Systematic and genomic studies in the genus *Aubrieta* (Brassicaceae)

A thesis submitted to the University of Leicester for the degree of Doctor of Philosophy

By

Jotyar Jassim Muhammed M.Sc.

November 2016

Department of Genetics

Abstract

The study focuses on the herbaceous perennial plant genus *Aubrieta* (Brassicaceae). Distributed from Armenia through the Levant and Anatolia to Greece, the Balkans and Italy, the species have proved difficult to distinguish. About 12 species are currently recognised, although estimates range from only one up to about 20 or more. Furthermore, their evolutionary relationships are unknown. In order to remedy this situation molecular and cytogenetical studies were conducted.

Next Generation Sequencing (NGS) data were produced to generate a complete chloroplast genome for four species of *Aubrieta* in order to confirm the phylogenetic position of the genus in the family. Earlier suggestions that it belongs in tribe Arabdieae were confirmed. Details of plastome structure were analysed in *A. gracilis*, which was shown to have 88 protein-encoding, 37 transfer-RNA and eight ribosomal-RNA genes.

A phylogenetic analysis of the genus was conducted based on chloroplast and nuclear sequences as well as mitochondrial RFLPs. Five chloroplast regions (matK, trnD-trnT, ycf6-psbM, rps11-rpl36 and trnH-psbA) and two nuclear genes (duo1 and rbp2) were amplified and sequenced successfully. Six mitochondrial gene regions (Orf114, Nad9-ccmFN2, Orf25, matR, ccm FC, trnK-rps3) were studied by means of restriction enzymes. Data analyses show that *Aubrieta* comprises the annual, pan-Mediterranean *Arabis verna* plus perennial taxa that fall into one of five geographically delimited gene pools: i) Near East (Levant, Iraq and Iran); ii) Anatolia; iii) Aegean Basin; iv) Greece, Albania, Bulgaria (Pindus Mts and associated ranges); and v) Trans-Adriatic Sea and Sicily. There was some disagreement between the plastid and nuclear trees, which was attributed to hybridisation, chiefly affecting the taxa occupying the Aegean Basin.

Evolution in the genus appears to have proceeded largely at the diploid level (2n=2x=16). In order to see what changes at the chromosome level have accompanied speciation, fluorescence in situ hybridization (FISH) studies were conducted. The probes pTa71 and pTa794 were used to locate the position of 45S and 5S rDNA sites on the chromosomes. The number of 45S rDNA sites are 2, 4, 5, or 6, localized on short-arms, the centromere and on long-arms. The number of 5S rDNA sites is a constant two, located either on short-arms or long-arms. These rDNA sites (45S and 5S) are either located on different chromosomes or shared by one or two chromosomes. Speciation is accompanied (promoted?) by translocations and duplications. Hybridisation was confirmed in the genus.

The timing of the various bifurcations in the evolutionary tree were estimated from a study of the concatenated chloroplast sequences, but the major split into an annual lineage (*Arabis verna*) and a perennial lineage appears to date from 1.4 Mya. The hybridisation events involving the Aegean Basin taxa appear to date from the early to mid-Pleistocene (ca 600-800 Kya), a time when considerable parts of the Aegean were above sea level.

The taxonomy of the genus is still problematic, it being impossible to diagnose the five geographical genepools by means of morphological characters. Instead, a splitting approach is adopted whereby regional or local phenotypes are recognised as species. This can be unsatisfactory in some cases where there is considerable morphological, but not geographical, overlap. A total of 21 species, including *Arabis verna* which is recombined into *Aubrieta*, is recognised.

Declaration

I hereby declare that no part of this thesis has been previously submitted to this or any other university as part of the requirements for a higher degree. The content of this thesis is the result of my own experimentation and data analysis unless otherwise acknowledged in the text or by reference.

The work was conducted in the Department of Genetics, University of Leicester, during the period January 2013 to December 2016.

Signed …….....

Jotyar J. Muhammed

Acknowledgement

First and foremost, I would like to thank my two supervisors Dr Richard Gornall and Prof. Pat Heslop-Harrison, for their continuous guidance and advice during my PhD study. I would also like to say thanks to Dr Trude Schwarzacher and Dr John Bailey for their help and encouragement. I express my deepest gratitude to Dr Rob Hammond for being my internal departmental assessor and who was always willing to advise me.

I am deeply grateful to my sponsor (Ministry of Higher Education and Scientific Research, Kurdistan region of Iraq- HCDP program) for their funding of my PhD project.

I wish to thank my parents for their moral and financial support and also my brothers and sisters for being encouraging and supportive.

I reserve very special thanks to my wife Zakeia Mahmood for standing beside me throughout the study.

Finally, I would like to thank all my friends and, in particular, members of Lab 201.

CONTENTS

CHAPTER 1. INTRODUCTION

The Mediterranean Basin contains 11, 879 islands, which together constitute one of the world's largest archipelagos, fourth in rank behind the Caribbean, Indonesia and the South Pacific (Blondel *et al*. 2010). Perhaps more remarkably, over 80% of the islands are to be found in the Aegean Sea, a region that lies at the junction of the continents of Europe, Asia and Africa. This proximity to three very different floral and faunistic regions has led to a very high level of species diversity in the area, including exceptional levels of endemsim. All of this makes the region of great interest for biodiversity studies (Bittkau and Comes 2005; Lymberakis and Poulakakis 2010). On account of the diverse geological and climatic events since the late Tertiary, and the fact that some islands are continental whereas others are of tectonic origin, the area lends itself to studies of vicariance, dispersal, genepool fragmentation and speciation, including adaptive and non-adaptive radiation.

The fauna and flora of islands are vulnerable and susceptible to the exigencies of the environment. This is partly due to smaller population sizes and geographic ranges and comparatively simpler ecosystem relationships (Parent *et al*. 2008) but also to changes caused by human activities. Indeed the Aegean Basin has been the focus of human-induced change for over 8000 years. Playing host to several ancient civilisations, including the Minoan, Cycladic and Myceneaum, the landscape has been modified by human activity (Triantis and Mylonas 2009). Thus, the possibility of human-mediated dispersal must be taken into account in any evaluation of the phylogeographic history of biota in the Aegean archipelago.

Previous studies of the biogeography of the Aegean Basin and surrounding areas have focussed primarily on animals (Poulakakis *et al*. 2015), with more than 75 studies focussing on invertebrates (Douris *et al*. 1995; Gantenbein and Largiader 2002; Parmakelis *et al*. 2005, 2006a, b, 2008), or vertebrates (Beerli *et al*. 1996; Kasapidis *et al*. 2005a, b; Lymberakis *et al*. 2007; Poulakakis et al. 2003, 2005a, b, c, 2008a, b; Weisrock *et al*. 2001). Available evidence (Poulakakis *et al*. 2015) suggests that most terrestrial animals diversified towards the end of the Messinian Salinity Crisis (5.96—5.33 mya), when the Mediterranean basin dried up and that the formation (ca 12—9 mya) of the mid-Aegean Trench was not a factor in their evolution. But, to date, there has been relatively little attention paid to plants, with very few studies being available. The only comprehensive studies are by Bittkau and Comes (2005) and Comes *et al*. (2008) on *Nigella* (Ranunculaceae), although there have been others in a Medterranean-wide setting (Koch *et al*. 2006; Feliner 2014). It is in this context, that the present study has been conceived. The focus is on the genus *Aubrieta* Adanson, a genus of

between 1 and 22 species in the cabbage family (Brassicaceae), which are distributed across the Near East, Turkey, Greece, the Balkans and Italy, thus straddling the Aegean Basin. The uncertainty about the number of species highlights the fact that a critical taxonomic study of the genus is also needed.

To provide the necessary science for a resolution of the biogeography and systematics of the genus, the main focus of the study is on reconstructing a phylogeny. To provide the necessary taxonomic context, my account begins with a review of the phylogeny of the Brassicaceae and of the tribe Arabideae, to which *Aubrieta* has been claimed to belong. A description of the specific aims and objectives of the thesis then follows and completes the Introduction.

1.1 Brassicaceae: taxonomy and phylogeny

The Brassicaceae belongs to the angiosperm order Brassicales (Rosids, Malvanae). This is one of the largest taxonomic orders in the plant kingdom. It consists of approximately 17 families, 398 genera and 4765 species (Stevens 2001), as depicted in the phylogenetic tree shown in Fig. 1.1. There is some controversy regarding estimates of the age of the order. Gandolfo *et al*. (1998) proposed that some fossils, such as *Dressiantha*, were of Turonian age (ca 89.5 Mya); similarly, stem Brassicales taxa appear to date from 81–89 Mya and diversified approximately 10 million years later. In contrast, crown group Brassicales were estimated to have originated at 57—75 Mya (Wikstrom *et al*. 2001).

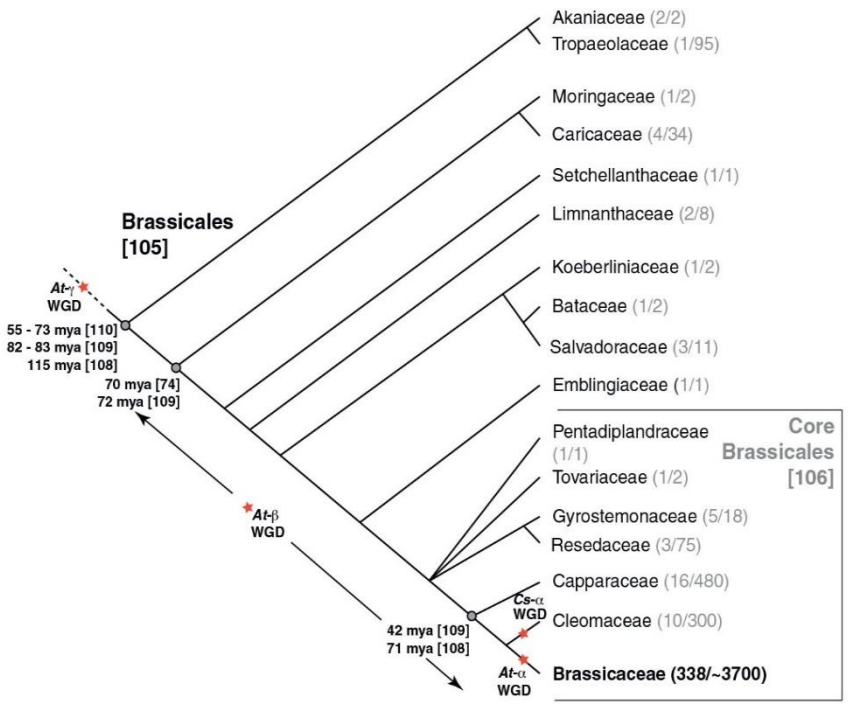


Figure 1.1. Phylogenetic tree of the Brassicales showing the constituent families (Stevens 2001) and some estimated divergence times (Wikström *et al.* 2001 [74]; Stevens 2001 [105]; Beilstein *et al.* 2010 [108]; Bell *et al.* 2010 [109]; Wang *et al.* 2009 [110]. Figure taken from Franzke *et al*. (2011).

By far the largest family in the Brassicales is the type family, Brassicaceae (= Cruciferae). It comprises 51 tribes, 321 genera and ca 3,700 species (Al-Shehbaz 2012; Al-Shehbaz *et al*. 2014; German and Friesen 2014), and includes many economically important species, such as cabbage and mustard, as well as ornamentals like stocks (*Mathiola*), wallflowers (*Erysimum*) and aubretia (*Aubrieta*) (Franzke *et al.* 2009), not to mention the laboratory model species *Arabidopsis thaliana*. Members of the family can be distinguished easily from other families by the cruciform shape made by the four petals, and by the tetradynamous stamens comprising an outer series of two short ones and an inner series of four taller ones. There are two united carpels that mature into a long, narrow pod (silique) or a short, wide pod (silicula).

Many authors have tried to classify the mustard family into tribes by using morphological characters, including fruits, embryos and cotyledons (Hayek 1911; Schulz 1936; Janchen 1942; Al-Shehbaz 1984). According to Schulz (1936) the Brassicaceae comprised ten tribes. However, six years later Janchen (1942) recognised an additional 19 tribes. By the end of the 1990s, numerous molecular studies focusing on DNA sequences of various regions (rbcL, ndhF, matK, trnL-F from the chloroplast genome and ITS, PHYA, adh, chs, from the nuclear genome) had been conducted in order to test these earlier classifications against the criterion of monophyly (=holophyly) (Price *et al*. 1994; Zunk *et al*. 1999; Beilstein *et al*. 2006; Bailey *et al*. 2005; Koch *et al*. 2007; Beilstein *et al*. 2008; Couvreur *et al*. 2010; Hohmann *et al*. 2015; Huang *et al*. 2015). In addition, more detailed morphological and anatomical studies, including of roots and stems, leaves, trichomes and seeds, have been conducted (Al-Shehbaz *et al*. 1999; Beilstein *et al*. 2006, 2008). The consequence of these studies is that the family is now recognised as containing 51 tribes. Furthermore, the limits of various genera have been re-drawn (Al-Shehbaz 2005; Al-Shehbaz *et al*. 2005, 2006).

In terms of the major clades in the Brassicaceae, Franzke *et al*. (2011) revealed that the family comprises three major lineages (I, II, and III; Fig. 1.2). These lineages were confirmed by a phylogenetic analysis of 213 orthologous genes (Kagale *et al*. 2014), although only 13 tribes were represented. Huang *et al.* (2015), using sequence data of 113 nuclear genes from 87 species representing 45 genera spanning 29 of the 51 tribes, produced a phylogenetic tree in which six clades were resolved (Fig. 1.3).

In the tree of Franzke *et al*. (2011), the genus *Aubrieta* is positioned in tribe Arabideae in 'expanded lineage II' (Fig. 1.2), a position supported by the analysis of Kagale *et al*. (2014). However, in the more recent and well resolved tree of Huang *et al.* (2015), expanded lineage II may not be monophyletic (Fig. 1.3), but tribe Arabideae was not sampled.

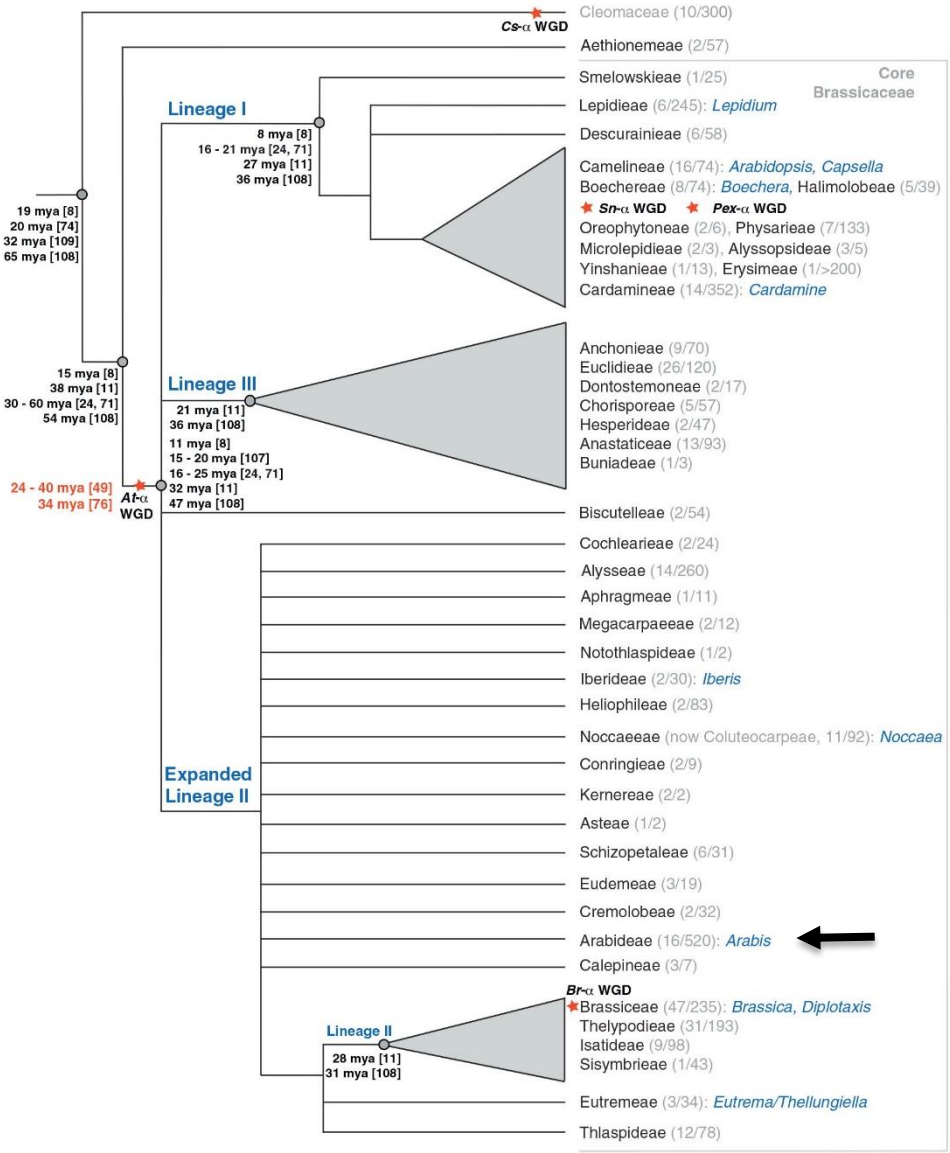


Figure 1.2. Phylogenetic tree of Brassicaceae showing its division into lineages I, II and III (Beilstein *et al*. 2006, 2008; Franzke *et al*. 2011), the relationships between the currently recognised 44 tribes (Franzke *et al*. 2011), and estimated divergence times between them (Franzke *et al.*2009 [8]; Couvreur *et al.*2010 [11]; Koch *et al.*2000 [24]; Koch *et al.* 2001 [71]; Wikström *et al.* 2001 [74]; Yang *et al.* 1999 [107]; Beilstein *et al.* 2010 [108]; Bell *et al.* 2010 [109]. Arrow indicates putative position of *Aubrieta*. Figure adapted from Franzke *et al*. (2011).

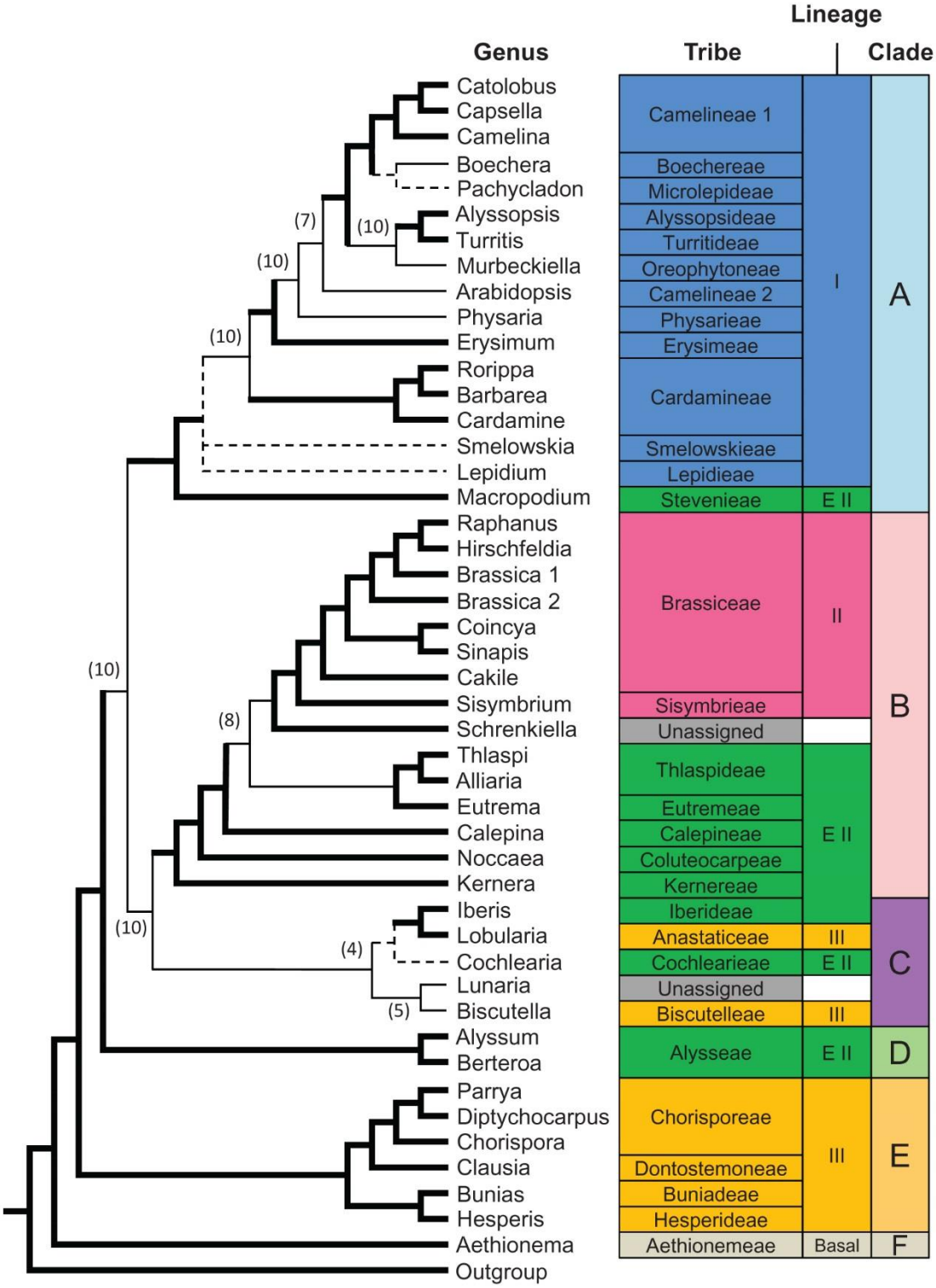


Fig 1.3. Brassicaceae phylogeny based on a concatenation of 113 nuclear genes sequences. Branching consistent in all reference trees are drawn in thick (all showing maximum support) or thin solid lines (with values showing the number of results with maximum support). Dashed lines indicate uncertainties. Figure taken from Huang *et al.* (2015).

1.2 Polyploid evolution in the Brassicaceae

All flowering plants show the genomic signature of at least two rounds of ancient polyploidy (Jiao *et al*. 2011; Amborella Genome Project 2013), and it has been estimated that almost 15% of angiosperm speciation events can be attributed to polyploidy (Wood *et al*. 2009). It has also been suggested that nearly 50% of species in the Brassicaceae are recently-evolved polyploids (Franzke *et al*. 2011) and it seems that repeated rounds of polyploidy have featured prominently in evolution within the family (Lysak and Koch 2011). In this context, polyploid events, or whole genome duplications (WGDs), have been categorised into three main types, based on their time of origin: i) neo-, ii) meso- and iii) paleopolyploids. The first are of recent origin and are characterised by increased genome size, higher chromosome number, redundant gene content and, if of hybrid origin (allopolyploids), extant diploid ancestors (Ramsey and Schemske 2002; Mandáková *et al*. 2010b). Over time, neopolyploids become diploidised through gene- and chromosome-loss and large-scale rearrangements (Song *et al*. 1995; Lynch and Conery 2000; Wolfe 2001; Mandáková *et al*. 2010a), and in so doing transition to the mesopolyploid and finally the palaeopolyploid state (Mandáková *et al*. 2010b). Parental subgenomes of meso-allopolyploids become identifiable only through comparative genetic and genomic approaches (Parkin *et al*. 2005; Mandáková *et al*. 2010b; Parkin *et al*. 2014), whereas in palaeo-allopolyploids the extent of genome restructuring is so widespread that the parental genomes become assimilated to the point where they are no longer identifiable (Mandáková *et al*. 2010b).

Within the Brassicaceae, all species show signs of three palaeopolyploid events, known as the α, β and γ WGDs (Haudry *et al*. 2013); these events were first revealed in studies of *Arabidopsis thaliana* (Bowers *et al*. 2003). Furthermore, an additional later mesopolyploidy (whole-genome triplication) event in diploid Brassicas was revealed through comparative genetic mapping (Parkin *et al*. 2005), cytogenetic studies (Lysak *et al*. 2005; Ziolkowski *et al*. 2006), and the whole-genome sequencing of *Brassica rapa* (Wang *et al*. 2011).

The designation as palaeo-, meso- or neopolyploids comes from Bayesian dating analysis and fossil information as age constraints. Thus the palaeopolyploid γ event at, which involved a triplication, has been dated at ca 135 mys; the β event, which was a duplication correlated with the origin of the Brassicales, is dated at 82—124 mya; and the α event, another duplication, correlated with the origin of the Brassicaceae, is dated at 40—47 mya (Lysak and Koch 2011; Kagale *et al*. 2014; Edger *et al*. 2015; Magallón *et al.* 2015). The age of the mesopolyploid triplication event in the diploid Brassica group has now been estimated as 22.5 mya (Beilstein

et al. 2010). Recent studies have revealed numerous neo- and mesopolyploid events in different parts of the family (Joly *et al*. 2009; Mandáková *et al*. 2010a, 2010b, 2012; Haudry *et al*. 2013; Kagale *et al*. 2014; Liu *et al*. 2014; Parkin *et al*. 2014), which implies an important role for recurrent polyloidy in the evolution of the family, and especially in the early evolution of the various tribes (Lysak and Koch 2011). Kagale *et al*. (2014) demonstrated at least five, relatively recent (7—12 mya) polyploid events and concluded that major diversification in the family took place in the Neogene period (0—23 mya), with species divergence and polyploidy events being correlated with prolonged, unstable climatic conditions. Kagale *et al*. (2014) suggested that genome duplication may have conferred higher adaptability and increased tolerance toward major environmental change, thus facilitating species radiation.

1.3 Tribe Arabideae

Based on molecular data analysed from the internal transcribed spacer region (ITS), Al-Shehbaz *et al*. (1999) and Koch *et al*. (1999) regarded early circumscriptions of the Arabideae as an unnatural (i.e. not monophyletic) group. Following more recent studies, however, and consequent taxonomic re-adjustments, the tribe Arabideae has now been made monophyletic. It contains 17 genera (*Abdra*, *Arabis*, *Arcyosperma*, *Athysanus*, *Aubrieta*, *Baimashania*, *Borodiniopsis*, *Botschantzevia*, *Dendroarabis*, *Draba*, *Drabella*, *Pachyneurum*, *Parryodes*, *Pseudodraba*, *Scapiarabis*, *Sinoarabis* and *Tomostima*) and nearly 500 species (Al-Shehbaz 2012). The two largest genera, *Draba* and *Arabis*, contain 363 and ca 100 species respectively. Both are reckoned to be non-monophyletic. According to a recent study (Karl and Koch 2013), *Aubrieta* is nested within a polyphyletic *Arabis*, along with *Arabis verna* (Fig. 1.4). A total of eight clades can be recognised within the tribe: 1) *Scapiarabis* clade; 2) main *Arabis* clade; 3) *Aubrieta* clade; 4) *Arabis aucheri* clade; 5) *Draba* clade; 6) *Arabis auriculata* clade; 7) *Arabis nordmanniana* clade; and 8) *Arabis alpina* clade (Fig. 1.4).

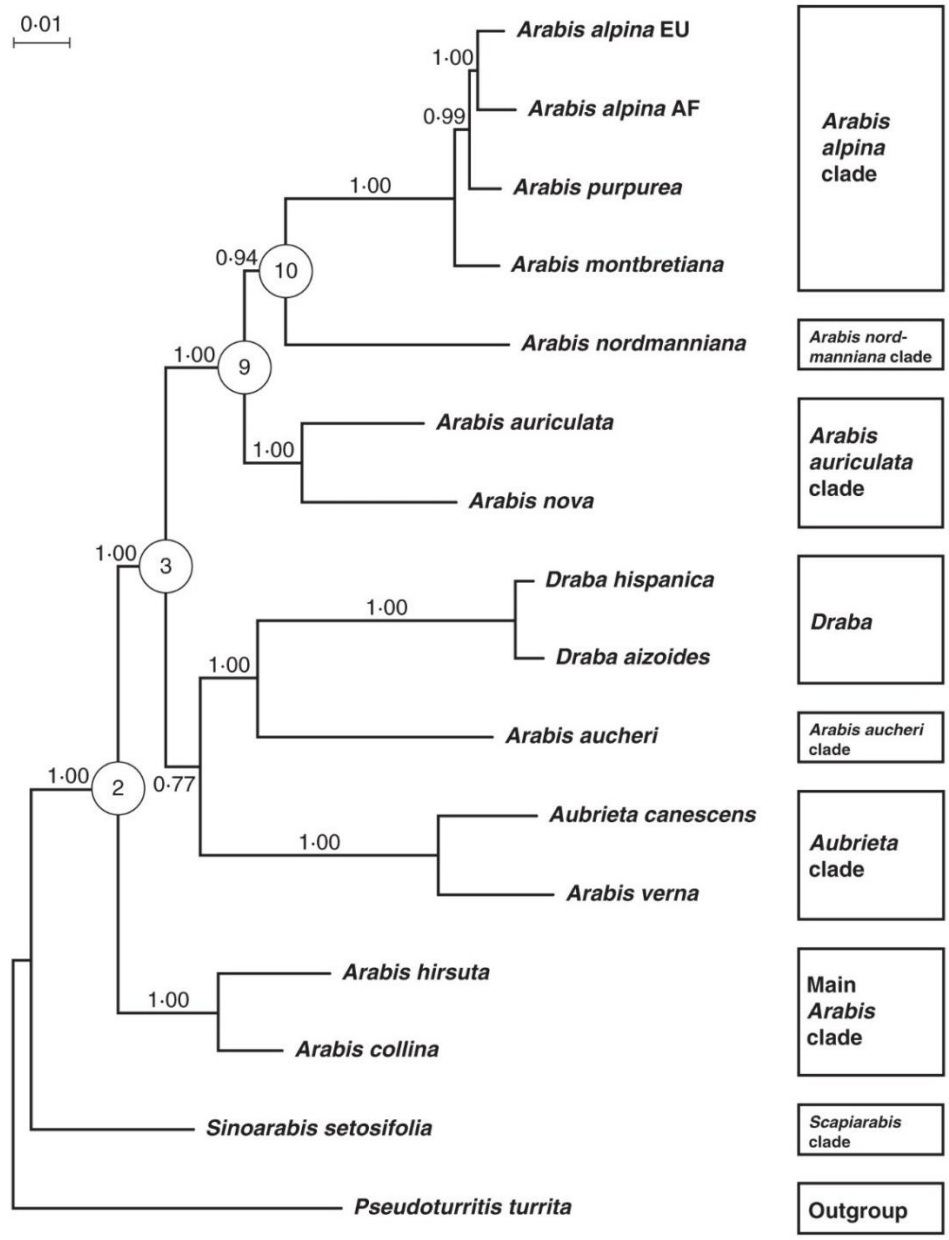


Fig 1.4. Phylogenetic tree of tribe Arabideae based on Bayesian analysis of nuclear ITS and chs and chloroplast trnL-F genes. The numbers on the nodes refers to posterior probability values. The boxes on the right refer to the clades recognised. Figure taken from Karl and Koch (2013).

1.4 The genus *Aubrieta*

1.4.1 Taxonomy

Aubrieta can be distinguished from its close relatives in *Arabis* and *Draba* by its stamen filaments, which are somewhat winged and the outer ones bear a tooth-like appendage; and even the basal leaves are sessile with a cordate base (Boissier 1867). The genus is widely regarded as being a difficult one taxonomically (Mattfeld 1939; Cullen 1965; Gustavsson 1986;

Akeroyd and Ball 1993). In the only monographic study of *Aubrieta* Adanson to date, Mattfeld (1939) recognised 13 species with 10 varieties. Perusal of recent regional accounts by Cullen (1965), Gustavsson (1986) and Phitos (2002) suggests that there may be about 20 species, although the comprehensive review of the family by Al-Shehbaz *et al*. (2006) credited only twelve 12 species and Voss (1896) despairingly lumped all taxa into a single species. Study of the *International Plant Names Index*, together with a careful reading of the regional Floras of Greece (Gustavsson 1986; Phitos 2002), Italy (Zangheri and Brilli-Cattarini 1976), Bulgaria (Stojanov and Stefanov 1924, 1933, 1948; Stojanov *et al*. 1966) and Turkey (Cullen 1965) shows that there are 122 names that have been attributed to taxa at various ranks from species to forma. For the purposes of this study, I have adopted the taxa recognised in *Atlas Flora Europaea* (Jalas and Suominen 1994) and in *Flora of Turkey* (Cullen 1965) as working hypotheses (Table 1), against which my findings can be tested. Any necessary revisions will be made at the end of the study.

1.4.2 Morphology

The plants are herbaceous perennials, usually forming mats or cushions. The species have an indumentum of three types of hair, stellate, furcate and simple. Leaves are alternate, simple, with margins entire or 1—7-toothed. The inflorescence is a compact, few-flowered raceme, which becomes much elongated in fruit. Sepals are erect and the inner pair are saccate at the base. Petals are cuneate at the base and rounded or toothed distally, and colour ranges from pink to violet, rarely white. Stamen filaments are slightly winged; the outer ones have an appendage. The gynoecium consists of a superior ovary containing biseriately-arranged ovules, a cylindrical style and a capitate stigma. Fruits vary in shape from long and narrow (silique) to short and wide (silicula) and from compressed to inflated.

Many *Aubrieta* species are not only polymorphic but also very similar morphologically and there is considerable dispute over species limits. In the past, a variety of morphological characters has been considered important in distinguishing species, including leaf shape (Mattfeld 1939), stamen filament morphology, fruit shape and size (Boissier 1867; Wettstein 1892), and anatomy of the fruit hairs (Bornmuller 1936).

Table 1 Provisional list of *Aubrieta* taxa; this working taxonomy adopts the taxa recognised in *Atlas Florae Europaeae* (Jalas and Suominen 1994) and *Flora of Turkey* (Cullen 1965) with the addition of one recently described new species (Yüzbaşıoğlu *et al*. 2015). For simplicity, however, the names used are at species rank where such are available.

No.	Taxon	Distribution
1	*Aubrieta italica* Boiss.	S. Italy (Monte Gargano, Basilicata)
2	*Aubrieta columnae* Guss.	Italy, Balkans
2.1	subsp. *columnae*	Italy: C. and S. Appennini
2.2	subsp. *bulgarica* Ančev	Bulgaria (Valley of Strouma River)
2.3	subsp. *pirinica* Assenov	S. Bulgaria (Pirin Plania)
3	*Aubrieta croatica* Schott	Albania, W. Jugoslavia, SW. Romania
4	*Aubrieta gracilis* Spruner ex Boiss.	C. Greece
5	*Aubrieta glabrescens* Turrill	Greece (N. Pindhos: Smolikas)
6	*Aubrieta scardica* (Wettst.) L.-Å. Gustavsson	S. Jugoslavia and N. Albania, S. to Sterea Ellas
7	*Aubrieta intermedia* (Boiss.) Bald.	S. Balkan Peninsula
7.1	var. *intermedia*	C. Greece
7.2	var. *macedonica* Adamović	Macedonia
8	*Aubrieta deltoidea* (L.) DC.	S. Greece, Aegean region; Sicily; naturalized in S. and W. Europe
8.1	var. *deltoidea*	Aegean region
8.2	var. *cithaeronea* Mattf.	E.C. Greece
8.3	var. *graeca* (Griseb.) Regel	Greece (Attic Mts and Peloponnesus)
8.6	var. *microphylla* Boiss.	Greece (Kefallinia, Karpathos)
8.7	var. *sporadum* (Phitos) Phitos	Greece (W. Aegean)
8.8	var. *sicula* Strobl	Sicily
9	*Aubrieta erubescens* Griseb.	Greece (Athos)
10	*Aubrieta thessala* H.Boissieu	Greece (Olimbos)
11	*Aubrieta scyria* Halácsy	Greece (Euboea, Skiros)
12	*Aubrieta olympica* Boiss.	Turkey (Bursa, Samsun)
13	*Aubrieta ekimii* Yüzb., Al-Shehbaz & M.A. Koch	Turkey (Istanbul)
14	*Aubrieta pinardii* Boiss.	Turkey (Afyon, Burdur, Konya)
15	*Aubrieta canescens* (Boiss.) Bornm.	Turkey (Adiyaman, Antalya, Cankiri, Kayseri, Konya, Maras Sivas, Tunceli)
15.1	subsp. *canescens*	Turkey (Antalya, Konya, Maras)
15.2	subsp. *cilicica* (Boiss.) Cullen	Turkey (Antalya, Icel, Seyhan)
15.3	subsp. *macrostyla* Cullen & Huber-Morath	Turkey (Adiyaman, Cankiri, Maras, Kayseri, Konya, Sivas, Tunceli)
16	*Aubrieta anamasica* H.Pesmen & A.Guner	Turkey (Isparta)
17	*Aubrieta vulcanica* Hayek & Siehe	Turkey (Nigde)
18	*Aubrieta parviflora* Boiss.	Turkey (Mardin); Iran (Shiraz); N.Iraq
19	*Aubrieta libanotica* Boiss.	Lebanon, Syria

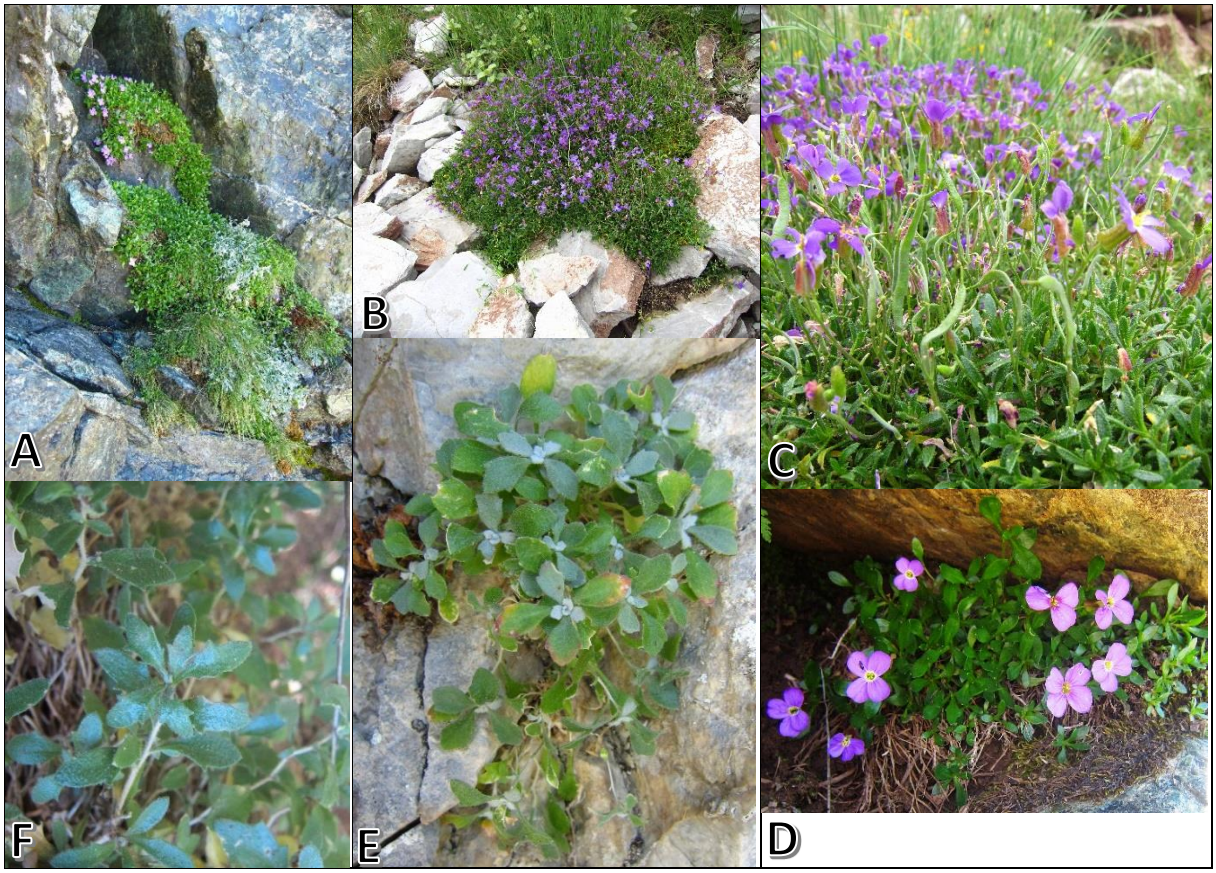


Figure 1.5. *Aubrieta* gross morphology. A&B: mat- or cushion-like habit; C&D: inflorescence structure. E&F: leaf shapes. Images taken by the author in Greece.

1.4.3 Geography

Aubrieta species are distributed from Anatolia through the Balkan Peninsula to Italy, including Sicily (Mattfeld 1939; Ančev and Goranova 2009), a geographical range illustrated in Fig. 1.6.

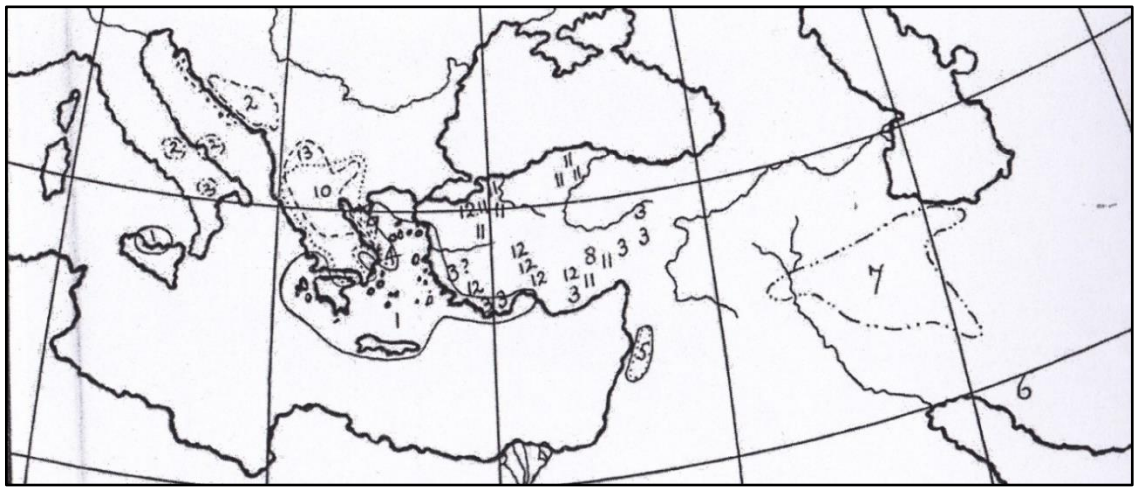


Figure 1.6. Geographical distribution of *Aubrieta* species (Mattfeld, 1939). 1: *Aubrieta deltoidea*; 2: *A. columnae*; 3: *A. canescens*; 4: *A. scyria*; 5: *A. libanotica*; 6: *A. parviflora*; 7: *A. kotschyi*; 8: *A. vulcanica*; 9: *A. erubescens*; 10: *A. intermedia*; 11: *A. olympica*; 12: *A. pinardii*; 13: *A. gracilis*.

Detailed distribution maps for European taxa have been presented by Jalas and Suominen (1994) and are shown below (Fig. 1.7). As can be seen, several taxa are narrow endemics and therefore may be of conservation concern.

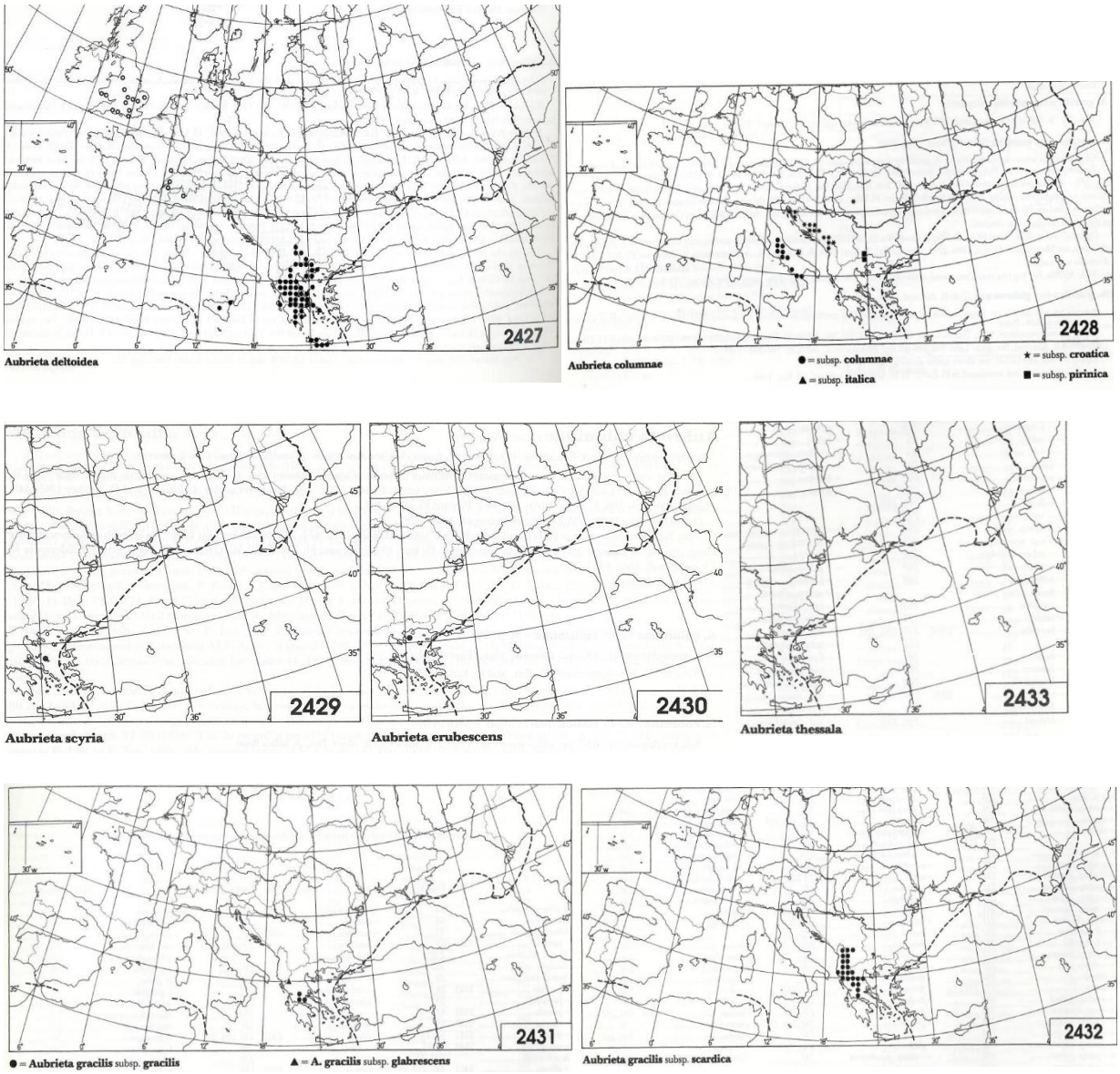


Figure 1.7. Distributions of European taxa of *Aubrieta* (Jalas and Suominen 1994).

1.4.4 Ecology

Aubrieta species grow in the mountains at altitudes ranging from 200—2900 m. They inhabit flushes on cliff-faces, rock crevices and scree slopes, in full sun or in the shade (Fig. 1.8). They occur on granite or limestone.

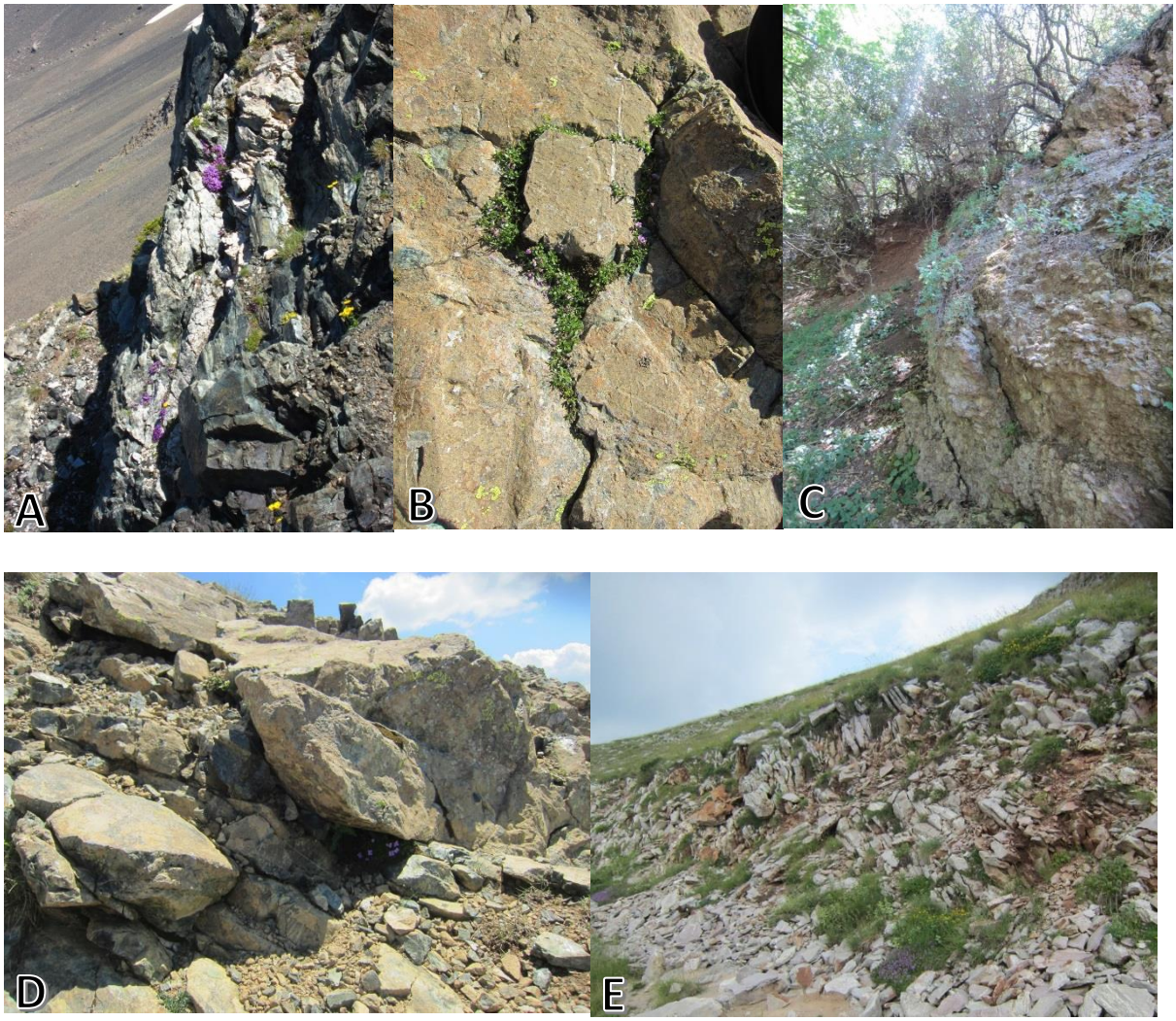


Figure 1.8. Habitats typical of *Aubrieta* species: A) cliff-faces; B) Rock crevices; C) Woodland clearings; D) Shady rocks; E) Scree slopes. Images taken by the author in Greece.

1.4.5 Reproductive biology and hybridisation

Very little has been published on the reproductive biology of the species and there has certainly been no systematic study of this topic in the genus. All species appear to reproduce sexually, via seeds, and vegetatively by means of rhizomes; sometimes the sprawling stems may also root at the nodes. *Aubrieta parviflora* has been reported to be protogynous (Al-Shehbaz 1977). Ančev and Goranova (2009) reported that in Bulgaria *A. columnae* subsp. *pirinica* and subsp. *bulgarica* as well as *A. scardica* are also protogynous and their flowers are visited by bees (Andrenidae) and by hover-flies (Syrphidae). My own observations in Greece of *Aubrieta*

gracilis show that a wide range of insects visit the flowers, including Lepidoptera (Erebidae), Diptera (Syrphidae), Hemiptera and Coleoptera (Fig. 1.9).

Figure 1.9. Flower visitors to *Aubrieta* species. A: Cinnabar moth (Lepidoptera: Erebidae); B, hover-fly (Diptera: Syrphidae); C, Hemiptera; D, small Coleoptera. Images taken by the author in Greece.

Reports of hybridisation in the genus are rare, although it has obviously been important in a horticultural context. In the wild, the hybrid between *A. gracilis* and *A. intermedia* has been reported and given the binomial *Aubrieta* × *hybrida* (Haussknecht 1893). In cultivation, *A. erubescens* can form synthetic hybrids with *A. thessala* (Phitos 2002) and Bergmans (1924) described *Aubrieta* × *cultorum* as a name to be applied to the myriad cultivars derived from undocumented horticultural crosses, real and hypothesised, between *A. deltoidea*, *A. libanotica*, *A. olympica*, *A. erubescens* and *A. kotschyi*. However, no scientific study of the parentage of the garden cultivars has ever been reported in the literature.

1.5 Aims and objectives of the study

The overall aim of the study is to provide data relevant to an understanding of how the genus has evolved and thereby recommend an appropriate taxonomy. Individual objectives are:

- To confirm the precise phylogenetic position of *Aubrieta* within the Brassicaceae by means of whole genome sequences (Chapter 3).
- To understand the evolutionary relationships of taxa within the genus *Aubrieta* (chapter 4).
- To investigate what chromosomal (karyotype) changes have accompanied speciation in the genus (chapter 5).
- To interpret the phylogeographical history of the genus, by identifying and dating vicariance and dispersal events (chapter 6).
- To draw taxonomic conclusions (chapter 7).

CHAPTER 2. MATERIALS AND METHODS

2.1 Plant material

The plant material used in the phylogenetic study was drawn largely from herbarium specimens and from silica-gel-dried material collected by me in the wild (Table 2.1). It was supplemented by fresh material from plants of diverse origins grown at the University of Leicester Botanic Garden (Table 2.1). Altogether 45 accessions, provisionally allocated to 18 taxa, from across the geographical range of the genus were investigated. As already mentioned the identification of taxa both in the wild and on herbarium sheets is extremely problematic owing to the complex patterns of morphological variation. Other taxonomists have had a similar experience; indeed, Yüzbaşıoğlu *et al*. (2015) noted that, frequently, material studied by them had been misidentified. Notably, however, the question of who is the arbiter of a correct identification was not addressed. In this study, I have done my best to assign an identity based on morphology, using the available keys in the various floristic accounts, but often have also provided this name together with the locality (country and province) from which the accession came. It was hoped that the pattern of morphological variation would begin to make more sense in a sound phylogeographic context.

Table 2.1 List of wild-collected, herbarium and cultivated material of *Aubrieta* species and their sources.

Pop No	***Country***	***Province / Prefecture***	***Locality***	***Latitude °N***	***Longitude °E***	***Collected as taxon***	***Herbarium voucher***	***Accession code***	***Collectors***	***Collection dates***
1	Spain	Jaen				*Arabis verna*	1032	S18	J. R. Edmondson	16/05/1972
2	Italy	Sicily	Palermo, Cefalu	38.116	13.366	sicula	63867	E19	D. Davis & S. Sutton	s.d.
3	Greece	Ioannina	Smolikas	40.089	20.925	glabrescens	2015-14	S7e	R.J. Gornall & J. Muhammed	12/07/2015
4	Greece	Ioannina	Smolikas	40.089	20.925	scardica	2015-15	S8a	R.J. Gornall & J. Muhammed	12/07/2015
5	Greece	Ioannina	Timfi	38.674	21.66	intermedia	2015-010	S5b	R.J. Gornall & J. Muhammed	09/07/2015
6	Greece	Ioannina	Timfi	38.936	21.808	gracilis	2015-011	S6c	R.J. Gornall & J. Muhammed	10/07/2015
7	Greece	Achaia	Kalavryta	38.033	22.116	intermedia	2015-003	S2b	R.J. Gornall & J. Muhammed	09/07/2015
8	Greece	Achaia	Chelmos	37.974	22.207	intermedia	13018	S50	O. Polunin & A.O. Chater	17/06/1975
9	Greece	Pella	Litochoro	40.1	22.5	thessala	8048	S17	A. Strid J. S. Andersen	02/05/1974
10	Greece	Phocis	Delphi	38.483	22.5	deltoidea	9852	S63	O. Polunin	1970
11	Greece	Phocis	Parnassus	38.535	22.624	intermedia	11222	S51	O. Polunin	02/08/1970
12	Greece	Attica	Attica	38.081	23.899	deltoidea	E00634033	S62	T.H.H. von Heldreich	22/04/1884
13	Greece	Kavala	Pangaeon	40.916	24.082	intermedia	38500	E15	K.H. Rechinger	15/07/1970
14	Greece	Euboea	Kymi	38.641	24.092	scyria	2015-17	S10e	R.J. Gornall & J. Muhammed	15/07/2015
15	Greece	Chalkidiki	Athos	40.157	24.326	erubescens		BG8	Ex hort. Copenhagen	11/07/2015
16	Greece	Crete	Omolo Plain	35.21	24.91	deltoidea	10046	S60	O. Polunin	1970
17	Greece	Lesvos	Lesbos	39.1	26.2	deltoidea	E.2236	E11	J.R. Edmondson & M.A.S. McClintock	28/04/1978
18	Greece	Samos	Mt Kerkis	37.75	26.833	deltoidea	1649k	E17	P.H. Davis	02/04/1940
19	Turkey	Aydin	Baba Dag	37.728	27.937	deltoidea	41624	S107	P.H. Davis	22/04/1965
20	Turkey	Bursa	Uludağ Bursa	40.15	29.02	olympica	4081	S105	J. Bornmüller	02/06/1899
21	Turkey	Antalya	Korkuteli	37.181	30.933	cilicica	780	S102	A. Manissadjian	s.d.
22	Turkey	Eskişehir	Sündiken	39.658	31.036	canescens	410	S100	T. Ekim	26/03/1971
23	Turkey	Konya	Sultan Dağları	38.333	31.333	pinardii	4082	E29	J.F.N. Bornmüller	14/06/1899

Table 2.1 cont'd.

