UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Morphology and genetic relationships in members of the *Solanum nigrum* L. complex used for jam production in the Highveld of South Africa

By

Erika Viljoen

Submitted in partial fulfilment of the requirements for the degree

MAGISTER SCIENTIAE: BIOTECHNOLOGY

In the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

June 2011

© University of Pretoria

Declaration

I, the undersigned, hereby declare that the thesis submitted herewith for the degree Magister Scientiae: Biotechnology to the University of Pretoria contains my own independent work as hitherto not been submitted for any degree at any university or faculty. I further declare that all sources cited or quoted are indicated and acknowledged by means of a comprehensive list of references.

_____

Signature: Erika Viljoen

_____

Date

© University of Pretoria

Table of Contents

Acknowledgements **7**

Preface **8**

Problem statement 9
Aims 10
Hypotheses 10
Outline of Thesis 11

Summary **12**

Abbreviations **14**

Tables and Figures **18**

Chapter 1 Literature Review **23**

1.1 Introduction 24
1.2 Origin of the *Solanum nigrum* complex 25
1.3 Ecology 27
1.4 Taxonomy 27
1.4.1 Morphology 27
1.4.2 Breeding 29
1.5 Molecular analysis 33
1.5.1 Gene regions 33
1.5.2 DNA barcoding 35
1.5.3 DNA marker analysis 39
1.5.4 Flow cytometry 43
1.6 Metabolomics 45
1.6.1 Nutritional content 45

© University of Pretoria

1.6.2 Toxicity 48
1.7 Medicinal properties 49
1.8 *Solanum nigrum* in Africa and South Africa 51

Chapter 2 Plant Material 54

2.1 Background 55
2.2 Farm A study area 55
2.3 Farm B study area 56
2.4 Plant material collected 58
2.5 Greenhouse trials 60
2.6 Herbarium material 60

Chapter 3 Morphological and Cytological Analysis 63

3.1 Introduction 64
3.2 Materials and methods 67
3.2.1 Measurements and photographs 67
3.2.2 Composite analyses 69
3.2.3 Scanning electron microscopy 71
3.2.4 Flow cytometry 71
3.3 Results 72
3.3.1 Whole plant 72
3.3.2 Leaves 81
3.3.3 Flowers 95
3.3.4 Berries 101
3.3.5 Seeds 107
3.3.6 Genomic ploidy 116
3.3.7 Composite analysis 119
3.4 Discussion 121
3.4.1. Whole plant 122
3.4.2. Leaves 122
3.4.3 Flowers 124

© University of Pretoria

3.4.4. Berries 126
3.4.5 Seeds 127
3.4.6. Genomic ploidy 128
3.4.7 Composite analysis 128

Chapter 4 Molecular Analysis 133

4.1 Introduction 134
4.2 Materials and methods 136
4.2.1 Single gene analysis 136
4.2.2 Diversity Arrays Technology 138
4.3 Results 145
4.3.1 Single gene analysis 145
4.3.2 Diversity Array Technology 158
4.4 Discussion 166
4.4.1 Single gene analysis 166
4.4.2 Diversity array technology 170
4.4.3 Concluding remarks 172

Chapter 5 General Discussion, Identification Keys and Conclusions 173

5.1 Cluster descriptions 175
5.2 Identification keys 180
5.3 Conclusion 185
5.4 Future prospects 186

References 187

Appendix A Herbarium Type Specimens 202

Appendix B Buffers, Solutions and Culture Media 208

Appendix C Statistical Analysis of Morphological data 211

© University of Pretoria

Appendix D Multiple Sequence Alignments 225

Appendix E IUPAC Nucleotide Code 243

Appendix F DArT Marker Frequencies 245

© University of Pretoria

Acknowledgements

"Ek is tot alles in staat deur Christus wat my krag gee"

Filippense 4:13

Above all else, thank you to God for His grace and mercy, and blessing me with the wonderful opportunity to complete my MSc degree.

Thank you to my supervisor, Prof Dave Berger, for his excellent guidance and patience. Your willingness to help at any time, and your enthusiasm towards my project is greatly appreciated. To my co-supervisor, Prof Braam van Wyk, a big thank you for the assistance with the intricate details of plant physiology without which the nightshades would still be indistinguishable to my eye. More gratitude towards my co-supervisor, Dr. Antoinette van Schalkwyk, your friendship, character and leadership are three aspects I will carry with me in future.

I am grateful to the EU-SOL project for research funding, without which this project would not have been possible. Thank you to Mr. Leon Strachan and Mr. Pierewiet Conradie for their kind donations of nightshade material from their fields, as well as knowledge and information on their farming practices. I would like to thank Chris van der Merwe, Allan Hall and Andre Botha for their assistance with scanning electron microscopy. Thank you also to Nanette Coetzer, for her valuable time and knowledge of data analysis. A special thank you to Dr. Andrzej Kilian and the rest of the team at Diversity Arrays Technology in Australia, who made my stay a wonderful and exciting adventure. An extra special note of thanks to the members of Lab 6-25, for all the mentoring, friendships formed, advice and support during this study.

And of course a big thank you to my wonderful family. My husband Chris, who has shown an immense amount of support and love through all the hard times, and without whom I would not have had the courage to endure. My parents, Gerhard and Sophia van der Walt, parents-in-law, Frik and Ingrid Viljoen, and my sister, Monica, for their patience, encouragement and enduring affection. Thank you to the rest of my family and friends for their encouragement and support.

© University of Pretoria

Preface

© University of Pretoria

Problem statement

Recently, the importance of supplementing daily diets with nutritional compounds in order to combat the effects of hidden hunger, or deficiencies arising from a lack of certain micronutrients, has become evident (Smith and Eyzaguirre, 2007). Indigenous crops have been identified a source of several micronutrients such as calcium, carotenes, vitamins and iron. South Africa is currently enjoying an increased awareness of the use of native crops in farming practices, as these plants have long been known to be indispensable ingredients in traditional additions to carbohydrate staples (Smith and Eyzaguirre, 2007).

The *Solanum nigrum* L. complex (black nightshades) form part of the Solanaceae family, which includes some of the most important commercial food species such as tomatoes, potatoes and peppers. Morphological data is available for black nightshade species found in European countries, the Americas and Asia, where the potential of these crops as food source was realized many years ago. Several metabolic and genetic studies have also been performed, mainly to obtain a method of effectively distinguishing between the species, as they are morphologically exceedingly similar.

The nightshades have also become increasingly popular as a source of leafy vegetables in the northern parts of South Africa, and the berries for jam production in the Highveld areas. Due to a high level of phenotypic plasticity, the need has arisen for a comprehensive study to determine the identity of species cultivated for these purposes. In this study, the focus will be on the characterization of black nightshades being used for jam production in terms of their morphological traits as well as genetic variability and diversity. With these results, the commercial farmer will have a better understanding of the species most suitable for nightshade jam cultivation.

© University of Pretoria

Aims

The focus of this study was the classification and identification of members of the *Solanum nigrum* complex (black nightshades) that are being grown for jam production in the Highveld region of South Africa, through morphological and molecular analysis. The main aims were to:

- Collect morphologically diverse nightshade plants utilized during the commercial production of jam from two farms located in the Highveld of South Africa (Harrismith and Bothaspas). Authentic material and type specimens of *Solanum retroflexum*, *Solanum chenopodioides*, *Solanum scabrum*, *Solanum americanum* and *Solanum burbankii* were included as reference accessions for each species.
- Record morphological and cytological traits of each plant and compare and portray observations with remaining accessions, herbarium accessions and type specimens in order to demarcate taxa and identify diagnostic characters.
- Infer phylogenetic relationships among accessions based on sequencing of one nuclear and three plastid gene regions.
- Determine genetic relatedness between accessions based on whole genome polymorphisms identified using Diversity Arrays Technology (DArT).
- Propose guidelines that could aid in the positive identification of species utilized for jam production, through the use of identification keys.

Hypotheses

- Morphological differences among accessions utilized during jam production are mirrored in the genetic and cytological comparisons of the accession.
- Plants cultivated on jam producing farms can be clustered into one of four species groups. These species groups are the putatively indigenous S. *retroflexum*, and the naturalized aliens *S. scabrum*, *S. americanum* and *S. chenopodioides*.
- *S. burbankii* is conspecific with *S. retroflexum*.

© University of Pretoria

Outline of Thesis

Chapter 1 gives a comprehensive literature overview on the status of research into black nightshades. The chapter deals with the origin of the species, their ecology, genetic analysis that has been performed to date, their metabolomic aspect and medicinal values. The last section will focus on the economical viability of these plants in African countries, with emphasis on South Africa.

Chapter 2 gives an overview of general materials used. Details on where the plants were collected, their putative identifications and how they were maintained in the greenhouses are discussed. The last section deals with type specimens for some of the scientific names for members of the *Solanum nigrum* complex from herbarium material.

Chapter 3 reports on comparative observations during the morphological and cytological investigations of 30 characters. This includes flow cytometry to determine the ploidy level of each accession.

Chapter 4 deals with genetic and molecular data, including results on the sequencing of four gene regions as well as genotyping outcomes obtained from DArT.

Chapter 5 gives a summary of all the results and arguments presented, and end with a final conclusion of which species had been collected and are therefore used for jam production in the Highveld regions of South Africa. Dichotomous and synoptic keys for the identification of species are presented. The impacts of the main findings of the study are discussed and prospects for future research are highlighted.

© University of Pretoria

Summary

© University of Pretoria

The family Solanaceae consists of many important crop species such as tomatoes, potatoes and peppers. The *Solanum nigrum* complex (black nightshades) also form part of this family and are an emerging, important food source in parts of Africa. Both leaves and ripe fruits are consumed. Additional advantages of the nightshades include several medicinal properties, high nutritional value and natural resistance mechanisms against many disease causing pathogens of the Solanaceae. The objectives of this study were to compare accessions used for "nastergal" jam production in the Highveld of South Africa to reference species of the *S. nigrum* complex using (i) morphological (ii) ploidy and (iii) molecular information. Twenty-nine morphological traits separated the farm-collected samples into three distinct groups. Flow cytometry indicated the presence of a polyploidy series (diploid, tetraploid and hexaploid entities), and the presence of possible hybridization events within the complex. ITS rRNA and three chloroplast genes (*mat*K, *ndh*F and *rbc*L) were sequenced for phylogenetic analysis. ITS yielded 19 informative sites across 600bp, while only six informative sites were identified across 3600bp of chloroplast gene sequences. Chloroplast genes yielded few phylogenetically informative sites, and were not able to distinguish between nightshade accessions. Diversity array technology, a whole genome genotyping approach on microarrays, produced 1145 polymorphic markers between the accessions. ITS rRNA gene sequencing and DArT data produced similar results, with three main groups distinguished, supported by morphological data. Our data suggests that farmers are using fruits of *Solanum retroflexum*, *Solanum chenopodioides* and *Solanum americanum* for jam production. Future work includes measuring levels of vitamin C, proteins, sugars and anthocyanins in the berries, as well as solanine content to shed light on assumptions concerning the toxicity of unripe berries. The results obtained could aid in the identification of germplasm which would be suitable to initiate breeding programs for enhanced physiological or metabolic traits for high quality jam production.

© University of Pretoria

Abbreviations

© University of Pretoria

2C	Amount of DNA in somatic cell
°C	Degrees centigrade
AIV	African indigenous vegetables
AFLP	Amplified Fragment Length Polymorphism
AG1	*S. americanum* group 1
AG2	*S. americanum* group 2
ANOVA	Analysis of Variance
ATP	Adenosine Triphosphate
BAC	Bacterial Artificial Chromosome
BSA	Bovine Serum Albumin
CAPS	Cleaved Amplified Polymorphic Sequences
CI	Consistency Index
cm	centimetre
DArT	Diversity Array Technology
ddH_2O	Double distilled water
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EG	Eglandular trichomes
EtOH	Ethanol
EU-SOL	European Union Solanaceae
FABI	Forestry and Agricultural Biotechnology Institute
g	gram
h	hour
ha	hectare
HI	Homoplasy Index

© University of Pretoria

HIV	Human Immunodeficiency Virus
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITS	Internal Transcribed Spacer region
IUPAC	International Union for Pure and Applied Chemistry
kDa	kilo Dalton
LB	Luria-Bertani broth
LC_{50}	Lethal Concentration 50% death
LG	Long glandular trichomes
LS accessions	Farm A
m	meter
M	molar
matK	maturaseK gene
µg	microgram
min	minutes
mg	milligram
mm	millimeter
mM	millimolar
NEB	New England Biolabs
PC accessions	Farm B
PCA	Principle Component Analysis
PCR	Polymerase Chain Reaction
PI	Propidium Iodide
PIC	Polymorphism Information Content
pg	pictogram
QTL	Quantitative trait loci
RAPD	Random Amplified Polymorphic Sequences

© University of Pretoria

RC	Rescaled Consistency Index
RFLP	Restriction Fragment Length Polymorphism
RI	Retention Index
SDS	Sodium Dodecyl Sulfate
sec	seconds
SEM	Scanning Electron Microscope
SG	Short glandular trichomes
SIBLC	Second International Barcoding of Life Conference
SNP	Single Nucleotide Polymorphism
SSC	Saline Sodium Citrate
SSR	Simple Sequence Repeats
TAU	Trypsin Activity Unit
U	Unit
USA	United States of America
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YAC	Yeast Artificial Chromosome

© University of Pretoria

Tables and Figures

© University of Pretoria

Tables

Table 1.1	General characteristics of black nightshades (Schilling, 1981; Defelice, 2003; Katambo, 2007)	28
Table 1.2	Crossing potential of South African black nightshades (Jacoby and Labuschagne 2006)	31
Table 1.3	Barcoding regions proposed at Second International Barcode of Life Conference	36
Table 1.4	Nutritive value of black nightshades per 100g fresh shoots/leaves	47
Table 2.1	Putative identities of black nightshades (*Solanum* spp.) from South Africa included in the present study	59
Table 2.2	Type specimens for species names of *Solanum* consulted in the present study	61
Table 2.3	Morphological characters of *Solanum nigrum* complex Type species	61
Table 3.1	Characters, character states and assessment techniques	67
Table 3.2	Stem indumentum types	80
Table 3.3	Summary of leaf morphological characters	93
Table 3.4	Leaf indumentum morphology	94
Table 3.5	Summary of floral morphological characters	100
Table 3.6	Summary of berry morphological characters	106
Table 3.7	Summary of seed morphological characters	115
Table 3.8	Averaged flow cytometry values for all nightshade accessions	117
Table 3.9	Putative identities of collective accessions based on morphology	121
Table 3.10	Summary of taxonomically most diagnostic morphological characters across all accessions	130
Table 4.1	Primer properties	137
Table 4.2	Nightshade extended library	139
Table 4.3	DArT nightshade library for polymorphism identification	141
Table 4.4	Hybridization design of 30 nightshade isolates	143
Table 4.5	Sequence statistics for four cloned gene regions	148
Table 4.6	SNP's identified per cloned gene region	149
Table 4.7	SNP's identified from direct sequencing of the ITS gene region	155
Table 4.8	Library producing maximum markers per accession of members of the *Solanum* spp.	162

© University of Pretoria

Table 5.1 Putative and revised identities of *Solanum* accessions 179

Figures

Figure 1.1 Gene regions from three different genomes in plants, most commonly investigated during barcoding studies 37
Figure 1.2 Overview of Diversity Arrays Technology 43
Figure 1.3 General flow cytometry procedure 44
Figure 1.4 Chemical structures of (I) solanine; (II) solasodine; (III) solanidine and (IV) solamargine 48
Figure 1.5 Relative distributions of black nightshade species groups across Africa 52

Figure 2.1 Average rainfall and annual temperature for Harrismith, South Africa. 55
Figure 2.2 Average rainfall and annual temperature for Bothaspas, South Africa. 57

Figure 3.1 Variations in leaf shape within the *Solanum nigrum* complex between seasons (A) and geographical locations (B) 65
Figure 3.2 Measurements of morphological features. 68
Figure 3.3 Diagrammatic representation of a boxplot. 70
Figure 3.4 Morphology of whole nightshade plants. 73
Figure 3.5 Boxplots indicating data distributions around the medium for plant height for each accession. 75
Figure 3.6 Morphology of nightshade plant stems. 76
Figure 3.7 Scanning electron microscope images for types of trichomes. 78
Figure 3.8 Scanning electron microscope images of stem indumentum. 79
Figure 3.9 Examples of leaf shapes and margins. 81
Figure 3.10 Leaf morphology for all nightshade accessions. 82
Figure 3.11 Boxplots indicating data distributions around the median for all accessions pertaining to indentation size. 83
Figure 3.12 Boxplots indicating data distributions around the median for all accessions pertaining to leaf lamina area. 85
Figure 3.13 Boxplots indicating data distributions around the median for all accessions pertaining to the ratio of leaf length: leaf width. 86
Figure 3.14 Boxplots indicating data distributions around the median for all accessions 87

© University of Pretoria

	pertaining to the ratio of leaf length: petiole length.	
Figure 3.15	Scanning electron microscopy images of leaf lamina.	89
Figure 3.16	Boxplots indicating data distributions around the median for all accessions pertaining to density of trichomes on lamina surface (number/mm^2).	91
Figure 3.17	Flower morphology for all nightshade accessions.	96
Figure 3.18	Corolla lobe fusion morphology.	97
Figure 3.19	Boxplots indicating data distributions around the median for all accessions pertaining to the ratio of corolla lobe length: corolla lobe width.	97
Figure 3.20	Style protrusions above anthers.	98
Figure 3.21	Peduncle and pedicel growth	99
Figure 3.22	Berry morphology for all nightshade accessions.	103
Figure 3.23	Boxplots indicating data distributions around the median for all accessions pertaining to berry size.	104
Figure 3.24	Scanning electron microscopy images of selected berry surfaces.	105
Figure 3.25	Seed morphology for all nightshade accessions.	108
Figure 3.26	Boxplots indicating data distributions around the median for all accessions pertaining to seed size.	109
Figure 3.27	Boxplots indicating data distributions around the median for all accessions pertaining to number of seeds per berry.	110
Figure 3.28	Boxplots indicating data distributions around the median for all accessions pertaining to ratio of seed number to berry size.	111
Figure 3.29	Graphical representation of seed relationships.	112
Figure 3.30	Seed SEM morphology for all nightshade accessions.	114
Figure 3.31	Graphical representation of 2C values for each nightshade accession.	118
Figure 3.32	PCA analysis of the most informative morphological character traits.	120
Figure 4.1	Gel electrophoresis following genomic DNA isolations of 25 samples collected from two commercial nightshade farms.	145
Figure 4.2	Gel electrophoresis following PCR amplification of 10 representative accessions from genomic DNA using ITS (A), *matK* (B), *ndhF* (C) and *rbc*L (D) gene specific primers.	146
Figure 4.3	Phylogenetic tree of various accessions of *Solanum* constructed using a combination of ITS, *matK*, *ndhF* and *rbcL* sequences.	151

© University of Pretoria

Figure 4.4	Direct sequencing for accessions of *Solanum* (A) LS7, (B) LS12 and (C) AME indicated conflicting base pair compositions at certain sites.	153
Figure 4.5	DNA sequence alignment for LS7, LS12 and AME haplotypes ("a" and "b") and cloned sequence (R).	154
Figure 4.6	Maximum likelihood analysis of 30 ITS accessions of *Solanum* together with "a" and "b" haplotypes of LS7, LS12 and AME. "R" indicates cloned sequences.	157
Figure 4.7	Marker dissimilarities between reference samples and unknown accessions.	159
Figure 4.8	Marker distribution from libraries.	161
Figure 4.9	Maximum parsimony dendogram indicating relatedness of 30 accessions of *Solanum* based on 1145 DArT marker analysis.	164
Figure 4.10	Maximum parsimony dendogram indicating relatedness of 30 accessions of *Solanum* based on the combination of DArT marker analysis and ITS sequence data.	165

© University of Pretoria

Chapter 1

Literature Review

© University of Pretoria

1.1 Introduction

Recently, native crops as a source of non-nutrient bioactive compounds (e.g. phenolic compounds, tannins) with a range of health promoting and protecting properties have received increased attention. The main reason is due to recent research indicating that fruits and vegetables indigenous to a region could contain the main dietary supplements to accompany carbohydrate staples (Smith and Eyzaguirre, 2007). Researchers in Sub-Saharan Africa have long been investigating leafy vegetables in traditional foods and medicine (Smith and Eyzaguirre, 2007). Africa is known for its large plant genetic diversity, with many indigenous food crops playing an important role in the food security of resource poor farmers. Poverty alleviation, agro biodiversity use and conservation are some of the aspects contributing to the recent interest in African indigenous vegetables (AIV's) (Mwai *et al*., 2007). Studies conducted in East Africa indicated a shift from the usage of alien leafy vegetables (such as cabbage, spinach or kale) to that of AIV's such as members of the *Solanum nigrum* complex and *Brassica oleraceae* (wild cabbage). This is mainly due to their abundant availability and low production costs compared to exotic vegetables (Mwai *et al*., 2007).

One of the economically important plant families produced worldwide is the Solanaceae. This family includes many popular crop species such as tomatoes, potatoes, peppers, paprika, deadly nightshade, tobacco, petunia and eggplant. Within this family, the genus *Solanum* includes between 1500 and 2000 species and contains both annual and perennial plants. The subgenus *Solanum sensu stricto* comprises of 20 recognized sections with section *Solanum* containing approximately 57 species (Edmonds and Chweya, 1997). Among these is the less known *S. nigrum* species complex, or black nightshades. A species complex is defined as group of species that are reproductively isolated from each other, but whose morphology is very similar and in some cases almost indistinguishable (Ridley, 2004). These species are known to be popular fruits and leafy vegetables in Africa, the Americas and South-east Asia (Jacoby and Labuschagne, 2006). *Solanum nigrum* is the type species in the complex as it was used by Linnaeus to describe the section *Solanum* and due to its high occurrence worldwide. It occurs in various habitats throughout the world, from tropical to temperate regions, from sea level to altitudes exceeding 3500 m and from irrigated land to moist environments (Edmonds and Chweya, 1997). *Solanum nigrum* is widely used as cooked leafy vegetables, sources of fruit and for various medicinal purposes (Edmonds and Chweya, 1997; Schippers, 2002). The species belonging to this complex are the second most important group of traditional leafy vegetables in many African countries after the commonly known pigweed, or *Amaranthus* species (Schippers, 2002). They are a reliable source of food and income for resource poor farmers, due to their relatively fast growth rate, even under adverse environmental conditions. Despite

© University of Pretoria

their advantages, the *S. nigrum* complex has received little research and development attention as a potential food crop, in comparison to more popular exotic leafy vegetables.

1.2 Origin of the *Solanum nigrum* complex

The species belonging to the *S. nigrum* complex show different ploidy levels, with diploid (24 chromosomes), tetraploid (48 chromosomes) and hexaploid (72 chromosomes) taxa found on most continents. The highest diversity of diploid species is in the New World, especially in South America, whereas central and North America has distinct morphological diploid species (Edmonds and Chweya, 1997). The highly variable diploid species *Solanum americanum* occurs throughout the Old and the New World, while the tetraploid species *Solanum villosum* and *Solanum retroflexum*, and the hexaploid species *S. nigrum* and *Solanum scabrum* are native to the Old World (Africa and Eurasia). It is noteworthy that in the Old World, the diploid species were introduced, and conversely the tetraploid and hexaploid species were introduced to the Americas and Australasia. However, species of all three ploidy levels occur in South America, and taking into account the morphological variability seen in the complex on this continent it was deduced that this is the origin of species for the section *Solanum* (Edmonds and Chweya, 1997). The hexaploid species *S. nigrum* does not occur naturally in South America, and is predominantly found in Europe and Asia. This species was introduced into other continents and regions, such as North America, Australia and New Zealand. It has by now been well established that *S. nigrum* was derived from the tetraploid *S. villosum* and the diploid *S. americanum* through the amphiploidy of a sterile triploid hybrid, having the complete chromosomal set from the diploid parent (Hawkes *et al*., 1979). Hawkes *et al*. (1979) demonstrated the relative ease by which the species could be synthesized artificially, with the vigorously growing F_2 and F_3 populations being morphologically similar to the naturally occurring species. The backcrossing of the F_3 plants to other accessions of *S. nigrum* resulted in a high percentage fruit set which was comparable to the natural parent accessions. The precise place and time of origin for *S. nigrum* is not known; though it almost certainly arose polytopically and polychronically (hybrids form simultaneously in different geographical areas). The progenitors of *S. nigrum* (*S. villosum* and *S. americanum*) vary extensively in morphology, thus explaining the variability in the triploid hybrids that formed between them. Hawkes *et al*. (1979) also mention that the origin of other tetraploid species is still not deduced. Hybridization work and numerical analysis suggest that the diploid species *S. chenopodioides*, *S. americanum* and *Solanum zahlbruckneri* are the most likely contributors to the South American *Solanum interandinum*, and that both the African tetraploid species *S. retroflexum* and *S. villosum* may have originated from the same or related ancestors (Rao and Kumar, 1981). The polyploid species of the section *Solanum* are thought to be alloploids (having two full sets of chromosomes from

© University of Pretoria

parents of different species), because cytologically most of the species studied so far show regular bivalent formation at meiosis. However, some of these polyploids may also have arisen from the functioning of unreduced gametes or somatic chromosome doubling. Edmonds (1977) recorded both such cases on several occasions where the hybridization of maternal diploids with paternal tetraploids has resulted in tetraploids, instead of the expected triploid progeny. This possession of similar genetic material may explain the similar ranges of morphological variation found in the section *Solanum*.

Many attempts have been made to decide whether a plant species is indigenous or alien to a specific region, and researchers tend to use different sources of evidence to label plants as such. This may lead to confusion in discussions of native or introduced species, which is why it is important to adopt a set of criteria when investigating the status of a certain species. Webb (1985) suggests the use of eight characteristics, and even if the use of only one is seldom completely decisive, when several point in the same direction the composite data can be considered reasonably conclusive. (i) Fossil record, where if a date can be attributed to the finding of a fossil between the last glaciations and the beginning of Neolithic agriculture, is an indication that the species is native. (ii) Historical evidence, (iii) known frequency of distribution and a (iv) record of possible means of introduction often gives an indication whether a plant is native or alien, as chronological details are important in determining where a certain species had originated from initially and when it was first observed in an area. In terms of (v) habitat and (vi) geographical distribution of a species, if a plant is growing remotely in man-made habitats, it is possibly alien. If, however, it grows extensively, and in large capacity in natural habitats it is likely to be native. It should be kept in mind that for some species their natural habitat has disappeared and they are now surviving in man-made habitats (classified as exceptions). (vii) Genetic diversity is an important factor to consider, because it seems reasonable to suppose that if weed populations of a species show obvious genetic differences from small populations found in natural habitats, the respective populations are more likely to be native than if the two populations are identical. This stems from the assumptions that the pressures on a plant to adapt itself to being a weed is more than the pressures of an escaped weed to adapt itself to a natural open habitat. Lastly (viii), it is assumed that most native plants are capable of reproducing by seed, and if a plant reproduces entirely vegetatively it is most probably an introduced species (Webb, 1985). Most literature refers to plants as being alien, introduced or adventives if it does not carry a native status. Adventive usually refers to a species that is not native to the environment, but is generally used to describe accidental introduction, which perseveres for a certain period and then disappears. It is also indiscriminate, and can be applied to species introduced from neighboring ecosystems. Naturalized species are introduced plants that have become established and self-sustaining in a new geographic area. It should be kept in mind that not all introduced species are naturalized, and not all naturalized species are pests (Haselwood and Motter, 1983)

© University of Pretoria

1.3 Ecology

The black nightshades are generally found in disturbed places such as roadsides, at the edges of cultivated fields and plantations, rubbish dumps, in areas around houses and trees, in gardens and on riverbanks (Edmonds and Chweya, 1997). The plants thrive in areas of high rainfall, and in nitrogen and phosphorous rich soils (Edmonds and Chweya, 1997; Van Averbeke *et al.*, 2007a). These species are cultivated on small scale in some countries in Africa and Indonesia, and are mainly utilized as a vegetable or fruit source from random plants in fields or gardens. Where these plants are grown as commodity, the seeds are usually extracted from the fruits by applying pressure to the berries. Seed are washed and dried, and can be stored at room temperature or at 4˚C. The seeds can remain viable for several years, depending on storage conditions and seed moisture content (Edmonds, 1977). Germination generally occurs at temperatures of 15–30˚C, although temperatures fluctuating between these two values break dormancy. Low light intensities favour germination, where intermittent light enhances germination at less favourable temperatures. Van Averbeke *et al.* (2007a) studied the growth and yield response of *S. retroflexum* to nitrogen availability in the soil, and what the effect of the nutrients are on seedlings in pots. The results indicated that optimum application rate was 150 kg N/ha. Excessive N application leads to yield loss, as the extra N can lead to osmotic imbalances in plant tissues. The application of corrected amount of N is important in the growth of *S. retroflexum*, and growers stand to benefit financially if the fertilizer application rates are optimized for production (Van Averbeke *et al.*, 2007a).

1.4 Taxonomy

1.4.1 Morphology

A great deal is known about the benefits of the black nightshades as food and medicine, but knowledge on their taxonomic classification is scarce. The *Solanum nigrum* complex has been the subject of many taxonomic revisions since Linnaeus first described these species in 1732 (Hawkes *et al.*, 1979). One of the first records of nightshades in South Africa is mentioned in Thunberg (1794), where *Solanum nigrum* has reportedly been found in the Paardeberg area, in the now Western Cape near Malmesbury. Whether it definitely was *S. nigrum* is an open debate, as morphological dissimilarities between members of the *S. nigrum* complex were not recognized at the time. Mainly the confusion is due to the phenotypic variation found within this section. Fruit size, leaf shape, flower morphology and plant height seem to be influenced by temperature and light intensity, making morphological classifications difficult (Baylis,

© University of Pretoria

1958; Edmonds, 1977). Floral and vegetative variations among these species are observed; a character may be genetically controlled in one species, but phenotypically plastic in another (Henderson, 1974). Although detailed morphological studies on the black nightshades have been performed in Eurasia (Schilling and Andersen, 1990), Australia (Henderson, 1974), New Zealand (Healy, 1974), North America (Schilling, 1981) and South America (Edmonds, 1972), the African species are still in need of comprehensive taxonomic revision. The black nightshade species generally display the morphological characteristics shown in Table 1.1, with variations occurring among species, or even plants within the same species.

Table 1.1 General characteristics of black nightshades (Schilling, 1981; Defelice, 2003; Katambo, 2007)

Roots	Generally taproot, with roots mainly fibrous
Stems	Cylindrical or angled, with ridges. Simple multicellular hairs with or without glandular heads. Often sparsely to densely pubescent
Leaves	Solitary or in triplet forms, simple and petiolate. Sizes range from 20–80 mm long, and 10–55 mm wide Hairs are glandular or eglandular, with the leaf margins ranging from ovate to lanceolate. Sinuate-dentate margins
Inflorescence	Peduncles simple or branched, usually cymose or cymose-umbellate
Flowers	Calyx bell shaped to stellate, sepal lobes triangular, ovate or lanceolate
Corolla	White or purple, lobes adherent, accrescent or reflexed when flower reaches maturity
Anthers	Vary between yellow and brown. Style straight or hooked, often protruding above the anthers
Fruits	Bi-locular, globose to ovoid, green to purple, black, yellow, orange and red. Black fruits shiny or dull
Seed	Tan, obovate, 1.7–2.4 mm long, flat in cross section, appear in numbers of 15–96 per fruit

Trichomes are hairs that occur on the aerial surfaces of a plant, and distinctions are made between unicellular and multicellular, simple and compound; as well as glandular and non-glandular types (Rodriguez *et al.,* 1984). Glandular trichomes have been found to secrete a wide variety of chemicals, many of which are important in the pesticide, pharmacological and fragrance industries. It is believed that some secretions may aid the plant in natural resistance against pests and diseases (Peter and Shanower,

© University of Pretoria

1998). In the genus *Solanum*, non-glandular as well as glandular trichomes are present on the foliar surface of *S. aculeastrum* (a species used medicinally). Short stalked glandular trichomes are abundant on the adaxial surface of the leaf, while single, multicellular stellate trichomes with 13–15 arms are abundant on the abaxial side of the leaves together with long-stalked glandular trichomes (Koduru *et al.,* 2006). The foliar micro-morphology of *S. pseudocapsicum* (winter cherry) shows short, multicellular glandular trichomes over the entire leaf surface (Aliero *et al.,* 2006). Trichomes have long been an important source of taxonomic evidence in the classification of the angiosperms (Adedeji *et al.,* 2007). They are easy to study, frequently present and often display variation patterns that are taxon specific. Adedeji *et al.,* (2007) reported in their study of trichomes within the Solanaceae, that it is possible to identify plants from different genera or species based on the micro-morphology of several plant organs such as leaves, stems, petals, sepals and petioles.

1.4.2 Breeding

The generally accepted definition of a species was proposed by Mayr (1942). Mayr defined a biological species as "... groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups." He later amended this definition to include an ecological component. In this form of the definition a species is "... a reproductive community of populations (reproductively isolated from others) that occupies a specific niche in nature." Recently, the longstanding controversy associated with species concepts have been reviewed, but no concrete definition of a species has yet been developed (Soltis and Soltis, 2009). Although the general classification of a biological species concept still adheres to the definition given by Mayr in 1942, additional species concepts have been developed. The morphology based taxonomic species concept states that a collection of morphological individuals, which differ from other individuals, based on the same traits are considered separate species groups (Soltis and Soltis, 2009). This is a very subjective system, as the differences observed are not always prescribed in an objective manner by the author, and may be influenced by external factors. The evolutionary species concept recognizes ancestral-descendant sequences of populations that evolve separately, and now occupy their own niches and have their own evolutionary tendencies (Wiley, 1978). All the above concepts have limitations, and classifications become increasingly difficult when hybrids or autoalloploids are produced, and the morphological or molecular traits of these "new" derivative species cannot be classified as distinct from their progenitors (Soltis and Soltis, 2009). Several attempts have been made to classify nightshades into species groups based on their morphology and recently the importance of possible hybrid formations has become evident, making the taxonomy of this species group more difficult.

© University of Pretoria

Jacoby and Labuschagne (2006) performed hybridization studies on the South African black nightshade species to determine if plants with different ploidy levels would produce fertile progeny. The results indicated that although most of the hybridization attempts seemed successful, the berries that formed contained either non-viable or no seed. The cross combinations are shown in Table 1.2. The authors hypothesize that the ploidy level may potentially limit natural hybridizations between the species (Jacoby and Labuschagne, 2006), indicating the separation of plants into different species groups.

© University of Pretoria

Table 1.2 Crossing potential of South African black nightshades (Jacoby and Labuschagne, 2006)

Cross (♀ x ♂)		Nr of berries	Nr of seed	Nr of germination events	Progeny (%)
S. americanum (2n)	*S. burbankii* (4n)	1	2	0	
S. americanum (2n)	*S. chenopodioides* (2n)	6	0	0	
S. americanum (2n)	*S. retroflexum* (4n)	5	0	0	
S. americanum (2n)	*S. scabrum* (6n)	7	250	85	Fertile (90%)
S. burbankii (4n)	*S. americanum* (2n)	2	0	0	
S. burbankii (4n)	*S. chenopodioides* (2n)	6	49	78	Sterile
S. burbankii (4n)	*S. retroflexum* (4n)	5	55	49	Fertile (60%)
S. burbankii (4n)	*S. scabrum* (6n)	2	0	0	
S. chenopodioides (2n)	*S. americanum* (2n)	6	19	0	
S. chenopodioides (2n)	*S. burbankii* (4n)	5	15	38	Sterile
S. chenopodioides (2n)	*S. retroflexum* (4n)	9	121	80	Sterile
S. chenopodioides (2n)	*S. scabrum* (6n)	9	51	10	Fertile (10%)
S. retroflexum (4n)	*S. americanum* (2n)	6	0	0	
S. retroflexum (4n)	*S. chenopodioides* (2n)	4	43	84	Sterile
S. retroflexum (4n)	*S. burbankii* (4n)	6	86	70	Fertile (70%)
S. retroflexum (4n)	*S. scabrum* (6n)	10	0	0	
S. scabrum (6n)	*S. americanum* (2n)	5	51	80	Fertile (80%)
S. scabrum (6n)	*S. burbankii* (4n)	3	0	0	
S. scabrum (6n)	*S. chenopodioides* (2n)	8	0	0	
S. scabrum (6n)	*S. retroflexum* (4n)	3	0	0	

It is well known from literature that members of the *S. nigrum* complex self-fertilize (Edmonds and Chweya, 1997; Schippers, 1998; Kumar and Pushpangadan, 2005; Jacoby and Labuschagne, 2006).

© University of Pretoria

However, nearly all species can be induced to cross in an artificial setting, even if differing ploidy levels are involved (Edmonds, 1979). The method of crossing may be very different in cultivation than it is in nature, but it is clear that several species are capable of interspecific hybridizations, especially among the smaller diploid genome plants. These plants are usually separated geographically, but can be introduced into a mutual setting where crosses take place. If the initial cross is successful, the progeny can also backcross with one of the parent lines, leading to introgression. Several authors have now published their results on the hybridizations of nightshade species, such as Henderson (1974), who has made several crosses between the hexaploid *S. nigrum* and the tetraploid *S. villosum*, as well as between the hexaploid *S. scabrum* and the diploid *S. nodiflorum*. Such crosses are beneficial when desired traits are consolidated in the progeny, if fertile offspring can be produced. Ganapathi and Rao (1986) successfully crossed South African tetraploid *S. retroflexum* and mostly European, hexaploid *S. nigrum*, if the higher ploidy plant is used as the female parent. If the female parent is the smaller chromosomal size, fertility barriers exist in the form of mechanical isolation (Ganapathi and Rao, 1986). Accessions of similar floral size are more readily crossable than those of different floral sizes, as is the case when different ploidy levels are present. The pollen tubes of the hexaploid species could be too thick to enter the style of the relatively slender diploid plants, which leads to ovule abortion before fertilization could take place. In the case where fertilization was successful, the inability of the endosperm to differentiate and function as a nutrient source for the developing embryo could result in abortion of the seed. The cause may have been the action of sterility genes or discordance between parental genomes (caused by incorrect balance of ploidy levels between embryo and endosperm), and would lead to seedless berries and empty seed (Edmonds, 1977).

In the broader Solanaceae, Eijlander and Stiekema (1993) have attempted to cross diploid potato (*Solanum tuberosum*) with the hexaploid *S. nigrum*. Even though *S. nigrum* was used as the female parent, the successful hybrids produced sterile seed. However, when somatic hybrids were produced between the cultivated potato diploid hybrid clone (ZEL-1136) and the hexaploid *S. nigrum* through fusion of protoplasts, the successful transfer of *Phytophthora infestance* resistance genes from *S. nigrum* to potato was observed (Szczerbakowa *et al.,* 2003). Jacoby and Labuschagne (2006) attempted reciprocal crosses between members of the nightshade complex (*S. americanum*, *S. chenopodioides*, *S. burbankii*, *S. retroflexum* and *S. scabrum*) and two cocktail tomato cultivars (SUN and JOS, supplied by a seed company, 'Mayfords'). Their findings show that if the tomatoes are used as the female parent, fertilized flowers resulted in berry sets of which all the seed were sterile. If the tomatoes are used as the male parent, small berries formed which aborted before maturity. These studies suggest that if the plants can be hybridized in an artificial setting (e.g. by protoplast fusions) as opposed to natural hybridizations, the transfer of economically important genes to the progeny could have a major effect in raising the

© University of Pretoria

commercial value of several Solanaceous crops. Traits pertaining to health benefits, medicinal properties, resistance to insect, fungi and bacteria could be traded between species with the genes of interest.

1.5 Molecular analysis

1.5.1 Gene regions

Molecular data has played a significant role in the fields of plant systematics and phylogenetic investigations. In recent years it has become apparent that different molecular tools are needed when investigating different hypotheses, due to varying rates of sequence evolution among genes and gene regions (Small *et al.,* 1998). The ease with which gene sequence data is generated since the introduction of next-generation sequencing platforms (Roche/454 platform and improvements in Sanger sequencing technologies), has led to a dramatic increase in molecular phylogenetic analyses (Gao *et al.,* 2010). The genome wide detection of sequence variation within species groups is the first step when linking genotypic variation and phenotypes. If single nucleotide polymorphisms detected in sequences can be converted into genetic markers, agronomical important crop species can be subjected to cost effective marker assisted selection strategies, whole genome fingerprinting, map based gene cloning and population studies. Next generation sequencing platforms are critical in achieving these goals, allowing for much quicker and cheaper variant discovery and leading to an extensive view of the genome or transcriptome (Deschamps and Cambell, 2010). The possible evolutionary relationships between species can be studied genetically by comparing the DNA sequences of specific genes or intergenic regions. Evolutionary change follows regular patterns enabling mathematical models to formulate these changes and compare DNA from distantly related organisms.

Mitochondrial gene sequences are not frequently used in plant systematics, due to their extremely low rate of nucleotide changes and fast rate of structural evolution (Soltis *et al.,* 1992, Small *et al.,* 1998). Low copy nuclear genes have been suggested as possible regions with high phylogenetic signal, but due to primer design, amplification, sequencing constraints as well as wide variety of nucleotide substitution rates between closely related species, it is generally accepted as a poor choice to sequence for distinguishing species groups (Hughes *et al*., 2006). The popularity of the ribosomal RNA (rRNA) genes and their spacer regions has increased, and may be attributed to the fact that they serve the same basic function in all living organisms and are therefore relatively conserved. In eukaryotes, the small 18S ribosomal unit combines with the large ribosomal unit (5S-5.8S-28S) and together with 49 proteins unite to form the ribosome involved during translation of mRNA into proteins (Mader, 2001). These ribosomal

© University of Pretoria

subunits represent some of the most conserved sequences in eukaryotes, and are a result of strong selection during the evolutionary process against loss-of-function mutations in their components (Poczai and Hyvönen, 2010). The internal transcribed spacers (ITS) are found between the main ribosomal units, and evolve more quickly than the coding regions. These are used for phylogenetic inference, as the variation is suitable to distinguish plants at species and genus levels. Several advantages exist for using the ITS region in conjunction with the less variable coding regions, and these include (i) high copy number; (ii) biparentally inherited, as opposed to the maternally inherited chloroplast genome; (iii) easy PCR amplification, with universal primers available for most species; (iv) smaller size (approximately 500–750bp), allowing direct sequencing without the need of laborious cloning steps; (v) the availability of a model ITS RNA transcript secondary structure, which is a novel method used during plant phylogenetic modelling; (vi) faster evolutionary rate than chloroplast DNA, leading to a higher amount of phylogenetic informative characters available and (vii) incongruence between plant datasets aid in the identification of hybridizations, introgression or lineage sorting (Small *et al.,* 1998, Poczai and Hyvönen, 2010). However, care should be taken when sequencing multigene families such as ITS due to variation among the different repeats in a single genome. Even though the different structures of the ITS within one genome may lead to homogenization of alleles (concerted evolution) where one copy of the sequence will be predominant when sequencing from a PCR mix, cloning from such a mix can give misleading results (Small et al., 1998). Buckler *et al.* (1997) have provided evidence that this variation among ITS sequences are found mostly in organisms that are hybrids or polyploids. Such variation should become evident through direct sequencing, where the differing sites will give signals of two nucleotides on a chromatogram (Poczai and Hyvönen, 2010). In addition, ITS sequences may not distinguish effectively between closely related taxa, especially where recent divergence has taken place (Kress *et al.,* 2005).

Chloroplast DNA has been most commonly used for inferring plant phylogeny, whether it is through restriction enzyme or DNA sequence analysis (Small *et al.,* 1998). The chloroplast is responsible for converting light energy and carbon dioxide into oxygen and carbohydrates in plants (Mader, 2001). There are multiple chloroplast copies per cell, and contains coding as well as non-coding DNA sequences. Advantages of using chloroplast DNA during phylogenetic inference includes (i) high copy number; (ii) relative ease of amplification and sequencing; (iii) uniparentally inherited (primarily maternally); (iv) due to the maternally inherited genome, the chances of paralogues inclusion is reduced and genes can be considered as unique loci; (v) the absence of DNA recombination; (vi) currently, at least 190 complete chloroplast genome sequences are available on the Genbank database, aiding in efficient primer design and (vii) significant species level genetic variability and divergence (Small *et al.,* 1998; Gao *et al.,* 2010, Knoop, 2010).

© University of Pretoria

However, exceptions exist to these generalizations. From literature it has become evident that while plastid genomes are assumed to be maternally inherited in angiosperms and paternally in gymnosperms, exceptions exist where it has been proven that organellar genomes can be inherited both paternally and biparentally (Reboud and Zeyl, 1994; Birky, 1995). If it is assumed that maternal inheritance has taken place, the resulting cell would only show half of the parentage in a polyploid or hybrid plant. In addition, coding regions (used to distinguish at familial level) generally evolve at slower rates than non-coding DNA (used at generic or species level), due to selective constraints. Because of the slower rate of chloroplast evolution, even non-coding regions may fail to distinguish among species (Shinozaki *et al.,* 1986; Maier *et al.,* 1995; Small *et al.,* 1998).

1.5.2 DNA barcoding

DNA barcoding can be defined as a diagnostic technique to identify species based on a short, standardized DNA region (Lahaye *et al*., 2008). DNA barcoding is hoped to be a high-throughput, standardized procedure to effectively recognize and identify eukaryotic species by comparing short sequences from a standard DNA fragment from an unknown accession to a database of reference sequences of known organisms (Shneyer, 2009). An ideal barcode should adhere to the following criteria; (i) easy to obtain suitable primers and amplify; (ii) easily sequenced; (iii) sufficient variance to discriminate between sister species; (iv) have a high confidence that no additional regions will be amplified other than the target region and (v) easily annotated, to ensure sequence errors can be easily detected (Ford *et al*., 2009). It has proven to be difficult to identify such a region, one which has to have such variation to be able to distinguish among species but still retain low variation within species to define inter and intra-specific genetic variations.

Additionally, the use of a single barcode to distinguish among closely related species is further hampered by additional considerations pertaining to species definitions and sequence variations. As there is no definite definition of a species, Spooner (2009) highlights the potential pitfalls when classifying an unknown organism to a specific species name. Most species on earth have not been identified (only about 2 million of the estimated 10 million [Wilson, 2003]), thus a comparison to the database will not necessarily bring forth the correct species name. This may lead to less reliance on barcoding regions, as they would be based on narrow criteria and poorly circumscribed species. It would be of value to test all species within a specific genus, to perceive the species specific nature of the barcode (Spooner, 2009). Many barcoding studies also lack robust assessments of polymorphisms within a species group. It should

© University of Pretoria

be kept in mind that intraspecific and intrapopulational DNA polymorphisms are widespread, especially in populations in the same geographical areas which are still undergoing hybridizations and backcrossings (Spooner, 2009). If and when such barcodes are identified, the advantages are numerous. Firstly, and most importantly, species identifications are essential in several industries and fields of study (Steele and Pires, 2011). Some of the main advantages include the analysis of plant material in forensic sciences to resolve the involvement of a certain plant in a crime or death, restoration ecologists need to positively identify plants suitable for re-establishing vanishing ecosystems, law enforcements need to accurately identify plants to prevent illegal trafficking, protecting native or threatened species in an ecosystem and understanding biological diversity (Steele and Pires, 2011).

In animals, such a region has been identified as a portion of the mitochondrial cytochrome C oxidase gene, but the quest for a universal plant DNA barcode is still ongoing (Lahaye *et al*., 2008). Mitochondrial gene regions were also investigated for plants, but their slow substitution rate had researchers move towards identifying a barcode in the chloroplast gene pool (Hollingsworth, 2008). It is now generally agreed that a multilocus approach based on plastid and nuclear data is the most suitable way to barcode plants (Fazekas *et al*., 2008). Table 1.3 and Figure 1.1 show gene regions proposed at the Second International Barcode of Life Conference (SIBLC) (Hollingsworth, 2008).

Table 1.3 Barcoding regions proposed at Second International Barcode of Life Conference

Research group	Barcoding region proposed
Chase *et al.,* 2007 (Consortium led by Royal Botanic Garden, Kew, UK)	*rpo*C1+*rpo*B+*mat*K OR *rpo*C1+*mat*K+*trn*H-*psb*A
Lee *et al.,* 2007 (Korea University, Korea)	*mat*K+*atp*F-H+*psb*K-I OR *mat*K+*atp*F-H+*trn*H-*psb*A
Kress and Erickson, 2007 (Smithsonian Institute, USA)	*rbc*L+*trn*H-*psb*A

© University of Pretoria

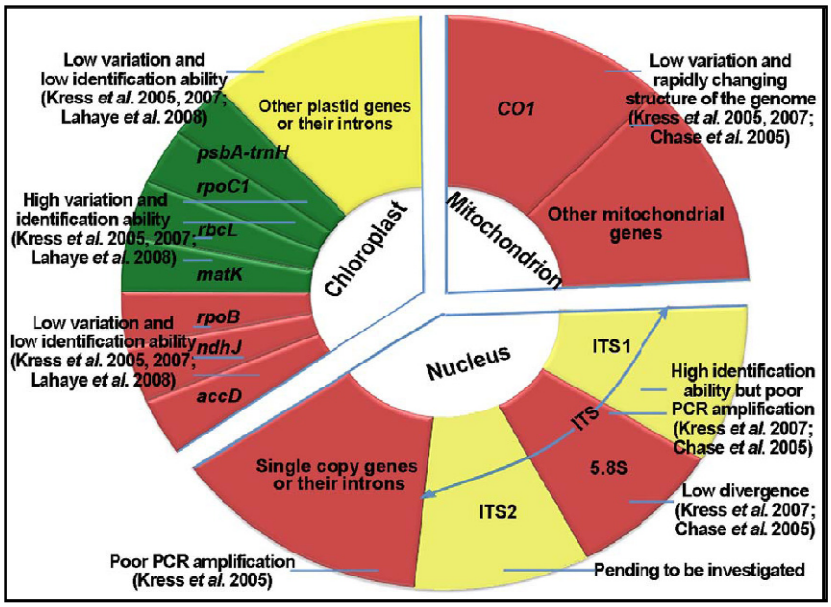


Figure 1.1 Gene regions from three different genomes in plants, most commonly investigated during barcoding studies (adapted from Chen *et al.,* [2010]).

In a recent study by Lahaye *et al.* (2008), the flora of the Kruger National Park in South Africa was examined using eight candidate barcoding regions. The *mat*K gene region (primers adapted from Cuenoud *et al.* [2002]) was the preferred option, as their analysis resulted in species discrimination of > 90%, where multiple individuals of species were resolved as monophyletic. The authors propose to use the *mat*K data in conjunction with sequence data of the *trn*H-*psb*A gene region which showed the highest interspecific divergence of the eight barcoding regions (Hollingsworth, 2008; Lahaye *et al*., 2008). Kress and Erickson (2007) have suggested the use of a region of the *rbc*L gene in conjunction with *trn*H-*psb*A. Even though *rbc*L has not readily been used for species discrimination in the past, it does have the advantage of being easily amplified and sequenced in most land plants (Kress and Erickson, 2007). Liu *et al.* (2010) investigated the suitability of ten different barcoding regions for discriminating among species of Bryophyta (mosses) from northeast China. The proposed regions included *atpF-atpH,* ITS2*, matK, psbK-psbI, rbcL, rpoB, rpoC1, trnH-psbA, rps4* and *trnL-trnF*. After sequencing of more than 50 moss species, the result showed that ITS2 exhibited high inter- and intra- specific divergence, but that amplifying high quality sequences from some taxa was hampered by the presence of paralogous copies. Their results show that *rbcL*, *rpoC1*, *rps4*, *trnH-psbA*, and *trnL-trnF* are the best candidate regions for barcoding mosses. The suggestion of the CBOL Plant Working group (2010) was to use *rbc*L in conjunction with *mat*K, as a mid way for solving individual problems with universality, discrimination, sequence quality and cost. This combination has been shown to successfully discriminate among 907

© University of Pretoria

samples from 550 species with a 72% probability level (Chen *et al*., 2010). This group further suggests using more non-coding plastid gene regions for especially difficult species resolutions, such as *trnH–psbA*, *atpF–atpH*, *psbK–psbI*, and the *trnL* intron together with the fast evolving internal transcribed spacer regions of the nuclear genome, ITS. Barcoding is especially useful when identifications need to be made from small amounts of tissue, sterile, fragmentary or juvenile materials, where conventional morphological identification is not possible (CBOL Plant Working Group, 2010). Some of the most commonly used genes utilized during barcoding initiatives include the following:

Nuclear 5.8S and intergenic spacer regions: In eukaryotes, the 18S, 5.8S and 23S ribosomal genes are separated by ITS-1 and ITS-2. These ITS regions encode signals needed for the processing of rRNA transcripts, and contain sub-repeating units that also serves as enhancers of transcription. One of the main reasons why rDNA is so useful during phylogenetic studies is the varying evolutionary rates associated with the different regions of the rDNA repeats. The 18S, 5.8S and 23S gene regions tend to be more conserved due to their important function in the cell, in contrast to the non-coding ITS regions which have significant variability for phylogenetic analysis (Hillis and Dixon, 1991). Under optimal conditions the regions of DNA used should have between 70 and 100 percent similarity in order to genotype species.

Chloroplast matK gene region: In plant phylogeny, the *matK* gene has emerged as an important gene due to its high phylogenetic signal compared to other genes (Müller *et al.*, 2006). Most green plants have a single copy of the 1500bp *matK* gene, which is nested in the group II intron between the 5' and 3' exon of the *trnK* gene (Barthet and Hilu, 2007). A homology search has indicated that a 102 amino acid carboxyl terminus is structurally similar to portions of maturase-like polypeptides and may be involved in splicing group II introns (Hilu and Liang, 1997). When comparing a data set composed of *matK*, *rbcL* and the *trnL-trnF* sequences, it was observed that *matK* contributes maximum parsimonious information and significantly more phylogenetic structure on average per parsimony informative site than the more commonly used *trnL-trnF* spacer region (Clegg, 1993, Müller *et al.*, 2006). The molecular information obtained through sequencing of the *matK* gene has been used to resolve shallow and deep taxonomic discrepancies and phylogenetic relationships (Johnson and Soltis, 1994; Steele and Vilgalys, 1994).

Chloroplast ndhF gene region: ndhF is a chloroplast gene encoding a component of a putative chloroplast respiratory pathway. In tobacco, this gene is approximately 2223bp in length, which makes it around 50% longer than the more commonly used *rbc*L (ribulose-1,5-bisphosphate carboxylase/oxygenase) gene (Clegg, 1993). This increase in size suggests a higher substitution rate than in *rbcL,* which in turn provides more informative characters that can be used during phylogenetic research (Olmstead and Sweere, 1994). After comparing *ndhF* sequences for tobacco, rice and broad bean, Olmstead and Sweere

© University of Pretoria

(1994) designed various primers spanning different areas of the total gene sequence. The authors further investigated species relationships between members of the Solanaceae family, and found that *ndhF* gave the highest resolution of clades.

rbcL: This gene encodes for the large subunit of the ribulose-1,5-bisphosphate carboxylase (RUBISCO) gene involved in photosynthesis, and extensive variation in synonymous and non-synonymous rates of substitution has been observed for more than 50 sequences spanning bryophytes, conifers, dicots and monocot taxa (Bousquet *et al.,* 1992). In comparison to plant nuclear encoded genes, the rate of nucleotide change was relatively slow for *rbcL*, adding to the effectiveness of this gene being used to infer phylogenies as they are fairly conserved. The conserved rate of substitution has the additional advantage of the design of synthetic oligonucleotide primers that can be applicable to all flowering plant taxa. This differs from nuclear encoded genes in that the chances of DNA-DNA hybridization of similar primers is reduced (Clegg, 1993). Richardson *et al.* (2000) has emphasized the effectiveness of *rbcL* in distinguishing between 90 angiosperm families on supra-generic levels, while inter species relationships among flowering plants have also shown successful separation.

1.5.3 DNA marker analysis

DNA polymorphisms between species can be identified through molecular marker techniques such as restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), simple sequence length polymorphisms (SSR), random amplified polymorphic sequences (RAPD), cleavable amplified polymorphic sequences (CAPS) and single nucleotide polymorphisms (SNP) (Francia *et al*., 2005). It has now been established that SNP's are the most abundant marker type utilized, with the promise of nearly unlimited amounts of markers identified. Although the advantages of high-throughput sequencing and SNP discovery has been well established in humans and some model crops, the high cost associated with assay development and marker discovery has limited the applicability of this technique in 'orphan' crops and polyploid species (Supriya *et al.,* 2011). AFLP is a DNA fingerprinting technique, which is based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA (Vos *et al*., 1995). With this method, sets of restriction fragments can be visualized on an agarose gel without any prior knowledge of DNA sequence. AFLP data can subsequently reveal genetic sequence differences and used to construct high density genetic maps of genome segments. AFLP's can also be used to fingerprint cloned DNA segments like cosmids, bacteriophages, bacterial artificial chromosomes (BAC's) and yeast artificial chromosomes (YAC's) (Vos *et al*., 1995).

© University of Pretoria

AFLP is a popular method for determining the genetic relationships between nightshades, and there are many examples found in literature to support the success of this technique. Previously, African isolates *S. nodiflorum* and *S. americanum* has received increased attention, seen as one species by various authors, while regarded as separate species by others (Edmonds and Chweya, 1997). Morphological inconsistencies have led to the supplementation with genetic data for more accurate classifications, and the lumping of these two groups into one species group, was subsequently not supported by this technique (Manoko *et al.,* 2007). Mwai *et al.* (2007) have investigated 48 vegetative and reproductive characters to delimit, characterize and identify certain *S. nigrum* complex species collected from East, West and South Africa. Nine species were distinguished, and the results indicated a polyploid series, comprising of two diploids (*S. americanum* and *S. chenopodioides*), five tetraploids (*S. retroflexum*, *S. villosum*, *S. florulentum*, *S. grossidentatum* and *S. tarderemotum*) and two hexaploids (*S. nigrum* and *S. scabrum*). In a separate study, Manoko *et al.* (2008) collected black nightshade samples identified as either *S. nigrum* or *S. scabrum* from different African countries including Cameroon, Ghana, Tanzania, South Africa, Uganda, Kenya and Zimbabwe. The objective was to apply a genetic marker technique (AFLP) to study the genetic diversity in *S. scabrum* accessions and its relationship to its geography and morphology, as well as to investigate the possible existence of subspecies within *S. scabrum* and *S. nigrum*. The results showed morphological diversity between *S. scabrum* accessions, but a low genetic variability, which is hypothesized to be due to human bidirectional selection. The AFLP clustering pattern did not segregate the recognized subspecies in *S. nigrum* and *S. scabrum* into separate sub-clusters. In Uganda, Olet (2004) had used morphological characterizations and AFLP analysis to distinguish between species occurring in this country. The 84 species used in this study represented accessions of *S. americanum*, *Solanum florulentum*, *S. scabrum*, *Solanum hirsutum*, *S. nigrum*, *Solanum sarrachoides*, *Solanum tarderemotum* and *S. villosum.* In the results, *S. villosum* (represented by 36 accessions) and *S. americanum* (29 accessions) clustered separately, while *S. sarrachoides* and *S. hirsutum* clustered together, as did *S. scabrum* with *S. nigrum*, and *S. florulentum* with *S. tarderemotum*. The previously eight species recognized in Uganda could be narrowed down to only five, based on combined morphological and molecular data.

The latest study based on morphological characteristics of most commonly found nightshade species in South Africa (*S. burbankii*, *S. scabrum*, *S. americanum*, *S. chenopodioides* and *S. retroflexum*) was performed by Jacoby *et al.,* (2003). AFLPs were used to determine the genetic relationships between the species groups collected. Based on the resulting data, the accessions clearly clustered into two major groups with two sub clusters each and the dendogram obtained through clustering algorithms indicate the underlying genetic differences between the accessions. The first major clade was composed of *S. retroflexum*, *S. americanum* and *S. scabrum* accessions. All the *S. retroflexum* accessions grouped together in one sub-clade, while *S. americanum* and *S. scabrum* formed a sister sub-clade. *S.*

© University of Pretoria

chenopodioides is more distantly related to these groups and form a separate major clade (Jacoby *et al.,* 2003). AFLP's have been successfully employed in phylogenetic studies within the section *Solanum* by Kardolus *et al*. (1998), Edmonds and Chweya (1997), Jacoby *et al.* (2003), Olet *et al*. (2004) and Manoko *et al*. (2007, 2008). However, AFLP is known to be a labour intensive technology, with relatively low genome coverage (Huttner *et al.,* 2005). There are several ongoing investigations into novel marker technologies which aim to overcome the limitations of conventional whole genome genotyping tools.

Marker technologies such as AFLP pose several limitations, the most important being the low amount of polymorphic markers being identified due to limited throughput of gel based systems. This leads to less than optimal genome coverage and marker density. The size of bands (and therefore different alleles) is difficult to assess based on a gel picture, and incorrect interpretations can be made based solely on the sizes of these bands. The identification of new varieties can be hampered further by the sequential process of cloning the polymorphic bands and the sequencing thereof, as errors may occur during the many amplification and sequencing reactions (Huttner *et al*., 2005). Although these limitations can be partially overcome by utilizing specialized hardware such as high-throughput capillary electrophoresis machines to improve allelic discriminations, there are still limitations related to the sequential nature of the technology, as well as the cost. In addition, known DNA sequences are required to expand currently available marker toolkits (Supriya *et al.,* 2011).

Diversity Array Technology (DArT) is a new, microarray hybridization based, genotyping technology. From a pool of genomes, a complexity reduction method is applied where the genomic DNA is cut with a frequent cutter restriction enzyme as well as a rare cutter. This results in a representation of thousands of fragments, which is subsequently cloned. These fragments are usually relatively well distributed across the genome (depending on the specific endonucleases used during digestion) and target areas of expressed genes (Supriya *et al.,* 2011). Random clones are chosen, and the insert produced through amplification is fixed to a microarray slide resulting in a "discovery array". Individual genomes are subjected to the same complexity reduction methods, labelled and hybridized to the array (Figure 1.2). Each genome will give a unique hybridization pattern, and these patterns usually vary between genomes. In this way, it would be possible to identify several hundred genetic markers to be used for genotyping different accessions. Polymorphic markers are categorized by identifying clones which show variable hybridization intensities between specific individuals. These clones are re-assembled into a "genotyping array", for routine genotyping of additional individuals from that germplasm (Huttner *et al.,* 2005). Applications of DArT are numerous, the most important being the identification of hundreds of polymorphic markers with one experiment. In addition, all types of DNA variation can be identified, from single nucleotide polymorphisms to copy number variations and methylation status (Howard *et al.,* 2011). Comprehensive

© University of Pretoria

genomic profiles are easily assembled, regardless of the amount of information available for that crop. The time plant breeders spend on identifying specific Quantitative Trait Loci (QTL's) is significantly reduced, and the introgression of a specific region of interest into an elite background may happen in a shorter amount of time, and with less effort (Jaccoud *et al.,* 2001). This is a sequence independent technology, which makes it applicable to most plant and fungi species with limited DNA sequence available. This technology has several applications such as genome profiling, genome background screening, rapid construction of genetic linkage maps, identification of quantitative trait loci (QTL), simultaneous marker-assisted selection for several traits and evaluation of genetic diversity (www.diversityarrays.com). In contrast to other marker technologies, a sufficiently large amount of markers to cover most of the genome can be identified for genotyping. In DArT, the genomes of several accessions are combined for marker identification, and cover the diversity across the whole species group. DArT is routinely used to determine genetic diversity within a population, but markers do not give an indication of phylogenetic relatedness. Marker sequences span several different genes with varying evolutionary rates, and a molecular clock can therefore not be applied. However, coupled with recent advanced in next-generation sequencing, it would be possible to identify polymorphic markers within certain genes of which the evolutionary rate is known that can be used for phylogenetic systematics.

To date, DArT has been used to genotype and create genetic linkage maps of several varieties of species, including sorghum (Mace *et al* 2009); wheat (Semagn *et al*., 2006), barley (Wenzl *et al.,* 2004), pigeonpea (Yang *et al.,* 2006), cassava (Xia *et al.,* 2005), apple (Soriano *et al.,* 2009) and *Arabidopsis thaliana* (Wittenberg *et al.,* 2005). Species with haploid, diploid and tetraploid chromosome numbers have also been used successfully to elucidate genetic relatedness, as was the case for banana accessions (Amorim *et al.,* 2009). James *et al* (2008) has successfully applied DArT to the evolutionary studies of non-model organisms, *Asplenium viride* (a diploid fern) and *Garovaglia elegans* (a haploid moss). Their results show that DArT can be effectively utilized to detect sufficient genomic variation to be useful across a large taxonomic spectrum. DArT holds real potential to detect biological correlations, hybridizations, mixed environmental samples, geographical distributions, gene discovery and QTL mapping for diverse species groups (James *et al.,* 2008). An additional advantage is the ability to sequence the DArT clones generated for the discovery array. This information can be used to convert the amplified products into PCR-based markers used in basic screening of genes, and for linkage group alignment with other species (Supriya *et al.,* 2011).

© University of Pretoria

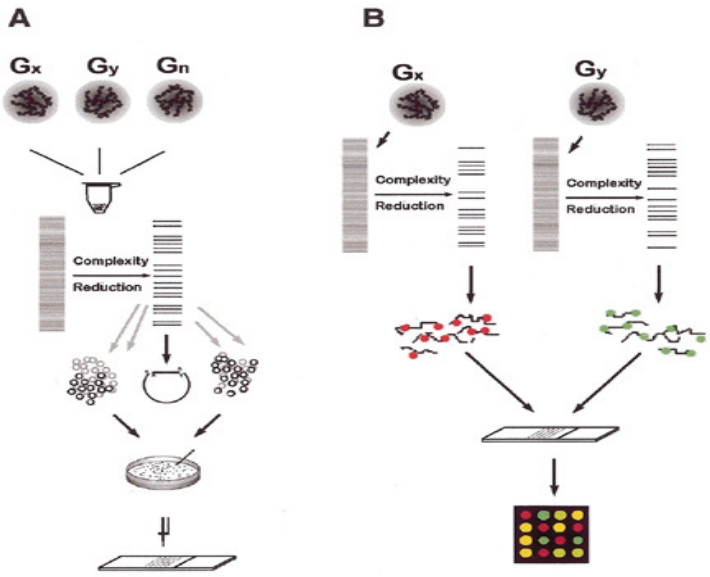


Figure 1.2 Overview of Diversity Arrays Technology. (A) Generation of diversity array from pooled genomic DNA representing a specific germplasm. The complexity of fragments is reduced and the fragments are cloned. Colonies are randomly selected, insert amplification is performed and targets are purified. The targets are fixed onto a solid support slide. (B) Two contrasting genomes used in DArT. Individual genomes are subjected to the same complexity reduction methods as in (A). Each genomic representation is labelled with either a green or red fluorescent dye and hybridized to the array. A difference in hybridization intensities for each feature (and thus the representative genomic fragment) indicates a polymorphic fragment between the two genomes (Adapted from Jaccoud *et al.,* 2001).

1.5.4 Flow cytometry

Flow cytometry can be broadly defined as a system which measures and then analyzes the signals that are produced when particles flow in a liquid stream through a beam of light, such as DNA molecules (Doležel *et al*., 2007). This technology is not limited to cell analysis, but also deals with chromosomes, latex beads and any other particles that are suspended in liquid. Flow cytometry had its origin in the mid-1960s when Louis Kamentsky in New York (USA) developed a microscope-based spectrophotometer which measured and recorded UV absorption and the scatter of blue light from cells which passed a microscope objectives at rates greater than 500 cells per second (Givan, 1992). It is possible to permeabilize the cell wall with a type of alcohol, after which propidium iodide can enter the nuclei of cells and intercalate between DNA

© University of Pretoria

bases with no sequence preference. The red fluorescent intensity obtained from the nuclei is more or less proportional to the amount of DNA present in a cell. Flow cytometric analysis of the red fluorescence from the nuclei will result in a single, very narrow peak on a histogram because nuclei emit nearly the same amount of red fluorescence. Cells with an unusual amount of DNA will lead to histograms with more than one peak, indicating cells with more or less DNA than would normally be expected. Flow cytometry has also been successfully utilized in the analysis of chromosomes. Particles flowing through the system are chromosomes released from cells during metaphase. These released chromosomes are stained (with e.g. propidium iodide) and sent through the flow cytometer (Figure 1.3). The resulting histograms of fluorescence intensity reveal peaks whose position along the x-axis is proportional to the amount of DNA in the chromosome and whose area is proportional to the number of chromosomes with that particular amount of DNA (Givan, 1992).

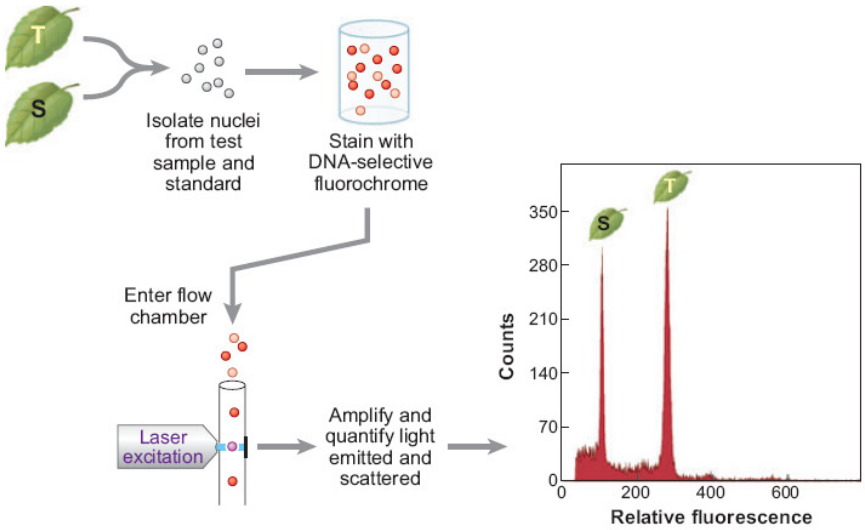


Figure 1.3 General flow cytometry procedure. Nuclei or cells are isolated from leaves of a tested individual (T) and an individual to be used as a standard (S). DNA is stained with a DNA-selective fluorochrome and the cells in suspension are passed individually past a laser beam. The light scatter and fluorescence is measured, and the output is usually in the form of a scatterplot or histogram (Kron *et al.,* 2007).

