

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

A survey of the genetic diversity in
populations of White Clover,
Trifolium repens with a focus on
South-Western Europe

A thesis presented in partial fulfilment of the requirements for the
degree of Masters of Science in Plant Molecular Biology
at Massey University, Palmerston North, New Zealand.

Roanna Richards 2011



Trifolium repens, white clover

Acknowledgements

Much appreciation is extended to my supervisor Professor Warren Williams, and Doctor Nick Ellison, for all their time and assistance with this survey of genetic diversity. A sincere thank you to the Forage Improvement Section of AgResearch, Grasslands for their assistance with numerous tasks.

Many thanks, to AgResearch, Grasslands for the opportunity to conduct this research.

Abstract

Trifolium repens (white clover), is an allotetraploid, formed from the cross of putative parents *T. pallescens* and *T. occidentale*. The purpose of this study was to analyse intra-species genetic diversity of *T. repens* among and within populations from the Mediterranean, UK, continental Europe, North Africa, and the near East. This allowed light to be shed on the origins and radiation of the species. It was hypothesized from current distributions of the putative parents, and *T. repens* that an origin for the species was within South-Western Europe.

A survey of genetic diversity was conducted through PCR and direct sequencing of the chloroplast trnL intron, and a chloroplast internal repeat region (IRB21NE) (300 plants) and from the nuclear rDNA, ITS regions of the nucleolar organizing region, (NOR), (327 plants). The interpretation of results was complicated by the spread of commercial varieties over the top of natural populations of *T. repens*.

Diversity was greater for ITS pattern types (45) than for chloroplast pattern types (5). Maximum chloroplast diversity was found in Portugal/Spain. Pockets of ITS diversity were observed in, Portugal/Spain, North Africa, UK (Kent landrace) and Northern/North-Western Europe. Within the natural distribution zone for the species two chloroplast lineages and three main nuclear lineages described the majority of the observed variation. However a novel nuclear lineage was observed in north/north-western Europe. Spain and Portugal differed in their predominant ITS lineages, as did Morocco and Algeria from each other.

The diversity in chloroplast data was consistent with an origin in Portugal and Spain. ITS diversity also supported this, but the presence of other areas of ITS diversity, gave a complex pattern that requires further study.

Abbreviations

PCR: Polymerase chain reaction

ITS: Internally transcribed spacer region

NOR: Nucleolar organising region

IRB21NE: Internal repeat region 21, NE- Nick Ellison.
(a shortened version of the full length IRB21)

FISH: fluorescence *in situ* hybridization

rDNA: ribosomal DNA

TCS (version 1.2.1): a software program that does phylogenetic estimation by using statistical parsimony, (Clement et al., 2000).

Table of Contents

Abstract	1
Abbreviations	2
Table of Contents	3-6
1 Literature review	1-21
1.1 <i>T. repens</i> or White Clover	1-7
1.1.2 <i>T. repens</i> : a rare hybrid of the Trifolium genus	
1.1.3 Parentage of <i>T. repens</i>	
1.1.4 Morphology of white clover	
1.1.5 Self incompatibility	
1.1.6 Diversity in white clover DNA	
1.1.7 Distribution of white clover	
1.1.8 Breeding history of <i>T. repens</i>	
1.2 Clover in general	7-10
1.2.1 The Trifolium genus	
1.2.2 Origins of clover	
1.2.3 Spread of clover	
1.2.4 Agricultural use	
1.3 Appropriate Genetic markers for phylogeny & diversity	10-17
1.3.2 Polyploidy	
1.3.3 Polyploid models for comparison to white clover	
1.3.4 Concerted Evolution, an influence on DNA sequence diversity	
1.3.4.2 Bidirectional Concerted Evolution	
1.3.4.3 Factors affecting the observation of and effect of concerted evolution	
1.3.5 Reticulate evolution	
1.4 Glaciations and affect on species and populations	17-21
1.4.2 Glacial hypotheses	
1.4.3 Interpreting species genetic diversity in relation to glaciations	

1.5 Hypotheses and aims	22
1.5.1 Aims	
1.5.2 Hypotheses	
2 Methods	23-43
2.1 Accession selection and growing of plants	23-30
- A list of used accessions and numbers of surviving plants	
- Seed germination techniques	
- Growing conditions and methods	
2.2 PCR	31-33
- Reagents and kits used in PCR	
- PCR primers	
- Gel purification protocol	
2.3 Sequencing	34-37
- Reagents for sequencing	
- Sequencing protocol	
2.4 Chromatogram analysis and contig assembly	37-39
2.5 Diversity analysis	39
2.6 Morphology study	40-42
2.7 Comparison of DNA and morphology	43
2.8 Novel Algerian accession C20859	43
3 Results from DNA sequence data	44-125
3.1 Results from chloroplast perspective	44-60
- Universal CD primers (<i>trn L</i> intron)	
- Modified IRB21(NE) primers	
3.2 Results from ITS nuclear pattern types	61-89
- EC1/EC2 (ITS) primers	
3.3 Results for the combined pattern types generated from the combination of the chloroplast and nuclear pattern types	90-104
3.4 TCS Analyses of pattern types	105-118

3.5	Combined Results and Morphological characteristics	119-125
4	Discussion DNA sequence data	126-167
4.1	Observations of chloroplast sequences	126-137
4.1.1	Geographic distributions of chloroplast sequences	
4.1.2	Comments on chloroplast sequence variation from selected ecotypes, ecotype selections and mixed cultivars	
4.1.3	Comparison of chloroplast diversity in Zone 1 to Zone 2	
4.1.4	Relationships between observed chloroplast pattern types	
4.1.5	Chloroplast pattern type origins	
4.2	Observations of Nuclear ITS sequences	138-152
4.2.1	Geographic distributions of nuclear ITS sequences	
4.2.2	Discussion of main and variants of main nuclear ITS pattern types	
4.2.3	Infrequently observed Nuclear ITS pattern types	
4.2.4	Hypotheses of dispersal and origin with a focus on common ITS types	
4.2.5	Parental or Biparental inheritance of nuclear ITS pattern types in <i>T. repens</i> ?	
4.3	Discussion of the combined perspective	153-160
4.3.1	Geographic distributions of combined (<i>trn L</i> , IRB21NE, and ITS) pattern combinations.	
4.3.2	Chloroplast & nuclear results do they agree: or disagree?	
4.3.3	Diversity and models of origin, from a combined perspective	
4.3.4	Discussion of the perspective gained from TCS networks	
4.3.5	Discussion of Combined DNA sequence pattern types and Morphology	
4.4	Discussion of the perspective gained from TCS networks	161-162
4.5	Discussion of Combined DNA analyses and Morphology	162-163
4.6	Concluding remarks	163
4.7	Limitations of the study	163-164
4.8	Future directions, extensions from this work	164-166
4.8.1	Extension of the sampled area	
4.8.2	Single copy markers	
4.8.3	Chromosome specific markers	
4.8.4	Phylogenetic analysis	

4.9 Conclusion	167
5 Novel Algerian species C20859	168-177
- Introduction	
- results	
- Discussion	
References	178-187

Appendices

A: Chloroplast DNA sequence data	188-199
B: Nuclear DNA sequence data	200-217
C: Combined chloroplast and nuclear DNA sequence data	218-229
D: Combined DNA sequence data and morphology scoring	230-241
E: TCS analyses	242
F: Chi Squared analyses	243-251

Alignments

1. <i>trn L</i> alignment (universal CD primers)	252-299
2. IRB21NE alignment (IRB21NE primers)	300-338
3. ITS alignment (EC1/EC2 (ITS) primers)	339-392

Chapter 1 Literature review

1.1 *T. repens*, or White clover

T. repens, or white clover, is an allotetraploid member of the *Trifolium* genus so called for the characteristic three green leaflets observed primarily in all species of the genus (Williams, 1987a; Ellison et al., 2006). It is classified in section *Trifolium*, (Ellison et al., 2006) but was formerly considered part of section *Lotoidea* (Zohary and Heller, 1984). It is of vital importance to agriculture worldwide, as “the most common cultivated clover species” (Ellison et al., 2006). The ability of white clover to fix nitrogen has led to its use in treating nitrogen deficiency, in classical (Kjaegaard, 2003) and contemporary agriculture. This has led to a designation as the most important species of *Trifolium* (Abberton, 2007).

1.1.2 *T. repens*: a rare hybrid of the *Trifolium* genus

Incongruence between phylogenetic topologies of chloroplast and nuclear sequence data provides evidence of hybridization (Ellison et al., 2006), but within the *Trifolium* genus this rarely occurs (Watson et al., 2000; Ellison et al., 2006; Abberton, 2007). Only five cases of hybrid speciation were observed in the comprehensive study of 218 species of the genus *Trifolium* by Ellison et al. (2006). In hybrids, the detection of putative parents is often complicated by evolution of DNA sequence, destruction or loss of ancestral populations or intermediate species or forms and inadequate samplings of potential areas of interest for species origins (Ellison et al., 2006).

Current evidence suggests a hybrid origin for *T. repens*, from the diploid ($2n=2x=16$) parents within the *Trifolium* section. The maternal parent is *T. pallescens* or a near relative (maternal) and the paternal parent is suggested to be *T. occidentale* (Ellison et al., 2006). Species formation is thought to have occurred through unidirectional hybridization, possibly involving *unreduced gametes*. Speciation resulted in a chromosome complement of $2n=4x=32$. To date, no individual has been found that shows paternal inheritance from *T. pallescens* or maternal inheritance from *T. occidentale* (Professor Warren Williams, pers. com, 2009-2010).

1.1.3 Parentage of *T. repens*

T. repens: mother's story

The maternal contributor to *T. repens* has remained elusive, and to date the closest relative found is the genetically diverse *T. pallescens*. It has been located at high altitude in areas of both the Austrian Alps (Raffl et al., 2008), and the Pyrenees (Larena et al., 2002), which as glacial refugia, contain a richness of species genetic diversity (Larena et al., 2002; Raffl et al., 2008). Variation in Amplified Fragment Length Polymorphism (AFLP) profiles has been reported for *T. pallescens* individuals from different populations from different glacial valleys of the Austrian Alps (Raffl et al., 2008). While *T. pallescens* is most similar in sequence to *T. repens*, it does differ at several positions in the chloroplast *trn L* intron sequence (Table 1.1). Assessed populations of *T. pallescens* from the Pyrenees Mountains are more similar to *T. repens* relative to *T. pallescens* populations from the Austrian Alps (Table 1.1). However, since there has been no comprehensive sampling of populations of *T. pallescens* and there is likely to be genetic variation within the species on a valley and/or glassier basis, significant variation within *T. pallescens* is probable. Hence, a chloroplast sequence match to *T. repens* and therefore the maternal contributor species/population may yet to be found.

Table 1.1 Sequence differences for the *trn L* intron for *-T. repens* (cv. Huia), *T. occidentale*, and *T. pallescens* from the Austrian Alps and the Pyrenees.

Trifolium																				
Individuals	Reference positions																			
	145	146	147	148	288	289	290	291	292	293	-	-	-	-	-	294	295	296	297	
NZ Huia																				
Reference	A	A	A	A	T	A	T	T	A	T	-	-	-	-	-	G	T	A	G	
<i>T. occidentale</i>																				
DQ3118521	A	A	A	A	-	-	-	-	-	-	-	-	-	-	-	T	G	T	A	G
<i>T. pallescens</i>																				
AZ4856	-	-	-	-	T	-	T	T	A	T	-	-	-	-	-	G	T	A	G	
<i>T. pallescens</i>																				
DQ311859	-	-	-	-	T	-	T	T	A	T	-	-	-	-	-	G	T	A	G	
<i>T. pallescens</i>																				
AZ4837	A	A	A	A	T	A	A	T	A	T	A	T	T	A	T	G	T	A	G	
<i>T. pallescens</i>																				
Jaca	A	A	A	A	T	A	A	T	A	T	A	T	T	A	T	G	T	A	A	

Note: *T. pallescens*: Austrian Alps- AZ4856, Pyrenees AZ4837 and Jaca, independent herbarium sample from Instituto Pirenaico de Ecología, C.S.I.C., Jaca, Spain. DQ311859 unknown location possibly Austrian Alps, Direct submissions to genbank: DQ3118521, DQ311859, source; Ellison et al., 2006, Molecular phylogenetics of the clover genus (*Trifolium*- Leguminosae), Molecular Phylogenetics and Evolution Vol. 39 Issue 3, 688-705. Sequences for individuals Jaca, AZ4856, and AZ4837 were obtained from Dr Nick Ellison and Pr Warren Williams, 2009-2010.

T. repens: the fathers story, *T. occidentale*

Assessment of paternal heritage in white clover is based on the ITS region, an internally transcribed spacer region, that is non-coding but is transcribed along with the adjacent 18S and 5.8S, (ITS1) and 5.8S and 26S, (ITS2) ribosomal genes within the NOR (nucleolar organising region). This forms a gene cluster that is repeated many times within a tandem repeat/array (Ansari et al., 1999; Alvarez, 2003; Ellison et al., 2006). In white clover, two of the expected four nuclear organizing regions (NORs) the assumed *T. pallescens* forms, are absent. Only one signal type is observed by cytological and molecular works. DNA sequence analysis of the ITS region, typically returns a single DNA sequence type, almost identical to *T. occidentale* (Ellison et al., 2006), supporting the hypothesis that information from NORs probably only reflects the paternal ancestral contribution.

FISH (fluorescence in situ hybridization) expression patterns of 18S-5.8S-26SrRNA and the separate 5S rRNA also show a strong similarity of the chromosomes of *T. repens* to *T. occidentale*. In an assessment of *T. repens* and seven closely related species, the FISH banding patterns of 18S-26S and 5S rDNA were similar in the NORs between *T. repens*, and *T. occidentale* (2), *T. uniflorum* (4), and *T. nigrescens ssp. nigrescens* (2) (Ansari et al., 1999). *T. nigrescens ssp. petrisavii* and *T. ambiguum* showed some similarity for banding patterns of 5S rDNA (Ansari et al., 1999, Fig 3, P. 203).

Earlier evidence suggests a role of *T. nigrescens* in the ancestry of *T. repens*. Analysis of the *Trifolium nigrescens* species complex in 2001 found three subspecies (*T. nigrescens ssp. nigrescens*- from Western Europe to Greece; *ssp. petrisavii* and *ssp. meneghinianum* both from Eastern Europe and Western Asia, including coastal Mediterranean, coastal Turkey, Lebanon, Israel, Iraq and Caucasus) (Williams et al., 2001). Each of these was crossed to *T. repens* (Grasslands Huia) to assess the likelihood of producing viable offspring, indicative of consistent ancestral parental contribution of *T. nigrescens* (Williams et al., 2001). Only *ssp. nigrescens* produced a full seed set, suggesting a close relationship to *T. repens* (Williams et al., 2001). An ancestral contribution is also supported by a cross of Portuguese *T. nigrescens* and Irish *T. repens* (Marshall et al., 1995).

However, the ITS sequence of *ssp. nigrescens* is less similar to *T. repens* than is *T. occidentale* (Ellison et al., 2006). Similarity of *ssp. nigrescens* to *T. repens* is indirect

and likely caused by introgression, hybridization and an ancestral relationship between *T. nigrescens* and *T. occidentale* populations (Williams et al., 2001; Williams et al., 2008).

Proof of concept synthetic white clover

The ability of *T. repens* to cross with its putative paternal species *T. occidentale* is consistent with the parental contribution of *T. occidentale* to *T. repens*. This has allowed the breeding of a synthetic white clover plant, from the crossing of *T. pallescens* (mother) to *T. occidentale* (father) and the identification of the likely parents (Professor Warren Williams, pers. com, 2010).

1.1.4 Morphology of white clover

Diverse morphology is associated with variation in habitat (Davies et al., 1967; Caradus et al., 1990). A gradient of leaflet size is observed across the Mediterranean basin; this is linked with changes in temperature and soil moisture. In the north tiny/small sized leaflets are observed in the selections S.100 and S.184 and the Kent landrace from the British Isles (Davies et al., 1967; Caradus et al., 1990), while ladino white clovers from the Po Valley in Northern Italy represent the larger leaflet sizes seen in Southern Europe and the Mediterranean (Caradus et al., 1990).

Characteristic leaflet marking is largely genetically controlled (Williams, 1987c). The presence of the white V shape on each leaflet is dominant over the recessive absence of a marking (Corkill, 1971; Williams, 1987c). This character is polymorphic with multiple allelic forms and regulatory genes for expression (Corkill, 1971; Williams, 1987c). These control variation in the intensity of white or yellow colouration in the V mark, the position of the V mark relative to the leaflet apex and the shape of the V mark- presence/absence of the point and arms of the V and single or double V mark presence (Williams, 1987c).

Variation is observed in leaflet shape, ranging from ovate to obovate to obcordate with variations in between (Williams, 1987c). Red anthocyanin patterning under the control of the R locus alleles is observed on the surface of leaflets (Caranahan et al., 1955; Hovin and Gibson, 1961; Davies et al., 1963; Corkill, 1971). Phenotypes vary including; R^l (red leaf), R^m (red midrib), R^{ld} (diffuse red leaf), and R^f (red fleck),

(Williams, 1987c). In addition, some associations between the V locus and the R locus are known, for example, the halo markings determined by V^x_2 (Hovin and Gibson, 1961) and V^h_2 (Corkill, 1971). In my study the different alleles of the V and R loci were ignored as this was beyond the scope of the study.

1.1.5 Self incompatibility

White clover as an out crossing, insect pollinated species (Williams, 1987c) uses an S allele self incompatibility system to prevent self pollination (Atwood, 1942). The system works by stylar recognition of the S alleles present in each male gamete and inhibition of the pollen tube prior to fertilisation when both the style and pollen carry the same S allele (Williams, 1987c).

1.1.6 Diversity in white clover DNA

White clover can be separated from its close relatives by interspecies differences in chloroplast and nuclear markers (Watson et al., 2000; Ellison et al., 2006). However, intra-species DNA sequence diversity of white clover is not widely studied. Preliminary research suggests that sequence variation in the chloroplast: (*trnL* intron, and non-coding variable internal repeat region), and nuclear rDNA ITS, can identify intra species variation (Dr Nick Ellison and Pr Warren Williams, pers. com, 2010).

1.1.7 Distribution of white clover

Contemporary distributions consist of temperate or subtropical climates in the northern and southern hemispheres (Williams, 1987a; Ellison et al., 2006). The natural distribution encompasses the Mediterranean Basin, including North Africa, as well as northern Europe, Asia Minor, the Caucasus region and West Asia, extending possibly as far as North-West China. Farming practice and commerce have extended this distribution to other temperate and sub-tropical regions, especially in the southern hemisphere. No species of clover are found naturally in South East Asia or Australia (Ellison et al., 2006).

1.1.8 Breeding history of *T. repens*

White clover came to New Zealand via numerous imports of seed, from throughout Europe, but where exactly this clover seed came from is not known (Williams, 1983).

A wide assessment of New Zealand clover was conducted in 1928-1930 (Williams, 1983). This survey showed diversity in agreement with varied European origins and local population adaptations (Williams, 1983). The clover indentified was classed into four types. Type 1 was found to be preferable for farming use throughout New Zealand, and was certified in 1930 (Williams, 1983). Continuous improvements were made to this line at the Grasslands Research Centre in Palmerston North, until 1957 when a selection was made from the line and named Grasslands Huia (Williams, 1983). Initially little interest was paid to overseas clover resources, however around the late 1950s, interest was reinspired in use of overseas genetic material for clover improvement and the white clover breeding programme began assessment of populations from the Mediterranean region, including the Iberian Peninsula and the USA. (Williams, 1983; Professor Warren Williams, pers. com, 2009-2010).

1.2 Clover in General

1.2.1 The *Trifolium* genus

The *Trifolium* genus, composed of about 255 identified species (Ellison et al., 2006; Zohary and Heller, 1984) has a derived position within tribe *Trifolieae*, and is among the largest genera of the legume family (*Leguminosae*) (Ellison et al., 2006).

Assessment of chromosome number for consideration of the ancestral relationships between the species and the range of, and amount of, polyploidy in *Trifolium*, has found the basal or ancestral chromosome number is $x=8$ (Williams, 1987b; Ellison et al., 2006).

The genus has undergone several classifications, largely geographical that treated European, North American, African, Oriental and Russian species separately (Ascherson, 1906-1910; Komarov, 1934-1964; Gillett, 1952; Coombe, 1968). The most influential internationally recognized (but now replaced) classification has been the largely morphological classification of Zohary and Heller (1984). This classification grouped the species of the genus into eight, mainly European sections, *Lotoidea*, *Chronosemium*, *Paramesus*, *Trifolium*, *Trichocephalum*, *Mistyllus*, *Vesicaria*. Section *Involucrarium* was restricted to America and *Lotoidea* had some species in America. African distributions were noted for some species from sections *Trifolium* and *Mistyllus*.

This prior classification has been challenged by a comprehensive phylogenetic (DNA) analysis of the genus (218 species) using parsimony and Bayesian approaches (Ellison et al., 2006). This study found support for a monophyletic genus, and suggested a new intra-genus structure based on differences in nuclear ITS and chloroplast DNA sequences. This classification moved white clover to a new section *Trifoliastrum* and divided up the large section *Lotoidea*, relocating many species to different sections, without clearly resolving the subsection order of *Lotoidea* (Ellison et al., 2006). Support for this revision of the genus is found in the prior work of Steele and Wojciechowski (2003) who observed stronger support for a monophyletic nature of *Trifolium*.

In their assessment of 23 *Trifolium* species, conservation of multiple relationships between the species agreed with those obtained in the Ellison et al 2006 study (Steele and Wojciechowski., 2003; Ellison et al., 2006).

It is now suggested that individuals previously placed within section *Lotoidea*, and suggested by Zohary and Heller (1984) to form a primitive ancestral grouping, likely represent the ancestral species within the reclassified clades and sections. This is supported by the observed molecular heterogeneity and wide distribution (America, Africa and Europe) (Williams, 1987a) of the species previously allocated to the section (Ellison et al., 2006).

Section *Trifoliastrum*

The study by Ellison et al (2006) found sufficient DNA diversity among the clades of *Trifolium* that the *Trifoliastrum* clade (Fig. 1 ~20 Eurasian species) was deemed of sufficient diversity to be raised to section status (Fig. 1).

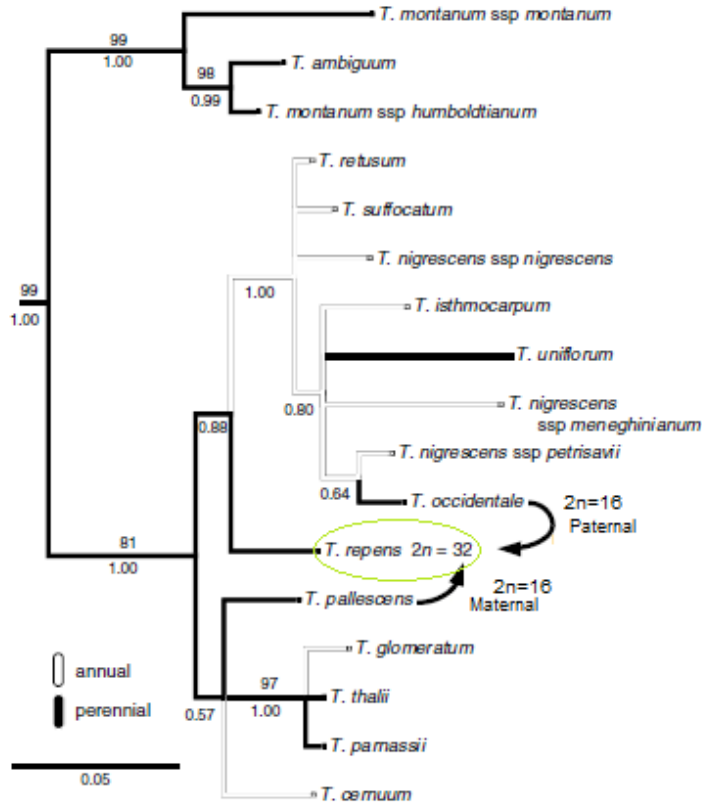


Figure.1 Adapted from Ellison et al. (2006), *Molecular phylogenetics of the clover genus (Trifolium-Leguminosae)*, Figure 6. The placement of *T. repens* within the section *Trifolium* is shown. Parental contributing species are labelled and diploid chromosome numbers are stated.

1.2.2 Origins of clover

Recent literature supports an origin for the *Trifolium* genus within the Mediterranean region, possibly during the early Miocene (Ellison et al., 2006). The Mediterranean region has the highest level of diversity within the genus, and over half of the *Trifolium* species (Williams, 1987a). It is suggested that North and South American species came from a shared origin, while sub-Saharan African species originated via three independent dispersal events (Ellison et al., 2006).

1.2.3 Spread of clover

Early domestication of clover (red and white clover) occurred in Southern Spain about 1000 AD (Kjaegaard, 2003). Wild populations of clover have been found throughout Europe prior to, and since, domestication. Clover was observed in Moorish Andalusia, Spain during Medieval times (Bodlens, 1981). By the mid 16th century clover was grown in Lombardy and the Netherlands (Kjaegaard, 2003). Domestication in France

occurred around 1583 (Kjaegaard, 2003). Clover was recorded in the British Isles around 1620, from a probable source in the Netherlands. By the end of the 17th century clover was widespread throughout England (Kjaegaard, 2003). However, domestication in Northern Europe was slower, with initial attempts failing in Denmark (1710 Fehmarn Island, 1730 Fehmard). In 1735 clover became first established in Fehmard and by 1740 it was widespread throughout the country (Kjaegaard, 2003). From Denmark, clover spread to Sweden and Finland, with domestication in Germany around 1645 (Kjaegaard, 2003).

1.2.4 Agricultural use

Clover (red and white) has been hailed as the “agricultural equivalent of coal” (Kjaegaard, 2003). This is because white clover through nitrogen fixation can increase nitrogen content in the soil, leading to healthier soil, better crops, better turf growth, healthier animals and healthier humans (Kjaegaard, 2003; Allen, 2008). This improvement in human health has been observed historically in Europe, during the 17th and 18th centuries where through use as fodder crop in English farming practice as part of the Norfolk rotation (turnips-barley-clover-wheat) white clover helped to produce better yield and healthier crops (Allen, 2008). Wide use in 19th century Europe allowed further agricultural benefits (Overton, 2002).

However during the agricultural revolution in 17th Century Europe, the development of the Haber-Bosch ammonia synthesis method for producing cheap nitrogen, saw farmers becoming increasingly reliant on a cost effective alternative; nitrogen fertilizer (Frame, 1987; Kjaegaard, 2003). Despite this, clover-based pastoral systems were continually developed and refined, especially in New Zealand and Australia. These still prevail, but have recently been largely replaced by Nitrogen fertilizer-based systems in dairy farming (Hemingway, 1999).

1.3 Appropriate genetic markers for use in phylogeny and diversity assessment

The main concern with phylogenetic and diversity assessment is the use of genetic markers that evolve at a rate of evolution appropriate for assessment of the history of the organism. The utility of non-coding, nuclear rDNA ITS repeats and chloroplast genetic regions (*trnL* intron, etc) have been widely known for some time (Eidesn et al., 2007; Volkov et al., 2007; Poczai and Hyvonen, 2010).

ITS

ITS is widely used for its ability to resolve inter and intra-species relationships. However, interpretation of organism history from the ITS regions is complicated by the inadequate understanding of the processes and actions of concerted evolution (Zimmer et al., 1980; Wendel et al., 1995; Buckler et al., 1997). Homogenization of parental forms, as in polyploidy, can obscure the organism history (Lihová et al., 2006) creating problems for assessment of inter and intra-specific diversity (Zimmer et al., 1980; Wendel et al., 1995; Buckler et al., 1997). Further, inadequate understanding of parental expression of the ITS region, knowledge of contributing parental relations (Volkov et al., 2007) and the role of molecular process during and post polyploid formation inhibits our understanding of the role of concerted evolution.

The ITS region appears to show uniparental expression despite biparental inheritance in white clover. No DNA sequences indicative of expression of *T. pallescens* NOR's have to date been observed in *T. repens* populations (Dr Nick Ellison, pers. com, 2009). Evidence from FISH results (Ansari et al., 1999) show NOR-bearing chromosomes of *T. repens* have banding patterns consistent with a *T. occidentale* origin, supporting the rare property of white clover to express only the ITS region from the suggested paternal parent. Furthermore, the allotetraployploid nature of white clover, means that finding a single copy gene for assessment of species history is difficult (Dr Nick Ellison, pers. com, 2009-2011), hence the ITS region is currently the most appropriate candidate single copy nuclear gene for evaluation of white clover species history.

Chloroplast genetic regions, *trn L* intron

The chloroplast *trnL* intron genetic region is well researched and accepted for diversity and phylogenetic inter and intra species assessment across numerous taxa. In general, haplotypes based on chloroplast DNA are advantageous for deduction of patterns suggestive of glacial influenced history (Tremblay and Schoen 1999; Abbott et al., 2000; Petit et al., 2002; Alsos et al., 2005; Koch et al., 2006). Recent work has proven the *trnL* intron sequence diverse and useful for understanding intra-species relationships in white clover (Doctor Nick Ellison, pers. com, 2009). Genetic assessment within European white oaks has shown the advantage of chloroplast DNA for assessment of post-glacial recolonizations and relationships (Kremer et al., 2002).

In out-crossing species such as clover, species history can occasionally be blurred as through chloroplast genetic markers being affected by inter-specific gene flow when introgression of foreign chloroplast DNA occurs through hybridization (Rieseberg and Soltis, 1991; Petit et al., 2002), and/or by the sharing of cytoplasm between unrelated sympatric species (Belahbib et al., 2001). However, it is likely that in white clover this contamination of maternal DNA lineages is an occasional event that may have featured in the history of the organism.

To avoid bias from the assessment of a single perspective of species history it is standard practice to look at multiple genetic regions and genetic perspectives, as this allows an impression of variation that is indicative of multiple time points in the evolutionary continuum of the species (Ellison et al., 2006; Volkov et al., 2007). The combined picture of chloroplast and nuclear perspectives allows assessment at typically shallow and deeper levels (Eidesen et al., 2007), while including exclusive maternal lineage assessment through chloroplast DNA. It permits the identification of multiple possible inter-specific hybridization events, through disagreement of chloroplast and nuclear data. This is widely and commonly observed in many polyploid species, for example, *Trifolium* (Ellison et al., 2006), *Cardamine asarifolia* (Lihová et al., 2006), *Vaccinium uliginosum* or Billberry (Eidesen et al., 2007), *Cardamine martima* group (Kucera et al., 2010).

1.3.2 Polyploidy

Hybridization often leads to the formation of a new polyploid species. There are two main types of polyploidy, autopolyploid and allopolyploid that refer to the difference in genomic chromosome composition of the individual (Soltis et al., 2003). An autopolyploid is produced by the somatic doubling or by inheritance of unreduced gametes from two individuals of the same species (Soltis et al., 2003). Allopolyploids are produced by gametes from two different species that then undergo doubling in the F1 generation, or by the coming together of unreduced gametes from different species (Soltis et al., 2003), as in *T. repens* (Williams, 1987b). Recognised examples of allopolyploids include the members of the *Brassicacae* genus, *B. napus*, *B. juncea* and *B. carinata* (Liu and Wang, 2006), relic and endemic Pyrenees *Borderea pyrenaica* (Gabriel et al., 2007) and recently formed Tragopogon allotetraploids *T. mirus* and *T. miscellus* (Soltis et al., 2003; Matyásek et al., 2007; Koh et al., 2010).

It is suggested that polyploidy is an important mechanism for speciation, and that it is more common than previously thought (Soltis et al., 2003) and is probably a part of the evolution of most organisms (Soltis et al., 2003; Volkov et al., 2007). This implies that the majority of species are the product of recent or ancient polyploidization, having undergone genome duplication at least once, (Salmon and Ainouche, 2010). It suggests that diploids are the product of polyploidization, at some prior point along their evolutionary trajectory (Soltis et al., 2003).

Polyploidy can obscure the species history and complicate genetic analysis through the doubling up of genetic complement. This is seen in recent genomic studies, where large amounts of gene redundancy have occurred leading to increased numbers of pseudo genes through suppression and inactivation of gene copies via methods such as DNA methylation. Polyploidization can also give rise to novel variation through the stimulation of transcriptional activity of retrotransposons and transposons (Zimmer et al., 1980; Dover and Tautz, 1986; Soltis et al., 2003).

As each polyploid system has its own unique way of responding and evolving under the process of polyploidization (Lihová et al., 2006), each polyploid species represents one way in which the process can affect speciation. The nature of polyploidization means *that* many polyploid species have the potential to have multiple origins (Soltis et al., 2003). Hence consideration must be given to what and who the parental and more distantly related contributing relatives were likely to have been (Marhold et al., 2004).

1.3.3 Polyploid models for comparison to white clover

No single species model could be found that might help to elucidate the processes of hybridization, radiation, post-glacial expansion and diversification of white clover DNA sequence patterns. However, from recent comparative studies of genetic variation, several species affected by glaciations and/or of known allopolyploid nature can be used for comparison to white clover, for construction of hypotheses around:

- **Formation/hybridization:** *Rorippa* (Marhold and Lihová 2006) recently formed *Tragopogon* species (Matyásek et al., 2007; Koh et al., 2010; Symonds et al., 2010), *Draba* species (Brochmann et al., 1992a, b; Koch and Al-Shehbaz, 2002)

Diplotaxis muralis (Warwick and Anderson 1997; Martí'n and Sa'nchez-Ye'lamo, 2000; Eschmann-Grupe et al., 2004).

- **Radiation:** *Borderea Pyrenaica* (Segarra-Moragues et al., 2007), *Diplotaxis muralis* (Eschmann-Grupe et al., 2004).
- **Post glacial expansion:** European white oaks (Petit et al., 2002), Billberry (Alsos et al., 2005), *Draba species* (Brochmann et al., 1992a, b; Koch, and Al-Shehbaz., 2002) *Diplotaxis* (Warwick and Anderson 1997; Martí'n and Sa'nchez-Ye'lamo, 2000).
- **Diversification:** *Rorippa* (Marhold and Lihová 2006), *Tragopogon* species (Matyásek et al., 2007; Koh et al., 2010, Symonds et al., 2010), *Diplotaxis muralis* (Eschmann-Grupe et al., 2004).

1.3.4 Concerted Evolution, an influence on DNA sequence diversity

Concerted evolution is the process by which parental inherited forms of DNA regions, for instance ITS rDNA, are homogenized via the mechanisms of recombination and gene conversion to produce the same DNA sequence throughout all the copies of the DNA region within the organism (Sang et al., 1995; Kovarik et al., 2005; Matyásek et al., 2007; Koh et al., 2010). The process occurs through gene regulation maintaining the different parental forms, followed by a progressive conversion of one parental type into that of the other. A gradient affect of gene conversion is observable where certain regions of the gene are more prone to conversion (Sang et al., 1995). It is thought the process of gene conversion occurs by the formation of heteroduplexes, which start in the 26S rRNA and randomly move in either direction along the DNA sequence, occasional creating errors, that result in areas of variation (Sang et al., 1995). Similarity of repeats within NORs is greater than between NORs, suggesting that intra-locus homogenization is of higher efficiency than inter-locus homogenization (Lichiterkonecki et al., 1994).

This means that maintenance of original inherited forms is under very strong influence, implying that concerted evolution acts very strongly locally, at least initially.

The process is commonly unidirectional and seemingly non-random. But the amount of homogenization and speed of occurrence in DNA sequence that can be expected, both within and between species is unknown (Matyásek et al., 2007). In addition, uniqueness

of each polyploid system; the kind of organism (natural or synthetic), and polyploid level affects the amount of variation observed. More importantly, it seems that this balance effect of keeping parental forms separate is lost at some point in the heterozygote and allopolyploid, when inter-locus homogenization occurs.

1.3.4.2 Bidirectional Concerted Evolution

Bidirectional concerted evolution, found in hybrids, is where both parental forms exist in the population (Kovarik et al., 2004; Matyásek et al., 2007). However, as the commonest parental form can vary between populations and over time, a degree of randomness or some regulatory mechanism is suggested (Matyásek et al., 2007). The allopolyploid species *Tragopogon mirus* shows evidence of population specific variation in occurrences of the parental *T. dubius* rDNA repeat, and reversal of parental favouritism to the *T. porrifolius* form (two individuals *T. mirus* population 2602) (Matyásek et al., 2007). Evidence of this is also observed in synthetic hybrids of *Nicotiana tomentosiformis* and *N. sylvestris* (Kovarik et al., 2004).

1.3.4.3 Factors affecting the observation of and effect of concerted evolution

Rate of concerted evolution

The rate at which concerted evolution acts in each organism affects the amount of homogeneity that occurs over time. In a recent study of nrITS regions from 33 species of the *Paeonia* genus, the long length of generation time and vegetative mode of reproduction are suggested to have slowed down the rate of nucleotide substitution resulting in a lower frequency of variation, causing high homogeneity in ITS sequences (Sang et al., 1995). Faster rates of change have been observed in studies of synthetic interspecific polyploids and hybrids (Kovarik et al., 2005; Marhold and Lihoviá, 2006), for instance, in the Brassica genus (Sang et al., 1995; Marhold and Lihoviá, 2006), in allopolyploids *Cardamine insueta* and *C. schulzii* (Franzke and Mummenho, 1999).

Organism age

There are several considerations for how the age of the polyploid species affects the action of concerted evolution and thus the nature of observable DNA sequence pattern types. For instance, it is suggested that if concerted evolution acts at a roughly

consistent rate across species within families and organisms of similar lifestyles, the more recent the age of the organism the less time concerted evolution has had to act and hence the more heterogeneous the organism may be. Under this scenario a population with variable levels of two or more parental DNA sequence types changes over time as is observed for the comparison of herbarium and natural populations of recently (within the last 80 years) and recurrently formed allopolyploids (*T. mirus* and *T. miscellus*). Here, a difference in bias of expression of parental forms of ITS sequences was observed in old and recent populations. Old herbarium hybrid samples had similar numbers of both parental rDNA repeats, while in present day populations of both allopolyploids, one different parental rDNA repeat dominated in each hybrid, suggesting a progression of selection toward parental rDNA *T. pratensis* and *T. porrifolius* forms and hence a decrease in heterogeneity (Kovarik et al., 2005). However this variation in present-day populations was observed in conjunction with incomplete homogenization suggesting the populations are still in a state of change (Kovarik et al., 2005).

If a population of species x starts with limited parental variation or is produced in a uniparental direction as in white clover, over time, errors in molecular processes in conjunction with subsequent interactions with other species or populations, might result in an increase in heterogeneity, producing species populations that are more diverse rather than showing signs of specific parental lineages and suggestion of concerted evolution. However, to achieve this, the rate of concerted evolution would have to be outweighed by the amount of variation observed across the assessed sample.

Time of origin

The origin of the species is the point from which concerted evolution can act to homogenise the species, hence, knowing when the species originated leads to understanding of the time frame over which concerted evolution has been able to affect the species. It has been shown that the allopolyploids of *Arabidopsis* and *Brassica*, of recent origin, show no signs of rDNA unit loss (Matyásek et al., 2007), suggesting that these polyploids are already highly homozygous, or that their rate of evolution is significantly faster, so that either concerted evolution is highly efficient (Matyásek et al., 2007) or insufficient time has passed for the effect of converted evolution to be observed (Bennett and Smith, 1991; Okane, 1996).

Epigenetic regulation and parental selection

Which parental form becomes favoured and which is lost/eliminated is an important part of the process speciation and evolution of the species (Volkov et al., 2007). Concerted evolution in allopolyploids is greatly affected by epigenetic regulation (Volkov et al., 2007). This occurs as “cross-talk”, when the inherited parental sub-genomes try to function together (Kato et al., 2005; Wendel and Adams, 2005; Volkov et al., 2007). The main importance of epigenetic regulation is the process of nucleolar dominance, the determination of which parental gene copies will be expressed and which will be lost or silenced (Chen and Pikaard, 1997; Matyásek et al., 2007). It has been suggested that favouritism of a specific parental form occurs post allopolyploid formation through epigenetic regulation, of nucleolar dominance, gene copy number and transcription regulation through DNA modifications that affect expression of parental rDNA (Volkov et al., 2007). This is seen in allotetraploid *T. mirus* and *T. miscellus* where it is thought that epigenetic mechanisms up-regulated the expression of the *T. dubius* parental form to compensate for low copy number (Matyásek et al., 2007). In T5 *Glycine tomentella* multiple generations are required before parental expression becomes fixed (Joly et al., 2004). This raises the question of whether homogenization is ever complete or if there always are a proportion of alternative forms present (Volkov et al., 1996, Volkov et al., 1999; Volkov et al., 2003).

1.3.5 Reticulate evolution

Reticulate evolution, the process by which repeated hybridization of and formation of polyploids occurs presents a further concern for assessment of genetic diversity in the polyploid. Extensive and possibly ancient reticulate evolution (Stebbins, 1938, 1948; Tzandoudakis, 1983) is observed in *Paeonia* (Sang et al., 1995), and also in tetraploid *Tragopogon* allopolyploids *T. mirus* and *T. miscellus*, with respective parents, *T. dubius* with *T. pratensis* and *T. porrifolius* (Kovarík et al., 2005).

1.4 Glaciations and affect on species and populations

The last glacial period, the quaternary glaciations, or the Pleistocene glaciations, is the period from about 2.5mya to the present. During the colder periods of these glaciations most of Europe was covered by ice. The Iberian, Balkan (Lihoviá et al., 2000; Brewer et al., 2002; Marhold and Lihoviá, 2006) and Italian peninsulas (Hewitt, 2000; Petit et al., 2002; Brewer et al., 2002) are suggested to have been European refugia. It is thought

that parts of the European Alps, within France, Austria and Italy experienced, an isolated ice age, induced by cold climate and altitude (Hughes, 2010). It has been suggested that the Iberian refugia, being geographically separated from the Italian and Balkan refugia, should be genetically distinctive (Brewer et al., 2002).

Evidence of these glacial refugia comes from genetic assessment of European white oak chloroplast lineages is observed in Italian and Balkan refugia (Petit et al., 2002). It is suggested that pollen flow during warmer climates reduced the nuclear genetic differences between populations of European white oaks from eastern and western refugia (Kremer et al., 2002). This study also found that eastern lineages showed more allelic variation but lower heterozygosity compared to western refugia (Kremer et al., 2002). Further evidence for Iberian and Balkan Peninsula as glacial refugia is found among relic populations of subspecies of *Cardamine amara*, Iberian (subsp. *olotensis* and subsp. *pyrenaea*), Balkan subsp. *balcanica* (Lihová et al., 2000) and supported by high diversity in polyploid *Micrathlaspi perfoliatum* (Koch, and Bernhardt, 2004).

Throughout the interglacial periods, it is hypothesised that the retreat of ice led to range expansion across Europe, and consequently species expansion. Many species of Northern Europe show evidence of large scale migrations and recolonization of territories previously inaccessible due to glaciations (Larena et al., 2002). Species within Europe appear to have survived and diversified without the same degree of geographical spread, implying the existence of numerous glacial refugia (Larena et al., 2002) in agreement with the warmer temperatures observed in Southern Europe (Petit et al., 2002).

As species migrate during warmer periods, change their altitude and colonize new areas, zones of secondary contact form. Within areas of secondary contact, species able to interbreed experience a “melting pot” scenario (Petit et al., 2003; Marhold and Lihoviá, 2006), where gene flow between the species changes the genetic variation within each population, altering the species history (Marhold and Lihoviá, 2006; Parisod and Besnard, 2007).

This change in genetic variation/diversity can lead to the formation of new hybrid species (Lihoviá et al., 2000; Marhold and Lihoviá, 2006). The Pyrenees endemic hybrid, *Armeria filicaulis* ssp. *nevadersis* shows commonality in chloroplast haplotypes

between *A. spendens*, *A. villosa ssp. bernisii*, and *A. filicaulis ssp. nevadersis* that suggest altitude migrations and secondary contact as causative for the hybrid formation (Larena et al., 2002). Ploidity in the *Brassicaceae* (Marhold and Lihoviá, 2006) and wide introgression in *Brassicaceae* (*Cardamine*, *Rorippa*, Marhold and Lihoviá, 2006), and in European white oak species (Petit et al., 2002) provide further examples.

It is hypothesised that altitude and/or latitude migrations caused by glaciations brought either the recent progenitors or the parents of modern white clover into contact in a glacial refugia, allowing the formation of the hybrid species. During early interglacial periods, it is suggested that meeting in secondary contact zones contributed further genetic variation to this novel hybrid species and/or resulted in the formation of, or repeated formation of, hybrid individuals. During this time there was potential for introgression from related species such as *T. nigrescens*. This resultant hybrid species has since diversified and expanded beyond the preferred altitude/latitudes of its parental species, evolving to become the modern white clover species.

1.4.2 Glacial hypotheses

When considering how species are affected by glaciations, two hypotheses are essential to discussion of glacial survival and postglacial colonization; these are the Nunatak and the Tabula rasa hypotheses.

The Nunatak Hypothesis

The Nunatak hypothesis refers to species/population's survival and evolution *in situ*, in isolated, non-glaciated areas surrounded by a glacial environment (Bauert et al., 1998; Stehlik, 2000; Gugerli & Holderegger, 2001; Stehlik et al., 2001, 2002a; Gabriel et al 2007). It is suggested that due to their isolation, species within these areas typically show low genetic diversity but high population differentiation, associated with inbreeding, recurrent bottlenecks and lineage sorting leading to higher homogeneity (Gabriel et al 2007). These populations are expected to evolve freely, becoming more differentiated from the original population. However, the Nunatak hypothesis seems unfounded for white clover, as it would require that the coastal dwelling *T. occidentale* move into cooler environments during periods of climatic cooling.

The Tabula rasa hypothesis

The Tabula rasa hypothesis describes the scenario of postglacial migration of species to environments not previously available. It is thought populations in recently recolonized areas show a decrease in genetic diversity, relative to the glacial refugia from which they have come (Gabriel et al 2007). The level of heterogeneity in populations along the line of migration, of geographical descent, can be used as a measure of age of the relative populations (Gabriel et al 2007). Northern Pyrenees endemic populations of *Borderea pyrenaica* show decreased genetic variation relative to older southern/Pre-Pyrenees populations in line with Tabula rasa hypothesis. The common beech (*Fagus sylvatica* L) is another example of a species that shows the affect of the tabula rasa hypothesis (Demesure et al., 1996).

1.4.3 Interpreting species genetic diversity in relation to glaciations

In order to interpret the genetic diversity within a species affected by glaciations and draw inference about species origins, several factors need to be considered. Is the history of a species describable by a single nunatak or tabula rasa hypothesis? The best description of species history is likely to be a combination of hypotheses that act in species-specific ways to give a unique history (Berg, 1963; Brochmann et al., 2003).

It is expected that in areas of glacial refugia “regional hot spots of allelic diversity” will be observed (Widmer and Lexer, 2001). A gradient of genetic diversity “a long-lasting genetic signature” (Parisod and Besnard, 2007), is expected to be observable from the area of glacial refuge to current populations. The level of genetic diversity should decrease with the increased distance travelled from the glacial refuge (Hewitt and Ibrahim 2001; Parisod and Besnard, 2007), as observed for *Borderea pyrenaica* (Gabriel et al 2007).

However, this gradient is often not clearly observable being affected by the time past since the glacial refuge, the species mating system, migratory habits, contact with other species, dispersal habits and to what extent human intervention has aided the spread of the species. For instance, *Cardamine insueta* is a recent hybrid for which farming practice has greatly extended its range (Franzke and Mummenho, 1999). The selection and subsequent distribution of numerous seemingly successful white clover cultivars (e.g. Huia, Crusader, Tammisto), and ecotype selections (Crau, Lodi and Kent) around

the world in recent times, and dispersal of white clover seed during the beginning of agriculture in Europe (Kjaegaard, 2003), represents a serious concern for assessment of “natural” genetic diversity in white clover.

Identification of species histories and paths of radiation/descent is obscured by the degree of species interaction (breeding with one another) along the migration gradient, their contact in secondary zones, and occurrence of “melting pots” (Petit et al., 2003; Parisod and Besnard, 2007), for example in European oaks (Petit et al., 2002). It should be noted that areas of high genetic diversity can falsely identify species as having been affected by glaciations (Petit et al., 2002).

Furthermore, observations and interpretations of concerted evolution are affected by the effects of climatic changes on population structure, heterozygosity vs homozygosity, population size and genetic variability (Willis and Niklas, 2004; Hampe and Petit, 2005). Geographical barriers such as mountain ranges prevent dispersal to new environments, prevent admixing of species or populations and consequently the formation of novel species (Gugerli et al., 2001; Mátyá and Sperisen, 2001), further obscuring interpretations.

For assessment of the radiation of white clover and hypotheses of origin, the Mountain ranges throughout Europe, namely the Pre-Pyrenees, Pyrenees (Larena et al., 2002; Segarra-Moragues et al., 2007) and Alps (Raffl et al., 2008), present physical barriers to be crossed and are areas of interest. The assessment of diversity gradients within these and extending regions as well as altitudinal and latitudinal movement will be required to provide answers for radiation assessment of white clover.

1.5 Aims and Hypotheses

1.5.1 Aims

This survey of genetic diversity set out to achieve the following aims:

- ❖ To determine whether there is measurable intra-species diversity among and within natural and commercial populations of *T. repens*, from Zone 1: Iberian Peninsula Portugal and Spain, and from Zone 2: Mediterranean, Continental Europe, UK, Near East, and North Africa.
- ❖ To attempt to formulate a model of species radiation and origin by:
 - Determining the diversity and spread of DNA sequence variation in maternally inherited chloroplast lineages;
 - Determining the diversity and spread of ribosomal nuclear ITS regions of paternally expressed *T. occidentale* (NOR's); and
 - Comparing the diversity and spread of the combination of the chloroplast and nuclear pattern types.

1.5.2 Hypotheses

- ❖ Chloroplast & nuclear genetic markers can give a picture of the radiation and origins of *T. repens*.
- ❖ The rates of evolution of the genetic markers used are suitable for resolution of intra species genetic diversity.
- ❖ Observed diversity reflects population geography and origins.
- ❖ Natural and commercial populations of *T. repens* can be separated based on DNA sequence variation.
- ❖ The formation of *T. repens* was unidirectional from *T. palleescens* (maternal) crossed with *T. occidentale* (paternal).
- ❖ If multiple origins had occurred evidence of *T. palleescens* ITS sequences would be expected, in some populations.
- ❖ Comparison of the chloroplast and nuclear data taken together will give a more comprehensive indication of the radiation and potential origins of the *T. repens*, than either the chloroplast or nuclear data taken separately.

Chapter 2 Methods

2.1 Accession selection and growing of plants

For the purposes of this experiment, forty-one accessions of *Trifolium repens* were obtained from the Margot Forde Germplasm Centre, AgResearch, Palmerston North. Of these, thirty-three were freshly grown from seed and the remainder were available as mature plants at the Grasslands Research Centre. The sources of the accessions and numbers of plants are given in Table 2.1.

Twenty-seven accessions were local collections from South Western Europe (France, Spain and Portugal) and North Africa (Algeria and Morocco).

Ten commercial populations were included to represent known populations from defined geographical regions. These included Morso - (Denmark), Tamar and Haifa - (Israel), Barbian - (The Netherlands), Ovcak - (Czech Republic), Regal - (USA Lodi), S.184 - (Britain), Lodi gigante Lodigiano - (Italy), Crau - (France) and Kent White/Wild White - (England). Four commercial cultivars termed “mixed” were used. Three of these cultivars contain genetic material from two separate geographic populations; Aran- (France/Israel), Crusader – (France-Crau/Syrian), and Blanca- (Belgium/Kersey). The fourth mixed cultivar, Dusi, produced from 78 lines introduced into South Africa; was used as an out-group commercial population. Inclusion of these commercial populations allowed assessment of the genetic variation in DNA sequences in populations bred for high agronomic performance across a range of environments. From the assessment and comparison of genetic diversity of commercial populations (geographic and mixed) and natural populations, ideas of radiation and evolutionary history of the species of white clover were explored. Commercial populations provided evidence of more recently distributed populations and thus more recent patterns in genetic diversity.

Table 2.1 Natural populations, non-commercial populations.

Accession Number	Country	Place of origin	No. Plants grown	No. plants sampled
C20859	Algeria	Algeria	10	10
C1751	Algeria	Algeria	30	10
C20499	Algeria	Miliana- Affreville rd	1	1
C6452	Algeria	-	5	5
C6452	Algeria	-	10	10
C1761	Morocco	-	10	7
C5897	Italy	Milan, Italy	18	18
C8979	Portugal	Ribatejo, Grândola	10	10
C8976	Portugal	Algarve	9	9
C9514	Portugal	Trans Os Montes	11	11
C7582	Portugal	Beira litoral	10	10
C8985	Portugal	Coimbra	13	13
C20477	Portugal	Agus de Mouco	7	7
C7576	Portugal	Estremadura	9	9
C9516	Portugal	Tras Os Montes, Vila Verde	10	10
C9024	Portugal	Tras Os Montes, Espinho	10	10
AZ4838	Portugal	Minho, near Braga, Ferreiras	10	4
AZ4839	Portugal	Tras Montes, near Baragonça	10	5
C7600	Spain	Pyrenees Mts	8	8
C8994	Spain	Asturias	12	12
C9034	Spain	Galicia	13	13
C9025	Spain	Extremadura	14	14
C8975	Spain	Andalucía	10	6
C9037	Spain	Galicia	11	11
C9531	Spain	Ponferrada, León & Castile	10	10
C9027	Spain	Pontevedra, Galicia	10	10
C9046	France	Mouriès, Bouches du Rhône	11	11
C7690	France	Entressen, Bouches du Rhône	8	8
C6374	England	(Kent White), Kent, Suffolk	10	10
C6304	England	(Wild White), Kent, Suffolk	10	10

Table 2.2 Commercial clover populations used in this study

Commercial names	Accession number	Classification	No. plants sampled	Represented Location
Morso	C2095	Ecotype	12	Old Danish ecotype
Tamar	C4143	Selection	11	Israeli populations
Tammisto	C960	Ecotype selection	9	Finish ecotypes
Crau	C18951	Ecotype selection	11	Southern France Crau
Barbian	24247	Ecotype selection	12	Old Dutch ecotype, (Yamada et al., 1987)
Ovcak	C7555	Ecotype selection	9	Czech Republic, (Southern Bohemian ecotype)
Aran	-	Mixed cultivar	10	France and Israel
Crusader	-	Mixed cultivar	10	South France (Crau) and Syrian ecotype Mader Village, Zebdani district
Dusi	C16511	Mixed cultivar	6	South Africa, (local lines + 78 introduced cultivars)
Blanca	C6360	Mixed cultivar	10	Belgium/Kersey cross with subsequent English contribution
Regal	C15111	Ecotype selection	7	USA adapted Italian Lodi
S.184 (ABERYSTWYTH)	C21105	Mixed cultivar	2	Wild British populations
Haifa	C5892	Ecotype	5	Huleh District North East Israel
(Italian Lodi) Giganteum	C5897	Ecotype	14	Po Valley, Northern Italy
Lodigiano				
Huia (Grasslands) or New Zealand certified Mother	-	Cultivar	1	New Zealand
Kent	C6374/C6304	Ecotype	19	Kent Suffolk, England

Table 2.3 Accessions included as DNA sequence data from Dr Nick Ellison

Accession Number	Country	Place of origin or commercial name	No. Plants sampled
C7648	Portugal	Northern Portugal	7
C11519	Portugal	Vila Pouca de Aguiar, Trás-os-Montes District	3
C7651	Portugal	Bobadelo, Tras Os Montes	1
C9024	Portugal	Tras Os Montes Espinho	3
C9516	Portugal	Tras Os Montes, Vila Verde	3
C9027	Spain	Pontevedra, Galicia	3
C9531	Spain	Ponferrada, León & Castile	3
C2418	France	Le Conquet, Finistère department, Bretagne	3
C2095	Denmark	Morso	3
TR14	Norway	-	1
TR08	Germany	Hausen, Baden-Wurttemberg	1
TR10	Greece	Pogoniani, Ioannina, Epirus	1
TR06	Romania	near Domnesti	1
N9	Romania	Barro Trro	1
C6452	Algeria	-	3
C25464	Morocco	Marrakech, near the Atlas Mountains	1
C25465	Morocco	Chefchaouen, Rif Mountains	1
C25466	Morocco	Bab-Berred, Chefchaouen provence	1
C25467	Morocco	Ketama	1
C25468	Morocco	South of Taza	1
C25469	Morocco	South of Azrou, Meknès-Tafilalet region	1
C25470	Morocco	Ain-Leuh, Ifrane Province, Meknès-Tafilalet	1
C25471	Morocco	Oukaïmeden	1
		South of Bou-Laouane, near the Atlas	
C25472	Morocco	mouuntains.	1
C25473	Morocco	South of Asni, near Marrakech	1
C25474	Morocco	North of Asni, near Marrakech	1
		North of Irhem, provience Ouarzazate, Souss-	
C25475	Morocco	Massa-Draâ	1
C25476	Morocco	Taddert	1
C25477	Morocco	Taddert	1
C25478	Morocco	Setti Fatma, near Marrakech	1
C25479	Morocco	Imlil, high in the Atlas Mountains	1
C25480	Morocco	South of Asni, near Marrakech	1
TR04	Morocco	near Marrakech	1

Table 2.3 (continued) Accessions included as DNA sequence data from Dr Nick Ellison

	Country	Place of origin	No. Plants sampled
C5897	Italy	Milan, Po Valley	3
TR11	Italy	Pasubio	1
TR01	Italy	-	1
C6304	England	Kent, Suffolk	3
Totals			66

Zone 3			
Individuals			
CR138941	Costa Rica	Costa Rica	1
CR	Costa Rica	Costa Rica	1
TR03	USA	-	1
WP unknown	USA	Georgia	1
TR07	Ethiopia	-	1
TR76	China	near Urumqi, Xinjiang	1
TR05	China	-	1
TR09	Japan	-	1
TR12	Kazakhstan	near Alma-Ata, Alma-Ata	1
Totals			9

2.1.1 Seed germination techniques

Fifteen seeds from each of the accessions were scarified on sand paper to crack the outer seed coat using the methods of Williams, (2001). Seeds were placed on moist filter paper in petri dishes. If the scarifying process had worked sufficiently, it was expected that after an hour the seeds would have increased in size by roughly five times. Consequently, any seeds that had not changed in size after an hour were rescarified.

Petri dishes were placed on a warm window sill and left to germinate. It was expected that germination would take between 2 and 10 days. The difference in time taken to germinate was attributed to the quality and age of the seed.

2.1.2 Growing conditions and methods

Once seedlings were about one centimetre in length and had two green cotyledons they were transferred to trays of potting mix, with 15 seedlings per half tray. Each tray contained seedlings from two accessions. Seedlings grew in the trays for about 4 months after which they were transferred to individual pots of potting mix and labelled.

All plants were grown in the glasshouse under ambient photoperiod and temperatures. After about 7 months post germination plants were transferred outside, with the exception of the six plants from accession C20859 from Algeria. This accession was transferred to a separate glasshouse and maintained in a warm environment, allowing for crossing experiments to be carried out.

Cultivars assessed in the second round of sampling were grown in the glasshouse under the same conditions as initial plants for about a month post germination.

A morphology scoring was performed on all plants (initial accessions and commercial varieties). At this time point the initial accessions were 8.5 months and commercial varieties were 6 weeks post germination. After morphology scoring all plants of the commercial varieties were placed outside with those of the initial accessions.

DNA Extraction

- Reagents used in the DNA extraction process

Table 2.4 Reagents used in the DNA extraction Protocol

Reagents	Amount used per extraction of an individual sample	
Stewart buffer	400 μ L	Laboratory produced
Chloroform/ Octanol (24:1)	300 μ L	Laboratory produced from stock solution
Isopropanol	300-350 μ L	Laboratory supplied
Ethanol (80%)	500 μ L	Laboratory produced
Milli Q water	50 μ L	Laboratory produced

2.1.2 DNA extraction Protocol

White clover plantlets were left to grow until each plantlet had at least two leaves that were no smaller than half the size of a five cent piece (approximately 1cm in diameter), about 6 weeks post germination in the glasshouse.

The DNA extraction method was as used by (Ellison et al., 2006) for fresh or dried leaf material, with the following modifications.

Fresh leaf samples were ground by use of two ceramic grinding beads in the presence of 400uL of Stewart extraction buffer, using a mini prep machine, two rounds of grinding were conducted for 30 seconds at a speed of 4.

Centrifuge tubes were placed in a hot block at 65°C for as little as 10 minutes and up to a few hours. No difference was observed in results with the length of the incubation period. To wash the DNA pellets 500ul of 80% ethanol was used. DNA pellets were re-suspended by vortexing for 10 seconds before and after 5 minutes heating in a hot block at 65°C. All samples were checked to see that no remnants of white DNA pellet were visible. These DNA extraction products were then used as templates for the PCR reactions.

2.2 PCR

- PCR primers

Table 2.5 PCR primers and their source

Primer Name	Sequence	Region Amplified	Source
Chloroplast			
Universal C	CGAAATCGGTAGACGCTACG	Non coding DNA region, trn L intron	Taberlet et al., (1991)
Universal D	GGGGATAGAGGGACTTGAAC	Non coding DNA region, trn L intron	Taberlet et al., (1991)
IRB21NE-F	GAGGTCTGGTTCAAGTCCAGGATG	Shortened sequence from internal repeat region	Dr. Nick Ellison, personal communication (2009)
IRB21NE-R	ATAAGCGGACTCGAACCGCTGACATC	Shortened sequence from internal repeat region	Dr. Nick Ellison, personal communication (2009)
Nuclear			
Ec1	GAGGAAGGAGAAGTCGTAAC	(ITS1- 5.8s- ITS2), rDNA internally transcribed spacer region	Williams et al., (2001)
Ec2	GTTCGCTCGCCGTTACTAAG	(ITS1- 5.8s- ITS2), rDNA internally transcribed spacer region	Williams et al., (2001)

Table 2. 6 PCR Reagents

Reagents	Amount used per 20 μ L PCR	Manufacturer
Reaction Buffer No MgCl ₂	2 μ L	BIOLINE Pharmacia Bio Tech
dNTPS 10mM (working solution)	0.4 μ L	(4x individual stock solutions)
MgCl ₂ (25mM)	1.2 μ L	Thermo Scientific
Taq (BIOTAQ Red DNA Pol.) 1 μ / μ L	0.2 μ L	BIOLINE
OR Taq (Red HO DNA Pol.)	0.2 μ L	ABGENE.com Pharmacia Bio Tech
Primer mix	0.4 μ L	(4x individual stock solutions)
Milli Q H ₂ O	14.8 μ L	Laboratory produced
DNA template	1 μ L	Processed gel purified DNA products

2.2.1 PCR protocol

The three sets of primers were used (Table 3) in 20 μ L PCR reactions containing 1 μ L of DNA template to 19 μ L of PCR reaction mix. This reaction mix was made from 14.8 μ L of milli-Q water, 2 μ L buffer, 1.2 μ L of MgCl₂, 0.4 μ L of dNTPs, 0.4 μ L of primer and 0.2 μ L of Taq polymerase. Two forms of Taq polymerase were used in the experiments due to laboratory circumstances. Initial PCR's were performed with Red HOT DNA polymerase from ABGENE.Com. The majority of PCR's were carried out using BIO TAQ Red DNA Polymerase at a concentration of 1 μ / μ L, (BIOLINE).

In subsequent PCR's MgCl₂ was included in the buffer, the amount of water used was increased to 16 μ L, to maintain the same concentration of MgCl₂ as in prior reactions. Both variations in methodology produced equal quality of PCR product.

All PCR reactions were performed on either the PTC-200, Peltier Thermal cycler, MJ Research, Gene Works, South Australia or the Applied Biosystems model 27,00 thermal cycler. The same PCR program ITS of 2 ¼ hours duration was used for all three sets of reactions. This programme was designed by Dr Nick Ellison to give optimal results for the primer set Ec1/ Ec2 and was known to work sufficiently for other primers sets including the remaining two used in this experiment.

The ITS program used one cycle at 96°C and 39 subsequent cycles of amplification corresponding to; 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute, followed by an extension phase of 7 minutes at 72°C.

2.2.2 Gel purification protocol

The gel purification protocol used was the same as employed in Williams et al., (2001) . However, the addition of 6x loading dye and the subsequent centrifugation step was omitted when using one brand of Taq polymerase, (BIO TAQ Red DNA Polymerase) as this Taq polymerase fluoresced without addition of the 6x loading dye. Observed bands showed no difference in intensity of colour from those treated with loading dye and those treated without, all were of the expected sizes.

The New Zealand white clover cultivar, Grasslands Huia was processed alongside samples in PCR reactions and gel purification.

Table 2.7 Reagents used in the gel purification protocol

Reagent	Manufacturer
TTE Buffer	Laboratory produced
SeqPlaque® agarose	Cambrex, Bio Science, Rockland, USA
Liquid Nitrogen	

2.3 Sequencing

-Reagents for sequencing

Table 2.8 Reagents used in the sequencing protocol

Reagents	Amount per 10 μ L sequencing reactions	Manufacturer
TM 5x buffer	2.5 μ L	Laboratory produced
DMSO	1 μ L	Aldrich Chemical company, Inc, USA
Primer	1 μ L	Laboratory produced
Seq Mix	0.5 μ L	Massey University Sequencing service
DNA template	5 μ L	Gel purified DNA extraction product

Each sequencing reaction contained 2.5 μ L buffer, 1 μ L of primer, 1 μ L of DMSO, 0.5 μ L sequencing mix. This sequencing reaction mix was then multiplied by the just over the number of samples to be processed to allow for pipetting error. For example if 30 samples were to be processed then the mix was made for 35 samples. Grasslands Huia DNA was included as a negative control in all sequencing test runs and as checking tool for reactions where sequencing was questionable.

2.3.1 Sequencing Protocol

Step 1: Generating sequence reactions

Sequencing reactions were performed on SeqPlaque® low melting point agarose. (Cambrex, Bio Science, Rockland, USA) gel purified PCR products (see section 2.2 gel purification), at a ratio of 50% sequence reaction mix to 50% PCR product template. Reactions were placed in Qiagen PCR reaction stripes or 0.2 mL eppendorf tubes. Sequencing reactions were then carried out on either the PTC-200, Peltier Thermal cycler, MJ Research, Gene Works, South Australia or the Applied Biosystems model 27,00 thermal cycler, under one of the programmes Big dye x99 or Seq 100.

