THE TAXONOMY OF *SENECIO* SECT. *SENECIO*: HYBRIDISATION AND SPECIATION

Mark A. Chapman

A thesis submitted to the
University of St Andrews for
The degree of Doctor of Philosophy

School of Biology
University of St Andrews
August 2004

ABSTRACT

The role of hybridisation as an evolutionary force was examined within a group of species of *Senecio* sect. *Senecio* L (Asteraceae). Conspecific pollen advantage as a reproductive isolating mechanism was found to be low or absent in two naturally hybridising *Senecio* species (*S. aethnensis* and *S. chrysanthemifolius*) on Mt. Etna, Sicily, indicating that other forces such as environmental-dependent selection maintain species integrity at this site. Hybrid material from the hybrid zone on Mt. Etna is thought to have given rise to British *S. squalidus* within the past 300 years, and the role of one of these hybridising species, *S. chrysanthemifolius*, as a parent of *S. squalidus* was confirmed by the analysis of sequence variation of two nuclear genes (*Scyc2* and *PgiC*). DNA sequence variation of these two genes also confirms that *S. squalidus* was the diploid parent of the British hexaploid *S. cambrensis*. Comparison of DNA sequence variation obtained from *S. cambrensis* and a related hexaploid, *S. teneriffae*, provide good evidence that *S. vulgaris* was the tetraploid parent of both; however, *S. leucanthemifolius* was implicated as the diploid parent of *S. teneriffae*. In addition, the analysis of sequence variation supported the hypothesis that *S. vulgaris* is an allotetraploid and that its parent species were likely to have been *S. vernalis* and a species closely related to *S. leucanthemifolius*. In a related investigation, the parentage of tetraploid *S. mohavensis* was also investigated using sequence variation of *Scyc2* and *PgiC*. *S. mohavensis* exhibits a morphology very similar to *S. flavus*, possibly indicating an autotetraploid origin; however the results produced firm evidence that it was an allotetraploid of *S. flavus* and *S. glaucus*. A survey of RAPD and chloroplast DNA variation in diploid *Senecio* material of uncertain taxonomic status that occurs on sand dunes in southern Sicily, and also in related material, confirmed that it was an isolated form of *S. glaucus*. Morphological divergence is likely to be due to natural selection and/or genetic drift.

DECLARATION

I, Mark Chapman, hereby certify that this thesis, which is approximately 33,000 words in length, has been written by me, and that it is a record of work carried out by me, and that it has not been submitted in any previous application for a higher degree.

Mark Chapman

August, 2004

STATEMENT

I was admitted as a research student in October 2000 and as a candidate for the degree of PhD in October 2001; the higher study for which this is a record was carried out in the University of St Andrews between 2000 and 2004.

Mark Chapman

August, 2004

CERTIFICATE

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this in application for that degree.

Richard J. Abbott

August, 2004

COPYRIGHT

ACKNOWLEDGEMENTS

First and foremost I'd like to thank Richard Abbott for getting me through the last four years of this rollercoaster ride known as a PhD! Richard has provided some well-needed encouragement through the gels and PCRs that didn't (and possibly wouldn't ever....) work. Huge thanks also go to Max, Amanda and Richard M. for the technical advice and teaching of lab techniques which got me up and running. Also, to Dave Forbes, whom unselfishly genotyped about 4,000 plants for me over a period of about eight weeks. This apparently is a new world record.

A big thank you goes to all the people at the conferences I've met who've been interested in my research, often more so after a couple of drinks – Hannah, Holly and Alex in Nova Scotia, and Rob and Andy at EGG. Rob and I came up with a revolutionary technique which could save hundreds of people and thousands of pounds. Unfortunately this had been published six months before. Rob, we will one day come up with a technique that no-one else has thought of.

Outside the lab, I couldn't possibly thank everyone who's made the last four years so much fun, but id like to personally mention Mireille for scientific, financial, emotional (and physical) support when needed, and also Holly and Belle, without whom Id probably not have gone outside during the last few months of this thesis! Also, to Kimberley for constant amusement and interest in my thesis; Ruth for the Irish stew and Guinness; Hugo for University Challenge and being a top housemate; Ali, Tom, Betsy and Derwent for Balgove House parties and free rhubarb; and the second best quiz team in St Andrews - Susi, Gordon, Nathan and Kirsten and the rest of the team. One day we shall win.....

CONTENTS

CHAPTER 1 - GENERAL INTRODUCTION

1.1 HISTORICAL PERSPECTIVES ON HYBRIDISATION

1.1.1 *The importance of hybridisation in evolution*

The contribution that hybridisation has made to the origin and evolution of species has been the subject of considerable debate over many years (Arnold, 1997). Early opinion, that interspecific hybridisation is detrimental and evolutionarily insignificant (Dobzhansky, 1940; Mayr, 1942), was derived from observations of hybrid unfitness and sterility, and the presence in species of an array of mechanisms that reduce or prevent the production of hybrid offspring. This view, however, which was primarily adopted by zoologists, was challenged by botanists who advanced the alternative viewpoint that crosses between individuals identified as distinct species can sometimes produce viable, fertile offspring, which may in turn create new genetic variation in an existing species, and/or be a step in the origin of new species (Anderson and Stebbins, 1954; Stebbins, 1959; Grant, 1981).

Many botanists in the first half of the 20th century recognised that hybridisation was common and widespread in the plant kingdom, and was potentially a major force in plant evolution (Anderson and Hubricht, 1938; Anderson, 1949; Heiser, 1949; Grant, 1953; Anderson and Stebbins, 1954). Lotsy (1916, 1931, in Arnold, 1997) argued that evolutionary change was possible through, if not primarily caused by, hybridisation. Lotsy (1925) also introduced a term to describe a group of taxa recognised as separate species that were interfertile, the *syngameon*, defined by Grant (1957, in Grant, 1981, p. 54) as "the sum total of species or semispecies linked by frequent or occasional hybridisation in nature; [hence] a hybridising group of species..."

During the same period, zoological opinion remained that interspecific hybridisation was maladaptive and gave rise to unfit progeny that were lost before becoming established (Dobzhansky, 1937; Mayr, 1942). Dobzhansky (1940) argued that interspecific hybridisation was so detrimental that it led to the reinforcement of reproductive isolating mechanisms between species due to the reduced fitness of

hybrids. The contrasting views of botanists (such as Grant, Stebbins and Anderson) and zoologists (Dobzhansky and Mayr) were largely maintained until the 1990s.

More recently, theoretical and empirical studies have shown that hybrids are not uniformly less fit than parental genotypes; instead, a subset of hybrid genotypes can sometimes be as fit as the parents and, under certain conditions, may become established. This has meant that a significant role for hybridisation in evolution is now recognised in both the animal and plant kingdoms (Arnold, 1992; Dowling and Secor, 1997; Rieseberg, 1997). The consequences of hybridisation are also of interest to the general public who may be concerned with the threat to native flora and fauna posed by introduced species (Chornesky and Randall, 2003), and the creation of 'franken-weeds' following the escape of transgenes from genetically-modified crops (Jorgensen *et al.*, 1996; Stewart *et al.*, 2000).

1.1.2 *Species concepts and hybridisation*

A satisfying, all-encompassing, definition of 'a species' has eluded evolutionary biologists since the time of Darwin (Mallet, 1995). Darwin's classification (Darwin, 1859) was materialistic, categorising individuals on the basis of shared morphological characteristics that exhibited discrete differences from individuals of other species. This did not account for examples of morphological indistinguishable species that are completely reproductively isolated from each other. Thus a species concept based on an absence of interbreeding was advanced by Dobzhansky (1937) and later refined by Mayr (1942) as the biological species concept (BSC). This defined species as "... groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups."

The BSC (Mayr, 1942) is undoubtedly the most utilised of species concepts, potentially due to the ease with which species can be defined. Further concepts have been proposed as improvements to the BSC, however none have been universally adopted (Hey, 2001). Under some species concepts, including the BSC, interspecific hybridisation is, by definition, not possible, due to complete reproductive isolation from other similarly-defined species, hence if two individuals can cross, they are conspecific. Other species concepts take into account the occurrence of interspecific hybridisation in nature, but relegate the species involved to the status of 'bad species',

as hybridisation is seen to violate the basic assumptions of divergent evolution (e.g. Templeton, 1989).

The prevalence of interspecific hybridisation in nature (reviewed in Arnold, 1992; Dowling and Secor, 1997; Rieseberg, 1997), however, suggests that either (1) the above species concepts do not suitably define species, or that (2) many species are 'bad species' (i.e. are not reproductively isolated). With regard to the BSC, Stace (1989, p. 145) has argued that in the plant kingdom "...there is now little adherence to such a naïve definition of a species." and that species capable of hybridising "...are the *usual* situation, and cannot be considered in any way abnormal." As highlighted by Rieseberg and Carney (1998), hybrid zones (see section 1.4.2) act as barriers to large-scale gene flow, and hence the taxa involved, although not reproductively isolated, are genetically isolated, and hence may be considered 'good' species.

1.2 DETECTION OF HYBRIDS

1.2.1 *Morphology*

A hybrid individual inherits morphological traits from both parent species and may be expected to be intermediate to them in character mean. This logical assumption, however, is not always borne out in practice. From a compilation of studies which measured morphological characters in hybrid individuals, Rieseberg and Ellstrand (1993) reported that an F1 hybrid was no more likely to be intermediate than parental-like in character mean. In fact, morphologically intermediate populations may not always be products of hybridisation because such populations can be ancestral and morphologically intermediate to the 'parent' species in question. Moreover, a hybrid individual may converge in morphology similar to that of one parent and therefore appear to be non-hybrid based on morphology. Also, a population that is genetically isolated from other populations of the same species could, through genetic drift and natural selection (Wright, 1978; Slatkin, 1985), diverge morphologically to produce a group of individuals that are morphologically intermediate to the parent species and another species, thus indicating wrongly a hybrid origin.

Coleman and Abbott (2003) recently investigated the possibility that in Morocco, populations of *Senecio leucanthemifolius* var. *casablancae* at the southern

margins of the species range exhibited leaf morphology intermediate to parapatric *S. glaucus* ssp. *coronopifolius* due to hybridisation. However, no evidence for hybridisation was revealed by a survey of nuclear genetic marker variation. Therefore, although morphology might act as an indicator that hybridisation played a role in the formation of a population, it cannot be relied upon as proof of hybrid ancestry. Hypotheses concerning the hybrid origin of the diploid species *Helianthus anomalus* (Heiser, 1947) and *Iris nelsonii* (Randolph, 1966) were proposed on morphological grounds and were later confirmed by the results of analyses of molecular marker variation (*H. anomalus*, Rieseberg, 1991; *I. nelsonii*, Arnold, 1993). However, in some cases morphology may have indicated a non-hybrid origin which was later disproved by molecular marker analysis (e.g. Watson *et al.*, 1991; Bleeker *et al.*, 1999).

1.2.2 *Molecular markers*

The advent of molecular techniques has revolutionised the study of hybrids and hybridisation, primarily due to their ability to detect hybrids, which are assumed to inherit molecular markers from both parents and display an additive marker profile. Many different molecular techniques exist for determining the presence of hybridisation in the phylogenetic history of a plant species (Table 1.1). For example, isozymes have been used to confirm the hybrid origin of *Stephanomeria diegensis* (Gallez and Gottlieb, 1982) and the introgressive origin of *Senecio vulgaris* var. *hibernicus* (Abbott *et al.*, 1992). Similarly, random PCR-generated markers have been used to determine the parentage of diploid (Arnold, 1993; Crawford *et al.*, 1993; Wolfe *et al.*, 1998) and polyploid hybrid species (Brochmann *et al.*, 1998; Ayres and Strong, 2001). More recently, DNA sequencing has allowed investigations of hybrid ancestry to resolve the parentage of a number of allopolyploids (e.g. Doyle *et al.*, 1999; Ford and Gottlieb, 1999; Baumel *et al.*, 2001).

Table 1.1 – A sample of studies of homoploid (diploid) and polyploid species which have utilised molecular markers to determine hybrid ancestry.

Origin	Species	Markers[1]	Reference
Homoploid	*Encelia virginensis*	PCR, cpR	Allan *et al.*, 1997
Homoploid	*Gossypium bickii*	I, cpR, rR	Wendel *et al.*, 1991
Homoploid	*Helianthus anomalus*	rR	Rieseberg, 1991
Homoploid	*Helianthus deserticola*	rR	Rieseberg, 1991
Homoploid	*Helianthus paradoxus*	rR	Rieseberg, 1991
Homoploid	*Iris nelsonii*	PCR, cpR	Arnold *et al.*, 1991; Arnold, 1993
Homoploid	*Margyracaena skottsbergii*	PCR	Crawford *et al.*, 1993
Homoploid[2]	*Paeonia officinalis*	nS	Ferguson and Sang, 2001
Homoploid	*Penstemon clevelandii*	PCR	Wolfe *et al.*, 1998
Homoploid	*Pinus densata*	I, cpR	Wang *et al.*, 1990; Wang and Szmidt, 1994
Homoploid	*Senecio squalidus*	I, PCR	James, 1999; Abbott *et al.*, 2000, 2002
Homoploid	*Stephanomeria diegensis*	I	Gallez and Gottlieb, 1982
Polyploid	*Arabidopsis suecica*	I, cpR, rS	Mummenhoff and Hurka, 1995; O'Kane *et al.*, 1996
Polyploid	*Arachis hypogaea*	nR	Kochert *et al.*, 1996
Polyploid	*Cardamine insueta*	rS	Franzke and Mummenhoff, 1999
Polyploid	*Cardamine schulzii*	rS	Franzke and Mummenhoff, 1999
Polyploid	*Clarkia gracilis*	nS	Ford and Gottlieb, 1999
Polyploid	*Glycine* spp.	nS	Doyle *et al.*, 2004
Polyploid	*Glycine tabacina*	nS	Doyle *et al.*, 1999
Polyploid	*Gossypium gossypioides*	rS	Wendel *et al.*, 1995b
Polyploid	Hawaiian silverswords	nS	Barrier *et al.*, 1999
Polyploid	*Lepidium* spp.	nS	Lee *et al.*, 2002
Polyploid	*Marshallia mohrii*	I	Watson *et al.*, 1991
Polyploid	*Nasturtium microphyllum*	I	Bleeker *et al.*, 1999
Polyploid	*Saxifraga rivularis*	PCR, cpS, rS	Brochmann *et al.*, 1998
Polyploid	*Saxifraga svalbardensis*	PCR, cpS, rS	Brochmann *et al.*, 1998
Polyploid	*Senecio cambrensis*	I, cpR	Ashton and Abbott, 1992a; Lowe and Abbott, 1996
Polyploid	*Senecio teneriffae*	I, cpR	Lowe and Abbott, 1996
Polyploid	*Spartina anglica*	I, PCR, cpR, rR	Raybould *et al.*, 1991; Ayres and Strong, 2001; Baumel *et al.*, 2001
Polyploid	*Tragopogon* spp.	I, cpR	Roose and Gottlieb, 1976; Soltis and Soltis, 1989

[1] I, Isozymes; PCR, random PCR-generated RAPDs or ISSRs; cp, chloroplast DNA; r, ribosomal DNA; n, nuclear (single- or low-copy) DNA; S, sequencing; R, RFLP

[2] Parents were allotetraploids

The choice of marker will largely depend on the application and the funds available. Analysis of more than one type of genetic marker is recommended, especially if hybridisation is considered to be likely in the group of species analysed (Small *et al.*, 1998; Soltis and Soltis, 1998). For example, 'alien' cytoplasmic markers (chloroplast (cp) and/or mitochondrial (mt) DNA) have been discovered in a wide range of plant species and have been attributed to ancient hybridisation and the retention of non-recombining cytoplasmic DNA from another species (e.g. Soltis *et al.*, 1991; Soltis and Kuzoff, 1995; VanRaamsdonk *et al.*, 1997; Jackson *et al.*, 1999). In these studies, analysis of a single nuclear or cytoplasmic marker may not have revealed the true phylogenetic relationships between taxa (reviewed in Wendel and Doyle, 1998).

1.3 REPRODUCTIVE ISOLATING MECHANISMS

For a species to maintain its genetic integrity it is necessary for it to be sufficiently reproductively isolated from other species (Grant, 1981; Ridley, 1993). Taxa separated geographically may be fully interfertile when crossed under glasshouse conditions because barriers to gene flow have not evolved in the wild. On the other hand, taxa that occur in close proximity are likely to have evolved characters that prevent the formation of ill-adapted hybrids. If these characters establish a pre-zygotic barrier to interspecific hybridisation, then additional intrinsic barriers to hybridisation may be weak and such taxa might also be crossed experimentally.

Closely related species which are capable of hybridising can co-exist provided that assortative mating, i.e. the preferential mating of conspecifics, occurs. This can be achieved in several ways. For example, heterospecific pollen transfer will be prevented if each species has its own particular pollinator (ethological isolation), if each species flowers at a different time (temporal isolation), or if one or both species reproduce by selfing.

Commonly a plant species will not be totally reproductively isolated from related species by assortative mating, and thus some interspecific pollen transfer may occur. Hence, mechanisms are often present which prevent the production of hybrids following the arrival of heterospecific pollen on the stigma. These mechanisms may

also be pre-zygotic, i.e. acting before gamete fusion, or be post-zygotic, i.e. occurring after the fusion of gametes.

1.3.1 *Pre-zygotic reproductive isolating mechanisms*

Geographical separation of two species eliminates the chance for heterospecific pollen transfer. In addition, ecological isolation of two species may provide a strong barrier to hybridisation. Anderson and Hubricht (1938) suggested that certain *Tradescantia* species were once sympatric, but reproductively isolated due to ecological differentiation; however, human activity disrupted barriers and has lead to hybridisation. Additionally, introgression in Louisiana iris species (Arnold *et al.*, 1990a) and *Cypripedium* orchids (Klier *et al.*, 1991) appears to have occurred as a direct result of human disturbance of environments.

Seasonal (and temporal) isolation occurs between species, which although parapatric or sympatric, do not cross simply because flowers are not open at the same time of year (or day). Limited overlap in flowering periods reduces the potential for gene flow between diploid and tetraploid forms of *Arrhenatherum elatius* (Petit *et al.*, 1997) and between *Iris fulva* and *I. brevicaulis* (Cruzan and Arnold, 1994). In addition, the flowers of two interfertile species may differ sufficiently in structure, preventing pollen being transported by a vector from one species to another (e.g. Goldblatt *et al.*, 1998; P. Wolf *et al.*, 2001). If flowers of different species open at the same time, the behaviour of pollinators may reduce or prevent heterospecific pollen arriving on the stigma of an individual. Grant (1992) suggests that *Aquilegia formosa* and *A. pubescens* are isolated due to ethology, the former species being pollinated mainly by hummingbirds and the latter almost exclusively by hawkmoths (although see Chase and Raven, 1975). To understand the effect of pollinator preference in a natural hybrid zone, Emms and Arnold (2000) introduced flowering stems of *Iris hexagona*, *I. fulva* and F1 hybrids into populations that were monospecific for a wild-growing *Iris*. Pollinators preferred the native species; the least visited was the other species, with F1 hybrids intermediate in attracting pollinators. Moreover, inter-individual pollinator movement indicated that pollen was more often transported from parent species to F1s than it was between species or between F1 individuals (Emms and Arnold, 2000). This was concordant with the results of a study into the rate of

natural formation of F1, F2 and backcross progeny in which backcross progeny formed at a rate much higher than did F1 hybrids (Hodges *et al.*, 1996).

A comprehensive study of the causes and genetic control of reproductive isolation between the monkeyflowers *Mimulus cardinalis* and *M. lewisii* has been carried out using a QTL (quantitative trait locus) mapping strategy. Strong reproductive isolation between these two species results from pollinator preference (Schemske and Bradshaw, 1999). *M. lewisii* is bumblebee pollinated with pink flowers, yellow nectar guides and a small volume of nectar; whereas hummingbird pollinated *M. cardinalis* has bright red petals and produces a large volume of nectar. QTL analysis of the traits associated with this pollinator preference revealed at least one major QTL, which accounted for more than 25% of the phenotypic variance, for each of nine of the 12 traits studied (Bradshaw *et al.*, 1998). Potentially, speciation could have occurred through mutations occurring in three of these major loci: the first increasing carotenoid deposition in the petals, the second increasing nectar volume and the third making the corolla shape more efficient for hummingbird-mediated pollen transfer (Bradshaw *et al.*, 1995). The magnitude of the effect of the locus controlling the yellow pigmentation was studied by breeding lines of each species containing the alternative allele for that locus. Lines of *M. cardinalis* containing the *M. lewisii* allele received 74-fold more bumblebee visitors than the wildtype, while lines of *M. lewisii* with the *M. cardinalis* allele received 68-fold more hummingbird visits (Bradshaw and Schemske, 2003).

Many plant species produce flowers that are not served by a single pollinator species and, by relying on generalist animal pollinators or wind pollination, receive pollen loads comprising a mixture of conspecific (i.e. same species) and heterospecific (i.e. another species) pollen. Such plants may exhibit a mechanism to distinguish, and preferentially accept, conspecific pollen at the expense of heterospecific pollen resulting in what is called conspecific pollen advantage (CPA; Alarcón and Campbell, 2000). This may be caused by heterospecific pollen exhibiting (1) reduced germination on the stigma, (2) retarded pollen tube growth in the style or (3) decreased fertilisation of ovules, relative to conspecific pollen.

CPA has been shown to reduce significantly the frequency of hybrid seed formation following mixed pollinations in a number of interfertile species pairs studied (Arnold *et al.*, 1993; Hauser *et al.*, 1997; Klips, 1999). In addition, the

strength of CPA may be asymmetric, i.e. differ depending on which species acted as the maternal parent (e.g. Kiang and Hamrick, 1978; Carney *et al.*, 1994; Rieseberg *et al.*, 1995a; Emms *et al.*, 1996; Carney and Arnold, 1997; Diaz and MacNair, 1999). In a study concerning two naturally hybridising *Ipomopsis* species, CPA did not act to reduce the proportion of hybrid seed formed in mixed pollinations (Alarcón and Campbell, 2000), however hybrid pollen was less successful in siring seeds than parental pollen, indicating a mechanism by which the formation of later-generation hybrids is reduced (Campbell *et al.*, 2003).

1.3.2 *Post-zygotic reproductive isolating mechanisms*

Hybrid unfitness and/or sterility appear to play a significant role in the maintenance of species boundaries (e.g. Coyne, 1996; Lopez *et al.*, 2000; Milne *et al.*, 2003). Hybrid incompatibility may be manifested as a high frequency of ovule abortion or the production of unfit or sterile offspring. In some studies of conspecific pollen advantage, an increase in the proportion of heterospecific pollen applied was accompanied by a decrease in seed set, indicating non-random abortion of hybrid zygotes (Carney *et al.*, 1994; Hauser *et al.*, 1997).

The evolution of post-zygotic isolation between species was proposed by Dobzhansky (1937) to involve a gene from one species interacting negatively in the genetic background of another, rendering the hybrid inviable or sterile. This 'inviability' gene may be repressed in the source species; however, should the gene introgress, repression is removed and the deleterious effect is expressed (Wu and Palopoli, 1994). The alternative view is that hybrid sterility is caused by the cumulative effect of several genes of relatively small effect interacting negatively in a hybrid genome (Wu and Palopoli, 1994). In rice, a small number of 'supergenes' appear to be responsible for hybrid sterility in the F1 generation (Li *et al.*, 1997). However, a larger number of genic factors are responsible for hybrid breakdown in later generations following the break-up of coadapted gene complexes (Li *et al.*, 1997).

The genic model of hybrid inviability was independently modelled by Bateson, Dobzhansky and Muller (BDM model; reviewed in Burke and Arnold, 2001). Essentially, an ancestral population becomes subdivided and in each a novel

allele arises which is incompatible with the new allele in the other population. Hence, following secondary contact, the viability of hybrids is reduced. The model has been developed further to account for duplicate genes (Lynch and Force, 2000) or polyploid species (Werth and Windham, 1991). In each of these two recent models, one copy of a locus becomes silenced in the first subpopulation and the other in the second subpopulation. If the two populations come into secondary contact, a portion of gametes or offspring will inherit both non-functional copies of the locus and show reduced viability/fertility.

Evidence supporting the negative interaction of genes from different parent species in interspecific hybrids comes from studies of *Iris* (Burke *et al.*, 1998) and *Gossypium* (Jiang *et al.*, 2000). Burke *et al.* (1998) generated two F2 populations following crosses between *Iris fulva* and *I. brevicaulis*. Molecular analysis of the F2 populations revealed a significant deficit in the proportion of intermediate genotypes and an excess in the numbers of parental-like genotypes, indicating intrinsic (i.e. genic) selection against hybrid genotypes (Burke *et al.*, 1998). In the study by Jiang *et al.* (2000), 24 backcross families were generated from an initial cross between *G. hirsutum* and *G. barbadense* and genotyped for 262 RFLP markers. A deficit from the expected number of introgressed *G. barbadense* markers was found; indeed, seven chromosomal regions were entirely absent in the backcross families (Jiang *et al.*, 2000).

Further evidence that the recombination of genes or linkage groups that distinguish species will result in lowered fitness in hybrids comes from studies on *Lens* and *Helianthus*. A cross between two *Lens* species revealed that hybrids heterozygous for a translocation that differentiated the species had pollen viability <65%, whereas hybrid plants homozygous for the translocation had pollen viabilities > 85 % (Tadmor *et al.*, 1987). A similar result was obtained from a cross between *Helianthus argophyllus* and *H. annuus* in which three genetic intervals correlated with pollen viability and meiotic abnormalities (Quillet *et al.*, 1995). Rieseberg *et al.* (1995b, 1996a) compared rates of introgression in an interspecific backcross population between *H. annuus* and *H. petiolaris*. Previous analysis of the two species identified seven chromosomes which are collinear, and another ten that differ in gene order (Rieseberg *et al.*, 1995c).

Introgression in the backcross population was much greater in the collinear portions of the genome: 40% of the genome from the collinear chromosomes introgressed, whereas only 2.4% from the structurally diverged chromosomes introgressed (Rieseberg *et al.*, 1995b, 1996a). The reduction in introgression of structurally divergent chromosomes is most likely due to selection acting against recombinant gametes; hence, introgressive hybridisation is reduced in these species due to chromosomal structural differences. The finding that 60% of the structurally collinear linkage groups also failed to introgress suggests that incompatible genetic factors are also contained within regions on which selection will act (Rieseberg *et al.*, 1995b).

Further analysis by Rieseberg *et al.* (1999b) of three natural hybrid zones revealed 26 *H. petiolaris* chromosome segments that were negatively selected against in an *H. annuus* background. Of these, 16 were significantly associated with pollen sterility (Rieseberg *et al.*, 1999b).

Despite findings cited above of significant negative interactions in interspecific hybrids between genes that distinguish the parents, some positive interactions between such genes were also detected. This highlights the potential for interspecific hybridisation to generate a portion of recombinant hybrid individuals that do not show reduced fitness, relative to the parents. In this respect, Arnold and Hodges (1995) compiled a list of 44 studies which compared hybrid and parental fitness. In most studies (24), hybrids achieved an overall average fitness equal to or higher than the parent species. In only 13 cases did hybrids exhibit fitness lower than both parents. However, Arnold and Hodges (1995) acknowledge that usually only one or two hybrid classes were measured for one trait (e.g. survivorship), hence the results cannot be considered conclusive. For example, classification of *Iris* hybrids into several genotypic classes (8) revealed that relative fitness of the different classes was variable, and that some hybrid classes were more fit than the parent species, a result that would not have been recognised if all hybrids were combined into one class (Arnold and Hodges, 1995).

It is also clear that hybrid fitness may vary in different environments (Stebbins and Daly, 1961; Cruzan and Arnold, 1993, 1994; Arnold, 1997). For example, Grant and Grant (1993) demonstrated that a change in the size of seeds available following

an El Niño event resulted in greater fitness of a finch hybrid relative to its parent species in the Galapagos Islands.

1.4 THE CONSEQUENCES OF HYBRIDISATION

The outcome of interspecific hybridisation may range from the formation of a single unfit/sterile individual, which is evolutionarily insignificant, to the production of hybrid species or the extinction of one or both parent species. Modelling of the potential outcomes of interspecific hybridisation has revealed that F1 fertility and habitat availability are key factors (McCarthy *et al.*, 1995). In a recent simulation, parent species (one common, one rare) were present in two habitats (one for each parent species) and F1 fertility, habitat selection and the gap between the parent species were varied (Buerkle *et al.*, 2003). In 84% of simulations a stable hybrid zone was produced; however if F1 fertility was high the rarer species tended to be replaced by hybrids, usually through adaptive trait introgression. Rarely, i.e. in 2.1% of simulations, hybrid speciation occurred, and was always accompanied by the extinction of the rare species (Buerkle *et al.*, 2003). In an earlier simulation, the likelihood of hybrid speciation was found to be higher due to the availability of a third, intermediate and unoccupied, habitat (Buerkle *et al.*, 2000).

1.4.1 *Extinction by means of hybridisation*

Introduced species may pose a risk to the native flora and fauna of a region if a level of cross-compatibility is present (Levin *et al.*, 1996; Rhymer and Simberloff, 1996; D. Wolf *et al.*, 2001), especially if interspecific hybrids show high fertility (Buerkle *et al.*, 2003).

The Catalina Island mahogany, *Cercocarpus traskiae*, is under threat of extinction through hybridisation with a much more common mahogany following considerable habitat degradation by goats and sheep. Of the 12 remaining adult trees in 1995, five were hybrids; however of the 70 or so seedlings in a herbivore-free area, most were pure *C. traskiae* indicating that the population may be recovering (Rieseberg and Gerber, 1995). *Spartina foliosa* is native to San Francisco bay and is under threat from introduced *S. alterniflora*, not because the former is rare, but

because the latter produces over 20-fold more pollen and is more vigorous than the native species (Anttila *et al.*, 1998). Interspecific *S. foliosa* x *S. alterniflora* hybrids have also been shown to be fertile and vigorous and it is predicted that the native species will be extinct within 3-20 generations (D. Wolf *et al.*, 2001).

1.4.2 *Maintenance of hybrid zones*

Hybrid zones have been described as "natural laboratories for evolution" (Barton and Hewitt, 1989) as they can allow us to examine the genetic differences responsible for speciation, levels of gene flow, and the spread of adaptations. Such zones may have existed in the wild for many hundreds or thousands of generations, in comparison to experimental populations, which have seldom existed for more than a few generations (Rieseberg *et al.*, 1999b). The dynamics and evolutionary fate of natural hybrid zones depend primarily upon gene flow and fitness of hybrid progeny (Harrison, 1993; Arnold, 1997). If hybrids form easily and show high fitness, then the hybrid zone may well be extensive, perhaps resulting in the fusion of the two parental taxa. However, strong reproductive barriers and/or reduced hybrid fitness can lead to very narrow zones with relatively few hybrid individuals present (Barton and Hewitt, 1985).

Essentially, models used to differentiate the factors affecting the genetic structure of hybrid zones can be classified into two groups. The first of these, the 'Tension Zone' or 'Dynamic Equilibrium' model (Barton and Hewitt, 1985), assumes that morphological clines along a transect that passes through a hybrid zone represent a balance between dispersal and endogenous selection against hybrid genotypes. Hybrid inviability or sterility arises if the parental genomes are incompatible or if recombination between the parental genomes disrupts coadapted gene complexes (Barton and Hewitt, 1985; Arnold and Hodges, 1995), reducing the diffusion of genes between the parental taxa. Clines develop wherever parental taxa come into contact regardless of environmental conditions (Barton and Hewitt, 1989). The second model, termed the Environment-Dependent Selection model, assumes that environmental gradients, or ecotones, produce the observed clines in morphology and allele frequencies. The hybrid zone in this instance is considered to represent equilibrium between gene flow and differential selection along the cline (Endler, 1973). Diffusion of genes is prevented by environmental selection rather that by intrinsic

incompatibilities of parental alleles as is assumed in the Tension Zone model (Barton and Hewitt, 1985; 1989).

Three versions of the environment-dependent model have been proposed. In the first, the 'Bounded Hybrid Superiority' model (Moore, 1977), hybrid fitness is taken to be greater than parental fitness in intermediate habitats, but lower in parental habitats. The second type considers 'mosaics' of habitats within the environment and not a simple gradation from one habitat into another (Harrison, 1986; Rand and Harrison, 1989). The most recent version, termed the 'Hybrid Novelty' model by Arnold (1997), assumes that hybridisation will give rise to a subset of hybrid genotypes with fitness greater than the parents in certain environments (including the parental habitat).

1.4.3 *Introgression*

Introgression, the permanent incorporation of interspecific DNA in an organism through hybridisation (Anderson and Hubricht, 1938; Anderson, 1949), has major adaptive potential and is a widespread phenomenon (Abbott, 1992; Arnold, 1992; Rieseberg and Wendel, 1993; Dowling and Secor, 1997). Hybrid formation tends to be rare and hence when hybrid individuals arise they will be in the numerical minority. If they reproduce by outcrossing they will tend to cross with individuals of the more common parent species and not other hybrid individuals, leading to the production of backcross progeny. Also, if the hybrid arises in the same population as one parent, natural selection will favour progeny of the hybrid which are more similar to the parent species (i.e. backcrosses) as they will be more suited to the existing habitat conditions (Grant, 1981). This appears to be the case in natural *Iris* hybrid zones in which F1 hybrid formation is rare (Arnold *et al.*, 1990b, 1991); however, once F1s have formed they act as a bridge to later hybrid generations (Hodges *et al.*, 1996). An important result of repeated backcrossing of the hybrid lineage is the likelihood of introgression.

Alleles likely to introgress are those which confer an advantage to the host organism, plus alleles at closely linked loci. Thus, an analysis of genetic markers in hybrid populations can reveal the chromosomal locations of introgressed loci (Rieseberg *et al.*, 2000).

A number of studies have reported that populations or subspecies which are introgressive in origin have an altered ecological range relative to the non-introgressed populations. However, direct evidence that introgression has provided the genetic substrate for ecological diversification is lacking. The Texan subspecies of wild sunflower, *Helianthus annuus* ssp. *texanus*, is of introgressive origin, containing chloroplast DNA (cpDNA) and ribosomal DNA (rDNA) of *H. debilis* ssp. *cucumerifolius* and it is thought that introgression has allowed the spread of this subspecies further south than the non-introgressed *H. annuus* (Rieseberg *et al.*, 1990). Introgression has also been suggested to have altered the environmental tolerance of introgressed populations of the orchid *Cypripedium pubescens* (Klier *et al.*, 1991) and British *Rhododendron ponticum* (Milne and Abbott, 2000).

1.4.4 *Speciation by hybridisation*

For natural hybridisation to create a successful hybrid derivative it is necessary for a fraction of recombinant genotypes to be able to outperform the parental species in at least some habitats (Arnold, 1997) and for there to be a degree of genetic isolation between the hybrid and its parent species (Charlesworth, 1995). Theoretical studies have shown that the presence of strong natural selection for those hybrids with high fertility and viability increases the chance of the hybrid lineage becoming established (Templeton, 1981; McCarthy *et al.*, 1995).

The formation of species by hybridisation can occur with or without a change in chromosome number (Table 1.1). Speciation together with an increase in chromosome number means instant reproductive isolation from the parents, whereas if no increase in chromosome number occurs then the neospecies must become isolated in some other way. Below, the two pathways to speciation by hybridisation - polyploid and homoploid - are discussed.

1.4.4.1 *Polyploid speciation*

Polyploid species are characterised by an increased chromosome number relative to the parent species. This increase may arise within a species (autopolyploid speciation) or via a cross between species (allopolyploid speciation). It is thought that most

polyploids arise via a triploid intermediate genotype ('triploid bridge') and not as a result of somatic chromosome doubling (Ramsey and Schemske, 1998). Importantly, a neospecies arising via polyploidy will attain immediate reproductive isolation from the parent species, hence preventing genetic swamping by matings with the parental taxa.

Polyploidy has played an important role in the evolution of plants, with an estimated 70% of angiosperms (Whitham *et al.*, 1991; Masterson, 1994) and 95% of pteridophytes (Grant, 1981) having experienced at least one round of polyploidy. Although polyploidy is most widely apparent in the plant kingdom, it is not solely restricted to plants - it has been suggested that two major polyploidisation events have occurred for all vertebrates (Spring, 1997), and an entire genome duplication for yeast (Wolfe and Schields, 1997; Kellis *et al.*, 2004).

Early opinion, that polyploidy was an evolutionary 'dead-end' (e.g. Wagner, 1983), was based on predictions that (1) a polyploid species will originate only once and hence exhibit less genetic variation than the diploid progenitor species, and that (2) the polyploid is likely to arise in a population of diploids and consequential inter-cytotype matings will result in sterile triploid offspring. However, it has been proven using molecular markers that the majority of polyploid species have originated independently several times (e.g. Roose and Gottlieb, 1976; Soltis and Soltis, 1989; Ashton and Abbott, 1992a; Cook *et al.*, 1998; Doyle *et al.*, 1999; Segraves *et al.*, 1999; reviewed in Soltis and Soltis, 1993, 1999). As a consequence, the polyploid species is likely to harbour genetic variation which is necessary for their establishment. Additionally, triploids are not uniformly sterile and produce haploid, diploid and triploid gametes at low frequency facilitating the propagation of the polyploid lineage (Ramsey and Schemske, 1998).

The success of polyploids is thought to be in part due to increased heterozygosity (Roose and Gottlieb, 1976; Levin, 1983; Soltis and Soltis, 1993, 1995). Allopolyploids are fixed heterozygotes that show disomic inheritance (Roose and Gottlieb, 1976), whereas autopolyploids are likely to form multivalents at meiosis and exhibit polysomic inheritance (Moody *et al.*, 1993). Another consequence of polyploidy appears to be greater ecological tolerances relative to the diploid progenitors, potentially contributing to the expansion of the polyploid into a new niche (Levin, 1983; Thompson and Lumaret, 1992; Soltis and Soltis, 1995). In his

review, Levin (1983) reports examples of polyploid species with increased tolerance to pathogens, nutrient stress and drought. Further to alteration in ecological tolerances, polyploid species may exhibit changes in the interaction with herbivores and pollinator species, leading to potential diversification of the plant and/or insect (e.g. Segraves and Thompson, 1999; Nuismer and Thompson, 2001).

Following allopolyploid formation, genomic changes may take place to stabilise the hybrid genome (Rieseberg, 2001). These genomic changes are often considerable, involving translocations and recombination between chromosomes inherited from different parents, and may occur to such an extent that the genome behaves like a diploid at meiosis. Maize and wheat have been shown to be ancient polyploids, despite their genomes being extensively 'diploidised' (Chen *et al.*, 1997; Gaut and Doebley, 1997).

Chromosomal repatterning tends to occur within a few generations of polyploid formation, e.g. in Brassica (Song *et al.*, 1995) and cotton (Adams *et al.*, 2003). Data from the formation of artificial allopolyploid wheat indicates that major chromosomal changes (i.e. the loss of DNA sequences and altered methylation patterns) occur in first generation hybrid and early generation allopolyploid genomes (Galili and Feldman, 1984; Ozkan *et al.*, 2001; Shaked *et al.*, 2001).

In addition to the above large-scale rearrangements and elimination of genome-specific sequences, changes can also occur to individual loci (reviewed in Wendel, 2000). In a neopolyploid each locus is present in duplicate, these are not homologous in an allopolyploid due to inheritance from two parents and so are termed 'homoeologues'. Over time the fate of these homoeologues may vary. In some instances one copy may become silenced, develop a new function, or develop a new pattern of expression (Wendel, 2000; Adams *et al.*, 2003; Osborn *et al.*, 2003). Such changes in gene expression in a polyploid genome may not result from changes in DNA sequence of the loci involved, but instead result from epigenetic changes, such as DNA methylation (Wolffe and Matzke, 1999; Comai, 2000). Resynthesis of the allotetraploid *Arabidopsis suecica* caused silencing of a number of loci, which could be reversed by treating the plant with a demelthylating agent (Madlung *et al.*, 2002).

1.4.4.2 *Homoploid hybrid speciation*

Homoploid (diploid) hybrid species arise as a result of interspecific hybridisation without an increase in chromosome number. Whereas an increase in chromosome number affords immediate reproductive isolation from the parent species (i.e. polyploid speciation), the homoploid neospecies has to be reproductively isolated in other ways. The simplest mechanism by which a homoploid hybrid lineage can become established is through spatial separation, whereby geographical isolation from the parent species prevents the hybrid from being swamped by gene flow from the parents (e.g. Sang *et al.*, 1995; Abbott *et al.*, 2000). However, hybrid individuals are most likely to arise within, or at close proximity to, one of the parent species. In this situation, the hybrid will often be swamped by gene flow, and for the hybrid(s) to persist, some level of reproductive isolation from its parents is necessary (Charlesworth, 1995). Theoretical models suggest that the most critical factors involved in hybrid speciation are strong natural selection for the most fertile/viable progeny (McCarthy *et al.*, 1995), rapid chromosomal evolution (Rieseberg *et al.*, 1995c) and/or an environment available in which the hybrid individuals are well-suited (Templeton, 1981; Arnold, 1997). Without niche divergence of the neospecies, the likelihood of the homoploid hybrid lineage becoming established is very low (Buerkle *et al.*, 2000). This is supported by empirical data in which most hybrid derivatives exhibit some form of ecological divergence from their parental species (Abbott, 1992; Arnold, 1997; Rieseberg, 1997).

Chromosomal rearrangements have been predicted to play a role in the reproductive isolation of homoploid hybrid species (Grant, 1981). This is based on observations of hybrids which exhibit karyotypic differences when compared to the parent species (e.g. Randolph, 1966; Gallez and Gottlieb, 1982; Rieseberg *et al.*, 1995c). In an early study, Grant (1966a) demonstrated that selection for fertility over nine hybrid generations in an interspecific cross resulted in the complete restoration of fertility, vigour and regular chromosome pairing, despite the F1 generation showing almost complete sterility and irregular chromosome pairing. Accompanied by this was the attainment of a high degree of reproductive isolation between one of the hybrid lines and the parent species due to chromosomal recombination. This was not due to

selection for reproductive isolation but rather due to selection for increased fertility of the hybrid (Grant, 1966b).

More recently, a series of investigations into the homoploid hybrid origin of the sunflower, *Helianthus anomalus*, have provided an insight into the process of diploid hybrid speciation (Rieseberg *et al.*, 1995c, 1996b; Rieseberg, 2000). Genetic mapping of *H. anomalus* and the parent species revealed that extensive recombination had taken place in the hybrid species although this was unlikely to be due to selection for reproductive isolation from the parent species (Rieseberg *et al.*, 1995c). Rieseberg *et al.* (1996b) produced three hybrid lineages following a cross between the putative parent species and comparative genetic mapping of these lineages indicated that they had converged in gene order. Moreover, this was statistically concordant with the gene order of the ancient hybrid species, indicating that selection, rather than chance, had governed the genetic composition of the hybrid species (Rieseberg *et al.*, 1996b). Analysis of the three artificial hybrid lines showed that pollen fertility in each of three hybrid lineages had increased from only ~5% in the F1 hybrid to ~90% over five generations and was attributed to genomic reorganisation (Rieseberg *et al.*, 1996b). Subsequent crossing experiments between the three hybrid lineages, the parent species and the ancient hybrid revealed that the artificial lineages showed a high degree of reproductive isolation from the parents and were partly fertile with each other and the ancient hybrid lineage (Rieseberg, 2000).

The question arises as to how such hybrid species achieve ecological isolation from their parents in order to avoid the minority type disadvantage immediately following their origin (Abbott, 2003). Transgressive segregation may provide the key.

Transgressive segregation, the presence in hybrid offspring of traits which are extreme to the trait values of both parents, appears to be common. A recent review found that of 171 studies reporting phenotypic variation in segregating hybrid populations, 155 reported at least one transgressive trait and that 44% of traits were transgressive (Rieseberg *et al.*, 1999a). Explanations of how these characters become transgressive include an increased mutation rate, complementary gene action, epistasis, fixation of recessive alleles from the parents, reduced developmental stability and overdominance (Rieseberg and Carney, 1998). The complementary action of genes from the parent species, however, is the most widely accepted

explanation for transgressive phenotypes (e.g. de Vincente and Tanksley, 1993; Schwarzbach *et al.*, 2001; Lexer *et al.*, 2003b).

In *Helianthus*, three hybrid sunflower species are derived from the same parent species (Rieseberg, 1991). These three sunflowers are morphologically and ecologically diverged from the parents (Heiser, 1947; Rieseberg, 1991), and a number of morphological and physiological traits associated with their ecological tolerances are transgressive (Rosenthal *et al.*, 2002; Table 1.2). The majority of the transgressive traits exhibited in the three diploid hybrid species could be recreated by artificial hybridisation of the parent species (Rieseberg *et al.*, 2003).

Evidence concerning the adaptive potential of transgressive traits is, however, limited to *H. paradoxus*, the hybrid sunflower found in saline marshes. Initial greenhouse-based growth experiments demonstrated that the *H. paradoxus* was over five times more fit, in terms of biomass, survivorship and root growth than either parent under high salt conditions (Welch and Rieseberg, 2002). Following this, individuals of the two parent species, the ancient hybrid species *H. paradoxus* and a large backcross population, were transplanted into a typical *H. paradoxus* habitat (i.e. brackish marsh) and revealed that leaf succulence and mineral ion uptake were under strong directional selection (Lexer *et al.*, 2003a). Additionally, for five of the seven elemental uptake traits, antagonistic QTL were detected (Lexer *et al.*, 2003a, b), a prerequisite for the development of a transgressive phenotype.

These studies show that ecological divergence, the most likely means by which a homoploid hybrid species will become established (Buerkle *et al.*, 2000), can be achieved by the sorting of antagonistic genetic factors present in the parent species.

Table 1.2 – Transgressive traits that may be associated with ecological divergence of three ancient hybrid sunflower species.

Species	*H. anomalus*	*H. deserticola*	*H. paradoxus*
Typical Habitat	Sand dunes	Desert	Brackish marshes
Transgressive traits (relative to parental species)	• Increased leaf succulence • Larger achenes • Greater air-leaf temperature difference	• Early bud/flowering • Decreased leaf size	• Increased leaf succulence • Increased Na concentration in leaves
References	Schwarzbach *et al.*, 2001; Rosenthal *et al.*, 2002	Rosenthal *et al.*, 2002	Rosenthal *et al.*, 2002; Welch and Rieseberg, 2002

1.5 HYBRIDISATION IN *SENECIO*

Senecio sect. *Senecio* offers a group of plant species in which to investigate the causes and consequences of interspecific hybridisation. The taxonomy of the diploid members is complex, with species limits often being poorly defined. This is in part due to high levels of phenotypic plasticity (Alexander, 1979) and the close relationship between most of the diploids due to a recent radiation in the group (Comes and Abbott, 2001). Additionally, species interfertility is high between some members, and recent hybridisation and interspecific gene flow have been documented (Abbott, 1992; Abbott *et al.*, 1995; Comes and Abbott, 1999; James, 1999). Hybridisation has given rise to the homoploid hybrid species, *S. squalidus* (James, 1999; Abbott *et al.*, 2000, 2002), a number of allopolyploid species (Ashton and Abbott, 1992a; Lowe and Abbott, 1996, 2000, 2003; Coleman *et al.*, 2003) and an introgressant form of *S. vulgaris* (Abbott *et al.*, 1992; Lowe and Abbott, 2000). The relationships between known hybrid taxa and their parent species are shown in Fig 1.1, and further details regarding some of the species are given below.

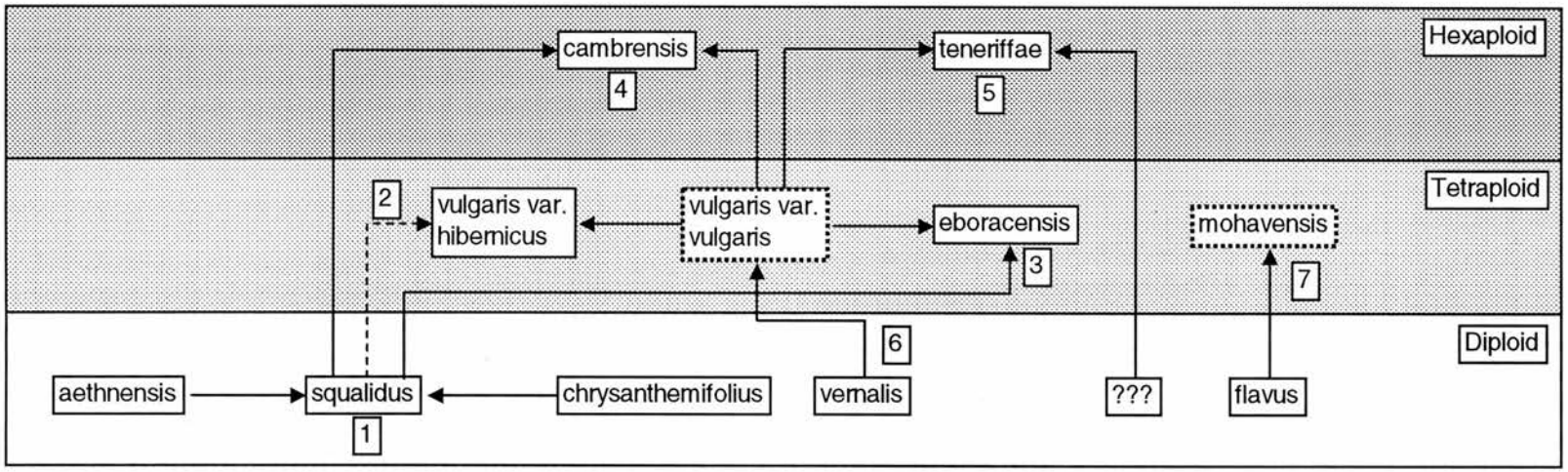


Figure 1.1 - Relationships between known hybrid *Senecio* taxa and their parental species. Solid lines between taxa indicate parentage, dashed lines indicate introgression. Taxa surrounded by dashed lines may be auto- or allopolyploid. ??? denotes unknown parent. Numbers refer to references: (1) Abbott *et al*., 2000, 2002; (2) Abbott *et al.*, 1992; (3) Irwin and Abbott, 1992; Lowe and Abbott, 2000, 2003; (4) Ashton and Abbott, 1992a; Harris and Ingram, 1992a; (5) Lowe and Abbott, 1996; (6) Comes *et al.*, 1997; Harris and Ingram, 1992b; (7) Coleman *et al.*, 2003; Comes and Abbott, 2001.

1.5.1 *Senecio squalidus*

Diploid *S. squalidus* L. (2*n*=20) is thought to have been introduced to southern England from Sicily in the early 18th century (Harris, 2002). After ~90 years, the species was noted growing outside the gardens in which it was originally planted, and has subsequently spread throughout much of the UK aided by its ability to colonise disturbed habitat, including railways and motorway verges (Abbott, 1992; Abbott *et al.*, 2000 and references therein).

Crisp (1972) postulated that British *S. squalidus* was atypical of other populations of *S. squalidus* from throughout Europe and believed that there was no evidence that material from elsewhere in Europe contributed to the British populations. Instead, Crisp (1972) proposed that British *S. squalidus* was a hybrid species derived from a zone of hybridisation between *S. chrysanthemifolius* Poiret and *S. aethnensis* DC. on Mt. Etna, Sicily, and showed that some hybrid individuals collected from Sicily were morphologically very similar to British *S. squalidus*. Further morphometric and allozyme analysis has confirmed the similarity of *S. squalidus* to natural *S. chrysanthemifolius* x *S. aethnensis* hybrids (James, 1999; Abbott *et al.*, 2000). *S. chrysanthemifolius* and *S. aethnensis* are fixed for alternative alleles at the *Acp-2* locus, and British *S. squalidus* is polymorphic for these alleles indicating that *S. squalidus* is not simply a derivative of one of the Etnan species (Abbott *et al.*, 2000). Further examination of *S. squalidus* with PCR-generated molecular markers which are specific to *S. chrysanthemifolius* or *S. aethnensis* has revealed that *S. squalidus* exhibits an additive profile for these markers (James, 1999).

Walters (1963) highlighted the morphological similarity between British *S. squalidus* and central and southern European *S. rupestris* Waldst. and Kit., leading Alexander (1979), in a revision of the section, to classify *S. rupestris* as a variety of *S. squalidus*. However, analysis of allozymes suggests that *S. rupestris* is a distinct species and was not involved in the origin of British *S. squalidus* (Abbott *et al.*, 2002).

1.5.2 *Senecio vulgaris* var. *hibernicus* and *S. eboracensis*

Tetraploid *Senecio vulgaris* L. var. *vulgaris* is widespread throughout Europe and represents the type species of sect. *Senecio* (Alexander, 1979). *S. vulgaris* var.

vulgaris is non-radiate and is thought to be derived from radiate *S. vulgaris* ssp. *denticulatus* (O. F. Muell.) P. D. Sell (Comes *et al.*, 1997). *Senecio vulgaris* var. *hibernicus* Syme and *S. eboracensis* Abbott & Lowe, in contrast, are both radiate and thought to be derived following hybridisation between *S. vulgaris* var. *vulgaris* and introduced (radiate) *S. squalidus*. Hybridisation between *S. vulgaris* and *S. squalidus* occurs very rarely in the wild and results in the production of the sterile triploid *S.* x *baxteri* (2*n*=30) (Marshall and Abbott, 1980). However, backcrosses to *S. vulgaris* occasionally produce near-tetraploid offspring (Ingram *et al.*, 1980).

A parallel spread of *S. squalidus* and var. *hibernicus* throughout the UK has been proposed as evidence for the origin of var. *hibernicus* by hybridisation (Crisp, 1972). Introgression has also been shown to have occurred from *S. squalidus* into *S. vulgaris* var. *hibernicus* by the presence of an *Aat-3c* allele in var. *hibernicus* that is present in *S. squalidus* and absent from populations of *S. vulgaris* that are monomorphic for the non-radiate capitulum (Abbott *et al.*, 1992). Controlled crosses between *S. vulgaris* var. *vulgaris* and *S. squalidus* can also result in backcrosses which are tetraploid and resemble var. *hibernicus* in morphology (Ingram *et al.*, 1980; Lowe and Abbott, 2000). Only a very small amount of genetic material is thought to have been transferred from *S. squalidus* into var. *hibernicus* via introgression. Of nine *S. squalidus*-specific PCR-generated molecular markers, none were found to be present in var. *hibernicus* (Abbott *et al.*, 2003).

Senecio eboracensis is a recently originated hybrid derivative that is restricted to York in the North of England. It is morphologically intermediate to the two parent species, *S. squalidus* and *S. vulgaris*, although exhibits some novel characters (Irwin and Abbott, 1992; Lowe and Abbott, 2003). Lowe and Abbott (2000) demonstrated that plants resembling *S. eboracensis* could be generated following crosses between *S. vulgaris* and *S. squalidus*. Analysis of molecular markers diagnostic of *S. vulgaris* and *S. squalidus* revealed that *S. eboracensis* possesses nine of 12 *S. vulgaris* markers and six of nine *S. squalidus* markers, indicating that a much greater proportion of the *S. squalidus* genome is present in *S. eboracensis* than in *S. vulgaris* var. *hibernicus* (Abbott *et al.*, 2003).

1.5.3 *Senecio cambrensis* and *S. teneriffae*

A third hybrid derivative of *S. vulgaris* and *S. squalidus* is the allohexaploid (2*n*=60) *S. cambrensis* Rosser (Rosser, 1955). This species is believed to have originated after chromosome doubling of the sterile triploid *S.* x *baxteri*, and synthetic hexaploids have been produced by treating triploid hybrids with colchicine (Weir and Ingram, 1980). Surveys of allozyme variation in material from Wales and Edinburgh showed that the species originated independently at these two locations and that there has been possibly more than one origin within Wales (Ashton and Abbott, 1992a). In a survey of cpDNA variation, Harris and Ingram (1992a) detected a 330bp insertion in the cp genome of Welsh *S. cambrensis* that was absent from plants in the Edinburgh population, lending further support to the hypothesis of independent origins. This cpDNA insertion is present in *S. vulgaris* var. *hibernicus* at low frequency (Lowe and Abbott, 1996).

The Canary Islands' endemic hexaploid *Senecio teneriffae* Schultz Bip. is morphologically similar to, and interfertile with, *S. cambrensis* (Lowe and Abbott, 1996). An analysis of cpDNA revealed that the 330bp insertion present in Welsh *S. cambrensis* is also present in *S. teneriffae* (Lowe and Abbott, 1996). The close relationship between these species has been attributed to similar origins. Whereas *S. cambrensis* has been proven to be the allohexaploid of *S. vulgaris* and *S. squalidus*, it is thought that *S. teneriffae* is possibly the allohexaploid of *S. vulgaris* and another diploid species related to *S. squalidus*, possibly *S. glaucus* (Lowe and Abbott, 1996).

1.5.4 *Senecio mohavensis*

The tetraploid (2n=40) *S. mohavensis* A. Gray is represented as two disjunct subspecies displaying an Old World - New World disjunction (Coleman *et al.*, 2001). The type subspecies is found in Arizona and California, whereas ssp. *breviflorus* (Kadereit) M. Coleman is found from Egypt, through the Near East and Arabia to the Thar Desert in Pakistan. Morphologically, *S. mohavensis* is very similar to diploid *S. flavus* (Dechne.) Sch. Bip. except for slight differences in leaf morphology and the presence of radiate capitula in *S. mohavensis* which are absent from *S. flavus*. Both of these taxa are morphologically distinct from other Mediterranean members of sect.

Senecio and it was thought that ssp. *breviflorus* was a subspecies of *S. flavus* until it was confirmed that ssp. *breviflorus* was tetraploid (Coleman *et al.*, 2001).

Morphologically, *S. mohavensis* appears to be the autotetraploid of *S. flavus* and this is supported by a study of random amplified polymorphic DNA (RAPDs) which revealed that *S. mohavensis* and *S. flavus* are closely related and distinct from other diploid *Senecio* species (Comes and Abbott, 2001). Under this scenario, the presence of radiate capitula in *S. mohavensis* is likely to be due to introgression.

However, evidence for an autotetraploid origin of *S. mohavensis* from *S. flavus* is not supported by the analysis of cpDNA and nuclear internal transcribed spacer (ITS) rDNA sequence variation. Liston and Kadereit (1995) found that the cpDNA haplotypes present in *S. flavus* and *S. mohavensis* are very different, with *S. mohavensis* having a haplotype much more similar to diploid *S. squalidus* which had been used as the outgroup in the analysis. Further study by Comes and Abbott (2001) identified the *S. mohavensis* cpDNA haplotype as being identical to one found in the widespread *S. glaucus*. An analysis of the sequence of the rDNA ITS region also revealed that *S. flavus* was distinct from other members of sect. *Senecio* and that the sequence obtained for *S. mohavensis* was most closely related to the ITS sequence of *S. glaucus* (Comes and Abbott, 2001; Coleman *et al.*, 2003).