Advantages of flow cytometry include easy sample preparation, rapid analysis (~500–1000 nuclei/second) and the material need not be actively dividing when analyzed (Kron *et al*., 2007). On a taxonomic level, flow cytometry has the advantage of expanding the amount of characters available for taxonomic

© University of Pretoria

analysis, most notably DNA content (ploidy and genome size) and to a lesser extent base pair composition (AT/GC ratio). Consequently, significant developments in systematics attributable to flow cytometry include (i) improved detection and delineation of species (detection of polyploidy, origin of polyploidy, information on hybrid states), (ii) improved data which can be related to the input of chromosomal speciation to species diversity, and (iii) increased character availability to infer phylogenetic relationships (Kron *et al*., 2007) This technology is useful in determining the ploidy levels of unknown nightshades, aiding in species classifications. Many studies on ploidy levels have been conducted for different species groups, e.g. potatoes (Jiménez *et al*., 2008), roses (Roberts *et al*., 2009), *Passiflora* spp. (Souza *et al*., 2008) and *Chenopodium* spp. (Palomino *et al*., 2008) as well as in the Solanaceae (Szczerbakowa *et al.,* 2003).

1.6 Metabolomics

1.6.1 Nutritional content

Approximately 10% of the world's population lives in sub-Saharan Africa, where one in three people are chronically hungry. Natural disasters, poverty, low employment rate and the escalating incidence of the human immunodeficiency virus (HIV) are factors which influence the livelihood of these rural people (UNAIDS, 2007). An estimated 33.3% (or 190 million) of children under the age of 5 years are at risk of vitamin A deficiency worldwide. In Africa and South-East Asia the highest numbers occur, with 44.4% and 49.9% of children respectively suffering from vitamin A deficiencies (WHO, 2009). The importance of indigenous leafy vegetables cannot be underestimated, as these are often the only source of nutritional food, and essential micronutrients, for many of these people. The patterns of consumption of these vegetables differ among households in different countries. In South Africa, this pattern is influenced by several factors such as poverty, distance to fresh produce markets, degree of urbanization and the season of the year (Jansen van Rensburg *et al.,* 2007). Several studies have shown micronutrient status of theses indigenous vegetables are as high as or higher than conventional leafy vegetables such as cabbage or spinach (Edmonds and Chweya, 1997). Many people in sub-Saharan Africa depend on staple foods prepared from cereal crops (millet, sorghum and maize), tubers (yam, cassava) and plantains (a vegetable similar to bananas) to fulfil their daily requirements of micronutrients (Oniang'o *et al.,* 2003). However, these crops were found to not meet the WHO/FAO recommended daily intake, and for most countries the intake did not reach half the recommended intake (Ruel *et al.,* 2005). A diversified diet is needed to fulfil the requirement of all micronutrients, and a number of indigenous African vegetable species have been

© University of Pretoria

found to have the potential to contribute significantly to helping to alleviate chronic problems associated with the lack of micronutrients.

Very few studies have been performed where the metabolite content of *S. nigrum* leaves, berries or roots were investigated. Edmonds and Chweya (1997) state that these plants are nutritious vegetables, where the leaves provide appreciable amounts of protein, amino acids, minerals including calcium, iron and phosphorous, vitamins A and C, fat and fiber as well as a scarce vegetable amino acid, methionine. The nutritional content varies with the age of the crop, the species, whether it was grown during the rainy or dry season, the soil fertility and preparation method (Schippers, 1998). Mudzwiri (2007) indicated that the levels of antinutrients found in leaves of *S. retroflexum* species (collected in KwaZulu Natal, South Africa) to be as follows: phytic acid (0.04 mg/ml), saponins (1.33 mg/ml), oxalic acid (798 mg/ml), cyanogenic glycosides (undetected), alkaloids (undetected) and trypsin inhibitors (4045 Trypsin Activity Units [TAU]). The range of nutritive values in *S. nigrum* leaves (adapted from Edmonds and Chweya, 1997 and Uusiko *et al.,* 2010) are given in Table 1.4. *Solanum nigrum* leaves are high in moisture content (83–91%), which is comparable to the 94% observed for the exotic leafy vegetable *Spinacea oleracea* (spinach). Vitamin A and C concentrations were found to be 1.7–11.6 mg/100g fresh tissue and 20–158 mg/100g fresh tissue respectively, and this is more than what had been determined for spinach (4 mg/100g vitamin A and 36 mg/100g vitamin C). Although spinach contains more of the mineral zinc (6 mg) in comparison to black nightshade (3.5 mg), the latter again has the advantage of being high in iron content (4.3 mg) as opposed to 3.5 mg in spinach (Singh *et al.,* 2001).

The berries of members of the *S. nigrum* complex contain considerable amounts of vitamin C, iron, calcium, carotenes and vitamin B (Edmonds and Chweya, 1997). In addition, it was recently found that these berries contain appreciable amounts of vitamin A. Mongolian gerbils were vitamin A depleted for 5 weeks, but retained normal vitamin A levels through the bioconversion of β-carotene to retinol, after being fed small quantities of *S. nigrum* berries (Ejoh *et al.,* 2010). Using a chloroform extraction method, Loganayaki *et al.* (2010) obtained a yield of approximately 500 mg/100 g fresh weight phenolic content in the berries of *S. nigrum,* with the implication that the fruits of this plant could be used as a natural supplement of antioxidants in the human diet. The phenolic content was found to be comparable to other crops such as blackberry (361-555 mg/100 g fresh weight), blueberry (181-473 mg/100 g fresh weight), raspberry (114-517 mg/100 g fresh weight) and strawberry (102-443 mg/100 g fresh weight) (Vasco, 2009). Other crops include garlic (2 mg/100 g fresh weight), grapes (0.5 mg/100 g fresh weight) (Oviasogie et al., 2009); eggplant (106 mg/100g fresh weight) (Nisha *et al.,* 2009); potatoes (5 mg/100 g fresh weight) (Taskeen *et al.,* 2011); tomato (25-35 mg/100 g fresh weight) (Luthria *et al.,* 2006).

© University of Pretoria

Table 1.4 Nutritive value of black nightshades per 100 g fresh shoots/leaves

Nutrient	Measure	Range of values (Edmonds and Chweya, 1997)	Range of values (Uusiko *et al.*, 2010)
Water	%	83–91	83–90
Crude protein	g	2.8–5.8	3–5
Crude fiber	g	0.6–1.4	1
Fat	g	0.8	0.6
Carbohydrate	g	3.3–5.0	2
Calories	kcal	38	55–58
Etheral extract	g	38–44	-
Total Ash	g	3.3–8.8	-
Iron	mg	1.0–4.3	8.5–12.8
Calcium	mg	90–442	278–310
Phosphorous	mg	75	-
Beta carotene	mg	1.7–11.6	-
Ascorbic acid	mg	20–158	2
Oxalate	mg	58.8–98.5	-
Nitrate-N	mg	29–400	-
Vitamin A	mg	-	1070
Folate	ug	-	404
Magnesium	mg	-	84
Zinc	mg	-	0.8–3.5
Antioxidant activity	%	-	78–92

© University of Pretoria

1.6.2 Toxicity

In addition to the nutritional benefits, it is known that members of the Solanaceae contain various toxins. The amount of toxin in the plant varies between species, part of the plant, environment and stage of growth (Defelice, 2003). The steroidal alkaloids of the Solanaceae fall into two groups. The first is the solanidine type (characterized by an indolizidine moiety) and the second is the spirosolane type (where the nitrogen atom is part of an oxo-azaspirodecane structure). They occur as glycosides where the sugar is most often an oligosaccharide (Bruneton, 1995). These steroidal glycoalkaloids are an important group of secondary metabolites found in *Solanum* species, and the major glycosides are solanine and solamargine (Figure 1.4). It is known that these compounds can be toxic or inhibitory to many organisms (Eltayeb *et al.*, 1997). The main mechanism, by which these compounds work, is thought to be opening of the inner and outer membrane channels of the mitochondria, leading to a release of Ca^{2+} against the membrane potential. The elevated levels of Ca^{2+} within the cells trigger the mechanism and occurrence of apoptosis (Gao *et al.,* 2006). After a few hours latency, the intoxication symptoms in humans include nausea, headaches, vomiting, diarrhoea, hallucination, nervousness and convulsions (Bruneton, 1995), as well as loss of speech, fever, sweating, reduced heartbeat, tachycardia, blindness, mental confusion, coma, and if large quantities are consumed, death.

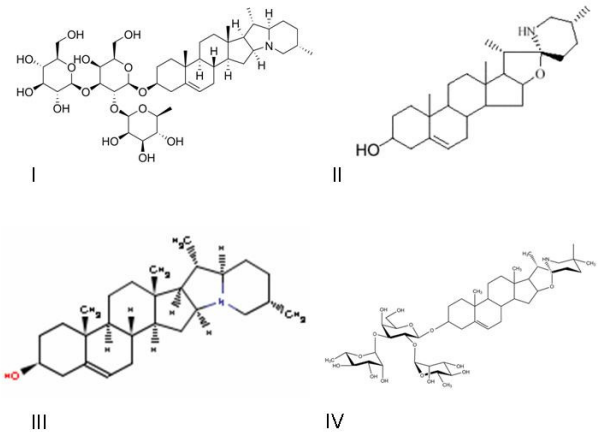


Figure 1.4 Chemical structures of (I) solanine (www.answers.com)**; (II) solasodine** (www.cyberlipid.org)**; (III) solanidine** (www.mdidea.com) **and (IV) solamargine** (www.glycomix.co.uk).

Eltayeb *et al.* (1997) conducted a study where the changes in solasodine profiles during the development of *S. nigrum* were monitored across the whole plant. Their findings show that the highest amount of solasodine occurs in the young leaves and the concentration decline during leaf development. The small,

© University of Pretoria

unripe fruits of *S. nigrum* contain a high concentration of solasodine, but the concentration declines as the fruit mature. Clearly, consumption of unripe fruits or young leaves can lead to poisoning or death due to the increased concentration of this glycoalkaloid. The ripe, fully matured black nightshade berries pose a smaller risk of solasodine poisoning (due to lower concentrations), and they are popularly used in pies and preserves in the United States as well as in most parts of South Africa (Watt and Breyer-Brandwijk, 1962, Defelice, 2003).

Many compounds are toxic if ingested in large quantities, but *in vivo* lethality tests in a simple zoological organism can be used as a rapid and simple pre-screen of the toxicity level of a certain compound concentration (McLaughlin *et al*., 1993). The eggs of the brine shrimp can be purchased at a low cost, and remain viable for a number of years when in a dry state. If placed in a brine solution, the eggs hatch within 48 hours, and different concentrations of the potentially toxic substance is applied to the many newly hatched nauplii. The lethality data is processed to estimate the rate of death of the shrimps, and can be used to determine the LC_{50} values with 95% confidence intervals to obtain statistically significant potency values (MacLaughlin *et al*., 1993). Mudzwiri (2007) performed toxicity tests on aqueous and organic extracts from *S. retroflexum* leaves (collected in KwaZulu Natal, South Africa) using the brine shrimp assay. The findings indicated that aqueous extracts made from this species is non-toxic to the brine shrimps (less than 50% shrimp died). Aqueous extracts are therefore not poisonous, which is beneficial to consumers who boil the leaves in water for consumption. At a concentration of 1 mg/ml organic extraction however, almost exactly 50% of shrimps died. The authors speculate that even though toxicity is assumed if an extract caused more than 50% brine shrimp deaths, and the *S. retroflexum* concentration was just below the threshold, this species should still be handled cautiously (Mudzwiri, 2007).

1.7 Medicinal properties

Plants synthesize primary and secondary metabolites which function as chemical compounds in normal metabolic activities. The primary metabolites, such as sugars and fats, are found in all plants and constitute basal cell functions. Important secondary compounds not involved in basal cell functions such as alkaloids and terpenes are most often found only in a specific genus or species. These compounds could be involved in a plant defence strategies, or survivability of that plant. The functions of secondary metabolites include defence against microorganisms and attack by herbivores (Hopkins and Hüner, 2004). Examples of secondary metabolites include alkaloids, glycosides, terpenes and phenolic compounds.

© University of Pretoria

Alkaloids and glycosides occur in several black nightshade species (Eltayeb *et al*., 1997), and may be the key to their use in medicinal medicine.

Rural villagers in many countries often do not have access to health care practices in their region, due to distance or cost of medical attention. Therefore, they have returned to knowledge accumulated by their forefathers on plant remedies. In China, leaves and berry extracts of *S. nigrum* have been used for treating fever and in India for treating gonorrhea and urinary infections. Local healers in Israel employ members of the *S. nigrum* complex to treat a wide variety of ailments, including wounds, toothache, swellings, burns, constipation, heart and liver disorders; extracts is also used as a sedative (Defelice, 2003). In Kenya, it is proposed that pregnant women who eat boiled leaves of *S. nigrum* will have a shortened labour and that post-birth recovery will be significantly faster (Edmonds and Chweya, 1997). Other medicinal uses include a remedy for convulsions, a sleeping agent (Europe), raw fruit is chewed and swallowed for treatment of stomach ulcers or general abdominal upsets (East Africa). The Zulu in South Africa use an infusion as an enema for abdominal upsets in children while the Xhosa squeezes the juice of unripe fruits onto teeth and gums of infants to reduce the pain during teething (Juma, 2006). Additional findings include the potential of *S. nigrum* to have anti-seizure activity (Wannang *et al.,* 2008). A study by Lee and Lim (2006) showed that aqueous leaf extracts of members of the *S. nigrum* complex produces significant dose dependent protection against electrically-induced seizure in chicks and rats, pentylenetetrazole-induced seizure in mice and rats and picrotoxin-induced seizure in mice and rats. Results obtained in this study suggest that the leaves of this plant may possess anti-convulsant properties in chicks, mice and rats. *S. nigrum* has traditionally been used as a remedy in cancer treatment and a diuretic and antipyretic in Korea and China (Lee and Lim, 2006). Recently, a 150 kDa glycoprotein has been isolated from a *S. nigrum* accession collected in South Korea, which consists of a carbohydrate content of 69% and a protein content of 30%, with pharmacological activities such as anti-tumor, anti-cancer and antioxidative effects (Lee and Lim, 2006, Loganayaki *et al.,* 2010). An ethanol extract from dried *S. nigrum* berries was fed to rats, and the results indicated an effect on general behaviour patterns, suppressed aggressive behaviour and affected locomotive activity. These observations suggest that fruits of *S. nigrum* plants can have a central nervous system depressant action (Perez *et al.,* 1998).

Various parts of plants from the *S. nigrum* complex are known to have inhibitory properties to fungi, bacteria and viruses. *Penicillium italicum* is a fungus known to cause post-harvest disease in most types of citrus fruit. A recent study by Kanan and Al-Najar (2009) show a complete inhibition of growth against this fungus, from a hexane extraction from *S. nigrum* fruits at very low IC_{50} values of 37–80 µg/ml. Plants from this complex is known to have inhibitory actions against other fungi such as *Altenaria brassicicola* (Muto *et al.,* 2006) and *Phytophthora infestans* (Lebecka, 2009), bacteria such as *Klepsella*

© University of Pretoria

pneumoniae (Kannabiran *et al.,* 2009), *Aeromonas hydrophila* (Rajendiran *et al.,* 2008) and *Vibrio cholereae* (Acharyya *et al*., 2009). Edmonds and Chweya (1997) report activity against *Staphyllococcus aurens* and *Escherichia coli*, as well as antiviral, larvicidal, molluscicidal and paraciticidal properties.

1.8 *Solanum nigrum* in Africa and South Africa

Few studies have been conducted on the diversity of black nightshades in Africa (Edmonds, 1977; Jacoby and Labuschagne, 2006; Manoko *et al.*, 2007; Mwai *et al.,* 2007), although the distributions of different species groups have been well documented (Edmonds and Chweya, 1997; Figure 1.5 of this thesis).

The nightshades are being cultivated in several highland communities in Africa, and somewhat less in the Central and West African lowlands like Democratic Republic of Congo. However, when moving north into Africa, in Cameroon it is a popular cash crop, while in southern Nigeria it is an important subsistence crop. In Cameroon, the yields of nightshade plants are between 7 and 13 tons/ha for the first harvest, 12–20 tons/ha for the second harvest with the yields decreasing with every harvest thereafter. A farmer can obtain an average of 50 tons/ha in a growing season if correct management practices are applied. When the bundles of nightshade leaves are sold on the local market, the farmer will set a reasonable price per bag of dried leaves. Nightshades are grown on a much more limited scale in the eastern parts of Africa, such as in southern Uganda, and parts of Kenya (Schippers, 1998). Three varieties of nightshade are grown on a commercial basis in Kenya, but proper species identification of these has not been completed. These plants are mostly grown during June, July and August, and leaves are harvested fortnightly (Onyang'o, 1993). Several studies on mode of propagation, effects of plant density and responses to fertilizer application and deflowering have been conducted. General results show propagation from seed give the highest nutritional value compared to plants regenerated using shoot cuttings, and that a higher density of plant growth leads to an overall increase in plant yield (Mwafusi, 1992). The application of 5 g N/plant also significantly increased beta-carotene, crude protein, nitrates and phenolic leaf contents, at the same time leading to increased amount of leaf yield per plant. By deflowering the plants the energy spent on reproductive organs is re-directed to vegetative organs, again leading to significant increases in leaf yield (Mwafusi, 1992; Onyang'o, 1993).

© University of Pretoria

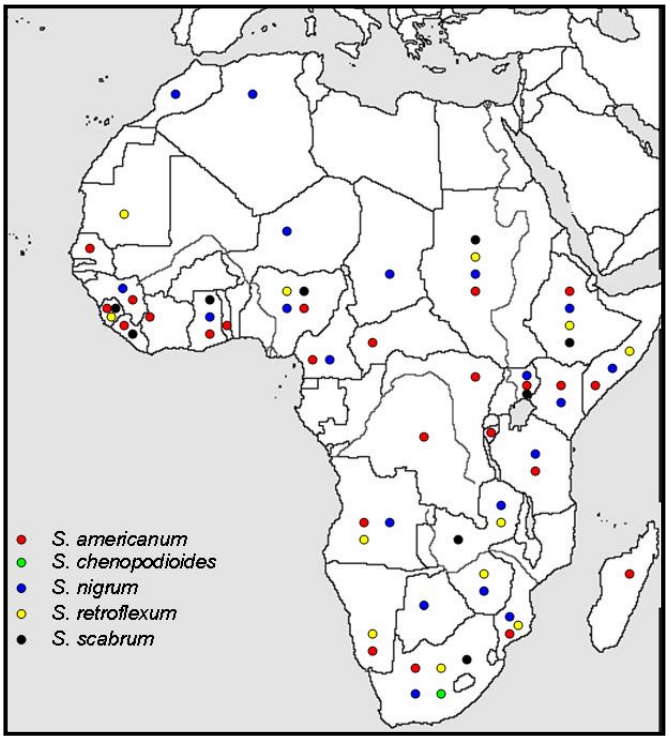


Figure 1.5 Relative distributions of black nightshade species groups across Africa (adapted from Edmonds and Chweya, 1997).

South Africa is a country of self-sufficiency when it comes to food, and several food crops are also exported. Nonetheless, malnutrition is still found in several rural communities across the country (WHO, 2009). Conventional vegetables are difficult to grow under resource poor and harsh climatic conditions and traditional leafy vegetables are of considerable importance in alleviating the problem of nutrient and calorie deficiencies in diets (Shippers, 1998). Chronic hunger is an enduring problem in several third world countries, but the efficient use of traditional, locally grown leafy vegetables can assist in enhancing the quality of foods to pave the way in preventing chronic hunger.

Relatively little information is available for black nightshade consumption and commercialization in South Africa, with isolated farmers of this relatively unknown leafy vegetable. *S. retroflexum* forms part of the black nightshade complex, and is sometimes considered indigenous to South Africa (Edmonds and Chweya, 1997). White cabbage (*Brassica oleraceae* L. var. *capitata*) is an introduced species to South Africa and is one of the most commonly produced and consumed leafy vegetable among rural populations. Together with *S. retroflexum,* these two species are rich in nutrients per unit fresh mass, and this makes them potentially suitable elements for inclusion in approaches to combat micronutrient deficiencies in South Africa. In total, there are approximately 50 irrigation farmers in the Vhembe District

© University of Pretoria

(Limpopo province) who actively produce *S. retroflexum* for the commercial market (Van Averbeke *et al*., 2007b). It should be noted however that these farmers concentrate on the leaves of *S. retroflexum* as commodity, instead of the berries due to consumer demand in that area (Van Averbeke, personal communication). Similar use of *S. scabrum* has been reported in the Eastern Cape and KwaZulu Natal in South Africa, but only the districts in Vhembe are producing these species at a commercial level (Juma, 2006). Small scale traders control the *S. retroflexum* market, and since 2005 the plant has also been supplied to large supermarket chain stores. This development clearly indicates that the *S. retroflexum* commodity system is expanding in the Vhembe area (Van Averbeke *et al*., 2007b). There is little literature on the cultivation of berries by farmers in South Africa, but it is known that at least two farms located in the Free State and KwaZulu Natal are utilizing the berries in the production of preserves such as jam.

Based on the abovementioned data, it is proposed that the black nightshades form an agriculturally important component to many rural farmers in South Africa. The numerous advantages associated with these plants prove that subsequent research on their morphological and genetic diversity as well as metabolic properties could significantly enhance the economical and agricultural sectors of South Africa. In the present study, it was aimed to identify black nightshade species being grown commercially for jam production in the eastern highlands of South Africa. One of the major objectives was to integrate findings on morphological and molecular evidence, to produce an identification key for easy identification of the various species in the field. General recommendations on diagnostic characters are given, as well as broad-spectrum suggestions to the farmers about which species are the best to use in commercial areas.

© University of Pretoria

Chapter 2

Plant Material

© University of Pretoria

2.1 Background

The *Solanum nigrum* L. complex (black nightshades) forms part of the Solanaceae, a family that includes some of the most important commercial food species such as tomatoes, potatoes and peppers. Morphological data is available for black nightshade species found in Europe, the Americas and Asia, where the potential of these crops as food source has been exploited for some time. Several metabolic and genetic studies have been performed, mainly to establish a method of effectively distinguishing between the species, as they are morphologically exceedingly similar. South African black nightshade species, however, are in need of comprehensive classifications into distinct species groups. During the present study, the focus was on plants collected from two farms on which nightshades are grown commercially for jam production.

2.2 Farm A study area

Farm A produces "Annies" nightshade jam for the commercial market. The farm is located in Harrismith in the Free State, South Africa, at an altitude of 1661 meters above sea level. Mr. Leon Strachan and his wife, Elsa, grow many crops including strawberries, blackberries, maize, tomatoes and potatoes together with the black nightshades. They also keep cattle, and run a guesthouse on the farm. Nightshades have been grown here since 1994 on approximately 10 ha, with the farmer relying heavily on rainfall, but a supplementary irrigation system had been installed in recent years. This is a high rainfall area, which keep the soil moist during growth (Figure 2.1).

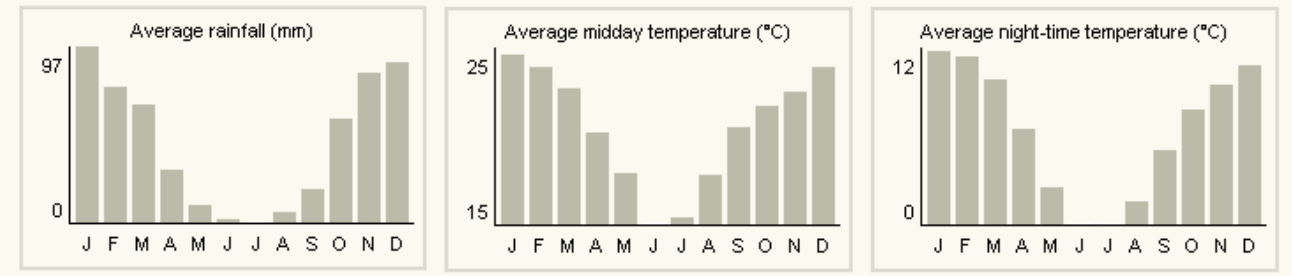


Figure 2.1 Average rainfall and annual temperature for Harrismith, South Africa. The region receives approximately 529mm of rain per year, with most rainfall occurring mainly during mid-summer. The chart (left) shows the average rainfall values for Harrismith per month. It receives the lowest rainfall (0 mm) in July and the highest (97 mm) in January. The monthly distribution of average daily maximum temperatures (centre chart) shows that the average midday temperatures range from 14.7°C in June to 24.6°C in January, and the minimum temperatures (right) were recorded as 0.1°C on average during the evening in July (www.saexplorer.co.za).

© University of Pretoria

Mr. Strachan does not use synthetic fertilizer and applies fresh chicken manure in an effort to produce organic products. Field workers consist predominantly of women from surrounding settlements, with ages ranging from 21–55 years. The workers obtain a brief introduction on the handling of the plants and berries prior to commencing fieldwork. Around 11 100 plants are grown in a season (starting in September), with many berries being produced per plant from January until early May. Frost is the main reason for plant death at the end of the growing season, with new seed germinated at the start of each season. Natural pests include several species of beetle, red spider mites and birds. The occurrence of root-rot fungi has also been recorded. The first nightshade species to be grown on a large scale in the 1990s was found in a garden on this farm, and was thought to be *S. retroflexum*. Mr. Strachan does not subject the plants to selective breeding programs, and no specific selection pressure is applied. Today, it seems that several different species types within the black nightshade complex are grown on the farm.

2.3 Farm B study area

Farm B produces “Umsobo” nightshade jam for the commercial market. The farm is located in Bothaspas (close to Newcastle) in KwaZulu Natal, South Africa. Mr. Pierewiet Conradie and his wife, Elize, are avid nightshade enthusiasts. Their focus on the farm is the growth, maintenance and commercial production of black nightshade berry preserves, but they also have a few cattle and limited plantings of maize. Mrs. Conradie manages a business (Kontreiwinkel) where the jam is sold, together with fig jam, marmalade and atchar, which she produces herself. Bothaspas is located near the town of Newcastle, which has an elevation of 1203m above sea level, and the annual weather and climate data are shown in Figure 2.2. Currently, 13 ha are used for black nightshade production, and approximately 10 tons of berries are obtained from these fields in one growing season. The plants are germinated in seedling trays in greenhouses at the beginning of August and transplanted into the fields during early September, with the second and third batches reaching the field during October and November, respectively. Harvesting of berries starts early in January and continues until early in May.

© University of Pretoria

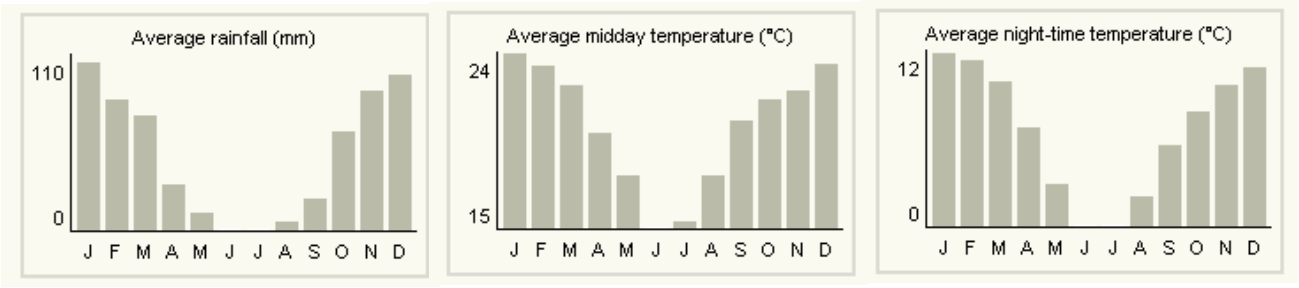


Figure 2.2 Average rainfall and annual temperature for Bothaspas, South Africa. The region receives approximately 569 mm of rain per year, with most rainfall occurring mainly during mid-summer. The chart (left) shows the average rainfall values for Bothaspas per month. It receives the lowest rainfall (0 mm) in June and the highest (105 mm) in January. The monthly distribution of average daily maximum temperatures (centre chart) shows that the average midday temperatures range from 14.9°C in June to 23.9°C in January, and the minimum temperatures (right) were recorded as 0.2°C on average during the evening in July (www.saexplorer.co.za).

Approximately 100–120 workers are needed for the berry picking, being mainly women from surrounding towns. Mr. Conradie has an irrigation system in place, which waters the plants at least once a week. Fertilization is administered through the water irrigation system, with including nitrogen, phosphorous and potassium (2:3:2), as well as $(NH_4)_2SO_4$. The biggest problem for the growth of these plants is invading weeds and Mr. Conradie approaches this problem by planting the nightshade plants in the middle of plastic sheets with their roots protruding through holes into the soil. In this way, weeds are not able to grow close to the nightshades. Alternatively, they are planted on a hay patch which is elevated, to firstly keeping the weeds from suffocating the nightshades and secondly to keep the root environment moist for optimal growth. The first species to be used on the farm was found in a vegetable garden in the Bothaspas area and is alleged to be *S. retroflexum*. At the moment, Mr. Conradie is still growing only one morphological species group. He keeps a close eye on the plants, and although no conscious breeding programs are in place, he selects seed from plants that seem to have slightly bigger berries to propagate in the next season. From the 10 tons of berries harvested, approximately 20 000 bottles of jam are produced, and dispatched to bakeries and private home industries across the country, with the most consumers in Gauteng.

© University of Pretoria

2.4 Plant material collected

Randomly selected plants were harvested from different planting fields on each of the two farms. The plants were visually identified to species based on distinguishing characters found in the literature (Table 2.1) and **these provisional identities (whether eventually shown to be correct or incorrect) will be followed through the course of the manuscript**. The plants were chosen based on their morphological diversity and putative identifications in the field. At least 15 leaf blades as well as 20 berries per plant were harvested. The material was placed in airtight plastic bags and kept in a portable refrigerator at 4˚C. Seed from a previously identified representative of *S. retroflexum*, *S. chenopodioides*, *S. scabrum*, *S. americanum* and *S. burbankii* was kindly donated by Jacoby *et al.* (2003) and will be included as reference material for each species group. The leaves were stored at -80˚C until DNA extractions were performed. The berries were stored at -20˚C prior to seed extraction.

© University of Pretoria

Table 2.1 Putative identities of black nightshades (Solanum spp.) from South Africa included in the present study

Sample[1]	Location	Area	Province	Leaves collected	Berries collected	Putative ID[2]
LS1	Farm A	Harrismith	Free State	Yes	Yes	*S. chenopodioides*
LS2	Farm A	Harrismith	Free State	Yes	Yes	*S. americanum*
LS3	Farm A	Harrismith	Free State	Yes	Yes	*S. chenopodioides*
LS4	Farm A	Harrismith	Free State	Yes	Yes	*S. chenopodioides*
LS5	Farm A	Harrismith	Free State	Yes	Yes	*S. americanum*
LS6	Farm A	Harrismith	Free State	Yes	Yes	*S. retroflexum*
LS7	Farm A	Harrismith	Free State	Yes	Yes	*S. americanum*
LS8	Farm A	Harrismith	Free State	Yes	Yes	*S. retroflexum*
LS9	Farm A	Harrismith	Free State	Yes	Yes	*S. americanum*
LS10	Farm A	Harrismith	Free State	Yes	Yes	*S. chenopodioides*
LS11	Farm A	Harrismith	Free State	Yes	Yes	*S. retroflexum*
LS12	Farm A	Harrismith	Free State	Yes	Yes	*S. americanum*
LS13	Farm A	Harrismith	Free State	Yes	Yes	*S. retroflexum*
LS14	Farm A	Harrismith	Free State	Yes	Yes	*S. chenopodioides*
LS15	Farm A	Harrismith	Free State	Yes	Yes	*S. chenopodioides*
PC1	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC2	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC3	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC4	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC5	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC6	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC7	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC8	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC9	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
PC10	Farm B	Bothaspas	KwaZulu Natal	Yes	Yes	*S. retroflexum*
S. americanum	University of the Free State[3]		Free State	No	Yes	*S. americanum*
S. burbankii	University of the Free State[3]		Free State	No	Yes	*S. burbankii*
S. retroflexum	University of the Free State[3]		Free State	No	Yes	*S. retroflexum*
S. scabrum	University of the Free State[3]		Free State	No	Yes	*S. scabrum*
S. chenopodioides	University of the Free State[3]		Free State	No	Yes	*S. chenopodioides*
Total accessions:		30				

[1]Abbreviations assigned to specific accessions.

[2]Based on field observations and diagnostic characters from literature

[3] QwaQwa Campus

© University of Pretoria

2.5 Greenhouse trials

Berries collected from the farm were fermented at room temperature for approximately 7 days, to ensure ripeness of seeds. Rotten berries were cut with a scalpel to remove the seeds within. These seeds were washed several times with ddH_2O, to remove all debris and placed on tissue paper, after which they were placed in an oven at 37˚C for 48 hours. Storage was at room temperature in eppendorf tubes (Laboratory and Scientific Equipment Company (Pty) Ltd., South Africa). Seeds were sterilized for 15 min in 1.5% sodium hypochlorite ("Jik") and washed at least three times in ddH_2O water for 10 min each. Three seeds of each accession were placed in pots containing a 50:50 mixture of potting soil and top soil (Earth-to-Earth Potting Soil, South Africa); with germination occurring approximately 20–25 days post planting. A plant mineral supplement (Multifeed Classic®) was applied once a week. Mature plants were grown in a phytotron in the Forestry and Agricultural Biotechnology (FABI) institute, where the conditions were 26˚C for 16 hrs and 23˚C for 8 hrs, with natural day and night light intensities respectively. Temperature and humidity were controlled within the greenhouse, with the humidity being approximately 30%. Plants were monitored and watered twice a day.

2.6 Herbarium material

Herbarium samples of several members of the *S. nigrum* complex, with emphasis on authentically identified collections such as the Type specimens, were consulted in different herbaria across the world. The scanned "Type species" images of each of the five species in this complex, which occurs in South Africa, were obtained (Table 2.2, Appendix A). Leaf type, leaf margin, leaf length to leaf width ratio, leaf length to petiole length ratio, berry size, indentation size and nr of pedicles were investigated where possible from type images, and are summarized in Table 2.3.

Samples collected during the present study were compared morphologically and geographically to identified herbarium type specimens, in order to make proposals concerning the true identities of the unknown plants. Morphological traits associated with type specimens are further discussed in Chapter 3.

© University of Pretoria

Table 2.2 Type specimens for species names of *Solanum* consulted in the present study

Species name	Reference	Image
S. americanum	The Gardeners Dictionary, ed 8 (1768)	Royal Botanical Gardens, Kew Herbarium *S. nodiflorum* var γ-*acuminatum*
S. chenopodioides	Tableau encyclopédique et méthodique des trois Règnes de la Naure. Botanique 2:18 (1794)	Herb. Lamarck s. n. Diet no. 35
S. nigrum	Species Plantarum 186 (1753)	Herb. LINN 248.18
S. retroflexum	A.P. De Candolle, Prodromus Systematis naturalis Regni vegetrabilis 13 (1): 50 (1852)	Royal Botanical Gardens, Kew Herbarium 7864α and β
S. scabrum	The Gardeners Dictionary, ed 8 (1768)	Royal Botanical Gardens, Kew Herbarium *S. scabrum* (JM Edmonds); Herb Lamarck Diet no. 34

Table 2.3 Morphological characters of *Solanum nigrum* complex Type species

	Leaf shape	Leaf margin	Leaf L:W[1]	Leaf L:PL[2]	Berry size (mm)	Lamina Indentations	Nr of pedicels
S. nigrum	Ovate to rhomboid	Sinuate-dentate	1.4	1.7	5.5	1.5	4-6
S. retroflexum	Ovate to rhomboid	Serrate	1.4	2.1	7.3	2	4-6
S. americanum	Lanceolate	Entire	3.3	2.5	5.0	0	4-6
S. chenopodioides	Lanceolate	Entire	1.7	3.4	3.8	0	4-6
S. scabrum	Ovoid	Entire	1.2	1.6	9.8	0	-

[1] Ratio of leaf length (mm) to leaf width (mm)

[2] Ratio of leaf length (mm) to petiole length (mm)

© University of Pretoria

In the present study, one representative plant from each identified morphological group collected, was harvested at the five flower stage and placed in a plant press and dried at 65˚C for 48 hours. For each specimen the sample location, putative identity and general morphology were recorded. The samples were deposited in the H.G.J.W. Schweickerdt Herbarium [PRU] at the University of Pretoria as follows: *S. americanum* (Harrismith) – PRU 115535.0; *Unknown species* (suspected hybrid species) (Harrismith) – PRU 115536.0; *S. chenopodioides* – PRU 115537.0; *S. retroflexum* (Harrismith) – PRU 115538.0; *S. retroflexum* (Bothaspas) – PRU 115539.0 and *S. scabrum* – PRU 115540.0.

The abbreviations as listed in this chapter will be applied in this manuscript.

© University of Pretoria

Chapter 3

Morphological and Cytological Analysis

© University of Pretoria

3.1 Introduction

The black nightshades (*Solanum nigrum* complex) belong to the economically important Solanaceae, a family of which tomatoes, potatoes and peppers are well-known members. Human societies require a diverse range of plant species, cultivated across the world as staple crops for food supply. However, it is important that the contributions of minor crop species as sources of food should not be underestimated. These underutilized or neglected crops have failed to attract significant research funding in the past, but are currently receiving increased attention. An important advantage of these minor crops is their significant genetic diversity, which enables them to adapt to various environmental conditions. Many of these plants retain their ability to sustain growth in areas where conventional crops fail, such as arid areas and salt-affected soils (Edmonds and Chweya, 1997). Their natural and genetic diversity could be exploited providing the addition of genetic information. A shortage of this type of information considerably restricts the development and sustainable conservation of these crops.

As a minor crop, black nightshades are mainly considered a leafy vegetable (Venkateswarlu and Rao, 1972), although its potential as a source of edible berries is becoming increasingly important, especially in more affluent societies. It is cultivated in certain parts of the world as an alternate food resource to more exotic leafy vegetables such as spinach or cabbage. In South Africa, the leaves are generally used as a relish or as a substitute for other expensive or rare leafy vegetables. The berries are either utilized fresh or in commodities such as preserves (jam). The phenotypic variation within this group of plants hampers identifications based on gross morphology. Ogg *et al.* (1981) summarized different forms of nightshade species growing in several regions in North America. The authors found differing growth forms between seasons, as well as between geographical locations (Figure 3.1). These difficulties have been studied morphologically, molecularly and numerically (comparing patterns of differentiation) to aid in explaining why this should be the case (Venkateswarlu and Rao, 1972; Jardine and Edmonds, 1974; Edmonds 1972, 1977, 1979).

© University of Pretoria

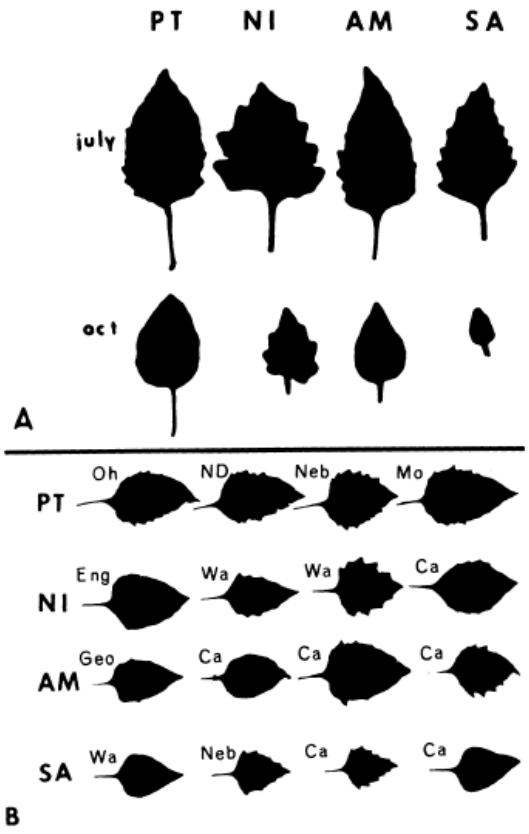


Figure 3.1 Variations in leaf shape within the *Solanum nigrum* complex between seasons (A) and geographical locations (B). Leaf sizes vary mostly between seasons, while leaf margins shows several morphological variations between locations; Ohio (Oh), North Dakota (ND), Nebraska (Neb), Montana (Mo), England (Eng), Washington (Wa), California (Ca) and Georgia (Geo), USA. Species investigated include *S. ptycanthum* (PT), *S. nigrum* (NI), *S. americanum* (AM) and *S. sarrachoides* (SA). Adapted from Ogg *et al.* (1981).

The black nightshades have been subjected to numerous classifications since their original description by Linnaeus in 1753. Due to the insufficient diagnostic characters to differentiate among species, several "new" taxa are in fact mere morphological variations of species previously described (Edmonds and Chweya, 1997). The significant phenotypic plasticity and variation of several vegetative features contributes to the difficulties with identification. It has been noted that senescence is accompanied by smaller berries, and fewer, smaller flowers, implying that such character states are not necessarily species specific traits. Anthocyanin production in flowers which leads to a purple stripe has been shown to be dependent on light intensities, implying its expression is probably not species specific either (Edmonds, 1977). Based on these observations it is difficult to assign morphological traits as being environmentally controlled or genetically fixed. The species belonging to this complex constitute a polyploidy series, with

© University of Pretoria

diploid, tetraploid and hexaploid forms occurring across the entire geographical range previously assigned to these crops (Edmonds and Chweya, 1997). Cases of natural hybridizations (plants normally self-fertilize) have been reported, which resulted in morpho-genetically complex population variations (Venkateswarly and Rao, 1972). Ploidy levels have previously been used to aid in distinguishing among species, but recent classifications based primarily on chromosomal numbers have rendered the effective identification based solely on chromosome counts ineffective. Due to possible hybridization events, the ploidy amount of a certain accession may not truly reflect the identity.

The *S. nigrum* complex has many characters to consider for their potential taxonomic significance, but it is proposed that a combination of several character states be employed in order to achieve positive identification. In this chapter the taxonomic significance of 30 morphological characters as well as ploidy levels are assessed and applied to demarcate taxonomic entities within accessions of the *S. nigrum* complex from South Africa. These accessions were compared to five reference samples previously identified as *Solanum americanum*, *Solanum burbankii*, *Solanum chenopodioides*, *Solanum retroflexum* and *Solanum scabrum* by Dr. A van Biljon (née Jacoby) at the University of the Free State. Isolates which were unsuccessfully matched to the known samples were tentatively identified based on the available literature.

© University of Pretoria

3.2 Materials and methods

3.2.1 Measurements and photographs

Samples were collected from two farms (A and B) in April 2009. Information concerning the maturity of the plants and berries were collected from the farmers (See chapter 2 for background and putative identities of accessions). Morphological characters or traits as well as their scoring methods are indicated in Table 3.1.

Table 3.1 Characters, character states and assessment techniques

Character	Technique	Character States
Plant Height (m)	Ruler	Quantitative
Plant growth form	Observe	Erect; spreading
Stem color	Observe	Dark green; light green; slightly purple
Stem texture	Scanning Electron Microscope	Density of trichomes
Leaf shape	Observe	Lanceolate; Rhomboid to ovate; ovate
Leaf margin	Observe	Entire; entire to sinuate; serrate
Leaf area (mm^2)	Assess	Quantitative
Leaf indentation (mm)	Assess	Quantitative
Leaf length:leaf width	Quantitative	Quantitative
Petiole length:leaf length	Assess	Quantitative
Leaf texture (trichomes)	Scanning Electron Microscope	Density of trichomes
Trichomes	Scanning Electron Microscope	Long glandular; short glandular; eglandular
Flower outline	Assess	Elongated; triangular; stellate
Corolla length:corolla width	Quantitative	Quantitative
Flower colour	Observe	White; White with purple stripe
Anther colour	Observe	Yellow; brown
Style exertion	Observe	Yes; No
Berry size (mm)	Ruler	Quantitative
Berry color	Observe	Green; dull purple; shiny purple
Berry surface	Scanning Electron Microscope	Dull; shiny
Berries per bunch	Count	Quantitative
Inflorescence orientation	Observe	Erect; deflexed
Calyx orientation	Observe	Reflexed; adherent
Pedicel number	Count	Quantitative
Pedicel origin	Observe	Same point; not same point
Seed size (mm)	Assess	Quantitative
Number of seeds per fruit	Count	Quantitative
Number of seeds: berry size	Quantitative	Quantitative
Seed surface	Scanning Electron Microscope	Density of trichomes
Ploidy level	Flow cytometry	Diploid; tetraploid; hexaploid

Plants were classified as mature if more than five open flowers were present per plant. Intact leaves and flowers of mature plants were harvested and scanned (HP ScanJet G4050, Microsoft, South Africa). The

© University of Pretoria

images were analyzed using the computer software program Assess (Version 2.0, Image Analysis Software for Plant Disease Quantification, 2002-2008 American Phytopathological Society). Colours observed for flowers, stems and berries were classified according to character states described in Table 3.1. Plant height, leaves, fruit and seed size were measured using a standard 30 cm ruler (Figure 3.2). Fruits per bunch and seeds per fruit were counted. Individual accessions were photographed for the whole plant, stems, leaves, flowers and berries with a Canon digital camera. Photographs of seed were taken with a stereo microscope (Nikon SMZ800, Nikon Instech Co., Kanagawa, Japan). All plant structures were harvested at approximately 6-week old stage, and were collected from non-reproductive nodes.

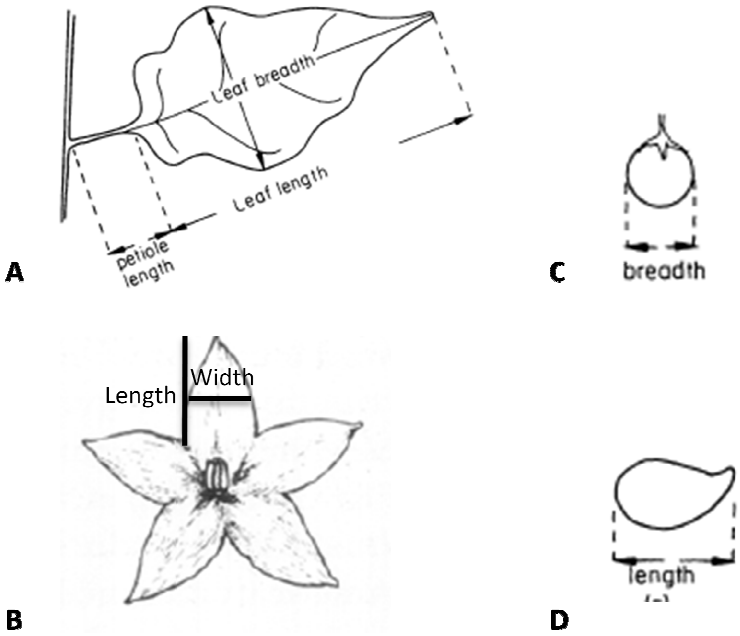


Figure 3.2 Measurements of morphological features. (A) Leaf lamina length, lamina width and petiole length. (B) Corolla lobe length and width. (C) Berry diameter. (D) Seed diameter.

© University of Pretoria

Standard descriptive terminology was used to describe the morphology of the plants. For a useful glossary on the meaning of the various terms, see Beentje (2010). The terms "erect" and "spreading" used to describe growth form (habit) requires further explanation. Erect refers to main stems with a tendency to be orientated more or less erect; habit of plants relatively tall, slender and upright. Spreading indicates main stems with a tendency to spread sideways; habit of plants relatively short, more rounded and bushy. Each accession was linked to its **putative identity** mentioned in Chapter 2 (Table 2.1), and data results are shown in graphs for easy identification in this chapter; *S. americanum* = green; *S. chenopodioides* = pink; *S. retroflexum* = blue; *S. burbankii* = orange and *S. scabrum* = yellow.

3.2.2 Composite analyses

The analytical software program, R was designed for data manipulation, calculations and graphical display (Gentleman *et al.,* 2004). Normality of the data was tested in a one way Analysis of Variance (ANOVA), to test if the sample means are significantly different to each other for every datapoint. The data was found to be not normally distributed, and therefore a non-parametic (Kruskal-Wallis) test was performed. In the present study, the null hypothesis (H_0) is that medians, rather than the means, are equal across all samples.. Boxplots were produced through R package software (R Version 2.10.1, 2009 The R Foundation for Statistical Computing), to compare the values of each accession graphically for each trait assessed. The boxplots represent sample values in four quartiles, where the upper two quartiles are separated from the lower two by the sample median value. Each quartile signifies 25% of the data, and the box area represents 50% of the data. The ends of vertical lines extending from the box represents maximum and minimum observed values, unless there are outliers (small, circular structures), in which case the vertical lines represent values which are 1.5 times the inter quartile range (boxed area on plot – 50% of data) (Figure 3.3).

© University of Pretoria

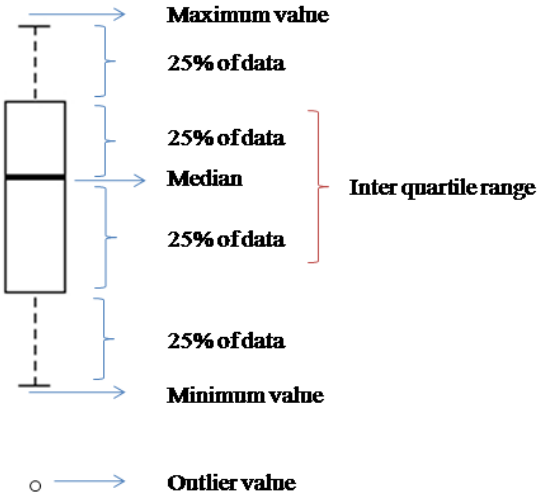


Figure 3.3 Diagrammatic representation of a boxplot. The maximum and minimum values of the dataset are indicated, and the box area represents 50% of total data (inter quartile range), 25% above median and 25% below median. Outliers are indicated by circles.

Traits investigated through R analysis include plant height, indumentum size, ratio of leaf length to leaf width, ratio of leaf length to petiole length, ratio of corolla length to corolla width, berry size, seed size, amount of seed per berry, ratio of amount of seeds to berry size and amount of trichomes on abaxial and adaxial leaf surfaces. Analysis of variance (ANOVA) can be used for testing equality of multiple independent population means. Before applying this test, the normality of the distributions for each population (group) within each trait was calculated. A Shapiro-Wilks test was used to determine a P-value which is indicative of the data's distribution. A P-value less than 0.05 (95% confidence) is deemed sufficient to reject the null hypothesis and assume the data does not fit a normal distribution. The data for each trait was found to be not normally distributed, and a Kruscal-Wallis non-parametric test was performed to test which population medians differ significantly from the rest of the populations. In addition, principle component analysis was performed, using standardized morphological quantitative data to obtain a graphical representation of the relatedness of the 30 accessions. PCA was performed using The Unscrambler X 10.0 analysis program (2010 CAMO Software, Woodbridge, USA).

© University of Pretoria

3.2.3 Scanning electron microscopy

Fresh leaf material was harvested near the top of the plant and sectioned into 5 mm: 5 mm parts. These were subsequently placed in 2.5% gluteraldehyde in 0.075 M phosphate buffer (pH 7.4) and were kept at 4˚C until use. The leaf material was washed three times, 10 min each wash, in 0.075 M phosphate buffer, and fixed in 0.5% aqueous osmium tetroxide for 2 hours. Afterwards, the samples were washed three times, 15 min each wash, with distilled water. The samples were dehydrated by sequential treatment of 10 min each in an ethanol series of 30%, 50%, 70%, 90% and 100%. The treatment with 100% was repeated three times to ensure complete dehydration. The leaf samples were critical point dried with liquid CO_2 in an Autosampler 810 critical point drier and sputter-coated with gold particles. Both the adaxial and abaxial surfaces were observed with a JSM 840 Electron Microscope (JEOL, Tokyo, Japan). Images were captured, stored as .tiff files and analyzed.

3.2.4 Flow cytometry

Flow cytometry experiments were performed by Dr. Jan Suda (Laboratory of Flow Cytometry, Institute of Botany, Prague, CZ) using fresh, young leaves of selected accessions, shipped in moist tissue paper. A modified method of Doležel *et al*. (2007) was followed. Lamina tissues were harvested and 20 mg was placed in a sterile petri-dish (Merck). One ml ice cold nuclei isolation buffer (Appendix B) was added and the tissue was immediately cut with a sterile razor blade or scalpel. The homogenate was mixed and filtered through a 42 µm nylon mesh. The sample was visually inspected for any debris that may clog the machine in subsequent readings. Propidium iodide (PI) fluorochrome (50 µg/ml and 50 µg/ml RNase) was added and incubated on ice for an hour with occasional shaking. The suspension of stained nuclei was introduced into the flow cytometer and processed for a few seconds. The sample rate was adjusted (usually 20–50 nuclei per second), and the DNA peak(s) were positioned to the required positions in the abscissa. A threshold to discard undesirably low channel signals derived from cell debris was set. Approximately 5000 particles were measured, and the data analyzed by an appropriate software program to estimate the ploidy levels of each sample. To accurately interpret the flow cytometer results, it is necessary to include a reference standard of known characteristics (e.g. ploidy level, genome size), and *Bellis perennis* (2C = 3.38pg) was used as standard.

© University of Pretoria

3.3 Results

For each of the 30 accessions collected (see Chapter 2, Table 2.1), three replicates were grown under uniform greenhouse conditions. Four measurements were collected per data point per plant. This resulted in twelve replicate measurements for each data point per accession. In the following section, the morphological characteristics pertaining to the whole plant, leaves, flowers, berries and seed will be individually discussed.

3.3.1 Whole plant

3.3.1.1 Height

Height variation was observed for the accessions at approximately six weeks post germination. Stem growth showed an erect (>95 cm) versus a spreading form (<90 cm) and petiolated leaves are alternately arranged on the stems for all accessions (Figure 3.4). The tallest growth form was LS12 (175 cm), while the shortest plant was BUR (39 cm). The median plant height across all accessions was 77.64 cm. Seven accessions (AME, SCA, LS7, LS12, PC3, PC7 and PC9) showed an erect, tall growth form above the average starting at 98 cm to 175 cm (Figure 3.4 AME, SCA, g, l, v, x). The remaining accessions were classified as having a spreading growth form, below 90 cm in height (Figure 3.4 a, b, c, d, e, f, h, i, j, k, m, n, o, p, q, r, s, t, u, w, CHEN). Accessions LS1, LS2, LS3, LS4, LS5, LS6, LS8, LS9, LS10, LS11, LS13, LS14, LS15, PC1, PC2, PC4, PC5, PC6, PC8, PC10, BUR and CHEN differed in height by small increments with height varying between 39 cm and 89 cm (Figure 3.5).

© University of Pretoria

Figure 3.4 Morphology of whole nightshade plants (Continued to next page). (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14 and (o) LS15. Scale bar = 100 mm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

Figure 3.4 (Cont.) Morphology of whole nightshade plants. (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum;* (BUR) *S. burbankii;* (CHEN) *S. chenopodioides;* (RET) S. *retroflexum* and (SCA) *S. scabrum.* The plants were all grown in similar conditions within a glasshouse, and have an average age of six weeks. Scale bar = 100 mm. BUR and RET accessions had difficulty in growing to maturity, and therefore images are absent. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

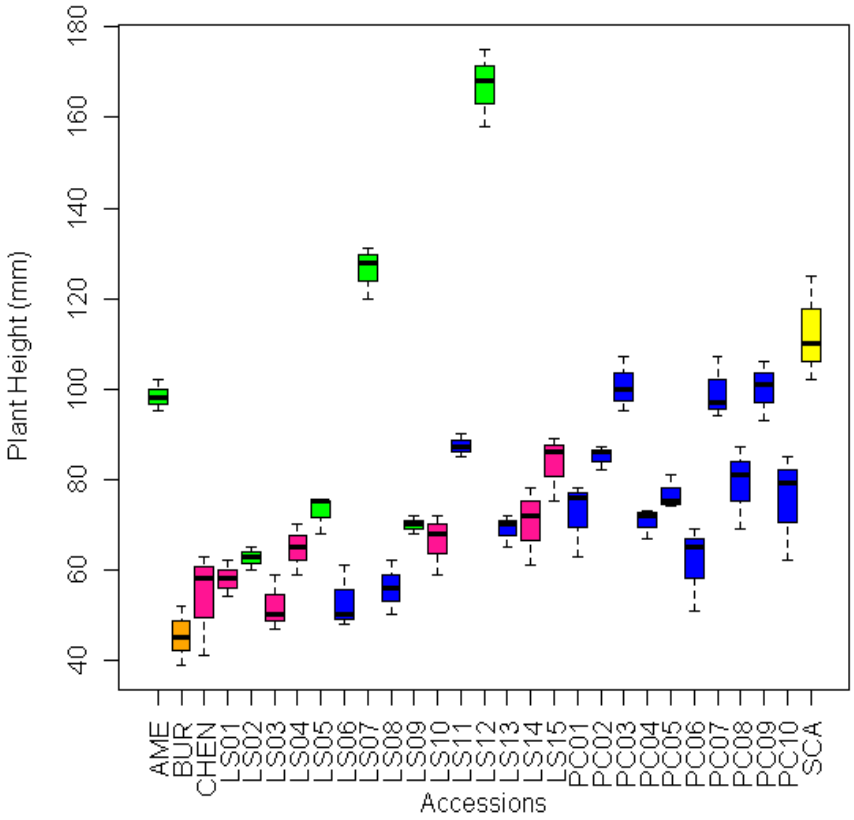


Figure 3.5 Boxplots indicating data distributions around the medium for plant height for each accession. Boxplot calculations were performed in R statistical analysis software. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

3.3.1.2 Stem

Stems are coloured either green or purplish green, and have varying indumentums (degree of hairiness) (Figure 3.6). Close observations showed that LS5, LS7, LS9, LS12 and PC3 had a purplish green stem with only slight indumentum (Figure 3.6 e, g, i, l, r). PC4 and PC10 were the only accessions which showed a moderate indumentum together with a purplish green stem (Figure 3.6 s, y). No accessions had a high indumentum, and purple tinged stem. The majority of accessions, LS2, LS6, LS8, PC1, PC2, PC9, AME and SCA, showed a green stem with slight indumentum (Figure 3.6 b, f, h, p, q, x, AME and SCA). A moderately hairy green stem was observed for BUR, LS11, LS13, PC5, PC6, PC7 and PC8 (Figure 3.6 k, n, t, u, v and BUR). High indumentum green stems were observed for accessions LS1, LS3, LS4, LS10, LS14, LS15 and CHEN (Figure 3.6 a, c, d, j, n, o and CHEN).

© University of Pretoria

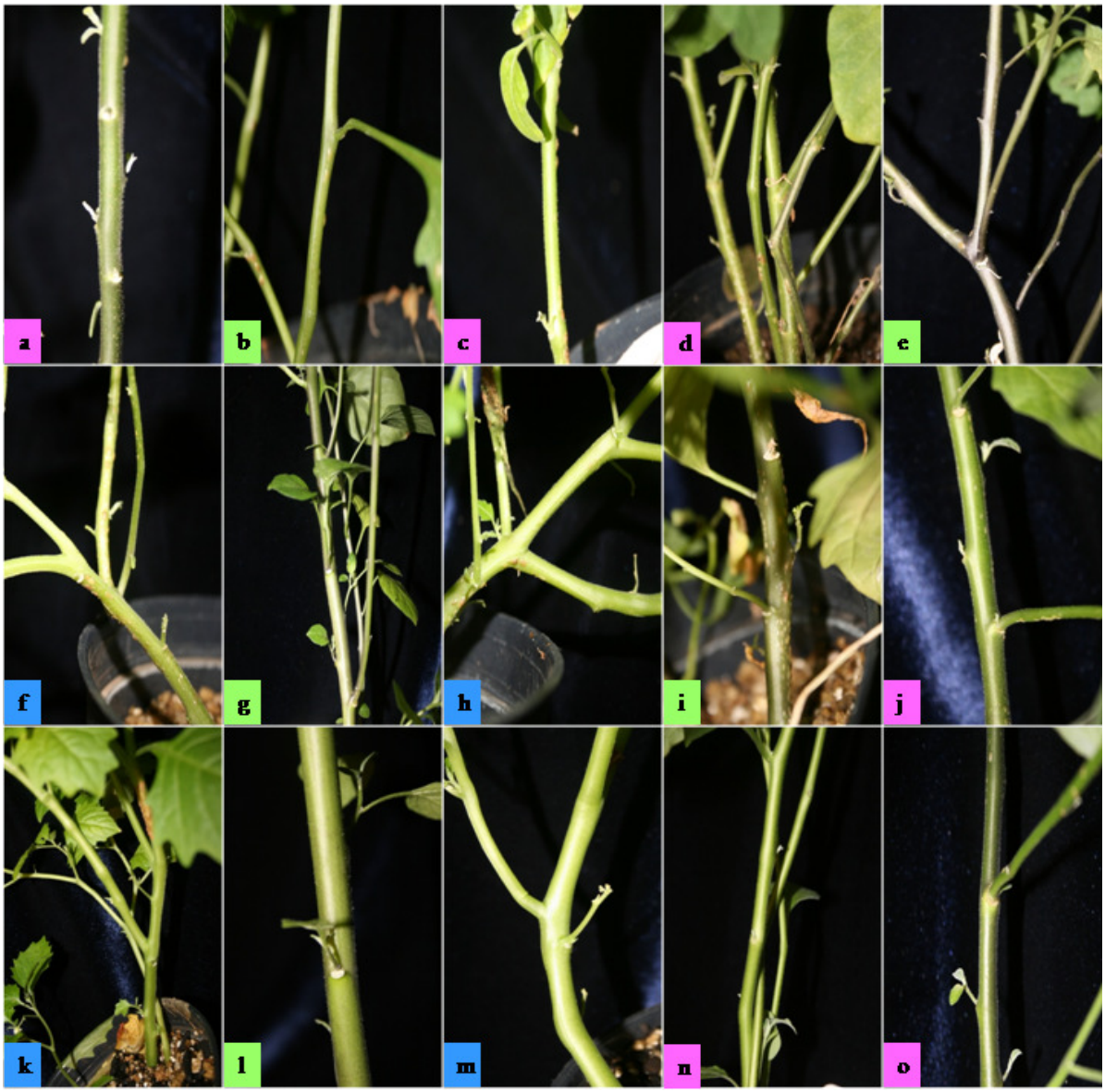


Figure 3.6 Morphology of nightshade plant stems (Continued to next page). (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14 and (o) LS15. The plants were all grown in similar conditions within a glasshouse, and have an average age of six weeks. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

Figure 3.6 (Cont.) Morphology of nightshade plant stems. (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum;* (BUR) *S. burbankii;* (CHEN) *S. chenopodioides;* (RET) S. *retroflexum* and (SCA) *S. scabrum*. The plants were all grown in similar conditions within a glasshouse, and have an average age of six weeks. BUR and RET accessions had difficulty in growing to maturity, and therefore images are absent. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.1.3 Scanning electron microscopy of stem

Cells on stem surface are arranged in longitudinal lines, with a striate appearance in surface view. However, in comparison to the other accessions, SCA had randomly arranged cells with a polygonal outline, and are not clearly axially elongated. Further investigations may reveal if this trait is possibly diagnostic for *S. scabrum*. The stems were slightly (< 2 trichomes/mm^2), moderately (5–20 trichomes/mm^2) to highly pubescent herbaceous (> 20 trichomes/mm^2). Three types of trichomes occur on the surface of the stem, as well as leaf. These were short glandular (SG), characterized by a short (± 50 µm), thick base, a stalk and the head (gland) at the top. Longer (±150 µm), multi-cellular trichomes with a glandular head, referred to as long glandular (LG), occurred with less frequency. The most prevalent was a long, multi-cellular trichome where the gland was absent (eglandular - EG) (Figure 3.7). SG, LG and EG trichomes were all present on the stem. As maturity was reached, the number of LG trichomes diminishes. SG trichomes were present on all accessions.

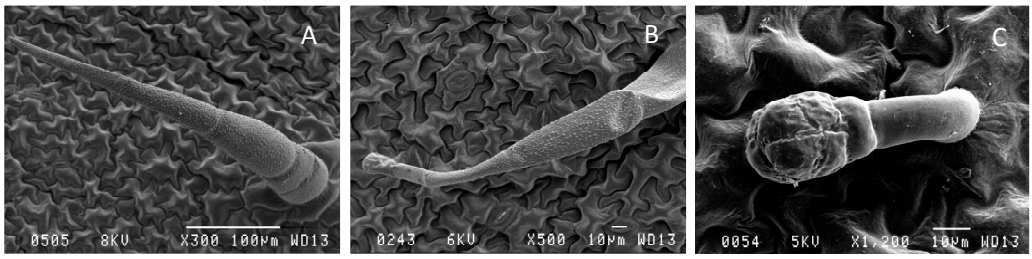


Figure 3.7 Scanning electron microscope images for types of trichomes. (A) eglandular (B) long glandular (C) short glandular

AME, LS2, LS5, LS7, LS9 and LS12 were slightly pubescent, while BUR, SCA, LS6, LS8, LS11, LS13 and PC1–PC10 were moderately pubescent. CHEN, LS1, LS3, LS4, LS10, LS14 and LS15 were highly pubescent, and had a dense indumentum of eglandular trichomes occurring on the whole stem (representative images in Figure 3.8). Stem pubescence was variable across accessions as well as within accessions. The most prevalent indumentum types are summarized in Table 3.2.

© University of Pretoria

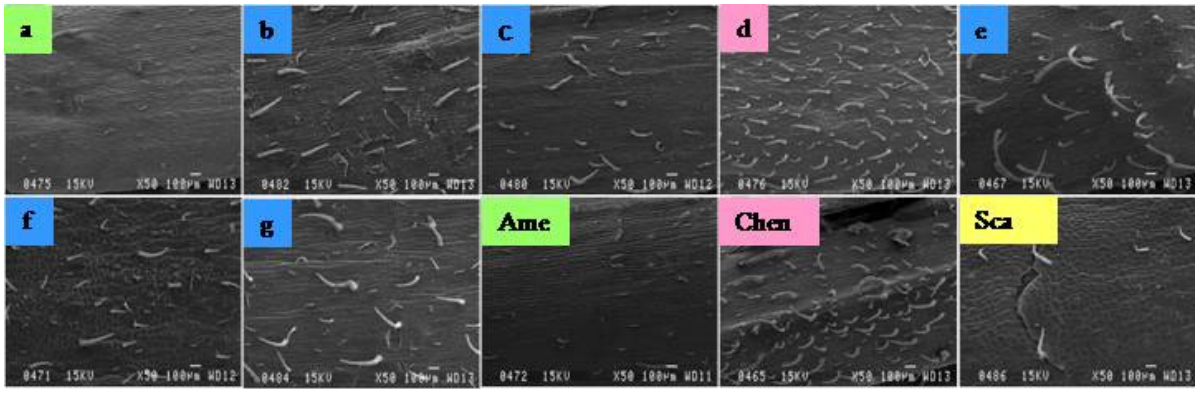


Figure 3.8 Scanning electron microscope images of stem indumentum. (a) LS7; (b) LS11; (c) LS13; (d) LS14; (e) PC5; (f) PC7; (g) PC10; (h) AME; (i) CHEN and (j) SCA. Slight pubescence = < 2 trichomes/mm^2, moderate pubescence = 2–20 trichomes/mm^2 and highly pubescent = > 20 trichomes/mm^2. Scale bar = 100 µm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

3.3.1.4 Character summary for whole plant, stem and stem trichomes

Plant growth form seems to distinguish effectively between two main groups of plants, namely a tall, erect growth form versus a shorter, spreading growth form. Plant height is too unstable to be further diagnostic in spreading growth forms. It was anticipated that only SCA would have a tall, erect growth form, but additional accessions were observed to have an erect instead of spreading growth form. The reason is unclear, but may be due to placing within the growth-room, leading to an irregular result due to environmental influences. Stem indumentum divides accessions into three groups, with the stems being slightly, moderately to highly pubescent. The putative identities were reflected in the results noted for the whole plant, stem and stem trichomes.