Seq 100 programme:

96°C 14 seconds, 50°C for 15 seconds, 60°C for 4 minutes.

This was repeated for ninety-nine cycles and run for 7½ hours duration.

BIG Dye x 99 programme:

96°C 10 seconds, 50°C for 5 seconds, 60°C for 4 minutes

This was repeated for ninety-nine cycles and ran for about 7hours.

The difference between the programs was considered irrelevant for analysis and was ignored.

Step 2: Preparation for sequencing

This used the method of Ellison, et al., (2006) with the modification that 96 column multi screen plate (Millipore Corporation) with a 45 µL column loader was used to load GE Healthcare illustra™ Sephadex™ G-50 Fine DNA grade (GE Healthcare UK Limited little Chalfont Buckinghamshire HP7 9NA UK) into the 96 column multi screen plate. 300 µL of milli Q water was added to each column, and the plate was covered with parafilm and placed in the fridge for a minimum of three hours or overnight prior to processing.

This system allows for processing of up to 96 individual samples or sequence generated from both primers for 48 individual samples.

This plate was then spun in a Centrifuge 5810 R, (Eppendorf, Germany) at 910 G for 5 minutes, to prepare columns for use. 10 µL of cresol red was added to each individual

sequencing reaction. These sequencing reactions were spun for 40 seconds at 1000rpm in the Centrifuge 5810 R, (Eppendorf, Germany). Samples were carefully transferred according to a designed grid formation to the spun Sephadex column multi screen plate and this plate placed on a 96 well collection plate. These two plates were spun at 910G for 5 minutes to allow for collection of purified DNA sequence reaction products.

The collection plate and a balance were spun in a DNA Speedy Vac set on high temperature for an hour. The drying time was increased to 2-3hours when slipping of the column plate relative to the collection plate occurred during the centrifugation process. This slippage results in uneven distribution of the collected liquid between the wells, requiring further drying to remove the excess liquid from these wells. The used Sephadex columns were discarded and the 96 column multi screen plate washed and reused.

Step 3: Resuspension protocol

10 μ L of formide was added to each well of the dried collection plate that contained the sequenced DNA to be read. To plate wells not containing DNA 10 μ L of Milli Q water was added. The collection plate was covered with a lightweight sticky plastic seal and spun at 910G for 40 seconds in the Centrifuge 5810 R, (Eppendorf, Germany), vortexed for 10 seconds on a plate vortex and spun at 1000rpm for a further 40 seconds.

Step 4: Reading the DNA sequence

The DNA sequences were read from the collection plate containing resuspended sequenced DNA from a plate record by an ABI 3100 genotyper (Applied Bio systems Easter City, California) based at AgResearch, using a POP7 polymer in a 22cm array. The run module parameters were a Base caller-3100POP4UR, with default settings for Basecaller setting and Dye set/Primer file: DT3730POP7(BDv3).

Table 2.7 Run module and parameters for the sequencing runs

Run Module Seq22_POP7_40si_prerun_def	
Parameter Name	Value
Run Temperature	60
Cap Fill Volume	88
Current Tolerance	100
Run Current	100
Voltage Tolerance	0.6
Pre Run Voltage	7
Pre Run Time	60
Injection Voltage	1
Injection Time	40
Run Voltage	7
Number of Steps	5
Voltage Step Interval	5
Data Delay Time	400
Run Time	1300

The raw files generated from the sequence reading were extracted using the provided software *Sequence analysis*, (Applied Bio systems Foster City, California). This showed that the average spacing between peaks on the extracted sequence file was ten sample points.

2.4 Chromatogram analysis and contig assembly

The extracted sequence data were processed via the software Sequencer v. 4.6, (Gene Codes Corporation, Ann Arbor, MI, 2008). The chromatogram of the sequence run generated from each primer for each sample was analyzed prior to contig assembly to assess the likelihood of a good quality contig being produced.

The sequence reads of satisfactory quality were assembled via the automatic contig function of Sequencer v. 4.6. A default setting of 85% similarity and 25% minimum overlap under the dirty sequence setting were used as a further quality test of the generated sequence. These parameters sufficed for sequences of the highest quality. Where sequences were of poorer quality the above parameters were decreased sequentially by 5% respectively as required e.g. (85%, 25%) decreased to (80%, 20%),

to (75%, 15%) etc to allow sequence reads to be aligned. A cut off of 65% similarity and 10% minimum overlap was taken as the minimum standard of acceptable quality of sequence data for analysis. Those sequence reads that failed to meet the standard of sequence quality required were resequenced and reread.

Where resequencing failed to produce sufficient quality of sequence and other likely causes of poor quality were ruled out, PCR and gel purification steps and sequencing were repeated for these samples.

All contigs were individually and manually processed to remove sequence ambiguities and aligned to a reference sequence. The reference sequences used were from Grasslands Huia for the universal C/D primers and Ec1/Ec2 primers. The *Trn L* sequence from

T. repens strain="USDA NPGS PI 376882/ C2478 (Grasslands) was used as the IRB21NE reference, (See Alignment IRB21NE, on the CD) All generated contigs were manually checked in alignments generated in Sequencer. Comparison was made of differences in DNA sequence across multiple individuals per accession and across different accessions.

Initially sequences generated from the nuclear Ec1/Ec2 primer pair proved more difficult to obtain and were of poorer quality. Several small experiments were conducted to resolve this. However despite obtaining relatively bright PCR product bands, the quality of these initial DNA extractions was the likely cause for poor sequencing, possibly coupled with sensitivity of the nuclear genetic region, (amount of DNA variation and nucleotide length sequenced ~ 700 bps) relative to more homogeneous chloroplast regions. Hence where samples produced poor quality for all three genetic regions, the DNA extractions and subsequent processing was repeated.

2.4.1 Alignments and tabulation

To visualize DNA diversity within the data set and to check for DNA sequence ambiguities the consensus sequence from each contig was imported into the program BioEdit (Hall, T., Ibis Therapeutics, a division of Isis Pharmaceuticals Carlsbad, CA, 92008). The colour and the conservation function of this program allowed easy viewing of ambiguities across the data set, providing a further stage of checking for true DNA diversity (Appendices on the CD for alignments). All observed diversity was recorded in an excel spreadsheet by accession and plant number, per genetic marker, per variable site, (Appendices A-C).

2.5 Diversity analysis

2.5.1 TCS 1.2.1 Phylogenetic network estimation using statistical parsimony

The software TCS, (Clement et al., 2000) was used to carry out a series of statistical parsimony analyses of the genetic variation within the dataset. Analyses were performed on combinations of shortened contiguous alignments of all three genetic regions as well as separate analyses of the chloroplast regions (*trn L* and IRB21NE) and nuclear (ITS) genetic region. Each analysis was performed under both options offered by the software for treatment of gaps, (gaps as 5th state or gaps as missing data). The outputs were compared to see which gap treatment gave the most sensible explanation of the genetic relationships between the pattern types. The most sensible of each of these analyses is presented in the results section.

2.6 Morphological study

2.6.1 Introduction

A morphological study was conducted on 30 accessions. Initially 20 accessions were grown from seed, and about 7 months later 10 accessions corresponding to known cultivars, were added. All plants were scored for qualitative morphological characters at 8.5 months post germination (initial accessions) or 6 weeks for cultivars. For each population the aim was to assess at least 10 plants, (Van de Wouw, 1999).

The purpose of the morphological study was to assess if distinct phenotypes composed of qualitative characters largely independent of the environment, could be found and whether these could be linked to geography, for instance if large leaflet size was observed in plants grown from Lodi and Regal. The aim of this was to compare morphology with DNA sequence and thus to detect any connection between specific phenotypes and specific DNA sequence variation.

All plants were considered of sufficient maturity at time of analysis for a reliable impression of physical appearance to be scored (Caranahan, 1955). Younger age of cultivars was considered insignificant as the traits assessed were qualitative not quantitative.

2.6.2 Description of leaflet characters

Assessed were the leaflet characteristics,

Leaflet shape: (Ovate, Orbicular, Retuse obovate and Obcordate)

Relative leaflet sizes (Tiny- T, Small- S, Small-medium- S-m, Medium-small -M-s, Medium-M, Medium- Large and Large- L).

Presence or absence of the dominant white V mark, in single or double form

Appearance of the white V mark: faint, bright, or normal.

Width and shape of the white V mark: thick, thin, squat or incomplete.

Note: Lower case letter indicate decreased presence of character.

Figure 2.1, photographs of prominent leaf morphology



Ovate

Orbicular

Retuse obovate

Obcordate

White V mark

Descriptions of leaflet shapes

Ovate- These leaflets were longer than they were wide, with elongation lengthwise as opposed to expansion sideways. “Egg-shaped, attached by the broad end” (Allan, 1982). The nick in the leaflet apex was absent or reduced, relative to retuse obovate and obcordate shaped leaflets, (see figures 38 and 39, Ps. 995-996, Allan, 1982)

Orbicular – These leaflets were more round than elongated in appearance, being approximately circular. “Length = width” (Allan, 1982). Lack of extreme slope in side/edges of the leaflet and a reduction of the nick in the apex separated these leaflets from those of the retuse obovate and obcordate shaped leaflets, (see figure 38, P. 995, Allan, 1982).

Retuse obovate- These leaflets showed less extremity in shape, being more rounded. Retuse: “rounded apex with a small notch” (Allan, 1982) The term obovate in relation to this class refers to the inversion of a egg-like shape (ovate) where the attached end of the leaflet comes to a point rather than a rounded end, as is observed in ovate shaped leaflets, (see figures 38 and 39, Ps. 995-996, Allan, 1982).

Obcordate- The leaflets of this shape were a more extreme version of the retuse obovate shape. They had a prominent heart like appearance, (see figure 39, P. 996, Allan, 1982).

2.6.3 Descriptions of leaflet size classes

A subjective measurement of leaflet size was made through comparison of leaflet sizes to plants grown from recognized ecotype/selection standards. Italian Lodi was used as a representation of large sized leaflets (Caradus et al., 1990) and ecotype Kent and selection S.184 for tiny/small (Davies et al., 1967; Caradus et al., 1990) sized leaflets.

2.6.4 Descriptions of the classes of white V mark

Thin - A thin white V mark that was variable in width but clearly distinguishable from that of a double V mark composed of two thin V marks.

Thick- V mark was thick.

Squat- A flat top or levelling off to the point of the V was observed in absence a peak of the V. The appearance was observed of a slight m like shape.

Incomplete - Absence of a complete V or M like shaped white marking. Most of these plants were scored as having white “bits” on their leaflets. In general it was just the side parts of the V that were present, the point of the V missing.

Faint - This ranged from just visible, to a defined but very light V marking.

Bright - White V markings were clearly discernable, and had no yellow colour.

Normal- White V marking appeared indistinguishable from the average appearance observed when looking across all plants assessed.

Double white V mark- Two V marks, one inside the other.

2.6.5 Recording and tabulation

The accession number, plant number and morphology scoring were recorded by myself at the same time over the course of two afternoons. To avoid any effects due to difference in light levels in the glasshouse the analyses were carried out at the same time of day on both occasions. Data was tabulated in Microsoft Excel.

2.7 Comparison of DNA and Morphology

A comparison was made to assess whether there were any correlations between DNA patterns and morphology, or morphology and geographic origin.

Chi squared assessments were used to test for statistical significance of any associations.

2.8 The novel Algerian accession C20859

During the course of the experiment it was found that the morphological characteristics of the Algerian accession C20859 were unlike those of *T. repens* and that it appeared to have some characteristics similar to *T. occidentale* and *T. ambiguum*.

This prompted the hypothesis of a novel *Trifolium* species.

To verify the nature of this accession PCR and direct sequencing was performed as for *T. repens* individuals. A chromosome count on ~250 metaphase cells was conducted by Dr Wajid Hussain, (2009). Crossing experiments, both self and between *Trifolium* species closely related to *T. repens* were performed by Prof Warren Williams, and Isabelle Williams (2009).

Chapter 3 Results

3.1 Diversity Results from Sequencing of Chloroplast Genetic Markers, in White Clover, *T. repens*.

3.1.1 Introduction

Two sets of primers, the universal CD primers (Taberlet, 1991) and IRB21NE (Dr Nick Ellison, 2009) amplify two independent regions from the maternally inherited chloroplast genome. These are respectively the non-coding regions of the; *trn L* intron and a shortened sequence from the variable internal repeat region, (IRB). Previous work by Ellison et al., (2006 and unpublished data) has shown these primers were sufficient to detect diversity within the *Trifolium* genus. Subsequent work by Dr Nick Ellison, (Pers comm., 2009) had suggested, and confirmation has been provided by this study, that there is adequate diversity among and within populations of white clover from different geographic locations for scientific study of genetic patterns.

Pattern types

There are three recognised variable positions in the sequences, C/T variation at positions 433 and 456 (CC, CT, TC) in the *trn L* intron (non-coding region) amplified by the universal primers, and the presence/absence (+/-) of a penta-nucleotide repeat, (CAAAA) in the IRB21NE sequence. The combination of these gives five pattern types that can be derived from one another by a single mutational step (Fig. 3.1).

Three chloroplast pattern combinations CC+, CT+ and TC- have been previously found in white clover populations (Nick. Ellison, pers. com, 2009). This study found evidence of a fourth pattern type CC-, confirmed for two individuals. No evidence was found of the most likely variations on previously observed pattern combinations, TC+ and CT-. A further maternal variant CC+* with a novel insert in the *trn L* intron was observed infrequently. Throughout this report the chloroplast pattern combinations are coloured as follows; CT+ = red, CC+ = blue, CC- = purple, CC+* = pink and TC- = dark green, (Fig. 3.1).

Pattern type frequencies

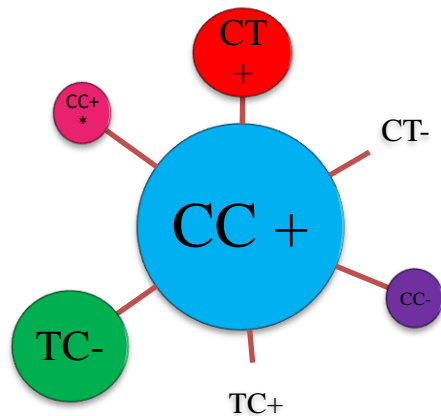


Figure 3.1 The four observed pattern types and their approximate relative occurrences depicted by balloon sizes. CT+ = red, CC+ = blue, CC- = purple, CC+* = pink and TC- = dark green. Other possible, but not observed, pattern types TC+- and CT- are not coloured.

From the sampling of 300 individuals, it was observed, as shown in Table 3.1, that pattern type CC+ was the most frequent pattern, predominating throughout Europe, and found in Israel and Africa, but with limited distribution in Portugal and Spain.

By contrast, TC-, the second most frequent pattern was very infrequent outside of Portugal and Spain. CT+ was the least common of the three main pattern types and had a slightly narrower distribution than CC+. Very limited occurrence was observed of pattern type, CC-.

Table 3.1 Relative occurrences and places of prominence for the four observed chloroplast pattern types.

C/D	IRB21NE	Classification	Relative Occurrence	Place of Prominence
CT	+	CT+	11.7% = (53/300)	France, (Aran/Crusader)
CC	+	CC+	64.3% = (193/300)	All Europe, Africa, Israel Portugal, Spain, Italy,
CC	-	CC-	0.7% = (2/300)	Algeria
TC	-	TC-	23.3% = (70/300)	Spain, Portugal

Note: 300 individuals were assessed for the presence of DNA variation at the three variable sites within, IRB21NE and *trn L*. Chloroplast pattern types are coloured according to type.

The percentage occurrence of each pattern type across the complete data set is given in column 4 and the areas of prominence are shown in column 5.

3.1.2 Portuguese accessions of white clover, showed the prominence of type, TC-. The pattern types CC+ and CT+ were also found.

Table 3.2 Percentage occurrence of chloroplast pattern types in Portuguese accessions

Accessions & Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C9516, Tras Os Montes, Vila Verde, (NW)	CT+	5.9	4
C9024, Tras Os Montes Esphino, (NE)	CC+	1.5	1
C9514, Trans Os Montes, (NE)	CC+	10.3	7
AZ4839, Baragonça, Tras Montes, (NE)	CC+	2.9	2
C7576, Estremadura, (CN)	CC+	8.8	6
C8985, Coimbra, (C)	CC+	16.2	11
C11519, Vila Pouca de Agu, (N)	CC+	4.4	3
C20477, Águas de Moura, Pamela, (SW)	CC+	4.4	3
C8979, Ribatejo, Grândola, (SW)	CC+	2.9	2
C8979, Ribatejo, Grândola, (SW)	CC-	1.5	1
AZ4839, Baragonça, Tras Montes, (NE)	TC-	2.9	2
AZ4838, Minho, Braga, Ferreiras, (NW)	TC-	1.5	1
C7651, Bobadelo, Tras Os Montes, (NE)	TC-	1.5	1
C9514, Trans Os Montes, (NE)	TC-	1.5	1
C9024, Tras Os Montes Esphino, (NE)	TC-	7.4	5
C9516, Tras Os Montes, Vila Verde, (NW)	TC-	1.5	1
C7582, Beira littoral, (NW)	TC-	8.8	6
C8976, Algarve, (S)	TC-	5.9	4
C20477, Águas de Moura, Pamela, (SW)	TC-	2.9	2
C8979, Ribatejo, Grândola, (SW)	TC-	7.4	5
Totals	4	100	68

Note: Chloroplast pattern types are coloured according to type and as in accompanying pie charts. A total of 68 individual plants were assessed in this category. The percentage occurrence of each pattern type per accession and the number of individuals is shown in columns 3 and 4 respectively. (CN)- Central North, (NE)-North-East, (NW)- North-West, (S)- South, (SW)-South-West.

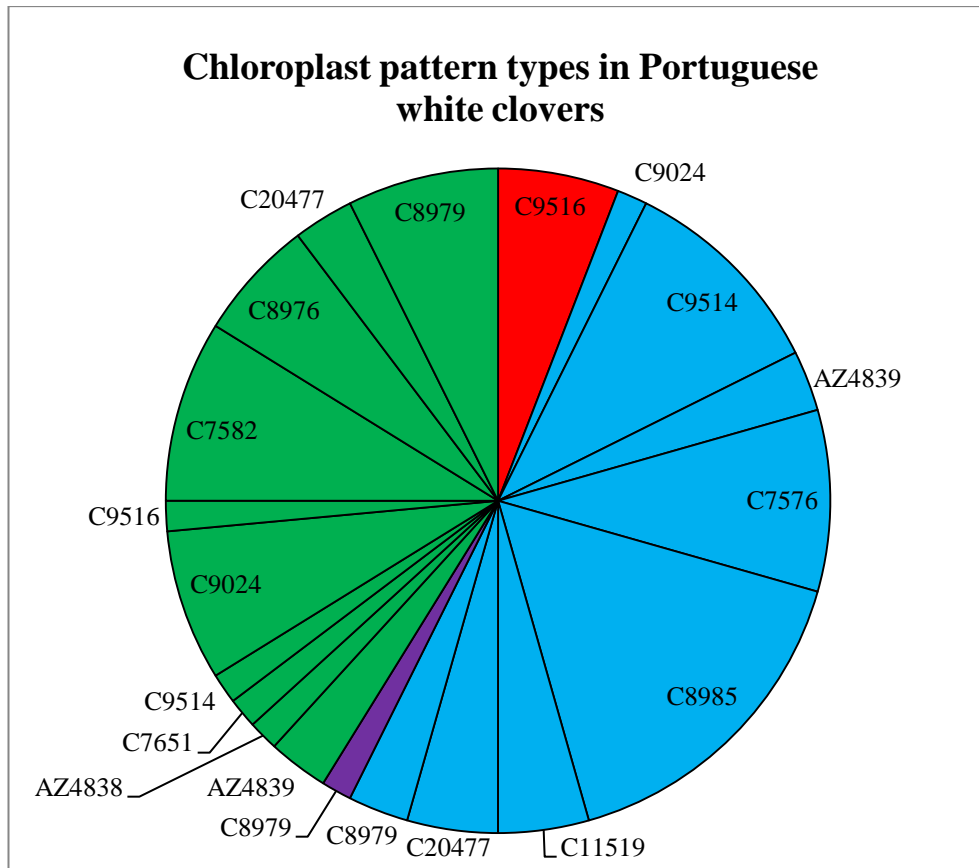


Figure 3.2 Distribution of chloroplast pattern types per accession for plants sampled from Portuguese accessions. Chloroplast pattern types, CT+ = red, CC+ = blue, CC- = purple and TC- = dark green. A total of 68 individuals were sampled. North = C11519, Central North = C7576. North-East = C9514, AZ4839, C7651, C9024. North-West = C9516, Az4838, C7582. South = C8976. South-West = C20477, C8979. Central = C89585.

The main feature of the Portuguese dataset was the prominence of the pattern type TC-. It was observed at 41.2% of Portuguese individuals sampled. Three separate pattern types made up the remainder of the observed variation. The relative proportion of pattern type CC+ was similar to that found throughout the remainder of the dataset, while CT+ was very rare, occurring only in accession C9516 from Tras Os Montes, Vila Verde north west Portugal. Pattern type CC- was observed in one individual from accession C8979, Ribatejo, Grândola, Setúbal south-west Portugal.

3.1.3 Spanish accessions showed a strong prominence of pattern type TC-. Pattern types CC+, CT+ and CC- were observed at a much lesser extent.

Table 3.3 Occurrence of chloroplast pattern types in Spanish accessions.

Accessions & Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C9027, La Estrada, Galicia, (NW)	CT+	1.6	1
C9034, Galicia, (NW)	CT+	3.2	2
C9027, La Estrada, Galicia, (NW)	CC+	3.2	2
C7600, Spanish Pyrenees	CC+	9.7	6
C9025, Extremadura, (W)	CC+	4.8	3
C9531, Ponferrada, León & Castile (NW)	CC+	4.8	3
C8975, Andalucía, (S)	CC+	6.5	4
C9025, Extremadura, (W)	CC-	1.6	1
C9025, Extremadura, (W)	TC-	1.6	1
C9531, Ponferrada, León & Castile (NW)	TC-	9.7	6
C7600, Spanish Pyrenees	TC-	1.6	1
C9027, La Estrada, Galicia, (NW)	TC-	6.5	4
C9037, Galicia, (NW)	TC-	12.9	8
C9034, Galicia, (NW)	TC-	12.9	8
C8994, Asturias, (N)	TC-	19.4	12
Totals	5	100	62

Note: Chloroplast pattern types are coloured as in accompanying pie charts. 62 individual plants were sampled from Spain. Note: Chloroplast pattern types are coloured according to type, and as in the accompanying pie chart. The percentage occurrence of each pattern type for each accession or selection and the number of individuals is shown in columns 3 and 4 respectively. (N) - North, (NW) - North-West, (W) - West, (S) - South.

The interesting feature of the Spanish dataset was the prominence of the chloroplast pattern type TC-. This occurrence was higher relative to the situation observed in Portugal, where TC- was also prominent. Type CT+ was observed in only two individuals from different parts of Galicia. In the Spanish data the pattern type CC- was observed in one individual from accession C9025, from Extremadura, western Spain.

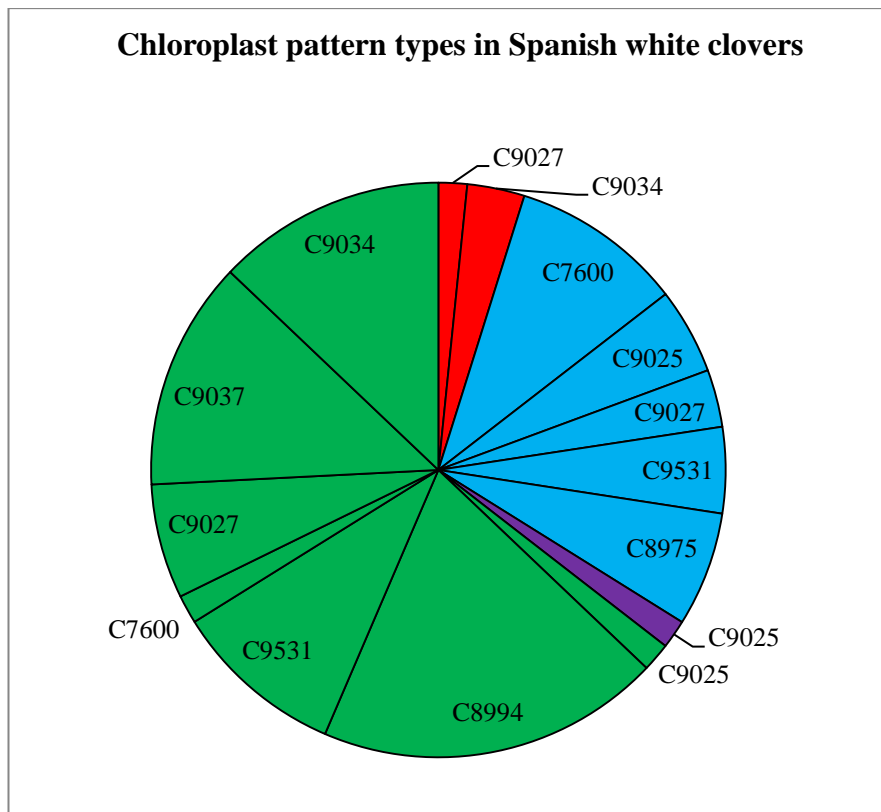


Figure 3.3 Distribution of chloroplast pattern types per accession for individual Spanish plants. Chloroplast pattern types, CT+ = red, CC+ = blue, CC- = purple, TC- = dark green. A total of 30 individuals were assessed. North = C8994, North-West = C9034, C9027, 9531. West = 9025. South = 8975.

3.1.4 Comparisons with *T. pallescens* chloroplast sequences

The majority of the data set showed greater similarity in *trn L* intron sequence to Grasslands Huia. The three *T. pallescens* individuals included in this study showed two forms of *trn L* sequence these forms were from different locations in and around the Pyrenees Mountains and Austria (Table 3.4).

Four individual *T. repens* plants classed as CC* were observed that had a unique *trn L* chloroplast sequence. This sequence differs from both the Huia reference sequence and the two *T. pallescens* forms. These four individuals are more similar to *T. pallescens* forms from the Pyrenees Mountains, to which they are different by only an A/T inversion, at positions 288, 289 in the Huia reference sequence (Table 3.4).

These four individuals came from three accessions, Portuguese C7576 from Estremadura- Central Southern Portugal, and Spanish accessions C8975, Andalucía-

Southern Spain and C9025, Extremadura -Central Southern Spain. Other *T. repens* individuals sampled from the same accessions and neighbouring locations did not show this unique sequence. Sequencing of IRB21NE for two of these individuals revealed presence of the CAAAA, insert; these individuals were classed as CC+*. IRB21NE Sequence could not be obtained for the other two individuals.

Table 3.4 Similarity of unique Portuguese and Spanish individuals to *T. pallescens* individuals from the Pyrenees and Austria

Trifolium Individuals	Reference positions												
	145	146	147	148	288	289	290	293	-	-	-	-	294
NZ Huia Reference	A	A	A	A	T	A	T	T	-	-	-	-	G
<i>T. pallescens</i> AZ4856	-	-	-	-	T	-	T	T	-	-	-	-	G
<i>T. pallescens</i> AZ4837	A	A	A	A	T	A	A	T	A	T	T	A	T
<i>T. pallescens</i> Jaca	A	A	A	A	T	A	A	T	A	T	T	A	T
C7576 No.11B	A	A	A	A	A	T	A	T	A	T	T	A	T
C7576 No. 34	A	A	A	A	A	T	A	T	A	T	T	A	T
C8975 No. 52	A	A	A	A	A	T	A	T	A	T	T	A	T
C9025 No. 73	A	A	A	A	A	T	A	T	A	T	T	A	T

Note: Huia, (Grasslands Palmerston North white clover cultivar), *T. pallescens*; AZ4856 - Tirol, Sölden/Obergurgl, Rotmoos Valley, Austria. Pyrenees forms: AZ4837, from Nuria, north of Barcelona, Spain; Jaca, independent herbarium sample from Instituto Pirenaico de Ecología, C.S.I.C., Jaca, Spain. Sequences for individuals Jaca, AZ4856, and AZ4837 were obtained from Dr Nick Ellison, 2009-2010. Accessions: Portugal individuals numbered 11B & 34 –Extremadura, Spanish accessions C8975- number 52, and C9025 number 73- Extremadura. The one A/T inversion difference between novel individuals and *T. pallescens* Pyrenees and *T. repens* Huia reference sequence is shown in blue (T) and red (A). (+/-) indicates presence/absence of differences.

3.1.5 White clover from French origins showed near equal prominence of two maternal chloroplast pattern types, CC+ and CT+.

The chloroplast data from French derived plants showed an unequal distribution between the two most common chloroplast pattern types, CT+ and CC+. Both were present in all selections, mixed cultivars and accessions, although in accessions C2418 and C9046, only CC+ was observed.

Small sampling sizes influenced the observed pattern distributions in France, as 24 individuals were sampled from southern France and only three individuals were sampled from the accession C2418, (north-west). However to gain a better knowledge of the distribution of chloroplast types around France more sampling from northern and central France is required. Hence this hypothesis of a CC+ consensus pattern type in northern France is tentative.

A high prominence was observed of the CT+ pattern type in southern France, due to the high prominence of the type in geographically representative selection Crau, in 7 out of 9 plants, and in 3 out of 7 plants of Crusader (also containing Crau parentage).

Table 3.5 Occurrence per accession/cultivar of chloroplast pattern types in French derived white clovers.

Accession/Selection Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C18951, French Crau, (S)	CT+	25.6	11
C7690, Entressen, Bouches du Rhône (SE)	CT+	7.0	3
Aran, France & Israel cross	CT+	7.0	3
Crusader, French Crau & Syrian parentage	CT+	7.0	3
C7690, Entressen, Bouches du Rhône (SE)	CC+	2.3	1
C9046, Mouriès, Bouches du Rhône, (S)	CC+	14.0	6
C2418, Le Conquet, (NW)	CC+	7.0	3
Aran, France & Israel cross	CC+	16.3	7
Crusader, French Crau & Syrian parentage	CC+	14.0	6
Totals	3	100	43

Note: Chloroplast pattern types are coloured as according to type and as in the accompanying pie chart. The percentage occurrence of each pattern type for each accession or selection and the number of individuals is shown in columns 3 and 4 respectively. (S) = South, (SE) = South East, (NW) = North-West.

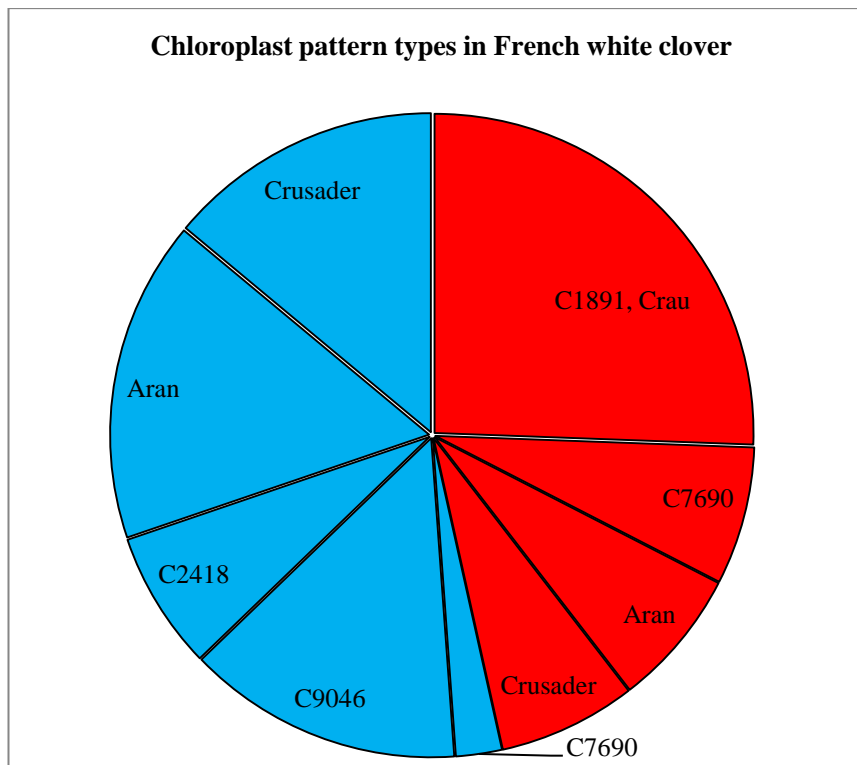


Figure 3.5 Distribution of chloroplast pattern types per accessions/selection and mixed cultivars for plants sampled with French origins. Chloroplast pattern types, CT+ = red, CC+ = blue. Absent; CC- = purple and TC- = dark green. Assessed were 43 individuals. South = C9046, 18951. North-West = C2418. South-East = C7690.

3.1.6 White clover individuals from British cultivars have chloroplast pattern type CC+. Pattern type CT+ is less common and is restricted to landrace Kent Wild White and mixed cultivar S.184.

The majority of plants showed presence of the chloroplast pattern type CC+. This pattern type was shared by both of the Kent landraces and mixed cultivar Blanca. It was absent from the cultivar S.184. It is interesting to note the difference in pattern composition between individuals of the Kent populations. The two Kent populations show a difference in relative proportions of pattern types CT+, CC+ and TC-.

Table 3.6 Distribution of chloroplast pattern types in British white clovers

Landraces, Selections & Mixed cultivar Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C6304, Kent Wild White	CT+	13.0	3
C6374, Kent White	CT+	4.4	1
C21105, S. 184	CT+	4.4	1
C6360, Blanca	CC+	30.4	7
C6304, Kent Wild White	CC+	17.4	4
C6374, Kent White	CC+	26.1	6
C6374, Kent White	TC-	4.4	1
Totals	3	100	23

Note: Sampled were individuals from Kent landraces, Kent White C6374, Kent Wild White C6304 and cultivars S.184 derived from wild British White Clover populations and mixed cultivar Blanca bred from Kersey (English) x Belgian populations. Chloroplast pattern types are coloured as in accompanying pie charts. A total 23 individuals were assessed.

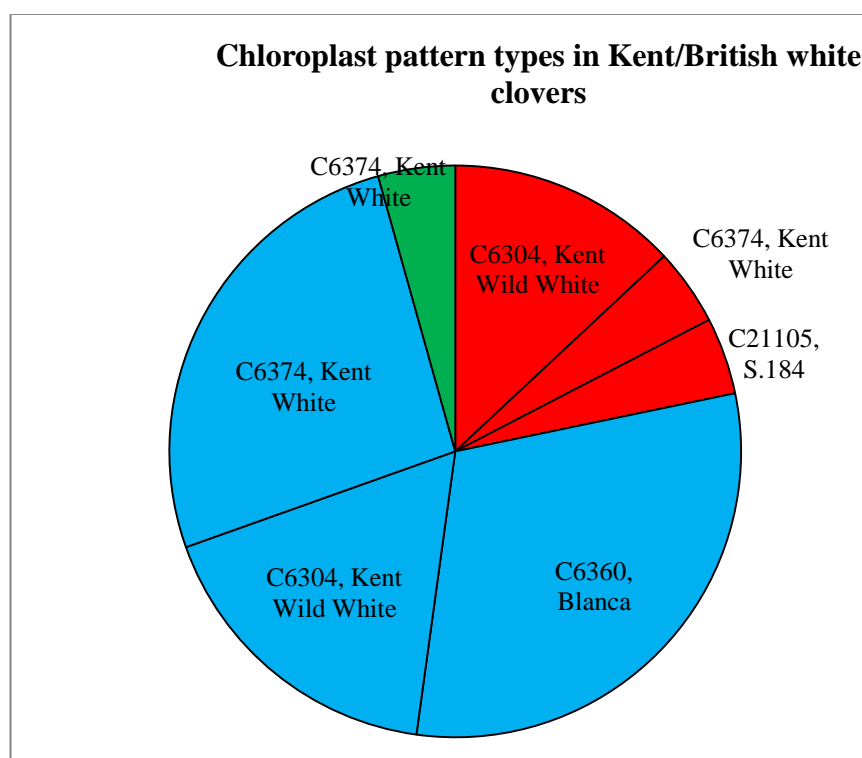


Figure 3.6 Distribution of chloroplast pattern types per landraces/mixed Cultivar for plants sampled with British origins. Origins: Kent landrace (White C6374, Wild White C6304), Kent Suffolk England, Blanca C6360 (Belgium, Kersey mix) and S.184, C21105, (Wild British populations). Chloroplast pattern types, CT+ = red, CC+ = blue, TC- = dark green. Absent are; CC- = purple. Assessed were 23 individuals.

3.1.7 White Clover from the Italian ecotype Lodi, and USA ecotype cultivar Regal show the prominence of the same maternally inherited pattern type CC+.

The results showed that despite growth in two separate geographic locations the ecotype selection Regal selected from populations of imported Italian Lodi locally grown in south eastern states of the USA showed the same pattern type as the parental source population. An individual of Italian Lodi showed the presence of pattern type TC-; this type was not observed in ecotype selection Regal.

Table 3.7 Occurrence of chloroplast pattern types in plants from ecotype Italian Lodi and USA ecotype selection Regal

Landrace/Ecotype selection	Chloroplast pattern type	% Occurrence	No. Individuals
C5897, Giganteum Lodigiano, Italy	CC+	54.5	12
C15111, Regal, Alabama, USA	CC+	40.1	9
C5897, Giganteum Lodigiano, Italy	TC-	4.6	1
Totals	3	100	22

Note: Lodi = Ladino Lodigiano Gigante, (C5897) from the Northern Italian Po Valley. Chloroplast pattern types are coloured as in accompanying pie charts. 22 individual plants were assessed.

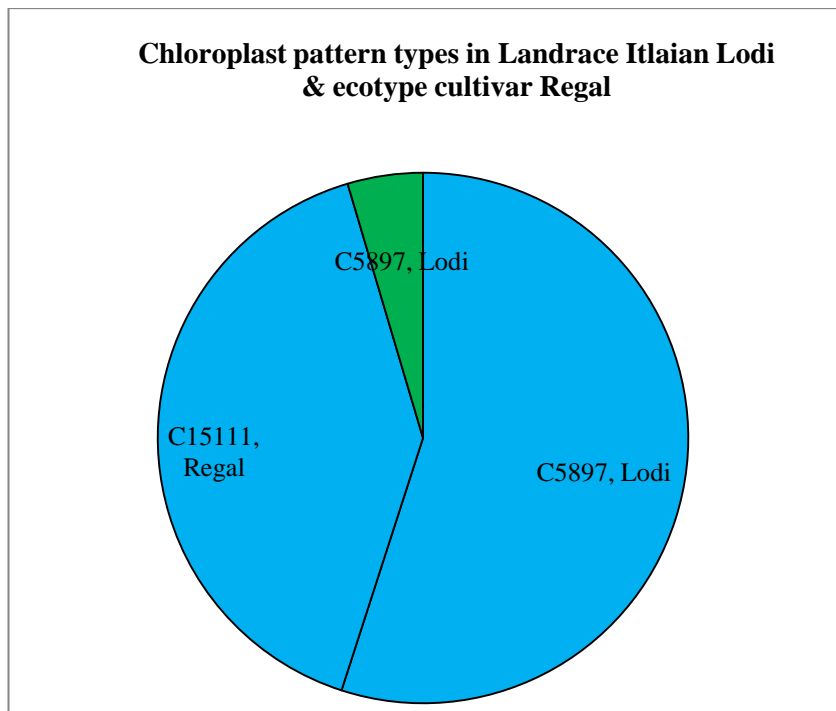


Figure 3.8 Distribution of chloroplast pattern types per landrace/cultivar for plants sampled from Landrace Italian Lodi C5897 and USA Ecotype cultivar Regal C15111. Chloroplast pattern types CC+ = blue, TC- = dark green. Absent; CT+ = red, CC- = purple. A total of 22 individuals were assessed.

3.1.8 White Clover individuals from northern European populations show near complete conservation of pattern type CC+.

The results from the northern European populations show an overwhelming presence of the chloroplast pattern type CC+, 87.5%. Only limited presence was observed for pattern type CT+, which was found in selection Barbian from the Netherlands.

Table 3.8 Occurrence of chloroplast pattern types in northern and north-western European white clover selections.

Selections, Cultivars & Landrace Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C24247, Barbian, The Netherlands	CT+	6.7	2
C24247, Barbian, The Netherlands	CC+	23.3	7
C960, Tammisto, Finland	CC+	30.0	9
C2095, Morso, Denmark	CC+	40.0	12
Totals		100	30

Note: Chloroplast Pattern types CT+ to CC-, per individual for northern European cultivars north-west –Barbian, North- Tammisto, Morso, Chloroplast pattern types are coloured as in the accompanying pie chart. Assessed were a total of 30 individuals. The percentage occurrence of each pattern type for each selection or landrace and the number of individuals is shown in columns 3 and 4 respectively.

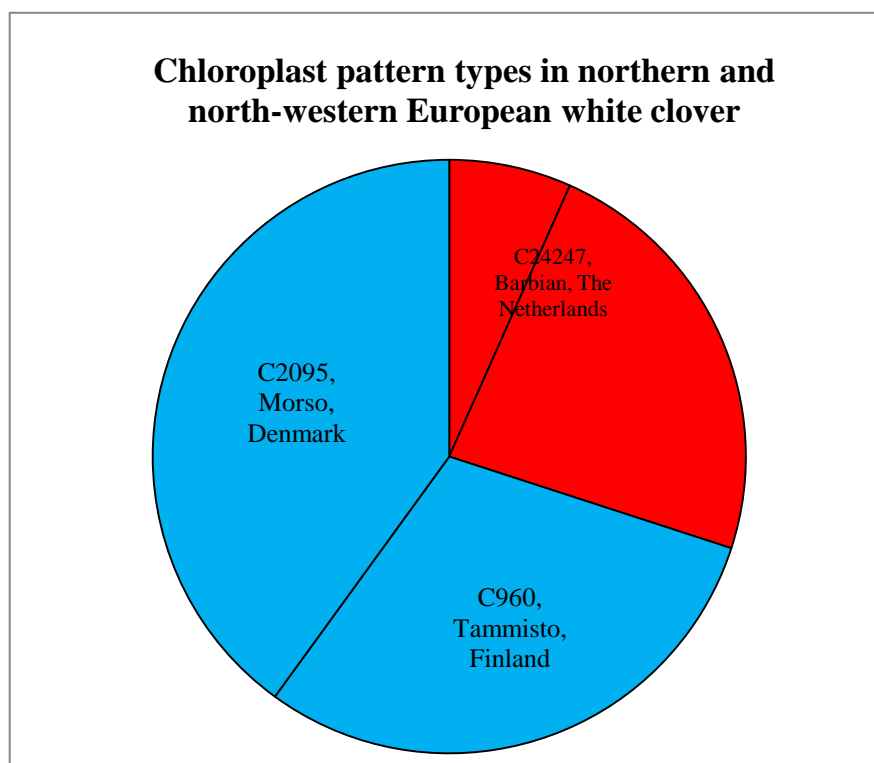


Figure 3.8 Distribution of chloroplast pattern types listed per cultivar for individual plants from northern European selections. Chloroplast pattern types, CT+ = red, CC+ = blue. Absent are CC- = purple and TC- = dark green. A total of 30 individuals were sampled. North = Finland-Tammisto. North-West = Denmark- Morso and the Netherlands- Barbian.

3.1.9 Pattern type CC+ is prominent in central Europe and western Europe mixed cultivar Blanca, CT+ was observed at limited presence.

As was observed for the northern European individuals, CC+ is the predominating pattern type. CT+ was observed only once, in the Czech Republic, cultivar Ovcak.

Table 3.9 Occurrence of chloroplast pattern types in central European selection Ovcak and in mixed cultivar Blanca.

Selection & Mixed cultivar Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C7555, Ovcak, Czech Republic, (C)	CT+	8.3	1
C7555, Ovcak, Czech Republic, (C)	CC+	33.3	4
C6360, Blanca, Belgium Kersey mix (W)	CC+	58.3	7
Totals		100	12

Note: Percentage occurrence of chloroplast pattern types listed for plants from Central European Czech Republic selection and Mixed origin cultivar Blanca. Chloroplast pattern types, CT+ = red, CC+ = blue. Absent are CC- = purple and TC- = dark green. A total of 12 individuals were sampled. (C)- Central Europe, (W)- Western Europe.

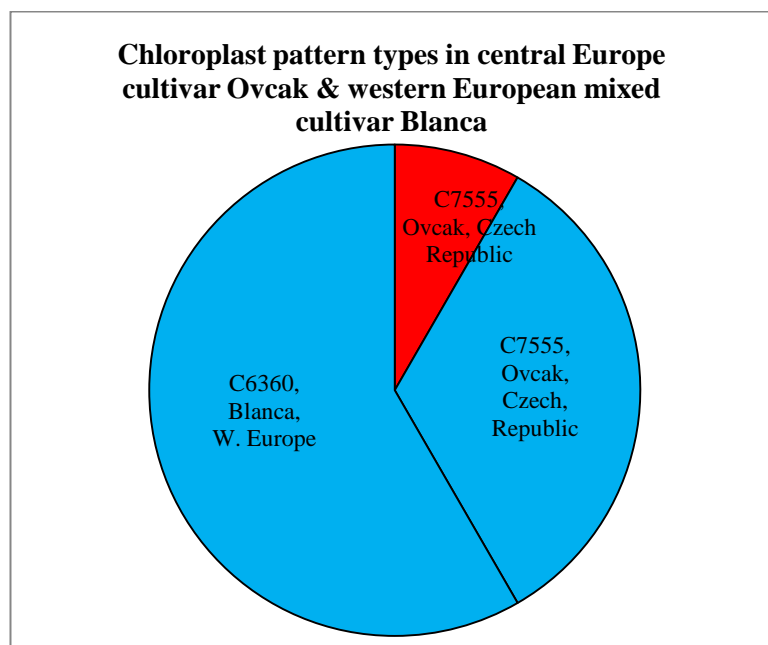


Figure 3.9 Distribution of chloroplast pattern types listed for plants from Central European Czech Republic selection and Western European mixed origin cultivar Blanca. Chloroplast pattern types, CT+ = red, CC+ = blue. Absent are CC- = purple and TC- = dark green. A total of 12 individuals were sampled.

3.1.10 White clover plants from Israeli cultivars show only pattern type CC+. Mixed origin cultivar Aran shows the presence of pattern types CT+ and CC+, in line with its French ancestry.

The results from plants sampled from cultivars with Israeli origins show an interesting split in distribution of pattern types. Those derived from just Israeli origin showed 100% occurrence of the commonest chloroplast pattern type, type CC+. Those individuals from the mixed cultivar Aran with breeding history of French and Israeli parentage showed the presence of the moderately common pattern type, CT+, in three out of nine individuals.

Table 3.10 Occurrence of chloroplast pattern types in cultivars with Israeli ancestry

Mixed cultivars & Selection Origins	Chloroplast pattern type	% Occurrence	No. Individuals
Aran, French and Israeli cross	CT+	11.1	3
C5892, Haifa, Huleh district N.E Israel	CC+	16	4
C4143, Tamar, Israel	CC+	44	11
Aran, French and Israeli cross	CC+	28	7
Totals	2	100	25

Note: Chloroplast Pattern types, per individual for plants from Israeli cultivars, Haifa- Huleh district North- East Israel, and Tamar- Israel. Chloroplast pattern types are coloured as per and as in the accompanying pie chart. Assessed were total of 25 individuals. The percentage occurrence of each pattern type for each mixed cultivar or selection and the number of individuals is shown in columns 3 and 4 respectively.

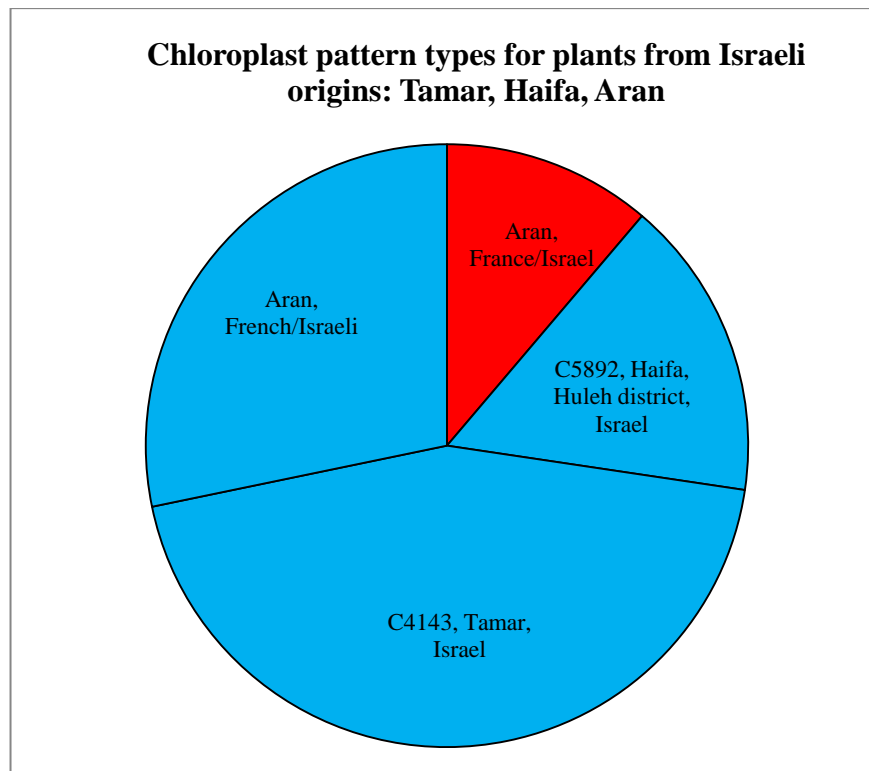


Figure 3.10 Distribution of chloroplast pattern types for cultivars derived from Israeli origins. Included are: Pure Israeli cultivars Haifa and Tamar, mixed origin cultivar Aran (France/Israel). Chloroplast pattern types, CT+ = red, CC+ = blue. Absent are; CC- = purple and TC- = dark green. A total of 25 individuals from were assessed.

3.11 North African white clover accessions show conservation of pattern type CC+.

The results for individuals of North African origin, (Algeria and Morocco) showed complete conservation of the common pattern type, CC+.

Table 3.11 Occurrences of chloroplast pattern types in White Clover from North African Accessions.

Accessions & Cultivar Origins	Chloroplast pattern type	% Occurrence	No. Individuals
C1761, Morocco	CC+	24.1	7
C6452, Algeria	CC+	38.0	11
C1751, Algeria	CC+	34.5	10
C20499, Algeria	CC+	3.4	1
Totals	1	100	29

Chloroplast pattern types are coloured as per type. Assessed were 29 individuals. The percentage occurrence of each pattern type for each accession and the number of individuals is shown in columns 3 and 4 respectively. Absent were pattern types CT+, TC- and CC-.

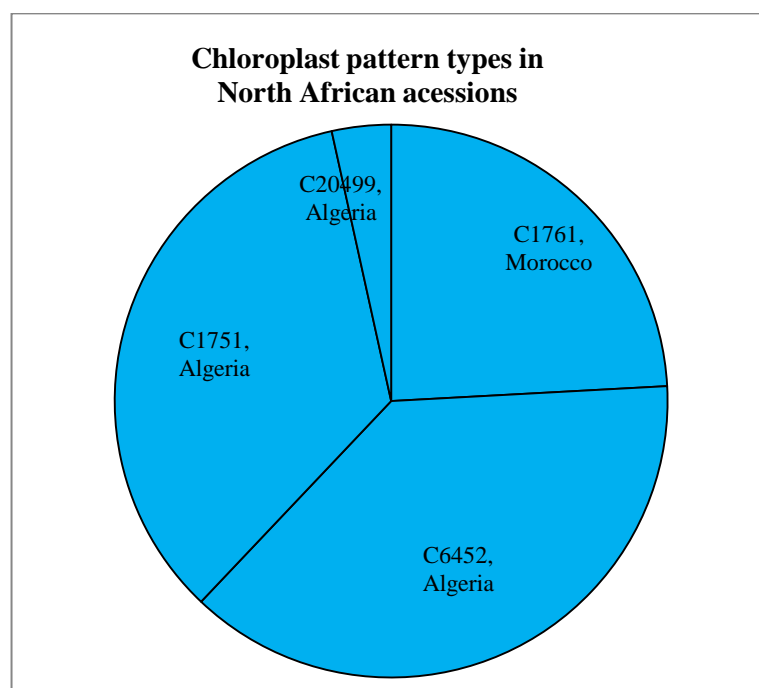


Figure 3.11 Distribution of chloroplast pattern types listed per accession/cultivar for plants from African origins. Chloroplast pattern types, CC+ = blue. Absent are; CT+ = red, CC- = purple and TC- = dark green. A total of 29 individuals were sampled.

3.12 South African out-group population, (Zone 3), Dusi showed only pattern type CC+.

Three individuals were sampled from the South African population of Dusi, treated as the out-group of the chloroplast data set. These three individuals showed the same conservation of the chloroplast type CC+ as was observed in the North African individuals.

3.2 ITS Nuclear pattern types

Table 3.12 Main ITS nuclear pattern types

Nuclear pattern type	50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C
5	T	C	G	G	C	C	C	T	C
5'	T	C	G	G	C	C	C'	T	C
30	T	Y	G	G	C	C	C	T	C
29	T	C	G	G	C	C	C	T	Y
8	T	C	G	G	T	C	C	T	C
32	T	C	G	G	Y	C	C	T	C
44	W	C	G	G	Y	C	C	T	C
34	W	C	G	G	C	C	C	T	C
7	T	T	G	G	C	C	C	T	C
10	T	C	G	T	C	C	C	T	C
12	A	C	G	G	C	C	C	T	C
9	A	C	G	G	T	C	C	T	C
31	A	C	G	G	Y	C	C	T	C
32A	N	C	G	G	Y	C	C	T	C
15	T	C	G	G	C	T	C	T	C
15'	T	C	G	G	C	T	C	T	C
15A	W	C	G	G	C	T	C	T	C
35	T	C	G	G	Y	T	C	T	C
24	T	C	G	G	C	T	C	T	G
23	T	C	G	G	C	T	A	T	G
17	T	C	G	G	C	T	A	T	C
17'	T	C	G	G	C	T	A'	T	C
18	T	C	G	G	C	T	M	T	C
26	T	C	G	G	C	T	C	C	T
36	T	C	G	G	C	T	C	Y	C
27	A	C	G	G	T	T	C	T	C
48	G	G	G	G	C	T	G	Y	C
20	T	C	G	G	C	Y	C	T	C
20'	T	C	G	G	C	Y'	C	T	C
39	T	C	G	G	Y	Y	C	T	C
40	T	C	G	K	C	Y	C	T	C
11	T	C	G	G	C	Y	C	T	T
19	T	C	G	G	Y	Y	M	T	C
38	T	T	G	G	C	Y	C	T	C
21	A	C	G	G	Y	Y	C	T	C
41	W	C	G	K	C	Y	C	T	C

Table 3.13 Rarer ITS nuclear pattern types

Nuclear pattern type	50 T/a	62 C/T	111 G/a	216G	486 C/T	509 T/C	571 C/a/m	613 T	654C	% Occurrence	No. Individuals
32	T	C	G	G	Y	C	C	T	C	1.5	5
39	T	C	G	G	Y	Y	C	T	C	1.8	6
30	T	Y	G	G	C	C	C	T	C	1.8	6
7	T	T	G	G	C	C	C	T	C	1.8	6
10	T	C	G	T	C	C	C	T	C	1.2	4
9	A	C	G	G	T	C	C	T	C	1.5	5
<hr/>											
2	T	C	A	G	C	C	C	T	C	1.2	4
3	T	C	A	G	C	C	C	C	T	0.3	1
25	T	C	A	G	C	Y	C	T	C	0.3	1
28	W	C	A	G	C	Y	C	T	C	0.9	3
1	A	C	A	G	C	C	C	T	C	0.6	2
45	T	C	R	C	C	T	C	T	C	0.3	1
43	T	C	R	G	C	Y	C	T	C	0.3	1
<hr/>										Totals	
13										13.5%	45

Introduction

In this study 45 nuclear pattern types were observed in 327 individuals. Three main types 5, 15, 17 and a heterozygous pattern type 20 explained the majority of diversity, across the nuclear data set. Type 15, was the most common and wide spread across all sampled populations. The heterozygote type 20 was more common than main type 5 or type 17 (Table 3.12, 3.14). Variants of these main types 5', and 15' were found at lower occurrence than the main form. Variant types 17' and 20' were more frequent than types 17 and 20, (Table 3.14). A second but less frequent case of heterozygosity was observed from the mixed inheritance of pattern types 15 and 17 that generated pattern 18 (Table 3.12, 3.14).

Table 3.14 Main ITS nuclear pattern types and their positions of variability

Nuclear pattern type	509, T/C	571, C/a/m	No. individuals	Occurrences
5	C	C	33	10.10%
15	T	C	105	32.10%
20	Y	C	53	16.20%
17	T	A	5	1.5%
5'	C	C'	1	0.31%
15'	T'	C	6	1.8%
20'	Y'	C	11	3.60%
17'	T	A'	13	3.9%
18	T	M	1	0.31%
Totals			228	69.8%

Note: Shaded boxes indicate positions of variability between the four main pattern types. All other positions of variability are conserved between the main pattern types, (See Table 3.12), for listed positions of variation.

Main pattern types 5, 15, 17, and heterozygotes 20 and 18.

The main types 5 and 15 that predominate throughout the data set, differed by the presence of a C or T at position 509 in the ITS nuclear reference sequence. Individuals, for which at this position in the DNA sequence a C was observed in one ITS region and a T in the ITS region of the other homologue, were classified as pattern type 20. This variation is denoted in tables and appendices by the recognised phylogenetic Y symbol for inability to distinguish a consensus form between bases C and T. Where ever type 20 occurred, type 15 is generally present. However the presence of pattern type 5 was usually limited, reflecting a lower allelic frequency within the nuclear data set.

Pattern type 17 differs from types 5 and 15 by the presence of an A at position 571 in the reference sequence (Tables 3.12 & 3.14) and was infrequent within the data set, (Table 3.14). The heterozygote, pattern type 18 was observed once.

Pattern types 20' and 17'

The term, type 20' was used to distinguish individuals where a C was clearly observable at position 509 on one strand while on the opposing DNA strand the base present could not be distinguished between a T or a C, denotation Y'. These individuals were treated as a separate class and coloured light purple to show their similarity to pattern type 20.

Pattern type 17' observed as the presence of a slight C peak in the presence of dominating A peaks in both DNA strands at position 571. This position variability is denoted by an A' in tables, (Tables 3.13 and 3.14).

Distribution of Main nuclear pattern types, 5, 15, 17, 20 and 18

Pattern type 15 shown in medium blue was the most wide spread pattern type, (Fig. 3.12) A reduced distribution was observed for type 5, (orange-red), relative to type 15 (blue), (Fig. 3.12). The heterozygous types 20/20', shown in purple, were only observed in accessions from across Europe and the Iberian Peninsula (France, Italy, landrace Kent White of England, and Portugal/Spain), (Figs. 3.13- 3.22 and Tables 3.14- 3.18). Of these pattern types, 20 was more numerous and of widespread occurrence. A prominence of type 20 was observed in Portugal and France, (Fig. 3.12).

Pattern types 17 and 17' (medium green and grass green) were observed in Northern Europe, (Figs. 3.23 and 3.24). Grouped heterozygote type 18 and type 19, shown in dark green were restricted to Denmark and the Czech Republic, (Fig. 3.12)



Fig. 3.12 A depiction of the relative approximate proportions from sampled populations with country locations for just the main ITS pattern types, as shown by circles.

Each circle represents:

5- orange-red, 5'- dark red,

15- medium blue, 15'- pale blue,

20- purple, 20'-pale purple,

17'-grass green, 17- medium green,

18/19-dark green,

7-pink, 10-yellow, 30-peach, 35-blue,

36-navy and 39-blue

Rarer and different pattern types

Table 3.15 Individuals with rare nuclear ITS pattern types 1, 2, 3, 25 and 28

Nuclear pattern type	50 T/a	62 C/T	111 G/a	216G	486 C/T	509 T/C	613 T	654C	% Occurrence	No. Individuals
32	T	C	G	G	Y	C	T	C	1.5	5
39	T	C	G	G	Y	Y	T	C	1.8	6
30	T	Y	G	G	C	C	T	C	1.8	6
7	T	T	G	G	C	C	T	C	1.8	6
10	T	C	G	T	C	C	T	C	1.2	4
9	A	C	G	G	T	C	T	C	1.5	5
2	T	C	A	G	C	C	T	C	1.2	4
3	T	C	A	G	C	C	C	T	0.3	1
25	T	C	A	G	C	Y	T	C	0.3	1
28	W	C	A	G	C	Y	T	C	0.9	3
1	A	C	A	G	C	C	T	C	0.6	2
45	T	C	R	C	C	T	T	C	0.3	1
43	T	C	R	G	C	Y	T	C	0.3	1
Totals										
13									13.5%	45

The pattern types 1, 2, 3 and 25 were observed rarely within the nuclear data set. These pattern types all have a conservation of an A at position 111 in the ITS reference sequence, which makes this group distinct from the rest of the observed nuclear pattern types, (Table 3.12 and 3.15).

Other pattern types

A selection of pattern types were observed, that appeared as variations within the complete dataset. The low occurrence of these pattern types indicates either origin due to sequencing error or natural low frequency. Collectively these pattern types were found in regions/populations of diversity, and are commented on further in the appropriate geographic sections.

Identification (colour coding)

A colouring scheme was employed to identify which of the main pattern types an individual pattern types was most similar to. A range of red to pink shades was used for pattern types most similar to main type 5. Blues for pattern types most similar to 15. Purples were used for types similar to heterozygotes 20 and 20'.

Shades of green were used to identify the various forms of 17, (17/17'), heterozygote 18 and the related type 19. The pattern types 17'/15' have been grouped in with 17 and 15 respectively. Types 18 and 19 have been treated as one group.

A selection of individuals with an A at position 111 in the ITS reference sequence (See Table 3.13) were identified by the colours pink- 25, peach- 28, lime green-1, yellow- 2 and orange -3. Other rare patterns types are shaded according to the main type to which they were most similar.

Nuclear ITS Results for plants based on geographic distinctions

3.2.1 Portuguese white clover plants show presence of main pattern types 5, 15 and 20 as well as a diverse collection of minor pattern types.

In general the main types 15 and 5 were observed in all accessions but at different frequencies, type 5 was less frequent than type 15. An absence of type 15 was observed for accessions, C11519 (Vila Pouca de Agu, Portugal) and C9024 (Tras Os Montes Espinho), C7648, and C20477 (South-West). Type 5 was prominent in accessions C7582, C9024, AZ4839 C8985, C8979, and C11519.

The heterozygote type 20 was detected in accessions except for; C8976 (Algarve the southernmost region), C9514 (North-East), C20477 (South-West), and C8985 (Coimbra). The variant type 20' was observed twice in accessions, C7582, and C9516. Neither of the variant forms, 5', 15' or 17' was detected.

Seven rare pattern types were observed in Portuguese individuals. Pattern types 1, 2, 3, and 28, were found, these have in common an A at position 111 in the ITS reference sequence, (see, shaded boxes, Table 3.14). These types, 1, 2, 3, and 28 were observed in accession C7648. Pattern type 2 showed a distribution in three accessions, C8979, C8976 and C20477. Types 7, 32 and 17 (predominantly European), were also observed. Novel heterozygous forms 45, 28 (also shaded) were found in C20477, and heterozygote type 32 was found in accession C9514 (north-eastern) Portugal.

Table 3.16 Nuclear pattern types in Portuguese accessions

Nuclear pattern type	Accessions	Locations	% Occurrence	No. Individuals
5	C7648, AZ4839, AZ4838, RT, C8985, C20477, C8976, C7582, C9024, C9514, C8979, C9516, C11519	(NE), On road to Baragonça, Tras Montes, (NE), Minho, near Braga, Ferreiras, (NE), Portugal (?), Coimbra (Central), Águas de Moura, Pamela, (SW), Algarve (S), Beira litoral, (NW), Tras Os Montes, Espinho, (NE), Tras Os Montes, (NE), Ribatejo, Grândola, Setúbal (SW), Villa Verde (NW), Villa Pouca de Agu (CN)	27.0	24
15	AZ4838, C8985, C8976, C7582, C7576, C9514, C8979, C9516	Minho, near Braga, Ferreiras, (NE), Coimbra (Central), Algarve (S), Beira litoral (NW), Estremadura (C), Trans Os Montes (NE), Ribatejo, Grândola, Setúbal (SW), Vila Verde (NW)	10.1	9
20	C9024, AZ4838, C7582, C9024, C8979, C9516, C11519,	Tras Os Montes, Espinho, (NE), Minho, Braga, Ferreiras (NE), Beira litoral (NW), Tras Os Montes Espinho (NE), Ribatejo, Grândola ,Setúbal (SW), Vila Verde, (NW), Vila Pouca de Agu, (CN)	40.4	36
20'	AZ4839, C7582, C9516	On road to Baragonça, Tras Montes (NE), Beira litoral (NW), Vila Verde (NW)	7.9	7
2	C20477	Águas de Moura, Pamela, (SW)	1.1	1
2	C8976	8976, Algarve, (S)	1.1	1
2	C8979	Ribatejo, Grândola	1.1	1
17	C8985	Coimbra, (Central)	1.1	1
3	C20477	Águas de Moura, Pamela, (SW)	1.1	1
28	C7648	(NE)	3.4	3
7	AZ4838	Minho, near Braga, Ferreiras, (NE)	1.1	1
32	C9514	Trans Os Montes, (NE)	2.2	2
1	C7648	(NE)	1.1	1
45	C20477	Águas de Moura, Pamela, (SW)	1.1	1
Total			100	89

Note: 89 individuals from different populations were assessed. (NW) North-West: C7582, C9516, AZ4838. (NE) North-East: C9024, C7651, C9514, C7648, AZ4839. (C): C8985, C7576. (CN): C11519. (SW) South-West: C20477, C8979. (S) Southernmost region: C8976. Shaded pattern types represent novel pattern types found in Portuguese white clovers.