Hence two possibilities for the origin of *S. mohavensis* exist (Comes and Abbott, 2001). The first, an autotetraploid origin, is supported by the morphological similarity to *S. flavus* and the results of a RAPD survey. This would have been accompanied by introgression of genetic material from a species closely related to *S. glaucus* to account for the cpDNA and ITS results. The second potential route of origin of *S. mohavensis* is via an allotetraploid event involving *S. flavus* and a species related to *S. glaucus*. Subsequent evolution towards a *S. flavus*-like morphology may then have occurred.

1.5.5 *Outline of thesis chapters*

This thesis contains four studies concerning hybridisation in sect. *Senecio*, each of which is outlined here:

Chapter 3

The first study reports an investigation into the strength of conspecific pollen advantage (CPA) exhibited between two interfertile species, *S. aethnensis* and *S. chrysanthemifolius*, which form a series of hybrid swarms on Mt. Etna, Sicily. CPA is caused by conspecific pollen preferentially fertilising the ovules of a species at the expense of heterospecific pollen. To investigate this, greenhouse-grown individuals of both species were pollinated with a pollen replacement series (i.e. 100:0, 75:25, 50:50, 25:75, 0:100 conspecific:heterospecific) and the resultant progeny genotyped to determine if they were products of conspecific or heterospecific fertilisation. In addition, the proportion of seed set and proportion of seed germinated were recorded to determine at which point the conspecific pollen might exhibit an advantage (i.e. on the stigmatic surface or in the style, at fertilisation of the ovules or due to hybrid seed abortion).

Chapter 4

A second study analysed morphology and random nuclear genetic marker and cpDNA RFLP variation to define the origin and taxonomic status of two populations of *Senecio* of uncertain taxonomic status that occur on sand dunes in the south of Sicily. Morphologically the populations are intermediate between certain species of sect. *Senecio*, which might suggest a hybrid ancestry of the populations. Alternatively, these populations could be an isolated island form of a more widespread taxon, with morphological divergence brought about by natural selection and/or genetic drift. Previous studies have indicated that the Sicilian material is either a derivative of *S. glaucus* x *S. gallicus*, or an isolated form of *S. glaucus* or *S. leucanthemifolius*.

Chapter 5

The aim of the third study was to determine the origin and parentage of the tetraploid species *S. mohavensis*. Previously, the mode of origin (auto- or allopolyploidy) has

not been confirmed, due to conflicting results based on morphology, RAPD, cpDNA and ITS DNA sequence variation. Hence, a survey was conducted of DNA sequence variation for two single-copy nuclear genes in a range of Mediterranean *Senecio* species that included *S. mohavensis*. This variation was then subjected to phylogenetic analysis.

Chapter 6

The final study conducted was similar to the investigation carried out in chapter 5. The same two nuclear loci were sequenced, but this time to determine the parentage of the hexaploid *S. teneriffae*. This species is morphologically very similar to the British hexaploid *S. cambrensis*, and isozyme evidence suggests that tetraploid *S. vulgaris* was a parent of both hexaploids. The diploid parent of *S. cambrensis* is known to be *S. squalidus*, whereas differences in allozymes between *S. teneriffae* and *S. cambrensis* imply that *S. squalidus* was not a parent of *S. teneriffae*. This investigation therefore was carried out to clarify the parentage of both species. In addition, the origin of *S. vulgaris* var. *hibernicus* is investigated. This taxon is thought to contain introgressed DNA from *S. squalidus*, including the gene responsible for formation of ray florets (the *RAY* gene). One of the genes sequenced in this study is tightly linked to *RAY* and hence close similarity between *S. squalidus* and *S. vulgaris* var. *hibernicus* at this gene would provide additional evidence for the introgressive origin of var. *hibernicus*.

CHAPTER 2 – MATERIALS AND METHODS

2.1 PLANT GROWTH CONDITIONS

To aid seed germination, seeds were placed on damp filter paper in a Petri dish for one week at 4°C. After this period, dishes were transferred to the University of St Andrews greenhouse for germination with day/night temperatures of 20°C/12°C and a 16 h daylength supplemented by 400W metal halide lamps. Following germination, seedlings were transferred singly to 9cm diameter pots containing 3:1 Levington's M2 compost:gravel and raised in the greenhouse in a randomised block design.

2.2 MORPHOMETRIC ANALYSIS

Plants were measured for 13 characters on the day of apical anthesis. Seven characters were descriptors of the apical inflorescence and six were descriptors of the midleaf, as follows:

1. Inflorescence length (mm) – Length of the apical stem node measured from the branch point to the stigma of the apical capitulum
2. Capitulum length (mm) – Measured from the point at which the peduncle widens to become the capitulum and the stigma of the apical capitulum
3. Capitulum Base Width (mm) - Diameter of the base of the apical capitulum
4. Mean number of calyculus bracts – Total number of bracts attached to the apical capitulum
5. Number of ray florets – Counted on the apical capitulum
6. Mean ray floret length (mm) – Mean floret length of five ray florets
7. Mean ray floret width (mm) – Mean floret width of five ray florets
8. Midleaf length (mm) – Measured from the point at which the leaf attaches to the stem to the apex
9. Midleaf area (mm^2)
10. Midleaf perimeter (mm)
11. Standardised midleaf area – Square root of midleaf area divided by the midleaf length

12. Midleaf dissection – Midleaf perimeter divided by square root of midleaf area
13. Standardised leaf perimeter - Midleaf perimeter divided by midleaf length

2.3 STARCH GEL ELECTROPHORESIS AND STAINING OF ISOZYMES

Starch gels were prepared by heating 33 g starch (Sigma or StarchArt) in 300 ml Lithium borate gel buffer (appendix Ia) until dissolved. Air bubbles were removed using a vacuum pump and the gel was poured into a mould and left to set.

A small amount of young leaf tissue (approx. 0.5 x 0.5 cm) was ground in one drop of extraction buffer (appendix Ia) in a well of a microtitre plate using a glass rod. The extract was absorbed onto a small (3 x 5 mm) filter paper wick and placed into a well of the gel. Marker dye was added to the end wells. The gel was then covered in Clingfilm and placed in a gel tank containing electrode buffer (appendix Ia). Sponges were used to provide a contact between the buffer and the gel. The apparatus was placed in a fridge and a current of 200 V applied. The gel was run for approximately three hours until the marker dye had moved ~7 cm. After this time the gel was removed from the tank and cut 1 cm below the origin and 1 cm above the solvent front (where the marker dye had travelled to). The gel was then placed on a glass plate and, using fishing wire and plastic spaces, sliced into three. The top slice was removed and the next slice was used for the staining for acid phosphatase (Acp).

Gel slices were placed in 0.4 M acetate buffer (pH 5.0) for 30 minutes at 4°C. They were then transferred to 50ml Acp stain buffer (appendix Ia) for ~30 min (Fig. 2.1).

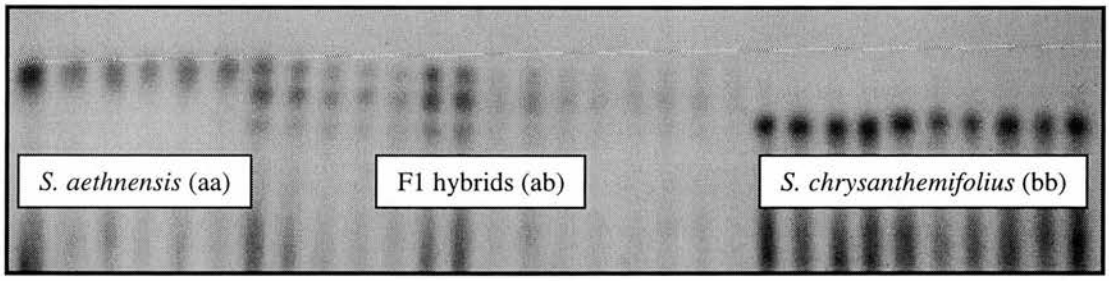


Figure 2.1 – Isozyme gel stained for acid phosphatase. Lanes 1-6 *S. aethnensis*; 7-21 *S. aethnensis* x *S. chrysanthemifolius* F1 hybrids; 22-31 *S. chrysanthemifolius*.

2.4 DNA EXTRACTION

DNA was extracted by one of two methods. Large-scale extractions were carried out on ~2g leaf tissue to obtain large amounts of DNA (e.g. for cpDNA RFLP). Further smaller extractions were carried out on ~200mg leaf tissue giving suitable amounts of DNA for PCR-based analysis. These protocols are adapted from Doyle and Doyle (1990).

2.4.1 *Large-scale DNA extraction*

Two grams of leaf tissue were removed from the plant and left overnight at 4°C, as this was shown to reduce the amount of starch and other impurities in the final extract. The leaf tissue was ground to a fine powder using liquid nitrogen and then added to 20 ml prewarmed (60°C) 2 × CTAB (appendix Ib) + 200µl 2-β-mercaptoethanol, and extracted for 30 min at 60°C. Following cooling for 10 min, 10 ml chloroform: isoamyl alcohol (24:1) was added and mixed thoroughly. The samples were centrifuged at 10,000 × g for 10 min at 20°C. The supernatant was removed and placed in a 50 ml tube, to which 30 ml 96% ethanol (ice-cold) was added. The DNA precipitated at the interface between the solutions and was removed using a glass hook and placed in 5 ml 76% ethanol, 0.2M sodium acetate for 1 hour. The hooks were then dipped in 76% ethanol, 10mM ammonium acetate and placed in 1ml TE (appendix Ib) to resuspend the DNA. To each sample 30 µl RNase A (10 mg/ml) was added and incubated at 37°C for 1 hr. Following incubation 50 µl 3 M sodium acetate was added and the DNA precipitated with 96% ethanol. Samples were centrifuged (10 min at 13,000 rpm), the supernatant poured off and the tubes left to air dry for 30 min.

2.4.2 *Small-scale DNA extraction*

Two hundred mg of fresh leaf material were placed in a 2 ml conical microtube, frozen using liquid nitrogen and ground to a fine powder. To each sample was then added 1 ml preheated (60°C) 2 × CTAB and 20 µl 2-β-mercaptoethanol. Samples were then incubated at 60°C for 30 min. After cooling at room temperature for 10 min, samples were centrifuged for 10 min at 13,000 rpm and the supernatant removed

and placed in a clean 2 ml microtube. To each sample was added 600 µl dichloromethane and mixed by gentle inversion for 20 min. Samples were then centrifuged (10 min at 13,000 rpm), the supernatant was placed in a clean 2 ml microtube, and the dichloromethane step followed by centrifugation was repeated. The resulting supernatant was placed in a clean 1.5 ml microtube, 600 µl freezer-cold 100% isopropanol added to precipitate the DNA, and centrifuged (10 min at 13,000 rpm). The supernatant was poured off and the tube inverted and left to air dry for 30 min. The DNA pellet was then dissolved in 0.5 ml TE and treated with 3 µl (10 mg/ml) RNase at 37°C for 1 hour. Following incubation, to each sample was added 50 µl 3 M sodium acetate and 600 µl freezer-cold 96% ethanol. Samples were centrifuged (10 min at 13,000 rpm), the supernatant was poured off and the tubes were left to air dry for 30 min.

DNA from both extraction protocols was dissolved in 200 µl TE and the concentration determined by running a 2 or 5 µl aliquot on a 1% agarose gel against 50 and 100 ng standards.

Agarose gels were prepared by heating 50 or 300 ml of 0.5 x TBE (appendix Ib) containing 1 - 1.5% agarose, until boiling. The mix was then left to cool for 10-30 min on a rotary shaker, at which point 25 or 100 µl ethidium bromide (10mg/ml) was added before pouring the gel into a mould using combs to make wells.

2.5 POLYMERASE CHAIN REACTION (PCR)

The polymerase chain reaction (PCR) is a powerful technique enabling amplification of sections of DNA using primers complementary to the immediate flanking regions. Random primers, e.g. RAPD and ISSR primers, allow genetic analysis without prior DNA sequence knowledge, whereas the amplification of a gene or intron requires primers to be designed based on DNA sequences in the taxa of interest.

2.5.1 *Random Amplified Polymorphic DNA (RAPD) PCR*

RAPD primers are random decamers used to amplify sections of the genome (Williams *et al.*, 1990). The DNA sequence of the PCR products are not important, however it is presumed that two bands migrating on a gel the same distance are

homologous. This has been quantified by Rieseberg (1996) who found that 9.1% of comigrating bands were non-homologous based on the production of different restriction fragment bands after digestion with three endonucleases. The principle behind RAPD analysis is that closely related individuals will exhibit more similar banding patterns than distantly related individuals.

Each RAPD PCR (25 µl total volume) contained 2.5 µl 10 × reaction buffer (supplied with enzyme; 160 mM $(NH_4)_2SO_4$; 670 mM Tris-Hcl (pH 8.8); 0.1% Tween-20), 2.5 mM $MgCl_2$, 0.2 mM each dNTP, 0.2 µM primer, one unit of BioTaq™ DNA polymerase (BioLine, UK) and 5 ng genomic DNA. The thermal cycling profile began with a denaturation at 94°C for 3 min, followed by 44 cycles of 30 s at 94°C, 45 s at 35°C and 90 s at 72°C, and a final elongation step of 4 min at 72°C. Amplification products were resolved in 1.2% agarose gels for ~3 hours at 100 V and visualised using UV transillumination.

To test for homology between PCR products generated by the same primer in different individuals, a restriction enzyme-based test was carried out following Rieseberg (1996). PCR products were excised from the gel for one or two individuals per taxon and purified using a gel extraction kit following the manufacturer's protocol (Qiagen). PCR amplification was carried out on ~5 ng gel-extracted DNA using the same primer and conditions as above. These products were restricted using three 4 bp-cutter restriction enzymes, *Alu*I, *Hpa*II and *Taq*I, following the manufacturer's protocol (Promega), separated on 8% polyacrylamide gels and stained with ethidium bromide. Products that showed identical restriction patterns for at least two of the enzymes were considered homologous between taxa (Rieseberg, 1996; Fig. 2.2). Non-homologous bands were not scored. In some instances, the product of interest was of a similar size to another produced by the same primer and could not be excised cleanly from the gel, hence could not be tested for homology using this protocol.

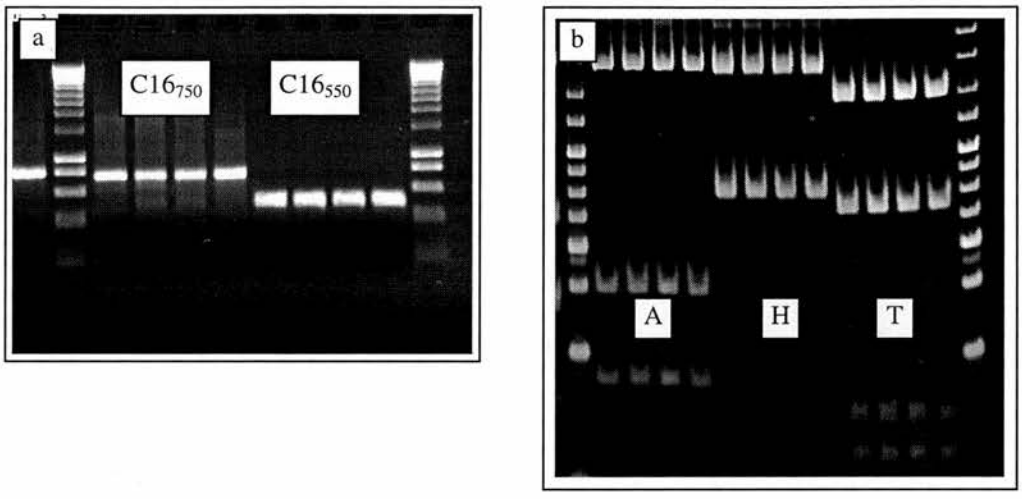


Figure 2.2 – (a) Gel-extracted and re-amplified RAPD bands generated with primer C16. (b) RAPD band $C16_{550}$ digested with restriction enzymes *Alu*I (A), *Hpa*II (H) and *Taq*I (T). In this example all bands were shown to be homologous.

2.5.2 *Single-copy gene amplification*

Two single-copy gene sequences were amplified using PCR. These were, (1) part of the gene encoding cytosolic phosphoglucose isomerase (*PgiC*) and (2) part of a *cycloidea* homologue isolated from *Senecio vulgaris* (*Scyc2*).

The first sequence examined (*PgiC*) is a portion of the nuclear gene encoding the cytosolic form of phosphoglucose isomerase (PGIC; EC 5.3.1.9). The locus was fully characterised in *Clarkia lewisii* by Thomas *et al.* (1992), and comprises 22 introns and 23 exons and encodes a protein of 569 amino acids (Fig. 2.3). The portion of the gene amplified contains exons 11 to 16 (Fig. 2.3). The second sequence studied here (*Scyc2*) is a portion of a gene that was isolated from *Senecio vulgaris*, and is homologous to the *cycloidea* gene in *Antirrhinum* that controls floral zygomorphy (Luo *et al.*, 1996). *Scyc2* encodes a protein of 303 amino acids in *S. vulgaris*, and contains the highly conserved TCP- and R-domains (see Cubas *et al.*, 1999a; Cubas, 2002). The portion amplified in this study is ~718 bp long, beginning at the start codon and containing both highly conserved domains (Fig. 2.3).

The polymerase chain reactions (PCR) conditions for amplifying the two sequences were identical: 94°C for 4 min, 35 cycles of 94°C for 30 s, 55°C for 1 min and 72°C for 2 min; followed by a final 10 min elongation at 72°C. PCR was carried

out in 20 µl reaction volumes containing 2 µl 10 x buffer (Qiagen; containing Tris-Cl, KCl, $(NH_4)_2SO_4$), 2 mM $MgCl_2$, 0.1 mM dNTPs, 0.5 mM each primer (Table 2.1), 0.5 Units *Taq* DNA polymerase (Qiagen), 4 µl Q solution (Qiagen) and ~5 ng genomic DNA. Reactions were carried out in a Perkin Elmer GeneAmp 9700 thermal cycler and resolved on 1.2% agarose gels stained with ethidium bromide.

Table 2.1 – Primers used to amplify *PgiC* and *Scyc2*. Positions of primers for *Scyc2* are based on the sequence for *S. vulgaris* (relative to the start codon). In most instances S2-START and S2-08 were used for amplification of *Scyc2*, however for the tetraploid *S. mohavensis*, primer S2-5' was used in place of S2-START due to a mutation in the priming site. Primer sequences for *PgiC* were supplied by L.D. Gottlieb, University of California, Davis.

Primer	Locus	Sequence (5'-3')	Position
S2-START	*Scyc2*	CTA ATG TTT TCC A	-3 to +10
S2-08	*Scyc2*	TAA ACA CTC TTT GAC TCG AT	+691 to +713
S2-5'	*Scyc2*	GAA CAT TAA TTA ATG TCT AT	-33 to -16
AA11F	*PgiC*	TTY GCN TTY TGG GAY TGG GT	Exon 11. 37-56
AA16R	*PgiC*	CCY TTN CCR TTR CTY TCC AT	Exon 16. 7-26

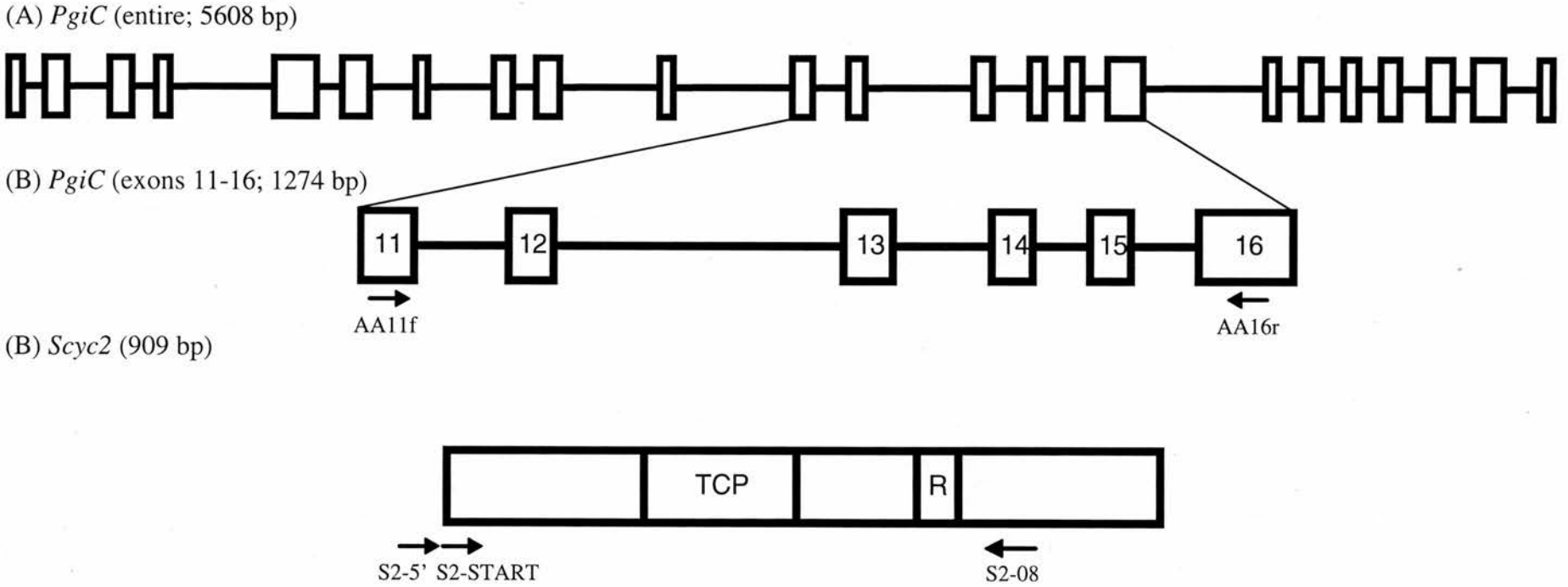


Figure 2.3 – Schematic diagrams of the *PgiC* and *Scyc2* genes. Primer annealing positions are denoted by arrows. (A) The entire length of *PgiC* (drawn to scale of *Clarkia lewisii*) is made up of 23 exons and 22 introns (Thomas *et al.*, 1992), (B) the portion amplified in this study (drawn to scale of *C. lewisii*). Primer AA11f anneals to positions 37-56 of exon 11 and AA16r to positions 7-26 of exon 16. (C) The structure of *Scyc2* with the highly conserved TCP- and R-domains marked. Primer S2-START overlaps the start codon (positions -3 to +10), primer S2-5' anneals to positions -33 to -16 and S2-08 anneals to positions 691-713. *Scyc2* primer positions are based on the sequence obtained for *S. vulgaris*.

2.6 CLONING AND SEQUENCING OF PCR PRODUCTS

2.6.1. *Ligation of PCR products into vector*

PCR products were ligated into pGEM-TEasy vectors (Promega) using the protocol supplied with the kit. Each ligation reaction contained 5 µl 2× rapid ligation buffer, 1 µl T4 DNA ligase, 1 µl vector (50 ng) and 3 µl PCR product. Ligation reactions were mixed gently, but thoroughly and left overnight at 4°C.

2.6.2. *Transformation of competent cells*

JM109 competent cells (Promega) were placed on ice for 5 min until just thawed. Fifty µl were gently pipetted into 1.5 ml microtubes, one per ligation reaction. To the side of the tube was added 7.5 µl of the ligation mix (from 2.6.1) and gently tapped down into the cells. Tubes were left on ice for 20 min, heat shocked at 42°C for 90 seconds then placed back on ice for 2 min. To each tube was then added 500 µl LB growth medium (appendix Ic) and placed at 37°C for one hour with shaking. Three agar plates (appendix Ic) per reaction were then plated with 50 µl, 200 µl or the rest of the mix, left for 10min, then incubated upside down overnight at 37°C.

2.6.3 *Assessing clones for insert size*

A sterile toothpick was used to select an individual colony, plated out in duplicate on fresh agar plates and then the toothpick was dipped into PCR mix (see 2.5.2, but 10µl reaction volume). Twelve to 24 colonies were screened per ligation. PCR conditions follow those stated above. Colonies plated in duplicate were incubated overnight. PCR products were resolved on 1.5% agarose gels to ensure the insert size was the same as the initial PCR product.

2.6.4 *Plasmid extraction*

Colonies to be sequenced were cultured in 5 ml volumes overnight. A colony was removed from the agar plate using a sterile toothpick and placed in 5 ml LB

[containing 2.5 µl ampicillin (100 mg/ml v/v in dH_2O; filter-sterilised)] in a 30 ml universal glass tube. Cultures were grown up overnight at 37°C with rotation.

Plasmid DNA was extracted from 2 ml of the overnight culture using the Perfectprep mini kit (Eppendorf) and the manufacturer's protocol was followed:

One ml of the overnight culture was centrifuged for 20 s at 13,000 rpm in 1.5 ml microfuge tubes to pellet the cells. The liquid was removed and another 1 ml culture added, and the step repeated. Cell pellets were re-suspended in 100 µl solution 1 by vigorous vortexing. To this was added 100 µl of solution 2, mixed gently, then 100 µl solution 3 were added and mixed vigorously. Tubes were centrifuged at 13,000 rpm for 30 s and the supernatant placed in a spin column in a collection tube. To this was then added 450 µl DNA binging matrix, before mixing and centrifuging at 13,000 rpm for 30 s. The filtrate was decanted and the spin column placed back into the collection tube. To the column were then added 400 µl of diluted purification solution (1:1 purification solution: 96% ethanol) and shaken briefly. Columns were centrifuged twice at 13,000 rpm for 1 min, with the filtrate removed between spins. The column was then placed in a fresh collection tube and plasmid was eluted by adding 50 µl ddH_2O, before vortexing and centrifuging for 1 min. Plasmid concentration was then determined on 1.2% agarose gels using 25, 50 and 100 ng uncut λ DNA as standards.

2.6.5 *Cycle sequencing*

A plasmid preheat step was included according to the DTCS QuickStart kit (Beckman Coulter) protocol: Fifty ng plasmid DNA was placed in a 0.2 ml PCR tube and made up to 10 µl using ddH_2O. This was then heated to 96°C for 1 min and allowed to cool to room temperature.

Each cycle-sequencing reaction contained 10 µl DNA/dH_2O (above), 8 µl DTCS (Beckman Coulter) and 2 µl (F or R) universal pUC/M13 primer (2 pmol/µl). PCR conditions followed that of the kit protocol: 30 cycles of 96°C for 20s, 50°C for 20s and 60°C for 4 min, followed by holding at 4°C.

Following PCR, samples were ethanol precipitated as detailed in the DTCS kit (Beckman Coulter) protocol. All following procedures took place on ice. Into 0.5 ml microfuge tubes (one per sequencing reaction) were placed 5 µl stop solution [2 µl

3M sodium acetate (pH 5.4), 2 µl 100 mM Na_2-EDTA (pH 8.0), 1 µl glycogen (20 mg/ml; supplied with DTCS kit)]. Sequencing reactions were then added to the stop solution and mixed.

To each tube 60 µl 96% ethanol (from -20°C freezer) were added, mixed well, and spun at 13,000 rpm at 4°C for 15 min. The supernatant was removed carefully using a pipette. Pellets were then rinsed twice with 200 µl 70% ethanol (from -20°C freezer). Following each rinse, tubes were centrifuged for 10 min and the ethanol was removed. Pellets were vacuum-dried for ~10 min and finally re-suspended in 40 µl sample loading solution (supplied with kit). Samples were then loaded into reaction plates and overlaid with one drop of mineral oil. Sequencing was carried out in a Beckman Coulter CEQ using the CEQ8000 software and default parameters.

2.7 CHLOROPLAST DNA RESTRICTION FRAGMENT LENGTH POLYMORPHISM

This technique allows the detection of specific sequences of chloroplast (cp) DNA following restriction enzyme-digestion of whole genome DNA. DNA fragments are separated on agarose gels and transferred to a nylon membrane. This membrane is then probed using labelled cpDNA and differences in restriction patterns (i.e. RFLPs) are detected.

2.7.1 *Digest of genomic DNA and running of gels*

Restriction digests of 500 ng genomic DNA were performed according to the enzyme manufacturer's instructions (Promega). Digested DNA was run through agarose gels at 30 V in 0.5 x TBE buffer as follows. For *Pst*I and *EcoR*I digests, samples were loaded onto 1% agarose gels and were run for 16-20 h. For *Cla*I and *Cfo*I digests, samples were loaded onto 1.4% agarose gels and run for ~40 h to increase resolution of large and similar-sized fragments. Size standards (Hyperladder I, Bioline, UK) were run alongside the restricted DNA.

Gels were removed from the gel tanks and the surplus gel plus the lanes containing the size standards were cut off. The portion containing the size standard was soaked in a solution of 1 µg/µL ethidium bromide for 30 min, visualised with UV and photographed against a ruler so that band sizes on the blots could be determined

by measuring distance from origin and comparison with the photograph of the size standards.

2.7.2 *Southern blotting of DNA onto nylon membrane*

The main portion of the gel (containing the restricted DNA) was soaked in 200 ml denaturation solution (appendix Id) for 20 min then transferred to 200 ml neutralisation solution (appendix Id) for 1.5 h. DNA from the gel was then transferred to a nylon membrane using the Southern blotting technique (Southern, 1975). Gels were rinsed and placed upside down on 3-ply Whatman 3 mm chromatography paper. The chromatography paper acted as a wick by placing the ends in a tray of 20 x SSC (appendix Id). On top of the gel was placed a sheet of nylon membrane (Hybond NX, Amersham) the same size as the gel, then two pieces of chromatography paper the same size as the gel, followed by a stack (approx. 8 cm thick) of laboratory paper towels. On top of this was placed a sheet of Perspex (1 cm thick) and a weight. The apparatus was left overnight.

2.7.3 *Preparation of DIG-labelled probes*

CpDNA probes were prepared by labelling cloned fragments of the *Lactuca sativa* cp genome (Jansen and Palmer, 1987) with digoxygenin (DIG) using the DIG-High Prime non-radioactive DNA-labelling and detection kit II (Roche Molecular Biochemicals, Germany).

One µg of cloned cpDNA fragment was made up to 16 µl with ddH_2O, boiled for 10 min then quenched on ice. To this was added 4 µl of DIG-High Prime (containing random primers, Klenow polymerase, dNTPs, DIG-11-dUTP and reaction buffer in 50% glycerol; Roche). The mix was centrifuged briefly and incubated overnight at 37°C. Following incubation, 2 µl 0.2 M EDTA (pH 8.0) were added and heated to 65°C for 10 min to stop the reaction.

The concentration of the DIG-labelled probe was determined by preparing a dilution series of the labelled probe and the control DIG-labelled DNA (Roche) in DNA dilution buffer (50 µg/ml fish sperm DNA in 10 mM Tris-HCl, 1 mM EDTA; pH 8.0). One µl spots of each dilution were applied to a piece of nylon membrane

(Hybond NX, Amersham) and fixed by exposure to UV for 2 min. A scaled-down version of the protocol in section 2.7.4 was then used to determine the concentration of labelled probe.

2.7.4 *Detection of cpDNA haplotypes*

The following protocol is suitable for a nylon membrane of ~200 cm^2. For larger or smaller membranes the protocol can be scaled up or down, respectively.

The blotted nylon membranes from 2.7.2 were placed individually in hybridisation cylinders containing 20 ml DIG Easy Hyb (Roche) preheated to 40°C. Cylinders were placed in a hybridisation oven at 40°C and rotated gently for 30 min. During incubation, the probe (from 2.7.3) was denatured by boiling for 5 min before quenching on ice. Following incubation, DIG Easy Hyb was removed and ~500ng denatured probe in 7 ml fresh DIG Easy Hyb was added to the cylinder and left to hybridise overnight.

Following hybridisation, the DIG Easy Hyb containing the probe was removed and stored at -20°C for future use. Membranes were placed flat in glass trays and rinsed twice in 100 ml of 2 × SSC, 0.1% sodium dodecyl sulphate (SDS) at room temperature with agitation. Membranes were transferred to hybridisation cylinders containing 75 ml of 0.5 × SSC, 0.1% SDS preheated to 65°C, before placing in the hybridisation oven at 65°C for 15 min. This wash was repeated with fresh 0.5 × SSC, 0.1% SDS.

The following immunological detection was carried out in glass trays at 37°C with agitation, except where stated. After wash steps, the membranes were rinsed for 5 min in DIG wash buffer (appendix Id), then incubated sequentially in the following solutions:

200 ml Blocking solution [1:5 blocking solution (Roche): maleic acid buffer (appendix Id)] for 30 min;

50 ml Antibody solution (5 µl Anti-digoxygenin-AP (Roche) in 50ml diluted blocking solution) for 30 min;

100 ml DIG wash buffer for 2 x 15 min;

40 ml DIG detection buffer (appendix Id) for 5 min.

Membranes were then transferred, DNA side up, to hybridisation bags (Roche) and 2 ml CSPD (Roche) were spread over the surface. Immediately following this the bag was sealed. The bag and its contents were incubated for 5 min at room temperature, then 10 min at 37°C. Each membrane was placed in a developing folder and exposed to Kodak X-OMAT AR-5 X-ray film for 10-20 min. Films were developed, and the resultant RFLPs were inspected before assigning a haplotype to each individual.

CHAPTER 3 - POLLEN COMPETITION AMONG TWO SPECIES OF *SENECIO* (ASTERACEAE) THAT FORM A HYBRID ZONE ON MT. ETNA, SICILY

ABSTRACT

Hybridisation between interfertile, sympatric or parapatric, plant species can be reduced significantly by conspecific pollen advantage (CPA), whereby conspecific pollen exhibits an advantage over heterospecific pollen in terms of ovule fertilisation. An examination of CPA in two interfertile species of *Senecio*, *S. aethnensis* and *S. chrysanthemifolius* (Asteraceae), which form a hybrid zone on Mt. Etna, Sicily was carried out. Individuals of both species were pollinated with pollen mixtures containing 0, 25, 50, 75 or 100% heterospecific pollen and offspring were genotyped to determine if they were products of conspecific or heterospecific pollen fertilising the ovules. The mean proportion of hybrid offspring produced on *S. aethnensis* plants was not significantly different to that expected based on the proportion of heterospecific pollen applied to the flower head. However, *S. chrysanthemifolius* mother plants showed moderate CPA, with the proportion of hybrid offspring significantly less than expected. No reduction in seed set or seed germination was observed, hence the CPA found for *S. chrysanthemifolius* acts before ovule fertilisation. The consequences of asymmetry in CPA on the reproductive isolation of *S. aethnensis* are briefly discussed, along with other mechanisms that may play a role in the maintenance of the hybrid zone on Mt. Etna.

3.1 INTRODUCTION

Interfertile plant species that occur sympatrically or parapatrically in the wild may be reproductively isolated due to the action of one or more isolating mechanisms (Grant, 1992; Ramsey *et al.*, 2003). Pollination by another species may be prevented by species specific pollinators (e.g. Grant, 1994; P. Wolf *et al.*, 2001) or differences in peak-flowering periods (e.g. Gottlieb and Pilz, 1976; Cruzan and Arnold, 1994). However, for species that rely on generalist animal pollinators and/or wind pollination, the likelihood of receiving pollen loads comprising a mixture of conspecific (i.e. same species) and heterospecific (i.e. another species) pollen is expected to be high. In these species, the production of potentially unfit hybrid offspring can be greatly reduced if conspecific pollen fertilises more ovules than expected based on its proportion in the mixed pollen load. This has been termed conspecific pollen advantage (CPA; Alarcón and Campbell, 2000) and is believed to be a common isolating mechanism in the plant kingdom (Stace, 1989). CPA may result from heterospecific pollen exhibiting reduced germination on the stigma, retarded heterospecific pollen tube growth in the style and decreased fertilisation of ovules, when compared to conspecific pollen. Conspecific pollen advantage at any of these stages will result in less hybrid offspring produced than would be expected.

Conspecific pollen advantage has been shown to reduce significantly the frequency of hybrid seed formation following mixed pollinations in several interfertile species pairs examined (Arnold *et al.*, 1993; Carney *et al.*, 1994; Hauser *et al.*, 1997; Klips, 1999). Strong CPA will decrease the potential for introgression, and if the strength of CPA is asymmetric (as is often the case: Kiang and Hamrick, 1978; Rieseberg *et al.*, 1995a; Emms *et al.*, 1996; Carney and Arnold, 1997; Diaz and MacNair, 1999), introgression will occur more readily from one species into the other. Asymmetric introgression has been documented in some hybridising populations by the analysis of the distribution of species-specific molecular markers through the hybrid zone (Paige *et al.*, 1991; Hardig *et al.*, 2000). Asymmetric CPA could potentially have implications for the extinction of species which lack or exhibit weak CPA (Levin *et al.*, 1996; Buerkle *et al.*, 2003).

Even in cases of strong CPA, e.g. in *Helianthus* (Rieseberg *et al.*, 1995a) and *Iris* (Arnold *et al.*, 1993; Carney *et al.*, 1994; Carney and Arnold, 1997), some hybrid

progeny may be formed in areas of sympatry/parapatry. Although these may initially be of reduced fitness relative to parent species, their fitness may increase over generations such that in certain circumstances they may evolve into a stabilised introgressant (Rieseberg *et al.*, 1990, 1991b; Arnold *et al.*, 1991; Abbott, 1992) or a new hybrid species (Rieseberg, 1991; Abbott, 1992; Arnold, 1993; Rieseberg *et al.*, 2003).

In the study reported here it has been examined whether CPA occurs in two interfertile species of *Senecio*, *S. chrysanthemifolius* Poiret and *S. aethnensis* Jan. ex DC., that form a hybrid zone on Mt. Etna, Sicily. On Mt. Etna, *S. chrysanthemifolius* occurs at altitudes below 1000m, whereas *S. aethnensis* is endemic to high altitude sites (>1600m). At intermediate sites hybrid swarms are common between the two species (Crisp, 1972; James, 1999). The two species are morphologically distinguishable with *S. aethnensis* exhibiting relatively large flower heads (capitula) and entire, glaucous leaves, while *S. chrysanthemifolius* produces smaller capitula and highly dissected, non-glaucous leaves (Abbott *et al.*, 2000). Hybrid material is generally intermediate for these characters (Abbott *et al.*, 2000); although a wide range of hybrid and backcross phenotypes is evident in some hybrid swarms (Abbott, pers. obs.). Detailed surveys of allozyme, RAPD and cpDNA variation have detected a number of species-specific markers, and broad clines in marker frequency are exhibited between the two species indicating extensive hybridisation and gene flow across the hybrid zone (James, 1999). The distribution of these markers suggests that *S. aethnensis* is more introgressed than *S. chrysanthemifolius* over much of its distribution on Mt. Etna (James, 1999; Abbott *et al.*, 2000). The hybrid zone on Mt. Etna is believed to be the source of material that gave rise to the homoploid hybrid species, *Senecio squalidus* L., in the British Isles. Support for this hypothesis has come from a recent comparative morphometric and allozyme analysis of all three *Senecio* species and material from the hybrid zone (Abbott *et al.*, 2000, 2002).

3.2 MATERIALS AND METHODS

3.2.1 *Plant growth conditions and genotyping according to allozyme phenotype*

Seed (achenes) were collected from two populations of each of *S. aethnensis* and *S. chrysanthemifolius* occurring on Mt Etna, Sicily (Table 3.1). Approximately 25 seed from each of ten different mother plants per population were placed on damp filter paper in Petri dishes and placed in the dark at 4°C for one week. Dishes were then transferred to a growth cabinet with day/night temperatures of 20°C/12°C and a 16 h daylength supplied by 40 W fluorescent tubes. Two days after germination, seedlings were transferred to pots of 3:1 compost to gravel and grown to maturity in the same cabinet. After approximately three weeks, plants were genotyped at a locus (*Acp*-2) controlling acid phosphatase variation using starch gel-electrophoresis (see section 2.3). Previous work (James, 1999; Abbott *et al.*, 2000) has shown that *S. aethnensis* and *S. chrysanthemifolius* are fixed for different alleles at the *Acp*-2 locus. Parents of crosses were therefore selected such that *S. aethnensis* parents were homozygous for the *Acp*-2a allele, and *S. chrysanthemifolius* parents homozygous for the *Acp*-2b allele (Fig. 2.1). After genotyping, plants were grown to maturity in the same cabinet.

Table 3.1 - Locations on Mt Etna, Sicily, of populations of *Senecio aethnensis* and *S. chrysanthemifolius* examined (from James, 1999).

Population	Species	Location	Lat. 00°00' N	Long. 00°00'E	Altitude
C0	*S. chrysanthemifolius*	Northern Catania	37.32	15.05	<50 m
C1	*S. chrysanthemifolius*	Pedara	37.37	15.04	600 m
VBU	*S. aethnensis*	Near Cisternazza	37.44	15.01	2600 m
BB	*S. aethnensis*	Bocca Superiore	37.44	14.56	2525 m

3.2.2 *Crossing design*

From ten plants per population, five were chosen randomly as pollen recipients and donors, and five as pollen donors only. Prior to anthesis, flower heads (capitula) were covered with small bags made from lens tissue, preventing stray pollen arriving on the

stigmas. No attempt was made to remove all self-pollen from florets in a capitulum as both species are highly self-incompatible (pers. obs.). However, in most cases, the majority of self-pollen was removed from a capitulum prior to pollination for use in other crosses. Pollen used in pollinations was collected and pooled from each population. This was done by gently tapping fully-opened capitula from at least four plants over a piece of silver foil, to release pollen, which was then mixed with a wooden toothpick.

Five mother plants per population were subjected to the same pollination treatments. Five different ratios of conspecific to heterospecific pollen were applied to each mother plant using a small paintbrush. Each pollen ratio mixture was applied to one capitulum per plant and one capitulum per mother was left unpollinated (and bagged) to measure level of selfing. Pollen mixtures comprising the following pollen ratios by mass to the nearest 0.1 mg were produced using a Mettler-Toledo AB104-S balance: 100:0, 75:25, 50:50, 25:75, 0:100 conspecific:heterospecific pollen. The ratios containing 100:0 and 0:100 conspecific:heterospecific pollen are referred to hereafter as 'pure conspecific' and 'pure heterospecific' pollinations, respectively. The other three ratios are referred to as 'mixed' pollinations. Following pollination capitula were re-bagged.

To minimise the chance of donors and recipients sharing self-incompatibility (SI) alleles conspecific pollen was taken from the other population of the same species. In addition, pollen mixtures were prepared using pollen from at least four individuals per species.

3.2.3 *Offspring analysis*

Once seeds had matured, capitula were removed and the numbers of filled and unfilled seed were counted to calculate seed set (proportion of filled seed). All seed were allowed to germinate (as above) and the proportion that germinated was recorded. All seedlings resulting from mixed pollinations were genotyped electrophoretically at the *Acp*-2 locus to identify the relative proportions of offspring that resulted from conspecific and heterospecific fertilisation. In addition, offspring from the pure heterospecific pollinations were genotyped to determine if the presence of heterospecific pollen on stigmas had induced some selfing to occur, due to the

mentor effect (Richards, 1986). In total, 3755 progeny were genotyped. Hybrids were easily identified by their heterozygote banding pattern at the *Acp*-2 locus (Fig 2.1).

3.2.4 *Statistical analysis*

Data for proportion of seed set, proportion of seeds germinated and proportion of hybrids were transformed using the arcsine square-root of the ratio (y +3/8)/(N+3/4), where y was either the number of seed set, seed germinated or hybrid offspring recorded and N was either the potential seed set (i.e. total number of filled and unfilled achenes), number of seed set, or number of offspring genotyped, respectively. Transformed data were analysed using the GLM option of SAS 8.2 (SAS Institute Inc., Cary, NY, USA) with the main factors being pollen mixture, species, population (nested within species), and individual (nested within population). For proportion of hybrid offspring, the results of pure conspecific and pure heterospecific pollinations were omitted from analysis.

χ^2 tests were also conducted to determine if the proportion of hybrid progeny differed significantly from that expected based on the ratio of pollen mixture applied.

3.3 RESULTS

3.3.1 *Selfing*

No seeds were produced in capitula that were bagged and left to self, thus confirming that the two species are strongly self-incompatible. However, a few non-hybrid progeny (eight out of 895) were formed from pure heterospecific pollinations and were assumed to be products of selfing. Six of these were produced on *S. chrysanthemifolius* plants and two on *S. aethnensis* plants.

3.3.2 *Seed set*

Seed set (i.e. the proportion of filled achenes) was significantly lower when *S. aethnensis* acted as the maternal plant [$\overline{x} = 0.494 \pm 0.028$ (SE)] than when *S.*

chrysanthemifolius was the maternal parent ($\bar{x}$ = 0.768 ± 0.015; $P < 0.0001$; Fig. 3.1; Table 3.2). Seed set varied significantly between individuals within populations ($P < 0.0001$), but not between populations within species. Pollen mixture did not affect seed set. Seed set of *S. aethnensis* mother plants tended to decrease as the proportion of heterospecific pollen in the pollen mixture increased (Fig. 3.1), but the pollen x species interaction was not significant (Table 3.2).

3.3.3 *Germination*

The proportion of seed to germinate remained high and was not significantly different across pollen mixture treatments [$\bar{x}$ = 0.935 ± 0.005 (SE); range = 0.711-1.000; Fig. 3.2]. Germination percentage did not differ significantly between species, populations within species or individuals within populations (Table 3.2).

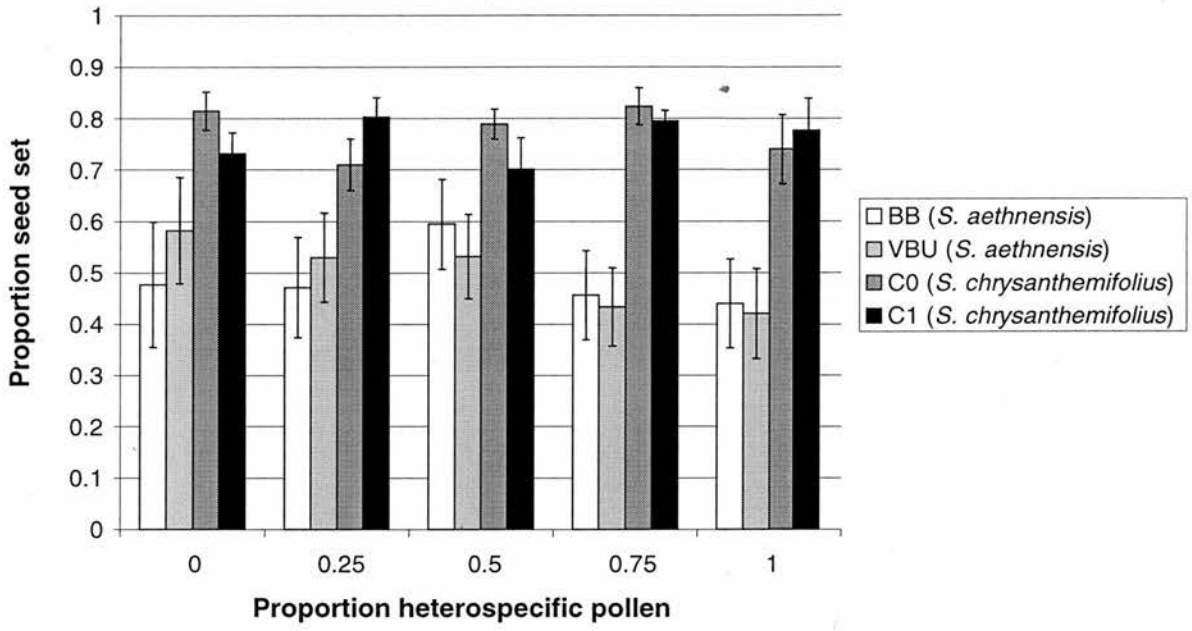


Figure 3.1 – Mean seed set (± SE) across pollen ratio treatments.

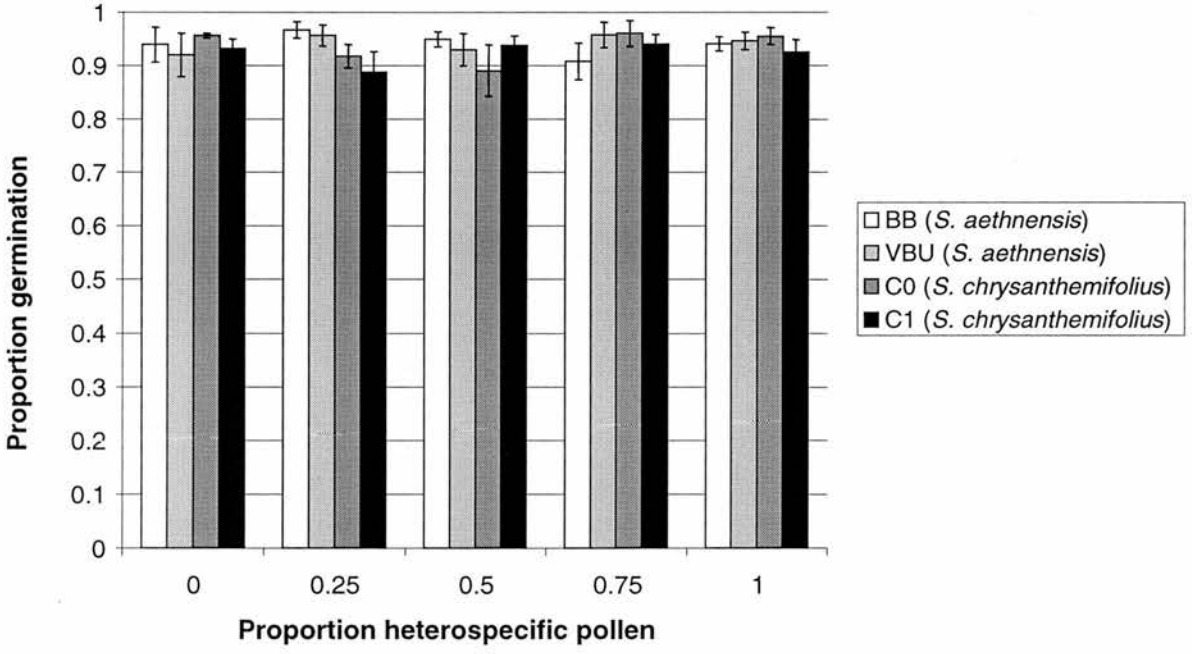


Figure 3.2 – Mean proportion of seeds germinated (± SE) across pollen ratio treatments.

Table 3.2 – Analysis of variance of (a) seed set, (b) seed germination and (c) proportion of hybrid progeny. Pollen mix refers to the ratio of conspecific:heterospecific pollen applied to the capitulum. For proportion of hybrid progeny, only the three pollen mixes containing both pollen types were analysed (i.e. 25:75, 50:50 and 75:25). Error mean squares for testing significant differences are indicated in the table as subscripts of the *F* ratio. Type III sums of squares were used throughout.

Source of Variation	(a) Seed set				(b) Seed germination				(c) Proportion of hybrid progeny			
	df	MS	*F* ratio	*P* value	df	MS	*F* ratio	*P* value	df	MS	*F* ratio	*P* value
(1) Pollen mix	4	0.0120	$0.64_{(1,6)}$	0.6381	4	0.0023	$0.20_{(1,6)}$	0.9353	2	1.0201	$163.78_{(1,6)}$	<0.0001
(2) Species	1	2.1809	$27.42_{(2,5)}$	<0.0001	1	0.0017	$0.14_{(2,6)}$	0.7050	1	0.2123	$34.08_{(2,6)}$	<0.0001
(3) Pollen mix × Species	4	0.0317	$1.69_{(3,6)}$	0.1624	4	0.0144	$1.25_{(3,6)}$	0.2962	2	0.0018	$0.29_{(3,6)}$	0.7468
(4) Population (species)	2	0.0023	$0.03_{(4,5)}$	0.9714	2	0.0063	$0.55_{(4,6)}$	0.5793	2	0.0126	$2.02_{(4,6)}$	0.1477
(5) Individual (population)	16	0.0795	$4.23_{(5,6)}$	<0.0001	16	0.0119	$1.04_{(5,6)}$	0.4249	16	0.0063	$1.02_{(5,6)}$	0.4608
(6) Error	72				72				36			

Abbreviations: df, degrees of freedom; MS, mean squares.

3.3.4 *Offspring genotype ratios*

Applying an increased proportion of heterospecific pollen to capitula of both species caused a significant increase in the proportion of hybrid offspring produced (Fig. 3.3; Table 3.2). *Senecio aethnensis* produced a higher proportion of hybrid progeny than *S. chrysanthemifolius* at all three pollen mix ratios. The species x pollen mixture interaction was not significant, indicating that the trend of an increase in proportion of hybrid offspring in each species was correlated. No significant difference was observed in the proportion of hybrid offspring formed between populations within species or between individuals within populations (Table 3.2).

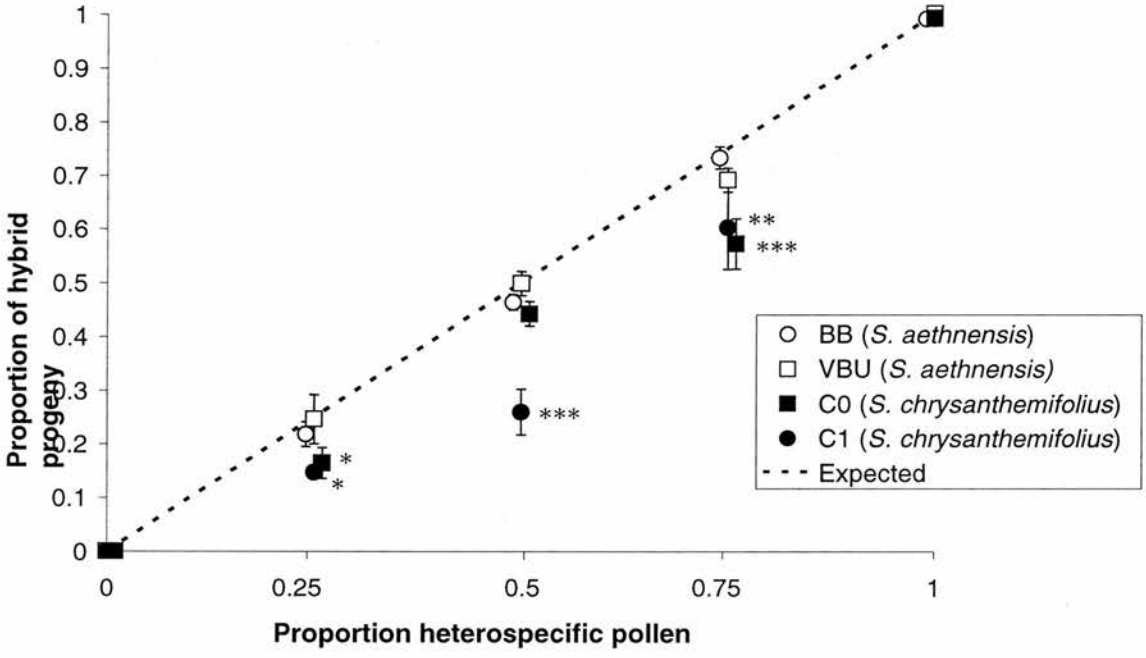


Figure 3.3 – Mean proportion of hybrid offspring (±SE) obtained from mixed pollinations. Data points are staggered horizontally for clarity. Those values that differ significantly (χ^2 analysis) from expected values based on pollen ratio applied (dashed line) are indicated * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$.

When data were pooled over the three pollen mixtures within species, a χ^2 test showed that *S. aethnensis* mother plants did not produce significantly fewer hybrid offspring than expected (Table 3.3). The same was found for each population of *S. aethnensis* (BB and VBU) analysed separately. Conversely, *S. chrysanthemifolius*

mothers produced significantly less hybrid offspring than expected when populations were analysed together or separately (Table 3.3).

Table 3.3 – χ^2 and *P* values for comparisons of observed and expected numbers of hybrid progeny obtained in the pollen mix treatments. Data for pure conspecific and pure heterospecific pollen loads were omitted from the analysis.

Population/Species	χ^2	d.f.	*P*
BB	3.944	14	0.996
VBU	8.375	14	0.869
S. aethnensis	12.318	28	0.997
C0	35.842	14	0.001
C1	53.804	14	<0.001
S. chrysanthemifolius	89.646	28	<0.001

χ^2 tests carried out separately on each of the three different pollen mixtures, revealed significant departures from expected proportions of hybrid offspring in five of the six *S. chrysanthemifolius* groups examined. The one exception was for population C0 treated with a 50:50 mixture of conspecific:heterospecific pollen (Fig. 3.3). The mean for this was 0.442 ± 0.023 (SE), whereas the equivalent treatment of C1 plants produced a mean proportion of hybrid offspring of 0.260 ± 0.043 (SE). Similar χ^2 tests of *S. aethnensis* data showed that both populations of *S. aethnensis* produced the expected proportions of hybrid offspring at each pollen treatment ($0.191 \leqslant P \leqslant 0.978$; Fig. 3.3).

The results indicate that CPA reduces hybrid formation when mixed pollen loads arrive on *S. chrysanthemifolius* plants, but the same is not true for *S. aethnensis*, which produces hybrid progeny directly proportional to the fraction of *S. chrysanthemifolius* pollen in the pollen mix.

3.4 DISCUSSION

For most plant species, several mechanisms act in concert to prevent or reduce hybridisation with other species (Grant, 1992; Ramsey *et al.*, 2003). One such mechanism is conspecific pollen advantage (CPA), which results from conspecific pollen having an advantage over heterospecific pollen in fertilisation and the production of offspring. In the present study of two interfertile species of *Senecio* that form a hybrid zone on Mt Etna, Sicily, it is shown that whereas CPA is not exhibited by the high altitude species, *S. aethnensis*, a moderate level of CPA is exhibited by the low altitude species, *S. chrysanthemifolius*. This is an example, therefore, of asymmetrical CPA between two species.

Several other studies have measured CPA between interfertile species using a similar approach to that employed in the present study. Often, the proportion of hybrid offspring produced from mixed pollinations is reported to be considerably less than the proportion of heterospecific pollen in the pollen mixture (Arnold *et al.*, 1993; Carney *et al.*, 1994; Rieseberg *et al.*, 1995a; Klips, 1999). Only in one study, involving two species of *Ipomopsis*, was CPA found to be absent (Alarcón and Campbell, 2000). However, Campbell *et al.* (2003) later demonstrated that conspecific pollen of the same two species had an advantage over F_1 pollen in terms of number of seeds sired. Thus, CPA can sometimes prevent the formation of backcross progeny without affecting the formation of F_1 hybrids.