© University of Pretoria

Table 3.2 Stem indumentum types

	Stem texture			Trichomes present on stem		
	Not pubescent (< 2 trich/mm^2)	Moderately pubescent (5–20 trich/mm^2)	Pubescent (> 20 trich/mm^2)	Short glandular (50µm)	Long glandular (150 µm)	Eglandular
LS1	-	-	XXX	XXX	-	XXX
LS2	XXX	-	-	XXX	-	XXX
LS3	-	-	XXX	XXX	-	XXX
LS4	-	-	XXX	XXX	-	XXX
LS5	XXX	-	-	XXX	-	XXX
LS6	-	XXX	-	XXX	X	XXX
LS7	XXX	-	-	XXX	-	XXX
LS8	-	XXX	-	XXX	X	XXX
LS9	XXX	-	-	XXX	-	XXX
LS10	-	-	XXX	XXX	-	XXX
LS11	-	XXX	-	XXX	X	XXX
LS12	XXX	-	-	XXX	-	XXX
LS13	-	XXX	-	XXX	X	XXX
LS14	-	-	XXX	XXX	-	XXX
LS15	-	-	XXX	XXX	-	XXX
PC1	-	XXX	-	XXX	X	XXX
PC2	-	XXX	-	XXX	X	XXX
PC3	-	XXX	-	XXX	X	XXX
PC4	-	XXX	-	XXX	X	XXX
PC5	-	XXX	-	XXX	X	XXX
PC6	-	XXX	-	XXX	X	XXX
PC7	-	XXX	-	XXX	X	XXX
PC8	-	XXX	-	XXX	X	XXX
PC9	-	XXX	-	XXX	X	XXX
PC10	-	XXX	-	XXX	X	XXX
S. americanum	XXX	-	-	XXX	-	XXX
S. burbankii	-	XXX	-	XXX	-	XXX
S. chenopodioides	-	-	XXX	XXX	-	XXX
S. retroflexum	-	XXX	-	XXX	-	XXX
S. scabrum	-	XXX	-	XXX	-	XXX

XXX Usually; XX Occasionally; X Rarely; - Never

© University of Pretoria

3.3.2 Leaves

For this assessment the leaf will be divided into lamina blade, lamina margin, lamina size and petiole followed by the observations of scanning electron microscopy.

3.3.2.1 Lamina blade

Based on the literature, in nightshades the leaf blade shape can be separated into three states. The first is lanceolate, characterized by a thin, elongated leaf form which is pointed at both ends. The second is ovate, being egg-shaped and wide at the base. The third is a combination of ovate and rhomboid, or having an egg shape leaning towards a slight diamond shape (Figure. 3.9) (Edmonds and Chweya, 1997).

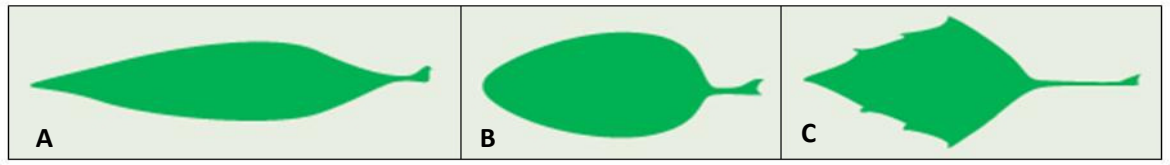


Figure 3.9 Examples of leaf shapes and margins. (A) Lanceolate (margin entire); (B) ovate (margin entire) and (C) rhomboid (margin serrate) (zipcodezoo.com).

CHEN, LS1, LS3, LS4, LS10, LS 14 and LS15 fit into the first state, with a clearly elongated, lanceolate shape (Figure 3.10 a, c, d, j, n, o and CHEN). The second category is ovate, with AME, SCA, LS7 and LS12 assigned to this group (Figure 3.10 g, l, AME and SCA). Although the rest of the accessions (BUR, RET, LS2, LS5, LS6, LS8, LS9, LS11, LS13 and PC1–PC10) all seemed to have an ovate shape as well, there was a particular rhomboid, or diamond shape, characteristic to this group (Figure 3.10 b, e, f, h, I, k, m, p, q, r, s, t, u, v, w, x, y, BUR and RET). Leaf shape appeared to be a distinguishing character, dividing the accessions into clear groups. The apices of the leaf blades could also be divided into two groups, namely pointed or cuspidate. However, observations show either leaf tip can be seen in replicates within the same accession, and is not considered a taxonomically informative morphological trait.

© University of Pretoria

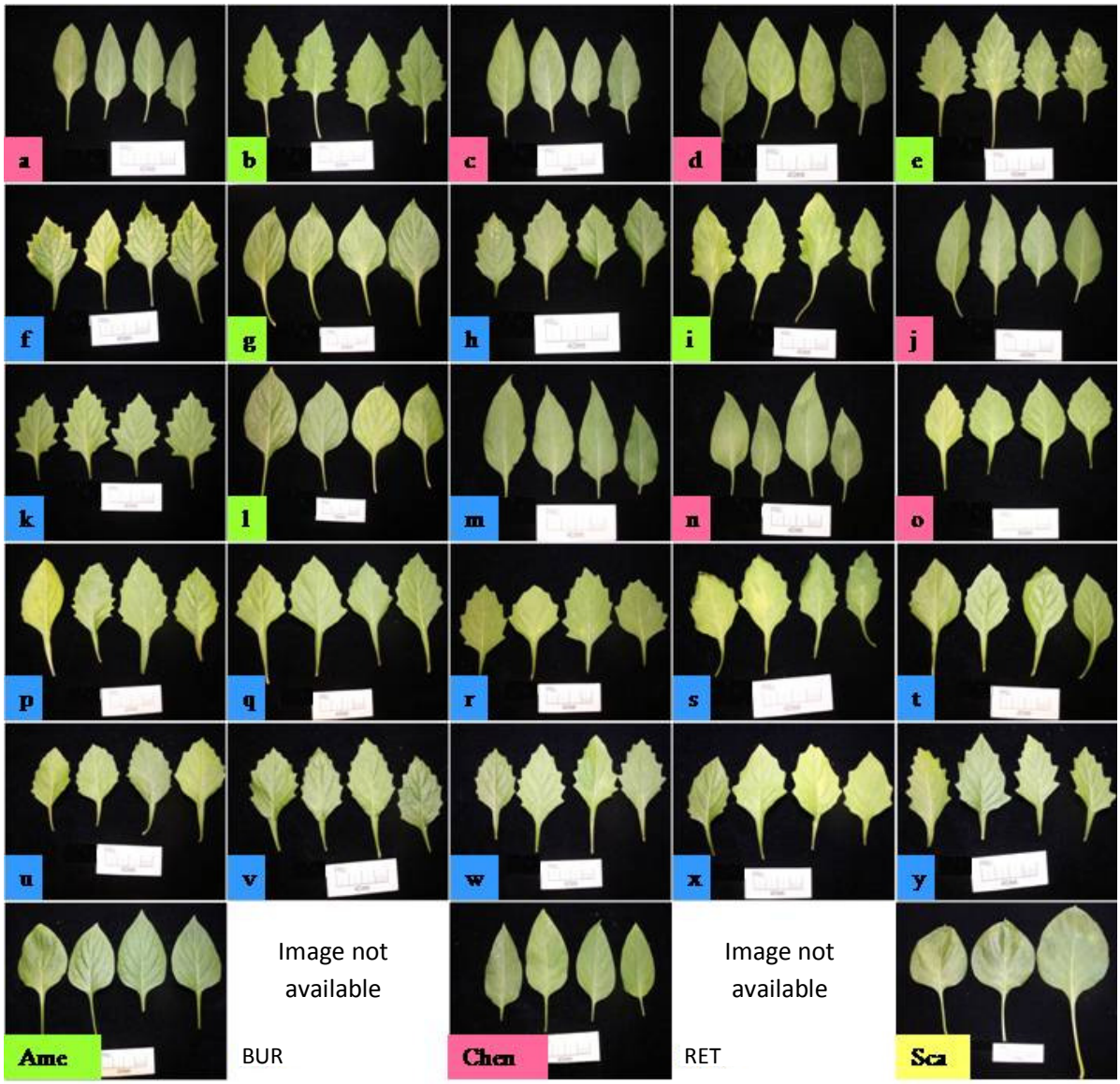


Figure 3.10 Leaf morphology for all nightshade accessions. (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) S. *retroflexum* and (SCA) *S. scabrum*. The plants were all grown under similar conditions within a glasshouse, and leaves were harvested from plants with an average age of six weeks. BUR and RET accessions had difficulty in growing to maturity, and therefore images are absent. Scale bar = 40 mm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.2.2 Lamina margin

Leaf margins of black nightshades can be classified as either entire or serrate (teeth forward-pointing) (Ogg *et al.,* 1981). Two groups could be distinguished with CHEN, LS1, LS3, LS4, LS7, LS10, LS12, LS14, LS15, SCA and AME (Figure 3.10 a, c, d, g, j, l, n, o, CHEN, SCA and AME) all having entire margins. In contrast BUR, RET, LS2, LS5, LS6, LS8, LS9, LS11, LS13, and PC1–PC10 had serrate leaf edges (Figure 3.10 b, e, f, h, i, k, m, p, q, r, s, t, u, v, w, x, y, BUR and RET). The second group could be split further based on the depth of indentations, with LS2, LS5 and LS9 having higher median teeth sizes (> 3 mm long) compared to BUR, RET, LS6, LS8, LS11, LS13 and PC1–PC10 (< 3 mm long) (Figure 3.11).

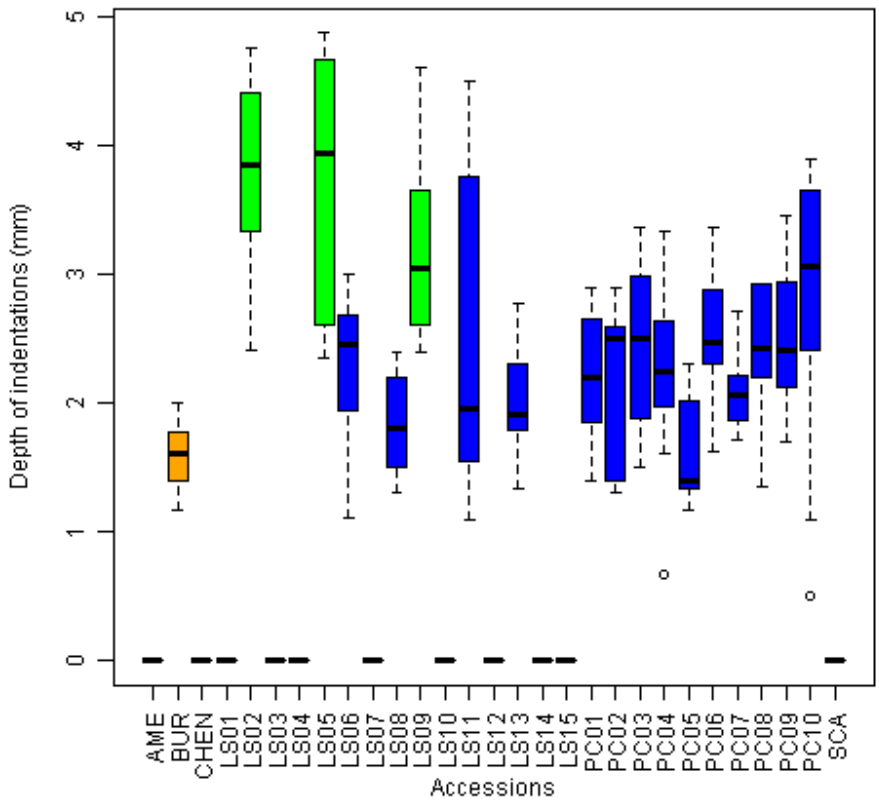


Figure 3.11 Boxplots indicating data distributions around the median for all accessions pertaining to indentation size. A value of 0 indicates an entire leaf margin while the remaining accessions' indentations are measured in mm. Boxplot calculations were performed in R statistical analysis software. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.2.3 Lamina size

It was difficult to determine whether leaf area is a diagnostic morphological trait, as the amount of shade, water and age of the plant all influenced lamina area (Edmonds and Chweya, 1997). The largest mature leaf surface area was observed to be 7887 mm^2 for SCA, and the smallest was 200 mm^2 for CHEN. The median leaf surface area across all accessions was 1881 mm^2. Generally, accessions with larger leaf surface areas (>3200 mm^2) were SCA, LS7 and LS12, whiles LS5, LS9 and LS11 had leaf areas of between 2000 mm^2 and 3200 mm^2. Lamina surface area of LS1, LS2, LS3, LS4, LS6, LS8, LS10, LS13, LS14, LS15, PC1–PC10, AME and BUR varied between 1000 mm^2 and 2000 mm^2. CHEN had the smallest average leaf surface area, less than 1000 mm^2 (Figure 3.12). These measurements were based on average values across the twelve replicates of each accession.

Furthermore, lamina length to width ratio was calculated by dividing the length of the leaf (mm) by the width of the leaf (mm). The highest ratio was observed for LS5 (3.6), the lowest ratio for BUR (0.6) and a median of 1.5 across all accessions. CHEN, LS1, LS2, LS3, LS4, LS5, LS9, LS10, LS14 and LS15 had a ratio of more than 1.5 times the leaf length to leaf width, and could be linked to the lanceolate shape of the leaf. Accessions with approximately equal leaf length to leaf width were AME, LS6, LS7, LS8, LS11, LS12, LS13 and PC1-PC10, and are linked to the ovoid to rhomboid shapes of the leaves. BUR had leaves which are wider than they are long, with a ratio of less than 1 (Figure 3.13).

3.3.2.4 Petiole

The ratio of leaf lamina length to petiole length was determined by dividing leaf length (mm) by petiole length (mm) (Figure 3.14). The highest ratio was recorded for CHEN (11.3) and the lowest ratio for AME (0.9). A median ratio across accessions was calculated as 2.6. For CHEN, LS1, LS3, LS4, LS10, LS14 and LS15 the ratio was calculated as above 3, indicating that leaf length was three times the length of the leaf petiole, again linking to the lanceolate shape of the leaf. For the rest of the accessions, the ratio differed by small increments below 3, and do not differentiate these accessions effectively. The shape of the petiole seemed to be an informative characteristic, with accessions AME, CHEN, SCA, LS1, LS2, LS3, LS4, LS5, LS7, LS9, LS10, LS12, LS14 and LS15 having thin, un-winged petioles (Figure 3.10 a, b, c, d, e, g, I, j, l, n, o, AME, CHEN and SCA) while BUR, RET, LS6, LS8, LS11, LS13 and PC1–PC10 have thicker, winged petioles (Figure 3.10 f, h, k, m, p, q, r, s, t, u, v, w, x, y, BUR and RET).

© University of Pretoria

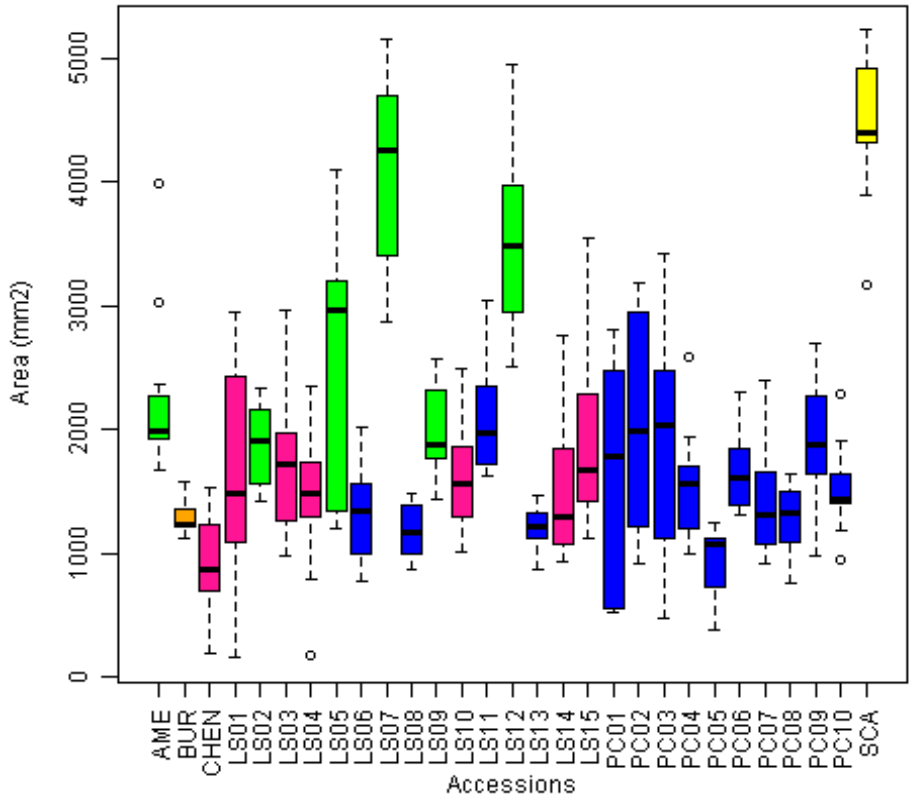


Figure 3.12 Boxplots indicating data distributions around the median for all accessions pertaining to leaf lamina area. Leaves were harvested near the top of approximately six-week old plants. Results are shown as mm^2. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

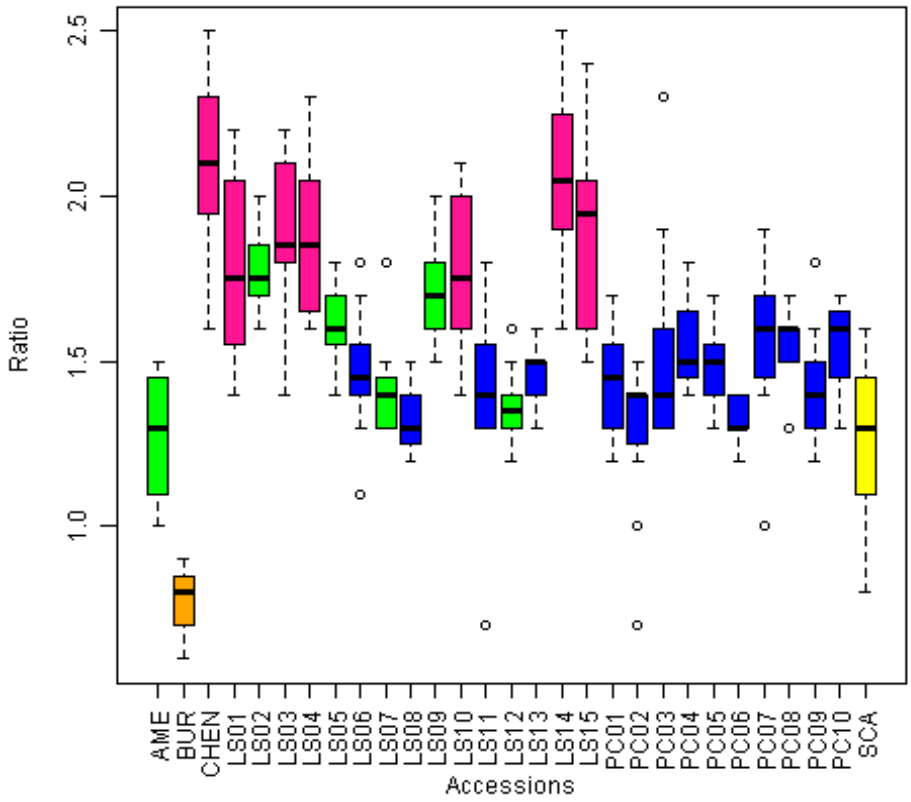


Figure 3.13 Boxplots indicating data distributions around the median for all accessions pertaining to the ratio of leaf length: leaf width. Ratios were calculated by dividing leaf length (mm) by leaf width (mm). Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

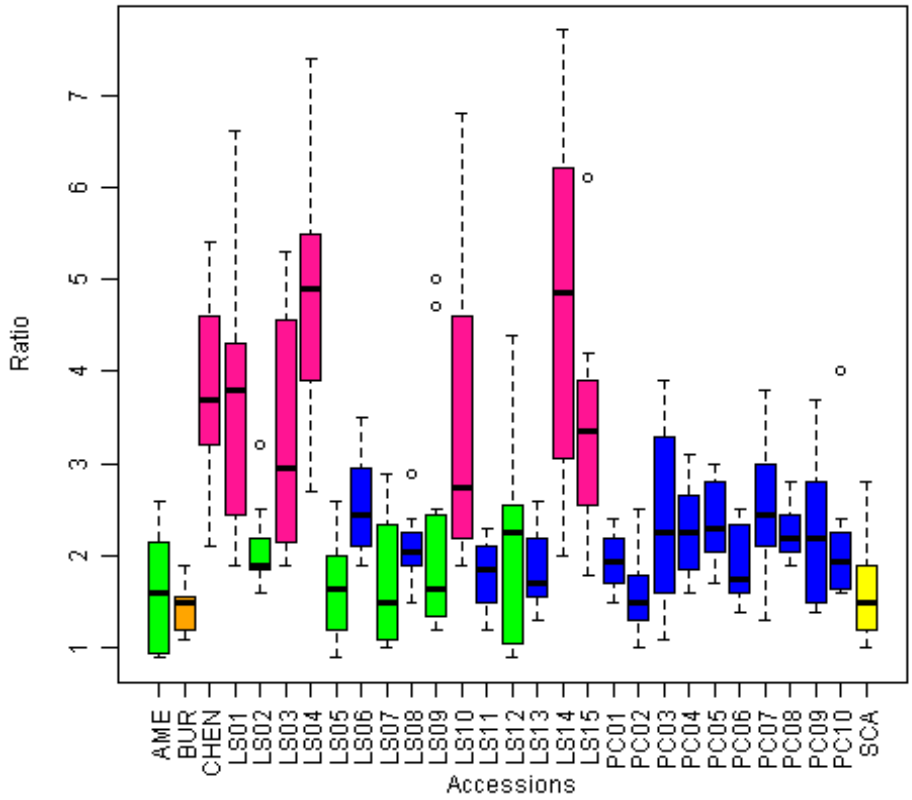


Figure 3.14 Boxplots indicating data distributions around the median for all accessions pertaining to the ratio of leaf length: petiole length. Ratios were calculated by division of leaf length (mm) by petiole length (mm). Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.2.5 Leaf SEM observations

In surface view epidermal cells of all accessions were polygonal in outline, with anisocytic stomata occurring more abundantly on abaxial than adaxial surfaces (Aliero *et al.,* 2006). CHEN, LS1, LS3, LS4, LS10, LS14 and LS15 were highly pubescent adaxially and abaxially for EG and SG (Figure 3.15 a, c, d, j, n, o, CHEN). The remaining accessions could be split into two groups with slight to moderate pubescence adaxially and abaxially. Adaxially, LS2, LS6, LS8, LS11, LS13, PC1, PC7, PC8, PC10 and BUR were slightly pubescent (Figure 3.15 b, f, h, k, m, p, v, w, y and BUR), while LS5, LS7, LS9, LS12, PC2, PC3, PC4, PC5, PC6, PC9, AME and SCA are moderately pubescent (Figure 3.15 e, g, i, l, q, r, s, t, u, x, AME and SCA). The trichome density on the adaxial side of the lamina was approximately mirrored on the abaxial side of the lamina. LS2, LS5, LS6, LS7, LS8, LS9, LS11, LS12, LS13, PC1, PC2, PC6, PC7, PC10 and AME were slightly pubescent, while PC3, PC4, PC5, PC8, PC9, BUR and SCA were moderately so. CHEN, LS1, LS3, LS4, LS10, LS14 and LS15 are highly pubescent. LG trichomes occurred sporadically on both surfaces of all accessions, but no definite pattern of occurrence could be determined. The sizes of the trichomes do not discriminate between the groups, and is probably dependant on environmental factors and/or the age of the leaf (Figure 3.16).

© University of Pretoria

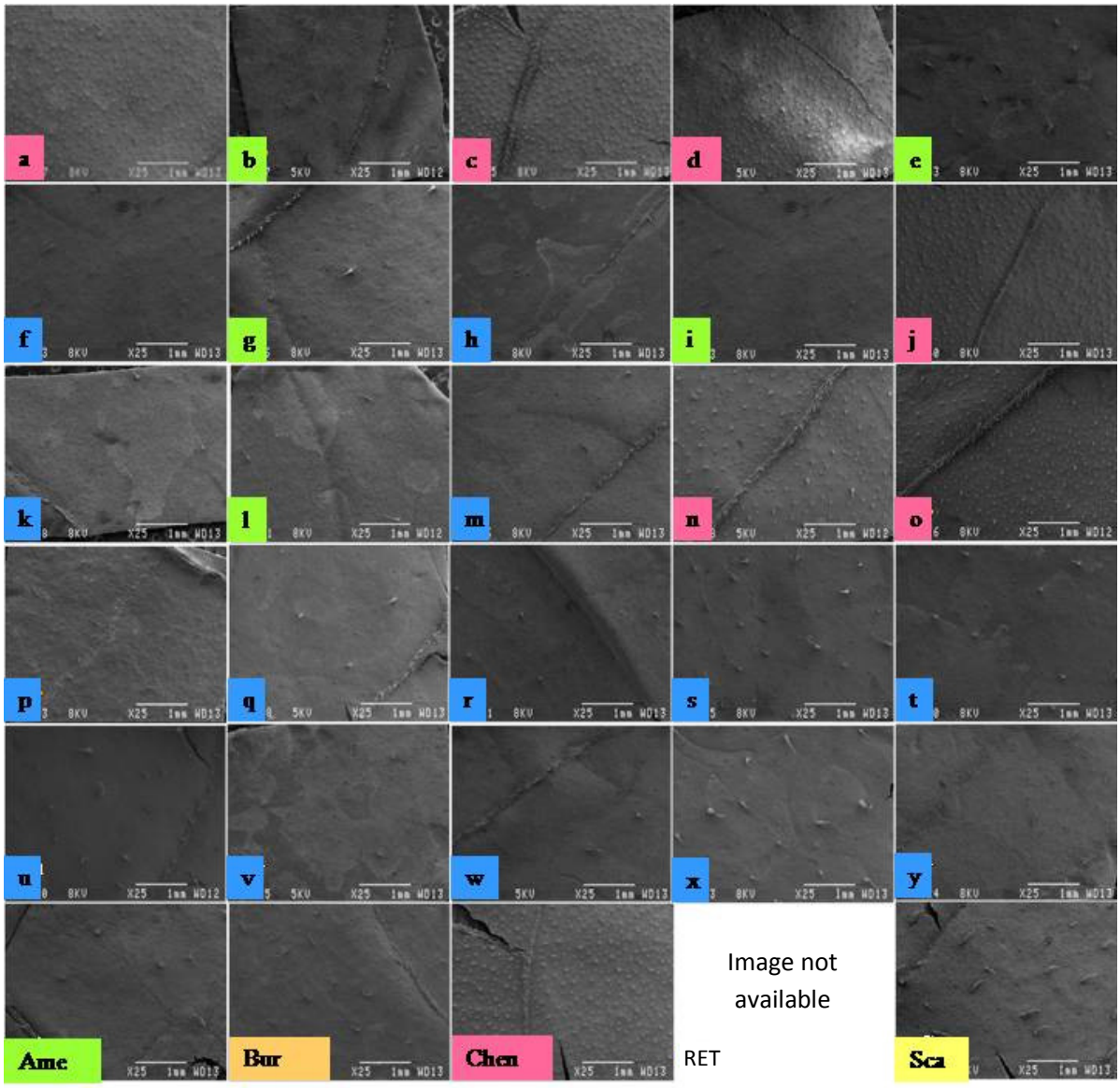


Figure 3.15 Scanning electron microscopy images of the leaf lamina (Continued to next page). Adaxial side (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) S. *retroflexum* and (SCA) *S. scabrum*. All plants were grown in similar environmental conditions, and are approximately six weeks old. All images were taken at a 25x magnification. Scale bar = 1 mm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

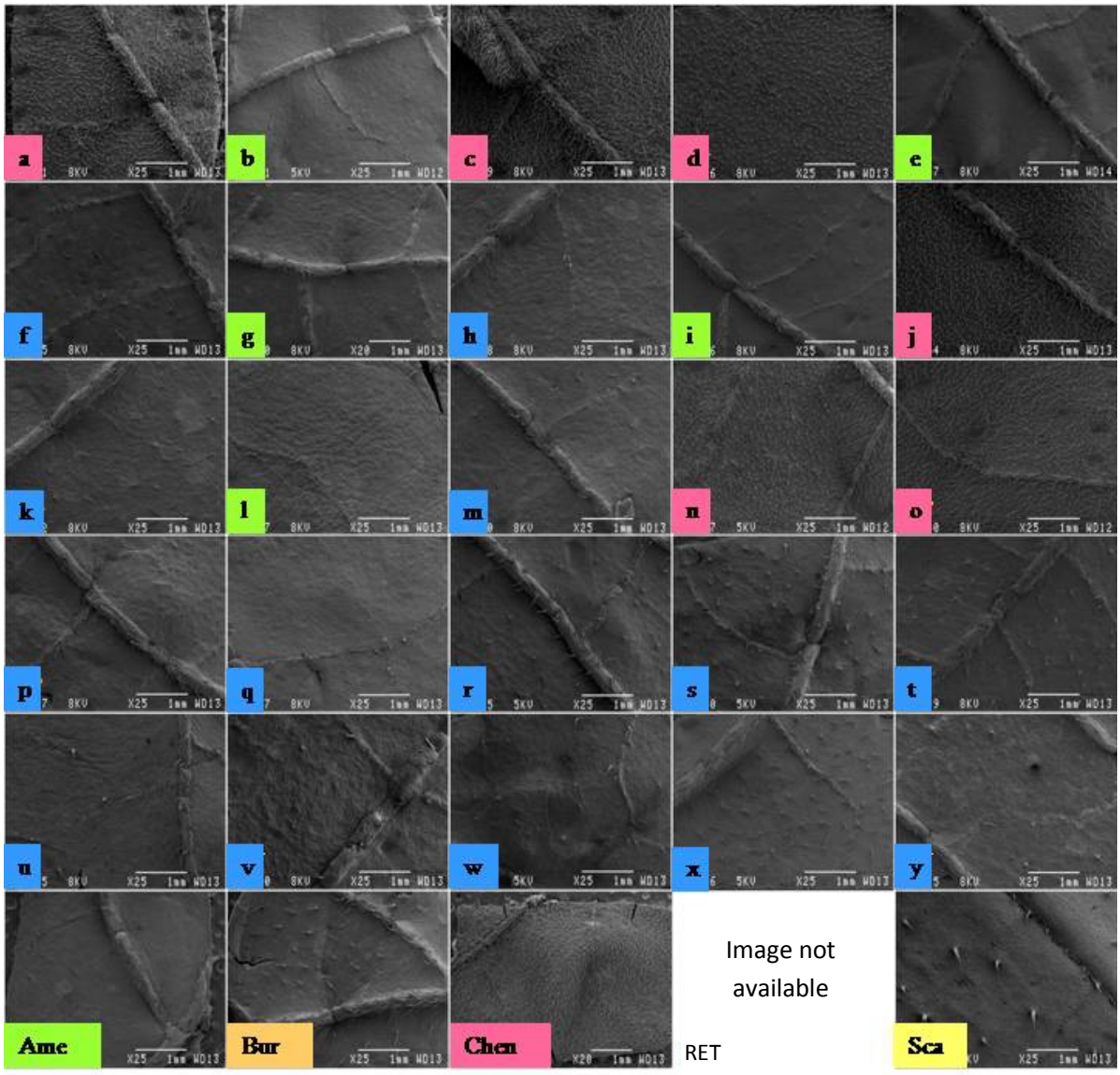


Figure 3.15 (Cont.) Scanning electron microscopy images of the leaf lamina. Abaxial side (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) S. *retroflexum* and (SCA) *S. scabrum*. All plants were grown in similar environmental conditions, and are approximately six weeks old. All images were taken at a 25x magnification. Scale bar = 1 mm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

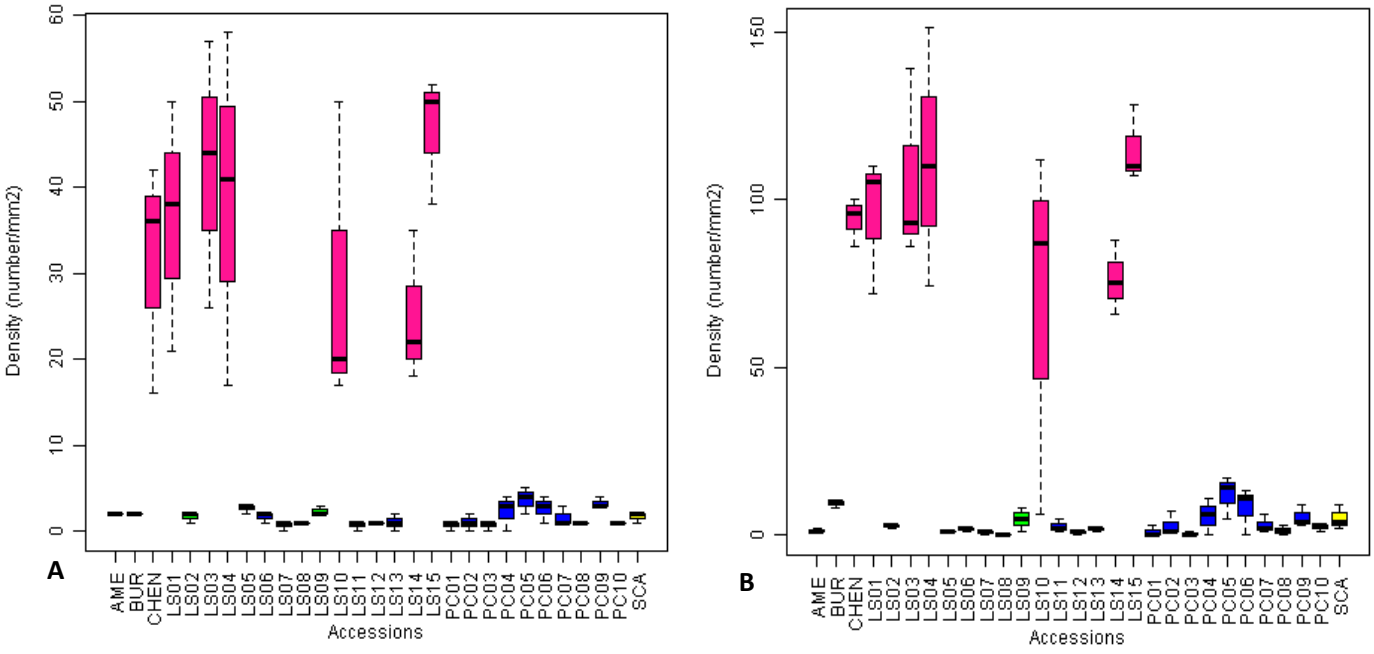


Figure 3.16 Boxplots indicating data distributions around the median for all accessions pertaining to density of trichomes on lamina surface (number/mm^2). (A) Adaxial surface (B) Abaxial surface. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

3.3.2.6 Character summary for leaves

Table 3.3 summarizes the measured and observed leaf traits as discussed above. Due to small amounts of possible environmental variation within and between plants, the table indicates the traits which appear most commonly for specific accessions as well as more occasional or rare traits. Leaf size, leaf shape and leaf margin type are informative characteristics which enable the separation of accessions into discrete groups. A similar distinction based on plant growth form alone could not be made. These visual inspections of the leaves are easy and quick to perform, and valuable for the purposes of identification.

The variation in indumentum type as well as density is summarized in Table 3.4. EG and SG trichomes occur stably across abaxial and adaxial sides of the leaf. LG may be environmentally or developmentally controlled, and occur sporadically on both sides of the leaf. Pubescence seems to group the accessions based on the density of trichomes. Even though the level of pubescence could be an informative morphological characteristic, no apparent pattern was observed for trichome sizes. The observed trichome densities corresponds well to putative identifications, with *S. chenopodioides* accessions (pink) are very

© University of Pretoria

clearly separated from the rest of the accessions, and this is an excellent trait for the identification of this species group. Varying densities of trichomes were observed for the remaining accessions, which is consistent with results based on literature (Edmonds and Chweya, 1997).

© University of Pretoria

Table 3.3 Summary of leaf morphological characters

	Leaf Shape			Leaf margin			Leaf Area				Leaf length:leaf width				Leaf length: petiole length	
	Lanceolate	Rhomboid to ovate	Ovate	Entire	Entire to sinuate	Serrate	< 1000mm^2	1000mm^2 < 2000mm^2	2000mm^2 < 3200mm^2	> 3000mm^2	<1	1<1.5	1.5<2	>2	<3	>3
LS1	XXX	-	-	XXX	-	-	XX	XXX	-	-	-	-	XXX	XX	-	XXX
LS2	-	XXX	-	-	XXX	-	XX	XXX	-	-	-	-	XXX	XX	XXX	-
LS3	XXX	-	-	XXX	-	-	X	XXX	-	-	-	-	XXX	XX	-	XXX
LS4	XXX	-	-	XXX	-	-	X	XXX	-	-	-	-	XXX	XX	-	XXX
LS5	-	XXX	-	-	XXX	-	-	-	XXX	XX	-	-	XXX	XX	XXX	X
LS6	-	XXX	-	-	-	XXX	X	XXX	-	-	-	XXX	-	-	XXX	XX
LS7	-	-	XXX	XXX	-	-	-	X	-	XXX	-	XXX	-	-	XXX	-
LS8	-	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	-	-	XXX	-
LS9	-	XXX	-	-	XXX	-	XX	-	XXX	-	-	-	XXX	-	XXX	XX
LS10	XXX	-	-	XXX	-	-	XX	XXX	-	-	-	XX	XXX	XX	-	XXX
LS11	-	XXX	-	-	-	XXX	X	-	XXX	-	-	XXX	-	-	XXX	XX
LS12	-	-	XXX	XXX	-	-	-	X	-	XXX	-	XXX	-	-	XXX	-
LS13	-	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	-	-	XXX	-
LS14	XXX	-	-	XXX	-	-	-	XXX	-	-	-	-	XXX	-	-	XXX
LS15	XXX	-	-	XXX	-	-	-	XXX	-	-	-	-	XXX	-	-	XXX
PC1	-	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XX	-	XXX	-
PC2	-	XXX	-	-	-	XXX	-	XXX	XX	-	-	XXX	XX	-	XXX	-
PC3	-	XXX	-	-	-	XXX	-	XXX	XX	-	-	XXX	XX	-	XXX	-
PC4	-	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XX	-	XXX	-
PC5	-	XXX	-	-	-	XXX	XX	XXX	-	-	-	XXX	XX	-	XXX	-
PC6	-	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XX	-	XXX	-
PC7	-	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XX	-	XXX	-
PC8	-	XXX	-	-	-	XXX	XX	XXX	-	-	-	XXX	XX	-	XXX	-
PC9	-	XXX	-	-	-	XXX	-	XXX	XX	-	-	XXX	XX	-	XXX	-
PC10	-	XXX	-	-	-	XXX	-	XXX	XX	-	-	XXX	XX	-	XXX	-
AME	-	-	XXX	XXX	-	-	-	XXX	XX	-	XXX	XX	-	-	XXX	-
BUR	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX	-	-	-	XXX	-
CHEN	XXX	-	-	XXX	-	-	XXX	XX	-	-	-	-	-	-	XXX	X
RET	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SCA	XX	-	XXX	XXX	-	-	-	-	-	XXX	-	-	-	-	XXX	-

XXX Usually; XX Occasionally; X Rarely; - Never

© University of Pretoria

Table 3.4 Leaf indumentum morphology

	Leaf Texture						Trichomes present on leaf						Length of eglandular trichomes					
	Not pubescent		Moderately pubescent		Pubescent		Short glandular		Long glandular		Eglandular		< 0.1 mm		0.1mm < 0.2mm		> 0.2mm	
	Ad*	Ab*	Ad	Ab	Ad	Ab	Ad	Ab	Ad	Ab	Ad	Ab	Ad	Ab	Ad	Ab	Ad	Ab
LS1	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-
LS2	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	XXX	XXX	-
LS3	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-
LS4	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	X	XXX	XXX	X	-
LS5	-	XXX	XXX	-	-	-	XXX	XXX	-	XXX	XXX	X	-	-	-	XXX	XXX	-
LS6	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	XX
LS7	-	XXX	XXX	-	-	-	XXX	XXX	-	-	X	XXX	-	-	-	-	XXX	XXX
LS8	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	X	-	-	X	-	XXX	-
LS9	-	XXX	XXX	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	XXX	XXX	X
LS10	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	X	-
LS11	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	XXX	XXX	-
LS12	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
LS13	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	X	XXX	XXX
LS14	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	X
LS15	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-
PC1	XXX	XXX	-	-	-	-	XXX	XXX	XXX	-	XXX	XXX	-	-	-	-	XXX	XXX
PC2	-	XXX	XXX	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
PC3	-	-	XXX	XXX	-	-	XXX	XXX	XXX	XXX	XXX	X	-	-	-	-	XXX	XXX
PC4	-	-	XXX	XXX	-	-	XXX	XXX	XXX	XXX	XXX	XXX	-	-	-	XXX	XXX	-
PC5	-	-	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
PC6	-	XXX	XXX	-	-	-	XXX	XXX	XXX	XXX	XXX	XXX	-	-	-	XXX	XXX	-
PC7	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
PC8	XXX	-	-	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
PC9	-	-	XXX	XXX	-	-	XXX	XXX	XXX	-	XXX	XXX	-	-	-	-	XXX	XXX
PC10	XXX	XXX	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
S. americanum	-	XXX	XXX	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	XXX	XXX	-
S. burbankii	XXX	-	-	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX
S. chenopodioides	-	-	-	-	XXX	XXX	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-
S. retroflexum	-	-	-	-	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	-	-
S. scabrum	-	-	XXX	XXX	-	-	XXX	XXX	-	-	XXX	XXX	-	-	-	-	XXX	XXX

XXX Usually; XX Occasionally; X Rarely; - Never
*Ad – Adaxial side of lamina *Ab – Abaxial side of lamina

© University of Pretoria

3.3.3 Flowers

3.3.3.1 Corolla

All accessions had white corollas, with a slight purple stripe in accessions PC1–PC10. The stripe was not observed for every flower on each plant. Based on existing literature, the presence of a stripe could be attributed to variations in light intensities within the glasshouse or in the field and it is not generally accepted as a diagnostic morphological trait (Edmonds, 1972).

Flower corollas were distinguished between different accessions based on size and shape (Figure 3.17). Corolla ratios were determined by dividing corolla lobe length (mm) by corolla lobe width (mm). The ratio of corolla lobe length to corolla lobe width separated the accessions into two groups (Figure 3.19). The highest corolla ratio was observed for LS15 (3.54), while the lowest calculated ratio was AME (0.83). Corolla ratios across all accessions had a median of 2.2. The first group was characterized by a ratio < 2 indicating a relatively symmetrical corolla type. The second group has a ratio of > 2 where the length is more than twice the width of the corolla lobes, and is generally larger than the first group. The first group consisted of LS2, LS5, LS7, LS9, LS11, LS12, LS13, PC1-PC10, AME, BUR and SCA (Figure 3.17 b, e, g, i, k, l, m, p, q, r, s, t, u, v, w, x, y, AME, BUR and SCA). The second group included LS1, LS3, LS4, LS6, LS8, LS10, LS14, LS15 and CHEN (Figure 3.17 a, c, d, f, h, j, n, o and CHEN) (Figure 3.19). The corollas of all accessions could be assigned to the stellate type, but could be clustered into two groups based on whether the corolla lobes fuse at, or away from, the point of stamen insertion. CHEN, LS1, LS2, LS3, LS4, LS5, LS9, LS10, LS14 and LS15 had corolla lobes fused at the point of anthers insertion, whiles all the other accessions had corolla lobes fused halfway to two-thirds above the point of anther insertion (Figure 3.18).

© University of Pretoria

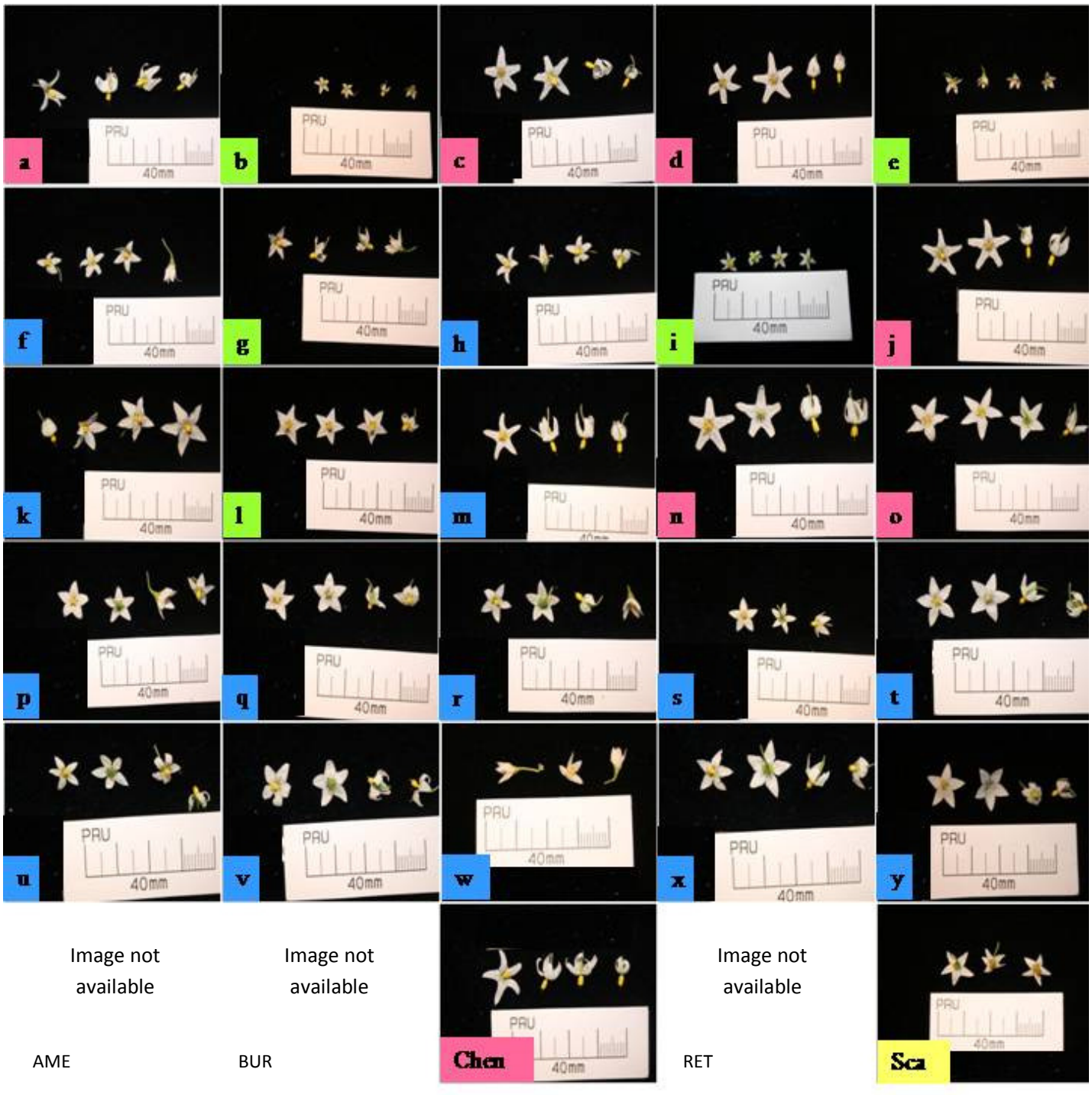


Figure 3.17 Flower morphology for all nightshade accessions. (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x): PC9; (y) PC10; (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) S. *retroflexum* and (SCA): *S. scabrum*. All plants were grown in similar environmental conditions, and are approximately six weeks old. Scale bar = 40 mm. Two flowers were photographed from the top and two from the side. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

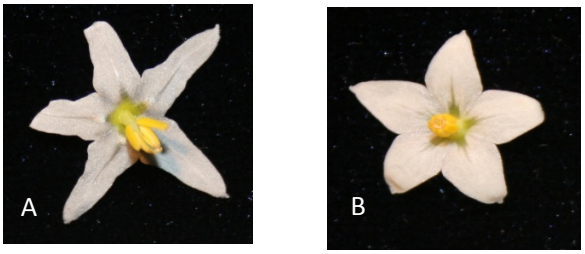


Figure 3.18 Corolla lobe fusion morphology. (A) Corollas fused close to the origin of the anthers. (B) Corolla lobes fused halfway to two thirds to the anthers.

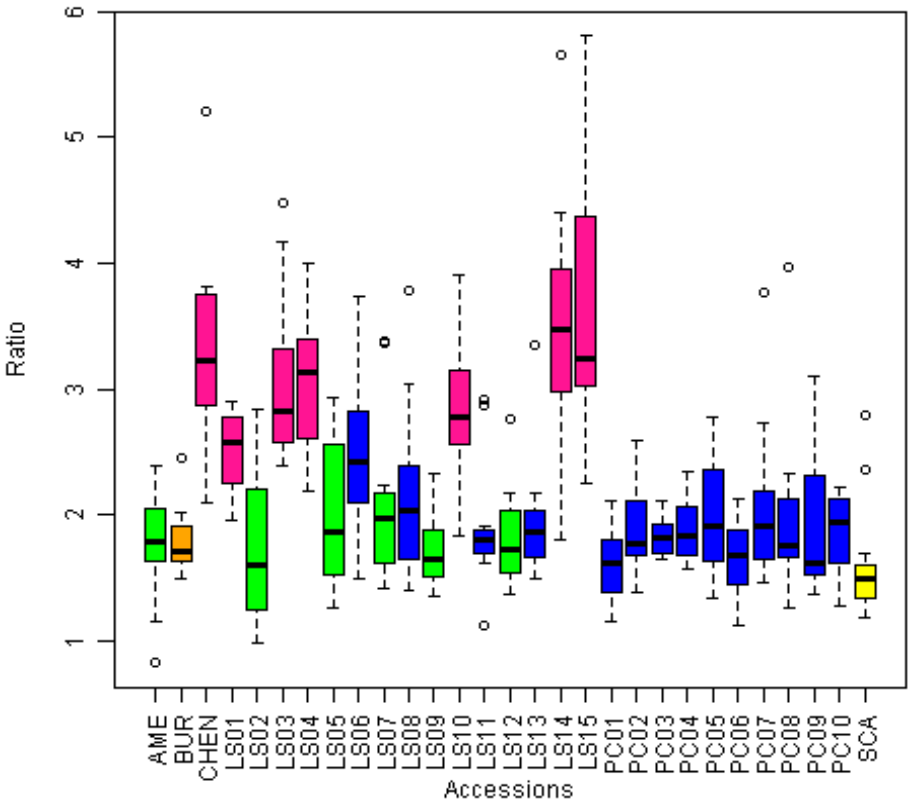


Figure 3.19 Boxplots indicating data distributions around the median for all accessions pertaining to the ratio of corolla lobe length: corolla lobe width. Ratios were calculated by the division of corolla length by corolla width. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.3.2 Style and anthers

Style exertion above the level of the anthers had been suggested as a possible diagnostic morphological trait, and was therefore included in the present study (Edmonds, 1977). Accessions SCA, LS2, LS5 and LS9 did not display the style exertion at any developmental stage (Figure 3.20 A) whiles the remaining accessions had clear style extensions. AME, BUR, LS6, LS7, LS8, LS11, LS12, LS13 and PC1–PC10 had styles exerted for approximately 1 mm above the anthers (Figure 3.20 B). Prominent style exertions were observed for CHEN, LS1, LS3, LS4, LS10, LS14 and LS15, with roughly 2 mm extended beyond the anthers (Figure 3.20 C). A prominent yellow colour was observed for the anthers of all accessions (Figure 3.20 B and C), except SCA (Figure 3.20 A). The latter was visibly different from the rest of the accessions in having brown coloured anthers.

Figure 3.20 Style protrusions above anthers. (A) no protrusion (B) slight protrusion (±1 mm) (C) prominent protrusion (±2 mm)

3.3.3.3 Inflorescences

Peduncle and pedicel morphology were diagnostic in this complex. CHEN, LS1, LS2, LS3, LS4, LS5, LS9, LS10, LS14 and LS15 had peduncles growing straight up from the stem apex, as well as an erect pedicel type (Figure 3.21 A). AME, BUR, LS6, LS7, LS8, LS11, LS12, LS13 and PC1-PC10 had deflexed peduncles growing out of the stem apex, and reflexed pedicels (Figure 3.21 B). SCA was distinguished from the other accessions in having erect, branched peduncle growing out of the stem apex (Figure 3.21 C). Usually there were 2–3 peduncle branches.

© University of Pretoria

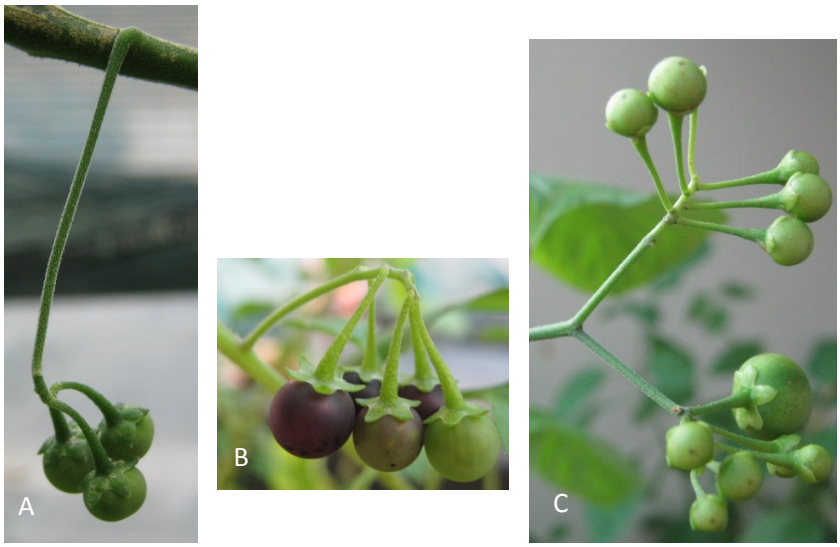


Figure 3.21 Peduncle and pedicel growth (A) straight peduncle growing out of the stem apex, erect pedicel type (B) deflexed peduncles out of the stem apex, reflexed pedicel type (C) branched, erect peduncle out of stem apex, erect pedicel type.

3.3.3.4 Character summary for flowers

Black nightshade flowers seemed to provide a rich amount of diagnostic information. The size of the flowers (length to width ratio) could be utilized as an initial descriptor, whilst the style, anther and inflorescence provide subsequent separation into possible species groups. Table 3.5 indicates the variation within an accession, with occasional and rare corolla shapes, style exertions and inflorescence types occurring on an individual plant. The results observed for flower morphology was consistent with the putative identification of accessions, except for LS7, LS12 and AME which were expected to have no style protrusion.

© University of Pretoria

Table 3.5 Summary of floral morphological characters

	Flower shape		Corolla color		Anther color		Style protruding		Corolla Length: width		Inflorescence type	
	Stellate; Deeply lobed	Stellate; Shallowly lobed	White	White with purple stripe	Yellow	Brown	Above anthers	Not above anthers	< 2	> 2	Drooping cymes	Erect cymes
LS1	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
LS2	XXX	-	XXX	-	XXX	-	-	XXX	XXX	X	-	XXX
LS3	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
LS4	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
LS5	XXX	-	XXX	-	XXX	-	-	XXX	XXX	XX	-	XXX
LS6	-	XXX	XXX	-	XXX	-	XXX		XX	XXX	XXX	-
LS7	-	XXX	XXX	-	XXX	-	XXX	X	XXX	XX	XXX	-
LS8	-	XXX	XXX	-	XXX	-	XXX	-	XX	XXX	XXX	-
LS9	XXX	-	XXX	-	XXX	-	-	XXX	XXX	X	-	XXX
LS10	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
LS11	-	XXX	XXX	-	XXX	-	XXX	-	XXX	XX	XXX	-
LS12	-	XXX	XXX	-	XXX	-	XXX	X	XXX	XX	XXX	-
LS13	-	XXX	XXX	-	XXX	-	XXX	-	XXX	XX	XXX	-
LS14	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
LS15	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
PC1	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC2	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC3	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC4	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC5	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC6	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC7	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC8	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC9	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
PC10	-	XXX	XXX	XX	XXX	-	XXX	-	XXX	X	XXX	-
AME	-	XXX	XXX	-	XXX	-	XXX	-	XXX	XX	XXX	-
BUR	-	XXX	XXX	-	XXX	-	XXX	-	XXX	X	XXX	-
CHEN	XXX	-	XXX	-	XXX	-	XXX	-	X	XXX	-	XXX
RET	-	-	-	-	-	-	-	-	-	-	-	-
SCA	-	XXX	XXX	-	-	XXX	-	XXX	XXX	X	-	XXX

XXX Usually; XX Occasionally; X Rarely; - Never

© University of Pretoria

3.3.4 Berries

3.3.4.1 Fruit colour and size

All fruits (berries) were ovoid and green when immature, gradually changing during ripening from green to slightly purple to dark purple-black at maturity. The berries differed with respect to shininess of their cuticles. LS2, LS5 and LS9 were extremely shiny (Figure 3.22 b, e, i), while AME, LS7 and LS12 were moderately shiny (Figure 3.22 g, l, AME). LS1, LS3, LS4, LS6, LS8, LS10, LS11, LS13, LS14, LS15, PC1-PC10, CHEN, BUR and SCA had a dull cuticle (Figure 3.22 a, c, d, f, h, j, k, m, n, o, p, q, r, s, t, u, v, w, x, y, CHEN, BUR and SCA). Berry diameter sizes ranged from a minimum of 4 mm to a maximum of 15 mm, with a median across accessions of 6.67 mm. CHEN, LS1, LS2, LS3, LS4, LS5, LS10, LS14 and LS15 had berry diameter sizes of less than 6 mm (Figure 3.22 a, b, c, d, e, j, n, o, CHEN), while LS6, LS7, LS8, LS9, LS11, LS12, LS13, PC1-PC10, AME and BUR had berry sizes of between 6 mm and 10 mm (Figure 3.22 f, g, h, I, k, l, m, p, q, r, s, t, u, v, w, x, y, AME and BUR). SCA had berry diameter sizes of between 10 mm and 15 mm, and was easily distinguished from the rest of the accessions by their significantly larger size (Figure 3.23).

3.3.4.2 Origin and number of pedicels

The branching pattern and number of pedicels originating per peduncle was observed to be a possible diagnostic trait. In terms of the origin of the pedicels, CHEN, LS1, LS2, LS3, LS4, LS5 LS9, LS10, LS14 and LS15 had pedicel branches originating from a central point on the peduncle (Figure 3.21 A). LS6, LS7, LS8, LS11, LS12, LS13, PC1–PC10, AME, SCA and BUR had pedicels which arise from different points, alternately down the peduncle (Figure 3.21 B). The number of pedicels differed per inflorescence, with LS2, LS5, LS8, LS11, LS13 having less than four and LS7, LS12, AME, SCA having more than seven pedicels per peduncle. The number of pedicels was not a diagnostic trait, as LS1, LS3, LS4, LS6, LS9, LS10, LS14, LS15, PC1–PC10, BUR and CHEN varied in having between four and seven pedicels per peduncle.

© University of Pretoria

3.3.4.3 Calyx

The calyx was either reflexed or adherent. CHEN, AME, LS1, LS3, LS4, LS7, LS10, LS12, LS14 and LS15 have the adherent state, with the calyx fitting tightly around the mature berry (Figure 3.22 a, c, d, g, j, l, n, o, CHEN and AME). The calyx was reflexed away from the mature berry in SCA, BUR, LS2, LS5, LS6, LS8, LS9, LS11, LS13 and PC1-PC10 (Figure 3.22 b, e, f, h, i, k, m, p, q, r, s, t, u, v, w, x, y, BUR and SCA). All the berries except for SCA detached from their calyx when ripe. The latter remained adhered upon maturity.

© University of Pretoria

Figure 3.22 Berry morphology for all nightshade accessions. (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x): PC9; (y) PC10; (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) S. *retroflexum* and (SCA): *S. scabrum*. All plants were grown in similar environmental conditions, and are approximately six weeks old. Scale = 40 mm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

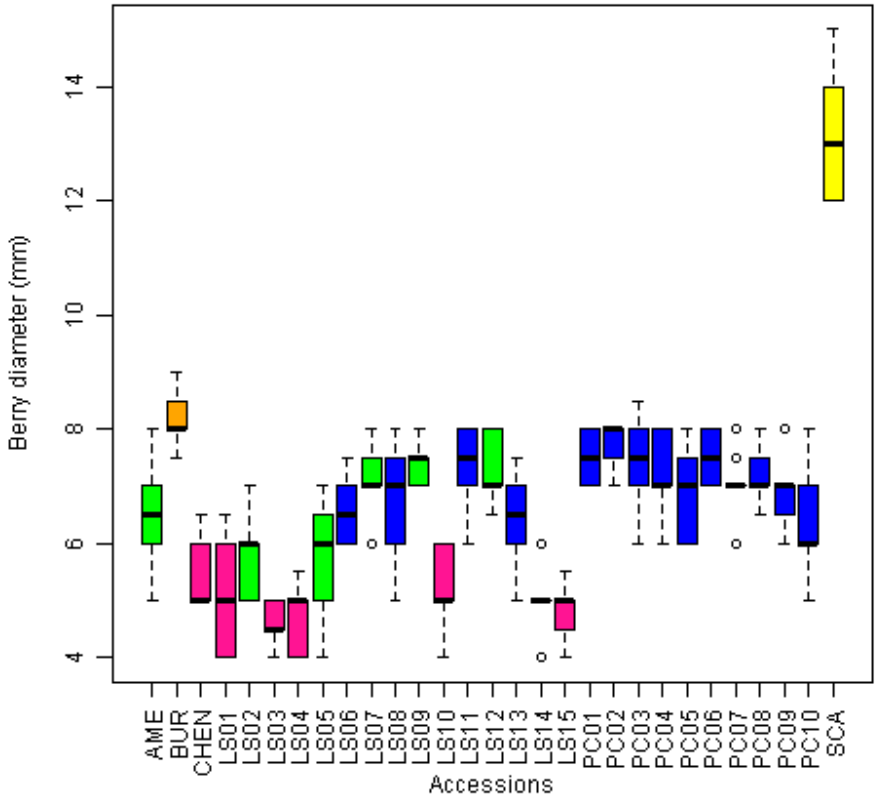


Figure 3.23 Boxplots indicating data distributions around the median for all accessions pertaining to berry size. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.4.4. Berry SEM morphology

Accessions were grouped based on the glossiness of the berries by visual interpretation. SEM of berry surfaces was only done for one representative from each of these groups. The first group was classified as dull (Figure. 3.24 a), the second as moderately shiny (Figure. 3.24 b) and the third group as very shiny (Figure. 3.24 c). The cuticle surface for dull berry types indicated the presence of a dense "mat"-like structure, which may be absorbing light instead of reflecting it, thus creating a dull appearance. Moderately shiny has the general striated exterior seen in dull berries, but seems to be ordered into polygonal shapes. Shiny berries have an undulating, smooth appearance, which may reflect the most amount of light resulting in the more polished look.

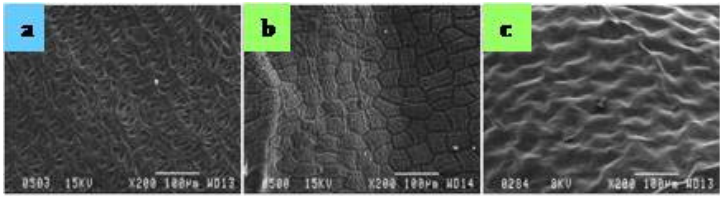


Figure 3.24 Scanning electron microscopy images of selected berry surfaces. (a) PC5–Dull (b) LS7–Moderately shiny (c) LS9–Shiny. All plants were grown in identical environmental conditions, and were approximately six weeks old. All images were taken at a 200x magnification. Scale bar = 100 µm. Putative identities are colour coded; blue = *S. retroflexum,* green = *S. americanum*

3.3.4.5 Character summary

The cuticle appearance of the berries was a diagnostic morphological trait, which formed three distinct groups (shiny, moderately glossy and dull). The size of the berry was also useful and could be utilized as an informative characteristic to classify small berries into one group (putatively *S. chenopodioides*), while SCA was different from the rest of the accessions in having the largest berries. Pedicel arrangement on the peduncle, as well as the amount of pedicels per peduncle was not deemed a diagnostic characteristic with significant variation occurring within and between accessions. Calyx adherence could be a primary diagnostic trait; even though it did not separate all species into groups it generates initial identifications. Additional observations are needed to determine if berry surface morphology can be deemed a significant diagnostic trait. Common, occasional and rare forms of these morphological traits are indicated for each accession in Table 3.6.

© University of Pretoria

Table 3.6 Summary of berry morphological characters

	Polish of mature fruit			Fruit size			Nr of pedicels			Arrangement of pedicels		Calyx orientation	
	Dull	Moderately shiny	Shiny	< 6mm	6mm < 10mm	> 10mm	< 4	4–7	> 7	Same origin	Alternate origins	Reflexed	Adherent
LS1	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
LS2	-	-	XXX	XXX	XX	-	XXX	XX	-	XXX	-	XXX	-
LS3	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
LS4	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
LS5	-	-	XXX	XXX	XX	-	XXX	XX	-	XXX	-	XXX	-
LS6	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
LS7	-	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	XXX
LS8	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XXX	-
LS9	-	-	XXX	-	XXX	-	-	XXX	-	XXX	-	XXX	-
LS10	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
LS11	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XXX	-
LS12	-	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	XXX
LS13	XXX	-	-	-	XXX	-	XXX	-	-	-	XXX	XXX	-
LS14	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
LS15	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
PC1	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC2	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC3	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC4	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC5	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC6	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC7	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC8	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC9	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
PC10	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
S. americanum	-	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	XXX
S. burbankii	XXX	-	-	-	XXX	-	-	XXX	-	-	XXX	XXX	-
S. chenopodioides	XXX	-	-	XXX	-	-	-	XXX	-	XXX	-	-	XXX
S. retroflexum	-	-	-	-	-	-	-	-	-	-	-	-	-
S. scabrum	-	XXX	-	-	-	XXX	-	-	XXX	-	XXX	XXX	-

XXX Usually; XX Occasionally; X Rarely; - Never

© University of Pretoria

3.3.5 Seeds

3.3.5.1 Seed size

Seeds from all accessions had colours ranging from tan to light brown. The diameter size ranged from 0.68 mm to 2.34 mm, with a median across accessions of 1.52 mm. LS4 had the smallest seeds of 0.98 mm, while PC10 had the largest seeds with 2.34 mm. CHEN, LS1, LS2, LS3, LS5, LS9, LS10, LS14 and LS15 had seed sizes between 1 mm and 1.3 mm (Figure 3.25 a, b, c, e, i, j, n, o and CHEN). LS7, LS12 and AME had seed sizes of on average, between 1.3 mm and 1.6 mm (Figure 3.25 g, l and AME). The largest seeds ranged between 1.6 and 2.3 mm, and included LS6, LS8, LS11, LS13, PC1-PC10, BUR and SCA (Figure 3.25 f, h, k, m, p, q, r, s, t, u, v, w, x, y, BUR and SCA). Based on seed size, three clear groups could be distinguished (Figure 3.26).

3.3.5.2 Seeds per berry

The number of seeds per berry was inversely correlated to the seed size. There are fewer seeds per berry when they are larger than 1.5 mm and more seeds per berry if they are smaller than 1.5 mm. Number of seeds per berry ranged from 5–98, with a median of 24 seeds per berry across all accessions. LS6, LS8, LS11, LS13, PC1, PC3, PC4, PC5, PC6, PC7, PC8, PC9, PC10 and BUR had the least amount of seed per berry, with less than 20. LS4, LS7, LS12, PC2, AME and CHEN generally contained 20–30 seeds per berry, while LS1, LS2, LS3, LS5, LS9, LS10, LS14 and LS15 contained between 30–40 seeds per berry. SCA had the largest number of seeds per berry, with 98 in a single berry (Figure 3.27), but does not have an inverse correlation with berry size. This leads to the hypothesis that certain species have adapted their berry size to accommodate larger, although fewer, seeds.

3.3.5.3 Ratio of seed number vs. berry size

Ratios of seed number to berry diameter size were calculated by dividing the number of seeds by the berry size (mm). Ratios ranged from 0.6–10, with an overall median of 3.8. A ratio of < 2 was recorded in LS6, LS8, LS11, LS13, PC1, PC3, PC4, PC6, PC7, PC8, PC9, PC10 and BUR. A higher ratio of 2–4 was determined for CHEN, AME, LS7, LS12, PC2 and PC5. The highest ratio was calculated for LS1, LS2,

© University of Pretoria

LS3, LS4, LS5, LS9, LS10, LS14 and LS15, adding confidence to the hypothesis that smaller berries have more seeds which are smaller in size. Lower ratios for all the other accessions showed larger berries have fewer seeds of larger size (Figure 3.28). SCA is the exception, where seed size seems to be directly correlated to the amount of seed per berry and the berry size. Relationships between berry sizes, seed sizes and amount of seed per berry are indicated in Figure 3.29.

