THE UNIVERSITY
of EDINBURGH

Conserving the biodiversity of Kuwait through DNA barcoding the flora

A thesis submitted for the degree of Doctor of Philosophy

Mansour Abdullah

Institute of Molecular Plant Sciences
School of Biological Sciences
University of Edinburgh
&
Royal Botanic Garden Edinburgh

July 2017

Declaration:

I hereby declare that the work contained in this thesis is my own, unless otherwise acknowledged and cited. This thesis has not in whole or part been previously presented for any degree.

Mansour Abdullah
25th July 2017

Abstract

Biodiversity across the globe is threatened. Rapid surveying and monitoring techniques are required to understand the origin of the threats to biodiversity and to enable conservation actions to be undertaken. Kuwait is an arid desert country with a small flora of only 402 species. This flora is endangered by environmental factors, overgrazing, and human activities. DNA barcoding the flora and using Next Generation Sequencing (NGS) technologies allowed us to identify plants to species level, conduct a molecular taxonomic revision, and distinguish plant diversity found in soil environmental DNA samples. After investigating the discriminatory power of five commonly used DNA markers from plastid (*matK*, *rbcL*, *trnH-psbA*, *trnL*) and a nuclear genome (ITS2) on four largest genera of the flora using phylogenetics reconstruction tree based methods, two barcoding markers (*rbcL* and ITS2) were assigned to build a DNA reference library of the flora. Furthermore, the DNA reference library was tested to identify the plant diversity found below-ground level and comparing it with that above-ground, using environmental soil samples collected from both species rich and poor habitats in Kuwait by applying high-throughput sequencing methods. The DNA database provided in this study could be used as a reference library for the identification process and contribute towards the future of molecular taxonomy, biodiversity and ecological research in Kuwait.

Acknowledgments

I would like to give my sincere thanks and gratitude to my supervisors, Catherine Kidner, Tony Miller and Pete Hollingsworth, who have been supportive throughout my project and providing invaluable knowledge, skills and assistance which added a lot to my academic experience. It was a great privilege to have worked under their supervision.

I am grateful to CMEP-RBGE, especially Sabina Knees, Sophie Neale and Alan Forrest for their kind support and professional guidance throughout my project.

Special thanks to STS-RBGE, Michelle Hart, Laura Forrest, Ruth Hollands and Frieda Christie for expert advice, training and support in the lab.

I would like to express my appreciation and thanks to the RBGE Herbarium staff, especially David Harris, for the support, guidance and for processing my herbarium specimens and digitizing.

Many thanks to Stephan Helfer, for guiding me through the permit process to import my environmental samples.

I owe my deepest gratitude to K.T. Mathew for his kindness, encouragement and support during my field visits and sampling specimens from Kuwait University Herbarium.

Many thanks to all staff members and Ph.D. students at RBGE and UoE who have given help and support throughout my academic years. Special thanks to the Government of Kuwait and PAAET for financial support and funding my Ph.D. at the University of Edinburgh.

I am eternally grateful to my kind parents, Fauzia and Taleb, for the support and prayers of encouragement and faith throughout my entire life. Many thanks to my sister, Alia, and brothers, AbdulNasser, Mohammad, Manaf and Ali for the motivation, inspiration and the support they always provided.

Infinite thanks to my wife, Eman and my sons, Adam and Jassem for their support and encouragement and being patient while I'm away pursuing further studies.

Many thanks to all my cousins and friends for their spiritual support, especially Ali Dashti, Musaed Abdullah, Mohammad Al-Mousa, Meshal Abdullah, Taher Al Baghli, Hasan Al-Abdullah, Ahmad Al-Hasan, Essa Buhamad.

Above all, I thank God 'Allah', the most gracious and the most merciful, for providing me with knowledge, strength and energy to complete my Ph.D.

Table of Contents

Chapter 3 - Choice of gene regions

Chapter 4 - DNA barcoding the flora of Kuwait

Chapter 5 - Next-generation sequencing for molecular reconstruction of plant diversity using eDNA

List of Figures

List of Tables

List of Appendices

Chapter 1 Introduction

In recent years the loss of biodiversity has become of increasing concern. The degradation of natural habitats, ecosystems and land occurring throughout the world is principally due to climate change, the invasion of alien species and human activities (IUCN, 2013). A healthy biodiversity provides ecosystem services, biological resources, and social benefits (Kok et al., 2016). Conservation towards maintaining biodiversity has received much attention lately, with researchers aiming towards maximising ecosystem services. In Kuwait the loss of biodiversity is mainly caused by climate change, overgrazing and human activities which lead to land degradation and loss of vegetation (Al-Awadhi et al., 2005). Immediate action is required for the conservation of biodiversity to avoid further loss and to maximise ecosystem services; this will aid in the stability of the land and vegetation. Rapid assessments and ecological surveys using advanced molecular methods can be used to understand poorly known areas and to enable us to conserve the current biodiversity of Kuwait.

1.1 Aims of the thesis and chapter overview

The principle objective of this research is to contribute towards the conservation of Kuwait's plant biodiversity through the use of advanced molecular techniques. The current chapter gives an environmental and ecological background and literature review of Kuwait. Topics include changes in vegetation distribution, natural resources and land management and conservation and threats to the environment.

Chapter 2 is a revision and analysis of the flora of Kuwait. This chapter will summarise the current status of knowledge of the plants of Kuwait and update the checklist of the flora by reviewing relevant literature sources and the results from field studies.

Chapter 3 will be the introductory chapter on DNA barcoding application and assessment of five barcoding markers and assigning two for the entire flora. In this chapter, I will be evaluating the performance of five DNA barcoding markers (*rbcL, matK, trnL, trnsH-psbA*, ITS2) on four largest genera of the flora (9 *Astragalus* spp., 7

Plantago spp., 4 *Helianthemum* spp., and 4 *Launea* spp.). The DNA markers will be evaluated based on how easy they PCR amplify, produce bidirectional sequences and show maximum discriminatory power. Molecular methods including DNA extraction, amplification and sequencing are outlined in this chapter.

Chapter 4 considers building a DNA reference library of the entire flora. This chapter will be based on the choice of barcoding markers tested in Chapter 3; I will establish a DNA barcoding reference library for the flora of Kuwait by evaluating the discriminatory power across sampled species, genera and family of the flora. Followed by a discussion on how efficient DNA barcoding could be to identify species of the flora.

Chapter 5, the purpose of this chapter is to explore the past and present plant communities by analysing the organic remains in environment DNA (eDNA) soil samples collected from both species rich and species poor habitats of Kuwait by applying Next-generation sequencing (NGS) methods.

In the final, Chapter 6, I will be discussing how the identification tools can contribute towards plant identification, ecological research and conservation plans in Kuwait. Also, I will review DNA barcoding and NGS methods and assess their potential role in contributing towards plants species identification tools appropriate for vegetation restoration in Kuwait and similar desert environments.

This study has the following aims:

1. To revise and update the checklist of the flora of Kuwait
2. To test and evaluate five standard universal markers from the genomic regions of the plastids (*matK*, *rbcL*, *trnL*, *trnH-psbH*) and one nuclear ribosomal (ITS2) on four largest genera and assigning two markers for the entire flora.
3. To build a reference library of DNA barcodes for the entire flora of Kuwait.
4. To assess the plant diversity in the soil seed bank collected from rich and poor habitats from Um Neqa, North-east of Kuwait, by applying NGS methods, with a view to reconstruct the vegetation in highly degraded areas.

Literature review

1.2 Geography, Topography and Geology, and Soils of Kuwait

1.2.1 Geography

The State of Kuwait is situated in the north-eastern part of the Arabian Peninsula between longitudes 46° 33' E to 48° 35' E and latitudes 28° 45' N to 30° 06' N. It has a total land area of 17,820 km^2 including nine offshore islands. It is bounded to the north and north-west by Iraq (with a 240 km border), to the south by Saudi Arabia (with a 220 km border), and to the east by the Arabian Gulf coastal shoreline which extends 170 km from north to south. The nine offshore islands (listed in Table 1.1) are all uninhabited. Bubiyan and Warba islands (located North-east of Kuwait) are the largest and consists of mud flats and large areas of salt marsh. These islands are situated in the estuary of the Tigris and Euphrates rivers which reach the Gulf via Shat Al-Arab channel, which runs along the southern border of Iraq and Iran. The geographical region of the Arabian Peninsula and the position of Kuwait in relation to the neighbouring countries are shown in Figure 1.1.

Kuwait has a population of 4,239,000 million, only one-third (1.4 million) are Kuwaitis, and the rest are expatriates (PACI, 2015). The Kuwait mainland is split into six capital governorates: Al-Asma 'The Capital' (population of 546,400 in 31 areas), Al-Ahmadi (populated by 878,400 in 29 areas), Al-Jahra (populated by 517,500 in 22 areas), Hawalli (populated by 915,500 in 18 areas), Al-Farwaniya (populated by 1,133,500 in 21 areas), and Mubarak Al-Kabeer (populated by 242,500 in 10 areas) (PACI, 2015). The three principal agricultural areas are Al-Wafra (in the south), Al-Abdali (in the north), and Sulaibiya (towards the centre). They cover a total area of approximately 350 km^2; that is some 2% of the total land area of Kuwait. Kuwait's primary natural resources are petroleum, natural gas and fishing. The primary source of fresh water is from desalinated and treated wastewater which is produced at six desalination plants with a maximum capacity of 950,000 m^3/day. It is mixed with ground water to make it suitable for drinking.

Table 1.1 Islands of the State of Kuwait

No.	Island name	Area
1	Warbah	37 km^2
2	Bubiyan	863 km^2
3	Miskan	750 m^2
4	Failaka	20 km^2
5	Auhah	0.34 km^2
6	Um Al-Namil	0.30 km^2
7	Kubbar	0.11 km^2
8	Qaruh	0.035 km^2
9	Um Al-Maradim	0.65 km^2

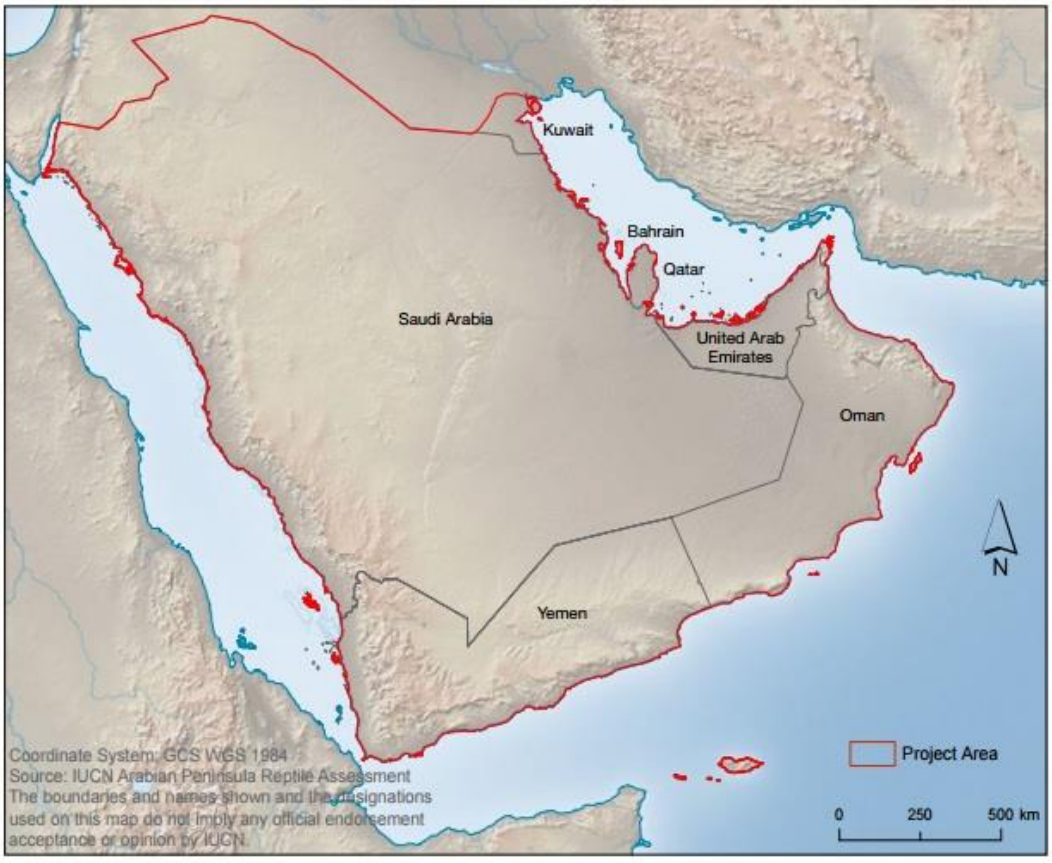


Figure 1.1 Map of the Arabian Peninsula *(source IUCN, 2012)*

1.2.2 Topography and Geology

Kuwait's surface topography is characterised by flat deserts which rise gradually from the coastal areas in the east to Wadi Al-Batin in the south-west where the highest point of Kuwait is found at 306 m (Figure 1.2). The east of the country is bounded by a coastal plain which runs from north to south across Kuwait bay. There are small coastal inlets, with highly saline water, forming salt flats (sabkhahs), along the northern bay which overlooks the Arabian Gulf. In the deserts, the otherwise often flat and featureless aspect is relieved by topographical features including sand dunes, alluvial fans, seasonal watercourses (known as 'wadis') and isolated hills and escarpments.

Sand dunes and small ridges of hills are composed of sand occur along the northern and southern part of the coastal plain. There are no permanent rivers in the country but the largest drainage (wadi) systems in the country are Wadi Al-Batin and Wadi Um Ar-Rimam which flow only after prolonged rain. Wadi Al-Batin is the largest wadi and occupies a land area of 60,000 km^2 across the northern part of the Arabian Peninsula. It extends from Hafar Al-Batin in Saudi Arabia and covers western and north-western parts of Kuwait. Wadi Al-Batin comprises a major alluvial fan ecosystem which is considered to be the largest non-active fan in the Peninsula (Al-Sulaimi and Pitty, 1995). It is the primary source of gravel for the upper 'Dibdibba' formation of gravel deposits (Al-Sulaimi and Pitty, 1995). The other major Wadi in Kuwait is the Wadi Um Ar-Rimam depression which is located north of Kuwait bay. It is approximately 8 km wide and 60 km long. However, perhaps the most distinctive topographical feature in the otherwise rather flat landscapes of Kuwait is the Jal Al-Zor Escarpment, which runs along the northern shore of Kuwait bay and rises to 145 m above sea level. A part of it which extends from the east to west has been protected from grazing by a fence for more than 20 years (Figure 1.2).

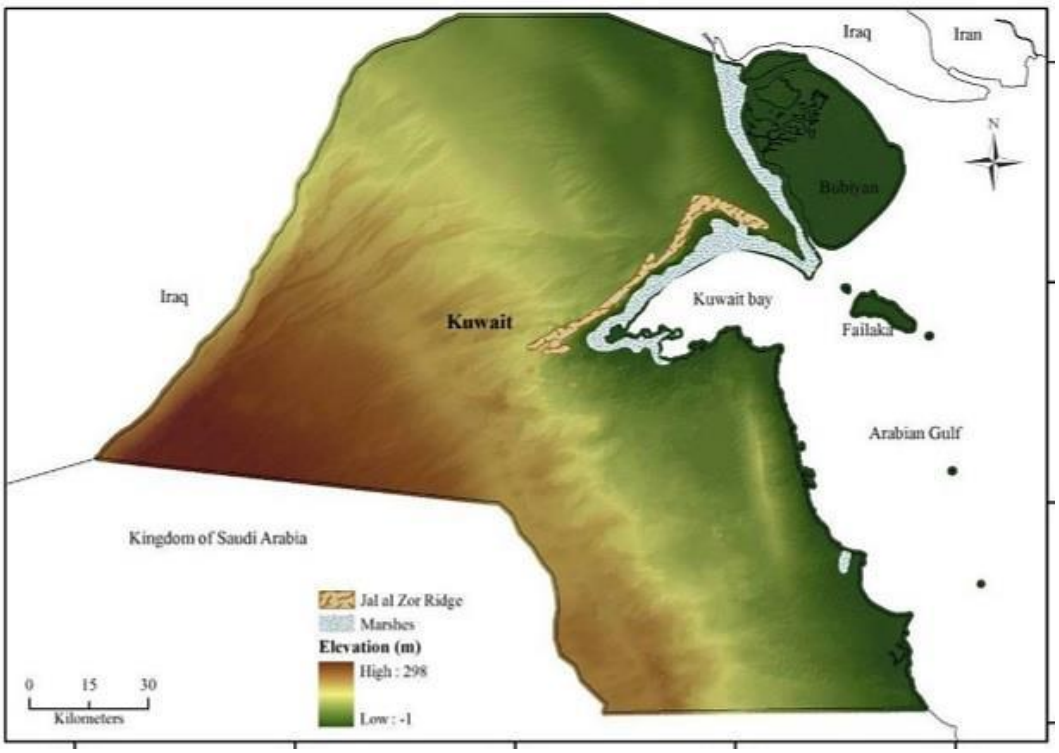


Figure 1.2 Topographical features of Kuwait *(source: Al-Shehabi, 2012)*

Kuwait's surface geological formations belong to the Neogene and Quaternary, and consist of clastic deposits with large amounts of limestone and gypsum. Two main formations represent the solid geology of Kuwait: the Ghar (from the Oligocene to lower Miocene periods) and the Fars (from the lower to middle Miocene periods). These rock formations are overlain by the Dibdibba formation (Al-Sulaimi and Pitty, 1995) which according to Ergun (1969) is divided into an upper layer composed of pebbles and cobbles derived from siliceous rocks and a lower layer of sandstones, which is exposed in the west and north-central Kuwait (Figure 1.3), and is dominated by calcareous sandstone, fine-grained limestone, and muddy sand with minor quantities of granules and scattered pebbles and gravel.

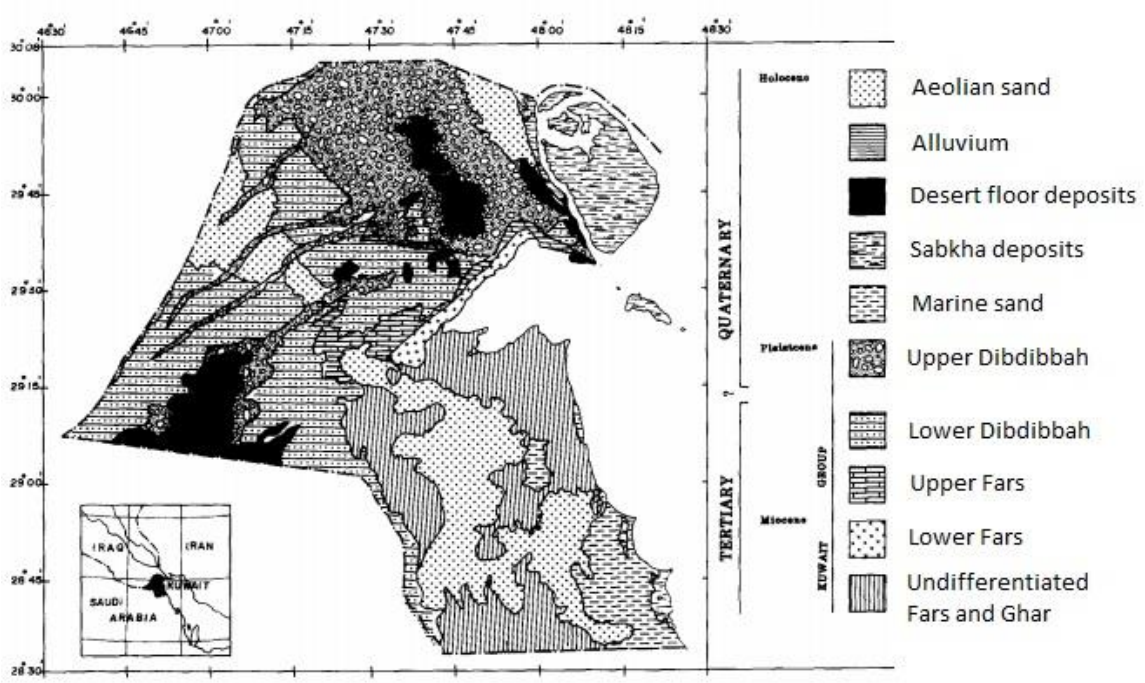


Figure 1.3 Geological map of Kuwait. *(source: Al-Sulaimi and Pitty, 1995)*

1.2.3 Soils

Kuwait is characterised by arid soils which are poor in organic matter, very low in moisture and shallow, varying from a few centimetres to up to two meters in depth. The dominant soil types are sandy, gravelly and intertidal muddy. The soil texture is classified as being mainly sandy or sandy to loamy (Ergun, 1969).

Due to the high evaporation rate and very low rainfall the levels of salinity are high at the upper layer of the soil surface. In general, vegetation is denser in areas with deep soils than in areas with shallow soils because of the larger storage capacity for water from precipitation in deeper soils. These deeper soils provide a continuous supply of moisture to perennial plants with deep root systems (Batanouny, 1983).

A detailed study of the soils in Kuwait was undertaken in 1999 by the Kuwait Institute for Scientific Research (KISR) and the Public Authority for Agriculture and Fisheries (PAAF) in collaboration with international consultants. The project covered the entire State of Kuwait. A reconnaissance survey was followed by sampling and field mapping, with guidance on the taxonomy and classification provided by the United States Department of Agriculture - Natural Resources Conservation Services (USDA-NRCS). The survey classified the soils of Kuwait into two orders: Aridisols occupying 70.8 % and Entisols occupying 29.2 % of the surveyed area (KISR, 1999). Aridisols or arid soils are mainly composed of calcium carbonate and characterised as being dry most of the year with limited leaching. Entisols are soils of recent origin. A detailed survey map of the soils of Kuwait (using classification to group level) is shown in Figure 1.4. Seven major soil groups are found: Torripsamments, Haplocalcids, Aquisalids, Calcigypsids, Petrocalcids, Petrogypsids, and Torriorthents, and one minor group - Haplogypsids soil. A brief description of each group is provided in Table 1.2 (Omar et al., 2001). The major types of soils are Petrogypsids (sandy to loamy soils overlying gypsic hardpans) and Torripsamments (deep sandy soils).

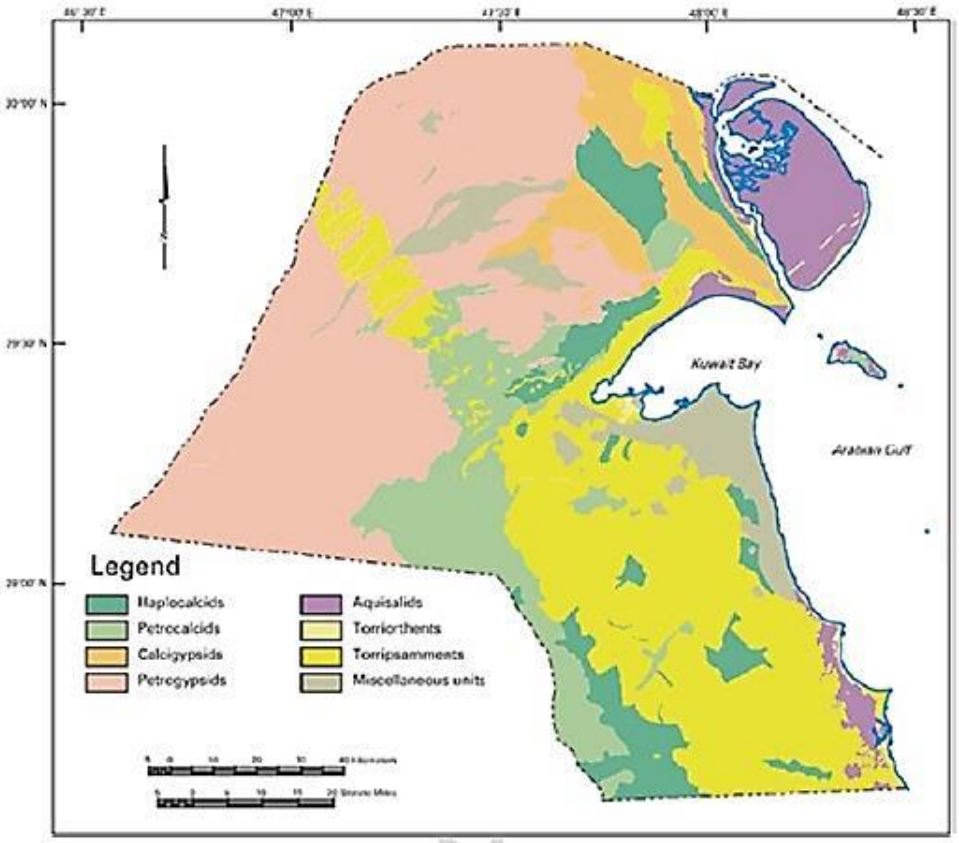


Figure 1.4 Distribution of soil groups in Kuwait *(source: Omar et al., 2001)*

Table 1.2 Description of soil groups in Kuwait

Map unit*	Area (%)	Description†
Torripsamments	27	Well to somewhat excessively drained, deep or very deep sandy soils
Petrocalcids	11	Well drained or moderately drained, shallow or moderately deep, sandy to loamy soils overlying a calcic hardpan. When upper soil is truncated, it may appear at surface
Haplocalcids	8	Well drained, deep or very deep, sandy to loamy soils, which have a layer of carbonate masses and nodules in the profile
Haplogypsids‡	<1	Well drained, deep or very deep, sandy to loamy soils, which have a layer of gypsum crystals in the profile
Aquisalids	7	Poorly or somewhat poorly drained, deep or very deep, sandy to clayey soils. Within the soil there is a layer of salt accumulation that usually occurs near the surface
Calcigypsids	6	Well drained, deep or very deep, sandy to loamy soils containing a layer of carbonate masses and nodules and a layer of gypsum crystals within the profile
Petrogypsids	33	Well drained, shallow or moderately deep, sandy to loamy soils overlying a gypsic hardpan. Hardpan may be exposed at surface, when upper soil is truncated
Torriorthents	1	Excessively drained to well drained, moderately deep or very deep, sandy soils. Within the soil profile there is a high content of shell fragments and some gypsum accumulations

1.3 Climate

Kuwait's climate is considered to be one of the aridest and hottest in the Middle-east. The four seasons are categorised as follows:

(1) Winter season, occurs from December to February (the first half is known as 'Murba'ania' and is extremely cold and second half is a period with milder temperatures). The winds during winter are predominantly from the north-western and bring cold air causing temperatures to drop to around 6 °C. The lowest temperature recorded was -4 ^{0}C in Kuwait City in January 1964.

(2) Spring season, occurs from February to May (It is split into an early period of moderate temperatures and ends with a warm period). The winds are predominantly warm and southerly.

(3) Summer season occurs from May to October. It starts with a transition period, which is followed by the dry, hot summer and ends with a period of high humidity. The summer is very hot and dry with midday temperature ranges from 42^{0}C to 46 °C. The highest temperature recorded was 53.8 C in Sulaibiyah on 31st July 2012.

(4) Autumn seasons spans about 1 month from November to December. The winds are predominantly from the south-east. November was the month with the highest rainfall (14.4 mm in 1997) recoded.

(Metrological data from the Kuwait Met Office, 2016 and UNFCCC, Kuwait report, 2012).

Relative humidity is highest in January (43 % to 84 %) and lowest in June (13 % to 40 %) (El-Sheikh et al., 2010). The Average evaporation rates vary from 3 to 13 mm per day with maximum levels ranging from 20 to 48 mm per day. The rainy season starts from November and ends in April. Rainfall is very low and sporadic varying from 75 to 150 mm a year with an average rainfall of 116 mm a year. Minimum annual levels have been recorded as low as 31 mm while the maximum reached was 242 mm.

Dust and sandstorms are common during the summer season, particularly during the month of July with wind speeds of up to 50 km/hr and reaching a maximum speed of 100 km/hr (Kuwait Meteorological Centre - KMC, 2016).

Recent climate change research concludes that climate change is already impacting on biodiversity and that immediate action is needed (UNFCCC, 2012). Based on the UNFCCC report on climate changes in Kuwait (2012), temperatures measured over a 48 years period (1962-2010) had an average annual temperature of 26.1^{0}C and had risen by 1.6^{0}C. The future predictions for average annual temperatures in Kuwait continue to increase from the current average annual temperature of 26.1^{0}C to 28.7^{0}C by 2035 (an increase of 2.6^{0}C) (UNFCCC, 2012).

1.4 Phytogeography

To understand the present-day phytogeography of Kuwait, it is important to take into account the floristic origins and vegetational history of the region going back over geological times. At the start of the Miocene period, the area of present-day Kuwait lay at the bottom of a large ocean known as Tethys. During the Miocene (≈ 25 million years ago) the most important Paleo-geographic event in the region occurred: the splitting of the African Plate from the Arabian Shield (Hijaz mountains) and the formation of the Red Sea occurred (Watts and Al-Nafie, 2003). The separation of these two plates played a major role in providing various conditions for plant growth and mixing elements of the Palaeotropical floras (from the south) and Holarctic floras (from the north). These events have a major influence on the distribution of the current Arabian flora (Miller and Cope, 1996; Ghazanfar and Fisher, 1998).

Zohary's (1973) researches on the classification and origins of the vegetation of the Middle East have dominated the interpretation of the phytogeography of the region for almost 50 years. He divided Arabia into two floral regions: the north and central regions belonged to the Saharo-Sindian and the southern region to the Sudano-Deccan (Zohary, 1973). The first major change to this interpretation was when White and

Leonard (1991) extended their interpretation and classification of African vegetation to cover SW Asia by revising Zohary's concept of the Saharo-Sindian region regarding both its eastern limits and its boundaries in the Arabian Peninsula (White and Leonard, 1991). According to White and Leonard (1991) Kuwait lays in the Saharo-Arabian region (represented as SS2 on Figure 1.5) with its vegetation being similar to that of the northern part of Arabian Peninsula but with significant elements from Irano-Turanian region to the north also being present e.g. *Astragalus tribuloides, Achillea fragrantissima, Arnebia decumbens, Artemisia herba-alba, Allium sindjarensis, Plantago ovata, Bassia indica* and *Bienertia cycloptera* (Miller and Cope, 1996; Ghazanfar and Fisher, 1998; Al-Nafie, 2008), and a small but very interesting tropical elements e.g. *Aizoon canariense, Caylusea hexagyna, Dinebra retroflexa, Halophila ovalis, Halodule uninervis, Hibiscus trionum, Leptochloa fusca, Panicum antidotale* (Al-Nafie, 2008).

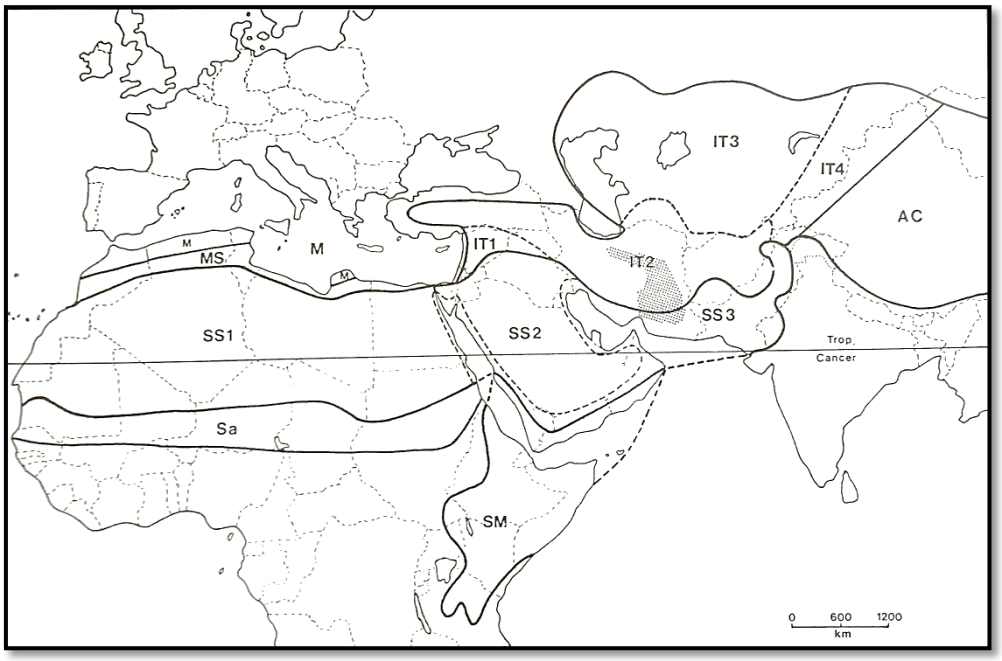


Figure 1.5. Main phytogeographical regions covering Asia and Africa (White and Leonard, 1991). Regional zones are represented by: AC= Central Asiatic; SS= Saharo-Sindian, Sa= Sahel, SM=Somalia-Masai; M= Mediterranean; MS = Mediterranean-Sahara; IT = Irano-Turanian.

1.5 Vegetation of Kuwait

The native vegetation of Kuwait is of high values as it represents semi-arid and arid vegetation of the region and contains highly adaptable genes to the harsh environment, extreme drought, poor soil and organic matter and tolerates high salinity (Zaman et al., 2009).

The vegetation types of Kuwait are mainly defined by geomorphological changes, soil types and climatic conditions. Four main vegetation types are recognised in Kuwait: (1) sand dunes vegetation, (2) salt marshes 'sabkha' and saline vegetation, (3) desert plains vegetation (4) and desert plateau vegetation (Halwagy and Halwagy, 1974).

Sand dunes vegetation: comprises a series of low coastal dunes which extends along the southern to northern coastal strips. The soil is composed of loose coarse sand. It is usually dominated by *Zygophyllum qatarense* and/or *Seidlitzia rosmarinus*, with occasional by *Atriplex leucoclada* and *Nitraria retusa*. Common associates are *Lycium shawii* and *Pennisetum divisum*.

Salt marshes 'sabkha' and saline vegetation: The marshes are influenced by tidal action and the shallow saline water table. Soil ranges from loamy sand to sandy clay. This vegetation type dominates in areas of Kuwait bay coasts and Khor Al-Sabiyah. It is also found in Bubiyan and Warba Islands. Chenopodiaceae dominates the vegetation with *Halocnemum strobilaceum* common near the shore and *Nitraria retusa* and *Zygophyllum qatarensis* further inland.

Desert plains vegetation: this type dominates most of the desert land area (west of the coastal region) and is represented by a number of communities:

a. **Cyperus steppe**: dominated by *Cyperus conglomeratus* and common to the south and south-west of Kuwait City. Commonly associated with *Panicum turgidum*. The soil is of deep, moderately loose, coarse sand.
b. **Rhanterium steppe**: dominated by *Rhanterium epapposum* and found in the central and north-east of Kuwait. Common associates are

Convolvulus oxyphyllus, *Moltkiopsis ciliata* and *Stipagrostis plumosa*. Found on shallow to moderately deep soils with a calcareous hardpan.

c. **Haloxylon steppe**: dominated by *Haloxylon salicornicum* found mainly in northern areas of Kuwait. Found on shallow soils with a hardpan.

The desert plateau vegetation: Found mainly in the extreme west of Kuwait; dominated by *Haloxylon salicornicum* with *Citrullus colocynthis* also occuring frequently. Annual plants densely cover areas where perennial plants are absent. These include *Arnebia* spp. *Helianthemum* spp. *Astragalus* spp. and *Schismus barbatus. Zilla spinosa* dominates Wadi Al-Batin area. Found on thin soils consisting of generally of few inches of coarse or soft loamy sand, often mixed with pebbles and gravel, over a hardpan.

1.5.1 Distribution of the vegetation

The distribution of vegetation across Kuwait has received attention from several authors (Dickson, 1955; Halwagy and Halwagy, 1974; Omar et al., 2001). Earlier, Dickson (1955) published an illustrated map of Kuwait's vegetation showing four potential plant communities widely distributed across Kuwait: (1) *Haloxylon* 'Hamdh' shrubland, (2) *Rhanterium* 'Arfaj' shrubland, (3) *Panicum* 'Thammam' grass-shrubland, and (4) *Cyperus* 'Thunda' sedges. Later, Halwagy and Halwagy (1974) produced a map recognising five plant communities, the four mentioned by Dickson (1955), and with the addition *Zygophyllum qatarensis* community. Their map shows the *Haloxylon* community is the most widely distributed across Kuwait, followed by the *Rhanterium* community (Figure 1.6).

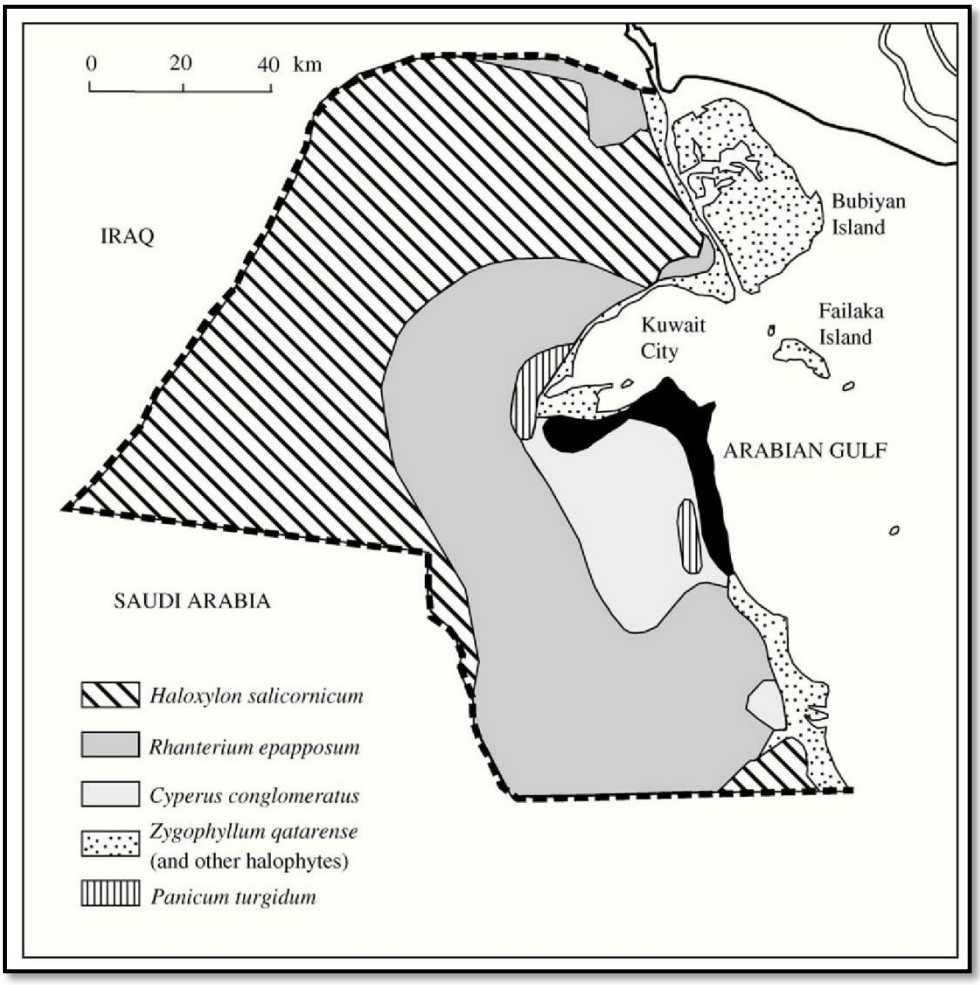


Figure 1.6 Vegetation map of Kuwait showing five plant communities
(Source: Halwagy and Halwagy, 1974 ed. Brown, 2003)

1.5.2 Present vegetation

Recently, Omar et al. (2001), using modern Geographical Information system (GIS), produced a map using integrated soil and vegetation data. Their survey covered the entire land area of Kuwait and involved a total of 8351 survey points covering measurements of vegetation and soil (Figure 1.7). The resulting map illustrates eight vegetation associations: Centropodietum, Cyperetum, Halophyletum, Haloxyletum, Panicetum, Rhanterietum, Stipagrostietum, and Zygophylletum which are described as follows:

1. **Haloxyletum unit** dominated by *Haloxylon salicornicum* and associated with shrubs such as *Astragalus spinosus* and *Chrozophora* spp.
2. **Rhantereitum unit** dominated by *Rhanterium epapposum* in association with *Convolvulus oxyphyllus* and *Moltkiopsis ciliata*.
3. **Cypertum unit** dominated by *Cyperus conglomeratus* and associated with *Astragalus annularis, Brassica tournefortii* and *Plantago boissieri*.
4. **Stipagrostietum unit** dominated by *Stipagrostis plumosa* and associated with *M. ciliata, P. boissieri* and *Stipa capensis*.
5. **Zygophylletum unit** dominated by *Zygophyllum qatarense* and associated with *Salsola imbricata, Cressa critica* and *Aizoanthemum hispanicum*.
6. **Centropodietum unit** dominated by *Centropodia forsskalii* and usually associated with *Stipagrostis plumosa*.
7. **Panicetum unit** dominated by *Panicum turgidum* and associated with *Aeluropus lagopoides* and *Pennisetum divisum*.
8. **Halophyletum unit** includes many halophytic communities and dominated by Chenopodiaceae with *Tamarix aucheriana, Nitraria retusa, Halocnemum strobilaceum,* and *Seidlitzia rosmarinus* occurring commonly (Omar et al., 2001).

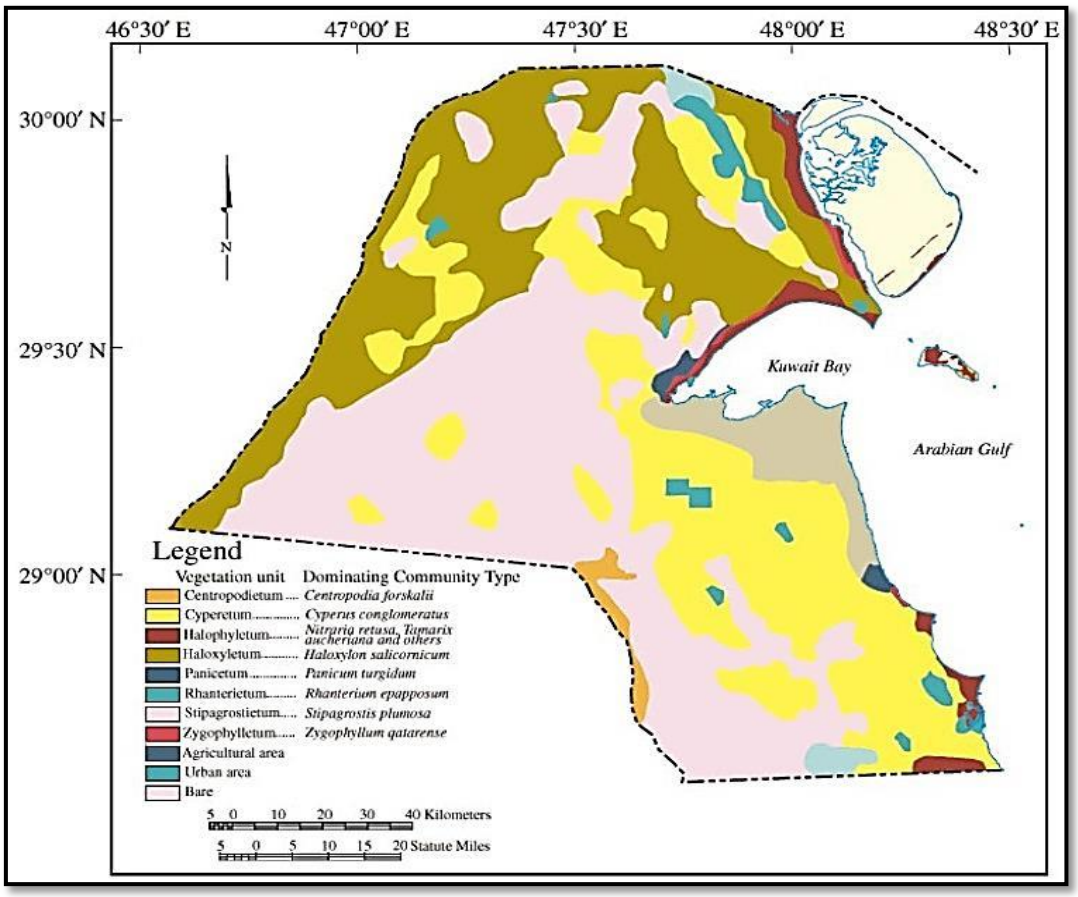


Figure 1.7. Current vegetation map of Kuwait (source: Omar et al., 2001)

1.5.2.1 Changes to plant communities

Omar et al. (2001) analysed the current vegetation distribution and compared it with Halwagy and Halwagy (1974) which shows significant changes to the vegetation cover (Omar et al., 2001). They concluded that over 25 years, some plant communities decreased in distribution, while others had increased (Figure 1.7). The following changes were recorded:

1. *Cyperus conglomeratus* community increased from 10 % to 27 %, with potential to expand over time.
2. *Rhanterium epapposum* community decreased from 30 % to 2 %, indicating that this community has considerably retreated from the open desert areas due to overgrazing.

3. *Haloxylon salicornicum* community decreased from 52 % to 23 %, and
4. *Zygophyllum qatarense* community decreased from 4 % to 0.3 %. (Omar et al., 2001).

The current status of the vegetation has shown intensive land degradation and decline of plant communities, e.g. *Rhanterium* spp. and *Haloxylon* spp., while grasses and sedges increased e.g. *Stipagrostis* spp. and *Cyperus* spp. The main changes in vegetation distribution were caused mainly by environmental factors (e.g. climate change and drought) and human interference such as the uprooting of shrubs, gravel quarrying, spring camping, off-road driving, livestock over-grazing and military activities (Al-Awadhi et al., 2001; Omar, 2007; Brown, 2003).

1.6 Conservation and threats

The desert of Kuwait is under severe pressure and heading towards extreme desertification. The United Nations Environment Programme (UNEP) has defined desertification as 'land degradation in arid, semiarid and dry sub-humid areas, resulting mainly from adverse human impact and partially from climatic factors' (UNCCD, 1994). In Kuwait, several interacting factors increasing land degradation and mainly caused by anthropogenic activities, overgrazing, and other natural processes.

After Gulf War I, due to war related damages, Kuwait was rewarded compensation for environmental related damages on five major claims to be spent towards the restoration and remediation of land and coastal areas, a total amount of $3,003,666,082 USD under decision No. 258 on 8th Dec 2005 (UNCC, 2005). The amount is to be spent on five major claims concerning the remediation of damages to underground water resources, areas damages by military activities, damages at open burning and open detonation sites, damages caused by oil (such as oil lakes, oil

contaminated piles, oil trenches, and oil spills) and the revegetation of damages occurred to the desert environment.

It is important to spend the funds on the rehabilitation of the degrading desert ecosystem and conserve the biodiversity. In Kuwait, areas where most of the above ground vegetation is lost or removed, a rehabilitation programme is required. The government of Kuwait (who owns the land), policy makers, and other scientific institutes should choose between natural regeneration and active restoration. Natural regeneration involves the removal of disturbing factors that cause land degradation and allowing natural regeneration of vegetation to maintain the cover and stabilize the soil. Active restoration involves the cultivation of the same type of vegetation that existed pre-disturbance (Ott et al., 2011).

1.6.2 Natural regeneration of vegetation

Natural regeneration is considered inexpensive but often slow and requires favourable environmental conditions (such as continuous good rainy seasons). In Kuwait, several studies indicate that natural regeneration and recovery is possible. Brown and Al-Mazrooei (2003) reported that rapid vegetation regeneration was possible after only four years of fence protection, avoiding grazers and human activities, in a severely degraded *Rhanterium epapposum* community in northern Kuwait (Brown and Al-Mazrooei, 2003).

Another example of natural recovery is the demilitarised zone (DMZ), located across the border line between Kuwait and Iraq. A preliminary field survey took place in June 2009 to study plant and soil conditions inside the DMZ and compare it with conditions in the open rangeland (simply grazed and disturbed), revealed that by avoiding drivers of disturbance a total of 74 plant species (annuals and perennials) were recorded and the vegetation cover varied from 35-47 % across the DMZ with *Haloxylon salicornicum* community being dominant (Kuwait National Focal Point, unpublished data).

Such findings suggest that by avoiding grazers and human disturbance it is possible to regenerate vegetation and maintain its coverage over a period of time. Factors that affect the natural recovery rates of a disturbed desert ecosystem are: 1) Climatic

conditions and soil compaction, 2) frequency, severity, size and type of disturbance, 3) vascular plant structure, and 4) the recovery of nitrogen and carbon fixation after disturbance (Webb et al., 2009).

1.6.3 Active restoration of vegetation

Active restoration is generally expensive and often unsuccessful (Webb et al., 2009). It can be implemented in several ways depending on the scale of the project and the availability of funds, tools (e.g. irrigation system) and access to facilities (for seed germination) (Weigand & Rodgers, 2009).

If the aim of active restoration is to revegetate the same type of plants that once existed pre-disturbance, then it should include a seed collection strategy and consider studies on seed germination for mass production and implementation in large-scale restoration programmes.

Desert seeds are difficult to germinate, and each type has its optimal favourable conditions to break its dormancy. In an attempt to break desert seeds dormancy Zaman et al. (2006) studied the effect of different salinity levels on the germination of *H. salicornicum*, *Z. qatarense* and *Tamarix aucheriana* seeds. The study revealed that the seed germination showed different tolerance to high salinity concentration and other techniques needed to be applied such as mechanical scarification and dry heat treatment was necessary for the germination of some desert seeds (Zaman et al., 2006).

Developing methods for cell and tissue culture is necessary for germplasm conservation to ensure the survival of endangered valuable plant species and possible save time for rapid mass propagation and revegetation in large scale programmes. *Rhanterium epapposum*, *Gynandiris sisyrinchium*, *Haloxylon salicornium*, *Lycium shawii*, *Nitraria retusa*, *Ochradenus baccatus* and *Gypsophila capillaris*, were selected, collected and tissue cultured because of their importance and potential use in desert revegetation (AboEl-Nil, 1997).

To achieve a successful restoration programme the choice of plants and their function is important. Choosing plants that can tolerate high salinity, temperature and drought

would increase the success rate of the restoration method. Also, including plants able to stabilise the soil by their rooting system; in return lowering soil erosion and possibly reduces the amount of dust storms. Some examples of potential plants for restoration projects in Kuwait may include *Rhanterium epapposum, Haloxylon salicornicum, Nitraria retusa, Lycium shawii, Tamarix aucheriana, Helocnemum strobiliaceum, Seidlitzia rosmarinum, Calligonum comosum, Frasetia aegyptiaca, Panicum turgidum, Pennisetum divisum* and *Cyperus conglomeratus.*

1.6.4 Biological soil crusts

An important natural process that should be considered while studying patterns of vegetation is the formation of biological soil crusts (also known as microbiotic soil crusts). This topic lacks proper documentation in Kuwait and the Arabian Peninsula although it has been well documented in other regions (Eldridge , 2000; Schulz et al., 2016). Biological soil crusts are the association of a variety of unrelated microorganisms in the soil surface and are represented by cyanobacteria, algae, fungi and lichens (West, 1990). Biological soil crusts possess an important effect on vascular plants in the top soil layer by retaining soil moisture for longer periods of time, increase nutrient uptake through nitrogen fixation and reduces wind and water erosion of soil surface (Belnap et al., 2001). Vascular plants growing on biological soil crusts showed higher concentrations of nutrients (organic matter and nitrogen) than plants growing without soil crusts (Harper and Pendleton, 1993). However, others reports on biological soil crusts showed an inhibiting effect on the germination of vascular plants (Prasse and Bornkamm, 2000). Li et al. (2002) showed that soil crust development leads to a change in vegetation from shrubs to herbs due to decreased soil moisture reaching deeper soil layers affected by the formation of soil crusts enriched with mosses, algae, and liverworts (Li et al., 2002).

This biological association of microorganisms leading to the formation of soil crusts, need to be identified and studied in details to understand the effect of the “living” crust on vascular plants (Belnap et al., 2001). It’s difficult to study microorganisms in the field when there is a lack of specialists, e.g. bryologists and phycologists. This

suggests potentially productive area of research in the Arabian Peninsula with the aid of DNA barcoding techniques to identify the various microorganisms.

1.6.5 Conservation through protected areas

A protected area is defined by IUCN (2008) as 'a clearly defined geographical space, recognised, dedicated and managed, through legal or other effective means, to achieve the long term conservation of nature with associated ecosystem services and cultural values' (IUCN Definition 2008).

Due to land degradation, loss of vegetation and the damage caused by human activities, it is considered a priority to conserve and protect the biodiversity of Kuwait in protected areas. The current and future (proposed) protected areas of Kuwait are shown in Figure 1.8. The main protected areas in Kuwait include Sabah Al-Ahmed Nature Reserve (SSNR), the Al-Jahra Pool Reserve, the Doha Nature Reserve, and the Sulaibiya Field Station (SFS).

1. **Sabah Al Ahmed Nature Reserve (SSNR)** Previously known as the Jal Az-Zor National Park was established in 1991 and is the largest protected area in Kuwait. SSNR is strategically located north of Kuwait bay and covers an area of 330 km^2. The protected area includes excellent examples of two vegetation types: salt marshes 'sabkha' and the desert vegetation due to its. It represents several geomorphological features such as Jal Al-Zor escarpments and Wadi Um Ar-Rimam depression. After 14 years of protection, a field survey by El-Sheikh and Abbadi (2004) was conducted on the vegetation within SSNR. This survey recorded 139 species belonging to 32 families which is an estimate of 30 % of the entire flora of Kuwait. The SSNR protected area adds great value to the conservation and protection of native vegetation and has good examples of natural regeneration.
2. **Al-Jahra Nature Reserve** was established in 1987. It is located at the western end of Kuwait Bay and covers an area of 3.5 km^2. The reserve

contains a human-made pond formed from treated sewage water flowing across sandy ‘sabkha’ into the sea (Omar, 2007). The reserve is dominated by halophytes, such as *Tamarix* sp. and *Suaeda aegyptiaca,* which provide a natural shelter for many migratory birds, appropriate microclimates and breeding and feeding resource (Al-Saqer , 2003).

3. **Doha Nature Reserve** was established in 1988. It is located on the south side of Kuwait Bay and covers an area of 4.5 km^2. The reserve supports salt tolerant vegetation and is used by migratory birds for breeding and feeding (Al-Tamimi , 2010).
4. **Sulaibiya Field Station** also known as Kabd Station was established in 1979. It is located south-west of Kuwait City and occupies an area of 40 km^2. The vegetation is dominated by the perennial dwarf shrub *Rhanterium epapposum* associated with *Cyperus conglomeratus*. A good flush of annuals also appears in each rainy season with such species as *Plantago boissieri*, *Schimpera arabica*, *Cutandia memphitica*, *Lotus halophilus*, and *Horwoodia dicksoniae* (Omar, 2007).

Figure 1.8 Current and proposed protected areas of Kuwait *(source: EPA Kuwait, 2013)*
Green: current/ adapted protected areas, Light green: future proposed areas

1.7 Natural resources and land management

For many generations, Bedouins inhibited the Arabian deserts. It provides them with an important source of food, medicinal plants and forages for livestock. They continuously travelled across the desert areas looking for food and water, known as desert dwellers. They managed to survive the hot climates and sand storms and developed their skills to make use of the wildlife. Oasis and wells dug underground are the primary source of fresh drinking water. Experienced people identified the presence of underground water by certain plants such as *Haloxylon salicoricum* known to dominate shallow soils with easy reach of water (Mandaville, 2011). The main source of food comes from raising livestock, agriculture and catching fish. Animals mainly, sheep and goats were used for food, wool and dairy products. Camels used as the main source for long distance transportation across the desert. They also used skilled falcons to hunt for birds and other wildlife animals.

The desert plants also provide a nutritious source of food summarised below adopted from Mandaville (2011) with his experience and collection of information from locals for more than 15 years working in the Arabian Peninsula (Mandaville, 2011).

Roots, tubers and bulbs: *Dicadi erythraeum* (for bulbs juicy and sweet), *Emex spinosa* (taproot sweet carrot-like), *Allium sphaerocephalum* (wild onions, green leaves are eaten raw), *Allium sindjarense* (edible bulbs), *Gagea reticulate* (bulb consumed raw), *Calligonum comosum* (edible stems), *Cynomorium coccineum* (parasitic plant, fleshy parts edible), and *Orobanchaceae* spp. (eaten baked).

Green parts eaten raw: mainly leaves eaten raw *Anisosciadium lanatum*, *Aaronsohnia factorovskyi*, *Rumex pictus*, *Rumex vesicarius*, *Sisybrium irio*, *Launaea capitate*, *Leptaleum filifolium*, *Senecio glaucus*, *Erodium* spp. (used for spicing food), *Schimpera arabica*, and *Scabiosa palaestina*.

Fruits and flowers: *Nitraria retusa*, *Lycium shawii*, *Salvadora persica*, *Neurada procumbens*, and *Ziziphus spina-christi*.

Seeds and grains: *Mesembryanthemum forskahlii*, *M. nodiflorum*, *Aizoon canariense*, and *Panicum turgidum.*

Gums and other Exudates: *Convolvulus oxyphyllus* (chewing gum) and *Haloxylon salicornicum* (sweet exudates).

Desert Truffles: underground fungus 'fagaa' known as Bedouins wild food, *Tirmania nivea*, and *T. pinoyi*.

The Bedouins were alerted by livestock and camels to avoid toxic and noxious plants. Most common plants used as forage are *Rhanterium epapposum, Stipagrostis plumosa, Salsola* spp., *Traganum* spp. also other salty shrubs were grazed by camels such as *Haloxylon* spp, *Halocnemum strobilaceum*, *Suaeda* spp., and *Seidlitzia rosmarinus*.

The desert also provide a valuable resources of traditional medicinal plants extracted from leaves, seeds, roots, or even entire plants e.g. teeth cleanser and mouth hygiene, *Salvadora persica* and *Pistacia lentiscus* (Ghazanfar and Al-Sabahi, 1993; Saganuwan, 2010), for abdominal and digestive problems *Thymus serphyllum*, and *Ziziphus vulgaris* (Nawash and Al-Haroni, 2011) and for fever and chest pain remedies *Ziziphus spina-christi* (Ghazanfar and Al-Sabahi, 1993).

From ancient medicine to current clinical based medicine e.g. *Tephrosia apolllinea*, *Curcuma longa*, *Zingiber officinalis*, *Vitis vinifera*, and *Nerium oleander* are used in bronchitis disease (Ghazanfar and Al-Sabahi, 1993). *Haplophylum tuberculatum*, *Pulicaria crispa*, *Ononis serrata*, and *Achillea beiberstenii* are used in the control of cancer diseases (Kuete et al., 2013). *Calligonum comosum*, *Rumex pictus*, *Euphorbia cuneate*, and *Chrozophora oblongifolia* are used for their antimicrobial activities (Rahman et al., 2004).

After the survival of the Bedouins for many generations in the desert and evolving with the available natural resources comes the blooming of the oil industry. The settlement changed from living in tents and simple housing created from mud and clay until the present day large modern houses and tall buildings made of massive amounts of steel and concrete which require plenty of building materials e.g. sand, cement and gravel excavation, adding more pressure to the land. Population increase resulted in more demand for livestock which increased land misuse and led to many of the flora and fauna extinction.

Kuwait's land use and biodiversity have been affected dramatically, mainly by environmental conditions, overgrazing and human activities due to its small size and easy to travel around by off-road vehicles. Figure 1.9 represents an image of the past and the present biodiversity of Kuwait, published by Al Arabi Magazine and presented by Mr. Saleh Al Mesbah (Omar, 2007) describing the biodiversity situation of early 1960's until the current day after the blooming of oil industry. The figure is reflecting the wildlife in Kuwait showing sheep and camel grazing the land and gazelles. Foxes and wolves all over the desert area, sadly the gazelles become extinct. Also, reptiles, birds and marine life reflecting the past biodiversity with present (Figure 1.9). In addition, Mr. Al Mesbah described several oil platforms polluting the environment with flares and smoke expressing the present day pollution (Figure 1.9).

An increase in land degradation also occurred in Kuwait during Gulf War I and II (1991 and 2003) due to the military activities and heavy machinery which added a massive amount of pressure on the desert's ecosystem. As a result of Gulf War I, a total of 608 oil wells were damaged and set on fire which caused public health problems mainly asthma and bronchitis diseases. The formation of dry and wet oil lakes covering huge areas of the desert and penetrated through the lower layers of soil still exists and continues to pollute the environment. Also, minefields and unexploded munitions (about 20,000 units) left over and cleared at a later date (Al-Damkhi, 2007).

Figure 1.9 Past and present land use in Kuwait.
(Published by Al Arabi Magazine, collection of Mr. Saleh Al-Misbah, Omar, 2007)

Chapter 2 Checklist of the flora of Kuwait

This chapter summarises the current status of knowledge of the plants of Kuwait. As a prerequisite to applying the molecular methods discussed in the following chapters, it is necessary to generate a checklist with revised orders and family names in line with Angiosperm Phylogeny Group (APG IV, 2016) and provide an update on nomenclatures to understand the status of the flora. The following revision includes orders and family names, nomenclatures, authorities, plant status and a floristic analysis. Plant status refers to each species of the flora as being either native (establishes naturally in the wild), cultivated (by humans) or introduced as weeds (could not establish it's life cycle in the wild). The floristic analyses will include the grouping of plants at different taxonomic levels (i.e. order, family, genera and species level).

For the purpose of conserving the knowledge of the plant biodiversity of the flora of Kuwait, it is very important to clarify the status of all species and to distinguish between natives, those non-natives that have become established in the wild, weedy and cultivated species.

2.1 Current knowledge of the flora of Kuwait

Boulos and Al-Dosari published the most recent checklist of the flora of Kuwait (1994) based on a review of more than 20 publications including Daoud, (1985), Al-Rawi (1987), Dickson and Macksad (1973), Halwagy and Macksad (1972). In this publication 374 species in 55 families are recorded. The largest families are Gramineae (Poaceae) (70 species), Compositae (Asteraceae) (47 species), Cruciferae (Brassicaceae) (34 species), Leguminosae (Fabaceae) (29 species), Amaranthaceae (including Chenopodiaceae) (27 species) and Caryophyllaceae (20 species). The largest genera are *Astragalus* (9 species), *Plantago* (8 species), *Bromus* (5 species) *Erodium* (5 species) *Helianthemum* (4 species) and Launea (4 species).

An analysis of life-forms show that there are 256 annual species; 83 herbaceous perennials; 34 under-shrubs and shrubs, and only one tree, *Vachellia gerrardii* (=*Acacia pachyceras*) which is locally known as 'Talha' (Boulos and Al-Dosari, 1994; Daoud, 1985; Al-Rawi, 1987). This single specimen of *Vachellia* survives today inside a fenced protected area, Sabah Al-Ahmad Nature Reserve (SSNR), which lies in the north-east of Kuwait (Figure 2.1).

Figure 2.1 *Vachellia gerrardii* **'Talha'** (Source: *M Abdullah,* 16-02-2012)

The outstanding contribution to the flora of Kuwait to date is the two volumes of the Flora of Kuwait published by Daoud (1985) and Al-Rawi (1987). Volume one covers the dicots in part (Daoud, 1985) and volume two covers Compositae and the monocots (Al-Rawi, 1987). Following the two volumes, Boulos (1988) published 'The Weed Flora of Kuwait' which included 84 species of weedy plants.

2.1.2 Botanical Exploration of Kuwait

A biographical index of botanists and plant collectors in the Arabian Peninsula has been compiled and presented by Blatter (1913-1936) and Wickens (1982). These include 20 individuals who have made collections in Kuwait. Table 2.1 shows a list of collector's names, year of collection and herbarium where their collections are deposited (Table. 2.1). Sadly, the collections deposited at Kuwait University Herbarium (KTUH) were looted at the time of the Iraqi invasion during the first Gulf War in 1990 and had not been recovered. Before the invasion, KTUH held about 22,000 collections covering the entire Arabian Peninsula. In the years following the invasion extensive field surveys, covering the whole country, were initiated and now the herbarium holds more than 15,500 well-documented specimens covering most species recorded from Kuwait and collected during the period from 1991 to the present day.

The earliest botanical collection from Kuwait is believed to have been made by Lewis Pelly with William Colvill (1865) who collected 60 specimens from their journey to Kuwait. Their collections were published by Kew Gardens (Burtt and Lewis, 1949). Sir Percy Cox made another collection known to include 100 specimens from Kuwait with Lt. Col. S. G. Knox and Lt. Col. Stuart George in the Jal Al-Zor Hills in about 1907 and was later published by Gilbert Carter (1917).

The largest collections were made by Violet Penelope Dickson, who collected about 600 herbarium specimens from Kuwait and neighbouring countries and sent them to Kew Gardens where Mr. A.R. Horwood made determinations. Dickson published 'The Wild Flowers of Kuwait and Bahrain' (1955). She lists 270 species from Kuwait and includes a number of maps and illustrations (Dickson, 1955).

Table 2.1 List of botanists and plant collectors made from Kuwait

Collectors name	Expedition Year	Herbarium (acronym)
Lt. Col. Sir Lewis Pelly with William Henry Colvill	1865	Royal Botanic Gardens (K)
Gen. Sir Percy Cox with Lt. Col S.G. Knox, Lt. Col. Stuart George	1907	The Natural History Museum (BM)
William Henry Irvine Shakespeare	1914	The Natural History Museum (BM)
Hazem Sulaiman Daoud	1930-1976	Cairo University
Violet Penelope Dickson	1935-65	Royal Botanic Gardens (K)
T. Wilson	1935	Royal Botanic Gardens (K)
Henry Field	1950	Peabody Museum
A.L. Temple	1950	The Natural History Museum (BM)
Assad Macksad	1961-1970	Kuwait University (KTUH)
Malcolm Dixon Kernick	1961-65	Kuwait University (KTUH)
Mrs Katherine J. Macintyre	1971	Royal Botanic Gardens (E)
Loutfy Boulos	1973-95	Royal Botanic Gardens (K)
Mohammed Nazir Sankary	1976-77	Royal Botanic Gardens (E)
Mohammed Halwagy	1967-70	Royal Botanic Gardens (K), Kuwait University (KTUH)
Ali Al-Rawi	1979-89	Royal Botanic Gardens (E) and (K), Geneva Herbaria catalogue (G)
Peter Show Green	1981	Royal Botanic Gardens (E) and (K)
Ian Charleson Hedge	1981	Royal Botanic Gardens (E) and (K)
Samira Omar	1982-	Kuwait University (KTUH)

2.2 Methods

The checklist is based on a review of all the major publications which have contributed towards the flora of Kuwait and the results of recent fieldwork. It provides an up to date nomenclature, authorities, synonyms and plant status. Classification of the orders and families follows Angiosperm Phylogeny Group (APG IV, 2016). Species names to follow largely the accepted names and synonyms as treated under Kew's World Checklist series (http://apps.kew.org/wcsp/), Tropicos (www.tropicos.org) and The Plant List (www.theplantlist.org). For some taxa, the Euro+Med PlantBase - the information resource for Euro-Mediterranean plant diversity (http://ww2.bgbm.org/EuroPlusMed/) and International Legume Database and Information Services – ILDIS (http://www.ildis.org/). For Salsoleae (Amaranthaceae) using Akhani et al. (2015). For families of monocotyledonous plants, accepted names and synonyms as given in eMonocot (http://e-monocot.org/). Also, other publications from surrounding regions have been consulted to bring regional consistency to the taxonomy. Author names format will be cited following Brummitt and Powell index (1992). The checklist covers all species of vascular plants (flowering plants and ferns) recorded growing in the State of Kuwait. Including all native plants, introduced plants (intentionally or accidentally by a human) which are now naturalised. The list also includes weeds and commonly cultivated species that are not yet known to be widely naturalised but might become so in the future. The checklist excludes crops and ornamental species since they are mainly restricted to agricultural farms and plant nurseries for economical production. The following numbered references [1-14] are referred to in the checklist (Table 2.2).

Table 2.2 Major references referred to in the checklist of the flora of Kuwait

Referral No.	References
1	Halwagy, R. and Macksad, A. (1972) A contribution towards a Flora of the State of Kuwait and the Neutral Zone. *Bot. J. Linn. Soc.*, 65: 61-79.
2	Daoud, H. S. (1985) *Flora of Kuwait vol. 1: Dicotyledoneae*. London, KPI and Kuwait University. Pg. 224.
3	Al-Rawi, A. (1987) *Flora of Kuwait vol. 2: Compositae and Monocotyledoneae*. Kuwait University. Pg. 225-455
4	Boulos, L. (1988) *The Weed flora of Kuwait*. Kuwait University pg. 175
5	Boulos, L. and Al-Dosari, M. (1994) Checklist of the flora of Kuwait. *J. Univ. Kuwait (Sci.)* 21: 203-217
6	Mathew, K.T. et al. (2012) Eleven new weeds in Kuwait. *Kuwait J. Sci. Eng.* 39 (1A) pp. 169-192
7	Miller, A.G. and Cope, T.A. (1996) *Flora of the Arabian Peninsula and Socotra*. Vol. 1, Edinburgh University Press. Pp 586
8	Cope, T.A. (2007) *Flora of the Arabian Peninsula and Socotra* vol. 5. Edinburgh University Press. Pp 387
9	Akhani, H. (2015) Plants and Vegetation of North-West Persian Gulf. University of Tehran Press. Pp 502
10	World Checklist of Selected Plant Families. (2016) Facilitated by the Royal Botanic Gardens, Kew. (online: http://apps.kew.org/wcsp/)
11	Tropicos (2016) Missouri Botanical Garden. (online: www.tropicos.org)
12	The Plant List (2013). Version 1.1 Published on the internet (www.plantlist.org)
13	Euro+Med PlantBase (2006) The information resources for Euro-Mediterranean plant diversity (online: http://ww2.bgbm.org/EuroPlusMed/)
14	eMonocot (2017). A web-based treatment for monocot plants of the world. (http://e-monocot.org/)

Several questions should be raised when dealing with alien species: 1) whether the taxon is native or alien to the region? 2) What is its position in the invasion process? 3) What is the degree of its naturalisation and possible invasion?

To understand the plant status of native origin, invasive aliens, and naturalised plants, the following definitions were adopted from Pysek et al. (2004):

Native plants: Taxa found established in the wild at a geographical area for many years without altering its condition and without the involvement or intervention of humans.

Naturalised plants: Alien plants that adapt in the wild and capable of independent growth without human involvement.

Weeds or invasive aliens: Taxa in a given area arrived or introduced and could not establish its growth without aid and its presence restricted to a limited environment (i.e. receives water by irrigation, located near agricultural areas, gardens, etc.).

Naturalised weed by cultivation: Taxa introduced by the involvement of human outside its normal geographical range and later able to adapt in the wild and undergoes its life cycles without any aid.

Cultivated plants: Taxa introduced by human involvement and requires continuous source of water and fertilisation.

In the checklist the following categories of plant status have been used: native (N), naturalised weed (NW), naturalised weed by cultivation (NWC), weeds (W) or cultivated plants (C). This part of categorising plant status may be irrelevant for the purpose of DNA barcoding the flora of Kuwait (which contains all the plants that are present regionally) but certainly considered in conservation studies and restoration ecology projects.

2.3 Results

The following checklist is arranged alphabetically by the family names with angiosperms first followed by a single gymnosperm (*Ephedra alata*) and a single fern (*Ophioglossum polyphyllum*) (Table 2.3). The currently accepted scientific names for each species, with their authority citations followed by synonyms (which have been used in Kuwaiti publications) are listed in Table 2.3. Numbers inside the square brackets refer to references listed in Table 2.2 which provide the source of the record. The collectors' number and herbarium where the specimen is deposited is given for new records, not covered in the references. Finally, the current plant status is represented by the following abbreviations: N, NW, NWC, C and W (Table 2.3).

Table 2.3. Checklist of the flora of Kuwait

MAGNOLIOPHYTA [ANGIOSPERMAE]	
ACANTHACEAE	
Avicennia marina (Forssk.) Vierh. [MTA329[E]]	**C***
AIZOACEAE	
Aizoanthemum hispanicum (L.) H.E.K. Hartmann [12]. Syn. Aizoon hispanicum L. [2,5,7]	**N**
Aizoon canariense L. [2,5,7]	**N**
Mesembryanthemum nodiflorum L. [1,2,5,7]	**N**
AMARANTHACEAE [CHENOPODIACEAE]	
Aerva javanica (Burm.f.) Juss. Ex J.A. Schult. [6,7]	**W**
Agathophora iraqensis Botsch. [11, 12, 13]. Syn. Agathophora alopecuroide (Delile) Fenzl ex Bunge [5,7]; Halogeton alopecuroides Moq. [13]	**N**
Amaranthus graecizans L. [4,5,7]. Syn. Amaranthus angustiloius Lam. [4]	**W**
Amaranthus hybridus L. [5,7]	**W**
Amaranthus lividus L. [4,5,7]	**W**
Anabasis lachnantha Aellen & Rech.f. [5,7]	**N**
Anabasis setifera Moq. [1,2,5,7]	**N**
Atriplex dimorphostegia Kar. & Kir. [2,5,7]	**N**
Atriplex leucoclada Boiss. var. inamoena (Aellen) Zohary [1,2,5,7,11,13]. Syn. Atriplex inamoena Aellen [13]	**N**
Bassia eriophora (Schrad.) Asch. [1,4,5,7]	**W**
Bassia muricata (L.) Asch. [1,2,4,5,7]	**NW**
Bassia scoparia (L.) A.J. Scott [5,7]. Syn.Kochia scoparia (L.) Schrad. [5]	**W**
Beta vulgaris L. [4,5,7]	**W**
Bienertia sinuspersici Akhani [9,11,12]. Syn. Bienertia cycloptera auctt. non Bunge ex. Boiss. [1,2,5,7]	**N**
Caroxylon cyclophyllum (Baker) Akhani & Roalson [13]. Syn.Salsola	**N**

cyclophylla Baker [1,5,7]	
Caroxylon imbricatum (Forssk,) Akhani & Roalson [11,12]. Syn. Salsola imbricata Forssk. [5,7]; Chenopodium baryosmum Roem. & Schult. [11,12]; Salsola baryosma (Roem. & Schult.) Dandy [2]	**N**
Caroxylon jordanicola (Eig.) Akhani & Roalson [11,12,13]. Syn. Salsola jordanicola Eig [1,2,5,7]	**N**
Chenopodiastrum murale (L.) S.Fuentes, Uotila & Borsch [11,13]. Syn. Chenopodium murale L. [2,5,7]	**NW**
Chenopodium album L. [4,5,7]	**W**
Chenopodium ficifolium Sm. [6]	**W**
Chenopodium glaucum L. [5,7]	**W**
Chenopodium opulifolium Schrad. ex Koch & Ziz [5,7]	**W**
Cornulaca aucheri Moq. [2,5,7]. Syn. Cornulaca leucacantha Charif & Aellen [2]	**N**
Cornulaca monacantha Delile [5,7]	**N**
Halocnemum strobilaceum (Pall.) M. Beib. [2,5,7]	**N**
Halopeplis perfoliata (Forssk.) Asch. & Bunge [11,12]	**N**
Halothamnus iraqensis Botsch. [5,7]	**N**
Hammada salicornica (Moq.) Iljin [5,11,13]. Syn. Caroxylon salicornicum Moq. [11,12,13]; Haloxylon salicornicum (Moq.) Bunge ex Boiss. [2,5,7].	**N**
Salicornia perennans Willd. [11,12,13]. Syn. Salicornia europaea auct. non L. [2,5,7]; Salicornia herbacea L. [2,11,12,13]	**N**
Seidlitzia rosmarinus Bunge ex Boiss. [2,5,7]	**N**
Suaeda aegyptiaca (Hasselq.) Zohary [5,7]. Syn. Schanginia aegyptiaca (Hasselq.) Allen [1,2]	**N**
Suaeda vermiculata Forssk. Ex J.F. Gmel. [2,5,7]. Syn. Suaeda fruticosa Forssk. ex J.F. Gmel. [5]	**N**
Traganum nudatum Delile [1,2,5,7]	**N**
AMARYLLIDACEAE	
Allium longisepalum Bertol. [11,12,13,14]. Syn. Allium laceratum Boiss. & Noe. [11,12,13,14]	**N**
Allium sindjarense Boiss. & Hausskn. ex Regel [3,5,14]	**N**
Allium sphaerocephalum L. [3,5,12]	**N**
APIACEAE [UMBELLIFERAE]	
Ammi majus L. [4,5]	**W**
Anisosciadium isosciadium Bornm. [5]	**N**
Anisosciadium lanatum Boiss. [2,5]	**N**
Bupleurum semicompositum L. [2,5]	**N**
Ducrosia anethifolia (DC.) Boiss. [5]	**N**
Pituranthos triradiatus (Hochst.) Asch. & Schweinf. [11,13]. Syn. Deverra triradiata Hochst. ex Boiss. [5,11,13].	**N**
APOCYNACEAE	
Calotropis procera (Aiton) W. T. Aiton [5]	**N**
ASPARAGACEAE	
Bellevalia saviczii Woronow [5]	**N**
Dipcadi erythraeum Webb & Berth. [3,5]	**N**
ASPHODELACEAE	
Asphodelus tenuifolius Cav. [3,5]	**N**
Asphodelus viscidulus Boiss. [1,3,5]	**N**

ASTERACEAE [COMPOSITAE]

Aaronsohnia factorovskyi Warb. & Eig [3,5]	**N**
Acantholepis orientalis Less. [3,5]	**W**
Achillea fragrantissima (Forssk.) Sch. Bip. [3,5]	**N**
Anthemis melampodina Delile subsp. Deserti Eig. [12,13]. Syn. Anthemis deserti Boiss. [3,5,13]	**N**
Anthemis pseudocotula Boiss [3,4,5]	**NW**
Anvillea garcinii (Burm.f.) DC. [3,5]	**N**
Artemisia sieberi Besser [11,12,13]. Syn. Seriphidium sieberi (Besser) K. Bermer & Humphries ex Y.R.Ling [13]; Artemisia herba-alba Asso var. laxiflora Boiss. [3,5,13]	**N**
Artemisia scoparia Waldst. & Kit. [1,3,5]	**N**
Aster squamatus (Spreng.) Hieron. [4,5]	**NW**
Atractylis cancellata L. [3,5]	**N**
Atractylis carduus (Forssk.) C. Chr. [3,5]	**N**
Calendula arvensis L. [3,4,5]	**NW**
Calendula tripterocarpa Rupr. [5]	**N**
Carduus pycnocephalus L. [3,5]	**N**
Carthamus oxyacantha M. Bieb. [3,5]	**N**
Centaurea bruguierana (DC.) Hand.-Mazz. [3,5]	**N**
Centaurea mesopotamica Bornm. [5]	**N**
Centaurea pseudosinaica Czerep. [3,5]	**N**
Chrysanthemum coronarium L. [5]	**W**
Cichorium endivia L. [5]	**W**
Erigeron bonariensis L. [12,12,13]. Syn. Conyza bonariensis (L.) Cronquist [3,4,5]	**W**
Echinops polyceras Boiss. [11,12,13]. Syn. Echinops blancheanus Boiss. [3,5]	**N**
Filago pyramidata L. [3,5]	**N**
Flaveria trinervia (Spreng.) C. Mohr [5]	**W**
Gnaphalium uliginosum L. [6]	**W**
Gymnarrhena micrantha Desf. [3,5]	**N**
Ifloga spicata (Forssk.) Sch. Bip. [3,5]	**N**
Koelpinia linearis Pall. [3,5]	**N**
Lactuta serriola L. [5]	**W**
Launaea angustifolia (Desf.) Kuntze [1,3,5]	**N**
Launaea capitata (Spreng.) Dandy [1,3,5]	**N**
Launaea mucronata (Forssk.) Muschl. [3,5]	**NW**
Launaea nudicaulis (L.) Hook.f. [3,5]	**N**
Leontodon laciniatus (Bertol.) Widder [1,3,5]	**N**
Matricaria aurea (Loefl.) Sch. Bip. [3,5]	**W**
Picris babylonica Hand.-Mazz. [3,5]	**N**
Pallenis hierochuntica (Michon) Greuter [11,12]. Syn. Asteriscus hierochunticus (Michon) Wiklund [5]; Asteriscus pygmaeus (DC.) Coss. & Durieu [1,3]	**N**
Pulicaria undulata (L.) C.A. Mey. [5]. Syn. Pulicaria crispa (Forssk.) Oliv. [3,5]	**N**
Reichardia tingitana (L.) Roth [3,5]. Syn. Picridim tingitanum (L.) Desf. [1]	**N**
Rhanterium epapposum Oliv. [3,5]	**N**
Scorzonera papposa DC. [3,5]	**N**
Scorzonera tortuosissima Boiss. [3,5]	**N**
Senecio glaucus subsp. Coronopifolius (Maire) Alexander [11,12,13]. Syn. Senecio desfontainei Druce [1,3]	**NW**
Senecio vulgaris L. [5]	**W**
Sonchus oleraceus L. [3,4,5]	**W**

Sonchus tenerrimus L. [3,5]	**W**
Urospermum picroides (L.) F.W. Schmidt [4,5]	**W**
Xanthium strumarium L. [5]	**W**
BORAGINACEAE	
Arnebia decumbens (Vent.) Coss. & Kralik [2,5]	**N**
Arnebia linearifolia DC. [5]	**N**
Arnebia tinctoria Forssk. [5]. syn. Arnebia tetrastigma Forssk. [2]	**N**
Echium angustifolium Mill. Subsp. Serceum (Vahl) Klotz. [5,11,12,13]	**N**
Gastrocotyle hispida (Forssk.) Bunge [11,12,13] Syn. Anchusa hispida Forssk. [5,11,12]	**N**
Heliotropium bacciferum Forssk. [5,11,12,13]. Syn. Heliotropium ramosissimum (Lehm.) DC. [12,13].	**N**
Heliotropium kotschyi (Bunge) Gürke [5]	**N**
Heliotropium europaeum L. [12,13]. Syn. Heliotropium ellipticum Ledeb. [12]; Heliotropium lasiocarpum Fisch. & C.A.Mey. [5]	**W**
Lappula spinocarpos (Forssk.) Asch., Verhandl. [1,2,5]	**N**
Moltkiopsis ciliata (Forssk.) I.M. Johnst. [1,2,5]	**N**
Neatostema apulum (L.) I.M. Johnst. [11,12,13]. Syn. Myosotis apula L. [13]; Lithospermum apulum (L.) Vahl. [13].	**N**
Ogastemma pusillum (Coss. & Durieu ex Bonnet & Barratte) Brummitt [5,11,12,13]. Syn. Megastoma pusillum Coss. & Durieu ex Bonnet & Barratte [5,11,13]	**N**
BRASSICACEAE [CRUCIFERAE]	
Alyssum homalocarpum (Fisch. & Mey.) Boiss. [2,5,7]	**N**
Alyssum linifolium Steph. Ex Willd. [2,5,7]	**N**
Anastatica hierochuntica L. [1,2,5,7]	**N**
Brassica juncea (L.) Czern. & Coss. [2,5,7]	**C**
Brassica tournefortii Gouan [2,4,5,7]	**NW**
Cakile arabica Velen. & Bornm. [1,2,5,7]	**N**
Capsells bursa-pastoris (L.) Medik. Pflan. [6]	**W**
Cardaria draba (L.) Desv. [4,7]	**W**
Carrichtera annua (L.) DC. [2,5,7]	**N**
Coronopus didymus (L.) Sm. [4,5,7]	**NW**
Descurainia sophia (L.) Webb ex Prantl [5,7]	**W**
Diplotaxis acris (Forssk.) Boiss. [1,2,5,7]	**N**
Diplotaxis harra (Forssk.) Boiss. [1,2,5,7]	**N**
Eremobium aegyptiacum (Spreng.) Asch. & Schwienf. Ex Boiss. [5]	**N**
Eruca sativa Mill. [2,4,5,7].	**NWC**
Farsetia aegyptia Turra [2,5,7]	**N**
Farsetia burtoniae Oliv. [2,5,7]	**N**
Horwoodia dicksoniae Turrill [1,2,5,7]	**N**
Lepedium sativum L. [4,5,7]	**C**
Lepidium aucheri Boiss. [2,5,7]	**N**
Leptaleum filifolium (Willd.) DC. [2,5,7]	**N**
Malcolmia africana (L.) R. Br. [5,7]	**W**
Malcolmia grandiflora (Bunge) Kuntze [5,7]	**N**
Maresia pygmaea (Delile) O.E. Schultz [2,7]. Syn. Malcolmia pygmaea (Delile) Boiss.[5]	**N**
Matthiola longipetala (Vent.) DC. [2,5,7]	**N**
Neotorularia torulosa (Desf.) Hedge & J.Léonard [5,7]. Syn. Torularia torulosa (Desf.) O.E. Schultz [2]	**N**
Notoceras bicorne (Aiton) Amo [1,2,5,7]	**N**

Savignya parviflora (Delile) Webb [1,2,5,7]	**N**
Schimpera arabica Hochst. & Steud. ex Boiss. [2,5,7]	**N**
Sinapis arvensis L. [4,5,7]	**W**
Sisymbrium erysimoides Desf. [2,5,7]	**W**
Sisymbrium irio L. [2,4,5,7]	**N**
Sisymbrium orientale L. [2,4,5,7]	**N**
Sisymbrium septulatum DC. [2,5,7]	**N**
Zilla spinosa (Turra) Prantl. [2,5]	**N**
CAPRIFOLIACEAE	
Lomelosia olivieri (Coult.) Greuter & Burdet [11,12]. Syn.Scabiosa olivieri Coult. [1,2,5,11,12]	**N**
Lomelosia palaestina (L.) Raf. [11,12]. Syn.Scabiosa palaestina L. [2,5,11,12]	**N**
Valerianella dufresnia Bunge ex. Boiss. [2,5]	**N**
CARYOPHYLLACEAE	
Gymnocarpos sclerocephalus (Decne.) Thulin. [11,12]. Syn. Sclerocephalus arabicus Boiss [1,2,5,7,12]; Sclerocephalus aucheri Walp. [12].	**N**
Gypsophila capillaris (Forssk.) C. Chr. [2,5,7]. Syn. Gypsophila antari Post & Beauverd [1]	**N**
Herniaria hemistemon J. Gay [2,5,7]	**N**
Herniaria hirsuta L. [2,5,7]	**N**
Loeflingia hispanica L. [2,5,7]	**N**
Paronychia arabica (L.) DC. [2,5,7]	**N**
Polycarpaea repens (Forssk.) Asch. & Schweinf. [2,5,7]	**N**
Polycarpaea robbairea (Kuntze) Greuter & Burdet [5,7]. Syn. Robbairea delileana Milne-Redhead [1,2]	**N**
Polycarpon tetraphyllum (L.) L. [2,5,7]	**N**
Pteranthus dichotomus Forssk. [2,5,7]	**N**
Silene arabica Boiss. [1,2,5,7]	**N**
Silene arenosa C. Koch [5,7]	**N**
Silene conoidea L. [4,5,7]	**W**
Silene villosa Forssk. [1,2,5,7]	**N**
Spergula fallax (Lowe) H.L. Krause [1,2,5,7]	**N**
Spergularia diandra (Guss.) Boiss. [7]. Syn. Spergularia diandra (Guss.) Heldr. & Sart. [1,2,5]	**N**
Spergularia marina (L.) Besler [5,7]. Syn. Spergularia marina (L.) Griseb. [5]	**N**
Stellaria media (L.) Vill. [4,7]	**W**
Telephium sphaerospermum Boiss. [5]	**N**
Vaccaria hispanica (Mill.) Rauschert [4,5,7]. Syn.Vaccaria pyramidata Medik. [1]	**NW**
CISTACEAE	
Helianthemum kahiricum Delile [2,5]	**N**
Helianthemum ledifolium (L.) Mill. [2,5]	**N**
Helianthemum lippii (L.) Dum. Cours. [2,5]	**N**
Helianthemum salicifolium (L.) Mill. [2,5]	**N**
CONVOLVULACEAE	
Convolvulus arvensis L. [4,5]	**W**
Convolvulus cephalopodus Boiss. [5]. Syn. Convolvulus buschiricus Bornm. [1,2,5]	**N**
Convolvulus oxyphyllus Boiss. [2,5]	**N**
Convolvulus pilosellifolius Desr. [1,2,4,5]	**NW**

Convolvulus prostratus Forssk. [11,12]. Syn. Convolvulus microphyllus Sieb. Ex Spreng. [11,12]; Convolvulus desertii Hochst. & Steud [11,12]	**NW**
Cressa cretica L. [2,5]	**N**
Cuscuta planiflora Ten. [1,2,5]	**N**
CRASSULACEAE	
Crassula alata (Viv.) A. Berger [5,7]	**N**
CUCURBITACEAE	
Citrullus colocynthis (L.) Schard. [2,5] Cucumis colocynthis L. [2]	**N**
CYMODOCEACEAE	
Halodule uninervis (Forssk.) Asch. [5]	**N**
CYNOMORIACEAE	
Cynomorium coccineum L. [2,5]	**N**
CYPERACEAE	
Cyperus conglomeratus Rottb. [5] Cyperus aucheri Jaub. & Spach [3]	**N**
Cyperus rotundus L. [4,5]	**W**
EUPHORBIACEAE	
Chrozophora tinctoria (L.) Raf. [5] Chrozophora verbascifolia (Willd.) A. Juss.[1,2]	**N**
Euphorbia densa Schrenk [1,2,5]	**N**
Euphorbia granulata Forssk. [1,2,5]	**N**
Euphorbia grossheimii (Prokh.) Prokh. [5]. Syn. Euphorbia isthmia Täckh [2]	**W**
Euphorbia helioscopia L. [4,5]	**W**
Euphorbia hirta L. [4,5]	**W**
Euphorbia indica Lam. [4,5]	**W**
Euphorbia peplus L. [4,5]	**W**
Euphorbia serpens Kunth [5]	**N**
FABACEAE [LEGUMINOSAE]	
Alhagi graecorum Boiss. [5]. Syn. Alhagi maurorum Medik. [2]	**N**
Astragalus annularis Forssk. [2,5]	**N**
Astragalus arpilobus Kar. & Kir. Subsp. Hauarensis (Boiss.) Podlech [11,12] Syn. Astragalus hauarensis Boiss. [2,5]	**N**
Astragalus bombycinus Boiss. [2,5]	**N**
Astragalus corrugatus Bertol. [1,2,5,12]. Syn. Astragalus tenuirugis Boiss. [12]	**N**
Astragalus dactylocarpus Boiss. [1,2]	**N**
Astragalus hamosus L. [MTA280[E]]	**N**
Astragalus schimperi Boiss. [2,5]	**N**
Astragalus sieberi DC. [5]	**N**
Astragalus spinosus (Forssk.) Muschl. [2,5]	**N**
Astragalus tribuloides Delile [2,5]	**N**
Coronilla scorpioides (L.) Koch [5]	**W**
Hippocrepis areolata Desv. [5]. Syn. Hippocrepsis bicontorta Loisel. [2]	**N**
Hippocrepis unisiliquosa L. [2,5]	**N**
Lathyrus aphaca L. [6]	**W**
Lotus halophilus Boiss. & Spruner [2,5]	**N**
Medicago laciniata (L.) Mill. [2,5]	**N**
Medicago polymorpha L. [4,5]	**W**
Medicago rotata Boiss. [5]	**NW**

Medicago sativa L. [MTA077[E]]	**C**
Melilotus alba Medik. [6]	**W**
Melilotus indica (L.) All. [1,2,5]. Syn. Melilotus parviflora Desf. [2]	**N**
Onobrychis ptolemaica (Delile) DC. [2,5]	**N**
Ononis reclinata L. [5]	**N**
Ononis serrata Forssk. [2,5]	**N**
Prosopis fracta (Banks & Sol.) J.F.Macbr. [5]	**C***
Prosopis juliflora (Sw.) DC. [5]	**C***
Scopiurus muricatus L. [4,5]	**W**
Trifolium lappaceum L. [5]	**W**
Trifolium resupinatum L. [5]	**W**
Trigonella anguina Delile [2,5]	**N**
Trigonella hamosa L. [2,5]	**N**
Trigonella stellata Forssk. [2,5]	**N**
Vachellia gerrardii (Benth.) P.J.H. Hurter [11,12]. Syn. Acacia pachyceras O. Schwartz [5,12]	**N**
Vicia sativa L. [5]	**C**
FRANKENIACEAE	
Frankenia pulverulenta L. [2,5]	**N**
GERANIACEAE	
Erodium bryoniifolium Boiss. [2,5]	**N**
Erodium ciconium (L.) L'Hér. [2,5]	**N**
Erodium cicutarium (L.) L'Hér. [2,5]	**N**
Erodium glaucophyllum (L.) L'Hér. [1,2,5]	**N**
Erodium laciniatum (Cav.) Willd. [1,2,5]	**N**
Monsonia nivea (Decne.) Decne ex Webb [2,5]	**N**
HYDROCHARITACEAE	
Halophila ovalis (R. Br.) Hook.f. [5]	**N**
IRIDACEAE	
Gladiolus italicus Mill. [3,4,5]	**NW**
Moraea sisyrinchium (L.) Ker Gawl. [12,13]. Syn. Gynandriris sisyrinchium (L.) Parl. [3,5];. Iris sisyrinchium L. [3,5]	**N**
IXIOLIRIACEAE	
Ixiolirion tataricum (Pall.) Schult. & Schult. f. [3,5]	**N**
JUNCACEAE	
Juncus rigidus Desf. [3,5]. Syn. Juncus arabicus (Asch. & Buch.) Adams [1]	**N**
LAMIACEAE [LABIATAE]	
Lallemantia royleana (Benth.) Benth. [5]	**W**
Salvia aegyptiaca L. [2,5]	**N**
Salvia lanigera Poir. [2,5]	**N**
Salvia spinosa L. [2,5]	**N**
Teucrium oliverianum Ging. Ex Benth. [1,2,5]	**N**
Teucrium polium L. [1,2,5]	**N**
LILIACEAE	
Gagea reticulata (Pall.) Schult. & Schult.f. [3,5]	**N**

MALVACEAE	
Althaea ludwigii L. [2,5]	**N**
Hibiscus trionum L. [6]	**W**
Malva nicaeensis All. [5]	**W**
Malva parviflora L. [2,4,5]	**N**
NITRARIACEAE	
Nitraria retusa (Forssk.) Asch. [2,5]	**N**
Peganum harmala L. [1,2,5]	**N**
NEURADACEAE	
Neurada procumbens L. [2,5,13]	**N**
NYCTAGINACEAE	
Boerhavia diffusa L. [6,7]	**W**
OROBANCHACEAE	
Cistanche phelypaea (L.) Cout. [1,2,5]. Syn. Lathraea phelypaea L [12,13]; Orobanche tinctoria Forssk. [12,13]; Cistanche tinctoria (Forssk.) Beck [12,13]; Cistanche tubulosa (Schenk) Hook.f. [1,2,5]	**N**
Orobanche aegyptiaca Pers. [2,5]	**N**
Orobanche cernua Loefl. [1,2,5]	**N**
Orobanche minor Sm. [5]	**W**
Orobanche ramosa L. [4,5]	**W**
OXALIDACEAE	
Oxalis corniculata L. [4,5]	**W**
PAPAVERACEAE	
Glaucium corniculatum (L.) Rudolph [5,7]	**N**
Papaver rhoeas L. [5,7]	**C**
Roemeria hybrida (L.) DC. [2,5,7]	**N**
Fumaria parvifiora Lam. [4,5,7]	**NW**
Hypecoum littorale Wulfen [5]. Syn. Hypecoum geslinii Coss. & Kralik [2,5,7]	**N**
Hypecoum pendulum L. [2,5,7]	**N**
PHYLLANTHACEAE	
Andrachne telephioides L. [2,5]	**N**
PLANTAGINACEAE	
Linaria albifrons (Sibth. & Sm.) Spreng. [5,12,13]. Syn. Antirrhinum albifrons Sm. [13]	**N**
Linaria simplex Desf. [5,13] Syn. Antirrhinum simplex Wild. [13];	**N**
Plantago amplexicaulis Cav. [2,5]	**N**
Plantago boissieri Hausskn. & Bornm. [2,5]	**N**
Plantago ciliata Desf. [2,5]	**N**
Plantago coronopus L. [2,5]	**N**
Plantago lanceolata L. [4,5]	**N**
Plantago notata Lag. [2,5]	**N**
Plantago ovata Forssk. [2,5]	**N**
Plantago psammophila Agnew & Chal.-Kabi [2,5]	**N**

PLUMBAGINACEAE	
Limonium carnosum (Boiss.) Kuntze [2,5]	**N**
Limonium lobatum (L.f.) Kuntze [2,5,12,13]. Syn. Statice thouninii Viv. [12,13]; Limonium thouini (Viv.) Kuntze [2,5,12,13]	**N**
Psylliostachys spicatus (Willd.) Nevski [2,5]. Syn. Statice spicata Willd. [2,5]	**N**
POACEAE [GRAMINEAE]	
Aegilops bicornis (Forssk.) Jaub. & Spach. [3,5,8,14]	**N**
Aegilops kotschyi Boiss. [3,5,8,14]	**N**
Aegilops triuncialis L. [3,5,8,14]	**N**
Aeluropus lagopoides (L.) Trin. ex Thwaites [1,3,5,8]	**N**
Aeluropus littoralis (Gouan) Parl. [1,3,5,8]	**N**
Ammochloa palaestina Boiss. [3,5,8]	**N**
Avena barbata Pott ex Link [3,4,5,8]	**NW**
Avena fatua L. [1,3,5,8]	**NW**
Avena sativa L. [5,8]	**W**
Avena sterilis L. [4,5,8]	**NW**
Brachypodium distachyum (L.) P. Beauv. [5,8]. Syn. Trachynia distachya (L.) Link [1,3]	**N**
Bromus catharticus Vahl [5,8]	**W**
Bromus danthoniae Trin. [3,5,8]	**N**
Bromus madritensis L. [1,3,4,5,8]	**NW**
Bromus sericeus Drobov [3,4,5,8]	**NW**
Bromus tectorum L. [3,5,8]	**N**
Cenchrus ciliaris L. [3,5,8]	**NW**
Cenchrus setigerus Vahl [3,5,8]	**W**
Centropodia forskalii (Vahl) Cope [5,8] Syn. Asthenatherum forskalii (Vahl) Nevski [1,3]	**N**
Cutandia dichotoma (Forssk.) Batt. & Trab. [3,5,8]	**N**
Cutandia memphitica (Spreng.) K. Richt. [1,3,5,8]	**N**
Cymbopogon commutatus (Steud.) Stapf [5,8]. Syn. Cymbopogon parkeri Stapf [3]	**N**
Cynodon dactylon (L.) Pers. [3,4,5,8]	**NWC**
Dactyloctenium aegyptium (L.) Willd. [4,5,8]	**W**
Dactyloctenium aristatum Link [5,8]	**W**
Dichanthium annulatum (Forssk.) Stapf [3,4,5,8]	**NW**
Dichanthium foveolatum (Delile) Roberty [2,3,5,8]	**W**
Digitaria ciliaris (Retz.) Koeler [5,8]	**N**
Digitaria sanguinalis (L.) Scop. [4,5,8]	**W**
Dinebra retroflexa (Vahl) Panzer [5,8]	**N**
Echinochloa colona (L.) Link [3,4,5,8]	**W**
Eleusine indica (L.) Gaertn. [6,8]	**W**
Eragrostis barrelieri Daveau [4,5,8]	**NW**
Eragrostis minor Host [5,8]	**NW**
Eremopoa persica (Trin.) Rosch. [5,8]	**N**
Eremopyrum bonaepartis (Spreng.) Nevski [1,3,5,8]	**N**
Eremopyrum distans (C. Koch) Nevski [1,3,5,8]	**N**
Hordeum marinum Huds [3,5,8]	**N**
Hordeum murinum L. [5,8]	**N**
Imperata cylindrica (L.) Raeusch. [3,5,8]	**N**
Lasiurus scindicus Henrard [5,8]. Syn. Lasiurus hirsutus (Forssk.) Boiss. [1,3]	**N**
Leptochloa fusca (L.) Kunth [4,5,8]	**W**
Lolium multiflorum Lam. [5,8]	**N**

Lolium rigidum Guadin [1,3,5,8]	**N**
Lolium temulentum L. [1,3,4,5,8]	**NW**
Panicum antidotale Retz. [4,5,8]	**C**
Panicum turgidum Forssk. [3,5,8]	**N**
Parapholis incurva (L.) C.E. Hubb. [3,5,8]	**N**
Pennisetum divisum (J.F. Gmel.) Henrard [1,3,5,8]	**N**
Phalaris minor Retz. [1,3,4,5]	**W**
Phalaris paradoxa L. [1,5,8]	**W**
Phragmites australis (Cav.) Trin. ex Steud. [3,4,5,8]	**N**
Poa annua L. [4,5,8]	**W**
Poa infirma Kunth [3,4,5,8]	**W**
Poa siniaica Steud. [3,4,5,8]	**W**
Polypogon monspeliensis (L.) Desf. [1.3,4,5,8]	**NW**
Rhynchelytrum repens (Willd.) C.E. Hubb [3,5,8]	**W**
Rostraria cristata (L.) Tzvelev [5,8]. Syn. Lophochloa phleoides (Vill.) Rchb. [1,3]	**N**
Rostraria pumila (Desf.) Tzvelev [4,5,8] Syn. Lophochloa pumila (Desf.) Bor [1,3]	**N**
Schismus arabicus Nees [3,5,8]	**N**
Schismus barbatus (L.) Thell. [1,3,5,8]	**N**
Setaria verticillata (L.) P. Beauv. [4,5,8]	**W**
Setaria viridis (L.) P. Beauv. [3,5,8]	**W**
Sorghum halepense (L.) Pers. [4,5,8]	**W**
Sphenopus divaricatus (Gouan) Rchb. [3,8]	**N**
Sporobolus arabicus Boiss. [1,3,5,8]	**N**
Stipa capensis Thunb. [1,3,5,8]	**N**
Stipagrostis ciliata (Def.) de Winter [3,5,8]	**N**
Stipagrostis drarii (Täckh.) de Winter [5,8]	**N**
Stipagrostis obtusa (Delile) Nees [3,5,8]	**N**
Stipagrostis plumosa (L.) Munro ex T. Anders. [1,3,5,8]	**N**
Trisetaria linearis Forssk. [5,8]	**W**
POLYGONACEAE	
Calligonum comosum L'Hér. [2,7]. Syn. Calligonum polygonoides subsp. comosum (L'Hér.) Soskov [5]	**N**
Emex spinosus (L.) Campd. [1,2,4,5,7]	**N**
Polygonum argyrocoleum Steud. ex. Kunze [4,5]. Syn. Polygonum patulum M.Bieb. [4,5]	**W**
Rumex pictus Forssk. [2,5,7]	**N**
Rumex vesicarius L. [2,5,7]	**N**
PORTULACEAE	
Portulaca oleracea L. [5,7]	**W**
PRIMULACEAE	
Anagallis arvensis L. [4,5]	**W**
RANUNCULACEAE	
Adonis dentata Delile [2,5,7]	**N**
RESEDACEAE	
Caylusea hexagyna (Forssk.) M.L. [2,5,7]	**N**
Ochradenus baccatus Delile [2,5,7]	**N**
Oligomeris linifolia (Vahl) J.F. Macbr. [1,5,7]. Syn. Oligomeris subulata Webb. [7]	**N**

Reseda arabica Boiss. [2,5,7]	**N**
Reseda decursiva Forssk. [2,5]. Syn. Reseda alba L. [5]	**N**
Reseda muricata C. Presl [1,2,5,7]	**N**
RHAMNACEAE	
Ziziphus nummularia (Burm.f.) Wight & Arn. [KTM 5392 [KUTH]]	**N**
Ziziphus spina-christi (L.) Desf. [MTA567[E]]	**C***
RUBIACEAE	
Crucianella membranacea Boiss. [2,5]	**N**
Galium tricornutum Dandy [5]	**N**
RUTACEAE	
Haplophyllum tuberculatum (Forssk.) A. Juss. [1,2,5]	**N**
SALVADORACEAE	
Salvadora persica L. [MTA356[E]]	**C***
SCROPHULARIACEAE	
Scrophularia desertii Delile [1,2,5]	**N**
SOLANACEAE	
Datura innoxia Mill. [4,5]	**W**
Hyoscyamus muticus L. [5]	**N**
Hyoscyamus pusillus L. [2,5]	**N**
Lycium shawii Roem. & Schult. [1,2,5]	**N**
Physalis angulata L. [6]	**W**
Solanum nigrum L. [4,5]	**W**
Withania somnifera (L.) Dunal [5]	**W**
TAMARICACEAE	
Tamarix aphylla (L.) H.Karst. [MTA603[E]]	**C***
Tamarix aucheriana (Decne.) B. R. Baum [2,5]	**N**
THYMELAEACEAE	
Thymelaea mesopotamica (C. Jeffrey) B. Peterson [2,5]	**N**
TYPHACEAE	
Typha domingensis (Pers.) Poir. ex Steud. [4,5]. Syn. Typha angustifolia L. [3]	**N**
URITACEAE	
Urtica urens L. [4,5,7]	**W**
VERBENACEAE	
Phyla nodiflora (L.) Greene [4,5]	**W**
Verbena tenuisecta Briq. [6]	**W**
ZYGOPHYLLACEAE	
Fagonia bruguieri DC. [2,5]	**N**
Fagonia glutinosa Delile [2,5]	**N**
Fagonia indica Burm.f. [5]	**N**
Fagonia olivieri DC. [1,2]	**N**
Seetzenia lanata (Willd.) Bullock [2,5,12]. Syn. Seetzenia orientalis Decne. [2,5,12]	**N**

Tribulus macropterus Boiss. [5]	**W**
Tribulus terrestris L. [2,4,5]	**NW**
Tribulus pentandrus Forssk. []. Syn. Tribulus longipetalus Viv. []	**N**
Tetraena simplex (L.) Beier & Thulin. [12]. Syn. Zygophyllum simplex L. [2,5]	**N**
Tetraena qatarensis (Hadidi) Beier & Thulin [12]. Syn. Zygophyllum qatarense Hadidi [5]	**N**
GYMNOSPERMAE	
EPHEDRACEAE	
Ephedra alata Decne. [5,7]	**N**
PTERIDOPHYTA	
OPHIOGLOSSACEAE	
Ophioglossum polyphyllum A. Braun [5,7]. Syn. Ophioglossum aitchisonii (C.B.Cl.) J.D. Almeida [5]	**N**

Asterisk (*) cultivated plants introduced in restoration projects and along local streets.
Plant status: N- native, NW- naturalised weed, NWC- naturalised weed by cultivation, W- weeds, C- cultivated plants. **Source reference [1-8]:**

[1] Halwagy, R. & Macksad, A. (1972) A contribution towards a Flora of the State of Kuwait and the Neutral Zone. *Bot. J. Linn. Soc*., 65: 61-79.

[2] Daoud, H. S. (1985) *Flora of Kuwait vol. 1: Dicotyledoneae*. London, KPI and Kuwait University. Pg. 224.

[3] Al-Rawi, A. (1987) *Flora of Kuwait vol. 2: Compositae and Monocotyledoneae*. Kuwait University. Pg. 225-455

[4] Boulos, L. (1988) *The Weed flora of Kuwait. Kuwait University* pp. 175

[5] Boulos, L. & Al-Dosari, M. (1994) Checklist of the flora of Kuwait. *J. Univ. Kuwait (Sci.)* 21: 203-217.