Asymmetrical CPA has been demonstrated between several species (Kiang and Hamrick, 1978; Rieseberg *et al.*, 1995a; Emms *et al.*, 1996; Carney and Arnold, 1997; Diaz and MacNair, 1999), and could be an important cause of asymmetrical gene flow (Rieseberg *et al.*, 1991a; Dorado *et al.*, 1992). Absence of CPA in *S. aethnensis* suggests that introgression is more likely to occur from *S. chrysanthemifolius* into *S. aethnensis*, than in the opposite direction, and this is borne out by the results of an analysis of molecular marker variation across the hybrid zone on Mt Etna (James 1999). James (1999) showed that whereas a large part of the distribution of *S. chrysanthemifolius* is not introgressed with markers diagnostic of *S. aethnensis*, *S. aethnensis* is introgressed with *S. chrysanthemifolius* markers except at the highest altitudes of its distribution on Mt Etna.

The CPA exhibited by *S. chrysanthemifolius* is likely to result from reduced germination of *S. aethnensis* pollen relative to *S. chrysanthemifolius* pollen on *S. chrysanthemifolius* stigmas and/or retarded growth of *S. aethnensis* pollen tubes in the *S. chrysanthemifolius* style. Selective abortion of hybrid ovules in *S. chrysanthemifolius* is unlikely to be a cause of the CPA as there was no reduction in seed set as the ratio of heterospecific pollen was increased in a pollination treatment. Additionally, pollen mix had no effect on seed germination, so the possibility that hybrid seeds show reduced germination relative to seeds sired by conspecific pollen can be excluded as a cause of the CPA exhibited by *S. chrysanthemifolius*.

In several studies of CPA, the growth of conspecific and heterospecific pollen tubes has been compared in the styles of both parent species (e.g. Carney *et al.*, 1994; Rieseberg *et al.*, 1995a; Emms *et al.*, 1996; Carney and Arnold, 1997; Diaz and MacNair, 1999). However, pollen tube growth has been shown to be a poor predictor of CPA in *Helianthus* and *Iris*, in terms of the proportion of hybrid progeny produced (Rieseberg *et al.* 1995a; Emms *et al.*, 1996), possibly due to pollen tube growth being a dynamic process and varying in different parts of the style (Walsh and Charlesworth, 1992). Consequently, no attempt was made to relate CPA to pollen tube growth in the present study.

The finding that CPA is absent in *S. aethnensis* and of only moderate strength in *S. chrysanthemifolius* raises the question as to what are the major factors that maintain the taxonomic identity of these two species on Mt Etna. Both species have similar flower heads, although capitula of *S. aethnensis* are larger than those of *S. chrysanthemifolius* (Abbott *et al.*, 2000), so pollinators are likely to be generalists and not specific to a particular species (Proctor, 1978; Schmitt, 1980; Comes and Kadereit, 1990). *S. chrysanthemifolius* flowers early in the year and most plants begin to die-back by mid-June. In contrast, *S. aethnensis* flowers later with peak flowering period occurring towards the end of August (pers. obs.). However, a series of hybrid populations connect the two species and flower at intermediate times of the year, thus allowing gene flow to occur between the species across the hybrid zone. It is feasible that the hybrid zone on Mt Etna is a tension-zone (Barton and Hewitt, 1985, 1989) and that endogenous selection against hybrids limits gene flow between the two species. However, although hybrid fitness has not been quantified in the wild, it was found that the hybrid seed produced in this study germinated as well as seed of

intraspecific crosses. In addition, hybrids have been observed to be vigorous and fertile in the wild and also when cultivated in a greenhouse.

The most likely factor that maintains the taxonomic identity of *S. aethnensis* and *S. chrysanthemifolius* as distinct species on Mt Etna, despite their interfertility, is environmental-dependent selection. Growing at altitudes greater than 1600 m, *S. aethnensis* will be subject to a higher level of ultraviolet-B (UV-B) irradiation (Caldwell and Robberecht, 1980), and reduced partial pressure of CO_2 ($ppCO_2$) and mean temperature (Fitter and Hay, 1987), relative to *S. chrysanthemifolius*. It is likely, therefore, that the two species are adapted to environmental conditions at the extremes of the altitudinal gradient on Mt Etna, and it is this that maintains their identity as different species in the face of interspecific gene flow. This now needs to be tested along with the possibility that hybrids exhibit higher fitness than parent species at intermediate sites along the gradient.

CHAPTER 4 - TAXONOMIC STATUS AND ORIGIN OF A SICILIAN *SENECIO* (ASTERACEAE) BASED ON MORPHOLOGY, RAPD AND CHLOROPLAST DNA VARIATION.

ABSTRACT

The taxonomy of diploid Mediterranean *Senecio* sect. *Senecio* (Asteraceae) is complex, due to a recent species radiation, high morphological plasticity and occasional interspecific hybridization and gene flow. In this study, two populations of *Senecio* of uncertain taxonomic status from sand dunes in southern Sicily, Italy were analysed. Plants in these populations have been described previously as morphologically intermediate to *S. gallicus* and *S. glaucus* ssp. *coronopifolius*, indicating a possible hybrid origin, or as a variant of *S. leucanthemifolius*. A survey of morphological, random amplified polymorphic DNAs (RAPDs), and cpDNA RFLP variation was carried out to help resolve the taxonomic status and origin of these plants. Plants raised in a greenhouse were morphologically intermediate to *S. glaucus* ssp. *coronopifolius* and *S. leucanthemifolius*, but were also similar to some cultivated individuals of *S. gallicus*. No evidence for a hybrid origin was obtained from RAPD analysis, which demonstrated that Sicilian material is closely allied to Tunisian *S. glaucus* ssp. *coronopifolius*. Sicilian plants were also found to be polymorphic for the same set of cpDNA haplotypes present in Tunisian *S. glaucus* ssp. *coronopifolius*; some of these haplotypes were absent from material surveyed of *S. leucanthemifolius* and *S. gallicus.* In conclusion, the Sicilian *Senecio* of uncertain taxonomic status is a variant form of North African *S. glaucus* ssp. *coronopifolius,* which most probably dispersed to sand dunes in southern Sicily in the relatively recent past. The presence of several cpDNA haplotypes in this material indicates that there have been multiple introductions of the species to Sicily, but RAPD variation is low and populations remain restricted to sand dunes close to the ports of Licata and Pozallo.

4.1 INTRODUCTION

In rapidly evolving species groups, taxonomic boundaries can be blurred because there are few diagnostic characters to distinguish taxa, high levels of phenotypic plasticity make species identification difficult, and occasional interspecific hybridisation and gene flow may erode species differences in certain areas. Thus, within such complexes, some populations may be difficult to assign with certainty to particular taxa. Under these circumstances, taxonomic resolution might be aided by comparing material raised in a common environment and/or by examining material at the molecular level. This approach can also resolve the origins of taxonomically ambiguous material.

Mediterranean *Senecio* sect. *Senecio* (Asteraceae) are taxonomically complex and species limits are often poorly defined (Alexander, 1979; Abbott *et al.*, 1995; Coleman, 2003), especially among the diploids. This is partly due to most diploids being products of a recent radiation in the group (Comes and Abbott, 2001), and also to high levels of phenotypic plasticity and species interfertility (Alexander, 1979; Abbott *et al.*, 1995). An important consequence of species interfertility is that hybridization may occur at locations where species co-exist leading to interspecific gene flow that can confound molecular phylogenetic investigations (e.g. Soltis and Kuzoff, 1995; Ferguson and Jansen, 2002; Manen *et al.*, 2002; Semerikov *et al.*, 2003). Examples of recent hybridisation and gene flow have been documented in the section (Abbott, 1992; Abbott and Lowe, 1996; Comes and Abbott, 1999; James, 1999) and have resulted in the origin of the diploid hybrid species *S. squalidus* (Abbott *et al.*, 2000, 2002), several allopolyploid species (Ashton and Abbott, 1992a; Lowe and Abbott, 1996, 2000; Coleman *et al.*, 2003), and an introgressant form of *S. vulgaris* (Abbott *et al.*, 1992).

The purpose of this investigation was to determine the taxonomic status and derivation of a type of *Senecio* found on sand dunes in the region of Licata and Pozallo in southern Sicily, Italy. In the most recent revision of Mediterranean *Senecio* sect. *Senecio*, Alexander (1979) described the plant as intermediate in morphology to *S. gallicus* Vill. and *S. glaucus* L. ssp. *coronopifolius* (Maire) Alexander, indicating a possible hybrid origin. However, in a study of genetic variation in the widespread *S. glaucus*, Coleman (2003) showed that the plant is closely allied to Tunisian *S. glaucus* ssp. *coronopifolius* material, although no other species were included in his analysis

and, therefore, a hybrid origin cannot be ruled out. Finally, in a recent large-scale examination of molecular variation among Mediterranean *Senecio* sect. *Senecio*, Comes and Abbott (2001) referred to the material as *S. leucanthemifolius* Poiret, a species commonly found in Sicily.

All three species that may be involved in the ancestry of the Licata/Pozallo material, i.e. *S. gallicus*, *S. glaucus* ssp. *coronopifolius*, and *S. leucanthemifolius*, are diploid ($2n$=20) and widespread throughout the Mediterranean (Alexander, 1979; Fig. 4.1). In addition to describing Licata/Pozallo *Senecio* material as intermediate to *S. glaucus* and *S. gallicus*, Alexander (1979) also described material from elsewhere in the Mediterranean as intermediate in morphology to *S. glaucus* and *S. leucanthemifolius* (Libya, Fig. 4.1), and to *S. gallicus* and *S. leucanthemifolius* (Gibraltar, Fig. 4.1). In each case such material might be of hybrid origin or a variant form of one particular species. These plants occur at the edge of the geographical ranges of the species postulated to be parents and might result from divergence due to selection and/or drift (Slatkin, 1985). Recently, Coleman and Abbott (2003) examined the possibility that atypical material of *S. leucanthemifolius* var. *casablancae* from the southern part of its range in Morocco originated following hybridisation with the more southerly distributed *S. glaucus* ssp. *coronopifolius*. Although some evidence of past hybridization was provided by chloroplast DNA markers, there was no evidence that nuclear DNA introgression had occurred and consequently it was concluded that the atypical morphology of the southern material was most probably caused by selection and/or drift.

The aim of the present study was to determine if the *Senecio* material that occurs on sand dunes near Licata and Pozallo in southern Sicily is a hybrid derivative or a variant form of *S. gallicus, S. glaucus* or *S. leucanthemifolius*. To this end, a detailed morphometric analysis was conducted on cultivated material from Licata and Pozallo, and material of the three potential parent species. In addition, material was surveyed for random amplified polymorphic DNA (RAPD) and chloroplast (cp) DNA variation.

Although RAPDs are frequently used in studies of plant population genetic structure and hybridisation, there can be problems with interpretation due to uncertain reproducibility and homology of co-migrating fragments (Williams *et al.*, 1993; Rieseberg, 1996; Adams and Rieseberg, 1998; Harris, 1999). However, tests for reproducibility and homology can reduce these problems. Chloroplast DNA markers

are also often used in studies of hybridization due to the chloroplast genome being a large non-recombining unit and hence the effects of past reticulation events are potentially retained for long periods of evolutionary time (Doyle, 1992; Wendel and Doyle, 1998). Previous studies by Comes and Abbott (1998, 1999, 2001) have extensively analysed cpDNA variation within and among Mediterranean *Senecio* sect. *Senecio*, and provided a solid background for further comparative studies of cpDNA variation in the section.

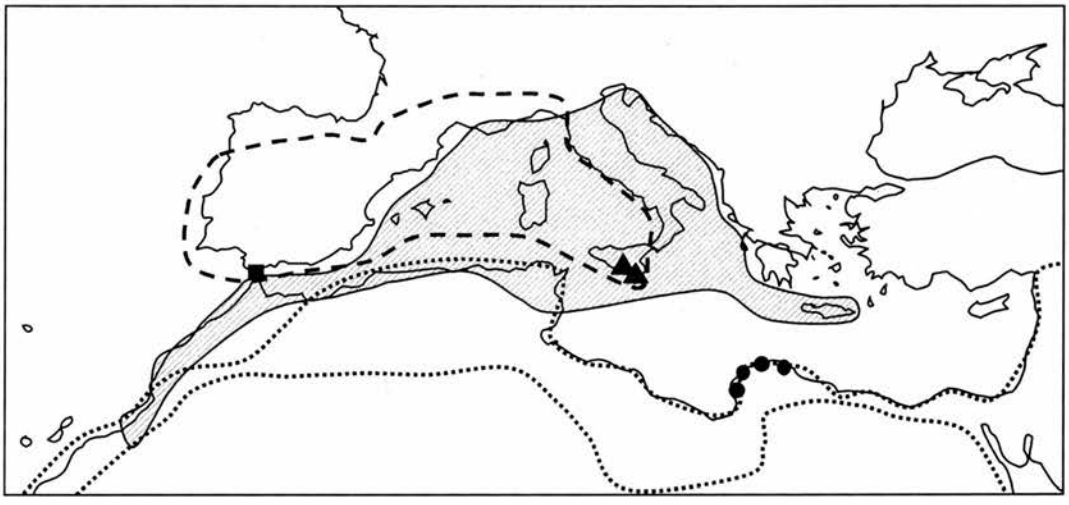


Figure 4.1 – Distributions of *Senecio gallicus* (dashed line), *S. glaucus* ssp. *coronopifolius* (dotted line) and *S. leucanthemifolius* var. *leucanthemifolius* (shaded region) in the Mediterranean. Three populations of intermediate morphology (as recognised by Alexander, 1979) are indicated: *S. gallicus* × *S. leucanthemifolius* (square), *S. leucanthemifolius* × *S. glaucus* (circles) and *S. gallicus* × *S. glaucus* ssp. *coronopifolius* (triangles; this study).

4.2 MATERIALS AND METHODS

4.2 1 *Plant material*

Seeds were collected from approximately 30 different plants from each of two populations of the *Senecio* of uncertain taxonomic status that occur on sand dunes near Licata and Pozallo in southern Sicily (Fig. 4.2; Table 4.1). In addition, seed was obtained from a similar number of plants from each of several populations of *S.*

gallicus, S. glaucus ssp. *coronopifolius* and *S. leucanthemifolius* var. *leucanthemifolius* that occur in the Mediterranean region (Fig. 4.2; Table 4.1). Seed samples were placed on damp filter paper in Petri dishes at 4°C in a refrigerator for one week, and then transferred to a greenhouse for germination with day/night temperatures of 20°C/12°C and a 16 h daylength supplemented by 400W metal halide lamps. Following germination, seedlings were transferred singly to 9cm diameter pots containing 3:1 Levington's M2 compost: gravel and raised in the greenhouse in a randomised block design.

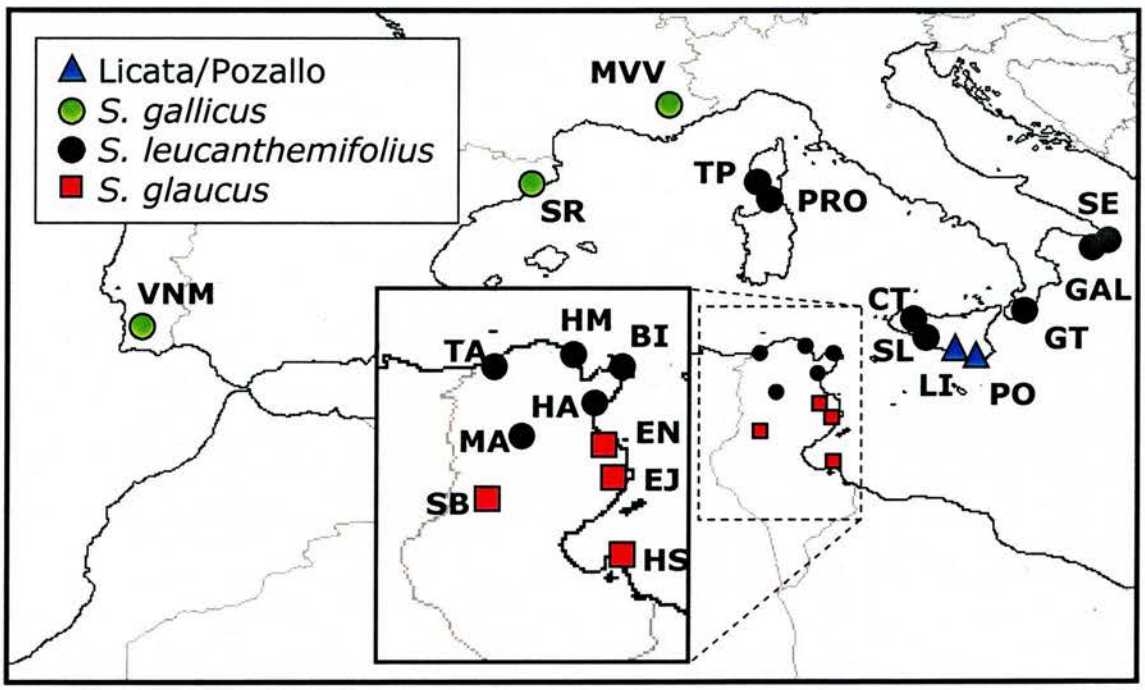


Figure 4.2 – Sampling locations for *Senecio* material analysed in this study. Population codes are given in table 4.1.

Table 4.1 – Populations analysed in this study. Number refers to number of individuals analysed in the morphometric (morph) and RAPD investigations. NA = not analysed.

Species	Code	Name	Lat 00°00'	Long 00°00'	Collector[1] and date	Number (morph)	Number (RAPD)
Unknown	LI	Licata, Sicily	37 07 N	13 57 E	HPC &RJA (1995)	6	10
Unknown	PO	Pozallo, Sicily	36 44 N	14 15 E	HPC &RJA (1995)	6	9
S. glaucus ssp. *coronopifolius*	EN	Enfida Plage, Tunisia	36 03 N	10 27 E	MC & RJA (1999)	4	6
S. glaucus ssp. *coronopifolius*	EJ	El Jem, Tunisia	35 15 N	10 15 E	MC & RJA (1999)	5	5
S. glaucus ssp. *coronopifolius*	SB	Sbeitla, Tunisia	35 13 N	09 30 E	MC & RJA (1999)	5	7
S. glaucus ssp. *coronopifolius*	HS	Houmt-Souk, Tunisia	33 55 N	10 52 E	MC & RJA (1999)	7	8
S. l. var. *leucanthemifolius*	MA	Makthar, Tunisia	35 50 N	09 12 E	MC & RJA (1999)	NA	5
S. l. var. *leucanthemifolius*	TA	Tabarka, Tunisia	36 55 N	08 45 E	MC & RJA (1999)	7	6
S. l. var. *leucanthemifolius*	BI	Bizerte, Tunisia	37 18 N	09 52 E	MC & RJA (1999)	5	8
S. l. var. *leucanthemifolius*	HA	El Haouaria, Tunisia	37 03 N	11 00 E	MC & RJA (1999)	5	8
S. l. var. *leucanthemifolius*	HM	Hammamet, Tunisia	36 25 N	10 40 E	MC & RJA (1999)	5	5
S. l. var. *leucanthemifolius*	SL	San Leone, Sicily	37 21 N	13 35 E	HPC & RJA (1995)	5	6
S. l. var. *leucanthemifolius*	CT	Castellammare, Sicily	38 02 N	12 52 E	HPC & RJA (1995)	NA	5
S. l. var. *leucanthemifolius*	GT	Gioia Tauro, Italy	38 26 N	15 55 E	HPC & RJA (1995)	4	6
S. l. var. *leucanthemifolius*	SE	Torre San Emiliano, Italy	40 06 N	18 20 E	HPC & RJA (1995)	NA	6
S. l. var. *leucanthemifolius*	GAL	Galipoli, Italy	40 03 N	17 59 E	HPC & RJA (1995)	3	6
S. l. var. *leucanthemifolius*	TP	Tour de la Parata, Corsica	41 54 N	08 36 E	CT (2001)	5	6
S. l. var. *leucanthemifolius*	PRO	Propriano, Corsica	41 40 N	08 54 E	CT (2001)	7	6
S. gallicus	MVV	Mont Ventoux, France	44 09 N	05 16 E	HPC & RJA (1995)	6	6
S. gallicus	VNM	Villa Nova, Portugal	37 43 N	08 47 W	HPC & RJA (1995)	6	3
S. gallicus	SR	Sa Riera, Spain	41 57 N	03 12 E	HPC & RJA (1995)	6	6

[1] Collections were made by R. J. Abbott (RJA), M. Coleman (MC

4.2.2 *Morphometric analysis*

In general, five to seven plants per population were measured for a range of morphological traits, except in populations EN, GT and GAL where only three or four plants were measured. On the day of full anthesis of the apical capitulum, plants were measured for 13 characters. Seven characters were descriptors of the apical inflorescence, i.e. inflorescence length, capitulum length and width, mean ray floret length and width, number of ray florets and number of calyculus bracts, and six were descriptors of the midleaf, i.e. length, area, perimeter, standardised midleaf area (square root of leaf area divided by the leaf length), midleaf dissection (leaf perimeter divided by square root of leaf area) and standardised leaf perimeter (leaf perimeter divided by leaf length). Further details of the characters measured are given in section 2.2.

4.2.3 *Genetic analysis*

DNA extraction

DNA extraction was conducted using the CTAB method of Doyle and Doyle (1990) with minor modifications (section 2.4). Large-scale extractions (~2g of leaf material) were conducted for the cpDNA RFLP analysis, while small-scale extractions (~200µg leaf material) were carried out to obtain DNA for the RAPD study.

RAPD Analysis

Initially seven DNA samples (two from each of *S. gallicus*, *S. glaucus* and *S. leucanthemifolius*, and one from a Licata plant) were amplified using 60 random decamer oligonucleotides (Operon Technologies sets A, B and C) following conditions stated in section 2.5.1.

From these 60 primers, 17 were chosen based on the production of clear banding patterns, and a repeat PCR was conducted using the same seven DNA samples plus a further six individuals of the parent species (two each from *S. gallicus*, *S. glaucus* and *S. leucanthemifolius*). Potential bands that did not amplify in both PCR reactions using the first seven DNAs were not scored. Those bands which amplified in more than one species were tested for homology using a restriction enzyme-based protocol similar to that used by Rieseberg (1996) following the procedure stated in

section 2.5.1. Products that showed identical restriction patterns for at least two of the enzymes were considered homologous between taxa (Rieseberg, 1996). A number of potential bands proved non-homologous and were consequently not scored in the final analysis. Eight primers were chosen for the full analysis of samples based on their ability to produce reproducible bands that were easy to score and were homologous between species. Of the 40 fragments examined in the total analysis, nine were found in only one or other of the three parent species and were assumed to be homologous between individuals within the particular species. Of the remaining 31 bands scored, 19 were successfully tested for homology. Ideally all bands shared between taxa should have been tested for homology; however, for the remaining 12 bands, the product of interest was of a similar size to another produced by the same primer and could not be excised cleanly from the gel.

A full analysis of RAPD variation was carried out on 133 DNA samples using eight primers (Table 4.2). Because all 133 DNA samples could not be amplified in one PCR plate, it was necessary to ensure that repeatable banding patterns were produced by different PCR amplifications with the same primer. This was tested by including four control DNAs and two controls that lacked DNA in each PCR plate. Bands were scored as present (1) or absent (0) and recorded in a binary data matrix for statistical analysis (appendix II). Fragment names were specified to reflect the primer used and approximate size of the amplification product, e.g. $A4_{1050}$ refers to a band produced by primer A4 that was approximately 1050 bp in length.

Table 4.2 – Details of RAPD primers used in the analysis.

Primer	Primer sequence (5'-3')	Number of bands scores	Band Size (bp)
A4	AATCGGGCTG	5	400, 500, 800, 1050, 1500
A10	GTGATCGCAG	6	200, 300, 350, 450, 600, 800
A19	CAAACGTCGG	5	300, 400, 500, 600, 650
B4	GGACTGGAGT	5	300, 500, 550, 600, 700
B7	GGTGACGCAG	4	350, 400, 500, 1000
C15	GACGGATCAG	6	200, 300, 400, 500, 600, 750
C16	CACACTCCAG	6	200, 300, 550, 600, 750, 1150
C17	TTCCCCCCAG	3	550, 750, 800

cpDNA RFLP

A previous study of cpDNA variation in Mediterranean species of *Senecio* identified 16 cpDNA haplotypes based on 27 restriction fragment length polymorphisms (Comes and Abbott, 2001). Samples in the present study were analysed for only those haplotypes recorded previously by Comes and Abbott (2001) in *S. gallicus*, *S. glaucus* and *S. leucanthemifolius* (Table 4.3A). This involved the detection of seven cpDNA haplotypes with the use of four enzyme/probe combinations (Table 4.3B).

Sixty-four individuals were initially analysed for cpDNA RFLP haplotype. This included 12 individuals from the Licata population, 12 from Pozallo, 20 Tunisian *S. leucanthemifolius* individuals (four from each of five different populations, Table 4.1) and 20 Tunisian *S. glaucus* ssp. *coronopifolius* individuals (five from each of four different populations, Table 4.1). However, six Tunisian *S. leucanthemifolius* individuals produced one or more polymorphisms that were difficult to interpret and were not analysed further.

Restriction digestion of genomic DNA, Southern blotting, preparation of probes and detection of haplotypes followed the protocol in section 2.7. Scoring of haplotypes was based on the sizes of fragments obtained for the four probe/enzyme combinations (Table 3B). Each polymorphism was scored as present or absent in an individual and a haplotype assigned accordingly.

Table 4.3 - Chloroplast DNA polymorphisms studied in the *Senecio* material. A – Haplotypes detected in a previous study of Mediterranean *Senecio* and their frequencies in *S. gallicus* (GAL), *S. glaucus* (GLAUC) and *S. leucanthemifolius* (LEUC) (Comes and Abbott, 2001), B – Details of polymorphisms necessary to distinguish the haplotypes present in table A. Character scores were polaried as ancestral (0) if identical to that found in the outgroup (*S. malacitanus*) of Comes and Abbott (2001).

A.	Polymorphism							Species		
Haplotype	8	9	11	12	13	14	15	GAL	GLAUC	LEUC
A	1	0	0	0	1	0	0	12	5	13
B	1	1	0	0	1	0	0	56	5	10
C	0	0	1	0	1	0	1		118	2
F	0	0	0	0	1	0	1	6	4	4
G	1	1	0	0	1	1	0	38		
H	1	1	0	1	1	0	0	3		
J	0	0	0	0	0	0	0		20	1

B.		Mutation (fragment size in kb)		
Polymorphism	Enzyme/probe	0	1	Type
8	*Cla*I/6	4.36	4.03	Deletion
9	*Cla*I/6	3.1 (+0.2)	3.3	Site loss
11	*Pst*I/1-3	2.8	Absent	?
12	*Pst*I/1-3	-	-0.05	Deletion
13	*Cfo*I/6	8.4	7.0 + 1.4	Site gain
14	*Cfo*I/6	-	+0.12	Insertion
15	*Eco*RI/8-11	2.0 (+0.4)	2.4	Site loss

4.2.4 *Statistical analysis*

Morphometric data analysis

In the morphometric analysis the widespread species *S. leucanthemifolius* was subdivided into three groups on the basis of geographic origin and each group was

treated as if it were a separate taxon. Hence the six 'taxon' groups compared were (1) Licata and Pozallo *Senecio*, (2) *S. gallicus*, (3) *S. glaucus* ssp. *coronopifolius*, (4) *S. leucanthemifolius* (Italy/Sicily), (5) *S. leucanthemifolius* (Corsica) and (6) *S. leucanthemifolius* (Tunisia). Two-way Analysis of Variance (ANOVA) was carried out using the GLM option of SAS 8.2 (SAS Institute Inc., Cary, NY, USA). Tukey tests were employed to determine whether the difference in trait values between each pair of taxa was significant. Characters that were not normally distributed were transformed into natural logarithms. Ray floret number and calyculus bract number were not normally distributed after transformation and were consequently not subjected to ANOVA. Instead, mean number of calyculus bracts was calculated for each of the six groups to allow inter-group comparison. Ray floret numbers were found to cluster around the values 8 and 13 in the Fibonacci series – 8, 13, 21 etc., and, therefore, ray number for a given individual that was not a Fibonacci number was assigned to the closest Fibonacci number before making comparisons between frequencies in each group. Prior to principal component analysis (PCA) using Minitab (ver. 13.31; Minitab Inc., 2000), untransformed data were standardised by subtracting the trait mean from each value and dividing by the trait standard deviation.

RAPD phenotype analysis

Lynch and Milligan (1994) recommend that bands absent in <3/N individuals (where N = sample size) are discarded from genetic analysis to prevent biased estimates of genetic parameters. For the full analysis (i.e. all 133 individuals), no bands were absent in <3/N individuals and hence all were included for the following analyses.

The complete data matrix (appendix II) was analysed to estimate the percentage of polymorphic loci per population, i.e. the number of polymorphic bands in the population divided by the total number of bands scored, and the Shannon-Weaver (1949) index of within population phenotypic diversity (*I*). *I* is calculated using the formula $I = -\Sigma p_i \log_2 p_i$, where p_i is the frequency of the *i*th band (King and Schaal, 1989). Shannon's index is a measure of diversity that is suitable for use with dominant markers such as RAPDs due to (i) non-reliance on Hardy-Weinberg equilibrium (Chalmers *et al.*, 1992; Yeh *et al.*, 1995; Bussell, 1999), and (ii) relative insensitivity to the inability to differentiate 1/0 heterozygotes from 1/1 homozygotes (Dawson *et al.*, 1995).

Pairwise genetic distances were estimated as Euclidean distance (Excoffier *et al.*, 1992) defined for RAPDs by Huff *et al.* (1993). Principal coordinate analysis (PCO) was conducted on the distance matrix using GenAlEx (Peakall and Smouse, 2002). Cluster analysis was also performed on the matrix using the unweighted pair group method with arithmetic averages (UPGMA) in PAUP* (Swofford, 2001). An assignment test was conducted following the method of Paetkau *et al.* (1995) using the online programme Assignment Calculator (Brzustowski, 2002). This assigns individuals (in this case the Licata and Pozallo individuals) to one of the pre-defined populations (i.e. the remaining 19 populations examined in this study) based on similarity of RAPD phenotype.

Analysis of Molecular Variance (AMOVA; Excoffier *et al.*, 1992) was performed using ARLEQUIN 2.001 (Schneider *et al.*, 2000) on the same pairwise squared Euclidean distance matrix. An initial AMOVA compared five groups of individuals, based on clusters observed in the PCO analysis, i.e. (1) Licata/Pozallo, (2) *S. gallicus*, (3) *S. glaucus*, (4) *S. leucanthemifolius* (Tunisia/Italy/Sicily) and (5) *S. leucanthemifolius* (Corsica). A second AMOVA was carried out after assigning Licata and Pozallo populations to the same group as *S. glaucus*. Finally, a third AMOVA was conducted to examine variation between Licata/Pozallo material and *S. glaucus*. For the third AMOVA, two RAPD bands ($A10_{200}$ and $C16_{750}$) were found to be absent in <3/N individuals (N = 45 individuals) and hence were excluded from analysis. Traditional *F* statistics (Wright, 1978) cannot be calculated from RAPD phenotypes due to the inability to distinguish 1/0 heterozygotes from 1/1 homozygotes. Instead, ARLEQUIN calculates analogous Φ-statistics. Pairwise Φ_{ST} values were calculated, first to measure differentiation between all populations, and then between groups of populations. The significance of these values were calculated based on permutation procedures (10,000 replicates).

cpDNA haplotype variation

The cpDNA haplotypes detected in the present study were combined for analysis with appropriate data from Comes and Abbott (2001) who had previously recorded cpDNA variation in *S. gallicus*, *S. glaucus* and *S. leucanthemifolius*, including some of the populations examined here (i.e. VNM, GAL, SE, GT, SE, CT; Table 4.1).

4.3 RESULTS

4.3.1 *Morphometric variation*

Analysis of Variance (Table 4.4) showed that there were significant differences between the means of 'taxa' for all 11 characters analysed; however, for most traits, there is considerable overlap between means. Significant variation was also observed among populations within taxa, except for capitulum length and mean ray floret length. The Licata and Pozallo material most closely resembles *S. leucanthemifolius* in leaf morphology, but is similar to all three species (*S. gallicus*, *S. glaucus* and *S. leucanthemifolius*) in capitulum type. For only one character (mean ray floret length) was Licata/Pozallo material significantly different from all three of these species. *S. glaucus* and *S. gallicus* produced the least number of calyculus bracts (2.95 ± 0.39 (SE) and 1.17 ± 0.38 per individual respectively), whereas Italian/Sicilian and Tunisian *S. leucanthemifolius* had the greatest means for this trait (14.50 ± 1.67) and 14.77 ± 1.15, respectively). Corsican *S. leucanthemifolius* produced on average 5.75 ± 0.39 calyculus bracts per capitulum, which was very similar to the mean number produced by Licata/Pozallo material (5.17 ± 0.58). Number of ray florets was calculated as frequencies of occurrence of eight or 13 ray florets per capitulum. Eleven of the 12 individuals from Licata and Pozallo had 13 ray florets per capitulum, as did most individuals of *S. gallicus*, *S. glaucus* and Corsican *S. leucanthemifolius*. However, over half of the Italian and Tunisian *S. leucanthemifolius* individuals produced only eight ray florets per capitulum (seven of 12 and 15 of 22 respectively).

The first two axes of the PCA plot (Fig. 4.3) account for 34.0 and 24.1% of the total variation. PCA1 separates *Senecio glaucus* and *S. gallicus* from *S. leucanthemifolius,* while the second axis separates *S. gallicus* from *S. glaucus*. No additional structuring was observed when PCA3 was plotted (not shown). The Licata/Pozallo material appears intermediate to *S. glaucus* and *S. leucanthemifolius*, but also similar in morphology to some individuals of *S. gallicus*. Corsican material of *S. leucanthemifolius* is most similar morphologically to Licata/Pozallo material and tends to be different from other material of *S. leucanthemifolius* examined.

Table 4.4 – Mean (± SE) of 13 morphological characters measured on three species of *Senecio* and the material of unknown origin from Licata and Pozallo, Sicily. The widespread species *S. leucanthemifolius* is divided into three groups based on geographic origin. N is the number of individuals sampled per group. ANOVA was carried out on all variables except mean number of calyculus bracts and number of ray florets. Shared letters in superscript implies non-significant ($P > 0.05$) difference in character mean (Tukey test).

	Licata/Pozallo	*S. gallicus*	*S. glaucus*	*S. leuc* (Italy)	*S. leuc* (Corsica)	*S. leuc* (Tunisia)	*P* value	
Character	*N*=12	*N*=18	*N*=21	*N*=12	*N*=12	*N*=22	Among populations within taxa	Among taxa
Inflorescence length (mm)	28.37 ± 3.26^{a,b}	33.92 ± 1.58^{b}	50.69 ± 3.86^{c}	21.19 ± 1.39^{a,b}	60.74 ± 7.20^{c}	17.09 ± 0.78^{a}	0.023	0.011
Capitulum length (mm)	9.95 ± 0.31^{d,e}	9.34 ± 0.28^{c,d}	10.52 ± 0.14^{e}	8.32 ± 0.29^{a,b}	7.62 ± 0.16^{a}	8.67 ± 0.21^{b,c}	0.189	0.021
Capitulum Base Width (mm)	5.22 ± 0.17^{b,c}	3.90 ± 0.14^{a}	5.39 ± 0.10^{c}	4.19 ± 0.14^{a}	4.89 ± 0.16^{b}	4.04 ± 0.11^{a}	0.027	0.001
Mean ray floret length (mm)	11.54 ± 0.32^{b}	8.46 ± 0.43^{a}	13.13 ± 0.42^{c}	7.06 ± 0.30^{a}	7.23 ± 0.43^{a}	8.06 ± 0.29^{a}	0.117	0.001
Mean ray floret width (mm)	3.35 ± 0.18^{b}	2.25 ± 0.08^{a}	2.82 ± 0.08^{a,b}	2.58 ± 0.08^{a,b}	2.54 ± 0.11^{a,b}	2.55 ± 0.09^{a,b}	0.024	0.049
Midleaf length (mm)	57.38 ± 2.83^{a}	84.27 ± 6.51^{b}	74.40 ± 1.89^{b}	71.01 ± 3.99^{a,b}	55.78 ± 2.67^{a}	61.75 ± 2.94^{a}	<0.001	0.004
Midleaf area (mm^2)	575.6 ± 28.6^{b,c}	1439.0 ± 266.0^{e}	556.9 ± 50.0^{a,b,c}	798.9 ± 32.9^{b,c,d}	331.0 ± 28.3^{a,b}	889.4 ± 88.9^{b,c,d}	<0.001	<0.001
Midleaf perimeter (mm)	277.3 ± 12.6^{a}	684.4 ± 52.2^{c}	434.6 ± 26.7^{b}	193.3 ± 14.7^{a}	165.3 ± 9.0^{a}	202.0 ± 18.3^{a}	0.013	<0.001
Standardised midleaf area[1,5]	0.43 ± 0.02^{b}	0.43 ± 0.02^{b}	0.31 ± 0.01^{a}	0.41 ± 0.02^{b}	0.32 ± 0.01^{a}	0.49 ± 0.01^{c}	<0.001	<0.001
Midleaf dissection[2]	11.72 ± 0.63^{b}	21.05 ± 1.98^{d}	18.55 ± 0.74^{c}	6.83 ± 0.45^{a}	9.40 ± 0.65^{a,b}	7.06 ± 0.56^{a}	<0.001	<0.001
Standardised leaf perimeter[3]	4.99 ± 0.37^{b}	8.56 ± 0.69^{c}	5.85 ± 0.33^{b}	2.70 ± 0.09^{a}	3.00 ± 0.17^{a}	3.34 ± 0.21^{a}	<0.001	0.001
Mean number calyculus bracts	5.17 ± 0.58	1.17 ± 0.38	2.95 ± 0.39	14.50 ± 1.67	5.75 ± 0.39	14.77 ± 1.15		
Number of ray florets[4]	1 • 11	3 • 15	3 • 18	7 • 5	1 • 11	15 • 7		

[1] Square root of midleaf area divided by the midleaf length; [2] Midleaf perimeter divided by square root of midleaf area; [3] Midleaf perimeter divided by midleaf length; [4] x • y where x = number of individuals with 8 ray florets, y = number of individuals with 13 ray florets; [5] This character was natural logarithm transformed to achieve normality prior to ANOVA

Table 4.4 to go here **(morphometrics ANOVA)**

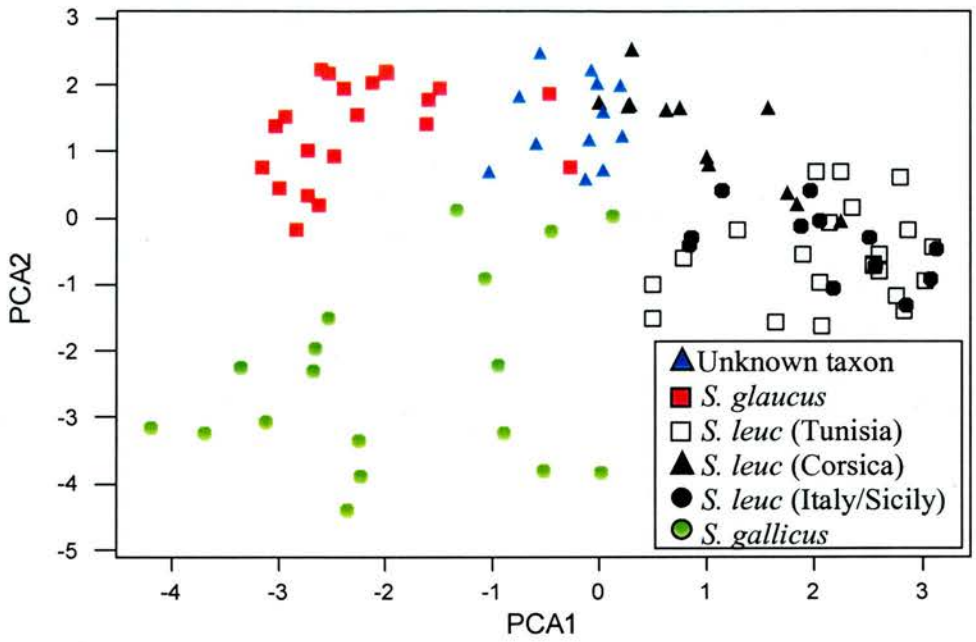


Figure 4.3 – Principal components analysis (PCA) plot of the first two axes based on 11 morphological variables in *Senecio glaucus*, *S. leucanthemifolius*, *S. gallicus* and the populations of taxonomic uncertainty.

4.3.2 *RAPD variation*

Values of percentage polymorphic loci and Shannon's diversity index (Table 4.5) in Licata/Pozallo material were lower than in Tunisian *S. glaucus*, but similar to those in the groupings of *S. leucanthemifolius* from Italy/Sicily and Tunisia. The lowest values for these two measures of RAPD diversity were recorded in Corsican *S. leucanthemifolius* and *S. gallicus* populations.

Seven of the 40 RAPD fragments surveyed were, to some degree, private (i.e. found in one taxon only at a frequency of 0.1 to 0.4) or specific (i.e. have a frequency of >0.4 in one taxon only) to *S. gallicus, S. glaucus* or *S. leucanthemifolius* (Table 4.6). One marker specific to *S. gallicus* ($A4_{1050}$) was absent from Licata/Pozallo material, while of the two taxon-specific markers and two private bands found in *S. glaucus*, three were present in Licata/Pozallo material. Two private bands, but no taxon-specific bands were recorded in *S. leucanthemifolius*, both of which were absent from the Licata and Pozallo populations. Licata/Pozallo material contained no unique bands.

Table 4.5 – Percentage polymorphic loci (% P L) and Shannon's index of diversity (*I*) calculated among 21 populations of *Senecio*.

Taxon	Population	% P L	*I*
Unknown	LI	32.5	0.180
	PO	35.0	0.161
S. glaucus	EN	42.5	0.244
	EJ	52.5	0.297
	SB	35.0	0.188
	HS	37.5	0.201
S. leucanthemifolius (Tunisia)	MA	30.0	0.180
	TA	32.5	0.183
	BI	35.0	0.184
	HA	45.0	0.220
	HM	25.0	0.147
S. leucanthemifolius (Italy/Sicily)	SL	35.0	0.193
	CT	32.5	0.184
	GT	30.0	0.166
	SE	45.0	0.238
	GAL	37.5	0.212
S. leucanthemifolius (Corsica)	TP	15.0	0.098
	PRO	15.0	0.097
S. gallicus	MVV	20.0	0.110
	SR	20.0	0.110
	VNM	7.5	0.047

Table 4.6 – Summary of seven RAPD products which were private (P) or taxon-specific (TS) to either *S. gallicus*, *S. leucanthemifolius* or *S. glaucus* and their frequencies in each species and the material from Licata and Pozallo.

RAPD product	Category	*S. gallicus*	*S. glaucus*	*S. leucanth.*	Licata/Pozallo
$A4_{1050}$	TS	**0.667**	0.000	0.000	0.000
$A10_{200}$	TS	0.000	**0.923**	0.000	**1.000**
$A10_{300}$	P	0.000	**0.269**	0.000	0.000
$A19_{400}$	P	0.000	**0.192**	0.000	**0.158**
$C15_{750}$	TS	0.000	**0.692**	0.000	**1.000**
$A19_{300}$	P	0.000	0.000	**0.205**	0.000
$C15_{300}$	P	0.000	0.000	**0.288**	0.000

In the plot of principal coordinates (Fig. 4.4), PCO1 separated *S. glaucus* and Licata/Pozallo material from *S. gallicus* and *S. leucanthemifolius,* and *S. gallicus* from Corsican *S. leucanthemifolius*. Genetic substructuring within *S. leucanthemifolius* is evident, with individuals from Italy/Sicily and Tunisia separated from Corsican material along PCO2. Moreover, a plot of PCO1 against PCO3 further subdivided Tunisian and Italian/Sicilian *S. leucanthemifolius*. UPGMA analysis of RAPD phenotypes revealed a similar pattern of divergence to that exhibited by the PCO (Fig. 4.5). What is apparent from both analyses is the close similarity of Licata/Pozallo material to Tunisian *S. glaucus* ssp. *coronopifolius*. In addition, it is evident that Corsican *S. leucanthemifolius* is dissimilar to other material of the species surveyed.

The results of the assignment test conducted on Licata/Pozallo individuals were of particular interest in that 17 of 19 individuals from this source were assigned to an inland population of *S. glaucus* at Sbeitla (SB), Tunisia, while the two remaining individuals were assigned to other *S. glaucus* populations.

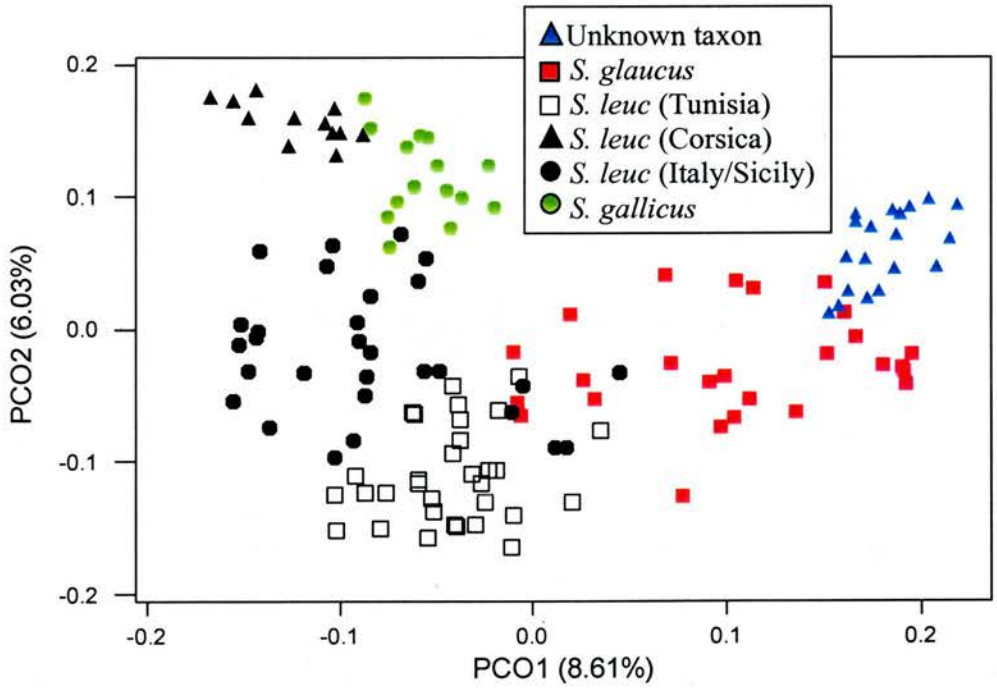


Figure 4.4 – Principal coordinates analysis (PCO) plot based on 40 RAPD products in *Senecio glaucus*, *S. leucanthemifolius* (*S. leuc*), *S. gallicus* and the populations of taxonomic uncertainty.

Figure 4.5 – UPGMA dendrogram calculated using the distance measure defined by Huff *et al.* (1993) based on 40 RAPD products in *Senecio glaucus*, *S. gallicus*, *S. leucanthemifolius*, and the populations of taxonomic uncertainty.

Analysis of Molecular Variance (AMOVA) revealed that Φ_{ST} values were lower within than between groups (Table 4.7). All pairwise Φ_{ST} values calculated between populations were significant except for two, between populations BI/HA and BI/HM of Tunisian *S. leucanthemifolius*, which were separated by only ~100 km. Between group Φ_{ST} values show that the Licata/Pozallo material is most similar to *S. glaucus* ssp. *coronopifolius* (Φ_{ST} = 0.183) and least similar to *S. gallicus* (Φ_{ST} = 0.557) and Corsican *S. leucanthemifolius* (Φ_{ST} = 0.565).

An AMOVA conducted on groups corresponding to the five clusters evident in the PCO plot, i.e. Licata/Pozallo material, *S. gallicus*, *S. glaucus*, Corsican *S. leucanthemifolius* and other *S. leucanthemifolius* (Table 4.8A) showed that variation between groups (28.7% of total), among populations within groups (26.2%) and within populations (45.1%) was in each case significant (P<0.001). A second AMOVA in which Licata/Pozallo material was grouped with *S. glaucus* (Table 4.8B) revealed a similar partitioning of variation. Finally, a third AMOVA, which compared Licata/Pozallo material with *S. glaucus*, showed that the variation between these two groups (16.0%) was not quite significant (P=0.068; Table 4.8C).

Table 4.7 – Interpopulation Φ_{ST} values (left of diagonal) calculated among 21 populations of *Senecio*. All Φ_{ST} values were significant ($P<0.05$) except those underlined. Values to the right of the diagonal are among group Φ_{ST} values.

Taxon	Unknown		*S. glaucus*				*S. leucanthemifolius* (Tunisia)					*S. leucanthemifolius* (Sicily/Italy)					*S. leuc* (Corsica)		*S. gallicus*		
Pop	LI	PO	EN	EJ	SB	HS	MA	TA	BI	HA	HM	SL	CT	GT	SE	GAL	TP	PRO	MVV	SR	VNM
LI	0.000																				
PO	0.187	0.000			0.183				0.468					0.401			0.565			0.557	
EN	0.359	0.386	0.000																		
EJ	0.336	0.417	0.197	0.000																	
SB	0.190	0.273	0.239	0.222	0.000				0.232					0.296			0.472			0.399	
HS	0.274	0.347	0.141	0.225	0.262	0.000															
MA	0.571	0.589	0.261	0.379	0.470	0.374	0.000														
TA	0.542	0.594	0.318	0.370	0.506	0.405	0.177	0.000													
BI	0.513	0.547	0.246	0.340	0.406	0.303	0.241	0.194	0.000					0.225			0.468			0.428	
HA	0.526	0.548	0.239	0.353	0.422	0.345	0.246	0.187	0.015	0.000											
HM	0.509	0.546	0.247	0.270	0.372	0.315	0.191	0.268	0.026	0.178	0.000										
SL	0.577	0.606	0.493	0.486	0.531	0.557	0.424	0.380	0.438	0.449	0.422	0.000									
CT	0.463	0.525	0.370	0.425	0.481	0.389	0.458	0.396	0.395	0.413	0.361	0.347	0.000								
GT	0.556	0.611	0.434	0.448	0.538	0.479	0.445	0.485	0.404	0.442	0.433	0.351	0.459	0.000			0.343			0.365	
SE	0.614	0.673	0.490	0.420	0.610	0.567	0.521	0.406	0.499	0.429	0.499	0.487	0.458	0.521	0.000						
GAL	0.565	0.622	0.452	0.395	0.539	0.549	0.494	0.421	0.463	0.445	0.436	0.403	0.409	0.425	0.233	0.000					
TP	0.702	0.731	0.627	0.639	0.707	0.696	0.685	0.651	0.648	0.623	0.706	0.694	0.726	0.636	0.627	0.539	0.000				
PRO	0.688	0.704	0.606	0.612	0.688	0.664	0.566	0.651	0.658	0.654	0.666	0.636	0.694	0.607	0.671	0.644	0.698	0.000		0.498	
MVV	0.642	0.684	0.504	0.571	0.627	0.542	0.589	0.621	0.552	0.548	0.612	0.676	0.663	0.544	0.692	0.632	0.714	0.745	0.000		
SR	0.635	0.678	0.555	0.546	0.627	0.557	0.628	0.636	0.569	0.577	0.612	0.575	0.613	0.547	0.667	0.602	0.782	0.728	0.456	0.000	
VNM	0.602	0.663	0.465	0.457	0.602	0.487	0.560	0.501	0.443	0.453	0.580	0.598	0.612	0.522	0.583	0.525	0.709	0.755	0.495	0.597	0.000

Table 4.8 – Results of the analysis of molecular variance (AMOVA). (A) – based on five groups [1-Licata/Pozallo, 2-*S. glaucus*, 3-*S. gallicus*, 4-*S. leucanthemifolius* (Tunisia and Italy/Sicily), 5-*S. leucanthemifolius* (Corsica)]. (B) – Four groups, as (A) except Licata/Pozallo and *S. glaucus* combined. (C) Two groups, Licata/Pozallo and *S. glaucus*. For (C) only 38 of the 40 scored products were used in the analysis (see text).

	Source of variation	d.f.[a]	Sum of squares	Variance components	% variation	P^b
A	Among groups	4	243.434	1.93866	28.70	<0.001
	Among populations within groups	16	224.154	1.76982	26.20	<0.001
	Within populations	112	341.321	3.04751	45.11	<0.001
	Total	132	808.910	6.75599		
B	Among groups	3	213.212	1.93620	28.19	<0.001
	Among populations within groups	17	254.376	1.88554	27.45	<0.001
	Within populations	112	341.321	3.04751	44.36	<0.001
	Total	132	808.190	6.86925		
C	Among groups	1	30.222	0.86132	16.04	0.068
	Among populations within groups	4	40.998	0.93609	16.80	<0.001
	Within populations	39	137.046	3.51401	67.15	<0.001
	Total	44	208.267	5.31142		

[a] Degrees of freedom. [b] Level of significance based on 10000 permutations

4.3.3 *Chloroplast DNA variation*

The results of the previous survey of cpDNA variation in *S. gallicus*, *S. leucanthemifolius* and *S. glaucus* conducted by Comes and Abbott (2001) showed that haplotypes B and G occurred at high frequency in *S. gallicus*, while haplotype A was present at high frequency in Italian and Sicilian *S. leucanthemifolius*. In addition, *S. glaucus* from Israel contained haplotype C at high frequency and haplotypes A, B, F and J at lower frequency. The present analysis established that Licata/Pozallo material

contains haplotypes A, B, C, F, H and J, all of which are also present in Tunisian *S. glaucus* (Fig. 4.6). In contrast, haplotype B is absent from material surveyed of *S. leucanthemifolius*, and haplotype C is absent from *S. gallicus* and only present in two individuals of *S. leucanthemifolius* examined.

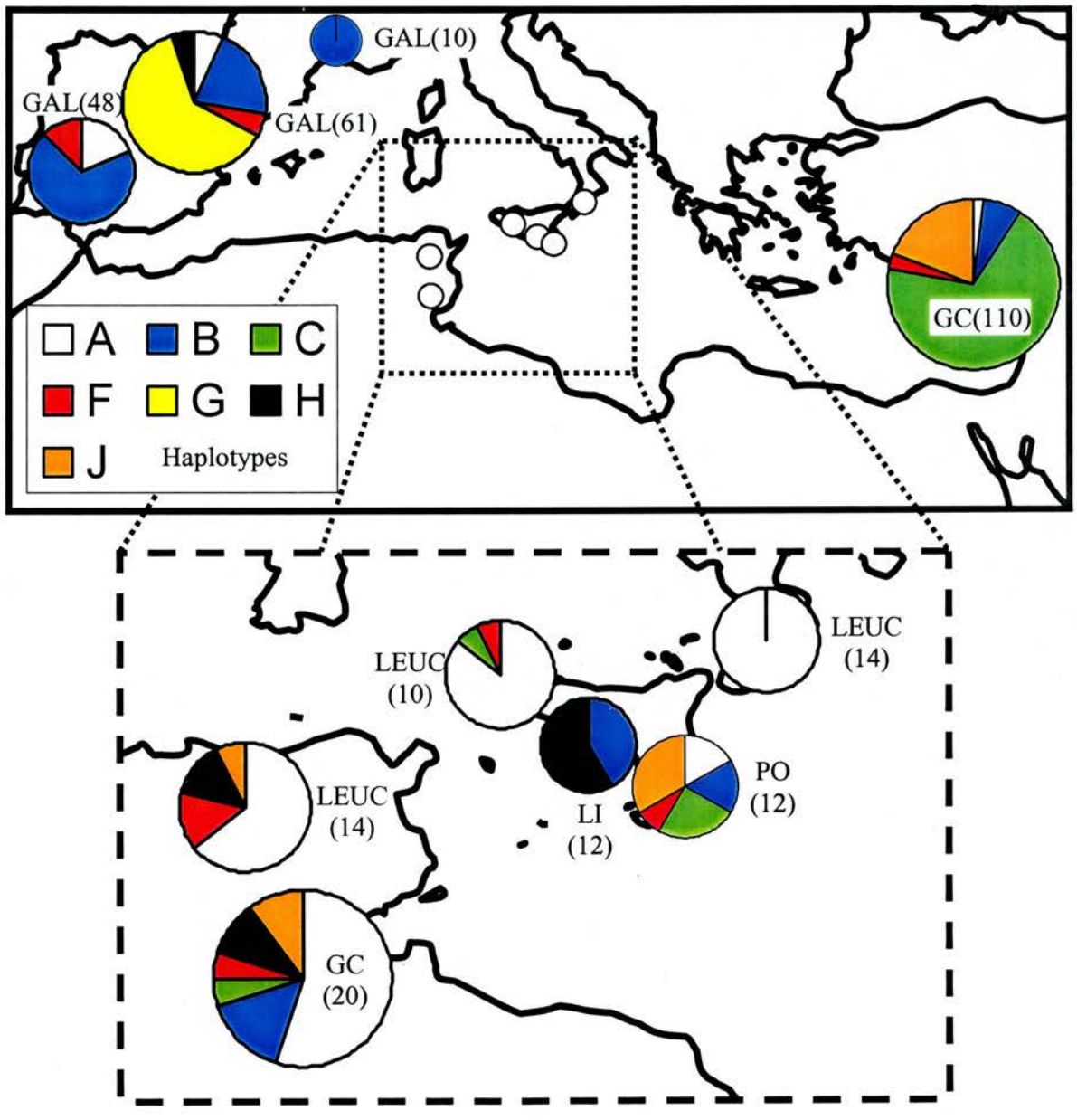


Figure 4.6 – cpDNA Haplotypes present in the material from Licata (LI), Pozallo (PO) and several groups of *S. gallicus* (GAL), *S. leucanthemifolius* (LEUC) and *S. glaucus* ssp. *coronopifolius* (GC) populations around the Mediterranean. Number of individuals analysed per population is given in brackets.