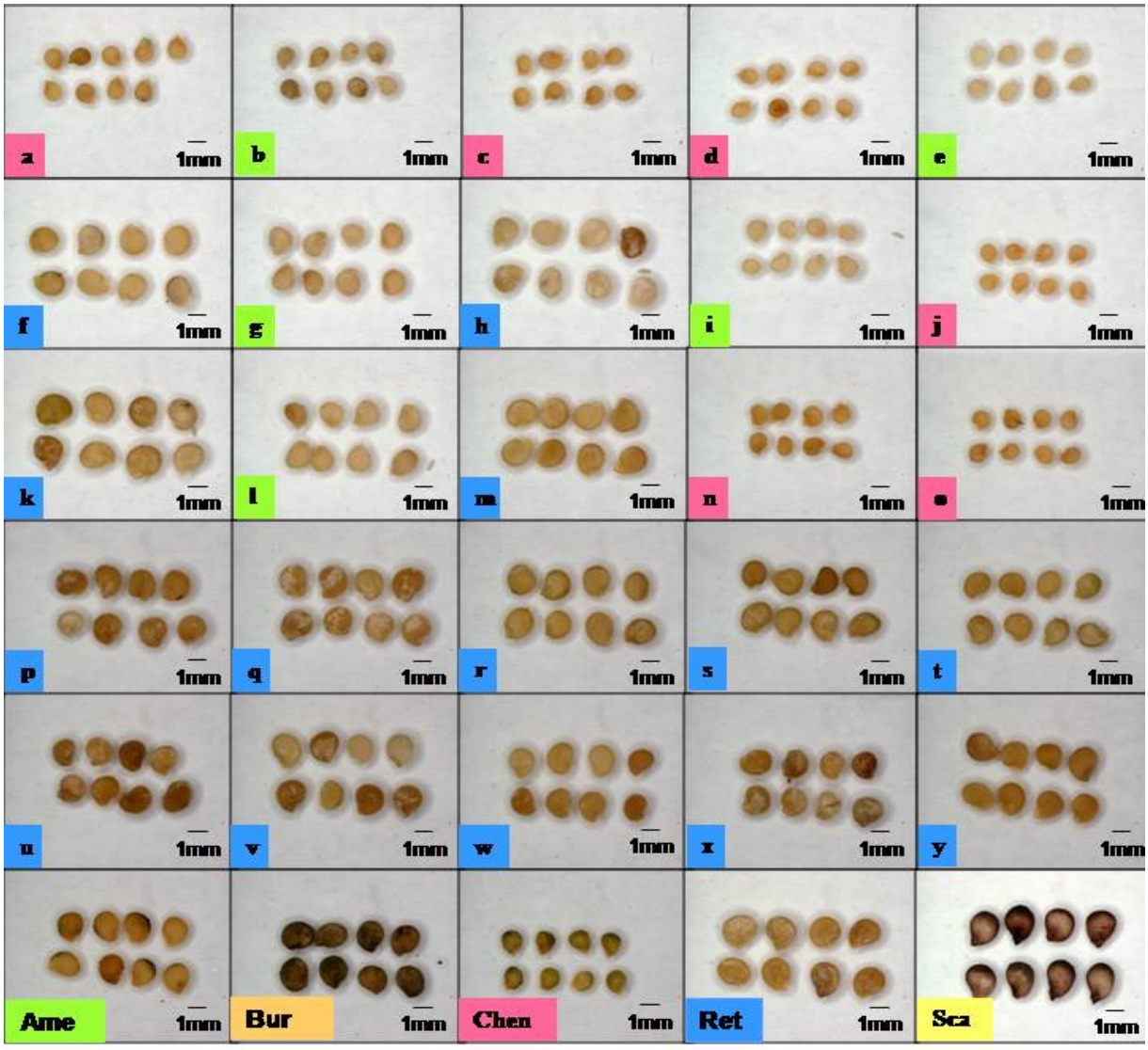


Figure 3.25 Seed morphology for all nightshade accessions. (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum;* (BUR) *S. burbankii;* (CHEN) *S. chenopodioides;* (RET) S. *retroflexum* and (SCA) *S. scabrum*. Seeds were washed with distilled water and air dried overnight at 37°C. Scale = 1 mm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

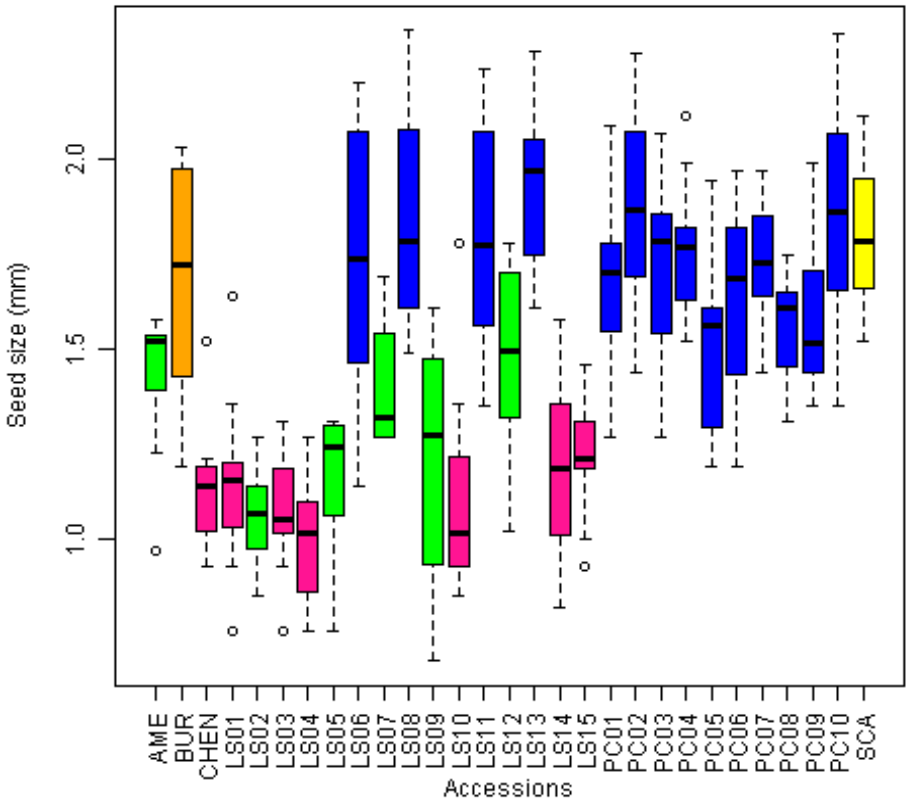


Figure 3.26 Boxplots indicating data distributions around the median for all accessions pertaining to seed size. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

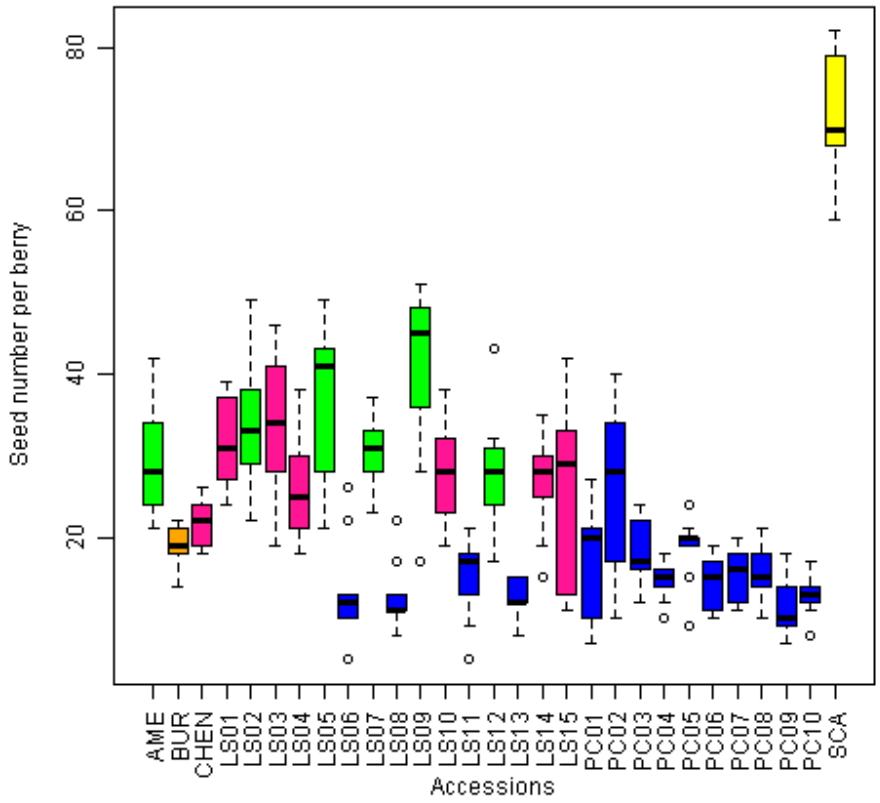


Figure 3.27 Boxplots indicating data distributions around the median for all accessions pertaining to number of seeds per berry. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

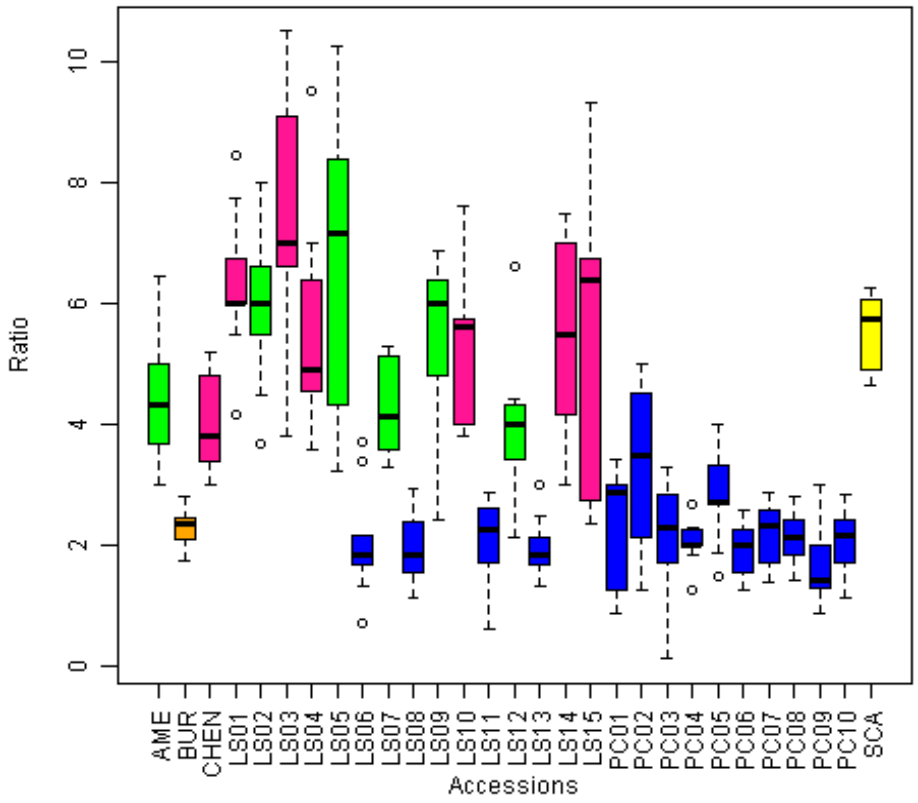


Figure 3.28 Boxplots indicating data distributions around the median for all accessions pertaining to ratio of seed number to berry size. Ratios were calculated as amount of seed per berry divided by the size of the berry. Circles indicate outlier values. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

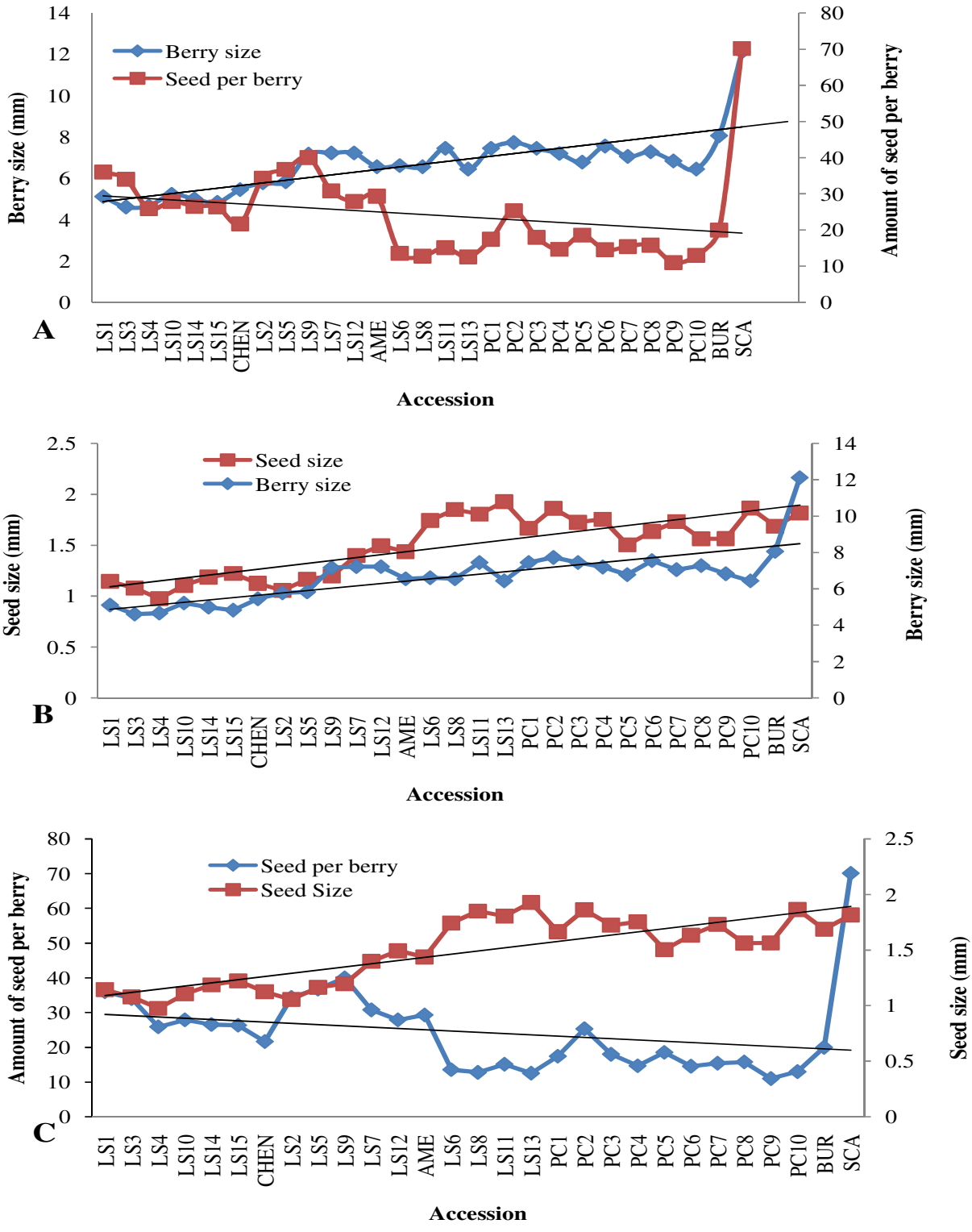


Figure 3.29 Graphical representations of seed relationships. (A) Berry size is inversely correlated to the number of seed per berry (B) The seed size is directly correlated to berry size (C) Amount of seed per berry is inversely correlated to seed size.

© University of Pretoria

3.3.5.4 Seed SEM morphology

Seed coat patterns under SEM were uniform across accessions. The seeds could be described as flattened and biconvex in section. The strands surrounding each epidermal pore support a thin outer cell wall. The strand-like structures could have been formed by the mechanical removal of the outer cell wall (Figure 3.30). Seed coat patterns were not informative in taxonomic classifications of the accessions.

3.3.5.5 Character summary

Seed size is as an important trait to distinguish between nightshade accessions. Based on sizes, three distinct groups are demarcated. The number of seeds per berry is correlated to seed size as well as berry size. Smaller seeds tend to appear in larger numbers in smaller berries, whereas larger berries contain fewer, but larger seeds. This pattern is also mirrored in the ratios calculated for the number of seeds per berry. The exception is SCA, which shows a direct correlation of amount of seed per berry and seed size. Seed sizes observed correlated with putative accession identifications. SEM observations of seed coat micro-structures did not reveal any taxonomically significant patterns. Common, occasional and rare seed traits are summarized for each accession in Table 3.7.

© University of Pretoria

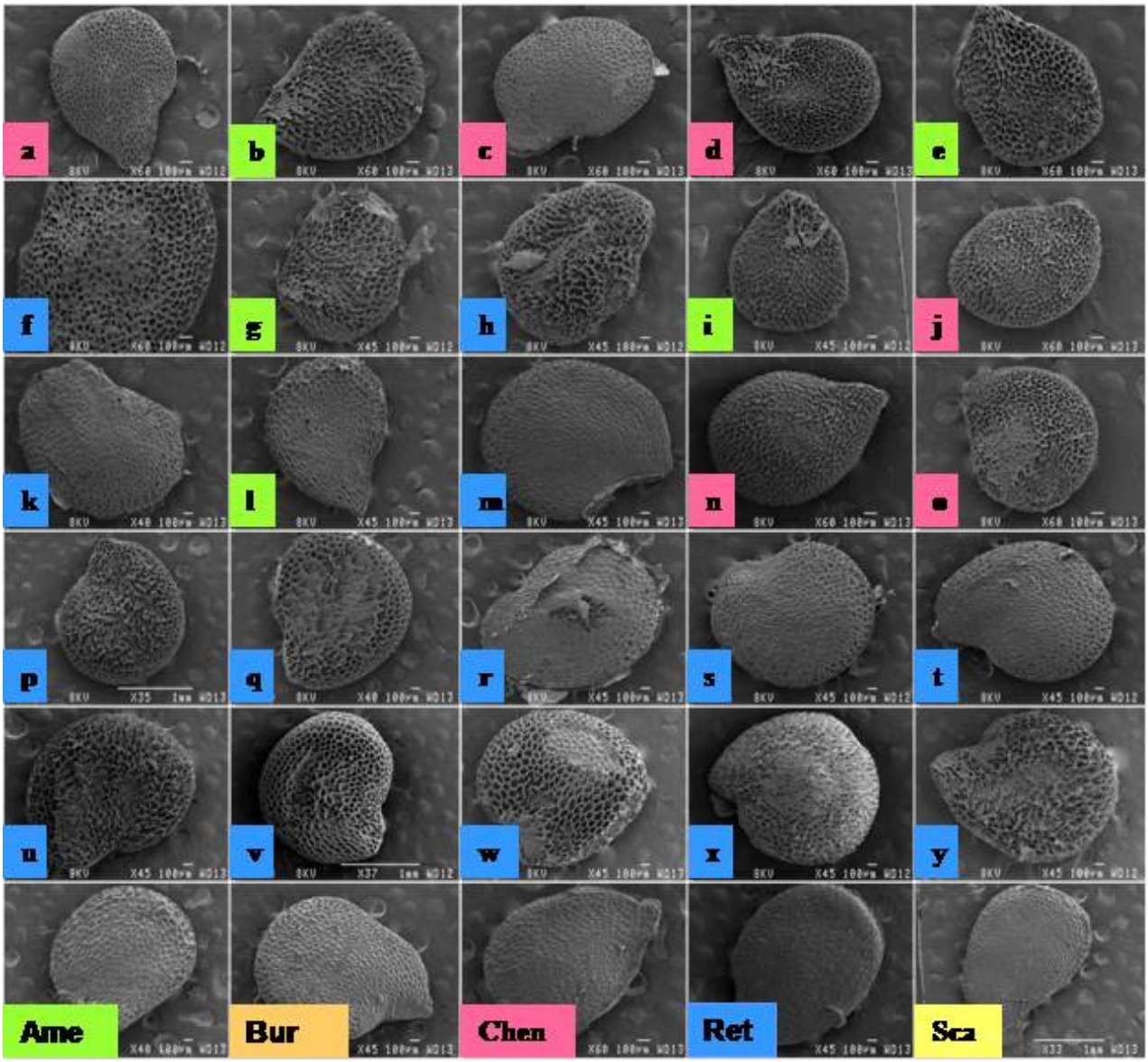


Figure 3.30 Seed SEM morphology for all nightshade accessions. (a) LS1; (b) LS2; (c) LS3; (d) LS4; (e) LS5; (f) LS6; (g) LS7; (h) LS8; (i) LS9; (j) LS10; (k) LS11; (l) LS12; (m) LS13; (n) LS14; (o) LS15; (p) PC1; (q) PC2; (r) PC3; (s) PC4; (t) PC5; (u) PC6; (v) PC7; (w) PC8; (x) PC9; (y) PC10; (AME) *S. americanum;* (BUR) *S. burbankii;* (CHEN) *S. chenopodioides;* (RET) S. *retroflexum* and (SCA) *S. scabrum*. Scale = 100 µm. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

Table 3.7 Summary of seed morphological characters

	Seed size				Number of seed per berry				Seed number: berry size			
	< 1	1 < 1.3	1.3 < 1.6	1.6 < 1.9	< 20	20 < 30	30 < 40	> 40	< 2	2 < 4	4 < 6	> 6
LS1	-	XXX	X	-	-	X	XXX	-	-	-	XXX	X
LS2	XX	XXX	-	-	-	X	XXX	-	-	XX	XXX	-
LS3	XX	XXX	-	-	-	-	XXX	-	-	-	XXX	X
LS4	XXX	XX	-	-	XX	XXX	-	-	-	-	XXX	-
LS5	XX	XXX	-	-	-	X	XXX	-	-	XX	XXX	-
LS6	-	-	-	XXX	XXX	-	-	-	XXX	XX	-	-
LS7	-	XX	XXX	-	-	XXX	XX	-		XXX	XX	-
LS8	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
LS9	-	XXX	-	-	-	X	XXX	-	-	XX	XXX	-
LS10	XX	XXX	-	-	-	X	XXX	-	-	-	XXX	X
LS11	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
LS12	-	-	XXX	-	XX	XXX	-	-		XXX	-	-
LS13	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
LS14	-	XXX	-	-	-	XX	XXX	-	-	-	XXX	-
LS15	-	XXX	-	-	-	XX	XXX	-	-	-	XXX	-
PC1	-	-	-	XXX	XXX	-	-	-	XXX	X	-	-
PC2	-	-	-	XXX	X	XXX	-	-	X	XXX	X	-
PC3	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
PC4	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
PC5	-	-	XX	XXX	XXX	-	-	-	X	XXX	-	-
PC6	-	-	XX	XXX	XXX	-	-	-	XXX	X	-	-
PC7	-	-	-	XXX	XXX	-	-	-	XXX	X	-	-
PC8	-	-	XX	XXX	XXX	-	-	-	XXX	-	-	-
PC9	-	-	XX	XXX	XXX	-	-	-	XXX	-	-	-
PC10	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
S. americanum	-	-	XXX	-	-	XXX	-	-	-	XXX	XX	-
S. burbankii	-	-	-	XXX	XXX	-	-	-	XXX	-	-	-
S. chenopodioides	-	XXX	-	-	-	XXX	XX	-	-	XXX	XX	-
S. retroflexum	-	-	-	-	-	-	-	-	-	-	-	-
S. scabrum	-	-	-	XXX	-	-	XX	XXX	-	-	XXX	XX

XXX Usually; XX Occasionally; X Rarely; - Never

© University of Pretoria

3.3.6 Genomic ploidy

The possible ploidy of each accession was calculated using flow cytometry as described in section 3.2.4. Values for LS2, LS5, LS6, LS7, LS10, LS12, LS15, PC2, PC5, AME and CHEN were determined in triplicate (three leaves used independently for analysis). The remaining accessions were analyzed once. The results indicated that CHEN, LS1, LS2, LS3, LS4, LS5, LS9, LS10, LS14 and LS15 were diploid (two sets of chromosomes) with 2C values of approximately 2.4 pg. BUR, RET, LS6, LS8, LS11, LS13 and PC1–PC10 were tetraploid (four sets of chromosomes) with 2C values of approximately 4.5 pg, while AME, SCA, LS7 and LS12 were hexaploid (six sets of chromosomes) with 2C values of approximately 6.8 pg. According to the existing literature, *S. chenopodioides* and *S. americanum* are diploid (2n), *S. retroflexum* and *S. burbankii* are tetraploid (4n) and *S. scabrum* and *S. nigrum* are hexaploid (6n) where n is equal to 12 (Edmonds and Chweya, 1997). *Bellis perennis* was used as a control sample, and has a 2C diploid value of 3.4pg. The calculated values are depicted in Table 3.8, and graphically represented in Figure 3.31.

© University of Pretoria

Table 3.8 Averaged flow cytometry values for all nightshade accessions

Accession	2C value (pg)	Replication
Bellis perennis	3.3800	None
LS1	2.4032	None
LS2	2.3143	Triplicate
LS3	2.3930	None
LS4	2.3795	None
LS5	2.3361	Triplicate
LS6	4.5096	Triplicate
LS7	6.7946	Triplicate
LS8	4.5089	None
LS9	2.3457	None
LS10	2.3731	Triplicate
LS11	4.5191	None
LS12	6.7600	Triplicate
LS13	4.5258	None
LS14	2.3660	None
LS15	2.3782	Triplicate
PC1	4.4515	None
PC2	4.4592	Triplicate
PC3	4.4244	None
PC4	4.4346	None
PC5	4.5076	Triplicate
PC6	4.4278	None
PC7	4.4346	None
PC8	4.4481	None
PC9	4.4278	None
PC10	4.4447	None
AME	6.8102	Triplicate
BUR	4.5258	None
CHEN	2.3832	Triplicate
RET	4.4751	None
SCA	6.6789	None

© University of Pretoria

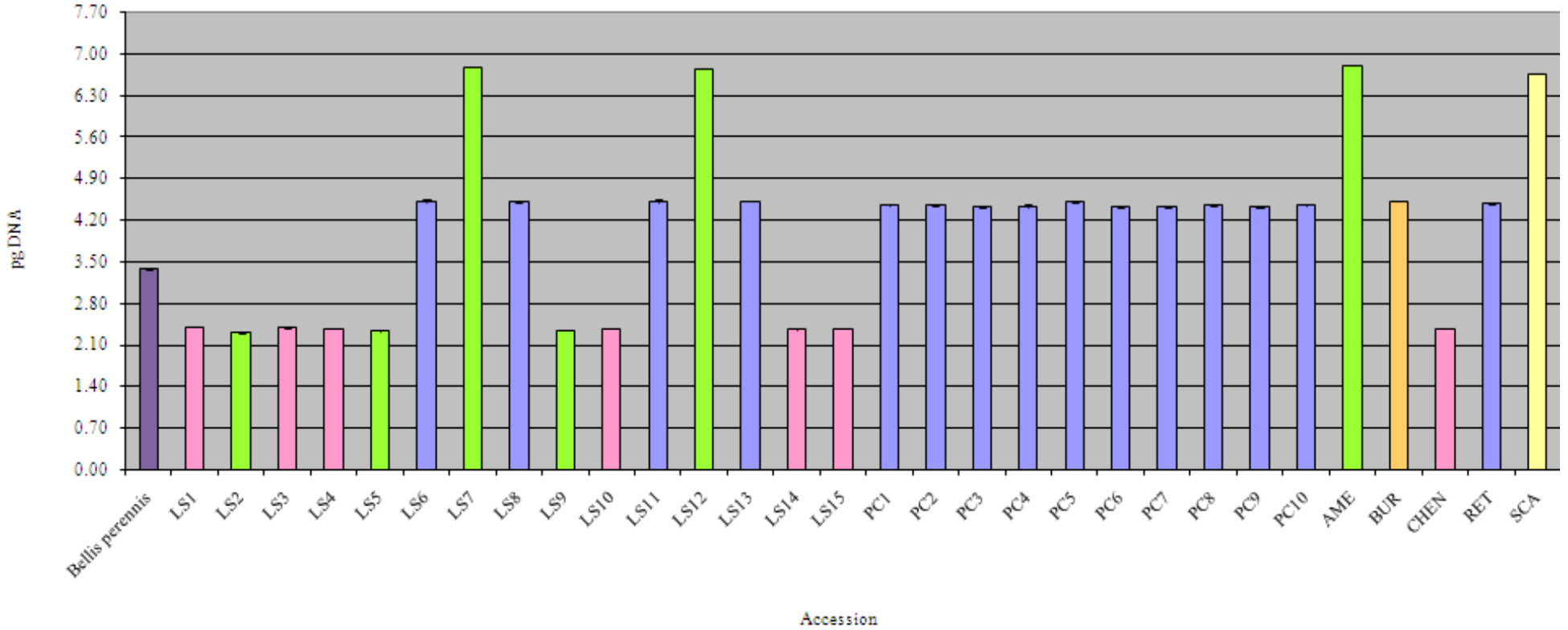


Figure 3.31 Graphical representation of 2C values for each nightshade accession. *Bellis perennis* was used as a control with 2C=3.4pg. Values for LS2, LS5, LS6, LS7, LS10, LS12, LS15, PC2, PC5, AME and CHEN were determined in triplicate. Putative identities are colour coded; green = *S. americanum*, orange = *S. burbankii*, pink = *S. chenopodioides*, blue = *S. retroflexum* and yellow = *S. scabrum*.

© University of Pretoria

3.3.7 Composite analysis

PCA analysis was performed using selected informative morphological datasets indicating distinctive separation of groups. Based on morphological discriminations and p-values, the following characters were identified as most informative: plant height, leaf area, petiole length, leaf margin indentation size, ratio of leaf length to width, ratio of leaf length to petiole length, ratio of corolla length to width, ratio of seed number to berry size, berry size, number of seeds per berry, seed size, ratio of number of seeds to size of seeds, density of trichomes on adaxial surface and density of trichomes on abaxial surface. PC 1 accounts for 39% variance between accessions, while PC 2 accounts for 24% variance. Morphological features of herbarium type specimens were analyzed (Table 2.3) and included in the PCA analysis. Four separate clusters were distinguished, and were putatively identified with significant support from the clustering of "type" herbarium specimens. Cluster I consisted of the *S. chenopodioides* type specimen, CHEN, LS1, LS3, LS4, LS10, LS14 and LS15. Cluster II consisted of LS2, LS5 and LS9, together with a *S. americanum* herbarium specimen. This was not the type specimen, but rather a slight morphological variation found of *S. americanum*. Cluster III comprised of the *S. retroflexum* type specimen, RET, BUR, LS6, LS8, LS11, LS13 and PC1–PC10. Cluster IV contained the remaining accessions, including *S. scabrum* type species, SCA, AME, LS7 and LS12 (Figure 3.32).

© University of Pretoria

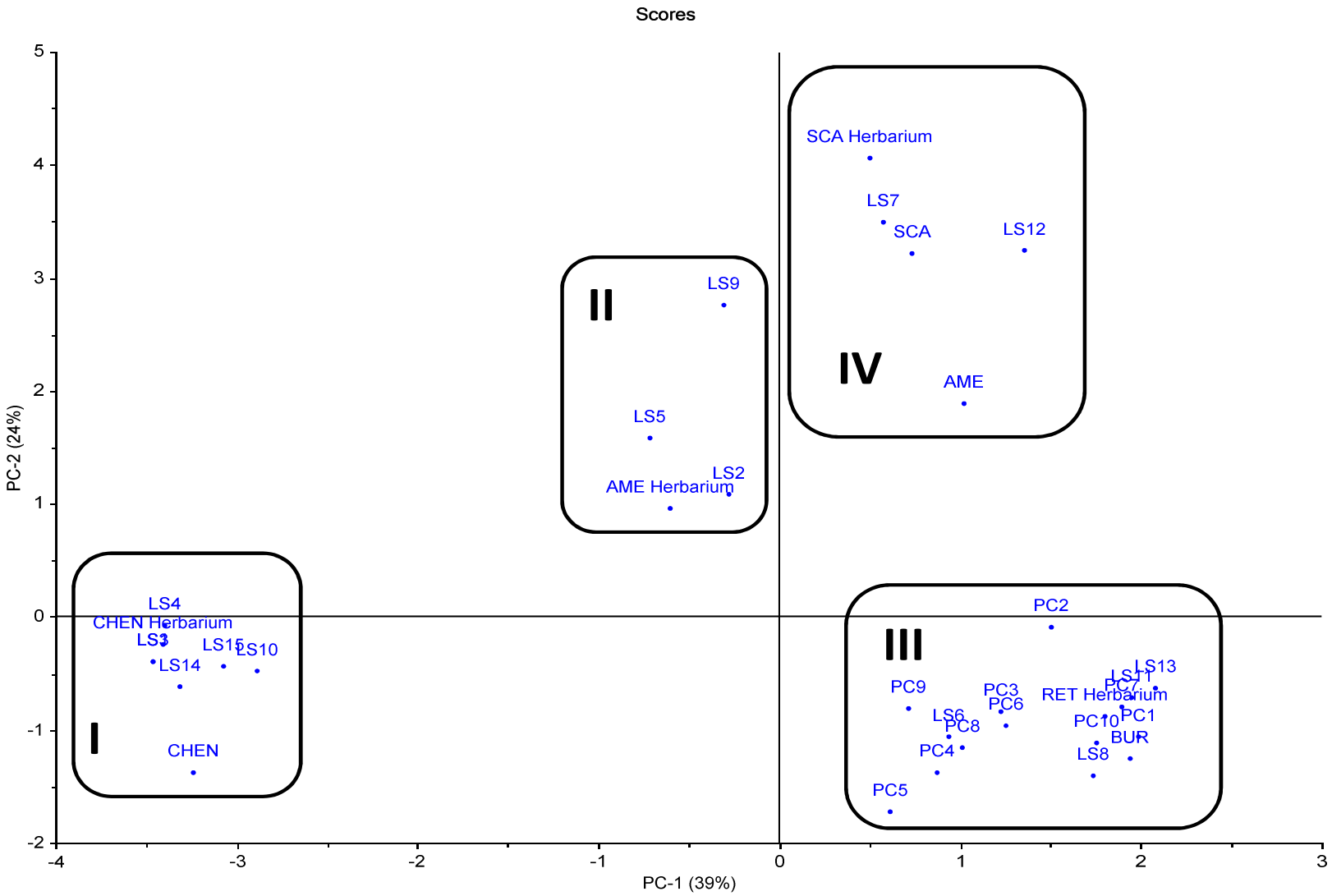


Figure 3.32 PCA analysis of the most informative morphological character traits. Four clusters were identified, and are considered putative species groups based on the high confidence when grouping with authentically identified herbarium specimens.

© University of Pretoria

3.4 Discussion

Delineation of species belonging to the *S. nigrum* complex based on morphological characteristics have been attempted by previous studies, only to encounter constraints and limitations due to large amounts of variation. Ogg *et al.* (1981) have listed a number of characters which they consider to provide the best resolution between species. These were plant growth form, berry morphology, seed morphology and flower morphology. They further mention that morphological characteristics were subjective and that very few obvious indicators for a specific species exist. Edmonds (1972) proposed to extend this list, further including leaf characteristics, stem morphology, and inflorescence morphology. Henderson (1974) concluded during his investigation of the nightshade species of Australia that unfavourable environmental conditions can have a major effect on the morphology of all vegetative features. It was assumed that as the plants used in the present study were all grown under strictly controlled and essentially identical environmental conditions, the morphological differences recorded were due to taxon specific differences. However, it should be kept in mind that the precise position of individual plants in the greenhouse (with more or less sunshine; changes in temperatures) as well as the amount of water received by each plant may have caused small variations in the morphology. The present study uses a combination of these characters in order to delineate the South African nightshade accessions. The taxonomic importance of each vegetative feature will be discussed below. Individual accessions will be referred to according to their initial identifications (Chapter 2), and as indicated by the colour coding in this chapter (Table 3.9).

Table 3.9 Putative identities of collective accessions based on morphology

Accession	Putative identity	Colour coding
LS1, LS3, LS4, LS10, LS14, LS15, CHEN	*S. chenopodioides*	Pink
LS2, LS5, LS9	*S. americanu*m - AG1	Green
LS7, LS12, AME	*S. americanum* - AG2	Green
LS6, LS8, LS11, LS13, PC1-PC10, RET, BUR	*S. retroflexum*	Blue, orange
SCA	*S. scabrum*	Yellow

© University of Pretoria

3.4.1. Whole plant

The outward appearance of each plant differed from spreading to erect, with the erect plants being significantly taller than spreading forms. The growth type was an informative characteristic for initial discriminations, but did not seem to distinguish between all accessions or species groups. Similar conclusions were reached by Ogg *et al.* (1981), where *S. nigrum* and *S. sarrachoides* were found to have a spreading growth form, while *S. americanum* and *S. ptycanthum* have an erect growth form. Poczai and Hyvönen (2010) and Schilling (1981) both emphasize the importance of plant growth form as a diagnostic trait. In the present study, *S. scabrum* and three *S. americanum* accessions illustrated an erect growth form, while the remaining accessions had a spreading growth form. The median plant height was 77.74 cm, with seven accessions above this value. The remaining 23 accessions could not be distinguished from each other, as their height increments did not vary with more than 20–30 mm. This may be due to growth rate relative to the position in the glasshouse. The shorter plants had the spreading growth form in common. Plant height successfully separates the three *S. americanum* and three *S. retroflexum* accessions from the rest. Stem colour varies between green to purplish green, but seemed to be under environmental control as both colouring types were present in replicates of the same accession. Stem pubescence was an informative character, with three main groups being distinguished based on this trait. Slight pubescence characterizes accessions putatively identified as *S. americanum* and a few *S. retroflexum*, while moderate pubescence was recorded for the majority of the *S. retroflexum* accessions. The highly pubescent state of the stem was limited to accessions identified as *S. chenopodioides*. Plant height characteristics as well as pubescence correlate extremely well with putative identifications assigned upon collection.

3.4.2. Leaves

According to the literature, the leaf shape of *S. americanum* may vary between ovate and rhomboid (Ogg *et al.,* 1981), or ovate to lanceolate (Edmonds and Chweya, 1997) with entire margins. *S. retroflexum* have ovate to rhomboid leaves, with deeply lobed margins. *S. chenopodioides* and *S. scabrum* have lanceolate and ovate leaves respectively, but both have entire margins (Edmonds and Chweya, 1997).

Lamina outline for nightshades in the present study could be divided into three main states, ovate, ovate to rhomboid and lanceolate. The *S. chenopodioides* accessions all grouped together with lanceolate leaves, while the *S. retroflexum* accessions had an ovate to rhomboid form. Significantly, LS2, LS5 and LS9 putatively identified as *S. americanum* groups had a rhomboid form (hereafter referred to as *S.*

© University of Pretoria

americanum Group 1 – AG1), while AME, LS7 and LS12 *S. americanum* accessions clearly have an ovate leaf blade (hereafter referred to as *S. americanum* Group 2 – AG2). In addition, *S. scabrum* had ovate leaves. Two types of leaf margin divisions were recorded in the present study, namely entire and serrated. Serrated leaf margins were limited to *S. retroflexum* and AG1. Further investigations indicated the indentation depths to be greater for AG1, effectively separating them from the *S. retroflexum* group. Leaf shapes in association with leaf margin were informative, observable traits which were relatively diagnostic, and were consistent across replicates and within a plant when grown in controlled greenhouse conditions.

Overall leaf size and area were too variable to be a trait of any taxonomic significance. In contrast, the length to width ratio was more informative and could be utilized to separate the accessions into two groups. The average leaf area was 1881 mm^2, with the majority of the samples measuring across this average between 1000 mm^2–2000 mm^2. The samples in this range consisted of *S. retroflexum*, *S. chenopodioides* as well as *S. americanum*, indicating that these species groups were not effectively separated from one another based on leaf area. The ratio of leaf length to leaf width indicates that the *S. chenopodioides* accessions group with AG1 and have a ratio of above 1.5 while the AG2 group with the *S. retroflexum* accessions and have a more uniform leaf size. This observation mirrored the results obtained from studies performed with several South American accessions of nightshades (Edmonds, 1972).

Previous studies have emphasized the importance of the ratio between leaf length and petiole length (Edmonds, 1972; Henderson, 1974). In contrast, the leaf: petiole ratio in the present study divided the accessions into only two groups and was thus not considered a useful trait. The *S. chenopodioides* accessions clearly group separately from the rest in having a very short petiole (and therefore a high ratio), whilst the remaining accessions differ by small increments and could therefore not be effectively separated based on this character. Petiole morphology could possibly aid in the identification process, with *S. americanum* and *S. chenopodioides* being separated from *S. retroflexum* in having a thin, un-winged petiole type, in contrast to thicker, winged petioles in the former.

Leaf pubescence is a relatively informative trait for the demarcation of species groups. *S. chenopodioides* was distinctly identified in having a very dense indumentum. SEM images were utilized to divide the remaining accessions into separate groups (slightly and moderately pubescent). However, the density of trichomes also seemed to be under environmental influence as the *S. americanum, S. scabrum* and *S. retroflexum* groups observed were not grouping according to initial identifications. Trichome types did not appear to be restricted to a specific species, with eglandular and short glandular trichomes occurring on both adaxial and abaxial sides of all the leaves. Sporadic occurrences of long glandular trichomes were

© University of Pretoria

observed. The relative trichome sizes did not play a role in the separation of accessions. The sparse pubescence of *S. americanum*, and dense pubescence of *S. chenopodioides* have been previously described by Henderson (1974).

In terms of leaf morphological characteristics, the accessions mostly adhere to putative identifications (colour coding) assigned at the beginning of the study. The *S. retroflexum* accessions had a serrated, rhomboid to ovate leaf type and were observed for all the putative *S. retroflexum* accessions. The same could be said for the S. *chenopodioides* accessions with their lanceolate, entire margined leaves. The *S. scabrum* reference sample also adhered to the characteristics described in literature, being very large, ovate leaves with an entire margin. AG1 adhere to the characters described in literature, having a slightly rhomboid leaf shape and a serrated leaf margin. However, AG2 showed different results, with ovoid leaf shapes and entire leaf margins. Their morphology does not coincide with *S. americanum*, and are suspected to be either a different or hybrid species. To summarize; leaf margin and leaf shape is relatively plastic when grown under controlled conditions, but care should be taken in the field as these traits appear to be under environmental control, which was in concordance to the observations of Edmonds (1972) and Ogg *et al*. (1981). Leaf shape, margin division and leaf length to width ratio were informative characters to use as a first step towards species identification, but could not be considered conclusive on their own.

3.4.3 Flowers

The most noticeable character in this section was the colour of the flowers and anthers. The species collected here all had white, stellate type corollas. Although occasional purple flowers have been noted for *S. americanum* and *S. scabrum*, this colour was not recorded in the present study (Edmonds and Chweya, 1997). *S. retroflexum* has been described as having a purple stripe along the centre of each corolla lobe and this was observed in several of the accessions. Edmonds (1977) proposed that the purple stripe is not species–specific and may be due to environmental influences. Henderson (1974) further speculates that the stripe is due to the presence of anthocyanins, and may be a dominant-recessive state governed by alleles of a small number of genes. The actual cause of the stripe has not yet been confirmed to date. *S. scabrum* was the only accession which had brown to brownish orange anthers, in comparison to the bright yellow anthers displayed by all the other accessions. This trait may not be able to distinguish between all accessions, but clearly separates *S. scabrum* from the remaining accessions.

© University of Pretoria

The ratio of corolla length to corolla width gives two principle groups. The first group was typified by small ratios (below 2) and had the *S. americanum* accessions (AG1 and AG2), together with selected *S. retroflexum* accessions and *S. scabrum* belonging to this group. The second group with larger ratios (above 2), was comprised of *S. chenopodioides* and several other *S. retroflexum* accessions. The *S. americanum* group was clearly separate from the *S. chenopodioides* group, but it appeared as if *S. retroflexum* was not as plastic in this character and was therefore found in both these two groups. Based on this trait, it could be argued that *S. retroflexum* is sensitive to environmental influences. Overall, AG1 had noticeably smaller flowers, and this group is therefore separated once more from AG2. Another useful trait was the fusion type of the petals. For the *S. chenopodioides* group and AG1, the fusion of the lobes occurred at a deep level to the anthers, with a short corolla lobe and a long corolla tube. AG2, *S. retroflexum* and *S. scabrum* had shallow fusions, resulting in corolla lobes being approximately equal in length to the corolla tube. Similar to this study, style exertion has previously been utilized as an informative characteristic for the nightshades (Henderson 1974, Edmonds, 1977; Ogg *et al.,* 1981). *S. scabrum* and AG1 had no style exertion at any developmental stage, while *S. chenopodioides* had a prominent exertion of more than 2 mm throughout the development of the plants. *S. retroflexum* and AG2 also had projection of their styles, but in a less prominent manner with only approximately 1 mm exerted beyond the anthers. The relative position of style to anthers has a possible influence on the breeding characteristics of the different species (Edmonds, 1977).

Pedicel and peduncle morphology provides informative traits. Noticeably, *S. scabrum* was the only accession with an erect, branched peduncle. *S. chenopodioides* and AG1 had an erect peduncle, with straight pedicels. In contrast, *S. retroflexum* and AG2 had a deflexed peduncle type and reflexed pedicels. According to Henderson (1974), pedicel growth form may be influenced by plant orientation, and there seems to be a genetic difference between species with negative geotropism and plants with positive geotropism. It was furthermore hypothesized that the orientation of the pedicels are also under genetic control (Henderson, 1974). In conclusion, flower characters are relatively stable, and not significantly influenced by environmental factors.

The accessions putatively identified as *S. retroflexum* and *S. chenopodioides* adhere to specific characteristics known in those groups. AG1 shows morphological characteristics associated with *S. americanum*, but AG2 characteristics diverge from original *S. americanum* descriptions. Their flowers are larger (in ratio), slight style protrusion and deflexed peduncles were observed, which are conflicting with documented *S. americanum* characteristics.

© University of Pretoria

3.4.4. Berries

Fruit size and colour was an informative diagnostic trait to investigate in order to demarcate the nightshade species. Average fruit size measured in the present study was 6.67 mm in diameter. *S. chenopodioides* and AG1 had berry sizes below average, while in *S. retroflexum* and AG2 the berry sizes are above the average values. Once again, *S. scabrum* was completely distinct from the rest of the accessions in having large berries (> 10 mm in diameter). To separate the accessions assigned to individual clusters based on berry size, the berry polish added additional distinguishing information. Based on the berry surfaces, three groups with shiny, moderately shiny and dull cuticle types could be identified. *S. chenopodioides*, *S. scabrum* and *S. retroflexum* had dull cuticles, AG2 had moderately glossy cuticles and AG1 had very glossy cuticles. Upon investigation of surface morphology of Solanaceous berries by other authors, it became evident that the morphological difference in berry polish was not due to the presence of a waxy layer or bloom as previously assumed. Waxy layers consists of a thin, flexible membrane made of cutin and solvent-soluble lipids (e.g. esters, fatty acids, alcohol, aldehydes and alkenes) (Vogg *et al*., 2004) and generally appear as films, layers or complex forms like rods and tubules (Ensikat *et al*., 2000). A structural difference in the epidermal cell layer is suspected, but additional replicates as well as investigations with a transmission electron microscope will be required in order to confirm this hypothesis.

The calyx was of diagnostic value, having either an adherent or reflexed form. *S. chenopodioides* and AG2 had adherent calyxes, while *S. retroflexum, S. scabrum* and AG1 had reflexed calyces. The number of pedicels per peduncle is potentially an informative diagnostic trait, but seemed to be under slight environmental and developmental control. Depending on where on the plant stem the peduncle originates, there may be more or less pedicels than was observed over the rest of the plant. Generally, *S. scabrum* and AG2 had more than seven pedicels per peduncle, while AG1 and some *S. retroflexum* accessions had less than four. The rest of the *S. retroflexum* accessions had varying numbers of pedicels, usually between four and seven per peduncle. Visual observation of the berries was useful in the initial identification of the species.

The observations support the putative identities allocated to the accessions during collection, except for group AG2. *S. retroflexum* and *S. chenopodioides* accessions have dull cuticle types as expected, and reflexed and adherent cuticle types respectively. AG1 have the anticipated shiny cuticle type and reflexed cuticles. However, once again AG2 differs from the expected morphological traits in having moderately shiny cuticles and adherent calyxes, indicating their putative identity to possibly be incorrect.

© University of Pretoria

3.4.5 Seeds

Seed colour and shape were not diagnostic morphological traits, as all seeds from all accessions were tan or light brown in colour and had a flattened biconvex shape. However, seed size did divide the accessions into three groups. The smallest seeds belonged to *S. chenopodioides* and AG1, falling below the average of 1.5 mm. AG2 had seed sizes of around the overall average (1.3–1.6 mm), while *S. scabrum* and *S. retroflexum* had the largest seeds of above 1.6 mm. The inverse was seen when observing the number of seeds per berry. *S. scabrum* was completely isolated from the rest of the accessions in having the most seeds per berry, with cases exceeding 95 seed per single berry. *S. chenopodioides* and AG1 have approximately 30–40 seeds per berry; AG2 generally has around 20–30 seeds per berry, while *S. retroflexum* had the fewest seeds per berry (< 20). It is suggested that berry size was inversely correlated to seed size, except for SCA which shows a direct correlation. Larger berries lead to fewer seeds of a larger size per berry, and smaller berries had more seeds of a smaller size per berry. The amount of seed per berry is inversely correlated to the seed size, thus smaller seeds lead to more seeds per berry. These findings were mirrored in the ratio calculated for number of seeds to berry size, where a low ratio (< 2) was observed for *S. retroflexum* and AG2 which had a ratio of between 2–4, but *S. chenopodioides* and AG1 having had the highest ratio of above four. These observations support the putative identities and colour coding assigned during collection.

Surface morphology of the seeds was characterized by a covering of hair-like structures, which has been described as “hairs”, “pseudohairs” and “rod like fibrous structures” by previous studies, and is characteristic of the section *Solanum* (Edmonds, 1983). It is possible that the thin epidermal outer layers has been torn apart (environmentally or during washing procedures) and thus reveal these hair-like structures. It has been hypothesized that these hair-like structures are composed of lignin and probably cellulose (Edmonds, 1983). However, the epidermal surface morphology does not appear to be a diagnostic at species level within the black nightshade complex. Additional seed traits suggested to aid in the delineation of nightshade species include weight, hilum size and shape, seed margin shape, seed dimensions, endosperm thickness and number of seed coat layers (Edmonds, 1983), but these have not been investigated in the present study.

© University of Pretoria

3.4.6. Genomic ploidy

Previous authors have indicated the presence of polyploidy within the black nightshade complex. The plants are composed of diploid, tetraploid and hexaploid accessions (Henderson, 1974, Edmonds and Chweya, 1997). These results were mirrored in the present study, with *S. chenopodioides* identified as diploid (2n=24), *S. retroflexum* as tetraploid (2n=48) and *S. scabrum* as hexaploid (2n=72). Significantly, the initially identified *S. americanum* species were continuously separated into groups AG1 and AG2 based on several morphological traits, the same distinction was found when investigating their cytology. AG1 accessions (LS2, LS5 and LS9) were all diploid, while AG2 accessions (LS7, LS12 and AME) were identified as hexaploid. These findings contradict Jacoby *et al.* (2006) where the accession of AME was originally described. From literature it appears that only *S. nigrum* and *S. scabrum* has hitherto been reported as a hexaploid species within this complex (Edmonds, 1977). As the morphological traits do not correspond to either *S. nigrum* or *S. scabrum* species, it is assumed that either a novel species has been identified or the entity is the product of natural hybridizations.

3.4.7 Composite analysis

Morphological comparisons across all traits for each accession are summarized in Table 3.10. Each accession was compared to the following and a P-value was calculated for that trait (Appendix C). The data did not follow a normal distribution pattern and P-values of below 0.05 indicated that the H_0 hypothesis was to be rejected (H_0 stated that the two accessions being compared was the same). When each trait was analyzed individually, discriminations could be made between species groups. However, the discriminations were not consistent across all traits, indicating the need to consolidate all findings in order to obtain more robust species groups.

PCA analysis was performed and included all character traits discussed above. Each trait separately was inadequate to discriminate completely between the different species groups. When all the traits were combined, the traits could be consolidated into principle components which separated the accessions into four clear clusters. The addition of traits from herbarium type specimens further aided in the effective naming of the different species groups. Cluster I comprised of the *S. chenopodioides* type specimen, as well as all accessions putatively identified as *S. chenopodioides*. Cluster II comprised of LS2, LS5 and LS9, which were putatively identified as *S. americanum* during collection and grouped with a *S. americanum* herbarium specimen. Several variations of the *S. americanum* type have been described since

© University of Pretoria

the typification thereof in 1768 and this may be the reason why the original *S. americanum* type specimen could not be used for identification during the present study. Cluster III consisted of RET, BUR and all accessions which have previously been identified as *S. retroflexum*. The grouping of the herbarium type specimen within this cluster added confidence that this was indeed the correct identification of these accessions. Cluster IV was the most noteworthy; it contains the *S. scabrum* type specimen, SCA reference sample, LS7 and LS12. The general assumptions were that LS7 and LS12 could be identified as *S. scabrum*, but several inconsistencies exist when analyzing the morphological data separately. It was of importance that these accessions be investigated on a molecular basis.

In conclusion, the putative identities observed for *S. retroflexum,* AG1 and *S. chenopodioides* accessions appeared to be correct for the morphological traits observed. AG1 was putatively identified as *S. americanum* based on morphological comparisons to herbarium samples (also Edmonds and Chweya, 1997, Dr. Sandra Knapp, personal communication). The AG2 group differed from the other groups in several characters, and it was hypothesized that this group could be a completely separate species group, or possibly a hybrid. The hybrid theory was further supported by the elevated ploidy level observed, as *S. scabrum* and *S. nigrum* are the only two species known to be hexaploid in nature (Edmonds, 1977). This hypothesis will be further investigated through molecular analysis in the following chapter.

© University of Pretoria

Table 3.10 Summary of taxonomically most diagnostic morphological characters across all accessions

Sample	Plant Height (cm)	Area (mm^2)	Leaf indentation average	Ratio leaf length: width	Ratio length of leaves: petiole length	Corolla width: length ratio	Seed number: berry size	Berry size (mm)	Seed number per berry	Seed size	Density of trichomes Adaxial	Density of trichomes Abaxial
LS1	58 ± 4	1660.9 ± 846.2	0	1.8 ± 0.3	3.7 ± 1.6	0.4 ± 0.1	7.3 ± 1.7	5.1 ± 1	36 ± 4.9	1.1 ± 0.2	36 ± 14.6	95.7 ± 20.6
LS2	62.7 ± 2.5	1881.8 ± 314.2	3.8 ± 0.7	1.8 ± 0.1	2.1 ± 0.5	0.5 ± 0.1	5.9 ± 1.3	5.7 ± 0.7	34 ± 9.2	1.1 ± 1.1	2 ± 0.6	2.7 ± 0.6
LS3	52 ± 6.2	1729 ± 597.9	0	1.9 ± 0.3	4.3 ± 2.9	0.3 ± 0.1	7.5 ± 2.2	4.6 ± 0.4	34 ± 8.5	1 ± 0.2	42.3 ± 15.6	106 ± 28.8
LS4	64.7 ± 5.5	1433.5 ± 549	0	1.9 ± 0.2	5 ± 1.7	0.3 ± 0.1	5.7 ± 1.8	4.7 ± 0.6	25.9 ± 6.7	1 ± 0.1	38.7 ± 20.6	111.7 ± 38.5
LS5	66.3 ± 15	2588.2 ± 1024.1	3.7 ± 1	1.8 ± 0.6	1.7 ± 0.5	0.5 ± 0.2	6.6 ± 2.6	5.8 ± 1	36.8 ± 10.5	1.2 ± 0.2	2.7 ± 0.6	1 ± 0
LS6	53 ± 7	1306.8 ± 364.4	2.3 ± 0.5	1.5 ± 0.2	2.5 ± 0.5	0.4 ± 0.1	2.1 ± 0.9	6.6 ± 0.5	13.6 ± 6.5	1.7 ± 0.4	1.7 ± 0.6	1.7 ± 0.6
LS7	126.3 ± 5.7	4047.5 ± 755.9	0	1.4 ± 0.1	1.7 ± 0.7	0.5 ± 0,1	4.3 ± 0.8	7.2 ± 0.6	30.8 ± 4.5	1.4 ± 0.2	1 ± 0.6	0.7 ± 0.6
LS8	56 ± 6	1181.1 ± 206.4	1.9 ± 0.4	1.3 ± 0.1	2.8 ± 2.4	0.5 ± 0.1	2 ± 0.6	6.6 ± 1.1	12.8 ± 4.2	1.9 ± 0.3	1 ± 0	0.0
LS9	70 ± 2	2001.5 ± 357	3.2 ± 0.7	1.7 ± 0.1	2.8 ± 2.3	0.6 ± 0.1	5.6 ± 1.5	7.2 ± 0.8	40 ± 11.3	1.2 ± 0.3	2.3 ± 0.6	4.7 ± 3.5
LS10	66.3 ± 6.7	1598.4 ± 410.8	0	1.8 ± 0.2	5.3 ± 3.4	0.4 ± 0.1	5.4 ± 1.4	5.2 ± 0.7	27.9 ± 6.4	1.1 ± 0.3	29 ± 18.2	86.7 ± 25.5
LS11	87.3 ± 2.5	2102.7 ± 465.8	2.5 ± 1.2	1.4 ± 0.3	1.8 ± 0.4	0.6 ± 0.1	2 ± 0.8	7.4 ± 0.7	15.1 ± 5.3	1.8 ± 0.3	1 ± 0.6	2.7 ± 2.1
LS12	167 ± 8.5	3505.5 ± 734.2	0	1.4 ± 0.1	2.1 ± 1	0.6 ± 0.1	3.9 ± 1.3	7.2 ± 0.6	27.9 ± 7.7	1.5 ± 0.2	1 ± 0	0.7 ± 0.6
LS13	69 ± 3.6	1212.1 ± 161.1	2 ± 0.4	1.5 ± 0.1	1.9 ± 0.5	0.5 ± 0.1	2 ± 0.5	6.4 ± 0.8	12.6 ± 2.3	1.9 ± 0.2	2.7 ± 3.8	1.7 ± 0.6
LS14	70.3 ± 8.6	1498.5 ± 560	0	2 ± 0.3	4.7 ± 1.9	0.3 ± 0.1	5.4 ± 1.6	5 ± 0.7	26.6 ± 6.4	1.2 ± 0.2	25 ± 8.9	76.3 ± 11.1
LS15	83.3 ± 7.4	1920.2 ± 744	0	1.9 ± 0.3	3.4 ± 1.2	0.3 ± 0.1	5.5 ± 2.5	4.8 ± 0.7	26.3 ± 11.5	1.2 ± 0.2	46.7 ± 7.6	115 ± 11.4
PC1	72.3 ± 8.1	1630.6 ± 899.2	2.2 ± 0.5	1.5 ± 0.1	2 ± 0.3	1.2 ± 0.9	2.4 ± 1	7.4 ± 0.5	17.4 ± 7.5	1.7 ± 0.2	1 ± 0.6	1 ± 0.7
PC2	85 ± 2.6	2094.7 ± 858.1	2.1 ± 0.6	1.3 ± 0.2	1.6 ± 0.4	0.5 ± 0.1	3.3 ± 1.5	7.7 ± 0.4	25.3 ± 10.5	1.9 ± 0.3	1 ± 0.1	3 ± 0.5
PC3	100.7 ± 6	1985.2 ± 1095.1	2.4 ± 0.7	1.5 ± 0.3	2.4 ± 1	0.6 ± 0	2.4 0.6	7.4 ± 0.8	18 ± 4.1	1.7 ± 0.2	1 ± 0.6	0.3 ± 0
PC4	70.7 ± 3.2	1544.7 ± 433.6	2.2 ± 0.7	1.6 ± 0.1	2.3 ± 0.5	1.4 ± 1.2	2 ± 0.4	7.2 ± 0.7	14.7 ± 2.5	1.7 ± 0.2	2.3 ± 2.1	5.7 ± 1.4
PC5	76.7 ± 3.8	941.8 ± 273.2	1.6 ± 0.4	1.5 ± 0.1	2.4 ± 0.4	0.5 ± 0.1	2.8 ± 0.8	6.8 ± 0.8	18.5 ± 4.3	1.5 ± 0.2	3.7 ± 1.5	12 ± 6.2
PC6	61.7 ± 9.5	1667.6 ± 312.7	2.6 ± 0.5	1.3 ± 0.1	2 ± 0.4	0.6 ± 0.1	1.9 ± 0.5	7.6 ± 0.5	14.5 ± 3.3	1.6 ± 0.2	2.7 ± 1.5	8 ± 4.1
PC7	99.3 ± 6.8	1440.7 ± 451.9	2.1 ± 0.3	1.6 ± 0.2	2.5 ± 0.7	0.5 ± 0.1	2.2 ± 0.5	7.1 ± 0.5	15.4 ± 3.3	1.7 ± 0.2	1.7 ± 1.2	3 ± 0.8
PC8	79 ± 9.2	1277.1 ± 286.6	2.4 ± 0.5	1.6 ± 0.1	2.2 ± 0.3	0.5 ± 0.1	2.2 ± 0.4	7.3 ± 0.5	15.8 ± 3.6	1.6 ± 0.1	1 ± 0	1.3 ± 0.3
PC9	100 ± 6.6	1938.3 ±512.8	2.5 ± 0.5	1.4 ± 0.2	2.2 ± 0.7	0.6 ± 0.1	1.7 ± 0.7	6.8 ± 0.6	11 ± 3.7	1.6 ± 0.2	3.3 ± 0.6	5.3 ± 3.2
PC10	75.3 ± 11.9	1516.6 ± 336.9	2.8 ± 1.1	1.6 ± 0.1	2 ± 0.7	0.6 ± 0.1	2.1 ± 0.6	6.4 ± 0.9	13 ± 2.8	1.8 ± 0.3	1 ± 0	2.3 ± 1.2
CHEN	54 ± 11.5	904.5 ± 317.9	0	2.1 ± 0.3	2.6 ± 2.9	0.3 ± 0.1	4 ± 0.8	5.5 ± 0.7	21.67 ± 3	1.1 ± 0.2	31.3 ± 13.6	94 ± 7.2
AME	98.3 ± 3.5	1934 ± 518.7	0	1 ± 0.2	1.6 ± 0.7	0.6 ± 0.2	4.5 ± 1.2	6.6 ± 0.8	29.3 ± 6.9	1.4 ± 0.2	2 ± 0	1.3 ± 0.6
SCA	81 ± 14.9	4743.3 ±1132.5	0	1.2 ± 0.3	1.6 ± 0.5	0.7 ± 0.1	5.8 ± 1.2	12.1 ± 1.9	70.1 ± 16.1	1.8 ± 0.2	1.7 ± 0.6	5 ± 1.4
BUR	45.3 ± 6.5	1276.1 ± 137.7	1.6 ± 0.2	0.8 ± 0.1	1.4 ± 0.2	0.6 ± 0.1	2.5 ± 0.6	8.0 ± 0.6	20 ± 4.5	1.7 ± 0.3	2 ± 0	9.3 ± 1.2

© University of Pretoria

Table 3.10 (Cont.) Summary of taxonomically most diagnostic morphological characters across all accessions

Sample	Plant growth form	Stem indumentum (density)	Stem color	Leaf shape	Margin type	Leaf pubescence (density)		Corolla color	Anther color	Style exertion from anthers
						Adaxial	Abaxial			
LS1	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
LS2	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White	Yellow	None
LS3	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
LS4	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
LS5	Spreading	Slightly	Purplish green	Ovate to rhomboid	Serrate	Moderately	Slightly	White	Yellow	None
LS6	Spreading	Slightly	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White	Yellow	< 2 mm
LS7	Erect	Slightly	Purplish green	Ovate	Entire	Moderately	Slightly	White	Yellow	< 2 mm
LS8	Spreading	Slightly	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White	Yellow	< 2 mm
LS9	Spreading	Slightly	Purplish green	Ovate to rhomboid	Serrate	Moderately	Slightly	White	Yellow	None
LS10	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
LS11	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White	Yellow	< 2 mm
LS12	Erect	Slightly	Purplish green	Ovate	Entire	Moderately	Slightly	White	Yellow	< 2 mm
LS13	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White	Yellow	< 2 mm
LS14	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
LS15	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
PC1	Spreading	Slightly	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White with purple stripe	Yellow	< 2 mm
PC2	Spreading	Slightly	Green	Ovate to rhomboid	Serrate	Moderately	Slightly	White with purple stripe	Yellow	< 2 mm
PC3	Erect	Slightly	Purplish green	Ovate to rhomboid	Serrate	Moderately	Moderately	White with purple stripe	Yellow	< 2 mm
PC4	Spreading	Moderately	Purplish green	Ovate to rhomboid	Serrate	Moderately	Moderately	White with purple stripe	Yellow	< 2 mm
PC5	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Moderately	Moderately	White with purple stripe	Yellow	< 2 mm
PC6	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Moderately	Slightly	White with purple stripe	Yellow	< 2 mm
PC7	Erect	Moderately	Green	Ovate to rhomboid	Serrate	Slightly	Slightly	White with purple stripe	Yellow	< 2 mm
PC8	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Slightly	Moderately	White with purple stripe	Yellow	< 2 mm
PC9	Erect	Slightly	Green	Ovate to rhomboid	Serrate	Moderately	Moderately	White with purple stripe	Yellow	< 2 mm
PC10	Spreading	Moderately	Purplish green	Ovate to rhomboid	Serrate	Slightly	Slightly	White with purple stripe	Yellow	< 2 mm
CHEN	Spreading	Highly	Green	Lanceolate	Entire	Highly	Highly	White	Yellow	> 2 mm
AME	Erect	Slightly	Green	Ovate	Entire	Moderately	Moderately	White	Yellow	< 2 mm
SCA	Erect	Slightly	Green	Ovate	Entire	Moderately	Moderately	White	Brown	None
BUR	Spreading	Moderately	Green	Ovate to rhomboid	Serrate	Slightly	Moderately	White	Yellow	< 2 mm

© University of Pretoria

Table 3.10 (Cont.) Summary of taxonomically most diagnostic morphological characters across all accessions

Sample	Orientation of peduncle	Fruit polish	Arrangement of pedicels	Nr of pedicels	Calyx orientation
LS1	Erect	Dull	Central point	4–6	Adherent
LS2	Erect	Shiny	Central point	< 4	Reflexed
LS3	Erect	Dull	Central point	4–6	Adherent
LS4	Erect	Dull	Central point	4–6	Adherent
LS5	Erect	Shiny	Central point	< 4	Reflexed
LS6	Deflexed	Shiny	Alternating	4–6	Reflexed
LS7	Deflexed	Moderately shiny	Alternating	> 7	Adherent
LS8	Deflexed	Dull	Alternating	< 4	Reflexed
LS9	Erect	Shiny	Central point	4–6	Reflexed
LS10	Erect	Dull	Central point	4–6	Adherent
LS11	Deflexed	Dull	Alternating	< 4	Reflexed
LS12	Deflexed	Moderately shiny	Alternating	> 7	Adherent
LS13	Deflexed	Dull	Alternating	< 4	Reflexed
LS14	Erect	Dull	Central point	4–6	Adherent
LS15	Erect	Dull	Central point	4–6	Adherent
PC1	Deflexed	Dull	Alternating	4–6	Reflexed
PC2	Deflexed	Dull	Alternating	4–6	Reflexed
PC3	Deflexed	Dull	Alternating	4–6	Reflexed
PC4	Deflexed	Dull	Alternating	4–6	Reflexed
PC5	Deflexed	Dull	Alternating	4–6	Reflexed
PC6	Deflexed	Dull	Alternating	4–6	Reflexed
PC7	Deflexed	Dull	Alternating	4–6	Reflexed
PC8	Deflexed	Dull	Alternating	4–6	Reflexed
PC9	Deflexed	Dull	Alternating	4–6	Reflexed
PC10	Deflexed	Dull	Alternating	4–6	Reflexed
CHEN	Erect	Dull	Central point	4–6	Adherent
AME	Deflexed	Moderately shiny	Central point	> 7	Adherent
SCA	Branched	Dull	Alternating	> 7	Reflexed
BUR	Deflexed	Dull	Central point	4–6	Reflexed

© University of Pretoria

Chapter 4

Molecular Analysis

© University of Pretoria

4.1 Introduction

The *S. nigrum* (black nightshade) complex is one of the most widespread and variable species groups within the Solanaceae. These plants are considered an emerging food crop in Africa, but to date no cultivars have been investigated or developed for nutritional enhancement using traditional plant breeding techniques (Poczai and Hyvönen, 2010). Even though it occurs on five of the continents (Africa, North America, South America, Europe, Asia and Australia), the morphological and genetic information relating to this group is lacking (Edmonds and Chweya, 1997). Morphological data has been the primary method to study diversity within this complex, but phenotypic traits have the critical drawback of being subjected to influences of the environment, and are not stringent enough to distinguish between plants growing in different regions or different growing seasons (Jacoby *et al.,* 2003). The regulation of genes responsible for morphological traits may be influenced by environmental factors, but the genetic makeup responsible for the morphological trait stays virtually the same regardless of where the plant is grown. Physiological data is considered ambiguous, and molecular traits evolve in a more regular pattern compared to morphological traits. Therefore, many researchers have turned to studying the more stable aspect of genetic variation between different plants. Polymorphic DNA markers are considered a practical approach compared to morphology and have had widespread applications in the delineation of different species groups such as tomato (Haanstra *et al.,* 1999), potatoes (Kardolus *et al.,* 1998), eggplant (Mace *et al.,* 1999), maize (Marsan *et al.,* 1998) and barley (Becker *et al.,* 1995).

The Barcode of Life initiative aims to discriminate between plant species using one or a combination of short selected gene regions, but to date no one gene region fulfilling this criterion has been described (Lahaye *et al*., 2008; CBOL Plant Working Group, 2010). Several authors have claimed success in using various nuclear and chloroplast gene regions to determine relatedness between members of numerous plant species in different flowering plant families. Examples are the intergenic spacers or genes in the chloroplast genome (*atpF–atpH* spacer, *rbcL* gene, *matK* gene, *rpoB* gene, *rpoC1* gene, *psbK–psbI* and *trnH–psbA* spacers) as well as nuclear intergenic regions (ITS1 and ITS2) (Kress *et al.,* 2005, CBOL 2010, Chen *et al.,* 2010). Scepticism remains about whether one of these regions or a combination of them, will be sufficient for identifications of plants to species level (Hebert and Gregory, 2005). AFLP's has emerged as a popular method for distinguishing between less explored plant species, for which little to no sequence information is available. Researchers are able to fingerprint most of the genome of a plant, without any prior knowledge on the sequence of the genes (Vos *et al.,* 1995).

