & Walker, A.W. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC biology*, 12(1):1-12.
- SANBI: South African National Biodiversity Institute. 2012. *Vegetation Map App* [Vector] 2012. Available from the Biodiversity GIS website, downloaded on 11 November 2022

- Sankaran, M., Hanan, N.P., Scholes, R.J., Ratnam, J., Augustine, D.J., Cade, B.S., Gignoux, J., Higgins, S.I., Le Roux, X., Ludwig, F. & Ardo, J. 2005. Determinants of woody cover in African savannas. *Nature*, 438(7069):846-849.
- Scasta, J.D., Jorns, T., Derner, J.D., Lake, S., Augustine, D.J., Windh, J.L. & Smith, T.L. 2019. Validation of DNA metabarcoding of fecal samples using cattle fed known rations. *Animal Feed Science and Technology*, 255. <https://doi.org/10.1016/j.anifeedsci.2019.114219> Date of access: 4 Aug. 2022.
- Scasta, J.D., Jorns, T., Derner, J.D., Stam, B., McClaren, M., Calkins, C. & Stewart, W. 2020. Toxic plants in sheep diets grazing extensive landscapes: Insights from Fecal DNA metabarcoding. *Livestock Science*, 236:104002. <https://doi.org/10.1016/j.livsci.2020.104002> Date of access: 26 Sept. 2022.
- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J. & Sahl, J.W. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*, 75(23):7537-7541.
- Scott-Shaw, R. & Morris, C.D. 2015. Grazing depletes forb species diversity in the mesic grasslands of KwaZulu-Natal, South Africa. *African Journal of Range & Forage Science*, 32(1): 21-31.
- Shackleton, C.M. 2000. Comparison of plant diversity in protected and communal lands in the Bushbuckridge lowveld savanna, South Africa. *Biological Conservation*, 94(3):273-285.
- Sheidai, M., Tabaripour, R., Talebi, S.M., Noormohammadi, Z. & Koohdar, F. 2019. Adulteration in medicinally important plant species of Ziziphora in Iran market: DNA barcoding approach. *Industrial Crops and Products*, 130:627-633.
- Shipley, L.A., Forbey, J.S. & Moore, B.D. 2009. Revisiting the dietary niche: when is a mammalian herbivore a specialist? *Integrative and comparative biology*, 49(3):274-290.
- Shuai, L.Y., Cao, C., Xin, X.P., Liu, Z.T. & Zeng, Z.G. 2019. Feeding preference and diet overlap at different taxonomic scales: the implications for coexistence of two small herbivores. *Journal of Mammalogy*, 100(2):544-551.
- Sickel, W., Ankenbrand, M.J., Grimmer, G., Holzschuh, A., Härtel, S., Lanzen, J., Steffan-Dewenter, I. & Keller, A. 2015. Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC ecology*, 15(1):1-9.
- Siebert, F. 2012. *A phytosociological synthesis of Mopanieveld vegetation at different spatial scales using various classification methods*. Potchefstroom: NWU. Potchefstroom: NWU. (Thesis – Phd).

- Siebert, F. & Dreber, N. 2019. Forb ecology research in dry African savannas: Knowledge, gaps, and future perspectives. *Ecology and Evolution*, 9(13): 7875-7891.
- Siebert, F. & Eckhardt, H.C. 2008. The vegetation and floristics of the Nkhuhlu exclosures, Kruger National Park. *Koedoe: African Protected Area Conservation and Science*, 50(1):126-144.
- Siebert, F. & Scogings, P. 2015. Browsing intensity of herbaceous forbs across a semi-arid savanna catenal sequence. *South African Journal of Botany*, 100:69-74.
- Siebert, F., Siebert, S.J. & Eckhardt, H.C. 2010. The vegetation and floristics of the Letaba exclosures, Kruger National Park, South Africa. *Koedoe: African Protected Area Conservation and Science*, 52(1):1-12.
- Silveira, F.A., Arruda, A.J., Bond, W., Durigan, G., Fidelis, A., Kirkman, K., Oliveira, R.S., Overbeck, G.E., Sansevero, J.B.B., Siebert, F. & Siebert, S.J. 2020. Myth-busting tropical grassy biome restoration. *Restoration Ecology*, 28(5):1067-1073.
- Soininen, E.M., Valentini, A., Coissac, E., Miquel, C., Gielly, L., Brochmann, C., Brysting, A.K., Sønstebo, J.H., Ims, R.A., Yoccoz, N.G. & Taberlet, P. 2009. Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in zoology*, 6(1):1-9.
- Stapleton, T.E., Weinstein, S.B., Greenhalgh, R. & Dearing, M.D. 2022. Successes and limitations of quantitative diet metabarcoding in a small, herbivorous mammal. *Molecular Ecology Resources*, 22(7):2573-2586.
- Taberlet, P., Bonin, A., Coissac, E. & Zinger, L. 2018. *Environmental DNA: For biodiversity research and monitoring*. United Kingdom: Oxford University Press.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular ecology*, 21(8):2045-2050.
- Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., Vermet, T., Corthier, G., Brochmann, C. & Willerslev, E. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic acids research*, 35(3). <https://doi.org/10.1093/nar/gkl938> Date of access: 17 Mar. 2020.
- Taberlet, P., Gielly, L., Pautou, G. & Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant molecular biology*, 17(5):1105-1109.

- Tan, G., Opitz, L., Schlapbach, R. & Rehrauer, H. 2019. Long fragments achieve lower base quality in Illumina paired-end sequencing. *Scientific reports*, 9(1):1-7.
- Tanabe, A.S. & Toju, H. 2013. Two new computational methods for universal DNA barcoding: a benchmark using barcode sequences of bacteria, archaea, animals, fungi, and land plants. *PLos One*, 8(10). <https://doi.org/10.1371/journal.pone.0076910> Date of access: 4 Aug. 2022.
- Ter Schure, A.T., Pillai, A.A., Thorbek, L., Bhavani Shankar, M., Puri, R., Ravikanth, G., de Boer, H.J. & Boessenkool, S. 2021. eDNA metabarcoding reveals dietary niche overlap among herbivores in an Indian wildlife sanctuary. *Environmental DNA*, 3(3):681-696.
- Thomas, A.C., Deagle, B.E., Eveson, J.P., Harsch, C.H. & Trites, A.W. 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. *Molecular ecology resources*, 16(3):714-726.
- Thomas, A.C., Jarman, S.N., Haman, K.H., Trites, A.W. & Deagle, B.E. 2014. Improving accuracy of DNA diet estimates using food tissue control materials and an evaluation of proxies for digestion bias. *Molecular ecology*, 23(15):3706-3718.
- Thomsen, P.F. & Willerslev, E. 2015. Environmental DNA—An emerging tool in conservation for monitoring past and present biodiversity. *Biological conservation*, 183:4-18.
- Timpano, E.K., Scheible, M.K. & Meiklejohn, K.A. 2020. Optimization of the second internal transcribed spacer (ITS2) for characterizing land plants from soil. *PLos One*, 15(4). <https://doi.org/10.1371/journal.pone.0231436> Date of access: 13 May 2022.
- Tsukamoto, Y., Yonezawa, S., Katayama, N. & Isagi, Y. 2021. Detection of endangered aquatic plants in rapid streams using environmental DNA. *Frontiers in Ecology and Evolution*, 8. <https://doi.org/10.3389/fevo.2020.622291> Date of access: 13 May 2022.
- Turunen, J., Mykrä, H., Elbrecht, V., Steinke, D., Braukmann, T. & Aroviita, J. 2021. The power of metabarcoding: Can we improve bioassessment and biodiversity surveys of stream macroinvertebrate communities? *Metabarcoding and Metagenomics*, 5:99-110.
- Valentini, A., Miquel, C., Nawaz, M.A., Bellemain, E.V.A., Coissac, E., Pompanon, F., Gielly, L., Cruaud, C., Nascetti, G., Wincker, P. & Swenson, J.E. 2009. New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources*, 9(1):51-60.
- Valentini, A., Pompanon, F. & Taberlet, P. 2008. DNA barcoding for ecologists. *Trends in ecology and evolution*, 24(2):110-117.

- Van Coller, H., Klem, J. & Siebert, F. 2021. Drought tolerant forb flora of a semi-arid protected savanna in the Lowveld of South Africa. *Bothalia-African Biodiversity & Conservation*, 51(1):1-6.
- Van Coller, H., Siebert, F. & Siebert, S.J. 2013. Herbaceous species diversity patterns across various treatments of herbivory and fire along the sodic zone of the Nkuhlu exclosures, Kruger National Park. *Koedoe: African Protected Area Conservation and Science*, 55(1):1-6.
- Van Coller, H., Siebert, F., Scogings, P.F. & Ellis, S. 2018. Herbaceous responses to herbivory, fire and rainfall variability differ between grasses and forbs. *South African Journal of Botany*, 119:94-103.
- Van der Merwe, J. & Marshal, J.P. 2012. Hierarchical resource selection by impala in a savanna environment. *Austral Ecology*, 37(3):401-412.
- Van Der Sluijs, J., Kaiser, M., Beder, S., Hosle, V., Kemelmajer de Carlucci, A., & Kinzig, A. 2005. *The precautionary principle*. Place de Fontenoy, Paris: Unesco.
- Weigand, H., Beermann, A.J., Čiampor, F., Costa, F.O., Csabai, Z., Duarte, S., Geiger, M.F., Grabowski, M., Rimet, F., Rulík, B. & Strand, M. 2019. DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment*, 678:499-524.
- Weiskopf, S.R., Kachel, S.M. & McCarthy, K.P. 2016. What are snow leopards really eating? Identifying bias in food-habit studies. *Wildlife Society Bulletin*, 40(2):233-240.
- Wessels, K.J., Prince, S.D., Frost, P.E. & Van Zyl, D. 2004. Assessing the effects of human-induced land degradation in the former homelands of northern South Africa with a 1 km AVHRR NDVI time-series. *Remote sensing of environment*, 91(1):47-67.
- Westcott, S.L. & Schloss, P.D. 2015. De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ*, 3. <https://doi.org/10.7717/peerj.1487> Date of access: 8 Jul. 2022.
- Wickham, H., Chang, W. & Wickham, M.H. 2016. Package ‘ggplot2’. Create elegant data visualisations using the grammar of graphics. *Version*, 2(1):1-189.
- Wilkinson, M.J., Szabo, C., Ford, C.S., Yarom, Y., Croxford, A.E., Camp, A. & Gooding, P. 2017. Replacing Sanger with Next Generation Sequencing to improve coverage and quality of reference DNA barcodes for plants. *Scientific reports*, 7(1):1-11.

- Willerslev, E., Davison, J., Moora, M., Zobel, M., Coissac, E., Edwards, M.E., Lorenzen, E.D., Vestergård, M., Gussarova, G., Haile, J. & Craine, J. 2014. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature*, 506(7486):47-51.
- Wright, E.S. 2016. Using DECIPHER v2. 0 to analyze big biological sequence data in R. *R Journal*, 8(1):352.
- Xiong, F., Shu, L., Gan, X., Zeng, H., He, S. & Peng, Z. 2022. Methodology for fish biodiversity monitoring with environmental DNA metabarcoding: the primers, databases and bioinformatic pipelines. *Water Biology and Security*, 1(1). <https://doi.org/10.1016/j.watbs.2022.100007> Date of access: 4 Aug. 2022.
- Yang, C., Wang, X., Miller, J.A., de Blécourt, M., Ji, Y., Yang, C., Harrison, R.D. & Douglas, W.Y. 2014. Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. *Ecological Indicators*, 46:379-389.
- Yoccoz, N.G., Bråthen, K.A., Gielly, L., Haile, J., Edwards, M.E., Goslar, T., von Stedingk, H., Brysting, A.K., Coissac, E., Pompanon, F. & Sørensen, J.H. 2012. DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular ecology*, 21(15):3647-3655.
- Zaloumis, N.P. & Bond, W.J. 2016. Reforestation or conservation? The attributes of old growth grasslands in South Africa. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1703). <https://doi.org/10.1098/rstb.2015.0310> Date of access: 14 Apr. 2022.
- Zhu, S., Liu, Q., Qiu, S., Dai, J. & Gao, X. 2022. DNA barcoding: an efficient technology to authenticate plant species of traditional Chinese medicine and recent advances. *Chinese Medicine*, 17(1):1-17

Annexures

Annexure 1: Supplementary tables supporting Chapter 2

Supplementary Table 1: Complete species list used in reference database.

Phylum	Class	Order	Family	Genus	Species
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	Asystasia	<i>Asystasia mysorensis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	Asystasia	<i>Asystasia retrocarpa</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	Barleria	<i>Barleria affinis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	Barleria	<i>Barleria crossandriiformis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Barleria</i>	<i>Barleria elegans</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	Barleria	<i>Barleria oxyphylla</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	Barleria	<i>Barleria senensis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Barleria</i>	<i>Barleria prionitis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Blepharis</i>	<i>Blepharis innocua</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Blepharis</i>	<i>Blepharis integrifolia</i>

Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Blepharis</i>	<i>Blepharis subvolubilis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Crabbea</i>	<i>Crabbea velutina</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Crossandra</i>	<i>Crossandra mucronata</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Dyschoriste</i>	<i>Dyschoriste fischeri</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Ecbolium</i>	<i>Ecbolium glabratum</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Elytraria</i>	<i>Elytraria acaulis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Hygrophila</i>	<i>Hygrophila auriculata</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Hypoestes</i>	<i>Hypoestes forskoolii</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Justicia</i>	<i>Justicia anagalloides</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Justicia</i>	<i>Justicia betonica</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Justicia</i>	<i>Justicia flava</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Justicia</i>	<i>Justicia matammensis</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Justicia</i>	<i>Justicia protracta</i>

Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Monechma</i>	<i>Monechma debile</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Peristrophe</i>	<i>Peristrophe cernua</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Rhinacanthus</i>	<i>Rhinacanthus xerophilus</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Ruellia</i>	<i>Ruellia patula</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Ruellia</i>	<i>Ruellia prostrata</i>
Tracheophyta	Magnoliopsida	Lamiales	Acanthaceae	<i>Thunbergia</i>	<i>Thunbergia alata</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Aizoaceae	<i>Aizoon</i>	<i>Aizoon canariense</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Aizoaceae	<i>Trianthema</i>	<i>Trianthema salsoloides</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Aizoaceae	<i>Trianthema</i>	<i>Trianthema triquetra</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Aizoaceae	<i>Zaleya</i>	<i>Zaleya pentandra</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Achyranthes</i>	<i>Achyranthes aspera</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Achyroopsis</i>	<i>Achyroopsis leptostachya</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Aerva</i>	<i>Aerva leucura</i>

Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Alternanthera</i>	<i>Alternanthera pungens</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Alternanthera</i>	<i>Alternanthera sessilis</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Amaranthus</i>	<i>Amaranthus hybridus</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Amaranthus</i>	<i>Amaranthus praetermissus</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Amaranthus</i>	<i>Amaranthus thunbergii</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Celosia</i>	<i>Celosia trigyna</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Chenopodium</i>	<i>Chenopodium ambrosioides</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Chenopodium</i>	<i>Chenopodium album</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Cyathula</i>	<i>Cyathula lanceolata</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Gomphrena</i>	<i>Gomphrena celosioides</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Guilleminea</i>	<i>Guilleminea densa</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Kyphocarpa</i>	<i>Kyphocarpa angustifolia</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Pupalia</i>	<i>Pupalia lappacea</i>

Tracheophyta	Magnoliopsida	Caryophyllales	Amaranthaceae	<i>Sericorema</i>	<i>Sericorema remotiflora</i>
Tracheophyta	Liliopsida	Asparagales	Amaryllidaceae	<i>Ammocharis</i>	<i>Ammocharis coranica</i>
Tracheophyta	Liliopsida	Asparagales	Amaryllidaceae	<i>Boophone</i>	<i>Boophone disticha</i>
Tracheophyta	Liliopsida	Asparagales	Amaryllidaceae	<i>Crinum</i>	<i>Crinum buphanoides</i>
Tracheophyta	Liliopsida	Asparagales	Amaryllidaceae	<i>Crinum</i>	<i>Crinum macowanii</i>
Tracheophyta	Liliopsida	Asparagales	Amaryllidaceae	<i>Crinum</i>	<i>Crinum moorei</i>
Tracheophyta	Liliopsida	Asparagales	Amaryllidaceae	<i>Tulbaghia</i>	<i>Tulbaghia leucantha</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Lannea</i>	<i>Lannea discolor</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Lannea</i>	<i>Lannea edulis</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Lannea</i>	<i>Lannea schweinfurthii</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa albicans</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa engleri</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa insignis</i>

Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa paniculosa</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Sclerocarya</i>	<i>Sclerocarya birrea</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Searsia</i>	<i>Searsia dentata</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Searsia</i>	<i>Searsia gueinzii</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Searsia</i>	<i>Searsia pentheri</i>
Tracheophyta	Magnoliopsida	Sapindales	Anacardiaceae	<i>Searsia</i>	<i>Searsia pyroides</i>
Tracheophyta	Magnoliopsida	Magnoliales	Annonaceae	<i>Hexalobus</i>	<i>Hexalobus monopetalus</i>
Tracheophyta	Magnoliopsida	Magnoliales	Annonaceae	<i>Monanthotaxis</i>	<i>Monanthotaxis caffra</i>
Tracheophyta	Magnoliopsida	Apiales	Apiaceae	<i>Centella</i>	<i>Centella asiatica</i>
Tracheophyta	Magnoliopsida	Apiales	Apiaceae	<i>Pappea</i>	<i>Pappea capensis</i>
Tracheophyta	Magnoliopsida	Apiales	Apiaceae	<i>Pappea</i>	<i>Pappea capensis</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Adenium</i>	<i>Adenium multiflorum</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Adenium</i>	<i>Adenium swazicum</i>

Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Carissa</i>	<i>Carissa bispinosa</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Carissa</i>	<i>Carissa edulis</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Catharanthus</i>	<i>Catharanthus roseus</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Cryptolepis</i>	<i>Cryptolepis oblongifolia</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Cynanchum</i>	<i>Cynanchum viminale</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Fockea</i>	<i>Fockea angustifolia</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Gomphocarpus</i>	<i>Gomphocarpus tomentosus</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Huernia</i>	<i>Huernia hystrix</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Pentarrhinum</i>	<i>Pentarrhinum insipidum</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Pergularia</i>	<i>Pergularia daemia</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Pergularia</i>	<i>Pergularia daemia</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Raphionacme</i>	<i>Raphionacme elata</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Riocreuxia</i>	<i>Riocreuxia torulosa</i>

Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Secamone</i>	<i>Secamone parvifolia</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Stapelia</i>	<i>Stapelia gigantea</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Stomatostemma</i>	<i>Stomatostemma monteiroae</i>
Tracheophyta	Magnoliopsida	Gentianales	Apocynaceae	<i>Tacazzea</i>	<i>Tacazzea apiculata</i>
Tracheophyta	Liliopsida	Alismatales	Araceae	<i>Stylochaeton</i>	<i>Stylochaeton natalensis</i>
Tracheophyta	Liliopsida	Arecales	Arecaceae	<i>Hyphaene</i>	<i>Hyphaene coriacea</i>
Tracheophyta	Liliopsida	Arecales	Arecaceae	<i>Phoenix</i>	<i>Phoenix reclinata</i>
Tracheophyta	Magnoliopsida	Piperales	Aristolochiaceae	<i>Aristolochia</i>	<i>Aristolochia elegans</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Albuca</i>	<i>Albuca setosa</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus africanus</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus burchellii</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus crassicladius</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus exuvialis</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus falcatus</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus plumosus</i>

Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus retrofractus</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Asparagus</i>	<i>Asparagus suaveolens</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Chlorophytum</i>	<i>Chlorophytum galpinii</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Dipcadi</i>	<i>Dipcadi viride</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Drimia</i>	<i>Drimia uniflora</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Eriospermum</i>	<i>Eriospermum flagelliforme</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Ledebouria</i>	<i>Ledebouria revoluta</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Ornithogalum</i>	<i>Ornithogalum seineri</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Sansevieria</i>	<i>Sansevieria hyacinthoides</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Sansevieria</i>	<i>Sansevieria pearsonii</i>
Tracheophyta	Liliopsida	Asparagales	Asparagaceae	<i>Urginea</i>	<i>Drimia altissima</i>
Tracheophyta	Liliopsida	Asparagales	Asphodelaceae	<i>Aloe</i>	<i>Aloe chabaudii</i>
Tracheophyta	Liliopsida	Asparagales	Asphodelaceae	<i>Aloe</i>	<i>Aloe marlothii</i>
Tracheophyta	Liliopsida	Asparagales	Asphodelaceae	<i>Bulbine</i>	<i>Bulbine abyssinica</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Acanthospermum</i>	<i>Acanthospermum hispidum</i>

Tracheophyta	Magnoliopsida	Asterales	Asteraceae	Ageratina	<i>Ageratina altissima</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Ageratum</i>	<i>Ageratum conyzoides</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Ambrosia</i>	<i>Ambrosia artemisiifolia</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Aspilia</i>	<i>Aspilia mossambicensis</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Baccharoides</i>	<i>Baccharoides adoensis</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Bidens</i>	<i>Bidens bipinnata</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Bidens</i>	<i>Bidens biternata</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Bidens</i>	<i>Bidens pilosa</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Cirsium</i>	<i>Cirsium vulgare</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Conyza</i>	<i>Conyza scabrida</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Dicoma</i>	<i>Dicoma galpinii</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Emilia</i>	<i>Emilia transvaalensis</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Felicia</i>	<i>Felicia bechuanica</i>

Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Geigeria</i>	<i>Geigeria burkei</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Geigeria</i>	<i>Geigeria ornativa</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Helichrysum</i>	<i>Helichrysum candolleanum</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Helichrysum</i>	<i>Helichrysum miconiifolium</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Hirpicium</i>	<i>Hirpicium bechuanense</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Nidorella</i>	<i>Nidorella resedifolia</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Parthenium</i>	<i>Parthenium hysterophorus</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Pegolettia</i>	<i>Pegolettia senegalensis</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Pluchea</i>	<i>Pluchea bojeri</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Pseudoconyza</i>	<i>Pseudoconyza viscosa</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Pseudognaphalium</i>	<i>Pseudognaphalium luteoalbum</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Schkuhria</i>	<i>Schkuhria pinnata</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Sphaeranthus</i>	<i>Sphaeranthus incisus</i>

Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Tagetes</i>	<i>Tagetes minuta</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Tridax</i>	<i>Tridax procumbens</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Vernonia</i>	<i>Vernonia oligocephala</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Vernonia</i>	<i>Vernonia glabra</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Vernonia</i>	<i>Vernonia fastigiata</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Vernonia</i>	<i>Vernonia poskeana</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Vernonia</i>	<i>Vernonia colorata</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Xanthium</i>	<i>Xanthium strumarium</i>
Tracheophyta	Magnoliopsida	Asterales	Asteraceae	<i>Zinnia</i>	<i>Zinnia peruviana</i>
Tracheophyta	Magnoliopsida	Lamiales	Bignoniaceae	<i>Kigelia</i>	<i>Kigelia africana</i>
Tracheophyta	Magnoliopsida	Lamiales	Bignoniaceae	<i>Rhigozum</i>	<i>Rhigozum zambesiaceum</i>
Tracheophyta	Magnoliopsida	Sapindales	Burseraceae	<i>Commiphora</i>	<i>Commiphora africana</i>
Tracheophyta	Magnoliopsida	Sapindales	Burseraceae	<i>Commiphora</i>	<i>Commiphora angolensis</i>

Tracheophyta	Magnoliopsida	Sapindales	Burseraceae	<i>Commiphora</i>	<i>Commiphora edulis</i>
Tracheophyta	Magnoliopsida	Sapindales	Burseraceae	<i>Commiphora</i>	<i>Commiphora glandulosa</i>
Tracheophyta	Magnoliopsida	Sapindales	Burseraceae	<i>Commiphora</i>	<i>Commiphora mollis</i>
Tracheophyta	Magnoliopsida	Sapindales	Burseraceae	<i>Commiphora</i>	<i>Commiphora pyracanthoides</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Cactaceae	<i>Cereus</i>	<i>Cereus peruvianus</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Cactaceae	<i>Opuntia</i>	<i>Opuntia ficus-indica</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Cactaceae	<i>Opuntia</i>	<i>Opuntia stricta</i>
Tracheophyta	Magnoliopsida	Asterales	Campanulaceae	<i>Wahlenbergia</i>	<i>Wahlenbergia krebsii</i>
Tracheophyta	Magnoliopsida	Asterales	Campanulaceae	<i>Wahlenbergia</i>	<i>Wahlenbergia undulata</i>
Tracheophyta	Magnoliopsida	Rosales	Cannabaceae	<i>Trema</i>	<i>Trema orientale</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Boscia</i>	<i>Boscia albitrunca</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Boscia</i>	<i>Boscia angustifolia</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Boscia</i>	<i>Boscia mossambicensis</i>

Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Capparis</i>	<i>Capparis sepiaria</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Capparis</i>	<i>Capparis tomentosa</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Capparis</i>	<i>Capparis tomentosa</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Maerua</i>	<i>Maerua angolensis</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Maerua</i>	<i>Maerua juncea</i>
Tracheophyta	Magnoliopsida	Brassicales	Capparaceae	<i>Maerua</i>	<i>Maerua parvifolia</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Caryophyllaceae	<i>Pollichia</i>	<i>Pollichia campestris</i>
Tracheophyta	Magnoliopsida	Celastrales	Celastraceae	<i>Elaeodendron</i>	<i>Elaeodendron transvaalense</i>
Tracheophyta	Magnoliopsida	Celastrales	Celastraceae	<i>Gymnosporia</i>	<i>Gymnosporia heterophylla</i>
Tracheophyta	Magnoliopsida	Celastrales	Celastraceae	<i>Gymnosporia</i>	<i>Gymnosporia polyacantha</i>
Tracheophyta	Magnoliopsida	Celastrales	Celastraceae	<i>Gymnosporia</i>	<i>Gymnosporia senegalensis</i>
Tracheophyta	Magnoliopsida	Celastrales	Celastraceae	<i>Hippocratea</i>	<i>Pristimera longipetiolata</i>
Tracheophyta	Magnoliopsida	Celastrales	Celastraceae	<i>Maytenus</i>	<i>Maytenus acuminata</i>

Tracheophyta	Magnoliosida	Celastrales	Celastraceae	<i>Mystroxydon</i>	<i>Mystroxydon aethiopicum</i>
Tracheophyta	Magnoliosida	Celastrales	Celastraceae	<i>Mystroxydon</i>	<i>Mystroxydon aethiopica</i>
Tracheophyta	Magnoliosida	Celastrales	Celastraceae	<i>Robsonodendron</i>	<i>Robsonodendron eucleiforme</i>
Tracheophyta	Magnoliosida	Brassicales	Cleomaceae	<i>Cleome</i>	<i>Cleome hirta</i>
Tracheophyta	Magnoliosida	Brassicales	Cleomaceae	<i>Cleome</i>	<i>Cleome monophylla</i>
Tracheophyta	Magnoliosida	Brassicales	Cleomaceae	<i>Gynandropsis</i>	<i>Gynandropsis gynandra</i>
Tracheophyta	Liliopsida	Liliales	Colchicaceae	<i>Camptorrhiza</i>	<i>Camptorrhiza strumosa</i>
Tracheophyta	Liliopsida	Liliales	Colchicaceae	<i>Gloriosa</i>	<i>Gloriosa superba</i>
Tracheophyta	Magnoliosida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum apiculatum</i>
Tracheophyta	Magnoliosida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum collinum</i>
Tracheophyta	Magnoliosida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum erythrophyllum</i>
Tracheophyta	Magnoliosida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum hereroense</i>
Tracheophyta	Magnoliosida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum imberbe</i>

Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum microphyllum</i>
Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum molle</i>
Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum mossambicense</i>
Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Combretum</i>	<i>Combretum zeyheri</i>
Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Terminalia</i>	<i>Terminalia phanerophlebia</i>
Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Terminalia</i>	<i>Terminalia prunioides</i>
Tracheophyta	Magnoliopsida	Myrtales	Combretaceae	<i>Terminalia</i>	<i>Terminalia sericea</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Commelina</i>	<i>Commelina africana</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Commelina</i>	<i>Commelina benghalensis</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Commelina</i>	<i>Commelina diffusa</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Commelina</i>	<i>Commelina eckloniana</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Commelina</i>	<i>Commelina erecta</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Commelina</i>	<i>Commelina forskaolii</i>

Tracheophyta	Liliopsida	Commelinales	Commelinaceae	Commelina	<i>Commelina subulata</i>
Tracheophyta	Liliopsida	Commelinales	Commelinaceae	<i>Murdannia</i>	<i>Murdannia simplex</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Convolvulus</i>	<i>Convolvulus farinosus</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Convolvulus</i>	<i>Convolvulus sagittatus</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Evolvulus</i>	<i>Evolvulus alsinoides</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea cairica</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea coptica</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea crassipes</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea lapathifolia</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea obscura</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea plebeia</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea robertsiana</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea sinensis</i>

Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea transvaalensis</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Jacquemontia</i>	<i>Jacquemontia tamnifolia</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Merremia</i>	<i>Merremia kentrocaulos</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Merremia</i>	<i>Merremia palmata</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Seddera</i>	<i>Seddera capensis</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Seddera</i>	<i>Seddera suffruticosa</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Seddera</i>	<i>Senecio speciosus</i>
Tracheophyta	Magnoliopsida	Solanales	Convolvulaceae	<i>Xenostegia</i>	<i>Xenostegia tridentata</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Corbichoniaceae	<i>Corbichonia</i>	<i>Corbichonia decumbens</i>
Tracheophyta	Magnoliopsida	Saxifragales	Crassulaceae	<i>Cotyledon</i>	<i>Cotyledon barbeyi</i>
Tracheophyta	Magnoliopsida	Saxifragales	Crassulaceae	<i>Kalanchoe</i>	<i>Kalanchoe paniculata</i>
Tracheophyta	Magnoliopsida	Cucurbitales	Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia adoensis</i>
Tracheophyta	Magnoliopsida	Cucurbitales	Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia rehmannii</i>

Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia sessilifolia</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	Cucumis	<i>Cucumis africanus</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis anguria</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis hirsutus</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis metuliferus</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis zeyheri</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Kedrostis</i>	<i>Kedrostis africana</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Kedrostis</i>	<i>Kedrostis foetidissima</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Kedrostis</i>	<i>Kedrostis leloja</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Lagenaria</i>	<i>Lagenaria sphaerica</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Momordica</i>	<i>Momordica balsamina</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Momordica</i>	<i>Momordica boivinii</i>
Tracheophyta	Magnolio psida	Cucurbitales	Cucurbitaceae	<i>Momordica</i>	<i>Momordica charantia</i>

Tracheophyta	Magnoliopsida	Cucurbitales	Cucurbitaceae	<i>Momordica</i>	<i>Momordica foetida</i>
Tracheophyta	Magnoliopsida	Cucurbitales	Cucurbitaceae	<i>Zehneria</i>	<i>Zehneria scabra</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Bulbostylis</i>	<i>Bulbostylis burchellii</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Bulbostylis</i>	<i>Bulbostylis hispidula</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus castaneus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus compressus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus cyperoides</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus esculentus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus indecorus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus obtusiflorus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus pseudovestitus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus rotundus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus rubicundus</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus rupestris</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus schinzii</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Cyperus</i>	<i>Cyperus sexangularis</i>

Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Fimbristylis</i>	<i>Fimbristylis complanata</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Rhynchospora</i>	<i>Rhynchospora brownii</i>
Tracheophyta	Liliopsida	Poales	Cyperaceae	<i>Schoenoplectus</i>	<i>Schoenoplectus corymbosus</i>
Tracheophyta	Liliopsida	Dioscoreales	Dioscoreaceae	<i>Dioscorea</i>	<i>Dioscorea cotinifolia</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Diospyros</i>	<i>Diospyros mespiliformis</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Euclea</i>	<i>Euclea crispa</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Euclea</i>	<i>Euclea divinorum</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Euclea</i>	<i>Euclea natalensis</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Euclea</i>	<i>Euclea schimperii</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Euclea</i>	<i>Euclea undulata</i>
Tracheophyta	Magnoliopsida	Ericales	Ebenaceae	<i>Royena</i>	<i>Royena lycioides</i>
Tracheophyta	Magnoliopsida	Boraginales	Ehretiaceae	<i>Cordia</i>	<i>Cordia caffra</i>
Tracheophyta	Magnoliopsida	Boraginales	Ehretiaceae	<i>Cordia</i>	<i>Cordia grandicalyx</i>

Tracheophyta	Magnoliopsida	Boraginales	Ehretiaceae	<i>Cordia</i>	<i>Cordia monoica</i>
Tracheophyta	Magnoliopsida	Boraginales	Ehretiaceae	<i>Ehretia</i>	<i>Ehretia amoena</i>
Tracheophyta	Magnoliopsida	Boraginales	Ehretiaceae	<i>Ehretia</i>	<i>Ehretia rigida</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	Acalypha	<i>Acalypha indica</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	Acalypha	<i>Acalypha villicaulis</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Croton</i>	<i>Croton gratissimus</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Croton</i>	<i>Croton megalobotrys</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	Dalechampia	<i>Dalechampia galpinii</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia inaequilatera</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia neopolycnemoides</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia prostrata</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia cyathophora</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia hirta</i>

Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia indica</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia ingens</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia neopolycnemoides</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia tirucalli</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Jatropha</i>	<i>Jatropha spicata</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Jatropha</i>	<i>Jatropha zeyheri</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Ricinus</i>	<i>Ricinus communis</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Spirostachys</i>	<i>Spirostachys africana</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Spirostachys</i>	<i>Spirostachys africana</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Tragia</i>	<i>Tragia dioica</i>
Tracheophyta	Magnoliopsida	Malpighiales	Euphorbiaceae	<i>Tragia</i>	<i>Tragia incisifolia</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Abrus</i>	<i>Abrus precatorius</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Acacia</i>	<i>Acacia borleae</i>

Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Acacia	<i>Acacia burkei</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Acacia	<i>Acacia caffra</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Acacia	<i>Acacia erubescens</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Acacia	<i>Acacia robusta</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Acacia	<i>Acacia senegal</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Acacia	<i>Acacia senegal</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	Aeschynomene	<i>Aeschynomene indica</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia anthelmintica</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia forbesii</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia harveyi</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia harveyi</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia petersiana</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia petersiana</i>

Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Albizia</i>	<i>Albizia versicolor</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Alysicarpus</i>	<i>Alysicarpus glumaceus</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Alysicarpus</i>	<i>Alysicarpus rugosus</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Bauhinia</i>	<i>Bauhinia galpinii</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Bolusanthus</i>	<i>Bolusanthus speciosus</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Cassia</i>	<i>Cassia abbreviata</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Chamaecrista</i>	<i>Chamaecrista absus</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Chamaecrista</i>	<i>Chamaecrista mimosoides</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Colophospermum</i>	<i>Colophospermum mopane</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Crotalaria</i>	<i>Crotalaria laburnifolia</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Crotalaria</i>	<i>Crotalaria monteiroi</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Crotalaria</i>	<i>Crotalaria podocarpa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Crotalaria</i>	<i>Crotalaria schinzii</i>

Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Crotalaria</i>	<i>Crotalaria sphaerocarpa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Crotalaria</i>	<i>Crotalaria virgulata</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Dalbergia</i>	<i>Dalbergia melanoxylon</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Dalbergia</i>	<i>Dalbergia melanoxylon</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Dichrostachys</i>	<i>Dichrostachys cinerea</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Dolichos</i>	<i>Dolichos junodii</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Dolichos</i>	<i>Dolichos trilobus</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Dumasia</i>	<i>Dumasia villosa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Erythrina</i>	<i>Erythrina humeana</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Erythrina</i>	<i>Erythrina latissima</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigastrum</i>	<i>Indigastrum costatum</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera arrecta</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera astragalina</i>

Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera bainesii</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera galpinii</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera heterotricha</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera lupatana</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera lydenburgensis</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera rhytidocarpa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera schimperi</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera spicata</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera tinctoria</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Indigofera</i>	<i>Indigofera viciodes</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Leobordea</i>	<i>Leobordea stipulosa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Macrotyloma</i>	<i>Macrotyloma maranguense</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Macrotyloma</i>	<i>Macrotyloma uniflorum</i>

Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Mundulea</i>	<i>Mundulea sericea</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Neorautanenia</i>	<i>Neorautanenia amboensis</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Ormocarpum</i>	<i>Ormocarpum trichocarpum</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Peltophorum</i>	<i>Peltophorum africanum</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Philenoptera</i>	<i>Philenoptera violacea</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Pterocarpus</i>	<i>Pterocarpus angolensis</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Pterocarpus</i>	<i>Pterocarpus rotundifolius</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia fleckii</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia minima</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia monophylla</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia totta</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Schotia</i>	<i>Schotia brachypetala</i>
Tracheophyta	Magnolio psida	Fabales	Fabaceae	<i>Schotia</i>	<i>Schotia brachypetala</i>

Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Schotia</i>	<i>Schotia capitata</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Schotia</i>	<i>Schotia capitata</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senegalia</i>	<i>Senegalia galpinii</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senegalia</i>	<i>Senegalia nigrescens</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senegalia</i>	<i>Senegalia schweinfurthii</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senegalia</i>	<i>Senegalia welwitschii</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senegalia</i>	<i>Senegalia mellifera</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senna</i>	<i>Senna bicapsularis</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senna</i>	<i>Senna didymobotrya</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senna</i>	<i>Senna italica</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senna</i>	<i>Senna obtusifolia</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senna</i>	<i>Senna occidentalis</i>
Tracheophyta	Magnoliosida	Fabales	Fabaceae	<i>Senna</i>	<i>Senna petersiana</i>

Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Sesbania</i>	<i>Sesbania bispinosa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Sphenostylis</i>	<i>Sphenostylis angustifolia</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Stylosanthes</i>	<i>Stylosanthes fruticosa</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tephrosia</i>	<i>Tephrosia longipes</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tephrosia</i>	<i>Tephrosia multijuga</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tephrosia</i>	<i>Tephrosia polystachya</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tephrosia</i>	<i>Tephrosia purpurea</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tephrosia</i>	<i>Tephrosia rhodesica</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tephrosia</i>	<i>Tephrosia uniflora</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Teramnus</i>	<i>Teramnus labialis</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Tylosema</i>	<i>Tylosema fassoglense</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia exuvialis</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia gerrardii</i>

Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia grandicornuta</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia karroo</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia luederitzii</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia nilotica</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia tortilis</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vachellia</i>	<i>Vachellia xanthophloea</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vigna</i>	<i>Vigna frutescens</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vigna</i>	<i>Vigna luteola</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vigna</i>	<i>Vigna unguiculata</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vigna</i>	<i>Vigna unguiculata</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Vigna</i>	<i>Vigna vexillata</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Xanthocercis</i>	<i>Xanthocercis zambesiaca</i>
Tracheophyta	Magnoliopsida	Fabales	Fabaceae	<i>Zornia</i>	<i>Zornia glochidiata</i>

Tracheophyta	Magnoliopsida	Gentianales	Gentianaceae	<i>Enicostema</i>	<i>Enicostema axillare</i>
Tracheophyta	Magnoliopsida	Geraniales	Geraniaceae	<i>Monsonia</i>	<i>Monsonia angustifolia</i>
Tracheophyta	Magnoliopsida	Geraniales	Geraniaceae	<i>Monsonia</i>	<i>Monsonia burkeana</i>
Tracheophyta	Magnoliopsida	Geraniales	Geraniaceae	<i>Monsonia</i>	<i>Monsonia emarginata</i>
Tracheophyta	Magnoliopsida	Geraniales	Geraniaceae	<i>Monsonia</i>	<i>Monsonia glauca</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Gisekiaceae	<i>Gisekia</i>	<i>Gisekia africana</i>
Tracheophyta	Magnoliopsida	Boraginales	Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium ciliatum</i>
Tracheophyta	Magnoliopsida	Boraginales	Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium ovalifolium</i>
Tracheophyta	Magnoliopsida	Boraginales	Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium steudneri</i>
Tracheophyta	Magnoliopsida	Boraginales	Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium strigosum</i>
Tracheophyta	Liliopsida	Asparagales	Hypoxidaceae	<i>Hypoxis</i>	<i>Hypoxis hemerocallidea</i>
Tracheophyta	Liliopsida	Asparagales	Iridaceae	<i>Lapeirousia</i>	<i>Lapeirousia sandersonii</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Kewaceae	<i>Kewa</i>	<i>Kewa bowkeriana</i>

Tracheophyta	Magnoliopsida	Caryophyllales	Kewaceae	<i>Kewa</i>	<i>Kewa salsoloides</i>
Tracheophyta	Magnoliopsida	Sapindales	Kirkiaceae	<i>Kirkia</i>	<i>Kirkia wilmsii</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Acrotome</i>	<i>Acrotome hispida</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Acrotome</i>	<i>Acrotome inflata</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Clerodendrum</i>	<i>Clerodendrum glabrum</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Clerodendrum</i>	<i>Clerodendrum ternatum</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Endostemon</i>	<i>Endostemon tereticaulis</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Leonotis</i>	<i>Leonotis nepetifolia</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Leonotis</i>	<i>Leonotis ocymifolia</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Leucas</i>	<i>Leucas capensis</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Leucas</i>	<i>Leucas glabrata</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Leucas</i>	<i>Leucas martinicensis</i>
Tracheophyta	Magnoliopsida	Lamiales	Lamiaceae	<i>Leucas</i>	<i>Leucas neuflizeana</i>

Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Leucas</i>	<i>Leucas sexdentata</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Ocimum</i>	<i>Ocimum filamentosum</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Ocimum</i>	<i>Ocimum americanum</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Ocimum</i>	<i>Ocimum gratissimum</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Ocimum</i>	<i>Ocimum labiatum</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Plectranthus</i>	<i>Plectranthus caninus</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Stachys</i>	<i>Stachys hyssopoides</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Syncolostemon</i>	<i>Syncolostemon elliottii</i>
Tracheophyta	Magnoliosida	Lamiales	Lamiaceae	<i>Tetradenia</i>	<i>Tetradenia riparia</i>
Tracheophyta	Magnoliosida	Laurales	Lauraceae	<i>Cassytha</i>	<i>Cassytha filiformis</i>
Tracheophyta	Magnoliosida	Caryophyllales	Limeaceae	<i>Limeum</i>	<i>Limeum fenestratum</i>
Tracheophyta	Magnoliosida	Caryophyllales	Limeaceae	<i>Limeum</i>	<i>Limeum viscosum</i>
Tracheophyta	Magnoliosida	Malpighiales	Linaceae	<i>Linum</i>	<i>Linum sulcatum</i>

Tracheophyta	Magnoliosida	Gentianales	Loganiaceae	<i>Strychnos</i>	<i>Strychnos madagascariensis</i>
Tracheophyta	Magnoliosida	Gentianales	Loganiaceae	<i>Strychnos</i>	<i>Strychnos spinosa</i>
Tracheophyta	Magnoliosida	Malpighiales	Malpighiaceae	<i>Sphedamnocarpus</i>	<i>Sphedamnocarpus pruriens</i>
Tracheophyta	Magnoliosida	Malpighiales	Malpighiaceae	<i>Triaspis</i>	<i>Triaspis hypericoides</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Abutilon</i>	<i>Abutilon angulatum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Abutilon</i>	<i>Abutilon austroafricanum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Abutilon</i>	<i>Abutilon grandiflorum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Abutilon</i>	<i>Abutilon mauritianum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Abutilon</i>	<i>Abutilon ramosum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Abutilon</i>	<i>Abutilon sonneratianum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Cienfuegosia</i>	<i>Cienfuegosia hildebrandtii</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Corchorus</i>	<i>Corchorus asplenifolius</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Corchorus</i>	<i>Corchorus confusus</i>

Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Corchorus</i>	<i>Corchorus trilocularis</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Dombeya</i>	<i>Dombeya rotundifolia</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Gossypium</i>	<i>Gossypium herbaceum</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia bicolor</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia flava</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia flavescens</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia hexamita</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia monticola</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia retinervis</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Grewia</i>	<i>Grewia villosa</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Hermannia</i>	<i>Hermannia glanduligera</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Hermannia</i>	<i>Hermannia modesta</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus calyphyllus</i>

Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus cannabinus</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus engleri</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus lunariifolius</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus micranthus</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus physaloides</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus praeteritus</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus pusillus</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus sabiensis</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus sidiformis</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus trionum</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus vitifolius</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Malvastrum</i>	<i>Malvastrum coromandelianum</i>
Tracheophyta	Magnoliopsida	Malvales	Malvaceae	<i>Melhania</i>	<i>Melhania acuminata</i>

Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Melhania</i>	<i>Melhania didyma</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Melhania</i>	<i>Melhania prostrata</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Pavonia</i>	<i>Pavonia burchellii</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Pavonia</i>	<i>Pavonia columella</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Sida</i>	<i>Sida alba</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Sida</i>	<i>Sida cordifolia</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Sida</i>	<i>Sida ovata</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Sida</i>	<i>Sida rhombifolia</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Sida</i>	<i>Sida spinosa</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Sterculia</i>	<i>Sterculia rogersii</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Triumfetta</i>	<i>Triumfetta annua</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Triumfetta</i>	<i>Triumfetta pentandra</i>
Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Triumfetta</i>	<i>Triumfetta rhomboidea</i>

Tracheophyta	Magnoliosida	Malvales	Malvaceae	<i>Waltheria</i>	<i>Waltheria indica</i>
Tracheophyta	Magnoliosida	Sapindales	Meliaceae	<i>Ekebergia</i>	<i>Ekebergia capensis</i>
Tracheophyta	Magnoliosida	Sapindales	Meliaceae	<i>Melia</i>	<i>Melia azedarach</i>
Tracheophyta	Magnoliosida	Sapindales	Meliaceae	<i>Trichilia</i>	<i>Trichilia emetica</i>
Tracheophyta	Magnoliosida	Sapindales	Meliaceae	<i>Turraea</i>	<i>Turraea obtusifolia</i>
Tracheophyta	Magnoliosida	Ranunculales	Menispermaceae	<i>Cissampelos</i>	<i>Cissampelos mucronata</i>
Tracheophyta	Magnoliosida	Ranunculales	Menispermaceae	<i>Cocculus</i>	<i>Cocculus hirsutus</i>
Tracheophyta	Magnoliosida	Caryophyllales	Molluginaceae	<i>Glinus</i>	<i>Glinus oppositifolius</i>
Tracheophyta	Magnoliosida	Caryophyllales	Molluginaceae	<i>Mollugo</i>	<i>Mollugo nudicaulis</i>
Tracheophyta	Magnoliosida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus abutilifolia</i>
Tracheophyta	Magnoliosida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus capreifolia</i>
Tracheophyta	Magnoliosida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus ingens</i>
Tracheophyta	Magnoliosida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus salicifolia</i>

Tracheophyta	Magnoliopsida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus sur</i>
Tracheophyta	Magnoliopsida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus sycomorus</i>
Tracheophyta	Magnoliopsida	Rosales	Moraceae	<i>Ficus</i>	<i>Ficus thonningii</i>
Tracheophyta	Magnoliopsida	Fagales	Myricaceae	<i>Morella</i>	<i>Morella serrata</i>
Tracheophyta	Magnoliopsida	Myrtales	Myrtaceae	<i>Heteropyxis</i>	<i>Heteropyxis natalensis</i>
Tracheophyta	Magnoliopsida	Myrtales	Myrtaceae	<i>Psidium</i>	<i>Psidium guajava</i>
Tracheophyta	Magnoliopsida	Myrtales	Myrtaceae	<i>Syzygium</i>	<i>Syzygium guineense</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Nyctaginaceae	<i>Boerhavia</i>	<i>Boerhavia diffusa</i>
Tracheophyta	Magnoliopsida	Malpighiales	Ochnaceae	<i>Ochna</i>	<i>Ochna inermis</i>
Tracheophyta	Magnoliopsida	Malpighiales	Ochnaceae	<i>Ochna</i>	<i>Ochna natalitia</i>
Tracheophyta	Magnoliopsida	Malpighiales	Ochnaceae	<i>Ochna</i>	<i>Ochna natalitia</i>
Tracheophyta	Magnoliopsida	Lamiales	Oleaceae	<i>Jasminum</i>	<i>Jasminum fluminense</i>
Tracheophyta	Magnoliopsida	Lamiales	Oleaceae	<i>Jasminum</i>	<i>Jasminum stenolobum</i>

Tracheophyta	Magnoliopsida	Myrtales	Onagraceae	<i>Ludwigia</i>	<i>Ludwigia adscendens</i>
Tracheophyta	Magnoliopsida	Myrtales	Onagraceae	<i>Ludwigia</i>	<i>Ludwigia octovalvis</i>
Tracheophyta	Magnoliopsida	Myrtales	Onagraceae	<i>Ludwigia</i>	<i>Ludwigia stolonifera</i>
Tracheophyta	Liliopsida	Asparagales	Orchidaceae	<i>Ansellia</i>	<i>Ansellia africana</i>
Tracheophyta	Magnoliopsida	Lamiales	Orobanchaceae	<i>Striga</i>	<i>Striga bilabiata</i>
Tracheophyta	Magnoliopsida	Lamiales	Orobanchaceae	<i>Striga</i>	<i>Striga elegans</i>
Tracheophyta	Magnoliopsida	Lamiales	Orobanchaceae	<i>Striga</i>	<i>Striga gesnerioides</i>
Tracheophyta	Magnoliopsida	Oxalidales	Oxalidaceae	<i>Oxalis</i>	<i>Oxalis latifolia</i>
Tracheophyta	Magnoliopsida	Oxalidales	Oxalidaceae	<i>Oxalis</i>	<i>Oxalis obliquifolia</i>
Tracheophyta	Magnoliopsida	Oxalidales	Oxalidaceae	<i>Oxalis</i>	<i>Oxalis smithiana</i>
Tracheophyta	Magnoliopsida	Ranunculales	Papaveraceae	<i>Argemone</i>	<i>Argemone mexicana</i>
Tracheophyta	Magnoliopsida	Malpighiales	Passifloraceae	<i>Adenia</i>	<i>Adenia digitata</i>
Tracheophyta	Magnoliopsida	Malpighiales	Passifloraceae	<i>Tricliceras</i>	<i>Tricliceras glanduliferum</i>

Tracheophyta	Magnoliopsida	Malpighiales	Passifloraceae	<i>Tricliceras</i>	<i>Tricliceras laceratum</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Ceratotheca</i>	<i>Ceratotheca triloba</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Dicerocaryum</i>	<i>Dicerocaryum eriocarpum</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Dicerocaryum</i>	<i>Dicerocaryum senecioides</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Dicerocaryum</i>	<i>Dicerocaryum senecioides</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Harpagophytum</i>	<i>Harpagophytum zeyheri</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Holubia</i>	<i>Holubia saccata</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Pterodiscus</i>	<i>Pterodiscus luridus</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Sesamum</i>	<i>Sesamum alatum</i>
Tracheophyta	Magnoliopsida	Lamiales	Pedaliaceae	<i>Sesamum</i>	<i>Sesamum triphyllum</i>
Tracheophyta	Magnoliopsida	Malpighiales	Peraceae	<i>Clutia</i>	<i>Clutia pulchella</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Bridelia</i>	<i>Bridelia cathartica</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Bridelia</i>	<i>Bridelia mollis</i>

Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Flueggea</i>	<i>Flueggea virosa</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus asperulatus</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus incurvus</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus maderaspatensis</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus nummulariifolius</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus pentandrus</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus reticulatus</i>
Tracheophyta	Magnoliopsida	Malpighiales	Phyllanthaceae	<i>Thecacoris</i>	<i>Acalypha glabrata</i>
Tracheophyta	Magnoliopsida	Piperales	Piperaceae	<i>Peperomia</i>	<i>Peperomia blanda</i>
Tracheophyta	Magnoliopsida	Lamiales	Plantaginaceae	<i>Bacopa</i>	<i>Bacopa floribunda</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Plumbaginaceae	<i>Plumbago</i>	<i>Plumbago zeylanica</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Andropogon</i>	<i>Andropogon chinensis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Andropogon</i>	<i>Andropogon gayanus</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida adscensionis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida bipartita</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida canescens</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida congesta</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida meridionalis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida mollissima</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida scabrivalvis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida stipitata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida stipitata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Aristida</i>	<i>Aristida vestita</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa bladhii</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa insculpta</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa radicans</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Brachiaria</i>	<i>Brachiaria eruciformis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Brachiaria</i>	<i>Brachiaria nigropedata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Brachiaria</i>	<i>Brachiaria xantholeuca</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Brachiaria</i>	<i>Brachiaria deflexa</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Brachiaria</i>	<i>Brachiaria serrata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Cenchrus</i>	<i>Cenchrus ciliaris</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Chloris</i>	<i>Chloris roxburghiana</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Chloris</i>	<i>Chloris gayana</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Chloris</i>	<i>Chloris pycnothrix</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Chloris</i>	<i>Chloris virgata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Cymbopogon</i>	<i>Cymbopogon excavatus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Cymbopogon</i>	<i>Cymbopogon plurinodis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Cynodon</i>	<i>Cynodon dactylon</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Dactyloctenium</i>	<i>Dactyloctenium aegyptium</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Dactyloctenium</i>	<i>Dactyloctenium australe</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Dactyloctenium</i>	<i>Dactyloctenium giganteum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Dichanthium</i>	<i>Dichanthium annulatum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Digitaria</i>	<i>Digitaria debilis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Digitaria</i>	<i>Digitaria eriantha</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Digitaria</i>	<i>Digitaria ternata</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Digitaria</i>	<i>Digitaria velutina</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Diheteropogon</i>	<i>Diheteropogon amplectens</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Diplachne</i>	<i>Diplachne fusca</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Echinochloa</i>	<i>Echinochloa colona</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Elionurus</i>	<i>Elionurus muticus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Enneapogon</i>	<i>Enneapogon cenchroides</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Enneapogon</i>	<i>Enneapogon scoparius</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Enteropogon</i>	<i>Enteropogon macrostachyus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis aspera</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis biflora</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis capensis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis chloromelas</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis cilianensis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis curvula</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis cylindriflora</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis gummiiflua</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis heteromera</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis inamoena</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis lehmanniana</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis nindensis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis pallens</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis racemosa</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis rigidior</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis rotifer</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis stapfii</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis superba</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eragrostis</i>	<i>Eragrostis trichophora</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eriochloa</i>	<i>Eriochloa meyeriana</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Eustachys</i>	<i>Eustachys paspaloides</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Fingerhuthia</i>	<i>Fingerhuthia africana</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Hemarthria</i>	<i>Hemarthria altissima</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Heteropogon</i>	<i>Heteropogon contortus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Hyparrhenia</i>	<i>Hyparrhenia filipendula</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Hyperthelia</i>	<i>Hyperthelia dissoluta</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Megathyrsus</i>	<i>Megathyrsus maximus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Melinis</i>	<i>Melinis nerviglumis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Melinis</i>	<i>Melinis repens</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Microchloa</i>	<i>Microchloa caffra</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Oropetium</i>	<i>Oropetium capense</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Panicum</i>	<i>Panicum coloratum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Panicum</i>	<i>Panicum deustum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Panicum</i>	<i>Panicum infestum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Panicum</i>	<i>Panicum maximum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Panicum</i>	<i>Panicum natalense</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Paspalum</i>	<i>Paspalum distichum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Paspalum</i>	<i>Paspalum scrobiculatum</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Perotis</i>	<i>Perotis patens</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Phragmites</i>	<i>Phragmites australis</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Phragmites</i>	<i>Phragmites mauritanus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Pogonarthria</i>	<i>Pogonarthria squarrosa</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Pogonarthria</i>	<i>Pogonarthria squarrosa</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Schmidtia</i>	<i>Schmidtia pappophoroides</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Setaria</i>	<i>Setaria incrassata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Setaria</i>	<i>Setaria pumila</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Setaria</i>	<i>Setaria sagittifolia</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Setaria</i>	<i>Setaria sphacelata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Setaria</i>	<i>Setaria verticillata</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sorghum</i>	<i>Sorghum bicolor</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sorghum</i>	<i>Sorghum bicolor</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sorghum</i>	<i>Sorghum versicolor</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus africanus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus festivus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus fimbriatus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus ioclados</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus nitens</i>

Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus pectinatus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus pyramidalis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Sporobolus</i>	<i>Sporobolus stapfianus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Stipagrostis</i>	<i>Stipagrostis hirtigluma</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Themeda</i>	<i>Themeda triandra</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Trachypogon</i>	<i>Trachypogon spicatus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Tragus</i>	<i>Tragus berteronianus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Tricholaena</i>	<i>Tricholaena monachne</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Trichoneura</i>	<i>Trichoneura grandiglumis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Tripogon</i>	<i>Tripogon minimus</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Urochloa</i>	<i>Urochloa brachyura</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Urochloa</i>	<i>Urochloa mosambicensis</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Urochloa</i>	<i>Urochloa oligotricha</i>
Tracheophyta	Liliopsida	Poales	Poaceae	<i>Urochloa</i>	<i>Urochloa panicoides</i>

