Variation among native and alien populations of hoary mustard, Hirschfeldia incana (L.) Lagreze-Fossat, and the application of DNA melting analysis to investigate microsatellite (SSR) variation.

Smith, Melvin N. E

Melvin N. E. Smith B.S.c. (HONS.)

Variation Among Native and Alien Populations of Hoary Mustard, *Hirschfeldia incana* (L.) Lagreze-Fossat, and the Application of DNA Melting Analysis to Investigate Microsatellite (SSR) Variation.

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy.

Swansea University
2010

Acknowledgements

I would like to thank Dr PLM Lee and Dr CR Hipkin for their advice and guidance throughout my PhD.

I would like to thank the BSBI for the research grant that allowed me to undertake fieldwork in Crete.

I would like to thank Dr R John for assistance with fieldwork in Crete

I would like to thank Dr S Forbes-Robertson for the use of the Roche LightCycler, and Dr T Butt for the use of the Corbett Rotorgene.

Summary:

H. incana is a native species of the Mediterranean and Middle East. As a neophyte (alien) it has undergone a large range expansion in Northern Europe, the Americas, Asia and Australasia. Casual field observations suggested that within its native range, the dominant life strategy of *H.incana* was annual, whereas in the British flora it was predominantly perennial.

Populations from native and alien ranges were studied in the field and in common garden experiments. Phenotypic differences in morphological and physiological characteristics were compared. Plants derived from neophyte British populations made larger leaf rosettes, flowered later (> 140 days) and exhibited a perennial life cycle. Plants from native, North African and Southern European populations (excepting those from montane Spain) made smaller rosettes, flowered early (< 110 days) and died after flowering once. Neophyte populations from California were similar to native populations. Some native populations (e.g. Cypress) did not survive a British winter. Unlike native populations, initiation of flowering in neophyte British populations was stimulated by a period of vernalisation. These results suggest that life strategy changes have occurred in neophyte populations of *H. incana* as this species expanded its range northwards, and implies possible genetic differences.

Ten microsatellite primers, previously described for related Brassicaceae species, were therefore investigated for potential use in the assessment of *H. incana* population genetic structure. Five primers successfully amplified a product of expected size, of which 3 were subscequently sequenced to confirm the presence of the SSR.

The application of real-time PCR DNA melting analysis to identify SSR variation was investigated using Roche SYBR green and Corbett HRM platforms. SSR variation could be detected using DNA melt analysis, but due to difficulty identifying the composition of heterozygous SSR's the technique could not be sufficiently refined to investigate population diversity. However, preliminary results indicated possible SSR variation between isolated populations.

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ..... (candidate)

Date ....05/07/10.....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ..... (candidate)

Date ....05/07/10.....

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .. .. (candidate)

Date ....05/07/10.....

NB: *Candidates on whose behalf a bar on access has been approved by the University (see Note 7), should use the following version of Statement 2:*

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access approved by the Swansea University.**

Signed ..... (candidate)

Date ....05/07/10.....

Contents

List of Figures

List of Tables

List of Abbreviations

AFLP; amplified fragment length polymorphism.

BBSRC; Biotechnology and Biological Sciences Research Council.

bp; base pairs.

cm; centimetre.

CO_2; carbon dioxide.

Conc; concentrated.

C-S-R model; competitors, stress tolerators and ruderal species.

dHPLC; denaturing high pressure liquid chromatography.

dH_2O; distilled water.

DNA; deoxyribonucleic acid.

Fo; fluorescence under dark conditions.

Fv; variable fluorescence.

Fv/Fm; the maximum quantum efficiency of photosystem II.

f.wt; fresh weight.

g; gram.

Hr; hour.

HRM; high resolution melt.

JNCC; Joint Nature Conservation Council.

L; litre.

Lat; latitude.

Long; longitude.

m; $1x10^{-3}$.

M; metre.

ml; millilitre.

Mol; moles.

mm; millimetre.

N; nitrogen.

nad4; nicotinamide adenine dinucleotide.

NaOH; sodium hydroxide.

nm; nanometre.

NR; nitrate reductase.

NVC; National Vegetation Classification.

P; phosphorus.

PAGE; polyacrylamide gel.

PCR; polymerase chain reaction.

Pers.Comm; personal communication.

Pers. obs.; personal observation.

ppm; parts per million.

PI; performance indicator.

PPFD; photosynthetic photon fluence density.

RAPD; random amplification of polymorphic DNA.

S; second.

SD; standard deviation.

SNP; single nucleotide polymorphism.

SSCP; single-strand DNA conformational polymorphism.

SSR; simple sequence repeat.

UV; ultraviolet.

UK; United Kingdom.

V; volt.

$; United States Dollar.

£; Pound Sterling.

"; inches.

°C; degrees centigrade.

%; percent.

μ; $1x10^{-6}$.

Chapter 1

General introduction

1.1 The concept of invasive species

Species invasion poses a major threat to biodiversity and is rated second only to habitat destruction as one of the main threats to global biodiversity, (Wilcove *et al.* 1998). In addition to ecological impacts such as native species loss, invasive species can have direct negative effects on humans including acting as vectors of disease and adding pressure on agricultural and other managed habitats (Colautti & MacIsaac 2004). Invasive species cost the USA an estimated $137 billion annually, e.g. losses to agriculture, species control costs (Pimentel *et al.* 2000). In order to minimise impacts by invasive species a thorough knowledge of their biology is required to provide sustainable management plans (McNeely *et al.* 2001).

There has been a global increase in the human transportation of species, both intentionally and accidentally. Historically, human travel has influenced species distribution by providing new routes to previously unoccupied areas. Many crop, food and ornamental species have been deliberately introduced into new regions, whereas vast numbers of species have been unintentionally introduced as a result of trade and travel. Definitions and terminology used in describing and discussing invasion biology is highly variable and sometimes contentious. For example, the following are terminology used in Sax *et al.* 2005:

- 'Species invasion' refers to any species occupying a region in which it was not previously present, either historically or prior to some reference point.

- 'Non-native', 'exotic', 'alien' and 'introduced': all interchangeable and used to refer to species that are not indigenous (native) to the region of interest. These do not necessarily mean species that are established.

- 'Naturalised' are non-native species that have become established (i.e. self-sustaining populations).

• 'Invasive' species are naturalised species that have spread widely within the newly occupied region. Usually these also cause ecological or economic damage.

A native species may be considered as a species that is indigenous to an area. Natural colonisation occurs when such species disperse to new regions (i.e. areas previously unoccupied by that species) without anthropogenic intervention. Conversely, 'non-native' generally refers to a species whose presence in an area is the result of anthropogenic activities. This can be a contentious subject because it is often difficult to be certain about historic species movements. A distinction is often made based on the duration of a species' presence in the area (Manchester & Bullock 2000). In addition to this there is possible speculation that other human influences such as climate change may have assisted what would appear to be natural colonisation. Colautti's 2004 review of invasive species terminology indicated possible confusion with the wide use of terms such as 'exotic', 'introduced', 'invasive', 'alien' and 'naturalised' among others. Where appropriate, terminology will be kept constrained in this thesis and will be expanded upon where required to minimise confusion. Both 'exotic' and 'alien' will refer to non-native species, 'introduced' will mean the human transportation of plant or propagule across a geographical border (Richardson *et al.* 2000). Naturalised (non-native) plant species are viewed as self sustaining by reproduction without further input by humans. These are in contrast to 'casuals' that do not persist without repeated introductions.

The term 'invasive' when referred to in ecology means to spread into areas away from sites of introduction. Richardson *et al.* (2000) produced a schematic representation of barriers that introduced species must overcome in the processes that lead to invasiveness. These consist initially of geographical barriers that control introduction, local environmental and reproductive barriers that may limit naturalisation, and dispersal and environmental barriers for both disturbed and natural habitats that determine if a species will become invasive. It is important to note that native species can have invasive properties. This is most commonly observed in modified or managed habitats where competition is reduced or nutrients are abundant. *Urtica dioica* (stinging nettle), *Pteridium aquilinum* (bracken), *Hedera helix* (common ivy) and *Rubus spp.* (bramble) are all native to the UK, but are among the most

invasive species in the British flora. Alpert *et al.* (2000) highlighted the importance of invasibility of a habitat that can promote invasion by either natives or non-natives. In order to better understand the dynamics involved in invasive ecology, knowledge can be transferred from natural dispersal models and combined with information on habitat structure.