4.4 Discussion

The results of the morphometric analysis showed that Licata/Pozallo *Senecio* material is morphologically intermediate to *S. glaucus* ssp. *coronopifolius* and *S. leucanthemifolius* var. *leucanthemifolius*, but is also similar to some individuals of *S. gallicus* examined. Consequently, the material is not easy to assign to a particular taxon even when compared with related material raised under the same set of conditions. It is not surprising, therefore, that there has been taxonomic confusion with regard to these plants in the past (Alexander, 1979; Comes and Abbott, 2001; Coleman, 2003). Plants from Licata/Pozallo were most similar in leaf morphology to *S. leucanthemifolius*, but could not easily be distinguished from each of the three species it was compared with, based on capitulum type. A further finding to emerge from the morphometric analysis was that Corsican material of *S. leucanthemifolius* var. *leucanthemifolius* is divergent from Tunisian and Italian/Sicilian material and might be regarded as a distinctive island form of this species.

Although Licata/Pozallo material could not easily be assigned to a particular taxon based on its morphology, a survey of RAPD variation showed it to be very similar to Tunisian *S. glaucus* ssp. *coronopifolius*. Licata/Pozallo material clustered closely with Tunisian *S. glaucus* ssp. *coronopifolius* in both a PCO plot and UPGMA tree constructed from the RAPD data set. In addition, Licata/Pozallo material contained three of four private/taxon-specific RAPD bands present in Tunisian *S. glaucus* ssp. *coronopifolius*. In contrast, the only taxon-specific band found in *S. gallicus* was absent from Licata/Pozallo material as were both private bands present in *S. leucanthemifolius*. No RAPD bands were recorded that were specific to Licata/Pozallo material indicating that this material is most probably of recent origin and has not had sufficient time to accumulate new RAPD alleles. Assignment tests indicated that all 19 Licata/Pozallo plants could be assigned to populations of Tunisian *S. glaucus* ssp. *coronopifolius* with 17 individuals being assigned to the inland population at Sbeitla. Taken overall the RAPD results strongly indicate that Licata/Pozallo material is a variant form of *S. glaucus* ssp. *coronopifolius* derived from material introduced to Sicily from North Africa.

The results of the survey of cpDNA haplotype variation showed that Licata/Pozallo material was highly polymorphic containing haplotypes A, B, C, F, H and J. Haplotypes A, F, H and J were recorded in at least two of the putative parent

species; however, only *S. glaucus* was shown to contain all six haplotypes present in Licata/Pozallo plants. Haplotype B was absent from *S. leucanthemifolius*, and C and J were absent from *S. gallicus* (and rare in *S. leucanthemifolius*). This finding provides a further pointer that Licata/Pozallo material is derived from *S. glaucus* ssp. *coronopifolius* introduced to Sicily from North Africa.

The high level of cpDNA variation found in Licata/Pozallo material stands in contrast to the relatively low levels of RAPD diversity recorded in this material. In this respect, it is worth noting that Coleman (2003) also observed lower levels of RAPD and ISSR diversity in Licata/Pozallo material relative to two Tunisian populations of *S. glaucus* that were analysed in a more widespread study of the species. The presence of five cpDNA haplotypes in Licata/Pozallo material indicates that there have been multiple introductions of the species to Sicily. In contrast, the low levels of RAPD diversity suggest few introductions and/or material has passed through a bottleneck during or following colonization. This paradox might be resolved by invoking the possibility that Licata/Pozallo material was subjected to intense selection after its introduction resulting in a selective sweep of the nuclear genome that caused considerable loss of allelic variation at RAPD loci linked to genes subject to selection. In contrast, the cpDNA genome was not subject to selection and consequently diversity for this genome has been maintained among Licata/Pozallo plants. This, of course, is a highly speculative explanation, and it is difficult to envisage that intense selection may have acted across the nuclear genome in the manner that would be required. For the time being, therefore, the markedly different levels of cpDNA and RAPD variation resolved in Licata/Pozallo material remains as an unresolved conundrum.

Conclusions

From the results presented it is concluded that Licata/Pozallo material is a variant form of *Senecio glaucus* ssp. *coronopifolius*. This variant, which for certain morphological characters is not easily distinguished from either *S. gallicus* or *S. leucanthemifolius*, was most probably introduced to Sicily from North Africa in the relatively recent past. It is not possible to rule out that the material is long-standing in Sicily and a relict of a wider distribution of *S. glaucus* in the northern part of the Mediterranean basin. However, the species is currently not known from elsewhere in

this part of the Mediterranean and an absence of unique RAPD alleles argues against the possibility of it being long-standing in Sicily and isolated from the distribution of *S. glaucus* in North Africa. The likelihood of a relatively recent introduction, possibly as a consequence of human activity, is given additional weight by the fact that the material is restricted to sand dunes next to the southern ports of Licata and Pozallo and is found nowhere else in Sicily or mainland Italy. It is feasible that since its introduction the material has been subject to convergent selection for certain characters, which may explain its current intermediate morphology to *S. glaucus*, *S. leucanthemifolius* and *S. gallicus*.

CHAPTER 5 - ALLOPOLYPLOID ORIGIN OF TETRAPLOID *SENECIO MOHAVENSIS* (ASTERACEAE): EVIDENCE FROM DNA SEQUENCE OF TWO SINGLE-COPY NUCLEAR GENES

ABSTRACT

Polyploidy is an important mechanism in the evolution of many plant taxa; however the processes that contribute to the success of polyploid species are still poorly understood. These process will differ fundamentally depending on whether the species arose through auto- or allopolyploidy. DNA sequence variation of two nuclear genes was examined to establish the polyploid origin of two disjunct subspecies of the tetraploid *Senecio mohavensis* (Asteraceae). This species is morphologically very similar to diploid *S. flavus*, but differs from it in producing flower heads (capitula) that are radiate. It is uncertain whether the species is an allotetraploid or is of autopolyploid origin. Two distinct copies of a *cycloidea* homologue (*Scyc2*) and also of *PgiC* were found to be present in both subspecies of *S. mohavensis* indicating an allotetraploid origin. Phylogenetic analysis of these sequences and those of related species provides strong evidence that *S. flavus* and the widespread *S. glaucus* are the parents of *S. mohavensis*. Reasons why *S. mohavensis* closely resembles *S. flavus* in overall morphology, despite being an allotetraploid, are discussed.

5.1 INTRODUCTION

Polyploidy is of great importance in plant evolution, with ~30% (Stebbins, 1950) to 70% (Masterson, 1994) of angiosperms, and potentially up to 95% of pteridophytes (Grant, 1981), having at least one round of polyploidy in their phylogenetic histories. Despite this importance, we are only beginning to understand the processes that promote the origin, establishment and success of polyploids (Soltis and Soltis, 1999, 2000; Soltis *et al.*, 2004).

Polyploids arise through an increase in chromosome number before or after interspecific (allopolyploids) or intraspecific (autopolyploids) hybridisation, or, rarely, via a somatic mutation. Most polyploids are thought to arise via a triploid intermediate which produces viable gametes at low frequency (Ramsey and Schemske, 1998).

Traditionally it was presumed that a polyploid species would originate once, and thus genetic variation would be low, suggesting that polyploidy was an evolutionary dead-end (Wagner, 1983). In contrast, the majority of polyploid species studied in detail have been proven to be of multiple origin (reviewed in Soltis and Soltis, 1993), and hence genetic variation may be relatively high in the neospecies, potentially increasing its chance of establishment, especially if matings between independently-derived populations subsequently occur, thus creating new genetic combinations (Soltis and Soltis, 1993; Soltis *et al.*, 1995).

The success of polyploids has been attributed to increased heterozygosity and allelic diversity relative to the diploid parents (Roose and Gottlieb, 1976; Levin, 1983; Soltis and Soltis, 1993, 1995). Polyploid species sometimes have broader ecological tolerances than their parent species (Levin, 1983; Thompson and Lumaret, 1992; Soltis and Soltis, 1995) and may also exhibit new interactions with herbivores (Nuismer and Thompson, 2001) and pollinators (Segraves and Thompson, 1999).

Following polyploid formation, genomic changes (i.e. reorganisation by translocations and recombination between chromosomes from the different parents) can be rapid, and has been reported in a number of crop species (Song *et al.*, 1995; Leitch and Bennett, 1997; Ozkan *et al.*, 2001; Shaked *et al.*, 2001). Early generation polyploids inherit two copies of each locus (homoeologues), and over time the two copies may exhibit different fates, such as silencing, development of a new function

(sub-functionalisation), or tissue-specific expression (Comai, 2000; Wendel, 2000; Adams *et al.*, 2003; Osborn *et al.*, 2003). Such alterations can be so extensive that the polyploid eventually behaves like a diploid (Chen *et al.*, 1997; Gaut and Doebley, 1997).

Clearly, the mechanisms which contribute to the success of polyploid species (increased heterozygosity and alteration in ecological tolerances) and the genomic changes which can take place (chromosomal rearrangements, gene silencing, sub-functionalisation) have major evolutionary implications, and are currently the subject of considerable analysis. However, these mechanisms will fundamentally differ depending on the mode of origin, i.e. through auto- or allopolyploidy. Both auto- and allopolyploids will experience gene dosage effects due to the presence of multiple copies of each gene (e.g. Galili *et al.*, 1986; Guo *et al.*, 1996), whereas for allopolyploids, the interactions between two differentiated genomes are likely to have important consequences (McClintock, 1984; Wendel, 2000; Comai *et al.*, 2003; Riddle and Birchler, 2003).

This study was carried out to investigate the origin of tetraploid *Senecio mohavensis* A. Gray (Asteraceae). Morphologically, this species closely resembles diploid *S. flavus* (Dechne.) Sch. Bip. and is distinct from other members of sect. *Senecio* which suggests that it may have arisen from *S. flavus* by autopolyploidy. However, it possesses radiate flower heads, unlike *S. flavus* which is discoid (Fig. 5.1 A, B). The acquisition of ray florets by *S. mohavensis* may have resulted from an allopolyploid origin, by mutation, or introgression after an autopolyploid origin. Florets in the *Senecio* capitulum are pentamerous. In disc florets all five parts are fused to form a corolla, whereas in ray florets, three parts are elongated to form a ligule. Non-radiate capitula contain only disc florets whereas in radiate capitula, the outer whorl is made up of ray florets.

Senecio mohavensis is an annual tetraploid (2n=40) species of section *Senecio* that is represented as two distinct subspecies displaying an Old World - New World disjunction (Coleman *et al.*, 2001). The type subspecies occurs in Arizona and California, whereas *S. mohavensis* subsp. *breviflorus* (Kadereit) M. Coleman is found from Egypt, through the Near East and Arabia to the Thar Desert in Pakistan (Fig. 5.2). The closely related diploid species, *S. flavus* (2n=20) also has a disjunct

distribution being found in Namibia and across North Africa (Fig. 5.2). Both species are highly self-compatible and produce sterile triploid hybrids when crossed.

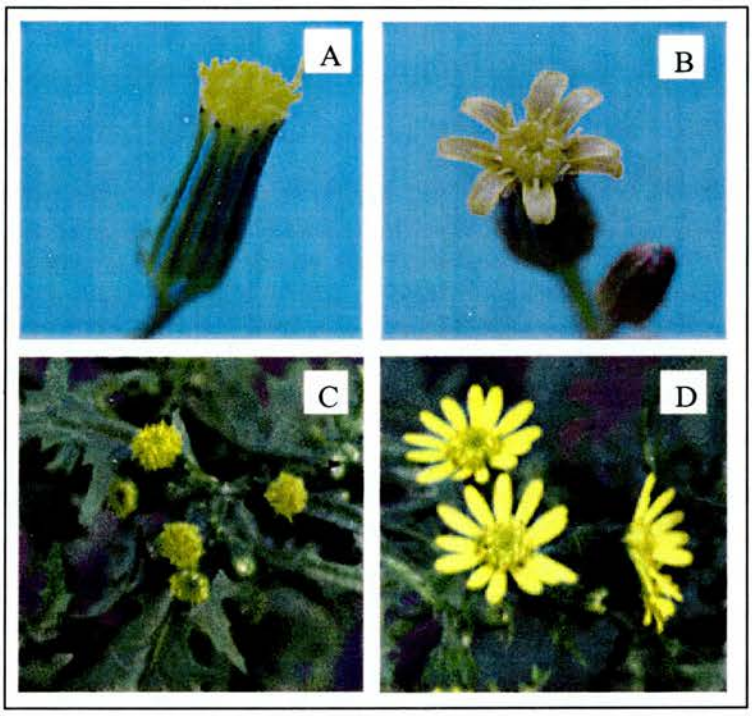


Figure 5.1 – Non-radiate (A, C) and radiate (B, D) capitulum types in *Senecio*. Species shown are *Senecio flavus* (A), *S. mohavensis* subsp. *breviflorus* (B) and *S. vulgaris* var. *vulgaris* (C) and *S. vulgaris* var. *hibernicus* (D).

Until recently, it was thought that subsp. *breviflorus* was a subsp. of the diploid *S. flavus*, due to its very similar gross morphology (Kadereit, 1984a). However, the finding that subsp. *breviflorus* was genetically more similar (based on isozyme polymorphism) to *S. mohavensis* than it was to *S. flavus* subsp. *flavus* (Liston *et al.*, 1989), and confirmation that subsp. *breviflorus* is tetraploid, led to it being transferred to *S. mohavensis* as a new subspecies (Coleman *et al.*, 2001).

Senecio flavus has been considered distinct from other annual Mediterranean *Senecio* species, both morphologically (Alexander, 1979), and genetically, with its closest relative being South African *S. englerianus* O. Hoffm., suggesting a southern African origin (Coleman *et al.*, 2003). DNA sequence variation of the internal transcribed spacer (ITS) of nuclear ribosomal DNA (rDNA) places the divergence time between north and south African *S. flavus* at 0.15 ± 0.22 million years ago

(Coleman *et al.*, 2003), consistent with a late Pleistocene migration, possibly along an east African arid track linking north and south Africa as suggested by Liston *et al.* (1989) and Liston and Kadereit (1995). *Senecio mohavensis* subsp. *mohavensis* appears to lack close relatives in North America being most closely related to the Mediterranean subsp. *breviflorus* and *S. glaucus* (Coleman *et al.*, 2003). This suggests that it arose in North America following long-distance dispersal from south-west Asia (Liston and Kadereit, 1995). The absence of ITS sequence divergence between the two subspecies of *S. mohavensis* provides an estimated divergence within the last 0.15 million years (Myr) (Coleman *et al.*, 2003). Long distance dispersal mediated via birds has been proposed for the disjunction (Liston *et al.*, 1989; Liston and Kadereit, 1995) and is supported by the discovery of pappus characters (grapple-like tips and firm attachment to the fruit) that suggest adaptation to epizoochory (Coleman *et al.*, 2003). Similarly, long distance dispersal by birds is also a possible explanation of the disjunction of *S. flavus* in Africa (Coleman *et al.*, 2003).

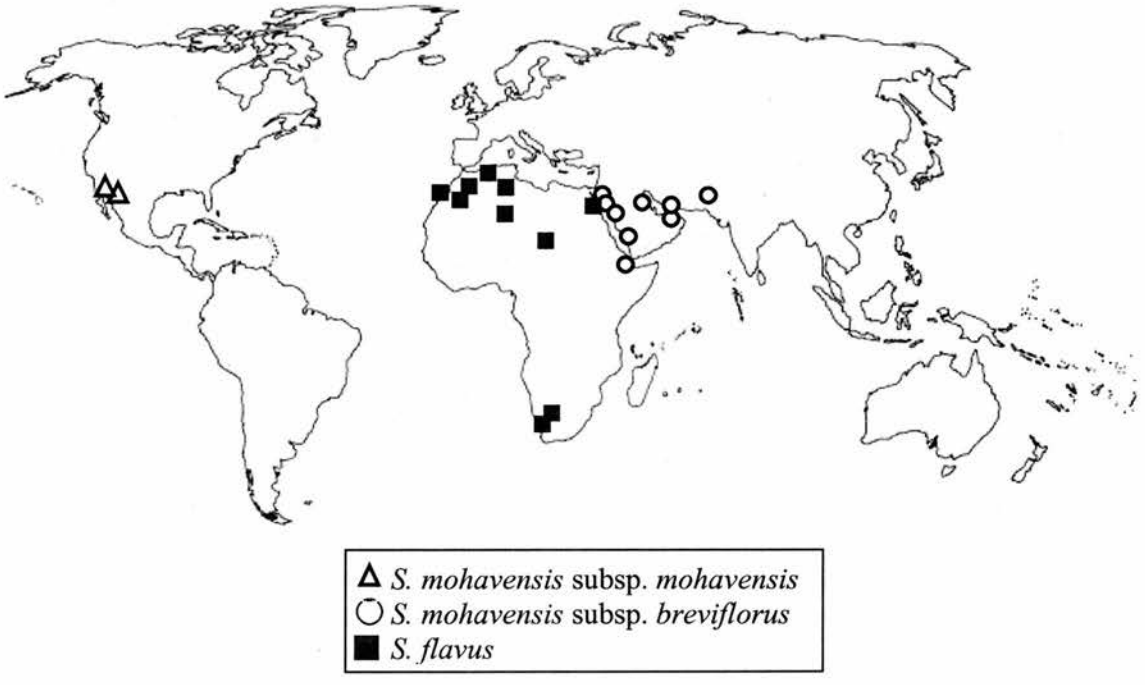


Figure 5.2 – The disjunct distributions of *Senecio mohavensis* and closely related *S. flavus*.

The similarity in ITS sequence between *S. mohavensis* and *S. glaucus* subsp. *coronopifolius*, which contrast with that of *S. flavus* (Comes and Abbott, 2001; Coleman *et al.*, 2003), suggests that *S. mohavensis* is of allotetraploid origin, then concerted evolution of the ITS repeats has possibly occurred towards the *S. glaucus*-like parent. Coleman (pers. comm.) failed to record a *S. flavus*-like ITS sequence in *S. mohavensis* despite sequencing several cloned ITS repeats from this species. An interspecific hybrid is expected to inherit the ITS repeats from both parents (e.g. Sang *et al.*, 1995), but over time homogenisation of repeat types towards one parent may occur (e.g. Hillis *et al.*, 1991; Wendel *et al.*, 1995a; Franzke and Mummenhoff, 1999).

Additional evidence for an allopolyploid origin of *Senecio mohavensis* comes from two surveys of chloroplast DNA (cpDNA) restriction fragment length polymorphism (RFLP). Firstly, Liston and Kadereit (1995) found that both subspecies of *S. mohavensis* possess the same haplotype which is markedly different from that possessed by *S. flavus*. The cpDNA of *S. mohavensis* was in fact very similar to that of *S. squalidus*, a species included in the analysis as outgroup. More recently, a study of cpDNA variation among 18 Mediterranean *Senecio* species revealed that subsp. *breviflorus* possessed a haplotype that was identical to that found in *S. glaucus* subsp. *coronopifloius* (Comes and Abbott, 2001).

However, an allotetraploid origin of *S. mohavensis* is not supported by the result from a study of random amplified polymorphic DNA (RAPDs), which indicated that *S. flavus* and subsp. *breviflorus* are very similar genetically (Comes and Abbott, 2001). Additionally, the observation that, morphologically, *S. mohavensis* is not intermediate between *S. flavus* and *S. glaucus*, but is much more similar to *S. flavus* (Kadereit, 1984a; Coleman *et al.*, 2001) argues against an allopolyploid origin.

Therefore, an alternative hypothesis is that introgression of nuclear and plastid genetic material of *S. glaucus* took place before or after autotetraploidisation of *S. flavus* that gave rise to *S. mohavensis*. In this scenario, the cpDNA and ITS of *S. mohavensis* is derived from *S. glaucus* or a close relative, which must have acted as the maternal parent as cpDNA is maternally inherited in *Senecio* (Harris, 1990). This hypothesis has been advanced by Comes and Abbott (2001) as an alternative to the hypothesis of allotetraploid origin.

The aim of the present study was to obtain additional nuclear DNA sequence variation which might help to distinguish between an auto- and allotetraploid origin of

S. mohavensis. The two nuclear sequences examined were: (1) part of the gene encoding cytosolic phosphoglucose isomerase (*PgiC*); and (2) a homologue of the gene *cycloidea* (*Scyc2*) that has been isolated and characterised from *Senecio*.

The *PgiC* gene was fully characterised in *Clarkia lewisii* by Thomas *et al.* (1992) and comprises 22 introns and 23 exons. It encodes a protein of 569 amino acids. The portion of the gene sequenced in this study contains exons 11 to 16. The second sequence studied here is a portion of *Scyc2*, one of six *cycloidea* (*cyc*) homologues isolated from *Senecio vulgaris* by A. C. M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott (unpubl. data). *Scyc2* is homologous to the *cyc* gene in *Antirrhinum* that controls floral zygomorphy (Luo *et al.*, 1996). *Scyc2* encodes a protein of 303 amino acids in *S. vulgaris*, and contains the highly conserved TCP- and R-domains (see Cubas, 2002). The TCP domain is thought to be involved in DNA-binding and dimerisation and hence *cyc* in *Antirrhinum* is considered to be a transcription factor (Cubas *et al.*, 1999a). In *S. vulgaris* the *Scyc2* gene is tightly linked to the *RAY* locus which controls presence/absence of ray florets in the species (A.C.M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott, unpubl. data).

If *S. mohavensis* is allotetraploid it is expected to contain two divergent copies of each of the *PgiC* and *Scyc2* genes. One copy of each gene is expected to have been inherited from *S. flavus*, while the other would have come from the second species, possibly *S. glaucus*. Alternatively, if *S. mohavensis* is the autotetraploid of *S. flavus*, the presence of ray florets in the capitulum may have resulted from the introgression of a gene that causes ray floret production. Consequently, the *Scyc2* gene, which is tightly linked to the *ray* gene (at least in *S. vulgaris*), may also have been introgressed, causing *S. mohavensis* to contain two divergent copies of this sequence. In contrast, two very similar copies of *PgiC* inherited from *S. flavus* would be expected to be recovered if *S. mohavensis* was autotetraploid.

5.2 MATERIALS AND METHODS

5.2.1 *Plant material*

The species examined in this study and the locations from which seed were collected are listed in table 5.1. Other than *Senecio flavus* and the two subspecies of *S.*

mohavensis, the species include most of the widespread diploid species in *Senecio* section *Senecio* (Alexander, 1979) from the Mediterranean (where *S. mohavensis* is thought to have originated).

Table 5.1 – Species surveyed in the analysis of DNA sequence variation

Species	Location of sample	Collector[1]
S. aethnensis	Cisternazza, Sicily	JKJ
S. chrysanthemifolius	Catania, Sicily	JKJ
S. flavus	Tafraoute, Morocco	HPC/RJA
S. gallicus	Sa Riera, Spain	RJA
S. glaucus subsp. *coronopifolius*	Houmt Souk, Tunisia	MC/RJA
S. leucanthemifolius	Gioia Tauro, Italy	HPC/RJA
S. mohavensis subsp. *breviflorus*	Khirbet Mezin, Israel	HPC/RJA
S. mohavensis subsp. *mohavensis*	Zzyzyx, California, USA	AL
S. rupestris	Mgna della Maiella, Italy	HPC/RJA
S. squalidus	Oxford, UK	RJA
S. vernalis	Western Galilee, Israel	HPC/RJA

[1] Collectors – R. J. Abbott, M. Coleman, H.-P. Comes, J. K. James, A. Liston.

Plants were raised from seed in the glasshouse at the University of St. Andrews, Scotland. DNA was extracted from leaf material according to a CTAB extraction procedure (Doyle and Doyle, 1990) with some minor modifications (see section 2.4). DNA concentration was quantified by agarose gel electrophoresis alongside standards of 25, 50 and 100ng λ DNA.

5.2.2 *Genetic analysis*

The polymerase chain reaction (PCR) conditions for amplifying nuclear sequences of *PgiC* and *Scyc2* were identical (section 2.5.2). The primers used are listed in table 2.1 and their annealing positions are shown in Fig. 2.3. Primers for the *Scyc2* partial gene sequence were designed based on sequence data obtained from

Senecio vulgaris (A.C.M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott, unpublished). In most cases the S2-START and S2-08 primers were used to amplify the sequence. Use of these primers amplifies the first ~711 bp of *Scyc2*. However, for *S. mohavensis* it was found that, after cloning, the S2-START primer only yielded one copy of the sequence, but amplification further downstream suggested two copies were present (data not shown). Therefore the S2-5' primer was designed, and amplification using this and S2-08 yielded two copies of the sequence in *S. mohavensis*, one of which was found to be modified at the priming site for primer S2-START. The S2-5' primer was tested in *S. flavus* and the product amplified was found to have the same sequence as that amplified by the S2-START primer.

Universal *PgiC* primer sequences were supplied by L. Gottlieb (University of California, Davis). These primers were designed based on *PgiC* sequences from several phylogenetically divergent plant species and have been used to successfully amplify *PgiC* in a number of taxa including five members of the Asteraceae (L. Gottlieb, pers. comm.). In the initial PCR of *Senecio* genomic DNA using these primers, two products of different size were obtained from several diploid species. Subsequent sequencing of the two products from two diploid species, *S. vernalis* and *S. gallicus*, revealed that one product was ~90% identical in its amino acid sequence to a partial *PgiC* sequence in *Arabidopsis thaliana* (GenBank accession number X69195). The second product was shorter, and potentially non-functional, containing a deletion of 57 bp, 43 bp of which comprise the 3' end of exon 13. This second product was not considered further in the analysis.

PCR products of both sequences were ligated into the pGEM-Teasy vectors (Promega) and transformed into JM109 competent *E. coli* cells according to manufacturer's protocol with some minor modifications (section 2.6). For each diploid species two or three positive clones were selected for sequencing. From each tetraploid taxon eight to 12 clones containing the correct size insert were selected and another PCR was conducted with the same primers used to generate the initial PCR product. The products of this amplification were digested with three restriction enzymes according to manufacturer's instructions (Promega) in an attempt to distinguish different divergent copies of the gene prior to sequencing. Clones exhibiting two different restriction patterns were found in both subspecies of *S.*

mohavensis. Three products of each gene which showed each of the two different restriction patterns were selected for sequencing.

Selected clones were then cultured overnight and plasmid extraction and insert sequencing were carried out using the protocol in section 2.6. Sequence data were then analysed using Chromas ver. 1.0.0.1 (Technelysium Pty Ltd., 1998-2001) and aligned using Genedoc ver. 2.6.001 (Nicholas and Nicholas, 1997).

5.2.3 *Phylogenetic analysis*

Phylogenetic analysis was first carried out on the two nuclear DNA sequences separately. This was followed by an analysis of the combined dataset. Insertions/ deletions (indels) were recorded as either present of absent and added as additional characters in the sequence analysis. Maximum parsimony (MP) trees were generated separately for each dataset and the combined dataset using the heuristic search option of PAUP* (Swofford, 2001) with 1000 replicates, and the options TBR, MULTREES and COLLAPSE(max). *Senecio flavus* was designated outgroup based on sequencing of rDNA ITS (Coleman *et al.*, 2003). The model of DNA substitution was determined using the program Modeltest (Posada and Crandall, 1998) and the resultant parameters were employed in a subsequent maximum likelihood (ML) analysis using the heuristic search option of PAUP* (Swofford, 2001) with the same options in effect as for the MP analysis.

Prior to combining the datasets of the *PgiC* and *Scyc2* sequences for analysis, an Incongruence Length Difference (ILD) test was conducted using the Partition Homogeneity function (1000 replicates) of PAUP* to ensure that these sequences from each individual were congruent in phylogenetic signal. It was established that the two sequences were incongruent in phylogenetic signal (ILD: P=0.003). However, when the introns of *PgiC* were excluded, it was found that the *Scyc2* sequences and the exon sequences of *PgiC* examined were congruent in phylogenetic signal (ILD: P=0.080). Thus, these two datasets were combined for MP and ML analysis.

5.3 RESULTS

5.3.1 *Comparisons of* Senecio *sequences with those published for other taxa*

The nucleotide sequences obtained for *PgiC* and *Scyc2* from *Senecio flavus* and *S. squalidus* were used to search the GenBank database using the BLAST search tool (Altschul *et al.*, 1997). In addition, comparisons were made with the protein sequences of *PgiC* in *Arabidopsis thaliana* and *Clarkia lewisii* (Thomas *et al.*, 1993) and members of the Asteraceae (L.D. Gottlieb, unpubl. data).

PgiC *exons 11 to 16*

The BLAST search was carried out using the nucleotide sequence of *PgiC* obtained for *S. flavus* and *S. squalidus.* Greatest homology was found with *PgiC* in *Clarkia* (Onagraceae). Nucleotide sequence similarity was low in the introns and alignments were difficult to compile, hence comparisons were made with the *Senecio* amino acid sequence of exons 11-16 and the equivalent sequences obtained from the GenBank database. The two *Senecio* amino acid sequences were 90.5-92.6% identical to *PgiC* in *C. dudleyana* and *C. unguiculata* (Genbank accession numbers AJ437278 and AJ312370). With regard to the unpublished amino acid sequence data for *PgiC* in the Asteraceae, the two *Senecio* sequences were most similar to *Munzothamnus blairii* (Asteraceae; subfamily Cichoriodeae) with 94.44 - 95.56% identity (Fig. 5.3).

The length of the five introns in *Senecio* were shorter than previously reported for *Clarkia lewisii* or *Arabidopsis thaliana* (Thomas *et al.*, 1993), and shorter than those amplified in other members of the Asteraceae (Fig. 5.3). Total length of the five introns in *Senecio* ranged from 462-469 nucelotides (nts), whereas the total length of these introns in *C. lewisii* was 833-1109 nts and in *A. thaliana* 760 nts (Thomas *et al.*, 1993; Fig. 5.4). The greatest total length of these five introns that has so far been recorded in a particular species occurs in *Stephanomeria cichoriacea* (up to 1718 nts; L.D. Gottlieb, unpubl. data). The relative length of each of the introns varied between taxa. In *Senecio* all five introns were between 84 and 122 nts long, whereas in a single individual of *M. blairii* intron length varied from 88 nts (intron 15) to 892 nts (intron 13) (L.D. Gottlieb, unpubl. data; Fig 5.4)

```
      Exon 11          Exon 12                    Exon 13
Ssqu  FWDWVGGRYSG 88   CSAVGVLPLSLQYGFPVIEK 110   FLKGARSIDQHFHTAPFEKNIP 92
Sfla  FWDWVGGRYSG 88   CSAVGVLPLSLQYGFSVIEK 122   FLKGARSIDQHFHSAPFEKNIP 90
Atha  FWDWVGGRYSV 76   CSAVGVLPLSLQYGFSMVEK 280   FLKGASSIDQHFQSTPFEKNIP 197
Clew  FWDWVGGRYSV 129  CSAVGVLPLSLQYGFAVVEK 392   FLQGAHSIDQHFSSAPFEKNIP 135
Pspi  ?????GGRYSV 92   CSAVGVLPLSLQYGFAVVEK 201   FLQGAHSIDQHFSSASFEKNIP 123
Mbla  ?????GGRYSV 98   CSAVGVLPLSLQYGFPVVEK 96    FLKGARSIDQHFLSAPFEKNIP 892
Rcal  ?????GGRYSV 100  CSAVGVLPLSLQYGFPVVEK 96    FLKGARSIDQNFLSAPFEKNIP 615
Seco  ?????GGRYSV 433  CSAVGVLPLSLQYGFPVVEK 93    FLKGARSIDQQFLSAPFEKNIP 627
Scic  FWDWVGGRYSV 482  CSAVGVLPLSLQYGFPVVEK 96    FLKGARSIDQQFLSAPFEKNIP 632

      Exon 14                  Exon 15                  Exon 16
Ssqu  VLLGLLSVWNVSLLGYPAR 98   AILPYTQALEKLAPHIQQ 81   VSMES
Sfla  VLLGLLSVWNVSFLGYPAR 84   AILPYTQALEKLAPHIQQ 78   VSMES
Atha  VLLGLLSVWNVSFLGYPAR 89   AILPYSQALEKFAPHIQQ 118  VSMES
Clew  VLLGLLSVWNVSFLGYPAR 81   AILPYSQALEKLAPHIQQ 96   VSMES
Pspi  VLLGLLSVWNVSFLGYPAR 86   AILPYSQALEKLAPHIQQ 90   VS???
Mbla  VLLGLLSVWNVSFLGYPAR 443  AILPYTQALEKLAPHIQQ 88   VSMES
Rcal  VLLGLLSVWNVSFLEYPAR 385  AILPYTQALEKLAPHIQQ 92   VSMES
Seco  VLLGLLSVWNVSFLEYPAR 385  AILPYTQALEKLAPHIQQ 57   VSMES
Scic  VLLGLLSVWNVSFLEYPAR 396  AILPYTQALEKLAPHIQQ 92   VSMES
```

Figure 5.3 – Amino acid alignment for the exons of *PgiC* in *Arabidopsis thaliana* (Atha; Brassicaceae), *Clarkia lewisii* (Clew; Onagraceae) (Thomas et al 1993) and some members of the Asteraceae: *Munzothamnus blairii* (Mbla), *Pleiacanthus spinosus* (Pspi), *Rafinesquia californica* (Rcal), *Stephanomeria cichoriacea* (Scic), *Stephanomeria exigua* subsp. *coronaria* (Seco), *Senecio flavus* (Sfla) and *Senecio squalidus* (Ssqu). The 3' end of exon 11 and the 5' end of exon 16 are shown. Regions of complete homology between species are shaded. Intron length (bp) is indicated between exons.

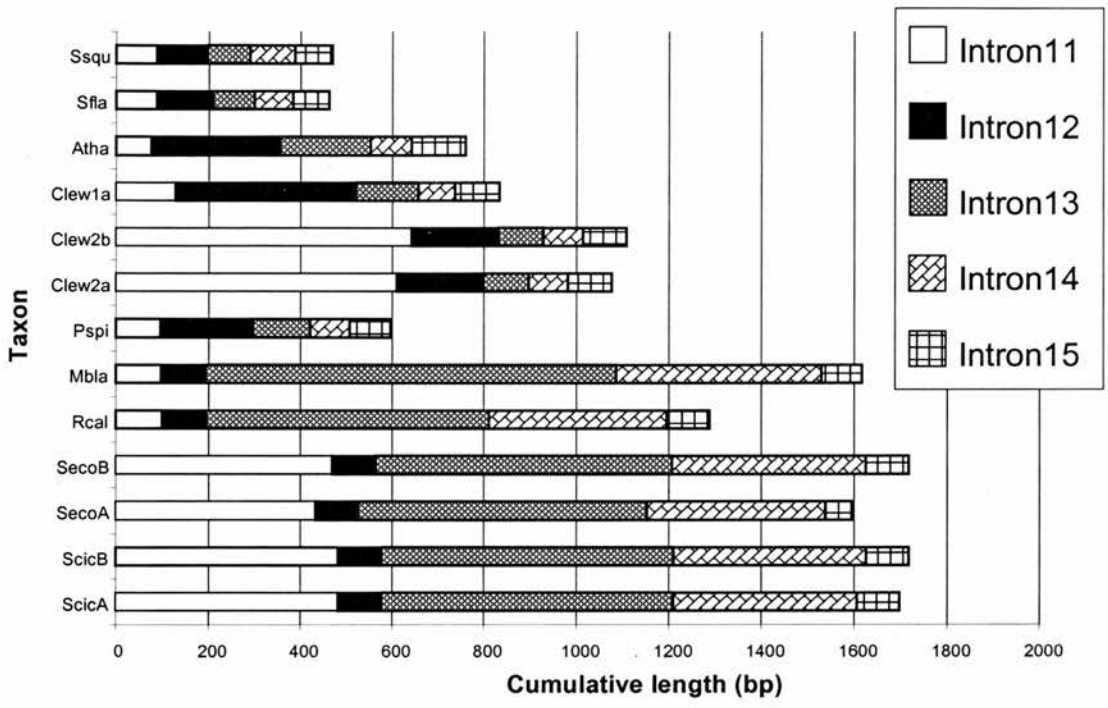


Figure 5.4 – Variation in length of introns 11 to 15 of *PgiC* in various taxa. Taxon abbreviations follow Fig. 5.3. A and B following taxon abbreviations represent different individuals. Three related copies were sequenced in *C. lewisii* (designated 1a, 2a and 2b; see Thomas *et al.*, 1993).

Scyc2

Significant similarity was recorded between both *Senecio* sequences of *Scyc2* and *cyc* homologues of a range of plant taxa. Greatest homology was with *cyc* homologues in the Fabaceae and Plantaginaceae, including *Acosmium subelegans* (GenBank accession AY225829), *Lupinus angustifolius* (AY225839), *Linaria vulgaris* (AF161252), *Digitalis purpurea* (AF512065) and *Antirrhinum leptaleum* (AF512078). Protein homology was particularly high for the conserved TCP- and R-domains (Fig. 5.5).

A – TCP domain

```
                    1         1         1         1         1
          9         0         1         2         3         4
          0         0         0         0         0         0
Sfla HTSKKDHHSKIHTAQGPRDRRVRLSIEVAKKFFYLQDLLGFDKASKTLDWLFNKSKIPINELVQTKKQ
Ssqu HSPKKDHHSKIHTAQGTRDRRVRLSIEVAKKFFYLQDLLGFDKASKTLDWLFDKSKIPINELVQTKKQ
Alep HTPKKDRHSKIHTAQGPRDRRVRLSIGIARKFFDLQEMLGFDKPSKTLDWLLTKSKEAIKELVQSKSA
Lvul STVKKDRHSKIYTAQGPRDRRVRLSIGIARKFFDLQEMLGFDKPSKTLDWLLTKSKTAIKELVQSKST
Dpur VTKKDRRHSKIYTAQGPRDRRVRLSIGIARKFFDLQELLGFDKPSKTLDWLLSKSKAAIKELVQSKTT
```

B – R domain

```
      2         2         2
      0         1         2
      0         0         0
Sfla VNQSRADARARARERTKEKLNF
Ssqu VNQSRADARARARERTKEKLNS
Alep AKESRAKARARARERTKEKMCI
Lvul VKESRAKARARARERTKEKMCI
Dpur AKESRAKARARARERTKEKMCI
```

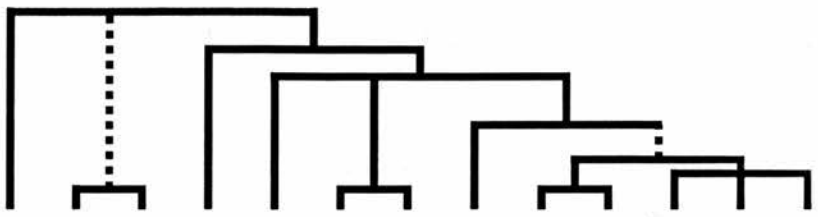
Figure 5.5 – Protein alignment for two conserved regions of the *cycloidea* locus (A) the TCP domain, (B) the R domain. Species are Sfla – *Senecio flavus*, Ssqu – *Senecio squalidus* (this study), Alep – *Antirrhinum leptaleum* (GenBank AAP46519), Lvul – *Linaria vulgaris* (AF161252), Dpur – *Digitalis purpurea* (AAP03352). The three GenBank accessions were chosen due to high sequence similarity. Amino acid positions are taken from the *S. flavus* sequence. Shaded regions show complete homology between species.

5.3.2 *Sequence alignment and phylogenetic analysis*

The sequence alignment for *Scyc2* (appendix IIIa) contained 744 nucleotides and five insertions/deletions (indels). Of the 749 characters, 664 (88.65%) were constant and 55 (7.34%) were variable and parsimony informative. Four of the parsimony informative characters were indels. The *PgiC* alignment (appendix IIIb) contained 772 nucleotides and 14 indels. All of the indels were found in the introns. Of the 786 total characters, 660 (83.96%) were constant and 78 (9.92%) parsimony informative. Of the 78 informative characters, eight were indels, 13 were nucleotides in the exons and the remaining 57 were nucleotides in the introns.

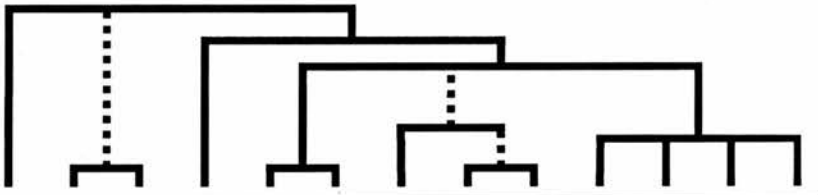
Sequence difference (measured as uncorrected *p* distance) for the *Scyc2* data (Fig. 5.6A) ranged from 0.14% (*S. squalidus* - *S. chrysanthemifolius/S. rupestris*) to 7.09% (*S. gallicus* – *S. mohavensis*B). For *PgiC* (Fig. 5.6B), least sequence divergence was shown between *S. squalidus, S. aethnensis* and *S. rupestris* (0.52%) and greatest between *S. mohavensis*A and *S. mohavensis*B (9.78%). Two distinct copies of both DNA sequences were isolated from both subsp. of tetraploid *S. mohavensis*. Mean sequence divergence between 'A' and 'B' copies was 5.83% for *Scyc2* and 9.58% for *PgiC*.

A



	1	2	3	4	5	6	7	8	9	10	11	12	13
(1) flavus													
(2) breviflorusB	0.29												
(3) mohavensisB	1.30	0.57											
(4) vernalis	5.35	5.10	6.07										
(5) glaucus	5.35	5.10	6.07	1.12									
(6) mohavensisA	5.78	5.48	6.04	1.43	1.14								
(7) breviflorusA	5.93	5.62	6.19	1.58	1.29	0.69							
(8) leucanthemifolius	5.61	5.37	6.35	1.85	1.71	2.05	2.20						
(9) gallicus	6.36	6.13	7.09	1.96	1.82	2.17	2.31	2.13					
(10) aethnensis	6.61	6.31	6.91	2.28	2.13	2.46	2.60	2.45	0.56				
(11) chrysanthemifolius	6.16	5.86	6.46	2.13	1.99	2.30	2.45	2.18	2.13	2.41			
(12) rupestris	6.12	5.83	6.42	2.12	1.97	2.30	2.44	1.87	2.12	2.42	0.28		
(13) squalidus	5.98	5.69	6.29	1.98	1.84	2.16	2.30	2.02	1.98	2.27	0.14	0.14	

B



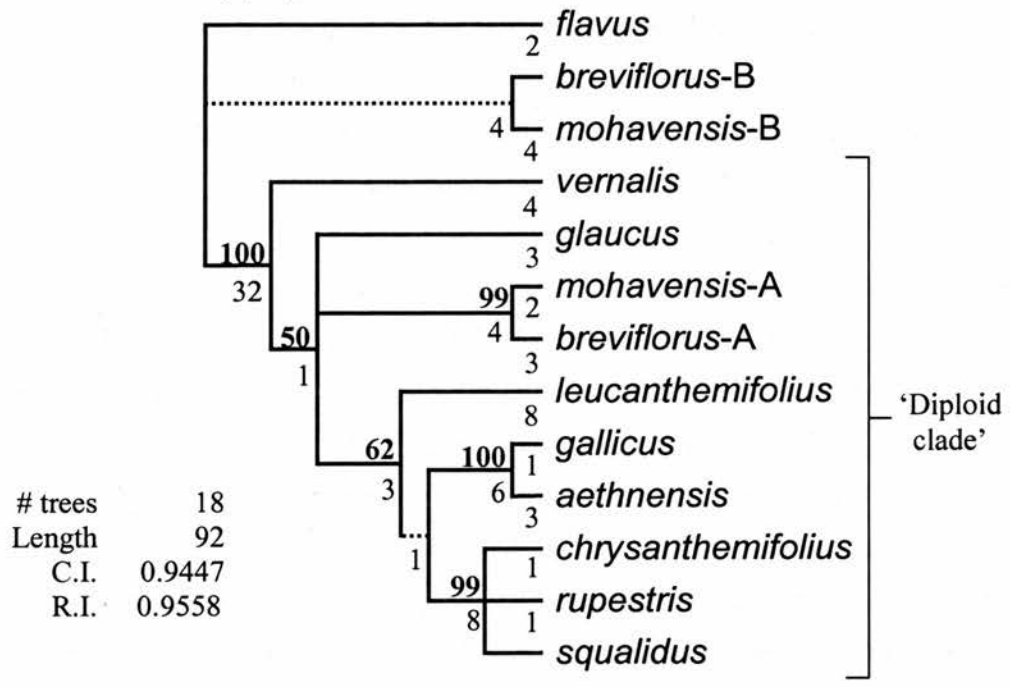
	1	2	3	4	5	6	7	8	9	10	11	12	13
(1) flavus													
(2) breviflorusB	0.78												
(3) mohavensisB	1.44	1.44											
(4) vernalis	9.28	9.42	9.68										
(5) mohavensisA	9.38	9.51	9.78	4.66									
(6) breviflorusA	9.25	9.38	9.38	4.27	0.64								
(7) leucanthemifolius	8.96	9.22	9.76	4.85	3.91	3.78							
(8) gallicus	8.91	9.04	9.57	4.17	2.98	2.85	2.76						
(9) glaucus	8.72	8.85	9.38	3.76	2.58	2.45	2.48	1.56					
(10) chrysanthemifolius	9.01	9.14	9.40	3.77	3.23	2.84	3.14	2.21	2.07				
(11) aethnensis	8.87	9.01	9.54	4.16	3.23	3.10	3.01	2.34	1.94	0.65			
(12) rupestris	8.74	8.87	9.40	4.16	3.23	3.10	3.01	2.34	2.07	0.65	0.52		
(13) squalidus	8.89	9.02	9.55	4.16	3.23	3.10	3.01	2.34	2.07	0.65	0.52	0.52	

Figure 5.6 – Matrix of pairwise sequence differences (uncorrected *p* distances) calculated between individuals. (A) *Scyc2* (B) *PgiC*. The selected MP tree (see Fig. 5.7) is shown above each matrix.

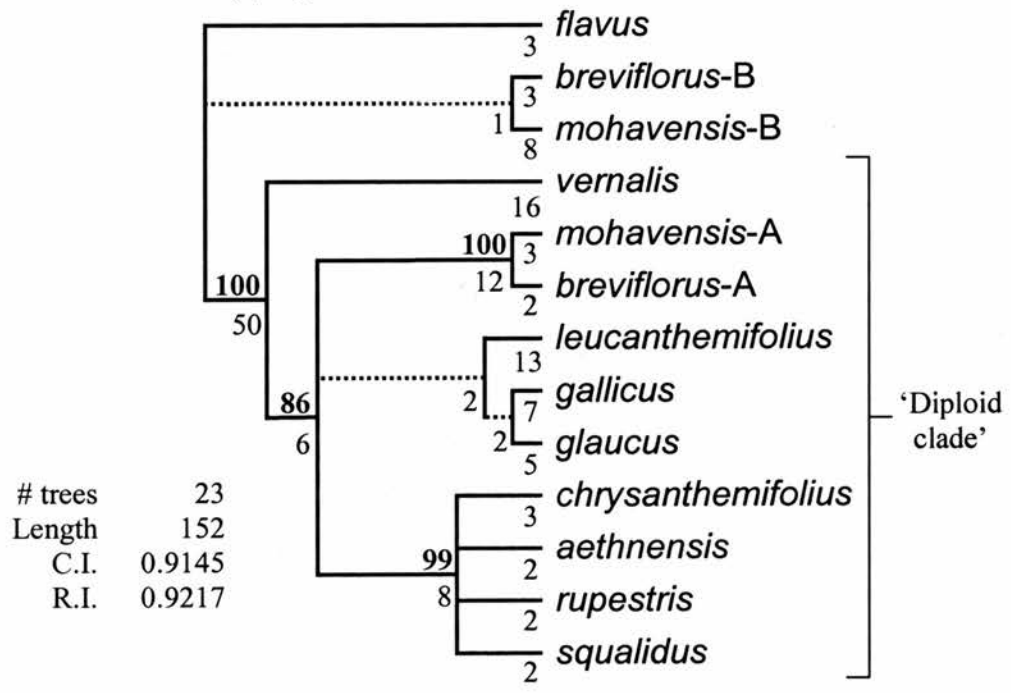
The phylogenetic trees shown in Fig. 5.7 represent the most parsimonious tree for each data set which agreed with the topology of the single ML tree. The two copies obtained from each subsp. of *S. mohavensis* are designated A and B, respectively. For each sequence (*PgiC* and *Scyc2*) two major clades were resolved in MP trees. All of the diploid species examined occurred within one clade (the 'diploid clade') with the exception of *S. flavus*, which possesses a highly divergent sequence for *PgiC* and *Scyc2* relative to the rest of the diploid species. There were differences between the two trees in the topology of the diploid clade, primarily due to the relative positioning of *S. glaucus* and *S. gallicus*. *S. vernalis* was resolved as the most basal member of the diploid clade in both trees. For each sequence in *S. mohavensis*, one copy is resolved within the diploid clade (copy 'A') and the other clusters with *S. flavus* (the 'B' copy). The diploid hybrid species *S. squalidus* was resolved with high bootstrap support in a sub-clade with one of its putative parents, *S. chrysanthemifolius*, and the closely related *S. rupestris* in the *Scyc2* tree and also with its other putative parent, *S. aethnensis*, in the *PgiC* tree.

Figure 5.7 (following page) – MP trees generated for the *Scyc2* (A) and *PgiC* (B) sequence data. The two copies present in *S. mohavensis* subsp. *mohavensis* and subsp. *breviflorus* are designated A and B. Trees shown agreed with the topology of the ML trees (–lnL = 1591.18 (*Scyc2*) and –lnL = 1827.74 (*PgiC*)). Branches which collapse under strict concensus are indicated by dotted lines. Bootstrap percentages (≥50%; 1000 replicates) are indicated above branches, number of nucleotide changes below branches. Below each tree is given the number of MP trees generated (# trees), tree length (length), consistency index (C.I.) and retention index (R.I.).

(A) *Scyc2*



(B) *PgiC*



The sequence alignment for the combined analysis of *Scyc2* and the exons of *PgiC* contained 1034 characters. Five of these were indels which were found in the *Scyc2* sequence, whereas all of the *PgiC* indels were located in the introns and hence were omitted from the combined analysis. Of the 1034 characters, 925 (89.46%) were constant and 68 (6.61%) variable and parsimony informative. Fifty-five of these informative sites were located in *Scyc2* and the remaining 13 located in *PgiC*.

An MP tree obtained from the analysis of *Scyc2* and *PgiC* sequences combined which matched the topology of the ML tree is shown in Fig. 5.8. The least divergence between these two sequences combined was between *S. rupestris* and *S. squalidus* (0.20%), while greatest divergence occurred between *S. gallicus* and *S. mohavensis*B (6.35%). Concordant with the analysis of single sequences, two clades are resolved in the MP and ML trees produced. As was the case in the single sequence trees, one clade represents all of the diploid species investigated, with the exception of *S. flavus* which is highly diverged. As with the single sequence analysis, the two copies present in *S. mohavensis* resolve in the different clades (Fig. 5.8). The copy of combined sequences which clusters in the diploid clade (copy 'A') is most similar to the sequence obtained for *S. glaucus* (1.12-1.32% sequence divergence), and is supported by high bootstrap support (72%).

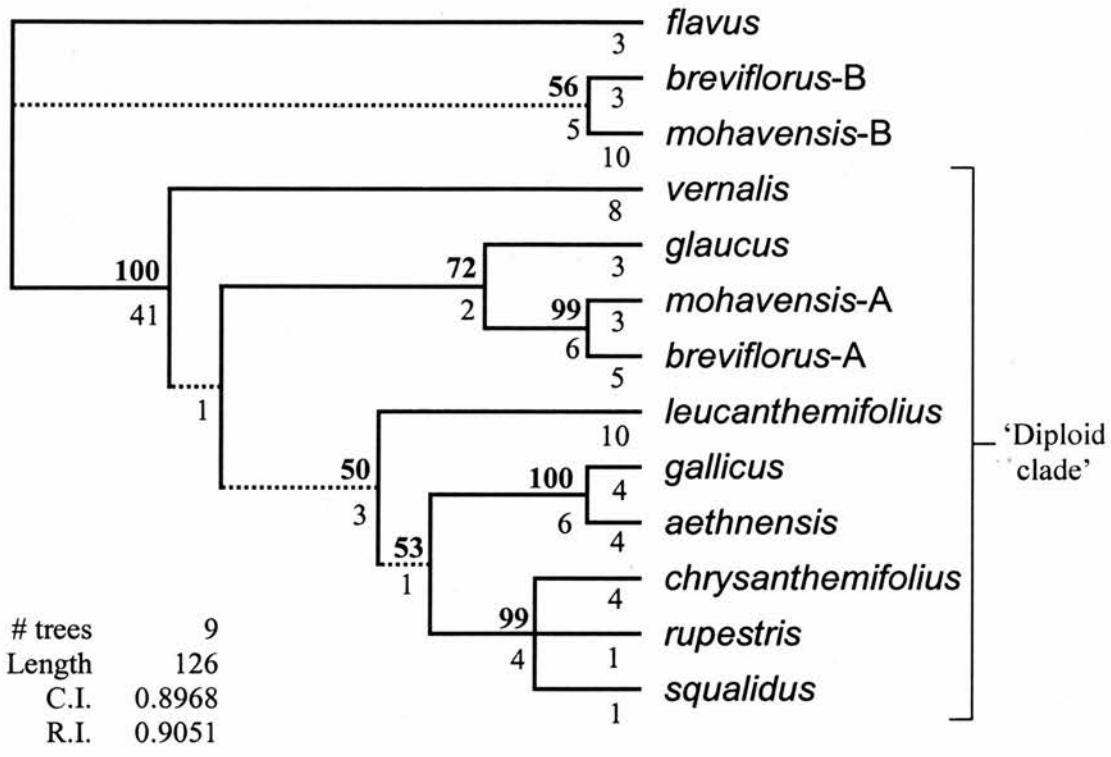


Figure 5.8 – MP tree generated for the *Scyc2/PgiC* combined sequence data (*PgiC* introns excluded – see text). The two copies present in *S. mohavensis* subsp. *mohavensis* and subsp. *breviflorus* are designated A and B. Tree shown agreed with the topology of the ML tree (-lnL = 2211.53). Branches which collapse under strict concensus are indicated by dotted lines. Bootstrap percentages (≥50%; 1000 replicates) are indicated above branches, number of nucleotide changes below branches. Below the tree is given the number of MP trees generated (# trees), tree length (length), consistency index (C.I.) and retention index (R.I.).

5.4 DISCUSSION

Polyploidy is a prominent mechanism of speciation in the plant kingdom (Stebbins, 1950; Grant, 1981; Soltis and Soltis, 1993), and polyploids have been well studied in terms of their origin (Ramsey and Schemske, 1998; Soltis and Soltis, 1999), and their biology relative to parent taxa (Thompson and Lumaret, 1992; Segraves and Thompson, 1999; Nuismer and Thompson, 2001). However, the genomic changes that may accompany or follow the process of polyploidisation are only now being subjected to detailed analysis (reviewed in Soltis and Soltis, 1999, 2000; Comai, 2000; Wendel, 2000; Liu and Wendel, 2003; Osborn *et al.*, 2003). The manner in which a polyploid species originates, either through auto- or allopolyploidy, will have an impact on the type of genomic changes likely to take place, for example with regard to the generation and maintenance of genetic variation in the species, and levels of gene silencing and subfunctionalisation.

In this study the origin of the tetraploid *Senecio mohavensis* was investigated by analysing nucleotide variation of two nuclear DNA sequences. *S. flavus* had previously been considered a parent of *S. mohavensis* due to close morphological similarity (Kadereit, 1984a). This similarity may suggest an autotetraploid origin followed by minor morphological divergence, and this is supported by the results of a study of RAPDs in which these two species clustered together in a dendrogram, apart from other diploids studied (Comes and Abbott, 2001). *S. mohavensis*, however, has been shown to possess a cpDNA type and nuclear ITS DNA sequence very different from *S. flavus* and similar to those possessed by *S. glaucus* (Comes and Abbott, 2001; Coleman *et al.*, 2003). Hence two possibilities exist for the origin of *S. mohavensis*, (1) autotetraploidisation of *S. flavus* followed by introgression of cpDNA and some nuclear genes from *S. glaucus* or a similar diploid species, or (2) an allotetraploid origin following a cross between *S. flavus* and *S. glaucus* and subsequent evolution towards a *S. flavus*-like morphology.

The approach employed here, to analyse DNA sequence variation of a portion of the nuclear genes *PgiC* and *Scyc2*, was adopted as a means to distinguish between the two possible origins hypothesised for *S. mohavensis*. In a tetraploid species, duplicate copies of each nuclear locus are expected to be present. In an autotetraploid species these will be identical, or very similar in sequence depending on whether the

species is of recent or ancient origin. In contrast, in an allotetraploid the two copies of a gene are likely to show a much greater difference in sequence as each copy is derived from a different parent. Previous unpublished work (A.C.M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott, unpubl. data) has shown that the *Scyc2* sequence analysed here is tightly linked to the gene that controls presence/absence of ray florets in *S. vulgaris*. Hence, if it were shown from an analysis of variation that radiate *S. mohavensis* is most likely the autotetraploid of non-radiate *S. flavus*, then sequence variation for *Scyc2* in *S. mohavensis* would indicate which other diploid species was involved as donor of introgressed genetic material that produces radiate flower heads. If, however, two very different copies of both the *PgiC* sequence and *Scyc2* sequence were found in *S. mohavensis* this would suggest an allotetraploid origin and indicate the two most likely parents of the species.

The results of this study are concordant with an allopolyploid origin of *S. mohavensis*, following hybridisation between the diploid species *S. flavus* and *S. glaucus*. The *PgiC* and *Scyc2* sequences were amplified in several diploid species of Mediterranean *Senecio* section *Senecio* and in both subspecies of *S. mohavensis*. Two or three clones per diploid species were sequenced, as were three clones of each homoeologue from the tetraploid. Maximum parsimony and maximum likelihood phylogenetic analysis of the sequences obtained for the diploid species showed that most are relatively closely related; with the exception of *S. flavus* which is much more distantly related. *S. vernalis* resolved as the most basal member of the 'diploid clade' as reported by Comes and Abbott, (2001) and Coleman *et al.* (2003). In addition, the diploid hybrid species *S. squalidus* resolves in the same sub-clade as *S. chrysanthemifolius*, one of its putative parent species (Abbott *et al.*, 2000, 2002), in both the *Scyc2* and *PgiC* trees.