3.2.2 Predominance of type 15 in Spanish white clover plants.

Diversity was observed in ITS types of Spanish individuals, with prominence of pattern type 15, (Figs 3.15 and 3.16). Pattern type 5 was observed to a much lesser extent than 15 and heterozygote 20, and was not observed in accessions C9531 and C8975. Type 20 was absent from accessions C7600 (Spanish Pyrenees), C9027 (La Estrada, Galicia) and C8975 (Andalucía). Variant pattern types 15' and 20' were observed at low frequency. Additional diversity above these main types was observed in single plants. The rare type 34 was found in three out of five individuals from the North-West accession C9531 from Ponferrada. While the less frequent pattern types 8 and 26 were observed in accessions C7600 (Spanish Pyrenees) and accession C9034 (Galicia) respectively.

Table 3.17 Occurrence of nuclear pattern types in White Clover plants sampled from Spanish accessions.

Nuclear pattern type	Spain	Locations	% occurrence	No. Individuals
5	C9027, C7600, C8994, C9034, C9037	La Estrada (NW), Galicia (NW), Spanish Pyrenees Mts, Asturias (N-Coastal), Galicia (NW)	15.2	7
15	C9027, C9531, TR13, C8994, C9025, C9034, C9037, C8975,	Galicia: La Estrada (NW), Galicia (NW), Ponferrada, León & Castile (NW), Asturias (N-Coastal), Extremadura (W), Andalucía (S)	45.6	21
15'	C7600, C9034, C9037	Spanish Pyrenees Mts, Galicia (NW),	8.7	4
20	C9531, C8994, C9034, C9037, C9025	Ponferrada, León & Castile (NW), Asturias (N-Coastal), Extremadura (W), Galicia (NW)	15.2	7
20'	C8994, C9037	Asturias (N-Coastal), Galicia (NW)	4.3	2
26	C9034	Galicia (NW)	2.2	1
34	C9531	Ponferrada, León & Castile (NW)	6.5	3
9	C7600	Spanish Pyrenees Mountains	2.2	1
Totals	9		100	46

Note: A total of 46 individuals were assessed. Accession origins: (NW) North-West: C9027, C9034, C9037 and C9531. (N) North: C8994. (W) West: C9025. (S) South: C8975. Spanish Pyrenees Mountains: C7600. (C) Central: C8975. (N-Coastal): C8994.

Shaded pattern types present novel pattern types found in Spanish white clovers.

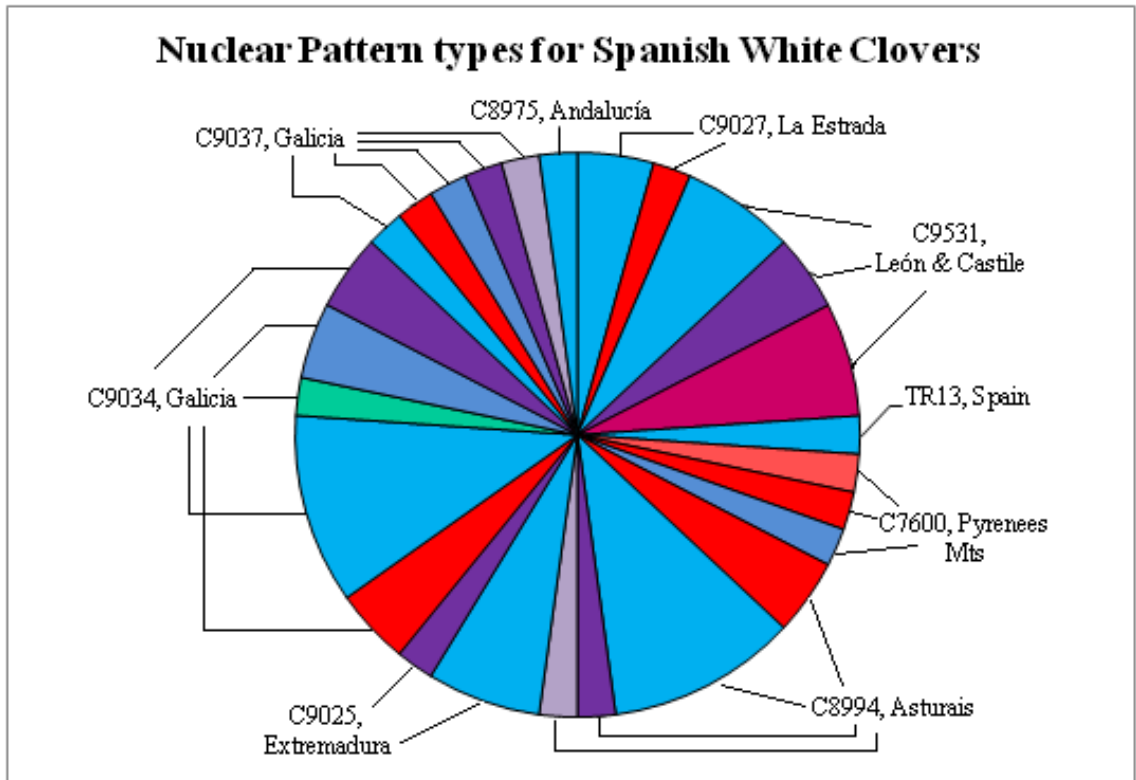


Figure 3.15 Distribution of nuclear pattern types listed per accession for plants with Spanish origins. Note: A total of 46 individuals were assessed. Accession origins: North-West: C9027, C9034, C9037 and C9531. North: C8994. West: C9025. South: C8975. Spanish Pyrenees Mountains: C7600.

As can be seen in the adjacent figure 3.16, just over half the pattern types observed in Spain were type 15, blue. While types 5, red and 20 collectively accounted for about a third of observed individuals.

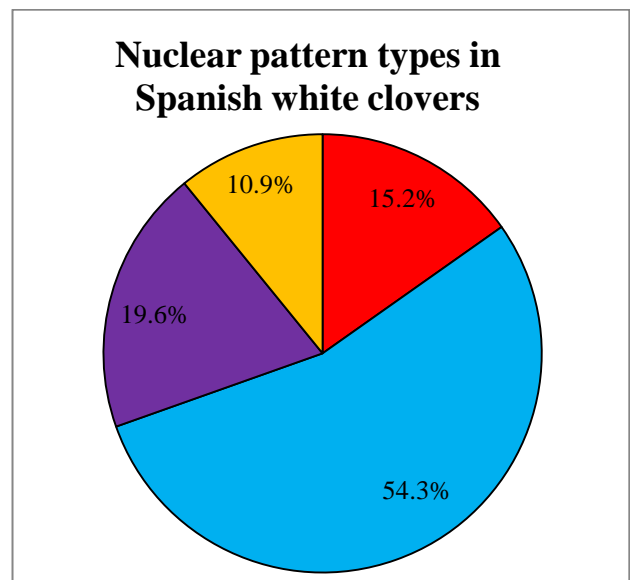


Figure 3.16 A summary of the main nuclear pattern types in Spanish white clovers. Pattern type colours blue 15, red 5, purple 20, orange others.

3.2.1 French White Clover shows a prominence of nuclear pattern type 15

The results for individuals from French origins show a prominence of pattern type 15, 59%. This was observed in all of the three cultivars and the three accessions sampled, particularly so in Crau, and to a lesser extent in Aran.

Pattern type 5 was only observed in the mixed cultivar Crusader, accounting for ~5% of the observed ITS types. Interestingly the population of plants from Crusader proved to be the most variable in nature.

Pattern type 20 showed wide spread presence and at 25.6% accounted for over half the remaining individuals, within the group, (Figs. 3.16 & 3.17). It predominates in accession C9046 and Crusader but was absent from the accessions C7690 (Bouches du Rhone, Entressen) South France and C2418 (Finistère, Bretagne) in North-West France. The variant pattern type 20' was observed at 7.7% and was more common in accession C9046 than was type 20. This brings the total class 20/20' to 33.3% occurrence, (Table 3.18). The presence of one individual with rare pattern type 7 was noted, (Table 3.18).

Table 3.18 Nuclear pattern types in individual French White Clover plants

Nuclear pattern type	Accessions	Locations	% Occurrence	No. Individuals
5	Crusader	French Crau & Syrian parentage	5.1	2
15	C7690, C9046, C2418, Crau, Aran, Crusader	Entressen, Bouches du Rhône (SE), Mouriès (S), Le Conquet (NW), Crau (SE), (French x Israeli), (French x Syrian- (Mader Village SW Syria)	59.0	23
20	C9046, Aran, Crau, Crusader	Mouriès (S), (France x Israel), Crau (SE), (French x Syrian- (Mader Village SW Syria)	25.6	10
20'	C9046, C18951	Mouriès (S), Crau (SE)	7.7	3
7	C18951	Crau (SE)	2.6	1
Totals	6		100	39

Assessed were a total of 39 individuals, from: (NW) North-West France – C2418, (SE) South-East: mixed cultivar Crusader (French Crau (SE) France x Mader Village, Zebdani District, Syria), C7690. (S) Southern France: C9046. Aran (mixed cultivar French and Israeli parentage). Shading represents rare pattern type 7.

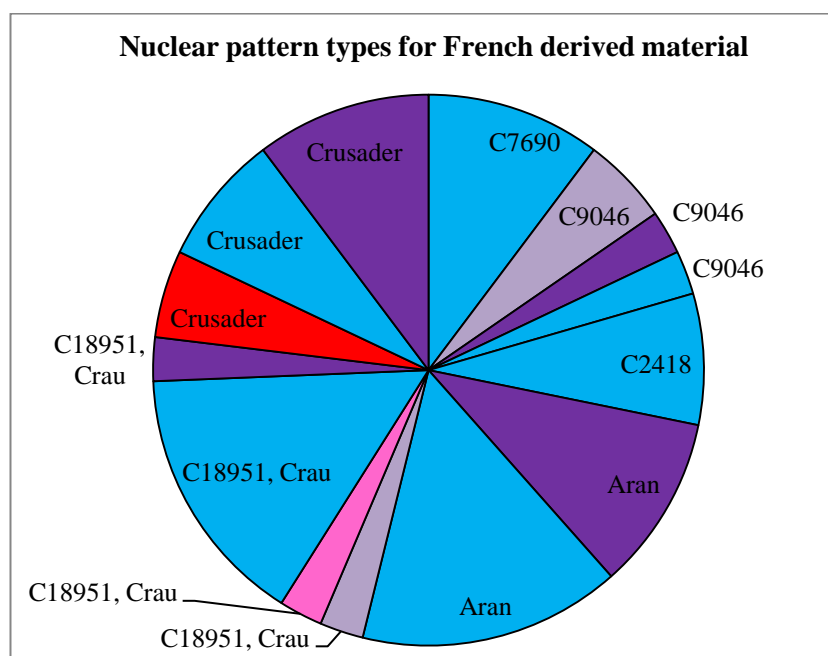


Figure 3.17 Distribution of nuclear pattern types listed per accession/cultivar for a total of 39 individuals from 3 accessions, two mixed origin cultivars and Ecotype Crau. Origins: North-West France – C2418, South-East France - mixed cultivar Crusader (French Crau South – East France and Mader Village, Zebdani District, Syria), and C7690. Southern France – C9046. Mixed cultivar Aran (Unknown French and Israeli locations).

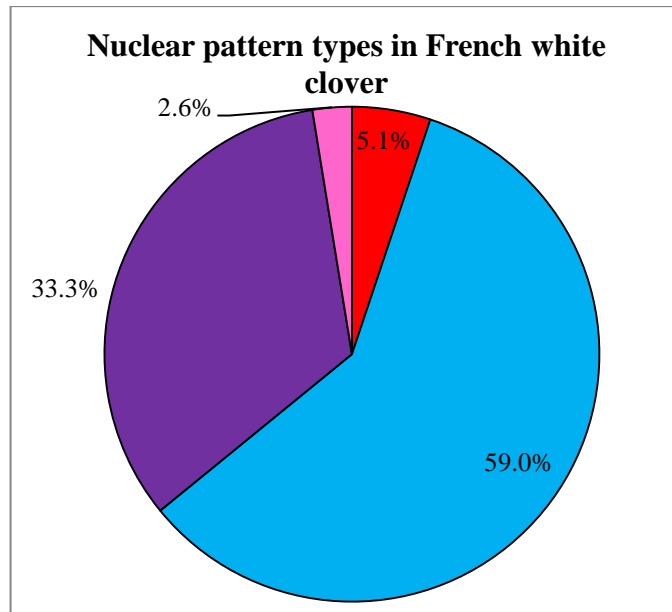


Figure 3.18 A summary of the relative approximate occurrences of main nuclear pattern types in French material. Pattern type colours: blue 15, red 5, purple 20, pink others =7.

3.2.6 Ecotype Italian Lodi, ecotype selection American Regal and accessions from Italy and USA, show similar prominence of main types 5 and 15.

It was observed that individuals from the Italian ecotype *Giganteum Lodigiano* or Lodi of Northern Italy and ecotype selection Regal showed similar presence of main pattern types 5, and 15. Heterozygous type 20 and 20' were observed in both, (Table 3.17 and Figs. 3.19 & 3.20). Slightly more diversity was observed in the Italian ecotype, with observation of 20' and two rare pattern types 36 and 48. One rare pattern type was observed in Regal, (Tables 3.17 and Figs 3.19 & 3.20). The two individuals of separate Italian origins showed the presence of the main pattern types.

Table 3.19 Nuclear pattern types in individual White Clover plants sampled from Italian Lodi, Ecotype cultivar Regal and accessions from Italy and USA

Nuclear pattern type	Italy/Regal	Locations	% occurrence	No. Individuals
5	C5897, TR11, C15111	Italian Giganteum Lodigiano (N), Pasubio (NE), Regal (SE, USA)	29.1	7
15	C5897, TR01, C15111	Italian Giganteum Lodigiano (N), Italy (?), Regal (SE, USA)	37.5	9
20	C5897, C15111	Italian Giganteum Lodigiano (N), Regal (SE, USA)	12.5	3
20'	C5897	Italian Giganteum Lodigiano (N)	4.2	1
36	C5897	Italian Giganteum Lodigiano (N)	8.3	2
48	C5897	Italian Giganteum Lodigiano (N)	4.2	1
9	C15111	Regal, USA Ladino Ecotype, (SE, USA)	4.2	1
Totals	4	4	100	24

Note: 24 individuals were assessed. Accession and cultivar origins: (NE) North East Italy: Pasubio – TR11. (N) Northern Italy: C5897- Po Valley. TR01(Italy unknown). (SE) South-Eastern USA: Alabama- Regal. Shaded pattern types present novel found pattern types.

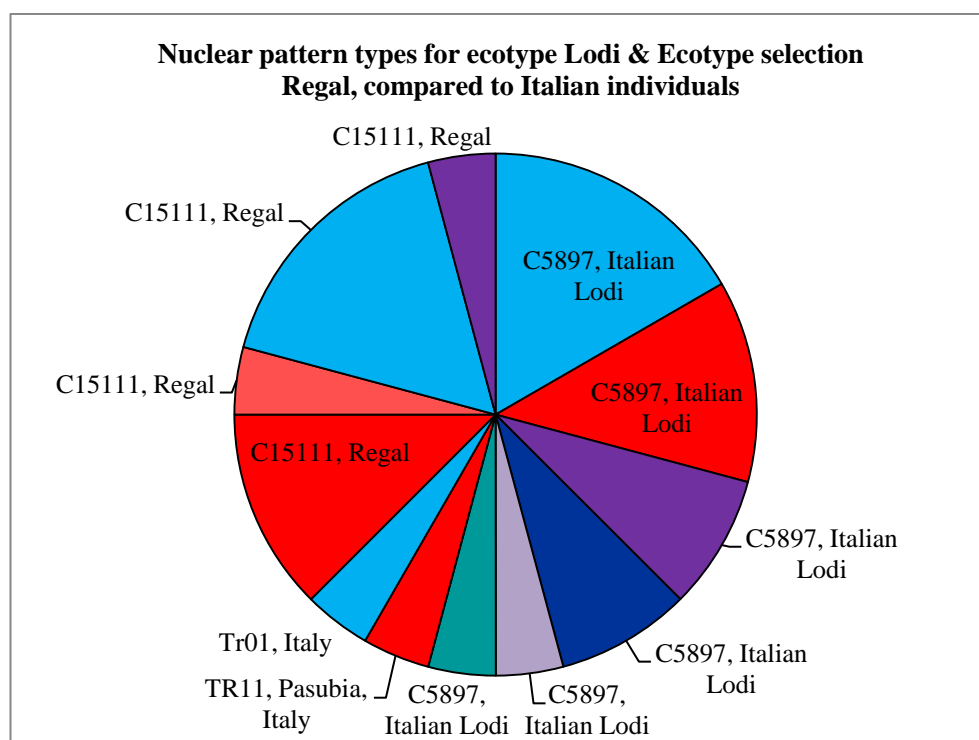
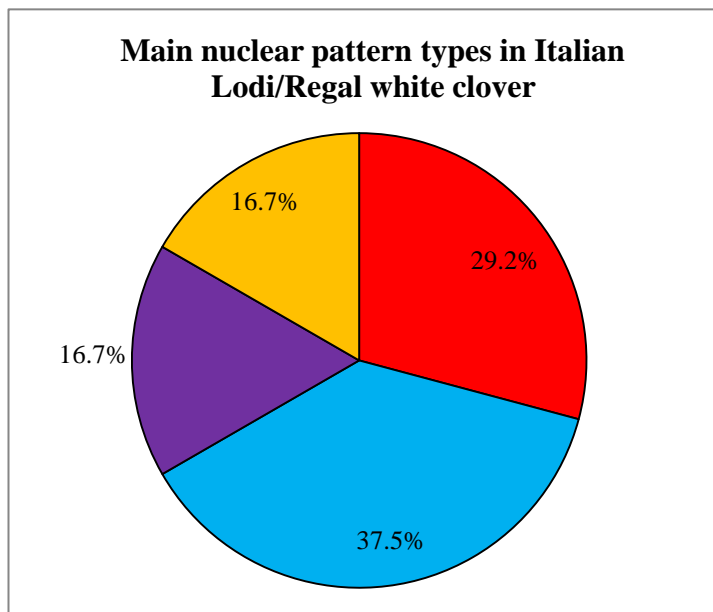


Figure 3.19 Distribution of nuclear pattern types listed per cultivar/accession for plants with Italian Lodi origins, the cultivars Italian Lodi and USA Ecotype cultivar Regal. Accession and cultivar origins: North East Italy: Pasubio –TR11. Northern Italy: C5897- Po Valley. South-Eastern USA: Alabama- Regal.



As can be seen from Fig. 3.19 types 5 and 15 accounts for more than half the observed pattern types seen. Pattern type 20 is no more common than the infrequent pattern types.

Figure 3.20 A summary of the relative distributions of the main nuclear pattern types. red- 5, blue- 15, purple -20, and orange- others.

3.2.2 British White Clover plants show a diverse collection of ITS types

The white clover from British origins showed a lot of variability in ITS types. Main type 5 was observed in the Kent landrace (C6374/C6304) and mixed origin cultivar Blanca. Type 15 occurred in Kent Wild White, (C6304) and S.184 (C21105). The variant form 15' was observed in Kent White (C6374) and in Blanca (C6360). 20' was observed in Kent White (C6374).

Nine rare pattern types were observed, (Table 3.18), of which three (31, 42, and 43) showed similarity to pattern types 5 and 15, and were unique to this group. Type 31 is heterozygous at two positions and seems to be produced from the inheritance of types 34 and something similar to type 6, (Table 3.12). Type 43 is unique in having position 509 as heterozygous, designated as (R). Type 42 observed once is heterozygous at positions 509 (Y), and 50 (W). Variation at this second position is what makes the type 42 different from the common type 20.

The rare pattern types 10, 32, 34, 35, and 39 were also observed in accessions from outside of Britain.

Table 3.20 Occurrence of nuclear pattern types in White Clovers plants from British origins.

Nuclear pattern type	Accessions	British Cultivars/Landraces	% Occurrence	No. Occurrences
5	C6374, C6304, C6360	Kent White	17.9	5
15	C6304, C21105	Wild White Kent, S.184-Wild Britain	10.7	3
15'	C6374, C6360	Kent White, Blanca	7.1	2
20	C6304, C6360	Wild White Kent, Blanca	21.4	6
20'	C6374	Kent White	3.6	1
32	C6374, C21105	Kent White, S.184-Wild Brittan	7.1	2
17'	C6360	Blanca- Belgium Kersey mix	3.6	1
34	C6304	Wild White Kent	3.6	1
39	C6304	Wild White Kent	3.6	1
31	C3754	Wild White Kent	3.6	1
35	C6374	Kent White	3.6	1
42	C6374	Kent White	3.6	1
10	C6374	Kent White	3.6	1
43	C6360	Blanca- Belgium Kersey mix	3.6	1
6	C6304	Wild White Kent	3.6	1
Totals			100	28

Note: Included are landrace Kent (Wild White, Kent White), cultivar S.184 and mixed origin cultivar (Belgium/Kersey) Blanca. A total of 28 individuals were sampled. Shaded pattern types present novel found pattern types

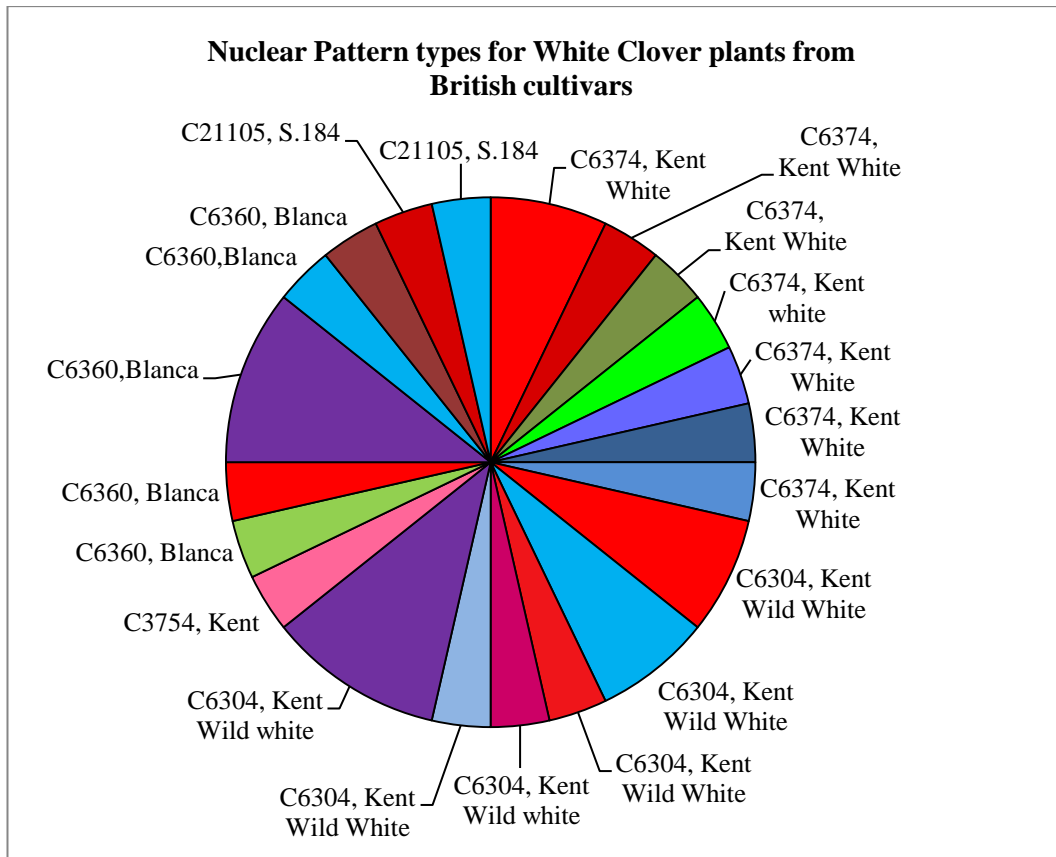


Figure 3.21 Distribution of nuclear pattern types listed per cultivar for plants from the landrace Kent (Wild White, Kent White), cultivar S.184 and mixed origin cultivar (Belgium/Kersey) Blanca. A total of 26 individuals were sampled.

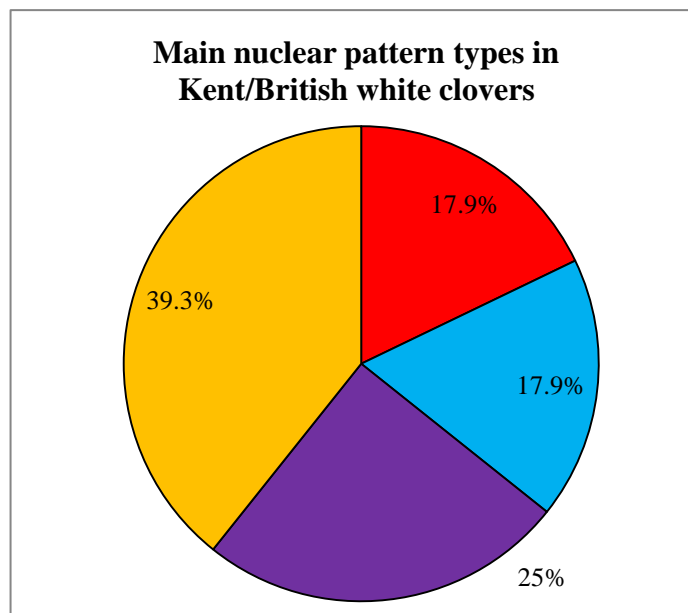


Figure 3.22 A summary of the relative distributions of the main nuclear pattern types. red- 5, blue- 15, purple -20, and orange- others.

3.2.3 White clover plants of Northern European origin, show prominence of 15, 20 and novel nuclear pattern type 17.

The results from the European group individuals showed the prominence of main type 15, the heterozygote 20, and low occurrence of pattern type 5 in Morso, Barbian and Tammisto. A fourth pattern type, 17, which appears largely restricted to Northern Europe, had the second highest occurrence.

This European pattern type 17 was shown to predominate in Morso, and with equal a frequency to type 15 in Barbian. It appears that where pattern type 17 predominates, occurrence of 20 is reduced, while the converse is also true, as is observed in Tammisto. Five rare pattern types were also observed in cultivars Morso and Tammisto.

Table 3.21 Occurrence of nuclear pattern types in Northern European cultivars and accessions

Nuclear pattern type	Accession Numbers	Cultivars & Selection Origins	% Occurrence	No. Individuals
5	C2095, C24247, TR14	Morso (N), Barbian (NW), Norway (N)	11.1	4
5'	C24247	Barbian (NW)	2.8	1
15	C2095, C24247, C960, C21105	Morso (N), Barbian (NW), Tammisto (N)	30.5	11
20	C24247, C960	Barbian (NW), Tammisto (N)	8.3	3
17	C2095	Morso (N)	5.5	2
17'	C2095, C24247, C960	Morso (N), Barbian (NW), Tammisto (N)	27.8	10
18	C2095	Morso (N)	2.8	1
19	C2095	Morso (N)	2.8	1
32	C2095	Morso (N)	5.5	2
8	C2095	Morso (N)	2.8	1
40	C960	Tammisto (N)	2.8	1
24	C960	Tammisto (N)	2.8	1
Totals	4	5	100	36

Note: Origins: Tammisto – (Finland), Morso – (Denmark), Barbian – (The Netherlands).

A total of 36 individuals were sampled.

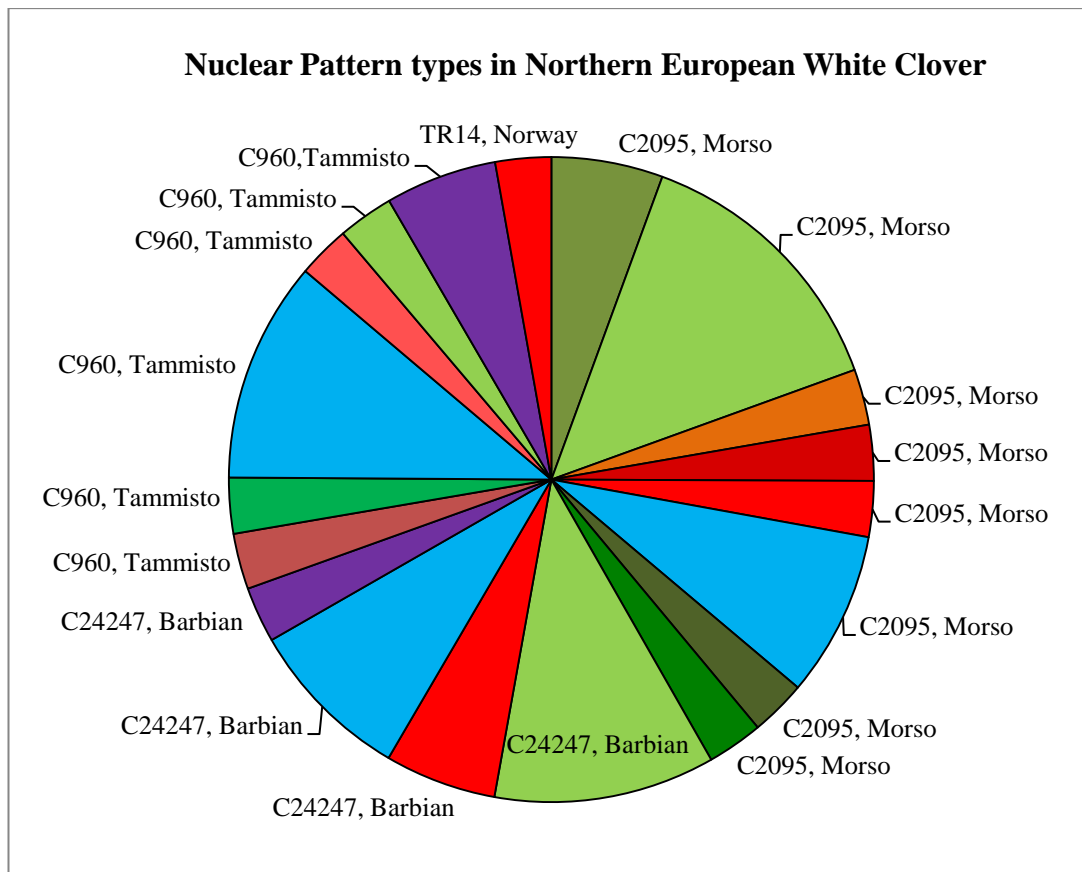


Figure 3.23 Distribution of nuclear pattern types listed per ecotype selection/accession for plants with origins in Northern Europe. Included are cultivars Tammisto – (Finland), Morso – (Denmark), Barbian – (The Netherlands). A total of 38 individuals were sampled.

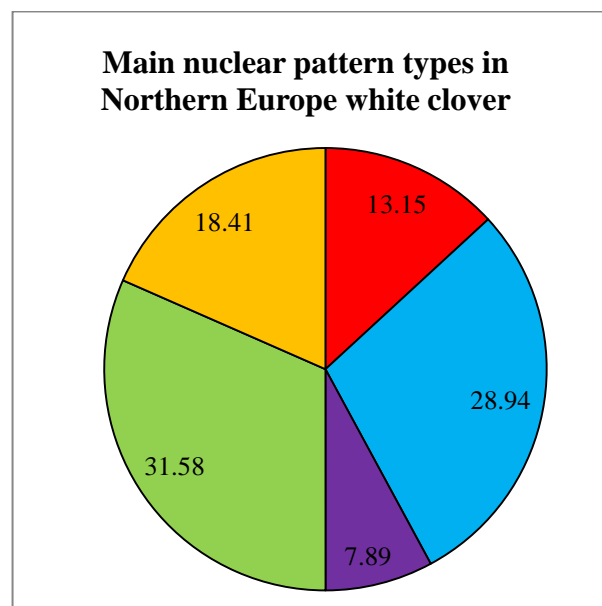


Figure 3.24 A summary of the relative portions of nuclear pattern types in Northern European white clovers. Pattern type colours: blue 15, red 5, purple 20, green 17/17'.

3.2. Nuclear pattern types in Central, and Western European white clover, including mixed cultivar Blanca are predominated by patterns types 15 and 20.

It was observed that in this selection of individuals the most common type was 15, 38.9%, being observed in Blanca, Romania and Germany. It was observed that in central European populations of white clover the predominant nuclear pattern types were 15 and 20. Main pattern type 5 was only observed in Blanca. Evidence of the primarily Northern European pattern type 17, was observed in Blanca, with the related type 19 being seen in Ovack.

Three rare pattern types were observed, 39, 43; these occurred in Ovack, Blanca and an individual from Barro Trro, Romania in central South West Europe.

Table 3.22 Occurrences of Nuclear pattern types in central, western and eastern Europe

Nuclear pattern type	Accessions	Accession & Cultivar Origins	% Occurrence	No. Individuals
5	C6360	Blanca (W)	5.6	1
15	C7555, TR06, TR08	Ovack (C), Domnesti-Romania (E), Hausen, Baden-Wurttemberg, Germany (C)	33.3	6
15'	C6360	Blanca (W)	5.6	1
20	TR10, C6360	(Pogoniani, Ioannina, Epirus)-Greece (SE) Blanca (W)	22.2	4
17'	C6360	Blanca (W)	5.6	1
19	C7555	Ovack (C)	5.6	1
39	C7555	Ovack (C)	11.1	2
43	C6360	Blanca (W)	5.6	1
44	N9	Barro Trro (SC2), Romania, (E)	5.6	1
Totals 9	5	5	100	18

Note: Assessed were 18 individuals. (C)- Central Europe, (E)- East Europe, (W)- West Europe, (SE) South- East Europe. Origins: Cultivars, Ovack, Czech Republic- Central Europe, TR06 near Domnesti, Romania, TR08, Hausen, Baden-Wurttemberg, Germany, TR10 Pogoniani Ioannina Epirus North-West Greece.

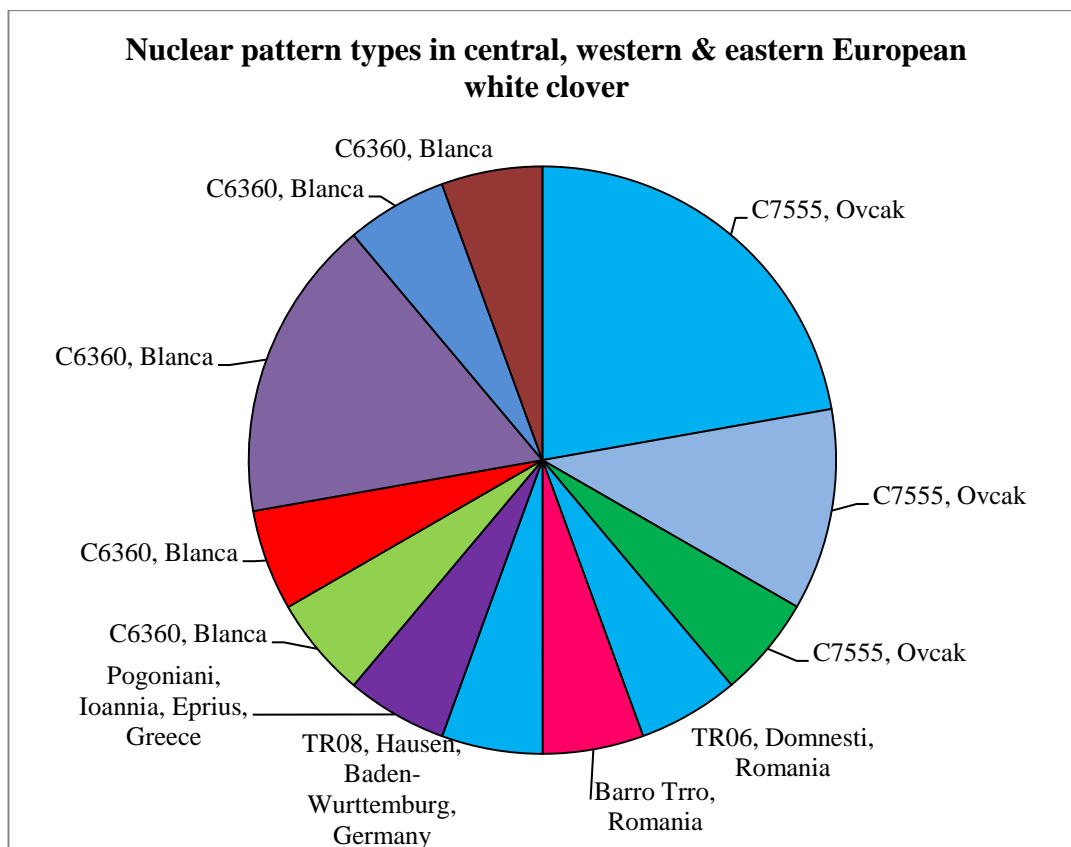


Figure 3.25 Occurrence of nuclear pattern types in populations from Central Europe and Western Europe, mixed cultivar Blanca. Origins: Blanca – (Belgium x Kersey mix- Continental Europe), Ovcak- (Czech Republic) – Central Europe, TR06 near Domnesti, (Romania), TR08, Hausen, Baden-Wurttemberg- (Germany), TR10 Pogoniani Ioannina Epirus North-West (Greece). Assessed were 18 individuals.

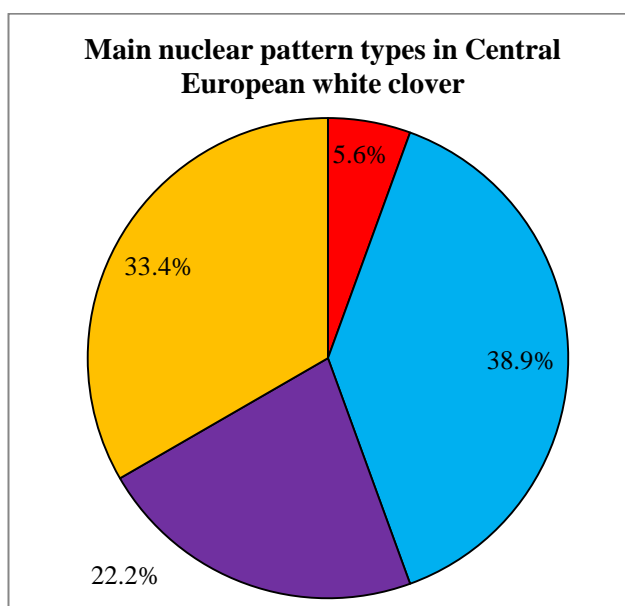


Figure 3.26 A summary of the main nuclear pattern types in Central European white clover populations and from mixed cultivar Blanca (Continental Europe). Pattern type colours: blue 15, red 5, purple 20, orange others.

3.2.7 White clover of Israeli origins showed the prominence of nuclear pattern type 15. Greater variation in pattern types was observed in Tamar and Haifa.

The results for plants containing Israeli White Clover germplasm showed a prominence of pattern type 15, 53.6% and secondary prominence for type 20 21.4%. Type 5 was of low occurrence and was observed only in Haifa and Tamar. A clear difference in observable variation was seen between the Israeli cultivars and mixed cultivar Aran; both Haifa and Tamar showed more variability in ITS types. However pattern type 20 showed significantly higher occurrence in Aran than either Haifa or Tamar.

In this group a fourth more prominent pattern type, 10 was observed. This was scored in only Haifa, but showed greater frequency than type 5 for the cultivar and for the group. Other less frequent pattern types observed were, 23 and 16G, (Table 3.22).

Table 3.23 Percentage occurrence of nuclear pattern types in White Clover plants with Israeli origins

Nuclear pattern type	Cultivars	Locations	% occurrence	No. Individuals
5	C4143, C5892	Tamar, Haifa (NE)	7.1	2
15	C4143, C5892, Aran, TR02	Tamar, Haifa, Aran, TR02	53.6	15
20	C4143, C5892, Aran	Tamar, Haifa (NE), Aran	21.4	6
10	C5892	Haifa (NE)	10.7	3
16G	C4143	Tamar	3.6	1
23	C4143	Tamar	3.6	1
Totals	4	4	100	28

Note: A total of 28 individuals were sampled from cultivars Tamar, Haifa (NE) - North East Israel, cultivar of mixed origin Aran (French and Israeli parentage) and individual TR02 -Israeli. Shaded boxes indicate novel pattern types.

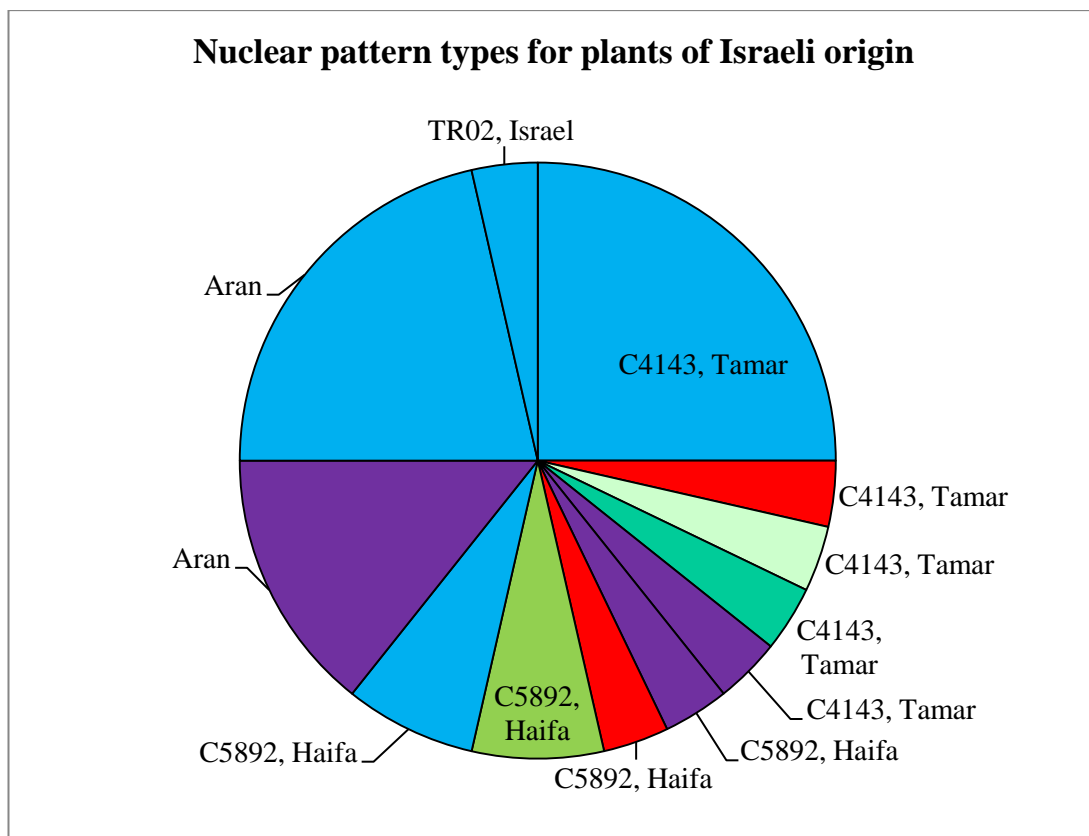


Figure 3.27 Distribution of nuclear pattern types listed per cultivar/accession for plants with Israeli origins. Included are cultivars Tamar, Haifva and mixed origin (French and Israeli) cultivar Aran. Also included is individual TR02 of Israeli origin. A total of 26 individuals were sampled.

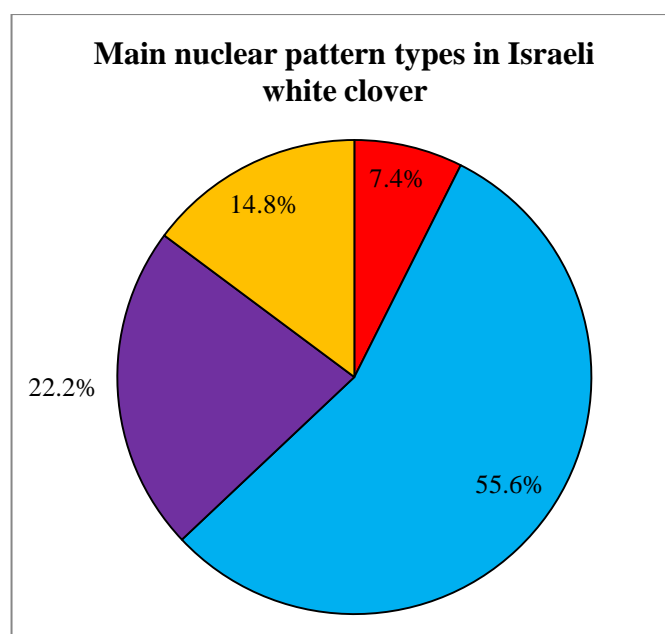


Figure 3.28 a summary of nuclear pattern types in white clover populations from Israeli origins. Pattern types; 15- blue, 5- red, 20 purple and others orange.

3.2.8 Nuclear Results for plants from North African accessions show a division in prominence of ITS type.

A division was observed in the North African accessions, where type 5 was more frequent than type 15 in Algerian accessions and showed further prominence of pattern types most similar to type 5. Individual examples from Moroccan accessions showed prominence of types 15 and 35. Twelve different rare types specific to each country were observed. Rare type 9 was the only pattern type observed in both accessions from Algeria C6452 and Morocco C25478 (Setti Fatma) and C25480 (south of Asni).

Twelve pattern types, (1, 7, 9, 11, 12, 21, 25, 27, 30, 35, 38, and 39) rare type 35 were found across different accessions from Morocco. Novel heterozygous pattern type 38, similar to type 7 was observed in Algerian accessions C6452 and C1751. Of these, type 30 and type 39 were found in Algerian and Moroccan accessions respectively. Types 1, 11, 12, 21, 25, and 27 were each observed once across North Africa (Table 3.23).

Table 3.24 Nuclear pattern types in North African, Algerian and Moroccan accessions

Nuclear pattern type	Accession numbers	Locations	% Occurrence	No. Individuals
5	C6452, C1751	Algeria, Algeria	7.3	3
15	C6452, C1761, C25466, C25468, C25469, C25470, C25471, C25472	Algeria, Morocco: Bab-Berred (NW), Taza (N), Azrou (NW), Ain-Leuh-Ifrane (CN), Oukaimeden (SW), Bou-Laouane (N)	24.4	10
35*	C25467, C25473, C25474, C25476,	Morocco: Ketama (Central N), South of Asni (SW), North of Asni (SW), Taddert (SW)	9.8	4
39*	C25464, C25475, C25479	Morocco: Marrakech (SW), Irhem (SC), Imlil (C)	7.3	3
9*	C6452, C25478, C25480	Algeria, Morocco: Setti Fatma (SW), South of Asni (SW)	7.3	3
12	C6452, C1751	Algeria, Algeria	4.9	2
30*	C6452, C1751	Algeria, Algeria	14.6	6
11	C1751	Algeria	2.4	1
38	C1751	Algeria	2.4	1
7*	C6452	Algeria	9.8	4
25	C1761	Morocco	2.4	1
21	TR04	near Marrakech (SW)	2.4	1
27	C25477	Taddert (SW)	2.4	1
1	C25465	Chefchaouen (NW)	2.4	1
Totals 14	22		100	41

Note: Assessed were 41 individuals from throughout North Africa, (N) North: C25468, C25472. (CN) Central North: C25467. (NW) North-West: C25466, C25469, C25465. (C) Central: C25479 (SW) South-West: C25471, C25473, C25474, C25476, C25464, C25478, TR04, and C25477. (SC) South Central: C25475. (*) indicates rare pattern types that show higher occurrence in North Africa. Shaded boxes indicate rare pattern types present in North African white clover.

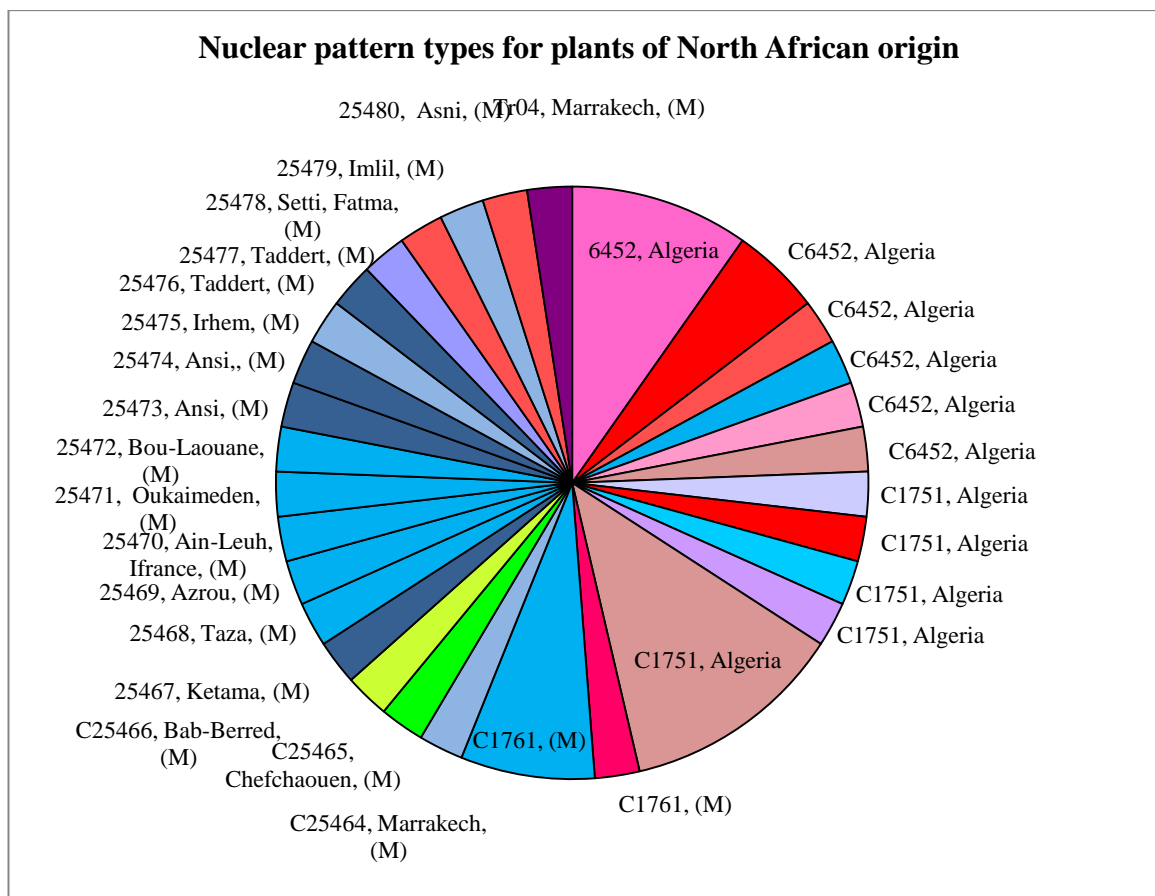


Figure 3.29 Distribution of nuclear pattern types listed per cultivar/accession for plants with African origins, including accessions from Morocco and Algeria. A total of 41 individuals were assessed. Note: (M) = Moroccan accession.

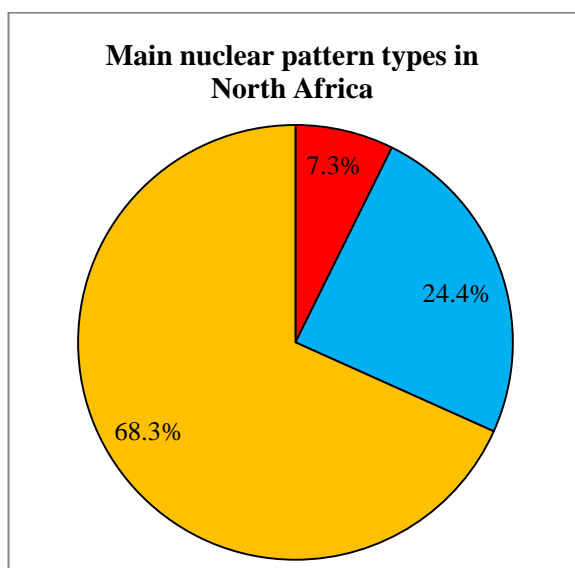


Figure 3.30 A summary of main nuclear pattern types in North African white clovers, from Algeria and Morocco. Pattern type colours: blue 15, red 5, purple 20, orange others.

3.2. 11 White clover populations from outside the natural distribution zone of white clover showed prominence of the main pattern types, 5, 15, and 20.

A selection of 15 individuals were sampled from zone 3 and treated as out-group individuals. These were from; South Africa, Costa Rica/USA, New Zealand, Kazakhstan and China. In this group the main nuclear pattern types 5, 15, and 20 are evident. Type 5 was observed in individuals from Costa Rica and the Republic of Kazakhstan. Type 15 was observed in individuals from the South African cultivar Dusi, and from China. Pattern type 20 was observed in the New Zealand Huia reference, and in three separate individuals, from Dusi, Ethiopia, and Japan. Two rare pattern types, 41 (Georgia, south eastern USA) and 12 (Dusi) were observed.

Table 3.25 Nuclear pattern types in White clover populations outside of the natural distribution zone

Nuclear pattern type	Reference Number	Locations/Origins	% Occurrence	No. Individuals
5	CR138941, TR03, C16511, N103	Costa Rica (C/USA), USA, Dusi (SA), Alma-Ata, Alma-Ata (K)	33.3	5
15	Reference sequence, N64, C16511, N107, N96	Huia (Ref) (NZ), Costa Rica, Dusi (SA), Urumqi-Xinjiang (C), China (C)	33.3	5
20	C16511, N98, N100	Dusi (SA), Ethiopia, Japan	20.0	3
41	N109	Georgia (SE/USA)	6.7	1
12	C16511	Dusi (SA)	6.7	1
Totals 5	11	11	100	15

Note: 15 individuals were included in this selection of white clover populations from outside the natural distribution zone of white clover. (C/USA) Central USA: Costa Rica. (SA) South Africa: Dusi. (NZ) New Zealand: Huia Reference sequence. (C) China. (SE/USA) South East USA: N109. (K) Republic of Kazakhstan, Eurasia.

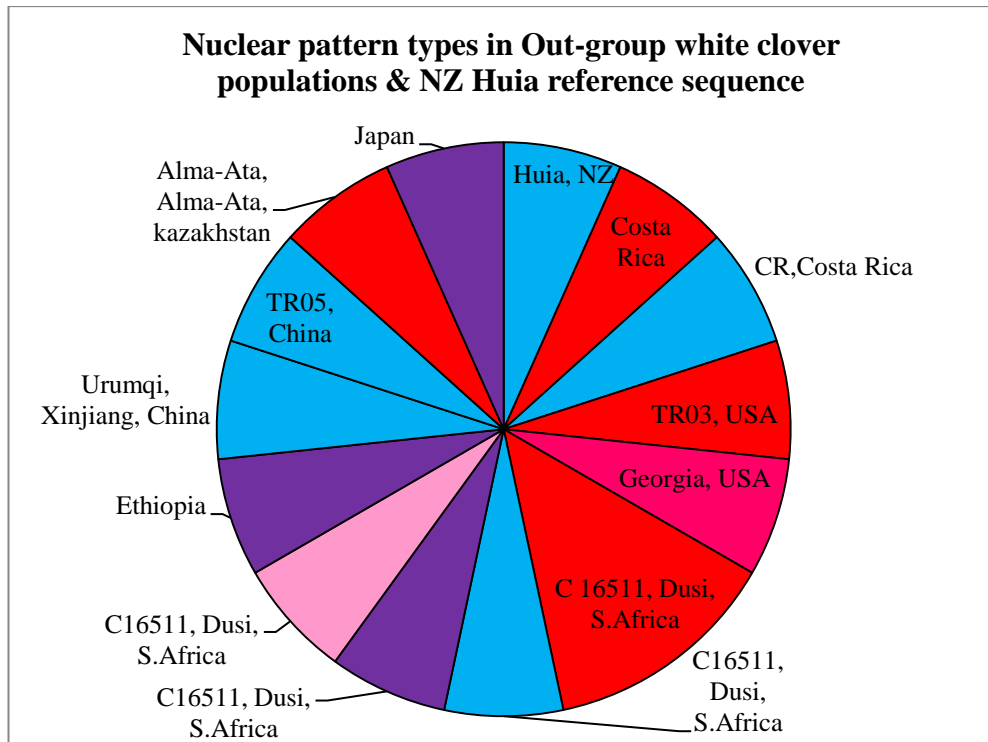


Figure 3.31 Out-group populations of white clover, from outside the natural distribution zone. Shown are representatives from, Costa Rica, USA, China, Japan, Kazakhstan and South Africa.

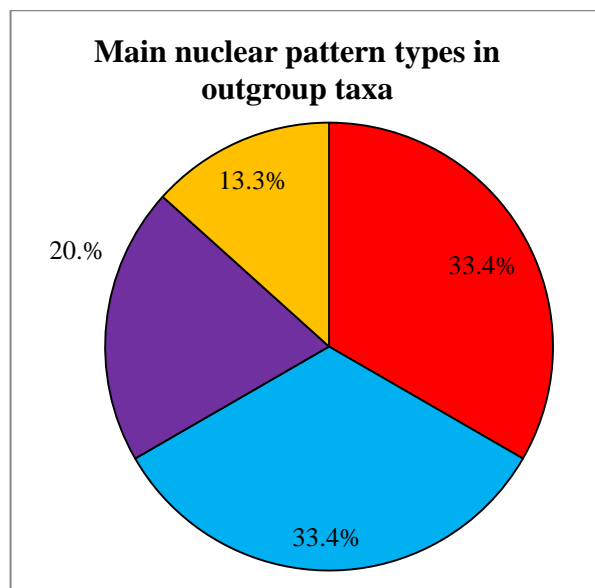


Figure 3.32 A summary of the main nuclear pattern types in Out-group populations of white clover, from outside the natural distribution zone. Included are representatives from, Costa Rica, USA, China, Japan, Kazakhstan and South Africa. Pattern type colours blue 15, red 5, purple 20 and orange other (12 and 41).

3.3 Results for the combined pattern types generated from the combination of the chloroplast and nuclear pattern types.

3.3.1 The complete picture, analysis of 218 individuals

Introduction

Data from the two chloroplast genetic regions and the nuclear ITS region was collected from 218 individuals. These were from different numbers of individuals per population, for a range of accessions, and commercial lines. Samples were also from different geographic regions. The picture shown by these individuals is presented here, firstly as a comparison across the data set and then by different geographic regions, as in prior sections. A complete table of the 218 individuals for which combined analysis was carried out can be found in Appendix E. For each geographic section, the most common pattern combinations were indentified.

Chloroplast and main ITS pattern types

The data showed that the most common chloroplast pattern type, CC+, was typically associated with the nuclear pattern type 15, and this combination was the most frequent pattern combination across the data set, 42 individuals or 19%. The main nuclear pattern type 5, (14%) and the common heterozygous type 20 (9%) were the next most common types observed in conjunction with chloroplast type CC+. Collectively the main pattern types classes 15, 5, and 20 including variant forms 5', 15' and 20' accounted for the largest number of individuals, 92, or 42% of the data set.

The less common chloroplast pattern type CT+ was observed most prominently in conjunction with nuclear pattern type 15, accounting for 9% or 20 of the 218 individuals.

The restricted chloroplast pattern type TC-, which was most prolific in Portuguese and Spanish populations, showed the highest association to nuclear pattern type 15/15', equating to 15 individuals or 7%.

Minor pattern combinations

41 individuals or 19% of the data set combined rare nuclear types observed with the common chloroplast pattern type CC+.

The very infrequent pattern type CC- was observed with rare nuclear forms 3 and 5, in 3 of the 218 individuals 1%.

Table 3.24 Occurrences of pattern combinations containing the main nuclear pattern types and the commonest chloroplast pattern type CC+.

Chloroplast pattern type	Nuclear pattern type	No. Occurrences	% Occurrence
CC+	15	42	19.3%
CC+	5	29	13.3%
CC+	20	14	6.4%
CC+	17	2	0.9%
CC+	17'	7	3.2%
CC+	20'	5	2.3%
CC+	15'	1	0.5%
CC+	5'	1	0.5%
Totals	8	101	46.4%
Rest of dataset	41	117	53.6%

Table 3.25 Occurrences of pattern combinations containing the chloroplast type TC-.

Chloroplast pattern type	Nuclear pattern type	No. Occurrences	% Occurrence
TC-	15	14	6.5%
TC-	5	11	5.1%
TC-	20	7	3.2%
TC-	20'	4	1.8%
TC-	15'	2	0.9%
TC-	34	2	0.9%
TC-	2	1	0.5%
TC-	10	1	0.5%
Totals	8	42	19.4%

Note: The distribution was primarily limited to Spain and Portugal, with additional individuals from French Crau, Kent White, Italian Lodi, and Algerian accession C6452.

Table 3.26 Occurrences of pattern combinations containing the chloroplast pattern type CT+.

Chloroplast pattern type	Nuclear pattern type	No. Occurrences	% Occurrence
CT+	15	18	8.3
CT+	20	6	2.8
CT+	5	1	0.5
CT+	15'	1	0.5
CT+	17'	1	0.5
CT+	7	1	0.5
CT+	39	1	0.5
CT+	32C	1	0.5
CT+	34	1	0.5
Total		31	14.60%

5.1.1

Table 3.27 Occurrences of nuclear pattern types with the rare chloroplast pattern type CC-.

Chloroplast pattern type	Nuclear pattern type	No. Occurrences	% Occurrence
CC-	15	1	0.4%
CC-	5	1	0.4%
CC-	3	1	0.4%
Totals	3	3	1.2%

3.3.2 Chloroplast and nuclear pattern combinations as listed by geographic distribution

3.3.2.1 Multiple common pattern types, primarily containing chloroplast type TC- were observed in Portuguese accessions.

This group of Portuguese individuals showed diversity at both the nuclear and chloroplast levels. Among the 42 Portuguese individuals, 20 had chloroplast pattern type TC- and 15 had type CC+. Four pattern combinations, CC+/5 (7 plants), TC-/5 (8 plants), TC-/15 (4 plants) and TC-/20 (6 plants) were most common. Three individuals were observed that had chloroplast type CT+, in conjunction with main types 15 and 20. The CC- combination was observed in conjunction with rare nuclear type 3 and main type 5 in two individuals of accessions C8979 and C20477.

Other rare pattern combinations observed were the heterozygote type CC+/45, TC-/2, CC+/2, CC+/32 and CC+/34.

Eighteen different pattern combinations were observed in the Portuguese populations, often in single individuals. The accessions C20477, C8979 and C9516 each showed four pattern combinations in four, five and five individuals respectively.

Table 3.28 Occurrence and pattern type for combined pattern types for Portuguese White Clover accessions.

Accession Numbers	Pattern	Occurrences
C7576, Estremadura, (C)	CC+/15	1/1
C7582, Beira littoral, (NW)	TC-/5	3/6
C7582, Beira littoral, (NW)	TC-/20	1/6
C7582, Beira littoral, (NW)	TC-/20'	1/6
C7582, Beira littoral, (NW)	CC+/15	1/6
C8976, Algarve, (S)	TC-/2	1/3
C8976, Algarve, (S)	TC-/5	1/3
C8976, Algarve, (S)	TC-/15	1/3
C8979, Ribatejo, Grandola, (SW)	CC+/2	1/5
C8979, Ribatejo, Grandola, (SW)	CC-/5	1/5
C8979, Ribatejo, Grandola, (SW)	TC-/5	2/5
C8979, Ribatejo, Grandola, (SW)	TC-/15	1/5
C8985, Coimbra, (C)	CC+/5	2/4
C8985, Coimbra, (C)	CC+/15	1/4
C8985, Coimbra, (C)	CC+/17	1/4
C9514, Tras Os Montes, (NE)	CC+/5	1/4
C9514, Tras Os Montes, (NE)	CC+/32	2/4
C9514, Tras Os Montes, (NE)	TC-/15	1/4
C9516, Tras Os Montes, Vila Verde,	CT+/15	2/5
C9516, Tras Os Montes, Vila Verde,	CT+/20	1/5
C9516, Tras Os Montes, Vila Verde,	TC-/5	1/5
C9516, Tras Os Montes, Vila Verde,	CC+/34	1/5
C11519, Vila Pouca de Agu, (CN)	CC+/5	2/3
C11519, Vila Pouca de Agu, (CN)	CC+/20	1/3
C20477, Águas de Moura, Pamela, (SW)	TC-/5	1/4
C20477, Águas de Moura, Pamela, (SW)	CC+/2	1/4
C20477, Águas de Moura, Pamela, (SW)	CC+/45	1/4
C20477, Águas de Moura, Pamela, (SW)	CC-/3	1/4
AZ4838, Minho, near Braga, Ferreiras,	TC-/15	1/1
AZ4839, near Baragonca, Tras Montes,	CC+/5	1/2
AZ4839, near Baragonca, Tras Montes,	CC+/20'	1/2
C9024, Tras Os Montes, Esphino, (NE)	CC+/5C	1/6
C9024, Tras Os Montes, Esphino, (NE)	TC-/20	5/6

Note: A total of 44 individuals from Portuguese accessions were assessed. Origins: (N) North, (NW) North-West, (NE) North-East, (CN) Central North, (C) Central: (SW) South-West, (S) Southernmost region. The second column shows the components of the pattern type. The first number in the third column is the number of individuals showing that pattern combination, the second number is the total of individuals assessed for that accession.

3.3.6 The combination of chloroplast type TC- and nuclear type 15 was prominent in Spanish accessions

The pattern combination TC-/15 was the most frequent combination in Spanish individuals. Collectively the TC- and main nuclear forms (5, 15, and 20) had the highest occurrences across the group, being found in five of seven accessions and in 15 of 22 individuals. Chloroplast types CC-, CT+ and CC+ were also observed in combination with the main nuclear pattern types, 5, 15, and 20. Three rare pattern combinations were observed with chloroplast type TC- and main types 5, 15, and 20.

The group of Spanish individuals showed thirteen different pattern combinations across seven accessions including representation of each of the four chloroplast types.

Table 3.29 Percentage occurrences for combined pattern types of sampled Spanish accessions.