In the present study, genetic characterization was carried out in two parts. Firstly, four genes were chosen for phylogenetic analysis based on previous barcoding recommendations (Kress *et al.,* 2005, Chen *et al*.,

© University of Pretoria

2010). These were the 18S intergenic spacer region (ITS), maturaseK (*matK*), a component of a putative chloroplast respiratory pathway (*ndhF*) and *rbcL,* which encodes for the large subunit of the ribulose–1,5–bisphosphase carboxylase gene involved in photosynthesis. The second part of this chapter entailed identification of polymorphic markers utilizing a whole genome array approach, DArT. These genetic markers are used in genotyping many different species, using a low-cost, high-throughput reliable technology (Jaccoud *et al.,* 2001). The use of DArT in the present study eliminates the need for any sequence information, primer design, or complex gel based analysis methods. It was aimed to group 30 provisionally identified accessions into specific species groups, based on their relatedness to previously identified reference samples obtained from Dr A. van Biljon (née Jacoby), University of the Free State. The methods of single gene analysis and a whole genome approach were compared, and phylogenetic inferences made about possible identities of unknown samples.

© University of Pretoria

4.2 Materials and methods

4.2.1 Single gene analysis

DNA isolation. DNA was isolated for each of the 30 plant accessions using the DNeasy Plant Mini DNA Isolation kit (Qiagen, Valencia CA, USA) as described by the manufacturers, with minor modifications. In short, fresh leaf tissue (± 200 mg) was homogenized using a mortar and pestle. DNA was eluted by the addition of 50 µl buffer AE (10 mM Tris-Cl, 0.5 mM EDTA pH 9) to concentrate the DNA. DNA quality was assessed by visualization on a 1% (w/v) agarose gel stained with 0.5 µg/ml ethidium bromide and concentrations determined using a nanodrop (Nanodrop® ND 1000 Spectrophotometer, Nanodrop Technologies, USA). Domesticated tomato (*Solanum lycopersicum* – 'M82') was included as an out-group during single gene analysis.

PCR amplification. Primers for ITS gene regions were designed using CLC Bio Version 5.5, and *mat*K, *ndh*F and *rbc*L primer sequences were obtained from reference material. Polymerase Chain Reactions (PCR's) was performed for each plant using four sets of primers for *ITS*, *mat*K, *ndh*F and *rbc*L respectively, the properties and references of which are listed in Table 4.1.

ITS, *mat*K and *ndh*F: 5 mM of forward and reverse primers (Inqaba Biotechnical Industries Pty. Ltd, Pretoria, South Africa); 1x buffer (10x NH_4), 1 mM $MgCl_2$, 2 mM dNTPs (Bioline, MA, USA) and 2.5 U Taq (KapaBiosystems, Cape Town, South Africa) in a 25 µl final volume. The following PCR conditions were followed on an Applied GeneAmp Thermal Cycler (Applied Biosystems), 95˚C for 1 min; 30 cycles of 94˚C for 30 sec; 55˚C for 40 sec; 72˚C for 1 min and a final extension of 72˚C for 7 min.

*rbc*L: PCR components were used in final concentrations of 5 mM of both forward and reverse primers (Inqaba Biotec); 1x buffer, 0.5 mM $MgCl_2$, 2 mM dNTPs (Bioline) and 1 U Taq (Kapa) in a 25 µl final volume. PCR conditions were as follows: 95˚C for 4 min; five cycles of 94˚C for 30 sec; 55˚C for 1 min; 72˚C for 1 min; 20 cycles of 94˚C for 30 sec; 54˚C for 30 sec; 72˚C for 1 min and a final extension of 72˚C for 10 min.

© University of Pretoria

Table 4.1 Primer properties

Gene	Primer sequence	Annealing temp (°C)	Amplicon size (bp)	Reference
ITS	ITS_F 5'-GCGGAAGGATCATTGTCG-3' ITS_R 5'-GCGACCCCAGGTCAGGC-3'	55	557	GENBANK^ EF108406, AY863063, AF244720
*mat*K	MatK_F 5'-CACTATGTATCATTTGATAACCCTC-3' MatK_R 5'-TATTACAATCAACATTTCAGAATAG-3'	55	1601	Aoki and Ito (2000)
*ndh*F	NdhF_F 5'-TTGTAACTAATCGTGTAGGGGA-3' NdhF_R 5'-CGATTATATGACCAATCATATA-3'	55	1397	Olmstead and Sweere (1994)
*rbc*L	rbcL_F 5'-ATGTCACCACAAACAGAGACTAAAGC-3' rbcL_R 5'-GTAAAATCAAGTCCACCRCG-3'	54	527	PM. Hollingsworth, personal communication
M13	M13_F 5'-GTTTTCCCAGTCACGACGTTG-3' M13_R 5'-TGACCGGATAACAATTCACACAG-3'	55	*	www.promega.com

^ www.ncbi.nlm.nih.gov/genbank/

Cloning and sequencing. Cloning was performed using the InsTA Clone™ T/A Cloning Kit (Fermentas Life Sciences, Lithuania) according to manufacturer's instructions. PCR product was added to 55 ng/µl vector pTZ57RT and the reaction was placed at 4˚C overnight. Competent JM109 *E. coli* cells were prepared (Appendix B) and 2.5 µl ligation mixture was added to 100 µl cells. The mixture was incubated on ice for 30 min, after which it was subjected to heat shock at 37˚C for 5 min. After 400 µl LB medium (Appendix B) was added, cell growth took place at 37˚C for 1 hour and cells were plated onto LB-agar supplemented with 100 µg/ml ampicillin, 80 µg/ml X-gal and 0.5 mM IPTG (Appendix B). Blue-white colony selection allowed the identification of recombinant plasmids containing the insert of interest. Colony PCR was performed by transferring the white colonies with a sterile toothpick and suspending it in the PCR mixture. PCR was performed using 5 mM of M13-F and M13-R primers (Inqaba), (Table 4.1). The amplicons were visualized on a 1.5% agarose gel. Colonies containing recombinant plasmid (of correct insertion size) were subsequently cultured overnight in LB broth and plasmid was isolated using

© University of Pretoria

the Invisorb Spin Plasmid Mini Kit Two (Invitek, Glasgow, Scotland) as stated by the manufacturers. Plasmid (10 µl) was sent to Inqaba Biotechnical Industries (Pty) Ltd. and sequencing was performed in both directions with M13-F and M13-R on the ABI 3130 XL sequencer (CD Genomics, New York, USA). Amplicons were also sequenced without prior T/A cloning.

Sequence analysis. The ITS, *matK, ndhF* and *rbcL* sequences obtained for each accession was compared to those obtained from reference samples. Basic sequence alignments were performed using CLC Bio, and there were 557 (ITS), 1601 (*matK*), 1397 (*ndhF*) and 627 (*rbcL*) base pair positions in the final dataset after vector and primer sequences were removed. Sequences were analysed using PAUP 1.0b10* (Phylogenetic Analysis Using Parsimony software, D.L. Swofford, Florida State University, USA). Gaps were treated as "missing" data and phylograms were obtained using heuristic searches based on maximum parsimony. Confidence intervals using 1000 bootstrap replicates were calculated. Sequence data from *S. lycopersicum* was used to root the tree with respect to the remaining accessions. For the ITS analysis, a Bayesian framework was used to determine possible evolutionary substitution rates of nucleotides. Gamma rate heterogeneity was set for ITS, and codon specific sites were specified with a site specific substitution rate. A partition homogeneity test was used to determine if the sequence datasets could be combined for the four different genome regions investigated (based on 10 representative accessions). For maximum likelihood analysis, PhyML Version 3.0 was used.

4.2.2 Diversity Arrays Technology

The DArT experiments were performed at DArT (Pty) Ltd in Yarralumla, Australia. A 4428 feature polymorphic array had been previously constructed by the University of Pretoria as part of work-package 2.2 of the EU-SOL Tomato project (www.eu-sol.net), investigating nutritional qualities of wild and domesticated tomato. The polymorphic array contains fragments generated through digestion using *Pst*I and *Taq*I followed by selective PCR using the genomes of domesticated tomato (*S. lycopersicum* – nine cultivars) and various wild tomato species (*S. arcanum, S. chmielewskii, S. galapagense, S. habrochaites, S. neorickii, Solanum pennellii* and *S. pimpinellifolium*). In addition to these, genomic representations of 12 unknown wild South African *S. nigrum* complex species are also present on the array. In this study, the array was expanded to include 384 additional clone representatives of each unknown collected plant group (Table 4.2), leading to a 1920 clone expansion of the current library.

© University of Pretoria

Table 4.2 Nightshade extended library

Library	DNA	Origin	384-well	Clones
S. retroflexum	LS6	Farm A	1	384
	LS8	Farm A		
	LS11	Farm A		
	LS13	Farm A		
	RET	UFS[1]		
	PC1	Farm B	1	384
	PC2	Farm B		
	PC3	Farm B		
	PC4	Farm B		
	PC5	Farm B		
	PC6	Farm B		
	PC7	Farm B		
	PC8	Farm B		
	PC9	Farm B		
	PC10	Farm B		
S. americanum	LS2	Farm A	1	384
	LS5	Farm A		
	LS7	Farm A		
	LS9	Farm A		
	LS12	Farm A		
	AME	UFS[1]		
S. chenopodioides	LS1	Farm A	1	384
	LS3	Farm A		
	LS4	Farm A		
	LS10	Farm A		
	LS14	Farm A		
	LS15	Farm A		
	CHEN	UFS[1]		
S. scabrum	SCA	UFS[1]	1	384
S. burbankii	BUR	UFS[1]		
Total			5	1920

[1]UFS – University of the Free State

4.2.2.1 Library construction

Complexity reduction. Genomic DNA was isolated as stated in section 5.2.1 and quality and quantity analysis performed. A DNA dilution of 100 ng/µl for each accession was made and 1 µl thereof was digested with 2 U of *Pst*I 6bp restriction enzyme (New England Biolabs (NEB), Beverly, MA, USA) and *Taq*I 4bp restriction enzyme (NEB), in a reaction containing 1x RE buffer (100 mM Tris-OAc, 500 mM KOAc, 100 mM $Mg(OAc)_2$ and 50 mM dithiothreitol (DTT) at a pH of 7.8), 1x BSA (NEB), 1 mM ATP, 0.05 µM *Pst*I+0 adaptor (5' CACGATGGATCCAGTGCA3' and 5'CTGGATCCATCGTGCAG3'), 6 U T4 DNA ligase (NEB) and ddH_2O to a final volume of 10 µl. Samples were incubated at 37˚C for 90 min

© University of Pretoria

and at 60˚C for 90 min. Heat inactivation was performed at 80˚C for 20 min and samples were stored at -20˚C until use.

Amplification from Pst*I*+0 *adaptor.* One µl of *Pst*I/*Taq*I digestion/ligation reaction used as template in a subsequent 49 µl PCR mix. The PCR mix consisted of 1x PCR buffer (100 mM Tris-Cl, pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.1% gelatine (Sigma, NSW, Australia)), 0.5 mM dNTPs (Bioline), 20 µM PstI+0 primer (5' GATGGATCCAGTGCAG 3', Sigma), 2 U Taq polymerase (KAPPA) and ddH_2O. PCR conditions were 94˚C for 1 min; 35 cycles of 94˚C for 20 sec; 58˚C for 40 sec; 72˚C for 1 min; a final extension of 7 min at 72˚C and hold at 10˚C, in a PTC200-Thermal Cycler (GMI Inc. Adelaide, Australia). Analysis of PCR product on a 1.2% agarose gel confirmed a homogeneous smear of fragments between 500 and 1000 bp.

Cloning. The amplified restriction fragments of individual genomes were pooled in a 1:1 ratio and subsequently 4 µl was ligated into 10 ng of pCR2.1 TOPO Vector (Invitrogen, CA, USA). The salt concentration was adjusted to 300 mM NaCl and 15 mM $MgCl_2$ and the reaction incubated for 15 min at room temperature. The ligated mixture (3 µl) was added to 50 µl ice-thawed electro-competent *E. coli* cells. The transformation was performed by transferring 25 µl bacteria/ligation solution to an electrocuvette (Life Technologies, Gaithersburg, USA) and subjecting it to 2.48kV voltage in a CELL-Porator® *E. coli* Pulser (Life Technologies). The cells were transferred to SOC medium (Invitrogen), and incubated at 37˚C for 40 min. Approximately 120 µl of the transformed *E. coli* was plated onto an agar plate containing 100 µg/ml ampicillin and 40 µg/ml X-gal. The bacterial plates were incubated overnight at 37˚C. Blue-white selection ensured identification of putative clones containing inserts. White colonies were picked using sterile toothpicks and transferred to a 384 well plate containing freezing medium (Appendix B) with 100 mg/L ampicillin and 100 mg/L kanamycin. The 384-well plate was incubated overnight at 37˚C. In total, 1920 colonies were selected to fill five 384-well plates (Eppindorf AG, Germany) for library extension.

Insert PCR. Insert amplification was performed in a reaction mixture consisting of 1x Possum Taq buffer (in house produced by DArT), 200 µM dNTP's, 0.2 µM M13-F primer, 0.2 µM M13-R primer and 0.35 µl Possum Taq. Small aliquots from the 384 well culture plates were transferred to the 384 well PCR plate using a 384-well replicator (V&P Scientific, CA, USA) and the inserts were amplified using the following program; 95˚C for 4 min; 57˚C for 35 sec; 72˚C for 1 min; 35 cycles of 94˚C for 35 sec; 52˚C for 35 sec; 72˚C for 1 min and a final extension of 72˚C for 7 min. Quality control of the amplicons were performed by visualizing 2 µl of the product on a 1.2% agarose gel. The PCR products were precipitated by drying overnight and washed in 77% EtOH. The EtOH was completely removed from the plate and the

© University of Pretoria

DNA was dried overnight. The pellets were re-suspended in 25 µl DArT spotting buffer (prepared by DArT).

Printing of library. A 20 plate array was used for slide printing, consisting of tomato and nightshade clones (Table 4.3). The Microgrid BioRobotics printer (Isogen Life Science, The Netherlands) deposits approximately ~1 nl of each amplicon as a single spot on a SuperChip poly-L-lysine slides (Erie Scientific, Portsmouth, USA). The array was composed of 7680 features, spotted in duplicate. After printing, the slides were heated to 80°C for 2 hours, incubated in hot water (95°C) for 2 min and dried by centrifugation.

Table 4.3 DArT nightshade library for polymorphism identification

Name	Species used in construction	Number of 384-well plates
RET1	Nightshade - *S. retroflexum* (Farm A)	1
RET2	Nightshade - *S. retroflexum* (Farm B)	1
AME	Nightshade - *S. americanum*	1
CHEN	Nightshade - *S. chenopodioides*	1
SB	Nightshade - *S. scabrum* and *S. burbankii*	1
Nightshade	Wild South African nightshade species	2
M82	Domesticated tomato	3
S. pennellii (LA0716)	Wild tomato	3
S. galapagense (LA0483)	Wild tomato	2
Hawaii7998	Domesticated tomato with wild introgressions	1
West Virginia007	Domesticated tomato with wild introgressions	1
Re-arrayed Solanaceae	Polymorphic tomato clones	3
Total plates		20

© University of Pretoria

4.2.2.2 Target Preparation

Complexity Reduction. Target representations were prepared similarly to representations for library construction. Each individual genome (30 accessions) was digested by *Pst*I and *Taq*I restriction enzymes, *Pst*I specific adapters were ligated and PCR amplified. The targets were precipitated by the addition of 1 volume isopropanol and centrifuged for 40 min at 4000 rpm (30˚C), after which the supernatant was discarded and washed with EtOH. Target DNA was dried overnight and stored at -20˚C until use.

Labelling of targets. Targets were labelled by mixing 1x NEB Buffer 2, 50 µM random decamers (in house produced by DArT) and labelling dNTPs (2 mM dATP, 2 mM dCTP, 2 mM dGTP and 0.2 mM dTTP) to a final volume of 5 µl. The labelling reaction was added to the dried pellet of precipitated target and mixed gently. The samples were denatured at 95˚C for 3 min and the temperature was decreased to 25˚C, after which 5 µl Cy-dye mix (1x NEB buffer 2, 25 µM Cy3-dUTP or Cy5 dUTP (Amersham Biosciences, UK) and 2.5 U Klenow exo–enzyme (NEB) was added. Plates were wrapped with aluminium foil and incubated at 37˚C for 3 hours. Each target combination was hybridized together with an M13-F and M13-R amplified product originating from an empty TOPO plasmid labelled with pFAM dye (Amersham) which served as an internal control.

Hybridization. The 30 accessions were hybridized in the order indicated in Table 4.4. One aliquot of DArT-hybridizer (pFAM labelled pCR2.1 vector polylinker (approximately 20 ng per slide), ExpressHyb (Clontech), 10 mg/µl herring sperm DNA (Promega) and 2 mM EDTA (pH 8.0) was thawed and 60 µl thereof was mixed with the labelled targets. The targets were denatured 95˚C for 3 min, 56˚C for 5 min and incubated at 55˚C until further use. The denatured products were carefully dripped onto the microarray slides, and cover slips (60 x 24 mm; Lomb Scientific (Pty) Ltd., NSW, Australia) placed over each slide. Slides were placed into a hybridization chamber (together with 40 µl ddH_2O in the wells of the chamber). Samples were allowed to hybridize at 62˚C overnight (12-16 hours) in a waterbath.

Washing of slides. Four wash solutions were prepared; Wash 1: 1x SSC, 0.1% SDS; Wash 2: 1x SSC; Wash 3: 0.2x SSC; Wash 4: 0.02x SSC (Roche, Indianapolis, USA). To each of these wash solutions 0.5 mM DTT was added and mixed. The cover slips were removed from the slides and the latter were placed in wash solution 1 for 5 min with gentle agitation. They were subsequently washed in all four wash buffers for 5, 2 and 4 min respectively. Slides are blotted between each wash step. After wash 4, the slides were submerged into double distilled water containing 0.5 mM DTT for 90 sec. The slides were dried by centrifugation at 30˚C for 7 min at 500g, followed by 30 min incubation in a light-protected desiccator connected to a vacuum pump. Slides were scanned individually on a fluorescent microarray LS300

© University of Pretoria

scanner (TECAN Group Ltd., Austria) and images subsequent images for each fluorescent dye were generated using an appropriate filter/laser combination.

Table 4.4 Hybridization design of 30 nightshade isolates

Nr	ID Cy3 (543)		Nr	ID Cy5 (633)
1	LS1	X	30	SCA
2	LS2	X	29	RET
3	LS3	X	28	CHEN
4	LS4	X	27	BUR
5	LS5	X	26	AME
6	LS6	X	25	PC10
7	LS7	X	24	PC9
8	LS8	X	23	PC8
9	LS9	X	22	PC7
10	LS10	X	21	PC6
11	LS11	X	20	PC5
12	LS12	X	19	PC4
13	LS13	X	18	PC3
14	LS14	X	17	PC2
15	LS15	X	16	PC1
16	PC1	X	15	LS15
17	PC2	X	14	LS14
18	PC3	X	13	LS13
19	PC4	X	12	LS12
20	PC5	X	11	LS11
21	PC6	X	10	LS10
22	PC7	X	9	LS9
23	PC8	X	8	LS8
24	PC9	X	7	LS7
25	PC10	X	6	LS6
26	AME	X	5	LS5
27	BUR	X	4	LS4
28	CHEN	X	3	LS3
29	RET	X	2	LS2
30	SCA	X	1	LS1

© University of Pretoria

Data analysis. The images generated for each slide, was entered into DArTdb, which is a laboratory information system facilitating barcoded sample tracking, data storage and data management. Each batch of TIFF image pairs generated are automatically analyzed by DArTSoft (Version 7), which is an in-house developed software package. The software localized spots according to their reference signal, and rejected spots with a weak signal. The relative hybridization intensities were normalized using the formula log (cy3target/cy5reference). The values for log (cy3target/FAMreference) and log (Cy5target/FAMreference indicates the amount of hybridization signal per amount of DNA spotted onto the array and is used for quality control purposes (Amorim *et al.,* 2009). DArTSoft computed several quality parameters including the between-allelic-states variance of relative target hybridization intensity as a percentage of the total variance (known as the P-value), the median value for replicate spots, the fraction of concordant calls for replicate assays and identified polymorphic clones by using a combination of ANOVA (variance partitioning algorithm) and fuzzy K-means clustering. The alleles were classified as being present (1) or absent (0) in the representation, based on selective fragment hybridizations by the different genotypes. This algorithm also calculated a probability for each individual genotype call. A clone was considered polymorphic if the probability of it belonging to one of the two classes (0 or 1) was above 0.95 averaged across all slides of the batch of images, and if it's P-value was above 0.95 in at least 90% of the total number of slides. Clones which either showed conflicting scores between replicates or could not be scored in any of the replicates were assigned as "unknown".

Markers that did not adhere to the aforementioned criteria were removed while the remaining 0/1 scoring table was imported to PAUP analysis software for the construction of phylogenetic relationships (phylograms) between the accessions. Gaps were treated as "missing" data and trees were obtained using heuristic searches based on maximum parsimony. Confidence intervals using 1000 bootstrap replicates were calculated.

© University of Pretoria

4.3 Results

4.3.1 Single gene analysis

Total genomic DNA was isolated from 30 black nightshade accessions; 15 accessions from Farm A, 10 accessions from Farm B and five reference accessions obtained from the University of the Free State. The quality of genomic DNA was visually analyzed on a 1% agarose gel, while the quantity was measured spectrophotometrically using a Nanodrop (Nanodrop® ND 1000 Spectrophotometer, Nanodrop Technologies USA). The concentrations varied between 200 ng/µl–800 ng/µl. The genomic DNA was observed as a single large, bright band with insignificant degradation patterns observed in certain samples (Figure 4.1). The ITS gene region of each accession (all 30) was amplified in duplicate, while *matK*, *ndhF* and *rbcL* were only amplified in ten representative samples respectively (Figure 4.2).

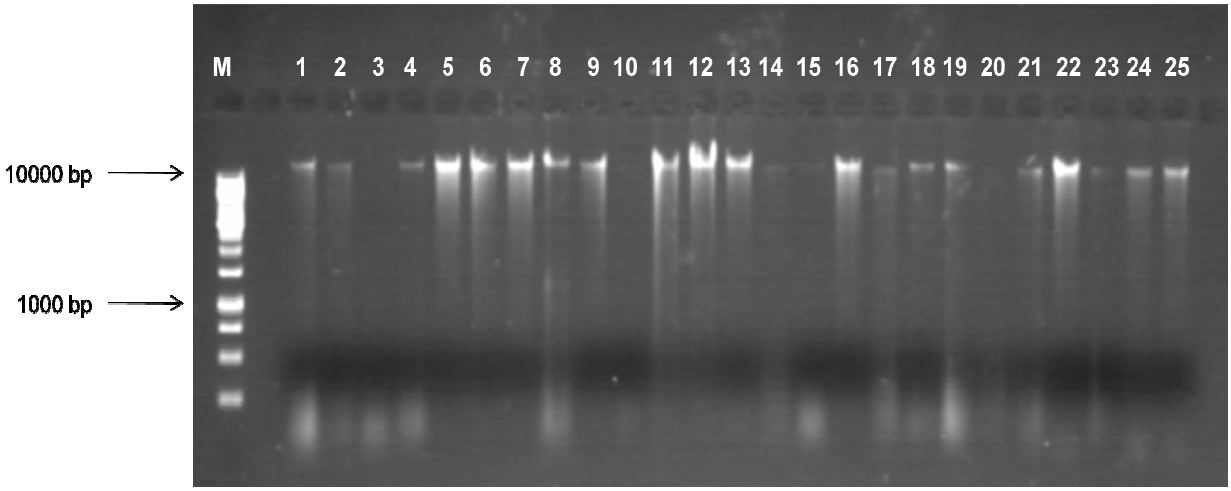


Figure 4.1 Gel electrophoresis following genomic DNA isolations of 25 samples collected from two commercial nightshade farms. M represents a 1 kb O'Generuler DNA ladder (0.1 µg/ml, Fermentas). Lane 1 to 15: LS1-LS15 from Farm A. Lane 16-25: PC1–PC10 from Farm B. DNA quality was deemed sufficient with little to no degradation and RNA contamination. Failed samples were repeated.

© University of Pretoria

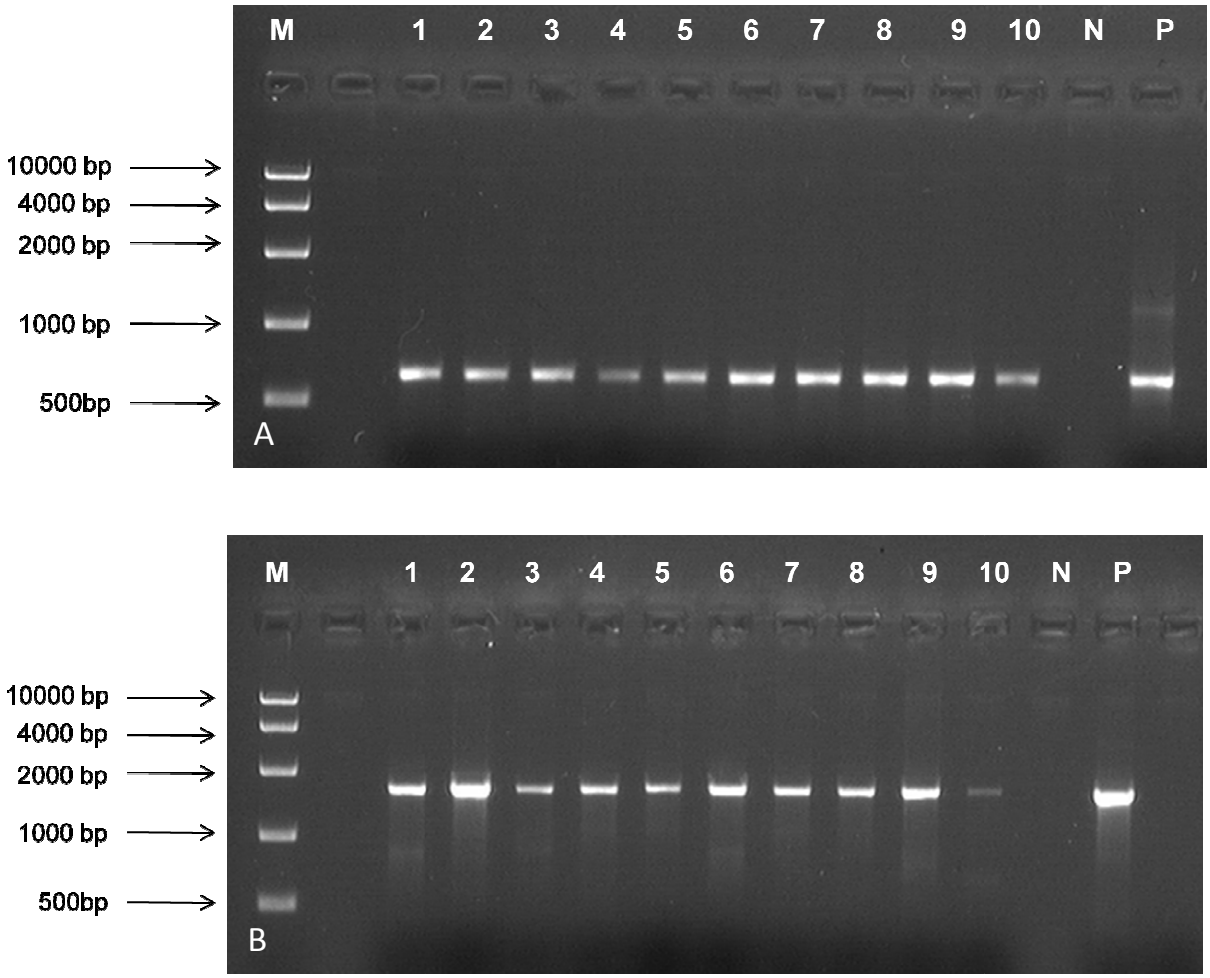


Figure 4.2 (Continued to next page) Gel electrophoresis following PCR amplification of 10 representative accessions from genomic DNA using ITS (A) and *matK* (B) gene specific primers. M represents a Fastruler High Range DNA marker (20 ng/µl, Fermentas). (1) LS1; (2) LS4; (3) LS5; (4) LS7; (5) LS8; (6) LS14; (7) PC3; (8) PC4; (9) PC5, (10) PC9 (N) Negative water control and (P) Positive control of *S. lycopersicum* genomic DNA. ITS amplicon sizes are 557bp, while *matK* shows 1601bp amplicons.

© University of Pretoria

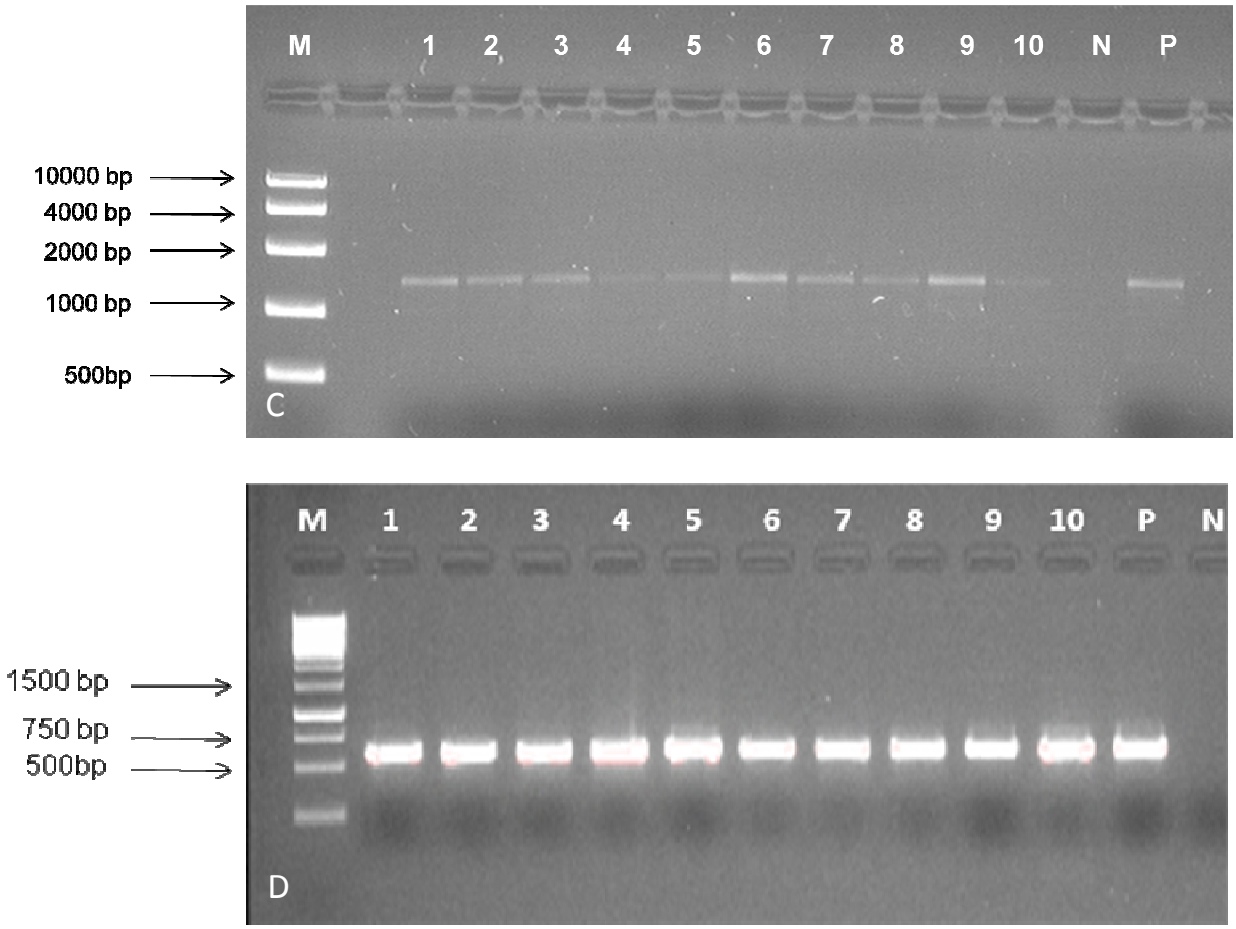


Figure 4.2 (Cont.) Gel electrophoresis following PCR amplification of 10 representative accessions from genomic DNA using *ndhF* (C) and *rbcL* (D) gene specific primers. (C) M represents a Fastruler High Range DNA marker (20 ng/µl, Fermentas). (D) M represents a Generuler 1kb DNA marker (0.1 µg/µl, Fermentas). Both (C) and (D): (1) LS1; (2) LS4; (3) LS5; (4) LS7; (5) LS8; (6) LS14; (7) PC3; (8) PC4; (9) PC5, (10) PC9 and (N) Negative water control (P) Positive control of *S. lycopersicum* genomic DNA. n*dhF* amplicon sizes are 1397bp, while *rbcL* shows 523bp amplicons.

© University of Pretoria

4.3.1.1. Cloned ITS, *matK*, *ndhF* and *rbcL*

Amplification of the genomic ITS gene region resulted in an amplicon size of ~600bp, while amplification of *matK, ndhF* and *rbcL* chloroplast gene regions generated amplicon sizes of approximately 1600bp, 1400bp and 600bp respectively. Each amplicon was cloned and visually selected based on the blue-white screening method. White colonies, which are indicative of an interrupted *lac* operon and therefore the most probable to contain recombinant plasmids, were utilized as templates in subsequent PCR reactions using M13 forward and reverse primers. Recombinant plasmids containing the correct insert sizes, as confirmed by gel electrophoresis, were subjected to sequence analysis. The sequence information of 30 accessions was generated for the ITS regions, while *matK*, *ndhF* and *rbcL* were sequenced in a subset of 10 accessions (LS7, LS5, LS8, LS1, LS4, LS14, PC3, PC4, PC5, PC9) together with the five reference samples (AME, BUR, CHEN, RET, SCA). Sequencing data was analysed using CLC Bio Version 5. This implies the removal of flanking vector and primer sequences followed by the construction of alignments (Appendix D). Sequence statistics are shown in Table 4.5. If one base pair is identical across two or more accessions, but differs from the remaining accessions, it was deemed a polymorphic site separating those species from the remaining accessions. The ITS gene region contained 17 SNP markers, while only three were identified for *matK*, two for *ndhF* and one for *rbcL*. The ITS gene region possess twice the GC to AT content, compared to *matK* and *ndhF* where the opposite was true. *rbcL* showed an equal distribution of each nucleotide across the gene region investigated. Table 4.6 shows the location of single nucleotide polymorphic sites between the different accessions.

Table 4.5 Sequence statistics for four cloned gene regions

			Nucleotide frequencies (%)				
Gene region	Total sequence length	Parsimony informative sites (%)	A	C	G	T	Estimated protein size (kDa)
ITS	557	3.05	17	33	30	17	180.1
*mat*K	1601	0.18	30	17	15	36	512.7
*ndh*F	1397	0.14	28	16	17	37	448.1
*rbc*L	523	0.19	26	21	23	29	168.6

© University of Pretoria

Table 4.6 SNP's identified per cloned gene region

Accession	Position on gene sequence																						
	ITS																	*mat*K			*ndh*F		*Rbc*L
	4	**13**	**22**	**36**	**46**	**48**	**55**	**80**	**106**	**129**	**168**	**377**	**386**	**435**	**522**	**520**	**553**	**293**	**828**	**1538**	**799**	**1016**	**125**
LS 1	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	T	G	T	G	T	T
LS 2	C	A	C	T	C	C	C	C	A	C	C	A	A	T	C	C	A	.	.	.	.	.	.
LS 3	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	.	.	.	.	.	.
LS 4	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	T	G	T	G	T	T
LS 5	C	A	C	T	C	C	C	C	A	C	C	A	A	T	C	C	A	T	G	T	G	T	T
LS 6	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
LS 7	C	A	C	T	C	C	C	C	A	C	C	A	A	T	C	C	A	G	C	C	T	C	C
LS 8	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
LS 9	C	A	C	T	C	C	C	C	A	C	C	A	A	T	C	C	A	.	.	.	.	.	.
LS 10	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	.	.	.	.	.	.
LS 11	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
LS 12	C	A	C	T	C	C	C	C	A	C	C	A	A	T	C	C	A	.	.	.	.	.	.
LS 13	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
LS 14	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	T	G	T	G	T	T
LS 15	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	.	.	.	.	.	.
PC 1	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
PC 2	T	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
PC 3	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
PC 4	T	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
PC 5	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
PC 6	T	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
PC 7	T	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
PC 8	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
PC 9	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
PC 10	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	.	.	.	.	.	.
AME	C	A	C	T	C	C	C	C	A	C	C	A	A	T	C	C	A	G	C	C	T	C	C
BUR	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
CHEN	C	C	C	C	C	T	T	T	G	T	C	C	T	C	C	C	A	T	G	T	T	C	T
RET	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C
SCA	C	C	T	T	T	C	C	C	G	T	T	C	A	C	A	T	T	G	G	C	T	C	C

© University of Pretoria

Combined tree: A combination phylogram based on concatenated ITS, *matK*, *ndhF* and *rbcL* gene sequences was constructed (Figure 4.3). Maximum parsimony analysis was performed, as each gene evolves at a different rate and no evolutionary model could be enforced. Partition homogeneity tests gave a congruence p-value of 0.086, which is above the minimum required value of 0.05 indicating that the sequences could effectively be combined for phylogenetic analysis. The combined sequence length of the four gene regions was 3538 base pairs. The data was analyzed to find the trees with the shortest overall distance, and resulted in six most parsimonious trees, with a tree length of 150, a consistency index (CI) of 0.986; retention index (RI) of 0.975 and 1000 bootstrap confidence replicates were performed. Nucleotide frequencies were 0.177 (A), 0.189 (T/U), 0.333 (C), and 0.301 (G). The transition/transversion rate ratios were $k_1 = 2.398$ (purines) and $k_2 = 0.174$ (pyrimidines).

A grouping is considered robust when bootstraps are at and above 70% (Hillis and Bull, 1993). Three putative species groups, (I) *S. chenopodioides,* (II) *S. americanum* and (III) *S. retroflexum* could be distinguished from the phylogram. The three groups were separated based on sound bootstrap confidence values, and the putative identities were assigned based on the close clustering with previously identified reference samples. Domesticated tomato, *S. lycopersicum*, ITS sequence was included as an outgroup to root the tree.

© University of Pretoria

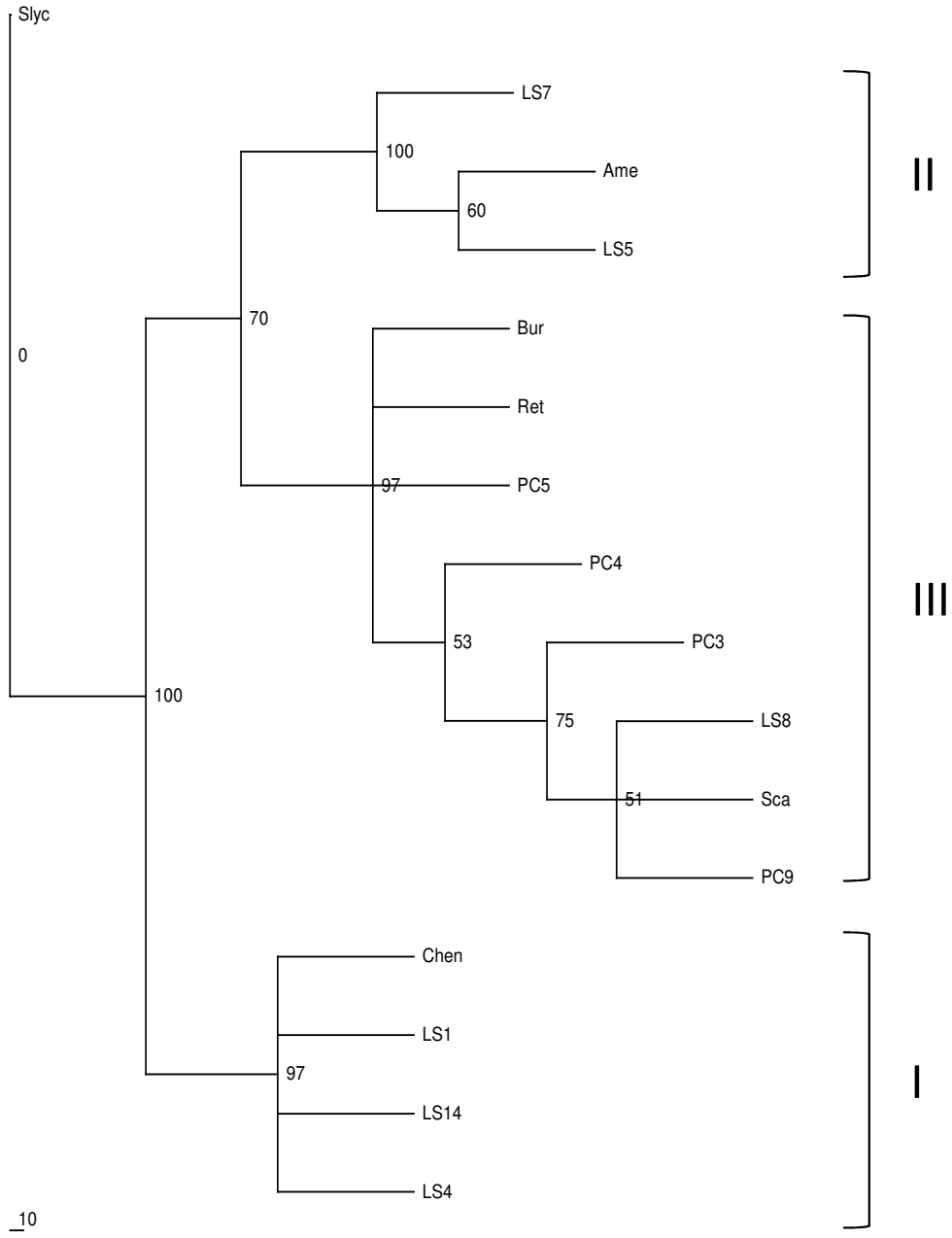


Figure 4.3 Phylogenetic tree of various accessions of *Solanum* constructed using a combination of ITS, *matK*, *ndhF* and *rbcL* sequences. The tree was generated with four gene sequences from 15 genotypes. Reference samples included: (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) *S. retroflexum*; (SCA) *S. scabrum*. Putative close associations of accessions were: (LS7 and LS5) *S. americanum*; (LS8, PC3, PC4, PC5 and PC9) *S. retroflexum*; (LS1, LS4 and LS14) *S. chenopodioides*. *S. lycopersicum* ITS sequence was used as an outgroup.

© University of Pretoria

4.3.1.2 Sequencing analysis of ITS from amplicon

Duplicate amplicons of the ITS gene region generated from the 30 accessions were directly subjected to sequence analysis. These sequences were similar to the sequences of the cloned accessions with the exception of LS7, LS12 and AME. These accessions possessed SNP's at eight different base pair positions on the sequence. The two different haplotypes were not equally distributed based on the size of the primary peak compared to the smaller peak on the chromatogram (Figure 4.4). This indicates the presence of two different haplotypes for these three accessions. These conflicting characters are represented by a degenerate code that corresponds to the standard nomenclature used by the International Union for Pure and Applied Chemistry (IUPAC). The standard IUPAC nucleotide code describes ambiguous sites in a DNA sequence motif, where a single character may represent more than one nucleotide (Appendix E). Plasmids containing inserts of both haplotypes were sequenced and an alignment was constructed in CLC Bio indicating both the haplotypes ("a" and "b") as well as the cloned sequence obtained for LS7, LS12 and AME (Figure 4.5). Seventeen SNP's identified during cloned sequence analyses were common to direct sequencing analysis and two additional SNP's were identified at base pair positions 120 and 427 (Table 4.7).

© University of Pretoria

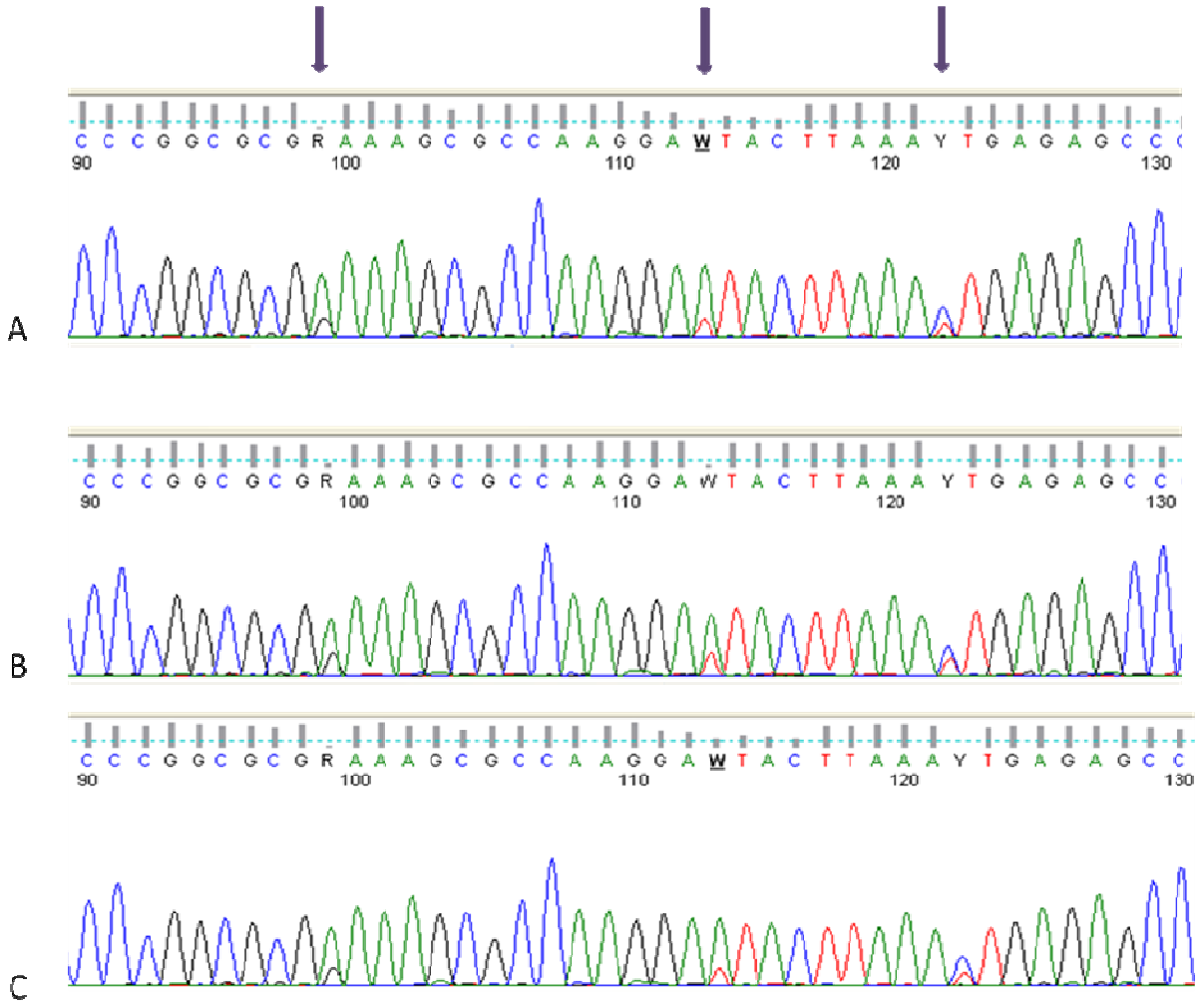


Figure 4.4 Direct sequencing for accessions of *Solanum* (A) LS7, (B) LS12 and (C) AME indicated conflicting base pair compositions at certain sites. These were subsequently scored according to the degenerate IUPAC nucleotide code system. The double peak indicates the presence of two haplotypes for these accessions.

© University of Pretoria

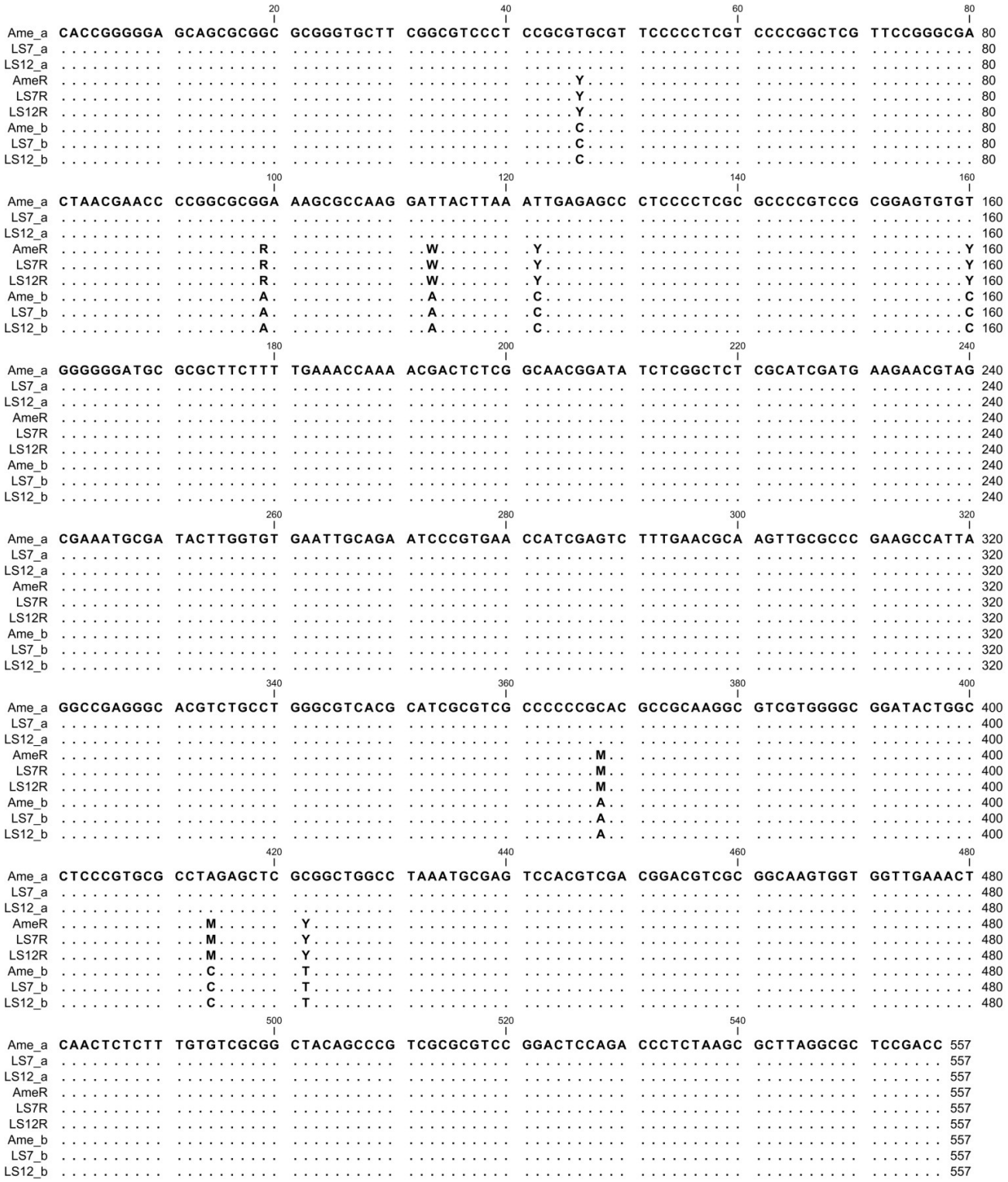


Figure 4.5 DNA sequence alignment for LS7, LS12 and AME haplotypes ("a" and "b") and cloned sequence (R). Eight nucleotide sites were identified as having conflicting base pairs based on the chromatogram.

© University of Pretoria

Table 4.7 SNP's identified from direct sequencing of the ITS gene region

Accession	Position on gene sequence																										
	ITS																					*mat*K			*ndh*F		*Rbc*L
	4	13	22	36	46	48	55	80	106	120	129	168	377	386	427	435	522	535	553	293	828	1538	799	1016	125
LS 1	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	T	G	T	G	T	T
LS 2	C	A	C	T	C	C	C	C	A	A	C	C	A	A	C	T	C	C	A	.	.	.	.	.	.
LS 3	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	.	.	.	.	.	.
LS 4	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	T	G	T	G	T	T
LS 5	C	A	C	T	C	C	C	C	A	A	C	C	A	A	C	T	C	C	A	T	G	T	G	T	T
LS 6	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
LS 7	C	A	C	T	Y	C	C	C	R	W	Y	Y	M	A	M	Y	C	C	A	G	C	C	T	C	C
LS 8	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
LS 9	C	A	C	T	C	C	C	C	A	A	C	C	A	A	C	T	C	C	A	.	.	.	.	.	.
LS 10	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	.	.	.	.	.	.
LS 11	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
LS 12	C	A	C	T	Y	C	C	C	R	W	Y	Y	M	A	M	Y	C	C	A	.	.	.	.	.	.
LS 13	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
LS 14	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	T	G	T	G	T	T
LS 15	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	.	.	.	.	.	.
PC 1	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
PC 2	T	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
PC 3	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
PC 4	T	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
PC 5	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
PC 6	T	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
PC 7	T	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
PC 8	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
PC 9	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
PC 10	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	.	.	.	.	.	.
AME	C	A	C	T	Y	C	C	C	R	W	Y	Y	M	A	M	Y	C	C	A	G	C	C	T	C	C
BUR	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
CHEN	C	C	C	C	C	T	T	T	G	A	T	C	C	T	C	C	C	C	A	T	G	T	T	C	T
RET	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C
SCA	C	C	T	T	T	C	C	C	G	A	T	T	C	A	C	C	A	T	T	G	G	C	T	C	C

© University of Pretoria

ITS sequences for the 30 accessions together with both the respective haplotype sequences of LS7, LS12 and AME were entered into PhyML 3.0 for maximum likelihood analysis. Maximum likelihood enforces an evolutionary model on the dataset and is the preferred method for single gene sequence analysis. The evolutionary model used to infer phylogenetic relationships between 30 accessions based on ITS sequence was found to be GTR+G, where the base frequencies were calculated to be freqA = 0.17438, freqC = 0.34610, freqG = 0.30428, freqT = 0.17524. Gamma shape parameter was determined to be 1.7 and 1000 bootstrap replicates were performed. From the 577 characters used, 511 were constant, 47 variable characters were parsimony un-informative and 19 were parsimony informative. The data was analyzed to find the trees with the shortest overall distance, and resulted in 11 maximum parsimonious trees, with a tree length of 68 (Figure 4.3), a CI of 0.985, a RI of 0.994, RC of 0.979 and HI of 0.015.

rRNA *ITS* gene sequences from the NCBI database Genbank for *S. nigrum* (FJ980391.1 - NIG) and *S. americanum* (GU323361.1 - Same) were included in phylogenetic analysis. Four main clusters could be observed, which are separated by bootstrap values of above 65% (Figure 4.6). The putative species clusters were assigned possible identities, (I) *S. chenopodioides*, (II) accessions LS2, LS5 and LS9 (putatively identified as *S. americanum*) together with "b" haplotype of LS7, LS12 and AME as well as the Genbank *S. americanum* sequence. The *S. nigrum* Genbank sequence was separated from this group with 98% confidence values. Group (III) was identified as *S. retroflexum* and (IV) "a" haplotype of unknown accessions LS7, LS12 and AME. The groups were separated from one another based on sound bootstrap confidence values and the putative identities assigned based on the close clustering with previously identified reference samples. Domesticated tomato, *S. lycopersicum*, ITS sequence was included as an outgroup to root the tree.

© University of Pretoria

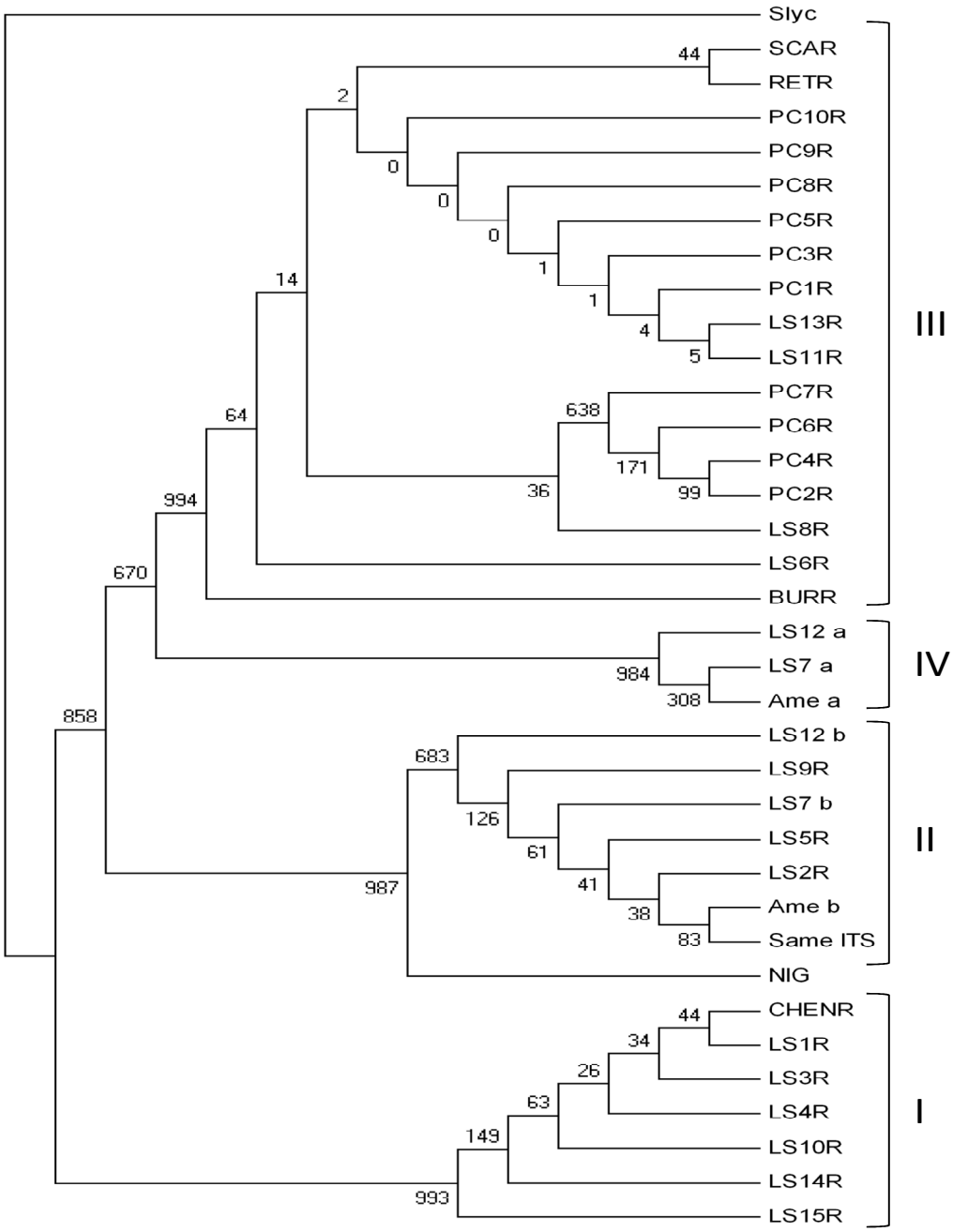


Figure 4.6 Maximum likelihood analysis of 30 rRNA ITS accessions of *Solanum* together with "a" and "b" haplotypes of LS7, LS12 and AME. "R" indicates cloned sequences. Reference samples: (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) *S. retroflexum*; (SCA) *S. scabrum*; (Same ITS) *S. americanum* GENBANK accession; (NIG) *S. nigrum* GENBANK accession. Both haplotypes "a" and "b" for LS7, LS12 and AME were used during tree construction. Putative identities of accessions: (LS7, LA12, AME) Unknown hybrid species; (LS2, LS5, LS9) *S. americanum*; (LS6, LS8, LS11, LS13, PC1–PC10) *S. retroflexum*; (LS1, LS3, LS4, LS10, LS14 and LS15) *S. chenopodioides*. *S. lycopersicum* ITS sequence was used as an outgroup. Bootstrap confidence intervals (based on 1000 bootstraps) above 65% are shown on branch divergence.

© University of Pretoria

4.3.2 Diversity Array Technology

In order to identify markers polymorphic between the 30 black nightshade accessions, individual hybridizations to the diversity array containing fragments from tomato and nightshades were performed in duplicate. The images were scanned and polymorphic markers were identified using a fuzzy logic clustering algorithm (DArTsoft). Duplicate hybridization experiments as well as the duplicate spots on each array were compared between the 30 individuals under investigation.

A total of 1145 markers polymorphic between the accessions were identified (Appendix F). The markers adhered to specific quality control criteria, which included call rate (above 95%), PIC value (average of 0.37), reproducibility (95%) and P-value (average of 85.2 based on DArTSoft quality score). The "0" or "1" status of each individual polymorphic marker was compared between each reference sample (AME, BUR, CHEN, RET and SCA) and each remaining accessions individually. If the status was found to be the same (at that marker position) e.g. CHEN and LS1, the two accessions were considered the same. Comparisons were performed for all 1145 markers. Figure 4.7 indicates the percentage difference between the specific reference sample and the rest of the accessions. The CHEN reference samples were most closely related to LS1, LS3, LS4, LS10, LS14 and LS15, with less than 10% marker differences between them. AME was found to be most similar to LS7 and LS12 (approximately 28% difference), while LS6, LS8, LS11, LS13, PC1–PC10, BUR and RET were also found to be relatively closely related to AME (approximately 35–41% differences). RET, BUR and SCA were closely related to LS6, LS8, LS11, LS13, PC1–PC10, with less than 10% difference in polymorphisms between the unknown accessions and RET/BUR and around 30% difference between the unknown accessions and SCA.

The percentage of markers originating from the respective libraries was calculated (Figure 4.8). The highest amount of markers is present in the libraries constructed from previously investigated wild nightshade species (12.7%), as well as the library constructed from a wild tomato variety, *S. pennellii* (11.7%). Libraries which yielded the least number of markers were constructed from domesticated tomato variants with introgressed segments from wild tomato species, WestVirginia (3.3%) and Hawaii (3.6%). Overall, 57.8% of the total markers were

© University of Pretoria

obtained from libraries of black nightshades, while 42.2% of markers were shown to originate from the tomato libraries. Table 4.8 indicates from which library the maximum number of markers for each accessions originated. The results indicate that the majority of markers for each accession originated from their contributing species library.

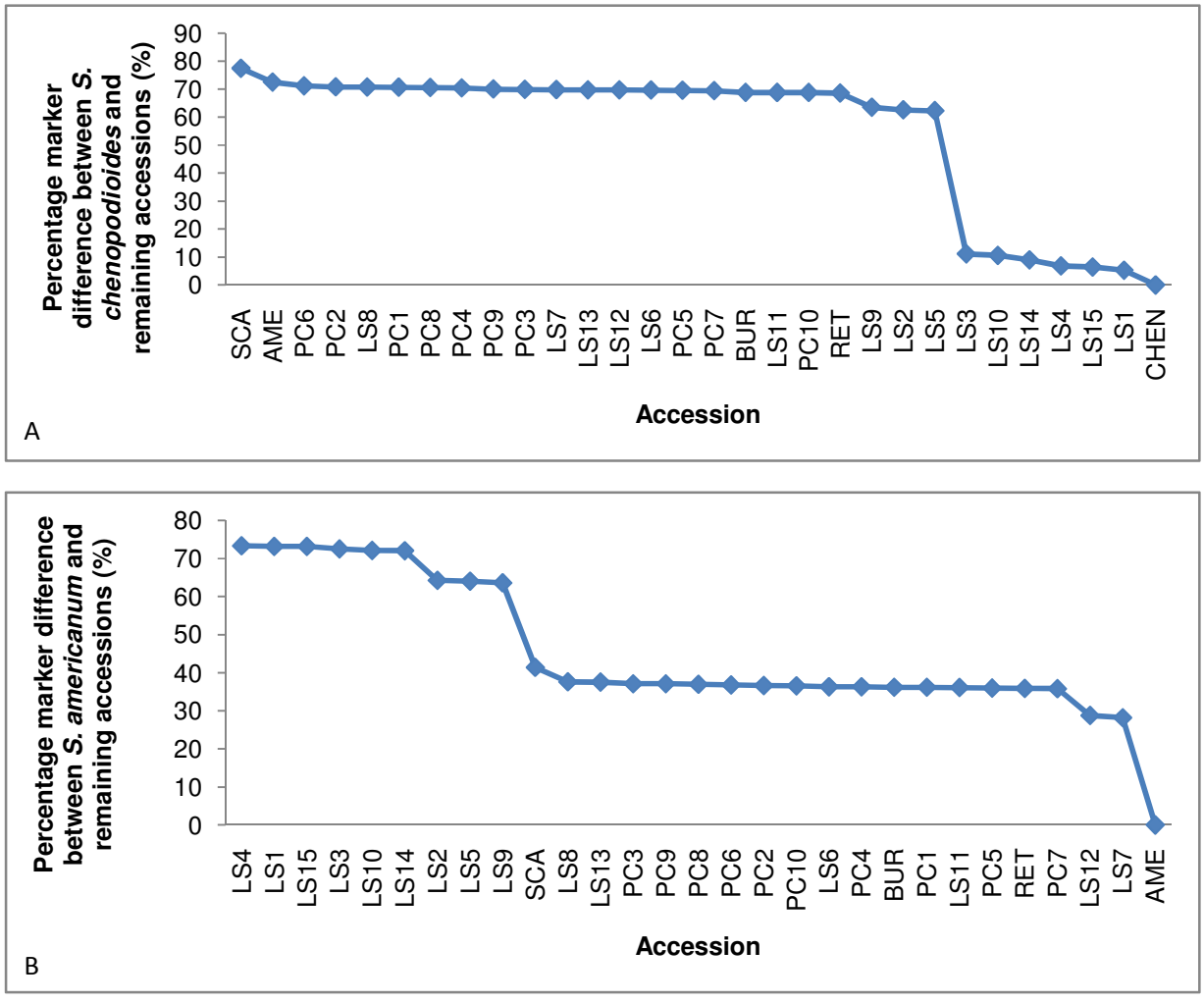


Figure 4.7 (Continued to next page) Marker dissimilarities between reference samples and unknown accessions. Percentage difference markers between (A) *S. chenopodioides* and the remaining accessions and (B) *S. americanum* and the remaining accessions.

© University of Pretoria

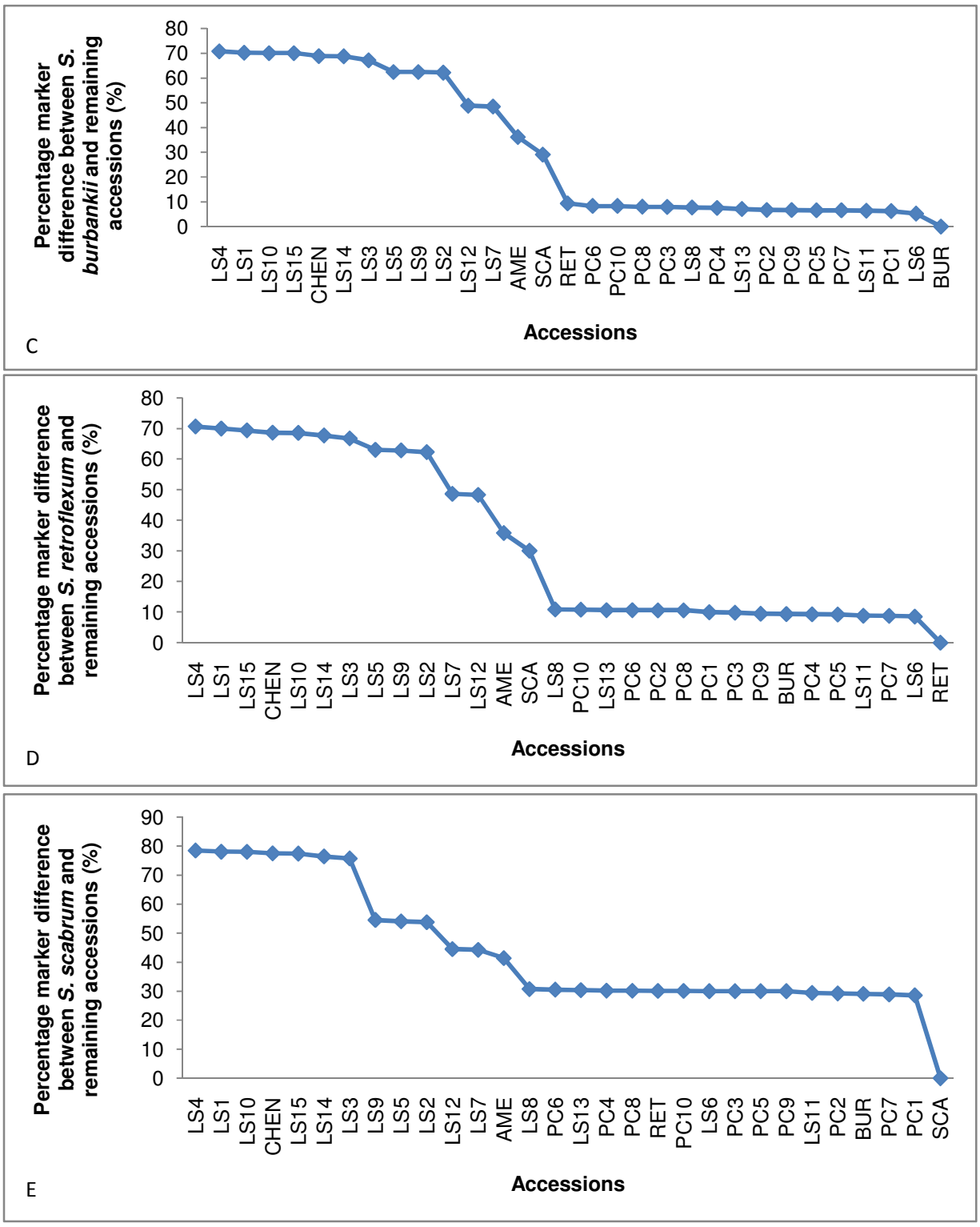


Figure 4.7 (Cont.) Marker dissimilarities between reference samples and unknown accessions. Percentage marker differences between (C) *S. burbankii* and the remaining accessions; (D) *S. retroflexum* and the remaining accessions and (E) *S. scabrum* and the remaining accessions.