1.2 Barriers to migration and transportation of non-native invasive species

Millions of years of bio-geographical isolation have lead to high levels of biodiversity, both at the species and population level. Natural barriers to plant species migration or gene flow must prevent the transportation of either pollen or propagules. Obvious barriers include the oceans, mountain ranges, deserts and areas of extremes of climate. However, over geographical time scales these barriers are constantly in a state of flux and can split or join populations and cause changes in species associations.

In contrast, human mediated species transportation, in evolutionary timescales, is a relatively new occurrence originating c. 6000 BC. Human migration and trade with a massive expansion in human movement between continents is particularly apparent from approximately AD 1500 (Manchester & Bullock 2000). As well as increasing in frequency in more recent times, species introductions have taken a greater variety of transportation routes. High volume trafficking of goods has resulted in increased numbers of biological stowaways across an ever increasing diversity of locations. A well-known issue is the exchange of ballast water from loading and unloading ships that is causing ecological problems due to the transportation and release of non-native planktonic animals and algae that then become invasive and ecologically damaging in new locations (Cariton & Geller 1993). Invasive species have long been intentionally introduced by humans without consideration for ecological consequences, for agricultural use (crops, timber, aquaculture and fur) or as ornamentals or botanical collections.

Among non-native invasive organisms, plants are an important group. Although sessile, plants are often easily dispersed as seeds and are therefore often transported by humans across natural barriers either intentionally or unintentionally. For example, mixed seed

often contains a diverse range of species and when produced cheaply may include undesirable invasive species. Aggressively invasive species of plants can have devastating effects on the environment, in both economic and conservation terms.

1.3 Characteristic traits of non-native invasive species

It has been a goal of ecologists to form a framework with which potentially invasive species could be rapidly identified at early stages to allow management strategies to be developed to ensure speedy eradication or restricted introduction. This would minimise ecological damage in the most cost effective manner. One approach has been to attempt to identify traits that are common features to invasive species, or that are likely to promote invasiveness (Goodwin *et al.* 1999, Radford & Cousens 2000, Lloret *et al.* 2005). This provides a good starting point, but the influence of the environment and also individual floras should be considered. It is not possible to review here all possible characteristics in great detail. The following are briefly just some of those commonly believed to promote invasiveness in non-native species.

1.3.1 ***Competitiveness.*** The newly introduced species needs to be able to out-compete species native to the habitat (Levine & D'Antonio 2003). This demonstrates a better competitive ability of the invader for an individual niche. Competitive traits provide an advantage over the other species in the same habitat. Whilst not always easy to predict the outcome of species traits in a new habitat (Rejmanek & Richardson 1996) it is an aspect that must be considered for assessing risk of invasiveness. A striking example of direct competitive interaction among plants is 'allelopathy': root exudates that have negative biochemical effects (phytotoxic) on other plants. One hypothesis explaining how certain introduced plants can be competitively dominant in a new community, but relatively benign at 'home', is that they have a more strongly allelopathic effect on species they have not co-evolved with than they do with species in their 'home' communities. Evidence for this 'novel weapons hypothesis' has been found for some plants invasive in non-native settings, including the North American invaders, knapweed (*Centaurea*) (Ridenour *et al.* 2008) and garlic mustard (*Alliaria petiolata*) (Callaway *et al.* 2008). An alternative, but equally compelling, hypothesis is 'enemy-release': where the introduced species is free of its

normal predators, diseases and other normal 'enemies' in the new habitat and as a consequence, becomes invasive (Keane & Crawley 2002)

1.3.2 ***High turnover.*** High fecundity will promote rapid spread of a species. Higher seed production, higher rate or frequency of seed production and greater viability of seed will increase the probability of a species success (high propagule pressure; Kolar & Lodge 2001). A rapid rate of germination can reduce generation times, allowing the rapid turnover of new individuals. For example, demographic models based on data from common garden experiments with Commelinaceae species demonstrated that invasive species outperformed non-invasive congeners, especially under high-nutrient conditions (Burns 2008). The study found that greater fecundity, shorter time to first reproduction, and greater vegetative reproduction all contributed to the superior performance of the invasive species. In contrast, specialist strategies, such as specific environmental requirements to break dormancy with cold treatment or fire regimes, are unlikely to be advantageous unless the species is colonizing a new region with very similar climatic stimuli as found in its native region.

1.3.3 ***Dormancy.*** Long term viability in a dormant state (e.g., in the seed bed) will aid survival of a population during periods of adverse conditions, which the organisms may not be able to tolerate. Domancy enables the building up of a seed 'bank'. Cues for dormancy termination and germination requirements are ways by which seeds in the bank sense 'gaps' of favourable conditions for growth, which significantly contribute to the naturalization and invasion processes by assuring a successful seedling establishment in environments of high competition (Martinez-Ghersa & Ghersa. 2006). Dormancy may also enhance long distance transport by humans, and may even evolve as a result of introduction. For example, *Cardamine hirsuta* is a European annual weed that was only recently introduced and naturalized in Japan. It was found that the Japanese plants generally had stronger seed dormancy and greater suppressed germination in higher temperature regimes than the original European strains (Kudoh *et al.* 2007).

1.3.4 ***Dispersal.*** Seed dispersal for an adventive species is obviously important for its spread. One definition of species invasion specifies that a species must produce reproductive offspring in areas distant from the site of introduction, with an approximate scale of over 100 m within 50 years (Richardson *et al.* 2000). It is likely that seeds spread by wind and by generalist animal vectors will become more invasive than those plants that require specialist vectors that may not be present outside their native range. For example, a frequent observation is that exotic and invasive species are common along roadsides (Gelbard & Belnap 2003). It was shown that long-distance seed dispersal via transport by vehicles along roads was significantly more frequent for non-native than native species (Von der Lippe & Kowarik 2007). A recent study of seeds that are released by vehicles at motorway tunnels further showed that the magnitude of seed deposition and species richness were higher in lanes leading out of the city, indicating that traffic may indeed foster invasion processes starting from cities and spreading to rural habitats (Von der Lippe & Kowarik. 2008).

1.3.5 ***Asexual reproduction.*** The ability to self fertilise or propagate vegetatively greatly increases the chance of survival for an individual or a small founder population. For example the highly invasive *Alliaria petiolata* (garlic mustard) is capable of both cross and self-fertilisation, although pollination and stigma maturity occur before the flower opens; indicating autogamy as a key breeding system (Anderson *et al.* 1996). This potentially allows the propagation of an invasive population from the introduction of a single individual. This is a similar strategy in many island floras. The ability to self pollinate and self fertilise increases the chance of species success from long range dispersal (Baker 1955, 1967). A higher proportion of self-compatible species are found on islands than in comparable mainland floras (Schueller 2004). In the UK the highly invasive *Fallopia japonica* is a single clone, reproducing predominantly through vegetative growth or obligate hybridization with other *Fallopia* species (Grimsby *et al.* 2007). As one of the most aggressive alien species in the UK, this provides an excellent example of clonally reproducing invasive species. *Fallopia japonica* is also an invasive species of North America; however in the US there is greater evidence for sexual reproduction. The final determining factor in the plant life cycle for invasiveness is generation time, i.e. time taken

for a seed to grow to into a mature plant and reproduce either sexually or clonally. Species with rapid generation times are likely to be more invasive than those with relatively slow generation times. *Capsella bursa-pastoris*, an invasive species of North America has been observed to have up to three generations in a single year (Askoy *et al.* 1998).

1.3.6 ***Rate of growth.*** Plant growth rate is an important component of competitive ability, whether in open ground such as urban gap or ruderal habitats, or in structured established vegetation. For a species to become invasive, it would be expected that growth rates would be higher than for a non-invasive species in the same habitat. There are several aspects to growth rates relevant to invasive success, including rate of biomass accumulation, seedling relative growth rates and specific leaf areas. Rogers & Hartemink (2000) note the threat posed to native biodiversity by invasive species with aggressive biomass accumulation rates and dominance of the seed bed. Grotkopp & Rejmanek (2007) describes comparisons between invasive and non-invasive seedling relative growth rates and specific leaf areas. It was found that invasive (woody species) had significantly larger root biomass allocation than less invasive species. In addition fast seedling growth associated with opportunistic resource acquisition (high specific leaf areas) may be important for successful invasion in Mediterranean regions.

1.3.7 ***Climatic adaptions.*** Wide climatic tolerances are useful traits if a species is to invade a variety of habitats, or is to become widespread e.g. over a wide latitudinal or altitudinal range. If a species is not narrowly adapted, a tolerance to a range of environmental pressures may increase chances of survival. Some tolerance to frost, drought, fire, water-logging, salinity, pests or pathogens would increase the range of habitats and environmental conditions that a non-native species could survive and in which it could potentially prosper.