Accession Numbers	Pattern Type	Occurrences
C7600, Pyrenees Mts (NE)	CC+/ 5	1/1
C8994, Asturais (N)	TC-/5	2/8
C8994, Asturais (N)	TC-/15	5/8
C8994, Asturais (N)	TC-/20'	1/8
C9025, Extremadura, (W)	CC-/15	1/1
C9027, La Estrada, Galicia, (NW)	CT+/5	1/2
C9027, La Estrada, Galicia, (NW)	CC+/15	1/2
C9034, Galicia, (NW)	CT+/15	1/3
C9034, Galicia, (NW)	TC-/15	1/3
C9034, Galicia, (NW)	TC-/15'	1/3
C9037, Galicia, (NW)	TC-/5	1/4
C9037, Galicia, (NW)	TC-/15	1/4
C9037, Galicia, (NW)	TC-/20'	1/4
C9037, Galicia, (NW)	TC-/15'	1/4
C9531, Leon Castilla, Leon Ponferrada, (NW)	TC-/15	2/5
C9531, Leon Castilla, Leon Ponferrada, (NW)	TC-/20	1/5
C9531, Leon Castilla, Leon Ponferrada, (NW)	TC-/34	2/5

Note 24 individuals were sampled. For footnote see table 3.28.

(N)- North, (NW)- North West, (NE)- North East.

3.3.7 Combined pattern types of novel variant in *T. repens* individuals from Portugal and Spain

Of the four individuals that showed a novel chloroplast sequence, only the individual from Spanish accession C8975 was scored for ITS variation. This individual had the most common nuclear pattern type 15.

3.3.1 Pattern combination type CC+/15 was the most common in French material.

The most common pattern type combination was chloroplast type CC+/15. This occurred in at least one individual in each accession and cultivar. In the three individuals of accession C2418 it was the only pattern type observed. Each accession or cultivar showed the presence of a predominant pattern combination. A three way split in pattern types was observed for nine individuals from mixed cultivar Aran. Six of these individuals contained chloroplast pattern type CC+, three with pattern type 15 and three with pattern type 20. The remaining 3 individuals combined chloroplast type CT+ and nuclear pattern type 15.

Uncertain pattern types

One example of a minor nuclear pattern type, type 7, was observed in conjunction with the less frequent chloroplast pattern type CT+.

In nine cases pattern combinations were observed (instances where the pattern combination was observed in one individual). In eight of these cases the combination involved either chloroplast pattern types CT+ or CC+ in conjunction with one of the main nuclear types 15, 20, and 20', or 5. The combination CC+/15 was the most common.

Table 3.30 Combined pattern types listed per accession or cultivar for white clover from French origins.

Accession No. & Cultivar names	Pattern Type	Occurrences
C7690, Bouches du Rhône, (SE)	CT+/15	3/4
C7690, Bouches du Rhône, (SE)	CC+/15	1/4
C9046, Mouries, Bouches du Rhône, (S)	CC+/15	1/4
C9046, Mouries, Bouches du Rhône, (S)	CC+/20	1/4
C9046, Mouries, Bouches du Rhône, (S)	CC+/20'	2/4
C18951, Crau, (SE)	CT+/7	1/9
C18951, Crau, (SE)	CT+/15	6/9
C18951, Crau, (SE)	CT+/20	1/9
C18951, Crau, (SE)	CT+/20'	1/9
Aran	CT+/15, (3- way)	3/9
Aran	CC+/15, (3- way)	3/9
Aran	CC+/20, (3- way)	3/9
Crusader	CT+/15	1/7
Crusader	CT+/20	2/7
Crusader	CC+/5	1/7
Crusader	CC+/15	2/7
Crusader	CC+/20	1/7
C2418, Le Conquet, (NW)	CC+/15	3/3

Note: 36 plants were assessed for combined pattern types. Columns as for Table 3.28. Origins: (SE), South- East, (S), South and (NW), North-West.

The cultivars, French Crau and Crusader were variable, having 5 pattern types each. Of which four were various combinations of the main chloroplast and nuclear pattern types. Throughout the rest of the accessions and cultivars, the common pattern combinations were observed as; CC+/15, CC+/20 and CT+/15.

In the Crusader cultivar, the pattern combinations CT+/20 and CC+/15 were most common, while in French Crau, CT+/15 was observed in 6 out of 9 individuals, the remaining 3 each showing a different pattern combination.

The mixed cultivar Aran of (French and Israeli ancestry) was also variable having a three way split in pattern combinations, derived from CC+/15 or 20 and CT+/15.

3.3.2 White clover from Kent and British cultivars showed considerable variation in combined pattern types.

This group of British or part-British white clovers showed considerable variation in pattern combinations. Across the group, only combination CC+/5 was observed repeatedly. CC+ was frequently observed in conjunction with one of the main nuclear pattern types.

In Blanca, CC+/20 was prominent, while a two way split CC+/5 and CT+/20 in pattern types was observed in Kent Wild White. The other Kent population, Kent White (C6374) was quite variable.

One rare combination TC- and nuclear type 10 was observed in Kent White, C6374. The combinations CC+/15, CC+/15' and CC+/17' were observed in single individuals, as were the less frequent nuclear types 34 and 43, with chloroplast pattern types CT+ and CC+ respectively.

Of the four assessed Kent/British cultivars, Kent White was the most variable with seven different pattern combinations in seven individuals.

Blanca and Kent Wild White each had four different pattern combinations.

Table 3.31 Combined pattern types of Kent landrace white clover and the cultivar Blanca (English x Belgian).

Cultivar Numbers & Names	Pattern Type	Occurrences
C21105, S. 184, (British)	CT+/32C	1/1
C6304, Kent Wild White	CC+/5	2/6
C6304, Kent Wild White	CT+/20	2/6
C6304, Kent Wild White	CT+/34	1/6
C6304, Kent Wild White	CC+/15	1/6
C6374, Kent White	TC-/10	1/7
C6374, Kent White	CT+/15	1/7
C6374, Kent White	CC+/5	1/7
C6374, Kent White	CC+/6	1/7
C6374, Kent White	CC+/20'	1/7
C6374, Kent White	CC+/35	1/7
C6374, Kent White	CC+/42	1/7
C6360, Blanca, (W)	CC+/15'	1/6
C6360, Blanca, (W)	CC+/17'	1/6
C6360, Blanca, (W)	CC+/20	3/6
C6360, Blanca, (W)	CC+/43	1/6

Note: 20 individuals from Kent and Blanca were assessed. Columns as for Table 3.28.

Blanca = (Mixed cultivar (Kersey x Belgium), (W) - Western Europe).

3.3.4 Ecotype selections Italian Lodi (ladino clover) and American Regal, showed similarity in combined pattern types.

In this group the chloroplast type CC+ was observed with one of the main nuclear pattern types, 5, 15 or 20, in all but two individuals. Two rare pattern types were observed, (TC-/15) in Italian Lodi and (CC+/9) in Regal.

Italian Lodi showed a slightly higher level of diversity, with three common and one rarer pattern combination.

Table 3.32 Combined pattern types for white clover ecotype selections Lodi and American ladino, Regal.

Accession Numbers & Names	Pattern Type	Occurrences
C15111, Regal	CC+/5	3/7
C15111, Regal	CC+/9	1/7
C15111, Regal	CC+/15	3/7
C5897, Lodi	CC+/5	3/10
C5897, Lodi	CC+/15	3/10
C5897, Lodi	CC+/20	3/10
C5897, Lodi	TC-/15	1/10

Note: 17 individuals of Italian Lodi (C5897, *Giganteum Lodigiano*) and USA grown ladino clover-derived cultivar, Regal were assessed. See footnote, Table 3.28.

3.3.3 In North-Western European white clovers combination type CC+/15 was the most common. Combination CC+/17 was largely restricted to this region.

All common pattern types were composed of combinations of chloroplast types CC+ and CT+ with one of the main nuclear types. The most common pattern combinations were CC+/15 and the new combination CC+/17'.

Twelve pattern combinations containing chloroplast pattern type CC+ were recorded as rare. These combinations all had nuclear pattern types that occurred at least twice in the nuclear data set and are recognised as being true SNP variants. Four ambiguous pattern combination types containing CT+ were observed. They contained nuclear pattern types with single occurrences within the nuclear data set.

This group was diverse in showing 10 nuclear types in combination with two chloroplast types, making 15 different pattern type combinations. In Morso and Barbian three new and common pattern combinations were observed that included nuclear types 17' and 5'. The highest number of combinations was found in ecotype Morso from Denmark.

Table 3.33 Combined pattern types listed per cultivar for North and North Western European white clover populations.

Cultivar accession Numbers & Names	Pattern Type	Occurrences
C2095, Morso, Denmark, (NW)	CC+/5	1/11
C2095, Morso, Denmark, (NW)	CC+/8	1/11
C2095, Morso, Denmark, (NW)	CC+/15	2/11
C2095, Morso, Denmark, (NW)	CC+/17	1/11
C2095, Morso, Denmark, (NW)	CC+/17'	3/11
C2095, Morso, Denmark, (NW)	CC+/18	1/11
C2095, Morso, Denmark, (NW)	CC+/19	1/11
C2095, Morso, Denmark, (NW)	CC+/32	1/11
C24247, Barbian, The Netherlands, (NW)	CT+/15	2/10
C24247, Barbian, The Netherlands, (NW)	CC+/5	2/10
C24247, Barbian, The Netherlands, (NW)	CC+/15	1/10
C24247, Barbian, The Netherlands, (NW)	CC+/17'	2/10
C24247, Barbian, The Netherlands, (NW)	CC+/20	1/10
C24247, Barbian, The Netherlands, (NW)	CT+/17'	1/10
C24247, Barbian, The Netherlands, (NW)	CC+/5'	1/10
C960, Tammisto, Finland, (N)	CC+/20	1/9
C960, Tammisto, Finland, (N)	CC+/15	4/9
C960, Tammisto, Finland, (N)	CC+/17'	1/9
C960, Tammisto, Finland, (N)	CC+/20	1/9
C960, Tammisto, Finland, (N)	CC+/24	1/9
C960, Tammisto, Finland, (N)	CC+/40	1/9

Note: 30 individuals were assessed. See footnote Table 3.28, (N), Northern Europe, Finland-Tammisto, (NW), North-Western Europe, Denmark- Morso, and the Netherlands- Barbian.

3.3.3 Two chloroplast lineages and rare nuclear types contributed to a prominence of rare pattern type combinations in Central and Western European white clovers.

Nine combined pattern types were found in this group. These were combinations of two chloroplast lineages (CC+ and CT+) with seven separate nuclear forms. The common types observed in Ovcak were CC+/15 and CC+/20. Four nuclear pattern types that rarely occurred across the complete dataset were observed in this group.

Table 3.34 Occurrences of combined pattern types listed per cultivar for central European white clover cultivar Ovcak and western mixed cultivar Blanca.

Cultivar accession Numbers	Pattern Type	Occurrences
C7555, Ovcak, (C)	CC+/15	3/6
C7555, Ovcak,(C)	CC+/19	1/6
C7555, Ovcak, (C)	CC+/39	1/6
C7555, Ovcak, (C)	CT+/9	1/6
C6360, Blanca, (W)	CC+/15'	1/6
C6360, Blanca, (W)	CC+/17'	1/6
C6360, Blanca, (W)	CC+/20	3/6
C6360, Blanca, (W)	CC+/43	1/6

Note: 6 individuals, (12 in total) were sampled from each of two populations, representing Central and Western Europe. See Table 3.28 for footnote.

3.3.5 Maternal lineage CC+ was observed with diverse nuclear types in white clover with Israeli parentage.

CC+ was observed exclusively in cultivars from Israel, with main nuclear types common to both cultivars as well as with rare nuclear types that were cultivar specific; Tamar CC+/23 in Haifa CC+/10. Mixed origin cultivar Aran, showed more common European pattern combinations.

Table 3.35 Occurrences of combined pattern types in cultivars with Israeli parentage.

Cultivar Numbers & Names	Pattern Type	Occurrences
C4143, Tamar	CC+/5	1/8
C4143, Tamar	CC+/15	6/8
C4143, Tamar	CC+/23	1/8
C5892, Haifa	CC+/5	1/4
C5892, Haifa	CC+/10	2/4
C5892, Haifa	CC+/15	1/4
Aran, (French & Israeli)	CT+/15	3/9
Aran, (French & Israeli)	CC+/15	3/9
Aran, (French & Israeli)	CC+/20	3/9

Note: 21 individuals were assessed from mixed cultivar Aran (French and Israeli parentage), and pure Israeli parentage cultivars, Haifa and Tamar. See Table 3.28 for footnote.

3.3.11 Combination pattern types in accessions from North Africa showed conservation of chloroplast form CC+ associated with variable nuclear types.

All combination pattern types contained the chloroplast type CC+ lineage. In Algerian accessions it occurred with main and less common nuclear forms. In C1761, from Morocco main nuclear type 15 and the rarer form 25 were observed. Each Algerian accession showed an accession specific combination, CC+/7 (C6452) or CC+/30 in (C1751).

Seven combinations were observed that contained a rare nuclear pattern type. These were pattern combinations that had CC+ with the nuclear types 7, 9, 12, and 30, and once with each of 11, 25, and 38.

Algerian individuals showed greater variation in nuclear forms and hence diversity in pattern combinations than were observed for Moroccan individuals, even though more Moroccan individuals were sampled.

Table 3.36 Occurrence and combined pattern types of North African (Algerian & Moroccan) accessions.

Accessions No. & Origins	Pattern Type	Occurrences
C1761, Morocco	CC+/15	3/4
C1761, Morocco	CC+/25	1/4
C6452, Algeria	CC+/5	1/10
C6452, Algeria	CC+/7	4/10
C6452, Algeria	CC+/9	1/10
C6452, Algeria	CC+/12	1/10
C6452, Algeria	CC+/15	1/10
C6452, Algeria	CC+/30	1/10
C6452, Algeria	CC+/5	1/10
C1751, Algeria	CC+/5	1/9
C1751, Algeria	CC+/11	1/9
C1751, Algeria	CC+/12	1/9
C1751, Algeria	CC+/30	5/9
C1751, Algeria	CC+/38	1/9

Note: 23 plants were assessed. See Table 3.28 for footnote.

3.3.12 Combination pattern types South African cultivar Dusi out-group (Zone 3),

Combination pattern types from the two plants sampled from the out-group cultivar Dusi showed one chloroplast lineage CC+ and two nuclear pattern types 5 and 20, (Table 3.36).

Table 3.37 Combined pattern types of two plants of South African cv. Dusi

Cultivar Number & origin	Pattern Type	Occurrences
C16511, Dusi, South Africa	CC+/5	1/2
C16511, Dusi, South Africa	CC+/20	1/2

Note: 2 plants were sampled. See Table 3.28 for footnote.

3.4 TCS Analyses of pattern types

Introduction

The software TCS was used to perform analyses, singularly on the nuclear ITS sequence data and on the combination of the chloroplast sequence data (IRB21NE and *trn L*). This gave an indication of the relationships between the individuals of the dataset based on paternal and maternal pictures. In addition, the three genetic regions were assembled into a combined alignment and were analysed to give a picture of the overall relationship between the different DNA sequence combinations and their geographic locations.

TCS

The TCS software uses a statistical parsimony algorithm to estimate the differences between aligned sequences in order to assess the relation between haplotypes and to build a network diagram that reflects these relationships. It works by initially calculating a parsimony limit; this is the estimated maximum number of differences among the haplotypes likely to have occurred through single substitutions, under a 95% confidence criterion. Then the number of differences between the sequences is estimated and the haplotypes are sequentially connected together, starting with those that have one difference then those with 2 differences and so on. The process continues until the parsimony limit is reached or all the haplotypes have been connected within the network, (Posada and Crandall, 2001).

3.4.1 TCS analysis of combined chloroplast (*trn L* and IRB21NE) DNA patterns

3.4.1.1 TCS analysis of joined chloroplast DNA sequences - gaps in the alignment treated as fifth-state variants (Figure 3.33)

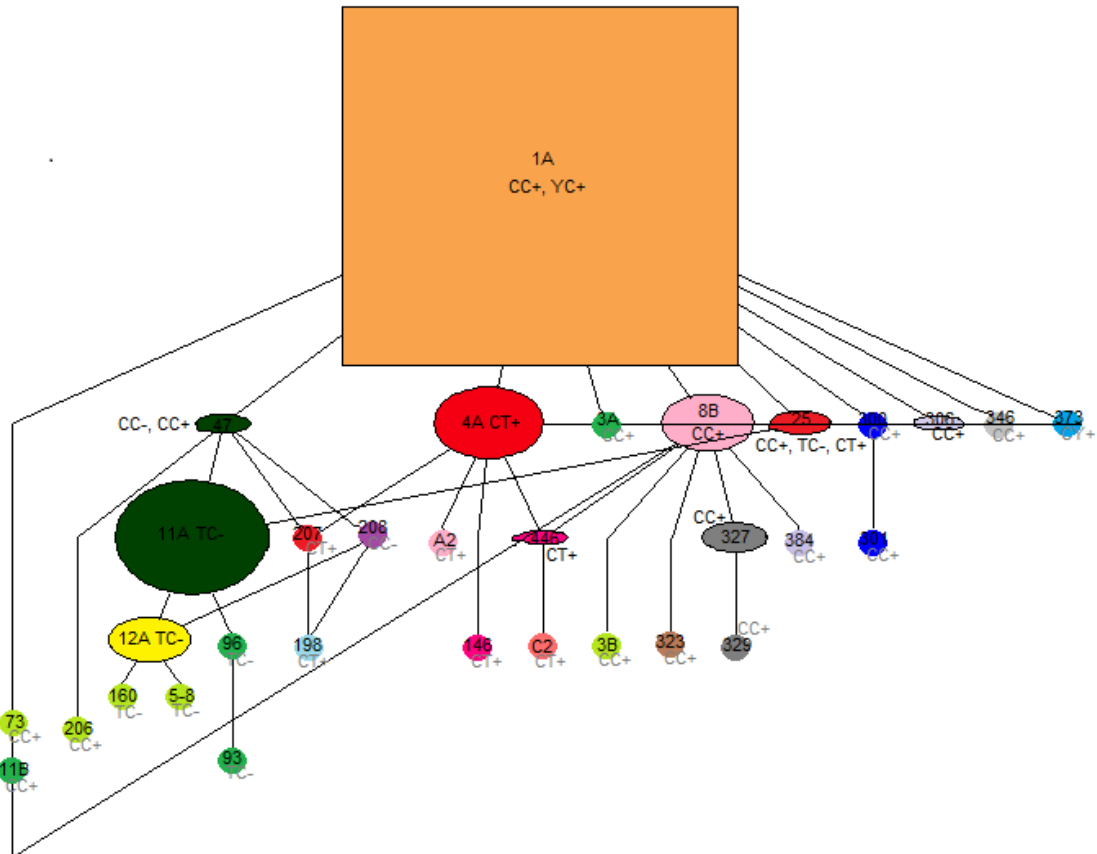


Figure 3.33 TCS analysis of joined chloroplast DNA sequences from 272 individuals including commercial and non commercial plants from across the sampling area. Gaps in the alignment were treated as 5th state differences. Origins of groupings (square/ovals): pale orange (131) - all populations excluding Crusader, Crau, Morso, baby pink- Portugal, Kent, Tamar, Morso, Barbian, Blanca, Ovcak, Regal, Crusader, Huia reference, vermilion (24)- France, Kent, Spain, Portugal, Barbian, Crau, Aran, dark green- Portugal and Spain, red (6)- France, Spain, Portugal, dark grey (8)- Tamar and Morso, Yellow- (14) Portugal, Spain and Kent, lime green (14)- Spain, raspberry pink (2)- Crau, pale purple (2)- Tammisto, Blanca. Individuals: lime green- Spain- (3B (S), 160 (NW), 5-8 (NW), 206 (W), 73 (W)), medium green- Portugal- (3A (C), 11B (C), 93/96 (NW)), pale blue- Kent. medium blue- Tammisto (Finland), red- France, purple- Lodi- (Italy), lavender- Regal (USA/Italy), brown- (Algeria), peach- Crusader (France Syria), raspberry pink- Crau (France), baby pink- Aran (France/Israel), dark grey- Tamar (Israel). Note: the numbers in the brackets refer to the number of individuals sampled. These individuals present in these groupings can be found in Appendix F.

The TCS analysis of the combined chloroplast regions showed geographic separation for the chloroplast pattern types CC+, CT+ and TC-, with a strong separation of TC- from CC+ and CT+ pattern types. Pattern type CC+ was the most commonly observed and geographically spread, with a placement as the possible pattern type source for types CT+, TC-, and CC-.

The CT+ pattern combination was largely confined to European populations of France, Kent, commercial Northern, Central European cultivars, (groups 4A, 446, 25), and mixed cultivars Aran, Crusader. Individual plants with this pattern type showed closer proximity to those with the CC+ type.

The analysis showed two clusters of TC- individuals 11A (Portuguese and Spanish) and group 12A; which was largely Spanish, but also included were two individuals, one from Portugal and one from Kent. Four Spanish and Portuguese individuals were shown to separate from these two groups. TC- was observed only once outside of this cluster of groups and individuals. Other Portuguese and Spanish individuals with pattern types CT+ and CC+ were observed throughout the diagram space.

Included but largely ignored were two positions of variation in the IRB21NE sequenced region; inserts of a T at position 458 and an A at position 464 in the Grasslands Huia reference sequence. This analysis has placed 14 individuals that have the T insert and the TC- pattern type in group 12A. These were from several populations Spanish: (C9025, C9034, C9037), Portuguese: (C8976), Kent: C6374, Italy: (C5897). The A insert occurred across a large portion of the data set, and was possibly an artefact of sequencing.

Twenty-two individual plants (circles) fell outside of the main groupings. Those with the type CT+ show a close relationship to group 4A, and are mostly of French origin. Individuals of chloroplast type CC+ showed a geographic distribution reflective of individuals within the clusters 1A and 8B, to which they are proximate. These twenty-two individuals had minor variants in DNA sequence that, by using the gaps as 5th state setting for the analysis were separated from the general pattern type groups.

Two individuals 73 and (11B) of Spanish and Portuguese origin that have the additional insert in the *trn L* intron sequence were connected by multiple changes to the central group 1A.

3.4.1.2 TCS analysis of joined chloroplast DNA sequences - gaps in the alignment treated as missing data

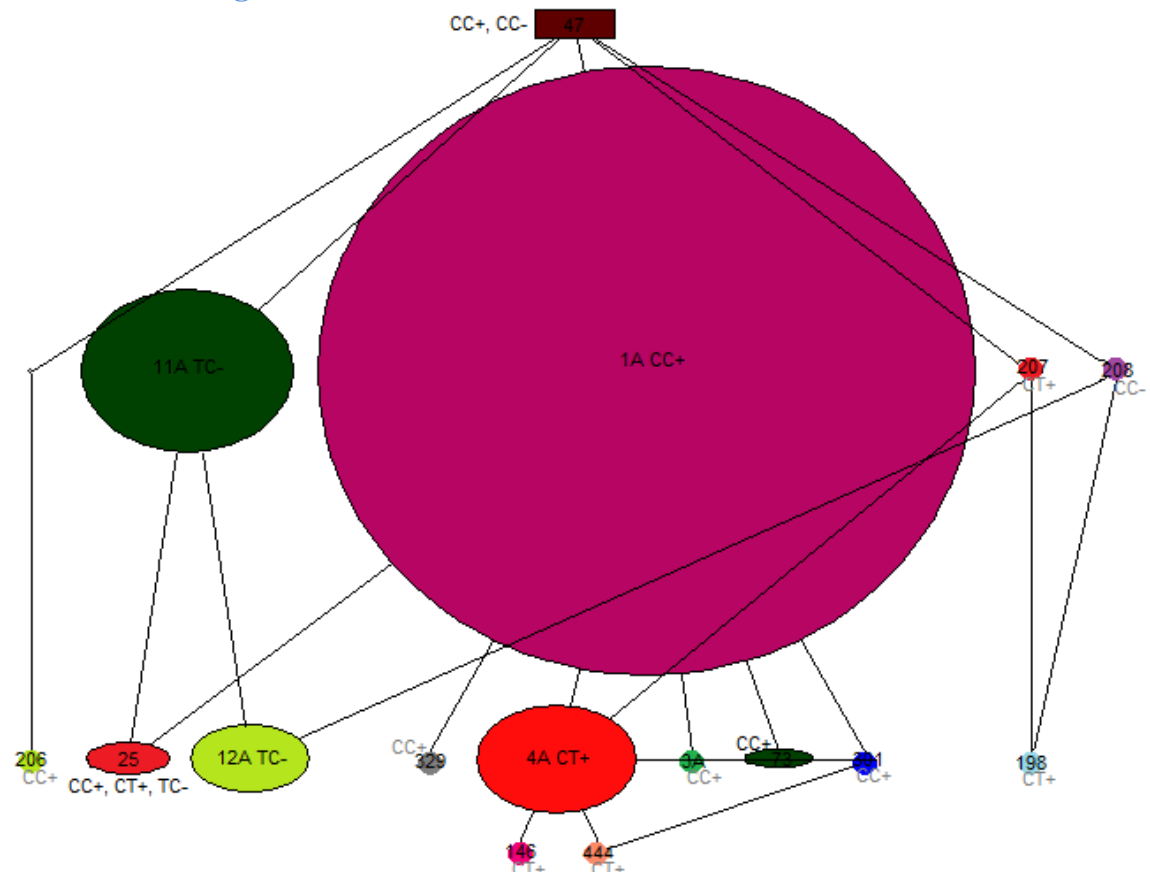


Figure 3.34 TCS analysis of joined chloroplast DNA sequences from 272 individuals including commercial and non commercial plants from across the sampling area. Gaps in the alignment were treated as missing data. Origins of groupings (rectangle/ovals): 1A: mulberry (131)- individuals from every sampled population, 4A: vermilion (27)- France, Kent, Spain, Portugal, Barbican, Crau, Aran, 11A: dark green (42)- Portugal and Spain, 12A: lime green- Spain, 25: red (6)- France, Spain, Portugal, 47: dark brown (5)- Portugal, Spain, Algeria, dark green (73). Individuals: lime green- Spain – (206, W), medium green- Portugal- (3A, C), pale blue- Kent, medium blue- Tammisto (Finland), raspberry pink- Crau (France), peach- Crusader (France/Israel), red- France, purple- Lodi (Italy), brown- (Algeria), dark grey- Tamar (Israel). Note: the numbers in the brackets refer to the number of individuals sampled. These individuals present in these groupings can be found in Appendix F.

This analysis showed that, by treating gaps in the alignment as missing data, six groupings were resolved, of which group 47 containing pattern types CC+ and CC-, and was placed prominently at the head of the diagram. This grouping is closely related to 1A but is also related to group 11A, (TC-).

The CC+ pattern type was the most prominent and present in individuals from most sampled populations, although it was heavily represented in cultivated and commercial populations. In addition the two individuals in the data set with the insert in the *trn L* intron sequence grouped together as group 73, and showed connection to groups 4A (CT+) and 1A (CC+).

Two separate groupings of TC- individuals (12A Portuguese/Spanish and 11A Spanish) were clearly separated from the central group. Group 12A included 16 individuals with the T insert in IRB21NE.

Groups containing the CT+ pattern type showed close proximity to CC+ group 1A. Connection between French, Spanish and Portuguese populations was observed in group 25.

Nine separate individuals were observed; those of pattern type CT+ largely had French origins, being from Crusader and Crau. While CC+ showed Northern Europe and Portuguese/Spain origins. An individual of Lodi type CC- was noted from Italy.

3.4.2 TCS analysis of ITS DNA pattern types

3.4.2.1 Commercial and non-commercial populations

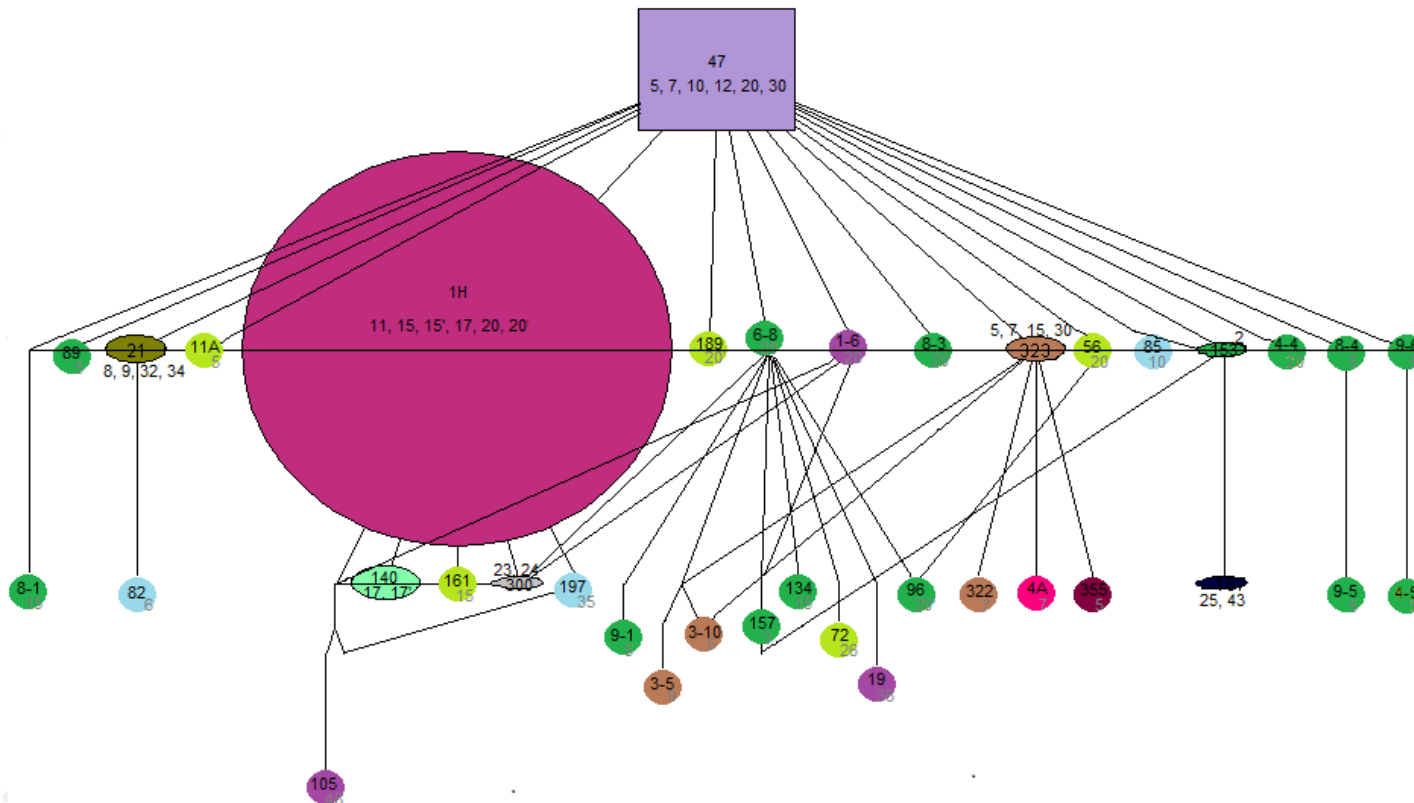


Figure 3.35 TCS analysis of ITS patterns from 265 individuals including commercial and non commercial plants from across the sampling area. The analysis was run with the setting gaps as missing data. Origins of groupings: mulberry- individuals from every sampled population, lavender- Spain, Portugal, Kent, Europe (not France or Aran), Algeria, S. Africa, Crusader, Israel, Italy, (Lodi/Regal), dark green- Portugal and Spain, Mouldy green- Portugal, Spain, Regal, Morso, Kent, grey/blue- Portugal, Tammisto, Barbian, Morso and Blanca, mint green- Portugal, Europe, Blanca, navy blue- Morocco and Blanca, grey- Tammisto, Tamar, brown- Algeria. Individuals: lime green- Spain – (11A (N), 56 (NW), 72 (NW), 161 (NW), 189 (N)), medium green- Portugal- (8-1, 8-3, 8-4 (NW), 6-8 (NW), 96 (NW), 4-4, 4-5 (NE), 134 (C), 157 (SW), 9-1, 9-5, 9-6 (NE)), pale blue- Kent, raspberry pink- Crau (France), purple- Lodi (Italy), brown- (Algeria), maroon- Blanca (Western Europe). Note: the numbers in the brackets refer to the number of individuals sampled.

This analysis of 265 individuals (Fig. 3.35) showed two main groups (47 and 1H) that indicate a clear split across the dataset for the two main ITS pattern types 5 and 15. Individuals of group 47 (46 individuals) most closely resemble type 5 while those individuals of group 1H (156 individuals) had pattern types most like pattern type 15. The majority of individuals from cultivated and commercial populations were found in group 1H. Individuals from French accessions and cultivar Aran were excluded from group 47.

Six small groupings were observed, five of which showed geographic placement of pattern types related to types 5 and 15. Groups 300 and 167 showed the separation of four less frequent pattern types. Group 323 shows the connection of a group Algerian of individuals with rarer pattern types 7 and 30 to group 47. Individuals of group 21 showed the presence of variable pattern types, present in populations from Spain, Portugal, Kent, Morso, Regal and S.184. Grouping 153 represents the presence of rare pattern type 2 in three Portuguese accessions, C8976, C8979, and C20477. The sixth group 140, indicates the separation of a group of 11 individuals, 10 from Northern and Western Europe and one from C8985 Portugal, that have ITS types 17 or 17'.

This analysis separated out 28 individuals, of which the majority were of Portuguese or Spanish origin. The remainder were from; Algeria, Italian Lodi, landrace Kent, Blanca, and France.

3.4.2.2 Analysis of non commercial populations, Figure 3.36

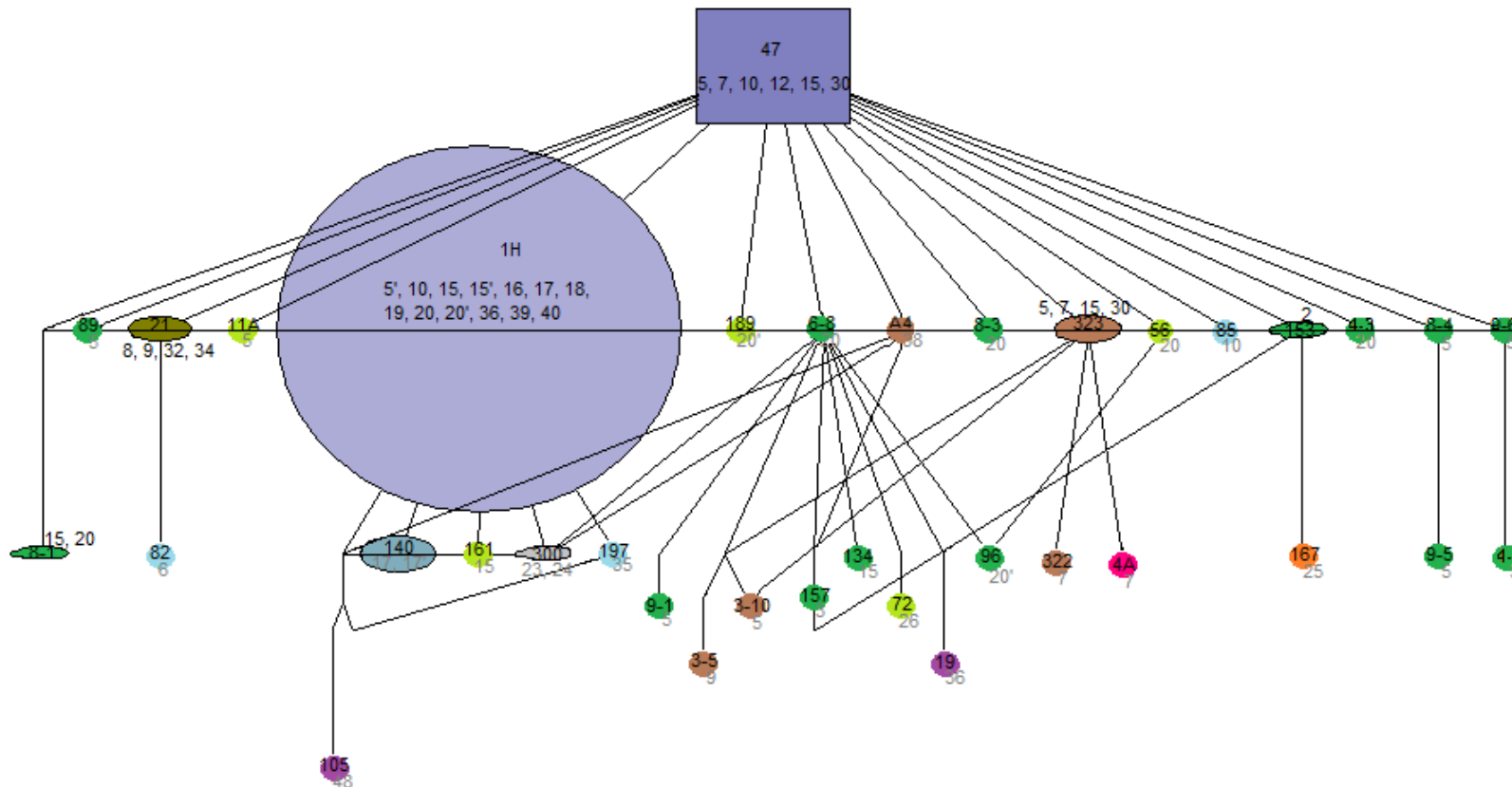


Figure 3.36 TCS analysis of ITS patterns of 224 individuals from non commercial populations. The analysis was run with the setting gaps as missing data. Origins of groupings: dark lavender-Portugal, Spain, Kent, Italy-Lodi, Morso, Barbian, Haifa, Algeria, lavender- all sampled populations except Morocco, French Aran, Mouldy green- Portugal, Spain, Regal, Morso, Kent, grey/blue- Portugal, Tammisto, Barbian, Morso and Blanca, grey- Tammisto, Tamar, brown- Algeria. Individuals: lime green- Spain- (11A (N), 56 (NW), 72 (NW), 161 (NW), 189 (N)) medium green- Portugal- (89 (NW), 96 (NW), 134 (C), 157 (SW), 6-8 (NW), 4-3, 4-5 (NE), 8-3,8-4 (NW), 9-1, 9-5, 9-6 (NE), pale blue- Kent, raspberry pink- Crau (France), purple- Lodi (Italy), brown- (Algeria), orange- (Morocco). Note: the numbers in the brackets refer to the number of individuals sampled.

An analysis excluding commercial populations was performed for a data set of 224 individuals. Gaps in the alignment file were treated as missing data. This analysis showed that the general shape of the diagram was little changed but that the 5, 15 pattern type separations were less obvious. The same two main groupings (47, 1H) were observed but both groups contained individuals of 5, 15 and related pattern types. Four of the six small groupings (21, 153, 300, and 323) were conserved in this analysis. A fifth Portuguese grouping 8-1 was observed. Retention of group 140 showed that the absence of individuals of Blanca in the alignment did not affect the placement of the group or change the connections of the group.

Group 167 (Fig. 3.35) was reduced to individual 167 Morocco, ITS pattern type 25 through absence of Blanca, but retained the same placement.

On the whole the same 28 individuals that fell out of the full analysis (Fig. 3. 35) were observed here. Portuguese and Spanish individuals were most frequent. Fewer individuals were observed from Italian Lodi and Algeria.

3.4.3 TCS analysis of combined pattern types, with alignment order, ITS-CD-IRB21NE.

This analysis based on 200 individuals of both commercial and non commercial accessions from throughout the sampled areas was performed under the setting gaps as missing data. The data set contained the two previously stated points of variation in IRB21NE noted in section 3.4.1, a C/T presence in place of the presence or absence of the indentified insert in IRB21NE, and an A/T presence at the end of IRB21NE sequence (464 in the Huia reference).

The presentation of the data by the TCS software shows that the relationships within the data based on DNA patterns can be broken down into 13 groupings of two or more individuals. Clear separation of chloroplast and ITS pattern combinations is observable. Of these the group 1H is the largest and contains a collection of individuals with commonality in ITS type 15, and variants on this pattern type, caused by the presence of ambiguous positions and rare variants. The disproportionately large size of this group is attributable to the large number of commercial varieties present in the group and

indicates the effect of commerce on the spread of type 15, (Appendix F). Individuals in this group contain both chloroplast types CC+ and CT+ although prominence of CC+ is observed suggesting a high contribution of this lineage to commercial varieties, (Appendix G). Two groups (40 and 81), show a connection of ITS type 5 with CC+ and related variant ITS types in numerous individuals.

A significant separation of Portuguese and Spanish individuals was observed as separation of group 24, and the number of individual green circles. Two further groups 65 and 51, identified connections between Portugal, Spain and the landrace Kent. Further Portuguese and Spanish individuals were observed as part of the smaller groups 40 and 140 suggestive of the diversity observable in these countries, and their potential as source populations for wide spread diversity. This was further supported by their presence in group 81 which contained a mix of Kent, Spanish, Portuguese, Barbican, Morso, Regal, Lodi, Dusi, Crusader and Algerian individuals.

A small grouping of individuals from both assessed Algerian accessions showed a placement separate from the majority of the data set, through multiple changes, indicating a potentially separate origin or a different evolutionary trajectory for Algerian germplasm. The closeness of this Algerian group to Portuguese individuals also supports the prominence of one dominant ITS sequence type, type 5, in both countries (Figs: 3.13, 3.14, 3.29, 3.30. 4.4 and Tables: 3.14, 3.23). The similarity in prevalence of type 15 as observed in the combined and ITS results (Figs: 3.15, 3.16, 3.29, 3.30, 4.4 and Tables: 3.15, 3.23, 3.29, 3.36) for Morocco and Spain for was not obvious in this analysis.

In this analysis a large number of individuals were observed with DNA pattern types separate from the main groupings. The minor groupings 167, 300, 342, and 309 show chloroplast pattern types with CC+ or CT+ in connection with rarer ITS patterns types. These ambiguities could not be resolved through resequencing and are treated by TCS as missing data hence; these groupings may not be correct.

However, what was most noticeable about the individuals in this analysis was the spatial arrangement of; Spanish individuals with their close proximity to Portuguese individuals that were more closely placed with European individuals. Portuguese

individuals are wider spread throughout the diagram, and showed close placement to individuals and groupings of European and/or Algerian individuals. Derived positions for individuals from France, (207) C7690, Bouches du Rhône, Entressen, in the Cote d'Azur South East France, Crau (4A) South France, and Crusader (444), support a suggestive role of France in the region of origin of white clover, (Figs: 3.17, 3.18, 4.44 and Table 3.1.

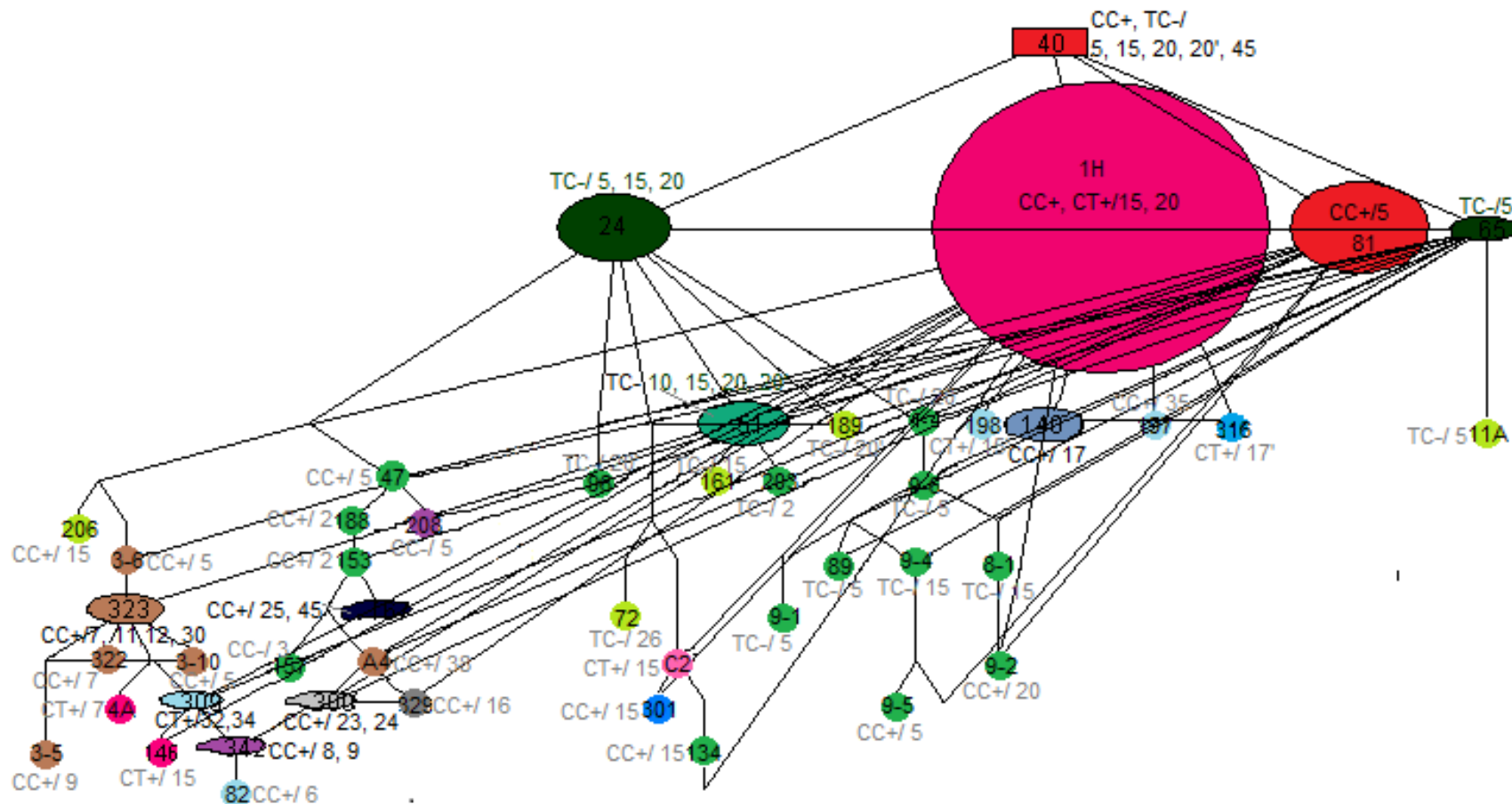


Figure 3.37 Combined pattern types TCS analysis with C/T and A/T, with the alignment ordered ITS-*trn L*- IRB21NE and gaps treated as missing data. Groupings: raspberry pink- Huia Reference, Spain, Portugal, France, Crau, Aran, Crusader, Italy, Tammisto, Tamar, Barbian, Morso, Kent, Regal, Blanca, Hafia, S. Africa- Dusi, and Morocco, dark green- Portugal and Spain, turquoise- Portugal, Spain and Kent, grey/blue- Portugal, Tammisto, Barbian, Morso and Blanca, purple- Lodi and Regal, maroon- Portugal and Italy, navy blue- Morocco and Blanca, orange-red- Portugal, Spain, Kent, Barbian, Morso, Regal, Italy, Algeria, and Dusi, brown- Algeria C6452 and C1751. Individuals: lime green- Spain- (11A (N), 72 (NW), 161 (NW), 189 (N), 206 (W)), medium green- Portugal- (47 (SW), 89 (NW), 134 (C), 153/157 (SW), 188 (SW), 4-4 (NE), 8-1 (NW), 9-1, 9-2, 9-4, 9-5 (NE)), medium blue- Tammisto, pale blue- Kent, raspberry pink- Crau (France), medium pink- Crusader (France/Syria), dark purple- Lodi (Italy), dark grey- Tamar (Israel), brown- (Algeria). Note: the numbers in the brackets refer to the number of individuals sampled.

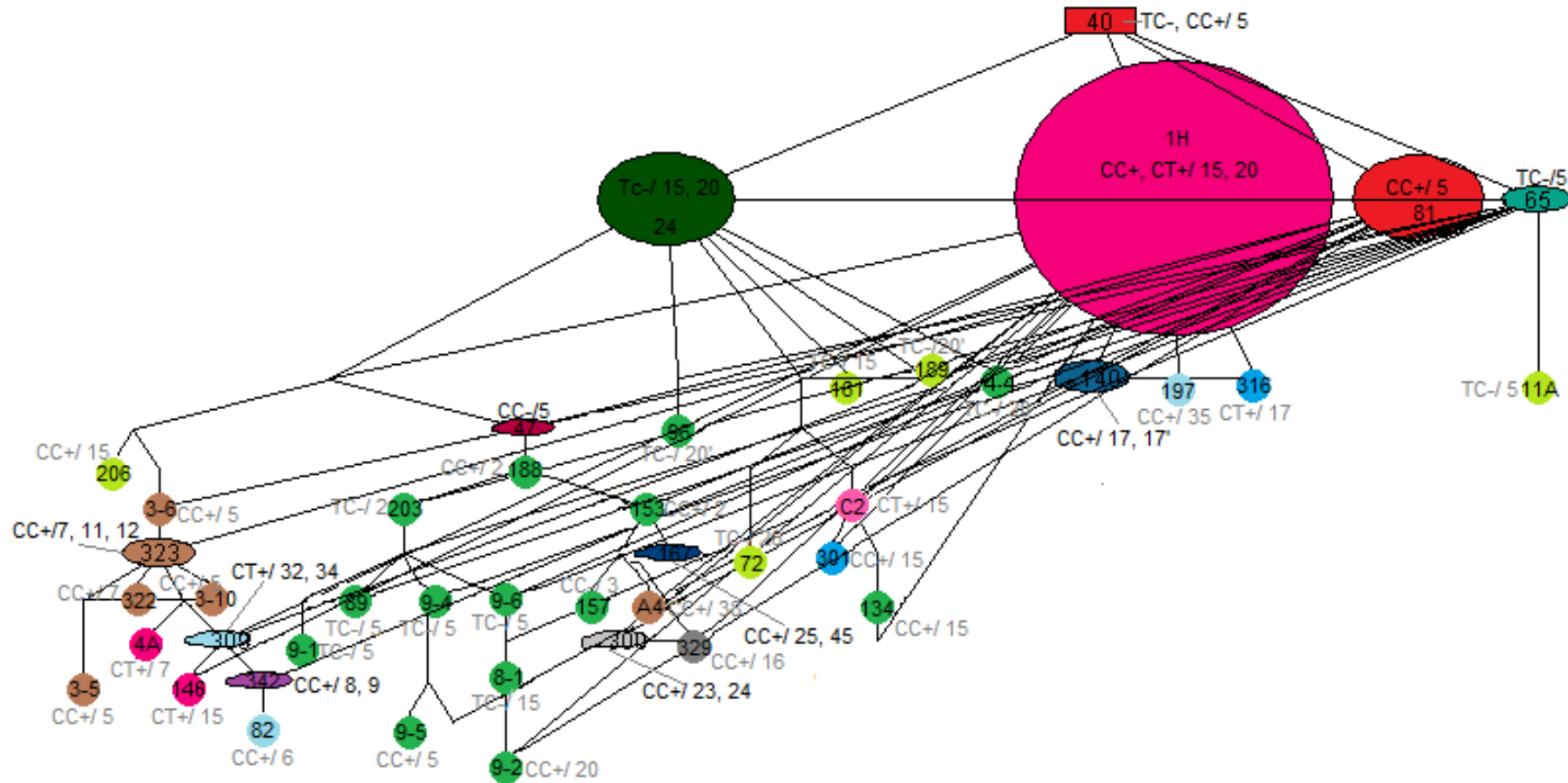


Figure 3.38 Combined pattern types TCS analysis with C/T, but with A/T treated as non-existent, with the alignment ordered ITS-*trn L*- IRB21NE and gaps treated as missing data. Groupings: raspberry pink- Huia Reference, Spain, Portugal, France, Crau, Aran, Crusader, Italy, Tammisto, Tamar, Barbian, Morso, Kent, Regal, Blanca, Hafia, S. Africa- Dusi, and Morocco, dark green- Portugal/ Spain, turquoise- Portugal, Spain and Kent, grey/blue- Portugal, Tammisto, Barbian, Morso and Blanca, purple- Lodi and Regal, maroon- Portugal and Italy, navy blue- Morocco and Blanca, orange-red- Portugal, Spain, Kent, Barbian, Morso, Regal, Italy, Algeria, and Dusi, brown- Algeria 6452 and 1751. Individuals: lime green- Spain- (11A (N), 72 (NW), 161 (NW), 189 (N), 206 (W)), medium green- Portugal- (89 (NW), 96 (NW), 134 (C), 153/157 (SW), 188 (SW), 203 (S), 4-4 (NE), 8-1 (NW), 9-1, 9-2, 9-4, 9-5, 9-6 (NE)), medium blue- Tammisto, pale blue- Kent, raspberry pink- Crau (France), medium pink- Crusader (France/Syria), dark purple- Lodi (Italy), dark grey- Tamar (Israel), brown- (Algeria). Note: the numbers in the brackets refer to the number of individuals sampled.

Figure 3.38 analysis of combined pattern types with absence of the T insert in IRB21NE

The treatment of the T insert in IRB21NE as irrelevant did not change the general space of the diagram. Clear separation was observed for CC+/CT+ relative to TC- groupings, this separation was evidently more heavily based on ITS pattern types than chloroplast, with separation falling into the same four groups ((40/81- CC+/5), (1H- CC+, CT+/15, 20), TC-/15, 20), (65 TC-/ 5)) as in the prior analysis, (Fig. 3.37). At the head of the diagram group 47 showed connections to the three other prominent groups. Placement of group 140 with ITS type 17 was conserved. The analysis showed better geographic separation within the dataset for groups of Portuguese and Algerian individuals. Connections to the group 24, (TC-) were more observable for the cluster of Portuguese and Spanish individuals, while clear division within the Portuguese and Spanish populations was prominent.

This analysis showed a new group, 47, which shows separation of 2 individuals with the rare pattern type combination CC-/5, from Portugal and Italian Lodi that are treated as individuals in the prior analysis. This second analysis also showed fewer individuals with variable pattern types, 31 relative to 35.

3.4.4 TCS analysis of combined pattern types, with alignment order, CD-IRB21NE-ITS.

To check the validity of the relationships presented in the prior analysis and to test if the order of genetic regions affected the shape of the diagram, the combined alignment was ordered as *trn L* intron- IRB21NE- ITS and run under the setting gaps as missing data. This produced a diagram identical to prior analyses, lending weight to the relationships presented, (Appendix E, Fig. E.1)

3.5 Combined Results and Morphological characteristics

Overview

Variation was observed in the appearance of white V marks, leaflet shapes, and leaflet sizes with DNA combination pattern types (Appendices D). **No association was found of specific distinctive morphological characters with DNA pattern combination types. DNA combination pattern types were independent of morphological appearance (Appendix D).**

Populations for which no complete morphological data was scored have been omitted from this comparison. In total, 130 plants were compared (Appendix D). A few general trends were noted;

3.5.1 Leaflet Markings

The study showed that leaf mark variation occurred within and between cultivars, mixed cultivars, ecotypes, and ecotype selections. A few individuals with the recessive green marking were observed, but again showed no typical association with other morphological traits or DNA character. Three analyses were performed for the comparison of each of the traits against the DNA pattern types, firstly comparison against the combined pattern types was assessed, and comparisons were then made against chloroplast and nuclear pattern types separately.

3.5.1.1 Chloroplast pattern types against V mark variation

The total observed chi squared value 5.18 giving no significant associations between V markings on leaflets and chloroplast pattern type classes, (Table 3.38).

Table 3.38 Chloroplast pattern types against presence of V mark character

	Chloroplast pattern type	Leaflet character			Totals
		Presence	Partial	Green	
Observed	CC+	50	7	14	71
Expected	CC+	53.99	5.92	11.09	71
χ^2	CC+	0.3	0.2	0.76	1.26
Observed	CT+	7	0	0	7
Expected	CT+	5.32	0.58	1.09	6.99
χ^2	CT+	0.52	0.58	1.09	2.19
Observed	TC-	16	1	1	18
Expected	TC-	13.69	1.5	2.81	18
χ^2	TC-	0.39	0.17	1.17	1.73
Observed Totals		73	8	15	N= 96

For this comparison 4 degrees of freedom were used and the analysis was performed for (N= 96) 96 individuals. Categories: Presence-Thick, Thin, normal; Partial- squat, incomplete; Green-absence of V mark- recessive genotype. Characters bright and faint were excluded from this analysis due to their reflection of intensity of pigmentation rather than presence or absence of marking. The total χ^2 value was 5.18.

3.5.1.2 Nuclear pattern types against variation in presence of V mark character

A total chi squared value of 9.8 was observed, no significant associations were obtained for comparison of nuclear ITS pattern types to morphological variation in V marking of leaflets (Table 3.39).

Table 3.39 Nuclear pattern types against V mark character

	Nuclear ITS type	Leaflet character			Totals
		Presence	Partial	Green	
Observed	5, 5'	15	0	3	18
Expected	5, 5'	13.46	1.46	3.09	18.01
χ^2	5, 5'	0.18	1.46	2.68x10 ⁻	1.64
Observed	15, 15'	31	2	4	37
Expected	15, 15'	27.66	2.99	6.35	37
χ^2	15, 15'	0.40	0.33	0.87	1.6
Observed	17, 17, 18	7	1	3	11
Expected	17, 17, 18	8.22	0.89	1.89	11
χ^2	17, 17, 18	0.18	0.01	0.65	0.84
Observed	20, 20'	11	1	3	15
Expected	20, 20'	11.21	1.21	2.58	15
χ^2	20, 20'	4.01x10 ⁻	0.04	0.07	0.11
Observed	Others	10	4	4	18
Expected	Others	13.46	1.46	3.09	18.01
χ^2	Others	0.89	4.45	0.27	5.61
Observed Totals		74	8	17	N= 99

For this analysis 8 degrees of freedom were used for assessment of (N= 99), 99 individuals. Categories: Presence-Thick, Thin, normal; Partial- squat, incomplete; Green- absence of V mark- recessive genotype. Characters bright and faint were excluded from this analysis due to their reflection of intensity of pigmentation rather than presence or absence of marking. A total χ^2 value of 9.8 was gained.

3.5.1.3 Chloroplast and Nuclear pattern types against variation in pooled presence of V mark character, presence or absence.

The data from categories presence and partial presence from tables 3.38 and 3.39 was individually pooled and two analyses preformed; one per parental perspective, chloroplast table 3.40 and nuclear table 3.41. Neither analysis showed significance, for obtained χ^2 values.

Table 3.40 Chloroplast pattern types against presence or absence of V marking

	Chloroplast pattern type	Leaflet character Presence	Green	Totals
Observed	CC+	57	14	71
Expected	CC+	59.91	11.09	71
χ^2	CC+	0.14	0.76	0.9
Observed	CT+	7	0	7
Expected	CT+	5.91	1.09	7
χ^2	CT+	0.2	1.09	1.29
Observed	TC-	17	1	18
Expected	TC-	15.19	2.81	18
χ^2	TC-	0.22	1.17	1.39
	Totals	81	15	N= 96

For this analysis 2 degrees of freedom were used, for assessment of (N= 96) 96 individuals.

Categories: Presence-Thick, Thin, normal; squat, incomplete; Green- absence of V mark- double recessive genotype. Characters bright and faint were excluded from this analysis due to their reflection of intensity of pigmentation rather than presence or absence of marking.

The total χ^2 value observed was 3.58

Table 3.41 Nuclear pattern types against presence or absence of V marking

	ITS pattern type	Leaflet character Presence	Green	Totals
Observed	5, 5'	15	3	18
Expected	5, 5'	14.91	3.09	18
χ^2	5, 5'	5.43×10^{-4}	2.62×10^{-3}	3.16×10^{-3}
Observed	15, 15'	33	4	37
Expected	15, 15'	30.65	6.35	37
χ^2	15, 15'	0.18	0.87	1.05
Observed	17, 17, 18	8	3	11
Expected	17, 17, 18	9.11	1.89	11
χ^2	17, 17, 18	0.14	0.65	0.79
Observed	20, 20'	12	3	15
Expected	20, 20'	12.42	2.58	15
χ^2	20, 20'	0.01	0.07	0.08
Observed	Others	14	4	18
Expected	Others	14.91	3.09	18
χ^2	Others	0.06	0.27	0.33
Observed Totals		82	17	N= 99

This analysis used 4 degrees of freedom for assessment of (N=99) 99 individuals. Categories: Presence-Thick, Thin, normal, squat, incomplete; Green- absence of V mark- double recessive genotype. Characters bright and faint were excluded from this analysis due to their reflection of intensity of pigmentation rather than presence or absence of marking.

The total χ^2 value obtained was 2.25.

3.5.2 Leaflet Shape

Variation in leaflet shape forms showed two main forms, ovate (ov) and orbicular (or). These had contrasting occurrences, but were often observed together within the same population, or geographic group, (Appendix D).

3.5.2.1 Geographic comparison against leaflet Shape

Given the focus of this study of potential origins within zone 1 for white clover, a chi squared analysis of leaflet shape for plants from Zone 1 relative to Zone 2 was conducted.

This analysis showed no significance in associations between leaflet shapes and geographic locations between zones 1 and 2 with a total χ^2 value of 2.99, (Table 3.42).

Table 3.42 Geographic comparison against leaflet shape

Grouping		Leaflet shape		Totals
		Ov/Or	Ro/Ob	
Portugal/Spain	Observed	31	0	31
Portugal/Spain	Expected	28.96	2.04	31
Portugal/Spain	χ^2	0.14	2.04	2.18
Others	Observed	75.5	7.5	83
Others	Expected	77.54	5.46	83
Others	χ^2	0.05	0.76	0.81
Totals		106.5	7.5	N=114

For of a 114 individuals a total n= 114 was used as a total of character observance. Degrees of freedom were = 1. A total χ^2 value of 2.99 was obtained. Leaflet shape groupings: Ov/Or: ovate (Ov)/orbicular (Or) and Ro/Ob: retuse obovate (Ro)/ Obcordate (Ob).
Others: Individuals sampled from Zone 2.

3.5.3 Leaflet Sizes

Leaflets sizes were observed to largely follow the recognised Mediterranean gradient of leaflet size, where in the Northern Mediterranean leaflets are typically in the tiny to small size range, (Davies et al., 1967; Caradus et al., 1990). Leaflets in the tiny-small size range are also found at high altitude throughout the sampling area. Larger leaflet sizes are observed in Italian Lodi and in southern parts of Europe and the Mediterranean, (Caradus et al., 1990). The data for DNA and morphology comparisons can be found in Appendix D. Additional results of χ^2 analyses are included in Appendix H Chi squared.

As the focus of this study was on the areas of Portugal and Spain as a region of origin chi squared analyses was performed for pooled plants from Portuguese and Spanish origin to test for associations in leaflet size relative to plants sampled from zones 2 and 3 (Table 3. 43).

Significance was found for an absence of larger leaflets in the Portuguese and Spanish group, $\chi^2 = 7.88$. The results suggest that in this study individuals within this class were underrepresented, as the χ^2 value suggests that between 5% and 10% of individuals from Portugal and Spain should have leaflets medium or bigger (Table 3.43).

Table 3.43 Leaflet sizes of Portuguese and Spanish individuals compared to those from Zones 2 and 3

Grouping		Leaflet sizes				Totals
		T, T-S	S	S-M	Others	
Portugal/Spain	Observed	8	13	10	0	31
Portugal/Spain	Expected	6.61	9.91	6.61	7.88	31.01
Portugal/Spain	χ^2	0.29	0.96	1.74	7.88	10.87
Others	Observed	18	26	16	31	91
Others	Expected	19.39	29.09	19.39	23.12	90.99
Others	χ^2	0.1	0.33	0.59	2.69	3.71
Observed totals		26	39	26	31	N= 122

A total of (N=122) 122 individuals were assessed with 3 degrees of freedom. Leaflet classes were grouped as; (T, T-S): tiny, tiny-small, relative to small (S), (S-M) small- Medium compared to larger leaflets (medium or greater) was performed compared to the rest of the dataset. A total χ^2 value of 14.58 was obtained. Others: Individuals sampled from Zones 2 and 3.

Chapter 4 Discussion

The focus of this study was assessment of the diversity within the hypothesised area of origin for white clover; Portugal and Spain, (Zone 1) relative to the area of natural distribution, Europe, the near east and North Africa, (Zone 2) and to the outer distribution, South Africa (Zone 3). The geographic distributions and occurrences of these sequences give an impression of the diversity and spread of white clover throughout the areas sampled. The maternal perspective was assessed from analysis of *trn L* intron (Taberlet, 1991) and the shortened IRB21NE sequence (Dr Nick Ellison, 2009). The paternal history was assessed by analysis of nuclear ITS sequences (Williams et al., 2001).

4.1. Observations of chloroplast sequences

4.1.1 Geographic distributions of chloroplast sequences

Five species lineages (pattern combinations) represented by five maternal DNA variants, (CC+, CC+*, CC-, CT+ and TC-) were observed. These correspond to C/T variation at positions 433 and 456 in the Grasslands Huia *trn L* intron and the presence or absence of the insert (CAAAA) at position 239 in the IRB21NE sequence of Grasslands Huia. The fifth novel maternal DNA variant CC+*, observed infrequently shows further variation with an penta-nucleotide insert at position 288 of the Grasslands Huia *trn L* intron sequence.

From Zone 1 130 individuals were sampled, 68 from Portugal and 62 from Spain. A larger number of individuals (174) were sampled from across Zone 2, . These were from France (43), Northern Europe (31), Central Europe (12), Italy (22), Britain (23), Israel (25) and North Africa (29). From Zone 3, the man-influenced distribution, 2 individuals from cultivar Dusi were assessed. In zone 2, 11 individuals were counted twice based on geographic location, giving a total of 185 individuals.