[6] Mathew, K.T. et al. (2012) Eleven new weeds in Kuwait. *Kuwait J. Sci. Eng.* 39 (1A) pp. 169-192.

[7] Miller, A.G. & Cope, T.A. (1996) *Flora of the Arabian Peninsula and Socotra. vol. 1*, Edinburgh University Press. Pp 586.

[8] Cope, T.A. (2007) *Flora of the Arabian Peninsula and Socotra vol. 5*. Edinburgh University Press, Pp 387

[9] Akhani, H. (2015) Plants and Vegetation of North-West Persian Gulf. University of Tehran Press. Pp 502

[10] World Checklist of Selected Plant Families (2016). Facilitated by the Royal Botanic Gardens, Kew. (http://apps.kew.org/wcsp/)

[11] Tropicos (2016) Missouri Botanical Garden. (online: www.tropicos.org)

[12] The Plant List (2013). Version 1.1 Published on the internet (www.plantlist.org)

[13] Euro+Med PlantBase (2006) The information resources for Euro-Mediterranean Plant Diversity (online: http://ww2.bgbm.org/EuroPlusMed/)

[14] eMonocot (2017). A web-based treatment for monocot plants of the world. (online: http://e-monocot.org/)

2.3.1 Floristic analyses of the flora of Kuwait

The largest families in the flora are Poaceae [Gramineae] 72 spp, Asteraceae [Compositae] 48 spp, Brassicaceae [Cruciferae] and Fabaceae [Leguminosae] 35 spp, Amaranthaceae [Chenopodiaceae] 33 spp and Caryophyllaceae 20 spp (Table 2.3 and Figure 2.2).

The largest genera are *Astragalus* 10 spp (all native), *Euphorbia* 8 spp (3 native and 5 weeds), *Plantago* 8 spp (all native), *Bromus* 5 spp (2 native, 2 naturalised weeds and 1 weed), *Convolvulus* 5 spp (3 native and 1 weed), *Erodium* 5 spp (all native), *Avena* 4 spp (3 naturalised weed and 1 weed), *Chenopodium* 4 spp (1 naturalised weed and 4 weeds), *Helianthemum* 4 spp (all native), *Launaea* 4 spp (3 native and 1 naturalised weed), *Orobanche* 4 spp (2 native and 2 weeds), *Silene* 4 spp (3 native and 1 weed), *Sisymbrium* 4 spp (3 native and 1 weed) and *Stipagrostis* 4 spp (all native) (Table 2.3). Eighteen species of the flora have changes to their family names according to APG IV (2016) classification, listed in Table 2.4.

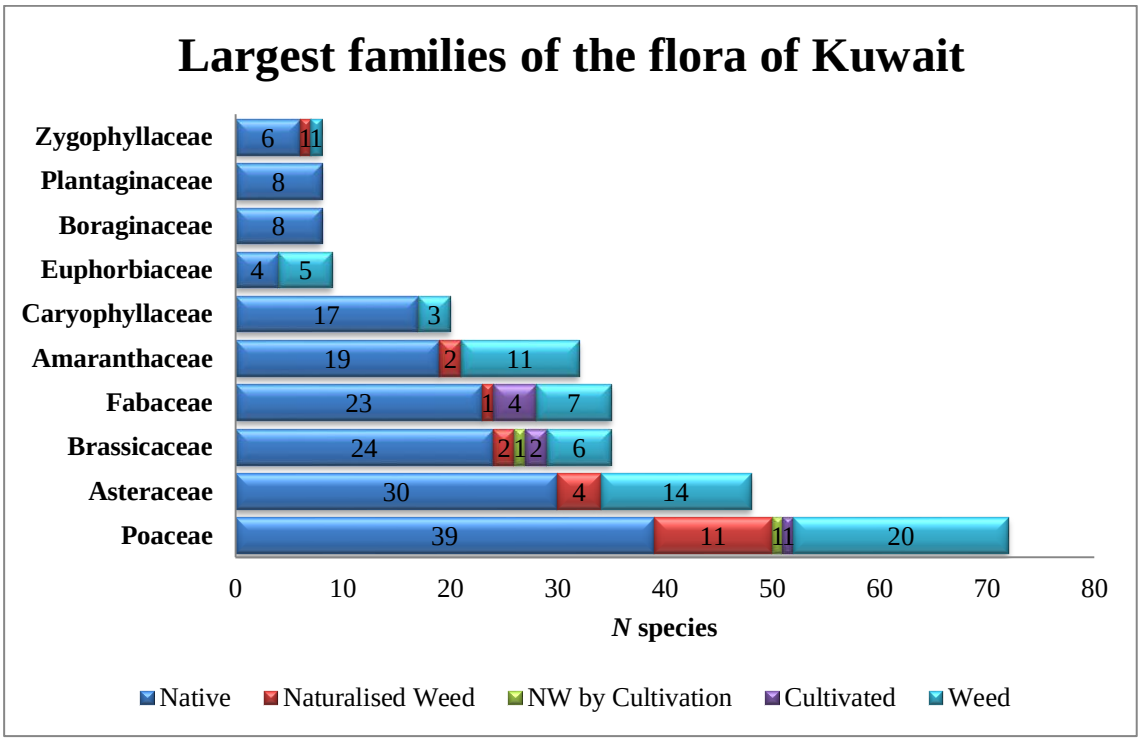


Figure 2.2 Bar chart showing largest families of the flora of Kuwait

Table 2.4 Plant species with changes to their families according to APG IV classification

Species	Family -present	Family -former
Allium sindjarense	Amaryllidaceae	Liliaceae
Allium longisepalum	Amaryllidaceae	Liliaceae
Allium sphaerocephalum	Amaryllidaceae	Liliaceae
Bellevalia saviczii	Asparagaceae	Liliaceae
Dipcadi erythraeum	Asparagaceae	Liliaceae
Asphodelus tenuifolius	Asphodelaceae	Liliaceae
Asphodelus viscidulus	Asphodelaceae	Liliaceae
Calotropis procera	Apocynaceae	Asclepiadaceae
Lomelosia olivieri	Caprifoliaceae	Dipsacaceae
Lomelosia palaestina	Caprifoliaceae	Dipsacaceae
Valerianella dufresnia	Caprifoliaceae	Valerianaceae
Cuscuta planiflora	Convolvulaceae	Cuscutaceae
Nitraria retusa	Nitrariaceae	Zygophyllaceae
Peganum harmala	Nitrariaceae	Zygophyllaceae
Fumaria parvifiora	Papaveraceae	Fumariaceae
Hypecoum littorale	Papaveraceae	Fumariaceae
Hypecoum pendulum	Papaveraceae	Fumariaceae
Andrachne telephioides	Phyllanthaceae	Euphorbiaceae

The total number of species of flowering plants, gymnosperms and ferns in the State of Kuwait is 402 species belonging to 256 genera of which 273 species are truly native, and 25 species are naturalised either by cultivation or naturally, 90 species are weeds and 12 are plants of cultivation (Figure 2.3). The species are found in 60 families and 26 orders, including gymnosperms and ferns (Figure 2.4). A chart representing the flora of Kuwait classified by major groups of angiosperms (monocot and dicot), gymnosperms and ferns are shown in Figure 2.4. The floristic analysis of the main groups of the flora of Kuwait is presented in Table 2.5. The Angiosperms are the largest group representing 254 genera and 400 species of which the dicots contain 310 species in 198 genera belonging to 45 families and 20 orders and the monocots 90 species in 56 genera in 13 families and 4 orders (Table 2.5 and Figure 2.4). A single species each represents the gymnosperms and ferns.

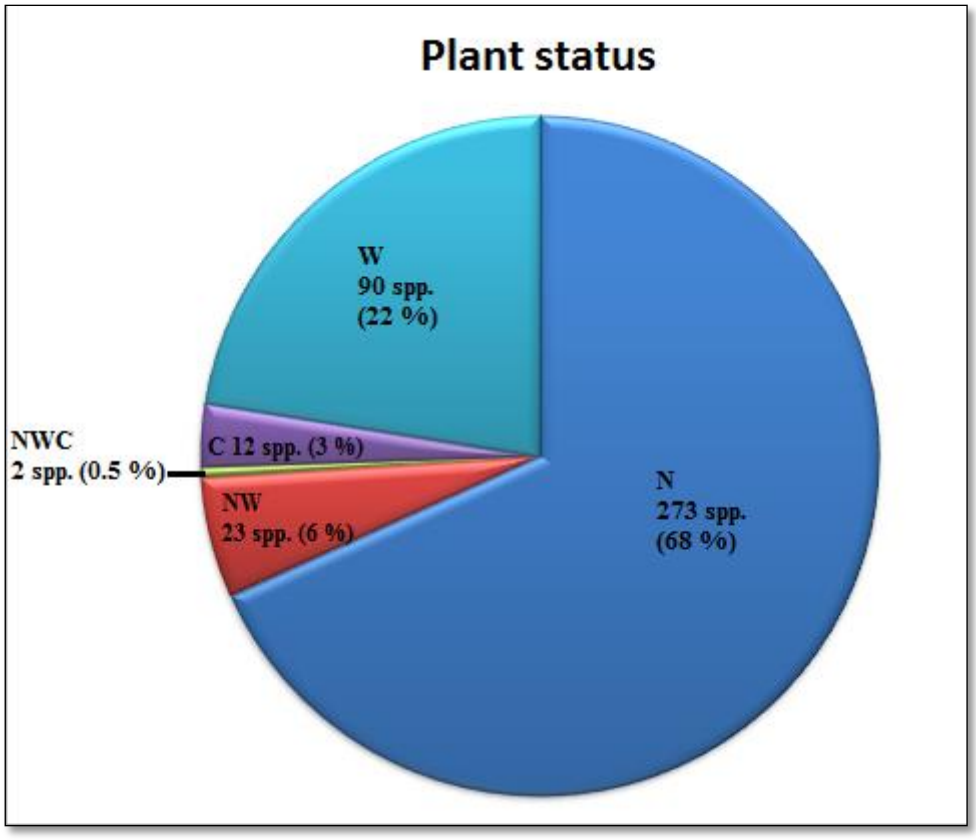


Figure 2.3 Pie chart showing the plant status of the Kuwaiti flora
N: native plants, NW: naturalised weed, NWC: naturalised weed by cultivation, C: cultivated plants, W: weedy plants

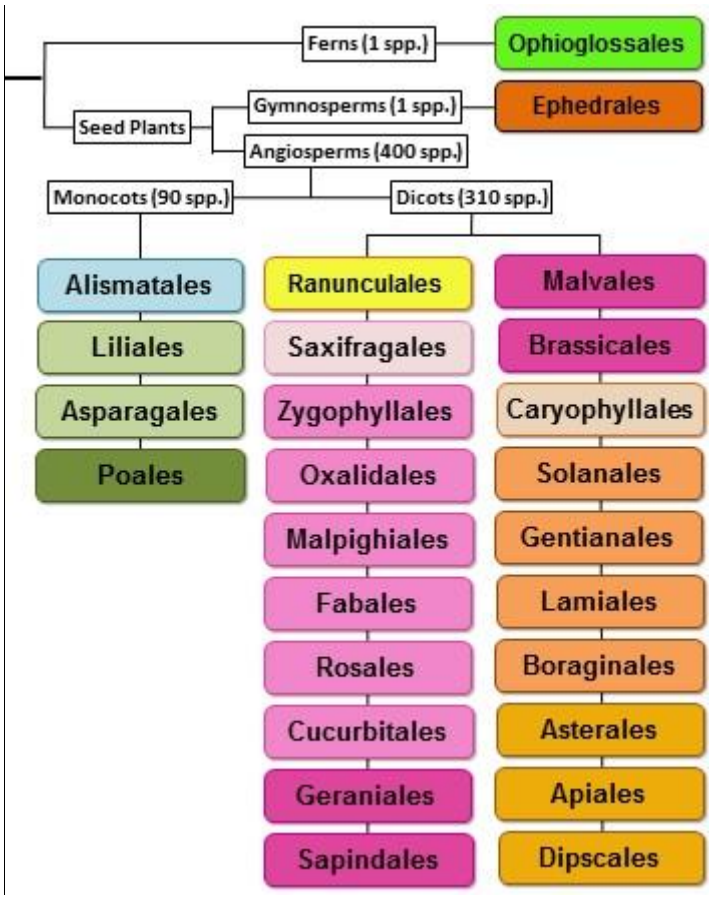


Figure 2.4 Illustration showing the classification of all plant orders of the flora of Kuwait according to APG IVclassification

Table 2.5 Floristic analysis of the flora of Kuwait

	Orders	Families	Genera	Species
Monocots	4	13	56	90
Dicots	20	45	198	310
Angiospermae total	**24**	**58**	**254**	**400**
Pteridophyta	**1**	**1**	**1**	**1**
Gymnospermae	**1**	**1**	**1**	**1**
Overall	**26**	**60**	**256**	**402**

2.3.2 Native and naturalised vascular plants of Kuwait

This part of the floral analysis is more focused on the native and naturalised vascular plants of the flora of Kuwait and is of interest to plant ecologist involved in conservation studies and restoration ecology.

If only truly native and naturalised vascular plants of the flora (excluding weeds and cultivated plants) are considered, the flora is represented by 298 species in 182 genera arranged in 52 families and 25 orders, including the gymnosperms and ferns each of which is represented by a single species (Table 2.6). The Dicots are represented by 232 species in 148 genera related to 39 families and 19 orders and monocots 64 species in 32 genera related to 11 families and 4 orders (Table 2.6). An illustration representing the native species and species established in the wild, classified by order and family names according to APGIV (2016) is shown in Figure 2.5.

Table 2.6 Floristic analysis of native and natuarlized plants of Kuwait

	Orders	Families	Genera	Species
Pteridophyta	**1**	**1**	**1**	**1**
Gymnospermae	**1**	**1**	**1**	**1**
Angiospermae total	**23**	**50**	**180**	**296**
Monocots	4	11	32	64
Dicots	19	39	148	232
Overall	**25**	**52**	**182**	**298**

The largest families of native vascular plants and those established in the wild are Poaceae [Gramineae] 51 spp, Asteraceae [Compositae] 34 spp, Brassicaceae [Cruciferae] 27 spp, Fabaceae [Leguminosae] 24 spp, Amaranthaceae [Chenopodiaceae] 21 spp and Caryophyllaceae 17 spp (Table 2.3 and Figure 2.5). Largest genera of native vascular plants and those established in the wild are *Astragalus* (10 spp all native), *Plantago* (8 spp all native), *Erodium* (5 spp all native), *Helianthemum* (4 spp all native), *Launaea* (4 spp: 3 native and 1 naturalised weed), *Arnebia* (4 spp all native), *Stipagrostis* (4 spp all native), and *Bromus* (4 spp 2 spp native, 2 spp naturalised weeds) (Table 2.3).

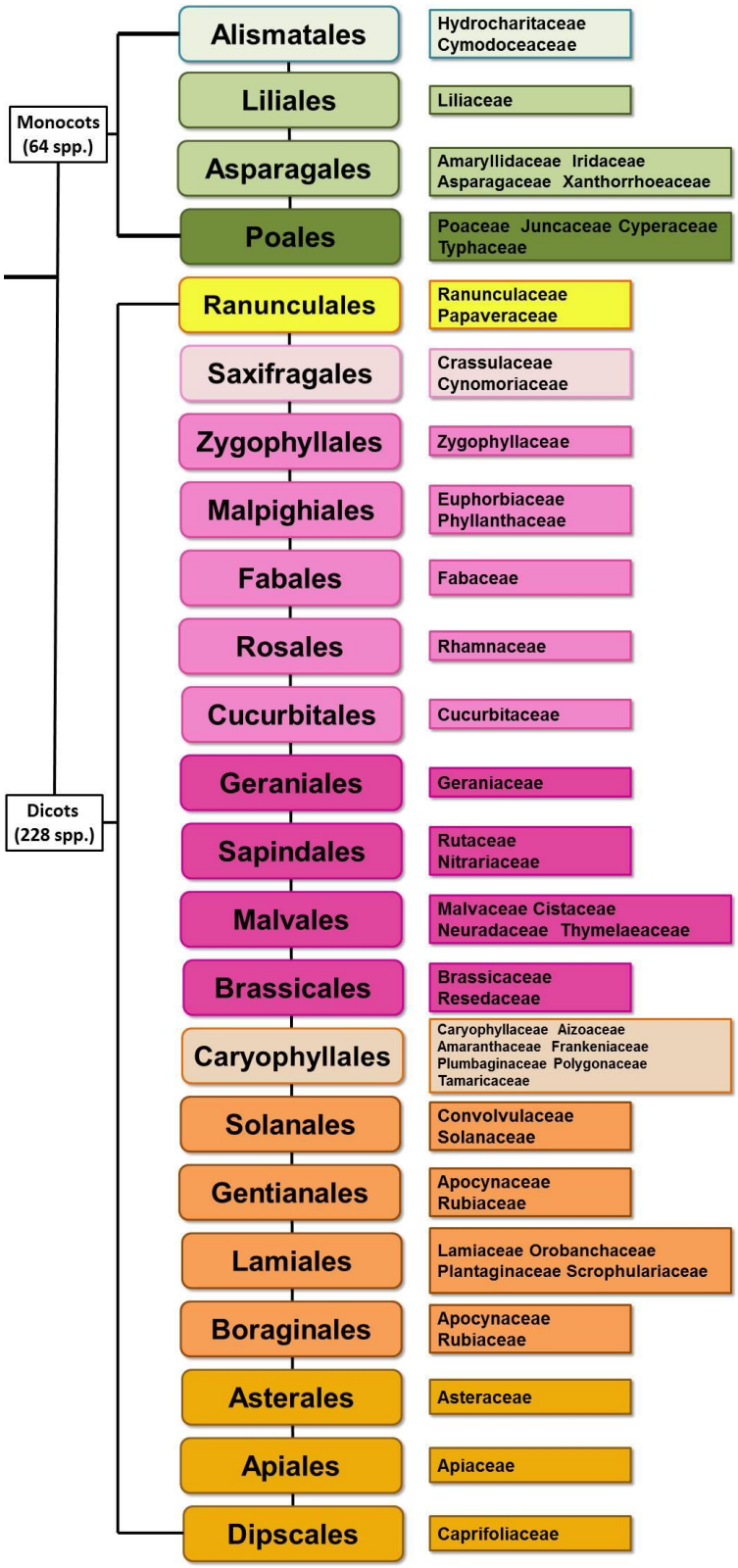


Figure 2.5 Illustration showing the ranking of 23 orders and 50 families of native and natuarlised plants of the flora of Kuwait according to APGIV (2016) classification.

2.4 Discussion

The present checklist provides an update of the flora of Kuwait and aids in conserving the knowledge of the plant diversity of the Kuwaiti flora. It will mainly act as a guide for DNA barcoding the entire flora of Kuwait in chapter 4. The plant checklist (Table 2.3) includes updates to nomenclature, authorities, synonyms and plant status. Also, the classification of the orders and families in line with APG IV (2016). The present checklist of the flora of Kuwait comprises 402 species of which 273 species are native, and 25 species are naturalised either by cultivation or naturally, 12 species are plants of cultivation and 90 species are weedy plants. The previous checklist presented by Boulos and Al-Dosari (1994) represented 374 species related to 55 families, comparing it with the current checklist comprises 402 species in 60 families. The addition of 28 species are represented by weeds (15 spp), cultivated plants (5 spp.) and eight native species (*Allium longisepalum*, *Astragalus hamosus*, *Convolvulus prostratus*, *Halopeplis perfoliata*, *Linaria albifrons*, *Neatostema apulum*, *Sphenopus divaricatus*, and *Ziziphus nummularia*) (Table 2.3).

2.4.1 Native plants

Climate change, land degradation, human activities, and overgrazing have gradually altered the pattern of native plant communities in Kuwait and several plants are being threatened and facing the danger of extinction. Restoration and revegetation programmes are necessary to conserve the biodiversity of these important ecosystems. *Haloxylon salicornicum* has gradually replaced *Rhanterium eppaposum* that once dominated the northern areas of Kuwait as a result of overgrazing and soil erosion (Brown, 2003). Omar et al. (2001) also reported a decrease in several plant communities' distribution: *Haloxylon salicornicum*, *Zygophyllum qatarense*, and *Rhanterium epapposum* (Omar et al., 2001).

Rhanterium epapposum is a perennial shrub usually found northern part of Kuwait is on the verge of extinction due to overgrazing and use as a source of firewood (Brown, 2003; Al-Salameen et al., 2014). *Rhanterium* plants have the ability to build up a large seed bank in the soil and stay viable for up to four years (Zaman, 2006) and have the

potential for vegetative regeneration from stumps in heavily grazed sandy areas (Omar and Bhat, 2008).

Important native plants that have the potential to build up the soil seed bank and stabilize moving sand by forming small sandy hillock 'Nabkha' should be considered in restoration projects e.g. *Lycium shawii, Panicum turgidum, Cyperus conglomerates, Astragalus spinosus, Rhanterium epapposum, Haloxylon salicornicum, Halocnemum strobilaceum, Salicornia europaea, Tamarix aucheriana* and *Nitraria retusa* (Ahmed et al., 2016).

Other important plants of the flora that has the potential to re-establish and restore vast desert areas represented by *Calligonum polygonoides, Heliotropium bacciferum, Arnebia decumbens, Convolvulus oxyphyllus, Gynandiris sisyrinchium* and *Ochradenus baccatus* (Abo El-Nil, 1997).

2.4.2 Cultivated plants

Cultivated plants in Kuwait represented by a small portion of the flora (3 %). The following discussion will include some trees grown and used in restoration projects in Kuwait. An interesting small tree re-introduced once again to the flora is mangrove (*Avicennia marina*) which once occurred naturally more than 70 years ago (mentioned by Dickson, 1955). The native mangrove species of the flora is extinct due to human usage of the plants as firewood and charcoal. Mangrove was re-introduced earlier (1991) by propagule cultivation from neighbouring countries, Bahrain and the United Arab Emirates, in the intertidal zones of Shuwaikh and Sulaibiya coastal areas of Kuwait. After only seven years of growth, a study by Abo El-Nil (2001) showed that mangrove successfully established from propagules, flowering and producing viable seeds (Abo El-Nil, 2001). At present, mangrove is considered to be a major plant in restoration projects along the shoreline of the State of Kuwait.

Another successful tree used in restoration projects is *Prosopis juliflora*. Having the ability to tolerate high salinity and temperature (El-Keblawy and Al-Rawai, 2005) and provide large vegetation cover, although its foliage contains water-soluble allelopathic

chemicals (e.g. tannins, flavonoids, steroids, and alkaloids) which are known to inhibit the germination and growth of other plants in its vicinity (Al-Humaid & Warrag , 1998; Pasiecznik et al., 2001). In my opinion, I see the use of this tree as an ornamental plant (e.g. planted along highways and local streets), avoiding it in large-scale desert restoration projects which are costly, since it requires large amounts of water for the first few years before it becomes established in the wild.

Tamarix aphylla tree is usually cultivated in Kuwait as an ornamental and windbreaker for being a fast-growing tree, extremely tolerant to drought and salinity and mainly used by birds for shelter and fodder while providing a source of firewood (Le Houerou, 1984). It is not recommended to be included in large-scale desert restoration projects due to its ability in spreading fast and competing with other native plants for space and water (Griffin et al., 1989).

2.4.3 Weedy plants

Weeds make up 24 % of the flora, which has the potential to spread and adapt outside their normal range of distribution. Frequently weeds in Kuwait are found growing near agricultural lands and gardens. Therefore, it is important to document plants and monitor them over a period of time, since they have the potential to become established in the wild without the aid of agricultural soils or irrigated water. In the present checklist, I documented eleven new weeds recently published (Mathew et al., 2012) not included in previous lists or floral publications of Kuwait. Such observations are important to document since I will be applying molecular analysis in the following chapters.

2.4.4 Conclusion

When considering the conservation of plant biodiversity, it is essential to understand the plant status in a given flora and also to record whether or not the species is naturalised by cultivation (or naturally), occur as a weed or plant of cultivation. This sort of classification can be used as a guide in prioritizing species for use in vegetation restoration programmes and other conservation initiatives. In the following

chapters, I will be using the presently updated checklist (Table 2.3) to guide me through DNA barcoding the entire flora of Kuwait and build a local DNA reference library.

Chapter 3 Choice of gene regions

3.1 Introduction

DNA barcoding provides a fast and reliable way to identify individuals by sequencing short region of its genome and comparing it with a DNA reference database. The cytochrome oxidase I (COI) mitochondrial gene, used widely in DNA barcoding animals, is not appropriate for land plants because of its slower rate of evolution in plants (Hebert et al., 2003; Kress et al., 2005).

Difficulties in finding a single DNA region that could serve as a barcode across land plants encouraged researchers to explore more gene regions (Tables 3.1and 3.2). The use of multi-loci in combination could increase sequence variation and identification ability. Initial investigations in previous studies (Kress et al., 2005; Fazekas et al., 2008; Ford et al., 2009; CBOL, 2009) have made use of a variety of genes from both nuclear (ITS) and plastid regions (*atpB-rbcL, psbM-trnD, trnC-ycf6, trnH-psbA, trnL-F, trnk-rps16, trnV-atpE, rpl36-rps8, ycf6-psbM, rpoB, rpoC1, rpoC2, rbcL, matK, 23S rDNA, atpF-atpH, psbK-psbI, accD, ndhA, ndhJ, ndhK, rpl22, ycf2, ycf5,* and *ycf9*). Different gene regions are capable of resolving species relationships at different levels depending on their mutation rate (slow or fast). The criteria used by most research groups in evaluating DNA barcoding regions for land plants must be: (1) routinely amplifiable (universality using single primer pair), (2) easily sequenced (producing bi-directional quality sequences) and (3) discrimination power (maximum discrimination of species) (Kress et al., 2005; Fazekas et al., 2008; Ford et al., 2009; CBOL, 2009). To standardise the selection of a plant barcode, the Consortium for the Barcode of Life (CBOL) initiated the formation of a working group with representatives from different research groups from the molecular systematics community that had proposed or tested the seven leading candidate barcoding markers (*atpF-atpH, matK, rbcL, rpoB, rpoC1, psbK-psbI, trnH-psbA*) (CBOL, 2009). A number of markers (e.g. *rpoC1* and *rpoB*) were eliminated from the proposed candidates due to lower discriminatory power. The final recommendation of the CBOL Plant Working group was based on applying a core-barcode consisting of portions of two plastids coding regions, *rbcL+matK*, to be supplemented with

additional markers (Table 3.1) as required (CBOL, 2009; Hollingsworth et al., 2011). The recommendation of *rbcL+matK* as a barcode region for *rbcL* was based on the ease of amplification, sequencing, and aligning in most land plants and *matK* showing great discriminatory power (although difficult to amplify using one set of primers) (CBOL, 2009). In spite of discriminatory power, *rbcL* region resulted in modest discriminatory power amongst species (48 % to 68 %), unlike *matK* showed higher resolution (65 % to 80 %) (Fazekas et al., 2008; CBOL, 2009; Liu et al, 2012).

Many studies used combinations of gene regions approaches using variable non-coding and conserved coding regions of the plastid genome to DNA barcode land plants. Table 3.2 summarises the studies published which involved the comparisons of multilocus DNA barcode candidates.

Table 3.1. Markers that have routinely been used as plant DNA barcodes.

Marker	Genomic source	Type
nrITS	Nuclear	Transcribed spacers and 5.8S gene
nrITS2	Nuclear	Transcribed spacer
atpF-H	Plastid	Inter-genic spacer
matK	Plastid	Protein coding
psbK-I	Plastid	Inter-genic spacer
rbcL	Plastid	Protein coding
rpoB	Plastid	Protein coding
rpoC1	Plastid	Protein coding
trnH-psbA	Plastid	Inter-genic spacer
trnL-F	Plastid	Intron and inter-genic spacer
trnL **(P6)**	Plastid	Intron

(Source: Hollingsworth et al., 2011)

Table 3.2 Summary of studies comparing DNA barcoding regions in plants

Study	Regions used	Sampling material	Regions / species resolution (%)	Barcode recommendation
Kress et al., 2005	*atpB-rbcL*, ITS, *psbM-trnD, trnC-ycf6, trnH-psbA, trnL-F, trnk-rps16, trnV-atpE, rpl36-rps8, ycf6-psbM*	19 species belonging to 7 angiosperm families	*trnH-psbA, rpl136-rpf8, trnL-F* 100% each, *trnC-ycf6, ycf6-psbM* 90% each, other regions 73-80%	ITS and *trnH-psbA*
Fazekas et al., 2008	*rpoB, rpoC1, rbcL, matK, 23S rDNA, trnH-psbA, atpF-atpH, psbK-psbI*	92 species belonging to 32 diverse genera of land plants	*rpoB* 43%, *rpoC1* 29%, *rbcL* 48%, *matK* 65%, *23S rDNA* 7%, *trnH-psbA* 59%, *atpF-atpH* 45%, *psbK-psbI* 44%	Combination of multi locus = *rbcL, rpoB, matK* with *trnH-psbA, atpF-atpH*
Ford et al., 2009	*accD, matK, ndhA, ndhJ, ndhK, rpl22, rpoB, rpoC1, rpoC2, ycf2, ycf5, ycf9,*	98 land plant taxa: 4 liverworts, 6 pteridophytes, 6 gymnosperms 28 monocotyledons 54 angiosperms	Based on amplification success: 85-94%	matK, rpoB, rpoC1, ndhJ, ycf5, accD
CBOL 2009	*atpF-atpH, matK, rbcL, rpoB, rpoC1, psbK-psbI, trnH-psbA*	Total 907 samples representing 550 species of angiosperms, gymnosperms and cryptogams	*rpoC1* 43%, *psbK-psbI* and *trnH-psbA* (68-69%), *rbcL* 61% *matK* 66%	Combination of *rbcL* + *matK* 72%
Hollingsworth et al., 2009	*rpoC1; rpoB, rbcL, matK, trnH-psbA, atpF-atpH, psbK-psbI*	Three groups: Inga 26 species (N = 44) angiosperms, gymnosperms; liverworts	Inga single loci = *matK* 31 %; dual locus = *matK* + *trnH-psbA* 57.7 %; three locus = *rbcL* + *matK* + (*rpoC1* or *trnH-psbA*) 69 %	Combination of multi-locus for land plants *rbcL, rpoC1, matK, trnH-psbA*
Luo et al., 2010	*trnH-psbA, ycf5. rpoC1, rbcL,* ITS2, ITS	192 species belonging to 72 genera (Rutaceae) Total 300 samples	ITS2 89%, *trnH-psbA* 83%, *rbcL* 78%, *rpoC1* 40%, *ycf5* 79%	ITS2 and *trnH-psbA*
Pettengill and Neel, 2010	*matK, rbcL, rpoB, rps2, trnT-trnL, trnL-intron, trnL-trnF, trnH-psbA,*	29 species (N = 92) in genus *Agalinis* (Orobanchaceae)	*trnL-trnF 67%, psbA-trnH* 65 %, *matK* 62%, other regions between 50 – 61%	*trnH-psbA* and *trnT-trnL*
Liu et al., 2011	*rbcL, matK, trnH-psbA, trnL-F*, ITS	47 samples belonging to 7 genera of *Taxus* in Eurasia	*rbcL* 46% , *matK* 80% , *trnH-psbA* 64%, *trnL-F* 100, ITS 100%	ITS and *trnL-F* alone or in combination to identify *Taxus* in Eurasia
Li et al., 2012	*rbcL, matK, trnH-psbA, psbK-psbI, atpF-atpH*, ITS	63 species in *Ficus* (Moraceae) of China (total samples 228)	ITS 72%, *psbK-psbI* 21%, *trnH-psbA* 19%, *atpF-atpH* 18%, *matK* 16%	Single loci ITS for DNA barcoding *Ficus*

One difficulty with making the decision has been that the level of discrimination varies between studies. From the widely applied gene regions, *matK* demonstrated high levels of discrimination among angiosperms in some studies, e.g. 88.8 % (Burgess et al., 2011) and 79.4 % (de Vere et al., 2012); in other studies it resulted in low resolution, e.g. 31 % (Hollingsworth et al., 2009) and 16 % (Li et al., 2012). This type of variation is also common for other gene regions, e.g. *rbcL*, *trnH-psbA*, *atpF-atpH* (Bolson et al., 2015; Saarela et al., 2013; Liu et al., 2012; Pettengill and Neel, 2010; Luo et al., 2010; Fazekas et al., 2008), due to very low levels of variation among closely related species. Several factors are responsible for reducing the power of species discrimination in particular lineages investigated in plant DNA barcoding studies as a result of e.g. hybridization, polyploidy, slow mutation rates (Meyer and Paulay, 2005; Hollingsworth et al., 2011). Amongst closely related species shared mitochondrial sequences have been observed as a result of hybridization which restricts the identification of species (Hebert and Gregory, 2005).

DNA barcoding is mainly used as an effective tool for identifying unknown plant specimens and compared with a library of reference barcode sequences derived from individuals of known identity (Hajibabaei et al., 2007). Many applications have been developed, e.g. DNA barcoding has been conducted on the verification of species of medicinal plants (Gong et al., 2016; Michel et al., 2016; Zhang et al., 2015), monitoring invasive plants (Sciver et al., 2015; Zhang et al., 2013), understanding herbivore diet (Soininen et al., 2013; Meheust et al., 2015), discovering cryptic species (Hernandez et al., 2015; Nigro et al., 2016; Guarnizo et al., 2015), identifying forest trees (Costion et al., 2016; Nithaniyal et al., 2014), Identification of below-ground diversity (Partel et al., 2012; Valverde-Barrantes et al., 2013; Kesanakurti et al., 2011), and the reconstruction of past vegetation history (Sonstebo et al., 2010).

The purpose of the present chapter is to evaluate the performance of five DNA barcoding regions for developing a DNA reference library of the flora of Kuwait. The choice of regions will follow the standard criteria mentioned earlier by CBOL (2009) for choosing DNA barcodes. The region must easily PCR amplify, produce bi-directional quality sequencing and show maximum discriminatory power (CBOL, 2009).

The regions tested are four plastid regions (*matK*, *rbcL*, *trnH-psbA*, *trnL-F*), and one nuclear (ITS2). The choice of tested regions was based on the recommendation of several previous studies showing high species resolution amongst angiosperms (CBOL, 2009; de Vere et al., 2012; Kress et al., 2005; Luo et al., 2010; Chen et al., 2010; Liu et al., 2011).

The evaluation included, PCR amplification success and universality of primers, high-quality sequencing, testing combined regions of successful barcodes, discrimination by similarity based on BLASTn (Basic Local Alignment Search Tool: Altschul et al., 1990) and generating monophyly tree based tests using Neighbour Joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) applied in DNA barcoding studies by Lahaye et al., 2008; Hollingsworth et al., 2009; Kress et al., 2009; Pettengill and Neel, 2010. Some of the questions that may be raised here are directly linked to the molecular identification and discrimination of species: 1) is it possible for the proposed markers to amplify and sequence using one set of universal primers each? 2) Which marker demonstrates the greatest level of species discrimination? 3) Which of the gene region combinations performs best in discriminating species?

3.2 Materials and methods

3.2.1. Sampling material

For the evaluation of DNA regions, a total of 49 individuals were sampled belonging to 25 spp. (listed in Table 3.3 and Appendix 3.1); 27 from herbarium specimens collected from Kuwait University herbarium (KTUH) and 22 from living plants collected throughout Kuwait. The samples used here belong to 4 largest genera of the flora of Kuwait: *Astragalus* (9 spp), *Plantago* (8 spp.), *Launaea* (4 spp.), and *Helianthemum* (4 spp.) belonging to the families Fabaceae, Plantaginaceae, Asteraceae, and Cistaceae, respectively (Table 3.3 and Appendix 3.1). The largest genera were chosen in an attempt to study the sequence diversity and barcode

discrimination amongst closely related species and investigated whether they are closely or distantly related within each genus. All species included two accessions for each species, except *Plantago notata* where only one herbarium sample was available (Table 3.3 and Appendix 3.1).

Table 3.3 List of plant species used in the evaluation of five DNA markers

Fabaceae	**Plantaginaceae**	**Asteraceae**	**Cistaceae**
***Astragalus* spp.**	***Plantago* spp.**	***Launaea* spp.**	***Helianthemum* spp.**
A. annularis	*P. amplexicaulis*	*L. angustifolia*	*H. kahiricum*
A. bombycinus	*P. boissieri*	*L. capitata*	*H. ledifolium*
A. corrugatus	*P. ciliata*	*L. mucronata*	*H. lippii*
A. hamosus	*P. coronopus*	*L. nudicaulis*	*H. salicifolium*
A. hauarensis	*P. lanceolata*		
A. schimperi	*P. notata*		
A. sieberi	*P. ovata*		
A. spinosus	*P. psammophila*		
A. tribuloides			

Fresh plant vouchers were determined using several floristic publications on the flora of Kuwait (Daoud, 1985; Al-Rawi, 1987; Boulos, 1988) and reconfirmed by consulting an expert, the herbarium curator at Kuwait University Herbarium (KTUH), Dr. K.T. Mathew. Herbarium vouchers made for freshly collected plants were deposited in KTUH and Royal Botanic Garden Edinburgh (RBGE) Herbarium (E), mounted and digitised.

In addition to the freshly sampled material (to complete the data set) herbarium samples were collected from KTUH specimens, based on the following criteria:

1. Specimens which have been determined by an expert in addition to the authors;

2. Small samples of leafy tissue were collected without destroying the value of the herbarium specimen;
3. Three individuals (where available) from separate geographical locations within the boundary of Kuwait;
4. Recent collections preferably; and
5. Referenced and linked to KTUH herbarium specimens.

A list of species used, the source of material (fresh/herbarium), unique Edinburgh DNA numbers (EDNA) for each DNA prep and name of collectors are presented in Appendix 3.1.

3.2.2 DNA extraction, amplification and sequencing

DNA extraction

DNA was extracted from leafy material (freshly collected silica dried or herbarium material) using Qiagen Plant DNeasy kits following the manufacturer's protocol: DNeasy Plant Handbook. A small amount (~ 20 mg) of dry, healthy leaf material was selected and loaded in 2.0 ml Eppendorf tubes with one 5 mm stainless steel bead. Samples were ground using TissueLyser II (Qiagen, Ltd.) until the material was powdered (frequency 20 Hz x 2 x ~ 30 sec).

All DNA extractions were given a unique EDNA accession number (Appendix 3.1) and banked for long term storage at RBGE.

PCR amplification

DNA fragments were amplified via standard polymerase chain reaction (PCR). All primers synthesised for PCR and sequencing are listed in Table 3.4 and the reaction conditions for each region are presented in Table 3.5.

Table 3.4. DNA regions and primers used

Region	Primer	Genome	Direction	Primer Sequence	Reference
rbcL	aaf	Plastid	Forward	ATGTCACCACAAACAGAGACTAAAGC	Kress & Erickson, 2007
rbcL	ajf634R	Plastid	Reverse	GAAACGGTCTCTCCAACGCAT	Fazekas et al., 2008
trnH-psbA	psbA3'f	Chloroplast	Forward	GTTATGCATGAACGTAATGCTC	Sang et al., 1997
trnH-psbA	trnHf	Chloroplast	Reverse	CGCGCATGGTGGATTCACAATCC	Tate & Simpson, 2003
ITS2	S2F	Nuclear	Forward	ATGCGATACTTGGTGTGAAT	Chen et al., 2010
ITS2	S3R	Nuclear	Reverse	GACGCTTCTCCAGACTACAAT	Chen et al., 2010
matK	Xf	Chloroplast	Forward	TAATTTACGATCAATTCATTC	Ford et al., 2009
matK	MALPR1	Chloroplast	Reverse	ACAAGAAAGTCGAAGTAT	Dunning & Savolainen, 2010
matK	1RKIM-f	Chloroplast	Forward	ACCCAGTCCATCTGGAAATCTTGGTTC	Ki-Joong Kim, pers. comm
matK	3FKIM-r	Chloroplast	Reverse	CGTACAGTACTTTTGTGTTTACGAG	Ki-Joong Kim, pers. comm
trnL	C	Chloroplast	Forward	CGAAATCGGTAGACGCTACG	Taberlet et al., 1991
trnL	D	Chloroplast	Reverse	GGGGATAGAGGGACTTGAAC	Taberlet et al., 1991

PCR for *rbcL*, *psbA-trnH*, *trnL*, and ITS2 regions were performed using one set of primer for each locus (forward/ reserve) in 20 µl reactions containing 1.5 Unit Biotaq (Bioline), 1 x PCR Buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.5 mM of each primer, 1 x Combinatorial PCR Enhancer Solution (CES) and 1.0 µl (30-50 ng/µl) genomic DNA.

PCR for *matK* was performed using three sets of primers in 10 µl reactions containing 1.5 Unit Biotaq, 1 x PCR Buffer, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.5 mM of each primer, 1 M Betaine and 1 µl (30-50 ng/µl) genomic DNA. PCR cycles used for each region are presented in Tables 3.4. Positive and negative controls were included in each PCR run to check for any contamination and help with troubleshooting.

Additives such as Betaine, Trehalose and/ or CES were included to enhance the PCR amplification. CES includes 2.7 M betaine, 6.7% dimethyl sulfoxide (DMSO) and 50 mg/ml bovine serum albumin (BSA).

Betaine and DMSO aid in reducing secondary structure of GC-rich templates and improves yield during PCR amplification (Jensen et al., 2010). BSA is effective when amplifying old/ ancient DNA contain PCR inhibitors such as phenolic compounds (Farell and Alexandre, 2012). CES is a combination of all enhancers' betaine, DMSO and BSA (Ralser et al., 2006).

Table 3.5. PCR cycles used for each region

Region	PCR cycle
rbcL	94 C 1 min
	94 C 45 sec, 51 C 45 sec, 72 C 2 min, 40 cycles
	72 C 7 min
psbA-trnH	94 C 3 min
	94 C 45 sec, 50 C 45 sec, 72 C 1 min, 2 cycles
	94 C 45 sec, 45 C 45 sec, 72 C 1 min, 30 cycles
	72 C 5 min
ITS2	95 C 4 min
	94 C 1 min, 55 C 1 min, 72 C 45 sec, 30 cycles
	72 C 5 min
matK	94 C 1 min
	94 C 30 sec, 55 C 30 sec, 72 C 1 min, 10 cycles
	88 C 30 sec, 50 C 30 sec, 72 C 1 min, 25 cycles
	72 C 10 min
trnL	95 C 4 min
	94 C 1 min, 50-55 C 1 min, 72 C 2 min, 35 cycles
	72 C 5 min

Gel electrophoresis

DNA extractions and PCR amplifications samples were run on an agarose gel (1 %). SYBR Safe DNA gel stain was added at a concentration of 5 µl per 100 ml to allow visualisation of DNA.

Agarose gels were visualised using GeneSyn software and a Gene Genius UV trans-illuminator system. Band brightness relative to the 1 kb+ ladder determined how

much PCR product to use in the sequencing PCR reaction – for the same brightness of the ladder 1 µl was used, for brighter bands 0.5 µl and fainter bands 1.5-2 µl.

PCR purification

PCR products were cleaned up using ExoSAP IT. ExoSAP IT contains two hydrolytic enzymes, Exonuclease I to degrade single-stranded primers and Shrimp Alkaline Phosphatase to remove dNTPs. 2 µl ExoSAP IT was mixed with 5 µl PCR products in 0.2 ml reaction tubes or 96 well plates and incubated in a thermocycler at 37 °C for 15 minutes followed by heating at 80 °C for 15 minutes to inactivate enzymes.

Sequencing Protocol

BigDye Sequencing was performed in 10 µl reaction containing 0.5 µl BigDye Terminator v3.1 Cycle Sequencing kit, 2 µl of 5 x BigDye Buffer, 0.32 µl of 10mm Primer, 6.68 µl H_2O and 0.5 µl of purified PCR.
BigDye sequencing cycles incubated in a thermocycler: 25 cycles of 95 ˚C for 30 sec, 50˚C for 20 sec, and 60˚C for 4 mins.

3.2.3 Sequence editing, alignment and molecular analysis

Geneious software (ver. 6.1.8, Biomatters Ltd., Kearse et al., 2012) was used to trim ends (using a 25 bp window segments with > 2 bp showing QV < 20 removed) and assemble the sequences into contigs. Every contig was checked for base call disagreements and ambiguities and manually edited where necessary. Poor quality sequences that were not amenable to manual editing, those with low overlap (less than 50 %) were removed. Also, the number of contigs meeting the criteria for high quality sequences according to the CBOL Plant Working Group (2009) was determined. The CBOL (2009) define high quality sequences as those in which both the forward and reverse reads have a minimum length of 100 bp, a minimum mean QV of 30 and the post-trim lengths are > 50% of the original read length; the assembled contig have > 50% overlap in the alignment of the forward and reverse reads with < 1% low-quality bases (<20 QV) and < 1% internal gaps and substitutions when aligning the forward and reverse reads.

Quality statistics including the amount of bi-directional read, mean QV of sequences, the percentage of high (QV > 30) and low quality (QV < 20) bases were calculated for each contig using Geneious software. Multiple sequence alignments (MSA) were performed by MUSCLE alignment (Edgar, 2004) using Molecular Evolutionary Genetics Analysis software version 7.0 (MEGA7: Kumar et al., 2015). For the non-coding regions (ITS2 and *trnL*) the settings for Gap penalties were adjusted to generate MSA with fewer Gaps in the final alignment (the default settings: Gap open -400 with Gap extend at 0. Gap penalties for ITS2 region was set at: Gap open -800 with Gap extend -1; for *trnL* Gap open was set at -900 with Gap extend -1. The MEGA software was also used to describe the genetic variability of each marker by calculating the mean length of base pairs, total aligned base pairs, variable sites (%), parsimony informative sites (%), and singleton sites (%).

A list of specimens with Barcode of Life Data Systems (BOLD) and GenBank accessions representing successfully barcoded sequences for *rbcL*, *trnL* and ITS2 are represented in Appendix 3.2.

BLASTn searches of barcoded sequences

A BLASTn search was performed to test the barcode sequences for similarity, available online by National Center for Biotechnology Information (NCBI) (Altschul et al., 1990).

Sequences for each region were queried against the NCBI non-redundant database through Geneious software using BLASTn search tool (built in plugin). The DNA barcodes were blasted for sequence similarities and best match based on sequences already banked online at NCBI, which automatically includes a search through the following databases: GenBank, Reference Sequence (RefSeq), European Molecular Biology Laboratory- European Bioinformatics Institute (EMBL-EBI), DNA DataBank of Japan (DDBJ) and Protein Data Bank (PDB) (accessed on 10th April 2017 using Geneious ver. R10, Kearse et al., 2012). The query excluded human sequences by using the following command: all[filter] NOT human[orgn], and the maximum hits was set at 100. The cut-off was determined by a match ≥ 99 % identity.

Sequence discrimination

The level of species discrimination of barcoded sequences was calculated by the generation of a monophyly tree based test using Neighbour-joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) (Saitou and Nei, 1987; Fitch, 1971; Felsenstein, 1981, respectively).

Tree-based analysis: NJ, ML and MP trees were generated for each region using MEGA7 software (Kumar et al., 2015). The genetic distance model for all tree-based analyses was computed using the Kimura 2-parameter method (Kimura, 1980) with gaps/ missing data treatment adjusted using pairwise deletion with bootstrap support set at 1000 replicates.

The tree-based methods (NJ, ML, and MP) were used to evaluate which tree produced greater species resolution and whether the barcode sequences form monophyletic groups, in addition, to calculate the percentage resolution of species-specific clusters for each.

3.3 Results

3.3.1 DNA recoverability, amplification and sequences success.

Comparing the plant material types (fresh/ herbarium) for all tested accessions showed DNA extracted from freshly silica-dried leaves has the highest percentage of specimens for which DNA sequences recoverable (95 %) against herbarium specimens (80 %). Freshly collected material produced higher levels of amplification and sequencing success rates were more consistent across the samples (Table 3.6).

Table 3.6 summarises PCR amplification and sequencing success rates for all markers tested. PCR amplification and sequencing from silica-dried and herbarium plant material was highest for *rbcL* and nrITS2 (each with 98 % success), followed by *trnL* (88 %), *matK* (80 %) and *trnH-psbH* (65 %) (Table 3.6).

Table 3.6. Summary of the proportion of samples successfully amplified and sequenced for five barcoding regions using fresh and herbarium plant material

DNA Region/ Collection type	No of individuals sampled	Sequence efficiency (%)	Amplification failure (%)	Sequence failure (%)
rbcL	**49**	**48 (98)**	**1 (2)**	**0**
Herbarium	27	26 (96)	1 (4)	0
Fresh	22	22 (100)	0	0
trnL	**49**	**42 (86)**	**6 (12)**	**1 (2)**
Herbarium	27	21 (78)	5 (19)	1 (4)
Fresh	22	21 (95)	1 (5)	0
nrITS2	**49**	**48 (98)**	**0**	**1 (2)**
Herbarium	27	26 (96)	0	1 (4)
Fresh	22	22 (100)	0	0
matK	**49**	**39 (80)**	**8 (16)**	**2 (4)**
Herbarium	27	19 (70)	7 (26)	1 (4)
Fresh	22	20 (91)	1 (5)	1 (5)
trnH-psbA	**49**	**32 (65)**	**10 (20)**	**7 (14)**
Herbarium	27	14 (52)	8 (30)	6 (22)
Fresh	22	18 (82)	2 (9)	2 (9)

The five barcoding markers tested showed differences in amplification and sequencing success. Amongst the regions tested *rbcL* and ITS2 resulted in the highest amplification and sequencing success rate (both 98 %), followed by *trnL* (86 %), *matK* (80 %), and the lowest was *trnH-psbA* (65 %) (Table 3.6). Thus, the amplification and sequence failures were highest for *trnH-psbA* followed by *matK*, 20 % and 14 %, and 16 % and 4 %, respectively (Table 3.6). *matK* and *trnH-psbA* amplified better using freshly collected material than herbarium specimens. *rbcL* and ITS2 showed the greatest amplification success using one set of primer each. *trnL* and *trnH-psbA* were also amplified using one set of primers each but showed very low amplification success compared to *rbcL* and ITS2 regions.

For *matK*, not all specimens amplified and sequenced from the first run and the locus was tested using three different sets of *matK* primers. First set: forward primer *matK*-Xf (Ford et al., 2009), with reverse primer *matK*-MALPR1 (Dunning & Savolainen, 2010). Second set: forward primer *matK*-1RKIM-f and reverse primer *matK*-3FKIM-r

(Kim et al., 2010). Third set: forward primer *matK*-1RKIM-f and reverse primer *matK*-MALPR1 (See Table 3.4. for primers list). Overall, for *matK* region, after several amplification attempts, only 39 out of 49 species (80%) successfully sequenced and barcoded, 8 samples belonging to *Helianthemum* spp. failed to amplify.

matK and *trnH-psbA* regions were excluded from any further analyses due to high amplification and sequencing failures (20 % and 35 %, respectively).

Single regions (*rbcL*, ITS2, *trnL*) and combined regions (*rbcL* + ITS2, *rbcL* + *trnL*, and *trnL* + ITS2) were further analysed and tested for levels of species discrimination.

3.3.2 Sequence quality and alignment

Single region sequences: The multiple sequence alignment (MSA) lengths of all tested accessions for *rbcL*, *trnL* and ITS2, barcodes are 532, 488 and 408 bp, respectively. For *rbcL*, the amplicon sizes of the two primer pairs were 532 bp, followed by *trnL*, 342 bp to 453 bp and ITS2 barcodes ranged from 378 bp to 399 bp. The percentage of gaps present in the non-coding regions was highest in *trnL*, 16 % followed by ITS2, with only 7 % (Table 3.7). The coding region, *rbcL*, showed no gaps in the MSA.

The mean percentage of high-quality bases within the sequences (defined as a QV score greater than 30) ranged from 98.3 % to 99.3 % for all three regions with *rbcL* containing the highest quality of 99.3 % (Table 3.7). Mean low quality (QV < 20) was less than 1 % for all three regions. The ITS2 matrix contained the most variable sites 53 % and parsimony informative sites 52 % followed by *trnL* with variable sites of 52 % and parsimony informative sites of 50 % followed by *rbcL* with variable sites of 20 % and parsimony informative sites of 19 % (Table 3.7).

Manually exploring the variation in the multiple sequence alignments (MSA) based on a genus-by-genus basis, it is clear that sequences representing a genus showing a unique pattern of sequence and gaps that differentiate them from the other genera. Looking at *rbcL* MSA, although the sequence variation amongst all the sites was low

(20 %), the four different genera each contained unique sequences that distinguished the groups in the MSA. For example, unique nucleotide bases noted for *Astragalus* spp. T-thymine present at positions 184, 308, 332, 333, 497; G-guanine present at 137, 329; A-adenine present at 191, 530; C-cytosine present at 257, 299, 464).

Unlike the coding region *rbcL*, the non-coding regions ITS2 and *trnL*, contain 7 % and 16 % alignment gaps, respectively (Table 3.7). For the non-coding regions, the variation amongst the four genera resulted from a combination of nucleotide bases present at different positioning along the sequences and also the presence of alignment gaps. For example, in ITS2 MSA, *Plantago* spp were distinguishable by: T-thymine present at positions 138, 185, 230, 247, 351, 402; G-guanine present at 241, 217,238; A-adenine present at 178, 233, 312; C-cytosine present at 25, 86, 169, 192, 234, 286; Alignment Gaps present at the following positions: 88-97, 138-149, 221-224, 274-281).

Combined regions sequences: The multiple aligned sequence lengths for *rbcL* + ITS2, *rbcL* + *trnL* and *trnL* + ITS2 barcodes are 940, 1020, and 896 bp, respectively. *trnL* + ITS2 matrix contained the most variable sites of 467 (52 %) and parsimony informative sites 454 (51 %), followed by *rbcL* + ITS2 with variable sites of 332 (35 %) and parsimony informative sites 319 (34 %), and *rbcL* + *trnL* with variable sites of 360 (35 %) and parsimony informative sites 350 (34 %) (Table 3.7). The percentage of alignment gaps was highest for the combined region *trnL* + ITS2 (12 %), followed by *rbcL*+*trnL* (7.7 %) and *rbcL*+ITS2 representing the fewest alignment gaps (3.2 %).

Table 3.7.Alignment metrics for single regions and combined regions

	Single DNA region			Combined DNA regions		
	rbcL	**ITS2**	***trnL***	***rbcL*+ITS2**	***rbcL*+*trnL***	***trnL*+ITS2**
***N* species/ total individuals (%)**	48/49 (98)	48/49 (98)	42/49 (90)	47/49 (96)	42/49 (86)	41/49 (86)
Aligned sequence length bp	532	408	488	940	1020	896
Minimum sequence length bp	-	378	342	880	874	739
Maximum sequence length bp	-	399	453	931	985	840
Number of Gaps (%)	0	30 (7)	79 (16)	30 (3.2)	79 (7.7)	109 (12)
Variable sites bp (%)	105 (20)	217 (53)	255 (52)	332 (35)	360 (35)	467 (52)
Parsim-inform sites bp (%)	105 (19)	214 (52)	245 (50)	319 (34)	350 (34)	454 (51)
Singleton sites (%)	0	3 (0.7)	10 (2)	3 (0.3)	10 (0.9)	13 (1.5)
GC content (%)	43.30	45.40	33.6	49.0	38.9	44.0
Mean high quality bases QV>30 (%)	99.30	98.90	98.30			
Mean low quality bases QV<20 (%)	0.30	0.40	0.70			

3.3.3 BLASTn searches

BLASTn searches were applied to provide valuable insights into understanding how well the 25 tested species belonging to the flora of Kuwait are represented by comparing them to similar sequences banked at NCBI database. Matches to NCBI database of *rbcL*, ITS2, and *trnL* sequences were determined by BLASTn cut-off value ≥ 99 % identity for a top match.

NCBI database (http://blast.ncbi.nlm.nih.gov/Blast) was accessed by Geneious plugin (on 10th April 2017). By blasting each barcoded sequence against the NCBI database, and comparing the sequences using BLASTn cut-off value ≥ 99 % identity, a list of similarities was compiled according to genus and species match.

All sequences representing *rbcL*, *trnL* and ITS2 barcodes of the 25 tested species matched with a similar sequence present in the NCBI database at the family, genus level and 16 matched to species level (Appendix 3.3). ITS2 and *trnL* barcodes showed the highest number of 24 sequences matching at the species level (represented by six species for each region), 24 % and 25 %, respectively. *rbcL* barcodes matched four

species only (16 %), other sequences for the three markers matched to genus level with few to the family level only.

In order to understand how well the sequences of the 25 tested species of the flora of Kuwait (Table 3.3) are well represented in the NCBI database, I compiled a list of Genbank accessions (Appendix 3.3) representing the three markers (*rbcL*, *trnL*, ITS2) by performing a search by species names using NCBI taxonomy database and comparing the accessions list with that generated by BLASTn matches of the barcode sequences, using the following web page accessed on 12th April 2017, NCBI taxonomy database: https://www.ncbi.nlm.nih.gov/taxonomy

In Appendix 3.3, the Genbank accessions in bold text represent sequence matches to *rbcL*, *trnL*, and ITS2 barcoded sequences of the 25 tested species of the Kuwaiti flora using BLASTn; the other listed Genbank accessions (normal text Appendix 3.3) are representatives of *rbcL*, *trnL* and ITS2 accessions searched by species name through the NCBI taxonomy database (see web page above), (Appendix 3.3).

The largest number of GenBank accessions (using NCBI taxonomy database search by species name) compiled for ITS2 followed by *trnL* and *rbcL*, 33, 18, and 13, respectively (Appendix 3.3). A total of 33 ITS2 accessions representing 18 species present in the NCBI database, from which only 6 species matched (using BLASTn) the ITS2 barcodes of the 25 tested species belonging to the Kuwaiti flora. A total of 18 *trnL* accessions belonging to 9 species from which only 6 species matched the *trnL* barcodes of the 25 tested species. A total of 13 *rbcL* accessions were compiled representing 6 species from which only 4 species matched the barcode sequences of the 25 tested species (Appendix 3.3).

3.3.4 Species discrimination

Monophyly tree based analyses

Phylogenetic tree based analysis using Neighbour Joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) trees were reconstructed to evaluate the four genera for the three barcode regions (*rbcL*, ITS2, *trnL*) and three combinations (*rbcL*+ITS2, *rbcL*+*trnL*, *trnL* + ITS2), are resolvable as monophyletic groups and to determine the levels of species discrimination.

Single region tree based analysis

All tested trees resolved all four genera (*Astragalus* spp., *Plantago* spp., *Helianthemum* spp., *Launaea* spp.) for *rbcL*, ITS2 and *trnL* as monophyletic groups with ≥ 99 % clades support using bootstrap of 1000 replicates. The comparison of three phylogenetic tree methods (NJ, MP, ML) generated very similar topologies and species resolution shown in Table 3.8. and Figures 3.1-3.3. Species-specific clustering for NJ resulted in the highest resolution for two markers *rbcL* and ITS2, 60 % and 64 %, respectively. NJ and MP resulted in equal resolution for *trnL* marker, 42 % compared to ML with 38 % species resolution (Table 3.8).

Comparing the three phylogenetic trees and three markers together, ITS2 resulted in the highest species-specific clustering, followed by *rbcL* and *trnL*. For ITS2 region the levels of species discrimination was greatest, 56-64 %, using all tree methods (Table 3.8). *rbcL* region showed the highest species resolution using NJ, 60 %, followed by MP and ML with 52 % each. The *trnL* region resulted in the lowest species resolution 38-42 % for all tree methods (Table 3.8).

Table 3.8 Percentage of species-specific clusters using phylogenetic reconstruction tree for a single and combined DNA regions.

DNA region(s)	*N* species/ total individuals	Species-specific clusters		
		NJ (%)	MP (%)	ML (%)
rbcL	25/ 48	15 (60)	13 (52)	13 (52)
ITS2	25/ 48	16 (64)	14 (56)	14 (56)
trnL	24/ 43	10 (42)	10 (42)	9 (38)
***rbcL* + ITS2**	25/ 46	16 (64)	16 (64)	16 (64)
rbcL* + *trnL	24/ 43	12 (50)	12 (50)	10 (42)
***trnL* + ITS2**	24/ 42	12 (50)	10 (42)	10 (42)

NJ trees are presented here for all three regions and three combinations for further analyses (MP and ML trees are provided in Appendix 3.4. due to lower species resolution than NJ). The topology generated by NJ demonstrated the greatest support values for ITS2 region (Figure 3.1) which was capable of recovering 64 % of species-specific clusters, followed by *rbcL* 60 % (Figure 3.2) and *trnL* with only 42 % (Figure 3.3 and Table 3.8). Based on these results, I will explore in more detail *rbcL* and ITS2 NJ analyses.

NJ analysis for *rbcL* resulted in slightly better monophyletic species clade support, ranging from 64 to 100 % compared to ITS2, 61 to 100 %. Paraphyletic clades were present across both regions, *rbcL* showed four paraphyletic clades and ITS2, showed three clades (Figures 3.1-3.2).

Results of species-specific clusters for the four major plant groups (Table 3.9) shows great species resolution for ITS2 compared to *rbcL* region. ITS2 discriminated 100 % of *Launaea* spp., followed by *rbcL* (only 50 %) and 50 % for *Plantago* spp., while *rbcL* showed only 38 % species discrimination. *rbcL* showed a better species resolution with *Astragalus* spp., 78 % compared with ITS2, slightly lower, 67 % (Table 3.9). *Helianthemum* spp. for both regions (*rbcL* and ITS2) resulted in an equal discriminatory power, 75 % each, (Table 3.9 and Figures 3.1 and 3.2).

Table 3.9 Percentage of species-specific clusters for major groups using NJ for single DNA regions and different combinations

	Species-specific clusters showing major groups using NJ ***N*** **species/ total individuals per clade (NJ %)**			
DNA region(s)	*Astragalus* spp.	*Plantago* spp.	*Helianthemum* spp.	*Launaea* spp.
rbcL	7/9 (78)	3/8 (38)	3/4 (75)	2/4 (50)
ITS2	6/9 (67)	4/8 (50)	3/4 (75)	4/4 (100)
rbcL **+ ITS2**	6/9 (67)	3/8 (38)	3/4 (75)	4/4 (100)

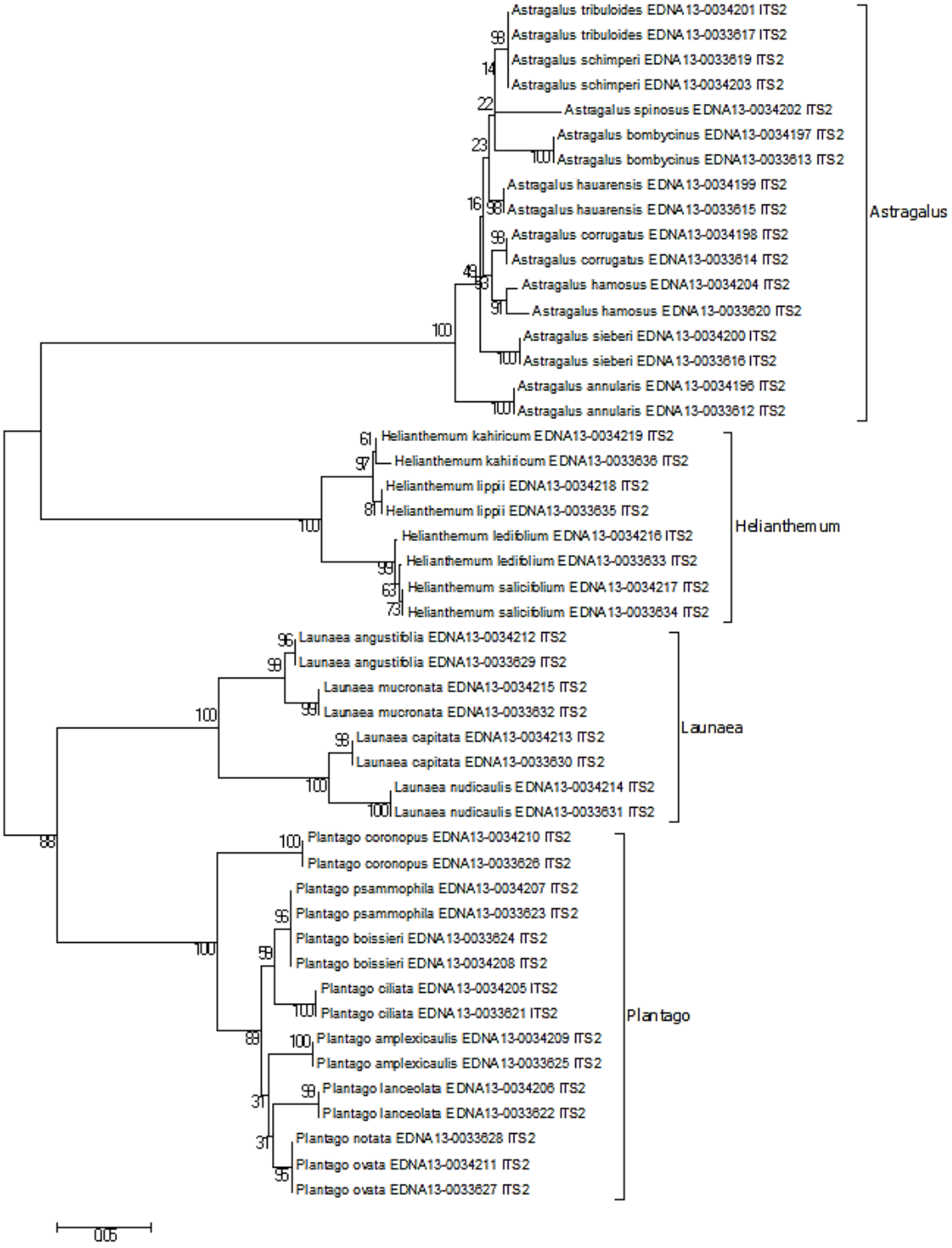


Figure 3.1 Neighbour joining phylograms for ITS2 barcodes illustrating the four largest genera of the flora of Kuwait (values represent % boot strap support with 1000 replicates)

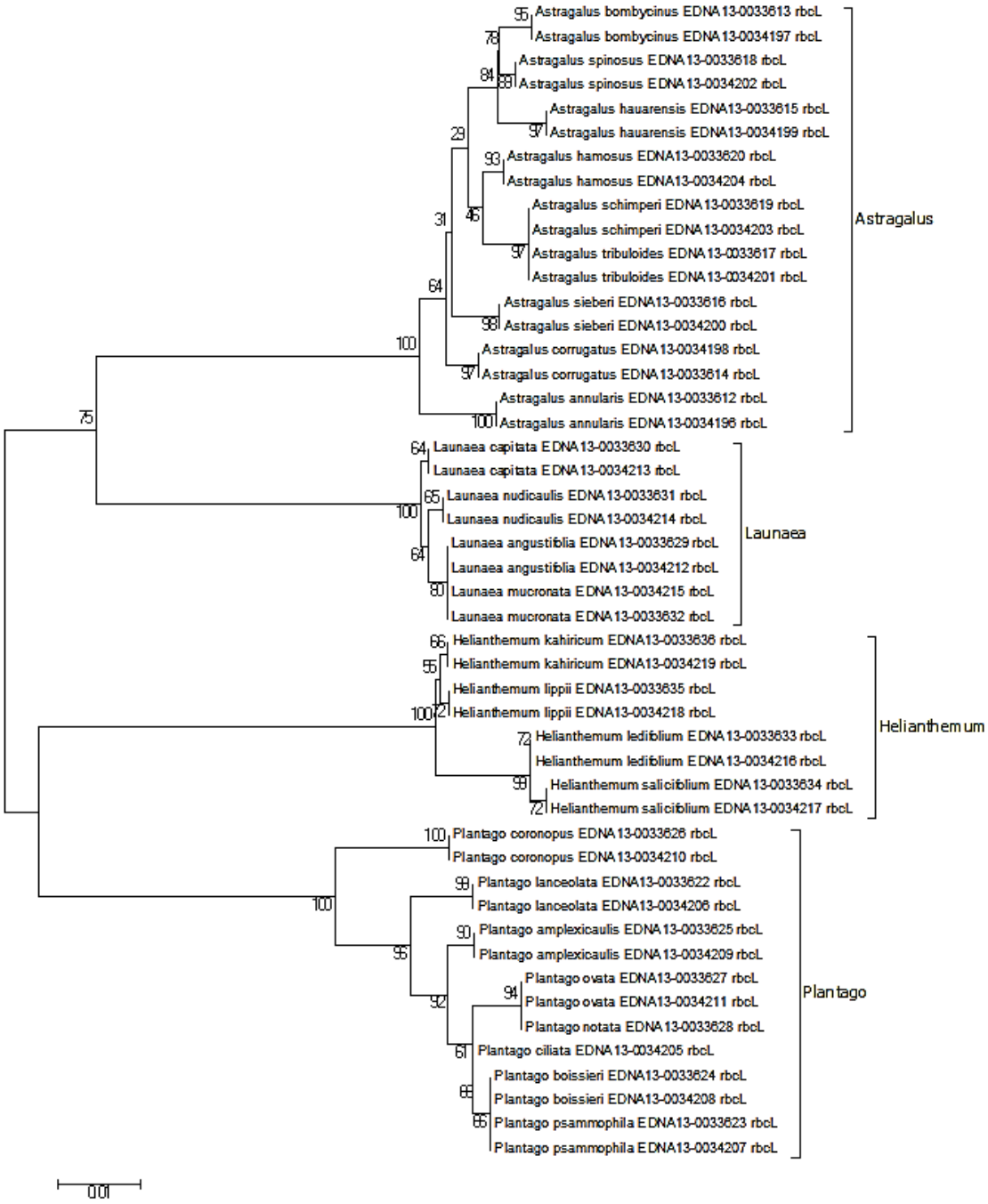


Figure 3.2 Neighbour joining phylograms for *rbcL* barcodes illustrating the four largest genera of the flora of Kuwait (values represent % boot strap support with 1000 replicates)

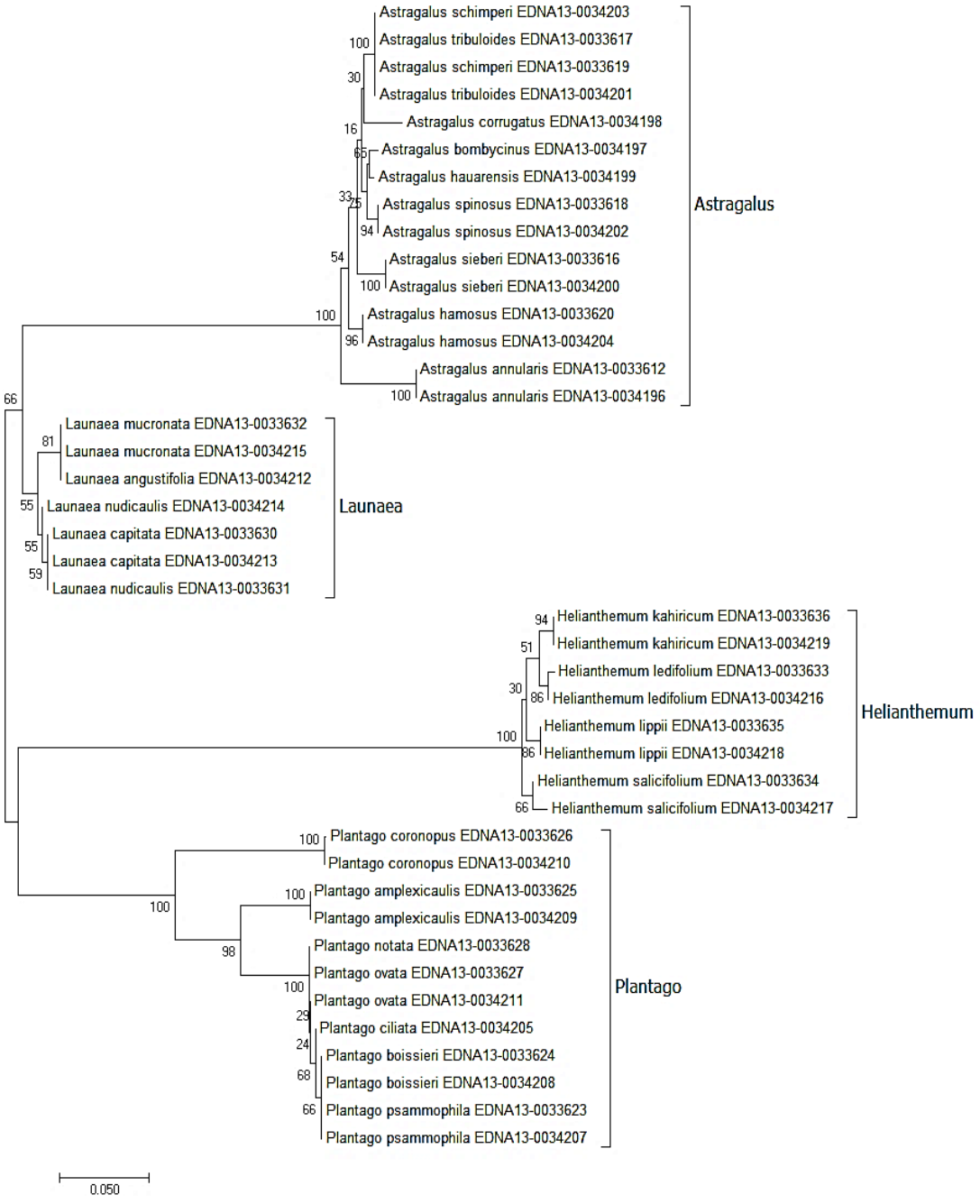


Figure 3.3 Neighbour joining phylograms for *trnL* barcodes illustrating the four largest genera of the flora of Kuwait (values represent % boot strap support with 1000 replicates)

Combined regions tree based analysis:

NJ, MP and ML trees resolved all four genera for combined *rbcL* + ITS2, *rbcL* + *trnL* and *trnL* + ITS2 as monophyletic groups with higher clade support than single regions (% support applying bootstrap, 1000 replicates).

For *rbcL* + ITS2, the combined region resolved the most species-specific clusters of 64 % (for all tree methods NJ, MP, ML) (Table 3.8 and Figure 3.4), followed by *rbcL* + *trnL*, 50 % for NJ and MP (Table 3.8) and the lowest species resolution was 42 % for the combined regions *trnL* + ITS2 using MP and ML tree, while with NJ the resolution was slightly higher 50 % (Table 3.8). Therefore, based on these findings, I choose to explore in more details *rbcL* + ITS2 regions using NJ tree methods.

The combined region *rbcL* + ITS2 showed the greatest species resolution (64 %) and resolved 16 of 25 species into monophyletic clades (Table 3.8 and Figure 3.4). In addition, the combined regions (*rbcL*+ ITS2) resulted in 100 % species-specific resolution for one of the four genera, *Launaea* spp. (Table 3.9 and Figure 3.4). However, the combined regions contained paraphyletic clades which included *Astragalus tribuloides* with *A. schimperi*; *Plantago boissieri* with *P. psammophila*; *P. notata* with *P. ovata* (Figure 3.4).

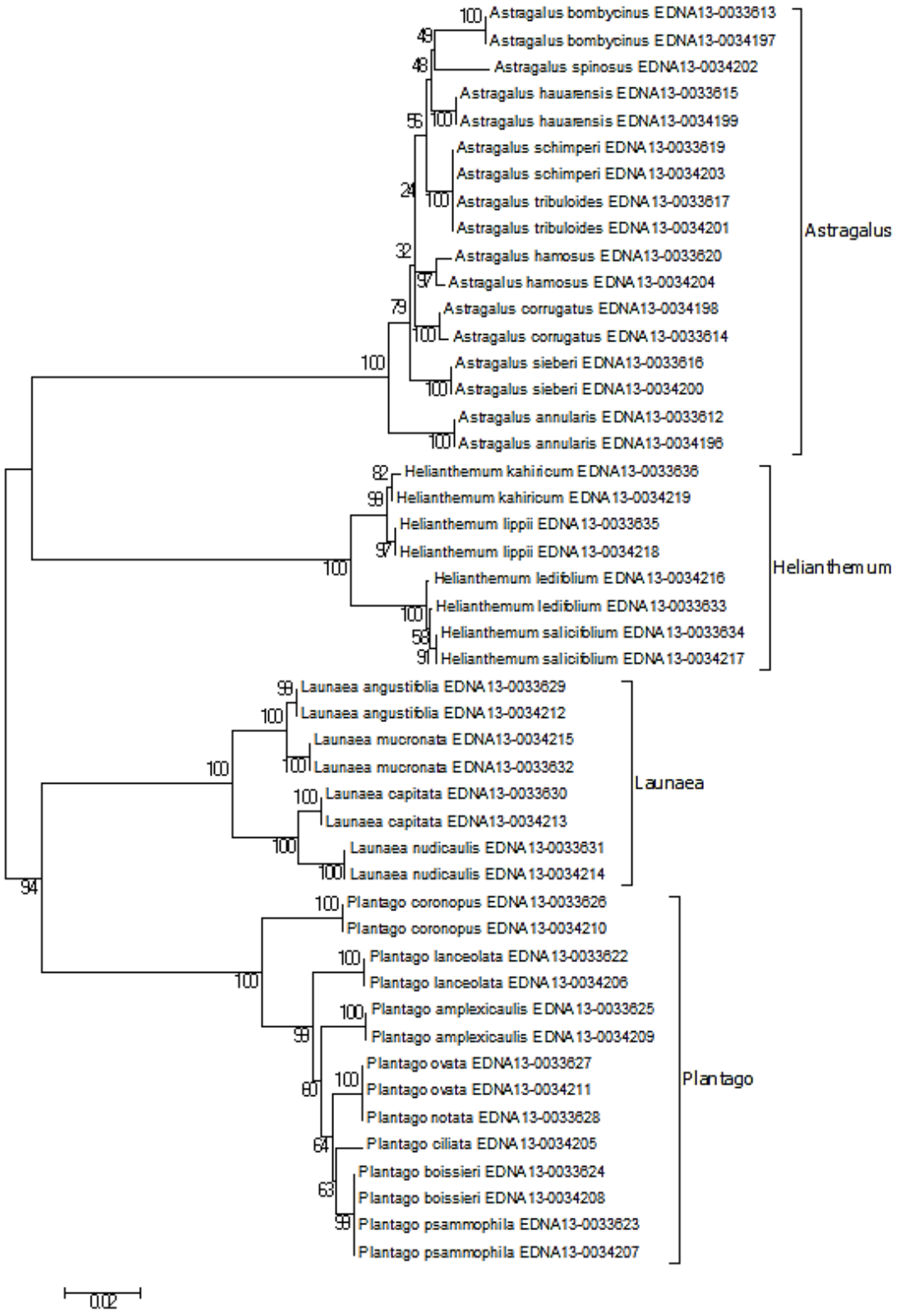


Figure 3.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes illustrating the four largest genera of the flora of Kuwait (values represent % boot strap support with 1000 replicates)

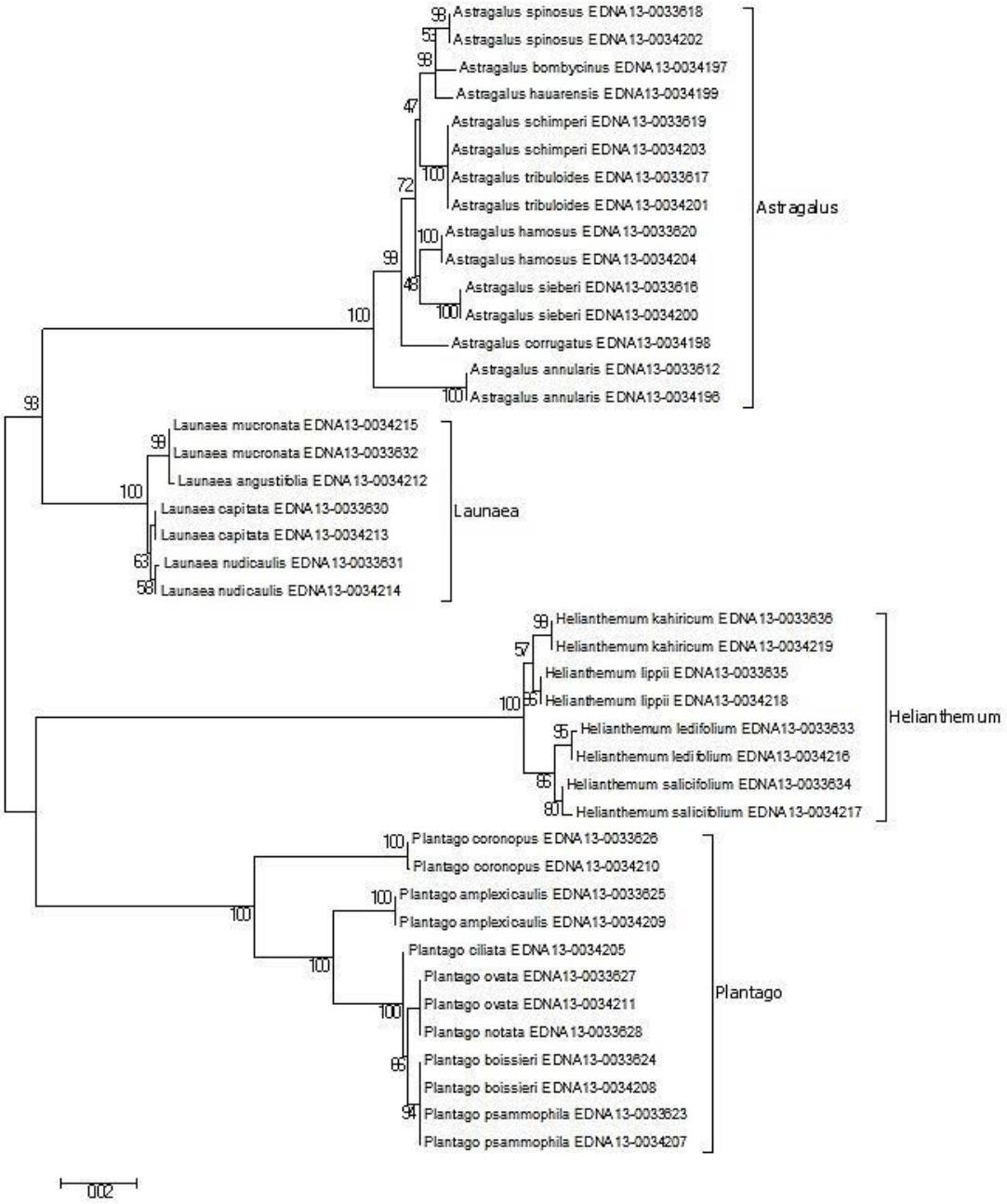


Figure 3.5 Neighbour joining phylograms for combined *rbcL* + *trnL* barcodes illustrating the four largest genera of the flora of Kuwait (values represent % boot strap support with 1000 replicates)

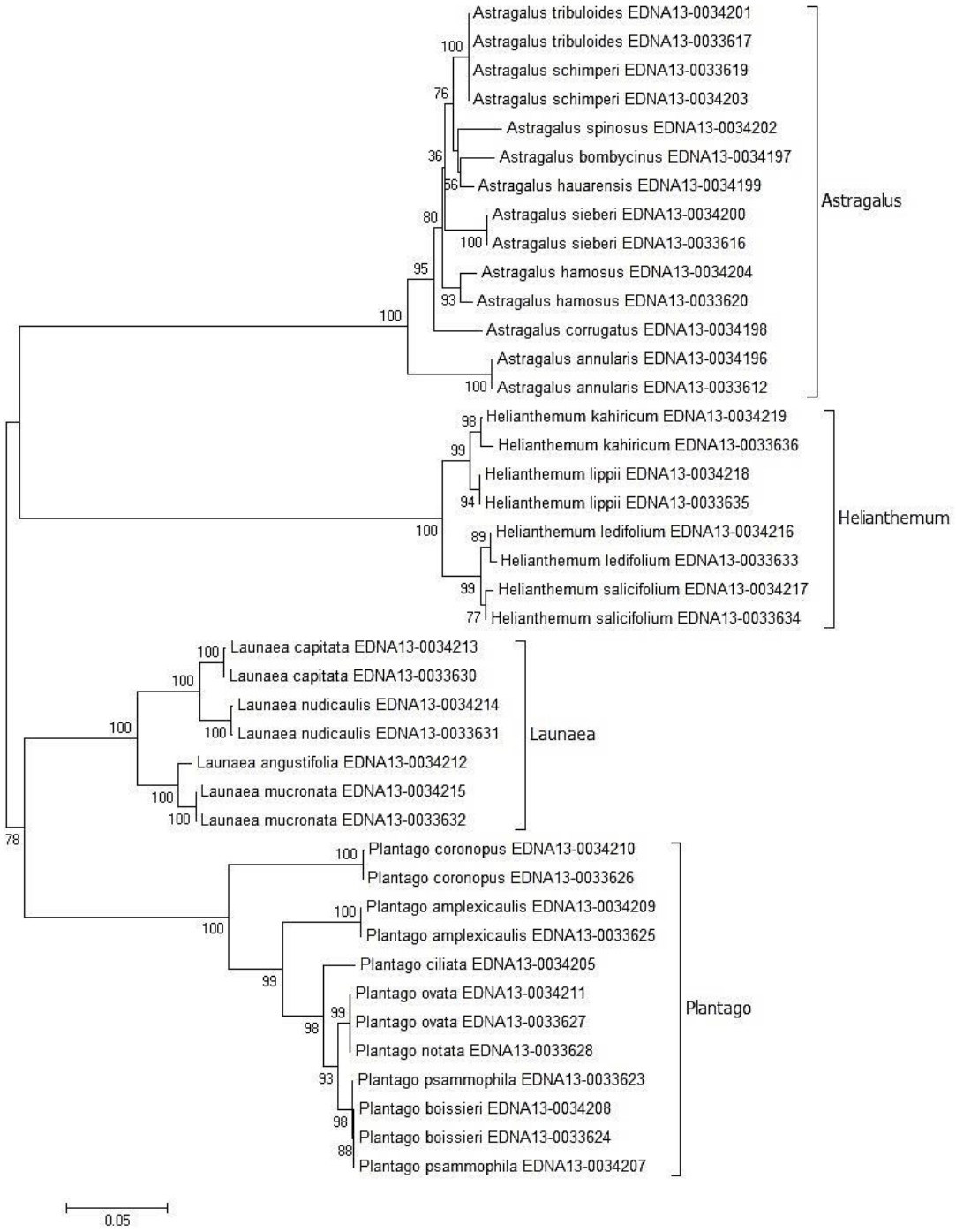


Figure 3.6 Neighbour joining phylograms for *trnL* + ITS2 barcodes illustrating the four largest genera of the flora of Kuwait (values represent % boot strap support with 1000 replicates)

3.4 Discussion

Universality, sequence quality and bidirectional sequence coverage, and levels of species discrimination are essential to consider while evaluating barcode regions (CBOL, 2009). The gene regions *rbcL* and ITS2 showed the highest success in PCR amplification using a single set of primers each, also showed excellent sequence assemblage (98 % each), resolved monophyletic groups using tree-based analysis and resulted in good discrimination power amongst closely related species, for *rbcL* (60 %) and ITS2 (64 %).

3.4.1 DNA recoverability, amplification and sequencing quality

Amongst the five evaluated barcode regions for DNA recoverability, amplification and sequencing success, *rbcL* and ITS2 showed the greatest rate of success (98 % each), followed by *trnL* (88 %). *matK* and *trnH-psbA* excluded from further evaluation due to high failures of amplification (mainly *matK*) and sequencing rates (*trnH-psbA*). Although, three sets of primers were applied for *matK* region, none of the individuals for *Helianthemum* spp. amplified. Such issues with DNA regions and amplification problems are important to detect at early stages to avoid complications later while establishing a DNA reference library. The *matK* region is well documented for having problems due to amplification failure even while testing several pairs of primers (Li et al., 2012; Saarela et al., 2013; Bolson et al., 2015). However, despite amplification problems, *matK* barcodes show great levels of species discrimination amongst closely related groups (CBOL, 2009; Burgess et al., 2011; de Vere et al., 2012).

rbcL and ITS2 showed the best performance and universality with 98 % of the samples successfully amplified and sequenced using only one pair of primers for each. Although, previous studies have used more than one pair of primers to successfully sequence ≥ 80 % of the samples (de Vere et al., 2012; Bafeel et al., 2012; Yao et al., 2010).

rbcL was easily amplified and produced high-quality bidirectional sequences (98 %) which make it an ideal barcode region in agreement with many previously published studies (CBOL, 2009; de Vere et al, 2012; Fazekas et al., 2008; Saarela et al., 2013). *rbcL* was the most conserved sequence compared to the other four regions tested.

ITS2 also showed high amplification, and sequencing success (98 %) and several studies considered it as an ideal barcoding marker (Han et al., 2013; Chen et al., 2010; Yao et al., 2010; Braukmann et al., 2017). There are three major potential problems with using the entire ITS region: 1) fungal contamination, 2) paralogous gene copies, and 3) difficult to amplify and sequence (Hollingsworth, 2011). Using a partial region ITS2 makes it more reliable and easier to amplify and sequence (Han et al., 2013; Luo et al., 2010; Chen et al., 2010).

3.4.2 Discrimination of species

Phylogenetic analysis using tree-based methods is an important approach to determine the DNA region and evaluate its ability to verify whether it can identify and detect species-specific clusters of species from the same genus.

In this study, NJ analysis produced phylogenetic trees with better resolution for all tested barcodes. ITS2 showed the greatest percentage of species-specific clusters (64 %) followed by *rbcL* (60 %). The *rbcL* and ITS2 barcodes alone demonstrated interesting findings on closely related species of 9 *Astragalus* spp. of the flora, by resolving monophyletic clades of 78 % (*rbcL*) and 67 % (ITS2).Two of the *Astragalus* species (*A. schimperi* and *A. tribuloides*) were paraphyletic across all three tested regions (*rbcL*, *trnL*, ITS2) due to sequence similarities and both species belong to the section, Sesamei (Sharawy and Badr, 2014). Sharawy and Badr (2014) managed to differentiate amongst five closely related *Astragalus* spp. in the section Sesamei (including *A. schimperi* and *A. tribuloides*) based on morphological variation and molecular polymorphisim using RAPD and ISSR fingerprinting analyses (Sharawy and Badr, 2014). *Astragalus* is the largest genus of flowering plants, belonging to Fabaceae with over 2500 species and more than 250 sections (Lock and Simpson, 1991; Mabberley, 1997). They are known to be difficult to discriminate because of the

reduced levels of sequence divergence as reported by several publications (Dizkirici et al., 2014; Naderi Safar et al., 2014; Javanmardi et al., 2012). Another paraphyletic relationship was noted for the species *Plantago* (between *P. psammophila* / *P. boisseri* and *P. ovata* / *P. notata*). *Plantago* species comprises 200 species in 19 sections were described in 2 subgenera (*Euplantago* and *Psyllium*) also known to have low levels of variation amongst closely related species which makes it difficult to discriminate even with combined regions given in this study (Tutel et al., 2005; Ronsted et al., 2002).

The findings in the current study agrees with recent published research evaluating ITS2 as a DNA barcode marker for land plants (Braukmann et al., 2017; Kuzmina et al., 2012; Chen et al., 2010; Yao et al., 2010) and for medicinal plants identification (Han et al., 2013; Pang et al., 2013; Zhang et al., 2015), these studies demonstrated ITS2 barcode species resolution variation from 63 % to 92 %.

The evaluation of three combinations of barcode markers demonstrated best species resolution by combining *rbcL* + ITS2 (64 %) and the percentage of bootstrap support for monophyletic clades increased when combined which resulted in higher resolution and discrimination amongst closely related species. *rbcL* + ITS2 resolved one of the four genera into monophyletic clades with 100 % bootstrap support with all individuals of species clustered correctly (*Lanuaea* spp.) while using a single region *rbcL* only resolved it to 50 % resolution. This demonstrates the reliability of combined regions in discriminating closely related species into species-specific clusters.

Exploring DNA barcoding approaches on arid plants in the Arabian Peninsula is rare. A few studies applied it to identify medicinal plants, rare species, and cultivated plants; none have yet applied it to an entire flora. In Saudi Arabia, Bafeel et al. (2012) used *rbcL* marker to study 12 species belonging to diverse families of arid regions and their findings showed very low species resolution (17 %) (Bafeel et al., 2012). Findings in the current study disagree with Bafeel et al. (2012) where *rbcL* region alone managed to resolve up to 60 % of 25 tested species belonging to four families of the arid flora. Another study in the United Arab Emirates, by Enan and Ahmed (2014) applied *matK* and *rpoC1* gene regions on 11 different date palm cultivars (*Phoenix*

dactylifera) and resulted in high species discrimination using *matK* alone (≥ 95 %) followed by *rpoC1* (≥ 88 %) (Enan and Ahmed, 2014).

A number of DNA barcoding publications in the Arabian Peninsula used BLASTn search tool of NCBI database to test for sequences similarities and matches. The findings below will help in understanding how well arid plants sequences match with NCBI database. A study by Al-Hemaid et al., (2015) used ITS for the identification of endemic species to Saudi Arabia, *Echinops mandavillei*, ITS region showed 98 % sequence identity to genus level with *Echinops glaberrimus*, matched from BLASTn search (NCBI database). A similar study based on BLASTn search similarities using total genomic sequence of medicinal plants known in Saudi Arabia (*Nepeta deflersiana*) investigated by Al-Qurainy et al. (2014) showed sequence similarities to the same genera level of the same family (92-97 % for ITS, *rbcL*, *rps16*, *rpoC* and *psbA-trnH*) (Al-Qurainy et al., 2014). A study by Al-Qurainy et al. (2013) used the ITS region to study 5 rare species of Saudi Arabia. This resulted in sequences identifying samples to the genera level but none to species level (Al-Qurainy et al., 2013). These findings indicate that sequences of arid plants not well represented in the NCBI database. The current study shows the ability of *rbcL*, ITS2 and *trnL* barcodes to match a number of sequences with high similarity using NCBI database BLASTn searches (Appendix 3.3). Searching the NCBI taxonomy database by species name and DNA region, ITS/ ITS2 sequences are well represented and banked at NCBI database, showing 18 out of the 25 species of the current study, of which only 6 species matched the NCBI database (Appendix 3.3). Therefore, reliance on BLASTn limits the utility of barcode data to well known and sampled flora and restricts their use on unknown samples or poorly known floras.

3.4.3 Factors limiting the discriminatory power in plant DNA barcodes

The differences in species resolution that results in DNA barcoding failure amongst closely related species are certainly influenced by one or more of the following factors: slow rate of molecular evolution, hybridization, polyploidy, introgression and incomplete sorting of ancestral polymorphisims (Hollingsworth et al., 2011; Kress et al., 2015). These factors are not evenly distributed among the plant groups; therefore,

species resolution level will be reasonably good in some closely related groups and quite weak in others. The rate of species discrimination is expected to be greater when studying a group of plants limited to a geographical region which usually restricts the number of closely related species.

An example of hybridization is found amongst closely related species in the Canadian flora, comprising of 171 families, Salicaceae showed the lowest species discrimination, largely due to the very limited genetic variation among the 90 species of Salix; the lack of variation was noticed in seven plastid regions and was linked to frequent hybridization and incomplete lineage sorting (Percy et al., 2014).

Phylogenetic studies targeting nuclear markers or whole plastids genomes are necessary to understand the process driving the unusual low divergences in Salix. Furthermore, polyploidization (or whole genome duplication) limits the discriminatory power of plant DNA barcodes and commonly found across angiosperms associated with the rearrangements of chromosomes, DNA mutation, duplicate gene deletion, gene expression and epigenetic (Soltis et al., 2009; Adams and Wendel, 2005).

Due to hybridization, polyploidy and low mutation rate in plant plastid genomes, it is not possible to establish a single standard DNA barcoding region to barcode all land plants (Kress et al., 2005); therefore multiple DNA regions with species variation for discriminatory power are necessary to provide adequate information, e.g. Kress and Erickson (2007) recommended the use of two DNA regions, *rbcL* in combination with *trnH-psbA*, which showed an increase of discriminatory power from 79 %, for a single region, up to 88 %, for combining the two regions (Kress and Erickson, 2007).

Another important issue that certainly affects the quality of DNA barcodes and limits its discriminatory power is the use of misidentified voucher specimens which results in the production of misidentified barcodes that could end up in public DNA databases. Sampling DNA material from voucher specimens identified by an expert in the field is recommended to overcome the issue of taxonomic misidentification that possibly leads to barcode failure. Furthermore, DNA extraction from herbarium material requires greater attention than freshly collected material, due to contamination and misidentification, de Vere et al. (2012) found that some orders of

flowering plants did not sequence using herbarium material for either *rbcL* or *matK* regions and recommended the use of fresh plant material to overcome the issue of degraded DNA and / or any contaminants present (de Vere et al., 2012).

The discriminatory power amongst closely related plant species is limited by the size and choice of gene regions. Small portions of the barcode region known as mini-barcodes may be used in place of full-length barcodes to overcome DNA degradation for samples with poor DNA preservation (Little, 2014). Mini-barcodes such as the P6 loop in the *trnL* intron (ca 10-143 bp) for species identification is widely used by ecologists studying highly degraded DNA and using next generation sequencing technologies to assess the diversity of complex environmental samples e.g. studying faecal samples from leaf-feeding monkey (Srivathsan et al., 2015); identifying past vegetation from degraded DNA material amplified from soil samples (Yoccoz et al., 2012); detecting ingested plant DNA in soil-living insect larvae (Staudacher et al., 2011).

Despite its ability in DNA barcoding highly degraded samples, the *trnL* intron is known to have difficulty discriminating species in some families such as Poaceae, Cyperaceae, and Asteraceae (Hiiesalu et al., 2012). Using a DNA reference library of a local flora, the P6 loop provided only 50 % species resolution (Valentini et al., 2009). In another study, the P6 loop was able to identify only 47.2 % of 106 species belonging to the Arctic plant collection using a local DNA database (Taberlet et al., 2007). Although well documented for being powerful to amplify from degraded DNA material, the short region of *trnL* P6 loop provides a low resolution and limited discriminatory power when compared with other barcoding regions ≥ 200 bp, e.g. *rbcL*, *matK*, *trnH-psbA*, and ITS2. Taberlet et al. (2007) evaluated the resolution of *trnL* (UAA) intron (254-767 bp) and the shorter fragment, the P6 loop (10-143 bp) and found that the main limitation of both fragments is the low species resolution, *trnL* intron 67.3 % and the P6 loop 19.5 % (Taberlet et al., 2007).

Exploring another mini-barcode region, the shorter nuclear region of the internal transcribed spacer, ITS (ca. 571-1153) is ITS2 with size ca. 163-311 bp (Chen et al., 2010). ITS2 region is used as a DNA minibarcode candidate for its ability to show

high discriminatory power amogst closely related species, ranging from 71 % to 93 % (Chen et al., 2010; Han et al., 2013; Raveendar et al., 2015; Li et al., 2016; Wang et al., 2016) and also used in phylogenetic studies at both genus and species levels (Schultz et al., 2005; Miao et al., 2008; Hribova et al., 2011; Sharma et al., 2012; Heeg and Wolf, 2015).

Chen et al. (2010) showed high PCR amplification efficiency of ITS2 sequences, 93.8 %, across 400 samples belonging to a wide range of plant taxa represented by 326 species in 98 families including dicots, monocots, gymnosperms, and ferns. Also, ITS2 correctly identified 90.3 % and 99.7 % of the samples using BLAST1 method at the species and genus level, respectively (Chent et al., 2010).

Another study by Luo et al. (2010) showed high species resolution using ITS2 region amongst six evaluated DNA barcoding candidates (*trnH-psbA, ycf5, rpoC1, rbcL*, ITS2 and ITS). The study included 300 samples represented by 192 species belonging to 72 genera of Rutaceae family and recommended the application of ITS2 and *trnH-psbA* for their ability to discriminate closely related species by 89 % and 83 %, respectively (Luo et al., 2010).

Other routinely used barcode regions were discussed in the introduction section such as *rbcL* (the small portion ca. 500 bp of the full-length ca. 1,400 bp, Kress et al., 2005) applied for the ease of amplification, sequencing, and aligning in most land plants and *matK* (the small portion ca. 800 bp of the full-length ca 1,500 bp, Li et al., 2011) showing great discriminatory power, despite their high amplification and sequencing failure (CBOL, 2009).

In the meanwhile, many drawbacks and problems with DNA barcodes cannot be neglected, including the failure of amplification and sequencing, difficulties in finding universal primers, lack of barcoding gap, hybridization and introgression in some plant groups. Also, finding a single DNA barcode region for land plants is still a challenge, and the necessity of applying a combined two or more barcode regions supports the level of discriminatory power amongst closely related species.

3.4.4. Conclusion

The main objective of this chapter is to evaluate five DNA barcodes to provide a practical molecular identification tool for establishing a DNA reference library of the flora of Kuwait in the following chapter. The barcoding region *rbcL* was easy to amplify and provided high-quality sequences; also, the region demonstrated a good resolution of species discrimination amongst closely related plant groups which makes it a practical barcoding region to apply. ITS2 region showed the highest levels of species discrimination amongst tested regions. Also, in an attempt to support the discriminatory power amongst closely related species, combining *rbcL* with ITS2 barcodes demonstrated greater support of monophyletic clades which increases its reliability in the identification process.

Based on the findings in the current chapter, the choice lies on the use of combining *rbcL* and ITS2 barcode regions for the molecular identification of species and building a DNA reference library of the flora of Kuwait, as it demonstrated the highest level of identification success and a better resolution of species-specific clusters in phylogenetic trees generated by Neighbour Joining method. Considering the small size of the flora of Kuwait (ca 400 species) the proposed molecular regions *rbcL* + ITS2 can assist in identifying unknown individuals of the flora mainly to species level by developing a localized barcoding library. This method can serve efficiently in plant molecular identification of the local flora and in future can be introduced in vegetation monitoring and large-scale ecological surveys in an attempt to conserve the biodiversity of the flora.