© University of Pretoria

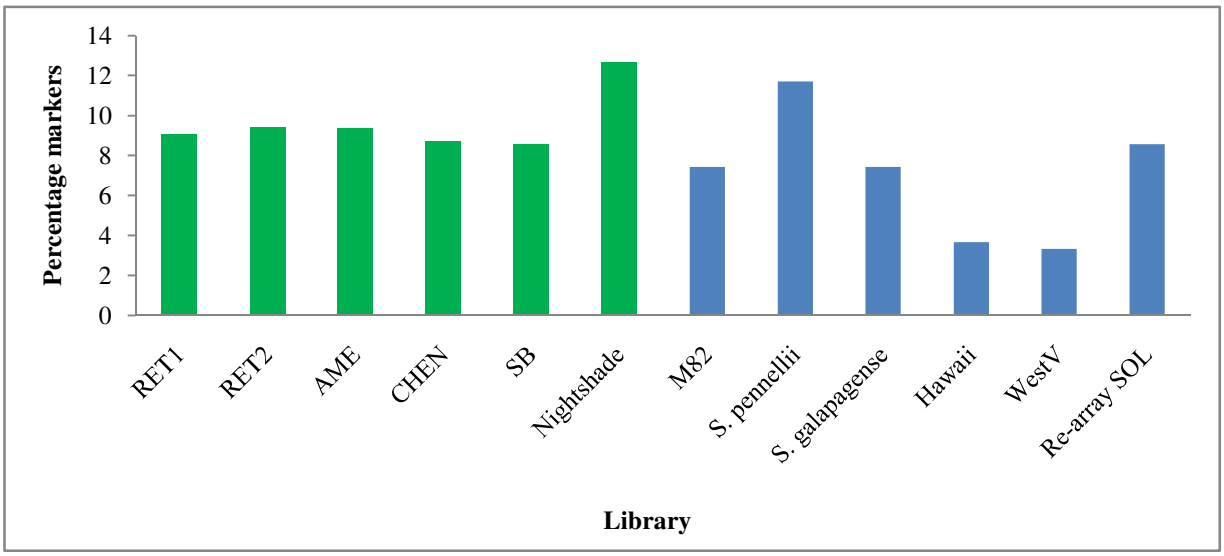


Figure 4.8 Marker distribution from libraries. The highest percentage of markers were obtained from the wild nightshade (12.7%) and *S. pennellii* (11.7%) libraries, while the lowest percentage was obtained from the West Virginia (3.3%) and Hawaii (3.6%) tomato varieties. Green bars represent nightshade libraries while blue represents tomato libraries.

© University of Pretoria

Table 4.8 Library producing maximum markers per accession of members of the *Solanum* spp.

Accession	Putative identity upon collection	Library*
LS1	*S. chenopodioides*	CHEN
LS2	*S. americanum*	AME
LS3	*S. chenopodioides*	CHEN
LS4	*S. chenopodioides*	CHEN
LS5	*S. americanum*	AME
LS6	*S. retroflexum*	RET1 / RET2
LS7	*S. americanum*	RET1 / RET2 + AME
LS8	*S. retroflexum*	RET1 / RET2
LS9	*S. americanum*	AME
LS10	*S. chenopodioides*	CHEN
LS11	*S. retroflexum*	RET1 / RET2
LS12	*S. americanum*	RET1 / RET2 + AME
LS13	*S. retroflexum*	RET1 / RET2
LS14	*S. chenopodioides*	CHEN
LS15	*S. chenopodioides*	CHEN
PC1	*S. retroflexum*	RET1 / RET2
PC2	*S. retroflexum*	RET1 / RET2
PC3	*S. retroflexum*	RET1 / RET2
PC4	*S. retroflexum*	RET1 / RET2
PC5	*S. retroflexum*	RET1 / RET2
PC6	*S. retroflexum*	RET1 / RET2
PC7	*S. retroflexum*	RET1 / RET2
PC8	*S. retroflexum*	RET1 / RET2
PC9	*S. retroflexum*	RET1 / RET2
PC10	*S. retroflexum*	RET1 / RET2
AME	*S. americanum*	RET1 / RET2 + AME
BUR	*S. burbankii*	RET1 / RET2
CHEN	*S. chenopodioides*	CHEN
RET	*S. retroflexum*	RET1 / RET2
SCA	*S. scabrum*	SB

* Refer to table 4.3 for Library construction

The markers were scored as present (1) or absent (0) in individual accessions, and the scoring matrix was imported into PAUP analysis software to construct a maximum parsimony dendogram, indicating the relationships between different accessions (Figure 4.9). Heuristic searches were employed based on parsimony and 1000 bootstrap intervals were performed.

Four main clusters could be identified, and were separated by high bootstrap confidence values. Cluster I contained LS1, LS3, LS4, LS10, LS14, LS15 and the CHEN reference sample, indicating an identity of *S.*

© University of Pretoria

chenopodioides. Cluster II consisted of LS2, LS5 and LS9 which had a putative identity of *S. americanum* (based on morphology). These two clusters can be separated from each other based on a high bootstrap value (100%), adding confidence that these are two definitive species groups. The third cluster (III) consisted of LS6, LS8, LS11, LS13, PC1–PC10 and the RET and BUR reference samples, indicating an identity of *S. retroflexum*. Cluster IV was composed of LS7, LS12 and AME which formed a stepwise out-grouping with high bootstrap values (100%).

To increase the amount of parsimony informative sites, a combination parsimony tree was constructed using DArT polymorphic data (0 and 1) together with ITS sequence data. In total, 2711 characters were used, of which 576 were constant, 44 were parsimony-uninformative and 2091 were parsimony informative. A tree-bisection-reconnection (TBR) branch swapping algorithm was used and 1000 partition-homogeneity test replicates performed. The p-value obtained was 0.84 for the combination tree and 1000 bootstrap replicates were completed. The resulting maximum parsimonious dendogram is represented in Figure 4.10. The combination tree was highly similar to the DArT dendogram, with only minor differences in the bootstrap confidence values.

© University of Pretoria

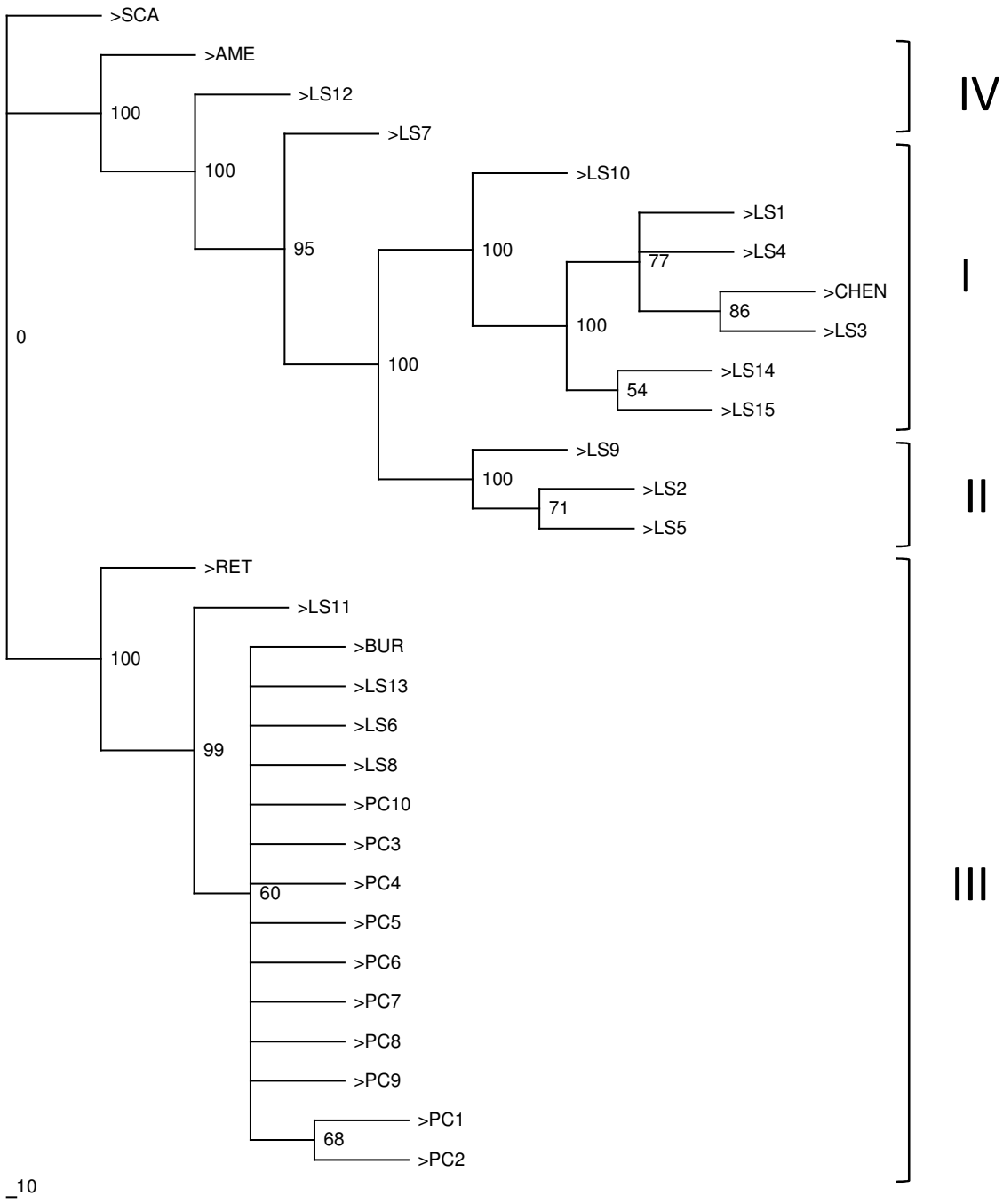


Figure 4.9 Maximum parsimony dendogram indicating relatedness of 30 accessions of *Solanum* based on 1145 DArT marker analysis. Reference samples: (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) *S. retroflexum*; (SCA) *S. scabrum*. Putative identities of accessions: (LS2, LS5, LS7, LS9, LS12) *S. americanum*; (LS6, LS8, LS11, LS13, PC1–PC10) *S. retroflexum*; (LS1, LS3, LS4, LS10, LS14 and LS15) *S. chenopodioides*. The phylogram was obtained using the heuristic search option based on parsimony, and SCA was treated as an outgroup. Bootstrap confidence intervals are shown on branch divergence, and the scale bar indicates number of nucleotide changes per unit of branch length.

© University of Pretoria

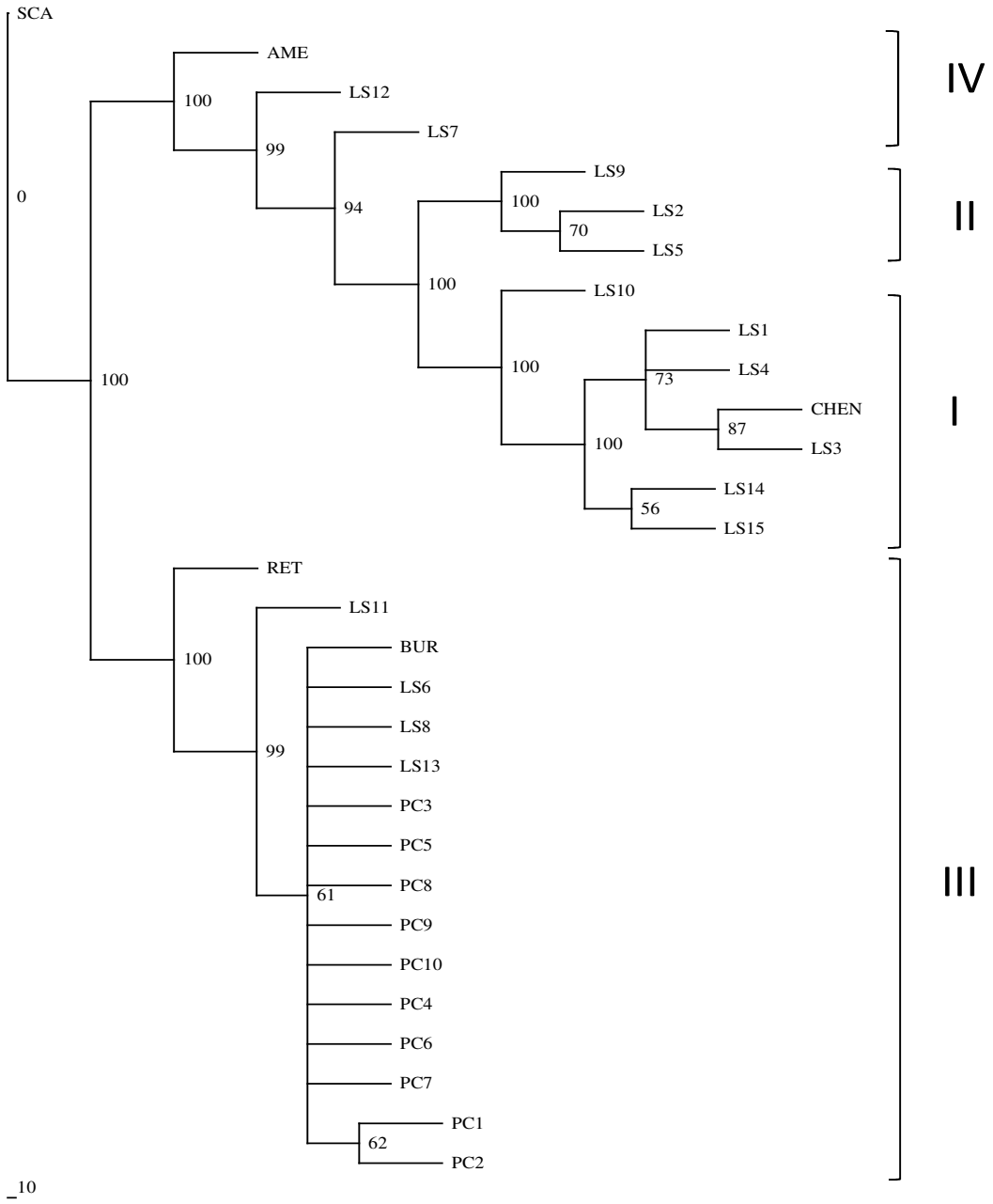


Figure 4.10 Maximum parsimony dendogram indicating relatedness of 30 accessions of *Solanum* based on the combination of DArT marker analysis and ITS sequence data. Reference samples: (AME) *S. americanum*; (BUR) *S. burbankii*; (CHEN) *S. chenopodioides*; (RET) *S. retroflexum*; (SCA) *S. scabrum*. Putative identities of accessions: (LS2, LS5, LS7, LS9, LS12) *S. americanum*; (LS6, LS8, LS11, LS13, PC1–PC10) *S. retroflexum*; (LS1, LS3, LS4, LS10, LS14 and LS15) *S. chenopodioides*. The phylogram was obtained using the heuristic search option based on parsimony, and SCA was treated as an outgroup. Bootstrap confidence intervals are shown on branch divergence, and the scale bar indicates number of nucleotide changes per unit of branch length.

© University of Pretoria

4.4 Discussion

The aim of this chapter was to determine the genetic relationships between members of the *S. nigrum* complex, utilizing a single gene as well as a whole genome approach. The hypothesis stated that the unknown accessions could be identified based on their associations with previously identified reference samples, *S. americanum*, *S. chenopodioides*, *S. burbankii*, *S. retroflexum* and *S. scabrum*.

4.4.1 Single gene analysis

Wolfe *et al.* (1987) performed various statistical analysis on plant nuclear and plastid genomes to determine the relative rate of evolution. His findings indicate mitochondrial DNA to evolve at a rate less than one-third of chloroplast DNA, which in turn evolve at a rate less than half the nuclear DNA. Evolutionary rates were calculated as $0.2–1x10^{-9}$, $3–30x10^{-9}$ and $1–3x10^{-9}$ substitutions per site per year for mitochondrial, nuclear and chloroplast genes respectively. In the present study, nuclear as well as chloroplast gene sequences were investigated and it became apparent that plastid genes were not informative on species level within this complex. Where the ITS gene sequence showed a sufficient amount of informative characters for phylogenetic analysis (~3%), the chloroplast showed less than 0.2% informative characters across the three genes investigated. This mirrors the hypothesis by Wolfe *et al.* (1987) that chloroplast DNA evolved at a slower rate than nuclear DNA.

Kress *et al.* (2005) have emphasized the importance of using multiple gene regions for the barcoding of species, as demonstrated in their study of several angiosperm species. The nuclear and chloroplast regions used during the present study were based on suggested sequences by the Barcode of Life initiative, where they have been effectively used to discriminate between other species groups (Chen *et al.,* 2010). The low number of nucleotide substitutions observed in the chloroplast genome indicates a young, closely related species group, as diversification of plastid genomes were at a low level. As the chloroplast is maternally inherited, the expected sequence variation is slower than the nuclear gene regions, and this was the case for the *S. nigrum* complex (Palmer, 1987). Based on the above results, it is hypothesized that the chloroplast gene regions should not be used during the Barcoding initiative for the *S. nigrum* complex. In addition, it is proposed that additional nuclear genes and chloroplast intergenic regions should be used during the multi-gene approach to positively identify these plants.

© University of Pretoria

4.4.1.1 Concatenated ITS, *matK*, *ndhF* and *rbcL*

A combined phylogram was constructed based on nuclear ITS and chloroplast *matK*, *ndhF* and *rbcL* gene regions. The congruence between datasets was found to be 0.086, which is higher than the minimum required value of 0.05, therefore effectively supporting the data to be combined. The resulting phylogenetic tree shows the formation of two major clusters with *S. lycopersicum* forming an expected outgroup to the black nightshade accessions. The reference samples, *S. americanum*, *S. burbankii*, *S chenopodioides*, *S. retroflexum* and *S. scabrum* were used as initial references in the tree topology to assign putative identities to unidentified accessions. Three main clusters were identified, and separated based on sound bootstrap confidence values of above 100% and 70% respectively. Cluster I was composed of *S. chenopodioides* (reference sample), LS1, LS4 and LS14, while cluster II was composed of *S. americanum* (reference sample), together with LS7 and LS5 (putatively identified as *S. americanum*). Cluster III consisted of the reference samples *S. retroflexum*, *S. burbankii* and *S. scabrum* together with unidentified accessions LS8, PC3, PC4, PC8 and PC9. These samples were putatively identified as *S. retroflexum* during collection based on morphology and this assumption was supported by the combined phylogenetic data. The tree provides strong support that *S. retroflexum* is closely related to *S. americanum*, with *S. chenopodioides* forming a separate, distant clade. Cluster I and II are separated from cluster III by a bootstrap value of 100%, supporting the hypothesis that these were very likely separate species groups, while clusters I and II were further separated by a 70% bootstrap confidence value indicating a robust possibility of separate species groups. The differences observed in the combined tree were most likely deduced from the differences observed in the ITS sequences, as this was the major contributor of parsimony informative sites compared to the chloroplast gene regions.

4.4.1.2 Direct sequencing of ITS

The presence of possible hybrid nightshades was suspected for certain samples based on inconsistent morphological character states. The ITS gene region for each accession was directly sequenced from PCR amplicons in duplicate, and resulting chromatograms analyzed. Sequences obtained for accessions LS1, LS2, LS3, LS4, LS5, LS6, LS8, LS9, LS10, LS11, LS13, LS14, LS15, PC1–PC10, BUR, RET, CHEN and SCA were significantly similar to the cloned sequence and no abnormalities could be distinguished based on chromatogram investigations. Accessions LS7, LS12 and AME had informative chromatogram results, with primary and secondary peaks at eight base pair sites indicating the presence of two allelic

© University of Pretoria

variants (haplotypes). It became apparent that these accessions are heterozygous at this allele and were possibly composed of genomes from two different homozygous parent lines. Further support for hybridization theory is the unexpected chromosomal number for these accessions. The anticipated diploid nature of the reference, AME, was not supported as described by Jacoby *et al.,* (2003). Instead a hexaploid chromosomal number was determined by flow cytometry (Discussed in Chapter 3). Hexaploid ploidy numbers were also recorded for accessions LS7 and LS12.

Basic sequence alignments of the allelic variants together with the rest of the accessions indicated that the "a" variant corresponds to tetraploid accessions (LS6, LS8, LS11, LS13, PC1–PC10, RET and BUR), while the "b" variant corresponds to the sequence observed for the diploid accessions (LS2, LS3, LS4, LS5, LS9, LS10, LS14, LS15 and CHEN) based on the eight polymorphic sites identified (Table 4.9). The diploid accessions were further subdivided into two groups based on sequence and morphological data. This implies that LS2, LS5 and LS9 were separated from LS1, LS3, LS4, LS10, LS14, LS15 and CHEN based on four of the eight SNP sites described in the ±600bp sequence. The latter group has been putatively identified as *S. chenopodioides*, while the first group was putatively identified as *S. americanum* (LS2, LS5 and LS9). The unique SNP markers identified indicated two allelic variants (haplotypes), possibly originating from two different parents. Comparisons of SNP identities from the first haplotype indicated a putative identity of *S. retroflexum* (LS6, LS8, LS11 and LS13), while the second haplotype could either be inferred from the putative *S. americanum* (LS2, LS5 or LS9) or the *S. chenopodioides* groups (LS1, LS3, LS4, LS10, LS14 and LS15). It was established that the "b" variant of LS7, LS12 and AME corresponds to the sequence of LS2, LS5 and LS9 at the four nucleotide positions (106, 129, 377 and 435), indicating that *S. americanum* was a probable donor during a previous hybridization event. Two different alleles of one locus within one plant may indicate the presence of a hybridization event between two different species. Hybridizations between these groups (*S. retroflexum* (tetraploid) and *S. americanum* (diploid)) would explain the hexaploid nature of the hybrids through alloploidy. Hybridization events are considered rare in natural habitats, but do occur sporadically under ideal conditions. These events are typified by possible backcrossing to a parent line, leading to introgression of the new alleles (Edmonds, 1979). These in turn may affect the morphology of the plants. This event is of importance, as the possibility of breeding programs in artificial settings become possible. It has been previously reported that taxa formed by hybridizations are linked by splits (or intermediate branch divergences) to their potential parents (Poczai and Hyvönen, 2010).

The haplotype sequences of LS7, LS12 and AME were included in the maximum likelihood phylogenetic analysis of ITS sequences, and the resulting tree shows the formation of four clusters. LS1, LS3, LS4,

© University of Pretoria

LS10, LS14, LS15 and CHEN forms cluster I and have a putative identity of *S. chenopodioides* based on the groupings with the reference sample. Cluster II, III and IV were separated from cluster I by ~80% bootstrap confidence. Cluster II consisted of LS2, LS5 and LS9 which were identified as *S. americanum* based on morphological data and the grouping together with previously identified *S. americanum* Genbank database sequence. The "b" haplotypes of LS7, LS12 and AME also form part of this group. This cluster was supported by a high bootstrap value of 99%, indicating the identity of the "b" haplotype to possibly be a diploid *S. americanum*. The *S. nigrum* Genbank sequence is closely related to this group, but separates from the cluster with 98% confidence values. This relationship of *S. nigrum* to *S. americanum* is in contrast to what was found by Poczai *et al.* (2008), and may be due to the under-representation of *S. nigrum* accessions in the present study. The possibility of *S. nigrum* being the progenitor of the "b" haplotype was disregarded due to the high confidence separation from *S. americanum*, as well as morphological discrepancies.

Cluster III was supported by a 99% bootstrap confidence value and consisted of RET, BUR, SCA, LS6, LS8, LS11, LS13 and PC1–PC10. The accessions within the cluster were putatively identified as *S. retroflexum*. This cluster included accessions from both farms, indicating the close relationships between the two varieties of *S. retroflexum* based on a single gene. Further investigations are needed to confirm this hypothesis. Cluster IV was composed of the "a" haplotypes of LS7, LS12 and AME. These accessions were hypothesized to be closely related to *S. retroflexum* based on direct sequence alignments, and this was mirrored in the phylogenetic tree where they were weakly separated with a bootstrap confidence value of 60%. However, the tree topology suggests the close relationship between the "a" haplotype and *S. retroflexum* and distant relatedness to *S. chenopodioides*. Analysis using only ITS sequence data closely mirrored the sequence alignments obtained during multiple sequence analysis. The close relationship observed between the unknown accessions (LS7, LS12 and AME) and *S. americanum* (LS2, LS5 and LS9) and *S. retroflexum* (LS6, LS8, LS11 and LS13; from the same farm) may indicate a hybridization event between these parental lines. *S. chenopodioides* was more distantly related to the unknown LS7, LS12 and AME and not likely to be a potential diploid progenitor.

These results obtained from ITS and chloroplast gene sequencing was to our knowledge the first report for the *S. nigrum* species complex in South Africa. DNA sequence analysis has not been so widely used in this group, but Poczai *et al.* (2010) used ITS data to determine the origin of an ambiguous *S. scabrum* accession from the 18th century collected by Paulus Kitaibel. Even though ITS gives adequate resolution between unknown accessions, chloroplast gene sequences are of little practical use during classifications or phylogenetic analysis. It is therefore suggested that in contrast to utilizing only plastid regions as proposed by the CBOL group (2010), nuclear gene(s) should be employed for species determination in the black nightshade complex. This proposal was supported by Chen *et al.* (2010), where the authors

© University of Pretoria

suggested the use of ITS2 as a universal barcode following the investigation of more than 6600 closely related medicinal plant species. Their results indicated that by using only the ITS2 gene region it was possible to effectively distinguish between 92.7% of the plants on species level.

4.4.2 Diversity array technology

Diversity array technology allows the simultaneous investigation of large numbers of polymorphic characters for phylogeny inference. Markers adhering to selection criteria of a high call rate (above 95%), high reproducibility (above 95%) and high polymorphism information content (PIC) values (above 0.37) were utilized for genotyping. Markers not adhering to these quality control measures were removed from the analysis. A total of 1145 polymorphic DArT markers were identified and utilized in subsequent phylogenetic analysis using PAUP software.

The most parsimonious dendogram resulted in the formation of four major clusters. *S. scabrum* was used as an out-group, as it is morphologically the most different to the unknown accessions. Cluster I and II are separated by a bootstrap confidence value of 100%. Cluster I consisted of the reference sample *S. chenopodioides* together with collected accessions putatively identified as *S. chenopodioides*. Cluster II consisted of accessions LS2, LS5 and LS9 which were putatively identified as *S. americanum* during collection from the farms. Cluster III consisted of the reference samples *S. retroflexum* and *S. burbankii* in conjunction with putatively identified *S. retroflexum* accessions. The IVth cluster (LS7, LS12 and AME) formed step-wise additive branches outwards from clusters I and II with high confidence bootstrap values of 95%–100%. The position of these accessions indicated that they are not forming a discrete cluster by themselves, but are rather forming intermediate branches indicating a close relationship between members in cluster I or II, as well as cluster III. The results could be indicative of a hybrid that therefore falls between better supported species groups.

The DArT library origin of each marker was determined and the results indicate that the majority of markers for each accession originated from their contributing species library as expected. The majority of the markers present in LS1, LS3, LS4, LS10, LS14 and LS15 originate from the *S. chenopodioides* library; whilst LS2, LS5 and LS9 originates from the *S. americanum* library and LS6, LS8, LS11, LS13, and PC1–PC10 originate from the *S. retroflexum* libraries. Significantly, the unknown AME, LS7 and LS12 accessions show equal numbers of markers from the *S. retroflexum* and *S. americanum* libraries which may indicate the presence of several fragments from both these species groups. Furthermore, each polymorphic marker identity was compared between accessions. If the marker is shared between two

© University of Pretoria

accessions, it was assumed that the two accessions are related. All 1145 markers were compared between all 30 accessions. It is clear that the CHEN reference accession matches the polymorphic markers for LS1, LS3, LS4, LS10, LS14 and LS15 with <10% marker discrepancies. The AME reference closely relates to LS7 and LS12, with >90% marker identity matches. BUR and RET reference accessions closely mirror markers from LS6, LS8, LS11, LS13 and PC1–PC10, having more than 89% marker similarity. SCA is relatively closely related to LS6, LS8, LS11, LS13 and PC1–PC10 with more than 70% marker similarity, but if taken into account that just less than 30% of the markers are dissimilar, it can be concluded that although they may be closely related, they are not the same species. These results match the putative identities assigned during collection of the plant accessions.

The DArT tree mirrors results obtained for ITS sequence data, where the putative *S. retroflexum* accessions from two different farms group together. This emphasizes the result that regardless of geographical location, the molecular composition of the plants remains the same. The results obtained during the present study are in contrast to results previously reported. Poczai *et al.* (2008) investigated the diversity of *Solanum* accessions using RAPD technique, and the resulting dendogram based on 235 polymorphic markers illustrated a closer relationship between *S. retroflexum* and *S. chenopodioides*, which form a cluster away from *S. americanum*. In a separate study, simple sequence repeat techniques were used to determine the generic relationship between members of the black nightshade complex in South Africa, based on 232 polymorphic markers (Van Biljon *et al.,* 2010). The authors showed a closer relationship between *S. chenopodioides* and *S retroflexum*, while *S. americanum* forms a separate sister cluster. In the present study, when investigating DArT polymorphic markers *S. chenopodioides* is closely related to *S. americanum*, and are more distantly related to *S. retroflexum*. This result was in contrast to previous studies. It should always be kept in mind that identifications of taxa might not have been congruent in the different studies. Even though DArT shows a highly similar result to other marker technologies, the phenomenon of increased resolution could be attributed to the fact that the DArT technology utilized more of the genome due to its high throughput nature (thus more polymorphic markers identified) than other PCR or restriction enzyme based technologies. The technique was therefore identifying additional differences or similarities between the groups causing them to group in a different way.

A combination tree utilizing DArT polymorphic markers and ITS sequence data was constructed using parsimony calculations. The resulting tree is extremely similar to the dendogram obtained for DArT polymorphic data only. This fact could be due to the low amount of informative sites which are being contributed by the ITS sequence (557bp with 19 informative sites) in comparison to the DArT polymorphic markers (1145 informative), which also explains the small amount of bootstrap variation which could be observed between the DArT tree and the DArT/ITS combination tree.

© University of Pretoria

4.4.3 Concluding remarks

A highly similar tree topology was observed when comparing the phylogram obtained through gene sequencing data and the dendogram constructed using DArT polymorphic markers. The main difference observed between the two datasets was the close relationship between the putative *S. retroflexum* and *S. americanum* sub-clusters, with *S. chenopodioides* forming a separate main cluster in the ITS results. In contrast, when analyzing the DArT polymorphic markers, *S. americanum* groups together with *S. chenopodioides*, while *S. retroflexum* forms an independent main cluster. All species groupings relative to reference samples closely mirror initial putative identifications from the field with the exception of LS7, LS12 and AME. The possibility of these accessions being the same entity formed through a single hybridization event was confirmed by the tree topology observed for both ITS sequence and DArT polymorphic data. The hypothesis is further supported by morphological character analysis, as these three accessions are identical to each other, but separated from the remaining accessions. These findings contradict observations of Jacoby *et al.,* (2003), from whom these reference samples had been obtained. Jacoby *et al.* (2003) reported the identity of AME to be *S. americanum*, supported by morphological, AFLP and ploidy data. In the present study however, AME was found to be hexaploid and based on the secondary peaks observed in sequence data, more probably a hybrid developed from two different parental lines. An increased resolution of accession relationships is observed for the DArT data, indicating the importance of an increased number of informative sites for species resolution. However, for initial species identification, the ITS gene region was sufficient and could effectively be used during the Barcoding of Life strategy to create a universal signature for each species. These results indicate the difficulties encountered by basing species identification solely on one molecular technique. Although delineations seem relatively straightforward, the supplementation of molecular data by morphological characteristics may shed additional light on true identities, especially if unknown hybridization events had occurred.

© University of Pretoria

Chapter 5

General Discussion, Identification Keys and Conclusions

© University of Pretoria

The *S. nigrum* complex comprises a taxonomically complicated assemblage of species, with no clear set of morphological or molecular data for effective distinctions among them. Several authors have attempted to delineate species within this complex, with varying success. Morphological keys have been developed for Eurasian (Schilling and Andersen, 1990), Australian (Henderson, 1974), New Zealand (Healy, 1974), North American (Schilling, 1981) and South American (Edmonds, 1972) members of this complex. These authors' conclusions are in broad agreement, and they all emphasized the difficulties in scoring traits when environmental factors influence the characters. Recently, increased attention has been given to studying the underlying genetic differences between the plants, as the molecular traits are considered a more reliable approach. Whole genome marker studies seem to be the most useful of molecular approach which is employed to study the nightshades, while single gene analysis methods are not readily used.

Even though several studies have been performed on the black nightshade species complex across the world, information related to African, and specifically the South African species, is severely lacking. This species group holds several advantages to humans, the most important being highly nutritious, rich in important metabolites, easily cultivated, extensive medicinal properties and having potential resistance mechanisms against several plant pathogens. However, because little information concerning the morphology and genetics of the species grown and naturalized in South Africa are available, it is almost impossible to determine which entities would be the most useful to use in future breeding programs to exploit these advantages. A few research initiatives in South Africa start to address these deficiencies and during the present study it was attempted to shed additional light on the complex taxonomic relationships within the black nightshade species group. The previous chapters dealt in detail with i) morphological aspects and ploidy levels (Chapter 3) and ii) molecular aspects (Chapter 4) of members of the *S. nigrum* complex used for commercial jam production in the Highveld region of South Africa. The main conclusions reached in each chapter are summarized below.

Ecologically, the accessions from South Africa mirror what has been recorded in literature for field grown plants. Germination generally occurred at temperatures of 15–30˚C at moderate light intensities and regular watering was needed for effective growth. Black nightshade plants are under heavy environmental influences and it is hypothesized that most morphological characters shows variation in response to varying temperature, light and water levels (Ogg *et al.,* 1981). The current study was performed in a controlled greenhouse environment and the environmental fluctuations were kept to a minimum. It was noted that the position of each plant within the glasshouse may have had a small effect, as the light intensities were not precisely equal throughout the glasshouse. Thirty different morphological characters were measured, observed or quantitatively determined. The traits included several morphological and

© University of Pretoria

cytological structures such as the entire plant, leaves, flowers, berries, seed and ploidy levels. When all characters were combined and principle component analysis performed, four possible species groups were identified. Their identities were inferred based on their groupings together with previously identified reference samples obtained from the University of the Free State; *S. americanum*, *S. burbankii*, *S. chenopodioides*, *S. retroflexum* and *S. scabrum*. In addition, type specimens of certain species names lend confidence to the putative identifications.

Molecular traits were investigated based on four gene region analyses, as well as a whole genome approach. The Barcoding of Life initiative aims to identify a single gene region for the effective identification of plant species across the world. The advantages of such a barcode cannot be overly emphasized for conservation and diversity studies and it was aimed to try and identify members of the *S. nigrum* complex based on suggested gene regions. One nuclear (ITS) and three chloroplast genes (*mat*K, *ndh*F, *and rbcL*) were explored to aid in the barcoding initiative. It was found that the nuclear ITS region provides sufficient parsimonious informative sites to enable phylogenetic separation within this complex. However, chloroplast analysis yielded insufficient informative sites and was ineffective in discriminating among members of this closely related species complex. A whole genome approach such as diversity array technology was more useful, as an increased amount of informative sites were obtained. DArT has the additional advantages of using a microarray method, where the need for sequence information and complex gel–based analysis methods are eliminated. DArT identifies polymorphic marker fragments spanning the whole genome and therefore increases the resolution of separation during phylogenetic analysis. Each species group identified based on the aforementioned methods is described below.

5.1 Cluster descriptions

S. chenopodioides group

Accessions LS1, LS3, LS4, LS10, LS14 and LS15 were highly similar to the CHEN (*S. chenopodioides*) reference samples and to the Type specimen for this name, based on morphological characters. These plants were characterized by a spreading, rather short habit (< 90 cm) with extensively pubescent, green stems. The leaves were highly pubescent adaxially and abaxially, elongated with an entire leaf margin, and had a very short petiole (< 30 mm). The flower corollas were also easily distinguishable as large, with long and narrow lobes which were deeply lobed. Style exertion was prominent, with more than 2 mm projection beyond the anthers. Berries were small (< 6 mm) and dull black and arranged on an erect peduncle growing straight from the stem apex. The calyx adhered to the berry until maturity and berries

© University of Pretoria

fall from the plant when mature. Seeds were small (< 1.5 mm), with large amounts of seed occurring per berry (20–40). Flow cytometry analysis indicated a diploid ploidy for these accessions. During ITS gene analysis, these accessions form a separate clade together with the CHEN reference sample. The *S. chenopodioides* group forms a sister clade to the *S. americanum* and *S. retroflexum* groups (99% bootstrap). DArT analysis yielded a highly similar result; with the CHEN clade separated from the sister clade *S. americanum* with 100% bootstrap confidence.

S. americanum group

Accessions LS2, LS5 and LS9 were not matching any of the reference samples. They did, however, morphologically closely resemble herbarium specimens identified as *S. americanum,* and were visually identified as *S. americanum* by Dr. Sandra Knapp (Natural History Museum, London, personal communication). It was therefore hypothesized that these accessions could be African (or South African) variants of the form of *S. americanum* presented by the Type specimen of the name. These plants were characterized by a spreading, short habit (< 90 cm) with slightly pubescent, greenish to purple stems. The leaves were slightly to moderately pubescent adaxially and abaxially, rhomboid to ovate in shape with a serrated leaf margin and had a very long petiole (> 50 mm). The flower corollas were also easily distinguishable as very small, with short, broad lobes fused at the base of the anthers. Style exertion was absent with no projection beyond the anthers. Berries were small (< 6 mm), shiny and purplish-black, and arranged on erect peduncle growing straight from the stem apex. The calyx reflexed from the berry upon maturity and fall together with the berries from the plant when mature. Seeds were small (< 1.5 mm), with many seeds (20–50) per berry. Flow cytometry analysis indicated a diploid ploidy for these accessions. ITS analysis yielded an informative result, where the LS2, LS5 and LS9 accessions formed a clade which was separate from *S. chenopodioides* and *S. retroflexum*. Within this clade, the accessions grouped together with the *S. americanum* Genbank accession (emphasizing its putative identity) and “b” haplotype of LS7, LS12 and AME which were not positively identified. It is hypothesized that *S. americanum* could be the original progenitor of the “b” haplotype of LS7, LS12 and AME. Polymorphic analysis with DArT suggested a closer relationship between LS2, LS5 and LS9 and the *S. chenopodioides* clade, while *S. retroflexum* was a further related group.

S. retroflexum group

Accessions LS6, LS8, LS11, LS13 and PC1–PC10 were similar to the RET (*S. retroflexum*) and BUR (*S. burbankii*) reference samples, as well as the Type specimens for these names when morphological traits

© University of Pretoria

are compared. *S. retroflexum* and *S. burbankii* are considered synonyms for the same taxonomic entity, due to the identical morphological characters observed. These two species are thought to be the "Sunberry" which was promoted by the plant breeder Luther Burbank in North America at the beginning of the century. Burbank claimed success in crossing *S. scabrum* with *S. villosum* to create *S. burbankii*, but this was later disproved and it is thought that *S. burbankii* is in fact *S. retroflexum* (claimed to be native to sub-Saharan Africa), imported to North America (Edmonds and Chweya, 1997). These plants were characterized by spreading, moderately short plants (< 90 cm) with slightly pubescent, greenish to purple stems. The leaves were moderately pubescent adaxially and abaxially, rhomboid to ovate in shape with a serrated leaf margin and had a reasonably short petiole (20–40 mm). The flower corollas were moderate in size with short, broad lobes which fuse halfway to the base of the anthers. A faint purple stripe was present on the corolla lobes of flowers, but was not found constantly across all flowers of all accessions. Edmonds (1977) and Henderson (1974) have also commented on the presence of a purple stripe, and the authors concluded that it may be caused by variations in anthocyanin content due to differing light intensities. In the present study, the stripes could be due to the relative positions of plants within the glasshouse. Style exertion was observed with slight (up to 1 mm) projection beyond the anthers. Berries were large (6–8 mm), dull black and arranged on a deflexed peduncle growing from the stem apex. The calyx reflexed from the berry upon maturity and fell from the plant together with the berries when mature. Seeds were very large (1.5–2 mm), with few seeds per berry. Flow cytometry analysis yielded a tetraploid ploidy for these accessions. This group was closely related to the *S. americanum* cluster based on ITS region analysis. *S. burbankii* and *S. scabrum* also form part of the *S. retroflexum* group as was also observed by Jacoby *et al.* 2003. Whole genome DArT analysis indicates a similar result, but the *S. retroflexum* clade was distantly related to *S. americanum* and *S. chenopodioides,* the latter of which fall into the same clade. All putatively identified *S. retroflexum* accessions grouped into the same clade, together with the positively identified reference sample.

S. scabrum and an unidentified hexaploid species group

Based on morphology, the fourth species group consisted of the reference SCA (*S. scabrum*), as well as three other accessions, LS7, LS12 and AME (which were classified as *S. americanum* during collection). SCA was characterized by tall (> 1 m), erect plants with slightly pubescent, green stems. The leaves were very slightly pubescent adaxially and abaxially, ovate in shape with an entire margin and had a long petiole (> 50 mm). The flower corollas were moderate in size with short, broad lobes which fused halfway to the base of the anthers. Style exertion was not observed, but the anthers showed a characteristic brown colour compared to the bright yellow anther colours of the remaining accessions. Berries were very large (> 13 mm), dull black, and arranged on a branched, erect peduncle growing from

© University of Pretoria

the stem apex (characteristic of *S. scabrum*; Edmonds and Chweya, 1997). The calyx reflexed from the berry upon maturity, but berries remained on the plant for their entire life span. Seeds were large (1.5–2 mm), with many seeds (> 90) per berry. Flow cytometry analysis yielded a hexaploid chromosomal amount for SCA.

Accessions LS7, LS12 and AME shared many characteristics with SCA, but differed substantially in berry, flower and peduncle morphology. LS7, LS12 and AME shared a similar morphology for plant growth type (tall and erect, > 1 m), slightly pubescent stems and leaves (slightly pubescent, ovate shape and entire margin). The differences from *S. scabrum* accessions included having a moderately shiny berry type, with the calyx remaining adhered to the mature fruit and having a deflexed peduncle away from the stem apex. In terms of flowers, LS7, LS12 and AME differed from SCA in having yellow anthers and having slight style protrusion (up to 1 mm). Flow cytometry analysis yielded a hexaploid ploidy for LS7, LS12 and AME. Even though they were highly similar to SCA, the morphological discrepancies indicate the possible presence of a different, closely related, species group; or hybrid formations from two other parental lines which resulted in averaged morphological traits. ITS gene sequence analysis indicated the presence of two haplotypes ("a" and "b") for these three accessions, where primary and secondary peaks were observed at several positions within the 550bp sequence. ITS phylogenetic analysis yielded four main clusters, with the "a" haplotype forming one cluster (IV) and haplotype "b" falling into another cluster (II). The "b" haplotype clustered together with LS2, LS5 and LS9 (putatively identified as *S. americanum*), while the "a" haplotype did not cluster with additional accessions. The "a" cluster seemed to be closer related to the *S. retroflexum* species group and it was hypothesized that LS7, LS12 and AME could possibly be a hybrid between LS2, LS5 or LS9 and an accession within the *S. retroflexum* group. The possibility that these hexaploid accessions could be the native form of *S. nigrum* was disregarded due to morphological inconsistencies between the accessions and the type specimen. In addition, the ITS tree shows a separation of 98% bootstrap of *S. nigrum* from the "b" haplotypes of LS7, LS12 and AME. It is possible that these three accessions are the progeny of a single hybridization event, due to their identical morphology, hexaploid chromosomal number and phylogenetic clustering. These findings were further supported by the DArT results, where LS7, LS12 and AME formed a step-wise out-grouping from the LS2, LS5 and LS9 sub-cluster and the *S. retroflexum* main cluster with 100% bootstrap confidence values.

The abovementioned groupings and characteristics are in line with what Edmonds and Chweya (1997) published as a guideline to the effective identification of the black nightshade species. Morphological traits which could be observed from herbarium "type" specimens of names concur with what had been established in the present study. Based on morphology, ploidy and molecular analysis, putative species

© University of Pretoria

identities were assigned to the individual accessions (Table 5.1). Discrepancies between initial and final identifications are indicated in bold font.

Table 5.1 Putative and revised identities of *Solanum* accessions

Accession	Putative identity upon collection	Revised identity*
LS1	*S. chenopodioides*	*S. chenopodioides*
LS2	*S. americanum*	*S. americanum*
LS3	*S. chenopodioides*	*S. chenopodioides*
LS4	*S. chenopodioides*	*S. chenopodioides*
LS5	*S. americanum*	*S. americanum*
LS6	*S. retroflexum*	*S. retroflexum*
LS7	*S. americanum*	**Solanum spp.**
LS8	*S. retroflexum*	*S. retroflexum*
LS9	*S. americanum*	*S. americanum*
LS10	*S. chenopodioides*	*S. chenopodioides*
LS11	*S. retroflexum*	*S. retroflexum*
LS12	*S. americanum*	**Solanum spp.**
LS13	*S. retroflexum*	*S. retroflexum*
LS14	*S. chenopodioides*	*S. chenopodioides*
LS15	*S. chenopodioides*	*S. chenopodioides*
PC1	*S. retroflexum*	*S. retroflexum*
PC2	*S. retroflexum*	*S. retroflexum*
PC3	*S. retroflexum*	*S. retroflexum*
PC4	*S. retroflexum*	*S. retroflexum*
PC5	*S. retroflexum*	*S. retroflexum*
PC6	*S. retroflexum*	*S. retroflexum*
PC7	*S. retroflexum*	*S. retroflexum*
PC8	*S. retroflexum*	*S. retroflexum*
PC9	*S. retroflexum*	*S. retroflexum*
PC10	*S. retroflexum*	*S. retroflexum*
AME	*S. americanum*	**Solanum spp.**
BUR	*S. burbankii*	***S. retroflexum* = *S. burbankii***
CHEN	*S. chenopodioides*	*S. chenopodioides*
RET	*S. retroflexum*	*S. retroflexum*
SCA	*S. scabrum*	*S. scabrum*

* Based on morphological traits of greenhouse grown plants, molecular analysis and comparisons of Type specimens.

The origin of the *S. nigrum* complex is claimed to be South America (Edmonds and Chweya, 1997) and *S. americanum* and *S. chenopodioides* are considered to be native to South America. These species are

© University of Pretoria

thought to be naturalized in South Africa and can now be viewed as established and self sustaining in this region. *S. retroflexum* is claimed to be native to South Africa (Edmonds and Chweya, 1997; Jacoby *et al.,* 2003; Van Averbeke *et al.,* 2007a), while *S. scabrum* is possibly native to several African countries. Very little accurate data on the distribution and origin of *S. scabrum* has been published, leaving doubt as the precise region of origin. However, this species was probably introduced relatively recently into South Africa and has since become naturalized. Very little or no information is available on the fossil record and historical evidence of black nightshade species in South Africa. The oldest record to date for a black nightshade in South Africa is that by Thunberg (1794) where an unknown black nightshade had been collected at Paardeberg area, near Malmesbury in the Western Cape, South Africa. Inadequate morphological characterizations of this accession make the correct identification of this plant difficult, but it is thought that it may be *S. retroflexum*. No known frequency of distribution or records of introduction exists for this species group. However, recently it has been recorded that *S. retroflexum* is most commonly found growing extensively in all provinces in South Africa (indicating its possibly native status), while the sporadic occurrences of other nightshade species possibly indicate that they are alien. There is relatively low genetic diversity between accessions of the same nightshade species group, indicating very recent diversification events and that the introduction of these species may have occurred in the recent past. Investigating the genetic differences among small populations in natural habitats and weed populations from undisturbed natural areas may also give an indication of whether the plants are native, but no conclusions can be reached at this stage. Based on available information, *S. retroflexum* is the only species likely to be native to South Africa. The remaining species have all been introduced quite recently and are in the process of becoming naturalized.

5.2 Identification keys

Variations in morphological features are often poorly represented in pictures and for that reason taxonomists prefer to use a taxonomic key in written form. The key is a more comprehensive summary of all morphological features. In a dichotomous key, the statements featured lead to multiple alternative identities. By following the features which apply to your plant, the identity can be determined as in a flow chart. For multi-access keys, statements (character states) may be accessed in any order; simply skip statements you cannot answer based on the available plant material. Only pay attention to statements that apply to your plant; ignore those that do not. In the case of an applicable statement, compare your plant only with the one or more species listed under that statement. If a statement fits your plant and only a

© University of Pretoria

single species is listed for it, then you plant is most probably that species. If more than one statement applies to your plant, only consider those species common to all the positive statements.

In the keys below *Solanum* sp. refers to accessions LS7, LS12 and AME. Based on the results of morphological, cytological and molecular analysis, these three accessions are provisionally not considered a separate species, but a possible hybrid between *S. americanum* and *S. retroflexum*.

Dichotomous keys:

All characters

1a Plants diploid (2n = 24)……..……...……..…..**2**

1b Plants polyploid..……..……...……..**3**

2a Leaf margin entire……...…….***S. chenopodioides***

2b Leaf margin serrate……...…….……..***S. americanum***

3a Plants tetraploid (2n = 48)…….….***S. retroflexum***

3b Plants hexaploid (2n = 72)……..……..**4**

4a Anthers brown to brownish orange…..…….……..***S. scabrum***

4b Anthers bright yellow…..…...……...……***Solanum* sp.**

Vegetative characters

1a Stems densely hairy……...……..…….***S. chenopodioides***

1b Stems sparsely hairy to glabrous……..**2**

2a Leaf margin entire……***S. scabrum, Solanum* sp.**

2b Leaf margin serrate……...……**3**

© University of Pretoria

3a Stems purplish……..……..……..……..……***S. americanum***

3b Stems green……..……***S. retroflexum***

Vegetative and reproductive characters

1a Leaf margin entire……..……..……..……..

4a Seed < 1.5 mm in diameter…….***S. chenopodioides***

4b Seed > 1.8 mm in diameter……..…….***S. retroflexum***

Multi-access key

1. Plants spreading, usually < 900 mm tall
    a. *Solanum americanum*
    b. *Solanum chenopodioides*
    c. *Solanum retroflexum*

2. Plants erect, usually > 900 mm tall
    a. *Solanum scabrum*
    b. *Solanum* sp.

3. Stems densely hairy, green
    a. *Solanum chenopodioides*

4. Stems sparsely hairy to glabrous, green
    a. *Solanum americanum*
    b. *Solanum retroflexum*

5. Stems sparsely hairy to glabrous, purplish
    a. *Solanum* sp.

6. Leaf margin entire
    a. *Solanum chenopodioides*
    b. *Solanum scabrum*
    c. *Solanum* sp.

7. Leaf margin serrate
    a. *Solanum americanum*
    b. *Solanum retroflexum*

8. Leaf shape lanceolate
    a. *Solanum chenopodioides*

9. Leaf shape rhomboid to ovate
    a. *Solanum americanum*
    b. *Solanum retroflexum*
    c. *Solanum scabrum*
    d. *Solanum* sp.

10. Leaf lamina ratio length to width > 1.6
    a. *Solanum americanum*
    b. *Solanum chenopodioides*

11. Leaf lamina ratio length to width < 1.5
    a. *Solanum retroflexum*
    b. *Solanum scabrum*
    c. *Solanum* sp.

12. Leaf lamina densely hairy
    a. *Solanum chenopodioides*

13. Leaf lamina sparsely hairy to glabrous
    a. *Solanum americanum*
    b. *Solanum retroflexum*
    c. *Solanum scabrum*
    d. *Solanum* sp.

14. Corolla lobe ratio length to width > 2
    a. *Solanum chenopodioides*

15. Corolla lobe ratio length to width < 2
    a. *Solanum americanum*
    b. *Solanum retroflexum*
    c. *Solanum scabrum*
    d. *Solanum* sp.

16. Style project beyond anthers
    a. *Solanum chenopodioides*
    b. *Solanum retroflexum*
    c. *Solanum* sp.

17. Style does not project beyond anthers
    a. *Solanum americanum*
    b. *Solanum scabrum*

18. Anthers bright yellow
    a. *Solanum americanum*
    b. *Solanum chenopodioides*
    c. *Solanum retroflexum*
    d. *Solanum* sp.

19. Anthers brown to brownish orange
    a. *Solanum scabrum*

20. Pedicels erect
    a. *Solanum americanum*
    b. *Solanum chenopodioides*

© University of Pretoria

21. Pedicels deflexed
    a. *Solanum retroflexum*
    b. *Solanum* sp.

22. Peduncle branched
    a. *Solanum scabrum*

23. Number of pedicels per peduncle < 4
    a. *Solanum americanum*

24. Number of pedicels per peduncle 4–7
    a. *Solanum chenopodioides*
    b. *Solanum retroflexum*

25. Number of pedicels per peduncle > 7
    a. *Solanum scabrum*
    b. *Solanum* sp.

26. Fruit size < 5 mm in diameter
    a. *Solanum chenopodioides*

27. Fruit size 5–8 mm in diameter
    a. *Solanum americanum*
    b. *Solanum retroflexum*
    c. *Solanum* sp.

28. Fruit size > 8 mm in diameter
    a. *Solanum scabrum*

29. Fruits dull
    a. *Solanum chenopodioides*
    b. *Solanum retroflexum*
    c. *Solanum scabrum*

30. Fruits moderately shiny
    a. *Solanum* sp.

31. Fruits very shiny
    a. *Solanum americanum*

32. Calyx lobes reflexed from base of fruit
    a. *Solanum americanum*
    b. *Solanum retroflexum*

33. Calyx lobes adherent to base of fruit
    a. *Solanum chenopodioides*
    b. *Solanum scabrum*
    c. *Solanum* sp.

34. Seed size > 1.6 mm in diameter
    a. *Solanum retroflexum*
    b. *Solanum scabrum*
    c. *Solanum* sp.

35. Seed size < 1.5 mm in diameter
    a. *Solanum americanum*
    b. *Solanum chenopodioides*

36. Seeds number per berry < 20
    a. *Solanum retroflexum*

37. Seeds number per berry 20–30
    a. *Solanum* sp.

38. Seeds number per berry 30–40
    a. *Solanum chenopodioides*
    b. *Solanum americanum*

39. Seeds number per berry > 40
    a. *Solanum scabrum*

40. Plants diploid (2n = 24)
    a. *Solanum americanum*
    b. *Solanum chenopodioides*

41. Plants tetraploid (2n = 48)
    a. *Solanum retroflexum*

42. Plants hexaploid (2n = 72)
    a. *Solanum scabrum*
    b. *Solanum* sp.

© University of Pretoria

5.3 Conclusion

The principle aim of the present study was to identify members of the *S. nigrum* complex which is grown for jam production in the Highveld region of South Africa. The putative species groupings of 30 diverse accessions based on morphology closely mirrored the groupings based on molecular analysis. The various accessions could be assigned species names with high degrees of confidence based on the variety of techniques used. Putative identities initially assigned during collecting on farms closely corresponded to the final identities based on all evidence considered during the present study. Species groups identified were *S. americanum*, *S. retroflexum*, *S. chenopodioides* and an entity possibly representing plants of hybrid origin (LS7, LS12 and AME). It was suggested that the hybrid originated from successful crosses between a member of the *S. americanum* and *S. retroflexum* species groups. The putative hybrid species poses the possibility for future breeding programs between members of the complex with conflicting ploidy levels. If hybridization occurs naturally, it should also be possible to initiate hybridizations in an artificial setting.

Informative morphological traits included stem and leaf texture, leaf shape, leaf margin, leaf indumentum, shape of flower, colour of anthers, style exertion, berry size, berry colour, calyx orientation, origin of pedicels, seed size, number of seeds per fruit and chromosomal numbers. Traits with low taxonomic value included plant height, plant growth form, stem colour, leaf area, corolla colour, number of berries per inflorescence, number of pedicels per inflorescence, seed surface texture and seed colour. Genetic diversity was resolved quite well by nuclear ITS gene analysis. Chloroplast gene analysis yielded very few informative sequence sites and was therefore not suggested to be used during the identification of a universal barcode for the members of the black nightshade complex. Diversity array technology effectively identified numerous polymorphic markers without the need for sequence data. Increased resolution on phylogenetic relationships was obtained for the DArT technology, but initial classifications can effectively be performed using ITS sequence data.

Farm A is growing three species groups, with the presence of a possible natural hybrid. *S. retroflexum*, *S. americanum* and *S. chenopodioides* are being cultivated in unison and therefore gave rise to the opportunity for hybridizations. No preference is given to which species is being cultivated, leading to interbreeding opportunities within the field. Farm B exclusively grows *S. retroflexum*. *S. retroflexum* is thought to be indigenous and is therefore well adapted to the South Africa climate, leading to relatively easy cultivation. However, it may be suggested to farmers that *S. scabrum* (which is also claimed to be

© University of Pretoria

native to several African countries) is a more useful species for farming for several reasons including, (i) the plants grow erect and take up less space in the field; (ii) the berries are larger; (iii) the berries grow on an erect stem apex, making harvesting easier and faster for field hands. However, metabolic analysis and consumer evaluations should be performed, to ensure that the jam taste is not compromised by using a different nightshade species.

5.4 Future prospects

Due to the limited information available for the black nightshade species complex in South Africa, the potential for many future studies exist. Additional accessions should be collected from domesticated farm areas as well as undisturbed natural habitats to investigate the degree of naturalization. These results in turn could shed additional light on the time from introduction into South African habitats. Additional morphological traits can be included, spanning in depth characterization of berry micro-morphology and seed traits. Molecular data can be expanded to include investigations into additional nuclear gene and intergenic (ITS2) regions as well as intergenic chloroplast regions. Intergenic regions can possibly yield additional informative sites, as the genic regions are usually subjected to evolutionary constraints. DArT polymorphic markers can be sequenced to create specific CAPS markers, which can be used to quickly and effectively identify unknown nightshade species without the need to investigate unstable morphological characters. Gene regions which are identified by DArT markers can also be used to start mapping desired traits (morphological or metabolic). The metabolic aspects of the different species groups may shed light on essential nutritional attributes of each cultivar and breeding programs can be initiated to exploit enhanced nutritional benefits, physiological attributes or resistance mechanisms. Additional studies will have to be performed to determine breeding potential between accessions with differing ploidy levels. In the shorter term, several physiological traits can be investigated to aid the farmers with optimal growth conditions and ideal harvesting techniques. The potential toxicity of the berries also requires investigation, together with research pertaining to the medicinal value of these plants. The isolation and positive identification of active compounds can reveal the potential of these plants as possible natural remedies against several diseases.

© University of Pretoria

References

© University of Pretoria

Acharyya S, Patra A, Bag PK. 2009. Evaluation of the antimicrobial activity of some medicinal plants against enteric bacteria with particular preference to multi-drug resistant *Vibrio cholera*. *Tropical Journal of Pharmaceutical Research* **8**: 231-237

Adedeji O, Ajuwon OY, Babawale OO. 2007. Foliar epidermal studies, organographic distribution and taxonomic importance of trichomes in the family Solanaceae. *International Journal of Botany* **3**: 276-282

Aliero AA, Grierson DS, Afolayan AJ. 2006. The foliar micro-morphology of *Solanum pseudocapsicum*. *Flora* **201**: 326-330

Amorim EP, Vilarinhos AD, Cohen KO, Amorim VBO, dos Santos-Serejo J, e Silva SO, Pestana KN, dos Santos VJ, Paes NS, Monte DC, dos Reis RV. 2009. Genetic diversity of carotenoid-rich bananas evaluated by Diversity Arrays Technology (DArT). *Genetics and Molecular Biology* **32**: 96-103

Aoki S, Ito M. 2000. Molecular phylogeny of Nicotiana (Solanaceae) based on the nucleotide sequence of the *mat*K gene. *Plant Biology* **2**: 316-324

Barthet MM, Hilu KW. 2007. Expression of *mat*K: Functional and evolutionary implications. *American Journal of Botany* **94**: 1402-1412

Baylis GTS. 1958. A Cytogenetical study of New Zealand forms of *Solanum nigrum* L., *S. nodiflorum* Jacq. and *S. gracile* Otto. *Transactions of the Royal Society of New Zealand*. **85**: 379-385

Beentje H. 2010. The Kew plant glossary: an illustrated dictionary of plant identification terms. *Royal Botanic Gardens, Kew*.

Becker J, Vos P, Kuiper M, Salamini F, Heun M. 1995. Combined mapping of AFLP and RFLP markers in barley. *Molecular and General Genetics* **249**: 65-73

Birky CW. 1995. Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution. *Proceedings of the National Academy of Sciences USA* **92**: 11331-11338

Bousquet J, Strauss SH, Doerksen AH, Price RA. 1992. Extensive variation in evolutionary rate of *rbc*L gene sequences among seed plants. *Proceedings of the National Academy of Sciences USA* **89**: 7844-7848

Bruneton J. 1995. *Pharmacognosy Phytochemistry Medicinal Plants*. Lavoisier Publishing, 875-878

© University of Pretoria

Buckler IV ES, Ippolito A, Holtsford TP. 1997. The evolution of ribosomal DNA divergent paralogues and phylogenetic implications. *Genetics Society of America* **145**: 821-832

CBOL Plant Working Group. 2010. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences USA* **106** 12794-12797

Chase MW, Cowan RS, Hollingsworth PM, Van Der Berg C, Madriñán S, Peterson G, Seberg O, Jørgensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M. 2007. A proposal for a standardized protocol to barcode all land plants. *Taxon* **56**: 295-299

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One* **5**: e8613

Clegg MT. 1993. Chloroplast gene sequences and the study of plant evolution. *Proceedings of the National Academy of Sciences USA* **90**: 363-367

Cuenoud P, Savolainen V, Chatrou LW, Powell M, Grayer RJ, Chase RJ. 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid *rbc*L, *atp*B and *mat*K DNA sequences. *American Journal of Botany* **89**: 132-244

Defelice MS. 2003. The black nightshades, *Solanum nigrum* L. *et al.* – Poison, Poultice, and Pie. *Intriguing World of Weeds* **17**: 421-427

Deschamps S, Campbell MA. 2010. Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. *Molecular Breeding* **25**: 553-570

Doležel J, Greilhuber J, Suda J. 2007. Flow cytometry with plants - an overview. In J Doležel, J Greilhuber, J Suda, eds., Flow *Cytometry with Plant Cells. Analysis of Genes, Chromosomes and Genomes*, Wiley–VCH Verlag GmbH & Co. KGaA, Weinheim, 41–65

Edmonds JM. 1972. A Synopsis of the taxonomy of *Solanum* Sect. *Solanum* (Maurella) in South America. *Kew Bulletin* **27**: 95-114

Edmonds JM. 1977. Taxonomic studies on *Solanum* section *Solanum* (Maurella). *Botanical Journal of the Linnaean Society* **75**: 141-178

© University of Pretoria

Edmonds JM. 1979. Biosystematics of *Solanum* L., section *Solanum* (Maurella) In: The *Biology and Taxonomy of Solanaceae*. Hawkes, Lester and Skelding. Academic Press, 529-548

Edmonds JM. 1983. Seed coat structure and development in *Solanum* L. section *Solanum* (Solanaceae). *Botanical Journal of the Linnaean Society* **87**: 229-246

Edmonds JM, Chweya JA. 1997. Black nightshade *Solanum nigrum* L. and related species. Promoting the conservation and use of underutilized and neglected crops. **15.** Institute of Plant Genetics and Crop Plant Research, Glatersleben/ *International Plant Genetic Resources Institute*, Rome.

Eijlander R, Stiekema WJ. 1993. Biological containment of potato *(Solanum tuberosum)*: outcrossing to the related wild species black nightshade *(Solanum nigrum)* and bittersweet *(Solanum dulcamara)*. *Sexual Plant Reproduction* **7**: 29-40

Ejoh RA, Dever JT, Mills JP, Tanumihardjo SA. 2010. Small quantities of carotenoid-rich tropical green leafy vegetables indigenous to Africa maintain vitamin A status in Mongolian gerbils (*Meriones unguiculatus*). *British Journal of Nutrition* **1**: 1-8

Eltayeb EA, Al-Ansari AS, Roddick JG. 1997. Changes in the steroidal alkaloid solasodine during development of *Solanum nigrum* and *Solanum incanum*. *Phytochemistry* **46**: 489-494

Ensikat HJ, Neinhuis C, Barthlott W. 2000. Direct access to plant epicuticular wax crystals by a new mechanical isolation method. *International Journal of Plant Science* **161**: 143-148

Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, Percy DM, Hajibabaei M, Barrett CH. 2008. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* **3**: e2802

Ford CS, Ayres KL, Toomey N, Haider N, Van Alphen Stahl J, Kelly LJ, Wikström N, Hollingsworth PM, Duff RJ, Hoot SB, Cowan RS, Chase MW, Wilkinson MJ. 2009. Selection of candidate coding DNA barcoding regions for use on land plants. *Botanical Journal for the Linnaean Society* **159**: 1-11

Francia E, Tacconi G, Crosatti C, Barabaschi D, Bulgarelli D, Dall'aglio E, Valè G. 2005. Marker assisted selection in crop plants. *Plant Cell, Tissue and Organ Culture* **82**: 317-342

Ganapathi A, Rao GR. 1986. The crossability and genetic relationship between *Solanum retroflexum* Dun. and *S. nigrum* L. *Cytologia* **57**: 757-762

© University of Pretoria

Gao S-Y, Wang Q-J, Ji Y-B. 2006. Effect of solanine on the membrane potential of mitochondria in $HepG_2$ cells and [Ca^{2+}] in the cells. *World Journal of Gastroenterology* **12**: 3359-3367

Gao L, SU Y-J, Wang T. 2010. Plastid genome sequencing, comparative genomics, and phylogenomics: Current status and prospects. *Journal of Systematics and Evolution* **48**: 77-93

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**: R80

Givan AL. 1992. *Flow Cytometry First Principles.* John Wiley & Sons, Inc, 104-110; 121-127

Haanstra JPW, Wye C, Verbakel H, Meijer-Dekens F, Van den Berg P, Odinot P, Van Heusden AW, Tanksley S, Lindhout P, Peleman J. 1999. An integrated high-density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* x *L. pennellii* F_2 populations. *Theoretical and Applied Genetics* **99**: 254-271

Haselwood EL, Motter GG. 1983. *Handbook of Hawaiian weeds,* 2nd ed., revised and expanded by R.T. Hirano. Harold L. Lyon Arboretum. Honolulu: Univ. Hawaii Pr.

Hawkes JG, Lester RN, Skelding AD. 1979. *The Biology and Taxonomy of the Solanaceae*. Academic Press, 529-548

Healy AJ. 1974. Identification of nightshades (*Atropa* and *Solanum*) in New Zealand. *Proc. 27th N.Z. Weed and Pest Control Conference,* 41-58

Hebert PDN, Gregory TR. 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology* **54**: 852-859

Henderson RJF. 1974. *Solanum nigrum* L. (Solanaceae) and related species in Australia. *Contributions from the Queensland Herbarium* **16**: 1-78

Hillis DM, Bull JJ. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* **42**:182-192

Hillis DM, Dixon MT. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. *The Quarterly Review of Biology* **66**: 411-453

© University of Pretoria

Hilu KW, Liang H. 1997. The *mat*K gene: Sequence variation and application in plant systematics. *American Journal of Botany* **84**: 830-839

Hollingsworth PM. 2008. Progress and outstanding questions. *Heredity* **101**:1-2

Hopkins WG, Hüner NPA. 2004. *Plant Physiology* 3rd Edition. John Wiley & Sons Inc. pp. 493-510

Howard EL, Whittock SP, Jakše J, Carling J, Matthews PD, Probasco G, Henning JA, Darby P, Cerenak A, Javornik B, Kilian A, Koutoulis A. 2011. High-throughput genoptying of hop (*Humulus lupulus* L.) utilizing diversity arrays technology (DArT). *Theoretical and Applied Genetics* **122**: 1265-1280

Hughes CE, Eastwood RJ, Bailey CD. 2006. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions of the Royal Society: Biological Sciences* **361**: 211-225

Huttner E, Wenzl P, Akbari M, Caig V, Carling J, Cayla C, Evers M, Jaccoud D, Peng K, Patarapuwadol S, Uszynski G, Xia L, Yang S, Kilian A. 2005. Diversity Arrays Technology: A novel tool for harnessing the genetic potential of orphan crops. In: Serageldin I, Persley GJ (eds). *Discovery to Delivery: BioVision* Alexandria 2004, Proceedings of the 2004 Conference of The World Biological Forum. CABI Publishing: UK, 145–155.