Pop No	*Country*	*Province / Prefecture*	*Locality*	*Latitude °N*	*Longitude °E*	*Collected as taxon*	*Herbarium voucher*	*Accession code*	*Collectors*	*Collection dates*
24	Turkey	Isparta	Isauria (Bozkır)	37.187	32.24	canescens	15942	E47	P.H. Davis	01/08/1949
25	Turkey	Nevşehir	Cappadocia	38.658	34.853	pinardii	252	S106	W. Siehe	1912
26	Turkey	Samsun	Kizilirmak	41.734	35.956	olympica	1803	E33	C. Tobey	20/05/1967
27	Turkey	Kahramanmaraş	Marash	37.585	36.926	macrostyla	20053	S101	P.H. Davis, J.G. Dodds & R. Çetik	14/07/1952
28	Turkey	Nigde	Nigde	37.915	34.693	vulcanica		S49		
29	Turkey	Mardin	Cudi dag	37.363	40.908	parviflora	42773	S103	P.H. Davis	10/05/1966
30	Syria	Homs	Al Quaryatayn	34.23	37.23	libanotica	0.5614	E21	P.H. Davis	13/03/1943
31	Iraq	Dahuk	Amadiya, Hiror	37.253	43.231	parviflora		S77	J. Muhammed	01/07/2015
32	Iran	Kurdistan	Sanandaj	35.16	47.08	parviflora	49	S104	ETU	17/05/2005
33	Lebanon	Béqaa	Hasharimara (W. of TalatMousa)	34.008	36.145	libanotica	6543A	E25	P.H. Davis	23/06/1963
34	Ex hort.	n/a	n/a	n/a	n/a	gracilis		BG1	Neuchatel BG	
35	Ex hort.	n/a	n/a	n/a	n/a	scardica		BG2	Leicester BG	
36	Ex hort.	n/a	n/a	n/a	n/a	hesperidiflora		BG3	Leicester BG	
37	Ex hort.	n/a	n/a	n/a	n/a	columnae		BG4	Udine BG	
38	Ex hort.	n/a	n/a	n/a	n/a	macedonica		BG5	Vienna (Belvedere) BG	
39	Ex hort.	n/a	n/a	n/a	n/a	intermedia		BG6	Leicester BG	
40	Ex hort.	n/a	n/a	n/a	n/a	intermedia		BG7	Leicester BG	
41	Ex hort.	n/a	n/a	n/a	n/a	intermedia		BG16	Berlin-Dahlem BG	
42	Ex hort.	n/a	n/a	n/a	n/a	columnae		BG9	Vienna (Belvedere) BG	
43	Ex hort.	n/a	n/a	n/a	n/a	anamasica		BG12	Trieste BG	
44	Ex hort.	n/a	n/a	n/a	n/a	croatica		BG13	Leicester BG	
45	Ex hort.	n/a	n/a	n/a	n/a	deltoidea		BG14	Berlin-Dahlem BG	

2.2 DNA extraction

Fresh leaves of *Aubrieta* plants were used to extract genomic DNA following the CTAB method of Doyle and Doyle (1990) with slight modifications. 1 g of young leaves were put in small plastic bags with silica gel (1:20) for 24 hours at 60°C. A pinch of polyvinyl-polypyrrolidone (PVPP) was added to the dried leaf material and put in a1.5 ml microcentrifuge tube (eppendorff) and shaken for 30—60 s by (Silamat Plus) shaker until the samples became a powder.1ml of pre-heated CTAB/ β-merceptoethanol buffer (1:200) at 65°C for 30 minutes added to each samples and incubated 30 minutes at same degree. 500µl of chloroform/isoamyl alcohol (IAA) (24:1) was added to each tube subsequently and gently shacked by orbital shaker for 10-20 minutes. Three different layers were appeared after centrifuging the samples at 13000 rpm for 10 minutes. The supernatant layers which contain DNA collected from the samples and mixed with chloroform: iso-amyl-alcohol (IAA) 24:1 again. In the clean 1.5ml micro-centrifuge tube, supernatant solutions were precipitated with 600 ml of iso-propanol overnight in the freezer. Small pellets of DNA were spin down in the tubes after centrifuging at 13000 rpm for 10 minutes, subsequently 500µl of wash buffer was added to each samples directly. Upper layer solution discarded without losing the pellets of DNA in the bottom of tubes. The tubes inverted for 30 minutes to dry out DNA pellets from the washing buffer. 50µl of TE buffer added to sample in order to dissolve DNA completely. The concentration and quality of DNA was tested by agarose gel electrophoresis as well as a Nano-drop spectrophotometer. A typical 260/680 ratio of a highly purified sample would be 1.8 – 2.0 and a 230/260 ratio ranges from 2.0 – 2.2. When the first ratio is lower than 1.8, it suggests that there is a phenol residue or other reagent used in the extraction. When the latter ratio is lower than 2.0, it suggests that there is carbohydrate carryover, while when the ratio is above 2.2 it suggests that there is a technical problem (Desjardins and Conklin 2010).

2.3 Targeted DNA regions

Phylogenetic and systematic studies of angiosperms have largely focused on the chloroplast genome (e.g. rbcL, matK, ndhF, trnL-F) and nuclear-derived rDNA (18S, 26S, ITS, ETS). This is a very limited sample compared with the thousands of genes available. Notable because of their infrequent use are low- or single-copy nuclear genes. The inclusion of such regions in a dataset alongside chloroplast and ribosomal DNA markers should improve the resolution and statistical support of phylogenies by increasing the proportion of informative characters (i.e. increasing phylogenetic signal) (Sang 2002).

In this study I have attempted to include a range not only of chloroplast DNA regions, but also of single-copy nuclear regions, as well as ITS; mitochondrial RFLPs were also studied.

2.3.1 Chloroplast regions

Based on the survey by Shaw *et al*. (2005), chloroplast regions were selected for sequencing because they were either barcoding sequences or comprised other sequences that evolve rapidly and are usually informative at the species level (Taberlet *et al.* 1991; Demesure *et al.* 1995; Sang *et al.* 1997; Wakasugi *et al.* 1998; Hamilton 1999; Friesen *et al.* 2000; Ohsako and Ohnishi 2000, 2001; Azuma *et al.* 2001; Hamilton *et al.* 2003). The following five chloroplast gene regions were selected, amplified (Fig. 2.1) and sequenced successfully: matK, trnD-trnT, ycf6-psbM, rps11-rpl36 and trnH-psbA. Their structure is shown in Fig. 2.2.

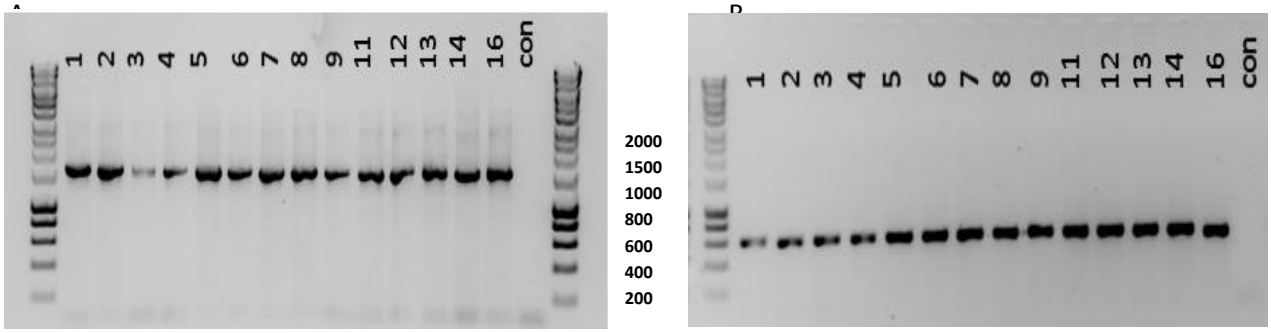


Figure 2.1 Amplified chloroplast DNA fragments: A) trnD-trnT; B) rps11-rpl36 gene regions. The codes on the top of pictures A&B refer to the *Aubrieta* accession numbers (see Table 2.1).

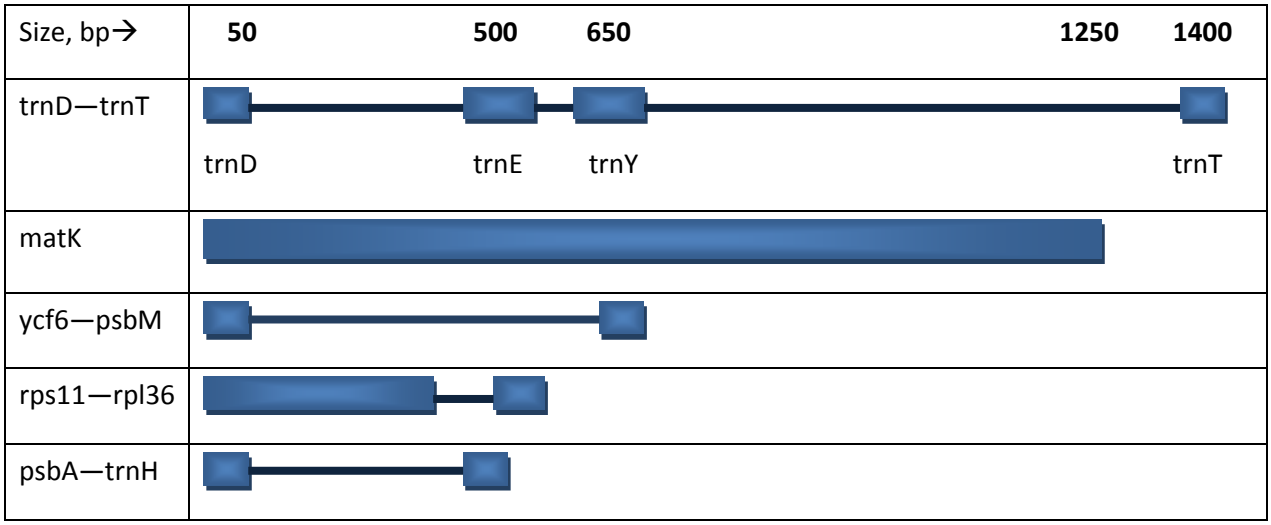


Figure 2.2 Structure and sizes of chloroplast fragments successfully used in the study. Solid blue = coding region; black line = non-coding spacer region.

The phylogenetic utility of these plastid regions has been demonstrated by a number of studies. Thus, matK is one of the barcode genes and is widely used to reconstruct the evolutionary relationships among plants. Its chief advantage at generic level is its relatively high mutation rate. Examples of its successful use include studies of the Saxifragaceae (Johnson and Soltis 1994), Polemoniaceae (Steele and Vilgalys 1994) and Orchidaceae Freudenstein and Chase (2015).

Demesure *et al.* (1995) used trnD-trnT and ycf6-psbM as universal primers to amplify non coding chloroplast regions, which were subsequently shown to be phylogenetically informative at a range of taxonomic levels (Friesen *et al.* 2000; *Lu et al.* 2001; Hahn 2002; Potter *et al.* 2002). Yang *et al.* (2002) used this gene in Brassicaceae to reconstruct phylogenetic relationships between *Brassica*, *Raphanus* and their relatives. The ycf6-psbM region was used by Demsure *et al.* (1996) to study phylogeography of *Fagus* using PCR-RFLP techniques.

Shimamoto *et al*. (1992) conducted research to detect polymorphisms in the rps11-rpl36 region in *Glycine soja* using RFLP depends. Phylogenetic relationships between soybean and related wild species were investigated using the same region by Xu *et al.* (2000) and Sakai *et al.* (2003).

Since 1990, trnH-psbA was used as a remarkable marker for plant phylogeny. Aldrich *et al.* (1988) noted that indels (insertions and deletions) are highly abundant in the trnH-psbA intergenic spacer region, even between closely related taxa. Furthermore, trnH-psbA is tremendously variable compare to other regions (Sang *et al.*1997). Successful application of this region to phylogenetic questions has been demonstrated by Azuma *et al*. (1999) in *Magnolia* (Magnoliaceae), Cyperaceae (Tate and Simpson 2003), Poaceae tribe Triticeae (Bieniek and Mizianty 2014) and commercial woody angiosperms, e.g. *Quercus* (Bolson *et al*. 2015).

2.3.2 Mitochondrial regions

The use of PCR-generated Restriction Fragment Length Polymorphisms (RFLPs) is a powerful tool to detect genetic variation among species. Particularly studied are chloroplast and mitochondrial DNAs after treating with an appropriate restriction enzyme (Tsumura *et al.* 1995; Dumolin *et al*. 1995. The mitochondrial genome sequences in plants mutate about four times slower on average than in the chloroplast genome (Wilfe *et al*. 1987; Palmer 1992). Demesure *et al*. (1995) and Dumolin *et al*. (1997) designed universal primers to amplify mitochondrial-coding and non-coding gene regions, which allowed researchers to apply the

PCR-RFLP technique more easily. As an example, Dumolin *et al*. (1997) reported on the polymorphisms detected by this method in oak (*Quercus*) mitochondrial DNA. A similar approach was taken in a phylogenetic study of Papaya (*Carica papaya*) by Van Droogenbroeck *et al*. (2004).

In my study, I designed primers specific to *Aubrieta* (see chapter 3) in order to generate mitochondrial products by PCR. These were then cut with restriction enzymes (MseI, ecoRI, BstNI, HinfI, mbol, taqI) (Table 2.4).

2.3.3 Nuclear regions

Although coded by the nuclear genome, ribosomal DNA sequences are liable to incomplete concerted evolution, which can lead to divergent paralogous copies within individuals; this makes such sequences potentially unreliable for reconstructing phylogenies especially if they involve ancient hybridisation events, including allopolyploidization (Sang 2002). Owing to the primers used, another concern is that of fungal contamination (Álvarez and Wendel 2003), particularly in cases where plants contain endophytic fungi. In contrast, singe- or low-copy number nuclear genes are able to detect hybridisation, including introgression and allopolyploid events, and are not likely to be confounded by fungal contaminants. Such DNA regions thus potentially provide an important and particular window into phylogenetic reconstructions.

Several candidate single- or low-copy number genes have been used in previous studies of angiosperms, including the Phy genes (phytochomes), ACCase, ADH, AGB1, GAP3DH, GBSSI, GIGANTEA, GPA1, GPAT, LEAFY, ncpGS, petD, PGK, PPR, RBP2 and TPI. Additional genes are sought (Mort and Crawford 2004; Duarte *et al.* 2010). In *Aubrieta,* ADH, and Chs (*chalcone synthase*) have been used (Koch *et al.* 2012).

Following a review of the literature, I chose the following six nuclear genes or regions as potential sources of phylogenetic information on the genus *Aubrieta*. Only the first three listed could be reliably and consistently amplified and sequenced.

1. ITS. This is an rDNA region comprising the 5.8S coding region and the spacer region to either side (ITS1 and ITS2). It is one of the now standard bar-coding supplementary regions and is known often to show greater discrimination than chloroplast regions at low taxonomic levels (Hollingsworth *et al*. 2011). In *Aubrieta,* the entire length amplified to produce one band, which could be sequenced either directly or after cloning.

2. Duo1. This gene contributes to the control of male gamete formation across a wide range of land plants and, as such, it appears to be evolutionarily conserved and defines a new subfamily of pollen-specific MYB genes (Borg *et al.* 2011). This gene has never been screened for phylogenetic use before. In *Aubrieta*, it amplified with one band (Fig. 2.3D,E), and was sequenced in all taxa after cloning.

3. RNA polymerase II subunit 2 (rpb2). In eukaryotes, nuclear RNA polymerases consist of three classes: I, II and III. Each of them has two large subunits (> 100 kDa) and many small ones. All subunits are encoded by a single gene (Oxelman and Bremer 2000). In RNA polymerase II the two large subunits are 205 and 140 kDa in size (Guilfoyle 1983). These two subunits are considered to be useful for evolutionary studies at a range of taxonomic levels because of their ubiquity and the fact that they comprise a number of exons and introns; the latter property bestows both a high level of conservation (in exons) and a high level of variability (in introns) (Allison *et al*. 1985; Iwabe *et al*. 1991). The second subunit of RNA polymerase II (known as rpb2) catalyses messenger RNA synthesis (Allison *et al*. 1985; Iwabe *et al*. 1991) and is also responsible for protein encoded gene transcription (Kolodzieje*t al*. 1990; Sawadogo and Sentenac 1990).

The whole gene sequence of rpb2 from *Arabidopsis thaliana* is well known: it consists of 25 exons (total length ca 3564 bp) and 24 introns (each of which varies from 75 to 776 bp) (Larkin and Guilfoyle 1993). The regions between exons 11-23 are determined to be plant-specific (Denton *et al*. 1998).

Rpb2 behaves either as a single gene, as in *Arabidopsis* (Larkin and Guilfoyle 1993), tomato (Warrilow and Symons 1996), *Rhododendron* and *Hordeum vulgare* (Denton *et al*. 1998), and several diploid Triticeae: H genome (*Hordeum*), St genome (*Pseudoroegneria*), P genome (*Agropyron*) and W genome (*Australopyrum*) (Sun *et al*. 2008), or as a duplicated locus with two copies, as in *Mostuea* (Gelsemiaceae) and *Kopsia* (Apocynaceae) (Oxelman and Bremer 2000).

Few phylogenetic studies have been conducted using the rpb2 gene, but there are some examples of its successful application, e.g. Asterids (Oxelman*et al*. 2004); *Hordeum* (Sun *et al*. 2009); *Elymus* (Sun *et al*. 2008); and *Rhododendron* (Ericaceae) (Goetsch *et al*. 2005).

In *Aubrieta*, the rpb2 region exon16-exon23 (three parts of the plant-specific region, including both exons and introns) could be amplified (Fig. 2.3A, B, C) and sequenced either directly or

after cloning. Preliminary trials with rpb2 (intron 23) amplified with one band (Fig. 2.4B), but the sequences were unreadable, so work on this fragment was discontinued.

4. Cytosolic phosphoglucose isomerase (pgiC). This gene amplified with two bands (Fig. 2.4A). The upper band was cut, purified and sent for sequencing. The results were unreadable. So DNA fragments were cloned and sent again for sequencing. The resulting sequences were checked by blasting in gene bank and were found to match certain bacteria. It was concluded that contamination was present and the gene was not investigated further.

5. Phytochrome C (phyC). This region amplified with many bands, despite experiments involving changes of temperature (Fig. 2.3C). Work on this was therefore discontinued.

6. Leafy (Lfy). This region amplified in some taxa with one band, while others did not amplify at all (Fig. 2.3D). Amplified products produced either good sequences or none at all. Since even the good sequences proved to contain very few SNPs, work on this region was also abandoned.

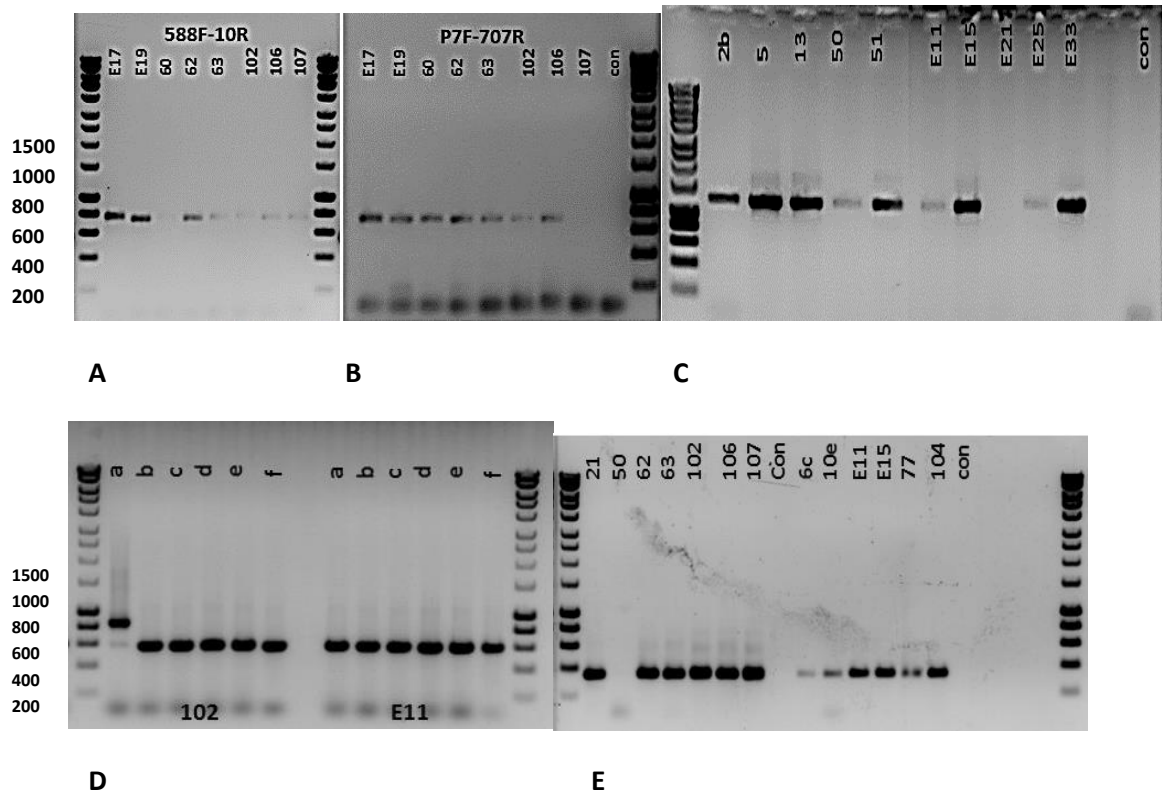


Figure 2.3 Amplified nuclear DNA fragments: A) rpb2 (P7F-10R); B) rpb2 (P7F-707R); C) rpb2 (588F-10R); D) duo1 PCR clones; E) duo1 (duo1-512F- duo1-873R). The codes on the top of pictures A-D refer to the *Aubrieta* accession numbers (see Table 2.1).

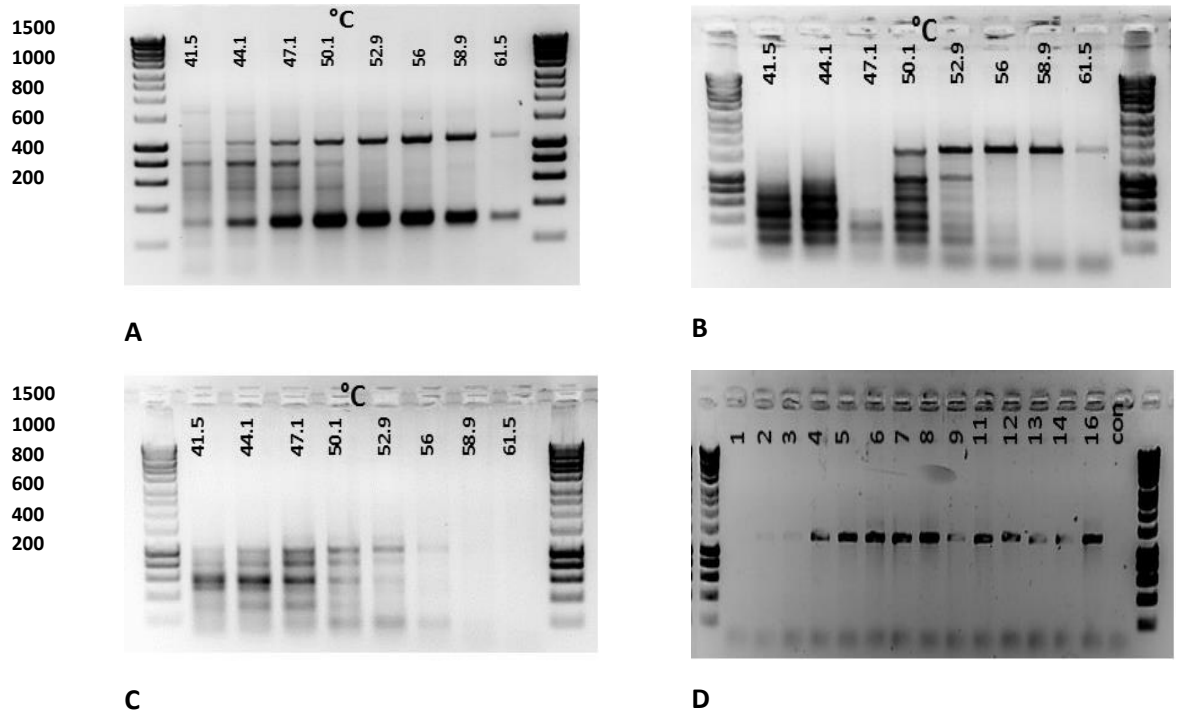


Figure 2.4 Amplified nuclear DNA fragments: A) pgiC; B) rpb2 (intron 23); C) phyC; D) lfy genes. The codes on the top of picture D refer to *Aubrieta* accession numbers (see Table 2.1).

2.4 Polymerase chain reaction (PCR) amplification

Total genomic DNA was diluted by double distilled and deionized water (SIGMA) to 100ng/µl in order to use it in the PCR reaction. In a clean1.5ml micro centrifuge tube, 2µl of 10x KAPA Taq buffer A (1.5 mM Mg at 1X) was mixed with 1.3µl 25mM $MgCl_2$, 1.2µl dNTP mix, 1µl forward and reverse primers (Tables 2.2, 2.3 & 2.4) and 0.1µl KAPA Taq DNA polymerase (5U/ µl), and suspended in water to a volume of 19µl. 1µl of DNA (100ng/µl) was added to a clean PCR tube and 19µl of master mix added to each tube. A short centrifuge spin was required to be sure that the solutions were properly mixed.

For the amplification of chloroplast regions, the PCR program was as follows: 95°C for 2 minutes followed by 30 cycles at 95°C for 1 minute for denaturation. The annealing temperature ranged from 50° to 60°C depending on the fragment being amplified, duration 1 minute with an extension period of 2 minutes, then 10 minutes at 72°C. Finally, the reaction was blocked at 16°C. 2µl of PCR product was pipetted into the agarose gel (1%) with 2µl of 1x loading buffer. 5µl Hyper Ladder I (BIOLINE) was used to identify DNA quality and concentration (Fig. 3.03). Subsequent purification was effected by loading 18µl of the product onto Nucleo-Spin (MACHEREY-NAGEL) or Cycle Pure Kit clean-up kit columns OMEGA bio-tek).

For amplification of nuclear regions, the PCR program was as follows: 93°C for 2 minutes followed by 35 cycles at 93°C for 1 minute for denaturation. The annealing temperature was a constant 60°C, duration 1 minute with an extension period of 1 minute, then 10 minutes at 72°C. Finally, the reaction was blocked at 16°C. 2µl of PCR product was pipetted into the agarose gel (1%) with 2µl of 1x loading buffer. 5µl Hyper Ladder I (BIOLINE) was used to identify DNA quality and concentration. Subsequent purification was effected by loading 18µl of the product onto Nucleo-Spin (MACHEREY-NAGEL) or Cycle Pure Kit clean-up kit columns (OMEGA bio-tek).

Table 2.2 Primers used to PCR-amplify chloroplast DNA regions.

DNA region	*Primer name*	*Primer sequence*	*Reference*
Chloroplast			
matK	3F_KIM 1R_KIM	5'-CGTACAGTACTTTTGTGTTTACGAG-3' 5'-ACCCAGTCCATCTGGAAATCTTGGTTC-3'	CBOL (2009)
matK	2F-matK 2R-matK	5'-AAAGGAGAAGACCGCTTGCG-3' 5'-TTAGAGTTCGATGGGGCTCG-3'	This study
trnD-trnT*	trnD-F trnT-R	5'-ACCAATTGAACTACAATCCC-3' 5'-CTACCACTGAGTTAAAAGGG-3'	Shaw *et al.* (2005)
*trnD-trnE	trnD-F trnE-R	5'-CTACCACTGAGTTAAAAGGG-3' 5'-AGGACATCTCTCTTTCAAGGAG-3'	Shaw *et al.* (2005)
*trnY-trnT	trnY-F trnT-R	5'-CCGAGCTGGATTTGAACCA-3' 5'-CTACCACTGAGTTAAAAGGG-3'	Shaw *et al.* (2005)
trnH-psbA	trnH-F psbA-R	5'-ACTGCCTTGATCCACTTGGC-3' 5'-CGAAGCTCCATCTACAAATGG-3'	CBOL (2009)
rps11-rpl36	rps11-F rpl36-R	5'-TAGGGGTTACATCTCGTACG-3' 5'-AAAATAAGGGCTTCTGTTCG-3'	This study
ycf6-psbM	ycf6-F psbM-R	5'-ATGGATATAGTAAGTCTYGCTTGGGC-3' 5'-ATGGAAGTAAATATTCTYGCATTTATTGCT-3'	Shaw *et al.* (2005)

Table 2.3 Primers used to PCR-amplify nuclear DNA regions.

DNA region	*Primer name*	*Primer sequence*	*Reference*
Nuclear			
ITS	ITS5 ITS4	5'-GGAAGGAGAAGTCGTAACAAGG-3' 5'-TCCTCCGCTTATTGATATGC-3'	CBOL (2009)
duo1 (exon 1-2)	DUO1-1F DUO1-1R DUO1-512F DUO1-873R	5'-TGGAGCTBATTCGRTCCAAAGG-3' 5'-AGRACYGAYAAYGAYGTRAAGAA-3' 5'ATCCTAGCGAGTCTCTTTTGTC-3' 5'TGTCAATCGATACGGTCCTCG-3'	D. Twell (unpub.) This study
rpb2 (exon 15-23)	P7F(PRNTYQSA) P10R(QLIECIMG)	5'-CCYCGTAATACWTAYCARTCWGC-3' 5'-CCCATRATACACTCAATGAGYTG-3'	Denton *et al.* (1998)
rpb2 (exon 16-23)	20F 707R 588F 1249R	5'-CCATGGGGAAGCAAGCAATG-3' 5'-CAACCATGGCGCATACCCTG-3' 5'-CGCCCAGACAGAGGAAGTAC-3' 5'-CCGGAGTGACCCCTTCAATC-3'	This study
rpb2 intron 23	RPB2 INT 23F RPB2-INT-23R	CAACTTATTGAGTGCATCATGG CCACGCATCTGATATCCAC	Roncal*et al.* (2005)
Lfy intron2	LFsx1-2F LFtxR	CACCCACGACCITTYTIGTIACIGARCCIGGIGA CCTGCCIACRTARTGICKCATYTTIGGYTT	Frohlich and Meyerowit (1997)
PgiC exons 11-16	PgiC AA11F PgiC AA16R	TTYGCNTTYTGGGAYTGGGT CCYTTNCCRTTRCTYTCCAT	Ford *et al.* (2006)
Phyc exon 1	Phyc 503F Phyc 1705R	TCVGGGAAGCCSTTYTAYGC GRATWGCATCCATYTCAACATC	Russell *et al.* (2010)

*There is a poly-A/T in the middle part of the trnD-trnDT region which prevents single-step amplification in some taxa. To overcome this in the taxa affected, the region was sequenced in two parts, trnD—trnE and trnY—trnT.

Table 2.4 RFLP analysis of six mitochondrial DNA regions.

Region	*Primer name*	*Primer sequence*	*Product length*	*Restriction enzymes*	*Restriction Sites*
Orf114	7,539 F 8,145 R	ATTTAGCTTTCGACCACGGG TCCATTGCACCCTCATGTGG	607	MseI	130
Nad9-ccmFN2	24,445 F 24,869 R	CACCTTCGCCTGAGACGG AGTACTTACCTGAGAGTGAGAGAGG	425	EcoRI	257
Orf25	34,516 F 35,210 R	TCTTCTTTATTCTTAAGAGAAACCCCC TCAGGGCCTTTCTCGCTGG	693	BstNI	396
matR	44,603 F 45,162 R	TGAGAGGGGCCCTTACTCC TAACTAGTATCGCCTCCCCGC	560	HinfI	133
ccm FC	70,327 F 71,286 F	CGCACCATTACTGCTGGGG GTTCTTCTGTCGTGCGACCC	960	mboI	220
trnK-rps3	137,567 F 137,973 R	AAAAAGAACGCCGCCGCC ATCATTTGTTTGACAGCTTAAACTAGG	407	taqI	242

2.5 Sequencing

Purified PCR and plasmid DNA fragments were sent to two different companies for Sanger sequencing: 1) *Source Bioscience* company. 20µl (15-25 ng/ µl) of DNA or 20µl (60-80 ng/ µl) of plasmid DNA with 20 µl (1µM) of each primer were put into 1.5 ml Eppendorf tube separately.2) *GATC* company. 5µl (15-25 ng/ µl) of DNA or 5µl (60-80 ng/ µl) of plasmid DNA were premixed with 5 µl (5µM) of each primers into 1.5 ml Eppendorf tube. The gene regions ≤ 1000 bp were sequenced in only one direction whereas regions > 1000 bp were sequenced in both directions.

2.6 Cloning

Cloning was only needed in the case of heterozygous nuclear DNA regions. Purified PCR fragments were cloned in pGEM®-T Easy vectors, using pGEM®-T Easy Vector System I kit (Promega) following manufacturer's protocol with little modification. pGEM®-T Easy vector has a single overhanging 3′ deoxythymidine (T) nucleotide that can be ligated to a single base deoxyadenosine (A) to the 3′ end of the PCR products generated by Taq polymerase.

Ligation reactions of 10µl were set up in a small 300µl tube, comprised of 5µl of 2x Rapid Ligation Buffer [60mM Tris-HCL pH 7.8, 20mM $MgCl_2$, 20mM DTT, 2mM ATP, 10% PEG (Promega)], 0.9µl of the pGEM-Teasy vector, 1.2µl of T4 DNA Ligase and 5.4µl of purified PCR product were mixed and incubated at room temperature (RT) for 1 hr, then at 4ºC overnight. The insert: vector ratio was calculated below following guidelines provided by Promega (http://www.promega.com/). For transformation, 5µl of the ligation reaction was added to 50µl of the competent *E. coli* cells (α -Select Bronze Efficiency, Bioline), kept on ice for 20 min, then heated for 45 second at 42ºC, followed by 10 min on ice again. Pre warmed 750µl of Super Optimal Broth media SOB added to each reaction tube on ice and then incubated at 37ºC for 3 hr in orbital shaker at 230 rpm to allow transformed competent cells to grow properly. After this incubation, 100µl & 200 µl of this media added to LB agar plates, containing 100µg/ml ampicillin, 40µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-gal) and 500µM isopropyl-β-Δ-thiogalactopyranoside (IPTG). These plates were incubated at 37ºC overnight. To calculate the weight of the target DNA required, the following equation was used:

$$\text{ng of target DNA (insert)} = \frac{\text{ng of vector} \times \text{kb size of insert} \times \text{insert: vector molar ratio}}{\text{kb size of vector}}$$

2.7 Mitochondrial RFLPs

Next generation sequencing (NGS) data were assembled for four *Aubrieta* accessions (Chapter 3), using *Geneious* software (Kearse *et al*. 2012). One hundred contigs were generated. Up to ten of them were mitochondrial DNA. These contigs were mapped against the whole annotated *Brassica napus* mitochondrial genome (Genbank accession number AP006444) as a reference. As a result, 57 genes, including introns and spacers were assembled and annotated from the reference. The four *Aubrieta* mitochondrial genomes (57 genes) were aligned using MAFFT Alignment (Katoh *et al* 2002) with the following settings: a) Algorithim: Auto; b) scoring matrix: 2000PAM/k=2; c) gap open penalty: 1.53, offset value 0.123. Enzyme restriction sites were identified on these aligned sequences using *Geneious* software. DNA regions that varied between accessions with respect to restriction sites were selected. Primers were designed to cover these sites.

The PCR amplification conditions and procedures were similar to those applied to chloroplast genes. In small PCR tube, 2 µl of PCR product were diluted with 3.8 µl of deionized water and mixed with 1 µl of either buffer 2 (50mM NaCl, 10mM Tris-HCl, 10mM $MgCl_2$, 1mM DTT) or CutSmart Buffer (50mM Potassium Acetate, 20mM Tris-acetate, 10mM Magnesium Acetate, 100µg/ml BSA), and 0.2 µl of restriction enzyme. The mixture was either incubated at 37 or 65C for 45 minutes based on enzyme incubation temperature. The whole amount of product, together with 2µl of 5x loading buffer (BIOLINE), was pipetted into an agarose gel (2%).

Based on the restriction sites, a network tree was generated using *Network* software (Bandelt and Dress 1992). The presence of a restriction site was coded as 1, while the samples without it were coded as 0. Exceptionally, the trnK-rps3 gene spacer was cut more than once. The codes were giving as follows; a) 0 when there was one restriction site, b) 1 when there were two restriction sites.

CHAPTER 3. WHOLE GENOMES: CHLOROPLAST AND MITOCHONDRION

3.1 Introduction

In plants, species identification, classification and phylogenetic relationships are often problematic, especially when only morphological data is available. These phenotypic characters are often polymorphic (Duminil and Michele 2009). In the last few decades, molecular techniques have been widely applied to resolve these taxonomic and evolutionary questions; the studies have involved a range of molecular markers. For example, Sang *et al*. (1997) conducted a study using a few chloroplast non-coding gene regions to determine phylogenetic relationships among *Paeonia* (Paeoniaceae) species; Choi *et al*. (2006) used RAPD markers to identify the diversity and phylogeny of the genus *Cymbidium* (Orchidaceae); Koch *et al*. (2010) classified North American *Arabis* (Brassicaceae) species using the nuclear internal transcribed spacer region (ITS) and chloroplast (trnL-F) genes; and Wang *et al*. (2011) used ISSR markers to detect genetic variation within *Cymbidium ensifolium.* Although there is no modern review of these approaches, many of them have been documented and explained by Bateman *et al.* (1999).

However, despite the invention of DNA sequencing technologies and their regular application, their cost was still a barrier for studies that ideally would have focussed on whole genomes. Instead the limitations obliged researchers to select a few DNA regions and markers with relatively few informative loci (Choi *et al*. 2006; Sharma *et al.* 2012), leading to the development of the barcoding project (Chase *et al*. 2007; Hollingsworth *et al*. 2009) primarily aimed at aiding species identification.

In recent years, due to DNA sequencing prices dropping and the development of next-generation sequencing technologies, it is now possible to sequence complete genomes (nuclear, mitochondria and chloroplast). This has allowed us to refine considerably our approaches to population genetics and species identification and our study of phylogenetic relationships (Handa 2003; Mardis 2008; Shendure and Ji 2008; Cronn *et al*. 2008; De Pristo *et al*. 2011; Hohmann *et al*. 2015).

3.1.1 Next Generation sequencing (NGS)

In the late seventies and after the Sanger chain termination technology invention, scientists were able to start DNA sequencing more efficiently. Up to date, DNA sequencing techniques have been developed considerably (Sanger 1988; Hutchinson 2007). In the late 1980s, DNA was sequenced by semi-automated methods based on the Sanger biochemistry capillary system (Swerdlow *et al.* 1990; Hunkapiller *et al.* 1991). Two decades later, the Sanger biochemistry output product reached up to 1000 bp per sequencing reaction (Shendure and Ji 2008). Subsequently, cyclic-array technology was developed, also known as Next Generation Sequencing (NGS). Goodwin *et al*. (2016) divided NGS into short-read and long-read approaches as summarised in Table 3.1.

Table 3.1 Classification of NGS platforms according to Goodwin *et al.* (2016).

Platform	Read Length (bp)	Throughput	Reads	Runtime
Short-read approaches				
Sequencing by ligation				
SOLiD	50-75	80-320 Gb	Up to 1.4 B	6 days
Complete Genomics	50-100	8-200Gb	NA	24hrs
Sequencing by synthesis: CRT				
Illumina	25-300	540Mb-900Gb	Up to 3B	Up to 11 days
Qiagen GeneReader	NA	12 genes; 1,250 mutations	NA	Several days
Sequencing by synthesis: SNA				
454 pyrosequencing	Up to 1000	35-700Mb	Up to 1 M	Up to 23hrs
Ion Torrent	200-400	30Mb-15Gb	0.04-80M	2.5-23hrs
Long-read approaches				
Single-molecule real-time reads				
PacBio	8-200kb	500mb-7Gb	55000-350000	0.5-6hrs
Oxford Nanopore	Up to 200kb	Up to 4Tb	NA	48hrs
Synthetic long reads				
Illumina Synthetic Long-Read	100kb	9-500Gb	300m-4B	Up to 6 days
10X Genomics	Up to 100 Kb	9-500Gb	300m-4B	Up to 6 days

Short read NGS

Short-read sequencing approaches correspond to a) sequencing by ligation (SBL) and b) sequencing by synthesis (SBS). With SBL, a probe sequence is bound to a fluorophore and then hybridised to a fragment of DNA. The complex is then ligated to an adjacent oligo-nucleotide for imaging. The base or sequence of bases, complementary to positions within the probe, are identified by the emission spectrum of the fluorophore. With SBS, a polymerase is used, and the addition of a base to an elongating strand is identified by a signal, such as a fluorophore or a change in ionic concentration.

In this study, the *Sequencing by synthesis: CRT (Illumina)* approach (Table 3.1) was used. This procedure uses terminator molecules to prevent elongation by blocking the ribose 3'-OH groups, similar to the Sanger protocol. To start, a DNA template is primed by a sequence complementary to an adapter region. This initiates polymerase-binding to the double-stranded DNA (dsDNA) region. A mixture of the four deoxynucleotides, which are individually labelled and 3'-blocked, are added during each cycle. After each single dNTP is incorporated into the elongating complementary strand, unbound dNTPs are removed and the surface is imaged to identify which dNTP was incorporated at each cluster. The image signal relies on the fact that each dNTP is bound to a single fluorophore specific to its base. After each cycle, the fluoro-phore and blocking group are removed and a new cycle begins. Imaging is conducted by means of total internal reflection fluorescence (TIRF) microscopy, using four imaging channels.

Long-read sequencing

Long-read sequencing delivers reads more than several kilobases long, which can span large structural features and permit their resolution, avoiding ambiguity in their size and position. Long reads are also useful for transcriptomic studies, because they span whole mRNA transcripts. This permits the connectivity of exons to be identified and any isoforms to be established. There are two main kinds of long-read protocols:

a) single-molecule real-time sequencing which, unlike existing short-read technologies, do not rely on a clonal population of amplified DNA fragments to produce a signal, nor do they need a chemical cycling stage for each dNTP added.

b) synthetic protocols, which compile long-reads in silico and are based on existing short-read technologies; such protocols do not produce long-reads themselves, but instead they generate libraries which reference barcodes, thus allowing the assembly of larger fragments.

3.1.2 Chloroplast genome

Chloroplasts are one of the fundamental organelles in the plant cells. Ultimately derived from photosynthetic bacteria by means of endosymbiosis (Cavalier-Smith 2003), these organelles play a functional role by participating in both photosynthesis and other metabolic pathway processes. The distinct features of these organelles are that they are circular, they show *maternal inheritance in most angiosperms*, their genes are single copy, inherited as a linkage group and are highly conserved (Palmer *et al*. 1988; Tian 2002; Dyall *et al*. 2004; Tetlow *et al*. 2009; Wu *et al*. 2010; Liu *et al.* 2013). Umesono *et al*. (1984) were the first to publish the sequence of a complete chloroplast genome, of *Marchantia polymorpha*. To date (September 2016), there are 7,524 complete chloroplast genomes available on the NCBI website (www. ncbi.nlm.nih.gov / genomes / ORGANELLES / organelles.html).

In Angiosperms, chloroplast genomes vary in size from 120 to 160 kb. Each chloroplast genome comprises two inverted repeats (IRs) 20-28 kb in length, located in between the large single copy (LSC) region, 80–90-kb, and the small single copy (SSC) region, 16–27-kb; four ribosomal rRNAs genes; 30 transfer tRNAs; and genes encoding ca 80 specific proteins (Bendich 2004; Chumley *et al*. 2006).

The chloroplast genome is a versatile tool to facilitate answering questions of plant identification and phylogeny. Moore *et al*. (2007) conducted a study to clarify the relationships of basal angiosperms by examining 61 plastid genes. Similarly, the chloroplasts of 37 *Pinus* species were studied by Parks *et al*. (2007) to elucidate their relationships. Moore *et al*. (2010) analysed 83 chloroplast rDNA and protien-coding genes phylogenetically to determine the early diversification of eudicots. Likewise, 73 coding genes were used by Hohmann *et al*. (2015) to determine the evolutionary relationships and divergence time estimation of 29 Brassicaceae species. In terms of using entire genomes, Nock *et al.* (2011) constructed a phylogenetic tree for five grass species.

3.1.3 Mitochondrial genome

Mitochondria are subcellular organelles responsible for producing energy through the oxidative phosphorylation process. The origin, like chloroplasts, is also from an endosymbiotic event involving bacteria (Cavalier-Smith 2003). As results of the interactions between bacteria and the host cells, the organellar genome was reduced due to either the genes lost or translocation between the two interacting organisms.

In angiosperms, the mtDNAs genomes expand or reduce in size due to either the duplication in intergenic spacer sequences or exchanging sequences with chloroplast and nuclear genomes (Palmer 1992; Unseld *et al*. 1997; Marienfeld *et al*. 1999). In plants, mitochondrial genes lose and/or transfer sequences more frequently than in other organisms (Palmer *et al*. 2000).

The mtDNA genome in angiosperms varies in size from 200kb to 20000 kb (Palmer 1990 and 1992). For example, some members of the Cucurbitaceae family have mtDNA genomes six-fold bigger than others in size (Ward *et al*. 1981). Oda *et al*. 1992 found that there were extra genes and expansion of the *Marchantia polymorpha* liverwort genome which resulted from the enlargement of introns and spacers. In *Arabidopsis* (Brassicaceae), the mtDNA genome is about 367 kb in, which is two times bigger than that in the related *Brassica hirta,* the variation being ascribed to mitochondria-nuclear genomes interactions (Unsel *et al*. 2000).

3.1.4 Aims

The present study aims to:

- assemble complete sequences for the chloroplast and mitochondrial genomes in *Aubrieta*. This will further our understanding of genome organization and structure in the Brassicaceae in particular, and aid future studies of phylogeny.
- confirm the phylogenetic position of *Aubrieta* within the Brassicaceae by analysing whole chloroplast genome sequences alongside equivalent data from forty-three other species.

3.2 Materials and methods

3.2.1 DNA isolation from Aubrieta species

Total genomic DNA was extracted from fresh leaves of four *Aubrieta* species; *Aubrieta gracilis* (BG1), *A. scardica* (BG2), *A. erubescens* (BG8) and *A. anamasica* (BG12), following the standard cetyl-trimethylammonium bromide (CTAB) method with some modifications; for more details see Chapter 2. DNA samples were sent to the Interdisciplinary Centre for Biotechnology Research, University of Florida, USA to sequence these samples commercially. Hiseq500 2x150bp reads were used to sequence the all accessions following the manufacturer's procedure.

3.2.2 Data analysis: procedure for assembling the chloroplast genome of Aubrieta gracilis

Geneious 7 software (Kaerse *et al*. 2012) was used to analyse and assemble the total genomic data. Firstly, the bi-directional raw reads of the DNA sequences were assembled in their complementary pairs. Then the de novo assembly was selected to generate 100 contigs. The sizes of the contigs varied from (1000-80,000bp). Each of the contigs was blasted in the NCBI website in order to know the genome type (nuclear, chloroplast, mitochondrial). The chloroplast contigs were mapped against *Arabis alpina* (Genbank accession number NC_023367) as a reference genome, then mapped again to *Aubrieta gracilis* raw reads DNA sequences. In order to get the complete circular, quadripartite organization of the *Aubrieta gracilis* plastome, 4,185,439 reads with sequences 151bp long were assembled from a total of 52,967,592 reads.

The circular chloroplast genome was annotated using Geneious (Kaerse *et al*.2012) by transferring the annotation from the reference genome *A. alpina*. Then this annotation was compared with published complete chloroplast genomes using DOGMA, Dual Organellar Genome Annotator (Wyman *et al*. 2004). After that, this annotation was drawn using the online software, GenomeVX (Conant and Wolf 2008).

Complete Brassicaceae plastid genomes are available for 39 species (Hohmann *et al*. 2015; Seol *et al*. 2015; Novikova *et al*. 2016). The eight phylogenetically closest were selected for comparative analysis (*Arabidopsis thaliana*, ***Aubrieta gracilis***, *Arabis alpina, Arabis hirsuta*, *Brassica napus*, *Capsella bursa-pastoris Draba nemorosa*, *Raphenus sativus*). The plastomes were aligned using MAFFT Alignment (Katoh *et al.* 2002) with the following settings: a) Algorithim: Auto; b) scoring matrix: 2000PAM/k=2; c) gap open penalty: 1.53, offset value 0.123. The above eight genomes were saved as FASTA files, then analysed by the mVISTA online program (Dubchak and Ryaboy 2006) using the Multi-LAGAN (Limited Area Global Alignment of Nucleotides) option (Brudno *et al*. 2003).

In addition all of the 39 complete plastid genomes were subjected to phylogenetic Bayesian analysis to test the position of the genus *Aubrieta*.

3.3 Results

3.3.1 Chloroplast genome assembly

The data analysis shows that there were 4,101 copies of the plastome which covered ca 7.9% of the complete genome. General features of the chloroplast genome are summarised in Table 3.2 and Fig. 3.1. The length of the plastome of *A. gracilis* is 154,103 bp, whereas the average length in Brassicaceae ranges from 152,860 to 155,611 bp. The GC content of *A. gracilis* is 36.5% while the average in Brassicaceae is 36.4 ± 0.1%.

The *A. gracilis* plastome has two inverted repeats with a total length of 26,159 bp, and two unique single-copy regions (Fig. 3.1): small (SSC), 17,596 bp long, and long (LSC), 84188bp long. The plastome comprises 133 genes, whose lengths are summarised in Table 3.3, and which are categorised below:

1) 88 protein-encoding genes, of which eight are duplicated on the inverted repeat region (ndhB, rpl2, rpl23, rps7, rps12, ycf1, ycf2, ycf15).

2) 37 transfer-RNA genes, seven of which are duplicated on the inverted repeat (IR) region (trnA-UGC, trnI-CAU, trnI-GAU,trnL-CAA, trnN-GUU,trnR-ACG, trnV-GAC).

3) Eight ribosomal-RNA genes, of which four are on the IR1 (rrn4.5, rrn5, rrn16, rrn23) and the other four are duplicated on the IR2 (rrn4.5, rrn5, rrn16, rrn23).

Among these 133 genes, 19 include introns which occur in twelve protein-encoding genes and seven transfer-RNA genes (Fig. 2). The clp and ycf3 genes were connected by two introns. Furthermore, the rps12 gene is trans-spliced all around the circular plastome. The 5′ exon localization is in the LSC region whereas the two copies of 3′ exon are localised in the inverted repeat regions.

Table 3.2 General features of the *Aubrieta gracilis* chloroplast genome.

Features	Chloroplast
Genome size (bp)	154,103
GC content (%)	36.5
Number of protein-coding gene	88
Number of transfer-RNA genes	37
Number of ribosomal-RNA genes	8
Number of genes with introns	19

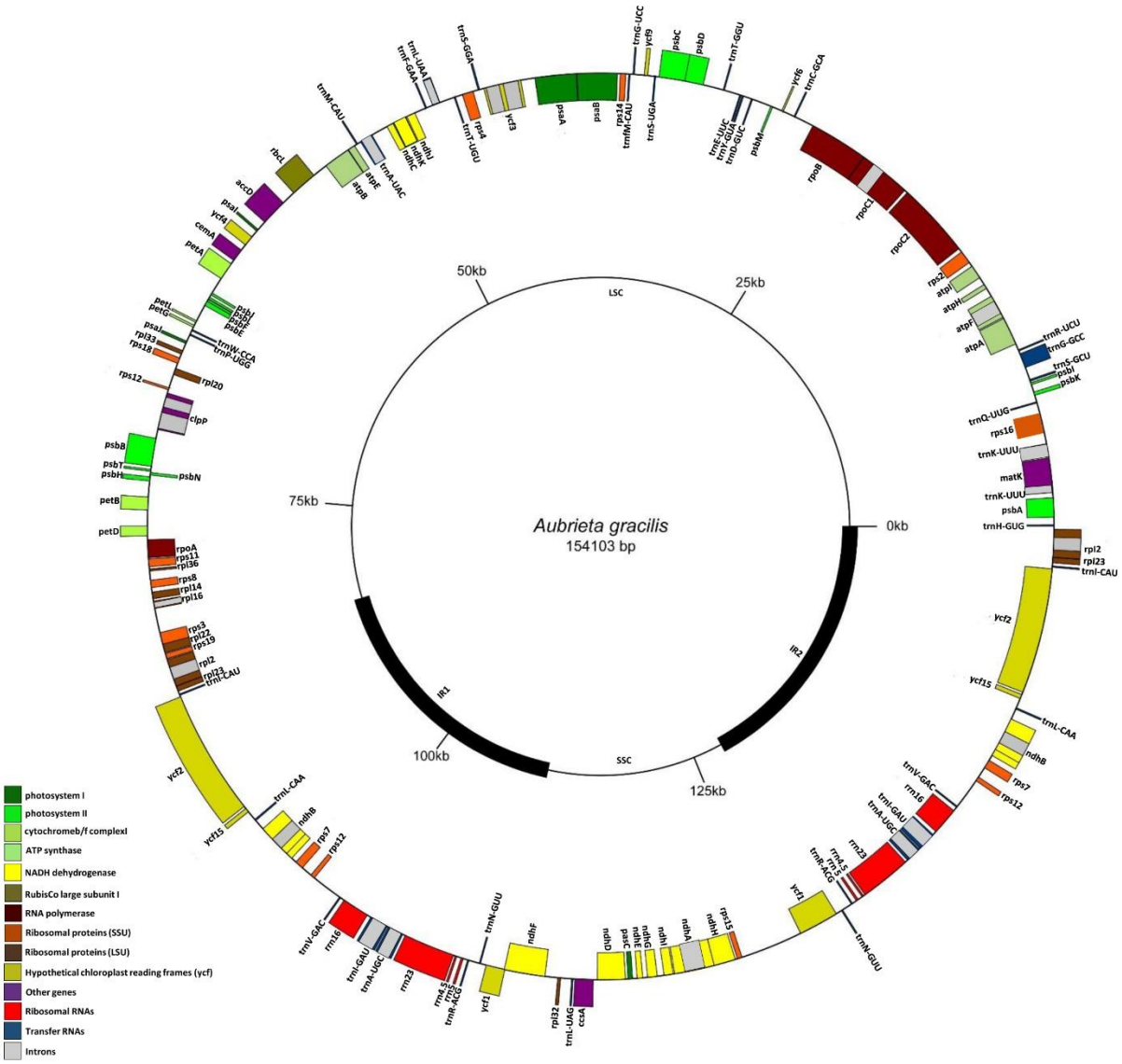


Figure 3.1 Chloroplast gene map of *Aubrieta gracilis*. The genes marked inside the circle are transcribed counter clockwise, while the outer ones are transcribed clockwise. The colour-codes represent the genes based on their functions as per the key.