Tracheophyta	Magnoliopsida	Fabales	Polygalaceae	<i>Polygala</i>	<i>Polygala erioptera</i>
Tracheophyta	Magnoliopsida	Fabales	Polygalaceae	<i>Polygala</i>	<i>Polygala producta</i>
Tracheophyta	Magnoliopsida	Fabales	Polygalaceae	<i>Polygala</i>	<i>Polygala sphenoptera</i>
Tracheophyta	Magnoliopsida	Fabales	Polygalaceae	<i>Polygala</i>	<i>Polygala uncinata</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Polygonaceae	<i>Oxygonum</i>	<i>Oxygonum dregeanum</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Polygonaceae	<i>Oxygonum</i>	<i>Oxygonum sinuatum</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Polygonaceae	<i>Persicaria</i>	<i>Persicaria decipiens</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Portulacaceae	<i>Portulaca</i>	<i>Portulaca hereroensis</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Portulacaceae	<i>Portulaca</i>	<i>Portulaca kermesina</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Portulacaceae	<i>Portulaca</i>	<i>Portulaca oleracea</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Portulacaceae	<i>Portulaca</i>	<i>Portulaca quadrifida</i>
Tracheophyta	Polypodiopsida	Polypodiales	Pteridaceae	<i>Cheilanthes</i>	<i>Cheilanthes viridis</i>
Tracheophyta	Polypodiopsida	Polypodiales	Pteridaceae	<i>Pellaea</i>	<i>Pellaea calomelanos</i>

Tracheophyta	Magnoliosida	Malpighiales	Putranjivaceae	<i>Drypetes</i>	<i>Drypetes gerrardii</i>
Tracheophyta	Magnoliosida	Ranunculales	Ranunculaceae	<i>Clematis</i>	<i>Clematis brachiata</i>
Tracheophyta	Magnoliosida	Rosales	Rhamnaceae	<i>Berchemia</i>	<i>Berchemia discolor</i>
Tracheophyta	Magnoliosida	Rosales	Rhamnaceae	<i>Berchemia</i>	<i>Berchemia zeyheri</i>
Tracheophyta	Magnoliosida	Rosales	Rhamnaceae	<i>Ziziphus</i>	<i>Ziziphus mucronata</i>
Tracheophyta	Magnoliosida	Rosales	Rhamnaceae	<i>Ziziphus</i>	<i>Ziziphus mucronata</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Agathisanthemum</i>	<i>Agathisanthemum bojeri</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Agathisanthemum</i>	<i>Agathisanthemum bojeri</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Breonadia</i>	<i>Breonadia salicina</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Canthium</i>	<i>Canthium armatum</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Canthium</i>	<i>Canthium ciliatum</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Catunaregam</i>	<i>Catunaregam spinosa</i>
Tracheophyta	Magnoliosida	Gentianales	Rubiaceae	<i>Catunaregam</i>	<i>Catunaregam spinosa</i>

Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Conostomium</i>	<i>Conostomium natalense</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Gardenia</i>	<i>Gardenia volkensii</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Gardenia</i>	<i>Gardenia volkensii</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Hyperacanthus</i>	<i>Hyperacanthus amoenus</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kohautia</i>	<i>Kohautia amatymbica</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kohautia</i>	<i>Kohautia caespitosa</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kohautia</i>	<i>Kohautia caespitosa</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kohautia</i>	<i>Kohautia cynanchica</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kohautia</i>	<i>Kohautia cynanchica</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kohautia</i>	<i>Kohautia virgata</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kraussia</i>	<i>Kraussia floribunda</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Kraussia</i>	<i>Kraussia floribunda</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Oldenlandia</i>	<i>Oldenlandia herbacea</i>

Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Pavetta</i>	<i>Pavetta catophylla</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Pavetta</i>	<i>Pavetta gardeniifolia</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Pavetta</i>	<i>Pavetta lanceolata</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Pavetta</i>	<i>Pavetta schumanniana</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Pentodon</i>	<i>Pentodon pentandrus</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Psydrax</i>	<i>Psydrax obovata</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Pyrostria</i>	<i>Pyrostria hystrix</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Richardia</i>	<i>Richardia brasiliensis</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Spermacoce</i>	<i>Spermacoce senensis</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Tricalysia</i>	<i>Tricalysia junodii</i>
Tracheophyta	Magnoliales	Gentianales	Rubiaceae	<i>Vangueria</i>	<i>Vangueria infausta</i>
Tracheophyta	Magnoliales	Sapindales	Rutaceae	<i>Clausena</i>	<i>Clausena anisata</i>
Tracheophyta	Magnoliales	Sapindales	Rutaceae	<i>Ptaeroxylon</i>	<i>Ptaeroxylon obliquum</i>

Tracheophyta	Magnoliosida	Sapindales	Rutaceae	<i>Vepris</i>	<i>Vepris carringtoniana</i>
Tracheophyta	Magnoliosida	Sapindales	Rutaceae	<i>Zanthoxylum</i>	<i>Zanthoxylum humile</i>
Tracheophyta	Magnoliosida	Malpighiales	Salicaceae	<i>Dovyalis</i>	<i>Dovyalis caffra</i>
Tracheophyta	Magnoliosida	Malpighiales	Salicaceae	<i>Flacourtia</i>	<i>Flacourtia indica</i>
Tracheophyta	Magnoliosida	Malpighiales	Salicaceae	<i>Scolopia</i>	<i>Scolopia zeyheri</i>
Tracheophyta	Magnoliosida	Sapindales	Sapindaceae	<i>Allophylus</i>	<i>Allophylus decipiens</i>
Tracheophyta	Magnoliosida	Sapindales	Sapindaceae	<i>Cardiospermum</i>	<i>Cardiospermum corindum</i>
Tracheophyta	Magnoliosida	Sapindales	Sapindaceae	<i>Cardiospermum</i>	<i>Cardiospermum corindum</i>
Tracheophyta	Magnoliosida	Sapindales	Sapindaceae	<i>Cardiospermum</i>	<i>Cardiospermum halicacabum</i>
Tracheophyta	Magnoliosida	Ericales	Sapotaceae	<i>Manilkara</i>	<i>Manilkara mocharia</i>
Tracheophyta	Magnoliosida	Lamiales	Scrophulariaceae	<i>Aptosimum</i>	<i>Aptosimum lineare</i>
Tracheophyta	Magnoliosida	Solanales	Solanaceae	<i>Datura</i>	<i>Datura stramonium</i>
Tracheophyta	Magnoliosida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum aculeastrum</i>

Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum catombelense</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum delagoense</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum incanum</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum lichtensteinii</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum nigrum</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum panduriforme</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum seafortianum</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Solanum</i>	<i>Solanum tettense</i>
Tracheophyta	Magnoliopsida	Solanales	Solanaceae	<i>Withania</i>	<i>Withania somnifera</i>
Tracheophyta	Magnoliopsida	Lamiales	Stilbaceae	<i>Nuxia</i>	<i>Nuxia oppositifolia</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Talinaceae	<i>Talinum</i>	<i>Talinum arnotii</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Talinaceae	<i>Talinum</i>	<i>Talinum caffrum</i>
Tracheophyta	Magnoliopsida	Caryophyllales	Talinaceae	<i>Talinum</i>	<i>Talinum portulacifolium</i>

Tracheophyta	Magnoliopsida	Caryophyllales	Talinaceae	<i>Talinum</i>	<i>Talinum tenuissimum</i>
Tracheophyta	Magnoliopsida	Malvales	Thymelaeaceae	<i>Gnidia</i>	<i>Gnidia rubescens</i>
Tracheophyta	Magnoliopsida	Rosales	Urticaceae	<i>Pouzolzia</i>	<i>Pouzolzia mixta</i>
Tracheophyta	Magnoliopsida	Vahliales	Vahliaceae	<i>Vahlia</i>	<i>Vahlia capensis</i>
Tracheophyta	Magnoliopsida	Vitales	Vataceae	<i>Rhoicissus</i>	<i>Rhoicissus revoilii</i>
Tracheophyta	Magnoliopsida	Vitales	Vataceae	<i>Rhoicissus</i>	<i>Rhoicissus tridentata</i>
Tracheophyta	Liliopsida	Pandanales	Velloziaceae	<i>Xerophyta</i>	<i>Xerophyta retinervis</i>
Tracheophyta	Liliopsida	Pandanales	Velloziaceae	<i>Xerophyta</i>	<i>Xerophyta retinervis</i>
Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Chascanum</i>	<i>Chascanum pinnatifidum</i>
Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Lantana</i>	<i>Lantana camara</i>
Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Lantana</i>	<i>Lantana rugosa</i>
Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Lippia</i>	<i>Lippia javanica</i>
Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Priva</i>	<i>Priva cordifolia</i>
Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Priva</i>	<i>Priva meyeri</i>

Tracheophyta	Magnoliopsida	Lamiales	Verbenaceae	<i>Verbena</i>	<i>Verbena bonariensis</i>
Tracheophyta	Magnoliopsida	Malpighiales	Violaceae	<i>Hybanthus</i>	<i>Hybanthus enneaspermus</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cissus</i>	<i>Cissus cornifolia</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cissus</i>	<i>Cissus quadrangularis</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cissus</i>	<i>Cissus quadrangularis</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cissus</i>	<i>Cissus rotundifolia</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma cirrhosum</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma puberulum</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma schlechteri</i>
Tracheophyta	Magnoliopsida	Vitales	Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma woodii</i>
Tracheophyta	Magnoliopsida	Santalales	Ximeniaceae	<i>Ximenia</i>	<i>Ximenia americana</i>
Tracheophyta	Magnoliopsida	Santalales	Ximeniaceae	<i>Ximenia</i>	<i>Ximenia caffra</i>
Tracheophyta	Magnoliopsida	Zygophyllales	Zygophyllaceae	<i>Balanites</i>	<i>Balanites maughamii</i>

Tracheophyta	Magnoliopsida	Zygophyllales	Zygophyllaceae	<i>Tribulus</i>	<i>Tribulus terrestris</i>
Tracheophyta	Magnoliopsida	Zygophyllales	Zygophyllaceae	<i>Tribulus</i>	<i>Tribulus terrestris</i>
Tracheophyta	Magnoliopsida	Zygophyllales	Zygophyllaceae	<i>Tribulus</i>	<i>Tribulus zeyheri</i>

Supplementary Table 2: Plants from which leaf material was collected between 26 March and 6 April 2015.

Family	Species
Malvaceae	<i>Abutilon austroafricanum</i>
Euphorbiaceae	<i>Acalypha indica</i>
Apocynaceae	<i>Adenium multiflorum</i>
Amaranthaceae	<i>Aerva leucura</i>
Rubiaceae	<i>Agathisanthemum bojeri</i>
Scrophulariaceae	<i>Aptosimum lineare</i>
Acanthaceae	<i>Barleria elegans</i>
Acanthaceae	<i>Barleria prionitoides</i>
Asteraceae	<i>Bidens bipinnata*</i>
Acanthaceae	<i>Blepharis innocua</i>
Acanthaceae	<i>Blepharis integrifolia</i>
Pedaliaceae	<i>Ceratotheca triloba</i>
Fabaceae	<i>Chamaecrista mimosoides</i>

Combretaceae	<i>Combretum zeyheri</i>
Burseraceae	<i>Commiphora africana</i>
Acanthaceae	<i>Crabbea velutina</i>
Acanthaceae	<i>Ecbolium glabratum</i>
Celastraceae	<i>Elaeodendron transvaalense</i>
Convolvulaceae	<i>Evolvulus alsinoides</i>
Euphorbiaceae	<i>Flueggea virosa</i>
Asteraceae	<i>Geigeria burkei</i>
Amaranthaceae	<i>Gomphrena celosiodes*</i>
Malvaceae	<i>Grewia bicolor</i>
Boraginaceae	<i>Heliotropium ovalifolium</i>
Malvaceae	<i>Hibiscus sabiensis</i>
Molluginaceae	<i>Hypertelis bowkeriana</i>
Fabaceae	<i>Indigofera tinctoria</i>
Convolvulaceae	<i>Ipomoea crassipes</i>
Acanthaceae	<i>Justicia flava</i>
Acanthaceae	<i>Justicia protracta</i>
Lamiaceae	<i>Leucas glabrata</i>
Sterculiaceae	<i>Melhania prostrata</i>
Sterculiaceae	<i>Melhania prostrata</i>
Molluginaceae	<i>Mollugo nudicaulis</i>

Phyllanthaceae	<i>Phyllanthus asperulatus</i>
Phyllanthaceae	<i>Phyllanthus incurvus</i>
Portulacaceae	<i>Portulaca hereroensis</i>
Portulacaceae	<i>Portulaca quadrifida</i>
Bignoniaceae	<i>Rhigozum zambesiaceum</i>
Fabaceae	<i>Rhynchosia confusa</i>
Convolvulaceae	<i>Seddera capensis</i>
Solanaceae	<i>Solanum delagoense</i>
Fabaceae	<i>Tephrosia purpurea</i>
Euphorbiaceae	<i>Tragia dioca</i>
Tiliaceae	<i>Triumfetta pentandra</i>
Sterculiaceae	<i>Waltheria indica</i> *
Convolvulaceae	<i>Xenostegia tridentata</i>
Asteraceae	<i>Zinnia peruviana</i> *

Annexure 2: Supplementary tables supporting Chapter 3

Supplementary Table 3: Taxonomy and accession numbers (GenBank) and/or Process ID's (BOLD) included in the rbcL and trnL reference datasets.

Family	Genus	Species	Barcode	Accession/Process ID		
Acanthaceae	<i>Asystasia</i>	<i>Asystasia mysorensis</i>	trnL	MK261330	MK261517	
Acanthaceae	<i>Asystasia</i>	<i>Asystasia mysorensis</i>	rbcL	MK285106		
Acanthaceae	<i>Asystasia</i>	<i>Asystasia sp.</i>	trnL	KR738111	KR737908	KR737825
Acanthaceae	<i>Asystasia</i>	<i>Asystasia sp.</i>	rbcL	KR736918	KR736699	KR736619
Acanthaceae	<i>Barleria</i>	<i>Barleria elegans</i>	trnL	MZ461549		
Acanthaceae	<i>Barleria</i>	<i>Barleria elegans</i>	rbcL	MZ461573		
Acanthaceae	<i>Barleria</i>	<i>Barleria prionitis</i>	trnL	KT075002	MZ461550	
Acanthaceae	<i>Barleria</i>	<i>Barleria prionitis</i>	rbcL	JQ231001	MZ461574	
Acanthaceae	<i>Blepharis</i>	<i>Blepharis innocua</i>	trnL	MZ461551		

Acanthaceae	<i>Blepharis</i>	<i>Blepharis innocua</i>	rbcL	MZ46157 5		
Acanthaceae	<i>Blepharis</i>	<i>Blepharis integrifolia</i>	trnL	DQ05486 1	DQ05486 0	
Acanthaceae	<i>Blepharis</i>	<i>Blepharis integrifolia</i>	rbcL	MZ46157 6		
Acanthaceae	<i>Blepharis</i>	<i>Blepharis subvolubilis</i>	trnL	DQ05486 4		
Acanthaceae	<i>Crabbea</i>	<i>Crabbea velutina</i>	trnL	MK26164 9	MK26168 2	MZ46155 5
Acanthaceae	<i>Crabbea</i>	<i>Crabbea velutina</i>	rbcL	MZ46157 9		
Acanthaceae	<i>Ecboium</i>	<i>Ecboium glabratum</i>	trnL	MZ46155 6		
Acanthaceae	<i>Ecboium</i>	<i>Ecboium glabratum</i>	rbcL	MZ46158 0		
Acanthaceae	<i>Elytraria</i>	<i>Elytraria acaulis</i>	rbcL	GENG10 64- 15.rbcL	GENG108 5-15.rbcL	
Acanthaceae	<i>Hygrophila</i>	<i>Hygrophila auriculata</i>	rbcL	GENG23 5-14.rbcL	TCHDC0 29- 19.rbcLa	TRM288- 15.rbcL
Acanthaceae	<i>Hypoestes</i>	<i>Hypoestes forskalii</i>	trnL	MK26165 8	MK26173 8	MK26147 8
Acanthaceae	<i>Hypoestes</i>	<i>Hypoestes forskalii</i>	rbcL	MK28517 3	MK28522 3	KR73699 5

Acanthaceae	<i>Justicia</i>	<i>Justicia anagalloides</i>	trnL	KR73788 5	KR73799 6	KR73809 2
Acanthaceae	<i>Justicia</i>	<i>Justicia anagalloides</i>	rbcL	KR73667 9	KR73679 5	KR73742 0
Acanthaceae	<i>Justicia</i>	<i>Justicia betonica</i>	trnL	KP74416 2	KT07501 1	
Acanthaceae	<i>Justicia</i>	<i>Justicia betonica</i>	rbcL	KF66939 0		
Acanthaceae	<i>Justicia</i>	<i>Justicia flava</i>	trnL	MK28237 4	MZ46156 2	
Acanthaceae	<i>Justicia</i>	<i>Justicia flava</i>	rbcL	MZ46158 4		
Acanthaceae	<i>Justicia</i>	<i>Justicia protracta</i>	trnL	MZ46156 3		
Acanthaceae	<i>Justicia</i>	<i>Justicia protracta</i>	rbcL	MZ46158 5		
Acanthaceae	<i>Monechma</i>	<i>Monechma debile</i>	trnL	KR73859 8	KR73871 0	KR73758 8
Acanthaceae	<i>Monechma</i>	<i>Monechma debile</i>	rbcL	KR73664 0	KR73744 6	KR73756 5
Acanthaceae	<i>Peristrophe</i>	<i>Peristrophe cernua</i>	rbcL	AM23478 2		
Acanthaceae	<i>Rhinacanthus</i>	<i>Rhinacanthus xerophilus</i>	trnL	MK28240 0		

Acanthaceae	<i>Ruellia</i>	<i>Ruellia patula</i>	trnL	KR73868 5	KR73821 5	KR73824 7
Acanthaceae	<i>Ruellia</i>	<i>Ruellia patula</i>	rbcL	KR73697 8	KR73703 8	KR73707 0
Acanthaceae	<i>Ruellia</i>	<i>Ruellia prostrata</i>	trnL	KR73871 3	KR73776 2	KR73778 1
Acanthaceae	<i>Ruellia</i>	<i>Ruellia prostrata</i>	rbcL	MK28513 0	KR73756 8	KR73654 9
Acanthaceae	<i>Thunbergia</i>	<i>Thunbergia alata</i>	rbcL	KR73637 1	KR73700 0	KR73709 9
Aizoaceae	<i>Aizoon</i>	<i>Aizoon canariense</i>	trnL	KX19755 3	HE58509 1	
Aizoaceae	<i>Aizoon</i>	<i>Aizoon canariense</i>	rbcL	MF69465 9	KX28251 5	MK09715 1
Aizoaceae	<i>Zaleya</i>	<i>Zaleya pentandra</i>	trnL	KR73849 3		
Aizoaceae	<i>Zaleya</i>	<i>Zaleya pentandra</i>	rbcL	KR73732 8		
Amaranthaceae	<i>Achyranthes</i>	<i>Achyranthes aspera</i>	trnL	MK18695 9	KR73855 6	KR73811 9
Amaranthaceae	<i>Achyranthes</i>	<i>Achyranthes aspera</i>	rbcL	MF69465 3	AY27004 8	MK95920 1

Amaranthace ae	Achyroopsis	<i>Achyroopsis leptostachya</i>	trnL	LT99302 8		
Amaranthace ae	Aerva	<i>Aerva leucura</i>	rbcL	PNG209- 18.rbcL	PNG499- 18.rbcL	
Amaranthace ae	<i>Alternanthera</i>	<i>Alternanthera pungens</i>	trnL	MK26172 7	KR73825 6	
Amaranthace ae	<i>Alternanthera</i>	<i>Alternanthera pungens</i>	rbcL	KR73721 6	KR73707 9	KR73642 7
Amaranthace ae	<i>Alternanthera</i>	<i>Alternanthera sessilis</i>	trnL	MK26154 6		
Amaranthace ae	<i>Alternanthera</i>	<i>Alternanthera sessilis</i>	rbcL	IPSUB10 8-18.rbcL	MK28521 6	KJ773257
Amaranthace ae	<i>Amaranthus</i>	<i>Amaranthus hybridus</i>	trnL	MK26172 1	MK26136 5	MK26138 9
Amaranthace ae	<i>Amaranthus</i>	<i>Amaranthus hybridus</i>	rbcL	MK28520 6	MK28518 6	
Amaranthace ae	<i>Amaranthus</i>	<i>Amaranthus praetermissus</i>	trnL	MG68546 8		
Amaranthace ae	<i>Amaranthus</i>	<i>Amaranthus thunbergii</i>	trnL	MG68548 7	MG68548 6	
Amaranthace ae	<i>Celosia</i>	<i>Celosia trigyna</i>	trnL	MK26154 2	MK26131 6	MK18700 5
Amaranthace ae	<i>Chenopodium</i>	<i>Chenopodium (Dysphania) ambrosioides</i>	trnL	MF07367 0	MF07366 8	MF07366 7

Amaranthaceae	<i>Chenopodium</i>	<i>Chenopodium (Dysphania) ambrosioides</i>	rbcL	MH65869 3	MK52573 5	MK28508 3
Amaranthaceae	<i>Chenopodium</i>	<i>Chenopodium album</i>	trnL	MF07360 3	MF07360 5	MF07363 6
Amaranthaceae	<i>Chenopodium</i>	<i>Chenopodium album</i>	rbcL	JX848451	HM84988 8	MF69468
Amaranthaceae	<i>Cyathula</i>	<i>Cyathula lanceolata</i>	trnL	LT99303 5	LT995128	
Amaranthaceae	<i>Gomphrena</i>	<i>Gomphrena celosioides</i>	trnL	LT99303 8	MF03968 0	MK18707 3
Amaranthaceae	<i>Gomphrena</i>	<i>Gomphrena celosioides</i>	rbcL	MH04989 1	MH04989 3	MH04989 4
Amaranthaceae	<i>Guilleminea</i>	<i>Guilleminea densa</i>	trnL	LT99306 1	EF688764	
Amaranthaceae	<i>Guilleminea</i>	<i>Guilleminea densa</i>	rbcL	AY27009 1		KP14953 6
Amaranthaceae	<i>Kyphocarpa</i>	<i>Kyphocarpa angustifolia</i>	trnL	LT99301 7	EF688778	
Amaranthaceae	<i>Kyphocarpa</i>	<i>Kyphocarpa angustifolia</i>	rbcL	JQ693464		
Amaranthaceae	<i>Pupalia</i>	<i>Pupalia lappacea</i>	trnL	LT99305 0	LT993042	LT993040
Amaranthaceae	<i>Pupalia</i>	<i>Pupalia lappacea</i>	rbcL	AY27012 2	GENG454 -14.rbcL	PNG503- 18.rbcL

Amaranthaceae	<i>Sericorema</i>	<i>Sericorema remotiflora</i>	trnL	LT99303 1	EF688788	
Amaryllidaceae	<i>Ammocharis</i>	<i>Ammocharis coranica</i>	trnL	JX464335	AY13915 2	
Amaryllidaceae	<i>Boophane</i>	<i>Boophane disticha</i>	trnL	JX464336		
Amaryllidaceae	<i>Boophone</i>	<i>Boophone disticha</i>	rbcL	JQ025023		
Amaryllidaceae	<i>Crinum</i>	<i>Crinum buphanoides</i>	trnL	AY13915 7	EF119732	EU52377 4
Amaryllidaceae	<i>Crinum</i>	<i>Crinum macowanii</i>	trnL	KR73810 2	KR73770 6	KR73760 4
Amaryllidaceae	<i>Crinum</i>	<i>Crinum macowanii</i>	rbcL	KR73691 0	KR73648 5	KR73637 5
Amaryllidaceae	<i>Crinum</i>	<i>Crinum moorei</i>	trnL	JX464342		
Amaryllidaceae	<i>Crinum</i>	<i>Crinum moorei</i>	rbcL	KY62730 3		
Amaryllidaceae	<i>Tulbaghia</i>	<i>Tulbaghia leucantha</i>	trnL	KU69210 8	KU69210 7	
Anacardiaceae	<i>Lannea</i>	<i>Lannea discolor</i>	trnL	PNG347- 18.trnL-F		
Anacardiaceae	<i>Lannea</i>	<i>Lannea discolor</i>	rbcL	JF265496	PNG347- 18.rbcL	

Anacardiaceae	<i>Lannea</i>	<i>Lannea edulis</i>	rbcL	JF265497	JQ025059	JX572710
Anacardiaceae	<i>Lannea</i>	<i>Lannea schweinfurthii</i>	trnL	AY594552		
Anacardiaceae	<i>Lannea</i>	<i>Lannea schweinfurthii</i>	rbcL	GU935430	JF265498	JX572711
Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa albicans</i>	rbcL	KF147498		
Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa engleri</i>	rbcL	JF265536	JX572820	
Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa insignis</i>	trnL	AY594444		
Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa insignis</i>	rbcL	KU568163		
Anacardiaceae	<i>Ozoroa</i>	<i>Ozoroa paniculosa</i>	rbcL	JX572822		
Anacardiaceae	<i>Sclerocarya</i>	<i>Sclerocarya birrea</i>	trnL	KR081762	KR081763	KC479214
Anacardiaceae	<i>Sclerocarya</i>	<i>Sclerocarya birrea</i>	rbcL	MN166719	MN216516	MN216531
Anacardiaceae	<i>Searsia</i>	<i>Searsia gueinzii</i>	rbcL	EU213504	EU213506	EU213505
Anacardiaceae	<i>Searsia</i>	<i>Searsia pentheri</i>	rbcL	JX572976	JF265592	
Anacardiaceae	<i>Searsia</i>	<i>Searsia pyroides</i>	trnL	AY640467	FJ946002	

Anacardiaceae	<i>Searsia</i>	<i>Searsia pyroides</i>	rbcL	JF265593	JX572929	JX572977
Annonaceae	<i>Hexalobus</i>	<i>Hexalobus monopetalus</i>	rbcL	JF265472	JX572672	
Annonaceae	<i>Monanthes</i>	<i>Monanthes affinis</i>	rbcL	MF353750	JF265520	
Apiaceae	<i>Centella</i>	<i>Centella asiatica</i>	rbcL	MK905059	MK905037	MK905036
Apiaceae	<i>Papaya</i>	<i>Papaya edulis</i>	trnL	KF180245	KF180246	KF180320
Apiaceae	<i>Papaya</i>	<i>Papaya edulis</i>	rbcL	AM235130	JF265540	JX572827
Apocynaceae	<i>Adenium</i>	<i>Adenium multiflorum</i>	rbcL	JF265270	JX572243	
Apocynaceae	<i>Adenium</i>	<i>Adenium swazicum</i>	rbcL	JF265271	JX572244	
Apocynaceae	<i>Carissa</i>	<i>Carissa bispinosa</i>	trnL	AF214172		
Apocynaceae	<i>Carissa</i>	<i>Carissa bispinosa</i>	rbcL	JF265326	JX572375	AJ419738
Apocynaceae	<i>Carissa</i>	<i>Carissa edulis</i>	rbcL	JX572376	JF265327	MT231387
Apocynaceae	<i>Catharanthus</i>	<i>Catharanthus roseus</i>	trnL	AF214175		
Apocynaceae	<i>Catharanthus</i>	<i>Catharanthus roseus</i>	rbcL	MN125628	MH549765	MH069755

Apocynaceae	<i>Cryptolepis</i>	<i>Cryptolepis oblongifolia</i>	trnL	DQ916872	AF214422	KT280301
Apocynaceae	<i>Cynanchum</i>	<i>Cynanchum viminale</i>	trnL	LN896759	HG530580.2	LN896724
Apocynaceae	<i>Cynanchum</i>	<i>Cynanchum viminale</i>	rbcL	JQ025085	KR736494	KR736504
Apocynaceae	<i>Fockea</i>	<i>Fockea angustifolia</i>	trnL	AM233365	AM233366	AM233367
Apocynaceae	<i>Gomphocarpus</i>	<i>Gomphocarpus tomentosus</i>	trnL	AM295753		AM295755
Apocynaceae	<i>Pentarrhinum</i>	<i>Pentarrhinum insipidum</i>	trnL	AJ410233	AJ410232	AJ410234
Apocynaceae	<i>Pentarrhinum</i>	<i>Pentarrhinum insipidum</i>	rbcL	HG530565		
Apocynaceae	<i>Pergularia</i>	<i>Pergularia daemia</i>	trnL	AJ290891	AJ290893	
Apocynaceae	<i>Pergularia</i>	<i>Pergularia daemia</i>	rbcL	JQ025071	API105-12.rbcL	GENG635-14.rbcL
Apocynaceae	<i>Raphionacme</i>	<i>Raphionacme elata</i>	trnL	AJ581816	AJ581814	
Apocynaceae	<i>Riocreuxia</i>	<i>Riocreuxia torulosa</i>	trnL	LT595694	LT595695	LT595681
Apocynaceae	<i>Riocreuxia</i>	<i>Riocreuxia torulosa</i>	rbcL	AM234841		
Apocynaceae	<i>Secamone</i>	<i>Secamone parvifolia</i>	trnL	HE805520		
Apocynaceae	<i>Stapelia</i>	<i>Stapelia gigantea</i>	trnL	KF678326	KF678044	

Apocynaceae	<i>Stapelia</i>	<i>Stapelia gigantea</i>	rbcL	JQ025094		
Apocynaceae	<i>Stomatostemma</i>	<i>Stomatostemma monteiroae</i>	trnL	AJ431780	AJ581833	AF214428
Apocynaceae	<i>Tacazzea</i>	<i>Tacazzea apiculata</i>	trnL	DQ221206	AY899972	DQ221164
Apocynaceae	<i>Tacazzea</i>	<i>Tacazzea apiculata</i>	rbcL	AJ419764		
Arecaceae	<i>Hyphaene</i>	<i>Hyphaene coriacea</i>	trnL	AM903235		
Arecaceae	<i>Hyphaene</i>	<i>Hyphaene coriacea</i>	rbcL	AM903195	AY012470	MG437594
Arecaceae	<i>Phoenix</i>	<i>Phoenix reclinata</i>	trnL	AJ241272	MK261397	
Arecaceae	<i>Phoenix</i>	<i>Phoenix reclinata</i>	rbcL	AJ404767	MG437546	GU135280
Aristolochiaceae	<i>Aristolochia</i>	<i>Aristolochia elegans</i>	trnL	AY689164	AY781544	
Aristolochiaceae	<i>Aristolochia</i>	<i>Aristolochia elegans</i>	rbcL	INB163-12.rbcLa		
Asparagaceae	<i>Albuca</i>	<i>Albuca setosa</i>	rbcL	AM902292		
Asparagaceae	<i>Asparagus</i>	<i>Asparagus africanus</i>	trnL	MH092175	MK186979	KJ774039
Asparagaceae	<i>Asparagus</i>	<i>Asparagus africanus</i>	rbcL	JQ014166	PNG388-18.rbcL	PNG539-18.rbcL
Asparagaceae	<i>Asparagus</i>	<i>Asparagus crassycladus</i>	rbcL	JQ014167		

Asparagaceae	<i>Asparagus</i>	<i>Asparagus falcatus</i>	trnL	KR73797 8	AF508514	MH09217 4
Asparagaceae	<i>Asparagus</i>	<i>Asparagus falcatus</i>	rbcL	MH63026 0	KR73744 2	KR73702 0
Asparagaceae	<i>Asparagus</i>	<i>Asparagus plumosus (setaceus)</i>	trnL	KJ774038		
Asparagaceae	<i>Asparagus</i>	<i>Asparagus plumosus (setaceus)</i>	rbcL	JQ734483	JQ734482	KX78383 1
Asparagaceae	<i>Asparagus</i>	<i>Asparagus retrofractus</i>	trnL	KJ774035		
Asparagaceae	<i>Asparagus</i>	<i>Asparagus retrofractus</i>	rbcL	KP11020 2		
Asparagaceae	<i>Asparagus</i>	<i>Asparagus suaveolens</i>	rbcL	JQ014168		
Asparagaceae	<i>Chlorophytum</i>	<i>Chlorophytum galpinii</i>	trnL	EU00003 1		
Asparagaceae	<i>Dipcadi</i>	<i>Dipcadi viride</i>	trnL	AM90244 6		
Asparagaceae	<i>Dipcadi</i>	<i>Dipcadi viride</i>	rbcL	AM90230 0		
Asparagaceae	<i>Eriospermum</i>	<i>Eriospermum flagelliforme</i>	rbcL	HM64047 5		
Asparagaceae	<i>Ledebouria</i>	<i>Ledebouria revoluta</i>	trnL	AJ507944	AJ507945	Z99145
Asparagaceae	<i>Ledebouria</i>	<i>Ledebouria revoluta</i>	rbcL	MF34964 5		

Asparagaceae	<i>Ornithogalum</i>	<i>Ornithogalum seineri</i>	trnL	AM90246 1		
Asparagaceae	<i>Ornithogalum</i>	<i>Ornithogalum seineri</i>	rbcL	AM90231 4		
Asparagaceae	<i>Sansevieria</i>	<i>Sansevieria hyacinthoides</i>	trnL	MH09221 7		
Asparagaceae	<i>Sansevieria</i>	<i>Sansevieria hyacinthoides</i>	rbcL	GU13519 5	JQ025084	GENG74 7-14.rbcL
Asparagaceae	<i>Sansevieria</i>	<i>Sansevieria pearsonii</i>	trnL	MH09222 0		
Asparagaceae	<i>Urginea</i>	<i>Drimia altissima</i>	trnL	AM41090 9		
Asparagaceae	<i>Urginea</i>	<i>Drimia altissima</i>	rbcL	JQ273914	HM64050 5	HM64051 5
Asphodelaceae	<i>Aloe</i>	<i>Aloe chabaudii</i>	rbcL	KNPA97 8- 09.rbcLa		
Asphodelaceae	<i>Aloe</i>	<i>Aloe marlothii</i>	rbcL	JF265285		
Asphodelaceae	<i>Bulbine</i>	<i>Bulbine abyssinica</i>	trnL	MK26163 8		
Asphodelaceae	<i>Bulbine</i>	<i>Bulbine abyssinica</i>	rbcL	MK28520 4	MK28520 5	
Asteraceae	<i>Acanthospermum</i>	<i>Acanthospermum hispidum</i>	trnL	MK26161 5		

Asteraceae	<i>Acanthospermum</i>	<i>Acanthospermum hispidum</i>	rbcL	MG224060	MG221964	
Asteraceae	<i>Ageratina</i>	<i>Ageratina altissima</i>	rbcL	KP643669	KJ841084	HQ590095
Asteraceae	<i>Ageratum</i>	<i>Ageratum conyzoides</i>	trnL	KR738616	KR738644	KR738674
Asteraceae	<i>Ageratum</i>	<i>Ageratum conyzoides</i>	rbcL	MG707331	KR737496	KR737527
Asteraceae	<i>Ambrosia</i>	<i>Ambrosia artemisiifolia</i>	trnL	MG709415		
Asteraceae	<i>Ambrosia</i>	<i>Ambrosia artemisiifolia</i>	rbcL	MN812527	MT642638	MK900737
Asteraceae	<i>Aspilia</i>	<i>Aspilia mossambicensis</i>	trnL	MK186981		
Asteraceae	<i>Aspilia</i>	<i>Aspilia mossambicensis</i>	rbcL	KR736520	KR737379	KR736733
Asteraceae	<i>Baccharoides</i>	<i>Baccharoides adoensis</i>	trnL	EF155833		
Asteraceae	<i>Bidens</i>	<i>Bidens biternata</i>	rbcL	MK903548	MH168544	
Asteraceae	<i>Bidens</i>	<i>Bidens pilosa</i>	rbcL	MF688949	MH658694	KU958561
Asteraceae	<i>Cirsium</i>	<i>Cirsium vulgare</i>	trnL	MN314896	MN919154	AF129826
Asteraceae	<i>Cirsium</i>	<i>Cirsium vulgare</i>	rbcL	HM849899	MT189198	HQ590038

Asteraceae	<i>Conyza</i>	<i>Conyza scabrida</i>	rbcL	AM23486 1		
Asteraceae	<i>Geigeria</i>	<i>Geigeria burkei</i>	trnL	LN60709 8		
Asteraceae	<i>Geigeria</i>	<i>Geigeria ornativa</i>	trnL	LN60710 5		
Asteraceae	<i>Hirpicium</i>	<i>Hirpicium bechuanense</i>	trnL	EU52728 8	JQ220177	MG83320 4
Asteraceae	<i>Parthenium</i>	<i>Parthenium hysterophorus</i>	trnL	MH01792 0	AY21621 5	MH01789 5
Asteraceae	<i>Parthenium</i>	<i>Parthenium hysterophorus</i>	rbcL	AY21515 5	JQ933433	KJ773721
Asteraceae	<i>Pegolettia</i>	<i>Pegolettia senegalensis</i>	trnL	LN60714 9		
Asteraceae	<i>Pluchea</i>	<i>Pluchea bojeri</i>	trnL	LN60715 5		
Asteraceae	<i>Pseudoconyza</i>	<i>Pseudoconyza viscosa</i>	trnL	LN60718 3	EF211021	
Asteraceae	<i>Pseudognaphal ium</i>	<i>Pseudognaphalium luteoalbum</i>	trnL	FN64576 3		
Asteraceae	<i>Pseudognaphal ium</i>	<i>Pseudognaphalium luteoalbum</i>	rbcL	KT62676 9	HM85003 4	KP14953 1
Asteraceae	<i>Schkuhria</i>	<i>Schkuhria pinnata</i>	trnL	KR73852 1	KR73845 8	
Asteraceae	<i>Schkuhria</i>	<i>Schkuhria pinnata</i>	rbcL	KR73728 9	KR73736 2	

Asteraceae	<i>Schkuhria</i>	<i>Schkuhria pinnata</i>	rbcL	KR737362	KR737289	
Asteraceae	<i>Sphaeranthus</i>	<i>Sphaeranthus incisus</i>	trnL	LN607208		
Asteraceae	<i>Tagetes</i>	<i>Tagetes minuta</i>	trnL	KR738171	KR738696	KR738034
Asteraceae	<i>Tagetes</i>	<i>Tagetes minuta</i>	rbcL	MF694781	KR736990	KR736837
Asteraceae	<i>Tridax</i>	<i>Tridax procumbens</i>	rbcL	JQ933511	KJ773956	MH767506
Asteraceae	<i>Vernonia</i>	<i>Vernonia (Hilliardiella) oligocephala</i>	trnL	EF155870		
Asteraceae	<i>Vernonia</i>	<i>Vernonia (Hilliardiella) oligocephala</i>	rbcL	JQ025103	JQ025104	MF688955
Asteraceae	<i>Vernonia</i>	<i>Vernonia (Linzia) glabra</i>	trnL	JN715856	JN715857	
Asteraceae	<i>Vernonia</i>	<i>Vernonia (Parapolydora) fastigiata</i>	trnL	EF155885	EF155905	
Asteraceae	<i>Vernonia</i>	<i>Vernonia (Polydora) poskeana</i>	trnL	JN837198		
Asteraceae	<i>Vernonia</i>	<i>Vernonia colorata</i>	trnL	JN715853		
Asteraceae	<i>Vernonia</i>	<i>Vernonia colorata</i>	rbcL	JF265649	PNG290-18.rbcL	

Asteraceae	<i>Xanthium</i>	<i>Xanthium strumarium</i>	trnL	MK26133 3	MK26138 7	MH98548 6
Asteraceae	<i>Xanthium</i>	<i>Xanthium strumarium</i>	rbcL	MK28508 7	MK28508 6	MH04996 3
Asteraceae	<i>Zinnia</i>	<i>Zinnia peruviana</i>	rbcL	MT18926 5	MH02887 2	
Bignoniaceae	<i>Kigelia</i>	<i>Kigelia africana</i>	trnL	AF03488 0	EF105072	
Bignoniaceae	<i>Kigelia</i>	<i>Kigelia africana</i>	rbcL	JN114810	JN114811	KY98831 6
Bignoniaceae	<i>Rhigozum</i>	<i>Rhigozum zambesiaceum</i>	trnL	KNPA55 2- 09.rbcLa	JF265571	JX572924
Burseraceae	<i>Commiphora</i>	<i>Commiphora africana</i>	trnL	MK26149 2		
Burseraceae	<i>Commiphora</i>	<i>Commiphora africana</i>	rbcL	JX572461	JF265357	
Burseraceae	<i>Commiphora</i>	<i>Commiphora angolensis</i>	trnL	KM51681 2		
Burseraceae	<i>Commiphora</i>	<i>Commiphora edulis</i>	trnL	FJ466480	GU24609 6	
Burseraceae	<i>Commiphora</i>	<i>Commiphora edulis</i>	rbcL	FJ466630	JF265358	JX572462
Burseraceae	<i>Commiphora</i>	<i>Commiphora glandulosa</i>	rbcL	JF265359		
Burseraceae	<i>Commiphora</i>	<i>Commiphora mollis</i>	rbcL	JF265362	JX572464	

Burseraceae	<i>Commiphora</i>	<i>Commiphora pyracanthoides</i>	rbcL	JX572465	JF265364	
Cactaceae	<i>Opuntia</i>	<i>Opuntia ficus-indica</i>	trnL	MK261524	JF712757	JF712756
Cactaceae	<i>Opuntia</i>	<i>Opuntia ficus-indica</i>	rbcL	MK285242	SDH786-14.rbcLa	KR737121
Cactaceae	<i>Opuntia</i>	<i>Opuntia stricta</i>	trnL	JF712826		
Cactaceae	<i>Opuntia</i>	<i>Opuntia stricta</i>	rbcL	KJ773705	JQ412402	
Campanulaceae	<i>Wahlenbergia</i>	<i>Wahlenbergia krebsii</i>	trnL	KC147797	HQ823521	
Campanulaceae	<i>Wahlenbergia</i>	<i>Wahlenbergia undulata</i>	trnL	HQ823547		
Cannabaceae	<i>Trema</i>	<i>Trema orientale</i>	trnL	KY827124	KY827125	AY488689
Cannabaceae	<i>Trema</i>	<i>Trema orientale</i>	rbcL	MH778733	MH767744	KY827166
Capparaceae	<i>Boscia</i>	<i>Boscia albitrunca</i>	rbcL	JF265307	KNPA495-09.rbcLa	JX572338
Capparaceae	<i>Boscia</i>	<i>Boscia angustifolia</i>	trnL	KR738286	KR738042	KR737829
Capparaceae	<i>Boscia</i>	<i>Boscia angustifolia</i>	rbcL	KR736845	KR736623	KR736458
Capparaceae	<i>Boscia</i>	<i>Boscia mossambicensis</i>	rbcL	JX572340	JF265310	MN166695

Capparaceae	<i>Capparis</i>	<i>Capparis sepiaria</i>	rbcL	JF265324	KY55852 9	JX572373
Capparaceae	<i>Capparis</i>	<i>Capparis tomentosa</i>	trnL	KR73781 3	KR73833 3	KR73777 5
Capparaceae	<i>Capparis</i>	<i>Capparis tomentosa</i>	rbcL	JF265325	JX572374	JQ898576
Capparaceae	<i>Maerua</i>	<i>Maerua angolensis</i>	trnL	KR73805 3	KR73801 6	KR73775 2
Capparaceae	<i>Maerua</i>	<i>Maerua angolensis</i>	rbcL	KR73685 9	KR73681 6	KR73653 9
Capparaceae	<i>Maerua</i>	<i>Maerua juncea</i>	rbcL	JF265510	JX572747	
Capparaceae	<i>Maerua</i>	<i>Maerua parvifolia</i>	rbcL	JF265511		
Caryophyllaceae	<i>Pollichia</i>	<i>Pollichia campestris</i>	trnL	KR73814 2	KR73806 8	KR73785 7
Caryophyllaceae	<i>Pollichia</i>	<i>Pollichia campestris</i>	rbcL	KR73687 5	KR73720 0	KR73665 1
Celastraceae	<i>Elaeodendron</i>	<i>Elaeodendron transvaalense</i>	trnL	MZ46155 7		
Celastraceae	<i>Elaeodendron</i>	<i>Elaeodendron transvaalense</i>	rbcL	AM23495 3	JF265407	JX572547
Celastraceae	<i>Gymnosporia</i>	<i>Gymnosporia heterophylla</i>	rbcL	JF265458		
Celastraceae	<i>Gymnosporia</i>	<i>Gymnosporia senegalensis</i>	trnL	EU32878 2	EU32883 4	EU32886 0
Celastraceae	<i>Gymnosporia</i>	<i>Gymnosporia senegalensis</i>	rbcL	JF265463	JX572661	AY38035 3

Celastraceae	<i>Hippocratea</i>	<i>Pristimera longipetiolata</i>	trnL	HM23023 2	HM23027 6	
Celastraceae	<i>Hippocratea</i>	<i>Pristimera longipetiolata</i>	rbcL	JQ025074	JF265558	JX572876
Celastraceae	<i>Maytenus</i>	<i>Maytenus acuminata</i>	rbcL	JX572760		
Celastraceae	<i>Mystroxylon</i>	<i>Mystroxylon aethiopicum</i>	trnL	MK26162 7	DQ21749 0	
Celastraceae	<i>Mystroxylon</i>	<i>Mystroxylon aethiopicum</i>	rbcL	AM23495 8	JF265523	JX572790
Celastraceae	<i>Robsonodendron</i>	<i>Robsonodendron eucleiforme</i>	trnL	DQ21746 8	DQ21749 5	
Celastraceae	<i>Robsonodendron</i>	<i>Robsonodendron eucleiforme</i>	rbcL	JX572935		
Cleomaceae	<i>Cleome</i>	<i>Cleome hirta</i>	trnL	KR73778 7	KR73789 8	KR73785 2
Cleomaceae	<i>Cleome</i>	<i>Cleome hirta</i>	rbcL	KU73962 0	KR73669 0	KR73657 7
Cleomaceae	<i>Cleome</i>	<i>Cleome monophylla</i>	trnL	KR73828 0	KR73835 8	KR73777 8
Cleomaceae	<i>Cleome</i>	<i>Cleome monophylla</i>	rbcL	KT58879 2	KT58879 3	KR73710 4
Cleomaceae	<i>Gynandropsis</i>	<i>Cleome (Gynandropsis) gynandra</i>	trnL	KR73819 8		

Cleomaceae	<i>Gynandropsis</i>	<i>Cleome</i> (<i>Gynandropsis</i>) <i>gynandra</i>	rbcL	KU73963 0	KR73701 7	
Colchicaceae	<i>Camptorrhiza</i>	<i>Camptorrhiza</i> <i>strumosa</i>	rbcL	KC79686 4		
Colchicaceae	<i>Gloriosa</i>	<i>Gloriosa superba</i>	trnL	AJ551362	EU04468 6	AJ551361
Colchicaceae	<i>Gloriosa</i>	<i>Gloriosa superba</i>	rbcL	KC79686 9	KR73742 4	KR73735 3
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>apiculatum</i>	rbcL	KC15850 3	EU33814 3	EU33814 2
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>collinum</i>	rbcL	EU33815 8	EU33815 9	FJ381790
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>erythrophyllum</i>	rbcL	EU33813 3	JF265347	JX572439
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>hereroense</i>	rbcL	KC15853 0	EU21345 7	EU33813 8
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>imberbe</i>	trnL	PNG471- 18.trnL-F		
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>imberbe</i>	rbcL	EU33815 5	EU33815 4	JF265349
Combretaceae	<i>Combretum</i>	<i>Combretum</i> <i>microphyllum</i>	rbcL	EU33813 0	JF265351	KC15853 1
Combretaceae	<i>Combretum</i>	<i>Combretum molle</i>	trnL	MK26134 7		
Combretaceae	<i>Combretum</i>	<i>Combretum molle</i>	rbcL	EU33814 6	EU33814 7	JF265352

Combretaceae	<i>Combretum</i>	<i>Combretum mossambicense</i>	rbcL	EU338131	JF265353	JX572448
Combretaceae	<i>Combretum</i>	<i>Combretum zeyheri</i>	trnL	MZ461553		
Combretaceae	<i>Combretum</i>	<i>Combretum zeyheri</i>	rbcL	EU338166	JF265356	MZ461578
Combretaceae	<i>Terminalia</i>	<i>Terminalia phanerophlebia</i>	rbcL	EU338119	JX573038	JF265624
Combretaceae	<i>Terminalia</i>	<i>Terminalia prunioides</i>	rbcL	EU338120	JF265625	
Combretaceae	<i>Terminalia</i>	<i>Terminalia sericea</i>	rbcL	EU338121	JF265626	JX573041
Commelinaceae	<i>Commelina</i>	<i>Commelina africana</i>	trnL	KR738448	EF092863	
Commelinaceae	<i>Commelina</i>	<i>Commelina africana</i>	rbcL	JQ025036	KR737249	KR737279
Commelinaceae	<i>Commelina</i>	<i>Commelina benghalensis</i>	trnL	KR738418	KR738678	KR737616
Commelinaceae	<i>Commelina</i>	<i>Commelina benghalensis</i>	rbcL	KR737245	MH767487	JF941286
Commelinaceae	<i>Commelina</i>	<i>Commelina diffusa</i>	rbcL	KJ773395	KU564782	
Commelinaceae	<i>Commelina</i>	<i>Commelina eckloniana</i>	trnL	EF092851	EF092853	
Commelinaceae	<i>Commelina</i>	<i>Commelina erecta</i>	trnL	KR737593	EF092858	

Commelinace ae	<i>Commelina</i>	<i>Commelina erecta</i>	rbcL	KY62676 7	MK52558 5	JQ591125
Commelinace ae	<i>Murdannia</i>	<i>Murdannia simplex</i>	trnL	EF092845		
Commelinace ae	<i>Murdannia</i>	<i>Murdannia simplex</i>	rbcL	GENG12 53- 15.rbcL		
Convolvulace ae	<i>Convolvulus</i>	<i>Convolvulus farinosus</i>	rbcL	HM84991 0	KC52920 8	KC52921 0
Convolvulace ae	<i>Convolvulus</i>	<i>Convolvulus sagittatus</i>	rbcL	KC52927 1		
Convolvulace ae	<i>Evolvulus</i>	<i>Evolvulus alsinoides</i>	trnL	KR73796 4	KR73865 3	KR73856 1
Convolvulace ae	<i>Evolvulus</i>	<i>Evolvulus alsinoides</i>	rbcL	KR73750 5	KR73740 9	MH67008 0
Convolvulace ae	<i>Ipomoea</i>	<i>Ipomoea cairica</i>	trnL	KR73869 2	KR73851 7	KR73827 8
Convolvulace ae	<i>Ipomoea</i>	<i>Ipomoea cairica</i>	rbcL	MK28510 2	MK28510 1	KR73754 6
Convolvulace ae	<i>Ipomoea</i>	<i>Ipomoea coptica</i>	rbcL	GENG16 55- 16.rbcL	PNG087- 18.rbcL	PNG461- 18.rbcL
Convolvulace ae	<i>Ipomoea</i>	<i>Ipomoea crassipes</i>	trnL	MZ46156 1		
Convolvulace ae	<i>Ipomoea</i>	<i>Ipomoea obscura</i>	trnL	AY10107 7	KP236630	KR73775 8

Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea obscura</i>	rbcL	KR73746 6	MH76751 5	KR73703 4
Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea plebeia</i>	rbcL	MT08454 5	MT08454 6	MT08454 7
Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea sinensis</i>	trnL	KP23663 9	KR73837 8	KR73855 5
Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea sinensis</i>	rbcL	KR73720 2	KR73740 3	KR73741 7
Convolvulaceae	<i>Jacquemontia</i>	<i>Jacquemontia tamnifolia</i>	trnL	MK18711 4	MK18711 3	
Convolvulaceae	<i>Jacquemontia</i>	<i>Jacquemontia tamnifolia</i>	rbcL	KY62748 1	PNG089-18.rbcL	PNG156-18.rbcL
Convolvulaceae	<i>Merremia</i>	<i>Merremia kentrocaulos</i>	trnL	KP23666 8		
Convolvulaceae	<i>Merremia</i>	<i>Merremia palmata</i>	trnL	KP23667 4		
Convolvulaceae	<i>Xenostegia</i>	<i>Xenostegia tridentata</i>	trnL	MH82572 9	MH82572 6	MH82572 8
Convolvulaceae	<i>Xenostegia</i>	<i>Xenostegia tridentata</i>	rbcL	MT05343 4	MT05343 5	
Corbichoniaceae	<i>Corbichonia</i>	<i>Corbichonia decumbens</i>	rbcL	FN82447 5	KF724212	MH01458 3
Crassulaceae	<i>Cotyledon</i>	<i>Cotyledon barbeyi</i>	trnL	MH50321 7	LN87883 9	
Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia adoensis</i>	trnL	HQ60840 3	HQ60839 6	

Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia adoensis</i>	rbcL	HQ60849 9	HQ60849 8	HQ60850 9
Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia rehmannii</i>	trnL	DQ53679 9	HQ62549 6	HQ60842 3
Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia rehmannii</i>	rbcL	DQ53579 3		
Cucurbitaceae	<i>Coccinia</i>	<i>Coccinia sessilifolia</i>	rbcL	AY96852 0		
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis anguria</i>	trnL	DQ78586 9	MK79719 1	KY43444 1
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis anguria</i>	rbcL	DQ78582 7	KY43439 7	KY62750 2
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis hirsutus</i>	trnL	DQ53680 4		
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis hirsutus</i>	rbcL	DQ53579 9		
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis metuliferus</i>	trnL	DQ78587 6	DQ78587 7	DQ53680 5
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis metuliferus</i>	rbcL	DQ53580 1	DQ78583 4	DQ78583 5
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis zeyheri</i>	trnL	DQ53680 7		
Cucurbitaceae	<i>Cucumis</i>	<i>Cucumis zeyheri</i>	rbcL	DQ53580 3		
Cucurbitaceae	<i>Kedrostis</i>	<i>Kedrostis africana</i>	rbcL	AJ235782		

Cucurbitacea e	<i>Kedrostis</i>	<i>Kedrostis foetidissima</i>	trnL	MK26158 0	MK26136 9	MK26141 2
Cucurbitacea e	<i>Lagenaria</i>	<i>Lagenaria sphaerica</i>	rbcL	GQ24862 7	EF590539	
Cucurbitacea e	<i>Momordica</i>	<i>Momordica balsamina</i>	trnL	GQ16322 8	GQ16323 0	GQ16322 7
Cucurbitacea e	<i>Momordica</i>	<i>Momordica balsamina</i>	rbcL	GENG10 47- 15.rbcL	GENG122 8-15.rbcL	GENG91 1-15.rbcL
Cucurbitacea e	<i>Momordica</i>	<i>Momordica boivinii</i>	trnL	GQ16323 1	GQ16323 2	GQ16323 3
Cucurbitacea e	<i>Momordica</i>	<i>Momordica charantia</i>	trnL	DQ50126 9	GQ16324 6	GQ16324 7
Cucurbitacea e	<i>Momordica</i>	<i>Momordica charantia</i>	rbcL	DQ53576 0	KT77931 4	MF07147 9
Cucurbitacea e	<i>Momordica</i>	<i>Momordica foetida</i>	trnL	GQ16327 6	GQ16327 8	GQ16327 9
Cucurbitacea e	<i>Momordica</i>	<i>Momordica foetida</i>	rbcL	DQ53582 9		
Cucurbitacea e	<i>Zehneria</i>	<i>Zehneria scabra</i>	trnL	KY52339 3	KY52339 2	
Cucurbitacea e	<i>Zehneria</i>	<i>Zehneria scabra</i>	rbcL	AM23497 0		
Cyperaceae	<i>Bulbostylis</i>	<i>Bulbostylis burchellii</i>	rbcL	AM99978 0		
Cyperaceae	<i>Bulbostylis</i>	<i>Bulbostylis hispidula</i>	trnL	MG84424 4		

Cyperaceae	<i>Bulbostylis</i>	<i>Bulbostylis hispidula</i>	rbcL	AM99978 1		
Cyperaceae	<i>Cyperus</i>	<i>Cyperus compressus</i>	trnL	MK79720 0		
Cyperaceae	<i>Cyperus</i>	<i>Cyperus compressus</i>	rbcL	MH65880 3	KY62725 2	
Cyperaceae	<i>Cyperus</i>	<i>Cyperus cyperoides</i>	trnL	JX644800	KR73782 2	KR73858 8
Cyperaceae	<i>Cyperus</i>	<i>Cyperus cyperoides</i>	rbcL	AB36994 7	JX644655	KR73661 4
Cyperaceae	<i>Cyperus</i>	<i>Cyperus esculentus</i>	rbcL	HM84993 7	KJ773425	LK02990 2
Cyperaceae	<i>Cyperus</i>	<i>Cyperus rotundus</i>	trnL	JX644807		
Cyperaceae	<i>Cyperus</i>	<i>Cyperus rotundus</i>	rbcL	KJ773433	JX644667	
Cyperaceae	<u><i>Cyperus</i></u>	<u><i>Cyperus rubicundus</i></u>	rbcL	KR73675 1	KR73702 7	KR73711 6
Cyperaceae	<i>Fimbristylis</i>	<i>Fimbristylis complanata</i>	trnL	JX644781	MG84423 2	
Cyperaceae	<i>Fimbristylis</i>	<i>Fimbristylis complanata</i>	rbcL	JX644687		
Cyperaceae	<i>Rhynchospora</i>	<i>Rhynchospora brownii</i>	rbcL	DQ05835 3		
Cyperaceae	<i>Schoenoplectus</i>	<i>Schoenoplectus corymbosus</i>	trnL	EF178607	KM46253 2	
Cyperaceae	<i>Schoenoplectus</i>	<i>Schoenoplectus corymbosus</i>	rbcL	EF178570		