Two distinct copies of both *PgiC* and *Scyc2* were present in both subspecies of the tetraploid *S. mohavensis*. One of the copies of each sequence (copy 'B') was very similar and closely related to the respective sequence present in *S. flavus*, whereas the other copy (copy 'A') comprised a sequence that was positioned in the clade containing all the remaining diploid species sequences. This phylogenetic pattern for both sequences reflects the contribution of two distinct diploid parent species in the origin of *S. mohavensis*. One parent is *S. flavus*, and in the combined analysis high

bootstrap support (72%) indicated that the second parent is likely to have been *S. glaucus*, as was postulated by Comes and Abbott (2001) and Coleman *et al.* (2003).

The findings presented here do not entirely exclude the possibility of an autotetraploid origin of *S. mohavensis* followed by introgression. Introgression appears to be common phenomenon in plants (Rieseberg and Wendel, 1993), especially with regard to 'chloroplast capture', i.e. the introgression of the cpDNA of one species into another with little or no accompanying introgression of nuclear material (e.g. Soltis *et al.*, 1991; Soltis and Kuzoff, 1995; VanRaamsdonk *et al.*, 1997; Jackson *et al.*, 1999). Nuclear introgression based on ITS sequence variation has also been documented in a variety of plant species (e.g. Jobst *et al.*, 1998; Setoguchi and Watanabe, 2000; Pardo *et al.*, 2004).

Introgression involves a hybrid repeatedly backcrossing to one parent, such that the majority of the recipient species' genome is that of the backcross parent, with some of the second parents' genome incorporated. A scenario involving the possession of three introgressed *S. glaucus* nuclear genes (*PgiC*, *Scyc2* and rDNA ITS repeats) in *S. mohavensis* is in theory possible, but would suggest that their combined presence is caused either by the fact that each is favoured by selection in *S. mohavensis*, or that they are tightly linked as a single unit. The linkage hypothesis seems unlikely as, even if the genes were physically linked in *S. glaucus*, polyploid genome evolution often involves large-scale chromosomal rearrangements potentially disrupting such linkage blocks (Levin, 2002). Further analysis of DNA sequence variation at additional loci is required to determine how often two distinct copies of a gene inherited from *S. flavus* and *S. glaucus*, respectively, are found in *S. mohavensis*. If this is found to be the case for most genes examined then this would be very strong evidence for an allotetraploid origin.

Although little is known about the possibility of hybrid formation in the wild, the two species, *S. flavus* and *S. glaucus*, exhibit strong barriers to hybridisation in the greenhouse, and so far no hybrids have been produced between these species despite several attempts (M. Coleman and D.G. Forbes, University of St Andrews). Indeed, *S. flavus* could not be crossed with any member of *Senecio* section *Senecio* studied here. In contrast, *S. mohavensis* can be crossed to *S. flavus* and *S. glaucus* resulting in sterile triploid offspring. Strong reproductive isolation is not uncommon between putative parents of allopolyploid species. For example, two polyploid species of

Senecio, *S. cambrensis* (Ashton and Abbott, 1992a) and *S. eboracensis* (Lowe and Abbott, 2000) have recently evolved despite a strong reproductive barrier between the parents *S. squalidus* and *S. vulgaris* (Abbott and Lowe, 2004).

The fact that *S. mohavensis* is morphologically very similar to *S. flavus* despite its likely allopolyploid origin is somewhat surprising. However, interspecific hybrids are not always morphologically intermediate to their parents, but instead may exhibit a mosaic of parental traits (Rieseberg and Carney, 1998). Moreover, there are other examples of allopolyploids resembling one parent more so than the other parent (Watson *et al.*, 1991; Bleeker *et al.*, 1999). The morphological similarity between *S. mohavensis* and *S. flavus* might reflect convergent adaptation to very similar environments in which they grow. However, the geographic distributions of the species are not concordant with this hypothesis. *S. glaucus* and *S. flavus* are morphologically very different, yet their distributions are largely overlapping (Alexander, 1979), and the two species have been reported to co-occur at Tafraoute, Morocco, and at Dhahab, Sinai (R.J. Abbott, pers. obs.). In contrast, the distributions of the morphologically very similar *S. flavus* and *S. mohavensis* subsp. *breviflorus* are virtually non-overlapping (Fig. 5.2). Considering the potential for bird dispersal of these taxa (Coleman *et al.*, 2003), one might expect that unless ecological differentiation is present, the species would co-occur. It seems unlikely therefore that *S. mohavensis* exhibits morphological similarity to *S. flavus* due to ecological reasons.

An alternative cause of the similar morphologies of *S. flavus* and *S. mohavensis* may be due to unequal expression in *S. mohavensis* of the genes originating from the two parents, with a bias towards expression of genes inherited from *S. flavus*. Recently, Adams *et al.* (2003) investigated the relative expression of several pairs of homoeologous genes in allopolyploid cotton. It was shown that expression levels were very variable, with one homoeologue silenced in some organs, but fully functional in others (Adams *et al.*, 2003). In contrast, in the parents the genes were expressed in the entire suite of organs investigated (Adams *et al.*, 2003). Changes in gene expression in polyploid wheat have been well studied and probably involve DNA elimination and alteration in methylation patterns as early as the F1 hybrid allopolyploid generation (Galili and Feldman, 1984; Ozkan *et al.*, 2001; Shaked *et al.*, 2001). Genomic changes associated with polyploidisation (i.e. silencing or divergence of function of homoeologues) may have had a similar effect in *S.*

mohavensis, with a bias towards the expression of *S. flavus* genes and possibly the silencing (or elimination) of many *S. glaucus* homoeologues. Nothing is currently known about the relative expression of any genes from the two distinct parental genomes in the tetraploid, but should expression of the homoeologues from *S. flavus* be greater overall than expression of homoeologues from *S. glaucus*, then it is likely that a *S. flavus*-like morphology would be exhibited. If this were the case, then it might be hypothesised that natural selection favoured a particular differential pattern of homoeologue expression in the region where *S. mohavensis* is found. Clearly this argument deserves further investigation and *S. mohavensis* would seem an interesting allotetraploid in which to examine gene silencing and elimination, and the possible consequences.

The maintenance of radiate capitula in *S. mohavensis* (Fig. 5.1 A, B) despite its overall morphology converging on that of *S. flavus* is of further interest. British *S. vulgaris* is polymorphic for presence/absence of ray florets (Fig. 5.1 C, D), and the radiate form (var. *hibernicus*) has an increased outcrossing rate relative to non-radiate var. *vulgaris* (Marshall and Abbott, 1982, 1984). It is possible that the presence of radiate capitula in *S. mohavensis* also increases this species' outcrossing rate relative to *S. flavus*. However, some individuals of *S. mohavensis* subsp. *mohavensis* from North America are non-radiate, or, if the rays are present, they are poorly formed and shorter than in subsp. *breviflorus*. It would be worth conducting a detailed analysis of the genetic control of presence/absence of ray florets in capitula of *S. mohavensis* and compare the control mechanism to what is known for *S. vulgaris*. It is possible that the same genes are involved in the development of ray florets in these species.

In conclusion, evidence has been obtained which strongly indicates that *Senecio mohavensis* is an allotetraploid of *S. flavus* and *S. glaucus*. It may exhibit a morphology most similar to *S. flavus* due to changes in relative gene expression of the homoeologous loci from the parents. However, further analysis of expression patterns of duplicated copies of genes in *S. mohavensis* is required to test this hypothesis.

CHAPTER 6 - ORIGINS OF, AND RELATIONSHIPS BETWEEN, SEVERAL *SENECIO* HYBRID TAXA BASED ON DNA SEQUENCE VARIATION IN TWO NUCLEAR GENES

ABSTRACT

Hybridisation and polyploidy are evident in the evolutionary history of many plant species complexes, and have played important roles in the evolution of British and Mediterranean *Senecio* sect. *Senecio* (Asteraceae). Here DNA sequence variation within two nuclear genes (*Scyc2* and *PgiC*) is used to determine the parentage and relationships between the British neohexaploid *S. cambrensis* and the Canary Islands' endemic hexaploid *S. teneriffae*. The close relationship of two homoeologues of each gene from each of the hexaploids to two homoeologues of each sequence present in *S. vulgaris* confirms the hypothesis that tetraploid *S. vulgaris* was one parent of *S. cambrensis* and *S. teneriffae*. However, sequence analysis suggests that the diploid parent differs between the hexaploid species. *S. squalidus* is confirmed as the diploid parent of *S. cambrensis*, whereas the diploid parent of *S. teneriffae* is likely to be *S. leucanthemifolius* based on *Scyc2* variation, although this is not well-supported by the analysis of *PgiC* variation. The results also support the hypothesis that *S. vulgaris* is an allotetraploid, containing two divergent homoeologues of both *PgiC* and *Scyc2*. *S. vernalis* and *S. leucanthemifolius* are inferred as possible parents of *S. vulgaris*, although the parental role of each of these species is only supported by one of the two sequences examined. In addition, the radiate morph of *S. vulgaris* (var. *hibernicus*) contains a *Scyc2* DNA sequence identical to that found in *S. squalidus*, supporting the hypothesis of an introgressant origin of *S. vulgaris* var. *hibernicus*. Surprisingly, this particular sequence was not found in the parents of *S. squalidus* (i.e. *S. aethnensis* and *S. chrysanthemifolius*) and its possible origin in *S. squalidus* is discussed.

6.1 INTRODUCTION

Speciation via hybridisation can occur at the diploid or polyploid level. Polyploid species are characterised by an increased chromosome number relative to the parent species whereas diploid (homoploid) hybrid species exhibit the same chromosome number as the parents. Early opinion (e.g. Wagner, 1983) regarded polyploid species as transient and evolutionarily insignificant due to predictions that (1) polyploid species arise once, hence have low genetic variation, and (2) a polyploid will not find a genetically compatible mate, and crosses with diploids will produce sterile triploid offspring (minority cytotype disadvantage; Fowler and Levin, 1984). However, these predictions have been shown to be ill-founded. Firstly, most polyploids analysed have originated more than once (e.g. Ashton and Abbott, 1992a; Cook *et al.*, 1998; Doyle *et al.*, 1999; Segraves *et al.*, 1999; reviewed in Soltis and Soltis, 1993, 1999), and secondly, triploid individuals are not completely sterile, and produce, at low frequency, diploid and triploid gametes (Ramsey and Schemske, 1998). In fact, it is thought that most polyploids arise via a 'triploid bridge' (Ramsey and Schemske, 1998).

It is known that polyploidy has played a major role in the evolution of plants. Up to 70% of angiosperms and 95% of ferns are thought to be polyploid in origin (Grant, 1981; Whitham *et al.*, 1991; Masterson, 1994) and speciation by polyploidy probably accounts for 2-4% of speciation events (Otto and Whitton, 2000). Polyploid species tend to exhibit greater heterozygosity than their diploid progenitors and this may be a reason for their success (Roose and Gottlieb, 1976; Soltis and Soltis, 1993). In addition, polyploid species are often observed to have altered ecological tolerances relative to their parent species that assists in their ability to become established (Levin, 1983; Thompson and Lumaret, 1992).

Genomic changes following polyploid formation can be extensive, involving inter-genome recombination, gene silencing and subfunctionalisation (Soltis and Soltis, 2000; Wendel, 2000; Adams *et al.*, 2003; Osborn *et al.*, 2003). Experiments in which a polyploid is artificially resynthesised have revealed that some alterations (e.g. elimination of some loci or changes in methylation patterns) can be observed immediately following the polyploidisation event (Song *et al.*, 1995; Ozkan *et al.*,

2001; Shaked *et al.*, 2001; Madlung *et al.*, 2002), although this is not always the case (e.g. Liu *et al.*, 2001; Baumel *et al.*, 2002).

Hybridisation and polyploidy have played significant roles in the evolution of Mediterranean and European members of *Senecio* sect. *Senecio*. The most recent revision of Mediterranean *Senecio* sect. *Senecio* recognises ~24 diploid and polyploid species ($2n$=20, 40, 60; Alexander, 1979). Species limits within the diploid species are poorly defined due to phenotypic plasticity, interfertility and close genetic relationships due to a recent radiation (Comes and Abbott, 2001) and hybridisation has given rise to both homoploid and polyploid hybrid species in the group (Abbott, 1992; Ashton and Abbott, 1992a; Harris and Ingram, 1992a; Lowe and Abbott, 1996, 2000, 2003; Abbott *et al.*, 2000, 2002; Coleman *et al.*, 2003; Fig. 6.1).

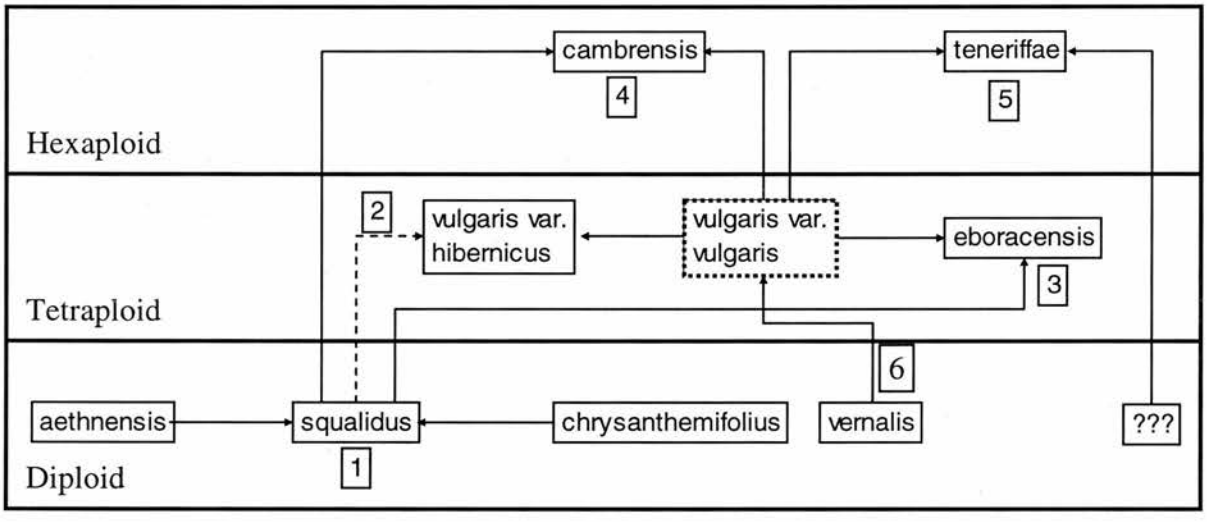


Figure 6.1 - Relationships between hybrid *Senecio* taxa and their parent species. Solid lines between taxa indicate parentage, dashed lines indicate introgression. *S.* vulgaris var. *vulgaris*, surrounded by dashed lines, may be auto- or allopolyploid. ??? denotes unknown parent. Numbers refer to references: (1) James, 1999; Abbott *et al.*, 2000, 2002; (2) Abbott *et al.*, 1992; (3) Irwin and Abbott, 1992; Lowe and Abbott, 2000, 2003; (4) Ashton and Abbott, 1992a; Harris and Ingram, 1992a; (5) Lowe and Abbott, 1996; (6) Comes *et al.*, 1997; Harris and Ingram, 1992b.

In the British Isles, three *Senecio* taxa have evolved following hybridisation between introduced *S. squalidus* L. and native *S. vulgaris* L. *Senecio squalidus* (2*n*=20) is a homoploid hybrid species, derived from material collected from a hybrid zone between *S. aethnensis* DC. and *S. chrysanthemifolius* Poiret on Mt. Etna, Sicily (Crisp, 1972; Abbott *et al.*, 2000). Material from which this species is derived was introduced to the UK in the early 18th century and, following a lag phase of approximately 90 years, the species began to spread throughout much of the UK and to parts of Ireland (Abbott, 1992; Harris, 2002). *Senecio squalidus* is frequently found to co-occur on disturbed ground with native *S. vulgaris* (2*n*=40) and crosses between the species occur at low frequency in the wild (Marshall and Abbott, 1980). The resultant triploid *S.* x *baxteri* is highly sterile, but produces balanced gametes (x = 10, 20 or 30) on occasion and controlled crosses with *S. vulgaris* can produce fertile tetraploid plants (Ingram *et al.*, 1980; Lowe and Abbott, 2000). The three taxa that have formed as a result of hybridisation between *S. squalidus* and *S. vulgaris* (see Fig. 6.1) are as follows:

(1) *Senecio vulgaris* var. *hibernicus* Syme (2*n*=40) – A stabilised introgressant of typical var. *vulgaris* L. distinguished by the presence of ray florets which are absent from var. *vulgaris*. Plants morphologically similar to var. *hibernicus* can be synthesised following crosses between var. *vulgaris* and *S. squalidus* (Ingram *et al.*, 1980; Lowe and Abbott, 2000), and natural var. *hibernicus* contains an *Aat-3* allozyme that is present in *S.* squalidus but absent from var. *vulgaris* suggesting an introgressive origin (Abbott *et al.*, 1992).

(2) *Senecio eboracensis* Abbott & Lowe (2*n*=40) – A tetraploid derivative first recorded in York, north England, in 1979, and only known from there (Irwin and Abbott, 1992; Lowe and Abbott, 2003). As is the case for *S. vulgaris* var. *hibernicus*, plants resembling *S. eboracensis* can be synthesised by crossing the parental taxa (Lowe and Abbott, 2000). *Senecio eboracensis* is distinct from *S. vulgaris* var. *hibernicus* as it contains much more of the *S. squalidus* genome, thus it has been shown to exhibit a number of RAPD markers diagnostic of either *S. vulgaris* or *S. squalidus*, whereas none of the *S. squalidus* RAPD markers were found in *S. vulgaris* var. *hibernicus* (Abbott *et al.*, 2003).

(3) *Senecio cambrensis* Rosser ($2n$=60) – A hexaploid first recorded in Wales (Rosser, 1955) and more recently in Edinburgh (Abbott *et al.*, 1983). These populations arose independently as verified by allozyme and chloroplast (cp) DNA analysis (Ashton and Abbott, 1992a; Harris and Ingram, 1992a; Lowe and Abbott, 1996). *S. cambrensis* is thought to have arisen via chromosome doubling of the triploid hybrid *S. vulgaris* x *S. squalidus* (Weir and Ingram, 1980).

A survey of cpDNA variation in *S. cambrensis* revealed a 330 base pair (bp) insertion in the chloroplast genome of Welsh *S. cambrensis* that was absent from the Edinburgh population (Harris and Ingram, 1992a). This insertion was initially not found in *S. squalidus* or *S. vulgaris* (Abbott *et al.*, 1995; Harris and Ingram, 1992a, b) leading Lowe and Abbott (1996) to investigate the hypothesis that Welsh *S. cambrensis* was an introduced species, possibly derived from the closely related Canary Islands' endemic hexaploid *S. teneriffae* Schultz Bip. It was, however, confirmed that Welsh *S. cambrensis* was not a derivative of *S. teneriffae*. Firstly, a more detailed survey showed that the cpDNA insertion was present at low frequency in *S. vulgaris* var. *hibernicus*, and secondly, there were some differences in allozymes expressed in *S. cambrensis* and *S. teneriffae*. This led to the conclusion that *S. vulgaris* was the tetraploid parent of both hexaploids, but the diploid parent of *S. teneriffae* was not *S. squalidus*. Instead it was proposed that *S. glaucus*, or another related species, was the diploid parent of *S. teneriffae*, based on the expression of allozymes in *S. glaucus* that are also expressed by *S. teneriffae* but not by *S. cambrensis* or *S. squalidus* (King, 1994; Lowe and Abbott, 1996).

A number of questions remain unanswered with respect to the origin of certain hybrid taxa in *Senecio* sect. *Senecio*. This study was initially designed to investigate the origin of hexaploid *S. teneriffae*, however during the course of the investigation several additional findings have emerged with regard to the origin of diploid *S. squalidus* and tetraploid *S. vulgaris* var. *vulgaris* and var. *hibernicus*.

The use of DNA sequence variation of low-copy number nuclear genes for phylogeny reconstruction (including studies of polyploid and hybrid species) is becoming more popular, as the ability to isolate and characterise these genes increases (reviewed in Small *et al.*, 2004). In polyploid species, genes at homoeologous loci will be more similar in DNA sequence to the locus in the parent species than they are

to each other. The evolutionary dynamics of low-copy nuclear genes appear to be more suitable for analyses of hybrid or polyploid species than other genetic markers. For example, the use of cpDNA to infer relationships may be unsuitable due to the frequent instances of 'chloroplast capture' via introgression (e.g. Soltis *et al.*, 1991; Soltis and Kuzoff, 1995; VanRaamsdonk *et al.*, 1997; Jackson *et al.*, 1999) and the ability to only detect one parent of a hybrid species due to uniparental inheritance of cpDNA in most angiosperms.

The sequence of the internal transcribed spacer (ITS) of nuclear ribosomal DNA (rDNA) is frequently used in phylogenetic analyses due to a relatively fast rate of evolution and the availability of universal primers to amplify the sequence (Baldwin *et al.*, 1995; Soltis and Soltis, 1998). The ITS region is present in hundreds or thousands of copies in one or more arrays in the higher plant genome (Hillis and Dixon, 1991). However, within polyploid species, the evolutionary dynamics of the ITS region are variable. In some polyploids, the copies inherited from the different parents are maintained (Sang *et al.*, 1995; Baumel *et al.*, 2001), whereas in other polyploid species, concerted evolution may eliminate the copies inherited from one parent (e.g. Wendel *et al.*, 1995a; Franzke and Mummenhoff, 1999; reviewed in Alvarez and Wendel, 2003). Additionally, interspecific DNA sequence variation of the ITS region is very low between diploid Mediterranean *Senecio* species (Comes and Abbott, 2001) making it inappropriate for this study.

Low-copy nuclear genes are biparentally inherited and multiple independent loci can be sampled if reticulate evolution is suspected. Perhaps more importantly, concerted evolution of low-copy nuclear genes does not appear to occur (reviewed in Small *et al.*, 2004), although the possibility of this phenomenon has only been well studied in cotton (Cronn *et al.*, 1999; Cedroni *et al.*, 2003; Senchina *et al.*, 2003). For these reasons, low-copy nuclear gene sequences are becoming a popular choice of genetic marker for determining the parentage of hybrid/polyploid plant taxa (Small *et al.*, 1998; Cronn *et al.*, 1999; Doyle *et al.*, 1999, 2002; Ford and Gottlieb, 1999; Ge *et al.*, 1999; Sang and Zhang, 1999; Small and Wendel, 2000; Mason-Gamer, 2001).

In this investigation, DNA sequence variation of portions of two single-copy nuclear genes is used to ascertain the parentage of the hexaploid *S. teneriffae* and to compare this with that of the British neohexaploid *S. cambrensis*. In addition, DNA sequence variation is used to infer the parentage of, and relationships between, several

other related hybrid taxa. The two genes are (1) a portion of *Scyc2*, a *cycloidea* (*cyc*) homologue, and (2) a portion of *PgiC* which encodes the cytosolic form of phosphoglucose isomerase. *Scyc2* is one of six *cycloidea* homologues isolated from *S. vulgaris* and is tightly linked to the gene controlling presence/absence of ray florets in *S. vulgaris* (A.C.M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott, unpublished). *Cycloidea* is necessary for floral dorso-ventral asymmetry in *Antirrhinum majus* (Luo *et al.*, 1996) and *Linaria vulgaris* (Cubas *et al.*, 1999b) and is homologous to the *tb1* gene controlling apical dominance in Maize (Doebley *et al.*, 1997). The *Scyc2* gene is 909 base pairs (bp) long in *S. vulgaris* and the sequence amplified here contains the first ~711 bp of the gene (Fig. 2.3). *PgiC* was fully characterised in *Clarkia lewisii* by Thomas *et al.* (1992). The gene comprises 23 exons and 22 introns, and exons 11 to 16 (~770 bp) are amplified in this study (Fig. 2.3).

6.2 MATERIALS AND METHODS

6.2.1 *Plant material*

Seed were collected from populations of the species shown in table 6.1. The species investigated include the hexaploid species *S. cambrensis* and *S. teneriffae*, the tetraploid *S. vulgaris* (var. *vulgaris* and var. *hibernicus*) and several diploid members of *Senecio* sect. *Senecio*. *S. flavus* was included as outgroup based on previous ITS sequence analysis (Comes and Abbott, 2001; Coleman *et al.*, 2003). A previous analysis of the homoploid hybrid species *S. squalidus* revealed polymorphism for the *Scyc2* DNA sequence in both Edinburgh and Oxford populations (M.A. Chapman and R.J. Abbott, unpubl. data). Hence, two individuals from the Edinburgh population that were known to differ in *Scyc2* sequence were selected for analysis, and in addition, four UK populations of *S. squalidus* were analysed using PCR-RFLP to determine whether other populations of this species are polymorphic for these two sequences.

Plants were raised from seed in a greenhouse at the University of St Andrews. DNA was extracted from ~200 mg leaf material using the CTAB-based extraction protocol detailed in section 2.4.2. DNA samples were diluted to ~5 ng/µl and stored at -20°C until use.

Table 6.1 – Species surveyed in the analysis of DNA sequence variation

Species	Location of sample	Collector[1]
S. aethnensis	Cisternazza, Sicily	JKJ
S. cambrensis	Edinburgh, UK	RJA
S. chrysanthemifolius	Catania, Sicily	JKJ
S. flavus	Tafraoute, Morocco	RJA
S. gallicus	Sa Riera, Spain	RJA
S. glaucus subsp. *coronopifolius*	Houmt Souk, Tunisia	MC/RJA
S. leucanthemifolius	Gioia Tauro, Italy	HPC/RJA
S. rupestris	Mgna della Maiella, Italy	HPC/RJA
S. squalidus-1	Edinburgh, UK	RJA
S. squalidus-2	Edinburgh, UK	RJA
S. teneriffae	Tenerife, Spain	JD
S. vernalis	Western Galilee, Israel	HPC/RJA
S. vulgaris var. *hibernicus*	Edinburgh	RJA
S. vulgaris var. *vulgaris*	Edinburgh	RJA

[1] Collectors – R. J. Abbott, M. Coleman, H.-P. Comes, J. Dickson, J. K. James.

6.2.2 *Genetic Analysis*

The polymerase chain reaction (PCR) cycling parameters were identical for the amplification of *Scyc2*and *PgiC*. These were 94°C for 4 min, followed by 35 cycles of 94°C for 30 s, 55°C for 1 min, 72°C for 2 min and a final 10 min extension at 72°C. Primers used to amplify the portion of the two genes were as follows: for *Scyc2*: S2-START: CTAATGTTTTCCA; S2-08: TAAACACTCTTTGACTCGAT, and for *PgiC*: AA11f: TTYGCNTTYTGGGAYTGGGT; AA16r: CCYTTNCCRTTRCTYT-CCAT, where Y = C or T; R = A or G; N = A, C, G or T. Primer annealing positions for the two genes are shown in Fig. 2.3. Primers for *Scyc2* were designed based on the sequence obtained from *S. vulgaris* whereas the primer sequences for *PgiC* were provided by L. Gottlieb (University of California, Davis) and had previously been used to amplify successfully exons 11-16 of *PgiC* in some other members of the Asteraceae (L. Gottlieb, pers. comm.).

PCR was carried out in 20 µl reactions containing 2 µl 10 x PCR buffer (Qiagen; supplied with enzyme), 2 mM $MgCl_2$, 0.1 mM dNTPs, 0.5 mM each primer, 0.5 Units *Taq* DNA polymerase (Qiagen), 4 µl Q solution (Qiagen) and ~5 ng genomic DNA. PCR was carried out in a Perkin Elmer GeneAmp 9700 thermal cycler

and products were resolved on 1.2% agarose gels stained with ethidium bromide. PCR amplification of *PgiC* sometimes revealed two products differing in size by ~50bp. Upon sequencing (see below), the shorter of the two products was found to have a deletion at the end of exon 13 and was hence regarded as non-functional. This product was not considered further in this investigation.

PCR products of *Scyc2* and *PgiC* were ligated into pGEM-Teasy vectors and transformed into *E. coli* cells following the manufacturers' protocol (Promega) with some minor modifications (section 2.6). Clones containing the correct size insert were identified by performing a 10 µl PCR with a small amount of colony material from 12 to 24 clones per sequence per individual and resolving products in 1.5% agarose gels. PCR conditions and primers were as above. Two or three positive clones from each of the diploid species for *PgiC* and *Scyc2* were sequenced. In an attempt to identify the different homoeologues in the tetraploid and hexaploid species a repeat PCR (20 µl) was carried out on ten to 12 positive clones and the products were digested with three restriction enzymes (*Alu*I, *Hpa*II, *Hind*III, using 5 µl PCR product per restriction) following the enzyme manufacturers' protocol (Promega). The digested PCR products were separated in 8% acrylamide gels and two to three clones that exhibited different restriction patterns for each enzyme were sequenced. For *Scyc2*, three homoeologues (judged by different restriction patterns) were recovered from *S. teneriffae* and two from *S. vulgaris*. However, only two of such homoeologues were recovered from *S. cambrensis* (see below). For *PgiC* three homoeologues were recovered from *S. teneriffae* and *S. cambrensis* and two from *S. vulgaris*.

The clones selected for sequencing were cultured overnight in 5 ml volumes, and plasmid DNA was extracted using the PerfectPrep kit (Eppendorf). Insert sequence was determined in both directions using pUC/M13 vector-specific primers and Dye-Terminator Cycle Sequencing (DTCS) kits (Beckman-Coulter). Products were precipitated with 96% ethanol and washed two or three times using 70% ethanol before being resuspended in 40 µl of sample loading solution (Beckman-Coulter). Products were the sequenced using a Beckman-Coulter CEQ8000 automated sequencer. Electropherograms were analysed using Chromas ver. 1.0.0.1 (Technelysium Pty Ltd, 1998-2001) and sequences were aligned manually using GeneDoc ver. 2.6.001 (Nicholas and Nicholas Jr., 1997).

Sequencing of *Scyc2* products from *S. cambrensis* that differed in restriction patterns revealed a single *EcoR*I restriction enzyme cut site in one homoeologue that was absent in the other. A subsequent restriction digestion of twelve cloned *Scyc2* sequences from *S. cambrensis* using *EcoR*I revealed that nine clones contained the cut site and three did not. Further sequencing of four additional clones that possessed the *EcoR*I cut site revealed a third sequence that was almost identical to the original sequence that contained the cut site. A possible reason for presence of two near-identical homoeologues is discussed below.

The previous finding that two divergent copies of the *Scyc2* sequence are present in two populations of *S. squalidus* was investigated further by performing a PCR-RFLP analysis of *Scyc2* sequences in four populations of *S. squalidus*. The *squalidus*-1 sequence contained two *Hinf*I cut sites (G ↓ ANTC) at positions 235 and 694. The cut site at position 235 was absent from *squalidus*-2 (GAGTC in *squalidus*-1, GAGAC in *squalidus*-2). *Scyc2* was amplified by PCR (as above) in six individuals of *S. squalidus* from four UK populations (Wigan, Cardiff, Edinburgh and St. Helens). PCR products were subsequently digested with *Hinf*I, following the enzyme manufacturer's protocol (Promega), separated in 8% acrylamide gels and stained with ethidium bromide. A genotype was assigned to each individual based on the banding pattern observed. Individuals homozygous for the *squalidus*-1 sequence exhibited bands of size 459, 235 and 13 bp whereas individuals homozygous for the *squalidus*-2 sequence exhibited bands of size 694 and 13 bp. Heterozygous individuals exhibited bands of 694, 459, 235 and 13 bp.

The nucleotide sequences obtained for *Scyc2* and *PgiC* from *S. flavus* and *S. squalidus* were used to search the GenBank database using the BLAST search tool (Altschul *et al.*, 1997). In addition, amino acid sequences of exons 11-16 of *PgiC* from some members of the Asteraceae were provided by L. Gottlieb (unpubl. data) for comparison.

6.2.3 *Phylogenetic analysis*

An incongruence length difference (ILD) test conducted using the partition homogeneity function of PAUP* (Swofford, 2001) revealed that *Scyc2* and *PgiC* were not congruent in their phylogenetic signal ($P = 0.023$). This is likely to be due to reticulate evolution in the taxa investigated here. Hence, the two sequences (*Scyc2* and *PgiC*) were not combined for analysis.

The different homoeologues obtained from each polyploid species were designated A, B or C. Insertions or deletions (indels) were binary coded and polarised relative to the outgroup (*S. flavus*) and added to the alignment matrix as additional characters. Maximum parsimony (MP) analysis was carried out using the heuristic search option of PAUP* with 1000 replicates and the options TBR, MULTREES and COLLAPSE(max) in effect. Maximum likelihood (ML) analysis was carried out in PAUP*, with the same options in effect as for the MP analysis, using the model of DNA substitution and parameters determined using Modeltest (Posada and Crandall, 1998). Bootstrap analysis (1000 replicates) was performed in PAUP* to assess confidence in tree topologies.

6.3 RESULTS

6.3.1 *Comparisons of* Senecio *sequences with those published for other taxa*

Scyc2

The *Scyc2* sequences of *S. flavus* and *S. squalidus* used to search the GenBank database showed significant homology with a range of species. Greatest similarity was found for *cycloidea* homologues isolated from *Acosmium subelegans* (Fabaceae; GenBank accession number AY225829) and *Antirrhinum leptaleum* (Plantaginaceae; AF512078). Sequence homology between species was particularly high for the conserved TCP and R domains.

PgiC exons 11-16

The BLAST search using the nucleotide sequences obtained from *S. flavus* and *S. squalidus* revealed greatest homology with *PgiC* from *Arabidopsis lyrata* (GenBank

accession AY174553) and members of the Onagaceae such as *Fuchsia cylindrica* (AJ550741), *Clarkia dudleyana* (AJ437278) and *C. unguiculata* (AJ312370). However, nucleotide alignments of *PgiC* were difficult to assemble for comparison of sequences obtained from *Senecio* with those obtained from the BLAST search due to high variation in intron length between species and also low sequence similarity in the introns. Further comparisons were therefore made between amino acid sequences of exons 11-16 as these were highly conserved and identical in length. These comparisons showed that the *Senecio* amino acid sequences were ~ 90% identical to those of exons 11-16 in *Fuchsia* and *Clarkia*. With regard to unpublished amino acid sequences for *PgiC* in other Asteraceae (L. Gottleb, pers. comm.), the two *Senecio* sequences were most similar to *Munzothamnus blairii* showing ~ 95% identity. Total length of the five introns in these *Senecio* ranged from 462-469 nucleotides (nts), which is shorter than previously recorded for *A. thaliana*, *C. lewisii* (Thomas et al. 1993) or any other member of the Asteraceae (L. Gottlieb, unpubl. data).

6.3.2 *PCR-RFLP analysis of* Scyc2 *variation in four populations of* S. squalidus

All four populations analysed were polymorphic for *Scyc2* sequence based on the survey of presence/absence of a *Hinf*I restriction enzyme cut site (Table 6.2). However, in only one population (from St. Helens) were all three genotypes present among the six individuals analysed.

Table 6.2 – Frequencies of *Scyc2* genotypes in four populations of *S. squalidus*. Six individuals were analysed per population by PCR-RFLP. 1 and 2 refer to the two sequences present in *S. squalidus* (see text for details).

	Genotype		
Population	1/1	1/2	2/2
Wigan	-	3	3
Cardiff	4	2	-
Edinburgh	-	4	2
St. Helens	1	3	2
TOTAL	5	12	7

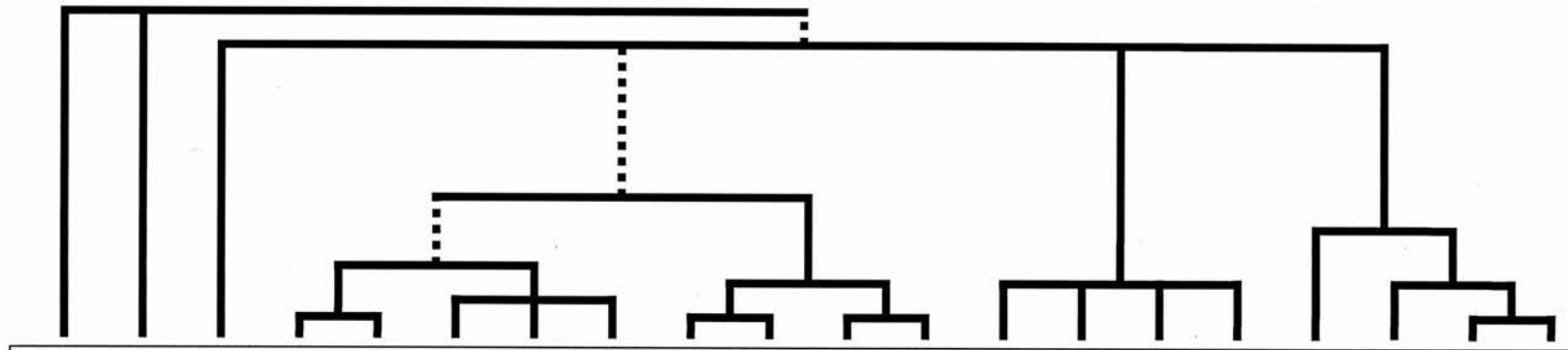
6.3.3 *Sequence alignment and analysis*

Scyc2

The alignment of *Scyc2* in *Senecio* included 711 nucleotides (nts) and four indels (appendix IIIc). Of the 715 characters, 625 (87.4%) were constant and 34 (4.76%) were variable and parsimony informative. Sequence divergence (measured as uncorrected *p* distance) between outgroup and ingroup taxa ranged from 5.40% (*S. flavus* and *S. glaucus*/*S. vernalis*) to 6.61% (*S. flavus* and *S. aethnensis*) (Figure 6.3A). Between ingroup taxa, sequence divergence was least between *S. vulgaris*R-A and *S. squalidus*-1 (0.0%) and greatest between *S. vulgaris*R-B and *S. aethnensis* (2.98%).

PgiC

The alignment of *PgiC* comprised 772 nts and 14 indels (appendix IIId). All indels were located in the introns. Of the 786 characters, 650 (82.7%) were constant and 44 (5.6%) were variable and parsimony informative. Of the parsimony informative characters, nine were nucleotides located in the exons, 30 were nucleotides in the introns and five were indels in the introns. Sequence divergence was greatest between *S. flavus* and *S. vernalis* (8.61%). Between ingroup taxa greatest sequence divergence was between *S. squalidus*-1 and *S. leucanthemifolius* (4.41%) (Figure 6.3B).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
(1) flavus																				
(2) vernalis	5.40																			
(3) glaucus	5.40	1.13																		
(4) gallicus	6.43	1.98	1.84																	
(5) aethnensis	6.61	2.28	2.13	0.57																
(6) chrysanthemifolius	6.16	2.13	1.99	2.13	2.41															
(7) rupestris	6.13	2.12	1.98	2.12	2.42	0.28														
(8) squalidusB	5.98	1.98	1.84	1.98	2.27	0.14	0.14													
(9) leucanthemifolius	5.83	2.02	1.88	2.30	2.59	2.32	2.02	2.17												
(10) teneriffaeC	5.67	1.73	1.59	2.01	2.30	2.03	1.73	1.88	0.29											
(11) vulgarisNa	5.62	1.57	1.43	1.87	2.15	1.85	1.85	1.71	1.32	1.03										
(12) teneriffaeA	6.12	2.12	1.98	2.41	2.70	2.41	2.40	2.26	1.88	1.59	0.86									
(13) squalidusA	5.42	1.13	0.85	1.84	2.13	1.42	1.42	1.28	1.88	1.59	1.42	1.98								
(14) vulgarisRa	5.48	1.15	0.86	1.87	2.16	1.43	1.43	1.28	1.90	1.61	1.43	2.00	0.00							
(15) cambrensisC	5.54	1.27	0.99	1.98	2.27	1.56	1.55	1.41	2.02	1.73	1.57	2.12	0.14	0.14						
(16) cambrensisA	5.69	1.41	1.13	2.12	2.42	1.70	1.69	1.55	2.17	1.88	1.71	2.26	0.28	0.29	0.42					
(17) cambrensisB	6.13	1.98	1.84	2.54	2.84	2.98	2.97	2.82	2.89	2.60	2.28	2.97	1.84	1.86	1.98	2.12				
(18) teneriffaeB	5.70	1.55	1.41	2.40	2.70	2.55	2.54	2.40	2.46	2.17	1.85	2.54	1.42	1.43	1.55	1.69	0.71			
(19) vulgarisNb	5.93	1.72	1.57	2.58	2.87	2.71	2.71	2.57	2.63	2.34	2.00	2.71	1.57	1.57	1.71	1.86	0.85	0.14		
(20) vulgarisRb	6.02	1.84	1.70	2.69	2.98	2.84	2.84	2.70	2.76	2.47	2.14	2.83	1.70	1.72	1.85	1.99	0.99	0.28	0.14	

Figure 6.3 A - Matrix of pairwise sequence differences (uncorrected *p* differences) calculated between individuals for the *Scyc2* sequences.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
(1) flavus																				
(2) leucanthemifolius	8.49																			
(3) gallicus	7.93	3.57																		
(4) glaucus	7.97	3.86	1.46																	
(5) teneriffaeC	7.93	3.83	1.84	2.12																
(6) aethnensis	8.07	3.44	1.97	1.99	0.66															
(7) rupestris	7.80	3.83	1.97	2.12	0.53	0.66														
(8) cambrensisC	7.81	3.70	1.84	1.99	0.40	0.53	0.40													
(9) squalidusA	7.87	4.41	2.26	2.41	2.66	2.80	2.66	2.53												
(10) squalidusB	7.31	3.46	1.99	2.01	2.12	2.25	2.12	1.99	2.67											
(11) chrysanthemifolius	8.22	3.43	2.37	2.39	1.32	1.19	1.31	1.19	2.94	2.40										
(12) cambrensisA	8.08	3.30	2.24	2.26	0.92	1.05	1.18	1.05	2.80	2.26	0.40									
(13) vulgarisRa	7.81	3.70	1.84	1.99	0.40	0.53	0.40	0.00	2.53	1.99	1.19	1.05								
(14) teneriffaeA	7.94	3.97	1.98	1.86	1.05	0.92	1.06	0.92	2.54	2.12	1.06	0.93	0.92							
(15) vulgarisNa	8.34	3.71	2.37	2.52	0.66	1.05	0.92	0.79	3.07	2.52	1.19	0.79	0.79	1.45						
(16) vernalis	8.61	3.97	2.63	2.79	1.19	0.79	1.19	1.05	3.34	2.78	1.72	1.59	1.05	1.72	1.58					
(17) vulgarisNb	8.48	1.72	3.43	3.72	3.57	3.17	3.56	3.44	4.14	3.33	3.30	3.17	3.44	3.84	3.44	3.44				
(18) vulgarisRb	8.35	1.85	3.04	3.59	3.44	3.30	3.43	3.31	4.00	3.33	3.17	3.03	3.31	3.70	3.31	3.58	0.92			
(19) teneriffaeB	8.22	1.72	3.43	3.72	3.70	3.57	3.56	3.57	4.27	3.33	3.30	3.17	3.57	3.83	3.57	4.10	1.85	1.72		
(20) cambrensisB	8.22	1.72	3.43	3.72	3.70	3.57	3.56	3.57	4.27	3.33	3.30	3.17	3.57	3.83	3.57	4.10	1.85	1.72	0.00	

Figure 6.3 B - Matrix of pairwise sequence differences (uncorrected *p* differences) calculated between individuals for the *PgiC* sequences.

6.3.4 *Phylogenetic analysis*

Scyc2

Fitch parsimony resulted in four most parsimonious trees of length 103 steps with a consistency index (CI) of 0.9223 and retention index (RI) of 0.9149 (Fig. 6.4). The model of evolution was determined as HKY + Γ using Modeltest (Posada and Crandall, 1998), which follows the substitution model of Hasegawa *et al.*, (1985) and gamma (Γ)-distributed rate variation. The HKY model assumes unequal base frequencies, transition ratio is not equal to transversion ratio, and unequal rates among sites. The single ML tree (-lnL = 1646.0522) was identical in topology to the MP tree shown in Fig. 6.4.

Two indels characterised the outgroup *S. flavus* (a 21 bp and a 3 bp deletion), and a further 15bp deletion occurred only in the *S. leucanthemifolius* and *S. teneriffae*-C *Scyc2* sequence. One other indel, a 3bp insertion at position 153-5 was present in only *S. leucanthemifolius*, *S. teneriffae*-C, *S. aethnensis* and *S. gallicus* sequences and would appear to have had two independent origins according to the tree (Fig. 6.4).

The two different copies of *Scyc2* obtained from individuals of *S. squalidus* (*squalidus*-1 and *squalidus*-2) are positioned in different clades of the MP tree (Fig. 6.4). The *squalidus*-2 sequence clusters with the sequence obtained from *S. chrysanthemifolius*, one of the parent species of *S. squalidus* (Abbott *et al.*, 2000). In contrast, the *squalidus*-1 sequence clusters with one sequence obtained from *S. vulgaris* var. *hibernicus* (*vulgaris*R-A) and *S. cambrensis*. Visual inspection of the *squalidus*-1 sequence indicates that the latter ~360 bp are identical in sequence to the same part of the *Scyc2* sequence obtained from *S. chrysanthemifolius*, whereas the first ~350 bp differ considerably and contain ten nucleotide substitutions, eight of which are non-synonymous resulting in changes in protein sequence (Fig. 6.5A). The entire *squalidus*-2 sequence, however, is only one nucleotide different from that of *S. chrysanthemifolius* (a non-synonymous substitution at position 319; Fig. 6.5A). When MP analysis is carried out on the first half of the *Scyc2* sequences (Fig. 6.5B), the *squalidus*-1 and *squalidus*-2 sequences are found in distinct clades. However, MP analysis of the second half of the *Scyc2* sequences results in both *S. squalidus* sequences and those obtained from *S. chrysanthemifolius* and *S. vulgaris* var. *hibernicus* (*vulgaris*R-A) forming a single clade (Fig. 6.5C).

Two copies of the *Scyc2* sequence are present in both varieties of *S. vulgaris* (var. *vulgaris* and var. *hibernicus*) examined. In addition, *S. cambrensis* and *S. teneriffae* both contain two copies of the sequence (A and B) which cluster closely with the A and B copies present in *S. vulgaris*. With regard to the third copy of *Scyc2* sequence (copy C) found in the hexaploids, the copy obtained from *S. teneriffae* is closely related to that found in *S. leucanthemifolius* (99% bootstrap support), whereas the C copy obtained from *S. cambrensis* is found in the same sub-clade as *vulgaris*R-A, *cambrensis*-A and *squalidus*-1.

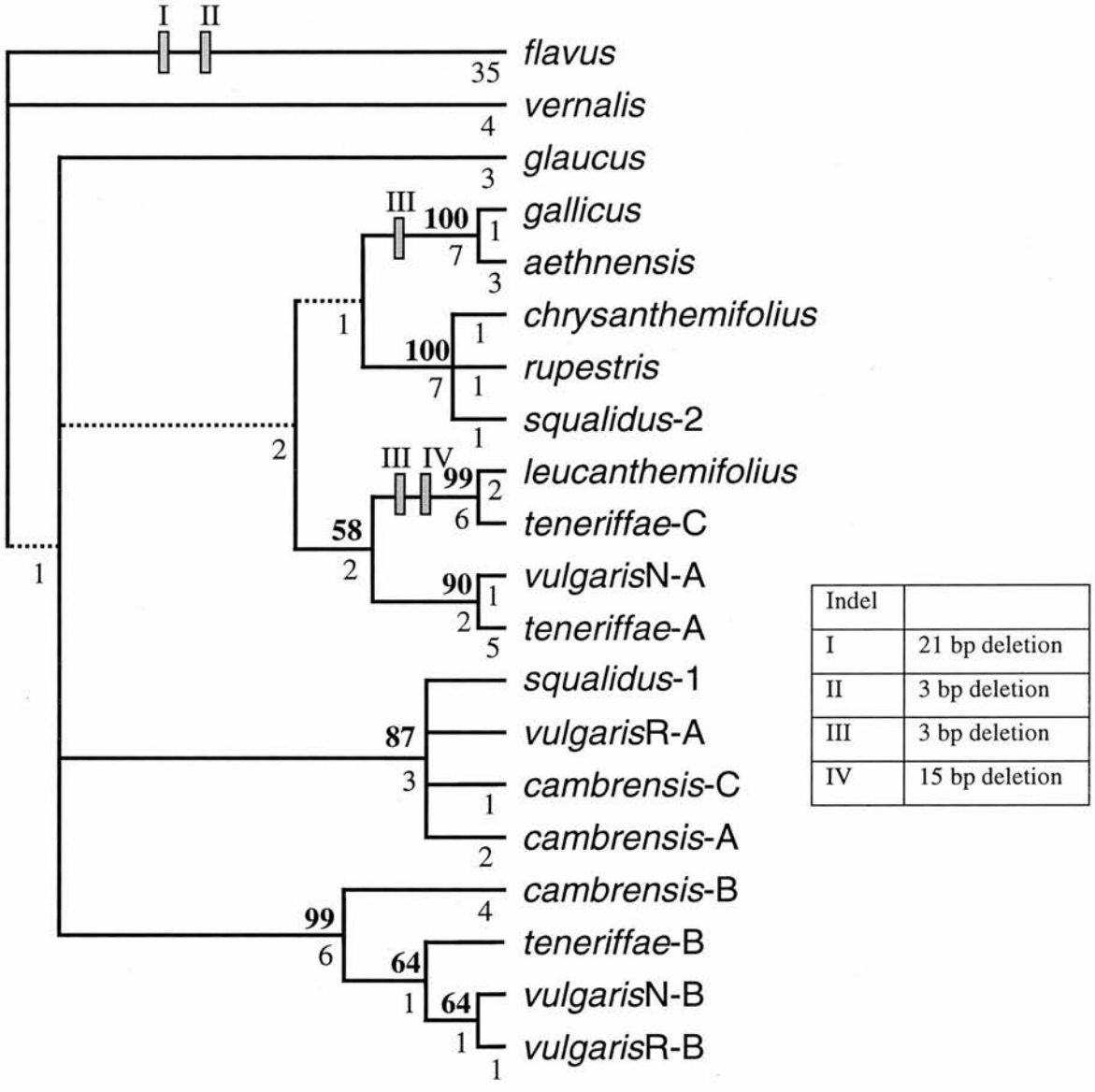


Figure 6.4 – One of four equally parsimonious MP trees generated from *Scyc2* sequence data. The tree shown agrees with the topology of the ML tree. Multiple copies of *Scyc2* present in the polyploid species are designated A, B and C. The two varieties of *S. vulgaris* (Radiate var. *hibernicus*, R and non-radiate var. *vulgaris*, N) are indicated. Branches that collapse under strict consensus are indicated by dotted lines. Bootstrap percentages (≥50%; 1000 replicates) are indicated above branches. Number of nucleotide changes are indicated below branches. Selected insertions/deletions (indels) are shown by shaded boxes along branches and are labelled with roman numerals.

A

Nt position[1]	75	76	112	127	129	130	132	153	158	160	162	164	167	222	241	243	319	373	381	478	480	513	570	618	673	680
Amino Acid	25	27	38	43	43	44	44	51	53	54	54	55	56	74	81	81	107	125	127	160	160	171	190	206	225	227
vulgarisN-A	T	T	T	G	A	T	G	-	C	C	G	G	G	T	T	A	A	C	A	G	T	A	G	G	T	T
vulgarisR-A	T	T	C	C	A	T	C	-	C	A	G	A	A	T	T	A	A	C	T	G	T	A	C	A	T	C
squalidus-1	T	T	C	C	A	T	C	-	C	A	G	A	A	T	T	A	A	C	T	G	T	A	C	A	T	C
chrysanth.	T	C	T	A	A	C	C	-	C	A	A	G	G	C	A	A	G	C	T	G	T	A	C	A	T	C
squalidus-1	T	C	T	A	A	C	C	-	C	A	A	G	G	C	A	A	A	C	T	G	T	A	C	A	T	C
aethnensis	A	T	T	C	T	C	C	CTC	T	A	G	G	G	T	T	T	A	T	T	T	G	G	G	A	C	T
Substitution[2]	S	N	N	N	N	N	N	I	N	S	S	N	N	S	N	N	N	S	N	N	N	S	N	S	S	N

[1]Relative to start of alignment; [2] Synonymous (S), non-synonymous (N) or insertion (I)

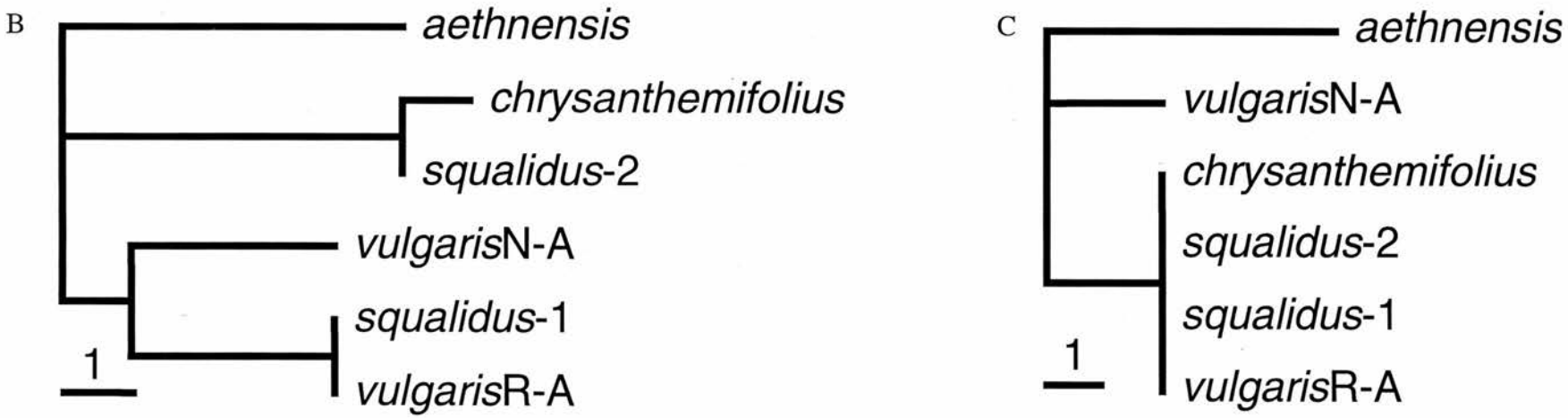


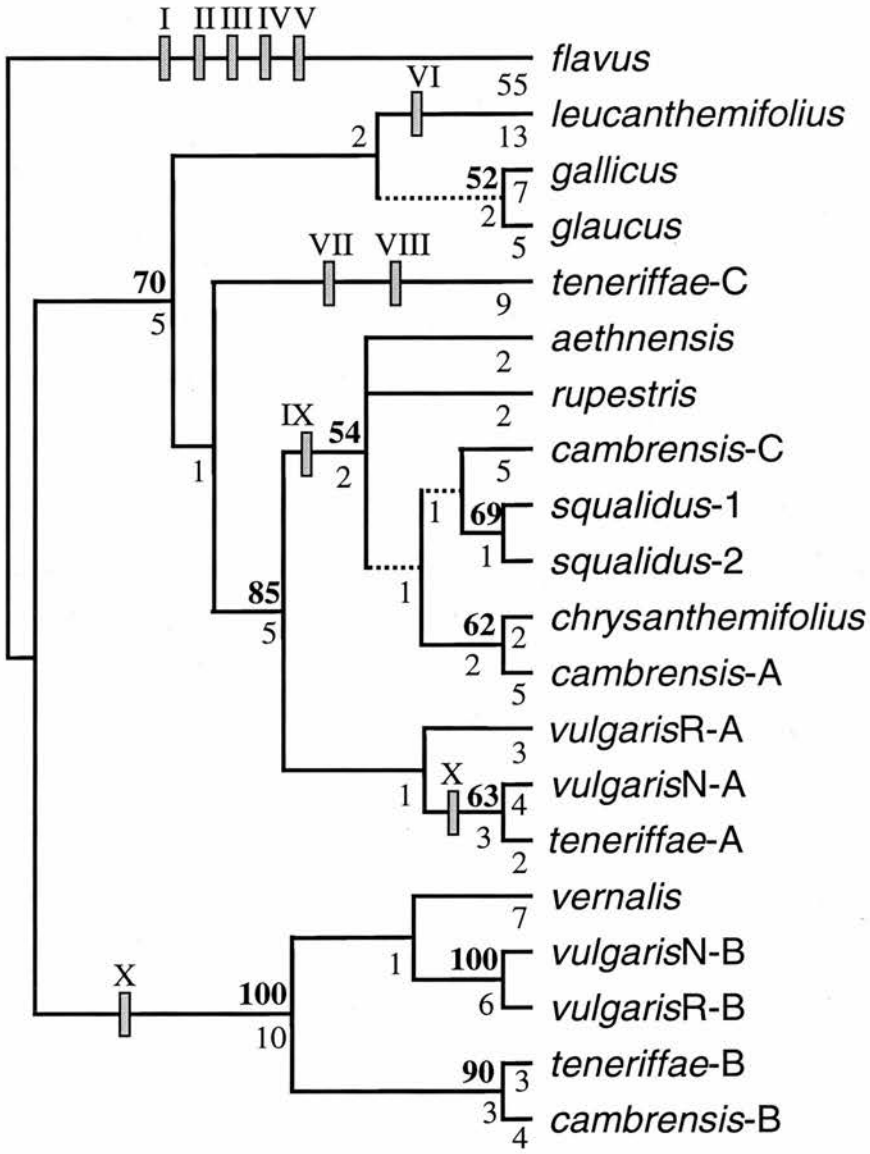
Figure 6.5 – (A) Table summarising the nucleotide (nt) differences present in the sequence alignment of *Scyc2* for six sequences. The double vertical line indicates the two halves of the sequence. Shaded regions show identical nucleotides between *S. squalidus* and *S. chrysanthemifolius*. (B) Single MP tree (length 18 steps) generated using the first 350 nts (includes one indel) of the *Scyc2* alignment. (C) Single MP tree (length 9 steps) generated using nts 351-711 of the *Scyc2* alignment. See text for details.

PgiC

Maximum parsimony analysis of *PgiC* sequences resulted in 22 equally parsimonious trees of length 174 steps with a CI and RI of 0. 8448 (Fig. 6.6). The same model of evolution (HKY + Γ) was determined using Modeltest (Posada and Crandall, 1998) as for the *Scyc2* sequence alignment. Actual parameters (e.g. transition and transversion ratio, base frequencies) were not identical for the *Scyc2* and *PgiC* analysis. ML analysis resulted in one tree (-lnL = 1938.8931) which was identical in topology to the MP tree shown in Fig. 5. Five indels separated *S. flavus* from the other taxa and one 3-bp deletion (indel X) appears to have had two independent origins (Fig. 6.6).