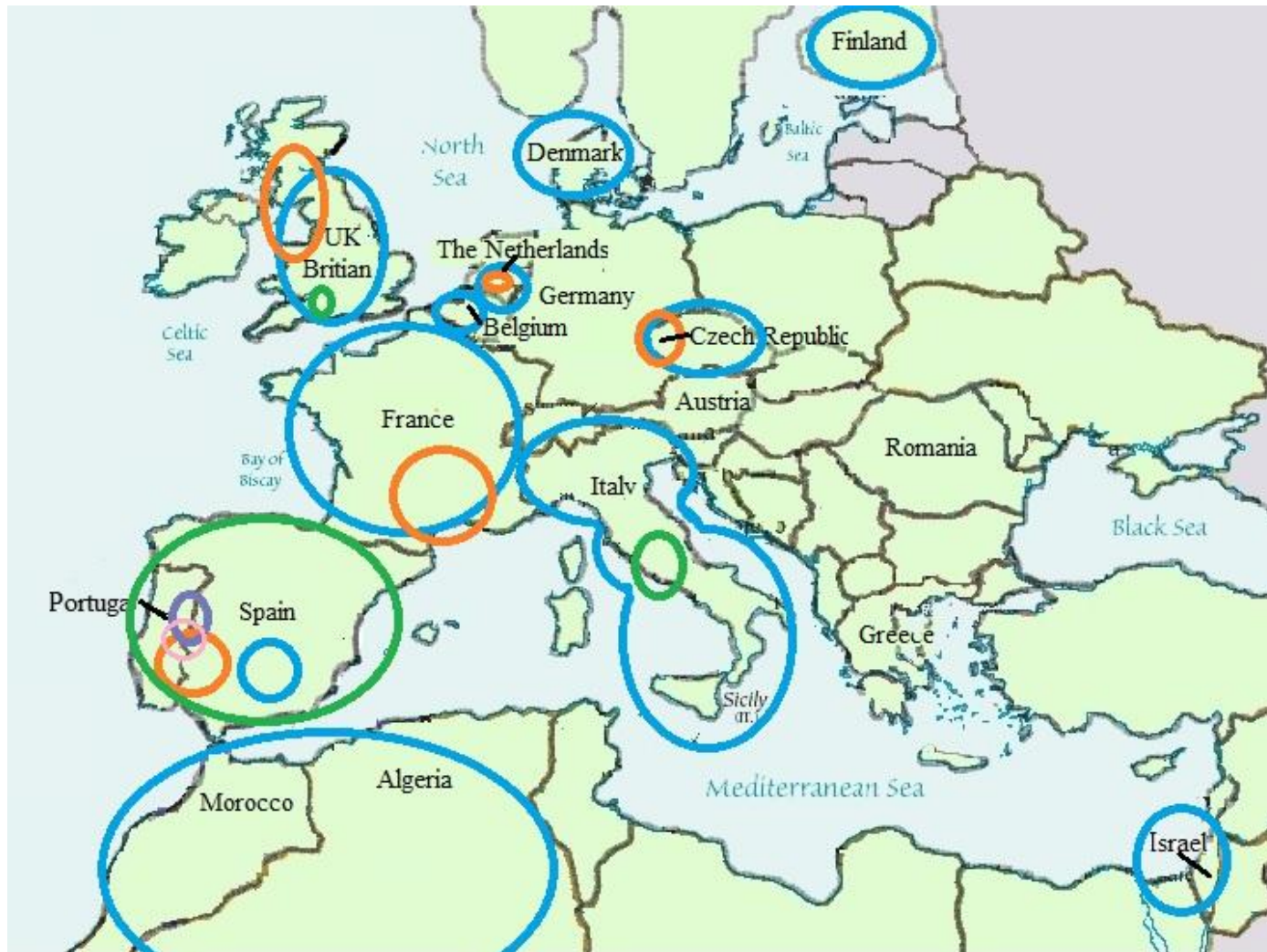


Figure 4.1 Europe and the Mediterranean region.

Countries shown in green represent countries of interest. Zone 1 includes Portugal and Spain. Zone 2, represents the natural distribution of white clover. Countries sampled from are labelled.

Circles show the approximate distributions of the different chloroplast pattern types.

Chloroplast pattern types:
 blue = CC+, red = CT+, green = TC-,
 purple = CC-,
 pink = novel CC+* *T. pallescens* variant

Diversity of white clover from Portugal and Spain, (Zone 1)

All four chloroplast types were observed within accessions from Portugal and Spain (Zone 1). This high level of diversity could be explained by species diversification in relative isolation and in the absence of commercial interest, suggesting that the natural diversity within the zone has been maintained and allowed to evolve freely with little competition from commercial germplasm. The restriction of pattern type TC- to Portugal and Spain (Figs. 3.2 and 3.3 and Tables 3.2 and 3.3) and the prominence of this lineage indicate an area of diversification and potential origin for white clover. The observed distribution of this pattern type suggests that it is wide spread throughout Spain particularly in Northern Spain (accessions C8994, Asturias and C9531 Ponferrada, León & Castile) and also has a significant distribution in Portugal. The virtual absence of TC- outside of Portugal and Spain supports the designation of Zone 1, as an area of high diversity and potential origin.

The restriction of CT+ to accessions from northern Portugal C9516, Vila Verde in Trás-os-Montes in the North West and in Spain C9027, C9034 from Galicia in the north-west, suggests a lack of dispersal of the lineage throughout Portugal and Spain.

CC+ was less frequent in the samples from Spain and Portugal, although proportionally the pattern combination was more common in Portugal than in Spain (Fig. 4.1, blue). Sampling suggested that the pattern type has about half the frequency throughout the populations in the North and Andalucía in the South of Spain. The pattern type was frequent in accession C7600 from the Spanish side of the Pyrenees Mountains. This fits with two hypotheses; that the lineage recently reached Spain from French or wider European origin, with limited dispersal or that the lineage is ancestral to the Zone 1 but has been overridden by recent pattern type TC- and/or by commercial influences and become relic populations in the North and South of Zone 1. Data from Spanish white clover populations suggests that there is enough variation to warrant a more in depth assessment.

Novel Portuguese and Spanish individuals

The observance of the previously undetected pattern type CC- in Portuguese and Spanish accessions further emphasises the greater diversity in Zone 1. This pattern type was found in individuals from geographically separated regions (South-West Portugal -

Grândola, Setúbal District, Águas de Moura, Pamela and Extremadura in Spain) which suggests that this maternal lineage may initially have had a wider distribution in the south western Iberian Peninsula and over time has become isolated in separate populations. However as the sample size was small it may be that this study failed to capture the true extent of the distribution of the lineage.

The fifth observed maternal lineage CC+* (Pink in Fig. 4.1) was observed solely in a central region of Portugal and Spain in the two adjacent areas Estremadura/Extremadura and the in the neighbouring southern Spanish region Andalucía. This insert at position 288 in the Grasslands Huia *trn L* intron sequence, denoted as * was also found in the two independent samples of *T. pallescens* from the Pyrenees, (Jaca, and AZ4837) included in this study (Table 3.5). This suggests that there may be a pocket of diverse maternal lineages within the Pyrenees as is observed in populations of *T. pallescens* within different glacial valleys in the Austrian Alps (Raffl et al., 2008) and is noted for the endemic Pyrenees *Borderea pyrenaica* (Gabriel et al., 2007). Further presence in Southern Spain does suggest a possibly wide distribution for the pattern type. These observations imply either that diversification has occurred since the original formation of white clover, with several lines of the maternal types existing in the populations, or speciation has occurred repeatedly within this region, producing *T. repens* individuals with novel *trn L* intron sequences.

However as the sample size was small, the results are likely give an inaccurate representation of the natural distribution of these pattern forms. Consequently further sampling of Portugal and Spain may find greater occurrence and wider distribution of these novel maternal types, as well as others. Increased sampling should also help resolve the connection between observed lineages and may even show a connection between the CC- and CC* types.

Natural distributions of white clover Zone 2, outside of the hypothetical area of origin

The wide distribution of the chloroplast type CC+ (Fig. 4.1, blue) suggests a common maternal lineage that extends throughout Europe, including (Britain, Europe, France, and Italy), North Africa and Israel. The pattern type TC- (Fig. 4.1, green) was virtually absent in populations from Zone 2.

However given the limited number of samples collected and the few places sampled from in Zone 2, in particular the numerous European countries not sampled from, a true picture of the diversity and hence of comparison of African and near eastern white clover and their relationship to European or other Mediterranean white clovers, cannot be determined. Further, sampling is required.

4.1.2 Comments on chloroplast sequence variation from selected ecotypes, ecotype selections and mixed cultivars

Israeli derived Aran, Haifa and Tamar

The more limited distribution of the pattern type CT+ (Fig. 4.1, red) across all of the sample area indicates a separately derived or a related maternal chloroplast type. The CT+ lineage was observed in the mixed cultivar Aran (France x Israel) but was absent from pure Israeli derived Haifa and Tamar; suggesting the CT+ type in Aran may be further evidence for prominence of the CT+ type in France. This supports the idea that multiple pattern types can be expected within a mixed cultivar, based on the mixed parentage of the founding parental populations. It proposes a narrower, more European based distribution for the chloroplast type CT+. However, the absence of sampling of the French and Israeli populations used in Aran, and better coverage of European populations inhibits further comment.

Expectations of ecotype Kent

It was expected that sampling of two populations of the Kent ecotype/landrace, (C6304 and C6374) known to be from specific fields in Kent, Suffolk, England would produce similar results for diversity in chloroplast pattern composition. However it appears as if the Kent ecotypes are more variable than might otherwise be expected, having CC+, CT+ and TC- sequence combinations. A thorough investigation the Kent landrace by sampling from different areas within the fields designated as Kent and would be needed to determine the extent of this diversity.

Regal (USA) and Lodi (Italy)

Regal is an ecotype selection derived from Italian Lodi, the inclusion of this of Regal allowed the comparison of DNA sequence between a natural population and a selection derived from it. The prominence of CC+ and absence of novel variation in both the

Italian Lodi and Regal suggested absence of diversification in the common maternal ancestry. The results agree with the hypothesis that the genetic variation present in Regal is a subset of the natural variation observed in Italian Lodi of the Po Valley, Northern Italy and that this ecotype has likely remained true to form.

Dusi, South Africa- Zone 3

Tentative assessment of the out group South African cultivar Dusi, (C16511) was made from the three individuals. These showed the wide spread common commercial CC+ pattern type, conducive with the cultivar being produced from introduced populations. Despite sampling size absence of CT+ and TC- suggests that the introduced populations from which Dusi was produced are from well recognized European and Mediterranean populations and commercial varieties. Further absence of TC- implies these populations were not from a shared origin with populations from the Iberian Peninsula.

4.1.3 Comparison of chloroplast diversity in Zone 1 to Zone 2

Given the wide geographic spread over which the sampled populations were from, comparisons between designated Zones 1 and 2 are tentative. The relative absence of maternal diversity in Zone 2 may be caused by factors, such as: random sampling, and the sample size, the sampling of this study was too sparse to capture the actual diversity present in Zone 2. Additionally absence of diversity could be explained by a natural lack of variation in maternal populations due to recent origins for white clover within Zone 2, for instance dispersal of only a limited range of white clover out of Zone 1, and/or influence of commerce such as a large sweep through in recent times of more genetically uniform commercial varieties, in conjunction with farming practice.

Chloroplast pattern types CC+, CT+ and TC- in Zone 1

The presence of diverse chloroplast sequence types in white clover from Portugal and Spain could be explained by movement of diverse ancestral populations of *T. pallescens* from separate glacial valleys into different areas throughout the region, (Gutiérrez et al., 2002; Raffl et al., 2008) (Figs. 4.2 and 4.3). Alternatively it may be that the three main types (CC+, CT+, TC-) arose independently in separate *T. pallescens* lineages (Model 1, Fig 4.1) and that by adaptation and natural selection, the maternal TC- lineage is prominent in Portugal and Spain. Or that the three lineages are derived from one another, (Model 2, Fig 4.2).

The virtual absence of TC- outside of Portugal and Spain supports an origin for the lineage within geographically isolated populations of zone 1. It raises the question of whether this pattern type is an example of another maternal ancestor within this region, or is this sequence the result of a recent mutation event?

As clover was first domesticated in Spain around 1000AD, (Kjaegaard, 2003) it can be assumed that Spanish clover populations are likely truer in form to old ancestral domesticated populations, and that relative isolation, helped in part by lack of international commercial interest, and domestication has played an important role in maintaining diversity within these populations. Perhaps little attention has been paid in international breeding programmes to Portuguese and Spanish white clover, meaning that deliberate spread of Portuguese and Spanish white clover and hence the pattern type TC- has been minimized outside of Portugal and Spain. Excluding uneven sampling between accessions, the simplest hypothesis is that the confined distribution of TC- implies it has moved little.

Further, the locations in which this pattern type was found suggest that there is a central region that loosely follows the border between Portugal and Spain and the mountain/hill ridges within these areas. The low frequency in Pyrenees accession C7600 could be taken as a suggestion of the outer limit of the distribution of TC-. If this is the case then local adaptation may be responsible for maintaining the boundaries of this pattern type.

France, a transition zone?

The large landmass and geographic location of France (between Spain and greater Europe), in conjunction with commerce, explain why France could be considered as a transition zone. Had the sampling been broader and the population sample sizes bigger this study may have shown a gradient like distribution of pattern types across the country. The results of this study lend support to this theory, with an observed prevalence of CT+ in Southern France (Fig. 4.1, red), but the observed distributions of CC+ and CT+ convincingly neither support nor negates this.

The absence of TC- pattern type is not unexpected given its absence from elsewhere in Europe. However, wider and more intense sampling throughout France and particularly in regions in South West France adjacent to the border with Spain is required.

Sampling should be conducted in the southern towns and surrounding areas of Carcassonne, Toulouse, and Pau within the regions of the Midi-Pyrénées and that of Aquitaine. From this sampling, greater variation in maternal lineages may be found. Finding pattern type TC- within the French foothills of the Pyrenees and adjacent localities would support the hypothesis of a transition zone. Further, based on geography and altitude the populations from Andorra in the eastern Pyrenees Mountains between South France and North- East Spain may contain a diverse collection of *T. pallescens* populations and hence potential maternal lineages not currently known.

4.1.4 Relationships between observed chloroplast pattern types

In this study three main maternal lineages were evident in *T. repens*, CC+, CT+ and TC- (Fig. 4.2). Pattern type CC+ was very common and had a wide distribution across all areas sampled. The next most widespread type, outside Portugal and Spain, was CT+. TC- was the most prominent pattern type in Zone 1.

Widespread of pattern combinations CC+ and CT+ across the complete data set

The two lineages CC+ (Fig. 4.1, blue) and CT+ (Fig. 4.1, red) generally occur together across all the areas sampled in Zone 2, with significant presence in Zone 1. The two pattern types differ by one SNP mutation, suggesting they could be closely related. The proportions of CC+ and CT+ varied widely between accessions. The purposeful planting of white clover in a progressive manner across Europe (Kjaegaard, 2003), is likely to have affected allele frequencies in wild populations of clover. In some regions, these populations could have been over thrown and superseded by better agronomic seed, such that the pattern types picked up in subsequent collections might not represent the original patterns.

The relative frequencies of CC+ and CT+ would have been affected by the amount of crossing between different populations and the amount of genetic variation and hence adaptation of each population. The introduced and native varieties may have come together to produce a more diverse population.

A better understanding of the effect of commercial farming practices in obscuring genetic diversity and mixing of the wild genetic diversity present in wild populations of white clover will be gained by assessing a wide range of commercial clover types

selected from different geographic localities. Extensive sampling may then reveal the rarer original sequence types.

Pattern combination CC+

CC+ was very common and had a wide spread distribution across Europe, the Mediterranean and the Near East (Fig. 4.1, blue). Thus it seems this type could represent an ancestral population that has become wide spread with time. Natural selection and farming practice are likely to have influenced its distribution. The dominance of the CC+ type would have likely been influenced by, individuals of the CC+ lineage with improved survival properties due to natural selection. If genetically improved individuals carrying the CC+ type were incorporated into commercial varieties this would act as a further promotion mechanism for spread of the CC+ pattern type, via commerce.

The greater proportion of the CT+ pattern type relative to CC+ in the Kent landrace, and in the ecotype selection Crau, is exceptional to patterns elsewhere, although this maybe due to limited sampling.

Pattern combination CT+

How did CT+ come to have such a wide distribution? Is CT+ a natural variant from the CC+ pattern type, which arose spontaneously as a diversification event in multiple places, many times? (Model 2, Fig 4.2). Or is its dispersal throughout most of Europe and North Africa due to co-dispersal with CC+ through recent farming practices. Further sampling particularly; of untouched fields and isolated populations throughout Zone 2 may find a wider distribution of pattern type CT+, than was shown by this study.

Discussion of the narrow distributions of chloroplast pattern types TC- and CC-, restricted to Portugal and Spain.

The rare occurrence of CC- could be evidence of recent lineage in *T. repens* populations (Fig. 4.1, purple) although the origin for this maternal lineage is unclear. The few individuals in which CC- was observed were found in populations that also had individuals with pattern type TC-. Hence it is possible that one sequence perhaps was derived from the other. The low frequency of CC- neither negates nor supports this connection. It could be that the sequence is a relic from ancestral populations that gave

rise to *T. repens*, but which was out competed by CC+, CT+ and/or TC- in the sampled populations (Fig. 4.1).

4.1.4 Chloroplast pattern type origins

One possibility is that the chloroplast lineages: CC+, CT+, and TC- may have arisen in three separate glacial events. Model 1, (Fig. 4.2) proposes that based on relative occurrence and distribution the pattern type CC+ is the oldest lineage. The wide distribution but low frequency (influenced by sampling limitations) of the CT+ lineage suggests this type arose next. The alternative suggestion, outlined in Model 2, (Fig. 4.3) suggests that the three less frequent pattern types CT+, TC- and CC- were derived from the most wide spread, common lineage CC+.

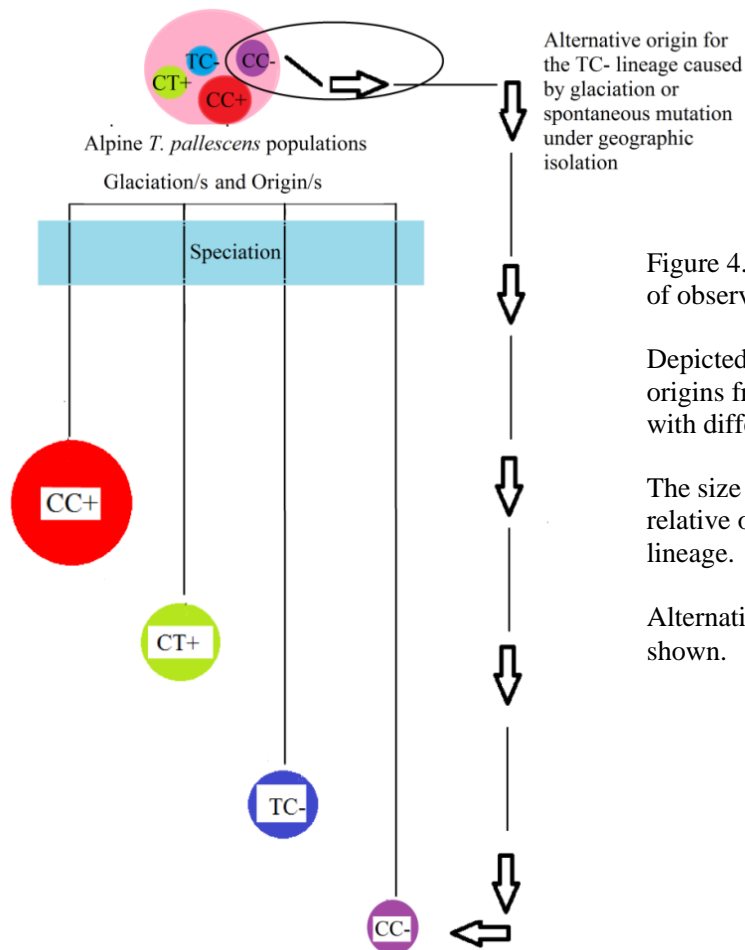


Figure 4.2 The suggested origins of observed chloroplast pattern lineages.

Depicted are possible hypothetical origins from different maternal ancestors with different chloroplast sequences.

The size of the circles depicts the relative observed proportions of each lineage.

Alternative routes for lineage TC- are shown.

In Model 1(Fig. 4.2) it is suggested that the three most commonly observed pattern types represent three maternal lineages that were introduced into current populations of *T. repens* through three separate glacial events. An alternative pathway is shown for the

maternal lineage TC-, where a transition of a C to T in the rare CC- maternal lineage could have produced the TC- lineage in current populations.

The hypothesis is that each pattern type represents a maternal lineage that arose during a glacial event that brought the parental species together at different times and/or in different places, two routes are suggested for the origin of the TC- lineage.

These are that TC- arose spontaneously during a recent glacial period in which the TC- type was inherited from a *T. pallescens* population that was forced down to the Portuguese and Spanish plains where intercrossing with *T. occidentale* led to speciation and appearance of the pattern type in current *T. repens* populations. Alternatively, the lineage may have arisen by mutation from the CC- form in *T. repens* populations. The occurrence of both types in Portugal and Spain is consistent with this. However, as discussed above, it is also possible that CC- is recently derived from the TC- lineage.

Relationships between chloroplast pattern types, Model 2, Figure 4.3

The alternative suggestion, outlined in Model 2, (Fig. 4.3) suggests that the three less frequent pattern types CT+, TC- and CC- were derived from the most wide spread, common lineage CC+. However the hypothetical transitional forms (crossed out) for conversion between the three main types were not found in this study. As the sampling of this study was insufficient to cover the recognised zones of diversity, (Fig. 4.1) wider sampling may yet find transitional forms and shed light on the geographic distributions of the indentified lineages and further geographic species diversity.

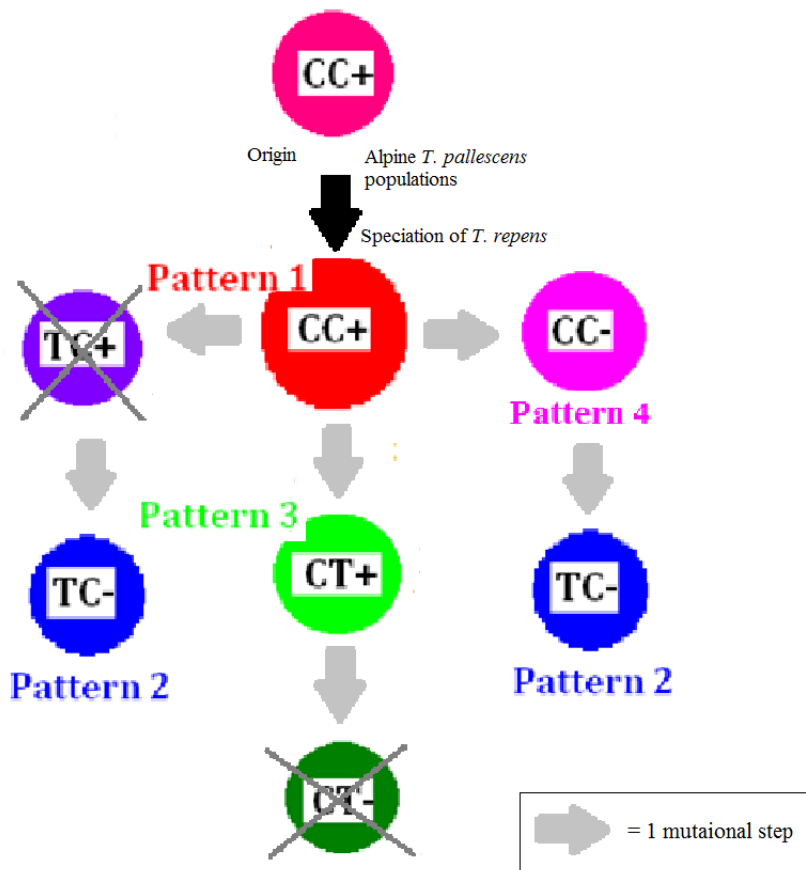


Figure 4.3 Relationships between chloroplast pattern types, a suggestion of the evolutionary directions of the main maternal lineages. The assumption of a maternal CC+ ancestor is made, with this lineage being the transferred to *T. repens* upon hybridization. Mutational steps required to change between lineages are shown. Non-observed intermediates are crossed out.

T. pallescens vs *T. repens*

It is known that *T. pallescens* is polymorphic and accessions differ at several chloroplast DNA positions. Polymorphism in *T. pallescens* was also reported in a recent study by Raffl et al., (2008) which showed AFLP diversity among Austrian alpine populations of *T. pallescens*. Raffl et al., (2008) stated that the amount of diversity in DNA sequence was sufficient to identify differences between populations growing in separate, even adjacent glacial valleys within the Alps. They argued that movement up and down the glaciated valleys, in conjunction with glacial and warmer interglacial periods has led to the observed diversity (Raffl et al., 2008).

The difference between CC+ and CC- is the presence or absence of a penta-nucleotide insert (CAAAA) in the IRB21NE sequence. The CC- variant could have arisen before speciation in the maternal ancestor or post the speciation of white clover. To assess this,

sampling and sequence analysis should be made of current *T. pallescens* populations. Sequence comparisons can then be made with *T. repens* populations.

Further assessment of the occurrence of the CC- pattern type in *T. repens* populations would help to gauge how wide spread the CC- lineage is. Sequencing error has been ruled out due to independent verification of this CC- pattern type in studied individuals (Dr Nick Ellison, pers. com, 2010).

4.2 Observations of Nuclear ITS Sequences

This study assessed the variation in nuclear ITS sequence differences with a focus on the genetic diversity Zone 1 (Portugal and Spain; 135 individuals) relative to Zone 2, the natural distribution of white clover (Europe, the Mediterranean, Near East and North Africa). Further comparison was made to Zone 3, the outer main influenced distribution. Fifteen individuals were included from Zone 3, from: NZ (Grasslands Huia), South Africa, Romania, Ethiopia, Japan, China, Kazakhstan, Costa Rica, and the USA. In total, 327 individuals were assessed (Results Section 3.2). This produced 45 different nuclear ITS types, with three common and one separate, strongly Northern European-based, ITS pattern; type 17 (Tables 3.12 and 3.14).

4.2.1 Geographic distributions of nuclear ITS sequences



4.2.1 Geographic distributions of nuclear ITS sequences

Figure 4.5 Approximate relative proportions from sampled populations with country locations of the main ITS pattern types.

Main types: 5- red, 5'- dark red; 15- medium blue, 15'- pale blue; 20- medium purple, 20'- pale purple; 17 medium green, 17'- grass green, 18/19- dark green.

Rarer types: 7- pink, 10- yellow, 30- peach, 35- blue, 36- navy and 39- blue.

Each circle represents one pattern type.

Infrequent pattern types are excluded, so that circles represent the approximate relative proportions of each main type.

Genetic diversity in main pattern types 5, 15 and 20 in Zones 1 and 2

Differences were observed for the distributions and frequency of pattern types 5 and 15 across Zone 1, the area of origin, and Zone 2, the natural distribution zone of white clover. The difference between these pattern types of a single SNP at position 509 of the ITS (Grasslands Huia) reference sequence shows their similarity (Section 3.2, Table 3.12).

The results of this study suggest there is higher prominence of; ITS 5 in Portugal and of ITS 15 in Spain (Figs. 3.13-3.16). This suggests: different modifying influences (e.g. glacial events), different founding populations, and possibly different contributing paternal lineages of *T. occidentale* in each country.

Interestingly a similar situation occurs in North Africa Where pattern type 5 predominates in Algeria and type 15 predominates in Morocco.

In contrast this distinction was not observed in Zone 2, (the rest of Europe) where, both pattern types were generally observed in the same areas, albeit at different allelic frequencies (Fig. 4.5). Type 5 was very rare in France.

The heterozygous ITS type 20 was prevalent throughout most of sampled area, but not in Denmark or Algeria. It is likely that very low allelic frequencies of parental type 5 in Denmark, and type 15 in Algeria prevented observance of expected ratios under Mendelian genetic principles (Griffiths et al., 2004).

The third lineage

The third most common ITS pattern type was type 17. This is observed primarily in populations sampled from Northern and North-Western Europe and has a distinguishing A at position 571 of the Grasslands Huia reference sequence (Tables 3.12 and 3.14).

The limited distribution and frequency of type 17 suggests two hypothesis of origin: Firstly that this pattern type represents a separate, third paternal *T. occidentale* lineage, (Fig. 4.5), and secondly the closeness of this lineage to type 15 suggests that the lineage 17 could have arisen through a transversion from the 15 lineage and subsequently become spread.

Furthermore the observation of a heterozygous individual type 18 with similarity to types 15 and 17 is in agreement with parental roles of types 15 and 17 for type 18. However, lack of detection of individuals of type 18 may be due to, limited sampling and the class distinction of 17'. It may in fact be that the observed frequency of 18 individuals in this study is higher and approaching Mendelian genetic ratios more readily than thought (Griffiths et al., 2004).

Diversity in rare pattern types

The frequency of rare nuclear pattern types differed across Zones 1 and 2. Diversity was greater in Zone 2, with 28 rare pattern types among 216 plants 13%, while in the Zone 1, 11 different pattern types were observed for 135 plants 8%. This higher variation in Zone 2 is likely due to the larger landmass covered by the natural distribution. It is probable that this study did not capture the full extent of the diversity in Zone 2 populations. This variation is likely due to more recent dispersal and selection under different environmental constraints allowing adaptation. Two potential pockets of diversity were observed in Zone 2: these fell within the regions of North Africa and Southern England, (Kent landrace) (Figs. 3.21, 3.22, Table 3.18). In North Africa, 12 out of 28 rare pattern types were observed, with eight unique to the region. The Kent landrace showed more diversity than expected for a standardized commercial line, with the 9 rarer ITS pattern types.

Of the rare pattern types observed, types 7, 9, 10, 30, 32, 34 and 39 showed multiple occurrences within and across populations (Fig. 4.5). This suggests that further sampling may show these rare types to have wider occurrences and account for a reasonable amount of natural diversity in white clover. Of particular interest are the heterozygote type 30, unique to Algeria, and type 10, which has prominence in the Israeli cultivar Haifa (Table 3.18, Figs. 3.21 and 3.22).

Man influenced distribution, Zone 3

The observations from the 15 out-group individuals showed the presence of the same three main pattern types: 5, 15, and 20. This suggests that Zone 3 shows the same main sequences as in the natural distribution zones of white clover and that none of the minor sequences were disproportionately represented. This supports the assumption that the white clover in Zone 3 was derived by commercial agricultural interests from the

common forms of white clover with origins in Zone 2, and potentially from Zone 1 also. However, initial limited commercial interest in clover from Zone 1 has likely inhibited the impact that clover from this region could have had relative to that of deliberately spread commercial varieties.



4.2.2 Discussion of main and variants of main nuclear ITS pattern types

Figure 4.7 Part of Northern and Central Europe, showing the approximate relative proportions of main pattern types:

5-Red, 5'- orange,

15- medium blue, 15'- pale blue,

20- purple,

17- forest green, 17'- grass green,
18/19- dark green

Pattern type 20'

The distribution of and reduced frequency relative to 5, 15, 20 of this pattern type (pale purple in Fig. 4.5) limited to Portugal and Spain, France and Italy, suggests in mixed populations there is a proportion of incomplete heterozygotes. The sequence from these individuals appears to have unequal contributions of each sequence type that are seemingly caught in a state of partial recombination or incomplete homogenization (Fig. 4.6). In these instances it seems the mechanisms of recombination and concerted evolution are yet to completely transform all ITS units in the tandem repeats of the NOR region into a specific class, (Song et al., 1995; Soltis et al., 2003; Kovarik et al., 2004; Kovarik et al., 2005; Matyásek et al., 2007).

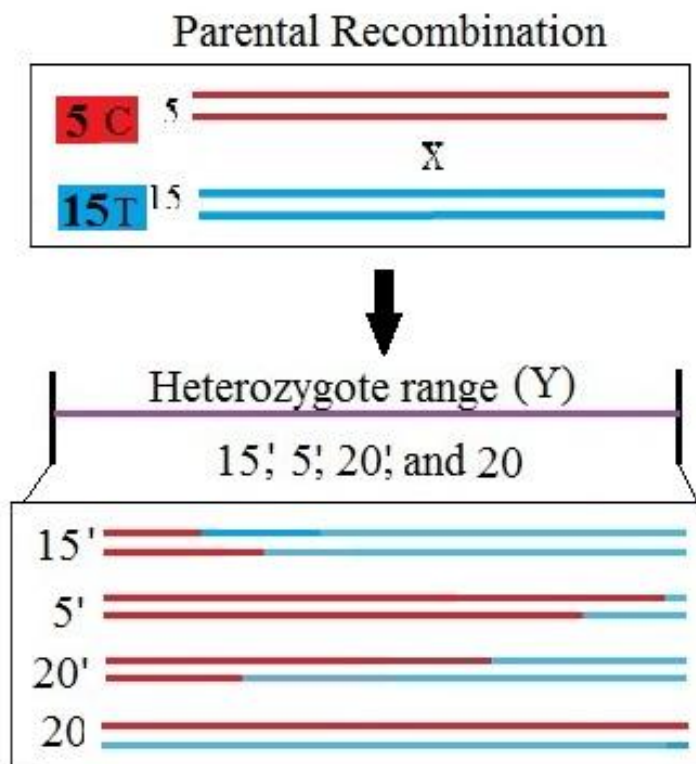


Figure 4.6 A schematic depiction of the proposed heterozygote class continuum for type 20/20' individuals, formed in mixed populations, from sequence types 5 and 15. Because white clover is cross fertilised, many plants would have pairs of NORs carrying mixtures of these types, creating a continuum.

Pattern type group 17, 17', heterozygote 18 and type 19

As can be seen from Figure 4.5, the distribution of pattern type 17 and the closely related types 18 and 19, represented in dark green, showed a clear northern Europe distribution.

The distribution of the European geographical restricted heterozygous group, 17, 17', 18 and related pattern type 19, indicates a pocket of diversity in north-west Europe/Scandinavia, Tammisto (Finland), Barbian (the Netherlands) and Morso (Denmark) (Figs. 4.5 and 4.7).

Pattern types 17 and 18

- Type 17 is most similar to type 15, but differs by an A/C SNP at position 571.
- Type 18 is the heterozygote of types 15 and 17.

The connection of the Netherlands and Denmark through north-west Germany could be hypothesized to have allowed for the development of a relative commonality between the three countries. Populations containing types 17' and 15 were observed in Belgium, The Netherlands and Finland. The absence of sampling in this study, of populations from Germany prevented further comment on the distributions of types 17, 15, and 18. However, in populations from Denmark and the Czech Republic there was an absence of type 5 and type 20 but presence of type 17 and type 18 (Fig. 4.7). This could be interpreted to mean that the presence of type 15 is crucial for the formation of both heterozygote classes (17 and 20). This interpretation is supported by prominence of type 20 when type 5 is more prominent than type 17, and by prominence of type 18 when type 17 is frequent. The absence of wider detection of type 18 may be due to small sampling sizes.

The occurrence of type 18 and related type 19 in Ovcak from the Czech Republic, but no detection of type 17, suggests that in central Europe type 17 may be infrequent and/or has been missed in sampling. Wider and more intense sampling is needed to clarify this.

The nature of type 17'

The uncertain nature of type 17' is due to the observation of an A at position 571 in one NOR region and a slight indication of a C at the same position in the NOR region of the other homolog. Three possible explanations could be considered.

- **Sequencing error**

This variable position 571, while near the end of the sequence, was always observed in a region of good quality sequence and all sequence runs were clean within this region. Individuals with 17 or 17' were compared to all other individuals sequenced from the same populations and across these populations to verify the nature of their variability. Verification via resequencing would be required as further proof of nature.

- **Mixed inheritance, actually type 18**

The heterozygosity of type 18 is shown by the presence of a C (15) in one NOR region and an A (17) in the NOR region of the other homolog. This variation is coded for by an M at position 571 (Huia reference sequence). The ambiguous nature of the type 17' individuals suggests they may be type 18. Differences were observed in chromatogram peak heights at position 571 (Huia reference sequence): these are indicative of differences in the relative contributions of each parental ITS type to the heterozygote. This variation across the class of individuals 17' suggests that the class spans different positions along the continuum of heterozygosity between the two assumed parental types, 15 and 17 (Fig. 4.8). Therefore, the class 17' does not contain first generation heterozygotes, as was proposed for 5'/15'/20'. This variation in pattern types can also be interpreted as the incomplete homogenization of all the ITS repeats in the individual into one parental type (Soltis et al., 2003; Kovarik et al., 2004; Kovarik et al., 2005; Matyásek et al., 2007).

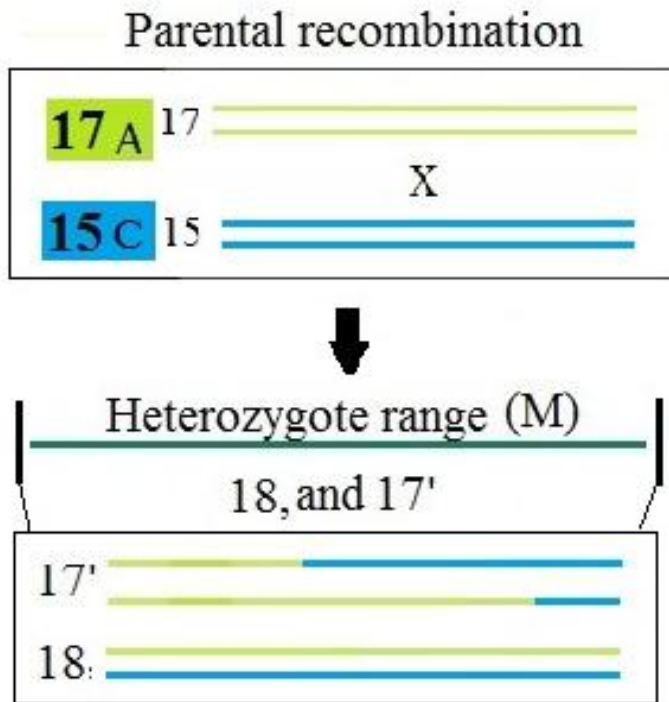


Figure 4.8 The heterozygote class continuum for type 18 individuals suggested to be formed from parental ITS types 17 and 15.

- **Bidirectional concerted evolution**

From the literature, it is known that in the case of concerted evolution the process favours the homogenization of ITS repeats to produce a homozygous situation of parental variation, rather than the conservation of the heterozygote form. This is usually represented by favouritism toward one parental type. However, it has been observed in that the opposite parental type sometimes becomes favoured (Wendel et al., 1995; Soltis et al., 2003). It is not known why and to what extent this occurs (Wendel et al., 1995; Soltis et al., 2003). Nevertheless, should this be the case, the transition of some of these 17' individuals to type 17 as opposed to type 15 should be observed, but to what degree, if any, these individuals contribute to the pool of type 17 individuals can only be speculated upon. This suggests that there is a mechanism acting that buffers both the parental classes as well as those variations of class that fall within the hypothetical space between the parental classes. This means that individuals, and hence populations, can change in either direction, to varying degrees.

4.2.3 Infrequently observed Nuclear ITS pattern types



Figure 4.9 A depiction of the approximate relative proportions of infrequent nuclear ITS pattern types.

Individual circles are numbered to show the corresponding ITS pattern type.

Colours coincide with those used in the pie charts and tables of the results.

The distribution and frequency of pattern types found between two and six times across the data set (Tables 3.12, 3.15), showed support for two regions of diversity: Morocco/Algeria, and Portugal (Fig. 4.9). Interestingly, about half the 22 pattern types in these two regions were unique to these regions, including the finding of the rare heterozygous pattern type 30, which was unique to Algeria. This pattern type appears to be the product of inheritance of pattern types 5 and 7, and both of these were found in Algerian populations. Limited observance and small sample size may mean that low detection of these ITS types inadequately reflects the natural diversity within the area.

4.2.4 Hypotheses of dispersal and origin with a focus on common ITS types

More than one model can be proposed to explain the variation in ITS sequence distributions. First, based on frequencies of occurrence of the main pattern types 5, 15, 17, and heterozygotes 20, and 18, a hypothesis of different origins and dispersal events in conjunction with glaciations and inter glacial movements is suggested (Fig. 4.10). This first model, Model 1 (Fig. 4.10) shows a suggestion of the origins of the main nuclear pattern types based on the assumption that the relative frequencies depict the order of origin for the pattern types, and that the three main lineages represent three separate *T. occidentale* lineages. It is suggested that the frequency of pattern type 15 indicates: an earlier origin than the other pattern types and hence has more time to spread, particularly throughout Zone 2. Under this assumption that frequency is related to time since origin, it is proposed that type 5 is the next most common pattern type. The rarer and apparently Northern European restricted pattern type 17 is suggested, (in the absence of more thorough sampling), to have originated later or remained largely restricted to Northern Europe.

Heterozygous pattern types

The two heterozygous types 20 and 18 formed from these main types are shown in pale blue boxes (Fig. 4.10). Based on observed occurrence and distribution of these types and the parental types it is suggested that type 20 is older than type 18. The initial point of origin in time or location for each of these heterozygotes could not be determined from this study. In fact both heterozygotes were not always observed when both parental forms were detected and in limited cases were observed in the absence of detection of one of the parents in individual populations.

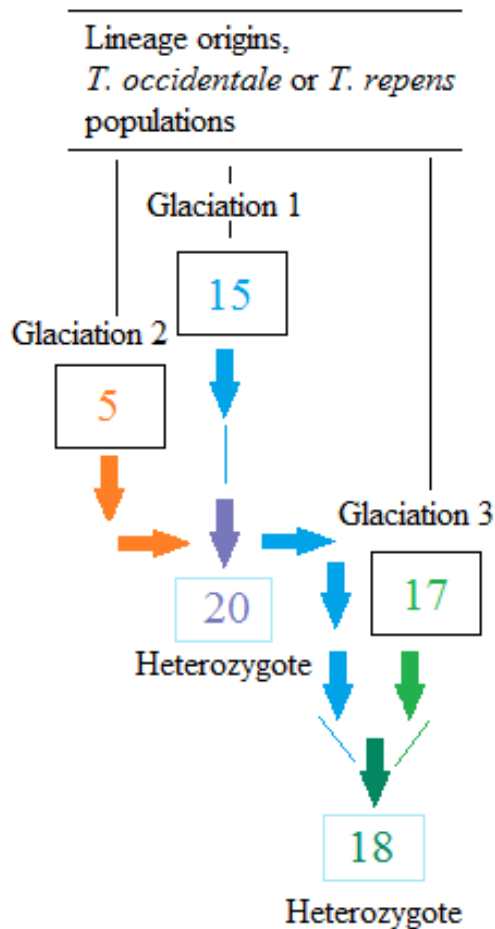


Figure 4.10 (Model 1) describes a suggestion of independent origins during different glaciations for the main ITS pattern types 5, 15, and 17. Potential parental inheritance leading to heterozygous types 20 and 18 is shown.

One ancestral paternal ITS type, type 15 (Model 2)

An alternative hypothesis (Model 2) is illustrated in Figure 4.11. This hypothesis is based on the observation of both parental types 5 and 15 occurring in close proximity, and often in the same country and region. This model suggests that the main types all originated from a *T. occidentale* plant bearing the most common type 15 which was transferred to *T. repens* upon speciation. It is suggested that the types 5 and 17 evolved from type 15 through spontaneous mutation in *T. repens*.

The difference in shading of arrows leading to the formation of types 5 and 17 represents the suggestion that the two lineages have significant differences in distribution and frequency and hence type 17 may represent a lineage produced by a later mutation from type 15.

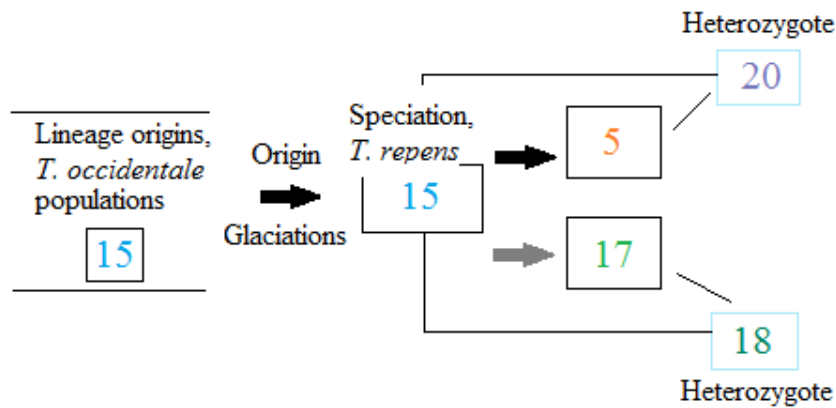


Figure 4.11 Model 2, an alternative hypothesis for origin of the main pattern types from a *T. occidentale* individual carrying pattern type 15 ITS sequence, from which it spread through the population. Note pattern types are coloured as in graphs and tables. Heterozygote types 20 and 18 are shown in pale blue boxes.

However, the observation of principally separate pattern types from different ancestral lineages in Portugal, Spain, Algeria and Morocco, suggests that type 5 is not derived directly from type 15. If Portugal and Spain are treated as one region and the population of type 5 through chance became isolated in Portugal, while type 15 became isolated in Spain. It is possible that type 5 was derived from type 15 and either spread widely through commerce or arose spontaneously multiple times (Figs. 4.5 and 4.11) across Europe.

However, determination of whether the types 5, and 17 represent independent lineages that evolved from type 15 (Hypothesis 2) or independently, but from a common ancestral pattern type (Hypothesis 1) that subsequently gave rise to three *T. occidentale* plants or post-speciation in *T. repens*, requires further research. Hence, while Model 1 (Fig. 4.10) may be the preferred hypothesis, there is insufficient data to determine the most likely of the two hypotheses.

A third model is suggested as another possible explanation for the observed variation in ITS pattern types. This is the introgression model (Fig. 4.12).

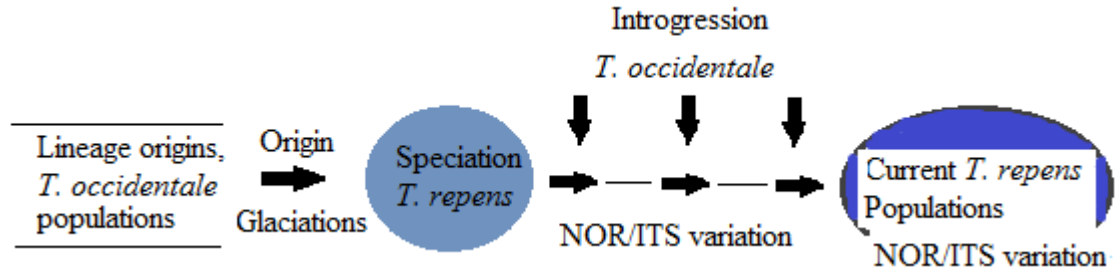


Figure 4.12 Introgression by *T. occidentale* populations. This model demonstrates how post-speciation; an influence via introgression from *T. occidentale* populations could have introduced novel variation into the modern *T. repens* populations.

The introgression by *T. occidentale* model (Fig. 4. 12) is the third alternative hypothesis suggested here; further research is required before this model can be excluded. This model suggests that variation in modern *T. repens* could have been produced from introgression over time from separate populations of *T. occidentale*. The occurrence and distribution of type 17 in northern Europe could be an example of a lineage caused by the introgression from *T. occidentale* through either inheritance of a novel *T. occidentale* lineage or through introgression.

Similarly, in the eastern Mediterranean region, in the Israeli cultivar Haifa, the pattern type 10 was observed. As the position of variability of (K) at 216 was not observed outside of this pattern type, and as little diversity was observed in the small sample from the eastern Mediterranean region, the possibility that pattern type 10 represents a case of rare introgression cannot be excluded.

Placing a molecular clock on the rate of evolution of the ITS region in both *T. occidentale* and *T. repens* would resolve the age of the observed variation. Further identifying species specific sequence features that can identify between NOR types from *T. repens* and *T. occidentale* would allow for estimation of when a pattern type originated.

4.2.5 Parental or Biparental inheritance of nuclear ITS pattern types in *T. repens*?

As white clover is a tetraploid (Williams, 1987a) derived from two divergent species (Ellison et al., 2006), ITS sequences from the maternal *T. pallescens* ancestor as well as the paternal *T. occidentale* parent might be expected. However, previous work on the NOR regions (Ansari et al., 1999) and sequencing of ITS sequences from *T. repens* has failed to find this. Furthermore, it is recognised that the *T. occidentale* and *T. pallescens* ITS sequences are clearly discernible, being different between ~7 positions and ~30 positions depending on the accession. These are further positions to those scored as variable within this study of *T. repens* individuals.

This study failed to detect any individuals that had an ITS sequence indicative of a *T. pallescens* origin, although it is possible that at some point during speciation process recombination between the ITS forms of *T. occidentale* and *T. pallescens* produced variation in ITS pattern types, some of which may have been observed in this study. However, it is more likely that variation observed in ITS pattern types is due to variation caused by mutation, natural selection, and rapid changes in rDNA (Kovarik et al., 2004) in the NOR regions inherited from *T. occidentale*.

4.3 Discussion of the combined DNA sequencing work from ITS and chloroplast results.

The interpretations of the results from the combined data set (*trn L*, IRB21NE, and ITS) have to be tempered by its small size. The accuracy of the combined results section is limited by the sub-sampling from the complete data set. For a true reflection of combined diversity a larger data set collected from a broader geographic range of white clover is required. However despite the limitations of the data set it still provides an insight into the diversity and origin/s of white clover.

4.3.1 Geographic distributions of combined (*trn L*, IRB21NE, and ITS) pattern combinations.

It was found that the majority (77%) of pattern type combinations were explained by combinations containing the four recognised chloroplast lineages (CC+, CT+, TC-, CC-) with the main nuclear pattern types 5, 15, 17, and 20, and their

variants (Tables 3.24-3.27). Of this variation over half (46%) was CC+ (Table 3.24) supporting this lineage as the most wide spread and potentially the oldest. The above results support the idea of free intermixing between out-crossing individuals, allowing chloroplast and nuclear pattern types to become recombined. This modifies the diversity in the populations in a way proportional to the initial level of diversity.

Combinations containing rare nuclear pattern types collectively accounted for 22% of observed combinations (Tables 3.24-3.27). The CC+ lineage occurred in 82% of these (Tables 3.24, 3.25-3.27).

These results seem to suggest that lineages CC+/15 and CC+/5 are among the oldest. The frequencies of these lineages suggest there is a mechanism for promoting their wide distribution. As these patterns predominated in the commercial populations, it is likely that commercial activity has been a major factor in their spread. This agricultural spread probably changed allele frequencies, swamped the local variation with CC+/15 and CC+/5 and created new combinations through intermixing and recombination.

4.3.2 Chloroplast & nuclear results do they agree: or disagree?

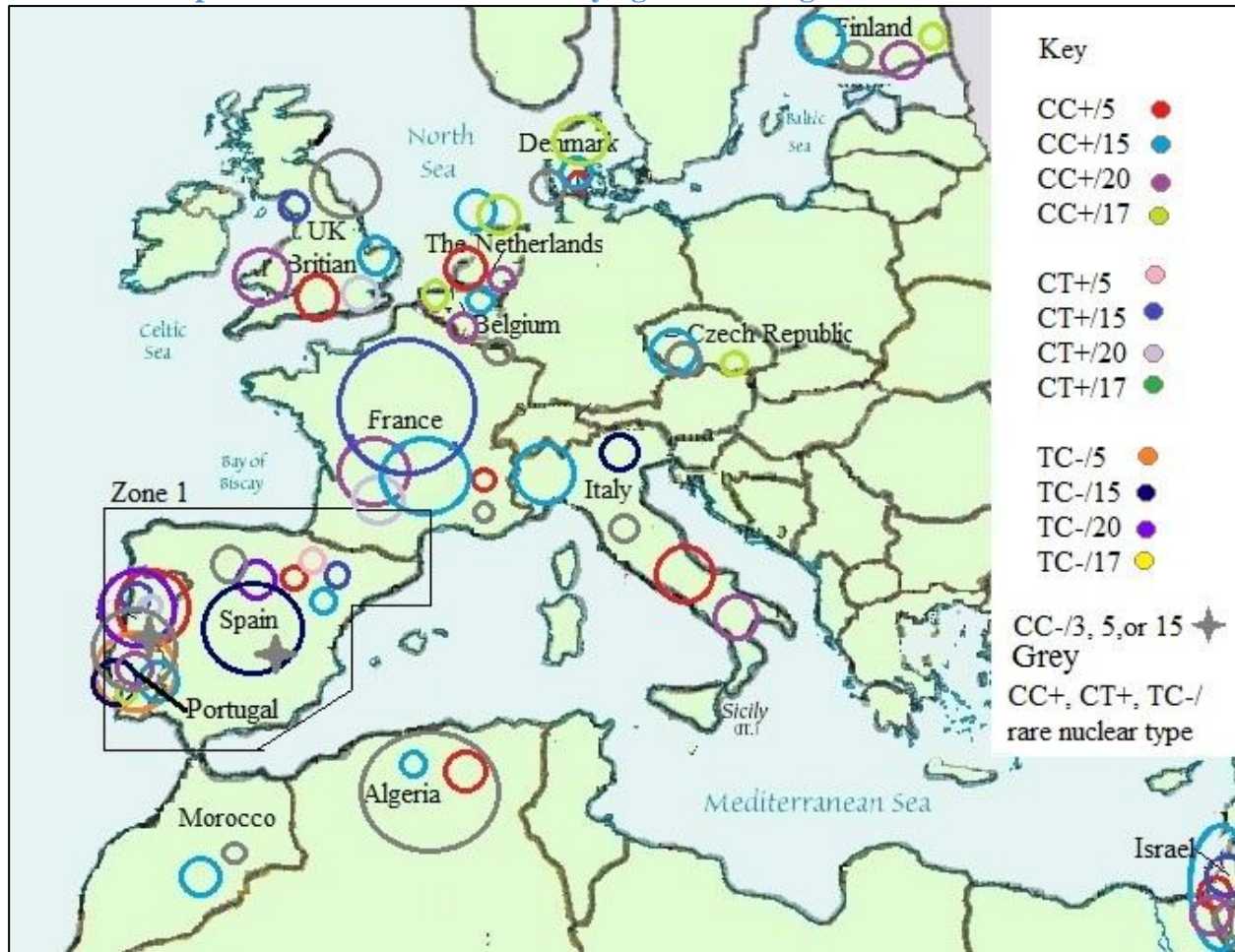


Figure 4.13 The approximate relative pattern combinations, shown as circles and coloured to show different combinations.

Note grey circles represent combinations with rare nuclear pattern types.

Grey stars (Portugal and Spain) represent the rare chloroplast lineage CC-.

Distribution of combined pattern types, diversity in the regions

Figure 4.13 showed that diversity in pattern types was greater in Zone 1, Portugal and Spain, than in Zone 2. Within Zone 2, two regions, Northern Europe and the UK had the greatest number of pattern combinations. This differed from the results observed for the individual nuclear and chloroplast data sets.

The regions of diversity

Portugal and Spain, Zone 1

The nuclear and chloroplast results both suggest Portugal as an area of diverse pattern types but the variation observed in Spain is less prominent (Fig. 4.13). However, as nearly twice as many individuals were sampled for ITS from Portugal as from Spain further sampling may show equivalent or greater diversity in Spanish populations.

North Africa

In the nuclear data North Africa was one of the most diverse regions. Although the absence of chloroplast variation in North Africa suggested singular or limited *T. pallescens*-like lineages within the region. The prominence of rare nuclear pattern types supports the hypothesis of multiple ancestral paternal lineages. However as the sample size was small, the results are tentative.

Northern and North-Western Europe

Northern and North-Western European populations showed a limited amount of chloroplast diversity, but included the novel nuclear ITS lineage 17. This geographically restricted lineage was observed with each of the chloroplast types CC+, CT+ and TC-, creating combinations not observed anywhere else. The combined results suggested that despite limited chloroplast diversity, the amount of variation in nuclear pattern types, even in absence of rare nuclear pattern types, indicates considerable diversity within the region (Figs. 4. 5 and 4.12).

UK

The UK showed multiple combinations based on chloroplast and nuclear variation. The high diversity within the Kent landrace populations could be explained by several factors: natural mutation over time, introgression from nearby *T. repens* or *T. occidentale* populations.

4.3.3 Diversity and models of origin, from a combined perspective

In order to evaluate models of origin of white clover it is necessary to consider the picture gained from a combined perspective, of nuclear and chloroplast data.

The main problem with this is that the different genetic markers, not only give different parental perspectives, but also evolve at different rates time (Eidesn et al., 2007; Volkov et al., 2007; Poczai and Hyvonen, 2010). The chloroplast genetic regions are thought to be under stronger constraint so evolve slower. This means that putting the two different perspectives together gives a better analogy of the actual story of origin.

An eastern origin, western origin or both for white clover?

Reasonable sequence diversity was observed in the pure Israeli sample, entirely through variation in ITS sequences and mostly involving combinations that also occurred to the west. Additionally, the high chloroplast diversity in South-West Europe and relative lack of chloroplast diversity in Central and Eastern Europe is consistent with a western region of origin. An eastern origin or influence cannot be ruled out under a “Melting pot” scenario (Petit et al., 2003; Marhold and Lihoviá, 2006; Parisod and Besnard, 2007). Interaction between the progenitor species speciation could have resulted in speciation. *T. pallescens* is known to occur in sporadic and widely scattered alpine populations in the eastern Mediterranean mountain ranges (Macedonia, Greece) (Professor Warren Williams, pers. com 2010).

In this study the population sampling strongly emphasised the western Mediterranean and western Europe, because this is where *T. occidentale* now grows. *However* had *T. occidentale* previously had a wider distribution in the east, a hypothesis for an eastern origin could be invoked.

Further what of separate eastern and western origins for *T. repens*? Given the amount of diversity recently observed within *T. pallescens* populations between glacial valleys in the Austrian Alps (Raffl et al., 2008) potential for island populations and diversity gradients, (Hewitt and Ibrahim 2001; Widmer and Lexer, 2001; Gabriel et al 2007; Parisod and Besnard, 2007), it is reasonable to assume that different populations of *T. pallescens*, could have contributed variation through separate speciation events to *T. repens*. This argues for the likelihood of multiple species origins.

It is assumed that was this, the case populations from the east and west should show significant genetic differences, due to different founding parental populations. To test this possibility further it will be necessary to carry out more intensive sampling of white clover populations from the eastern Mediterranean region, as well as characterisation of *T. pallescens* populations from that region. Currently none are available in seed banks world-wide and it will initially be necessary to analyse herbarium specimens (Professor Warren Williams, pers. com, 2010).

Discussion of Chloroplast and Nuclear Models in the light of combined data

Chloroplast Models

Model 1: (Fig. 4.2) suggested that each of the main chloroplast lineages was the product of glaciations bringing different *T. pallescens* populations into contact with *T. occidentale* populations, sequentially resulting in several speciation events.

Model 2: (Fig. 4.3) suggests that CC+ is the oldest lineage, with the three less frequent pattern types being derived from it (Fig. 4.4).

Nuclear Models

Model 1: Glaciations led to origins (Fig. 4.10), this model suggests that each of the three separate nuclear lineages 5, 15, and 17 resulted when glacial and interglacial movements brought *T. pallescens* into contact with different *T. occidentale* populations, bringing about separate events of speciation.

Model 2: an alternative hypothesis for origin (Fig. 4.11). This is the hypothesis that pattern type 15 was the ancestral nuclear sequence and that types 5 and 17 arose from it.

Model 3: The *T. occidentale* introgression model. The key point of this model is that variation observed in the nuclear data could be the result of more recent introgression from different *T. occidentale* populations after the formation of *T. repens*. The evidence excludes any introgression from *T. pallescens*, (Fig. 4.12).

Commerce

A further modifying influence is commerce, the effect of which is important and influential, but hard to establish. The argument is that commercial varieties have been spread over the top of natural populations, potentially obscuring natural variation in native populations.

Around the time of domestication of clover in southern Spain about the year 1000 AD, clover was widespread throughout Europe occurring in natural populations (Kjaegaard, 2003). It is likely that across Europe clover remained in a natural variable state until cultivation and domestication reached Lombardy and the Netherlands in the 16th century (Kjaegaard, 2003). Influence from commercial clover will likely have resulted in the wide spread of largely uniform varieties adding to the natural genetic variation present.

Northern and North-Western Europe

The influences on clover in Northern and North-Western Europe are not well known beyond, climatic influences and the similarity in French and Northern European white clover populations cannot be blamed wholly on mans influence through farming practice. The presence of the class of novel pattern types (17, 17' and 18) in Northern European populations and the novel pattern combinations extending from these suggest that there has been sufficient time within this area to give rise to a lineage that is geographically restricted. However, without adequate sampling of the surrounding countries it cannot be determined if this observed distribution is evidence of an old or recent origin for the lineage. Sampling white clover from Germany where it was established in 1645 (Kjaegaard 2003) will help establish if the CC+/15 and others were introduced to this area, or were part of the existing natural variation prior to commerce and where lineage 17 is most frequent and likely to have originated from.

Further, this sampling may show that Germany acts a transition zone between Denmark, the Netherlands and Belgium. By sampling populations along and across the border between France and Germany, an indication of the extent and direction of the spread of white clover in these countries can be gained.

Multiple origins for white clover?

The models outlined in the chloroplast and nuclear sections collectively show the complexity of the problem for deciphering the events that caused the level of intra species diversity observed in current populations of white clover.

The data clearly show the greater contribution to diversity observed from the *T. occidentale* populations relative to the chloroplast contribution. It is clear that the ITS sequences of *T. pallescens* and *T. occidentale* are quite different and only *T. occidentale* like sequences were found in all the individuals sampled in this study. The present evidence suggests that if multiple speciation events have occurred; then the maternal NOR regions must have been lost or suppressed every time. FISH results on *T. repens* individuals are consistent with this, repeatedly showing no detectable signal from the 18S gene of the NOR on *T. pallescens* like chromosomes, (Ansari et al., 1999). If multiple origins have occurred, then cross talk between the parental genomes has consistently resulted in loss of the same parental pair of NORs, as is observed in some other allopolyploids (Kato et al., 2005; Wendel et al., 2005; Volkov et al., 2007).

Other allopolyploids

Alternatively, it would seem that the ground is set for a single origin for *T. repens*, with current distribution of diversity indicating this origin as within Zone 1 Portugal and Spain. Natural dispersal could then be responsible for the distribution of white clover throughout Zone 2 and, more recently, human influence has spread the species throughout Zone 3.

Under this hypothesis all of the observed diversity in nuclear pattern types and multiple chloroplast lineages would be due to recent mutation, natural selection and environmental influences. This hypothesis fits with the chloroplast model 2 in which all the chloroplast lineages are derived from CC+ and nuclear models 1 and 3 that suggest variation post speciation and the possibility of introgression for further changes in allelic frequencies and diversity of DNA sequence.

4.4 Discussion of the perspective gained from TCS networks

The TCS analysis of diversity patterns in DNA sequence showed in general the same distribution in space and connections of pattern combination groups to each other for all three sets of analyses: (1) the two chloroplast (*trn L* and IRB21NE) regions; (2) the nuclear ITS region; and (3) the combined analysis of all three (Figs. 3.33-3.38).

However, the analyses shed little light on hypotheses of origin. The highly conserved spatial arrangement of the networks across the analyses is constant with a singular origin for white clover. The amount of individual variation observed is due in part to irregularities in sequencing and novel DNA sequence variants.

For instance:

- Group 12A of primarily of Spanish origin but also including, individuals from Portugal, Lodi and Kent. This group is unique in their presence of a T at position 458 (Grasslands Huia).
- CC+* individuals

The network diagrams did tentatively suggest;

- The chloroplast network diagrams showed clear separation for the three main chloroplast lineages and the two included CC+* individuals.
- The nuclear networks suggested relationship between the main ITS types 5, 15, 17 and 20 and gave indications about regions of genetic diversity.

Regions of diversity

The TCS network diagrams showed closeness of individuals from Algeria and Portugal. Further connections were observed between this group and a second grouping of Algerian and Portuguese individuals. The placement of these individuals within the space of the diagram further supports the observation of similarity in prominence of Pattern type between these countries. Given the nature of the *T. occidentale* contribution to *T. repens* and the involvement of glacial events in the speciation of *T. repens* it is suggested that the similarity here is indicative of spread of old *T. occidentale* or *T. repens* lineages between these areas across the Gibraltar strait during interglacial periods; as has been observed for populations of Mediterranean conifers (Jaramillo-Correa et al., 2010).

However proper conclusions could not be drawn from the TCS analyses and further work is required before a clearer picture is gained. Assessment of further genetic regions, further points of variability and broader geographic and increased population sizes will add this.

4.5 Discussion of Combined DNA sequence pattern types and Morphology

This study found that there is no correlation of morphological characters with specific DNA sequence combinations, within populations, across or within geographic regions that were not influenced by sampling bias. This was supported by chi squared assessment (Section 3.5). As the three genetic regions assessed (chloroplast *trn L* and IRB21NE, nuclear ITS) have no apparent connection or influence on morphological character and sample sizes varied and were small 10, or less plants per population, any correlations are tentative and likely due to chance sampling.

The only morphological trait to really show a consistent trend was that of leaflet size. This trend fitted with recognized expectations from the literature of gradients of leaflet size across the Mediterranean basin, where tiny to small sized leaflets are observed with increase in altitude and latitude, while more Southern populations and those at lower altitude have larger sized leaflets (Davies et al., 1967; Caradus et al., 1990). As leaflet size is environmentally and genetically determined (Caranahan et al., 1955; Davies et al., 1967; Corkill, 1971), it is a reliable trait with which to compare DNA sequence variation by chi squared analyses.

To further test for associations between DNA sequence variation and morphological appearance Chi squared analyses were also preformed for comparisons of variation in V mark character (Tables 3.38 – 3.41) and leaflet shape. In both these cases no associations between characteristics were observed. Assessment of leaflet shape found no significant associations of characteristic paternal *T. occidentale* obcordate shaped leaflets (Pers. Com. Professor Warren Williams 2009) (Table 3.42); most leaflets were ovate or orbicular in shape (Table 3.42).

However significance was observed for an absence of large sized leaflets in the pooled sample of Portuguese and Spanish individuals (Table 3.43). This is interesting in that it

suggests: A bias in sampling from populations of higher altitudes, where environmental effect is likely to strongly influence leaflet size. Further has both the parental species are typically small leaved (Pers. Com. Professor Warren Williams, 2009-2010) this observation lends weight to a hypothesis of a primitive *T. repens* type being present in this region, and hence supporting the involvement of this area, in the origin of *T. repens*.

4.6 Concluding remarks

The single origin model is a simpler and more readily explainable model of origin given the present knowledge of the distribution of genetic diversity.

- The level of diversity found within the study
- Highest diversity was observed in Portugal- five chloroplast types, 19 nuclear
- Four rare individuals from Portugal and Spain with chloroplast DNA sequences more like the Pyrenees *T. pallescens* than previously observed.
- Diversity gradients, across, Portugal/Spain and Morocco/Algeria
- The presence of variant forms 5', 15', 20', and 17'.