Jaccoud D, Peng K, Feinstein D, Kilian A. 2001. Diversity Arrays: A solid state technology sequence information independent genotyping. *Nucleic Acids Research* **29**: e25

Jacoby A, Labuschagne MT, Viljoen CD. 2003. Genetic relationships between Southern African *Solanum retroflexum* Dun. and other related species measured by morphological and DNA markers. *Euphytica* **132**: 109-113

Jacoby A, Labuschagne MT. 2006. Hybridization studies of five species of the *Solanum nigrum* complex found in South Africa and two cocktail tomato cultivars. *Euphytica* **149**: 303-307

James KE, Schneider H, Ansell SW, Evers M, Robba L, Uszynski G, Pedersen N, Newton AE, Russel SJ, Vogel JC, Kilian A. 2008. Diversity Arrays Technology (DArT) for Pan-genomic evolutionary studies of non-model organisms. *PLoS One* **3**: e1682

Jansen van Rensburg WS, Van Averbeke W, Slabbert R, Faber M, Van Jaarsveld P, van Heerden I, Wenhold F, Oelofse A. 2007. African leafy vegetables in South Africa. *SA Water* **33**: 317-326

© University of Pretoria

Jardine M, Edmonds JM. 1974. The use of numerical methods to describe population differentiation. *New Phytology* **73**: 1259-1277

Jiménez JP, Brenes A, Fajardo D, Salas A, Spooner DM. 2008. The uses and limits of AFLP data in the taxonomy of polyploid wild potato species in *Solanum* series *Conicibaccata*. *Conservation Genetics* **9**: 381-387

Johnson LA, Soltis DE. 1994. *Mat*K DNA sequences and phylogenetic reconstruction in Saxifragaceae *s. str.* *Systematic Botany* **19**: 143-156

Juma KA. 2006. *Response of* Solanum retroflexum *Dun. to nitrogen, phosphorous and potassium in pots.* MSc Thesis. Tshwane University of Technology

Kanan GJM, Al-Najar RAK. 2009. *In vitro* and *in vivo* activity of selected plant crude extracts and fractions against *Penicillium italicum*. *Journal of Plant Protection Research* **49**: 341-352

Kannabiran K, Mohankumar T, Gunaseker V. 2009. Evaluation of antimicrobial activity of saponin isolated from *Solanum xanthocarpum* and *Centella asiatica*. *International Journal of Natural and Engineering Sciences* **3**: 22-25

Kardolus R, Van Eck HJ, Van Den Berg RG. 1998. The potential of AFLPs in biosystematics: a first application in Solanum taxonomy (Solanaceae). *Plant Systematics and Evolution* **210**: 87-103

Katambo MLM. 2007. *A Systematic study of African* Solanum *L. section* Solanum *(Solanaceae).* PhD Thesis. Radboud University of Nijmegen.

Knoop V. 2010. Looking for sense in the nonsense: a short review of non-coding organellar DNA elucidating the phylogeny of bryophytes. *Tropical Bryology* **31**: 51-60

Koduru S, Grierson DS, Afolayan AJ. 2006. The foliar micro-morphology of *Solanum aculeastrum*, a medicinal plant of South Africa. *Pakistan Journal of Biological Sciences* **9**: 1056-1060

Kress WJ, Erickson DL. 2007. A two-locus global DNA barcode for land plants: the coding *rbc*L gene complements the non-coding *trn*H-*psb*A spacer region. *PLoS ONE* **2**: e508

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences USA* **102**: 8369-8374

Kron P, Suda J, Husband BC. 2007. Applications of flow cytometry to evolutionary and population biology. *Annual Review of Ecology, Evolution and Systematics* **38**: 847-876

© University of Pretoria

Kumar A, Pushpangadan P. 2005. Molecular systematic study of variants of *Solanum nigrum* L. in India. *Cytologia* **70**: 33-38

Lahaye R, Van Der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences* **105**: 2923-2928

Lebecka R. 2009. Inheritance of resistance in *Solanum nigrum* to *Phytophthora infestans*. *European Journal of Plant Pathology* **124**: 345-348

Lee S-J, Lim K-T. 2006. 150Kd glycoprotein isolated from *Solanum nigrum* Linne stimulates caspase-3 activation and reduces inducible nitric oxide production in HCT-116 cells. *Toxicology in Vitro* **20**: 1088-1097

Lee H-L, Yi D-K, Kim J-S, Kim K-J. 2007. Development of plant DNA barcoding markers from the variable non-coding regions of chloroplast genome. *Abstract presented at the Second International Barcode of Life Conference*. Academia Sinica, Taipei, Taiwan. September 18–20,

2007. http://www.bolinfonet.org/conferences/assets/files/conference_abstract_book.pdf.

Liu Y, Yan H-F, Cao T, Ge X-J. 2010. Evaluation of 10 plant barcodes in Bryophyta (mosses). *Journal of Systematics and Evolution* **48**: 36-46

Loganayaki N, Siddhuraju P, Manian S. 2010. Antioxidant activity of two traditional Indian vegetables: *Solanum nigrum* L. and *Solanum torvum* L. *Food Science and Biotechnology* **19**: 121-127

Luthria DL, Mukhopadhyay S, Krizek DT. 2006. Content of total phenolics and phenolic acids in tomato (*Lycopersicon esculentum* Mill.) fruits as influenced by cultivar and solar UV radiation. *Journal of Food Composition and Analysis* **19**: 771–777

Mace ES, Lester RN, Gebhardt CG. 1999. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). *Theoretical and Applied Genetics* **99**: 626-633

Mace ES, Rami J-F, Bouchet S, Klein PE, Klein RR, Kilian A, Wenzl P, Xia L, Halloran K, Jordan DR. 2009. A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. *BMC Plant Biology* **9**:13

© University of Pretoria

MacLaughlin JL, Chang C, Smith DL. 1993. *Simple bench-top assays (brine shrimp and potato discs) for the discovery of plant antitumor compounds.* In: Human medicinal agents from plants, Chapter 9, edited by AD. Kinghorn and MF. Balandrin. American Chemical Society, Washington, D.C, 112-137

Mader SS. 2001. *Biology* 6th Edition. McGraw-Hill Publishing, Chapter 4

Maier RM, Neckermann K, Igloi GL, Kössel H. 1995. Complete sequence of the maize chloroplast genome: Gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *Journal of Molecular Biology* **251**: 614-628

Manoko MLK, Van Den Berg RG, Feron RMC, Van Der Weerden GM, Mariani C. 2007. AFLP markers support separation of *Solanum nodiflorum* from *Solanum americanum sensu stricto* (Solanaceae). *Plant Systematic Evolution* **267**: 1-11

Manoko MLK, Van Den Berg RG, Feron RMC, Van Der Weerden GM, Mariani C. 2008. Genetic diversity of the African hexaploid *Solanum scabrum* Mill. and *Solanum nigrum* L. (Solanaceae). *Genetic Resources in Crop Evolution* **55**: 409-418

Marsan PA, Castiglioni P, Fusari F, Kuiper M, Motto M. 1998. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. *Theoretical and Applied Genetics* **96**: 219-227

Mayr, E. 1942. *Systematics and the Origin of Species.* Columbia University Press, New York, 120

Mudzwiri M. 2007. *Evaluation of traditional South African leafy plants for their safety in human consumption.* MSc Thesis, Durban University of Technology, Durban, South Africa

Müller KF, Borsch T, Hilu KW. 2006. Phylogenetic utility of rapidly evolving DNA at high taxonomical levels: Contrasting *mat*K, *trn*T-F, and *rbc*L in basal angiosperms. *Molecular Phylogenetics and Evolution* **41**: 99-117

Muto M, Mulabagal V, Huang H-C, Takahashi H, Tsay H-S, Huang J-W. 2006. Toxicity of black nightshade (*Solanum nigrum*) extracts on *Alternaria brassicicola*, causal agent of black leaf spot of Chinese cabbage (*Brassica pekinensis*). *Journal of Phytopathology* **154**: 45-50

Mwafusi CN. 1992. *Effects of propagation method and deflowering on vegetative growth, leaf yield, phenolic and glycoalkaloid contents of three black nightshade selections used as vegetable in Kenya.* MSc Thesis, University of Nairobi

© University of Pretoria

Mwai GN, Onyango JC, Abukusta-Onyango M. 2007. Taxonomic identification and characterization of African nightshades (*Solanum* L. Section *Solanum*). *African Journal of Food, Agriculture, Nutrition and Development* **7**

Nisha P, Abdul Nazar P, Jayamurthy P. 2009. A comparative study on antioxidant activities of different varieties of *Solanum melongena*. *Food and Chemical Toxicology* **47**: 2640-2644

Ogg AG, Rogers BS, Schilling EE. 1981. Characterization of black nightshade (*Solanum nigrum*) and related species in the United States. *Weed Science*. **29**: 27-32

Olet EA. 2004. *Taxonomy of* Solanum *L. section* Solanum *in Uganda.* PhD Thesis, Agricultural University of Norway.

Olmstead RG, Sweere JA. 1994. Combining data in phylogenetic systematics: An empirical approach using three molecular data sets in the Solanaceae. *Systematic Biology* **43**: 467-481

Onyang'o MA. 1993. *Effect of plant density and harvesting frequency on the yield and vegetable quality of four variants of black nightshade (*Solanum nigrum *L.).* MSc. Thesis, University of Nairobi, Kenya

Oniang'o RK, Mutuku JM, Malaba SJ. 2003. Contemporary African food habits and their nutritional and health implications. *Asia Pacific Journal of Clinical Nutrition* **12**: 231-236

Palmer JD. 1987. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *The American Naturalist* **130**: S6-S29

Palomino G, Hernándes LT, De La Cruz Torres E. 2008. Nuclear genome size and chromosome analysis in *Chenopodium quinoa* and *C. berlandieri* subsp. *nuttalliae*. *Euphytica* **164**: 221-230

Perez RM, Perez JA, Garcia LM, Sossa H. 1998. Neuropharmacological activity of *Solanum nigrum* fruit. *Journal of Ethnopharmacology* **62**: 43-48

Peter AJ, Shanower TG. 1998. Plant glandular trichomes, chemical factories with many potential uses. *Resonance* **3**: 41-45

Poczai P, Taller J, Szabó I. 2008. Analysis of phylogenetic relationships in the genus *Solanum* (Solanaceae) as revealed by RAPD markers. *Plant Systematic Evolution* **275**: 59-67

Poczai P, Hyvönen J. 2010. Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reports* **37**: 1897-1912

© University of Pretoria

Poczai P, Mátyás K, Taller J, Szabó I. 2010. Study of the origin of the rarely cultivated edible *Solanum* species: morphological and molecular data. *Biologia Plantarum* **54**: 543-546

Rajendiran A, Natarajan E, Subramanian P. 2008. Control of *Aeromonas hydrophila* infection in spotted snakehead, *Channa punctatus*, by *Solanum nigrum* L., a medicinal plant. *Journal of the World Aquaculture Society* **39**: 375-383

Rao GR, Kumar A. 1981. Cytology of *Solanum nigrum* L., *S. retroflexum* Dun, and their hybrids. *Proceedings in Indian Academy of Science* **90**: 227-230

Reboud X, Zeyl C. 1994. Organelle inheritance in plants. *Heredity* **72**: 132-140

Richardson JE, Fay MF, Cronk QCB, Bowman D, Chase MW. 2000. A phylogenetic analysis of Rhamnaceae using *rbc*L and *trn*L-F plastid DNA sequences. *American Journal of Botany* **87**: 1309-1324

Ridley M. 2004. Species concepts and Intraspecific Variation In: *Evolution*, 3rd Edition. Wiley-Blackwell Publishing, USA, 347-381

Roberts AV, Gladis TH, Brumme H. 2009. DNA amounts of roses (*Rosa* L.) and their use in attributing ploidy levels. *Plant Cell Reports* **28**: 61-71

Rodriguez E, Healey PL, Mehta I. 1984. *Biology and Chemistry of Plant Trichomes.* Plenum Press, New York, 1-15

Ruel MT, Minot N, Smith L. 2005. Patterns and determinants of fruit and vegetable consumption in sub-Saharan Africa. Background paper for the Joint FAO/WHO Workshop on Fruit and Vegetables for Health, 1-3 September 2004, Kobe, Japan

Schilling EE. 1981. Systematics of *Solanum* sect *Solanum* (Solanaceae) in North America. *Systematic Botany* **6**: 172-185

Schilling EE, Andersen RN. 1990. The black nightshades (*Solanum* section *Solanum*) of the Indian subcontinent. *Botanical Journal of the Linnaean Society* **102**: 253-259

Schippers R. 1998. Notes on Huckleberry, *Solanum scabrum*, and related black nightshade species. *Natural Resources Institute University of Greenwich* pp. 1-17

Schippers R. 2002. *African indigenous vegetables: an overview of the cultivated species*. Rev. ed. [CD-ROM]. Chatham, UK: Natural Resources Institute/ACP-EU Technical Centre for Agricultural and Rural Cooperation.

© University of Pretoria

Semagn K, Bjørnstad Å, Skinnes H, Marøy AG, Tarkegne Y, William M. 2006. Distribution of DArT, AFLP, and SSR markers in a genetic linkage map of a doubled-haploid hexaploid wheat population. *Genome* **49**: 545-555

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayshida N, Matsubayasha T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa A, Kata A, Tohdoh N, Shimada H, Sugiura M. 1986. The complete nucleotide sequence of the tobacco chloroplast genome. *Centre for Gene Research, Nagoya University, Chikusa, Nagoya 464, Japan.*

Shneyer VS. 2009. DNA Barcoding is a new approach in comparative genomics of plants. *Russian Journal of Genetics* **45**: 1268-1278

Singh G, Kawatra A, Sehgal S. 2001. Nutritional composition of selected green leafy vegetables, herbs and carrots. *Plant Foods for Human Nutrition* **56**: 359-364

Small RL, Ryburn JA, Cronn RC, Seelanan T, Wendel JF. 1998. The tortoise and the hare: Choosing between non-coding plastome and nuclear *ADH* sequences for phylogenetic reconstruction in a recently diverged plant group. *American Journal of Botany* **85**: 1301-1315

Smith FI, Eyzaguirre P. 2007. African leafy vegetables: Their role in the world health organization's global fruit and vegetable initiative. *African Journal of Food, Agriculture, Nutrition and Development* **7**: 3

Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annual Reviews in Plant Biology* **60**: 561-588

Soltis PS, Soltis DE, Doyle JJ. 1992. *Molecular systematics in plants*. Springer, 36-42

Soriano JM, Joshi SG, Van Kaauwen M, Noordijk Y, Groenwold R, Henken B, Van de Weg WE, Schouten HJ. 2009. Identification and mapping of the novel apple scab resistance gene Vd3. *Tree Genetics and Genomes* **5**: 475-482

Souza MM, Pereira TNS, Vieira MLC. 2008. Cytogenetic studies in some species of *Passiflora* L. (Passifloraceae): A review emphasizing Brazilian species. *Brazilian Archives of Biology and Technology* **51**: 247-258

Spooner DM. 2009. DNA barcoding will frequently fail in complicated groups: An example in wild potatoes. *American Journal of Botany* **96**: 1177-1189

© University of Pretoria

Steele KP, Vilgalys R. 1994. Phylogenetic analysis of Polemoniaceae using nucleotide sequences of the plastid gene *mat*K. *Systematic Botany* **19**: 126-142

Steele PR, Pires JC. 2011. Biodiversity assessment: State-of-the-art techniques in phylogenomics and species identification. *American Journal of Botany* **98**: 415-425

Supriya A, Senthilvel S, Nepolean T, Eshwar K, Rajaram V, Shaw R, Hash CT, Kilian A, Yadav RC, Narasu ML. 2011. Development of a molecular linkage map in pearl millet intergrating DArT and SSR markers. *Theoretical and Applied Genetics*

Szczerbakowa A, Maciejewska U, Zimnoch-Guzowska E, Wielgat B. 2003. Somatic hybrids *Solanum nigrum* (+) *S. tuberosum*: morphological assessment and verification of hybridity. *Plant Cell Reports* **21**: 577-584

Tangkanakul P, Auttaviboonkul P, Niyomwit B, Lowvitoon N, Charoenthamawat P, Trakoontivakorn G. 2009. Antioxidant capacity, total phenolic content and nutritional composition of Asian foods after thermal processing. *International Food Research Journal* **16**: 571-580

Taskeen A, Naeem I. 2011. Identification and quantitation of flavonoid in fresh and processed potatoes. *Electronic Journal for Environmental, Agricultural and Food Chemistry* **10**: 2206-2215

Thunberg CP. 1794. *Prodromus plantarum Capensium: quas in promontorio Bonæ Spei Africes, annis* 1772-1775. Upsaliæ: J Edman, 1794-1800

UNAIDS. 2007. Sub-Saharan Africa. http://www.unaids.org/en/CountryResponses/Regions/SubSaharanAfrica.asp

Uusiku NP, Oelofse A, Duodu KG, Bester MJ, Faber M. 2010. Nutrtitional value of indigenous vegetables of sub-Saharan Africa and their potential contribution to human health: A review. *Journal of Food Composition and Analysis* **23**: 499-509

Van Averbeke W, Juma KA, Tshikalange TE. 2007 a. Yield response of African leafy vegetables to nitrogen, phosphorous and potassium: The case of *Brassica rapa* L. subsp. *chinensis* and *Solanum retroflexum* Dun. *Water SA* **33**: 355-362

Van Averbeke W, Tshikalange TE, Juma KA. 2007 b. The commodity systems of *Brassica rapa* L. subsp. *chinensis* and *Solanum retroflexum* Dun. in Vhembe, Limpopo province, South Africa. *Water SA* **33**: 349-354

© University of Pretoria

Van Biljon A, Labuschagne M, Koen E. 2010. Microsatellite-based assessment of five Solanum nigrum complex species and their progeny. *Acta Agriculturae Scandinavica Section B – Plant Soil Science* 1-6, iFirst Article

Venkateswarlu J, Rao MK. 1972. Breeding system, crossability relationships and isolating mechanisms in the *Solanum nigrum* complex. *Cytologia* **37**: 317-326

Vogg G, Fischer S, Leide J, Emmanuel E, Jetter R, Levy AA, Riederer M. 2004. Tomato fruit cuticular waxes and their effects on transpiration barrier properties: functional characterization of a mutant deficient in a very-long-chain fatty acid b-ketoacyl-CoA synthase. *Journal of Experimental Botany* **55**: 1401-1410

Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407-4414

Wannang NN, Anuka JA, Kwanashie HO, Gyang SS, Auta A. 2008. Anti-seizure activity of the aqueous leaf extract of *Solanum nigrum* Linn (Solanaceae) in experimental animals. *African Health Sciences* **8**: 74-79

Watt JM, Breyer-Brandwijk MG. 1962. *Medicinal and poisonous plants of Southern and Eastern Africa*. 2nd Edition. E. S. Livingstone Ltd., 998-1000

Webb DA. 1985. What are the criteria for presuming native status? *Watsonia* **15**: 231-236

Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A. 2004. Diversity Arrays Technology (DArT) for whole genome profiling of barley. *Proceedings of the National Academy of Sciences* **101**: 9915-9920

Wiley EO. 1978. The evolutionary species concept reconsidered. *Systematic Zoology* **27**: 17-26

Wilson , EO. 2003. The encyclopedia of life. *Trends in Ecology &Evolution* **18**: 77 – 80 .

Wittenberg AHJ, Van der Lee T, Cayla C, Kilian A, Visser RGF, Schouten HJ. 2005. Validation of the high throughput marker technology DArT using the model plant *Arabidopsis thaliana*. *Molecular Genetics and Genomics* **274**: 30-39

WHO. 2009. World Health Statistics, http://www.who.int/whosis/whostat/2009/en/index.html

© University of Pretoria

Wolfe KH, Li W-H, Sharp MP. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proceedings of the National Academy of Sciences* **84**: 9054-9058

Yang S, Pang W, Ash G, Harper J, Carling J, Wenzl P, Huttner E, Zong X, Kilian A. 2006. Low level of genetic diversity in cultivated Pigeonpea compared to its wild relatives is revealed by diversity arrays technology. *Theoretical and Applied Genetics* **113**: 585-595

Xia L, Peng K, Yang S, Wenzl P, De Vicente MC, Fregene M, Kilian A. 2005. DArT for high-throughput genotyping of Cassava (*Manihot esculenta*) and its wild relatives. *Theoretical and Applied Genetics* **110**: 1092-1098

© University of Pretoria

Appendix A

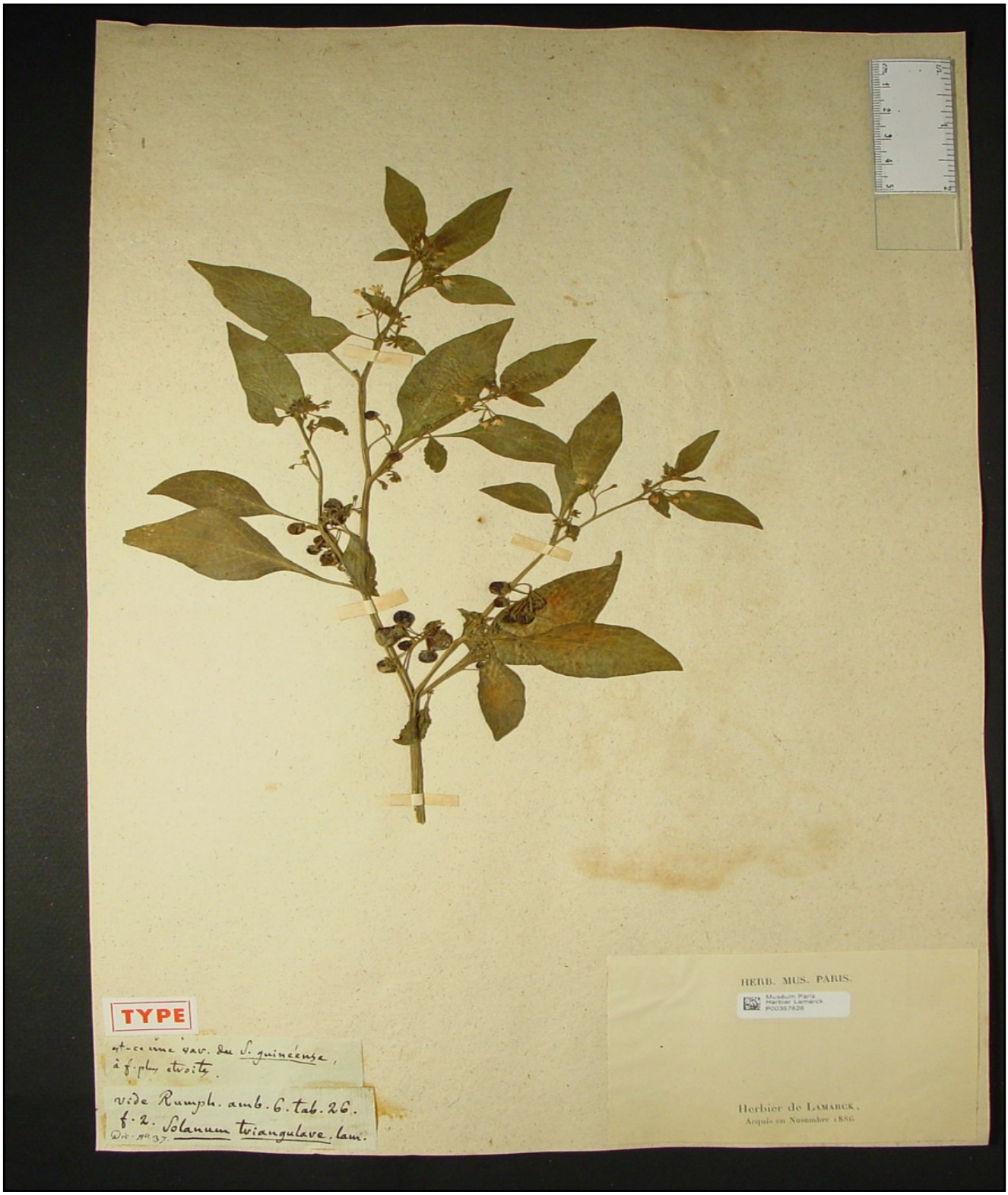
Herbarium Type Specimens

© University of Pretoria

Solanum americanum

© University of Pretoria

Solanum chenopodioides



© University of Pretoria

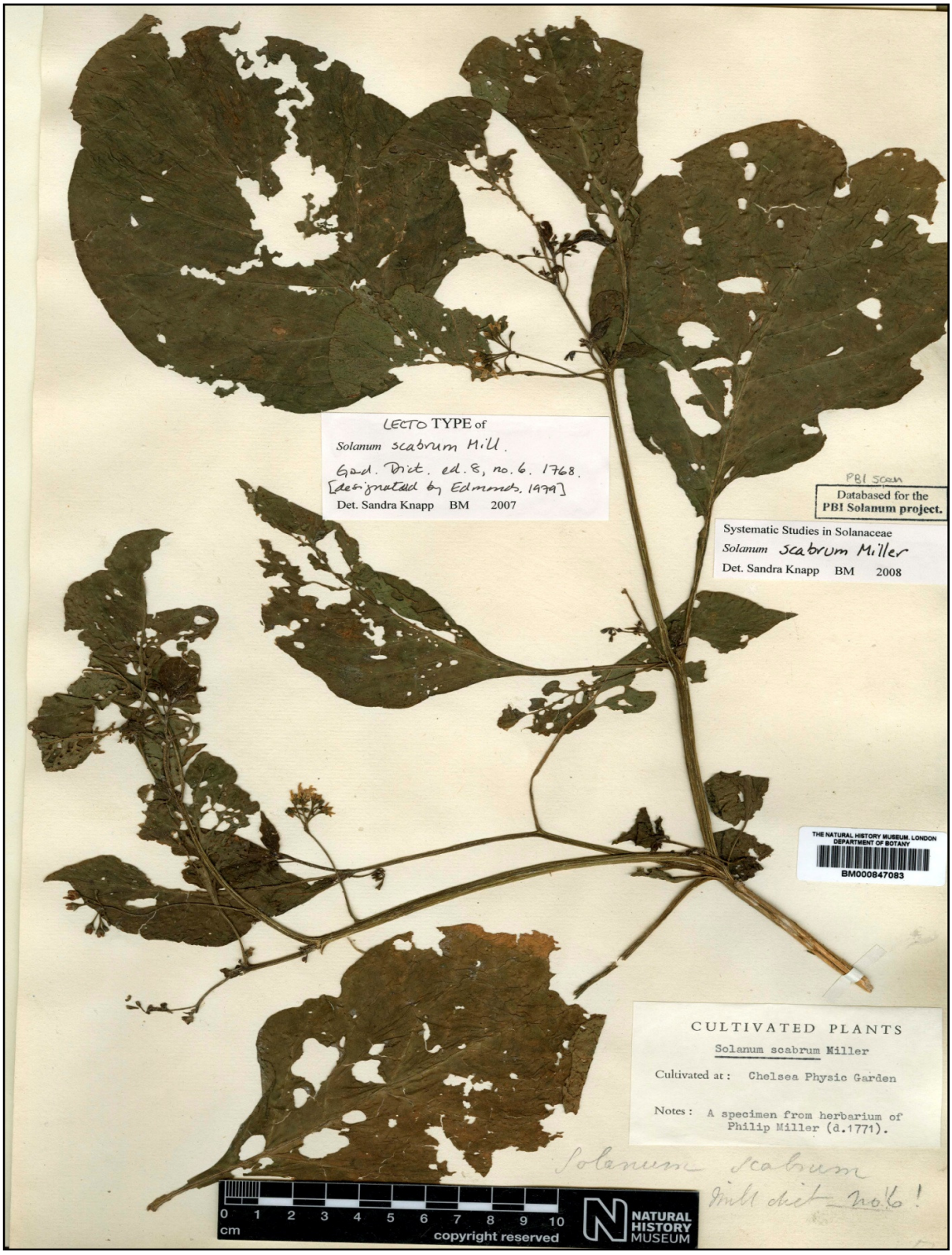
S. nigrum

© University of Pretoria

Solanum retroflexum

© University of Pretoria

Solanum scabrum



© University of Pretoria

Appendix B

Buffers, Solutions and Culture Media

© University of Pretoria

Nuclei Isolation Buffer:

Nuclei Isolation buffer consists of three different buffers; LB01, Galbraith's buffer and Tris.$MgCl_2$ buffer.

LB01: 15mM Tris, 2mM Na_2EDTA, 0.5mM spermine tetrahydrochloride, 80mM KCl, 20mM NaCl, 0.1% (vol/vol) Triton X-100. pH adjusted to 7.5 with 1 M NaOH. Pass through a 0.22um filter; add β-mercapto–ethanol to 15mM. Store at -20˚C in 10ml aliquots. Do no refreeze after thawing.

Galbraith's buffer: 45mM $MgCl_2$, 20mM MOPS, 30mM sodium citrate, 0.1% (vol/vol) Triton X-100. Adjust pH to 7 with 1 M NaOH and filter through a 0.22um filter. Store at -20˚C in 10ml aliquots. Do no refreeze after thawing.

Tris.$MgCl_2$ buffer: 200mM Tris, 4mM $MgCl_2 \cdot 6H_2O$, 0.5% (vol/vol) Triton X-100. Adjust pH to 7.5 with 1 N HCl and filter through a 0.22um filter. Store at 4˚C.

Bacterial JM109 Competent Cells:

- Grow JM109 cells in 10ml LB broth overnight at 37˚C
- Inoculate 1ml overnight culture in 100ml LB broth
- Grow at 37˚C for 3-4 hours until the optical density is 0.94 (600nm)
- Place 50ml in two grinder tubes, and spin in centrifuge for 5 min at 5000rpm (4˚C)
- Discard supernatant and re-suspend pellet in ice cold 10ml 80mM $CaCl_2$, 50mM $MgCl_2$
- Place on ice for 10 min
- Spin for 5 min at 5000rpm (4˚C), and discard supernatant
- Resuspend pellet in ice cold 10ml 80mM $CaCl_2$, 50mM $MgCl_2$
- Place on ice for 10 min
- Spin for 5 min at 5000rpm (4˚C), and discard supernatant
- Resuspend in ice cold 5ml 100mM $CaCl_2$
- Add 1.5ml 50% glycerol, and aliquot 200ul into cold 1.5ml eppendorf tubes
- Snap freeze in liquid nitrogen, and store at -80C.

© University of Pretoria

Luria-Bertani Medium (LB), (Promega, Madison, USA)

LB Medium: 10g Bacto®-tryptone, 5g Bacto®-yeast extract, 5g NaCl. Adjust pH to 7.0 with NaOH and add ddH_2O until 1 L.

IPTG stock solution (0.1M): 1.2g Isopropyl β-D-1-thiogalactopyranoside, add water to 50ml final volume. Filter-sterilize and store at 4°C.

X-Gal (2ml): 100mg 5-bromo-4-chloro-3-indolyl-β-D-galactoside, Dissolve in 2ml N,N´-dimethylformamide. Cover with aluminium foil and store at –20˚C.

LB plates with ampicillin: Add 15g bacteriological agar to 1 litre of LB medium and autoclave. Allow the medium to cool to 50˚C before adding ampicillin to a final concentration of 100µg/ml. Pour 30–35ml of medium into 90mm petri dishes. Let the agar harden. Store at 4˚C for up to 1 month or at room temperature for up to 1 week.

LB plates with ampicillin/IPTG/XGal: Make one litre of LB medium and add ampicillin to a final volume of 100ug/ml. The LB medium is supplemented with 0.5mM IPTG and 80µg/ml X-Gal, and 30-35ml is poured into 90mm petri dishes. Let the agar harden.

Cell Culture Freezing Medium (1L)

1 x LB Medium, 4.4% glycerol, 8.21 g/L K_2HPO_4, 1.80 g/L KH_2PO_4, 0.50 g/L Na_3-citrate, 0.10 g/L $MgSO_4$ x $7H_2O$, 0.90 g/L $(NH_4)_2SO_4$, 100 mg/L ampicillin and 100 m/L kanamycin. Autoclave before glycerol and antibiotics are added.

© University of Pretoria

Appendix C

Statistical Analysis of Morphological data

© University of Pretoria

R Computational Software Script:

```
setwd("\\Erika_data_analise")

getwd()

filename = "For_ANOVA_Shortlist.txt"

d <- read.table(filename, header=T)


# y-axis names in the correct order:

yaxis = c("","Plant Height (cm)","Horizontal (mm)","Vertical (mm)","Area (mm2)","Petiole length
(mm)","Indumentum size (mm)","Ratio","Ratio","Corolla length","Corolla width","Ratio","Ratio","Ratio","Size
(mm)","Amount","Size (mm)","Ratio","Ratio","Size (mm)","Amount","Ratio","Ratio","Size
(mm)","Amount","Ratio","Ratio","Ratio","Ratio")


for (i in 2:29) {

cat(i," ")

X11()

boxplot(d[,i]~d[,1],main=names(d[i]),xlab = "Accessions",ylab = yaxis[i], col =
c("green","orange","deeppink","deeppink","green","deeppink","deeppink","green","blue","green","blue","green","d
eeppink","blue","green","blue","deeppink","deeppink","blue","blue","blue","blue","blue","blue","blue","blue","blue
","blue","yellow"),

las=3)

extension = ".jpg"

f = paste(names(d[i]),extension,sep="")

dev.print(device=jpeg,file=f,width=480,height=480)

}

save.image("C:\\Erika_data_analise\\Data analise")


#Shapiro Wilk normality test

area <- data.frame(d[,3],d[,1])

a <- unstack(area)

names(a)
```


© University of Pretoria

```
shapiro.test(a$AME) #p<0.05, conclude reject H0, thus assume normality


for (i in 1:29){
        s = shapiro.test(a[,i])
        print(s)}


#Kruscal Wallis Non-parametric test
kruskal.test(d[,3]~d[,1])
#p-val < 0.05, reject H0, thus one or more of the medians are different


# Non-parametric post-hoc tests
install.packages("npmc")
library(npmc)
area <- data.frame(d[,3],d[,1])
a <- area
names(a) = c("var","class")
npmc(a)
```

P-values were calculated between accessions for most informative traits.

P-values >0.05 = Pink = accept H0; Accessions are the same

P-values <0.05 = Blue = reject H0; Accessions differ

© University of Pretoria

Plant height

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0
BUR		0	0.99	0	0	0.99	0	0	0.99	0	0.78	0	0	0	0	0	0	0	0	0	0	0	0	0.78	0	0	0	0	0
CHEN			0	1	0.95	1	0.78	0	1	0	1	0	0.78	0	0	0	0.78	0	0.26	0	0	0	0	0.99	0	0	0	0.78	0
LS1				0	0.78	0.99	0.78	0	0.99	0	1	0	0.78	0	0	0	0.78	0	0	0	0	0	0	1	0	0	0	0.26	0
LS2					0	0	1	0	0.78	0	0.78	0	1	0	0	0.26	0.99	0	0.95	0	0	0	0	1	0	0	0	0.99	0
LS3						0	0.26	0	1	0	0.99	0	0.26	0	0	0	0	0	0	0	0	0	0	0.78	0	0	0	0	0
LS4							0	0.78	0.78	0	0.78	0.95	1	0	0	0.99	0.99	0	0.99	0	0	0.78	0	1	0	0.78	0	0.99	0.26
LS5								0	0	0	0	0.99	0.95	0	0	0.99	1	0.62	1	0	0	0.99	0.99	0.78	0	0.99	0	1	1
LS6									0	0	0.99	0	0.78	0	0	0	0.26	0	0	0	0	0	0	0.78	0	0	0	0	0
LS7										0	0	0	0	0	0	0	0	0	0.01	0	0	0	0	0	0	0	0	0	0
LS8											0	0	0.78	0	0	0	0.78	0	0	0	0	0	0	0.99	0	0	0	0.26	0
LS9												0	0.99	0	0	1	1	0	1	0	0	1	0	0.78	0.01	0.99	0	1	0.95
LS10													0	0	0	1	0.99	0	0.99	0	0	0.99	0	0.99	0	0.78	0.01	0.99	0.78
LS11														0	0	0	0	0.99	0	0.99	0	0	0	0	0	0.95	0	0.26	0
LS12															0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LS13																0	1	0	1	0	0	0.99	0	0.95	0	0.99	0	1	0.95
LS14																	0	0.78	1	0	0	1	0.99	0.99	0	0.99	0	0.99	0.99
LS15																		0	0.99	1	0	0	0.95	0	0	0.99	0	0.99	0.95
PC1																			0	0	0	1	1	0.99	0	0.99	0	1	1
PC2																				0	0	0	0	0	0	0.99	0	0.78	0.26
PC3																					0	0	0	0	1	0	1	0	0
PC4																						0	0	0.78	0	0.99	0	1	0.99
PC5																							0	0	0	1	0	1	1
PC6																								0	0	0.26	0	0.99	0
PC7																									0	0	1	0	0
PC8																										0	0	0.99	1
PC9																											0	0	0
PC10																												0	1
SCA																													0

© University of Pretoria

Leaf area

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0	0.005	0.005	0.98	1	0.79	0.06	1	0.02	0.02	0.005	1	0.24	1	0.07	0.01	0.32	0.92	1.00	1	1	0.11	0.01	0.36	0.25	0.01	1	0.09	0.01
BUR		0	0.79	1	0.01	0.97	0.96	0.46	1	0.01	1	0.01	0.79	0.01	0.01	1	1	0.25	1	0.98	1	1.00	0.06	0.11	1	1	0.13	0.94	0.01
CHEN			0	0.65	0.01	0.1	0.51	0.04	0.7	0.01	0.94	0.01	0.09	0.01	0.01	0.92	0.46	0.04	1.00	0.15	0.46	0.18	1	0.02	0.56	0.80	0.02	0.09	0.01
LS1				0	1	1	1	0.83	1	0.01	0.99	1	1	0.99	0.02	1	1	1	1	1	1	1	0.84	1	1	1	1	1	0.01
LS2					0	1	0.83	1	0.28	0.01	0.01	1	0.99	1	0.01	0.01	0.75	1	1	1	1	0.93	0.01	1.00	0.80	0.08	1	0.71	0.01
LS3						0	1	0.97	0.99	0.01	0.55	1	1	0.99	0.01	0.79	1	1	1	1	1	1	0.08	1	1	0.96	1	1	0.01
LS4							0	0.94	1	0.01	0.94	0.46	1	0.28	0.01	0.87	1	1	1	1.00	1	1	0.28	1	1	1	0.93	1	0.01
LS5								0	0.55	0.18	0.18	1	0.97	1	0.94	0.18	0.61	1	0.84	1	0.98	0.87	0.01	1.00	0.66	0.56	1.00	0.98	0.02
LS6									0	0.01	1	0.06	0.99	0.03	0.01	1	1	0.87	1	0.95	1.00	1	0.93	0.95	1	1	0.29	1	0.01
LS7										0	0.01	0.01	0.01	0.01	0.99	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	1
LS8											0	0.01	0.46	0.01	0.01	1	1	0.13	1	0.75	0.98	0.80	1.00	0.11	1	1	0.08	0.41	0.01
LS9												0	0.89	1	0.01	0.01	0.84	1.00	1	1	1	0.51	0.01	0.9	0.32	0.013	1	0.29	0.01
LS10													0	0.6	0.01	0.65	1	1.00	1	1	1	1	0.06	1	1	0.99	0.99	1	0.01
LS11														0	0.02	0.01	0.66	1.00	1	1	1	0.37	0.01	0.80	0.25	0.01	1	0.11	0.01
LS12															0	0.01	0.01	0.06	0.02	0.16	0.06	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.36
LS13																0	1	0.13	1	0.90	1.00	0.96	0.41	0.03	1	1	0.08	0.41	0.01
LS14																	0	0.99	1	1.00	1	1	0.61	1	1	1	0.99	1	0.01
LS15																		0	1	1	1	1	0.01	1	0.99	0.56	1	1	0.01
PC1																			0	1	1	1	1.00	1	1	1	1	1	0.01
PC2																				0	1	1.00	0.25	1	0.99	0.93	1	1	0.01
PC3																					0	1	0.51	1	1	1	1	1	0.01
PC4																						0	0.09	1	1	1.00	0.97	1	0.01
PC5																							0	0.01	0.61	0.66	0.03	0.04	0.01
PC6																								0	1.00	0.46	1	1	0.01
PC7																									0	1	0.87	1	0.01
PC8																										0	0.15	1.00	0.01
PC9																											0	0.90	0.01
PC10																												0	0.01
SCA																													0

© University of Pretoria

Lamina length x lamina width ratio

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.72	1.00	1.00	0.00	0.00	0.99	1.00	0.46	0.00	0.00	0.86	1.00	0.98	0.01	0.15	1.00	0.10	0.00	0.97	0.02	1.00
BUR		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CHEN			0.00	0.30	0.06	0.93	0.69	0.00	0.00	0.00	0.00	0.01	0.15	0.00	0.00	0.00	1.00	0.86	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS1				0.00	1.00	1.00	1.00	1.00	0.11	0.00	0.00	1.00	1.00	0.05	0.00	0.03	0.82	1.00	0.03	0.00	0.49	0.73	0.14	0.00	0.94	0.84	0.02	0.78	0.00
LS2					0.00	0.99	1.00	0.41	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.53	1.00	0.00	0.00	0.47	0.06	0.00	0.00	0.36	0.00	0.00	0.01	0.00
LS3						0.00	1.00	0.19	0.00	0.00	0.00	0.85	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.23	0.05	0.01	0.00	0.12	0.04	0.00	0.02	0.00
LS4							0.00	0.22	0.00	0.00	0.00	0.96	1.00	0.00	0.00	0.00	0.99	1.00	0.00	0.00	0.20	0.01	0.00	0.00	0.10	0.00	0.00	0.02	0.00
LS5								0.00	0.64	0.05	0.00	1.00	1.00	0.51	0.00	0.04	0.01	0.89	0.17	0.00	0.93	1.00	0.41	0.00	1.00	1.00	0.18	1.00	0.00
LS6									0.00	1.00	0.61	0.07	0.20	1.00	0.96	1.00	0.00	0.01	1.00	0.86	1.00	1.00	1.00	0.37	1.00	0.99	1.00	1.00	0.84
LS7										0.00	1.00	0.00	0.01	1.00	1.00	0.97	0.00	0.00	1.00	1.00	1.00	0.28	0.86	0.98	0.65	0.32	1.00	0.79	1.00
LS8											0.00	0.00	0.00	0.98	1.00	0.06	0.00	0.00	0.89	1.00	0.98	0.00	0.02	1.00	0.10	0.00	0.96	0.02	1.00
LS9												0.00	1.00	0.05	0.00	0.00	0.12	1.00	0.01	0.00	0.65	0.57	0.02	0.00	0.98	0.61	0.01	0.77	0.00
LS10													0.00	0.08	0.00	0.10	0.62	1.00	0.05	0.00	0.54	0.79	0.28	0.00	0.96	0.86	0.03	0.83	0.00
LS11														0.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	0.99	1.00	0.94	0.98	0.98	1.00	1.00	1.00
LS12															0.00	0.69	0.00	0.00	1.00	1.00	1.00	0.04	0.38	1.00	0.25	0.03	1.00	0.20	1.00
LS13																0.00	0.00	0.00	1.00	0.45	1.00	0.98	1.00	0.01	0.94	0.42	1.00	0.95	0.76
LS14																	0.00	1.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS15																		0.00	0.00	0.00	0.16	0.11	0.01	0.00	0.46	0.35	0.00	0.43	0.00
PC1																			0.00	0.99	1.00	0.99	1.00	0.88	0.97	0.89	1.00	0.97	0.93
PC2																				0.00	1.00	0.01	0.18	1.00	0.11	0.00	1.00	0.10	1.00
PC3																					0.00	0.99	1.00	0.96	1.00	1.00	1.00	1.00	0.99
PC4																						0.00	1.00	0.00	1.00	1.00	0.89	1.00	0.07
PC5																							0.00	0.00	1.00	0.99	1.00	1.00	0.38
PC6																								0.00	0.06	0.00	0.92	0.03	1.00
PC7																									0.00	1.00	0.92	1.00	0.15
PC8																										0.00	0.69	1.00	0.01
PC9																											0.00	0.94	0.99
PC10																												0.00	0.05
SCA																													0.00

© University of Pretoria

Lamina length x petiole length ratio

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	1.00	0.00	0.00	1.00	0.04	0.00	1.00	0.07	1.00	1.00	1.00	0.01	1.00	1.00	1.00	0.00	0.00	1.00	1.00	0.84	0.66	0.33	1.00	0.13	0.63	0.85	1.00	1.00
BUR		0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.83	0.00	0.31	1.00	0.27	0.00	0.00	0.00	1.00	0.64	0.00	0.00	0.03	0.00	0.00	0.14	0.00	1.00
CHEN			0.00	1.00	0.00	1.00	0.82	0.00	0.10	0.00	0.00	0.40	1.00	0.00	0.04	0.00	1.00	1.00	0.00	0.00	0.12	0.01	0.02	0.00	0.16	0.01	0.02	0.00	0.00
LS1				0.00	0.02	1.00	0.90	0.00	0.94	0.01	0.15	0.38	1.00	0.00	0.49	0.00	1.00	1.00	0.01	0.00	0.69	0.39	0.83	0.00	0.91	0.60	0.29	0.09	0.00
LS2					0.00	0.05	0.00	1.00	0.30	1.00	1.00	1.00	0.03	1.00	1.00	1.00	0.00	0.03	1.00	0.11	1.00	1.00	0.80	1.00	0.72	0.88	1.00	1.00	0.71
LS3						0.00	0.34	0.00	1.00	0.11	0.40	0.73	1.00	0.01	0.84	0.03	0.97	1.00	0.05	0.00	0.97	0.89	1.00	0.08	1.00	0.94	0.81	0.13	0.00
LS4							0.00	0.00	0.00	0.00	0.00	0.00	0.84	0.00	0.00	0.00	1.00	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS5								0.00	0.01	1.00	0.73	1.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	0.86	0.26	0.03	1.00	0.08	0.12	0.95	0.99	1.00
LS6									0.00	0.35	0.41	0.96	1.00	0.01	1.00	0.11	0.66	0.91	0.12	0.00	1.00	1.00	1.00	0.19	1.00	1.00	1.00	0.47	0.01
LS7										0.00	1.00	1.00	0.07	1.00	1.00	1.00	0.01	0.01	1.00	1.00	0.97	0.95	0.82	1.00	0.30	0.99	0.91	1.00	1.00
LS8											0.00	1.00	0.26	0.98	1.00	0.99	0.04	0.12	1.00	0.08	1.00	1.00	0.98	1.00	0.87	1.00	1.00	1.00	0.66
LS9												0.00	0.55	1.00	1.00	1.00	0.05	0.62	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
LS10													0.00	0.00	0.75	0.01	1.00	1.00	0.03	0.00	0.96	0.88	1.00	0.03	1.00	0.89	0.73	0.07	0.00
LS11														0.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	0.55	0.10	1.00	0.15	0.22	1.00	1.00	1.00
LS12															0.00	1.00	0.15	0.53	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS13																0.00	0.00	0.00	1.00	0.98	1.00	0.79	0.29	1.00	0.53	0.53	1.00	1.00	1.00
LS14																	0.00	0.92	0.00	0.00	0.14	0.20	0.43	0.01	0.46	0.29	0.05	0.01	0.00
LS15																		0.00	0.03	0.00	0.83	0.28	0.56	0.01	0.90	0.49	0.26	0.11	0.00
PC1																			0.00	0.39	1.00	0.98	0.57	1.00	0.45	0.88	1.00	1.00	0.91
PC2																				0.00	0.88	0.02	0.00	0.85	0.02	0.00	0.48	0.36	1.00
PC3																					0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.95
PC4																						0.00	1.00	0.98	1.00	1.00	1.00	1.00	0.19
PC5																							0.00	0.67	1.00	1.00	1.00	0.86	0.08
PC6																								0.00	0.52	0.96	1.00	1.00	0.94
PC7																									0.00	1.00	1.00	0.86	0.07
PC8																										0.00	1.00	0.97	0.27
PC9																											0.00	1.00	0.77
PC10																												0.00	0.78
SCA																													0.00

© University of Pretoria

Indentation size

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	0.00	1.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
BUR		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	1.00	0.00	0.00	0.97	0.00	0.25	0.00	0.00	0.33	1.00	0.02	0.17	1.00	0.00	0.00	0.02	0.00	0.61	0.00
CHEN			0.00	1.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS1				0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS2					0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.95	0.00	0.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.51	0.00
LS3						0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS4							0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS5								0.00	0.13	0.00	0.00	1.00	0.00	0.35	0.00	0.00	0.00	0.00	0.02	0.06	0.24	0.03	0.00	0.28	0.00	0.22	0.11	0.99	0.00
LS6									0.00	0.00	0.76	0.12	0.00	1.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.37	1.00	1.00	1.00	1.00	0.98	0.00
LS7										0.00	0.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS8											0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.99	1.00	0.83	0.98	1.00	0.05	1.00	0.37	0.08	0.71	0.00
LS9												0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.04	0.05	0.90	0.10	0.00	0.69	0.00	0.37	0.68	1.00	0.00
LS10													0.00	0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS11														0.00	0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	1.00	1.00	1.00	0.00
LS12															0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS13																0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.89	0.41	1.00	0.62	0.62	0.83	0.00
LS14																	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
LS15																		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
PC1																			0.00	1.00	1.00	1.00	0.17	1.00	1.00	1.00	1.00	0.94	0.00
PC2																				0.00	1.00	1.00	0.82	1.00	1.00	1.00	1.00	0.91	0.00
PC3																					0.00	1.00	0.08	1.00	1.00	1.00	1.00	1.00	0.00
PC4																						0.00	0.47	1.00	1.00	1.00	1.00	0.97	0.00
PC5																							0.00	0.00	0.47	0.02	0.00	0.64	0.00
PC6																								0.00	0.56	1.00	1.00	1.00	0.00
PC7																									0.00	0.81	0.94	0.78	0.00
PC8																										0.00	1.00	1.00	0.00
PC9																											0.00	1.00	0.00
PC10																												0.00	0.00
SCA																													0.00

© University of Pretoria

Trichome density adaxial

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.89	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.89	0.00	0.00	0.00	0.89	0.00	1.00	0.89	1.00	1.00	0.00	0.00	0.00	1.00
BUR		0.00	0.00	0.00	1.00	0.00	0.00	0.89	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.89	0.00	0.00	0.00	0.89	0.00	1.00	0.89	1.00	1.00	0.00	0.00	0.00	1.00
CHEN			0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS1				0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS2					0.00	0.00	0.00	0.61	1.00	0.61	0.89	0.94	0.00	0.61	0.89	1.00	0.00	0.00	0.61	1.00	0.61	1.00	0.61	1.00	1.00	0.89	0.00	0.89	1.00
LS3						0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.78	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS4							0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS5								0.00	0.61	0.00	0.00	1.00	0.00	0.00	0.00	0.26	0.00	0.00	0.00	0.26	0.00	1.00	1.00	1.00	1.00	0.00	0.94	0.00	0.61
LS6									0.00	0.61	0.89	0.94	0.00	0.61	0.89	1.00	0.00	0.00	0.61	1.00	0.61	1.00	0.61	1.00	1.00	0.89	0.00	0.89	1.00
LS7										0.00	1.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.61	0.94	1.00	0.00	1.00	0.61
LS8											0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.89	1.00	1.00	0.00	1.00	0.89
LS9												0.00	0.00	0.00	0.00	0.61	0.00	0.00	0.00	0.61	0.00	1.00	1.00	1.00	1.00	0.00	0.61	0.00	0.94
LS10													0.00	0.00	0.00	0.00	1.00	0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS11														0.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.61	0.94	1.00	0.00	1.00	0.61
LS12															0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.00	0.89	1.00	1.00	0.00	1.00	0.89
LS13																0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.26	0.95	1.00	1.00	0.00	1.00	1.00
LS14																	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS15																		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PC1																			0.00	1.00	1.00	1.00	0.00	0.61	0.94	1.00	0.00	1.00	0.61
PC2																				0.00	1.00	1.00	0.26	0.95	1.00	1.00	0.00	1.00	1.00
PC3																					0.00	1.00	0.00	0.61	0.94	1.00	0.00	1.00	0.61
PC4																						0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
PC5																							0.00	1.00	0.77	0.00	1.00	0.00	0.61
PC6																								0.00	1.00	0.89	1.00	0.89	1.00
PC7																									0.00	1.00	0.61	1.00	1.00
PC8																										0.00	0.00	1.00	0.89
PC9																											0.00	0.00	0.00
PC10																												0.00	0.89
SCA																													0.00

© University of Pretoria

Trichome density abaxial

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	0.00	0.00	0.00	0.26	0.00	0.00	1.00	1.00	0.94	0.00	1.00	0.00	1.00	0.94	1.00	0.00	0.00	1.00	1.00	0.61	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.26
BUR		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.26	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.77	0.00	0.77
CHEN			0.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS1				0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS2					0.00	0.00	0.00	0.00	0.61	0.00	0.00	1.00	0.00	1.00	0.00	0.61	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	1.00	1.00	0.61	1.00	1.00
LS3						0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.77	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS4							0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS5								0.00	0.89	1.00	0.00	0.89	0.00	0.89	1.00	0.89	0.00	0.00	1.00	1.00	0.89	1.00	0.00	1.00	0.89	1.00	0.00	0.89	0.00
LS6									0.00	0.61	0.00	1.00	0.00	1.00	0.61	1.00	0.00	0.00	1.00	1.00	0.25	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.61
LS7										0.00	0.89	0.61	0.00	0.61	1.00	0.61	0.00	0.00	1.00	0.94	1.00	1.00	0.00	1.00	0.61	1.00	0.00	0.61	0.00
LS8											0.00	0.00	0.00	0.00	0.89	0.00	0.00	0.00	1.00	0.00	1.00	0.89	0.00	0.89	0.00	0.89	0.00	0.00	0.00
LS9												0.00	0.77	1.00	0.61	1.00	0.00	0.00	0.77	1.00	0.26	1.00	0.95	1.00	1.00	0.95	1.00	1.00	1.00
LS10													0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.77	0.00	0.77	1.00	1.00	0.00	0.00	0.77	0.00	0.77
LS11														0.00	0.61	1.00	0.00	0.00	1.00	1.00	0.26	1.00	0.26	1.00	1.00	1.00	1.00	1.00	1.00
LS12															0.00	0.61	0.00	0.00	1.00	0.94	1.00	1.00	0.00	1.00	0.61	1.00	0.00	0.61	0.00
LS13																0.00	0.00	0.00	1.00	1.00	0.26	1.00	0.00	1.00	1.00	1.00	0.00	1.00	0.61
LS14																	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS15																		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PC1																			0.00	1.00	1.00	1.00	0.00	1.00	1.00	1.00	0.26	1.00	0.77
PC2																				0.00	0.61	1.00	0.77	1.00	1.00	1.00	1.00	1.00	1.00
PC3																					0.00	1.00	0.00	1.00	0.26	1.00	0.00	0.26	0.00
PC4																						0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
PC5																							0.00	1.00	0.77	0.00	0.77	0.00	0.77
PC6																								0.00	1.00	1.00	1.00	1.00	1.00
PC7																									0.00	1.00	1.00	1.00	1.00
PC8																										0.00	0.26	1.00	0.77
PC9																											0.00	0.61	1.00
PC10																												0.00	1.00
SCA																													0.00

© University of Pretoria

Corolla length x corolla width ratio

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	1.00	2.59	0.00	1.00	0.00	0.00	1.00	0.12	1.00	1.00	1.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
BUR		0.00	0.00	2.60	1.00	0.00	0.00	1.00	0.16	1.00	1.00	1.00	0.00	1.00	1.00	1.00	0.00	0.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.60
CHEN			0.00	0.16	2.23	1.00	1.00	0.00	0.42	0.01	0.05	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00
LS1				0.00	0.17	0.89	0.46	0.81	1.00	0.79	0.95	1.50	1.00	0.39	0.01	0.10	0.42	0.02	0.00	0.01	0.00	0.00	0.20	0.00	0.40	0.09	0.43	0.00	0.01
LS2					0.00	0.00	0.00	1.00	0.36	1.00	1.00	1.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS3						0.00	1.00	0.02	0.88	0.09	0.09	0.00	1.00	0.01	0.00	0.00	1.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.00
LS4							0.00	0.00	0.76	0.10	0.08	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.00	0.00
LS5								0.00	0.97	1.00	1.00	1.00	0.14	1.00	1.00	1.00	0.01	0.00	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.92
LS6									0.00	0.81	1.00	0.03	0.99	0.57	0.23	0.45	0.66	0.17	0.00	0.32	0.12	0.32	0.72	0.01	0.97	0.77	0.78	0.18	0.04
LS7										0.00	1.00	1.00	0.33	1.00	1.00	1.00	0.01	0.01	0.49	1.00	1.00	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.78
LS8											0.00	0.93	0.38	1.00	1.00	1.00	0.04	0.00	0.23	1.00	1.00	1.00	1.00	0.75	1.00	1.00	1.00	1.00	0.32
LS9												0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	0.99	1.00	1.00	1.00	0.96	1.00	1.00	1.00	0.99
LS10													0.00	0.06	0.00	0.04	0.98	0.84	0.00	0.00	0.00	0.00	0.01	0.00	0.14	0.07	0.02	0.00	0.00
LS11														0.00	1.00	1.00	0.00	0.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.74
LS12															0.00	1.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.84
LS13																0.00	0.00	0.00	0.71	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.56
LS14																	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.02	0.00	0.00	0.00
LS15																		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PC1																			0.00	0.86	0.87	0.76	0.71	1.00	0.45	0.96	1.00	0.69	1.00
PC2																				0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.70
PC3																					0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.58
PC4																						0.00	1.00	0.98	1.00	1.00	1.00	1.00	0.62
PC5																							0.00	0.98	1.00	1.00	1.00	1.00	0.76
PC6																								0.00	0.92	1.00	1.00	0.99	1.00
PC7																									0.00	1.00	1.00	1.00	0.52
PC8																										0.00	1.00	1.00	0.83
PC9																											0.00	1.00	0.88
PC10																												0.00	0.97
SCA																													0.00

© University of Pretoria

Berry size

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	0.00	0.29	0.10	0.80	0.00	0.00	0.99	1.00	0.91	1.00	0.45	0.05	0.67	0.95	1.00	0.01	0.00	0.37	0.07	0.70	0.97	1.00	0.21	0.99	0.78	1.00	1.00	0.00
BUR		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.01	0.11	0.00	0.67	0.18	0.00	0.00	0.00	0.30	0.99	0.89	0.23	0.01	0.67	0.00	0.07	0.00	0.00	0.00
CHEN			0.00	1.00	1.00	0.06	0.60	1.00	0.07	0.00	0.69	0.00	1.00	0.00	0.00	0.53	1.00	0.92	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.01	0.60	0.00
LS1				0.00	1.00	1.00	1.00	0.99	0.03	0.00	0.37	0.00	1.00	0.00	0.00	0.24	1.00	1.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.36	0.00
LS2					0.00	0.00	0.03	1.00	0.32	0.00	0.99	0.00	0.98	0.00	0.00	0.91	0.71	0.10	0.00	0.00	0.00	0.01	0.31	0.00	0.01	0.00	0.10	0.97	0.00
LS3						0.00	1.00	0.34	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS4							0.00	0.42	0.00	0.00	0.01	0.00	0.96	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS5								0.00	0.95	0.02	1.00	0.00	0.99	0.01	0.01	1.00	0.87	0.61	0.00	0.00	0.02	0.06	0.95	0.00	0.04	0.00	0.48	1.00	0.00
LS6									0.00	0.80	1.00	0.13	0.00	0.41	0.87	1.00	0.00	0.00	0.10	0.00	0.56	0.92	1.00	0.03	0.98	0.52	1.00	1.00	0.00
LS7										0.00	1.00	1.00	0.00	1.00	1.00	0.84	0.00	0.00	1.00	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.85	0.00
LS8											0.00	0.99	0.37	0.90	1.00	1.00	0.12	0.02	0.96	0.29	0.96	1.00	1.00	0.77	1.00	1.00	1.00	1.00	0.00
LS9												0.00	0.00	1.00	1.00	0.34	0.00	0.00	1.00	0.99	1.00	1.00	0.99	1.00	1.00	1.00	0.71	0.45	0.00
LS10													0.00	0.00	0.00	0.08	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00
LS11														0.00	1.00	0.36	0.00	0.00	1.00	1.00	1.00	1.00	0.94	1.00	0.99	1.00	0.89	0.53	0.00
LS12															0.00	0.80	0.00	0.00	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.88	0.00
LS13																0.00	0.02	0.00	0.24	0.01	0.51	0.85	1.00	0.07	0.99	0.63	1.00	1.00	0.00
LS14																	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00
LS15																		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
PC1																			0.00	1.00	1.00	1.00	0.96	1.00	0.99	1.00	0.64	0.37	0.00
PC2																				0.00	1.00	0.98	0.29	1.00	0.42	0.93	0.14	0.07	0.00
PC3																					0.00	1.00	0.98	1.00	1.00	1.00	0.94	0.57	0.00
PC4																						0.00	1.00	1.00	1.00	1.00	1.00	0.90	0.00
PC5																							0.00	0.77	1.00	1.00	1.00	1.00	0.00
PC6																								0.00	0.86	1.00	0.39	0.21	0.00
PC7																									0.00	1.00	1.00	0.98	0.00
PC8																										0.00	0.98	0.71	0.00
PC9																											0.00	1.00	0.00
PC10																												0.00	0.00
SCA																													0.00

© University of Pretoria

Amount of seed per berry

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	0.00	0.24	1.00	1.00	1.00	1.00	0.99	0.00	1.00	0.00	0.74	1.00	0.00	1.00	0.00	1.00	1.00	0.01	1.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
BUR		0.00	0.85	0.00	0.00	0.00	0.28	0.00	0.95	0.00	0.34	0.09	0.00	0.97	0.27	0.00	0.42	1.00	1.00	1.00	1.00	0.13	1.00	0.30	0.67	0.94	0.00	0.00	0.00
CHEN			0.00	0.00	0.00	0.06	1.00	0.07	0.61	0.00	0.01	0.34	0.53	0.04	0.91	0.00	0.92	1.00	1.00	1.00	0.83	0.00	1.00	0.00	0.00	0.02	0.00	0.00	0.00
LS1				0.00	1.00	1.00	0.94	1.00	0.00	1.00	0.00	0.92	1.00	0.00	1.00	0.00	1.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS2					0.00	1.00	0.74	1.00	0.00	1.00	0.00	1.00	0.99	0.00	0.99	0.00	0.95	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS3						0.00	0.82	1.00	0.00	1.00	0.00	0.99	0.99	0.00	0.99	0.00	0.91	1.00	0.04	0.95	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00
LS4							0.00	0.64	0.07	0.98	0.00	0.59	1.00	0.00	1.00	0.00	1.00	1.00	0.80	1.00	0.21	0.00	0.66	0.00	0.00	0.00	0.00	0.00	0.00
LS5								0.00	0.00	1.00	0.00	1.00	0.96	0.00	0.98	0.00	0.91	0.96	0.00	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS6									0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.01	0.33	1.00	0.39	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
LS7										0.00	0.00	0.88	1.00	0.00	1.00	0.00	1.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS8											0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.20	1.00	0.37	0.33	1.00	0.81	1.00	0.96	0.99	1.00	1.00	0.00
LS9												0.00	0.66	0.01	0.67	0.00	0.47	0.51	0.04	0.31	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00
LS10													0.00	0.00	1.00	0.00	1.00	1.00	0.22	1.00	0.01	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00
LS11														0.00	0.02	0.99	0.04	0.98	1.00	0.96	1.00	1.00	0.98	1.00	1.00	1.00	0.98	1.00	0.00
LS12															0.00	0.00	1.00	1.00	0.50	1.00	0.04	0.00	0.53	0.00	0.00	0.01	0.00	0.00	0.00
LS13																0.00	0.00	0.75	1.00	0.19	0.07	0.96	0.23	1.00	0.97	0.90	1.00	1.00	0.00
LS14																	0.00	1.00	0.51	1.00	0.20	0.00	0.58	0.01	0.02	0.03	0.00	0.00	0.00
LS15																		0.00	0.88	1.00	1.00	1.00	1.00	0.93	0.99	1.00	0.09	0.83	0.00
PC1																			0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	1.00	0.00
PC2																				0.00	1.00	0.60	1.00	0.79	0.98	0.93	0.03	0.36	0.00
PC3																					0.00	0.96	1.00	0.98	1.00	1.00	0.09	0.35	0.00
PC4																						0.00	0.64	1.00	1.00	1.00	0.70	1.00	0.00
PC5																							0.00	0.68	0.87	1.00	0.04	0.27	0.00
PC6																								0.00	1.00	1.00	0.73	1.00	0.00
PC7																									0.00	1.00	0.43	1.00	0.00
PC8																										0.00	0.34	0.97	0.00
PC9																											0.00	1.00	0.00
PC10																												0.00	0.00
SCA																													0.00

© University of Pretoria

Seed size

	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	SCA
AME	0.00	0.96	0.06	0.29	0.00	0.01	0.00	0.04	0.99	1.00	0.01	0.86	0.17	0.24	1.00	0.00	0.42	0.09	0.28	0.01	0.18	0.00	1.00	0.95	0.01	0.89	1.00	0.04	0.00
BUR		0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.52	1.00	0.02	0.00	1.00	0.99	0.73	0.01	0.01	1.00	1.00	1.00	1.00	0.97	1.00	1.00	1.00	1.00	0.99	1.00
CHEN			0.00	1.00	1.00	1.00	0.87	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.02	0.00	1.00	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS1				0.00	1.00	1.00	0.92	1.00	0.00	0.11	0.00	1.00	1.00	0.00	0.07	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
LS2					0.00	1.00	1.00	0.94	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	1.00	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS3						0.00	1.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	1.00	0.90	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS4							0.00	0.48	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.82	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LS5								0.00	0.00	0.34	0.00	1.00	1.00	0.00	0.03	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00
LS6									0.00	0.65	1.00	0.05	0.00	1.00	0.99	1.00	0.01	0.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LS7										0.00	0.00	1.00	0.21	0.01	0.99	0.00	0.78	0.24	0.15	0.00	0.08	0.00	1.00	0.56	0.00	0.34	0.81	0.00	0.00
LS8											0.00	0.00	0.00	1.00	0.09	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.16	1.00	1.00	0.31	0.20	1.00	1.00
LS9												0.00	1.00	0.00	0.68	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.50	0.03	0.00	0.02	0.21	0.00	0.00
LS10													0.00	0.00	0.13	0.00	1.00	0.97	0.01	0.00	0.00	0.00	0.06	0.00	0.00	0.04	0.02	0.00	0.00
LS11														0.00	0.57	1.00	0.00	0.00	1.00	1.00	1.00	1.00	0.38	1.00	1.00	0.86	0.88	1.00	1.00
LS12															0.00	0.00	0.16	0.08	0.97	0.06	0.53	0.23	1.00	0.99	0.68	1.00	1.00	0.07	0.07
LS13																0.00	0.00	0.00	0.53	1.00	0.98	0.90	0.00	0.36	0.63	0.00	0.00	1.00	1.00
LS14																	0.00	1.00	0.00	0.00	0.00	0.00	0.11	0.01	0.00	0.00	0.01	0.00	0.00
LS15																		0.00	0.00	0.00	0.00	0.00	0.04	0.01	0.00	0.00	0.00	0.00	0.00
PC1																			0.00	0.93	1.00	1.00	0.97	1.00	1.00	1.00	1.00	0.93	0.99
PC2																				0.00	1.00	1.00	0.08	0.84	1.00	0.14	0.15	1.00	1.00
PC3																					0.00	1.00	0.89	1.00	1.00	0.93	0.85	1.00	1.00
PC4																						0.00	0.27	1.00	1.00	0.43	0.36	1.00	1.00
PC5																							0.00	1.00	0.24	1.00	1.00	0.10	0.03
PC6																								0.00	1.00	1.00	1.00	0.84	0.99
PC7																									0.00	0.53	0.72	1.00	1.00
PC8																										0.00	1.00	0.27	0.02
PC9																											0.00	0.34	0.16
PC10																												0.00	1.00
SCA																													0.00