Appendix 3.1 List of specimens from which DNA was extracted and tested on four genera of the flora of Kuwait

Species	EDNA Numbers	Collection ID	Collector and number	Year	Collection type	Locality / region
Astragalus annularis	EDNA13-0033612	KUTH179	M Al Dosari MD5757	2005	Herbarium	Al Wafra Farm
Astragalus annularis	EDNA13-0034196	KUTH180	M Al Dosari MD2535	1997	Herbarium	Al Nuwaseeb border station
Astragalus bombycinus	EDNA13-0034197	MTA470	M Abdullah MTA470	2013	Fresh	Al Salmi near border
Astragalus bombycinus	EDNA13-0033613	KUTH182	M Al Dosari MD5537	2004	Herbarium	Failaika Island
Astragalus corrugatus	EDNA13-0034198	MTA205	M Abdullah MTA205	2013	Fresh	Kabd KISR Station
Astragalus corrugatus	EDNA13-0033614	KUTH184	M Al Dosari MD5034	2001	Herbarium	Al Salmi near border
Astragalus hamosus	EDNA13-0033620	MTA281	M Abdullah MTA281	2013	Fresh	Nature reserve SSNR
Astragalus hamosus	EDNA13-0034204	MTA280	M Abdullah MTA280	2013	Fresh	Sabah Al Ahmad Natural Reserve
Astragalus hauarensis	EDNA13-0033615	KUTH186	R Halwagy RH792	1971	Herbarium	Sulaibiya station
Astragalus hauarensis	EDNA13-0034199	KUTH187	M Al Dosari MD1097	1990	Herbarium	Jal Az-Zor
Astragalus schimperi	EDNA13-0034203	MTA376	M Abdullah MTA376	2013	Fresh	Doha outside Entertainment City
Astragalus schimperi	EDNA13-0033619	KUTH189	KT Mathew KTM3380	1998	Herbarium	Al Abdali
Astragalus sieberi	EDNA13-0034200	MTA461	M Abdullah MTA461	2013	Fresh	Al Salmi near border
Astragalus sieberi	EDNA13-0033616	KUTH190	L Boulos LB18053	1993	Herbarium	Jal Az-Zor
Astragalus spinosus	EDNA13-0033618	MTA454	M Abdullah MTA454	2013	Fresh	Al Salmi near border
Astragalus spinosus	EDNA13-0034202	MTA212	M Abdullah MTA212	2013	Fresh	Sabah Al Ahmad Natural Reserve
Astragalus tribuloides	EDNA13-0034201	MTA345	M Abdullah MTA345	2013	Fresh	Doha outside Entertainment City
Astragalus tribuloides	EDNA13-0033617	KUTH192	M Al Dosari MD5653	2005	Herbarium	Kabd KISR Station
Helianthemum kahiricum	EDNA13-0033636	KUTH052	I Ibrahim (IB1045)	1990	Herbarium	Um Al Rimam
Helianthemum kahiricum	EDNA13-0034219	KUTH053	KT Mathew KTM4754	2000	Herbarium	Um Al Rimam
Helianthemum ledifolium	EDNA13-0033633	KUTH054	M Al Dodari MD4074	1999	Herbarium	Al Salmi near border
Helianthemum ledifolium	EDNA13-0034216	KUTH55	M Al Dosari MD6438	2009	Herbarium	Al Abdali
Helianthemum lippii	EDNA13-0033635	MTA451	M Abdullah MTA451	2013	Fresh	Al Salmi near border
Helianthemum lippii	EDNA13-0034218	MTA371	M Abdullah MTA371	2013	Fresh	Doha outside Entertainment City

Species	EDNA Number	Collection ID	Collector and number	Year	Collection type	Locality / region
Helianthemum salicifolium	EDNA13-0033634	KUTH057	M Al Dosari MD4572	2000	Herbarium	Al Salmi near border
Helianthemum salicifolium	EDNA13-0034217	KUTH58	R Halwagy RH1057	1972	Herbarium	Um Gudair
Launaea angustifolia	EDNA13-0033629	KUTH092	R Halwagy RH1009	1972	Herbarium	Al-Khafji rd
Launaea angustifolia	EDNA13-0034212	KUTH093	KT Mathew KTM2888	1996	Herbarium	Al Mutlaa
Launaea capitata	EDNA13-0033630	KUTH095	M Al Dosari MD3773	1999	Herbarium	Al Salmi near border
Launaea capitata	EDNA13-0034213	KUTH96	M AL Dosari MD2561	1997	Herbarium	Al Nuwaseeb border station
Launaea mucronata	EDNA13-0033632	MTA383	M Abdullah MTA383	2013	Fresh	Failaika Island
Launaea mucronata	EDNA13-0034215	MTA601	M Abdullah MTA601	2013	Fresh	Um Niqa
Launaea nudicaulis	EDNA13-0034214	MTA467	M Abdullah MTA467	2013	Fresh	Al Salmi near border
Launaea nudicaulis	EDNA13-0033631	KUTH098	KT Mathew KTM5473	2007	Herbarium	Nature reserve SSNR
Plantago amplexicaulis	EDNA13-0033625	MTA503	M Abdullah MTA503	2013	Fresh	Al Abdali
Plantago amplexicaulis	EDNA13-0034209	MTA244	M Abdullah MTA244	2013	Fresh	Kabd KISR Station
Plantago boissieri	EDNA13-0033624	MTA305	M Abdullah MTA305	2013	Fresh	Al Liyah
Plantago boissieri	EDNA13-0034208	MTA200	M Abdullah MTA200	2013	Fresh	Nature reserve SSNR
Plantago ciliata	EDNA13-0033621	KUTH243	M Al Dodari MD1768	1996	Herbarium	Al Salmi near border
Plantago ciliata	EDNA13-0034205	KUTH244	KT Mathew KTM4424	1999	Herbarium	Kathma
Plantago coronopus	EDNA13-0033626	MTA388	M Abdullah MTA388	2013	Fresh	E Kuwait: Failaika Island
Plantago coronopus	EDNA13-0034210	MTA499	M Abdullah MTA499	2013	Fresh	N Kuwait: Al Abdali
Plantago lanceolata	EDNA13-0033622	KUTH246	M Al Dosari MD5229	2001	Herbarium	Khaldiyah
Plantago lanceolata	EDNA13-0034206	KUTH247	M Al Dosari MD3963	1999	Herbarium	Gulf street
Plantago notata	EDNA13-0033628	KUTH249	R Halwagy RH74/10	1974	Herbarium	Al Shaqq khabrat Um Omara
Plantago ovata	EDNA13-0033627	MTA391	M Abdullah MTA391	2013	Fresh	E Kuwait: Failaika Island
Plantago ovata	EDNA13-0034211	MTA298	M Abdullah MTA298	2013	Fresh	W Kuwait : Al Liyah
Plantago psammophila	EDNA13-0033623	KUTH250	M AL Dosari MD2019	1997	Herbarium	Al-Khuwaisat
Plantago psammophila	EDNA13-0034207	KUTH251	L Boulos LB18124	1993	Herbarium	W of Al-Jahra

Appendix 3.2 Specimen information with BOLD and GenBank accessions

			GenBank accessions		
Species	**EDNA Number**	**BOLD ID**	***rbcL***	**ITS2**	***trnL***
Astragalus annularis	EDNA13-0033612	KWT001-17	KY951666	KY951573	KY951621
Astragalus annularis	EDNA13-0034196	KWT002-17	KY951667	KY951574	KY951622
Astragalus bombycinus	EDNA13-0034197	KWT003-17	KY951668	KY951575	KY951623
Astragalus bombycinus	EDNA13-0033613	KWT004-17	KY951669	KY951576	KY951624
Astragalus corrugatus	EDNA13-0034198	KWT005-17	KY951671	KY951578	KY951626
Astragalus corrugatus	EDNA13-0033614	KWT006-17	KY951670	KY951577	NA
Astragalus hamosus	EDNA13-0033620	KWT007-17	KY951672	KY951579	KY951627
Astragalus hamosus	EDNA13-0034204	KWT008-17	KY951673	KY951580	KY951628
Astragalus hauarensis	EDNA13-0033615	KWT009-17	KY951674	KY951581	NA
Astragalus hauarensis	EDNA13-0034199	KWT010-17	KY951675	KY951582	KY951630
Astragalus schimperi	EDNA13-0034203	KWT011-17	KY951676	KY951583	KY951631
Astragalus schimperi	EDNA13-0033619	KWT012-17	KY951677	KY951584	KY951632
Astragalus sieberi	EDNA13-0034200	KWT013-17	KY951679	KY951586	KY951634
Astragalus sieberi	EDNA13-0033616	KWT014-17	KY951678	KY951585	KY951633
Astragalus spinosus	EDNA13-0033618	KWT015-17	KY951680	KY951587	KY951635
Astragalus tribuloides	EDNA13-0034201	KWT016-17	KY951681	KY951588	KY951636
Astragalus tribuloides	EDNA13-0033617	KWT017-17	KY951682	KY951589	KY951637
Helianthemum kahiricum	EDNA13-0033636	KWT018-17	KY951684	KY951591	KY951639
Helianthemum kahiricum	EDNA13-0034219	KWT019-17	KY951683	KY951590	KY951638
Helianthemum ledifolium	EDNA13-0033633	KWT020-17	KY951685	KY951592	KY951640
Helianthemum ledifolium	EDNA13-0034216	KWT021-17	KY951686	KY951593	KY951641
Helianthemum lippii	EDNA13-0033635	KWT022-17	KY951688	KY951595	KY951643
Helianthemum lippii	EDNA13-0034218	KWT023-17	KY951687	KY951594	KY951642
Helianthemum salicifolium	EDNA13-0033634	KWT024-17	KY951689	KY951596	KY951644
Helianthemum salicifolium	EDNA13-0034217	KWT025-17	KY951690	KY951597	KY951645
Launaea angustifolia	EDNA13-0033629	KWT026-17	KY951691	KY951598	NA
Launaea angustifolia	EDNA13-0034212	KWT027-17	KY951692	KY951599	KY951647
Launaea capitata	EDNA13-0033630	KWT028-17	KY951693	KY951600	KY951648
Launaea capitata	EDNA13-0034213	KWT029-17	KY951694	KY951601	KY951649
Launaea mucronata	EDNA13-0033632	KWT030-17	KY951696	KY951603	KY951651
Launaea mucronata	EDNA13-0034215	KWT031-17	KY951695	KY951602	KY951650
Launaea nudicaulis	EDNA13-0034214	KWT032-17	KY951697	KY951604	KY951652
Launaea nudicaulis	EDNA13-0033631	KWT033-17	KY951698	KY951605	KY951653
Plantago amplexicaulis	EDNA13-0033625	KWT034-17	KY951699	KY951606	KY951654
Plantago amplexicaulis	EDNA13-0034209	KWT035-17	KY951700	KY951607	KY951655
Plantago boissieri	EDNA13-0033624	KWT036-17	KY951701	KY951608	KY951656
Plantago boissieri	EDNA13-0034208	KWT037-17	KY951702	KY951609	KY951657
Plantago ciliata	EDNA13-0033621	KWT038-17	KY951703	KY951610	NA
Plantago ciliata	EDNA13-0034205	KWT039-17	NA	KY951611	KY951658
Plantago coronopus	EDNA13-0033626	KWT040-17	KY951704	KY951612	KY951659
Plantago coronopus	EDNA13-0034210	KWT041-17	KY951705	KY951613	KY951660

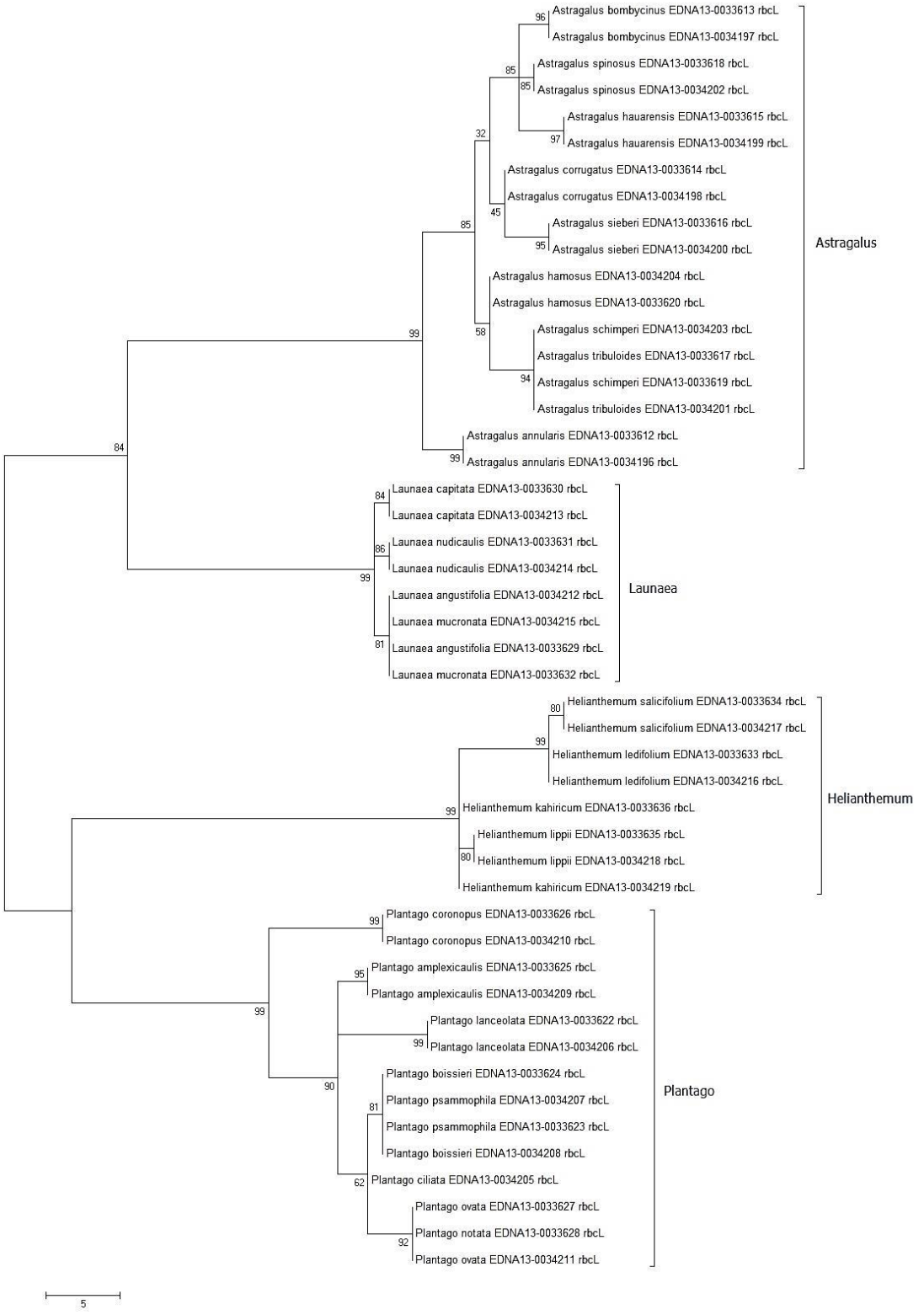
			GenBank accessions		
Species	**EDNA Number**	**BOLD ID**	***rbcL***	**ITS2**	***trnL***
Plantago lanceolata	EDNA13-0033622	KWT042-17	KY951707	KY951615	NA
Plantago lanceolata	EDNA13-0034206	KWT043-17	KY951706	KY951614	NA
Plantago notata	EDNA13-0033628	KWT044-17	KY951708	KY951616	KY951661
Plantago ovata	EDNA13-0033627	KWT045-17	KY951709	KY951617	KY951662
Plantago ovata	EDNA13-0034211	KWT046-17	KY951710	KY951618	KY951663
Plantago psammophila	EDNA13-0033623	KWT047-17	KY951711	KY951619	KY951664
Plantago psammophila	EDNA13-0034207	KWT048-17	KY951712	KY951620	KY951665

Appendix 3.3. GenBank accessions showing twentyfive species belonging to the four largest genera of the flora of Kuwait. GenBank accessions in bold showing barcoded sequence match using BLASTn.

		GenBank accessions		
Family	**Species**	***rbcL***	**ITS2**	***trnL***
Fabaceae	*Astragalus annularis*	NA	AB051912.1	**AB485924.1**
Fabaceae	*Astragalus bombycinus*	NA	AB051929.1	NA
Fabaceae	*Astragalus corrugatus*	NA	HQ613378.1 L10775.1 L10774.1	NA
Fabaceae	*Astragalus hamosus*	KX421132.1	AB051936.1 L10779.1	**AB485945.1**
Fabaceae	*Astragalus hauarensis*	NA	NA	NA
Fabaceae	*Astragalus schimperi*	NA	NA	NA
Fabaceae	*Astragalus sieberi*	NA	KF815519.1	NA
Fabaceae	*Astragalus spinosus*	NA	KF805110.1	NA
Fabaceae	*Astragalus tribuloides*	NA	NA	**AB485929.1**
Cistaceae	*Helianthemum kahiricum*	**FJ492029.1**	**GU327666.1** GU327667.1	FJ492003.1
Cistaceae	*Helianthemum ledifolium*	**FJ492023.1**	NA	FJ491995.1
Cistaceae	*Helianthemum lippii*	NA	KF805114.1	NA
Cistaceae	*Helianthemum salicifolium*	NA	NA	NA
Asteraceae	*Launaea angustifolia*	NA	NA	NA
Asteraceae	*Launaea capitata*	NA	**KF850550.1**	NA
Asteraceae	*Launaea mucronata*	NA	KF805121.1	NA
Asteraceae	*Launaea nudicaulis*	NA	L48148.1 L48147.1	NA
Plantaginaceae	*Plantago amplexicaulis*	NA	**AY101900.1** KF850592.1	**AY101954.1**
Plantaginaceae	*Plantago boissieri*	NA	KF815500.1	NA
Plantaginaceae	*Plantago ciliata*	NA	AY101906.1	NA
Plantaginaceae	*Plantago coronopus*	**AJ389600.1** HM850263.1 HQ593827.1 HQ593826.1	**KX167680.1** AY101882.1 AJ548987.1 HQ593833.1	**AY101937.1** AF486419.1 HQ593818.1 HQ593817.1
Plantaginaceae	*Plantago lanceolata*	**KT695487.1** KJ204385.1 HQ644063.1 L36454.1	**KF454409.1** AJ548984.1 AB281171.1 KP278481.1	HM590326.1 AY101952.1 KU600401.1 KU600344.1
Plantaginaceae	*Plantago notata*	NA	NA	NA
Plantaginaceae	*Plantago ovata*	GQ248675.1 EF590563.1	**KX534375.1** AJ548973.1 EU347721.1 AY101903.1	**AY101957.1** EU036271.1 EU036270.1 EU036269.1
Plantaginaceae	*Plantago psammophila*	NA	AB051913.1	NA

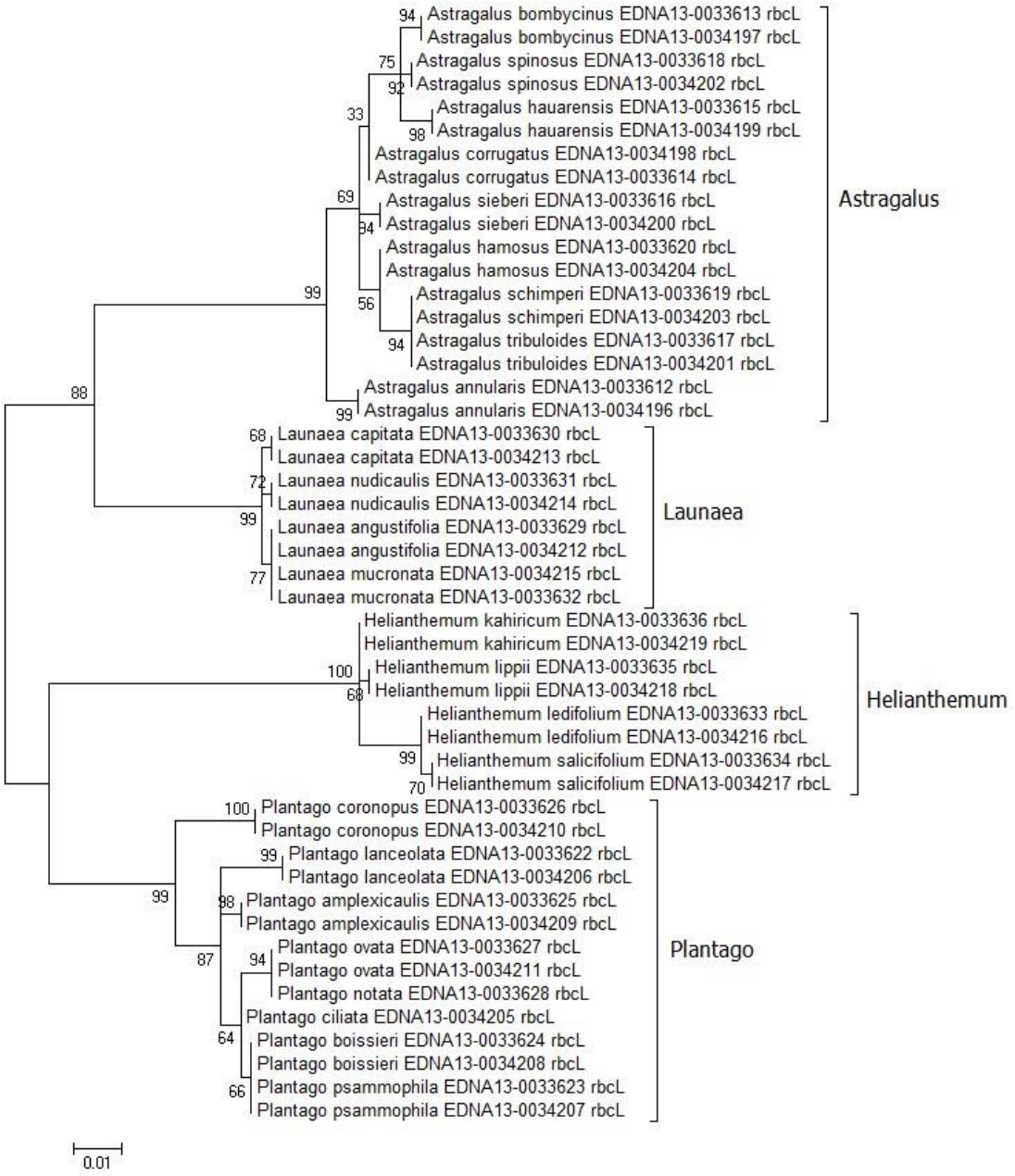
NA – accession not available in GenBank

Appendix 3.4 MP and ML trees for single and combined barcode regions



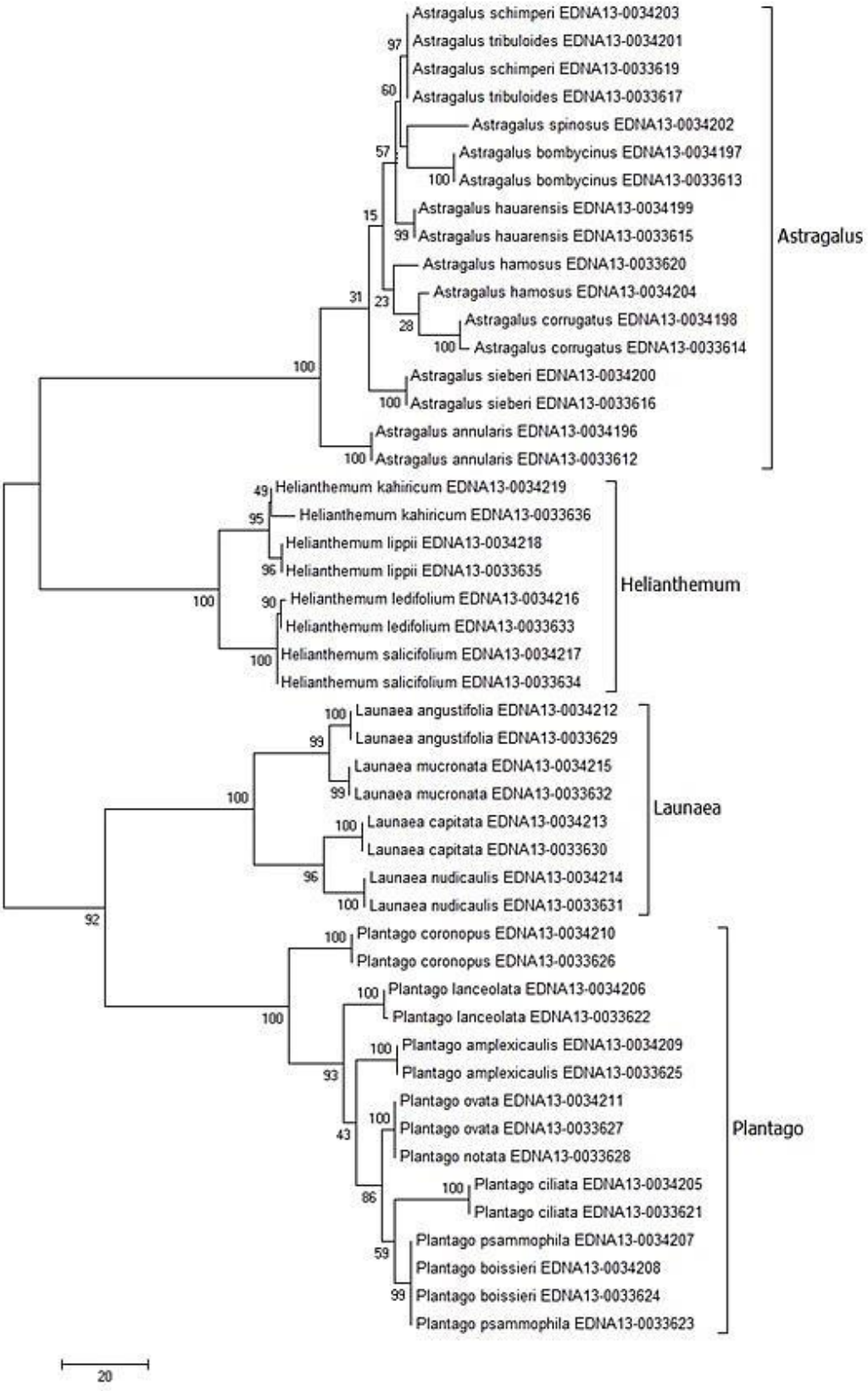
Maximum parsimony phylograms for *rbcL* barcodes
(values represent % boot strap support with 1000 replicates)

(Continued)

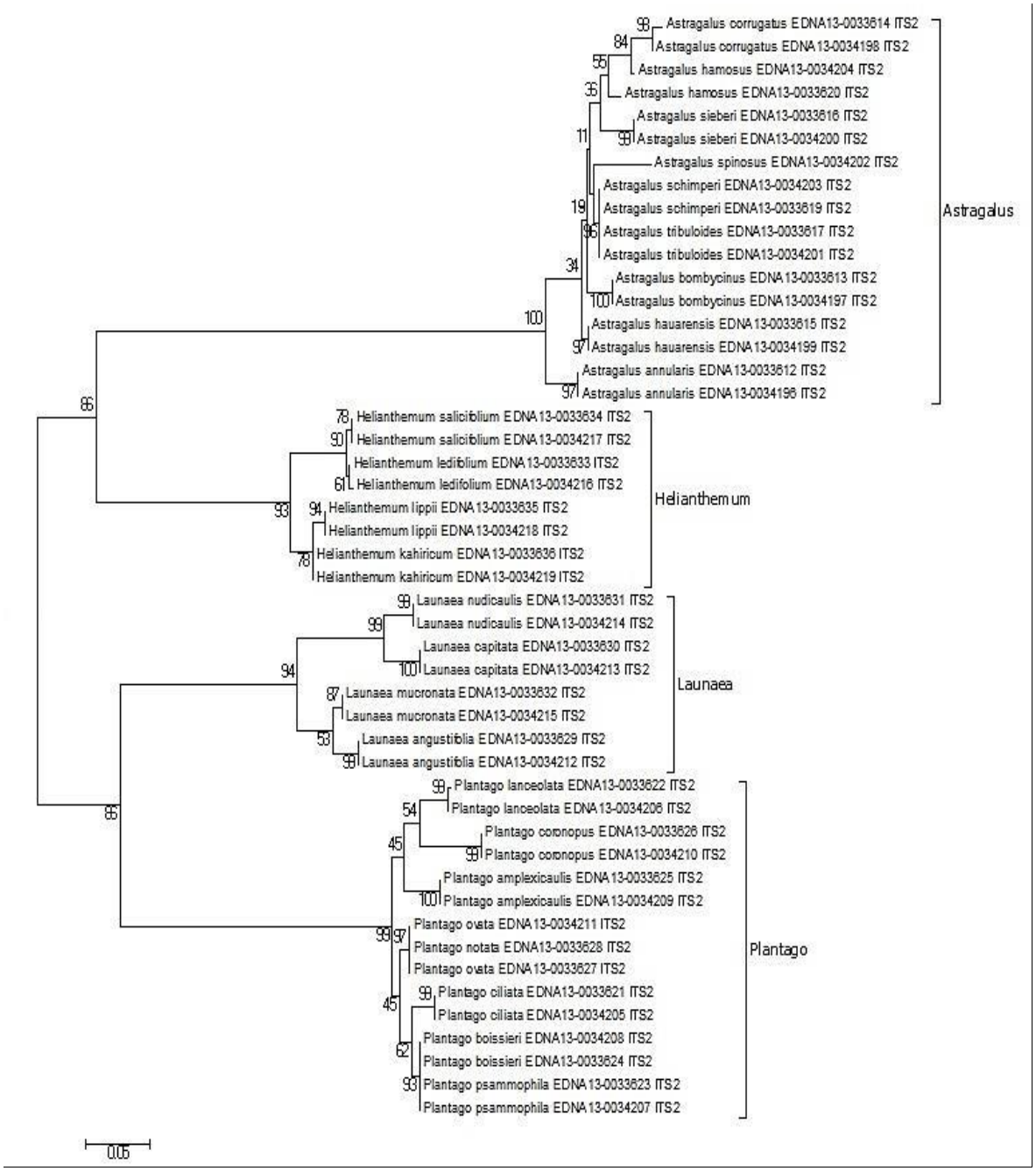


Maximum likelihood phylograms for *rbcL* barcodes
(values represent % boot strap support with 1000 replicates)

(Continued)

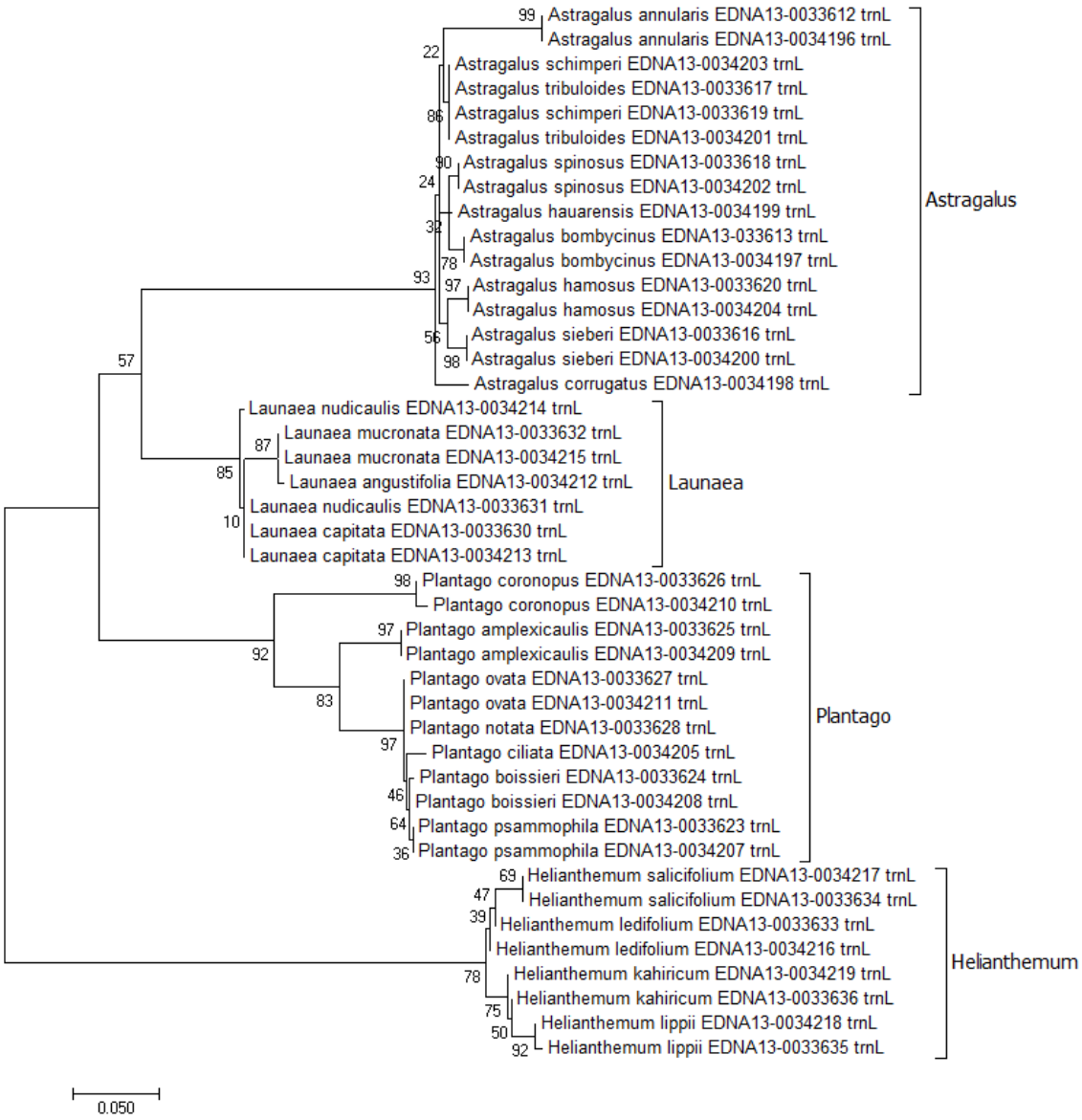


Maximum parsimony phylograms for ITS2 barcodes
(values represent % boot strap support with 1000 replicates) *(Continued)*



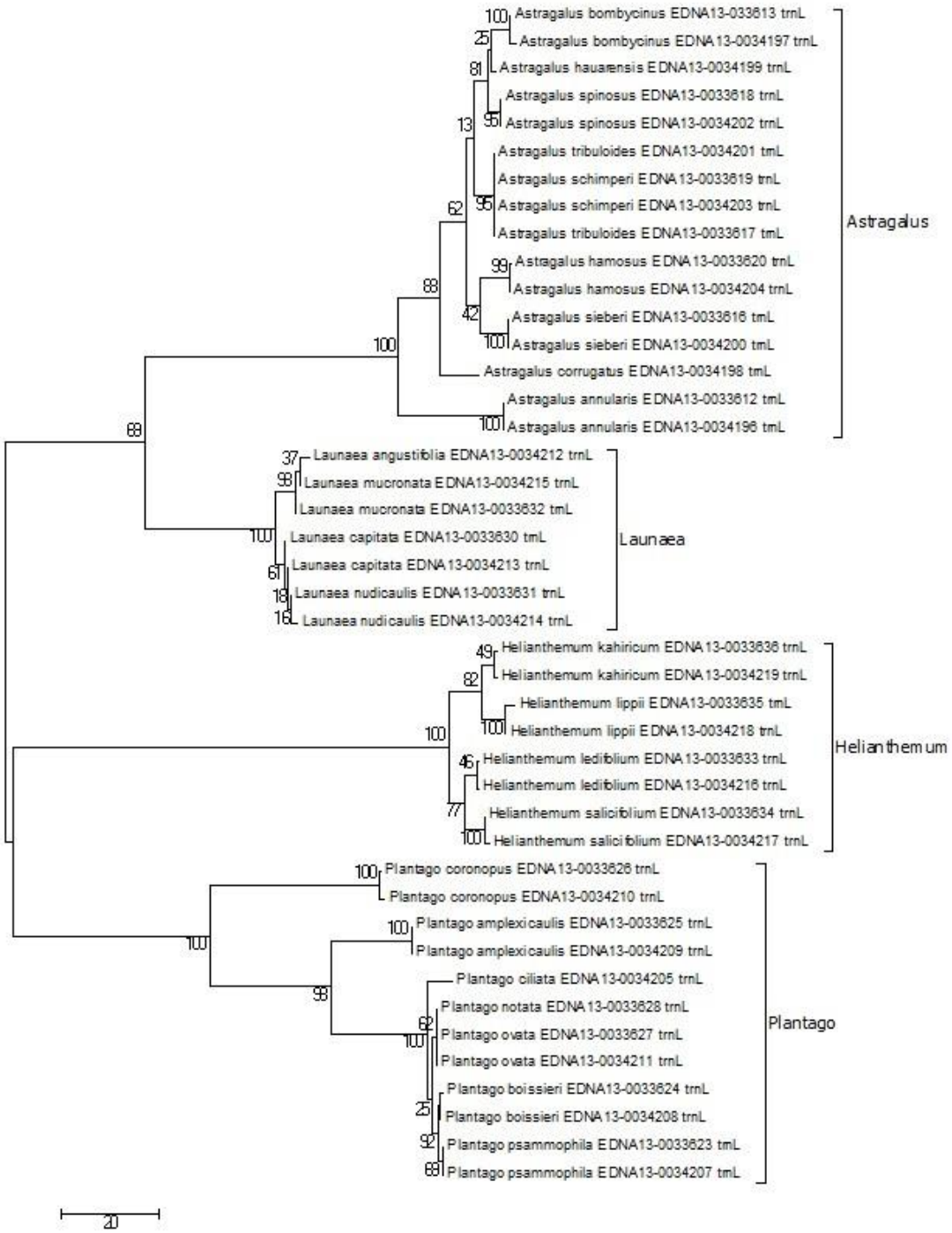
Maximum Likelihood for ITS2 barcodes
(values represent % boot strap support with 1000 replicates)

(Continued)



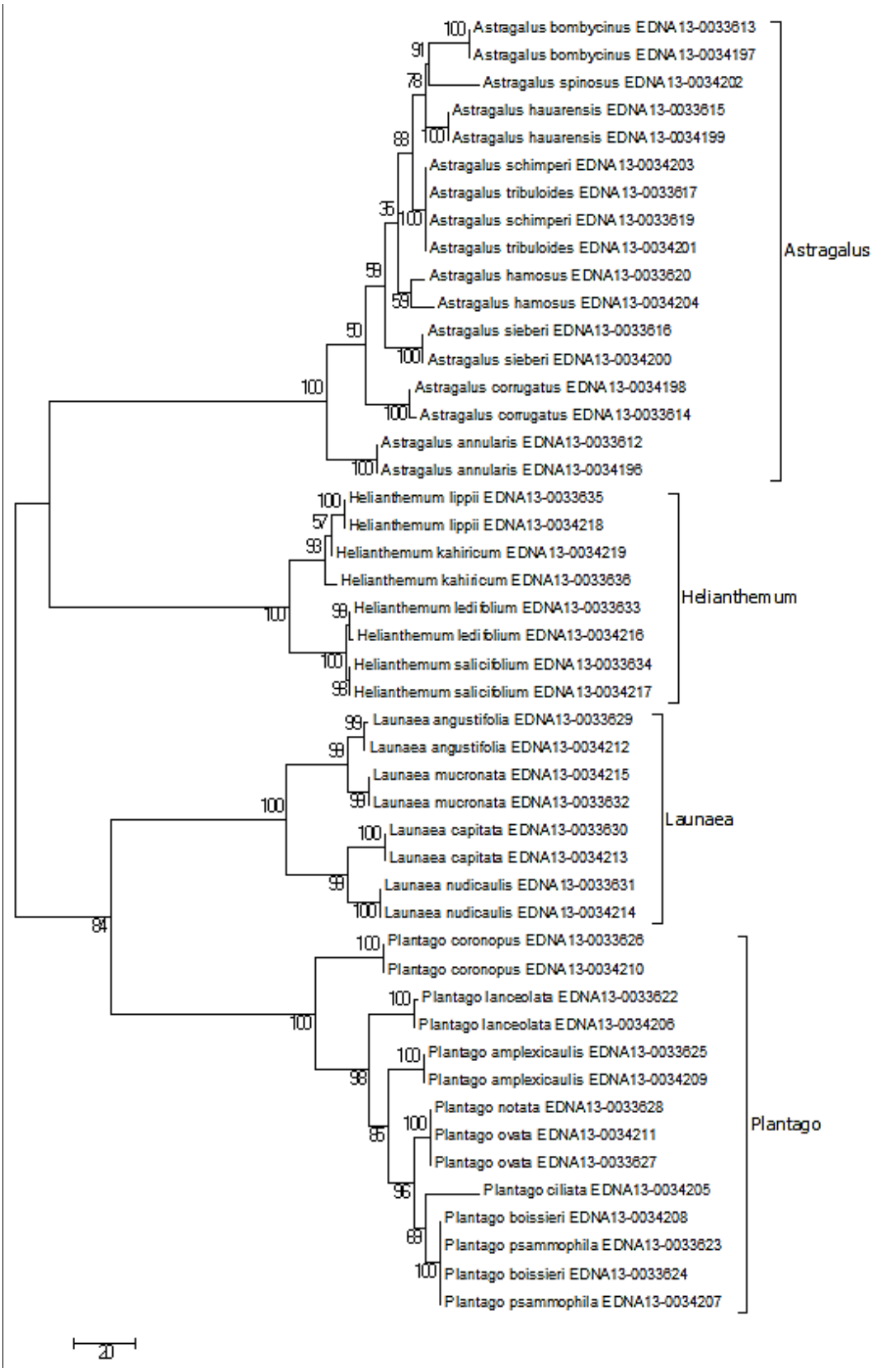
Maximum likelihood phylograms for *trnL* barcodes
(values represent % boot strap support with 1000 replicates)

(Continued)



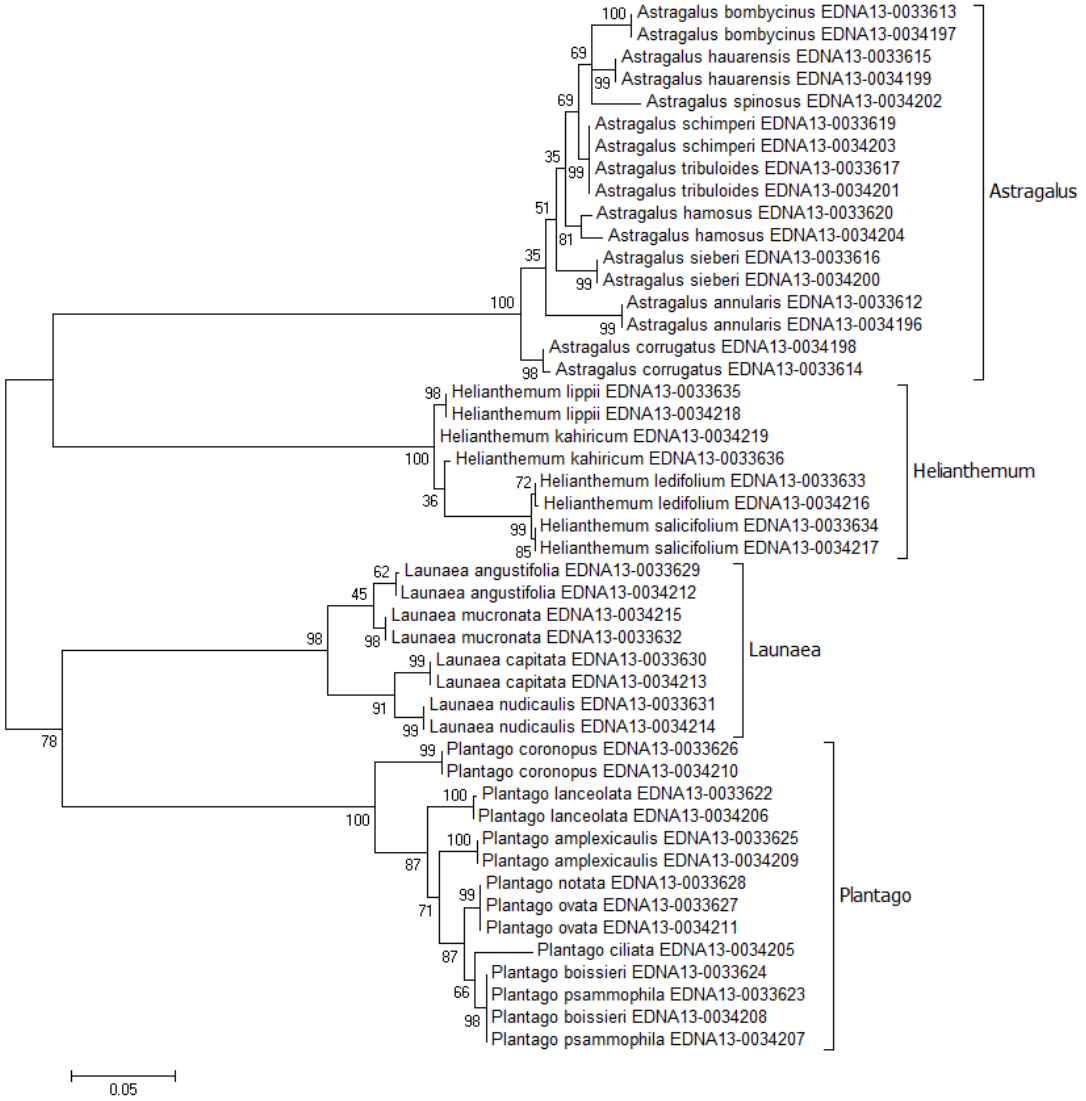
Maximum parsimony phylograms for *trnL* barcodes
(values represent % boot strap support with 1000 replicates)

(Continued)



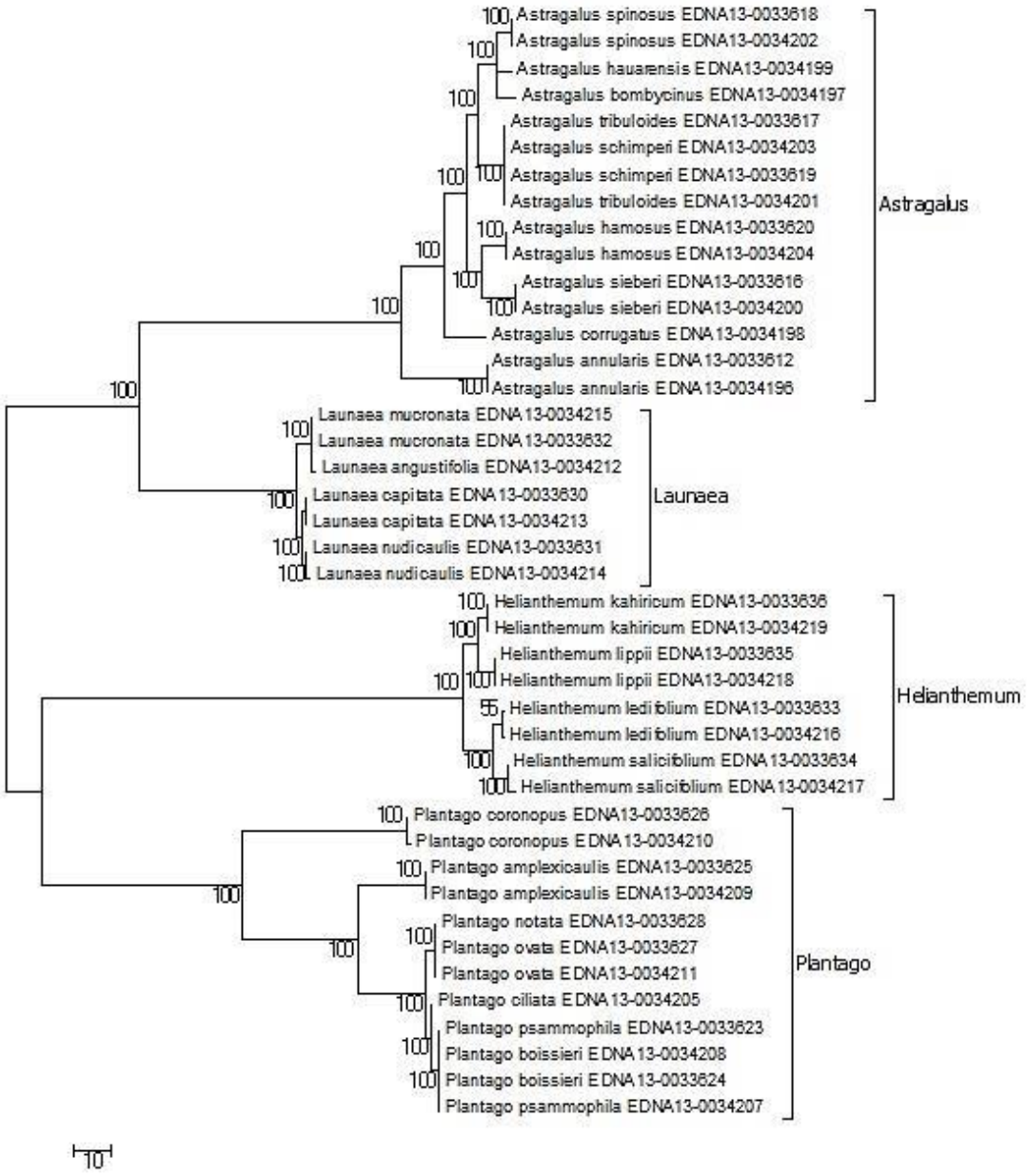
Maximum parsimony phylograms for combined *rbcL* + ITS2 barcodes (values represent % boot strap support with 1000 replicates)

(Continued)



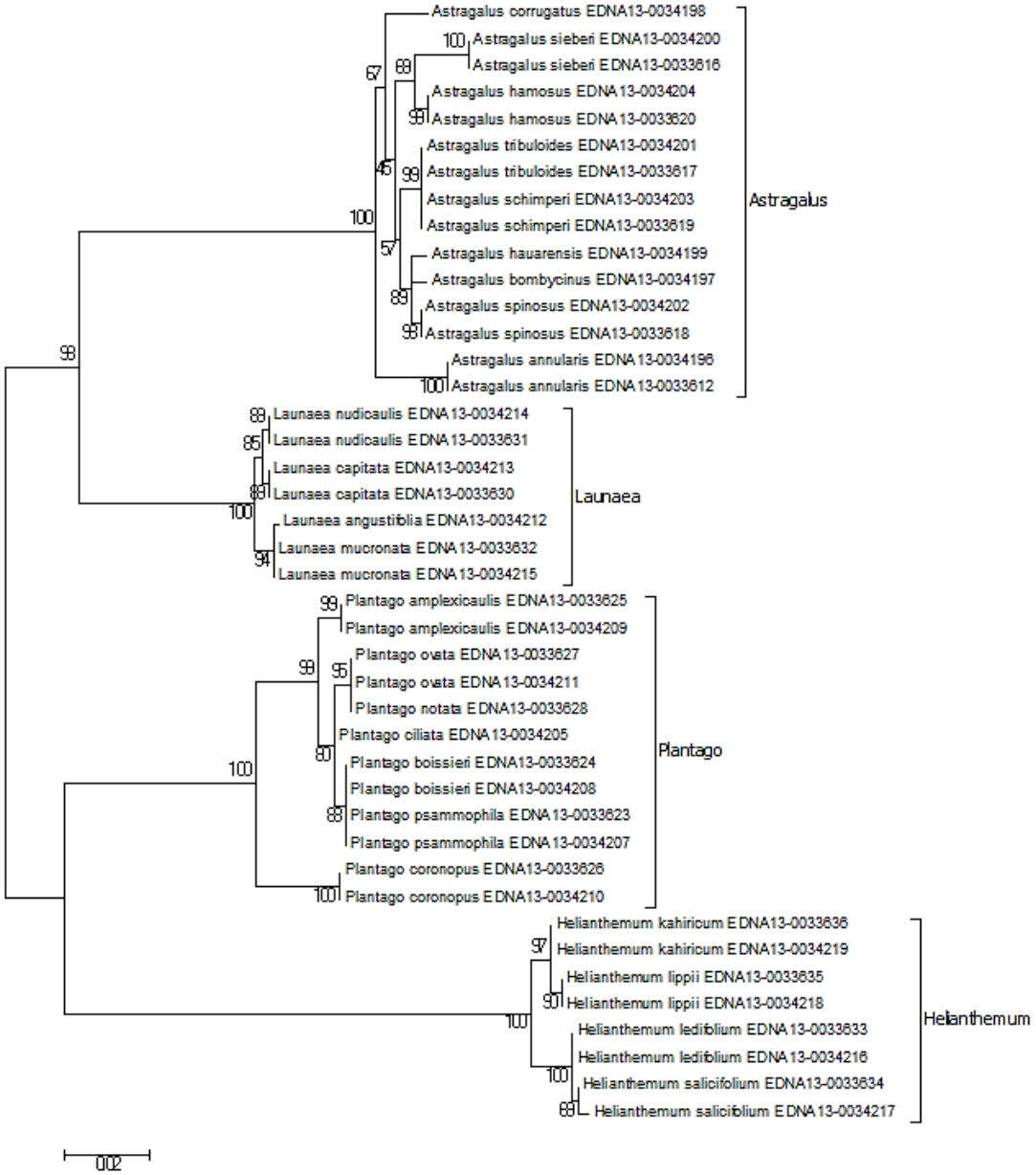
Maximum likelihood phylograms for combined *rbcL* + ITS2 barcodes (values represent % boot strap support with 1000 replicates)

(Continued)



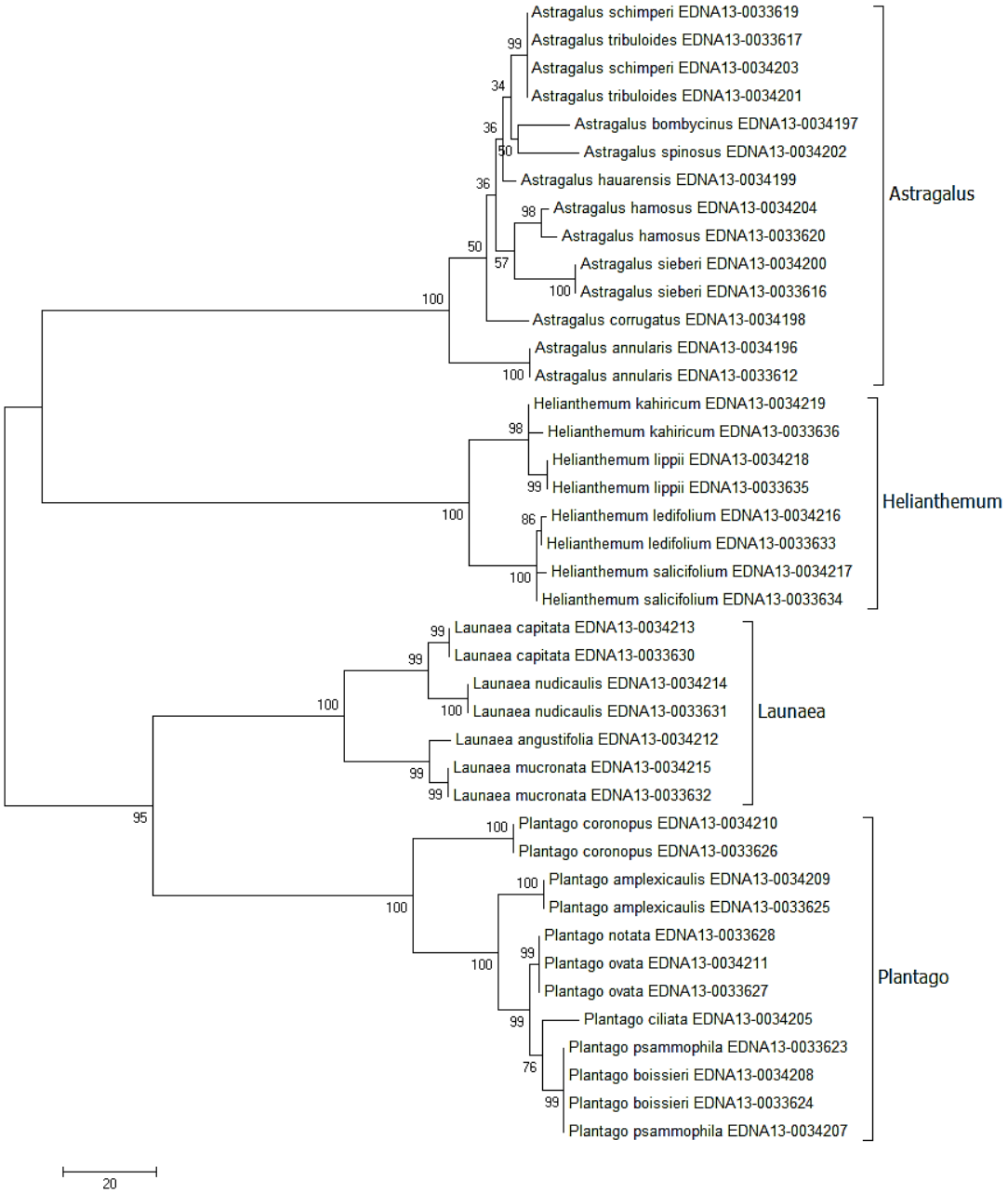
Maximum parsimony phylograms for combined *rbcL* + *trnL* barcodes (values represent % boot strap support with 1000 replicates)

(Continued)



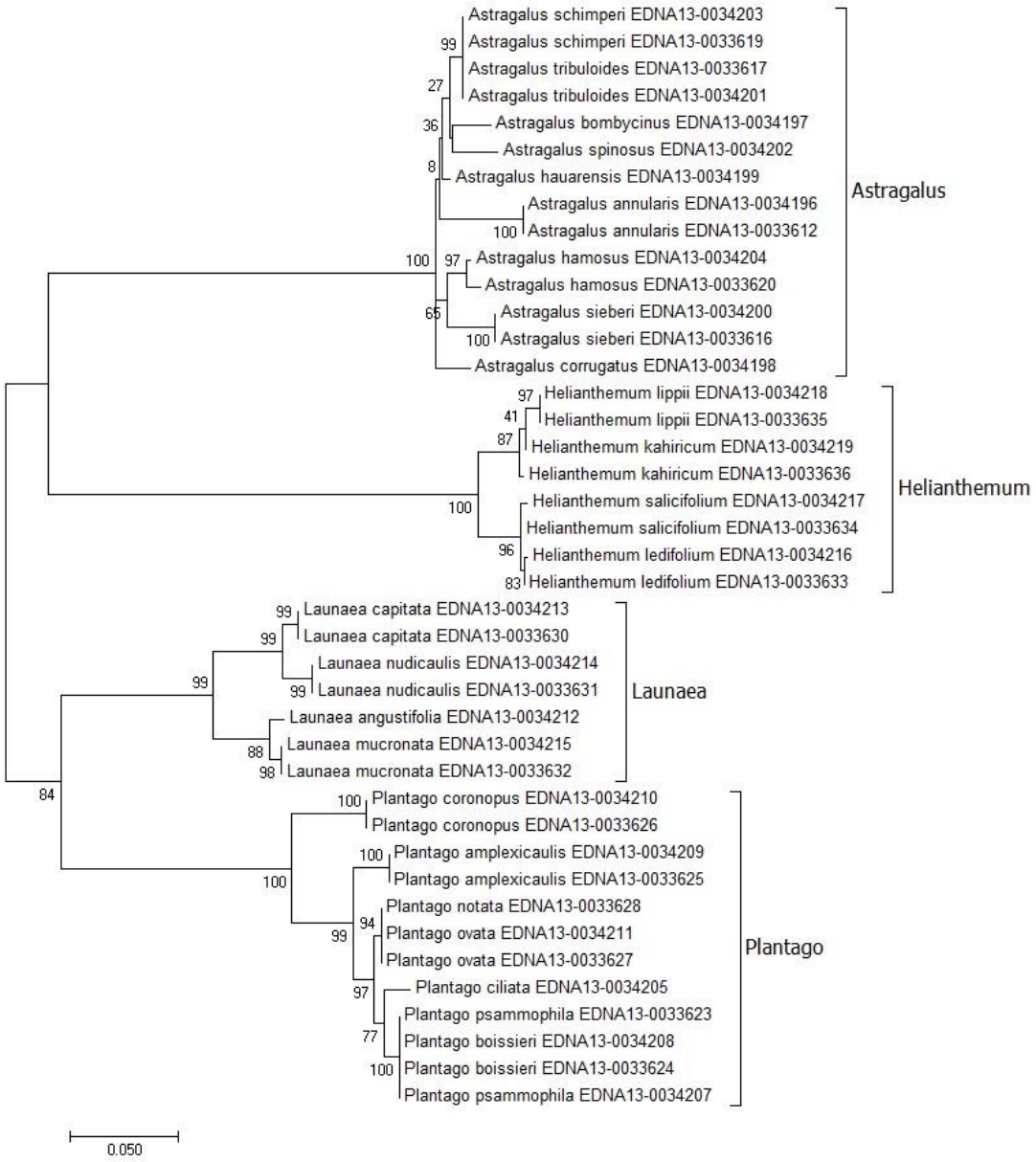
Maximum likelihood phylograms for combined *rbcL* + *trnL* barcodes (values represent % boot strap support with 1000 replicates)

(Continued)



Maximum parsimony phylograms for combined *trnL* + ITS2 barcodes (values represent % boot strap support with 1000 replicates)

(Continued)



Maximum likelihood phylograms for *trnL* + ITS2 barcodes
(values represent % boot strap support with 1000 replicates)

(End of Appendix 3.4)

Chapter 4 DNA barcoding the flora of Kuwait

4.1 Introduction

The ecological communities and plant biodiversity of Kuwait is seriously threatened and rapid changes have been reported in recent years due to climate change, land degradation, human interference, and overgrazing (Brown and Schoknecht, 2001; Brown, 2003; Al-Awadhi et al., 2005). The advancement of DNA barcoding and building a reference library for the Kuwaiti flora will produce floristic information about the current biodiversity, and aid in plant identification for taxonomic research, vegetation monitoring, and possible applications in large-scale ecological restoration projects (Sonstebo et al., 2010; Yoccoz et al., 2012).

DNA barcoding and building a reference library of the flora will contribute towards a quick identification method for identifying a large number of unknown plants from field surveys, replacing the lengthy time consuming traditional methods which require specialised skills. (i.e. using floristic key to identify one specimen with many possibilities to compare). DNA barcoding works by applying simple molecular techniques and protocols and is relatively straightforward to use.

Applications of DNA barcoding in Kuwait include identifying unknown plants from fragments of roots, seeds, leaves or even mixture of plants sampled up to the reconstruction of vegetation of a degraded site to study which plants have the potential to naturally regenerate or become reintroduced through restoration projects e.g. DNA based belowground root identification with a DNA barcode database to compared with current above ground vegetation (Jones et al., 2011; Kesananakurti et al., 2011; Wilson et al., 2014). Overgrazing in desert areas throughout Kuwait is common, plants trimmed showing only stem parts, in some cases only roots are buried in the soil, and it's difficult to identify using traditional taxonomic methods, while easily determined through simple steps by applying DNA molecular methods.

Wales (20,800 km^2), about the size of Kuwait, became the first country in the world to DNA barcode all of its 1,143 native flowering plants and conifers (de Vere, et al.,

2012). Their choice of DNA regions based on the recommended markers, *rbcL* and *matK* (CBOL, 2009). de Vere et al. (2012) managed to assemble 97.7 % sequence coverage for *rbcL*, 90.2% for *matK*, and a combined regions (*rbcL* + *matK*) for 89.7% of the Welsh flora (de Vere et al., 2012). The samples used mainly extracted from herbarium specimens (85 %). The freshly collected samples gained higher amplification success compared with the herbarium samples due to the degradation levels of DNA (de Vere, et al., 2012). They assessed the ability of DNA barcodes to identify species by the phylogenetic reconstruction of monophyletic groups using Neighbour-joining (NJ) trees which resulted in 69.4 to 74.9 % species resolution for combined regions *rbcL+matK* (de Vere et al., 2012). Also assessed the scope for improving species discrimination by looking at the resolution at different spatial scales, 10 x 10 km^2 (81.6 %) and 2 x 2 km^2 (93.3 %) (de Vere et al., 2012). Comparing the native flora of Kuwait (only 269 species) with Wales (1143 species), which covers less than one-quarter of the flora of Wales and almost the same land size, DNA barcoding approach is possible and manageable due to the small size of the flora of Kuwait. Other studies covered local geographical scale flora, using single or multiple gene regions are summarised in Table 4.1

Table 4.1 DNA barcoding approaches of floristic studies at local geographical scale.

Study	Geographical range/ sampling size	Markers used (% spp. resolution)
Lahaye et al, 2008	**Biodiversity hotspots:** (Mesoamerica and southern Africa). N= 1084 species	Single *matK* (90.6%) or Combined *matK* + *trnH-psbA* (90.9%)
Burgess et al., 2011	**Local temperate flora:** Canada (436 plant species, N= 513 samples)	Combined *rbcL* + *matK* (92.7%)
Theodoridis et al., 2012	**Country Greece:** native plants of Lamiaceae; N= 80 samples	Combined *matK* + *trnH-psbA* (81%)
Sosa et al., 2013	**Endangered plants** 20 spp. Orchids and 36 spp. Bamboo	*Orchids*= *matK* (100%); Bamboo= *matK* + *psbI-K* (71%)
Saarela et al., 2013	**Canadian Artic flora** (N= 490 vascular plant species)	Combined *rbcL* + *matK* (54%)
Liu et al., 2015	**Country China:** 531 local trees (N= 971 samples)	Combined *rbcL* + ITS2 71%)

The DNA barcoding database of the flora of Kuwait will include all plants mentioned in the updated checklist (Chapter 2, Table 2.3), excluding crops and cultivated plants. In compiling the specimens for the DNA database of the 402 species of Kuwait, I have managed to collect 614 voucher specimens belonging to 162 species from the field throughout Kuwait and 524 specimens were sampled belonging to 155 species from KTUH to complete the missing data set required for this project. Freshly collected leaf samples were preferable over herbarium material due to DNA degradation and amplification issues with herbarium material (Staats et al., 2011; Sarkinen et al., 2012). A total number of 30 native species missing from the collection (specimens not available).

At least three accessions for each species were chosen (some species represented by only one accession based on the availability of sampling material) from a widespread geographical location within Kuwait, following de Vere et al. (2012) sampling strategy. Other DNA barcoding studies dealing with local floras also sampled across

separate geographical location between 1 to 5 accessions per species (Lahaye et al., 2008; Sosa et al., 2013; Liu et al., 2015).

This study will present for the first time, locally in Kuwait and regionally in the Arabian Peninsula, a new DNA barcode approach representing about ¾ of the flora of Kuwait. Here, I will establish a barcoding database of the flora based on the evaluation of five barcodes in Chapter 3, and will apply *rbcL* and ITS2 barcodes and evaluate their discriminatory power across species, genera, and families of the flora.

4.2 Methods

4.2.1 Sampling material

In total 721 individuals were sampled; 388 from herbarium specimens (collected from KTUH) and 333 from living plants collected throughout Kuwait. Information for each DNA extracted individual including species names, Edinburgh DNA numbers, Collection ID, Collectors names and numbers, year of collection, collection type (herbarium or fresh collection) and locality/ region of the collection along with BOLD and GenBank accession numbers representing *rbcL* and ITS2 sequences are presented in Appendix 4.1. All individuals sampled represent 264 species belonging to 44 families and 22 orders of the flora of Kuwait (mean = 2.7 individual sequenced per species). Most plant material sampled from KTUH herbarium specimens collected between 1991 and 2012 (range of 21 years) are considered to be of young age. Fresh material from plant leaves collected and silica-gel dried from 20 different locations throughout Kuwait during late spring 2012 and early spring of 2013. Plant collection sites included fence protected areas, such as Sabah Al-Ahmad Nature Reserve (SSNR), Kabd protected area, and the Al-liyah natural reserve, which are considered natural examples of the flora with minimal impacts from grazing and anthropogenic activities. Other visited field sites include Al-Abdaly and Um Neqa Demilitarized zone (DMZ), along the North border line with Iraq; and several unfenced desert areas with minimal disturbance such as Al-Salmi and Nuwaiseeb. Also, Coastal areas such as Sulaibikhat and Khairan were visited. A collection of 60 accessions, representing

52 species, was also made from Failaka Island 20 km off the coast of Kuwait City. All necessary permits were obtained before collection of specimens.

Fresh plant vouchers were determined using several floristic publications on the flora of Kuwait (Daoud, 1985; Al-Rawi, 1987; Boulos, 1988) and reconfirmed by consulting the herbarium curator at KTUH, Dr. K.T. Mathew. Herbarium vouchers made from freshly collected plants deposited in KTUH and Royal Botanic Garden Edinburgh (RBGE) Herbarium (E), mounted and digitized. The herbarium specimens deposited in E are accessible online by searching RBGE herbarium catalogue under my name M Abdullah following:
http://elmer.rbge.org.uk/bgbase/vherb/bgbasevherb.php

4.2.2 DNA extraction, amplification, and sequencing

DNA extraction

A small amount (~ 20 mg) of dry, healthy leaf material was selected from herbarium and freshly collected specimens and loaded in 2.0 ml Eppendorf tubes carried on 96-well plate, with one 5 mm stainless steel bead. Samples were ground using TissueLyser II (Qiagen, Ltd.) until the material was powdered (frequency 20 Hz x 2 x ~ 30 sec). DNA was extracted using the fine powdery plant material through QIAxtractor automated instrument (Qiagen Ltd.) following QIAxtractor user defined protocol which yields high-quality DNA from plant tissues suitable for a wide variety of downstream applications. The procedure is divided into two parts: lysis/ digestion and extraction.

The lysis/ digestion part was conducted manually by the preparation of the digestion buffer using a mixture of the following reagents: 400 µl RNase, 400 µl DX enzyme and 40.4 ml of DXT reagent; loading 420 µl for each sample using lysis tubes and mixed using multi-channel pipette with care to avoid cross contamination. The samples placed into the Thermomixer for 1 hour at 65°C and 800 rpm (once complete samples can be left for few hours up to overnight at room temperature).

The automatic extraction part requires a quick setup of the QIAxtractor robot by launching the QIAxtractor software (RBGE user defined, run file: QXT Liquid DNA V1.QSP), after loading the QIAxtractor platform with the appropriate reagents (loading the reservoirs with DX Binding, DX Wash, DX Final, and DX Elution buffer) , pipetting tips and capturing plates. Fresh DXB reagent is prepared by adding 1.4 g of binding additives (DX) to the whole bottle. The lysed/ digested samples were centrifuged at 2000 rpm for 10 mins; for each lysed sample 220 µl (110 µl at a time) of the supernatant is transferred to the QIAxtractor 96-well plate (with square wells) and placed into position inside the robot extractor. By closing the extractor lid, the following automated cycles will take place: 440 µl of DX Binding with DX Binding Additive is added to the lysis plate. The lysate is then mixed and incubated at room temperature for 5 min. 600 µl of the lysate is added into the capture plate with a vacuum of 35 kPa is applied for 5 min. 200 µl of DX Binding with DX Binding additive is loaded into the capture plate with a vacuum of 35 kPa is applied for 5 min. 600 µl of DX Wash is loaded into the capture plate with a vacuum of 25 kPa applied for 1 min (repeated twice). A vacuum of 25 kPs is applied for 5 min to dry the plate. The carriage is moved to the elution chamber, and 100 µl of DX Elution buffer is loaded into the capture plate. The samples are incubated for 5 min followed by 1 min vacuum at 35 kPa ending the extraction cycle. The tubes were removed and frozen at -20°C for long-time preservation.

PCR amplification and sequencing

Two regions *rbcL* and ITS2 were used to build a DNA reference library of the flora of Kuwait. Due to smeary bands with previously used reverse primer *rbcL*-ajf634R (in chapter 3). The following primers were evaluated for the *rbcL* region:

- *rbcL* forward primer: *rbcL*-aF, primer sequence 5`-ATGTCACCACAAACAGAGACTAAAGC-3` (Kress and Erickson 2007).
- *rbcL* reverse primers: *rbcL*-aRev, primer sequence 5`-GTAAAATCAAGTCCACCRCG-3` with a mean size of 554 bp (Kress et al., 2009), and *rbcL*-724R, primer sequence 5`-TCGCATGTACCTGCAGTAGC-3` with size of 702-883 bp (Kress et al., 2005).

For *rbcL* region the chosen primers, *rbcL*-aF and *rbcL*-aRev with a mean size of 554 bp (Kress et al., 2009) were used to DNA barcode the flora of Kuwait.

The same primers described in Chapter 3 were used for the ITS2 region:

- ITS2 forward primer: ITS-S2F, primer sequence 5`-ATGCGATACTTGGTGTGAAT-3` with a mean size of 226 bp (Chen et al., 2010).
- ITS2 reverse primer: ITS-S3R, primer sequence 5`-GACGCTTCTCCAGACTACAAT-3` with size of 163-311 bp (Chen et al., 2010).

PCR amplification and sequencing methods used here follow Chapter 3 Section 3.2.2. PCR conditions for each region are presented in Table 3.4. Also, sequence editing, alignment, and molecular analysis used follow previous methods in Chapter 3 Section 3.2.3.

4.2.3. Sequence editing, alignment, and molecular analysis

The methods used here follow section 3.2.3 in Chapter 3. Trimming and assemblage of sequences into contigs was performed using Geneious software (ver. 6.1.8, Kearse et al., 2012). Each contig was checked for base call disagreements and ambiguities and manually edited where necessary.

Multiple sequence alignments (MSA) were performed using MUSCLE alignment (Edgar, 2004), followed by manual sequence editing where necessary using Molecular Evolutionary Genetics Analysis software version 7.0 (MEGA7: Kumar et al., 2015). The coding region, *rbcL* sequences were aligned using the default settings of MUSCLE provided by MEGA7 software. For the non-coding region, ITS2, was aligned and manually edited by comparing species with multiple accessions together and/ or by genus for those genera containing more than two species and the settings for gap penalties were adjusted to generate MSA with fewer gaps in the final alignment; the default settings: gap open -400 with gap extension at 0, adjusted to gap open at -800 and gap extension at -1. The method of using MUSCLE alignment with

ITS2 region was applied previously by several studies (Pang et al., 2013; Mishra et al., 2016; Yu et al., 2016). The use of Neighbour-joining tree based distance method to evaluate the ability of ITS2 sequences to delimit the species into discrete clades or monophyletic groups is well documented in several studies (Hribova et al., 2011; Zhang et al., 2012; Raveendar et al., 2015; Mishra et al., 2016).

The MEGA7 software was also used to describe the genetic variability of each marker by calculating the maximum and minimum length of base pairs, total aligned matrix, variable sites (%), parsimony informative sites (%), and singleton sites (%).

A list of specimens with Barcode of Life Data Systems (BOLD) and GenBank accessions representing successfully barcoded sequences for *rbcL* and ITS2 are listed in Appendix 4.1.

4.3 Results

4.3.1 DNA recoverability, amplification and sequence success

For both tested regions (*rbcL* and ITS2), DNA extracted from silica-dried fresh leaves showed the highest percentage of DNA sequence recoverability (84-96 %) compared to herbarium specimens (63-71 %). The number of low DNA barcode sequence recoverability using herbarium material was mainly due to lower PCR amplification success (Table 4.2). Although PCR was performed twice on those individuals with amplification problems for each region, it was not possible to obtain sequences from all individuals sampled.

For the core barcode regions 1,117 sequences were obtained, 594 and 523 for *rbcL* and ITS2, respectively. The percentage of sequences which were of high quality varied for each region amplified. Sequencing efficiency was highest for *rbcL* (82 %) and lowest for ITS2 (72.5 %). Table 4.2 summarizes PCR amplification and sequencing rates for each region. The most efficient percentage of DNA extraction was from fresh plant material for the *rbcL* region (96 %) and the lowest was herbarium material for the ITS2 region (63 %). The highest amplification and sequence failure was shown by ITS2, 18.6 % and 8.8 %, respectively (Table 4.2).

Table 4.2 Summary of *rbcL* and ITS2 successfully amplified and sequenced using fresh and herbarium plant material

Region/ Collection type	No of individuals	Sequence efficiency (%)	Amplification failure (%)	Sequence failure (%)
***rbcL*- Total**	**721**	**594 (82)**	**99 (14)**	**28 (3.8)**
rbcL Herbarium	388	273 (71)	87 (22)	28 (7.2)
rbcL Fresh	333	321 (96)	12 (3.6)	0
ITS2 - Total	**721**	**523 (72.5)**	**134 (18.6)**	**64 (8.8)**
ITS2 Herbarium	388	246 (63)	97 (25)	45 (11.5)
ITS2 Fresh	333	280 (84)	37 (11)	16 (4.8)

4.3.2 Sequence quality and alignment

The multiple sequence alignment (MSA) lengths of *rbcL*, ITS2, and the combination of *rbcL* + ITS2 barcodes are 496, 467, and 963 bp, respectively (Table 4.3). For *rbcL*, the final matrix contained 594 sequences of length 496 bp; ITS2 final matrix contained 523 sequences with minimum sequence length range from 354 bp up to a maximum length of 417 bp. The final matrix of ITS2 MSA including gaps showed a length of 467 bp (Table 4.3). The MSA length of the final matrix for the combined regions *rbcL* + ITS2 was 963 bp with a minimum sequence length range from 850 bp up to a maximum length of 913 bp.

The mean percentage of bases within the sequences (QV > 30) was highest for *rbcL* (96.9 %), slightly lower for ITS2 region (91.8 %), and 94.3 % for the combined regions *rbcL* + ITS2. The ITS2 matrix contains the most variable sites 77 % and parsimony informative sites 73 % compared to *rbcL* with 44 % variable sites and 43 % parsimony informative sites (Table 4.3). The combined matrix of *rbcL*+ITS2 contains 59 % of variable sites and 57 % of parsimony informative sites. Analyses were performed with the final matrix composed of 594 sequences of *rbcL*, 523 of ITS2, and 480 sequences of combined *rbcL* + ITS2 (Table 4.3).

Table 4.3 Sequence quality and size of DNA barcodes for *rbcL*, ITS2 and combination of *rbcL* + ITS2

	rbcL	**ITS2**	***rbcL* + ITS2**
Multiple sequence alignment length (bp)	496	467	963
Minimum sequence length (bp)	496	354	850
Maximum sequence length (bp)	496	417	913
Number of Gaps (%)	0	79 (17)	71 (7)
Variable sites bp (%)	220 (44)	358 (77)	565 (59)
Parsim-informative sites (%)	215 (43)	343 (73)	546 (57)
Singleton sites (%)	5 (1)	13 (2.8)	18 (1.9)
GC content (%)	41.50	32.50	35.80
Mean high quality bases QV>30 (%)	96.9	91.8	94.3
Mean low quality bases QV<20 (%)	1.3	4.1	2.7
Efficiency of PCR amplification (%)	622/721 (86)	587/721 (81)	
Success rate of sequencing (%)	594/622 (95)	523/587 (89)	480*
Total Missing Sequences (%)	127/721 (18)	198/721 (27)	241/721 (33)

* Number of combined *rbcL* + ITS2 sequences

4.3.3 BLASTn searches

A nucleotide BLASTn searches were carried out in GenBank against all organisms in the NCBI database to investigate sequences similarity at the species, genus and family level using *rbcL* and ITS2 barcoded sequences of the flora of Kuwait.

Sequences of *rbcL* and ITS2 were tested by top-scoring hits using BLASTn searches through NCBI database (accessed on 14/05/2017 via Geneious software ver. 10.1.3, BLAST plugin). Figure 4.1 shows the percentage of barcoded sequences matching the NCBI database using BLASTn searches tool and the comparison between *rbcL* and ITS2 barcodes are shown to species, genus and family level (Figure 4.1). The percentage of barcodes matching NCBI database to species level was highest for ITS2 (47 %) and lowest for *rbcL* (31 %). Highest percentage matches to the genus level are shown by *rbcL* barcodes 42 % followed by ITS2 37 %; at the family level *rbcL* matched 27 % and ITS2 16 % (Figure 4.1). Sequences of ITS2 and *rbcL* regions that did not match to at least family level of the NCBI database are considered as a failure.

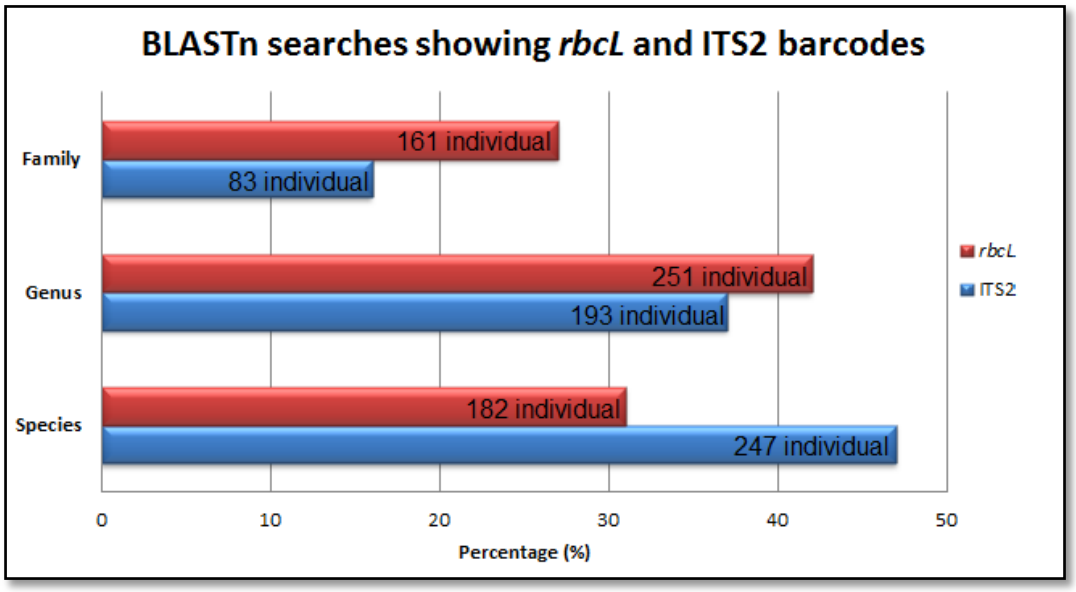


Figure 4.1. Percentage of barcoded sequences matching family, genus, and species level to similar sequences from NBCI database using BLASTn searches. Total barcoded sequences for *rbcL* = 594 individual; ITS2 = 523 individual.

4.3.4 Species discrimination

The determination of species resolution (percentage) was based on the analyses using neighbour joining (NJ) tree method. A species considered resolved if all members had barcodes sequences that more related to each other than to members of other species.

Monophyly tree based tests

Determination of species delineation was based on the phylogenetic reconstruction of NJ trees for *rbcL*, ITS2 and combined regions of *rbcL* + ITS2. NJ tree based analysis was used to determine the species resolution (%) whether the multiple alignments of a single region (*rbcL*, ITS2) and combined *rbcL* + ITS2 barcoded sequences resolve as species-specific clusters. The NJ trees reconstructed for ITS2, *rbcL*, and *rbcL* + ITS2 are represented as phylograms to show branch order and branch length with clades labeled indicating major plant groupings.

Single region tree based analysis

The resolution generated by the NJ tree demonstrated the greatest support values for ITS2 region (Table 4.4 and Figure 4.2), which was capable of recovering 69 % of species-specific clusters (with bootstrap support, 100 replicates). For *rbcL* region species resolution was lower, 58.2 % of species-specific clusters (Table 4.4 and Figure 4.3). The analysis included all barcoded sequences, *rbcL* = 594 individual and ITS2 = 523 individual.

Combined regions tree based analysis

Combining the multiple aligned sequences of *rbcL* + ITS2 (Figure 4.4) NJ tree was capable of recovering 70.5 % of species-specific clusters (with bootstrap support, 100 replicates) (Table 4.4 and Figure 4.4). The variation and resolution of species-specific clusters for the combined *rbcL* + ITS2 regions increased by 1.5 % (70.5 %) when compared with ITS2 region alone (69 %) and increased greatly by 12.3 % when compared with *rbcL* region (58.2 %) (Table 4.4).

Table 4.4 Percentage of species-specific clusters using Neighbour Joining tree ≥ 50 % bootstrap support for single DNA regions and combined.

DNA regions	*N* species/ total individuals	Species-specific clusters
		N species/(NJ %)
rbcL	244 / 594	142 (58.2 %)
ITS2	228 / 523	158 (69 %)
***rbcL* + ITS2**	217 / 480	153 (70.5 %)

All accessions included

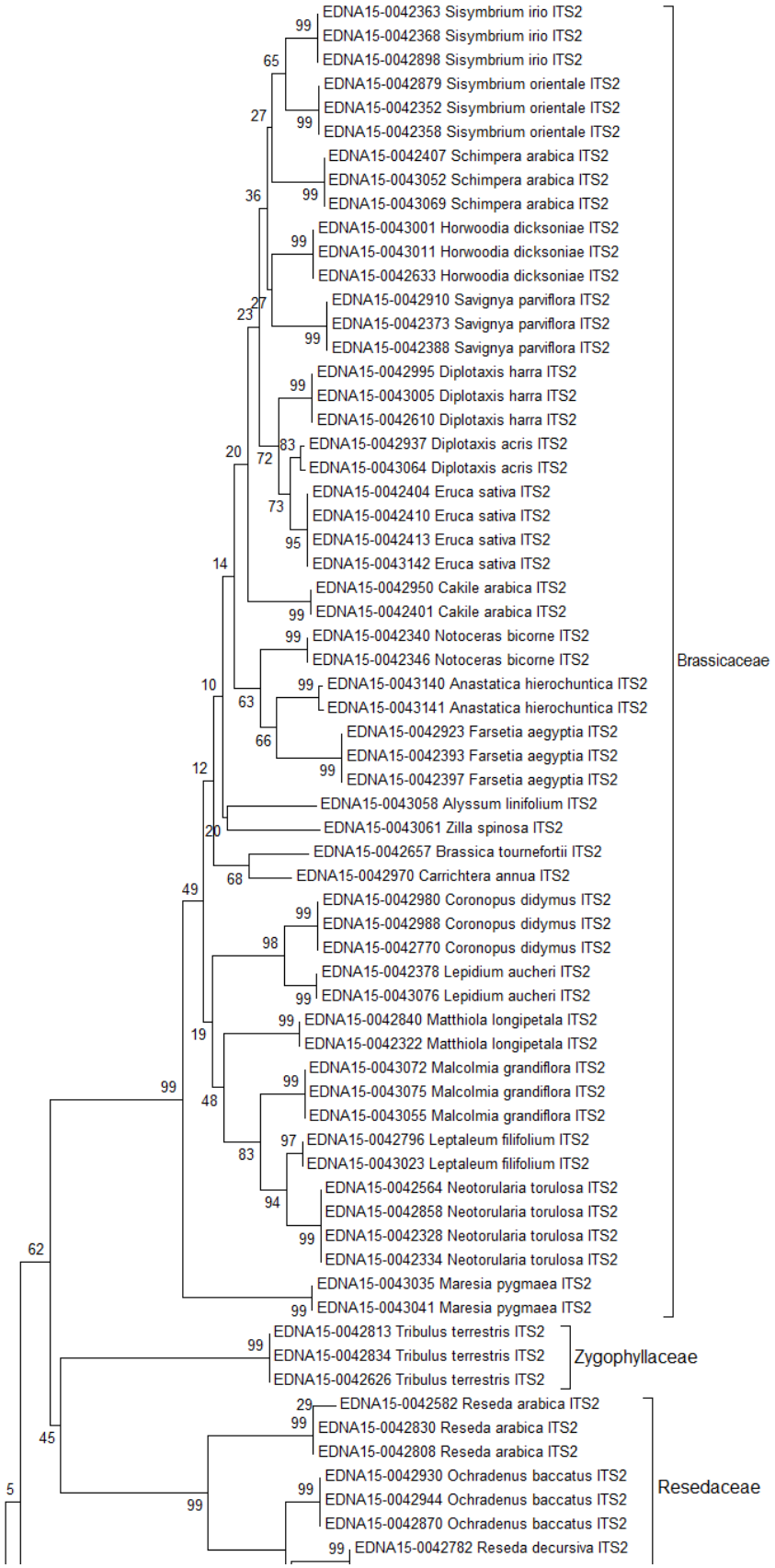


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 1/8]

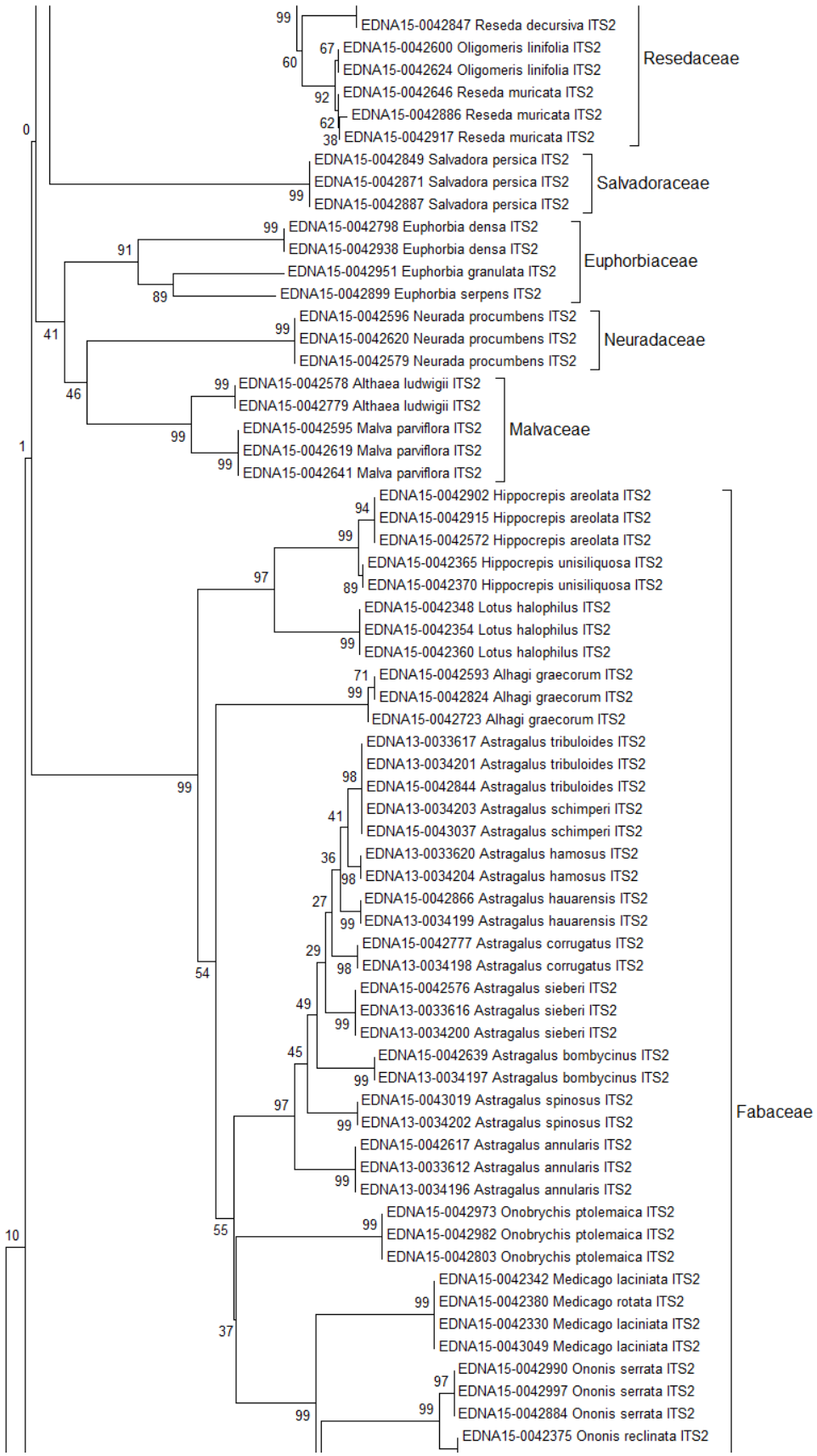


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 2/8]

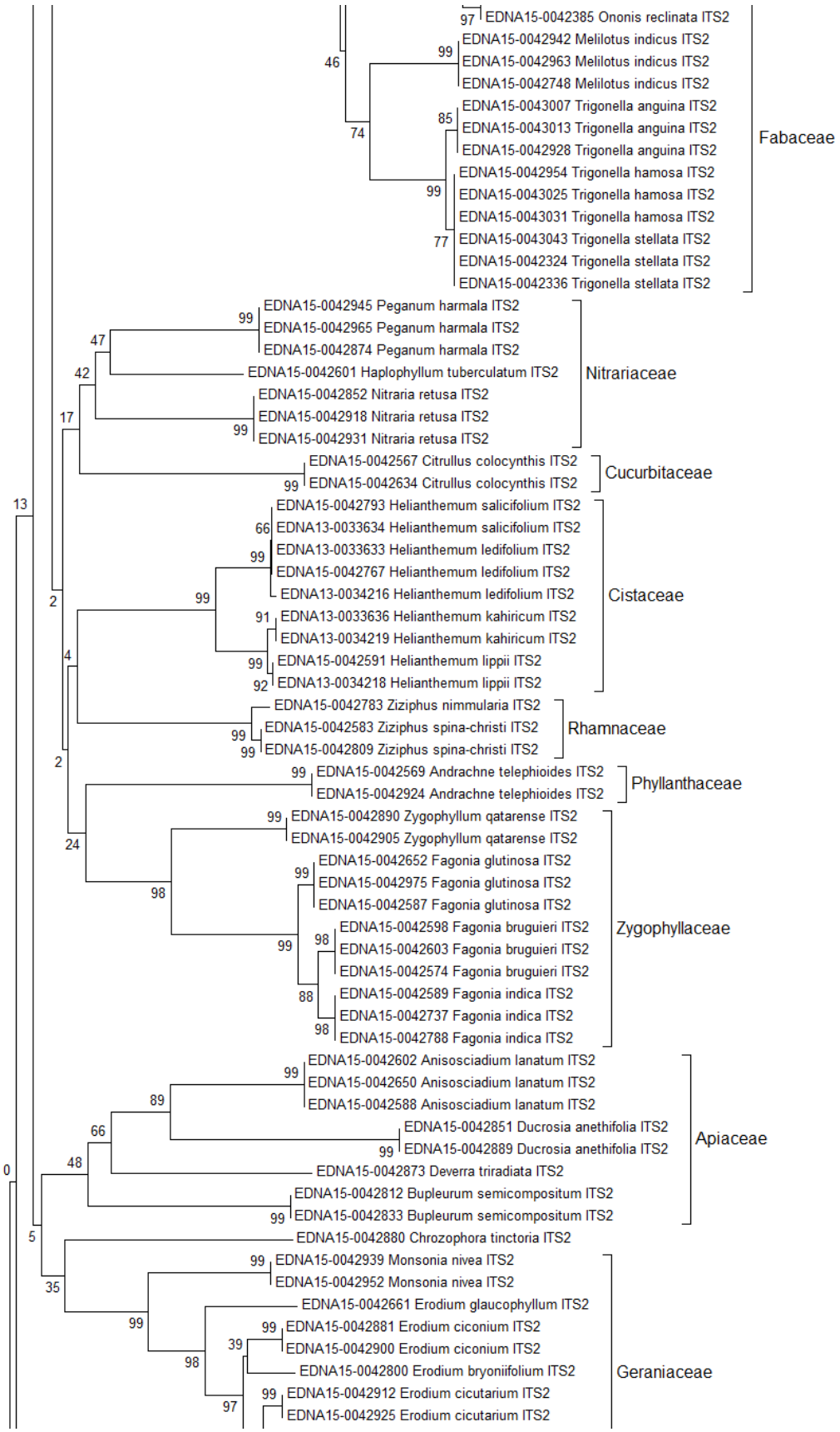


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 3/8]

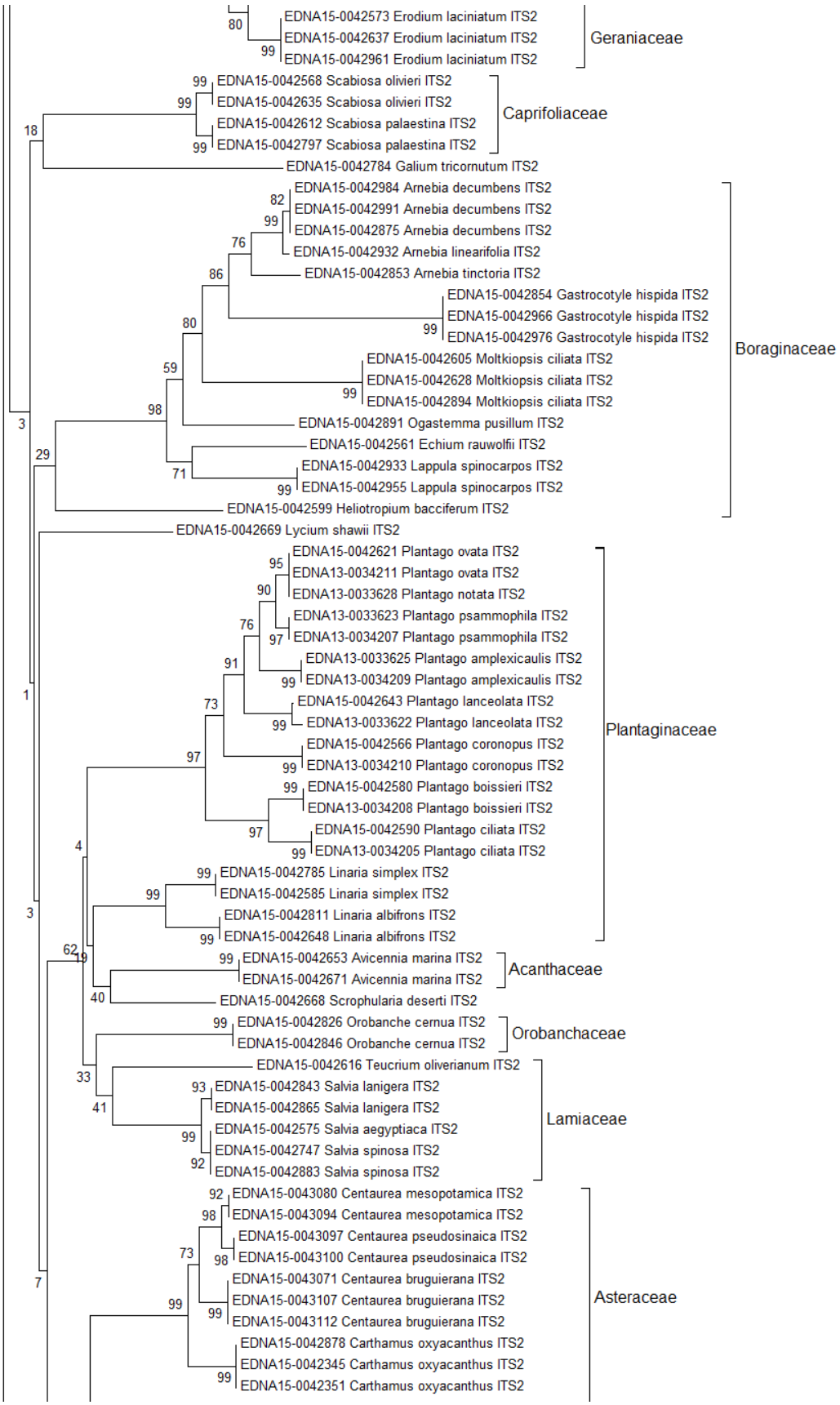


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 4/8]

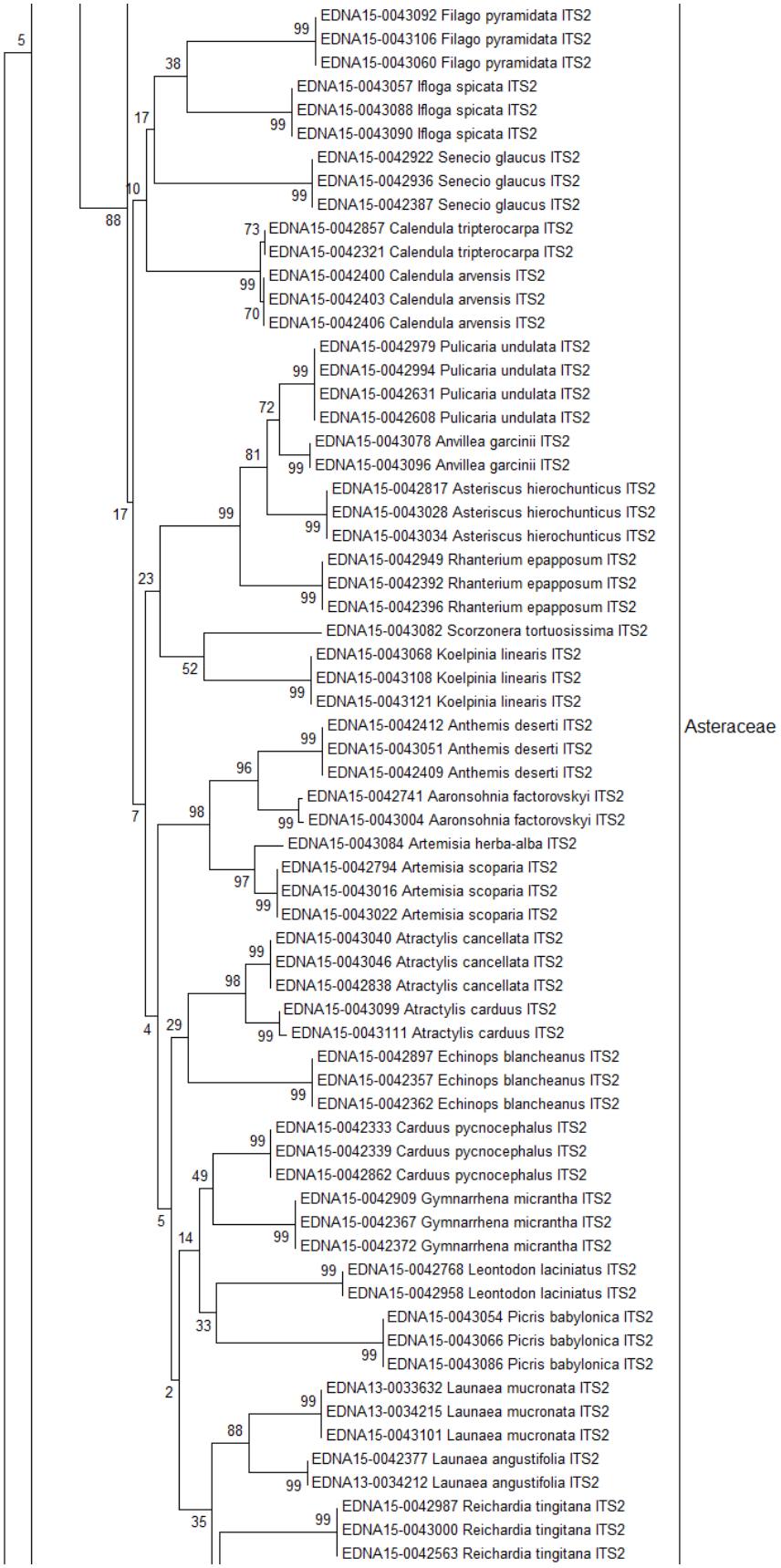


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 5/8]

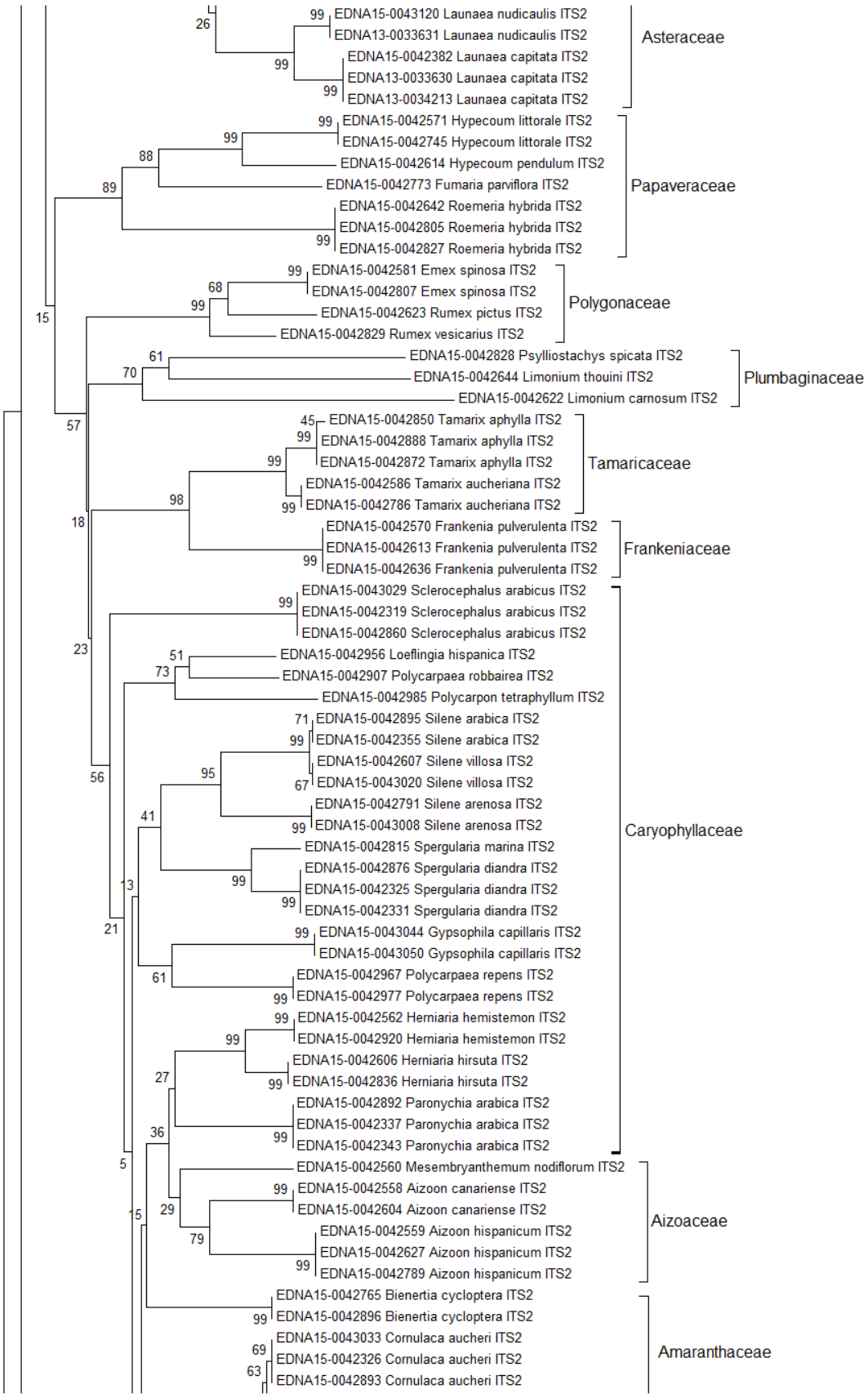


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 6/8]

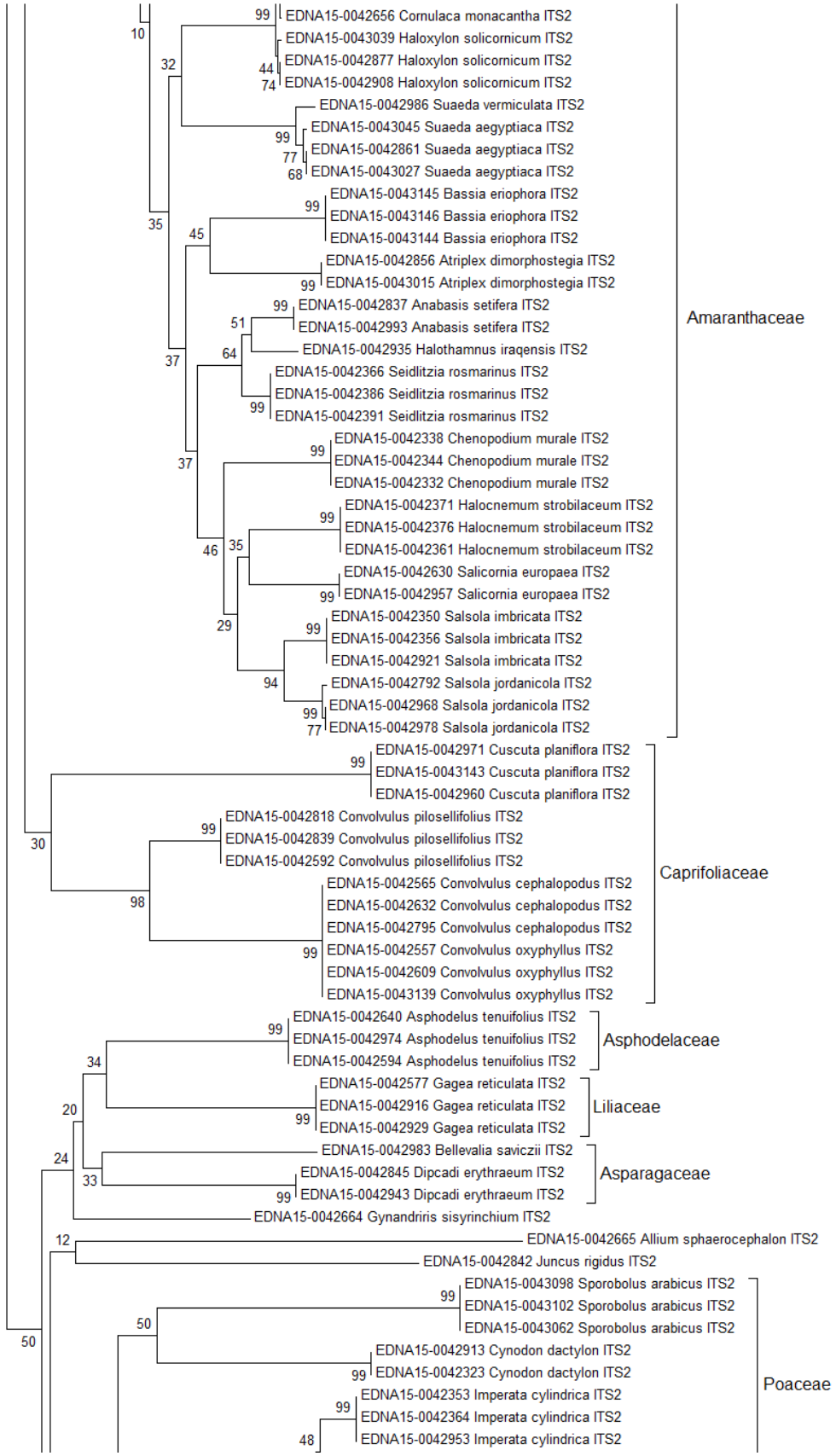


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [Cont. 7/8]