Table 3.3 Complete list of 133 chloroplast gene exons in *Aubrieta gracilis*, showing their names, start point, end point and length.

name	start	end	length	name	start	end	length	name	start	end	length
trnH	11	83	73	trnM	52398	52470	73	rps7	97498	97965	468
psbA	499	1560	1062	atpE	52604	53002	399	rps12	98571	98813	243
trnK	1839	4536	72	atpB	53023	54495	1473	trnV	100522	100593	72
matK	2197	3726	1530	rbcL	55325	56764	1440	rrn16S	100825	102315	1491
rps16	5228	6351	1124	accD	57546	59001	1456	trnI	102580	103560	77
trnQ	6762	6833	72	psaI	59534	59647	114	trnA-UGC	103625	104497	73
psbK	7198	7383	186	ycf4	60059	60613	555	rrn23S	104650	107459	2810
psbI	7802	7912	111	cemA	60974	61663	690	rrn4.5S	107558	107660	103
trnS	8003	8090	88	petA	61894	62856	963	rrn5S	107908	108028	121
trnG	8786	9576	72	psbJ	63779	63901	123	trnR	108279	108352	74
trnR	9740	9811	72	psbL	64045	64161	117	trnN	108929	109000	72
atpA	10164	11687	1524	psbF	64183	64302	120	ycf1	109324	110355	1032
atpF	11754	13064	555	psbE	64312	64563	252	ndhF	110323	112557	2235
atpH	13562	13807	246	petL	65505	65600	96	rpl32	113257	113415	159
atpI	14247	14996	750	petG	65782	65895	114	trnL	114006	114085	80
rps2	15241	15951	711	trnW	66032	66105	74	ccsA	114194	115180	987
rpoC2	16205	20311	4107	trnP	66293	66366	74	ndhD	115398	116918	1521
rpoC1	20486	23329	2055	psaJ	66733	66867	135	psaC	117062	117307	246
rpoB	23356	26574	3219	rpl33	67288	67488	201	ndhE	117559	117864	306
trnC	27652	27722	71	rps18	67721	68026	306	ndhG	118127	118657	531
ycf6	28358	28447	90	rpl20	68318	68671	354	ndhI	119019	119522	504
psbM	28979	29083	105	rps12	69430	69543	114	ndhA	119598	121790	1083
trnD	30142	30215	74	clpP	69722	71764	591	ndhH	121792	122973	1182
trnY	30660	30743	84	psbB	72262	73788	1527	rps15	123090	123356	267
trnE	30803	30875	73	psbT	73963	74064	102	ycf1	123688	128982	5295
trnT	31746	31817	72	psbN	74128	74259	132	trnN	129306	129377	72
psbD	32931	33992	1062	psbH	74365	74586	222	trnR	129954	130027	74
psbC	33940	35361	1422	petB	74721	76159	1439	rrn5S	130278	130398	121
trnS	35543	35635	93	petD	76375	77556	1182	rrn4.5S	130646	130748	103
ycf9	35978	36166	189	rpoA	77767	78750	984	rrn23S	130847	133656	2810
trnG	36706	36776	71	rps11	78820	79236	417	trnA-UGC	133809	134681	73
trnfM	36928	37001	74	rpl36	79355	79468	114	trnI	134746	135767	72
rps14	37163	37465	303	rps8	79967	80371	405	rrn16S	135991	137481	1491
psaB	37614	39818	2205	rpl14	80618	80986	369	trnV	137713	137784	72
psaA	39844	42096	2253	rpl16	81110	81511	402	rps12	139493	139735	243
ycf3	42873	44915	509	rps3	82850	83506	657	rps7	140341	140808	468
trnS	45150	45236	87	rpl22	83491	83973	483	ndhB	141122	143347	2226
rps4	45520	46125	606	rps19	84030	84308	279	trnL	143936	144016	81
trnT	46526	46598	73	rpl2	84363	85870	828	ycf15	144554	144787	234
trnL	47416	47821	85	rpl23	85889	86170	282	ycf2	144924	151805	6882
trnF	48186	48258	73	trnI	86339	86412	74	trnI	151894	151967	74
ndhJ	48887	49363	477	ycf2	86501	93382	6882	rpl23	152136	152417	282
ndhK	49424	50145	722	ycf15	93519	93752	234	rpl2	152436	153943	1508
ndhC	50194	50556	363	trnL	94290	94370	81				
trnA	51535	52215	84	ndhB	94959	97184	1539				

3.3.2 Comparison with eight related Brassicaceae chloroplast genomes

Comparison between the eight studied species of Brassicaceae demonstrated that the chloroplast genome identity is ca 96.6%, which indicates that the chloroplast genome among these species of Brassicaceae is highly conserved. This conservation is more pronounced in coding regions comparing with non-coding regions, which are highly variable (Fig. 3.2). The alignment of the eight studied plastomes shows that *Aubrieta gracilis* differs primarily in the intergenic spacer regions. For example, there is a huge insertion between the trnT-trnE genes; there is a small deletion in the psbE-petN region; and there is a 53bp deletion between the trnR-trnN genes. In addition, the rps16 protein-coding gene, which exists in each of *Aubrieta gracilis, Arabidopsis thaliana, Arabis hirsute, Brassica napus, Capsella bursa-pastoris* and *Raphenus sativus*, is absent from *Draba nemorosa* and *Arabis alpina* (Fig. 3.2).

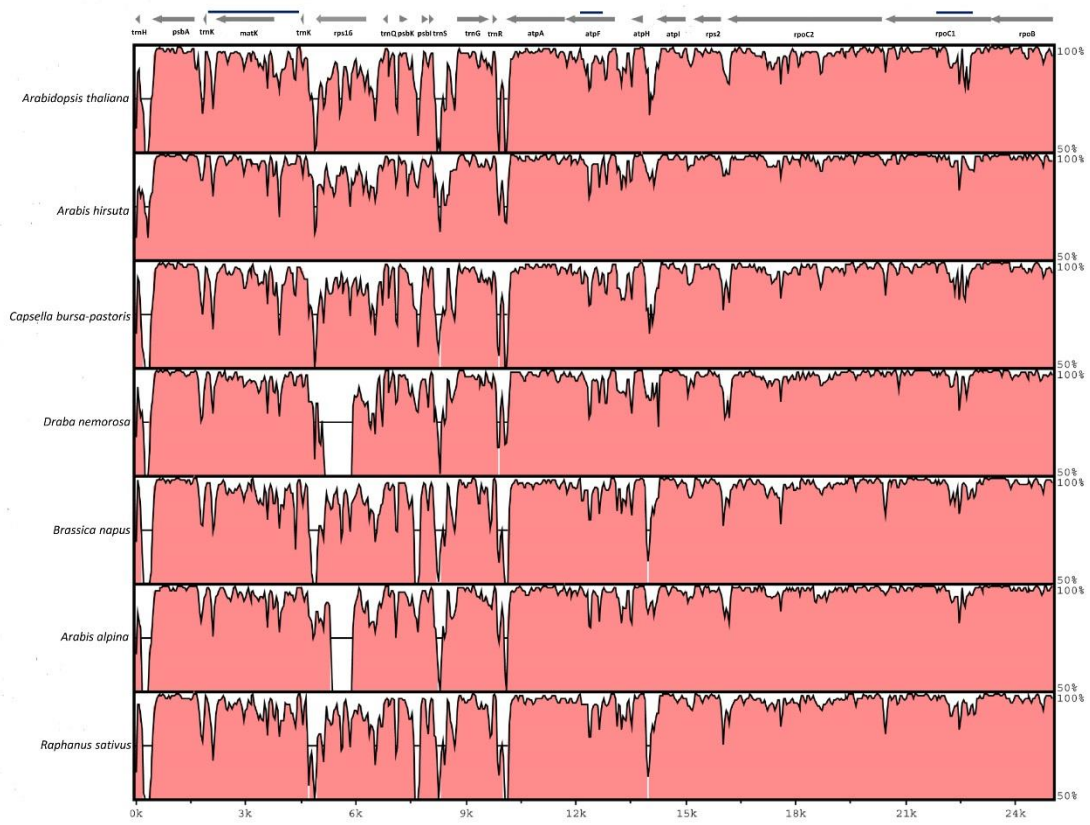


Figure 3.2 Graphic view of eight aligned Brassicaceae cp genomes. The sequences identified using VISTA plot identity between the *Aubrieta gracilis*, *Arabidopsis thaliana* (NC_000932), *Arabis hirsuta* (NC_009268), *Capsella bursa-pastoris* (NC_009270), *Draba nemorosa* (NC_009272), *Brassica napus* (NC_016734), *Arabis alpina* (NC_023367), *Raphanus sativus* (NC_024469) plastomes. The X axis is *A. gracilis* cp genome coordination, while the Y axis is the sequences variety proportion (50% & 100%). The grey arrows refer to the genes and their directions. The dark blue lines refer to introns.

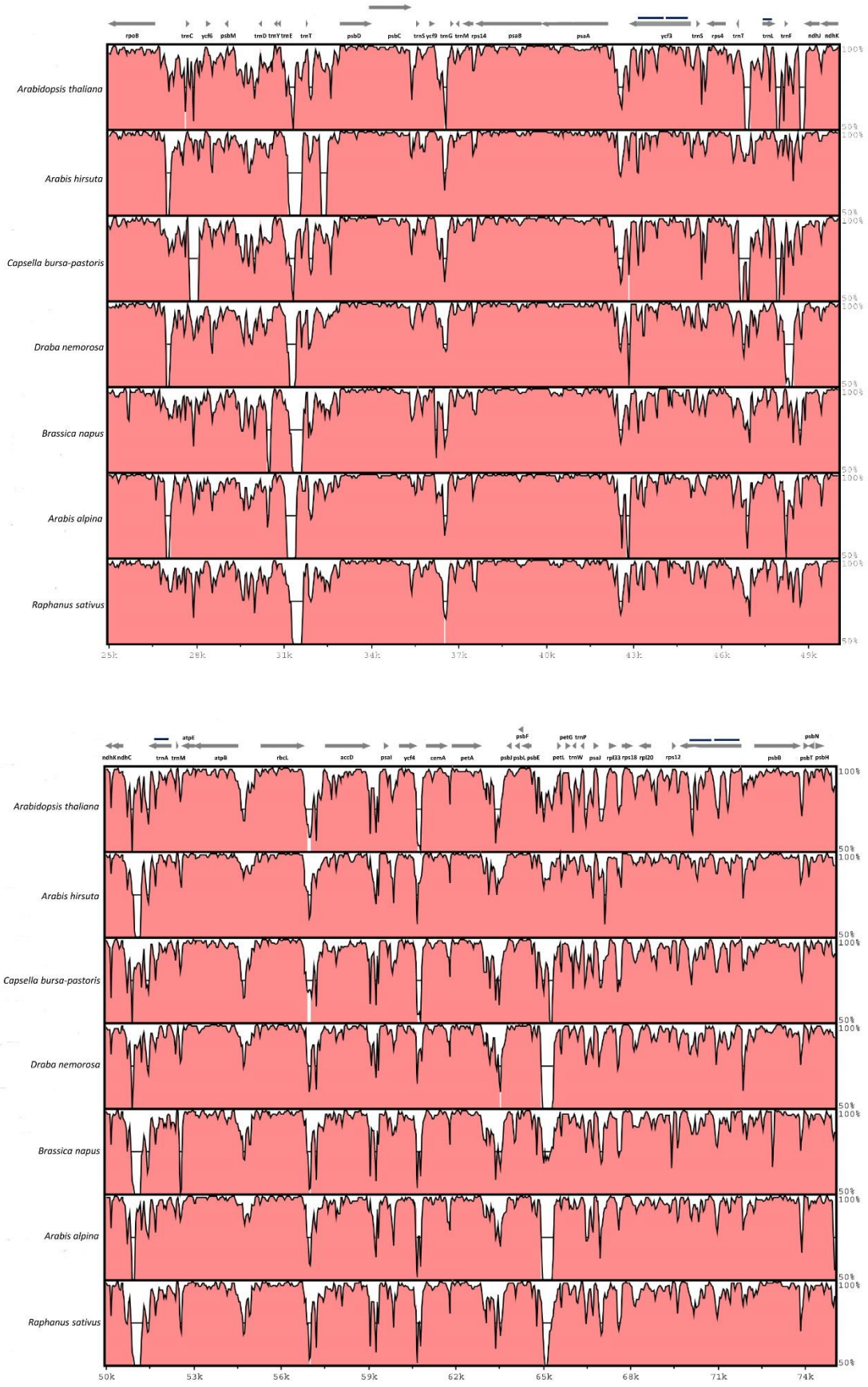


Figure 3.2 continued.

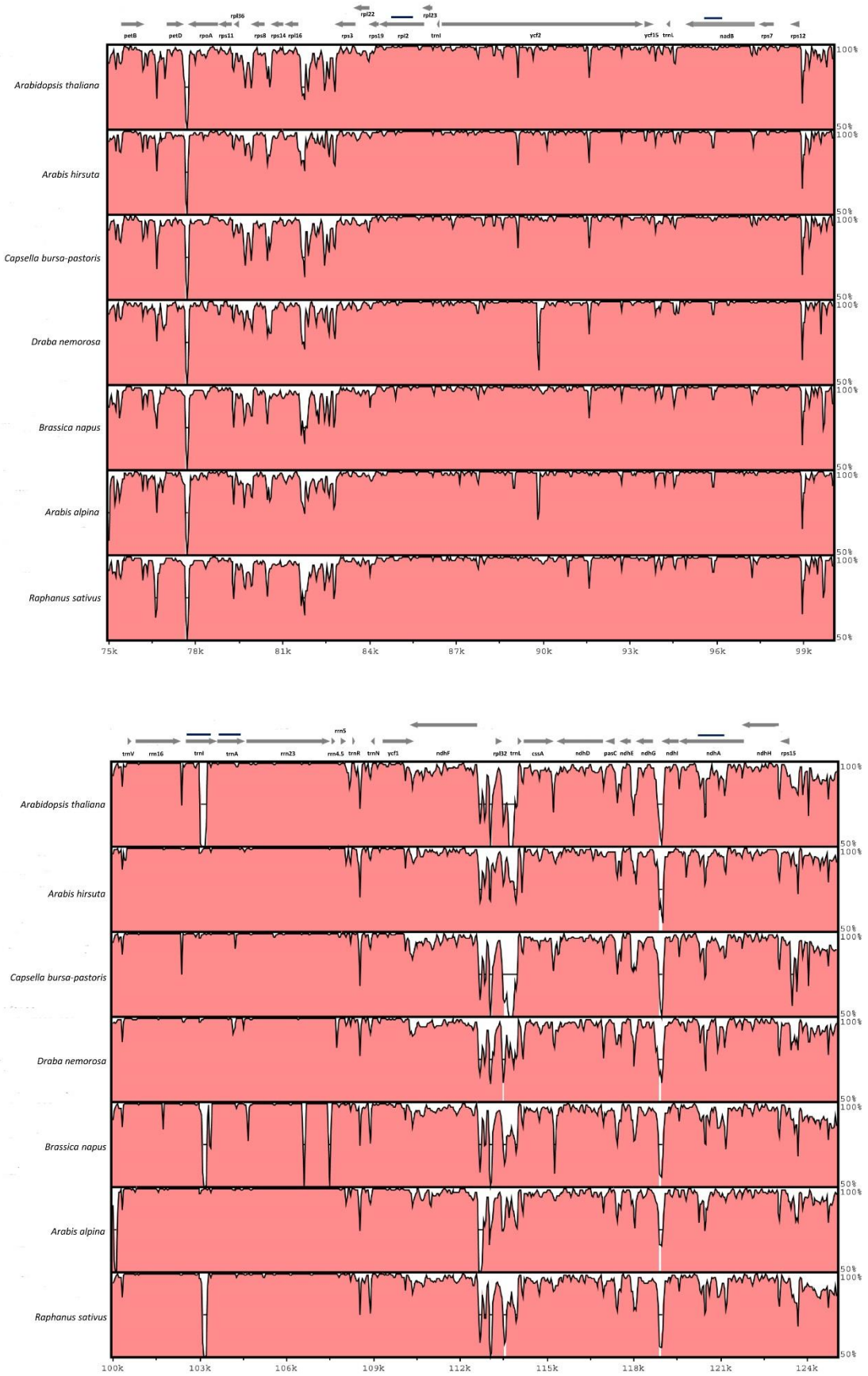


Figure 3.2 continued.

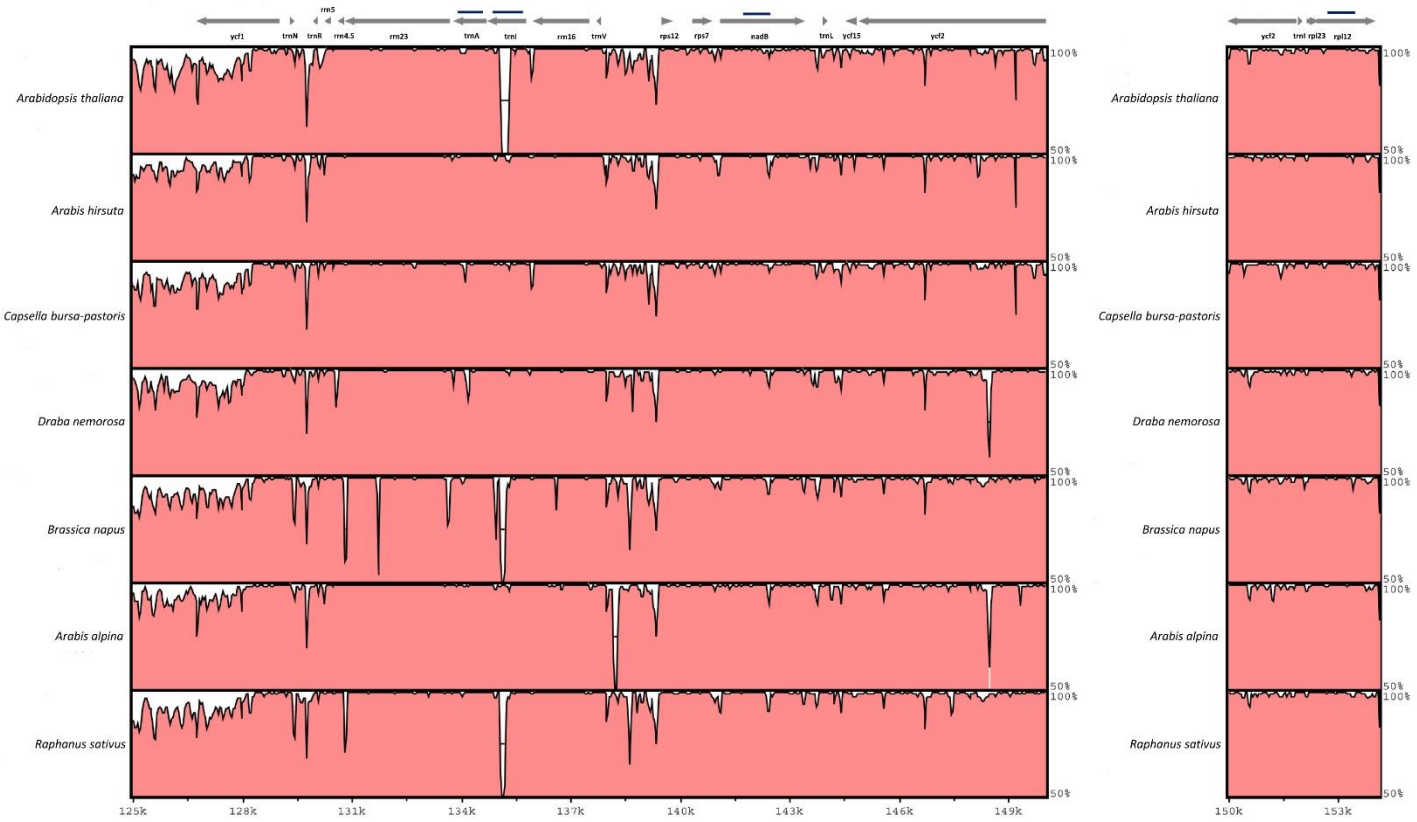


Figure 3.2 continued.

3.3.3 Phylogeny of whole chloroplast genomes in Brassicaceae: position of Aubrieta

Bayesian phylogenetic analysis of 43 whole chloroplast Brassicaceae, including four from *Aubrieta*, genomes produced the tree shown in Fig. 3.3. This tree demonstrates that *Aubrieta* is monophyletic, forming a clade with pp = 1.00. Almost all other clades shown by the tree also have perfect support (pp = 1.00), the only exceptions being within *Arabidopsis* and within tribe Arabideae where the clade containing *Arabis hirsuta* and *Draba nemorosa* is shown as sister to *Aubrieta* but with a posterior probability of 0.83.

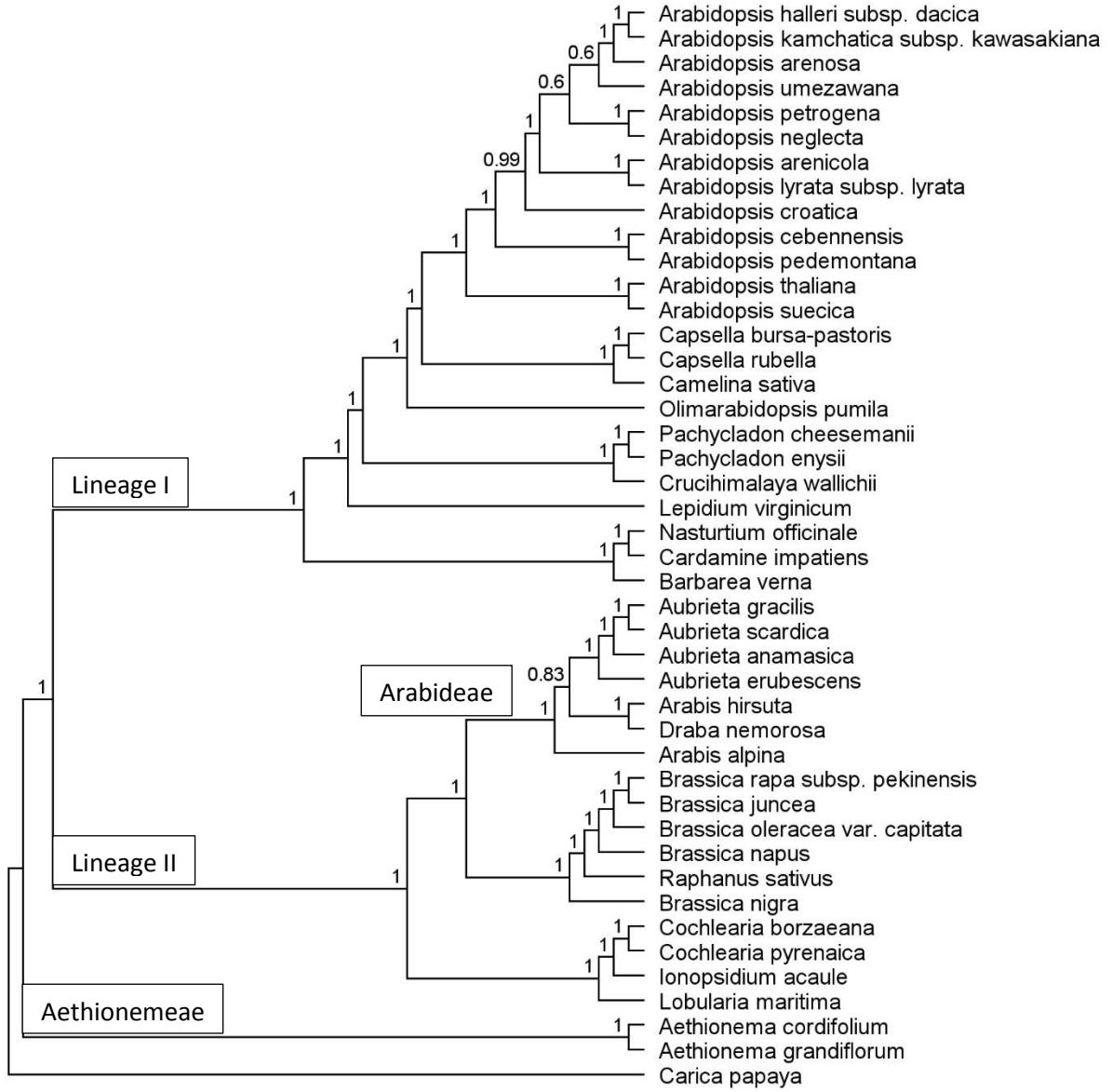


Figure 3.3 Bayesian phylogenetic tree for 43 species of Brassicaceae, including four of *Aubrieta*, based on 133 chloroplast genes plus introns and inter-genic spacers, from whole genome sequencing.

3.3.4 Mitochondrial genome assembly

Of the 100 contigs that were assembled for *Aubrieta gracilis*, only twenty were of mitochondrial origin. It is unlikely therefore that the total mitochondrial genome was sequenced completely. Nevertheless, the twenty contigs that were identified were mapped to two reference genomes: *Brassica napus* (Genbank accession number AP006444) of 221,853 bp (Handa 2003), and *Arabidopsis thaliana* (Genbank accession number NC_001284) of 366,924bp (Unseld *et al*. 1997). The sequence output was 151,680 bp in length. This consensus

sequence was mapped again to 52,967,592 whole raw reads from *Aubrieta gracilis*. A 151,830 bp- long sequence was assembled from 1,278,015 reads.

From the two reference taxa, 61 genes were identified in *Aubrieta gracilis*. These comprised 46 protein-encoding genes, of which one (nad2) is duplicated and another (nad1) is triplicated; 13 transfer-RNA genes, of which the trnS gene has three copies distributed all around the annotated mitochondrial genome. The total length of the exons is 36,832 bp. Individually they vary from 22 bp (nad5) to 3,169 bp (rrn26). See Table 3.4 for more details.

Table 3.4 List of mitochondrial genes in *Aubrieta gracilis*, with type and sequence length.

gene name	gene type	length	gene name	gene type	length
atp1	complex V	1524	orf120	open reading frame genes	363
atp6		792	orf147		446
atp9		225	orf159		478
ccmB	cytochrome-c-biogenesis	621	orf161		489
ccmC		744	orf25		579
ccmFC		1329	orfB		477
ccmFN1		1146	orfX		810
ccmFN2		636	rpl16	ribosomal proteins	540
cob	complex III	1182	rpl2		1047
cox1	complex IV	1584	rpl5		558
cox2		783	rps12		378
cox3		798	rps14		303
matR	other gene	2034	rps3		1677
nad1	complex I	591	rps4		1089
nad2		1467	rps7		445
nad3		357	rrn26	ribosomal RNAs	3169
nad4		1488	rrn5		119
nad4L		303	trnC	transfer RNAs	71
nad5		22	trnE		72
nad6		618	trnfM		74
nad7		1184	trnG		72
orf100a	open reading frame genes	302	trnI		85
orf100b		291	trnK		73
orf101b		308	trnM		73
orf106a		327	trnN		72
orf106b		322	trnP		75
orf113a		343	trnQ		72
orf114		344	trnS		262
orf115a		348	trnW		74
orf115b		342	trnY		83
orf116		352			

After alignment and analysis, I found that there were some genes missing from the sequence of *A. gracilis*, for example, from complex I: nad9; from complex II: sdhB, sdhC, sdhD; from ribosomal-RNA: rrn16; and from ribosomal proteins: rps1, rps2, rps11, rps13 and rps19 genes.

3.4 Discussion

3.4.1 Chloroplast genome

This is the first complete chloroplast genome sequence available for *Aubrieta*. In general, land plants have highly conserved chloroplast genomes (Kim and Jansen 1995), but their lengths vary due to the expansion or contraction of the inverted repeats and single copy regions (Wang *et al*. 2016). The plastome of *Aubrieta gracilis*, at 154103 bp, is of medium size in comparison with other Brassicaceae species, e.g. *Arabis alpina* 152866 bp (Melodelima and Lobréaux 2013) and *Cardamine impatiens* 155611 bp (Hu *et al.* 2015).

Although its 36.5% GC content would qualify *Aubrieta gracilis* as an AT-rich plastome, the GC value is comparable to other Brassicaceae plastomes, e.g. 36.3% in *Arabidopsis thaliana* (Sato *et al*. 1999*), Brassica napus* (Hu *et al*. 2011), *Cardamine impatiens* (Hu *et al.* 2015), and 36.45% in *Arabis alpina* (Melodelima and Lobréaux 2013).

In plants, the number of chloroplast genes ranges from 42 in the parasitic *Epifagus virginiana* to 110, 113 and 114 in rice, cucumber and tomato, respectively. These genes are protein-coding, transfer- and ribosomal-RNA. The annotated *Aubrieta gracilis* plastome comprises 88 protein-coding and 37 transfer-RNA genes which is similar to *Arabidopsis thaliana*, *Arabis hirsuta*, *Capsella bursa-pastoris*, *Brassica napus* and *Raphanus sativus*. *Aubrieta* has the rps16 gene, which is missing from the related *Arabis alpina* and *Draba nemososa* where it is believed to have been transferred to the nuclear genome during evolution (Melodelima and Lobréaux 2013). Similar transferrals of plastome genes to the nucleus have been recorded in *Populus alba* and *Medicago truncatula* by Ueda *et al.* (2008), who suggested that the plastome of these species is highly targeted by the nuclear genome.

In the comparison of the *Aubrieta* plastome with those in seven other species of Brassicaceae, selected on the basis of their phylogenetic proximity, there were fewer indels in *Aubrieta gracilis* in the trnT-trnE, psbE-petN and trnR-trnN intergenic spacer regions. This difference may be due to random mutations, including DNA sequence slipped-strand mispairing and illegitimate recombination events. Such explanations have been suggested previously in the context of other taxonomic groups by Levinson and Gutman (1987), Ogihara *et al*. (1998) and Kelchner (2000). In addition, in all eight species, these indels were especially abundant on the LSC or SSC regions when compared with inverted repeat regions. These results are very similar to the findings in other angiosperms taxa by Wolfe *et al*. (1989), Palmer (1991), Kim and Lee (2004), Kim *et al*. (2009), Yamane *et al*. (2006), Terakami *et al.* (2012) and Wang *et al.* (2016).

3.4.2 Phylogenetic topology of the Brassicaceae and the position of Aubrieta

Bayesian analysis of 133 genes plus their introns and inter-genic spacers from the whole chloroplast genome produced a tree (Fig. 3.3) that shows a large degree of congruence with the Brassicacaeae family phylogeny shown in Fig. 1.3 (see Introduction). Thus the basal clade corresponds to *Aethionema* (tribe Aethionemeae). The next most basal clade (Lineage III) was not included in my analysis (whole genome sequences unavailable), but lineages II and I were recovered with perfect support (p.p. = 1.00).

With respect to the genus *Aubrieta*, Fig. 3.3 shows that the genus falls in a clade which includes *Draba* and a polyphyletic *Arabis*. This confirms the position of *Aubrieta* in tribe Arabideae. Both the genus and the tribe appear to be monophyletic (p.p. = 1.00).

3.4.3 Mitochondrial genome

After alignment and analysis of the mitochondrial contigs, several genes appeared to be missing from the *Aubrieta gracilis* mitochondrial genome, for example, complex I: nad9; complex II: sdhB, sdhC, sdhD; ribosomal-RNA: rrn16; and ribosomal protein-coding genes: rps1, rps2, rps11, rps13, rps19. The explanation for this is that the total mitochondrial genome was not sequenced completely. Each single missing gene was unsuccessfully mapped to raw sequence data, indicating its complete absence rather than it being exported to the nuclear genome.

Although, the *Aubrieta gracilis* mitochondrial genome was not sequenced in its entirety, there are some features that appear to be homologous to those in *Arabidopsis thaliana* and *Brassica napus*. For instance, the ccmFN gene is split into two parts (ccmFN1-ccmFN2). This feature is considered to be typical of Brassicaceae (Handa *et al*. 1996; Unseld *et al*. 1997). And, the rps14 gene in *Aubrieta gracilis* is very similar to that in *Brassica napus*, where it is an open reading frame gene (Handa 2003), whereas in *Arabidopsis thaliana* it is a pseudo-gene (Brandt *et al*. 1993).

Lastly, the mitochondrial sequences were used to identify restriction sites whose occurrence and position differed between the four species of *Aubrieta*. These sites then formed the basis of an RFLP study conducted as part of the phylogenetic analysis of the genus described later (Chapter 4).

CHAPTER 4. PHYLOGENY OF *AUBRIETA*

4.1 Introduction

Although the genus *Aubrieta* is reasonably simple to identify, owing to its outer stamen filaments with a linear tooth below the anther, the identification of species within the genus is notoriously difficult. Authors of floristic accounts are unanimous in describing the genus as taxonomically difficult or critical (Mattfeld 1939; Cullen 1965; Gustavsson 1986). The morphological features used to try to distinguish the taxa comprise: stem and leaf pubescence; leaf shape (including the number of teeth); leaf hair anatomy; sepal size; sepal shape (inner saccate or not); petal size; petal colour; fruit size; fruit shape (outline and compression); fruit venation; and fruit hair anatomy. By using these characters, between 12—20 species have been recognised in modern floristic accounts (Cullen 1965; Gustavsson 1986; Phitos 1970, 2002; Al-Shehbaz *et al*. 2006).

Not only are the species hard to diagnose, very little is known about their evolutionary relationships. Karl *et al*. (2012) discovered from ITS and trnL-F sequences that the pan-Mediterranean *Arabis verna* was sister to *Aubrieta* and separated from other *Arabis* species; only 2—3 species of *Aubrieta* were included in their analyses, however. Karl and Koch (2013) increased their taxonomic sampling and reported trnL-F and ITS sequences for 13 species (14 accessions) of *Aubrieta* in their wider survey of the Arabideae. They produced plastid and nuclear trees for *Aubrieta* but many of the clades had no or only very weak support and there appeared to be significant differences between the trees (Fig. 4.1). Since the focus of their study was on the tribe Arabideae, Karl and Koch (2013) did not attempt to analyse or discuss the *Aubrieta* results further. Later, following a test for congruence between the trees, the combined dataset (ITS + trnL-F) was used to produce a phylogeny in which the position of a new species, *A. ekimii*, was identified (Fig. 4.1). However, with the exception of *A. canescens*, only one accession per species was used. Given the huge variability in the genus and especially the difficulty in diagnosing species, a more detailed study is needed in order a) to confirm and amplify the results of Karl and Koch (2013), and b) to investigate further the reality of the claimed congruence between chloroplast- and nuclear-derived datasets.

Against this background, I aimed to generate sequences from five plastid DNA regions (matK, trnH-psbA, trnD-trnT (=trnD-trnE + spacer + trnY-trnT), rps11-rpl36 and ycf6-psbM), RFLPs from six mitochondrial regions(ccmFC intron, matR, Nad9-ccmFN2, Orf25, Orf114 and trnK-rps3) and sequences from seven nuclear regions(duo1, ITS, lfy, pgiC, phytC, rpb2 and rpb2

intron23) in order to produce a robust phylogeny for the genus. A total of 45 accessions were studied.

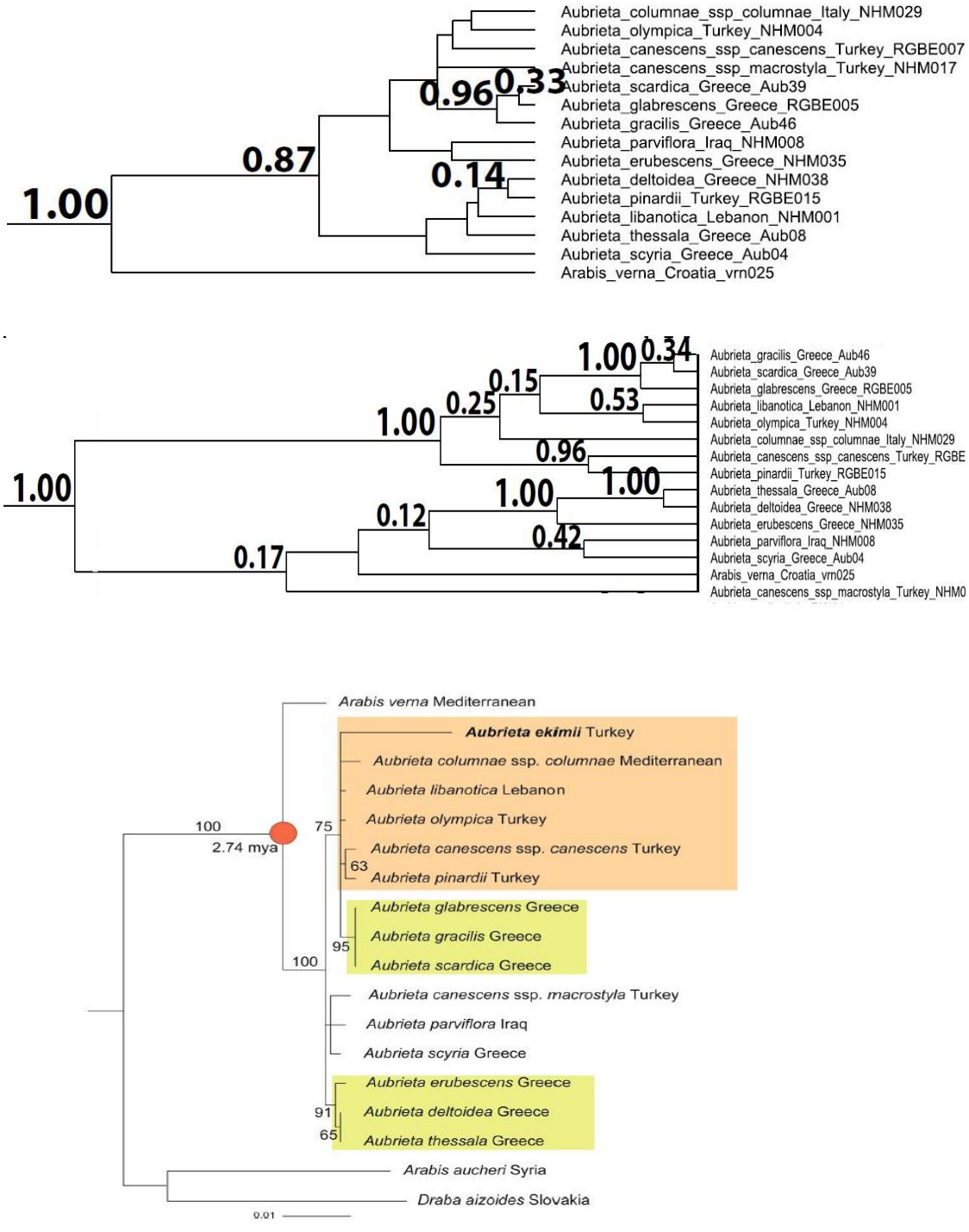


Figure 4.1 Phylogenetic trees of *Aubrieta* published by Karl and Koch (2013) and Yüzbaşıoğlu *et al*. (2015).

4.2 Materials and methods

See Chapter 2 for full laboratory details.

4.2.1 Phylogenetic analyses, including choice of outgroup

The data were analysed using the software packages *Mega 6* (Tamura *et al*. 2013) and *Geneious 7* (Kearse *et al*. 2012). Plastid alignments were conducted using *Geneious 7* with the following settings: a) alignment type: global alignment with free end gaps or ClustalW; b) cost matrix: 93% similarity; c) gap open penalty: 12; and d) gap extension penalty: 3. with nuclear analyses a model of evolution is required, obtainable only via *Mega 6*. Hence nuclear sequences were aligned using *Mega 6*, with the following settings: a) alignment type: ClustalW with free end gaps; b) cost matrix CLUSTALW; and c) gap open and extension cost: 6.

For reconstructing the phylogenetic trees, three methods were employed: a) maximum parsimony (MP); b) maximum likelihood (ML); and Bayesian inference (BI). For the latter two, models of evolution are required. These were selected using *Mega 6* software. The best fitting nucleotide substitution models were chosen using Modeltest v3.7 (Posada and Crandall 1998) which relied on the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The best model of evolution for each of the five chloroplast regions was Tamura 1992 (+G); second best was HKY+G, but this option was unavailable in the Bayesian programme in *Geneious 7*; instead the next best available model was chosen (HKY+G), its BIC score differing by 16. For the concatenated 5-region sequence the best model was Tamura 1992 (G+I), but this option was also unavailable in the Bayesian programme in *Geneious 7*; instead the next best model was chosen (HKY+G+I), its BIC score differing only by 2. To test the effect of using suboptimal models, ML trees using the optimal models were obtained using the software *Mega 6*. Their topologies (not shown) were congruent with Bayesian trees, although unresolved in a few places.

The best models of evolution for duo1 was K2+G+I, but since this is not available in *Geneious 7*, the next best one available was used, viz. GTR+G+I. For rpb2 and the combined nuclear set the best model was HKY+G+I and this was used.

Maximum parsimony analysis was conducted using *Mega 6* with the following settings: Test of phylogeny: bootstrap method, No. of bootstrap replications 1000, gaps/missing data treatment: use all sites, MP search method: subtree-pruning-regrafting (SPR), number of initial trees: 10, MP search level 1, max. No. of tree to retain: 100

Maximum likelihood analysis was also conducted using *Mega 6*, with the following settings: Test of phylogeny: Bootstrap method, no. of bootstrap replications 1000, gaps/missing data treatment: use all sites, ML heuristic method: subtree-pruning-regrafting (SPR) Initial tree for ML: make initial tree automatically (NJ), number of Threads: 1

For Bayesian inference, aligned sequences were analysed using MrBayes (Ronquist and Huelsenbeck 2003), using the following settings: chain length 20000, heated chains 4, rate variation (G+I), gamma categories 4, random seed 11451 and unconstrained branch length: exponential 10.

The out-group for the chloroplast analyses was chosen as *Arabis hirsuta*. This was because it is closely related to the *Aubrieta* clade (Karl and Koch 2013) and because the whole chloroplast sequence was available, allowing the individual regions used in this study to be compared.

4.3 Results

4.3.1 Chloroplast sequence data

Details of the five chloroplast DNA regions as found in the *Aubrieta* taxa studied here are shown in Table 4.1.

Table 4.1 Details of the chloroplast DNA regions sequenced.

DNA region	*matK*	*trnD-trnT*	*trnH-psbA*	*rps11-rpl36*	*ycf6-psbM*
Length (range), bp	1235-1257	1380-1439	354-518	539-551	606-632
G-C content %	30.9-32.8	30.9-32.3	12.9-20.8	39.7-41.3	25.8-29.7
Aligned length, sites	1241	1542	770	553	652
Informative sites	134	342	581	72	170
% polymorphic sites	10.8	22.2	75.5	13	26.1

In the following descriptions, the trees presented are derived from Bayesian analysis. They are extremely similar to those produced by MP and ML analyses so, in order to avoid cluttering up the account, they are not shown.

matK tree (Fig. 4.2)

The length of the matK gene among *Aubrieta* species ranges from 1235 to 1257 bp. Bayesian analysis shows that the taxa form a basal polytomy of five clades (Fig. 4.2), four of which have high posterior probabilities (pp). They correspond to clade A: *Aubrieta scyria* (S10e) (unresolved); clade B: a group of three accessions comprising *A. canescens* (S107), *A. pinardii* (E29) and *Arabis verna* (S18) (pp=0.75); clade C: a group of seven accessions from the Anatolia and the Near East (pp=1.0); clade D: accessions from the Aegean basin (pp=1.0); and clade E: a large group of 21 accessions (pp=1.0) within which there is a well-supported sub-clade of eight accessions from the Greek mountains.

trnD-trnT tree (Fig. 4.3)

The trnD-trnT region comprises three spacer regions and two genes (trnY and trnE). The combined length among *Aubrieta* taxa ranges from 1380-1439 bp. Although this is the longest sequence studied, it did not produce a well resolved tree by itself. The Bayesian tree shows an initial split into a basal lineage and a polytomy of eleven more unresolved lineages, including one large well-supported clade (A) of 21 accessions (pp=1.0) containing two further well-supported sister groups: one (B) of 13 Greek accessions (pp=1.0) and one (C) of eight Anatolian accessions (pp=1.0). Another clade (D) (pp=0.55) includes a subclade (E) of seven well-supported (pp=1.0) Near East accessions.

trnH-psbA tree (Fig. 4.4)

The trnH-psbA intergenic spacer region, including part of both genes (trnH and psbA) is comparatively short at 354-518 bp. This region is highly variable with many indels (insertions and deletions) among all species. The Bayesian tree is largely unresolved with none of the basal clades showing significant support.

rps11-rpl36 tree (Fig. 4.5)

The average length of rps11-rpl36 gene region is about 539-551bp. The few polymorphisms generate three basal lineages in the Bayesian tree, only one of which (A) has highly significant support (pp=1.0) and this comprises six accessions from Greece and two from cultivation. The second clade (B) (pp=0.74) comprises largely unresolved accessions from throughout the study area.

ycf6-psbM tree (Fig. 4.6)

The non-coding ycf6-psbM spacer is 606-632 bp. Analysis of this region resulted in a Bayesian tree with a reasonable amount of resolution but with rather weak support near the base. The tree splits into one basal group (pp=0.53), which includes a clade (A) of seven Anatolian accessions (pp=1.0), and a group (pp=0.58) of four other lineages. Among these, the better supported are (B) a clade comprised of ten Greek accessions (pp=0.93); (C) a clade containing two subgroups: (D) a Near East one (pp=0.91) and (E) an Aegean one (pp=0.86).

Combined chloroplast sequence tree (Fig. 4.7)

The tree produced from a Bayesian analysis of concatenated sequences from all five chloroplast regions is much better resolved than the individual 'gene' trees and has higher overall levels of support. It comprises six mostly well supported basal clades. The first clade is a basal group (A), which comprises *Aubrieta canescens* (S107-Aydin), *A. pinardii* (E29-Konya) plus *Arabis verna* (S18-Spain). The second clade (B) is a Near East clade, which comprises *Aubrieta libanotica* (E25-Lebanon and E21-Syria), *A. parviflora* (S77-Iraq, S104-Iran and S103-eastern Anatolia), *A. vulcanica* (S49-Nigde), *A. canescens* var. *macrostyla* (S101-Maras) as well as *A. scyria* (S10e-Kymii). The third clade (C) comprises taxa from the Aegean basin, comprising 13 accessions including *Aubrieta deltoidea* (S60-Crete, E17-Samos, E11-Lesvos islands, S62-Attica, S63-Phocis and BG14, BG6-botanic garden), *Aubrieta intermedia* (S50-Achaia), *Aubrieta intermedia* var. *macedonica* (E15-Kavalla), *A. thessala* (S17-Mt Olympos), *A. erubescens* (BG8-Mt Athos) and *A. intermedia* (BG7, BG16). The fourth clade (D) consists of accessions from western Anatolia, comprising *Aubrieta canesens* (E49-Asparta), *A. olympica* (E33-Samsun), *A.pinardii* (S106-Nevşehir), *A. olympica* (S105-Bursa), *A. canescens* var. *clilicica* (S102-Antalya), *A. canescens var. canescens* (S100-Eskişehir), *A. anamasica* (BG12) and *A. hesperidiflora* (BG3). The fifth clade (E) comprises accessions from the Greek mountains, viz. *Aubrieta gracilis* (BG1, S6c-Ioannina), *A. scardica* (BG2, S8a-Ioannina), *A. glabrescens* (S7e-Ioannina), *A.intermedia* (S2b-Achaia, S5b-Ioannina and S51-Phocis) and *A. macedonica* (BG5). The sixth clade (F) comprises accessions from Italy, Sicily and the Adriatic Balkan states, viz. *A. croatica* (BG13), *A. columnae* (BG4, BG9) and *A. sicula* (E19-Sicily).

Figure 4.2 Bayesian **matK** tree of *Aubrieta* taxa. The number above each branch is the posterior probability if it has > 0.5 support.

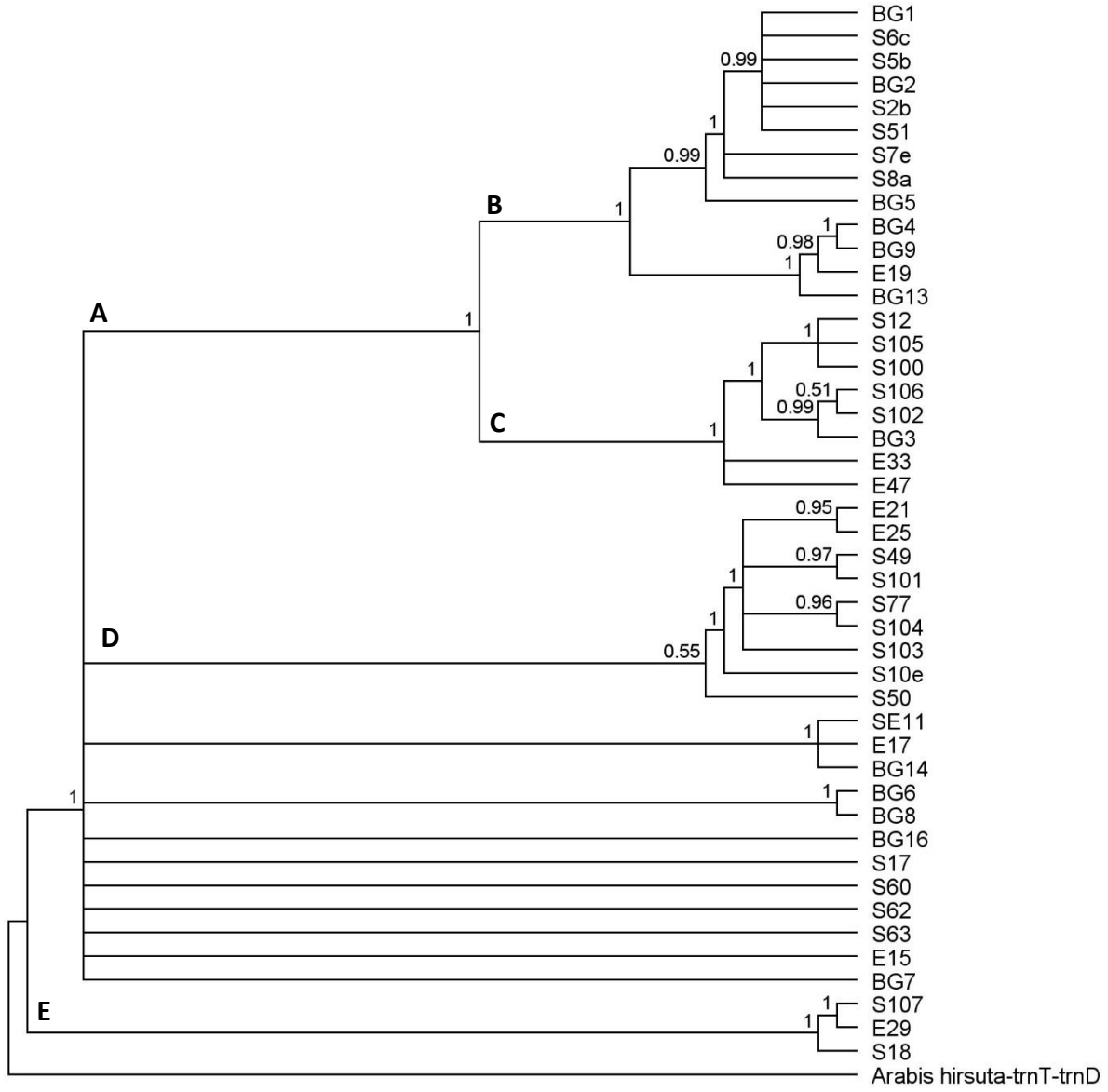


Figure 4.3 Bayesian **trnD-trnT**tree of *Aubrieta* taxa. The number above each branch is the posterior probability if it has > 0.5 support.

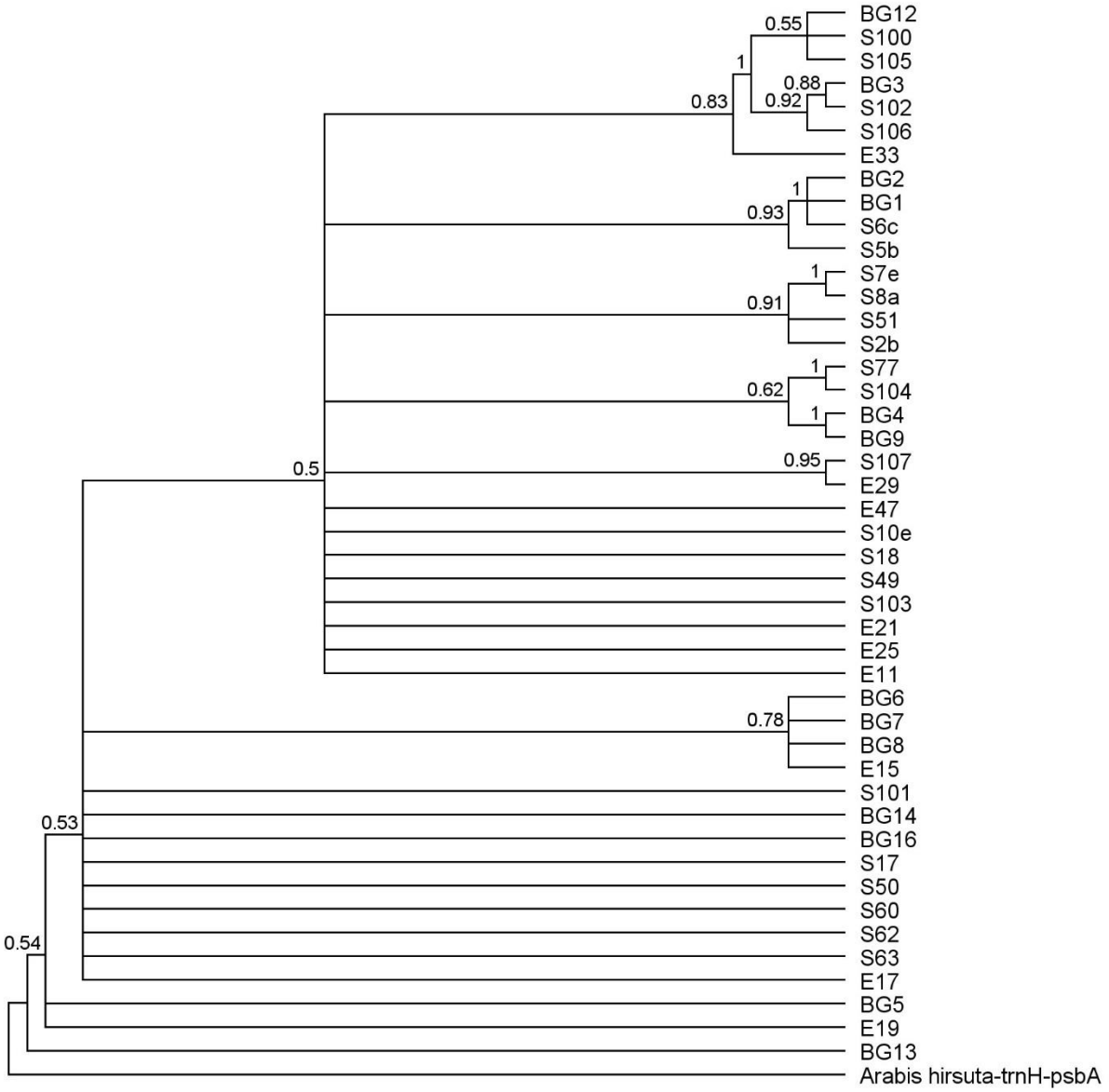


Figure 4.4 Bayesian **trnH-psbA**tree of *Aubrieta* taxa. The number above each branch is the posterior probability if it has > 0.5 support.

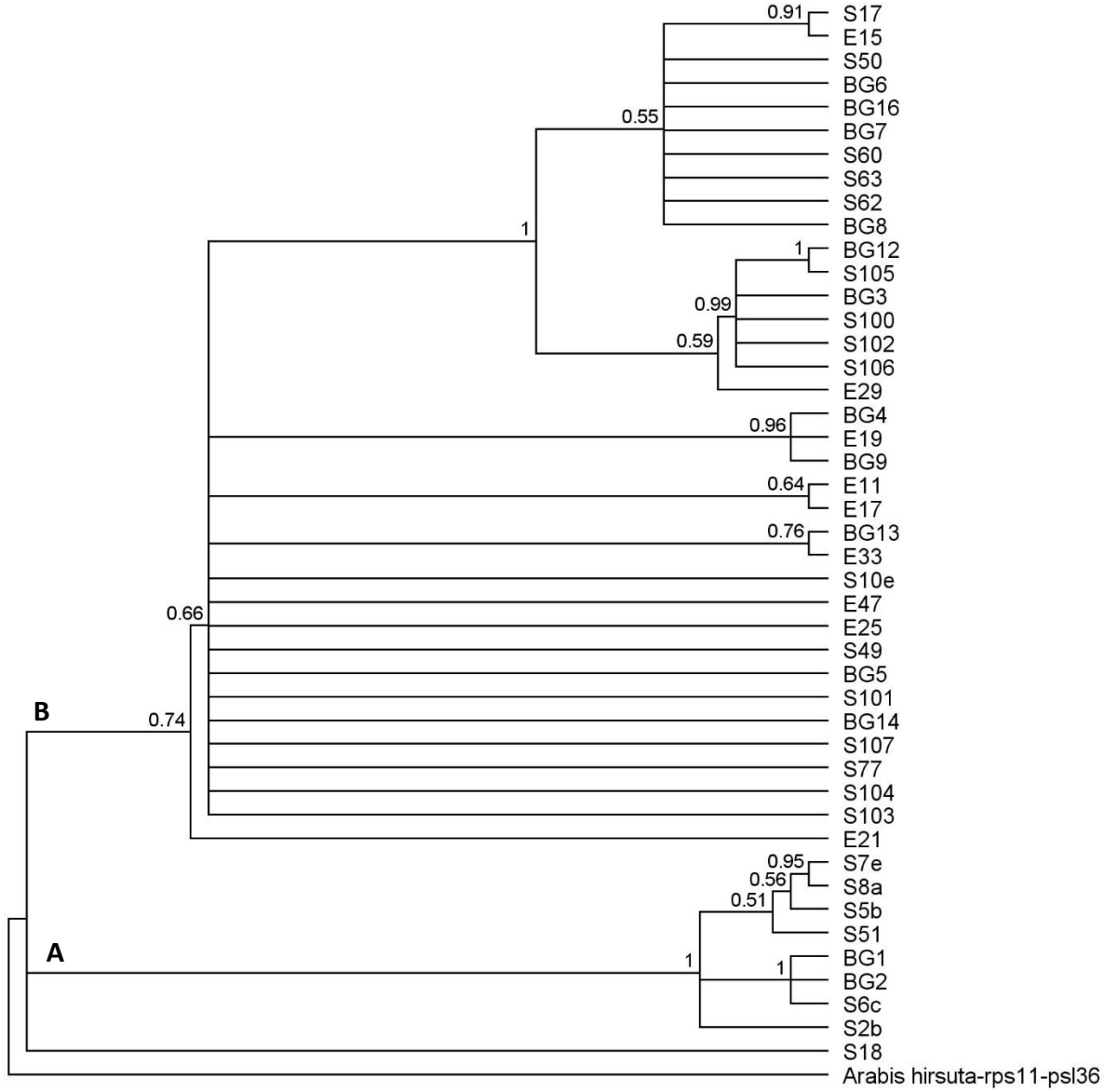


Figure 4.5 Bayesian **rps11-rpl36** tree of *Aubrieta* taxa. The number above each branch is the posterior probability if it has > 0.5 support.