Dioscoreaceae	<i>Dioscorea</i>	<i>Dioscorea cotinifolia</i>	trnL	KR07086 3	KR07086 2	KR07086 4
Dioscoreaceae	<i>Dioscorea</i>	<i>Dioscorea cotinifolia</i>	rbcL	EF614219	KR08703 4	KR08703 5
Ebenaceae	<i>Diospyros</i>	<i>Diospyros mespiliformis</i>	trnL	DQ92425 6	FJ238259	
Ebenaceae	<i>Diospyros</i>	<i>Diospyros mespiliformis</i>	rbcL	EU98071 2	JF265390	KP99685 4
Ebenaceae	<i>Euclea</i>	<i>Euclea crispa</i>	trnL	DQ92429 0	FJ238241	
Ebenaceae	<i>Euclea</i>	<i>Euclea crispa</i>	rbcL	EU98078 9	JX572574	
Ebenaceae	<i>Euclea</i>	<i>Euclea divinorum</i>	trnL	DQ92429 1	KR73767 9	KR73772 9
Ebenaceae	<i>Euclea</i>	<i>Euclea divinorum</i>	rbcL	EU98079 0	KR73657 2	KR73731 3
Ebenaceae	<i>Euclea</i>	<i>Euclea natalensis</i>	trnL	DQ92429 2	FJ238243	
Ebenaceae	<i>Euclea</i>	<i>Euclea natalensis</i>	rbcL	EU98079 1	JF265419	JF265420
Ebenaceae	<i>Euclea</i>	<i>Euclea schimperi</i>	trnL	FJ238246		
Ebenaceae	<i>Euclea</i>	<i>Euclea schimperi</i>	rbcL	JF265421	JF265422	
Ebenaceae	<i>Euclea</i>	<i>Euclea undulata</i>	trnL	DQ92429 3	FJ238248	
Ebenaceae	<i>Euclea</i>	<i>Euclea undulata</i>	rbcL	EU98079 2	EU04217 9	JF265423

Ebenaceae	<i>Royena</i>	<i>Royena lycioides</i>	trnL	FJ238254		
Ebenaceae	<i>Royena</i>	<i>Royena lycioides</i>	rbcL	EU980802	JF265389	JX572516
Ehretiaceae	<i>Cordia</i>	<i>Cordia caffra</i>	rbcL	JF265366		
Ehretiaceae	<i>Cordia</i>	<i>Cordia grandicalyx</i>	rbcL	JF265367		
Ehretiaceae	<i>Cordia</i>	<i>Cordia monoica</i>	trnL	KF158202	KR737768	KR738540
Ehretiaceae	<i>Cordia</i>	<i>Cordia monoica</i>	rbcL	KF158112	KR736555	KR737206
Ehretiaceae	<i>Ehretia</i>	<i>Ehretia amoena</i>	trnL	PNG303-18.trnL-F		
Ehretiaceae	<i>Ehretia</i>	<i>Ehretia amoena</i>	rbcL	MN166723	JX572543	JF265404
Ehretiaceae	<i>Ehretia</i>	<i>Ehretia rigida</i>	trnL	KF673278		
Ehretiaceae	<i>Ehretia</i>	<i>Ehretia rigida</i>	rbcL	EU042183	JF265405	JX572544
Euphorbiaceae	<i>Acalypha</i>	<i>Acalypha indica</i>	rbcL	API024-12.rbcL	KF381097	GENG676-14.rbcL
Euphorbiaceae	<i>Acalypha</i>	<i>Acalypha villicaulis</i>	rbcL	PNG352-18.rbcL		
Euphorbiaceae	<i>Croton</i>	<i>Croton gratissimus</i>	trnL	AY794696		
Euphorbiaceae	<i>Croton</i>	<i>Croton gratissimus</i>	rbcL	EU213460	EU213461	EU213459

Euphorbiaceae	<i>Croton</i>	<i>Croton megalobotrys</i>	trnL	KP878430		
Euphorbiaceae	<i>Croton</i>	<i>Croton megalobotrys</i>	rbcL	EU213463	EU213464	EU213462
Euphorbiaceae	<i>Euphorbia</i>	<i>Chamaesyce (Euphorbia) inaequilatera</i>	trnL	KR738541	KR737625	
Euphorbiaceae	<i>Euphorbia</i>	<i>Chamaesyce (Euphorbia) inaequilatera</i>	rbcL	KR737384	KR737306	KR737195
Euphorbiaceae	<i>Euphorbia</i>	<i>Chamaesyce (Euphorbia) neopolycnemoides</i>	trnL	HQ645613		
Euphorbiaceae	<i>Euphorbia</i>	<i>Chamaesyce (Euphorbia) prostrata</i>	trnL	MH490421	HQ645634	
Euphorbiaceae	<i>Euphorbia</i>	<i>Chamaesyce (Euphorbia) prostrata</i>	rbcL	MH767543	HM849994	
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia cyathophora</i>	trnL	JN249622		
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia cyathophora</i>	rbcL	MH014775		
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia hirta</i>	trnL	MH490420	HQ645584	PNG215-18.trnL-F
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia indica</i>	trnL	HQ645653	HQ645655	

Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia indica</i>	rbcL	GENG1063-15.rbcL	GENG1256-15.rbcL	
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia ingens</i>	trnL	JN207678		
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia tirucalli</i>	trnL	JN207723	JN249733	JN249734
Euphorbiaceae	<i>Euphorbia</i>	<i>Euphorbia tirucalli</i>	rbcL	JQ952252	JQ952253	JQ952254
Euphorbiaceae	<i>Jatropha</i>	<i>Jatropha zeyheri</i>	rbcL	JQ025058		
Euphorbiaceae	<i>Ricinus</i>	<i>Ricinus communis</i>	trnL	MK261654	MK261515	MK261385
Euphorbiaceae	<i>Spirostachys</i>	<i>Spirostachys africana</i>	trnL	AY794621		
Euphorbiaceae	<i>Spirostachys</i>	<i>Spirostachys africana</i>	rbcL	AY794838	AB267960	JF265602
Fabaceae	<i>Abrus</i>	<i>Abrus precatorius</i>	trnL	KX268134	EF543423	MK186948
Fabaceae	<i>Abrus</i>	<i>Abrus precatorius</i>	rbcL	GU135185	MN216761	PNG357-18.rbcL
Fabaceae	<i>Acacia</i>	<i>Acacia senegal</i>	trnL	KX268139	GQ872293	AF522976
Fabaceae	<i>Acacia</i>	<i>Acacia senegal</i>	rbcl	KNPA133-08.rbcLa	JX572227	JF265258

Fabaceae	<i>Aeschynomene</i>	<i>Aeschynomene indica</i>	rbcL	AF308701	MN726221	MN726220
Fabaceae	<i>Albizia</i>	<i>Albizia anthelmintica</i>	trnL	JQ230209		
Fabaceae	<i>Albizia</i>	<i>Albizia anthelmintica</i>	rbcL	JF265275	JX572254	
Fabaceae	<i>Albizia</i>	<i>Albizia forbesii</i>	rbcL	JF265277	JX572255	
Fabaceae	<i>Albizia</i>	<i>Albizia harveyi</i>	trnL	EU439977	PNG168-18.trnL-F	
Fabaceae	<i>Albizia</i>	<i>Albizia harveyi</i>	rbcL	JX572257	JF265278	PNG168-18.rbcL
Fabaceae	<i>Albizia</i>	<i>Albizia petersiana</i>	trnL	JQ230213	JQ230212	
Fabaceae	<i>Albizia</i>	<i>Albizia petersiana</i>	rbcL	JF265279	JX572258	
Fabaceae	<i>Albizia</i>	<i>Albizia versicolor</i>	trnL	MK186969	JQ230218	
Fabaceae	<i>Albizia</i>	<i>Albizia versicolor</i>	rbcL	JF265281	JX572260	
Fabaceae	<i>Alysicarpus</i>	<i>Alysicarpus glumaceus</i>	trnL	KX268145		
Fabaceae	<i>Alysicarpus</i>	<i>Alysicarpus rugosus</i>	trnL	KX268147		
Fabaceae	<i>Alysicarpus</i>	<i>Alysicarpus rugosus</i>	rbcL	KX119259		
Fabaceae	<i>Bauhinia</i>	<i>Bauhinia galpinii</i>	trnL	PNG353-18.trnL-F	FJ801056	FJ801055
Fabaceae	<i>Bauhinia</i>	<i>Bauhinia galpinii</i>	rbcL	JF265301	KNPA557-09.rbcLa	AM234262

Fabaceae	<i>Bolusanthus</i>	<i>Bolusanthus speciosus</i>	rbcL	JF265305		
Fabaceae	<i>Cassia</i>	<i>Cassia abbreviata</i>	rbcL	JX572384	JX572385	
Fabaceae	<i>Chamaecrista</i>	<i>Chamaecrista absus</i>	trnL	FJ009886		
Fabaceae	<i>Chamaecrista</i>	<i>Chamaecrista mimosoides</i>	trnL	MK18700 7	KX26816 2	PNG171- 18.trnL-F
Fabaceae	<i>Chamaecrista</i>	<i>Chamaecrista mimosoides</i>	rbcL	GENG10 14- 15.rbcL	GENG592 -14.rbcL	
Fabaceae	<i>Colophospermum</i>	<i>Colophospermum mopane</i>	trnL	AF54926 0		
Fabaceae	<i>Colophospermum</i>	<i>Colophospermum mopane</i>	rbcL	JF265343	JX572425	KY98833 3
Fabaceae	<i>Crotalaria</i>	<i>Crotalaria laburnifolia</i>	rbcL	JF265373	JX120599	JX120600
Fabaceae	<i>Crotalaria</i>	<i>Crotalaria monteiroi</i>	rbcL	JQ041241		
Fabaceae	<i>Crotalaria</i>	<i>Crotalaria podocarpa</i>	rbcL	JQ067598		
Fabaceae	<i>Crotalaria</i>	<i>Crotalaria sp.</i>	trnL	KR73790 9	KR73839 4	KR73815 2
Fabaceae	<i>Crotalaria</i>	<i>Crotalaria sphaerocarpa</i>	rbcL	JQ041269	JQ067599	
Fabaceae	<i>Crotalaria</i>	<i>Crotalaria virgulata</i>	rbcL	JQ041275		
Fabaceae	<i>Dalbergia</i>	<i>Dalbergia melanoxylon</i>	rbcL	KM51027 1	JF265384	JX572502

Fabaceae	<i>Dichrostachys</i>	<i>Dichrostachys cinerea</i>	trnL	KX26818 2	EU43999 3	PNG507- 18.trnL-F
Fabaceae	<i>Dichrostachys</i>	<i>Dichrostachys cinerea</i>	rbcL	MN59116 6	KX11929 0	JF265387
Fabaceae	<i>Dumasia</i>	<i>Dumasia villosa</i>	trnL	MK50122 0	MK50122 1	MK50122 2
Fabaceae	<i>Dumasia</i>	<i>Dumasia villosa</i>	rbcL	EU71726 5	MK50109 5	
Fabaceae	<i>Erythrina</i>	<i>Erythrina humeana</i>	rbcL	JF265413	JX571824	MK23890 0
Fabaceae	<i>Erythrina</i>	<i>Erythrina latissima</i>	rbcL	JF265414	MK23885 0	MK23895 1
Fabaceae	<i>Indigastrum</i>	<i>Indigastrum costatum</i>	rbcL	KNPA56 4- 09.rbcLa		
Fabaceae	<i>Indigofera</i>	<i>Indigofera arrecta</i>	trnL	KX26819 0	MK26145 4	MK26142 9
Fabaceae	<i>Indigofera</i>	<i>Indigofera arrecta</i>	rbcL	KX11929 6	KM89423 0	
Fabaceae	<i>Indigofera</i>	<i>Indigofera astragalina</i>	trnL	MK18710 6	PNG041- 18.trnL-F	
Fabaceae	<i>Indigofera</i>	<i>Indigofera astragalina</i>	rbcL	GENG12 58- 15.rbcL	PNG041- 18.rbcL	PNG074- 18.rbcL
Fabaceae	<i>Indigofera</i>	<i>Indigofera lupatana</i>	rbcL	MH63025 4	MN16668 0	

Fabaceae	<i>Indigofera</i>	<i>Indigofera rhytidocarpa</i>	rbcL	JX905964		
Fabaceae	<i>Indigofera</i>	<i>Indigofera schimperi</i>	trnL	MK26146 4	MK26141 4	
Fabaceae	<i>Indigofera</i>	<i>Indigofera spicata</i>	trnL	MK18710 7	PNG081- 18.trnL-F	
Fabaceae	<i>Indigofera</i>	<i>Indigofera spicata</i>	rbcL	KX11930 3	KJ773588	KY62723 1
Fabaceae	<i>Indigofera</i>	<i>Indigofera tinctoria</i>	trnL	MK18710 8	PNG482- 18.trnL-F	
Fabaceae	<i>Indigofera</i>	<i>Indigofera tinctoria</i>	rbcL	JF265485	KJ939238	KJ939239
Fabaceae	<i>Macrotyloma</i>	<i>Macrotyloma uniflorum</i>	trnL	EU71734 1		
Fabaceae	<i>Macrotyloma</i>	<i>Macrotyloma uniflorum</i>	rbcL	EU71726 9		
Fabaceae	<i>Mundulea</i>	<i>Mundulea sericea</i>	rbcL	JF265522	JQ025063	JX572788
Fabaceae	<i>Ormocarpum</i>	<i>Ormocarpum trichocarpum</i>	rbcL	JF265535	JX572810	KNPA542 -09.rbcLa
Fabaceae	<i>Peltophorum</i>	<i>Peltophorum africanum</i>	trnL	AY89971 2		
Fabaceae	<i>Peltophorum</i>	<i>Peltophorum africanum</i>	rbcL	AY90439 9	JF265546	JX571871
Fabaceae	<i>Philenoptera</i>	<i>Philenoptera violacea</i>	trnL	EU71735 7	MK18716 7	MK18716 8
Fabaceae	<i>Philenoptera</i>	<i>Philenoptera violacea</i>	rbcL	JF265547		

Fabaceae	<i>Pterocarpus</i>	<i>Pterocarpus angolensis</i>	trnL	JN083646		
Fabaceae	<i>Pterocarpus</i>	<i>Pterocarpus angolensis</i>	rbcL	JN083717	JN083716	JF265564
Fabaceae	<i>Pterocarpus</i>	<i>Pterocarpus rotundifolius</i>	trnL	JN083679	JN083677	JN083678
Fabaceae	<i>Pterocarpus</i>	<i>Pterocarpus rotundifolius</i>	rbcL	JN083748	JN083749	JN083751
Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia fleckii</i>	trnL	MT54752 2		
Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia fleckii</i>	rbcL	MZ46159 2		
Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia minima</i>	trnL	KR73816 0	KR73835 4	MT54749 5
Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia minima</i>	rbcL	KR73756 7	MH76769 1	MH54996 6
Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia monophylla</i>	rbcL	KF14751 0		
Fabaceae	<i>Schotia</i>	<i>Schotia brachypetala</i>	trnL	AY23275 2	GQ40514 2	
Fabaceae	<i>Schotia</i>	<i>Schotia brachypetala</i>	rbcL	JF265583	JQ025087	
Fabaceae	<i>Schotia</i>	<i>Schotia capitata</i>	trnL	GQ40516 7		
Fabaceae	<i>Schotia</i>	<i>Schotia capitata</i>	rbcL	JF265584		

Fabaceae	Senegalia	<i>Acacia (Senegalia) galpinii</i>	trnL	EU440000	GQ87226	KY68858
				0	8	9
Fabaceae	Senegalia	<i>Acacia (Senegalia) galpinii</i>	rbcL	JX572194		
Fabaceae	Senegalia	<i>Acacia (Senegalia) nigrescens</i>	trnL	KY68863	KY68863	
				0	1	
Fabaceae	Senegalia	<i>Acacia (Senegalia) nigrescens</i>	rbcL	EU21343	EU21343	JX518103
				8	9	
Fabaceae	Senegalia	<i>Acacia (Senegalia) schweinfurthii</i>	trnL	AF52297	EU43997	GQ87229
				9	6	0
Fabaceae	Senegalia	<i>Acacia (Senegalia) schweinfurthii</i>	rbcL	JF265257	JX572225	
Fabaceae	Senegalia	<i>Acacia (Senegalia) welwitschii</i>	trnL	KY68866	KY68866	GQ87229
				8	7	9
Fabaceae	Senegalia	<i>Acacia (Senegalia) welwitschii</i>	rbcL	JF265262	KNPA125	KNPA144
					-08.rbcLa	-08.rbcLa
Fabaceae	Senegalia	<i>Acacia burkei</i>	trnL	GQ87226		
				0		
Fabaceae	Senegalia	<i>Acacia burkei</i>	rbcL	JX572186		
Fabaceae	Senegalia	<i>Acacia caffra</i>	trnL	GQ87226	KY68858	
				1	5	
Fabaceae	Senegalia	<i>Acacia caffra</i>	rbcL	JX572187	JF265246	
Fabaceae	Senegalia	<i>Acacia erubescens</i>	trnL	GQ87226	KY68858	KY68858
				6	7	6
Fabaceae	Senegalia	<i>Acacia erubescens</i>	rbcL	JF265248		

Fabaceae	<i>Senegalia</i>	<i>Senegalia mellifera</i>	trnL	KR73782 6	KR73801 0	KR73825 8
Fabaceae	<i>Senegalia</i>	<i>Senegalia mellifera</i>	rbcL	JQ025017	JX572211	JX572212
Fabaceae	<i>Senna</i>	<i>Senna bicapsularis</i>	rbcL	MF13538 2	LC385911	LC38591 4
Fabaceae	<i>Senna</i>	<i>Senna didymobotrya</i>	trnL	MK26134 3		
Fabaceae	<i>Senna</i>	<i>Senna italica</i>	trnL	KX26821 1		
Fabaceae	<i>Senna</i>	<i>Senna italica</i>	rbcL	KY46411 2	KY46411 3	KY62335 9
Fabaceae	<i>Senna</i>	<i>Senna obtusifolia</i>	trnL	KX26821 2		
Fabaceae	<i>Senna</i>	<i>Senna obtusifolia</i>	rbcL	KJ773878	MF13541 9	GQ43673 3
Fabaceae	<i>Senna</i>	<i>Senna occidentalis</i>	trnL	KX26821 3	AF365030	EU36183 6
Fabaceae	<i>Senna</i>	<i>Senna occidentalis</i>	rbcL	KJ773879	MH76747 4	
Fabaceae	<i>Sesbania</i>	<i>Sesbania bispinosa</i>	trnL	KX26821 7		
Fabaceae	<i>Sesbania</i>	<i>Sesbania bispinosa</i>	rbcL	MH76769 3		

Fabaceae	<i>Stylosanthes</i>	<i>Stylosanthes fruticosa</i>	rbcL	KR73638 2	KR73637 8	KR73694 6
Fabaceae	<i>Tephrosia</i>	<i>Tephrosia purpurea</i>	rbcL	MG94686 8	MG94686 9	MH76770 0
Fabaceae	<i>Tephrosia</i>	<i>Tephrosia rhodesica</i>	rbcL	EU71728 8		
Fabaceae	<i>Teramnus</i>	<i>Teramnus labialis</i>	trnL	LC31510 6	EU71732 9	LC31512 0
Fabaceae	<i>Teramnus</i>	<i>Teramnus labialis</i>	rbcL	API314- 14.rbcL	GENG114 6-15.rbcL	GENG17 08- 16.rbcL
Fabaceae	<i>Tylosema</i>	<i>Tylosema fassoglense</i>	trnL	FJ801091	FJ801124	EU36174 3
Fabaceae	Vachellia	<i>Acacia (Vachellia) exuvialis</i>	trnL	GQ87226 7		
Fabaceae	Vachellia	<i>Acacia (Vachellia) exuvialis</i>	rbcL	EU21343 7	EU21343 6	EU21343 5
Fabaceae	Vachellia	<i>Acacia (Vachellia) gerrardii</i>	trnL	KR73857 2	KR73823 7	KR73852 4
Fabaceae	Vachellia	<i>Acacia (Vachellia) gerrardii</i>	rbcL	JF265250	KR73706 1	KR73736 4
Fabaceae	Vachellia	<i>Acacia (Vachellia) grandicornuta</i>	trnL	EU44002 6	GQ87227 1	
Fabaceae	Vachellia	<i>Acacia (Vachellia) grandicornuta</i>	rbcL	JF265251	JX572197	

Fabaceae	Vachellia	<i>Acacia (Vachellia) karroo</i>	trnL	AF52297 2	GQ87227 4	
Fabaceae	Vachellia	<i>Acacia (Vachellia) karroo</i>	rbcL	EU04217 8	JX572203	AM23500 3
Fabaceae	Vachellia	<i>Acacia (Vachellia) luederitzii</i>	trnL	GQ87227 8	JQ230220	JQ23021 9
Fabaceae	Vachellia	<i>Acacia (Vachellia) luederitzii</i>	rbcL	JX572208	JX572207	JF265253
Fabaceae	Vachellia	<i>Acacia (Vachellia) nilotica</i>	trnL	PNG427- 18.trnL-F	AF522973	KR73866 5
Fabaceae	Vachellia	<i>Acacia (Vachellia) nilotica</i>	rbcL	JF265255	JX572217	MH56045 2
Fabaceae	Vachellia	<i>Acacia (Vachellia) tortilis</i>	trnL	HM02082 2	AF522974	KR73817 9
Fabaceae	Vachellia	<i>Acacia (Vachellia) tortilis</i>	rbcL	JF265261	JX572233	MN59249 2
Fabaceae	Vachellia	<i>Acacia (Vachellia) xanthophloea</i>	trnL	KR73856 6	KR73844 2	KR73786 4
Fabaceae	Vachellia	<i>Acacia (Vachellia) xanthophloea</i>	rbcL	JF265263	JX572235	KR73727 2
Fabaceae	Vachellia	<i>Acacia borleae</i>	rbcL	JF265243	JX572185	
Fabaceae	Vachellia	<i>Acacia robusta</i>	trnL	MK18695 6		
Fabaceae	Vachellia	<i>Acacia robusta</i>	rbcL	JF265256	JX572223	PNG466- 18.rbcL

Fabaceae	<i>Vigna</i>	<i>Vigna luteola</i>	rbcL	KJ773979	MH55004 0	
Fabaceae	<i>Vigna</i>	<i>Vigna unguiculata</i>	trnL	AB30407 4	GQ41171 5	AB30402 3
Fabaceae	<i>Vigna</i>	<i>Vigna unguiculata</i>	rbcL	EU71726 6	GQ41165 9	API148- 12.rbcL
Fabaceae	<i>Vigna</i>	<i>Vigna vexillata</i>	trnL	AB30402 4		
Fabaceae	<i>Xanthocercis</i>	<i>Xanthocercis zambesiaca</i>	rbcL	JX573092		
Gentianaceae	<i>Enicostema</i>	<i>Enicostema axillare</i>	rbcL	API204- 12.rbcL	GENG158 9-16.rbcL	GENG79 0-14.rbcL
Geraniaceae	<i>Monsonia</i>	<i>Monsonia angustifolia</i>	trnL	KR73761 2	KR73769 6	KR73773 7
Geraniaceae	<i>Monsonia</i>	<i>Monsonia angustifolia</i>	rbcL	KR73755 3	KR73672 3	KR73705 4
Geraniaceae	<i>Monsonia</i>	<i>Monsonia emarginata</i>	trnL	HE79547 3		
Gisekiaceae	<i>Gisekia</i>	<i>Gisekia africana</i>	trnL	HE58508 6		
Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium ovalifolium</i>	trnL	KY76455 9	MK18708 5	PNG267- 18.trnL-F
Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium ovalifolium</i>	rbcL	KU71159 1	KU71159 2	KU71159 0
Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium strigosum</i>	trnL	MK26162 0		

Heliotropiaceae	<i>Heliotropium</i>	<i>Heliotropium strigosum</i>	rbcL	MF694729		
Hypoxidaceae	<i>Hypoxis</i>	<i>Hypoxis hemerocallidea</i>	trnL	HM459507		
Kewaceae	<i>Kewa</i>	<i>Kewa bowkeriana</i>	trnL	MK261354	MK261495	MK261437
Kewaceae	<i>Kewa</i>	<i>Kewa bowkeriana</i>	rbcL	FN824476		
Kewaceae	<i>Kewa</i>	<i>Kewa salsoloides</i>	trnL	KX197538		
Kewaceae	<i>Kewa</i>	<i>Kewa salsoloides</i>	rbcL	FN824477	FN824478	
Kirkiaceae	<i>Kirkia</i>	<i>Kirkia wilmsii</i>	rbcL	JF265493		
Lamiaceae	<i>Acrotome</i>	<i>Acrotome hispida</i>	trnL	EU138376	EU138299	
Lamiaceae	<i>Acrotome</i>	<i>Acrotome inflata</i>	trnL	EU138378	EU138377	EU138301
Lamiaceae	<i>Clerodendrum</i>	<i>Clerodendrum glabrum</i>	rbcL	JF265341	JX572414	
Lamiaceae	<i>Clerodendrum</i>	<i>Clerodendrum ternatum</i>	rbcL	KF147460		
Lamiaceae	<i>Leonotis</i>	<i>Leonotis nepetifolia</i>	trnL	EU138385	EU138386	KR737719
Lamiaceae	<i>Leonotis</i>	<i>Leonotis nepetifolia</i>	rbcL	JF265501	KR736498	KR736355

Lamiaceae	<i>Leonotis</i>	<i>Leonotis ocymifolia</i>	trnL	EU13838 8	EU13838 9	EU13838 7
Lamiaceae	<i>Leucas</i>	<i>Leucas capensis</i>	trnL	EU13839 9	EU13832 2	
Lamiaceae	<i>Leucas</i>	<i>Leucas glabrata</i>	trnL	MK26155 6	EU13840 7	EU13840 8
Lamiaceae	<i>Leucas</i>	<i>Leucas martinicensis</i>	trnL	EU13841 6	MK26166 1	EU13833 9
Lamiaceae	<i>Leucas</i>	<i>Leucas martinicensis</i>	rbcL	JQ673540		
Lamiaceae	<i>Leucas</i>	<i>Leucas neuflizeana</i>	trnL	EU13842 1	MK26152 9	EU13834 4
Lamiaceae	<i>Leucas</i>	<i>Leucas sexdentata</i>	trnL	EU13842 4	EU13834 7	
Lamiaceae	<i>Ocimum</i>	<i>Becium (Ocimum) filamentosum</i>	trnL	KR73870 4	KR73826 7	KR73777 0
Lamiaceae	<i>Ocimum</i>	<i>Becium (Ocimum) filamentosum</i>	rbcL	KR73755 9	KR73744 0	KR73709 0
Lamiaceae	<i>Ocimum</i>	<i>Ocimum americanum</i>	trnL	AJ505464	MK26154 1	MK26170 7
Lamiaceae	<i>Ocimum</i>	<i>Ocimum americanum</i>	rbcL	MF46818 6	MF46818 7	MF46818 8
Lamiaceae	<i>Ocimum</i>	<i>Ocimum gratissimum</i>	trnL	AJ505467	AJ505468	AJ505469
Lamiaceae	<i>Ocimum</i>	<i>Ocimum gratissimum</i>	rbcL	MF46819 2	MF46819 3	MF46819 4.2

Lamiaceae	<i>Ocimum</i>	<i>Ocimum labiatum</i>	trnL	AJ505470	AJ505471	
Lamiaceae	<i>Ocimum</i>	<i>Ocimum labiatum</i>	trnL	EU54314 2	EU54314 3	
Lamiaceae	<i>Plectranthus</i>	<i>Plectranthus caninus</i>	trnL	MK26150 9	MK26151 1	
Lamiaceae	<i>Stachys</i>	<i>Stachys hyssopoides</i>	trnL	FJ854319	FJ854218	
Lamiaceae	<i>Tetradenia</i>	<i>Tetradenia riparia</i>	trnL	MK26137 2		
Lamiaceae	<i>Tetradenia</i>	<i>Tetradenia riparia</i>	rbcL	MK28523 1		
Lamiales	Bignoniaceae	<i>Rhigozum zambesiacum</i>	rbcL	MZ46159 1		
Loganiaceae	<i>Strychnos</i>	<i>Strychnos madagascariensis</i>	trnL	MK18722 5	MK18722 6	KM36513 2
Loganiaceae	<i>Strychnos</i>	<i>Strychnos madagascariensis</i>	rbcL	EU21352 5	EU21352 4	EU21352 3
Loganiaceae	<i>Strychnos</i>	<i>Strychnos spinosa</i>	rbcL	EU21352 6	EU21352 8	EU21352 7
Malpighiaceae	<i>Triaspis</i>	<i>Triaspis hypericoides</i>	rbcL	HQ24759 8	HQ24759 6	JF265632
Malvaceae	<i>Abutilon</i>	<i>Abutilon angulatum</i>	trnL	PNG488-18.trnL-F		
Malvaceae	<i>Abutilon</i>	<i>Abutilon angulatum</i>	rbcL	JX572177	JF265241	PNG020-18.rbcL
Malvaceae	<i>Abutilon</i>	<i>Abutilon austroafricanum</i>	rbcL	MZ46157 1		

Malvaceae	Abutilon	<i>Abutilon grandiflorum</i>	trnL	MK26150 6		
Malvaceae	Abutilon	<i>Abutilon mauritianum</i>	trnL	KR73814 6	KR73788 9	KR73774 9
Malvaceae	Abutilon	<i>Abutilon mauritianum</i>	rbcL	KR73716 8	KR73709 3	KR73668 2
Malvaceae	Abutilon	<i>Abutilon ramosum</i>	rbcL	GENG12 44- 15.rbcL	GENG480 -14.rbcL	
Malvaceae	Abutilon	<i>Abutilon sonneratianum</i>	rbcL	JX572178		
Malvaceae	<i>Cienfuegosia</i>	<i>Cienfuegosia hildebrandtii</i>	rbcL	PNG143- 18.rbcL		
Malvaceae	<i>Corchorus</i>	<i>Corchorus trilocularis</i>	trnL	MK26136 1	MK26174 7	
Malvaceae	<i>Corchorus</i>	<i>Corchorus trilocularis</i>	rbcL	MK28515 7		
Malvaceae	<i>Dombeya</i>	<i>Dombeya rotundifolia</i>	rbcL	MH63025 6	JF265396	JQ025044
Malvaceae	<i>Gossypium</i>	<i>Gossypium herbaceum</i>	rbcL	JF265447	JN243187	JN243188
Malvaceae	<i>Grewia</i>	<i>Grewia bicolor</i>	trnL	MK26149 4	MK26147 3	MK26154 7
Malvaceae	<i>Grewia</i>	<i>Grewia bicolor</i>	rbcL	EU21348 5	EU21348 6	MZ46158 2

Malvaceae	<i>Grewia</i>	<i>Grewia flava</i>	rbcL	KM89422 1		
Malvaceae	<i>Grewia</i>	<i>Grewia flavescens</i>	trnL	MK18707 7		
Malvaceae	<i>Grewia</i>	<i>Grewia flavescens</i>	rbcL	EU21348 7	EU21348 8	JF265450
Malvaceae	<i>Grewia</i>	<i>Grewia hexamita</i>	rbcL	JF265452		
Malvaceae	<i>Grewia</i>	<i>Grewia monticola</i>	trnL	MK18708 0		
Malvaceae	<i>Grewia</i>	<i>Grewia monticola</i>	rbcL	JF265454	JX572641	
Malvaceae	<i>Grewia</i>	<i>Grewia villosa</i>	rbcL	EU21349 1	EU21348 9	EU21349 0
Malvaceae	<i>Hermannia</i>	<i>Hermannia sp.</i>	rbcL	AM23502 5		
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus calyphyllus</i>	trnL	KR73785 9	KR73798 1	KR73858 0
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus calyphyllus</i>	rbcL	KR73681 9	KR73730 7	KR73742 5
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus cannabinus</i>	trnL	MK18708 7	JQ625353	AY72722 7
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus cannabinus</i>	rbcL	DBHI004 -16.rbcL	DBHI005- 16.rbcL	DBHI006 -16.rbcL
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus lunariifolius</i>	trnL	KR73815 3	KR73798 9	KR73807 7
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus lunariifolius</i>	rbcL	KR73677 2	KR73678 8	KR73696 9

Malvaceae	<i>Hibiscus</i>	<i>Hibiscus micranthus</i>	trnL	MK26136 4	MK26151 9	MK26164 2
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus micranthus</i>	rbcL	JF265474	JX572675	KR73728 5
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus physaloides</i>	trnL	PNG058- 18.trnL-F		
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus physaloides</i>	rbcL	PNG058- 18.rbcL		
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus pusillus</i>	rbcL	KBGPP0 16- 18.rbcLa	KBGPP02 9- 18.rbcLa	KBGPP05 6- 18.rbcLa
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus sabiensis</i>	rbcL	MZ46158 3		
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus trionum</i>	trnL	KR73759 5	KR73789 9	KR73867 0
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus trionum</i>	rbcL	HM85005 1	MH65826 7	MG24650 9
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus vitifolius</i>	trnL	MK18709 1		
Malvaceae	<i>Hibiscus</i>	<i>Hibiscus vitifolius</i>	rbcL	PNG350- 18.rbcL		
Malvaceae	<i>Malvastrum</i>	<i>Malvastrum coromandelianum</i>	rbcL	MH05003 3		
Malvaceae	<i>Melhania</i>	<i>Melhania prostrata</i>	rbcL	MZ46158 7		

Malvaceae	<i>Pavonia</i>	<i>Pavonia burchellii</i>	rbcL	KNPA55 1- 09.rbcLa		
Malvaceae	<i>Sida</i>	<i>Sida alba</i>	trnL	MK26154 5	MK26152 6	
Malvaceae	<i>Sida</i>	<i>Sida cordifolia</i>	trnL	MK18721 3	MK18721 2	
Malvaceae	<i>Sida</i>	<i>Sida cordifolia</i>	rbcL	KJ773888	MH76764 0	
Malvaceae	<i>Sida</i>	<i>Sida ovata</i>	trnL	KR73769 4	KR73841 5	KR73863 8
Malvaceae	<i>Sida</i>	<i>Sida ovata</i>	rbcL	KR73749 0	KR73724 2	KR73647 1
Malvaceae	<i>Sida</i>	<i>Sida rhombifolia</i>	rbcL	HM85035 3	KJ773889	
Malvaceae	<i>Sida</i>	<i>Sida spinosa</i>	trnL	MK26154 5	MK26152 6	
Malvaceae	<i>Sida</i>	<i>Sida spinosa</i>	rbcL	DQ00610 5	MG24596 9	MG24724 2
Malvaceae	<i>Triumfetta</i>	<i>Triumfetta annua</i>	rbcL	JQ933514		
Malvaceae	<i>Triumfetta</i>	<i>Triumfetta pentandra</i>	rbcL	GENG12 54- 15.rbcL		
Malvaceae	<i>Triumfetta</i>	<i>Triumfetta rhomboidea</i>	rbcL	KX52708 0	GENG118 6-15.rbcL	

Malvaceae	<i>Waltheria</i>	<i>Waltheria indica</i>	trnL	MK79772 9	KR73805 9	KR73772 4
Malvaceae	<i>Waltheria</i>	<i>Waltheria indica</i>	rbcL	KR73686 5	MH76773 7	MF13541 0
Meliaceae	<i>Ekebergia</i>	<i>Ekebergia capensis</i>	trnL	KU93911 6		
Meliaceae	<i>Ekebergia</i>	<i>Ekebergia capensis</i>	rbcL	AJ402947	AY12822 8	KM89421 5
Meliaceae	<i>Melia</i>	<i>Melia azedarach</i>	trnL	FM17953 6	FN599481	AB05753 5
Meliaceae	<i>Melia</i>	<i>Melia azedarach</i>	rbcL	MG94685 1	AY12823 4	FN59945 3
Meliaceae	<i>Trichilia</i>	<i>Trichilia emetica</i>	trnL	MK18723 4	KU93912 5	
Meliaceae	<i>Trichilia</i>	<i>Trichilia emetica</i>	rbcL	U39082.2	AY12824 4	JX307354
Meliaceae	<i>Turraea</i>	<i>Turraea obtusifolia</i>	rbcL	JF265641		
Menispermaceae	<i>Cissampelos</i>	<i>Cissampelos mucronata</i>	rbcL	PNG372-18.rbcL		
Menispermaceae	<i>Cocculus</i>	<i>Cocculus hirsutus</i>	trnL	MT15383 3		
Menispermaceae	<i>Cocculus</i>	<i>Cocculus hirsutus</i>	rbcL	MT15374 8	MT15375 0	MK48269 3
Molluginaceae	<i>Glinus</i>	<i>Glinus oppositifolius</i>	trnL	MK18707 0	MK18707 1	

Molluginaceae	<i>Glinus</i>	<i>Glinus oppositifolius</i>	rbcL	FN824414	FN824415	FN824416
Molluginaceae	<i>Mollugo</i>	<i>Mollugo nudicaulis</i>	rbcL	FN824444	FN824445	FN824446
Moraceae	<i>Ficus</i>	<i>Ficus abutilifolia</i>	rbcL	EU213477	EU213475	EU213476
Moraceae	<i>Ficus</i>	<i>Ficus capreifolia</i>	rbcL	JX572601		
Moraceae	<i>Ficus</i>	<i>Ficus ingens</i>	trnL	MK261593		
Moraceae	<i>Ficus</i>	<i>Ficus ingens</i>	rbcL	MK285222		
Moraceae	<i>Ficus</i>	<i>Ficus salicifolia</i>	rbcL	JF265436	JX572609	
Moraceae	<i>Ficus</i>	<i>Ficus sur</i>	trnL	GQ504506		
Moraceae	<i>Ficus</i>	<i>Ficus sur</i>	rbcL	JF265438		
Moraceae	<i>Ficus</i>	<i>Ficus sycomorus</i>	trnL	MK261538	MK261530	GQ504508
Moraceae	<i>Ficus</i>	<i>Ficus sycomorus</i>	rbcL	EU213481	EU213482	EU213483
Moraceae	<i>Ficus</i>	<i>Ficus thonningii</i>	trnL	GQ504484	GQ504493	
Moraceae	<i>Ficus</i>	<i>Ficus thonningii</i>	rbcL	JF265432	JX571841	
Myrtaceae	<i>Heteropyxis</i>	<i>Heteropyxis natalensis</i>	rbcL	AM235662	JF265471	JQ025054
Myrtaceae	<i>Psidium</i>	<i>Psidium guajava</i>	trnL	MN443100	KC428590	

Myrtaceae	<i>Psidium</i>	<i>Psidium guajava</i>	rbcL	KX52709 7	KX02112 7	MH06983 5
Myrtaceae	<i>Syzygium</i>	<i>Syzygium guineense</i>	rbcL	JX573018	JX573019	JX573020
Nyctaginaceae	<i>Boerhavia</i>	<i>Boerhavia diffusa</i>	trnL	KU17253 6		
Ochnaceae	<i>Ochna</i>	<i>Ochna inermis</i>	rbcL	JF265529		
Ochnaceae	<i>Ochna</i>	<i>Ochna natalitia</i>	rbcL	JF265530	KF147495	KF26336 6
Oleaceae	<i>Jasminum</i>	<i>Jasminum fluminense</i>	trnL	KR73805 2	KR73834 7	KR73818 8
Oleaceae	<i>Jasminum</i>	<i>Jasminum fluminense</i>	rbcL	GU13517 7	MH54989 1	JF265486
Oleaceae	<i>Jasminum</i>	<i>Jasminum stenolobum</i>	rbcL	JF265488	JX572699	
Onagraceae	<i>Ludwigia</i>	<i>Ludwigia adscendens</i>	trnL	KU94199 3		
Onagraceae	<i>Ludwigia</i>	<i>Ludwigia adscendens</i>	rbcL	KX52747 0	PNG070- 18.rbcL	
Onagraceae	<i>Ludwigia</i>	<i>Ludwigia octovalvis</i>	trnL	KX16817 6	KX16817 2	KX16817 3
Onagraceae	<i>Ludwigia</i>	<i>Ludwigia octovalvis</i>	rbcL	MH05004 3	JF265505	
Onagraceae	<i>Ludwigia</i>	<i>Ludwigia stolonifera</i>	trnL	KR73812 4	KR73792 6	
Onagraceae	<i>Ludwigia</i>	<i>Ludwigia stolonifera</i>	rbcL	KR73693 6	KR73671 9	

Orchidaceae	<i>Ansellia</i>	<i>Ansellia africana</i>	rbcL	EU21344 4	EU21344 5	MH74884 4
Orobanchaceae	<i>Striga</i>	<i>Striga bilabiata</i>	trnL	EU25664 2		
Orobanchaceae	<i>Striga</i>	<i>Striga elegans</i>	rbcL	EU21351 9		
Orobanchaceae	<i>Striga</i>	<i>Striga gesnerioides</i>	trnL	EU26419 5		
Orobanchaceae	<i>Striga</i>	<i>Striga gesnerioides</i>	rbcL	PNG110- 18.rbcL		
Oxalidaceae	<i>Oxalis</i>	<i>Oxalis latifolia</i>	trnL	JN639556	JN639573	JN639574
Oxalidaceae	<i>Oxalis</i>	<i>Oxalis latifolia</i>	rbcL	EU00228 2	HM85022 3	
Papaveraceae	<i>Argemone</i>	<i>Argemone mexicana</i>	trnL	PNG151- 18.trnL-F	AY32824 8	
Papaveraceae	<i>Argemone</i>	<i>Argemone mexicana</i>	rbcL	GQ43673 0	GENG647 -14.rbcL	MG24917 4
Passifloraceae	<i>Adenia</i>	<i>Adenia digitata</i>	trnL	MK18696 3		
Passifloraceae	<i>Tricliceras</i>	<i>Tricliceras laceratum</i>	rbcL	AB53657 3		
Pedaliaceae	<i>Ceratotheca</i>	<i>Ceratotheca triloba</i>	trnL	KJ743178	AF482611	
Pedaliaceae	<i>Ceratotheca</i>	<i>Ceratotheca triloba</i>	rbcL	JQ025031		
Pedaliaceae	<i>Dicerocaryum</i>	<i>Dicerocaryum eriocarpum</i>	trnL	KJ743186		

Pedaliaceae	<i>Dicerocaryum</i>	<i>Dicerocaryum senecioides</i>	rbcL	PNG392-18.rbcL		
Pedaliaceae	<i>Harpagophytum</i>	<i>Harpagophytum zeyheri</i>	trnL	KJ743192		
Pedaliaceae	<i>Harpagophytum</i>	<i>Harpagophytum zeyheri</i>	rbcL	KT717180	KT717179	KT717181
Pedaliaceae	<i>Holubia</i>	<i>Holubia saccata</i>	trnL	KJ743197		
Pedaliaceae	<i>Sesamum</i>	<i>Sesamum alatum</i>	trnL	KJ743201		
Pedaliaceae	<i>Sesamum</i>	<i>Sesamum triphyllum</i>	trnL	KJ743202		
Peraceae	<i>Clutia</i>	<i>Clutia pulchella</i>	rbcL	JX572416	AM234976	
Phyllanthaceae	<i>Bridelia</i>	<i>Bridelia cathartica</i>	trnL	MK186997	FJ440011	
Phyllanthaceae	<i>Bridelia</i>	<i>Bridelia cathartica</i>	rbcL	JF265314	JX572350	
Phyllanthaceae	<i>Bridelia</i>	<i>Bridelia mollis</i>	trnL	FJ440006		
Phyllanthaceae	<i>Bridelia</i>	<i>Bridelia mollis</i>	rbcL	JX572351	KNPA342-09.rbcLa	JF265316
Phyllanthaceae	<i>Flueggea</i>	<i>Flueggea virosa</i>	trnL	MK187066		
Phyllanthaceae	<i>Flueggea</i>	<i>Flueggea virosa</i>	rbcL	AY663585	GQ436328	MZ461581
Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus asperulatus</i>	rbcL	MZ461588		

Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus incurvus</i>	trnL	MZ461567		
Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus incurvus</i>	rbcL	MZ461589		
Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus maderaspatensis</i>	trnL	MK261613		
Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus maderaspatensis</i>	rbcL	MK244691	GENG858-14.rbcL	
Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus nummulariifolius</i>	rbcL	AY663609		
Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus reticulatus</i>	trnL	MK187174		
Phyllanthaceae	<i>Thecacoris</i>	<i>Acalypha glabrata</i>	rbcL	JF265264	JX572237	JX572238
Piperaceae	<i>Peperomia</i>	<i>Peperomia blanda</i>	trnL	EU519741		
Piperaceae	<i>Peperomia</i>	<i>Peperomia blanda</i>	rbcL	KM895679		
Plantaginaceae	<i>Bacopa</i>	<i>Bacopa floribunda</i>	rbcL	LC214993		
Plumbaginaceae	<i>Plumbago</i>	<i>Plumbago zeylanica</i>	trnL	AJ312248	MH561153	
Plumbaginaceae	<i>Plumbago</i>	<i>Plumbago zeylanica</i>	rbcL	MG946923	MH582834	LC637763

Poaceae	<i>Andropogon</i>	<i>Andropogon chinensis</i>	rbcL	KNPA380-09.rbcLa		
Poaceae	<i>Andropogon</i>	<i>Andropogon gayanus</i>	rbcL	JQ593287	JQ593288	MF998308
Poaceae	<i>Aristida</i>	<i>Aristida adscensionis</i>	trnL	KR737828	GQ924358	GQ924356
Poaceae	<i>Aristida</i>	<i>Aristida adscensionis</i>	rbcL	EF423002	AM849349	HE573453
Poaceae	<i>Aristida</i>	<i>Aristida congesta</i>	trnL	KR738288	GQ924369	GQ924368
Poaceae	<i>Aristida</i>	<i>Aristida congesta</i>	rbcL	FR821334	KR737225	KR737111
Poaceae	<i>Aristida</i>	<i>Aristida meridionalis</i>	trnL	GQ924405	GQ924295	
Poaceae	<i>Aristida</i>	<i>Aristida meridionalis</i>	rbcL	KNPA217-09.rbcLa	KNPA255-09.rbcLa	
Poaceae	<i>Aristida</i>	<i>Aristida mollissima</i>	rbcL	HG970259		
Poaceae	<i>Aristida</i>	<i>Aristida scabrivalvis</i>	rbcL	KNPA259-09.rbcLa	KNPA275-09.rbcLa	KNPA320-09.rbcLa
Poaceae	<i>Aristida</i>	<i>Aristida stipitata</i>	trnL	GQ924418	GQ924417	
Poaceae	<i>Aristida</i>	<i>Aristida stipitata</i>	rbcL	KNPA236-09.rbcLa	MK570614	

Poaceae	<i>Aristida</i>	<i>Aristida vestita</i>	trnL	GQ92442 5	GQ92442 4	
Poaceae	<i>Aristida</i>	<i>Aristida vestita</i>	rbcL	HG97026 5		
Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa bladhii</i>	trnL	DQ00496 0		
Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa bladhii</i>	rbcL	MF99832 1		
Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa insculpta</i>	trnL	KR73847 5	KR73815 8	KR73771 3
Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa insculpta</i>	rbcL	KNPA31 5- 09.rbcLa	LN90789 1	KNPA456 -09.rbcLa
Poaceae	<i>Bothriochloa</i>	<i>Bothriochloa radicans</i>	rbcL	KNPA30 7- 09.rbcLa	KNPA426 -09.rbcLa	KNPA431 -09.rbcLa
Poaceae	<i>Brachiaria</i>	<i>Brachiaria (Moorochloa) eruciformis</i>	trnL	GU59455 0	KR73835 1	KR73804 7
Poaceae	<i>Brachiaria</i>	<i>Brachiaria (Moorochloa) eruciformis</i>	rbcL	KNPA42 8- 09.rbcLa	KR73685 3	KR73667 1
Poaceae	<i>Brachiaria</i>	<i>Brachiaria (Urochloa) nigropedata</i>	trnL	GU59455 6		
Poaceae	<i>Brachiaria</i>	<i>Brachiaria (Urochloa) xantholeuca</i>	trnL	MK26150 1		

Poaceae	<i>Brachiaria</i>	<i>Brachiaria</i> (<i>Urochloa</i>) <i>xantholeuca</i>	rbcL	HE57584 7	MF99817 3	
Poaceae	<i>Brachiaria</i>	<i>Brachiaria deflexa</i>	trnL	KR73823 0	GU59457 1	KR73807 9
Poaceae	<i>Brachiaria</i>	<i>Brachiaria deflexa</i>	rbcL	KNPA21 6- 09.rbcLa	KNPA280 -09.rbcLa	MF99817 5
Poaceae	<i>Brachiaria</i>	<i>Brachiaria serrata</i>	trnL	GU59455 8		
Poaceae	<i>Brachiaria</i>	<i>Brachiaria serrata</i>	rbcL	KNPA23 0- 09.rbcLa	MF99817 6	KJ740990
Poaceae	<i>Cenchrus</i>	<i>Cenchrus ciliaris</i>	trnL	EU94000 6	MK26138 8	AF49916 8
Poaceae	<i>Cenchrus</i>	<i>Cenchrus ciliaris</i>	rbcL	LN90790 4	IPSUB36 9-18.rbcL	KNPA197 -09.rbcLa
Poaceae	<i>Chloris</i>	<i>Chloris</i> (<i>Tetrapogon</i>) <i>roxburghiana</i>	trnL	MK26149 9	MK26150 0	MK26160 7
Poaceae	<i>Chloris</i>	<i>Chloris</i> (<i>Tetrapogon</i>) <i>roxburghiana</i>	rbcL	KNPA19 5- 09.rbcLa	KNPA196 -09.rbcLa	MF99836 8
Poaceae	<i>Chloris</i>	<i>Chloris gayana</i>	trnL	KR73835 5	MK26170 0	KR73826 4
Poaceae	<i>Chloris</i>	<i>Chloris gayana</i>	rbcL	MF99838 3	AM84940 9	KR73717 8

Poaceae	<i>Chloris</i>	<i>Chloris pycnothrix</i>	trnL	KR73774 0	KR73762 3	KR73856 9
Poaceae	<i>Chloris</i>	<i>Chloris pycnothrix</i>	rbcL	KR73652 2	KR73639 8	KR73741 5
Poaceae	<i>Chloris</i>	<i>Chloris virgata</i>	trnL	EF137560	DQ65585 5	DQ65585 6
Poaceae	<i>Chloris</i>	<i>Chloris virgata</i>	rbcL	KNPA25 4- 09.rbcLa	KNPA305 -09.rbcLa	HE57582 1
Poaceae	<i>Cymbopogon</i>	<i>Cymbopogon excavatus (caesius)</i>	rbcL	LN90791 1	KJ740991	
Poaceae	<i>Cymbopogon</i>	<i>Cymbopogon plurinodis (pospischilii)</i>	rbcL	KNPA25 1- 09.rbcLa	KNPA314 -09.rbcLa	
Poaceae	<i>Cynodon</i>	<i>Cynodon dactylon</i>	trnL	EF156685	DQ65585 4	MG70945 2
Poaceae	<i>Cynodon</i>	<i>Cynodon dactylon</i>	rbcL	MF99836 6	AM84939 3	
Poaceae	<i>Dactyloctenium</i>	<i>Dactyloctenium aegyptium</i>	trnL	EF156686	DQ65586 4	PNG506- 18.trnL-F
Poaceae	<i>Dactyloctenium</i>	<i>Dactyloctenium aegyptium</i>	rbcL	EF125106	AM88787 7	MK28514 6
Poaceae	<i>Dactyloctenium</i>	<i>Dactyloctenium giganteum</i>	rbcL	KNPA28 9- 09.rbcLa		
Poaceae	<i>Dichanthium</i>	<i>Dichanthium annulatum</i>	trnL	DQ00497 3	DQ00497 5	DQ00497 4

Poaceae	<i>Dichanthium</i>	<i>Dichanthium annulatum</i>	rbcL	MF99832 2	MK16679 6	KX28269 4
Poaceae	<i>Digitaria</i>	<i>Digitaria debilis</i>	rbcL	MF99828 3		
Poaceae	<i>Digitaria</i>	<i>Digitaria ternata</i>	trnL	MK26145 0	MK26165 3	MK26163 1
Poaceae	<i>Digitaria</i>	<i>Digitaria ternata</i>	rbcL	MF99828 4		
Poaceae	<i>Digitaria</i>	<i>Digitaria velutina</i>	trnL	KR73770 4	KR73780 0	KR73853 1
Poaceae	<i>Digitaria</i>	<i>Digitaria velutina</i>	rbcL	LN90792 2	KR73648 3	KR73737 3
Poaceae	<i>Diheteropogon</i>	<i>Diheteropogon amplexans</i>	trnL	LN90792 3	FN870803	
Poaceae	<i>Diplachne</i>	<i>Diplachne fusca</i>	trnL	EF156708	DQ65586 2	
Poaceae	<i>Diplachne</i>	<i>Diplachne fusca</i>	rbcL	HE57582 7	KP056888	KP05688 9
Poaceae	<i>Echinochloa</i>	<i>Echinochloa colona (crus-galli)</i>	trnL	KP87910 3	MK79723 7	KF16357 5
Poaceae	<i>Echinochloa</i>	<i>Echinochloa colona (crus-galli)</i>	rbcL	HE57337 7	KT36528 6	KT36528 7
Poaceae	<i>Elionurus</i>	<i>Elionurus muticus</i>	rbcL	MF99831 4	KC12335 8	FN87081 2
Poaceae	<i>Enneapogon</i>	<i>Enneapogon cenchroides</i>	trnL	DQ65589 4	KR73822 4	KR73762 0

Poaceae	<i>Enneapogon</i>	<i>Enneapogon cenchroides</i>	rbcL	MF99841 8	LN90792 9	KR73751 2
Poaceae	<i>Enneapogon</i>	<i>Enneapogon scoparius</i>	trnL	DQ65589 5	GU99036 7	DQ65589 3
Poaceae	<i>Enteropogon</i>	<i>Enteropogon macrostachyus</i>	trnL	DQ65586 3	KR73776 9	KR73802 0
Poaceae	<i>Enteropogon</i>	<i>Enteropogon macrostachyus</i>	rbcL	KR73655 6	KR73660 0	KR73727 1
Poaceae	<i>Eragrostis</i>	<i>Eragrostis aspera</i>	trnL	GU99037 4		
Poaceae	<i>Eragrostis</i>	<i>Eragrostis aspera</i>	rbcL	MF99841 5		
Poaceae	<i>Eragrostis</i>	<i>Eragrostis biflora</i>	trnL	DQ65587 9		
Poaceae	<i>Eragrostis</i>	<i>Eragrostis capensis</i>	trnL	DQ65588 1	DQ65588 2	
Poaceae	<i>Eragrostis</i>	<i>Eragrostis capensis</i>	rbcL	MF99840 5		
Poaceae	<i>Eragrostis</i>	<i>Eragrostis chloromelas</i>	trnL	DQ65587 6	DQ65587 7	
Poaceae	<i>Eragrostis</i>	<i>Eragrostis cilianensis</i>	trnL	GU99038 1	KP711154	MK26144 3
Poaceae	<i>Eragrostis</i>	<i>Eragrostis cilianensis</i>	rbcL	MF59676 5	MG22805 1	MG22823 3
Poaceae	<i>Eragrostis</i>	<i>Eragrostis cylindriflora</i>	trnL	MK26140 4		

Poaceae	<i>Eragrostis</i>	<i>Eragrostis cylindriflora</i>	rbcL	KNPA23 9- 09.rbcLa	KNPA313 -09.rbcLa	KNPA318 -09.rbcLa
Poaceae	<i>Eragrostis</i>	<i>Eragrostis gummiflua</i>	trnL	KNPA27 0- 09.rbcLa		
Poaceae	<i>Eragrostis</i>	<i>Eragrostis heteromera</i>	trnL	GU99039 3	MK26156 2	MK26153 9
Poaceae	<i>Eragrostis</i>	<i>Eragrostis lehmanniana</i>	trnL	DQ65588 3	DQ65587 5	GU99039 8
Poaceae	<i>Eragrostis</i>	<i>Eragrostis lehmanniana</i>	rbcL	KNPA28 1- 09.rbcLa	KNPA298 -09.rbcLa	KNPA432 -09.rbcLa
Poaceae	<i>Eragrostis</i>	<i>Eragrostis racemosa</i>	trnL	DQ65587 8	KR73828 1	KR73866 3
Poaceae	<i>Eragrostis</i>	<i>Eragrostis racemosa</i>	rbcL	LN90794 0	KR73751 7	KR73710 5
Poaceae	<i>Eragrostis</i>	<i>Eragrostis rigidior</i>	trnL	MK26133 6	MK26144 8	MK26171 6
Poaceae	<i>Eragrostis</i>	<i>Eragrostis rotifer</i>	rbcL	KNPA28 3- 09.rbcLa		
Poaceae	<i>Eragrostis</i>	<i>Eragrostis superba</i>	trnL	DQ65588 0	MG70944 4	KR73779 8
Poaceae	<i>Eragrostis</i>	<i>Eragrostis superba</i>	rbcL	KR73658 8	KR73668 5	
Poaceae	<i>Eragrostis</i>	<i>Eragrostis trichophora</i>	trnL	GU99038 6		

Poaceae	<i>Eragrostis</i>	<i>Eragrostis trichophora</i>	rbcL	KNPA250-09.rbcLa	KNPA383-09.rbcLa	KNPA383-09.rbcLa
Poaceae	<i>Eriochloa</i>	<i>Eriochloa meyeriana</i>	trnL	GU594583		
Poaceae	<i>Eustachys</i>	<i>Eustachys paspaloides</i>	trnL	DQ655857	MK261614	
Poaceae	<i>Eustachys</i>	<i>Eustachys paspaloides</i>	rbcL	AM235058	KNPA206-09.rbcLa	KNPA228-09.rbcLa
Poaceae	<i>Fingerhuthia</i>	<i>Fingerhuthia africana</i>	trnL	EF156701	DQ655891	
Poaceae	<i>Fingerhuthia</i>	<i>Fingerhuthia africana</i>	rbcL	HE575825	AM235059	KNPA442-09.rbcLa
Poaceae	<i>Hemarthria</i>	<i>Hemarthria altissima</i>	rbcL	FN870839	GU135209	
Poaceae	<i>Heteropogon</i>	<i>Heteropogon contortus</i>	trnL	KY991322	DQ004984	KR737831
Poaceae	<i>Heteropogon</i>	<i>Heteropogon contortus</i>	rbcL	HE575844	AM235061	KR737296
Poaceae	<i>Hyparrhenia</i>	<i>Hyparrhenia filipendula</i>	trnL	KR737880	KR737709	KR738081
Poaceae	<i>Hyparrhenia</i>	<i>Hyparrhenia filipendula</i>	rbcL	KR736673	KR736488	KR736888
Poaceae	<i>Hyperthelia</i>	<i>Hyperthelia dissoluta</i>	trnL	DQ004985		