1.3.8 ***Reproductive timing.*** For plants, the timing of flowering is an important trait for a successful reproduction. Especially in the case of vector pollinated flowers, it is critical that flowers are open while pollinators are present. In the UK rapid flowering followed by a long flowering period is likely to be advantageous to maximise pollination potential.

Species flowering too early or late in the season are likely to risk production of fewer seeds and be negatively selected (O'Neil 1999). In a recent review, Barrett *et al.* (2007) suggests "evolutionary modifications to reproductive systems promote the colonizing ability of invading populations and that reproductive timing is an important target of selection during range expansion". They note that there have been several reports of clinal variation in life-history traits among invasive plants, with variation in timing of flowering one of the most important of these traits in determining range expansion (Barrett *et al.* 2007).

1.3.9 ***Genetic diversity.*** In changing environments, genetic diversity is an important component of adaptive potential. The genetic diversity of colonising alien species may have an influence on their ability to survive and propagate in a new environment (Lee 2002). High genetic diversity provides adaptive potential and may be used to give an indication of population fitness (Reed & Frankham 2003). This may be dependent on the breeding strategy of the species and other factors such as environmental pressures and competition, but generally high diversity is considered to be advantageous. An extreme example is the hypothesis that hybridisation could lead to evolution of greater invasiveness (Ellstrand & Schierenbeck 2000). According to this idea, multiple introductions of non-native species from different locations mean that there is opportunity for diverse genotypes to mix, which would not occur naturally. Selection on the hybrids produced from the admixture will result in the genotypes best adapted to the new environment surviving and forming invasive populations in the non-native setting. At one extreme would be hybridisation between different species, but hybridisation between individuals from previously separated and genetically distinct populations could also lead to the evolution of invasive forms. Ellstrand and Schierenbeck (2000) suggest that an example of the former is *Spartina anglica* in Britain. This is an allopolyploid hybrid of the native *Spartina maritima* and the introduced *Spartina alterniflora*. The hybrid species is a dominant plant in the marshes it invades, whereas the progenitor introduced species remains limited in distribution.

Plants that are successful invaders do not necessarily have all the characteristics above, but a combination of some of these characteristics may be expected. For example Perrins *et al.*

(1993) considered the characteristics that enabled *Impatiens glandulifera* (Himalayan Balsam) to be such a successful invader of the British Isles. Key traits were identified as frost tolerance, high seed output, rapid germination and growth rate, and the ability to form dense mono-specific stands.

1.4 Anthropogenic activities promoting invasive species success

Habitat modification is the leading cause of species extinctions and biodiversity loss (Wilcove *et al.* 1998, Tylianakis *et al.* 2007). Ecosystems are based on complex interactions between species that are usually in a fine balance. Habitat modification generally leaves an area more susceptible to invasion than a closed or undisturbed natural ecosystem. Species have evolved to survive in their particular native habitat or ecosystem. Typically the environmental processes that have shaped natural ecosystems occurred over long periods of time - e.g. climatic changes, sea level changes, movement of tectonic plates, and mountain building and erosion. Human mediated impacts on habitats are very rapid in comparison, e.g. deforestation, selective harvesting, road building, construction and mining. As well as the physical destruction causing species loss or displacement, habitat fragmentation and the formation of gaps, edges and corridors cause further instability to the system. Invasive species, often being opportunist species or generalists can thrive in these conditions, in contrast to more ecologically specialized species (Marvier *et al.* 2004). An example of species invasion in Mauritius as a result of habitat destruction is provided by Lorence & Sussman (1986). The majority of native forests have been felled and the relicts of native evergreen wet forest are located in small nature reserves. Whilst the majority of woody species remain natives (95%) seedling plots were dominated by exotics with up to 22% of the species and 97% of the individuals being non-natives.

1.5 Environmental impacts of non-native invasive species

The most serious impacts of non-native invasive species include species extinctions, large scale ecosystem change and spread of disease. The UK has been fortunate in many cases with regard to impact from invasive species. That is not to say that it is without problems, but the more serious examples have been observed in other parts of the world. As a relatively recent flora exists in the UK due to post-glacial re-colonization, most of the

species are also represented in continental Europe. In effect the UK has a relatively low number of endemic species (i.e. species unique to that area) especially when compared to isolated oceanic islands. Island floras are typically not as diverse as continental floras, but islands with a history of isolation often have a high proportion of endemism. Two examples include Hawaii and New Zealand. Long periods of isolation have resulted in these islands developing a highly unique flora and fauna. The biodiversity of both islands has been seriously impacted upon by human activities. This includes impacts resulting from the introduction, both intentionally and accidental, of non-native species in addition to habitat loss. A lack of predatory species in New Zealand resulted in the evolution of several flightless birds. Human introduction of non-native mammals such as rats and the domestic cat caused devastation to many of these birds The Stephen Islands wren *Traversia lyalli* provides one example of many endemic species extinctions on the island.

Continents are also prone to impact by invasive species. Species displacement is a serious problem in some habitats e.g. in cases where species are dependent on a certain vegetation type. European purple loosestrife (*Lythrum salicaria*) is an invasive species of US mainland wetlands and estimated to be spreading at a rate of 115, 000 hectares a year. It reduces the abundance of native plants in this habitat; this in turn has negative impacts on the animals (some endangered) that are associated with them (Pimentel *et al.* 2000). Pimentel *et al.* 2000 also highlight major ecological changes by invasive species that out-compete native species and dominate ecosystems on continents. The yellow star thistle (*Centaurea solstitalis*) has invaded, and now dominates, over 15 million hectares of Northern Californian grassland. It reduces grazing potential, crop yields and displaces native plants, thereby reducing native plant and animal diversity. European cheatgrass (*Bromus tectorum*) has had serious impacts on the shrub-steppe habitat of the Great Basin, Idaho, dramatically altering its natural fire regime. This has created a novel ecosystem, with more frequent fires preventing the re-growth of native shrubs, resulting in the displacement of animals that depended on the native habitat. European cheatgrass was identified as early as 1949 as an 'ecologic invader' of North America, where it had become the dominant species on over 4 million hectares (Stewart & Hull 1949).

Closed habitats are usually considered to be more resistant to invasion than modified habitats e.g. shrub land or woodland. However, woodlands are still prone to invasion from aggressive species. *Rhododendron ponticum* is a well known invader of acidic woodland and heaths. Ecological problems associated with this species have been identified particularly in hazel and oak woodlands of western Scotland, where rare bryophyte and lichen communities are threatened (Long & Williams 2007).

Some non-native species used in commercial forestry cause major problems as invasive species in natural and semi natural ecosystems (Richardson 1998). The most successful invasive species from forestry includes *Pinus*, *Acacia* and *Eucalyptus* species. *Pinus* species are particularly troublesome in the southern hemisphere, where they can form dense stands and exclude other vegetation. *Eucalyptus* species in North America contribute to the dangers of forest fires, their high flammability aiding spread of fire.

Aquatic invaders are also becoming a widespread problem. In the UK, problem species include: parrot's feather (*Myriophyllum aquaticum*), New Zealand pigmy weed (*Crassula helmsii*), floating pennywort (*Hydrocotyle rununculoides*) and water primrose (*Ludwigia grandiflora*). Common impacts of invasive aquatic plants are the formation of dense mats excluding light and oxygen from the water, resulting in unsuitable conditions for fish and invertebrates. In addition, choked waterways can cause problems for watercraft.

It is not an easy task to attribute species extinctions directly to introduction of invasive species without considering all the other factors involved. However, ecosystem changes are more easily observed. Certainly the rising trend of species translocation is likely to be associated with increasing threats to native biodiversity. A greater understanding of invasion biology is still required for early identification and eradication of introduced species that are potentially damaging to the environment. The UK has an obligation to address invasive non native species under the Convention on Biological Diversity. The JNCC (Joint Nature Conservation Council) are preparing a document as an indicator of number and impacts of invasive non-native species, due in 2008. The aim is also to

quantify the costs associated with controlling and documenting the range of non native species.

An example of the economic cost of invasive species can be demonstrated by *Rhododendron ponticum*. *R. ponticum* affects approximately 52,000 hectares of land in the UK, including 30,000 hectares in nature reserves. During 2001, 1275 hectares of *R. ponticum* was controlled at a cost of £670, 924 (Dehnen-Schmutz *et al.* 2003). If this was extended to the control of all reported *R. ponticum* at this cost per hectare, approx £27,360,000 would be required for this individual species.