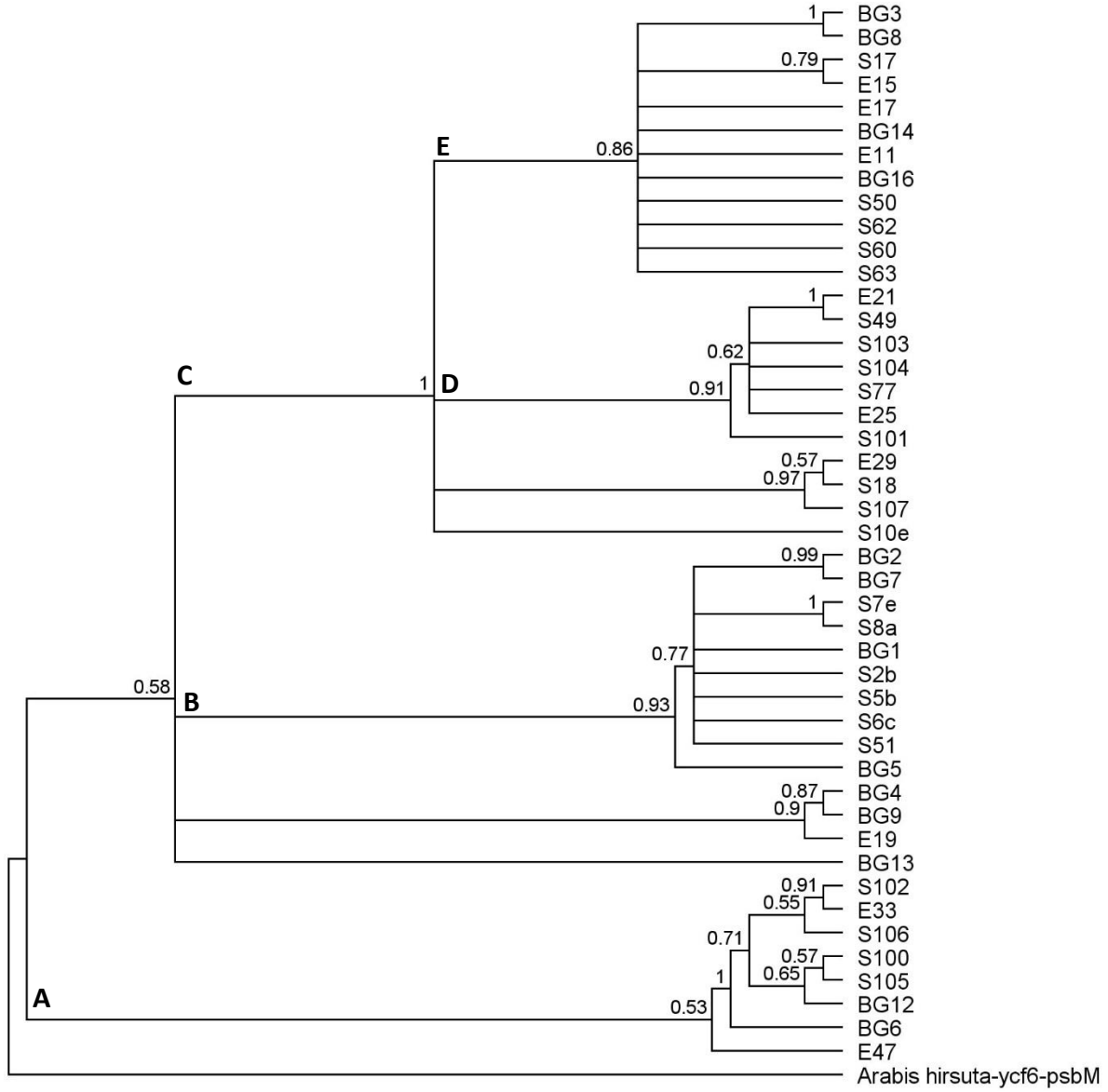


Figure 4.6 Bayesian **ycf6-psbM** tree of *Aubrieta* taxa. The number above each branch is the posterior probability if it has > 0.5 support.

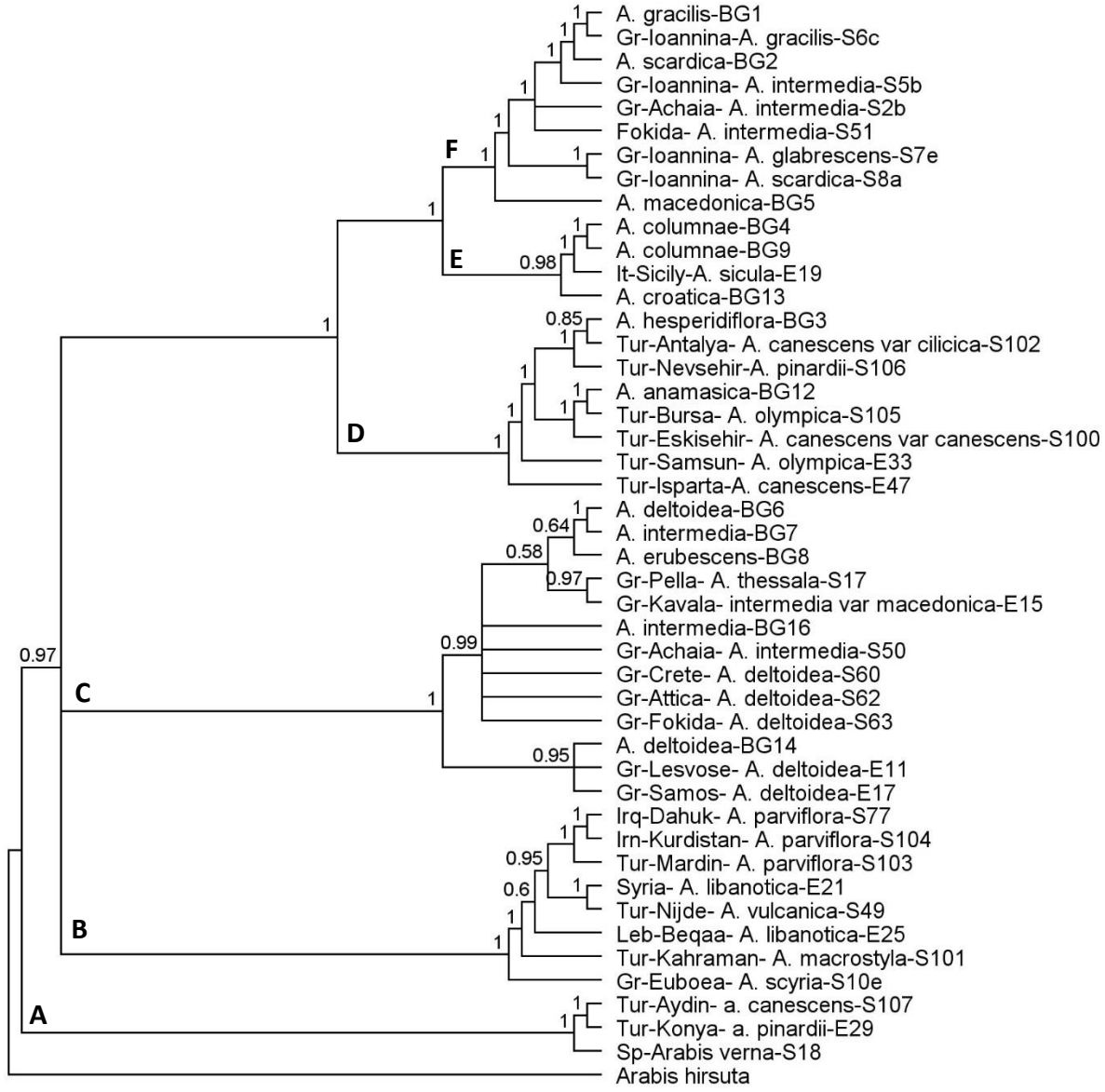


Figure 4.7 Bayesian tree of *Aubrieta* species using a concatenated chloroplast data set comprising the regions matK,trnT-trnD,ycf6-psbM, rps11-rpl36 and trnH-psbA. The number above each branch is the posterior probabilityif it has > 0.5 support. Accessions are coded according to their geographical origin: Gr = Greece, Irn = Iran, Iq = Iraq, Leb = Lebanon, Sp = Spain, Tur = Turkey; BG = of Botanic Garden origin.

4.3.2 Mitochondrial RFLP data

Four of the six restriction enzymes (BSTN1, EcoR1, Hinf1, Mse1) cut their respective DNA regions once or not at all. Enzyme mbo1 cut two or three times, and taq1 cut zero, one or two times (Figs. 4.8, 4.9, Table 4.2). These polymorphisms resulted in single-enzyme phenotypes as summarised in Table 4.3.

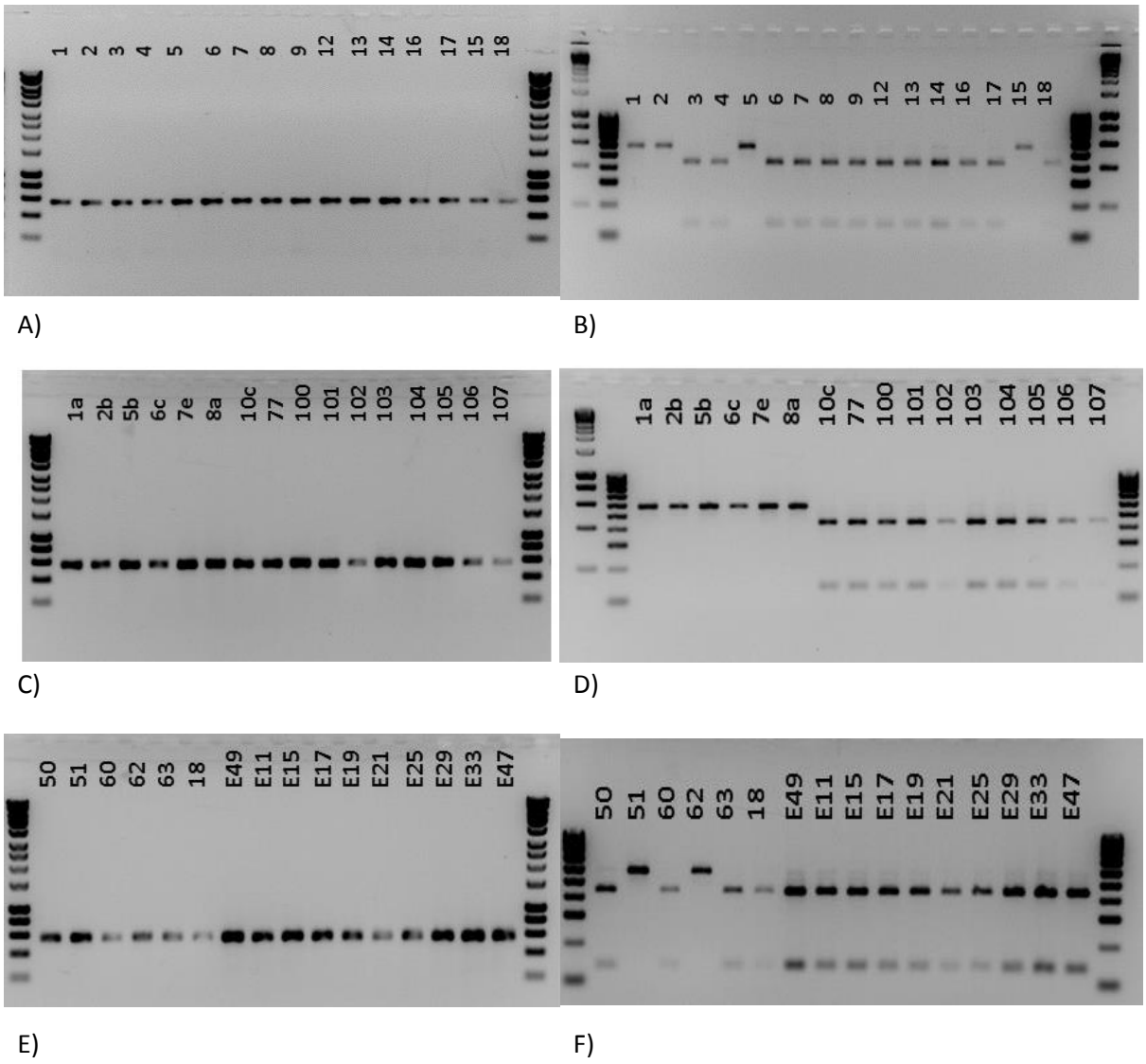
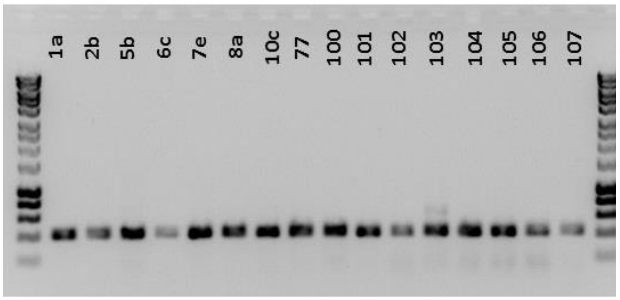
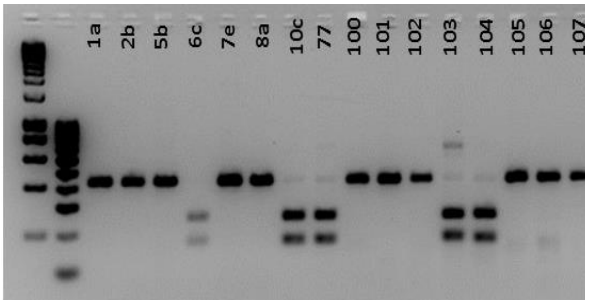


Figure 4.8 Amplified matR mitochondrial gene region: A, C, E) PCR products; B, D, F) respective fragments cut by hinfI restriction enzyme. The codes on the top of pictures A-F refer to *Aubrieta* accession numbers (see Table 2.1).

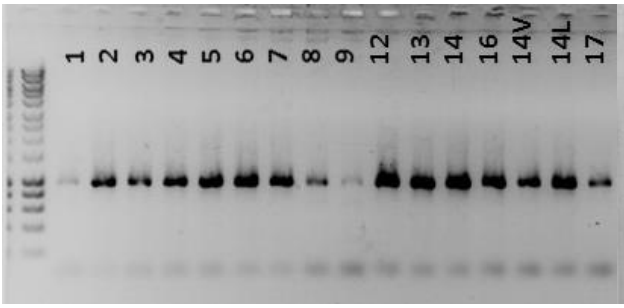
Figure 4.9 (overleaf) Amplified mitochondrial gene regions: A,C,E,G,I) PCR products of nad9-ccmFN2, Ccm-FC, Orf25, orf114 and trnK-rps3; B,D,F,H,J) fragments cut by EcoRI, Mbol, BSTNI, MseI, taqI restriction enzymes, respectively. Codes on the top of all pictures refer to *Aubrieta* accession numbers (see Table 2.1).


1a
2b
5b
6c
7e
8a
10c
77
100
101
102
103
104
105
106
107

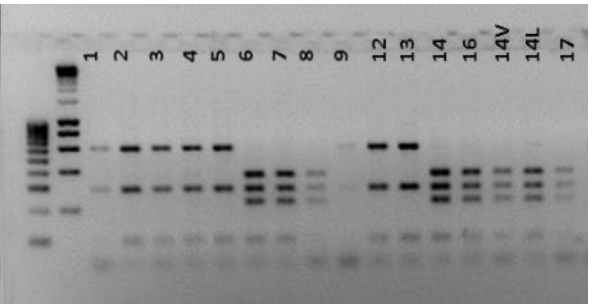
A)


1a
2b
5b
6c
7e
8a
10c
77
100
101
102
103
104
105
106
107

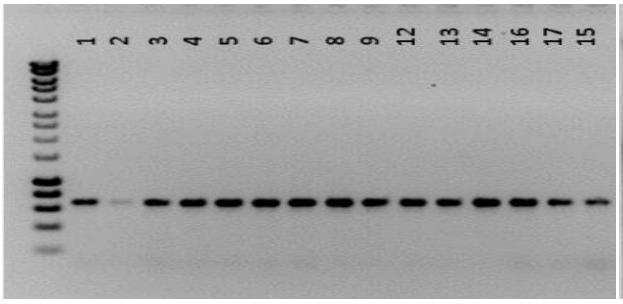
B)


1
2
3
4
5
6
7
8
9
12
13
14
16
14V
14L
17

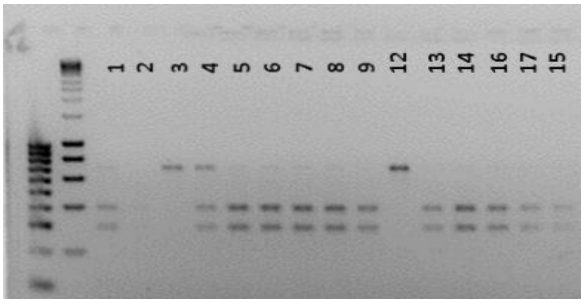
C)


1
2
3
4
5
6
7
8
9
12
13
14
16
14V
14L
17

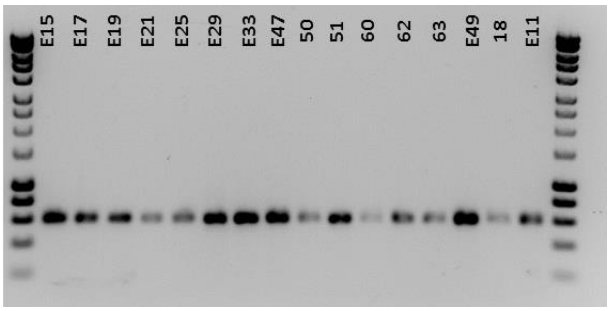
D)


1
2
3
4
5
6
7
8
9
12
13
14
16
17
15

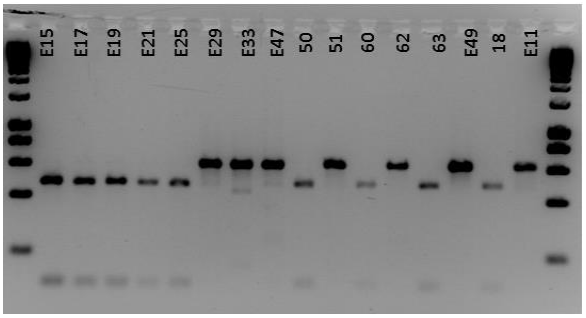
E)


1
2
3
4
5
6
7
8
9
12
13
14
16
17
15

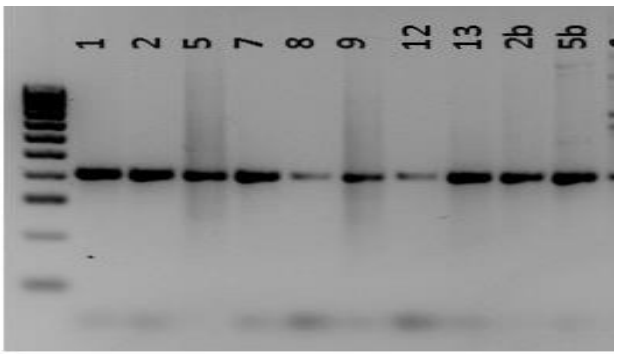
F)


E15
E17
E19
E21
E25
E29
E33
E47
50
51
60
62
63
E49
18
E11

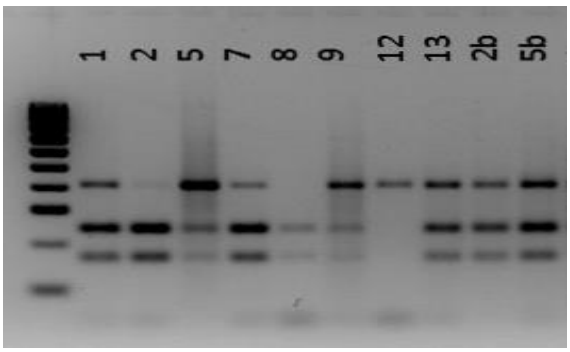
G)


E15
E17
E19
E21
E25
E29
E33
E47
50
51
60
62
63
E49
18
E11

H)


1
2
5
7
8
9
12
13
2b
5b

I)


1
2
5
7
8
9
12
13
2b
5b

J)

Table 4.2 Fragment number and sizes of mitochondrial single-enzyme RFLP phenotypes.

RFLP phenotype	Restriction enzyme	Fragment sizes, bp
A	MboI	93,282,585
B	MboI	93,219,282,366
C	HinfI	560
D	HinfI	128,432
E	MseI	607
F	MseI	127,480
G	EcoRI	425
H	EcoRI	174,251
I	BstNI	693
J	BstNI	303,390
K	taqI	407
L	taqI	167,240
M	taqI	167,240,407

Table 4.3 Mitochondrial RFLP single enzyme phenotypes: numbers refer to bands produced by each gene/enzyme combination.

Acces-sion	Mitochondrial gene fragment / restriction enzyme combinations											
	ccm-FC /MboI	Pheno-type	matR/ HinfI	Pheno-type	orf114/ MseI	Pheno-type	Nad9-ccmFN2 /EcoRI	Pheno-type	orf25/ BstNI	Pheno-type	trnKrps3/ taqI	Pheno-type
BG1	3	A	1	C	1	E	2	H	2	J	3	M
BG2	3	A	1	C	1	E	2	H	2	J	2	L
BG3	3	A	2	D	1	E	2	H	1	I	1	K
BG4	3	A	2	D	2	F	1	G	2	J	3	M
BG5	3	A	1	C	1	E	1	G	2	J	3	M
BG6	4	B	2	D	2	F	1	G	2	J	3	M
BG7	4	B	2	D	2	F	1	G	2	J	2	L
BG8	4	B	2	D	2	F	1	G	2	J	2	L
BG9	3	A	2	D	2	F	1	G	2	J	3	M
BG12	3	A	2	D	1	E	1	G	1	I	1	K
BG13	3	A	2	D	2	F	1	G	2	J	3	M
BG14	4	B	2	D	2	F	2	H	2	J	2	L
Bg16	4	B	2	D	2	F	1	G	2	J	2	L

Acces-sion	Mitochondrial gene fragment / restriction enzyme combinations											
	ccm-FC /MboI	Pheno-type	matR/ HinfI	Pheno-type	orf114/ MseI	Pheno-type	Nad9-ccmFN2 /EcoRI	Pheno-type	orf25/ BstNI	Pheno-type	trnKrps3/ taqI	Pheno-type
S17	4	B	2	D	2	F	1	G	2	J	2	L
S18	3	A	2	D	2	F	1	G	2	J	2	L
S2b	3	A	1	C	1	E	1	G	2	J	3	M
S5b	3	A	1	C	1	E	1	G	2	J	3	M
S6c	3	A	1	C	1	E	2	H	2	J	3	M
S7e	3	A	1	C	1	E	1	G	2	J	3	M
S8a	3	A	1	C	1	E	1	G	2	J	3	M
S10e	4	B	2	D	1	E	2	H	2	J	3	M
S49	3	A	2	D	1	E	1	G	2	J	2	L
S50	4	B	2	D	2	F	1	G	2	J	2	L
S51	3	A	1	C	1	E	1	G	2	J	2	L
S60	4	B	2	D	2	F	1	G	2	J	2	L
S62	3	A	1	C	1	E	1	G	n/a	n/a	2	L
S63	4	B	2	D	2	F	1	G	n/a	n/a	2	L
S77	3	A	2	D	2	F	2	H	n/a	n/a	n/a	n/a
S100	3	A	2	D	n/a	n/a	1	G	1	I	1	K
S101	4	B	2	D	1	E	1	G	2	J	2	L
S102	3	A	2	D	1	E	1	G	n/a	n/a	1	K
S103	3	A	2	D	2	F	2	H	2	J	2	L
S104	3	A	2	D	2	F	2	H	2	J	2	L
S105	3	A	2	D	1	E	1	G	1	I	1	K
S106	3	A	2	D	1	E	1	G	1	I	1	K
S107	n/a	n/a	2	D	2	F	1	G	n/a	n/a	2	L
E11	3	A	2	D	1	E	1	G	2	J	2	L
E15	4	B	2	D	2	F	1	G	2	J	2	L
E17	4	B	2	D	2	F	2	H	2	J	2	L
E19	3	A	2	D	2	F	1	G	2	J	2	L
E21	n/a	n/a	2	D	2	F	2	H	2	J	2	L
E25	3	A	2	D	2	F	2	H	2	J	2	L
E29	4	B	2	D	1	E	2	H	2	J	2	L
E33	3	A	2	D	1	E	1	G	2	J	1	K
E47	3	A	2	D	1	E	2	H	2	J	2	L

Altogether 13 multi-enzyme phenotypes (= mitotypes) were recovered. Accessions shared mitotypes as follows: mitotype (1) comprises *A. erubescens* (BG8), *A. intermedia* (BG7, BG16), *A. intermedia* var. *macedonica* (E15), *A. thessala* (S17), *A. deltoidea* (S50, S60, S63, BG6) and *A. canescens* (S107); mitotype (2) comprises only *A. canescens* var. *macrostyla* (S101); mitotype (3) comprises *A. deltoidea* (E11) and *A. vulcanica* (S49); mitotype (4)

comprises two accessions of *A. deltoidea* (BG14, E17); mitotype (5) comprises *A. columnae* (BG4, BG9), *A. croatica* (BG13), *Arabis verna* (S18) and *A. sicula* (E19); mitotype (6) comprises *A. scyria* (S10e), and *A. pinardii* (E29); mitotype (7) comprises *A. intermedia* (S2b, S5b, S51), *A. deltoidea* (S62), *A. macedonica* (BG5), *A. glabrescens* (S7e) and *A. scardica* (S8a); mitotype (8) comprises *A. gracilis* (BG1, S6c) and *A. scardica* (BG2); mitotype (9) comprises only *A. canescens* (E47); mitotype (10) comprises only *A. olympica* (E33); mitotype (11) comprises *A. parviflora* (S77, S103, S104) and *A. libanotica* (E21, E25); mitotype (12) comprises *A. anamasica* (BG12), *A. canescens* var. *canescens* (S100), *A. canescens* var. *cilicica* (S102), *A. olympica* (S105) and *A. pinardii* (S106); and mitotype (13) comprises only *A. hesperidiflora* (BG3). These groups are summarised in Table 4.4.

Table 4.4 Distribution of *Aubrieta* accessions among mitochondrial multi-enzyme haplotypes (mitotypes), chloroplast clades and nuclear clades.

Accession	Mitotype	Chloroplast clade	Nuclear clade	Accession	Mitotype	Chloroplast clade	Nuclear clade
				S2b	7	1	1
BG6	1	3	1	S5b		1	4
BG7		3	5	S62		3	5
BG8		3	1	S7e		1	1
BG16		3	-	S8a		1	1
E15		3	3	S51		1	1
S17		3	1	BG5		1	4
S50		3	4	BG1	8	1	1
S60		3	1	BG2		1	5
S63		3	1	S6c		1	1
S107		5	4	E47	9	2	2
S101	2	4	2	E33	10	2	4
E11	3	3	5	S77	11	4	2
E49		4	2	S103		4	2
BG14	4	3	1	S104		4	2
E17		3	4	E21		4	2
BG4	5	1	1	E25		4	2
BG9		1	1	BG12	12	2	4
S13-BG		1	1	S100		2	4
S18		5	4	S102		2	2
SE19		1	1	S105		2	4
S10e	6	4	1	S106		2	4
E29		5	4	BG3	13	2	4

Relationships between mitotypes, as revealed by Network analysis, are shown in Fig. 4.10.

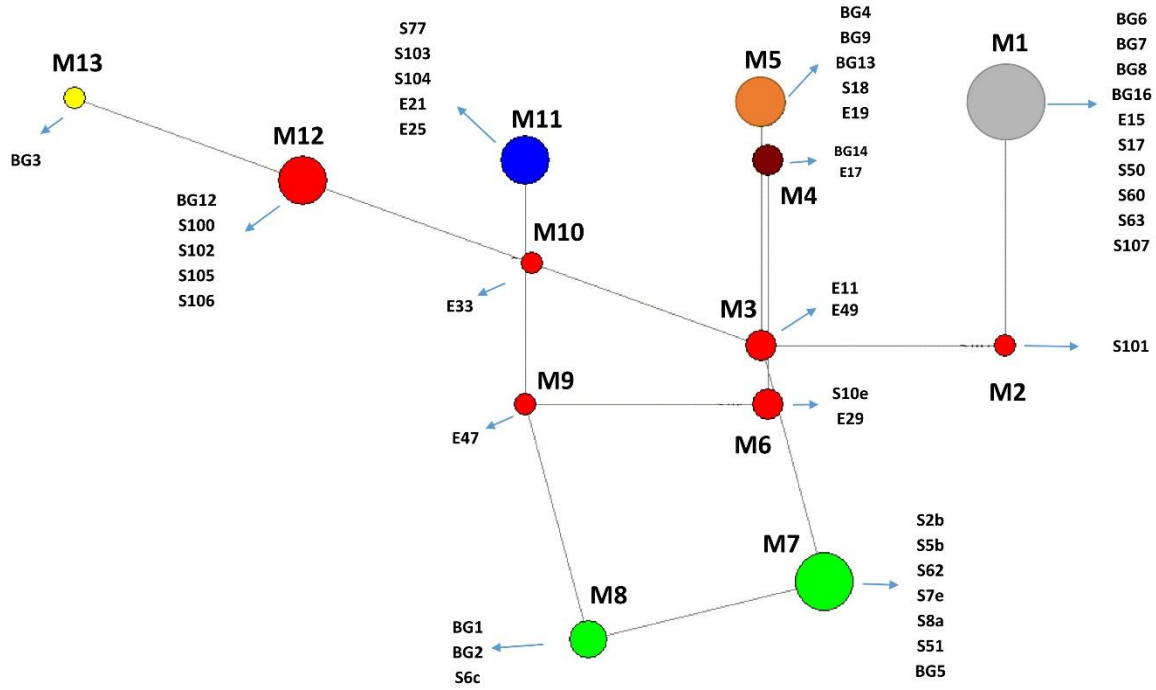


Figure 4.10 Network showing relationships between the 13 mitotypes (designated M) and the 45 accessions of *Aubrieta*. Internode lengths are proportional to the number of mutations.

4.3.3 Nuclear sequence data

As with the chloroplast data, the trees shown in the following account are derived from Bayesian analysis. Those produced using MP or ML are extremely similar and so are not shown here.

duo1 tree (Fig. 4.11)

The tree produced by Bayesian analysis of the duo1 data showed little resolution and such clades as were identifiable mostly had low levels of support.

rpb2 tree (Fig. 4.12)

The tree produced by Bayesian analysis of the rpb2 data showed a significant amount of resolution.

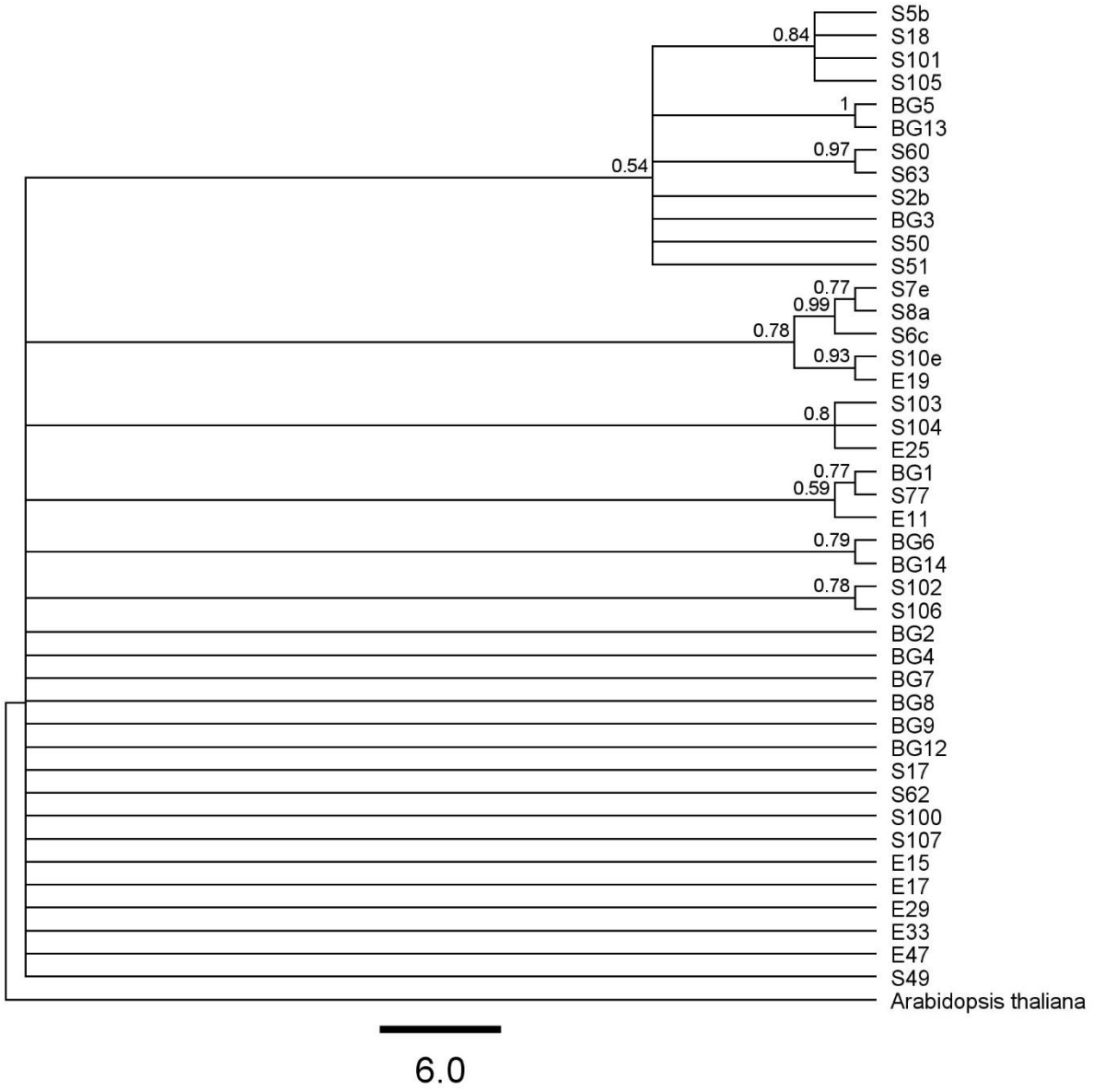


Figure 4.11 Bayesian tree of *Aubrieta* species using the nuclear gene, duo1. The number above each branch is the posterior probability if it has > 0.5 support. Accessions are coded as in Table 2.1.

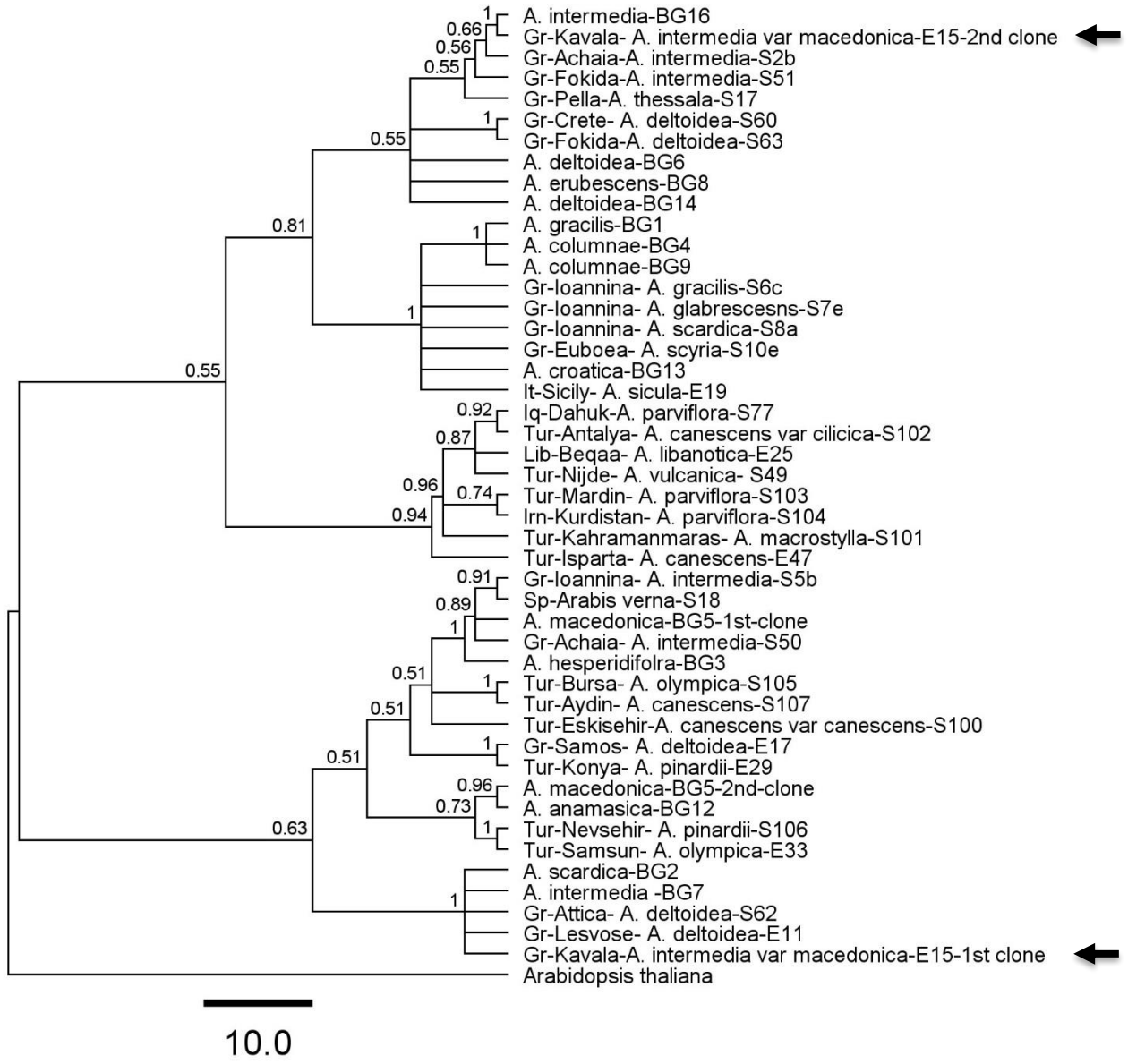


Figure 4.12 Bayesian tree of *Aubrieta* species using the nuclear gene, rbp2. The number above each branch is the posterior probabilityif it has > 0.5 support. Accessions are coded according to their geographical origin: Gr = Greece, Irn = Iran, Iq = Iraq, Leb = Lebanon, Sp = Spain, Tur = Turkey; BG = of Botanic Garden origin.

Combined nuclear sequence tree (Fig. 4.13)

The Bayesian analysis tree generated from concatenated nuclear sequences (rpb2 and duo1) is quite well resolved but support for the branches varies, being rather weak in some places, especially near the base. Clade A accessions are weakly grouped together (pp=0.54), and contain taxa from Anatolia, the Aegean basin and mainland Greece, comprising *A. deltoidea* (E11-Lesvos, S62-Attica, E17-Samos), *A. intermedia* (BG7, S50-Achaia, S5b-Ioannina), *A. scardica* (BG2), *A. anamasica* (BG12), *A. olympica* (E33-Samsun, S105-Bursa), *A. pinardii* (S106-Nevsehir, E29-Konya), *A. canescens* (S107-Aydin, S100-Eskisehir), *A. hesperidiflora* (BG3), *A. macedonica* (BG5), *Arabis verna* (Spain). Clade B has high support (pp=0.91) and comprises a group of eight accessions from Anatolia and the Near East, including *Aubrieta libanotica* (E25-Lebanon), *A. parviflora* (S77-Iraq, S104-Iran, and S103-eastern Anatolia), *A. vulcanica* (S49-Nigde), *A. canescens* (E49-Isparta), *A. canescens* var. *cilicica* (S102-Antalya) and *A. canescens* var. *macrostyla* (S101-Maras). Clade C, also well supported (pp=0.96), comprises 17 accessions from Sicily, the Greek mountains and the Aegean basin, including *Aubrieta glabrescens* (S7e-Ioannina), *A. scardica* (S8a-Ioannina), *S. gracilis* (S6c-Ioannina), *A. scyria* (S10e-Euboea), *A. erubescens* (BG8-Athos), *A. deltoidea* (S63-Phocis, S60-Crete, BG6, BG14), *A. thessala* (S17-Mt Olympos), *A. intermedia* (S2b-Achaia, S51-Phocis), *A. sicula* (E19-Sicily), *A. croatica* (BG13-Croatia) and *A. columnae* (BG4, BG9).

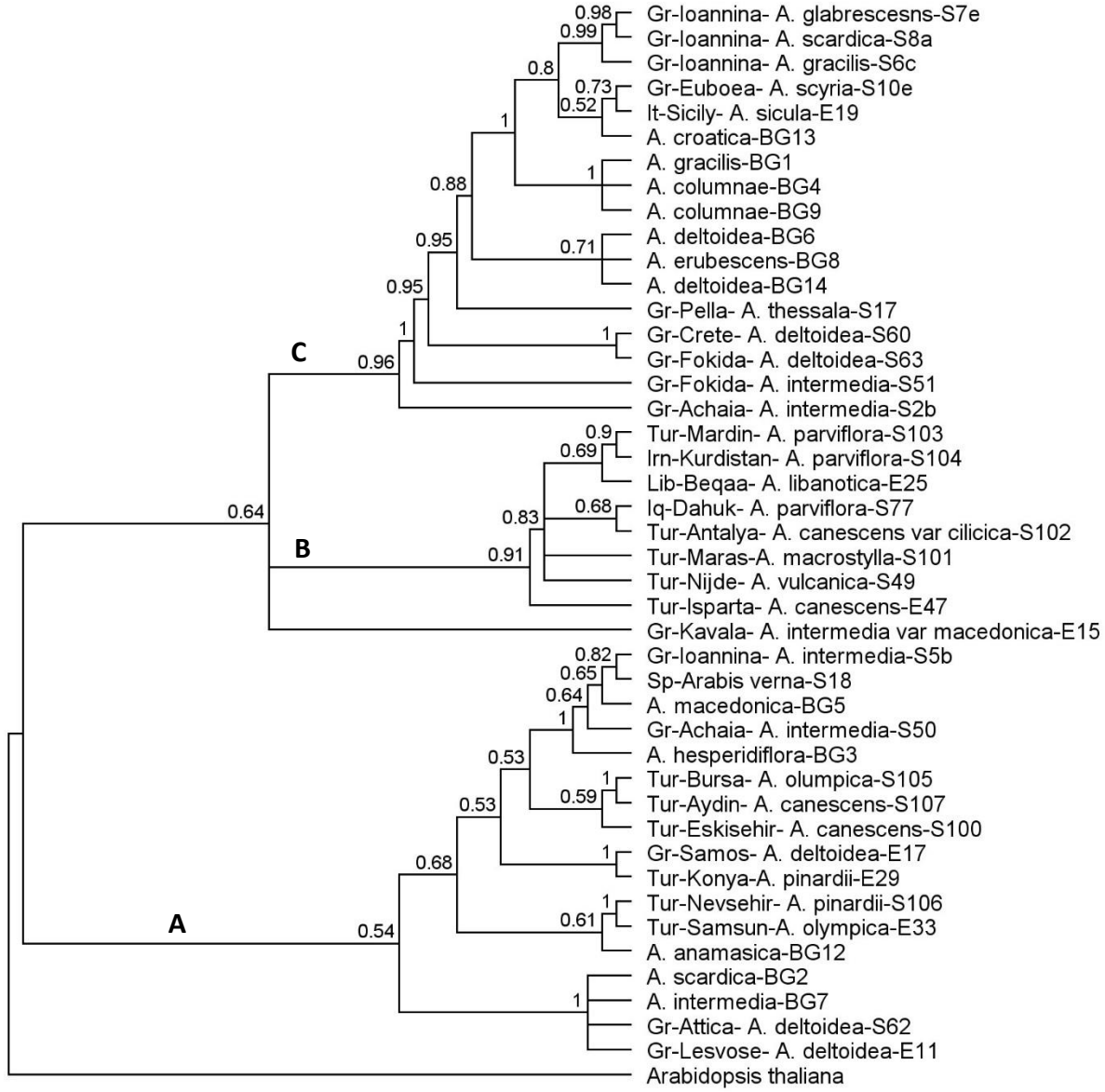


Figure 4.13 Bayesian tree of *Aubrieta* species using a concatenated nuclear data set comprising the regions (rbp2 & duo1). The number above each branch is the posterior probabilityif it has > 0.5 support. Accessions are coded according to their geographical origin: Gr = Greece, Irn = Iran, Iq = Iraq, Leb = Lebanon, Sp = Spain, Tur = Turkey; BG = of Botanic Garden origin.

4.3.4 Heterozygosity

Evidence of significant heterozygosity was detected among *Aubrieta* taxa, based on analysis of NGS data and on study of sequence data from the low copy number nuclear gene regions, rpb2 and duo1 (Table 4.5).

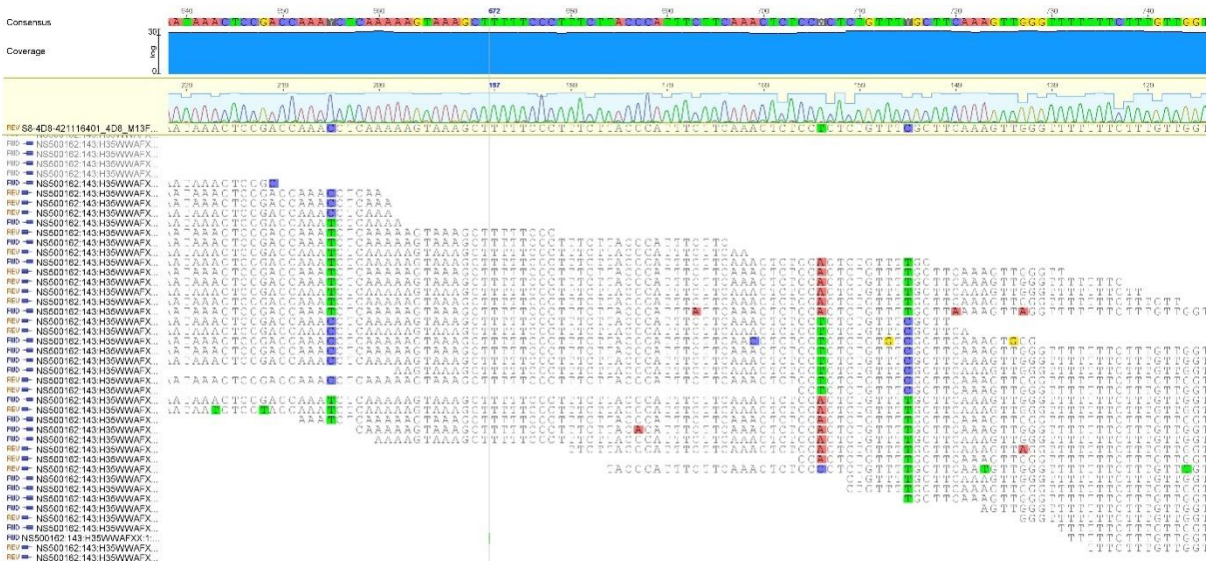
Table 4.5 List of accessions that show heterozygosity

Taxon	Accession	Evidence
Arabis verna	S18	Cloned alleles
Aubrieta canescens	S100	Cloned alleles
Aubrieta canescens	S107	Cloned alleles
Aubrieta canescens var. *cilicica*	S102	Cloned alleles
Aubrieta deltoidea	E11	Cloned alleles
Aubrieta deltoidea	E17	Cloned alleles
Aubrieta deltoidea	S60	Direct sequence: double peaks
Aubrieta deltoidea	S62	Direct sequence: double peaks
Aubrieta erubescens	BG8	NGS: double reads
Aubrieta intermedia	S5b	Cloned alleles
Aubrieta intermedia var. *macedonica*	E15	Cloned alleles
Aubrieta pinardii	S106	Cloned alleles
Aubrieta olympica	S105	Cloned alleles
Aubrieta scyria	S10e	Direct sequence: double peaks
Aubrieta thessala	S17	Cloned alleles

A. erubescens (BG8) was sequenced by NGS methods and showed clear evidence of heterozygosity in both rpb2 and duo1 gene regions (Fig. 4.14). Other taxa shown in Table 4.5 also showed significant evidence of heterozygosity as inferred from direct sequencing and cloning results. For example, direct sequencing of *A. scyria* (S10e), *A. deltoidea* (S60) and *A. deltoidea* (S62) produced double electropherogram peaks at many sites in the highly variable introns (Fig. 4.15). Some accessions, however, proved resistant to direct sequencing because of the extent of heterozygosity, and fully readable sequences could only be obtained through the cloning procedure (see Methods). The accessions concerned were *A. intermedia* (S5b), *A. deltoidea* (E11), *A. intermedia var macedonica* (E15), *A. deltoidea* (E17), *A. thessala* (S17), *A.*

canescens (S100 and S107), *A. canescens var cilicica* (S102), *A. olympica* (S105) and *Arabis verna* (S18). Cloning of duo1 products resulted in the recovery of two allele sequences in each case. Cloning of rpb2 products resulted in the recovery of two allele sequences in some cases but only one in others. See Fig. 4.15 B-F, for the electropherograms in which the direct sequence profile is contrasted with the profiles of the two allelic clones.

A



B

Figure 4.14 NGS assembled data of *A. erubescens* (BG8) featuring the genes rpb2 (A) and duo1 (B).

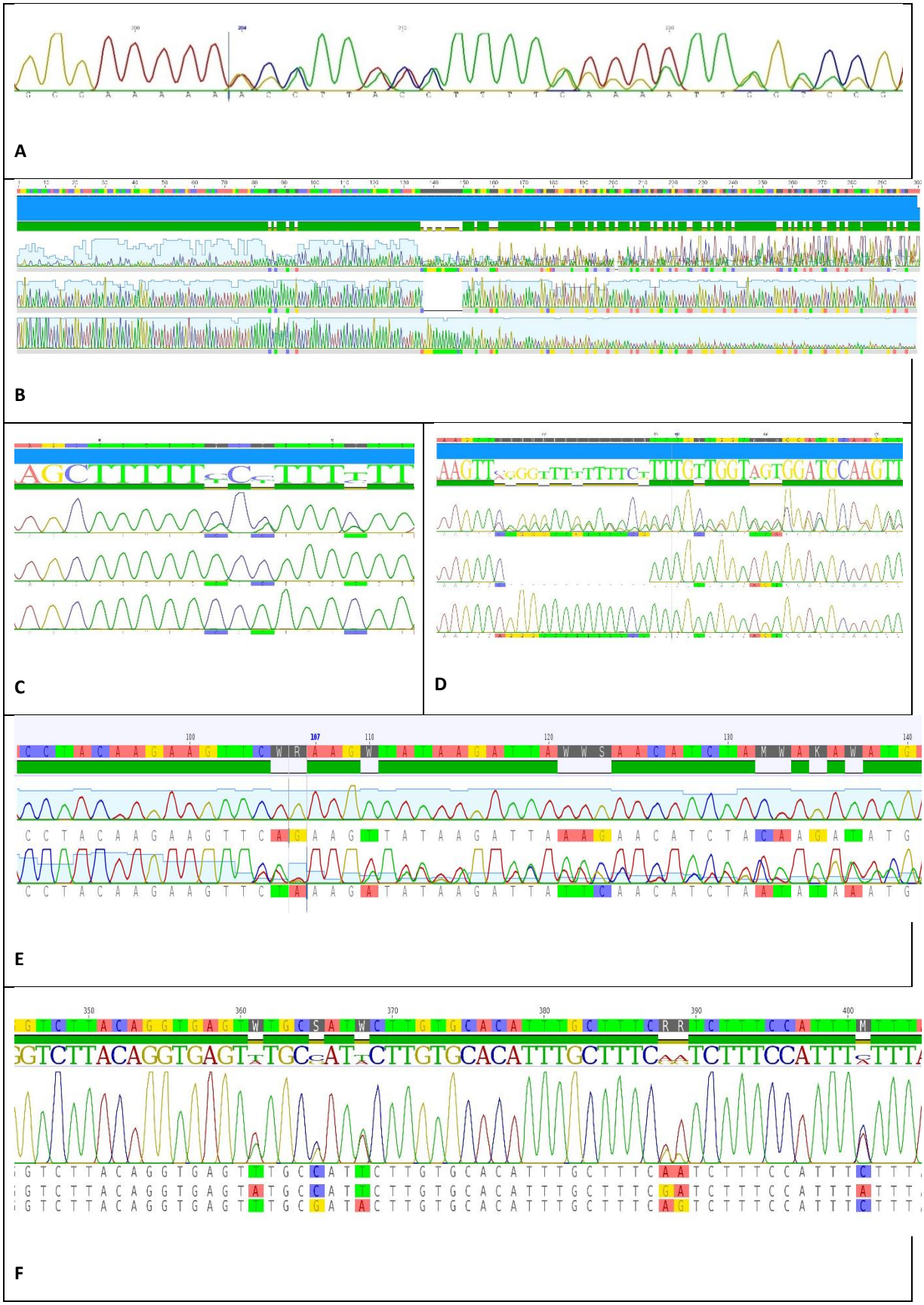


Figure 4.15 Electropherograms showing double peaks. A: direct sequence of duo1 in *A. scyria* (S10e) showing extensive heterozygosity. B-D: duo1 direct (top) and cloned (bottom pair) sequences from *A. canescens* (S107). E: rpb2 direct (bottom) and cloned (top) sequences from *A. canescens* (S107). F: rpb2 direct (top) and cloned (bottom pair) sequences from *A. intermedia* var. *macedonica* (E15).

Depending on the allele present, the sequences from the same accession cluster differently. Thus, the Bayesian phylogenetic tree derived from the rpb2 sequences (Fig. 4.12) showed that one allele (clone) of *A. intermedia* var. *macedonica* (E15) clusters with *A. deltoidea* from Lesvos Island (E11) and Attica (S62) and the other allele (clone) clusters with *A. intermedia* from south-central Greece (e.g. S51). In nearly all cases the allelic variation was located in intron sequences. In a few cases in rpb2, however, allelic variation ocurred as SNPs in the exon region. In these cases, the mutations were always in the third position and thus did not affect the protein produced.

Similarly, the Bayesian phylogenetic tree derived from the duo1sequences demonstrated that there was evidence of significant heterozygosity among Turkish taxa as shown in Fig. 4.16. Here, one allele (clone) of *A. canescens* (S107) clusters with each of *A. canescens* (S100, E47) and the other allele (clone) clusters with *A. canescens* var. *macrostylla* (S101) and *A. olympica* (S105). Likewise, the two alleles of *A. canescens* var. *cilicica* (S102) split into two different clades: one of them clustered with *A. pinardii* (S106) and the other one clustered with the Near-eastern clade (Fig. 4.16). And similarly in *A. pinardii* (S106) the two alleles are separated. (Fig. 4.16).