4.5 Limitations of this study

There were three main limitations to this study, the size of the sample set and number of individuals, the scattered sampling of locations from Zone 2, and the method of direct sequencing of PCR products.

- **Sample set size**
 - The number of individuals sampled per individual population.
 - Number of populations sampled.
 -
- **Distribution of sampled populations sampled**

Desired regions for sampling from Zone 2 are:

- Eastern Mediterranean
- Western Mediterranean
- Central Mediterranean
- UK- Scotland, Ireland, Wales and England.

- Inadequate sampling from populations in the north and middle region of France and hence inadequate assessment of a potential diversity gradient across France was observed.

- **Direct sequencing technique**

This technique is recognised as not being as accurate as cloning (Rauscher et al., 2004). Recent literature (Rauscher et al., 2004) also states how the technique particularly, when PCR with general primers is used, can amplify minor sequence types that cloning and resequencing methods fail to detect (Rauscher et al., 2004). The ratio of the observed nucleotides at a particular position of variability is of particular concern for consideration of the variant classes of the main ITS pattern types 5', 15', 17' 20'. This is because when the ITS region is in a state of incomplete homogeneity any difference between the points of variation in the ITS copies within and between homologs has a chance of being picked up by PCR and copied many times. This means that the actual ratio of different nucleotides at a SNP may not be a true reflection of the nature of the individual.

4.8 Future directions, extensions from this work

Immediate Future directions of this research are;

4.8.1 Extension of the sampled area

This study has shown that there is a remarkable amount of intra species diversity both between and within geographic regions, towns and even populations of white clover, of both natural and commercial origin.

However the area sampled from captured only a snapshot of the potential variation in white clover. Extending the sample area will allow better assessment of the natural and commercially influenced diversity in white clover.

4.8.2 Single copy markers

The analysis of single copy genetic markers is advantageous as these sorts of markers are less prone to homogenization so provide a good picture of parental inheritance in polyploid individuals; allowing identification of phylogenetic relationships.

Finding *Trifolium* species specific markers would help strengthen the analysis of white clover formation. Assessing variants of these single copy genetic markers may allow for population-specific identification of individuals. As has been observed in *Cardamine asarifolia* ref A recent study in this species looked at the Chalcone synthase (CHS) gene, and found that through primer specific sequencing of the exon 2 parental variations could be determined. Combining single copy genetic variation with geographic locations will help to further sort and assess the relationships and likely points of origin within the white clover species.

4.8.3 Chromosome specific markers

Another application for identifying phylogenetic closeness is the technique of chromosome specific markers. Identifying chromosome specific markers and relating this to individual species will allow assessment of the nature of the chromosome complement in hybrid species, like white clover. By having these markers assessment can be made of the content of the different chromosomes and whether they contain regions with sequences indicative of another species. Hence, indicating the presence of hybridization, introgression and a hybrid nature.

- Commercial Application

Another application of single copy and chromosome specific markers is the application of these to the assessment of commercial varieties. If markers can be found that are specific to certain populations used in the backgrounds of today varieties then this would serve as a screening method for quality control and genetic variation assessment for commercial varieties.

4.8.4 Phylogenetic analysis

A phylogenetic analysis of the chloroplast, and nuclear ITS DNAs sequences collected in this study would give an assessment idea of the relationships, between the different chloroplast and nuclear pattern types. Further a combined analysis of the three regions of genetic diversity, with comparison to geography would help to breakdown of intra species diversity, and locate potential areas of origin for each of the chloroplast, nuclear lineages and those combination types created by their combination.

Further this analysis may also help resolve the level of diversity in the region designated as zone 1 the area of origin, (Portugal and Spain) and potentially Southern France.

4.8.5 Testing models

Another area of research is the testing of hypothesised models of origin and spread. New scientific practices and techniques will allow further exploration of the models outlined in the chloroplast and nuclear sections, (Figs. 4.1-4.3, 4.9-4.11).

- Paternal variation, by Introgression model

To test for maternal ITS variation FISH analysis of *T. repens* individuals with infrequent combination pattern types or infrequent nuclear ITS pattern types should be conducted to confirm that these individuals have expression of the paternal NOR region. If any individuals are found that have maternal ITS expression, from the maternal NOR region then assessing the frequencies of different chloroplast pattern types in these individuals and the geographic location of these will help to determine if there has been different events of speciation of white clover in different places. Application of chromosome specific markers in conjunction with FISH will allow assessment of whether these unusual pattern types are the product of inheritance of chromatin rather than whole chromosomes from a *T. pallescens* or *T. pallescens* like ancestor or even sub species variation.

- Glaciations, origins and speciation

Further through use of these techniques that allow closer inspection of the composite DNA of the individuals, collective points of DNA variation can be gained that will allow the application of a molecular clock. The use of a molecular clock would allow estimation of the time of origin of the species and to assist in determining if there have been multiple sequential origins, in line with geological events.

4.9 Conclusion

The aim of this survey was to assess the radiation and origins of white clover, through assessment of genetic diversity. The level of intra species diversity was so great that despite, the scattered and limited sampling within zone 2, (the natural distribution zone), two main regions of diversity Portugal-Spain and North Africa, and two other regions of diversity Northern/North-Western Europe and UK were indentified. Of these Portugal appears as a “hot spot” of variation, being more numerous in both chloroplast and nuclear lineages. This was further supported by the results of the combined analysis. Although, there are multiple explanations for the diversity observed across the data set, there is strong support for a hypothesis of a single origin, within the mountainous regions of Portugal.

Speciation within this region is assumed to have been influenced by glacial and interglacial periods, while relative absence of commercial involvement, has maintained the natural diversity within the area. The absence of a correlation of morphological character and DNA pattern combinations across the data set, adds no strength to this argument. Instead it indicates the interaction of the environment and genetics in producing diverse morphological characteristics.

Hence while it is suggested that the area of origin is within Portugal, the inconclusive nature of this survey and presence of other areas of diversity does not preclude an alternative origin or multiple origins for *T. repens*.

Chapter 5 Novel Algerian accession C20859

5.1 Introduction

One Algerian population C20859 was different from white clover in both morphology (Warren Williams and Isabelle Williams, pers. comm..) and in the trnL intron region of the chloroplast sequence amplified by the CD Universal primers, (Taberlet et al., 1991). This unusual population was examined further and appeared to be a *Trifolium* species of unknown identity.

Results

5.2 Chloroplast DNA Sequence differences

The sequences generated from the CD universal primers for the chloroplast non-coding region (Table 5.1) showed two SNPs at positions 234, and 281 of the Huia reference sequence. At these two positions a T nucleotide was observed in the Algerian species C20859 (Table 5.1). This was not seen in the Huia reference sequence or any of the sampled populations of white clover. A further five bp deletion (TATTA) were observed in this Algerian species, corresponding to positions 288 to 292 in the Huia reference sequence (Table 5.1). This deletion was not observed in any of the white clover individuals sampled.

Comparison of the chloroplast sequences derived using the shortened version IRB21NE primers (Dr Nick Ellison, pers. comm. 2009) showed conservation of the *T. repens* DNA sequence. The only polymorphism was the presence or absence of the same penta-nucleotide repeat, (CAAAA) as was observed in white clover populations. Of the five individuals assessed four had the repeat, while in the fifth plant the repeat was absent.

Table 5.1 Sequence differences in the *trn L* intron chloroplast regions of Grasslands Huia reference sequence and *Trifolium* species C20859.

Genetic region	DNA position	Grasslands Huia <i>T. repens</i>	C20859, Algeria
C/D (Universal primers)	207	T	T
C/D (Universal primers)	234	A	T
Non-coding			
C/D (Universal primers)	281	G	T
Non-coding	288-292		
C/D (Universal primers)	TATTA insert	+	-
Non-coding			
C/D (Universal primers)	433	C or T	C
Non-coding			
C/D (Universal primers)	456	C or T	C

5.2.2 BLAST (Basic Alignment Search Tool)

BLAST search results based on the chloroplast CD sequence returned an E value of 0.0 to at least 98 hits of species of the *Trifolium* genus. All of these were to the same region of the tRNA Leucine intron. Max identity ranged from 100% to 95%, with the highest max score of 1070 for *T. occidentale*. In this alignment no gaps or differences in nucleotide sequence were scored for the length of the inquiry of 579 nucleotides (Table 5.2). Similarity was observed to all the species sampled but these were all different from the sequences of *T. occidentale* and C20859 (Table 5.2).

The BLAST results of the comparison of the IRB21NE sequence against the nucleotide data base (nr) found only one hit to a *Trifolium* species - that of *T. subterraneum* whole genome chloroplast sequence. This is likely due to an absence IRB21NE sequences from *Trifolium* species in Genbank. The next BLAST hits were to species of *Medicago*. Whilst 100% coverage was achieved, the max score was about 891 for *T. subterraneum* and 848 for the *Medicago* species. Further comparison of *T. repens*- Huia and C20859 showed that both species are dimorphic for the single recognised point of variation +/- of the pentanucleotide repeat at positions 226-230.

Table 5.2 Chloroplast sequence differences between C20859 and related *Trifolium* species, produced by CD primers

Genbank Accession No's	Species	Positions of variation												
		146	149 insert	149 5A's	207	234	281	Deletion	291 insert	358	361	262- 264	442	
	Huia, (<i>T. repens</i> NZ)	A	-	-	T	A	G	-	-	T	C	TTT	C	
DQ311961.1	<i>T. repens</i>	A	-	-	T	A	G	-	-	T	C	TTT	C	
DQ311895.1	<i>T. retusum</i>	A	-	-	T	T	T	-	ATA	T	C	TTT	C	
DQ311811.1	<i>T. isthmocarpum</i>	T	A	-	T	T	T	290	ATA	T	C	TTT	C	
DQ311927.1	<i>T. suffocatum</i>	A	-	-	T	T	T	-	A-A	T	A	TTT	C	
DQ311939.1	<i>T. uniflorum</i>	A	-	-	T	T	T	282/287	ATA	T	C	AAA	C	
DQ311852.1	<i>T. occidentale</i>	A	-	-	T	T	T	288-292	-	T	C	TTT	C	
N/A	C20859	A	-	-	T	T	T	288-292	-	T	C	TTT	C	
DQ311931.1	<i>T. thalii</i>	A	-	-	T	A	T	288-290	-	T	C	TTT	C	
DQ311862.1	<i>T. parnassii</i>	A	-	-	C	A	T	288-290	-	T	C	TTT	C	
DQ311796.1	<i>T. glomeratum</i>	A	-	-	C	A	T	288-290	-	G	C	TTT	T	
DQ311848.1	<i>T. n. subsp. petrisavii</i>	A	-	+	C	T	T	288-290	-	T	C	TTT	C	
DQ311859.1	<i>T. pallescens</i>	A	-	-	C	A	G	145-148	-	T	C	TTT	C	

Note: Nucleotides highlighted in red represent SNP's, deletions and inserts. Nucleotides highlighted in orange indicate a site where two different nucleotides are found. All named species shown are from section *Trifoliastrum*.

5.3 Nuclear DNA sequences

The BLAST alignment results for the nuclear ITS sequence showed that there was high similarity of species C20859 to *T. occidentale* and *T. repens*. The best four hits were to these two species. *T. pallescens* was among the top 17 hits, but had a markedly lower max score and total score, 1225, than was observed for *T. repens* 1269-1258 or *T. occidentale* 1262, (Table 5.3). The subspecies *T. nigrescens* subsp. *meneghinianum* was more distant than *T. pallescens*. Comparison to a second *T. pallescens* accession showed even more variation. These more variable accessions were excluded from Table 5.3.

From the Table 5.3, it was observed that species C20859 showed two different ITS sequences that differed by the presence of an A or G at position 111. This variation in one individual of C20859 and a deletion at position 199 in *T. occidentale* were the only two points of variance between species C20859 and *T. occidentale*. Ambiguous variation was observed within and between the three accessions of *T. repens*. These ambiguities were not observed in species C.20859.

BLAST searches suggested closeness outside of *T. occidentale*, of *T. nigrescens* subsp. *petrisavii*, *T. retusum* and *T. suffocatum*, to species C20859. Of these three species, *T. retusum* was the most similar differing by a C at position 139. *T. suffocatum* and *T. nigrescens* subsp. *petrisavii* were the next most similar differing by two positions each. Of these two species *T. nigrescens* subsp. *petrisavii* was the only one to show a T at position 139, which is the same as *T. repens*, *T. occidentale* and species C20859.

Table 5.3 SNP variations in ITS sequences between species C20859 closely related *Trifolium* species.

Species	Positions in the Huia reference sequence																	
	50	88	101	111	132	136	139	140	199	234	455	509	512	515	535	564	596	644
Huia, (<i>T.repens</i> , NZ)	T	T	T	G	A	C	T	G	T	T	T	T	T	C	C	G	G	C
<i>T. occidentale</i>	T	T	T	G	A	C	T	G	-	A	T	C	T	C	C	G	G	C
20859 (14/60)	T	T	T	G	A	C	T	G	T	A	T	C	T	C	C	G	G	C
20859 (97)	T	T	T	A	A	C	T	G	T	A	T	C	T	C	C	G	G	C
<i>T. repens</i> ₁	W	T	T	G	A	C	T	G	T	A	T	Y	T	C	C	G	G	C
<i>T. repens</i> ₂	A	T	T	G	A	C	T	G	T	A	T	C	C	C	C	G	G	C
<i>T. retsum</i>	T	T	T	G	A	C	C	G	T	A	T	C	T	C	C	G	G	C
<i>T. n.</i> subsp. <i>petrisavii</i>	T	T	A	G	A	C	T	G	T	A	T	C	T	G	C	A	G	C
<i>T. n.</i> subsp. <i>nigrescens</i>	T	G	T	G	A	C	C	G	T	A	T	C	T	C	C	A	G	T
<i>T. suffocatum</i>	T	T	T	G	A	T	C	G	T	A	T	C	T	C	C	A	G	C
<i>T. pallescens</i>	T	T	T	G	T	C	C	A	T	A	A	C	T	C	T	A	A	C

Note: Nucleotides highlighted in red represent SNPs. Nucleotides highlighted in orange suggest a site where two nucleotides are found widely. Genebank accession numbers are AF053168.1, F053172.1, DQ311962.1, DQ312145.1, DQ312103.1, DQ312101.1, DQ312218.1, DQ312102.1 and DQ312111.1. Novel Species C20859 is shown in dark purple. Included to show dissimilarity is the most similarly *T. pallescens* individual, DQ312111.1. Subspecies *T. nigrescens*. subsp. *meneghinianum* has been excluded due to dissimilarity. All named species shown are from section *Trifoliastrum*.

Table 5.4 An example of the species results from the BLAST search of the ITS region of C20859 against the nr database.

Genbank Accession No's	Species	Max Score	Max Identity
F053172.1 ¹	<i>T. repens</i>	1269	99%
DQ311962.1 ³	<i>T. repens</i>	1264	99%
AF053168.1 ¹	<i>T. occidentale</i>	1262	99%
EU036988.1 ²	<i>T. repens</i>	1258	99%
DQ312145.1 ³			
AF053173.1 ¹	<i>T. retusum</i>	1258	99%
DQ312103.1 ³	<i>T. n. subsp. petrisavii</i>	1258	99%
DQ312218.1 ³	<i>T. suffocatum</i>	1253	99%
DQ312101.1 ³	<i>T. n. subsp. nigrescens</i>	1242	99%
DQ312111.1 ²	<i>T. pallescens</i> AZ1895	1225	98%

Note: 35S rRNA, = 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1 (ITS1), 5.8S ribosomal RNA gene and internal transcribed spacer 2 (ITS2), complete sequence; and 26S ribosomal RNA gene, partial sequence (VOLKOV 2004). (¹)= direct submission by Ellison, N.W., and Williams, W. M., (1998), Phylogenetic relationships in the genus *Trifolium* based on sequences of the internal transcribed spacers of nuclear ribosomal DNA, (unpublished).

(²)= direct submission by (Vink, C. J., and Phillips, C. B., (2007), First record of *Sitona discoideus* Gyllenhal 1834 (Coleoptera: Curculionidae) on Norfolk Island, New Zealand Zoology Journal, **34**, 283-287.

(³) = Ellison, N. W., Liston, A., Steiner, J.J., Williams, W. M., and Taylor, N.L, (2006) Molecular Phylogenetics and Evolution, **39**, 688-705.

5.4 Morphology of species C20859

In the glasshouse it was observed that the flowering pattern of C20859 was distinctly different from that of *T. repens*. Instead of single unbranched axillary peduncles bearing flowers (as in *T. repens*), branched axillary inflorescences formed instead. This is more reminiscent of the annual species, *T. nigrescens*. The leaflet shape in young plants was typically heart shaped, but this changed to primarily orbicular, but also ovular leaflet shapes in older plants. Leaflets were tiny - small, (Table 5.5).



Figure 5.1

Heart shaped leaflet



Figure 5.2

Orbicular shaped leaflet



Figure 5.3

Ovular shaped leaflet

Table 5.5 Combined pattern types and corresponding morphological characters of the novel Algerian *Trifolium* species 20859.

Individuals	Accession		Leaflet size	Leaflet shape
	Number	White V-mark		
1 (14)	20859	v. faint	T-S	or
2 (59) & 3 (60)	20859	absent	T-S	or
4 (179)	20859	absent	T-S	ov

Note: Four individuals were scored. Numbers in brackets represent plant reference numbers. Leaflet shape: ov (ovular), or (orbicular). T-S: Tiny- small leaflet size.

5.5 Subsequent analyses on species C20859

As an initial step of identification, chromosome counts were performed by Dr Wajid Hussain (2009). It was found that the species is diploid ($2n=16$) as compared to $2n=32$ for white clover.

Crossing experiments were performed by Professor Warren Williams and Isabelle Williams, 2009/2010. These crosses showed that species C20859 is self incompatible and crosses to *T. occidentale* and, possibly to *T. repens* (W. Williams pers. comm.).

Further attempts to identify this species, were conducted by other members of the research team. This included FISH assessment of metaphase chromosomes by Dr. Helal Ansari, (2010).

Dr Ansari's work showed that in initial assessments of one individual the new species appears to have a similar rDNA hybridisation pattern to *T. occidentale* and *T. repens*. (Ansari et al. 1999).

Cyanogenesis assessment has also been performed on C20859 by Michelle Williamson, (2010). Individuals of C20859 show the recessive ac/ac li/li genotype.

5.6 Interpretation of collective results of species C20859

Collectively the results seem to indicate that species C20859 has more characteristics in common with *T. occidentale*, than any other known *Trifolium* species. The chloroplast sequences obtained with the CD universal primers (Taberlet, et al., 1991) were similar to that of *T. occidentale* and differed from those of *T. repens*.

It is hypothesized that, as this species is diploid and not tetraploid, it is a subspecies of *T. occidentale*, or at least a very close relative of *T. occidentale*.

5.7 Discussion of the Algerian species C20859

Initial thoughts

Initially, given the DNA sequence similarity of C20859 to *T. occidentale* and *T. nigrescens subsp. petrisavii* and the similar reproductive morphology to the latter, it was hypothesized that a contribution from *T. nigrescens subsp. petrisavii* to species C20859 was conceivable. Further given this closeness in DNA identity between *T. occidentale* and species C20859, and the presence of a large amount of variation in ITS pattern types found in Algeria, it maybe that diversification of the line that gave rise to *T. occidentale* has undergone further diversification in Algeria producing a novel sister species, C20859. If this species was present during the time of glaciations and speciation of *T. repens*, then a contribution to *T. repens* could easily have occurred.

Although given that C20859 is diploid (Dr Wajid Hussain, pers. comm., 2009), the hypothesis of a hybrid of *T. nigrescens subsp. petrisavii* and *T. occidentale* is unlikely to be correct. However a contribution to *T. repens* cannot yet be ruled out. Identification and application of chromosome/species specific markers will help to identify areas of

recombination in chromosomes of Algerian *T. repens* individuals that could find introgression from species C20859.

The hypothesis of a “*T. repens* in reverse” e.g. the cross of *T. occidentale* female to *T. pallescens* male, was rejected for several reasons, including diploidy, a strong similarity to *T. occidentale* ITS sequence, and because eleven better BLAST hits were obtained before similarity was found to the ITS sequences of *T. pallescens* and close relatives. This left the hypothesis of a unique species, with some traits similar to both *T. nigrescens* and *T. occidentale* but different from either.

The FISH results (Dr. Helal Ansari, pers. com, 2010) also supported high similarity to *T. occidentale*. Alternatively, the maternal contribution to species C20859 could be from an as yet unidentified common ancestor of the *Trifolium nigrescens* group and *T. occidentale* (Fig. 5.4).

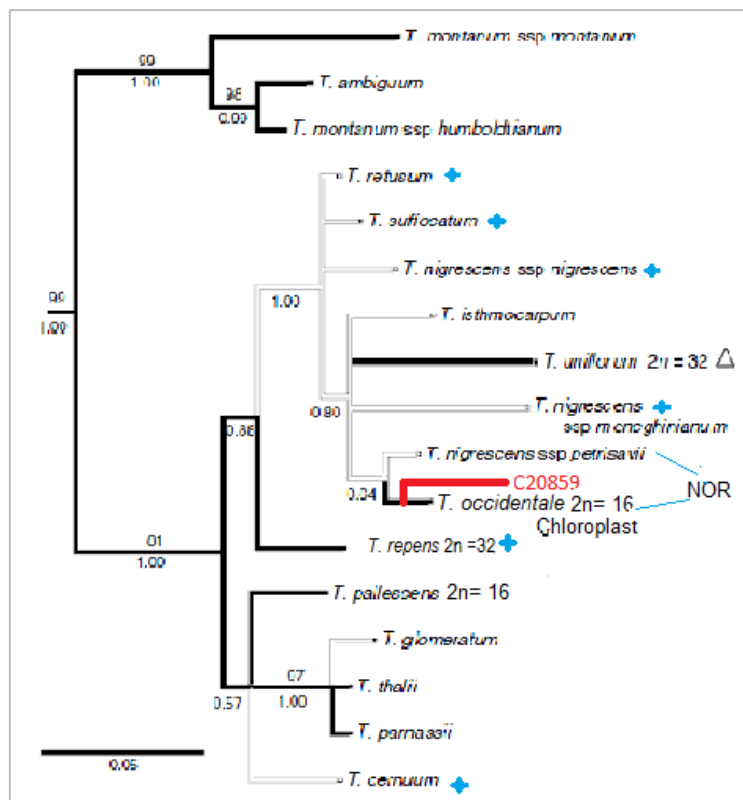


Figure 5.4, Current phylogeny of the *Trifolium* section *Trifolium* species (Ellison et al., 2006). Shown by blue stars are species found by BLAST search to share close ITS sequence similarity with C20859, Black lines indicate perennial life history and grey lines indicate annual life history. Figure adapted from ELLISON, N. W., LISTON, A., STEINER, J.J., WILLIAMS, W. M., AND TAYLOR, N.L, 2006 Molecular phylogenetics of the clover genus (*Trifolium*-Leguminosae). *Molecular Phylogenetics and Evolution* **39**: 688-705.

5.8 Conclusions on Novel Algerian species C20859

From the research done to date on species C20859, it seems that two alternative hypotheses can be suggested as likely explanations of origin. First, it could be that C20859 is a subspecies of *T. occidentale*, that has arisen at high altitude in an isolated environment, but which due to glacial events has been brought in contact with and subsequently been influenced by populations of *T. repens*, *T. occidentale*, and *T. nigrescens*. Second, and more likely, it is an ancient sister species to *T. occidentale* that has remained in an area of isolation that has prevented its discovery before now.

Further the difference in morphology is in line with the effect of evolutionary forces promoting adaptation to the local environment, with differences and similarity being invoked through the mechanism based on different and common or similar environmental influences in each of the species.

To test these hypotheses it is suggested that knowledge be gained of the *Trifolium* species present in the area and that wider sampling be carried out to determine the range of this species and its rarity or prominence particularly in Algeria. Further sampling in mountainous areas for instance in Morocco/Algeria, Portugal/Spain, Andorra and Eastern Mediterranean, may show a wider distribution of the species.

References

Articles and Books

- Abberton, M. T. (2007). Interspecific hybridization in the genus *Trifolium*. [Review] *Plant Breeding* 126, 337-342.
- Abbott, R. J., Smith, L. C., and Milne, R. I., *et al*, (2000). Molecular analysis of plant migration and refugia in the Arctic, *Science*, 289, 1343–1346.
- Abbott, R. J., James, J. K., Milne, R. I., and Gillies, A. C. M. (2003). Plant introductions, hybridization and gene flow. *Philosophical Transactions of the Royal Society London*, 358, 1123-1132.
- Allen, R. C. (2008). The Nitrogen Hypothesis and the English Agricultural Revolution: A Biological Analysis. *JOURNAL OF ECONOMIC HISTORY* 68, (1), 182-210.
- Allan, H. H. (1982). Flora of New Zealand, Volume 1, INDIGENOUS TRACHEOPHYTA PSILOPSIDA, LYCOPSIDA, FILICOPSIDA, GYMNOSPERMAE, DICOTYLEDONES, 987-988, 990, P. D. Hasselberg Government Printer, Wellington, New Zealand.
- Alsos, I. G., Engelskjøn, T., Gielly, L., Taberlet, P., Brochmann, C., (2005). Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Molecular Ecology*, 14, 2739–2753.
- Alvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, 29, (3), 417-434.
- Ansari, H. A., Ellison, N. W., and Williams, W. M. (2007). Molecular and cytogenetic evidence for an allotetraploid origin of *Trifolium dubium* (Leguminosae). *Chromosoma*.
- Ansari, H. A., Ellison, N. W., Griffiths, G. a., and Williams, W. M. (2004). A lineage-specific centromeric satellite sequence in the genus *Trifolium*. *Chromosome Research*, 12, 357-367.

- Ansari, H. A., Ellison, N. W., Reader, S. M., Badaeva, E. D., Friebes, B., Miller, T. E. a., et al. (1999). Molecular Cytogenetic Organization of 5S and 18S-26S rDNA loci in white clover (*Trifolium repens* L.) and related species. *Annals of Botany*, 83, 199-206.
- Ascherson, P. a. G., P. (1906-1910). Synopsis der Mitterleuropaischern Flora. *Teil*, 2, 472-617.
- Atwood, S. S. (1942). Genetics of self-compatibility in *Trifolium repens*. *Journal of American Society of Agronomics*, 34, 353-364.
- Bao, Y., Wendel, J. F., and Ge, S. (2010). Multiple patterns of rDNA evolution following polyploidy in *Oryza*. *Molecular Phylogenetics and Evolution* 55, 136-142.
- Bauert, M. R., Källin, M., Baltisberger, M., and Edwards, P. J., (1998). No genetic variation detected within isolated relict populations of *Saxifraga cernua* in the Alps using RAPD markers. *Molecular Ecology*, 7, 1519–1527.
- Bennett, R. L., and Smith, A. G. (1991). Use of genomic clone for ribosomal RNA from *Brassica oleracea* in RLFP analysis of *Brassica* species. *Plant Molecular Biology*, 16, 685-688.
- Berg, R. Y., (1963). Disjunksjoner i Norges fjellflora og de teorier som er framsatt til forklaring av dem. *Blyttia* 21: 133–177. [In Norwegian with English summary.]
- Birky, W. C. (1995). Uniparental inheritance of mitochondrial and chloroplast genes: Mechanisms and evolution [Review]. *Proceedings of the National Academy of Sciences of the United States of America*, 92, 11331-11338.
- Bodlens, L. (1981). Agronomes andalous du Moyen-Age.
- Brochmann, C., Soltis, P. S., Soltis, D. E., (1992a), Recurrent formation and polyphylogeny of Nordic polyploids in *Draba* (Brassicaceae). *American Journal of Botany*, 79, 673–688.
- Brochmann, C., Stedje, B., and Borgen, L., (1992b). Gene flow across ploidal levels in *Draba* (Brassicaceae). *Evolutionary Trends in Plants*, 6, 125–134.
- Brochmann, C, Gabrielsen, T. M, Nordal, I, Landvik, J. Y, and Elven, R., (2003), Glacial survival or *tabula rasa*? The history of North Atlantic biota revisited, *Taxon*, 52, 417–450.
- Buckler, E. S., Ippolito, A., Holtsford, T. P., (1997). The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications, *Genetics*, 145, 821–832.

- Caradus, J. R., Forde., M. B., Wewala., S., and Mackay., A. C. (1990). Description and classification of a white clover (*Trifolium repens* L.) germplasm collection from southwest Europe. *New Zealand Journal of Agricultural Research*, 33, 367-375.
- Caranahan, H. L., Hill, H. D., Hanson, A. A., and Brown, K. G. (1955). Inheritance and frequencies of leaf markings in white clover. *Heredity*, 46, 109-114.
- Chen, Z. J., Pikaard, C. S., (1997), Transcriptional analysis of nucleolar dominance in polyploid plants: Biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*, *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, 94, 7, 3442-3447.
- Chen, J. Z., & Ni, Z. (2006). Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. [Review article]. *BioEssays*, 28, 240-252.
- Clement, M., Posada, D., and Crandall, K., (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, 9, 10, 1657-1660.
- Coombe, D. E. (1968). *Trifolium Flora Europaea*, 2, 157-172.
- Corkill, L. (1942). Cyanogenesis in white clover (*Trifolium repens* L.) V. The inheritance of cyanogenesis. *New Zealand Journal of Science Technology*, 23, 178-179.
- Corkill, L. (1971). Leaf markings in white clover. *Heredity*, 62(5), 307-310.
- Daday, H. (1965). Gene frequencies in wild populations of *Trifolium repens* L. 4 Mechanism of Natural Selection. *Heredity* 20, 355-365.
- Dadejova, M., Yoong Lim, K., Souckova-Skalicka, K., Matyasek, R., Grandbastien, M.-A., Leitch, A., et al. (2007). Transcription activity of rRNA genes correlates with a tendency intergenomic homogenization in *Nicotiana* allotetraploids. *New Phytologist*, 174, 658-668.
- Davies, W. E. (1963). *Leaf markings in Trifolium repens*: C. D . Darlington & A. D. Bradshaw.
- Davies, W. E., and Young, N. R. (1967). The characteristics of European, Mediterranean and other populations of white clover (*Trifolium repens* L.). *Euphytica*, 16, 33-340.

- Demesure, B., Comps, B., and Petit, R. J. (1996). Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L) in Europe, *EVOLUTION*, 50 (6), 2515-2520.
- Dhingra, A., and Folta, K. M. (2005). ASAP: Amplification, sequencing & annotation of plastomes. *BMC Genomics*, 6(176).
- Dover, G. A., and Tautz, D. (1986). Conservation and divergence in multigene families - alternatives to selection and drift *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY OF LONDON SERIES B-BIOLOGICAL SCIENCES*, 312(1154), 275-289.
- Eidesen, P. B., Alsos, I. G., Popp, M., Stensurd, Ø., Suda, J., and Brochmann, C., (2007). Nuclear vs. plastid data: complex Pleistocene history of a circumpolar key species, *Molecular Ecology*, 16, 3902–3925.
- Ellison, N. W., Liston, A., Steiner, J. J., Williams, W. M., and Taylor, N. L. (2006). Molecular phylogenetics of the clover genus (*Trifolium*-Leguminosae). *Molecular Phylogenetics and Evolution*, 39, 688-705.
- Eschmann-Grupe, G., Hurka, H., and Neuffer, B., (2004). Extent and structure of genetic variation in two *Diplotaxis* species (Brassicaceae) with contrasting breeding systems. *Plant Systematics and Evolution*, 244, 31–43.
- Frame, J., (1987). *THE ROLE OF WHITE CLOVER IN UNITED-KINGDOM PASTURES, OUTLOOK ON AGRICULTURE*, 16(1), 28-34.
- Franzke, A., and Mummenho, K. (1999). Recent hybrid speciation in *Cardamine* (*Brassicaceae*) conversion of nuclear ribosomal ITS sequences in *statu nascendi*, *Theoretical Applied Genetics*.
- Gaeta, R. T., and Pires, C. (2010). Homoeologous recombination in allopolyploids: the polyploid ratchet. [Research review]. *New Phytologist*, 186, 18-28.
- George, J., Sawbridge, T. I., Cogan, N. O. I., Gendall, A. R., Smith, K. F., Spangenberg, G. C., et al. (2008). Comparison of genome structure between white clover and *Medicago truncatula* supports homoeologous group nomenclature based on conserved synteny. *Genome*, 51, 905-911.
- Gillett, J. B. (1952). The genus *Trifolium* in Southern Arabia and in Africa south of the Sahara. *Kew Bulletin* (3), 367-404.
- Goldblatt, P. (1981). *Advanced Legume Systematics* 2, 427.

- Griffiths, A. J. F., Wessler, S. R., Lewontin, R. C., Gelbart, W. M., Suzuki, D. T., and Miller, J. H. (2004). Chapter 2: Patterns of inheritance, 27-72, in *Introduction to genetic analysis*, 8th Edition, W. H Freeman and Company, New York, USA.
- Gugerli, F., Sperisen, C., Büchler, U., *et al.* (2001). Haplotype variation in a mitochondrial tandem repeat of Norway spruce (*Picea abies*) populations suggests a serious effect during postglacial recolonization of the Western Alps. *Molecular Ecology*, 10, 1255–1263.
- Hampe, A, and Petit, R. J., (2005). Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, 8, 461–467.
- Hartl, D. L., and Clark, A. G. (1989). *Principles of population genetics, Second Edition* (2 ed.). Massachusetts, USA: Sinauer Associates, INC. Publishers.
- Hemingway, R. G., (1999). The effect of changing patterns of fertilizer applications on the major mineral composition of herbage in relation to the requirements of cattle: a 50-year review, *Animal Science*, 69, 1-18.
- Hewitt, G. M (2000). The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.
- Hewitt, G. M., Ibrahim, K.M. (2001). Inferring glacial refugia and historical migrations with molecular phylogenies. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J), 271–294, Blackwell Science, Oxford.
- Hovin, A. W., and Gibson, P. B. (1961). A red leaf marking in white clover *Journal of Heredity*, 52, 295-296.
- Hughes, P. D. (2010). Little Ice Age glaciers in the Balkans: low altitude glaciation enabled by cooler temperatures and local topoclimatic controls. *SURFACE PROCESSES AND LANDFORMS* 35(2), 229-241.
- Jaramillo-Correa, J. P., Grivet, D., Terrab, A., Kurt, Y., De-Lucas, A. I., Wahid, N., Vendramin, G. G., and Gonza Lez-Martínez, S. G., (2010). The Strait of Gibraltar as a major biogeographic barrier in Mediterranean conifers: a comparative phylogeographic survey *Molecular Ecology*, 19, 5452–5468.
- Joly, S., Rauscher, J. T., Sherman-broyles, S. I., Brown, A. H. D., and Doyle, J. J. (2004). Evolutionary dynamics and preferential expression of homoelogenous 18S-5.8S-26S nuclear ribosomal genes in natural and artificial *Glycine* allopolyploids. *Molecular Biology and Evolution* 21, 1409-1421.

- Kakes, P., and Chardonnens, A. N. (1994).
- Kakes, P., and Chardonnens, A. N. (2000). Cyanotypic frequencies in adjacent and mixed populations of *Trifolium occidentale* Coombe and *Trifolium repens* L. are regulated by different mechanisms. *Biochemical Systematics and Ecology* 28, 633-649.
- Kato, A., Lamb, J. C., Han, F., Vega, J., and Birchler, J. A. (2005). MOLECULAR ANALYSIS OF MAIZE CHROMOSOMES. *MAYDICA* 50 (3-4), 311-320.
- Kjaegaard, T. (2003). A plant that Changed the World: the rise and fall of clover 1000-2000. *Landscape Research*, 28(1), 41-49.
- Koch, M., and Al-Shehbaz, I. A., (2002). Molecular data indicate complex intra- and intercontinental differentiation of American *Draba* (Brassicaceae), *Annals Missouri Botanic Gardens*, 89, 88–109.
- Koch, M., and Bernhardt, K.-G., (2004). Comparative biogeography of the cytotypes of annual *Microthlaspi perfoliatum* (Brassicaceae) in Europe using isozymes and cpDNA data: refugia, diversity centres, and postglacial colonization, *American Journal of Botany*, 91, 115–124.
- Koch, M. A., Kiefer, C., Ehrich, D., *et al.* (2006). Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae), *Molecular Ecology*, 15, 825–839.
- Koh, J., Soltis., P. S., and Soltis., D. E. (2010). Homelog loss and expression changes in natural populations of the recently and repeatedly formed allotetraploid *Tragopogon mirus* (Asteraceae). *Genomics*, 11(97), 1-16.
- Komarov, V. L., *et al.* (1934-1964). *Flora URSS*, 1-30.
- Kovarik, A., Matyasek, R., Lim, K. Y., Skalicka, K., Koukalova, B., Knapp, S., *et al.* (2004). Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. *Biological Journal of the Linnean Society* 82, 615=625.
- Kovarik, A., Pries, J. C., Leitch, A. R., Lim, K. Y., Sherwood, A. M., Matyasek, R., *et al.* (2005). Rapid concerted evolution of Nuclear Ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics*, 169, 931-944.
- Kucera, J., Marhold, K., Lihova, J., (2010). *Cardamine maritima* group (*Brassicaceae*) in the amphi-Adriatic area: A hotspot of species diversity revealed by DNA sequences and morphological variation, *TAXON*, 59, (1), 148-164.

- Larena, G. B., Aguilar, F. J., and Feliner, N. G., (2002). Glacial-induced altitudinal migrations in *Armeria* (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing, *Molecular Ecology*, *11*, 1965–1974.
- Lichiterkonecki, U., Schlotter, M., and Konecki, D. S. (1994). DNA-SEQUENCE POLYMORPHISMS IN EXONIC AND INTRONIC REGIONS OF THE HUMAN PHENYLALANINE-HYDROXYLASE GENE AID IN THE IDENTIFICATION OF ALLELES. *HUMAN GENETICS*, *94*(3), 307-310.
- Lihová, J. A., Marhold, K., and Neuffer, B., (2000). Taxonomy of *Cardamine amara* (Cruciferae) in the Iberian Peninsula, *Taxon*, *49*, 4, Golden Jubilee Part 2, 747-763.
- Lihová, J. A., Kentaro, K., Shimizu, K. K., Marhold, K., (2006). Allopolyploid origin of *Cardamine asarifolia* (Brassicaceae): Incongruence between plastid and nuclear ribosomal DNA sequences solved by a single-copy nuclear gene, *Molecular Phylogenetics and Evolution*, *39*, 759–786.
- Liu, A. H., and Wang, J. B. (2006). Genomic evolution of Brassica allopolyploids revealed by ISSR marker. *Genetic Resources and Crop Evolution* *53*(3), 603-611.
- Marshall, A. H., Michaelson-Yeates, T. P. T., and Meredith, A. P. (1995). Reproductive characters of interspecific hybrids between *Trifolium repens* L. and *T. nigrescens* Viv. *Heredity*, *74*, 136-145.
- Martín, J. P., and Sañchez-Ye'lamo, M. D, (2000). Genetic relationships among species of the genus *Diplotaxis* (Brassicaceae) using inter-simple sequence repeat markers. *Theoretical Applied Genetics*, *101*, 1234–1241.
- Marhold, K., Lihová, J. A., Perny, M. A. N. A., and Bleeker, W., (2004). Comparative ITS and AFLP Analysis of Diploid *Cardamine* (Brassicaceae) Taxa from Closely Related Polyploid Complexes, *Annals of Botany*, *93*, 507- 520.
- Marhold, K., Lihová, J. A., (2006). Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae, *Plant Systematics and Evolution*, *259*, 143–174.
- Matyásek, R., Tate J. A, Lim Y. K, Srubarová H, Koh J, Leitch A. R, et al. (2007). Concerted evolution of rDNA in recently formed Tragopogon allotetraploids is typically associated with an inverse correlation between gene copy number and expression *Genetics*, *176*, 2509-2519.

- Matyásek, R., Tate, J. A., Lim, Y. K., Srubarova, H., Koh, J., et al., (2007). Concerted evolution of rDNA in recently formed Tragopogon allotetraploids is typically associated with an inverse correlation between gene copy number and expression, *Genetics*, 176, 2509-2519.
- Navashin, M. (1934). *CYTOLOGIA*, 5.
- Okane, S. L. (1996). The origins of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences *SYSTEMATIC BOTANY* 21 559-566.
- Parisod, C., and Besnard, G., (2007). Glacial *in situ* survival in the Western Alps and polytopic autopolyploidy in *Biscutella laevigata* L. (Brassicaceae), *Molecular Ecology*, 16.
- Petit, R. J., Brewer, S., Bordács, S *et al.* (2002). Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence, *Forest Ecology and Management*, 156, 49–79.
- Petit, R.J., Aguinagalde, I., De Beaulieu, J.L *et al.* (2003). Glacial refugia: hotspots but not melting pots of genetic diversity, *Science*, 300, 1563–1565.
- Poczai, P., and Hyvonen, J. (2010). Nuclear ribosomal spacer regions in plant phylogenetics: problems and prospects. *Molecular Biology Reporter*, 37, 1897-1912.
- Raffl, C., Holderegger, R., Parson, W., and Erschbamer, B. (2008). Patterns in genetic diversity of *Trifolium pallescens* populations do not reflect chronosequence on alpine glacier forelands. *Heredity*, 100 (5), 526-532.
- Rauscher, J. T., Doyle, J. J., and Brown, A. H. D. (2004). Multiple Origins and nrDNA Internal Transcribed Spacer Homeologue Evolution in the *Glycine tomentella* (Leguminosae) Allopolyploid Complex. *Genetics*, 166, 987-998.
- Rieseberg, L. H., Soltis, D. E., (1991). Phylogenetic consequences of cytoplasmic gene flow in plants, *Evolutionary Trends in Plants*, 5, 65–84.
- Salmon, A., and Ainouche, M. L. (2010). Polyploidy and DNA methylation: new tools available. *Molecular Ecology* 19(2), 213-215.
- Sang, T., Crawford, D. J., and Stuessy, T. F. (1995). Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer of nuclear ribosomal DNA: Implications for biogeography and concerted evolution *proceedings of the National Academy of Sciences of the United States of America Evolution*, 92, 6813-6817.

- Segarra-Moragues, G. J., Palop-Esteban, M., González-Candelas, F., and Catalán, P., (2007). Nunatak survival vs. tabula rasa in the Central Pyrenees: a study on the endemic plant species *Borderea pyrenaica* (Dioscoreaceae), *Journal of Biogeography*, *34*, 1893–1906.
- Soltis, D. E., Soltis, P. S., and Tate, J. A. (2003). Advances in the study of ployploidy since *Plant speciation*. [Research Review]. *New Phytologist*, *161*, 173–191.
- Song, K., Lu, P., Tang, K., and Osborn, T. C. (1995). Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 7719–7723.
- Stehlik, I. (2000). Nunatak and peripheral refugia for alpine plants during Quaternary glaciation in the middle parts of the Alps, *Botanica Helvetica*, *110*, 25–30.
- Stehlik, I., Schneller, J. J., and Bachmann, K., (2001) Resistance or emigration: Response of the high-alpine plant *Eritrichum nanum* (L.) Gaudin to the ice age within the European Alps. *Molecular Ecology*, *10*, 357–370.
- Stehlik, I., Blattner, F. R., Holderegger, R., and Bachmann, K., (2002) Nunatak survival of the high Alpine plant *Eritrichum nanum* (L.) Gaudin in the Central Alps during the ice ages. *Molecular Ecology*, *11*, 2027–2036.
- Symonds, V. V., Soltis, P. S., Soltis, D. E., (2010), Dynamics of Polyploid formation in *Tragopogon* (*Asteraceae*) recurrent formation, gene flow, and population structure, *Evolution*, *64*(7), 1984–2003.
- Taberlet, P., Gielly, L., Pautou, G., and Bouvet, J., (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, *17*, 1105–1109.
- Vizintin, L., Javornik, B., and Bohanec, B. (2006). Genetic characterization of selected *Trifolium* species as revealed by nuclear DNA content and ITS rDNA region analysis. *Plant Science*, *170*, 859–866.
- Volkov, R., Kostishin, S., Ehrendorfer, E., and Schweizer, D. (1996). Molecular organization and evolution of the external transcribed rDNA spacer region in two diploid relatives of *Nicotiana tabacum* (Solanaceae), *Plant Systematics and Evolution*, *201*, (1–4), 117–129.
- Volkov, R. A., Borisjuk, N. V., Panchuk, I. I., Schweizer, D., and Hemleben, V. (1999). Elimination and rearrangement of parental rDNA in the allotetraploid *Nicotiana tabacum*. *Molecular Biology and Evolution*, *16* (3), 311–320.

- Volkov, R. A., Komarova, N. Y., and Hemleben, V. (2007). Ribosomal DNA in plant hybrids: inheritance, rearrangement, expression. *Systematics and Biodiversity*, 5, (3), 261-276.
- Volkov, R. A., Komarova, N. Y., Panchuk, I. I., and Hemleben, V. (2003). Molecular evolution of rDNA external transcribed spacer and phylogeny of sect. *Petota* (genus *Solanum*). *Molecular phylogenetics and Evolution*, 29 (2), 187-202.
- Warwick, S. I., and Anderson, J. K., (1997a). Isozyme analysis of parentage in allopolyploid *Diplotaxis muralis* (L.) DC. (Brassicaceae). *Eucarpia Cruciferae, Newslett*, 19, 35–36.
- Watson, L. E., Sayed-Ahmed., H., and Badr., A. (2000). Molecular phylogeny of Old World *Trifolium* (Fabaceae) based on plastid and nuclear markers. *Plant Systematics and Evolution*, 224, 153-171.
- Wendel, J. F., Schnabel, A., and Seelanan, T. (1995). Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences of the United States of America*, 92, 280-284.
- Wendel, J. F., and Adams, K. L. (Eds.). (2005). *Genome evolution and gene silencing in polyploid cotton* (Vol. 143).
- Widmer, A., and Lexer, C., (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology and Evolution*, 16, 267–269.
- Williams, W. M. (1987a), Chapter 8 Adaptive variation pgs. 299-323, in, Williams, W. M. and Baker, M. J., (Eds.). (1987). *White Clover*. Oxon: C.A.B International Wallingford, Oxon, United Kingdom.
- Williams, W. M. (1987b), Chapter 9 White clover taxonomy and biosystematics, pgs 323-342 in, Williams, W. M. and Baker, M. J., (Eds.). (1987). *White Clover*. Oxon: C.A.B International Wallingford, Oxon, United Kingdom.
- Williams, W. M. (1987c), Chapter 10 Genetic and breeding pgs 343-419 in, Williams, W. M. and Baker, M. J., (Eds.). (1987). *White Clover*. Oxon: C.A.B International Wallingford, Oxon, United Kingdom.
- Williams, W. M. Chapter, 25 White Clover, Pgs 221-228, (1983) in , *Plant Breeding in New Zealand*, edditted by Wratt, G. S., and Smith, H. C. (Eds.). Butterworths of New Zealand (Ltd) in association with D.S.I.R.

- Williams, W. M., and Williamson, M. L. (1998). Genetic analysis of shikimate dehydrogenase allozymes in *Trifolium repens* L. *Theoretical and Applied Genetics*, 98, 859-868.
- Williams, W. M., and Williamson, M. L. (2001, B). Genetic polymorphism for cyanogenesis and linkage at the linamarase locus in *Trifolium nigrescens* Viv. subsp. *nigrescens*. *Theoretical and Applied Genetics*, 103, 1211-1215.
- Williams, W. M., Andari, H. A., Ellison, N. W., and Hussain, S. W. (2001). Evidence of three Subspecies in *Trifolium nigrescens* Viv. *Annals of Botany*, 87, 683-691.
- Willis, K. J., Niklas, K. J., (2004). The role of Quaternary environmental change in plant macroevolution: the exception or the rule? *Philosophical Transactions of the Royal Society of London Series B Biological Sciences*, 359, 159–172.
- Wissemann, V. (2007). Plant evolution by means of hybridization *Systematics and Biodiversity*, 5(3), 243-253.
- Wolfe, A., and Randle, C. P. (2004). Recombination, Heteroplasmy, Haplotype Polymorphism, and Paralogy in Plastid Genes: Implications for Plant Molecular Systematics. *Systematic Botany*, 29(4), 1011-1020.
- Zhang, Q. (2010). Why does biparental plastid inheritance revive in angiosperms? *Journal of Plant Research*, 123, 201-206.
- Zimmer, E. A., Martin, S. L., Beverley, S. M., Kan, Y. W., and Wilson, A. C. (1980). RAPID DUPLICATION AND LOSS OF GENES-CODING FOR THE ALPHA-CHAINS OF HEMOGLOBIN. *Proceedings of the National Academy of Sciences of the United States of America*, 77, 2158-2162.
- Zohary, M., and Heller, D. (1984). *The genus Trifolium* Jerusalem: The Israel Academy of Sciences and humanities.

Websites

Overton, M., *Agricultural Revolution in England 1500 - 1850* (19 September 2002), Cambridge University Press. ISBN 0-521-56859-5, from http://www.bbc.co.uk/history/british/empire_seapower/agricultural_revolution_01.shtml

TCS: Software program from David Posada's Lab <http://darwin.uvigo.es/>

Appendix A: Chloroplast pattern combination data

Table A1 Chloroplast pattern types in Zone 1

Ref Number	Accession Number	Country of origin	Place of origin	433	456	CAAAA insert
Zone 1	Area of origin Portugal					
152	C20477	Portugal	Águas de Moura, Palmela	T	C	-
154	C20477	Portugal	Águas de Moura, Palmela	C	C	+
157	C20477	Portugal	Águas de Moura, Palmela	C	C	-
153	C20477	Portugal	Águas de Moura, Palmela	C	C	+
155	C20477	Portugal	Águas de Moura, Palmela	C	C	-
156	C20477	Portugal	Águas de Moura, Palmela	T	C	-
133	C8985	Portugal	Coimbra	-	-	+
134	C8985	Portugal	Coimbra	C	C	+
3A	C8985	Portugal	Coimbra	C	C	+
137	C8985	Portugal	Coimbra	C	C	+
138	C8985	Portugal	Coimbra	C	C	+
139	C8985	Portugal	Coimbra	C	C	+
140	C8985	Portugal	Coimbra	C	C	+
141	C8985	Portugal	Coimbra	C	C	+
132	C8985	Portugal	Coimbra	C	C	+
12B	C8985	Portugal	Coimbra	C	C	+
13	C8985	Portugal	Coimbra	C	C	+
135	C8985	Portugal	Coimbra	C	C	+
136	C8985	Portugal	Coimbra	C	C	+
203	C8976	Portugal	Algarve	T	C	-
35	C8976	Portugal	Algarve	-	-	+
64	C8976	Portugal	Algarve	T	C	-
65	C8976	Portugal	Algarve	T	C	-
36	C8976	Portugal	Algarve	T	C	-
164	C8976	Portugal	Algarve	T	C	-
204 / 61	C8976	Portugal	Algarve	T	C	-
62	C8976	Portugal	Algarve	T	C	-
93	C7582	Portugal	Beira litoral	T	C	-
96	C7582	Portugal	Beira litoral	T	C	-
191	C7582	Portugal	Beira litoral	T	C	-
91	C7582	Portugal	Beira litoral	T	C	-
94	C7582	Portugal	Beira litoral	T	C	-
24	C7582	Portugal	Beira litoral	T	C	-
92	C7582	Portugal	Beira litoral	-	-	+
95	C7582	Portugal	Beira litoral	C	C	-
89	C7582	Portugal	Beira litoral	T	C	-
90	C7582	Portugal	Beira litoral	-	-	-
31/177	C7576	Portugal	Estremadura	C	C	+
32	C7576	Portugal	Estremadura	-	-	+
33	C7576	Portugal	Estremadura	C	C	+

2A	C7576	Portugal	Estremadura	C	C	+
10B	C7576	Portugal	Estremadura	C	C	+
11B	C7576	Portugal	Estremadura	C	C	+
34	C7576	Portugal	Estremadura	C	C	+
200	C7576	Morocco	Morocco	C	C	-
4-1	C9024	Portugal	Tras Os Montes, Esphino	A	C	-
4-10	C9024	Portugal	Tras Os Montes, Esphino	C	C	+
4-3	C9024	Portugal	Tras Os Montes, Esphino	T	C	-
4-4	C9024	Portugal	Tras Os Montes, Esphino	T	C	-
4-5	C9024	Portugal	Tras Os Montes, Esphino	T	C	-
4-6	C9024	Portugal	Tras Os Montes, Esphino	-	-	+
4-7	C9024	Portugal	Tras Os Montes, Esphino	-	-	-
4-8	C9024	Portugal	Tras Os Montes, Esphino	T	C	-
4-9	C9024	Portugal	Tras Os Montes, Esphino	-	-	-
128	C9514	Portugal	Trans Os Montes	C	C	+
124	C9514	Portugal	Trans Os Montes	C	C	+
99	C9514	Portugal	Trans Os Montes	T	C	-
125	C9514	Portugal	Trans Os Montes	C	C	+
126	C9514	Portugal	Trans Os Montes	C	C	+
127	C9514	Portugal	Trans Os Montes	C	C	+
129	C9514	Portugal	Trans Os Montes	-	-	+
130	C9514	Portugal	Trans Os Montes	C	C	+
8A	C9514	Portugal	Trans Os Montes	C	C	-
186	C8979	Portugal	Ribatejo, Grândola	T	C	-
187	C8979	Portugal	Ribatejo, Grândola	T	C	-
196	C8979	Portugal	Ribatejo, Grândola	T	C	-
188	C8979	Portugal	Ribatejo, Grândola	C	C	+
4B / 195	C8979	Portugal	Ribatejo, Grândola	C	C	+
45	C8979	Portugal	Ribatejo, Grândola	T	C	-
46	C8979	Portugal	Ribatejo, Grândola	T	C	-
47	C8979	Portugal	Ribatejo, Grândola	C	C	-
175	C8979	Portugal	Ribatejo, Grândola	T	C	-
48	C8979	Portugal	Ribatejo, Grândola	T	C	-
6-1	C9516	Portugal	Tras Os Montes, Vila Verde	C	T	+
6-9	C9516	Portugal	Tras Os Montes, Vila Verde	-	-	+
6-10	C9516	Portugal	Tras Os Montes, Vila Verde	T	C	/
6-2	C9516	Portugal	Tras Os Montes, Vila Verde	-	-	/
6-3	C9516	Portugal	Tras Os Montes, Vila Verde	C	T	/
6-4	C9516	Portugal	Tras Os Montes, Vila Verde	-	-	+
6-5	C9516	Portugal	Tras Os Montes, Vila Verde	C	T	/
6-6	C9516	Portugal	Tras Os Montes, Vila Verde	C	T	/
6-7	C9516	Portugal	Tras Os Montes, Vila Verde	C	T	/
6-8	C9516	Portugal	Tras Os Montes, Vila Verde	C	T	+
8-5	AZ4838	Portugal	Portugal	-	-	-
8-3	AZ4838	Portugal	Portugal	T	C	/
8-1	AZ4838	Portugal	Portugal	T	C	-
8-4	AZ4838	Portugal	Portugal	C	C	/
9-2	AZ4839	Portugal	Portugal	C	C	+

9-5	AZ4839		Portugal	C	C	+
9-1	AZ4839		Portugal	T	C	-
9-4	AZ4839		Portugal	T	C	-
9-6	AZ4839		Portugal	T	C	-
Zone 1: Spain						
28	C8975	Spain	Andalucía	C	C	+
27	C8975	Spain	Andalucía	-	-	+
26	C8975	Spain	Andalucía	C	C	+
3B	C8975	Spain	Andalucía	C	C	+
52	C8975	Spain	Andalucía	C	C	+
7-10	C9531	Spain	Leon & Castile, Ponferrada	-	-	-
7-1	C9531	Spain	Leon Castile Leon, Ponferrada	T	C	-
7-2	C9531	Spain	Leon Castile Leon, Ponferrada	C	C	+
7-3	C9531	Spain	Leon Castile Leon, Ponferrada	C	C	-
7-4	C9531	Spain	Leon Castile Leon, Ponferrada	T	C	-
7-5	C9531	Spain	Leon Castile Leon, Ponferrada	T	C	-
7-6	C9531	Spain	Leon Castile Leon, Ponferrada	T	C	-
7-7	C9531	Spain	Leon Castile Leon, Ponferrada	C	C	+
7-8	C9531	Spain	Leon Castile Leon, Ponferrada	C	C	+
7-9	C9531	Spain	Ponferrada	T	C	-
21	C7600	Spain	Pyrenees Mountains	C	C	-
117	C7600	Spain	Pyrenees Mountains	C	C	+
118	C7600	Spain	Pyrenees Mountains	C	C	+
9A	C7600	Spain	Pyrenees Mountains	C	C	+
20	C7600	Spain	Pyrenees Mountains	C	C	-
119	C7600	Spain	Pyrenees Mountains	C	C	+
120	C7600	Spain	Pyrenees Mountains	C	C	+
145	C7600	Spain	Pyrenees Mountains	C	C	+
122	C7600	Spain	Pyrenees Mountains	T	C	-
123	C7600	Spain	Pyrenees Mountains	C	C	-
106	C8994	Spain	Asturias	T	C	-
111	C8994	Spain	Asturias	T	C	-
189	C8994	Spain	Asturias	T	C	-
199	C8994	Spain	Asturias	T	C	-
110/ 210	C8994	Spain	Asturias	T	C	-
113/ 190	C8994	Spain	Asturias	T	C	-
114/ 211	C8994	Spain	Asturias	T	C	-
22/107	C8994	Spain	Asturias	T	C	-
11A	C8994	Spain	Asturias	T	C	-
108	C8994	Spain	Asturias	-	-	-
112	C8994	Spain	Asturias	T	C	-
115	C8994	Spain	Asturias	T	C	-
116	C8994	Spain	Asturias	T	C	-

109/ 209	C8994	Spain	Asturias	T	C	+
74	C9025	Spain	Extremadura	C	C	+
77 / 206	C9025	Spain	Extremadura	C	C	+
76	C9025	Spain	Extremadura	-	-	+
49	C9025	Spain	Extremadura	C	C	
25	C9025	Spain	Extremadura	T	C	-
50	C9025	Spain	Extremadura	T	C	-
51	C9025	Spain	Extremadura	T	C	-
73	C9025	Spain	Extremadura	C	C	+
75	C9025	Spain	Extremadura	C	C	+
66	C9034	Spain	Galicia	T	C	-
71	C9034	Spain	Galicia	T	C	-
178	C9034	Spain	Galicia	T	C	-
88	C9034	Spain	Galicia	C	T	+
161	C9037	Spain	Galicia	T	C	-
193	C9037	Spain	Galicia	T	C	-
194	C9037	Spain	Galicia	T	C	-
55	C9034	Spain	Galicia	C	T	+
67	C9034	Spain	Galicia	T	C	-
68	C9034	Spain	Galicia	C	C	-
69 / 205	C9034	Spain	Galicia	T	C	-
70	C9034	Spain	Galicia	T	C	-
72	C9034	Spain	Galicia	T	C	-
57	C9034	Spain	Galicia	T	C	-
58	C9034	Spain	Galicia	T	C	/
158	C9034	Spain	Galicia	T	C	-
1B	C9037	Spain	Galicia	-	-	-
30/ 201	C9037	Spain	Galicia	T	C	-
12A	C9037	Spain	Galicia	T	C	-
159	C9037	Spain	Galicia	T	C	-
160	C9037	Spain	Galicia	T	C	-
29	C9037	Spain	Galicia	T	C	-
5-10	C9027	Spain	La Estrada, Galicia, Pontevedra	Y	C	+
5-2	C9027	Spain	La Estrada, Galicia, Pontevedra	C	C	+
5-3	C9027	Spain	La Estrada, Galicia, Pontevedra	-	-	+
5-4	C9027	Spain	La Estrada, Galicia, Pontevedra	T	C	-
5-5	C9027	Spain	La Estrada, Galicia, Pontevedra	C	C	-
5-6	C9027	Spain	La Estrada, Galicia, Pontevedra	T	C	-
5-8	C9027	Spain	La Estrada, Galicia, Pontevedra	T	C	-
5-9	C9027	Spain	La Estrada, Galicia, Pontevedra	T	C	-

Table A2 Chloroplast pattern types in North Africa

Ref Number	Accession Number	Place of origin	Country of origin	Positions		
1A	C20499	Algeria	Algeria	C	C	+
322	C6452	Algeria	Algeria	C	C	+
323	C6452	Algeria	Algeria	C	C	+
3-10	C6452	Algeria	Algeria	C	C	+
3-1	C6452	Algeria	Algeria	C	C	+
3-3	C6452	Algeria	Algeria	C	C	+
3-4	C6452	Algeria	Algeria	C	C	-
3-5	C6452	Algeria	Algeria	C	C	+
3-6	C6452	Algeria	Algeria	C	C	+
3-7	C6452	Algeria	Algeria	C	C	+
3-8	C6452	Algeria	Algeria	C	C	+
3-9	C6452	Algeria	Algeria	C	C	+
A1	C 1751	Algeria	Algeria	C	C	+
A10	C 1751	Algeria	Algeria	C	C	+
A2	C 1751	Algeria	Algeria	C	C	+
A3	C 1751	Algeria	Algeria	C	C	+
A4	C 1751	Algeria	Algeria	C	C	+
A5	C 1751	Algeria	Algeria	C	C	+
A6	C 1751	Algeria	Algeria	C	C	+
A7	C 1751	Algeria	Algeria	C	C	+
A8	C 1751	Algeria	Algeria	C	C	+
A9	C 1751	Algeria	Algeria	C	C	+
170	C1761	Morocco	Morocco	C	C	+
165	C1761	Morocco	Morocco	C	C	+
171	C1761	Morocco	Morocco	C	C	+
166	C1761	Morocco	Morocco	C	C	+
167	C1761	Morocco	Morocco	C	C	+
168	C1761	Morocco	Morocco	C	C	+
169	C1761	Morocco	Morocco	C	C	+

Table A3 Chloroplast pattern types in Zone 2: Israel

Ref No.	Accession Number	Place of origin	Country of origin	Commercial Name	Positions		CAAAA insert
					433	456	
Zone 2 Near East, Israel							
381	C5892	Huleh district	Israel	Haifa	C	C	-
321	C5892	Huleh district	Israel	Haifa	C	C	+
345	C5892	Huleh district	Israel	Haifa	C	C	+
346	C5892	Huleh district	Israel	Haifa	C	C	+
347	C5892	Huleh district	Israel	Haifa	C	C	+
348	C5892	Huleh district	Israel	Haifa	C	C	+
324	C4143	Israel	Israel	Tamar	C	C	+
325	C4143	Israel	Israel	Tamar	C	C	+
326	C4143	Israel	Israel	Tamar	C	C	+
327	C4143	Israel	Israel	Tamar	C	C	+
328	C4143	Israel	Israel	Tamar	C	C	+
329	C4143	Israel	Israel	Tamar	C	C	+
330	C4143	Israel	Israel	Tamar	C	C	+
331	C4143	Israel	Israel	Tamar	C	C	+
332	C4143	Israel	Israel	Tamar	C	C	+
333	C4143	Israel	Israel	Tamar	C	C	-
370	C4143	Israel	Israel	Tamar	C	C	-
393	C4143	Israel	Israel	Tamar	C	C	+

Table A4 Chloroplast pattern types in Zone 2 Europe

Ref No.	Accession Number	Place of origin	Ref Number	Accession Number	Positions		CAAAA insert
					433	456	
Zone 2 Europe							
84	C6374	England	84	C6374	C	C	+
197	C6374	England	197	C6374	C	C	+
198	C6374	England	Kent Suffolk	Kent	C	T	+
87	C6374	England	Kent Suffolk	Kent	C	C	+
8B	C6374	England	Kent Suffolk	Kent	C	C	+
81	C6374	England	Kent Suffolk	Kent	C	C	+
82	C6374	England	Kent Suffolk	Kent	C	C	+
85	C6374	England	Kent Suffolk	Kent	T	C	-
86	C6374	England	Kent Suffolk	Kent	C	C	-
2-Oct	C6304	England	Kent Suffolk	Kent	C	T	+
2-1	C6304	England	Kent Suffolk	Kent	-	-	+
2-3	C6304	England	Kent Suffolk	Kent	C	C	+
2-4	C6304	England	Kent Suffolk	Kent	C	C	+
2-5	C6304	England	Kent Suffolk	Kent	C	C	-
2-6	C6304	England	Kent Suffolk	Kent	C	T	+
2-7	C6304	England	Kent Suffolk	Kent	-	-	+
2-8	C6304	England	Kent Suffolk	Kent	-	-	+
2-2	C6304	England	Kent Suffolk	Kent	C	C	+
2-9	C6304	England	Kent Suffolk	Kent	C	T	+

308	C21105	Britain	Wild British populations	S. 184	-	-	+
309	C21105	Britain	Wild British populations	S. 184	C	T	+
334	C2095	Denmark	Denmark	Morso	-	-	+
335	C2095	Denmark	Denmark	Morso	C	C	+
336	C2095	Denmark	Denmark	Morso	C	C	+
337	C2095	Denmark	Denmark	Morso	-	-	+
338	C2095	Denmark	Denmark	Morso	C	C	+
339	C2095	Denmark	Denmark	Morso	C	C	+
340	C2095	Denmark	Denmark	Morso	C	C	+
341	C2095	Denmark	Denmark	Morso	C	C	+
342	C2095	Denmark	Denmark	Morso	C	C	+
343	C2095	Denmark	Denmark	Morso	C	C	+
344	C2095	Denmark	Denmark	Morso	C	C	+
377	C2095	Denmark	Denmark	Morso	C	C	+
378	C2095	Denmark	Denmark	Morso	C	C	+
300	C24247	Finland	Finland	Tammisto	C	C	+
301	C24247	Finland	Finland	Tammisto	C	C	+
302	C24247	Finland	Finland	Tammisto	C	C	+
303	C24247	Finland	Finland	Tammisto	C	C	+
304	C24247	Finland	Finland	Tammisto	C	C	+
305	C24247	Finland	Finland	Tammisto	C	C	+
306	C24247	Finland	Finland	Tammisto	C	C	+
307	C24247	Finland	Finland	Tammisto	C	C	+
369	C24247	Finland	Finland	Tammisto	C	C	+
310	C960	The Netherlands	The Netherlands	Barbian	C	C	+
311	C960	The Netherlands	The Netherlands	Barbian	C	C	+
312	C960	The Netherlands	The Netherlands	Barbian	C	C	+
313	C960	The Netherlands	The Netherlands	Barbian	-	-	+
314	C960	The Netherlands	The Netherlands	Barbian	C	T	+
315	C960	The Netherlands	The Netherlands	Barbian	C	C	+
316	C960	The Netherlands	The Netherlands	Barbian	C	T	+
317	C960	The Netherlands	The Netherlands	Barbian	C	C	+
349	C960	The Netherlands	The Netherlands	Barbian	C	C	+
373	C960	The Netherlands	The Netherlands	Barbian	C	Y	+
374	C960	The Netherlands	The Netherlands	Barbian	C	C	+
360	C7555	Czech Republic	Czech Republic	Ovack	C	C	+
361	C7555	Czech Republic	Czech Republic	Ovack	C	C	+
362	C7555	Czech Republic	Czech Republic	Ovack	C	T	+
363	C7555	Czech Republic	Czech Republic	Ovack	C	C	+
364	C7555	Czech Republic	Czech Republic	Ovack	C	C	+
365	C7555	Czech Republic	Czech Republic	Ovack	C	C	+