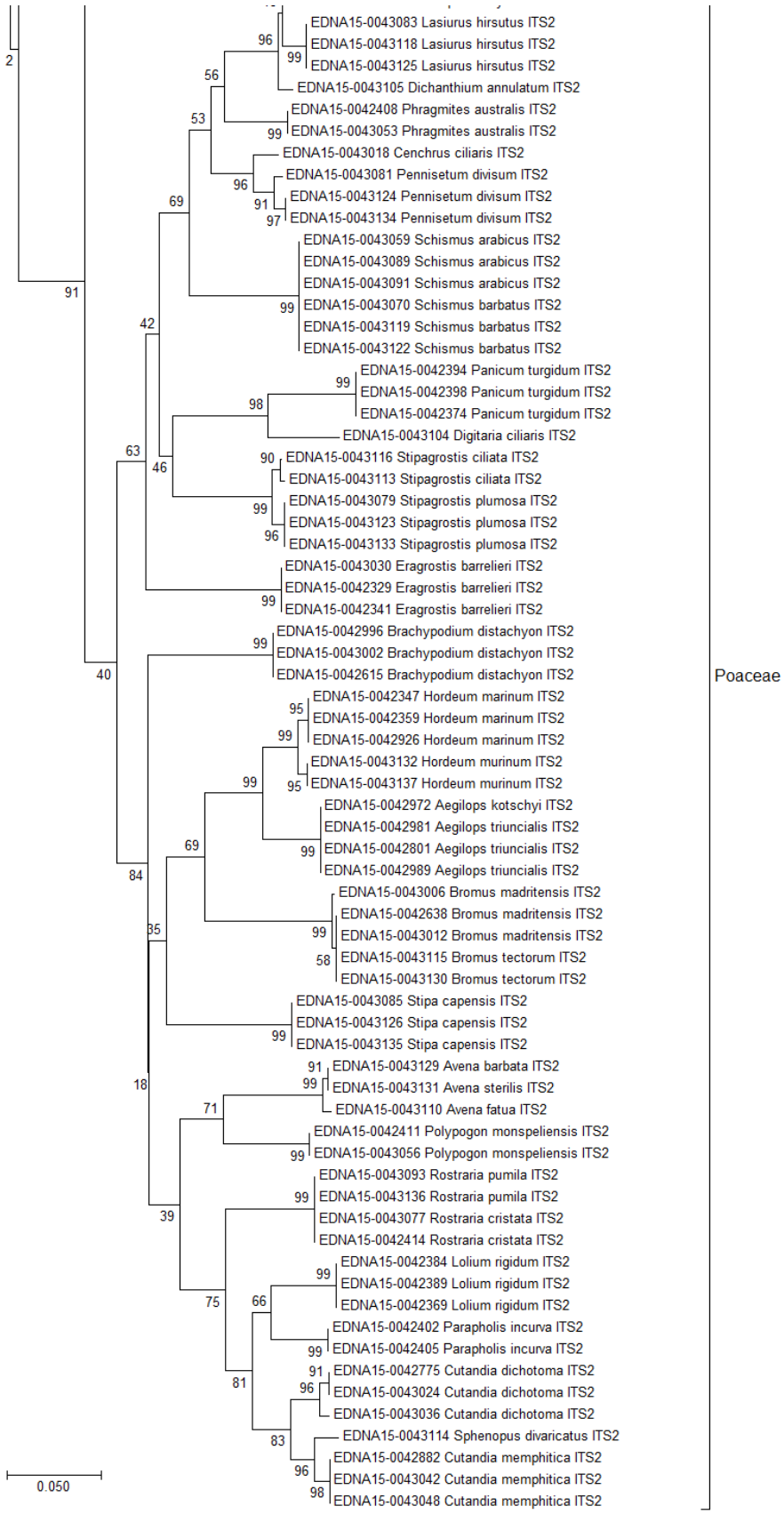


Figure 4.2 Neighbour joining phylograms for ITS2 barcodes representing 523 sequences
(values represent % boot strap support with 100 replicates) [End. 8/8]

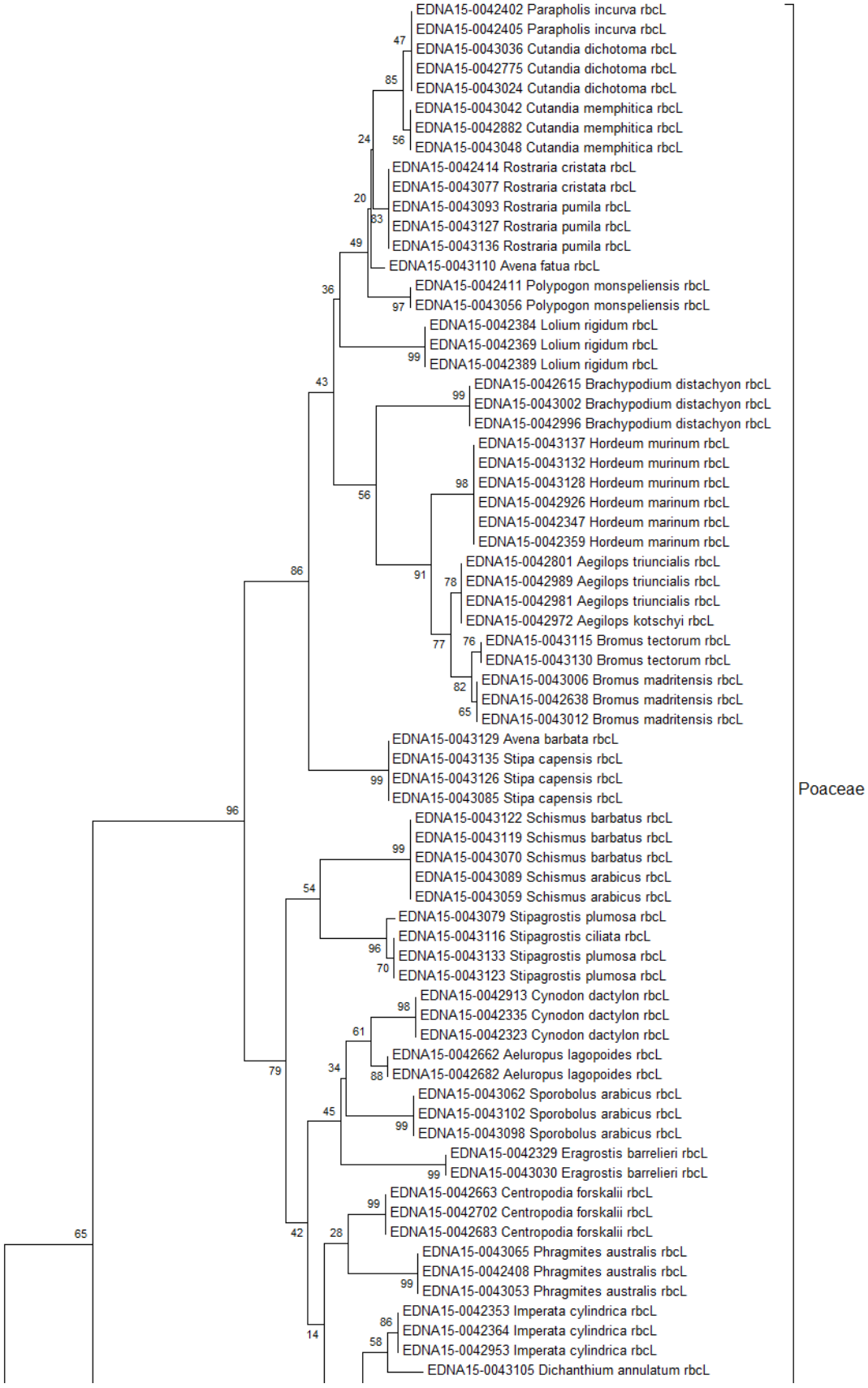


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 1/9]

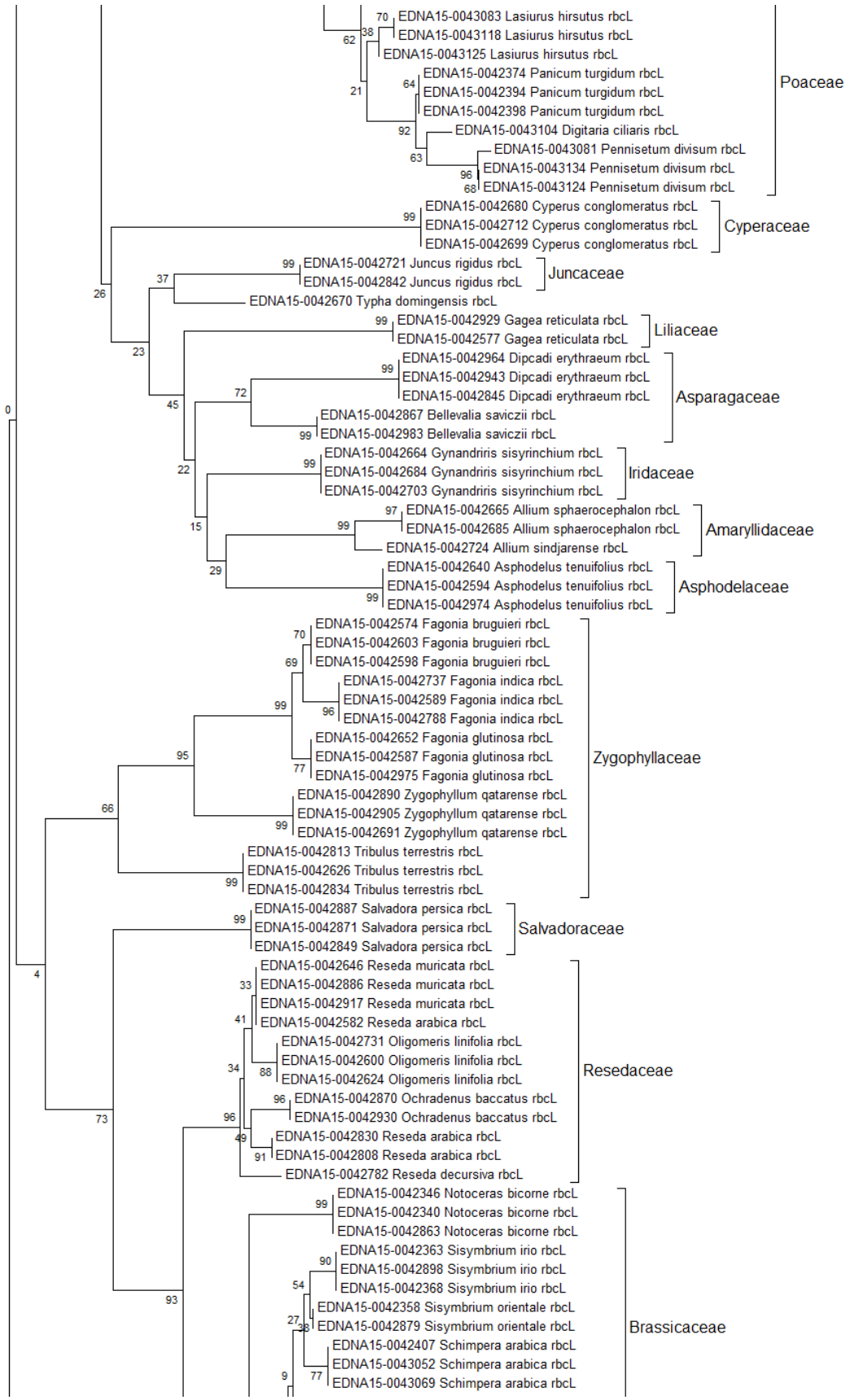


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 2/9]

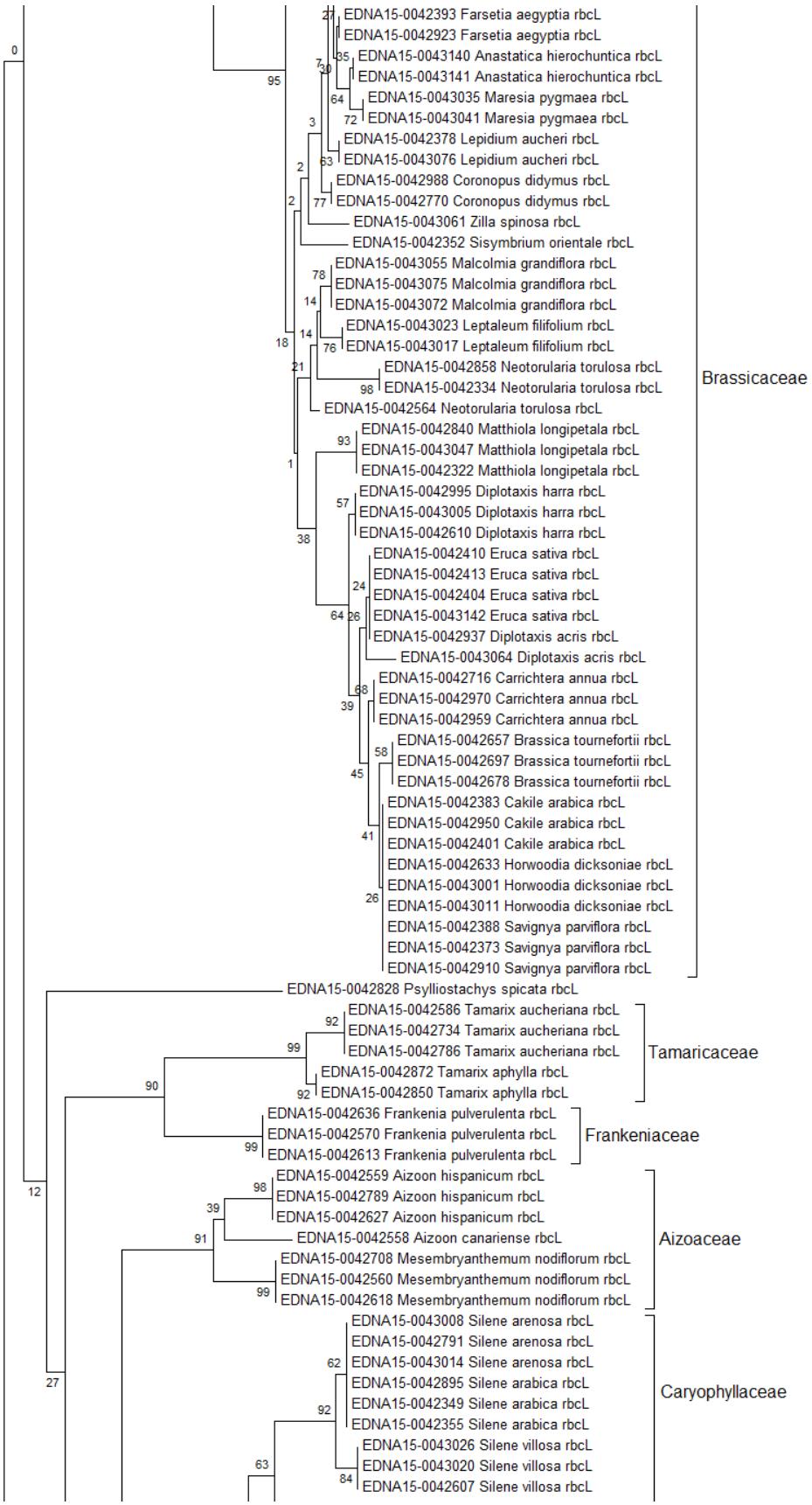


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 3/9]

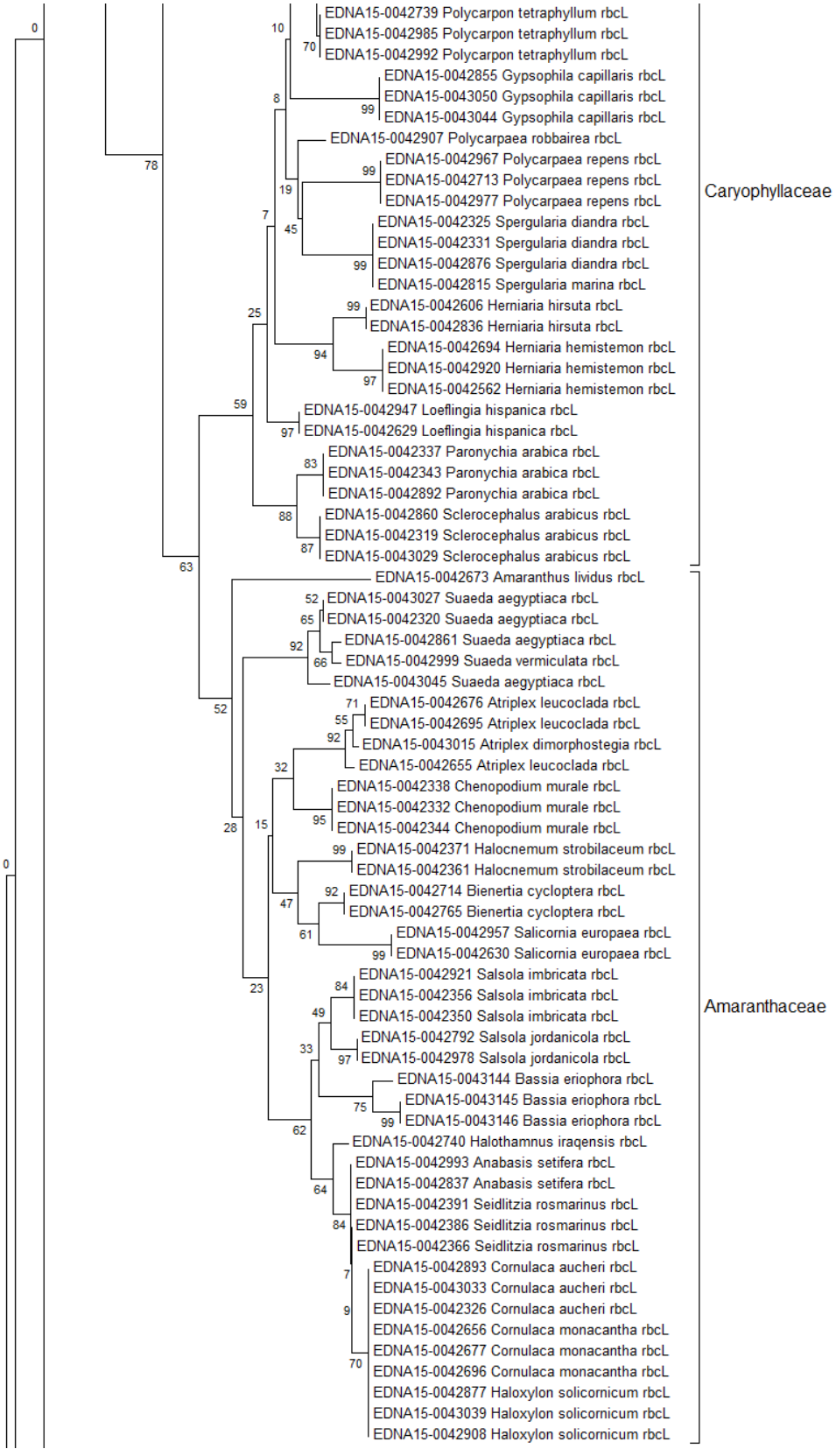


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 4/9]

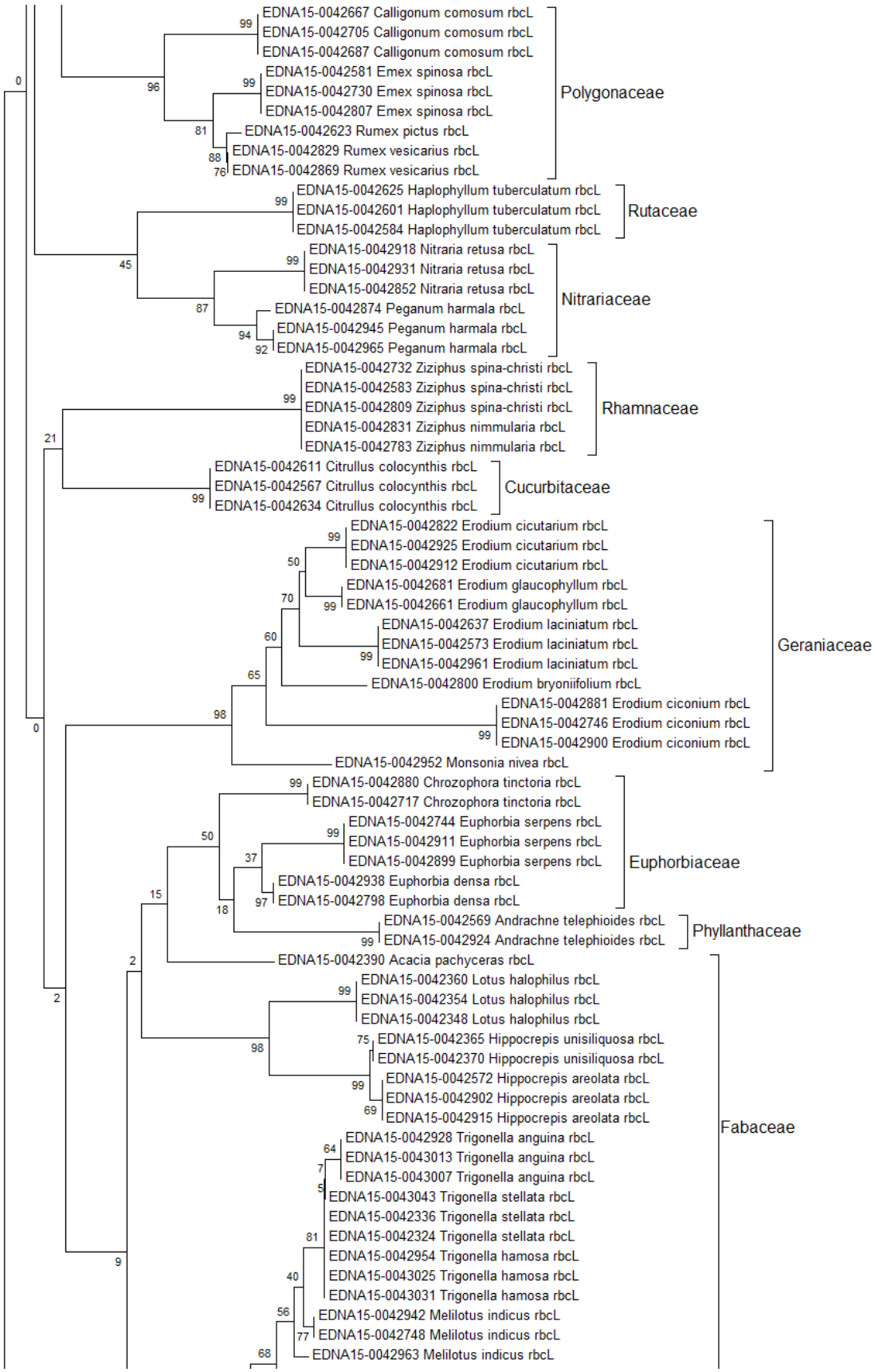


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 5/9]

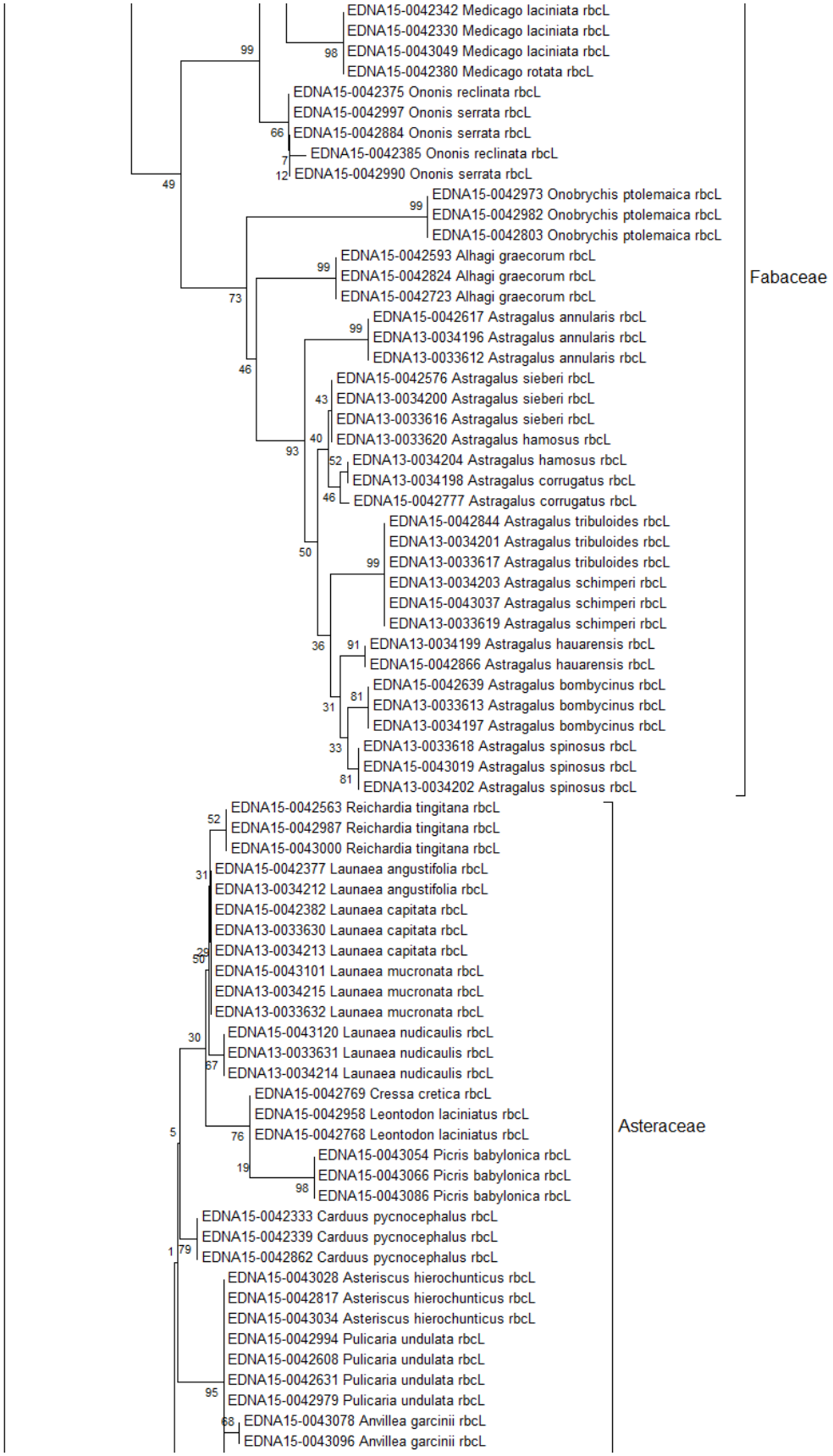


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 6/9]

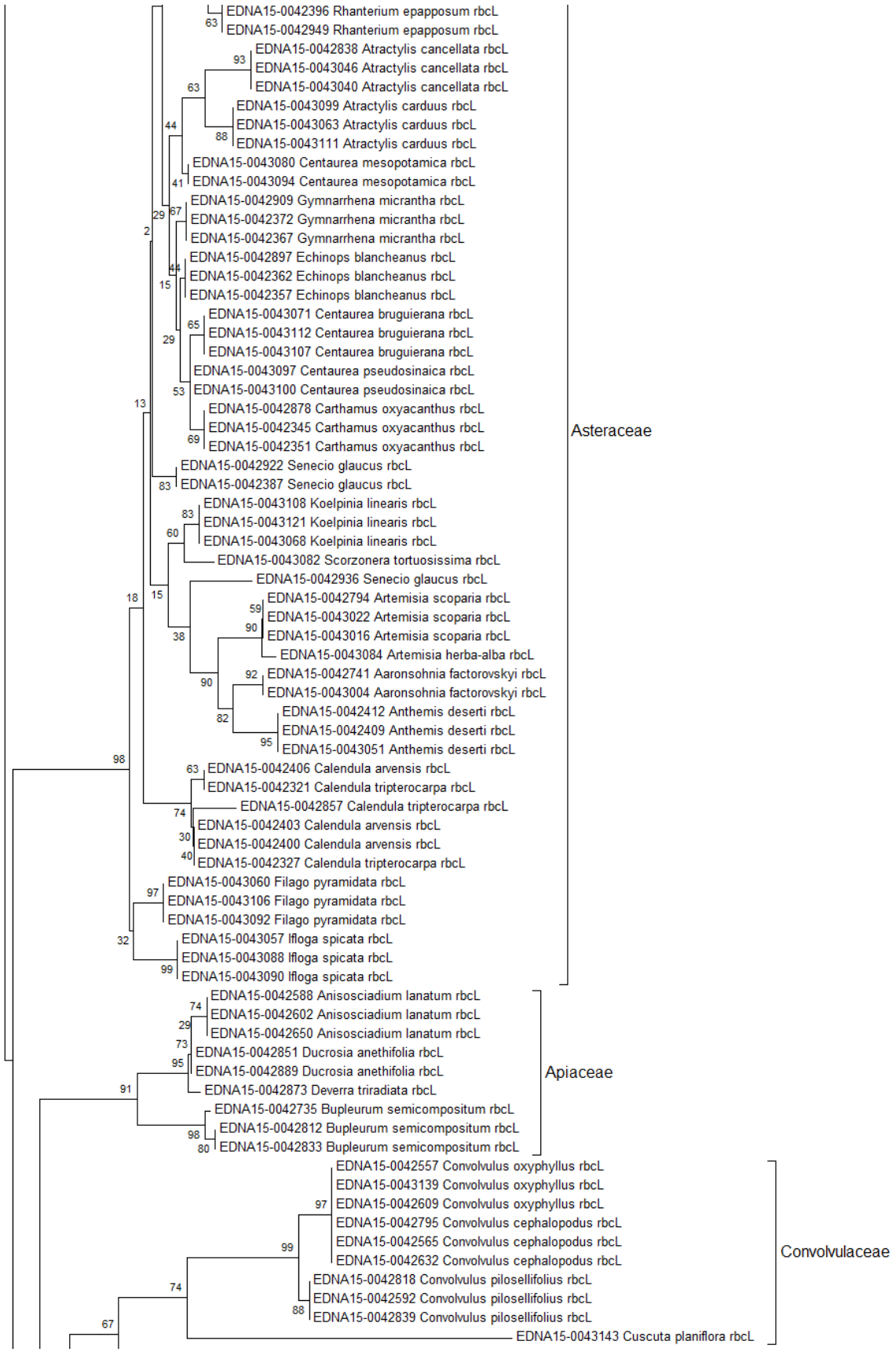


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 7/9]

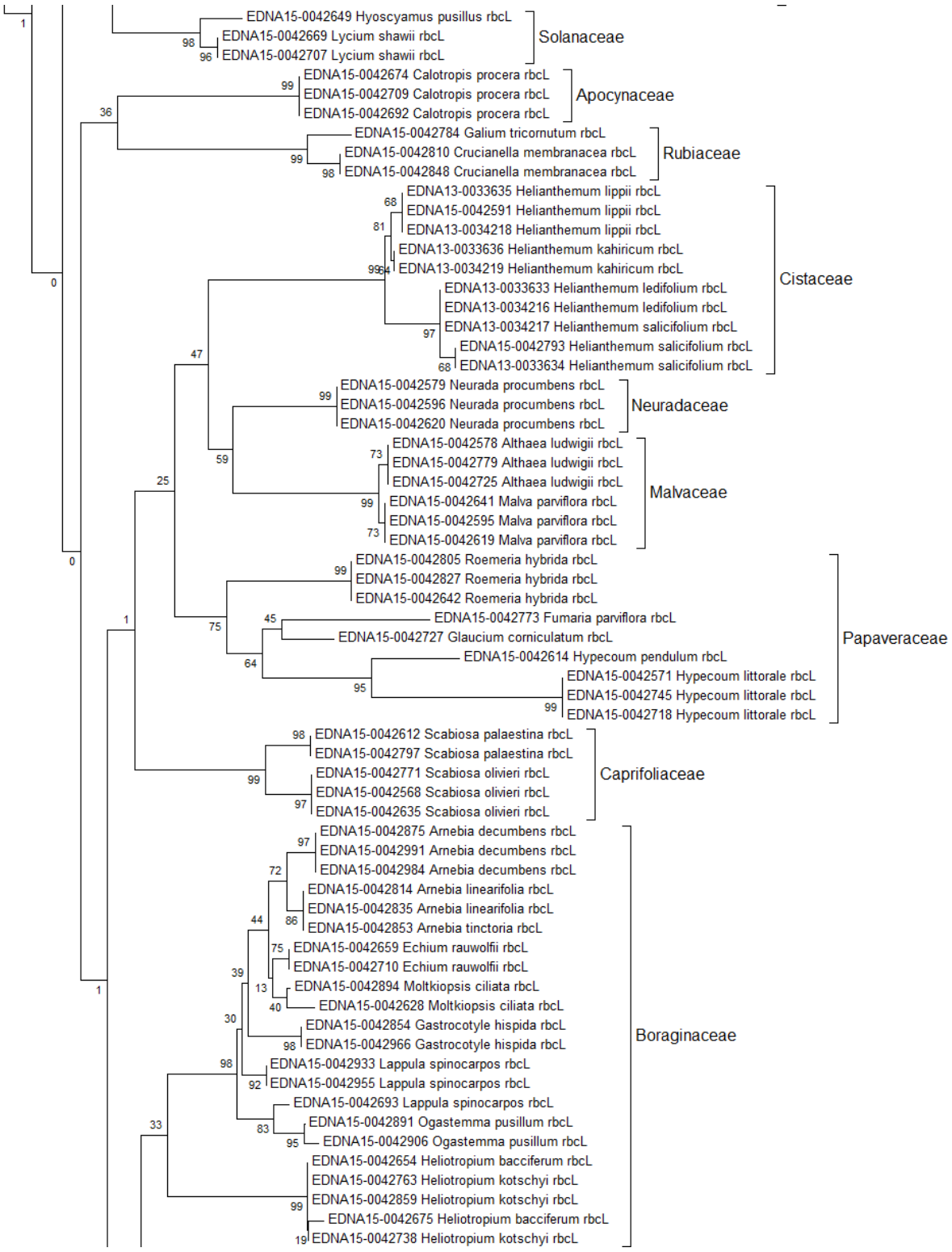


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [Cont. 8/9]

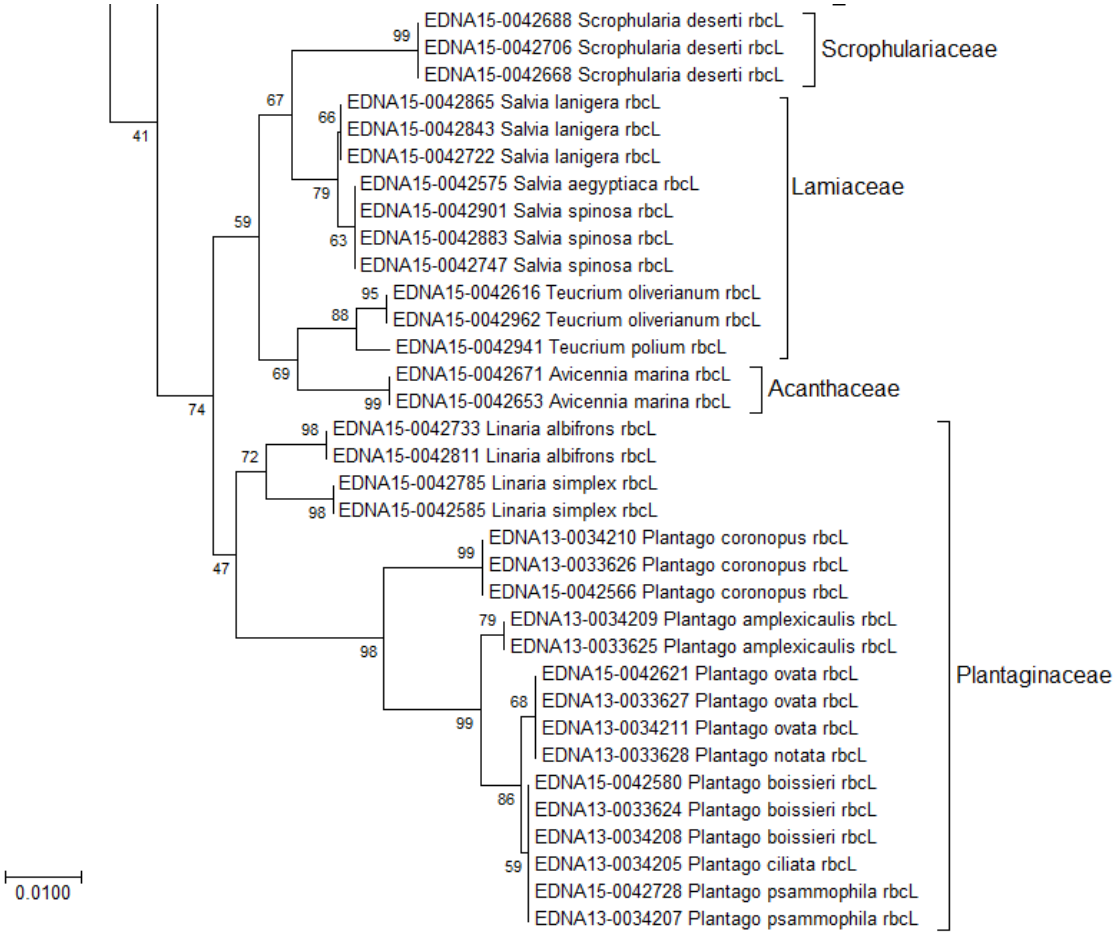


Figure 4.3 Neighbour Joining phylograms for *rbcL* barcodes representing 594 sequences
(values represent % boot strap support with 100 replicates) [END 9/9]

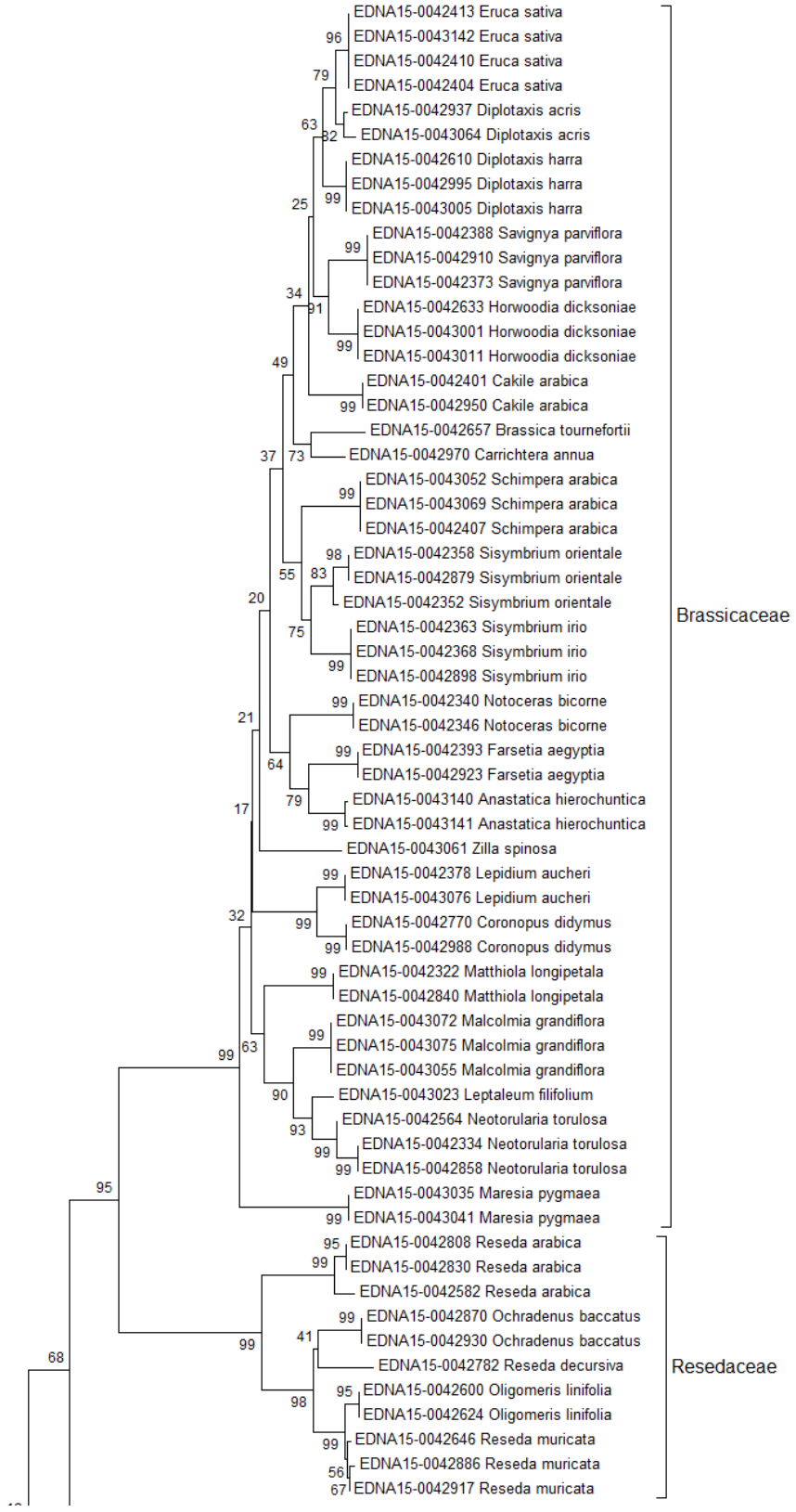


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 1/8]

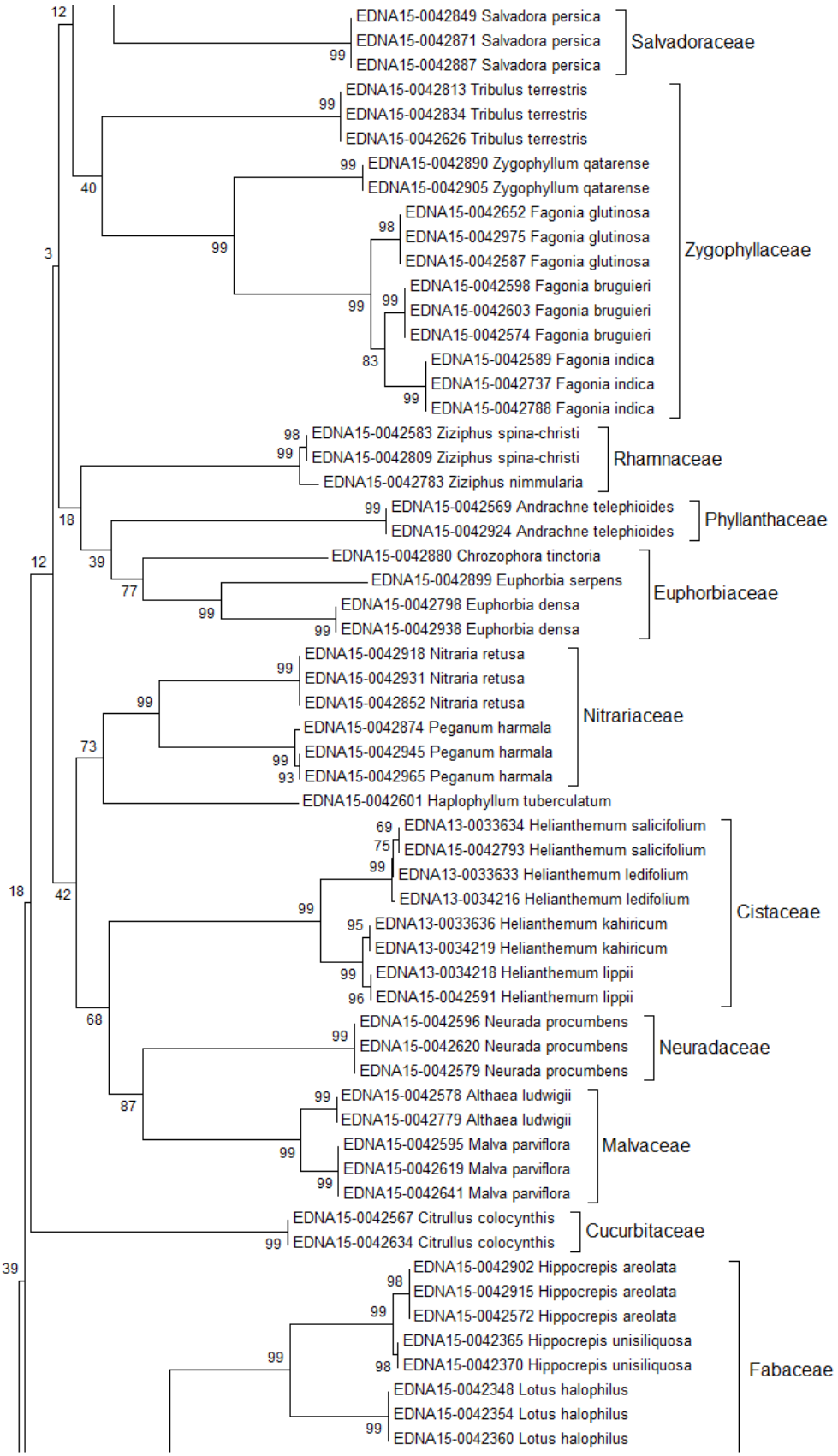


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 2/8]

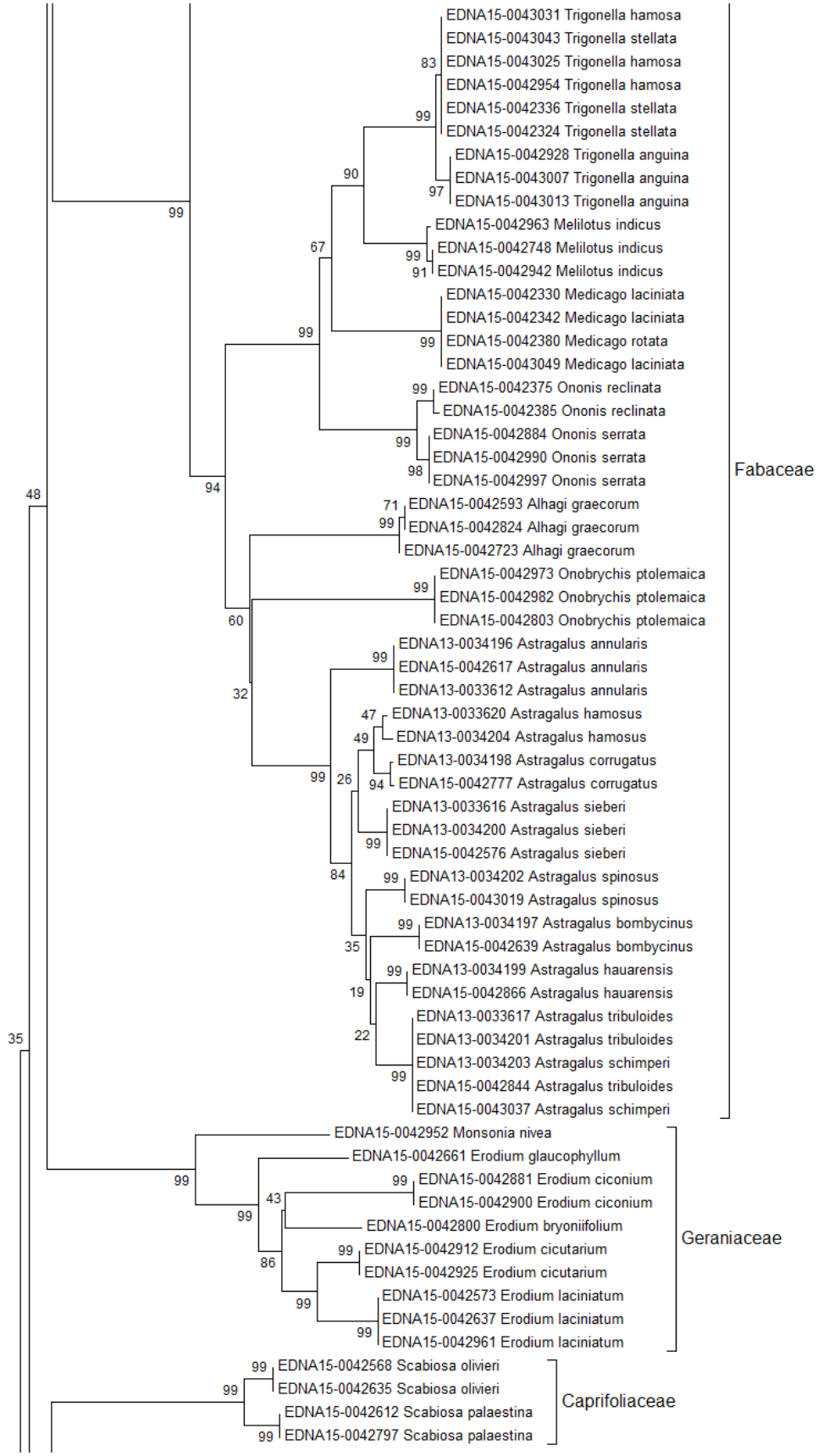


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 3/8]

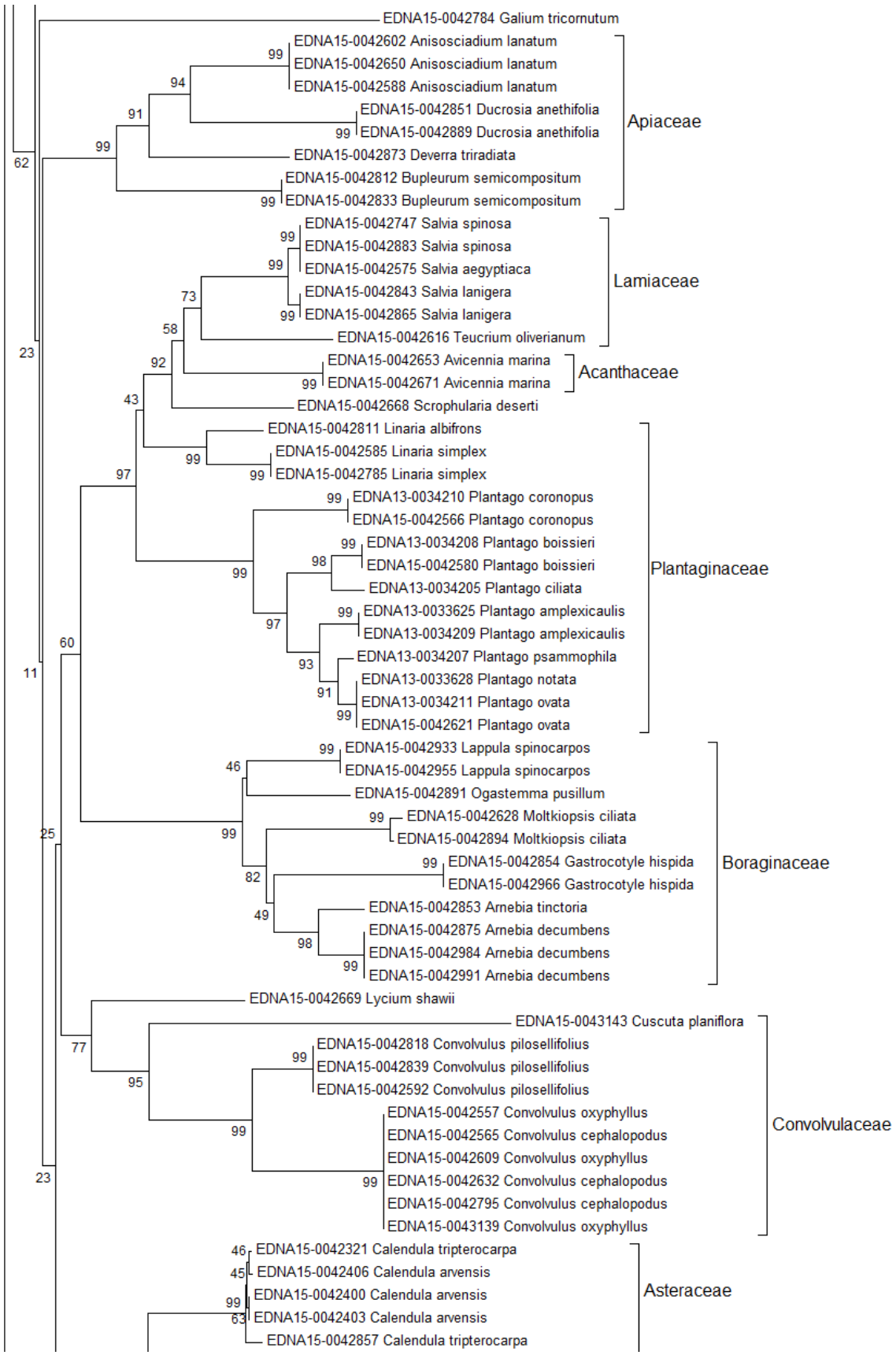


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 4/8]

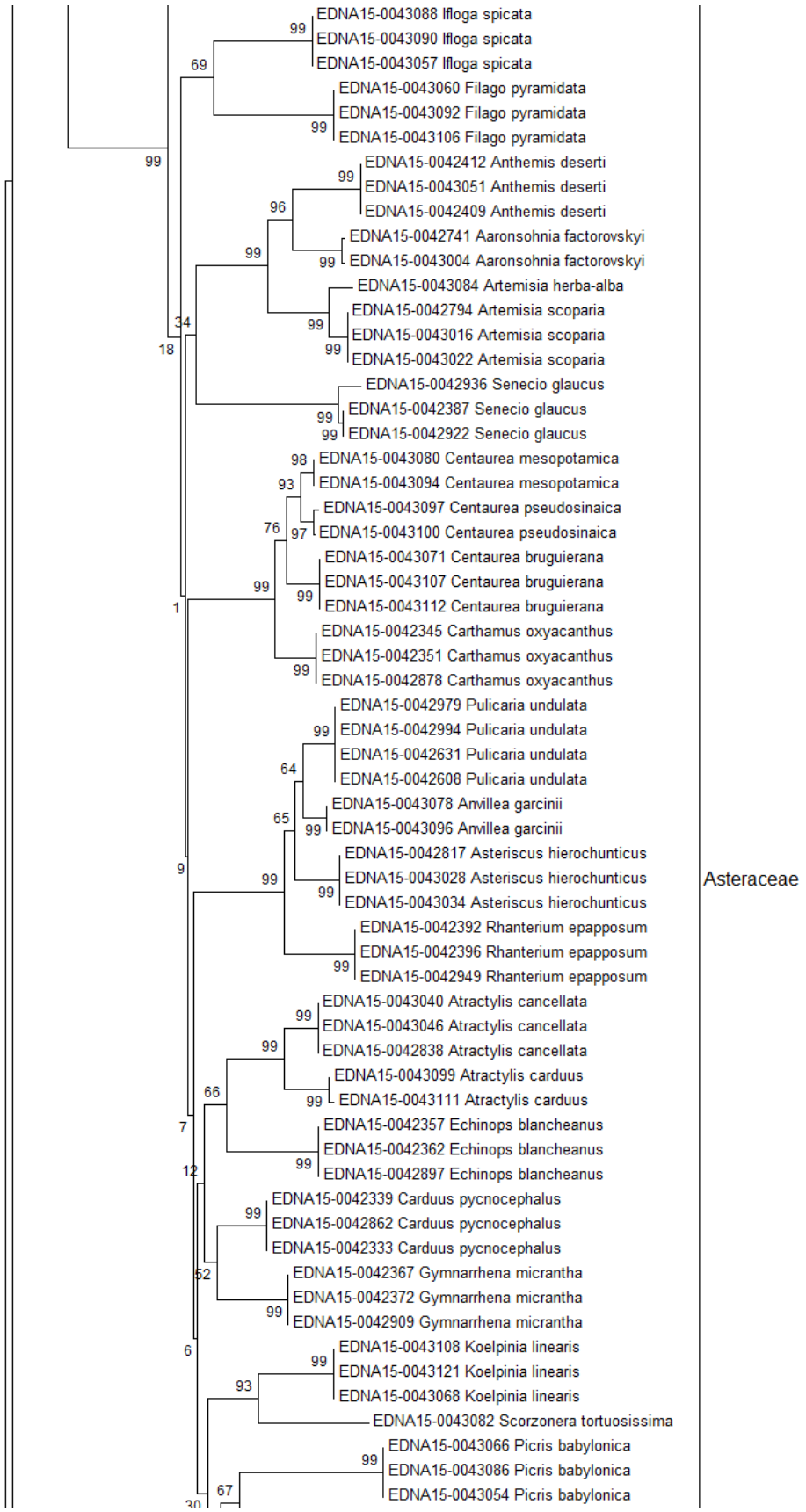


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 5/8]

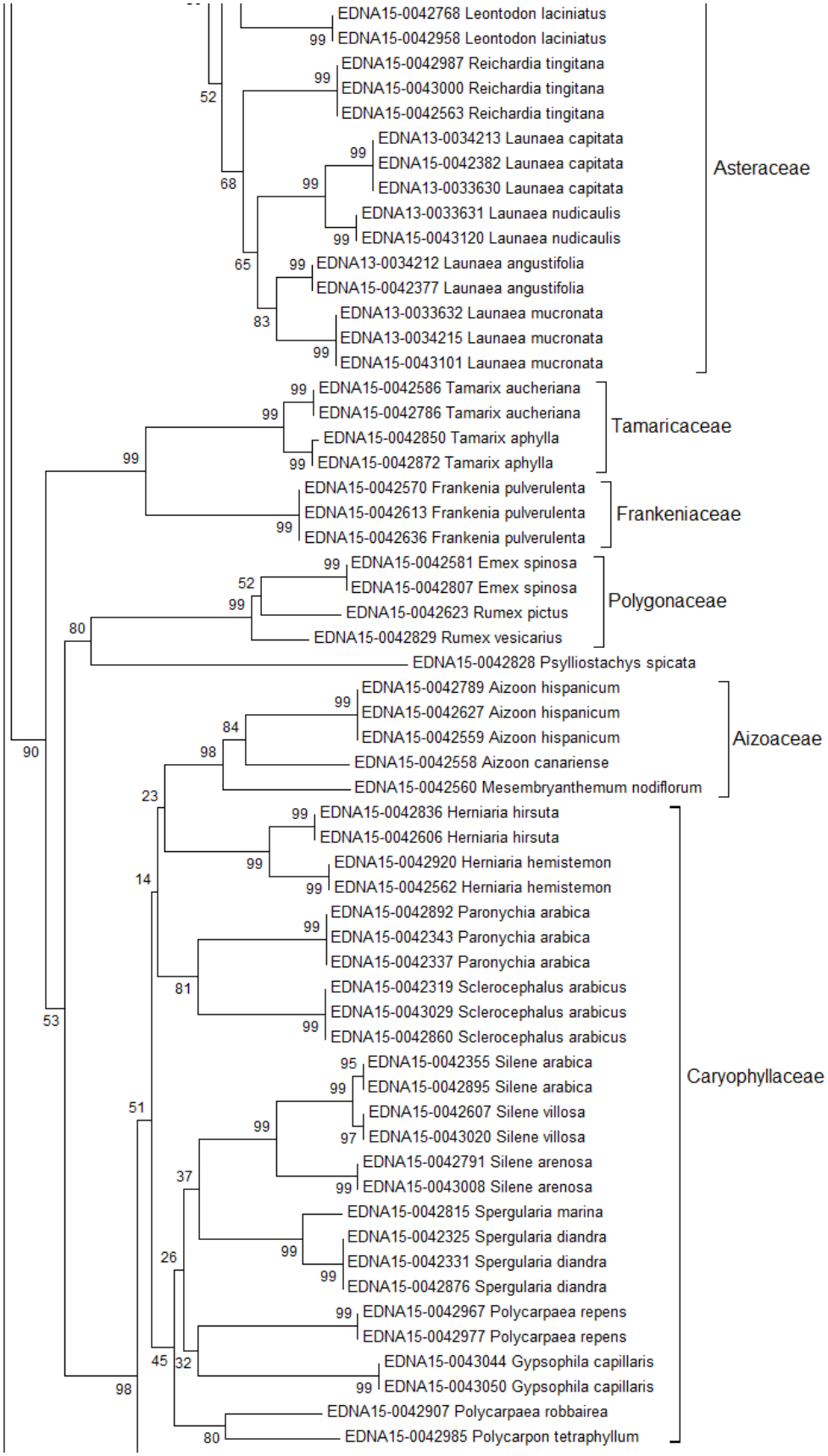


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 6/8]

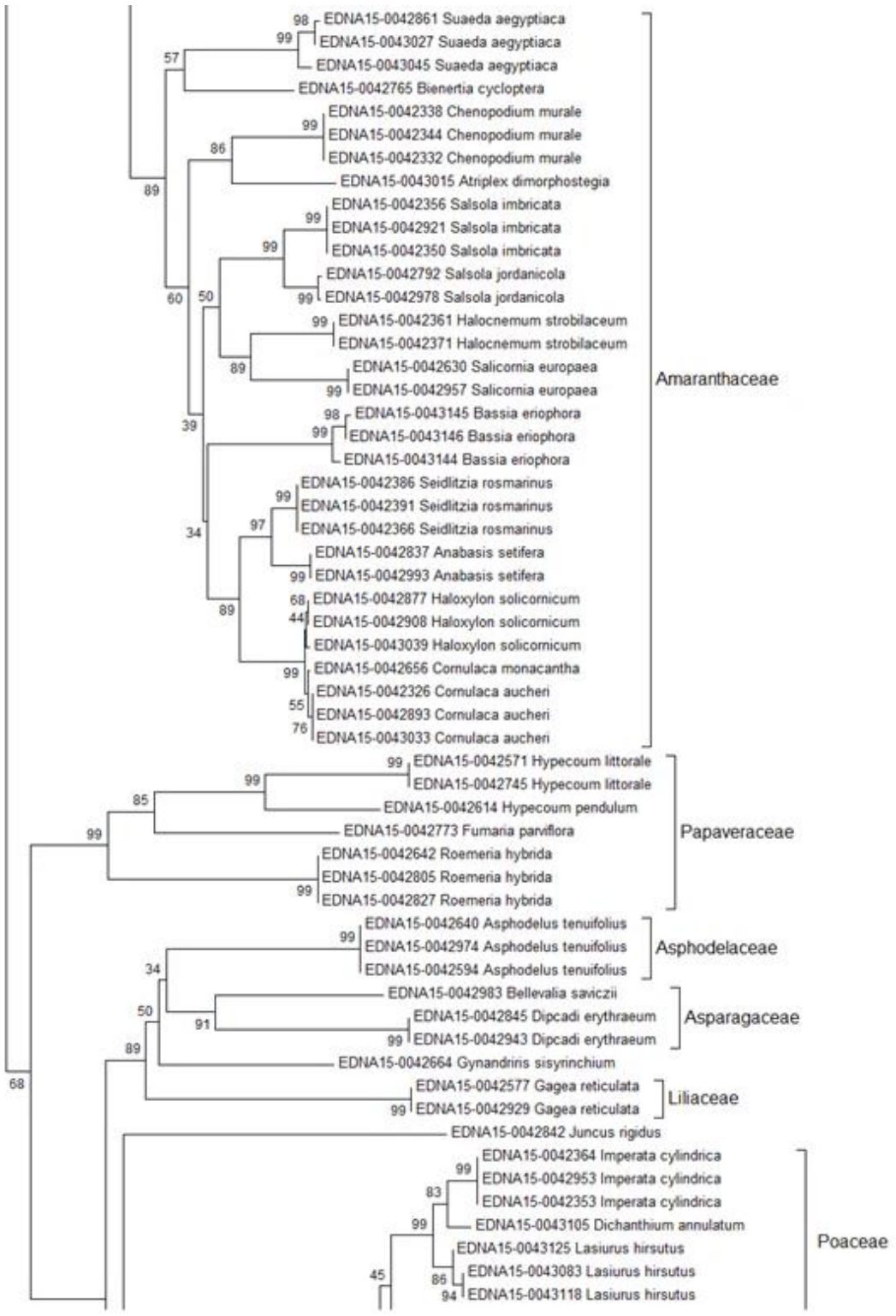


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [Cont. 7/8]

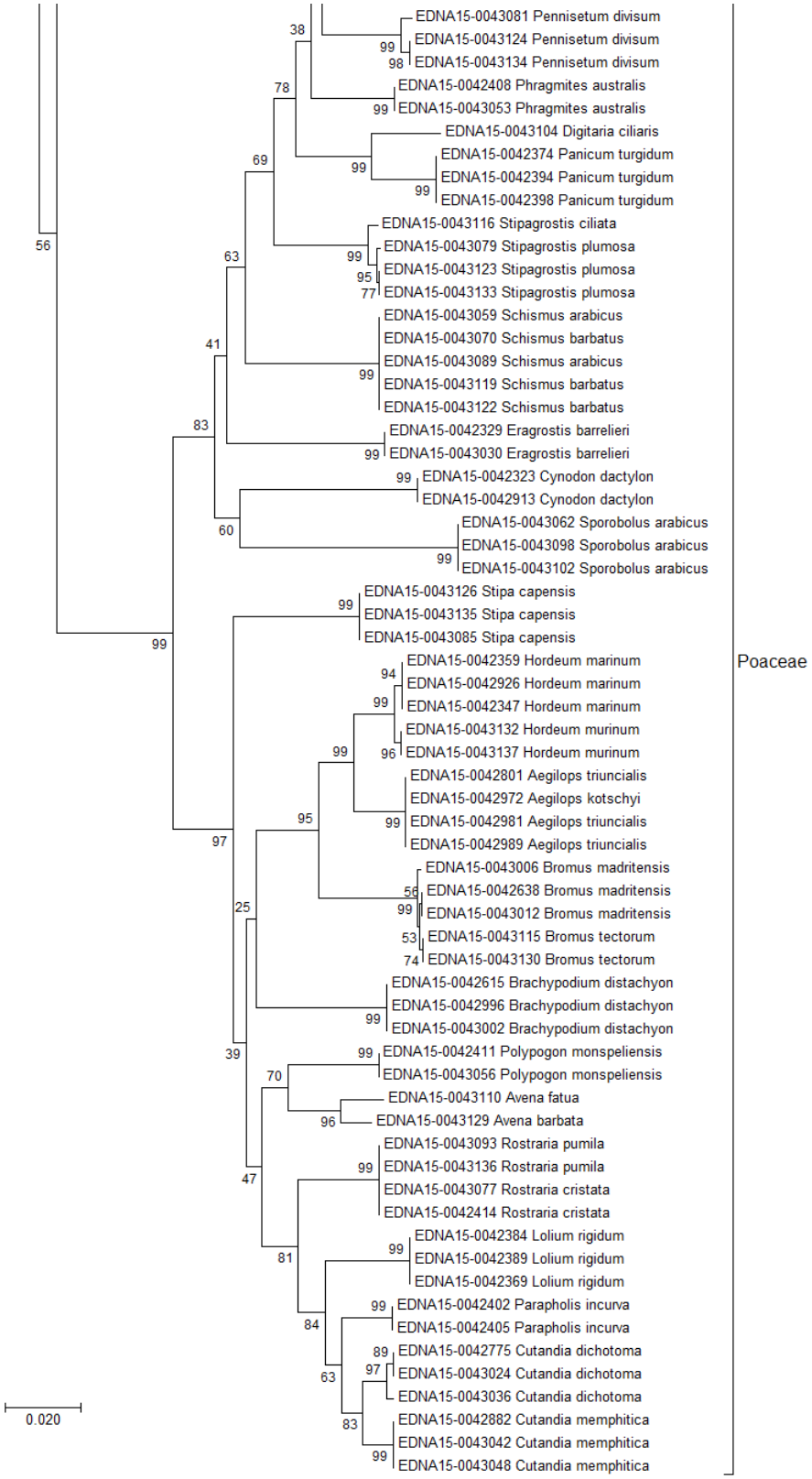


Figure 4.4 Neighbour joining phylograms for combined *rbcL* + ITS2 barcodes representing 480 sequences (values represent % boot strap support with 100 replicates) [End 8/8]

4.3.5 Resolution at the family level

In an attempt to study the DNA barcode regions used and how they represent major plant groups, species resolution for each family was calculated separately from NJ trees (Figures 4.2 to 4.4) representing 19 largest families of the flora (species ≥ 1 included) presented in Figure 4.5. At the family level, species resolution for all combinations (*rbcL*, ITS2, *rbcL*+ITS2) ranged from 25 % for two families (Convolvulaceae and Lamiaceae) up to 100 % for 4 families (Caprifoliaceae, Malvaceae, Nitrariaceae, Zygophyllaceae) (Figure 4.5).

The combined regions *rbcL* + ITS2 resolved greater species clustering at the family level (12 out of 19 families), followed by ITS2 alone (9 out of 19) and *rbcL* (5 out of 19) (Figure 4.5).

For all three combinations (*rbcL*, ITS2, *rbcL*+ITS2) the highest species resolution was recovered by Caprifoliaceae, Malvaceae, Nitrariaceae, and Zygophyllaceae (100% each) (Figure 4.5).

The families Amaranthaceae, Boraginaceae, Caryophyllaceae, Cistaceae, and Poaceae were best resolved by combining *rbcL*+ITS2, 79, 66, 76, 75 and 65 %, respectively (Figure 4.5). Asteraceae is best resolved by ITS2, 93 %. Two families showed better resolution using *rbcL* region, Geraniaceae and Lamiaceae, 66 and 50 %, respectively (Figure 4.5).

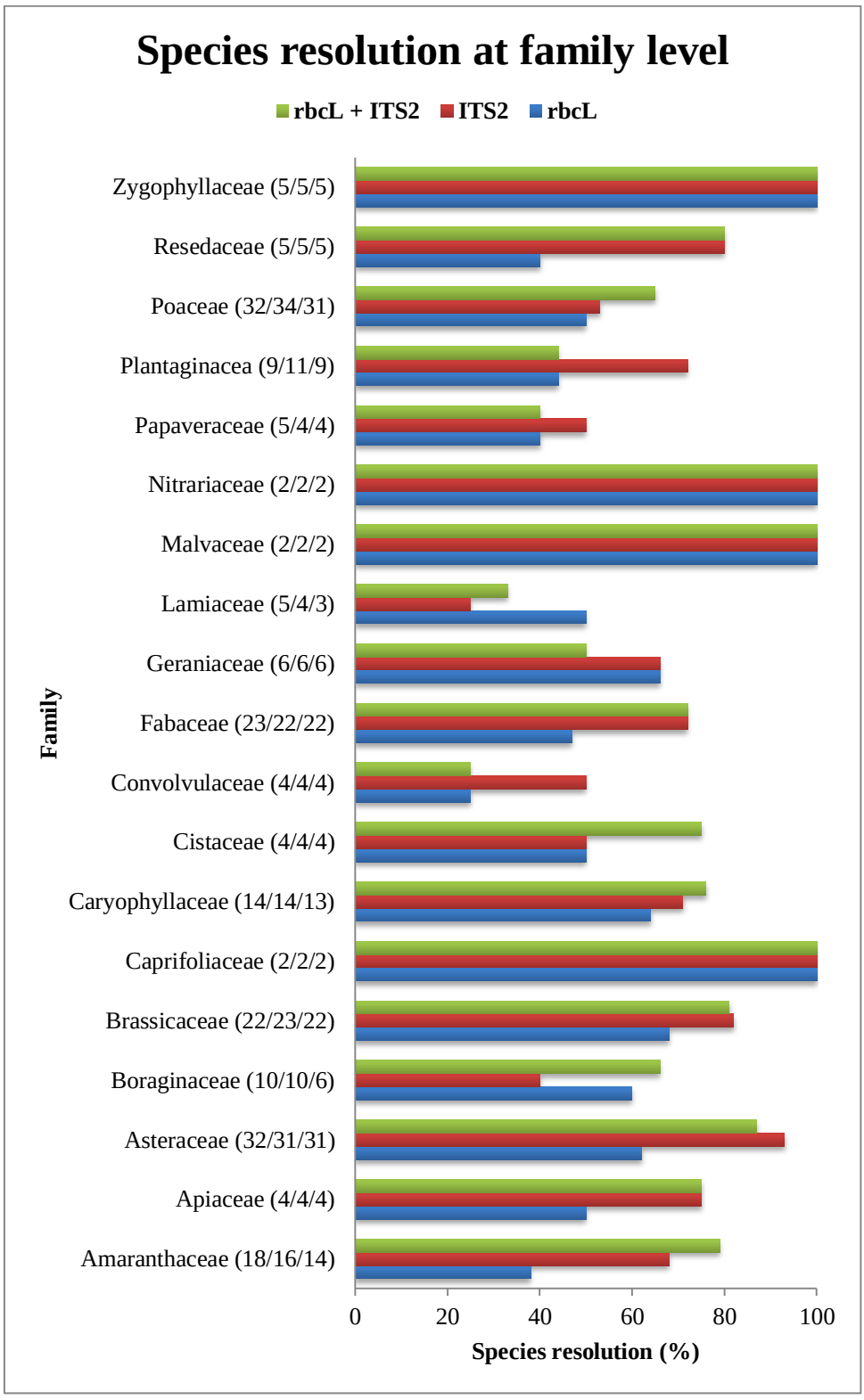


Figure 4.5 Species resolution (%) at family level using NJ trees for *rbcL*, ITS2, and *rbcL* + ITS2 representing 19 families of the flora of Kuwait. Numbers in parentheses refer to the numbers of species for which barcode data were recovered for *rbcL* (blue), ITS2 (red), and *rbcL* + ITS2 (green), respectively.

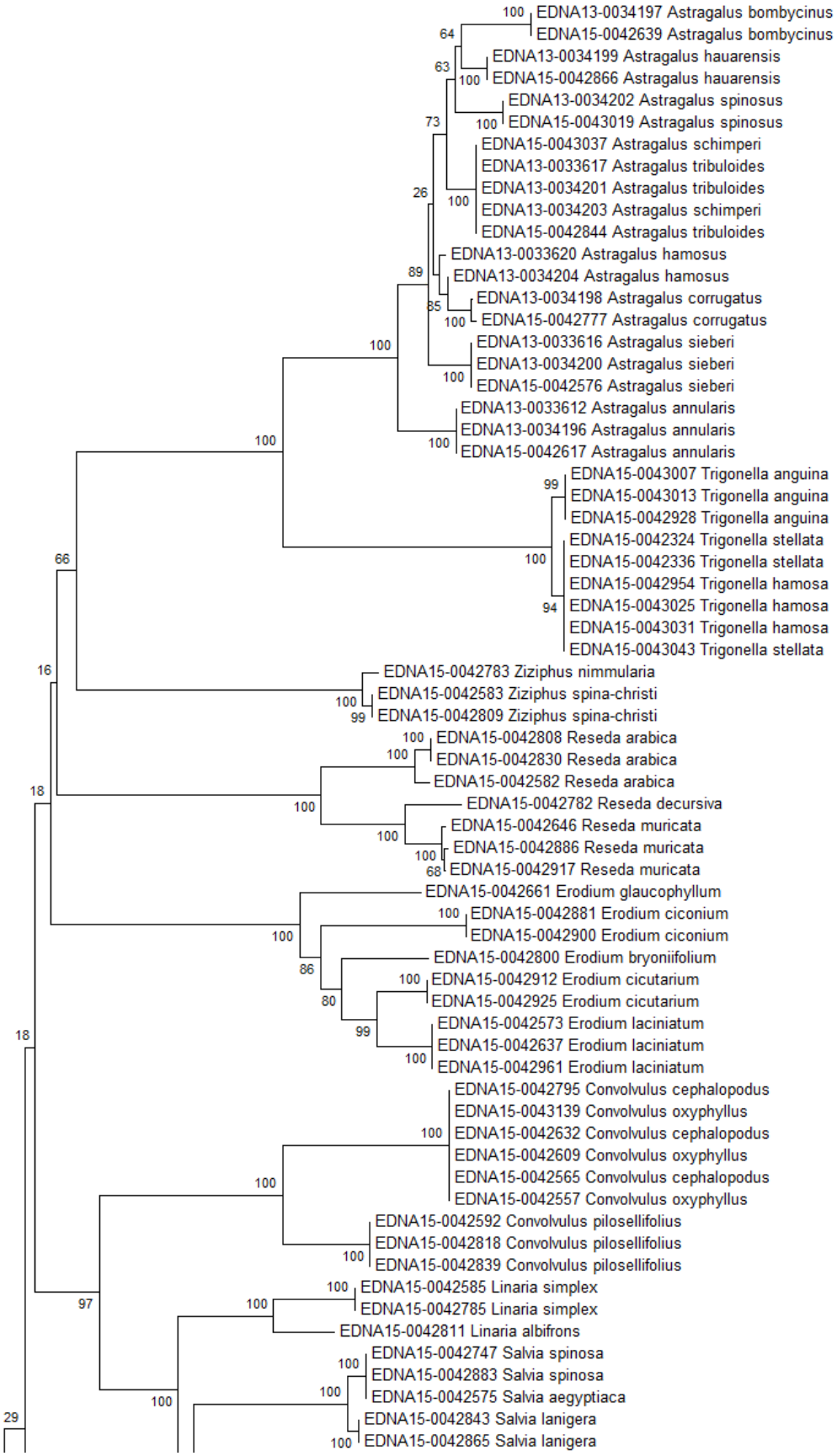
4.3.6 Resolution at the genus level

A total of 94 species represented by 34 genera (including species ≥ 1 per genus) belonging to 17 families of the flora were evaluated to understand the resolution at the genus level. Sixty-five species (69 %) resolved into their genera by *rbcL* + ITS2 region (Table 4.5). The following 16 genera showed 100 % genus resolution: 2 *Salsola* spp. (Amaranthaceae), 2 *Artemisia* spp., 2 *Atractylis* spp., 3 *Centaurea* spp., 4 *Launeae* spp. (Asteraceae), 2 *Diplotaxis* spp., 2 *Sisymbrium* spp. (Brassicaceae), 2 *Scabiosa* spp. (Caprifoliaceae), 2 *Herniaria* spp., 3 *Silene* spp. (Caryophyllaceae), 2 *Hippocrepis* spp., *Ononis* spp. (Fabaceae), 2 *Bromus* spp., 2 *Cutandia* spp., 2 *Hordeum* spp. (Poaceae), and 3 *Fagonia* spp. (Zygophyllaceae) (Table 4.5). Further analysis was performed on genera < 100 % resolutions. A phylogenetic reconstruction using NJ tree method presented in Figure 4.6 showing the relationship of unresolved genera. Three genera belonging to the family Poaceae did not show any resolution due to paraphyletic relationships: *Aegilops* spp., *Rostraria* spp. and *Schismus* spp. (Table 4.5 and Figure 4.6). Other genera with < 100 % resolution is either due to paraphyletic relationships between two or more species of the same genus (e.g. paraphyletic relation between *Astragalus schimperi* and *A. tribuloides*; *Trigonella stellata* and *T. hamosa*; *Convolvulus oxyphyllus* and *C. cephalopodus*) or due to some species being represented by only one individual per genus which lowers the % of species resolution (Table 4.5 and Figure 4.6).

Table 4.5 Genera in which more than one species resolved by *rbcL* + ITS2

Family	Genus	Number of species sampled/ resolved *rbcL* + ITS2
Aizoaceae	Aizoon	2/1*
Amaranthaceae	Cornulaca	2/1*
Amaranthaceae	Salsola	2/2
Asteraceae	Artemisia	2/2
Asteraceae	Atractylis	2/2
Asteraceae	Centaurea	3/3
Asteraceae	Launaea	4/4
Brassicaceae	Diplotaxis	2/2
Brassicaceae	Sisymbrium	2/2
Caprifoliaceae	Scabiosa	2/2
Caryophyllaceae	Herniaria	2/2
Caryophyllaceae	Silene	3/3
Caryophyllaceae	Spergularia	2/1*
Cistaceae	Helianthemum	4/3*
Convolvulaceae	Convolvulus	3/1*
Fabaceae	Astragalus	9/7*
Fabaceae	Hippocrepis	2/2
Fabaceae	Ononis	2/2
Fabaceae	Trigonella	3/1*
Geraniaceae	Erodium	5/3*
Lamiaceae	Salvia	3/1*
Papaveraceae	Hypecoum	2/1*
Plantaginaceae	Linaria	2/1*
Plantaginaceae	Plantago	7/3*
Poaceae	Aegilops	2/0*
Poaceae	Bromus	2/2
Poaceae	Cutandia	2/2
Poaceae	Hordeum	2/2
Poaceae	Rostraria	2/0*
Poaceae	Schismus	2/0*
Poaceae	Stipagrostis	2/1*
Resedaceae	Reseda	3/2*
Rhamnaceae	Ziziphus	2/1*
Zygophyllaceae	Fagonia	3/3

Asterisk (*) Genus resolution < 100 %


EDNA13-0034197 Astragalus bombycinus
EDNA15-0042639 Astragalus bombycinus
EDNA13-0034199 Astragalus hauarensis
EDNA15-0042866 Astragalus hauarensis
EDNA13-0034202 Astragalus spinosus
EDNA15-0043019 Astragalus spinosus
EDNA15-0043037 Astragalus schimperi
EDNA13-0033617 Astragalus tribuloides
EDNA13-0034201 Astragalus tribuloides
EDNA13-0034203 Astragalus schimperi
EDNA15-0042844 Astragalus tribuloides
EDNA13-0033620 Astragalus hamosus
EDNA13-0034204 Astragalus hamosus
EDNA13-0034198 Astragalus corrugatus
EDNA15-0042777 Astragalus corrugatus
EDNA13-0033616 Astragalus sieberi
EDNA13-0034200 Astragalus sieberi
EDNA15-0042576 Astragalus sieberi
EDNA13-0033612 Astragalus annularis
EDNA13-0034196 Astragalus annularis
EDNA15-0042617 Astragalus annularis
EDNA15-0043007 Trigonella anguina
EDNA15-0043013 Trigonella anguina
EDNA15-0042928 Trigonella anguina
EDNA15-0042324 Trigonella stellata
EDNA15-0042336 Trigonella stellata
EDNA15-0042954 Trigonella hamosa
EDNA15-0043025 Trigonella hamosa
EDNA15-0043031 Trigonella hamosa
EDNA15-0043043 Trigonella stellata
EDNA15-0042783 Ziziphus nimmularia
EDNA15-0042583 Ziziphus spina-christi
EDNA15-0042809 Ziziphus spina-christi
EDNA15-0042808 Reseda arabica
EDNA15-0042830 Reseda arabica
EDNA15-0042582 Reseda arabica
EDNA15-0042782 Reseda decursiva
EDNA15-0042646 Reseda muricata
EDNA15-0042886 Reseda muricata
EDNA15-0042917 Reseda muricata
EDNA15-0042661 Erodium glaucophyllum
EDNA15-0042881 Erodium ciconium
EDNA15-0042900 Erodium ciconium
EDNA15-0042800 Erodium bryoniifolium
EDNA15-0042912 Erodium cicutarium
EDNA15-0042925 Erodium cicutarium
EDNA15-0042573 Erodium laciniatum
EDNA15-0042637 Erodium laciniatum
EDNA15-0042961 Erodium laciniatum
EDNA15-0042795 Convolvulus cephalopodus
EDNA15-0043139 Convolvulus oxyphyllus
EDNA15-0042632 Convolvulus cephalopodus
EDNA15-0042609 Convolvulus oxyphyllus
EDNA15-0042565 Convolvulus cephalopodus
EDNA15-0042557 Convolvulus oxyphyllus
EDNA15-0042592 Convolvulus pilosellifolius
EDNA15-0042818 Convolvulus pilosellifolius
EDNA15-0042839 Convolvulus pilosellifolius
EDNA15-0042585 Linaria simplex
EDNA15-0042785 Linaria simplex
EDNA15-0042811 Linaria albifrons
EDNA15-0042747 Salvia spinosa
EDNA15-0042883 Salvia spinosa
EDNA15-0042575 Salvia aegyptiaca
EDNA15-0042843 Salvia lanigera
EDNA15-0042865 Salvia lanigera

(Continued)

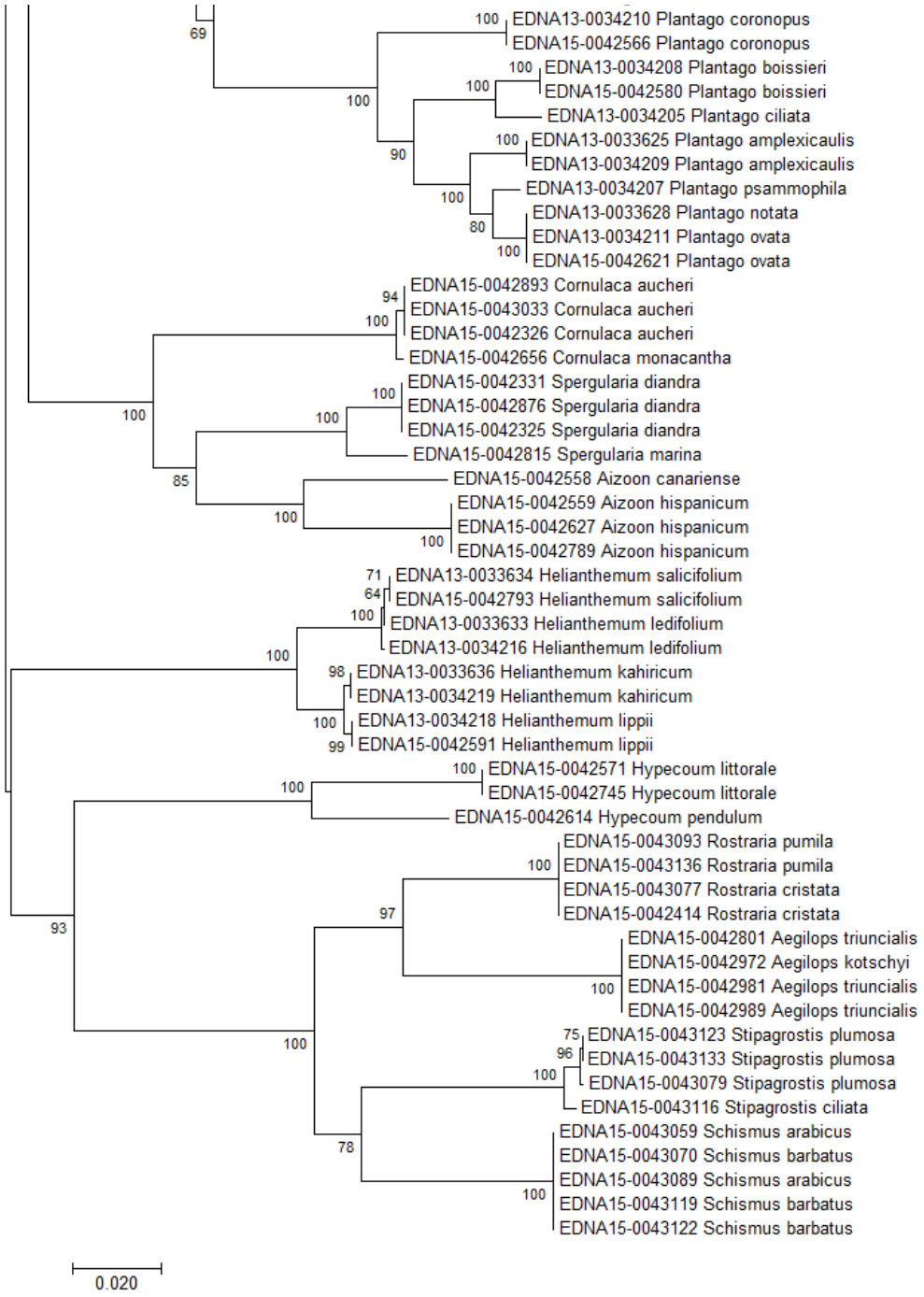


Figure 4.6 Neighbour joining phylograms representing 19 genera < 100 % resolution for combined regions *rbcL* + ITS2 barcodes
(values represent % boot strap support with 1000 replicates)

4.3.7 Results summary

The main findings which arise from this set of analyses are as follows:

1. The highest percentage of DNA sequence recoverability is represented by DNA extracted from silica dried fresh leaves (84 to 96 %).
2. Sequencing efficiency using one set of universal primers was highest for *rbcL* region (82 %) and lowest for ITS2 (72.5 %).
3. High levels of sequence divergence were detected using ITS2 region (69 %) compared to *rbcL* with lower resolution (58.2 %).
4. The combined regions, *rbcL* + ITS2 enhanced the resolution and resolved greater species-specific clusters (70.5 %) than used singly.
5. Resolution at the family level was best resolved using the combined region (12 out of 19 families showed higher sequence divergence).
6. At the genus level, 69 % of species determined into there genera by combining *rbcL* + ITS2 regions.

4.4 Discussion

This study represents one of very few attempts to barcode an entire flora of a country or regional floras, e.g. Sareela, et al., 2013; de Vere et al., 2012; Bruni et al., 2012; Kuzmina et al., 2012; Burgess et al., 2011. The results here are based on two barcode regions (*rbcL* and ITS2) used singly and in combination. DNA extracted from freshly collected plant material performed greater than using herbarium specimens for both regions. This is a common problem being confronted by many DNA barcoding studies working with herbarium material due to damage and high rates of specimen degradation (Staats et al., 2011). The sequence recovery from DNA extractions for *rbcL* and ITS2 using freshly collected material was quite high (96 % and 84 %, respectively) compared to herbarium material (71 % and 63 %, respectively). The determination of whether a DNA fragment can serve as a barcode is to evaluate its universality by showing high rates of PCR and quality of bidirectional sequences recovered from all individuals sampled (CBOL, 2009).

4.4.1 Sequence recovery

Across the two DNA regions, 1,117 sequences were recovered representing both regions by 77.5 % of the total accessions sampled (1442). *rbcL* produced greater amplification success using one pair of primers and resulted in high quality of sequences (82 %) and for ITS2 (72.5 %), consistent with the results of several other studies that sampled the same regions broadly across land plants (CBOL, 2009; Chen et al., 2010; Burgess et al, 2011; de Vere et al., 2012; Liu et al., 2015). The *rbcL* primers applied here (*rbcL*-aF and *rbcL*-aRev) were used previously by Kress et al. (2009) showed higher sequencing recovery (93 %) (Kress et al., 2009). In addition, ITS2 primers (ITS2-S2F/ S3R) applied by Chen et al. (2010) recovered 93.8 % of the sequences (Chen et al., 2010). Mainly, DNA extractions from herbarium specimens resulted in lower amplification success and sequence recovery (herbarium material: *rbcL*= 71 % and ITS2= 63 %) due to the degraded level of DNA specimens. DNA extraction from herbarium material requires greater caution than freshly collected material, to avoid any contaminants from other organisms (e.g. fungal, algae), which is difficult to detect when sampling but may easily be limited by using more specific primers, or in the case of fungi, working on plastidial regions. Furthermore, de Vere et al. (2012) noticed that some orders of flowering plants did not sequence well using herbarium material for either *rbcL* or *matK* regions and recommended the use of fresh plant material for the following plant orders: Oxalidales, Liliales, Myrtales, Saxifragales and Asparagales (de Vere et al., 2012). Also, obtaining high-quality bidirectional sequences with barcoding regions is important, e.g. *trnH-psbA* demonstrated good amplification across land plants; however the limitation was obtaining high-quality bidirectional sequences (CBOL, 2009; Mahadani et al., 2013).

4.4.2 Species resolution

Levels of species discrimination are critical in evaluating barcode regions. The two markers demonstrated here provided good species resolution using the phylogenetic reconstruction tree based method (NJ) across all successfully sequenced plants, 58 % for *rbcL* and 69 % for ITS2 (Table 4.4). The low species resolution was caused by high percentage of PCR amplification failure (*rbcL* = 14 % and ITS2 = 19 %) which

resulted in fewer accessions per species. A representative of one barcode sequence could not resolve species-specific clustering and possibly create paraphyletic relationships with closely related species. Therefore, to improve the species resolution, rather than using one pair of primers per region, (although universality is necessary to consider in DNA barcoding studies), a second set of primers should be considered in resolving PCR amplification failures. In situations where it is not even possible to amplify, (i.e. due to degraded DNA samples) a new DNA extraction from the accession should be considered.

The combined regions (*rbcL*+ITS2) increased species resolution up to 70.5 % (12.3 % higher than *rbcL* alone and 1.5 % higher than ITS2) and the findings here in combination with the findings in Chapter 3 (testing 5 different DNA regions) support the designation of the combined region *rbcL*+ ITS2 as the main barcoding regions for identification purposes of the flora of Kuwait. The combination of plastid (*rbcL*) and nuclear (ITS2) information increases the power of taxonomic identification amongst closely related species. Findings in the current study agree with the recently published study by Liu et al. (2015), which showed similar species resolution of 71 % across a broad range of angiosperms (531 species) by combining *rbcL*+ITS2 regions after the evaluation of five DNA barcode regions, *rbcL*, *matK*, ITS, ITS2 and *trnH-psbA*. (Liu et al, 2015). Similar molecular analysis methods used in the current study (MUSCLE alignment and neighbour joining tree based method) were applied by Liu et al. (2015).

Several local DNA barcoding libraries have been constructed previously which mainly focused on a limited region (Table 4.1). Theodoridis et al. (2012) barcoded native plants (80 species) of Lamiaceae family belonging to Greece and Turkey and showed that some species of the family are partial to fully unresolvable (e.g. Salvia, Mentha) due to hybridization (Theodorids et al., 2012).

Burgess et al. (2011) established a barcoding database for a local temperate of Canadian flora on 436 species based on two combined regions (*rbcL+matK*). Their study showed that several genera are fully not resolvable (e.g. *Agrostis, Crataegus, Verbena, Erigeron*) due to polytypic genera contained species (Burgess et al., 2011).

In the same country, Saarela et al. (2013) DNA barcoded 490 species of the Canadian Arctic flora using the same combined regions (*rbcL+matK*) showed that 42 genera

(belonging to 17 families) did not distinguish at all, comparing the families of undistinguishable genera with the current study (Table 4.5) several families matched e.g. Amaranthaceae, Asteraceae, Brassicaceae, Caryophyllaceae, Papaveraceae, Poaceae and Amaranthaceae and only one genus, *Suaeda* spp. (Saarela et al., 2013).

Accurate genus level identification is important for poorly described groups (Little, 2011). From the 34 largest genera of the flora 16 genera were capable of resolving all species with 100 % resolution using combined region *rbcL*+ITS2, 3 resolved between 60-70 %, 12 resolved between 40-50 %, and 3 genera did not show any resolution (*Aegilops, Rostaria,* and *Schismus*) due to paraphyletic relationships (Figure 4.6). Species recovery of four largest families in the flora was best resolved using the combined regions *rbcL* + ITS2, which showed species resolution range from 65 % to 95 %, represented by Brassicaceae (95 %), Asteraceae (93 %), Fabaceae (75 %), and Poaceae (65 %) (Figure 4.5). These findings indicate that the combined region is capable of resolving closely related species belonging to the largest families of the flora with higher resolution than using a single region alone (Figure 4.5).

The formations of paraphyletic relationships amongst closely related species within a genus are spotted within 18 genera (Table 4.5 and Figure 4.6) e.g. *Suaeda* (Amaranthaceae) did not resolve any species and showing paraphyletic relationships within two species represented by *S. aegyptiaca* and *S. vermiculata*. *Suaeda* is a taxonomically difficult genus, mainly because of the large numbers of species (ca. 110 species) divided into 2 subgenus (*Brezia* and *Suaeda*) and 9 sections and distinguishing of morphologically characters are usually few and present only after flowering as investigated by Schutze et al. (2003).

Aegilops is a genus in Poaceae family (generally known as ‘goat-grasses’), consists of more than 20 species and identification is challenging due to their vast morphological similarities (Badaeva et al., 2004; Keshavarzi et al., 2007). In Kuwait, *Aegilops* is represented by three species *A. triuncialis*, *A. bicornis* and *A. kotschyi*. Also, another complex genus from the grass family is *Hordeum* (barley) with more than 30 species, only two species represent the flora *H. marinum* (section *Marina*) and *H. murinum* (section *Trichostachys*) (Blattner et al., 2009).

The genus *Ononis* (Fabaceae) comprises of 69 species is also known to be difficult to identify due to high similarity in morphological features. *Ononis* consists of five major lineages referred to as clades I to V and both *O. reclinata* and *O. serrata* (representatives of the flora of Kuwait) are known to be classified in Clade V (Turini et al., 2010).

In such cases where the identification of unknown plant specimen could not be resolved to species level by the DNA database, and only resolves to the genus level a taxonomic key of the flora would be handy to identify the specimen down to the species level in less time, since an average of 2-3 species per genus are mainly represented in the flora.

Studying the distribution of species is an essential requirement for conservation management especially when alien species are present which could compete with local plants and possess a high risk to the biodiversity, a control measure is required. Therefore, applying simple DNA molecular techniques is quick and reliable for vegetation monitoring, by blasting unknown plant sequences against local DNA database (considering it covers all the native and naturalised plants of the flora) could efficiently identify whether the species match sequences from the local DNA database or else may be considered as an alien species.

The use of BLASTn searches resulted in low species matches for both barcode regions, 31 % *rbcL* and 47 % ITS2, compared to tree-based methods showing higher species resolution (58 % *rbcL* and 69 % ITS2). The entire flora contains a manageable number of species per genera (Table 4.5) which supports the variation amongst closely related species. Thus, increases the resolution of species using phylogenetic reconstruction tree-based methods. On the other hand, NCBI database using BLASTn (similar sequences blast) could be useful to determine whether contamination was an issue amongst barcode sequences of the flora.

This chapter represents a DNA reference library of almost 70 % the flora of Kuwait using the barcode regions *rbcL* and ITS2, and in the meanwhile, will act as the primary source of DNA database for local Kuwaiti plant molecular identifications. The DNA database is accessible to researchers through The Barcode of Life Data (BOLD) systems database under the title ‘Barcoding the flora of Kuwait using *rbcL*

and ITS2', freely available web platform database (Ratnasingham and Hebert, 2013 accessible online: http://www.boldsystems.org/). BOLD delivers an online database for the collection and management of specimens, distributional, and molecular data. Interestingly, BOLD shares an integrated data exchange pipeline with NCBI (GenBank) that allows for automatic submission of data to GenBank. Specimens with BOLD and GenBank accessions representing successfully barcoded sequences for *rbcL* and ITS2 are listed in Appendix 4.1.

4.4.3 Conclusion

The current study provides key information on the expected rates of sequence recovery and species resolution for the combined regions (*rbcL* + ITS2). Practically, the results described here directly provide information at species, genus and family level for researchers interested in identifying unknown plants and conducting biodiversity surveys on the local flora of Kuwait. The establishment of a local DNA barcoding library reference of the flora is valuable for a broad range of potential ecological applications, including the reconstruction of previous vegetation (Yoccoz et al., 2012), identifying invasive species (Van De Wiel et al., 2009), analysing the diets of mammals (Jurado-Rivera et al., 2009). In the next chapter, the DNA barcode database will be used in studying DNA extracted from soil samples collected from the rich and poor habitat of Kuwait and comparing plant patterns of below ground with above ground diversity by applying Next Generation Sequencing methods (Sonstebo et al., 2010).

4.4.4 Suggested for future research

- As in many DNA barcoding studies, there were some accessions which were consistently difficult to amplify and sequence for both regions (*rbcL* and ITS2). Therefore, continuing to generate a complete DNA database for all species would be ideal. Thus, by testing different primers or else start again by extracting DNA from the same sample or another specimen if available.

- Fresh plant material is always preferable due to the different age and degradation levels of herbarium specimens. Additionally, including more than three accessions per species (5-6 where possible) to avoid ending with one accession per species.

- The addition of a third marker would contribute to a more robust dataset and help resolve relationships between closely related species. I would suggest two markers, the coding region *matK* and/ or non-coding region *trnL*, both capable of showing high discriminatory power in previous studies.

Appendix 4.1 A List of Specimens from which DNA was extracted to establish DNA barcode library for the flora of Kuwait. Includes BOLD, GenBank accessions, EDNA numbers, collection ID, collector name and number, year of collection, and locality/ region of collection.