1.6 Occurrence of non-native invasive species

Just because individuals of a species have managed to be transported to a new area where the species has not previously occurred, it does not necessarily follow that these first founders will survive or that stable and persistent populations will form. Indeed, many newly arrived non-natives probably do not establish at all. A well known observation is that there is a lag period before non-native populations increase in size and become self-sustaining, or invasive – i.e., a lag period before naturalisation. Before this, populations are small and local extinctions would be common. One hypothesis explaining this lag phase is the 'Baldwin effect' (Whitfield 2008). This suggests that the organism relies on plasticity in its initial colonization – the traits that enable survival in the new environment are difficult for the organism, and it relies on plasticity to survive and therefore the founder population remains small and unstable. However, with time, natural selection fixes these favorable traits in the non-native population, and once these are 'hard-wired', they become the dominant traits.

One way to establish whether non-native populations are relying on trait plasticity to survive in the new environment, or whether genetic changes have occurred and the population has adapted and gone beyond the lag phase, is to conduct 'common garden' experiments. Propagules from both the native and non-native populations are grown together under identical conditions in a 'common garden'. If both natives and non-natives are plastic in traits, then all individuals will develop similarly in the common garden.

However, if genetic changes have occurred because of natural selection, then there will be differences in the traits concerned, when comparing native and non-native phenotypes (e.g. Blossey & Notzold 1995, Stinchcombe *et al.* 2004).

If there are phenotypic differences between natives and non-natives in the common garden, the experiment can be further extended by conducting reciprocal transplant experiments, where natives and non-natives are grown under both native and non-native environmental conditions. However, reciprocal transplant experiments are beyond the scope of this study and remain for future work. Here, we first focus on the common garden approach, which will establish whether there is a genetic basis underlying observed phenotypic differences between natural populations in their native and non-native settings.

1.7 Genetic data as research tools

High diversity may be advantageous to newly introduced species. Various factors can affect the genetic diversity of founder populations, including founder population size, and origin of source population(s). Molecular methods can be used to study founder events (e.g. Mithen *et al.* 1995) and determine routes of introduction of alien species (Amsellem *et al.* 2000). An indication of genetic diversity and adaptive potential can also be provided by the use of certain molecular techniques (Sakai *et al.* 2001). These are therefore useful tools to increase understanding of the biology of invasions and can compliment traditional phenotypic studies.

1.8 The study species *Hirschfeldia incana* (L.) Lagreze-Fossat

With expanding distribution and range, new populations of alien or invasive species provide an exciting research opportunity to investigate phenotypic variation that may have arisen over short evolutionary timescales. With increased species movements and habitat modification, the number of naturalised alien species and their abundance continually increases. Considerable research attention has been given to invasive plant species with the most obvious negative impacts. These impacts include; ecological changes to native habitats and species that cause structural damage, e.g. *Rhododendron ponticum* (e.g. Milne & Abbott 2001; Erfmeier & Bruelheide 2004), *Fallopia japonica* (Japanese Knotweed) (e.g. Hollingsworth *et al.* 1998; Hollingsworth & Bailey 2000), *Heracleum mantegazzianum* (Giant Hogweed) (e.g. Walker *et al.* 2003; Krinke *et al.* 2005) & *Impatiens glandulifera* (Himalayan Balsam) (e.g. Beerling 1993; Pysek & Prach 1995; Kollmann & Banuelos 2004). In addition, *Arabidopsis spp.* (e.g. Alonso-Blanco *et al.* 1999; Alonso-Blanco & Koornneef 2000; Riihimaki *et al.* 2005; Rutter & Fenster 2007) and *Capsella bursa-pastoris* (e.g. Hurka *et al.* 2003; Hintz *et al.* 2006) have been well studied as model species – in part due to the high volume of genetic resources for these species. *Capsella bursa-pastoris* has been given considerable attention as a weedy or invasive species in the USA.

An autecological approach, investigating the ecology of individuals and populations of a species, by studying habitat, distribution and life cycle, allows the description of the ecological niche of that species. Clapham (1956), during his presidential address to the British Ecological Society, recommended studying species both in the field and under controlled conditions for a greater understanding of the ecology of species, and to understand why plants occur in some habitats and are absent in others. Comparisons made from populations from contrasting habitats may identify critical characteristics that determine the species suitability to certain environments. Alternatively, it is possible to describe plant strategies based on their response to external factors such as stress and disturbance (e.g. Grime *et al.* 1988). This has been summarised using the C-S-R model,

which describes the equilibrium between stress, disturbance and competition, and can be used to categorise species based on their responses as competitors, stress-tolerators and ruderal species. Knowledge of these ecological strategies may assist in understanding potential threats to native habitats or biodiversity, when considering alien or invasive species.

This thesis will focus on one species (*Hirschfeldia incana*) that has been identified as rapidly expanding its range in the UK (Preston *et al.* 2000). Although it has not specifically been identified as an invasive species that is causing ecological problems in the UK (at present its ecological impact in the UK has not been documented), it has been classified as an invasive species and a noxious weed in other countries where it is present as a non-native. English Nature has listed it as 'strongly increasing in number' (Hill *et al.* 2007). The relatively recent expansion in its range and its increased frequency of occurrence makes it a suitable study species that provides an opportunity to investigate the biology of a potentially invasive species at an early stage in its colonisation history. However, there is currently little comparative information describing the life form strategies of native versus neophyte populations of *H. incana*. To have a better understanding of the mechanisms involved in the increasing spread of this species, its ecology in the UK should be studied in more detail. Autecological studies can be complimented with comparative molecular studies to explore the adaptive genetic potential of *H. incana* and possibly be used to model the routes of its introduction and spread. In combining information of habitat suitability and genetic population structure it may be possible to predict future range expansion of *H. incana*. In further increasing our understanding of how plants colonise new habitats, *H. incana* may offer valuable insights for outcrossing species at the critical transitional stage between the formation of founding colonies and naturalisation. As a UK invader, the study of *H. incana* will add to our existing knowledge of introduced organisms in the British landscape. Finally, as a plant with its origins in the Mediterranean and an apparent northwards range expansion in recent times, *H. incana* may in future prove to be a useful case study for understanding the biological consequences of predicted global climate change.

1.9 Aims and Objectives

The overall objective of this thesis is to provide an improved understanding of the biology of *H. incana* with specific focus on characteristics that promote its success as an invasive species. In addition considerations will be made to assess the threat that a species such as this may have to native floras (particularly the UK).

The biology and ecology of *H. incana* is presented in Chapter two using the guidelines supplied by the Journal of Ecology for the Flora of the British Isles.

Field studies in South Wales were undertaken to describe the habitats and community types occupied by this species in its alien range. These include description of habitat types and floristic surveys from Swansea and descriptions of the life history and characteristic traits of the species in this area. Comparative field studies were also undertaken in Crete, (a location where the species is considered to be native). This information forms the body of chapter three of the thesis.

A detailed comparison is then made of phenotypic traits of *H. incana* from a range of populations from its native and alien range. This information was derived from experiments where plants were grown from seed (obtained from each of these regions) under common greenhouse conditions. The aim of this was to identify if phenotypic differences are apparent between populations, and specifically between the native and alien range. This is reported in Chapter four.

Chapters five and six focus on developing methods to investigate the population genetic structure of *H. incana*. Chapter five describes the application of microsatellite (SSR) analysis using primers developed for closely related species and assesses the suitability of this approach. Chapter six describes the application of a novel method for SSR analysis using the primers previously described in the previous chapter and comments on the suitability of this method compared to traditional methods of SSR analysis and possible future directions for this technology and also the feasibility of full scale population studies of *H. incana* using SSR type analysis.

1.10 References

Alpert P, Bone E, Holzapfel C (2000) Invasiveness, invasibility and the role of environmental stress in the spread of non-native plants. Perspectives in Plant Ecology, Evolution and Systematics, 3, 52-66.

Amsellem L, Noyer JL, Le Bourgeois T, Hossaert-McKey M (2000) Comparison of genetic diversity of the invasive weed Rubus alceifolius Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. Molecular Ecology, 9, 443-455.

Anderson RC, Dhillion SS, Kelley TM (1996) Aspects of the ecology of an invasive plant, garlic mustard (Alliaria petiolata), in central Illinois. Restoration Ecology, 4, 181-191.

Askoy A, Dixon JM, Hale WHG (1998) Capsella bursa-pastoris (L.) Medikus (Thlaspi bursa-pastoris L., Bursa bursa-pastoris (L.) Shull, Bursa pastoris (L.) Weber). Journal of Ecology, 86, 171-186.

Baker HG (1955) Self-compatibility and establishment after "long distance" dispersal. Evolution, 9, 347-348.

Baker HG (1967) Support for Baker's law – as a rule. Evolution, 21, 853-856.