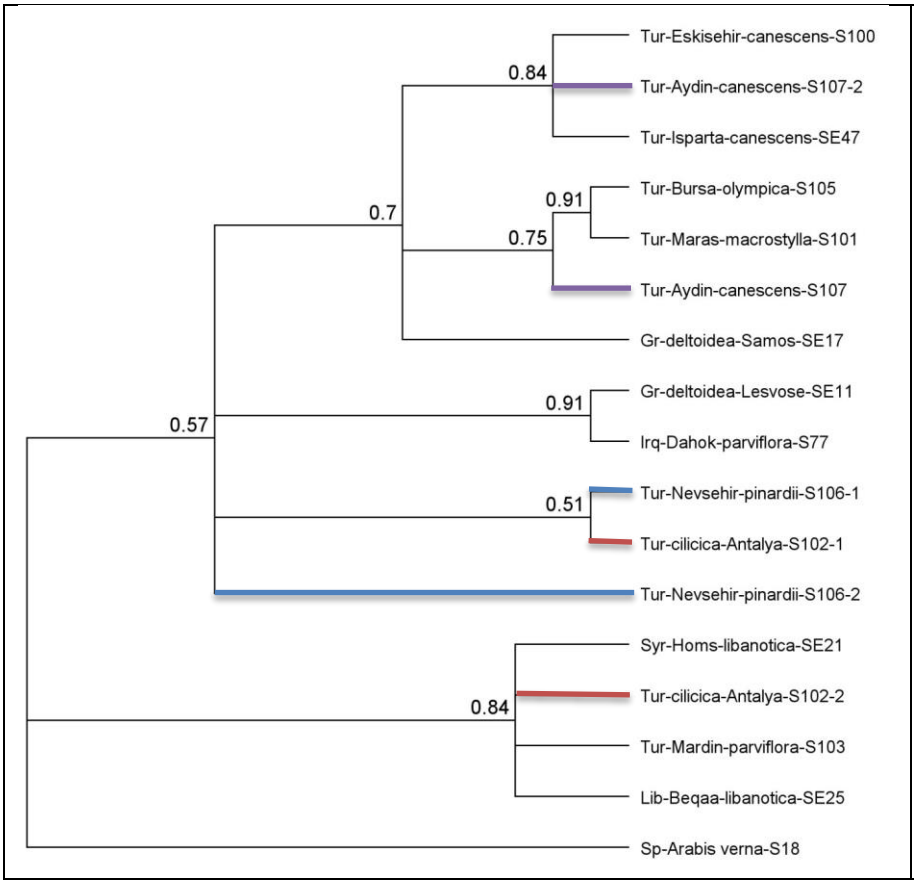


Figure 4.16 Bayesian tree of *Aubrieta* taxa using a nuclear gene (duo1). The number above each branch is the posterior probability if it has > 0.5 support. Coloured internodes indicate the pairs of alternative alleles. Accessions are coded according to their geographical origin: Gr = Greece, Irn = Iran, Iq = Iraq, Leb = Lebanon, Sp = Spain, Tur = Turkey.

It could be claimed that the double peaks and cloning results are consistent with gene duplication rather than heterozygosity. However, double peaks are confined to intron regions and occasionally to the third position of a codon. This is highly suggestive of heterozygosity rather than a duplication event. If it were the latter it would be expected that double peaks would occur also at all three positions in the codons (assuming duplication is not recent, allowing divergence to have occurred). The exons in both alleles have functional open-reading frames which are translated, again indicative of a heterozygous situation.

4.4 Discussion

4.4.1 Chloroplast phylogeny

Since the individual chloroplast regions are part of a non-recombining molecule and their respective topologies are congruent with each other, I shall confine my discussion of the results to the tree produced by the combined dataset. This tree shows six chief lineages which correspond in large measure to geographical regions (Fig. 4.18). Thus there is a small, basal clade of two Anatolian accessions together with the pan-Mediterranean *Arabis verna*. Clade 2 comprises taxa from the Near East, including SE. Anatolia and perhaps surprisingly, *A. scyria* from the Greek side of the Aegean basin. Clade 3 comprises taxa from the Aegean basin. Clade 4 is from western Anatolia. Clade 5 is largely from the Pindus Mountains of Greece. Clade 6 consists of accessions from Italy and Sicily. The western Anatolian clade is strongly supported as a sister group to the Greek Pindus plus Italy clade (pp=0.98). In contrast, the position of the Aegean basin clade is not well supported (pp=0.62) so its clustering with the Near East clade is not secure. Similarly, the clade containing *Arabis verna* is not well located on the tree.

The boundaries between the ranges of the different plastid clades correspond to geographical features. Thus the boundary between the Anatolian and Near East clades correlates well with the so-called Anatolian Diagonal (Davis 1971), a range of mountains that includes the Anti-Taurus and which runs north-east to south west. The boundary between the Anatolian and Aegean basin clades corresponds to the Thracian Plain and the Marmara and Aegean Seas, although there is obviously some overlap on and around the Aegean coast of Anatolia. To the west, the Aegean basin clade overlaps somewhat in its geographical distribution with that which occupies the Pindus mountain range, a chain of mountains running southwards down the western part of Greece to the Peloponnese.

An understanding of the origin and relationships of these clades requires further evidence from additional types of DNA sequence and these are described in the next sections. The phylogeographical story is taken up further in Chapter 6.

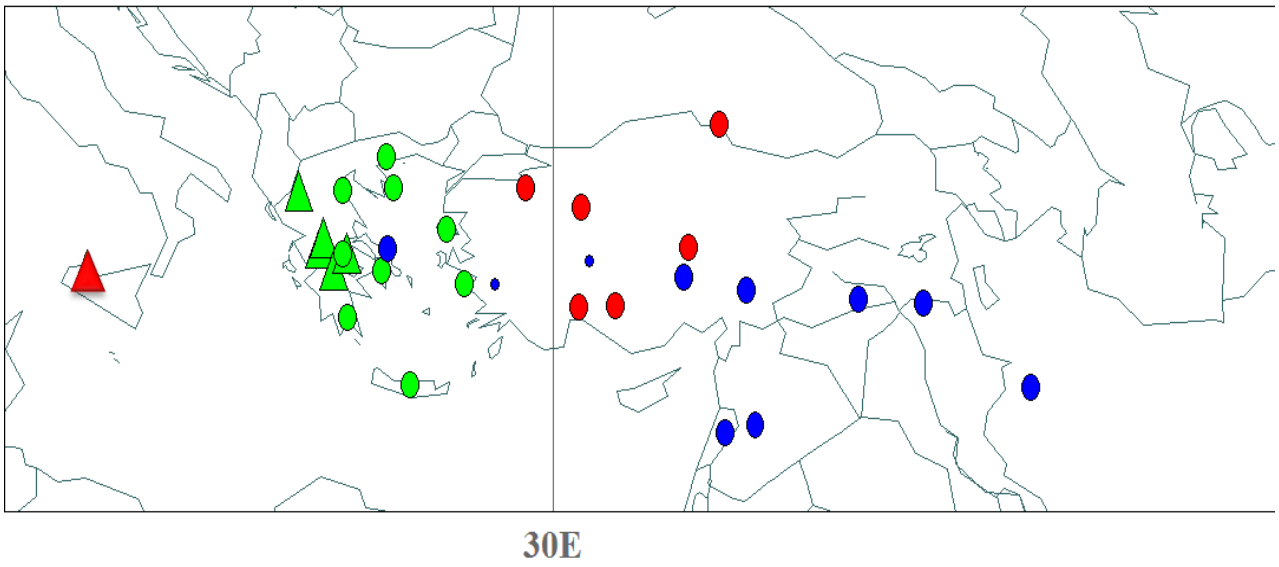


Figure 4.18 Geographic distribution of *Aubrieta* chloroplast clades based on a concatenated 5-region phylogeny. Large blue circles = Near Eastern clade; red circles = Anatolian clade; green circles = Aegean basin clade; green triangles = Greek mountain clade; red triangle = Italy/Sicily clade; small blue circles form a well-supported clade along with the pan-Mediterranean *Arabis verna*.

4.4.2 Mitochondrial phylogeny

The 13 mitochondrial haplotypes identified by PCR-RLFP methodology are related according to the network shown in Fig. 4.10. The branches correspond to geographical groupings (Fig. 4.19) as follows: i) an Italian (Sicily) mitotype, #5; ii) a Greek mountain mitotype group, #7 and 8; iii) an Aegean basin mitotype, #1; iv) an Anatolia/Aegean mitotype, #4; v) a chiefly Anatolian group, consisting of six related mitotypes, #2, 3, 6, 9, 10, 12; and vi) a Near-East group consisting of mitotype 11. Anatolia is clearly a centre of mitotyope diversity, with adjacent areas to the west and to the east/south-east having far less. This pattern is consistent with that found by Ansell *et al*. (2011) in their work on chloroplast and nuclear sequence variation in *Arabis alpina*.

The phylogenetic relationships indicated by the uniparentally-inherited mtDNA mirror to a large extent those shown by chloroplast DNA, pointing to a highly geographically-structured genepool for *Aubrieta*.

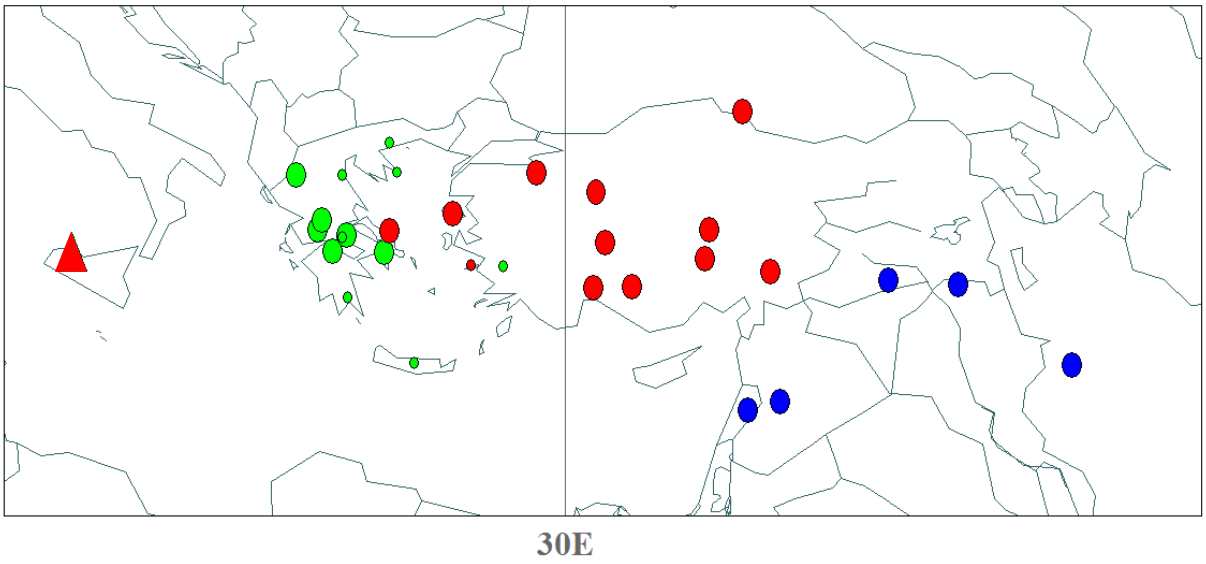


Figure 4.19 Geographic distribution of *Aubrieta* mitochondrial haplotype (mitotype) groups: i) Sicily mitotype, #5 (triangle); ii) Greek mainland mitotypes, #7 & 8 (big green circle); iii) Aegean Basin mitotype, #1 (small green circle); iv) Anatolia (chiefly) mitotype group, #2, 3, 6, 9, 10 & 12 (big red circle); v) Anatolia/Aegean mitotype, #4 (small red circle); vi) Near East mitotype, #11 (big blue circle).

4.4.3 Nuclear phylogeny

The tree derived from the combined nuclear sequences shows three main clades. The first, clade A, is not only in itself weakly supported but so also are almost all of its internal nodes. It contains a mixture of accessions from Anatolia, the Aegean basin and montane Greece. The second clade, B, is strongly supported (pp=0.91) and comprises accessions from Anatolia and the Near East. The third clade is also well supported (pp=0.96) and comprises accessions from Sicily, Greece and the Aegean basin. Although not nearly so pronounced as with the chloroplast and mitochondrial data, there is still a recognisable geographical structuring to the nuclear phylogeny which corresponds imperfectly with the plastid and mitochondrial trees (Fig. 4.20). Thus we can infer the existence of a Near-East genepool, a highly variable Anatolian genepool, another in the Aegean Basin, and another to the west. One of the reasons for the less than perfect match between the nuclear and plastid / mitochondrial trees is the rather variable position of accessions of *A. deltoidea* and *A. intermedia*. Some accessions of these taxa cluster in Clade A and others in Clade C. This is suggestive of a hybrid origin, evidence for which exists in the form of the heterozygous sequences described earlier.

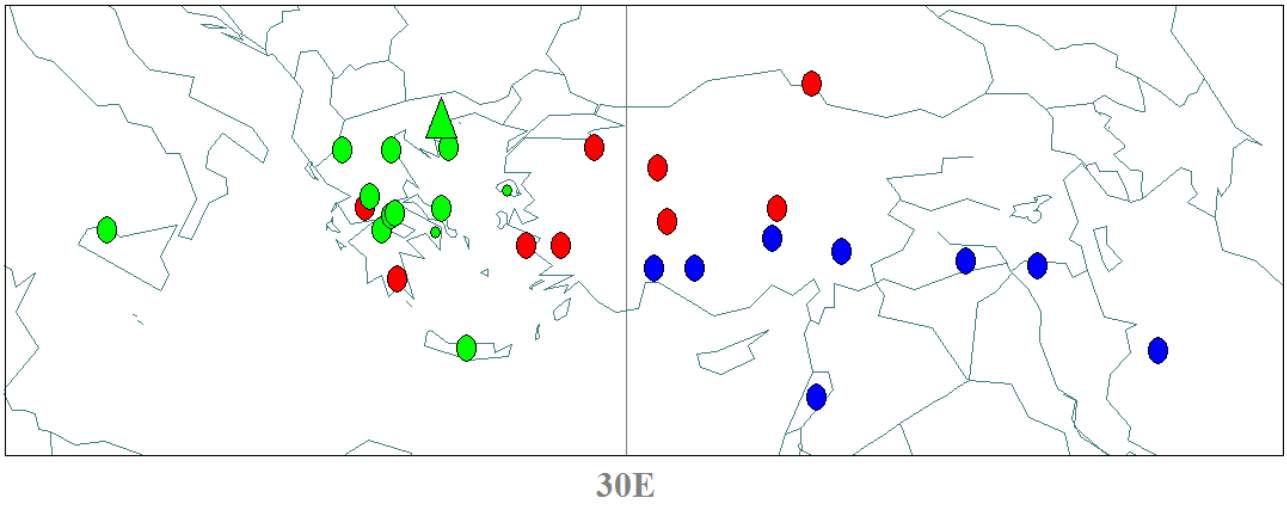


Figure 4.20 Geographic distribution of *Aubrieta* nuclear clades based on the phylogenetic analysis of concatenated rpb2 & duo1 gene regions. Large blue circles = Near East clade; red circles = Anatolian clade with part of Greece; green circles = Aegean coastal/western clade; green triangles = western clade; small green circles from Samos Island and Attica provence.

Geographically, the heterozygous accessions are mainly centred on the Aegean basin, but with some extensions eastwards into Anatolia and, to a lesser extent, westwards into central Greece (Fig. 4.21).

The explanation for this distribution is complex. There is no evidence for any recent hybridisation between the Greek and Anatolian genepools to create a hybrid Aegean centre. What seems to have happened is that the Aegean basin taxa, which form a clade in both the plastid and nuclear sequence phylogenetic trees, have inter-crossed among themselves and with neighbouring Greek and Anatolian populations. This inference is based on the evidence relating to the clustering behaviour of the respective alleles.

This association between the Aegean basin and frequent hybridisation contrasts with the situations to the west and to the east where hybridisation is rare or non-existent. This topic is taken up again in Chapter 5.

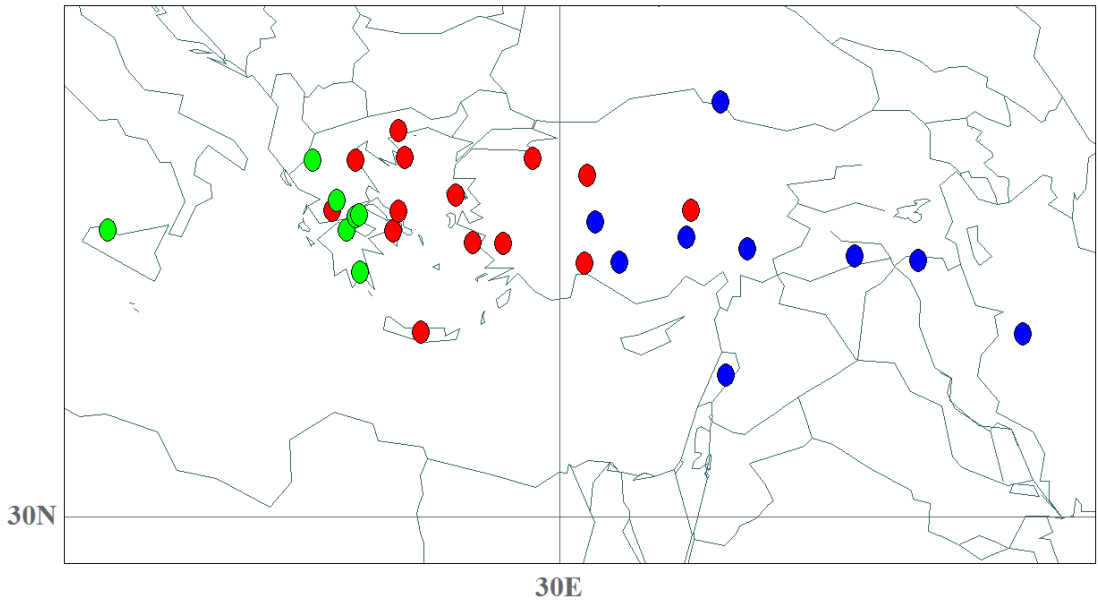


Fig. 4.21 Map showing the distribution of heterozygotic populations (in red) and non- heterozygotic populations (western in green, eastern in blue).

CHAPTER 5. CHROMOSOMAL EVOLUTION

5.1 Introduction

All species of *Aubrieta* that have been studied to date have a base chromosome number of x = 8 (Table 5.1; Warwick and Al-Shehbaz 2006). All are diploid, 2n = 16, except for one report of tetraploidy, 2n=32, in *A. scardica* (Franzen and Gustavsson 1983; Alden in GUSTAVSSON 1986). No information is available regarding karyotypes or meiotic behaviour.

Table 5.1. Published chromosome numbers of *Aubrieta* taxa.

Taxon	Chromosome number	Locality & Voucher	Reference
Aubrieta deltoidea	2n=16	unknown	Dvorak and Ddakova (1984a)
		Greece: Crete	Montmollin (1982)
		Greece: Crete	Montmollin *et al.* (1986)
		Greece: Menikion	Strid in Gustavsson (1986)
		Germany: Jena botanic garden	Koch *et al.* (1999)
var. cithaeronea	2n=16	Greece: Boeotia, Mt. Cithaeron; *Phitos 8802*	Phitos (1970)
var. deltiodea	2n=16	Greece: Peloponnesus, Mt. Erymanthos; *Phitos 7028, 7057*	Phitos (1970)
var. sporadum	2n=16	Greece: Sporades Is., Skiathos; *Phitos 360*	Phitos (1970)
var. graeca	2n=16	Greece: Attiki, Penteli; *Phitos s.n.*	Phitos (1970)
Aubrieta parviflora	n=8	Iraq: Bekhal IAS; *Al-Shehbaz & Al-Omar 76136*	Al-Shehbaz and Al-Omar (1982)
		Iraq: 17 km E of Jindiyan IAS; *Al-Shehbaz & Al-Omar 76119*	Al-Shehbaz and Al-Omar (1983)
		Iran	Aryavand (1975a)
		Iran	Kupfer (1980)
Aubrieta gracilis	2n=16	Greece: Mt Timfristos, 2 km NNE of Karpenision, 1750 m; *Gustavsson 1521* Greece: Mt Timfristos, 6 km ENE of Karpenision, 1500 m; *Gustavsson 1687*	Franzen and Gustavsson (1983)
		Greece: Mt Vardousia, S km WNW of AthanasiosDia- kos, 2000 m; *Gustavsson 3641*	Franzen and Gustavsson (1983)
Aubrieta scardica	2n=16	Bulgaria: Pirin Mt., near Vihren chalet, 2000 m; *Ancev A5268*	Ancev (1978)

Table 5.1 cont'd

Taxon	Chromosome number	Locality & Voucher	Reference
Aubrieta scardica	2n=32	Greece: Mt Timfristos, 4 km N of Karpenision, 1900 m, *Gustavsson 1590, 1593*	Franzen and Gustavsson (1983)
		Greece: Mt Timfristos, 4 km NNW of Karpenision, 2050m, *Gustavsson 4148*	
		Greece: Loupata, *Alden 1690*. Katarrachias, *Alden 5233*. Korouna, *Alden 2005*. Between Boustagani and Boutai, *Alden 3249*.	Alden in Gustavsson (1986)
Aubrieta intermedia	2n=16	Greece: Phocis, Mt. Parnassus; *Phitos 8556*	Phitos (1970)
		Greece: Phocis, Mt. Parnassus, 1500 m; *Van Loon & Oudemans 24527*	Van Loon and Oudemans (1982)
Aubrieta glabrescens	2n=16	Greece: Ipiros, Smolikas; *Phitos 235* (as *A. gracilis* ssp. *scardica* but identification later queried by Gustavsson (1986))	Phitos (1970); Gustavsson (1986)
Aubrieta erubescens	2n=16	Greece: Mt Athos [no voucher cited]	Papanicolaou in Gustavsson (1986)
Aubrieta scyria	2n=16	Greece: Sporades Is., Skyros; *Phitos 767*	Phitos (1970)
Aubrieta columnae subsp. pirinica	2n=16	Bulgaria: Slavjanka Mt., Ambardere above Paril, 1100 m; *Ancev A780*	Ancev (1978)
Aubrieta columnae subsp. *bulgarica*	2n=16	Bulgaria: Znepole region, Zemenska Planina; *Ancev A748 & 7161*	Ancev (1978)
Aubrieta ekimii	2n=16	Turkey, Kocaeli, Yuvacık, Soğuksu, Menekşe yaylası patikası, Şahin kaya; *Yüzbaşıoğlu 3207*	Yüzbaşıoğlu *et al.* (2015)
		Turkey, Serindere; *Yüzbaşıoğlu 3972*	

5.1.1 The plant nuclear genome

Nuclear genome size is variable from one plant species to another. It ranges from 63—49000 million base pairs (Mbp) in the double DNA helix (Bennett and Leitch 2011). Shranz *et al.* (2012) suggested that the variation in the nuclear genome belongs to a continuous multiplication inside the genome itself which mostly involved repetitive DNA sequences.

These repeats are key to an understanding of plant genomic functions and structures (Bennet and Smith 1976; Walbot and Golderg 1979; Preisler and Thompson 1981; Flavell 1980, 1982; Bennett and Leitch 2011). Repetitive DNAs inside the genome might be changed quickly during evolution, leading to genome divergence and speciation. Moreover, these repeats have a direct impact on chromosome packaging, recombination, segregation and behaviour.

Within the genome, the DNA sequences are present in the form of genes (exons and introns), gene regulators, high or low copy number repetitive motifs. The size of sequence motif varies from 2-10,000 bp (Heslop-Harrison and Schmidt 2012). Nuclear genomic DNA sequences are divided into three main groups (Fig. 5.1): 1) dispersed repeats including transposable elements, 2) structural components of chromosomes, and 3) tandem repeat, including 45S and 5S rRNA.

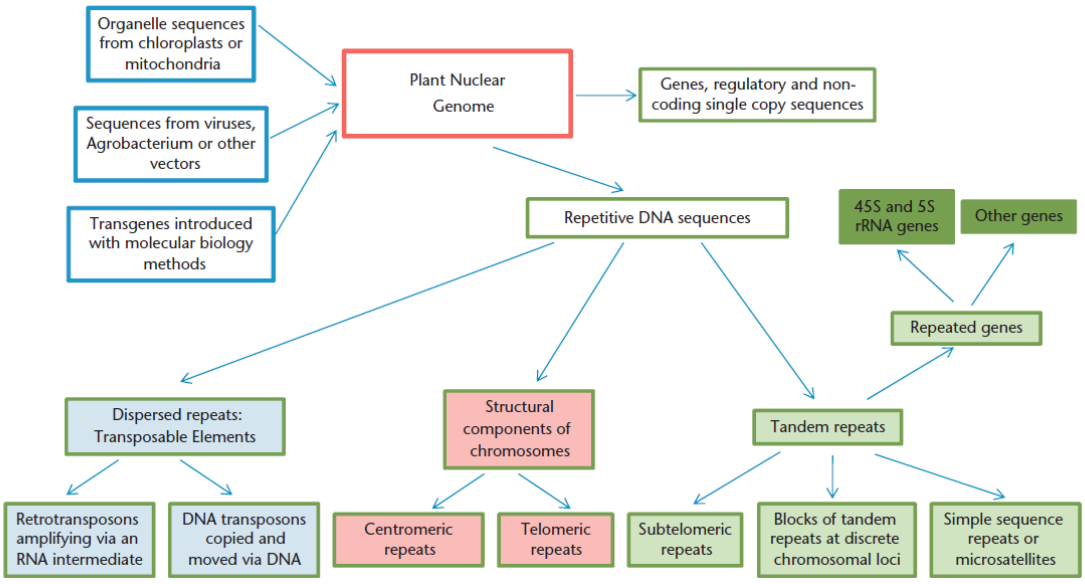


Figure 5.1 Major DNA components of the plant nuclear genome and their relationships. Figure taken from Heslop-Harrison and Schmidt (2012).

In plants, ribosomal DNA (rRNA or rDNA) genes are represented by many hundreds or thousands of copies and they are tandemly organised at one or more loci on the chromosome (Rogers *et al.* 1987; Maluszynska and Heslop-Harrison 1993; Lafontaine and Tollervey 2001; Garcia *et al.* 2012). rDNA comprises four genes (5S, 5.8S, 18S and 26S) and two spacers (internal transcribed spacer I and II (ITS1 and ITS2)). Together, these regions are also known as 45S rDNA. These genes are transcribed in two different orders, based on polymerase I and III activities, 18S-5.8S-26S and 5S respectively (Garcia *et al.* 2012).

Repetitive DNA sequences, in particular rDNA, are conserved within the genome by a process conventionally referred to as concerted evolution. Furthermore, due to the sequence homogeneity, these repeats can slightly evolve "in concert" after a series of gene conversions and random crossings over during recombination (Garrido-Ramos 2012). Although theoretical studies have been reported that demonstrate concerted evolution mechanisms (Arnheim 1983; Ohta and Dover 1983; Slatkin 1986), few studies have been able to test these hypothesises empirically, but mention may be made of *Gossypium* (Wendel *et al*. 1995); *Paeonia* (Zhang and Sang 1999); and *Lilium* (Chacon *et al.* 2012).

5.1.2 Chromosome karyotype and rearrangement

Each individual organism has its own somatic chromosome number, which is commonly constant within a species and may even be the same in closely related species. It tends to vary, however, in more distantly related taxa, e.g. genera and families (Appels *et al.* 1998). The chromosome complement can be described on the basis of what is called the karyotype. This is the physical appearance of the genome and reflects fundamental features such as the number and size of chromosomes, position of centromeres, arm ratios, secondary constrictions and other morphological features. The typical pattern is usually represented by a diagram called an ideogram.

In the early 1970s, researchers started to augment karyotypes with data relating to the staining properties of the constituent DNA. For instance, C- and G-banding (Dobel *et al.* 1973; Vosa 1973, 1976, 1977; Marks and Schweizer 1974); Q-banding (Sharma 1975, 1977), and Quinacrine Mustard (QM) (Modest and Sengupta 1973) and Fluorescent *In Situ* Hybridization (FISH) in different plants such as *Aegilops, Arabidopsis, Hordeum* and Trifolium (Castilo and Harrison 1995; Dadaeva *et al.* 1996; Fransz *et al.* 1998; Taketa *et al.* 1999; Ansari *et al.* 1999) respectively.

The fact that the karyotype varies between plant species (White 1978; Grant 1981) suggests that the variation accumulates by means of chromosomal rearrangements during the process of speciation (White 1978; King 1993). Furthermore, chromosomes may become rearranged in different ways. This mostly happens when chromosomes are separated or reconnected during meiosis, resulting in segments being deleted, duplicated, inverted or translocated between non-homologous chromosomes (Rieseberg 2001). Grant (1981) proposed that each rearranged

segment has a unique pattern, which may result in re-patterning during pairing at meiosis and this re-patterning can lead to speciation.

Many models for chromosome speciation have been reported last few years, as reviewed in King (1993), Rieseberg (2001) and Levin (2002). Chromosome re-patterning occurs in the context of geographical isolation, strong under-dominance, genetic drift and adaptation (Lewis 1966; White 1978; Grant 1981; Templeton 1981). For example, chromosome structure rearrangement has been found by detecting the repetitive DNA sequences in heterochromatin-euchromatin borders (Badaeva *et al.* 2007).

Heterochromatin comprises of different groups of tandem repeats (Sharma and Raina 2005) and transposable elements (TEs), mainly Class I retrotransposons (Lippman *et al.* 2004). Raskina *et al.* (2008) mentioned that the TEs have a direct impact on DNA and genome composition. For instance, the movement of rDNA clusters between homologous and non-homologous chromosomes is carried out by these elements. Moreover, Dubkovsky and Dvorak (1995) demonstrated that the rDNA sites in Triticeae genomes may move without undergoing chromosome rearrangement. Similarly, Schubert and Wobus (1995) showed that the movement of these sites is attributable to TE activity.

5.1.3 In situ hybridisation

In the late of 1980s, the chromosomal in situ hybridization technique was invented to map DNA sequences physically as well as to distinguish a specific repetitive DNA sequences in chromosomes, which were otherwise difficult to deal with by conventional cytogenetic techniques (Schwarzacher *et al.* 1989). The technique is based on hybridising DNA-probes to the karyotype and visualising the hybridisation sites by means of a UV-sensitive fluorescent dye.

Several studies were conducted in *Brassica* species. For instance, Hallden *et al.* (1987) identified the tandem repeat sequences of *Brassica* species. Similarly, Lakshimikumaran and Ranade (1990) focused their study on the highly repetitive DNA sequence characteristics of *B. campestris*. Maluzynska and Heslop-Harrison (1991) underlined *Arabidopsis thaliana* chromosome number and identified a particular repetitive gene in the chromosome. Two years later they localized the rDNA in *Arabidopsis* chromosomes. Maluszynska and Heslop-Harrison (1993) used rDNA probes to construct a physical map of rDNA loci in different

Brassica species in interphase nuclei and metaphase chromosomes. Furthermore, *Arabidopsis arenosa* repetitive DNA was analysed by Kamm *et al.* (1995).

Alix and Heslop-Harrison (2004) identified the diversity of reteroelements in *Brassica* species and Alix *et al.* (2005) determined the organisation of retroelements in *Brassica oleracea*. Hasterok *et al.* (2005) examined the *Brassica* genomes by using fluorescence in situ hybridization (FISH) and genomic in situ hybridization (GISH). Subsequently, the *Brassica juncea* 45S and 5S rDNA genes were determined by Maluszynska and Hasterok (2005) using pTa71 and pTa 794 probes.

5.1.4 Aims

The aims of this part of the study were:

- To establish karyotypes where possible and localise the rDNA sites;
- To interpret karyotype variation in terms of chromosome morphology, and in terms of the number and position of rDNA loci, in the context of the evolutionary tree, thereby gaining insight into the nature of chromosome change accompanying speciation;

In this study I investigate what chromosomal changes have accompanied speciation in the genus, including study of the nature and mechanisms of repetitive DNA evolution in the rDNA genes.

5.2 Materials and methods

5.2.1 Materials

Accessions used for cytological study are shown in Table 5.2. A total of 17 taxa was investigated. Plants were grown in a medium of 3 parts potting compost to 1 part grit in an unheated greenhouse.

5.2.2 Chromosome preparation

1-2cm white young root tips were collected from a seedling and pre-treated with 8-hydroxyquinoline for 24 hours at 4^0C and then directly transferred to Carnoy fixative (3:1 absolute ethanol : glacial acidic acid). Roots were washed with distilled water for 10 minutes and then 1X enzyme buffer diluted from 10X (100mM citric Acid and 100mM Tri-Sodium Citrate) for 10 minutes followed by digestion in an enzyme mix ((3% pectinase (Sigma, 450 units ml^{-1}), 1.8% cellulose (Calbiochem, 4000 units gm^{-1}) and 0.2% cellulose (Onozuka RS, 500 units gm^{-1})) for 40-60 minutes at 37^0C. Roots were then transferred to 1X enzyme buffer. One or two roots were put on a glass microscope slide and squashed in 45% or 60% glacial acidic acid as described by Schwarzacher and Heslop-Harrison (2000).

5.2.3 DNA probes and probe labelling

Two different types of probe were used to identify 45S and 5S rDNA sites in the chromosomes of *Aubrieta* species; pTa71 probes, contains 9 kb *Eco*R1 fragment isolated from *Triticum aestivum* as a repeat of 18S-5.8S-26S rDNA (Gerlach and Bedbrook 1979) and cloned in pUC19. pTa794 probes, contains 410kb of *Bam*H1 fragment isolated from embryo of *T. aestivum* as a repeat of 5S rDNA. 3 µl of pTa71 and pTa794 added to each sample of 1.5ml micro centrifuge tube (Eppendorf) separately and suspended by distilled water to 19µl. 16µl of 2.5X Random Primer Solution added to each tube then incubated in water bath for 5 minutes at 95^0C and transferred directly to ice for 5 minutes. 3µl of 10X dUTP Nucleotide Mixed with 1.8µl biotin-16-dUTP (50 nmol) and digoxigenin-11-dUTP (125 nmol), 0.8µl Exo-Klenow Fragment was added to the mixture and then followed by short spin down and incubated for 2 hours at 37^0C. Finally, 5µl of stop buffer was added to the each sample. NucleoSpin, Gel and PCR Clean-up kit (MACHEREY-NAGEL was used to purify the probe samples.

5.2.4 Fluorescence in situ hybridisation (FISH) technique

Previous prepared slides were pre-treated by ethanol/ acidic acid 3:1 for 20 minutes then absolute ethanol 5 minutes twice followed by RNase pre-treatment (100 µg/µl) in 2X SSC for 60 minutes at 37^0C, washed and dehydrated by 2X SSC and ethanol respectively. In clear Eppendorf tube, probes master mix were prepared on ice; 20µl 100% Formamide, 4µl 20xssc, 8µl Dextran sulphate, 1µl Salmon sperm DNA 1 µg/µl, 0.5µl 100mM EDTA (ethylene-diamine-tetra acetic acid), 0.5µl of 10%SDS (sodium dodecyl sulphate), 2µl of Bio and Dig probes then suspended by water to 40µl to each slide. Chromosomes and probes were denatured in hybridization chamber for 7 minutes at 70^0C then reduced to 37^0C for 15-20 hours. Hybridized slides were washed by low stringency stringent solution using 1X SSC at 42^0C. Molecular probes (2.0mg/ml streptavidin conjugated to Alexa595) and 4mg/ml of antidigoxigenin conjugated to FITC used to detect hybridization sites in the examined chromosomes. 100µl of DAPI solution (4mg/ml in Mcllvaine's buffer) was added to each slide then slides rinsed with detection buffer (100ml 20x ssc,1ml Tween 20, distil water) and mounted by 24*32 mm slides with one drop of antifade solution. Nikon Eclipse 80i epifluoroscence microscope with Nikon camera used to scan and taking pictures (Schwarzacher and Heslop-Harrison 2000).

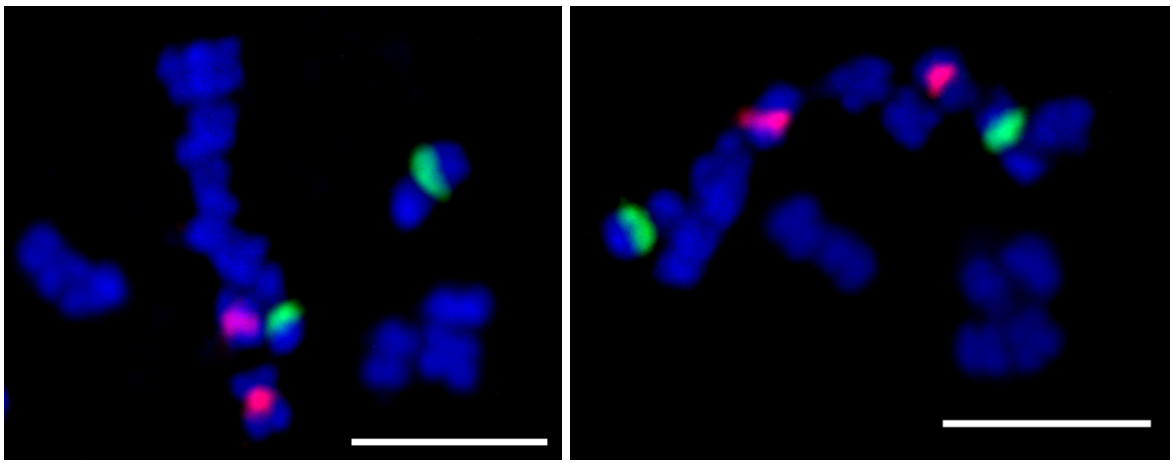
5.3 Results

Chromosome numbers were established for all of the taxa that were studied; all had a sporophytic number of 2n=16 (Table 5.2). The number of rDNA sites of fluorescent in-situ hybridisation was four, six, seven or eight (Table 5.2). Of these, 5S rDNA consistently showed two sites in all taxa examined. These two sites are always in an interstitial position in *Aubrieta*, but vary among taxa according to whether they are on a long or short arm of the chromsoome. Variation also occurs among the taxa with respect to the number of 45S rDNA sites (Table 5.2). *Aubrieta parviflora* (S77) shows only two sites. *A. gracilis* (S1), *A. columnae* (BG4 & BG9), *A. croatica* (BG13), and *A. deltoidea* (BG6 & BG14) have four sites. *A. hesperidiflora* (BG3), *A. erubescens* (BG8) and *A. intermedia* (S16) have five sites. *A. anamasica* (BG12), *A. intermedia* (S2b, S5b, BG7), *A. macedonica* (BG5) and *A. scardica* (BG2) have six sites. The locations of the rDNA sites in each of the taxa will now be discussed in turn.

Table 5.2 Chromosome numbers and details of the numbers and positions of the rDNA sites among the studied taxa.

Code	Taxon	Chrom. no.	rDNA sites		Total sites	45S+5S sites	45S rDNA sites			5S rDNA sites	
			45S	5S			centro-mere	short arm	long arm	short arm	long arm
S77	*A. parviflora*	2n=16	2	2	4	0	0	0	2	2	0
BG6	*A. deltoidea*	2n=16	4	2	6	1	0	2	2	0	2
BG14	*A. deltoidea*	2n=16	4	2	6	1	0	2	2	0	2
BG1	*A. gracilis*	2n=16	4	2	6	0	0	2	2	0	2
BG4	*A. columnae*	2n=16	4	2	6	0	0	2	2	0	2
BG9	*A. columnae*	2n=16	4	2	6	0	0	2	2	0	2
S10e	*A. scyria*	2n=16	4	2	6	0	0	2	2	0	2
BG13	*A. croatica*	2n=16	4	2	6	0	0	2	2	0	2
BG3	*A. hesperidiflora*	2n=16	5	2	7	1	1	2	2	0	2
BG8	*A. erubescens*	2n=16	5	2	7	1	1	2	2	0	2
BG16	*A. intermedia*	2n=16	5	2	7	2	1	2	2	2	0
S5	*A. macedonica*	2n=16	6	2	8	1	2	2	2	0	2
BG2	*A. scardica*	2n=16	6	2	8	2	2	2	2	2	0
S7	*A. intermedia*	2n=16	6	2	8	2	2	2	2	2	0
S2b	*A. intermedia*	2n=16	6	2	8	2	2	2	2	0	2
S5b	*A. intermedia*	2n=16	6	2	8	2	2	2	2	0	2
BG12	*A. anamasica*	2n=16	6	2	8	2	2	2	2	0	2

In *Aubrieta parviflora*, the single pair of 45S rDNA sites is located in a proximal position, adjacent the centromere, on the long arm of one of the sub-metacentric chromosome pairs (Fig. 5.2). The two 5S sites are located proximally on a short arm of a different sub-metacentric chromosome pair (Fig. 5.2).



A **B**

Figure 5.2 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes of ***Aubrieta parviflora***. A & B) Physical location of the two 45S rDNA sites, detected using a pTa71 probe labelled with digoxigenin (in green) and of the two 5S rDNA sites detected using a pTa794 probe labelled with biotin (in red). Scale bars = 5µm.

All the taxa with four 45S rDNA sites, viz. *Aubrieta gracilis*, *A. croatica*, *A. columnae* (BG4 & BG9) and *A. scyria* showed the same rDNA karyotype (Figs. 5.3, 5.4). Thus, two of the sites are in a proximal position on a short arm adjacent to the centromere and the two others are in a proximal position on a long arm adjacent to the centromere. Occasionally in *A. gracilis*, one pair of 45S signals was dispersed over a chromosome arm rather than being condensed in one position. This is interpreted as a possible artefact caused by some aspect of the procedure. In all four accessions, the two 5S rDNA sites are on different chromosomes from the 45S sites. They are proximal on the long-arm of a sub-metacentric pair (Figs. 5.3 & 5.4).

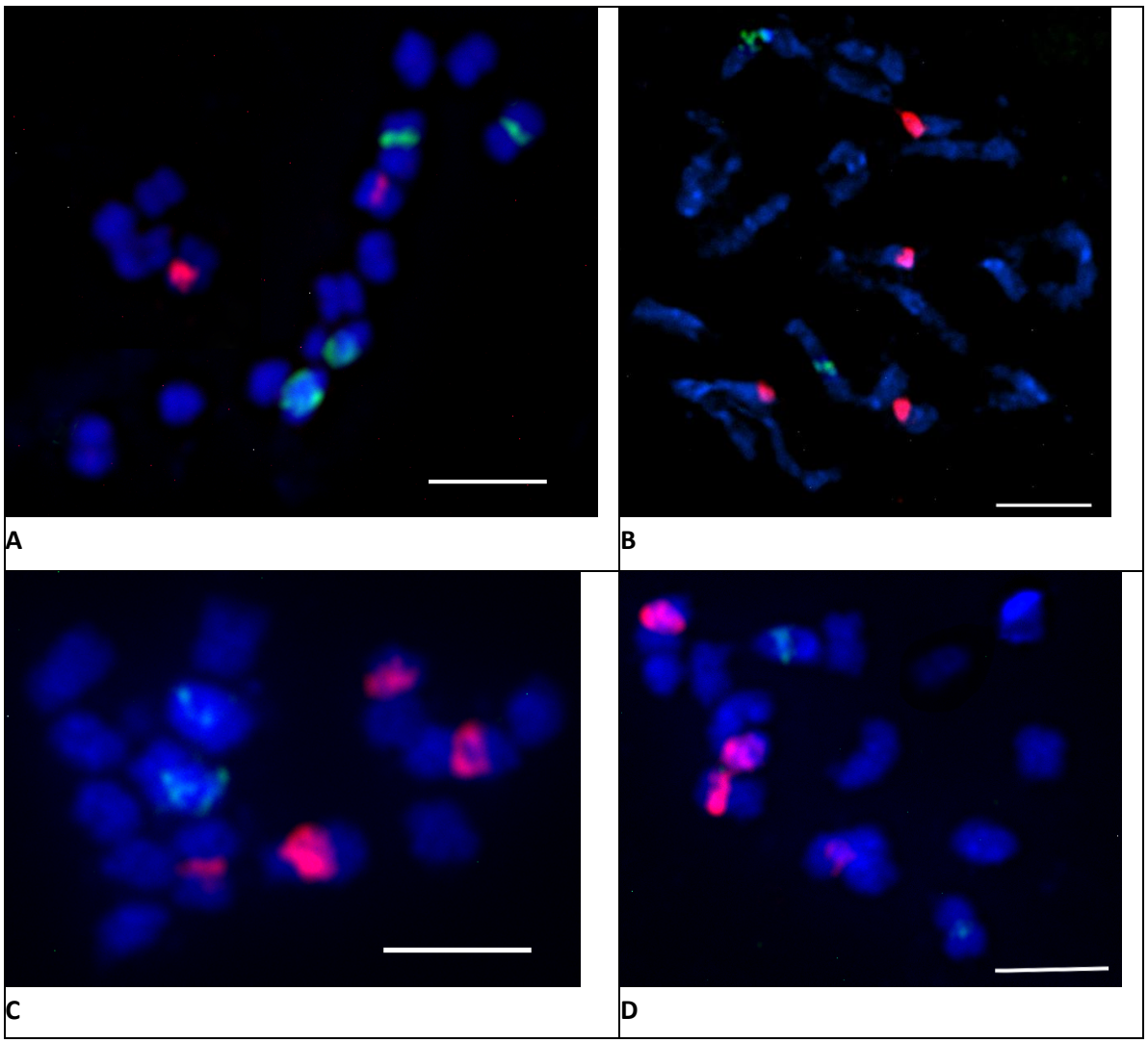


Figure 5.3 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes. A) ***Aubrieta gracilis*** (BG1): physical location of the four 45S rDNA sites, detected using a pTa71 probe labelled with digoxigenin (in green) and of the two 5S rDNA sites detected using a pTa794 probe labelled with biotin (in red). B) ***Aubrieta gracilis*** (S1): physical location of the four 45S rDNA sites, detected using a pTa794 probe labelled with biotin (in red), and of the two 5S rDNA sites, detected using a pTa71 probe labelled with digoxigenin (in green). C & D) ***Aubrieta columnae*** (BG4): physical location of the four 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green). Scale bars = 5µm.

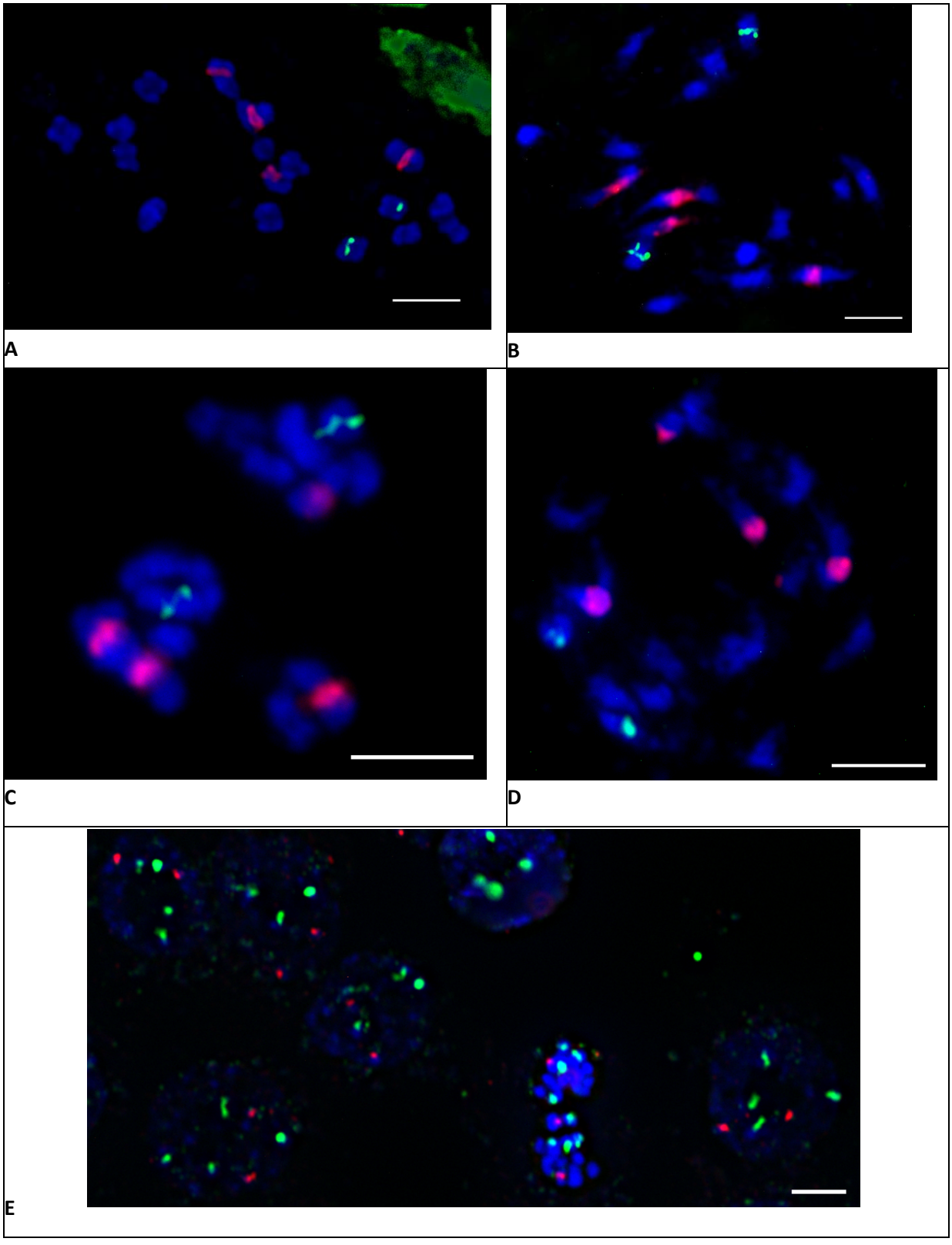


Figure 5.4 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes. Physical location of the four 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green). A & B) ***Aubrieta croatica*** (BG13). C & D) ***Aubrieta columnae*** (BG9). E) ***A. scyria*** (several cells shown). Scale bars = 5µm.

The chromosomes of *A. scyria* are the smallest of all the taxa I examined so it is difficult to be precise about the locations of the rDNA sites. They appear, however, to be as described above. Interphase preparations show that the four 45S rDNA sites are often stretched, peripheral to the nucleolus, unlike the two 5S rDNA sites (Fig. 5.4E).

In the accessions of *A. deltoidea* (BG6 & BG14) there are also four 45S rDNA sites and two 5S sites, but in these taxa, instead of the two types of rDNA being on different chromosomes, one of the 5S sites is on the same chromosome arm as one of the 45S sites, and immediately adjacent to it (Fig. 5.5). Two of the 45S sites are located proximally on the short arm of a submetacentric chromosome pair, and the other two are place similarly but on a long arm. The 5S sites are located proximally on a long arm of a submetacentric chromosome pair.

Six accessions had six 45S rDNA sites. These were *A. anamasica* (BG12), *A. intermedia* (S2b, S5b, BG7), *A. macedonica* (BG5) and *A. scardica* (BG2) (Figs. 5.6-5.9). The 45S rDNA karyotype here comprises two sites in a centromeric position, two proximally on the short arm of a chromosome pair and two proximally on a long arm. In four taxa *A. anamasica, A. intermedia* (S2b, S5b) and *A. macedonica*, the two 5S rDNA sites are located on the same long chromosome arms as two of the 45S sites (Figs. 5.6-5.8). In the other two taxa, *A.intermedia* (BG7) and *A.scardica*, the pair of 5S sites is located on the same short chromosome arms as two of the 45S sites (Fig. 5.9).

In *Aubrieta macedonica* interphase preparations show the 45S rDNA running around the nucleolus. At both interphase and prophase stages, the chromosomes show a small region of 45S rDNA satellites (Fig. 5.10).

Some preparations of *A. anamasica* and *A. intermedia* were good enough to establish morphological as well as rDNA karyotypes (Figs. 5.7 & 5.8). In *A. anamasica*, there is one long pair, five medium-sized pairs and two short pairs of metacentric to sub-metacentric chromosomes; the long pair has a secondary constriction at the nucleolar-organiser region (NOR) (Fig. 5.7). *Aubrieta intermedia* has six pairs of medium chromosomes and two pairs of smaller metacentric to sub-metacentric chromosomes. In both cases the rDNA loci are located on the medium chromosomes.

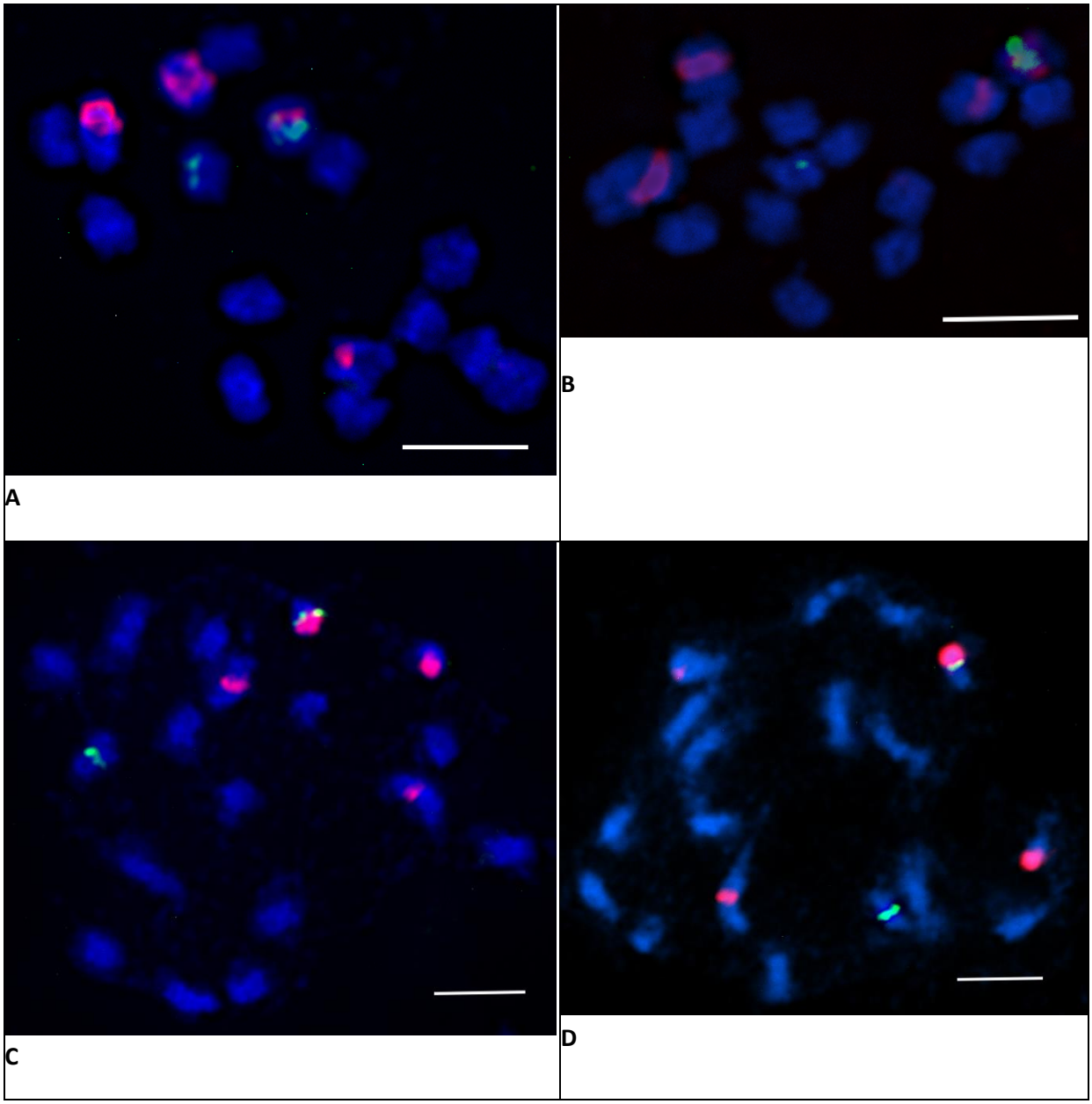


Figure 5.5 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes. Physical location of the four 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green). A & B) ***Aubrieta deltoidea*** (BG6). C & D) ***Aubrieta deltoidea*** (BG14). Scale bars = 5μm.