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Aaronsohnia factorovskyi	EDNA15-0043010	MTA003-16			KTUH060	R Halwagy 1087	1972	Herbarium	Al-Dibdibah 15 KM N of Salmy
Aaronsohnia factorovskyi	EDNA15-0043004	MTA001-16	KX282506	KX281954	MTA522	M Abdullah MTA522	2013	Fresh	Abdali
Aaronsohnia factorovskyi	EDNA15-0042741	MTA002-16	KX282507	KX281955	MTA568	M Abdullah MTA568	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Acacia pachyceras	EDNA15-0042390	MTA004-16	KX282508		MTA213	M Abdullah MTA213	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Aegilops kotschyi	EDNA15-0042720	MTA007-16	KX282509	KX281956	KTUH427	KT Mathew 2731	1996	Herbarium	Road to Ahmadi along King Fahad Highway
Aegilops kotschyi	EDNA15-0042940	MTA005-16			KTUH428	M Halwagy 1063	1972	Herbarium	Jal Az-Zor
Aegilops kotschyi	EDNA15-0042972	MTA006-16			KTUH429	M Al-Dosari 4773	2000	Herbarium	Um Neqa Ajayed farm
Aegilops triuncialis	EDNA15-0042801	MTA009-16	KX282510	KX281957	KTUH430	M Dib & M Al-Dosari	2001	Herbarium	Al-Khiran close to sea shore
Aegilops triuncialis	EDNA15-0042981	MTA008-16	KX282511	KX281958	KTUH431	M Al-Dosari 6147	2007	Herbarium	Failaka Island
Aegilops triuncialis	EDNA15-0042989	MTA010-16	KX282512	KX281959	KTUH432	M Al-Dosari 6000	2006	Herbarium	Al-Subbiyah Power station
Aeluropus lagopoides	EDNA15-0042682	MTA013-16			KTUH433	M Al-Dosari 3455	1998	Herbarium	Failaka Island
Aeluropus lagopoides	EDNA15-0042701	MTA011-16	KX282513		KTUH434	KT Mathew 3988	1998	Herbarium	Doha - Sulaibikhat road
Aeluropus lagopoides	EDNA15-0042662	MTA012-16	KX282514		MTA138	M Abdullah MTA138	2012	Fresh	Nuwaiseeb
Agathophora alopecuroide	EDNA15-0042399	MTA014-16			KTUH047	R Halwagy 81-78	1981	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Aizoon canariense	EDNA15-0042672	MTA016-16		KX281960	KTUH001	KT Mathew 5365	2004	Herbarium	Failaka Island Archeological ruins
Aizoon canariense	EDNA15-0042558	MTA015-16	KX282515	KX281961	MTA431	M Abdullah MTA431	2013	Fresh	Failaka Island
Aizoon canariense	EDNA15-0042604	MTA017-16			MTA547	M Abdullah MTA547	2013	Fresh	Subiyah
Aizoon hispanicum	EDNA15-0042627	MTA019-16	KX282516	KX281962	MTA302	M Abdullah MTA302	2013	Fresh	Al-Liyah
Aizoon hispanicum	EDNA15-0042559	MTA018-16	KX282517	KX281963	MTA387	M Abdullah MTA387	2013	Fresh	Failaka Island
Aizoon hispanicum	EDNA15-0042789	MTA020-16	KX282518	KX281964	MTA468	M Abdullah MTA468	2013	Fresh	Al-Salmi
Alhagi graecorum	EDNA15-0042593	MTA022-16	KX282519	KX281965	KTUH176	M Al-Dosari 4484	1999	Herbarium	Failaka Island
Alhagi graecorum	EDNA15-0042723	MTA021-16	KX282520	KX281966	KTUH177	KT Mathew 4093	1998	Herbarium	Omariyah Agricultural Research station
Alhagi graecorum	EDNA15-0042824	MTA023-16	KX282521	KX281967	KTUH178	L Boulos 17892	1992	Herbarium	Kuwait University Campus

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Allium sindjarense	EDNA15-0042724	MTA025-16			KTUH337	L Boulos 18098	1993	Herbarium	7 KM N of Ahmad Al-Jaber Air Base
Allium sindjarense	EDNA15-0042778	MTA024-16	KX282522		KTUH338	KT Mathew 5504	2007	Herbarium	Sabah Al-Ahmad Nature Reserve
Allium sindjarense	EDNA15-0042825	MTA026-16			KTUH339	G Al-Abbadi 339	1998	Herbarium	KISR - Sulaibiya Research Station
Allium sphaerocephalum	EDNA15-0042665	MTA027-16	KX282523	KX281968	MTA600	M Abdullah MTA600	2013	Fresh	Um-Neqa
Allium sphaerocephalum	EDNA15-0042685	MTA028-16	KX282524		MTA612	M Abdullah MTA612	2013	Fresh	Abdali
Althaea ludwigii	EDNA15-0042725	MTA030-16	KX282525		KTUH227	M Halwagy 1066	1972	Herbarium	Jal Az Zor
Althaea ludwigii	EDNA15-0042779	MTA029-16	KX282526	KX281969	KTUH228	M Al-Dosari 5585	2005	Herbarium	Al-Subiyah power station
Althaea ludwigii	EDNA15-0042578	MTA031-16	KX282527	KX281970	MTA272	M Abdullah MTA272	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Alyssum linifolium	EDNA15-0043058	MTA032-16		KX281971	KTUH382	R Halwagy 1307	1976	Herbarium	Al-Shaqayah police station
Amaranthus lividus	EDNA15-0042673	MTA033-16	KX282528		KTUH331	M Al-Dosari 6170	2007	Herbarium	Al-Wafra Al-Jouriah farm
Anabasis setifera	EDNA15-0043009	MTA035-16	KX282529	KX281972	KTUH048	R Halwagy 81-45	1981	Herbarium	Al Atraf 18 KM W of Jahra
Anabasis setifera	EDNA15-0042837	MTA034-16	KX282530	KX281973	MTA438	M Abdullah MTA438	2013	Fresh	Failaka Island
Anabasis setifera	EDNA15-0042993	MTA036-16			MTA588	M Abdullah MTA588	2013	Fresh	Um-Neqa
Anastatica hierochuntica	EDNA15-0042743	MTA038-16	KX282531	KX281974	KTUH383	M Al-Dosari 2198	1997	Herbarium	Al-Wafra farms
Anastatica hierochuntica	EDNA15-0043140	MTA037-16			KTUH384	M Al-Duleimi 1039	1990	Herbarium	Flora of Bahrain - Arabian Gulf University
Anastatica hierochuntica	EDNA15-0043141	MTA039-16	KX282532	KX281975	KTUH385	L Boulos 385	1990	Herbarium	Flora of Oman
Andrachne telephioides	EDNA15-0042924	MTA040-16	KX282533	KX281976	KTUH128	M Al-Dosari 5568	2004	Herbarium	Al-Zoor power station
Andrachne telephioides	EDNA15-0042569	MTA041-16	KX282534	KX281977	MTA400	M Abdullah MTA400	2013	Fresh	Failaka Island
Anisosciadium lanatum	EDNA15-0042602	MTA044-16	KX282535	KX281978	MTA134	M Abdullah MTA134	2012	Fresh	Nuwaiseeb
Anisosciadium lanatum	EDNA15-0042588	MTA042-16	KX282536	KX281979	MTA594	M Abdullah MTA594	2013	Fresh	Um-Neqa
Anisosciadium lanatum	EDNA15-0042650	MTA043-16	KX282537	KX281980	MTA610	M Abdullah MTA610	2013	Fresh	Al-Liyah
Anthemis deserti	EDNA15-0042409	MTA046-16	KX282538	KX281981	MTA203	M Abdullah MTA203	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Anthemis deserti	EDNA15-0042412	MTA045-16	KX282539	KX281982	MTA225	M Abdullah MTA225	2013	Fresh	Nuwaiseeb
Anthemis deserti	EDNA15-0043051	MTA047-16	KX282540	KX281983	MTA489	M Abdullah MTA489	2013	Fresh	Abdali
Anvillea garcinii	EDNA15-0043096	MTA048-16	KX282541	KX281984	KTUH061	M Al-Dosari 6372	2009	Herbarium	KISR - Sulaibiya Research Station

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Anvillea garcinii	EDNA15-0043078	MTA049-16	KX282542	KX281985	MTA459	M Abdullah MTA459	2013	Fresh	Al-Salmi
Arnebia decumbens	EDNA15-0042875	MTA051-16	KX282543	KX281986	MTA314	M Abdullah MTA314	2013	Fresh	KISR - Sulaibiya Research Station
Arnebia decumbens	EDNA15-0042991	MTA050-16	KX282544	KX281987	MTA390	M Abdullah MTA390	2013	Fresh	Failaka Island
Arnebia decumbens	EDNA15-0042984	MTA052-16	KX282545	KX281988	MTA447	M Abdullah MTA447	2013	Fresh	Al-Salmi
Arnebia linearifolia	EDNA15-0042814	MTA055-16			KTUH002	L Boulos LB18090	1993	Herbarium	Ahmed Al-Jabir Air Base - Sulaibiyah
Arnebia linearifolia	EDNA15-0042835	MTA053-16	KX282546		KTUH003	L Boulos LB18057	1993	Herbarium	Al-Bahra plateau - Jal Az Zor ridge
Arnebia linearifolia	EDNA15-0042919	MTA054-16	KX282547		KTUH003 b	L Boulos LB18057	1993	Herbarium	Al-Bahra plateau - Jal Az Zor ridge
Arnebia linearifolia	EDNA15-0042932	MTA056-16		KX281989	KTUH004	L Boulos LB18021	1993	Herbarium	Al-Mutla'a along the road of subbiyah
Arnebia tinctoria	EDNA15-0042853	MTA058-16			KTUH005	R Halwagy RH1175	1972	Herbarium	Abdali Sha'ab Abu Jarfan
Arnebia tinctoria	EDNA15-0042946	MTA057-16	KX282548	KX281990	KTUH006	L Boulos LB18147	1993	Herbarium	Al-salmi near Saudi Border
Artemisia herba-alba	EDNA15-0043084	MTA059-16	KX282549	KX281991	KTUH062	R Halwagy 81/77	1981	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Artemisia scoparia	EDNA15-0043016	MTA061-16	KX282550	KX281992	KTUH063	M Al-Dosari 5940	2005	Herbarium	Al-Khiran 8 KM from Nuwaiseeb border
Artemisia scoparia	EDNA15-0043022	MTA060-16	KX282551	KX281993	KTUH064	M Al-Dosari 4874	2000	Herbarium	Al-Khiran near the sea shore
Artemisia scoparia	EDNA15-0042794	MTA062-16	KX282552	KX281994	MTA136	M Abdullah MTA136	2012	Fresh	Nuwaiseeb
Asphodelus tenuifolius	EDNA15-0042974	MTA064-16	KX282553	KX281995	MTA328	M Abdullah MTA328	2013	Fresh	KISR - Sulaibiya Research Station
Asphodelus tenuifolius	EDNA15-0042594	MTA063-16	KX282554	KX281996	MTA427	M Abdullah MTA427	2013	Fresh	Failaka Island
Asphodelus tenuifolius	EDNA15-0042640	MTA065-16	KX282555	KX281997	MTA490	M Abdullah MTA490	2013	Fresh	Abdali
Asphodelus viscidulus	EDNA15-0042804	MTA067-16			KTUH340	M Al-Dosari 3019	1998	Herbarium	Mina Abdullah
Asphodelus viscidulus	EDNA15-0042885	MTA066-16			KTUH341	R Halwagy 1100	1972	Herbarium	Al-Shaqq
Asphodelus viscidulus	EDNA15-0042903	MTA068-16			KTUH342	M Al-Dosari 2795	1998	Herbarium	Sulaibiyah station KISR near water tank
Asteriscus hierochunticus	EDNA15-0042817	MTA070-16	KX282556	KX281998	KTUH065	KT Mathew 5443	2006	Herbarium	Subbiyah Power Station
Asteriscus hierochunticus	EDNA15-0043028	MTA069-16	KX282557	KX281999	KTUH066	M Al-Dosari 4476	1999	Herbarium	Failaka Island
Asteriscus hierochunticus	EDNA15-0043034	MTA071-16	KX282558	KX282000	KTUH067	DM Al-Awadi 120	2000	Herbarium	Um Al-Rimam
Astragalus annularis	EDNA13-0033612	MTA073-16	KX282559	KX282001	KTUH179	M AL-Dosari 5757	2005	Herbarium	Al-Wafra farms Al=Ameri farm
Astragalus annularis	EDNA13-0034196	MTA072-16	KX282560	KX282002	KTUH180	M Al-Dosari 2535	1997	Herbarium	Nuwaiseeb border station with Saudi Arabia

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Astragalus annularis	EDNA15-0042617	MTA074-16	KX282561	KX282003	KTUH181	KT Mathew 3047	1997	Herbarium	Subbiyah gulf shore
Astragalus bombycinus	EDNA13-0033613	MTA077-16	KX282562	KX282004	KTUH182	M Al-Dosari 5537	2004	Herbarium	Failaka Island
Astragalus bombycinus	EDNA15-0042639	MTA075-16	KX282563	KX282005	KTUH183	M Al-Dosari 1425	1995	Herbarium	Nuwaiseeb border station with Saudi Arabia
Astragalus bombycinus	EDNA13-0034197	MTA076-16	KX282564		MTA470	M Abdullah MTA470	2013	Fresh	Al-Salmi
Astragalus corrugatus	EDNA13-0033614	MTA078-16	KX282565	KX282006	KTUH184	M Al-Dosari 5034	2001	Herbarium	Al-Salmi border station near Saudi Arabia
Astragalus corrugatus	EDNA15-0042777	MTA079-16			KTUH185	M Al-Dosari 6133	2007	Herbarium	Subiyah Power Station
Astragalus corrugatus	EDNA13-0034198	MTA080-16	KX282566	KX282007	MTA205	M Abdullah MTA205	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Astragalus hamosus	EDNA13-0034204	MTA081-16	KX282567	KX282008	MTA280	M Abdullah MTA280	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Astragalus hamosus	EDNA13-0033620	MTA082-16	KX282568	KX282009	MTA281	M Abdullah MTA281	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Astragalus hauarensis	EDNA13-0033615	MTA085-16	KX282569	KX282010	KTUH186	R Halwagy 792	1971	Herbarium	Sulaibiyah power station
Astragalus hauarensis	EDNA13-0034199	MTA083-16	KX282570	KX282011	KTUH187	M Al-Dosari 1097	1990	Herbarium	Below Jal Az Zor along subiyah road
Astragalus hauarensis	EDNA15-0042866	MTA084-16			KTUH188	M Al-Dosari 1742	1996	Herbarium	15 KM from Al-Salmi
Astragalus schimperi	EDNA13-0033619	MTA087-16	KX282571		KTUH189	KT Mathew 3380	1998	Herbarium	Al-Abdaly 8.5 KM from border with Iraq
Astragalus schimperi	EDNA15-0043037	MTA086-16	KX282572	KX282012	MTA238	M Abdullah MTA238	2013	Fresh	KISR - Sulaibiya Research Station
Astragalus schimperi	EDNA13-0034203	MTA088-16	KX282573	KX282013	MTA376	M Abdullah MTA376	2013	Fresh	Doha Entertainment City
Astragalus sieberi	EDNA13-0033616	MTA090-16	KX282574	KX282014	KTUH190	L Boulos 18053	1993	Herbarium	Al-Bahra plateau - Jal Az Zor ridge
Astragalus sieberi	EDNA15-0042576	MTA089-16	KX282575	KX282015	MTA146	M Abdullah MTA146	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Astragalus sieberi	EDNA13-0034200	MTA091-16	KX282576	KX282016	MTA461	M Abdullah MTA461	2013	Fresh	Al-Salmi
Astragalus spinosus	EDNA13-0034202	MTA093-16	KX282577		MTA212	M Abdullah MTA212	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Astragalus spinosus	EDNA13-0033618	MTA092-16	KX282578	KX282017	MTA454	M Abdullah MTA454	2013	Fresh	Al-Salmi
Astragalus spinosus	EDNA15-0043019	MTA094-16	KX282579	KX282018	MTA531	M Abdullah MTA531	2013	Fresh	Abdali
Astragalus tribuloides	EDNA13-0033617	MTA096-16	KX282580	KX282019	KTUH192	M Al-Dosari 5653	2005	Herbarium	KISR - Sulaibiya Research Station
Astragalus tribuloides	EDNA15-0042844	MTA095-16	KX282581	KX282020	KTUH193	R Halwagy 1076	1972	Herbarium	Al-Dibdibah 15 KM N of Salmy
Astragalus tribuloides	EDNA13-0034201	MTA097-16	KX282582	KX282021	MTA345	M Abdullah MTA345	2013	Fresh	Doha Entertainment City
Atractylis cancellata	EDNA15-0042838	MTA099-16	KX282583	KX282022	KTUH068	Fatima F120	1998	Herbarium	Al-Subbiyah

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Atractylis cancellata	EDNA15-0043040	MTA098-16	KX282584	KX282023	KTUH069	M Al-Sdosari 4807	2000	Herbarium	Um Al-Rimam close to water catchment
Atractylis cancellata	EDNA15-0043046	MTA100-16	KX282585	KX282024	KTUH070	M Al-Dosari 4730	2000	Herbarium	Um Neqa near Abdali border
Atractylis carduus	EDNA15-0043063	MTA102-16	KX282586	KX282025	MTA012	M Abdullah MTA012	2012	Fresh	King Fahad High Way
Atractylis carduus	EDNA15-0043111	MTA101-16	KX282587		MTA373	M Abdullah MTA373	2013	Fresh	Doha Entertainment City
Atractylis carduus	EDNA15-0043099	MTA103-16	KX282588	KX282026	MTA533	M Abdullah MTA533	2013	Fresh	Subiyah
Atriplex dimorphostegia	EDNA15-0042856	MTA105-16	KX282589	KX282027	KTUH049	KT Mathew 3887	1998	Herbarium	Subiyah waterfront near power station enterance
Atriplex dimorphostegia	EDNA15-0043015	MTA104-16		KX282028	KTUH050	KT Mathew 3972	1998	Herbarium	Doha along the roadside
Atriplex dimorphostegia	EDNA15-0043021	MTA106-16			KTUH051	KT Mathew 4920	2000	Herbarium	Abdali Sheikh Ali Ajaib farm
Atriplex leucoclada	EDNA15-0042676	MTA108-16	KX282590		KTUH350	M Al-Dosari 3459	1998	Herbarium	Failaka Island
Atriplex leucoclada	EDNA15-0042695	MTA107-16	KX282591		KTUH352	M Al-Dosari 3537	1998	Herbarium	Subbiyah near power station
Atriplex leucoclada	EDNA15-0042655	MTA109-16	KX282592		MTA164	M Abdullah MTA164	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Avena barbata	EDNA15-0043129	MTA111-16			KTUH435	M Al-Dosari 2924	1998	Herbarium	Mina Abdullah along King Fahad road
Avena barbata	EDNA15-0043138	MTA110-16	KX282593	KX282029	KTUH436	M Bajwa 821-75	1975	Herbarium	Flora of KSA - khafji
Avena fatua	EDNA15-0043110	MTA112-16	KX282594	KX282030	KTUH437	M Al-Dosari 6104	2007	Herbarium	Subbiyah Power Station
Avena sterilis	EDNA15-0043117	MTA113-16			KTUH438	R Halwagy 1260	1974	Herbarium	Al-Dba'iyyah
Avena sterilis	EDNA15-0043131	MTA114-16		KX282031	KTUH439	M Al-Dosari 1709	1996	Herbarium	Jahra - Al-salmi road
Avicennia marina	EDNA15-0042671	MTA115-16	KX282595	KX282032	KTUH329	M Al-Dosari 6077	2006	Herbarium	KISR - Sulaibiya Research Station
Avicennia marina	EDNA15-0042653	MTA116-16	KX282596	KX282033	MTA195	M Abdullah MTA195	2012	Fresh	Sulaibiya - Coastal area
Bassia eriophora	EDNA15-0043146	MTA118-16	KX282597	KX282034	KTUH353	M Al-Dosari 3314	1998	Herbarium	Subbiyah near the sea shore
Bassia eriophora	EDNA15-0043144	MTA117-16	KX282598	KX282035	MTA350	M Abdullah MTA350	2013	Fresh	Sulaibiya - Coastal area
Bassia eriophora	EDNA15-0043145	MTA119-16	KX282599	KX282036	MTA478	M Abdullah MTA478	2013	Fresh	Abdali
Bellevalia saviczii	EDNA15-0042867	MTA120-16	KX282600		KTUH343	KT Mathew 3275	1998	Herbarium	Failaka Island salt depressions
Bellevalia saviczii	EDNA15-0042983	MTA121-16	KX282601	KX282037	KTUH344	KT Mathew 5451	2006	Herbarium	Subbiyah Power Station
Bienertia cycloptera	EDNA15-0042714	MTA124-16		KX282038	KTUH357	M Al-Dosari 1557	1995	Herbarium	Subbiyah near the sea shore
Bienertia cycloptera	EDNA15-0042765	MTA122-16	KX282602		KTUH358	KT Mathew 2947	1996	Herbarium	Al-Nuwaiseeb

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Bienertia cycloptera	EDNA15-0042896	MTA123-16	KX282603	KX282039	KTUH359	M Al-Dosari 3471a	1998	Herbarium	Failaka Island
Brachypodium distachyum	EDNA15-0042996	MTA126-16	KX282604	KX282040	KTUH440	M Al-Dosari 5999	2006	Herbarium	Al Subiyah Power Station
Brachypodium distachyum	EDNA15-0043002	MTA125-16	KX282605	KX282041	KTUH441	KT Mathew 4969	2001	Herbarium	Nuwaiseeb border station with Saudi Arabia
Brachypodium distachyum	EDNA15-0042615	MTA127-16	KX282606	KX282042	MTA416	M Abdullah MTA416	2013	Fresh	Failaka Island
Brassica tournefortii	EDNA15-0042678	MTA129-16	KX282607		KTUH386	M Al-Dosari 5028	2001	Herbarium	Al-Salmi border station near Saudi Arabia
Brassica tournefortii	EDNA15-0042697	MTA128-16	KX282608	KX282043	KTUH387	M Al-Dosari 3929	1999	Herbarium	Wadi Um Al-Rimam
Brassica tournefortii	EDNA15-0042657	MTA130-16	KX282609		MTA220	M Abdullah MTA220	2013	Fresh	Nuwaiseeb
Bromus madritensis	EDNA15-0043012	MTA132-16	KX282610	KX282044	KTUH444	KT Mathew 3783	1998	Herbarium	Al-Salmi border station near Saudi Arabia
Bromus madritensis	EDNA15-0042638	MTA131-16	KX282611	KX282045	MTA418	M Abdullah MTA418	2013	Fresh	Failaka Island
Bromus madritensis	EDNA15-0043006	MTA133-16	KX282612	KX282046	MTA501	M Abdullah MTA501	2013	Fresh	Abdali
Bromus tectorum	EDNA15-0043115	MTA134-16	KX282613	KX282047	KTUH442	M Al-Dosari 4660	2000	Herbarium	Wadi Um Al-Rimam
Bromus tectorum	EDNA15-0043130	MTA135-16	KX282614	KX282048	KTUH443	KT Mathew 3264	1997	Herbarium	Al-Salmi border station near Saudi Arabia
Bupleurum semicompositum	EDNA15-0042735	MTA137-16	KX282615	KX282049	KTUH308	M Al-Dosari 4471	1999	Herbarium	Failaka Island
Bupleurum semicompositum	EDNA15-0042812	MTA136-16	KX282616		KTUH309	M Al-Dosari 4740	2000	Herbarium	Um Neqa near Al-Abdali border
Bupleurum semicompositum	EDNA15-0042833	MTA138-16	KX282617	KX282050	KTUH310	KT Mathew 5331	2004	Herbarium	Al-Khiran
Cakile arabica	EDNA15-0042383	MTA140-16	KX282618		MTA218	M Abdullah MTA218	2013	Fresh	Nuwaiseeb
Cakile arabica	EDNA15-0042950	MTA139-16	KX282619	KX282051	MTA311	M Abdullah MTA311	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Cakile arabica	EDNA15-0042401	MTA141-16	KX282620	KX282052	MTA518	M Abdullah MTA518	2013	Fresh	Abdali
Calendula arvensis	EDNA15-0042400	MTA143-16	KX282621	KX282053	MTA201	M Abdullah MTA201	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Calendula arvensis	EDNA15-0042403	MTA142-16	KX282622	KX282054	MTA228	M Abdullah MTA228	2013	Fresh	Nuwaiseeb
Calendula arvensis	EDNA15-0042406	MTA144-16	KX282623	KX282055	MTA449	M Abdullah MTA449	2013	Fresh	Al-Salmi
Calendula tripterocarpa	EDNA15-0042857	MTA146-16	KX282624	KX282056	KTUH071	M Bajwa 170-75	1975	Herbarium	near Khafji 60 KM near Saudi Arabia
Calendula tripterocarpa	EDNA15-0042321	MTA145-16	KX282625	KX282057	KTUH072	R Halwagy 1136	1972	Herbarium	Wadi Um Al-Rimam close to water catchment
Calendula tripterocarpa	EDNA15-0042327	MTA147-16	KX282626		KTUH073	R Halwagy 82/5	1982	Herbarium	Jahra plantation
Calligonum comosum	EDNA15-0042705	MTA149-16	KX282627		KTUH258	KT Mathew 2920	1996	Herbarium	Jal Az-Zor

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Calligonum comosum	EDNA15-0042687	MTA148-16	KX282628		MTA039	M Abdullah MTA039	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Calligonum comosum	EDNA15-0042667	MTA150-16	KX282629		MTA153	M Abdullah MTA153	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Calotropis procera	EDNA15-0042709	MTA152-16	KX282630		KTUH332	M Al-Dosari 6380	2009	Herbarium	KISR - Sulaibiya Research Station
Calotropis procera	EDNA15-0042674	MTA151-16	KX282631		MTA144	M Abdullah MTA144	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Calotropis procera	EDNA15-0042692	MTA153-16	KX282632		MTA574	M Abdullah MTA574	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Carduus pycnocephalus	EDNA15-0042339	MTA155-16	KX282633	KX282058	KTUH074	M AL-Dosari 2882	1998	Herbarium	King Fahad High Way along the roadside
Carduus pycnocephalus	EDNA15-0042862	MTA156-16	KX282634	KX282059	MTA211	M Abdullah MTA211	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Carduus pycnocephalus	EDNA15-0042333	MTA154-16	KX282635	KX282060	MTA227	M Abdullah MTA227	2013	Fresh	Nuwaiseeb
Carrichtera annua	EDNA15-0042716	MTA158-16	KX282636		KTUH388	M Al-Dosari 2671	1998	Herbarium	Kuwait University Khaldiyah Campus
Carrichtera annua	EDNA15-0042959	MTA157-16	KX282637		KTUH389	KT Mathew 4414	1999	Herbarium	Um Al-Rimam
Carrichtera annua	EDNA15-0042970	MTA159-16	KX282638	KX282061	KTUH390	M Al-Dosari 1893	1997	Herbarium	Al-Salmi border station near Saudi Arabia
Carthamus oxyacantha	EDNA15-0042345	MTA160-16	KX282639	KX282062	KTUH075	Fatima F118	1998	Herbarium	Al-Khiran
Carthamus oxyacantha	EDNA15-0042351	MTA161-16	KX282640	KX282063	KTUH076	M Al-Dosari 5545	2004	Herbarium	Al-Abdali40 KM from border with Iraq
Carthamus oxyacantha	EDNA15-0042878	MTA162-16	KX282641	KX282064	MTA452	M Abdullah MTA452	2013	Fresh	Al-Salmi
Cenchrus ciliaris	EDNA15-0042864	MTA164-16		KX282065	KTUH445	L Boulos 17899	1992	Herbarium	Sabah Hospital Shuwailh waste ground
Cenchrus ciliaris	EDNA15-0043018	MTA163-16			KTUH446	M Al-Dosari 3741	1999	Herbarium	KISR - Sulaibiya Research Station
Centaurea bruguierana	EDNA15-0043112	MTA166-16	KX282642	KX282066	KTUH077	M Al-Dosari 5894	2005	Herbarium	Al-Salmi along jahra road
Centaurea bruguierana	EDNA15-0043071	MTA165-16	KX282643	KX282067	MTA081	M Abdullah MTA081	2012	Fresh	Sixth Ring Road
Centaurea bruguierana	EDNA15-0043107	MTA167-16	KX282644	KX282068	MTA529	M Abdullah MTA529	2013	Fresh	Abdali
Centaurea mesopotamica	EDNA15-0043080	MTA168-16	KX282645	KX282069	KTUH078	R Halwagy 14/83	1983	Herbarium	Al-Khiran coastal area near Saudi border
Centaurea mesopotamica	EDNA15-0043094	MTA169-16	KX282646	KX282070	KTUH079	KT Mathew 2674	1995	Herbarium	Nuwaiseeb near border with Saudi Arabia
Centaurea pseudosinaica	EDNA15-0043074	MTA171-16	KX282647	KX282071	MTA498	M Abdullah MTA498	2013	Fresh	Abdali
Centaurea pseudosinaica	EDNA15-0043097	MTA170-16			MTA608	M Abdullah MTA608	2013	Fresh	Al-Liyah
Centaurea pseudosinaica	EDNA15-0043100	MTA172-16	KX282648	KX282072	MTA609	M Abdullah MTA609	2013	Fresh	Abdali
Centropodia forskalii	EDNA15-0042683	MTA174-16	KX282649		KTUH448	M Al-Dosari 1479	1995	Herbarium	Al-Jahra - Al-salmi road

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Centropodia forskalii	EDNA15-0042702	MTA173-16	KX282650		KTUH449	M Al-Dosari 1534	1995	Herbarium	Jahra - Al-Subbiyah road
Centropodia forskalii	EDNA15-0042663	MTA175-16	KX282651		MTA589	M Abdullah MTA589	2013	Fresh	Um-Neqa
Chenopodium murale	EDNA15-0042332	MTA177-16	KX282652	KX282073	MTA072	M Abdullah MTA072	2012	Fresh	Sixth Ring Road
Chenopodium murale	EDNA15-0042338	MTA176-16	KX282653	KX282074	MTA394	M Abdullah MTA394	2013	Fresh	Failaka Island
Chenopodium murale	EDNA15-0042344	MTA178-16	KX282654	KX282075	MTA555	M Abdullah MTA555	2013	Fresh	Subiyah
Chrozophora tinctoria	EDNA15-0042717	MTA180-16			KTUH129	M Al-Dosari 5555	2004	Herbarium	Al-Retqa police station near Iraqi border
Chrozophora tinctoria	EDNA15-0042772	MTA179-16	KX282655		KTUH130	R Halwagy 19/83	1983	Herbarium	Al-Khiran near Saudi border station
Chrozophora tinctoria	EDNA15-0042880	MTA181-16	KX282656	KX282076	KTUH131	M Al-Dosari 5548	2004	Herbarium	Al-Abdaly 40 KM from border with Iraq
Cistanche tubulosa	EDNA15-0042666	MTA183-16			MTA379	M Abdullah MTA379	2013	Fresh	Doha Entertainment City
Cistanche tubulosa	EDNA15-0042686	MTA182-16			MTA380	M Abdullah MTA380	2013	Fresh	KISR - Sulaibiya Research Station
Cistanche tubulosa	EDNA15-0042704	MTA184-16			MTA577	M Abdullah MTA577	2013	Fresh	Al-Salmi
Citrullus colocynthis	EDNA15-0042567	MTA186-16	KX282657		MTA047	M Abdullah MTA047	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Citrullus colocynthis	EDNA15-0042611	MTA185-16	KX282658	KX282077	MTA560	M Abdullah MTA560	2013	Fresh	Subiyah
Citrullus colocynthis	EDNA15-0042634	MTA187-16	KX282659	KX282078	MTA565	M Abdullah MTA565	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Convolvulus cephalopodus	EDNA15-0042795	MTA189-16	KX282660	KX282079	KTUH118	KT Mathew 5496	2007	Herbarium	Sabah Al-Ahmad Nature Reserve
Convolvulus cephalopodus	EDNA15-0042565	MTA188-16	KX282661	KX282080	MTA113	M Abdullah MTA113	2012	Fresh	Nuwaiseeb
Convolvulus cephalopodus	EDNA15-0042632	MTA190-16	KX282662	KX282081	MTA535	M Abdullah MTA535	2013	Fresh	Subiyah
Convolvulus oxyphyllus	EDNA15-0042609	MTA192-16	KX282663	KX282082	MTA592	M Abdullah MTA592	2013	Fresh	Um-Neqa
Convolvulus oxyphyllus	EDNA15-0043139	MTA191-16	KX282664	KX282083	MTA604	M Abdullah MTA604	2013	Fresh	Um-Neqa
Convolvulus oxyphyllus	EDNA15-0042557	MTA193-16	KX282665	KX282084	MTA606	M Abdullah MTA606	2013	Fresh	Al-Liyah
Convolvulus pilosellifolius	EDNA15-0042592	MTA195-16	KX282666	KX282085	MTA572	M Abdullah MTA572	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Convolvulus pilosellifolius	EDNA15-0042818	MTA194-16	KX282667	KX282086	MTA602	M Abdullah MTA602	2013	Fresh	Um-Neqa
Convolvulus pilosellifolius	EDNA15-0042839	MTA196-16	KX282668	KX282087	MTA605	M Abdullah MTA605	2013	Fresh	Al-Liyah
Cornulaca aucheri	EDNA15-0042893	MTA197-16	KX282669	KX282088	MTA006	M Abdullah MTA006	2012	Fresh	King Fahad High Way
Cornulaca aucheri	EDNA15-0043033	MTA198-16	KX282670	KX282089	MTA131	M Abdullah MTA131	2012	Fresh	Nuwaiseeb -near Saudi Arabia border

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Cornulaca aucheri	EDNA15-0042326	MTA199-16	KX282671	KX282090	MTA517	M Abdullah MTA517	2013	Fresh	Abdali
Cornulaca monacantha	EDNA15-0042696	MTA200-16	KX282672	KX282091	KTUH365	M Al-Dosari 1602	1996	Herbarium	Along king Fahad Highway towards Ahmadi
Cornulaca monacantha	EDNA15-0042677	MTA201-16	KX282673		MTA367	M Abdullah MTA367	2013	Fresh	Doha Entertainment City
Cornulaca monacantha	EDNA15-0042656	MTA202-16	KX282674		MTA550	M Abdullah MTA550	2013	Fresh	Subiyah
Coronopus didymus	EDNA15-0042770	MTA205-16	KX282675	KX282092	KTUH391	M Al-Dosari 2067	1997	Herbarium	Kuwait University Khaldiyah Campus
Coronopus didymus	EDNA15-0042980	MTA203-16	KX282676	KX282093	KTUH392	I Ibrahim 1150	1990	Herbarium	Al-salmi near Saudi Border
Coronopus didymus	EDNA15-0042988	MTA204-16		KX282094	KTUH393	KT Mathew 5329	2004	Herbarium	Al-khiran plateau
Cressa cretica	EDNA15-0042715	MTA207-16			KTUH119	M Al-Dosari 3667	1998	Herbarium	Al-Wafra 43 KM from Nuwaiseeb Fire Station
Cressa cretica	EDNA15-0042742	MTA206-16			KTUH120	R Halwagy W-116	1971	Herbarium	Along Al-Istiqlal road sabkha area
Cressa cretica	EDNA15-0042769	MTA208-16	KX282677		KTUH121	M Al-Dosari 4786	2000	Herbarium	Failaka Island
Crucianella membranacea	EDNA15-0042647	MTA210-16	KX282678		KTUH283	M Al-Dosari 4437	1999	Herbarium	Nuwaiseeb border station with Saudi Arabia
Crucianella membranacea	EDNA15-0042810	MTA209-16			KTUH284	G Al-Abbadi 284	2001	Herbarium	KISR - Sulaibiya Research Station
Crucianella membranacea	EDNA15-0042848	MTA211-16	KX282679		KTUH285	M Al-Dosari 6076	2006	Herbarium	Al-Salmi
Cuscuta planiflora	EDNA15-0042960	MTA213-16		KX282095	KTUH521	M Al-Dosari 5212	2001	Herbarium	KISR - Sulaibiya Research Station
Cuscuta planiflora	EDNA15-0042971	MTA212-16		KX282096	KTUH522	R Halwagy 82/41	1982	Herbarium	AL-Khiran near border with Saudi Arabia
Cuscuta planiflora	EDNA15-0043143	MTA214-16	KX282680	KX282097	KTUH523	R Halwagy 82/58	1982	Herbarium	Ras Az Zor
Cutandia dichotoma	EDNA15-0043024	MTA216-16	KX282681	KX282098	KTUH450	KT Mathew 2766	1996	Herbarium	Al-Nuwaiseeb border station
Cutandia dichotoma	EDNA15-0043036	MTA215-16	KX282682	KX282099	KTUH451	G Al-Abbadi 451	1998	Herbarium	KISR - Sulaibiya Research Station
Cutandia dichotoma	EDNA15-0042775	MTA217-16	KX282683	KX282100	MTA412	M Abdullah MTA412	2013	Fresh	Failaka Island
Cutandia memphitica	EDNA15-0043042	MTA219-16	KX282684	KX282101	KTUH452	KT Mathew 2747	1996	Herbarium	Along King Fahad Motorway towards Ahmadi
Cutandia memphitica	EDNA15-0043048	MTA218-16	KX282685	KX282102	KTUH453	M Al-Dosari 1861	1997	Herbarium	Al-Abdaly
Cutandia memphitica	EDNA15-0042882	MTA220-16	KX282686	KX282103	MTA335	M Abdullah MTA335	2013	Fresh	KISR - Sulaibiya Research Station
Cynodon dactylon	EDNA15-0042913	MTA222-16	KX282687	KX282104	KTUH454	KT Mathew 5064	2001	Herbarium	Equestrian Club premises
Cynodon dactylon	EDNA15-0042323	MTA221-16	KX282688	KX282105	KTUH455	KT Mathew 4895	2000	Herbarium	Abdali
Cynodon dactylon	EDNA15-0042335	MTA223-16	KX282689		KTUH456	L Boulos 17856	1992	Herbarium	Failaka Island

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Cynomorium coccineum	EDNA15-0042679	MTA225-16			KTUH519	M Al-Dosari 4689	2000	Herbarium	Khadma salt marshes
Cynomorium coccineum	EDNA15-0042698	MTA224-16			KTUH520	KT Mathew 4695	2000	Herbarium	Jal- Az Zor SSNR
Cynomorium coccineum	EDNA15-0042658	MTA226-16			MTA581	M Abdullah MTA581	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Cyperus conglomeratus	EDNA15-0042680	MTA228-16	KX282690		MTA052	M Abdullah MTA052	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Cyperus conglomeratus	EDNA15-0042712	MTA227-16	KX282691		MTA091	M Abdullah MTA091	2012	Fresh	Sixth Ring Road
Cyperus conglomeratus	EDNA15-0042699	MTA229-16	KX282692		MTA430	M Abdullah MTA430	2013	Fresh	Failaka Island
Deverra triradiata	EDNA15-0042787	MTA230-16			KTUH311	R Halwagy 006/85	1985	Herbarium	Al-Salmi near border with Saudi Arabia
Deverra triradiata	EDNA15-0042873	MTA231-16	KX282693	KX282106	KTUH312	M Al-Dosari 5465	2003	Herbarium	Al-Dubaiah Resort
Dichanthium annulatum	EDNA15-0043105	MTA232-16	KX282694	KX282107	KTUH460	M Al-Dosari 3625	1998	Herbarium	Kuwait University Khaldiyah Campus
Digitaria ciliaris	EDNA15-0043104	MTA233-16	KX282695	KX282108	KTUH462	KT Mathew 4177	1998	Herbarium	Kuwait University Khaldiyah Campus
Dipcadi erythraeum	EDNA15-0042845	MTA234-16	KX282696	KX282109	KTUH347	KT Mathew 5323	2004	Herbarium	KISR - Sulaibiya Research Station
Dipcadi erythraeum	EDNA15-0042943	MTA236-16	KX282697		KTUH348	M Al-Dosari 5699	2005	Herbarium	Sabah Al-Ahmad Nature Reserve
Dipcadi erythraeum	EDNA15-0042964	MTA235-16	KX282698	KX282110	KTUH349	L Boulos 17958	1993	Herbarium	8 KM S of Al-Wafra
Diplotaxis acris	EDNA15-0043064	MTA238-16	KX282699	KX282111	KTUH395	R Halwagy 395	1974	Herbarium	Wadi Al-Batin
Diplotaxis acris	EDNA15-0042937	MTA237-16	KX282700	KX282112	MTA460	M Abdullah MTA460	2013	Fresh	Al-Salmi
Diplotaxis harra	EDNA15-0043005	MTA239-16	KX282701	KX282113	KTUH394	M Al-Dosari 6016	2006	Herbarium	Failaka Island
Diplotaxis harra	EDNA15-0042610	MTA240-16	KX282702	KX282114	MTA198	M Abdullah MTA198	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Diplotaxis harra	EDNA15-0042995	MTA241-16	KX282703	KX282115	MTA444	M Abdullah MTA444	2013	Fresh	Al-Salmi
Ducrosia anethifolia	EDNA15-0042851	MTA243-16	KX282704	KX282116	KTUH313	M Al-Dosari 2083a	1997	Herbarium	Al-Khiran gulf shore
Ducrosia anethifolia	EDNA15-0042889	MTA242-16	KX282705	KX282117	KTUH314	M Dib 314	1998	Herbarium	Salmiyah Abu Halifa
Echinops blancheanus	EDNA15-0042362	MTA245-16	KX282706	KX282118	KTUH087	M Al-Dosari 5886	2005	Herbarium	Al-Salmi along the roadside
Echinops blancheanus	EDNA15-0042897	MTA244-16	KX282707	KX282119	MTA171	M Abdullah MTA171	2012	Fresh	KISR - Sulaibiya Research Station
Echinops blancheanus	EDNA15-0042357	MTA246-16	KX282708	KX282120	MTA172	M Abdullah MTA172	2012	Fresh	KISR - Sulaibiya Research Station
Echium rauwolfii	EDNA15-0042561	MTA248-16	KX282709		KTUH008	KT Mathew 3006	1997	Herbarium	10 KM from Al-salmi Kuwait City road
Echium rauwolfii	EDNA15-0042710	MTA247-16		KX282121	KTUH009	KT Mathew 5294	2002	Herbarium	Jahra Al-Salmi roadside

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Echium rauwolfii	EDNA15-0042659	MTA249-16	KX282710		KTUH010	KT Mathew 2839	1996	Herbarium	Al-Salmi Kuwait City road
Emex spinosa	EDNA15-0042581	MTA252-16	KX282711	KX282122	MTA395	M Abdullah MTA395	2013	Fresh	Failaka Island
Emex spinosus	EDNA15-0042807	MTA250-16	KX282712	KX282123	MTA232	M Abdullah MTA232	2013	Fresh	Nuwaiseeb
Emex spinosus	EDNA15-0042730	MTA251-16	KX282713		MTA249	M Abdullah MTA249	2013	Fresh	KISR - Sulaibiya Research Station
Ephedra alata	EDNA15-0042660	MTA253-16			MTA580	M Abdullah MTA580	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Eragrostis barrelieri	EDNA15-0043030	MTA255-16	KX282714	KX282124	KTUH468	M Al-Dosari 2473	1997	Herbarium	Along king Fahad Highway near road no. 238
Eragrostis barrelieri	EDNA15-0042329	MTA254-16	KX282715	KX282125	KTUH469	M Al-Dosari 2576	1997	Herbarium	Nuwaiseeb near border with Saudi Arabia
Eragrostis barrelieri	EDNA15-0042341	MTA256-16		KX282126	KTUH470	M Al-Dosari 2334	1997	Herbarium	Al-Dbaiyyah near road 238
Eremopoa persica	EDNA15-0043103	MTA257-16			KTUH472	A Al-Yahya 166	1988	Herbarium	Rumaithiyah garden weed
Eremopyrum bonaepartis	EDNA15-0043095	MTA258-16			KTUH473	R Halwagy 74-473	1974	Herbarium	Al-Mutla'a
Erodium bryoniifolium	EDNA15-0042719	MTA260-16			KTUH147	R Halwagy 1182	1972	Herbarium	Khabrat Al-Awazem 40 KM W of Kuwait City
Erodium bryoniifolium	EDNA15-0042774	MTA259-16			KTUH148	KT Mathew 3679	1998	Herbarium	Jal Az Zor on the plateau
Erodium bryoniifolium	EDNA15-0042800	MTA261-16	KX282716	KX282127	KTUH149	R Halwagy 4/83	1983	Herbarium	Wadi Um-Al-Rimam
Erodium ciconium	EDNA15-0042746	MTA263-16	KX282717	KX282128	KTUH150	R Halwagy 1250	1974	Herbarium	Al-Dba'iyyah 55 KM SE Kuwait City
Erodium ciconium	EDNA15-0042881	MTA262-16	KX282718		KTUH151	KT Mathew 5330	2004	Herbarium	Khiran plateau
Erodium ciconium	EDNA15-0042900	MTA264-16	KX282719	KX282129	KTUH152	Linda Shuaib	1995	Herbarium	Um Al-Rimam
Erodium cicutarium	EDNA15-0042822	MTA266-16	KX282720	KX282130	KTUH153	M Al-Dosari 4106	1999	Herbarium	Al-Ahmadi waste ground near research station
Erodium cicutarium	EDNA15-0042912	MTA265-16	KX282721		KTUH154	R Halwagy 82/6	1982	Herbarium	Jahra plantation
Erodium cicutarium	EDNA15-0042925	MTA267-16	KX282722	KX282131	KTUH155	KT Mathew 4994	2001	Herbarium	Subbiyah water front near power station enterance
Erodium glaucophyllum	EDNA15-0042681	MTA269-16	KX282723		KTUH156	M Al-Dosari 2605	1998	Herbarium	Wadi Um-Al-Rimam
Erodium glaucophyllum	EDNA15-0042700	MTA268-16	KX282724	KX282132	KTUH157	M Al-Dosari 2353	1997	Herbarium	Al-Khiran near sea shore
Erodium glaucophyllum	EDNA15-0042661	MTA270-16			MTA472	M Abdullah MTA472	2013	Fresh	Al-Salmi
Erodium laciniatum	EDNA15-0042573	MTA272-16	KX282725	KX282133	MTA098	M Abdullah MTA098	2012	Fresh	Nuwaiseeb -near Saudi Arabia border
Erodium laciniatum	EDNA15-0042637	MTA271-16	KX282726	KX282134	MTA252	M Abdullah MTA252	2013	Fresh	Mina Abdullah
Erodium laciniatum	EDNA15-0042961	MTA273-16	KX282727	KX282135	MTA541	M Abdullah MTA541	2013	Fresh	Subiyah

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Eruca sativa	EDNA15-0043142	MTA277-16	KX282728	KX282136	MTA221	M Abdullah MTA221	2013	Fresh	Nuwaiseeb
Eruca sativa	EDNA15-0042413	MTA274-16	KX282729	KX282137	MTA342	M Abdullah MTA342	2013	Fresh	KISR - Sulaibiya Research Station
Eruca sativa	EDNA15-0042404	MTA275-16	KX282730	KX282138	MTA396	M Abdullah MTA396	2013	Fresh	Failaka Island
Eruca sativa	EDNA15-0042410	MTA276-16	KX282731	KX282139	MTA524	M Abdullah MTA524	2013	Fresh	Abdali
Euphorbia densa	EDNA15-0042798	MTA279-16	KX282732	KX282140	KTUH132	L Boulos 18003	1993	Herbarium	Al-Mutla'a along the road of subbiyah
Euphorbia densa	EDNA15-0042938	MTA278-16	KX282733	KX282141	KTUH133	M Al-Dosari 3212	1998	Herbarium	Subbiyah 40 KM from Kuwait City
Euphorbia granulata	EDNA15-0042821	MTA280-16			KTUH517	R Halwagy 81/42	1981	Herbarium	Al-Mutla'a along basra road
Euphorbia granulata	EDNA15-0042951	MTA281-16		KX282142	KTUH518	KT Mathew 2673	1995	Herbarium	Al-Nuwaiseeb border station
Euphorbia serpens	EDNA15-0042744	MTA283-16	KX282734	KX282143	KTUH143	M Al-Dosari 4871	2000	Herbarium	Al-Khiran near the sea shore
Euphorbia serpens	EDNA15-0042899	MTA282-16	KX282735		KTUH144	M Al-Dosari 2706	1998	Herbarium	Gulf road near KISR water front project
Euphorbia serpens	EDNA15-0042911	MTA284-16	KX282736		KTUH145	M Al-Dosari 1878	1997	Herbarium	near Al-Rawdatain water front
Fagonia bruguieri	EDNA15-0042598	MTA286-16	KX282737	KX282144	MTA090	M Abdullah MTA090	2012	Fresh	Sixth Ring Road
Fagonia bruguieri	EDNA15-0042574	MTA285-16	KX282738	KX282145	MTA240	M Abdullah MTA240	2013	Fresh	Al-Liyah
Fagonia bruguieri	EDNA15-0042603	MTA287-16	KX282739	KX282146	MTA479	M Abdullah MTA479	2013	Fresh	Abdali
Fagonia glutinosa	EDNA15-0042652	MTA289-16	KX282740	KX282147	MTA260	M Abdullah MTA260	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Fagonia glutinosa	EDNA15-0042587	MTA288-16	KX282741	KX282148	MTA297	M Abdullah MTA297	2013	Fresh	Al-Liyah
Fagonia glutinosa	EDNA15-0042975	MTA290-16	KX282742	KX282149	MTA339	M Abdullah MTA339	2013	Fresh	KISR - Sulaibiya Research Station
Fagonia indica	EDNA15-0042737	MTA292-16	KX282743	KX282150	KTUH324	Maha M0061	2000	Herbarium	Al-Abdali
Fagonia indica	EDNA15-0042788	MTA291-16	KX282744	KX282151	KTUH325	R Halwagy 109-76	1976	Herbarium	Al-khiran
Fagonia indica	EDNA15-0042589	MTA293-16	KX282745	KX282152	MTA432	M Abdullah MTA432	2013	Fresh	Failaka Island
Farsetia aegyptia	EDNA15-0042393	MTA295-16	KX282746	KX282153	KTUH396	R Halwagy 82/17	1982	Herbarium	Wadi Um Al-Rimam
Farsetia aegyptia	EDNA15-0042397	MTA294-16	KX282747	KX282154	KTUH397	R Halwagy 81/96	1981	Herbarium	Wadi Al-Batin
Farsetia aegyptia	EDNA15-0042923	MTA296-16		KX282155	MTA331	M Abdullah MTA331	2013	Fresh	KISR - Sulaibiya Research Station
Filago pyramidata	EDNA15-0043106	MTA298-16	KX282748	KX282156	MTA271	M Abdullah MTA271	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Filago pyramidata	EDNA15-0043060	MTA297-16	KX282749	KX282157	MTA403	M Abdullah MTA403	2013	Fresh	Failaka Island

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Filago pyramidata	EDNA15-0043092	MTA299-16	KX282750	KX282158	MTA511	M Abdullah MTA511	2013	Fresh	Abdali
Frankenia pulverulenta	EDNA15-0042570	MTA302-16	KX282751	KX282159	MTA115	M Abdullah MTA115	2012	Fresh	Nuwaiseeb
Frankenia pulverulenta	EDNA15-0042636	MTA300-16	KX282752	KX282160	MTA406	M Abdullah MTA406	2013	Fresh	Failaka Island
Frankenia pulverulenta	EDNA15-0042613	MTA301-16	KX282753	KX282161	MTA477	M Abdullah MTA477	2013	Fresh	Abdali
Fumaria parviflora	EDNA15-0042773	MTA303-16	KX282754	KX282162	KTUH146	M Al-Dosari 2726	1998	Herbarium	Failaka Island
Gagea reticulata	EDNA15-0042916	MTA306-16	KX282755	KX282163	KTUH345	M Al-Dosari 4634	2000	Herbarium	KISR - Sulaibiya Research Station
Gagea reticulata	EDNA15-0042929	MTA304-16	KX282756	KX282164	KTUH346	L Boulos 17952	1993	Herbarium	5 KM E of Abraq
Gagea reticulata	EDNA15-0042577	MTA305-16		KX282165	MTA215	M Abdullah MTA215	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Galium tricornutum	EDNA15-0042784	MTA307-16	KX282757	KX282166	KTUH286	R Halwagy 81/59	1981	Herbarium	Al-Atraf 18 KM W of Jahra
Gastrocotyle hispida	EDNA15-0042976	MTA310-16		KX282167	MTA362	M Abdullah MTA362	2013	Fresh	Doha Entertainment City
Gastrocotyle hispida	EDNA15-0042966	MTA308-16	KX282758	KX282168	MTA402	M Abdullah MTA402	2013	Fresh	Failaka Island
Gastrocotyle hispida	EDNA15-0042854	MTA309-16	KX282759	KX282169	MTA495	M Abdullah MTA495	2013	Fresh	Abdali
Glaucium corniculatum	EDNA15-0042727	MTA311-16	KX282760		KTUH524	L Shuaib 524	1993	Herbarium	Al-Salmi near border with Saudi Arabia
Gymnarrhena micrantha	EDNA15-0042372	MTA313-16	KX282761	KX282170	KTUH088	KT Mathew 3892	1998	Herbarium	Subbiyah near power station
Gymnarrhena micrantha	EDNA15-0042909	MTA312-16	KX282762	KX282171	MTA369	M Abdullah MTA369	2013	Fresh	Doha Entertainment City
Gymnarrhena micrantha	EDNA15-0042367	MTA314-16	KX282763	KX282172	MTA523	M Abdullah MTA523	2013	Fresh	Abdali
Gynandriris sisyrinchium	EDNA15-0042703	MTA317-16	KX282764		MTA320	M Abdullah MTA320	2013	Fresh	Nuwaiseeb
Gynandriris sisyrinchium	EDNA15-0042664	MTA315-16	KX282765	KX282173	MTA321	M Abdullah MTA321	2013	Fresh	Mina Abdullah
Gynandriris sisyrinchium	EDNA15-0042684	MTA316-16	KX282766		MTA404	M Abdullah MTA404	2013	Fresh	Failaka Island
Gypsophila capillaris	EDNA15-0042855	MTA318-16	KX282767		MTA044	M Abdullah MTA044	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Gypsophila capillaris	EDNA15-0043044	MTA320-16	KX282768	KX282174	MTA068	M Abdullah MTA068	2012	Fresh	Sixth Ring Road
Gypsophila capillaris	EDNA15-0043050	MTA319-16	KX282769	KX282175	MTA480	M Abdullah MTA480	2013	Fresh	Abdali
Halocnemum strobilaceum	EDNA15-0042376	MTA323-16		KX282176	KTUH366	M Al-Dosari 1421	1995	Herbarium	Between Ahmadi and Mina Abdullah salt marshes
Halocnemum strobilaceum	EDNA15-0042361	MTA322-16	KX282770	KX282177	MTA193	M Abdullah MTA193	2012	Fresh	Sulaibiya - Coastal area
Halocnemum strobilaceum	EDNA15-0042371	MTA321-16	KX282771	KX282178	MTA582	M Abdullah MTA582	2013	Fresh	Um-Neqa

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Halothamnus iraqensis	EDNA15-0042740	MTA324-16	KX282772		KTUH367	KT Mathew 4579	1999	Herbarium	Dubaiah resort enterance
Halothamnus iraqensis	EDNA15-0042935	MTA326-16			KTUH368	KT Mathew 3978	1998	Herbarium	Doha 1 KM from waterfront
Halothamnus iraqensis	EDNA15-0042948	MTA325-16		KX282179	KTUH369	R Halwagy 757	1971	Herbarium	Along Mina Abdullah to Wafra road
Haloxylon salicornicum	EDNA15-0043039	MTA327-16	KX282773	KX282180	MTA186	M Abdullah MTA186	2012	Fresh	KISR - Sulaibiya Research Station
Haloxylon salicornicum	EDNA15-0042877	MTA329-16	KX282774	KX282181	MTA473	M Abdullah MTA473	2013	Fresh	Al-Salmi
Haloxylon salicornicum	EDNA15-0042908	MTA328-16	KX282775	KX282182	MTA591	M Abdullah MTA591	2013	Fresh	Um-Neqa
Haplophyllum tuberculatum	EDNA15-0042601	MTA330-16	KX282776		MTA028	M AbdullahMTA028	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Haplophyllum tuberculatum	EDNA15-0042625	MTA332-16	KX282777		MTA128	M Abdullah MTA128	2012	Fresh	Nuwaiseeb
Haplophyllum tuberculatum	EDNA15-0042584	MTA331-16	KX282778	KX282183	MTA456	M Abdullah MTA456	2013	Fresh	Al-Salmi
Helianthemum kahiricum	EDNA13-0033636	MTA333-16	KX282779	KX282184	KTUH052	IK Ibrahim 1045	1990	Herbarium	Um Al-Rimam protected area
Helianthemum kahiricum	EDNA13-0034219	MTA334-16	KX282780	KX282185	KTUH053	KT Mathew 4754	2000	Herbarium	Um Al-Rimam into the flat wadi
Helianthemum ledifolium	EDNA13-0033633	MTA336-16	KX282781	KX282186	KTUH054	M Al-Dosari 4074	1999	Herbarium	Al-Salmi border station near Saudi Arabia
Helianthemum ledifolium	EDNA13-0034216	MTA335-16		KX282187	KTUH055	M Al-Dosari 6438	2009	Herbarium	Al-Abdali Wleed Al-Omery farm
Helianthemum ledifolium	EDNA15-0042767	MTA337-16	KX282782	KX282188	KTUH056	R Halwagy 14-76	1976	Herbarium	Wadi Um Al-Rimam
Helianthemum lippii	EDNA15-0042591	MTA340-16	KX282783	KX282189	MTA176	M Abdullah MTA176	2012	Fresh	KISR - Sulaibiya Research Station
Helianthemum lippii	EDNA13-0034218	MTA339-16	KX282784		MTA371	M Abdullah MTA371	2013	Fresh	Doha Entertainment City
Helianthemum lippii	EDNA13-0033635	MTA338-16	KX282785	KX282190	MTA451	M Abdullah MTA451	2013	Fresh	Al-Salmi
Helianthemum salicifolium	EDNA13-0033634	MTA343-16	KX282786		KTUH057	M Al-Dosari 4572	2000	Herbarium	Al-Salmi border station near Saudi Arabia
Helianthemum salicifolium	EDNA13-0034217	MTA341-16	KX282787	KX282191	KTUH058	R Halwagy 1057	1972	Herbarium	Um Gudayr 60 KM from Kuwait City
Helianthemum salicifolium	EDNA15-0042793	MTA342-16	KX282788	KX282192	KTUH059	R Halwagy 13-76	1976	Herbarium	Wadi Um-Al-Rimam
Heliotropium bacciferum	EDNA15-0042599	MTA345-16	KX282789		MTA029	M Abdullah MTA029	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Heliotropium bacciferum	EDNA15-0042654	MTA346-16	KX282790		MTA118	M Abdullah MTA118	2012	Fresh	Nuwaiseeb
Heliotropium bacciferum	EDNA15-0042675	MTA344-16		KX282193	MTA597	M Abdullah MTA597	2013	Fresh	Um-Neqa
Heliotropium kotschyi	EDNA15-0042738	MTA348-16	KX282791		KTUH011	R Halwagy RH111-76	1976	Herbarium	Al-Khiran near Saudi border station
Heliotropium kotschyi	EDNA15-0042763	MTA349-16	KX282792		KTUH012	L Boulos LB18223	1993	Herbarium	Al-Khiran 2 KM from Gulf Shore

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Heliotropium kotschyi	EDNA15-0042859	MTA347-16	KX282793		KTUH013	M Al-Dosari 1456	1995	Herbarium	Mina Abdullah Gulf shore
Herniaria hemistemon	EDNA15-0042694	MTA351-16	KX282794		KTUH017	KT Mathew 2550	1995	Herbarium	Mina Abdullah Gulf shore
Herniaria hemistemon	EDNA15-0042920	MTA352-16	KX282795	KX282194	KTUH018	KT Mathew 2634	1995	Herbarium	Jahra Al-Subbiyah road
Herniaria hemistemon	EDNA15-0042562	MTA350-16	KX282796	KX282195	MTA414	M Abdullah MTA414	2013	Fresh	Failaka Island
Herniaria hirsuta	EDNA15-0042836	MTA353-16	KX282797	KX282196	KTUH019	KT Mathew 3420	1998	Herbarium	Abdali near border station
Herniaria hirsuta	EDNA15-0042934	MTA354-16	KX282798	KX282197	KTUH020	KT Mathew 3371	1998	Herbarium	Abdali 8.5 KM near border station
Herniaria hirsuta	EDNA15-0042606	MTA355-16			MTA270	M Abdullah MTA270	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Hippocrepis areolata	EDNA15-0042915	MTA356-16	KX282799	KX282198	KTUH196	M Al-Dosari 2772	1998	Herbarium	KISR - Sulaibiya Research Station
Hippocrepis areolata	EDNA15-0042902	MTA357-16	KX282800	KX282199	MTA237	M Abdullah MTA237	2013	Fresh	KISR - Sulaibiya Research Station
Hippocrepis areolata	EDNA15-0042572	MTA358-16	KX282801	KX282200	MTA333	M Abdullah MTA333	2013	Fresh	KISR - Sulaibiya Research Station
Hippocrepis unisiliquosa	EDNA15-0042370	MTA359-16	KX282802	KX282201	KTUH198	M Al-Dosari 6002	2006	Herbarium	Al-Sulaibiyah power station
Hippocrepis unisiliquosa	EDNA15-0042365	MTA360-16	KX282803	KX282202	MTA407	M Abdullah MTA407	2013	Fresh	Failaka Island
Hordeum marinum	EDNA15-0042926	MTA363-16	KX282804	KX282203	KTUH474	M Al-Dosari 6062	2006	Herbarium	Sabah Al-Ahmad Nature Reserve
Hordeum marinum	EDNA15-0042347	MTA362-16	KX282805	KX282204	KTUH475	M Al-Dosari 5162	2001	Herbarium	Sabhan near the water station
Hordeum marinum	EDNA15-0042359	MTA361-16	KX282806	KX282205	KTUH476	M Al-Dosari 5779	2005	Herbarium	Al-Wafra farms Al-Ameri farm
Hordeum murinum	EDNA15-0043137	MTA364-16	KX282807		MTA255	M Abdullah MTA255	2013	Fresh	Mina Abdullah
Hordeum murinum	EDNA15-0043132	MTA366-16	KX282808	KX282206	MTA364	M Abdullah MTA364	2013	Fresh	Doha Entertainment City
Hordeum murinum	EDNA15-0043128	MTA365-16	KX282809	KX282207	MTA526	M Abdullah MTA526	2013	Fresh	Abdali
Horwoodia dicksoniae	EDNA15-0043001	MTA367-16	KX282810	KX282208	KTUH398	G Brown 965012	1997	Herbarium	Ali Al-Salem Air Base Salmi Road
Horwoodia dicksoniae	EDNA15-0043011	MTA368-16	KX282811	KX282209	KTUH399	I Ibrahim 1116	1990	Herbarium	Al-Mutla'a alomg the motorway to Al-Salmi
Horwoodia dicksoniae	EDNA15-0042633	MTA369-16	KX282812	KX282210	MTA318	M Abdullah MTA318	2013	Fresh	KISR - Sulaibiya Research Station
Hyoscyamus pusillus	EDNA15-0042649	MTA370-16	KX282813		KTUH298	R Halwagy 005/85	1985	Herbarium	Al-Salmi near border with Saudi Arabia
Hypecoum littorale	EDNA15-0042718	MTA371-16	KX282814	KX282211	KTUH161	KT Mathew 5429	2005	Herbarium	Sabah Al-Ahmad Nature Reserve
Hypecoum littorale	EDNA15-0042745	MTA373-16	KX282815	KX282212	KTUH162	KT Mathew 5299	2002	Herbarium	Jahra Al-Salmi roadside
Hypecoum littorale	EDNA15-0042571	MTA372-16	KX282816		MTA254	M Abdullah MTA254	2013	Fresh	Mina Abdullah

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Hypecoum pendulum	EDNA15-0042799	MTA374-16	KX282817	KX282213	KTUH163	G Brown 965050	1996	Herbarium	Al-Abraq
Hypecoum pendulum	EDNA15-0042614	MTA375-16			MTA338	M Abdullah MTA338	2013	Fresh	KISR - Sulaibiya Research Station
Ifloga spicata	EDNA15-0043090	MTA376-16	KX282818	KX282214	MTA246	M Abdullah MTA246	2013	Fresh	KISR - Sulaibiya Research Station
Ifloga spicata	EDNA15-0043057	MTA377-16	KX282819	KX282215	MTA257	M Abdullah MTA257	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Ifloga spicata	EDNA15-0043088	MTA378-16	KX282820	KX282216	MTA422	M Abdullah MTA422	2013	Fresh	Failaka Island
Imperata cylindrica	EDNA15-0042353	MTA379-16	KX282821	KX282217	KTUH480	M Al-Dosari 4881	2000	Herbarium	Al-Khiran near sea shore
Imperata cylindrica	EDNA15-0042364	MTA380-16	KX282822	KX282218	KTUH481	KT Mathew 4127	1998	Herbarium	Omariyah Agricultural Research station
Imperata cylindrica	EDNA15-0042953	MTA381-16	KX282823	KX282219	MTA546	M Abdullah MTA546	2013	Fresh	Subiyah
Juncus rigidus	EDNA15-0042721	MTA382-16	KX282824		KTUH335	M Al-Dosari 4685	2000	Herbarium	Khadma salt marshes
Juncus rigidus	EDNA15-0042842	MTA383-16	KX282825	KX282220	KTUH336	M Al-Dosari 6446	2009	Herbarium	Al-Jdaliat area
Koelpinia linearis	EDNA15-0043121	MTA384-16	KX282826	KX282221	MTA245	M Abdullah MTA245	2013	Fresh	KISR - Sulaibiya Research Station
Koelpinia linearis	EDNA15-0043108	MTA385-16	KX282827	KX282222	MTA450	M Abdullah MTA450	2013	Fresh	Al-Salmi
Koelpinia linearis	EDNA15-0043068	MTA386-16	KX282828	KX282223	MTA496	M Abdullah MTA496	2013	Fresh	Abdali
Lappula spinocarpos	EDNA15-0042693	MTA387-16	KX282829		MTA135	M Abdullah MTA135	2012	Fresh	Nuwaiseeb
Lappula spinocarpos	EDNA15-0042933	MTA388-16	KX282830	KX282224	MTA301	M Abdullah MTA301	2013	Fresh	Al-Liyah
Lappula spinocarpos	EDNA15-0042955	MTA389-16	KX282831	KX282225	MTA445	M Abdullah MTA445	2013	Fresh	Al-Salmi
Lasiurus hirsutus	EDNA15-0043125	MTA390-16	KX282832	KX282226	KTUH482	KT Mathew 2885	1996	Herbarium	Al-Mutla'a along the road of subbiyah
Lasiurus hirsutus	EDNA15-0043083	MTA391-16	KX282833	KX282227	MTA055	M Abdullah MTA055	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Lasiurus hirsutus	EDNA15-0043118	MTA392-16	KX282834	KX282228	MTA180	M Abdullah MTA180	2012	Fresh	KISR - Sulaibiya Research Station
Launaea angustifolia	EDNA13-0033629	MTA393-16	KX282835	KX282229	KTUH092	R Halwagy 1009	1972	Herbarium	Al-Khafji road 23 KM N of Al-Khafji
Launaea angustifolia	EDNA13-0034212	MTA394-16			KTUH093	KT Mathew 2888	1996	Herbarium	Al-Mutla'a - Al-Subiyah roadside
Launaea angustifolia	EDNA15-0042377	MTA395-16	KX282836	KX282230	KTUH094	KT Mathew 3652	1998	Herbarium	Jal Az Zor on the plateau
Launaea capitata	EDNA13-0033630	MTA396-16	KX282837	KX282231	KTUH095	M Al-Dosari 3773	1999	Herbarium	Al-Salmi near border with Saudi Arabia
Launaea capitata	EDNA13-0034213	MTA397-16	KX282838	KX282232	KTUH096	M Al-Dosari 2561	1997	Herbarium	Nuwaiseeb border station with Saudi Arabia
Launaea capitata	EDNA15-0042382	MTA398-16	KX282839	KX282233	KTUH097	M Al-Dosari 2611	1998	Herbarium	Al-Doha Sheikh Zaid preservative area

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Launaea mucronata	EDNA13-0033632	MTA399-16	KX282840	KX282234	MTA383	M Abdullah MTA383	2013	Fresh	Failaka Island
Launaea mucronata	EDNA15-0043101	MTA400-16	KX282841	KX282235	MTA505	M Abdullah MTA505	2013	Fresh	Abdali
Launaea mucronata	EDNA13-0034215	MTA401-16	KX282842	KX282236	MTA601	M Abdullah MTA601	2013	Fresh	Um-Neqa
Launaea nudicaulis	EDNA13-0033631	MTA402-16	KX282843	KX282237	KTUH098	KT Mathew 5473	2007	Herbarium	Sabah Al-Ahmad Nature Reserve
Launaea nudicaulis	EDNA15-0043120	MTA403-16	KX282844	KX282238	MTA202	M Abdullah MTA202	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Launaea nudicaulis	EDNA13-0034214	MTA404-16	KX282845		MTA467	M Abdullah MTA467	2013	Fresh	Al-Salmi
Leontodon laciniatus	EDNA15-0042958	MTA406-16	KX282846	KX282239	KTUH099	KT Mathew 5420	2005	Herbarium	Khor Al-Ami between the chalets and the mosque
Leontodon laciniatus	EDNA15-0042969	MTA405-16	KX282847	KX282240	KTUH100	R Halwagy 8-76	1976	Herbarium	Wadi Um Al-Rimam
Leontodon laciniatus	EDNA15-0042768	MTA407-16			MTA504	M Abdullah MTA504	2013	Fresh	Abdali
Lepidium aucheri	EDNA15-0042378	MTA408-16	KX282848	KX282241	KTUH400	M Halwagy 1156	1972	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Lepidium aucheri	EDNA15-0043076	MTA409-16	KX282849	KX282242	KTUH401	M Al-Dosari 5450	2003	Herbarium	KISR - Sulaibiya Research Station
Leptaleum filifolium	EDNA15-0042796	MTA412-16	KX282850	KX282243	KTUH405	R Halwagy 6-76	1976	Herbarium	Wadi Um Al-Rimam
Leptaleum filifolium	EDNA15-0043017	MTA411-16	KX282851		KTUH406	R Halwagy 1084	1972	Herbarium	Al-Dibdibah 15 KM N of Salmy
Leptaleum filifolium	EDNA15-0043023	MTA410-16		KX282244	KTUH407	R Halwagy 1120	1972	Herbarium	Al-Shaqq
Limonium carnosum	EDNA15-0042622	MTA413-16		KX282245	KTUH253	KT Mathew 3283	1998	Herbarium	Failaka Island salt depressions
Limonium carnosum	EDNA15-0042780	MTA414-16			KTUH254	R Halwagy 81/130	1981	Herbarium	Auhah Island
Limonium thouini	EDNA15-0042644	MTA416-16			KTUH255	M Dib & M Al-Dosari	2001	Herbarium	Al-Zor Power Station
Limonium thouini	EDNA15-0042806	MTA415-16		KX282246	KTUH256	KT Mathew 5311	2003	Herbarium	Al-Khiran near resort enterance
Linaria albifrons	EDNA15-0042648	MTA418-16	KX282852		KTUH289	R Halwagy 82/32	1982	Herbarium	Kadhmah Jal Az-Zor
Linaria albifrons	EDNA15-0042733	MTA417-16		KX282247	KTUH290	M Al-Dosari 2727	1998	Herbarium	Failaka Island
Linaria albifrons	EDNA15-0042811	MTA419-16	KX282853	KX282248	KTUH291	I Ibrahim 1044	1990	Herbarium	Um Al-Rimam protected area
Linaria simplex	EDNA15-0042785	MTA421-16	KX282854	KX282249	KTUH293	R Halwagy 004/85	1985	Herbarium	Al-Salmi near border with Saudi Arabia
Linaria simplex	EDNA15-0042832	MTA420-16	KX282855	KX282250	KTUH294	R Halwagy 82/33	1982	Herbarium	Kadhmah Jal Az-Zor
Linaria simplex	EDNA15-0042585	MTA422-16			MTA253	M Abdullah MTA253	2013	Fresh	Mina Abdullah
Loeflingia hispanica	EDNA15-0042947	MTA424-16	KX282856		KTUH021	KT Mathew 3668	1998	Herbarium	Jal Az Zor on the plateau

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Loeflingia hispanica	EDNA15-0042956	MTA423-16	KX282857		KTUH022	KT Mathew 3747	1998	Herbarium	19 KM from Al-Salmi - Jahra road
Loeflingia hispanica	EDNA15-0042629	MTA425-16		KX282251	MTA259	M Abdullah MTA259	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Lolium rigidum	EDNA15-0042384	MTA427-16	KX282858	KX282252	KTUH483	KT Mathew 3028	1997	Herbarium	Khuwaisat salt marshes
Lolium rigidum	EDNA15-0042389	MTA426-16	KX282859	KX282253	KTUH484	M Al-Dosari 2569	1997	Herbarium	Al-Nuwaiseeb border station
Lolium rigidum	EDNA15-0042369	MTA428-16	KX282860	KX282254	MTA010	M Abdullah MTA010	2012	Fresh	King Fahad High Way
Lotus halophilus	EDNA15-0042360	MTA429-16	KX282861	KX282255	MTA239	M Abdullah MTA239	2013	Fresh	KISR - Sulaibiya Research Station
Lotus halophilus	EDNA15-0042348	MTA431-16	KX282862	KX282256	MTA424	M Abdullah MTA424	2013	Fresh	Failaka Island
Lotus halophilus	EDNA15-0042354	MTA430-16	KX282863	KX282257	MTA551	M Abdullah MTA551	2013	Fresh	Subiyah
Lycium shawii	EDNA15-0042669	MTA433-16			MTA049	M Abdullah MTA049	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Lycium shawii	EDNA15-0042707	MTA432-16	KX282864	KX282258	MTA188	M Abdullah MTA188	2012	Fresh	KISR - Sulaibiya Research Station
Lycium shawii	EDNA15-0042689	MTA434-16	KX282865		MTA543	M Abdullah MTA543	2013	Fresh	Subiyah
Malcolmia grandiflora	EDNA15-0043072	MTA436-16	KX282866	KX282259	MTA291	M Abdullah MTA291	2013	Fresh	Al-Liyah
Malcolmia grandiflora	EDNA15-0043075	MTA435-16	KX282867	KX282260	MTA310	M Abdullah MTA310	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Malcolmia grandiflora	EDNA15-0043055	MTA437-16	KX282868	KX282261	MTA485	M Abdullah MTA485	2013	Fresh	Abdali
Malva parviflora	EDNA15-0042641	MTA440-16	KX282869	KX282262	MTA409	M Abdullah MTA409	2013	Fresh	Failaka Island
Malva parviflora	EDNA15-0042619	MTA438-16	KX282870	KX282263	MTA475	M Abdullah MTA475	2013	Fresh	Abdali
Malva parviflora	EDNA15-0042595	MTA439-16	KX282871	KX282264	MTA553	M Abdullah MTA553	2013	Fresh	Subiyah
Maresia pygmaea	EDNA15-0042819	MTA442-16	KX282872	KX282265	KTUH411	R Halwagy 1-76	1976	Herbarium	Al-Mutla'a 10 KM N of Jahra
Maresia pygmaea	EDNA15-0043035	MTA441-16			KTUH412	M Al-Dosari 4608	2000	Herbarium	Al-Funaitees area
Maresia pygmaea	EDNA15-0043041	MTA443-16	KX282873	KX282266	KTUH413	M Al-Dosari 5058	2001	Herbarium	Subbiyah opposite to the military camp
Matthiola longipetala	EDNA15-0042840	MTA446-16	KX282874	KX282267	KTUH414	R Halwagy 1103	1972	Herbarium	Al-Shaqq
Matthiola longipetala	EDNA15-0043047	MTA444-16	KX282875	KX282268	KTUH415	M Al-Dosari 4920	2001	Herbarium	Wafra Al-Amiri farm
Matthiola longipetala	EDNA15-0042322	MTA445-16	KX282876		KTUH416	L Boulos 17950	1993	Herbarium	5 KM E of Abraq
Medicago laciniata	EDNA15-0043049	MTA449-16	KX282877	KX282269	MTA261	M Abdullah MTA261	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Medicago laciniata	EDNA15-0042330	MTA448-16	KX282878	KX282270	MTA372	M Abdullah MTA372	2013	Fresh	Doha Entertainment City

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Medicago laciniata	EDNA15-0042342	MTA447-16	KX282879	KX282271	MTA420	M Abdullah MTA420	2013	Fresh	Failaka Island
Medicago rotata	EDNA15-0042380	MTA450-16	KX282880	KX282272	KTUH202	M Al-Dosari 3453	1998	Herbarium	Failaka Island
Melilotus indicus	EDNA15-0042748	MTA452-16	KX282881	KX282273	KTUH205	M Al-Dosari 3731	1999	Herbarium	KISR - Sulaibiya Research Station
Melilotus indicus	EDNA15-0042942	MTA451-16	KX282882	KX282274	KTUH206	M Al-Dosari 4962	2001	Herbarium	Kuwait University Campus
Melilotus indicus	EDNA15-0042963	MTA453-16	KX282883	KX282275	KTUH207	M Al-Dosari 3859	1999	Herbarium	Al-Wafra Al-Ameri farm
Mesembryanthemum nodiflorum	EDNA15-0042560	MTA456-16	KX282884		MTA007	M Abdullah MTA007	2012	Fresh	King Fahad High Way
Mesembryanthemum nodiflorum	EDNA15-0042708	MTA454-16	KX282885	KX282276	MTA393	M Abdullah MTA393	2013	Fresh	Failaka Island
Mesembryanthemum nodiflorum	EDNA15-0042618	MTA455-16	KX282886		MTA525	M Abdullah MTA525	2013	Fresh	Abdali
Moltkiopsis ciliata	EDNA15-0042605	MTA459-16	KX282887	KX282277	MTA294	M Abdullah MTA294	2013	Fresh	Al-Liyah
Moltkiopsis ciliata	EDNA15-0042894	MTA457-16		KX282278	MTA429	M Abdullah MTA429	2013	Fresh	Failaka Island
Moltkiopsis ciliata	EDNA15-0042628	MTA458-16	KX282888	KX282279	MTA471	M Abdullah MTA471	2013	Fresh	Al-Salmi
Monsonia nivea	EDNA15-0042841	MTA461-16		KX282280	KTUH158	KT Mathew 2891	1996	Herbarium	AL-Mutla'a
Monsonia nivea	EDNA15-0042939	MTA462-16	KX282889	KX282281	KTUH159	M Al-Dosari 4507	1999	Herbarium	Al-Dba'iyyah 55 KM SE Kuwait City
Monsonia nivea	EDNA15-0042952	MTA460-16			KTUH160	KT Mathew 5345	2004	Herbarium	Sabah Al-Ahmad Nature Reserve
Neotorularia torulosa	EDNA15-0042328	MTA465-16		KX282282	KTUH417	R Halwagy 1104	1972	Herbarium	Al-Shaqq
Neotorularia torulosa	EDNA15-0042334	MTA464-16	KX282890	KX282283	KTUH418	M Al-Dosari 5032	2001	Herbarium	Al-Salmi border station near Saudi Arabia
Neotorularia torulosa	EDNA15-0042858	MTA463-16	KX282891	KX282284	MTA278	M Abdullah MTA278	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Neotorularia torulosa	EDNA15-0042564	MTA466-16	KX282892	KX282285	MTA279	M Abdullah MTA279	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Neurada procumbens	EDNA15-0042620	MTA469-16	KX282893	KX282286	KTUH239	KT Mathew 3132	1997	Herbarium	Al-Nuwaiseeb border station
Neurada procumbens	EDNA15-0042596	MTA467-16	KX282894	KX282287	MTA327	M Abdullah MTA327	2013	Fresh	KISR - Sulaibiya Research Station
Neurada procumbens	EDNA15-0042579	MTA468-16	KX282895	KX282288	MTA552	M Abdullah MTA552	2013	Fresh	Subiyah
Nitraria retusa	EDNA15-0042852	MTA470-16	KX282896	KX282289	MTA123	M Abdullah MTA123	2012	Fresh	Nuwaiseeb -near Saudi Arabia border
Nitraria retusa	EDNA15-0042918	MTA471-16	KX282897	KX282290	MTA162	M Abdullah MTA162	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Nitraria retusa	EDNA15-0042931	MTA472-16	KX282898	KX282291	MTA187	M Abdullah MTA187	2012	Fresh	KISR - Sulaibiya Research Station

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Notoceras bicorne	EDNA15-0042863	MTA473-16	KX282899		KTUH419	M Halwagy 1158	1972	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Notoceras bicorne	EDNA15-0042340	MTA475-16	KX282900	KX282292	KTUH420	R Halwagy 1145	1972	Herbarium	Wadi Um Al-Rimam
Notoceras bicorne	EDNA15-0042346	MTA474-16	KX282901	KX282293	KTUH421	M Al-Dosari 5905	2005	Herbarium	Al-Salmi
Ochradenus baccatus	EDNA15-0042930	MTA478-16		KX282294	KTUH271	M Al-Dosari 6367	2009	Herbarium	KISR - Sulaibiya Research Station
Ochradenus baccatus	EDNA15-0042944	MTA476-16	KX282902	KX282295	KTUH272	KT Mathew 4403	1999	Herbarium	Um Al-Rimam open area in the Wadi
Ochradenus baccatus	EDNA15-0042870	MTA477-16	KX282903	KX282296	MTA578	M Abdullah MTA578	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Ogastemma pusillum	EDNA15-0042790	MTA479-16			KTUH014	KT Mathew 3923	1998	Herbarium	Doha Entertainment City
Ogastemma pusillum	EDNA15-0042891	MTA480-16	KX282904	KX282297	KTUH015	R Halwagy RH23-83	1983	Herbarium	Wadi Um Al-Rimam close to water catchment
Ogastemma pusillum	EDNA15-0042906	MTA481-16	KX282905		KTUH016	R Halwagy 1185	1972	Herbarium	Khabrat Al-Awazem 40 KM of Kuwait City
Oligomeris linifolia	EDNA15-0042624	MTA484-16	KX282906		KTUH273	M Al-Dosari 6398	2009	Herbarium	Al-Abdaly Waleed Al-Omori farm
Oligomeris linifolia	EDNA15-0042731	MTA482-16	KX282907	KX282298	KTUH274	KT Mathew 3572	1998	Herbarium	Sulaibikhat sea side on the way to Doha
Oligomeris linifolia	EDNA15-0042600	MTA483-16	KX282908	KX282299	MTA384	M Abdullah MTA384	2013	Fresh	Failaka Island
Onobrychis ptolemaica	EDNA15-0042973	MTA485-16	KX282909	KX282300	KTUH208	M Halwagy 1067	1972	Herbarium	Jal Az Zor on the plateau
Onobrychis ptolemaica	EDNA15-0042982	MTA486-16	KX282910	KX282301	KTUH209	M Al-Dosari 6091	2007	Herbarium	Subiyah power station
Onobrychis ptolemaica	EDNA15-0042803	MTA487-16	KX282911	KX282302	MTA464	M Abdullah MTA464	2013	Fresh	Al-Salmi
Ononis reclinata	EDNA15-0042375	MTA489-16	KX282912	KX282303	KTUH210	KT Mathew 3634	1998	Herbarium	AL-Jahra along Abdali road
Ononis reclinata	EDNA15-0042385	MTA488-16	KX282913	KX282304	KTUH211	M Al-Dosari 1437	1995	Herbarium	Al-Nuwaiseeb border station
Ononis serrata	EDNA15-0042884	MTA492-16	KX282914	KX282305	KTUH212	M Leo 171	1998	Herbarium	KISR - Sulaibiya Research Station
Ononis serrata	EDNA15-0042990	MTA490-16	KX282915	KX282306	KTUH213	KT Mathew 5341	2004	Herbarium	Al-Wafra Juwairiah farm
Ononis serrata	EDNA15-0042997	MTA491-16	KX282916	KX282307	KTUH214	M Al-Dosari 5727	2005	Herbarium	Sabah Al-Ahmad Nature Reserve
Orobanche aegyptiaca	EDNA15-0042868	MTA493-16			KTUH230	M Al-Dosari 6357	2009	Herbarium	KISR - Sulaibiya Research Station
Orobanche cernua	EDNA15-0042726	MTA495-16		KX282308	KTUH231	R Halwagy 1231	1972	Herbarium	Failaka Island
Orobanche cernua	EDNA15-0042826	MTA496-16		KX282309	KTUH232	M Al-Dosari 1533	1995	Herbarium	Jahra along Jahra - Subiyah road
Orobanche cernua	EDNA15-0042846	MTA494-16			KTUH233	R Halwagy 17/83	1983	Herbarium	AL-Khiran near border with Saudi Arabia
Panicum turgidum	EDNA15-0042398	MTA499-16	KX282917	KX282310	KTUH486	KT Mathew 3227	1997	Herbarium	Um Al-Heiman area along roadside

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Panicum turgidum	EDNA15-0042374	MTA497-16	KX282918	KX282311	MTA137	M Abdullah MTA137	2012	Fresh	Nuwaiseeb
Panicum turgidum	EDNA15-0042394	MTA498-16	KX282919	KX282312	MTA513	M Abdullah MTA513	2013	Fresh	Abdali
Parapholis incurva	EDNA15-0042379	MTA500-16			KTUH487	M Al-Dosari 1340	1994	Herbarium	Al-Wafra farms
Parapholis incurva	EDNA15-0042402	MTA501-16	KX282920	KX282313	KTUH488	M Al-Dosari 3321	1998	Herbarium	Subiyah near the sea shore
Parapholis incurva	EDNA15-0042405	MTA502-16	KX282921	KX282314	KTUH489	M Al-Dosari 4771	2000	Herbarium	Um Neqa Ajayed farm
Paronychia arabica	EDNA15-0042343	MTA503-16	KX282922	KX282315	MTA236	M Abdullah MTA236	2013	Fresh	KISR - Sulaibiya Research Station
Paronychia arabica	EDNA15-0042892	MTA504-16	KX282923	KX282316	MTA413	M Abdullah MTA413	2013	Fresh	Failaka Island
Paronychia arabica	EDNA15-0042337	MTA505-16	KX282924	KX282317	MTA469	M Abdullah MTA469	2013	Fresh	Al-Salmi
Peganum harmala	EDNA15-0042965	MTA506-16	KX282925	KX282318	KTUH326	M Al-Dosari 3380	1998	Herbarium	King Fahad Highway side of the road before Ahmadi
Peganum harmala	EDNA15-0042945	MTA507-16	KX282926	KX282319	MTA148	M Abdullah MTA148	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Peganum harmala	EDNA15-0042874	MTA508-16	KX282927	KX282320	MTA173	M Abdullah MTA173	2012	Fresh	KISR - Sulaibiya Research Station
Pennisetum divisum	EDNA15-0043081	MTA509-16	KX282928	KX282321	MTA108	M Abdullah MTA108	2012	Fresh	Nuwaiseeb -near Saudi Arabia border
Pennisetum divisum	EDNA15-0043124	MTA511-16	KX282929	KX282322	MTA168	M Abdullah MTA168	2012	Fresh	KISR - Sulaibiya Research Station
Pennisetum divisum	EDNA15-0043134	MTA510-16	KX282930	KX282323	MTA512	M Abdullah MTA512	2013	Fresh	Abdali
Phragmites australis	EDNA15-0043065	MTA514-16	KX282931		KTUH491	L Boulos 18226	1993	Herbarium	Along the motorway to Al-Nuwaiseeb salt marshes
Phragmites australis	EDNA15-0042408	MTA512-16	KX282932	KX282324	MTA347	M Abdullah MTA347	2013	Fresh	Sulaibiya - Coastal area
Phragmites australis	EDNA15-0043053	MTA513-16	KX282933	KX282325	MTA566	M Abdullah MTA566	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Picris babylonica	EDNA15-0043054	MTA515-16	KX282934	KX282326	MTA229	M Abdullah MTA229	2013	Fresh	Nuwaiseeb
Picris babylonica	EDNA15-0043086	MTA516-16	KX282935	KX282327	MTA381	M Abdullah MTA381	2013	Fresh	Failaka Island
Picris babylonica	EDNA15-0043066	MTA517-16	KX282936	KX282328	MTA486	M Abdullah MTA486	2013	Fresh	Abdali
Plantago amplexicaulis	EDNA13-0034209	MTA520-16	KX282937	KX282329	MTA244	M Abdullah MTA244	2013	Fresh	KISR - Sulaibiya Research Station
Plantago amplexicaulis	EDNA15-0042597	MTA518-16			MTA446	M Abdullah MTA446	2013	Fresh	Al-Salmi
Plantago amplexicaulis	EDNA13-0033625	MTA519-16	KX282938	KX282330	MTA503	M Abdullah MTA503	2013	Fresh	Abdali
Plantago boissieri	EDNA15-0042580	MTA521-16	KX282939	KX282331	MTA016	M Abdullah MTA016	2012	Fresh	King Fahad High Way
Plantago boissieri	EDNA13-0034208	MTA522-16	KX282940		MTA200	M Abdullah MTA200	2013	Fresh	Sabah Al-Ahmad Nature Reserve

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Plantago boissieri	EDNA13-0033624	MTA523-16	KX282941	KX282332	MTA305	M Abdullah MTA305	2013	Fresh	Al-Liyah
Plantago ciliata	EDNA13-0033621	MTA526-16	KX282942	KX282333	KTUH243 a	M Al-Dosari 1768	1996	Herbarium	Al-Salmi border station near Saudi Arabia
Plantago ciliata	EDNA13-0034205	MTA524-16		KX282334	KTUH244	KT Mathew 4424	1999	Herbarium	Kathma salt marshes
Plantago ciliata	EDNA15-0042590	MTA525-16			KTUH245	KT Mathew 3619	1998	Herbarium	AL-Jahra along Abdali road
Plantago coronopus	EDNA15-0042566	MTA527-16	KX282943	KX282335	MTA322	M Abdullah MTA322	2013	Fresh	KISR - Sulaibiya Research Station
Plantago coronopus	EDNA13-0033626	MTA529-16	KX282944	KX282336	MTA388	M Abdullah MTA388	2013	Fresh	Failaka Island
Plantago coronopus	EDNA13-0034210	MTA528-16	KX282945		MTA499	M Abdullah MTA499	2013	Fresh	Abdali
Plantago lanceolata	EDNA13-0033622	MTA530-16		KX282337	KTUH246	M Al-Dosari 5229	2001	Herbarium	Kuwait University Khaldiyah Campus
Plantago lanceolata	EDNA13-0034206	MTA532-16			KTUH247	M Al-Dosari 3963	1999	Herbarium	Gulf road near KISR water front project
Plantago lanceolata	EDNA15-0042643	MTA531-16		KX282338	KTUH248	M Al-Dosari 6457	2009	Herbarium	Al-Abdaly Boshihry farm
Plantago notata	EDNA13-0033628	MTA533-16	KX282946	KX282339	KTUH249	R Halwagy 74-1974	1974	Herbarium	Khabrat Um-Omara Al-Shaqq
Plantago ovata	EDNA15-0042621	MTA536-16	KX282947	KX282340	KTUH243 b	M Al-Dosari 1768b	1996	Herbarium	Al-Salmi border station near Saudi Arabia
Plantago ovata	EDNA13-0034211	MTA534-16	KX282948	KX282341	MTA298	M Abdullah MTA298	2013	Fresh	Al-Liyah
Plantago ovata	EDNA13-0033627	MTA535-16	KX282949		MTA391	M Abdullah MTA391	2013	Fresh	Failaka Island
Plantago psammophila	EDNA13-0033623	MTA537-16	KX282950		KTUH250	R Halwagy 33	1981	Herbarium	Sulaibiyah 13 KM SE Kuwait International Airport
Plantago psammophila	EDNA13-0034207	MTA539-16	KX282951	KX282342	KTUH251	L Boulos 18124	1993	Herbarium	Al-Jahra along the highway to Al-Salmi
Plantago psammophila	EDNA15-0042728	MTA538-16		KX282343	KTUH252	M Al-Dosari 2019	1997	Herbarium	Al-Khuwaisat salt marshes
Polycarpaea repens	EDNA15-0042967	MTA542-16	KX282952	KX282344	KTUH023	KT Mathew 3052	1997	Herbarium	Subiyah gulf shore along the coast
Polycarpaea repens	EDNA15-0042977	MTA541-16	KX282953	KX282345	KTUH024	KT Mathew 4517	1999	Herbarium	Al-Mutla'a police outpost Abdali road
Polycarpaea repens	EDNA15-0042713	MTA540-16	KX282954		MTA035	M Abdullah MTA035	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Polycarpaea robbairea	EDNA15-0042907	MTA543-16	KX282955	KX282346	KTUH025	KT Mathew 4860	2000	Herbarium	Failaka Island along the coastal side
Polycarpon tetraphyllum	EDNA15-0042739	MTA544-16	KX282956		KTUH026	KT Mathew 5007	2001	Herbarium	Subbiyah water front near power station enterance
Polycarpon tetraphyllum	EDNA15-0042985	MTA545-16	KX282957	KX282347	KTUH027	KT Mathew 5363	2004	Herbarium	Mischan Island
Polycarpon tetraphyllum	EDNA15-0042992	MTA546-16	KX282958		KTUH028	KT Mathew 5102	2001	Herbarium	Failaka Island
Polypogon monspeliensis	EDNA15-0043056	MTA547-16	KX282959	KX282348	KTUH495	M Al-Dosari 1987	1997	Herbarium	Al-Khuwaisat salt marshes

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Polypogon monspeliensis	EDNA15-0043067	MTA549-16			KTUH496	S Morshed 1018	1989	Herbarium	Al-Doha near irrigated area along roadside
Polypogon monspeliensis	EDNA15-0042411	MTA548-16	KX282960	KX282349	MTA025	M Abdullah MTA025	2012	Fresh	King Fahad High Way
Psylliostachys spicata	EDNA15-0042828	MTA550-16	KX282961	KX282350	KTUH257	KT Mathew 5107	2001	Herbarium	Failaka Island
Pteranthus dichotomus	EDNA15-0042764	MTA552-16			KTUH029	KT Mathew 5020	2001	Herbarium	Subbiyah Kuwait City road along the road
Pteranthus dichotomus	EDNA15-0042998	MTA553-16			KTUH030	KT Mathew 5349	2004	Herbarium	Sabah Al-Ahmad Nature Reserve top of the escarpments
Pteranthus dichotomus	EDNA15-0043003	MTA551-16			KTUH031	KT Mathew 4758	2000	Herbarium	Um Al-Rimam into the flat wadi
Pulicaria undulata	EDNA15-0042979	MTA557-16	KX282962	KX282351	KTUH103	R Halwagy 012/85	1985	Herbarium	Subiyah facing Bubiyan Island
Pulicaria undulata	EDNA15-0042994	MTA554-16	KX282963	KX282352	KTUH104	KT Mathew 4612	1999	Herbarium	KISR - Sulaibiya Research Station
Pulicaria undulata	EDNA15-0042608	MTA555-16	KX282964	KX282353	MTA045	M Abdullah MTA045	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Pulicaria undulata	EDNA15-0042631	MTA556-16	KX282965	KX282354	MTA046	M Abdullah MTA046	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Reichardia tingitana	EDNA15-0042987	MTA560-16	KX282966	KX282355	KTUH105	KT Mathew 2764	1996	Herbarium	Nuwaiseeb border station with Saudi Arabia
Reichardia tingitana	EDNA15-0043000	MTA559-16	KX282967	KX282356	KTUH106	KT Mathew 3396	1998	Herbarium	Al-Abdali
Reichardia tingitana	EDNA15-0042563	MTA558-16	KX282968	KX282357	MTA426	M Abdullah MTA426	2013	Fresh	Failaka Island
Reseda arabica	EDNA15-0042830	MTA563-16	KX282969	KX282358	KTUH275	G Al-Abbadi 275	1998	Herbarium	KISR - Sulaibiya Research Station
Reseda arabica	EDNA15-0042808	MTA562-16	KX282970	KX282359	MTA248	M Abdullah MTA248	2013	Fresh	KISR - Sulaibiya Research Station
Reseda arabica	EDNA15-0042582	MTA561-16	KX282971	KX282360	MTA482	M Abdullah MTA482	2013	Fresh	Abdali
Reseda decursiva	EDNA15-0042782	MTA564-16	KX282972	KX282361	KTUH276	M Al-Dosari 4756	2000	Herbarium	Um Neqa near the road side
Reseda decursiva	EDNA15-0042847	MTA565-16		KX282362	KTUH277	R Halwagy 1310	1976	Herbarium	Wadi Al-Batin
Reseda decursiva	EDNA15-0042904	MTA566-16			KTUH278	R Halwagy 27-76	1976	Herbarium	Wadi Um Al-Rimam
Reseda muricata	EDNA15-0042646	MTA569-16	KX282973	KX282363	MTA361	M Abdullah MTA361	2013	Fresh	Doha Entertainment City
Reseda muricata	EDNA15-0042886	MTA567-16	KX282974	KX282364	MTA481	M Abdullah MTA481	2013	Fresh	Abdali
Reseda muricata	EDNA15-0042917	MTA568-16	KX282975	KX282365	MTA534	M Abdullah MTA534	2013	Fresh	Subiyah
Rhanterium epapposum	EDNA15-0042396	MTA571-16	KX282976	KX282366	MTA296	M Abdullah MTA296	2013	Fresh	Al-Liyah
Rhanterium epapposum	EDNA15-0042949	MTA572-16	KX282977	KX282367	MTA330	M Abdullah MTA330	2013	Fresh	KISR - Sulaibiya Research Station
Rhanterium epapposum	EDNA15-0042392	MTA570-16	KX282978	KX282368	MTA599	M Abdullah MTA599	2013	Fresh	Um-Neqa

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Roemeria hybrida	EDNA15-0042642	MTA573-16	KX282979	KX282369	KTUH240	KT Mathew 5096	2001	Herbarium	Failaka Island near red palace building
Roemeria hybrida	EDNA15-0042805	MTA574-16	KX282980	KX282370	KTUH241	M Al-Dosari 6345	2008	Herbarium	Al-Abdali Agayep farm
Roemeria hybrida	EDNA15-0042827	MTA575-16	KX282981	KX282371	KTUH242	M Halwagy 1160	1972	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Rostraria cristata	EDNA15-0042414	MTA576-16	KX282982	KX282372	KTUH497	KT Mathew 3017	1997	Herbarium	Khuwaisat salt marshes
Rostraria cristata	EDNA15-0043077	MTA578-16			KTUH498	KT Mathew 4966	2001	Herbarium	Nuwaiseeb border station with Saudi Arabia
Rostraria cristata	EDNA15-0043087	MTA577-16	KX282983	KX282373	KTUH499	KT Mathew 5046	2001	Herbarium	Doha salt marshes
Rostraria pumila	EDNA15-0043127	MTA579-16	KX282984	KX282374	MTA223	M Abdullah MTA223	2013	Fresh	Nuwaiseeb
Rostraria pumila	EDNA15-0043136	MTA580-16	KX282985		MTA363	M Abdullah MTA363	2013	Fresh	Doha Entertainment City
Rostraria pumila	EDNA15-0043093	MTA581-16	KX282986	KX282375	MTA408	M Abdullah MTA408	2013	Fresh	Failaka Island
Rumex pictus	EDNA15-0042623	MTA582-16	KX282987	KX282376	KTUH263	M Al-Dosari 5141	2001	Herbarium	Wadi Um-Al-Rimam
Rumex pictus	EDNA15-0042645	MTA584-16			KTUH264	M Halwagy 1065	1972	Herbarium	Jal Az-Zor
Rumex pictus	EDNA15-0042729	MTA583-16			KTUH265	R Halwagy 265	1974	Herbarium	Al-Mutla'a
Rumex vesicarius	EDNA15-0042781	MTA586-16	KX282988	KX282377	MTA231	M Abdullah MTA231	2013	Fresh	Nuwaiseeb
Rumex vesicarius	EDNA15-0042869	MTA587-16	KX282989		MTA295	M Abdullah MTA295	2013	Fresh	Al-Liyah
Rumex vesicarius	EDNA15-0042829	MTA585-16			MTA434	M Abdullah MTA434	2013	Fresh	Failaka Island
Salicornia europaea	EDNA15-0042766	MTA590-16	KX282990	KX282378	KTUH370	S Morshed 1055	1989	Herbarium	Al-Doha salt marshes
Salicornia europaea	EDNA15-0042957	MTA589-16			KTUH371	M Al-Dosari 6079	2006	Herbarium	KISR - Sulaibiya Research Station
Salicornia europaea	EDNA15-0042630	MTA588-16	KX282991	KX282379	MTA348	M Abdullah MTA348	2013	Fresh	Sulaibiya - Coastal area
Salsola cyclophylla	EDNA15-0042711	MTA592-16			KTUH372	L Boulos 93	1993	Herbarium	Flora of UAE - Abu Dhabi
Salsola cyclophylla	EDNA15-0042381	MTA591-16		KX282380	KTUH373	M Bajwa 285-75	1975	Herbarium	Flora of KSA - Hufuf
Salsola imbricata	EDNA15-0042356	MTA594-16	KX282992	KX282381	MTA084	M Abdullah MTA084	2012	Fresh	Sixth Ring Road
Salsola imbricata	EDNA15-0042921	MTA593-16	KX282993	KX282382	MTA437	M Abdullah MTA437	2013	Fresh	Failaka Island
Salsola imbricata	EDNA15-0042350	MTA595-16	KX282994	KX282383	MTA476	M Abdullah MTA476	2013	Fresh	Abdali
Salsola jordanicola	EDNA15-0042792	MTA597-16		KX282384	KTUH374	KT Mathew 5383	2004	Herbarium	KISR - Sulaibiya Research Station
Salsola jordanicola	EDNA15-0042968	MTA596-16	KX282995	KX282385	KTUH375	L Boulos 18219	1993	Herbarium	Al-Khiran 2 KM from Gulf Shore

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Salsola jordanicola	EDNA15-0042978	MTA598-16	KX282996	KX282386	KTUH376	M Al-Dosari 5577	2004	Herbarium	Al-Sulaibiyah power station
Salvadora persica	EDNA15-0042871	MTA599-16	KX282997	KX282387	KTUH287	M Al-Dosari 6329	2008	Herbarium	Sabah Al-Ahmad Nature Reserve
Salvadora persica	EDNA15-0042887	MTA601-16	KX282998	KX282388	KTUH288	M Al-Dosari 5549	2004	Herbarium	40 KM from Al-Abdaly border
Salvadora persica	EDNA15-0042849	MTA600-16	KX282999	KX282389	MTA356	M Abdullah MTA356	2013	Fresh	Doha Entertainment City
Salvia aegyptiaca	EDNA15-0042776	MTA603-16			KTUH164	M Halwagy 1286	1976	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Salvia aegyptiaca	EDNA15-0042823	MTA602-16	KX283000	KX282390	KTUH165	M Al-Dosari 3475	1998	Herbarium	Failaka Island near old museum
Salvia aegyptiaca	EDNA15-0042575	MTA604-16			MTA466	M Abdullah MTA466	2013	Fresh	Al-Salmi
Salvia lanigera	EDNA15-0042722	MTA605-16	KX283001		KTUH166	KT Mathew 5278	2002	Herbarium	Al-Khiran near the village
Salvia lanigera	EDNA15-0042843	MTA606-16	KX283002	KX282391	KTUH167	KT Mathew 5421	2005	Herbarium	Khor Al-Ami between the chalets and the mosque
Salvia lanigera	EDNA15-0042865	MTA607-16	KX283003	KX282392	KTUH168	R Halwagy 1285	1976	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Salvia spinosa	EDNA15-0042747	MTA610-16	KX283004		KTUH169	M Leo 165	1998	Herbarium	KISR - Sulaibiya Research Station
Salvia spinosa	EDNA15-0042883	MTA608-16	KX283005	KX282393	KTUH170	KT Mathew 2849	1996	Herbarium	Al-Salmi along the roadside
Salvia spinosa	EDNA15-0042901	MTA609-16	KX283006	KX282394	KTUH171	R Halwagy 1270	1976	Herbarium	Al-Shagayah police station
Savignya parviflora	EDNA15-0042388	MTA613-16	KX283007	KX282395	MTA217	M Abdullah MTA217	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Savignya parviflora	EDNA15-0042373	MTA612-16	KX283008	KX282396	MTA234	M Abdullah MTA234	2013	Fresh	KISR - Sulaibiya Research Station
Savignya parviflora	EDNA15-0042910	MTA611-16	KX283009	KX282397	MTA443	M Abdullah MTA443	2013	Fresh	Al-Salmi
Scabiosa olivieri	EDNA15-0042771	MTA616-16	KX283010		KTUH126	M Al-Dosari 3126	1998	Herbarium	Al-Salmi border station near Saudi Arabia
Scabiosa olivieri	EDNA15-0042635	MTA614-16	KX283011	KX282398	MTA564	M Abdullah MTA564	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Scabiosa olivieri	EDNA15-0042568	MTA615-16	KX283012	KX282399	MTA596	M Abdullah MTA596	2013	Fresh	Um-Neqa
Scabiosa palaestina	EDNA15-0042820	MTA617-16	KX283013	KX282400	KTUH127	R Halwagy 1318	1976	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Scabiosa palaestina	EDNA15-0042612	MTA618-16	KX283014	KX282401	MTA595	M Abdullah MTA595	2013	Fresh	Um-Neqa
Scabiosa palaestina	EDNA15-0042797	MTA619-16			MTA611	M Abdullah MTA611	2013	Fresh	Al-Liyah
Schimpera arabica	EDNA15-0043069	MTA620-16	KX283015	KX282402	MTA293	M Abdullah MTA293	2013	Fresh	Al-Liyah
Schimpera arabica	EDNA15-0042407	MTA621-16	KX283016	KX282403	MTA317	M Abdullah MTA317	2013	Fresh	KISR - Sulaibiya Research Station
Schimpera arabica	EDNA15-0043052	MTA622-16	KX283017	KX282404	MTA441	M Abdullah MTA441	2013	Fresh	Al-Salmi

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Schismus arabicus	EDNA15-0043059	MTA625-16		KX282405	KTUH500	KT Mathew 3485	1998	Herbarium	Khuwaisat salt marshes
Schismus arabicus	EDNA15-0043089	MTA623-16	KX283018	KX282406	KTUH501	KT Mathew 5293	2002	Herbarium	Jahra along AL-Salmo road
Schismus arabicus	EDNA15-0043091	MTA624-16	KX283019	KX282407	KTUH502	M Al-Dosari 4644	2000	Herbarium	AL-Wafra farms Yousef Kamal farm
Schismus barbatus	EDNA15-0043119	MTA626-16	KX283020	KX282408	MTA075	M Abdullah MTA075	2012	Fresh	Sixth Ring Road
Schismus barbatus	EDNA15-0043070	MTA628-16	KX283021	KX282409	MTA120	M Abdullah MTA120	2012	Fresh	Nuwaiseeb
Schismus barbatus	EDNA15-0043122	MTA627-16	KX283022	KX282410	MTA415	M Abdullah MTA415	2013	Fresh	Failaka Island
Sclerocephalus arabicus	EDNA15-0042319	MTA629-16	KX283023	KX282411	KTUH032	KT Mathew 5350	2004	Herbarium	Sabah Al-Ahmad Nature Reserve top of the escarpments
Sclerocephalus arabicus	EDNA15-0042860	MTA631-16	KX283024	KX282412	MTA100	M Abdullah MTA100	2012	Fresh	Nuwaiseeb -near Saudi Arabia border
Sclerocephalus arabicus	EDNA15-0043029	MTA630-16	KX283025	KX282413	MTA386	M Abdullah MTA386	2013	Fresh	Failaka Island
Scorzonera tortuosissima	EDNA15-0043082	MTA632-16	KX283026	KX282414	KTUH107	KT Mathew 5307	2002	Herbarium	Al-Salmi border station near Saudi Arabia
Scrophularia desertii	EDNA15-0042706	MTA633-16	KX283027	KX282415	MTA095	M Abdullah MTA095	2012	Fresh	Nuwaiseeb -near Saudi Arabia border
Scrophularia desertii	EDNA15-0042668	MTA635-16	KX283028		MTA457	M Abdullah MTA457	2013	Fresh	Al-Salmi
Scrophularia desertii	EDNA15-0042688	MTA634-16	KX283029		MTA488	M Abdullah MTA488	2013	Fresh	Abdali
Seidlitzia rosmarinus	EDNA15-0042366	MTA638-16	KX283030	KX282416	MTA191	M Abdullah MTA191	2012	Fresh	Sulaibiya - Coastal area
Seidlitzia rosmarinus	EDNA15-0042391	MTA637-16	KX283031	KX282417	MTA538	M Abdullah MTA538	2013	Fresh	Subiyah
Seidlitzia rosmarinus	EDNA15-0042386	MTA636-16	KX283032	KX282418	MTA583	M Abdullah MTA583	2013	Fresh	Um-Neqa
Senecio glaucus	EDNA15-0042936	MTA639-16	KX283033	KX282419	MTA073	M Abdullah MTA073	2012	Fresh	Sixth Ring Road
Senecio glaucus	EDNA15-0042922	MTA641-16	KX283034	KX282420	MTA439	M Abdullah MTA439	2013	Fresh	Failaka Island
Senecio glaucus	EDNA15-0042387	MTA640-16	KX283035	KX282421	MTA497	M Abdullah MTA497	2013	Fresh	Abdali
Silene arabica	EDNA15-0042349	MTA644-16	KX283036	KX282422	MTA251	M Abdullah MTA251	2013	Fresh	Mina Abdullah
Silene arabica	EDNA15-0042355	MTA643-16	KX283037		MTA287	M Abdullah MTA287	2013	Fresh	Al-Liyah
Silene arabica	EDNA15-0042895	MTA642-16	KX283038	KX282423	MTA433	M Abdullah MTA433	2013	Fresh	Failaka Island
Silene arenosa	EDNA15-0042791	MTA647-16	KX283039		KTUH034	M Al-Dosari 6430	2009	Herbarium	Al-Abdali Wleed Al-Omery farm
Silene arenosa	EDNA15-0043008	MTA646-16	KX283040	KX282424	KTUH035	M Al-Saadi 106	2002	Herbarium	Failaka Island
Silene arenosa	EDNA15-0043014	MTA645-16	KX283041	KX282425	KTUH036	R Halwagy RH18-83	1983	Herbarium	AL-Khiran southern coatal area

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Silene villosa	EDNA15-0043020	MTA650-16	KX283042		KTUH038	M Al-Dosari 3259	1998	Herbarium	Um Qasr near Iraqi border
Silene villosa	EDNA15-0043026	MTA649-16	KX283043	KX282426	KTUH039	KT Mathew 2635	1995	Herbarium	Al-Subiyah along Jahra road
Silene villosa	EDNA15-0042607	MTA648-16	KX283044	KX282427	MTA233	M Abdullah MTA233	2013	Fresh	KISR - Sulaibiya Research Station
Sisymbrium irio	EDNA15-0042898	MTA653-16	KX283045	KX282428	MTA064	M Abdullah MTA064	2012	Fresh	Sixth Ring Road
Sisymbrium irio	EDNA15-0042368	MTA652-16	KX283046	KX282429	MTA219	M Abdullah MTA219	2013	Fresh	Nuwaiseeb
Sisymbrium irio	EDNA15-0042363	MTA651-16	KX283047	KX282430	MTA520	M Abdullah MTA520	2013	Fresh	Abdali
Sisymbrium orientale	EDNA15-0042879	MTA656-16	KX283048	KX282431	KTUH422	M Al-Dosari 1989	1997	Herbarium	Al-Khuwaisat salt marshes
Sisymbrium orientale	EDNA15-0042352	MTA655-16	KX283049	KX282432	KTUH423	M Al-Dosari 1715	1996	Herbarium	Al-Jahra - Al-salmi road
Sisymbrium orientale	EDNA15-0042358	MTA654-16	KX283050	KX282433	KTUH424	KT Mathew 4938	2000	Herbarium	Abdali Sheikh Ali Ajaib farm
Spergularia diandra	EDNA15-0042876	MTA658-16	KX283051	KX282434	MTA285	M Abdullah MTA285	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Spergularia diandra	EDNA15-0042331	MTA659-16	KX283052	KX282435	MTA337	M Abdullah MTA337	2013	Fresh	KISR - Sulaibiya Research Station
Spergularia diandra	EDNA15-0042325	MTA657-16	KX283053	KX282436	MTA385	M Abdullah MTA385	2013	Fresh	Failaka Island
Spergularia marina	EDNA15-0042815	MTA661-16			KTUH040	KT Mathew 3472	1998	Herbarium	Khuwaisat elevated plains leading to salt marshes
Spergularia marina	EDNA15-0043032	MTA662-16			KTUH041	KT Mathew 3041	1997	Herbarium	Subiyah gulf shore along the coast
Spergularia marina	EDNA15-0043038	MTA660-16	KX283054	KX282437	KTUH042	KT Mathew 4928	2000	Herbarium	Abdali Sheikh Ali Ajaib farm
Sphenopus divaricatus	EDNA15-0043073	MTA663-16			KTUH510	KT Mathew 2770	1996	Herbarium	Al-Khiran road no 285
Sphenopus divaricatus	EDNA15-0043109	MTA664-16			KTUH511	M Al-Dosari 4680	2000	Herbarium	Um Neqa near Al-Abdaly border
Sphenopus divaricatus	EDNA15-0043114	MTA665-16		KX282438	KTUH512	KT Mathew 5038	2001	Herbarium	Doha salt marshes by roadsides
Sporobolus arabicus	EDNA15-0043062	MTA666-16	KX283055	KX282439	KTUH506	R Halwagy 82/56	1982	Herbarium	Ras Az Zor
Sporobolus arabicus	EDNA15-0043098	MTA668-16	KX283056	KX282440	KTUH507	KT Mathew 4869	2000	Herbarium	Failaka Island coastal strip
Sporobolus arabicus	EDNA15-0043102	MTA667-16	KX283057	KX282441	KTUH508	M Al-Dosari 3524	1998	Herbarium	Subbiyah near police station by the sea shore
Stipa capensis	EDNA15-0043135	MTA671-16	KX283058	KX282442	MTA250	M Abdullah MTA250	2013	Fresh	KISR - Sulaibiya Research Station
Stipa capensis	EDNA15-0043126	MTA670-16	KX283059	KX282443	MTA435	M Abdullah MTA435	2013	Fresh	Failaka Island
Stipa capensis	EDNA15-0043085	MTA669-16	KX283060	KX282444	MTA528	M Abdullah MTA528	2013	Fresh	Abdali
Stipagrostis ciliata	EDNA15-0043113	MTA673-16	KX283061	KX282445	KTUH513	KT Mathew 4890	2001	Herbarium	Wadi Um Al-Rimam a closed water catchment

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Stipagrostis ciliata	EDNA15-0043116	MTA672-16		KX282446	KTUH514	M Leo 267	1998	Herbarium	Al-Khiran near the sea shore
Stipagrostis plumosa	EDNA15-0043133	MTA676-16	KX283062	KX282447	KTUH515	M Al-Dosari 4508	1999	Herbarium	Al-Dbaiyyah
Stipagrostis plumosa	EDNA15-0043079	MTA674-16	KX283063	KX282448	MTA031	M Abdullah MTA031	2012	Fresh	Sabah Al-Ahmad Nature Reserve
Stipagrostis plumosa	EDNA15-0043123	MTA675-16	KX283064	KX282449	MTA141	M Abdullah MTA141	2012	Fresh	Nuwaiseeb -near Saudi Arabia border
Suaeda aegyptiaca	EDNA15-0043045	MTA679-16	KX283065	KX282450	MTA129	M Abdullah MTA129	2012	Fresh	Nuwaiseeb
Suaeda aegyptiaca	EDNA15-0042861	MTA678-16	KX283066	KX282451	MTA189	M Abdullah MTA189	2012	Fresh	KISR - Sulaibiya Research Station
Suaeda aegyptiaca	EDNA15-0042320	MTA677-16	KX283067	KX282452	MTA190	M Abdullah MTA190	2012	Fresh	Sulaibiya - Coastal area
Suaeda aegyptiaca	EDNA15-0043027	MTA680-16	KX283068		MTA392	M Abdullah MTA392	2013	Fresh	Failaka Island
Suaeda vermiculata	EDNA15-0042816	MTA682-16		KX282453	KTUH377	L Boulos 18205	1993	Herbarium	Mina Abdullah salt marshes
Suaeda vermiculata	EDNA15-0042986	MTA681-16			KTUH378	L Boulos 18215	1993	Herbarium	Al-Qleiaa near Abdullah Al-Mubarak Mosque
Suaeda vermiculata	EDNA15-0042999	MTA683-16	KX283069		KTUH379	M Al-Dosari 6336	2008	Herbarium	West Jahra Bird Reserve
Tamarix aphylla	EDNA15-0042888	MTA686-16		KX282454	KTUH304	KT Mathew 3270	1997	Herbarium	Al-Salmi near border with Saudi Arabia
Tamarix aphylla	EDNA15-0042850	MTA685-16	KX283070	KX282455	MTA183	M Abdullah MTA183	2012	Fresh	KISR - Sulaibiya Research Station
Tamarix aphylla	EDNA15-0042872	MTA684-16	KX283071	KX282456	MTA603	M Abdullah MTA603	2013	Fresh	Um-Neqa
Tamarix aucheriana	EDNA15-0042586	MTA689-16	KX283072	KX282457	MTA149	M Abdullah MTA149	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Tamarix aucheriana	EDNA15-0042734	MTA687-16	KX283073	KX282458	MTA192	M Abdullah MTA192	2012	Fresh	Sulaibiya - Coastal area
Tamarix aucheriana	EDNA15-0042786	MTA688-16	KX283074		MTA544	M Abdullah MTA544	2013	Fresh	Subiyah
Teucrium oliverianum	EDNA15-0042914	MTA692-16	KX283075		KTUH172	R Halwagy 81/97	1981	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Teucrium oliverianum	EDNA15-0042616	MTA691-16			MTA458	M Abdullah MTA458	2013	Fresh	Al-Salmi
Teucrium oliverianum	EDNA15-0042962	MTA690-16	KX283076	KX282459	MTA462	M Abdullah MTA462	2013	Fresh	Al-Salmi
Teucrium polium	EDNA15-0042802	MTA695-16	KX283077		KTUH173	R Halwagy 761 a	1971	Herbarium	Wadi um Al-Rimam
Teucrium polium	EDNA15-0042927	MTA694-16			KTUH174	R Halwagy 761	1971	Herbarium	Al-Khiran
Teucrium polium	EDNA15-0042941	MTA693-16			KTUH175	M Dib & M Al-Dosari	2001	Herbarium	Zor power station Khiran
Traganum nudatum	EDNA15-0042395	MTA696-16			KTUH380	Linda Shuaib 380	1993	Herbarium	Al-Khiran near Gulf shore
Tribulus terrestris	EDNA15-0042834	MTA698-16	KX283078	KX282460	KTUH327	KT Mathew 4094	1998	Herbarium	Omariyah Agricultural Research station

Species	EDNA No.	BOLD ID	*rbcL* GenBank accession	ITS2 GenBank accession	Coll. ID	Collector and number	Year	Collection type	Locality/ region
Tribulus terrestris	EDNA15-0042626	MTA697-16	KX283079	KX282461	MTA066	M Abdullah MTA066	2012	Fresh	Sixth Ring Road
Tribulus terrestris	EDNA15-0042813	MTA699-16	KX283080	KX282462	MTA157	M Abdullah MTA157	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Trigonella anguina	EDNA15-0042928	MTA701-16	KX283081	KX282463	KTUH221	KT Mathew 2860	1996	Herbarium	Al-Salmi border station near Saudi Arabia
Trigonella anguina	EDNA15-0043007	MTA700-16	KX283082	KX282464	KTUH222	M Al-Dosari 5449	2003	Herbarium	KISR - Sulaibiya Research Station
Trigonella anguina	EDNA15-0043013	MTA702-16	KX283083	KX282465	KTUH223	M Al-Dosari 6406	2009	Herbarium	Sabah Al-Ahmad Nature Reserve
Trigonella hamosa	EDNA15-0043031	MTA705-16	KX283084	KX282466	KTUH224	M Al-Dosari 3978a	1999	Herbarium	Gulf road near KISR water front project
Trigonella hamosa	EDNA15-0043025	MTA704-16	KX283085	KX282467	MTA242	M Abdullah MTA242	2013	Fresh	KISR - Sulaibiya Research Station
Trigonella hamosa	EDNA15-0042954	MTA703-16	KX283086	KX282468	MTA276	M Abdullah MTA276	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Trigonella stellata	EDNA15-0042336	MTA707-16	KX283087	KX282469	MTA299	M Abdullah MTA299	2013	Fresh	Al-Liyah
Trigonella stellata	EDNA15-0043043	MTA708-16	KX283088	KX282470	MTA417	M Abdullah MTA417	2013	Fresh	Failaka Island
Trigonella stellata	EDNA15-0042324	MTA706-16	KX283089	KX282471	MTA474	M Abdullah MTA474	2013	Fresh	Abdali
Typha domingensis	EDNA15-0042690	MTA709-16	KX283090		KTUH516	KT Mathew 5217	2001	Herbarium	KISR - Sulaibiya Research Station
Typha domingensis	EDNA15-0042670	MTA710-16			MTA167	M Abdullah MTA167	2012	Fresh	PAAF Al-Rabiyah Nursery Kuwait
Valerianella dufresnia	EDNA15-0042651	MTA712-16			KTUH318	M Al-Dosari 5138	2001	Herbarium	Wadi Um Al-Rimam
Valerianella dufresnia	EDNA15-0042736	MTA711-16			KTUH319	M Al-Dosari 4994	2001	Herbarium	Al-Khiran near police station
Zilla spinosa	EDNA15-0043061	MTA713-16	KX283091	KX282472	KTUH425	M Halwagy 1154	1972	Herbarium	Wadi Al-Batin 18 KM N of Al-Salmi
Ziziphus nummularia	EDNA15-0042783	MTA714-16	KX283092	KX282473	KTUH279	KT Mathew 5392	2004	Herbarium	Al-Abdali near border station
Ziziphus nummularia	EDNA15-0042831	MTA715-16	KX283093		KTUH280	M Al-Dosari 5554	2004	Herbarium	Al-Retqa police station near Iraqi border
Ziziphus spina-christi	EDNA15-0042732	MTA718-16	KX283094	KX282474	KTUH281	KT Mathew 5424	2005	Herbarium	Al-Wafra farma cultivated tree
Ziziphus spina-christi	EDNA15-0042809	MTA717-16	KX283095		KTUH282	MA Raouf 1291	1996	Herbarium	Al-Salmi border station near Saudi Arabia
Ziziphus spina-christi	EDNA15-0042583	MTA716-16	KX283096	KX282475	MTA567	M Abdullah MTA567	2013	Fresh	Sabah Al-Ahmad Nature Reserve
Zygophyllum qatarense	EDNA15-0042691	MTA720-16	KX283097	KX282476	MTA132	M Abdullah MTA132	2012	Fresh	Nuwaiseeb
Zygophyllum qatarense	EDNA15-0042905	MTA721-16	KX283098	KX282477	MTA536	M Abdullah MTA536	2013	Fresh	Subiyah
Zygophyllum qatarense	EDNA15-0042890	MTA719-16	KX283099		MTA584	M Abdullah MTA584	2013	Fresh	Um-Neqa

End of Appendix 4.1

Chapter 5 Next-generation sequencing for molecular reconstruction of plant diversity using eDNA

5.1 Introduction

The ecosystems of Kuwait are under increasing pressure due to climate change, land degradation and habitat loss (Al-Awadhi et al., 2005). Rapid and reliable identification methods of taxa can help with the identification of the region's diversity, planning *in situ* conservation efforts and documenting the progress of restoration. Most field surveys only picture the actual aboveground plant growth and may fail to observe any species missing morphological characters such as leaves and flowers. DNA barcoding as a molecular method could identify specimens by sequencing a standard barcoding gene region and comparing it against a DNA database (e,g, Kress et al., 2005; CBOL, 2009; de Vere et al., 2012; Saarela et al., 2013; Liu et al., 2015); and creating a DNA barcoding reference library will provide a powerful tool for ecologists interested in studying below-ground floras (e.g. Kesanakurti et al., 2011; Hiiesalu et al., 2012; Lamb et al., 2016). Next-Generation Sequence (NGS) based techniques using metagenomics and metabarcoding methods can provide an in-depth and reliable source of identification and molecular information by assessing the biodiversity using environmental DNA (eDNA) samples (Yoccoz et al., 2012).

eDNA is a complex mixture of genetic material extracted from many different remains of organisms collected from the environment such as soil, water and air. Examples of eDNA biodiversity analysis include documenting microbial diversity in soil and water quality (Terrat et al. 2012 and Vierheilig et al. 2015), fungal ecology identification from forest soils (Baptista et al., 2015), herbivore and carnivore diet analysis studies (Pompanon et al., 2012; Shehzad et al., 2012), nematodes communities found in soil (Sapkota and Nicolaisen, 2015), below ground plant root diversity (Lamb et al., 2016), and airborne pollen monitoring (Kraaijeveld et al., 2015).