The topology of the *PgiC* phylogenetic tree (Fig. 6.6) is partly congruent with the tree obtained for *Scyc2* sequences. In both trees a close relationship is found between *S. rupestris*, *S. chrysanthemifolius* and *S. squalidus* (although not *squalidus*-1 in the *Scyc2* tree). However, the relative phylogenetic positions of *S. glaucus*, *S. gallicus* and *S. leucanthemifolius* vary between trees. In the *PgiC* tree these three sequences form a single sub-clade, whereas for the *Scyc2* sequences these three species resolve apart from each other. In addition, a well-supported clade (99-100% bootstrap support) containing the B copy of each sequence from *S. vulgaris* (var. *vulgaris* and var. *hibernicus*), *S. cambrensis* and *S. teneriffae* is resolved in both trees. *S. vernalis* is found in this clade for the *PgiC* phylogenetic tree, however is not for the *Scyc2* tree.

Both radiate and non-radiate *S. vulgaris* were found to possess two copies of *PgiC* sequence and *S. cambrensis* and *S. teneriffae* each contained a copy of *PgiC* sequence that is closely related to that of the B sequence in *S. vulgaris*. The A copy of *PgiC* obtained from *S. teneriffae* is closely related to that of the A copy from non-radiate *S. vulgaris* (*vulgaris*N-A), whereas the C copy is not closely related to any particular taxon and is not placed in the same clade as *S. leucanthemifolius* as it was in the *Scyc2* tree. The C copy obtained from *S. cambrensis* clusters with *S. squalidus* whereas the B copy is not resolved in the same clade as *S. vulgaris* as was found for *Scyc2*.



Indel	
I	8 bp insertion
II	1 bp insertion
III	2 bp deletion
IV	14 bp deletion
V	1 bp deletion
VI	9 bp deletion
VII	2 bp deletion
VIII	3 bp deletion
IX	2 bp deletion
X	3 bp deletion

Figure 6.6 – One of 22 equally parsimonious MP trees generated from *PgiC* sequence data. The tree shown agrees with the topology of the ML tree. Multiple copies of *PgiC* present in the polyploid species are designated A, B and C. The two morphs of *S. vulgaris* (Radiate var. *hibernicus*, R and non-radiate var. *vulgaris*, N) are indicated. Branches that collapse under strict consensus are indicated by dotted lines. Bootstrap percentages (≥50%; 1000 replicates) are indicated above branches. Number of nucleotide changes are indicated below branches. Selected insertions/deletions (indels) are shown by shaded boxes along branches and are labelled with roman numerals.

6.4 DISCUSSION

The origin of Senecio squalidus

Previous sequence analysis of *Scyc2* in *S. squalidus* has revealed that this species is polymorphic for two *Scyc2* DNA sequences (M.A. Chapman and R.J. Abbott, unpubl. data) and this is clearly shown in the phylogeny of *Scyc2*. The two individuals selected for analysis were not polymorphic for *PgiC*.

PCR-RFLP analysis of the *Scyc2* sequences in 24 individuals of *S. squalidus* reveals that both copies (based on the presence/absence of a single restriction site) are present in the four populations analysed. *S. squalidus* is thought to have been introduced into the UK only once, and hence the polymorphism in the four populations analysed suggests that the species was polymorphic for *Scyc2* before it expanded its range from its place of introduction and this polymorphism has not arisen through recent mutation or introgression.

One copy of the *Scyc2* sequence from *S. squalidus* (*squalidus*-2) resolved with *S. chrysanthemifolius* in the MP tree (100% bootstrap support), while the single *S. squalidus PgiC* sequence also resolved closely with *S. chrysanthemifolius*, confirming the role of this species in the origin of *S. squalidus* (James, 1999; Abbott *et al.*, 2000, 2002). However, the second *Scyc2* sequence obtained from *S. squalidus* (*squalidus*-1) was not recovered from either of the proposed parent species, *S. chrysanthemifolius* and *S. aethnensis*, and is highly divergent from both.

The *squalidus*-1 sequence could be present in either or both *S. chrysanthemifolius* and *S. aethnensis* at low frequency, however the phylogenetic placement of the *squalidus*-1 and *squalidus*-2 sequences indicates that divergence of these two alleles predated the divergence of several of the diploid species investigated (i.e. *S. gallicus*, *S. aethnensis*, *S. chrysanthemifolius*, *S. rupestris* and S. *leucanthemifolius*). An alternative possibility is that the divergent nucleotide sequence of *squalidus*-1 was introgressed into *S. squalidus* following the introduction of this species to the UK. *S. squalidus* hybridises with *S. vulgaris* at low frequency in the wild (Marshall and Abbott, 1980), and therefore the *squalidus*-1 sequence might have arisen by mutation in *S. vulgaris* and introgressed into *S. squalidus*. Additionally, *S. squalidus* is reported to occasionally hybridise with *S. viscosus* in the UK (Abbott and

Lowe, 1996), so this species is also a potential donor of the *squalidus*-1 sequence. However, we have no molecular data concerning *S. viscosus*.

On closer inspection of the *squalidus*-1 sequence, it was found that the second half of it is identical to the same part of the *Scyc2* sequence in *S. chrysanthemifolius*; however, the first ~350 bp of the sequence differs by ten base substitutions (Fig. 6.5). The possibility that mutation is a cause of these differences seems unlikely, as base substitutions caused by mutation would be expected to be distributed along the length of the sequence and not restricted to only the first part of the gene. Moreover, substitutions were not found to be limited to the first half of the gene when the sequences from other species were compared to *S. squalidus* (Fig 6.5A; appendix IIIc). It is suggested, therefore, that the pattern of nucleotide substitution recorded in the *squalidus*-1 sequence is concordant with a scenario in which intragenic recombination of *Scyc2* has occurred. The possibility that this particular sequence is a PCR-generated chimera (e.g. Bradley and Hillis, 1997; Cronn *et al.*, 2002; Ingram and Doyle, 2003; Feliner *et al.*, 2004), and therefore an artefact, is ruled out because very similar sequences were obtained from both *S. vulgaris* (*vulgaris*R-A) and *S. cambrensis* (*cambrensis*-A and *cambrensis*-C). Further evidence that this sequence is not a PCR-generated chimera is provided by an additional analysis of *S. squalidus* and *S. vulgaris* var. *hibernicus* (data not shown). This revealed that (1) both the A and B copies of *Scyc2* sequence are present in Oxford and Edinburgh populations of *S. squalidus*, and (2) the same sequence as *squalidus*-1 (i.e. *vulgaris*R-A) is present in all four individuals of *S. vulgaris* var. *hibernicus* from Edinburgh examined. Additionally, the results of the PCR-RFLP analysis of *Scyc2* in four populations of *S. squalidus* indicate that both the *squalidus*-1 and *squalidus*-2 sequences are present in other populations.

A phylogenetic analysis of the first ~350 bp of *Scyc2* sequence did not reveal a close similarity of *squalidus*-1 with any of the other *Scyc2* sequences obtained, with the exception of the *vulgaris*R-A, *cambrensis*-A and *cambrensis*-C sequences, which are thought to be derived from *S. squalidus* and not vice versa (see below). *Scyc2* is part of a family of six closely-related *cycloidea* homologues isolated from *S. vulgaris* (A.C.M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott, unpubl. data), however only *Scyc2* has been analysed in the species investigated here. Secondly, in several investigations into *cyc*-like genes in other species, a very closely related gene,

dichotoma (*dich*), has been isolated (e.g. Luo *et al.*, 1999; Shepard and Purugganan, 2002; Gubitz *et al.*, 2003; Hileman and Baum, 2003; Reeves and Olmstead, 2003). Therefore it is possible that the *Scyc2* sequence in *S. squalidus* is a chimera of two *Scyc* homologues or is a *cyc/dich* recombinant gene. At present, we can only offer this as a hypothesis; further work will be required to investigate these possibilities.

The *Scyc2* and *PgiC* sequence data presented here provide strong support for the hypothesis that *S. squalidus* is derived from *S. chrysanthemifolius* (Abbott *et al.*, 2000, 2002); however the recovery of a second, unusual, *Scyc2* sequence from *S. squalidus* cannot be explained by the current data.

The origin of S. vulgaris

Tetraploid *Senecio vulgaris* was hypothesised to be the autotetraploid of *S. vernalis* Waldst. & Kit. by Kadereit (1984b) on the basis of natural and artificial hybridisations and studies of morphological and geographic affinities. This relationship is supported by a recent survey of rDNA ITS sequence variation among several Mediterranean species of *Senecio* that confirmed a close relationship between *S. vulgaris* and *S. vernalis* (Comes and Abbott, 2001). Harris and Ingram (1992b), however, argued that *S. vernalis* was unlikely to be the only parent of *S. vulgaris*, due to differences in cpDNA and rDNA RFLP type, although only one *S. vernalis* individual was examined in their analysis. An allotetraploid origin of *S. vulgaris* was also proposed based on a survey of allozyme variation (Ashton and Abbott, 1992b).

In the present study, *S. vulgaris* var. *vulgaris* was found to possess two distinct copies (A and B) of both the *Scyc2* and *PgiC* sequences which is consistent with an allotetraploid origin. In the phylogeny constructed from the *PgiC* data, the B copy from *S. vulgaris* is resolved as sister to *S. vernalis*, but this relationship is not supported by the phylogeny based on *Scyc2* sequences. The second copy of *Scyc2* in *S. vulgaris* var. *vulgaris* (*vulgaris*N-A) was found to be closely allied to diploid *S. leucanthemifolius*, whereas the second copy of the *PgiC* sequence (copy A) was positioned in a clade (85% bootstrap support) that included the diploids *S. aethnensis*, *S. rupestris* and *S. chrysanthemifolius*, but not *S. leucanthemifolius*. Taken together we conclude that *S. vulgaris* is an allotetraploid, possibly of *S. vernalis* and another diploid species that is currently undetermined based on existing data.

Radiate *S. vulgaris* is thought to be of introgressive origin containing a small amount of genetic material of *S. squalidus*, including the gene responsible for the presence of ray florets (the *RAY* gene) in capitula (Abbott *et al.*, 1992; Lowe and Abbott, 2000; Abbott *et al.*, 2003). *Scyc2* is tightly linked to *RAY* (A.C.M. Gillies, P. Cubas, M.A. Chapman, E. Coen and R.J. Abbott, unpubl. data) and hence the presence of an identical *Scyc2* DNA sequences in *S. squalidus* (*squalidus*-1) and *S. vulgaris* (*vulgaris*R-A) might be taken as evidence for introgression. However, until the origin of the *squalidus*-1 copy of *Scyc2* is ascertained, we cannot confirm that this copy has been introgressed into *S. vulgaris*, rather than having originated *de novo* in *S. vulgaris* and then introgressing into *S. squalidus*.

***Confirmed origin of* Senecio cambrensis**

Previous studies into the origin of *Senecio cambrensis* have shown that *S. cambrensis* exhibits a genetic profile (rDNA, isozymes and RAPDs) that combines those of *S. squalidus* and *S. vulgaris* (Ashton and Abbott, 1992a; Harris and Ingram, 1992a; Lowe and Abbott, 1996; Abbott *et al.*, 2003). However, in these studies, no other species were included in the analysis, except for *S. teneriffae* and *S. glaucus* in the study by Lowe and Abbott (1996); hence it is possible that other closely-related species were involved in the formation of *S. cambrensis*. The present study, utilising DNA sequence information obtained from several additional diploid members of *Senecio* sect. *Senecio* has confirmed that *S. cambrensis* is the allohexaploid of *S. vulgaris* and *S. squalidus*. An allohexaploid is expected to exhibit three copies of each gene (homoeologues); however in the analysis of *Scyc2*, two sequences (*cambrensis*-A and *cambrensis*-C) were almost identical. This can be explained by the presence of an identical *Scyc2* sequence in both *S. squalidus* and radiate *S. vulgaris*. Hence, the A and C copies of *Scyc2* would seem to have been inherited from radiate *S. vulgaris* (*vulgaris*R-A) and *S. squalidus* (*squalidus*-1), in which case the form of *S. vulgaris* that gave rise to this population (Edinburgh) would have been var. *hibernicus*.

For the *PgiC* sequence, three homoeologues (A, B and C) were clearly identified. One of which (*cambrensis*-B) is closely related to the B copy from *S. vulgaris* and another (*cambrensis*-C) is closely related to *S. squalidus*. However, the third copy, *cambrensis*-A, does not resolve, as would be expected, with the A copy of *S. vulgaris*; instead it is positioned in the same clade as *S. squalidus*. It is possible

therefore that this particular sequence (*cambrensis*-A) is in fact a different allele of the *cambrensis*-C homoeologue of *PgiC*. If this is correct, then it means that a *vulgaris*-A-like sequence was not amplified successfully in the *S. cambrensis* individual examined.

***The origin of* Senecio teneriffae**

The results confirmed the role of *S. vulgaris* in the formation of *S. teneriffae*; for both *Scyc2* and *PgiC* the A and B copies from *S. teneriffae* were closely related to the A and B copies obtained from *S. vulgaris*. The C copy, however, differs in phylogenetic placement when one compares the result for *Scyc2* with that of *PgiC*. In the *Scyc2* phylogeny the C copy in *S. teneriffae* is very closely related to *S. leucanthemifolius*, differing by only two nucleotide changes and supported by a 99% bootstrap value. However, for *PgiC* the *teneriffae*-C copy is not closely related to the *PgiC* sequence obtained from *S. leucanthemifolius*, or indeed from any other species. For this gene, the diploid parent cannot be concluded; hence the involvement of *S. leucanthemifolius* as a parent of *S. teneriffae* cannot be confirmed. However, these results do not support the previous hypothesis that the diploid parent of *S. teneriffae* was *S. glaucus* (Lowe and Abbott, 1996).

Conclusions and Future Work

In this investigation, the use of DNA sequence variation in a portion of two single-copy nuclear genes has provided information regarding the origin of three polyploid species. Firstly, *S. vulgaris* is now confirmed to be an allotetraploid and *S. vernalis* and possibly *S. leucanthemifolius* are proposed as parents. Secondly, the parentage of the hexaploid *S. cambrensis* has been confirmed, with the radiate form of *S. vulgaris* being the more likely tetraploid parent. Thirdly, the hexaploid *S. teneriffae* has been shown to differ from *S. cambrensis* in that the diploid parent was not *S. squalidus*, and instead was possibly *S. leucanthemifolius*, although this is not supported by the analysis of the *PgiC* sequence data.

S. cambrensis is thought to have arisen independently in Wales and Edinburgh, and possibly more than once within Wales (Ashton and Abbott, 1992a; Harris and Ingram, 1992a). In addition, different lengths of ray florets are present in

each population; hence individuals with shorter rays may have originated from crosses between non-radiate *S. vulgaris* and *S. squalidus*. Further sequencing of *Scyc2*, which is tightly linked to the *RAY* gene, in individuals with different ray lengths will provide information regarding whether ray length is caused by a gene dosage effect.

An unexpected finding to emerge from this study is that two alleles of *Scyc2* are present in *S. squalidus*, one of which may be recombinant in origin. Further investigations into *cycloidea* and related homologues in *Senecio* are required to clarify the origin of this allele.

CHAPTER 7 – GENERAL DISCUSSION

The studies reported in this thesis examined different aspects of the role of hybridisation in the evolution of certain taxa of *Senecio* sect. *Senecio* (Asteraceae). Previous investigations carried out on members of sect. *Senecio* have documented cases of contemporary gene flow between species (Abbott and Lowe, 1996; James, 1999; Comes and Abbott 1999) and have verified the hybrid origin of several taxa (Abbott *et al.*, 1992; Ashton and Abbott, 1992; Harris and Ingram, 1992a; Irwin and Abbott, 1992; Lowe and Abbott, 1996, 2000, 2003; Abbott *et al.*, 2000, 2002). The results of the studies reported here further assert the importance of hybridisation in the evolution of sect. *Senecio*. The main findings are detailed below and include possible areas of future research.

Chapter 3

The role of conspecific pollen advantage (CPA) in maintaining a natural hybrid zone between *S. aethnensis* and *S. chrysanthemifolius* on Mt. Etna, Sicily, was investigated by performing control pollinations between these species. Five different pollen ratios (100:0, 75:25, 50:50, 25:75 and 0:100 conspecific: heterospecific) were applied to flower heads of individuals of the two species, and the resultant seed collected and allowed to germinate.

The proportion of hybrid offspring produced by the high altitude species, *S. aethnensis*, was not significantly different from the proportion of heterospecific pollen applied to the flower head. Conversely, the low altitude species, *S. chrysanthemifolius*, produced significantly less hybrid offspring than predicted based on the proportion of *S. aethnensis* pollen applied. In *S. chrysanthemifolius* an increase in the proportion of heterospecific pollen applied was not accompanied by a reduction in seed set or seed germination indicating that conspecific pollen exhibits an advantage over heterospecific pollen before ovule fertilisation takes place. This asymmetry in CPA suggests that gene flow is more likely to occur from *S. chrysanthemifolius* into *S. aethnensis*, and previous molecular analysis of the hybrid zone indicates that *S. aethnensis* may be more introgressed than *S. chrysanthemifolius* (James, 1999). If this is confirmed then *S. aethnensis*, which is endemic to Mt. Etna,

may be threatened with the risk of extinction due to gene flow from *S. chrysanthemifolius*.

Extensive hybrid swarms between *S. aethnensis* and *S. chrysanthemifolius* have been observed at intermediate altitudes on Mt. Etna, however taxonomic identity is maintained, despite the absence of strong CPA in either species. Additionally, from personal observations, it seems unlikely that pollinator preference or a reduction in hybrid fitness maintains the hybrid zone. Instead, it is possible that the two species are adapted to different environmental conditions (e.g. temperature and level of UV-B irradiation) at the extremes of the altitudinal transect. Further investigation of the ecological tolerances of *S. aethnensis* and *S. chrysanthemifolius* would be of value to understand the mechanisms that contribute to the maintenance of this hybrid zone.

Chapter 4

An investigation was carried out to determine the taxonomic status of certain Sicilian plants that occur on sand dunes in southern Sicily and which had previously been suggested to be a product of hybridisation between two widespread *Senecio* species. Morphological analysis confirmed that the material is morphologically intermediate to *S. glaucus* ssp. *coronopifolius* and *S. leucanthemifolius* var. *leucanthemifolius*, and also overlapped in morphology with individuals of a third species, *S. gallicus*. In contrast, an analysis of the distribution of 40 RAPD fragments provided no evidence for a hybrid origin of the Sicilian material. Instead, the results strongly suggest that this material is closely related to Tunisian *S. glaucus* ssp. *coronopifolius* and it is likely, therefore, that the morphological divergence exhibited by the material is associated with natural selection and/or genetic drift in Sicily and is not due to hybridisation. The results of a survey of cpDNA variation add support to these proposals.

The high level of cpDNA diversity recorded for the Sicilian material suggests that multiple introductions to Sicily from Tunisia have occurred, however RAPD diversity was lower in the Sicilian populations relative to the Tunisian populations, which appears to indicate a genetic bottleneck associated with long-distance dispersal. This may be explained by invoking a scenario in which intense selection following its introduction to Sicily resulted in a selective sweep of the nuclear genome causing loss

of nuclear variation. This explanation, however, is merely speculative and further analysis of genetic variation is necessary to clarify this seeming paradox.

Chapter 5

The tetraploid species *Senecio mohavensis* shows close morphological similarity to diploid *S. flavus* and a close genetic relationship was established in a previous study of RAPD variation between these two species (Comes and Abbott, 2001). However, previous analyses revealed that the two species were very different in cpDNA and ITS rDNA sequence, and that *S. mohavensis* possessed a cpDNA haplotype and ITS rDNA sequence most similar to those of *S. glaucus* (Comes and Abbott, 2001). It has been proposed, therefore that *Senecio mohavensis* could be the autotetraploid of *S. flavus* introgressed with the cpDNA and ITS repeats of *S. glaucus* or a closely related diploid, or alternatively, that it is an allotetraploid (Comes and Abbott, 2001; Coleman *et al.*, 2003). If the latter were the case, then it is presumed that evolution towards the morphology of *S. flavus* may have followed its origin. To investigate these alternative hypotheses, DNA sequence variation in a portion of two nuclear genes (*PgiC* and *Scyc2*) was surveyed in *S. mohavensis*, *S. flavus* and several other members of sect. *Senecio*. The presence of two distinct copies of both genes in *S. mohavensis* indicates an allotetraploid origin, and phylogenetic analysis suggests that the two parent species were *S. flavus* and *S. glaucus*. Morphological evolution towards a *S. flavus*-like morphology is unlikely to be due to adaptation to the same environmental conditions in the wild because the geographic distributions of *S. flavus* and *S. mohavensis* are largely non-coincident. Instead, it is possible that the expression of the genes inherited from the two parents is unequal with a bias towards expression of the genes inherited from *S. flavus*. The mechanisms by which polyploid genomes evolve are currently of great interest and subject to investigation in a range of species (reviewed in Soltis and Soltis, 2000; Comai, 2000; Wendel, 2000; Liu and Wendel, 2003; Osborn *et al.*, 2003). Further investigation of *S. mohavensis* would seem to be worthwhile, in that it is likely to provide an important insight into some of the processes involved in polyploid speciation, especially the modification of gene expression that results in unequal expression of parent genomes.

Chapter 6

In the final study undertaken, further analysis of DNA sequence variation in portions of the two nuclear genes, *PgiC* and *Scyc2*, was surveyed in several more diploid and polyploid species of sect. *Senecio*, with the aim of ascertaining the origin of the hexaploid *S. teneriffae*. In the event, several additional findings of interest were obtained from the analysis.

Firstly, the diploid hybrid species *S. squalidus* was shown to be polymorphic for two very different *Scyc2* DNA sequences. One of the sequences was closely related to that of the hypothesised parent species *S. chrysanthemifolius*, and the close relationship between these species was confirmed by phylogenetic analysis of *PgiC*. However, the second copy of the *Scyc2* sequence obtained from *S. squalidus* showed an unusual pattern of nucleotide variation. The latter half of the sequence (~360 bp) was identical to the second half of the first sequence obtained from *S. squalidus* and also to that part of the sequence possessed by *S. chrysanthemifolius*. However, the first half of the sequence (~350 bp) differed by ten nucleotide substitutions from the first sequence found in *S. squalidus* and was not found to be similar to the sequences of any of the other species studied. It was concluded that this unusual sequence was unlikely to be a PCR-mediated artefact as the same sequence was recovered in several different individuals of *S. squalidus* and was also found in the radiate variant of *S. vulgaris* (var. *hibernicus*). This allele may be a result of intragenic recombination with a closely related gene in *S. squalidus*, or a *Scyc2* sequence present in the parents that was not recovered in this analysis, however further investigation into the origin of this sequence is necessary.

One of the two copies of *Scyc2* obtained from *S. vulgaris* var. *hibernicus* was identical to the unusual *Scyc2* sequence obtained from *S. squalidus*. Previously it has been hypothesised that *S. vulgaris* var. *hibernicus* contains introgressed DNA from *S. squalidus* and this result supports this hypothesis; however, until the origin of this unusual allele in *S. squalidus* is confirmed, we cannot be sure that this allele did not arise in *S. vulgaris* and has introgressed into *S. squalidus*.

Non-radiate *S. vulgaris* (which is not thought to be of introgressive origin) was shown to possess two distinct copies of *Scyc2* and *PgiC* supporting the hypothesis that *S. vulgaris* is an allotetraploid. The *Scyc2* tree, which is poorly resolved except at the tips of the branches, shows a close relationship between one copy of *Scyc2* obtained

from *S. vulgaris* and the sequence obtained from *S. leucanthemifolius*, supported by 58% bootstrap support. The second copy of *Scyc2* from *S. vulgaris* is not closely related to the sequences obtained from any of the other diploid species. The *PgiC* analysis reveals a close similarity between one of the copies of *Scyc2* from *S. vulgaris* and the sequence obtained from *S. vernalis*, a species which has previously been hypothesised as a parent of *S. vulgaris*. The second copy of *PgiC* is found in a clade which contains *S. leucanthemifolius*. *Senecio leucanthemifolius* may not have been the diploid parent; however a species which gave rise to *S. leucanthemifolius* also appears to have given rise to the diploid species that was the second parent of *S. vulgaris*. It is feasible that *S. leucanthemifolius* was polymorphic for *PgiC* sequence and one of the variants (now extinct) was the parent of *S. vulgaris*.

Phylogenetic analysis of the DNA sequences obtained from hexaploid *S. teneriffae* and the closely related *S. cambrensis* revealed that *S. vulgaris* was the tetraploid parent of both. *Senecio squalidus* was confirmed as the diploid parent of *S. cambrensis*, whereas the diploid parent of *S. teneriffae* was shown to be possibly *S. leucanthemifolius*, although this is only supported by the *Scyc2* sequence variation. The *Scyc2* and *PgiC* sequence analysis do not support the role of *S. glaucus* in the origin of *S. teneriffae* as proposed by Lowe and Abbott (1996).

Further DNA sequence analysis would be desirable to confirm some of the findings above. Firstly, the origin of the unusual *Scyc2* sequence in *S. squalidus* requires further investigation. This sequence could be present in one or both of its parent taxa (*S. aethnensis* and *S. chrysanthemifolius*), or could be related to another *cycloidea* homologue in *S. squalidus*. In addition, the parentage of *S. vulgaris* and *S. teneriffae* could be investigated further in an analysis of other single-copy nuclear genes in these individuals, or by DNA sequence analysis of other individuals of these species.

REFERENCES

ABBOTT, R. J. 1992. Plant invasions, interspecific hybridisation and the evolution of new plant taxa. *Trends in Ecology and Evolution* 7: 401-405.

ABBOTT, R. J. 2003. Sex, sunflowers and speciation. *Science* 301: 1189-1190.

ABBOTT, R. J., P. A. ASHTON, AND D. G. FORBES. 1992. Introgressive origin of the radiate groundsel, *Senecio vulgaris* L. var. *hibernicus* Syme: *Aat-3* evidence. *Heredity* 68: 425-435.

ABBOTT, R. J., D. J. CURNOW, AND J. A. IRWIN. 1995. Molecular systematics of *Senecio squalidus* L. and its close diploid relatives. *In* D. J. N. Hind, C. Jeffrey and G. V. Pope, eds. Advances in Compositae Systematics, pp 223-237. Royal Botanic Gardens, Kew.

ABBOTT, R. J., R. INGRAM, AND H. J. NOLTIE. 1983. Discovery of *Senecio cambrensis* Rosser in Edinburgh. *Watsonia* 14: 407-408.

ABBOTT, R. J., J. K. JAMES, D. G. FORBES, AND H.-P. COMES. 2002. Hybrid origin of the Oxford ragwort, *Senecio squalidus* L.: morphological and allozyme differences between *S. squalidus* and *S. rupestris* Waldst. and Kit. *Watsonia* 24: 17-29.

ABBOTT, R. J., J. K. JAMES, J. A. IRWIN, AND H.-P. COMES. 2000. Hybrid origin of the Oxford ragwort, *Senecio squalidus* L. *Watsonia* 23: 123-138.

ABBOTT, R. J., J. K. JAMES, R. I. MILNE, AND A. C. M. GILLIES. 2003. Plant introductions, hybridization and gene flow. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 358: 1123-1132.

ABBOTT, R. J., AND A. J. LOWE. 1996. A review of hybridisation and evolution in British *Senecio*. *In* D. J. N. Hind and H. J. Beentje, eds. Compositae: Systematics. Proceedings of the International Compositae Conference, Kew, 1994, pp 679-689. Royal Botanic Gardens, Kew, London, UK.

ABBOTT, R. J., AND A. J. LOWE. 2004. Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biological Journal of the Linnean Society* 82: (in press).

ADAMS, K. L., R. CRONN, R. PERCIFIELD, AND J. F. WENDEL. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-

specific reciprocal silencing. *Proceedings of the National Academy of Sciences of the USA* 100: 4649-4654.

ADAMS, R. P., AND L. H. RIESEBERG. 1998. The effects of non-homology in RAPD bands on similarity and multivariate statistical ordination in *Brassica* and *Helianthus*. *Theoretical and Applied Genetics* 97: 323-326.

ALARCÓN, R., AND D. R. CAMPBELL. 2000. Absence of conspecific pollen advantage in the dynamics of an *Ipomopsis* (Polemoniaceae) hybrid zone. *American Journal of Botany* 87: 819-824.

ALEXANDER, J. C. M. 1979. The Mediterranean species of *Senecio* sections *Senecio* and *Delphinifolius*. *Notes from the Royal Botanic Garden Edinburgh* 37: 387-428.

ALLAN, G. J., C. CLARK, AND L. H. RIESEBERG. 1997. Distribution of parental DNA markers in *Encelia virginensis* (Asteraceae: Heliantheae), a diploid species of putative hybrid origin. *Plant Systematics and Evolution* 205: 205-221.

ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAFFER, J. H. ZHANG, Z. ZHANG, W. MILLER, AND D. J. LIPMAN. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389-3402.

ALVAREZ, I., AND J. F. WENDEL. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* 29: 417-434.

ANDERSON, E. 1949. Introgressive Hybridisation. John Wiley and Sons, New York.

ANDERSON, E., AND L. HUBRICHT. 1938. Hybridisation in *Tradescantia*. III. The evidence for introgressive hybridisation. *American Journal of Botany* 25: 396-402.

ANDERSON, E., AND G. L. STEBBINS. 1954. Hybridisation as an evolutionary stimulus. *Evolution* 8: 378-388.

ANTTILA, C. K., C. C. DAEHLER, E. N. RANK, AND D. R. STRONG. 1998. Greater male fitness of a rare invader (*Spartina alterniflora*, Poaceae) threatens a common native (*Spartina foliosa*) with hybridisation. *American Journal of Botany* 85: 1597-1601.

ARNOLD, M. L. 1992. Natural hybridisation as an evolutionary process. *Annual Review of Ecology and Systematics* 23: 237-261.

ARNOLD, M. L. 1993. *Iris nelsonii* (Iridaceae): Origin and genetic composition of a homoploid hybrid species. *American Journal of Botany* 80: 577-583.

ARNOLD, M. L. 1997. Natural Hybridisation and Evolution. Oxford University Press, Oxford, UK.

ARNOLD, M. L., B. D. BENNETT, AND E. A. ZIMMER. 1990a. Natural hybridisation between *Iris fulva* and *Iris hexagona*: Patterns of ribosomal DNA variation. *Evolution* 44: 1512-1521.

ARNOLD, M. L., C. M. BUCKNER, AND J. J. ROBINSON. 1991. Pollen-mediated introgression and hybrid speciation in Louisiana irises. *Proceedings of the National Academy of Sciences of the USA* 88: 1398-1402.

ARNOLD, M. L., J. L. HAMRICK, AND B. D. BENNETT. 1990b. Allozyme variation in Louisiana irises: a test for introgression and hybrid speciation. *Heredity* 84: 297-306.

ARNOLD, M. L., J. L. HAMRICK, AND B. D. BENNETT. 1993. Interspecific pollen competition and reproductive isolation in *Iris*. *Journal of Heredity* 84: 13-16.

ARNOLD, M. L., AND S. A. HODGES. 1995. Are natural hybrids fit or unfit relative to their parents? *Trends in Ecology and Evolution* 10: 67-71.

ASHTON, P. A., AND R. J. ABBOTT. 1992a. Multiple origins and genetic diversity in the newly arisen allopolyploid species, *Senecio cambrensis* Rosser (Compositae). *Heredity* 68: 25-32.

ASHTON, P. A., AND R. J. ABBOTT. 1992b. Isozyme evidence and the origin of *Senecio vulgaris* (Compositae). *Plant Systematics and Evolution* 179: 167-174.

AYRES, D. R., AND D. R. STRONG. 2001. Origin and genetic diversity of *Spartina anglica* (Poaceae) using nuclear DNA markers. *American Journal of Botany* 88: 1863-1867.

BALDWIN, B. G., M. J. SANDERSON, J. M. PORTER, M. F. WOJCIECHOWSKI, C. S. CAMPBELL, AND M. J. DONOGHUE. 1995. The ITS region of nuclear ribosomal DNA - a valuable source of evidence on Angiosperm phylogeny. *Annals of the Missouri Botanical Garden* 82: 247-277.

BARRIER, M., B. G. BALDWIN, R. H. ROBICHAUX, AND M. D. PURUGGANAN. 1999. Interspecific hybrid ancestry of a plant adaptive radiation: Allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Molecular Biology and Evolution* 16: 1105-1113.

BARTON, N. H., AND G. M. HEWITT. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* 16: 113-148.

BARTON, N. H., AND G. M. HEWITT. 1989. Adaptation, speciation and hybrid zones. *Nature* 341: 497-503.

BAUMEL, A., M. AINOUCHE, R. KALENDAR, AND A. H. SCHULMAN. 2002. Retrotransposons and genomic stability in populations of the young allopolyploid species *Spartina anglica* C.E. Hubbard (Poaceae). *Molecular Biology and Evolution* 19: 1218-1227.

BAUMEL, A., M. L. AINOUCHE, AND J. E. LEVASSEUR. 2001. Molecular investigations in populations of *Spartina anglica* C.E. Hubbard (Poaceae) invading coastal Brittany (France). *Molecular Ecology* 10: 1689-1701.

BLEEKER, W., M. HUTHMANN, AND H. HURKA. 1999. Evolution of hybrid taxa in *Nasturtium* R. Br. (Brassicaceae). *Folia Geobotanica* 34: 421-433.

BRADLEY, R. D., AND D. M. HILLIS. 1997. Recombinant DNA sequences generated by PCR amplification. *Molecular Biology and Evolution* 14: 592-593.

BRADSHAW, H. D., K. G. OTTO, B. E. FREWEN, J. K. MCKAY, AND D. W. SCHEMSKE. 1998. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics* 149: 367-382.

BRADSHAW, H. D., AND D. W. SCHEMSKE. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426: 176-178.

BRADSHAW, H. D., S. M. WILBERT, K. G. OTTO, AND D. W. SCHEMSKE. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature* 376: 762-765.

BROCHMANN, C., Q. Y. XIANG, S. J. BRUNSFELD, D. E. SOLTIS, AND P. S. SOLTIS. 1998. Molecular evidence for polyploid origins in *Saxifraga* (Saxifragaceae): The narrow arctic endemic *S. svalbardensis* and its widespread allies. *American Journal of Botany* 85: 135-143.

BRZUSTOWSKI, J. 2002. Doh assignment test calculator, http://www2.biology.ualberta.ca/jbrzusto/Doh.php.

BUERKLE, C. A., R. J. MORRIS, M. A. ASMUSSEN, AND L. H. RIESEBERG. 2000. The likelihood of homoploid hybrid speciation. *Heredity* 84: 441-451.

BUERKLE, C. A., D. E. WOLF, AND L. H. RIESEBERG. 2003. The origin and extinction of species through hybridisation. *In* C. A. Brigham and M. W. Schwarz, eds. Population viability in plants, pp 117-141. Springer-Verlag, Berlin Heidelberg.

BURKE, J. M., AND M. L. ARNOLD. 2001. Genetics and the fitness of hybrids. *Annual Review of Genetics* 35: 31-52.

BURKE, J. M., T. J. VOSS, AND M. L. ARNOLD. 1998. Genetic interactions and natural selection in Louisiana iris hybrids. *Evolution* 52: 1304-1310.

BUSSELL, J. D. 1999. The distribution of random amplified polymorphic DNA (RAPD) diversity amongst populations of *Isotoma petraea* (Lobeliaceae). *Molecular Ecology* 8: 775-789.

CALDWELL, M. M., AND R. ROBBERECHT. 1980. A steep latitudinal gradient of solar ultraviolet-B radiation in the arctic-alpine life zone. *Ecology* 61: 600-611.

CAMPBELL, D. R., R. ALARCÓN, AND C. A. WU. 2003. Reproductive isolation and hybrid pollen disadvantage in *Ipomopsis*. *Journal of Evolutionary Biology* 16: 536-540.

CARNEY, S. E., AND M. L. ARNOLD. 1997. Differences in pollen-tube growth rate and reproductive isolation between Louisiana irises. *Journal of Heredity* 88: 545-549.

CARNEY, S. E., M. B. CRUZAN, AND M. L. ARNOLD. 1994. Reproductive interactions between hybridising irises: analysis of pollen-tube growth and fertilization success. *American Journal of Botany* 81: 1169-1175.

CEDRONI, M. L., R. C. CRONN, K. L. ADAMS, T. A. WILKINS, AND J. F. WENDEL. 2003. Evolution and expression of MYB genes in diploid and polyploid cotton. *Plant Molecular Biology* 51: 313-325.

CHALMERS, K. J., R. WAUGH, J. I. SPRENT, A. J. SIMONS, AND W. POWELL. 1992. Detection of genetic variation between and within populations of *Gliricidia sepium* and *G. maculata* using RAPD markers. *Heredity* 69: 465-472.

CHARLESWORTH, D. 1995. Evolution under the microscope. *Current Biology* 5: 935-836.

CHASE, V. C., AND P. H. RAVEN. 1975. Evolutionary and ecological relationships between *Aquilegia formosa* and *A. pubescens* (Ranunculaceae), two perennial plants. *Evolution* 29: 474-486.

CHEN, M., P. SANMIGUEL, A. C. DE OLIVEIRA, S.-S. WOO, H. ZHANG, R. A. WING, AND J. L. BENNETZEN. 1997. Microlinearity in *sh2*-homologous regions of the maize, rice and sorghum genomes. *Proceedings of the National Academy of Sciences of the USA* 94: 3431-3435.

CHORNESKY, E. A., AND J. M. RANDALL. 2003. The threat of invasive alien species to biological diversity: Setting a future course. *Annals of the Missouri Botanical Garden* 90: 67-76.

COLEMAN, M. 2003. Studies in taxonomy, biogeography and evolution of *Senecio* sect. *Senecio* (Compositae). PhD thesis, University of St. Andrews, St. Andrews, UK.

COLEMAN, M., AND R. J. ABBOTT. 2003. Possible causes of morphological variation in an endemic Moroccan groundsel (*Senecio leucanthemifolius* var. *casablancae*): evidence from chloroplast DNA and random amplified polymorphic DNA markers. *Molecular Ecology* 12: 423-434.

COLEMAN, M., D. G. FORBES, AND R. J. ABBOTT. 2001. A new subspecies of *Senecio mohavensis* (Compositae) reveals Old-New World species disjunction. *Edinburgh Journal of Botany* 58: 389-403.

COLEMAN, M., A. LISTON, J. W. KADEREIT, AND R. J. ABBOTT. 2003. Repeat intercontinental dispersal and Pleistocene speciation in disjunct Mediterranean and desert *Senecio* (Asteraceae). *American Journal of Botany* 90: 1446-1454.

COMAI, L. 2000. Genetic and epigenetic interactions in allopolyploid plants. *Plant Molecular Biology* 43: 387-399.

COMAI, L., A. MADLUNG, C. JOSEFSSON, AND A. TYAGI. 2003. Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids? *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 358: 1149-1155.

COMES, H. P., AND R. J. ABBOTT. 1998. The relative importance of historical events and gene flow on the population structure of a Mediterranean ragwort, *Senecio gallicus* (Asteraceae). *Evolution* 52: 355-367.

COMES, H. P., AND R. J. ABBOTT. 1999. Population genetic structure and gene flow across arid versus mesic environments: A comparative study of two parapatric *Senecio* species from the Near East. *Evolution* 53: 36-54.

COMES, H. P., AND R. J. ABBOTT. 2001. Molecular phylogeography, reticulation and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). *Evolution* 55: 1943-1962.

COMES, H. P., AND J. W. KADEREIT. 1990. Aspects of hybridisation between the closely related *Senecio vulgaris* L. and *Senecio vernalis* Waldst. & Kit. *Flora* 184: 381-388.

COMES, H. P., J. W. KADEREIT, A. POHL, AND R. J. ABBOTT. 1997. Chloroplast DNA and isozyme evidence on the evolution of *Senecio vulgaris* (Asteraceae). *Plant Systematics and Evolution* 206: 375-392.

COOK, L. M., P. S. SOLTIS, S. J. BRUNSFELD, AND D. E. SOLTIS. 1998. Multiple independent formations of *Tragopogon* tetraploids (Asteraceae): evidence from RAPD markers. *Molecular Ecology* 7: 1293-1302.

COYNE, J. A. 1996. Speciation in action. *Science* 272: 700-701.

CRAWFORD, D. J., S. BRAUMER, M. B. COSNER, AND T. F. STUESSY. 1993. Use of RAPD markers to document the origin of the intergeneric hybrid x *Margyracaena skottsbergii* (Rosaceae) on the Juan Fernandez Islands. *American Journal of Botany* 80: 89-92.

CRISP, P. 1972. Cytotaxonomic studies in the Section *Annui* of *Senecio*. Ph.D. thesis, University of London, UK

CRONN, R., M. CEDRONI, T. HASELKORN, C. GROVER, AND J. F. WENDEL. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* 104: 482-489.

CRONN, R. C., R. L. SMALL, AND J. F. WENDEL. 1999. Duplicated genes evolve independently after polyploid formation in cotton. *Proceedings of the National Academy of Sciences of the USA* 96: 14406-14411.

CRUZAN, M. B., AND M. L. ARNOLD. 1993. Ecological and genetic associations in an *Iris* hybrid zone. *Evolution* 47: 1432-1445.

CRUZAN, M. B., AND M. L. ARNOLD. 1994. Assortative mating and natural selection in an *Iris* hybrid zone. *Evolution* 48: 1946-1958.

CUBAS, P. 2002. Role of TCP genes in the evolution of morphological characters in angiosperms. *In* Q. C. B. Cronk, R. M. Bateman and J. A. Hawkins, eds. Developmental genetics and plant evolution, pp 247-266. Taylor & Francis, London, UK.

CUBAS, P., N. LAUTER, J. DOEBLEY, AND E. S. COEN. 1999. The TCP domain: a motif found in proteins regulating plant growth and development. *The Plant Journal* 18: 215-222.

CUBAS, P., C. VINCENT, AND E. COEN. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401: 157-161.

DARWIN, C. 1859. On the origin of species by means of natural selection. John Murray

DAWSON, I. K., A. J. SIMONS, R. WAUGH, AND W. POWELL. 1995. Diversity and genetic differentiation among subpopulations of *Gliricidia sepium* revealed by PCR-based assays. *Heredity* 74: 10-18.

DE VINCENTE, M. C., AND S. D. TANKSLEY. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134: 585-596.

DIAZ, A., AND M. R. MACNAIR. 1999. Pollen tube competition as a mechanism of prezygotic reproductive isolation between *Mimulus nasutus* and its presumed progenitor *M. guttatus*. *New Phytologist* 144: 471-478.

DOBZHANSKY, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

DOBZHANSKY, T. 1940. Speciation as a stage in evolutionary divergence. *American Naturalist* 74: 312-321.

DOEBLEY, J., A. STEC, AND L. HUBBARD. 1997. The evolution of apical dominance in maize. *Nature* 386: 485-488.

DORADO, O., L. H. RIESEBERG, AND D. ARIAS. 1992. Chloroplast DNA and introgression in southern California sunflowers. *Evolution* 46: 566-572.

DOWLING, T. E., AND C. L. SECOR. 1997. The role of hybridisation and introgression in the diversification of animals. *Annual Review of Ecology and Systematics* 28: 593-619.

DOYLE, J. J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Systematic Botany* 17: 144-163.

DOYLE, J. J., AND J. L. DOYLE. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15

DOYLE, J. J., J. L. DOYLE, AND A. H. D. BROWN. 1999. Origins, colonisation, and lineage recombination in a widespread perennial soybean polyploid complex. *Proceedings of the National Academy of Sciences of the USA* 96: 10741-10745.

DOYLE, J. J., J. L. DOYLE, A. H. D. BROWN, AND R. G. PALMER. 2002. Genomes, multiple origins, and lineage recombination in the *Glycine tomentella* (Leguminosae) polyploid complex: Histone H3- D gene sequences. *Evolution* 56: 1388-1402.

DOYLE, J. J., J. L. DOYLE, J. T. RAUSCHER, AND A. H. D. BROWN. 2004. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytologist* 161: 121-132.

EMMS, S. K., AND M. L. ARNOLD. 2000. Site-to-site differences in pollinator visitation patterns in a Louisiana iris hybrid zone. *Oikos* 91: 568-578.

EMMS, S. K., S. A. HODGES, AND M. L. ARNOLD. 1996. Pollen tube competition, siring success, and consistent asymmetric hybridisation in Louisiana irises. *Evolution* 50: 2201-2206.

ENDLER, J. A. 1973. Gene flow and population differentiation. *Science* 179: 243-250.

EXCOFFIER, L., P. E. SMOUSE, AND J. M. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

FELINER, G. N., B. G. LARENA, AND J. F. AGUILAR. 2004. Fine-scale geographical structure, intra-individual polymorphism and recombination in nuclear ribosomal internal transcribed spacers in *Armeria* (Plumbaginaceae). *Annals of Botany* 93: 189-200.

FERGUSON, C. J., AND R. K. JANSEN. 2002. A chloroplast DNA phylogeny of eastern *Phlox* (Polemoniaceae): Implications of congruence and incongruence with the ITS phylogeny. *American Journal of Botany* 89: 1324-1335.

FERGUSON, D., AND T. SANG. 2001. Speciation through homoploid hybridisation between allotetraploids in peonies (*Paeonia*). *Proceedings of the National Academy of Sciences of the USA* 98: 3915-3919.

FITTER, A. H., AND R. K. M. HAY. 1987. Environmental physiology of plants. Academic Press, London, UK.

FORD, V. S., AND L. D. GOTTLIEB. 1999. Molecular characterization of PgiC in a tetraploid plant and its diploid relatives. *Evolution* 53: 1060-1067.

FOWLER, N. L., AND D. A. LEVIN. 1984. Ecological constraints on the establishment of a novel polyploid in competition with its diploid progenitor. *American Naturalist* 124: 703-711.

FRANZKE, A., AND K. MUMMENHOFF. 1999. Recent hybrid speciation in *Cardamine* (Brassicaceae) - conversion of nuclear ribosomal ITS sequences in statu nascendi. *Theoretical and Applied Genetics* 98: 831-834.

GALILI, G., AND M. FELDMAN. 1984. Intergenomic suppression of endosperm protein genes in common wheat. *Canadian Journal of Genetics and Cytology* 26: 651-656.

GALILI, G., A. A. LEVY, AND M. FELDMAN. 1986. Gene dosage compensation of endosperm proteins in hexaploid wheat *Triticum aestivum*. *Proceedings of the National Academy of Sciences of the USA* 83: 6524-6528.

GALLEZ, G. P., AND L. D. GOTTLIEB. 1982. Genetic evidence for the hybrid origin of the diploid plant *Stephanomeria diegensis*. *Evolution* 36: 1158-1167.

GAUT, B. S., AND J. F. DOEBLEY. 1997. DNA sequence evidence for the segmental allopolyploid origin of maize. *Proceedings of the National Academy of Sciences of the USA* 94: 6809-6814.

GE, S., T. SANG, B.-R. LU, AND D.-Y. HONG. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences of the USA* 96: 14400-14405.

GOLDBLATT, P., J. C. MANNING, AND P. BERNHARDT. 1998. Adaptive radiation of bee-pollinated *Gladiolus* species (Iridaceae) in southern Africa. *Annals of the Missouri Botanical Garden* 85: 492-517.

GOTTLIEB, L. D., AND G. PILZ. 1976. Genetic similarity between *Gaura longiflora* and its obligatory outcrossing derivative *G. demareei*. *Systematic Botany* 1: 181-187.

GRANT, B. R., AND P. R. GRANT. 1993. Evolution of Darwin finches caused by a rare climatic event. *Proceedings of the Royal Society of London Series B-Biological Sciences* 251: 111-117.

GRANT, V. 1953. The role of hybridisation in the evolution of the leafy-stemmed gilias. *Evolution* 7: 51-64

GRANT, V. 1966a. Selection for vigor and fertility in the progeny of a highly sterile species hybrid in *Gilia*. *Genetics* 53: 757-775.

GRANT, V. 1966b. The origin of a new species of *Gilia* in a hybridisation experiment. *Genetics* 54: 1189-1199.

GRANT, V. 1981. Plant Speciation. Columbia University Press, New York.

GRANT, V. 1992. Floral isolation between ornithophilous and sphingophilous species of *Ipomopsis* and *Aquilegia*. *Proceedings of the National Academy of Sciences of the USA* 89: 11828-11831.

GRANT, V. 1994. Modes and origins of mechanical and ethological isolation in angiosperms. *Proceedings of the National Academy of Sciences of the USA* 91: 3-10.

GUBITZ, T., A. CALDWELL, AND A. HUDSON. 2003. Rapid molecular evolution of CYCLOIDEA-like genes in antirrhinum and its relatives. *Molecular Biology and Evolution* 20: 1537-1544.

GUO, M., D. DAVIS, AND J. A. BIRCHLER. 1996. Dosage effects on gene expression in a maize ploidy series. *Genetics* 142: 1349-1355.

HARDIG, T. M., S. J. BRUNSFELD, R. S. FRITZ, M. MORGAN, AND C. M. ORIAN. 2000. Morphological and molecular evidence for hybridisation and introgression in a willow (*Salix*) hybrid zone. *Molecular Ecology* 9: 9-24.

HARRISON, R. G. 1986. Pattern and process in a narrow hybrid zone. *Heredity* 56: 337-349.

HARRISON, R. G. 1993. Hybrids and hybrid zones: historical perspective. *In* R. G. Harrison, ed. Hybrid Zones and the Evolutionary Process, pp 2-12. Oxford University Press, New York.

HARRIS, S. A. 1990. Molecular systematic studies in some members of the genus *Senecio* L. (Asteraceae). PhD thesis, University of St Andrews, St Andrews, UK.

HARRIS, S. A. 1999. RAPDs in systematics - a useful methodology? *In* P. M. Hollingsworth, R. M. Bateman and R. J. Gornall, eds. Molecular Systematics and Plant Evolution, pp 211-228. Taylor & Francis, London.

HARRIS, S. A. 2002. Introduction of the Oxford ragwort, *Senecio squalidus* L. (Asteraceae), to the United Kingdom. *Watsonia* 24: 31-43.

HARRIS, S. A., AND R. INGRAM. 1992a. Molecular systematics of the genus *Senecio* L. I: Hybridisation in a British polyploid complex. *Heredity* 69: 1-10.

HARRIS, S. A., AND R. INGRAM. 1992b. Molecular systematics of the genus *Senecio* L. II: The origin of *S. vulgaris*. *Heredity* 69: 112-121.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 21: 160-174.

HAUSER, T. P., R. B. JORGENSEN, AND H. OSTERGÅRD. 1997. Preferential exclusion of hybrids in mixed populations between oilseed rape (*Brassica napus*) and weedy *B. campestris* (Brassicaceae). *American Journal of Botany* 84: 756-762.

HEISER, C. B. 1947. Hybridisation between the sunflower species *Helianthus annuus* and *H. petiolaris*. *Evolution* 1: 249-262.

HEISER, C. B. 1949. Natural hybridisation with particular reference to introgression. *Botanical Review* 15: 645-687.

HEY, J. 2001. The mind of the species problem. *Trends in Ecology and Evolution* 16: 326-329.

HILEMAN, L. C., AND D. A. BAUM. 2003. Why do paralogs persist? Molecular evolution of CYCLOIDEA and related floral symmetry genes in antirrhineae (Veronicaceae). *Molecular Biology and Evolution* 20: 591-600.

HILLIS, D. M., AND M. T. DIXON. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. *Quarterly Review of Biology* 66: 411-453.

HILLIS, D. M., C. MORITZ, C. A. PORTER, AND R. J. BAKER. 1991. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science* 251: 308-310.

HODGES, S. A., J. M. BURKE, AND M. L. ARNOLD. 1996. Natural formation of Iris hybrids: Experimental evidence on the establishment of hybrid zones. *Evolution* 50: 2504-2509.

HUFF, D. R., R. PEAKALL, AND P. E. SMOUSE. 1993. RAPD variation within and among populations of outcrossing buffalograss (*Buchloë dactyloides* (Nutt.) Engelm.). *Theoretical and Applied Genetics* 86: 927-934.

INGRAM, A. L., AND J. J. DOYLE. 2003. The origin and evolution of *Eragrostis tef* (Poaceae) and related polyploids: Evidence from nuclear *waxy* and plastid *rps16*. *American Journal of Botany* 90: 116-122.

INGRAM, R., J. WEIR, AND R. J. ABBOTT. 1980. New evidence concerning the origin of inland radiate groundsel, *S. vulgaris* L. var. *hibernicus* Syme. *New Phytologist* 84: 543-546.

IRWIN, J. A., AND R. J. ABBOTT. 1992. Morphometric and isozyme evidence for the hybrid origin of a new tetraploid radiate groundsel in York, England. *Heredity* 69: 431-439.

JACKSON, H. D., D. A. STEANE, B. M. POTTS, AND R. E. VAILLANCOURT. 1999. Chloroplast DNA evidence for reticulate evolution in *Eucalyptus* (Myrtaceae). *Molecular Ecology* 8: 739-751.

JAMES, J. K. 1999. Molecular analysis of the hybrid origin of the Oxford ragwort, *Senecio squalidus* L., and related studies. Ph.D. thesis, University of St. Andrews, St. Andrews, UK.

JANSEN, R. K., AND J. D. PALMER. 1987. Chloroplast DNA from lettuce and *Barnadesia* (Asteraceae): structure, gene localisation and characterisation of a large inversion. *Current Genetics* 11: 553-564.

JIANG, C.-X., P. W. CHEE, X. DRAYE, P. L. MORRELL, C. W. SMITH, AND A. H. PATERSON. 2000. Multilocus interactions restrict gene introgression in interspecific populations of polyploid *Gossypium* (cotton). *Evolution* 54: 798-814.

JOBST, J., K. KING, AND V. HEMLEBEN. 1998. Molecular evolution of the internal transcribed spacers (ITS1 and ITS2) and phylogenetic relationships among species of the family Cucurbitaceae. *Molecular Phylogenetics and Evolution* 9: 204-219.

JORGENSEN, R. B., T. HAUSER, T. R. MIKKELSEN, AND H. OSTERGÅRD. 1996. Transfer of engineered genes from crop to wild plants. *Trends in Plant Science* 1: 356-358.

KADEREIT, J. W. 1984a. Some notes on annual species of *Senecio* L. (Asteraceae) from northern Africa and the Canary Islands. *Botanische Jahrbuch der Systematischen* 104: 509-517.

KADEREIT, J. W. 1984b. The origin of *Senecio vulgaris*. *Plant Systematics and Evolution* 145: 135-153.

KELLIS, M., B. W. BIRREN, AND E. S. LANDER. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428: 617-624.

KIANG, Y. T., AND J. L. HAMRICK. 1978. Reproductive isolation in the *Mimulus guttatus-M. nasutus* complex. *American Midland Naturalist* 100: 269-276.

KING, L. M., AND B. A. SCHAAL. 1989. Ribosomal-DNA variation and distribution in *Rudbeckia missouriensis*. *Evolution* 43: 1117-1119.

KING, R. A. 1994. Hybridisation and ecotypic differentiation in southern European *Senecio* (Asteraceae): a preliminary study. BSc thesis, University of St. Andrews, St. Andrews, UK.

KLIER, K., M. J. LEOSCHKE, AND J. F. WENDEL. 1991. Hybridization and introgression in White and Yellow Ladyslipper Orchids (*Cypripedium candidum* and *C. pubescens*). *Journal of Heredity* 82: 305-318.

KLIPS, R. A. 1999. Pollen competition as a reproductive isolating mechanism between two sympatric *Hibiscus* species (Malvaceae). *American Journal of Botany* 86: 269-272.

KOCHERT, G., H. T. STALKER, M. GIMENES, L. GALGARO, C. R. LOPES, AND K. MOORE. 1996. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). *American Journal of Botany* 83: 1282-1291.

LEE, J. Y., K. MUMMENHOFF, AND J. L. BOWMAN. 2002. Allopolyploidization and evolution of species with reduced floral structures in *Lepidium* L. (Brassicaceae). *Proceedings of the National Academy of Sciences of the USA* 99: 16835-16840.

LEITCH, I. J., AND M. D. BENNETT. 1997. Polyploidy in angiosperms. *Trends in Plant Science* 2: 470-476.

LEVIN, D. A. 1983. Polyploidy and novelty in flowering plants. *American Naturalist* 122: 1-25.

LEVIN, D. A. 2002. The Role of Chromosomal Change in Plant Evolution. Oxford University Press Inc, U.S.

LEVIN, D. A., J. FRANCISCO-ORTEGA, AND R. K. JANSEN. 1996. Hybridization and the extinction of rare plant species. *Conservation Biology* 10: 10-16.

LEXER, C., M. E. WELCH, J. L. DURPHY, AND L. H. RIESEBERG. 2003a. Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Molecular Ecology* 12: 1225-1235.

LEXER, C., M. E. WELCH, O. RAYMOND, AND L. H. RIESEBERG. 2003b. The origin of ecological divergence in *Helianthus paradoxus* (Asteraceae): Selection on transgressive characters in a novel hybrid habitat. *Evolution* 57: 1989-2000.

LI, Z., S. R. M. PINSON, A. H. PATERSON, W. D. PARK, AND J. W. STANSEL. 1997. Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* L.) population. *Genetics* 145: 1139-1148.

LISTON, A., AND J. W. KADEREIT. 1995. Chloroplast DNA evidence for introgression and long distance dispersal in the desert annual *Senecio flavus* (Asteraceae). *Plant Systematics and Evolution* 197: 33-41.

LISTON, A., L. H. RIESEBERG, AND T. S. ELIAS. 1989. Genetic similarity is high between intercontinental disjunct species of *Senecio* (Asteraceae). *American Journal of Botany* 76: 383-388.

LIU, B., C. L. BRUBAKER, G. MERGEAI, R. C. CRONN, AND J. F. WENDEL. 2001. Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome* 44: 321-330.

LIU, B., AND J. F. WENDEL. 2003. Epigenetic phenomena and the evolution of plant allopolyploids. *Molecular Phylogenetics and Evolution* 29: 365-379.

LOPEZ, G. A., B. M. POTTS, AND P. A. TILYARD. 2000. F1 hybrid inviability in *Eucalyptus*: the case of *E. ovata* x *E. globulus*. *Heredity* 85: 242-250.

LOTSY, J. P. 1925. Species or Linneon. *Genetica* 7: 487-506.

LOWE, A. J., AND R. J. ABBOTT. 1996. Origins of the new allopolyploid species *Senecio cambrensis* (Asteraceae) and its relationship to the Canary Island endemic *Senecio teneriffae*. *American Journal of Botany* 83: 1365-1372.