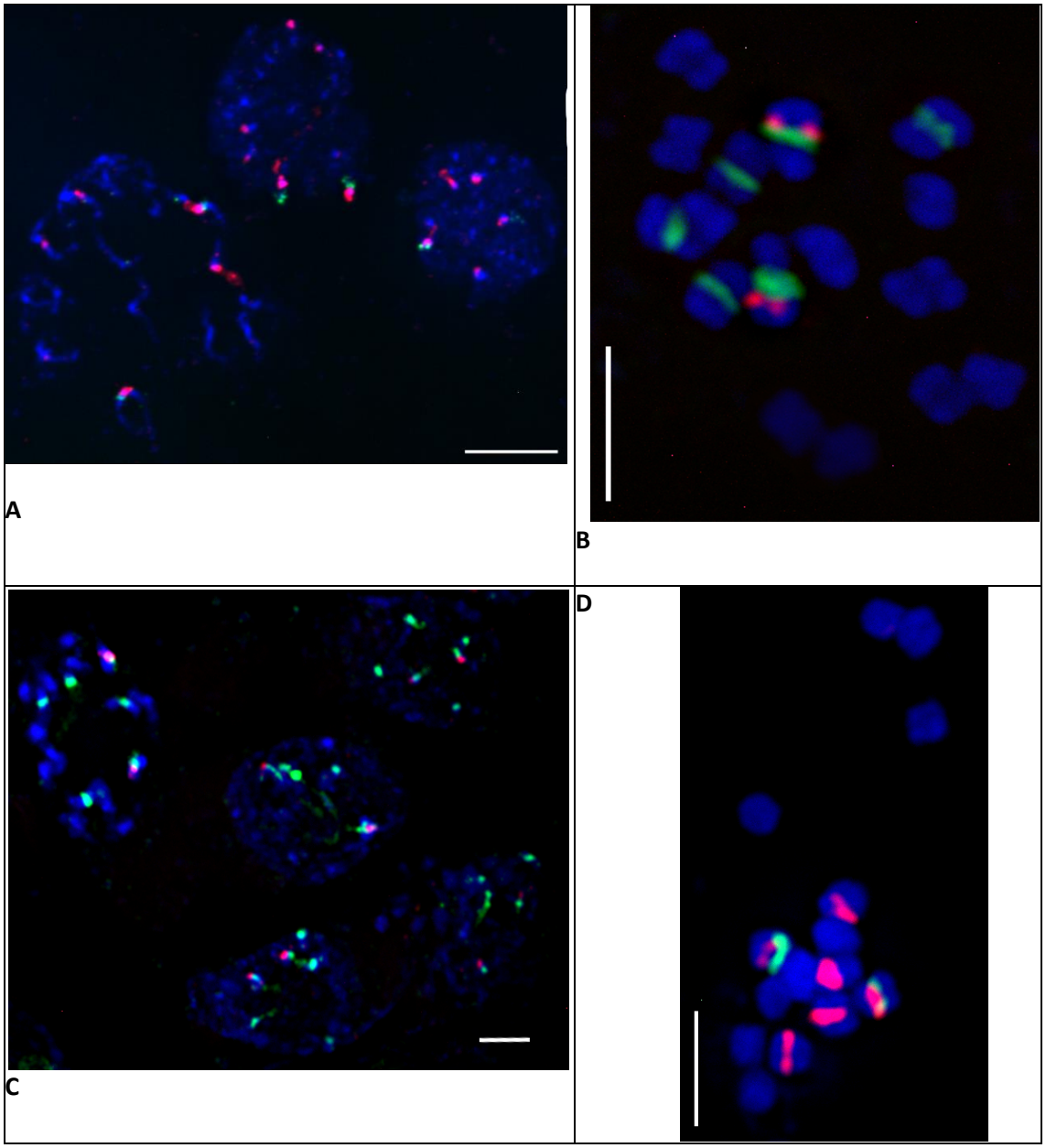
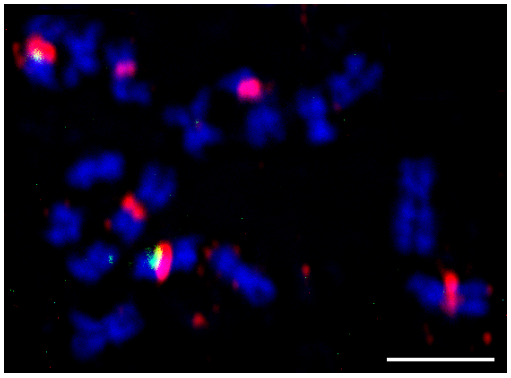
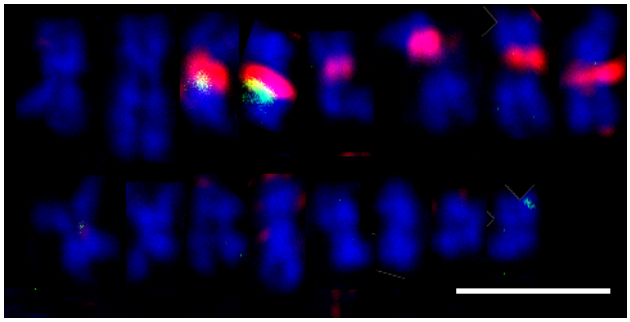


Figure 5.6 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes. A) ***Aubrieta anamasica***. Physical location of the six 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green). B) ***Aubrieta intermedia* (S2b)**. Physical location of the six 45S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green) and of the two 5S rDNA sites, detected using a pTa71 probe labelled with biotin (in red). C) ***Aubrieta macedonica*** (several cells). Physical location of the six 45S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green) and of the two 5S rDNA sites, detected using a pTa71 probe labelled with biotin (in red). D) ***Aubrieta macedonica*** (one cell). Physical location of the six 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green). Scale bars = 5µm.

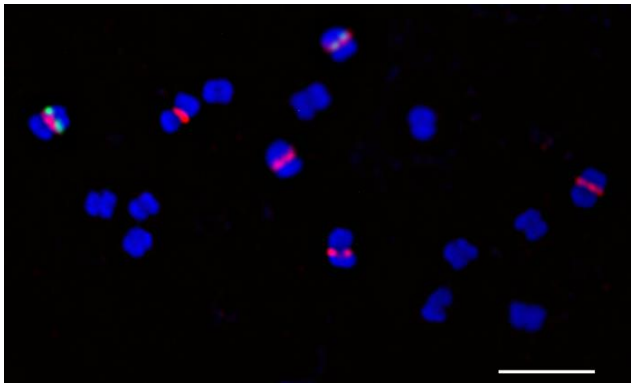


whole cell

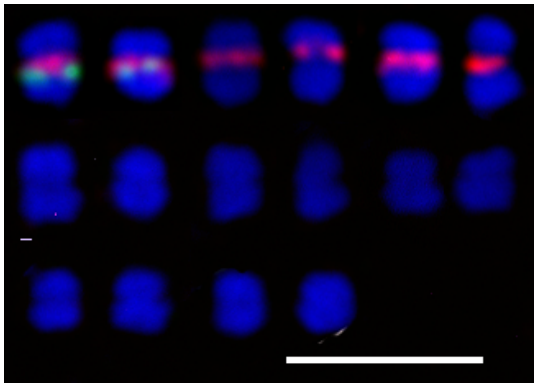


karyotype

Figure 5.7 Karyotype of ***Aubrieta anamasica***, showing a long pair of chromosomes with a secondary constriction, five medium pairs and two short pairs or metacentric to submetacentric chromosomes. Three of the medium sized pairs contain the 45S and 5S rDNA loci.



whole cell



karyotype

Figure 5.8 Karyotype of ***Aubrieta intermedia*** **(S5b)**, showing six pairs of medium chromosomes and two pairs of smaller metacentric to submetacentric chromosomes. Three of the medium sized pairs contain the 45S and 5S rDNA loci.

Finally, three taxa showed five 45S rDNA sites along with the two 5S sites. They are: *A. hesperidiflora, A. erubescens* and *A. intermedia* var. *macedonica*. The five 45S sites comprise one in a centromeric position, two proximal on a short arm and two proximal on a long arm of a sub-metcentric to metacentric chromosome. In *A. hersperidiflora*, the two 5S sites are on the long arm of a chromosome pair, whereas in *A. erubescens* and *A. intermedia* var. *macedonica* they are on a short arm.

Interphase preparations confirm that the two 5S sites (red) and five 45S sites (green) run around the nucleolus to form compact peripheral-on-nucleolus sites (Fig. 5.10).

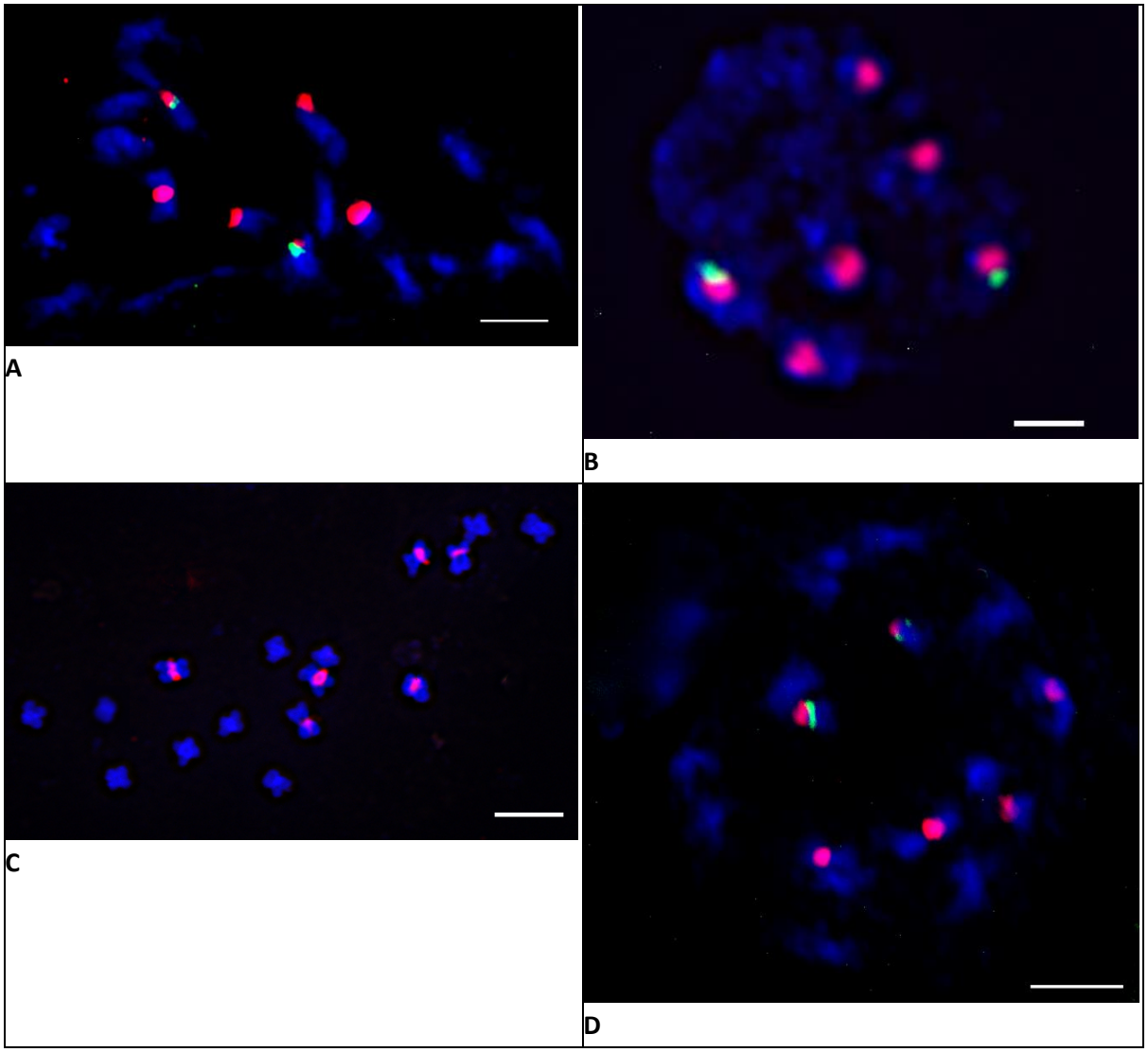


Figure 5.9 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes. Physical location of the six 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green). A & B) ***Aubrieta intermedia*** (BG7). C & D) ***A. scardica*** (BG2). Scale bars = 5µm.

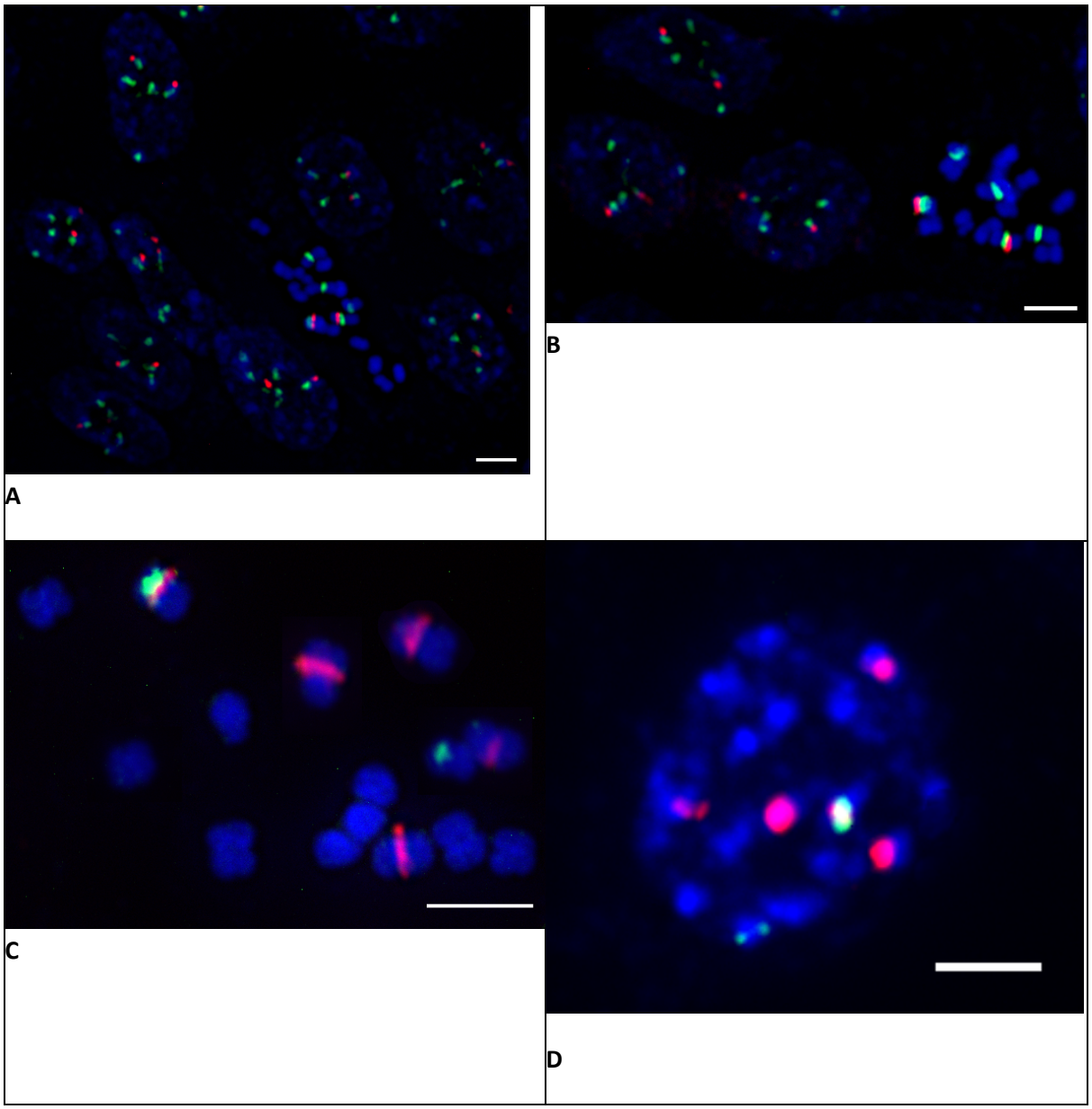


Figure 5.10 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes. A&B) ***Aubrieta intermedia*** (BG16). Physical location of the five 45S rDNA sites, detected using a pTa71probe labelled with digoxigenin (in green) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with biotin (in red). C&D) ***Aubrieta erubescens*** (BG8). Physical location of the five 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green) Scale bars = 5µm.

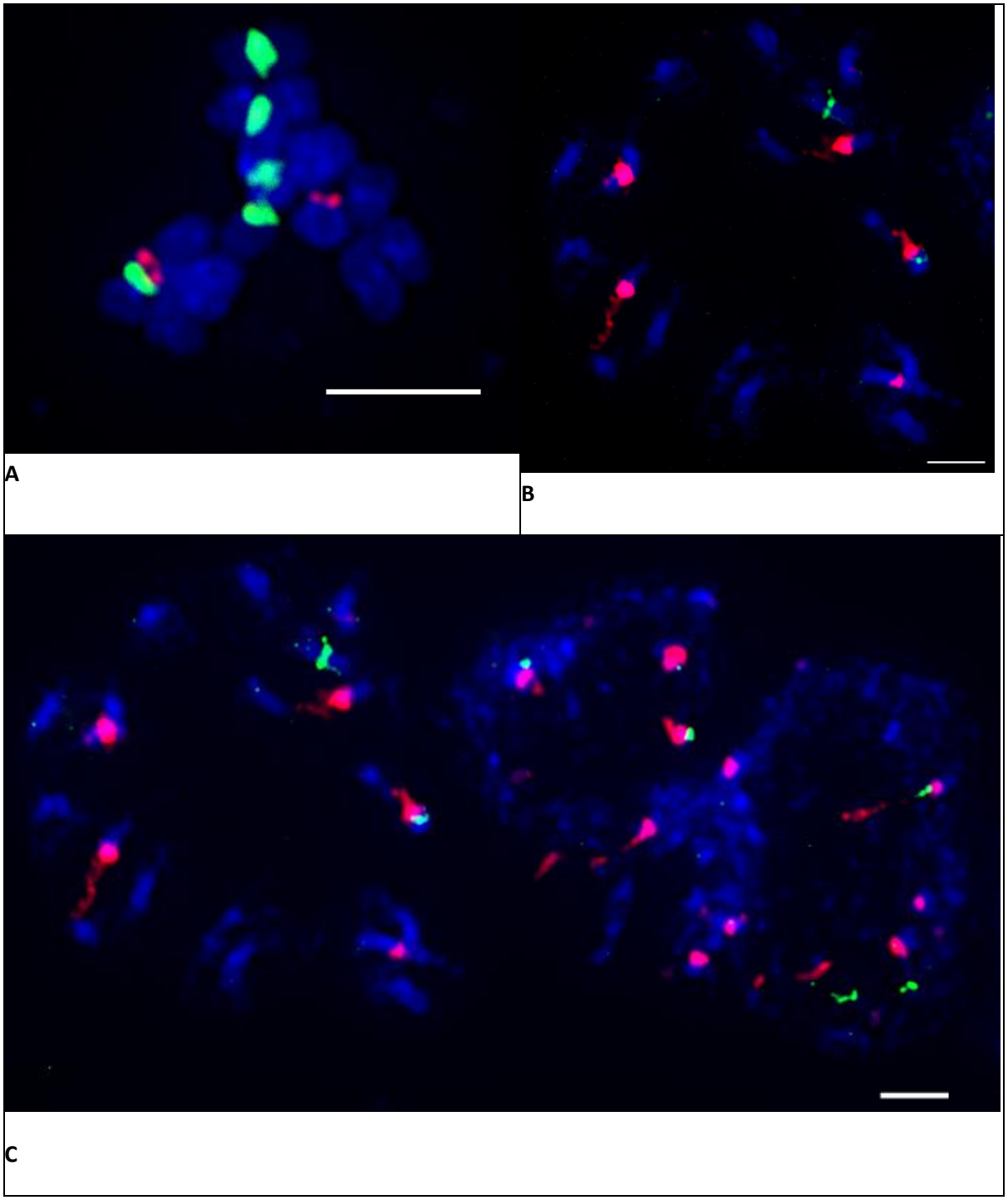


Figure 5.11 Fluorescent *in situ* hybridization (FISH) of metaphase root tip chromosomes of ***Aubrieta hesperidiflora*** (BG3). A) Physical location of the five 45S rDNA sites, detected using a pTa71probe labelled with digoxigenin (in green) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with biotin (in red). B&C) Physical location of the five 45S rDNA sites, detected using a pTa71 probe labelled with biotin (in red) and of the two 5S rDNA sites, detected using a pTa794 probe labelled with digoxigenin (in green) Scale bars = 5µm.

5.4 Discussion

5.4.1 Chromosome number

The genus *Aubrieta* appears to be uniform in its chromosome number, with all taxa studied to date displaying a sporophytic number of 2n=16. The findings reported here thus correspond to those recorded in the literature (Tables 5.1, 5.2). There are, however, reports of tetraploid counts, 2n=32, being made from five populations of *A. scardica* (Franzen and Gustavsson 1983; Alden in Gustavsson 1986; see Table 5.1).

5.4.2 Morphological karyotypes

Aubrieta chromosomes are uniform, small in size and metacentric to sub-metacentric with few if any secondary constructions or conspicuous NORs. As a consequence, morphological recognition of individual chromosomes is very difficult. Nevertheless, I managed to construct karyotypes for two taxa: *Aubrieta anamasica* (BG12) and *A. intermedia* (S5b). These are the first karyotypes to be established for any species in the genus *Aubrieta*.

5.4.3 rDNA karyotypes

Ribosomal rDNA is considered to be an effective marker that facilitates recognition of the chromosomal karyotype (Guerra 2012). It is also regarded as a good guide to indicate chromosomal re-arrangements that may occur during speciation (Castilho and Heslop-Harrison 1995). A total of seven rDNA karyotypes was recovered from the accessions studied (Fig. 5.12).

The physical location and number of rDNA sites vary among *Aubrieta* species. Regarding the number of rDNA sites, species of *Aubrieta* had a pair of 5S sites and 4—6 45S sites. In the Brassicaceae as a whole, the number of 5S sites ranges from two to 14, and of 45S sites from two to 16 (Ali *et al.* 2005). Within the tribe Arabideae, *Arabis alpina* (2n=16) and *Arabis glabra* (2n=12) are the only other species studied apart from the results presented here. Both these species are reported to show two 5S and four 45S rDNA sites (Ali *et al.* 2005). *Aubrieta* thus appears to be very similar to the closely related *Arabis* in these respects.

The location of the 5S rDNA sites in all accession studied is interstitial, whereas the 45S rDNA sites are either interstitial or centromeric. The position of the 5S sites matches the pattern found elsewhere in the Brassicaceae, where in 19 of the 45 taxa (including *Arabis alpina*) that have

been studied, reviewed by Ali *et al.* (2005), the 5S sites are interstitial and in the rest of the taxa (including *Arabis glabra*) they are terminal or sub-terminal. The centromeric positions of one or two of the 45S sites in *Aubrieta* is, however, unusual in the Brassicaceae where eight species (including *Arabis alpina*) have interstitial 45S rDNA loci and in the rest (including *Arabis glabra*) they are terminal (Ali *et al*. 2005).

In terms of location on chromosome arms, those accessions with four 45S sites have two in a proximal position on a short arm and two proximally on a long arm. Those with five sites have the extra site at the centromere. Those with six sites have both the extra sites at centromeres. With regard to the two 5S rDNA sites, both are either on a short arm or on a long arm, depending on the clade to which the accession belongs (see later discussion of phylogeny). In some cases, a chromosome arm has two rDNA sites, a 5S and a 45S, and again the details depend on the clade to which the accession belongs (see later discussion of phylogeny).

5.4.4 Phylogeny

The rDNA characters that vary are: number of 45S sites, the location of the 45S sites, the location of the 5S sites and the number of chromosomes that share a 5S and a 45S site. In order to understand better the evolution of the different characters, the character states were plotted on a phylogenetic tree constructed using sequence data from the nuclear duo1 and rpb2 genes (Fig. 5.12).

5.4.4.1 Number and location of 5S rDNA sites

The outgroup used for the nuclear phylogenetic tree was *Arabidopsis thaliana* (tribe Camelineae) and this species has four to six 5S rDNA sites (Ali *et al.* 2005). Within the tribe Arabideae, however, *Arabis alpina* has two 5S rDNA sites (Ali *et al.* 2005), like all accessions of *Aubrieta* studied here. On this basis, it seems reasonable to conclude that this is the ancestral number in *Aubrieta*.

Mapping the data regarding location of the 5S sites onto the duo1 and rpb2 tree (Fig. 5.12), shows that the position of the sites is variable, with some accessions having them on the short arm (subclade A1, clade B, some clade C) and some on the long arm (subclade A2, some clade C). Since Ali *et al.* (2005) said only that the sites in *Arabis alpina* were interstitial, it is not possible with the available evidence to conclude anything about the ancestral location. It seems likely that the difference between the two states is caused by a translocation. The translocation seems to be a frequent one, judging by the appearance of the rDNA karyotypes which show

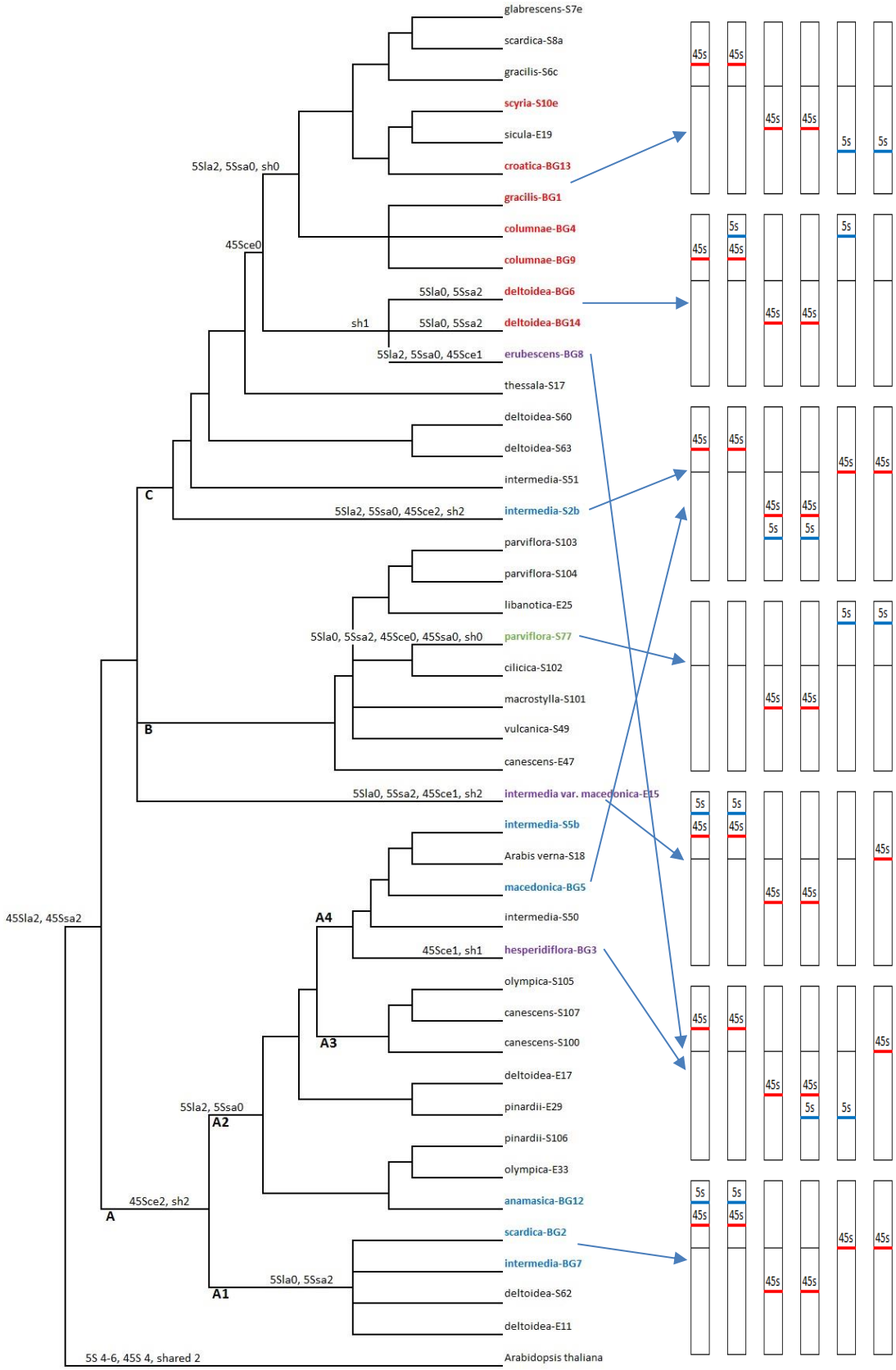


Figure 5.12 Phylogenetic tree, derived from Bayesian analysis of concatenated duo1 and rpb2 sequences, on to which are plotted the number of rDNA sites (taxon name blue = 8; purple = 7; red = 6; green = 4 sites), the number on a long chromosome arm (la), on a short arm (sa), and centromeric (ce) and the number of chromosomes bearing a pair of 5S and 45S sites (sh). rDNA karyotypes shown on the right.

that the 5S sites have moved around quite considerably during the course of evolution, in contrast to the 45S sites whose positions appear to be conserved.

5.4.4.2 Number and location of 45S rDNA sites

Both the outgroup *Arabidopsis thaliana* (tribe Camelineae) and *Arabis alpina* (tribe Arabideae) have four 45S rDNA sites (Ali *et al.* 2005), and on this basis, it seems reasonable to conclude that this is the ancestral number in *Aubrieta*.

When the number and locations of the 45S rDNA sites are plotted onto the phylogenetic tree (Fig. 5.12), it can be seen that the position of the 45S sites is invariant; only their number changes. Thus clade A contains accessions showing five or six 45S rDNA sites, with two located proximally on a short arm and two proximally on a long arm and the extra one or two being centromeric. Duplications would account for the occurrence of the additional centromeric sites. The case of the 5-site in *A. hesperidiflora* is indicative of a hybrid origin and will be discussed later.

In clade B, the only accession studied cytologically showed two 45S rDNA sites, both on a long arm, so here there is apparently a deletion of the two short arm sites. Clade C comprises accessions showing four, five or six sites, located on short-arm, centromere and long-arm. Thus the duplication producing the 6-site taxa has probably occurred at least twice. As before, the 5-site accession (*A. erubescens*, BG8) is likely a consequence of hybridisation and is discussed later. Similarly, *A. intermedia* var. *macedonica* (E15) also has five 45S sites and can be explained as a hybrid (see next section).

5.4.4.3 Chromosomes that contain both 5S and 45S rDNA sites

Chromosomes that have both 5S and 45S rDNA sites are known elsewhere in the Brassicaceae. Thus many genera have been shown to share a 5S and a 45S site on a pair of chromosomes (i.e. two chromosomes, each with a 5S plus a 45S site), including the related *Arabis alpina*; some genera even have four or six chromosomes with shared rDNA sites (Ali *et al*. 2005). Within *Aubrieta*, accessions had zero, one or two chromosomes each containing both a 5S and a 45S site. The most parsimonious explanation is that the ancestral condition is represented by two chromosomes each with a 5S and a 45S site (Fig. 5.12). If this is the case, then the taxa that either do not have any rDNA sites that share a chromosome, or have these but only on one chromosome, have experienced translocations.

Taxa with only one chromosome possessing a 5S and a 45S site (*A. hesperidiflora* BG3, *A. deltoidea* BG6 &14 and *A. erubescens* BG8) are putative hybrids as mentioned earlier and are discussed in the next section.

5.4.4.4 Hybridisation

The rDNA karyotype data support a hybrid origin for those taxa that i) have five, rather than four or six, 45S sites; and/or ii) have only a single chromosome, rather than none or two, on which a 5S/45S pair of rDNA sites is located. These taxa will now be discussed in turn.

The putative rDNA karyotype parentage of *A. erubescens* and *A. hesperidiflora* is the same (Fig. 5.13). In the former case, *A. erubescens* appears to be derived from a cross between *A. intermedia*, with six 45S rDNA sites (two on the same arm as two 5S sites), and *A. gracilis* or a close relative, with four sites, none of them on the same chromosome as a 5S site. The hybrid shows the expected five 45S sites, one of them on the same chromosome arm as a 5S site (Fig. 5.13). In the case of *A. hesperidiflora*, this taxon also appears to be derived from a cross involving *A. intermedia*, with six 45S rDNA sites (Fig. 5.13); judging from the phylogenetic tree, the other parent is likely to be an Anatolian species from clade A whose karyotype has yet to be studied but is likely to be similar to that found in *A. gracilis*.

The putative parentage of *A. intermedia* var. *macedonica* (E15) can be deduced by similar analysis. Thus, a taxon from clade A1, with six 45S sites, two of them sharing chromosome arms with two 5S sites, is suggested to have hybridised with a heterozygous population of *A. deltoidea* (e.g. BG 6 & 14), in which the respective 5S and 45S sites are inherited as shown in Fig. 5.13.

The heterozygous accessions of *A. deltoidea* are themselves likely to be of hybrid origin from parents yet to be investigated – a sampling issue.

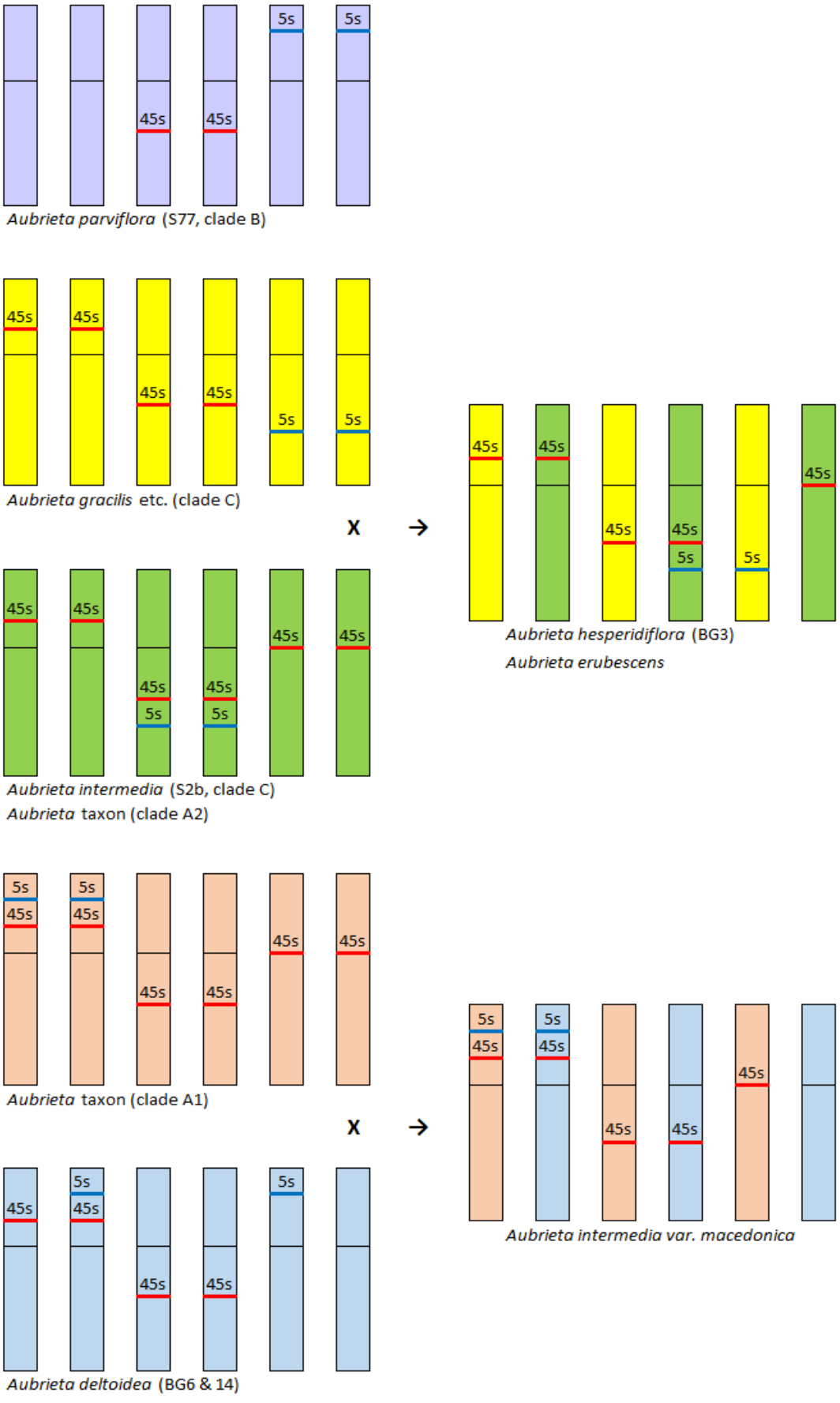


Figure 5.13 Putative rDNA karyotype relationships.

CHAPTER 6 PHYLOGEOGRAPHY OF *AUBRIETA*

6.1 Introduction

Many researchers have studied the biogeography of the Greece / Aegean / Anatolian area but these studies have mainly focused on animals, such as bats, turtles, rodents, Albinaria snails, Palearctic frogs and lacertid lizards (Michaux *et al*. 2004; Bilgin *et al*. 2006; Dubey *et al*. 2007; Gunduz *et al*. 2007; Flanders *et al*. 2009; Furman *et al*. 2009; Fritz *et al*. 2009; Douris *et al*. 1995 and 1998, Beerli *et al*. 1996, Poulakakis *et al*. 2003). There have been fewer studies of plants, and these mostly of trees or shrubs, such *Fagus spp*., *Pinus sylvestris* and *Hippophae rhamnoides* (Bartish *et al*. 2006; Gomory *et al*. 2007; Naydenov *et al*. 2007). Herbaceous species have been relatively neglected, although mention should be made of the comprehensive study on *Nigella* by Comes *et al*. (2008). It is in this context that the present study of Aubrieta is made. First I provide an overview of the geological history of the Aegean Basin and of Anatolia.

6.2 Geological history of the Aegean Basin

The geological history of the Aegean Basin is complex and incompletely known. The basin forms part of the continental shelf extending westwards from Anatolia and eastwards from Greece. Originally, it was a single land-mass, known as Aegeis. A variety of processes, comprising tectonic, volcanic and eustatic events, meant that from about 13—10 mya onwards Aegeis began to fragment, resulting in the formation of various islands and archipelagos (Dermitzakis 1990a). These comprise six main groups: Ionian Islands, Argo-Saronic Islands, the Cyclades, the Sporades, the north-east Aegean islands and the Dodecanese (Fig. 6.1).

First to become isolated as an archipelago was Crete, which formed a group of islands in the Upper Miocene, ca 10 mya (Dermitzakis 1990a). Shortly afterwards, the mid-Aegean trench formed (ca 9 mya), which caused the separation of the eastern islands, including some of the Dodecanese, and Asia Minor from the Cyclades and mainland Greece (Dermitzakis and Papanikolaou 1981). Karpathos, the second largest island of the Dodecanese, was isolated by ca 8 mya, although the isolation was intermittent and it probably retained a continental climate until the Pliocene (Meulenkamp 1971). Indeed, there was a land-bridge between Karpathos, Rhodes and Asia Minor at the beginning of Pliocene (Daams and Van de Weerd 1980), which

was broken in the mid- to late-Pliocene to render Karpathos isolated again (Kuss 1975). Rhodes finally became an island in the late Pliocene or early Pleistocene (Boger and Dermitzakis 1985).

Figure 6.1. Land-masses, archipelagos and islands of the Aegean basin.

The Messinian Salinity Crisis (5.96—5.33 mya), in which the whole Mediterranean Basin more or less dried up (Hsu 1972; Hsu *et al.* 1977; Krijgsman *et al.* 1999), obviously led to the reinstatement of land connections between the various island groups, although the deep canyons in the dried basin and the undoubtedly severe climatic and edaphic conditions presumably acted as a barrier to the dispersal of many plants and animals. Following the re-opening of the Straits of Gibraltar, the Mediterranean Basin refilled, taking about 1000 years to do so. Crete became permanently isolated (ca 5 mya). The Peloponnese Peninsula was separated from northern Greece and the island of Kythira through much of the Pliocene (4—3 mya) (Dermitzakis 1990a). At about this time and until the Upper Pleistocene, the southern part of the Cyclades archipelago formed an island, separated by a seaway from the northern Cyclades (Anastasakis and Dermitzakis 1990; Chatzimanolis *et al.* 2003). The eustatic cycles of the Pleistocene caused the breakage of land-bridges between the Cyclades islands, and between Alonnisos and northern Greece, Kythira and Antikythira. In contrast, the land connections between mainland Greece and the Peloponnese, and between the islands of the Dodecanese and the adjacent coast of Anatolia, were retained until the beginning of the Holocene.

In fact, recent studies have suggested that the degree of continuity between land masses during the Pleistocene has been underestimated (Lykousis 2009; Tourloukis and Karkanas 2012). Thus emergence of formerly submerged land did not only occur during glacial stages but also occurred during early inter-glacials, notably those of marine isotope stage (MIS) 9 and 11, and to a lesser extent MIS 7. Whilst the Aegean Sea persisted in much of the south, more than half was subaerially exposed during successive glacial and inter-glacial stages MIS 8, 9, 10, 11 and 12 (Fig. 6.2). Together, the sedimentological and biostratigraphic data indicate that the central Aegean would have been exposed land even before ca 500 Ka (Lykousis 2009). The exposure of about half of the Aegean for most of the early Pleistocene would have had a significant impact not only on opportunities for biotic dispersal (landbridge) but also on the availability of land for colonisation and establishment.

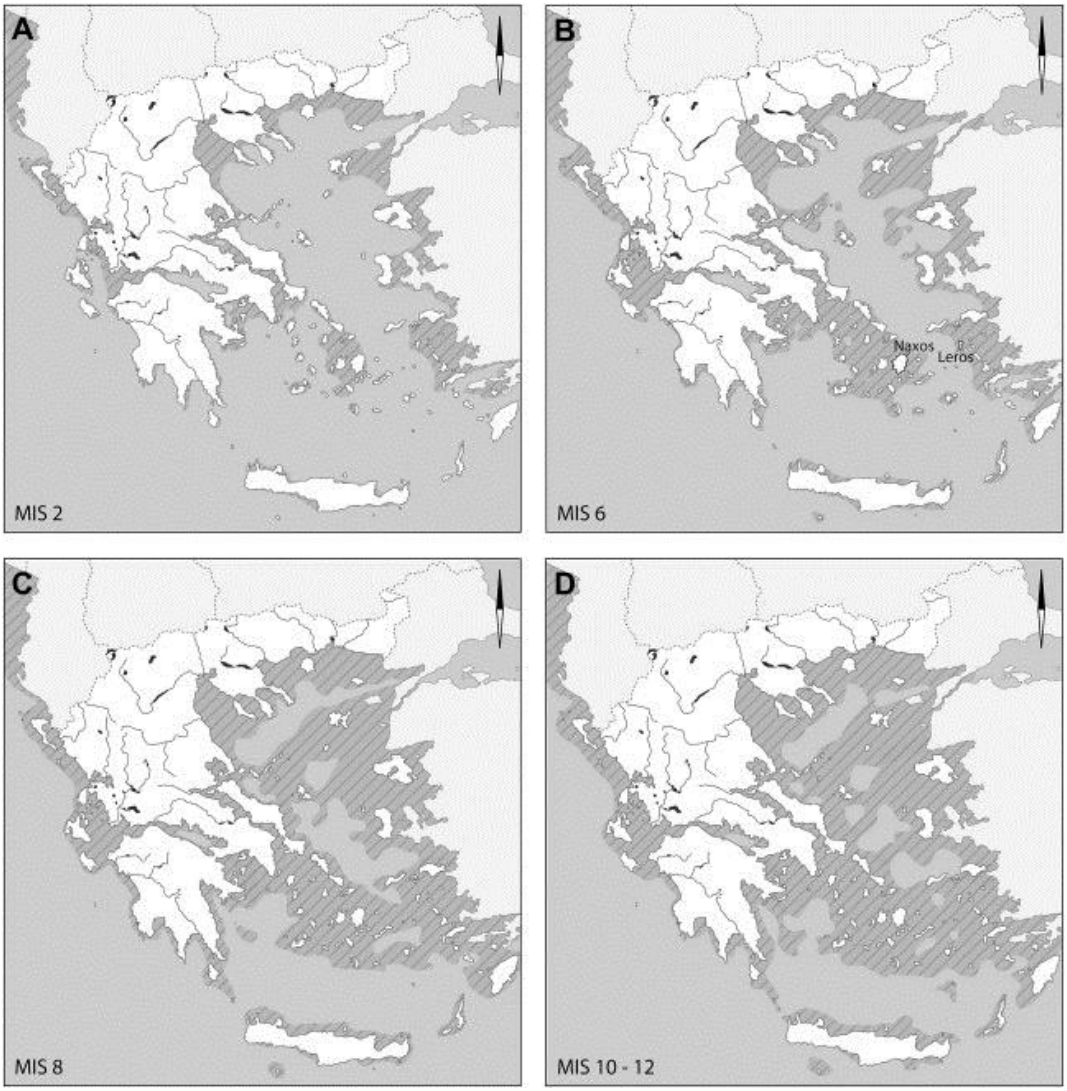


Figure 6.2 The Aegean Basin during successive glacial and inter-glacial stages (A: MIS2 = ca 11 Kya; B: MIS6 = 191 Kya; C: MIS8 = 300 Kya; D: MIS10—12 = 374—478 Kya). Hatched areas indicate dry land. Figure taken from Tourloukis and Karkanas (2012).

6.3 Geological history of Anatolia

Located between Europe and Asia, Anatolia is where the European and Turco-Iranian floras meet, overlapping in fact in the central part (Davis 1965). During the Pleistocene, lowland areas developed steppe vegetation (Michaux *et al.* 2004; Magyari *et al.* 2008) whilst the upland regions were less hospitable, developing glaciers on the higher peaks (Atalay 1996). There is considerable evidence that many temperate species survived at least the last glacial maximum, and perhaps earlier ones, in the lowland areas (Rokas *et al.* 2003; Dubey *et al.* 2006; Fritz *et al.* 2009). Connections with Greece via land bridges, especially across the Sea of Marmara and the Bosphorus Strait, were a frequent occurrence in the late Pleistocene, owing to the reduced sea levels (Magyari *et al.* 2008). These connections promoted the dispersal of biota, especially temperate species, in both directions between Europe and Asia (Kučera *et al.* 2006; Dubey *et al.* 2007; Gündüz *et al.* 2007). The only impediment to this biological traffic were the high mountains (>2000m) toward the eastern end of Anatolia, and which constitute the 'Anatolian Diagonal' (Davis 1971; Ekim and Güner 1986; Bilgin *et al.* 2006) and, more recently, the Marmara Sea and adjacent Thyracian Plain (Fig. 6.3). Nevertheless, their rugged and varied topography may well have allowed not only many species to survive there but also many others to speciate, as is evidenced by the high degree of endemism (Davis 1971; Çiplak 2003). As a consequence of its environmental history and its location, therefore, Anatolia is an important area for the understanding of the origin of both the European and the Asiatic floras and faunas.

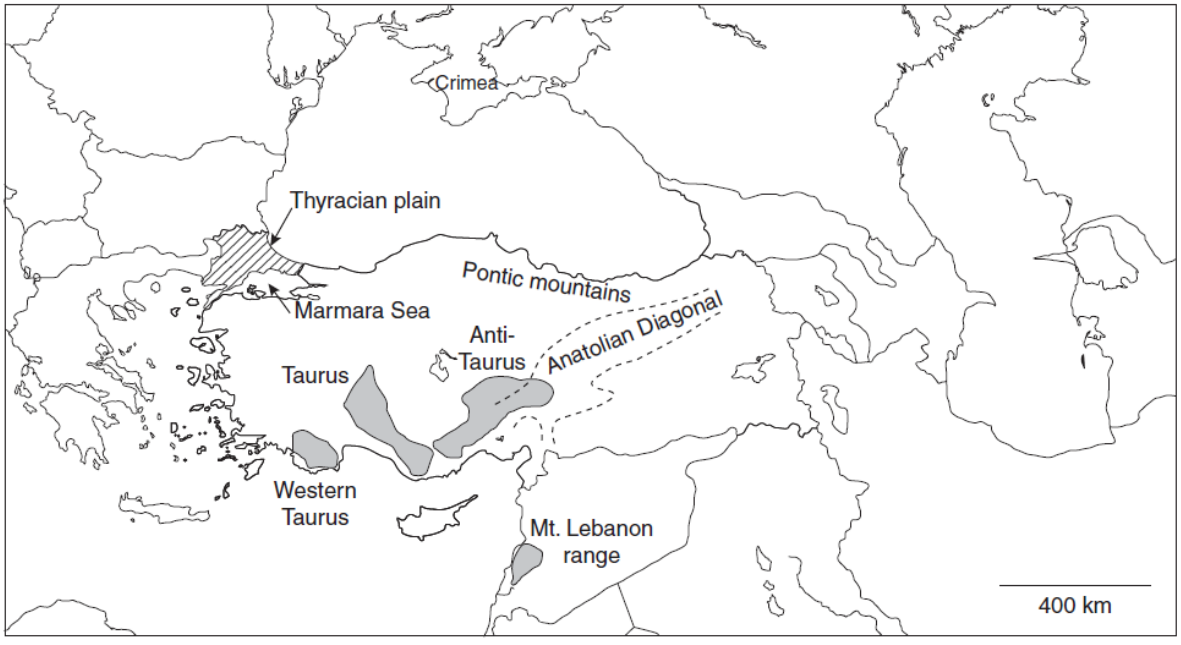


Figure 6.3 Main topographical features of Anatolia. Figure taken from Ansell *et al.* (2011).

6.4 Aims

The principal aims of the phylogeographical study were:

1) to elucidate the phylogeographical history of *Aubrieta*, chiefly by assigning dates on the clades of the phylogenetic trees produced and described in Chapter 4;
2) to interpret the ages and origins of the clades in the context of topography and the geological events described here.

6.5 Methods

The concatenated chloroplast gene sequences were aligned using *Geneious 7* software (Kearse *et al*. 2012) and then saved as a nexus file. The divergence times were estimated using BEAST v1.8.3 and v2 (Drummond *et al*. 2013; Bouckaert *et al*. 2014). The method of dating relied on the suggested date of origin of *Aubrieta*, viz. 5.19 million years ago (Mya), based on plastid trnL-F data, as proposed by Karl and Koch (2013) in their analysis of tribe Arabideae.

In BEAST, the aligned data file was imported to *BEAUti* software. The program parameters were set as follows; 1) date specified as: years before the present; 2) guess the date: everything after the past; 3) specify origin date: 5.19; 4) substitution Model: HKY; 5) site heterogeneity: gamma and invariant site; 6) gamma category: 4; 7) clock type: uncorrelated relaxed clock; 8) relaxed distribution: lognormal; 9) tree prior: speciation: birth-death process.

The *BEAUti* file was created and then saved as XML file. The XML file was imported into *BEAST* with random speed and an automated thread pool size. Then the trees file generated by *BEAST* was annotated by *Tree Annotator* and program parameters set as follows; 1) Burn-in (as states) 100; 2) posterior probability limits 0.5; 3) targeted tree type: maximum clade credibility tree; 4) node heights: median heights. *Fig tree v1.4.2* was used to view the generated tree by *Tree Annotator.*

6.6 Results

The results of the BEAST analysis are shown in Fig. 6.4, where key divergence dates of the various nodes are estimated.

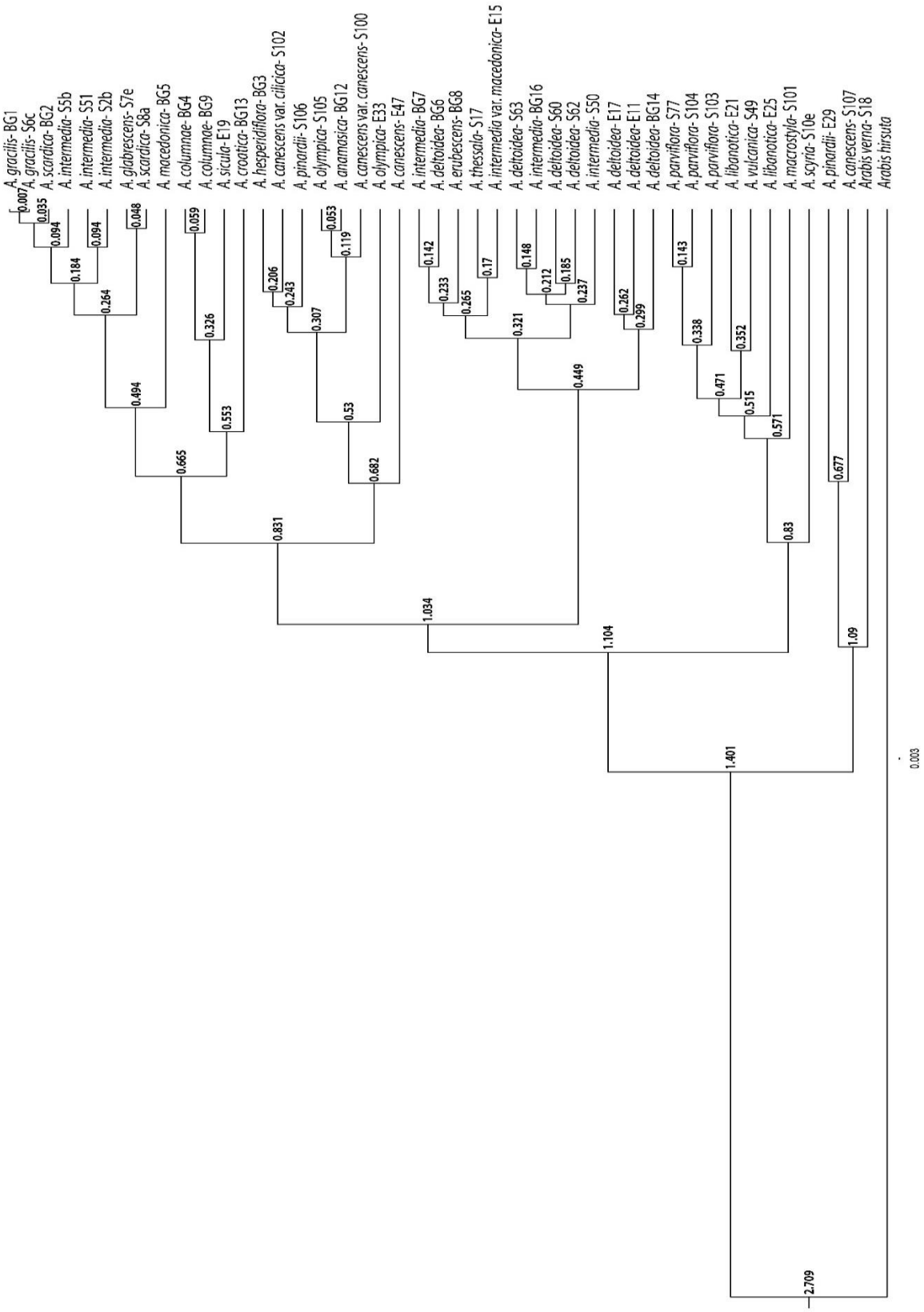


Figure 6.4 Bayesian phylogenetic tree produced by BEAST software, with nodes dated in millions of years.

6.7 Discussion

Hedge (1976) was of the opinion that the origin of tribe Arabideae lies in the Irano-Turanian floristic region, a massive area that runs from Mongolia and Tibet in the east to Iraq and Anatolia in the west. Within this region, more recent studies have proposed that ancestors of the tribe Arabideae grew in and around Anatolia and the Levant (Karl and Koch 2013), giving rise, in mid- to late-Miocene, to the tribe itself in Iran and the Caucasus.

The origin of the genus *Aubrieta* dates to about 5.19 Mya (Karl and Koch 2013) which is more or less contemporaneous with the Messinian Salinity Crisis (5.96 to 5.33 Mya), in which the Mediterranean Basin more or less dried up. It may be presumed that the great environmental changes at this time stimulated the origin of *Aubrieta* s.l., which eventually (1.4 Mya) split into a lineage of annual plants (*Arabis verna*) and a lineage of perennials (*Aubrieta*). The genetic control of the perennial vs annual habit has been shown to be regulated by the locus PEP1 in the perennial *Arabis alpina*, and by its orthologue, FLOWERING LOCUS C (FLC), in the annual *Arabidopsis thaliana* (Wang *et al*. 2009). PEP1 controls the return to vegetative growth by limiting the duration of flowering; it also prevents some shoots from undergoing floral transition, thereby allowing a polycarpic life history, and restricts flowering to the spring season because it is transiently repressed by low winter temperatures. The only difference between this situation and the operation of FLC in the annual *Arabidopsis thaliana* appears to be related to differences in chromatin regulation (Wang *et al*. 2009).

Karl and Koch (2013) suggested that annual taxa of tribe Arabideae showed a preference for drier, lowland habitats, whereas perennials showed a preference for more mesic, montane habitats (Schaffer and Gadgil 1975). The striking contrast between a single annual species and possibly 20 or more perennials in the genus can be linked to the greater topographical and climatic diversity of the montane habitats (Brown 2001; Badgley 2010), which leads to increased rates of diversification in response not only to selection in a wider range of ecological niches but also to reduced gene flow between populations (Rahbek and Graves 2001; Hughes and Eastwood 2006).

Based on the phylogenetic tree of the tribe Arabideae, it can be concluded that the ancestral fruit type has a strongly compressed, long-podded morphology. Taxa with this condition in *Aubrieta* are *Arabis verna* (pan-Mediterranean), *Aubrieta pinardii* (Turkey) and *A. gracilis* and *A. thessala* (Greece). Thus it seems that long-podded ancestors gave rise to the early diverging (ca 1.104 Mya) Near East group whose taxa are characterised by short pods.