Metagenomics is defined as the direct genetic sequencing of total genomic DNA within an environmental sample (Thomas et al., 2012). These genomes are fragmented and require building a sequence library and are subject to analysis based on available DNA database (Lam et al., 2015). A number of NGS-based applications using metagenomics approach are presented in Table 5.1.

Table 5.1 Examples of metagenomics analysis

eDNA Samples	Note on the analysis methodology	Platform	Reference
Organic matter collected from windshield of a moving vehicle	Galaxy online platform performed class-level phylogenetic profiling	454 FLX Roche Life Sciences	Pond et al., 2009
Airborne microbial communities containing bacterial, fungal and plant sequences	Replicate sequences removed using DUST and RepeatMasker; BLASTn method applied	454 Titanium	Yooseph et al., 2013
Banded leaf monkey (*Presbytis femoralis*) diet analysis in the rainforest, identifying 59 plant spp.	Sequence library constructed (fragment size 280-300 bp) MEGABLAST searches applied using custom plant barcode database	Illumina HiSeq and MiSeq	Srivathsan et al., 2016
Leaf-feeding monkey (*Pygathrix nemaeus*) diet analysis identifying 16 plant spp.	FASTQC performed followed by FASTA sequences matched against diet and plant barcode custom databases using BLASTn method	Illumina HiSeq	Srivathsan et al., 2015

The major advantages of metagenomic is that it provides analysis of the total diversity of organisms within eDNA sample and sequencing is directly performed using NGS-platforms avoiding the need for prior selection of barcode markers and amplification (Srivathsan et al., 2016). Also, sampling and applying NGS-based methods by molecular ecologists is not restricted to a period of time (i.e. spring season) and is flexible to apply throughout the year according to the monitoring plan (e.g. DNA samples can be collected from the environment any time of the year and processed for sequencing following NGS methods) (Yoccoz et al., 2012). Major Bioinformatics

challenges arise for metagenomics approach is the need of reliable, high-quality reference database to match and identify the unknown sequences of the large sequence data sets.

For the metagenomics approach, following Srivathsan et al. (2015) methodology, an initial assessment of quality scores across Illumina data was performed using FASTQC (www.bioinformatics.bbsrc.ac.uk/projects /fastqc/) and sequences were analysed either with or without assembly. For the assembly-free analyses, FASTQ sequences were converted to FASTA format and matched against custom plant barcode database downloaded from GenBank using BLASTn (Altschul et al., 1990). For the assembled reads, sequences were assembled using SOAPdenovo2 method from Luo et al. (2012) and were matched against the custom plant database using MEGABLAST with a 98 % identity threshold for identification.

DNA metabarcoding approach identifies large sets of taxa present in an environmental DNA sample by limiting the survey to the PCR products and requires careful selection of barcode markers and primers to amplify widely over the taxa of interest (Taberlet et al., 2012; Cristescu , 2014). The use of long DNA barcodes for the taxonomic identification in an environmental mixture containing degraded DNA samples is likely to result in few positive matches and many species not being amplified (Yoccoz et al., 2012). Previous studies used short DNA markers (Table 5.2), widely using the non-coding plastid *trn*L (UAA) intron P6 loop, which have the potential to amplify across degraded DNA environmental samples mainly due to its short length, 10-143 bp (Yoccoz et al., 2012; Taberlet et al., 2007), however, species resolution is minimal, e.g. the P6 loop was able to identify only 47.2 % of 106 species belonging to the Arctic plant collection using a local DNA database (Taberlet et al. 2007). In a recent study by Fahner et al., (2016) a large-scale monitoring survey of plants was evaluated through eDNA metabarcoding of soil samples using four DNA markers: *mat*K, *rbc*L, ITS2, and the commonly used *trn*L P6 loop. The authors highly recommended the use of *rbc*L and ITS2 for biodiversity assessment of vascular plants from soil eDNA based on sequence recovery, annotation, and resolution; the other tested markers, *matK* had the lowest taxonomic recovery and *trnL* P6 loop showed the

least taxonomic resolution of recovered vascular plant sequences (Fahner et al., 2016). A number of metabarcoding applications are presented in Table 5.2.

Table 5.2. NGS-based applications using metabarcoding analysis

Samples	Gene region	NGS Platform	Reference
Silty sediments eDNA Reconstruction of past Arctic vegetation	*trnL*, *trnL* P6 loop	Roche 454 FLX	Sonstebo et al, 2010
Soil eDNA 3 Sites: Boreal, Temperate & Tropical	*trnL* P6 loop	Roche GS FLX: Roche 454 FLX Illumina GA IIx	Yoccoz et al, 2012
Sediments eDNA Pollen based vegetation analysis	*trnL* P6 loop	Roche GS FLX Platform	Parducci et al, 2013
Mixed plant roots and leaves samples	*trnL*	Illumina MiSeq	Lamb et al., 2016
Pollen monitoring	*trnL*	Ion Torrent 314 chip	Kraaijeveld et al., 2015
Monitoring plants through metabarcoding of soil eDNA	*trn*L P6 loop, *rbcL*, *mat*K, ITS2	Illumina MiSeq	Fahner et al., 2016

The metabarcoding method consists of extracting DNA from environmental samples, amplifying DNA markers, sequencing the amplicons using NGS platforms (Table 5.2), analysing the sequences to assessing the taxonomic diversity of the analysed environmental sample and identifying the organisms using available DNA barcode reference libraries (Taberlet et al. 2012; Orgiazzi et al., 2015).

Although the eDNA metabarcoding approach is powerful in species identification, several obstacles exist which can result in its failure. First, false negatives (the present taxa remain undetected) occur in metabarcoding studies due to degradation of template DNA, failures during amplification and sequencing process. The reduction of false negatives can be performed by increasing the number of replicated PCRs which improves the reliability of results (Ficetola et al., 2014). Second, issues with

finding a suitable conserved and short metabarcode region, variable between species for optimal taxonomic resolution and allow amplification from degraded eDNA samples (Epp et al., 2012). Degraded DNA from environmental samples often prevents the recovery of PCR fragments longer than 200 bp, (Goldstein and Rob, 2007). Third, bioinformatics challenges arise from the need for a comprehensive, high-quality reference DNA database (Srivathsan et al., 2016). The primary goal of DNA metabarcoding is to link an unknown DNA sequence to a taxonomic name by comparing the sequences with a reliable reference DNA database that includes specimens properly identified by a taxonomist, as well as their DNA barcodes (Coissac et al., 2012).

This chapter describes a test of metagenomics and metabarcoding as a method to address restoration projects in Kuwait. I will address three questions:
1. Does a useable amount of sequenceable /amplifiable DNA survive in soil samples from arid environments?
2. Does a metagenomics approach identify vascular plants known to grow in the area from soil samples?
3. Does a metabarcoding approach identify vascular plants known to grow in the area from soil samples?

Both sequencing sets will be examined for the presence of sequences identifiable as species known to grow in Kuwait using the barcode database assembled in chapter 4. I will also use public databases (Genbank) to identify any other vascular plant sequences and determine if this matches with what is known to grow in the area and what may have grown there in the past.

5.2 Methods

5.2.1 Study area

Soil samples were collected for investigating the plant diversity below-ground from two sites located at Um Neqa in the north-eastern portion of Kuwait, N 30^0 00' and E 47^0 55' (Figure 5.1). Um Neqa is a demilitarised zone (DMZ) fence protected with a total area of 246 km^2. The DMZ area has been fenced since 1994 by the United Nations Iraq-Kuwait Observation Mission (UNIKOM). A total of 120 km^2 of Um Neqa is considered as an open desert area (or open rangeland), as it is relatively distant from residential areas (ca. 50 km from Kuwait City). A neighbouring area, within 3 km west to Um Neqa is Al-Abdaly farms (Figure 5.1), which represents one of three large agricultural areas in Kuwait (other two are Al-Wafra and Kabd). The open desert area of Um Neqa is currently used for camping and livestock grazing.

Site 1 is a DMZ fence protected area of the study area, and site 2 is a disturbed open desert area, which is located outside the DMZ area (ca. 3 km south) (Figure 5.1). United Nations Compensation Commission (UNCC) funds will be spent on restoration and protection of the extended area outside the DMZ of Um Neqa (Figure 1.8, Chapter 1). Um Neqa is proposed as a future protected area as it represents native plant communities: *Hammada* and *Rhanterium* communities (Figure 5.2). The restoration plan, management, and monitoring of these sites will be implemented and monitored by Public Authority for Agriculture and Fisheries (PAAF) with the help of other local institutes in Kuwait and international consultants.

The two study sites at Um Neqa are chosen according to their highly similar functions based on the following features: soil types, geology, geomorphology, and mineral resources (Omar et al., 2001).

The two sites were surveyed in March 2014 to study the plant diversity and the vegetation cover (qualitative data) using the Braun-Blanquet cover-abundance scale (Wikum & Shanholtzer, 1978).

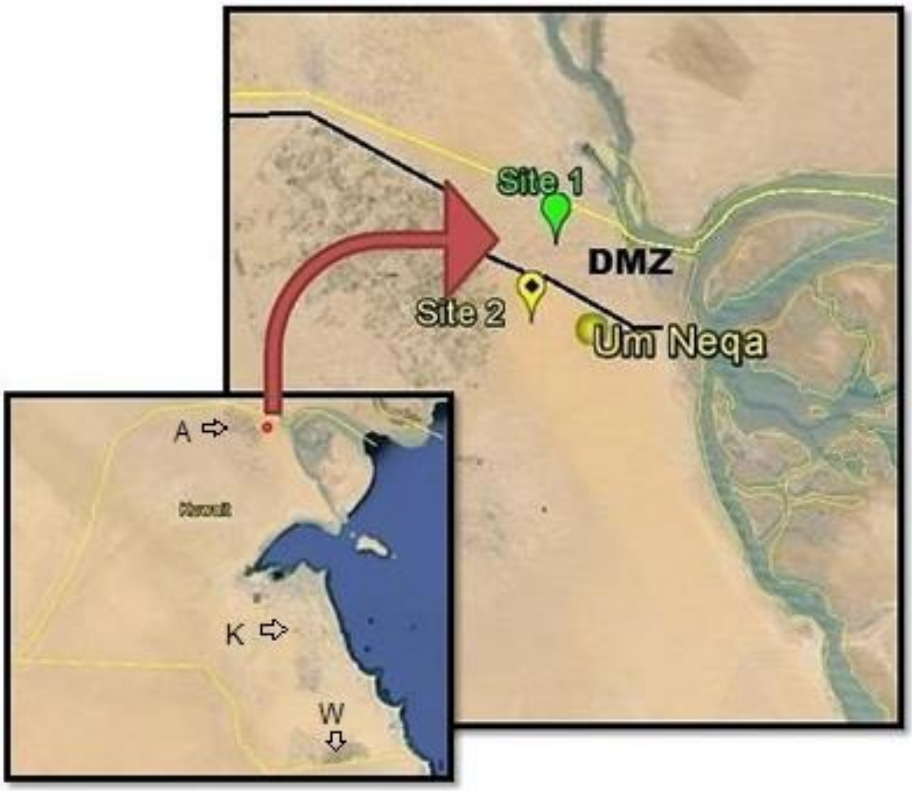


Figure 5.1. A map of Kuwait showing Um Neqa study area
(Green icon: Site 1 fence protected area and yellow icon: Site 2 open desert area).
(A) Al-Abdaly farms, (K) Kabd farms and (W) Wafra farms. (Between the black line and the yellow line is the DMZ fenced area) (*Google earth, 2015*).

Figure 5.2 Um Neqa (DMZ) protected area showing *Rhanterium* community
(Source: M Abdullah, 2014)

5.2.2. Soil sampling, eDNA extraction and NGS

A total of 40 soil samples representing 20 plots were collected from the two sites. Ten plots at each site divided into upper and lower soil layers. Soil corer metal cylinders were used for sampling (15 cm long x 5 cm in diameter; cleaned with ethanol to remove DNA before use). Rock, gravel and unwanted material removed either by hand or using a sieve of size 2-5 mm screen. For each plot, 30 g of soil from the top layer (0-5 cm depth) and 30 g deep layer (10-15 cm depth) were transferred into a tea bag separately. The tea bags were placed in clear plastic zip lock bags with silica gel granules to dry and preserve eDNA from any moisture before the extraction process following Yoccoz et al. (2012). For each soil sample, eDNA was extracted from 10 g of dry soil using the PowerMax™ Soil DNA isolation kit following the manufacture's protocol. The soil sampling and DNA extraction methods were adopted from several papers recently published using PowerMax™ Soil DNA isolation kit (Parducci et al, 2013; Yoccoz et al, 2012; Epp et al., 2012; Sonstebo et al, 2010). The soil samples were imported to the UK in April 2014 and kept in the Ferguson's Science Lab, RBGE, under the Soil Import Licence No. IMP/SOIL/29/2013.

5.2.2.1 Metagenomics MiSeq library preparation: The initial investigation of the eight eDNA soil extractions included the followings:

DNA quantification using Qubit® 2.0 Fluorometer using Qubit dsDNA High Sensitivity (HS) Assay Kit following the manufacturer's protocol (Catalogue no. Q32866, Invitrogen, UK), where necessary using DNA Speedvac vacuum concentrator, set at low heat for 1 hr.

DNA was visualised using 1 % agarose gel electrophoresis prepared in 100 ml of 1 x TBE buffer, mixed with 10 µl of SYBR Safe DNA gel stain (Invitrogen, UK). An amount of 5 µl of each sample was mixed with 3 µl of gel loading dye. DNA ladder of size 1 Kb plus (Invitrogen, UK) was loaded (3 µl) to determine the molecular size of the DNA bands. The solidified agarose gel was loaded with the samples and left to run for 50 mins at 80 V before being visualised using GeneSyn software and Gene Genius UV trans-illuminator system (Figure 5.3).

After the initial investigation, the eDNA samples were sent to Edinburgh Genomics, Ashworth Labs for Illumina MiSeq sequencing to generate 2 x 250 base paired-end sequencing from 1 pool of 8 samples of TruSeq Nano libraries (using TruSeq Nano DNA Library Prep. Kit, Illumina). At Ashworth Labs a quality control (QC) measurements were run on the samples which included a quality and quantity measurements followed by a recommendation of sample level requirements for sequencing. Metagenomics minimum DNA amount required is 250 ng. [Details of Edinburgh Genomics QC measurements and levels of sample requirements are available online at the following website: https://genomics.ed.ac.uk/resources/sample-requirements].

Table 5.3 Metagenomics: soil eDNA samples submitted for DNA library preparation and sequencing.

Soil Sample ID	Depth of Soil sampling (cm)	DNA Conc (ng/µl)	DNA Volume (µl)	Total DNA (ng)	DNA yield (µg)
Site 1:					
MA1	surface (0-5)	97.8	30	2934	2.9
MA2	surface (0-5)	73.45	30	2204	2.2
MA3	deep (15-20)	31.28	40	1251	1.3
MA4	deep (15-20)	23.46	40	938	1.0
Site 2:					
MA5	surface (0-5)	18.5	60	1110	1.1
MA6	surface (0-5)	48.75	55	2681	2.2
MA7	deep (15-20)	7.5	30	225	0.2 *
MA8	deep (15-20)	5.26	30	158	0.1 *

Site 1 – Fenced area, Site 2 – Open desert area
Asterisks (*) DNA quantity too low, Edinburgh Genomics did not recommend sequencing

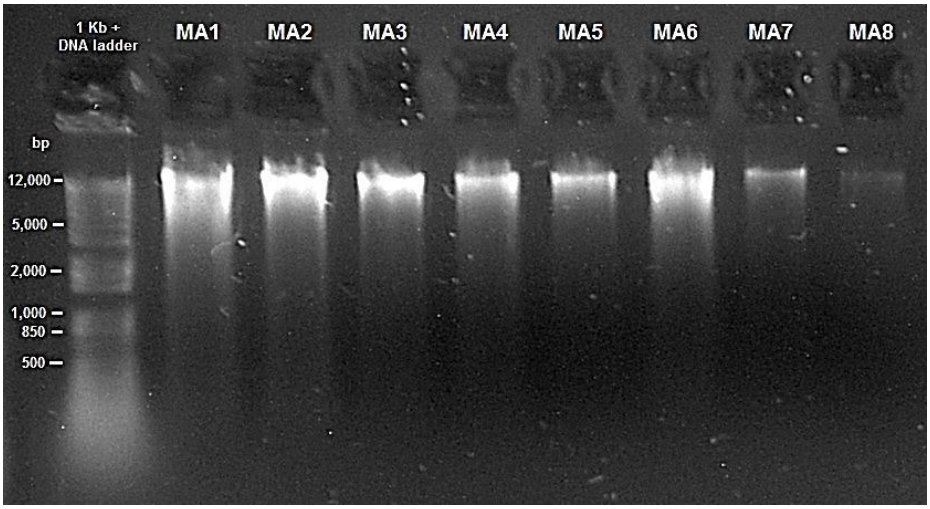


Figure 5.3 eDNA soil extractions.

Lane 1 represents 3 µl of size strands 1 Kb + DNA ladder (Invitrogen, UK), Lanes 2-9 represents samples MA1-MA8. MA1 and MA2 soil eDNA collected from site 1 surface soil (0-5 cm); MA3 and MA4 collected from site 1 deeper soil (10-15 cm); MA5 and MA6 collected from site 2 surface soil (0-5 cm); MA7 and MA8 collected from site 2 deeper soil (10-15 cm). For each sample 5 µl of eDNA extract was mixed with 3 µl of loading dye, loaded in the wells and run on 1 % gel agarose for 50 min at 80 V.

5.2.2.2 Metabarcoding MiSeq library preparation: The barcoding markers were chosen according to the DNA database of Kuwait, *rbcL* and ITS2 (Chapter 4). Other DNA markers were tested, *trnL* (UAA) intron and a shorter fragment of this intron the P6 loop. DNA markers and primers used to amplify eDNA from soil samples are listed in Table 5.4.

Table 5.4 Primer sequences and expected amplicon sizes for each DNA marker

DNA markers	Expected Size (bp)	Primers	Sequence (5' - 3')	Reference
***rbc*L**	500	aaF	ATGTCACCACAAACAGAGACTAAAGC	Kress & Erickson, 2007
***rbc*L**		rev	GTAAAATCAAGTCCACCRCG	Kress et al., 2009
ITS2	300-460	S2F	ATGCGATACTTGGTGTGAAT	Chen et al., 2010
ITS2		S3R	GACGCTTCTCCAGACTACAAT	Chen et al., 2010
***trn*L**	254-767	c	CGAAATCGGTAGACGCTACG	Taberlet et al., 1991
***trn*L**		d	GGGGATAGAGGGACTTGAAC	Taberlet et al., 1991
***trn*L P6 loop**	10-143	g	GGGCAATCCTGAGCCAA	Taberlet et al., 2007
***trn*L P6 loop**		h	CCATTGAGTCTCTGCACCTATC	Taberlet et al., 2007

Metabarcoding Polymerase Chain Reaction (PCR)

I tested a range of PCR additives such as betaine, bovine serum albumin (BSA), trehalose and / or Combinatorial PCR Enhancer Solution (CES) to optimise the PCR amplification. The primary PCR additive worked best was CES which includes a mixture of 2.7 M betaine, 6.7% dimethyl sulfoxide (DMSO) and 50 mg/ml BSA. Previous metabarcoding studies used mainly BSA solution as an enhancer for PCR (e.g. Sonstebo et al., 2010; EPP et al., 2012; Parducci et al., 2013) which is already included in the CES mix used in this study.

PCR for each DNA marker used was performed using one set of primers each (Table 5.4) in 50 µl reactions containing 1.5 Unit BIOTAQ DNA polymerase (Bioline, UK), 1 x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.5mN of each primer, 1 x

CES and 3.0 µl (15-20 ng/µl) genomic DNA. PCR mix used for each marker is presented in Table 5.5.

Table 5.5 PCR conditions for each DNA marker used

	DNA markers		
PCR reaction mix	*rbcL*	ITS2	*trnL*
PCR Buffer	1x	1x	1x
$MgCl_2$	1.5mM	1.5mM	1.5mM
dNTP mix	0.2mM	0.2mM	0.2mM
Forward Primer	1.5µM	1.5µM	1.5µM
Reverse Primer	1.5µM	1.5µM	1.5µM
CES	1x	1x	1x
BIOTAQ DNA Polymerase (Bioline)	0.15U/µL	0.15U/µL	0.15U/µL
DNA Template	3µL	3µL	3µL
Total Volume	50µL	50µL	50µL

I tested a range of PCR thermocycler conditions for each DNA marker. The range of annealing temperatures tested for *rbcL* was 48 ºC to 53 ºC and ITS2 was 50 ºC to 55 ºC, the best annealing temperature for each marker are listed in Table 5.6. The optimised PCR condition for *rbc*L and ITS2 are the ones that worked best in DNA barcoding the flora of Kuwait (see Chapter 4). For *trn*L (UAA) and the shorter fragment P6 loop a range of annealing temperatures was tested from 50 to 55 ºC. PCR programmes used for each DNA marker are listed in Table 5.6

Table 5.6 Thermocycler programmes used for each DNA marker

DNA marker	PCR cycle
***rbc*L**	94 C 1 min
	94 C 45 sec, 51 C 45 sec, 72 C 2 min, x 40 cycles
	72 C 7 min
	10 C forever
ITS2	95 C 4 min
	94 C 1 min, 55 C 1 min, 72 C 45 sec, x 30 cycles
	72 C 5 min
	10 C forever
***trn*L**	95 C 4 min
	94 C 1 min, 50-55 C 1 min, 72 C 2 min, x 35 cycles
	72 C 5 min
	10 C forever

The PCR products were visualised using 2.0 % agarose gel electrophoresis prepared in 100 ml of 1 x TBE buffer, mixed with 10 µl of SYBR Safe DNA gel stain (Invitrogen, UK). An amount of 5 µl of each sample was mixed with 3 µl of gel loading dye. DNA ladder of size 1 Kb plus (Invitrogen, UK) was loaded (3 µl) to determine the molecular size of the DNA bands. The solidified agarose gel was loaded with the samples and left to run for 50 mins at 80 V before being visualised using GeneSyn software and Gene Genius UV trans-illuminator system. Figure 5.4 to Figure 5.6 showing PCR products of *rbc*L and ITS2 for 40 metabarcoding samples including a positive (source: plant DNA) and negative control (Figure 5.4 – 5.6). The primers used for *trn*L (UAA) and the shorter fragment P6 loop did not amplify any of the soil samples.

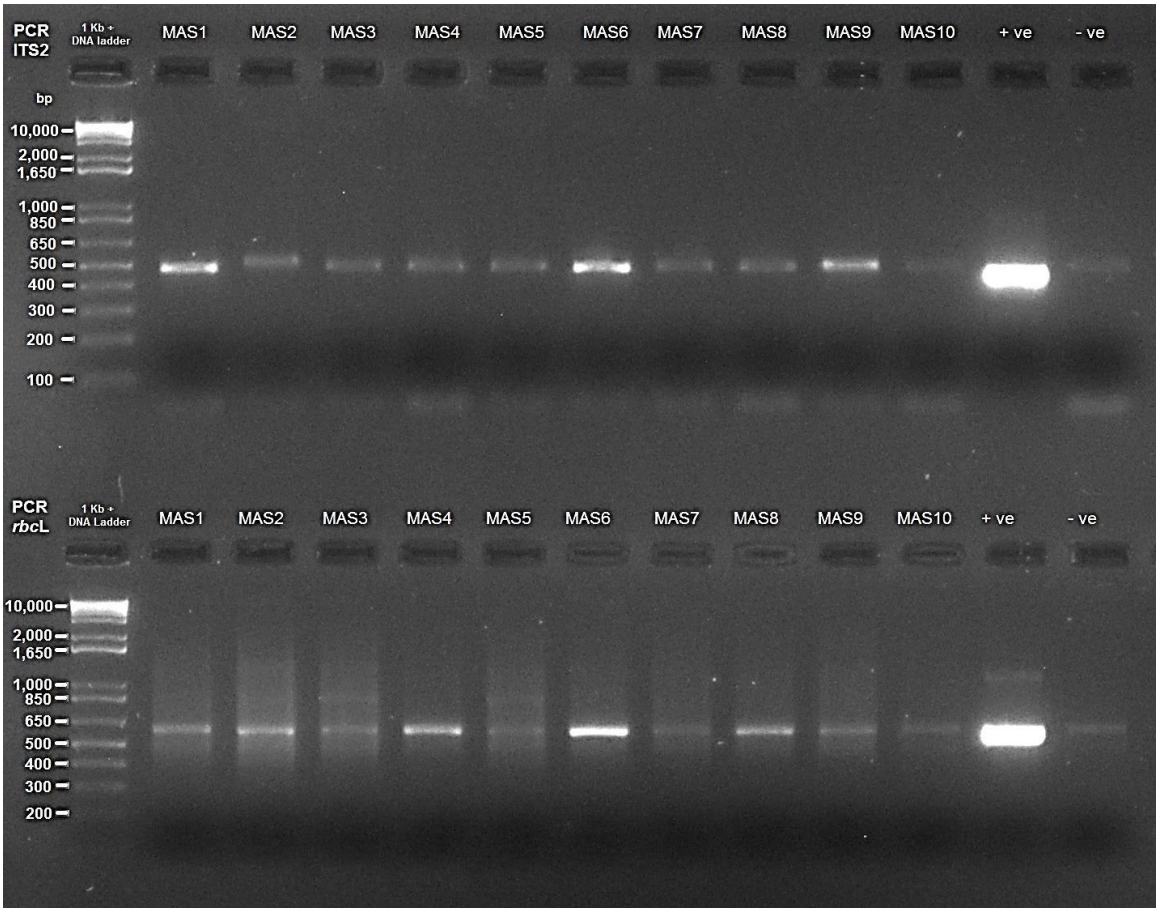


Figure 5.4 PCR products (*rbc*L and ITS2) for metabarcoding soil analyses

ITS2 PCR products (top) and *rbc*L (bottom). Lane 1 contains size strands of 1 Kb + DNA ladder (Invitrogen, UK), Lanes 2-11 representing samples MAS1-MAS10, Lane 12 representing a plant DNA extract as a positive control and Lane 13 a negative control. The PCR products were run on 2.0 % agarose gel for 50 min at 80 V.

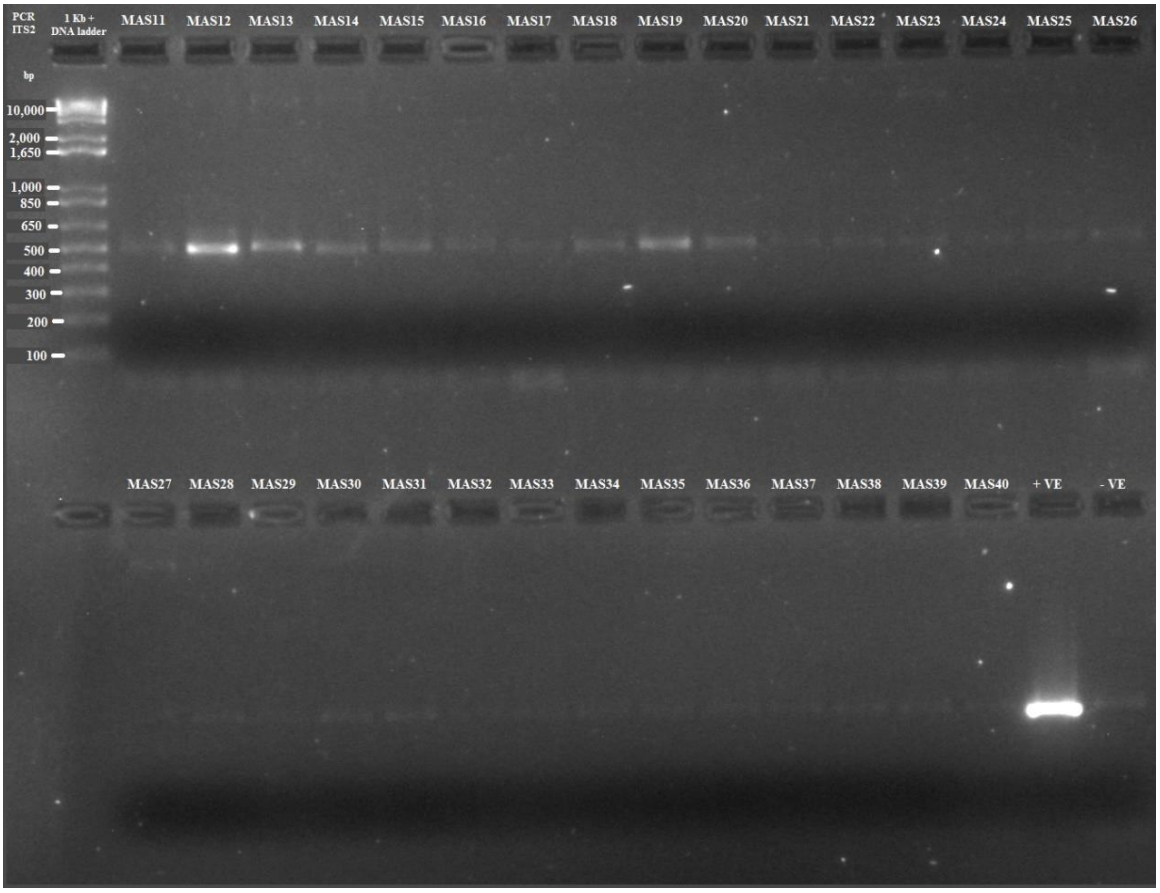


Figure 5.5 ITS2 PCR products for metabarcoding soil analyses.

Lane 1 (top) contains size strands of 1 Kb + DNA ladder (Invitrogen, UK), on top showing ITS2 PCR products for samples MAS11 to MAS26 and bottom MAS27 to MAS40 including plant DNA extract as a positive control followed by a negative control. Samples MAS12, MAS13 and MAS19 were chosen for further analyses. The PCR products were run on 2.0 % agarose gel for 50 min at 80 V.

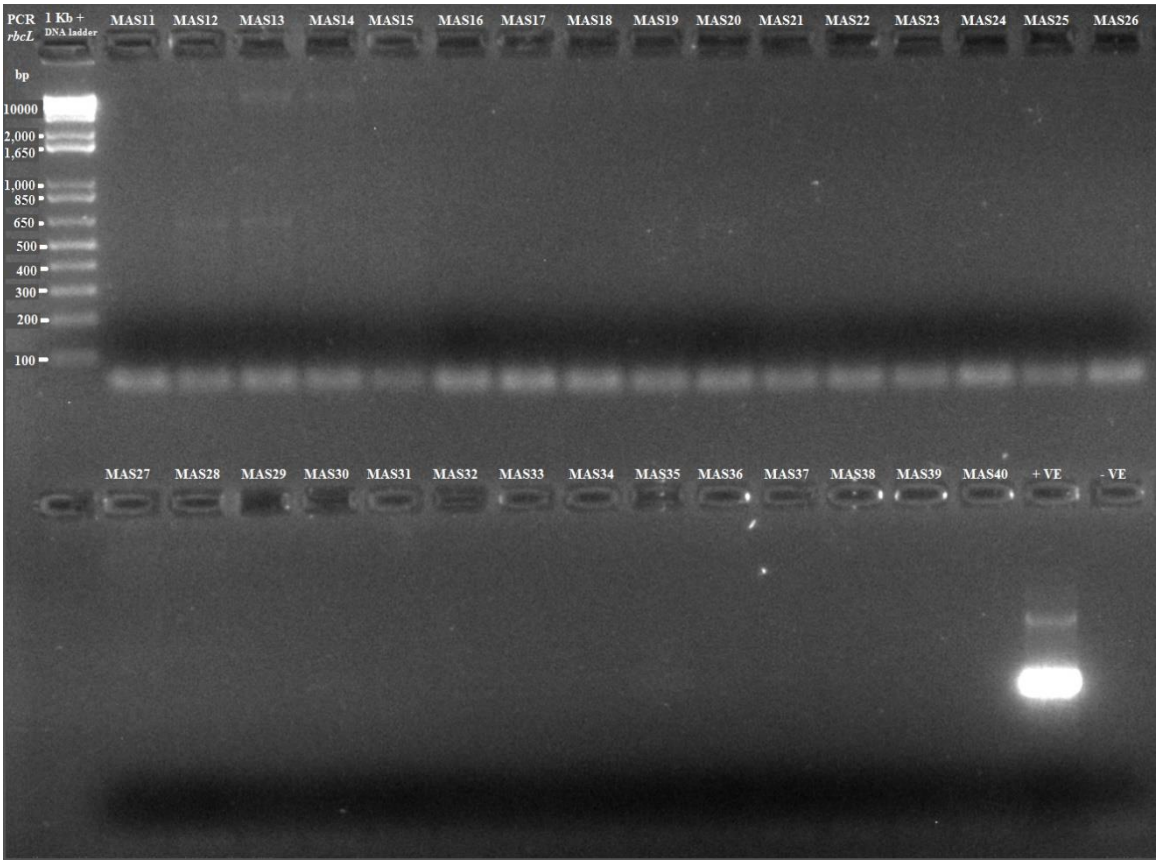


Figure 5.6 *rbcL* PCR products for metabarcoding soil analyses.

Lane 1 (top) contains size strands of 1 Kb + DNA ladder (Invitrogen, UK), on top showing *rbcL* PCR products for samples MAS11 to MAS26 and bottom MAS27 to MAS40 including plant DNA extract as a positive control followed by a negative control. Samples MAS12 and MAS13 were chosen for further analyses. The PCR products were run on 2.0 % agarose gel for 50 min at 80 V.

Amplified PCR products of *rbc*L and ITS2 samples were purified prior to submission for Illumina sequencing process using Illustra GFX PCR band purification kit (by GE Healthcare Life Sciences, UK) and quantified using Qubit dsDNA HS Assay Kit with Qubit® 2.0 Fluorometer. The Qubit quantification readings for the 12 metabarcoding samples including a negative control are listed in Table 5.7. Although the concentrations for *rbcL* and ITS2 amplicons were pooled from an equal column mix of the two PCR reactions prior to submission to Edinburgh Genomics lab., the fluorometer readings did not reflect that in the final readings of the equally pooled, *rbcL*+ITS2 concentration (Table 5.7). This might be due to a possible error in the lab during the preparation of the Qubit solutions, and sample preparations and / or variation of the pooled PCR reactions for each sample resulted in quantification error. Therefore, to avoid any further confusion, I have added readings generated by Edinburgh Genomics Lab. QC report using Agilent 2100 BioAnalyzer (Agilent Technologies) for the twelve metabarcoding samples present in Table 5.8. For the full QC report on metabarcoding samples generated by Edinburgh Genomics please refer to Appendix 5.1. Metabarcoding samples quantification for *rbcL* and ITS2 amplicons are illustrated in Figure 5.7 with the minimum recommendation of DNA concentration required by Edinburgh Genomics.

Table 5.7 Metabarcoding: eDNA samples amplified using two DNA barcoding markers (*rbcL* + ITS2)

Soil Sample ID	Depth of Soil Collection (cm)	*rbcL* (ng/ul)	ITS2 (ng/ul)	*rbcL*+ITS2 Conc (ng/ul)
Site 1: Fenced area				
MAS1	surface (0-5)	5.54	11.8	10.75
MAS2	surface (0-5)	19.1	5.66	11.25
MAS3	surface (0-5)	6.94	3.84	7.61
MAS4	surface (0-5)	17.7	3.16	9.37
MAS5	surface (0-5)	7.30	3.86	6.06
MAS6	surface (0-5)	41.8	12.3	17.0
MAS7	surface (0-5)	23.4	2.64	5.24
MAS8	surface (0-5)	14.6	2.24	7.49
MAS9	surface (0-5)	36.2	4.94	15.5
MAS12	deep (10-15)	3.02	6.72	7.36
MAS13	deep (10-15)	9.5	3.08	5.30
MAS19	deep (10-15)	2.46	3.10	5.77
Negative control	-	*	*	*

*Sample too low to measure with the Qubit® 2.0 Fluorometer (< 0.05 ng/µl)

PCR products for site 2 (open desert area) did not reach Edinburgh Genomics minimum requirements

Table 5.8. Metabarcoding samples quantification using Edinburgh genomics QC report

Soil Sample ID	Depth of Soil Collection (cm)	*rbcL* (ng/µl)	ITS2 (ng/µl)	*rbcL* + ITS2 (ng/µl)
Site 1: Fenced area				
MAS1	surface (0-5)	2.5	4.61	7.11
MAS2	surface (0-5)	2.77	2.84	5.61
MAS3	surface (0-5)	2.1	2.04	4.14
MAS4	surface (0-5)	5.02	2.58	7.6
MAS5	surface (0-5)	1.7	2	3.7
MAS6	surface (0-5)	5.27	6.06	11.33
MAS7	surface (0-5)	2	1.85	3.85
MAS8	surface (0-5)	3	1.94	4.94
MAS9	surface (0-5)	2.84	10	12.84
MAS12	deep (10-15)	0.5	5.3	5.8
MAS13	deep (10-15)	0.1	2.63	2.73
MAS19	deep (10-15)	0	4.74	4.74

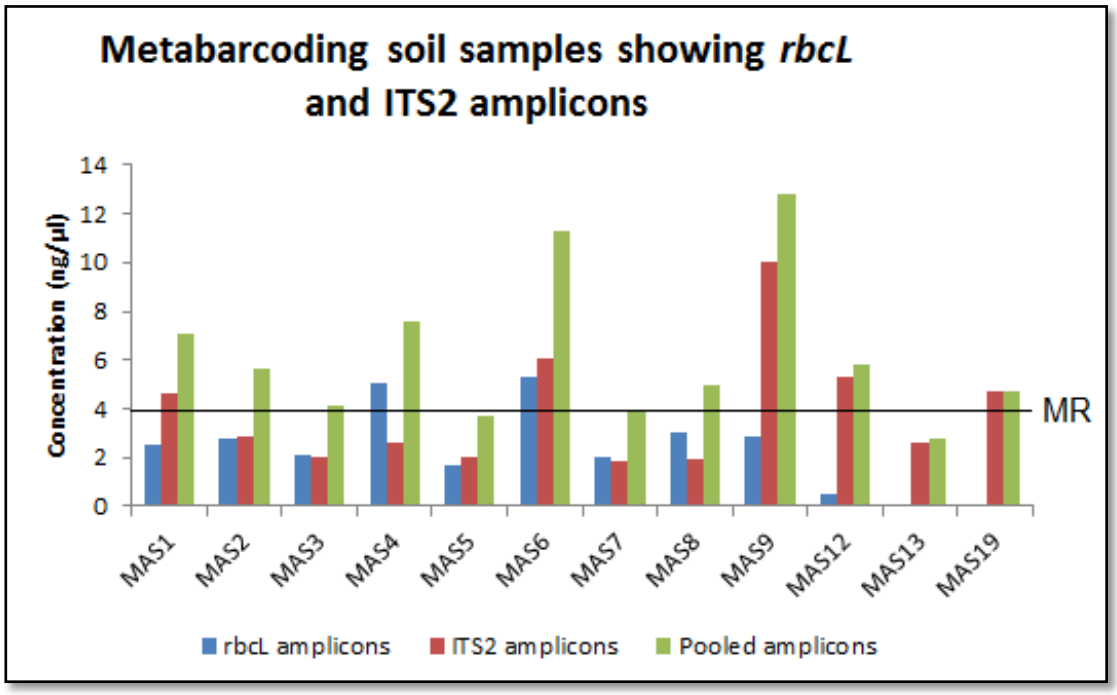


Figure 5.7 Metabarcoding soil samples showing quantification for *rbcL* and ITS2 amplicons based on Edinburgh Genomics QC report
MR- Minimum DNA concentration (4 ng/µl) for sequencing recommended by Edinburgh Genomics. Negative control was too low to measure (< 0.05 ng/µl).

The purified amplicons (*rbcL* + ITS2) were pooled in equal volumes (40 µl of each amplicon) instead of equimolar due to the base pair size variation and run on 2.0 % agarose gel for 50 min at 80 V represented in Figure 5.8. In the previous chapter, DNA barcoding the flora of Kuwait, the base pair length of the individuals for each of the two DNA markers varied in size, *rbcL* ranged from 532 to 600 bp and ITS2 from 354 to 417 base pair (see Chapter 4). Therefore, due to the variation in base pair lengths of both markers, molarity was not calculated; amplicons were pooled in equal volumes.

Library preparation and sequencing was performed on each sample containing the pooled amplicons (*rbcL* + ITS2) by Edinburgh Genomics. At Edinburgh Genomics, Illumina sequencing was performed using MiSeq Reagent v2 sequencing kits capable of producing paired-end sequencing (2 x 250 bp sequences). Individual libraries were generated per sample and high costs precluded the Illumina sequencing of the negative control.

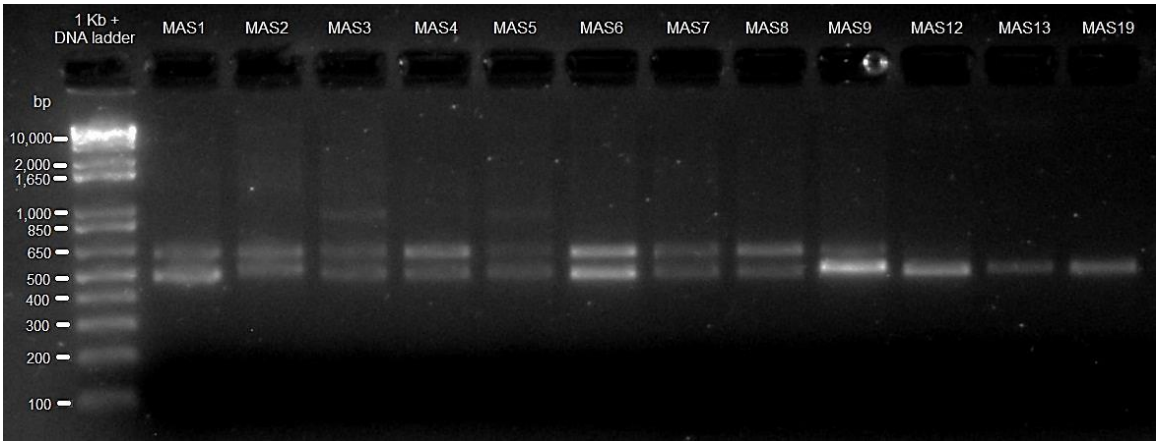


Figure 5.8 PCR amplicons (*rbcL* and ITS2) for Metabarcoding analysis run on 2 % agarose gel for 50 min at 80 V. Lane 1 contains size strands of 1 Kb + DNA ladder (Invitrogen, UK), Lanes 2-12 representing purified pooled amplicon samples, The bands represent the followings: MAS1 two bands *(rbc*L 650 bp; ITS2 500 bp), MAS2 two bands (*rbc*L 650 bp; ITS2 550 bp), MAS3 three bands (*rbc*L 650 bp, 950 bp; ITS2 500 bp), MAS4 two bands (*rbc*L 650 bp; ITS2 550 bp), MAS5 three bands (*rbc*L 650 bp, 950 bp; ITS2 500 bp), MAS6 two bands (*rbc*L 650 bp; ITS2 550 bp), MAS7 two bands *(rbc*L 650 bp; ITS2 500 bp), MAS8 two bands (*rbc*L 650 bp; ITS2 500 bp), MAS9 two bands (*rbc*L 650 bp; ITS2 500 bp), MAS12 two bands (fainted *rbc*L 650 bp; ITS2 500 bp), MAS13 and MAS19 each showing one band (ITS2 500 bp).

5.2.3 Next Generation Sequencing Bioinformatics

Bioinformatics metagenomics:

An initial assessment of quality scores and quality check across Illumina data was performed using FASTQC by Babraham Bioinformatics, (version 0.10.1) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequence assembly and alignment of overlapping reads was performed using Paired-end read merger (PEAR - ver. 0.8.1) Zhang et al. (2014) software used for trimming and assembling raw Illumina paired-end reads. Paired-end sequence reads were trimmed of low-quality data, and short reads were discarded. Given the low number of angiosperm sequence reads, further sequence assembly was not required.

FASTQ sequences were converted to FASTA for each sample, and paired-end reads were matched against three databases: complete plastid genomes database and Angiosperm plastid sequence database (both downloaded from the NCBI database, accessed on February 2016) and the Kuwaiti flora DNA database. Each sequence read was matched to the databases using BLASTn searches method (Altschul et al., 1990). The BLASTn matches were filtered using percentage sequence identity ≥ 98 % and mismatches < 1 %. An overview of the metagenomics sequence workflow is summarised in Figure 5.9.

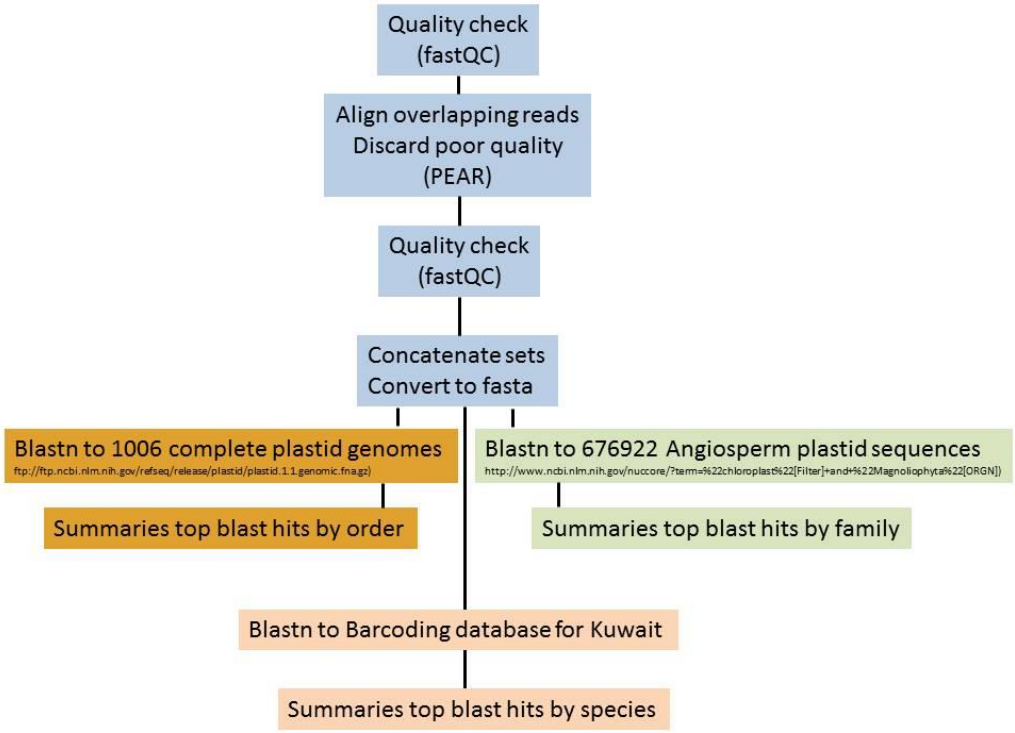


Figure 5.9 Workflow of Metagenomics sequences generated by Illumina platform

Bioinformatics metabarcoding:

For metabarcoding, an initial assessment of quality scores across the raw sequence data was performed using FASTQC by Babraham Bioinformatics (version 0.10.1) (http://www.bioinformatics.babraham.ac.uk/projects /fastqc/) providing a quick

overview and impression of whether the data contains any problems before proceeding to any further analysis. The removal of *rbcL* and ITS2 primers (Table 5.4) was performed using Cutadapt software v 1.4.2 (Martin, 2011). File conversion from FASTQ to FASTA used a python script from (BrianKnaus.com, 2009). Sequence assembly (forward and reverse reads) was performed using PEAR (PEAR software v 0.8.1 by Zhang et al. 2014) for paired-end reads to trim and align overlapping sequences followed by clustering in Qiime software (v 1.8.0, Edgar 2010). When paired-end reads did not overlap, forward and reverse reads were analysed separately. The PEAR assembled sequences were concatenated into a single file and the reads renamed to Qiime specific naming format. The *de novo* clustering of reads was performed using Uclust implemented in Qiime software using Atmosphere, CyVerse's cloud-computing platform (http://www.cyverse.org/atmosphere). Representative sequences from each *de novo* cluster with ≥ 100 reads were matched to the flora of Kuwait DNA database and NCBI-Angiosperm database using BLASTn with a cut-off value ≥ 99 % ID (Altschul et al., 1990).

5.2.4 Data accessibility

Raw sequences of metagenomics and metabarcoding received from Edinburgh Genomics were uploaded on the European Nucleotide Archive server (ENA: http://www.ebi.ac.uk/ena) for public accessibility and archived under the project accession numbers PRJEB12627 for metagenomics (Table 5.9) and PRJEB13939 for metabarcoding (Table 5.10).

In addition to the archiving of raw sequence at ENA, the data were also processed for metagenomics analyses by the European Bioinformatics Institute (EBI). The details of the pipeline used by EBI metagenomics for processing the raw sequences are documented by Hunter et al. (2014), and an overview of the pipeline is shown in Figure 5.10. Also, to access the metagenomics soil eDNA analyses and the summary for the taxonomic and functional analyses performed by EBI, please refer to the following web address which will direct you to the metagenomics project number ERP014120:
https://www.ebi.ac.uk/metagenomics/projects/ERP014120

The EBI pipeline predicts both rRNA coding and protein coding features. Predicted rRNAs are used for taxonomic analysis and predicted protein coding sequences (pCDS) are fed into the functional analysis steps (Figure 5.10). Ribosomal database project classifier (Cole et al., 2009) and the Greengenes reference database (DeSantis et al., 2006) were used for the classification of archaeal and bacterial species. In this chapter, I will discuss the taxonomic analysis generated by EBI pipeline. A table summarising phylum level taxonomy for the six metagenomics soil samples are represented in Appendix 5.2.

Table 5.9 Metagenomics eDNA samples archived by ENA

Soil Sample ID	ENA ID	Run ID
MA1	ERS1059009	ERR1260491
MA2	ERS1059010	ERR1260492
MA3	ERS1059011	ERR1260493
MA4	ERS1059012	ERR1260494
MA5	ERS1059013	ERR1260495
MA6	ERS1059014	ERR1260496

Table 5.10 Metabarcoding eDNA samples archived by ENA

Soil Sample ID	ENA ID	Run ID
MAS1	ERS1147245	ERR1406352
MAS2	ERS1147246	ERR1406353
MAS3	ERS1147247	ERR1406354
MAS4	ERS1147248	ERR1406355
MAS5	ERS1147249	ERR1406356
MAS6	ERS1147250	ERR1406357
MAS7	ERS1147251	ERR1406358
MAS8	ERS1147252	ERR1406359
MAS9	ERS1147253	ERR1406360
MAS12	ERS1147254	ERR1406361
MAS13	ERS1147255	ERR1406362
MAS19	ERS1147256	ERR1406363

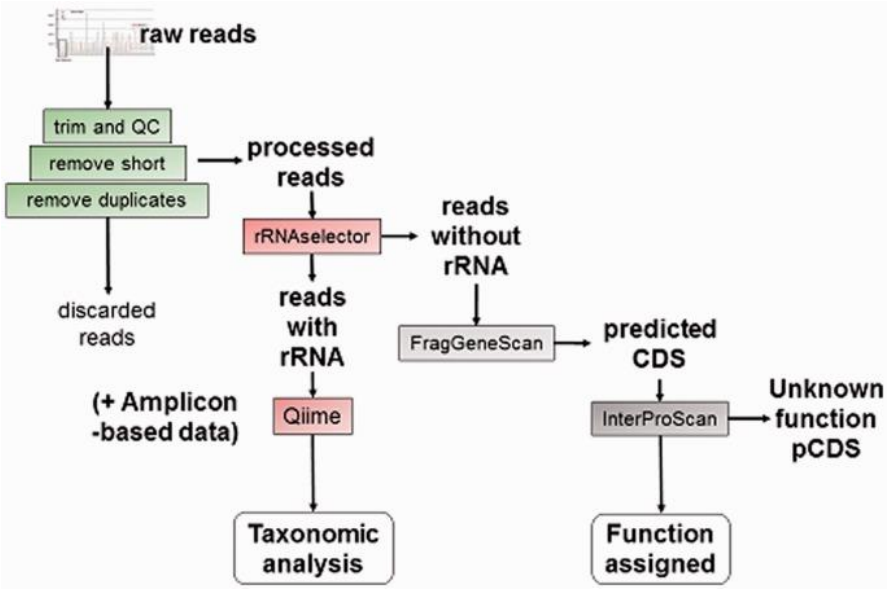


Figure 5.10 An overview of the pipeline used by European Bioinformatics Institute (EBI) metagenomics to process raw sequence files and predict the functions and taxa present in a given sample

5.3. Results

5.3.1 Plant diversity above ground

Surveying the vegetation above ground, site 1 (DMZ protected) represents rich plant diversity with 27 species compared to the disturbed site 2 with only 5 species present (Table 5.11). Plants listed in Table 5.11 will act as a reference checklist to compare current above-ground diversity of vascular plants with that found below ground. Comparing the two sites together it is clear that the open desert area (site 2) is exposed to grazers and other human disturbing activities, while the fenced area (site 1) is rich in plant diversity due to the protection over a period of 25 years avoiding any grazers and human interference.

Table 5.11. List of plant species present above ground in the two study sites at Um Neqa

Site/ Species present	Braun-Blanquet Cover Scale *
Site 1: Fence protected area	
Allium longisepalum Bertol.	1
Anisosciadium lanatum Boiss.	1
Asphodelus tenuifolius Cav.	1
Atractylis carduus (Forssk.) C.Chr.	1
Brassica tournefortii Gouan	2
Carduus pycnocephalus L.	1
Centaurea pseudosinaica Czerep.	+
Centaurea bruguierana (DC.) Hand.-Mazz.	1
Convolvulus oxyphyllus Boiss.	+
Cuscuta planiflora Ten.	1
Gypsophila capillaris (Forssk.) C.Chr.	2
Hammada salicornica (Moq.) Iljin.	4
Helianthemum lippii (L.) Dum.Cours.	1
Heliotropium bacciferum Forssk.	1
Koelpinia linearis Pall.	1
Launaea mucronata (Forssk.) Muschl.	1
Pennisetum divisum (Forssk. ex J.F.Gmel.) Henrard	2
Plantago boissieri Hausskn. & Bornm.	2
Plantago ovata Forssk..	1
Rhanterium epapposum Oliv.	4
Rumex vesicarius L.	1
Salvia aegyptiaca L.	1
Lomelosia olivieri (Coult.) Greuter & Burdet	1
Lomelosia palaestina (L.) Raf.	+ (rare plant)
Schismus barbatus (L.) Thell.	2
Senecio glaucus L.	2
Stipa capensis Thunb.	2
Site 2: Open desert area	
Hammada salicornica (Moq.) Iljin	+
Arnebia decumbens (Vent.) Coss. & Kralik	1
Astragalus schimperi Boiss.	1
Gymnarrhena micrantha Desf.	1
Moraea sisyrinchium (L.) Ker. Gawl.	1

***Braun-Blanquet Cover Scale** (+) sparsely present; cover very small, (1) plentiful, but of small cover value, (2) cover 5-20 % (3) cover 25-50 % (4) cover 50-75 % (5) Cover greater than 75 % (Wikum & Shanholtzer, 1978).

5.3.2 Comparison of eDNA quantification

5.3.2.1 Metagenomics eDNA yield quantification

All metagenomics eDNA extractions showed high molecular weight DNA with no visible low-weight DNA (Figure 5.3). Comparing DNA yield (Table 5.3) with the gel image representing eDNA soil extractions samples MA1-MA8 (Figure 5.3), the highest DNA yield was retrieved by Samples MA1 (2.9 µg), MA2 (2.2 µg) and MA6 (2.2 µg) shown on the gel image with high density. The lowest DNA yield resulted for soil samples MA7 (0.2 µg) and MA8 (0.1 µg) shown on the gel image with very low density (Table 5.2 and Figure 5.3).

To investigate the relationship of the amount of DNA yield recovered for each metagenomics sample with the number of trimmed and assembled sequence reads (Table 5.12) a graph illustrating the relation of the two factors are shown in Figure 5.11. There was no relationship between the number of sequences and the DNA yield, samples MA1 and MA2 with the highest DNA yield (2.9 and 2.2 µg, respectively) generated the lowest number of sequence reads (986,161 and 424,097, respectively) (Figure 5.11). Furthermore, samples MA3. MA4 and MA5 representing the lowest DNA yields generated the highest number of sequences (Table 5.12 and Figure 5.11)

Samples MA7 and MA8 did not qualify for Illumina sequencing and generated the lowest DNA yields 0.2 µg and 0.1 µg, respectively. Therefore, Edinburgh Genomics did not recommend sequencing for either MA7 or MA8 samples.

Table 5.12 Summary of eDNA metagenomic sequence reads and DNA yield representing two study sites

eDNA samples	Depth of Soil sampling (cm)	Number of trimmed, assembled sequence reads	DNA yield (µg)
Site 1			
MA1	surface (0-5)	986,161	2.9
MA2	surface (0-5)	424,097	2.2
MA3	deep (10-15)	2,909,966	1.3
MA4	deep (10-15)	2,851,722	1.0
Site 2			
MA5	surface (0-5)	2,734,219	1.1
MA6	surface (0-5)	2,812,619	2.2

Site 1: Fenced area; Site 2: open desert area

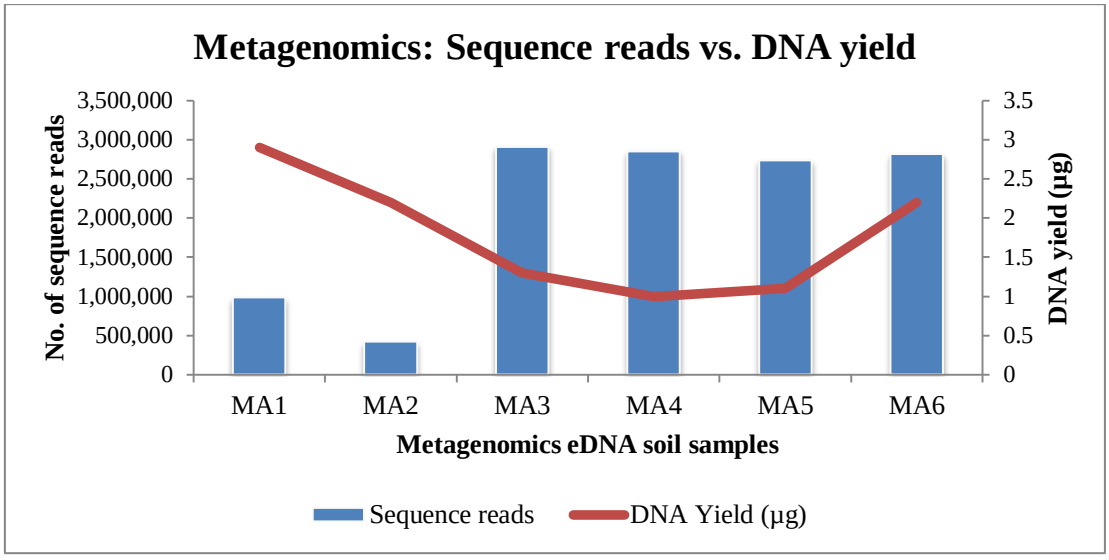


Figure 5.11 Metagenomics analysis illustrating the relation between sequence reads and DNA yields for six eDNA soil samples

5.3.2.2 Metabarcoding eDNA amplicon quantification

Metabacroding samples MAS1-MAS9 generated high molecular weight bands for *rbcL* (c. 600-650 bp) and ITS2 (c. 500-550 bp) amplicons compared with the 1 Kb + ladder (Figure 5.4). Samples MAS12 and MAS13 showed a very light band for *rbcL* amplicons while MAS19 was only represented by ITS2 amplicons (Figure 5.5 and Figure 5.6). Samples MAS3 and MAS5 are showing a third band at approximately 1,000 bp which is representing *rbcL* amplicon and showing on both gels, the PCR gel and the pooled amplicons gel (Figure 5.4 and Figure 5.8). Also, the visibility of a light positive band in the negative control showing in Figure 5.4 and Figure 5.5 indicates that there was some form of contamination. This indicates that any step of the PCR procedure and/ or the preparation of the master mix and/ or the stock reagents used could be contaminated. In case the source of contamination is not from cross contamination of the metabarcoding samples and the DNA source comes from plants floating around the lab and not present in the DNA reference library of the Kuwaiti flora, the issue will be minimised since the metabarcoding samples will only result in matches to plant sequences present in the Kuwaiti DNA database.

The amplicons, representing *rbcL* and ITS2, were pooled in equal volumes instead of equimolar due to the variation in bp size of the PCR products molarity was not calculated. Samples MAS9 and MAS6 generated the highest quantification of pooled amplicon concentration,12.84 ng/µl and 11.3 ng/µl, respectively (Table 5.7; Figures 5.8). Samples MAS5 and MAS7 represented the lowest quantification of pooled amplicon concentration, 3.7 ng/µl and 3.85 ng/µl, respectively (Table 5.7; Figures 5.8).

Evaluating amplicon concentration/ primer, ITS2 generated the highest concentration for MAS6 (6.06 ng/µl) and the lowest was *rbcL* for MAS12 and MAS13 (0.5 and 0.1 ng/µl, respectively) (Table 5.7; Figures 5.4-5.6 and Figure 5.8).

The comparison of the pooled amplicons concentration with the number of trimmed and assembled sequences showed that there is no match of sequences generated per sample compared with amplicons concentration, e.g. sample MAS13 with the lowest

amplicon concentration (2.73 ng/µl) generated more sequence reads (1,131,353) than MAS6 with high amplicon concentration (11.3 ng/µl) which generated only 891,348 sequence reads (Table 5.13).

Table 5.13 Summary of metabarcoding sequences and DNA yield representing 12 samples

Soil Sample ID	Trimmed, assembled sequence reads	*rbcL*+ITS2 Conc (ng/µl)
Site 1		
MAS1	716,735	7.11
MAS2	1,177,465	5.61
MAS3	708,364	4.14
MAS4	932,518	7.6
MAS5	752,954	3.7
MAS6	891,348	11.33
MAS7	993,762	3.85
MAS8	1,038,804	4.94
MAS9	909,167	12.84
MAS12	671,093	5.8
MAS13	1,131,353	2.73
MAS19	662,429	4.74

5.3.3 Plant diversity below ground using NGS technologies

The following sections of the results will be based on BLASTn searches using Genbank database (NCBI Resource Coordinators, 2016) and the Kuwaiti DNA library reference generated in chapter 4. I will first explore the sequenced data generated from metagenomics analyses, followed by PCR-based metabarcoding analyses.

Metagenomics analyses

5.3.3.1 High order level analyses using whole plastid genomes

In an attempt to study the diversity of green plant material present in the eDNA soil samples, a BLASTn of sequence reads was carried out against a complete plastid genome database. Of the top blast hits to a complete plastid genome database at order

level, (blast matches set at ≥ 98 % identity and mismatches < 1 % were excluded), the largest proportion of the 0.1 % of the total number of sequences in the datasets of green plants matches across all metagenomics samples (MA1 to MA6) are represented by algae with 81 % (8,931 sequences), followed by 12 % to bryophytes (1,252 sequences), 4 % to ferns (421 sequences) and only 3 % to flowering plants (372 sequences) (Tables 5.14 and Figure 5.12).

Metagenomics samples MA3 and MA4 from site 1 resulted in the greatest numbers of sequence matches to complete plastid genomes at order level, 2,435 and 2,410, respectively (Table 5.14). The lowest number of matches was represented by MA2 with 942 sequences only (Table 5.14).

Algae sequence reads resulted in the highest number of matches in the green plant complete plastid genome database with 42 orders (Table 5.14) represented by green algae (Chlamydomonadales and Chlorellales) and red algae (Cyanidiales and Bangiales). Bryophytes covered the greatest percentage of sequence matches (12 %) after algae (81 %) and are only represented in the data set by two orders: Orthotrichales (mosses) and Pelliales (liverworts). Ferns sequence matched to only 4 % of the data set and are represented by 3 orders: Polypodiales (the largest order of ferns), Isoetales (quillworts) and Schizaeales. Although magnoliophytes sequence reads matches are showing the lowest percentage (3 %) of the data set (Table 5.14), it matched to 6 major orders of the flora of Kuwait belonging to Asparagales, Caryophyllales, Cupressales, Fabales, Geraniales and Lamiales.

In an attempt to study the differences across the two soil depths (0-5 and 10-15 cm), for site 1 (MA1-MA4 are representatives of soil collected from site 1) the soil samples MA3 and MA4 (collected from 10-15 cm depth) resulted in greatest sequence matches compared to the upper surface of the soil, represented by samples MA1 and MA2 (Table 5.14). For site 2 (represented by samples MA5-MA8) it was not possible to compare amongst the two depths since samples MA7 and MA8 were not sequenced for not passing the minimum requirement recommended by Edinburgh Genomics. Comparing only the upper layer of soil collection depth (0-5 cm) across the two collection sites, samples MA5 and MA6 from site 2 (open desert area) showed very

high numbers of sequence matches compared to site 1 (fence protected area) represented by samples MA1 and MA2 (Table 5.14).

Table 5.14 Raw number of sequences with blast hit to a complete plastid genome at Order level (matches ≥ 98 % ID; mismatches < 1 % were excluded)

		Site 1				Site 2			
	No. of Order level matches	**MA1**	**MA2**	**MA3**	**MA4**	**MA5**	**MA6**	**Total No. Seq**	**Total (%)**
Soil sample collection depth (cm)		0-5	0-5	10-15	10-15	0-5	0-5		
Magnoliophyta	6	44	28	81	79	82	58	**372**	**3 %**
Ferns	3	44	26	90	87	105	69	**421**	**4 %**
Bryophytes	2	166	126	274	252	214	220	**1252**	**12 %**
Algae	42	1018	762	1990	1992	1534	1635	**8931**	**81 %**
Total number of green plant sequences (%)	**53**	**1272 (0.13)**	**942 (0.22)**	**2435 (0.1)**	**2410 (0.1)**	**1935 (0.1)**	**1982 (0.1)**	**10976 (0.1)**	
Total number of raw sequences		**986161**	**424097**	**2909966**	**2851722**	**2734219**	**2812619**	**12718784**	

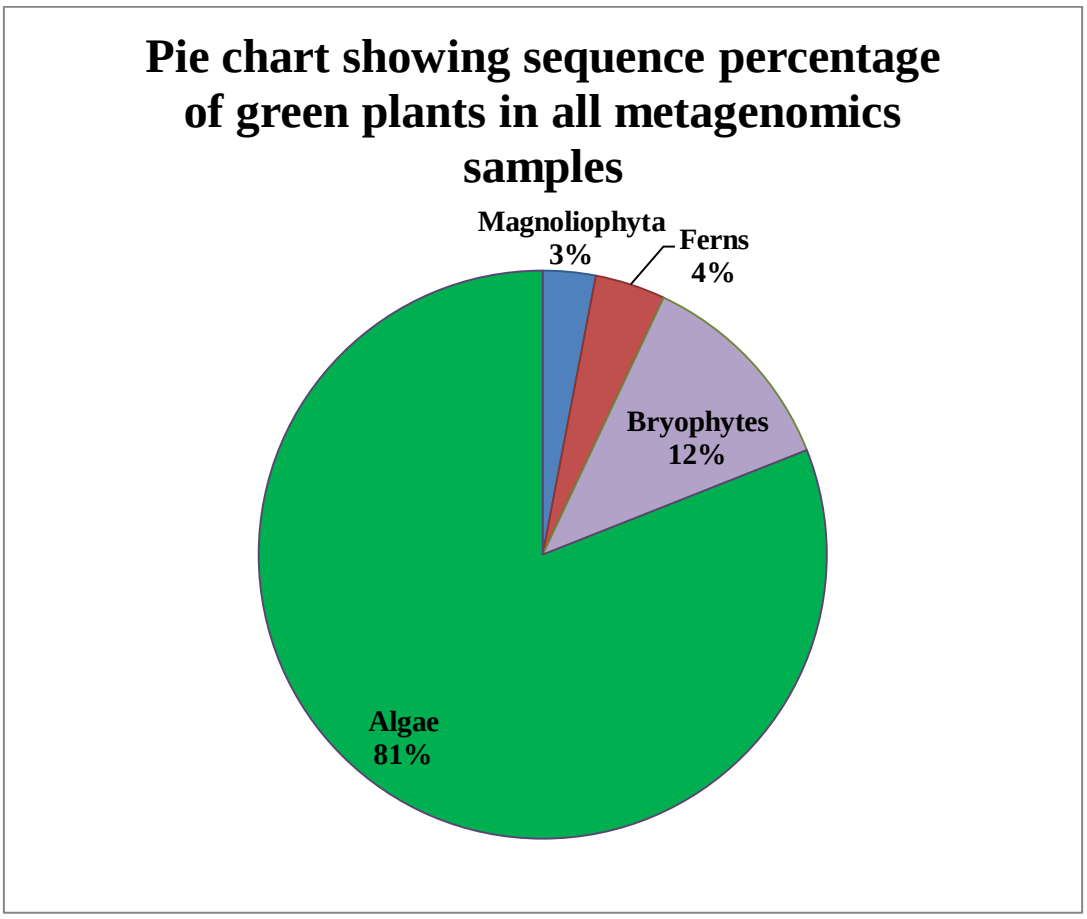


Figure 5.12 Percentage of green plants present in all metagenomic samples. BLASTn matches ≥ 98% ID to a complete plastid genome database at Order level

5.3.3.2 Family-level analyses using Angiosperm-NCBI plastid sequence database

In an attempt to show finer scale analyses, reads which matched the green plant plastid sequences were blasted to NCBI Angiosperm plastid database to identify the reads at the family level. The sequence matches set at ≥ 98 % ID and mismatches < 1 % were excluded. A summary of Blast hits to family level using plastid sequence database are presented in Table 5.15, single asterisks indicating families found in the flora of Kuwait.

The highest number of sequence matches at the family level was generated by samples MA6 (1,996 sequences) followed by MA3 (1,545) and the least was MA1 and MA2, 668 and 238 sequences, respectively (Table 5.15).

Several sequences matched ordinary families of the flora of Kuwait with sequence matches set at ≥ 99 % ID and mismatches < 1 % were excluded.and are representatives of Fabaceae, Caryophyllaceae, Convolvulaceae, Cucurbitaceae, Geraniaceae, Orobanchaceae, Poaceae and Verbenaceae (Table 5.14 and Table 2.3). Other families with a high percentage of sequence matches to family level and not familiar to the flora of Kuwait are represented by Cactaceae, Campanulaceae, Ericaceae, Platanaceae and Triuridaceae (Table 5.15).

Comparing the differences between the two soil collection depths, for samples MA1 to MA4 there was variation amongst soil sampled from the surface layer (0-5 cm) and soil sampled from the deep layer (10-15 cm). The top surface layer showed slightly higher percentage (55 - 68 %) of sequence matches than the deep soil layer (42 – 44 %) which indicates that, as expected, more plant DNA material is deposited in the top surface layer (Table 5.15).

Table 5.15 Percentage of Sequence blast hit to a complete plastid genome at Family level (Matches ≥ 98 % to Angiosperm-NCBI database; mismatches < 1 % were excluded)

		Metagenomics samples Percentage of sequence blast hit					
				Site 1		Site 2	
		MA1	MA2	MA3	MA4	MA5	MA6
Soil sample collection depth (cm)		0-5	0-5	10-15	10-15	0-5	0-5
Family	Prevalence in the flora of Kuwait						
Fabaceae *	Common, wide spread	11.0	9.0	7.4	6.6	6.9	9.8
Convolvulaceae *	Common, moderate spread	6.7	4.7	4.0	3.8	2.3	6.0
Poaceae *	Common, wide spread	6.3	4.5	3.4	4.0	2.9	5.3
Ericaceae	Limited to farms	5.9	5.4	2.4	2.8	2.2	4.1
Caryophyllaceae *	Common, wide spread	4.7	3.1	3.3	3.2	3.6	3.7
Triuridaceae	Uncommon	4.5	3.3	2.9	2.9	2.7	5.0
Geraniaceae *	Common, moderate spread	4.2	3.8	3.8	3.3	2.6	3.8
Cactaceae	Limited to farms	3.7	2.8	1.7	1.7	1.6	1.5
Platanaceae	Uncommon	3.4	4.5	3.0	2.4	1.9	3.4
Campanulaceae	Limited to farms	2.9	2.6	1.8	2.0	1.9	2.4
Musaceae	Limited to farms	2.4	1.7	1.4	1.2	1.6	2.3
Orobanchaceae *	Common, wide spread	2.3	0.9	2.2	2.3	2.2	2.6
Orchidaceae	Limited to farms	1.8	1.4	2.1	2.0	1.7	2.1
Chloranthaceae	Uncommon	1.7	1.4	0.4	0.4	0.4	1.1
Santalaceae	Uncommon	1.6	1.4	0.9	0.7	0.5	1.6
Verbenaceae *	Limited to farms	1.4	0.2	0.9	0.7	0.7	1.2
Viscaceae	Uncommon	1.1	1.4	1.0	0.9	0.5	1.5
Schisandraceae	Uncommon	1.1	1.2	0.9	0.8	0.3	1.1
Cucurbitaceae *	Moderate spread 1 spp. only	1.0	1.2	0.3	0.2	0.1	0.3
Total percentage		68	55	44	42	37	59
Total number of sequences		986161	424097	2909966	2851722	2734219	2812619
Total number of sequences with blast hit to NCBI database mismatches < 1 % were excluded		668	238	1180	1116	1017	1996

Asterisk (*) indicates families matching the flora of Kuwait

5.3.3.3. Species-level analyses using Angiosperm-NCBI plastid sequence database

The following analysis shows matches to species-level performed on metagenomics samples represented in the Angiosperm-NCBI database and later compared with the results of blast matches against the Kuwaiti DNA database. A total number of 20 species matched across all metagenomics samples (Table 5.16). The highest percentage of sequences matched across all samples is represented by *Cuscuta gronovii* (parasitic plant) (Table 5.16). The only species match to Angiosperm-NCBI database representing the flora of Kuwait is *Silene conica* found in samples MA3, MA4 and MA5. Soil eDNA metagenomics species-level matches to Angiosperm-NCBI plastid sequences are listed in Table 5.16.

The following is a summary of matches representing Table 5.16, organised by highest to lowest percentage of sequence matches:

MA1, a total of 48 % sequence matches, represented mainly by *Cuscuta gronovii*, followed by *Sciaphila densiflora, Silene conica* and *Enkianthus perulatus*.
MA6, a total of 39 % sequence matches mainly represented by *Cuscuta gronovii, Sciaphila densiflora, Platanus occidentalis* and *Enkianthus perulatus.*
MA2, a total of 38 % sequence matches, represented by *Cuscuta gronovii, Platanus occidentalis* and *Enkianthus perulatus.*
MA3, a total of 27 % sequence matches, mainly represented by *Cuscuta gronovii, Platanus occidentalis, Sciaphila densiflora* and *Silene conica.*
MA4 and **MA5,** a total of 21 % and 22 %, respectively, mainly represented by *Sciaphila densiflora, Silene conica,* and *Cuscuta gronovii.*

Comparing the differences between the two soil collection depths, for site 1: samples MA1 to MA4, the top surface layer (0-5 cm) showed almost double the percentage of sequences (38-48 %) of the blast hit than the deep soil layer (21 - 27 %). The results of species level are similar to that of the blast hit results at the family level, which also indicates that more DNA material are deposited in the top surface layer of the soil (Table 5.16).

Table 5.16 Percentage of sequence blast hits to NCBI-Angiosperm plastid sequences at species level (the number of raw sequence reads matched to the Genbank database were converted to percentage) BLASTn cut-off value ≥ 98 % ID.

		Metagenomics samples Percentage of sequence blast hit to NCBI database					
		Site 1				**Site 2**	
		MA1	**MA2**	**MA3**	**MA4**	**MA5**	**MA6**
Soil sample collection depth (cm)		0-5	0-5	10-15	10-15	0-5	0-5
Family	**Species**						
Convolvulaceae *	*Cuscuta gronovii **	7	5	4	2	2	6
Triuridaceae	*Sciaphila densiflora*	4	3	3	3	3	5
Ericaceae	*Enkianthus perulatus*	4	4	2	2	2	3
Caryophyllaceae *	*Silene conica **	4	2	2	2	2	2
Platanaceae	*Platanus occidentalis*	3	4	3	2	2	3
Fabaceae *	*Pisum sativum*	3	1	2	2	2	2
Poaceae *	*Hordeum vulgare **	2	2	1	1	1	2
Geraniaceae *	*Erodium carvifolium **	2	2	2	1	1	2
Fabaceae *	*Medicago truncatula **	2	2	1	1	1	2
Campanulaceae	*Trachelium caeruleum*	2	2	1	1	1	2
Musaceae	*Musa acuminata*	2	1	1	1	1	2
Chloranthaceae	*Chloranthus spicatus*	2	1	0	0	0	1
Santalaceae	*Osyris alba*	2	1	1	0	1	2
Poaceae *	*Oryza sativa*	1	2	0	0	0	1
Verbenaceae *	*Lippia sidoides*	1	0	1	0	1	1
Fabaceae *	*Astragalus nakaianus **	1	1	1	0	0	1
Viscaceae	*Viscum album*	1	1	1	0	1	1
Cactaceae	*Carnegiea gigantea*	1	1	0	0	0	0
Orobanchaceae *	*Orobanche purpurea **	1	0	1	1	1	1
Cucurbitaceae *	*Cucumis melo*	1	1	0	0	0	0
Total percentage		48	38	27	21	22	39
Total number of sequences		**986161**	**424097**	**2909966**	**2851722**	**2734219**	**2812619**

Asterisk (*) indicates family, genus and/ or species present in the flora of Kuwait

5.3.3.4 Metagenomics analyses using Kuwait DNA database

BLAST of trimmed and assembled reads of the six eDNA samples against the DNA database of Kuwait generated using *rbc*L and ITS2 barcodes assembled in chapter 4 resulted in a total match to 25 accessions belonging to 24 species (Table 5.17). Only ITS2 sequences returned with matches to the six samples and none for *rbcL*. Although several attempts were made to blast metagenomic sequences against the *rbcL* barcodes of the Kuwaiti DNA database, none returned with any sequence matches indicating that sequences are either not represented in the samples or the DNA region is highly degraded. Furthermore, ITS2 sequences were poorly represented in the data sets with a very low number of sequence matches varied from 1 to 9 sequences only per accession (Table 5.17).

The following will be a summary of findings representing Table 5.17, arranged by high to a low number of sequence matches:
MA6 resulted in the greatest sequence matches to 38 sequences, represented by 13 species (*Trigonella stellata, Trigonella anguina, Erodium cicutarium, Erodium laciniatum, Spergularia marina, Tribulus terrestris, Zygophyllum qatarense, Ducrosia anethifolia, Emex spinosus, Filago pyramidata, Monsonia nivea, Picris babylonica,* and *Silene arenosa*).
MA4 matched 11 sequences represented by 6 species (*Astragalus sieberi, Erodium cicutarium, Linaria simplex, Polycarpon tetraphyllum, Salvadora persica*, and *Spergularia marina*).
MA3 matched 8 sequences represented by 7 species (*Cressa cretica, Gymnarrhena micrantha, Leptaleum filifolium, Ochradenus baccatus, Polycarpon tetraphyllum, Salvadora persica*, and *Spergularia marina*).
MA1 matched 6 sequences represented by 4 species (*Loeflingia hispanica, Sisymbrium orientale, Spergularia marina* and *Trigonella stellate*).
MA5 matched 4 sequences represented by 4 species (*Ducrosia anethifolia, Spergularia marina, Suaeda aegyptiaca* and *Trigonella anguina*).
MA2 matched only 1 sequence and represented by *Loeflingia hispanica.*

The highest match to sequences and present across most samples (except MA2) is represented by *Spergularia marina* (Table 5.17). Only 1 species matched with the above ground plants (Table 5.11) and represented by *Gymnarrhena micrantha* for soil collected from site 1 represented by sample MA3 (Table 5.17).

Comparing the differences between the two soil collection depths for site 1, samples MA1 (collected from 0-5 cm) and MA3 (10-15 cm) resulted in the same number of total sequence matches, 6 and 8 sequences, respectively, indicating that DNA material present at the top surface layer is almost equal to that collected from a deeper layer (Table 5.17). On the other hand, sample MA2 with soil sampled from the surface layer showed very low sequence match (only 1 sequence) when compared to its deeper sample, MA4 with 11 sequence matches (Table 5.17).

In the open desert area (site 2) sample MA6 (collected from the surface layer 0-5 cm) showed a high number of sequence matches (38 sequences), unlike MA5, also collected from the surface layer, only represented by 4 sequences (Table 5.17). Sample MA6 will further be investigated using Principal Component Analysis (PCA) method to understand the variation of sequence matches amongst the metagenomics samples.

Table 5.17 Raw number of sequences BLAST hit to the Kuwaiti flora DNA database.
BLAST cut-off value ≥ 98 % ID

	Metagenomics samples Raw numbers of sequence matches					
	Site 1				**Site 2**	
	MA1	**MA2**	**MA3**	**MA4**	**MA5**	**MA6**
Soil sample collection depth (cm)	0-5	0-5	10-15	10-15	0-5	0-5
Species match to accessions from Kuwait DNA database						
Astragalus_sieberi_13-0033616_ITS	0	0	0	1	0	0
Cressa_cretica_15-0042769_ITS	0	0	1	0	0	0
Ducrosia_anethifolia_15-0042889_ITS	0	0	0	0	1	2
Emex_spinosus_15-0042807_ITS	0	0	0	0	0	1
Erodium_cicutarium_15-0042925_ITS	0	0	0	1	0	7
Erodium_laciniatum_15-0042637_ITS	0	0	0	0	0	4
Filago_pyramidata_15-0043060_ITS	0	0	0	0	0	1
Gymnarrhena_micrantha_15-0042372_ITS **	0	0	1	0	0	0
Leptaleum_filifolium_15-0043023_ITS	0	0	1	0	0	0
Linaria_simplex_15-0042585_ITS	0	0	0	1	0	0
Loeflingia_hispanica_15-0042956_ITS	1	1	0	0	0	0
Monsonia_nivea_15-0042939_ITS	0	0	0	0	0	2
Ochradenus_baccatus_15-0042944_ITS	0	0	1	0	0	0
Picris_babylonica_15-0043086_ITS	0	0	0	0	0	2
Polycarpon_tetraphyllum_15-0042739_ITS	0	0	1	1	0	0
Salvadora_persica_15-0042887_ITS	0	0	2	1	0	0
Silene_arenosa_15-0043008_ITS	0	0	0	0	0	1
Sisymbrium_orientale_15-0042352_ITS	1	0	0	0	0	0
Spergularia_marina_15-0042815_ITS	2	0	1	6	1	3
Suaeda_aegyptiaca_15-0043045_ITS	0	0	0	0	1	0
Tribulus_terrestris_15-0042834_ITS	0	0	0	0	0	1
Trigonella_anguina_15-0043007_ITS	0	0	0	0	1	1
Trigonella_anguina_15-0043013_ITS	0	0	0	0	0	3
Trigonella_stellata_15-0043043_ITS	2	0	0	0	0	9
Zygophyllum_qatarense_15-0042905_ITS	0	0	0	0	0	1
Total number of sequence matches	6	1	8	11	4	38

(**) Double asterisks represent plant species present above ground during field survey

5.3.3.5 PCA analyses

In an attempt to study the correlations amongst the six metagenomics samples using information from blast match of sequences to the DNA database of the Kuwaiti flora, I used Principal Component Analysis (PCA).

PCA showed that sites 1, 2 and 4 are most similar, with sites 6, 3 and 5 distant to that group and to each other. (Table 5.18; Figures 5.13 and 5.14).

Table 5.18 PCA analysis showing correlated matrix of sequence matches to Kuwait DNA database for six eDNA soil samples

Importance of components:	PC1	PC2	PC3	PC4	PC5	PC6
Standard deviation	6.18	4.70	3.60	2.93	1.80	8.27E-15
Proportion of Variance	0.45	0.26	0.15	0.10	0.04	0.00E+00
Cumulative Proportion	0.45	0.71	0.86	0.96	1.00	1.00E+00

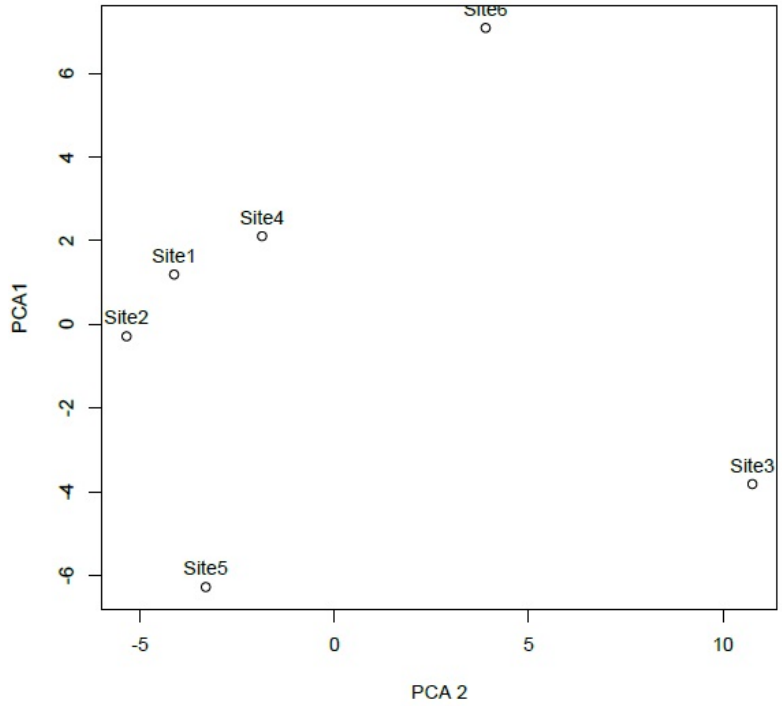


Figure 5.13 Principal component analyses, PCA1 against PCA2 of six eDNA soil samples against DNA barcode database of the flora of Kuwait (match with a score ≥ 100)

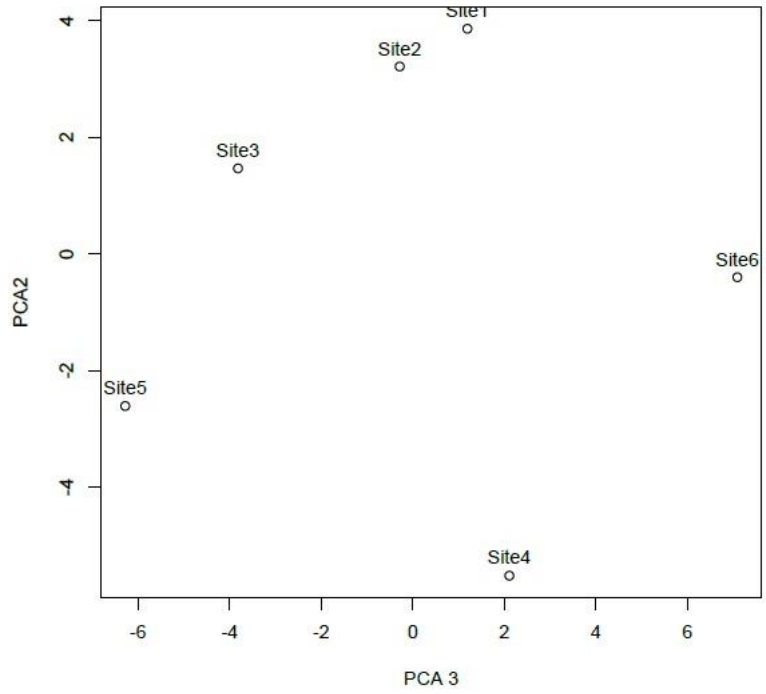


Figure 5.14 Principal component analyses, PCA2 against PCA3 of six eDNA soil samples against DNA barcode database of the flora of Kuwait (match with a score ≥ 100)

Metabarcoding analyses

5.3.3.6 Sequence reads for PCR-based metabarcoding samples

PCR amplification only worked for a total of 12 samples, using *rbcL* and ITS2 markers, out of the 40 soil eDNA samples (Table 5.8 and Figures 5.4-5.6). The samples are representatives of total DNA extracted from soil and amplified from site 1 only (fence protected area), none of the samples collected from site 2 (degraded desert area) produced any PCR products. The amplified samples are mainly representatives of soil eDNA collected from the upper soil layer, 0-5 cm (MAS1-MAS9) and only 3 samples are representing samples collected from the deeper soil layer, 10-15 cm (MAS12, MAS13 and MAS19) (Table 5.8 and Figures 5.4-5.6).

For metabarcoding raw sequences, an initial assessment of sequence quality was performed using FASTQC, which provides a quick overview and impression of whether the data contains any problems before proceeding to any further analysis. Sequence assembly was performed using PEAR software for paired-end reads to trim and align overlapping sequences (full details present in section 5.2.3).

Illumina sequencing produced a total number of 10,585,992 sequences of *rbcL* and ITS2 amplicons recovered across all 12 metabarcoding samples. The greatest number of sequence reads recovered by sample MAS2 (1,177,465) and the lowest was MAS19 (662,429) (Table 5.19). The percentage of total sequence length representing high-quality sequences of 473 bp length or better across all samples varied from 14 % (MAS5 and MAS7) to 19 % (MAS2, MAS6 and MAS9). The percentage of sequence length representing sequences of 250 bp length or better across all samples varied from 59 % (MAS19) to 68 % (MAS2, MAS4, and MAS8) (Table 5.19).

The *de novo* clustering of reads was performed using Uclust implemented in Qiime software using Atmosphere, CyVerse's cloud-computing platform. Uclust generated the greatest number of cluster counts for sample MAS6 with 26,161 clusters and the lowest for sample MAS19 with 11,191 (Table 5.19).

Table 5.19 Summary for metabarcoding samples showing representative sequences for the FASTA and de novo cluster counts

	Metabarcoding samples											
	MAS1	**MAS2**	**MAS3**	**MAS4**	**MAS5**	**MAS6**	**MAS7**	**MAS8**	**MAS9**	**MAS12**	**MAS13**	**MAS19**
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
FASTA sequence reads summary:												
Total length of sequences (bp)	217400379	290110239	198011657	237674808	208382092	214205228	268624645	271869941	207098128	206329677	271047175	176863223
Total number of sequences	716735	1177465	708364	932518	752954	891348	993762	1038804	909167	671093	1131353	662429
25 % of total sequence length ≥ 473 bp	114090 15 %	223580 19 %	103663 15 %	160280 17 %	108943 14 %	167031 19 %	141420 14 %	163198 16 %	172858 19 %	106092 16 %	183144 16 %	91412 14 %
50 % of total sequence length ≥ 250 bp	244288 34 %	513691 44 %	268594 38 %	397955 43 %	280844 37 %	381236 43 %	386239 39 %	435068 42 %	379956 42 %	213935 32 %	454192 40 %	216665 33 %
75 % of total sequence length ≥ 250 bp	461688 64 %	803801 68 %	466606 66 %	635629 68 %	489226 65 %	595442 67 %	654864 66 %	706938 68 %	589765 65 %	385281 60 %	726027 64 %	393529 59 %
Total GC count (bp)	103160638	129962547	95405451	110743239	100644816	104785239	120763685	126026424	113162172	114288767	148592069	103893701
GC (%)	47.45	44.8	48.18	46.59	48.3	48.92	44.96	46.36	54.64	55.39	54.82	58.74
BLASTn sequence matches to Kuwaiti DNA database:												
ITS2 total no. of sequences	141548	84738	110244	79220	125021	246453	89911	86164	462836	294326	484646	393200
rbcL total no. of sequences	230642	179075	222800	382783	237608	548782	258601	382037	109868	93016	47392	0
***de novo* sequence cluster counts:**	20234	22971	21513	14878	21832	26161	25014	20839	15659	19478	23178	11191

5.3.3.7 Metabarcoding analyses using Kuwaiti DNA database

BLASTn searches of MiSeq metabarcoding data using *de novo* clustered sequences conducted against the DNA database of Kuwait comprising of *rbcL* and ITS2 for all 12 samples resulted in a total match to 206 accessions representing 139 species with the number of reads in the *de novo* clusters varied from 100 to 127,629 reads (Table 5.19). The BLASTn matches were filtered using percentage sequence identity ≥ 99 % and only *de novo* sequence clusters ≥ 100 sequences in a cluster included. Also, due to the great number of *de novo* clusters matching multiple identical accessions representing a single species of the Kuwaiti DNA database, the accessions were reduced to a single accession per individual to represent a unique accession rather than a list of repeated accessions representing the same individual (e.g. sample MAS1 *de novo* clustered sequences resulted in a match to 15 identical accessions representing one sequence belonging to Plantago_ciliata_13-0034205_ITS, therefore, the matches were reduced to represent a single accession per individual).

The expected size of the primer set used for *rbcL* amplicons is 500 bp (Kress and Erickson, 2007; Kress et al., 2009) and ITS2 amplicons range from 300 to 460 bp (Chen et al., 2010). The metabarcoding samples amplified from soil eDNA and sequenced by Illumina, resulted in *rbcL* amplicon size range from 150 to 464 bp and ITS2 from 114 to 486 bp, 19 % of the amplicons represented by sequence size > 250 bp, with matches ≥ 99 % identity (*E*-value varied from 0.0E+00 to 2.00E-31). The sequence counts of *de novo* cluster reads for *rbcL* amplicons range from 100 to 153,051 counts and ITS2 from 100 to 168,889 counts (Table 5.19).

The BLASTn search results for each sample are presented in two tables showing *rbcL* and ITS2 matches separately, Table 5.20 is showing *rbcL* and Table 5.21 is showing ITS2 matches of *de novo* clustered sequences to the Kuwaiti DNA database. BLASTn matches for all samples represented by *rbcL* sequences resulted in a total match of 119 accessions belonging to 99 species (Table 5.20) and ITS2 resulted in 87 accessions belonging to 76 species (Table 5.21).

The greatest match per sample representing both DNA markers using *de novo* clustered sequences BLASTn to Kuwait DNA database varied from 87 accessions representing 76 species for sample MAS7 to only 30 accessions representing 22 species for sample MAS6 (Table 5.20 and Table 5.21).

The following summarizes BLASTn matches to Kuwaiti DNA database representing species with high sequence reads across all metabarcoding samples:

Representatives of *rbcL* barcodes: *Astragalus spinosus, Cuscuta planiflora, Gymnarrhena micrantha, Polypogon monspeliensis, Savignya parviflora, Trigonella hamosa,* and *Trigonella stellate* (Table 5.20), and ITS2 barcodes: *Astragalus annularis, Cuscuta planiflora, Fagonia glutinosa, Melilotus indicus, Plantago boissieri, Plantago notata* and *Senecio glaucus* (Table 5.21).

The following summarises the plants found present above ground with sequence matches to Kuwaiti DNA database across the metabarcoding samples:

Seventeen species are representatives of *rbcL* barcodes: *Anisosciadium lanatum, Asphodelus tenuifolius, Convolvulus oxyphyllus, Cuscuta planiflora, Gymnarrhena micrantha, Gypsophila capillaris, Helianthemum lippii, Heliotropium bacciferum, Koelpinia linearis, Launaea mucronata Launaea nudicaulis, Plantago boissieri, Plantago ovata, Rhanterium epapposum, Schismus barbatus, Senecio glaucus,* and *Stipa capensis* (Table 5.20), and 16 species are representatives of ITS2 barcodes: *Anisosciadium lanatum, Asphodelus tenuifolius, Atractylis carduus, Brassica tournefortii, Cuscuta planiflora, Gymnarrhena micrantha, Gypsophila capillaris, Helianthemum lippii, Koelpinia linearis, Launaea mucronata, Launaea nudicaulis, Plantago boissieri, Rhanterium epapposum, Lomelosia palaestina, Senecio glaucus,* and *Stipa capensis* (Table 5.21).