Poaceae	<i>Hyperthelia</i>	<i>Hyperthelia dissoluta</i>	rbcL	MF99832 5	KNPA191 -09.rbcLa	KNPA261 -09.rbcLa
Poaceae	<i>Megathyrsus</i>	<i>Megathyrsus maximus</i>	trnL	GU59454 7	KR73805 0	KR73842 3
Poaceae	<i>Megathyrsus</i>	<i>Megathyrsus maximus</i>	rbcL	LN90796 7	FR667682	MH05006 0
Poaceae	<i>Melinis</i>	<i>Melinis nerviglumis</i>	rbcL	KJ740994		
Poaceae	<i>Melinis</i>	<i>Melinis repens</i>	trnL	GU59456 1	KP879139	KR73770 5
Poaceae	<i>Melinis</i>	<i>Melinis repens</i>	rbcL	EF125136	HE57339 7	AM23506 4
Poaceae	<i>Microchloa</i>	<i>Microchloa caffra</i>	rbcL	EF125102	KNPA200 -09.rbcLa	
Poaceae	<i>Oropetium</i>	<i>Oropetium capense</i>	rbcL	HE57583 1	FN870887	
Poaceae	<i>Panicum</i>	<i>Panicum coloratum</i>	trnL	PNG449- 18.trnL-F		
Poaceae	<i>Panicum</i>	<i>Panicum coloratum</i>	rbcL	AM84941 5	KNPA282 -09.rbcLa	KNPA310 -09.rbcLa
Poaceae	<i>Panicum</i>	<i>Panicum deustum</i>	trnL	GU59459 1	MK26173 2	MK26175 2
Poaceae	<i>Panicum</i>	<i>Panicum deustum</i>	rbcL	KNPA22 3- 09.rbcLa	KNPA267 -09.rbcLa	
Poaceae	<i>Panicum</i>	<i>Panicum infestum</i>	rbcL	MF99817 7		

Poaceae	<i>Panicum</i>	<i>Panicum maximum</i>	trnL	KR73846 7	KR73805 0	KR73842 3
Poaceae	<i>Panicum</i>	<i>Panicum maximum</i>	rbcL	KR73729 9	KR73685 6	KR73695 3
Poaceae	<i>Paspalum</i>	<i>Paspalum distichum</i>	trnL	EU87110 0	KX37246 1	
Poaceae	<i>Paspalum</i>	<i>Paspalum distichum</i>	rbcL	FN87039 9	KC16432 8	KF16353 2
Poaceae	<i>Paspalum</i>	<i>Paspalum scrobiculatum</i>	rbcL	LN90799 4	API322- 14.rbcL	
Poaceae	<i>Perotis</i>	<i>Perotis patens</i>	trnL	DQ65586 9		
Poaceae	<i>Perotis</i>	<i>Perotis patens</i>	rbcL	MF99837 5	MF99837 7	FR821340
Poaceae	<i>Phragmites</i>	<i>Phragmites australis</i>	trnL	KC98961 6	EF137518	KX37246 9
Poaceae	<i>Phragmites</i>	<i>Phragmites australis</i>	rbcL	EF423005	KP056891	MF06528 3
Poaceae	<i>Phragmites</i>	<i>Phragmites mauritanus</i>	rbcL	LN90803 3	PNG153- 18.rbcL	
Poaceae	<i>Pogonarthria</i>	<i>Pogonarthria squarrosa</i>	trnL	EF156724	DQ65588 5	DQ65588 4
Poaceae	<i>Pogonarthria</i>	<i>Pogonarthria squarrosa</i>	rbcL	MF99840 3	JN681701	KNPA215 -09.rbcLa
Poaceae	<i>Schmidtia</i>	<i>Schmidtia pappophoroides</i>	trnL	GU99036 8	DQ65589 6	

Poaceae	<i>Schmidtia</i>	<i>Schmidtia pappophoroides</i>	rbcL	KNPA21 4- 09.rbcLa	KNPA257 -09.rbcLa	KNPA271 -09.rbcLa
Poaceae	<i>Setaria</i>	<i>Setaria incrassata</i>	rbcL	LN90800 8	KNPA233 -09.rbcLa	KNPA265 -09.rbcLa
Poaceae	<i>Setaria</i>	<i>Setaria pumila</i>	trnL	KX37248 7	MK26154 8	MK26157 6
Poaceae	<i>Setaria</i>	<i>Setaria pumila</i>	rbcL	KC16433 8	KF163543	KJ841567
Poaceae	<i>Setaria</i>	<i>Setaria sagittifolia</i>	trnL	PNG430- 18.trnL-F		
Poaceae	<i>Setaria</i>	<i>Setaria sagittifolia</i>	rbcL	MF99826 0	KNPA222 -09.rbcLa	KNPA296 -09.rbcLa
Poaceae	<i>Setaria</i>	<i>Setaria sphacelata</i>	trnL	GU56149 4	KR73763 1	KR73769 5
Poaceae	<i>Setaria</i>	<i>Setaria sphacelata</i>	rbcL	HE57788 0	LN90801 1	KR73647 2
Poaceae	<i>Setaria</i>	<i>Setaria verticillata</i>	trnL	KX37248 9	AF499154	KR73847 7
Poaceae	<i>Setaria</i>	<i>Setaria verticillata</i>	rbcL	LN90801 2	FR821349	HQ60036 6
Poaceae	<i>Sorghum</i>	<i>Sorghum bicolor</i>	trnL	KX37249 1	GQ86766 3	
Poaceae	<i>Sorghum</i>	<i>Sorghum bicolor</i>	rbcL	MG22751 3	FSSM057 -14.rbcLa	SDH2517 -14.rbcLa
Poaceae	<i>Sorghum</i>	<i>Sorghum versicolor</i>	rbcL	MG77277 2	KNPA234 -09.rbcLa	KNPA404 -09.rbcLa

Poaceae	<i>Sporobolus</i>	<i>Sporobolus africanus</i>	trnL	DQ65589 9	KR73794 6	KR73865 0
Poaceae	<i>Sporobolus</i>	<i>Sporobolus africanus</i>	rbcL	HM85038 3	KR73638 6	KR73750 2
Poaceae	<i>Sporobolus</i>	<i>Sporobolus festinus</i>	trnL	KR73815 0	KR73812 3	MK26155 7
Poaceae	<i>Sporobolus</i>	<i>Sporobolus festinus</i>	rbcL	AM84938 3	MF99839 7	KR73696 6
Poaceae	<i>Sporobolus</i>	<i>Sporobolus fimbriatus</i>	trnL	DQ65590 1	DQ65590 3	
Poaceae	<i>Sporobolus</i>	<i>Sporobolus fimbriatus</i>	rbcL	KNPA24 8- 09.rbcLa	KNPA276 -09.rbcLa	KNPA336 -09.rbcLa
Poaceae	<i>Sporobolus</i>	<i>Sporobolus ioclados</i>	trnL	KR73783 7	KR73789 3	KR73846 2
Poaceae	<i>Sporobolus</i>	<i>Sporobolus ioclados</i>	rbcL	KR73729 3	KR73714 9	KR73663 0
Poaceae	<i>Sporobolus</i>	<i>Sporobolus nitens</i>	rbcL	KNPA20 3- 09.rbcLa	KNPA213 -09.rbcLa	KNPA392 -09.rbcLa
Poaceae	<i>Sporobolus</i>	<i>Sporobolus pyramidalis</i>	trnL	DQ65590 2	KR73824 4	
Poaceae	<i>Sporobolus</i>	<i>Sporobolus pyramidalis</i>	rbcL	LN90801 7	MF99839 6	MH46434 8
Poaceae	<i>Sporobolus</i>	<i>Sporobolus stapfianus</i>	trnL	KR73819 0	KR73827 9	KR73767 7

Poaceae	<i>Sporobolus</i>	<i>Sporobolus stapfianus</i>	rbcL	KR73659 3	KR73700 9	KR73710 3
Poaceae	<i>Stipagrostis</i>	<i>Stipagrostis hirtigluma</i>	trnL	GQ92433 0	GQ92433 3	GQ92433 1
Poaceae	<i>Themeda</i>	<i>Themeda triandra</i>	trnL	KY99134 3	KY99132 3	KY99133 3
Poaceae	<i>Themeda</i>	<i>Themeda triandra</i>	rbcL	LN90802 2	KF163556	KF16355 7
Poaceae	<i>Trachypogon</i>	<i>Trachypogon spicatus</i>	rbcL	EF125123	LN90802 3	MG77277 3
Poaceae	<i>Tragus</i>	<i>Tragus berteronianus</i>	trnL	DQ65586 5	KR73771 0	KR73837 1
Poaceae	<i>Tragus</i>	<i>Tragus berteronianus</i>	rbcL	MF99838 8	MN18526 4	KR73680 1
Poaceae	<i>Tricholaena</i>	<i>Tricholaena monachne</i>	trnL	GU59457 4		
Poaceae	<i>Tricholaena</i>	<i>Tricholaena monachne</i>	rbcL	MF99818 4	HE57788 1	
Poaceae	<i>Trichoneura</i>	<i>Trichoneura grandiglumis</i>	rbcL	HE57583 5	KC12338 8	FN87094 7
Poaceae	<i>Tripogon</i>	<i>Tripogon minimus</i>	rbcL	AM84934 2		
Poaceae	<i>Urochloa</i>	<i>Urochloa brachyura</i>	trnL	KR73765 6	KR73814 4	KR73818 5
Poaceae	<i>Urochloa</i>	<i>Urochloa brachyura</i>	rbcL	MF99816 6	KR73643 5	KR73700 4

Poaceae	<i>Urochloa</i>	<i>Urochloa mosambicensis</i>	trnL	GU594532	PNG431-18.trnL-F	
Poaceae	<i>Urochloa</i>	<i>Urochloa mosambicensis</i>	rbcL	MF998168	KNPA247-09.rbcLa	KNPA278-09.rbcLa
Poaceae	<i>Urochloa</i>	<i>Urochloa panicoides</i>	rbcL	LT608220	MF998167	HE573318
Polygalaceae	<i>Polygala</i>	<i>Polygala erioptera</i>	trnL	AF366990	GQ889131	AM234331
Polygalaceae	<i>Polygala</i>	<i>Polygala erioptera</i>	rbcL	AM234219	KU551064	
Polygalaceae	<i>Polygala</i>	<i>Polygala sphenoptera</i>	trnL	KR737872	KR738419	KR738325
Polygalaceae	<i>Polygala</i>	<i>Polygala sphenoptera</i>	rbcL	KR736665	KR737150	KR737085
Polygalaceae	<i>Polygala</i>	<i>Polygala uncinata</i>	trnL	GQ889215	AM234318	GQ888867
Polygalaceae	<i>Polygala</i>	<i>Polygala uncinata</i>	rbcL	AM234192		
Polygonaceae	<i>Oxygonum</i>	<i>Oxygonum dregeanum</i>	trnL	JN161140		
Polygonaceae	<i>Oxygonum</i>	<i>Oxygonum sinuatum</i>	trnL	KR537752	KR738018	KR737905
Polygonaceae	<i>Oxygonum</i>	<i>Oxygonum sinuatum</i>	rbcL	KR736460	KR736970	KR736745
Polygonaceae	<i>Persicaria</i>	<i>Persicaria decipiens</i>	trnL	KR737613	KR738471	KR738613

Polygonaceae	<i>Persicaria</i>	<i>Persicaria decipiens</i>	rbcL	KR73746 1	KR73649 6	KR73730 3
Portulacaceae	<i>Portulaca</i>	<i>Portulaca oleracea</i>	rbcL	MG94685 8	HQ62134 0	AY87524 9
Portulacaceae	<i>Portulaca</i>	<i>Portulaca quadrifida</i>	trnL	MK26172 0	MK26157 7	KR73847 2
Portulacaceae	<i>Portulaca</i>	<i>Portulaca quadrifida</i>	rbcL	KR73701 0	KR73730 4	KR73695 7
Pteridaceae	<i>Cheilanthes</i>	<i>Cheilanthes viridis</i>	trnL	GU93560 3	GU93560 1	
Pteridaceae	<i>Pellaea</i>	<i>Pellaea calomelanos</i>	trnL	KR73766 9	KR73768 8	KR73767 5
Pteridaceae	<i>Pellaea</i>	<i>Pellaea calomelanos</i>	rbcL	KR73644 7	KR73646 5	KR73645 3
Rhamnaceae	<i>Berchemia</i>	<i>Berchemia discolor</i>	rbcL	KNPA18 1- 08.rbcLa	JF265302	JX572329
Rhamnaceae	<i>Berchemia</i>	<i>Berchemia zeyheri</i>	rbcL	JF265303	JX572330	
Rhamnaceae	<i>Ziziphus</i>	<i>Ziziphus mucronata</i>	trnL	KR08315 1	MK26155 1	
Rhamnaceae	<i>Ziziphus</i>	<i>Ziziphus mucronata</i>	rbcL	JF265666	JX573107	
Rubiaceae	<i>Agathisanthemum</i>	<i>Agathisanthemum bojeri</i>	trnL	EU54307 7		
Rubiaceae	<i>Agathisanthemum</i>	<i>Agathisanthemum bojeri</i>	rbcL	Z68787		

Rubiaceae	<i>Breonadia</i>	<i>Breonadia salicina</i>	rbcL	JF265313	AJ346967	JX572348
Rubiaceae	<i>Canthium</i>	<i>Canthium armatum</i>	trnL	AJ620155	JQ958229	
Rubiaceae	<i>Canthium</i>	<i>Canthium armatum</i>	rbcL	AM117259	JF265553	JX572859
Rubiaceae	<i>Canthium</i>	<i>Canthium ciliatum</i>	trnL	JQ958230	AJ620121	
Rubiaceae	<i>Canthium</i>	<i>Canthium ciliatum</i>	rbcL	JX572365		
Rubiaceae	<i>Catunaregam</i>	<i>Catunaregam spinosa</i>	rbcL	JF265332	KNPA153-08.rbcLa	KX910827
Rubiaceae	<i>Conostomium</i>	<i>Conostomium natalense</i>	trnL	EU543085		
Rubiaceae	<i>Conostomium</i>	<i>Conostomium natalense</i>	rbcL	EU542927		
Rubiaceae	<i>Gardenia</i>	<i>Gardenia volkensii</i>	rbcL	JF265446	JX572624	
Rubiaceae	<i>Hyperacanthus</i>	<i>Hyperacanthus amoenus</i>	trnL	KF965209		
Rubiaceae	<i>Hyperacanthus</i>	<i>Hyperacanthus amoenus</i>	rbcL	KF964840	JF265481	JX572687
Rubiaceae	<i>Kohautia</i>	<i>Kohautia amatymbica</i>	trnL	EU543117	GU951652	
Rubiaceae	<i>Kohautia</i>	<i>Kohautia caespitosa</i>	trnL	EU543118	GU951659	
Rubiaceae	<i>Kohautia</i>	<i>Kohautia caespitosa</i>	rbcL	GU951593	GU951592	FJ695364
Rubiaceae	<i>Kohautia</i>	<i>Kohautia cynanchica</i>	trnL	EU543120		

Rubiaceae	<i>Kohautia</i>	<i>Kohautia cynanchica</i>	rbcL	EU54296 5		
Rubiaceae	<i>Kohautia</i>	<i>Kohautia virgata</i>	trnL	EU54312 4		
Rubiaceae	<i>Kohautia</i>	<i>Kohautia virgata</i>	rbcL	EU54296 9		
Rubiaceae	<i>Kraussia</i>	<i>Kraussia floribunda</i>	trnL	KF96521 6	AM11736 8	KT21892 4
Rubiaceae	<i>Kraussia</i>	<i>Kraussia floribunda</i>	rbcL	KF96484 8		
Rubiaceae	<i>Pavetta</i>	<i>Pavetta catophylla</i>	rbcL	JF265542	JX572837	KNPA160 -08.rbcLa
Rubiaceae	<i>Pavetta</i>	<i>Pavetta lanceolata</i>	trnL	AM11737 7		
Rubiaceae	<i>Pavetta</i>	<i>Pavetta lanceolata</i>	rbcL	JF265544	JX572841	
Rubiaceae	<i>Pavetta</i>	<i>Pavetta schumanniana</i>	trnL	KM59213 5	JN053647	JN053646
Rubiaceae	<i>Pavetta</i>	<i>Pavetta schumanniana</i>	rbcL	JX572843	JF265545	KNPA106 -08.rbcLa
Rubiaceae	<i>Pentodon</i>	<i>Pentodon pentandrus</i>	trnL	EU54315 4		
Rubiaceae	<i>Psydrax</i>	<i>Psydrax obovata</i>	trnL	AJ620161	JQ958217	
Rubiaceae	<i>Psydrax</i>	<i>Psydrax obovata</i>	rbcL	HM16417 6	JX572899	
Rubiaceae	<i>Pyrostria</i>	<i>Pyrostria hystrix</i>	trnL	AJ620168	JQ958226	

Rubiaceae	<i>Pyrostria</i>	<i>Pyrostria hystrix</i>	rbcL	AM11726 2	JF265568	JX572915
Rubiaceae	<i>Richardia</i>	<i>Richardia brasiliensis</i>	rbcL	KJ773839	KY62689 3	
Rubiaceae	<i>Tricalysia</i>	<i>Tricalysia junodii</i>	trnL	KT21895 2		
Rubiaceae	<i>Tricalysia</i>	<i>Tricalysia junodii</i>	rbcL	JF265633	JX573060	KNPA169 -08.rbcLa
Rubiaceae	<i>Vangueria</i>	<i>Vangueria infausta</i>	trnL	JQ958289	KF488321	
Rubiaceae	<i>Vangueria</i>	<i>Vangueria infausta</i>	rbcL	DQ13179 4	JF265644	JX573076
Rutaceae	<i>Clausena</i>	<i>Clausena anisata</i>	trnL	MK26153 5		
Rutaceae	<i>Clausena</i>	<i>Clausena anisata</i>	rbcL	MK28524 0		
Rutaceae	<i>Ptaeroxylon</i>	<i>Ptaeroxylon obliquum</i>	rbcL	FR74783 8	JF265562	
Rutaceae	<i>Vepris</i>	<i>Vepris carringtoniana</i>	rbcL	KU56799 4		
Rutaceae	<i>Zanthoxylum</i>	<i>Zanthoxylum humile</i>	rbcL	JX573103	JF265664	
Salicaceae	<i>Dovyalis</i>	<i>Dovyalis caffra</i>	rbcL	JF265397	JQ412354	JX572533
Salicaceae	<i>Flacourtia</i>	<i>Flacourtia indica</i>	rbcL	MH33246 7	MH54985 9	GU13521 8
Salicaceae	<i>Scolopia</i>	<i>Scolopia zeyheri</i>	rbcL	MT23141 7	JF265587	

Sapindaceae	<i>Allophylus</i>	<i>Allophylus decipiens</i>	trnL	KX58499 2		
Sapindaceae	<i>Allophylus</i>	<i>Allophylus decipiens</i>	rbcL	JF265283		
Sapindaceae	<i>Cardiospermum</i>	<i>Cardiospermum corindum</i>	trnL	KX58500 7		
Sapindaceae	<i>Cardiospermum</i>	<i>Cardiospermum halicacabum</i>	trnL	AY59450 4	JN681469	MK18700 2
Sapindaceae	<i>Cardiospermum</i>	<i>Cardiospermum halicacabum</i>	rbcL	GU93544 8	KJ773336	MK34896 2
Sapotaceae	<i>Manilkara</i>	<i>Manilkara mochisia</i>	rbcL	JF265514	KNPA122 -08.rbcLa	PNG408- 18.rbcL
Scrophulariaceae	<i>Aptosimum</i>	<i>Aptosimum lineare</i>	trnL	MZ46154 8		
Scrophulariaceae	<i>Aptosimum</i>	<i>Aptosimum lineare</i>	rbcL	MZ46157 2		
Solanaceae	<i>Datura</i>	<i>Datura stramonium</i>	trnL	MK26136 3		
Solanaceae	<i>Datura</i>	<i>Datura stramonium</i>	rbcL	MK28523 5		
Solanaceae	<i>Solanum</i>	<i>Solanum aculeastrum</i>	trnL	DQ81210 2	HQ72189 6	HM01646 0
Solanaceae	<i>Solanum</i>	<i>Solanum aculeastrum</i>	rbcL	JQ025091		
Solanaceae	<i>Solanum</i>	<i>Solanum catombelense</i>	trnL	HQ72190 9		

Solanaceae	<i>Solanum</i>	<i>Solanum catombelense</i>	rbcL	JF265599		
Solanaceae	<i>Solanum</i>	<i>Solanum incanum</i>	trnL	EU17614 4	GQ16352 2	HQ72192 4
Solanaceae	<i>Solanum</i>	<i>Solanum lichtensteinii</i>	trnL	EU17614 6	EU17614 7	KU71979 8
Solanaceae	<i>Solanum</i>	<i>Solanum lichtensteinii</i>	rbcL	JF265600	SLNM001 -19.rbcL	SLNM02 1-19.rbcL
Solanaceae	<i>Solanum</i>	<i>Solanum nigrum</i>	trnL	MK26142 4	MK26169 1	MK26171 1
Solanaceae	<i>Solanum</i>	<i>Solanum nigrum</i>	rbcL	MK28527 9	MK28527 7	MK28528 0
Solanaceae	<i>Solanum</i>	<i>Solanum panduriforme</i>	trnL	EU17614 3	MK18721 7	MK18721 8
Solanaceae	<i>Solanum</i>	<i>Solanum panduriforme</i>	rbcL	EU21351 3	EU21351 4	EU21351 5
Solanaceae	<i>Solanum</i>	<i>Solanum seaforthianum</i>	rbcL	MH72238 0	SLNM016 -19.rbcL	
Solanaceae	<i>Solanum</i>	<i>Solanum tettense</i>	trnL	HQ72194 9		
Solanaceae	<i>Withania</i>	<i>Withania somnifera</i>	trnL	EU58106 9	KR73766 2	KR73823 2
Solanaceae	<i>Withania</i>	<i>Withania somnifera</i>	rbcL	MG94690 5	FJ914179	KY65670 8
Stilbaceae	<i>Nuxia</i>	<i>Nuxia oppositifolia</i>	trnL	AJ621117		
Stilbaceae	<i>Nuxia</i>	<i>Nuxia oppositifolia</i>	rbcL	JF265527	JX572793	

Talinaceae	<i>Talinum</i>	<i>Talinum caffrum</i>	trnL	KR73861 0	KR73865 5	
Talinaceae	<i>Talinum</i>	<i>Talinum caffrum</i>	rbcL	KR73680 0	KR73750 9	KR73671 5
Talinaceae	<i>Talinum</i>	<i>Talinum portulacifolium</i>	trnL	KM26195 8	MK26138 2	
Urticaceae	<i>Pouzolzia</i>	<i>Pouzolzia mixta</i>	rbcL	JQ025073		
Vahliaceae	<i>Vahlia</i>	<i>Vahlia capensis</i>	trnL	HQ41298 4	MK18724 0	
Vataceae	<i>Rhoicissus</i>	<i>Rhoicissus revoilii</i>	trnL	JQ182555		
Vataceae	<i>Rhoicissus</i>	<i>Rhoicissus revoilii</i>	rbcL	JQ182456	JX572926	JF265572
Vataceae	<i>Rhoicissus</i>	<i>Rhoicissus tridentata</i>	trnL	JF437341		
Vataceae	<i>Rhoicissus</i>	<i>Rhoicissus tridentata</i>	rbcL	JQ182443	JF265574	JQ025083
Velloziaceae	<i>Xerophyta</i>	<i>Xerophyta retinervis</i>	trnL	JN016936	JX287056	
Velloziaceae	<i>Xerophyta</i>	<i>Xerophyta retinervis</i>	rbcL	KT20481 7	JX032732	EU21353 0
Verbenaceae	<i>Chascanum</i>	<i>Chascanum pinnatifidum</i>	rbcL	KBGPP0 04- 18.rbcLa		
Verbenaceae	<i>Lantana</i>	<i>Lantana camara</i>	trnL	HM21663 3	AF231884	
Verbenaceae	<i>Lantana</i>	<i>Lantana camara</i>	rbcL	AF15673 6	HM85010 4	

Verbenaceae	<i>Lantana</i>	<i>Lantana rugosa</i>	trnL	HM21663 7		
Verbenaceae	<i>Lantana</i>	<i>Lantana rugosa</i>	rbcL	JF265500	JX572712	KBGPP07 9- 18.rbcLa
Verbenaceae	<i>Lippia</i>	<i>Lippia javanica</i>	trnL	HM21664 4	KR73764 7	KR73772 0
Verbenaceae	<i>Lippia</i>	<i>Lippia javanica</i>	rbcL	KR73651 2	JF265503	JQ025061
Verbenaceae	<i>Verbena</i>	<i>Verbena bonariensis</i>	trnL	EF571518	KR73805 7	KR73849 2
Verbenaceae	<i>Verbena</i>	<i>Verbena bonariensis</i>	rbcL	HM85044 3	KF724243	KR73732 7
Violaceae	<i>Hybanthus</i>	<i>Hybanthus enneaspermus</i>	trnL	MK26154 9	MK26158 2	
Violaceae	<i>Hybanthus</i>	<i>Hybanthus enneaspermus</i>	rbcL	AB35441 8	KC69958 4	GENG10 60- 15.rbcL
Vitaceae	<i>Cissus</i>	<i>Cissus cornifolia</i>	trnL	JF437308	JX476869	JX476870
Vitaceae	<i>Cissus</i>	<i>Cissus cornifolia</i>	rbcL	JX476449	KNPA583 -09.rbcLa	JF265339
Vitaceae	<i>Cissus</i>	<i>Cissus quadrangularis</i>	trnL	JX476898	JF437313	JX313444
Vitaceae	<i>Cissus</i>	<i>Cissus quadrangularis</i>	trnL	AJ419720	KX95108 4	MN12563 6
Vitaceae	<i>Cissus</i>	<i>Cissus rotundifolia</i>	trnL	JF437314	MK18701 7	KR73771 8

Vitaceae	<i>Cissus</i>	<i>Cissus rotundifolia</i>	rbcL	MF34968 6	GENG171 3-16.rbcL	PNG311- 18.rbcL
Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma cirrhosum</i>	trnL	MH08694 6		
Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma cirrhosum</i>	rbcL	MH08688 6	MH08683 5	
Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma schlechteri</i>	trnL	MH08693 6		
Vitaceae	<i>Cyphostemma</i>	<i>Cyphostemma schlechteri</i>	rbcL	MH08687 6	MH08682 6	
Ximeniaceae	<i>Ximenia</i>	<i>Ximenia americana</i>	rbcL	GQ99789 8	MH39066 8	DQ79014 9
Ximeniaceae	<i>Ximenia</i>	<i>Ximenia caffra</i>	rbcL	MN16669 3	MN21645 3	MN21648 7
Zygophyllace ae	<i>Balanites</i>	<i>Balanites maughamii</i>	rbcL	JX572324		
Zygophyllace ae	<i>Tribulus</i>	<i>Tribulus terrestris</i>	trnL	MN58721 5	KP087733	KT37726 6
Zygophyllace ae	<i>Tribulus</i>	<i>Tribulus terrestris</i>	rbcL	MG94687 4	AM23516 7	MN52577 9

Supplementary Table 4: Species sequenced at the NWU with the *rbcL* and *trnL* barcode which have been submitted to GenBank.

Barcode	Family	Genus	Species	Genbank accession
	Malvaceae	<i>Abutilon</i>	<i>Abutilon austroafricanum</i>	MZ461571

rbcL (TOTAL: 24)	Scrophulariaceae	<i>Aptosimum</i>	<i>Aptosimum lineare</i>	MZ461572
	Acanthaceae	<i>Barleria</i>	<i>Barleria elegans</i>	MZ461573
	Acanthaceae	<i>Barleria</i>	<i>Barleria prionitis</i>	MZ461574
	Acanthaceae	<i>Blepharis</i>	<i>Blepharis innocua</i>	MZ461575
	Acanthaceae	<i>Blepharis</i>	<i>Blepharis integrifolia</i>	MZ461576
	Fabaceae	<i>Chamaecrista</i>	<i>Chamaecrista mimosoides</i>	MZ461577
	Combretaceae	<i>Combretum</i>	<i>Combretum zeyheri</i>	MZ461578
	Acanthaceae	<i>Crabbea</i>	<i>Crabbea velutina</i>	MZ461579
	Acanthaceae	<i>Ecbolium</i>	<i>Ecbolium glabratum</i>	MZ461580
	Phyllanthaceae	<i>Flueggea</i>	<i>Flueggea virosa</i>	MZ461581
	Malvaceae	<i>Grewia</i>	<i>Grewia bicolor</i>	MZ461582
	Malvaceae	<i>Hibiscus</i>	<i>Hibiscus sabiensis</i>	MZ461583
	Acanthaceae	<i>Justicia</i>	<i>Justicia flava</i>	MZ461584
	Acanthaceae	<i>Justicia</i>	<i>Justicia protracta</i>	MZ461585
	Lamiaceae	<i>Leucas</i>	<i>Leucas glabrata</i>	MZ461586
	Malvaceae	<i>Melhania</i>	<i>Melhania prostrata</i>	MZ461587
	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus asperulatus</i>	MZ461588
	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus incurvus</i>	MZ461589

	Portulacaceae	<i>Portulaca</i>	<i>Portulaca quadrifida</i>	MZ461590
	Bignoniaceae	<i>Rhigozum</i>	<i>Rhigozum zambesiicum</i>	MZ461591
	Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia fleckii</i>	MZ461592
	Solanaceae	<i>Solanum</i>	<i>Solanum delagoense</i>	MZ461593
	Euphorbiaceae	<i>Tragia</i>	<i>Tragia dioica</i>	MZ461594
trnL (TOTAL: 24)	Malvaceae	<i>Abutilon</i>	<i>Abutilon austroafricanum</i>	MZ461547
	Scrophulariaceae	<i>Aptosimum</i>	<i>Aptosimum lineare</i>	MZ461548
	Acanthaceae	<i>Barleria</i>	<i>Barleria elegans</i>	MZ461549
	Acanthaceae	<i>Barleria</i>	<i>Barleria prionitis</i>	MZ461550
	Acanthaceae	<i>Blepharis</i>	<i>Blepharis innocua</i>	MZ461551
	Pedaliaceae	<i>Ceratotheca</i>	<i>Ceratotheca triloba</i>	MZ461552
	Combretaceae	<i>Combretum</i>	<i>Combretum zeyheri</i>	MZ461553
	Burseraceae	<i>Commiphora</i>	<i>Commiphora africana</i>	MZ461554
	Acanthaceae	<i>Crabbea</i>	<i>Crabbea velutina</i>	MZ461555
	Acanthaceae	<i>Ecboium</i>	<i>Ecboium glabratum</i>	MZ461556
	Celastraceae	<i>Elaeodendron</i>	<i>Elaeodendron transvaalense</i>	MZ461557

	Convolvulaceae	<i>Evolvulus</i>	<i>Evolvulus alsinoides</i>	MZ461558
	Asteraceae	<i>Geigeria</i>	<i>Geigeria burkei</i>	MZ461559
	Fabaceae	<i>Indigofera</i>	<i>Indigofera tinctoria</i>	MZ461560
	Convolvulaceae	<i>Ipomoea</i>	<i>Ipomoea crassipes</i>	MZ461561
	Acanthaceae	<i>Justicia</i>	<i>Justicia flava</i>	MZ461562
	Acanthaceae	<i>Justicia</i>	<i>Justicia protracta</i>	MZ461563
	Malvaceae	<i>Melhania</i>	<i>Melhania prostrata</i>	MZ461564
	Convolvulaceae	<i>Xenostegia</i>	<i>Xenostegia tridentata</i>	MZ461565
	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus asperulatus</i>	MZ461566
	Phyllanthaceae	<i>Phyllanthus</i>	<i>Phyllanthus incurvus</i>	MZ461567
	Fabaceae	<i>Rhynchosia</i>	<i>Rhynchosia fleckii</i>	MZ461568
	Convolvulaceae	<i>Seddera</i>	<i>Seddera capensis</i>	MZ461569
	Malvaceae	<i>Triumfetta</i>	<i>Triumfetta pentandra</i>	MZ461570

Annexure 3: Introducing an *rbcL* and *trnL* reference library to aid in the metabarcoding analysis of foraged plants from semi-arid eastern South African savannas

bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.06.511093>; this version posted October 7, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 Introducing an *rbcL* and *trnL* reference library to aid in the metabarcoding analysis of
2 foraged plants from semi-arid eastern South African savannas

3 Danielle Botha¹, Mornè du Plessis², Frances Siebert¹, Sandra Barnard¹

4 ¹Unit for Environmental Sciences and Management, North-West University,
5 Potchefstroom, South Africa.

6 ²Core Sequencing Facility, National Institute for Communicable Diseases of the
7 National Health Laboratory Service, Sandringham, Johannesburg, South Africa

8

9 *Corresponding author

10 E:mail: sandra.barnard@nwu.ac.za

11 &SB and FS conceived and designed the research and conducted field collections;

12 FS conducted taxonomic identifications; DB and MduP performed analyses. DB

13 drafted the manuscript with subsequent contributions from all authors.

Annexure 4: Supplementary tables supporting Chapter 4

Supplementary Table 5: Comparison criteria of the filtering methods applied to the rbcL mock community dataset. Values of the means for samples and their replicates are used to calculate the final value for the comparison indices. Methods appear in the order of highest precision, lowest RMSE and lastly highest sensitivity.

Method	Richness	Sensitivity	Precision	RMSE
DADA_1_RA_LULU	64,000000	0,195652	0,680455	11,213810
DADA_1_LULU	64,000000	0,195652	0,680455	11,213810
DADA_1_RA	69,000000	0,197826	0,641544	11,151909
DADA_1	69,000000	0,197826	0,641544	11,151909
DADA_10_LULU	103,000000	0,232609	0,614730	13,064257
DADA_10_RA_LULU	103,000000	0,232609	0,614730	13,064257
DADA_ALL_LULU	103,000000	0,232609	0,614730	13,064257
DADA_RA_LULU	103,000000	0,232609	0,614730	13,064257
DADA_10	672,000000	0,571739	0,361480	5,302324
DADA_ALL	672,000000	0,571739	0,361480	5,302324
DADA_RA	672,000000	0,571739	0,361480	5,302324
DADA_10_RA	672,000000	0,571739	0,361480	5,302324
VSEARCH_DN_Q_10_RA_LULU	297,000000	0,739130	0,063550	3,285872
VSEARCH_DN_Q_RA_LULU	297,000000	0,739130	0,063550	3,286073
VSEARCH_DN_Q_1_RA_LULU	297,000000	0,739130	0,063550	3,286073
VSEARCH_DN_Q_10_LULU	494,000000	0,739130	0,046289	3,286615

VSEARCH_DN_Q_RA	539,000000	0,869565	0,041357	2,679179
VSEARCH_DN_Q_1_RA	539,000000	0,869565	0,041357	2,679179
VSEARCH_DN_Q_10_RA	539,000000	0,869565	0,041357	2,679179
VSEARCH_DN_Q_LULU	2170,000000	0,739130	0,033721	3,287546
VSEARCH_DN_Q_1_LULU	961,000000	0,739130	0,033721	3,287546
VSEARCH_DN_Q_10	994,000000	0,869565	0,028137	2,679178
VSEARCH_DN_Q_1	2173,000000	0,869565	0,022238	2,679179
VSEARCH_DN_Q	5116,000000	0,869565	0,019114	2,679179
VSEARCH_DN_RA_LULU	1308,000000	0,956522	0,019087	2,093366
VSEARCH_DN_1_RA_LULU	1342,000000	0,956522	0,018678	2,092389
VSEARCH_DN_10_RA_LULU	1377,000000	0,956522	0,018269	2,092389
VSEARCH_DN_RA	1582,000000	0,956522	0,015921	4,480015
VSEARCH_DN_1_RA	1634,000000	0,956522	0,015479	2,115181
VSEARCH_DN_10_RA	1688,000000	0,956522	0,015056	2,115181
VSEARCH_DN_10_LULU	2941,000000	0,956522	0,010836	2,086349
VSEARCH_DN_1_LULU	6238,000000	0,956522	0,008472	2,084453
VSEARCH_DN_10	4530,000000	0,956522	0,007580	2,115181
VSEARCH_DN_ALL_LULU	19620,000000	0,956522	0,006739	2,077301
VSEARCH_DN_1	14656,000000	0,956522	0,004881	2,115181
VSEARCH_DN_ALL	69384,000000	0,956522	0,003036	2,115181

Supplementary Table 6: Comparison criteria of the filtering methods applied to the trnL mock community dataset. Values of the means for samples and their replicates are used to calculate the final value for the comparison indices. Methods appear in the order of highest precision, lowest RMSE and lastly highest sensitivity.

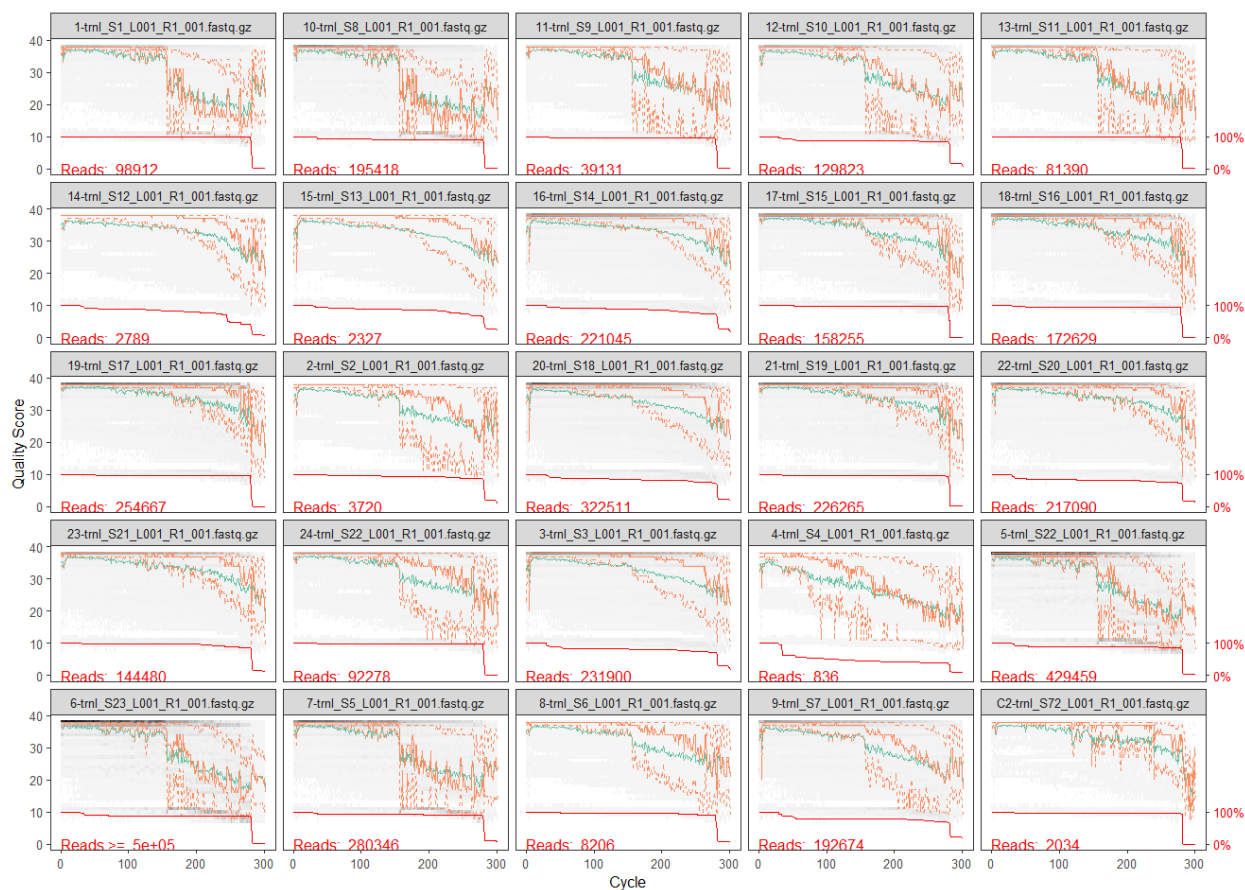
Method	Richness	Sensitivity	Precision	RMSE
DADA_1_LULU	132	0,361	0,649	9,628
DADA_1_RA_LULU	132	0,361	0,649	9,628
DADA_1	135	0,361	0,630	7,351
DADA_1_RA	135	0,361	0,630	7,351
DADA_ALL_LULU	512	0,720	0,528	4,172
DADA_RA_LULU	511	0,717	0,528	4,164
DADA_10_LULU	505	0,709	0,527	4,172
DADA_10_RA_LULU	505	0,709	0,527	4,172
DADA_ALL	1057	0,811	0,359	3,655
DADA_RA	1055	0,809	0,359	3,655
DADA_10_RA	1046	0,798	0,358	3,663
DADA_10	1046	0,798	0,358	3,663
VSEARCH_DN_Q_1_RA_LULU	228	1,000	0,105	1,215
VSEARCH_DN_Q_RA_LULU	228	1,000	0,105	1,215
VSEARCH_DN_Q_10_LULU	317	1,000	0,086	1,036
VSEARCH_DN_Q_10_RA_LULU	317	1,000	0,086	1,215
VSEARCH_DN_Q_1_RA	339	1,000	0,072	1,229

VSEARCH_DN_Q_RA	339	1,000	0,072	1,229
VSEARCH_DN_Q_LULU	1797	1,000	0,062	1,215
VSEARCH_DN_Q_1_LULU	601	1,000	0,062	1,215
VSEARCH_DN_Q_10	525	1,000	0,054	1,047
VSEARCH_DN_Q_10_RA	525	1,000	0,054	1,229
VSEARCH_DN_Q_1	1152	1,000	0,044	1,229
VSEARCH_DN_Q	3093	1,000	0,037	1,229
VSEARCH_DN_RA_LULU	996	1,000	0,026	2,351
VSEARCH_DN_1_RA_LULU	996	1,000	0,026	2,201
VSEARCH_DN_10_RA_LULU	1010	1,000	0,025	2,201
VSEARCH_DN_RA	1183	1,000	0,022	2,386
VSEARCH_DN_1_RA	1226	1,000	0,021	2,239
VSEARCH_DN_10_RA	1242	1,000	0,021	2,239
VSEARCH_DN_10_LULU	2158	1,000	0,015	2,195
VSEARCH_DN_1_LULU	4696	1,000	0,012	2,193
VSEARCH_DN_10	3286	1,000	0,011	2,240
VSEARCH_DN_ALL_LULU	21264	1,000	0,008	2,211
VSEARCH_DN_1	10532	1,000	0,007	2,240
VSEARCH_DN_ALL	59444	1,000	0,004	2,240

Annexure 5: Supplementary tables and figures supporting Chapter 5



Supplementary Figure 1: The quality profiles of Welverdiend sequence data barcoded with *rbcL* shows a heat map of the frequency of each quality score at each base position.



Supplementary Figure 2: The quality profiles of Welverdiend sequence data barcoded with *trnL* shows a heat map of the frequency of each quality score at each base position.



Supplementary Figure 3: The quality profiles of Syferkuil sequence data barcoded with *rbcl* shows a heat map of the frequency of each quality score at each base position.



Supplementary Figure 4: The quality profiles of Syferkuil sequence data barcoded with *trnL* shows a heat map of the frequency of each quality score at each base position.

Supplementary Table 7: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Welverdiend cattle barcoded with *rbcL*.

Sample	input	filtered	denoised	non-chim
1-rbcl	158873	104211	103918	793854
	8	8	4	
2-rbcl	782092	465583	460803	389150

3-rbcl	140436	82271	79496	70189
4-rbcl	156146	108225	67558	67345
5-rbcl	464615	313542	312365	233250
6-rbcl	894827	598888	595028	460319
7-rbcl	159321	99021	97426	80679
8-rbcl	110351	62246	58933	47635
9-rbcl	154321	87294	84625	79006
10-rbcl	151933	95184	93907	72060
11-rbcl	505817	318783	315645	253817
12-rbcl	231398	136871	133608	102253
13-rbcl	149392	93990	92468	70498
14-rbcl	187220	117172	116339	91726
15-rbcl	920415	524722	520362	463836
16-rbcl	105065 6	573194	564662	535490
17-rbcl	845520	531251	527997	356533
18-rbcl	175698	108227	106483	76474
19-rbcl	171358	110778	109004	73520
20-rbcl	228126	130116	127795	107824
21-rbcl	166863	109700	109021	73037
22-rbcl	652825	373260	369948	334174
23-rbcl	159142	88664	87235	80350

24-rbcl	154760	100578	99706	71821
C2-rbcl	7974	5439	5310	4710

Supplementary Table 8: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Welverdiend cattle barcoded with *trnL*.

Sample	input	filtered	denoised	non-chim
1-trnl	98912	25875	24940	24615
2-trnl	3720	1867	1551	1547
3-trnl	231900	152847	123335	122386
4-trnl	836	565	186	164
5-trnl	429459	184531	151731	143286
6-trnl	516328	217224	169509	164384
7-trnl	280346	102402	86655	85002
8-trnl	8206	4393	3979	3979
9-trnl	192674	101611	72659	71404
10-trnl	195418	54448	40201	39011
11-trnl	39131	17991	16463	16297
12-trnl	129823	57905	45907	45408
13-trnl	81390	31213	30029	29657
14-trnl	2789	2133	1711	1711

15-trnl	2327	1586	1266	1264
16-trnl	221045	156636	138206	137154
17-trnl	158255	104469	102643	98955
18-trnl	172629	118744	111170	106169
19-trnl	254668	186476	180850	166637
20-trnl	322511	211917	176429	174595
21-trnl	226265	162626	156219	148115
22-trnl	217090	158315	129170	128730
23-trnl	144480	97962	95012	94702
24-trnl	92278	48090	47104	45944
C2-trnl	2034	1376	1330	1326

Supplementary Table 9: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Syferkuil cattle barcoded with *rbcL*.

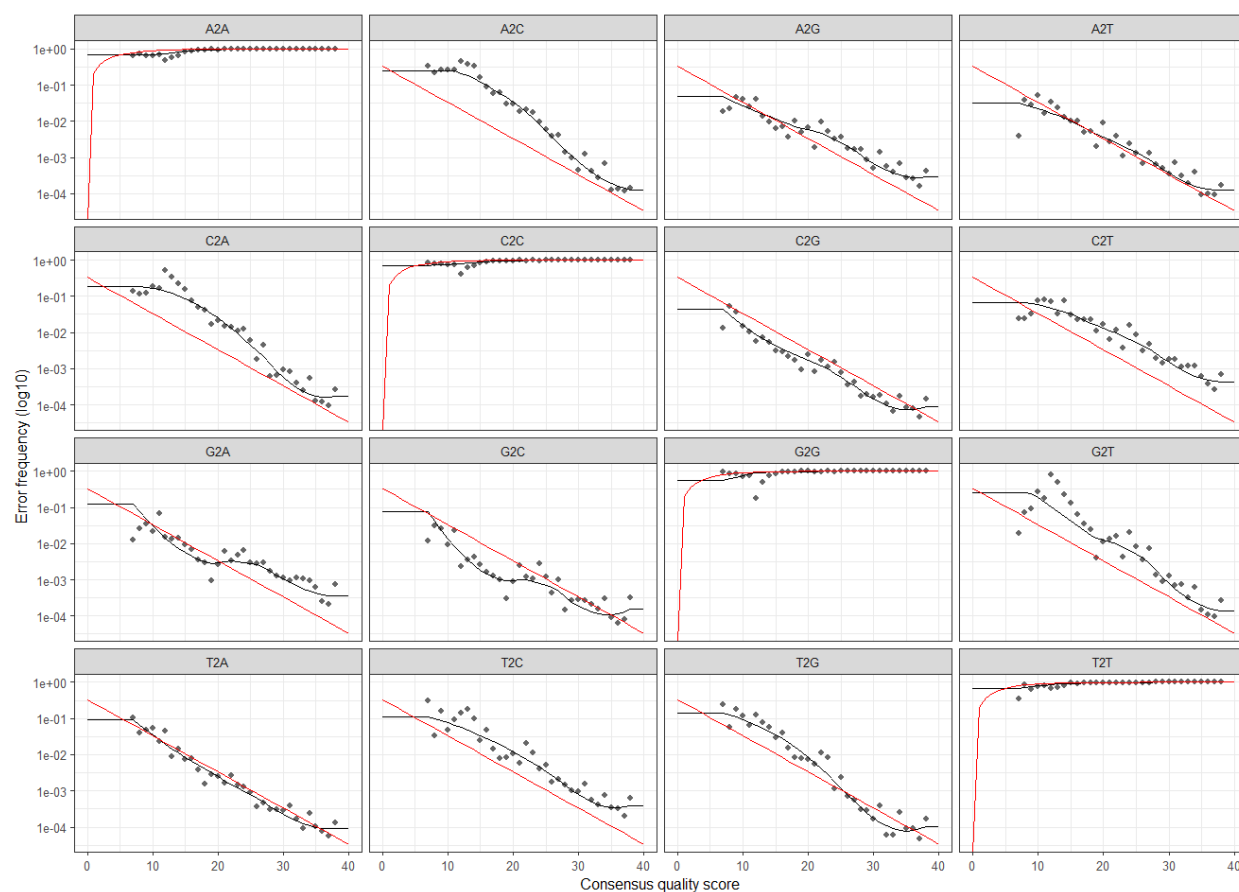
Sample	input	filtered	denoised	non-chim
25-rbcl	73867	63651	61545	59379
26-rbcl	101997	85760	84119	79801
27-rbcl	82403	70230	69009	64344
28-rbcl	82421	69400	67934	64238
29-rbcl	101655	85170	83665	76517

30-rbcl	79641	67998	66281	64059
31-rbcl	72183	61589	60197	57547
32-rbcl	64206	53552	52316	48419
33-rbcl	74427	63188	61833	57154
34-rbcl	69191	57820	56298	52712
35-rbcl	64846	55151	54007	50736
36-rbcl	69353	58812	57415	54578
37-rbcl	74571	62529	61396	56575
38-rbcl	79631	66226	65396	62782
39-rbcl	59923	51299	49938	48525
40-rbcl	83352	70419	69105	64429
41-rbcl	93151	79198	77228	68439
42-rbcl	86772	73650	71955	64893
43-rbcl	31708	24723	23781	23070
44-rbcl	82172	69593	68125	64877
45-rbcl	89295	75264	73834	68736
46-rbcl	96735	82909	81439	79586
47-rbcl	71923	58051	56527	54461
48-rbcl	72562	60934	59472	58287
C2-rbcl	140	44	2	2

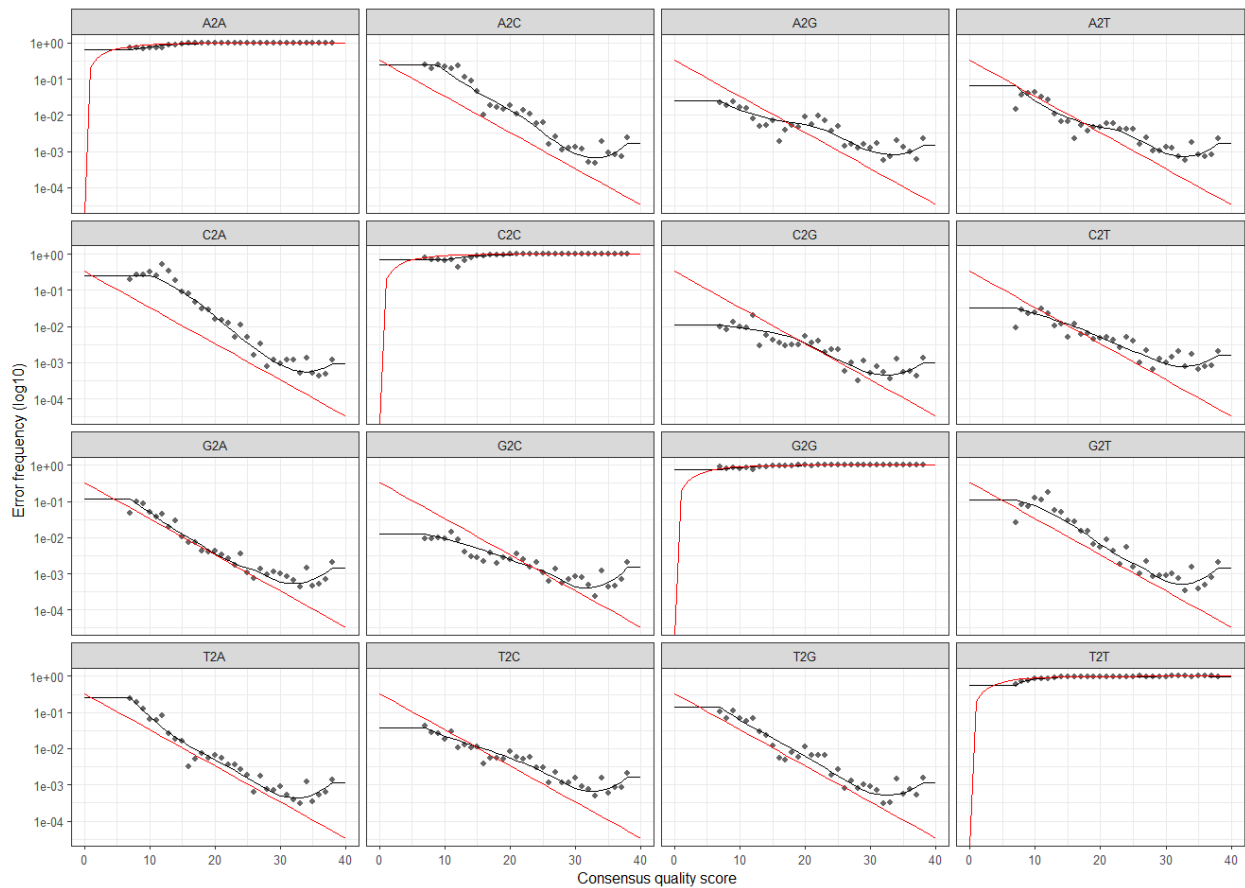
Supplementary Table 10: The number of sequences input into the DADA2 workflow per sample, the number left after filtering and trimming according to the presence of ambiguous bases, sequence lengths and quality scores, the number of sequences remaining after the inference of amplicon sequence variants as well as the number of sequences remaining after chimera removal (nonchim) for ASVs inferred from Syferkuil cattle barcoded with *trnL*.