367	C7555	Czech Republic	Czech Republic	Ovack	C	C	-
376	C6360	Western Europe	Belgium Kersey mix	Blanca	-	-	-
350	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
351	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
352	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
353	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
354	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
355	C6360	Western Europe	Belgium Kersey mix	Blanca	-	-	+
356	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
357	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
358	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
359	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+
375	C6360	Western Europe	Belgium Kersey mix	Blanca	C	C	+

Table A5 Chloroplast pattern types in Zone 2: in plants French origins

Ref No.	Accession No.	Country of origin	Place of origin	Commercial Name	Positions		CAAAA insert
					433	456	
Zone 2 France							
163	C9046	France	Mouriès		C	C	+
192	C9046	France	Mouriès		C	C	+
44/							
202	C9046	France	Mouriès		C	C	+
162	C9046	France	Mouriès		C	C	+
40	C9046	France	Mouriès		C	C	+
41	C9046	France	Mouriès		C	C	-
7A							
/2b	C9046	France	Mouriès		C	C	-
42	C9046	France	Mouriès		C	C	-
43	C9046	France	Mouriès		C	C	+
54	C7690	France	Bouches du Rhone, Entressen		C	T	+
78	C7690	France	Bouches du Rhone, Entressen		C	T	+
207	C7690	France	Bouches du Rhone, Entressen		C	T	+
53	C7690	France	Bouches du Rhone, Entressen		C	C	+
80	C7690	France	Bouches du Rhone, Entressen		-	-	+
			Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
121	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
150	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
151	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
4A	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
143	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+

144	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
131	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
146	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
147	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
148	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
149	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	+
R1	-	France/Israel	France & Israel cross	Aran	C	C	+
R10	-	France/Israel	France & Israel cross	Aran	C	T	+
R2	-	France/Israel	France & Israel cross	Aran	C	T	+
R3	-	France/Israel	France & Israel cross	Aran	C	T	+
R4	-	France/Israel	France & Israel cross	Aran	C	C	+
R5	-	France/Israel	France & Israel cross	Aran	C	C	+
R6	-	France/Israel	France & Israel cross	Aran	C	T	+
R7	-	France/Israel	France & Israel cross	Aran	C	C	+
R8	-	France/Israel	France & Israel cross	Aran	C	C	+
R9	-	France/Israel	France & Israel cross	Aran	C	C	+
C1	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	C	+
C10	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	C	+
C2	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	T	+
C3	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	Y	+
C4	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	T	+

C5	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	C	+
C6	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	-	-	+
C7	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	C	+
C8	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	C	+
C9	-	France/Syria	Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	T	+

Table A6 Chloroplast pattern types in Zone 2 for Italian Lodi and ecotype Selection Regal

Ref Number	Accession Number	Place of origin		Commercial	Positions		
		origin	Country of origin		Name	433	456
Zone 2 Italian Ecotype Lodi and USA Ecotype Selection Regal							
105	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	?
103	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-2	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-10	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-1	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
17	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
18	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
100	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-4	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-5	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-6	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
1-7	C5897	Italy	Milan, Po Valley	Italian Lodi	-	-	+
1-8	C5897	Italy	Milan, Po Valley	Italian Lodi	C	Y	+
1-9	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
104	C5897	Italy	Milan, Po Valley	Italian Lodi	C	C	+
391	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
390	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
385	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
383	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
388	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
386	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+

389	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
387	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+
384	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	C	+

Table A Zone 3 :South African Cultivar Dusi

Ref Number	Accession Number	Country of origin	Place of origin	Commercial Name	433	456	CAAAA insert
Zone 3, outer populations							
319	C16511	South Africa	Introduced mixed populations	Dusi	-	-	+
320	C16511	South Africa	Introduced mixed populations	Dusi	-	-	+
371	C16511	South Africa	Introduced mixed populations	Dusi	C	C	+
372	C16511	South Africa	Introduced mixed populations	Dusi	C	C	+
318	C16511	South Africa	Introduced mixed populations	Dusi	C	C	+

1. Appendix B Nuclear ITS sequence data

Table B1 Zone 1 Portugal and Spain, ITS pattern types

Ref No.	Accession number	Place of origin	Positions of variability									ITS Type
			50, T/a	62, C/T	111, G/a	21, G	486, C/T	509, T/C	57, 1, C/a /m	61, 3, T	65, 4, C	
Zone 1 area of origin, Portugal												
31/177	C7576	Estremadura	T	C	G	G	C	T	C	T	C	15
94	C7582	Beira litoral	T	C	G	G	C	C	C	T	C	5
95	C7582	Beira litoral	T	C	G	G	C	C	C	T	C	5
191	C7582	Beira litoral	T	C	G	G	C	C	C	T	C	5
24	C7582	Beira litoral	T	C	G	G	C	C	C	T	C	5
92	C7582	Beira litoral	/	C	G	G	C	C	C	T	C	5?
93	C7582	Beira litoral	T	C	G	G	C	Y	C	T	C	20
89	C7582	Beira litoral	T	C	G	G	C	C	C	T	C	5
96	C7582	Beira litoral	T	C	G	G	C	Y'	C	T	C	20'
N4	C7648	N. region	T	C	G	G	C	C	C	T	C	5
N6	C7648	N. region	A	C	A	G	C	C	C	T	C	1
N2	C7648	N. region	W	C	A	G	C	Y	C	T	C	28
N3	C7648	N. region	W	C	A	G	C	Y	C	T	C	28
N7	C7648	N. region	W	C	A	G	C	Y	C	T	C	28
203	C8976	Algarve	T	C	A	G	C	C	C	T	C	2
65	C8976	Algarve	T	C	G	G	C	C	C	T	C	5
204/ 61	C8976	Algarve	T	C	G	G	C	T	C	T	C	15
188	C8979	Ribatejo, Grândola	T	C	A	G	C	C	C	T	C	2
47	C8979	Ribatejo, Grândola	T	C	G	G	C	C	C	T	C	5

187	C8979	Ribatejo, Grândola	T	C	G	G	C	C	C	T	C	5
196	C8979	Ribatejo, Grândola	T	C	G	G	C	C	C	T	C	5
186	C8979	Ribatejo, Grândola	T	C	G	G	C	T	C	T	C	15
48	C8979	Ribatejo, Grândola	T	C	G	G	C	Y	C	T	C	20
140	C8985	Coimbra	T	C	G	G	C	T	A	T	C	17
139	C8985	Coimbra	T	C	G	G	C	C	C	T	C	5
141	C8985	Coimbra	T	C	G	G	C	C	C	T	C	5
134	C8985	Coimbra	T	C	G	G	C	T	C	T	C	15
4-6	C9024	Esphino	T	C	G	G	C	C	C	T	C	5
4-9	C9024	Esphino	T	C	G	G	C	C	C	T	C	5
4-10	C9024	Esphino	N	C	G	G	C	C	C	T	C	5C
4-4	C9024	Esphino	T	C	G	G	C	Y	C	T	C	20
4-3	C9024	Esphino	T	C	G	G	C	Y	C	T	C	20
4-7	C9024	Esphino	T	C	G	G	C	Y	C	T	C	20
4-5	C9024	Esphino	T	C	G	G	C	C	C	T	C	5
N49	C9024	Esphino	T	C	G	G	C	Y	C	T	C	20
N51	C9024	Esphino	T	C	G	G	C	Y	C	T	C	20
124	C9514	Trans Os Montes	T	C	G	G	C	C	C	T	C	5
127	C9514	Trans Os Montes	T	C	G	G	Y	C	C	T	C	32
130	C9514	Trans Os Montes	T	C	G	G	Y	C	C	T	C	32
99	C9514	Trans Os Montes	T	C	G	G	C	T	C	T	C	15
6-1	C9516	Vila Verde	T	C	G	G	C	T	C	T	C	15
6-2	C9516	Vila Verde	/	/	G	G	C	Y'	C	T	C	20'?
6-3	C9516	Vila Verde	T	C	G	G	C	C	C	T	C	5
6-4	C9516	Vila Verde	T	C	G	G	C	Y	C	T	C	20
6-5	C9516	Vila Verde	T	C	G	G	C	Y'	C	T	C	15
6-6	C9516	Vila Verde	T	C	G	G	C	Y	C	T	C	20'?
6-7	C9516	Vila Verde	T	C	G	G	C	Y	C	T	C	5
6-8	C9516	Vila Verde	T	C	G	G	C	Y	C	T	C	20
6-9	C9516	Vila Verde	T	C	G	G	C	T	C	T	C	20'

6-10	C9516	Vila Verde	T	C	G	G	C	Y	C	T	C	20
N56	C9516	Vila Verde	T	C	G	G	C	C	C	T	C	20
N57	C9516	Vila Verde	T	C	G	G	C	T	C	T	C	20
N55	C9516	Vila Verde	T	C	G	G	C	Y	C	T	C	15
N10	C11519	Vila Pouca de Agu	T	C	G	G	C	C	C	T	C	20
N11	C11519	Vila Pouca de Agu	T	C	G	G	C	C	C	T	C	5
N12	C11519	Vila Pouca de Agu	T	C	G	G	C	Y	C	T	C	15
153	C20477	Águas de Moura, Palmela	T	C	A	G	C	C	C	T	C	20
157	C20477	Águas de Moura, Palmela	T	C	A	G	C	C	C	C	T	5
152	C20477	Águas de Moura, Palmela	T	C	G	G	C	C	C	T	C	5
154	C20477	Águas de Moura, Palmela	T	C	R	C	C	T	C	T	C	20
8-5	AZ4838	Minho, Braga, Ferreiras	T	C	G	G	C	C	C	T	C	2
8-4	AZ4838	Minho, Ferreira	T	T	G	G	C	C	C	T	C	3
8-1	AZ4838	Minho, Ferreira	T	C	G	G	C	T	C	T	C	5
8-3	AZ4838	Minho, Ferreira	T	C	G	G	C	Y	C	T	C	45
9-5	AZ4839	Tras Montes, Baragonça	T	C	G	G	C	C	C	T	C	5
9-4	AZ4839	Tras Montes, Baragonça	T	C	G	G	C	C	C	T	C	7
9-6	AZ4839	Tras Montes, Baragonça	-	C	G	G	C	C	C	T	C	15
9-1	AZ4839	Tras Montes, Baragonça	T	C	N	G	C	C	C	T	C	20
9-2	AZ4839	Tras Montes, Baragonça	T	C	G	G	C	Y	C	T	C	5
N78	P	Portugal	T	C	G	G	C	C	C	T	C	5

Zone 1, Spain												
21	C7600	Pyrenees	N/a (A)	C	G	G	T	C	C	T	C	20'
117	C7600	Pyrenees	T	C	G	G	C	C	C	T	C	5
20	C7600	Pyrenees	T	C	G	G	C	T	C	T	C	5
52	C8975	Andalucía	T	C	G	G	C	T	C	T	C	
199	C8994	Asturias	-	C	G	G	C	C	C	T	C	9
11A	C8994	Asturias	T	C	G	G	C	C	C	T	C	5
106	C8994	Asturias	T	C	G	G	C	T	C	T	C	15'
110/ 210	C8994	Asturias	T	C	G	G	C	T	C	T	C	15
113/ 190	C8994	Asturias	T	C	G	G	C	T	C	T	C	5?
114/ 211	C8994	Asturias	T	C	G	G	C	T	C	T	C	5
22/107	C8994	Asturias	T	C	G	G	C	T	C	T	C	15
108	C8994	Asturias	T	C	G	G	C	Y	C	T	C	15
189	C8994	Asturias	T	C	G	G	C	Y'	C	T	C	15
74	C9025	Extremadura	T	C	G	G	C	T	C	T	C	15
75	C9025	Extremadura	T	C	G	G	C	T	C	T	C	15
77/206	C9025	Extremadura	T	C	G	G	C	T	C	T	C	20
76	C9025	Extremadura	T	C	G	G	C	Y	C	T	C	20'
N52	C9027	La Estrada, Galicia, Pontevedra	T	C	G	G	C	T	C	T	C	15
N54	C9027	La Estrada, Galicia, Pontevedra	W	C	G	G	C	T	C	T	C	15
N53	C9027	La Estrada, Galicia, Pontevedra	T	C	G	G	C	C	C	T	C	15
55	C9034	Galicia	T	C	G	G	C	C	C	T	C	20
68	C9034	Galicia	T	C	G	G	C	C	C	T	C	15?
58	C9034	Galicia	T	C	G	G	C	T	C	T	C	15C?
71	C9034	Galicia	T	C	G	G	C	T	C	T	C	5
88	C9034	Galicia	T	C	G	G	C	T	C	T	C	5
66/145	C9034	Galicia	T	C	G	G	C	T	C	T	C	5

69/205	C9034	Galicia	T	C	G	G	C	T	C	T	C	15
72	C9034	Galicia	T	C	G	G	C	T	C	C	T	15
67	C9034	Galicia	T	C	G	G	C	T	C	T	C	15
178	C9034	Galicia	T	C	G	G	C	T	C	T	C	15
70	C9034	Galicia	T	C	G	G	C	Y	C	T	C	15
56	C9034	Galicia	T	C	G	G	C	Y	C	T	C	26
193	C9037	Galicia	T	C	G	G	C	C	C	T	C	15'
161	C9037	Galicia	T	C	G	G	C	T	C	T	C	15'
194	C9037	Galicia	T	C	G	G	C	T	C	T	C	20
30/ 201	C9037	Galicia	T	C	G	G	C	Y	C	T	C	20
160	C9037	Galicia	T	C	G	G	C	Y	C	T	C	5
7-3	C9531	Ponferrada, León & Castile	T	C	G	G	C	T	C	T	C	15
7-4	C9531	Ponferrada, León & Castile	T	C	G	G	C	T	C	T	C	15'
7-5	C9531	Ponferrada, León & Castile	/	C	G	G	C	T	C	T	C	20
7-6	C9531	Ponferrada, León & Castile	T	C	G	G	C	Y	C	T	C	20'
7-7	C9531	Ponferrada, León & Castile	/	C	G	G	C	T	C	T	C	15
7-9	C9531	Ponferrada	T	C	G	G	C	Y	C	T	C	15
N58	C9531	Ponferrada	W	C	G	G	C	C	C	T	C	15?
N59	C9531	Ponferrada	W	C	G	G	C	C	C	T	C	20
N60	C9531	Ponferrada	W	C	G	G	C	C	C	T	C	15?
N104	TR13	Spain unknown	T	C	G	G	C	T	C	T	C	20

Table B2 Zone 2 ITS pattern types in North Africa

Ref No.	Accession number	Place of origin	Positions of variability								ITS Type
			50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	
Zone 2 N. Africa											
322	C6452	Algeria	T	T	G	G	C	C	C	T	7
323	C6452	Algeria	T	T	G	G	C	C	C	T	7
N45	C6452	Algeria	A	C	G	G	C	C	C	T	12
N47	C6452	Algeria	T	T	G	G	C	C	C	T	7
3-3	C6452	Algeria	T	C	G	G	C	T	C	T	15
3-5	C6452	Algeria	/	/	G	G	T	C	C	T	9
3-6	C6452	Algeria	T	N	G	G	C	C	C	T	5D
3-8	C6452	Algeria	T	Y	G	G	C	C	C	T	30
3-9	C6452	Algeria	T	T	G	G	C	C	C	T	7
3-10	C6452	Algeria	T	C	G	G	C	C	C	T	5
A1	C1751	Algeria	T	C	G	G	C	Y	C	T	11
A2	C1751	Algeria		T	C	G	G	C	C	C	5
A3	C1751	Algeria		A	C	G	G	C	C	C	12
A4	C1751	Algeria		T	T	G	G	C	Y	C	38
A5	C1751	Algeria		T	Y	G	G	C	C	C	30
A6	C1751	Algeria		T	Y	G	G	C	C	C	30
A7	C1751	Algeria		T	Y	G	G	C	C	C	30
A8	C1751	Algeria		T	Y	G	G	C	C	C	30
A10	C1751	Algeria		-	Y	G	G	C	C	C	30
Morocco											
167	C1761	Morocco		T	C	A	G	C	Y	C	25
168	C1761	Morocco		T	C	G	G	C	T	C	15
169	C1761	Morocco		T	C	G	G	C	T	C	15

170	C1761	Morocco	T	C	G	G	C	T	C	15
N20	C25464	Marrakech	T	C	G	G	Y	Y	C	39
N21	C25465	Chefchaouen	A	C	A	G	C	C	C	1
N22	C25466	Bab-Berred	T	C	G	G	C	T	C	15
N23	C25467	Ketama	T	C	G	G	Y	T	C	35
N24	C25468	South of Taza	T	C	G	G	C	T	C	15
N25	C25469	South of Azrou	T	C	G	G	C	T	C	15
N26	C25470	Ain-Leuh, Ifrane	T	C	G	G	C	T	C	15
N27	C25471	Oukaimeden	T	C	G	G	C	T	C	15
N28	C25472	South of Bou-Laouane	T	C	G	G	C	T	C	15
N29	C25473	South of Asni	T	C	G	G	Y	T	C	35
N30	C25474	North of Asni	T	C	G	G	Y	T	C	35
N31	C25475	North of Irhem	T	C	G	G	Y	Y	C	39
N32	C25476	Taddert	T	C	G	G	Y	T	C	35
N33	C25477	Taddert	A	C	G	G	T	T	C	27
N34	C25478	Setti Fatma	A	C	G	G	T	C	C	9
N35	C25479	Imlil	T	C	G	G	Y	Y	C	39
N36	C25480	South of Asni	A	C	G	G	T	C	C	9
N95	TR04	near Marrakech	A	C	G	G	Y	Y	C	21

Table B3 Zone 2 ITS pattern types in Israel

Ref No.	Accession No.	Commercial name	Positions of variability									ITS Type
			50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C	
Near East Israel												
324	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
325	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
326	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
327	C4143	Tamar	T	C	G	G	C	C	C	T	C	5
328	C4143	Tamar	T	C	G	G	C	T	A	T	G	23
329	C4143	Tamar	G	G	G	G	C	G	G	T	N	16G
330	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
331	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
332	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
333	C4143	Tamar	T	C	G	G	C	Y	C	T	C	20
370	C4143	Tamar	T	C	G	G	C	T	C	T	C	15
321	C5892	Haifa	T	C	G	K	C	Y	C	T	C	10
345	C5892	Haifa	T	C	G	G	C	C	C	T	C	5
347	C5892	Haifa	T	C	G	K	C	C	C	T	C	10
348	C5892	Haifa	T	C	G	G	C	T	C	T	C	15
381	C5892	Haifa	T	C	G	K	C	C	C	T	C	10
N93	TR02	-	T	C	G	G	C	T	C	T	C	15

Table B4 Zone 2 ITS pattern types in Northern and North-Western Europe

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Positions of variability									
					50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C	ITS Type
Northern/North-Western Europe														
344	C2095	Denmark	-	Morso	/	C	G	G	C	T	A	T	C	17?
337	C2095	Denmark	-	Morso	T	C	G	G	C	T	A'	T	C	17'
340	C2095	Denmark	-	Morso	T	C	G	G	C	T	A'	T	C	17'
343	C2095	Denmark	-	Morso	T	C	G	G	C	T	A'	T	C	17'
378	C2095	Denmark	-	Morso	T	C	G	G	C	T	A'	T	C	17'
341	C2095	Denmark	-	Morso	T	C	G	G	C	T	M	T	C	18
338	C2095	Denmark	-	Morso	T	C	G	G	Y	Y	M	T	C	19
N16	C2095	Denmark	-	Morso	T	C	G	G	C	T	A	T	C	17
342	C2095	Denmark	-	Morso	T	C	G	G	T	C	C	T	C	8
339	C2095	Denmark	-	Morso	T	C	G	G	Y	C	C	T	C	32
335	C2095	Denmark	-	Morso	T	C	G	G	C	C	C	T	C	5
334	C2095	Denmark	-	Morso	T	C	G	G	C	T	C	T	C	15
336	C2095	Denmark	-	Morso	T	C	G	G	C	T	C	T	C	15
N14	C2095	Denmark	-	Morso	T	C	G	G	C	T	C	T	C	15
N15	C2095	Denmark	-	Morso	T	C	G	G	C	T	A'	T	C	17'
312	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	A'	T	C	17'
313	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	A'	T	C	17'
315	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	A'	T	C	17'
316	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	A'	T	C	17'
311	C24247	The Netherlands	-	Barbian	T	C	G	G	C	C	C	T	C	5

310	C24247	The Netherlands	-	Barbian	T	C	G	G	C	C	C	T	-	5?
314	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	C	T	C	15
349	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	C	T	C	15
373	C24247	The Netherlands	-	Barbian	T	C	G	G	C	T	C	T	C	15
317	C24247	The Netherlands	-	Barbian	T	C	G	G	C	Y	C	T	C	20
374	C24247	The Netherlands	-	Barbian	T	C	G	G	C	C	C'	T	C	5'
301	C960	Finland	-	Tammisto	T	C	G	G	C	T	C	T	C	15
300	C960	Finland	-	Tammisto	T	C	G	G	C	T	C	T	G	24
301	C960	Finland	-	Tammisto	T	C	G	G	C	T	C	T	C	15
302	C960	Finland	-	Tammisto	T	C	G	G	C	Y	C	T	C	20
303	C960	Finland	-	Tammisto	T	C	G	-	-	T	C	T	-	15?
304	C960	Finland	-	Tammisto	T	C	G	K	C	Y	C	T	C	40
305	C960	Finland	-	Tammisto	T	C	G	G	C	T	C	T	C	15
306	C960	Finland	-	Tammisto	T	C	G	G	C	T	A'	T	C	17'
307	C960	Finland	-	Tammisto	T	C	G	G	C	T	C	T	C	15
369	C960	Finland	-	Tammisto	T	C	G	G	C	Y	C	T	C	20
N99	TR08	Germany	Hausen, Baden-Wurttemberg		T	C	G	G	C	T	C	T	C	15
N105	TR14	Norway			T	C	G	G	C	C	C	T	C	5
N101	TR10	Greece	Pogoniani, Ioannina, Epirus		T	C	G	G	C	Y	C	T	C	20
N97	TR06	Romania	Domnesti		T	C	G	G	C	T	C	T	C	15
N9	-	Romania	Barro Trro		W	C	G	G	Y	C	C	T	C	44

Table B5 Zone 2 ITS pattern types in Central and Western Europe

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Positions of variability									ITS Type
					50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C	
Central/Western Europe					50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C	ITS Type
361	C7555	Czech Republic		Ovack	T	C	G	G	C	T	C	T	C	15
364	C7555	Czech Republic		Ovack	T	C	G	G	C	T	C	T	C	15
365	C7555	Czech Republic		Ovack	T	C	G	G	C	T	C	T	C	15
379	C7555	Czech Republic		Ovack	T	C	G	G	C	T	C	T	C	15
360	C7555	Czech Republic		Ovack	T	C	G	G	Y	Y	C	T	C	39
362	C7555	Czech Republic		Ovack	T	C	G	G	Y	Y	C	T	C	39
363	C7555	Czech Republic		Ovack	T	C	G	G	Y	Y	M	T	C	19
353	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	G	G	C	T	A'	T	C	17'
355	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	G	G	C	C	C	T	C	5
354	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	G	G	C	Y	C	T	C	20
350	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	G	G	C	Y	C	T	C	20
356	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	G	G	C	Y	C	T	C	20
359	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	G	G	C	T	C'	T	C	15'
352	C6360	Western Europe	Belgium Kersey mix	Blanca	T	C	R	G	C	Y	C	T	C	43

Table B6 Zone 2 ITS pattern types in British plants

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Positions of variability										ITS Type
					50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C		
81	C6374	England	Kent, Suffolk	Kent White	T	C	G	G	C	C	C	T	C	5	
82	C6374	England	Kent, Suffolk	Kent White	A	C	G	G	T'	C	C	T	C	6	
84	C6374	England	Kent, Suffolk	Kent White	T	C	G	G	C	Y'	C	T	C	20'	
85	C6374	England	Kent, Suffolk	Kent White	T	C	G	K	C	C	C	T	C	10	
86	C6374	England	Kent, Suffolk	Kent White	N	C	G	G	C	C	C	T	C	5C	
87	C6374	England	Kent, Suffolk	Kent White	W	C	G	G	C	Y	C	T	C	42	
197	C6374	England	Kent, Suffolk	Kent White	T	C	G	G	Y	T	C	T	C	35	
198	C6374	England	Kent, Suffolk	Kent White	T	C	G	G	C	T'	C	T	C	15'	
N42	C6304	England	Kent, Suffolk	Kent Wild White	T	C	G	G	C	C	C	T	C	5	
N44	C6304	England	Kent, Suffolk	Kent Wild White	T	C	G	G	C	C	C	T	C	5	
2-5	C6304	England	Kent, Suffolk	Wild White Kent	T	C	G	G	Y	C	C	T	C	32	
2-6	C6304	England	Kent, Suffolk	Wild White Kent	W	C	G	G	C	C	C	T	C	34	
2-3	C6304	England	Kent, Suffolk	Wild White Kent	T	C	G	G	C	T	C	T	C	15	
2-4	C6304	England	Kent, Suffolk	Wild White Kent	T	C	G	G	Y	Y	C	T	C	39	
K-10	C6304	England	Kent, Suffolk	Wild White Kent	T	C	G	G	C	Y	C	T	C	20	
K-9	C6304	England	Kent, Suffolk	Wild White Kent	T	C	G	G	C	Y	C	T	C	20	

2-8	C6304	England	Kent, Suffolk	Wild White Kent	/	/	G	G	C	Y	C	/	/	20?
N75	C3754	England	Kent, Suffolk	Kent Wild White	A	C	G	G	Y	C	C	T	C	31
309	C21105	Britain	Wild British populations	S. 184	N	C	G	G	Y	C	C	T	C	32C

Table B7 Zone 2: ITS pattern types in plants from French origins

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Positions of variability										
					509, T/C	50, T/a	62, C/T	111, G/a	216, G	486, C/T	571, C/a/m	613, T	654, C	ITS type	
Zone 2 French origins															
53	C7690	France	Bouches du Rhone, Entressen		T	T	C	G	G	C	C	T	C	15	
54	C7690	France	Bouches du Rhone, Entressen		T	T	C	G	G	C	C	T	C	15	
78	C7690	France	Bouches du Rhone, Entressen		T	T	C	G	G	C	C	T	C	15	
207	C7690	France	Bouches du Rhone, Entressen		T	T	C	G	G	C	C	T	C	15	
40	C9046	France	Mouriès, Bouches du Rhône		Y'	T	C	G	G	C	C	T	C	20'	
44/ 202	C9046	France	Mouriès, Bouches du Rhône		Y	T	C	G	G	C	C	T	C	20	
163	C9046	France	Mouriès, Bouches du Rhône		Y'	T	C	G	G	C	C	T	C	20'	
192	C9046	France	Mouriès, Bouches du Rhône		T	T	C	G	G	C	C	T	C	15	
N1 7	C2418	France	Le Conquet, Finistère department, Bretagne		T	T	C	G	G	C	C	T	C	15	
N1 8	C2418	France	Le Conquet, Finistère department, Bretagne			T	C	G	G	C	C	T	C	15	
N1 9	C2418	France	Le Conquet, Finistère department, Bretagne		T	T	C	G	G	C	C	T	C	15	
R1	-	France/Israe l	France & Israel cross	Aran	Y	T	C	G	G	C	C	T	C	20	
R2	-	France/Israe l	France & Israel cross	Aran	T	T	C	G	G	C	C	T	C	15	
R3	-	France/Israe l	France & Israel cross	Aran	T	T	C	G	G	C	C	T	C	15	
R4	-	France/Israe l	France & Israel cross	Aran	T	T	C	G	G	C	C	T	C	15	
R5	-	France/Israe l	France & Israel cross	Aran	T	T	C	G	G	C	C	T	C	15	

R6	-	France/Israel	France & Israel cross	Aran	Y	T	C	G	G	C	C	T	C	20
R7	-	France/Israel	France & Israel cross	Aran	Y	T	C	G	G	C	C	T	C	20
R8	-	France/Israel	France & Israel cross	Aran	T	T	C	G	G	C	C	T	C	15
R9	-	France/Israel	France & Israel cross	Aran	Y	T	C	G	G	C	C	T	C	20
R10	-	France/Israel	France & Israel cross	Aran	T	T	C	G	G	C	C	T	C	15
121	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	Y'	T	C	G	G	C	C	T	C	20'
4A	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	C	T	T	G	G	C	C	T	C	7
144	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	T	T	C	G	G	C	C	T	C	15
146	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	T	T	C	G	G	C	C	T	C	15
147	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	T	T	C	G	G	C	C	T	C	15
148	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	T	T	C	G	G	C	C	T	C	15
149	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	T	T	C	G	G	C	C	T	C	15
151	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	T	T	C	G	G	C	C	T	C	15
150	C18951	France	Crau, Provence-Alpes-Côte d'Azur region	Crau	Y	T	C	G	G	C	C	T	C	20
C1	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	T	C	G	G	C	C	T	C	5
C2	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	T	T	C	G	G	C	C	T	C	15

C3	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	Y	T	C	G	G	C	C	T	C	20
C4	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	Y	T	C	G	G	C	C	T	C	20
C5	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	T	T	C	G	G	C	C	T	C	15
C6	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	C	T	C	G	G	C	C	T	C	5
C8	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	T	T	C	G	G	C	C	T	C	15
C9	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	Y	T	C	G	G	C	C	T	C	20
C10	-	France/Syria	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	Y	T	C	G	G	C	C	T	C	20

B8 Zone 2: ITS pattern types in Ecotype Lodi and ecotype selection Regal

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Positions of variability										ITS Type
					509, T/C	50, T/a	62, C/T	111, G/a	216, G	486, C/T	571, c/a/m	613, T	654, C		
Zone 2 Italy & USA															
N39	C5897	Italy	Milan, Po Valley	Lodi	T	T	C	G	G	C	C	T	C	15	
1-2	C5897	Italy	Milan, Po Valley	Lodi	C	T	C	G	G	C	C	T	C	5	
1-4	C5897	Italy	Milan, Po Valley	Lodi	C	T	C	G	G	C	C	T	C	5	
N40	C5897	Italy	Milan, Po Valley	Lodi	T	T	C	G	G	C	C	T	C	15	
1-6	C5897	Italy	Milan, Po Valley	Lodi	Y	T	C	G	G	C	C	T	C	20	
1-8	C5897	Italy	Milan, Po Valley	Lodi	Y	T	C	G	G	C	C	T	C	20	
1-9	C5897	Italy	Milan, Po Valley	Lodi	T	T	C	G	G	C	C	T	C	15	
18	C5897	Italy	Milan, Po Valley	Lodi	T	T	C	G	G	C	C	T	C	15	
19	C5897	Italy	Milan, Po Valley	Lodi	T	T	C	G	G	C	C	Y	C	36	
103	C5897	Italy	Milan, Po Valley	Lodi	Y'	-	C	G	G	C	C	T	C	20'?	
104	C5897	Italy	Milan, Po Valley	Lodi	T	T	C	G	G	C	C	Y	C	36	
105	C5897	Italy	Milan, Po Valley	Lodi	T	G	G	G	G	C	C	Y	C	48	
102	C5897	Italy	Milan, Po Valley	Lodi	C	T	C	G	G	C	C	T	C	5	
208	C5897	Italy	Milan, Po Valley	Lodi	C	T	C	G	G	C	C	T	C	5	
N102	TR11	Italy	Pasubia		C	T	C	G	G	C	C	T	C	5	
N92	TR01	Italy	Italy		T	T	C	G	G	C	C	T	C	15	
386	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	?	T	C	G	G	C	C	T	C	5?	
387	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	T	C	G	G	C	C	T	C	5	

390	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	T	C	G	G	C	C	T	C	5
385	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	C	A	C	G	G	T	C	T	C	9
383	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	T	T	C	G	G	C	C	T	C	15
388	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	T	T	C	G	G	C	C	T	C	15
389	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	T	T	C	G	G	C	C	T	C	15
391	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	T	T	C	G	G	C	C	T	C	15
384	C15111	USA/Italy	Alabama/Po, Valley, USA Ladino Ecotype	Regal	Y	T	C	G	G	C	C	T	C	20

Table B9 Zone 3: ITS pattern types from plants outside the natural distribution zone

Ref No.	Accession No.	Country of origin	Place of origin	Positions of variability										ITS Type
				50, T/a	62, C/T	111, G/a	216, G	486, C/T	509, T/C	571, C/a/m	613, T	654, C		
Zone 3 Outer populations														
N65	CR138941	Costa Rica	Costa Rica	T	C	G	G	C	C	C	T	C	5	
N64	CR	Costa Rica	Costa Rica	T	C	G	G	C	T	C	T	C	15	
N94	TR03	USA	USA	T	C	G	G	C	C	C	T	C	5	
N109	-	USA	Georgia	W	C	G	K	C	Y	C	T	C	41	
N98	TR07	Ethiopia	Ethiopia	T	C	G	G	C	Y	C	T	C	20	
N107	TR76	China	near Urumqi, Xinjiang	T	C	G	G	C	T	C	T	C	15	
N96	TR05	China	China	T	C	G	G	C	T	C	T	C	15	
N100	TR09	Japan	Japan	T	C	G	G	C	Y	C	T	C	20	
		New Zealand	Southland/Hawkes Bay	T	C	G	G	C	T	C	T	C	15	
N103	TR12	Kazakhstan	near Alma-Ata, Alma-Ata	T	C	G	G	C	C	C	T	C	5	
319	C16511	South Africa	Introduced mixed populations	T	C	G	G	C	C	C	T	C	5	
372	C16511	South Africa	Introduced mixed populations	T	C	G	G	C	C	C	T	C	5	
371	C16511	South Africa	Introduced mixed populations	T	C	G	G	C	T	C	T	C	15	
320	C16511	South Africa	Introduced mixed populations	T	C	G	G	C	Y	C	T	C	20	
318	C16511	South Africa	Introduced mixed populations	A	C	G	G	C	C	C	T	C	12	
Novel Algerian species														
14/98	C20859	Algeria	Algeria	T	C	G	G	C	C	C	T	C	5	
59	C20859	Algeria	Algeria	-	-	G	G	C	C	C	T	C	5?	
60	C20859	Algeria	Algeria	T	C	G	G	C	C	C	T	C	5	

97/179 C20859

Algeria

Algeria

T C A G C C C T C 2

Appendix C Combined DNA sequence data

Table C1 Combined pattern types in Zone 1 Portugal and Spain

Ref No.	Accession No.	Country of origin	Place of origin	Chloroplast Type	ITS Type
Zone 1, area of origin, Portugal					
31/177	C7576	Portugal	Estremadura	CC+	15
89	C7582	Portugal	Beira litoral	TC-	5
24	C7582	Portugal	Beira litoral	TC-	5
191	C7582	Portugal	Beira litoral	TC-	5
92	C7582	Portugal	Beira litoral	CC+	15
93	C7582	Portugal	Beira litoral	TC-	20
96	C7582	Portugal	Beira litoral	TC-	20'
124	C9514	Portugal	Trans Os Montes	CC+	5
127	C9514	Portugal	Trans Os Montes	CC+	32
130	C9514	Portugal	Trans Os Montes	CC+	32
99	C9514	Portugal	Trans Os Montes	TC-	15
N57	C9516	Portugal	Vila Verde	CT+	15
6-1	C9516	Portugal	Vila Verde	CT+	15
6-8	C9516	Portugal	Vila Verde	CT+	20
N56	C9516	Portugal	Vila Verde	TC-	5
N58	C9516	Portugal	Vila Verde	CC+	34
N10	C11519	Portugal	Vila Pouca de Agu	CC+	5
N11	C11519	Portugal	Vila Pouca de Agu	CC+	5
N12	C11519	Portugal	Vila Pouca de Agu	CC+	20
153	C20477	Portugal	Águas de Moura, Palmela	CC+	2
157	C20477	Portugal	Águas de Moura, Palmela	CC-	3
154	C20477	Portugal	Águas de Moura, Palmela	CC+	45

152	C20477	Portugal	Águas de Moura, Palmela	TC-	5
8-1	AZ4838	Portugal	Minho, Braga, Ferreiras	TC-	15
9-5	AZ4839	Portugal	Tras Montes, Baragonça	CC+	5
9-2	AZ4839	Portugal	Tras Montes, Baragonça	CC+	20'
4-10	C9024	Portugal	Tras Os Montes, Esphino	CC+	5C
4-4	C9024	Portugal	Tras Os Montes, Esphino	TC-	20
4-3	C9024	Portugal	Tras Os Montes, Esphino	TC-	20
N49	C9024	Portugal	Tras Os Montes, Esphino	TC-	20
N50	C9024	Portugal	Tras Os Montes, Esphino	TC-	20
N51	C9024	Portugal	Tras Os Montes, Esphino	TC-	20
139	C8985	Portugal	Coimbra	CC+	5
141	C8985	Portugal	Coimbra	CC+	5
134	C8985	Portugal	Coimbra	CC+	15
140	C8985	Portugal	Coimbra	CC+	17
203	C8976	Portugal	Algarve	TC-	2
65	C8976	Portugal	Algarve	TC-	5
204/ 61	C8976	Portugal	Algarve	TC-	15
188	C8979	Portugal	Ribatejo, Grândola	CC+	2
47	C8979	Portugal	Ribatejo, Grândola	CC-	5
187	C8979	Portugal	Ribatejo, Grândola	TC-	5
196	C8979	Portugal	Ribatejo, Grândola	TC-	5
186	C8979	Portugal	Ribatejo, Grândola	TC-	15
117	C7600	Spain	Pyrenees	CC+	5
52	C8975	Spain	Andalucía	CC+	15
11A	C8994	Spain	Asturias	TC-	5
113/ 190	C8994	Spain	Asturias	TC-	15
114/ 211	C8994	Spain	Asturias	TC-	15
110/ 210	C8994	Spain	Asturias	TC-	15
106	C8994	Spain	Asturias	TC-	15

22/107	C8994	Spain	Asturias	TC-	15
189	C8994	Spain	Asturias	TC-	20'
199	C8994	Spain	Asturias	TC-	5
77/206	C9025	Spain	Extremadura	CC+	15
N53	C9027	Spain	La Estrada, Galicia, Pontevedra	CT+	5
N52	C9027	Spain	La Estrada, Galicia, Pontevedra	CC+	15?
88	C9034	Spain	Galicia	CT+	15
71	C9034	Spain	Galicia	TC-	15
178	C9034	Spain	Galicia	TC-	15'
193	C9037	Spain	Galicia	TC-	5
161	C9037	Spain	Galicia	TC-	15
160	C9037	Spain	Galicia	TC-	20'
194	C9037	Spain	Galicia	TC-	15'
7-5	C9531	Spain	Ponferrada, León & Castile	TC-	15?
7-4	C9531	Spain	Ponferrada, León & Castile	TC-	15
7-6	C9531	Spain	Ponferrada, León & Castile	TC-	20
N59	C9531	Spain	Ponferrada, León & Castile	TC-	34
N60	C9531	Spain	Ponferrada, León & Castile	TC-	34

Table C2 Combined pattern types in Zone 2: North Africa and the Near East: Israel

Ref No.	Accession No.	Place of origin	Commercial name	Chloroplast Type	ITS Type
Zone 2, North Africa					
3-10	C6452	Algeria	-	CC+	5
322	C6452	Algeria	-	CC+	7
323	C6452	Algeria	-	CC+	7
N47	C6452	Algeria	-	CC+	7
3-9	C6452	Algeria	-	CC+	7
3-5	C6452	Algeria	-	CC+	9?
N45	C6452	Algeria	-	3 =(3-1)	12
3-3	C6452	Algeria	-	CC+	15
3-8	C6452	Algeria	-	CC+	30
3-6	C6452	Algeria	-	CC+	5D
A2	C1751	Algeria	-	CC+	5
A1	C1751	Algeria	-	CC+	11
A3	C1751	Algeria	-	CC+	12
A5	C1751	Algeria	-	CC+	30
A7	C1751	Algeria	-	CC+	30
A8	C1751	Algeria	-	CC+	30
A6	C1751	Algeria	-	CC+	30
A10	C1751	Algeria	-	CC+	30?
A4	C1751	Algeria	-	CC+	38
170	C1761	Morocco	-	CC+	15
168	C1761	Morocco	-	CC+	15
169	C1761	Morocco	-	CC+	15
167	C1761	Morocco	-	CC+	25

Near East, Israel					
345	C5892	Huleh district	Haifa	CC+	5
347	C5892	Huleh district	Haifa	CC+	10
321	C5892	Huleh district	Haifa	CC+	10
348	C5892	Huleh district	Haifa	CC+	15
327	C4143	Israel	Tamar	CC+	5
324	C4143	Israel	Tamar	CC+	15
326	C4143	Israel	Tamar	CC+	15
330	C4143	Israel	Tamar	CC+	15
331	C4143	Israel	Tamar	CC+	15
332	C4143	Israel	Tamar	CC+	15
325	C4143	Israel	Tamar	CC+	15
328	C4143	Israel	Tamar	CC+	23

Table C3 Combined pattern types in Zone 2: Europe

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Chloroplast Type	ITS Type
Northern/North-Western Europe						
335	C2095	Denmark	Denmark	Morso	CC+	5
342	C2095	Denmark	Denmark	Morso	CC+	8
N14	C2095	Denmark	Denmark	Morso	CC+	15
336	C2095	Denmark	Denmark	Morso	CC+	15
344	C2095	Denmark	Denmark	Morso	CC+	17?
340	C2095	Denmark	Denmark	Morso	CC+	17'
343	C2095	Denmark	Denmark	Morso	CC+	17'
378	C2095	Denmark	Denmark	Morso	CC+	17'
341	C2095	Denmark	Denmark	Morso	CC+	18
338	C2095	Denmark	Denmark	Morso	CC+	19
339	C2095	Denmark	Denmark	Morso	CC+	32
314	C24247	The Netherlands	The Netherlands	Barbian	CT+	15
373	C24247	The Netherlands	The Netherlands	Barbian	CT+	15
310	C24247	The Netherlands	The Netherlands	Barbian	CC+	5?
311	C24247	The Netherlands	The Netherlands	Barbian	CC+	5
349	C24247	The Netherlands	The Netherlands	Barbian	CC+	15
312	C24247	The Netherlands	The Netherlands	Barbian	CC+	17'
315	C24247	The Netherlands	The Netherlands	Barbian	CC+	17'
317	C24247	The Netherlands	The Netherlands	Barbian	CC+	20
316	C24247	The Netherlands	The Netherlands	Barbian	CT+	17'
374	C24247	The Netherlands	The Netherlands	Barbian	CC+	5'
302	C960	Finland	Finland	Tammisto	CC+	20

303	C960	Finland	Finland	Tammisto	CC+	15?
307	C960	Finland	Finland	Tammisto	CC+	15
305	C960	Finland	Finland	Tammisto	CC+	15
301	C960	Finland	Finland	Tammisto	CC+	15
306	C960	Finland	Finland	Tammisto	CC+	17'
369	C960	Finland	Finland	Tammisto	CC+	20
300	C960	Finland	Finland	Tammisto	CC+	24
304	C960	Finland	Finland	Tammisto	CC+	40
Central/Western Europe						
361	C7555	Czech Republic	Czech Republic	Ovcak	CC+	15
364	C7555	Czech Republic	Czech Republic	Ovcak	CC+	15
365	C7555	Czech Republic	Czech Republic	Ovcak	CC+	15
363	C7555	Czech Republic	Czech Republic	Ovcak	CC+	19
360	C7555	Czech Republic	Czech Republic	Ovcak	CC+	39
362	C7555	Czech Republic	Czech Republic	Ovcak	CT+	39
359	C6360	Western Europe	Belgium/Kersey cross	Blanca	CC+	15'
353	C6360	Western Europe	Belgium/Kersey cross	Blanca	CC+	17'
350	C6360	Western Europe	Belgium/Kersey cross	Blanca	CC+	20
354	C6360	Western Europe	Belgium/Kersey cross	Blanca	CC+	20
356	C6360	Western Europe	Belgium/Kersey cross	Blanca	CC+	20
352	C6360	Western Europe	Belgium/Kersey cross	Blanca	CC+	43
N42	C6304	England	Kent, Suffolk	Kent Wild White	CC+	5
2-2	C6304	England	Kent, Suffolk	Kent Wild White	CC+	5 (Note N44)
2-3	C6304	England	Kent, Suffolk	Kent Wild White	CC+	15
2-9	C6304	England	Kent, Suffolk	Kent Wild White	CT+	20
2-10	C6304	England	Kent, Suffolk	Kent Wild White	CT+	20

2-6	C6304	England	Kent, Suffolk	Kent Wild White	CT+	34
85	C6374	England	Kent, Suffolk	Kent White	TC-	10
198	C6374	England	Kent, Suffolk	Kent White	CT+	15'
81	C6374	England	Kent, Suffolk	Kent White	CC+	5
82	C6374	England	Kent, Suffolk	Kent White	CC+	6
84	C6374	England	Kent, Suffolk	Kent White	CC+	20'
197	C6374	England	Kent, Suffolk	Kent White	CC+	35
87	C6374	England	Kent, Suffolk	Kent White	CC+	42
309	C21105	Britain	Wild British populations	S.184	CT+	32C

Table C4 Combined pattern types in Zone 2: for plants from French origins

Ref No.	Accession No.	Place of origin	Commercial name	Chloroplast Type	ITS Type
Zone 2, French Origins					
78	C7690	Bouches du Rhone, Entressen	-	CT+	15
54	C7690	Bouches du Rhone, Entressen	-	CT+	15
207	C7690	Bouches du Rhone, Entressen	-	CT+	15
53	C7690	Bouches du Rhone, Entressen	-	CC+	15
192	C9046	Mouriès, Bouches du Rhône	-	CC+	15
44/ 202	C9046	Mouriès, Bouches du Rhône	-	CC+	20
40	C9046	Mouriès, Bouches du Rhône	-	CC+	20'
163	C9046	Mouriès, Bouches du Rhône	-	CC+	20'
4A	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	7
149	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	15
151	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	15
144	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	15
147	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	15
148	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	15
150	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	20
146	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	CT+	15
121	C18951	Crau, Provence-Alpes-Côte d'Azur region	Crau	TC-	20'
R10	-	France & Israel cross	Aran	CT+	15
R2	-	France & Israel cross	Aran	CT+	15
R3	-	France & Israel cross	Aran	CT+	15
R8	-	France & Israel cross	Aran	CC+	15
R4	-	France & Israel cross	Aran	CC+	15
R5	-	France & Israel cross	Aran	CC+	15
R7	-	France & Israel cross	Aran	CC+	20

R9	-	France & Israel cross	Aran	CC+	20
R1	-	France & Israel cross	Aran	CC+	20
R6	-	France & Israel cross	Aran	CT+	20
C2	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CT+	15
C4	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CT+	20
C9	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CT+	20
C1	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CC+	5
C5	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CC+	15
C8	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CC+	15
C10	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CC+	20
C3	-	French Crau & Syrian-(Mader Village, Zebdani District)	Crusader	CY+	20
N17	C2418	Le Conquet, Finistère department, Bretagne		CC+	15
N18	C2418	Le Conquet, Finistère department, Bretagne		CC+	15
N19	C2418	Le Conquet, Finistère department, Bretagne		CC+	15

Table C5 Combined pattern types in ecotype Lodi and ecotype selection Regal

Ref No.	Accession No.	Place of origin	Commercial name	Chloroplast Type	ITS Type
Zone 2, Italy &					

USA

1-2	C5897	Milan, Po Valley, Italy	Lodi	CC+	5
1-4	C5897	Milan, Po Valley, Italy	Lodi	CC+	5
1-1	C5897	Milan, Po Valley, Italy	Lodi	CC+	15 = (n39)
1-9	C5897	Milan, Po Valley, Italy	Lodi	CC+	15
18	C5897	Milan, Po Valley, Italy	Lodi	CC+	15
1-6	C5897	Milan, Po Valley, Italy	Lodi	CC+	20
1-10	C5897	Milan, Po Valley, Italy	Lodi	CC+	20?
103	C5897	Milan, Po Valley, Italy	Lodi	CC+	20'?
208	C5897	Milan, Po Valley, Italy	Lodi	CC-	5
N40	C5897	Milan, Po Valley, Italy	Lodi	TC-	15
386	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	5?
387	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	5
390	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	5
385	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	9
383	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	15
388	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	15
389	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	15
391	C15111	Alabama/Po, Valley, USA Ladino Ecotype	Regal	CC+	15

Table C6 Combined pattern types in Zone 3: outer populations from South Africa

Ref No.	Accession No.	Country of origin	Place of origin	Commercial name	Chloroplast Type	ITS Type
Zone 3, Outer populations						
318	C16511	South Africa	Introduced mixed populations	Dusi	CC+	5
371	C16511	South Africa	Introduced mixed populations	Dusi	CC+	20
372	C16512	South Africa	Introduced mixed populations	Dusi	CC+	5

Appendix D-2: Comparison between DNA pattern combinations and morphological appearance in white clover individuals.

D.1 Portuguese individuals

In this group there was a trend of leaflets of small size and ovate or orbicular shape, with a higher occurrence of the ovate leaflet shape. No apparent association between leaflet appearance (mark/colour) with any of leaflet size, shape or the combined pattern type was observed. No consensus of DNA combination and morphological traits was observed. The recessive unmarked leaf phenotype was not observed.

Table D.1 Combined pattern types and corresponding morphological characters in Portuguese accessions of white clover

Reference Number	Accession Number	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
31/177	C7576	thick	-	-	CC+/15
89	C7582	thin	S	ov	TC-/5
24	C7582	thin	S-M	or	TC-/5
191	C7582	norm	S-M	ov	TC-/5
93	C7582	thick	S-M	ov	TC-/20
96	C7582	norm	S	or	TC-/20'
124	C9514	-	S	ov	CC+/5
99	C9514	-	S	-	TC-/15
153	C20477	incomp/bright	S	ov	CC+/2
157	C20477	bright/squat	S	ov	CC-/15
154	C20477	thick	S	ov	CC+/45
152	C20477	bright	S	-	TC-/5

Note: 11 plants were assessed. Leaflet Mark/Colour refers to the presence or absence (green) of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick, squat/incomplete or of standard normal appearance. Leaflet size: S (Small), S-M (Small>Medium). Leaflet shape: ov (ovate), or (orbicular).incomp = incomplete V marking. (-) this characteristic was not scored.

D.2 Comparison of genetic character and morphological appearance in Spanish accessions of white clover.

In the group of Spanish accessions a trend of leaflets primarily of the tiny-small size range was observed. Medium size leaflets were observed mostly in conjunction with ovate shaped leaflets and were independent of combined pattern type and mark appearance. The association of unmarked leaves and small sized ovate shaped leaflets was observed in two individuals from separate accessions.

Table D.2 Combined pattern types and corresponding morphological characters in Spanish accessions

Reference Number	Accession Number	Leaflet mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
117	C7600, Spain	norm	S	ov	CC+/5
203	C8976, Spain	thin	S-M	ov	TC-/2
65	C8976, Spain	thin	S-M	-	TC-/5
204/ 61	C8976, Spain	norm	S	ov	TC-/15
188	C8979, Spain	norm	T	ov	CC+/2
47	C8979, Spain	faint	T-S	ov	CC-/5
187	C8979, Spain	norm	T-S	or	TC-/5
196	C8979, Spain	norm	T-S	or	TC-/5
186	C8979, Spain	norm	T-S	or	TC-/15
141	C8985, Spain	-	S-M	ov	CC+/5
134	C8985, Spain	green	S	ov	CC+/15
140	C8985, Spain	-	S	ov	CC+/17
11A	C8994, Spain	norm	S	ov	TC-/5
113/ 190	C8994, Spain	norm	S-M	ov	TC-/15
110/ 210	C8994, Spain	thick	T-S	ov	TC-/15
106	C8994, Spain	faint	S-M	or	TC-/15
22/ 107	C8994, Spain	green	-	ov	TC-/15
189	C8994, Spain	faint	S-M	ov	TC-/20'
77/206	C9025, Spain	faint	T-S	-	CC+/15

Note: 25 plants were assessed. Leaflet Mark/Colour refers to the presence or absence (green) of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick, incomplete or of standard normal appearance. Leaflet size: T (Tiny), T-S (Tiny- small), S (Small), S-M (Small>Medium). Leaflet shape: ov (ovate), or (orbicular). (-) this characteristic was not scored.

Table D.2 (continued), Combined pattern types and corresponding morphological characters in Spanish accessions

Reference Number	Accession Number	Leaflet mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
71	C9034, Spain	thick/incomp	T-S	ov	TC-/15
178	C9034, Spain	faint	-	-	TC-/15'
193	C9037, Spain	green	S	ov	TC-/5
161	C9037, Spain	norm	S	or	TC-/15
160	C9037, Spain	bright	S-M	ov	TC/20'
194	C9037, Spain	faint	S	or	TC-/15'

Note: 25 plants were assessed. Leaflet Mark/Colour refers to the presence or absence (green) of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick, incomplete or of standard normal appearance. Leaflet size: T (Tiny), T-S (Tiny- small), S (Small), S-M (Small>Medium). Leaflet shape: ov (ovate), or (orbicular). (-) this characteristic was not scored.

D.3 Comparison of genetic character and morphological appearance in French white clovers

French white clover showed no consensus, although there was a tendency for leaflets of ovate shape. Small to medium sized leaflets were observed in each accession and were larger in Crau. Landrace Crau (Southern France) was the most variable. Three instances were observed of the green phenotype, but there was no correlation with other morphological traits.

Table D.3 Combined pattern types and corresponding morphological characters in French white clovers

Reference Number	Accession Number	Leaflet		Leaflet shape	Combined pattern type	
		Mark & Colour	Leaflet size			
78	C7690, France	green	S	or	CT+/15	
54	C7690, France	faint	S	ov	CT+/15	
207	C7690, France	thin	S-M	ov/or	CT+/15	
53	C7690, France	thin/bright	S	ov	CC+/15	
192	C9046, France	thin/faint	S	ov	CC+/15	
44/ 202	C9046, France	green	S	ov	CC+/20	
40	C9046, France	norm	S	ov/or	CC+/20'	
163	C9046, France	-	S-M	-	CC+/20'	
4A	C18951, Crau	faint	S	ov	CT+/7	
149	C18951, Crau	thin/faint	S	ov	CT+/15	
151	C18951, Crau	thick	S-M	ov	CT+/15	
144	C18951, Crau	bright	S	ov	CT+/15	Note:
147	C18951, Crau	thin/faint	S-M-L	ov	CT+/15	17
148	C18951, Crau	norm	S-M	or	CT+/15	plant
150	C18951, Crau	norm	S	ov	CT+/20	s
146	C18951, Crau	green	M-L	ov	CT+/15	were
121	C18951, Crau	faint	M	ov	TC-/20'	asses

sed. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick or of standard normal appearance. Leaflet size: S (Small), S-M (Small>Medium), M (Medium), M-L (Medium>Large), S-M-L (Small-Medium-Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular). (-) this characteristic was not scored.

D.4 Comparison of genetic character and morphological appearance in British white clovers

In the British grouped individuals tiny to tiny-small sized leaflets were consistently observed. These were primarily ovate in shape (ov). No commonality was observed in either leaflet appearance (mark/colour) or combined pattern type, except in the mixed cultivar Blanca, where thick white V mark with ovate shaped leaflets occurred with the combined pattern type CC+/20.

Table D.4 Combined pattern types and corresponding morphological characters in British white clovers

Reference Number	Accession Number & Names	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
309	C21105, S. 184	v. faint	T-S	ov	CT+/32C
353	C6360, Blanca	thin	S-M	ov/or	CC+/17'
350	C6360, Blanca	thick	M-L	ov	CC+/20
354	C6360, Blanca	thick	S-M	ov	CC+/20
356	C6360, Blanca	thick	M-L	ov	CC+/20
352	C6360, Blanca	thick	M-L	ov	CC+/43
85	C6374, Kent White	-	T-S	ov/or	TC-/10
198	C6374, Kent White	faint	T-S	ov	CT+/15
81	C6374, Kent White	norm	T	ov/or 3:1	CC+/5
82	C6374, Kent White	squat	T	ov/or 1:3	CC+/6
84	C6374, Kent White	green	T	ov	CC+/20'
197	C6374, Kent White	faint	T-S	ov	CC+/35
87	C6374, Kent White	incomplete	T-S	ov	CC+/42

Note: 13 plants were assessed. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick, squat/incomplete or of standard normal appearance. Leaflet size: (T) Tiny, (T-S) Tiny-small, S-M (Small>Medium), M-L (Medium>Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular). (-) this characteristic was not scored.

D.5 Comparison of genetic character and morphological appearance of white clover ecotype Italian Lodi and USA ecotype selection Regal.

Morphological scoring of only two plants of *Lodi Giganteum Lodigiano* (Italian Lodi) inhibited comparison to the ecotype selection Regal. The two individuals sampled showed leaflets with standard white V mark and orbicular shape, with a tendency for small size. This trend was not reflected in the ecotype selection Regal, which showed a prominence of ovate leaflets with thick white V marks and medium to Large size, although variation was observed in the size of the leaflets.

Table D.5 Combined pattern types and corresponding morphological characters of ecotype Italian Lodi and USA ecotype selection Regal.

Reference Number	Accession No. & names	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
386	C15111, Regal	Thick	M-S	ov	CC+/5
387	C15111, Regal	Thick	M-L	ov	CC+/5
390	C15111, Regal	Thick	M-L	ov	CC+/5
385	C15111, Regal	double V	S-M-L	ov	CC+/9
383	C15111, Regal	v. faint	M-L	ov/or	CC+/15
388	C15111, Regal	Thick	S-M	ov	CC+/15
389	C15111, Regal	-	M-L	ov	CC+/15
18	C5897, Lodi	norm	S-M	or	CC+/15
208	C5897, Lodi	norm	S	or	CC-/5

Note: 9 plants were assessed. *Giganteum Lodigiano* (5897/Italian Lodi). Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is v. faint, faint/ bright, thin/thick, squat/incomplete or of standard (norm) appearance.

Leaflet size: S (Small), S-M (Small>medium), M-S (Medium>small), M-L (Medium>Large), S-M-L

(Small-Medium-Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular).

(-) indicates a characteristic that was not scored.

D.6 Comparison of genetic character and morphological appearance in northern and north-western European ecotype selections and ecotype Morso.

In this European group it was observed that leaflet shape and combined pattern form varied within and between the different sub-groups. Ecotype Morso showed two different phenotypic forms, thin/incomplete/T-S/ovate and green/S/obcordate. Barbian showed a phenotypic form of v. faint/S/ov/combined pattern type CC+/5. No consensus of morphology and combination pattern type was present in Tammisto. The group also showed variable leaflet shape, but had a prominence of orbicular shaped leaflets.

The recessive green genotype, was observed in three and two individuals in Morso/Barbian and Tammisto. The green morph showed a tendency of association with small or tiny-small sized leaflets, but was independent of leaflet shape.

Table D.6 Combined pattern types and corresponding morphological characters in northern and north-western European white clover

Reference Number	Accession No. & Names	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
335	C2095, Morso	norm	S	ov/or 3:1	CC+/5
342	C2095, Morso	green	S	ob	CC+/8
N14	C2095, Morso	-	-	-	CC+/15
336	C2095, Morso	thin incomp	T-S	ov	CC+/15
340	C2095, Morso	v. faint	S	ov/or	CC+/17'
343	C2095, Morso	thin incomp	T-S	ov/or 1:3	CC+/17'
378	C2095, Morso	green	S	ob	CC+/17'
341	C2095, Morso	v. faint	S	ro	CC+/18
338	C2095, Morso	thin incomp	T-S	-	CC+/19
339	C2095, Morso	green	S-M	ov/ro	CC+/32
314	C24247, Barbian	faint	S-M	or/ ro	CT+/15
373	C24247, Barbian	v. Faint	S	ov/ro	CT+/15
310	C24247, Barbian	v. faint	S	ov	CC+/5
311	C24247, Barbian	v. faint	S	ov	CC+/5
349	C24247, Barbian	norm	S	or/ ob 1:3	CC+/15
312	C24247, Barbian	green	S	or	CC+/17'
315	C24247, Barbian	faint /thin	S	or	CC+/17'
317	C24247, Barbian	bright/ thick	S-M	ov	CC+/20
316	C24247, Barbian	thick	S	ov/ob	CT+/17'
374	C24247, Barbian	green	S	ov	CC+/5'
303	C960, Tammisto	norm	T-S	ov/or	CC+/15
307	C960, Tammisto	green	T-S	or	CC+/15
305	C960, Tammisto	thick	T-S	ov/or 1:3	CC+/15
306	C960, Tammisto	green	S	ov	CC+/17'
369	C960, Tammisto	faint incomp	S	ov	CC+/20
304	C960, Tammisto	green	T-S	ob/ro	CC+/40

Note: 26 plants were assessed. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick, squat/incomplete. Leaflet size: T-S (Tiny-small), S (Small), S-M (Small>Medium), M (Medium), M-L (Medium>Large), S-M-L (Small-Medium-Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular), ob (obcorate), ro (retuse obviate). (-) this characteristic was not scored.

D.7 Comparison of genetic character and morphological appearance in Central European cultivar Ovcak and Western European mixed cultivar Blanca.

For the cultivar Ovcak and the mixed cultivar Blanca, there was some consistency in appearance of the white V markings; thick in Blanca and to lesser extent in Ovcak. Leaflet size was primarily in the medium range, with sizes either S-M or M-B in both cultivars. Leaflet shape was primarily ovate. No connection between leaflet morphology and DNA pattern combination was observed.

Table D.7 Combined pattern types and corresponding morphological characters in Central European cultivar Ovcak and Western European mixed cultivar Blanca.

Reference Number	Accession No. & Names	Leaflet			Combined pattern type
		Mark & Colour	Leaflet size	Leaflet shape	
361	C7555, Ovcak	-	S-M	ov/or 1:3	CC+/15
364	C7555, Ovcak	thick	S-M	or	CC+/15
365	C7555, Ovcak	thick	M-L	ov	CC+/15
363	C7555, Ovcak	v. faint	M-S	ov/or 1:3	CC+/19
360	C7555, Ovcak	v. faint	M-S	-	CC+/39
362	C7555, Ovcak	-	S-M-L	ov	CT+/9
353	C6360, Blanca	thin	S-M	ov/or	CC+/17'
350	C6360, Blanca	thick	M-L	ov	CC+/20
354	C6360, Blanca	thick	S-M	ov	CC+/20
356	C6360, Blanca	thick	M-L	ov	CC+/20
352	C6360, Blanca	thick	M-L	ov	CC+/39

Note: 11 plants were assessed. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick, squat/incomplete or of standard normal appearance. Leaflet size S-M (Small>Medium), M-L (Medium>Large), S-M-L (Small-Medium-Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular), ob (obcorate). (-) this characteristic was not scored.

D.8 Comparison of genetic character and morphological appearance of white clover ecotype Haifa, selection Tamar of Israeli origins

It was observed that individuals with Israeli origin collectively showed a trend for leaflets with thick white V marks and predominantly medium size and ovate shape. A suggestion was observed that the pattern type combination CC+/15, maybe observed with this, but given the small sample size, any tentative association is likely due to chance. Further sampling and statistical testing is required to proof this. Leaflet size and shape was variable across the two populations.

No observations were made of the green phenotype/recessive genotype.

Table D.8 Combined pattern types and corresponding morphological characters of white clover selections of Israeli origins Haifa and Tamar and mixed cultivar Aran.

Reference Number	Accession No. & Names	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
327	C4143, Tamar	thick	S-M-L	ov	CC+/5
324	C4143, Tamar	thick	M-L	ov	CC+/15
326	C4143, Tamar	thick	S-M-L	ov/ob 1:3	CC+/15
330	C4143, Tamar	thick	M-S	ov/or	CC+/15
331	C4143, Tamar	thick	M-L	or	CC+/15
332	C4143, Tamar	thick	S-M-L	ov/or	CC+/15
325	C4143, Tamar	thick/bright	-	-	CC+/15
328	C4143, Tamar	thick/bright	M-S	ov	CC+/23
345	C5892, Haifa	thick	M	ov	CC+/5
321	C5892, Haifa	thick	S-M	ov	CC+/17

Note: 10 plants were assessed. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick or of standard normal appearance. Leaflet size: S-M (Small>Medium), M (Medium), M-L (Medium>Large), S-M-L (Small-Medium-Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular), ob (obcorate). (-) this characteristic was not scored.

D.9 Comparison of genetic character and morphological appearance of North African accessions from Algeria and Morocco.

Within the Moroccan accession (C1761) a tendency for tiny-small sized leaflets of ovate shape was observed. Two individuals of the Algerian accession C6452 were scored. The same pattern combination and leaflet size tendency was observed, but leaflet marking and shape were different for these two individuals. One occurrence of the green phenotype was observed, in one of these individuals.

Table D.9 Combined pattern types and corresponding morphological characters North African accessions from Algeria and Morocco.