Table 5.20 *de novo* clustered sequence reads blasted against Kuwait DNA database comprising of *rbc*L barcodes (matches ≥ 99 % identity, *de novo* clustered reads ≥ 100 sequences and matches to multiple accessions of the same individual reduced to a single accession)

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Aegilops_triuncialis_15-0042989_rbcL_	107	0	223	140	0	0	0	0	0	0	0	0
Allium_sindjarense_15-0042724_rbcL_	0	0	570	408	0	0	1126	351	0	0	0	0
Andrachne_telephioides_15-0042569_rbcL_	0	0	0	0	171	0	292	0	0	0	0	0
Anisosciadium_lanatum_15-0042650_rbcL_**	0	0	0	0	0	0	0	711	0	233	3241	0
Artemisia_scoparia_15-0042794_rbcL_	335	0	0	0	0	0	0	0	0	0	0	0
Asphodelus_tenuifolius___15-0042594_rbcL_**	0	0	0	0	0	0	0	0	499	0	0	0
Astragalus_annularis_15-0042617_rbcL_	0	0	0	0	0	0	0	232	0	0	0	0
Astragalus_bombycinus___13-0034197_rbcL_	637	437	627	2000	0	0	0	2073	0	0	0	0
Astragalus_spinosus__13-0034202_rbcL_	89325	70004	84657	285435	18209	2568	8247	258745	20026	21851	6151	0
Astragalus_tribuloides___13-0034201_rbcL_	0	0	0	0	0	0	0	0	194	0	0	0
Astragalus_tribuloides__13-0033617_rbcL_	798	727	720	1069	0	0	285	1397	1540	0	0	0
Calotropis_procera_15-0042692_rbcL_	0	0	0	235	681	0	675	0	0	0	0	0
Convolvulus_oxyphyllus_15-0042609_rbcL_**	1211	389	394	695	0	0	613	3934	554	1746	4417	0
Cuscuta_planiflora_15-0043143_rbcL_**	0	6139	590	543	2191	520205	1781	831	1031	9697	2090	0
Dipcadi_erythraeum_15-0042964_rbcL_	0	0	0	0	0	0	0	0	0	4245	0	0
Diplotaxis_harra_15-0042610_rbcL_	1006	1805	0	2171	3689	114	6384	2449	0	777	0	0
Ducrosia_anethifolia_15-0042851_rbcL_	0	0	0	0	0	0	0	0	0	304	3277	0
Ducrosia_anethifolia_15-0042851_rbcL_	0	0	0	0	0	0	479	588	0	0	0	0
Echium_rauwolfii_15-0042659_rbcL_	0	0	147	148	241	0	236	209	0	0	0	0
Erodium_cicutarium_15-0042912_rbcL_	0	0	0	0	110	0	0	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Erodium_glaucophyllum_15-0042661_rbcL_	0	100	0	0	0	0	0	0	0	0	0	0
Erodium_glaucophyllum_15-0042681_rbcL_	0	0	0	0	0	0	0	0	0	128	0	0
Erodium_laciniatum_15-0042573_rbcL_	123	0	113	0	479	0	0	255	0	0	0	0
Erodium_laciniatum_15-0042637_rbcL_	126	116	0	0	284	0	0	143	0	0	0	0
Euphorbia_serpens_15-0042911_rbcL_	0	0	0	0	0	0	1012	381	0	0	0	0
Fagonia_glutinosa_15-0042587_rbcL_	0	0	0	0	0	0	0	0	164	0	0	0
Fagonia_glutinosa_15-0042652_rbcL_	0	0	235	0	0	0	0	312	289	0	0	0
Filago_pyramidata_15-0043060_rbcL_	1006	2057	0	0	354	0	0	613	0	383	873	0
Frankenia_pulverulenta_15-0042636_rbcL_	0	0	0	0	0	0	101	0	0	0	0	0
Gymnarrhena_micrantha_15-0042909_rbcL_**	23807	24618	33828	26260	52088	229	69769	29118	24927	0	0	0
Gypsophila_capillaris_15-0043050_rbcL_**	765	1069	2713	1145	3652	0	2940	1273	1342	0	0	0
Helianthemum_kahiricum___13-0034219_rbcL_	0	916	0	0	0	0	0	0	0	0	0	0
Helianthemum_lippii_15-0042591_rbcL_**	0	1047	0	0	0	0	0	0	0	0	0	0
Heliotropium_bacciferum_15-0042675_rbcL_**	0	0	0	0	285	0	0	120	0	0	1611	0
Heliotropium_kotschyi_15-0042859_rbcL_	0	0	0	0	131	0	0	0	0	0	1300	0
Herniaria_hemistemon_15-0042694_rbcL_	0	0	0	0	112	0	0	0	0	0	0	0
Hippocrepis_areolata_15-0042572_rbcL_	0	0	1385	0	1230	0	265	0	0	0	365	0
Hippocrepis_areolata_15-0042902_rbcL_	0	0	1398	0	1365	0	0	0	0	0	0	0
Hippocrepis_unisiliquosa_15-0042370_rbcL_	0	0	0	0	0	0	268	0	0	0	345	0
Hordeum_murinum_15-0043128_rbcL_	0	0	113	0	0	0	0	0	0	0	689	0
Hordeum_murinum_15-0043137_rbcL_	110	632	237	140	377	0	404	118	0	0	754	0
Hypecoum_littorale_15-0042571_rbcL_	0	0	0	0	620	0	849	0	0	0	0	0
Ifloga_spicata_15-0043057_rbcL_	370	4813	0	0	0	0	110	0	0	0	0	0
Koelpinia_linearis_15-0043108_rbcL_**	1053	0	0	0	0	0	152	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Lasiurus_hirsutus_15-0043083_rbcL_	436	258	0	317	1072	0	0	419	0	0	671	0
Lasiurus_hirsutus_15-0043125_rbcL_	0	295	0	0	0	0	246	0	0	0	710	0
Launaea_mucronata__13-0033632_rbcL_**	0	0	0	0	0	7908	0	0	0	0	0	0
Launaea_nudicaulis___13-0034214_rbcL_**	0	0	0	0	0	0	0	100	0	0	0	0
Launaea_nudicaulis__13-0033631_rbcL_**	0	0	0	0	0	0	106	0	0	0	0	0
Loeflingia_hispanica_15-0042629_rbcL_	0	305	0	0	2243	0	0	0	0	0	0	0
Lolium_rigidum_15-0042384_rbcL_	0	0	0	0	0	0	102	0	0	0	129	0
Lotus_halophilus_15-0042348_rbcL_	241	405	0	350	0	0	0	0	0	0	0	0
Lotus_halophilus_15-0042360_rbcL_	218	415	0	346	0	0	0	0	0	0	0	0
Malva_parviflora_15-0042595_rbcL_	0	0	0	0	234	0	0	0	0	0	0	0
Malva_parviflora_15-0042641_rbcL_	0	0	0	0	274	0	0	0	0	0	0	0
Medicago_rotata_15-0042380_rbcL_	0	902	3193	323	5197	0	669	0	221	17148	0	0
Mesembryanthemum_nodiflorum_15-0042560_rbcL_	0	0	0	0	0	0	114	0	0	0	0	0
Neotorularia_torulosa_15-0042564_rbcL_	0	0	0	0	0	0	0	0	0	0	697	0
Nitraria_retusa_15-0042931_rbcL_	313	0	0	247	796	0	1284	340	0	0	0	0
Ochradenus_baccatus_15-0042870_rbcL_	0	0	0	173	182	0	383	0	0	0	0	0
Onobrychis_ptolemaica_15-0042803_rbcL_	0	0	0	0	0	0	766	0	0	576	0	0
Ononis_serrata_15-0042997_rbcL_	285	358	516	0	419	0	0	353	344	0	0	0
Panicum_turgidum_15-0042398_rbcL_	0	0	0	0	0	0	0	0	249	0	0	0
Paronychia_arabica_15-0042343_rbcL_	0	0	0	0	1869	0	510	0	0	0	0	0
Peganum_harmala_15-0042965_rbcL_	110	0	0	0	137	0	0	0	0	0	0	0
Phragmites_australis_15-0043065_rbcL_	0	719	184	0	791	0	0	162	0	0	0	0
Picris_babylonica_15-0043054_rbcL_	0	0	0	0	131	0	0	0	0	961	6493	0
Picris_babylonica_15-0043066_rbcL_	0	0	0	0	203	0	0	0	0	0	5927	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Plantago_boissieri_15-0042580_rbcL_**	0	0	3441	3525	4526	1138	0	0	0	0	520	0
Plantago_coronopus_15-0042566_rbcL_	0	0	0	0	0	0	456	0	0	0	0	0
Plantago_notata_13-0033628_rbcL_	29988	0	0	0	0	0	2060	2135	1549	0	0	0
Plantago_ovata_15-0042621_rbcL_**	29279	0	0	0	0	0	3030	2149	1337	0	0	0
Plantago_psammophila___13-0034207_rbcL_	270	0	3561	3544	4420	1089	0	0	0	0	475	0
Polycarpaea_repens_15-0042967_rbcL_	0	0	0	0	0	0	0	122	0	0	0	0
Polycarpaea_repens_15-0042977_rbcL_	0	0	0	0	0	0	0	244	0	0	0	0
Polycarpaea_robbairea_15-0042907_rbcL_	0	119	0	0	0	0	0	0	0	0	0	0
Polypogon_monspeliensis_15-0042411_rbcL_	0	0	0	0	111	0	157	0	0	0	0	0
Polypogon_monspeliensis_15-0043056_rbcL_	37993	46087	65709	43474	113777	6500	116966	53532	42897	404	0	0
Pulicaria_undulata_15-0042979_rbcL_	0	0	135	0	134	0	0	0	0	0	0	0
Reichardia_tingitana_15-0043000_rbcL_	0	0	120	0	0	6848	200	0	0	0	0	0
Reseda_arabica_15-0042582_rbcL_	0	0	0	0	0	0	239	0	0	0	0	0
Reseda_muricata_15-0042646_rbcL_	327	121	0	176	0	0	0	0	0	0	0	0
Rhanterium_epapposum_15-0042392_rbcL_**	0	0	0	0	0	0	123	0	0	0	0	0
Rostraria_pumila_15-0043127_rbcL_	0	0	0	0	0	0	0	0	0	0	263	0
Salicornia_europaea_15-0042630_rbcL_	0	0	0	116	101	0	0	181	0	0	0	0
Salvadora_persica_15-0042871_rbcL_	0	0	0	0	0	0	1115	0	0	0	0	0
Salvia_spinosa_15-0042901_rbcL_	0	0	148	0	138	0	984	332	327	0	0	0
Savignya_parviflora_15-0042910_rbcL_	2482	2183	2946	2527	4438	0	7456	2658	1830	824	1876	0
Schimpera_arabica_15-0043052_rbcL_	103	0	0	0	0	0	185	0	0	0	0	0
Schismus_arabicus_15-0043059_rbcL_	0	0	0	0	0	0	0	0	644	0	235	0
Schismus_barbatus_15-0043122_rbcL_**	0	0	0	0	0	0	0	0	617	0	153	0
Sclerocephalus_arabicus_15-0042319_rbcL_	101	0	0	0	0	0	0	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Sclerocephalus_arabicus_15-0042860_rbcL_	198	155	843	0	0	0	0	772	309	0	0	0
Scrophularia_desertii_15-0042706_rbcL_	0	123	0	408	616	0	842	153	0	0	0	0
Seidlitzia_rosmarinus_15-0042366_rbcL_	1079	952	2230	566	1495	0	2773	1535	985	191	0	0
Senecio_glaucus_15-0042387_rbcL_**	0	0	420	206	767	552	362	0	0	206	1265	0
Senecio_glaucus_15-0042922_rbcL_**	0	0	448	185	580	0	500	140	134	1372	0	0
Silene_arenosa_15-0043008_rbcL_	1232	1875	1734	280	0	264	1109	2285	1700	0	538	0
Silene_villosa_15-0042607_rbcL_	0	2799	0	0	0	348	260	3075	2579	0	293	0
Sisymbrium_irio_15-0042368_rbcL_	0	0	105	0	0	0	0	0	0	0	0	0
Sisymbrium_irio_15-0042898_rbcL_	0	0	0	0	142	0	0	0	0	0	0	0
Sisymbrium_orientale_15-0042352_rbcL_	0	125	130	1011	156	0	985	497	0	0	313	0
Spergularia_marina_15-0042815_rbcL_	509	924	0	0	144	0	1558	0	107	0	827	0
Sporobolus_arabicus_15-0043062_rbcL_	0	140	0	0	0	0	0	0	0	0	0	0
Stipa_capensis_15-0043126_rbcL_**	0	1644	0	136	0	0	570	0	0	14801	102	0
Stipa_capensis_15-0043135_rbcL_**	0	103	0	0	0	0	0	0	0	16500	0	0
Suaeda_aegyptiaca_15-0043045_rbcL_	0	0	0	0	434	0	169	0	0	0	0	0
Suaeda_vermiculata_15-0042999_rbcL_	0	0	0	116	0	0	0	0	0	0	0	0
Teucrium_oliverianum_15-0042616_rbcL_	532	245	636	825	0	0	2619	497	0	392	0	0
Teucrium_polium_15-0042941_rbcL_	0	0	0	0	483	0	0	0	0	0	0	0
Tribulus_terrestris_15-0042626_rbcL_	0	0	0	0	0	0	0	124	0	0	101	0
Trigonella_hamosa_15-0042954_rbcL_	1637	1401	4352	1549	202	124	6400	1805	1600	0	0	0
Trigonella_stellata_15-0043043_rbcL_	1467	417	1719	1406	3578	108	6081	1829	1264	0	0	0
Typha_domingensis_15-0042670_rbcL_	0	0	0	0	403	0	0	0	0	0	0	0
Zilla_spinosa_15-0043061_rbcL_	1063	1136	1348	280	388	787	240	1365	609	277	0	0
Ziziphus_nummularia_15-0042831_rbcL_	0	0	0	0	0	0	106	119	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Ziziphus_spina-christi_15-0042583_rbcL_	0	0	112	0	0	0	349	338	0	0	0	0
Zygophyllum_qatarense_15-0042691_rbcL_	0	0	239	0	374	0	279	273	0	0	280	0
Zygophyllum_qatarense_15-0042890_rbcL_	0	0	581	308	384	0	260	650	0	0	411	0

(**) Double asterisks indicate plant species present above ground during field survey of April 2014.

Table 5.21 *de novo* clustered sequence reads blasted against Kuwait DNA database comprising of ITS2 barcodes (matches ≥ 99 % identity, *de novo* clustered reads ≥ 100 sequences and matches to multiple accessions of the same individual reduced to a single accession)

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Allium_sphaerocephalum__15-0042665_ITS_	440	475	956	698	1576	101	1513	741	0	0	0	0
Anisosciadium_lanatum___15-0042650_ITS_**	0	0	219	200	577	0	209	1314	143	0	94259	448
Arnebia_linearifolia_15-0042932_ITS_	0	0	141	0	0	0	0	0	0	0	0	0
Asphodelus_tenuifolius_15-0042640_ITS_**	0	0	0	0	0	0	329	0	298853	128	520	186421
Asphodelus_tenuifolius_15-0042974_ITS_**	0	0	0	0	0	0	0	0	104434	0	0	0
Astragalus_annularis_15-0042617_ITS_	1042	0	0	1998	544	0	1434	2423	0	3504	1471	1928
Astragalus_bombycinus_15-0042639_ITS_	0	0	0	160	0	0	0	0	0	0	0	0
Astragalus_corrugatus_13-0034198_ITS_	0	0	0	165	0	0	0	151	0	0	0	0
Astragalus_hamosus__13-0033620_ITS_	197	0	235	879	193	0	0	1534	0	1165	0	0
Astragalus_sieberi_13-0034200_ITS_	0	0	0	0	0	0	0	0	0	140	0	0
Astragalus_tribuloides___13-0033617_ITS_	0	0	0	0	0	0	0	0	0	117	0	0
Astragalus_tribuloides_13-0034201_ITS_	466	0	0	0	0	0	0	0	0	178	7591	1283
Astragalus_tribuloides_13-0034201_ITS_	0	183	837	0	822	294	525	534	129	0	0	0
Atractylis_carduus_15-0043099_ITS_**	0	0	0	0	0	246	347	0	0	0	0	0
Atriplex_dimorphostegia_15-0042856_ITS_	0	0	0	0	4051	0	0	0	0	0	0	0
Brassica_tournefortii___15-0042657_ITS_**	0	0	119	0	0	0	261	0	0	0	252	0
Bromus_tectorum_15-0043130_ITS_	0	0	0	0	0	0	0	0	0	0	164	0
Cakile_arabica_15-0042401_ITS_	0	0	0	0	0	0	0	0	0	201	0	0
Calendula_arvensis_15-0042400_ITS_	0	0	0	0	0	0	0	0	0	0	0	1018
Carrichtera_annua_15-0042970_ITS_	0	0	162	117	0	0	273	196	0	0	0	0
Citrullus_colocynthis_15-0042567_ITS_	0	0	181	0	0	0	0	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Cornulaca_monacantha___15-0042656_ITS_	0	0	147	0	135	0	150	0	0	0	0	0
Cuscuta_planiflora_15-0042960_ITS_**	0	2987	462	377	659	177397	461	489	126	3131	985	4015
Diplotaxis_acris_15-0043064_ITS_	0	0	0	198	0	0	0	0	0	0	0	0
Echinops_blancheanus_15-0042362_ITS_	0	0	0	0	0	0	0	0	0	0	0	125
Emex_spinosa_15-0042581_ITS_	0	0	0	0	0	0	0	0	0	0	0	7139
Erodium_glaucophyllum___15-0042661_ITS_	0	0	168	0	0	0	0	0	0	0	0	0
Euphorbia_densa_15-0042938_ITS_	0	0	0	0	0	0	0	0	0	0	0	153
Fagonia_glutinosa_15-0042587_ITS_	423	471	1035	1339	990	164	847	625	112	271	2191	46401
Farsetia_aegyptia_15-0042397_ITS_	0	0	332	162	280	0	102	319	0	129	0	247
Filago_pyramidata_15-0043060_ITS_	0	1321	0	0	0	0	327	221	0	197	0	0
Gypsophila_capillaris_15-0043050_ITS_**	0	0	1480	0	0	0	0	0	0	0	306	0
Haplophyllum_tuberculatum_15-0042601_ITS_	0	0	245	0	162	0	191	102	0	0	0	0
Helianthemum_kahiricum_13-0034219_ITS_	0	0	0	0	0	0	0	0	0	224	0	0
Helianthemum_lippii_13-0034218_ITS_**	0	327	0	0	0	0	0	0	0	0	0	0
Herniaria_hirsuta_15-0042606_ITS_	171	908	0	0	0	0	0	0	0	0	0	0
Hippocrepis_areolata_15-0042915_ITS_	0	0	1640	0	1903	0	0	147	0	0	0	0
Hippocrepis_unisiliquosa_15-0042370_ITS_	0	0	0	0	0	0	0	0	0	0	0	109
Hordeum_marinum_15-0042359_ITS_	0	0	0	0	0	0	0	0	0	444	2245	914
Ifloga_spicata_15-0043088_ITS_	0	58743	0	0	556	0	0	1423	430	4893	0	16798
Koelpinia_linearis_15-0043068_ITS_**	13948	0	0	0	0	0	0	0	0	0	0	154
Lappula_spinocarpos_15-0042955_ITS_	0	0	0	0	0	0	0	0	0	0	653	0
Launaea_capitata___13-0033630_ITS_	917	0	0	0	0	0	0	0	0	0	0	0
Launaea_mucronata___13-0033632_ITS_**	0	0	0	0	0	21295	0	264	0	132	0	697
Launaea_mucronata_13-0034215_ITS_**	0	0	0	0	0	10397	0	125	0	156	0	264

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Linaria_albifrons_15-0042811_ITS_	0	115	319	213	418	110	185	304	0	0	482	0
Linaria_simplex_15-0042585_ITS_	0	0	0	0	0	0	0	0	0	0	313	124
Loeflingia_hispanica_15-0042956_ITS_	0	1325	0	0	18573	0	0	0	0	121	0	0
Malva_parviflora_15-0042641_ITS_	0	0	0	0	0	0	0	0	0	0	214	0
Medicago_laciniata_15-0043049_ITS_	0	0	7642	161	6669	0	0	0	0	103546	162	1177
Melilotus_indicus___15-0042748_ITS_	812	1988	3127	2053	4845	151	1878	2565	432	0	0	0
Melilotus_indicus_15-0042942_ITS_	379	948	1235	960	2301	141	989	1198	200	0	0	111
Melilotus_indicus_15-0042963_ITS_	0	0	0	0	0	155	0	0	0	0	0	274
Mesembryanthemum_nodiflorum_15-0042560_ITS_	0	0	155	0	0	0	0	0	0	0	0	0
Moltkiopsis_ciliata___15-0042628_ITS_	0	0	0	305	451	0	472	401	0	0	0	0
Moltkiopsis_ciliata_15-0042894_ITS_	0	0	0	0	103	0	0	281	0	0	210	0
Neotorularia_torulosa_15-0042328_ITS_	343	0	0	407	103	0	196	858	0	327	214	1029
Nitraria_retusa_15-0042852_ITS_	0	0	0	0	1158	0	0	0	0	0	0	0
Ogastemma_pusillum_15-0042891_ITS_	0	0	241	0	224	0	0	0	0	0	0	0
Ononis_reclinata_15-0042375_ITS_	0	0	0	173	0	0	108	0	0	0	0	115
Panicum_turgidum_15-0042374_ITS_	0	0	0	175	0	0	0	187	0	117	170	552
Picris_babylonica_15-0043054_ITS_	0	0	0	0	6884	0	0	0	0	0	272045	670
Picris_babylonica_15-0043066_ITS_	0	0	0	0	0	378	0	0	0	0	0	119
Plantago_amplexicaulis___13-0033625_ITS_	0	0	0	184	0	0	209	273	0	0	0	0
Plantago_boissieri_15-0042580_ITS_**	0	0	77928	55450	61127	15045	0	0	0	0	29289	14600
Plantago_ciliata_13-0034205_ITS_	15548	0	0	0	0	0	0	0	0	0	0	138
Plantago_coronopus_13-0034210_ITS_	0	0	0	400	0	0	181	292	0	0	0	0
Plantago_notata__13-0033628_ITS_	103435	1998	0	0	0	0	75924	67868	13189	1355	0	0
Polycarpaea_repens_15-0042977_ITS_	0	0	0	0	0	0	0	0	0	0	185	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil sample collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Matches to Kuwait DNA database accessions												
Reichardia_tingitana_15-0042987_ITS_	0	0	4923	0	4274	0	0	0	0	0	0	0
Rhanterium_epapposum_15-0042396_ITS_**	0	0	0	0	0	0	0	250	0	0	0	222
Rostraria_pumila_15-0043136_ITS_	0	0	0	0	125	0	1362	0	181	0	0	193
Salsola_imbricata_15-0042350_ITS_	0	0	112	0	0	0	0	0	0	0	0	0
Salvia_spinosa_15-0042883_ITS_	0	0	0	0	0	0	102	0	0	0	0	0
Savignya_parviflora_15-0042388_ITS_	0	0	0	0	0	0	0	0	0	150	0	0
Lomelosia_palaestina_15-0042612_ITS_**	0	0	0	257	0	0	0	0	0	0	0	0
Schismus_arabicus_15-0043059_ITS_	0	0	0	0	0	0	0	0	21620	0	123	52422
Schismus_arabicus_15-0043091_ITS_	0	0	0	0	0	0	0	0	22503	0	203	50709
Senecio_glaucus_15-0042387_ITS_**	0	942	1704	7540	1247	13676	0	113	323	0	46432	607
Senecio_glaucus_15-0042922_ITS_**	0	491	879	3499	610	6903	0	0	161	0	23519	294
Stipa_capensis_15-0043135_ITS_**	0	11516	0	0	0	0	364	0	0	169049	0	211
Stipagrostis_plumosa_15-0043079_ITS_	0	0	0	118	0	0	0	0	0	0	0	0
Trigonella_stellata_15-0043043_ITS_	3026	0	0	0	0	0	0	0	0	4197	0	0

(**) Double asterisks indicate plant species present above ground during field survey of April 2014.

In an attempt to further explore the BLASTn results, a summary of the *de novo* clustering matches to Kuwait DNA database for each of the metabarcoding samples are presented in Table 5.22. The greatest number of *de novo* clustered sequence matches for *rbcL* was sample MAS7 with 59 accessions belonging to 56 species and for ITS2 was MAS19 with 37 accessions belonging to 31 species. The lowest number of matches for *rbcL* was sample MAS6 with 15 accessions belonging to 15 species and for ITS2 samples MAS1, MAS6 and MAS9 each matched 15 accessions belonging to 14, 11, 11 species, respectively. For MAS19, no *rbcL* sequence matches were scored and the sample was only represented by ITS2 amplicons which resulted in a match to 37 accessions belonging to 31 species (Tables 5.20 - 5.22).

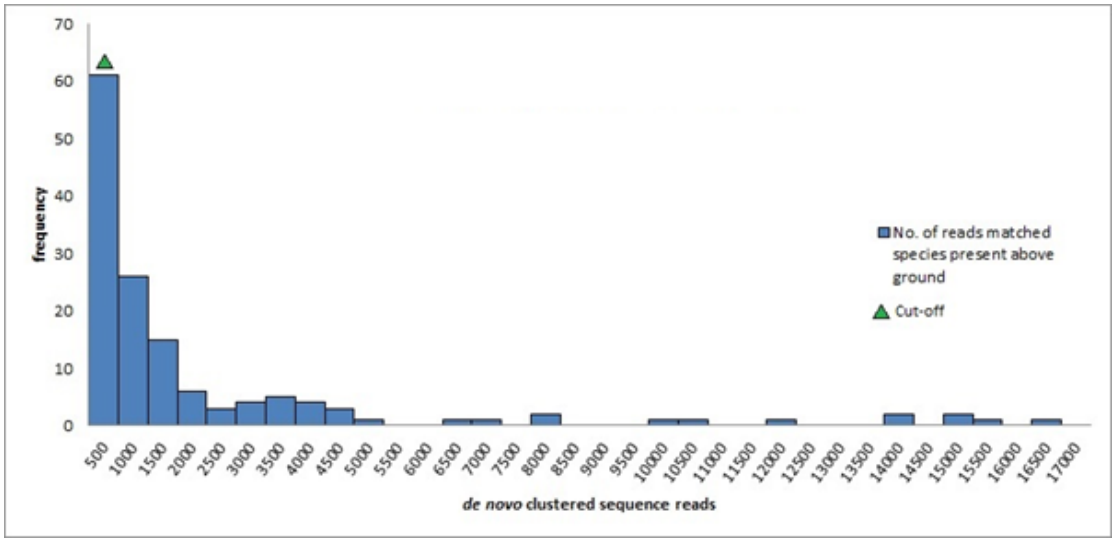
Table 5.22 Summary of *de novo* clustering matches to Kuwait DNA database

Metabarcoding sample	**Depth of soil collection (cm)**	**Number of accessions match**			**Species match**		
		*rbc*L	ITS2	Total *rbc*L + ITS2	*rbc*L	ITS2	Total *rbc*L + ITS2 *
Site 1							
MAS1	surface (0-5)	37	15	52	34	14	44
MAS2	surface (0-5)	41	16	57	38	14	45
MAS3	surface (0-5)	40	29	69	36	26	54
MAS4	surface (0-5)	38	27	65	36	25	56
MAS5	surface (0-5)	55	31	86	48	27	70
MAS6	surface (0-5)	15	15	30	15	11	22
MAS7	surface (0-5)	59	28	87	56	27	76
MAS8	surface (0-5)	49	29	78	46	26	65
MAS9	surface (0-5)	29	15	44	27	11	31
MAS12	deep (10-15)	21	25	46	19	23	37
MAS13	deep (10-15)	33	26	59	29	24	46
MAS19	deep (10-15)	0	37	37	0	31	31

(*) Astrik = total count of species excluding multiple accessions per species

de novo clustered sequence reads of *rbcL* and ITS2 amplicons were further investigated to find a cut-off value by including all reliable sequence reads of plants only known to be present above ground during the field survey (Table 5.11). The dataset from the BLASTn hit results representing all metabarcoding samples (Table 5.20 and Table 5.21) and present above ground (Table 5.11) were analysed using a histogram present in Figure 5.13. The histogram shows that the highest frequency of plants lies under the clustered sequence reads ≤ 500 (frequency = 61) followed by 501-1000 (frequency = 26) and 1001-1500 (frequency = 15). Therefore, I may conclude that a reasonable cut-off value of sequence reads ≤ 500 for showing the greatest frequency amongst plants present above ground (Figure 5.13). Furthermore, Table 5.20 and Table 5.21 are representatives of high *de novo* clustered sequence reads and best at representing the eDNA of plant diversity present in the soil mixture at the study area.

Figure 5.15 *de novo* clustered sequence reads matching plant species present above ground



5.3.4 A comparison of DNA markers represented by above ground plant diversity

The following comparison is to determine which DNA marker (*rbcL* or ITS2) is best at representing the diversity of plants using BLASTn matches to Kuwaiti DNA database and information collected during field survey which included recording the plants present above ground (Table 5.11). The highest numbers of sequence matches for both DNA markers are samples MAS7 and MAS8, 17 and 15 species match, respectively (Table 5.23). The lowest match was represented by sample MAS1 with only 6 species match to above ground plant species (Table 5.23).

Considering each DNA marker separately, *rbcL* marker represented the plant diversity above ground better by resulting in more matches to species present in the study area e.g. Samples MAS7 and MAS8 matched 10 species each for *rbcL* while ITS2 matched 7 species for sample MAS7 and 5 species for MAS8 (Table 5.23).

Table 5.23. Comparison of DNA markers showing matches to Kuwait DNA database using plant species present above ground during field survey

Metabarcoding sample	Depth of soil collection (cm)	No. of species match with above ground flora		Total number of species match
		*rbc*L	ITS2	*rbcL*+ITS2
Site 1				
MAS1	surface (0-5)	5	1	6
MAS2	surface (0-5)	7	4	11
MAS3	surface (0-5)	6	6	12
MAS4	surface (0-5)	7	5	12
MAS5	surface (0-5)	6	4	10
MAS6	surface (0-5)	5	6	11
MAS7	surface (0-5)	10	7	17
MAS8	surface (0-5)	10	5	15
MAS9	surface (0-5)	8	5	13
MAS12	deep (10-15)	5	5	10
MAS13	deep (10-15)	7	7	14
MAS19	deep (10-15)	0	9	9

Further investigation of sequence match to above ground plant species shows that ITS2 marker complements *rbcL* when studying the plant diversity found above ground. Table 5.24 is showing a list of plant species match for both DNA markers (*rbcL* and ITS2) using samples MAS7 and MAS8 as an example to investigate the results and compare between the two DNA markers. The DNA marker ITS2 for sample MAS7 resulted in four new species match (*Anisosciadium lanatum, Asphodelus tenuifolius, Atractylis carduus* and *Brassica tournefortii*) not present in *rbcL* datasets. Also for sample MAS8, ITS2 resulted in two new species matches (*Launaea mucronata* and *Rhanterium epapposum*) not present in *rbcL* data set (Table 5.24). Although *rbcL* sequences are showing the greatest number of species matches compared to ITS2, *rbcL* is being complemented by ITS2 in the identification process by adding new individuals to the data set (Table 5.24).

Table 5.24 Comparison of DNA markers (*rbcL* and ITS2) matching Kuwait DNA database using blastn results for samples MAS7 and MAS8

Sample MAS7 DNA markers		Sample MAS8 DNA markers	
rbcL	**ITS2**	***rbcL***	**ITS2**
Convolvulus oxyphyllus	***Anisosciadium lanatum***	*Anisosciadium lanatum*	*Anisosciadium lanatum*
Cuscuta planiflora	***Asphodelus tenuifolius***	*Asphodelus tenuifolius*	*Cuscuta planiflora*
Gypsophilla capillaris	***Atractylis carduus***	*Convolvulus Oxyphyllus*	***Launaea mucronata***
Koelpinia linearis	*Brassica tournefortii*	*Cuscuta planiflora*	***Rhanterium epapposum***
Launaea nudicaulis	*Cascuta planiflora*	*Gypsophila capillaris*	*Senecio glaucus*
Plantago ovata	*Launaea nudicaulis*	*Heliotipium bacciferum*	
Rhanterium epapposum	*Stipa capensis*	*Launaea nudicaulis*	
senecio glaucus		*Plantago ovata*	
Stipa capensis		*Senecio glaucus*	
Gymnarrhena micrantha		*Gymnarrhena micrantha*	

Species in bold text are showing individuals not present in *rbcL* data set

5.3.5 Metabarcoding analyses using Angiosperm-NCBI database

BLASTn searches of MiSeq metabarcoding samples represented by *rbcL* and ITS2 regions using *de novo* clustered sequences conducted against the Angiosperm-NCBI database for all 12 samples resulted in a total match to 363 species with percentage identity match ≥ 99 %. Table 5.25 is showing the results from BLASTn matches of *de novo* clustered sequences for all 12 metabarcoding samples conducted against Angiosperm-NCBI database, followed by a summary presented in Table 5.26 showing the number of species matching NCBI database for each sample.

The greatest number of matches to the Angiosperm-NCBI database was represented by 171 species for sample MAS7 and the smallest number of matches with 57 species was sample MAS6 (Table 5.25 and Table 5.26).

The greatest number of matches to the flora of Kuwait was represented by 13 species for sample MAS7 and the smallest number of matches with only 2 species, sample MAS13 (Table 5.25 and Table 5.26).

From the total matches of 363 species to Angiosperm-NCBI database, only 23 matched the flora of Kuwait and only 2 species are representatives of the present above ground flora at the study area, Um-Niqa. The two species are represented by *Cuscuta planiflora* (common parasitic plant) and *Schismus barbatus* (common grass) (Table 5.25 and Table 5.26).Although the sequences generated from the soil eDNA are not well represented in the Angiosperm-NCBI database at species-level, a high number of sequences matched to genus-level of the flora of Kuwait (Table 5.25 and Table 2.3).

Table 5.25 Blastn of metabarcoding *de novo* clustered sequence reads per species match to NCBI-Angiosperm database

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Acantholepis_orientalis	0	310	4923	0	4274	0	0	0	0	107	0	0
Achatocarpus_gracilis	0	0	0	0	0	0	0	0	0	108	0	246
Achillea_millefolium	0	0	218	317	1072	0	0	419	1147	0	0	0
Achnatherum_pekinense	0	0	152	0	143	0	0	0	460	0	0	0
***Aeluropus_littoralis* ***	0	0	0	0	0	0	200	0	0	0	0	0
Agrostemma_githago	0	0	153	115	288	0	0	108	344	0	0	0
Agrostis_clavata	0	0	0	1588	254	0	1592	764	0	697	0	0
Ajania_gracilis	0	389	394	0	0	0	456	0	554	4417	0	1211
Aldrovanda_vesiculosa	0	618	0	0	304	0	0	0	697	0	156	436
Alisma_gramineum	0	2799	100	0	0	264	0	122	1700	1265	0	0
Allium_altaicum	0	0	2005	0	0	0	0	0	254	0	0	0
Allium_ampeloprasum	288	0	0	736	254	0	2048	2619	0	0	0	0
Allium_tuberosum	280	0	0	0	6801	0	0	0	321	220	750	0
Alocasia_macrorrhizos	0	0	0	0	141	0	572	0	0	0	0	0
Alocasia_sanderiana	0	0	567	0	2331	0	500	274	0	0	0	119
Amorphophallus_albus	0	306	0	0	0	302	570	0	0	0	0	0
Amphiglossa_tomentosa	0	0	0	244	0	0	0	0	0	0	0	128
Andersonia_ sprengelioides	0	0	0	0	0	0	0	0	0	ASU79742271	0	0
***Andrachne_telephioides* ***	0	0	0	0	343	0	292	0	0	0	0	0
Anemone_patens	0	993	194	0	0	0	0	133	0	625	0	0
Angelica_anomala	0	140	0	0	0	0	193	0	0	438	0	0
Antennaria_parvifolia	0	0	135	0	0	0	0	0	0	111	0	0
Anthemis_arvensis	0	0	133	103	0	0	104	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Aptenia_cordifolia	0	0	2230	0	101	696	0	0	0	0	0	0
Arabidopsis_arenosa	0	0	0	0	0	0	239	0	0	0	0	0
Arceuthobium_azoricum	0	415	1398	350	0	0	110	0	0	0	271	0
Artemisia_annua	0	0	0	0	0	0	0	258	0	0	0	0
Artemisia_frigida	0	0	0	0	0	0	0	157	0	0	0	0
Artemisia_lactiflora	1044	0	631	1989	0	0	0	717	0	0	2523	141
Arytera_brackenridgei	0	0	0	0	119	0	0	541	0	0	0	0
Asclepias_nivea	0	0	0	0	1294	0	0	0	0	0	0	0
Asparagus_asparagoides	0	0	0	0	0	0	0	147	0	0	0	0
Asphodelus_aestivus	0	101	0	0	0	552	0	0	0	0	0	159
Asphodelus_albus	0	0	148	0	117	0	0	0	290	0	0	0
Aster_ageratoides	0	0	261	0	0	0	0	0	0	0	0	0
Aster_glehnii	3367	380	359	5093	9421	0	0	8077	0	0	0	1523
Aster_koraiensis	0	105	0	0	0	0	0	0	0	0	0	541
Aster_spathulifolius	0	0	0	0	0	351	0	0	0	0	0	0
Astragalus_agrestis	954	0	0	714	9533	0	0	0	0	0	349	0
Astragalus_bodinii	0	0	0	1304	0	0	269	0	0	0	0	0
Astragalus_drummondii	169065	11516	0	0	274	0	0	0	0	0	0	0
Astragalus_laxmannii	1128	0	0	0	0	0	0	0	0	0	0	0
Astragalus_nakaianus	1279	1415	1212	0	0	109	0	0	203	0	123	817
Astragalus_uliginosus	0	0	309	408	0	0	1126	351	0	0	0	0
Astragalus_villosus	0	0	0	26260	52415	0	70432	29304	0	0	0	0
Atractylodes_lancea	0	0	3561	3544	4526	0	0	2149	0	963	0	0
Atractylodes_lancea	116804	106	0	17257	32808	0	91162	84909	0	0	0	0
Atriplex_glauca	0	0	0	0	0	0	240	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Baptisia_alba	961	0	0	0	131	0	0	0	0	6493	0	0
Barbarea_verna	0	436	227	840	125	105	106	713	0	0	0	354
Bellevalia_romana	0	0	0	0	149	0	0	0	0	0	0	0
Bienertia_cycloptera*	0	0	0	0	285	0	169	0	123	0	0	0
Bosea_yervamora	0	341	0	0	0	306	510	0	432	0	0	0
Bowiea_volubilis	0	0	0	0	0	0	0	0	142	0	0	135
Brachyachne_ciliaris	0	0	188	284	0	0	0	0	0	0	0	0
Brachylaena_ilicifolia	43773	141521	172812	577535	37325	5404	17077	527095	41588	12266	0	180901
Brachypodium_ sylvaticum	0	116695	112	196	0	0	0	328	21445	491	0	0
Brassica_nigra	3306	EGU806794620	EGU806794139	0	363	0	0	969	0	EGU80679129	0	EGU80679240
Bupleurum_falcatum	0	295	997	0	0	0	0	0	602	0	0	0
Burnatia_enneandra	0	0	0	0	0	0	330	0	0	719	0	0
Calathea_guzmanioides	0	216	0	0	0	0	0	0	0	0	0	0
Calepina_irregularis	0	118	340	0	0	0	0	0	0	0	0	0
Calligonum_molle	146	0	0	406	0	0	0	372	0	0	388	0
Calopappus_acerosus	304	0	0	0	0	0	0	0	0	3277	0	0
Calotesta_alba	0	105	412	414	0	0	304	329	107	10366	0	0
Calotropis_gigantea	120	0	0	0	0	0	0	0	0	0	7139	3026
Calycolpus_moritzianus	10694	5675	0	0	0	0	0	0	0	0	0	0
Camellia_cuspidata	0	0	0	0	0	0	0	0	0	208	0	0
Capsella_bursa-pastoris	0	0	106	1522	2756	0	0	246	0	0	0	1079
Carludovica_palmata	0	106	0	0	0	0	0	0	0	0	0	0
Cassinopsis_ madagascariensis	0	258	119	0	0	307	0	0	0	671	0	340
Catharanthus_roseus	0	0	0	0	0	101	0	0	0	0	0	0
Cathedra_acuminata	0	0	188	0	0	0	1071	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Centropodia_forskalii *	121	0	0	0	0	0	122	0	68967	303	0	119
Chenopodium_murale *	1300	0	14727	11120	24924	0	369	11627	0	0	0	0
Chilopsis_linearis	0	0	0	0	0	0	0	201	0	256	0	0
Chimaphila_umbellata	0	289	0	0	0	0	0	0	0	0	0	0
Chlorophytum_comosum	0	0	77925	0	0	15045	106	119	0	29536	14738	241
Choisya_ternata	1765	0	291	760	377	0	0	3855	0	0	0	0
Cicer_arietinum	0	1325	302	746	0	0	0	1273	249	0	1435	0
Cirsium_arvense	0	0	0	0	0	0	236	0	0	0	111	0
Cladrastis_delavayi	0	0	0	0	0	0	0	0	260	0	0	0
Codonanthe_gracilis	0	262	113	110	0	0	103	236	166	0	0	0
Convolvulus_ulcinus	422	0	0	303	184	0	0	1726	0	0	0	0
Cortaderia_selloana	0	0	169	0	1194	0	1366	635	0	0	0	927
Crassula_tillaea	0	0	946	0	185	0	0	511	0	0	0	0
Ctenanthe_marantifolia	178	138	184	0	0	0	0	0	0	0	0	30171
Cupaniopsis_myrmoctona	0	0	0	0	0	696	160	0	0	347	0	0
Cuscuta_approximata	0	0	0	129	0	0	0	0	0	0	0	0
Cuscuta_planiflora **	299	0	0	116	1844	221	0	2041	547	541	0	0
Cynodon_dactylon *	0	0	0	0	0	0	981	331	0	0	0	0
Cypripedium_macranthos	0	0	0	0	0	0	207	0	0	0	0	0
Cypselea_humifusa	0	0	0	0	1821	0	0	0	0	738	0	0
Delosperma_cooperi	263	267	16951	130	11155	2701	0	1756	455	1878	17469	0
Desmazeria_rigida	404	0	32482	21344	0	0	0	26544	0	0	0	0
Dianthus_longicalyx	1168	48840	0	0	8557	0	53640	11738	0	257	0	0
Dichilus_pilosus	0	168	341	0	175	0	0	0	110	0	0	0
Dieffenbachia_seguine	278	160	378	0	0	0	1416	LMU79740774	5419	191038	174	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Dionysia_caespitosa	0	0	843	0	0	0	845	0	0	0	0	0
Dionysia_hissarica	0	0	0	0	0	0	0	0	0	252	0	0
Dionysia_involucrata	0	0	0	0	0	0	0	0	0	221	0	0
Dionysia_paradoxa	0	0	0	0	0	0	116	0	0	145	0	0
Dionysia_revoluta	0	0	0	0	0	0	0	0	0	0	0	116
Discaria_chacaye	0	0	382	305	121	0	1878	682	462	0	371	18824
Disparago_ericoides	140217	206	758	103	1390	173	1009	2230	31307	2526	1903	0
Dolichothrix_ericoides	0	0	0	308	384	0	260	650	0	0	863	0
Echinodorus_grandiflorus	0	448	0	0	0	6848	1760	0	0	0	0	0
Edmondia_sesamoides	8769	0	244	323	1060	124	0	0	1264	292	0	116
Elymus_enysii	0	305	0	0	0	0	0	0	0	0	0	0
Elytropappus_ rhinocerotis	0	0	0	0	129	0	0	0	0	0	0	0
Empetrum_nigrum	0	0	0	208	134	0	0	117	0	101	0	0
Enhalus_acoroides	523	0	621	246	862	0	273	630	181	0	1083	0
Epipactis_palustris	0	0	952	0	0	862	318	0	449	0	0	246
Eremopyrum_orientale	0	0	0	1145	3652	0	2940	1273	0	1786	0	0
Erigeron_breviscapus	0	221	0	0	EGU8067914187	0	EGU80679922	0	0	0	0	0
Erodium_glaucophyllum*	0	0	235	0	0	0	491	131	289	0	0	0
Erodium_laciniatum*	0	0	0	0	763	0	0	398	0	0	0	0
Erodium_moschatum	0	0	128	0	0	0	108	0	0	0	0	0
Eruca_vesicaria	2567	2493	4980	3222	7162	1306	5862	5565	1069	6624	5320	1447
Erythroxylum_areolatum	0	0	0	0	206	0	0	188	0	0	0	0
Euphorbia_maculata	0	0	0	0	0	0	521	250	0	0	0	0
Euphorbia_sp.*	0	0	1385	0	0	0	0	0	0	365	0	459
Fagonia_acerosa	0	0	0	0	0	0	0	312	164	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Fagonia_paulayana	168	263	535	1289	0	0	0	1559	104	0	0	580
Fagopyrum_crispatifolium	0	884	290	0	238	0	108139	0	0	255	0	0
Fagopyrum_esculentum	147	0	0	0	0	0	0	118	0	115	0	0
Fagopyrum_gracilipes	0	0	0	0	0	0	1045	0	198	0	0	0
Fagopyrum_tataricum	0	100	0	0	0	0	275	0	0	0	0	0
Fagopyrum_ wenchuanense	0	0	0	0	269	0	0	0	0	129	0	0
Festuca_arundinacea	0	6183	184	125	0	0	58571	177	0	619	0	0
Festuca_simensis	0	0	0	346	1365	119	0	0	0	345	0	0
Festuca_simensis	254	2057	0	0	354	0	0	613	0	873	0	1006
Festuca_simensis	604	1330	142	0	0	0	373	0	0	0	1113	0
Flaveria_pringlei	0	5977	780	1364	768	0	298	1238	0	0	0	0
Flueggea_suffruticosa	0	0	0	0	322	0	271	0	0	3241	0	0
Foeniculum_vulgare	0	110	0	0	0	0	0	0	0	0	0	0
Franklinia_alatamaha	0	0	0	0	0	0	0	0	0	0	0	5943
Galium_boreale	0	0	0	161	0	0	0	0	0	0	0	173
Geranium_traversii	0	0	1888	246	202	502	6081	0	0	417	0	0
Gibbaeum_pachypodium	4227	475	3034	198	3070	0	106	3085	0	272	385	468
Glycyrrhiza_glabra	388	794	652	303	1197	0	0	1076	0	925	0	0
Gossypium_thurberi	0	10932	25368	235	681	100047	809	0	161970	292815	174139	129218
Gypsophila_repens	1620	0	0	0	0	0	0	0	0	0	0	0
Halocnemum_ strobilaceum *	186	0	0	651	0	0	10199	433	0	0	0	332
Haloxylon_ammodendron	0	0	0	552	0	0	0	767	0	130752	670	0
Haloxylon_persicum	306	0	245	662	630	0	350	102	0	1947	162	0
Handroanthus_aureus	151	908	548	0	0	0	158	228	0	0	0	0
Hedysarum_tibeticum	583	9575	406	1980	3353	180	1154	395	0	3269	394	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Heliotropium_erosum	0	0	0	0	0	0	325	0	0	1300	0	0
Helminthotheca_echioides	0	155	0	0	113	0	3496	0	PENCARBOXL147	0	0	0
Heptacodium_ miconioides	0	1116	322	0	2065	0	0	260	0	102	0	0
Herbertia_darwinii	0	232	0	0	0	0	0	124	0	0	0	0
Herniaria_ciliolata	0	0	0	0	0	0	0	0	0	143	0	0
Hesperis_matronalis	0	0	0	0	0	0	401	0	0	0	0	19656
Hesperostipa_comata	0	278	0	0	0	0	0	0	0	0	0	896
Hippocrepis_comosa	0	115	135	181	106	0	0	179	0	0	0	0
Homalomena_speariae	0	0	0	183	0	0	114	194	0	0	0	0
Hordeum_secalinum	0	0	223	0	272	0	0	0	0	0	0	107
Horovitzia_cnidoscoloides	0	0	0	173	182	0	383	0	0	0	0	0
Hyacinthoides_hispanica	845	2183	3339	0	0	0	0	0	1830	3002	0	2749
Hydrilla_verticillata	0	10110	587	0	0	109	102	0	157	2267	1338	246
Hypochaeris_brasiliensis	0	0	112	0	0	0	516	0	0	0	0	0
Hypselodelphys_hirsuta	334	3015	0	0	131	0	9412	3268	69224	384	0	636
Ilex_asperula	0	542	0	0	0	1089	0	164	1337	447	0	0
Imperata_cylindrica *	0	0	269	0	0	0	0	0	0	241	485	0
Ionopsidium_abulense	191	150	1254	566	1495	0	2610	1535	248	4674	1690	0
Isatis_tinctoria	0	358	516	183	0	0	593	0	0	0	0	202
Jacobaea_vulgaris	0	603	0	162	0	0	671	0	0	0	944	0
Jarilla_chocola	1372	139	960	0	0	0	0	0	0	0	0	0
Kadsura_coccinea	0	21807	143	120	0	0	2105	0	0	0	0	10638
Kniphofia_linearifolia	0	CSXCPRBCL123	0	0	0	0	0	0	0	0	0	0
Lachnospermum_ imbricatum	0	0	0	0	0	0	0	0	125	4677	0	285
Lactuca_sativa	0	24723	33966	0	255	586	0	0	24927	1330	0	23807

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Lathyrus_palustris	0	0	3441	3525	778	0	0	2135	0	0	0	0
Leontopodium_leiolepis	0	0	121	0	136	0	2060	0	0	110	0	10635
Lepidium_didymum	0	0	471	0	0	0	0	3075	416	0	0	0
Lepturus_repens	0	435	237	140	0	0	0	118	0	710	0	0
Leucopogon_ microphyllus	0	0	0	0	0	0	111	1173	237	0	836	0
Levisticum_officinale	0	245	0	107	124	0	0	120	0	0	0	2816
Lilium_brownii	271	552	877	341	1613	0	1937	548	758	3231	77435	0
Lilium_lankongense	0	0	529	0	0	0	1115	0	0	0	0	0
Lilium_pensylvanicum	0	0	0	22250	334	271	571	27499	355	0	637	0
Limeum_arabicum	0	0	0	0	0	0	118	0	0	0	0	218
Logfia_gallica	0	0	148	0	292	0	1026	267	0	0	537	110
Lotus_creticus	0	3890	0	0	0	0	0	0	0	129	0	0
Lotus_japonicus	257	1988	0	0	771	0	0	2712	0	0	0	30079
Loxodiscus_coriaceus	254	0	114	0	0	21295	0	264	0	770	1144	14865
Luzula_rufescens	0	0	0	0	272	0	268	126	0	0	0	0
Malva_pusilla	0	0	0	0	0	0	156	0	4278	0	9468	0
Marantochloa_filipes	0	127	168	0	0	0	401	0	398	499	0	370
Medicago_lupulina	0	0	0	0	0	0	0	0	221	159	0	1467
Medicago_sativa *	777	1285	177	2171	3689	662	6384	2449	0	0	0	122
Medicago_truncatula	206	0	320	206	0	0	0	0	0	0	0	0
Melilotus_albus	0	0	0	130	245	0	629	208	0	0	0	0
Melilotus_officinalis *	1495	12460	1796	0	1230	0	0	0	0	772	0	0
Mertensia_virginica	1746	32715	10171	695	0	0	613	3934	1160	1316	0	47562
Mesembryanthemum_ crystallinum	128	0	0	0	0	0	211	0	144115	0	88777	0
Mesembryanthemum_ nodiflorum *	1588	2087	168	10402	172	229	0	14017	2864	1493	1582	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Metalasia_adunca	3150	0	0	11196	7578	0	6854	5733	0	689	0	0
Metalasia_aurea	0	0	0	0	4635	0	114	0	0	195	125	0
Metalasia_divergens	0	0	0	0	0	0	4308	0	0	132064	1026	19169
Metalasia_inversa	0	48456	0	129	438	0	0	0	0	0	0	610
Metalasia_oligocephala	0	39062	0	55450	61127	0	0	0	298	0	11221	0
Microchloa_caffra	0	0	1734	280	460	0	0	2285	0	0	0	0
Monotagma_ smaragdinum	16600	103	0	0	264	0	0	0	0	0	0	0
Moringa_rivae	0	0	0	0	57425	0	371	0	1562	309	0	0
Muilla_maritima	0	0	0	274	877	0	293	0	0	0	265	0
Muscari_comosum	0	153	113	359	0	0	0	0	0	0	0	0
Najas_browniana	0	0	0	0	20787	0	0	0	0	0	0	0
Najas_flexilis	0	115	3709	0	234	0	110	0	174	0	0	1347
Najas_gracillima	0	185	712	0	188	0	223	0	1600	0	0	0
Najas_marina	0	0	0	0	618	0	0	0	0	0	0	0
Nassella_viridula	0	1047	0	0	0	0	257	0	0	0	0	0
Navarretia_intertexta	0	0	887	0	103	0	0	0	0	0	0	0
Nicotiana_undulata	0	0	0	0	0	0	1412	223	0	393	0	0
Nothocissus_spicifera	233	0	0	0	0	0	0	738	0	0	0	0
Notothixos_leiophyllus	0	0	0	0	478	107	0	0	0	0	0	0
Oecopetalum_mexicanum	0	0	0	0	0	0	122	0	0	0	0	0
Olea_woodiana	0	0	0	0	121	0	0	0	0	0	0	0
Onobrychis_montana	0	125	0	0	0	0	0	0	0	0	0	0
Ononis_repens	0	24280	33397	0	332	0	0	105	0	14482	0	10425
Orixa_japonica	0	0	0	2172	0	0	114	2889	0	5502	0	229
Ormosia_emarginata	0	0	0	0	457	0	0	0	249	197	0	137

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Oxytropis_campestris	0	632	102	102	0	0	0	0	0	0	0	0
Oziroe_biflora	0	0	163	134	0	0	0	157	0	235	0	0
Panax_japonicus	0	0	0	176	0	0	30209	0	0	101	0	0
Panax_japonicus	0	0	957	908	776	0	1844	0	309	495	642	0
Panax_japonicus	0	0	0	CSXCPRBCL408	CSXCPRBCL616	0	CSXCPRBCL842	CSXCPRBCL153	0	0	0	0
Panax_japonicus	1835	19612	1019	0	0	0	24792	0	0	970	0	0
Panax_japonicus	2231	159517	11799	0	0	0	58900	0	0	1758	0	231
Panax_japonicus	4350	472	0	0	0	0	508	0	0	0	0	0
Panax_notoginseng	356	361	480	0	683	212	805	507	201	313	0	0
Panax_pseudoginseng	1016	937	815	0	0	108647	0	0	126	0	1015	440
Panax_stipuleanatus	128	116	113	0	0	0	0	0	0	0	0	249
Paronychia_canadensis	0	378	0	0	0	0	0	162	0	0	0	0
Pectis_papposa	0	0	0	0	0	0	435	772	0	0	0	0
Pennisetum_glaucum	0	0	0	140	0	0	0	0	248	0	0	191
Petalostemon_purpureus	0	0	0	0	766	0	105	PENCARBOXL797	0	5927	0	812
Petroselinum_crispum	0	0	0	0	203	0	102	0	0	0	0	0
Phalaris_arundinacea	0	0	0	0	0	0	114	588	0	139	0	0
Phragmites_australis*	0	1428	0	0	419	0	318	353	0	0	0	0
Phyllostachys_nigra	0	1805	0	0	110	114	0	0	0	0	0	1006
Pickeringia_montana	0	0	132	0	392	0	398	0	0	0	0	0
Pickeringia_montana	0	0	581	0	0	0	0	0	0	411	0	0
Pickeringia_montana	2312	59365	8928	4796	0	0	0	6867	0	0	10230	0
Pilea_verrucosa	0	0	0	0	6934	1138	301	0	1549	371	0	0
Pilosella_lactucella	0	0	106	0	338	0	0	0	0	1540	0	112
Pistia_stratiotes	0	200	0	0	0	0	214	0	0	206	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Plantago_lanceolata *	3880	0	0	840	1625	0	0	1388	0	0	0	867
Plantago_virginica	4092	0	0	748	656	0	1968	914	0	0	0	0
Platycrater_arguta	0	0	0	0	4420	3183	0	0	21555	520	0	0
Pleiostachya_pruinosa	0	0	0	0	0	0	472	0	0	475	284	0
Pleurospermum_cristatum	0	268	300	0	0	151	444	203	0	0	0	0
Polycarpon_tetraphyllum	0	0	0	579	563	0	0	0	0	0	108	0
Polypogon_fugax	14801	0	0	136	0	0	209	0	0	0	0	0
Primula_floribunda	0	0	0	0	56818	1421	0	0	8894	0	0	0
Primula_veris	0	0	0	0	148	1345	1158	0	8971	0	0	0
Primula_verticillata	0	0	0	0	0	0	131	0	0	0	0	327
Prunus_padus	0	1069	2713	667	103	0	121	0	1342	0	0	765
Prunus_pensylvanica	9697	183	0	543	2191	0	1781	831	0	0	0	0
Prunus_persica	8552	902	2949	789	0	108	364	0	0	0	0	1637
Prunus_ussuriensis	0	0	0	0	24889	0	221	106	0	9804	0	0
Prunus_yedoensis	288	0	0	0	0	0	0	0	0	0	0	0
Remusatia_vivipara	0	314	0	0	268	0	1081	0	0	0	916	0
Reseda_luteola	0	0	147	1549	3578	0	585	1805	0	0	254	1144
Sagina_apetala	0	0	0	0	116	0	0	0	438	0	0	0
Sagittaria_pygmaea	0	0	113	0	7005	0	265	5334	135	0	6877	0
Sagittaria_subulata	0	0	0	0	265	0	206	0	0	0	0	242
Salicornia_dolichostachya	0	0	0	569	0	0	103	2166	0	135	0	0
Salsola_soda	0	0	0	191	0	0	0	212	0	520	0	241
Salsola_vermiculata	0	0	0	0	0	0	0	0	0	0	0	110
Salvia_flava	0	0	368	0	322	0	0	0	753	0	0	222
Salvia_isensis	0	0	0	0	0	0	0	332	0	329	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Salvia_nipponica	0	0	0	397	1429	0	0	2726	150300	0	107170	0
Salvia_roemeriana	824	0	103	2707	4438	0	7456	2778	0	0	0	0
Sambucus_nigra	0	0	0	0	176	0	984	0	644	1611	0	791
Sapindus_mukorossi	0	0	0	0	316	0	0	0	0	0	0	0
Saponaria_officinalis	0	1653	0	0	0	1401	0	0	1519	0	166	256
Scadoxus_cinnabarinus	0	0	0	0	112	0	2619	0	0	0	0	0
Schismus_barbatus **	0	0	0	0	0	0	227	0	0	0	0	0
Schoenus_efoliatus	0	224	0	0	0	0	0	0	0	3746	0	173
Schumannianthus_ dichotomus	0	242	130	1011	156	0	985	340	0	313	0	0
Scorzonera_intricata	0	0	0	749	0	0	0	0	0	448	0	0
Sedum_oryzifolium	0	0	0	0	0	348	0	0	2579	0	0	1232
Senecio_sylvaticus	109	24404	35032	186	53071	7908	109	127	25146	155	0	25651
Sesuvium_hydaspicum	0	0	0	0	767	0	122	0	107	136	155	0
Silene_aprica	0	0	189	0	0	0	8391	0	0	0	0	0
Silene_gallica	0	0	0	111	4254	0	0	0	0	0	0	0
Silene_latifolia	0	0	105	0	0	0	0	0	0	538	0	0
Silene_noctiflora	0	0	0	105	0	0	0	26335	0	155	0	0
Silene_paradoxa	139	0	0	868	428	0	0	0	0	0	0	124
Sinapis_arvensis	0	0	0	0	0	0	128	0	0	389	0	478
Sisymbrium_orientale*	0	0	205	0	0	0	0	0	0	0	0	0
Solanum_lycopersicum	0	0	0	0	0	1284	1185	0	0	0	0	0
Solanum_pennellii	0	0	0	0	0	0	1109	0	0	0	0	0
Solidago_missouriensis	0	122	345	0	0	0	0	137	816	118	0	0
Sorghum_bicolor	117	0	0	308	0	0	0	187	0	241	0	0
Soroseris_erysimoides	0	0	0	0	589	0	0	0	0	0	0	0

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Spergularia_azorica	0	443	0	0	232	0	0	0	0	274	0	0
Spirodela_polyrhiza	0	377	0	0	0	0	0	0	0	0	0	0
Sporobolus_japonicus	0	0	0	0	0	0	0	0	0	0	0	143
Stachyphrynium_repens	0	0	354	0	0	101	0	0	0	0	0	3625
Stipa_lipskyi	0	727	720	1069	668	0	285	1392	1734	0	0	798
Stipulicida_setacea	0	0	0	0	0	107	0	0	143	2554	0	564
Streblus_ilicifolius	0	2903	7649	0	0	0	326	1002	1462	13786	11274	0
Styphnolobium_ burseroides	0	916	0	0	0	0	0	0	0	0	0	0
Styphnolobium_conzattii	2261	159848	12064	103	21073	313	108450	26697	1552	0	0	10725
Suaeda_aegyptiaca*	0	0	305	496	582	0	0	173	0	241	934	0
Suaeda_fruticosa*	0	454	0	0	401	0	0	0	0	0	0	1194
Syncarpha_canescens	586	189	4767	2053	0	104	868	534	0	275	383	0
Tetragonia_tetragonioides	0	103	0	0	0	0	0	151	0	0	0	0
Teucrium_viscidum	0	0	0	0	18396	0	0	0	220	616	0	0
Thalia_dealbata	0	0	0	0	483	0	112	0	0	132	0	0
Toddalia_asiatica	0	952	0	116	0	0	0	181	327	0	0	532
Trianthema_argentinum	0	0	0	1496	524	0	6520	0	700	361	5182	0
Trianthema_ ceratosepalum	292	0	636	825	285	0	0	600	0	0	0	0
Trianthema_clavatum	1248	6507	15113	11937	151	0	0	11887	0	909	0	0
Trianthema_corymbosum	0	0	0	0	927	0	0	1846	320	1464	0	0
Trianthema_oxycalyptrum	0	0	0	0	0	0	0	0	0	1094	0	0
Trianthema_ portulacastrum	0	413	581	634	547	201	16972	945	0	3853	3969	0
Trianthema_salsoloides	0	1875	369	0	1103	0	116	0	0	293	0	269
Trianthema_vleiense	0	0	0	0	0	0	238	0	0	312	0	1206

	MAS1	MAS2	MAS3	MAS4	MAS5	MAS6	MAS7	MAS8	MAS9	MAS12	MAS13	MAS19
Soil collection depth (cm)	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**0-5**	**10-15**	**10-15**	**10-15**
Tribulus_lanuginosus	0	509	0	0	0	0	362	0	0	0	0	632
Trifolium_glanduliferum	0	361	0	171	138	0	1220	183	0	184	0	0
Tripsacum_dactyloides	0	0	0	0	0	0	358	0	0	101	0	189
Trisetum_sibiricum	0	1042	2196	313	0	0	4130	0	1292	726	0	100
Triticum_aestivum	0	0	1480	1237	1460	0	0	0	0	0	477	0
Turbina_oblongata	739	1813	3316	0	620	812	0	0	0	0	0	0
Typha_latifolia	0	0	0	0	0	0	0	0	0	754	0	1063
Urginavia_altissima	0	0	104	0	0	0	313	0	0	0	0	0
Vaccinium_myrtillus	0	111	126	0	131	0	0	0	0	0	0	0
Vaccinium_uliginosum	121	0	0	0	0	0	1147	0	0	0	0	148
Vangueria_agrestis	0	0	0	0	927	0	240	128	0	0	0	0
Vangueria_bowkeri	0	0	150	0	398	0	0	354	0	304	0	0
Vasconcellea_quercifolia	4197	2919	3501	0	4051	0	0	0	0	3520	0	2349
Vasconcellea_ sphaerocarpa	0	121	0	0	0	0	29046	0	0	0	0	0
Wisteria_floribunda	0	0	0	0	0	0	128	0	0	13460	0	278
Wolffia_australiana	0	4216	0	0	0	0	0	0	0	928	0	0
Xanthorrhoea_resinosa	0	6139	590	0	0	520523	0	0	1158	2090	0	0
Yucca_glauca	0	405	0	0	0	0	0	0	0	482	0	0
Zaleya_pentandra	0	0	0	151	0	0	0	0	0	0	0	0
Zanthoxylum_nitidum	0	0	0	0	135	0	0	0	0	0	0	0
Zanthoxylum_ovalifolium	388	2876	2595	1742	1447	531	610	204	637	25377	2149	0
Zanthoxylum_piperitum	0	1401	4352	1406	4137	0	331	1829	113	0	0	0
Zanthoxylum_schinifolium	704	0	0	27228	100	0	71410	31123	0	0	0	116

(*) Single asterisks indicate species present in the flora of Kuwait, (**) Double astriks indicate species present above ground at Um-Niqa [*End of Table 5.25*]

Table 5.26 Summary of Angiosperm-NCBI database matches to the flora of Kuwait using *de novo* clustered sequences for all metabarcoding samples

		Blastn matching NCBI-Angiosperm database		
Metabarcoding samples	Depth of soil collection (cm)	No. of species matches to above ground flora	No. of species matches to the flora of Kuwait	Total No. of matches to NCBI-Angiosperm database
Site 1				
MAS1	Surface (0-5)	1	3	86
MAS2	Surface (0-5)	1	8	135
MAS3	Surface (0-5)	0	9	148
MAS4	Surface (0-5)	1	10	123
MAS5	Surface (0-5)	1	9	166
MAS6	Surface (0-5)	0	3	56
MAS7	Surface (0-5)	2	13	170
MAS8	Surface (0-5)	1	6	138
MAS9	Surface (0-5)	0	6	93
MAS12	Deep (10-15)	0	7	139
MAS13	Deep (10-15)	0	2	62
MAS19	Deep (10-15)	0	7	103

5.3.6 A summary of the NGS technologies match with the above ground plant diversity at Um Neqa study area

In an attempt to study the status of plant diversity below-ground with the above ground, the sequence match results from metagenomics and metabarcoding analyses to Kuwait DNA database were compiled in Table 5.27 and compared with the above ground plant diversity, represented by a total number of 30 species. The metabarcoding analysis resulted in 19 species match with the above ground plant diversity, while metagenomics resulted in only 1 species match and represented by *Gymnarrhena micrantha* (Table 5.27).

Table 5.27 Summary of NGS matches with the above ground plant diversity at Um Niqa

Species present above ground	NGS applications	
	Metagenomics	Metabarcoding
Site 1: Protected area		
Allium longisepalum	-	-
Anisosciadium lanatum	-	+
Asphodelus tenuifolius	-	+
Atractylis carduus	-	+
Brassica tournefortii	-	+
Carduus pycnocephalus	-	-
Centaurea pseudosinaica	-	-
Centaurea bruiguriana	-	-
Convolvulus oxyphyllus	-	+
Cuscuta planiflora	-	+
Gypsophila capillaris	-	+
Hammada salicornica	-	-
Helianthemum lippii	-	+
Heliotropium bacciferum	-	+
Koelpinia linearis	-	+
Launaea mucronata	-	+
Pennisetum divisum	-	+
Plantago boissieri	-	+
Plantago ovata	-	+
Rhanterium epapposum	-	+
Rumex vesicarius	-	-
Salvia aegyptiaca	-	-
Lomelosia olivieri	-	-
Lomelosia palaestina	-	+
Schismus barbatus	-	+
Senecio glaucus	-	+
Stipa capensis	-	+
Site 2: Open desert area		
Hammada salicornica	-	-
Arnebia decumbens	-	-
Astragalus schimperi	-	-
Gymnarrhena micrantha	+	-
Moraea sisyrinchium	-	-

(+) species present in data set (-) species absent in data set

5.4 Discussion

This study presents a unique NGS molecular exploration through environmental DNA soil samples of Kuwait. Based on the results generated from metagenomics and metabarcoding analyses, DNA was traceable in the soil of an arid environment. This study documents how both molecular approaches can be used to identify plants extracted from eDNA soil samples collected from rich and poor species habitats.

A technical issue must be considered here is that the DNA is mainly from degraded sources (i.e. DNA material from dead organisms and plant parts leaves, roots, pollens, etc.). Yoccoz et al. (2012) shown that crops cultivated up to 50 years ago could still leave recoverable and identifiable DNA in the soil, despite the level of DNA decay (Yoccoz, et al., 2012). Environmental DNA extracted from soil samples does not have a clear age limit; unknown how fast DNA degrades in soil and how much of it reflects the present vegetation compared with past vegetation, some taxa might be conserved for longer periods than others (Taberlet et al., 2012). However, we do not know how much of the eDNA comes from the local origin (i.e. movement of DNA by wind, water, animals). Such investigation requires calibration studies to specify how much stratification and clustering will result with the highest accuracy of data and information.

5.4.1 Metagenomics

Metagenomics approach was limited in providing identification for sequences at species-level using Angiosperm-NCBI database; this is a common issue for a poorly known flora, weakly represented in the GenBank. Angiosperm-NCBI database resulted in a good BLASTn match at the family level representing the two study sites. Interestingly, using the complete plastid genome database (to show the diversity of green plants) at the order-level, algae, bryophytes and ferns were found to have a high percentage of sequences in the eDNA soil samples compared to magnoliophytes.

Similarity sequence searches blasted against Angiosperm-NCBI database to match species-level showed a high number of matches to *Cuscuta gronovii* represented by

matches at ≥ 98 % ID. A known species to the Kuwaiti flora from the same genus is *Cuscuta planiflora* (dodder parasitic plant), a common parasitic plant present throughout Kuwait and during the field survey, it was observed attached to a dwarf-shrub *Rhanterium epapposum* (Figure 5.2). *Cuscuta* seeds are well documented to germinate independently of the presence of a host; seedlings emerge from the soil up to 7 cm (favourable temperature range 15 to 38 °C); after completing its life cycle on a host plant, it sheds it seeds to the soil and remain dormant for more than 10 years, depending on the environmental conditions (Lanini and Kogan, 2005). Another root-parasitic plant matching Kuwait DNA database is *Orobanche purpurea*, which is found in Kuwait under different species names *O. aegtptiaca* and *O. cernua* (Table 2.3). *Orobanche* spp. is found in sandy and loose soil and commonly growing and flowering throughout Kuwait between March to May (Daoud, 1985). *Sciaphila densiflora*, is also a parasitic flowering plant, not common species to an arid region and generally spreads in tropical and subtropical regions (Tsukaya and Suetsugu, 2014); its sequence source in the eDNA samples could be from the surrounding agricultural area, where it is being cultivated as an ornamental plant.

Another species-level match is *Silene conica* which is a synonym of *S. conoidea* and belongs to a common family Caryophyllaceae, present in the local flora along with 3 other species that did not return with any sequence matches: *Silene arabica, S. arenosa,* and *S. villosa* (Table 2.3).

The presence of sequences matching to crops, are possibly due to fragments of material (e.g. pollens, spores, seeds) transported by wind or animals from nearby farming areas. Al-Abdaly farms are only 3 km away. Before the DMZ area was established and fenced, the farming land extended to Um Neqa, the study area. Crop sequence matches are represented by *Pisum sativum* (green peas), *Hordeum vulgare* (barley), *Musa acuminate* (banana), *Oryza sativa* (rice), and *Cucumis melo* (melon). Peas, barley and banana plants are commonly cultivated inside the greenhouses at Al-Abdaly farms. Also *Hordeum vulgare*, the wild species is present in the flora as *H. marinum* and *H. muranum.* DNA material matching ornamental plants detected in the soil samples and possibly cultivated in nearby nurseries are represented by

Trachelium caeruleum (blue throatwort), *Carnegiea gigantean* (cactus), *Chloranthus spicatus* (pearl orchid) and *Enkianthus perulatus* (white enkianthus). (Table 2.3). From the 19 plant families detected using the NCBI database (Table 5.15), 8 are present in the flora of Kuwait. A high percentage of sequences matched to three common families of the flora represented throughout Kuwait belongs to Fabaceae, Poaceae and Convolvulaceae (Table 5.15 and Table 2.3). Other families with a high percentage of sequence matches to family level and not common to the flora of Kuwait are represented by Cactaceae, Campanulaceae, Ericaceae, Platanaceae and Triuridaceae (Table 5.15). The source of DNA and sequence matches to the families uncommon to the flora of Kuwait could be transported from the neighbouring agriculture area (Al-Abdaly farms) for example Musaceae (banana) are commonly cultivated in Al-Abdaly farms. Individuals belonging to Cactaceae, Campanulaceae, Orchidaceae and Verbenaceae are also widely grown and sold as ornamental plants in Kuwaiti nurseries (Table 5.15). Other sources of DNA material deposited in the Kuwaiti soil could be due to long distance wind dispersal pollen grains from neighbouring countries of the Arabian Peninsula. Saudi Arabia, located south of Kuwait, with a flora consisting of about 2,284 species belonging to 131 families (Thomas, 2017) could possibly be the answer to some sequence matches that are not representatives of the flora of Kuwait. Cactaceae, Campanulaceae, and Ericaceae are families documented in the flora of Saudi Arabia and could be the source of DNA material (i.e. pollen grains) found matching the sequence reads of the Kuwaiti soil samples. Long-distance pollen dispersal is well documented in several publications (e.g. Kuparinen et al., 2007; Albrecht et al., 2009; Millar et al., 2014).

Furthermore, exploring the species present in the flora of Saudi Arabia belonging to the families with sequence matches to the Kuwaiti DNA database, the two species of the flora of Saudi Arabia representing Cactaceae are *Opuntia ficus-indica* (sweet prickly pear) and *O. dillenii* (spiny pest pear). *Opuntia* species are known to reproduce and disperse by seeds and also vegetatively by its fleshy stem fragments (Reyes -Aguero et al., 2006). Stem fragments are either cultivated or spread by becoming attached to animals. Considering the wind transportation of pollen grains from Saudi Arabia to Kuwait, the average size of *Opuntia* sp. pollens is about 100 µm (ElBehi et al., 2015; Reyes-Aguero et al., 2006) which is larger than the average size

of wind transported pollens, 20-60 µm (Faegri & Van Der Pijl, 1979) and could possibly be difficult to elevate into the air to become airborne and travel by wind across long distances. Therefore, the source of DNA material of Cactaceae represented in the Kuwaiti soil samples could mainly be from the neighbouring nurseries and farming areas (Al-Abdaly farms), which commonly cultivate cactus for medicinal and ornamental purposes.

Campanulaceae is represented in the flora of Saudi Arabia by 3 species: *Campanula dulcis, C. edulis,* and *C. erinus*. Campanula in Latin meaning small bell and commonly known as the bell-shaped flowers. The Campanula flowers are either self-fertilised or pollinated mainly by bees (Schlindwein et al., 2005). The mean size of the pollen grains of Campanula sp. is about 30 µm (Khansari et al., 2012; Perveen and Qaiser, 1999) which is the perfect size representing wind transported pollens.

Another family from Saudi Arabia is Ericaceae represented by only one species, *Erica arborea* (tree heath), a shrub or small tree with height ranges from 1 - 4 m. Erica pollens mean size is about 22 µm (Sarwar and Takahashi, 2014) which also makes it easy to disperse across long distances by the wind. Thus, a portion of the DNA source representing the families Campanulaceae and Ericaceae in Kuwaiti soil could be due to the accumulation of long distance wind-transported pollen grains from Saudi Arabia and other neighbouring countries e.g. Iraq located north of Kuwait and Iran located north-west.

BLASTn using complete plastid genome matches revealed an interesting finding at the two study sites, for all eDNA samples at the order level, with a total number of sequence matching 81 % of algae, followed by 12 % bryophytes, 4 % ferns and only 3 % of magnoliophytes. From the results, the two largest orders of the sequence matches representing bryophytes and ferns are Orthotrichales and Polypodiales, respectively. Although bryophytes and ferns are well documented in the Arabian Peninsula (Kurschner, 2003; Kurschner & Ochyra, 2014; Rothfels et al., 2012), it lacks proper documentation and requires serious attention in Kuwait. The only documented species in the Kuwaiti flora representing ferns is *Ophioglossum polyphyllum* (Ophioglossales).

To understand the presence of the rich source of DNA material of algae, bryophytes and ferns in the Kuwaiti soil, I will further explore the possible mechanism behind the dispersal of spores and the ecology of biological soil crusts. Lonnell (2011) divided the dispersal of spores into three stages: 1) Abscission (the release, liberation and discharge of spores), 2) Transport (transportation by the wind, water, animal etc.) and 3) Deposition (capture) (Lonnell, 2011). Many algae, bryophytes and ferns have microscopic spores (< 50 µm) and are quickly dispersed by high wind speeds over relatively long distances (Lonnell et al, 2014). The spores of bryophytes have a diameter of 5-310 µm, but few are greater than 30 µm (Crum, 2001) and most fern spores have a diameter of 20-60 µm (Tryon 1970), like most wind-dispersed pollen grains having a diameter of also 20-60 µm (Faegri & Van Der Pijl 1979). A critical stage of wind dispersal of spores is for them to reach above the canopy where the wind speeds are higher and easier for the spores to be carried away to higher altitudes and longer horizontal distances (Lonnell et al, 2014). Another mechanism is for the spores to be lifted with warm air (thermal upheaval) (Tackenberg et al., 2003). Many plant species have spores that can tolerate longer periods of desiccation, high levels of Ultra Violet (UV) radiation and high temperature (Zanten, 1978). Under the right conditions, spores could be viable for long periods of time, e.g. bryophytes: *Blinda acuta* was viable after 4 years and *Racomitrium sudeticum* after 7 years (Crum, 2001); ferns: *Cheilanthes mysurensis* viable after 8.5 years (Wright , 1909) and *Dicksonia antarctica* viable after 22 years (Anony , 1910).

Another potential source of the high percentage of algal DNA material found in Kuwaiti soil samples could be due to the presence of biological soil crusts. Biological soil crusts are common in semi-arid and arid environments where the vegetation cover is sparse (Weber et al., 2016). Open spaces between the vegetation are usually covered in biological soil crusts associated with soil particles and cyanobacteria, algae, microfungi, lichens and bryophytes which live within, or immediately on top of the uppermost few centimetres of the soil surface (Belnap et al., 2001). Cyanobacteria and algae, bacteria and microfungi generally initiate the basic matrix of the soil crust, facilitating the colonisation of bryophytes, lichens and microfauna (Weber et al., 2016). Factors that maintain the type of biological soil crust present are various abiotic stresses including high temperature, UV, salinity, pH and low moisture (Zhang

et al., 2007). Biological soil crusts plays a major role in determining the soil structure and the morphology of the soil surface; they influence hydrological cycles and the capture and retention of resources such as soil, organic matter, seeds, and nutrient-rich dust (Weber et al., 2016). Biological crusts also have the ability to enhance the soil fertility by fixing atmospheric carbon and nitrogen (Sancho et al., 2016). Biological soil crust lacks proper field studies and documentation in Kuwait which makes it a potentially rich area for future research.

5.4.2 Metagenomics analyses using EBI pipeline

The taxonomic classification of the six metagenomics samples was studied using EBI pipeline (Figure 5.8). The observed number of OUT's (at 99 % similarity) from bacterial 16S rRNA gene sequences in all samples ranged from 1,115 to 2,925 reads (56 to 69 %) (Appendix 5.2). The phylum level taxonomy was primarily assigned to bacteria (63 % average) and archaea (0.9 %). At the phylum level, the bacterial sequence reads are mainly represented by Actinobacteria and Proteobacteria; archaea reads are only represented by Crenarchaeota (Appendix 5.2). At the family level, the phylum Actinobacteria is represented by Actinomycetaceae, Actinosynnemataceae, Cellulomonadaceae, Frankiaceae, Intrasporangiaceae, Microbacteriaceae, Micrococcaceae, Nocardiaceae, and Streptosporangiaceae; the phylum Proteobacteria is represented by Caulobacteraceae, Beijerinckiaceae, Eruthrobacteraceae, Hyphomicrobiaceae and Phyllobacteriaceae. Crenarchaeota is represented by only two families: Cenarchaeaceae and Nitrosophaeraceae.

Actinobacteria, gram-positive bacteria (known for producing antibiotics) are of significant important influence on human health and plays an important role towards their contribution to soil systems by behaving like fungi and help in decomposing organic matter of dead organisms and providing nutrients to new plants (Lewin et al., 2016). The gram-negative bacteria represented by Proteobacteria include a wide variety of human and plant pathogens, such as Escherichia, Salmonella, Vibrio, Helicobacter, and many other notable genera (Eckburg et al., 2005). Other gram-negative bacteria are free-living and include agriculturally important bacteria capable

of colonising the rhizosphere and responsible for nitrogen fixation in a symbiotic relationship with plants (Mendes et al., 2013). The phylum Crenarchaeota represents archaea; known to play an important ecological role in the nitrogen and carbon cycles in the soil and are found colonising young plant roots at high frequency (Simon et al., 2000).

The EBI metagenomic analyses performed on the Kuwaiti soil eDNA extracts gives us an insight of the microbiome present in the study sites. The presence of bacterial sequences in the soil samples, belonging to the Phylum: Actinobacteria and Proteobacteria, are well documented to be the most abundant bacteria in biological soil crusts (Liu et al., 2017; Blay et al., 2017). Also, the archaeal sequences represented by Crenarchaeota are reported as an abundant member of biological soil crust communities across large-scale arid lands in North America (Soule et al., 2009). Thus, indicates that the Kuwaiti soil samples at both sites are represented by microorganisms (represented by bacterial and archaeal sequences) that support the initiation of the basic matrix for biological soil crust community alongside with other plants represented by algae and bryophytes (biological soil crusts discussed earlier).

5.4.3 Comparison of the metagenomics eDNA soil sampling depths

It is well documented that the seed bank density is higher in the upper soil layers (< 5 cm) and the seed density decrease as the depth increase (Fenner & Thompson, 2005). Many ecologists sampled up to 2 cm in depth to study the viability and germination of seeds (Nelson & Chew, 1977; Al-Yemeni et al., 2000; Marone et al., 2004; Reichman, 1984; Marone & Horno, 1997), while others sampled up to 10 cm in depth (Bakker et al., 1996; Leicht-Young et al., 2008), and fewer studies sampled above 10 cm in depth (Erenler et al., 2010; Gross, 1990). In this study the experimental design was set to study and compare between the upper soil layers (≤ 5 cm) with the deeper layers (10-15 cm) of two sites, fence protected rich in plant diversity and highly grazed open desert area poorly representing the flora.

Studying the top soil layer (0-5 cm) of the two sites together using total genomic DNA extractions revealed that although at site 1 (fence protected area) the diversity of

plants above ground is rich with an advantage of having plenty of seasonal DNA material floating above ground and in the top soil layer, DNA material (sample MA6) sequenced from site 2 (a highly degraded open desert area) represented a high sequence matches at species, family, and order-level (Table 5.14 - 5.16). It was not expected to find plenty of DNA material in the soil from a highly disturbed site, represented by very few plants above ground. This indicates that the sequenced environmental DNA material, whether it is representing a viable, decayed or dead DNA source (e.g. spores, seeds, pollens and fragments of leaves, roots and stems), not being mirrored by the current situation of plant diversity found above ground, and rather it is reflecting the historical path of DNA material that once existed in the study area and/ or could possibly be due to the accumulation and deposition of long distance wind transported DNA material.

A comparison of the two soil sampling layers (between 0-5 and 10-15 cm) of the two sites was not possible due to samples MA7 and MA8 (with very low DNA yield) did not pass the minimum requirements for Illumina sequencing (Table 5.3). Although the samples were concentrated, the very low DNA yield generated for samples MA7 and MA8 indicates that as we go deeper into the soil (> 5 cm) of disturbed sites, we generate very low DNA yield. Thus, indicates that most of the DNA material for site 2 is floating on the top surface layer of the soil (≤ 5 cm depth), e.g. from all the metagenomics samples collected from site 1 and site 2, sample MA6 showed the highest species richness, 14 species, below ground (≤ 5 cm depth) compared with the above-ground diversity with only 4 species present during the field survey (Table 5.16).

Comparing the top soil sampling layer with the deeper layer from site 1 only, the upper layer (≤ 5 cm) was represented by fewer sequence reads and resulted in the highest percentage of sequence matches at species and family-level using Angiosperm-NCBI database (Table 5.15 and Table 5.16). For the order-level using the complete plastid genome database, the percentage was higher for the deeper layers (10-15 cm) compared with the upper soil layer, this could be due to the high presences of algal DNA material in the deeper layers (Table 5.14). Moreover, the BLASTn matches against the Kuwaiti DNA database showed that the deeper soil

layers represented the sequence matches better than the top layer (Table 5.17). This is due to the generation of a high number of trimmed and assembled sequence reads found in the DNA soil extracted from the deeper soil layers (MA3 and MA4). This could be due to the accumulation of large amount of plant DNA material (e.g. pollens, seeds, fragments of roots and leaves) are deposited and preserved over extended period of time in the deep soil layers.

5.4.4 Metabarcoding

The markers used for metabarcoding method, *rbcL* (500 bp) and ITS2 (300-460 bp) showed reliable matches to the Kuwaiti DNA database. The blast of *de novo* clustered sequence reads (matches ≥ 99 % ID) for both markers of the 12 samples BLASTn against the local database resulted in a large number of matches, 139 species, which represents about 35 % of the flora of Kuwait (Table 2.3). The size of *rbcL* sequence fragments amplified from soil samples range from 100 to 466 bp; ITS2 from 103 to 488 bp across all twelve samples. An average of 40 % of the sequence fragments for both regions was represented by matches > 250 bp of paired-end reads ≥ 99 % ID using Kuwaiti DNA database. The analyses were not restricted to full-length reads; the matches were filtered at ≥ 99 % ID. Few matches (5-16 across all metabarcoding samples) represented by *rbcL* and ITS2 amplicons were longer than 500 bp which returned with matches < 99 % ID. Also, amplicons ≥ 500 bp long resulted in matching percentage identity lower than 99 % ID and ranged from 78 to 82 % ID across all samples.

A large portion of the 139 species belonging to common families of the flora of Kuwait, represented by Asteraceae, Amaranthaceae, Brassicaceae, Caryophyllaceae, Fabaceae, and Poaceae (Table 5. 20, Table 5.21, and Table 2.3). Similar plant families matching Kuwaiti flora were found in a study by Parducci et al., (2013), who conducted molecular metabarcoding (using *trnL*) and pollen-based vegetation analysis in lake sediments from central Scandinavia. Parducci et al., (2013) concluded that metabarcoding analysis provides a complementary, but not an alternative tool to pollen analysis for investigating the past flora (Parducci et al., 2013), since pollen

analyses revealed a large number of taxa (46) when compared to that identified by metabarcoding with only 14 taxa, due to the short bp length of *trnL* barcodes providing a low taxonomic resolution (Parducci et al., 2013). On the other hand, Fahner et al., (2016) used alternative barcoding markers, *matK*, *rbcL*, and ITS2, along with the traditionally used marker, *trnL* P6 loop, revealed that the best taxonomic resolution at species-level was generated by ITS2 (Fahner et al., 2016).

Comparing the results of plant diversity found below ground with the above ground level, using *rbcL* and ITS2 barcodes, in total 19 species from the above ground level matched below ground diversity and represented by *Anisosciadium lanatum, Asphodelus tenuifolius, Atractylis carduus, Brassica tournefortii, Convolvulus oxyphyllus, Cuscuta planiflora, Gymnarrhena micrantha, Gypsophila capillaris, Helianthemum lippii, Heliotropium bacciferum, Koelpinia linearis, Launaea mucronata, Plantago boissieri, P. ovata, Rhanterium epapposum, Lomelosia palaestina, Schismus barbatus, Senecio glaucus* and *Stipa capensis* (Tables 5.11, 5.20 and Table 5.21). Most of the matches belonging to common plants of the flora and found present above ground during the field survey at Um-Neqa (study area) and are also common during the spring season throughout Kuwait, except for *Lomelosia palaestina,* considered to be a rare species and only spotted growing in Al-Abdaly, North of Kuwait City (Table 5.11).

Three widespread plants present across most metabarcoding samples and found above and below ground level are represented by *Cuscuta planiflora* (common parasitic plant belonging to Convolvulaceae), *Gymnarrhena micrantha* and *Senecio glaucus* (common plants of the flora belonging to Asteraceae found throughout Kuwait). Twelve species (listed in Table 5.27) were detected growing above ground during the field survey but are not represented in the data sets, e.g. *Hammada salicornica,* a perennial shrub belongs to Amaranthaceae family, well represented above ground with more than 50 % vegetation coverage in site 1, was not detectable in any of the eDNA samples (Table 5.17, Table 5.20 and Table 5.21). *Hammada* and other species remained undetected listed in Table 5.27, could be an example of plants with highly degradable eDNA material and not leaving a traceable amount of PCR amplification or direct sequencing. Another possibility for some species remaining undetectable is

either due to low numbers of sequence reads or absence in DNA reference libraries. In a similar study, Bell et al. (2017) could not identify 21 sequences with *rbcL* or ITS2 represented by *Ambrosia* spp., *Artemisia* spp., *Populus* spp., *Poa* spp., *Xanthium* spp and *Zea mays* (Bell et al., 2017).

The BLAST match using Angiosperm-NCBI database, matched to 363 species of *de novo* sequences across all metabarcoding samples, only 23 species matched the flora of Kuwait, and two species are representatives of the above ground flora at the study site. Hiiesalu et al. (2012) studied grassland plant species richness belowground and compared it to aboveground diversity using the chloroplast *trnL* (UAA) intron; the results showed below ground species richness was two times greater than above-ground abundance (Hiiesalu et al., 2012). Furthermore, nine species were detected only below ground during the study period (*Solidago* spp., *Artemisia* spp., *Turritis* spp.) and later in the season had been detected growing above ground (Hiiesalu et al., 2012). Such findings indicate that some species stay dormant for a period of time and remain undetected by aboveground field surveys.

The BLASTn matches showed a high number of matches at a lower taxonomic level (genus-level) and present in the flora of Kuwait such as *Allium, Artemisia, Astragalus, Lilium, Salsola, Salvia,* and *Silene* (Table 5.25 and Table 2.3). Matches represented by a high number of sequence reads are found belonging to Asteraceae family and represented by *Aster glehnii, Atractylodes lancea, Brachylaena ilicifolia*, and *Metalasia adunca*. Other matches are represented by *Delosperma cooperis* (Aizoaceae), *Trianthema clavatum* (Aizoaceae) and *Mertensia virginica* (Boraginaceae) (Table 5.25).

5.4.5 Comparison of metagenomics and metabarcoding

The level of taxonomic classification varied and each technique identified a different number of taxa. At the species level, using Kuwaiti DNA database, metagenomics samples only matched ITS2 sequences, while metabarcoding samples resulted in a total match to139 species represented by both *rbcL* and ITS2 sequences. Considering

the soil layers, the results from both approaches showed better DNA material representing plants are extracted from soil collected from the upper surface layer (0-5 cm) than the deeper layer (10-15 cm), likely due to new DNA sources accumulating seasonally on the top surface of the soil, which makes it rich in DNA material.

Metagenomics approach is more reliable when the aim of the project is to identify a wider range of organisms present in eDNA samples, while the metabarcoding approach is more focused on the identification of targeted organisms using PCR-based markers. A major drawback of metagenomics and metabarcoding is that both approaches are highly dependent on the available databases and limited by the amount of information that is present in these repositories. Sequences that do not have any similarity in a known database are believed to be 1) a consequence of sequencing errors and/ or reflect the inaccuracy of gene prediction tools, or 2) truly novel genes that have no sequence to known genes and may share higher order similarity in the form of protein folds (Oulas et al., 2015; Thomas et al. 2012).

In this study, the sequence BLASTn results against the Kuwaiti DNA database produced by metagenomics (Table 5.17) matched with broader reads generated by metabarcoding approach (with sequence reads varied from 100 to 1000) (Table 5.20 and Table 5.21). For example, *Spergularia marina* present across five metagenomic samples (except MA2) matching to ITS2 barcode sequences, was found in six out of twelve metabarcoding samples at lower sequence reads matching *rbcL* barcodes only. Raw sequence reads for *rbcL* varied from 144/ individual for sample MAS5 to 1558/ individual for sample MAS7. The results from metabarcoding analyses suggest that PCR-based approaches can amplify most plants identified by metagenomics using *rbcL* and ITS2 barcodes (Table 5.17, Table 5.20, and Table 5.21).

For metagenomics, the low read counts matching ITS2 database of Kuwaiti plants (no matches for *rbcL* database), caused some species to remain undetected because they only had matches to one barcode region, ITS2. Matches of metagenomic samples to *rbcL* Kuwaiti plants database failed to detect any species, possibly the chloroplast DNA region represented by *rbcL* genes across all samples was highly degraded (*rbcL*

did not show any resolution), and only metabarcoding samples generated by ITS2 sequence reads matched the Kuwaiti DNA database.

When comparing the results generated from metagenomics and metabarcoding approaches using the above ground plant diversity, metabarcoding approach produced 19 species matches while metagenomics matched only 1 species, *Gymnarrhena micrantha* (Tables 5.27). This indicates that metabarcoding approach is more reliable when comparing the plant diversity of below ground with that found above ground level. The main issue with metagenomics is the low read counts that detected only 19 species represented by ITS2 barcode and possibly other species remained undetected when blasted against *rbcL* Kuwaiti DNA database, basically due to the level of DNA degradation (Table 5.17); for metabarcoding the matches generated from few hundreds to several thousand reads (Table. 5.20 and Table 5.21). A similar finding was observed by Srivathsan et al. (2015) while investigating monkey diet analysis and comparing the effectiveness of metagenomics and metabarcoding results. They used two faecal samples to characterise the diet of two monkeys (*Pygathrix nemaeus*) that were fed known foliage, fruits, vegetables and cereals and produced 74 and 67 million paired reads for these samples using Illumina HiSeq platform (Sirvathsan et al., 2015). The sequences were matched against plant database containing all angiosperm barcodes in Genbank. The results were compared with metabarcoding using *trnL* P6 loop gene region. Metagenomics identified seven and nine of the likely 16 diet plants while six and five were identified by metabarcoding (Sirvathsan et al., 2015). A similar comparison was performed in this study, except that metagenomics identified only one species of the 31 species present above ground while metabarcoding identified 19 species (61 %).

Plants found in the eDNA soil samples that resulted in high sequence matches to Kuwaiti DNA database using both NGS methods, although not present above ground during the field survey are represented by *Astragalus sieberi, Spergularia marina, Trigonella stellate, Loeflingia hispanica* and *Polycarpon tetraphyllum* (Table 5.17, Table 5.20 and Table 5.21). The DNA source representing the sequences could come from the extraction of seeds which remained dormant and unable to germinate due to a combination of environmental factors, such as rain, light and temperature, which

favoured the germination of other non-dormant seeds. It is well known that seeds of desert plants germinate only after a threshold amount (10-15 mm) of rainfall (Gutterman, 2012), and only under favourable temperatures appropriate for growth, ranging from 25 to 30 °C (Lai et al., 2016). Most common plants remain dormant in the soil for very long period of time and could germinate when exposed to favourable environments, e.g. Beal's study of soil seed longevity showed that after 120 years burial in moist and well aerated sand, 23 seeds of *Verbascum blattaria* germinated and produced normal plants (Telewski and Zeevaart, 2002), and another study by Duvel included 107 crop and weed species which lasted 39 years (Toole and Brown, 1946). Other sources of DNA material representing the sequences could be typically the remains of plant tissues from previous seasons well preserved in the soil over a long period of time (e.g. roots, leaves, woody parts, etc.).

Historically, plant community distribution comes from the reconstruction of ecosystems using palaeoecological records such as fossil pollen and macrofossils accumulated in lakes and peat sediments (Birks, 2001). Plant macrofossils are represented by diaspores (spores, seeds, and fruits) and vegetative parts, such as leaves, buds, budscales, flowers, bulbils, rhizomes, roots, bark and wood. Sedimentary ancient DNA (*seda*DNA) also referred to as 'dirt DNA' has successfully been used to recover DNA traces of past environments, providing new information on former flora and faunal changes (Hofreiter et al., 2003; Haile et al., 2009; Thomsen et al., 2009), *seda*DNA and pollen reveal the composition of past vegetation in Late Quaternary permafrost sediments (Pedersen et al., 2013; Parducci et al., 2013; Zimmermann et al., 2017). Mosses and occasionally liverworts can be found as macrofossils (Souto et al., 2017), lichens are also preserved as fossils (Jahren et al., 2003), and marine algae (Rulin and Lifu, 1985). Under favourable conditions, well-preserved plant macrofossils could provide a good source of DNA material for NGS molecular analysis.

The comparison between metagenomics and metabarcoding results provide important information that can be summarized as follows: 1) the metagenomics and metabarcoding analyses resulted in few overlapping species/taxon identification, 2) the two analyses each detected likely a portion of the total DNA material extracted

from the soil, 3) the metabarcoding analysis allowed better identification at higher taxonomic level than metagenomics analysis, 4) there was a substantial difference between sequence matches to Angiosperm-NCBI database obtained from metabarcoding vs metagenomics approaches, 5) the molecular analysis failed to detect some major taxa (e.g. *Hammada salicornica, Astragalus schimperi, Arnebia decumbens, Moraea sisyrichium*) found above ground at the study sites and not represented by both NGS techniques, and 6) Metagenomics analyses returned with few reads, representing 24 species, matching ITS2 only of the Kuwaiti DNA database, while metabarcoding returning with a total of 139 species matching both barcodes *rbcL* and ITS2 barcodes.

The advantage of using metagenomics approach is its ability in identifying different organisms (e.g. plants, fungi, microbes, nematodes) with wider biodiversity coverage, while metabarcoding data generates better coverage for the identification of a specific group of organisms by choosing a reliable DNA marker. Overall, in this study metabarcoding approach remains reliable when a plant species is to be picked from eDNA samples and compared with currently known above ground flora from a small number of distantly related species.

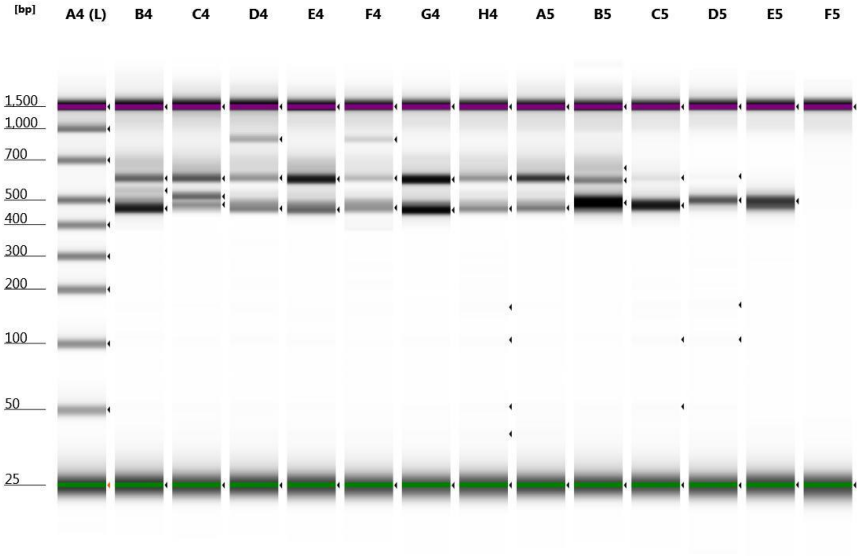
5.4.6 Conclusion

Metagenomics and metabarcoding approaches can significantly complement field-based research in Kuwait. It may help the advancement of field surveys by rapidly generating molecular data where urgent conservation intervention is required. Metabarcoding approach using PCR-based markers can efficiently identify DNA of unknown plant remains collected from soil samples of local origin by blasting them against the Kuwaiti DNA database. For metagenomics approach to avoid a small number of reads matching the Kuwaiti DNA database (since Kuwait DNA database is based on two markers *rbcL* and ITS2), a reliable mitochondrial and chloroplast reference genomes library is required to improve the resolution of identification to species level. Although in this study metagenomics data shed light on a wider ecological knowledge by identifying the presence of microorganisms such as bacteria

and archaea; also, green plants represented by algae, bryophytes, and ferns in the soil samples which require further investigation to understand and conserve the biodiversity of Kuwait.