Sample	input	filtered	denoised	non-chim
25-trnl	73007	62177	44837	44511
26-trnl	74721	64526	46906	46636
27-trnl	59795	49999	42945	42752
28-trnl	88963	75614	54710	54154
29-trnl	93440	79601	56863	56558
30-trnl	120172	103081	76312	75272
31-trnl	65449	55237	41855	41551
32-trnl	127572	112601	84617	83742
33-trnl	142943	125618	86868	86017
34-trnl	72892	63137	46037	45803
35-trnl	52535	44088	31850	31784
36-trnl	81315	69767	58219	58092
37-trnl	68405	60347	39893	39699
38-trnl	89166	75409	57052	56337
39-trnl	108845	92806	70954	66897
40-trnl	70974	60941	42088	41130
41-trnl	93380	79954	57144	56807

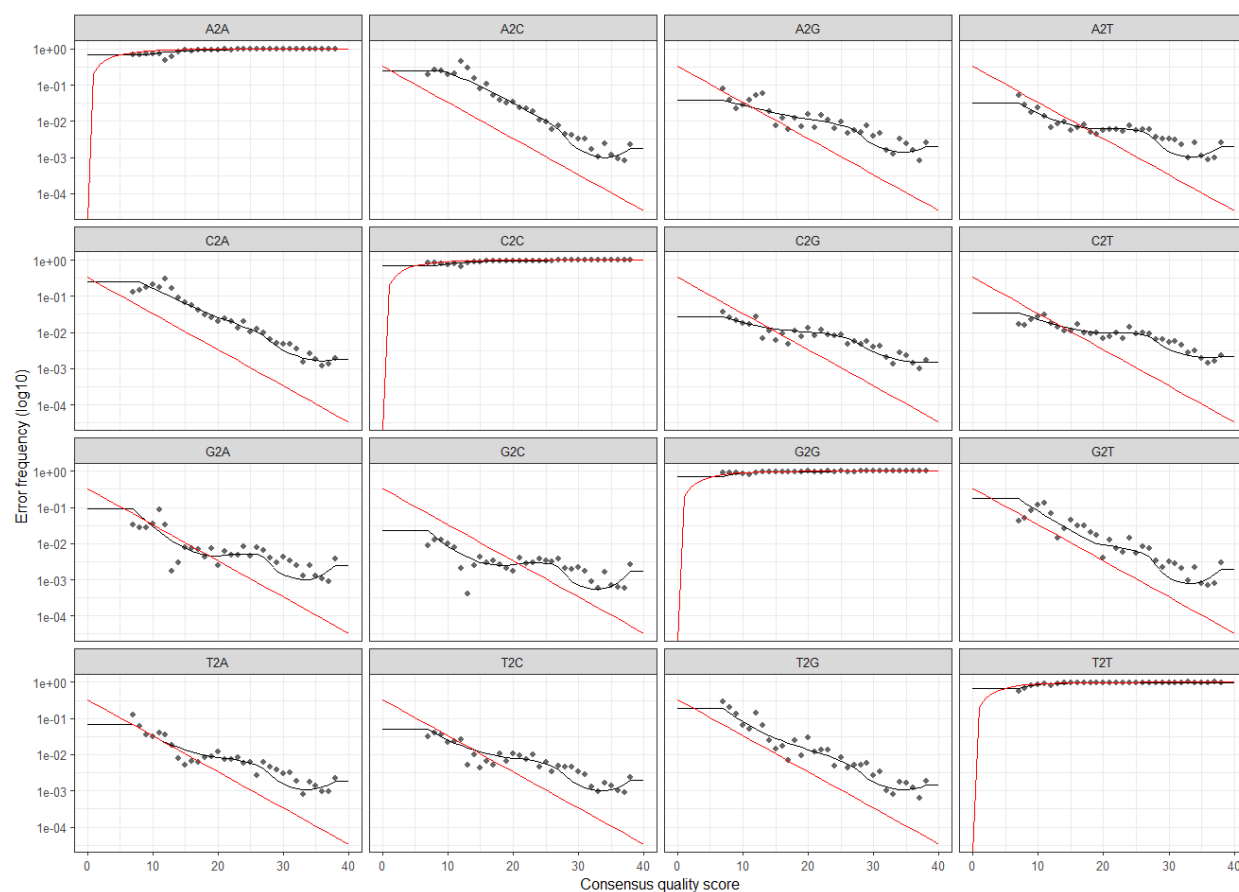
42-trnl	65241	56380	38745	38209
43-trnl	3085	1293	1289	1289
44-trnl	97883	82003	60642	59856
45-trnl	87921	77624	51736	51418
46-trnl	96712	82689	61914	60946
47-trnl	67458	57791	47664	47454
48-trnl	67693	55914	37387	37209
C2-trnl	16224	12018	10257	10257



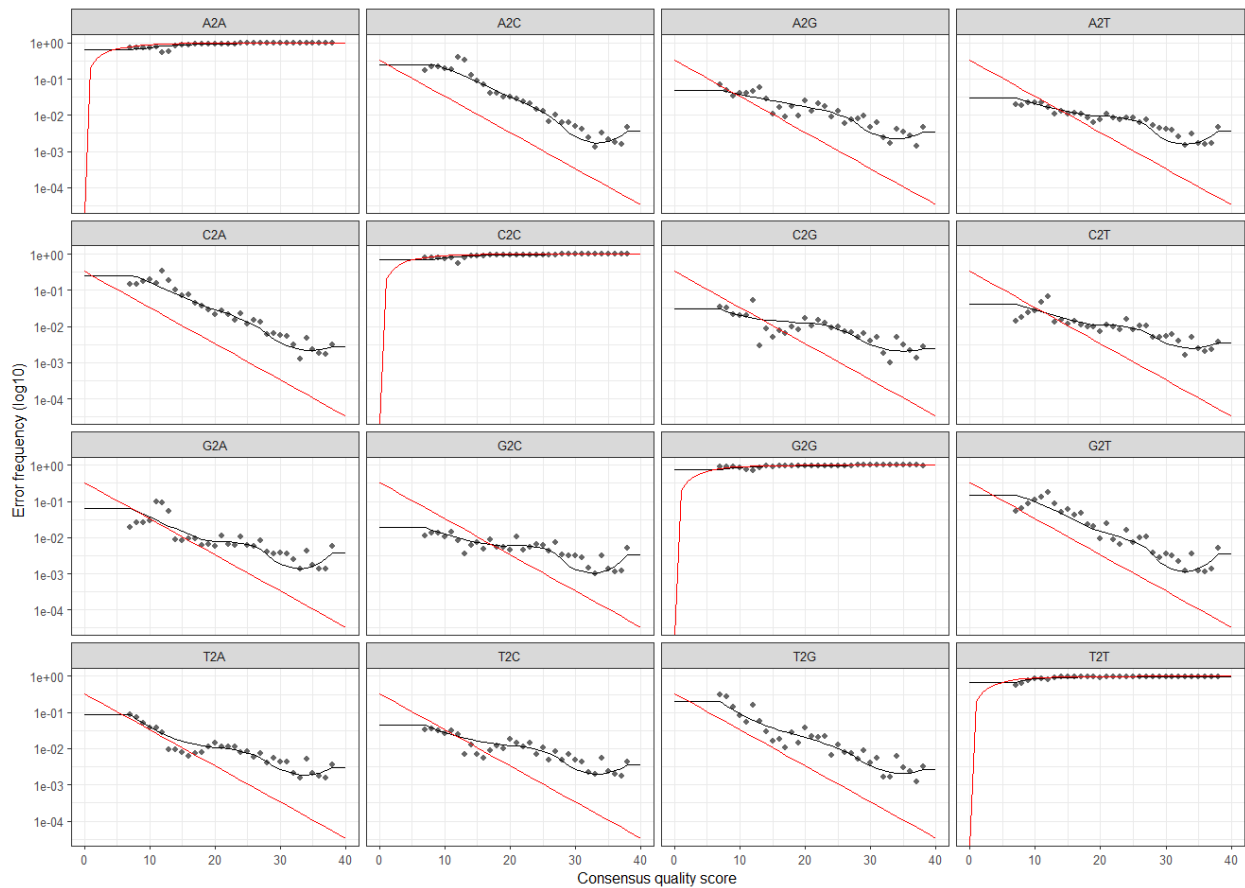
Supplementary Figure 5: The error rates plot for Welverdiend sequence data barcoded with *rbcL* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.



Supplementary Figure 6: The error rates plot for Welperdiend sequence data barcoded with *trnL* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.



Supplementary Figure 7: The error rates plot for Syferkuil sequence data barcoded with *rbcl* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.



Supplementary Figure 8: The error rates plot for Syferkuil sequence data barcoded with *trnL* illustrates the error rates for each possible base transition where points are the observed rates for each consensus quality score. Estimated error rates are represented by the black line, observed rates are represented by points and the red line represents the expected error rates according to the quality scores.

Supplementary Table 11: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Welverdiend *rbcL* dataset that could not be assigned by the compiled reference database of *rbcL* and *trnL* barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.

qseqid	accession	Description
3a498a2805f2813cb098c40be2f487ce9 59c5041	CP08400 7.1	<i>Flavonifractor plautii</i> strain NBRC 113805 chromosome, complete genome

8bc75a409334a233882eb064919178c9c246aa2e	CP08593 2.1	<i>Prevotella copri</i> DSM 18205 strain FDAARGOS_1573 chromosome, complete genome
9ac2e7f53589d1c8e593ed5cf876429d0a614c64	CP06049 0.1	<i>Oscillibacter hominis</i> strain NSJ-62 chromosome
a4fc14d2d1f34ea55b388e046f79a8b747e5feaa	AP02341 5.1	<i>Vescimonas fastidiosa</i> MM35 DNA, complete genome
b762bdf763fe11dd675703ddcfb67e71da7934f8	CP06049 0.1	<i>Oscillibacter hominis</i> strain NSJ-62 chromosome
d6b80a790845d1bb0c28a4f90e2beff7809fa40d	AP01853 2.1	<i>Oscillibacter</i> sp. PEA192 DNA, complete genome
d6b80a790845d1bb0c28a4f90e2beff7809fa40d	CP03441 3.3	<i>Dysosmobacter welbionis</i> strain J115 chromosome, complete genome
d902f37b6866417a76af22210cc2e09edaddc0b6	CP09171 8.1	MAG: <i>Oscillospiraceae</i> bacterium isolate nC33_bin.8.fa chromosome
d902f37b6866417a76af22210cc2e09edaddc0b6	CP09171 8.1	MAG: <i>Oscillospiraceae</i> bacterium isolate nC33_bin.8.fa chromosome

Supplementary Table 12: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Welperdiend *trnL* dataset that could not be assigned by the compiled reference database of *rbcL* and *trnL* barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.

qseqid	accession	Description
03a05891fb33fddd4d61bd363afcf128d7f5891	CP00087 5.1	<i>Herpetosiphon aurantiacus</i> DSM 785, complete genome

03a9e4224ff6537fd2303674b5af557bf8b450bc	OV78863 9.1	uncultured <i>Mollicutes</i> bacterium isolate min17_bin64 genome assembly, chromosome: 1
088e8ff643a045b3277fc3738d555dd64444dec3	CP04864 9.1	<i>Aminipila butyrica</i> strain DSM 103574 chromosome, complete genome
0ef98d0d98567c3554d95877d0f013db27ba4d23	AP02332 1.1	<i>Solibaculum mannosilyticum</i> 12CBH8 DNA, complete genome
1590773ae28fe467786edb480a79e9931f290e38	LR99447 3.1	uncultured <i>Prevotellamassilia</i> sp. isolate Prevotellamassilia canine HQ MAG genome assembly, chromosome: 1
2f9878cf3c26826763e28dd18ebcd6cc57d8431a	CP05870 7.1	MAG: <i>Bacteroidetes</i> bacterium isolate SSA2 chromosome
42c9ee2d966bd7dbaa84244a405a5ee99de98851	AP02342 0.1	<i>Pusillimonas faecalis</i> MM59 DNA, complete genome
52e7af6ded3261ecaaba211d5616f56043cea834	CP01420 4.2	<i>Clostridium baratii</i> strain CDC51267 chromosome, complete genome
52e7af6ded3261ecaaba211d5616f56043cea834	CP00690 5.1	<i>Clostridium baratii</i> str. Sullivan, complete genome
532da2e9a1a826ba331699bf41469bbd0bd9bbb4	OV78863 9.1	uncultured <i>Mollicutes</i> bacterium isolate min17_bin64 genome assembly, chromosome: 1
5af1fee9f8b150fa6f49e9ba260e054a1e6b1145	LK99601 7.1	<i>Desulfitobacterium hafniense</i> strain PCE-S genome assembly, scaffold: scaffold0
5af1fee9f8b150fa6f49e9ba260e054a1e6b1145	LR74649 6.1	Peptococcaceae bacterium INE isolate PacBioINE genome assembly, chromosome: 1
5af1fee9f8b150fa6f49e9ba260e054a1e6b1145	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
5af1fee9f8b150fa6f49e9ba260e054a1e6b1145	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome

5af1fee9f8b150fa6f49e9ba260e054a 1e6b1145	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
5af1fee9f8b150fa6f49e9ba260e054a 1e6b1145	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
5af1fee9f8b150fa6f49e9ba260e054a 1e6b1145	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
5af1fee9f8b150fa6f49e9ba260e054a 1e6b1145	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
5af1fee9f8b150fa6f49e9ba260e054a 1e6b1145	AP00823 0.1	<i>Desulfitobacterium hafniense</i> Y51 DNA, complete genome
69208ab38a7b385a14a6e1ddd06835 2abcda3888	CP04565 1.1	<i>Alistipes</i> sp. dk3624 chromosome, complete genome
818f8c7d4ae1a09c492a8eba08631a6 54cf64f49	CP09171 8.1	MAG: <i>Oscillospiraceae</i> bacterium isolate nC33_bin.8.fa chromosome
8567e123647e49854ed921c2ac5bc27 5f3490f35	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
8ebccd04a3029c35ff31f7635ac21dda 6b7c760d	CP04579 8.1	<i>Thermanaerosceptrum fracticalcis</i> strain DRI-13 chromosome, complete genome
c083091531718f0c37efd9078cd43ab 34f5df1eb	OV78863 9.1	uncultured <i>Mollicutes</i> bacterium isolate min17_bin64 genome assembly, chromosome: 1
cd5a3c2e07cdd90f99d056a5420ce49 3d423a47c	CP08046 7.1	<i>Alicyclobacillus acidoterrestris</i> strain DSM 3922 chromosome, complete genome
da0c6d6376b6c142d66d490a7eeec38 5a91339c1	OV78863 9.1	uncultured <i>Mollicutes</i> bacterium isolate min17_bin64 genome assembly, chromosome: 1

Supplementary Table 13: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Syferkuil *rbcL* dataset that could not be assigned by the compiled reference database of *rbcL*

and *trnL* barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI's non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.

qseqid	accession	Description
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10151 3.1	<i>Escherichia coli</i> strain HH194M chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10267 9.1	<i>Escherichia coli</i> strain XH989 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10267 5.1	<i>Escherichia coli</i> strain XH987 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07404 7.1	<i>Escherichia coli</i> strain PD7 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07404 2.1	<i>Escherichia coli</i> strain PI6 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07398 8.1	<i>Escherichia coli</i> strain MB10 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07397 6.1	<i>Escherichia coli</i> strain MB20 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07393 9.1	<i>Escherichia coli</i> strain MB53 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07393 6.1	<i>Escherichia coli</i> strain MB56 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07628 5.1	<i>Escherichia coli</i> strain 41 chromosome, complete genome

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07627 0.1	<i>Escherichia coli</i> strain Ceylon31 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07628 2.1	<i>Escherichia coli</i> strain 442/2 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07626 8.1	<i>Escherichia coli</i> strain D217-5 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07626 2.1	<i>Escherichia coli</i> strain 0-4 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07627 6.1	<i>Escherichia coli</i> strain SA31-1 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10192 0.1	<i>Escherichia coli</i> strain QA-1986 630 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10198 3.1	<i>Escherichia coli</i> strain STEC1096 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10197 1.1	<i>Escherichia coli</i> strain STEC1084 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07657 5.1	<i>Escherichia coli</i> strain CPWW_CT chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07571 6.1	<i>Escherichia coli</i> strain CPww7 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	AP02650 7.1	<i>Escherichia coli</i> TUM9868 DNA, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	AP02646 3.1	<i>Escherichia coli</i> TUM1887 DNA, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP10130 5.1	<i>Escherichia coli</i> strain 2002-3211 chromosome, complete genome

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09821 1.1	<i>Escherichia coli</i> strain Z0117EC0040 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09817 8.1	<i>Escherichia coli</i> strain Z0117EC0148 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07125 9.1	<i>Escherichia coli</i> strain MEI005 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	OW96811 1.1	<i>Escherichia coli</i> isolate 602 genome assembly, chromosome: main
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	OW96796 6.1	<i>Escherichia coli</i> isolate 648 genome assembly, chromosome: main
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06742 6.1	<i>Escherichia coli</i> O22:H8 strain 154 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08007 5.1	<i>Escherichia coli</i> strain 19SZHZ663Rt chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09719 8.1	<i>Escherichia coli</i> strain 2-1 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08622 0.1	<i>Escherichia coli</i> strain 779 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06093 3.1	<i>Escherichia coli</i> strain EC12 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06092 4.1	<i>Escherichia coli</i> strain EC14 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06091 9.1	<i>Escherichia coli</i> strain EC15 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06091 4.1	<i>Escherichia coli</i> strain EC18 chromosome, complete genome

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06089 9.1	<i>Escherichia coli</i> strain EC21 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02877 2.1	<i>Escherichia coli</i> strain A2 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02876 7.1	<i>Escherichia coli</i> strain A7 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02876 3.1	<i>Escherichia coli</i> strain A11 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02875 9.1	<i>Escherichia coli</i> strain A15 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02875 7.1	<i>Escherichia coli</i> strain A17 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02874 7.1	<i>Escherichia coli</i> strain A28 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02874 5.1	<i>Escherichia coli</i> strain A30 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02874 0.1	<i>Escherichia coli</i> strain A35 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02873 6.1	<i>Escherichia coli</i> strain A40 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02873 5.1	<i>Escherichia coli</i> strain A41 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP02873 3.1	<i>Escherichia coli</i> strain H2 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09263 9.1	<i>Escherichia coli</i> strain HS chromosome, complete genome

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05511 3.1	<i>Shigella flexneri</i> strain SWHIN_101 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05506 3.1	<i>Shigella dysenteriae</i> strain SWHE2 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05519 9.1	<i>Shigella flexneri</i> strain STEFF_3 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05519 4.1	<i>Shigella flexneri</i> strain STEFF_10 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05518 4.1	<i>Shigella flexneri</i> strain STEFF_16 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09225 2.1	<i>Escherichia coli</i> strain ISHS7 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP04382 1.1	<i>Escherichia coli</i> strain RM10740 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09183 9.1	<i>Escherichia coli</i> strain SWYHX138.261ft chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09183 6.1	<i>Escherichia coli</i> strain SWYHX138.33gt chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05466 2.1	<i>Escherichia coli</i> strain IGC_RcoliRes_1.1 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05419 1.1	<i>Escherichia coli</i> strain TH9F11 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05385 5.1	<i>Escherichia coli</i> strain 258E chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	OL79069 6.1	<i>Escherichia coli</i> strain H240 plasmid unnamed, complete sequence

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06126 4.1	<i>Escherichia coli</i> strain STEC1024 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06121 2.1	<i>Escherichia coli</i> strain STEC636 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06183 6.1	<i>Escherichia coli</i> strain STEC648 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP09027 7.1	<i>Escherichia coli</i> strain E-T207 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08757 8.1	<i>Escherichia coli</i> strain OSUCMP42NDM chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08563 8.1	<i>Escherichia coli</i> strain fEC.1 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08548 8.1	<i>Escherichia coli</i> strain VCSW8 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08518 3.1	<i>Escherichia coli</i> strain EC7518 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08326 3.1	<i>Escherichia coli</i> strain JNU chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP08000 1.1	<i>Escherichia coli</i> strain Colony157 chromosome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07857 0.1	<i>Escherichia coli</i> strain Colony199 chromosome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07859 1.1	<i>Escherichia coli</i> strain Colony139 chromosome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07859 0.1	<i>Escherichia coli</i> strain Colony160 chromosome

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07857 4.1	<i>Escherichia coli</i> strain Colony577 chromosome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07857 7.1	<i>Escherichia coli</i> strain Colony317 chromosome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07857 5.1	<i>Escherichia coli</i> strain Colony574 chromosome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	AP02447 8.1	<i>Escherichia coli</i> O121:H19 E15042 DNA, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	AP02447 5.1	<i>Escherichia coli</i> O121:H19 CEC14159 DNA, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	AP02447 3.1	<i>Escherichia coli</i> O121:H19 SE14002 DNA, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	AP02447 1.1	<i>Escherichia coli</i> O121:H19 51104 DNA, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07813 7.1	<i>Escherichia coli</i> strain Sx181-0-1 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	OU34984 9.1	<i>Escherichia coli</i> strain TA159 / Oneg:H31 / fimH31 / 3592 (ST Warwick) genome assembly, chromosome: 1
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	OU34983 9.1	<i>Escherichia coli</i> strain TA054 / Onovel10:H9 / fimH38-like / 3590 (ST Warwick) genome assembly, chromosome: 1
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP03588 2.1	<i>Escherichia coli</i> strain WS0115A chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07564 9.1	<i>Escherichia coli</i> strain STEC17-190 chromosome, complete genome

2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07548 4.1	<i>Escherichia coli</i> strain NT1N25 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07547 7.1	<i>Escherichia coli</i> strain NT1W25 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07506 2.1	<i>Escherichia coli</i> strain EC_J_9 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP05983 5.1	<i>Escherichia coli</i> strain tcmA_3 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP07280 2.1	<i>Escherichia coli</i> strain PK5074 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06723 2.1	<i>Escherichia coli</i> strain SBF45 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06351 8.1	<i>Escherichia coli</i> strain EH01-18-02 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06348 7.1	<i>Escherichia coli</i> strain EF7-18-51 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06970 6.1	<i>Escherichia coli</i> strain ECY44 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP04571 2.1	<i>Escherichia coli</i> strain FHI_NMBU_11 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06286 8.1	<i>Escherichia coli</i> strain Res13-Lact-PEB01-20 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06286 5.1	<i>Escherichia coli</i> O39:H21 strain Res13-Lact- PEB08-01 chromosome, complete genome
2c9c9d86fc68f6517d83c262c75e6abf 4c3be80d	CP06288 9.1	<i>Escherichia coli</i> O100:H21 strain Res13-Lact- ER07-18 chromosome, complete genome

3d4e6398e9d3d721d6d25d74b41f20 4751d44c7a	CP01620 2.1	<i>Eubacterium minutum</i> ATCC 700079 sequence
543b2225296f8e016cbd0484e7de0e0 1ae7cbfe3	AP01853 6.1	Lachnospiraceae bacterium GAM79 DNA, complete genome
5d71468941067d8e956858bd6cf9b4 539457e26e	LR99447 3.1	uncultured <i>Prevotellamassilia</i> sp. isolate <i>Prevotellamassilia canine</i> HQ MAG genome assembly, chromosome: 1
65943181f2cbc4e32e959353085615f c39c192ea	CP10307 2.1	<i>Bacteroides thetaiotaomicron</i> strain BFG-169 chromosome, complete genome
6fe36df589ef59d927b5d004b16e416 9ddc8e960	BK02810 7.1	MAG TPA_asm: <i>Bacteriophage</i> sp. isolate ctHX57, partial genome
79b1ec45340812143ed0cb73ca6dde9 73db222a8	CP05604 1.1	<i>Streptomyces chartreusis</i> strain NA02069 chromosome, complete genome
7b2da45a3925f2db38ea5128e9b815b 81a5eb315	CP03513 0.1	<i>Gudongella oleilytica</i> strain W6 chromosome, complete genome
90dc650514590458d57b70a3889c0e 3055bef61f	CP06942 1.1	<i>Clostridiales</i> bacterium strain WTE2008 chromosome, complete genome
990219c9c1be55abc1c6e05832acf6a 3a88649ab	CP08593 2.1	<i>Prevotella copri</i> DSM 18205 strain FDAARGOS_1573 chromosome, complete genome
a78b51824bb2537ed689aa78378647 6b86c90781	CP04920 7.1	<i>Agrobacterium rubi</i> strain W2/73 chromosome 2, complete sequence
c95573a1f4b285518205f081b521af1 c8f67ef04	LR21598 0.1	<i>Paraprevotella xylaniphila</i> YIT 11841 isolate <i>Paraprevotella xylaniphila</i> 82A6 genome assembly, chromosome: 1
cec9731b48fe35cc0168bbd6f7dab64 6b618b229	CP00912 9.1	<i>Planococcus</i> sp. PAMC 21323, complete genome

d2c5907b96d0cd4c0bf831e37ba62fd 75899d8c9	CP06049 0.1	<i>Oscillibacter hominis</i> strain NSJ-62 chromosome
fa0720e90e423769046fb15c3bda482 6084095f3	CP03231 2.1	<i>Pannonibacter phragmitetus</i> BB chromosome, complete genome

Supplementary Table 14: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Syferkuil *trnL* dataset that could not be assigned by the compiled reference database of *rbcL* and *trnL* barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_LULU workflow.

qseqid	accession	Description
026d60069d4306b97fe24347ddb4362 4baaed7bf	LR21598 3.1	<i>Eubacterium limosum</i> isolate Eubacterium limosum 81C1 genome assembly, chromosome: 1
026d60069d4306b97fe24347ddb4362 4baaed7bf	CP01996 2.1	<i>Eubacterium limosum</i> strain ATCC 8486 chromosome, complete genome
026d60069d4306b97fe24347ddb4362 4baaed7bf	CP09737 6.1	<i>Eubacterium limosum</i> strain B2 chromosome, complete genome
026d60069d4306b97fe24347ddb4362 4baaed7bf	CP00227 3.2	<i>Eubacterium callanderi</i> strain KIST612, complete genome
2d643aa53be053d82c020b965a17b1c3 a40e6091	CP09446 7.1	<i>Faecalibacterium</i> sp. I4179 chromosome, complete genome
438051be796fc996585ca41f79d6717fc e4eda1f	CP04843 6.1	<i>Flavonifractor plautii</i> strain JCM 32125 chromosome, complete genome
56cc02e406d83d7794083a24aafc5957 97194ea7	CP03620 0.1	<i>Shewanella maritima</i> strain D4-2 chromosome, complete genome

56cc02e406d83d7794083a24aafc5957 97194ea7	CP02047 2.1	<i>Shewanella japonica</i> strain KCTC 22435 chromosome, complete genome
56cc02e406d83d7794083a24aafc5957 97194ea7	CP06205 8.1	<i>Shewanella</i> sp. WPAGA9 chromosome, complete genome
56cc02e406d83d7794083a24aafc5957 97194ea7	CP06108 1.1	<i>Marinomonas arctica</i> strain BSI20414 chromosome, complete genome
5b14dbb816c456be039b9754de93bfe4 734f1231	CP09172 6.1	MAG: <i>Bacilli</i> bacterium isolate nC33_bin.244.fa chromosome
6c471fea63dbadd45355dcfa572ab692 3b4923cc	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
745fcc799306452ff8c50404c2f93f117 29ff784	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
745fcc799306452ff8c50404c2f93f117 29ff784	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
745fcc799306452ff8c50404c2f93f117 29ff784	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
8619aca1d2b99f77ba87cdab47889809 49602b59	LR13438 4.1	<i>Prevotella oris</i> strain NCTC13071 genome assembly, chromosome: 1
8b6776e752d9b1101e24b84082c1c69a 583ceb60	BK03792 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctxXH1, partial genome
9fdcb110f4845c025462b461e9c2b845 028eb72d	BK05729 4.1	MAG TPA_asm: <i>Podoviridae</i> sp. isolate ctI6w6, partial genome
a107bbf993cc141ae59995a0b1654570 6553be83	CP05468 9.1	Ectothiorhodospiraceae bacterium 2226 chromosome, complete genome
a69522903bd361ad83b5785db17cd53 5706766cf	CP04565 1.1	<i>Alistipes</i> sp. dk3624 chromosome, complete genome

b3b8f9962fbdc9ae2264a053044b3e72 e73ed801	CP09173 1.1	MAG: <i>Candidatus Melainabacteria</i> bacterium isolate nC33_lr_bin.1.fa chromosome
d89cf99f254f45cb06b767bf197d1719d cfb07a0	CP05491 9.1	<i>Kitasatospora</i> sp. NA04385 chromosome, complete genome
e088361219d29dd201347acce866cea5 cf74bde0	CP06382 2.1	<i>Clostridium botulinum</i> strain 1873 chromosome, complete genome
e088361219d29dd201347acce866cea5 cf74bde0	CP00241 0.1	<i>Clostridium botulinum</i> BKT015925, complete genome
e693066801c8f89691c81388893d8cb7 2b5694c2	BK05783 7.1	MAG TPA_asm: <i>Bacteriophage</i> sp. isolate ctWh23, partial genome
f2a186c5fec4318a8c98de6191906968 b8f50292	CP01420 4.2	<i>Clostridium baratii</i> strain CDC51267 chromosome, complete genome
ff4b4506c1517e6b07c2c5cc84a06d30 48b8de66	LN87950 2.1	<i>Protochlamydia naegleriophila</i> genome assembly PNK1, chromosome : cPNK
ff4b4506c1517e6b07c2c5cc84a06d30 48b8de66	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome

Supplementary Table 15: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Syferkuil *rbcL* dataset that could not be assigned by the compiled reference database of *rbcL* and *trnL* barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_RA_LULU workflow.

qseqid	accession	Description
0240b45054a035a71791223cfd0bc31 3f5f75238	CP08801 7.1	<i>Ruania halotolerans</i> strain M4N3S171 chromosome, complete genome

0341f05a53a86f4998df7f466f80afba 234e183d	CP01422 3.1	<i>Anaerotignum propionicum</i> DSM 1682 strain X2 chromosome, complete genome
080cf77a43e74a90e9b402b5e5759c6 1cb3fa987	CP09446 8.1	<i>Faecalibacterium</i> sp. I3389 chromosome, complete genome
09d7d69eddaabf331b996a70ab93785 c38c82873	CP01893 7.1	<i>Bacteroides fragilis</i> strain Q1F2 chromosome, complete genome
0cdf160637128eb5d9cafbcc94a56be 4d6cfd7	CP06944 0.1	<i>Phocaeicola coprophilus</i> strain FDAARGOS_1220 chromosome, complete genome
1619ef17ae2a402ea5b34175c2bedf58 ad3ede52	CP09603 4.1	<i>Gottfriedia acidiceris</i> strain LH14 chromosome, complete genome
17d74e68de3845e8fc9372069cd6ff17 b32543a3	AP02341 5.1	<i>Vescimonas fastidiosa</i> MM35 DNA, complete genome
2aa9456c3250303a82c60005d805b77 79555a51f	CP06941 9.1	<i>Clostridiales</i> bacterium strain XBB3002 chromosome, complete genome
352a2b68cf33eaac790e73839d44afa9 d71384d4	LM99741 2.1	<i>Peptoniphilus</i> sp. ING2-D1G genome assembly, chromosome: I
3aeb7de0ced80a9704bbffaae5ea9461 cd61105d	OV78863 9.1	uncultured <i>Mollicutes</i> bacterium isolate min17_bin64 genome assembly, chromosome: 1
3e72fde62fd6a666c1a0dab992648bb 09e72e2f2	CP04882 9.1	<i>Lachnospiraceae</i> bacterium MP1D12 chromosome, complete genome
425e92e64eb6235f67bdd6b14790b37 545b435da	CP03619 9.1	<i>Leclercia adecarboxylata</i> strain 16005813 chromosome, complete genome
425e92e64eb6235f67bdd6b14790b37 545b435da	CP08728 0.1	<i>Leclercia adecarboxylata</i> strain SH19PE29 chromosome, complete genome
45b443d5282887536eb8bf41379687 1d753b2e91	AP02528 4.1	<i>Amphritea atlantica</i> JCM 14776 DNA, complete genome

45fcee36c525bf273b97d55ad578a48a78aac588	CP00128 1.2	<i>Thauera</i> sp. MZ1T, complete genome
49d5fc8d165f52bf17ffdbe633fc8df815f965b6	CP04239 0.1	<i>Leuconostoc lactis</i> strain CBA3626 chromosome, complete genome
4b0522ac7cb76115937efa547b49e859eb47596b	CP00219 8.1	<i>Cyanothece</i> sp. PCC 7822, complete genome
518c28e6d60f611573b1229cb9cf2a7f483e8e5f	CP04842 7.1	<i>Rhizobium daejeonense</i> strain KACC 13094 chromosome, complete genome
6180da5b5e54ad56bbb6d0a085e6859a7c8ac3b5	CP02602 9.1	<i>Paenibacillus</i> sp. PK3_47 chromosome, complete genome
661eb5b4f56b05cfa130a4f3474980f80badcce2	CP09986 3.1	<i>Bacillus</i> sp. 1780r2a1 chromosome
6fe36df589ef59d927b5d004b16e4169ddc8e960	BK02810 7.1	MAG TPA_asm: <i>Bacteriophage</i> sp. isolate ctHX57, partial genome
71595bbd92513cd036e9dd4ea4dc148beb77202e	CP09537 6.1	<i>Amycolatopsis pretoriensis</i> strain FBCC-B4732 chromosome, complete genome
7326e23d7bd769b30f504f0075b1391861e3a4dc	CP07369 4.1	<i>Dysosmobacter</i> sp. Marseille-Q4140 chromosome
7b471929f8433513bfdb33ad74b54aea56041200	CP04842 7.1	<i>Rhizobium daejeonense</i> strain KACC 13094 chromosome, complete genome
7d45fd07335ced9838c13cc8f32fa0becb51654	CP03075 4.1	<i>Acinetobacter schindleri</i> strain H3 chromosome, complete genome
7d45fd07335ced9838c13cc8f32fa0becb51654	CP09483 4.1	<i>Acinetobacter schindleri</i> strain BDT2091 chromosome, complete genome
8229bd3e78c7b2101d3e7fd4ee02113dcf189656	CP05604 1.1	<i>Streptomyces chartreusis</i> strain NA02069 chromosome, complete genome

8cd21ef0e2a5ce1c3c66403293e23dbf 0360ec6b	CP04843 7.1	<i>Faecalibacterium duncaniae</i> strain JCM 31915 chromosome, complete genome
8cd21ef0e2a5ce1c3c66403293e23dbf 0360ec6b	CP02247 9.1	<i>Faecalibacterium prausnitzii</i> strain A2165 chromosome, complete genome
8f94f863b8d8fbdaf624844d9fb537e8 7304d79d	AP02341 8.1	<i>Vescimonas coprocola</i> MM50 DNA, complete genome
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP05146 8.1	<i>Cereibacter sphaeroides</i> strain CH10 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP04703 7.1	<i>Cereibacter sphaeroides</i> strain 2.4.1 substr. H2 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP04703 1.1	<i>Cereibacter sphaeroides</i> strain DSM 158 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP03641 9.1	<i>Cereibacter sphaeroides</i> strain HJ chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP03343 4.1	<i>Cereibacter sphaeroides</i> strain AB24 chromosome 1
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP03344 2.1	<i>Cereibacter sphaeroides</i> strain AB25 chromosome 1
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP03344 6.1	<i>Cereibacter sphaeroides</i> strain AB27 chromosome 1
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP03345 0.1	<i>Cereibacter sphaeroides</i> strain AB29 chromosome 1
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP03027 1.1	<i>Cereibacter sphaeroides</i> 2.4.1 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP01528 7.1	<i>Rhodobacter sphaeroides</i> strain MBTLJ-20 chromosome 1, complete sequence

975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP01521 0.1	<i>Cereibacter sphaeroides</i> strain MBTLJ-13 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP01296 0.1	<i>Rhodobacter sphaeroides</i> strain MBTLJ-8 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP00014 3.2	<i>Rhodobacter sphaeroides</i> 2.4.1 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	CP00115 0.1	<i>Rhodobacter sphaeroides</i> KD131 chromosome 1, complete sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	AF19512 2.1	<i>Rhodobacter sphaeroides</i> photosynthetic gene cluster, partial sequence
975ef4a4da896534ad251ebf80ab08e 2066cdee6	AJ01030 2.1	<i>Rhodobacter sphaeroides</i> photosynthetic gene cluster
99be3b97311bb88d3ce4298fb17a30d 468dc20a8	CP09603 4.1	<i>Gottfriedia acidiceris</i> strain LH14 chromosome, complete genome
9ae4719411c897ad61341becb5e4832 77f03737b	CP01240 1.1	<i>Azospirillum thiophilum</i> strain BV-S chromosome 1, complete sequence
a31e4e6b2dea628a0e9c8c0c0737316 6e27a87bd	CP09446 8.1	<i>Faecalibacterium</i> sp. I3389 chromosome, complete genome
aefac5da6e5a7f82d8f48ebb414b7bab 0cb6975f	CP01577 4.2	<i>Lelliottia amnigena</i> strain ZB04 chromosome, complete genome
b52ea16fed24d0a1f43485b69fe19bca 444aabb6	CP06293 9.1	<i>Bifidobacterium subtile</i> strain KCTC 3272 chromosome, complete genome
bb73a14595a371696110a0e75da5281 5f0fd26d6	CP03548 5.1	<i>Salicibacter halophilus</i> strain NKC3-5 chromosome, complete genome
c11d938fc5dc993240e5e4a9bf00c77 b75e69e08	CP06049 0.1	<i>Oscillibacter hominis</i> strain NSJ-62 chromosome

c6710913cfedb4ed085efb4033e442b69e752dc4	CP06941 9.1	<i>Clostridiales</i> bacterium strain XBB3002 chromosome, complete genome
c928f29ebc246e07df2f49b04a34f5788fba01c0	CP09728 2.1	<i>Flavonifractor plautii</i> strain NB2A-20-GAM chromosome
c95573a1f4b285518205f081b521af1c8f67ef04	LR21598 0.1	<i>Paraprevotella xylaniphila</i> YIT 11841 isolate <i>Paraprevotella xylaniphila</i> 82A6 genome assembly, chromosome: 1
ce715faf175d0c60f4fa088bad8bc2a68ceff931	CP04843 6.1	<i>Flavonifractor plautii</i> strain JCM 32125 chromosome, complete genome
ce715faf175d0c60f4fa088bad8bc2a68ceff931	CP09509 4.1	<i>Flavonifractor plautii</i> strain VE303-08 chromosome, complete genome
cec9731b48fe35cc0168bbd6f7dab646b618b229	CP09603 4.1	<i>Gottfriedia acidiceris</i> strain LH14 chromosome, complete genome
d316942fd1b2dcdb0284d67c5ce2e5d17db028da	CP07222 7.1	<i>Bacteroides eggerthii</i> strain CL11T00C20 chromosome, complete genome
d39ce176119ac08a87c3ac77ba2673252f4af2e5	CP01240 1.1	<i>Azospirillum thiophilum</i> strain BV-S chromosome 1, complete sequence
d7c2363232dff754db3ecaf4c166f3b56efabcc	CP06293 9.1	<i>Bifidobacterium subtile</i> strain KCTC 3272 chromosome, complete genome
e8f76a39fbcd395e1ef2f13a941aabe658067ae6	CP02961 8.1	<i>Streptomyces</i> sp. WAC 06738 chromosome, complete genome
e92053f66773c29922f77edf759e53aac8e432d8	CP09603 4.1	<i>Gottfriedia acidiceris</i> strain LH14 chromosome, complete genome
e94fa511e05609bec538e8d6b236e346b08c495e	CP05604 1.1	<i>Streptomyces chartreusis</i> strain NA02069 chromosome, complete genome
edfb7186a32b05d132568f6917d313533a08b5f0	CP08400 7.1	<i>Flavonifractor plautii</i> strain NBRC 113805 chromosome, complete genome

f070a64dd3b6e755f7df7662013b715 8c541bb35	CP07369 1.1	<i>Clostridiaceae</i> bacterium Marseille-Q4149 chromosome
f221e0529f46baf879fa82e1da3382ae 261c3855	AP02558 4.1	<i>Oscillospiraceae</i> bacterium CE91-St41 DNA, complete genome
f221e0529f46baf879fa82e1da3382ae 261c3855	AP02558 3.1	<i>Oscillospiraceae</i> bacterium CE91-St40 DNA, complete genome
f8052f0f23f60ac2393626ac0622051f ecf08e3b	CP09509 4.1	<i>Flavonifractor plautii</i> strain VE303-08 chromosome, complete genome
fc76bdaea4e4a926325ef6cf8e1fd8e7 60619aab	CP01654 0.2	<i>Planococcus</i> sp. L10.15, complete genome

Supplementary Table 16: Query sequence identities (“qseqid”) of amplicon sequence variants (ASVs) from the Syferkuil *trnL* dataset that could not be assigned by the compiled reference database of *rbcL* and *trnL* barcodes of potential foraging targets in an eastern semi-arid South African savanna but could be assigned to species other than plants by NCBI’s non-redundant nucleotide database with blast parameters: 80% sequence identity, 80% query coverage and e-values smaller than 0,001. These taxonomic assignments were made with ASVs resulting from the application of the DADA_1_RA_LULU workflow.

qseqid	accession	Description
0054ce76626f6054f75630cae2e2f 2c716122eed	BK05508 4.1	MAG TPA_asm: <i>Microviridae</i> sp. isolate ctSXA1, partial genome
0085add1b29a2a85a894ae1bb9b4 6416f6a43014	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
00db3deaeb0c1398df25603fea491 e9f48900066	CP07457 5.1	<i>Tenericutes</i> bacterium strain zrk6 chromosome, complete genome
00eac2ffe5f15f16e6707b66e1691 52b3e13af35	BK05593 9.1	MAG TPA_asm: <i>Herelleviridae</i> sp. isolate ctK5y9, partial genome

01223f347a3c44c8c62638d76687 1f53cee701c9	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
014a3bfca2605bc1dc6cdf3992ecc c160826373e	BK05022 1.1	MAG TPA_asm: <i>Bacteriophage</i> sp. isolate ctkor1, partial genome
0173281ccaddac0f409264e1f63ad 6e31bb58132	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
0176849b6f6be78942af64c230dd 2e75785fcebc	LR69898 1.1	<i>Clostridiales</i> bacterium isolate MGYG-HGUT-02340 genome assembly, chromosome: 1
01cab59108fab412afbd9bb6fdc9d eb5b97355a4	CP03027 8.1	<i>Ethanoligenens harbinense</i> strain W1 chromosome, complete genome
01cab59108fab412afbd9bb6fdc9d eb5b97355a4	CP02528 8.1	<i>Ethanoligenens harbinense</i> strain X-29 chromosome, complete genome
01cab59108fab412afbd9bb6fdc9d eb5b97355a4	CP02528 7.1	<i>Ethanoligenens harbinense</i> strain B49 chromosome, complete genome
01cab59108fab412afbd9bb6fdc9d eb5b97355a4	CP02528 6.1	<i>Ethanoligenens harbinense</i> YUAN-3 chromosome, complete genome
01cab59108fab412afbd9bb6fdc9d eb5b97355a4	CP00240 0.1	<i>Ethanoligenens harbinense</i> YUAN-3, complete genome
0233b891d16f264bfcda615c3a05e ec14d251222	CP01323 5.1	<i>Collimonas arenae</i> strain Ter282 chromosome, complete genome
0233b891d16f264bfcda615c3a05e ec14d251222	CP01323 3.1	<i>Collimonas arenae</i> strain Ter10 chromosome, complete genome
026d60069d4306b97fe24347ddb4 3624baaed7bf	CP02724 2.1	<i>Peptostreptococcaceae</i> bacterium oral taxon 929 strain W2294 chromosome, complete genome
026d60069d4306b97fe24347ddb4 3624baaed7bf	CP09664 9.1	<i>Fenollaria massiliensis</i> strain C0061C2 chromosome, complete genome

036c838f0c8e60c685616bb39dd6f 091cae84402	CP03456 2.1	<i>Flammeovirga pectinis</i> strain L12M1 chromosome 1, complete sequence
036c838f0c8e60c685616bb39dd6f 091cae84402	CP07612 8.1	<i>Flammeovirga kamogawensis</i> strain YS10 chromosome 1, complete sequence
03a9e4224ff6537fd2303674b5af5 57bf8b450bc	CP09172 3.1	MAG: <i>Christensenella</i> sp. isolate nC33_lr_bin.267.fa chromosome
03bafc3e48502d302cb3f522e5ce1 53c77b1be9d	CP04090 2.1	<i>Longibaculum</i> sp. KGMB06250 chromosome, complete genome
03bafc3e48502d302cb3f522e5ce1 53c77b1be9d	CP03183 7.1	<i>Clostridium sporogenes</i> strain YH-Raf3-2-5 chromosome, complete genome
03bafc3e48502d302cb3f522e5ce1 53c77b1be9d	AP01853 7.1	Erysipelotrichaceae bacterium GAM147 DNA, complete genome
03bafc3e48502d302cb3f522e5ce1 53c77b1be9d	AP02408 5.1	<i>Faecalibacillus intestinalis</i> 14EGH31 DNA, complete genome
061081f8384bc7d1fd46770de445 275e9b740b2a	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
067fcb08d31f9ecfa510016c68e02 01df5be83ca	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
067fdcfef58e72e3612b85583b9b 626cacb31c8	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
067fdcfef58e72e3612b85583b9b 626cacb31c8	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
06aa663f3b9b8b3617611394328a 15dab398730d	CP09446 8.1	<i>Faecalibacterium</i> sp. I3389 chromosome, complete genome
07094ea72875bf917a63365e38e0 9cbe839e6a18	CP10126 9.1	<i>Escherichia coli</i> strain EC21Z-048 chromosome, complete genome

07094ea72875bf917a63365e38e0 9cbe839e6a18	CP10054 4.1	<i>Escherichia coli</i> strain LH09-a chromosome, complete genome
07094ea72875bf917a63365e38e0 9cbe839e6a18	OW9682 88.1	<i>Escherichia coli</i> isolate 10 genome assembly, chromosome: main
07094ea72875bf917a63365e38e0 9cbe839e6a18	OW9681 24.1	<i>Escherichia coli</i> isolate 10 genome assembly, chromosome: main
07094ea72875bf917a63365e38e0 9cbe839e6a18	OW9679 75.1	<i>Escherichia coli</i> isolate 10 genome assembly, chromosome: main
07094ea72875bf917a63365e38e0 9cbe839e6a18	OW8493 77.1	<i>Escherichia coli</i> isolate 162 genome assembly, chromosome: main
07094ea72875bf917a63365e38e0 9cbe839e6a18	CP08006 6.1	<i>Escherichia coli</i> strain 19SZHZ803Rt chromosome, complete genome
07094ea72875bf917a63365e38e0 9cbe839e6a18	CP09719 8.1	<i>Escherichia coli</i> strain 2-1 chromosome, complete genome
07094ea72875bf917a63365e38e0 9cbe839e6a18	CP09683 4.1	<i>Escherichia coli</i> strain PM4-5 chromosome, complete genome
073a7c40845d2e89460682d3f4cc 8a47c0dec5b1	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
082502eef7ac2e7ed9cd80fe465fd 6573f82df5c	BK01738 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctPWM4, partial genome
0849f3cadd9e7def057b4d08075b4 b1390da9f86	CP02777 5.1	<i>Clostridium botulinum</i> strain MFBjulcb1 chromosome, complete genome
084a160fb19e81439a8ecdbeafa89 275d8071fee	CP03510 7.1	<i>Ornithobacterium rhinotracheale</i> strain FARPER-174b chromosome, complete genome
0975c4e34161c2fe45d5e64472fba 7bb574dcd7f	LR21598 1.1	<i>Ruthenibacterium lactatiformans</i> isolate <i>Ruthenibacterium lactatiformans</i> 82B1 genome assembly, chromosome: 1

0a742d1359539c165c58268937a9f8ffcd5e87b8	OV78864 0.1	uncultured <i>Tenericutes</i> bacterium isolate min17_bin11 genome assembly, chromosome: 1
0a90d58d5a5494a1e8ee56375043be07d9bbebf5	AP02281 0.1	<i>Desulfuromonas</i> sp. AOP6 DNA, complete genome
0b3d7ad61653e53b38070baed736b674df65e76a	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
0bf348cbe0594bb4821c5d6142a180de3c876cb8	BK01776 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct2EZ5, partial genome
0c4c66e3ba6192401d0adb56de4548ef5b3d0732	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
0cb235e1ae6258dfb2d18f6d8de1de5c5b2a803a	CP02614 3.1	<i>Bacillus methanolicus</i> strain DFS2 chromosome, complete genome
0cb235e1ae6258dfb2d18f6d8de1de5c5b2a803a	CP02614 3.1	<i>Bacillus methanolicus</i> strain DFS2 chromosome, complete genome
0d5e45ddf0758e2887b8032464c9c65d1fa482fe	CP05348 2.1	<i>Thiothrix subterranea</i> strain Ku-5 chromosome, complete genome
0da57bf4c2bcfce02269a3747820622b28b331d0	MH6229 43.1	MAG: <i>Myoviridae</i> sp. isolate ctbc_4, complete genome
0ef98d0d98567c3554d95877d0f013db27ba4d23	AP02332 1.1	<i>Solibaculum mannosilyticum</i> 12CBH8 DNA, complete genome
0f194855260959fdb86b5d457d4aeb4467842f11	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
0f8b5d3f424178cc836411e4ad3541a7b165edf4	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
0f9ab8b73a5754d1116c6596bb8d7e0f5b15a922	LS48347 6.1	<i>Bacillus lentus</i> strain NCTC4824 genome assembly, chromosome: 1

0fa6a88864381a7d2b1dce850f1d6 0d6ab65403d	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
10755496d920a0ae59bc8b1a6ef7 b3d8de6a5f8b	BK03525 2.1	MAG TPA_asm: CrAss-like virus sp. isolate ctwJE30, partial genome
10fc92a5cdceb109dd7eea6b2db0c 2f183c5791f	CP05166 6.1	<i>Marinobacterium</i> sp. LSUCC0821 chromosome, complete genome
10fc92a5cdceb109dd7eea6b2db0c 2f183c5791f	CP08528 4.1	<i>Acinetobacter haemolyticus</i> strain w12 chromosome, complete genome
111f25a840a4e808f1db362be8192 9cb75bdf0b2	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
111f25a840a4e808f1db362be8192 9cb75bdf0b2	AP02416 9.1	Lachnospiraceae bacterium TB5 DNA, complete genome
11a635b03239ccd1a6c24efc17bea 94bbf9c7022	BK04065 8.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrSv9, partial genome
11a635b03239ccd1a6c24efc17bea 94bbf9c7022	BK04401 9.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctjuG8, partial genome
11a635b03239ccd1a6c24efc17bea 94bbf9c7022	BK04521 7.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct64N4, partial genome
11a635b03239ccd1a6c24efc17bea 94bbf9c7022	BK04829 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctGEt9, partial genome
11a635b03239ccd1a6c24efc17bea 94bbf9c7022	BK03604 9.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct1jZ5, partial genome
11a635b03239ccd1a6c24efc17bea 94bbf9c7022	BK01933 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct6Ou7, partial genome
11b98fa5c01679cb2b2bfa7f7c7e6 51f0d4901c5	CP02257 9.1	<i>Oryzomicrobium terrae</i> strain TPP412 chromosome, complete genome

11b98fa5c01679cb2b2bfa7f7c7e6 51f0d4901c5	CP02218 8.1	<i>Azoarcus communis</i> strain TSNA42 chromosome, complete genome
11b98fa5c01679cb2b2bfa7f7c7e6 51f0d4901c5	CP02218 7.1	<i>Azoarcus communis</i> strain TSPY31 chromosome, complete genome
11d4c6dd6e7eea8d8aef49b1eefb0 daa15938a79	NC_0393 75.1	<i>Helixanthera parasitica</i> chloroplast, complete genome
11d4c6dd6e7eea8d8aef49b1eefb0 daa15938a79	MG8080 38.1	<i>Helixanthera parasitica</i> chloroplast, complete genome
11d4c6dd6e7eea8d8aef49b1eefb0 daa15938a79	MG9994 88.1	<i>Helixanthera parasitica</i> voucher B. Liu 2140 (PE) tRNA-Leu (<i>trnL</i>) gene, partial sequence; <i>trnL-trnF</i> intergenic spacer, complete sequence; and tRNA-Phe (<i>trnF</i>) gene, partial sequence; chloroplast
11d4c6dd6e7eea8d8aef49b1eefb0 daa15938a79	MN0807 18.1	<i>Helixanthera parasitica</i> chloroplast, complete genome
11d4c6dd6e7eea8d8aef49b1eefb0 daa15938a79	EU54450 7.1	<i>Socratina bemarivensis</i> voucher Nickrent 4179 tRNA-Leu (<i>trnL</i>) gene and <i>trnL-trnF</i> intergenic spacer, partial sequence; chloroplast
12570e7c6cfafe8608ddfabdbdfb7 7fbac8dd1b3	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
1269543f4e0df76dbf58719f25047 782a44f85ee	CP07369 4.1	<i>Dysosmobacter</i> sp. Marseille-Q4140 chromosome
12ac7ac3d915d1a769c0c2508370 04ded4f0a144	CP01620 2.1	<i>Eubacterium minutum</i> ATCC 700079 sequence
12d442cb9254334b6e85934b5e4d afd15aeaf5f5	CP02300 4.1	<i>Ereboglobus luteus</i> strain Ho45 chromosome
143613e100af35ea387639a4252b 584f81c96965	HE79668 3.1	<i>Fibrella aestuarina</i> BUZ 2 drat genome

146e35782378db5bdd320223a05ebb19eeabc62d	CP00193 6.1	<i>Thermoanaerobacter italicus</i> Ab9, complete genome
149ffc96cc3248d0de5634cd172812699bf1a206	CP06063 6.1	<i>Eubacterium hominis</i> strain NSJ-61 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07201 9.1	<i>Akkermansia muciniphila</i> strain Akk2680 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07202 1.1	<i>Akkermansia muciniphila</i> strain Akk2196 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07202 2.1	<i>Akkermansia muciniphila</i> strain Akk2190 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07202 5.1	<i>Akkermansia muciniphila</i> strain Akk2000 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07203 6.1	<i>Akkermansia muciniphila</i> strain Akk1573 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07203 8.1	<i>Akkermansia muciniphila</i> strain Akk1476 chromosome, complete genome
16297509ec8701fc469f8c0a5253a08361351324	CP07204 4.1	<i>Akkermansia muciniphila</i> strain Akk0580 chromosome, complete genome
16bb4909c0d5177e7cb9d917b5bf62934af2aebf	CP07069 6.1	MAG: Spirochaetaceae bacterium isolate bin59 chromosome
17acd829f2a3933b66c088af247f87ccb3ec6795	CP03881 7.1	<i>Haemophilus parahaemolyticus</i> strain NEB129 chromosome, complete genome
17acd829f2a3933b66c088af247f87ccb3ec6795	CP06951 0.1	<i>Haemophilus parahaemolyticus</i> strain FDAARGOS_1199 chromosome, complete genome
17c4d90b2b1f1fe9f59bc1f64f2e75dba9ad231b	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence

18868912e0d8b6f46cda5f15687ea 19722e7fc9a	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
19504e7851f665ea79b35a6712ab 08abe97e0eb6	KF05212 2.1	<i>Campylobacter fetus</i> strain ML7 16S ribosomal RNA gene, partial sequence; 16S-23S ribosomal RNA intergenic spacer, complete sequence; and 23S ribosomal RNA gene, partial sequence
19a987e663b76c86e6ee2ff234bb6 083f3737849	CP02298 3.1	<i>Cytobacillus kochii</i> strain BDGP4 chromosome, complete genome
1b9abe1caaaa93e686e5710e3e987 e1ce5618190	CP07457 5.1	<i>Tenericutes</i> bacterium strain zrk6 chromosome, complete genome
1cd88f319f7b238e7c0838a9d9000 3e14d71df72	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	LK99601 7.1	<i>Desulfitobacterium hafniense</i> strain PCE-S genome assembly, scaffold: scaffold0
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	LR74649 6.1	Peptococcaceae bacterium INE isolate PacBioINE genome assembly, chromosome: 1
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome

1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
1d45e4050bfa70a664e84a6afb4b8 754fa05bc9e	AP00823 0.1	<i>Desulfitobacterium hafniense</i> Y51 DNA, complete genome
1de20564558bf16589c1a3081927 0a952d73ed18	LS48344 7.1	<i>Porphyromonas crevioricanis</i> strain NCTC12858 genome assembly, chromosome: 1
1eb48eb289470ce7c49fee2655d24 959962643dc	LR79268 4.1	<i>Kyrpidia spormannii</i> isolate FAVT5 genome assembly, chromosome: 1
1eb48eb289470ce7c49fee2655d24 959962643dc	LR79268 3.1	<i>Kyrpidia spormannii</i> isolate COOX1 genome assembly, chromosome: 1
1eb48eb289470ce7c49fee2655d24 959962643dc	CP02495 5.1	<i>Kyrpidia spormannii</i> strain EA-1 chromosome, complete genome
1eb48eb289470ce7c49fee2655d24 959962643dc	CP02495 5.1	<i>Kyrpidia spormannii</i> strain EA-1 chromosome, complete genome
1eb48eb289470ce7c49fee2655d24 959962643dc	CP04797 2.1	<i>Kyrpidia spormannii</i> strain COOX1 chromosome, complete genome
1eb48eb289470ce7c49fee2655d24 959962643dc	CP04797 1.1	<i>Kyrpidia spormannii</i> strain FAVT5 chromosome, complete genome
1ee41b9bbbd002957f608fffc9edb 84ee79fc77e	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
1ef12a949fc40eb57f19385294aad 6248ab0199c	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
1f2b1055224acc0a46ccd8459999 05cb0aaff84a	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
20bbc039060694b1daac93ff2f81d ed787e2907e	BK01640 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctcu031, partial genome

20e50efba8a25f10b279b333677fb d0ddfa7bec1	LT82712 8.1	<i>Gordonibacter</i> sp. Marseille-P2775 genome assembly, chromosome: contig00001
213d3eaaad82f791949924a75fdb4 f1d32271c1b	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
2155e88417240b5d4048b97f4839 0e916e539681	CP04971 4.1	<i>Apibacter</i> sp. B2966 chromosome, complete genome
2155e88417240b5d4048b97f4839 0e916e539681	CP04971 5.1	<i>Apibacter</i> sp. B3706 chromosome, complete genome
2155e88417240b5d4048b97f4839 0e916e539681	CP04826 7.1	<i>Apibacter</i> sp. ESL0404 chromosome, complete genome
2155e88417240b5d4048b97f4839 0e916e539681	CP04826 6.1	<i>Apibacter</i> sp. ESL0432 chromosome, complete genome
21929b8ab8e5f9b5e77b417c1643 16379f41b9eb	CP05468 9.1	Ectothiorhodospiraceae bacterium 2226 chromosome, complete genome
229a0cb38518f5cbc8a2ba357a3f9 db470e7bbbf	AP02281 0.1	<i>Desulfuromonas</i> sp. AOP6 DNA, complete genome
22aa05ca07e1f01dc06b23f465351 7e22be373e1	LK00535 7.1	TPA: <i>Clostridium kluyveri</i> DSM 555 tRNA-Met-CAT-2-1 gene
22aa05ca07e1f01dc06b23f465351 7e22be373e1	CP00067 3.1	<i>Clostridium kluyveri</i> DSM 555, complete genome
236a7695d9398231ae76d8254edc 87b2d66b7913	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
23a722909088e95fe33caeb6887f8 717a7e33bcc	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
2411c60fbc117962d34d7b729302 6681646b965c	CP10111 2.1	<i>Moraxella</i> sp. FZLJ2109 chromosome, complete genome