Barrett SCH, Colautti RI, Eckert CG (2008) Plant reproductive systems and evolution during biological invasion. Molecular Ecology, 17,373-383.

Blossey B, Notzold R (1995) Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. Journal of Ecology, 83, 887-889.

Burns JH (2008) Demographic performance predicts invasiveness of species in the Commelinaceae under high-nutrient conditions. Ecological Applications, 18:335-346.

Cariton JT, Geller JB (1993) Ecological Roulette: The global transport of nonindigenous marine organisms. Science, 261, 78-82.

Callaway RM, Cipollini D, Barto K, Thelen GC, Hallett SG, Prati D, Stinson K, Klironomos J (2008) Novel weapons: Invasive plant suppresses fungal mutualists in America but not in its native Europe. Ecology 89, 1043-1055.

Colautti RI, MacIsaac HJ (2004) A neutral terminology to define 'invasive' species. Diversity and Distributions, 10, 135-141.

Dehnen-Schmutz K, Perrings C, Williamson M (2004) Controlling Rhododendron ponticum in the British Isles: an economic analysis. Journal of Environmental Management, 70, 323-332.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? Proceedings of the National Academy of Sciences of the United States of America, 97, 7043-7050.

Gelbard JL, Belnap J (2003) Roads as conduits for exotic plant invasions in a semiarid landscape. Conservation Biology, 17, 420-432.

Goodwin BJ, McAllister AJ, Fahrig L (1999) Predicting invasiveness of plant species based on biological information. Conservation Biology, 13, 422-426.

Grimsby JL, Tsirelson D, Gammon MA, Kesseli R (2007) Genetic diversity and clonal vs. sexual reproduction in Fallopia spp. (Polygonaceae). American Journal of Botany, 94, 957-964.

Grotkopp E, Rejmanek M (2007) High seedling relative growth rate and specific leaf area are traits of invasive species: Phylogenetically independent contrasts of woody angiospernis. American Journal of Botany, 94, 526-532.

Hill M, Baker R, Broad G, Chandler PJ, Copp GH, Ellis J, Jones D, Hoyland C, Laing I, Longshaw M, Moore N, Parrot D, Pearman D, Preston C, Smith RM, Waters R (2007) Audit of non-native species in England. English Nature reports, number 662.

Keane RM, Crawley MJ (2002) Exotic plant invasions and the enemy release hypothesis. Trends in Ecology and Evolution, 17, 164-170.

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution, 16, 199-204.

Kudoh H, Nakayama M, Lihova J, Marhold K (2007) Does invasion involve alternation of germination requirements? A comparative study between native and introduced strains of an annual Brassicaceae, Cardamine hirsuta. Ecological Research, 22, 869-875.

Levine JM, D'Antonio CM (2003) Forecasting biological invasions with increasing international trade. Conservation Biology, 17, 322-326.

Lloret F, Medail F, Brundu G, Camarda I, Moragues EVA, Rita J, Lambdon P, Hulme PE (2005) Species attributes and invasion success by alien plants on Mediterranean islands. Journal of Ecology, 93, 512-520.

Long D, Williams J (2007) Rhododendron ponticum: impact on lower plants and fungi communities on the west coast of Scotland. Working towards protecting internationally important bryophyte and lichen communities from Rhododendron ponticum invasion. Plant life.

Lorence D, Sussman RW (1986) Exotic species invasions into Mauritius wet forest remnants. Journal of Tropical Ecology, 2, 147-162.

Manchester SJ, Bullock JM (2000) The impacts of non-native species on UK biodiversity and the effectiveness of control. Journal of Applied Ecology, 37, 845-864.

Martinez-Ghersa, MA, and. Ghersa CM (2006) The relationship of propagule pressure to invasion potential in plants. Euphytica, 148, 87-96.

Marvier M, Kareiva P, Neubert MG (2004) Habitat destruction, fragmentation, and disturbance promote invasion by habitat generalists in a multispecies metapopulation. Risk Analysis, 24, 869-878.

McNeely JA, Mooney HA, Neville LE, Schei PJ, Waage JK (2001) Global strategy on invasive alien species. Gland, Switzerland: IUCN 50 p.

Mithen R, Raybould AF, Giamoustaris A (1995) Divergent selection for secondary metabolites between wild populations of Brassica oleracea and its implications for plant-herbivore interactions. Heredity, 75, 472-484.

O'Neil P (1999) Selection on flowering time: An adaptive fitness surface for nonexistent character combinations. Ecology, 80, 806-820.

Perrins J, Fitter A, Williamson M (1993) Population Biology and Rates of Invasion of 3 Introduced Impatiens Species in the British-Isles. Journal of Biogeography, 20, 33-44.

Pimentel D, Lach L, Zuniga R, Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. Bioscience, 50, 53-65.

Radford IJ, Cousens RD (2000) Invasiveness and comparative life-history traits of exotic and indigenous Senecio species in Australia. Oecologia, 125, 531-542.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conservation Biology, 17, 230-237.

Rejmanek M, Richardson DM (1996) What attributes make some plant species more invasive? Ecology 77, 1655-1661.

Richardson DM (1998) Forestry trees as invasive aliens. Conservation Biology, 12, 18-26.

Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta F, West C (2000) Naturalisation and invasion of alien plants: concepts and definitions. Diversity and Distributions, 6, 93-107.

Ridenour WM, Vivanco JM, Feng YL, Horiuchi J, Callaway RM (2008) No evidence for trade-offs: Centaurea plants from America are better competitors and defenders. Ecological Monographs, 78, 369-386.

Rogers HM, Hartemink AE (2000) Soil seed bank and growth rates of an invasive species, Piper aduncum, in the lowlands of Papua New Guinea. Journal of Tropical Ecology, 16, 243-251.

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Syndallas B, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. Annual Review of Ecology and Systematics, 32, 305-332.

Sax, D.F., Stachowicz, J.J., and Gaines, S.D., editors. (2005) Species Invasions: Insights into Ecology, Evolution and Biogeography. Sinauer Associates, Sunderland, MA.

Schueller SK (2004) Self-pollination in island and mainland populations of the introduced hummingbird-pollinated plant, Nicotiana glauca (Solanaceae). American Journal of Botany, 91, 672-681.

Stewart G, Hull AC (1949) Cheatgrass (Bromus tectorum L.) - an ecologic intruder in Southern Idaho. Ecology, 30, 58-74.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J (2004) A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene FRIGIDA. Proceedings of the National Academy of Sciences of the United States of America, 101, 4712-4717.

Tylianakis JM, Tscharntke T, Lewis OT (2007) Habitat modification alters the structure of tropical host-parasitoid food webs. Nature, 445, 202-205.

Von der Lippe M, Kowarik I (2007) Long-distance dispersal of plants by vehicles as a driver of plant invasions. Conservation Biology, 21, 986-996.

Von der Lippe, M., and I. Kowarik. (2008) Do cities export biodiversity? Traffic as dispersal vector across urban-rural gradients. Diversity and Distributions, 14, 18-25.

Whitfield J (2008) Postmodern evolution? Nature, 455, 281-284.

Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E (1998) Quantifying threats to imperiled species in the United States. Bioscience 48, 607-615.

Chapter 2

Hirschfeldia incana (L.) Lagrèze-Fossat; Biological Flora

This account reviews information on all aspects of the biology of *Hirschfeldia incana* that are relevant to understanding its ecological characteristics and behaviour. The main topics are presented within the standard framework of the *Biological Flora of the British Isles*: distribution, habitat, communities, responses to biotic factors, responses to environment, structure and physiology, phenology, reproductive characters, herbivores and disease, history, and conservation.

Hirschfeldia incana (Brassicaceae), Hoary Mustard is an annual or short-lived perennial herb. Stems are erect, simple or branched from the base, up to 180 cm high, pale green occasionally reddish on exposure to high sunlight conditions, with hairs on their lower sections. Stem leaves are petiolate, linear, lanceolate or simple, with entire, toothed or lobed margins. Basal leaves forming a rosette, to a diameter of 76 cm. Rosette leaves are petiolate, pinnate or pinnatifid with a large terminal lobe. Flowers are up to 15 mm in diameter with four petals; petals yellow, 5-12 mm, approximately twice the length of the sepals (2-5 mm), the petal's claw may make up half the petal's length. Fruits are 5-17 mm long, green/red when unripe, maturing to become a pale straw colour, adpressed to the flowering stem. Fruits contain up to 15 seeds, including one in the beak. Seed mass is between 0.19-0.53 mg.