These early events appear to have occurred somewhere in Anatolia, given that *Arabis verna* is sister to a couple of Anatolian accessions (Fig. 6.4) and that the next most basal lineage is one containing species which grow in the area occupied by the Anatolian diagonal or points to the east (Near East group). Originating ca 1.1 Mya, the distribution of this lineage is shown in Fig. 6.5, which is based on specimens seen by me together with those listed in major Floras (Cullen 1965; Hedge 1980; Fl. Iranica). An exceptional species in this lineage is *A. scyria*, from the Aegean coast of Greece. Its presence there is evidently a consequence of an early dispersal event (ca 830 Kya). It is hard to suggest a likely vector, but presumably birds might have been involved.

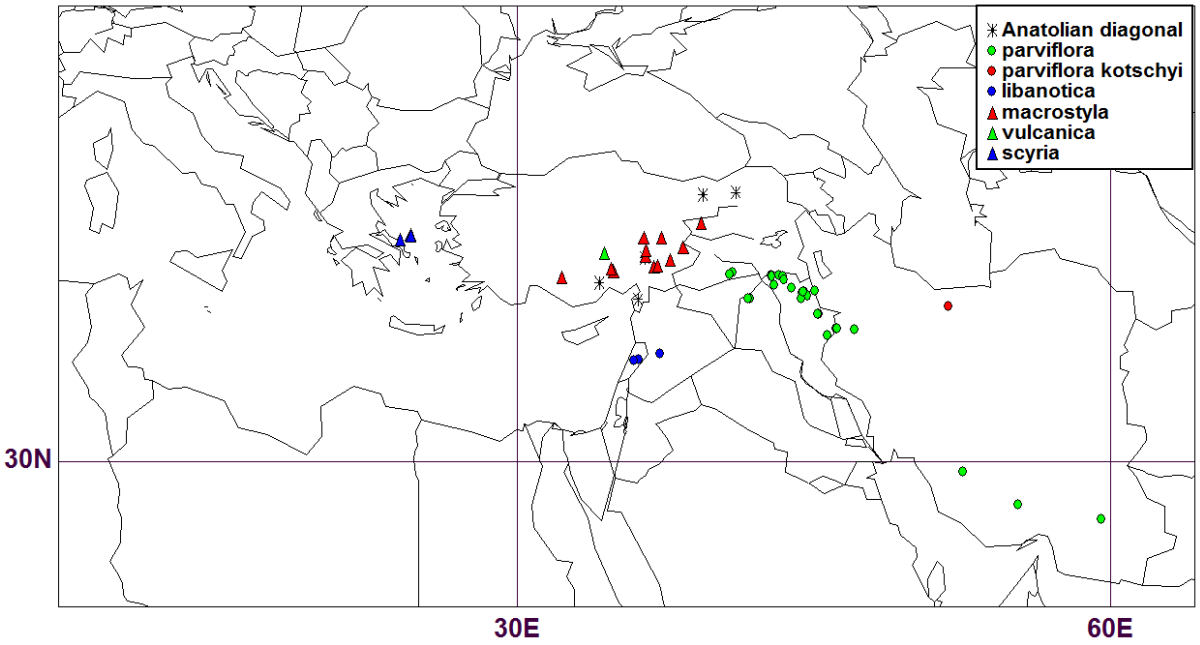


Figure 6.5 Geographical distribution of members of the Near East clade.

The next lineage to evolve (ca 1.034 Mya) is that which currently occupies the Aegean Basin, i.e. the western coast of mainland Greece, Crete and the western edge of Anatolia (Fig. 6.6). It sits rather puzzlingly in between the Anatolian and the Near East lineages and its mode of origin is obscure. It interrupts the otherwise step-by-step westwards geographical spread of the genus and may be a result of a dispersal event from the Anatolian Diagonal area to the western fringe of Anatolia. Only much later, at about 400 Kya, did the lineage diversify and spread across the Aegean. Its members are highly heterozygous and some are of inter-specific hybrid origin (Chapter 4) and this may go some way towards explaining its position in the phylogenetic analyses. The period of diversification of the lineage, including the hybrid origin of some taxa, correlates with the emergence of significant tracts of land across the Aegan Basin

during the early to mid- Pleistocene (Tourloukis and Karkanas 2012) when the falling sea-levels allowed formerly isolated genepools to come into genetic contact.

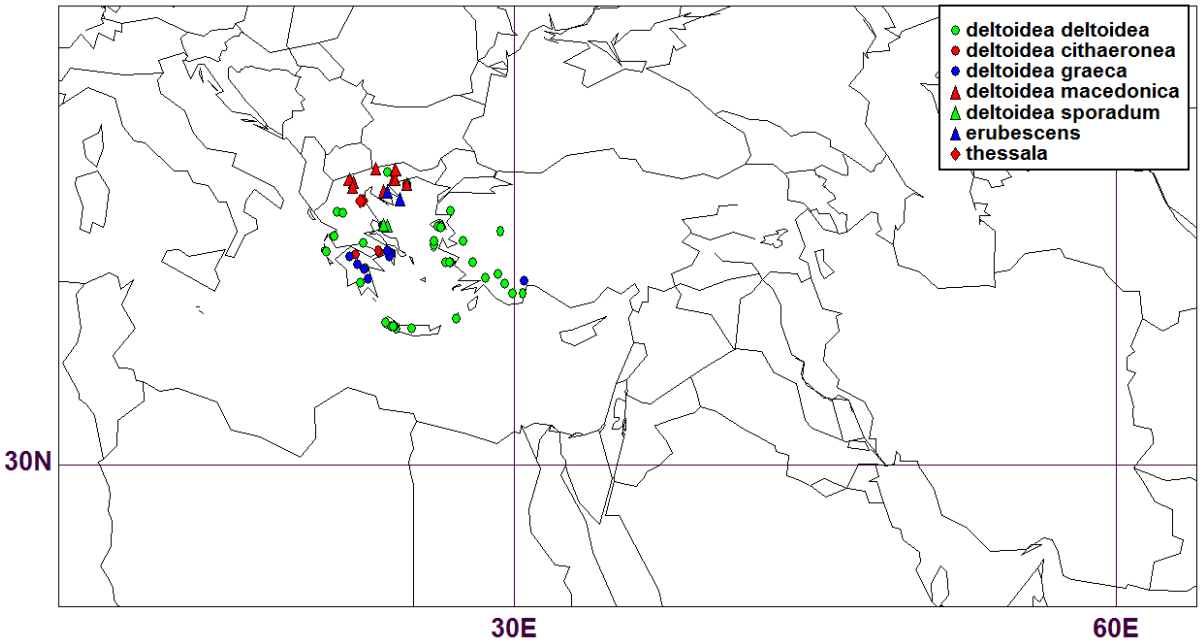


Figure 6.6 Geographical distribution of members of the Aegean Basin clade.

The third lineage to evolve was the Anatolian lineage (ca 831 Kya), the geographical distribution of whose members is shown in Fig. 6.7, based on specimens I have seen and published lists (Cullen 1965; Yüzbaşıoğlu *et al*. 2015). The boundary between the Anatolian and Near East clades correlates well with the Anatolian Diagonal (Figs. 6.5 and 6.7), a range of mountains that includes the Anti-Taurus and which runs north-east to south west. The land to the east of the Diagonal is at an average elevation of more than 2000m, consistently higher than it is to the west and thus is associated with a steep environmental gradient involving temperature seasonality, and therefore corresponds with a significant environmental barrier (Gür 2016). This suggests that selection against ill-adapted genotypes is primarily responsible for the distribution patterns rather than the inability of taxa to disperse. On one hand this boundary reduces the exchange of biota between Europe and the Near-East (Ekim and Guner 1986; Bilgin *et al*. 2006), but it also promotes the survival of indigenous organisms and the varied topography facilitates speciation, thereby increasing the levels of endemism (Davis 1971; Çiplak 2003). Ansell *et al*. (2011) demonstrated that genetic differentiation between populations of *Arabis alpina* was highest between central Anatolia and Anatolia-Iranian Plateaux. Likewise, the same result was observed with *Cardamine impatiens* (Kucera *et al*. 2006) when they studied taxonomy and phylogeography of this species.

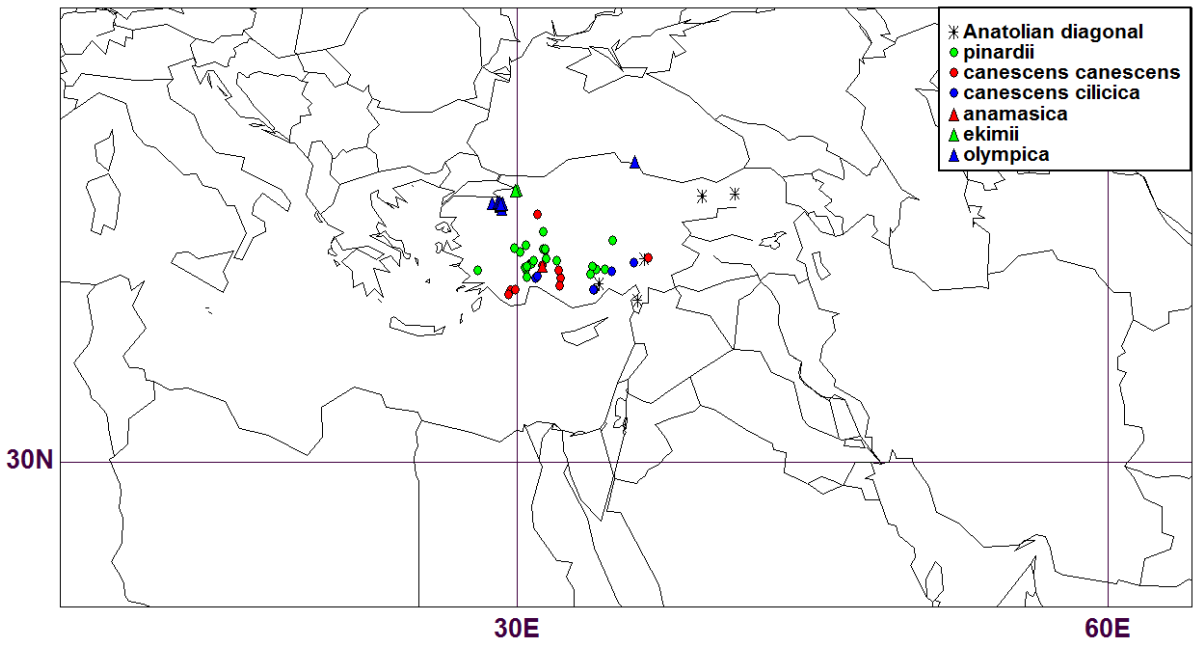


Figure 6.7 Geographical distribution of members of the Anatolian clade.

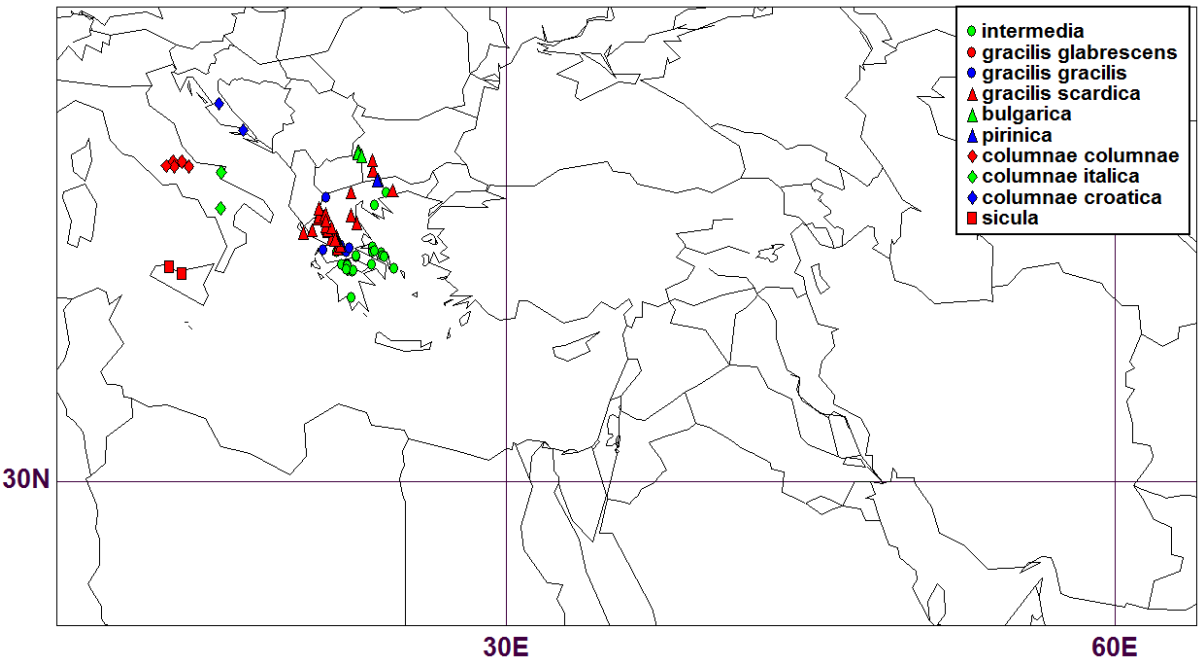


Figure 6.8 Geographical distributions of members of the Greece (montane) clade and the trans-Adriatic clade.

The final major split in *Aubrieta* took place about 665 Kya, to give rise to a fourth lineage, centred in the Pindus mountains of Greece but with extensions into Albania and Bulgaria, and its sister (fifth) lineage based around the Adriatic Sea (Italy and the Balkans). The geographical distributions of the two sister lineages are shown in Fig. 6.8, based on specimens I have seen and on published lists (Phitos 1970, 2002; Gustavsson 1986; Ančev and Goronova 2009). The boundary between the Anatolian and Aegean Coast clades corresponds to the Thyracian Plain

and Marmara and Aegean Seas, but at least by mid-Pleistocene the sea level in the Aegean Basin was at least about 130m lower than present level (Kerey *et al*. 2004), resulting in considerable areas of exposed land, and there were also land bridges connecting across the Marmara Sea and Bosporus Strait (Magyari *et al*. 2008). These connections allowed population exchange between the continents (Kucera *et al*. 2006; Dubey *et al.* 2007).

The trans-Adriatic differentiation of populations which are currently assigned to separate subspecies of *A. columnae*, viz. subsp. *columnae* and subsp. *italica* in Italy and subsp. *croatica* in Croatia and Bosnia-Hercegovinia, is reflected in differences in allozyme frequencies; indeed the genetic distances calculated between the three subspecies (ca 0.4—0.5) are very similar to those that separate *A. columnae* from *A. deltoidea* (Frizzi *et al*. 2013).

CHAPTER 7 CONCLUSIONS & TAXONOMY

As a consequence of the work described and discussed in the previous chapters, it is now possible to draw some conclusions regarding the structure of the *Aubrieta* genepool and how it might be classified.

7.1 Inferences from the molecular data

Plastid and mitochondrial data indicate the presence of *Arabis verna* plus five genepools that are strongly correlated with geography. Thus regional genepools can be identified as follows:

1) Near East (Levant, Iraq and Iran)
2) Anatolia
3) Aegean Basin
4) Greece, Albania, Bulgaria (Pindus Mts and associated ranges)
5) Trans-Adriatic Sea and Sicily

When sequence data from the nuclear genome is considered, the picture becomes somewhat less clear, owing to the effects of hybridisation, but nevertheless the same lineages can be discerned.

7.2 Morphological characters

A number of characters has been used to delimit the species of *Aubrieta*, including habit (caespitose vs straggling), leaf shape (principally the number of marginal teeth), petal shape, size and colour, stamen filament shape, fruit shape and size, and fruit indumentum. Petals and fruits feature heavily in modern identifications keys (Table 7.1). Early attempts to classify *Aubrieta* into species focussed on stamen filament shape (Boissier 1867). The filaments are more or less winged, such that, in the two short stamens, the wing extends upwards towards the apex to form a tooth of varying length, whereas in the four long stamens it either tapers distally or is truncated. These characters, however, have been shown to depend largely on the age of the flower (Wettstein 1892) and are therefore of little taxonomic use.

Habit and leaf shape vary strongly with environment within a species. Mattfeld (1939, p. 220) described in some detail how *A. intermedia*, for example, varies from short-stemmed variants with small, subentire leaves in sunny places in the Vouraikos Ravine, to long-stemmed variants with large, toothed leaves on nearby shady cliff faces, and back to small, entire-leaved,

cushion-forming phenotypes above the tree line. Bornmuller (1936) studied many of the taxa in the field and came to the conclusion that the size and colour of the petals, as well

Table 7.1 Morphological characters used to distinguish taxa in *Aubrieta*. Data from Cullen (1965), Rechinger (1968), Phitos (1970, 2002), Gustavsson (1976), Akeroyd (1993).

Taxon	Flower colour	Petal length, mm	Fruit			Fruit trichomes		
			length, mm	width, mm	cross-section	stellate	2-4-furcate	simple
verna	violet / white	5-8	25-70	1.5-2	strongly compressed	0	1	1
libanotica	violet	10-12	10	3.2-3.8	circular	1	1	1
parviflora	white to pale lilac/pale pink	7--8	4-7	3-4	ellipsoid	1	1	0
vulcanica	violet	6--7	8-10	3	ellipsoid	1	0	0
macrostyla	violet	11--19	9-17	1.5-2	ellipsoid	1	0	0
scyria	violet	(11-)13-18	(8-)10-15	2-3(-4)	semi-terete	1	1	0
pinardii	purple	18-19	22-35	2-2.5	strongly compressed	1	0	0
ekimii	violet	6.5-8	(24-)30-42(-45)	(2-)2.5-3(-3.7)	strongly compressed	1	1	0
canescens	violet	11--19	9-11	2-5	inflated or terete	1	0	0
cilicica	violet	11--19	12-13	2-5	semi-terete	1	0	0
anamasica	purple	13-16	17-18	2	compressed	1	0	0
olympica	violet	11--19	15-18	4.5-5.5	strongly compressed	1	0/1	0
deltoidea (Tu)	purple	15-17	9-13	3-3.5	inflated	1	1	1
deltoidea (Gr)	purple	11--29	(5-)7-16(-23)	1.8-4.7(-5.3)	compressed or not	1	1	1
erubescens	pink, white, sometimes violet	8--11	7-12	2-3.5	semi-terete	1	1	0
thessala	purple	12--20	14-26	2.2-4.1	strongly compressed	1	1	0
intermedia	purple	12(-15)-22	(12-)13-18(-20)	2.5-4(-5)	strongly compressed	1	1	0
scardica	purple	10.5-21	(10-)13-31(-35)	(2.1-) 2.5-3.7 (-4.3)	strongly compressed	1	0	0
bulgarica		18-25	(12-)14–25			1	1	0
glabrescens	purple /violet	13-15.5	13-32	1.6-2.8	strongly compressed	0/1	0	0
gracilis	purple	10.5-21.5	(15-)21-35(-45)	(1.7-) 2.0-3.4 (-3.9)	strongly compressed	1	0	0
sicula	red-purple or violet (white)	15-16	12-14	3-4.5	semi-terete	1	1	1
columnae	purple /violet	11-17	5-12	2-4.5	semi-terete	1	1	0
italica	purple /violet	15-18	8-11	3-4	semi-terete	1	1	0
croatica	purple /violet	12--18	7-16	2.5-4.5	semi-terete	1	1	0
pirinica		14-20	5-14		semi-terete	1	1	0

as the shape and size of the fruit and the length of the style vary between plants of the same taxon in the same geographical area and even within a single plant. It is tempting to suggest that many of the characters are to some extent probably under direct environmental control, i.e. phenotypically plastic, although the necessary cultivation experiments have yet to be conducted (Fig. 7.1).

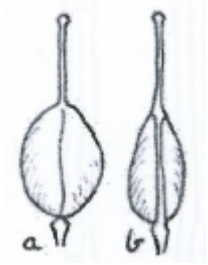
Figure 7.1 Variation between individuals of *A. parviflora* from the same wild population grown under different conditions.
Left: photoperiod; day time 16 hours, temperature 30 ℃; night time 8 hours, temperature 20 ℃.
Right: photoperiod: day time 14 hours, temperature 20 ℃; night time 10 hours, temperature 10 ℃.

It seems that fruit morphology and indumentum are the most reliable indicators of specific identity. Fruits vary in length and degree of compression and this feature appears prominently in the identification keys in modern Floras but, as has been noted, there is considerable and overlapping variation (Mattfeld 1939). See also Fig. 7.2 which shows some of the variation within and between clades. Problems arise especially when considering the genus over its whole distribution. Thus Cullen (1965) remarked that *A. columnae*, *A. intermedia* and *A. canescens*, with mid-length pods, are doubtfully distinct and "should perhaps be combined". Similarly the distinction between the Greek *A. gracilis* and the Turkish *A. pinardii*, both with long pods, is hard to make.

Much has been made of the indumentum on the fruits in the distinction of species. Thus three main types of fruit hair can be distinguished: i) short or subsessile stellate hairs; ii) long-stalked hairs with two to four branches; and iii) simple, unbranched, hairs. When hairs are present on the fruits (glabrous morphs occur sporadically) all taxa except *Arabis verna* have the short-stalked or subsessile stellate hairs. Some taxa have in addition, longer hairs that are branched and/or simple. In *Arabis verna* the fruits are usually glabrous but occasional populations bear long, branched or simple hairs (types ii and iii). Long, simple fruit hairs are apparently restricted to *Arabis verna*, *Aubrieta libanotica* and *A. deltoidea* (including var. *sicula,* but excluding *A. intermedia*).

Figure 7.2 Fruit shape examples from the different *Aubrieta* clades. Illustrations taken from Mattfeld (1939).

Near East clade



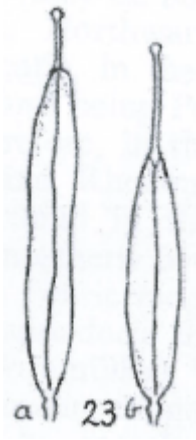
libanotica

kotschyi

parviflora

Anatolia clade

pinardii

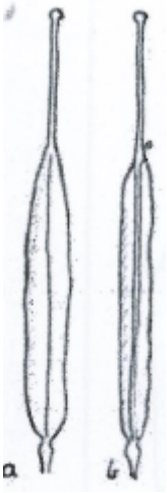


pinardii

olympica

olympica

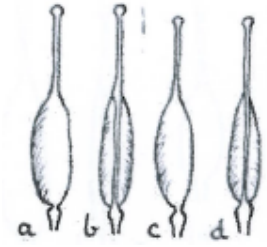
Aegean clade



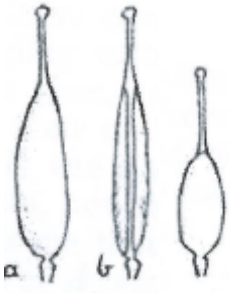
cithaeronia

graeca

deltoidea



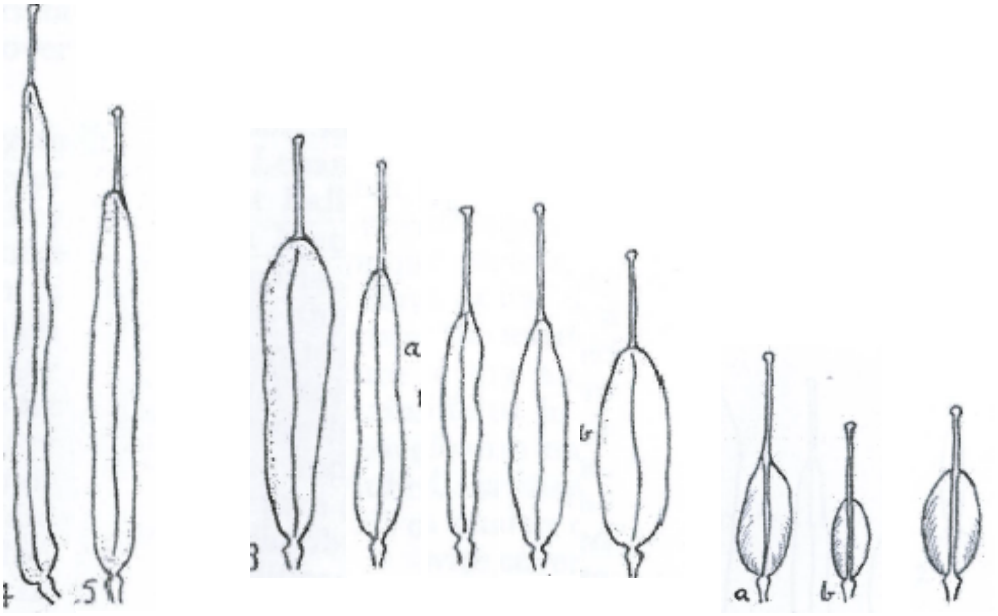
microphylla



erubescens

Fig. 7.2 cont'd

Greek and Trans-Adriatic clades



There have been study of pollen and of seed surfaces in species by Yüzbaşıoğlu *et al*. (2015) but there has been no systematic survey, so the range of variation present in these characters in the genus is still largely unknown. Preliminary studies suggest that variability is relatively low and that therefore such micro-morphological features may not contribute much towards solving the problem of species recognition.

A perusal of the synonymy of many of the taxa shows the uncertainty with which taxonomists have assessed relationships. For example, several taxa have been regarded as variants of both *A. deltoidea* and *A. intermedia* to the extent that the modern view is to combine these two taxa into one species, even though *A. intermedia* lacks the simple fruit hairs allegedly diagnostic of *A. deltoidea*. This uncertainty also extends to the identification of herbarium specimens, where many have been wrongly named (Yüzbaşıoğlu *et al*. 2015).

7.3 Taxonomy

The biology of the situation implies that we have *Arabis verna* plus five genetic lineages of *Aubrieta*, each with its own largely allopatric geographical distribution. But there are no morphological characters that can be used to diagnose these lineages. Each encompasses a range of morphologies, including of fruit type, which have clearly evolved in parallel. For example, if we assume that long fruits are ancestral, because they characterise *Arabis verna* and most other members of tribe Arabideae, then we can see that the evolution of shorter fruits has occurred at least once in each of the five lineages, leading to similar phenotypes in what are otherwise distinct clades, i.e. cryptic taxa.

There are perhaps two taxonomic solutions to this problem: i) recognise two species, i.e. *Arabis* (*Aubrieta*) *verna* and *Aubrieta deltoidea* sensu lato; or ii) recognise each diagnosable local population as a separate species. The problem with the latter approach is that some 'local' populations are not separable by morphology, e.g. *A. cansecens*, *A. columnae* and *A. intermedia*. In practice, this may only matter on a global scale. On a more local scale, where only one cryptic taxon is present, it will not be a problem. Thus in the flora of Turkey, for example, it is not necessary to consider the Italian *A. columnae* when recording *A. canescens*, and nor does the field botanist need to consider the Greek *A. gracilis* when recording *A. pinardii*. Notwithstanding the rather unsatisfactory nature of this proposal, it does at least have the merit that almost all of the names that would be needed are in existence. Furthermore, each of the species so recognised occupies a well-defined, largely allopatric geographical area and could be considered to be an evolutionary unit as envisaged under the phylogenetic species concept (e.g. Nixon and Wheeler 1990).

The alternative two species solution would require a significantly broader concept for *A. deltoidea* (the earliest name at species rank). Although such a species would be easy to recognise, its constituent variation would require an extensive infra-specific taxonomy employing the ranks of subspecies, variety and possibly also forma. Voss (1894) already adopted this solution and many of the required names exist, but not all. His approach, however, has not found favour in modern Floras.

On the basis of the assessments of both molecular and morphological data, I follow the second approach and recognise 21 species and associated infraspecific taxa as summarised in the following conspectus. Species are recognised on the basis of their genetic integrity (presumed in some cases, demonstrated in this study in others), supported by at least one morphological

character that can be used in keys for the purpose of identification. Genetic integrity and its morphological expression arises as a consequence of i) shared ancestry and descent; ii) reproductive coherence; and iii) occupation of a common ecological niche. For infraspecific taxa, I have followed the practice of Du Reitz (1930). Thus subspecies are regarded as regional variants of a species, where there is some element of morphological or genetic overlap between them. Varieties are treated as local variants, which may be polytopic in origin. Names at the level of forma are given to individuals whose diagnostic character is likely to be under single or oligogenic control, e.g. white petals; they may occur mixed in a population with more typical morphs.

7.3.1 Taxonomic changes

Several taxonomic changes are proposed as a consequence of the work presented here. Explanations for these changes are given below and the nomenclatural adjustments are made in the appropriate place in the conspectus which follows.

1) *Arabis verna* is basal or nearly so in *Aubrieta*. This being the case, it is desirable to formally include it in the genus and I make the necessary combination below. The alternative solution would be to erect a new, monospecific genus and there are insufficient grounds for doing so, especially given that the genetic control of the annual versus perennial habit appears to be relatively simple (Wang *et al*. 2009).
2) *Aubrieta canescens* was accorded three subspecies by Cullen (1965). The molecular phylogenies presented here, however, show that subsp. *macrostyla* belongs in the Near East clade, separate from the other two subspecies of *A. canescens*, which are in the Anatolian clade. Thus I propose to elevate subsp. *macrostyla* to species rank.
3) The populations of *Aubrieta* from Sicily have hitherto been recognised as *A. deltoidea* var. *sicula*. However, molecular phylogenetic analysis shows that they are more closely related to the trans-Adriatic group, i.e. *A. columnae* and its relatives. And this is despite the presence in var. *sicula* of long, simple fruit hairs, "albeit weak and not very numerous" (Mattfeld 1939), that are otherwise considered characteristic of *A. deltoidea*. This hair type also occurs, however, in *A. libanotica*, and so is not unique to *A. deltoidea*.

A summary of the proposed taxonomy is given in Table 7.2.

Table 7.2 Summary of the taxonomy proposed for *Aubrieta*

Mediterranean (Anatolia to Spain and N. Africa)

1) *A. verna*

Near East (Levant, Iraq, Iran)

2) *A. libanotica*
3) *A. parviflora*
4) *A. vulcanica*
5) *A. macrostyla*

Anatolia plus Euboea/Skiros

6) *A. pinardii*
7) *A. ekimii*
8) *A. canescens*
    a) subsp. *canescens*
    b) subsp. *cilicica*
9) *A. olympica*
10) *A. anamasica*
11) *A. scyria*

Aegean Basin

12) *A. thessala*
13) *A. deltoidea*
    a) var. *deltoidea*
    b) var. *graeca*
    c) var. *cithaeronea*
14) *A. erubescens*

Greece/Albania (Pindus Mts)

15) *A. gracilis*
16) *A. scardica*
    a) f. *scardica*
    b) f. *degeniana*
17) *A. intermedia*
    a) subsp. *intermedia*
    b) subsp. *macedonica*

Trans-Adriatic / Sicily

18) *A. columnae*
19) *A. croatica*
    a) ssp. *croatica*
    b) ssp. *pirinica*
20) *A. italica*
21) *A. sicula*

7.3.2 Taxonomic conspectus

Aubrieta Adans., Fam. Pl. 2: 420 (1763).

1. **Aubrieta verna** (L.) J. Muhammed, **comb. nov.**

   Hesperis verna L., Sp. Pl. {Linnaeus} 2: 664 (1753).

   Turritis purpurea Lam., Fl. Franç. {Lamarck} 2: 491 (1778 [>1779.iii.21]).

   Arabis violacea Moench, Methodus {Moench} 259 (1794).

   Arabis verna (L.) R.Br., Hort. Kew., ed. 2 {W.T. Aiton} 4: 105 (1812).

   Hesperis dauriensis Amo, Fl. Fan. Penins. Iberica 6: 495 (1873).

   Erysimum vernum (L.) Kuntze, non Mill., Revis. Gen. Pl. 2: 934 (1891).

   Turrita vernalis Bubani, Fl. Pyren. {Bubani} 3: 153 (1901).

   Arabis verna var. genuina Briq.

   Arabis verna var. dasycarpa Godr. ex Rouy & Foucaud.

2. **Aubrieta libanotica** Boiss., Diagn. Pl. Orient. ser. 1 8: 32 (1849).

   Aubrieta deltoidea subsp. libanotica (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

   Aubrieta schweinfurthiana Muschl., Repert. Spec. Nov. Regni Veg. Beih. 3: 212 (1906).

   Aubrieta antilibani hort., nom. nud., Bull. Alp. Gard. Soc. 7: 218 (1939).

3. **Aubrieta parviflora** Boiss., Diagn. Pl. Orient. ser. 1 6: 14 (1845, [1846]).

   Aubrieta kotschyi Boiss. & Hohen., Diagn. Pl. Orient. ser. 1 8: 32 (1849).

   Aubrieta edentula Boiss., Fl. Orient. {Boissier} 1: 254 (1867).

   Aubrieta parviflora var. kurdica Boiss., queried. Fl. Orient. {Boissier} 1: 253 (1867).

   Aubrieta elwendica Stapf, Denkschr. Akad. Wien 51: 301 (1886).

   Aubrieta deltoidea subsp. kotschyi (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 76 (1896).

   Aubrieta deltoidea subsp. parviflora (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 76 (1896).

   Aubrieta deltoidea subsp. edentula (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 76 (1896).

   Aubrieta kotschyi var. brachycarpa Bornm., Beih. Bot. Centralbl., Abt. 2. 28(2): 112 (1911).

4. **Aubrieta vulcanica** Hayek & Siehe, Ann. Nat. Hofmus. Wien 28: 154 (1914).
5. **Aubrieta macrostyla** (Cullen & Huber-Morath) Muhammed., comb. et stat. nov.

   Aubrieta macrostyla hort., nom. nud., Bull. Alp. Gard. Soc. 7: 163 (1939).

   Aubrieta canescens subsp. macrostyla Cullen & Huber-Morath, Notes Roy. Bot. Gard. Edinburgh 26: 191 (1965).
6. **Aubrieta pinardii** Boiss., Diagn. Pl. Orient. ser. 2 1: 37 (1854).

   Aubrieta pinardii var. integrifolia Boiss., Fl. Orient. {Boissier} 1: 252 (1867).

   Aubrieta deltoidea subsp. pinardii (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).
7. **Aubrieta ekimii** Yüzb., Al-Shehbaz & M. Koch, Pl. Syst. Evol. 301: 2046 (2015e).
8. **Aubrieta canescens** (Boiss.) Bornm., Repert. Spec. Nov. Regni Veg. Beih. 89: 44 cum descr. ampl. (1936).

   Aubrieta affinis Hausskn. ex Bornm. in obs., pro syn. Repert. Spec. Nov. Regni Veg. Beih. 89: 44 (1936).

   8a. subsp. **canescens**

   Aubrieta deltoidea var. canescens Boiss., Fl. Orient. {Boissier} 1: 252 (1867).

   Aubrieta deltoidea f. canescens (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

   8b. subsp. **cilicica** (Boiss.) Cullen, Notes Roy. Bot. Gard. Edinburgh 26: 191 (1965).

   Aubrieta deltoidea var. cilicica Boiss., Fl. Orient. {Boissier} 1: 252 (1867).

   Aubrieta deltoidea f. cilicica Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).
9. **Aubrieta olympica** Boiss., Fl. Orient. {Boissier} 1: 251 (1867).

   Aubrieta deltoidea subsp. olympica (Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

   Aubrieta olympica var. macrocarpa Haussknecht in Bornm., Repert. Spec. Nov. Regni Veg. Beih. 89: 44 (1936).
10. **Aubrieta anamasica** H. Pesmen & A. Guner, Notes Roy. Bot. Gard. Edinburgh 36(1): 35 (1978).
11. **Aubrieta scyria** Halácsy, Oesterr. Bot. Z. 60: 115 1910.

12. **Aubrieta thessala** H. Boissieu, Bull. Soc. Bot. France 43: 288 (1896).

    Aubrieta gracilis subsp. thessala (Boissieu) Maire & Petitmengin, Étud. Pl. Vasc. Gréce, (1908).

    Aubrieta gracilis var. thessala (Boissieu) Hayek, Prodr. Fl. Penins. Balkan. 1: 409 (1927).

    Aubrieta thessalica hort., nom. nud., sphalm.? Bull. Alp. Gard. Soc. 7: 222 (1939).

13. **Aubrieta deltoidea** (L.) DC., Syst. Nat. {DC.} 2: 294 1821.v (late) & in Mem. Mus. Par. 7: 232 (1821).

    Alyssum deltoideum L., Sp. Pl., 2nd ed. 2: 908 (1763).

    Draba hesperidiflora Lam., nom. illeg. (Alyssum deltoideum L. cited in synonymy), Encycl. {J. Lamarck & al.} 2: 328 (1786).

    Vesicaria deltoidea (L.) Poir., Encycl. {J. Lamarck & al.} 8: 572 (1808).

    Farsetia deltoidea (L.) R. Br., Hort. Kew., ed. 2 {W.T. Aiton} 4: 97 (1812).

    Aubrieta floribunda Spach, nom. inval.: (A. deltoidea cited in synonymy), Hist. Nat. Vég. {Spach} 6: 469 (1838).

    Aubrieta integrifolia Fisch. & C.A. Mey., Ann. Sci. Nat., Bot. sér. 4, 1: 35 [30?] (1854).

    Aubrieta deltoidea var. microphylla Boiss., Fl. Orient. {Boissier} 1: 252 (1867).

    ?Aubrieta purpurea hort. non (Sibth. & Sm.) DC., J. Horticulture, ser.3 16: 468-469 (1888).

    Aubrieta deltoidea f. microphylla Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

    13a. var. **deltoidea**

    Aubrieta deltoidea α. typica Paoletti, Fl. Anal. Ital. {Fiori & Paoletti} 1: 459 (1898).

    Aubrieta deltoidea subsp. normalis Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

    13b. var. **graeca** (Griseb.) Regel, Gartenfl. 20: 257 (1871).

    Aubrieta graeca Griseb., Spic. Fl. Rumel. 1: 268 (1843).

    Aubrieta deltoidea f. graeca (Griseb.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

    Aubrieta deltoidea subsp. sporadum Phitos, Phyton (Horn) 12: 116 (1967).

    Aubrieta deltoidea var. sporadum (Phitos) Phitos, Candollea 25: 80 (1970).

13c. var. **cithaeronea** Mattf., descr. German. Blätt. Staudenk. 1 (fol.5 recto) (1937); descr. Latin, Quart. Bull. Alp. Gard. Soc. 2: 180 (1939).

14. **Aubrieta erubescens** Griseb., Spic. Fl. Rumel. 1: 268 (1843).

Aubrieta erubescens f. latifolia Tocl & Rohlena, Sitzungsber. Königl. Böhm. Ges. Wiss. Prag. Math.-Naturwiss. Cl. 49: 7 (1902).

Aubrieta deltoidea subsp. erubescens (Griseb.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

15. **Aubrieta gracilis** Spruner ex Boiss., Diagn. Pl. Orient. ser. 1, 1: 74 (1843).

Aubrieta deltoidea subsp. gracilis (Spruner ex Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

15×17. **Aubrieta ×hybrida** Hausskn., Mitth. Thüring. Bot. Vereins, n.f. 3-4: 111 (1893) = A. gracilis × intermedia

16. **Aubrieta scardica** (Wettst.) L.-Å. Gustavsson, Mount. Fl. Greece {A.Strid} 1: 271 (1986).

Aubrieta croatica var. scardica Wettst., Biblioth. Bot. 26: 21 (1892).

Aubrieta gracilis var. elongata Hausskn., Mitth. Thüring. Bot. Vereins, n.f. 3-4: 111 (1893).

Aubrieta gracilis var. pirinica Stoj. & Acht., Izv. Tsarsk. Prir. Inst. Sofia 12: 184 (1939).

Aubrieta gracilis var. "pirinensis" Stoj. & Stef., Fl. Bulg., 3rd ed. 516 (1948).

Aubrieta gracilis f. "pirinensis" (Stoj. & Acht). Stoj., Stef. & Kitan., comb. inval. Fl. Bulg., 4th ed. 1: 475 (1966).

Aubrieta gracilis subsp. scardica (Wettst.) Phitos, Candollea 25: 84 (1970).

Aubrieta albanica F.K. Mey. & J.E. Mey., Haussknechtia Beih. 15: 62, f.8 (2011).

16a. f. **scardica**

16b. f. **degeniana** (Bald.) J. Muhammed, comb. nov.

Aubrieta deltoidea var. degeniana Bald., Nuovo Giorn. Bot. Ital., ser.2, 6: 15 (1899).

Aubrieta gracilis f. degeniana (Bald.) Hayek, Prodr. Fl. Penins. Balkan. 1: 409 (1927).

Aubrieta glabrescens Turrill, Gard. Chron., ser. 3, 95: 384 (1934).

Aubrieta gracilis var. degeniana (Bald.) Phitos, Candollea 25: 86 (1970).

Aubrieta gracilis subsp. glabrescens (Turrill) Akeroyd, Bot. J. Linn. Soc. 106(2): 100 (1991).

17. **Aubrieta intermedia** Heldr. & Orph. ex Boiss., Diagn. Pl. Orient. ser. 2 1: 36 (1854).

Aubrieta deltoidea var. intermedia (Boiss.) Bald., Malpighia 8: 73 (1894).

Aubrieta deltoidea subsp. intermedia (Heldr. & Orph. ex Boiss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea subsp. intermedia (Heldr. & Orph. ex Boiss.) Maire & Petitmengin, Étud. Pl. Vasc. Gréce, (1908). Isonym.

17a. subsp. **intermedia**

17b. subsp. **macedonica (**Adamović) J. Muhammed, comb. nov.

Aubrieta intermedia var. macedonica Adamović, Denkschr. Kaiserl. Akad. Wiss. Math.-Naturwiss. Kl. 74: 125 (1904).

Aubrieta deltoidea var. macedonica (Adamović) Phitos, Candollea 25: 80 (1970).

18. **Aubrieta columnae** Guss., Pl. Rar. 266 (1826).

Aubrieta deltoidea subsp. columnae (Guss.) Voss, Vilm. Blumengärtn., ed. 3. 1: 76 (1896).

Aubrieta deltoidea β. columnae (Guss.) Paoletti, Fl. Anal. Ital. {Fiori & Paoletti} 1: 459 (1898).

19. **Aubrieta croatica** Schott, Nyman & Kotschy, Analect. Bot. 47 (1854).

Aubrieta deltoidea f. croatica Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea var. croatica (Schott) Bald., Rivista della collezione botanica fatta nel 1896 in Albania (1899).

Aubrieta columnae subsp. croatica (Schott) Mattf., Blätt. Staudenk. 1 (1937).

19a. subsp. **croatica**

19b. subsp. **pirinica** (Assenov) J. Muhammed, **comb. nov.**

Aubrieta columnae subsp. pirinica {Stoj. ex} Assenov, Fl. Reipubl. Pop. Bulgaricae {Jordanov} 4: 474, 707 (1970).

Aubrieta intermedia var. pirinica Stoj., Izv. Bulg. Bot. Druzh. 1: 72 (1926), nom. nud.

Aubrieta columnae subsp. bulgarica Ančev, Phytol. Balcan. 13(2): 160 (2007).

20. **Aubrieta italica** Boiss., in nota, Fl. Orient. {Boissier} 1: 252 (1867).

Aubrieta deltoidea var. major Ten., Succ. Relaz. Viaggio Abruzzo (1830).

Aubrieta deltoidea β. italica (Boiss.) Fiori, Nuova Fl. Analitica d'Ital. 1: 609 (1924).

Aubrieta columnae subsp. italica (Boiss.) Mattf., Blätt. Staudenk. 1 (1937).

21. **Aubrieta sicula** (Strobl) J. Muhammed, **comb. et stat. nov**.

Aubrieta deltoidea var. sicula Strobl, Verh. Zool.-Bot. Ges. Wien 53: 458 (1903).

Aubrieta deltoidea subsp. sicula (Strobl) Phitos, Candollea 25: 76 (1970).

Aubrieta × cultorum Bergmans, Vaste Pl. Rotsheesters 97 (1924) = garden plants derived from hybridisation between *A. deltoidea*, *A. libanotica*, *A. olympica*, *A. erubescens* and *A. kotschyi*.

Aubrieta hendersonii E.G. Hend., Gard. Chron., n.s. 1: 2 (1874).

Aubrieta deltoidea f. hendersonii Voss, Vilm. Blumengärtn., ed. 3. 1: 75 1896

Aubrieta campbellii Hort. & Lem. Illustr. Hortic. xii. (1865) t. 455 Gard. Chron. 1133 1864

Aubrieta deltoidea f. campbellii Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. taurica (Schott) Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. foliis variegatis Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. violacea Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. conspicua Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. eyriesii Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. bougainvillei Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. morreniana Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta deltoidea f. richardii Voss, Vilm. Blumengärtn., ed. 3. 1: 75 1896

Aubrieta taurica Hort. nom. nud. Bull. Alp. Gard. Soc. 7: 174 1939

Aubrieta tauricola Hort. nom. nud. Bull. Alp. Gard. Soc. 7: 174 1939

Aubrieta leichtlinii Hort. ex Wien., Ill. Gartenz. 13: 325 (1888).

Aubrieta deltoidea f. leichtlinii Voss, Vilm. Blumengärtn., ed. 3. 1: 75 (1896).

Aubrieta hesperidiflora G. Don ex Loud., Hort. Brit. 257 (1830).

Excluded taxa

Aubrieta purpurea (Sibth. & Sm.) DC., Syst. Nat. {DC.} 2: 294 (1821) = Arabis purpurea Sibth. & Sm.

Incertae sedis

Aubrieta lepidioides Spreng., Syst. Veg. (ed. 16) {Sprengel} 2: 871 (1825).

Aubrieta froebeli hort. ex Guenthart, Beitr. Bluthenb. 22

Aubrieta diffusa Spreng., Syst. Veg. (ed. 16) {Sprengel} 2: 871 (1825).

Aubrieta deltoidea var. minor Ten., Succ. Relaz. Viaggio Abruzzo, (1830).

REFERENCES

Acosta, M. C. and Premoli, A. C. (2010) Evidence of chloroplast capture in South American *Nothofagus* (subgenus Nothofagus, Nothofagaceae). *Molecular phylogenetics and evolution,54*(1), pp. 235-242.

Akeroyd, J. and Ball, P. (1993) *Aubrieta* Adanson, in Tutin, T.G., Burges, N.A., Chater, A.O., Edmondson, J.R., Heywood, V.H., Moore, D.M., Valentine, D.H., Walters, S.M., Webb, D.A., eds. *Flora Europaea, 1*, pp. 356-358. Cambridge: Cambridge University Press.

Al-Shehbaz, I., Beilstein, M. and Kellogg, E. (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Plant Systematics and Evolution,259*(2-4), pp. 89-120.

Al-Shehbaz, I. A. (1977) Protogyny in the Cruciferae. *Systematic Botany,15*, pp. 327-333.

Al-Shehbaz, I. A. (1984) The tribes of Cruciferae (Brassicaceae) in the southeastern United States. *Journal of the Arnold Arboretum,65*(3), pp. 343-373.

Al-Shehbaz, I. A. (2005) Nomenclatural notes on Eurasian *Arabis* (Brassicaceae). *Novon*, pp. 519-524.

Al-Shehbaz, I. A. (2012) A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon*,61(5), pp. 931-954.

Al-Shehbaz, I. A. and Al-Omar, M.M. (1982) In IOPP chromosome number reports LXXVI. *Taxon,3*, pp. 587-589.

Al-Shehbaz, I. A. and Al-Omar, M.M. (1983) In IOPP chromosome number reports LXXX. *Taxon,32*, pp. 508-509.

Al-Shehbaz, I. A., German, D. A., Mummenhoff, K. and Moazzeni, H. (2014) Systematics, tribal placements, and synopses of the Malcolmia sl segregates (Brassicaceae). *Harvard Papers in Botany,19*(1), pp. 53-71.

Al-Shehbaz, I. A., O'Kane Jr, S. L. and Price, R. A. (1999) Generic placement of species excluded from *Arabidopsis* (Brassicaceae). *Novon,9*, pp. 296-307.

Al-Shehbaz, I. A. and Warwick, S. I. (2005) A synopsis of *Eutrema* (Brassicaceae). *Harvard Papers in Botany,10*(2), pp. 129-135.

Aldrich, J., Cherney, B. W. and Merlin, E. (1988) The role of insertions/deletions in the evolution of the intergenic region between psbA and trnH in the chloroplast genome. *Current genetics,14*(2), pp. 137-146.

Ali, H. B., Lysak, M. A. and Schubert, I. (2005) Chromosomal localization of rDNA in the Brassicaceae. *Genome,48*(2), pp. 341-346.

Alix, K. (2004) The diversity of retroelements in diploid and allotetraploid *Brassica* species. *Plant molecular biology,54*(6), pp. 895-909.

Alix, K., Ryder, C. D., Moore, J., King, G. J. and Heslop-Harrison, J. P. (2005) The genomic organization of retrotransposons in *Brassica oleracea*. *Plant molecular biology,59*(6), pp. 839-851.

Allison, L. A., Moyle, M., Shales, M. and Ingles, C. J. (1985) Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. *Cell,42*(2), pp. 599-610.

Álvarez, I. and Wendel, J. F. (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular phylogenetics and evolution,29*(3), pp. 417-434.

Amborella, G. P. (2013) The Amborella genome and the evolution of flowering plants. *Science (New York, NY),342*(6165), pp. 1241089.

Anastasakis, G. and Dermitzakis, M. (1990) Post-Middle-Miocene paleogeographic evolution of the Central Aegean Sea and detailed Quaternary reconstruction of the region. Its possible influence on the distribution of the Quaternary mammals of the Cyclades Islands. *Neues Jahrbuch für Geologie und Paläontologie, 1*, pp. 1-16.

Ančev, M. and Goranova, V. (2009) *Aubrieta* (Brassicaceae) in the Bulgarian flora. *Phytologia Balcanica,15*(1), pp. 43-50.

Ansari, H. A., Ellison, N. W., Readers, S. M., Badaeva, E. D., Friebe, B., Miller, T. E. and Williams, W. M. (1999) Molecular Cytogenetic Organization of 5S and 18S-26S rDNA Loci in White Clover (*Trifolium repens* L.) and Related Species. *Annals of Botany,83*(3), pp. 199-206.

Appels, R., Morris, R., Gill, B. S. and May, C. E. (2012) *Chromosome biology*. Berlin: Springer Science & Business Media.

Arnheim, N. (1983) Concerted evolution of multigene families. *Evolution of genes and proteins*, pp. *38-61*. Sunderland, MA: Sinauer Associates.

Aryavand, A. (1975) Contribution a l'etude Cytotaxinomique de quelques Cruciferes de l'Iran et de la Turquie. *Bulletin de la Société neuchâteloise des sciences naturelles,98*, pp. 43-58.

Atalay, Î. (1996) Palaeosols as indicators of the climatic changes during Quaternary period in S. Anatolia. *Journal of Arid Environments,32*(1), pp. 23-35.

Azuma, H., García-Franco, J. G., Rico-Gray, V. and Thien, L. B. (2001) Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. *American Journal of Botany,88*(12), pp. 2275-2285.

Azuma, H., Thien, L. B. and Kawano, S. (1999) Molecular phylogeny of *Magnolia* (Magnoliaceae) inferred from cpDNA sequences and evolutionary divergence of the floral scents. *Journal of Plant Research,112*(3), pp. 291-306.

Badaeva, E. D., Friebe, B. and Gill, B. S. (1996) Genome differentiation in *Aegilops*. 2. Physical mapping of 5S and 18S-26S ribosomal RNA gene families in diploid species. *Genome,39*(6), pp. 1150-1158.

Bailey, C. D., Koch, M. A., Mayer, M., Mummenhoff, K., O'Kane, S. L., Warwick, S. I., Windham, M. D. and Al-Shehbaz, I. A. (2006) Toward a global phylogeny of the Brassicaceae. *Molecular Biology and Evolution,23*(11), pp. 2142-2160.

Bandelt, H.-J. and Dress, A. W. (1992) Split decomposition: a new and useful approach to phylogenetic analysis of distance data. *Molecular phylogenetics and evolution,1*(3), pp. 242-252.

Bartish, I. V., Kadereit, J. W. and Comes, H. P. (2006) Late Quaternary history of *Hippophae rhamnoides* L.(Elaeagnaceae) inferred from chalcone synthase intron (Chsi) sequences and chloroplast DNA variation. *Molecular ecology,15*(13), pp. 4065-4083.

Beerli, P., Hotz, H. and Uzzell, T. (1996) Geologically dated sea barriers calibrate a protein clock for Aegean water frogs. *Evolution*, 50, pp. 1676-1687.

Beilstein, M. A., Al-Shehbaz, I. A. and Kellogg, E. A. (2006) Brassicaceae phylogeny and trichome evolution. *American Journal of Botany,93*(4), pp. 607-619.

Beilstein, M. A., Al-Shehbaz, I. A., Mathews, S. and Kellogg, E. A. (2008) Brassicaceae phylogeny inferred from phytochrome A and ndhF sequence data: tribes and trichomes revisited. *American Journal of Botany,95*(10), pp. 1307-1327.

Beilstein, M. A., Nagalingum, N. S., Clements, M. D., Manchester, S. R. and Mathews, S. (2010) Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences,107*(43), pp. 18724-18728.

Bell, C.D., Soltis, D.E. and Soltis, P.S. (2010) The age and diversification of the angiosperms re-revisited. *American journal of botany*, 97(8), pp.1296-1303.

Bendich, A. J. (2004) Circular chloroplast chromosomes: the grand illusion. *Plant Cell,16*(7), pp. 1661-1666.

Bennett, M. and Leitch, I. (2011) Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. *Annals of Botany,107*(3), pp. 467-590.

Bennett, M. D. and Smith, J. (1976) Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,274*(933), pp. 227-274.

Bergmans, J. (1924) *Vaste planten en rotsheesters*. Nv Drukkerij voorheen de erven Loosjes.

Bieniek, W., Mizianty, M. and Szklarczyk, M. (2015) Sequence variation at the three chloroplast loci (matK, rbcL, trnH-psbA) in the Triticeae tribe (Poaceae): comments on the relationships and utility in DNA barcoding of selected species. *Plant Systematics and Evolution,301*(4), pp. 1275-1286.

Bilgin, R., Karatas, A., Coraman, E., Pandurski, I., Papadatou, E. and Morales, J.C. (2006) Molecular taxonomy and phylogeography of *Miniopterus schreibersii* (Kuhl, 1817) (Chiroptera: Vespertilionidae), in the Eurasian transition. *Biological Journal of the Linnean Society*, *87*, pp. 577–582.

Bittkau C. and Comes H.P. (2005) Evolutionary processes in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. *Molecular Ecology*, *14*, pp. 4065–4083.

Blondel J, Aronson J, Bodiou J-Y. and Boeuf G. (2010) The Mediterranean Region. *Biological Diversity in Space and Time*. Oxford University Press Inc., New York, NY.

Boger, H. and Dermitzakis, D. M. (1985) Neogene paleogeography in the Central Aegean region. *VIIIth International Neogene Congress.* Budapest.

Boissier, E. (1867) *Aubrieta*. *Flora orientalis, 1*, pp. 249-254. Basel & Genève: H. Georg.

Bolson, M., de Camargo Smidt, E., Brotto, M. L. and 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), pp. e0143049.

Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M. and Twell, D. (2011) The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in *Arabidopsis*. *Plant Cell,23*(2), pp. 534-549.

Bornmüller, J.F.N. (1936) Aus der Pflanzenwelt des inner-iranischen Wüstengürtels. *Feddes Repertorium Specierum Novarum Regni Vegetabili*, 40, pp. 323-340, pl. 215.

Bowers, J. E., Chapman, B. A., Rong, J. and Paterson, A. H. (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature,422*(6930), pp. 433-438.

Brandt, P., Unseld, M., Eckert-Ossenkopp, U. and Brennicke, A. (1993) An rps14 pseudogene is transcribed and edited in *Arabidopsis* mitochondria. *Current genetics,24*(4), pp. 330-336.

Brudno, M., Do, C. B., Cooper, G. M., Kim, M. F., Davydov, E., Green, E. D., Sidow, A., Batzoglou, S. and Program, N. C. S. (2003) LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome research,13*(4), pp. 721-731.

Castilho, A. and Heslop-Harrison, J. (1995) Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. *Genome,38*(1), pp. 91-96.

Cavalier-Smith, T. (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). *Philosophical Transactions of the Royal Society of London B: Biological Sciences,358*(1429), pp. 109-134.

Chacón, J., Sousa, A., Baeza, C. M. and Renner, S. S. (2012) Ribosomal DNA distribution and a genus-wide phylogeny reveal patterns of chromosomal evolution in *Alstroemeria* (Alstroemeriaceae). *American Journal of Botany,99*(9), pp. 1501-1512.

Chase, M. W., Cowan, R. S., Hollingsworth, P. M., van den Berg, C., Madriñán, S., Petersen, G., Seberg, O., Jørgsensen, T., Cameron, K. M. and Carine, M. (2007) A proposal for a standardised protocol to barcode all land plants. *Taxon,56*(2), pp. 295-299.