2411c60fbc117962d34d7b729302 6681646b965c	CP10111 1.1	<i>Moraxella</i> sp. FZLJ2107 chromosome, complete genome
2411c60fbc117962d34d7b729302 6681646b965c	CP09996 0.1	<i>Moraxella</i> sp. FZFQ2102 chromosome, complete genome
2563c4109e35c360ae143e2442f5 b8f4301b01ca	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
2563c4109e35c360ae143e2442f5 b8f4301b01ca	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
2609fe8d6185b2916ad7f3dccf095 4d482f651b0	LR69898 1.1	<i>Clostridiales</i> bacterium isolate MGYG-HGUT-02340 genome assembly, chromosome: 1
26483f1181a1b3ac253e67159447 cb7098c0e1ea	CP01152 4.1	<i>Intestinimonas butyriciproducens</i> strain SRB-521-5-I chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP04887 5.1	<i>Streptomyces albus</i> strain CAS922 chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP02924 1.1	<i>Streptomyces</i> sp. NHF165 chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP02961 8.1	<i>Streptomyces</i> sp. WAC 06738 chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP01149 2.1	<i>Streptomyces</i> sp. CNQ-509 chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP09384 6.1	<i>Streptomyces</i> sp. 3025 chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP05094 8.1	<i>Iamia</i> sp. SCSIO 61187 chromosome, complete genome
26517e030b1b9c091bfd875e0e71 89862c866737	CP07760 4.1	<i>Aquihabitans</i> sp. G128 chromosome, complete genome

26517e030b1b9c091bfd875e0e71 89862c866737	CP04888 2.1	<i>Streptomyces bathyalis</i> strain ASO4wet chromosome, complete genome
270ad5d8e7a6d4832391470c9943 b5bd8d35e2a8	CP08528 4.1	<i>Acinetobacter haemolyticus</i> strain w12 chromosome, complete genome
270ad5d8e7a6d4832391470c9943 b5bd8d35e2a8	CP06701 9.1	<i>Acinetobacter</i> sp. CS-2 chromosome, complete genome
270ad5d8e7a6d4832391470c9943 b5bd8d35e2a8	CP01280 8.1	<i>Acinetobacter equi</i> strain 114 chromosome, complete genome
270ad5d8e7a6d4832391470c9943 b5bd8d35e2a8	CP05907 9.1	<i>Acinetobacter</i> sp. TTH0-4 strain 1BD1 chromosome, complete genome
270ad5d8e7a6d4832391470c9943 b5bd8d35e2a8	CP01260 8.1	<i>Acinetobacter</i> sp. TTH0-4, complete genome
270ecf760bc9958d9045b358c74d 59800953b07b	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
271b5d99af239bacbf592d13e41f0 0f8ca7e8dc9	HF54561 6.1	<i>Ruminococcus bicirculans</i> chromosome I, complete genome
2859dfb2ade60937d4a4c1737252 9e939f81d7ae	CP02636 5.1	<i>Neobacillus thermocopriae</i> strain DUT50_236 chromosome, complete genome
2859dfb2ade60937d4a4c1737252 9e939f81d7ae	CP02636 5.1	<i>Neobacillus thermocopriae</i> strain DUT50_236 chromosome, complete genome
2859dfb2ade60937d4a4c1737252 9e939f81d7ae	CP02081 5.1	<i>Anoxybacillus flavithermus</i> strain DSM 2641T chromosome, complete genome
2859dfb2ade60937d4a4c1737252 9e939f81d7ae	CP02081 5.1	<i>Anoxybacillus flavithermus</i> strain DSM 2641T chromosome, complete genome
2859dfb2ade60937d4a4c1737252 9e939f81d7ae	CP02183 8.1	<i>Anoxybacillus flavithermus</i> strain 52-1A chromosome, complete genome

2859dfb2ade60937d4a4c1737252 9e939f81d7ae	CP02183 8.1	<i>Anoxybacillus flavithermus</i> strain 52-1A chromosome, complete genome
29a8944edf2c0c04c8916a8570c30 8179b88c2dd	CP06499 3.1	MAG: <i>Candidatus Saccharibacteria</i> bacterium isolate Hirt_18-Q3-R61-65_MAXAC.330_cln chromosome, complete genome
29f64365a9f87b6b4b358bf30f282 1eaae5a07df	CP04987 1.1	<i>Sphingomonas sinipercae</i> strain HDW15C chromosome, complete genome
29f64365a9f87b6b4b358bf30f282 1eaae5a07df	CP05985 1.1	<i>Sandaracinobacteroides saxicola</i> strain M6 chromosome, complete genome
29f8895dcea41be3df5b754e0e2a0 fc2723c494b	AP02332 1.1	<i>Solibaculum mannosilyticum</i> 12CBH8 DNA, complete genome
2a3828919f41dc7a7f1a6537ab936 b9429c8832e	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
2a6a16019044322bff4f3dfc183ed 996f31e5cc5	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
2a6a16019044322bff4f3dfc183ed 996f31e5cc5	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
2b4a321f3a2c0fe67657e18249404 481472e4ad8	BK05022 1.1	MAG TPA_asm: <i>Bacteriophage</i> sp. isolate ctkor1, partial genome
2b7923085deb2435bf81154e596c ee1fc7cdf4e1	CP04163 5.1	<i>Mycetocola</i> sp. JXN-3 chromosome, complete genome
2c36da3aef2dbb1d4c84dd4196aa5 b3e82637260	CP00334 5.1	<i>Bernardetia litoralis</i> DSM 6794 chromosome, complete genome
2c65c7dd86a3a234c451b18981e7 ad952e6ec645	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
2c65c7dd86a3a234c451b18981e7 ad952e6ec645	CP03939 6.1	<i>Duncaniella dubosii</i> strain H5 chromosome

2c65c7dd86a3a234c451b18981e7 ad952e6ec645	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LR59585 8.1	<i>Streptococcus pseudoporcinus</i> strain NCTC10228 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LR59405 2.1	<i>Streptococcus porcinus</i> strain NCTC10924 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LR59405 0.1	<i>Streptococcus porcinus</i> strain NCTC10925 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LR59403 6.1	<i>Streptococcus porcinus</i> strain NCTC10710 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LR59403 5.1	<i>Streptococcus pseudoporcinus</i> strain NCTC5385 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP03240 1.1	<i>Streptococcus iniae</i> strain GX005 chromosome, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP03240 0.1	<i>Streptococcus iniae</i> strain YM011 chromosome, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LR13434 1.1	<i>Streptococcus pseudoporcinus</i> strain NCTC13786 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	LS48338 8.1	<i>Streptococcus porcinus</i> strain NCTC10999 genome assembly, chromosome: 1
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP02484 3.1	<i>Streptococcus iniae</i> strain FP5228 chromosome, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP02239 2.1	<i>Streptococcus iniae</i> strain QMA0248 chromosome, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP01795 2.1	<i>Streptococcus iniae</i> strain 89353 chromosome, complete genome

2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP01078 3.1	<i>Streptococcus iniae</i> strain YSFST01-82, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP00758 7.1	<i>Streptococcus iniae</i> strain ISNO, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP00758 6.1	<i>Streptococcus iniae</i> strain ISET0901, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP00594 1.1	<i>Streptococcus iniae</i> SF1, complete genome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP07023 7.1	<i>Streptococcus porcinus</i> strain Colony597 chromosome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP07023 6.1	<i>Streptococcus porcinus</i> strain Colony573 chromosome
2ceac2b9a7ea3b24a4ba1154ade50 d5b7c23c355	CP00247 1.1	<i>Streptococcus parauberis</i> KCTC 11537, complete genome
2d2d4b09d5b4333d8bc35cfa4afd d5fd675626c	AP01971 1.1	<i>Amedibacterium intestinale</i> DNA, complete genome
2d2d4b09d5b4333d8bc35cfa4afd d5fd675626c	AP01971 1.1	<i>Amedibacterium intestinale</i> DNA, complete genome
2d2d4b09d5b4333d8bc35cfa4afd d5fd675626c	AP01969 5.1	<i>Amedibacterium intestinale</i> JCM 30884 DNA, complete genome
2d2d4b09d5b4333d8bc35cfa4afd d5fd675626c	AP01969 5.1	<i>Amedibacterium intestinale</i> JCM 30884 DNA, complete genome
2d2d4b09d5b4333d8bc35cfa4afd d5fd675626c	AP02451 0.1	<i>Longicatena caecimuris</i> 3BBH23 DNA, complete genome
2d37decaf3febce01477701844c0a 46b4aeabe11	CP09172 6.1	MAG: <i>Bacilli</i> bacterium isolate nC33_bin.244.fa chromosome

2d37decaf3febce01477701844c0a46b4aeabe11	BK01993 2.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctDKc20, partial genome
2d38796709e4f25ddbef4ea1a95127995123454a	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
2e8d114c40aad2374c195d653bf1f55b23f4ca6c	AP01733 9.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F28
2e8d114c40aad2374c195d653bf1f55b23f4ca6c	AP01733 5.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F211
2e8d114c40aad2374c195d653bf1f55b23f4ca6c	AP01732 9.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F13
2ece04a8afd50144e3bc8921b77a53b1bcacc4a2	CP09171 8.1	MAG: <i>Oscillospiraceae</i> bacterium isolate nC33_bin.8.fa chromosome
2ee265033d0b7b9aa5dc877a99d7dd681c3ba24a	CP08362 9.1	<i>Porphyromonas endodontalis</i> strain FDAARGOS_1506 chromosome, complete genome
301f3ceda67bfec8366dd4446d9a30ad70ae61d8	CP06941 9.1	<i>Clostridiales</i> bacterium strain XBB3002 chromosome, complete genome
310c0b5561669a74b8db9d49b8fc8ecea4a95258	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
32776ba09969e30104bb5ca15aac a549d95604b4	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
32776ba09969e30104bb5ca15aac a549d95604b4	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
32776ba09969e30104bb5ca15aac a549d95604b4	OU94086 2.1	uncultured fungus genomic DNA sequence contains 18S rRNA gene, ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gene, isolate consensus02156

3313026134ff4624fcbacea3f39f93 76f104692a	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
35f138040a2c0abbffec1f5383222 eb3c469f43d	BK05508 4.1	MAG TPA_asm: <i>Microviridae</i> sp. isolate ctSXA1, partial genome
3669991686c4c2505c4ddca3b4c2 801921be8502	OW2037 21.1	<i>Yponomeuta sedellus</i> genome assembly, chromosome: Z
3750a159183bd6f938ce7b4c4637 cc73e670fd47	LS48344 7.1	<i>Porphyromonas crevioricanis</i> strain NCTC12858 genome assembly, chromosome: 1
3838b7f4eed86dae48ff7d364ff6d4 b0020ee2cd	CP00093 0.2	<i>Heliomicrobium modesticaldum</i> Ice1 chromosome, complete genome
3842b410643381f488c31d816a04 64f58dc7c284	CP04565 2.1	<i>Sphingobacterium</i> sp. dk4302 chromosome, complete genome
39df0cc0fe7a41cc8b6ed4a35994b 11368fa3337	CP02108 1.1	<i>Deinococcus ficus</i> strain CC-FR2-10 chromosome, complete genome
3a6d4ddc4117436db4e74fd8b746 aacb63382b90	AP02266 0.1	<i>Bacteroides thetaiotaomicron</i> F9-2 DNA, nearly complete genome
3a6d4ddc4117436db4e74fd8b746 aacb63382b90	CP04123 0.1	<i>Bacteroides xylanisolvens</i> strain H207 chromosome, complete genome
3a6d4ddc4117436db4e74fd8b746 aacb63382b90	CP03655 3.1	<i>Bacteroides fragilis</i> strain DCMOUH0067B chromosome, complete genome
3a6d4ddc4117436db4e74fd8b746 aacb63382b90	CP01540 1.2	<i>Bacteroides caecimuris</i> strain I48 chromosome, complete genome
3a88dcfcd2fb4c2450462eea2577f c02e4d52a01	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
3a88dcfcd2fb4c2450462eea2577f c02e4d52a01	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome

3afceb81d28e6a8816048f57a8e62 f4e9e863ff7	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP04765 2.1	<i>Bombella</i> sp. KACC 21507 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP04304 3.1	<i>Gluconobacter thailandicus</i> strain HD924 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP03248 5.1	<i>Neokomagataea tanensis</i> strain AH13 = NBRC 106556 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP03814 3.1	<i>Oecophyllibacter saccharovorans</i> strain Ha5 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	LT90033 8.1	<i>Gluconobacter oxydans</i> 621H isolate WT-DSMZ genome assembly, chromosome: 1
3b675d5344e24062959c52004dc2 3951b13fc75f	CP01468 9.1	<i>Gluconobacter albidus</i> strain TMW2.1191 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP01469 1.1	<i>Neoasaia chiangmaiensis</i> strain NBRC 101099 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP01468 1.1	<i>Kozakia baliensis</i> strain NBRC 16680, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP01467 4.1	<i>Kozakia baliensis</i> strain DSM 14400, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP01632 8.1	<i>Gluconobacter oxydans</i> strain 1.637 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP09273 4.1	Acetobacteraceae bacterium DM15PD chromosome cDM15PD, complete sequence
3b675d5344e24062959c52004dc2 3951b13fc75f	CP09268 9.1	<i>Gluconobacter frateurii</i> strain ML.ISBL3 chromosome, complete genome

3b675d5344e24062959c52004dc2 3951b13fc75f	CP00437 3.1	<i>Gluconobacter oxydans</i> DSM 3504, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP06841 9.1	<i>Gluconobacter sphaericus</i> strain SJF2-1 chromosome, complete genome
3b675d5344e24062959c52004dc2 3951b13fc75f	CP00000 9.1	<i>Gluconobacter oxydans</i> 621H, complete genome
3c11e23c4ee65bf8da5dc9516e433 e9dadffbd3	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
3c79125f1d72a98bd542719ae6bf5 293f0595d40	CP04642 7.1	<i>Phocaeicola dorei</i> strain JR04 chromosome, complete genome
3c79125f1d72a98bd542719ae6bf5 293f0595d40	CP00874 1.1	<i>Bacteroides dorei</i> isolate HS1_L_3_B_079, complete genome
3cfc363d4435f9d5002aa1c2a49f7 7eb08acff52	CP09171 8.1	MAG: Oscillospiraceae bacterium isolate nC33_bin.8.fa chromosome
3d51df56496a788a4aae12e1254fe 43abd76ff98	BK04065 8.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrSv9, partial genome
3d51df56496a788a4aae12e1254fe 43abd76ff98	BK04401 9.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctjuG8, partial genome
3d51df56496a788a4aae12e1254fe 43abd76ff98	BK04521 7.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct64N4, partial genome
3d51df56496a788a4aae12e1254fe 43abd76ff98	BK04829 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctGEt9, partial genome
3d51df56496a788a4aae12e1254fe 43abd76ff98	BK03604 9.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct1jZ5, partial genome
3d51df56496a788a4aae12e1254fe 43abd76ff98	BK01933 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct6Ou7, partial genome

3dbdc6faefd0afd0ae6047a657216ef357bfdcad	CP00326 1.1	<i>Clostridium pasteurianum</i> BC1, complete genome
3e1d569ff2c91e0cdae8da257e5ff95448a5d959	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22l31, partial genome
3e1d569ff2c91e0cdae8da257e5ff95448a5d959	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LR59585 8.1	<i>Streptococcus pseudoporcinus</i> strain NCTC10228 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LR59405 2.1	<i>Streptococcus porcinus</i> strain NCTC10924 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LR59405 0.1	<i>Streptococcus porcinus</i> strain NCTC10925 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LR59403 6.1	<i>Streptococcus porcinus</i> strain NCTC10710 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LR59403 5.1	<i>Streptococcus pseudoporcinus</i> strain NCTC5385 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LR13434 1.1	<i>Streptococcus pseudoporcinus</i> strain NCTC13786 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	LS48338 8.1	<i>Streptococcus porcinus</i> strain NCTC10999 genome assembly, chromosome: 1
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	CP07023 7.1	<i>Streptococcus porcinus</i> strain Colony597 chromosome
3e2aaa2eb406d68fa0b5f1a8297f49b9d7e99981	CP07023 6.1	<i>Streptococcus porcinus</i> strain Colony573 chromosome
3e4cbbae6d462cf4b1e7a789ff4c870f4eda5c29	CP04299 7.1	<i>Aquisphaera giovannonii</i> strain OJF2 chromosome, complete genome

3e7cef2099f9ee43fb93e2deb65ef5bea64cace9	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
3ee5def822c19f334fad944635c2053304b48c63	CP04914 3.1	<i>Roseimicrobium</i> sp. ORNL1 chromosome, complete genome
3ee5def822c19f334fad944635c2053304b48c63	CP00355 8.1	<i>Synechococcus</i> sp. PCC 6312, complete genome
3ef6749ed52adf91d2e32c27ccc9cc09ecd7b126	CP00260 6.1	<i>Hippea maritima</i> DSM 10411, complete genome
3fe39342085cb5989e1ecdd3513274ad47214f7a	CP05470 6.1	<i>Salicibibacter cibi</i> strain NKC21-4 chromosome, complete genome
40e06e627021ab738855fd0e913c1294134f90e7	CP04527 2.1	<i>Priestia megaterium</i> strain FDU301 chromosome, complete genome
40e06e627021ab738855fd0e913c1294134f90e7	CP04929 6.1	<i>Priestia megaterium</i> strain S188 chromosome, complete genome
40e06e627021ab738855fd0e913c1294134f90e7	CP04769 9.1	<i>Priestia megaterium</i> strain 5-3 chromosome, complete genome
40e06e627021ab738855fd0e913c1294134f90e7	CP04454 5.1	<i>Pradoshia</i> sp. D12 chromosome, complete genome
40e06e627021ab738855fd0e913c1294134f90e7	CP03252 7.2	<i>Priestia megaterium</i> NCT-2 chromosome, complete genome
40e06e627021ab738855fd0e913c1294134f90e7	CP03177 6.1	<i>Priestia megaterium</i> strain Ni2-3 chromosome
40e06e627021ab738855fd0e913c1294134f90e7	CP03177 7.1	<i>Priestia megaterium</i> strain Mn1-4 chromosome
40e06e627021ab738855fd0e913c1294134f90e7	CP04151 9.1	<i>Priestia aryabhattai</i> strain KNU10 chromosome, complete genome

40e06e627021ab738855fd0e913c 1294134f90e7	CP04106 6.1	<i>Priestia megaterium</i> strain KNU-01 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP03509 4.1	<i>Priestia megaterium</i> NBRC 15308 = ATCC 14581 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP02808 4.1	<i>Priestia megaterium</i> strain SGAir0080 chromosome
40e06e627021ab738855fd0e913c 1294134f90e7	CP02267 4.1	<i>Priestia megaterium</i> strain SR7 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP02479 4.1	<i>Bacillus</i> sp. Y-01 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP01887 4.1	<i>Priestia megaterium</i> strain JX285 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP01522 6.1	<i>Bacillus</i> sp. IHB B 7164 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP00992 0.1	<i>Bacillus megaterium</i> NBRC 15308 = ATCC 14581, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP08544 3.1	<i>Priestia megaterium</i> strain KF18 plasmid unnamed1, complete sequence
40e06e627021ab738855fd0e913c 1294134f90e7	CP08543 1.1	<i>Priestia megaterium</i> strain F20 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP08542 6.1	<i>Priestia megaterium</i> strain F16 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP08541 5.1	<i>Priestia megaterium</i> strain CK7 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP09861 0.1	<i>Priestia megaterium</i> strain Ag87 chromosome

40e06e627021ab738855fd0e913c 1294134f90e7	CP06427 5.1	<i>Priestia megaterium</i> strain B1 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP09238 7.1	<i>Priestia megaterium</i> strain BP01R2 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP05945 7.1	<i>Priestia megaterium</i> strain 2020WEIHUA_L chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP07247 4.1	<i>Priestia aryabhattai</i> strain LAD plasmid p-LAD1, complete sequence
40e06e627021ab738855fd0e913c 1294134f90e7	CP07146 6.1	<i>Priestia megaterium</i> strain H2 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP06960 6.1	<i>Priestia megaterium</i> strain CDC 2008724142 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP06928 8.1	<i>Priestia megaterium</i> strain ATCC 14581 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP04888 3.1	<i>Priestia megaterium</i> strain Mes11 chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP05826 2.1	<i>Priestia megaterium</i> strain BIM B-1314D chromosome, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP00198 2.1	<i>Bacillus megaterium</i> DSM319, complete genome
40e06e627021ab738855fd0e913c 1294134f90e7	CP00198 3.1	<i>Bacillus megaterium</i> QM B1551, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	AP02266 0.1	<i>Bacteroides thetaiotaomicron</i> F9-2 DNA, nearly complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP04137 9.1	<i>Bacteroides intestinalis</i> strain APC919/174 chromosome, complete genome

41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP04123 0.1	<i>Bacteroides xylanisolvens</i> strain H207 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP01540 1.2	<i>Bacteroides caecimuris</i> strain I48 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP06494 0.1	<i>Bacteroides intestinalis</i> strain APC919/174 substr. 8W chromosome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP06494 1.1	<i>Bacteroides intestinalis</i> strain APC919/174 substr. 8T chromosome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP08190 2.1	<i>Bacteroides salyersiae</i> strain BFG-256 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP08189 9.1	<i>Bacteroides salyersiae</i> strain BFG-288 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP08189 8.1	<i>Bacteroides thetaiotaomicron</i> strain BFG-55 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP07224 2.1	<i>Bacteroides thetaiotaomicron</i> strain CL06T03C18 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP07222 4.1	<i>Bacteroides thetaiotaomicron</i> strain CL11T00C24 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP07221 6.1	<i>Bacteroides xylanisolvens</i> strain CL11T00C03 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP07221 2.1	<i>Bacteroides xylanisolvens</i> strain CL11T00C41 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP06817 4.1	<i>Bacteroides caccae</i> strain FDAARGOS_1097 chromosome, complete genome
41040eaf26f1ff67df02c587c915c1 d3ea6657a8	CP05997 3.1	<i>Bacteroides luhongzhouii</i> strain HF-5141 chromosome, complete genome

41040eaf26f1ff67df02c587c915c1d3ea6657a8	CP05985 6.1	<i>Bacteroides zhangwenhongii</i> strain HF-5287 chromosome, complete genome
412babf8c5c6ef3444cb503dba44a4d3f9d7922	CP02636 5.1	<i>Neobacillus thermocopriae</i> strain DUT50_236 chromosome, complete genome
412babf8c5c6ef3444cb503dba44a4d3f9d7922	CP02636 5.1	<i>Neobacillus thermocopriae</i> strain DUT50_236 chromosome, complete genome
412babf8c5c6ef3444cb503dba44a4d3f9d7922	CP02081 5.1	<i>Anoxybacillus flavithermus</i> strain DSM 2641T chromosome, complete genome
412babf8c5c6ef3444cb503dba44a4d3f9d7922	CP02081 5.1	<i>Anoxybacillus flavithermus</i> strain DSM 2641T chromosome, complete genome
412babf8c5c6ef3444cb503dba44a4d3f9d7922	CP02183 8.1	<i>Anoxybacillus flavithermus</i> strain 52-1A chromosome, complete genome
412babf8c5c6ef3444cb503dba44a4d3f9d7922	CP02183 8.1	<i>Anoxybacillus flavithermus</i> strain 52-1A chromosome, complete genome
41a23ecba70d6367ad66da664c5b65ac07d03913	OV78864 0.1	uncultured <i>Tenericutes</i> bacterium isolate min17_bin11 genome assembly, chromosome: 1
41e1fda770cc1e4c2c85b0736f221724fe6ca9ef	BK05850 2.1	MAG TPA_asm: <i>Inoviridae</i> sp. isolate ctwJm14, partial genome
4235bff0e2dac3ff287fa9cc98f9b53267be7134	CP04792 1.1	<i>Candidatus Saccharibacteria</i> bacterium oral taxon 488 strain FS05P-B chromosome, complete genome
4235bff0e2dac3ff287fa9cc98f9b53267be7134	CP04792 0.1	<i>Candidatus Saccharibacteria</i> bacterium oral taxon 488 strain FS07P chromosome, complete genome
4351c384ad525087df328780a54d213332ff2bb3	CP09171 8.1	<i>Candidatus Saccharibacteria</i> bacterium oral taxon 488 strain FS07P chromosome, complete genome
43696ffc129f5129d47611368bd9d397d55b625c	CP09283 2.1	<i>Flavobacterium</i> sp. HJ-32-4 chromosome, complete genome

43c37fbbc8a2ac4012c8cc768134f9864630d0d2	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
43ed299db5a0d1f7c07d8e36c071c8c8ff6596c3	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
442e614fd0aa8df705c8a5959ef1807288f39952	CP00263 3.1	<i>Streptococcus suis</i> ST3, complete genome
44aea880792ad51e4a6950a811d2c9eb3a85b8b2	OV78864 0.1	uncultured <i>Tenericutes</i> bacterium isolate min17_bin11 genome assembly, chromosome: 1
44b626b8dd622bc8a69b79e4a752f046d0035dc8	AP02332 1.1	<i>Solibaculum mannosilyticum</i> 12CBH8 DNA, complete genome
45902a9b8f171f2ef4c479b327bed06db6a137cf	BK02513 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctmsk9, partial genome
46409344d0bc5604e234584edf91814134c8e9bc	BK02910 4.1	MAG TPA_asm: <i>Caudovirales</i> sp. isolate ctThO3, partial genome
479ff4eb81d8e8dbc5a645800b73501002beb46a	BK05850 2.1	MAG TPA_asm: <i>Inoviridae</i> sp. isolate ctwJm14, partial genome
47c89ea5b5964474895d80588258372a7ef1582e	CT97158 3.1	<i>Synechococcus</i> WH7803 complete genome sequence
47d46896acad200ae6c487e4f8e2b033466219a1	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
47f4f575fcf429dfa147369392578f8563fe208a	BK02908 5.1	AG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22l31, partial genome
47f4f575fcf429dfa147369392578f8563fe208a	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
4813590bdb4e30cc87d320c9bd04d67f45022c9a	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome

4813590bdb4e30cc87d320c9bd04 d67f45022c9a	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
4925163bb23891d033bbcc2d9b2f fd13ef4dfb41	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22l31, partial genome
4925163bb23891d033bbcc2d9b2f fd13ef4dfb41	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
4925163bb23891d033bbcc2d9b2f fd13ef4dfb41	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
496ad5b7832d0c78ae303f5e5fcb1 d675dfbd0b4	CP04986 0.1	<i>Erysipelothrix</i> sp. HDW6B chromosome, complete genome
4ad5c4ddfd81b9740a32a8b8328 448193594268	CP04246 6.1	<i>Geobacter</i> sp. FeAm09 chromosome, complete genome
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	LK99601 7.1	<i>Desulfitobacterium hafniense</i> strain PCE-S genome assembly, scaffold: scaffold0
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	LR74649 6.1	Peptococcaceae bacterium INE isolate PacBioINE genome assembly, chromosome: 1
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
4c1dfacee87be098b1a22c3dadc2a 37f843c23f5	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome

4c1dfacee87be098b1a22c3dad2a37f843c23f5	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
4c1dfacee87be098b1a22c3dad2a37f843c23f5	AP00823 0.1	<i>Desulfitobacterium hafniense</i> Y51 DNA, complete genome
4c9a8938e7fa7597dfa7ae963394e864e9830a27	CP00326 1.1	<i>Clostridium pasteurianum</i> BC1, complete genome
4caeb847954acc4cc162758563029a9def5699df2	FP92905 9.1	<i>Eubacterium siraeum</i> V10Sc8a draft genome
4caeb847954acc4cc162758563029a9def5699df2	FP92904 4.1	<i>Eubacterium siraeum</i> 70/3 draft genome
4d8b463de649db88c7af85317a3ce22d0ae5c608	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
4ebabe866a910a5cc2b0241ac6b99cf53cf8c48f	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
4f7f932f26e16fee38445c6cc407d21d221c0e89	CP04883 8.1	[<i>Clostridium</i>] <i>innocuum</i> strain ATCC 14501 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d21d221c0e89	CP04883 7.1	[<i>Clostridium</i>] <i>innocuum</i> strain LC-LUMC-CI-001 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d21d221c0e89	CP02272 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 genome
4f7f932f26e16fee38445c6cc407d21d221c0e89	CP01540 4.2	Erysipelotrichaceae bacterium I46 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d21d221c0e89	CP09494 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain VE303-07 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d21d221c0e89	AP02556 5.1	[<i>Clostridium</i>] <i>innocuum</i> CE91-St51 DNA, complete genome

4f7f932f26e16fee38445c6cc407d 21d221c0e89	BK03754 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctS5j6, partial genome
4f7f932f26e16fee38445c6cc407d 21d221c0e89	BK05411 2.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctWsN1, partial genome
4f7f932f26e16fee38445c6cc407d 21d221c0e89	CP07457 3.1	<i>Tenericutes</i> bacterium strain zrk4 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d 21d221c0e89	CP07457 3.1	<i>Tenericutes</i> bacterium strain zrk4 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d 21d221c0e89	CP04617 4.1	Erysipelotrichaceae bacterium 66202529 chromosome, complete genome
4f7f932f26e16fee38445c6cc407d 21d221c0e89	CP06532 0.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 chromosome
4fc83f8bb7af5f6351f8070e7c6722 5789fe7369	CT97158 3.1	<i>Synechococcus</i> WH7803 complete genome sequence
4fda729847256e5e723c27e6045c 587f4df66796	CP02468 4.1	<i>Bacillus wiedmannii</i> bv. <i>thuringiensis</i> strain FCC41 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP09233 5.1	<i>Bacillus</i> sp. OVS6 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP03164 2.1	<i>Bacillus anthracis</i> strain MCCC 1A02161 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP03236 5.1	<i>Bacillus wiedmannii</i> strain SR52 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08570 8.1	<i>Bacillus</i> sp. CMF21 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP09332 5.1	<i>Bacillus</i> sp. N5-665 strain N5/665 plasmid p1, complete sequence

4fda729847256e5e723c27e6045c 587f4df66796	CP04033 6.1	<i>Bacillus luti</i> strain FJ chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP04651 1.1	<i>Bacillus cereus</i> strain JHU chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP05456 8.1	<i>Bacillus thuringiensis</i> strain FDAARGOS_791 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP05393 1.1	<i>Bacillus cereus</i> strain FDAARGOS_797 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP05397 2.1	<i>Bacillus thuringiensis</i> strain FDAARGOS_796 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP05396 5.1	<i>Bacillus cereus</i> strain FDAARGOS_802 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP05398 1.1	<i>Bacillus thuringiensis</i> strain FDAARGOS_793 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP05185 8.1	<i>Bacillus thuringiensis</i> serovar israelensis strain BGSC 4Q7rifR chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP04868 7.1	<i>Bacillus paranthracis</i> strain MN1F chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP04198 1.1	<i>Bacillus paranthracis</i> strain NCCP 15910 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP04175 0.1	<i>Bacillus paranthracis</i> strain NCCP 14796 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP04708 5.1	<i>Bacillus paranthracis</i> strain BC307 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP04034 4.1	<i>Bacillus albus</i> strain DLOU-Yingkou chromosome, complete genome

4fda729847256e5e723c27e6045c 587f4df66796	CP10113 5.1	<i>Bacillus paranthracis</i> strain Bt C4 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP09980 1.1	<i>Bacillus paranthracis</i> strain SSBC101 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08541 3.1	<i>Bacillus paranthracis</i> strain KF11 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08540 9.1	<i>Bacillus thuringiensis</i> strain KF1 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08540 2.1	<i>Bacillus anthracis</i> strain CMF9 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08539 9.1	<i>Bacillus tropicus</i> strain CK18 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP09587 4.1	<i>Bacillus</i> sp. PGP15 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP09576 7.1	<i>Bacillus cereus</i> strain CDHWZ7 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP01531 0.1	<i>Bacillus thuringiensis</i> serovar <i>morrisoni</i> str. 4AA1 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP09445 5.1	<i>Bacillus</i> sp. ZJS3 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08952 1.1	<i>Bacillus thuringiensis</i> strain ZZQ-130 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	AP02300 0.1	<i>Bacillus cereus</i> 30077 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02299 4.1	<i>Bacillus cereus</i> 30075 DNA, complete genome

4fda729847256e5e723c27e6045c 587f4df66796	AP02298 6.1	<i>Bacillus cereus</i> 30052 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02297 8.1	<i>Bacillus cereus</i> 30043 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02297 5.1	<i>Bacillus cereus</i> 30040 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02297 0.1	<i>Bacillus cereus</i> J75 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02296 4.1	<i>Bacillus cereus</i> J62 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02295 6.1	<i>Bacillus cereus</i> J51 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02295 2.1	<i>Bacillus cereus</i> J39 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02293 4.1	<i>Bacillus cereus</i> J7 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02292 7.1	<i>Bacillus cereus</i> J2 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02292 1.1	<i>Bacillus cereus</i> J1 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02291 5.1	<i>Bacillus cereus</i> MRY14-0105 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02290 7.1	<i>Bacillus cereus</i> MRY14-0100 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02290 3.1	<i>Bacillus cereus</i> MRY14-0079 DNA, complete genome

4fda729847256e5e723c27e6045c 587f4df66796	AP02289 4.1	<i>Bacillus cereus</i> MRY14-0075 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02288 6.1	<i>Bacillus cereus</i> MRY14-0074 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02287 7.1	<i>Bacillus cereus</i> MRY14-0060 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02287 4.1	<i>Bacillus cereus</i> MRY14-0057 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	AP02285 7.1	<i>Bacillus cereus</i> MRY14-0045 DNA, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08974 3.1	<i>Bacillus paranthracis</i> strain SHOU-BC01 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08313 4.1	<i>Bacillus thuringiensis</i> strain BMP144 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP08312 9.1	<i>Bacillus thuringiensis</i> strain NB-176 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP07653 9.1	<i>Bacillus thuringiensis</i> strain GR007 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP06315 8.1	<i>Bacillus cereus</i> strain CF4-51 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP03599 9.1	<i>Bacillus mycoides</i> strain BPN37/1 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP07552 0.1	<i>Bacillus paranthracis</i> strain Colony502 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	CP07276 9.1	<i>Bacillus cereus</i> strain BC07 chromosome, complete genome

4fda729847256e5e723c27e6045c 587f4df66796	CP07276 6.1	<i>Bacillus cereus</i> strain BC06 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP07174 3.1	<i>Bacillus thuringiensis</i> strain TAND672 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	CP07033 9.1	<i>Bacillus cereus</i> strain VKM B-370 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP06924 4.1	<i>Bacillus paranthracis</i> strain BCCL1 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	CP06871 7.1	<i>Bacillus cereus</i> strain CH chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP06588 1.1	<i>Bacillus cereus</i> strain FDAARGOS_919 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	CP06565 0.1	<i>Bacillus cereus</i> strain FDAARGOS_918 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP06514 9.1	<i>Bacillus paranthracis</i> strain Gxun-30 chromosome, complete genome
4fda729847256e5e723c27e6045c 587f4df66796	CP06407 2.1	<i>Bacillus cereus</i> strain EFR-1 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	CP06407 9.1	<i>Bacillus paranthracis</i> strain EFR-4 chromosome
4fda729847256e5e723c27e6045c 587f4df66796	CP05365 6.2	<i>Bacillus cereus</i> strain CTMA_1571 chromosome, complete genome
4fec40e189aa872013892da2755f2 7ed71aaa395	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
501d5e941970e2fd0c903b0be336 460641c90eac	CP06026 1.1	MAG: Holosporaceae bacterium isolate HK-STAS-PROT-86 chromosome

508a2326622461402ae5cd917c8e fdfe721b77ec	AF16611 4.1	<i>Mesostigma viride</i> chloroplast DNA, complete genome
5129ddb0a258e1b6af7015a2c38b 64c72e62714d	CP02334 4.1	<i>Nibricoccus aquaticus</i> strain HZ-65 chromosome, complete genome
5129ddb0a258e1b6af7015a2c38b 64c72e62714d	LK01198 2.1	TPA: <i>Opitutus terrae</i> PB90-1 tRNA-Leu-GAG-1-1 gene
5129ddb0a258e1b6af7015a2c38b 64c72e62714d	CP00103 2.1	<i>Opitutus terrae</i> PB90-1, complete genome
5149792b3720cfe14bf96795896a 4f91aa778b91	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
515f52a541b838be67ca9dfc1e7af 5b842e69ac3	CP04033 6.1	<i>Bacillus luti</i> strain FJ chromosome, complete genome
515f52a541b838be67ca9dfc1e7af 5b842e69ac3	CP03599 9.1	<i>Bacillus mycoides</i> strain BPN37/1 chromosome, complete genome
515f52a541b838be67ca9dfc1e7af 5b842e69ac3	CP03799 2.1	<i>Bacillus mycoides</i> strain TH26 chromosome, complete genome
515f52a541b838be67ca9dfc1e7af 5b842e69ac3	CP03454 8.1	<i>Bacillus albus</i> strain PFYN01 chromosome, complete genome
520a081fa207789e0f25bf65d18cd 1a5da120e62	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
52e7af6ded3261ecaaba211d5616f 56043cea834	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	AP01853 6.1	Lachnospiraceae bacterium GAM79 DNA, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	AP01853 6.1	Lachnospiraceae bacterium GAM79 DNA, complete genome

532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP08593 8.1	<i>Lachnospira eligens</i> strain FDAARGOS_1570 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP08593 8.1	<i>Lachnospira eligens</i> strain FDAARGOS_1570 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP08593 8.1	<i>Lachnospira eligens</i> strain FDAARGOS_1570 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP06063 3.1	<i>Simiaoa sunii</i> strain NSJ-8 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP06063 3.1	<i>Simiaoa sunii</i> strain NSJ-8 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP00110 4.1	[<i>Eubacterium</i>] <i>eligens</i> ATCC 27750 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP00110 4.1	[<i>Eubacterium</i>] <i>eligens</i> ATCC 27750 chromosome, complete genome
532da2e9a1a826ba331699bf4146 9bbd0bd9bbb4	CP00110 4.1	[<i>Eubacterium</i>] <i>eligens</i> ATCC 27750 chromosome, complete genome
53bb1b62c2c6d2e31f59c0df93822 cc2575c9183	CP02914 5.1	<i>Hymenobacter nivis</i> strain NBRC 111535 chromosome, complete genome
53bb1b62c2c6d2e31f59c0df93822 cc2575c9183	CP01430 4.1	<i>Hymenobacter</i> sp. PAMC 26628 chromosome, complete genome
544605fe235d3e0efe29bedebd2e6 74d9ac61b5f	LR13051 8.1	<i>Staphylococcus aureus</i> strain BPH2986 genome assembly, chromosome: 1
544605fe235d3e0efe29bedebd2e6 74d9ac61b5f	CP02908 0.1	<i>Staphylococcus aureus</i> strain AR466 chromosome, complete genome
544605fe235d3e0efe29bedebd2e6 74d9ac61b5f	CP02110 5.1	<i>Staphylococcus aureus</i> strain CC5 chromosome, complete genome

544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP08055 6.1	<i>Staphylococcus aureus</i> strain HL17078 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP06814 3.1	<i>Staphylococcus aureus</i> strain FDAARGOS_1086 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP10112 3.1	<i>Staphylococcus aureus</i> strain MN8 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP10112 4.1	<i>Staphylococcus aureus</i> subsp. <i>aureus</i> RN4220 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP10042 8.1	<i>Staphylococcus aureus</i> strain RMSA24 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09152 5.1	<i>Staphylococcus aureus</i> strain N09CSA16 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09957 6.1	<i>Staphylococcus aureus</i> strain MRSA1369 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09949 5.1	<i>Staphylococcus aureus</i> strain JL28 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09949 7.1	<i>Staphylococcus aureus</i> strain 0213-M-4A chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09949 9.1	<i>Staphylococcus aureus</i> strain 0610-H-2A chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09950 2.1	<i>Staphylococcus aureus</i> strain JL42 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09950 4.1	<i>Staphylococcus aureus</i> strain S36 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09950 6.1	<i>Staphylococcus aureus</i> strain 0316-H-5A chromosome, complete genome

544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09950 8.1	<i>Staphylococcus aureus</i> strain S82 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 6.1	<i>Staphylococcus aureus</i> strain TCD2 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 7.1	<i>Staphylococcus aureus</i> strain TCD1 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 0.1	<i>Staphylococcus aureus</i> strain TCD10 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 9.1	<i>Staphylococcus aureus</i> strain TCD11 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 8.1	<i>Staphylococcus aureus</i> strain TCD12 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 2.1	<i>Staphylococcus aureus</i> strain TCD18 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 7.1	<i>Staphylococcus aureus</i> strain TCD13 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 3.1	<i>Staphylococcus aureus</i> strain TCD7 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 4.1	<i>Staphylococcus aureus</i> strain TCD4 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 5.1	<i>Staphylococcus aureus</i> strain TCD3 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 6.1	<i>Staphylococcus aureus</i> strain TCD14 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 1.1	<i>Staphylococcus aureus</i> strain TCD19 chromosome

544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 2.1	<i>Staphylococcus aureus</i> strain TCD8 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09866 8.1	<i>Staphylococcus aureus</i> strain TCD22 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 0.1	<i>Staphylococcus aureus</i> strain TCD20 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09868 1.1	<i>Staphylococcus aureus</i> strain TCD9 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 3.1	<i>Staphylococcus aureus</i> strain TCD17 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09866 7.1	<i>Staphylococcus aureus</i> strain TCD23 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09866 9.1	<i>Staphylococcus aureus</i> strain TCD21 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09867 5.1	<i>Staphylococcus aureus</i> strain TCD15 chromosome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02474 2.1	<i>Staphylococcus aureus</i> 59736 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02473 8.1	<i>Staphylococcus aureus</i> 59731 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02473 6.1	<i>Staphylococcus aureus</i> 59709 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02473 4.1	<i>Staphylococcus aureus</i> 59691 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02473 2.1	<i>Staphylococcus aureus</i> 59621 DNA, complete genome

544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02473 0.1	<i>Staphylococcus aureus</i> 59574 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02472 8.1	<i>Staphylococcus aureus</i> 59553 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02472 6.1	<i>Staphylococcus aureus</i> 59482 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	AP02472 3.1	<i>Staphylococcus aureus</i> 59458 DNA, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09466 3.1	<i>Staphylococcus aureus</i> strain 494 chromosome, complete genome
544605fe235d3e0efe29bedebd2e674d9ac61b5f	CP09825 7.1	<i>Staphylococcus aureus</i> strain M4 chromosome, complete genome
551ddd5617448cbeef10ed2bd35ccd219532a755	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
556387bee684223ed6e85b1451982e99db34cbb9	CP09446 7.1	<i>Faecalibacterium</i> sp. I4179 chromosome, complete genome
556387bee684223ed6e85b1451982e99db34cbb9	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
556387bee684223ed6e85b1451982e99db34cbb9	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
556387bee684223ed6e85b1451982e99db34cbb9	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome
556387bee684223ed6e85b1451982e99db34cbb9	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome
556387bee684223ed6e85b1451982e99db34cbb9	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome

556387bee684223ed6e85b145198 2e99db34cbb9	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
556387bee684223ed6e85b145198 2e99db34cbb9	BK02513 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctmsk9, partial genome
562553bece2f22fa72ab5658c06eb 7775e54acff	AP02332 1.1	<i>Solibaculum mannosilyticum</i> 12CBH8 DNA, complete genome
5675377618f9c55478f2391d0029 d629aad308f5	BK02132 2.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctjse2, partial genome
5677d64ac221ad83826f031d9e4c 63382e6d0b0e	CP00326 1.1	<i>Clostridium pasteurianum</i> BC1, complete genome
568d2d20a1f1465fb7eb0b36afeb1 5a1f1cd8ac5	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22l31, partial genome
568d2d20a1f1465fb7eb0b36afeb1 5a1f1cd8ac5	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
56d6260e517b025da0426b234c8d 220afbbeegee	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
575288a49e118cb4c02fbbd5db76 cbdada881852	CP09282 1.1	MAG: <i>Candidatus Nealonbacteria</i> bacterium isolate DGGOD1a chromosome, complete genome
578406580565cf9064fde6e23c722 6cbcf7beeb1	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22l31, partial genome
578406580565cf9064fde6e23c722 6cbcf7beeb1	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
5836f2bffedbd860cffb97801d754 827d366a458	BK02132 2.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctjse2, partial genome
590ac920be95687d67f1ff8825795 d92740bd7e2	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence

59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome
59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome
59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
59d5acafb3c8e0f1de9abec10f06ae 5df8a007ab	BK02513 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctmsk9, partial genome
5a9810b360b5331b06ba80a77e40 8a28ebf4a9ab	CP06499 3.1	MAG: <i>Candidatus Saccharibacteria</i> bacterium isolate Hirt_18-Q3-R61-65_MAXAC.330_cln chromosome, complete genome
5af4be93b1af185861b866e88b861 c958a452743	GQ87700 5.1	Uncultured organism clone VC1AI32TR genomic sequence
5b14dbb816c456be039b9754de93 bfe4734f1231	CP07457 3.1	<i>Tenericutes</i> bacterium strain zrk4 chromosome, complete genome
5d562be15b0e0c70adaba884eb20 3f00ae2c03d8	CP04985 9.1	<i>Erysipelothrix</i> sp. HDW6A chromosome, complete genome
5ec1844a43dfd11f833880d81d64a 1f15b250bec	CP07437 3.1	<i>Virgibacillus</i> sp. NKC19-16 chromosome, complete genome
5f9fb0517071602e1ebf43b1a6ea9 9919718a1e6	CP04151 9.1	<i>Priestia aryabhatai</i> strain KNU10 chromosome, complete genome

5f9fb0517071602e1ebf43b1a6ea9 9919718a1e6	CP04106 6.1	<i>Priestia megaterium</i> strain KNU-01 chromosome, complete genome
5fa8d87f1765321737cc59dec3539 bc241c1db64	LR77813 5.1	<i>Clostridiales</i> bacterium isolate BL-6 genome assembly, chromosome: CLOSBL6
5fc98a497b0feb87e3d5fc122b482 5ead79d5503	CP10111 2.1	<i>Moraxella</i> sp. FZLJ2109 chromosome, complete genome
5fc98a497b0feb87e3d5fc122b482 5ead79d5503	CP10111 1.1	<i>Moraxella</i> sp. FZLJ2107 chromosome, complete genome
5fc98a497b0feb87e3d5fc122b482 5ead79d5503	CP09996 0.1	<i>Moraxella</i> sp. FZFQ2102 chromosome, complete genome
5ff54e66b280e808472cd772ab4c8 c03ff0e8c63	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
607e28fdb0a692f927f8b081ee0f7 e341a19df05	LR99447 3.1	uncultured <i>Prevotellamassilia</i> sp. isolate <i>Prevotellamassilia</i> canine HQ MAG genome assembly, chromosome: 1
63c68333f22f3d46633e78413d78 9ded3d46ab42	AP01971 1.1	<i>Amedibacterium intestinale</i> DNA, complete genome
63c68333f22f3d46633e78413d78 9ded3d46ab42	AP01971 1.1	<i>Amedibacterium intestinale</i> DNA, complete genome
63c68333f22f3d46633e78413d78 9ded3d46ab42	AP01969 5.1	<i>Amedibacterium intestinale</i> JCM 30884 DNA, complete genome
63c68333f22f3d46633e78413d78 9ded3d46ab42	AP01969 5.1	<i>Amedibacterium intestinale</i> JCM 30884 DNA, complete genome
65b258fea9bca4a3dc25c73ec172c 130e0f347a4	CP07437 3.1	<i>Virgibacillus</i> sp. NKC19-16 chromosome, complete genome
66482c30735ee06da4f1105f2b435 5ee923640b4	CP04033 6.1	<i>Bacillus luti</i> strain FJ chromosome, complete genome

66482c30735ee06da4f1105f2b435 5ee923640b4	CP03599 9.1	<i>Bacillus mycoides</i> strain BPN37/1 chromosome, complete genome
66482c30735ee06da4f1105f2b435 5ee923640b4	CP03799 2.1	<i>Bacillus mycoides</i> strain TH26 chromosome, complete genome
66482c30735ee06da4f1105f2b435 5ee923640b4	CP03454 8.1	<i>Bacillus albus</i> strain PFYN01 chromosome, complete genome
66482c30735ee06da4f1105f2b435 5ee923640b4	CP00374 7.1	<i>Bacillus cereus</i> FRI-35, complete genome
674c498398e4db59609e6d2dd568 8bdac59b27f0	CP02298 3.1	<i>Cytobacillus kochii</i> strain BDGP4 chromosome, complete genome
68b8a132ff31468a355ee7ddba104 bb28263cc8a	GQ87700 5.1	Uncultured organism clone VC1AI32TR genomic sequence
69b8ccc8cb8a18ca3d73b7ede38ac b61a384f1f7	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
69b8ccc8cb8a18ca3d73b7ede38ac b61a384f1f7	AP02416 9.1	Lachnospiraceae bacterium TB5 DNA, complete genome
6a41b02d978837b530ddb7d178ac a5f99373d5d4	AP01450 9.1	<i>Thermotoga caldifontis</i> AZM44c09 DNA, complete genome
6a41b02d978837b530ddb7d178ac a5f99373d5d4	CP00235 1.1	<i>Pseudothermotoga thermarum</i> DSM 5069 chromosome, complete genome
6b444bb8aeda16876cc2e8706e8fd ead8c332e4e	AP01733 9.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F28
6b444bb8aeda16876cc2e8706e8fd ead8c332e4e	AP01733 5.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F211
6b444bb8aeda16876cc2e8706e8fd ead8c332e4e	AP01732 9.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F13

6b444bb8aeda16876cc2e8706e8fd ead8c332e4e	CP09417 7.1	<i>Helicobacter pylori</i> strain Hpfe062 chromosome, complete genome
6b444bb8aeda16876cc2e8706e8fd ead8c332e4e	KF05212 2.1	<i>Campylobacter fetus</i> strain ML7 16S ribosomal RNA gene, partial sequence; 16S-23S ribosomal RNA intergenic spacer, complete sequence; and 23S ribosomal RNA gene, partial sequence
6b520087eccc20b45d3bfcc1ba744 b5c1bedb46a	CP02300 4.1	<i>Ereboglobus luteus</i> strain Ho45 chromosome
6b7949e44296b6cafdcb01b224e2 34e2f0415d8b	AP01483 6.1	<i>Candidatus Nitrosoglobus terrae</i> DNA, complete genome
6b827e24f2bcc9ab56ac48309743 04eb07698c87	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
6b9f61894a11ccea21d9d364e44c3 8125fe8fcbf	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
6b9f61894a11ccea21d9d364e44c3 8125fe8fcbf	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
6bae88b5ee607fbf016536bb0ef4a e25605d1350	CP05515 6.1	<i>Adhaeribacter swui</i> strain KCTC 52873 chromosome, complete genome
6c471fea63dbadd45355dcfa572ab 6923b4923cc	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
6c9ba3320aa02b2d103b8d123973 0a700d179718	CP01577 2.1	<i>Niabella ginsenosidivorans</i> strain BS26 chromosome, complete genome
6c9ba3320aa02b2d103b8d123973 0a700d179718	CP09290 0.1	<i>Candidatus Comchoanobacter bicostacola</i> strain Comchoano-1 chromosome, complete genome
6c9ba3320aa02b2d103b8d123973 0a700d179718	LK01176 1.1	TPA: <i>Nitrosococcus oceani</i> ATCC 19707 tRNA-Leu-CAA-1-1 gene

6c9ba3320aa02b2d103b8d123973 0a700d179718	CP08338 7.1	<i>Arachidicoccus terrestris</i> strain 5GH13-10 chromosome, complete genome
6c9ba3320aa02b2d103b8d123973 0a700d179718	CP00703 5.1	<i>Niabella soli</i> DSM 19437, complete genome
6c9ba3320aa02b2d103b8d123973 0a700d179718	OA93366 6.1	<i>Notodromas monacha</i>
6c9ba3320aa02b2d103b8d123973 0a700d179718	CP00012 7.1	<i>Nitrosococcus oceani</i> ATCC 19707, complete genome
6cc8c8cdc7bab6e600dea70663af6 171c68c19bc	CP07457 5.1	<i>Tenericutes</i> bacterium strain zrk6 chromosome, complete genome
6d0af9578cf121903c9e19b910825 a6231412efa	CP08130 3.1	Prolixibacteraceae bacterium strain KCTC 25031 chromosome, complete genome
6d1964f7550827e19d44a12fc263 5962caea15cb	CP06026 1.1	MAG: Holosporaceae bacterium isolate HK-STAS- PROT-86 chromosome
6e18c24aee2c6de33aeea94df96cfe c870cbba5b	CP04640 1.1	<i>Maribellus comscasis</i> strain WC007 chromosome, complete genome
6e70427e527e130c9f96625215b8 5a8af06c3cec	CP04089 6.1	<i>Hymenobacter jejuensis</i> strain 17J68-5 chromosome, complete genome
6f51198c3a2e29d696b37495908b de6c6012512f	CP04306 0.1	<i>Pseudomonas</i> sp. J380 chromosome, complete genome
6f51198c3a2e29d696b37495908b de6c6012512f	CP02326 9.1	<i>Pseudomonas</i> sp. MYb193 chromosome, complete genome
6f51198c3a2e29d696b37495908b de6c6012512f	CP07708 3.1	<i>Pseudomonas salmasensis</i> strain SWRI126 chromosome, complete genome
7001b925c3adaeabbd1f1e71e8e7 75678ea14ac	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome

7001b925c3adaedbbd1f1e71e8e775678ea14ac	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
700c722f463f69164051d454505fd479e06f0c4e	NC_0557 33.1	<i>Aeromonas</i> phage Ah1, complete genome
700c722f463f69164051d454505fd479e06f0c4e	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
70f7647427a203b0af517c0b36b679c2fa6e8185	AP01733 9.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F28
70f7647427a203b0af517c0b36b679c2fa6e8185	AP01733 5.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F211
70f7647427a203b0af517c0b36b679c2fa6e8185	AP01732 9.1	<i>Helicobacter pylori</i> DNA, nearly complete genome, strain: F13
70f7647427a203b0af517c0b36b679c2fa6e8185	CP09417 7.1	<i>Helicobacter pylori</i> strain Hpfe062 chromosome, complete genome
70f7647427a203b0af517c0b36b679c2fa6e8185	KF05212 2.1	<i>Campylobacter fetus</i> strain ML7 16S ribosomal RNA gene, partial sequence; 16S-23S ribosomal RNA intergenic spacer, complete sequence; and 23S ribosomal RNA gene, partial sequence
715b564a86ee8e25c94db609ab610c192ecdcb7c	CP05491 9.1	<i>Kitasatospora</i> sp. NA04385 chromosome, complete genome
73a4eb2fbd087e6f46daea29e3c51155fc4e30d3	CP09173 1.1	MAG: <i>Candidatus Melainabacteria</i> bacterium isolate nC33_lr_bin.1.fa chromosome
73b3906ef6a50bf01fb456bc86fe3a1bc0f74378	CP06944 0.1	<i>Phocaeicola coprophilus</i> strain FDAARGOS_1220 chromosome, complete genome
741cc8b70d6bc1a345f3b4f89bcc54193580bbfe	CP04765 2.1	<i>Bombella</i> sp. KACC 21507 chromosome, complete genome

741cc8b70d6bc1a345f3b4f89bcc54193580bbfe	CP09273 4.1	Acetobacteraceae bacterium DM15PD chromosome cDM15PD, complete sequence
744cf270f64216fa2fa694958dd2fdac70735452	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
745fcc799306452ff8c50404c2f93f11729ff784	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
745fcc799306452ff8c50404c2f93f11729ff784	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
745fcc799306452ff8c50404c2f93f11729ff784	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
768a39353ae9dae622e5a88c47cfd9f4c1416404	FP92914 0.1	gamma proteobacterium HdN1 complete genome
7769d51ad7030487e1d6cac5cbdb2c955747fcad	LR77817 5.1	Chromatiaceae bacterium RW isolate NSCAC1 genome assembly, chromosome: 1
79d6d0a080be82abc801c15b5356fbd5508e3469	CP02722 8.1	<i>Mogibacterium diversum</i> strain CCUG 47132 chromosome, complete genome
79d6d0a080be82abc801c15b5356fbd5508e3469	CP09709 3.1	MAG: <i>Mogibacterium diversum</i> isolate JCVI-JB-Md32 chromosome, complete genome
7ae48566beaf592f0b3e8558c74ac227b0b3daec	KF05212 2.1	<i>Campylobacter fetus</i> strain ML7 16S ribosomal RNA gene, partial sequence; 16S-23S ribosomal RNA intergenic spacer, complete sequence; and 23S ribosomal RNA gene, partial sequence
7c8c164d5d2857af7b51b7313fc683e3ef66af56	CP08278 0.1	<i>Bacillus aquiflavi</i> strain 3H-10 chromosome, complete genome
7c8ec16387eea793e75cca6ad450479721f3d740	LR13452 5.1	<i>Neisseria meningitidis</i> strain NCTC10025 genome assembly, chromosome: 1

7c8ec16387eea793e75cca6ad4504 79721f3d740	CP02042 0.2	<i>Neisseria meningitidis</i> strain FDAARGOS_209 chromosome, complete genome
7c8ec16387eea793e75cca6ad4504 79721f3d740	CP06436 7.1	<i>Neisseria meningitidis</i> strain PartJ-Nmeningitidis- RM8376 chromosome, complete genome
7c8ec16387eea793e75cca6ad4504 79721f3d740	CP06565 3.1	<i>Neisseria meningitidis</i> strain FDAARGOS_914 chromosome, complete genome
7cbe88f2a85f4099d26fb6b8055a2 cb659afac5b	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
7cbe88f2a85f4099d26fb6b8055a2 cb659afac5b	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
7d8e1f36b184c27f8d83a1e70de66 9cafac7e55e	CP04090 2.1	<i>Longibaculum</i> sp. KGMB06250 chromosome, complete genome
7d8e1f36b184c27f8d83a1e70de66 9cafac7e55e	CP03183 7.1	<i>Clostridium sporogenes</i> strain YH-Raf3-2-5 chromosome, complete genome
7d8e1f36b184c27f8d83a1e70de66 9cafac7e55e	AP01853 7.1	Erysipelotrichaceae bacterium GAM147 DNA, complete genome
7d8e1f36b184c27f8d83a1e70de66 9cafac7e55e	AP02408 5.1	<i>Faecalibacillus intestinalis</i> 14EGH31 DNA, complete genome
7ddc7c329ae515ed89c070a8703af d284d6ef46b	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
7ddc7c329ae515ed89c070a8703af d284d6ef46b	CP03939 6.1	<i>Duncaniella dubosii</i> strain H5 chromosome
7ddc7c329ae515ed89c070a8703af d284d6ef46b	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
7deb487a5ccf27d82c3801c6d0eb5 5198e7a6465	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome

7ebbbfe95b6212dd61855b189a43a94697741b3e	KF05212 2.1	<i>Campylobacter fetus</i> strain ML7 16S ribosomal RNA gene, partial sequence; 16S-23S ribosomal RNA intergenic spacer, complete sequence; and 23S ribosomal RNA gene, partial sequence
7ec5eda5ac1a59e2be635e6795e27c6c76c7bdd3	OW3704 93.1	<i>Rickettsia endosymbiont</i> of <i>Oedothorax gibbosus</i> strain Oegibbosus-W744x776 genome assembly, chromosome: 1
7ef81c0ffbced4e7e047ccbf113f1e271e2c0dac	CP04009 3.1	<i>Caloramator</i> sp. E03 chromosome, complete genome
7f32ee7c073256f8b318a16af63c5676f207e319	OV78864 0.1	uncultured <i>Tenericutes</i> bacterium isolate min17_bin11 genome assembly, chromosome: 1
8014d14a00feac023737b9e33b0869dc75627650	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
8014d14a00feac023737b9e33b0869dc75627650	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
8014d14a00feac023737b9e33b0869dc75627650	OU94086 2.1	uncultured fungus genomic DNA sequence contains 18S rRNA gene, ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gene, isolate consensus02156
803b1ed6906413e23c5e26147b4203edf994042e	CP00239 0.1	<i>Filifactor alocis</i> ATCC 35896, complete genome
810b53d8e89434dfabf80e6688bf9a63395fe68a	LR99447 3.1	uncultured <i>Prevotellamassilia</i> sp. isolate <i>Prevotellamassilia canine</i> HQ MAG genome assembly, chromosome: 1
81642f2159f1b6d06cc834bed55c868aa68767d3	MH6229 43.1	MAG: <i>Myoviridae</i> sp. isolate ctbc_4, complete genome
8267473fe62f04bdffc9e238e402d41fa3b702cf	CP06063 6.1	<i>Eubacterium hominis</i> strain NSJ-61 chromosome, complete genome

82a20989c35c622b21bb903f1967bda70ba868ba	AP01737 2.2	<i>Halorhodospira halochloris</i> DNA, nearly complete genome, strain: DSM 1059
8308250b370541f43f30f694270c9143cf1bfb6a	OW2037 21.1	<i>Yponomeuta sedellus</i> genome assembly, chromosome: Z
83684404354a4cd3ad4d3585575c848b2e559446	CP09173 5.1	MAG: <i>Clostridium</i> sp. isolate nC29_bin.1.fa chromosome, complete genome
84412616d5a6d0895a3e5d722ea9258eabf2c8e2	CP00093 0.2	<i>Heliomicrobium modesticaldum</i> Ice1 chromosome, complete genome
852f16581404bf70e97c950ddc1d57ece17be5d3	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
888a8f10e9da6416e7e2f29af183a2a5e8dda361	CP03262 7.1	<i>Lactococcus allomyrinae</i> strain 1JSPR-7 chromosome, complete genome
89ab1b7cc67f6856454c963d9ffbb0474ac0d070	MW8221 45.1	<i>Streptomyces</i> phage TunaTartare, complete genome
89b0be9f17b0300b5dce38c22dcce8e549be6de1	AP02266 0.1	<i>Bacteroides thetaiotaomicron</i> F9-2 DNA, nearly complete genome
89b0be9f17b0300b5dce38c22dcce8e549be6de1	CP04569 6.1	<i>Sodaliophilus pleomorphus</i> strain Oil-RF-744-WCA-WT-10 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce8e549be6de1	CP04361 0.1	<i>Bacteroides fragilis</i> strain GUT04 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce8e549be6de1	CP04137 9.1	<i>Bacteroides intestinalis</i> strain APC919/174 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce8e549be6de1	CP04123 0.1	<i>Bacteroides xylanisolvens</i> strain H207 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce8e549be6de1	CP03655 5.1	<i>Bacteroides fragilis</i> strain CCUG4856T chromosome, complete genome

89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP03744 0.1	<i>Bacteroides fragilis</i> strain DCMOUH0085B chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP03655 3.1	<i>Bacteroides fragilis</i> strain DCMOUH0067B chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP03655 0.1	<i>Bacteroides fragilis</i> strain DCMOUH0042B chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP03654 6.1	<i>Bacteroides fragilis</i> strain DCMSKEJBY0001B chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP03653 9.1	<i>Bacteroides fragilis</i> strain DCMOUH0017B chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	LR21598 0.1	<i>Paraprevotella xylaniphila</i> YIT 11841 isolate <i>Paraprevotella xylaniphila</i> 82A6 genome assembly, chromosome: 1
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP09173 2.1	MAG: Porphyromonadaceae bacterium isolate nC29_bin.32.fa chromosome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP09173 2.1	MAG: Porphyromonadaceae bacterium isolate nC29_bin.32.fa chromosome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08029 6.1	<i>Bacteroides fragilis</i> strain IHMA_13 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08029 5.1	<i>Bacteroides fragilis</i> strain IHMA_4 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08594 2.1	<i>Prevotella melaninogenica</i> strain FDAARGOS_1566 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08592 6.1	<i>Prevotella buccalis</i> strain FDAARGOS_1588 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP06494 0.1	<i>Bacteroides intestinalis</i> strain APC919/174 substr. 8W chromosome

89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP06494 1.1	<i>Bacteroides intestinalis</i> strain APC919/174 substr. 8T chromosome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08192 2.1	<i>Bacteroides fragilis</i> strain BFG-1 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08190 2.1	<i>Bacteroides salyersiae</i> strain BFG-256 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08189 9.1	<i>Bacteroides salyersiae</i> strain BFG-288 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP08189 8.1	<i>Bacteroides thetaiotaomicron</i> strain BFG-55 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP05420 4.1	<i>Bacteroides uniformis</i> strain JCM13288 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07225 7.1	<i>Bacteroides fragilis</i> CL03T12C07 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07224 2.1	<i>Bacteroides thetaiotaomicron</i> strain CL06T03C18 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07223 6.1	<i>Bacteroides uniformis</i> strain CL06T06C18 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07222 4.1	<i>Bacteroides thetaiotaomicron</i> strain CL11T00C24 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07221 6.1	<i>Bacteroides xylanisolvens</i> strain CL11T00C03 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07221 2.1	<i>Bacteroides xylanisolvens</i> strain CL11T00C41 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07237 4.1	<i>Prevotella denticola</i> strain F0115 chromosome 1, complete sequence

89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07237 3.1	<i>Prevotella denticola</i> strain F0119 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07237 0.1	<i>Prevotella fusca</i> JCM 17724 strain W1435 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07236 7.1	<i>Prevotella histicola</i> strain F0411 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07236 5.1	<i>Prevotella jejuni</i> strain F0106 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07236 3.1	<i>Prevotella jejuni</i> strain F0697 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07236 2.1	<i>Prevotella melaninogenica</i> strain F0054 chromosome 2, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07235 9.1	<i>Prevotella melaninogenica</i> strain F0091 chromosome 2, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07234 9.1	<i>Prevotella melaninogenica</i> strain F0301 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07233 0.1	<i>Prevotella veroralis</i> strain F0319 chromosome 1, complete sequence
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP07219 8.1	<i>Bacteroides</i> sp. ZJ-18 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	AP02448 4.1	<i>Prevotella herbatica</i> WR041 DNA, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP04442 8.1	<i>Bacteroides fragilis</i> strain DK1985 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP06817 4.1	<i>Bacteroides caccae</i> strain FDAARGOS_1097 chromosome, complete genome

89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP06588 9.1	<i>Bacteroides uniformis</i> strain FDAARGOS_901 chromosome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP04246 4.1	<i>Prevotella copri</i> strain YF2 chromosome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP04246 4.1	<i>Prevotella copri</i> strain YF2 chromosome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP04063 0.1	<i>Bacteroides</i> sp. PHL 2737 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP05997 3.1	<i>Bacteroides luhongzhouii</i> strain HF-5141 chromosome, complete genome
89b0be9f17b0300b5dce38c22dcce 8e549be6de1	CP05985 6.1	<i>Bacteroides zhangwenhongii</i> strain HF-5287 chromosome, complete genome
8a6ee781157a1316142180ee57ad 98a14a9d5a3a	CP10131 5.1	<i>Bacillus amyloliquefaciens</i> strain KNU-28 chromosome, complete genome
8b257e203b460779b978d24b6ac4 f55f46025dc3	CP09172 6.1	MAG: <i>Bacilli</i> bacterium isolate nC33_bin.244.fa chromosome
8bab497c11c3f5403c166d7bafa37 349130232b0	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
8bab497c11c3f5403c166d7bafa37 349130232b0	CP03939 6.1	<i>Duncaniella dubosii</i> strain H5 chromosome
8bab497c11c3f5403c166d7bafa37 349130232b0	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
8c433c5d3170b39f51837543fbe1c 6ddf08cf2f5	CP00245 2.1	<i>Nitratifactor salsuginis</i> DSM 16511, complete genome
8c79eb907414bff5fccbd29cee1e5 75b4ff15a1b	CP08528 4.1	<i>Acinetobacter haemolyticus</i> strain w12 chromosome, complete genome

8c79eb907414bff5fccbd29cee1e5 75b4ff15a1b	CP06701 9.1	<i>Acinetobacter</i> sp. CS-2 chromosome, complete genome
8c79eb907414bff5fccbd29cee1e5 75b4ff15a1b	CP01280 8.1	<i>Acinetobacter equi</i> strain 114 chromosome, complete genome
8c79eb907414bff5fccbd29cee1e5 75b4ff15a1b	CP01260 8.1	<i>Acinetobacter</i> sp. TTH0-4, complete genome
8c79eb907414bff5fccbd29cee1e5 75b4ff15a1b	CP05907 9.1	<i>Acinetobacter</i> sp. TTH0-4 strain 1BD1 chromosome, complete genome
8c9e68c012e369e74109f7bf9fc dc21bc23a0fb	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
8c9e68c012e369e74109f7bf9fc dc21bc23a0fb	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
8c9e68c012e369e74109f7bf9fc dc21bc23a0fb	OU94086 2.1	uncultured fungus genomic DNA sequence contains 18S rRNA gene, ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gene, isolate consensus02156
8d11f1c30e47296b2d02c6e96b4d 26c74b1d3e56	LR21598 0.1	<i>Paraprevotella xylaniphila</i> YIT 11841 isolate <i>Paraprevotella xylaniphila</i> 82A6 genome assembly, chromosome: 1
8da36826938ecb58830d2e7bb1b4 740101ad01b2	AP01737 2.2	<i>Halorhodospira halochloris</i> DNA, nearly complete genome, strain: DSM 1059
8e67e1a079c179721fa364349724 7887551dd715	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
8f3c6e2d16843712feef7c4b9a652 5c55312f396	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
8fc7cb34ebdbeee5c703998d8f582 3cc459836cc	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome

901ec1d25078dd1fba359fb44ad3a5173f9935f	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
903c5925b6fe68fc8e3ea952548fc8cfbc4fc57b	CP04883 8.1	[<i>Clostridium</i>] <i>innocuum</i> strain ATCC 14501 chromosome, complete genome
903c5925b6fe68fc8e3ea952548fc8cfbc4fc57b	CP02272 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 genome
903c5925b6fe68fc8e3ea952548fc8cfbc4fc57b	CP01540 4.2	Erysipelotrichaceae bacterium I46 chromosome, complete genome
903c5925b6fe68fc8e3ea952548fc8cfbc4fc57b	CP09494 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain VE303-07 chromosome, complete genome
903c5925b6fe68fc8e3ea952548fc8cfbc4fc57b	BK03754 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctS5j6, partial genome
903c5925b6fe68fc8e3ea952548fc8cfbc4fc57b	CP06532 0.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 chromosome
904e9df4b31061b193904af0e59ec913451b0831	BK02788 7.1	MAG TPA_asm: <i>Myoviridae</i> sp. isolate ctpaF17, partial genome
904e9df4b31061b193904af0e59ec913451b0831	BK03999 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct9vp2, partial genome
904e9df4b31061b193904af0e59ec913451b0831	BK03002 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct5Qt3, partial genome
913e24dcbfd3697db9c0050726cef5eb86612968	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
9195b3b013342ff9858630da4588e7989262b01d	CP09172 6.1	MAG: <i>Bacilli</i> bacterium isolate nC33_bin.244.fa chromosome
91cbd137da14933c9a91296371e53f91bc62dcec	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome

925e301aad284c01367d8fa011ba b09ac653844c	CP04604 8.1	<i>Proteus mirabilis</i> strain HN2p chromosome HN2p, complete sequence
925e301aad284c01367d8fa011ba b09ac653844c	CP05361 4.1	<i>Proteus mirabilis</i> strain S74-1(++)-2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05371 8.1	<i>Proteus mirabilis</i> strain MPE4069 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05368 4.1	<i>Proteus mirabilis</i> strain MPE5139 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05368 5.1	<i>Proteus mirabilis</i> strain MPE5203 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05368 1.1	<i>Proteus mirabilis</i> strain M3-1-17 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05368 2.1	<i>Proteus mirabilis</i> strain MPE0156 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05368 3.1	<i>Proteus mirabilis</i> strain MPE0027 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05371 9.1	<i>Proteus mirabilis</i> strain MPE0346 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05361 5.1	<i>Proteus mirabilis</i> strain MPE0734 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05361 6.1	<i>Proteus mirabilis</i> strain MPE0767 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04525 7.1	<i>Proteus mirabilis</i> strain L90-1 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05389 8.1	<i>Proteus mirabilis</i> strain YPM35 chromosome, complete genome

925e301aad284c01367d8fa011ba b09ac653844c	CP05389 4.1	<i>Proteus mirabilis</i> strain JPM24 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05126 0.1	<i>Proteus mirabilis</i> strain STP3 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04994 2.1	<i>Proteus mirabilis</i> strain XH1569 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04994 1.1	<i>Proteus mirabilis</i> strain XH1568 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04975 3.1	<i>Proteus mirabilis</i> strain PmBR607 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04735 2.1	<i>Proteus mirabilis</i> strain ZA25 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04285 7.1	<i>Proteus mirabilis</i> strain 1701092 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04878 7.1	<i>Proteus mirabilis</i> strain CC15031 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04711 2.1	<i>Proteus mirabilis</i> strain SCBX1.1 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04553 8.2	<i>Proteus mirabilis</i> strain CRE14IB chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04413 4.1	<i>Proteus mirabilis</i> strain ENT1224 chromosome
925e301aad284c01367d8fa011ba b09ac653844c	CP04413 6.1	<i>Proteus mirabilis</i> strain ENT1157 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04413 5.1	<i>Proteus mirabilis</i> strain ENT1301 chromosome

925e301aad284c01367d8fa011ba b09ac653844c	CP04402 8.1	<i>Proteus mirabilis</i> strain K817 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04387 0.1	<i>Proteus mirabilis</i> strain S1959 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04333 2.1	<i>Proteus mirabilis</i> strain CRPM10 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04290 7.1	<i>Proteus mirabilis</i> strain VAC chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04139 1.1	<i>Proteus mirabilis</i> strain CLPM181223 chromosome
925e301aad284c01367d8fa011ba b09ac653844c	CP03409 0.1	<i>Proteus mirabilis</i> strain PmSC1111 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP03409 1.1	<i>Proteus mirabilis</i> strain PmBC1123 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	LR13420 5.1	<i>Proteus mirabilis</i> strain NCTC4199 genome assembly, chromosome: 1
925e301aad284c01367d8fa011ba b09ac653844c	CP02972 5.1	<i>Proteus mirabilis</i> strain AR_0029 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP02327 3.1	<i>Proteus mirabilis</i> strain CCUG 70746 chromosome
925e301aad284c01367d8fa011ba b09ac653844c	CP02913 3.1	<i>Proteus mirabilis</i> strain AR379 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP02835 6.1	<i>Proteus mirabilis</i> strain K670 chromosome
925e301aad284c01367d8fa011ba b09ac653844c	CP02605 1.1	<i>Proteus mirabilis</i> strain FDAARGOS_67 chromosome, complete genome

925e301aad284c01367d8fa011ba b09ac653844c	CP02658 1.1	<i>Proteus mirabilis</i> isolate GN2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP02657 1.1	<i>Proteus mirabilis</i> strain BC11-24 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP02605 9.1	<i>Proteus mirabilis</i> strain FDAARGOS_80 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP02604 4.1	<i>Proteus mirabilis</i> strain FDAARGOS_60 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP01708 5.1	<i>Proteus mirabilis</i> strain T18 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07178 0.1	<i>Proteus mirabilis</i> strain swupm3 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07177 7.1	<i>Proteus mirabilis</i> strain swupm2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07177 3.1	<i>Proteus mirabilis</i> strain swupm1 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09845 0.1	<i>Proteus mirabilis</i> strain FZP3115 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09844 7.1	<i>Proteus mirabilis</i> strain FZP2936 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09844 4.1	<i>Proteus mirabilis</i> strain FZP3105 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09844 6.1	<i>Proteus mirabilis</i> strain FZP2826 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09677 6.1	<i>Proteus mirabilis</i> strain HURS-186083 chromosome, complete genome

925e301aad284c01367d8fa011ba b09ac653844c	CP09677 5.1	<i>Proteus mirabilis</i> strain HURS-181823 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09576 5.1	<i>Proteus mirabilis</i> strain T1010 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP09265 2.1	<i>Proteus mirabilis</i> strain PM8762 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05509 5.1	<i>Proteus mirabilis</i> strain SWHIN_109 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05500 9.1	<i>Proteus mirabilis</i> strain STIN_74 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	OL79081 2.1	<i>Proteus mirabilis</i> strain H110 plasmid unnamed, complete sequence
925e301aad284c01367d8fa011ba b09ac653844c	OL79081 1.1	<i>Proteus mirabilis</i> strain H63 plasmid unnamed, complete sequence
925e301aad284c01367d8fa011ba b09ac653844c	CP08931 7.1	<i>Proteus mirabilis</i> strain PM1162 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP08637 7.1	<i>Proteus mirabilis</i> strain LB UEL H-11 chromosome
925e301aad284c01367d8fa011ba b09ac653844c	CP04869 2.1	<i>Proteus mirabilis</i> strain HNS2p chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06503 9.2	<i>Proteus mirabilis</i> strain XH1653 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07796 3.1	<i>Proteus mirabilis</i> strain 6Pmi283 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04758 9.1	<i>Proteus mirabilis</i> strain SNYG35 chromosome, complete genome

925e301aad284c01367d8fa011ba b09ac653844c	CP07324 5.1	<i>Proteus mirabilis</i> strain S74-3-2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07324 7.1	<i>Proteus mirabilis</i> strain S62-3-2-2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07324 8.1	<i>Proteus mirabilis</i> strain N292 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07324 6.1	<i>Proteus mirabilis</i> strain N639-2X chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07277 9.1	<i>Proteus mirabilis</i> strain 1035 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06683 3.1	<i>Proteus mirabilis</i> strain RGF134-1 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04604 9.1	<i>Proteus mirabilis</i> strain DY.F1.2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07057 2.1	<i>Proteus mirabilis</i> strain PM52808 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP07056 9.1	<i>Proteus mirabilis</i> strain PM52260 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04443 7.1	<i>Proteus mirabilis</i> strain C74 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04443 6.1	<i>Proteus mirabilis</i> strain C55 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06815 2.1	<i>Proteus mirabilis</i> strain FDAARGOS_1079 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06514 6.1	<i>Proteus mirabilis</i> strain PmBJ020-1 chromosome, complete genome

925e301aad284c01367d8fa011ba b09ac653844c	CP06514 5.1	<i>Proteus mirabilis</i> strain PmBJ023-2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06514 4.1	<i>Proteus mirabilis</i> strain PmBJ024-1 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06514 8.1	<i>Proteus mirabilis</i> strain PmBJ012-2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06514 7.1	<i>Proteus mirabilis</i> strain PmBJ015-2 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06344 0.1	<i>Proteus mirabilis</i> strain Yak 2019 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP04792 9.1	<i>Proteus mirabilis</i> strain ChSC1905 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP06214 6.1	<i>Proteus mirabilis</i> strain S012 chromosome, complete genome
925e301aad284c01367d8fa011ba b09ac653844c	CP05905 6.1	<i>Proteus mirabilis</i> strain CriePir89 chromosome, complete genome
92a7fbd3b5c14508a6ba1bf63e0a4 41808557123	BK01913 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctF907, partial genome
930b0de8d4cd082ea3a01d389cb4 54e25a19de50	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
930b0de8d4cd082ea3a01d389cb4 54e25a19de50	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
9379d32ac85f0e4fe2fd87fe3db4c 45fa0e2ff90	AP02441 2.1	<i>Mariniplasma anaerobium</i> DNA, complete genome
939efd3ed62f0790711460b9f2ca9 2c920919ab4	CP03938 1.1	<i>Ruminococcus bovis</i> strain JE7A12 chromosome, complete genome

93db6d86cc62fd8e7635aac62398 9b755c5f0114	CP09173 5.1	MAG: <i>Clostridium</i> sp. isolate nC29_bin.1.fa chromosome, complete genome
93ee78d56544ed8dc42e191395df 3c52abebebd9	CP01422 3.1	<i>Anaerotignum propionicum</i> DSM 1682 strain X2 chromosome, complete genome
941162f174c9f3302eae7101e01c3 61e21ed6b93	CP07049 1.1	<i>Methylococcus</i> sp. EFPC2 chromosome, complete genome
94a780e5748734d38a58ac302612 d06f084defcf	BK04987 3.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct7jy12, partial genome
95e7faa4757cbf1cb89aa7ced29c2 9603144fe34	CP04883 7.1	[<i>Clostridium</i>] <i>innocuum</i> strain LC-LUMC-CI-001 chromosome, complete genome
95e7faa4757cbf1cb89aa7ced29c2 9603144fe34	AP02556 5.1	[<i>Clostridium</i>] <i>innocuum</i> CE91-St51 DNA, complete genome
964129f3d3282adac9e537a25b76 9a668c6928e5	CP05937 8.1	<i>Clostridium intestinale</i> strain Lx1 chromosome, complete genome
9699302ec209ec14f8ae096b126e5 3d05c4144cb	CP00269 1.1	<i>Haliscomenobacter hydrossis</i> DSM 1100 chromosome, complete genome
9699302ec209ec14f8ae096b126e5 3d05c4144cb	GU26070 3.1	Uncultured bacterium 246 genomic sequence
96e535a7a2efb20404effe0cf95086 8167219496	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
97b6198a366b8b5bfd3b6c2926a2 988a231aa661	CP04986 0.1	<i>Erysipelothrix</i> sp. HDW6B chromosome, complete genome
9843aa024f5a5731108455116fddf 96a822d4ceb	CP00273 9.1	<i>Thermoanaerobacterium xylanolyticum</i> LX-11, complete genome
98d7d07d6757771b274ecbd7e902 a5a0ffda5a24	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome

98ec17506264443da32bc9bb95864b217c570bdf	CP09171 8.1	MAG: Oscillospiraceae bacterium isolate nC33_bin.8.fa chromosome
98ff6578011629170b72307fddc2e0ce5973fbea	KF15633 8.1	<i>Synechococcus</i> phage S-MbCM7, complete genome
996b43d7da4235887ffafe6ccc4b9527aa49da2	AP01930 9.1	Erysipelotrichaceae bacterium SG0102 DNA, complete genome
996b43d7da4235887ffafe6ccc4b9527aa49da2	LR99447 2.1	<i>Catenibacterium</i> sp. co_0103 isolate <i>Catenibacterium canine</i> HQ MAG genome assembly, chromosome: 1
99fc0482b5a3e3eff11ec5b827b3f6fffedc43cc	CP02334 4.1	<i>Nibricoccus aquaticus</i> strain HZ-65 chromosome, complete genome
99fc0482b5a3e3eff11ec5b827b3f6fffedc43cc	LK01198 2.1	TPA: <i>Opitutus terrae</i> PB90-1 tRNA-Leu-GAG-1-1 gene
99fc0482b5a3e3eff11ec5b827b3f6fffedc43cc	CP00103 2.1	<i>Opitutus terrae</i> PB90-1, complete genome
9a18f809be2272f94ffc936b5650cf5ab083bfbb	CP06063 6.1	<i>Eubacterium hominis</i> strain NSJ-61 chromosome, complete genome
9a8b80fd639c6dcd3c8c9c15f40cb00148934ca3	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
9adc5525b188ead4400a8e72629128346c443673	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
9e354232bdb13540b44b9708055996b1c8029401	AP01971 1.1	<i>Amedibacterium intestinale</i> DNA, complete genome
9e354232bdb13540b44b9708055996b1c8029401	AP01971 1.1	<i>Amedibacterium intestinale</i> DNA, complete genome
9e354232bdb13540b44b9708055996b1c8029401	AP01969 5.1	<i>Amedibacterium intestinale</i> JCM 30884 DNA, complete genome

9e354232bdb13540b44b97080559 96b1c8029401	AP01969 5.1	<i>Amedibacterium intestinale</i> JCM 30884 DNA, complete genome
9e448d134e6b5addc1b45c9ee601 7f2cd2a08d58	AP01930 9.1	Erysipelotrichaceae bacterium SG0102 DNA, complete genome
9e64079a8c2f334dcd425ff28d86f c4e64c47fa2	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
9f2a7818737d066476dc318842b5 b4ef55430dbe	CP01422 3.1	<i>Anaerotignum propionicum</i> DSM 1682 strain X2 chromosome, complete genome
9f5addeb8c9a25277e7191b6fe661 2d9b65c2fbb	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	LK99601 7.1	<i>Desulfitobacterium hafniense</i> strain PCE-S genome assembly, scaffold: scaffold0
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	LR74649 6.1	Peptococcaceae bacterium INE isolate PacBioINE genome assembly, chromosome: 1
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
9fc69aea35b1ebdbf62c1fdd489c1 d54e2643350	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome

9fc69aea35b1ebdbf62c1fdd489c1d54e2643350	AP00823 0.1	<i>Desulfitobacterium hafniense</i> Y51 DNA, complete genome
a05b30a1c7197f5ed2bb0cf0623d91158bc53de4	CP00245 2.1	<i>Nitratifractor salsuginis</i> DSM 16511, complete genome
a107bbf993cc141ae59995a0b16545706553be83	CP05468 9.1	Ectothiorhodospiraceae bacterium 2226 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP05018 3.1	<i>Bacillus thuringiensis</i> strain HER1410 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP05395 4.1	<i>Bacillus cereus</i> strain FDAARGOS_798 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP05395 5.1	<i>Bacillus tropicus</i> strain FDAARGOS_782 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP05395 1.1	<i>Bacillus cereus</i> strain FDAARGOS_799 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP05328 9.1	<i>Bacillus cereus</i> strain WPySW2 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP04997 8.1	<i>Bacillus</i> sp. RZ2MS9 chromosome, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP04901 9.1	<i>Bacillus tropicus</i> strain AOA-CPS1 chromosome
a143c0fe2330c7bd7941218de988192ddd9754f1	AP02264 3.1	<i>Bacillus wiedmannii</i> PL1 DNA, complete genome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP02800 9.1	<i>Bacillus cereus</i> strain SGAir0260 chromosome
a143c0fe2330c7bd7941218de988192ddd9754f1	CP02792 0.1	<i>Bacillus cereus</i> strain SGAir0263 chromosome, complete genome

a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04639 8.1	<i>Bacillus</i> sp. A260 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04034 4.1	<i>Bacillus albus</i> strain DLOU-Yingkou chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04034 2.1	<i>Bacillus cereus</i> strain DLOU-Weihai chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04034 0.1	<i>Bacillus cereus</i> strain DLOU-Tangshan chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04033 6.1	<i>Bacillus luti</i> strain FJ chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04227 0.1	<i>Bacillus thuringiensis</i> strain GA-A07 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04497 8.1	<i>Bacillus thuringiensis</i> strain BT62 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP03106 8.1	<i>Bacillus</i> sp. JAS24-2 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP03106 5.1	<i>Bacillus</i> sp. SH8-8 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP03106 2.1	<i>Bacillus</i> sp. AR4-2 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP03106 2.1	<i>Bacillus</i> sp. AR4-2 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04292 9.1	<i>Bacillus cereus</i> strain G1-1 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04287 4.1	<i>Bacillus cereus</i> strain 09 chromosome, complete genome

a143c0fe2330c7bd7941218de988 192ddd9754f1	CP03468 6.1	<i>Bacillus</i> sp. BD59S chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04107 1.1	<i>Bacillus tropicus</i> strain LM1212-W3 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP02945 4.1	<i>Bacillus cereus</i> strain FORC087 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09197 1.1	<i>Bacillus cereus</i> strain MB1 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP07269 1.1	<i>Bacillus thuringiensis</i> strain IMBL-B9 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08540 9.1	<i>Bacillus thuringiensis</i> strain KF1 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08540 6.1	<i>Bacillus bombysepticus</i> strain F12 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08539 9.1	<i>Bacillus tropicus</i> strain CK18 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09873 4.1	<i>Bacillus cereus</i> strain MV19 chromosome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09735 1.1	<i>Bacillus cereus</i> strain DQ01 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09587 4.1	<i>Bacillus</i> sp. PGP15 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09587 4.1	<i>Bacillus</i> sp. PGP15 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09537 7.1	<i>Bacillus cereus</i> strain SEM-15 chromosome, complete genome

a143c0fe2330c7bd7941218de988 192ddd9754f1	CP07808 1.1	<i>Bacillus tropicus</i> strain EMB20 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09462 4.1	<i>Bacillus thuringiensis</i> strain LX43 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09462 4.1	<i>Bacillus thuringiensis</i> strain LX43 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09440 7.1	<i>Bacillus thuringiensis</i> strain UFT038 chromosome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06550 9.1	<i>Bacillus</i> sp. N447-1 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06397 5.1	<i>Bacillus cereus</i> strain N435-1 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09392 8.1	<i>Bacillus cereus</i> strain AH7-7 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04558 5.1	<i>Bacillus thuringiensis</i> strain x023 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	AP02450 4.2	<i>Bacillus cereus</i> HT18 DNA, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP07471 4.1	<i>Bacillus thuringiensis</i> strain B13 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09042 1.1	<i>Bacillus cereus</i> strain NR1 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP09008 1.1	<i>Bacillus cereus</i> strain CPT56D-587-MTF chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	AP02298 1.1	<i>Bacillus cereus</i> 30048 DNA, complete genome

a143c0fe2330c7bd7941218de988 192ddd9754f1	AP02294 6.1	<i>Bacillus cereus</i> J10 DNA, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08315 6.1	<i>Bacillus thuringiensis</i> strain ABTS-1857 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08314 4.1	<i>Bacillus thuringiensis</i> strain GC-91 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08311 6.1	<i>Bacillus thuringiensis</i> strain SA11 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08310 1.1	<i>Bacillus thuringiensis</i> strain ABTS-351 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08308 5.1	<i>Bacillus thuringiensis</i> strain B401 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08549 8.1	<i>Bacillus cereus</i> strain A22 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08550 6.1	<i>Bacillus cereus</i> strain HD2.4 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08551 0.1	<i>Bacillus cereus</i> strain HD1.4B chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP08374 9.1	<i>Bacillus</i> sp. CRB-7 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP07277 4.1	<i>Bacillus cereus</i> strain BC33 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP05101 3.1	<i>Bacillus tropicus</i> strain LM1212-DB chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04323 4.1	<i>Bacillus thuringiensis</i> strain Bt-GS57 chromosome, complete genome

a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06871 9.1	<i>Bacillus cereus</i> strain 21155 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06871 9.1	<i>Bacillus cereus</i> strain 21155 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06872 4.1	<i>Bacillus cereus</i> strain AFA01 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06813 5.1	<i>Bacillus cereus</i> strain FDAARGOS_1084 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06737 5.1	<i>Bacillus</i> sp. TK-2 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06574 3.1	<i>Bacillus tropicus</i> strain FDAARGOS_897 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP06573 9.1	<i>Bacillus tropicus</i> strain FDAARGOS_920 chromosome, complete genome
a143c0fe2330c7bd7941218de988 192ddd9754f1	CP04396 6.1	<i>Bacillus cereus</i> strain H2 chromosome, complete genome
a176f3fdcc995a338270bfa8e5291 6bb636cc245	CP09734 9.1	<i>Cytobacillus oceanisediminis</i> strain BNO1 chromosome, complete genome
a1912453ffe14394a08c9a581562a a31fe074107	MH6229 43.1	MAG: <i>Myoviridae</i> sp. isolate ctbc_4, complete genome
a223683538555b6bd6f56c1bf825 1abcadfb396	CP09734 9.1	<i>Cytobacillus oceanisediminis</i> strain BNO1 chromosome, complete genome
a223683538555b6bd6f56c1bf825 1abcadfb396	CP02614 3.1	<i>Bacillus methanolicus</i> strain DFS2 chromosome, complete genome
a223683538555b6bd6f56c1bf825 1abcadfb396	CP02614 3.1	<i>Bacillus methanolicus</i> strain DFS2 chromosome, complete genome

a223683538555b6bd6f56c1bf825 1abcbdfbb396	CP00773 9.1	<i>Bacillus methanolicus</i> MGA3, complete genome
a223683538555b6bd6f56c1bf825 1abcbdfbb396	CP00773 9.1	<i>Bacillus methanolicus</i> MGA3, complete genome
a2462c9897f2600fca277ab8948c4 f13f66f9675	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
a381f4b3f37fc53e5c3e5701cb28a cb28930759d	LR69898 1.1	<i>Clostridiales</i> bacterium isolate MGYG-HGUT-02340 genome assembly, chromosome: 1
a3e10b631cff0303d4ca9919052a0 b4b4f9f2360	CP03510 7.1	<i>Ornithobacterium rhinotracheale</i> strain FARPER- 174b chromosome, complete genome
a548144df872ea64d0eecd3ceae23 c1f463fe034	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
a5be588e69b7b158302804816f74 25a48a1c171b	CP01152 4.1	<i>Intestinimonas butyriciproducens</i> strain SRB-521-5-I chromosome, complete genome
a620599610d0705c911ca067b326 9c7bfa891c0d	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04604 8.1	<i>Proteus mirabilis</i> strain HN2p chromosome HN2p, complete sequence
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05361 4.1	<i>Proteus mirabilis</i> strain S74-1(++)-2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05371 8.1	<i>Proteus mirabilis</i> strain MPE4069 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05368 4.1	<i>Proteus mirabilis</i> strain MPE5139 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05368 5.1	<i>Proteus mirabilis</i> strain MPE5203 chromosome, complete genome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP05368 1.1	<i>Proteus mirabilis</i> strain M3-1-17 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05368 2.1	<i>Proteus mirabilis</i> strain MPE0156 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05368 3.1	<i>Proteus mirabilis</i> strain MPE0027 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05371 9.1	<i>Proteus mirabilis</i> strain MPE0346 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05361 5.1	<i>Proteus mirabilis</i> strain MPE0734 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05361 6.1	<i>Proteus mirabilis</i> strain MPE0767 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04525 7.1	<i>Proteus mirabilis</i> strain L90-1 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05389 8.1	<i>Proteus mirabilis</i> strain YPM35 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05389 4.1	<i>Proteus mirabilis</i> strain JPM24 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05126 0.1	<i>Proteus mirabilis</i> strain STP3 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04994 2.1	<i>Proteus mirabilis</i> strain XH1569 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04994 1.1	<i>Proteus mirabilis</i> strain XH1568 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04975 3.1	<i>Proteus mirabilis</i> strain PmBR607 chromosome, complete genome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP04735 2.1	<i>Proteus mirabilis</i> strain ZA25 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04285 7.1	<i>Proteus mirabilis</i> strain 1701092 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04878 7.1	<i>Proteus mirabilis</i> strain CC15031 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04711 2.1	<i>Proteus mirabilis</i> strain SCBX1.1 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04553 8.2	<i>Proteus mirabilis</i> strain CRE14IB chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04413 4.1	<i>Proteus mirabilis</i> strain ENT1224 chromosome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04413 6.1	<i>Proteus mirabilis</i> strain ENT1157 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04413 5.1	<i>Proteus mirabilis</i> strain ENT1301 chromosome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04402 8.1	<i>Proteus mirabilis</i> strain K817 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04387 0.1	<i>Proteus mirabilis</i> strain S1959 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04333 2.1	<i>Proteus mirabilis</i> strain CRPM10 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04290 7.1	<i>Proteus mirabilis</i> strain VAC chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04139 1.1	<i>Proteus mirabilis</i> strain CLPM181223 chromosome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP03409 0.1	<i>Proteus mirabilis</i> strain PmSC1111 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP03409 1.1	<i>Proteus mirabilis</i> strain PmBC1123 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	LR13420 5.1	<i>Proteus mirabilis</i> strain NCTC4199 genome assembly, chromosome: 1
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02972 5.1	<i>Proteus mirabilis</i> strain AR_0029 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02327 3.1	<i>Proteus mirabilis</i> strain CCUG 70746 chromosome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02913 3.1	<i>Proteus mirabilis</i> strain AR379 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02889 7.1	<i>Dongshaea marina</i> strain DM2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02889 7.1	<i>Dongshaea marina</i> strain DM2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02835 6.1	<i>Proteus mirabilis</i> strain K670 chromosome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02605 1.1	<i>Proteus mirabilis</i> strain FDAARGOS_67 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02658 1.1	<i>Proteus mirabilis</i> isolate GN2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02657 1.1	<i>Proteus mirabilis</i> strain BC11-24 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02605 9.1	<i>Proteus mirabilis</i> strain FDAARGOS_80 chromosome, complete genome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP02604 4.1	<i>Proteus mirabilis</i> strain FDAARGOS_60 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP01708 5.1	<i>Proteus mirabilis</i> strain T18 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP01708 2.1	<i>Proteus mirabilis</i> strain T21, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP02185 2.1	<i>Proteus mirabilis</i> strain AR_0156 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07178 0.1	<i>Proteus mirabilis</i> strain swupm3 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07177 7.1	<i>Proteus mirabilis</i> strain swupm2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07177 3.1	<i>Proteus mirabilis</i> strain swupm1 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09845 0.1	<i>Proteus mirabilis</i> strain FZP3115 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09844 7.1	<i>Proteus mirabilis</i> strain FZP2936 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09844 4.1	<i>Proteus mirabilis</i> strain FZP3105 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09844 6.1	<i>Proteus mirabilis</i> strain FZP2826 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09677 6.1	<i>Proteus mirabilis</i> strain HURS-186083 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09677 5.1	<i>Proteus mirabilis</i> strain HURS-181823 chromosome, complete genome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP09576 5.1	<i>Proteus mirabilis</i> strain T1010 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP09265 2.1	<i>Proteus mirabilis</i> strain PM8762 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05509 5.1	<i>Proteus mirabilis</i> strain SWHIN_109 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05500 9.1	<i>Proteus mirabilis</i> strain STIN_74 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	OL79081 2.1	<i>Proteus mirabilis</i> strain H110 plasmid unnamed, complete sequence
a62ae4476d019c81ea031cf4fb659 41145764a13	OL79081 1.1	<i>Proteus mirabilis</i> strain H63 plasmid unnamed, complete sequence
a62ae4476d019c81ea031cf4fb659 41145764a13	CP08931 7.1	<i>Proteus mirabilis</i> strain PM1162 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP08637 7.1	<i>Proteus mirabilis</i> strain LB UEL H-11 chromosome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04869 2.1	<i>Proteus mirabilis</i> strain HNS2p chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06503 9.2	<i>Proteus mirabilis</i> strain XH1653 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07796 3.1	<i>Proteus mirabilis</i> strain 6Pmi283 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04758 9.1	<i>Proteus mirabilis</i> strain SNYG35 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07324 5.1	<i>Proteus mirabilis</i> strain S74-3-2 chromosome, complete genome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP07324 7.1	<i>Proteus mirabilis</i> strain S62-3-2-2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07324 8.1	<i>Proteus mirabilis</i> strain N292 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07324 6.1	<i>Proteus mirabilis</i> strain N639-2X chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07277 9.1	<i>Proteus mirabilis</i> strain 1035 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06683 3.1	<i>Proteus mirabilis</i> strain RGF134-1 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04604 9.1	<i>Proteus mirabilis</i> strain DY.F1.2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07057 2.1	<i>Proteus mirabilis</i> strain PM52808 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP07056 9.1	<i>Proteus mirabilis</i> strain PM52260 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04443 7.1	<i>Proteus mirabilis</i> strain C74 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04443 6.1	<i>Proteus mirabilis</i> strain C55 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06815 2.1	<i>Proteus mirabilis</i> strain FDAARGOS_1079 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06514 6.1	<i>Proteus mirabilis</i> strain PmBJ020-1 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06514 5.1	<i>Proteus mirabilis</i> strain PmBJ023-2 chromosome, complete genome

a62ae4476d019c81ea031cf4fb659 41145764a13	CP06514 4.1	<i>Proteus mirabilis</i> strain PmBJ024-1 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06514 8.1	<i>Proteus mirabilis</i> strain PmBJ012-2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06514 7.1	<i>Proteus mirabilis</i> strain PmBJ015-2 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06344 0.1	<i>Proteus mirabilis</i> strain Yak 2019 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP04792 9.1	<i>Proteus mirabilis</i> strain ChSC1905 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP06214 6.1	<i>Proteus mirabilis</i> strain S012 chromosome, complete genome
a62ae4476d019c81ea031cf4fb659 41145764a13	CP05905 6.1	<i>Proteus mirabilis</i> strain CriePir89 chromosome, complete genome
a69522903bd361ad83b5785db17c d535706766cf	CP09283 2.1	<i>Flavobacterium</i> sp. HJ-32-4 chromosome, complete genome
a6d1cb80edf55f1d63d335ba9a3fb c9eacca0524	LT62974 2.1	<i>Microterricola viridarii</i> strain DSM 21772 genome assembly, chromosome: I
a6d1cb80edf55f1d63d335ba9a3fb c9eacca0524	CP01414 5.1	<i>Microterricola viridarii</i> strain ERGS5:02 chromosome, complete genome
a71337c58d6b0ad05419dce7e753 a637bac1ee65	BK05508 4.1	MAG TPA_asm: <i>Microviridae</i> sp. isolate ctSXA1, partial genome
a74742032b2e4436f5f46bf215d3e 4d6370ea8db	HG99439 8.1	<i>Canis lupus</i> genome assembly, chromosome: 16
a78421b376ad7066ea4f65849ec6 d213785a1bed	CP04883 8.1	[<i>Clostridium</i>] <i>innocuum</i> strain ATCC 14501 chromosome, complete genome

a78421b376ad7066ea4f65849ec6 d213785a1bed	CP02272 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 genome
a78421b376ad7066ea4f65849ec6 d213785a1bed	CP01540 4.2	Erysipelotrichaceae bacterium I46 chromosome, complete genome
a78421b376ad7066ea4f65849ec6 d213785a1bed	CP09494 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain VE303-07 chromosome, complete genome
a78421b376ad7066ea4f65849ec6 d213785a1bed	BK03754 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctS5j6, partial genome
a78421b376ad7066ea4f65849ec6 d213785a1bed	CP06532 0.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 chromosome
a8b74912a0826baef77898bb07f46 b520603b8e1	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
a946b15d27661acb3e0b8c839829 da53ec3d562a	CP01422 3.1	<i>Anaerotignum propionicum</i> DSM 1682 strain X2 chromosome, complete genome
a961853d2d7198911d8b15198f7e 42f07538ef64	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
a96a324621bf18b8420ce07a2131 d68f422c8cba	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
a96ab756b007aed93851cfdde9d8a a9bf754b345	BK05508 4.1	MAG TPA_asm: <i>Microviridae</i> sp. isolate ctSXA1, partial genome
aa009a08c0e52cba0d121356b496 0c2a7ca27dfc	CP06499 3.1	MAG: <i>Candidatus Saccharibacteria</i> bacterium isolate Hirt_18-Q3-R61-65_MAXAC.330_cln chromosome, complete genome
aa205d30c8ebf0582086b018dcf93 c82b01619df	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
aa2da45ede5c8ace266834624c6af 3fb636fa502	CP04565 2.1	<i>Sphingobacterium</i> sp. dk4302 chromosome, complete genome

ab2369e0cf8d03abd72306936592 22be7f164638	CP09446 7.1	<i>Faecalibacterium</i> sp. I4179 chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP03077 7.1	<i>Faecalibacterium prausnitzii</i> strain APC918/95b chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP03077 7.1	<i>Faecalibacterium prausnitzii</i> strain APC918/95b chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP09447 3.1	<i>Faecalibacterium</i> sp. HTFF chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP09447 2.1	<i>Faecalibacterium</i> sp. IP118 chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP09446 9.1	<i>Faecalibacterium</i> sp. I3333 chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP09446 9.1	<i>Faecalibacterium</i> sp. I3333 chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	CP06538 2.1	<i>Faecalibacterium prausnitzii</i> strain Fp944 chromosome, complete genome
ab2369e0cf8d03abd72306936592 22be7f164638	FP92904 6.1	<i>Faecalibacterium prausnitzii</i> SL3/3 draft genome
ab2369e0cf8d03abd72306936592 22be7f164638	FP92904 6.1	<i>Faecalibacterium prausnitzii</i> SL3/3 draft genome
ab395af543d1a268812a1caee81cd 06f2e12677f	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
ab395af543d1a268812a1caee81cd 06f2e12677f	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
ab395af543d1a268812a1caee81cd 06f2e12677f	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome

ab395af543d1a268812a1caee81cd 06f2e12677f	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome
ab395af543d1a268812a1caee81cd 06f2e12677f	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
ab395af543d1a268812a1caee81cd 06f2e12677f	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
ab395af543d1a268812a1caee81cd 06f2e12677f	BK02513 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctmsk9, partial genome
abe23b94f7a6b2803d5c4290a490 2e7ffb82652d	LR13450 6.1	<i>Porphyromonas cangingivalis</i> strain NCTC12856 genome assembly, chromosome: 1
abe23b94f7a6b2803d5c4290a490 2e7ffb82652d	CP08191 3.1	<i>Bacteroides stercoris</i> strain BFG-121 chromosome, complete genome
ac19272556448b74a4eb3033b4ee 8c20986feb9b	AP01487 9.1	<i>Sulfuricaulis limicola</i> DNA, complete genome
ae31275f90a5c45513a14fe8bc919 f7291a464da	CP06063 3.1	<i>Simiaoa sunii</i> strain NSJ-8 chromosome, complete genome
ae32e73cf431a8491c06887d1b41 b81a45b89af9	CP07369 4.1	<i>Dysosmobacter</i> sp. Marseille-Q4140 chromosome
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84904 3.1	<i>Chlamydia pneumoniae</i> genome assembly YK41, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84725 7.1	<i>Chlamydia pneumoniae</i> genome assembly Wien3, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84725 4.1	<i>Chlamydia pneumoniae</i> genome assembly Wien2, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84724 5.1	<i>Chlamydia pneumoniae</i> genome assembly UZG1, chromosome : 1

af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84724 4.1	<i>Chlamydia pneumoniae</i> genome assembly U1271, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84724 1.1	<i>Chlamydia pneumoniae</i> genome assembly PB2, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84724 0.1	<i>Chlamydia pneumoniae</i> genome assembly PB1, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84723 5.1	<i>Chlamydia pneumoniae</i> genome assembly Panola, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84722 7.1	<i>Chlamydia pneumoniae</i> genome assembly MUL2216, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84722 1.1	<i>Chlamydia pneumoniae</i> genome assembly K7, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84719 0.1	<i>Chlamydia pneumoniae</i> genome assembly H12, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84705 1.1	<i>Chlamydia pneumoniae</i> genome assembly DC9, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84700 8.1	<i>Chlamydia pneumoniae</i> genome assembly GiD, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84700 4.1	<i>Chlamydia pneumoniae</i> genome assembly CWL029c, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84700 0.1	<i>Chlamydia pneumoniae</i> genome assembly CWL011, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84699 9.1	<i>Chlamydia pneumoniae</i> genome assembly CV15, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84699 6.1	<i>Chlamydia pneumoniae</i> genome assembly CV14, chromosome : 1

af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84699 5.1	<i>Chlamydia pneumoniae</i> genome assembly CM1, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LN84698 0.1	<i>Chlamydia pneumoniae</i> genome assembly Wien1, chromosome : 1
af4920f16c98817fbc368fb38bd35 142fd7f4b	LK00405 9.1	TPA: <i>Chlamydomonas reinhardtii</i> AR39 tRNA-Leu- CAG-1-1 gene
af4920f16c98817fbc368fb38bd35 142fd7f4b	CP00171 3.1	<i>Chlamydomonas reinhardtii</i> LPCoLN, complete genome
af4920f16c98817fbc368fb38bd35 142fd7f4b	AE00944 0.1	<i>Chlamydomonas reinhardtii</i> TW-183, complete genome
af4920f16c98817fbc368fb38bd35 142fd7f4b	AE00216 1.1	<i>Chlamydomonas reinhardtii</i> AR39, complete genome
af4920f16c98817fbc368fb38bd35 142fd7f4b	BA00000 8.3	<i>Chlamydomonas reinhardtii</i> J138 genomic DNA, complete sequence
af4920f16c98817fbc368fb38bd35 142fd7f4b	AE00136 3.1	<i>Chlamydomonas reinhardtii</i> CWL029, complete genome
afc745bd41bceb626aae44e4babe1 a7978f930af	AP01930 9.1	Erysipelotrichaceae bacterium SG0102 DNA, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP04843 7.1	<i>Faecalibacterium duncaniae</i> strain JCM 31915 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP04843 7.1	<i>Faecalibacterium duncaniae</i> strain JCM 31915 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP03077 7.1	<i>Faecalibacterium prausnitzii</i> strain APC918/95b chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP03077 7.1	<i>Faecalibacterium prausnitzii</i> strain APC918/95b chromosome, complete genome

afefcb078063908d1a7427f1b1094 13e28567ac9	LR69901 7.1	<i>Faecalibacterium prausnitzii</i> isolate MGYG-HGUT-02545 genome assembly, chromosome: 1
afefcb078063908d1a7427f1b1094 13e28567ac9	LR69901 7.1	<i>Faecalibacterium prausnitzii</i> isolate MGYG-HGUT-02545 genome assembly, chromosome: 1
afefcb078063908d1a7427f1b1094 13e28567ac9	CP02247 9.1	<i>Faecalibacterium prausnitzii</i> strain A2165 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP02247 9.1	<i>Faecalibacterium prausnitzii</i> strain A2165 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP02381 9.1	<i>Faecalibacterium prausnitzii</i> strain Indica chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP02381 9.1	<i>Faecalibacterium prausnitzii</i> strain Indica chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP10012 5.1	<i>Faecalibacterium duncaniae</i> strain NB2B-19-DCM chromosome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP10012 5.1	<i>Faecalibacterium duncaniae</i> strain NB2B-19-DCM chromosome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 3.1	<i>Faecalibacterium</i> sp. HTFF chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 3.1	<i>Faecalibacterium</i> sp. HTFF chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 2.1	<i>Faecalibacterium</i> sp. IP118 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 2.1	<i>Faecalibacterium</i> sp. IP118 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 1.1	<i>Faecalibacterium</i> sp. IP329 chromosome, complete genome

afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 1.1	<i>Faecalibacterium</i> sp. IP329 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 0.1	<i>Faecalibacterium</i> sp. I2392 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09447 0.1	<i>Faecalibacterium</i> sp. I2392 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09446 9.1	<i>Faecalibacterium</i> sp. I3333 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09446 9.1	<i>Faecalibacterium</i> sp. I3333 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09446 8.1	<i>Faecalibacterium</i> sp. I3389 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09446 7.1	<i>Faecalibacterium</i> sp. I4179 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09446 6.1	<i>Faecalibacterium</i> sp. I4384 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP09446 6.1	<i>Faecalibacterium</i> sp. I4384 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP06538 2.1	<i>Faecalibacterium prausnitzii</i> strain Fp944 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP06537 6.1	<i>Faecalibacterium prausnitzii</i> strain Fp360 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP06537 6.1	<i>Faecalibacterium prausnitzii</i> strain Fp360 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	CP06538 0.1	<i>Faecalibacterium prausnitzii</i> strain Fp28 chromosome, complete genome

afefcb078063908d1a7427f1b1094 13e28567ac9	CP06538 1.1	<i>Faecalibacterium prausnitzii</i> strain Fp1 chromosome, complete genome
afefcb078063908d1a7427f1b1094 13e28567ac9	FP92904 6.1	<i>Faecalibacterium prausnitzii</i> SL3/3 draft genome
afefcb078063908d1a7427f1b1094 13e28567ac9	FP92904 5.1	<i>Faecalibacterium prausnitzii</i> L2/6 draft genome
afefcb078063908d1a7427f1b1094 13e28567ac9	GQ87700 5.1	Uncultured organism clone VC1AI32TR genomic sequence
b00cc9621b5f9e0fde6bf36692821 26c3bbc4d4c	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
b00cc9621b5f9e0fde6bf36692821 26c3bbc4d4c	AP02416 9.1	Lachnospiraceae bacterium TB5 DNA, complete genome
b1355eeb5d645530d549a649fa7df bd52706b997	CP04642 7.1	<i>Phocaeicola dorei</i> strain JR04 chromosome, complete genome
b1355eeb5d645530d549a649fa7df bd52706b997	CP04361 0.1	<i>Bacteroides fragilis</i> strain GUT04 chromosome, complete genome
b1355eeb5d645530d549a649fa7df bd52706b997	CP03655 0.1	<i>Bacteroides fragilis</i> strain DCMOUH0042B chromosome, complete genome
b1355eeb5d645530d549a649fa7df bd52706b997	CP01270 6.1	<i>Bacteroides fragilis</i> strain S14 chromosome, complete genome
b1355eeb5d645530d549a649fa7df bd52706b997	CP00874 1.1	<i>Bacteroides dorei</i> isolate HS1_L_3_B_079, complete genome
b1355eeb5d645530d549a649fa7df bd52706b997	AP00684 1.1	<i>Bacteroides fragilis</i> YCH46 DNA, complete genome
b1370592fa99cafc10870f436690 e432a8f435e	CP04583 5.1	<i>Lysinibacillus pakistanensis</i> strain LZH-9 chromosome, complete genome

b1370592fa99cafc10870f436690e432a8f435e	CP02722 4.1	<i>Lysinibacillus</i> sp. B2A1 chromosome, complete genome
b1e0c80b4d7f2c5056e542f455d3e98977736cb6	CP09171 8.1	MAG: <i>Oscillospiraceae</i> bacterium isolate nC33_bin.8.fa chromosome
b2053ff4060490f2d6cc03185afd238f6fb25ae6	CP04864 9.1	<i>Aminipila butyrica</i> strain DSM 103574 chromosome, complete genome
b20f1b418efb53e2fdbe51b75d709c4151f3846a	CP04246 4.1	<i>Prevotella copri</i> strain YF2 chromosome
b22e080709bb3ec9c56df5c3b00e5b75df116222	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
b2e59c7343d4949dc8457e49b7c48702f978990a	LT90666 2.1	<i>Thermoanaerobacterium</i> sp. RBIITD genome assembly, chromosome: I
b31af5f911b1452fc12bf544d37e267a535b14e3	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
b31af5f911b1452fc12bf544d37e267a535b14e3	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
b3b8f9962fbdc9ae2264a053044b3e72e73ed801	CP09173 1.1	MAG: <i>Candidatus Melainabacteria</i> bacterium isolate nC33_lr_bin.1.fa chromosome
b3f8323bf3fb3f260fdff4a9514cb6edc5e2e2da	MN9907 34.1	Uncultured <i>Prevotella</i> sp. isolate LL70 clone UCI_7, sequence
b3f8323bf3fb3f260fdff4a9514cb6edc5e2e2da	BK04661 4.1	MAG TPA_asm: <i>Myoviridae</i> sp. isolate ctTeG6, partial genome
b3fa9150099b476a91db03aa2e9b36cb9a965441	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
b468b7c8f525f0d99720fb3670c5994b9a9bb902	OV78864 0.1	uncultured <i>Tenericutes</i> bacterium isolate min17_bin11 genome assembly, chromosome: 1

b590548ec560af4a0625a65f8ca5d780ad5745be	CP02257 9.1	<i>Oryzomicrobium terrae</i> strain TPP412 chromosome, complete genome
b590548ec560af4a0625a65f8ca5d780ad5745be	CP02218 8.1	<i>Azoarcus communis</i> strain TSNA42 chromosome, complete genome
b590548ec560af4a0625a65f8ca5d780ad5745be	CP02218 7.1	<i>Azoarcus communis</i> strain TSPY31 chromosome, complete genome
b5d4f88af6540b1febb85b1d32cb57e986932f65	CP06564 3.1	<i>Lysinibacillus macroides</i> strain FDAARGOS_925 chromosome, complete genome
b7ae6e6d26cbccd0eccd6ce88b13d1202de9526b	CP04422 7.1	<i>Flintibacter</i> sp. KGMB00164 chromosome, complete genome
b7bf47f567c29b2f374ebaea59e0616749bd2dbb	CP01260 1.1	<i>Bacillus</i> sp. FJAT-22090 genome
b7c730e07f87f45ae820253d29b72dcf272bbff7	LS48344 7.1	<i>Porphyromonas crevioricanis</i> strain NCTC12858 genome assembly, chromosome: 1
b80afc4d11629fda02b117a390c25eaf8705b089	CP09446 7.1	<i>Faecalibacterium</i> sp. I4179 chromosome, complete genome
b80afc4d11629fda02b117a390c25eaf8705b089	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
b80afc4d11629fda02b117a390c25eaf8705b089	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome
b80afc4d11629fda02b117a390c25eaf8705b089	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
b80afc4d11629fda02b117a390c25eaf8705b089	BK02513 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctmsk9, partial genome
b822d88a59fa918efc1179bff3ac711a41d2f045	BK02132 2.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctjse2, partial genome