Hirschfeldia incana is native to the Mediterranean and Middle East, but is naturalized in the UK. *H. incana* has been reported in the wild in the UK since the 18th century, but the species has undergone a huge range expansion since 1970 (Preston 2002).

Worldwide, *H. incana* is extremely variable in size, leaf form and life history, but it may be identified from other members of the *Brassicaceae* by its adpressed siliques that contain a seed in the beak.

2.1 Geographical and altitudinal distribution

H. incana is presumed to be native to the Middle East and Southern/ Mediterranean Europe (Table 2.1). It is an established alien of the UK, North Europe, USA (California Nevada & Oregon), South America and Australasia. It has undergone a rapid expansion in its UK range over the past 40 years. Between 1987-1999 it was recorded in 345 10 km squares (Figure 2.1) compared to 17 between 1970-1986 (Preston *et al.* 2002). In the UK, the range is concentrated south of Scotland, but is found as far north as Moray. It is also reported to occur in Norway at a latitude of approximately 63° (Figure 2.2).

Table 2.1 The native range of *H. incana* as considered by the USDA- United States Department of Agriculture 1997:

http://www.ars-grin.gov/cgi-bin/npgs/html/taxon.pl?316897

Africa
Macronesia: Portugal – Azores, Madeira Islands; Spain – Canary Islands
Northern Africa: Algeria; Libya; Morocco; Tunisia
Asia-Temperate
Western Asia: Cyprus; Iran; Iraq; Israel; Jordan; Lebanon; Syria; Turkey
Caucasus: Azerbaijan; Russian Federation - Dagestan
Europe
East Europe: Ukraine - Krym
Southeastern Europe: Albania; Greece (inc. Crete); Italy (inc.Sardinia, Sicily; Yugoslavia
Southwestern Europe: France (inc. Corsica) Portugal; Spain (inc. Baleares)

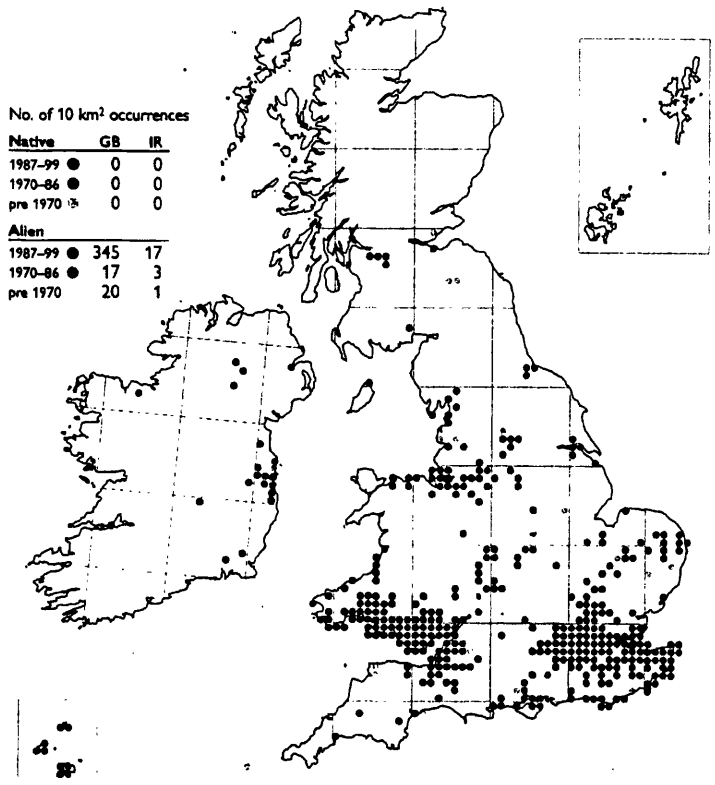


Figure 2.1 Distribution map of *H. incana* within the UK and Ireland. Reproduced from the Flora of the British Isles and Ireland (Preston *et al.* 2002).

H. incana occurs in temperate and Mediterranean zones (Figure 2.2). *H incana*'s native region of the Mediterranean and Middle East has a climate with distinct wet and dry seasons. The growing season of the plant in this region coincides with the availability of adequate water and typically falls between autumn and spring. It is therefore defined as a winter annual in these areas; i.e. a plant that germinates in the autumn, and produces flowers and seed the following spring, and completes its life cycle within one year.

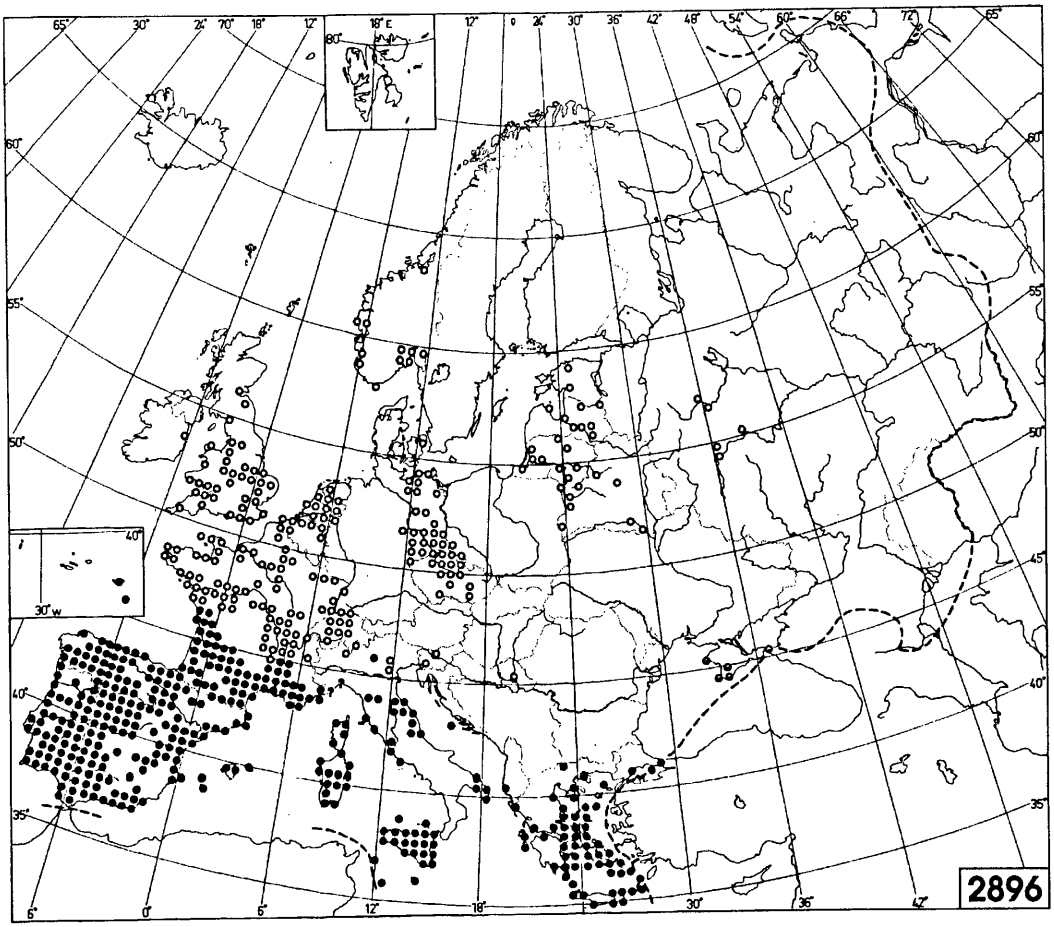


Figure 2.2 The European distribution of *Hirschfeldia incana* (data from 1990). The native range is indicated by the bold circles (•), the alien range is indicated by the unfilled circles (°). Reproduced from the Atlas Florae Europeae, distribution of vascular plants in Europe, 11. (Jalas & Suominen 1996).

2.2 Habitat

2.2.1 ***Climate and topographical limitations.*** *H. incana* has a predominantly southern distribution in Britain. However, this may reflect patterns of introduction and subsequent spread rather than climatic limitations. *H. incana* tolerates climates ranging from those found in the Mediterranean to Northern European temperate regions. The timing and length of the growing season varies across this range. Populations in the north of its range normally exhibit a longer growing season, from spring to autumn (in the UK *H. incana* flowers may be

observed for 8 months of the year, e.g. May-December) compared to those further south, where Mediterranean summer droughts terminate an earlier growing season that occurs between winter and spring (e.g. February-May).