LOWE, A. J., AND R. J. ABBOTT. 2000. Routes of origin of two recently evolved hybrid taxa: *Senecio vulgaris* var. *hibernicus* and York radiate groundsel. *American Journal of Botany* 87: 1159-1167.

LOWE, A. J., AND R. J. ABBOTT. 2003. A new British species of *Senecio* (Asteraceae), another hybrid derivative of *S. vulgaris* L. and *S. squalidus* L. *Watsonia* 24: 375-388.

LUO, D., R. CARPENTER, L. COPSEY, C. VINCENT, J. CLARK, AND E. S. COEN. 1999. Control of organ asymmetry in flowers of *Antirrhinum*. *Cell* 99: 367-376.

LUO, D., R. CARPENTER, C. VINCENT, L. COPSEY, AND E. COEN. 1996. Origin of floral asymmetry in *Antirrhinum*. *Nature* 383: 794-799.

LYNCH, M., AND A. G. FORCE. 2000. The origin of interspecific genomic incompatibility via gene duplication. *American Naturalist* 156: 590-605.

LYNCH, M., AND B. G. MILLIGAN. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3: 91-99.

MADLUNG, A., R. W. MASUELLI, B. WATSON, S. H. REYNOLDS, J. DAVISON, AND L. COMAI. 2002. Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant Physiology* 129: 733-746.

MALLET, J. 1995. A species definition for the modern synthesis. *Trends in Ecology and Evolution* 10: 294-299.

MANEN, J. F., M. C. BOULTER, AND Y. NACIRI-GRAVEN. 2002. The complex history of the genus *Ilex* L. (Aquifoliaceae): evidence from the comparison of plastid and nuclear DNA sequences and from fossil data. *Plant Systematics and Evolution* 235: 79-98.

MARSHALL, D. F., AND R. J. ABBOTT. 1980. On the frequency of introgression of the radiate (Tr) allele from *Senecio squalidus* L. into *Senecio vulgaris* L. *Heredity* 45: 133-135.

MARSHALL, D. F., AND R. J. ABBOTT. 1982. Polymorphism for outcrossing frequency at the ray floret locus in *Senecio vulgaris* L.: 1. Evidence. *Heredity* 48: 227-235.

MARSHALL, D. F., AND R. J. ABBOTT. 1984. Polymorphism for outcrossing frequency at the ray floret locus in *Senecio vulgaris* L.: 2. Confirmation. *Heredity* 52: 331-336.

MASON-GAMER, R. J. 2001. Origin of North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. *Systematic Botany* 26: 757-768.

MASTERSON, J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264: 421-424.

MAYR, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York.

MCCARTHY, E. M., M. A. ASMUSSEN, AND W. W. ANDERSON. 1995. A theoretical assessment of recombinational speciation. *Heredity* 74: 502-509.

MCCLINTOCK, B. 1984. The significance of responses of the genome to challenge. *Science* 226

MILNE, R. I., AND R. J. ABBOTT. 2000. Origin and evolution of invasive naturalized material of *Rhododendron ponticum* L. in the British Isles. *Molecular Ecology* 9: 541-556.

MILNE, R. I., S. TERZIOGLU, AND R. J. ABBOTT. 2003. A hybrid zone dominated by fertile F1s: maintenance of species barriers in *Rhododendron*. *Molecular Ecology* 12: 2719-2729.

MOODY, M. E., L. D. MUELLER, AND D. E. SOLTIS. 1993. Genetic variation and random drift in autotetraploid populations. *Genetics* 134: 649-657.

MOORE, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. *The Quarterly Review of Biology* 52: 263-277.

MUMMENHOFF, K., AND H. HURKA. 1995. Allopolyploid origin of *Arabidopsis suecica* (Fries) Norrlin - Evidence from chloroplast and nuclear genome markers. *Botanica Acta* 108: 449-456.

NICHOLAS, K. B., AND H. B. NICHOLAS JR. 1997. Genedoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author www.psc.edu/biomed/genedoc.

NUISMER, S. L., AND J. N. THOMPSON. 2001. Plant polyploidy and non-uniform effects on insect herbivores. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268: 1937-1940.

O'KANE, S. L., B. A. SCHAAL, AND I. A. AL SHEHBAZ. 1996. The origins of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences. *Systematic Botany* 21: 559-566.

OSBORN, T. C., J. C. PIRES, J. A. BIRCHLER, D. L. AUGER, Z. J. CHEN, H.-S. LEE, L. COMAI, A. MADLUNG, R. W. DOERGE, V. COLOT, AND R. A. MARTIENSSEN. 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics* 19: 141-147.

OTTO, S. P., AND J. WHITTON. 2000. Polyploid incidence and evolution. *Annual Review of Genetics* 34: 401-437.

OZKAN, H., A. A. LEVY, AND M. FELDMAN. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. *Plant Cell* 13: 1735-1747.

PAETKAU, D., W. CALVERT, I. STERLING, AND C. STROBECK. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* 4: 347-354.

PAIGE, K. N., W. C. CAPMAN, AND P. JENNETTEN. 1991. Mitochondrial inheritance patterns across a cottonwood hybrid zone: cytonuclear disequilibria and hybrid zone dynamics. *Evolution* 45: 1360-1369.

PARDO, C., P. CUBAS, AND H. TAHIRI. 2004. Molecular phylogeny and systematics of *Genista* (Leguminosae) and related genera based on nucleotide sequences of nrDNA (ITS region) and cpDNA (*trn*L-*trn*F intergenic spacer). *Plant Systematics and Evolution* 244: 93-119.

PEAKALL, R., AND P. E. SMOUSE. 2002. GenAlEx V5.04: Genetic Analysis in Excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia. http://www.anu.edu.au/BoZo/GenAlEx/.

PETIT, C., P. LESBROS, X. GE, AND J. D. THOMPSON. 1997. Variation in flowering phenology and selfing rate across a contact zone between diploid and tetraploid *Arrhenatherum elatius* (Poaceae). *Heredity* 79: 31-40.

POSADA, D., AND K. A. CRANDALL. 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* 14: 817-818.

PROCTOR, M. C. F. 1978. Insect pollination syndromes in an evolutionary and ecosystemic context. *In* A. J. Richards ed., The pollination of flowers by insects, pp 105-116. Academic Press, London, UK.

QUILLET, M. C., N. MADJIDIAN, T. GRIVEAU, H. SERIEYS, M. TERSAC, M. LORIEUS, AND A. BERVILLE. 1995. Mapping genetic factors controlling pollen viability in an interspecific cross in *Helianthus* section *Helianthus*. *Theoretical and Applied Genetics* 91: 1195-1202.

RAMSEY, J., H. D. BRADSHAW, AND D. W. SCHEMSKE. 2003. Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). *Evolution* 57: 1520-1534.

RAMSEY, J., AND D. W. SCHEMSKE. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics* 29: 467-591.

RAND, D. M., AND R. G. HARRISON. 1989. Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution* 43: 432-449.

RANDOLPH, L. F. 1966. *Iris nelsonii*, a new species of Louisiana iris of hybrid origin. *Baileya* 14: 143-169.

RAYBOULD, A. F., A. J. GRAY, M. J. LAWRENCE, AND D. F. MARSHALL. 1991. The evolution of *Spartina anglica* C. E. Hubbard (Gramineae): origin and genetic variability. *Biological Journal of the Linnean Society* 43: 111-126.

REEVES, P. A., AND R. G. OLMSTEAD. 2003. Evolution of the TCP gene family in Asteridae: Cladistic and network approaches to understanding regulatory gene family diversification and its impact on morphological evolution. *Molecular Biology and Evolution* 20: 1997-2009.

RHYMER, J. M., AND D. SIMBERLOFF. 1996. Extinction by hybridisation and introgression. *Annual Review of Ecology and Systematics* 27: 83-109.

RICHARDS, A. J. 1986. Plant breeding systems. George Allan and Unwin, London, UK.

RIDDLE, N. C., AND J. A. BIRCHLER. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. *Trends in Genetics* 19: 597-600.

RIDLEY, M. 1993. Evolution. Blackwell Scientific, Inc., Cambridge, Massachusetts.

RIESEBERG, L. H. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *American Journal of Botany* 78: 1218-1237.

RIESEBERG, L. H. 1996. Homology among RAPD fragments in interspecific comparisons. *Molecular Ecology* 5: 99-105.

RIESEBERG, L. H. 1997. Hybrid origins of plant species. *Annual Review of Ecology and Systematics* 28: 359-389.

RIESEBERG, L. H. 2000. Crossing relationships among ancient and experimental sunflower hybrid lineages. *Evolution* 54: 859-865.

RIESEBERG, L. H. 2001. Polyploid evolution: Keeping the peace at genomic reunions. *Current Biology* 11: R925-R928.

RIESEBERG, L. H., M. A. ARCHER, AND R. K. WAYNE. 1999a. Transgressive segregation, adaptation and speciation. *Heredity* 83: 363-372.

RIESEBERG, L. H., D. M. ARIAS, M. UNGERER, C. R. LINDER, AND B. SINERVO. 1996a. The effects of mating design on introgression between chromosomally divergent sunflower species. *Theoretical and Applied Genetics* 93: 633-644.

RIESEBERG, L. H., S. J. E. BAIRD, AND K. A. GARDNER. 2000. Hybridisation, introgression and linkage evolution. *Plant Molecular Biology* 42: 205-224.

RIESEBERG, L. H., S. BECKSTROM-STERNBERG, AND K. DOAN. 1990. *Helianthus annuus* ssp. *texanus* has chloroplast DNA and nuclear ribosomal RNA genes of *Helianthus debilis* ssp. *cucumerifolius*. *Proceedings of the National Academy of Sciences of the USA* 87: 593-597.

RIESEBERG, L. H., S. M. BECKSTROM-STERNBERG, A. LISTON, AND D. M. ARIAS. 1991a. Phylogenetic and systematic inferences from chloroplast DNA and isozyme variation in *Helianthus* sect. Helianthus (Asteraceae). *Systematic Botany* 16: 50-76.

RIESEBERG, L. H., AND S. E. CARNEY. 1998. Tansley review no. 102: Plant hybridisation. *New Phytologist* 140: 599-624.

RIESEBERG, L. H., H. C. CHOI, AND D. HAM. 1991b. Differential cytoplasmic versus nuclear introgression in *Helianthus*. *Journal of Heredity* 82: 489-493.

RIESEBERG, L. H., A. M. DESROCHERS, AND S. J. YOUN. 1995a. Interspecific pollen competition as a reproductive barrier between sympatric species of *Helianthus* (Asteraceae). *American Journal of Botany* 82: 515-519.

RIESEBERG, L. H., AND N. C. ELLSTRAND. 1993. What can morphological and molecular markers tell us about plant hybridisation? *Critical Reviews in Plant Sciences* 12: 213-241.

RIESEBERG, L. H., AND D. GERBER. 1995. Hybridization in the Catalina Island mountain mahogany (*Cercocarpus traskiae*) - RAPD evidence. *Conservation Biology* 9: 199-203.

RIESEBERG, L. H., C. R. LINDER, AND G. J. SEILER. 1995b. Chromosomal and genic barriers to introgression in *Helianthus*. *Genetics* 141: 1163-1171.

RIESEBERG, L. H., O. RAYMOND, D. M. ROSENTHAL, Z. LAI, K. LIVINGSTONE, T. NAKAZATO, J. L. DURPHY, A. E. SCHWARZBACH, L. A. DONOVAN, AND C. LEXER. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301: 1211-1216.

RIESEBERG, L. H., B. SINERVO, C. R. LINDER, M. C. UNGERER, AND D. M. ARIAS. 1996b. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* 272: 741-745.

RIESEBERG, L. H., C. VAN FOSSEN, AND A. M. DESROCHERS. 1995c. Hybrid speciation accompanied by genomic reorganisation in wild sunflowers. *Nature* 375: 313-316.

RIESEBERG, L. H., AND J. WENDEL. 1993. Introgression and its consequences in plants. *In* R. G. Harrison, ed. Hybrid Zones and the Evolutionary Process, pp 70-109. Oxford University Press, New York.

RIESEBERG, L. H., J. WHITTON, AND K. GARDNER. 1999b. Hybrid zones and the genetic architecture of a barrier to gene flow between two Sunflower species. *Genetics* 152: 713-727.

ROOSE, M. L., AND L. D. GOTTLIEB. 1976. Genetic and biochemical consequences of polyploidy in *Tragopogon*. *Evolution* 30: 818-830.

ROSENTHAL, D. M., A. E. SCHWARZBACH, L. A. DONOVAN, O. RAYMOND, AND L. H. RIESEBERG. 2002. Phenotypic differentiation between three ancient hybrid taxa and their parental species. *International Journal of Plant Sciences* 163: 387-398.

ROSSER, E. M. 1955. A new British species of *Senecio*. *Watsonia* 3: 228-232.

SANG, T., D. J. CRAWFORD, AND T. F. STUESSY. 1995. Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. *Proceedings of the National Academy of Sciences of the USA* 92: 6813-6817.

SANG, T., AND D. ZHANG. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. *Systematic Botany* 24: 148-163.

SCHEMSKE, D. W., AND H. D. BRADSHAW. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). *Proceedings of the National Academy of Sciences of the USA* 96: 11910-11915.

SCHMITT, J. 1980. Pollinator foraging behaviour and gene dispersal in *Senecio* (Compositae). *Evolution* 34: 934-943.

SCHNEIDER, S., D. ROESSLI, AND L. EXCOFFIER. 2000. Arlequin ver. 2.001: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. http://anthro.unige.ch/arlequin

SCHWARZBACH, A. E., L. A. DONOVAN, AND L. H. RIESEBERG. 2001. Transgressive character expression in a hybrid sunflower species. *American Journal of Botany* 88: 270-277.

SEGRAVES, K. A., AND J. N. THOMPSON. 1999. Plant polyploidy and pollination: Floral traits and insect visits to diploid and tetraploid *Heuchera grossulariifolia*. *Evolution* 53: 1114-1127.

SEGRAVES, K. A., J. N. THOMPSON, P. S. SOLTIS, AND D. E. SOLTIS. 1999. Multiple origins of polyploidy and the geographic structure of *Heuchera grossulariifolia*. *Molecular Ecology* 8: 253-262.

SEMERIKOV, V. L., H. Q. ZHANG, M. SUN, AND M. LASCOUX. 2003. Conflicting phylogenies of *Larix* (Pinaceae) based on cytoplasmic and nuclear DNA. *Molecular Phylogenetics and Evolution* 27: 173-184.

SENCHINA, D. S., I. ALVAREZ, R. C. CRONN, B. LIU, J. K. RONG, R. D. NOYES, A. H. PATERSON, R. A. WING, T. A. WILKINS, AND J. F. WENDEL. 2003. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. *Molecular Biology and Evolution* 20: 633-643.

SETOGUCHI, H., AND I. WATANABE. 2000. Intersectional gene flow between insular endemics of *Ilex* (Aquifoliaceae) on the Bonin Islands and the Ryukyu Islands. *American Journal of Botany* 87: 793-810.

SHAKED, H., K. KASHKUSH, H. OZKAN, M. FELDMAN, AND A. A. LEVY. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridisation and allopolyploidy in wheat. *Plant Cell* 13: 1749-1759.

SHANNON, C. E., AND W. WEAVER. 1949. The mathematical theory of communication. University of Illinois Press, Urbana.

SHEPARD, K. A., AND M. D. PURUGGANAN. 2002. The genetics of plant morphological evolution. *Current Opinion in Plant Biology* 5: 49-55.

SLATKIN, M. 1985. Gene flow in natural populations. *Annual Review of Ecology and Systematics* 16: 393-430.

SMALL, R. L., R. C. CRONN, AND J. F. WENDEL. 2004. Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* 17: 145-170.

SMALL, R. L., J. A. RYBURN, R. C. CRONN, T. SEELANAN, AND J. F. WENDEL. 1998. The tortoise and the hare: Choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. *American Journal of Botany* 85: 1301-1315.

SMALL, R. L., AND J. F. WENDEL. 2000. Copy number lability and evolutionary dynamics of the *Adh* gene family in diploid and tetraploid cotton (*Gossypium*). *Genetics* 155: 1913-1926.

SOLTIS, D. E., AND R. K. KUZOFF. 1995. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). *Evolution* 49: 727-742.

SOLTIS, D. E., AND P. S. SOLTIS. 1989. Allopolyploid speciation in *Tragopogon*: Insights from chloroplast DNA. *American Journal of Botany* 76: 1119-1124.

SOLTIS, D. E., AND P. S. SOLTIS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* 12: 243-273.

SOLTIS, D. E., AND P. S. SOLTIS. 1995. The dynamic nature of polyploid genomes. *Proceedings of the National Academy of Sciences of the USA* 92: 8089-8091.

SOLTIS, D. E., AND P. S. SOLTIS. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. *In* D. E. Soltis, P. S. Soltis and J. J. Doyle, eds. Molecular systematics of plants. II. DNA sequencing, pp 1-42. Kluwer Academic Publishers, Boston, MA.

SOLTIS, D. E., AND P. S. SOLTIS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology and Evolution* 14: 348-352.

SOLTIS, D. E., P. S. SOLTIS, T. G. COLLIER, AND M. L. EDGERTON. 1991. Chloroplast DNA variation within and among genera of the *Heuchera* group (Saxifragaceae) - Evidence for chloroplast transfer and paraphyly. *American Journal of Botany* 78: 1091-1112.

SOLTIS, D. E., P. S. SOLTIS, AND J. A. TATE. 2004. Advances in the study of polyploidy since Plant Speciation. *New Phytologist* 161: 173-191.

SOLTIS, P. S., G. M. PLUNKETT, S. J. NOVAK, AND D. E. SOLTIS. 1995. Genetic variation in *Tragopogon* species: additional origins of the allotetraploids *T. mirus* and *T. miscellus* (Compositae). *American Journal of Botany* 82: 1329-1341.

SOLTIS, P. S., AND D. E. SOLTIS. 2000. The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the USA* 97: 7051-7057.

SONG, K. M., P. LU, K. L. TANG, AND T. C. OSBORN. 1995. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proceedings of the National Academy of Sciences of the USA* 92: 7719-7723.

SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98: 503.

SPRING, J. 1997. Vertebrate evolution by interspecific hybridisation - Are we polyploid? *FEBS Letters* 400: 2-8.

STACE, C. A. 1989. Plant Taxonomy and Biosystematics. Cambridge University Press, Cambridge.

STEBBINS, G. L. 1950. Variation and Evolution in Plants. Columbia University Press, New York.

STEBBINS, G. L. 1959. The role of hybridisation in evolution. *Proceedings of the American Philosophical Society* 103: 231-251.

STEBBINS, G. L., AND G. K. DALY. 1961. Changes in the variation of a hybrid population of *Helianthus* over an eight year period. *Evolution* 15: 60-71.

STEWART, C. N., H. A. RICHARDS, AND M. D. HALFHILL. 2000. Transgenic plants and biosafety: Science, misconceptions and public perceptions. *Biotechniques* 29: 832.

SWOFFORD, D. L. 2001. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Sinauer, Sunderland, MA.

TADMOR, Y., D. ZAMIR, AND G. LADIZINSKY. 1987. Genetic mapping of an ancient translocation in the genus *Lens*. *Theoretical and Applied Genetics* 73: 883-892.

TECHNELYSIUM PTY LTD. 1998-2001. Chromas 2.12. www.technelysium.com.au.

TEMPLETON, A. R. 1981. Mechanisms and speciation - a population genetics approach. *Annual Review of Ecology and Systematics* 12: 23-48.

TEMPLETON, A. R. 1989. The meaning of species and speciation: a genetic perspective. *In* D. Otte and J. A. Endler, eds. Speciation and its Consequences, pp 3-27. Sinauer, Sunderland, MA.

THOMAS, B. R., V. S. FORD, E. PICHERSKY, AND L. D. GOTTLIEB. 1993. Molecular characterisation of duplicate cytosolic phosphoglucose isomerase genes in *Clarkia* and comparison to the single gene in *Arabidopsis*. *Genetics* 135: 895-905.

THOMAS, B. R., D. LAUDENCIA-CHINGCUANCO, AND L. D. GOTTLIEB. 1992. Molecular analysis of the plant gene encoding cytosolic phosphoglucose isomerase. *Plant Molecular Biology* 19: 745-757.

THOMPSON, J. D., AND R. LUMARET. 1992. The evolutionary dynamics of polyploid plants: Origins, establishment and persistence. *Trends in Ecology and Evolution* 7: 302-307.

VANRAAMSDONK, L. W. D., M. P. SMIECH, AND J. M. SANDBRINK. 1997. Introgression explains incongruence between nuclear and chloroplast DNA-based phylogenies in *Allium* section *Cepa*. *Botanical Journal of the Linnean Society* 123: 91-108.

WAGNER, W. H. J. 1983. Biosystematics and evolutionary noise. *Taxon* 19: 146-151.

WALSH, N. E., AND D. CHARLESWORTH. 1992. Evolutionary interpretations of differences in pollen tube growth rates. *The Quarterly Review of Biology* 67: 19-37.

WALTERS, S. M. 1963. *Senecio rupestris* Waldst. and Kit. and *Senecio squalidus* L. *Proceedings of the Botanical Society of the British Isles* 5: 582.

WANG, X.-R., AND A. E. SZMIDT. 1994. Hybridisation and chloroplast DNA variation in a *Pinus* complex from Asia. *Evolution* 48: 1020-1031.

WANG, X.-R., A. E. SZMIDT, A. LEWANDOWSKI, AND Z. R. WANG. 1990. Evolutionary analysis of *Pinus densata* (Masters) a putative Tertiary hybrid. I. Allozyme variation. *Theoretical and Applied Genetics* 80: 635-640.

WATSON, L. E., W. J. ELISENS, AND J. R. ESTES. 1991. Electrophoretic and cytogenetic evidence for allopolyploid origin of *Marshallia mohrii* (Asteraceae). *American Journal of Botany* 78: 408-416.

WEIR, J., AND R. INGRAM. 1980. Ray morphology and cytological investigations of *Senecio cambrensis* Rosser. *New Phytologist* 86: 237-241.

WELCH, M. E., AND L. H. RIESEBERG. 2002. Habitat divergence between a homoploid hybrid sunflower species, *Helianthus paradoxus* (Asteraceae), and its progenitors. *American Journal of Botany* 89: 472-478.

WENDEL, J. F. 2000. Genome evolution in polyploids. *Plant Molecular Biology* 42: 225-249.

WENDEL, J., AND J. J. DOYLE. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. *In* D. E. Soltis, P. S. Soltis and J. J. Doyle, eds. Molecular systematics of plants. II. DNA sequencing, pp 265-296. Kluwer Academic Publishers, Boston, MA.

WENDEL, J. F., A. SCHNABEL, AND T. SEELANAN. 1995. An unusual ribosomal DNA sequence from *Gossypium gossypioides* reveals ancient, cryptic intergenomic introgression. *Molecular Phylogenetics and Evolution* 4: 298-313.

WENDEL, J. F., J. M. STEWART, AND J. H. RETTIG. 1991. Molecular evidence for homoploid reticulate evolution among Australian species of *Gossypium*. *Evolution* 45: 694-711.

WERTH, C., AND M. WINDHAM. 1991. A model for divergent, allopatric speciation of polyploid pteridophytes resulting from silencing of duplicate-gene expression. *American Naturalist* 137: 515-526.

WHITHAM, T. G., P. A. MORROW, AND B. M. POTTS. 1991. Conservation of hybrid plants. *Science* 254: 779-780.

WILLIAMS, J. G. K., M. K. HANAFEY, J. A. RAFALSKI, AND S. V. TINGEY. 1993. Genetic analysis using random amplified polymorphic DNA markers. *Methods in Enzymology* 218: 704-740.

WILLIAMS, J. G. K., A. R. KUBELIK, K. J. LIVAK, J. A. RAFALSKI, AND S. V. TINGEY. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18: 6531-6535.

WOLF, D. E., N. TAKEBAYASHI, AND L. H. RIESEBERG. 2001. Predicting the risk of extinction through hybridization. *Conservation Biology* 15: 1039-1053.

WOLF, P. G., D. R. CAMPBELL, M. W. NICKOLAS, S. D. SIPES, T. R. TOLER, AND J. K. ARCHIBALD. 2001. Tests of pre- and postpollination barriers to hybridization between sympatric species of *Ipomopsis* (Polemoniaceae). *American Journal of Botany* 88: 213-219.

WOLFE, A. D., Q.-Y. XIANG, AND S. R. KEPHART. 1998. Diploid hybrid speciation in *Penstemon*. *Proceedings of the National Academy of Sciences of the USA* 95: 5112-5115

WOLFE, K. H., AND D. C. SCHIELDS. 1997. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 387: 708-713.

WOLFFE, A. P., AND M. A. MATZKE. 1999. Epigenetics: Regulation through repression. *Science* 286: 481-486.

WRIGHT, S. 1978. Evolution and the genetics of populations. Vol. 4. Variability within and among natural populations. University of Chicago Press, Chicago, USA.

WU, C.-I., AND M. F. PALOPOLI. 1994. Genetics of postmating reproductive isolation in animals. *Annual Review of Genetics* 27: 283-308.

YEH, F. C., D. K. X. CHONG, AND R. C. YANG. 1995. RAPD variation within and among natural populations of trembling aspen (*Populus tremuloides* Michx) from Alberta. *Journal of Heredity* 86: 454-460.

APPENDIX I – LABORATORY SOLUTIONS

Appendix Ia – Buffers used for allozyme analysis

- Lithium borate electrode buffer (per litre)

1.2 g	Lithium hydroxide
11.9 g	Boric acid
1 litre	dH_2O

pH to 8.3 with dry constituents

- Lithium borate gel buffer (per litre)

1.28 g	Citric acid (anhydrous)
5.45 g	Tris[1] base
100 ml	Electrode buffer (see above)
900 ml	dH_2O

pH to 8.3 with citric acid or 1M NaOH

- Protein extraction buffer

50 ml	Gel buffer
37 mg	KCl
10 mg	$MgCl_2$
18 mg	EDTA[2]
25 mg	PVPP[3]
0.5 ml	Triton X100
2 ml	2-mercaptoethanol

- Acp stain buffer

50 ml	0.2M acetate buffer (pH 5.0)
50 mg	α-naphthyl acid phosphate
40 mg	Fast Garnet GBC
1 ml	10% $MgCl_2$

Abbreviations

[1]Tris	tris (hydroxymethyl) methylamine
[2]EDTA	ethylene diaminetetra-acetic acid
[3]PVPP	poly (vinylpolypyrrolidone)

Appendix Ib - Buffers used for DNA extraction and preparing agarose gels

- 2 × CTAB (per l)

0.1 M	Tris[1] base
20 mM	EDTA[2]
1.4 M	NaCl
2 %	CTAB[3] (w/v)
2 %	PVP-40T[4] (w/v)

- 5 × TBE

0.45 M	Tris base
0.2 M	Boric acid
0.125 mM	EDTA

pH to 8.0

Prepare gels with a suitable volume of 0.5 × TBE and 1.0 - 1.5 % agarose

- TE

10 mM	Tris base
1 mM	EDTA

pH to 7.5

Abbreviations

[1]Tris	tris (hydroxymethyl) methylamine
[2]EDTA	ethylene diaminetetra-acetic acid
[3]CTAB	hexadecyltrimethylammonium bromide
[4]PVP-40T	poly (vinylpyrolidine)

Appendix Ic – Bacterial growth media

- Luria broth (per l)

10 g	Tryptone
5 g	Yeast extract
5 g	NaCl

pH to 7.2, autoclave

- Agar plates

250 ml	Luria broth
3.5 g	Bacteriological agar

autoclave

To prepare plates, melt in microwave, then, just before pouring, add:

800 µl	X-Gal[1] (20 mg/ml in dimethyl formamide)
100 µl	Ampicillin (100 mg/ml dH_2O; filter sterilised)

Abbreviations

[1]X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Appendix Id – Buffers used for cpDNA RFLP

- Denaturation buffer

1.5 M NaCl

0.5 M NaOH

- Neutralisation buffer

1.5 M NaCl

0.5 M Tris[1]-base

1 mM EDTA[2]

pH 7.2

- 20 × SSC

3 M NaCl

0.3 M Trisodium citrate

- DIG wash buffer

0.1 M Maleic acid

0.15 M NaCl

pH to 7.5, autoclave, then add 0.3% (v/v) Tween-20

- DIG maleic acid buffer

0.1 M Maleic acid

0.15 M NaCl

pH to 7.5, autoclave.

- DIG detection buffer

0.1 M Tris base

0.1 M NaCl

pH to 9.5, autoclave

Abbreviations

[1]Tris tris (hydroxymethyl) methylamine

[2]EDTA ethylene diaminetetra-acetic acid

APPENDIX II – BINARY MATRIX OF RAPD PHENOTYPES (CHAPTER 4).

```
Population LI - Licata, Sicily
LI1    00000010000100001001000010000100000000001
LI2    10000010100110001001000010001101000000101
LI3    10000010000110001001000010000101010000101
LI4    10000010100110001001000010001101100000001
LI5    10000010000110001001000010000001000000101
LI6    01000010100110001001000010000100111100001
LI7    01000010100110001001000010001001011000001
LI8    00000010100110001001000010000111010001001
LI9    10000010000110001001000010000111010001001
LI10   00000010000110001001000010001001010001001

Population PO - Pozallo, Sicily
POa8   00000010100110001001010010100101010000001
POa13  10000010000110001001010010010101000000001
POa15  00000010001110001001010010000101010100001
POa16  00000010100110001001010010000100010000001
POa17  11000010000100001001010010000100010000001
POa23  10000000000110001001010010010100000000001
POb6a  10000010000110001001010010000101010000001
POb6b  10000010000110001001010010000101100000000
POb8   00000010100100001001010010000101010001001

Population VNM - Vila Nova de Milfontes, Portugal
VNM6   00000000000001000000000010100111010000001
VNM9   00000000000001000001000011100011010000001
VNM9b  00000000000001000000000010100011010000001

Population MVV - Mont Ventoux, France
MVV1   00001100000001000000000011000011101010001
MVV2   00101100000001000000000010000111010000000
MVV3   00001100000000000000000010000111010100000
MVV4   00001100000000000000000010000111010000001
MVV5   00001100000001000000000010000011010100001
MVV6   00001000000001000000000011000111101000001

Population SR - Sa Riera, Spain
SR1    00001100000011100001000010100111010000001
SR2    00001100000000100001000010000111011000001
SR5    00001100000011100101000010010111011000001
SR6    00001100000001100001000010010111011000001
SR14   00001100000001100101000010010111010000001
SR17   00001000000011000001000010010111011000001

Population EN - Enfida Plage, Tunisia
EN7    00100000110100000001000010000001010100111
EN8    00100000011100001001000010000100010010001
EN12   00000010000110000001000010000100010000000
EN17   00100000001101000001000010000101010100000
EN18   00100000111101000001000010011100010100000
EN19   00100010000110000001000010000101010010000

Population EJ - El Jem, Tunisia
EJ1    00100000001110001001000000000010101101010
EJ2    00000000011011001001000000000001010000101
EJ3    00100000011111001001000000000110101011001
EJ4    01100010010111001001000000111110101101001
EJ14   01000010010111001001000000010010101000001
```

```
Population SB - Sbeitla, Tunisia
SB4    11000000000110001001000010000101011111000
SB11   00000010001110001001000010000101010000000
SB13   01000000000110001001000010001101110000000
SB14   00000010001111001001000010101101010001000
SB17   10000000000110001001000010000101001000000
SB18   01000010001110001001000010000101110000000
SB19   00000010001110001001000010000100010000001

Population HS - Houmt Souk, Tunisia
HS2    00100000000110000001100010000101010000001
HS5b   00100010000110000001000000000001001100000
HS5c   00100000000100001001000000000010111000000
HS6    00100000000100000001000000000001101100001
HS7a   00100010000110001001000010000000001000001
HS7b   00100010000001001000000010000100010000001
HS13   00100010000110001001000010000101011000001
HS15   00100000000110001001000010000101000000001

Population PRO - Propriano, Corsica
PRO3   10000001000011100011000001000101010100011
PRO4   10000001000011100011000011000101010100111
PRO5   10000001000011100011000010000101010100011
PRO6   10000001000011100011010001000101000100011
PRO8   10000001000011100011010010000101010100111
PRO10  10000001000011100011000011000101000100011

Population TP - Tour de la Parata, Corsica
TP10   10010001000001000000000001010010101011101
TP14   10010001000001000000000011110110101011101
TP16   10010001000000000000000011101101010100101
TP17   10010001000001000000010110100101010111101
TP18   10010001000000000000010111110110101010101
TP19   10010001000001000000000011010010101011101

Population GT - Gioia Tauro
GT1    00000000000000000000000000000010101110001
GT8    00000000100011100000000001101101011000001
GT9    00000000100010100000000010001101011100001
GT10   00000000100001100000000000010010101110001
GT11   00000000101000100001000010001101011100011
GT26   00000000000000010000000001000010101110001

Population SE - Torre San Emiliano, Italy
SE4    00110000101001110001000001100001101011101
SE6    00110000100011110001000001111001000111101
SE22   00100010101011010001000000100001011111101
SE27   00100000101011000001000011100001001111101
SE29   00110000101010010101000011100101010011101
SE31   00010000101011110001000000100001000111101

Population GAL - Galipoli, Italy
GAL1   00000000100011000101000111100101011010001
GAL2   00010000100011000001000111110101010011101
GAL5   00010000101011000001000000100101011111101
GAL9   00010000101011000001000100100101010111101
GAL14  00010000101011000001000011100101011101101
GAL19  00010000101011000101000110001001001100001
```

```
Population SL - San Leone, Sicily
SL4    10000000101010100101000010101001111000001
SL5    00000000101010100111010001010010101100001
SL6    10000000101010100101000010111101011100001
SL9    00000000101010100111000011110101011100001
SL11   00000000001010100011000010100101111100001
SL16   00000000101010100011010010110101010000001

Population CT - Castellamarre, Sicily
CT4    00100010101010100001000011000001010001010
CT13   00100010100010000101000011000101010000001
CT13b  00100010101010000101001011000101011000001
CT13c  00100010101010000101100110000101011000001
CT23   00000010101010100001110010100101011000001

Population MA - Makthar, Tunisia
MA1    00100000000010000011000000100001010100000
MA2    00100000001000000011000000100001010100010
MA6    00000000001000000011000010000101011100010
MA10   00100000001010000011100010000101010100011
MA11   00100000000010000011000010010100010100001

Population TA - Tabarka, Tunisia
TA2    00100010101010000011000010100101010011010
TA6    00100000000000000011000010100001010001010
TA9    00100000001001000011000000100000010010001
TA11   00100000001000000011000010100101011100001
TA12   00100000001010000011000010101101010110001
TA20   00100000101010000011000010101001010000001

Population BI - Bizerte, Tunisia
BI2    00000000001010000001000010101101000011010
BI3    00100000001000000001000010100101011000010
BI7    00100000001000000001000000000101010000000
BI9    00100000000000000001000000101100010000000
BI10   00100000001000000001000010001101011000010
BI12   00100000001000000001000010100101010000011
BI13   00000000001000000001000010100101010000001
BI15   00100000000010000001000000100001011000000

Population HA - El Haouaria, Tunisia
HA2    00101000001000000001001000100001000110000
HA5    00100000001010000001000010100100010110000
HA6    00100010001000000001000010101100000000000
HA7    00100000001000000001000010100101000000000
HA8    00100000001000000001000010100101011100000
HA10   00100000000000000001000000100101010000001
HA13   00100000001000000001000010110101001010001
HA14   00000000001000000001000010000001001011010

Population HM - Hammamet, Tunisia
HM1    00100000001010000001000000001101010100001
HM3    00100000001010000001000010100101010000000
HM6    00100000001010000001000000100101010000011
HM9    00000000001010000001000010000101010000000
HM16   00000000001010000001000000000100011001111
```

APPENDIX III - DNA SEQUENCE ALIGNMENTS OF *SCYC2* AND *PGIC* USED IN CHAPTERS FIVE AND SIX.

Appendix IIIa – *Scyc2* nucleotide alignment (chapter five).

Appendix IIIb – *PgiC* nucleotide alignment (chapter five).

Appendix IIIc – *Scyc2* nucleotide alignment (chapter six).

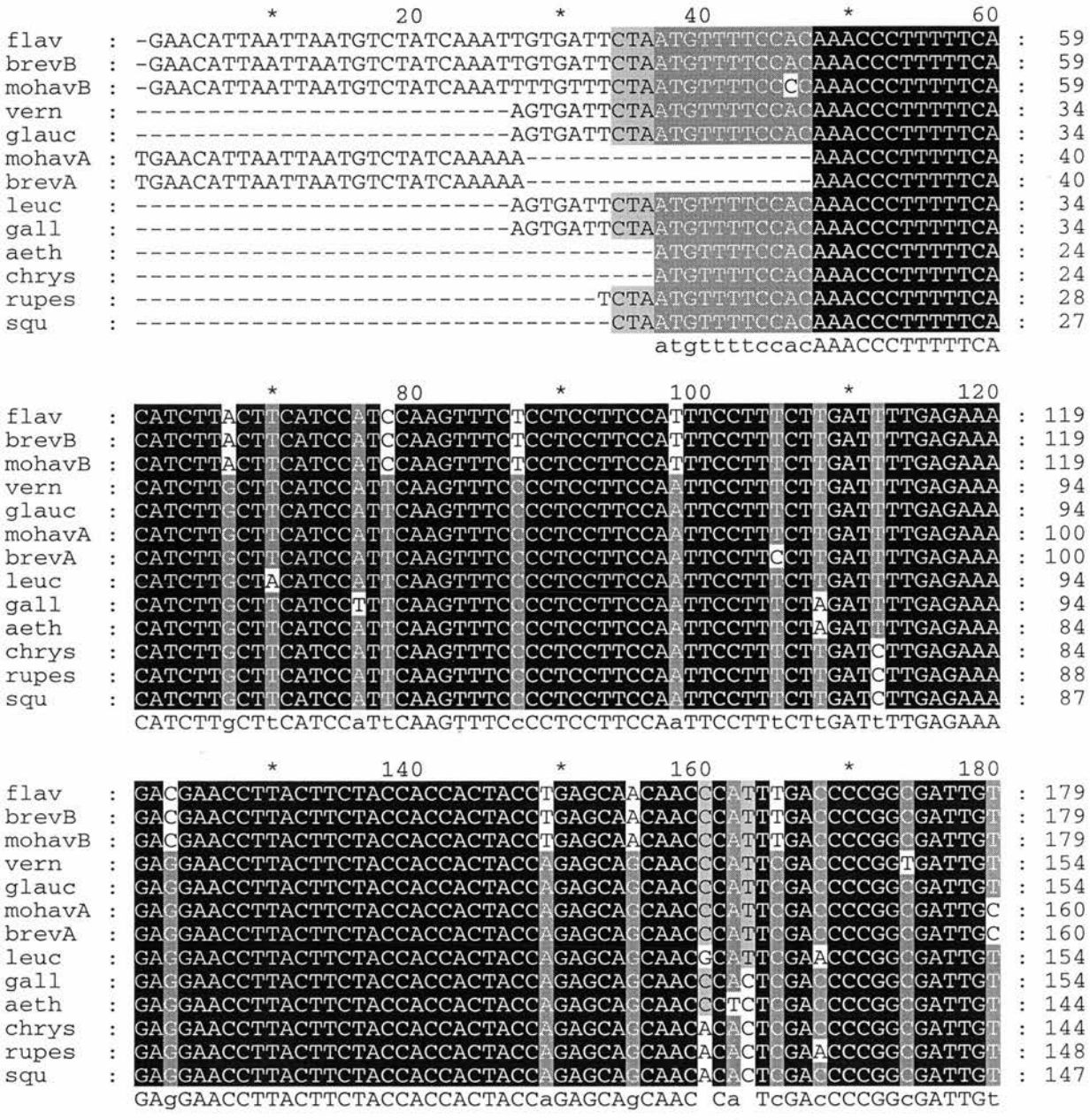
Appendix IIId – *PgiC* nucleotide alignment (chapter six).

Abbreviations are as follows:

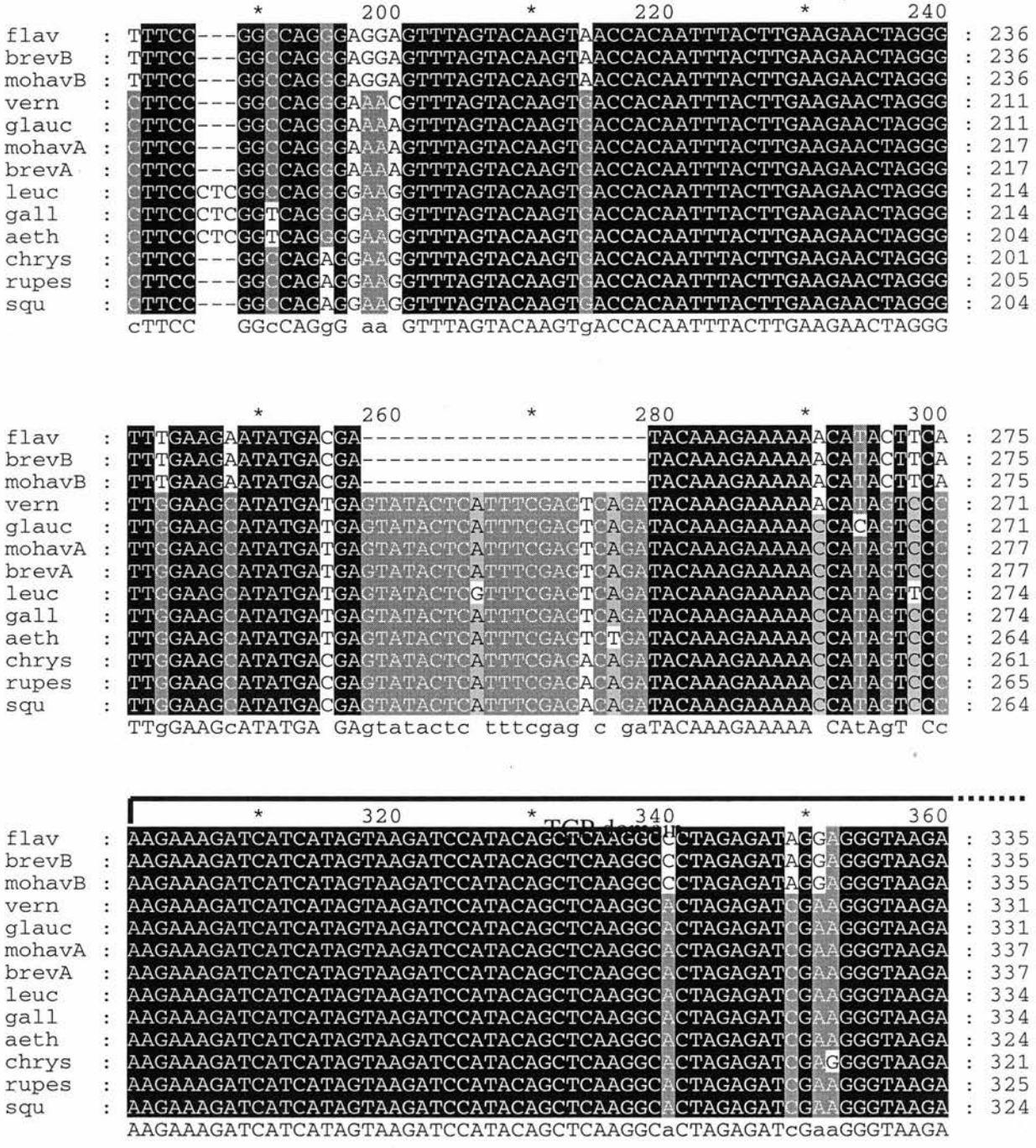
aeth	*Senecio aethnensis*
brev[1]	*S. mohavensis* ssp. *breviflorus*
cam[1]	*S. cambrensis*
chrys	*S. chrysanthemifolius*
flav	*S. flavus*
gall	*S. gallicus*
leuc	*S. leucanthemifolius*
mohav[1]	*S. mohavensis* ssp. *mohavensis*
rupes	*S. rupestris*
squ[2]	*S. squalidus*
ten[1]	*S. teneriffae*
vern	*S. vernalis*
vulN[1]	*S. vulgaris* var. *vulgaris* (non-radiate)
vulR[1]	*S. vulgaris* var. *hibernicus* (radiate)

[1] indicates that more than one sequence was amplified from these polyploid species, designated A, B or C. [2] Two individuals of *S. squalidus* were analysed in chapter 6.

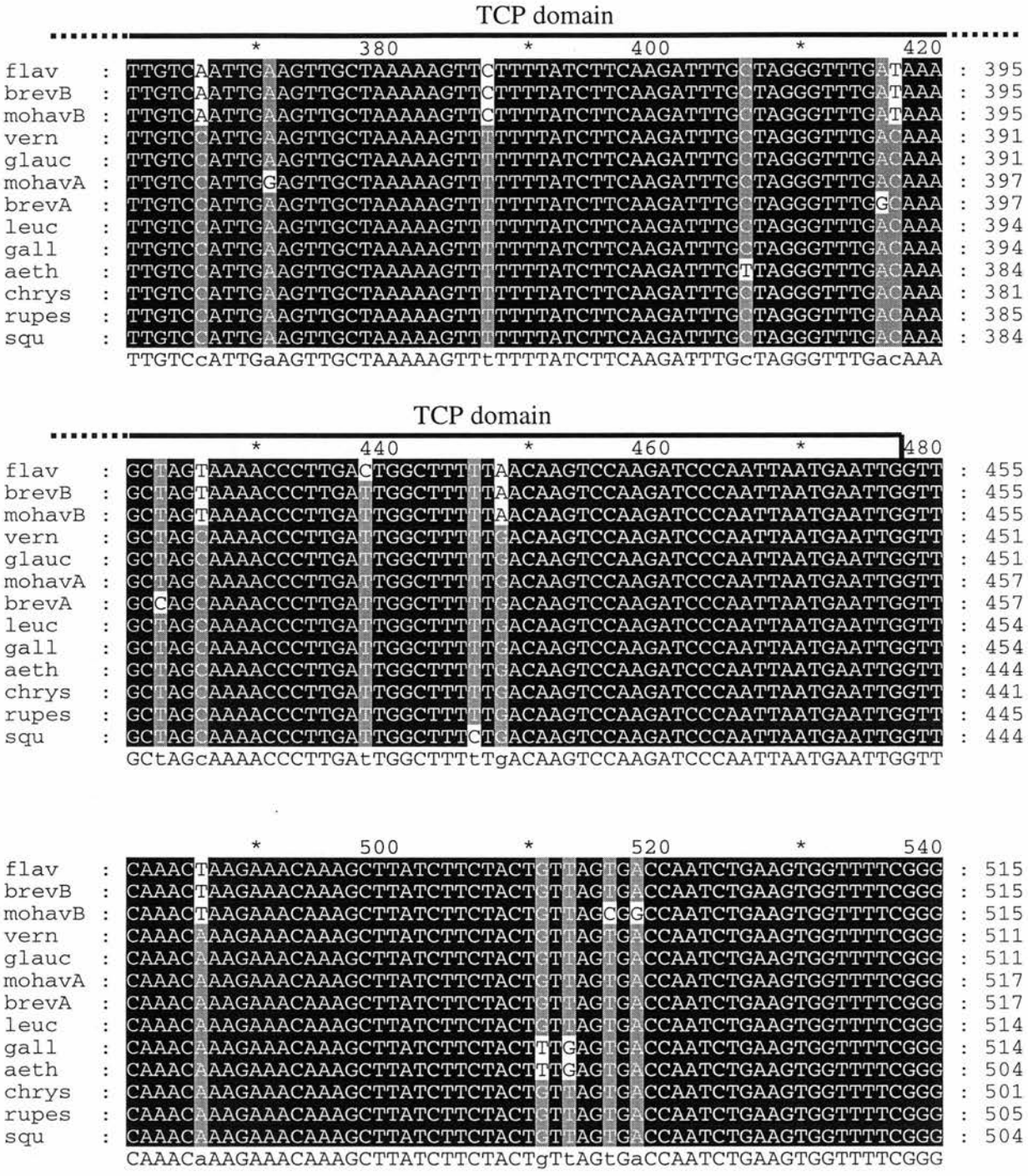
Appendix IIIa – *Scyc2* sequence alignment (chapter 5)



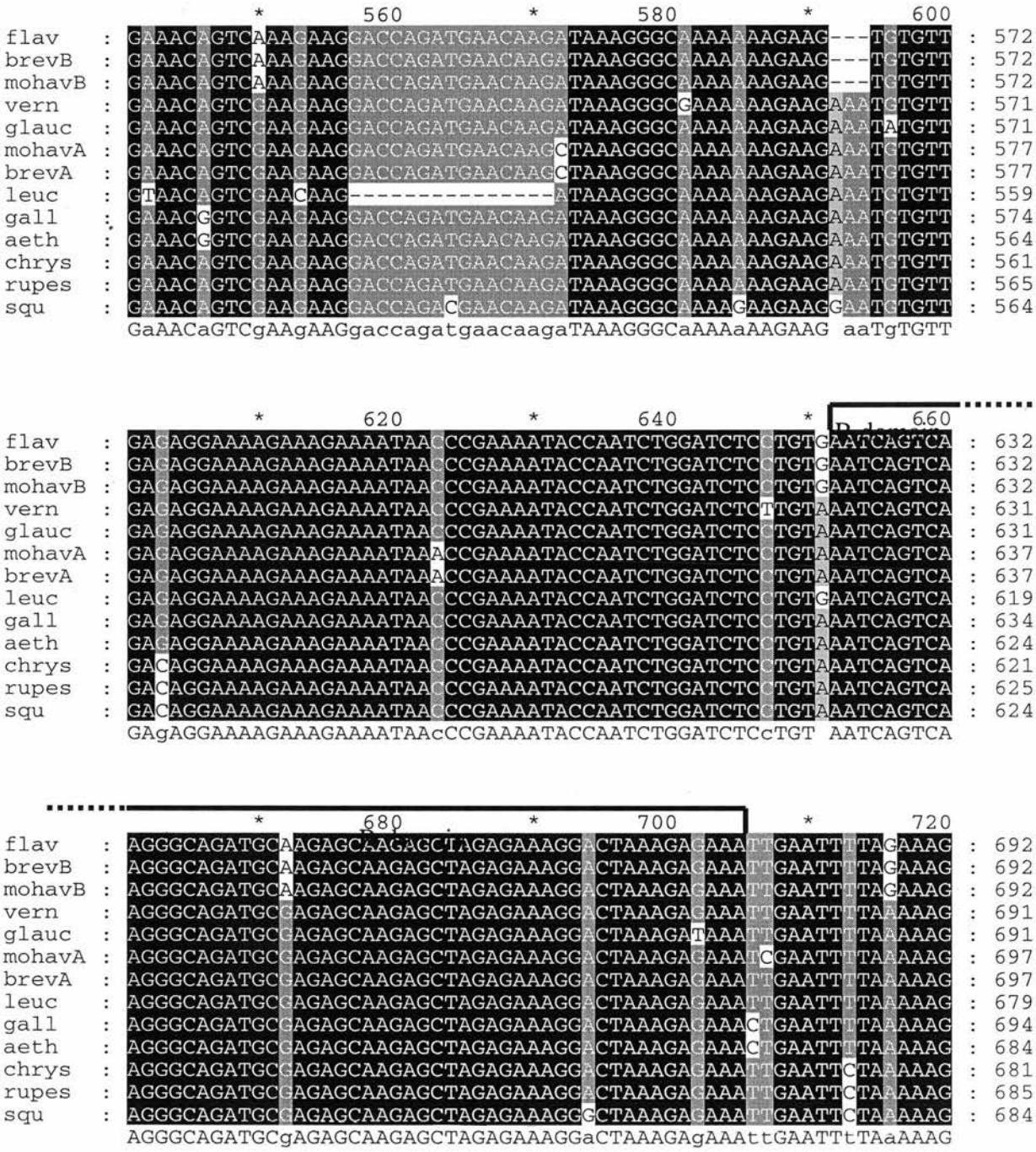
Appendix IIIa – *Scyc2* sequence alignment (chapter 5, continued)



Appendix IIIa – *Scyc2* sequence alignment (chapter 5, continued)



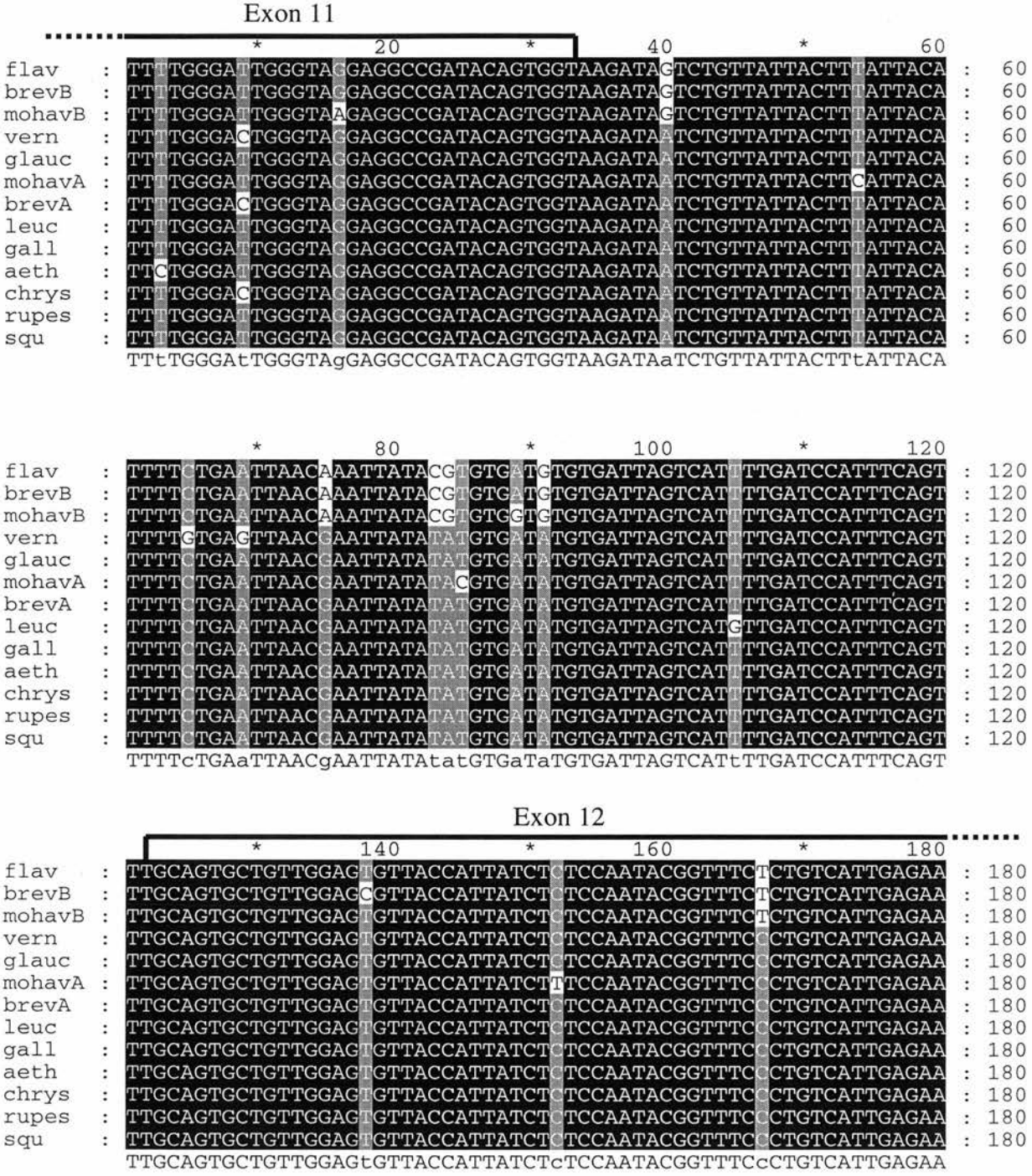
Appendix IIIa – *Scyc2* sequence alignment (chapter 5, continued)