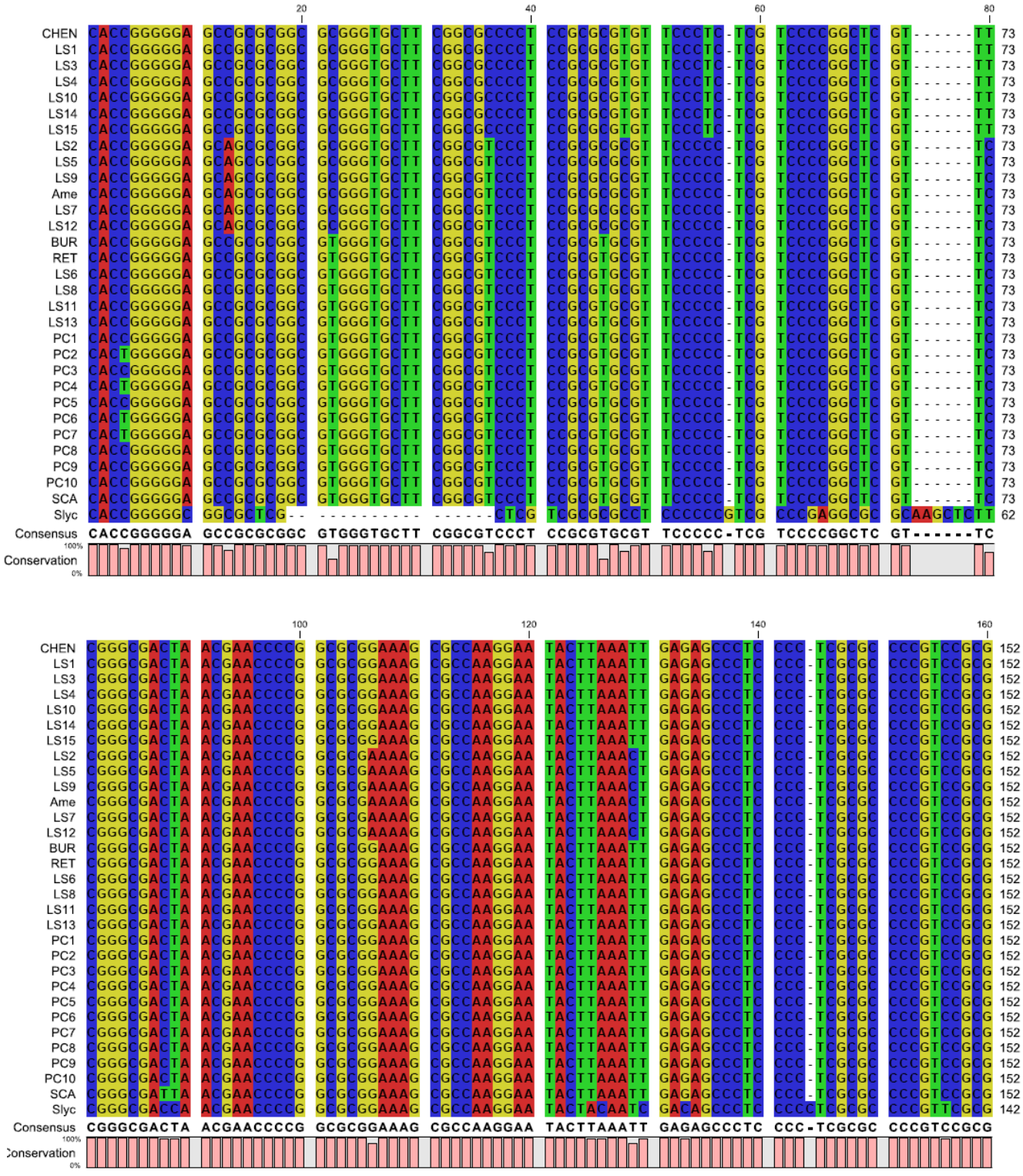
© University of Pretoria

Appendix D

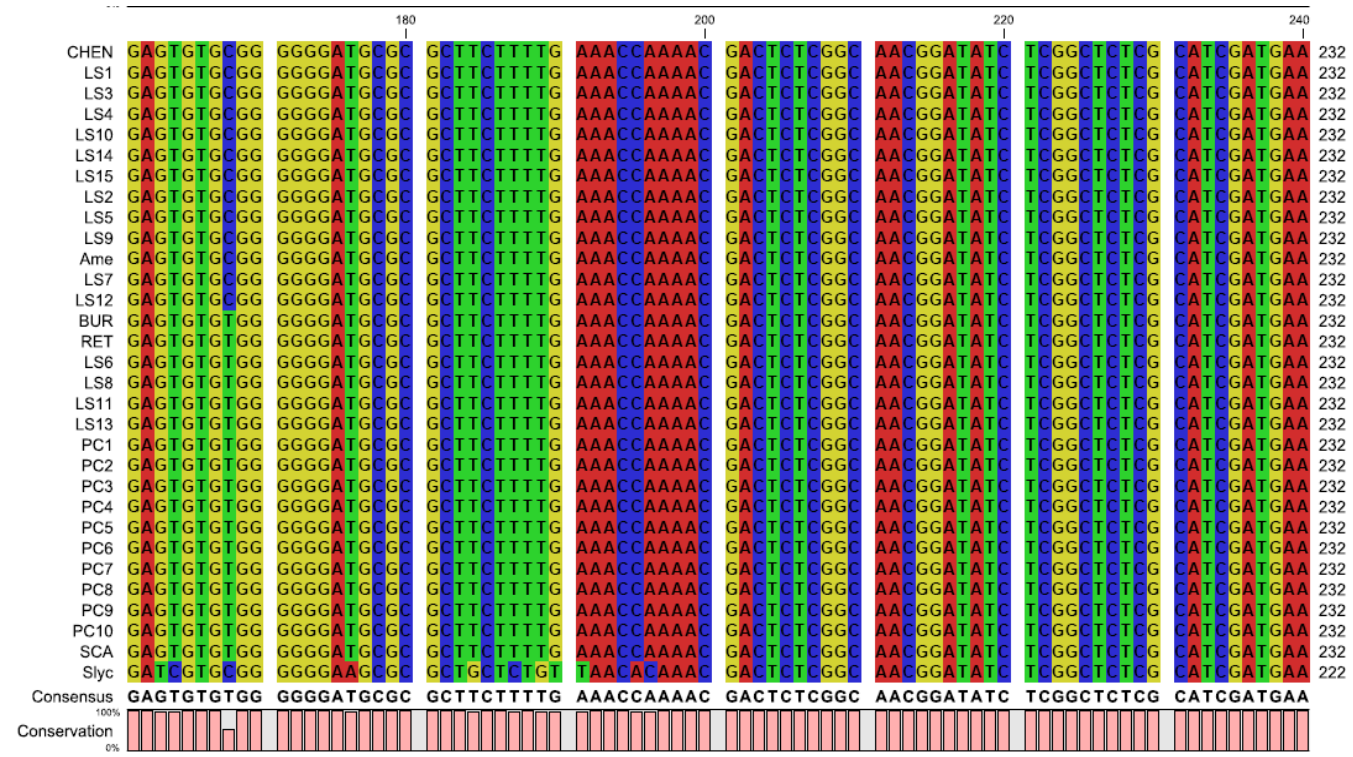
Multiple Sequence Alignments

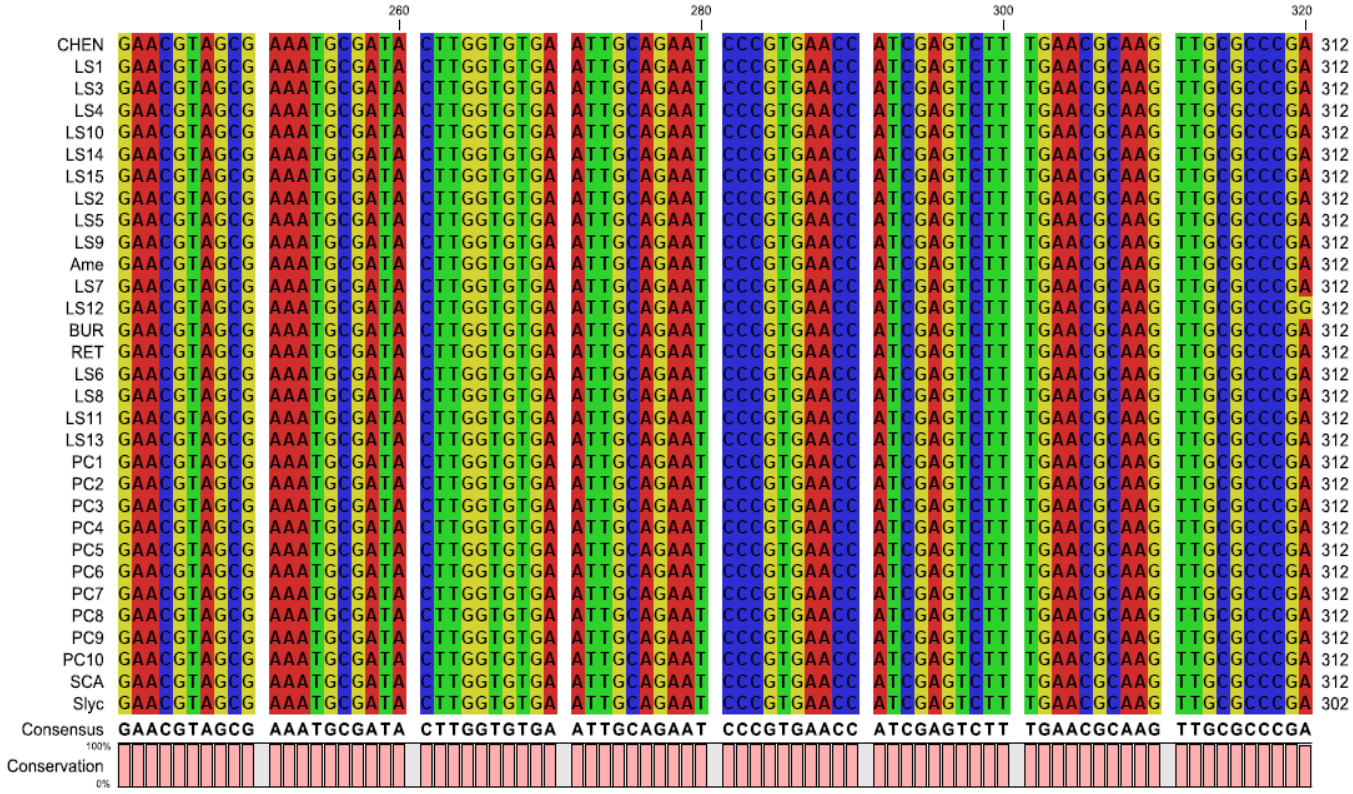
© University of Pretoria

ITS:

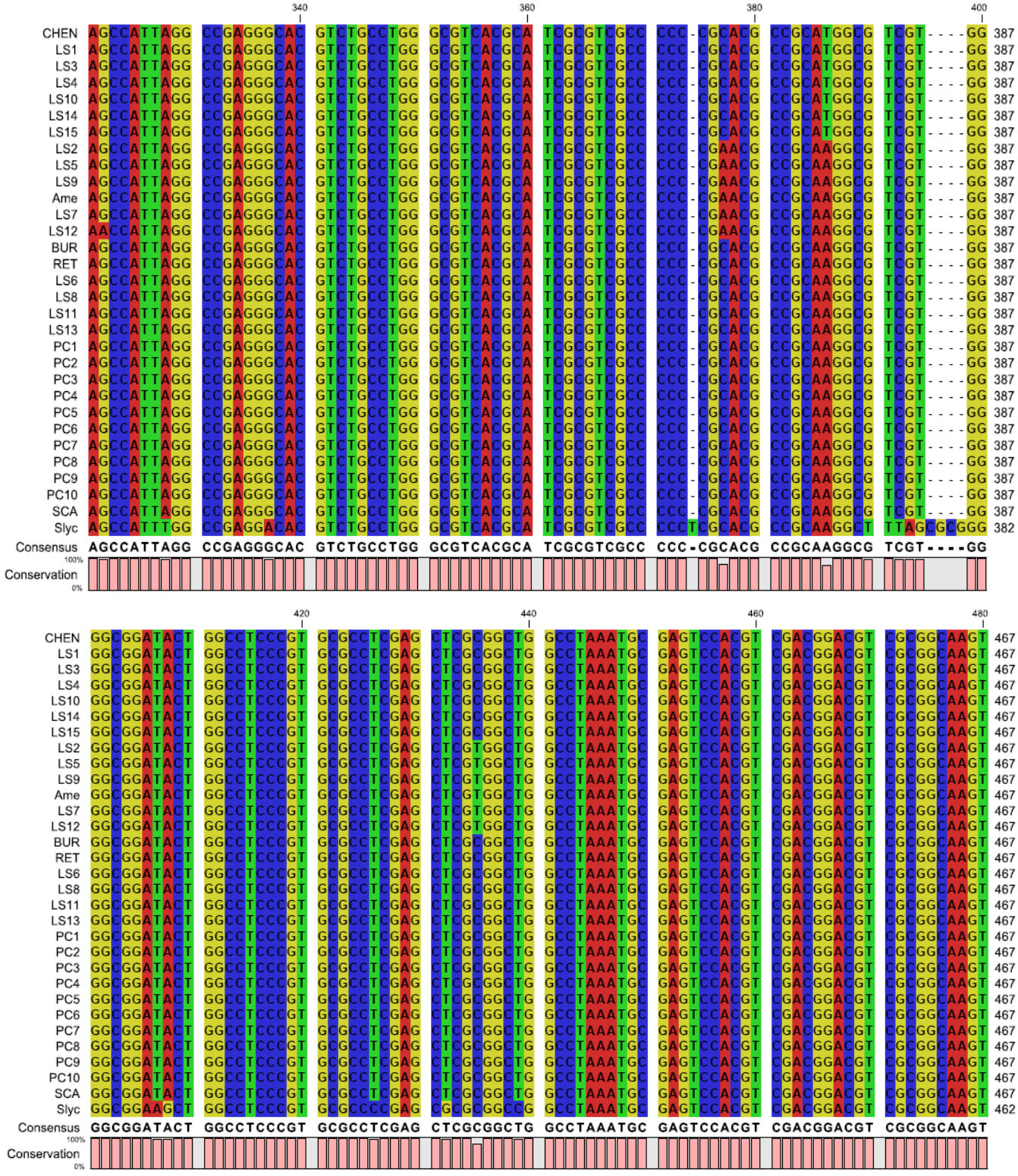


© University of Pretoria

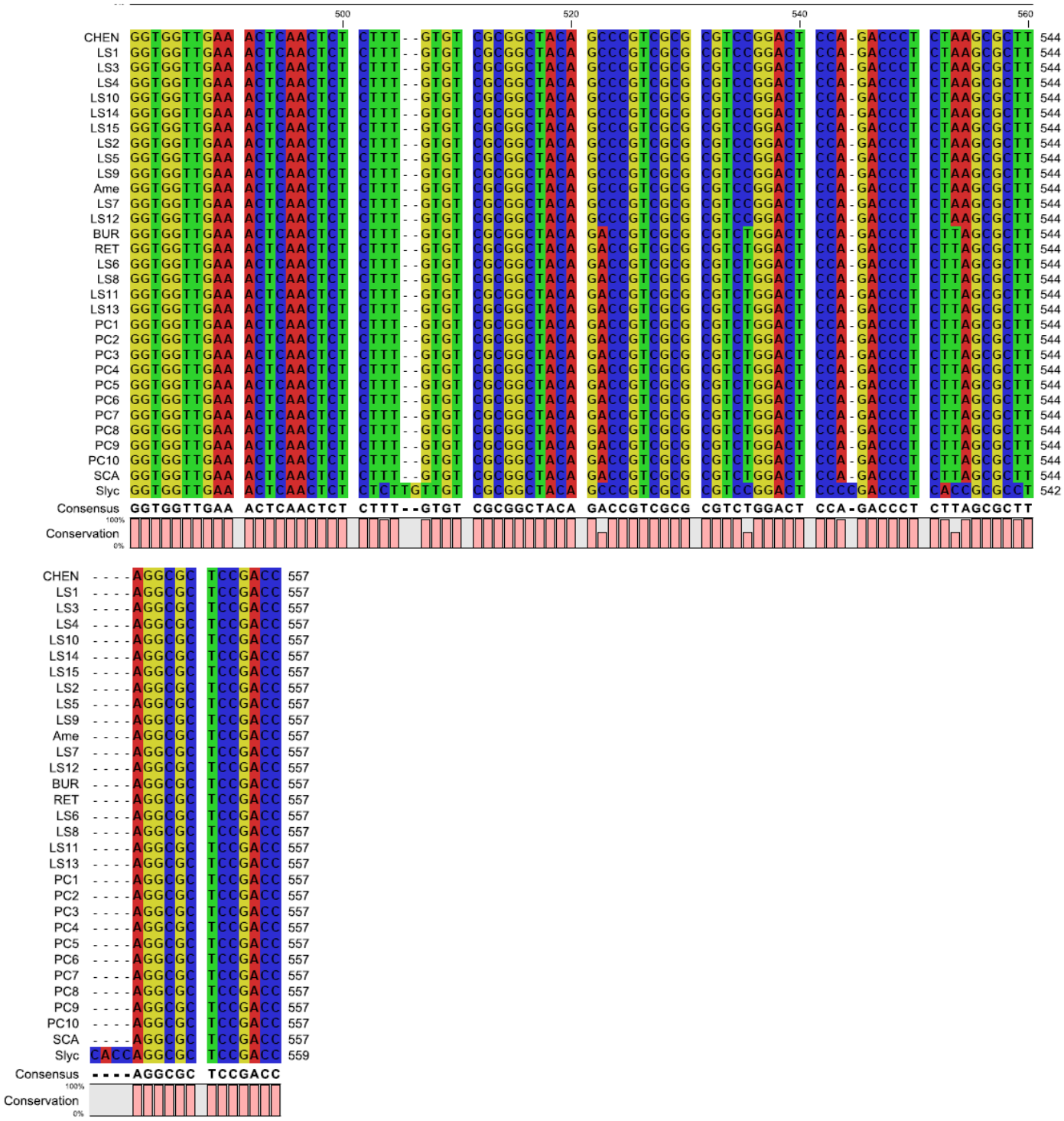




© University of Pretoria

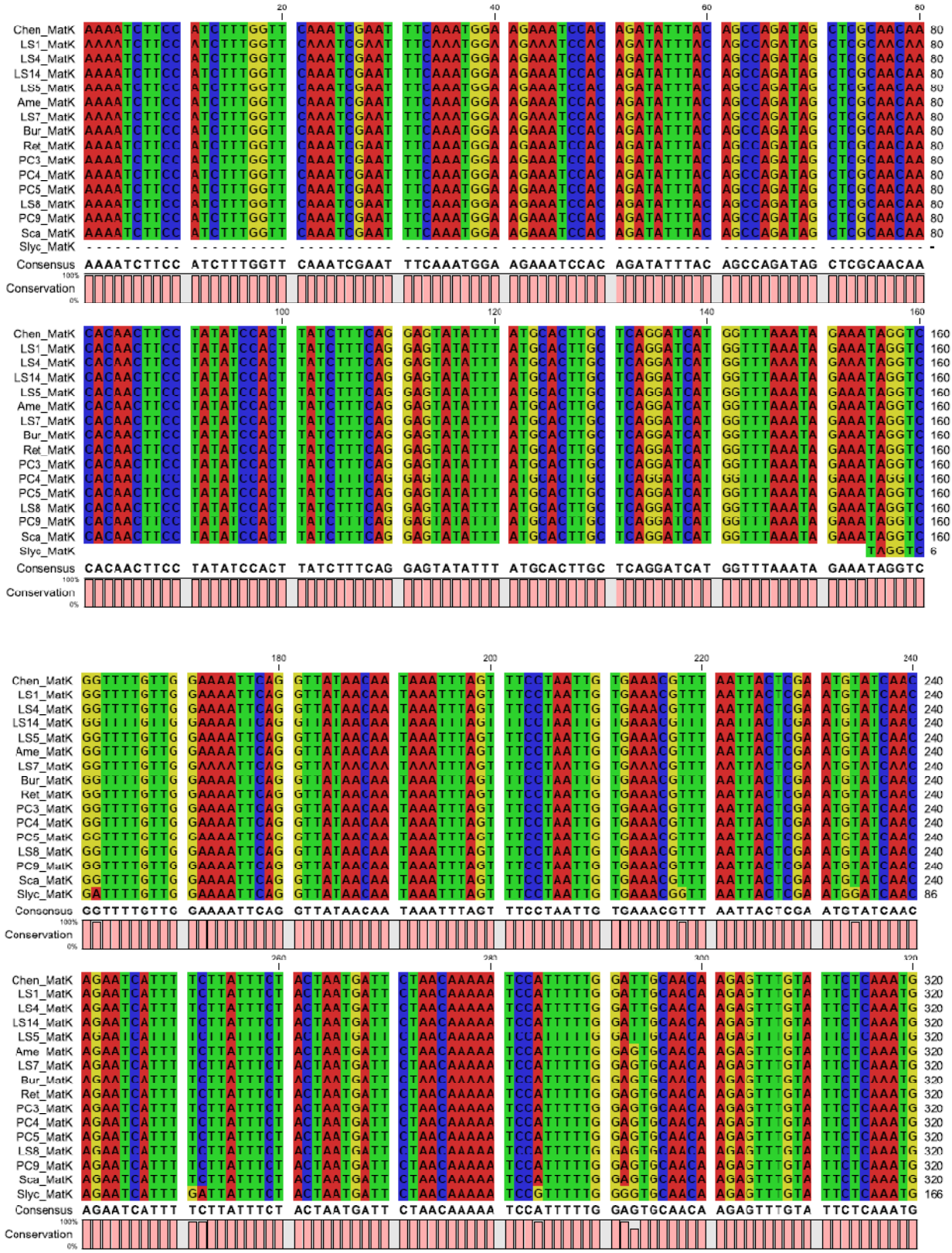


© University of Pretoria

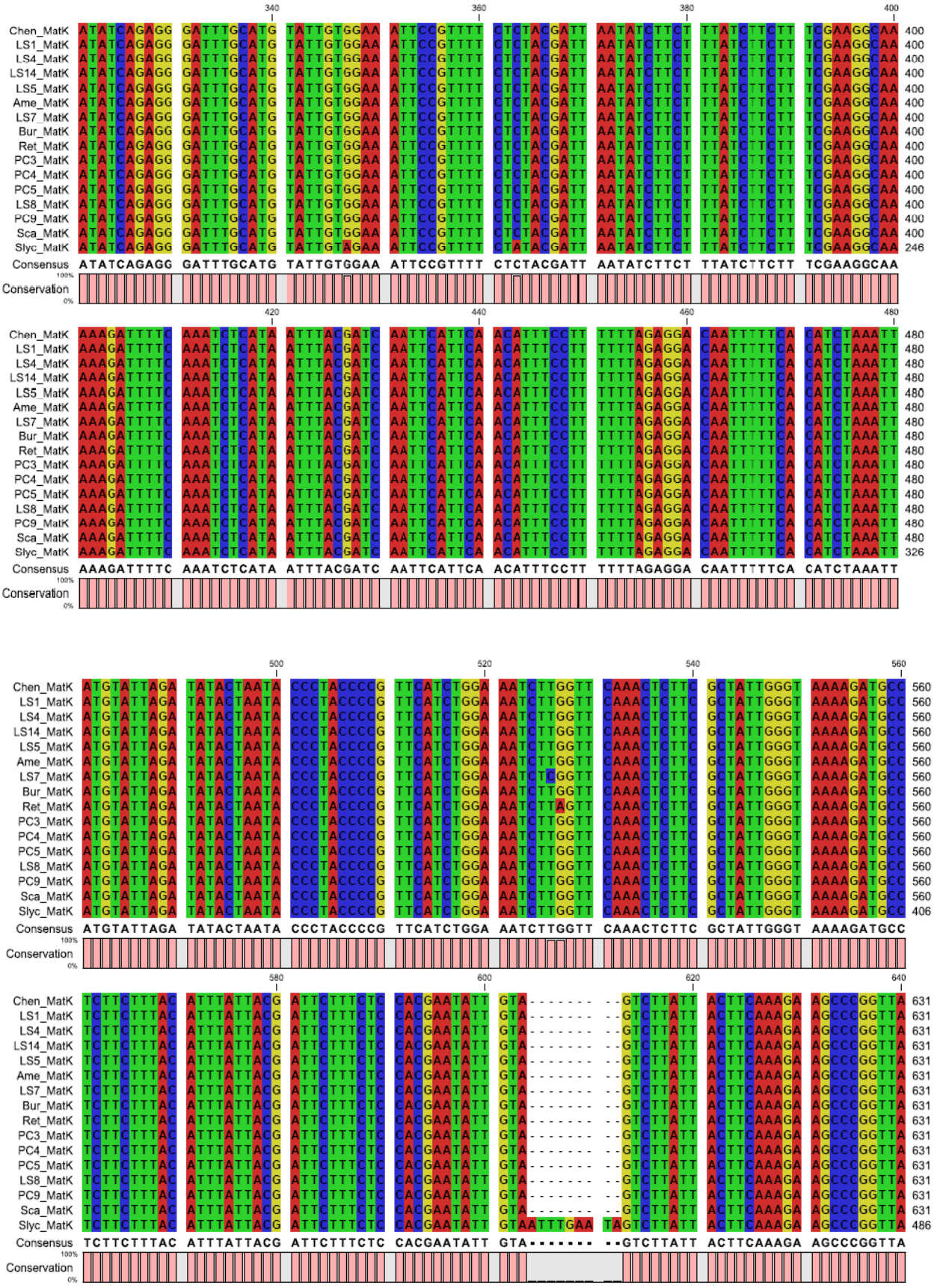


© University of Pretoria

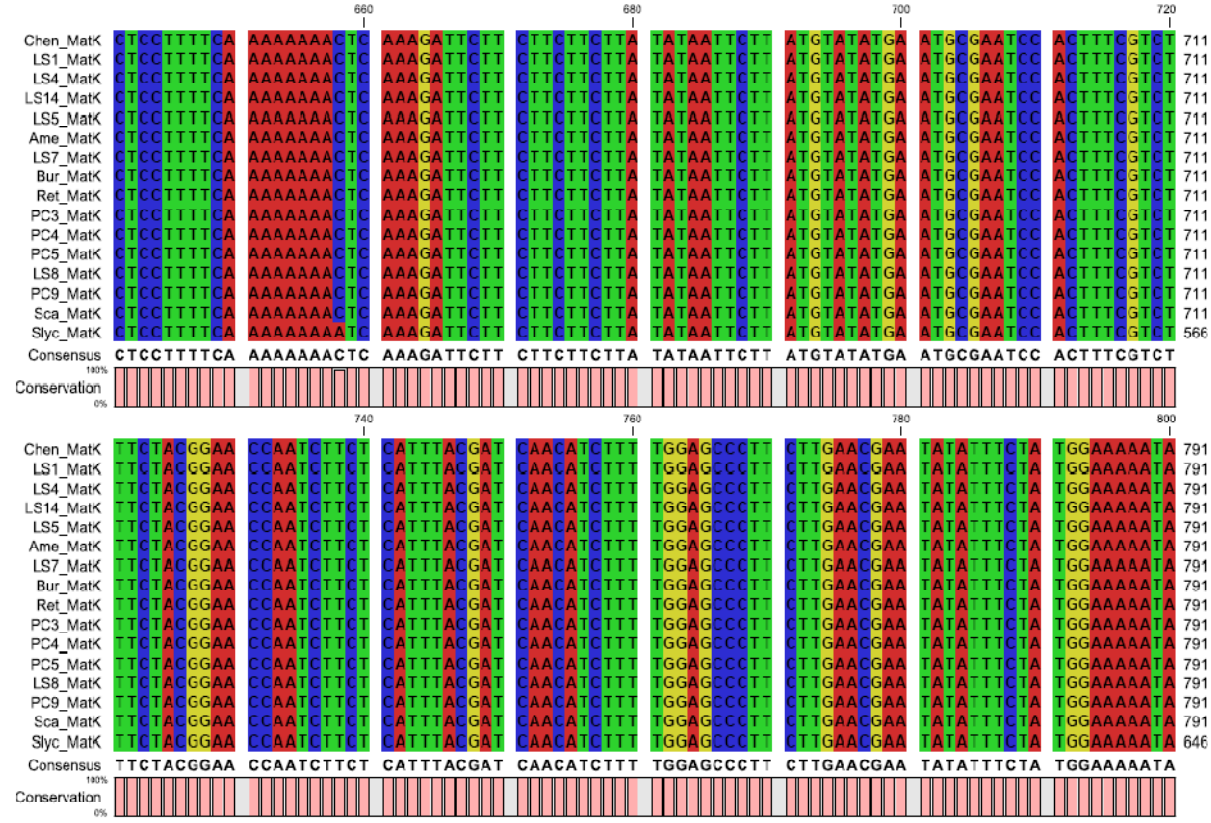
*mat*K:

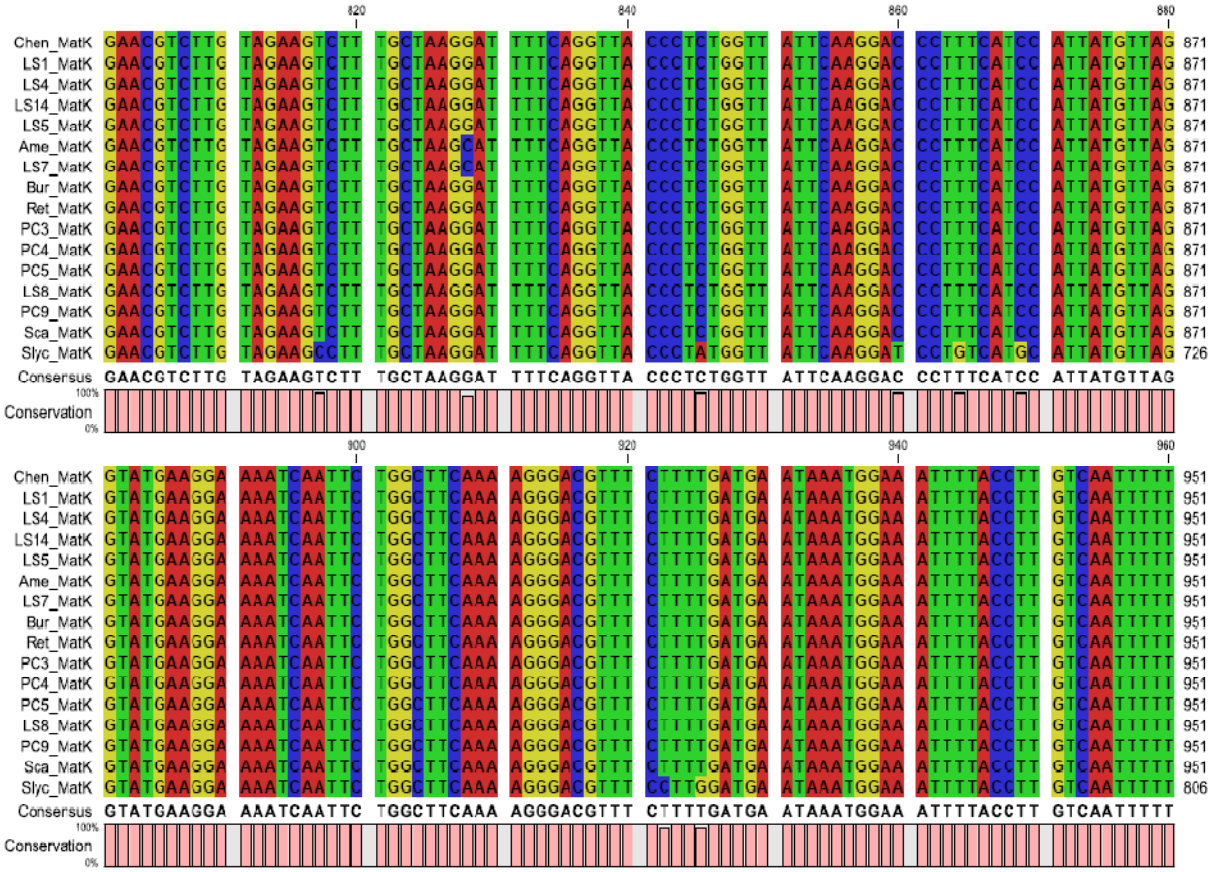


© University of Pretoria

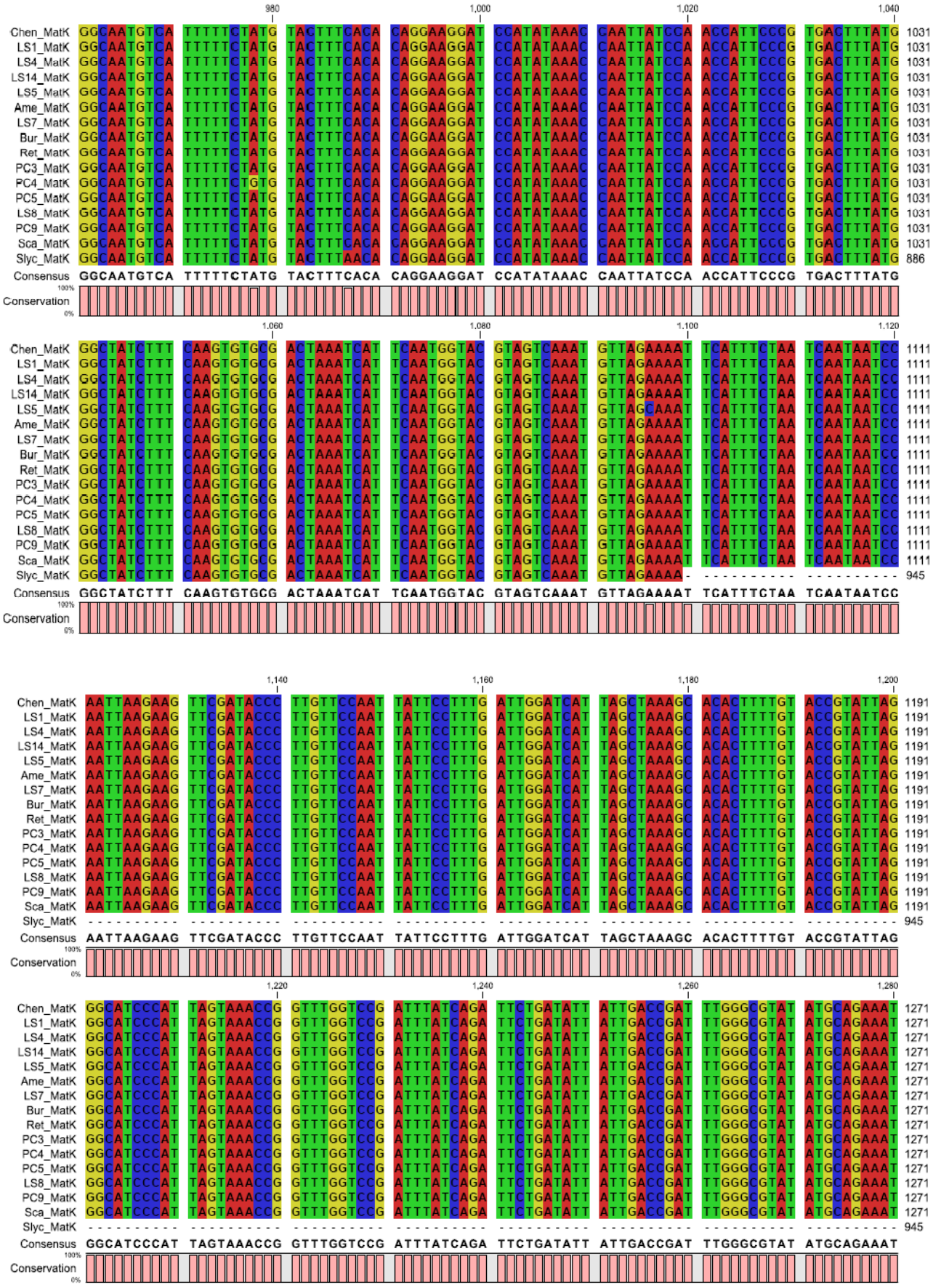


© University of Pretoria

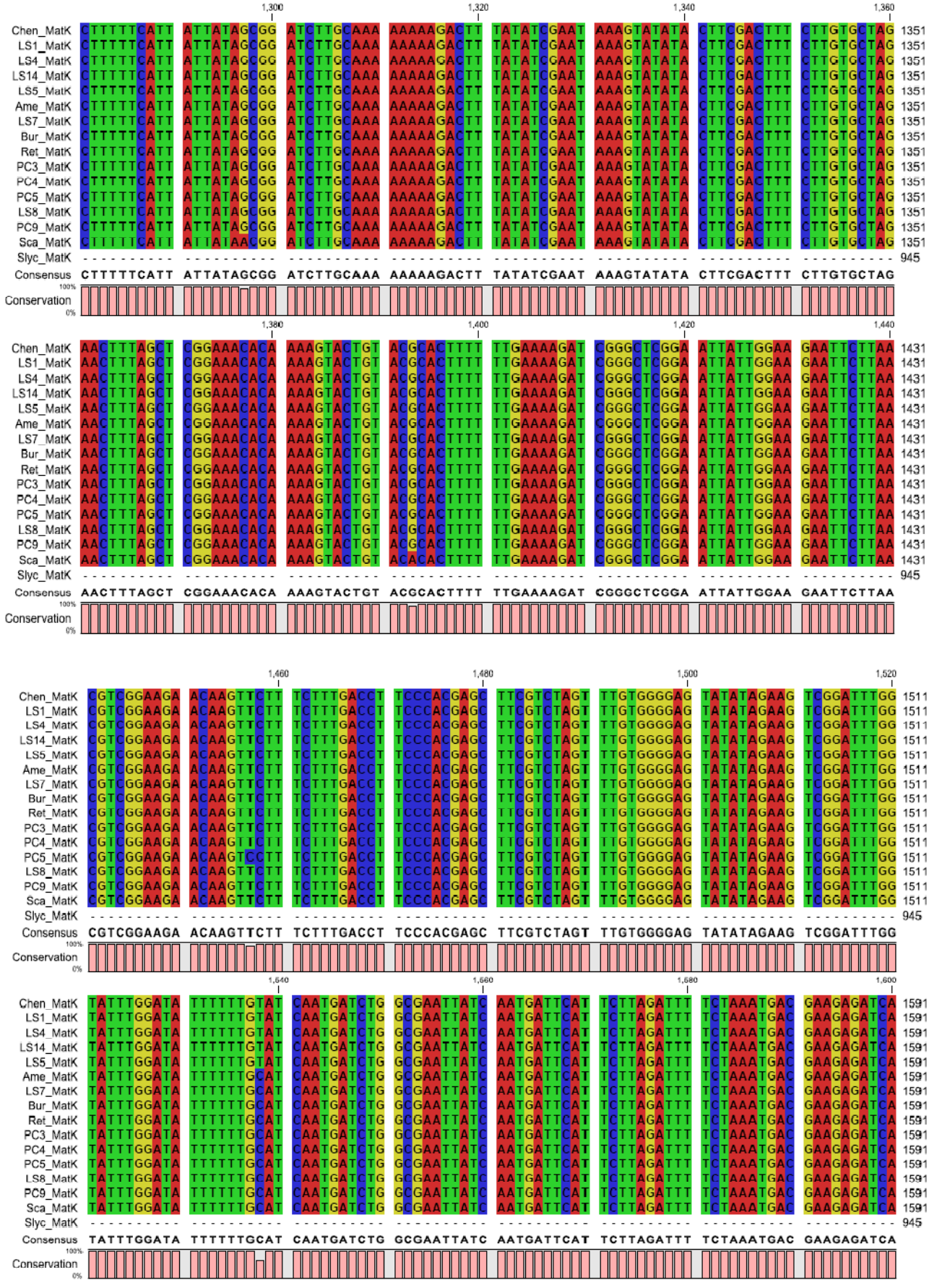




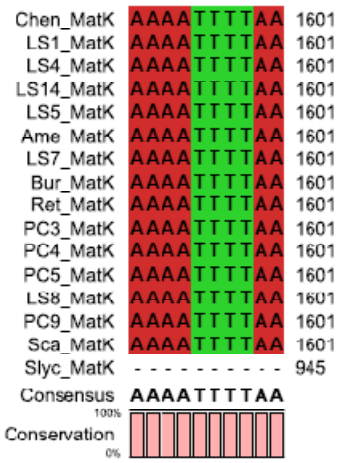
© University of Pretoria



© University of Pretoria

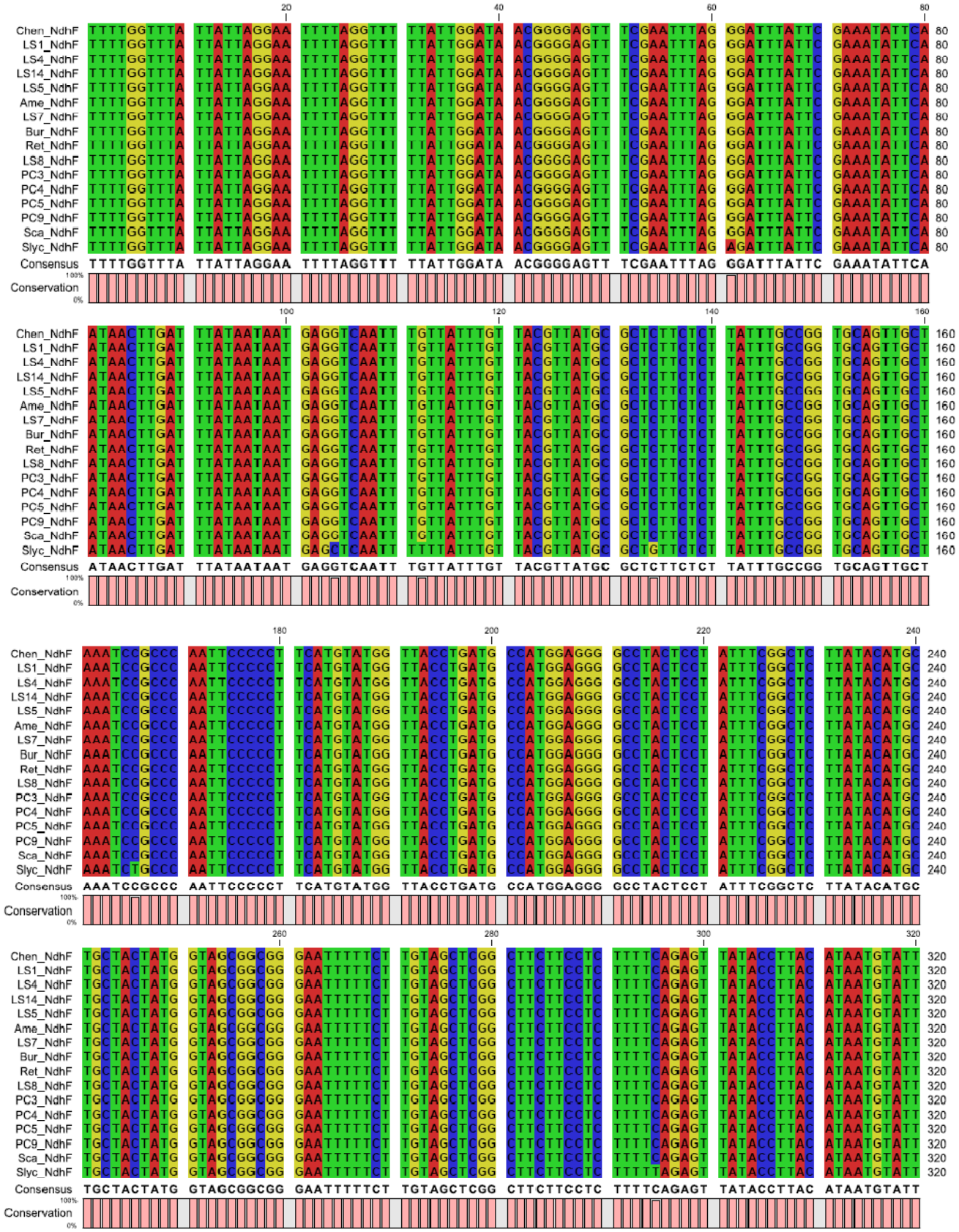


© University of Pretoria

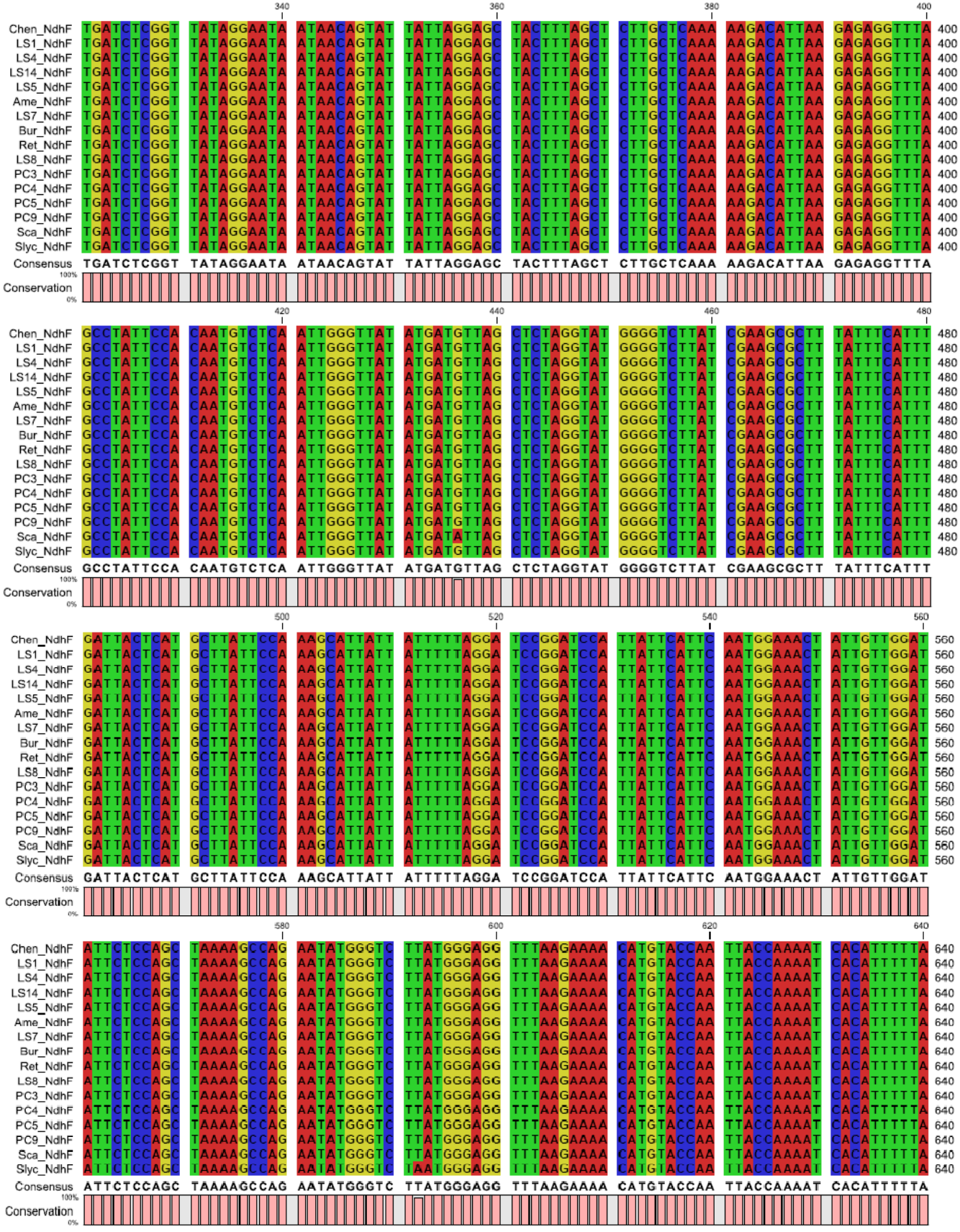


© University of Pretoria

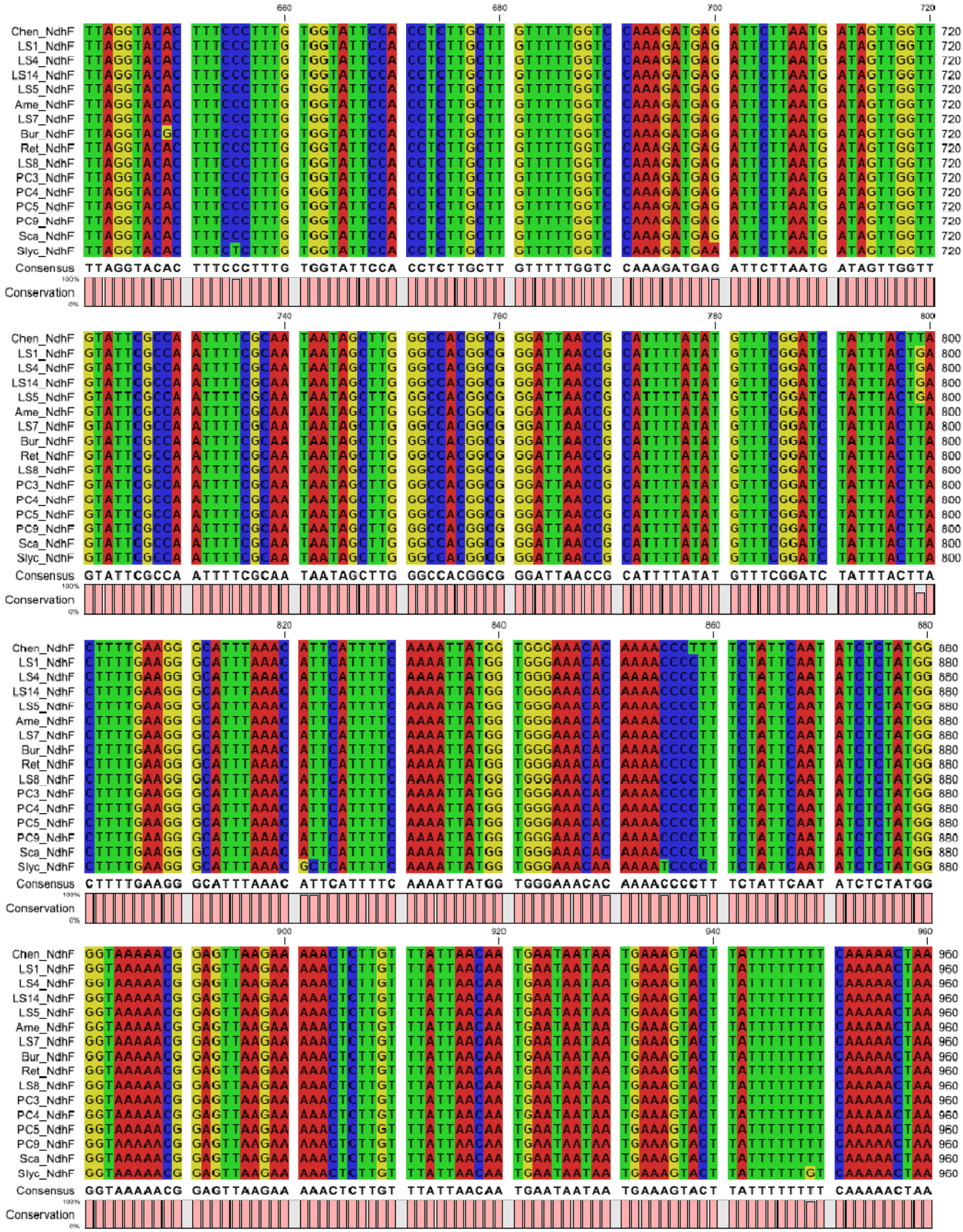
*ndh*F:



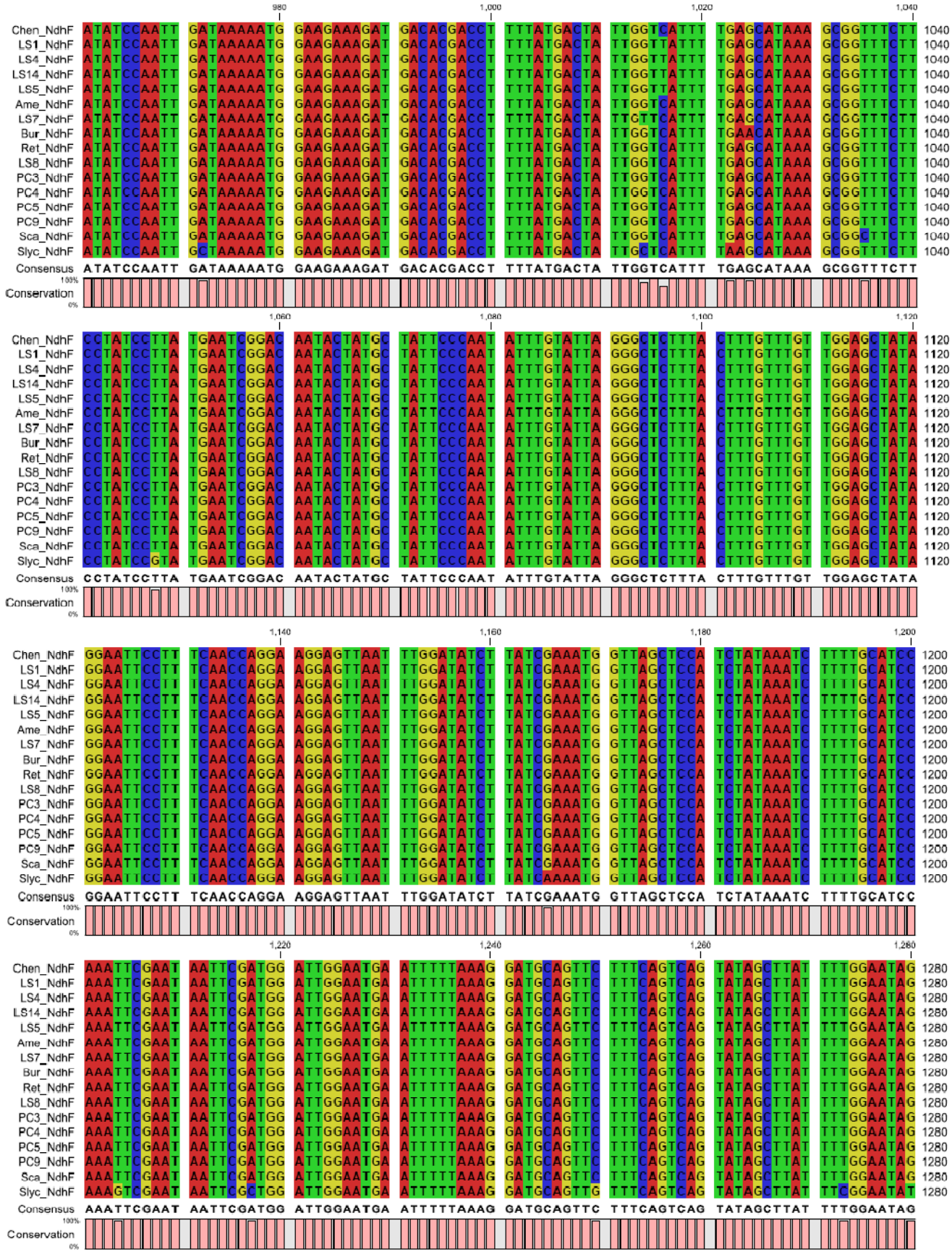
© University of Pretoria



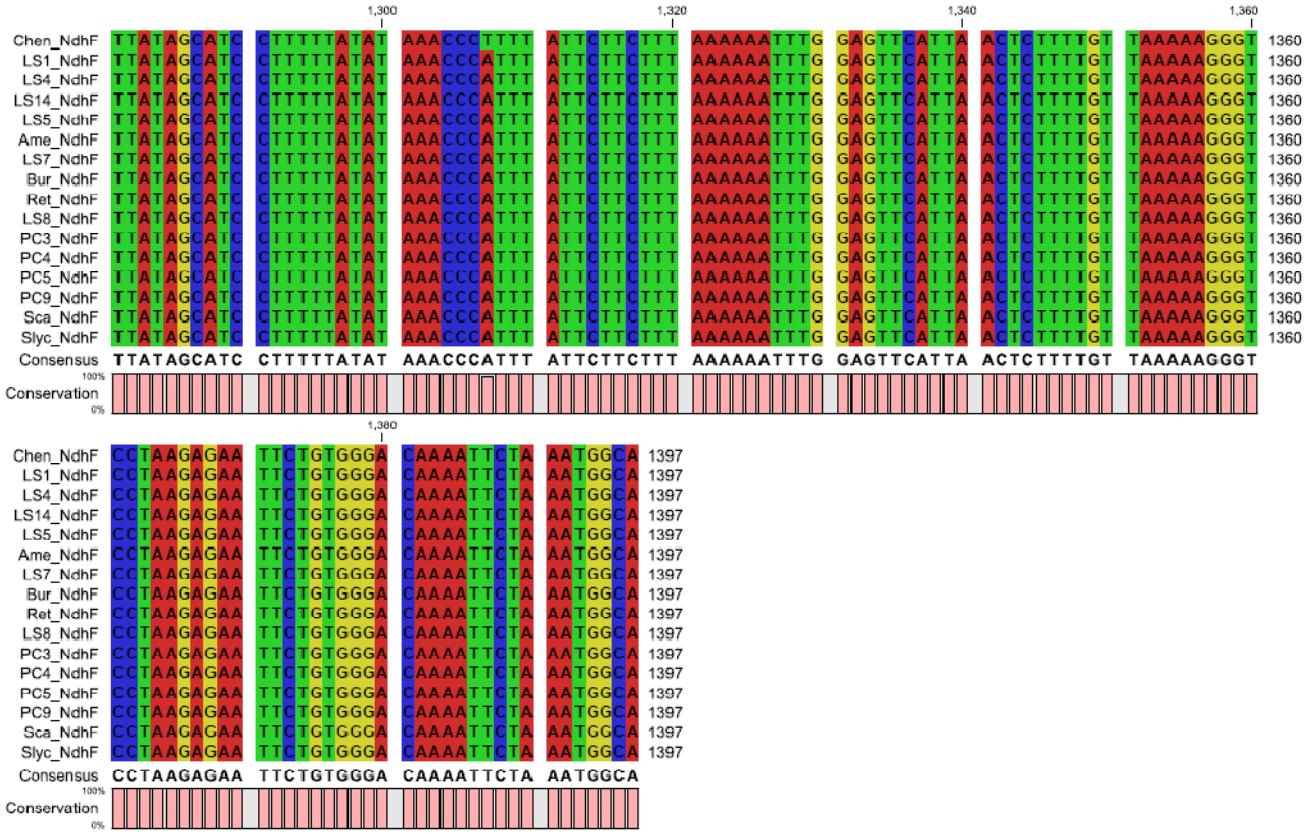
© University of Pretoria



© University of Pretoria

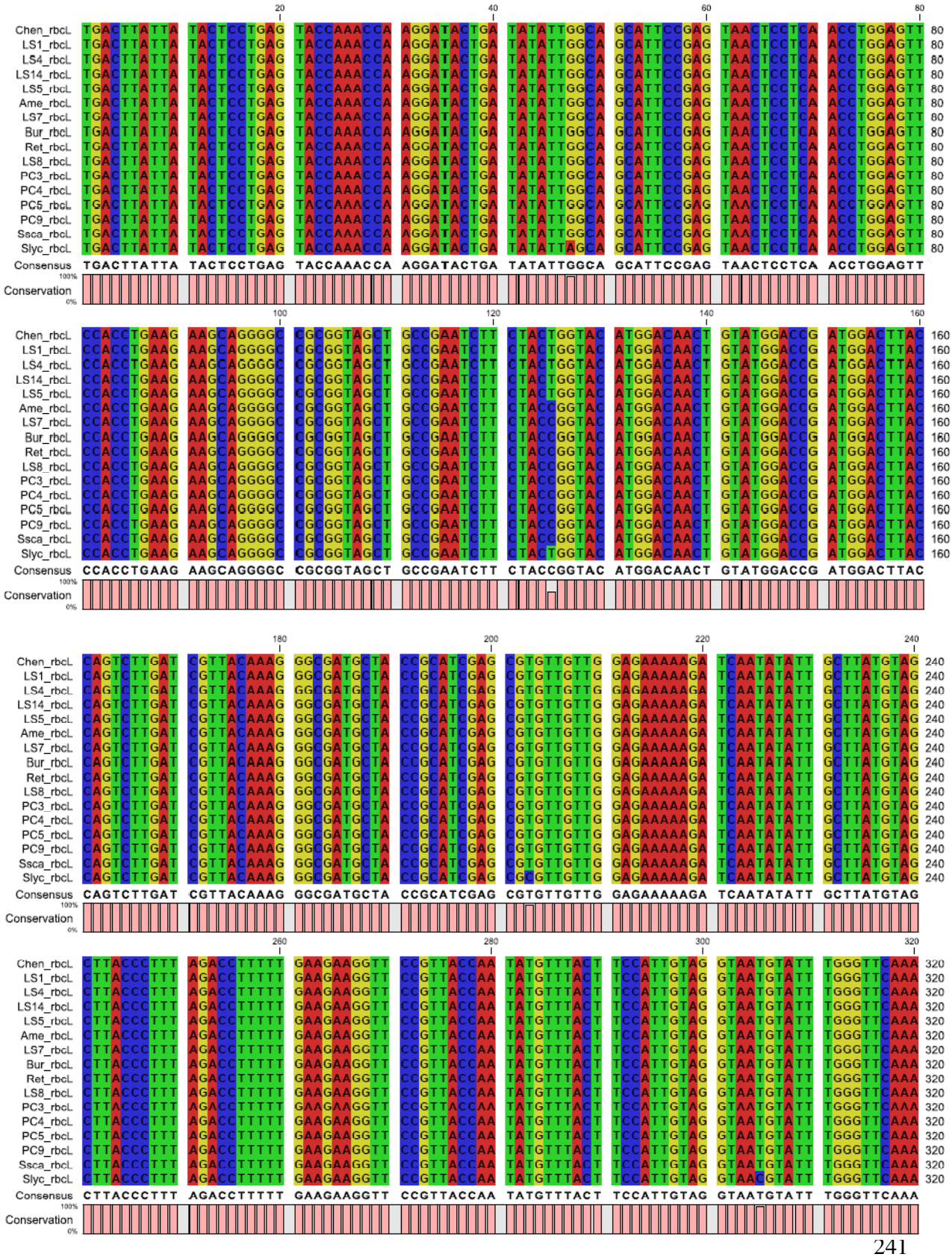


© University of Pretoria

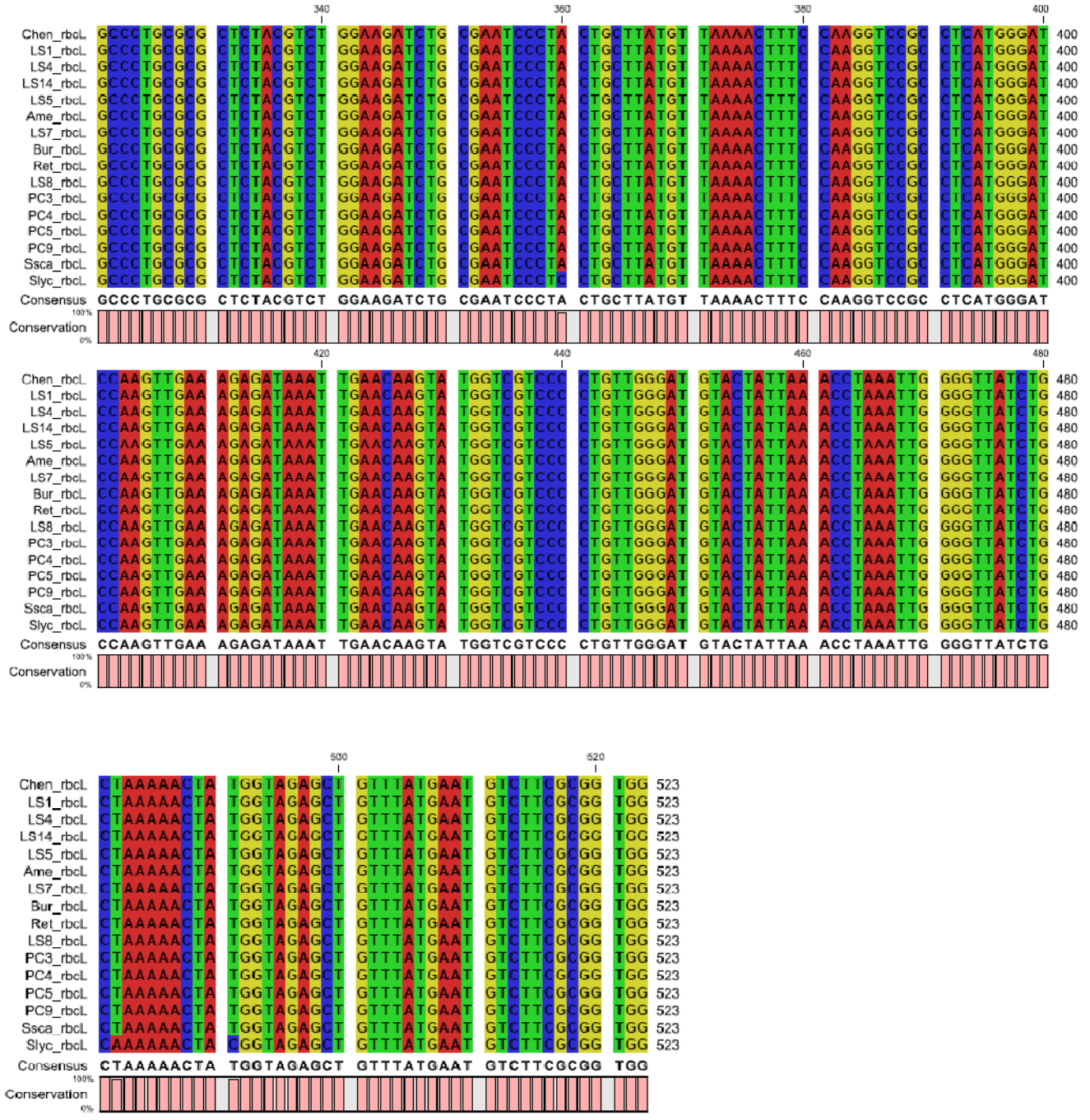


© University of Pretoria

*rbc*L:



© University of Pretoria



© University of Pretoria

Appendix E

IUPAC Nucleotide Code

© University of Pretoria

IUPAC Code	Meaning	Origin of Description
G	G	**G**uanine
A	A	**A**denine
T	T	**T**hymine
C	C	**C**ytosine
R	G or A	pu**R**ine
Y	T or C	p**Y**rimidine
M	A or C	a**M**ino
K	G or T	**K**etone
S	G or C	**S**trong interaction
W	A or T	**W**eak interaction
H	A or C or T	not-G, **H** follows G in the alphabet
B	G or T or C	not-A, **B** follows A in the alphabet
V	G or C or A	not-T (not-U), **V** follows U in the alphabet
D	G or A or T	not-C, **D** follows C in the alphabet
N	G or A or T or C	a**N**y

Adapted from:

Biochemical Journal **(1985)** 229, 281-286 281

Nomenclature Committee of the International Union of Biochemistry (NC-IUB)

Nomenclature for incompletely specified bases in nucleic acid sequences

Recommendations 1984

© University of Pretoria

Appendix F

DArT Marker Frequencies

© University of Pretoria

Percentage of markers that are polymorphic between samples on the array

Accession	AME	BUR	CHEN	LS1	LS2	LS3	LS4	LS5	LS6	LS7	LS8	LS9	LS10	LS11	LS12	LS13	LS14	LS15	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	RET	SCA
AME	0	9.3	74.6	76.3	64.9	74.6	76.8	64.6	9	2.6	9.4	63.4	69.5	8.9	1.9	9.4	73.8	75.3	9.6	9.8	9.4	9.4	9.2	10.5	9.7	9.4	9.6	9.2	8.9	6.5
BUR		0	74.8	76.6	70.9	74.1	77	70.8	0.8	14.9	1.8	70	71.1	1.6	14.4	1.6	74.8	75.9	1.6	1.8	2.1	1.8	1.5	2.9	1.6	1.6	1.4	1.2	2.7	11.5
CHEN			0	2.3	39.1	4.9	3.3	39	75.3	70.7	76.9	40.2	9.7	74.9	71.4	75.5	5.5	3.3	76.4	76.5	75.7	76.1	75.5	76.8	75.6	77.3	75.8	75.2	74.6	76.4
LS1				0	37.6	5.3	1.2	37.6	77	73	78.4	38.5	6.7	76.5	73.6	77.2	3.3	1.8	77.9	77.6	77.2	77.6	77.4	76.4	77.1	79	77.6	76.9	76.3	78.3
LS2					0	42.4	38.5	1.6	71.2	59.6	72.2	1.4	42	70.7	59.8	71.1	41	39.7	71.6	71.2	71.3	71.5	71.1	70.1	71.4	72.9	71.9	71.4	71.2	62.3
LS3						0	4.5	42.4	74.6	71.6	76.2	43.3	11.7	73.9	72.2	74.9	6.3	5.8	75.6	75.7	74.9	75.3	75	75.1	74.5	76.5	75	74.5	74	75.9
LS4							0	38.3	77.6	73.4	78.8	39	7.1	76.8	74	77.7	3.2	2.2	78.3	78.1	77.6	78	77.8	76.9	77.1	79.4	77.9	77.4	76.9	78.9
LS5								0	71	59.8	72.1	0.6	40.8	70.6	59.7	71	40.5	39.3	71.5	71	71.1	71.4	70.9	70	71.2	72.8	71.7	71.3	71.3	62.4
LS6									0	14.4	1.1	70.3	70.7	1.3	13.9	1.2	75	76.2	1.2	1.3	1.5	1.4	0.9	2.5	1.2	1.1	1	0.9	2.3	11
LS7										0	14.4	58.2	66.9	14.2	1.4	14.7	70.9	71.9	15	15.2	14.6	14.9	14.5	16.2	15.2	14.8	14.6	14.6	14.5	7.7
LS8											0	71.2	71.8	1.8	14.1	2.1	76.2	77.3	1.7	1.9	2.3	2	1.4	3.1	1.5	1	1.5	1.7	2.9	12
LS9												0	40.6	69.8	58.2	70.1	40.6	39.8	70.5	70.2	70.4	70.6	70	69.1	70.4	71.9	71	70.6	70.6	61.4
LS10													0	69.9	66.7	70.8	6.8	6.9	71.6	71	70.9	71.1	70.8	68.8	70.5	72.5	71.3	70.7	70.2	72.7
LS11														0	13.6	1.7	73.9	75.5	1.8	2	1.8	1.6	1.5	3	1.6	1.8	1.6	1.4	2.5	11.6
LS12															0	14.1	71.1	72.3	14.3	14.4	14	13.9	13.8	15.4	14.7	14.3	14.4	14.1	13.9	7.6
LS13																0	74.7	76.3	1.5	1.3	1.3	1.7	1.4	2.8	1.9	2	1.7	1.7	3.1	11.8
LS14																	0	2.2	75.4	75.1	74.8	75.5	75	74.2	74.9	76.8	75.2	74.7	74.2	76
LS15																		0	76.7	76.7	76.3	76.8	76.3	76.1	76.4	78	76.7	76	75.5	77.2
PC1																			0	0.8	1.6	1.4	1.2	2.9	1.7	1.6	1.5	1.5	2.9	11.6
PC2																				0	1.2	1.3	1.2	2.3	1.7	1.6	1.2	1.5	3.1	11.7
PC3																					0	1.4	1.6	2.6	1.8	1.8	1.3	1.7	2.6	11.6
PC4																						0	1.2	2.7	1.7	1	1.3	1.3	2.6	11.6
PC5																							0	2.5	1.5	1.1	1.3	1.4	2.7	11.5
PC6																								0	2.2	2.9	2.5	2.5	4	13.5
PC7																									0	1.5	1.3	1.3	2.8	11.8
PC8																										0	1	1.3	2.4	11.5
PC9																											0	1	2.5	11.5
PC10																												0	2.5	11.1
RET																													0	11.1
SCA																														0

© University of Pretoria