Chatzimanolis, S., Trichas, A., Giokas, S. and Mylonas, M. (2003) Phylogenetic analysis and biogeography of Aegean taxa of the genus *Dendarus* (Coleoptera: Tenebrionidae). *Insect Systematics & Evolution,34*(3), pp. 295-312.

Choi, S., Kim, M., Lee, J. and Ryu, K. (2006) Genetic diversity and phylogenetic relationships among and within species of oriental cymbidiums based on RAPD analysis. *Scientia Horticulturae,108*(1), pp. 79-85.

Chumley, T. W., Palmer, J. D., Mower, J. P., Fourcade, H. M., Calie, P. J., Boore, J. L. and Jansen, R. K. (2006) The complete chloroplast genome sequence of *Pelargonium*× *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Molecular Biology and Evolution,23*(11), pp. 2175-2190.

Çiplak, B. (2003) Distribution of Tettigoniinae (Orthoptera, Tettigoniidae) bush-crickets in Turkey: the importance of the Anatolian Taurus Mountains in biodiversity and implications for conservation. *Biodiversity & Conservation,12*(1), pp. 47-64.

Comes, H. P., Tribsch, A. and Bittkau, C. (2008) Plant speciation in continental island floras as exemplified by *Nigella* in the Aegean Archipelago. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,363*(1506), pp. 3083-3096.

Conant, G. C. and Wolfe, K. H. (2008) GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics,24*(6), pp. 861-862.

Couvreur, T. L., Franzke, A., Al-Shehbaz, I. A., Bakker, F. T., Koch, M. A. and Mummenhoff, K. (2010) Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). *Molecular Biology and Evolution,27*(1), pp. 55-71.

Cronn, R., Liston, A., Parks, M., Gernandt, D. S., Shen, R. and Mockler, T. (2008) Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Research,36*(19), pp. e122.

Cullen, J. (1965) Aubrieta Adans., in Davis, P.H., ed. *Flora of Turkey and the East Aegean Islands, 1*, pp. 444-447. Edinburgh: Edinburgh University Press.

Daams, R. and Van Der Weerd, A. (1980) Early Pliocene small mammals from the Aegean island of Karpathos (Greece) and their paleogeographic significance. *Geologie en Mijnbouw,59*(4), pp. 327-331.

Davis, P. (1971) Distribution patterns in Anatolia with particular reference to endemism, in Davis, P.H., Harper, P.C. and Hedge, I.C. *Plant life of South west Asia*. Edinburgh: Botanical Society of Edinburgh.

Davis, P.H., Mill, R.R., Tan, K. (1988) *Flora of Turkey and the East Aegean Islands (Supplement)*, vol 10, pp. 50, Edinburgh: Edinburgh University.

De Montmollin, B. (1986) Etude cytotaxonomique de la flore de la Crète. III: Nombres chromosomiques. *Candollea,41*(2), pp. 431-439.

Demesure, B., Comps, B. and Petit, R. J. (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution,50*(6), pp. 2515-2520.

Demesure, B., Sodzi, N. and Petit, R. (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular ecology,4*(1), pp. 129-134.

Denton, A. L., McConaughy, B. L. and Hall, B. D. (1998) Usefulness of RNA polymerase II coding sequences for estimation of green plant phylogeny. *Molecular Biology and Evolution,15*(8), pp. 1082-1085.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., Del Angel, G., Rivas, M. A. and Hanna, M. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics,43*(5), pp. 491-498.

Dermitzakis, D. and Papanikolaou, D. (1981) Paleogeography and geodynamics of the Aegean region during the Neogene. in *Annales geologiques des pays helleniques,30*, pp. 245-289.

Dermitzakis, M. (1990) The evolution of the Aegeis during the Late Cenozoic. *Geologica Balcanica,20*(3), pp. 3-16.

Desjardins, P. and Conklin, D. (2010) NanoDrop microvolume quantitation of nucleic acids. *Journal of Visualized Experiments*, *45*, pp. e2565-e2564.

Döbel, P., Rieger, R. and Michaelis, A. (1973) The Giemsa banding patterns of the standard and four reconstructed karyotypes of *Vicia faba*. *Chromosoma,43*(4), pp. 409-422.

Doorduin, L., Gravendeel, B., Lammers, Y., Ariyurek, Y., Chin-A-Woeng, T. and Vrieling, K. (2011) The complete chloroplast genome of 17 individuals of pest species *Jacobaea*

vulgaris: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. *DNA research*, *18*, pp. 93-105.

Douris, V., Cameron, R. A., Rodakis, G. C. and Lecanidou, R. (1998) Mitochondrial Phylogeography of the Land Snail *Albinaria* in Crete: Long-Term Geolgoical and Short-Term Vicariance Effects. *Evolution,52*, pp. 116-125.

Douris, V., Rodakis, G., Giokas, S., Mylonas, M. and Lecanidou, R. (1995) Mitochondrial DNA and morphological differentiation of *Albinaria populations* (Gastropoda: Clausiliidae). *Journal of Molluscan Studies,61*(1), pp. 65-78.

Doyle, J. and Doyle, J. (1990) Isolation of DNA from small amounts of plant tissues. *BRL focus,12*, pp. 13-15.

Duarte, J. M., Wall, P. K., Edger, P. P., Landherr, L. L., Ma, H., Pires, J. C. and Leebens-Mack, J. (2010) Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC Evolutionary Biology,10*(1), pp. 61.

Dubchak, I. and Ryaboy, D. V. (2006) VISTA family of computational tools for comparative analysis of DNA sequences and whole genomes. *Gene Mapping, Discovery, and Expression: Methods and Protocols,338*, pp. 69-89.

Dubcovsky, J. and Dvorák, J. (1995) Ribosomal RNA multigene loci: nomads of the Triticeae genomes. *Genetics,140*(4), pp. 1367-1377.

Dubey, S., Cosson, J. F., Vohralik, V., Krystufek, B., Diker, E. and Vogel, P. (2007) Molecular evidence of Pleistocene bidirectional faunal exchange between Europe and the Near East: the case of the bicoloured shrew (*Crocidura leucodon*, Soricidae). *Journal of evolutionary biology,20*(5), pp. 1799-1808.

Dubey, S., Zaitsev, M., Cosson, J.-F., Abdukadier, A. and Vogel, P. (2006) Pliocene and Pleistocene diversification and multiple refugia in a Eurasian shrew (*Crocidura suaveolens* group). *Molecular phylogenetics and evolution,38*(3), pp. 635-647.

Duminil, J. and Di Michele, M. (2009) Plant species delimitation: a comparison of morphological and molecular markers. *Plant Biosystems,143*(3), pp. 528-542.

Dumolin-Lapègue, S., Demesure, B., Fineschi, S., Le Come, V. and Petit, R. (1997) Phylogeographic structure of white oaks throughout the European continent. *Genetics,146*(4), pp. 1475-1487.

Dumolin, S., Demesure, B. and Petit, R. (1995) Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theoretical and applied genetics,91*(8), pp. 1253-1256.

Du Rietz, G. E. (1930) The fundamental units of biological taxonomy. *Svensk Botanisk Tidskrift*, 24, pp. 33-428.

Dvorak, F and Dadakova B. (1984a) Chromosome counts and chromosome morphology of some selected species. *Scripta Fac. Sci. Nat. Univ. Park. Brun. Biology*, 14, pp. 463-468.

Dyall, S. D., Brown, M. T. and Johnson, P. J. (2004) Ancient invasions: from endosymbionts to organelles. *Science,304*(5668), pp. 253-257.

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Glöckner, G., Platts, A. E., Heckel, D. G., Der, J. P., Wafula, E. K. and Tang, M. (2015) The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences,112*(27), pp. 8362-8366.

Ekim, T. and Güner, A. (1986) The Anatolian Diagonal: fact or fiction? *Proceedings of the Royal Society of Edinburgh. Section B. Biological Sciences,89*, pp. 69-77.

Elder Jr, J. F. and Turner, B. J. (1995) Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology,70*, pp. 297-320.

Feliner, G.N. (2014) Patterns and processes in plant phylogeography in the Mediterranean Basin. A review. *Perspectives in Plant Ecology, Evolution and Systematics*, *16*(5), pp. 265-278.

Flanders, J., Jones, G., Benda, P., Dietz, C., Zhang, S., Li, G., Sharifi, M. and Rossiter, S. J. (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrumequinum*: contrasting results from mitochondrial and microsatellite data. *Molecular ecology,18*(2), pp. 306-318.

Flavell, R. (1980) The molecular characterization and organization of plant chromosomal DNA sequences. *Annual Review of Plant Physiology,31*(1), pp. 569-596.

Ford, V., Lee, J., Baldwin, B. and Gottlieb, L. (2006) Species divergence and relationships in *Stephanomeria* (Compositae): PgiC phylogeny compared to prior biosystematic studies. *American Journal of Botany,93*(3), pp. 480-490.

Fransz, P., Armstrong, S., Alonso-blanco, C., Fischer, T. C., Torres-ruiz, R. A. and Jones, G. (1998) Cytogenetics for the model system *Arabidopsis thaliana*. *The Plant Journal,13*(6), pp. 867-876.

Franzen, R. and Gustavsson, L.-Å. (1983) Chromosome numbers in flowering plants from the high mountains of Sterea Ellas, Greece. *Willdenowia,13*, pp. 101-106.

Franzke, A., German, D., Al-Shehbaz, I. A. and Mummenhoff, K. (2009) *Arabidopsis* family ties: molecular phylogeny and age estimates in Brassicaceae. *Taxon,58*(2), pp. 425-437.

Franzke, A., Lysak, M. A., Al-Shehbaz, I. A., Koch, M. A. and Mummenhoff, K. (2011) Cabbage family affairs: the evolutionary history of Brassicaceae. *Trends in Plant Science,16*(2), pp. 108-116.

Freudenstein, J. V. and Chase, M. W. (2015) Phylogenetic relationships in Epidendroideae (Orchidaceae), one of the great flowering plant radiations: progressive specialization and diversification. *Annals of Botany,115*(4), pp. 665-681.

Friesen, N., Fritsch, R. M., Pollner, S. and Blattner, F. R. (2000) Molecular and morphological evidence for an origin of the aberrant genus *Milula* within Himalayan species of *Allium* (Alliacae). *Molecular phylogenetics and evolution,17*(2), pp. 209-218.

Fritz, U., Ayaz, D., Hundsdörfer, A. K., Kotenko, T., Guicking, D., Wink, M., Tok, C. V., Çiçek, K. and Buschbom, J. (2009) Mitochondrial diversity of European pond turtles (*Emys orbicularis*) in Anatolia and the Ponto-Caspian Region: Multiple old refuges, hotspot of extant diversification and critically endangered endemics. *Organisms Diversity & Evolution,9*(2), pp. 100-114.

Frohlich, M. W. and Meyerowitz, E. M. (1997) The search for flower homeotic gene homologs in basal angiosperms and Gnetales: a potential new source of data on the evolutionary origin of flowers. *International Journal of Plant Sciences,158*, pp. 131-142.

Furman, A., Çoraman, E., Bilgin, R. and Karataş, A. (2009) Molecular ecology and phylogeography of the bent-wing bat complex (*Miniopterus schreibersii*)(Chiroptera: Vespertilionidae) in Asia Minor and adjacent regions. *Zoologica Scripta,38*(2), pp. 129-141.

Gan, Y., Liu, F., Chen, D., Wu, Q., Qin, Q., Wang, C., Li, S., Zhang, X., Wang, Y. and Wang, K. (2013) Chromosomal locations of 5S and 45S rDNA in *gossypium* genus and its phylogenetic implications revealed by FISH. *PloS one,8*(6), pp. e68207.

Gandolfo, M., Nixon, K. and Crepet, W. (1998) A new fossil flower from the Turonian of New Jersey: *Dressiantha bicarpellata* gen. et sp. nov.(Capparales). *American Journal of Botany,85*(7), pp. 964-964.

Gantenbein, B. and Largiadèr, C.R. (2002) *Mesobuthus gibbosus* (Scorpiones: Buthidae) on the island of Rhodes—hybridization between Ulysses' stowaways and native scorpions?. *Molecular Ecology*, *11*(5), pp.925-938.

Garcia, S., Garnatje, T. and Kovařík, A. (2012) Plant rDNA database: ribosomal DNA loci information goes online. *Chromosoma,121*(4), pp. 389-394.

Gerlach, W. and Bedbrook, J. (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Research,7*(7), pp. 1869-1885.

German, D., Garden, S.-S. B. and Friesen, N. (2015) *Shehbazia* (Shehbazieae, Cruciferae), a new monotypic genus and tribe of hybrid origin from Tibet. *Turczaninowia*, *17*, pp. 17–23.

Goetsch, L., Eckert, A. J. and Hall, B. D. (2005) The molecular systematics of *Rhododendron* (Ericaceae): a phylogeny based upon RPB2 gene sequences. *Systematic Botany,30*(3), pp. 616-626.

Gömöry, D., Paule, L. and Vyšnŷ, J. (2007) Patterns of allozyme variation in western Eurasian *Fagus*. *Botanical Journal of the Linnean Society,154*(2), pp. 165-174.

Goodwin, S., McPherson, J. D. and McCombie, W. R. (2016) Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics,17*(6), pp. 333-351.

Graham, L. E., Cook, M. E. and Busse, J. S. (2000) The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proceedings of the National Academy of Sciences,97*(9), pp. 4535-4540.

Grant, V. (1981) *Plant speciation.* New York: Columbia University Press.

Gray, M. W. (1999) Evolution of organellar genomes. *Current opinion in genetics & development,9*(6), pp. 678-687.

Group, C. P. W., Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M. W., Cowan, R. S. and Erickson, D. L. (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences,106*(31), pp. 12794-12797.

Gündüz, İ., Jaarola, M., Tez, C., Yeniyurt, C., Polly, P. D. and Searle, J. B. (2007) Multigenic and morphometric differentiation of ground squirrels (Spermophilus, Scuiridae, Rodentia) in Turkey, with a description of a new species. *Molecular phylogenetics and evolution,43*(3), pp. 916-935.

Gür, H. 2016. The Anatolian diagonal revisited: Testing the ecological basis of a biogeographic boundary. *Zoology in the Middle East,* 62, 189-199.

Gustavsson, L. A. (1986) *Aubrieta* Adanson, in Strid, A., ed. *Mountain Flora of Greece*, 1, pp. 268-274. Cambridge University Press.

Hahn, W. J. (2002) A phylogenetic analysis of the Arecoid line of palms based on plastid DNA sequence data. *Molecular phylogenetics and evolution,23*(2), pp. 189-204.

Hallden, C., Bryngelsson, T., Säll, T. and Gustafsson, M. (1987) Distribution and evolution of a tandemly repeated DNA sequence in the family Brassicaceae. *Journal of molecular evolution,25*(4), pp. 318-323.

Hamilton, M. (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. *Molecular ecology,8*(3), pp. 521-523.

Hamilton, M. B., Braverman, J. M. and Soria-Hernanz, D. F. (2003) Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. *Molecular Biology and Evolution,20*(10), pp. 1710-1721.

Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Research,31*(20), pp. 5907-5916.

Handa, H., Bonnard, G. r. and Grienenberger, J.-M. (1996) The rapeseed mitochondrial gene encoding a homologue of the bacterial protein Ccl1 is divided into two independently transcribed reading frames. *Molecular and General Genetics,252*(3), pp. 292-302.

Hasterok, R., Ksiazczyk, T., Wolny, E. and Maluszynska, J. (2005) FISH and GISH analysis of *Brassica* genomes. *Acta Biologica Cracoviensia Botanica,47*(1), pp. 185-192.

Hasterok, R., Wolny, E., Hosiawa, M., Kowalczyk, M., Kulak-Ksiazczyk, S., Ksiazczyk, T., Heneen, W. K. and Maluszynska, J. (2006) Comparative analysis of rDNA distribution in chromosomes of various species of Brassicaceae. *Annals of Botany,97*(2), pp. 205-216.

Haudry, A., Platts, A. E., Vello, E., Hoen, D. R., Leclercq, M., Williamson, R. J., Forczek, E., Joly-Lopez, Z., Steffen, J. G. and Hazzouri, K. M. (2013) An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nature Genetics,45*(8), pp. 891-898.

Hausknecht, C. (1893) Symbolae ad floram graecam. Aufzählung der im Sommer 1885 in Griechenland gesammelten Pflanzen. *Mittheilungen des Thüringischen Botanischen Vereines* 3-4: 96.

Heslop-Harrison, J. and Schmidt, T. (2012) Plant nuclear genome composition. *eLS*. Chichester: John Wiley & Sons.

Hohmann, N., Wolf, E. M., Lysak, M. A. and Koch, M. A. (2015) A time-calibrated road map of Brassicaceae species radiation and evolutionary history. *Plant Cell,27*(10), pp. 2770-2784.

Hollingsworth, P. M., Graham, S. W. and Little, D. P. (2011) Choosing and using a plant DNA barcode. *PloS one,6*(5), pp. e19254.

Hsu, K. (1972) Origin of saline giants: a critical review after the discovery of the Mediterranean evaporite. *Earth-Science Reviews,8*(4), pp. 371-396.

Hsü, K. J., Montadert, L., Bernoulli, D., Cita, M. B., Erickson, A., Garrison, R. E., Kidd, R. B., Mèlierés, F., Müller, C. and Wright, R. (1977) History of the Mediterranean salinity crisis. *Nature,267*(5610), pp. 399-403.

Hu, S., Sablok, G., Wang, B., Qu, D., Barbaro, E., Viola, R., Li, M. and Varotto, C. (2015) Plastome organization and evolution of chloroplast genes in *Cardamine* species adapted to contrasting habitats. *BMC genomics,16*(1), pp. 1.

Hu, Z.-Y., Hua, W., Huang, S.-M. and Wang, H.-Z. (2011) Complete chloroplast genome sequence of rapeseed (*Brassica napus* L.) and its evolutionary implications. *Genetic Resources and Crop Evolution,58*(6), pp. 875-887.

Huang, C.-H., Sun, R., Hu, Y., Zeng, L., Zhang, N., Cai, L., Zhang, Q., Koch, M. A., Al-Shehbaz, I. and Edger, P. P. (2016) Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Molecular Biology and Evolution,33*(2), pp. 394-412.

Hunkapiller, T., Kaiser, R., Koop, B. and Hood, L. (1991) Large-scale and automated DNA sequence determination. *Science,254*(5028), pp. 59-67.

Hutchison, C. A. (2007) DNA sequencing: bench to bedside and beyond. *Nucleic Acids Research*,*35*(18), pp. 6227-6237.

Iwabe, N., Kuma, K. I., Kishino, H., Hasegawa, M. and Miyata, T. (1991) Evolution of RNA polymerases and branching patterns of the three major groups of archaebacteria. *Journal of molecular evolution*,*32*(1), pp. 70-78.

Jalas, J., Suominen, J. and Lampinen, R. (1994) Atlas Florae Europaeae 10 Cruciferae (*Sisymbrium* to *Aubrieta*). *Committee for Mapping the Flora of Europe, Helsinki*.

Janchen, E. (1942) Das system der Cruciferen. *Plant Systematics and Evolution*,*91*(1), pp. 1-28.

Jeanson, M. L., Labat, J.-N. and Little, D. P. (2011) DNA barcoding: a new tool for palm taxonomists? *Annals of Botany*,108, pp. 1445-1451.

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., Tomsho, L. P., Hu, Y., Liang, H. and Soltis, P. S. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature*,*473*(7345), pp. 97-100.

Johnson, L. A. and Soltis, D. E. (1994) matK DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany*,*19*, pp. 143-156.

Joly, S., Heenan, P. B. and Lockhart, P. J. (2009) A Pleistocene inter-tribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. *Molecular phylogenetics and evolution*,*51*(2), pp. 365-372.

Kagale, S., Robinson, S. J., Nixon, J., Xiao, R., Huebert, T., Condie, J., Kessler, D., Clarke, W. E., Edger, P. P. and Links, M. G. (2014) Polyploid evolution of the Brassicaceae during the Cenozoic era. *Plant Cell*,*26*(7), pp. 2777-2791.

Kamm, A., Galasso, I., Schmidt, T. and Heslop-Harrison, J. (1995) Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between *Arabidopsis* species. *Plant molecular biology*,*27*(5), pp. 853-862.

Karl, R., Kiefer, C., Ansell, S. W. and Koch, M. A. (2012) Systematics and evolution of arctic-alpine *Arabis alpina* (Brassicaceae) and its closest relatives in the eastern Mediterranean. *American Journal of Botany*,*99*(4), pp. 778-794.

Karl, R. and Koch, M. A. (2013) A world-wide perspective on crucifer speciation and evolution: phylogenetics, biogeography and trait evolution in tribe Arabideae. *Annals of Botany,112*, pp. 983-1001.

Kasapidis, P., Magoulas, A., Mylonas, M. and Zouros, E. (2005) The phylogeography of the gecko *Cyrtopodion kotschyi* (Reptilia: Gekkonidae) in the Aegean archipelago. *Molecular Phylogenetics and Evolution*, *35*(3), pp.612-623.

Kasapidis, P., Suchentrunk, F., Magoulas, A. and Kotoulas, G. (2005) The shaping of mitochondrial DNA phylogeographic patterns of the brown hare (*Lepus europaeus*) under the combined influence of Late Pleistocene climatic fluctuations and anthropogenic translocations. *Molecular Phylogenetics and Evolution*, *34*(1), pp.55-66.

Katoh, K., Misawa, K., Kuma, K. I. and Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research,30*(14), pp. 3059-3066.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S. and Duran, C. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,28*(12), pp. 1647-1649.

Kelchner, S. A. (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden,87*, pp. 482-498.

Kim, K.-J. and Jansen, R. K. (1995) ndhF sequence evolution and the major clades in the sunflower family. *Proceedings of the National Academy of Sciences,92*(22), pp. 10379-10383.

Kim, K.-J. and Lee, H.-L. (2004) Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA research,11*(4), pp. 247-261.

Kim, Y.-K., Park, C.-w. and Kim, K.-J. (2009) Complete chloroplast DNA sequence from a Korean endemic genus, *Megaleranthis saniculifolia*, and its evolutionary implications. *Molecules and cells,27*(3), pp. 365-381.

Koch, M., Bishop, J. and Mitchell-Olds, T. (1999) Molecular systematics and evolution of *Arabidopsis* and *Arabis*. *Plant Biology,1*(5), pp. 529-537.

Koch, M. A., Dobeš, C., Kiefer, C., Schmickl, R., Klimeš, L. and Lysak, M. A. (2007) Supernetwork identifies multiple events of plastid trnF (GAA) pseudogene evolution in the Brassicaceae. *Molecular Biology and Evolution,24*(1), pp. 63-73.

Koch, M.A., Haubold, B. and Mitchell-Olds, T. (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis, Arabis,* and related genera (Brassicaceae). *Molecular biology and evolution*, 17(10), pp.1483-1498.

Koch, M., Haubold, B. and Mitchell-Olds, T. (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *American Journal of Botany*, 88(3), pp.534-544.

Koch, M. A., Karl, R., German, D. A. and Al-Shehbaz, I. A. (2012) Systematics, taxonomy and biogeography of three new Asian genera of Brassicaceae tribe Arabideae: An ancient distribution circle around the Asian high mountains. *Taxon,61*, pp. 955-969.

Koch, M. A., Karl, R., Kiefer, C. and Al-Shehbaz, I. A. (2010) Colonizing the American continent: systematics of the genus *Arabis* in North America (Brassicaceae). *American Journal of Botany,97*(6), pp. 1040-1057.

Koch, M.A., Kiefer, C., Ehrich, D., Vogel, J., Brochmann, C. and Mummenhoff, K. (2006) Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology*, *15*(3), pp.825-839.

Kolodziej, P. A., Woychik, N., Liao, S. and Young, R. A. (1990) RNA polymerase II subunit composition, stoichiometry, and phosphorylation. *Molecular and Cellular Biology,10*(5), pp. 1915-1920.

Krijgsman, W., Hilgen, F., Raffi, I., Sierro, F. and Wilson, D. (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature,400*(6745), pp. 652-655.

Kučera, J., Lihova, J. and Marhold, K. (2006) Taxonomy and phylogeography of *Cardamine impatiens* and *C. pectinata* (Brassicaceae). *Botanical Journal of the Linnean Society,152*(2), pp. 169-195.

Kulak, S., Hasterok, R. and Maluszynska, J. (2002) Karyotyping of *Brassica* amphidiploids using 5S and 25S rDNA as chromosome markers. *Hereditas,136*(2), pp. 144-150.

Kupfer, P. (1980) Contribution a la cytotaxonomie de quelques orophytes iraniens. *Biologie et ecologie méditerranéenne,* 7.

Kuss, S. E. (1975) Die pleistozänen Hirsche der ostmediterranen Inseln Kreta, Kasos, Karpathos und Rhodos (Griechenland). *Berichte der naturforschenden Gesellschaft zu Freiburg im Breisgau,65*(1-2), pp. 25-79.

Lafontaine, D. L. and Tollervey, D. (2001) The function and synthesis of ribosomes. *Nature Reviews Molecular Cell Biology,2*(7), pp. 514-520.

Lakshmikumaran, M. and Ranade, S. (1990) Isolation and characterization of a highly repetitive DNA of *Brassica campestris*. *Plant molecular biology,14*(3), pp. 447-448.

Larkin, R. and Guilfoyle, T. (1993) The second largest subunit of RNA polymerase II from *Arabidopsis thaliana*. *Nucleic Acids Research,21*(4), pp. 1038.

Levin, D. A. (2002) *The role of chromosomal change in plant evolution.* Oxford: Oxford University Press on Demand.

Levinson, G. and Gutman, G. A. (1987) Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution,4*(3), pp. 203-221.

Lewis, H. (1966) Speciation in flowering plants. *Science,152*(3719), pp. 167-172.

Lippman, Z., Gendrel, A.-V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., Lavine, K., Mittal, V., May, B. and Kasschau, K. D. (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature,430*(6998), pp. 471-476.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., Zhao, M., Ma, J., Yu, J. and Huang, S. (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature communications,5*.

Liu, Y., Huo, N., Dong, L., Wang, Y., Zhang, S., Young, H. A., Feng, X. and Gu, Y. Q. (2013) Complete chloroplast genome sequences of Mongolia medicine *Artemisia frigida* and phylogenetic relationships with other plants. *PloS one,8*(2), pp. e57533.

Lu, S.-Y., Peng, C.-I., Cheng, Y.-P., Hong, K.-H. and Chiang, T.-Y. (2001) Chloroplast DNA phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of Taiwan. *Genome,44*(5), pp. 797-807.

Lykousis, V. (2009) Sea-level changes and shelf break prograding sequences during the last 400ka in the Aegean margins: subsidence rates and palaeogeographic implications. Continental Shelf Research,29(16), pp. 2037-2044.

Lymberakis, P. and Poulakakis, N. (2010) Three continents claiming an archipelago: the evolution of Aegean's herpetofaunal diversity. *Diversity*, *2*(2), pp. 233–255.

Lymberakis, P., Poulakakis, N., Manthalou, G., Tsigenopoulos, C.S., Magoulas, A. and Mylonas, M.(2007) Mitochondrial phylogeography of *Rana* (*Pelophylax*) populations in the Eastern Mediterranean region. *Molecular phylogenetics and evolution*, *44*(1), pp.115-125.

Lynch, M. and Conery, J. S. (2000) The evolutionary fate and consequences of duplicate genes. *Science,290*(5494), pp. 1151-1155.

Lysak, M. A. and Koch, M. A. (2011) Phylogeny, genome, and karyotype evolution of crucifers (Brassicaceae), in Schmidt, R. and Bancroft, I., eds. *Genetics and Genomics of the Brassicaceae*, pp. 1-31. Berlin: Springer.

Lysak, M. A., Koch, M. A., Pecinka, A. and Schubert, I. (2005) Chromosome triplication found across the tribe Brassiceae. *Genome research,15*(4), pp. 516-525.

Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L. L. and Hernández-Hernández, T. (2015) A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist,207*(2), pp. 437-453.

Magyari, E., Chapman, J., Gaydarska, B., Marinova, E., Deli, T., Huntley, J., Allen, J. and Huntley, B. (2008) The 'oriental'component of the Balkan flora: evidence of presence on the Thracian Plain during the Weichselian late-glacial. *Journal of Biogeography,35*(5), pp. 865-883.

Maluszynska, J. and Hasterok, R. (2005) Identification of individual chromosomes and parental genomes in *Brassica juncea* using GISH and FISH. *Cytogenetic and genome research,109*(1-3), pp. 310-314.

Maluszynska, J. and Heslop-Harrison, J. (1993) Physical mapping of rDNA loci in *Brassica* species. *Genome,36*(4), pp. 774-781.

Maluszynska, J. and Heslop-Harrison, J. (1991) Localization of tandemly repeated DMA sequences in *Arabidopsis thaliana*. *The Plant Journal,1*(2), pp. 159-166.

Mandáková, T., Heenan, P. B. and Lysak, M. A. (2010a) Island species radiation and karyotypic stasis in Pachycladon allopolyploids. *BMC Evolutionary Biology,10*(1), pp. 367.

Mandáková, T., Joly, S., Krzywinski, M., Mummenhoff, K. and Lysak, M. A. (2010b) Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. *Plant Cell,22*(7), pp. 2277-2290.

Mandáková, T., Mummenhoff, K., Al-Shehbaz, I. A., Mucina, L., Mühlhausen, A. and Lysak, M. A. (2012) Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). *Taxon,61*(5), pp. 989-1000.

Mardis, E. R. (2008) The impact of next-generation sequencing technology on genetics. *Trends in genetics,24*(3), pp. 133-141.

Margulis, L. and Bermudes, D. (1984) Symbiosis as a mechanism of evolution: status of cell symbiosis theory. *Symbiosis (Philadelphia, Pa.),1*, pp. 101-124.

Marienfeld, J., Unseld, M. and Brennicke, A. (1999) The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information. *Trends in Plant Science,4*(12), pp. 495-502.

Marks, G. and Schweizer, D. (1974) Giemsa banding: karyotype differences in some species of *Anemone* and in *Hepatica nobilis*. *Chromosoma,44*(4), pp. 405-416.

Mattfeld, J. (1939) The species of the genus *Aubrieta* Adanson. *Quarterly Bulletinof the Alpine Garden Society,7*, pp. 157-181.

Melodelima, C. and Lobréaux, S. (2013) Complete *Arabis alpina* chloroplast genome sequence and insight into its polymorphism. *Meta gene,1*, pp. 65-75.

Meulenkamp, J. (1971) The Neogene in the southern Aegean area. *Opera Botanica,30*, pp. 5-12.

Michaux, J., Libois, R., Paradis, E. and Filippucci, M.-G. (2004) Phylogeographic history of the yellow-necked fieldmouse (Apodemus flavicollis) in Europe and in the Near and Middle East. *Molecular phylogenetics and evolution,32*(3), pp. 788-798.

Modest, E. and Sengupta, S. (1973) Chemical correlates of chromosome banding, in Caspersson, T. and Zech, L., eds. *Chromosome identification—technique and applications in biology and medicine*, pp. 327-334. New York: Academic Press.

Montmollin, B. de (1982) Etude cytotaxonomique de la flore endemique de la Crete. I. Note preliminaire. *Bulletin de la Societe Neuchateloise des Sciences Naturelles. Quatrieme serie*, 105, pp. 65-77.

Moore, M. J., Bell, C. D., Soltis, P. S. and Soltis, D. E. (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences,104*(49), pp. 19363-19368.

Moore, M. J., Soltis, P. S., Bell, C. D., Burleigh, J. G. and Soltis, D. E. (2010) Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences,107*(10), pp. 4623-4628.

Mort, M. E. and Crawford, D. J. (2004) The continuing search: low-copy nuclear sequences for lower-level plant molecular phylogenetic studies. *Taxon,53*(2), pp. 257-261.

Muellner, A., Schaefer, H. and Lahaye, R. (2011) Evaluation of candidate DNA barcoding loci for economically important timber species of the mahogany family (Meliaceae). *Molecular Ecology Resources,11*(3), pp. 450-460.

Naydenov, K., Senneville, S., Beaulieu, J., Tremblay, F. and Bousquet, J. (2007) Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology,7*(1), pp. 233.

Nock, C. J., Waters, D. L., Edwards, M. A., Bowen, S. G., Rice, N., Cordeiro, G. M. and Henry, R. J. (2011) Chloroplast genome sequences from total DNA for plant identification. *Plant biotechnology journal,9*(3), pp. 328-333.

Novikova, P. Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., Muir, G., Guggisberg, A., Paape, T., Schmid, K. and Fedorenko, O. M. (2016) Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. *Nature Genetics, 488,* pp. 1077-1082.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y. and Kohchi, T. (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. *Journal of molecular biology,223*(1), pp. 1-7.

Ogihara, Y., Terachi, T. and Sasakuma, T. (1988) Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proceedings of the National Academy of Sciences,85*(22), pp. 8573-8577.

Ohsako, T. and Ohnishi, O. (2000) Intra-and interspecific phylogeny of wild *Fagopyrum* (Polygonaceae) species based on nucleotide sequences of noncoding regions in chloroplast DNA. *American Journal of Botany,87*(4), pp. 573-582.

Ohsako, T. and Ohnishi, O. (2001) Nucleotide sequence variation of the chloroplast trnK/matK region in two wild *Fagopyrum* (Polygonaceae) species, F. leptopodum and F. statice. *Genes & genetic systems,76*(1), pp. 39-46.

Ohta, T. and Dover, G. A. (1983) Population genetics of multigene families that are dispersed into two or more chromosomes. *Proceedings of the National Academy of Sciences,80*(13), pp. 4079-4083.

Oxelman, B. and Bremer, B. (2000) Discovery of paralogous nuclear gene sequences coding for the second-largest subunit of RNA polymerase II (RPB2) and their phylogenetic utility in Gentianales of the asterids. *Molecular Biology and Evolution,17*(8), pp. 1131-1145.

Oxelman, B., Yoshikawa, N., McConaughy, B. L., Luo, J., Denton, A. L. and Hall, B. D. (2004) RPB2 gene phylogeny in flowering plants, with particular emphasis on asterids. *Molecular phylogenetics and evolution,32*(2), pp. 462-479.

Palmer, J. D. (1990) Contrasting modes and tempos of genome evolution in land plant organelles. *Trends in genetics,6*, pp. 115-120.

Palmer, J. D. (1991) Plastid chromosomes: structure and evolution. *The molecular biology of plastids,7*, pp. 5-53.

Palmer, J. D. (1992) Comparison of chloroplast and mitochondrial genome evolution in plants. in Herrmann, R.G., ed. *Cell organelles*, pp. 99-133. Berlin: Springer.

Palmer, J. D., Adams, K. L., Cho, Y., Parkinson, C. L., Qiu, Y.-L. and Song, K. (2000) Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proceedings of the National Academy of Sciences,97*(13), pp. 6960-6966.

Palmer, J. D., Jansen, R. K., Michaels, H. J., Chase, M. W. and Manhart, J. R. (1988) Chloroplast DNA variation and plant phylogeny. *Annals of the Missouri Botanical Garden,75*, pp. 1180-1206.

Parent, C.E., Caccone, A. and Petren, K. (2008) Colonization and diversification of Galapagos terrestrial fauna: a phylogenetic and biogeographical synthesis. *Philosophical Transactions of The Royal Society B Biological Sciences*, 363, pp. 3347–3361.

Parkin, I. A., Gulden, S. M., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C. and Lydiate, D. J. (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with Arabidopsis thaliana. *Genetics,171*(2), pp. 765-781.

Parkin, I. A., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., Town, C. D., Nixon, J., Krishnakumar, V. and Bidwell, S. L. (2014) Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome biology,15*(6), pp. 1.

Parks, M., Cronn, R. and Liston, A. (2009) Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC biology,7*(1), pp. 1.

Parmakelis, A., Pfenninger, M., Spanos, L., Papagiannakis, G., Louis, C. and Mylonas, M. (2005) Inference of a radiation in Mastus (Gastropoda, Pulmonata, Enidae) on the island of Crete. *Evolution*, 59(5), pp.991-1005.

Parmakelis, A., Stathi, I., Chatzaki, M., Simaiakis, S., Spanos, L., Louis, C. and Mylonas, M. (2006a) Evolution of *Mesobuthus gibbosus* (Brullé, 1832)(Scorpiones: Buthidae) in the northeastern Mediterranean region. *Molecular Ecology*, *15*(10), pp.2883-2894.

Parmakelis, A., Stathi, I., Spanos, L., Louis, C. and Mylonas, M. (2006b) Phylogeography of *Iurus dufoureius* (Brulle, 1832)(Scorpiones, Iuridae). *Journal of Biogeography*, *33*(2), pp.251-260.

Parmakelis, A., Klossa-kilia , E., Kilias, G., Triantis, K.A. and Sfenthourakis, S. (2008) Increased molecular divergence of two endemic *Trachelipus* (Isopoda, Oniscidea) species from Greece reveals patterns not congruent with current taxonomy. *Biological Journal of the Linnean Society*, *95*(2), pp.361-370.

Phitos, D. (1970) Die Gattung *Aubrieta* in Griechenland. *Candollea*, 25, pp. 68-87.

Phitos, D. (2002) *Aubrieta* Adans., in Strid, A. and Tan, K., eds. *Flora Hellenica, 2*, pp. 192-195. Koenisgtein: Koeltz Scientific Books.

Posada, D. and Crandall, K. A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics,14*(9), pp. 817-818.

Potter, D., Gao, F., Baggett, S., McKenna, J. R. and McGranahan, G. H. (2002) Defining the sources of Paradox: DNA sequence markers for North American walnut (*Juglans L.*) species and hybrids. *Scientia Horticulturae,94*(1), pp. 157-170.

Poulakakis, N., Lymberakis, P., Antoniou, A., Chalkia, D., Zouros, E., Mylonas, M. and Valakos, E. (2003) Molecular phylogeny and biogeography of the wall-lizard *Podarcis erhardii* (Squamata: Lacertidae). *Molecular phylogenetics and evolution,28*(1), pp. 38-46.

Poulakakis, N., Kapli, P., Lymberakis, P., Trichas, A., Vardinoyiannis, K., Sfenthourakis, S. and Mylonas, M. (2015) A review of phylogeographic analyses of animal taxa from the Aegean and surrounding regions. *Journal of Zoological Systematics and Evolutionary Research*, *53*(1), pp.18-32.

Poulakakis, N., Lymberakis, P., Tsigenopoulos, C.S., Magoulas, A. and Mylonas, M. (2005a) Phylogenetic relationships and evolutionary history of snake-eyed skink *Ablepharus kitaibelii* (Sauria: Scincidae). *Molecular Phylogenetics and Evolution*, *34*(2), pp.245-256.

Poulakakis, N., Lymberakis, P., Valakos, E., Pafilis, P., Zouros, E. and Mylonas, M. (2005b) Phylogeography of Balkan wall lizard (*Podarcis taurica*) and its relatives inferred from mitochondrial DNA sequences. *Molecular Ecology*, *14*(8), pp.2433-2443.

Poulakakis, N., Lymberakis, P., Valakos, E., Zouros, E. and Mylonas, M. (2005c) Phylogenetic relationships and biogeography of *Podarcis* species from the Balkan Peninsula, by Bayesian and maximum likelihood analyses of mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, 37(3), pp.845-857.

Poulakakis, N., Antoniou, A., Mantziou, G., Parmakelis, A., Skartsi, T., Vasilakis, D., Elorriaga, J., De la Puente, J., Gavashelishvili, A., Ghasabyan, M., Katzner, T., McGrady, M., Batbayar, N., Fuller, M. and Natsagdorj, T. (2008a) Population structure, diversity, and phylogeography in the near-threatened Eurasian black vultures *Aegypius monachus* (Falconiformes; Accipitridae) in Europe: insights from microsatellite and mitochondrial DNA variation. *The Biological Journal of the Linnean Society, 95*,pp.859–872.

Poulakakis, N., Pakaki, V., Mylonas, M. and Lymberakis, P. (2008b) Molecular phylogeny of the Greek legless skink *Ophiomorus punctatissimus* (Squamata: Scincidae): the impact of the Mid-Aegean trench in its phylogeography. *Molecular Phylogenetics and Evolution, 47*, pp.396–402.

Preisler, R. S. and Thompson, W. F. (1981) Evolutionary sequence divergence within repeated DNA families of higher plant genomes. *Journal of molecular evolution,17*(2), pp. 85-93.

Price, R. A., Palmer, J. D. and Al-Shehbaz, I. A. (1994) 1 Systematic Relationships of *Arabidopsis*: A Molecular and Morphological Perspective. *Cold Spring Harbor Monograph Archive,27*, pp. 7-19.

Puertas, M., Badaeva, E., Dedkova, O., Gay, G., Pukhalskyi, V., Zelenin, A., Bernard, S. and Bernard, M. (2007) Chromosomal rearrangements in wheat: their types and distribution. *Genome,50*(10), pp. 907-926.

Ramsey, J. and Schemske, D. W. (2002) Neopolyploidy in flowering plants. *Annual review of ecology and systematics*, pp. 589-639.

Raskina, O., Barber, J., Nevo, E. and Belyayev, A. (2008) Repetitive DNA and chromosomal rearrangements: speciation-related events in plant genomes. *Cytogenetic and genome research,120*(3-4), pp. 351-357.

Ren, B. Q., Xiang, X. G. and Chen, Z. D. (2010) Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. *Molecular Ecology Resources,10*(4), pp. 594-605.

Rieseberg, L. H. (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution,16*(7), pp. 351-358.

Rogers, S. O. and Bendich, A. J. (1987) Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant molecular biology,9*(5), pp. 509-520.

Rokas, A., Atkinson, R. J., Webster, L., Csóka, G. and Stone, G. N. (2003) Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-Mediterranean oak gallwasp Andricus quercustozae. *Molecular ecology,12*(8), pp. 2153-2174.

Roncal, J., Francisco-Ortega, J., Asmussen, C. B., Lewis, C. E. and Hoot, S. B. (2005) Molecular phylogenetics of tribe Geonomeae (Arecaceae) using nuclear DNA sequences of phosphoribulokinase and RNA polymerase II. *Systematic Botany,30*(2), pp. 275-283.

Ronquist, F. and Huelsenbeck, J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics,19*(12), pp. 1572-1574.

Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U. M., Chaudhary, L. B., Datt, B., Bag, S. K., Singh, P. K. and Nair, N. K. (2010) Universal plant DNA barcode loci may not work in complex groups: a case study with Indian Berberis species. *PloS one,5*(10), pp. e13674.

Russell, A., Samuel, R., Klejna, V., Barfuss, M. H., Rupp, B. and Chase, M. W. (2010) Reticulate evolution in diploid and tetraploid species of *Polystachya* (Orchidaceae) as shown by plastid DNA sequences and low-copy nuclear genes. *Annals of Botany,106*(1), pp. 37-56.

Sakai, M., Kanazawa, A., Fujii, A., Thseng, F., Abe, J. and Shimamoto, Y. (2003) Phylogenetic relationships of the chloroplast genomes in the genus *Glycine* inferred from four intergenic spacer sequences. *Plant Systematics and Evolution,239*(1-2), pp. 29-54.

Sang, T. (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry and Molecular Biology,37*(3), pp. 121-147.

Sang, T., Crawford, D. and Stuessy, T. (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany,84*(8), pp. 1120-1120.

Sanger, F. (1988) Sequences, sequences, and sequences. *Annual review of biochemistry,57*(1), pp. 1-29.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Tabata, S. (1999) Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA research,6*(5), pp. 283-290.

Sawadogo, M. and Sentenac, A. (1990) RNA polymerase B (II) and general transcription factors. *Annual review of biochemistry,59*(1), pp. 711-754.

Schranz, M. E., Mohammadin, S. and Edger, P. P. (2012) Ancient whole genome duplications, novelty and diversification: the WGD Radiation Lag-Time Model. *Current opinion in plant biology,15*(2), pp. 147-153.

Schubert, I. and Wobus, U. (1985) In situ hybridization confirms jumping nucleolus organizing regions in *Allium*. *Chromosoma,92*(2), pp. 143-148.

Schulz, O. (1936) Cruciferae, in Engler, A. and Prantl, K., eds. *Die natürlichen Pflanzenfamilen, 2*, pp. 227-658. Leipzig: Engelmann.

Schwarzacher, T., Leitch, A., Bennett, M. and Heslop-Harrison, J. (1989) In situ localization of parental genomes in a wide hybrid. *Annals of Botany,64*(3), pp. 315-324.

Seol, Y.-J., Kim, K., Kang, S.-H., Perumal, S., Lee, J. and Kim, C.-K. (2015) The complete chloroplast genome of two *Brassica* species, *Brassica nigra* and *B. Oleracea*. *Mitochondrial DNA*, pp. 1-2.

Sharma, S. and Raina, S. (2005) Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. *Cytogenetic and genome research,109*(1-3), pp. 15-26.

Sharma, S. K., Dkhar, J., Kumaria, S., Tandon, P. and 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), pp. 10-15.

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., Siripun, K. C., Winder, C. T., Schilling, E. E. and Small, R. L. (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany,92*(1), pp. 142-166.

Shendure, J. and Ji, H. (2008) Next-generation DNA sequencing. *Nature biotechnology,26*(10), pp. 1135-1145.

Shimamoto, Y., Hasegawa, A., Abe, J., Ohara, M. and Mikami, T. (1992) *Glycine soja* germplasm in Japan: isozyme and chloroplast DNA variation. *Soybean genetics newsletter-US Department of Agriculture, Agricultural Research Service (USA),* 19 , pp. 73–77

Slatkin, M. (1986) Interchromosomal biased gene conversion, mutation and selection in a multigene family. *Genetics,112*(3), pp. 681-698.

Song, K., Lu, P., Tang, K. and Osborn, T. C. (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences,92*(17), pp. 7719-7723.

Steele, K. P. and Vilgalys, R. (1994) Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene matK. *Systematic Botany,19*, pp. 126-142.

Stevens, P. (2001>) Angiosperm Phylogeny Website. Version 12, July 2012 [and more or less continuously updated since]. http://www.mobot.org/MOBOT/research/APweb/

Stojanov, N. and Stefanov, B. (1924) *Flora of Bulgaria*. Vol 1. Sofia: State Printing House. (In Bulgarian).

Stojanov, N. and Stefanov, B. (1933) *Flora of Bulgaria*. Ed. 2. Sofia: Guttenberg Press. (In Bulgarian).

Stojanov, N. and Stefanov, B. (1948) *Flora of Bulgaria*. Ed. 3. Sofia: University Press. (In Bulgarian).

Stojanov, N., Stefanov, B. and Kitanov, B. (1966) *Flora of Bulgaria*. Ed. 4, vol. 1. Sofia: Nauka and Izkustvo. (In Bulgarian).

Sun, G., Ni, Y. and Daley, T. (2008) Molecular phylogeny of RPB2 gene reveals multiple origin, geographic differentiation of H genome, and the relationship of the Y genome to other genomes in *Elymus* species. *Molecular phylogenetics and evolution,46*(3), pp. 897-907.

Sun, G., Pourkheirandish, M. and Komatsuda, T. (2009) Molecular evolution and phylogeny of the RPB2 gene in the genus *Hordeum*. *Annals of Botany, 103*,pp. 975-983.

Swerdlow, H., Wu, S., Harke, H. and Dovichi, N. J. (1990) Capillary gel electrophoresis for DNA sequencing: laser-induced fluorescence detection with the sheath flow cuvette. *Journal of Chromatography A,516*(1), pp. 61-67.

Taberlet, P., Gielly, L., Pautou, G. and Bouvet, J. (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant molecular biology,17*(5), pp. 1105-1109.

Taketa, S., Harrison, G. and Heslop-Harrison, J. (1999) Comparative physical mapping of the 5S and 18S-25S rDNA in nine wild *Hordeum* species and cytotypes. *Theoretical and applied genetics,98*(1), pp. 1-9.

Tamura, K. (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. *Molecular Biology and Evolution,9*(4), pp. 678-687.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*,30, pp. 2725-2729.

Tate, J. A. and Simpson, B. B. (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany,28*(4), pp. 723-737.

Templeton, A. R. (1981) Mechanisms of speciation--a population genetic approach. *Annual review of ecology and systematics,12*, pp. 23-48.

Terakami, S., Matsumura, Y., Kurita, K., Kanamori, H., Katayose, Y., Yamamoto, T. and Katayama, H. (2012) Complete sequence of the chloroplast genome from pear (*Pyrus pyrifolia*): genome structure and comparative analysis. *Tree genetics & genomes,8*(4), pp. 841-854.

Tetlow, I. J., Rawsthorne, S., Raines, C. and Emes, M. J. (2005) 3 Plastid metabolic pathways. *Plastids,13*, pp. 60.

Tian, X. and Li, D. (2001) Application of DNA sequences in plant phylogenetic study. *Acta Botanica Yunnanica,24*(2), pp. 170-184.

Tourloukis, V. and Karkanas, P. (2012) The Middle Pleistocene archaeological record of Greece and the role of the Aegean in hominin dispersals: new data and interpretations. *Quaternary Science Reviews,43*, pp. 1-15.

Triantis, K.A. and Mylonas, M. (2009) Greek islands biology. In: Gillespie R,

Glague DA (eds), *Encyclopedia of Islands.* University of California Press, Berkeley, CA, pp. 388–392.

Tsumura, Y., Yoshimura, K., Tomaru, N. and Ohba, K. (1995) Molecular phytogeny of conifers using RFLP analysis of PCR-amplified specific chloroplast genes. *Theoretical and applied genetics,91*(8), pp. 1222-1236.

Turmel, M., Otis, C. and Lemieux, C. (2003) The mitochondrial genome of *Chara vulgaris*: insights into the mitochondrial DNA architecture of the last common ancestor of green algae and land plants. *Plant Cell,15*(8), pp. 1888-1903.

Ueda, M., Nishikawa, T., Fujimoto, M., Takanashi, H., Arimura, S.-i., Tsutsumi, N. and Kadowaki, K.-i. (2008) Substitution of the gene for chloroplast RPS16 was assisted by generation of a dual targeting signal. *Molecular Biology and Evolution,25*(8), pp. 1566-1575.

Umesono, K., Inokuchi, H., Ohyama, K. and Ozeki, H. (1984) Nucleotide sequence of Marchantia polymorpha chloroplast DNA: a region possibly encoding three tRNAs and three proteins including a homologue of *E. coli* ribosomal protein S14. *Nucleic Acids Research,12*(24), pp. 9551-9565.

Unseld, M., Marienfeld, J. R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924. *Nature Genetics,15*, pp. 57-61.

Van Droogenbroeck, B., Kyndt, T., Maertens, I., Romeijn-Peeters, E., Scheldeman, X., Romero-Motochi, J., Van Damme, P., Goetghebeur, P. and Gheysen, G. (2004) Phylogenetic analysis of the highland papayas (Vasconcellea) and allied genera (Caricaceae) using PCR-RFLP. *Theoretical and applied genetics,108*(8), pp. 1473-1486.

Van Loon, J. and Oudemans, J. (1982) IOPB chromosome number reports LXXV, ed. Á. Löve. *Taxon,31*, pp. 342-368.

von Hayek, A. (1911) Entwurf eines Cruciferen-systems auf phylogenetischer grundlage. Botanisches Zentralblatt,*27*, pp. 127-335.

Vosa, C. G. (1973) Heterochromatin recognition and analysis of chromosome variation in *Scilla sibirica*. *Chromosoma,43*(3), pp. 269-278.

Vosa, C. G. (1976) Heterochromatic banding patterns in *Allium*. *Chromosoma,57*(2), pp. 119-133.

Voss, A. (1896) *Aubrieta*, in Siebert, A. and Voss, A. eds. *Vilmorin's Blumengärtnerei. Beschreibung, Kultur und Verwendung des gesamten Pflanzenmaterials für deutsche Gärten*, vol. 1, pp. 75-76.

Wakasugi, T., Sugita, M., Tsudzuki, T. and Sugiura, M. (1998) Updated gene map of tobacco chloroplast DNA. *Plant Molecular Biology Reporter,16*(3), pp. 231-241.

Walbot, V. and Goldberg, R. (1979) Plant genome organization and its relationship to classical plant genetics. *Nucleic acids in plants,1*, pp. 3-40.

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R. and Soltis, D.E. (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences*, 106(10), pp.3853-3858.

Wang, R., Farrona, S., Vincent, C., Joecker, A., Schoof, H., Turck, F., Alonso-Blanco, C., Coupland, G. and Albani, M.C. (2009) PEP1 regulates perennial flowering in *Arabis alpina*. *Nature*, 459, pp. 423-427.

Wang, L., Wuyun, T.-n., Du, H., Wang, D. and Cao, D. (2016) Complete chloroplast genome sequences of *Eucommia ulmoides*: genome structure and evolution. *Tree genetics & genomes,12*(1), pp. 1-15.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.-H., Bancroft, I. and Cheng, F. (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nauret Genetics,43*(10), pp. 1035-1039.

Ward, B. L., Anderson, R. S. and Bendich, A. J. (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell,25*(3), pp. 793-803.

Warrilow, D. and Symons, R. H. (1996) Sequence analysis of the second largest subunit of tomato RNA polymerase II. *Plant molecular biology,30*(2), pp. 337-342.

Weisrock, D.W., Macey, J.R., Ugurtas, I.H., Larson, A. and Papenfuss, T.J. (2001) Molecular phylogenetics and historical biogeography among salamandrids of the "true" salamander clade: rapid branching of numerous highly divergent lineages in Mertensiella luschani associated with the rise of Anatolia. *Molecular phylogenetics and evolution, 18*(3), pp.434-448.

Wendel, J. F., Schnabel, A. and Seelanan, T. (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences,92*(1), pp. 280-284.

White, M. J. D. (1978) *Modes of speciation.* San Francisco: WH Freeman.

Wikström, N., Savolainen, V. and Chase, M. W. (2001) Evolution of the angiosperms: calibrating the family tree. *Proceedings of the Royal Society of London B: Biological Sciences,268*(1482), pp. 2211-2220.

Wolfe, K. H., Gouy, M., Yang, Y.-W., Sharp, P. M. and Li, W.-H. (1989) Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proceedings of the National Academy of Sciences,86*(16), pp. 6201-6205.

Wolfe, K. H., Li, W.-H. and Sharp, P. M. (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences,84*(24), pp. 9054-9058.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B. and Rieseberg, L. H. (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences,106*(33), pp. 13875-13879.

Wu, F.-H., Chan, M.-T., Liao, D.-C., Hsu, C.-T., Lee, Y.-W., Daniell, H., Duvall, M. R. and Lin, C.-S. (2010) Complete chloroplast genome of *Oncidium* 'Gower Ramsey' and evaluation of molecular markers for identification and breeding in Oncidiinae. *BMC plant biology,10*(1), pp. 68.

Wyman, S. K., Jansen, R. K. and Boore, J. L. (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics,20*(17), pp. 3252-3255.

Xu, D., Abe, J., Kanazawa, A., Gai, J. and Shimamoto, Y. (2001) Identification of sequence variations by PCR-RFLP and its application to the evaluation of cpDNA diversity in wild and cultivated soybeans. *Theoretical and applied genetics,102*(5), pp. 683-688.

Yamane, K., Yano, K. and Kawahara, T. (2006) Pattern and rate of indel evolution inferred from whole chloroplast intergenic regions in sugarcane, maize and rice. *DNA research,13*(5), pp. 197-204.

Yang, Y.W., Lai, K.N., Tai, P.Y. and Li, W.H. (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *Journal of Molecular Evolution*, 48(5), pp.597-604.

Yang, Y.W., Tai, P.-Y., Chen, Y. and Li, W.H. (2002) A study of the phylogeny of *Brassica rapa*, *B. nigra*, *Raphanus sativus*, and their related genera using noncoding regions of chloroplast DNA. *Molecular phylogenetics and evolution,23*(2), pp. 268-275.

Yüzbaşıoğlu, S., Koch, M. A. and Al-Shehbaz, I. I. (2015) Proof of a knowledge database concept. *Aubrieta ekimii* (Brassicaceae), a new species from NW Anatolia (Turkey): morphological and molecular support. *Plant Systematics and Evolution,301*(8), pp. 2043-2055.

Zhang, D. and Sang, T. (1999) Physical mapping of ribosomal RNA genes in *peonies* (Paeonia, Paeoniaceae) by fluorescent in situ hybridization: implications for phylogeny and concerted evolution. *American Journal of Botany,86*(5), pp. 735-740.

Ziolkowski, P. A., Kaczmarek, M., Babula, D. and Sadowski, J. (2006) Genome evolution in *Arabidopsis*/*Brassica*: conservation and divergence of ancient rearranged segments and their breakpoints. *The Plant Journal,47*(1), pp. 63-74.

Zunk, K., Mummenhoff, K. and Hurka, H. (2000) Phylogenetic relationships in tribe Lepidieae (Brassicaceae) based on chloroplast DNA restriction site variation. *Canadian Journal of Botany,77*(10), pp. 1504-1512.