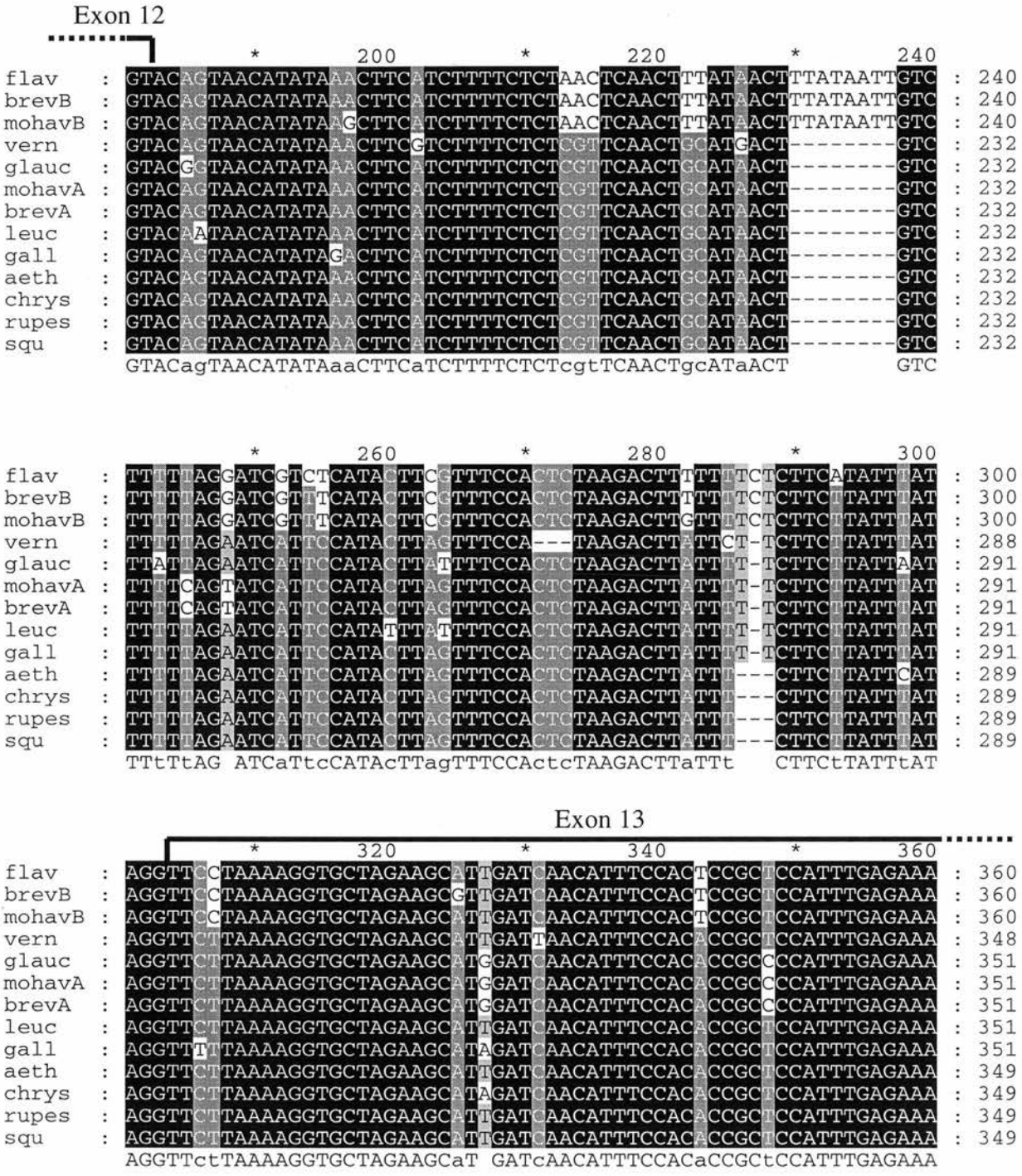
Appendix IIIa – *Scyc2* sequence alignment (chapter 5, continued)

```
                    *       740         *
flav    : CTCGATATCGAGCCAAAGAGTGTT---------- : 716
brevB   : CTCGATATCGAGTCAAAGAGTGTT---------- : 716
mohavB  : CTCGATGTCGAGTCAAAGAGTGTT---------- : 716
vern    : CTCGATATCGAGTCAAAGAGTGTT---------- : 715
glauc   : CTCGATATCGAGTCAAAGAGTGTT---------- : 715
mohavA  : CTCGATATCGAGTCAAAGAGTGTT---------- : 721
brevA   : CTCGATATCGAGTCAAAGAGTGTT---------- : 721
leuc    : CTCGATATCGAGTCAAAGAGTGTT---------- : 703
gall    : CTCGATATCGAGTCAAAGAGTGTT---------- : 718
aeth    : CTCGATATCGAGTCAAAGAGTGTT---------- : 708
chrys   : CTCGATATCGAGTCAAAGAGTGTT---------- : 705
rupes   : CTCGATATCGAGTCAAAGAGTGTT---------- : 709
squ     : CTCGATATCGAGTCAAAGAGTGTT---------- : 708
          CTCGATaTCGAGtCAAAGAGTGTT
```

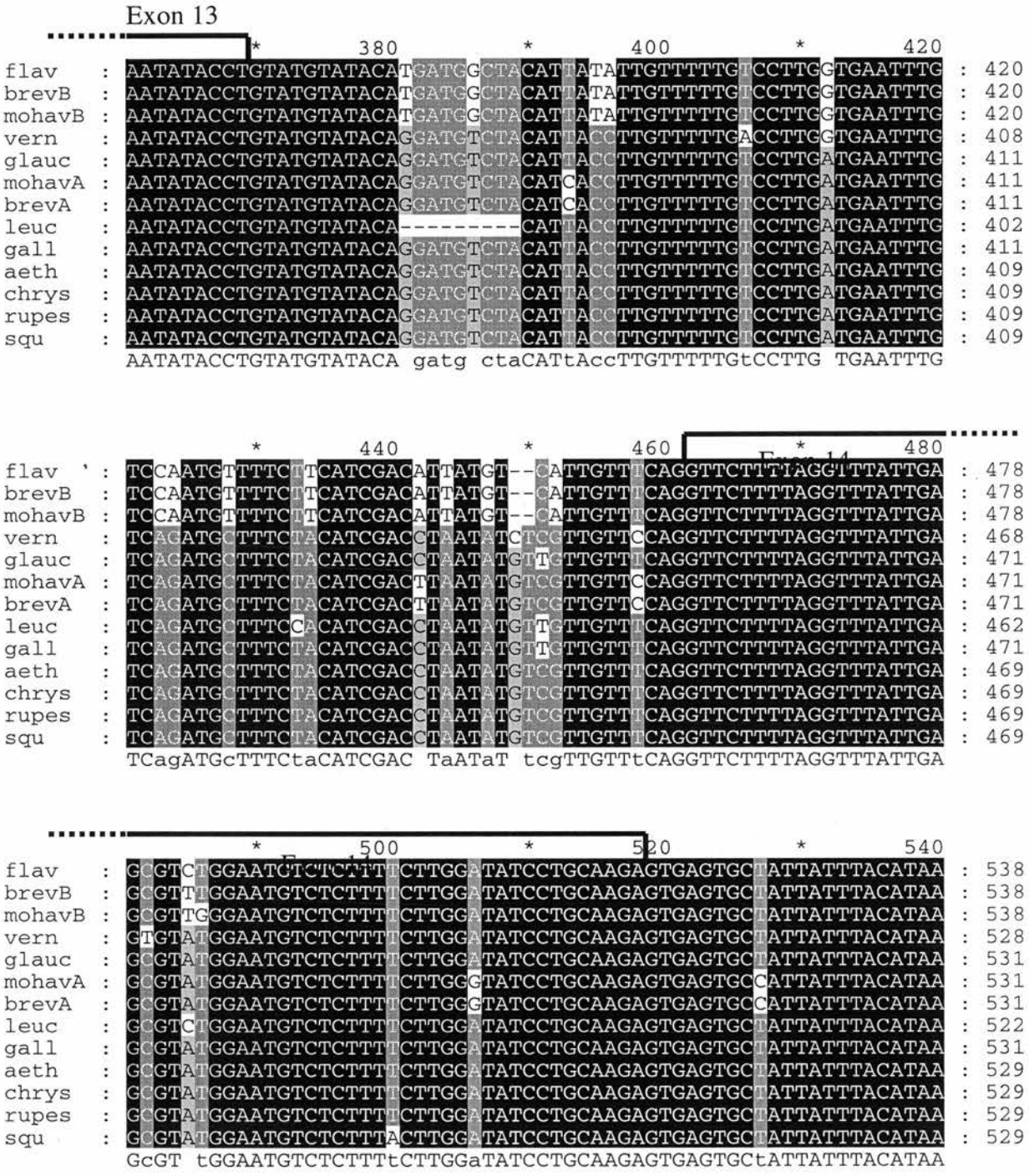
Appendix IIIb – *PgiC* sequence alignment (chapter 5)



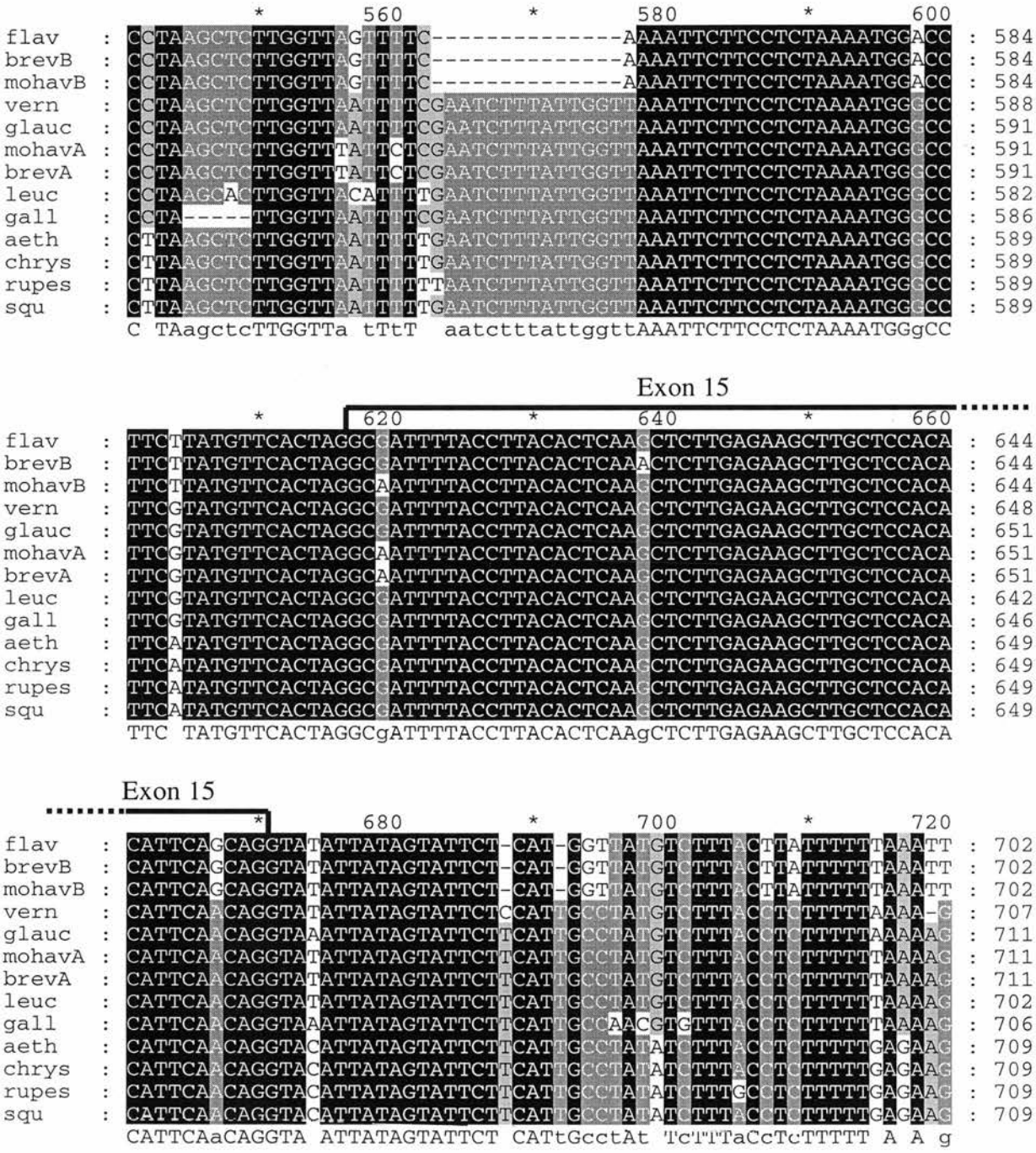
Appendix IIIb – *PgiC* sequence alignment (chapter 5, continued)



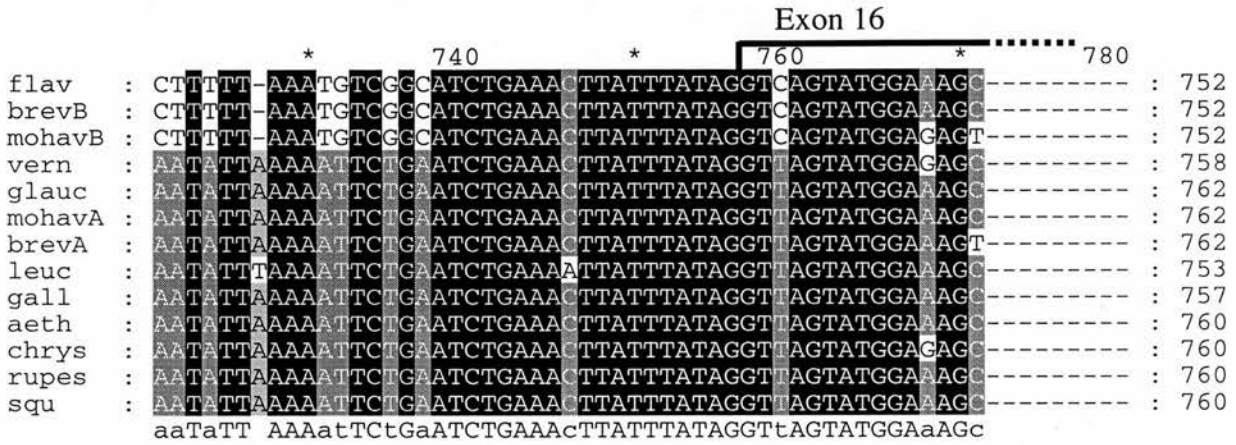
Appendix IIIb – *PgiC* sequence alignment (chapter 5, continued)



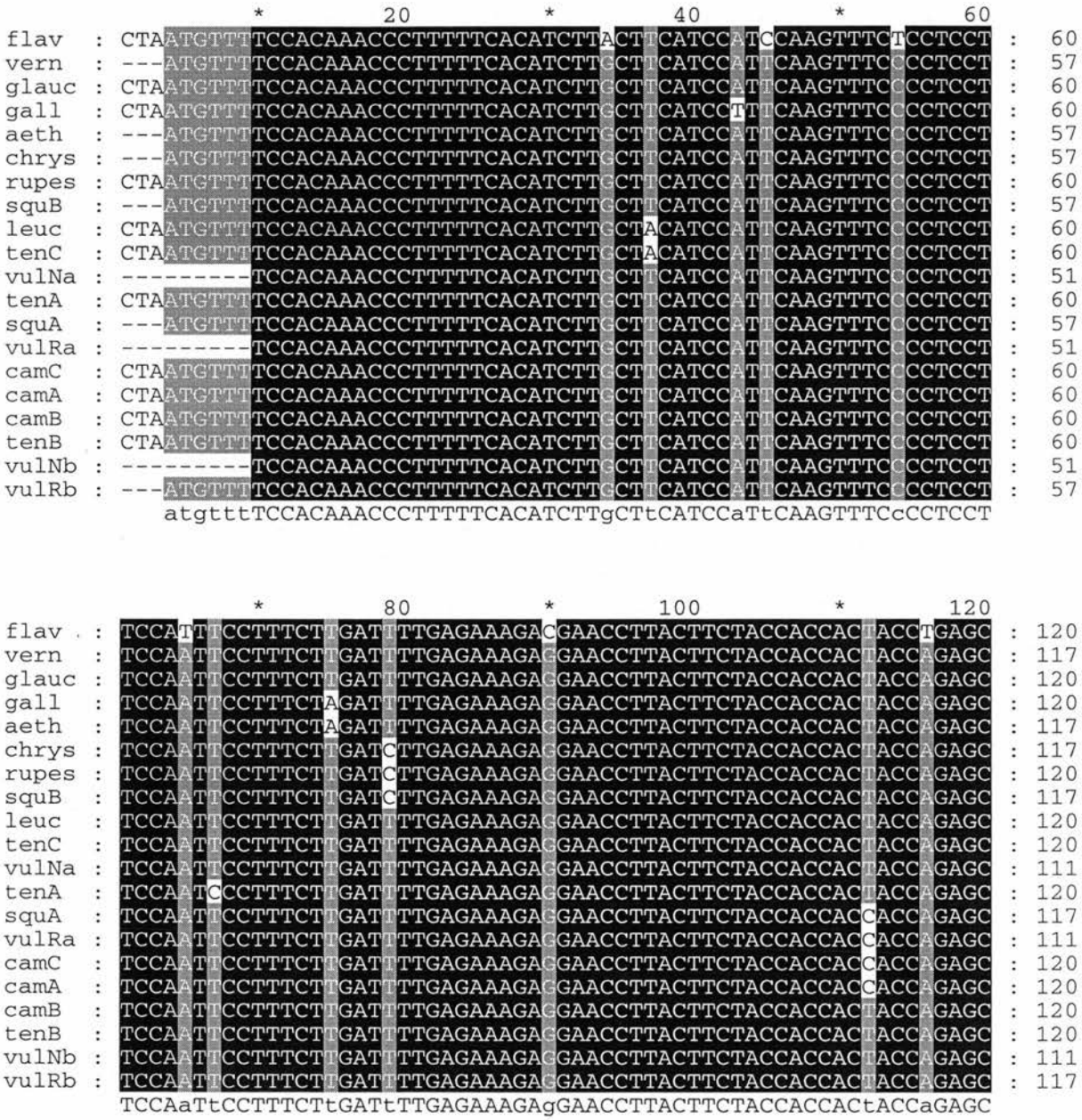
Appendix IIIb – *PgiC* sequence alignment (chapter 5, continued)



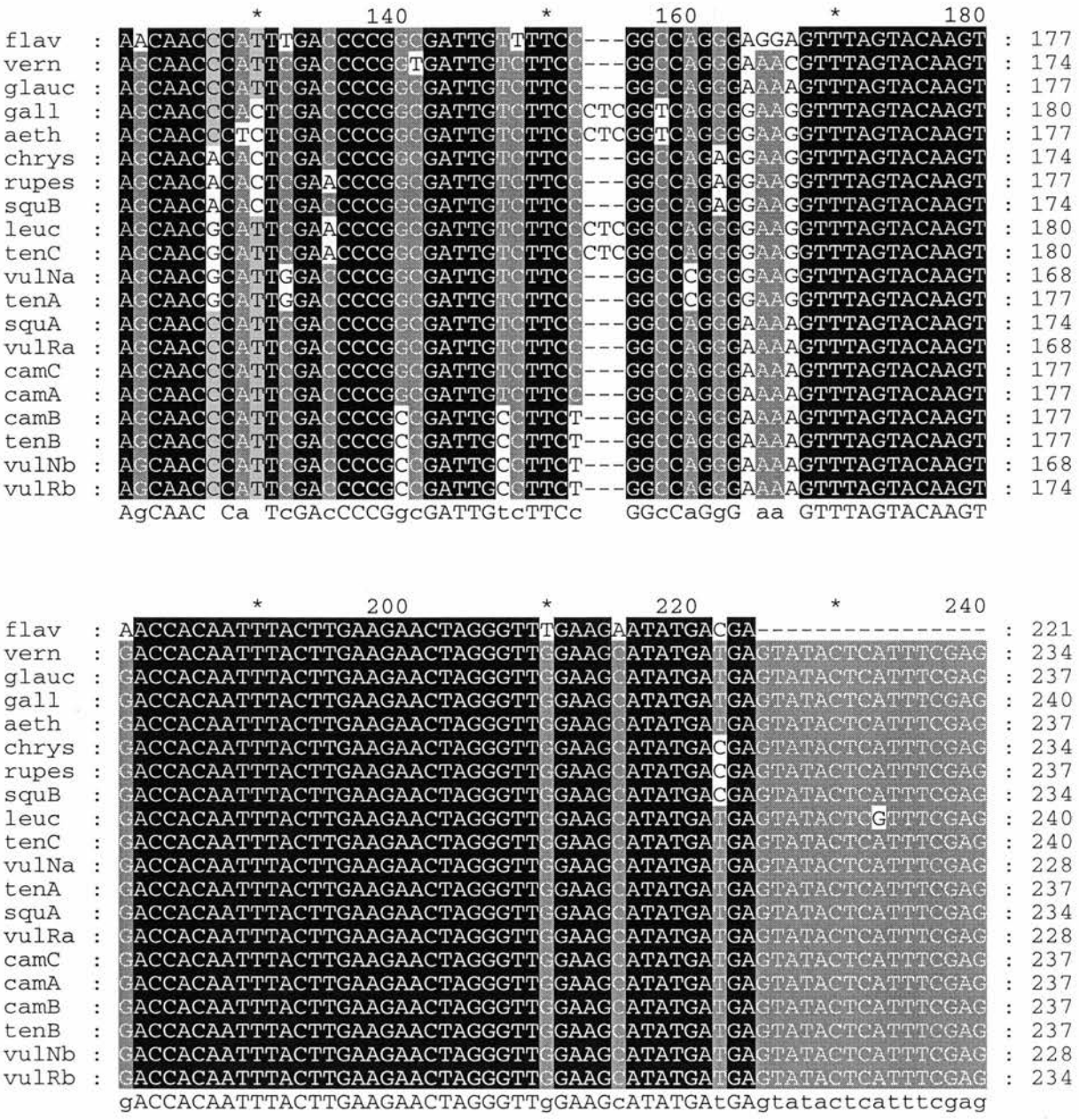
Appendix IIIb – *PgiC* sequence alignment (chapter 5, continued)



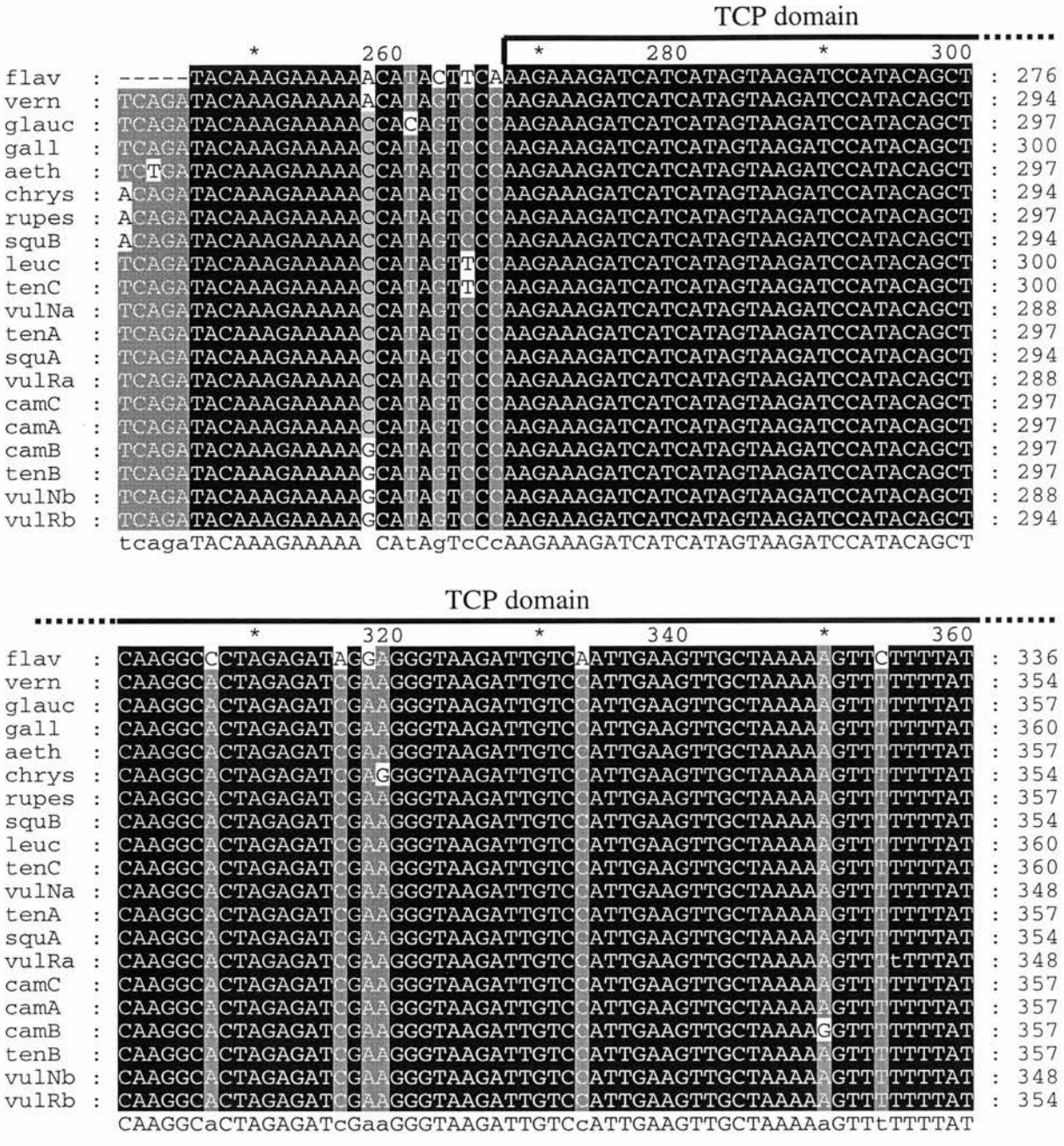
Appendix IIIc – *Scyc2* sequence alignment (chapter 6)



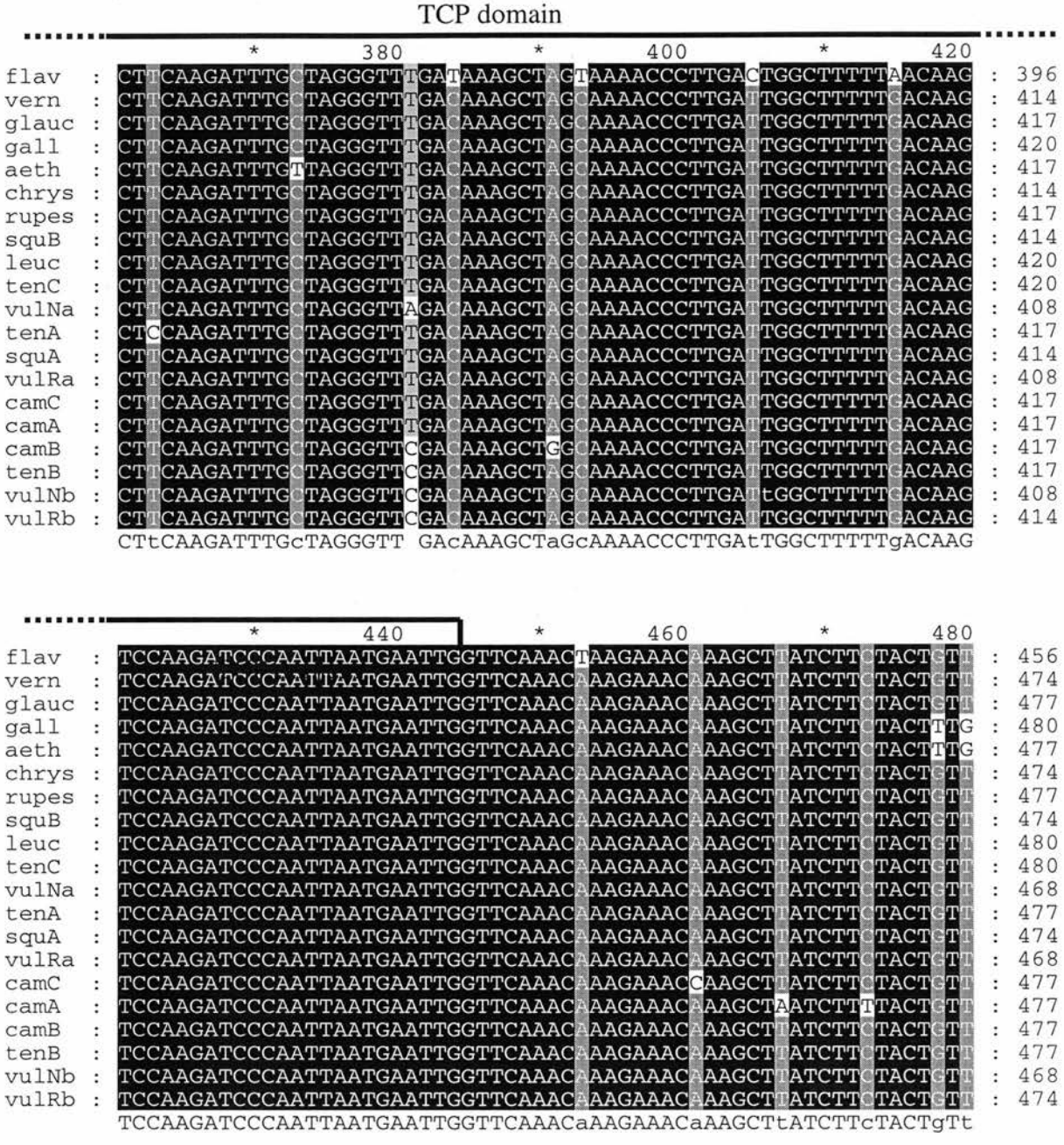
Appendix IIIc – *Scyc2* sequence alignment (chapter 6, continued)



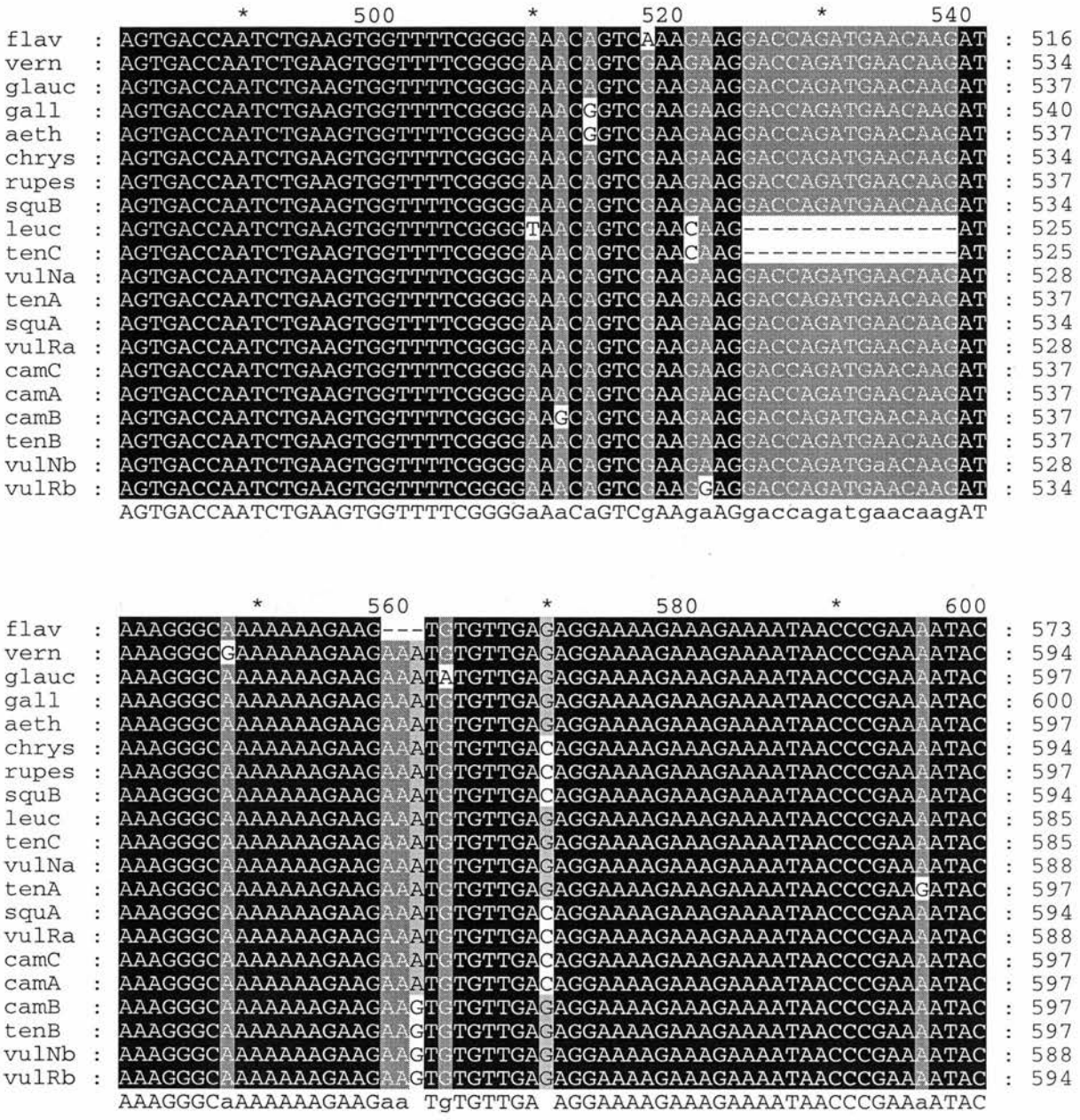
Appendix IIIc – *Scyc2* sequence alignment (chapter 6, continued)



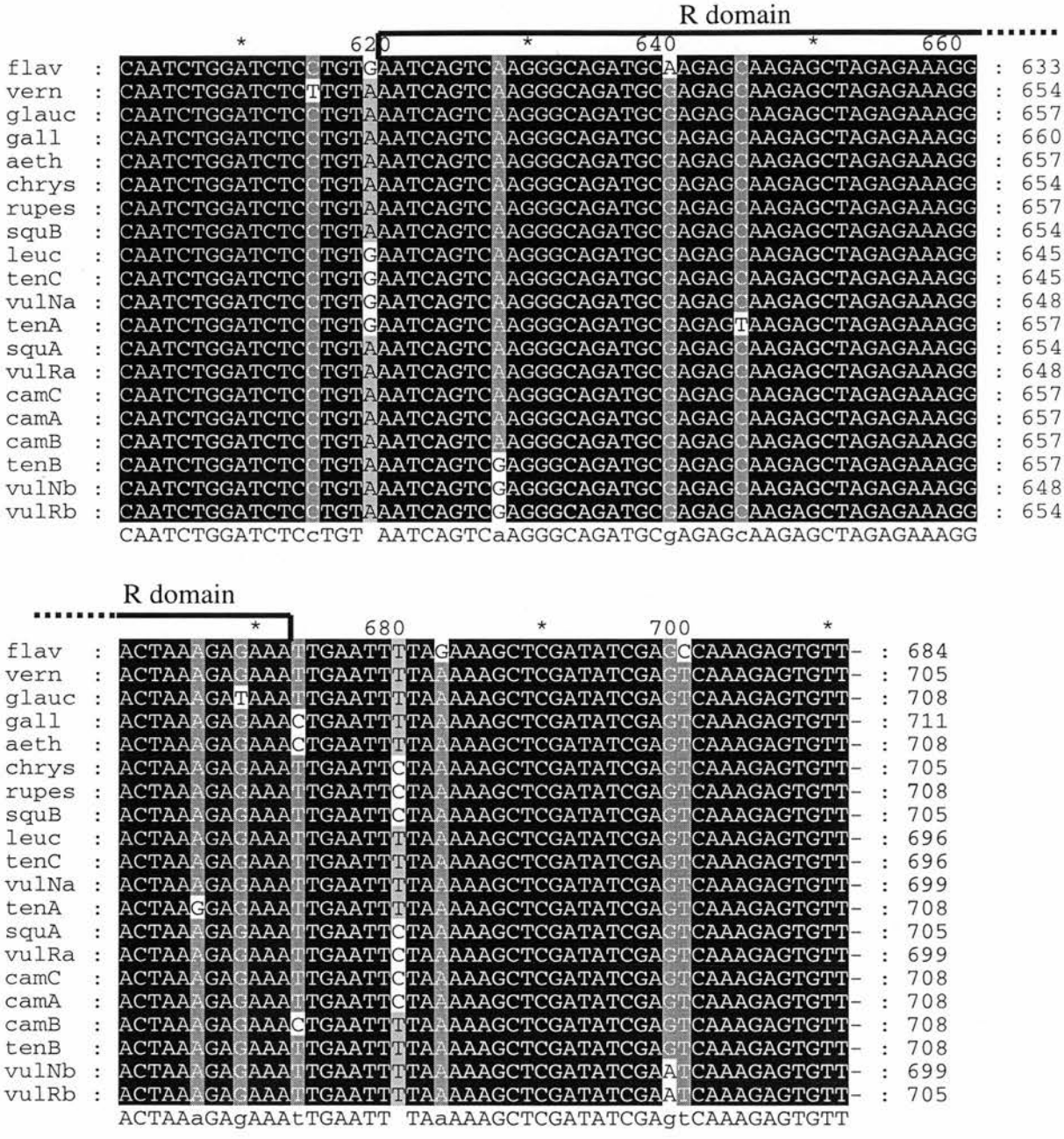
Appendix IIIc – *Scyc2* sequence alignment (chapter 6, continued)



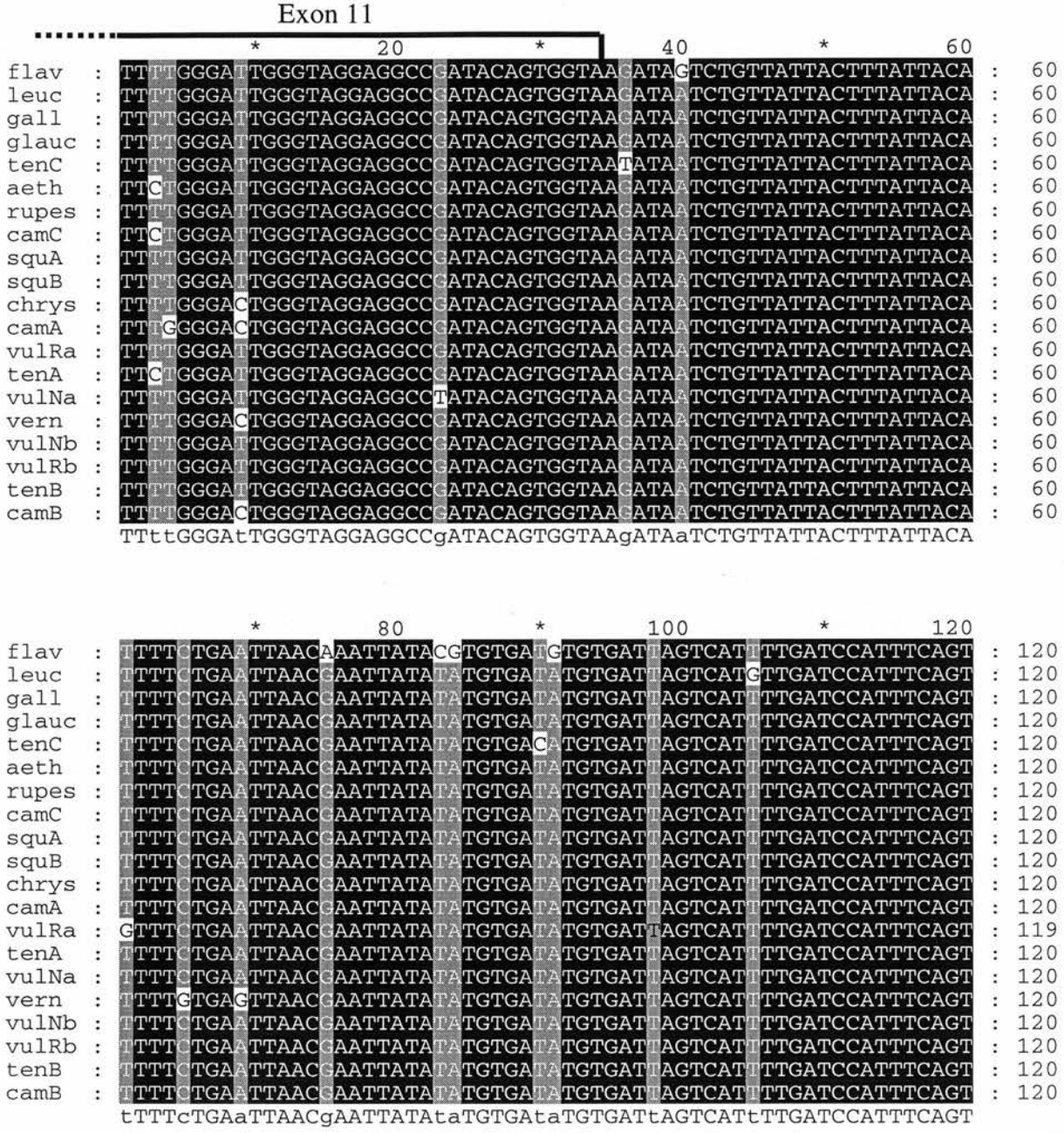
Appendix IIIc – *Scyc2* sequence alignment (chapter 6, continued)



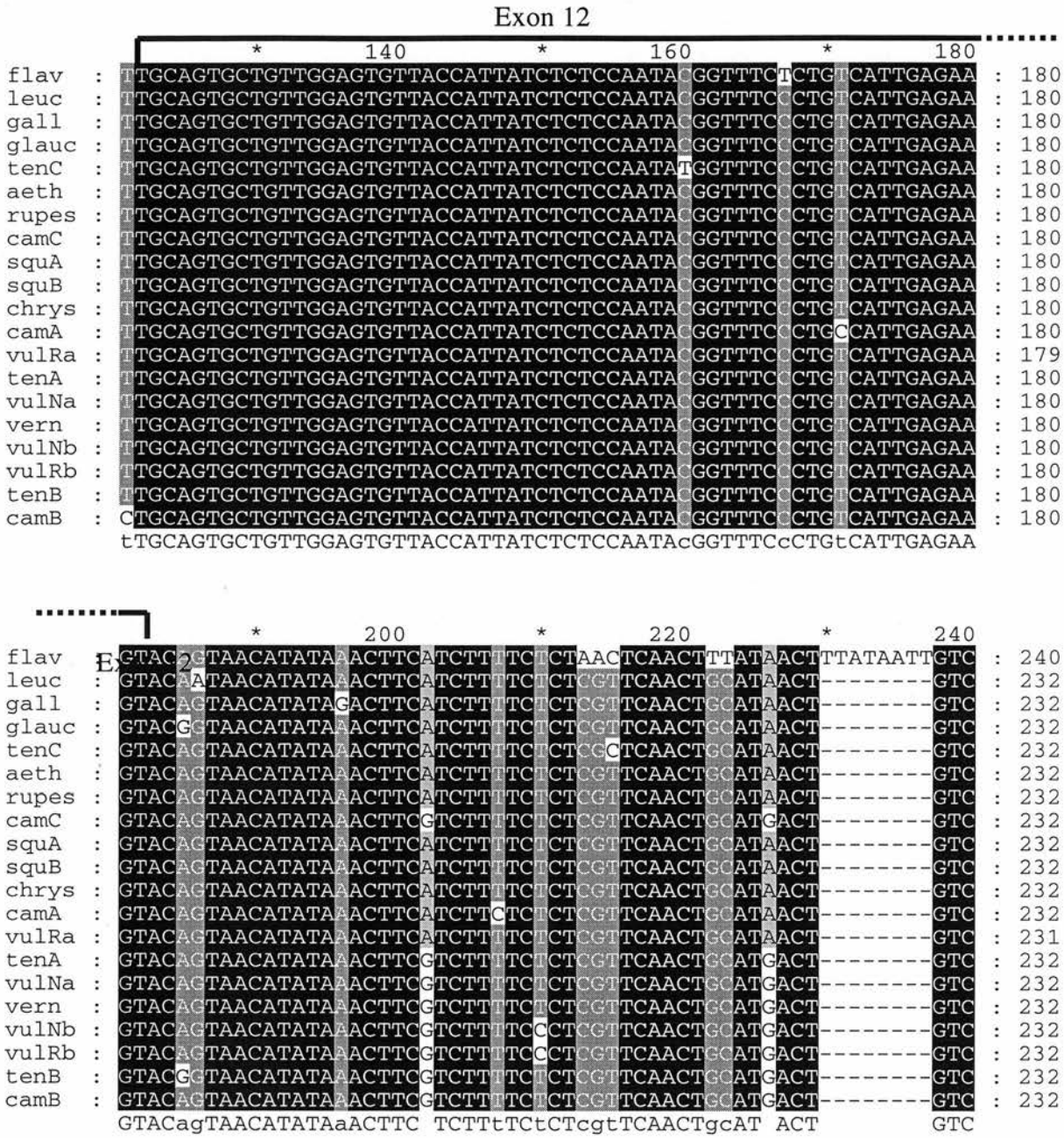
Appendix IIIc – *Scyc2* sequence alignment (chapter 6, continued)



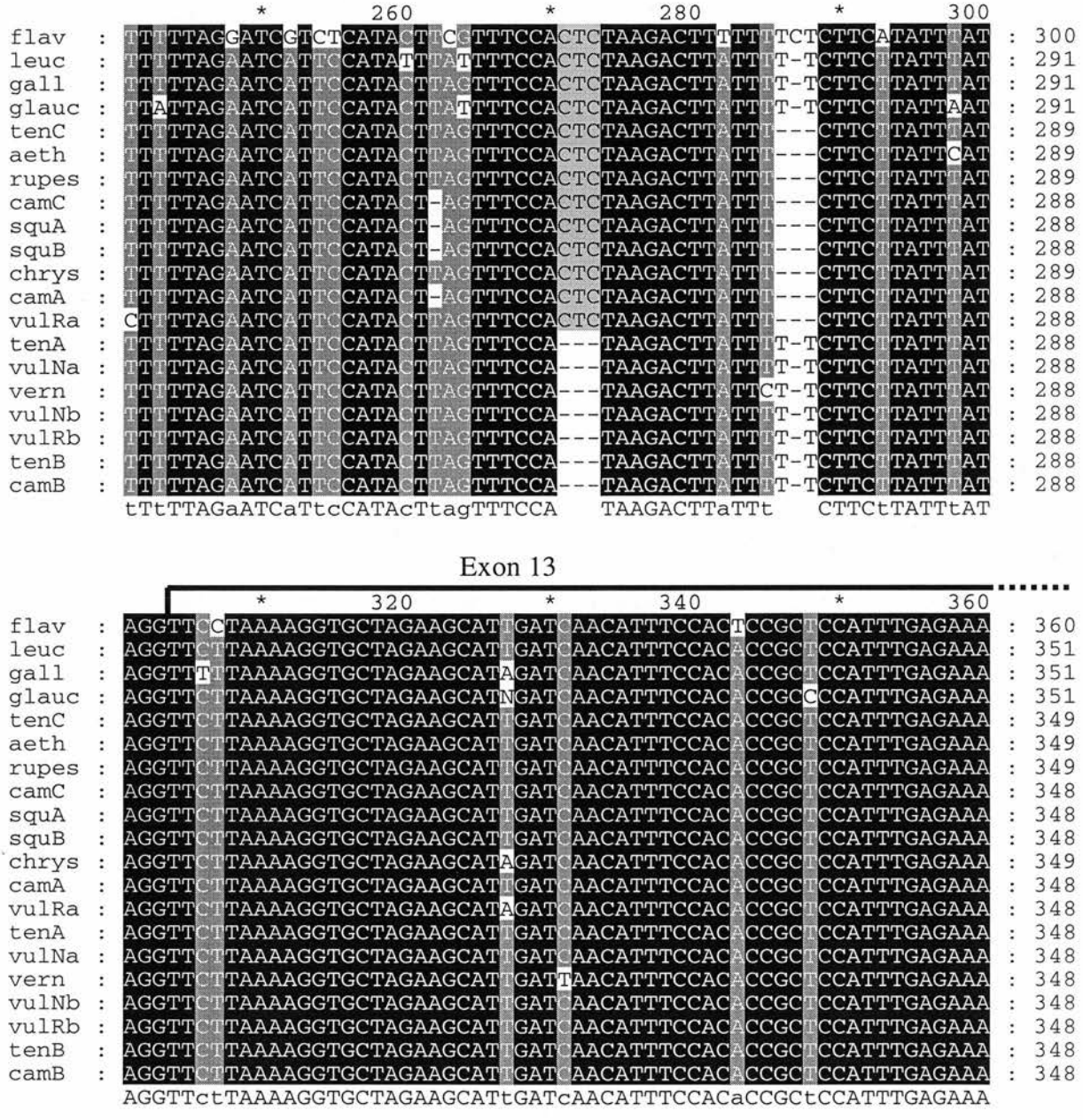
Appendix IIId – *PgiC* sequence alignment (chapter 6)



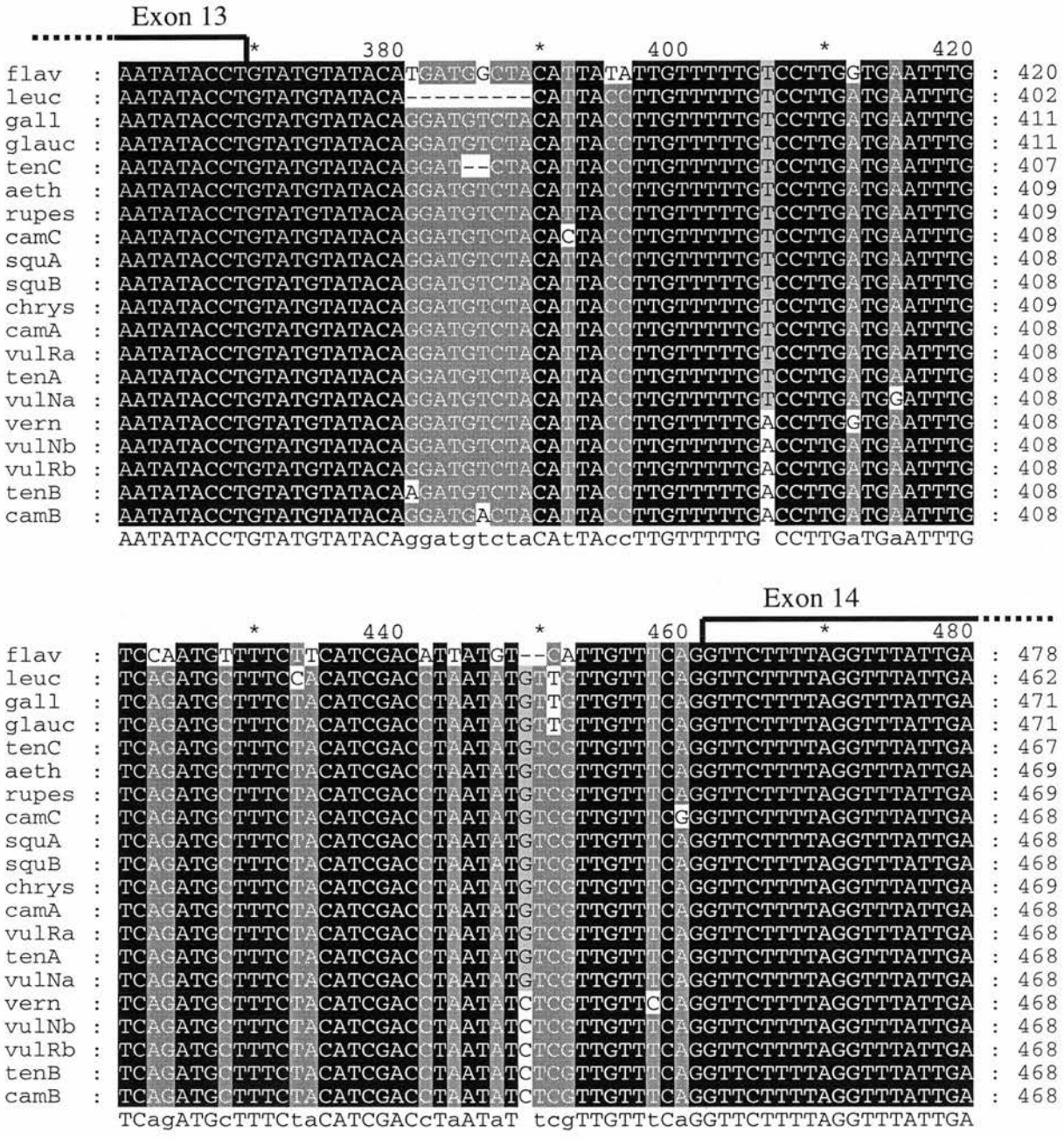
Appendix IIId – *PgiC* sequence alignment (chapter 6, continued)



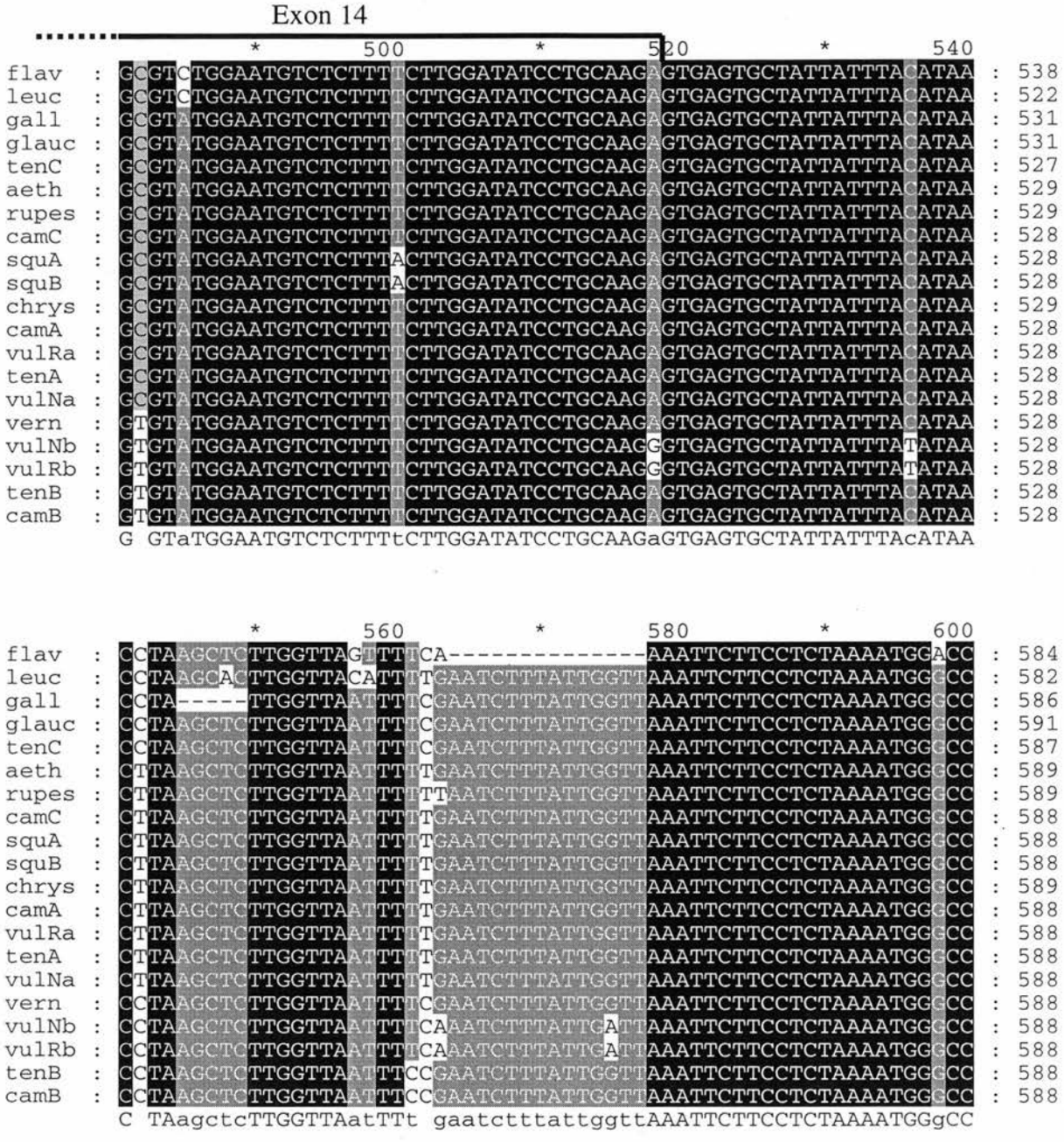
Appendix IIId – *PgiC* sequence alignment (chapter 6, continued)



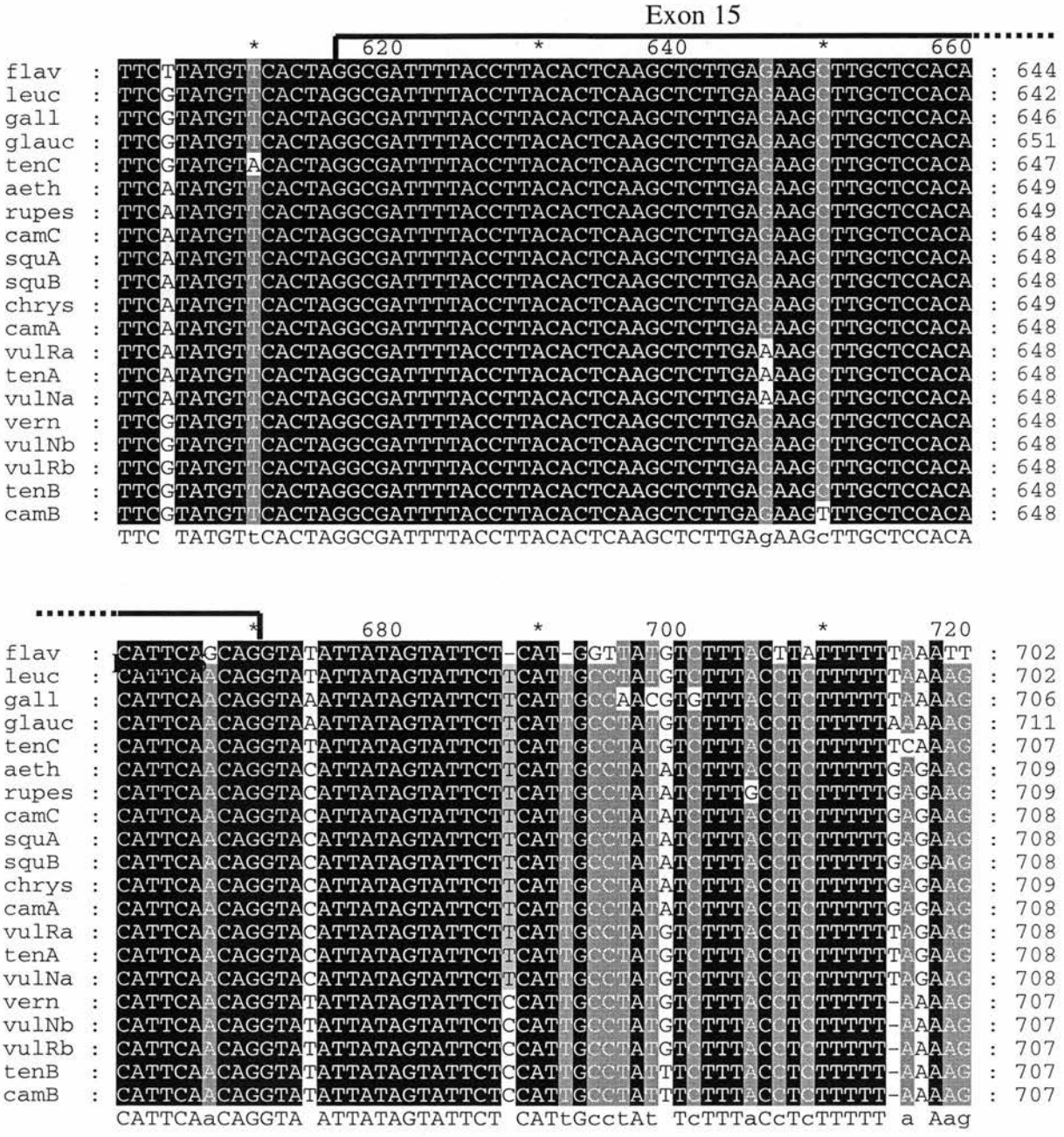
Appendix IIId – *PgiC* sequence alignment (chapter 6, continued)



Appendix IIId – *PgiC* sequence alignment (chapter 6, continued)



Appendix IIId – *PgiC* sequence alignment (chapter 6, continued)



Appendix IIId – *PgiC* sequence alignment (chapter 6, continued)