2.2.2 *Substratum.* *H. incana* is typically associated with early successional ruderal or urban gap habitats where it is widespread on disturbed, well drained soils, both in its native and alien range. It can be found at high densities on areas of urban development. Typically, these soils have a high proportion of gravel, sand or construction rubble. In addition *H. incana* can be found in pioneer communities e.g. sand dune systems, either close to areas of disturbance, or on recently disturbed sands. Ruderal urban gap sites are usually nutrient rich, although less nutrients would be expected in sand dune soils. The levels of nutrients for sites containing *H. incana* in Swansea were within the range of: Amonnia 2.8-16.7, phosphate 2-21.5, TON 3.8-13.9, nitrite 0.077-0.126, nitrate 3.7-13.7 (μgg^{-1} soil) (Chingombe pers. comm). *H. incana* is reported to be tolerant to high levels of copper in the soil and may have remedial value for copper contaminated land (Poschenrieder *et al.* 2001).

2.3 Communities

In the UK *H. incana* is most frequently considered a weed of urban gap areas, agriculture, and an early colonizer of disturbed sites. *H. incana* can colonise and temporarily dominate recently disturbed sites. It is most commonly seen on road side verges, landfill sites, railways, docklands, other man made habitats, but it also occurs occasionally on sand dunes and shingle which are natural habitats. The large number of seeds generated by a single plant allows rapid colonisation of open areas.

Species lists included in the community tables of the National Vegetation Classification (NVC) do not include *H. incana* (Rodwell 2000), but it is closely associated with the communities described as ruderal and disturbed communities within 'Other Vegetation' (Volume 5). Many communities that contain *H. incana* are dynamic in their content, between seasons and from year to year. However, *H. incana* is invariably associated with disturbed habitats containing other neophyte alien species and even where it occurs in grassland communities it is found where some disturbance has occurred e.g. near fence posts or newly

planted trees. Habitats occupied by *H. incana* have been described by Patel (2004). These habitat descriptions were based on species surveys from 17 sites in the South Wales area. Habitats occupied by *H. incana* and the habitat indicator species were described. Sites containing *H. incana* were divided into three distinct habitat types by Patel (2004), open habitats (less than 100% ground cover), closed habitats (ground cover of 100% or greater) and sand dunes. Open habitats containing *H. incana* were characterised by the indicator species *Mycelis muralis*. Closed habitats containing *H. incana* were characterised by the indicator species *Euphorbia peplus*. In this closed habitat *H. incana* was present in gaps among the existing vegetation. Sand dune communities containing *H. incana* were characterised by the indicator species *Ammophila arenaria*.

All populations of *H. incana* that were studied in the Swansea area were found on areas of modified habitat. The only habitats in the UK containing *H. incana* viewed by the author, not obviously modified include a pioneer sand dune (Blackpill, Swansea Bay) and a cobble beach (Beer, Dorset). In both these cases the species was present at low density (pers. obs.) and it is possible that both these locations had been subject to recent disturbance by human activities. The populations studied in Swansea in this instance were all at low altitude <20m. Species lists compiled from 4 locations in the Swansea area are listed in table 7.1-7.5 within the appendix; in addition more species associations are available in Patel 2004.

Community species associations for *H. incana* have been included in European phytosociological studies (Kehl 1995). In Antalya – Turkey *H. incana* occurs within the boundaries between trampling, segetal (field weeds) and ruderal zones – *H. incana*, *Sisymbrium officinale* and *Crepis foetida* being the most frequent species. Road and field-side communities contain high cover values of *Heliotropium hirsutissimum* and the presence of *Amaranthus albus*, *Marrubium vulgare, H. incana, Chrozophora tinctoria* and *Tribulus terrestris. H. incana* is a characteristic species of the *Hordeum leporinum – Chenopodium murale* community including: *Echium plantagineum, Nonea pulla, H. incana, Sisymbrium officinale, Lolium rigidum, Tribulus terrestris, Chenopodium opulifolium, Marrubium vulgare, Antirrhinum orontium.* Habitats containing shallow soils between travertine outcrops (non-marine calcium carbonate deposits), in parts covered with travertine rocks

were found to contain *H. incana, Plumbago europaea, Hypericum triquetrifolium, Centaurea solstitialis, Arum italicum, Pterocephalus plumosus, Peganum harmala, Alyssum strigosum, Marrubium vulgare, Delphinium peregrinum, Verbena officinalis, Echium plantagineum, Verbascum sinuatum, Quercus coccifera, Rhamnus oleoides and Asparagus acutifolius* (Kehl 1995). In addition, *H. incana* has been described as fitting into the subnitrophilous community of *Rapostro rugosi-Hirschfeldietum incanae.* For more details see, Chronopoulos *et al.* 2005.

2.4 Response to biotic factors

H. incana exploits early successional habitats e.g. recently exposed or disturbed sites with significant amounts of bare ground, where population density can reach high levels. Seedlings become established rapidly in the absence of competition. Under these conditions *H. incana* may become the dominant species (pers. obs.). However, *H. incana* competes less well as successional changes progress and a greater number of species occupy the previously bare areas. In such situations, *H. incana* decreases in frequency the community may become dominated by course perennial grasses (Patel 2004). The occurrence of *H. incana* in late successional habitats is likely to be from individuals that have perennated and survived successive seasons. Seed germination or survival of seedlings would appear to be lower among dense grass (pers. obs.).

H. incana has medium palatability to goats and may be tolerant to grazing damage (Simmonds *et al.* 2000). In a study in California, *H. incana* was found to be more prevalent in areas of higher cattle grazing pressure (Orre *et al.* 2005). Seeds from *H. incana* have been observed to germinate after digestion by horses (Ghosh *et al.* 2008), indicating that the effects of grazing may aid in seed dispersal. Its leaves are predated by slugs, snails and insects (e.g. cabbage white larvae (pers obs.)).

2.5 Responses to the environment

2.5.1 ***Gregariousness.*** After initial colonization *H. incana* is frequently found to occur in very dense localised populations, often as a dominant species for discrete periods of time (both in its native and alien range). Densities can be as high as 500 seedlings m^{-2} (pers. obs.

in SA1 development site Swansea) and mature plants have been recorded at 40 m^{-2} (pers. obs., Crete) and up to 30 m^{-2} at a site in South Wales (Patel 2004). At these high densities *H. incana* often forms a continuous stand, with aerial vegetation becoming entangled with neighbouring plants. As the habitat begins to close over with late successional progression, *H. incana* is found less frequently and as time progresses its occurrence may be as solitary individuals in populations as observed in Swansea, UK (Patel 2004, pers. obs.). No information is available for successional changes in habitats containing *H. incana* in its native range.

2.5.2 ***Performance in various habitats.*** Individual plant size, morphometric characters and life history are highly variable across populations of *H. incana* (see Chapters 3 and 4). This includes characteristic leaf shape (Figure 2.3), time taken from germination to flowering, flowering season and ability to perennate. Overall plant size can be reduced by limited root space, water or nutrient availability. However, flowering and seed production may still occur, but with a lower output.