b82fdc4e18b0119dcfc90a312e06e63b70621436	AP01930 9.1	Erysipelotrichaceae bacterium SG0102 DNA, complete genome
b935efc1b2ba214301524bd02f49d4886097384a	CP02108 1.1	<i>Deinococcus ficus</i> strain CC-FR2-10 chromosome, complete genome
b9adbafcd6564ff056dd92e53936f783a1892344	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
ba36ebf8f62f609f48a9e2ee4e5558bd71d1c1b0	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
bbd00dff35e48210021693d1dfdb450e6968cdea	AP02448 4.1	<i>Prevotella herbatica</i> WR041 DNA, complete genome
bbd591b0498c76e9ef1c1ccbb757d2cb5965b8bc	CP02218 8.1	<i>Azoarcus communis</i> strain TSNA42 chromosome, complete genome
bbd591b0498c76e9ef1c1ccbb757d2cb5965b8bc	CP02218 7.1	<i>Azoarcus communis</i> strain TSPY31 chromosome, complete genome
bccf53c3e029448606f837868a0140198de30926	CP02212 1.1	<i>Dehalobacterium formicoaceticum</i> strain DMC chromosome, complete genome
bd129709c14d3d435f3a3bef55529043eb9efec6	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
bd93abb1ffd55269b33f40cc065a7806d889fed3	CP09173 5.1	MAG: <i>Clostridium</i> sp. isolate nC29_bin.1.fa chromosome, complete genome
bdc89446f751fcaeeb921d165b56c4f6d4972e53	LR99447 2.1	<i>Catenibacterium</i> sp. co_0103 isolate <i>Catenibacterium canine</i> HQ MAG genome assembly, chromosome: 1
bf06677c904ed7b86e04a44a2ad3b5fe59a2184e	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
bf3b8924b7a23879d8825e842b6b4e5fdd5d7659	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome

bf3b8924b7a23879d8825e842b6b4e5fdd5d7659	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
bf3e53e92e204eff66f30b33bb73d48e3b1c80f9	CP04565 2.1	<i>Sphingobacterium</i> sp. dk4302 chromosome, complete genome
bf85da7f11021a56979e2244de46867827e884fa	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
bf85da7f11021a56979e2244de46867827e884fa	CP00703 2.1	<i>Desulfitobacterium metallireducens</i> DSM 15288, complete genome
bfd87d072238bffa30b61784647eeeb56873af	CP00245 2.1	<i>Nitratifractor salsuginis</i> DSM 16511, complete genome
bfeaa1ddef52fbdd79467c6c6cd6f44d9714d4f2	AP02556 1.1	Oscillospiraceae bacterium CE91-St42 DNA, complete genome
c0639fd43475df1dbe175a21d1a5bf63e9e811df	CP05418 2.1	MAG: <i>Ignavibacteriales</i> bacterium isolate H2_BAC3 chromosome, complete genome
c1cbdb92d4a2902bd7b99eff4a05f505b9c835a9	CP00334 5.1	<i>Bernardetia litoralis</i> DSM 6794 chromosome, complete genome
c286a9da559537b63e70a125c48c5cb5080e6fc8	CP06499 3.1	MAG: <i>Candidatus Saccharibacteria</i> bacterium isolate Hirt_18-Q3-R61-65_MAXAC.330_cln chromosome, complete genome
c2c713e13f9091d65a58240a8626848b1b3b2aec	AP02266 0.1	<i>Bacteroides thetaiotaomicron</i> F9-2 DNA, nearly complete genome
c2c713e13f9091d65a58240a8626848b1b3b2aec	CP04569 6.1	<i>Sodaliphilus pleomorphus</i> strain Oil-RF-744-WCA-WT-10 chromosome, complete genome
c2c713e13f9091d65a58240a8626848b1b3b2aec	CP04137 9.1	<i>Bacteroides intestinalis</i> strain APC919/174 chromosome, complete genome
c2c713e13f9091d65a58240a8626848b1b3b2aec	CP04123 0.1	<i>Bacteroides xylanisolvens</i> strain H207 chromosome, complete genome

c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP09173 2.1	MAG: Porphyromonadaceae bacterium isolate nC29_bin.32.fa chromosome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP09173 2.1	MAG: Porphyromonadaceae bacterium isolate nC29_bin.32.fa chromosome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP06494 0.1	<i>Bacteroides intestinalis</i> strain APC919/174 substr. 8W chromosome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP06494 1.1	<i>Bacteroides intestinalis</i> strain APC919/174 substr. 8T chromosome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP08190 2.1	<i>Bacteroides salyersiae</i> strain BFG-256 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP08189 9.1	<i>Bacteroides salyersiae</i> strain BFG-288 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP08189 8.1	<i>Bacteroides thetaiotaomicron</i> strain BFG-55 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP07224 2.1	<i>Bacteroides thetaiotaomicron</i> strain CL06T03C18 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP07222 4.1	<i>Bacteroides thetaiotaomicron</i> strain CL11T00C24 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP07221 6.1	<i>Bacteroides xylanisolvens</i> strain CL11T00C03 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP07221 2.1	<i>Bacteroides xylanisolvens</i> strain CL11T00C41 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP06817 4.1	<i>Bacteroides caccae</i> strain FDAARGOS_1097 chromosome, complete genome
c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP05997 3.1	<i>Bacteroides luhongzhouii</i> strain HF-5141 chromosome, complete genome

c2c713e13f9091d65a58240a8626 848b1b3b2aec	CP05985 6.1	<i>Bacteroides zhangwenhongii</i> strain HF-5287 chromosome, complete genome
c33d6eb0f223e3fcbb39d6015483f 626d2270467	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
c3ac11224ba485917e150c57a330 6b9010b49cf5	CP10111 2.1	<i>Moraxella</i> sp. FZLJ2109 chromosome, complete genome
c3ac11224ba485917e150c57a330 6b9010b49cf5	CP10111 1.1	<i>Moraxella</i> sp. FZLJ2107 chromosome, complete genome
c3ac11224ba485917e150c57a330 6b9010b49cf5	CP09996 0.1	<i>Moraxella</i> sp. FZFQ2102 chromosome, complete genome
c43d9463ce3e9f08b1df9b9e52b4b 13a9d5fbbc2	CP01620 2.1	<i>Eubacterium minutum</i> ATCC 700079 sequence
c4a754740948f837cfd5b9963ff29 b57244edf75	CP04764 7.1	<i>Hymenobacter busanensis</i> strain BT182 chromosome, complete genome
c555de279ea43243e231ed002abb a2f0d15f0f3b	AP02266 0.1	<i>Bacteroides thetaiotaomicron</i> F9-2 DNA, nearly complete genome
c555de279ea43243e231ed002abb a2f0d15f0f3b	CP04123 0.1	<i>Bacteroides xylanisolvens</i> strain H207 chromosome, complete genome
c555de279ea43243e231ed002abb a2f0d15f0f3b	CP03655 3.1	<i>Bacteroides fragilis</i> strain DCMOUH0067B chromosome, complete genome
c555de279ea43243e231ed002abb a2f0d15f0f3b	CP01540 1.2	<i>Bacteroides caecimuris</i> strain I48 chromosome, complete genome
c5cd9d4b95b3a555b3e33ee7704d 5df17048c8ca	CP09172 8.1	MAG: <i>Clostridium</i> sp. isolate nC33_bin.207.fa chromosome
c60bd77db885fc9a0af38cb539208 57a4b001435	CP02778 1.1	<i>Clostridium botulinum</i> strain MAP 5 chromosome, complete genome

c60bd77db885fc9a0af38cb539208 57a4b001435	CP02777 9.1	<i>Clostridium botulinum</i> strain RF5 chromosome, complete genome
c60bd77db885fc9a0af38cb539208 57a4b001435	CP02777 7.1	<i>Clostridium botulinum</i> strain Mfbjulcb8 chromosome, complete genome
c688ceca6f9b95e021be00c3b2582 a8a48e477de	LR13438 4.1	<i>Prevotella oris</i> strain NCTC13071 genome assembly, chromosome: 1
c688ceca6f9b95e021be00c3b2582 a8a48e477de	CP08593 2.1	<i>Prevotella copri</i> DSM 18205 strain FDAARGOS_1573 chromosome, complete genome
c688ceca6f9b95e021be00c3b2582 a8a48e477de	BK02984 9.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctcp6, partial genome
c688ceca6f9b95e021be00c3b2582 a8a48e477de	BK02894 0.1	MAG TPA_asm: <i>Bacteriophage</i> sp. isolate ctux411, partial genome
c688ceca6f9b95e021be00c3b2582 a8a48e477de	CP04246 4.1	<i>Prevotella copri</i> strain YF2 chromosome
c6beadbda3442132af61357b226c 525c43e81a58	GQ87700 5.1	Uncultured organism clone VC1AI32TR genomic sequence
c6c261889e3bb254a70126543533 cd77421116e7	CP03456 2.1	<i>Flammeovirga pectinis</i> strain L12M1 chromosome 1, complete sequence
c6c261889e3bb254a70126543533 cd77421116e7	CP07612 8.1	<i>Flammeovirga kamogawensis</i> strain YS10 chromosome 1, complete sequence
c725cb53df47fd37b90af3b9e4a42 247939db191	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
c772242f0db1ba5e50815ab77925 b1752ee23e62	CP06499 6.1	MAG: bacterium isolate Kalu_18-Q3-R12-55_BAT3C.2_fly chromosome, complete genome
c7e534b2b59d32c23972d4eb36be 0dc3c680b477	LR21598 1.1	<i>Ruthenibacterium lactatiformans</i> isolate <i>Ruthenibacterium lactatiformans</i> 82B1 genome assembly, chromosome: 1

c7eea2d12874dc6755c7b7845860 f2c650bd90b3	CP02267 4.1	<i>Priestia megaterium</i> strain SR7 chromosome, complete genome
c80b040e6308f1c046a1912b11a4f d40244bac2f	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
c80b5c0dac79a59b78814ae3a6ef5 3a4e247753e	CP02212 1.1	<i>Dehalobacterium formicoaceticum</i> strain DMC chromosome, complete genome
c860f7e12e39afff69146cb0c66b4 2172ad1e7fe	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
c860f7e12e39afff69146cb0c66b4 2172ad1e7fe	AP02416 9.1	Lachnospiraceae bacterium TB5 DNA, complete genome
c88b8f5818b6ecd3a598601c7038 b9d8845fd4be	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
c93dc71cfd3b6332d51bd9e2414d ec75903a867e	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
c93dc71cfd3b6332d51bd9e2414d ec75903a867e	OU94086 2.1	uncultured fungus genomic DNA sequence contains 18S rRNA gene, ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gene, isolate consensus02156
ca744fd44189628a5ac782d4f916b c891fd8f2d2	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
cb6307a509ec7762bafcca7e34789 b8ffd3d8b6d	CP00245 2.1	<i>Nitratifractor salsuginis</i> DSM 16511, complete genome
cb6b331f5e744447458d1b67d3f4 3e1396fb517a	CP03528 1.1	<i>Aminipila luticellarii</i> strain JN-18 chromosome, complete genome
cbd9bda8ff4c95520ac59ef956a9b e727595d66a	CP05338 9.1	Bacterium BFN5 strain BFN5 chromosome
cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07201 9.1	<i>Akkermansia muciniphila</i> strain Akk2680 chromosome, complete genome

cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07202 1.1	<i>Akkermansia muciniphila</i> strain Akk2196 chromosome, complete genome
cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07202 2.1	<i>Akkermansia muciniphila</i> strain Akk2190 chromosome, complete genome
cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07202 5.1	<i>Akkermansia muciniphila</i> strain Akk2000 chromosome, complete genome
cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07203 6.1	<i>Akkermansia muciniphila</i> strain Akk1573 chromosome, complete genome
cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07203 8.1	<i>Akkermansia muciniphila</i> strain Akk1476 chromosome, complete genome
cc20ed8b15d74c1c5088542cb54fe 0bd3399acf9	CP07204 4.1	<i>Akkermansia muciniphila</i> strain Akk0580 chromosome, complete genome
cce21810520c7b9a7eee49f98d051 07c4c831da2	OA93366 6.1	<i>Notodromas monacha</i>
cd7302acc525e972aa3220152a99 27b2518ac01b	CP04090 2.1	<i>Longibaculum</i> sp. KGMB06250 chromosome, complete genome
cd7302acc525e972aa3220152a99 27b2518ac01b	CP03183 7.1	<i>Clostridium sporogenes</i> strain YH-Raf3-2-5 chromosome, complete genome
cd7302acc525e972aa3220152a99 27b2518ac01b	AP01853 7.1	Erysipelotrichaceae bacterium GAM147 DNA, complete genome
cd7302acc525e972aa3220152a99 27b2518ac01b	AP02408 5.1	<i>Faecalibacillus intestinalis</i> 14EGH31 DNA, complete genome
cdbc1d04e3c9eb70b98d19d535ac 447c2ab979a8	CP00189 6.1	<i>Allochromatium vinosum</i> DSM 180 chromosome, complete genome
cdfc3ec6bbb4830a5e01471c94e28 838e8a86d36	CP00304 1.1	<i>Pseudomonas fluorescens</i> A506, complete genome

ce925e0a654e117a7a5a3c051ffbe bdc11d9e1fa	CP06063 6.1	<i>Eubacterium hominis</i> strain NSJ-61 chromosome, complete genome
ced836db448ada2463a7d6171504 901b31a7bb32	CP02636 5.1	<i>Neobacillus thermocopriae</i> strain DUT50_236 chromosome, complete genome
ced836db448ada2463a7d6171504 901b31a7bb32	CP02081 5.1	<i>Anoxybacillus flavithermus</i> strain DSM 2641T chromosome, complete genome
ced836db448ada2463a7d6171504 901b31a7bb32	CP02183 8.1	<i>Anoxybacillus flavithermus</i> strain 52-1A chromosome, complete genome
ced836db448ada2463a7d6171504 901b31a7bb32	CP01769 0.1	<i>Geobacillus thermodenitrificans</i> strain ID-1 chromosome, complete genome
ced836db448ada2463a7d6171504 901b31a7bb32	CP02003 0.1	<i>Geobacillus thermodenitrificans</i> strain T12 chromosome, complete genome
ced836db448ada2463a7d6171504 901b31a7bb32	CP06147 3.1	<i>Geobacillus</i> sp. 47C-IIb chromosome, complete genome
cee44f13eb3b60748c0564ec7cd82 fccd61b4833	CP09233 5.1	<i>Bacillus</i> sp. OVS6 chromosome, complete genome
cee44f13eb3b60748c0564ec7cd82 fccd61b4833	CP08570 8.1	<i>Bacillus</i> sp. CMF21 chromosome, complete genome
cee44f13eb3b60748c0564ec7cd82 fccd61b4833	CP01260 1.1	<i>Bacillus</i> sp. FJAT-22090 genome
cfa4542e2f1624eadb18e01af4f76f 4fa7cceedf	CP09172 5.1	MAG: <i>Bacteroidales</i> bacterium isolate nC33_bin.253.fa chromosome, complete genome
d020e0a9da28c22065e010c22220 309bf10e937d	CP03634 8.1	<i>Rosistilla carotiformis</i> strain Poly24 chromosome
d020e0a9da28c22065e010c22220 309bf10e937d	CP03631 8.1	<i>Rosistilla oblonga</i> strain Mal33 chromosome

d09116a5ec84d33af0cd153322a99242e4c5dd83	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
d0c793acec41aa58a56a933a1be9bee7190df4c6	CP01206 8.1	<i>Eubacterium sulci</i> ATCC 35585, complete genome
d0caa2bbdb3a94f42bf922e7bcde2cc908dec35f	CP00245 2.1	<i>Nitratifractor salsuginis</i> DSM 16511, complete genome
d15d8f959531b5dfdf8ad21eb150750dfd1f9393	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
d1badc1690aa0ebedb73c0ab66108765ad3a8974	BK03525 2.1	MAG TPA_asm: CrAss-like virus sp. isolate ctwJE30, partial genome
d1fedabf57c3fabaa028f64f09ec4e695694eb10	CP07808 8.1	<i>Allobaculum</i> sp. Allo2 chromosome, complete genome
d2b75fd99d3cce66c696d8e50e8d7d6c8491db80	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
d2d6b4d07cd7a30ab527fde1dd6cc467ed550bb7	LR21598 1.1	<i>Ruthenibacterium lactatiformans</i> isolate <i>Ruthenibacterium lactatiformans</i> 82B1 genome assembly, chromosome: 1
d3d47e6ade54b7163760e68af0e064721337f8a3	CP03510 7.1	<i>Ornithobacterium rhinotracheale</i> strain FARPER-174b chromosome, complete genome
d3d6704819d7a18a91f339f84e5c659ad72a2fd5	CP08568 9.1	<i>Lentisphaerae</i> bacterium WC36 chromosome, complete genome
d3e7be204b7ef6fd52ea854ea4bad871e4f2f9ec	CP00014 1.1	<i>Carboxydotherrmus hydrogenoformans</i> Z-2901, complete genome
d3e8ee7092f0062d756bc405e2a3e510f20615a5	LR69899 9.1	<i>Clostridiales</i> bacterium isolate MGYG-HGUT-02424 genome assembly, chromosome: 1
d411ec2c756db341af2ddf93bab7ffff6f21dcf52	CP04093 9.1	<i>Dolosigranulum pigrum</i> strain KPL1934_CDC4709-98 chromosome

d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04093 8.1	<i>Dolosigranulum pigrum</i> strain KPL1914 chromosome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04162 6.1	<i>Dolosigranulum pigrum</i> strain 83VPs-KB5 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04040 8.1	<i>Dolosigranulum pigrum</i> strain KPL3911 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04040 9.1	<i>Dolosigranulum pigrum</i> strain KPL3274 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 0.1	<i>Dolosigranulum pigrum</i> strain KPL3264 chromosome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 1.1	<i>Dolosigranulum pigrum</i> strain KPL3256 chromosome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 2.1	<i>Dolosigranulum pigrum</i> strain KPL3250 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 3.1	<i>Dolosigranulum pigrum</i> strain KPL3246 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 4.1	<i>Dolosigranulum pigrum</i> strain KPL3090 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 5.1	<i>Dolosigranulum pigrum</i> strain KPL3086 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 6.1	<i>Dolosigranulum pigrum</i> strain KPL3084 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 7.1	<i>Dolosigranulum pigrum</i> strain KPL3077 chromosome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 8.1	<i>Dolosigranulum pigrum</i> strain KPL3070 chromosome, complete genome

d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04041 9.1	<i>Dolosigranulum pigrum</i> strain KPL3069 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04042 0.1	<i>Dolosigranulum pigrum</i> strain KPL3065 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04042 1.1	<i>Dolosigranulum pigrum</i> strain KPL3052 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04042 2.1	<i>Dolosigranulum pigrum</i> strain KPL3050 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04042 3.1	<i>Dolosigranulum pigrum</i> strain KPL3043 chromosome, complete genome
d411ec2c756db341af2ddf93bab7f ff6f21dcf52	CP04042 4.1	<i>Dolosigranulum pigrum</i> strain KPL3033 chromosome, complete genome
d45565c435342a7bb91f00968e0f 198e83d930f4	CP01783 1.1	<i>Butyrivibrio hungatei</i> strain MB2003 chromosome I, complete sequence
d485ce9c10cf9c76a0ac3c45fca89 bfb44e0d8bb	CP06061 2.1	<i>Staphylococcus aureus</i> strain S6 chromosome
d4c91b31b4517389ab75049e14cd 369f4e03764f	LK99601 7.1	<i>Desulfitobacterium hafniense</i> strain PCE-S genome assembly, scaffold: scaffold0
d4c91b31b4517389ab75049e14cd 369f4e03764f	LR74649 6.1	Peptococcaceae bacterium INE isolate PacBioINE genome assembly, chromosome: 1
d4c91b31b4517389ab75049e14cd 369f4e03764f	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
d4c91b31b4517389ab75049e14cd 369f4e03764f	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
d4c91b31b4517389ab75049e14cd 369f4e03764f	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome

d4c91b31b4517389ab75049e14cd 369f4e03764f	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
d4c91b31b4517389ab75049e14cd 369f4e03764f	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
d4c91b31b4517389ab75049e14cd 369f4e03764f	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
d4c91b31b4517389ab75049e14cd 369f4e03764f	AP00823 0.1	<i>Desulfitobacterium hafniense</i> Y51 DNA, complete genome
d4e2f035a67a984142fd88f221ce4 1147e824915	CP09172 7.1	MAG: <i>Lentisphaeria</i> bacterium isolate nC33_lr_bin.207.fa chromosome
d5f5bd7762c22e88247e3911858d ee443d84b5cd	CP04653 2.1	<i>Mannheimia</i> sp. ZY171111 chromosome, complete genome
d719a1aab1a59bb60313e9e403e7 21b5ea45954c	CP03563 2.1	<i>Synechococcus</i> sp. RSCCF101 chromosome, complete genome
d8a3e9467faeb98f46790fbc4b09d 8cc50a348d5	OV78864 0.1	uncultured <i>Tenericutes</i> bacterium isolate min17_bin11 genome assembly, chromosome: 1
d8af25967e25cdc96064f2f464cb4 6f589f01731	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
d9b3de5c78a53e5299e96527c134 15e2f9206dac	CP04583 5.1	<i>Lysinibacillus pakistanensis</i> strain LZH-9 chromosome, complete genome
da0c6d6376b6c142d66d490a7eee c385a91339c1	U26683. 1	<i>Dichelobacter nodosus</i> rrnA operon 16S ribosomal RNA (rrs) and 23S ribosomal RNA (rrl) genes, partial sequence, and tRNA-Ile and tRNA-Ala genes, complete sequence
da6a8fad3cc4592e65ad56915e39a 94744c31a2e	CP04503 5.1	<i>Liquorilactobacillus mali</i> strain LM596 chromosome, complete genome

da72dabcec18a7eaac549531a615c c51b405ed40	CP10111 2.1	<i>Moraxella</i> sp. FZLJ2109 chromosome, complete genome
da72dabcec18a7eaac549531a615c c51b405ed40	CP10111 1.1	<i>Moraxella</i> sp. FZLJ2107 chromosome, complete genome
da72dabcec18a7eaac549531a615c c51b405ed40	CP09996 0.1	<i>Moraxella</i> sp. FZFQ2102 chromosome, complete genome
db27880e7c5e149d33bd488fdd09 e51e4afc2f47	AP01229 2.1	<i>Selenomonas ruminantium</i> subsp. lactilytica TAM6421 DNA, complete genome
db338f6c0fd8afc7cbe0f185caaa08 199ca3b335	CP05418 2.1	MAG: <i>Ignavibacteriales</i> bacterium isolate H2_BAC3 chromosome, complete genome
dd4b2d0b08a32baf5eef4e3fb8ad6 de28c5f8a6a	BK02503 3.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct6IQ4, partial genome
dd4b2d0b08a32baf5eef4e3fb8ad6 de28c5f8a6a	BK04688 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctckW2, partial genome
dd776b4b773b9f43334783f583a6 0581d880802d	CP09283 2.1	<i>Flavobacterium</i> sp. HJ-32-4 chromosome, complete genome
de151985af51a8e83fbc27bba330 21f6c8132a4	CP09283 2.1	<i>Flavobacterium</i> sp. HJ-32-4 chromosome, complete genome
df868f081a27adecd5f85fd5d8321 7083dc43394	CP01597 1.1	<i>Arachidicoccus</i> sp. BS20 chromosome, complete genome
dfecfda48bb287b5eb3950556b84c a8767dcbce5	LR13452 5.1	<i>Neisseria meningitidis</i> strain NCTC10025 genome assembly, chromosome: 1
dfecfda48bb287b5eb3950556b84c a8767dcbce5	CP02042 0.2	<i>Neisseria meningitidis</i> strain FDAARGOS_209 chromosome, complete genome
dfecfda48bb287b5eb3950556b84c a8767dcbce5	CP06436 7.1	<i>Neisseria meningitidis</i> strain PartJ-Nmeningitidis-RM8376 chromosome, complete genome

dfecfda48bb287b5eb3950556b84ca8767dcbce5	CP06565 3.1	<i>Neisseria meningitidis</i> strain FDAARGOS_914 chromosome, complete genome
e088361219d29dd201347acce866cea5cf74bde0	CP06382 2.1	<i>Clostridium botulinum</i> strain 1873 chromosome, complete genome
e088361219d29dd201347acce866cea5cf74bde0	CP00241 0.1	<i>Clostridium botulinum</i> BKT015925, complete genome
e15511dd44a82033e165fda4ab9ad0cd0b7ab112	BK05900 1.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctnvA6, partial genome
e1c3c0f1f925d6ff7a1e39d989d1d726c2f582f7	CP06063 2.1	<i>Wujia chipingensis</i> strain NSJ-4 chromosome, complete genome
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	LK99601 7.1	<i>Desulfitobacterium hafniense</i> strain PCE-S genome assembly, scaffold: scaffold0
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	LR74649 6.1	Peptococcaceae bacterium INE isolate PacBioINE genome assembly, chromosome: 1
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	CP00334 4.1	<i>Desulfitobacterium dichloroeliminans</i> LMG P-21439, complete genome
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	CP00334 8.1	<i>Desulfitobacterium dehalogenans</i> ATCC 51507, complete genome
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	CP00133 6.1	<i>Desulfitobacterium hafniense</i> DCB-2, complete genome
e2a4c6dcfe0a23d8eb16fd44803866c8c29f1d45	AP00823 0.1	<i>Desulfitobacterium hafniense</i> Y51 DNA, complete genome

e2a90f69f2d40d45f9d33191a70d43570c52bbc6	CP07808 8.1	<i>Allobaculum</i> sp. Allo2 chromosome, complete genome
e33e60778e063bc32e9d04ac4cc4586f13a280ca	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
e50ba5e885560308e5120a318c59f2d892f55d31	CP02777 5.1	<i>Clostridium botulinum</i> strain MFBjulcb1 chromosome, complete genome
e589a870e3f061deb50725b4b5ac2fc2ba21692a	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
e589a870e3f061deb50725b4b5ac2fc2ba21692a	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
e693066801c8f89691c81388893d8cb72b5694c2	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
e6de4974b3af0db88c1049f38471096091177dbb	CP03625 9.1	<i>Sporomusa termitida</i> strain DSM 4440 chromosome, complete genome
e6e9d773e80eb0709f1d2700ad74c4e6f59f982b	CP05491 9.1	<i>Kitasatospora</i> sp. NA04385 chromosome, complete genome
e72cfe229df93e6da4df3f3d151102e25c869b4e	LT63232 2.1	<i>Murdochiella</i> sp. Marseille-P2341 strain Marseille-P2341T genome assembly, chromosome: contig00001
e89c723c1838ebc3c3d955cfe9c6d0a43d1f9aea	CP05401 2.1	<i>Parabacteroides distasonis</i> strain FDAARGOS_759 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d0a43d1f9aea	CP05167 2.1	<i>Parabacteroides distasonis</i> strain CBBP-1 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d0a43d1f9aea	CP05095 6.1	<i>Parabacteroides distasonis</i> strain FDAARGOS_615 chromosome
e89c723c1838ebc3c3d955cfe9c6d0a43d1f9aea	AP01972 9.1	<i>Parabacteroides distasonis</i> NBRC 113806 DNA, complete genome

e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP04046 8.1	<i>Parabacteroides distasonis</i> strain CavFT-hAR46 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	LR21597 8.1	<i>Parabacteroides distasonis</i> ATCC 8503 isolate <i>Parabacteroides distasonis</i> 82G9 genome assembly, chromosome: 1
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP02275 4.1	<i>Parabacteroides</i> sp. CT06 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP09728 5.1	<i>Parabacteroides distasonis</i> strain NB2A-10-MRS chromosome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	LK01214 8.1	TPA: <i>Parabacteroides distasonis</i> ATCC 8503 tRNA- Leu-TAA-1-1 gene
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP08594 4.1	<i>Parabacteroides distasonis</i> strain FDAARGOS_1565 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP08597 5.1	<i>Parabacteroides johnsonii</i> strain FDAARGOS_1580 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP08592 7.1	<i>Parabacteroides merdae</i> strain FDAARGOS_1586 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP08190 6.1	<i>Parabacteroides goldsteinii</i> strain BFG-241 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP08190 8.1	<i>Parabacteroides distasonis</i> strain BFG-238 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP08190 1.1	<i>Parabacteroides merdae</i> strain BFG-280 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP07225 4.1	<i>Parabacteroides distasonis</i> CL03T12C09 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP07223 1.1	<i>Parabacteroides distasonis</i> strain CL06T03C10 chromosome, complete genome

e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP07222 9.1	<i>Parabacteroides merdae</i> strain CL06T03C08 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP07221 0.1	<i>Parabacteroides distasonis</i> strain CL11T00C22 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP04228 5.1	<i>Parabacteroides distasonis</i> strain APCS2/PD chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP06435 3.1	<i>Parabacteroides goldsteinii</i> strain MTS01 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP06943 1.1	<i>Parabacteroides distasonis</i> strain FDAARGOS_1234 chromosome, complete genome
e89c723c1838ebc3c3d955cfe9c6d 0a43d1f9aea	CP00014 0.1	<i>Parabacteroides distasonis</i> ATCC 8503, complete genome
e8cfdd52cc44d84452ff5d895080c 0c1937d0834	AP02556 8.1	<i>Emergencia timonensis</i> CE91-St49 DNA, complete genome
e8cfdd52cc44d84452ff5d895080c 0c1937d0834	AP02556 7.1	<i>Emergencia timonensis</i> CE91-St48 DNA, complete genome
e8e16b6a9530cf26af071a277e13b 5d5f5d7d832	CP04986 0.1	<i>Erysipelothrix</i> sp. HDW6B chromosome, complete genome
e9254a34d202000153d5e58e5036 69404509a85a	CP03528 1.1	<i>Aminipila luticellarii</i> strain JN-18 chromosome, complete genome
e92ba38a1b6bab480d7514d0384f 2ca632c30654	HF56360 9.2	<i>Tepidanaerobacter acetatoxydans</i> Re1 complete genome
e92ba38a1b6bab480d7514d0384f 2ca632c30654	CP00272 8.1	<i>Tepidanaerobacter acetatoxydans</i> Re1, complete genome
e9545bcd0e1bc241dc7cf956fdbdf f34c173888a	AP02556 8.1	<i>Emergencia timonensis</i> CE91-St49 DNA, complete genome

e9545bcd0e1bc241dc7cf956fdbdf f34c173888a	AP02556 7.1	<i>Emergencia timonensis</i> CE91-St48 DNA, complete genome
ea4e04290441277440dc788b928c 6f0d2cf49229	CP04012 1.1	<i>Duncaniella</i> sp. B8 chromosome, complete genome
ea4e04290441277440dc788b928c 6f0d2cf49229	CP03954 7.1	<i>Duncaniella</i> sp. C9 chromosome
eac78eaca91c37efdbce0478fa74bf e2241bcf3a	CP00239 0.1	<i>Filifactor alocis</i> ATCC 35896, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP04361 0.1	<i>Bacteroides fragilis</i> strain GUT04 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP03655 5.1	<i>Bacteroides fragilis</i> strain CCUG4856T chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP03744 0.1	<i>Bacteroides fragilis</i> strain DCMOUH0085B chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP03655 3.1	<i>Bacteroides fragilis</i> strain DCMOUH0067B chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP03655 0.1	<i>Bacteroides fragilis</i> strain DCMOUH0042B chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP03654 6.1	<i>Bacteroides fragilis</i> strain DCMSKEJBY0001B chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP03653 9.1	<i>Bacteroides fragilis</i> strain DCMOUH0017B chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP01893 7.1	<i>Bacteroides fragilis</i> strain Q1F2 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP01270 6.1	<i>Bacteroides fragilis</i> strain S14 chromosome, complete genome

eae919bf0695509970bfbd343fc76 7dc9ee361f4	LN87729 3.1	<i>Bacteroides fragilis</i> genome assembly BFBE1.1, chromosome : scaffold1
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP01107 3.1	<i>Bacteroides fragilis</i> strain BOB25, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP08029 6.1	<i>Bacteroides fragilis</i> strain IHMA_13 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP08029 5.1	<i>Bacteroides fragilis</i> strain IHMA_4 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP08192 2.1	<i>Bacteroides fragilis</i> strain BFG-1 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP07225 7.1	<i>Bacteroides fragilis</i> CL03T12C07 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP07219 8.1	<i>Bacteroides</i> sp. ZJ-18 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP04442 8.1	<i>Bacteroides fragilis</i> strain DK1985 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CP04063 0.1	<i>Bacteroides</i> sp. PHL 2737 chromosome, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	FQ31200 4.1	<i>Bacteroides fragilis</i> 638R genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	CR62692 7.1	<i>Bacteroides fragilis</i> NCTC 9343, complete genome
eae919bf0695509970bfbd343fc76 7dc9ee361f4	AP00684 1.1	<i>Bacteroides fragilis</i> YCH46 DNA, complete genome
ec07eb848603f7dd8afcd603451ab 022ca31efec	AP02341 5.1	<i>Vescimonas fastidiosa</i> MM35 DNA, complete genome

ec1d7382d0de06f5485bd130199a de9d3fa19bbd	LR21598 1.1	<i>Ruthenibacterium lactatiformans</i> isolate <i>Ruthenibacterium lactatiformans</i> 82B1 genome assembly, chromosome: 1
ecc3525845eb554acde7c7b0764b 424334c678cf	CP04883 8.1	[<i>Clostridium</i>] <i>innocuum</i> strain ATCC 14501 chromosome, complete genome
ecc3525845eb554acde7c7b0764b 424334c678cf	CP04883 7.1	[<i>Clostridium</i>] <i>innocuum</i> strain LC-LUMC-CI-001 chromosome, complete genome
ecc3525845eb554acde7c7b0764b 424334c678cf	CP02272 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 genome
ecc3525845eb554acde7c7b0764b 424334c678cf	CP01540 4.2	Erysipelotrichaceae bacterium I46 chromosome, complete genome
ecc3525845eb554acde7c7b0764b 424334c678cf	CP09494 2.1	[<i>Clostridium</i>] <i>innocuum</i> strain VE303-07 chromosome, complete genome
ecc3525845eb554acde7c7b0764b 424334c678cf	AP02556 5.1	[<i>Clostridium</i>] <i>innocuum</i> CE91-St51 DNA, complete genome
ecc3525845eb554acde7c7b0764b 424334c678cf	BK03754 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctS5j6, partial genome
ecc3525845eb554acde7c7b0764b 424334c678cf	CP06532 0.1	[<i>Clostridium</i>] <i>innocuum</i> strain I46 chromosome
ecc89ce10cb14d5e44267db0a488 da802efa9f04	CP02777 5.1	<i>Clostridium botulinum</i> strain MFBjulcb1 chromosome, complete genome
edbee3873a2515b82b47daad17c1f 918d7720f9a	CP08998 1.1	<i>Fulvivirga</i> sp. SS9-22 chromosome, complete genome
edbee3873a2515b82b47daad17c1f 918d7720f9a	CP07060 8.1	<i>Fulvivirga lutea</i> strain S481 chromosome, complete genome
edbee3873a2515b82b47daad17c1f 918d7720f9a	CP05919 1.1	<i>Blattabacterium cuenoti</i> strain PARATEMsp chromosome, complete genome

edbee3873a2515b82b47daad17c1f 918d7720f9a	CP05919 8.1	<i>Blattabacterium cuenoti</i> strain MEGALOSP chromosome
edbee3873a2515b82b47daad17c1f 918d7720f9a	CP05920 9.1	<i>Blattabacterium cuenoti</i> strain DYAKIKur chromosome
edbee3873a2515b82b47daad17c1f 918d7720f9a	CP05921 5.1	<i>Blattabacterium cuenoti</i> strain BEYBkur chromosome
edbee3873a2515b82b47daad17c1f 918d7720f9a	CP00148 7.1	<i>Blattabacterium</i> sp. (<i>Blattella germanica</i>) str. Bge, complete genome
ee17d4a861c18fc528fc32425454d 9496536edf0	BK04008 8.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctUCq5, partial genome
ee2a45dcfe7dda7905ab4bfc43547 c2881306fa1	CP03528 2.1	<i>Tissierella</i> sp. JN-28 chromosome, complete genome
eeef6c2698fdfe979ba3c1305e0b8 c34b39e8dd4	CP04090 2.1	<i>Longibaculum</i> sp. KGMB06250 chromosome, complete genome
eeef6c2698fdfe979ba3c1305e0b8 c34b39e8dd4	CP03183 7.1	<i>Clostridium sporogenes</i> strain YH-Raf3-2-5 chromosome, complete genome
eeef6c2698fdfe979ba3c1305e0b8 c34b39e8dd4	AP01853 7.1	Erysipelotrichaceae bacterium GAM147 DNA, complete genome
eeef6c2698fdfe979ba3c1305e0b8 c34b39e8dd4	AP02408 5.1	<i>Faecalibacillus intestinalis</i> 14EGH31 DNA, complete genome
ef01dd279f05e514bbac07268b073 28f38844823	CP06554 6.1	<i>Lysinibacillus</i> sp. JNUCC-52 chromosome, complete genome
ef91348ef93bd03bde4ff23c4909fd efa84cee13	CP06964 2.1	<i>Escherichia coli</i> strain SSGNw2 chromosome, complete genome
ef91348ef93bd03bde4ff23c4909fd efa84cee13	CP06963 9.1	<i>Escherichia coli</i> strain SSGNw1 chromosome, complete genome

ef91348ef93bd03bde4ff23c4909fd efa84cee13	CP06963 6.1	<i>Escherichia coli</i> strain SSGw2 chromosome, complete genome
ef91348ef93bd03bde4ff23c4909fd efa84cee13	CP06963 3.1	<i>Escherichia coli</i> strain SSGw1 chromosome, complete genome
f0028e407f501e371a48be4d481ca 7e473eda8f1	CP02298 3.1	<i>Cytobacillus kochii</i> strain BDGP4 chromosome, complete genome
f00e15e07e303ca1701ac38d2cae4 c1712bfeaa9	OX09718 0.1	MAG: <i>Candidatus Enterousia intestinigallinarum</i> isolate BRZ_NP__bin10 genome assembly, chromosome: 1
f1c1576ba29e6fa422137d3a1434d 358d5072f7d	CP05870 7.1	MAG: <i>Bacteroidetes</i> bacterium isolate SSA2 chromosome
f1c32fba28ced73721fc7de47ffb2e c41428ea8c	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22l31, partial genome
f1c32fba28ced73721fc7de47ffb2e c41428ea8c	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
f3157dfc9734c8f1c9d8816e634aa d8f5d23d305	GQ87700 5.1	Uncultured organism clone VC1AI32TR genomic sequence
f3515775947a3fa5552ae2971cb3d d7bf4f5c500	CP06063 6.1	<i>Eubacterium hominis</i> strain NSJ-61 chromosome, complete genome
f35b695d1a39664237d01e8963d3 a8b262c954f4	CP03634 8.1	<i>Rosistilla carotiformis</i> strain Poly24 chromosome
f35b695d1a39664237d01e8963d3 a8b262c954f4	CP03631 8.1	<i>Rosistilla oblonga</i> strain Mal33 chromosome
f3f5cd33f4fcea6fe52f2474a9323a 4fcd37aad3	CP06537 7.1	<i>Faecalibacterium prausnitzii</i> strain Fp77 chromosome, complete genome
f3f5cd33f4fcea6fe52f2474a9323a 4fcd37aad3	CP06537 8.1	<i>Faecalibacterium prausnitzii</i> strain Fp45 chromosome, complete genome

f3f5cd33f4fcea6fe52f2474a9323a4fcd37aad3	CP06537 9.1	<i>Faecalibacterium prausnitzii</i> strain Fp40 chromosome, complete genome
f3f5cd33f4fcea6fe52f2474a9323a4fcd37aad3	BK02513 0.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctmsk9, partial genome
f59b2c1cbd5e3bdf7c7e69bfa70e07b802403dca	BK02908 5.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ct22131, partial genome
f59b2c1cbd5e3bdf7c7e69bfa70e07b802403dca	BK03351 6.1	MAG TPA_asm: <i>Siphoviridae</i> sp. isolate ctrf111, partial genome
f67ad93a200363f836ca2cd42acb35bee6921b32	CP00326 1.1	<i>Clostridium pasteurianum</i> BC1, complete genome
f6ab150914ea8192569f515f024802c91bb5955d	CP04583 5.1	<i>Lysinibacillus pakistanensis</i> strain LZH-9 chromosome, complete genome
f6ab150914ea8192569f515f024802c91bb5955d	CP02722 4.1	<i>Lysinibacillus</i> sp. B2A1 chromosome, complete genome
f6c816b8580a91417eaa932f13187babb2dedf7f	CP06722 5.1	<i>Paracoccus</i> sp. MC1862 chromosome, complete genome
f6d893a6b9396ba1524b8a3c69740763372759e7	FJ858269 .1	<i>Micromonas pusilla</i> CCMP1545 chloroplast, partial genome
f718cdeb503bf2e847bde722e7447e01cc0da7b4	CP04925 8.1	<i>Nordella</i> sp. HKS 07 chromosome, complete genome
f718cdeb503bf2e847bde722e7447e01cc0da7b4	OU94086 2.1	uncultured fungus genomic DNA sequence contains 18S rRNA gene, ITS1, 5.8S rRNA gene, ITS2, 28S rRNA gene, isolate consensus02156
f7260f9a51c6d026af00b89f5802c5ea1afe46cf	CP04530 8.1	<i>Legionella longbeachae</i> strain B41211CHC chromosome, complete genome
f7260f9a51c6d026af00b89f5802c5ea1afe46cf	CP04530 6.1	<i>Legionella longbeachae</i> strain B1445CHC chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	AP02283 9.1	<i>Legionella antarctica</i> TUM19329 DNA, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP04861 8.1	<i>Legionella pneumophila</i> strain ERS1305867 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP04597 4.1	<i>Legionella pneumophila</i> strain AUSMDU00010536 isolate SBT211 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP04225 4.1	<i>Legionella longbeachae</i> strain B3526CHC chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02089 4.3	<i>Legionella longbeachae</i> strain F1157CHC chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP04098 7.1	<i>Legionella pneumophila</i> strain FDAARGOS_779 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13417 8.1	<i>Legionella sainthelensi</i> strain NCTC12450 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13417 6.1	<i>Legionella pneumophila</i> strain NCTC12179 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13393 3.1	<i>Legionella pneumophila</i> strain NCTC12180 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02549 1.2	<i>Legionella sainthelensi</i> strain LA01-117 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13438 8.1	<i>Legionella sainthelensi</i> strain NCTC11988 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13438 0.1	<i>Legionella pneumophila</i> subsp. pascullei strain NCTC12273 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13433 2.1	<i>Legionella pneumophila</i> strain NCTC11193 genome assembly, chromosome: 1

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LR13428 6.1	<i>Legionella oakridgensis</i> strain NCTC11531 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LS48341 2.1	<i>Legionella pneumophila</i> subsp. <i>pascullei</i> strain NCTC12272 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LS48341 0.1	<i>Legionella pneumophila</i> strain NCTC11404 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02956 3.1	<i>Legionella anisa</i> isolate UMCG_3A chromosome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01425 5.2	<i>Legionella pneumophila</i> subsp. <i>pascullei</i> strain F-4185 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02128 6.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Albuquerque 1 (D-7474) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02128 3.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Allentown 1 (D-7475) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02128 1.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Flint 2 (D-7477) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02127 9.1	<i>Legionella pneumophila</i> subsp. <i>fraseri</i> strain F-4198 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02127 7.1	<i>Legionella pneumophila</i> subsp. <i>fraseri</i> strain D-4058 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02127 4.1	<i>Legionella pneumophila</i> strain D-4040 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02127 2.1	<i>Legionella pneumophila</i> strain D-5265 chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02127 0.1	<i>Legionella pneumophila</i> strain D-7787 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 9.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Bellingham 1 (D-7473) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 8.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Birmingham 1 (D-7470) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 7.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Burlington 1 (D-7841) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 6.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain Knoxville 1 (D-7468) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 5.1	<i>Legionella pneumophila</i> subsp. <i>fraseri</i> strain Los Angeles 1 (D-7696) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 4.1	<i>Legionella pneumophila</i> subsp. <i>fraseri</i> strain D-5387 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 3.1	<i>Legionella pneumophila</i> subsp. <i>fraseri</i> strain D-3137 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 2.1	<i>Legionella pneumophila</i> subsp. <i>pascullei</i> strain U8W (D-7160) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 1.1	<i>Legionella pneumophila</i> strain NY23 (D-7705) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02126 0.1	<i>Legionella pneumophila</i> strain NY24 (D-7706) chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02125 9.1	<i>Legionella pneumophila</i> strain D-7708 chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02125 8.1	<i>Legionella pneumophila</i> subsp. fraseri strain D-5744 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02125 7.1	<i>Legionella pneumophila</i> subsp. fraseri strain Lansing 3 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02125 6.1	<i>Legionella pneumophila</i> strain D-4954 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP02041 2.2	<i>Legionella longbeachae</i> strain FDAARGOS_201 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01760 2.1	<i>Legionella pneumophila</i> strain D5945 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01760 1.1	<i>Legionella pneumophila</i> strain D6026 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT90647 6.1	<i>Legionella pneumophila</i> strain NCTC11286 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT90645 2.1	<i>Legionella pneumophila</i> strain NCTC11985 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT90644 2.1	<i>Legionella waltersii</i> strain NCTC13017 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01475 9.1	<i>Legionella pneumophila</i> strain F4468 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01374 2.1	<i>Legionella pneumophila</i> strain Philadelphia-1 isolate AE chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01476 0.1	<i>Legionella pneumophila</i> strain F4469 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01745 8.1	<i>Legionella pneumophila</i> subsp. fraseri strain Dallas 1E chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01745 7.1	<i>Legionella pneumophila</i> subsp. <i>fraseri</i> strain Detroit-1 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT63261 7.1	<i>Legionella pneumophila</i> strain ST42 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT63261 6.1	<i>Legionella pneumophila</i> strain ST37 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT63261 5.1	<i>Legionella pneumophila</i> strain ST23 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	LT63261 4.1	<i>Legionella pneumophila</i> strain ST62 genome assembly, chromosome: 1
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01687 6.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain FFI337 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01687 4.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain FFI329, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01687 3.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain FFI105 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01687 2.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain FFI104 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01687 0.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain FFI103, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01686 8.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain FFI102, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01595 6.1	<i>Legionella pneumophila</i> strain E9_O chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01595 5.1	<i>Legionella pneumophila</i> strain E8_O chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01595 4.1	<i>Legionella pneumophila</i> strain E7_O chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01595 3.1	<i>Legionella pneumophila</i> strain E6_N chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01595 1.1	<i>Legionella pneumophila</i> strain E5_N chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01595 0.1	<i>Legionella pneumophila</i> strain E4_N chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01594 9.1	<i>Legionella pneumophila</i> strain E3_N chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01594 7.1	<i>Legionella pneumophila</i> strain E2_N chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01594 6.1	<i>Legionella pneumophila</i> strain E1_P chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01594 5.1	<i>Legionella pneumophila</i> strain C11_O chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01594 4.1	<i>Legionella pneumophila</i> strain C10_S chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01594 1.1	<i>Legionella pneumophila</i> strain C9_S chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 9.1	<i>Legionella pneumophila</i> strain C8_S chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 8.1	<i>Legionella pneumophila</i> strain C7_O chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 7.1	<i>Legionella pneumophila</i> strain C6_S chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 6.1	<i>Legionella pneumophila</i> strain C5_P chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 5.1	<i>Legionella pneumophila</i> strain C4_S chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 4.1	<i>Legionella pneumophila</i> strain C3_O chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 3.1	<i>Legionella pneumophila</i> strain C2_S chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 2.1	<i>Legionella pneumophila</i> strain C1_S chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 1.1	<i>Legionella pneumophila</i> strain Philadelphia_4 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01593 0.1	<i>Legionella pneumophila</i> strain Philadelphia_3 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01592 9.1	<i>Legionella pneumophila</i> strain Philadelphia_2 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01592 8.1	<i>Legionella pneumophila</i> strain Philadelphia_1_CDC chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01592 7.1	<i>Legionella pneumophila</i> strain Philadelphia_1_ATCC chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01592 6.1	<i>Legionella pneumophila</i> strain E11_U chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP01592 5.1	<i>Legionella pneumophila</i> strain E10_P chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP07152 7.1	<i>Legionella lytica</i> strain PCM 2298 chromosome, complete genome

f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP06437 8.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain PartS-Lpneumophila-RM8376 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP09203 8.1	<i>Legionella pneumophila</i> strain 11052018-5 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP08285 2.1	<i>Legionella anisa</i> strain FDAARGOS_1480 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP08284 8.1	<i>Legionella pneumophila</i> strain FDAARGOS_1482 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP08285 0.1	<i>Legionella longbeachae</i> strain FDAARGOS_1481 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP08284 7.1	<i>Legionella pneumophila</i> strain FDAARGOS_1483 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	AP02496 1.1	<i>Legionella pneumophila</i> 80-045 DNA, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP06184 0.1	<i>Legionella pneumophila</i> subsp. <i>pneumophila</i> strain AW-13-4 chromosome, complete genome
f7260f9a51c6d026af00b89f5802c 5ea1afe46cf	CP05940 0.1	<i>Legionella</i> sp. PC1000 chromosome, complete genome
f7f8fe1ea992140d91c0e2dd4b10e 2e084e3070d	OW3704 93.1	<i>Rickettsia endosymbiont</i> of <i>Oedothorax gibbosus</i> strain Oegibbosus-W744x776 genome assembly, chromosome: 1
f8ee4ce8baed2a7650fa71a311520 d6e0799aef5	CP03304 4.1	<i>Mesobacillus foraminis</i> strain Bac44 chromosome, complete genome
fa6155e2fb4bcf2a45b88df2223ce 06e50ed75e6	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome

<i>faa102ecbf06521d53b2404da27ec5859d6fa24</i>	LN87943 0.1	<i>Herbinix</i> sp. SD1D genome assembly SD1D, chromosome : I
<i>fb43cea4f8e8b43b69781580b6a4436d9a105747</i>	AP02573 9.1	<i>Capsulimonas corticalis</i> AX-7 DNA, complete genome
<i>fb43cea4f8e8b43b69781580b6a4436d9a105747</i>	CP06500 3.1	MAG: <i>Acidobacteria</i> bacterium isolate Kalu_18-Q3-R12-55_MAXAC.185_cln chromosome, complete genome
<i>fdccf121db18a6a5ae08f8c24386952999d88f20</i>	CP09172 4.1	MAG: <i>Anaerotruncus</i> sp. isolate nC33_bin.255.fa chromosome
<i>fe43ecc5693aec91dc0f707c3d6974407ae5ce48</i>	CP02334 4.1	<i>Nibricoccus aquaticus</i> strain HZ-65 chromosome, complete genome
<i>fe43ecc5693aec91dc0f707c3d6974407ae5ce48</i>	LK01198 2.1	TPA: <i>Opitutus terrae</i> PB90-1 tRNA-Leu-GAG-1-1 gene
<i>fe43ecc5693aec91dc0f707c3d6974407ae5ce48</i>	CP00103 2.1	<i>Opitutus terrae</i> PB90-1, complete genome
<i>fe4d6fda530303f64101d570f0dc5abc450220e4</i>	CP01422 3.1	<i>Anaerotignum propionicum</i> DSM 1682 strain X2 chromosome, complete genome
<i>fe572b8d62a2c6ca7be6df27733ee4ed1a82941e</i>	CP01577 2.1	<i>Niabella ginsenosidivorans</i> strain BS26 chromosome, complete genome
<i>fe572b8d62a2c6ca7be6df27733ee4ed1a82941e</i>	CP09290 0.1	<i>Candidatus Comchoanobacter bicostacola</i> strain Comchoano-1 chromosome, complete genome
<i>fe572b8d62a2c6ca7be6df27733ee4ed1a82941e</i>	LK01176 1.1	TPA: <i>Nitrosococcus oceani</i> ATCC 19707 tRNA-Leu-CAA-1-1 gene
<i>fe572b8d62a2c6ca7be6df27733ee4ed1a82941e</i>	CP08338 7.1	<i>Arachidicoccus terrestris</i> strain 5GH13-10 chromosome, complete genome
<i>fe572b8d62a2c6ca7be6df27733ee4ed1a82941e</i>	CP00703 5.1	<i>Niabella soli</i> DSM 19437, complete genome

fe572b8d62a2c6ca7be6df27733ee4ed1a82941e	OA93366 6.1	<i>Notodromas monacha</i>
fe572b8d62a2c6ca7be6df27733ee4ed1a82941e	CP00012 7.1	<i>Nitrosococcus oceani</i> ATCC 19707, complete genome
fe6d324c4bac82f84cd8368bc7a590f501a95026	LR99447 3.1	uncultured <i>Prevotellamassilia</i> sp. isolate <i>Prevotellamassilia canine</i> HQ MAG genome assembly, chromosome: 1
fe995a18f29af15ee3a707aa1404b4bcab3be515	CP03122 6.1	<i>Pseudomonas amygdali</i> pv. <i>lachrymans</i> str. M301315 plasmid pMPPla107, complete sequence
ff241e8c8e95610a830e7707df7a03d59745b0a9	CP05166 6.1	<i>Marinobacterium</i> sp. LSUCC0821 chromosome, complete genome
ff241e8c8e95610a830e7707df7a03d59745b0a9	CP08528 4.1	<i>Acinetobacter haemolyticus</i> strain w12 chromosome, complete genome
ff46615c1c6d75137523826f5efa5cec188beda3	CP03122 2.1	<i>Aquirhabdus parva</i> strain HYN0046 chromosome, complete genome
ff4b4506c1517e6b07c2c5cc84a06d3048b8de66	LN87950 2.1	<i>Protochlamydia naegleriophila</i> genome assembly PNK1, chromosome : cPNK
ff4b4506c1517e6b07c2c5cc84a06d3048b8de66	FR87258 0.1	<i>Parachlamydia acanthamoebae</i> UV-7, complete genome
ff8b6f11db0c6f86d0a87cf562a2607b4e3884b1	CP04527 2.1	<i>Priestia megaterium</i> strain FDU301 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a2607b4e3884b1	CP04929 6.1	<i>Priestia megaterium</i> strain S188 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a2607b4e3884b1	CP04769 9.1	<i>Priestia megaterium</i> strain 5-3 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a2607b4e3884b1	CP03252 7.2	<i>Priestia megaterium</i> NCT-2 chromosome, complete genome

ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP03177 6.1	<i>Priestia megaterium</i> strain Ni2-3 chromosome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP03177 7.1	<i>Priestia megaterium</i> strain Mn1-4 chromosome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP04151 9.1	<i>Priestia aryabhattai</i> strain KNU10 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP04106 6.1	<i>Priestia megaterium</i> strain KNU-01 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP03509 4.1	<i>Priestia megaterium</i> NBRC 15308 = ATCC 14581 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP02808 4.1	<i>Priestia megaterium</i> strain SGAir0080 chromosome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP02267 4.1	<i>Priestia megaterium</i> strain SR7 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP02479 4.1	<i>Bacillus</i> sp. Y-01 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP01887 4.1	<i>Priestia megaterium</i> strain JX285 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP01522 6.1	<i>Bacillus</i> sp. IHB B 7164 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP00992 0.1	<i>Bacillus megaterium</i> NBRC 15308 = ATCC 14581, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP08544 3.1	<i>Priestia megaterium</i> strain KF18 plasmid unnamed1, complete sequence
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP08543 1.1	<i>Priestia megaterium</i> strain F20 chromosome, complete genome

ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP08542 6.1	<i>Priestia megaterium</i> strain F16 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP08541 5.1	<i>Priestia megaterium</i> strain CK7 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP09861 0.1	<i>Priestia megaterium</i> strain Ag87 chromosome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP06427 5.1	<i>Priestia megaterium</i> strain B1 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP09238 7.1	<i>Priestia megaterium</i> strain BP01R2 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP05945 7.1	<i>Priestia megaterium</i> strain 2020WEIHUA_L chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP07247 4.1	<i>Priestia aryabhattai</i> strain LAD plasmid p-LAD1, complete sequence
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP07146 6.1	<i>Priestia megaterium</i> strain H2 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP06960 6.1	<i>Priestia megaterium</i> strain CDC 2008724142 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP06928 8.1	<i>Priestia megaterium</i> strain ATCC 14581 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP04888 3.1	<i>Priestia megaterium</i> strain Mes11 chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP05826 2.1	<i>Priestia megaterium</i> strain BIM B-1314D chromosome, complete genome
ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP00198 2.1	<i>Bacillus megaterium</i> DSM319, complete genome

ff8b6f11db0c6f86d0a87cf562a260 7b4e3884b1	CP00198 3.1	<i>Bacillus megaterium</i> QM B1551, complete genome