Appendix 5.1 Metabarcoding QC report generated by Edinburgh Genomics

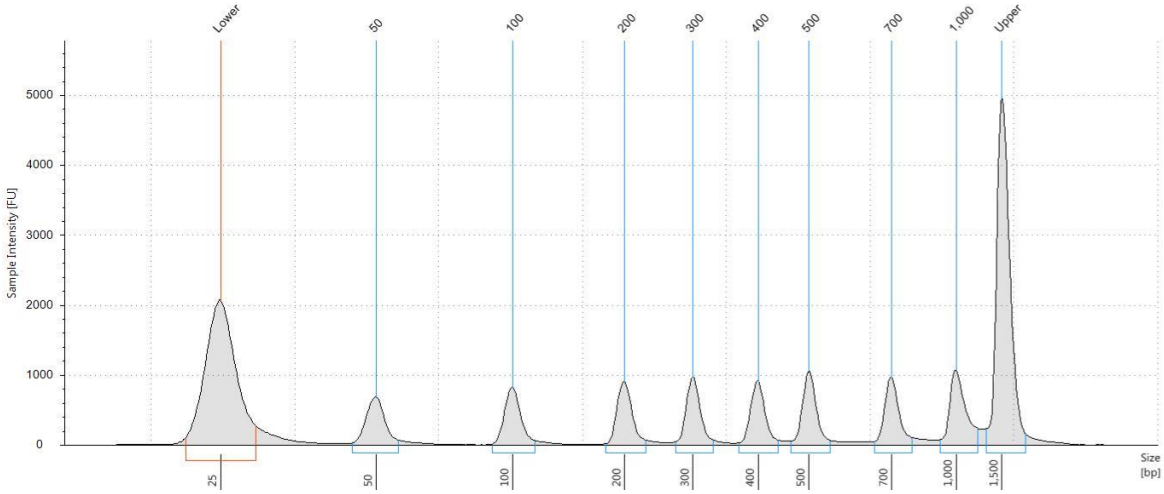
Gel Image



Sample Info

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
A4	11.2	Ladder		Ladder
B4	7.11	10342KC0001-MAS1		
C4	5.61	10342KC0002-MAS2		
D4	4.14	10342KC0003-MAS3		
E4	7.60	10342KC0004-MAS4		
F4	3.70	10342KC0005-MAS5		
G4	11.3	10342KC0006-MAS6		
H4	3.85	10342KC0007-MAS7		
A5	4.94	10342KC0008-MAS8		
B5	12.9	10342KC0009-MAS9		
C5	5.87	10342KC0010-MAS12		
D5	2.79	10342KC0011MAS13		
E5	4.74	10342KC0012-MAS19		
F5		blank		

A4: Ladder



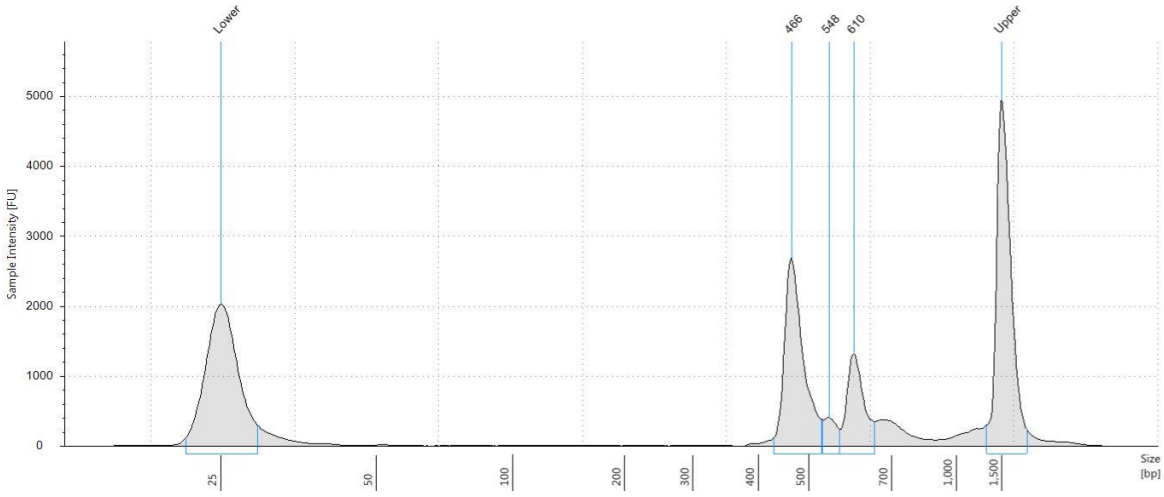
Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
A4	11.2	Ladder		Ladder

Peak Table

Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	6.29	-	387	-		Lower Marker
50	1.28	-	39.4	11.41		
100	1.33	-	20.5	11.86		
200	1.36	-	10.5	12.15		
300	1.38	-	7.08	12.29		
400	1.32	-	5.08	11.76		
500	1.46	-	4.49	13.01		
700	1.40	-	3.07	12.43		
1,000	1.69	-	2.61	15.08		
1,500	6.50	6.50	6.67	-		Upper Marker

B4: 10342KC0001-MAS1

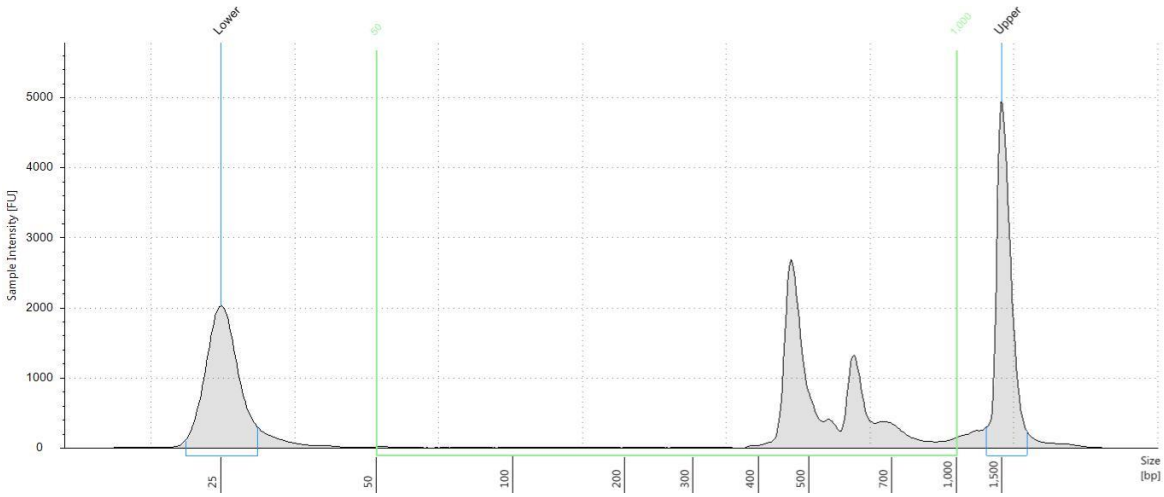


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
B4	7.11	10342KC0001		

Peak Table

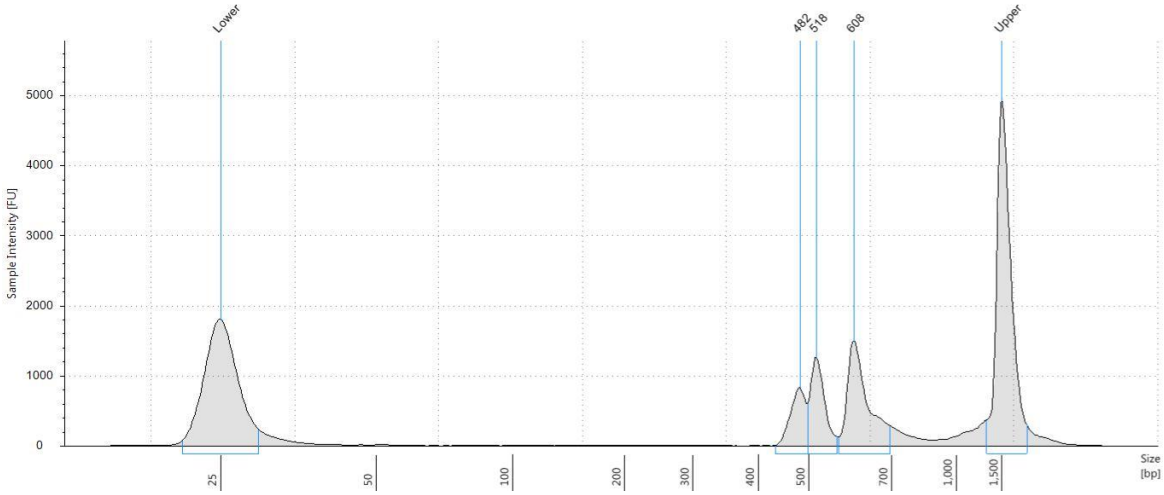
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	6.09	-	374	-		Lower Marker
466	4.61	-	15.2	64.87		
548	0.508	-	1.42	7.15		
610	1.99	-	5.02	27.98		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	551	8.53	25.6	85.48		

C4: 10342KC0002-MAS2

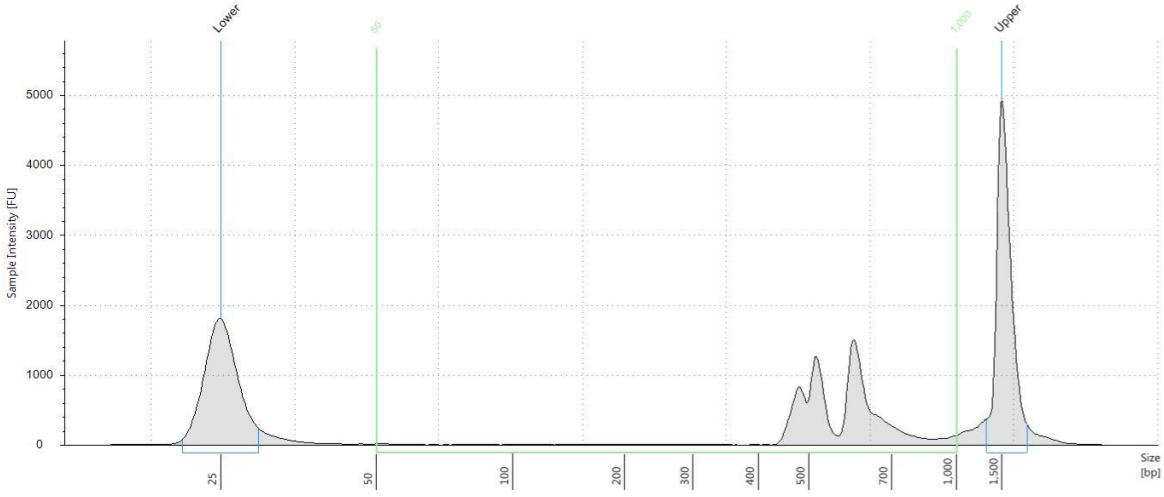


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
C4	5.61	10342KC0002		

Peak Table

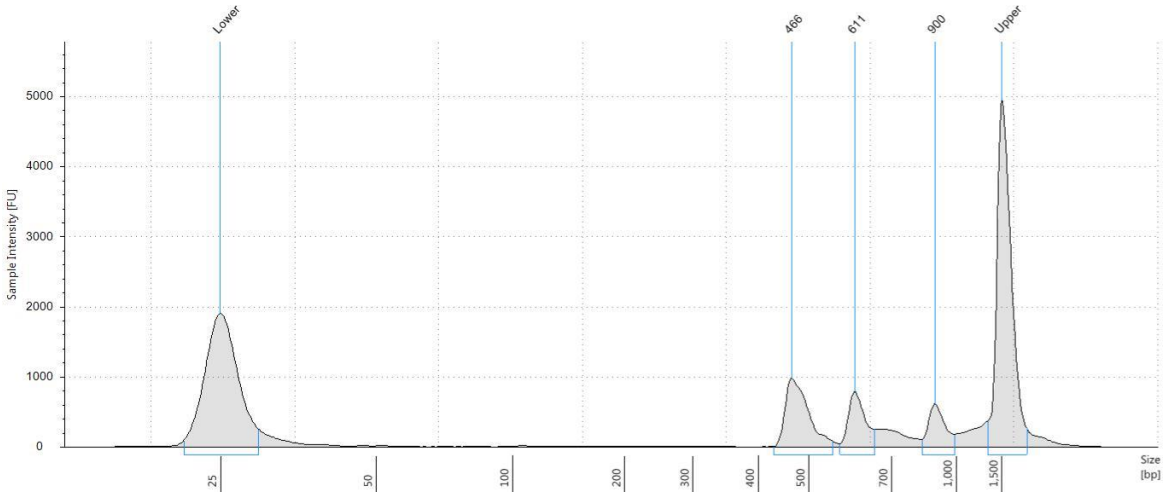
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.62	-	346	-		Lower Marker
482	1.17	-	3.72	20.78		
518	1.67	-	4.96	29.81		
608	2.77	-	7.01	49.41		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	589	6.33	17.9	81.73		

D4: 10342KC0003-MAS3

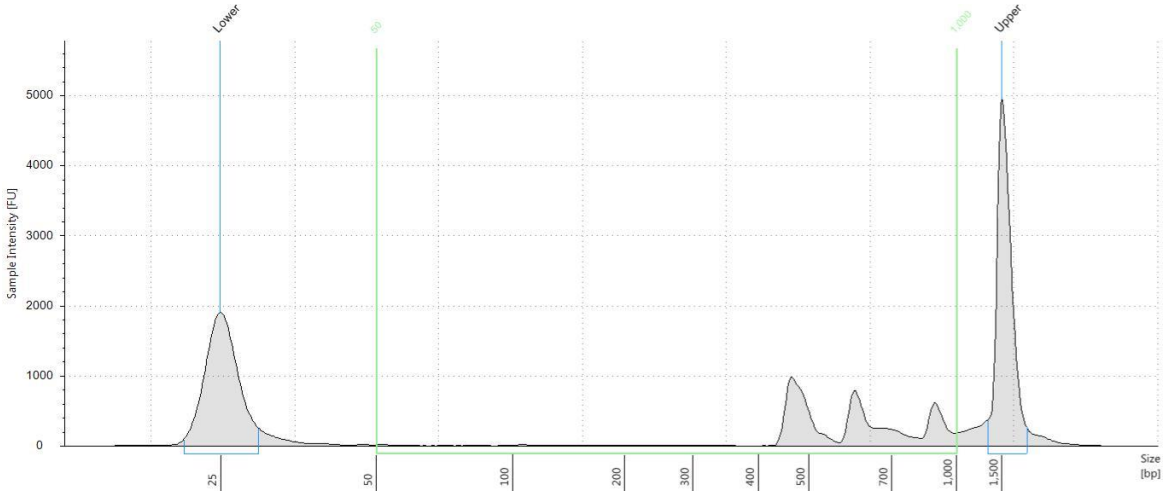


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
D4	4.14	10342KC0003		

Peak Table

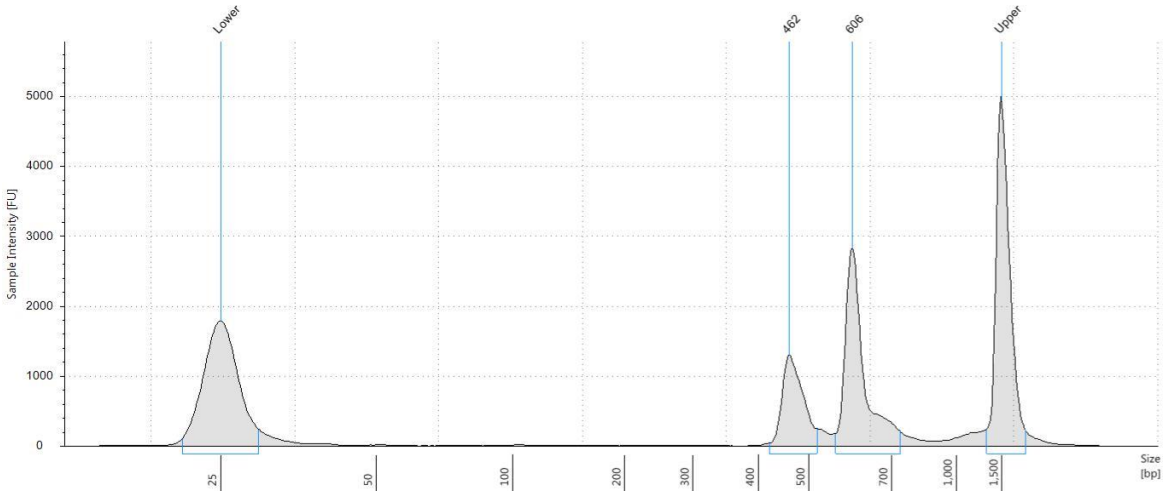
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.76	-	354	-		Lower Marker
466	2.05	-	6.76	49.45		
611	1.18	-	2.97	28.47		
900	0.914	-	1.56	22.08		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	622	4.99	14.1	77.26		

E4: 10342KC0004-MAS4

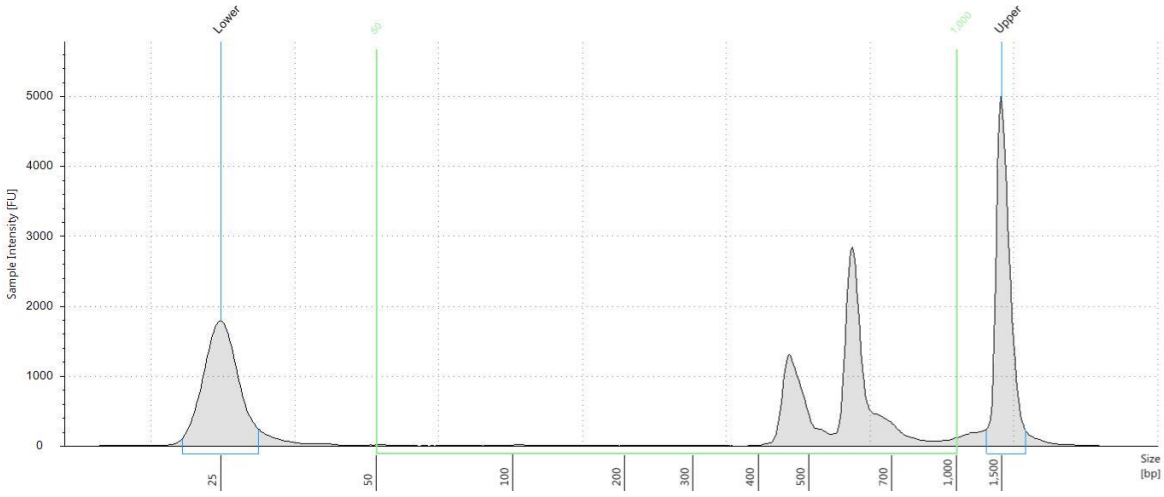


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
E4	7.60	10342KC0004		

Peak Table

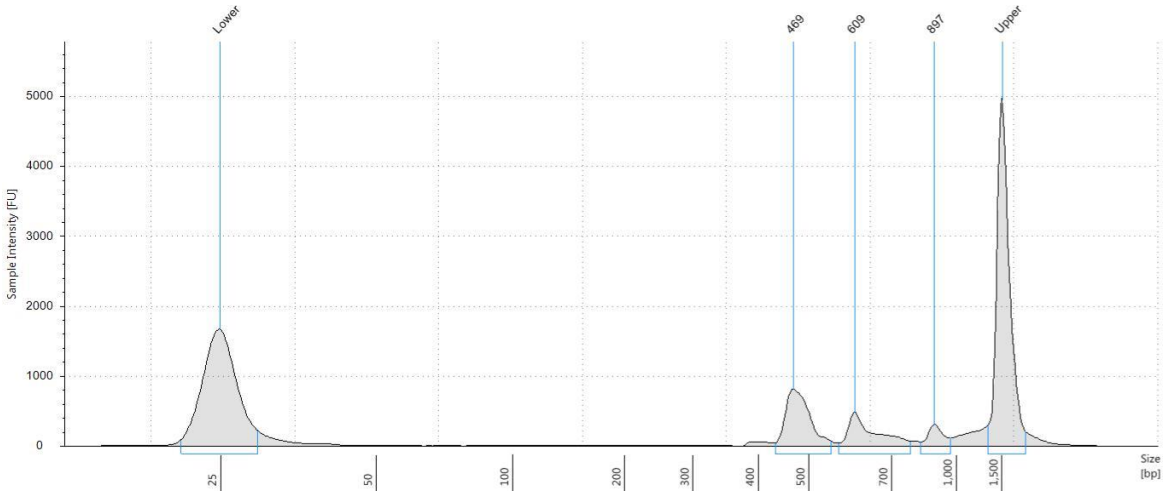
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.98	-	368	-		Lower Marker
462	2.58	-	8.61	34.00		
606	5.01	-	12.7	66.00		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	579	8.40	23.9	87.49		

F4: 10342KC0005-MAS5

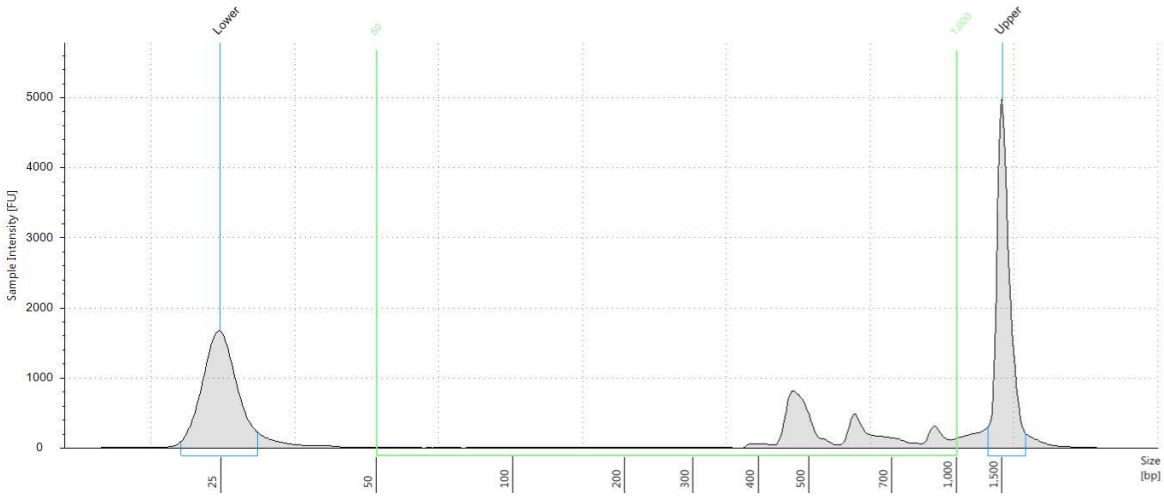


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
F4	3.70	10342KC0005		

Peak Table

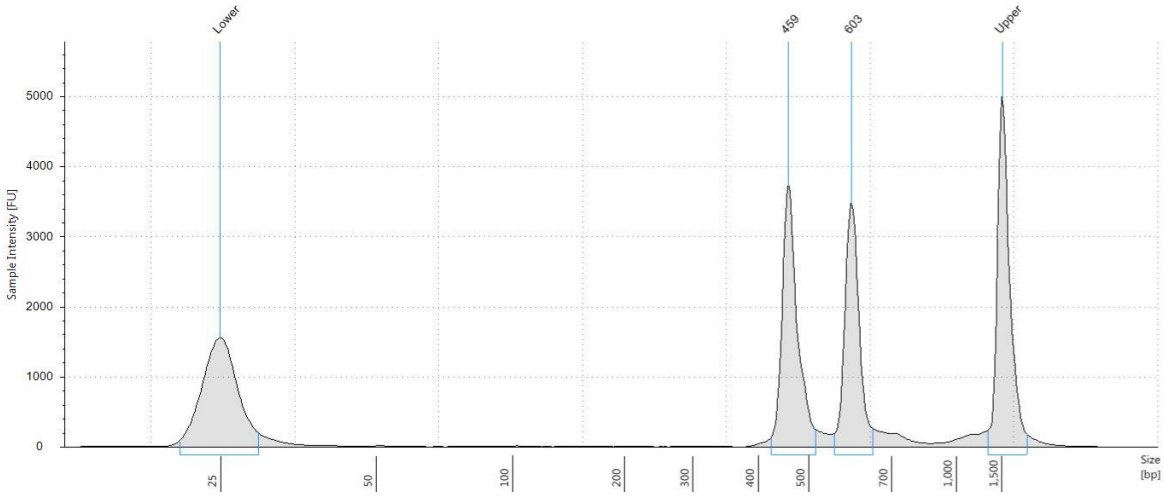
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.90	-	363	-		Lower Marker
469	2.00	-	6.56	54.05		
609	1.22	-	3.08	32.94		
897	0.482	-	0.826	13.01		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	584	4.05	12.2	75.56		

G4: 10342KC0006-MAS6

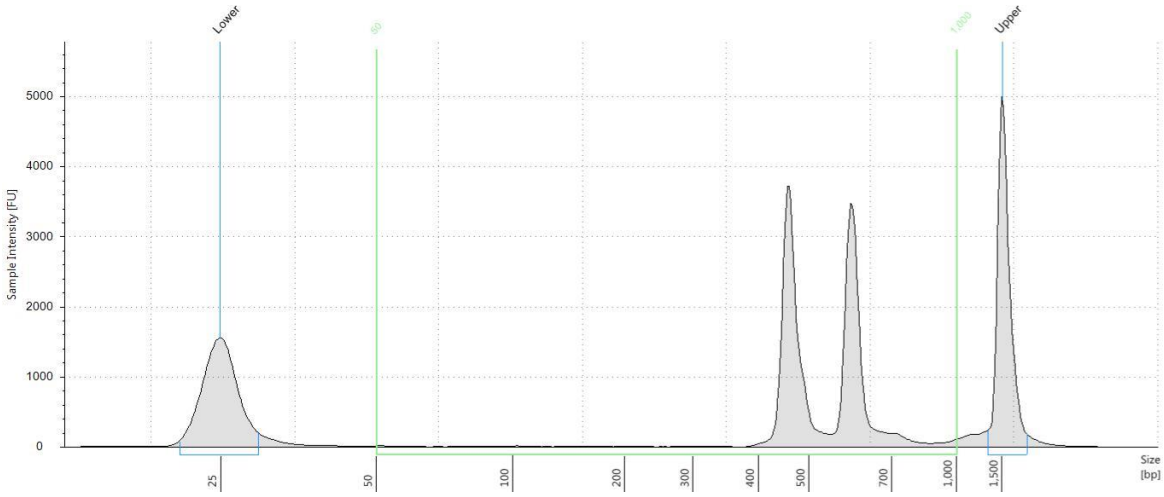


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
G4	11.3	10342KC0006		

Peak Table

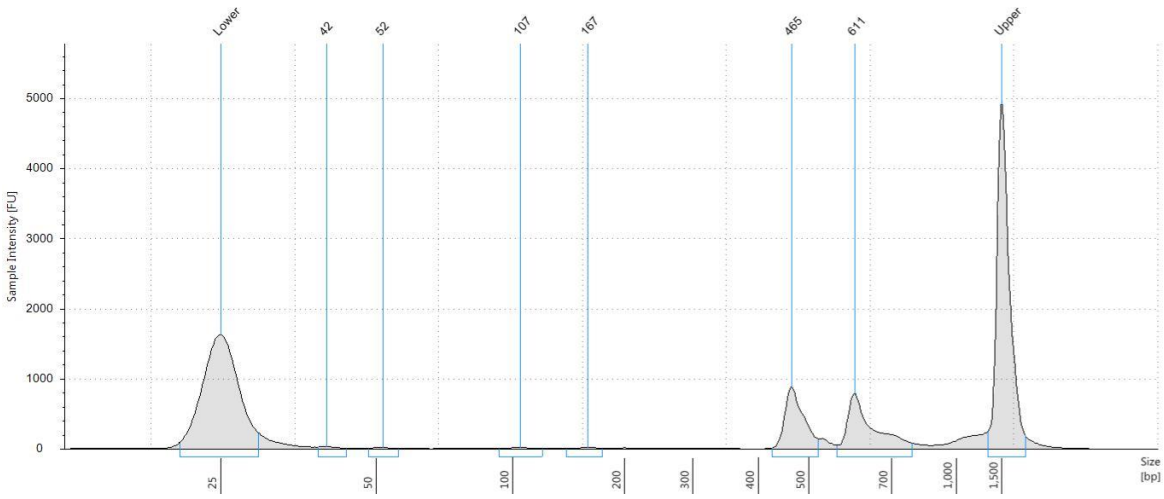
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.86	-	361	-		Lower Marker
459	6.06	-	20.3	53.48		
603	5.27	-	13.5	46.52		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	544	12.7	37.8	91.58		

H4: 10342KC0007-MAS7

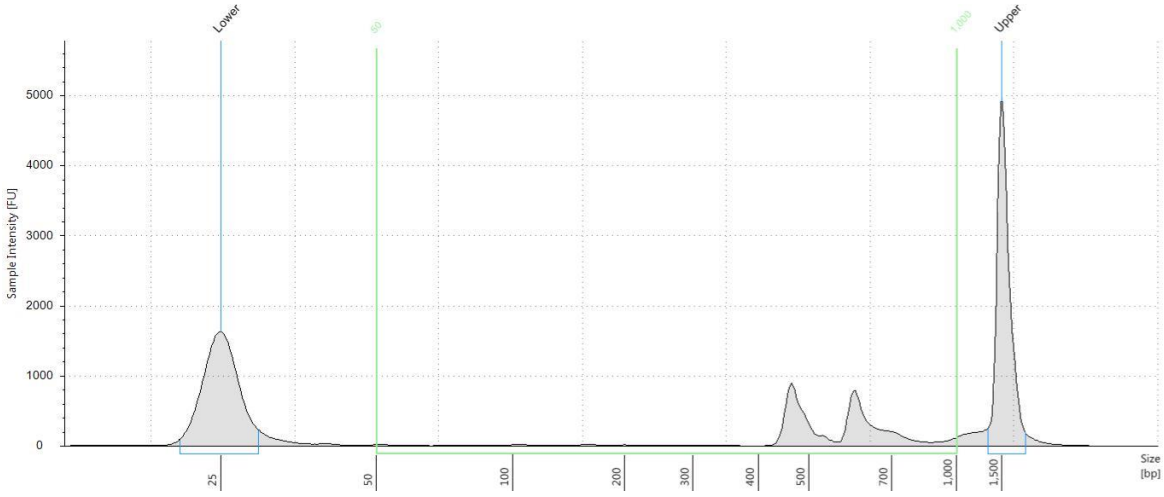


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
H4	3.85	10342KC0007		

Peak Table

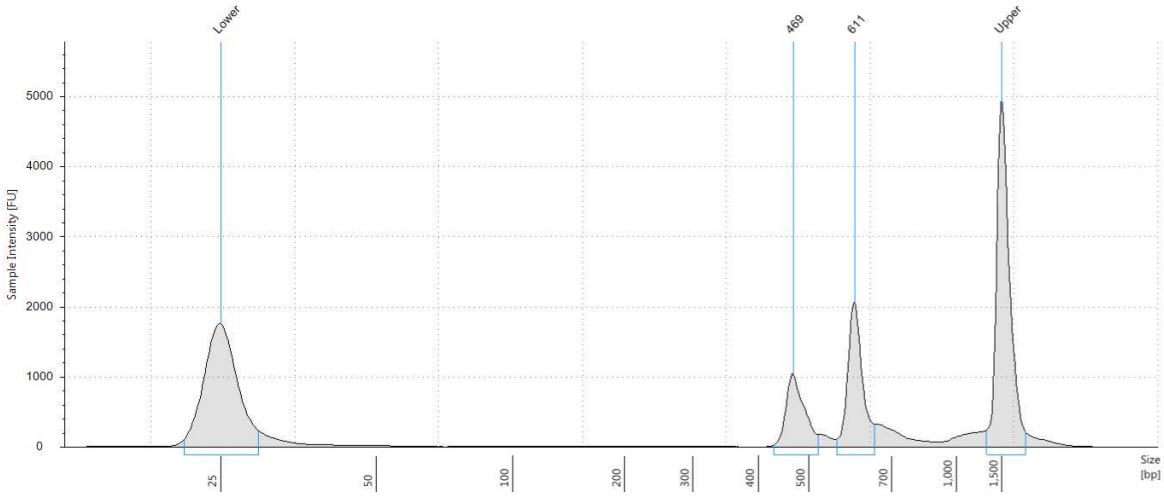
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	6.14	-	378	-		Lower Marker
42	0.0547	-	2.01	1.42		
52	0.0308	-	0.906	0.80		
107	0.0352	-	0.507	0.92		
167	0.0363	-	0.334	0.94		
465	1.72	-	5.68	44.63		
611	1.97	-	4.97	51.29		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	568	4.31	13.7	78.10		

A5: 10342KC0008-MAS8

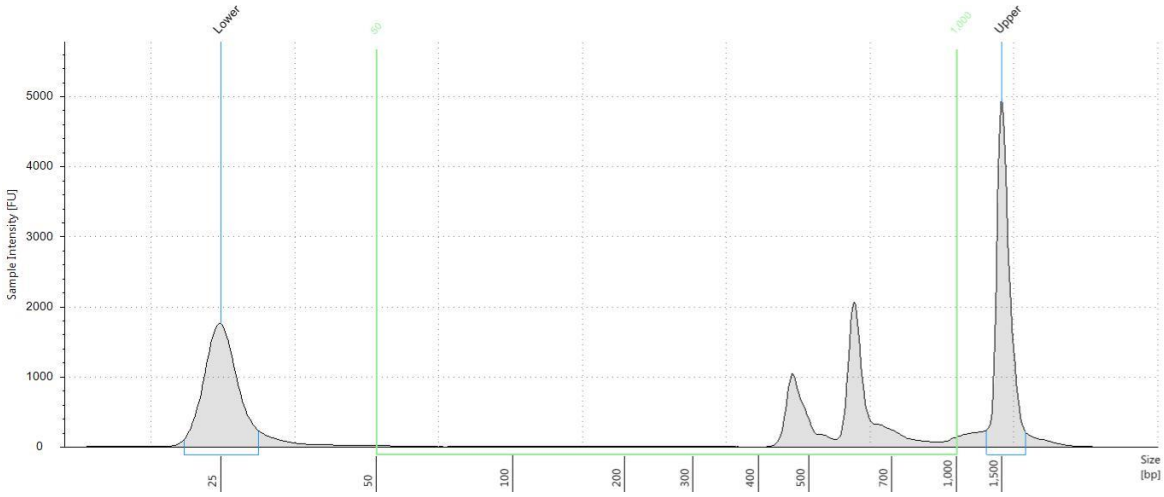


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
A5	4.94	10342KC0008		

Peak Table

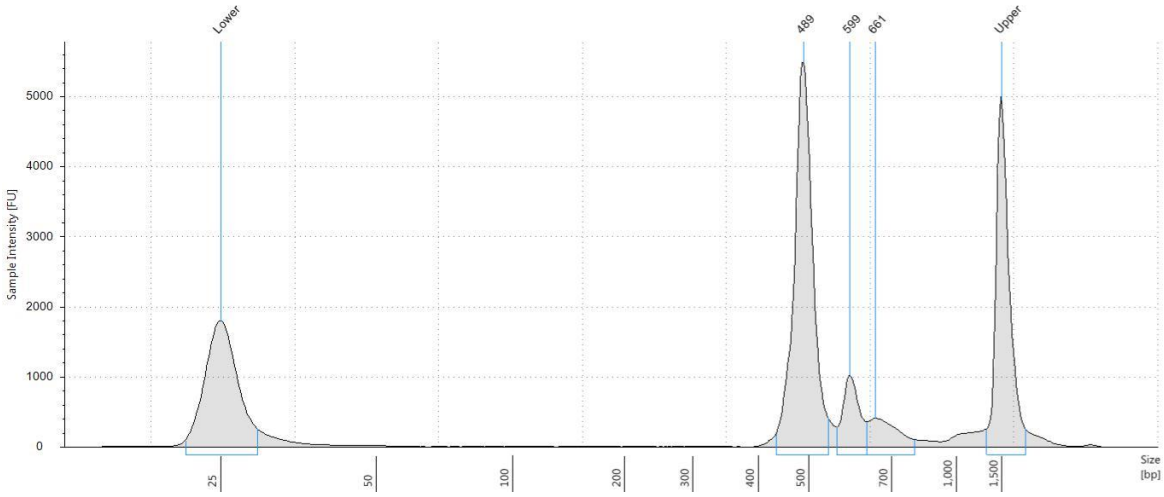
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.96	-	367	-		Lower Marker
469	1.94	-	6.37	39.26		
611	3.00	-	7.57	60.74		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	585	6.32	18.6	81.55		

B5: 10342KC0009-MAS9

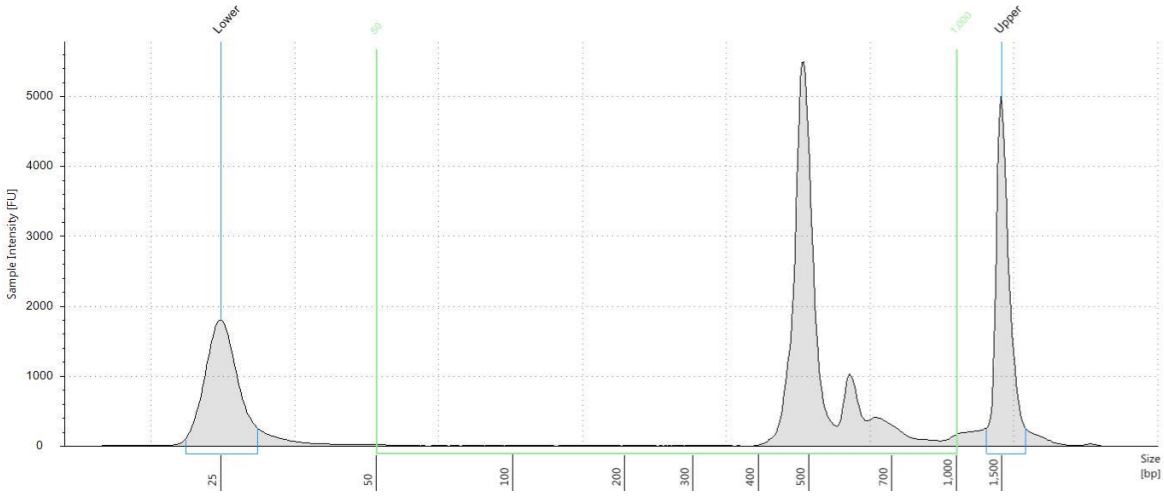


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
B5	12.9	10342KC0009		

Peak Table

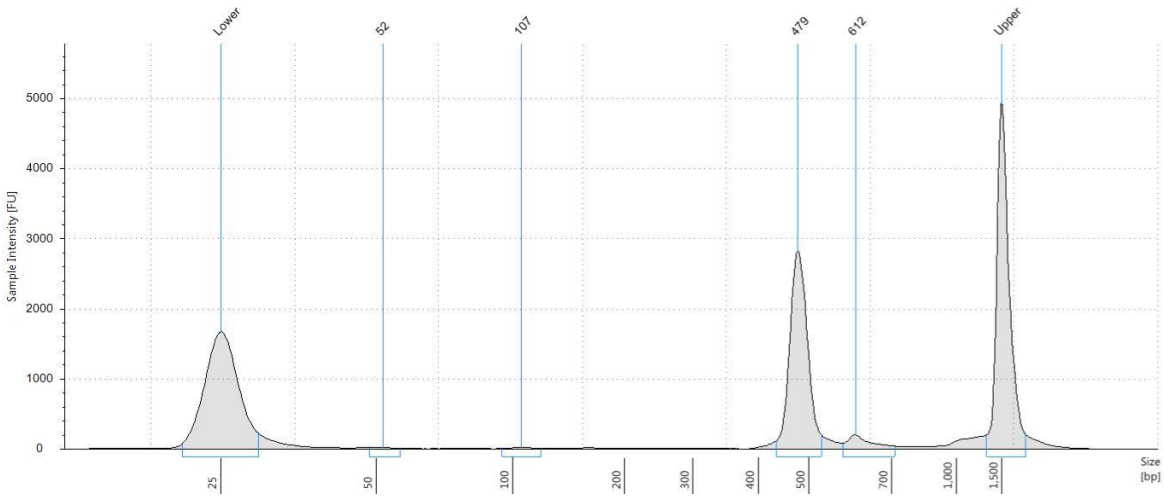
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.88	-	362	-		Lower Marker
489	10.0	-	31.6	77.98		
599	1.65	-	4.24	12.83		
661	1.18	-	2.76	9.19		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	531	13.6	41.1	89.79		

C5: 10342KC0010-MAS12

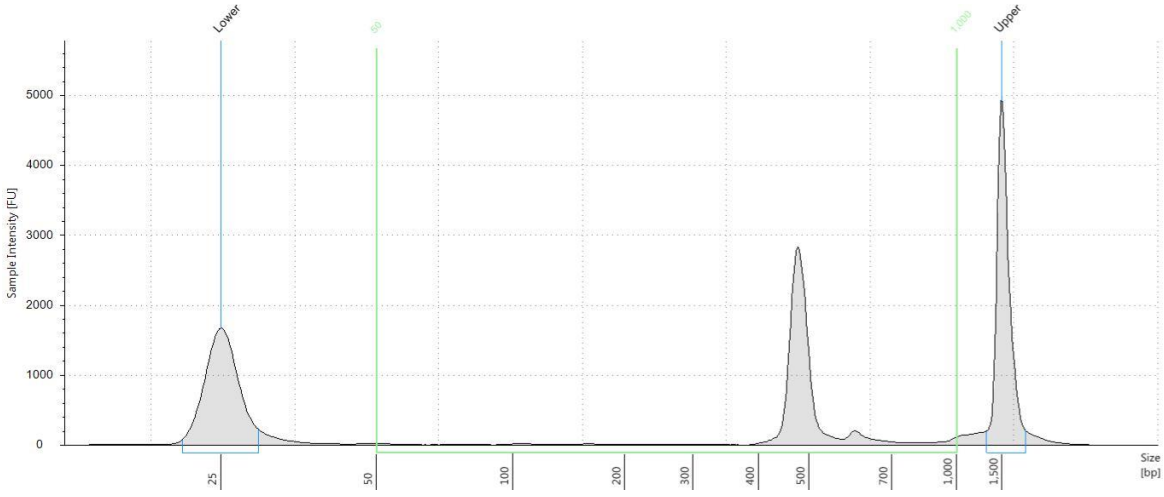


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
C5	5.87	10342KC0010		

Peak Table

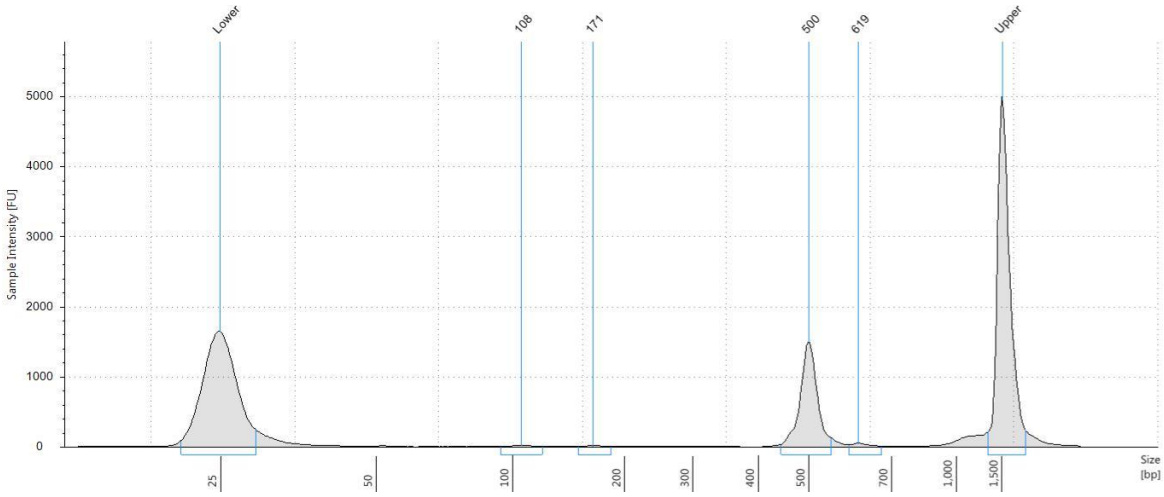
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	6.03	-	371	-		Lower Marker
52	0.0403	-	1.19	0.69		
107	0.0417	-	0.597	0.71		
479	5.31	-	17.1	90.47		
612	0.477	-	1.20	8.13		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	499	6.50	21.9	83.64		

D5: 10342KC0011-MAS13

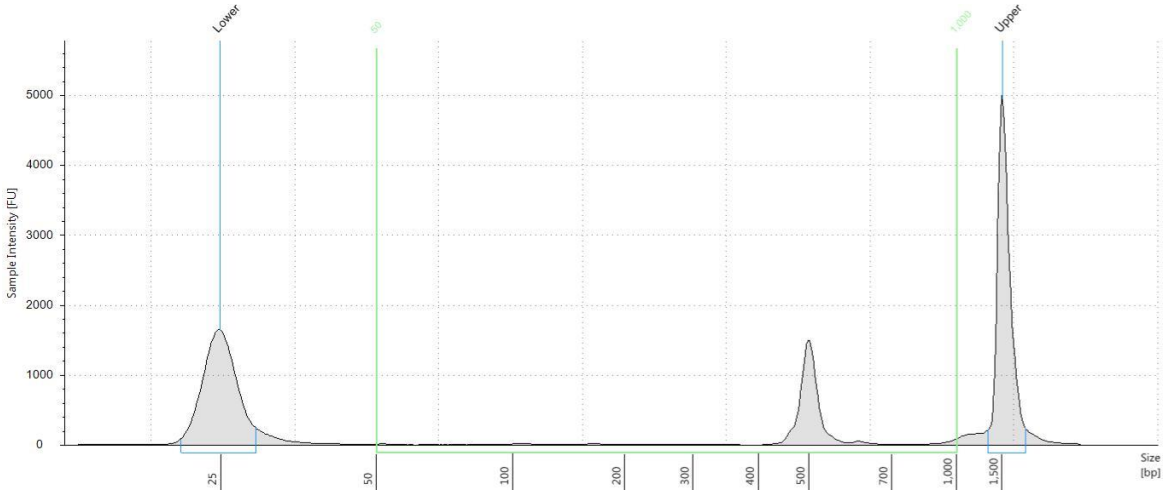


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
D5	2.79	10342KC0011		

Peak Table

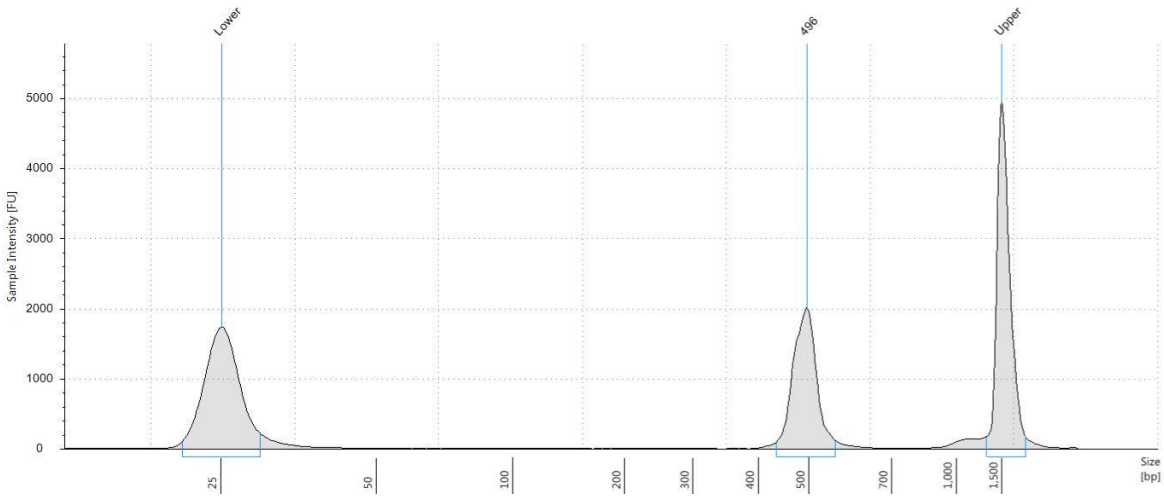
Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	5.88	-	362	-		Lower Marker
108	0.0409	-	0.584	1.47		
171	0.0302	-	0.271	1.08		
500	2.63	-	8.10	94.28		
619	0.0883	-	0.219	3.17		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	504	3.08	10.7	70.38		

E5: 10342KC0012-MAS19

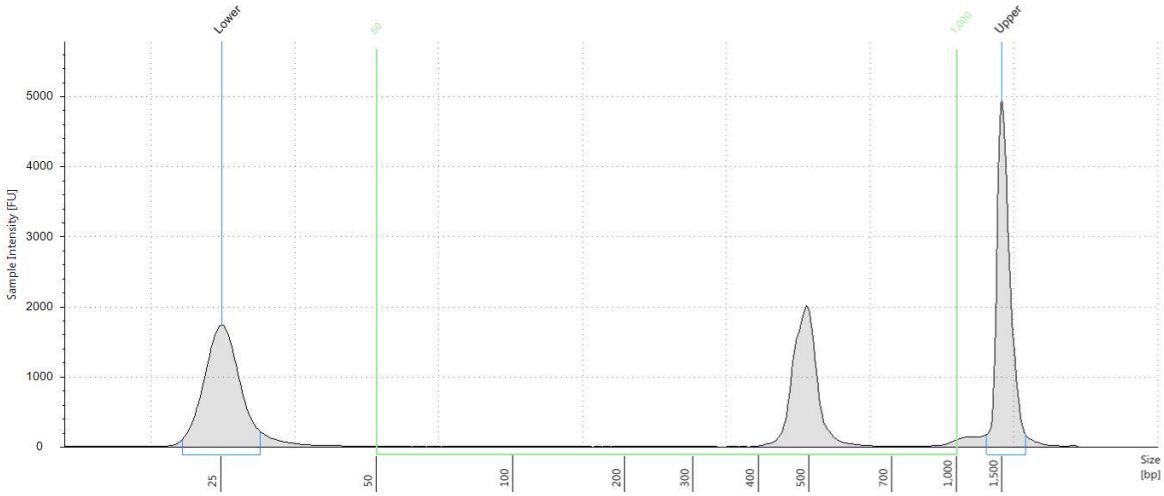


Sample Table

Well	Conc. [ng/µl]	Sample Description	Alert	Observations
E5	4.74	10342KC0012		

Peak Table

Size [bp]	Calibrated Conc. [ng/µl]	Assigned Conc. [ng/µl]	Peak Molarity [nmol/l]	% Integrated Area	Peak Comment	Observations
25	6.17	-	380	-		Lower Marker
496	4.74	-	14.7	100.00		
1,500	6.50	6.50	6.67	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [ng/µl]	Region Molarity [nmol/l]	% of Total	Region Comment	Color
50	1,000	502	5.08	16.1	83.35		

Appendix 5.2. A summary of phylum level taxonomy for six metagenomics soil samples generated by EBI pipeline

Kingdom	Phylum	Metagenomics sample ID					
		MA1	**MA2**	**MA3**	**MA4**	**MA5**	**MA6**
Unassigned	Unassigned Bacteria	997	1274	900	916	846	1385
Archaea	Crenarchaeota	7	14	35	45	21	23
Bacteria	Actinobacteria	1060	1528	917	758	777	1185
Bacteria	Proteobacteria	665	870	409	395	127	692
Bacteria	Bacteroidetes	105	126	52	43	13	133
Bacteria	Gemmatimonadetes	57	87	50	42	50	23
Bacteria	Acidobacteria	48	64	66	47	25	22
Bacteria	Planctomycetes	41	64	29	21	16	10
Bacteria	Chloroflexi	30	111	76	61	79	71
Bacteria	Firmicutes	25	19	83	97	14	139
Bacteria	Armatimonadetes	13	13	14	8	6	7
Bacteria	Verrucomicrobia	7	13	6	8	0	8
Bacteria	Cyanobacteria	6	4	1	1	1	0
Bacteria	Chlorobi	5	2	0	0	0	0
Bacteria	TM7	3	5	3	4	0	1
Bacteria	FBP	2	4	4	5	0	1
Bacteria	Nitrospirae	2	4	1	2	2	2
Bacteria	Thermi	2	8	0	0	2	4
Bacteria	WPS-2	2	0	0	0	0	0
Bacteria	Chlamydiae	1	0	0	0	0	0
Bacteria	NKB19	1	0	0	0	0	0
Bacteria	Fibrobacteres	0	0	1	0	2	0
Bacteria	OD1	0	0	2	1	1	0
Bacteria	OP11	0	3	1	0	0	1
Bacteria	Synergistetes	0	0	1	0	0	0
Bacteria	Tenericutes	0	0	2	1	0	0

Chapter 6 Discussion

This project was driven by major threats endangering the biodiversity of Kuwait. Immediate action is required to protect the environment and conserve endangered plant species. Land degradation and vegetation loss, leading to soil erosion and an increase in the intensity and frequency of dust-sand storms, are major issues across the Kuwaiti desert (Al-Dousari and Al-Awadhi, 2012). The main causes are the combined influences of climatological and geological processes compounded by an intense action of human activities in sandy areas (Al-Awadhi et al. 2003). Overgrazing, seasonal spring camping, off-road driving and the use of heavy machinery together with climatic change are leading to vegetation loss across the desert (Misak et al., 2002).

Kuwait received compensation from the United Nation Compensation Commission (UNCC) for the remediation and restoration of ecosystems damaged by war. Local researchers and institutes together with international consultants are planning to restore and remediate areas affected by war-related damages including areas contaminated by oil spills; desert surfaces damaged by military vehicles; remediation of groundwater; the opening of detonation sites and the revegetation of damaged desert areas (UNCC decision 258, December 2005).

My aim was to research the existing plans for the restoration of vegetation in Kuwait with a view to investigating the potential for applying molecular methods. Decision makers and researchers are at present deciding on the choice of plants to be used for introduction into damaged sites in need of restoration. However, successful restoration requires a number of elements to be in place including careful planning based on informed decision-making grounded in clear knowledge of the environmental history and current conditions in the target area as well as funding to be in place to enable maintenance and monitoring (Tolba , 2007).

Based on my field experience there are some preliminary stages that have to be carried out before deciding on the choice of plants. Firstly, fences must be erected around the areas (Figure 1.8) to protect them from livestock and human activities.

Secondly, the possibility of natural regeneration of vegetation should be considered before resorting using active restoration programmes. For instance, it was reported by Brown and Al-Mazrooei (2003) that active regeneration of plants occurred in northern Kuwait when grazing was removed for four years. This is particularly important as natural vegetation is an important source of sustainable benefits to society and provides an opportunity for livestock grazing and a source of valuable seed and genetic diversity (Omar and Bhat, 2007).

While natural recovery takes place, NGS technologies and DNA barcoding methods can be applied and contribute effectively towards the identification and monitoring of vegetation by sampling fragments of emerging seedlings up to study past biodiversity from eDNA samples.

DNA barcoding can identify unknown plant material collected from the field (i.e. fragments of roots, leaves, seeds, seedlings) by processing the samples following simple molecular techniques and identifying unknown sequences against the local DNA database. This method will help researchers to speed up the process of identification and possibly reduce the reliance on taxonomists in the future. As DNA services are advancing, the processing of samples using full plates in the lab are capable of generating sequences for 96 individual in 2-3 days.

The DNA database provided in this project is capable of identifying species up to 70.5 % resolution (*rbcL* + ITS2 barcodes). Considering paraphyletic relationships amongst closely related species within a genus (discussed earlier in Chapter 4), a list of species choice will be narrowed down and since the average species/genera of the flora is 2-3 species, it could be resolvable to species level by identifying the samples through standard taxonomic methods using conventional identification keys. Recent studies have demonstrated that DNA barcoding complements traditional taxonomic revision or have helped in identifying cryptic species of plants (Zhang et al., 2015; Liu et al., 2011).

DNA barcoding method could contribute towards monitoring vegetation in local restoration projects which requires a sampling strategy based on rapid assessment field surveys. For restoration programmes it is important to compare plant diversity across different sites, for instance: (1) sites actively restored by seedlings and/ or

seeds, (2) naturally recovered sites, (3) and open desert areas (outside the fenced area).

NGS technologies are advancing and getting cheaper in recent years (Black et al., 2015), applying them in restoration projects would increase the quality of knowledge of the biodiversity across different habitats. The NGS methods demonstrated in this study will help researchers understand the historical patterns of past vegetation and provide guidance on the choice of plants and plant communities best suited for restoration. It is important to note that not all plants identified by NGS from eDNA samples are suitable for restoration – for instance, the habitat may no longer be appropriate for the plant to flourish or a critical pollinator may now be absent. However, it will act as a reference point for making decisions and narrowing the choices of plants that might best be used.

In active restoration projects, irrigation systems are widely used and often represent the most substantial component of project expenditure (Weigand and Rodger, 2009). However, such systems present problems. Irrigation systems using plastic pipelines across the desert leaves behind a long-term contamination if not managed well, as observed at Al-liyah actively restored site (KISR, 2008). I would recommend that water tank sprinkler trucks might be a better solution if active irrigation is required.

The desert areas of Kuwait are mainly open rangeland with the vegetation of importance for livestock production. In proposed protected area controlled grazing should be considered as plant-animal interactions they are important in a number of ways: they aid in increasing species richness and improving forage quality (Denyer et al., 2010); they provide urea (organic compound) for the nitrogen cycle (Rufino et al., 2006); and livestock transform vegetation into decomposed organic matter which adds nutrients to the soil. To maintain the vegetation cover and species richness inside protected areas, it is highly recommended to include controlled, sustainable grazing. It is important that the number of livestock be properly balanced with the available forage resources and in years of drought, livestock numbers should be reduced to prevent vegetation degradation (Brown, 2003).

The aims of active restoration or natural revegetation methods should be compared and questioned. Fundamental questions need to be asked and answered. For instance,

why restore the desert's ecosystem? Is it to maximise ecosystem services and benefit the society and the biodiversity or just to reintroduce the natural vegetation that once dominated in "pre-disturbance" times that might be lost again due to human activities, drought or other environmental factors?

UNEP defines ecosystem services as follows: 'Ecosystem services are the benefits people obtain from ecosystems. These include provisioning services such as food and water; regulating services such as flood and disease control; cultural services such as spiritual, recreational, and cultural benefits; and supporting services, such as nutrient cycling, that maintain the conditions for life on Earth' (UNEP, 2011).

In my opinion, maximising ecosystem services is a primary consideration when restoring and protecting degraded ecosystems. This is particularly important in Kuwait in the current economic situation where oil prices are dropping, but alternative sources of income for the country are very limited. Therefore, maximising ecosystem services would provide many benefits such as clean water and air, the reduction of further soil erosion, renewable energy (by solar and wind), protection from extreme weather events (including dust storms), the maintenance and protection of cultural values and education for future generations to name but a few of the benefits which are of value both for the people and the biodiversity (Eastwood et al., 2016). In Kuwait maximising ecosystem services would be important in the future by providing alternative sources of income and for tourism.

At present, in Kuwait, there are examples of several protected areas where the natural biodiversity has been protected from livestock grazing by fences for more than 20 years. People are only allowed to enter these protected areas with a permit. Developing such areas by following ecosystem services criteria (Vollmer et al., 2016; Keller and Fournier, 2015) would contribute towards the protection of biodiversity as well as enhancing cultural and educational values.

An excellent example of the importance of ecosystem services being highlighted in the region is provided by the United Arab Emirates (UAE) under the slogan 'A Green Economy for Sustainable Development', which has earned the UAE the status of being one of the regional leaders in terms of the conservation of its environment and wildlife (UAE-MEW, 2014). The protection of its terrestrial and marine environment

is guaranteed by federal law which aims to conserve fragile ecosystems with particular attention being paid to endangered species. The country is also committed to the UN Convention to Combat Desertification (UNCCD) and aims to continue protecting its natural habitats and underground water. This perhaps provides a useful model for other countries in the region and a way forward for Kuwait.

References

Abo El-Nil, M. (1997). Role of Plant tissue culture propagation in the rehabilitation of desert environment. *Restoration and rehabilitation of the desert environment*. Amesterdam: Elsevier.

Abo El-Nil, M. (2001). Growth and estabishment of mangrove (*Avicennia marina*) on the coastlines of Kuwait. *Wetlands Ecology and Management, 9*, 421-428.

Adams, K. L., & Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology, 8*(2), 135-141.

Ahmed, M., Al-Dousari, N., & Al-Dousari, A. (2016). The role of dominant perennial native plant species in controlling the mobile sand encroachment and fallen dust problem in Kuwait. *Arabian Journal of Geosciences, 9*(2). doi:10.1007/s12517-015-2216-6

Al-Awadhi, J. M., Misak, R. F., & Omar, S. A. S. (2003). Causes and consequences of desertification in Kuwait: a case study of land degradation. *Bull Eng Geol Env*(62), 107-115. doi:10.1007/s10064-002-0175-0

Al-Awadhi, J. M., Omar, S. A., & Misak, R. F. (2005). Land degradation indicators in Kuwait. *Land Degradation & Development, 16*(2), 163-176. doi:10.1002/ldr.666

Al-Damkhi, A. M. (2007). Kuwait‟s oil well fires, 1991: environmental crime and war. *International Journal of Environmental Studies, 64*(1), 31-44.

Al-hemaid, F., Ali, A., Lee, J., Kim, S., & Rahman, M. (2015). Molecular evolutionary relationships of *Euphorbia cordifolia* jacq. within the genus inferred from analysis of internal transcribed spacer sequences. *Bangladesh J. Plant Taxon, 22*(2), 111-118.

Al-Humaid, A. L., & Warrang, M. O. (1998). Allelopathic effects of mesquite (Prosopis juliflora) foliage on seed germi- nation and seedling growth of Bermudagrass. *J. Arid Environ., 38*, 237-243.

Al-Nafie. (2008). Phytogeography of Saudi Arabia. *Saudi J Biol Sci, 1*(15), 159-176.

Al-Qurainy, F., Khan, S., Nadeem, M., Tarroum, M., & Gaafar, A. R. Z. (2014). Selection of DNA barcoding loci for Nepeta deflersiana Schweinf. ex Hedge from chloroplast and nuclear DNA genomes. *Genetics and Molecular Research, 13*(1), 1144-1151.

Al-Qurainy, F., Nadeem, M., Khan, S., Alansi, S., & Tarroum, M. (2013). Effecient regeneration of a potential medicinal plant Ochradenus baccatus Delile from cotyledon and shoot axis. *Pakistan Journal of Botany, 45*(2), 501-505.

Al-Rawi, A. (1987). *Flora of Kuwait vol. 2: Compositae and Monocotyledoneae*. Kuwait University.

Al-Salameen, F., Al-Amad, S., & Al-Hashash, H. (2014). Determination of genetic variation of Rhanterium epapposum in Kuwait desert using RAPD and SRAP DNA-based markers. *Kuwait J. Sci., 3*, 163-175.

Al-Saqer, N. A. (2003). Birds of the Jahra conserved area. *Journal of Arid Environments, 54*(1), 203-207. doi:10.1006/jare.2001.0878

Al-Sulaimi, J., & Pitty, A. (1995). Origin and depostional model of Wadi Al-Batin and its associated alluvial fan, Saudia Arabia and Kuwait. *Sedimentary Geology*(97), 203-229.

Al-Tamimi, H. A. (2010). Protected Areas in the State of Kuwait. *Environment Public Authority, Kuwait.* Retrieved from http://www.docs/24167108/Protected-Areas-in-the-State-of-Kuwait

Al-Yemeni, M. N., Al-Farraj, M. M., & Al-Farhan, A. (2000). Ecological studies on Rawdhat system in Saudi Arabia - II. Soil seed bank of Rawdhat Khorim. *Pakistan Journal of Botany, 32*(2), 273-282.

Al-Damkhi, A. M. (2007). Kuwait's oil well fires, 1991: environmental crime and war. *International Journal of Environmental Studies, 64*(1), 31-44. doi:10.1080/00207230601125036

Albrecht, M., Duelli, P., Obrist, M. K., Kleijn, D., & Schmid, B. (2009). Effective Long-Distance Pollen Dispersal in *Centaurea jacea*. *PLoS One, 4*(8), 1-8.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol., 215*(5), 403-410.

Anonymous. (1910). Longevity of fern spores. *The Garden, 74*, 63.

APGIV. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society, 181*, 1-20.

Badaeva, E. D., Amosova, A. V., Samatadze, T. E., Zoshchuk, S. A., Shostak, N. G., Chikida, N. N., . . . Gill, B. S. (2004). Genome differentiation in Aegilops . 4. Evolution of the U-genome cluster. *Plant Systematics and Evolution, 246*(1-2), 45-76. doi:10.1007/s00606-003-0072-4

Bafeel, S. O., Arif, I. A., Bakir, M. A., Al Homaidan, A. A., Al Farhan, A. H., & Khan, H. A. (2012). DNA barcoding of arid wild plants using rbcL gene sequences. *Genet Mol Res, 11*(3), 1934-1941. doi:10.4238/2012.July.19.12

Bakker, J. P., Poschlod, P., Strykstra, R. J., Bekker, R. M., & Thompson, K. (1996). Seed banks and seed dispersal: important topics in restoration ecology. *Acta Bot. Neerl., 45*(4), 461-490.

Baptista, P., Reis, F., Pereira, E., Tavares, R. M., Santos, P. M., Richard, F., . . . Lino-Neto, T. (2015). Soil DNA pyrosequencing and fruitbody surveys reveal contrasting diversity for various fungal ecological guilds in chestnut orchards. *Environ Microbiol Rep, 7*(6), 946-954.

Batanouny, K. H. (1983). Human Impact on Desert vegetation. *Junk, The Hague*, 139-149.

Bell, K. L., Loeffler, V. M., & Brosi, B. J. (2017). An *rbcL* reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Appl Plant Sci, 5*(3), 1-7. doi:10.3732/apps.1600110

Belnap, J., Rosentreter, R., Leonard, S., Kaltenecker, J., Williams, J., & Eldridge, D. (2001). Biological Soil Crusts: Ecology and Managment. *Denver, Colorado, Report*, 1-118.

Birks, H. H. (2001). Plant macrofossils. In J. P. Smol, H. J. B. Birks, & W. M. Last (Eds.), *Tracking Environmental Change Using Lake Sediments* (Vol. 3, pp. 49-74). Dordrecht: Kluwer

Black, J. S., Salto-Tellez, M., Mills, K. I., & Catherwood, M. A. (2015). The impact of next generation sequencing technologies on haematological research – A review. *Pathogenesis, 2*(1-2), 9-16.

Blattner, F. R. (2009). Progress in phylogenetic analysis and a new infrageneric classification of the barley genus Hordeum (Poaceae: Triticeae). *Breeding Science, 59*, 471-480.

Blay, E. S., Schwabedissen, S. G., Magnuson, T. S., Aho, K. A., Sheridan, P. P., & Lohse, K. A. (2017). Variation in Biological Soil Crust Bacterial Abundance and Diversity as a Function of Climate in Cold Steppe Ecosystems in the Intermountain West, USA. *Microb Ecol*. doi:10.1007/s00248-017-0981-3

Bolson, M., Smidt Ede, C., Brotto, M. L., & Silva-Pereira, V. (2015). ITS and trnH-psbA as Efficient DNA Barcodes to Identify Threatened Commercial Woody Angiosperms from Southern Brazilian Atlantic Rainforests. *PLoS One, 10*(12), e0143049. doi:10.1371/journal.pone.0143049

Boulos, L. (1988). *The Weed Flora of Kuwait*. Kuwait University.

Boulos, L., & Al-Dosari, M. (1994). Checklist of the flora of Kuwait. *J. Univ. Kuwait, 21*, 203-217.

Braukmann, T. W., Kuzmina, M. L., Sills, J., Zakharov, E. V., & Hebert, P. D. (2017). Testing the Efficacy of DNA Barcodes for Identifying the Vascular Plants of Canada. *PLoS One, 12*(1), e0169515. doi:10.1371/journal.pone.0169515

Brown, G. (2003). Factors maintaining plant diversity in degraded areas of northern Kuwait. *Journal of Arid Environments, 54*(1), 183-194. doi:10.1006/jare.2001.0880

Brown, G., & Al-Mazrooei, S. (2003). Rapid vegetation regeneration in a seriously degraded Rhanterium epapposum community in northern Kuwait after 4 years of protection. *Journal of Environmental Management, 68*(4), 387-395. doi:10.1016/s0301-4797(03)00107-5

Brown, G., & Schoknecht, N. (2001). Off-road vehicles and vegetation patterning in a degraded desert ecosystem in Kuwait. *Journal of Arid Environments, 49*(2), 413-427. doi:10.1006/jare.2000.0772

Brummitt, R. K., & Powell, C. E. (1992). *Authors of Plant Names*. RBG Kew: Whitstable Litho

Bruni, I., De Mattia, F., Martellos, S., Galimberti, A., Savadori, P., Casiraghi, M., . . . Labra, M. (2012). DNA barcoding as an effective tool in improving a digital plant identification system: a case study for the area of Mt. Valerio, Trieste (NE Italy). *PLoS One, 7*(9), e43256. doi:10.1371/journal.pone.0043256

Burgess, K. S., Fazekas, A. J., Kesanakurti, P. R., Graham, S. W., Husband, B. C., Newmaster, S. G., . . . Barrett, S. C. H. (2011). Discriminating plant species in a local temperate flora using the rbcL+matK DNA barcode. *Methods in Ecology and Evolution, 2*(4), 333-340. doi:10.1111/j.2041-210X.2011.00092.x

CBOL. (2009). A DNA barcode for land plants. *Proc Natl Acad Sci U S A, 106*(31), 12794-12797. doi:10.1073/pnas.0905845106

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., . . . Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One, 5*(1), e8613. doi:10.1371/journal.pone.0008613

Coissac, E., Riaz, T., & Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Mol Ecol, 21*(8), 1834-1847. doi:10.1111/j.1365-294X.2012.05550.x

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., & Tiedje, J. M. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research, 37*, D141-D145. doi:http://doi.org/10.1093/nar/gkn879

Cope, T. A. (2007). *Flora of the Arabian Peninsula and Socotra* (Vol. 5): Edinburgh University Press.

Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Trends Ecol Evol, 29*(10), 566-571.

Crum, H. (2001). *Structural Diversity of Bryophytes*. USA: University of Michigan Herbarium.

Daoud, H. S. (1985). *Flora of Kuwait vol. 1: Dicotyledoneae*. London: KPI and Kuwait University.

de Vere, N., Rich, T. C., Ford, C. R., Trinder, S. A., Long, C., Moore, C. W., . . . Wilkinson, M. J. (2012). DNA barcoding the native flowering plants and conifers of Wales. *PLoS One, 7*(6), e37945. doi:10.1371/journal.pone.0037945

Denyer, K. L., Hartley, S. E., & John, E. A. (2010). Both bottom-up and top-down processes contribute to plant diversity maintenance in an edaphically heterogeneous ecosystem. *Journal of Ecology, 98*(2), 498-508.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., & Andersen, G. L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology, 72*(7), 5069-5072. doi:http://doi.org/10.1128/AEM.03006-05

Dickson, V. (1955). *The Wild Flowers of Kuwait and Bahrain*. London: George Allen & Unwin Ltd.

Dizkirici, A., Ekici, M., & Kaya, Z. (2013). Comparative molecular phylogenetics of Astragalus L. sections from Turkey with New World Astragalus species using nrDNA ITS sequences. *Plant Systematics and Evolution, 300*(1), 163-175. doi:10.1007/s00606-013-0868-9

Dunning, l. T., & Savolainen, V. (2010). Broad-scale amplification of matK for DNA barcoding plants, a technical note. *Botanical Journal of the Linnean Society, 164*, 1-9.

Eastwood, A., Brooker, R., Irvine, R. J., Artz, R. E., Norton, L. R., Bullock, J. M., . . . Pakeman, R. J. (2016). Does nature conservation enhance ecosystem services delivery? *Ecosystem Services, 17*, 152-162.

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., . . . Relman, D. A. (2005). Diversity of the Human Intestinal Microbial Flora. *Science, 308*, 1635-1638.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*(5), 1792-1797.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics, 26*(19), 2460-2461.

El-Keblawy, A., & Al-Rawai, A. (2005). Effects of salinity, tem- perature and light on germination of invasive *Prosopis juliflora* (SW.). *J. Arid Env., 61*, 555-556.

El-Sheikh, M. A., Abbadi, G. A., & Bianco, P. M. (2010). Vegetation ecology of phytogenic hillocks (nabkhas) in coastal habitats of Jal Az-Zor National Park, Kuwait: Role of patches and edaphic factors. *Flora - Morphology, Distribution, Functional Ecology of Plants, 205*(12), 832-840. doi:10.1016/j.flora.2010.01.002

ElBehi, A. W., Orlandi, F., Bonofiglio, T., Romano, B., Inglese, P., Sortino, G., & Liguori, G. (2015). Pollen Morphology and Reproductive Performances in *Opuntia ficusindica* (L.) Mill. *Acta horticulturae, 1067*(1), 217-223.

Eldridge, D. (2000). Ecology and Management of Biological Soil Crusts. *The Bryologist, 103*(4), 742-747.

Enan, M. R., & Ahamed, A. (2014). DNA barcoding based on plastid matK and RNA polymerase for assessing the genetic identity of date (Phoenix dactylifera L.) cultivars. *Genet. Mol. Res., 13*(2), 3527-3536.

Epp, L. S., Boessenkool, S., Bellemain, E. P., Haile, J., Esposito, A., Riaz, T., . . . Brochmann, C. (2012). New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Mol Ecol, 21*(8), 1821-1833. doi:10.1111/j.1365-294X.2012.05537.x

Erenler, H. E., Ashton, P. A., Gillman, M. P., & Ollerton, J. (2010). Factors determining species richness of soil seed banks in lowland ancient woodlands. *Biodiversity and Conservation, 19*(6), 1631-1648. doi:10.1007/s10531-010-9793-1

Ergun, H. (1969). Reconnaissance Soil Survey: Kuwait. *FAOUN, ROME*, 104.

Faegri, K., & van der Pijl, L. (1979). *The Principles of Pollination Ecology* (3 ed.). New York: Pergamon Press.

Fahner, N. A., Shokralla, S., Baird, D. J., & Hajibabaei, M. (2016). Large-Scale Monitoring of Plants through Environmental DNA Metabarcoding of Soil: Recovery, Resolution, and Annotation of Four DNA Markers. *PLoS One, 11*(6), e0157505. doi:10.1371/journal.pone.0157505

Farell, E., & Alexandre, G. (2012). Bovine serum albumin further enhances the effects of organic solvents on increased yield of polymerase chain reaction of GC-rich templates. *BMC Research Notes, 5*, 257.

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., . . . Barrett, S. C. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS One, 3*(7), e2802. doi:10.1371/journal.pone.0002802

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol., 17*, 368-376.

Fenner, M., & Thompson, K. (2005). *The Ecology of Seeds* (pp. 31). doi:10.1017/cbo9780511614101

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., . . . Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Mol Ecol Resour, 15*(3), 543-556. doi:10.1111/1755-0998.12338

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool., 20*, 406-416.

Ford, C. S., Ayres, K. L., Toomey, N., Haider, N., Stahl, J. V. A., Kelly, l. J., . . . Wilkinson, M. J. (2009). Selection of candidate coding DNA barcoding regions for use on land plants. *Bot J. Linn. Soc., 159*(1), 1-11.

Ghazanfar, S., & Al-Sabahi, A. (1993). Medicinal Plants of Northern and Centeral Oman (Arabia). *Economic Botany, 47*(1), 89-98.

Ghazanfar, S. A., & Fisher, M. (1998). *Vegetation of the Arabian Peninsula*. Netherlands: Kluwer Academics Publishers.

Goldstein, P. Z., & Rob, D. (2003). Calibrating phylogenetic species formation in a threatened insect using DNA from historical specimens. *Mol Ecol, 12*, 1993-1998. doi:10.1046/j.1365-294X.2003.01860.x

Griffin, G. F., Stafford, D. M., Smith, G. R., & Morton, G. E. (1989). Status and implications of the invasion of Tamarisk (Tamarix aphylla) on the Finke River, northern territory, Australia. *J. Env. Mang., 29*, 297-315.

Gross, K. L. (1990). A Comparison of Methods for Estimating Seed Numbers in the Soil. *Journal of Ecology, 78*(4), 1079-1093.

Gutterman, Y. (2012). *Seed Germination in Desert Plants*: Springer Science & Business Media.

Haile, J., Froese, D. G., MacPhee, R. D. E., Roberts, R. G., Arnold, L. J., Reyes, A. V., ... (2009). Ancient DNA reveals late survival of mammoth and horse in interior Alaska *PNAS, 106*(52).

Hajibabaei, M., Singer, G. A., Hebert, P. D., & Hickey, D. A. (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet, 23*(4), 167-172. doi:10.1016/j.tig.2007.02.001

Halwagy, R., & Halwagy, M. (1974). Ecological studies on the desert of Kuwait II The vegetation. *J. Univ. Kuwait, 1*, 86-94.

Halwagy, R., & Macksad, A. (1972). A contribution towards a Flora of the State of Kuwait and the Neutral Zone. *Bot. J. Linn. Soc., 65*, 61-79.

Han, J., Zhu, Y., Chen, X., Liao, B., Yao, H., Song, J., . . . Meng, F. (2013). The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. *Biomed Res Int, 2013*, 1-7. doi:10.1155/2013/741476

Harper, K. T., & Pendleton, R. L. (1993). Cyanobacteria and cyanolichens: can they enhance availability of essential minerals for higher plants? *Great Basin Nat., 55*, 59-72.

Hebert, P. D., & Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Syst Biol, 54*(5), 852-859. doi:10.1080/10635150500354886

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. B., 270*, 313-321.

Heeg, J. S., & Wolf, M. (2015). ITS2 and 18S rDNA sequence-structure phylogeny of Chlorella and allies (Chlorophyta, Trebouxiophyceae, Chlorellaceae). *Plant Gene, 4*, 20-28. doi:10.1016/j.plgene.2015.08.001

Hernández, H. M., & Navarro, M. (2007). A new method to estimate areas of occupancy using herbarium data. *Biodiversity and Conservation, 16*(9), 2457-2470. doi:10.1007/s10531-006-9134-6

Hiiesalu, I., Opik, M., Metsis, M., Lilje, L., Davison, J., Vasar, M., . . . Partel, M. (2012). Plant species richness belowground: higher richness and new patterns revealed by next-generation sequencing. *Mol Ecol, 21*(8), 2004-2016. doi:10.1111/j.1365-294X.2011.05390.x

Hofreiter, M., Mead, J. I., Martin, P., & Poinar, H. N. (2003). Molecular Caving. *Current Biology, 13*(18), 693-695.

Hollingsworth, M. L., Andra Clark, A., Forrest, L. L., Richardson, J., Pennington, R. T., Long, D. G., . . . Hollingsworth, P. M. (2009). Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Mol Ecol Resour, 9*(2), 439-457. doi:10.1111/j.1755-0998.2008.02439.x

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS One, 6*(5), e19254. doi:10.1371/journal.pone.0019254

Hribova, E., Cizkova, J., Christelova, P., Taudien, S., de Langhe, E., & Dolezel, J. (2011). The ITS1-5.8S-ITS2 sequence region in the Musaceae: structure, diversity and use in molecular phylogeny. *PLoS One, 6*(3), e17863. doi:10.1371/journal.pone.0017863

Hunter, S., Corbett, M., Denise, H., Fraser, M., Gonzalez-Beltran, A., Hunter, C., . . . Sansone, S. A. (2014). EBI metagenomics--a new resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res, 42*(Database issue), D600-606. doi:10.1093/nar/gkt961

IUCN. (2008). Guidelines for Applying Protected Area Management Categories. *Gland, Switzerland*, 86.

IUCN. (2012). The Conservation Status and Distribution of Reptiles of the Arabian Peninsula. *Cambridge, UK and Gland, Switzerland: IUCN, and Sharjah, UAE: Environment and Protected Areas Authority.*, 38.

IUCN. (2013). Invasive alien species: the urban dimension, Case studies on strengthening local action in Europe. Brussels. *Brussels, Belgium*.

Jahren, A. H., Porter, S., & J., K. J. (2003). Lichen metabolism identified in Early Devonian terrestrial organisms. *Geological Society of America, 31*(2), 99-102.

Javanmardi, F., Kazempour Osaloo, S., Maassoumi, A. A., & Nejadsattrai, T. (2012). Molecular phylogeny of Astragalus section Alopecuroidei (Fabaceae) and its allies based on nrDNA ITS and three cpDNAs, matK, trnT-trnY and trnH-psbA sequences. *Biochemical Systematics and Ecology, 45*, 171-178. doi:10.1016/j.bse.2012.07.029

Jensen, J. S., & Hansen, J. K. (2010). Genetic variation in responses to different soil water treatments inQuercus roburL. *Scandinavian Journal of Forest Research, 25*(5), 400-411. doi:10.1080/02827581.2010.512873

Jones, F. A., Erickson, D. L., Bernal, M. A., Bermingham, E., Kress, W. J., Herre, E. A., . . . Turner, B. L. (2011). The roots of diversity: below ground species richness and rooting distributions in a tropical forest revealed by DNA barcodes and inverse modeling. *PLoS One, 6*(9), e24506. doi:10.1371/journal.pone.0024506

Jurado-Rivera, J. A., Vogler, A. P., Reid, C. A., Petitpierre, E., & Gomez-Zurita, J. (2009). DNA barcoding insect-host plant associations. *Proc Biol Sci, 276*(1657), 639-648. doi:10.1098/rspb.2008.1264

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform

for the organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-1649.
Keller, A. A., Fournier, E., & Fox, J. (2015). Minimizing impacts of land use change on ecosystem services using multi-criteria heuristic analysis. *Journal of Environmental Management, 156*, 23-30.
Kesanakurti, P. R., Fazekas, A. J., Burgess, K. S., Percy, D. M., Newmaster, S. G., Graham, S. W., . . . Husband, B. C. (2011). Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. *Mol Ecol, 20*(6), 1289-1302. doi:10.1111/j.1365-294X.2010.04989.x
Keshavarzi, M., Rahiminejad, M. R., & Kheradmandnia, M. (2007). Discriminative analysis of 10 species of Aegilops l. (Poaceae) native to Iran. *Iranian Journal of Science & Technology, 31*(A3), 321-324.
Khansari, E., Zarre, S., Alizadeh, K., Attar, F., Aghabeigi, F., & Salmaki, Y. (2012). Pollen morphology of Campanula (Campanulaceae) and allied genera in Iran with special focus on its systematic implication. *Flora - Morphology, Distribution, Functional Ecology of Plants, 207*(3), 203-211. doi:10.1016/j.flora.2012.01.006
Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol., 16*(111-120).
KISR. (1999). *Soil Survey for the State of Kuwait* (Vol. 2). KISR: AACM International, Adelaide.
KISR. (2008). Environmental Rehabilitation Study of Liyah Gravel Quarry North of Kuwait. *Kuwait Institute for Scietific Research*, 28.
KMC. (2016). Kuwait Meterological Center. Retrieved from http://www.met.gov.kw/
Kok, M. T. J., Kok, K., Peterson, G. D., Hill, R., Agard, J., & Carpenter, S. R. (2016). Biodiversity and ecosystem services require IPBES to take novel approach to scenarios. *Sustainability Science*, 1-5. doi:10.1007/s11625-016-0354-8
Kraaijeveld, K., de Weger, L. A., Ventayol Garcia, M., Buermans, H., Frank, J., Hiemstra, P. S., & den Dunnen, J. T. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol Ecol Resour, 15*(1), 8-16. doi:10.1111/1755-0998.12288
Kress, W. J., & Erickson, D. L. (2007). A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS One, 2*(6), e508. doi:10.1371/journal.pone.0000508
Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjur, O., & Bermingham, E. (2009). Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proc Natl Acad Sci U S A, 106*(44), 18621-18626. doi:10.1073/pnas.0909820106
Kress, W. J., Garcia-Robledo, C., Uriarte, M., & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends Ecol Evol, 30*(1), 25-35. doi:10.1016/j.tree.2014.10.008
Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A., & Janzen, D. H. (2005). Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci U S A, 102*(23), 8369-8374. doi:10.1073/pnas.0503123102
Kuete, V., Wiench, B., Alsaid, M. S., Alyahya, M. A., Fankam, A. G., Shahat, A. A., & Efferth, T. (2013). Cytotoxicity, mode of action and antibacterial activities of selected Saudi Arabian medicinal plants. *BMC Complementary and Alternative Medicine, 13*, 1-11.
Kumar, S., Stecher, G., & Tamura, K. (2015). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Ecol.*, 1-5.
Kuparinen, A., Markkanen, T., Riikonen, H., & Vesala, T. (2007). Modeling air-mediated dispersal of spores, pollen and seeds in forested areas. *Ecological Modelling, 208*(2-4), 177-188. doi:10.1016/j.ecolmodel.2007.05.023

Kurschner, H. (2003). Nineteen new records to the bryophyte flora of Socotra Island Additions to the Bryophyte Flora of the Arabian Peninsula and Socotra *Willdenowia, 33*(2), 445-458.

Kürschner, H., & Ochyra, R. (2014). Novelties in the moss flora of Oman, Saudi Arabia and Yemen, Including the most OutstandingVesicularia montagnei(Bél.) Broth. (Hypnaceae). *Cryptogamie, Bryologie, 35*(1), 93-97. doi:10.7872/cryb.v35.iss1.2014.93

Kuzmina, M. L., Johnson, K. L., Barron, H. R., & Hebert, P. D. N. (2012). Identification of the vascular plants of the Churchill, Manitoba, using DNA barcode library. *BMC Ecology, 12*, 25.

Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., . . . Savolainen, V. (2008). DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci U S A, 105*(8), 2923-2928. doi:10.1073/pnas.0709936105

Lam, K. N., Cheng, J., Engel, K., Neufeld, J. D., & Charles, T. C. (2015). Current and future resources for functional metagenomics. *Frontiers in Microbiology, 6*, 1-8.

Lamb, E. G., Winsley, T., Piper, C. L., Freidrich, S. A., & Siciliano, S. D. (2016). A high-throughput belowground plant diversity assay using next-generation sequencing of the trnL intron. *Plant and Soil*. doi:10.1007/s11104-016-2852-y

Le Houerou, H. N. (1984). Forage and fuel plants in the arid zone of North Africa the Near and Middle East. In G. E. G. Wickens, J.R. and Field, D.V. (Ed.), *Plants of Arid Lands* (pp. 117-141). London: George Allen and Unwin.

Leicht-Young, S. A., Pavlovic, N. B., Grundel, R., & Frohnapple, K. J. (2008). A comparison of seed banks across a sand dune successional gradient at Lake Michigan dunes (Indiana, USA). *Plant Ecology, 202*(2), 299-308. doi:10.1007/s11258-008-9484-0

Lewin, G. R., Carlos, C., Chevrette, M. G., Horn, H. A., McDonald, B. R., Stankey, R. J., . . . Currie, C. R. (2016). Evolution and Ecology of Actinobacteria and Their Bioenergy Applications. *Annu Rev Microbiol, 70*, 235-254. doi:10.1146/annurev-micro-102215-095748

Li, D.-Z., Liu, J.-Q., Chen, Z.-D., Wang, H., Ge, X.-J., Zhou, S.-L., . . . Chen, S.-L. (2011). Plant DNA barcoding in China. *Journal of Systematics and Evolution, 49*(3), 165-168. doi:10.1111/j.1759-6831.2011.00137.x

Li, H. Q., Chen, J. Y., Wang, S., & Xiong, S. Z. (2012). Evaluation of six candidate DNA barcoding loci in Ficus (Moraceae) of China. *Mol Ecol Resour, 12*(5), 783-790. doi:10.1111/j.1755-0998.2012.03147.x

Li, X. R., Wang, X. P., & Zhang, J. G. (2002). Microbiotic soil crust and its effects on vegetation and habit on artificially atabilized desert dunes in Tengger Desert North China. *Biol. Fert. Soil, 35*, 147-154.

Li, Y., Tong, Y., & Xing, F. (2016). DNA Barcoding Evaluation and Its Taxonomic Implications in the Recently Evolved Genus Oberonia Lindl. (Orchidaceae) in China. *Front Plant Sci, 7*, 1791. doi:10.3389/fpls.2016.01791

Little, D. P. (2011). DNA barcode sequence identification incorporating taxonomic hierarchy and within taxon variability. *PLoS One, 6*(8), e20552. doi:10.1371/journal.pone.0020552

Little, D. P. (2014). A DNA mini-barcode for land plants. *Mol Ecol Resour, 14*(3), 437-446. doi:10.1111/1755-0998.12194

Liu, J., Moller, M., Gao, L. M., Zhang, D. Q., & Li, D. Z. (2011). DNA barcoding for the discrimination of Eurasian yews (Taxus L., Taxaceae) and the discovery of cryptic species. *Mol Ecol Resour, 11*(1), 89-100. doi:10.1111/j.1755-0998.2010.02907.x

Liu, J., Provan, J., Gao, L. M., & Li, D. Z. (2012). Sampling strategy and potential utility of indels for DNA barcoding of closely related plant species: a case study in taxus. *Int J Mol Sci, 13*(7), 8740-8751. doi:10.3390/ijms13078740

Liu, J., Yan, H.-F., Newmaster, S. G., Pei, N., Ragupathy, S., Ge, X.-J., & Lowe, A. (2015). The use of DNA barcoding as a tool for the conservation biogeography of subtropical forests in China. *Diversity and Distributions, 21*(2), 188-199. doi:10.1111/ddi.12276

Liu, L., Liu, Y., Hui, R., & Xie, M. (2017). Recovery of microbial community structure of biological soil crusts in successional stages of Shapotou desert revegetation, northwest China. *Soil Biology and Biochemistry, 107*, 125-128. doi:10.1016/j.soilbio.2016.12.030

Lock, J. M., & Simpson, K. (1991). Legumes in West Asia: A Checklist. *RBG Kew*, 274.

Lonnell, N. (2011). *Plants & Ecology*. Stockholm: FMV Printcenter.

Lönnell, N., Jonsson, B. G., & Hylander, K. (2014). Production of diaspores at the landscape level regulates local colonization: an experiment with a spore-dispersed moss. *Ecography, 37*, 591-598.

Luo, K., Chen, S., Chen, K., Song, J., Yao, H., Ma, X., . . . Liu, Z. (2010). Assessment of candidate plant DNA barcodes using the Rutaceae family. *Sci China Life Sci, 53*(6), 701-708. doi:10.1007/s11427-010-4009-1

Luo, R., Liu, B., & Xie, Y. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience, 1*, 1-18.

Mabberley, D. J. (1997). *Mabberley's Plant-Book. A Portable Dictionary of the Vascular Plants*. Cape Town: Cambridge University Press.

Mahadani, P., Sharma, G. D., & Ghosh, S. K. (2013). Identification of ethnomedicinal plants (Rauvolfioideae: Apocynaceae) through DNA barcoding from northeast India. *Phcog Mag, 9*(35), 255-263.

Mandaville, J. P. (2011). *Bedouin Ethnobotany*. Tucson: The University of Arizona Press.

Marone, L., Cueto, V. R., Milesi, F. A., & de Casenave, J. L. (2004). Soil seed bank composition over desert microhabitats: patterns and plausible mechanisms. *Canadian Journal of Botany, 82*(12), 1809-1816.

Marone, L., & Horno, M. E. (1997). Seed reserves in the central Monte desert, Argentina: implications for granivory. *J. Arid Env., 36*, 661-670.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet, 17*, 10-12.

Mathew, K. T., Malallah, G., & Al-Dosari, M. (2012). Eleven new weeds in Kuwait. *Kuwait J. Sci. Eng., 39*(1A), 169-192.

Mendes, R., Garbeva, P., & Raaijmakers, J. M. (2013). The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiol Rev, 37*, 634-663.

Meyer, C. P., & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol, 3*(12), e422. doi:10.1371/journal.pbio.0030422

Miao, M., Warren, A., Song, W., Wang, S., Shang, H., & Chen, Z. (2008). Analysis of the Internal Transcribed Spacer 2 (ITS2) Region of Scuticociliates and Related Taxa (Ciliophora, Oligohymenophorea) to Infer their Evolution and Phylogeny. *Protist, 159*(4), 519-533.

Millar, M. A., Coates, D. J., & Byrne, M. (2014). Extensive long-distance pollen dispersal and highly outcrossed mating in historically small and disjunct populations of Acacia woodmaniorum (Fabaceae), a rare banded iron formation endemic. *Ann Bot, 114*(5), 961-971. doi:10.1093/aob/mcu167

Miller, A. G., & Cope, T. A. (1996). Flora of the Arabian Peninsula and Socotra. *Edinburgh University Press, 1*, 586.

Misak, R. F., Al-Awadhi, J. M., Omar, S. A., & Shahid, S. A. (2002). Soil degradation in Kabd area, southwestern Kuwait City. *Land Degradation & Development, 13*(5), 403-415. doi:10.1002/ldr.522

Mishra, P., Kumar, A., Rodrigues, V., Shukla, A. K., & Sundaresan, V. (2016). Feasibility of nuclear ribosomal region ITS1 over ITS2 in barcoding taxonomically challenging genera of subtribe Cassiinae (Fabaceae). *PeerJ, 4*, e2638. doi:10.7717/peerj.2638

Naderi Safar, K., Kazempour Osaloo, S., Maassoumi, A. A., & Zarre, S. (2014). Molecular phylogeny of Astragalus section Anthylloidei (Fabaceae) inferred from nrDNA ITS and plastid rpl32-trnL(UAG) sequence data. *Turkish Journal of Botany, 38*, 637-652. doi:10.3906/bot-1308-44

Nawash, O. S., & Al-Horani, A. S. (2011). The most important medicinal plants in Wadi Araba desert South West Jordan:. *Advances in Environmental Biology, 5*(2), 418-425.

NCBI. (2016). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research, 44*(Database issue), D7-D19.

Nelson, J. F., & Chew, R. M. (1977). Factors Affecting Seed Reserves in the Soil of a Mojave Desert Ecosystem, Rock Valley, Nye County, Nevada. *The AMerican Midland Naturalist, 97*(2), 300-320.

Omar, S. (2007). *Vegetation of Kuwait* (2 ed.): KISR.

Omar, S. A., Misak, R., King, P., Shahid, S. A., Abo-Rizq, H., Grealish, G., & Roy, W. (2001). Mapping the vegetation of Kuwait through reconnaissance soil survey. *Journal of Arid Environments, 48*(3), 341-355. doi:10.1006/jare.2000.0740

Omar, S. A. S., & Bhat, N. R. (2008). Alteration of theRhanterium epapposumplant community in Kuwait and restoration measures. *International Journal of Environmental Studies, 65*(1), 139-155. doi:10.1080/00207230701823332

Orgiazzi, A., Dunbar, M. B., Panagos, P., de Groot, G. A., & Lemanceau, P. (2015). Soil biodiversity and DNA barcodes: opportunities and challenges. *Soil Biology and Biochemistry, 80*, 244-250. doi:10.1016/j.soilbio.2014.10.014

Ott, J. E., McArthur, E. D., & Sanderson, S. C. (2011). Vegetation Dynamics at a Mojave Desert Restoration Site. Natural Resources and Environmental Issues*16*(1), 105-120. Retrieved from

Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G. A., Papanikolaou, N., Kotoulas, G., . . . Iliopoulos, I. (2015). Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinform Biol Insights, 9*, 75-88. doi:10.4137/BBI.S12462

PACI. (2015). The Public Authority For Civil Information Kuwait. https://www.paci.gov.kw/

Pang, X., Shi, L., Song, J., Chen, X., & Chen, S. (2013). Use of the potential DNA barcode ITS2 to identify herbal materials. *J Nat Med, 67*(3), 571-575. doi:10.1007/s11418-012-0715-2

Parducci, L., Matetovici, I., Fontana, S. L., Bennett, K. D., Suyama, Y., Haile, J., . . . Willerslev, E. (2013). Molecular- and pollen-based vegetation analysis in lake sediments from central Scandinavia. *Mol Ecol, 22*(13), 3511-3524. doi:10.1111/mec.12298

Pasiecznik, N. M., Felker, P., Harris, P. J. C., Harsh, L. N., Cruz, G., Tewari, J. C., . . . Maldonado, L. J. (2001). The Prosopis juliflora - Prosopis pallida Complex: A Monograph. *HDRA Coventry UK*, 172.

Pedersen, M. W., Ginolhac, A., Orlando, L., Olsen, J., Andersen, K., Holm, J., . . . Kjær, K. H. (2013). A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. *Quaternary Science Reviews, 75*, 161-168. doi:10.1016/j.quascirev.2013.06.006

Percy, D. M., Argus, G. W., Cronk, Q. C., Fazekas, A. J., Kesanakurti, P. R., Burgess, K. S., . . . Graham, S. W. (2014). Understanding the spectacular failure of DNA barcoding in

willows (Salix): does this result from a trans-specific selective sweep? *Mol Ecol, 23*(19), 4737-4756. doi:10.1111/mec.12837
Perveen, A., & Qaiser, M. (1999). Pollen Flora of Pakistan - XIII. Campanulaceae. *Tr. J. of Botany, 23*, 45-51.
Pettengill, J. B., & Neel, M. C. (2010). An evaluation of candidate plant DNA barcodes and assignment methods in diagnosing 29 species in the genus Agalinis (Orobanchaceae). *Am J Bot, 97*(8), 1391-1406. doi:10.3732/ajb.0900176
Pompanon, F., Deagle, B. E., Symondson, W. O., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: diet assessment using next generation sequencing. *Mol Ecol, 21*(8), 1931-1950. doi:10.1111/j.1365-294X.2011.05403.x
Pond, S. K., Wadhawan, S., Chiaromonte, F., Ananda, G., Chung, W. Y., Taylor, J., & Nekrutenko, A. (2009). Windshield splatter analysis with the Galaxy metagenomic pipeline. *Genome Research, 19*(11), 2144-2153. doi:10.1101/gr.094508.109
Prasse, R., & Bornkamm, R. (2000). Effect of microbiotic soil surface crusts on emergence of vscular plants. *Plant Ecology, 150*, 65-75.
Pysek, P., Richardson, D. M., Rejmánek, M., Webster, G. L., Williamson, M., & Kirschner, J. (2004). Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. *Taxon, 53*(1), 131-143.
Rahman, M. A., Mossa, J. S., Al-Said, M. S., & Al-Yahya, M. A. (2004). Medicinal plant diversity in the flora of Saudi Arabia 1: a report on seven plant families. *Fitoterapia, 75*(2), 149-161. doi:10.1016/j.fitote.2003.12.012
Ralser, M., Querfurth, R., Warnatz, H., Lehrach, H., Yaspo, M., & Krobitsch, S. (2006). An efficient and economic enhancer mix for PCR. *Biochemical and Biophysical Research Communications, 347*, 747-751.
Ratnasingham , S., & Hebert, P. D. N. (2013). A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS One, 8*(7), e66213.
Raveendar, S., Lee, J., Park, J., Lee, G., Jeon, Y., Lee, W., . . . Chung, J. (2015). Identification of Genus Vigna using ITS2 and matK as a Two-Locus DNA Barcode. *Plant Breeding and Biotechnology, 3*(2), 153-159. doi:10.9787/pbb.2015.3.2.153
Reichman, O. J. (1984). Spatial and temporal variation of seed distributions in Sonoran desert soils. *Journal of Biogeography, 11*, 1-11.
Reyes-Agüero, J. A., Aguirre R, J. R., & Valiente-Banuet, A. (2006). Reproductive biology of Opuntia: A review. *Journal of Arid Environments, 64*(4), 549-585. doi:10.1016/j.jaridenv.2005.06.018
Ronsted, N., Chase, M. W., Albach, D. C., & Bello, M. A. (2002). Phylogenetic relationships within *Plantago* (Plantaginaceae): evidence from nuclear ribosomal ITS and plastid *trnL-F* sequence data. *Botanical Journal of the Linnean Society, 139*, 323-338.
Rothfels, C. J., Gaya, E., Pokorny, L., Rothfels, P., Rothfels, P., & Feulner, G. R. (2012). Significant fern, lichen and bryophyte collections from the UAE and northern Oman, including five new records for the Arabian Peninsula. *Tribulus, 20*, 4-20.
Rufino, M., Rowe, E. C., J., D. R., & Giller, K. E. (2006). Nitrogen cycling efficiencies through resource-poor African crop–livestock systems. *Agriculture, Ecosystems & Environment, 112*(4), 261-282.
Rulin, D., & Lifu, T. (1985). Algal macrofossils from the Qingbaikou system in the Yanshan range of North China. *Precambrian Research, 29*(1-3), 5-14.
Saarela, J. M., Sokoloff, P. C., Gillespie, L. J., Consaul, L. L., & Bull, R. D. (2013). DNA barcoding the Canadian Arctic flora: core plastid barcodes (rbcL + matK) for 490 vascular plant species. *PLoS One, 8*(10), e77982. doi:10.1371/journal.pone.0077982
Saganuwan, A. S. (2010). Some medicinal plants of Arabian Peninsula. *Journal of Medicinal Plants Research, 4*(9), 766-788.

Saitou, M., & Nei, M. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees *Mol. Biol. Evol., 4*(4), 406-425.
Sancho, L. G., Belnap, J., Colesie, C., Raggios, B., & Weber, B. (2016). Carbon Budgets of Biological Soil Crusts at Micro-, Meso-, and Global Scales. In B. Weber, B. Budel, & J. Belnap (Eds.), *Biological Soil Crusts: An Organizing Principle in Drylands* (Vol. 226, pp. 287-304).
Sang, T., Crawford, D., & Stuessy, T. (1997). Chloroplast DNA phylogeny. Reticulate, Evolution and Biogeography. *American Journal of Botany, 84*(9), 1120-1136.
Sapkota, R., & Nicolaisen, M. (2015). An improved high throughput sequencing method for studying oomycete communities. *J Microbiol Methods, 110*(3), 33-39.
Sarkinen, T., Staats, M., Richardson, J. E., Cowan, R. S., & Bakker, F. T. (2012). How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS One, 7*(8), e43808. doi:10.1371/journal.pone.0043808
Sarwar, A. K. M. G., & Takahashi, H. (2014). Pollen morphology ofEricaL. and related genera and its taxonomic significance. *Grana, 53*(3), 221-231. doi:10.1080/00173134.2014.912346
Schlindwein, C., Wittmann, D., Martins, C. F., Hamm, A., Siqueira, J. A., Schiffler, D., & Machado, I. C. (2005). Pollination of *Campanula rapunculus* L. (Campanulaceae): How much pollen flows into pollination and into reproduction of oligolectic pollinators? *Plant Syst. Evol., 250*, 147-156.
Schultz J., Maisel, S., Gerlach, D., Muller, T., & Worlf, M. (2005). A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA, 11*, 361-364.
Schulz, K., Mikhailyum, T., DreBler, M., Leinweber, P., & Karsten, U. (2016). Biological Soil Crusts from Coastal Dunes at the Baltic Sea: Cyanobacterial and Algal Biodiversity and Related Soil Properties. *Soil Microbiology, 71*, 178-193.
Schutze, P., Freitag, H., & Weising, K. (2003). An integrated molecular and morphological study of the subfamily Suaedoideae Ulbr. (Chenopodiaceae). *Plant Syst. Evol., 239*, 257-286.
Sharawy, S. M., & Badr, A. (2014). Relationships of Astragalus L. in section Sesamei based on morphological criteria and molecular markers. *Bangladesh J. Plant Taxon, 21*(1), 1-12. doi:http://dx.doi.org/10.3329/bjpt.v21i1.19250
Sharma, S. K., Dkhar, J., Kumaria, S., Tandon, P., & Rao, S. R. (2012). Assessment of phylogenetic inter-relationships in the genus Cymbidium (Orchidaceae) based on internal transcribed spacer region of rDNA. *Gene, 495*(1), 10-15.
Shehzad, W., Riaz, T., Nawaz, M. A., Miquel, C., Poillot, C., Shah, S. A., . . . Taberlet, P. (2012). Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology, 21*, 1951-1965.
Simon, H. M., Dodsworth, J. A., & Goodman, R. M. (2000). Crenarchaeota colonize terrestrial plant roots. *Environmental Microbiology, 2*(5), 495-505.
Soltis, D. E., & Burleigh, J. G. (2009). Surviving the K-T mass extinction: new perspectives of polyploidization in angiosperms. *Proc Natl Acad Sci U S A, 106*(14), 5455-5456. doi:10.1073/pnas.0901994106
Sonstebo, J. H., Gielly, L., Brysting, A. K., Elven, R., Edwards, M., Haile, J., . . . Brochmann, C. (2010). Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Mol Ecol Resour, 10*(6), 1009-1018. doi:10.1111/j.1755-0998.2010.02855.x
Sosa, V., Mejía-Saules, T., Cuéllar, M. A., & Vovides, A. P. (2013). DNA Barcoding in Endangered Mesoamerican Groups of Plants. *The Botanical Review, 79*(4), 469-482. doi:10.1007/s12229-013-9129-4

Soule, T., Anderson, I. J., Johnson, S. L., Bates, S. T., & Garcia-Pichel, F. (2009). Archaeal populations in biological soil crusts from arid lands in North America. *Soil Biology and Biochemistry, 41*(10), 2069-2074. doi:10.1016/j.soilbio.2009.07.023

Souto, M., Castro, D., Pontevedra-Pombal, X., Garcia-Rodeja, E., & Fraga, M. I. (2017). Characterisation of Holocene plant macrofossils from North Spanish ombrotrophic mires: bryophytes. *Mires and Peat., 19*, 1-12. doi:10.19189/MaP.2016.OMB.263

Srivathsan, A., Ang, A., Vogler, A. P., & Meier, R. (2016). Fecal metagenomics for the simultaneous assessment of diet, parasites, and population genetics of an understudied primate. *Front Zool, 13*, 17. doi:10.1186/s12983-016-0150-4

Srivathsan, A., Sha, J. C., Vogler, A. P., & Meier, R. (2015). Comparing the effectiveness of metagenomics and metabarcoding for diet analysis of a leaf-feeding monkey (Pygathrix nemaeus). *Mol Ecol Resour, 15*(2), 250-261. doi:10.1111/1755-0998.12302

Staats, M., Cuenca, A., Richardson, J. E., Vrielink-van Ginkel, R., Petersen, G., Seberg, O., & Bakker, F. T. (2011). DNA damage in plant herbarium tissue. *PLoS One, 6*(12), e28448. doi:10.1371/journal.pone.0028448

Staudacher, K., Wallinger, C., Schallhart, N., & Traugott, M. (2011). Detecting ingested plant DNA in soil-living insect larvae. *Soil Biology and Biochemistry, 43*(2), 346-350. doi:10.1016/j.soilbio.2010.10.022

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology, 21*, 2045–2050.

Taberlet, P., Coissac, E., Pompanon, F., Gielly, L., Miquel, C., Valentini, A., . . . Willerslev, E. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res, 35*(3), e14. doi:10.1093/nar/gkl938

Taberlet, P., Gielly, L., Pautou, G., & Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 1105-1109.

Tackenberg, O., Poschlod, P., & Kahmen, S. (2003). Dandelion seed dispersal: The horizontal wind speed does not matter for long-distance dispersal - it is updraft! *Plant Biology, 5*, 451-454.

Tate, J. A., & Simpson, B. B. (2003). Paraphyly of Tarasa (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany, 28*, 723-737.

Telewski, F. W., & Zeevaart, J. A. D. (2002). The 120-yr period for Dr. Beal's seed viability experiment. *American Journal of Botany, 89*(8), 1285-1288.

Terrat, S., Christen, R., Dequiedt, S., Lelièvre, M., Nowak, V., Regnier, T., . . . Ranjard, L. (2012). Molecular biomass and MetaTaxogenomic assessment of soil microbial communities as influenced by soil DNA extraction procedure. *Microbial Biotechnology, 5*, 135-141.

Theodoridis, S., Stefanaki, A., Tezcan, M. C., Kokkini, S., & Vlachonasios, K. (2012). DNA barcoding in native plants of the Labiatae (Lamiaceae)family from Chios Island (Greece) and the adjacentC¸es¸me-Karaburun Peninsula (Turkey). *Molecular Ecology Resources, 12*, 620-633.

Thomas, J. (2017). Plant diversity of Saudi Arabia. Retrieved from http://plantdiversityofsaudiarabia.info/Biodiversity-Saudi-Arabia/Flora/Checklist/Cheklist.htm

Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. *Microb Inform Exp, 2*(1), 3. doi:10.1186/2042-5783-2-3

Thomsen, P. F., Elias, S., Gilbert, M. T. P., Haile, J., Munch, K., Kuzmina, S., . . . Willerslev, E. (2009). Non-destructive sampling of ancient insect DNA. *PLoS One, 4*, e5048.

Tolba, M. K. (2007). Harvest of Dust. In D. A. Bainbridge (Ed.), *Guide for Desert and Dryland Restoration* (pp. 1-12). U.S.A: Island Press.

Toole, E. H., & Brown, E. (1946). Final results of the Duvel buried seed experiment. *J Agri. Res., 72*, 201-210.

Tryon, R. (1970). Development and evolution of fern floras of Oceanic islands. *Biotropica, 2*, 76.

Tsukaya, H., & Suetsugu, K. (2014). Two New Species of Sciaphila (Triuridaceae) from Sarawak (Borneo, Malaysia). *Phytotaxa, 170*(4), 283-290.

Turini, F. G., Brauchler, C., & Heubl, G. (2010). Phylogenetic relationships and evolution of morphological characters in Ononis L. (Fabaceae). *Taxon, 59*(4), 1077-1090.

Tutel, B., Kandemir, I., Kus, S., & Kence, A. (2005). Classification of Turkish *Plantago* L. Species using Numerical Taxonomy. *Turkish Journal of Botany, 29*, 51-61.

UNCC. (2005). United Nation Compensation Commission Decision 258 (8 December, 2005). Retrieved from http://www.uncc.ch/decision/dec_258.pdf

UNFCCC. (2012). Kuwait Initial Natioanl Communications under the United Nations Framework Convention on Climate Change. *Environment Public Authority Kuwait*, 90.

Valentini, A., Miquel, C., Nawaz, M. A., Bellemain, E., Coissac, E., Pompanon, F., . . . Taberlet, P. (2009). New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Mol Ecol Resour, 9*(1), 51-60. doi:10.1111/j.1755-0998.2008.02352.x

Van De Wiel, C., Schoot, V. D., Van Valkenburg, J. C. H., Duistermaat, H., & Smudlers, M. J. M. (2009). DNA barcoding discriminates the noxious invasive plant species, floating pennywort (Hydrocotyle ranunculoides L.f.), from non-invasive relatives. *Molecular Ecology Resources, 9*, 1086-1091.

Vierheilig, J., Savio, D., Ley, R. E., Mach, R. L., Farnleitner, A. H., & Reischer, G. (2015). Potential applications of next generation DNA sequencing of 16S rRNA gene amplicons in microbial water quality monitoring. *Water Sci Technol, 72*(11), 1962-1972.

Vollmer, D., Regan, H., & Andelman, S. J. (2016). Assessing the sustainability of freshwater systems: A critical review of composite indicators. *Ambio., 45*, 765-780.

Wang, X., Liu, C., Huang, L., Bengtsson-Palme, J., Chen, H., Zhang, J., . . . Li, J. (2015). ITS1: a DNA barcode better than ITS2 in eukaryotes? *Molecular Ecology Resources, 15*(3), 573-586. doi:10.1111/1755-0998.12325

Watts, D., & Al-Nafie, A. H. (2003). Vegetation & Biogeography of Sand Sea of Saudi Arabia. *London: Kegan Paul*, 173.

Weber, B., Budel, B., & Jayne, B. (2016). *Biological Soil Crusts: An Organizing Principle in Drylands*. Switzerland: Springer International Publishing AG.

Weigand, J., & Rodgers, J. (2009). Active Restoration for the Mojave Desert. In R. H. Webb, Fenstermaker, L.F., Heaton, J.S., Hughson, D.L., Mcdonald, E.V., and Miller, D.M. (Ed.), *The Mojave Desert: Ecosystem Processes and Sustainability* (pp. 378-409). U.A.S: University of Nevada Press.

West, N. E. (1990). Structure and function of microphytic soil crusts in wildland ecosystems of arid to semi-arid regions. *Adv. Ecol. Res., 20*, 45-53.

White, F., & Leonard, J. (1991). Phytogeographical links between Africa and Southwest Asia. *Flora et Vegetatio Mundi, IX*, 229-246.

Wikum, D. A., & Shanholtzer, G. F. (1978). Application of the Braun-Blanquet cover-abundance scale for vegetation analysis in land development studies. *Environmental management, 2*(4), 323-329.

Wilson, J. J., Sing, K. W., Halim, M. R., Ramli, R., Hashim, R., & Sofian-Azirun, M. (2014). Utility of DNA barcoding for rapid and accurate assessment of bat diversity in Malaysia in the absence of formally described species. *Genet. Mol. Res., 13*(1), 920-925.

Wright, L. A. (1909). *Longevity of fern spores*: Kew.

Yamaguchi, T., Yano, S., & Tsukaya, H. (2010). Genetic framework for flattened leaf blade formation in unifacial leaves of Juncus prismatocarpus. *Plant Cell, 22*(7), 2141-2155. doi:10.1105/tpc.110.076927

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., . . . Chen, S. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS One, 5*(10). doi:10.1371/journal.pone.0013102

Yoccoz, N. G. (2012). The future of environmental DNA in ecology. *Molecular Ecology, 21*, 2031–2038.

Yoccoz, N. G., Brathen, K. A., Gielly, L., Haile, J., Edwards, M. E., Goslar, T., . . . Taberlet, P. (2012). DNA from soil mirrors plant taxonomic and growth form diversity. *Mol Ecol, 21*(15), 3647-3655. doi:10.1111/j.1365-294X.2012.05545.x

Yooseph, S., Andrews-Pfannkoch, C., Tenney, A., McQuaid, J., Williamson, S., Thiagarajan, M., . . . Venter, J. C. (2013). A metagenomic framework for the study of airborne microbial communities. *PLoS One, 8*(12), e81862. doi:10.1371/journal.pone.0081862

Yu, N., Gu, H., Wei, Y., Zhu, N., Wang, Y., Zhang, H., . . . Sun, A. (2016). Suitable DNA Barcoding for Identification and Supervision of Piper kadsura in Chinese Medicine Markets. *Molecules, 21*(9), 1-10. doi:10.3390/molecules21091221

Zaman, S., Padmesh, S., Bhat, N., & Tawfiq, H. (2009). Germination Characteristics and Storage Behavior of *Tamarix aucheriana* (Dence.) *Seeds European Journal of Scientific Research, 26*(4), 533-539.

Zanten, B. O. (1978). Experimental studies on transoceanic long-range dispersal of moss spores in the southern hemisphere. *The Journal of the Hattori Botanical Laboratory, 44*, 455-482.

Zhang, C. Y., Wang, F. Y., Yan, H. F., Hao, G., Hu, C. M., & Ge, X. J. (2012). Testing DNA barcoding in closely related groups of Lysimachia L. (Myrsinaceae). *Mol Ecol Resour, 12*(1), 98-108. doi:10.1111/j.1755-0998.2011.03076.x

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics, 30*(5), 614-620.

Zhang, W., Fan, X., Zhu, S., Zhao, H., & Fu, L. (2013). Species-specific identification from incomplete sampling: applying DNA barcodes to monitoring invasive solanum plants. *PLoS One, 8*(2), e55927. doi:10.1371/journal.pone.0055927

Zhang, W., Yuan, Y., Yang, S., Huang, J., & Huang, L. (2015). ITS2 Secondary Structure Improves Discrimination between Medicinal "Mu Tong" Species when Using DNA Barcoding. *PLoS One, 10*(7), e0131185. doi:10.1371/journal.pone.0131185

Zhang, Y. M., Chen, J., Wang, L., Wang, X. Q., & Gu, Z. H. (2007). The spatial distribution patterns of biological soil crusts in the Gurbantunggut Desert, Northern Xinjiang, China. *Journal of Arid Environments, 68*, 599-610.

Zimmermann, H. H., Raschke, E., Epp, L. S., Stoof-Leichsenring, K. R., Schwamborn, G., Schirrmeister, L., . . . Herzschuh, U. (2017). Sedimentary ancient DNA and pollen reveal the composition of plant organic matter in Late Quaternary permafrost sediments of the Buor Khaya Peninsula (north-eastern Siberia). *Biogeosciences, 14*(3), 575-596. doi:10.5194/bg-14-575-2017

Zohary, M. (1973). Geobotanical foundations of the Middle East. *Taylor & Francis, UK, 1*, 1-340.