Reference Number	Accession No. & Names	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
170	C1761, Morocco	thick	M-S	ov	CC+/15
168	C1761, Morocco	norm	T-S	ov	CC+/15
169	C1761, Morocco	norm	T-S	ov	CC+/15
167	C1761, Morocco	norm	T-S	ov	CC+/25
322	C6452, Algeria	green	S-M-L	or/ob 7:1	CC+/7
323	C6452, Algeria	thin/ faint	S-M	ov	CC+/7

Note: 6 plants were observed. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet and the appearance of this mark, whether it is faint/ bright, thin/thick or of standard normal appearance. Leaflet size: T-S (Tiny-small), S-M (Small>Medium), M-S (Medium.>small), M-L (Medium>Large), L (Large), S-M-L (Small-Medium-Large). Leaflet shape: ov (ovate), or (orbicular), ov/or (ovate & orbicular), ob (obcorate). (-) this characteristic was not scored.

D.10 Comparison of genetic character and morphological appearance of the out-group individuals, of the South African mixed cultivar Dusi.

An inconclusive survey was conducted on the out-group individuals, two were scored for morphological character. Both these individuals did however show a tendency of green leaflet appearance, large size leaflets and potentially ovate leaflet shape. No connection between morphological character and pattern combinations could be drawn.

Table D.10 Combined pattern types and corresponding morphological characters of the mixed cultivar Dusi (South Africa)

Reference Number	Accession No. & Names	Leaflet Mark & Colour	Leaflet size	Leaflet shape	Combined pattern type
318	C16511, Dusi , S. Africa	green	M-L	-	CC+/5
371	C16511, Dusi , S. Africa	green	L	ov	CC+/20

Note: 2 plants were scored. Leaflet Mark/Colour refers to the presence or absence of the white V mark on the clover leaflet. Green- no V mark observed. Leaflet size: M-L (Medium>Large), L (Large). Leaflet shape: ov (ovate). (-) this characteristic was not scored.

Appendix E TCS diagrams

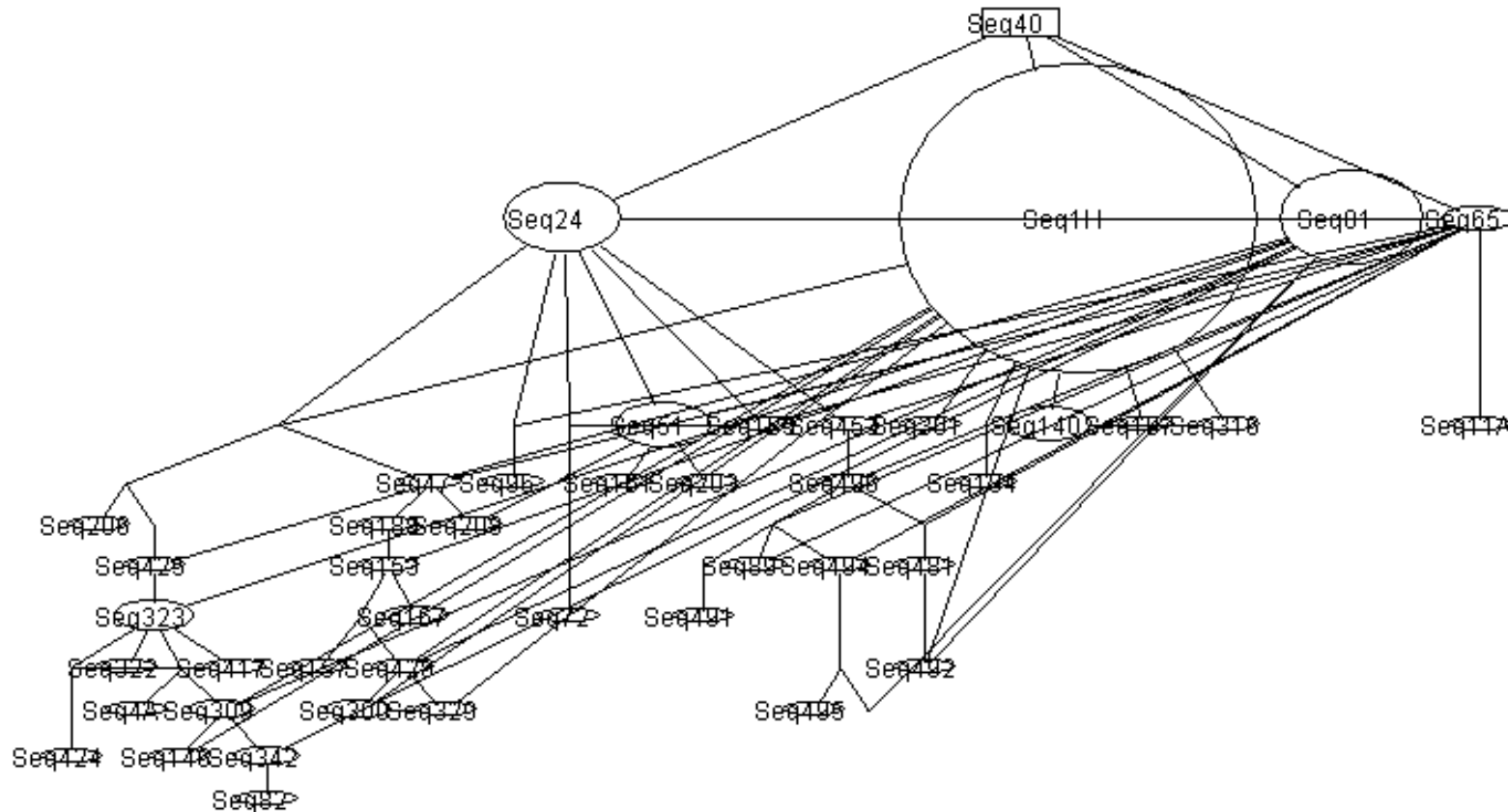


Figure E.1 Combined analysis with C/T and A/T, variations with the alignment ordered trnL intron- IRB21NE- ITS, and run under the setting gaps as missing data. Groupings: 1H: Huia Reference, Spain, Portugal, France, Crau, Aran, Crusader, Italy, Tammisto, Tamar, Barbian, Morso, Kent, Regal, Blanca, Hafia, S. Africa- Dusi, and Morocco, 24/65: Portugal/ Spain, turquoise- Portugal, Spain and Kent, grey/blue- Portugal, Tammisto, Barbian, Morso and Blanca, purple- Lodi and Regal, maroon- Portugal and Italy, navy blue- Morocco and Blanca, 81: Portugal, Spain, Kent, Barbian, Morso, Regal, Italy, Algeria, and Dusi, 323: Algeria 6452 and 1751. Individuals: same arrangement as prior analyses: Spain- (11A, 72, 453, 161, 189, 206), Portugal- (89, 96, 134, 146, 153, 157, 188, 203, 481, 491, 492, 494, 495, 496), Tammisto- (316, 301), Kent- (82, 197), Crau- (4A), Crusader- (444), Italy Lodi- (208), Tamar- (328), Algeria- (322, 417, 424, 425, 428).

Appendix F Chi squared analysis

1. Leaflet Markings

Chi squared analyses were performed for variation in white V mark (Thick, Thin, Norm, Squat, Green, Incomplete, Bright, Faint) against chloroplast pattern types (CC+, CT+, TC-) against, and against nuclear ITS pattern types, grouped according to prominence in the data set,(5-5', 15-15', 20-20', 17-17'-18, others).

1.1.1 Chloroplast patterns types against variation in V marking

For the analysis of chloroplast pattern types against variation in V marking a χ^2 value of 21.77 was obtained. No significance for comparisons made was observed, (Table F.1).

Table F.1 Chloroplast patterns types against variation in V marking

	Chloroplast pattern type	Leaflet character								Totals
		Thick	Thin	Normal	Green	Squat	Incomplete	Bright	Faint	
Observed	CC+	29	10	11	14	1	6	5	13	89
Expected	CC+	23	11.5	14.8	11.5	0.67	4.7	5.4	17.5	89.07
χ^2	CC+	1.56	0.2	0.98	0.54	0.16	0.36	0.03	1.16	4.99
Observed	CT+	2	3	2	0	0	0	1	8	18
Expected	CT+	4.6	2.3	3	2.3	0	0	1.1	3.8	17.1
χ^2	CT+	1.47	0.21	0.33	0.04	0	0	9.1 x 10 ⁻³	5.33	7.39
Observed	TC-	3	4	9	1	0	1	2	5	25
Expected	TC-	6.4	3.2	4.16	3.2	0	1.3	1.5	4.9	24.66
χ^2	TC-	1.81	0.2	5.63	1.51	0	0.07	0.17	2.04x10 ⁻³	9.39
	Observed Totals	34	17	22	17	1	7	8	26	N= 132

For this analysis 14 degrees of freedom were used for (N= 132) 132 individuals. The total χ^2 value was 21.77.

1.1.2 ITS types against variation in V marking

For assessment of nuclear ITS pattern types against variation in V marking of leaflets no associations of significance were observed, with a total χ^2 value of 28.20, (Table F.2).

Table F. 2 Nuclear pattern types against variation in white V marking

	Nuclear ITS type	Leaflet character								Totals
		Thick	Thin	Normal	Green	Squat	Incomplete	Bright	Faint	
Observed	5, 5'	5	3	7	3	0	0	2	2	22
Expected	5, 5'	6	2.75	3.95	2.92	0.17	1.2	1.2	3.78	21.97
X ²	5, 5'	0.17	0.23	2.36	2.19x10 ⁻³	0.17	1.2	0.53	0.84	5.50
Observed	15, 15'	15	6	10	4	0	2	3	9	49
Expected	15, 15'	13.4	6.13	8.8	6.51	0.38	2.68	2.68	8.42	49
X ²	15, 15'	0.19	2.55x10 ⁻³	0.16	0.97	0.38	0.17	0.04	0.4	2.31
Observed	17, 17, 18	2	4	1	3	0	1	0	2	13
Expected	17, 17, 18	3.55	1.63	2.34	1.73	0.10	0.71	0.71	2.23	13
X ²	17, 17, 18	0.68	3.47	0.77	0.93	0.10	0.12	0.71	0.02	6.8
Observed	20, 20'	8	0	3	3	0	1	1	2	18
Expected	20, 20'	4.92	2.25	3.23	2.39	0.14	0.98	0.98	3.09	17.98
X ²	20, 20'	1.93	2.25	0.02	0.16	0.14	2.6x10 ⁻⁴	2.6x10 ⁻⁴	0.38	4.88
Observed	Others	5	3	2	4	1	3	1	7	26
Expected	Others	7.11	3.25	4.67	3.45	0.20	1.42	1.42	4.47	25.99
X ²	Others	0.63	0.02	1.53	0.09	3.13	1.76	0.12	1.43	8.71
Observed	Totals	35	16	23	17	1	7	7	22	N= 128

For this analysis 28 degrees of freedom were used, with (N= 128) 128 individuals. The total χ^2 was 28.20.

1.1.2 Quantitative measurements of V mark character presence (complete, partial) or absence

1.2.1 Combined pattern types against variation in V mark character

Chi squared analyses were performed for the combinations of presence, partial V mark or green absence of V mark against combined pattern types. This analysis excluded the characteristics (faint, bright) intensity of whiteness of V mark, as this is a subjective quality indicative of pigmentation expression rather than presence or absence of marking. Instead this analysis groups the characteristics thick, thin and normal as full presence, squat and incomplete as partial presence and green as absence (Table F.3). No significant associations were observed for a total χ^2 of 21.68.

Table F. 3 Combined pattern types against the presence/ absence of V marking

	Combined Pattern types	Presence	Partial	Green	Totals
Observed	CC+/5, 5'	9	0	2	11
Expected	CC+/5, 5'	8.02	0.92	2.06	11
χ^2	CC+/5, 5'	0.12	0.92	1.92x10 ⁻³	1.04
Observed	CC+/15, 15'	19	1	2	22
Expected	CC+/15, 15'	16.04	1.83	4.13	22
χ^2	CC+/15, 15'	0.55	0.38	1.09	2.02
Observed	CC+/17, 17, 18	4	1	3	8
Expected	CC+/17, 17, 18	5.83	0.67	1.5	8
χ^2	CC+/17, 17, 18	0.58	0.17	1.5	2.25
Observed	CC+/20, 20'	8	1	3	12
Expected	CC+/20, 20'	8.75	1	2.25	12
χ^2	CC+/20, 20'	0.06	0	0.25	0.31
Observed	CC+/Others	8	4	4	16
Expected	CC+/Others	11.67	1.33	3	16
χ^2	CC+/Others	1.15	5.34	5.34	11.83

Note: this table is continued on the following page.

Table F. 3 (continued), Combined pattern types against variation in presence/absence of V marking

	Combined Pattern types	Presence	Partial	Green	Totals
Observed	CT+/5, 5'	0	0	0	0
Expected	CT+/5, 5'	0	0	0	0
χ^2	CT+/5, 5'	0	0	0	0
Observed	CT+/15, 15'	5	0	2	7
Expected	CT+/15, 15'	5.10	0.58	1.31	6.99
χ^2	CT+/15, 15'	2.12x10 ³	0.58	0.36	0.94
Observed	CT+/17, 17, 18	1	0	0	1
Expected	CT+/17, 17, 18	0.73	0.08	0.19	1
χ^2	CT+/17, 17, 18	0.10	0.08	0.19	0.37
Observed	CT+/20, 20'	1	0	0	1
Expected	CT+/20, 20'	0.73	0.08	0.19	1
χ^2	CT+/20, 20'	0.10	0.08	0.19	0.37
Observed	CT+/Others	0	0	0	0
Expected	CT+/Others	0	0	0	0
χ^2	CT+/Others	0	0	0	0
Observed	TC-/5, 5'	7	0	1	8
Expected	TC-/5, 5'	5.83	0.67	1.5	8
χ^2	TC-/5, 5'	0.23	0.67	0.17	1.07
Observed	TC-/15, 15'	5	1	1	7
Expected	TC-/15, 15'	5.10	0.58	1.31	6.99
χ^2	TC-/15, 15'	2.12x10 ³	0.3	0.07	0.37
Observed	TC-/17, 17, 18	0	0	0	0
Expected	TC-/17, 17, 18	0	0	0	0
χ^2	TC-/17, 17, 18	0	0	0	0
Observed	TC-/20, 20'	3	0	0	3
Expected	TC-/20, 20'	2.19	0.25	0.56	3
χ^2	TC-/20, 20'	0.30	0.25	0.56	1.11
Observed	TC-/Others	0	0	0	0
Expected	TC-/Others	0	0	0	0
χ^2	TC-/Others	0	0	0	0
Observed	Totals	70	8	18	96

This analysis used 24 degrees of freedom, with for (N=96) 96 individuals. The total χ^2 value observed was 21.68.

1.4 Leaflet Shape

Chi squared analyses were performed for chloroplast pattern types (CT+, CC+, TC-) against leaflet shape, and for nuclear ITS pattern types grouped as: (5-5', 15-15', 20-20', 17-17'-18, others).

1.4.1 Chloroplast pattern types against leaflet shape

The analysis of chloroplast pattern types against leaflet shape showed no significant associations. The total χ^2 was 9.96, (Table F.5).

Table F.5 Chloroplast pattern types against leaflet shape

	Chloroplast pattern type	Leaflet shape				Totals
		ov- ovate	or- orbicular	ro- retuse obovate	ob- obcordate	
Observed	CC+	64	24	2	6	96
Expected	CC+	63.3	25.2	2.72	4.77	95.99
X ²	CC+	7.7x10 ⁻³	0.06	0.19	0.32	0.58
Observed	CT+	15	4	2	1	22
Expected	CT+	14.51	5.77	0.62	1.09	20.92
X ²	CT+	0.02	0.54	3.06	7.43x10 ⁻³	3.63
Observed	TC-	14	9	0	0	23
Expected	TC-	15.2	6.04	0.65	1.14	23.03
X ²	TC-	0.09	1.46	3.06	1.14	5.75
	Observed Totals	93	37	4	7	N= 141

This analysis was performed with 6 degrees of freedom for 141 individuals (n=141). The total χ^2 was 9.96. Leaflet shape classes: ov- ovate, or -orbicular ro- retuse obovate, ob obcordate.

1.4.2 Nuclear pattern types against leaflet shape

Testing for associations of ITS pattern types (5-5', 15-15', 20-20', 17-17'-18, others) versus leaflet shape (ov- ovate, or -orbicular ro- retuse obovate, ob obcordate), found no significant associations.

Table F. 6 Nuclear pattern types against leaflet shape

	ITS pattern type	Leaflet shape ov- ovate	or- orbicular	ro- retuse obovate	ob- obcordate	Totals
Observed	5, 5'	17	5	0	0	22
Expected	5, 5'	14.51	5.77	0.04	1.09	21.41
χ^2	5, 5'	0.43	0.10	0.04	1.09	1.66
Observed	15, 15'	31	19	2	2	54
Expected	15, 15'	35.62	14.17	1.53	2.68	54
χ^2	15, 15'	0.6	1.65	0.14	0.17	2.56
Observed	17, 17, 18	8	6	0	2	16
Expected	17, 17, 18	10.55	4.19	0.45	0.79	15.98
χ^2	17, 17, 18	0.62	0.78	0.45	1.83	3.68
Observed	20, 20'	16	2	0	0	18
Expected	20, 20'	11.87	4.72	0.51	0.89	17.99
χ^2	20, 20'	1.44	1.57	0.51	0.89	4.41
Observed	Others	21	5	2	3	31
Expected	Others	20.45	8.14	0.88	1.54	31.01
χ^2	Others	0.02	1.21	1.43	1.38	4.04
Observed	Totals	93	37	4	7	N= 141

This analysis had 12 degrees of freedom and (n= 141) 141 individuals. Leaflet shape classes: ov- ovate, or -orbicular ro- retuse obovate, ob- obcordate. The total χ^2 value was 16.35.

1.4.3 Combined pattern types against leaflet shape

Chi squared analyses were not performed for this set of associations due to the lack of individuals observed within the Ro- Reutse obovate and Ob- obcordate classes. The observed numbers for leaflet shape against combined pattern types are presented in Table F.7.

Table F. 7 Observed leaflet shape against combined pattern types

Combined Pattern types	Leaflet shape				Totals
	ov-ovate	or-orbicular	ro-retuse obovate	ob-obcordate	
CC+/5, 5'	13	2	0	0	15
CC+/15, 15'	18	11	0	2	31
CC+/17, 17, 18	7	6	1	1	15
CC+/20, 20'	11	1	0	0	12
CC+/Others	14	4	1	2	15
CT+/5, 5'	0	0	0	0	0
CT+/15, 15'	9	4	2	0	15
CT+/17, 17, 18	1	0	0	1	2
CT+/20, 20'	2	0	0	0	2
CT+/Others	3	0	0	0	3
TC-/5, 5'	4	3	0	0	7
TC-/15, 15'	4	4	0	0	8
TC-/17, 17, 18	0	0	0	0	0
TC-/20, 20'	3	1	0	0	7
TC-/Others	3	1	0	0	4
Totals	92	37	4	6	139

Leaflet shape is noted of a total of 139 individuals. Leaflet shape classes: ov- ovate or-orbicular ro-retuse obovate, ob-obcordate.

1.5 Combined pattern types against leaflet size

It was observed from visual appearance and data collection that the recognised trends in leaflet size variation seen across the Mediterranean were observed in the data set and has the data set was too small to carry out a χ^2 assessment for full comparisons of each leaflet size category against combined pattern types no analysis was performed. Instead geographic association between zones 1 and 2 was tested for and this analysis is presented in the results section (Table 3.43).

DNA Sequences Alignments

Alignment 1: *trn L* intron

Alignment 2: IRB21NE

Alignment 3: ITS

Key:

Accession origins

A: Algeria

F: French

P: Portugal accession

S: Spanish accession

I: Italian accession

M: Morocco

Commercial varieties

B: Barbian

BL: Blanca

D: Dusi

H: Haifa

Mo: Morso

O: Ovack

TA: Tammistio

TR: Tamar

R: Regal

Aran: 1 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATG... ..

Aran: 2 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 3 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 4 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 5 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 6 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 7 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 8 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 9 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Aran: 10 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 1 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 2 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 3 CTCCACCACA G~YTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 4 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 5 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 7 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 8 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 9 CTCCACCACA G~TTT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

Crusader: 10 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 383 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 384 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 385 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 386 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 388 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 389 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 390 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

R: C15111: 391 CTCCACCACA G~TCT~GATG AATCTTTTGA ATAACTGATT AATCAGACGA GAATAAAGAT AGAGTCCCAT TCTACATGTC AATACCGACA

...|...| ...|...| ...|...| ...|...| ...|...| ...

545 555 565 575 585

REFERENCE SEQ HUIA

P: Estremadura: C7576: (2A) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Estremadura: C7576: (10B) TCAATGAAAT TTTTAGTAAG AGG..... ..

P: Estremadura: C7576: (177) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Estremadura: C7576: (200) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Estremadura: C7576: (11B)* TCAATGAAAT TTTTAGTAAG AGGAAAaTCC GTCGACTTTA GAAaTcGTGA GGG

P: Estremadura: C7576: (33) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Estremadura: C7576: (34)* TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (46) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GtCgACTTTa GAAATCGTga GGG

P: Ribatejo: C8979: (48) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GtCgACTTTa GAAATCGtGa GGG

P: Ribatejo: C8979: (47) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GtCgACTTTa GaAAATcGTGA GGG

P: Ribatejo: C8979: (175) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (188) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (186) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (187) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (195) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (196) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Ribatejo: C8979: (45) TCAATGAAAT TTTTAGTAAG AGGAAAATCC gTcGACTTta GAAATCGTGA GGG

P: Algarve: C8976: (62) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

P: Algarve: C8976: (64) TCAATGAAAT TTTTAGTA.. ..

P: Algarve: C8976: (36) TCAATGAAaT TTTTAGTAAG AGGAAAATCC GTCGACTTta GAAATCGTGA GGG
P: Algarve: C8976: (203) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Algarve: C8976: (204) TCAATGAAAT TTTTAGTAAG AGGAAAATCC
P: Algarve: C8976: (164) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Algarve: C8976: (65) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Beria literal: C7582: (89) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Beria literal: C7582: (91) TCAATGAAAT TTTTAGTAAG
P: Beria literal: C7582: (93) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCG...
P: Beria literal: C7582: (191) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Beria literal: C7582: (94) TCAATGAAAT TTTTAGT...
P: Beria literal: C7582: (95) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Beria literal: C7582: (96) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Beria literal: C7582: (24) TCAATGAAAT TTTTAGTAAG AGGAAAaTCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (12) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (13) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (99) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (12) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (12) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (12) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Trans Os Montes: C9514: (12) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (140) TCAATGAAAT TTTTAGTAAG AGGAAA-TCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (141) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (12B) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GaaATCGTGA GGG
P: Coimbra: C8985: (13B) TCAATGAAaT TTTtAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (132) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (134) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (135) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (136) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (137) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (138) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Coimbra: C8985: (3A) TCAATGAAAT TTTtAGTAAG AGGAAAaTCC GTCGACTTTA GAAaTCGTGA GGG
P: Coimbra: C8985: (139) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Agus de Mouco: C20477: (154) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Agus de Mouco: C20477: (153) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Agus de Mouco: C20477: (155) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Agus de Mouco: C20477: (156) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Agus de Mouco: C20477: (157) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-1 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-3 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-5 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-6 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-7 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-8 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Vila Verde: C9516-10 TCAATGAAAT TTTTAGTAAG AGGAAA~TCC GTCGACTTTA GAAATCGTGA GGG
P: Espfino: C9024-1 TCAATGAAAT TTTTAGTAAG AGGAAA~TCC GTCGACTTTA GAAATCGTGA GGG
P: Espfino: C9024-3 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Espfino: C9024-4 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Espfino: C9024-5 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Espfino: C9024-8 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
P: Espfino: C9024-10 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

S: Galicia: C9037: (29) TCAATGAAAT TTTtAgtaAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Andalucía: C8975: (28) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Andalucía: C8975: (26) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Andalucía: C8975: (52)* TCAATGAAAT TTTTAGTARG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Andalucía: C8975: (3B) TCAATGAAAT TTTTAGTAAG AG..... ..
S: Extremadura: C9025: (206) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GT..... ..
S: Extremadura: C9025: (49) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Extremadura: C9025: (50) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Extremadura: C9025: (51) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGAC.... ..
S: Extremadura: C9025: (73)* TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Extremadura: C9025: (75) TCAATGA... ..
S: Extremadura: C9025: (25) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTa GAAATCGTGA GGG
S: Ponferrada: C9531-1 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-2 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-3 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-4 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-5 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-6 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-7 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-8 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Ponferrada: C9531-9 TCAATGAAAT TTT~AGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-2 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-4 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-5 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-6 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-8 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-9 TCAATGAAAT TTTTAGTAAG AGGAAA~TCC GTCGACTTTA GAAATCGTGA GGG
S: Pontevedra: C9027-10 TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Mouriès: C9046: (162) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Mouriès: C9046: (163) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Mouriès: C9046: (192) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Mouriès: C9046: (202) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Mouriès: C9046: (40) TCAATGAAAT TTTTAGTAAG AGGAAAATCC gTcGA..... ..
F: Mouriès: C9046: (41) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTa GAAATCGTGA GGG
F: Mouriès: C9046: (42) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Mouriès: C9046: (43) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTa gaAATCGTGA GGG
F: Mouriès: C9046: (7A) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Entressen: C7690: (53) TCAATGAAAT TTTTAGTAAG AGGAAAATC..... ..
F: Entressen: C7690: (54) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Entressen: C7690: (78) TCAATGAAaT TTTtAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Entressen: C7690: (207) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (146) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GG.
F: Crau: C18951: (4A) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (143) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (144) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (147) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (148) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (149) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (150) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG
F: Crau: C18951: (151) TCAATGAAAT TTTTAGTAAG AGGAAAATCC GTCGACTTTA GAAATCGTGA GGG

F: Crau: C18951: (131)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
F: Crau: C18951: (121)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (8B)	TCAATGAAAT	TTTTAGTAAG	AGGAAAaTCC	gTCgACTTta	GaaATCGTGA	GGG
E: Kent: C6374: (81)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGA.....
E: Kent: *C6374: (82)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (84)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (85)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (86)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (87)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (197)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6374: (198)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6304-2	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6304-3	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6304-5	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6304-6	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
E: Kent: C6304-9
E: Kent: C6304-10
I: *C5897: (17)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897: (18)	TCAATGAAaT	TTTTaGTAAg	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897: (100)	TCAATGAAaT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTa	GAAATCGTGA	GGG
I: *C5897: (103)
I: C5897: (104)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897: (105)	TCAATGAAAT	TTTTAGTAAG	aGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897: (208)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897-1	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897-2	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897-4	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897-5	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897-6	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAA.....	...
I: C5897-9	TCAATGAAAT	TTTTAGGTAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
I: C5897-10	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C20499: (1A)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (165)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (166)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (167)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (168)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (169)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (170)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
M: C1761: (171)	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-1	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-3	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-4	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-5	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-6	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-7	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-8	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-9	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
A: C6452-10	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGR	GGG
TA: C960: 300	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
TA: C960: 301	TCAATGAAAT	TTTTAGTAAG	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG

R: C15111: 385	TCAATGAAAT	TTTtagtaag	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
R: C15111: 386	TCAATGAAAT	TTTtagtaag	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
R: C15111: 388	TCAATGAAAT	TTTtagtaag	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
R: C15111: 389	TCAATGAAAT	TTTtagtaag	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
R: C15111: 390	TCAATGAAAT	TTTtagtaag	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG
R: C15111: 391	TCAATGAAAT	TTTtagtaag	AGGAAAATCC	GTCGACTTTA	GAAATCGTGA	GGG

S: Galicia: C9037: 193	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
S: Galicia: C9037: 194	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
S: Galicia: C9037: 12A	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Galicia: C9037: 1B	ATTCTTATAG AGCAA~GAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Galicia: C9037: 201	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pontevedra: C9027-2	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGAAAGCTC TTCGTTCC.
S: Pontevedra: C9027-3	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pontevedra: C9027-4	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pontevedra: C9027-6	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pontevedra: C9027-8	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pontevedra: C9027-9	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pontevedra: C9027-10	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Ponferrada: C9531-1	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Ponferrada: C9531-2	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Ponferrada: C9531-4	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Ponferrada: C9531-5	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Ponferrada: C9531-6	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Andalucía: C8975: 26	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Andalucía: C8975: 27	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Andalucía: C8975: 28	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Andalucía: C8975: 3B	ATTCTTA~AG AGCAA~GAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pyrenees: C7600: 1A	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pyrenees: C7600: 145	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pyrenees: C7600: 123	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pyrenees: C7600: 117	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pyrenees: C7600: 120	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAgCTCT... ..
S: Pyrenees: C7600: 119	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTtCG TtCCCGGT.
S: Pyrenees: C7600: 118	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
S: Pyrenees: C7600: 122	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Beria literal: C7582: 24	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Beria literal: C7582: 89	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Beria literal: C7582: 90	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Beria literal: C7582: 91	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Beria literal: C7582: 92	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAG~~~CTCT TCgTTCC..
P: Beria literal: C7582: 96	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGG... ..
P: Beria literal: C7582: 191	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGA.
P: Estremadura: C7576: 177	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
P: Estremadura: C7576: 11B	ATTCTTA~AG AGCAA~GAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Estremadura: C7576: 10B	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Estremadura: C7576: 2A	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Estremadura: C7576: 32	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Estremadura: C7576: 33	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Estremadura: C7576: 34
P: Algarve: C8976: 164	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Algarve: C8976: 203	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
P: Algarve: C8976: 204	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Algarve: C8976: 65	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Algarve: C8976: 35	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Algarve: C8976: 36	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.

P: Ribatejo: C8979: 188	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
P: Ribatejo: C8979: 195	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGTT.
P: Ribatejo: C8979: 196	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
P: Ribatejo: C8979: 187	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGg .
P: Ribatejo: C8979: 175	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Ribatejo: C8979: 49	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Ribatejo: C8979: 47	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGTA
P: Ribatejo: C8979: 186	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Ribatejo: C8979: 175	ATTCTTATAG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
P: Coimbra: C8985: 13	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCG...
P: Coimbra: C8985: 132	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 133	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 141	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 140	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 139	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 138	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 137	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 136	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 135	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 134	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 12B	ATTCTTA~AG AGCAA~GAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Coimbra: C8985: 3A	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 99	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 124	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 125	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 126	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 127	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 128	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 129	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Trans Os Montes: C9514: 130	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Agus de Mouco: C20477: 152	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Agus de Mouco: C20477: 153	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCG...
P: Agus de Mouco: C20477: 154	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Agus de Mouco: C20477: 156	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Agus de Mouco: C20477: 157	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Vila Verde: C9516-1	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Vila Verde: C9516-4	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Vila Verde: C9516-8	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
P: Vila Verde: C9516-9	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Mouriès C9046: 43	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Mouriès C9046: 163	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Mouriès C9046: 162	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Mouriès C9046: 202	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Mouriès C9046: 40	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Mouriès C9046: 192	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
F: Entressen: C7690: 54	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGGT.
F: Entressen: C7690: 53	ATTCTTA~AG AGCAAAGAGT CGGGSSGAAA AAGGRGGGAG AAGCTCTTCG TTCCCGGT.
F: Entressen: C7690: 78	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
F: Entressen: C7690: 80	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.
F: Entressen: C7690: 207	ATTCTTA~AG AGCAAAGAGT CGGGCGGAAA AAGGGGGGAG AAGCTCTTCG TTCCCGgT.

354: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
355: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
375: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
378: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
377: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
355: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
357: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAR	AAGCTYTTTCG	TTCCCGGT.
358: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
359: C6360: Blanca	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
360: C7555: Ovack	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
361: C7555: Ovack	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
363: C7555: Ovack	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
362: C7555: Ovack	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
364: C7555: Ovack	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 300	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 301	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 302	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 303	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 304	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 305	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 306	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
B: C960: 307	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
S.184: C21105: 308	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
S.184: C21105: 309	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
A: C6452: 332	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
325: C4143: Tamar	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 383: C15111	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 384: C15111	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 385: C15111	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 386: C15111	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 387: C15111	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 388: C15111	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 389: C15111	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 390: C15111	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGRRAG	AAGCTCTTCG	TTCCCGGT.
R: 393: C15111	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
R: 391: C15111	ATTCTTA~AG	AGCAA~GAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.
D: 372: C16511	ATTCTTA~AG	AGCAAAGAGT	CGGGCGGAAA	AAGGGGGGAG	AAGCTCTTCG	TTCCCGGT.

Appendix: Alignment 3: ITS Nuclear DNA sequences

	5	15	25	35	45	55	65	75	85
GRASSLANDS HUIA	AAGGTTTCCG	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Estremadura: C7576: (177)	~~~~~	~~~~~CC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (96)	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (94)	~~~~~	~~~~~	~~~~~AAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (24)	~~~~~	~~~~~	~~~~~AAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (95)	~~~~~	~~~~~	~~~~~CGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (93)	~~~~~	~~~~~	~~~~~	~CATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (89)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~TTTGAACAC	ATAGGGTTGG
P: Beria literal: C7582: (191)	~~~~~	~~~~~AACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACA~	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Algarve: C8976: (204)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Algarve: C8976: (203)	~~~~~	~~~~~AACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Algarve: C8976: (65)	~~~~~	~~~~~	~~~~~CGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Ribatejo: C8979: (196)	~~~~~	~~~~~	~~~~~GGA	TCATTGTCGA	TGCCTTACCT	GCAGACCA~C	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Ribatejo: C8979: (47)	~~~~~	~~~~~	~~~~~GCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Ribatejo: C8979: (48)	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Ribatejo: C8979: (188)	~~~~~	~~~~~CC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Ribatejo: C8979: (186)	~~~~~	~~~~~	~~~~~CGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Ribatejo: C8979: (187)	~~~~~	~~~~~	~~~~~GGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Coimbra: C8985: (134)	~~~~~	~~~~~ACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Esphino: C9024-3	~~~~~	~~~~~	~~~~~	~~~~~GA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Esphino: C9024-4	~~~~~	~~~~~	~~~~~GGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Esphino: C9024-5	~~~~~	~~~~~	~~~~~GGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Esphino: C9024-6	~~~~~	~~~~~	~~~~~	~~~~~CGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Esphino: C9024-7	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Esphino: C9024-9	~~~~~	~~~~~	~~~~~	~~~~~	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	~TAGGGTTGG
P: Esphino: C9024-10	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	NGCCTTACAN	GCRRACCAAC	ACGTGAATCA	GTTTGAACAC	MTAGGGTTGG
P: Villa Verde: C9516-1	~~~~~	~~~~~	~~~~~GGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-10	~~~~~	~~~~~	~~~~~	~~~~~GA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-2	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	GTTTGAACAC	~TAGGGTTGG
P: Villa Verde: C9516-3	~~~~~	~~~~~C	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-4	~~~~~	~AGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-5	~~~~~	~~~~~	~~~~~AAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-6	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	~TAGGGTTGG
P: Villa Verde: C9516-7	~~~~~	~~~~~	~~~~~GGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-8	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Villa Verde: C9516-9	~~~~~	~~~~~	~~~~~GCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Agus de Mouco: C20477: (157)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAA	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Agus de Mouco: C20477: (152)	~~~~~	~~~~~	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Agus de Mouco: C20477: (153)	~~~~~	~~~~~TGAAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG

P: Agus de Mouco: C20477: (154	~~~~~	~~~~~C	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: C9514: Tras Os Montes: (124	~~~~~	~~~~~	~~~~GAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: C9514: Tras Os Montes: (127	~~~~~	~~~~~	~GCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: C9514: Tras Os Montes: (130	~~~~~	~~~~TGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: C8985: Coimbra: (139)	~~~~~	~~~~~	~~~~AAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: C8985: Coimbra: (140)	~~~~~	~~~~~	~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: C8985: Coimbra: (141)	~~~~~	~~~~~	~~~~AGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Tras Os Montes: C9514: (99)	~~~~~	~~~~~	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: Beira litoral: C7582: (92)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~CCAAC	ACGTGAATCA	GTTTGAACMC	ATAGGGTTGG
P: Villa Verde: 9516-1	~~~~~	~~~~~	~~~~~	~~~~ATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4838: Minho: No.1	~~~~~	~~~~~	~~~~GGA	TCATTGTCGA	TGCCTTNCNT	GCNGACCAAC	NCGTGAATCA	CTTTGAACAC	ATAGGGTTGG
P: AZ4838: Minho: No.3	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4838: Minho: No.4	~~~~~	~~~~~	~~~~GGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4838: Minho: No.5	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~CGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4389: Baragonça: NO.1	~~~~~	~~~~~	~~~~GGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4389: Baragonça: NO.4	~~~~~	~~~~~	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	STTTGAACAC	ATAGGGTTGG
P: AZ4389: Baragonça: NO.2	~~~~~	~~~~~	~GCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4389: Baragonça: NO.5	~~~~~	~~~~~	TGCGGAAGGA	TCATTGTCGA	TGCCTTACNT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
P: AZ4389: Baragonça: NO.6	~~~~~	~~~~~	~~~~~	~~~~~	~~~~ACAT	GCAGACCAAC	ACGTGAATCA	CTTTGAACAC	ATAGGGTTGG
S: Extremadura: C9025: (76)	~~~~~	~~~~~	~~~~~	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Extremadura: C9025: (75)	~~~~~	~~~~~	~~~~~	~~~~TCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Extremadura: C9025: (206)	~~~~~	~~~~~	~~~~AAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAaC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Andalucía: C8975: (27)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~TAGGGTTGG	
S: Andalucía: C8975: (52)	~~~~~	~~~~~ACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCA~C	ACGTGNATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (11A)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~ACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (190)	~~~~~	~~~~~	~CGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (189)	~~~~~	~~~~TGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (210)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (211)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (106)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~AC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (108)	~~~~~	~~~~~	~~~~GAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Asturias: C8994: (107)	~~~~~	~~~~~	~~~~~CGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG	
S: Asturias: C8994: (199)	~~~~~	~~~~GTGAACC	TGCGGAAGGA	TCATTGTCGA	TSCCTTACA~	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9037: (194)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCA~C	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9037: (193)	~~~~~	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCA~C	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9037: (201)	~~~~~	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9037: (161)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~TCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (71)	~~~~~	~~~~TGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (67)	~~~~~	~~~~~	~GCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (70)	~~~~~	~~~~ACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (55)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAaC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (58)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~T	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (66)	~~~~~	~~~~~	~~~~~	~~~~~GA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (178)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~ACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (205)	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
S: Galicia: C9034: (88)	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~TGAATCA	NTTTGAACAC	ATAGGGTTGG

S: Galicia: C9034: (68) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCA~C ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Galicia: C9034: (56) ~~~~~~CG TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACKTGAATCA GTTTGAACAC ATAGGGTTGG
S: Galicia: C9034: (72) AAGGTTTCCG TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Galicia: C9034: (56) ~~~~~~CGTA GGTGAACCTG CGGAAGGATC ATTGTCGATG CCTTACATGC AGACCAACAC KTGAATCAGT TTGAACACAC ATAGGGTTGG
S: Galicia: C9037: (160) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Galicia: C9037: (160) ~~~~~~CC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pyrenees: C7600: (21) ~~~~~~CC TGCGGAAGGA TCATTGTCGA TGCCTTACAN GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Extremadura: C9025: (51) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: C9531: (59n) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Ponferrada: C9531-3 ~~~~~~AAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ~TAGGGTTGG
S: Ponferrada: C9531-4 ~~~~~~GAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Ponferrada: C9531-5 ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Ponferrada: C9531-6 ~~~~~~ACC TGCGGAAGGA TCATTGTCR~ TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Ponferrada: C9531-7: ~~~~~~C TGCGGAAGGA TCATTGTCNA TNCCTTNCNT ~~~~ACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Ponferrada: C9531-9 ~~~~~~CC TGCGGAAGGA TCATTGTCNA TNCCTTNCNT CCGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-2 AAGGTTTCCG TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGNACAC ATAGGGTTGG
S: Pontevedra: C9027-3 ~~~~~~AAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-4 ~~~~~~TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-5 ~~~~~~AGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-6 ~~~~~~GGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-8 ~~~~~~TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-9 ~~~~~~GTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pontevedra: C9027-10 ~~~~~~GAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pyrenees: C7600: (117) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAG GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pyrenees: C7600: (20) ~~~~~~CGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
S: Pyrenees: C7600: (21) ~~~~~~CC TGCGGAAGGA TCATTGTCGA TGCCTTACAA GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Entressen: C7690: (53) ~~~~~~TGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Entressen: C7690: (54) ~~~~~~G TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Entressen: C7690: (207) ~~~~~~G TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAaC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Entressen: C7690: (78) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCA~C ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Mouriès: C9046: (202) ~~~~~~CC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Mouriès: C9046: (192) ~~~~~~G TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Mouriès: C9046: (40) ~~~~~~GAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Mouriès: C9046: (163) ~~~~~~C TGCGGAAGGA TCATTGTCNA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (150) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (4A) ~~~~~~AAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ATGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (151) ~~~~~~GTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (121) ~~~~~~CTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (144) ~~~~~~GAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (146) ~~~~~~GGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (147) ~~~~~~TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (148) ~~~~~~GGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
F: Crau: C18951: (149) ~~~~~~GAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (82) ~~~~~~AA GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (84) ~~~~~~C TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (197) ~~~~~~CTTACAT GCAGACCAAC NCGTGAATCA GTTTGACCAC ATAGGGTTGG

E: Kent: C6374: (85) ~~~~~~CG TAGGTGAACC TGCGGAAGGA TCATGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (198) ~~~~~~ ~~~GTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (81) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~CTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (87) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~ATTGTTCGA TGCCTTACAW GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (86) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~CC TGCGGAAGGA TCATTGTTCGA TGCCTTACAN GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6374: (85) ~~~~~~ ~~~~~~CG TAGGTGAACC TGCGGAAGGA TCATGTTCGA TSCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6304-3 ~~~~~~ ~~~~~~TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6304-4 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6304-5 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~AAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6304-6 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~AGGA TCATTGTTCGA TGCCTTACAW GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6304-9 ~~~~~~ TAGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
E: Kent: C6304-10 ~~~GTTTCCG TAGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
I: C5897-4 ~~~~~~ ~~~~~~TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
I: C5897-6 ~~~~~~ ~~~~~~ ~~~~~~CGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGGATCA GTTTGAACAC ATAGGGTTGG
I: C5897-8 ~~~~~~ ~~~~~~ ~~~~~~AGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
I: C5891-9 ~~~~~~ ~~~~~~ ~~~~~~GGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
I: C5897-10 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
I: C5897: (208) ~~~~~~ ~~~~~~G TAGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACaT GCAGACCA-C ACGTGAATCA GTTTGAACAC ATAGGGTTGG
I: C5897: (104) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~GTCCGAA TCCCTTACCT GCAGACCNAC ACGTGAATCN GTTTGAACAC ATAGGGTTGG
I: C5987: (105) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~TTGG
I: C5897: (19) ~~~~~~ ~~~~~~ ~~~~~~TTGTCGG ATGCCTTACA AGCAGGACCC AACACGTGAA TCCAGTTGA ACCACCATAG GGTGGTTGG
I: C5897: (102) ~~~~~~ TAGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAM ATAGGGTTGG
I: C5897-2n ~~~~~~ ~~~~~~AGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACCT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
I: C5897: (103) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
I: C5897: (18) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
M: C1761: (170) ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
M: C1761: (167) ~~~~~~ ~~~~~~ ~~~~~~AANN TGCGGAAGGA TCATTGTTCNA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
M: C1761: (168) ~~~~~~ ~~~~~~ ~~~~~~TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
M: C1761: (169) ~~~~~~ ~~~~~~ ~~~~~~TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
A: C6452-3 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
A: 6452-10 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
TA: C960: 300 ~~~~~~ ~~~~~~ ~~~~~~GAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 301 ~~~~~~ ~~~~~~ ~~~~~~ACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 302 ~~~~~~ ~~~~~~ ~~~~~~GTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 303 ~~~~~~ ~~~~~~ ~~~~~~AGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 304 ~~~~~~ ~~~~~~ ~~~~~~GTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 305 ~~~~~~ ~~~~~~ ~~~~~~AACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 306 ~~~~~~ ~~~~~~ ~~~~~~CC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
TA: C960: 307 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
TA: C960: 369 AAGGTTTCCG TAGGTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
B: C24247: 314 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
B: C24247: 315 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
B: C24247: 316 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
B: C24247: 317 ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~ ~~~~~~
B: C24247: 311 ~~~~~~ ~~~~~~ ~~~~~~AACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
B: C24247: 310 ~~~~~~ ~~~~~~ ~~~~~~GTGAACC TGCGGAAGGA TCATTGTTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG

364 C7555 Ovcak	~AGGTTTCCG	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
365 C7555 Ovcak	~~~~~	~~~~~	~~~~~AaGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran1	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran3	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran4	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran5	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran6	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran7	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran8	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran9	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCGA	TNCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Aran10	~~~~~	~~~~~	~GCGGAAGGA	TCATTNTCNA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader1	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~AC	ATAGGGTTGG
Crusader2	~~~~~	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader3	~~~~~	~~~GTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader4	~~~~~	~AGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader5	~~~~~	~~~GGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader6	~~~~~	~~~~~C	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader8	~~~~~	~~~~~	~~~~~A	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
Crusader9	~~~~~	~~~~~	~~~~~TCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG	
Cruasder10	~~~~~G	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 383	~~~~~	~~~~~	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 384	~~~~~	TAGGTGAACC	TGCGGAAGGA	TCATTGTCNA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 385	~~~~~	~AGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAA	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 387	~~~~~	~~~~~TGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 388	~~~~~	~~~GTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 389	~~~~~CG	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 390	~~~~~	~~~GTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
R: C15111: 391	~~~~~	~~~~~	~~~~~	~~~~~TTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG	
A: C6452: 322	~~~~~	~~~~~GAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ATGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C6452: 3232	~~~~~	~~~~~GAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ATGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C6452-5	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~GGGTGG	
A: C6452-6	~~~~~	~~~~~	~~~~~	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ANGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C6452-8	~~~~~	~~~~~	~GCGGAAGGA	TCATTGTCGA	TKCCTTMCAT	GCAGACCAAC	NYGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C6452-9	~~~~~	~~~~~	~GCGGAAGGA	TCATTGTCGA	TSCCTTWCAT	GCAGACCAAC	ATGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.1	~~~~~	~~~GGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.2	~~~~~	~~~~~TGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.3	~~~~~	~~~GGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAA	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.4	~~~~~	~~~~~AACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ATGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.5	~~~~~	~~~~~GGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	AYGTGAATCA	GTTTGAACAC	ATAGGGTTGG	
A: C17561 No.6	~~~~~	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	AYGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.7	~~~~~	~~~~~GAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	AYGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 No.8	~~~~~	~~~~~	~~~~~AGGA	TCATTGTCNA	TGCCTTACAT	GCAGACCAAC	AYGTGAATCA	GTTTGAACAC	ATAGGGTTGG
A: C17561 10	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~CAGACCAAC	AYGTGAATCA	GTTTGAACAC	ATAGGGTTGG
D: C16511: 372	~~~~~	~~~~~	~~~~~	~~~~~TTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
D: C16511: 319	~~~~~CG	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG
D: C16511: 320	~~~~~CCG	TAGGTGAACC	TGCGGAAGGA	TCATTGTCGA	TGCCTTACAT	GCAGACCAAC	ACGTGAATCA	GTTTGAACAC	ATAGGGTTGG

D: C16511: 318 ~~~~~~ ~~~~~~ ~GCGGAAGGA TCATTGTCGA TGCCTTACAA GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
D: C16511: 319 ~~~~~~ ~~~~~~ ~~~~~~GGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG
D: C16511: 371 ~AGGTTTCCG TAGGTGAACC TGCGGAAGGA TCATTGTCGA TGCCTTACAT GCAGACCAAC ACGTGAATCA GTTTGAACAC ATAGGGTTGG

....|....||....||....||....||....||....||....||....|
95 105 115 125 135 145 155 165 175

GRASSLANDS HUIA
P: Estremadura: C7576: (177) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (96) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (94) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (24) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (95) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (93) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (89) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Beria literal: C7582: (191) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Algarve: C8976: (204) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Algarve: C8976: (203) TTTGAGGTGT TCAACACCTC AGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Algarve: C8976: (65) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Ribatejo: C8979: (196) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Ribatejo: C8979: (47) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Ribatejo: C8979: (48) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Ribatejo: C8979: (188) TTTGAGGTGT TCAACACCTC AGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Ribatejo: C8979: (186) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Ribatejo: C8979: (187) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Coimbra: C8985: (134) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Esphino: C9024-3 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Esphino: C9024-4 TTTGAGGTGT TCCACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Esphino: C9024-5 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Esphino: C9024-6 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Esphino: C9024-7 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Esphino: C9024-9 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-1 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-10 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-2 ~TTGAGGTGT TCAACACCTC GGCTTSCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CRAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-3 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-4 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-5 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-6 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-7 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-8 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Villa Verde: C9516-9 TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Agus de Mouco: C20477: (157) TTTGAGGTGT TCAACACCTC AGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Agus de Mouco: C20477: (152) TTTGAGGTGT TCAACACCTC GGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Agus de Mouco: C20477: (153) TTTGAGGTGT TCAACACCTC AGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
P: Agus de Mouco: C20477: (154) TTTGAGGTGT TCAACACCTC RGCTTGCCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC

D: C16511: 319 TTTGAGGTGT TCAACACCTC GGCTTGCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC
D: C16511: 371 TTTGAGGTGT TCAACACCTC GGCTTGCCTC TGGTTCGGAG GAGGACGATG TTGTGCGTTC TCCTTTGTGC CAAAACACAA ACCCCGGCGC

....|....||....||....||....||....||....||....||....|
185 195 205 215 225 235 245 255 265

GRASSLANDS HUIA
P: Estremadura: C7576: (177) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Beria literal: C7582: (96) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Beria literal: C7582: (94) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Beria literal: C7582: (24) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Beria literal: C7582: (95) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACaTaaT
P: Beria literal: C7582: (93) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Beria literal: C7582: (89) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Beria literal: C7582: (191) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Algarve: C8976: (204) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Algarve: C8976: (203) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Algarve: C8976: (65) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Ribatejo: C8979: (196) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Ribatejo: C8979: (47) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Ribatejo: C8979: (48) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Ribatejo: C8979: (188) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Ribatejo: C8979: (186) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Ribatejo: C8979: (187) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Coimbra: C8985: (134) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-3 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-4 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-5 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-6 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-7 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-9 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Esphino: C9024-10 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-1 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-10 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-2 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-3 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-4 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-5 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-6 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-7 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-8 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Villa Verde: C9516-9 TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Agus de Mouco: C20477: (157) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Agus de Mouco: C20477: (152) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Agus de Mouco: C20477: (153) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: Agus de Mouco: C20477: (154) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT
P: C9514: Tras Os Montes: (124) TGAATGCGTC AAGGAATTTA AAATTTGCTC TGAGCGCACC TGCATGGCAC CGGAGACGGT TTTCGTGCGG GTTGTGTTCT GACACATAAT

P: Tras Os Montes: C9514: (99) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGC.

P: Beira litoral: C7582: (92) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCC....

P: Villa Verde: 9516-1 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATA..

P: AZ4838: Minho: No.1 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG

P: AZ4838: Minho: No.3 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CG.....

P: AZ4838: Minho: No.4 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGAAAAA

P: AZ4838: Minho: No.5 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGG....

P: AZ4389: Baragonça: NO.1 ACCCACATGC GTTTCGAAAC TCTCATGATG AGACCTCAGG TCAGGNGGGG CT.....

P: AZ4389: Baragonça: NO.4 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGG....

P: AZ4389: Baragonça: NO.2 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCATAAGC GGAGG....

P: AZ4389: Baragonça: NO.5 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGG....

P: AZ4389: Baragonça: NO.6 ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGG.....

S: Extremadura: C9025: (76) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG

S: Extremadura: C9025: (75) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGG....

S: Extremadura: C9025: (206) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Andalucía: C8975: (27) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TAT.....

S: Andalucía: C8975: (52) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Asturias: C8994: (11A) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCCTT AATTT.....

S: Asturias: C8994: (190) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Asturias: C8994: (189) ACCCACATGC GTTTCGAAAC GCTCGTGATG AAACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Asturias: C8994: (210) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Asturias: C8994: (211) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Asturias: C8994: (106) ACCCACATGC GTTTCGAAAC GCT.....

S: Asturias: C8994: (108) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Asturias: C8994: (107) ACCCACATGC GTTTCGAAAC GCT.....

S: Asturias: C8994: (199) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCCATAAG CGGAGGAA..

S: Galicia: C9037: (194) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAA..

S: Galicia: C9037: (193) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9037: (201) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGG....

S: Galicia: C9037: (161) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGG.....

S: Galicia: C9034: (71) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAG....

S: Galicia: C9034: (67) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAA.

S: Galicia: C9034: (70) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAA..

S: Galicia: C9034: (55) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9034: (58) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCATTAAG CGGAGGAA..

S: Galicia: C9034: (66) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAAT... ..

S: Galicia: C9034: (178) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9034: (205) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9034: (88) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG

S: Galicia: C9034: (68) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9034: (56) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9034: (72) ~CCCACATSC STTTCGAAAC STTTTTGWTS ARACCKNANG NCGGSGGGG CTNCCCGCWA AATTCASCA TWTCTC... ..

S: Galicia: C9034: (56) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Galicia: C9037: (160) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG GCTACCC... ..

S: Galicia: C9037: (160) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA

S: Pyrenees: C7600: (21) ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAA..


```

E: Kent: C6304-4      ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATA..
E: Kent: C6304-5      ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CG.....
E: Kent: C6304-6      ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG C.....
E: Kent: C6304-9      ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
E: Kent: C6304-10     ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
I: C5897-4            ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTA~GCA TATCAATAAG CGGAGGAAAA
I: C5897-6            ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TA.....
I: C5897-8            ACCCACATGC GTTTCGAAAC GCTCRTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA ATATCAATAA GCGGAGG...
I: C5891-9            ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTA~GCA TATCAATAAG CGG.....
I: C5897-10           ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCSCTG AATTTAAGCA TATC~~~~~ ~~~~~
I: C5897: (208)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
I: C5897: (104)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG GATTCAAGCA TNTATCTTAG CGGAGGGAAA
I: C5897: (105)       TACCCACATG CGTTTCGAAA CGCTYGTGAT GAGACCTCAG GTCAGGCGGG GCTNCCCGCW GGATTAAGCA ATNTAWCTTA GCGGAGGTA
I: C5897: (19)        TACCCACATG CGTTTCGAAA CGCTTGTGAT GAGACC....
I: C5897: (102)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGG....
I: C5897-2n           ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
I: C5897: (103)       ACCCACATGC GTTTCGAAAC GCTCRTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG RATTYAAGCA TRTCAA....
I: C5897: (18)        ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGA...
M: C1761: (170)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGAC~TCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGC.
M: C1761: (167)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
M: C1761: (168)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
M: C1761: (169)       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
A: C6452-3            ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA .....
A: 6452-10           ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
TA: C960: 300         ACCCACATGC GTTTCGAAAC GCTGGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AA.....
TA: C960: 301         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATTAG CGGAGGAAAA
TA: C960: 302         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCA.....
TA: C960: 303         .....
TA: C960: 304         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
TA: C960: 305         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
TA: C960: 306         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGA...
TA: C960: 307         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAA..
TA: C960: 369         ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
B: C24247: 314       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAA..
B: C24247: 315       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TAT.....
B: C24247: 316       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAG....
B: C24247: 317       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
B: C24247: 311       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATA..
B: C24247: 310       ACCCACATGC GTTTCGAAAC GCT.....
B: C24247: 312       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAA..
B: C24247: 313       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
B: C24247: 349       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAWTAAG CGGAGGAAAA
B: C24247: 373       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCATTAGC GGAGGAAAAG
B: C24247: 374       ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTTAAGCA TATCAATAAG CGGAGGAAAA
S.184: C21105: 308   .....
S.184: C21105: 309   ACCCACATGC GTTTCGAAAC GCTCGTGATG AGACCTCAGG TCAGGCGGGG CTACCCGCTG AATTAAGCAT .....

```


P: Estremadura: C7576: (177) GAAACTAACA AGGA
P: Beria literal: C7582: (96)
P: Beria literal: C7582: (94)
P: Beria literal: C7582: (24)
P: Beria literal: C7582: (95)
P: Beria literal: C7582: (93)
P: Beria literal: C7582: (89)
P: Beria literal: C7582: (191)
P: Algarve: C8976: (204) GAAACT.....
P: Algarve: C8976: (203) GAAAC.....
P: Algarve: C8976: (65)
P: Ribatejo: C8979: (196) GA.....
P: Ribatejo: C8979: (47)
P: Ribatejo: C8979: (48)
P: Ribatejo: C8979: (188) GAAACTAACA AG..
P: Ribatejo: C8979: (186) GAAAC.....
P: Ribatejo: C8979: (187) GAAACTAACA AGG1
P: Coimbra: C8985: (134) GAAACTAACA AGGA
P: Esphino: C9024-3
P: Esphino: C9024-4
P: Esphino: C9024-5
P: Esphino: C9024-6
P: Esphino: C9024-7
P: Esphino: C9024-9
P: Esphino: C9024-10
P: Villa Verde: C9516-1
P: Villa Verde: C9516-10 GAAACTAAC.
P: Villa Verde: C9516-2
P: Villa Verde: C9516-3 GAAACTAACA
P: Villa Verde: C9516-4
P: Villa Verde: C9516-5
P: Villa Verde: C9516-6
P: Villa Verde: C9516-7
P: Villa Verde: C9516-8
P: Villa Verde: C9516-9
P: Agus de Mouco: C20477: (157) GAAACTAACA AGG.
P: Agus de Mouco: C20477: (152)
P: Agus de Mouco: C20477: (153)
P: Agus de Mouco: C20477: (154)
P: C9514: Tras Os Montes: (124)
P: C9514: Tras Os Montes: (127) GAAACTA... ..
P: C9514: Tras Os Montes: (130) GAAACTA... ..
P: C8985: Coimbra: (139)
P: C8985: Coimbra: (140) GAA.....
P: C8985: Coimbra: (141) GAAACTAA... ..
P: Tras Os Montes: C9514: (99)

P: Beira litoral: C7582: (92)
P: Villa Verde: 9516-1
P: AZ4838: Minho: No.1
P: AZ4838: Minho: No.3
P: AZ4838: Minho: No.4 GAAACTAAC.
P: AZ4838: Minho: No.5
P: AZ4389: Baragonça: NO.1
P: AZ4389: Baragonça: NO.4
P: AZ4389: Baragonça: NO.2
P: AZ4389: Baragonça: NO.5
P: AZ4389: Baragonça: NO.6
S: Extremadura: C9025: (76)
S: Extremadura: C9025: (75)
S: Extremadura: C9025: (206) GAAACTAACA A...
S: Andalucía: C8975: (27)
S: Andalucía: C8975: (52)
S: Asturias: C8994: (11A)
S: Asturias: C8994: (190) GAAACTA...
S: Asturias: C8994: (189) GAAAC.....
S: Asturias: C8994: (210) GAAAC.....
S: Asturias: C8994: (211) GAAACTAACA
S: Asturias: C8994: (106)
S: Asturias: C8994: (108) G.....
S: Asturias: C8994: (107)
S: Asturias: C8994: (199)
S: Galicia: C9037: (194)
S: Galicia: C9037: (193) GAAAC.....
S: Galicia: C9037: (201)
S: Galicia: C9037: (161)
S: Galicia: C9034: (71)
S: Galicia: C9034: (67)
S: Galicia: C9034: (70)
S: Galicia: C9034: (55) GAAACTA...
S: Galicia: C9034: (58)
S: Galicia: C9034: (66)
S: Galicia: C9034: (178) GAAAC.....
S: Galicia: C9034: (205) GAAACT.....
S: Galicia: C9034: (88)
S: Galicia: C9034: (68) G.....
S: Galicia: C9034: (56) GA.....
S: Galicia: C9034: (72)
S: Galicia: C9034: (56) GA.....
S: Galicia: C9037: (160)
S: Galicia: C9037: (160) GAAACTA...
S: Pyrenees: C7600: (21)
S: Extremadura: C9025: (51)

S: C9531: (59n)
S: Ponferrada: C9531-3
S: Ponferrada: C9531-4	GAAACT....
S: Ponferrada: C9531-5
S: Ponferrada: C9531-6	GAAACTA....
S: Ponferrada: C9531-7:
S: Ponferrada: C9531-9	GAAACT....
S: Pontevedra: C9027-2
S: Pontevedra: C9027-3	GAAACTA....
S: Pontevedra: C9027-4	GAAACTAAC.
S: Pontevedra: C9027-5	GAAA.....
S: Pontevedra: C9027-6	GAAACTAAC.
S: Pontevedra: C9027-8
S: Pontevedra: C9027-9	GAAACTAACA AGGA
S: Pontevedra: C9027-10
S: Pyrenees: C7600: (117)
S: Pyrenees: C7600: (20)	GAA.....
S: Pyrenees: C7600: (21)
F: Entressen: C7690: (53)
F: Entressen: C7690: (54)
F: Entressen: C7690: (207)	GAAACTAACA AGG.
F: Entressen: C7690: (78)
F: Mouriès: C9046: (202)	GAAAC.....
F: Mouriès: C9046: (192)
F: Mouriès: C9046: (40)
F: Mouriès: C9046: (163)
F: Crau: C18951: (150)	GAAACTAACA AGGA
F: Crau: C18951: (4A)
F: Crau: C18951: (151)
F: Crau: C18951: (121)
F: Crau: C18951: (144)	GAAACTA....
F: Crau: C18951: (146)	G.....
F: Crau: C18951: (147)	GAAAC.....
F: Crau: C18951: (148)
F: Crau: C18951: (149)	GA.....
E: Kent: C6374: (82)	GAAACTA....
E: Kent: C6374: (84)
E: Kent: C6374: (197)	AAGAAA....
E: Kent: C6374: (85)	GAAACTAAC.
E: Kent: C6374: (198)	GA.....
E: Kent: C6374: (81)
E: Kent: C6374: (87)
E: Kent: C6374: (86)
E: Kent: C6374: (85)	GAAACTAAC.
E: Kent: C6304-3	G.....
E: Kent: C6304-4

E: Kent: C6304-5
E: Kent: C6304-6
E: Kent: C6304-9	GAAACTAAC.
E: Kent: C6304-10	GAAACTAACA AGGA
I: C5897-4	GAAACT.....
I: C5897-6
I: C5897-8
I: C5891-9
I: C5897-10
I: C5897: (208)	GAAACTAACA AGG.
I: C5897: (104)	ACTAACGTAA ACCG
I: C5987: (105)	CACAA.....
I: C5897: (19)
I: C5897: (102)
I: C5897-2n	GAAACTAACA AGGA
I: C5897: (103)
I: C5897: (18)
M: C1761: (170)
M: C1761: (167)	GAA.....
M: C1761: (168)	GAAACTAAC.
M: C1761: (169)	GAAACTAAC.
A: C6452-3
A: 6452-10	CAAACKAACA AGGA
TA: C960: 300
TA: C960: 301
TA: C960: 302
TA: C960: 303
TA: C960: 304	G.....
TA: C960: 305
TA: C960: 306
TA: C960: 307
TA: C960: 369	GAAACTAACA AGGA
B: C24247: 314
B: C24247: 315
B: C24247: 316
B: C24247: 317
B: C24247: 311
B: C24247: 310
B: C24247: 312
B: C24247: 313
B: C24247: 349	GAAACTAACA AGGA
B: C24247: 373	AAACTAACAA GGAT
B: C24247: 374	GAAACTAACA AGGA
S.184: C21105: 308
S.184: C21105: 309
TR: 324 C4143

TR: 325 C4143
TR: 326 C4143
TR: 328 C4143
TR: 329 C4143
TR: 330 C4143
TR: 331 C4143
TR: 332 C4143
TR: 333 C4143
TR: 370 C4143
334 C2095 Morso
335 C2095 Morso
336 C2095 Morso	GAAAC.....
337 C2095 Morso	AAACTAAC..
338 C2095 Morso	GAAA.....
338 C2095 Morso
339 C2095 Morso
340 C2095 Morso
341 C2095 Morso
342 C2095 Morso
343 C2095 Morso
378 C2095 Morso	GAAACTAACA AGGA
379 C2095 Morso	AAACTAACAA GGAT
345 C5892 Hafia
321 C5892 Hafia
347 C5892 Hafia
348 C5892 Hafia
381 C5892 Hafia	AAACTAACAA GGAT
350 C6360 Blanca
352 C6360 Blanca	GAAACTAACA AGGA
353 C6360 Blanca	AAACTAACAA GGAT
354 C6360 Blanca	GAAACTAACA AGGA
355 C6360 Blanca	GAAACTAACA AGGA
356 C6360 Blanca	GAAACTAACA AGGA
359 C6360 Blanca	GAAACTAACA AGGA
360 C7555 Ovcak	GAAACTAACA AGGA
361 C7555 Ovcak	GAAACTAACA AGGA
362 C7555 Ovcak	GAAACTAACA AGGA
363 C7555 Ovcak	GAAACTAACA AGGA
364 C7555 Ovcak	GAAACTAACA AGGA
365 C7555 Ovcak
Aran1	GAAACTAACA AGGA
Aran3	GAAACTAACA AGGA
Aran4	GAAACTAACA AGGA
Aran5	GAAACTAAC.
Aran6
Aran7	GAAA.....

Aran8
Aran9
Aran10
Crusader1
Crusader2
Crusader3	GAAAC.....
Crusader4
Crusader5
Crusader6	G.....
Crusader8
Crusader9
Cruasder10	GAAACTAACA AGGA
R: C15111: 383
R: C15111: 384	GAAACT....
R: C15111: 385	GAAACTAACA AGGA
R: C15111: 387	GAAACTAAC.
R: C15111: 388	GA.....
R: C15111: 389	GA.....
R: C15111: 390	GAAACTAACA
R: C15111: 391	GAAACTAAC.
A: C6452: 322	GAAACTAACA AG..
A: C6452: 3232
A: C6452-5
A: C6452-6
A: C6452-8
A: C6452-9
A: C17561 No.1
A: C17561 No.2
A: C17561 No.3	GAA.....
A: C17561 No.4
A: C17561 No.5
A: C17561 No.6
A: C17561 No.7
A: C17561 No.8
A: C17561 10
D: C16511: 372
D: C16511: 319
D: C16511: 320	GAAAC.....
D: C16511: 318	G.....
D: C16511: 319	G.....
D: C16511: 371	GAAACTAACA AGGA