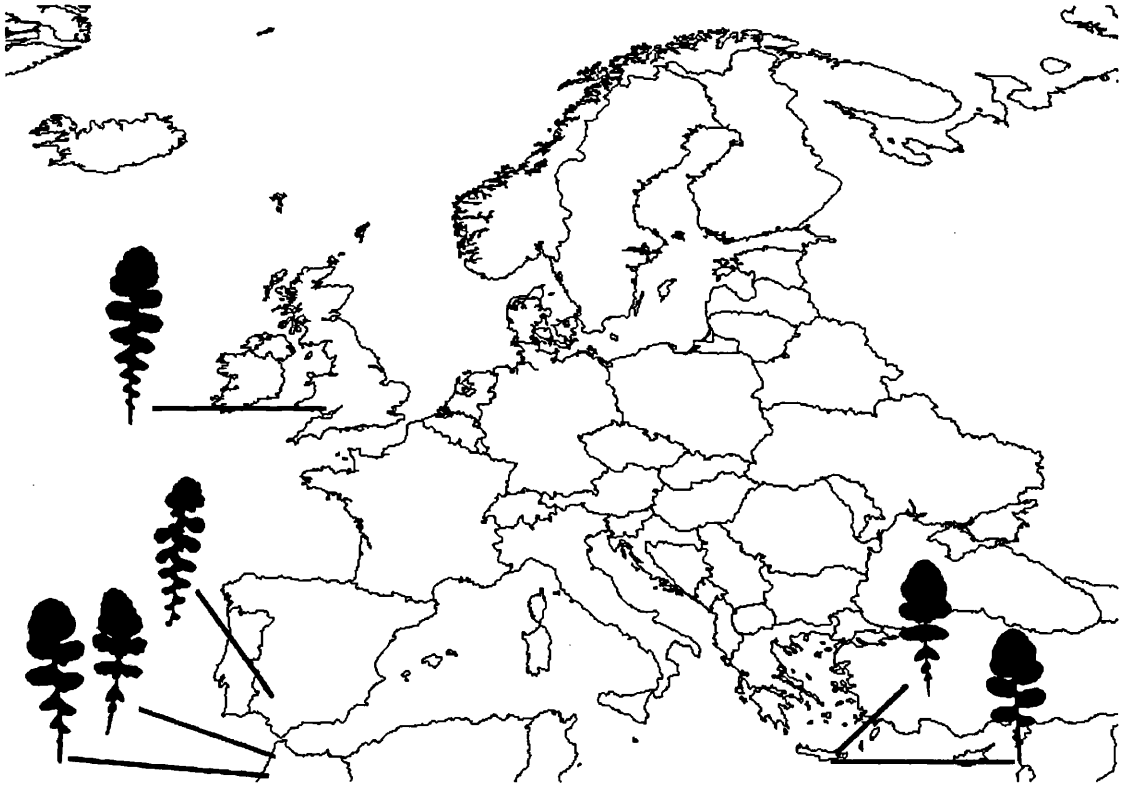


Figure 2.3 Variation in the leaf shape of *H. incana* originating from different populations across Europe. Outlines are scans of characteristic rosette leaves from plants grown under common garden conditions in a glasshouse at Swansea University. This range includes plants from: Swansea (UK), Bajar (Spain), Mohammedia (Morocco), Taza (Morocco), Elounda (Crete) and Irapetra (Crete).

A population study of *H. incana* from native and alien European populations (see Chapter 4 for addition information), indicated that time taken from germination to flowering was shorter in individuals from the south of their range (native range). In addition, perennation was more likely in individuals from more northerly latitudes (alien range). This resulted in a higher production of seed from two year old plants than from plants producing seed in their first year of growth (See Chapter 3).

2.6 Structure and physiology

2.6.1 ***Morphology.*** *H. incana* has been described as an annual or short-lived perennial herb, but this variation in life history is linked to population origin (pers obs. see chapter 4). It has an upright growth form with a stem up to 2 m high, with branched, simple hairs lower on the stem and is often glabrous and glaucous above. Basal leaves form rosettes of up to 760 mm in diameter (pers. obs.); leaves are petiolate, pinnate or pinnatifid with a large terminal lobe and up to 10 lateral lobes (number varies with developmental stage and source population pers. obs). Stem leaves are petiolate, linear, lanceolate or simple, margins may be entire, toothed or lobed. The inflorescence is crowded. Sepals are 2-5mm green-reddish, often hairy. Its petals are yellow 5-12 mm x 2-5 mm, obovate to thin elliptical. The petal's claw may be up to half of the petal's length. Fruit dimensions of 5-17 mm x c 1.5-2 mm are typical. Fruits are dehiscent, cylindrical and slightly tapering at both ends with a terminal beak. The fruits are divided into segments longitudinally, and are adpressed to the stem (hence its former name of *Brassica adpressa*). The fruits may contain up to 15 seeds per capsule, this includes one seed in the beak. Typical seed dimensions are 0.5-1.5mm, and seed weights between 0.19 and 0.53mg have been recorded (pers obs, see Chapter 4).

Variation in leaf morphology was observed in a population study of *H. incana*. A general pattern of fewer lateral lobes, in combination with a larger terminal lobe was observed from the majority of southern populations. In contrast, individuals from more northerly latitudes displayed a higher lateral lobe count and a proportionally smaller terminal lobe (Figure 2.3).

2.6.2 ***Mycorrhiza.*** *H. incana* has no known mycorrhizal associations (Harley & Harley 1987).

2.6.3 ***Perennation: Reproduction.*** There seem to be generalisations and inconsistencies with regard to observations and descriptions of perennation of *H. incana*. The literature lacks detailed information of the annual/perennial status across the majority of its range. From the results of common garden experiments (see chapter 4) and field observations (see chapter 3), it would appear that the capacity to perennate is dependent on source population, which in turn reflects the climatic pressures from these areas. For example, 100% of plants studied from Crete behave as strict annuals. This is in direct contrast to plants from Swansea (UK), which all displayed the ability to perennate. This may be a result of specialised strategies to deal with climatic pressures, e.g. perennation of *H. incana* would not be possible in Crete without specialist mechanisms to survive the dry season. However, within the UK sufficient precipitation is available year round so this is not an issue. In addition to survival as rosettes over-winter in the UK, *H. incana* may also produce aerial shoot rosettes that prevail through UK winter conditions (pers. obs.).

2.6.4 ***Chromosomes.*** *H. incana* is reported to be $2n = 14$ (Darmency and Fleury 2000). No data exists in the literature about ploidy levels of *H. incana* across its range.

2.6.5 ***Physiological Data***

Photosynthesis

Assimilation of CO_2 in *H. incana* is by the C_3 pathway (Canvin *et al.* 1980). Photosynthetic rates of *H. incana* at 20° C were found to peak at 25.8 $\mu molCO_2 m^{-2}s^{-1}$ at an irradiance of 1740 $\mu molm^{-2}s^{-1}$. The rate of dark respiration was -1.31 $\mu molCO_2 m^{-2}s^{-1}$ and a light compensation point of 12.9 $\mu molm^{-2}s^{-1}$ was observed. The initial slope of the light response curve, indicating photosynthetic efficiency or apparent maximum quantum yield, was 0.58 (Hipkin pers. comm. data not published). The rates of photosynthesis observed for *H. incana* compare with those obtained with sun plants. Typical rates of between 15-25 $\mu molm^{-2}s^{-1}$ are common for such plants (Radoglou *et al.* 1992, Bukhov *et al.* 1995). Rates in the upper part of this range might be expected for crop plants and fast growing weeds (McDowell 2002).

H. incana may be subject to photoinhibition stress under non-optimal growth conditions, which can result in the visible production of anthocyanin (e.g. high irradiance and low night-time temperatures, or high irradiance and drought conditions pers. obs). Greater effects of photoinhibition were observed in plants originating from Crete than from Swansea when investigated in a growth experiment in Swansea (Hipkin & Smith, data not published). In this case Fv/Fm ratios (indicating the maximum quantum efficiency of photosystem II) of 0.77 and 0.81 were observed for plants from Crete and Swansea respectively. These observations were made under conditions of high daytime irradiance (photosynthetic photon fluence density (PPFD) > 400 $\mu molm^{-2}s^{-1}$) following low night time temperatures (minimum temperature < 5°C) A typical range for non-photoinhibited plants would be > 0.80, but values of < 0.75 would indicate severe photoinhibition.

Inorganic nutrition

H. incana is found in habitats with contrasting levels of available nutrients e.g sand dunes and ruderal sites. Although *H. incana* tolerates nutrient poor soil it can be regarded as a nitrophile, displaying high gregariousness in nutrient rich soils commonly found in ruderal habitats (pers. obs.). This agrees with the high in vivo nitrate reductase activities observed in its tissues (1.13-1.56) (pers. obs., data not published). These NR values are comparable to those observed for *Cardamine hirsuta* (1.68-2.09) and *Coronopus didymus* (1.19-2.85) from the Brassicaceae (Al Gharbi & Hipkin 1984). These species, in common with *H. incana*, are frequently found on disturbed sites.

Biochemical Data

An investigation by Belkhiri & Lockwood (1994), indicated that no glucosinolates (secondary metabolites) could be found in suspension cells of *H. incana*. This is unexpected as they are found in almost all plants of the Brassicaceae, Capparidaceae and Caricaceae (Halkier & Gershenzon 2006).

Levels of ascorbic acid in *H. incana* have been investigated by Travers-Martin *et al.* 2008 and were found to be 0.69 μmol/g f.wt. this compares to values from *B. oleracea* ranging between 0.3-6.8 μmol/g f.wt. (Travers-Martin *et al.* 2008).

2.7 Phenology

Germination of *H. incana* in the UK occurs between spring and autumn, with the main period being spring. Seedlings can be seen into late summer/ early autumn. Plants begin to flower towards the beginning of May and flowering plants (with mild winters) may be observed through to December/ Early January (pers. obs.).

Floral initiation appears to differ between populations. Results from common garden experiments indicate that individuals from populations displaying an annual life history from Mediterranean regions commence flowering more rapidly than those from the UK (See Chapter 4).

2.8 Floral and seed characters

2.8.1 ***Floral biology.*** Darmency and Fleury (2000) investigated the mating system of *H. incana* in a study to determine the potential risk of hybridization with *B. napus*. They observed self-pollination showing two classes of response. One was the production of siliques containing very few viable seeds, with an average of 1 seed produced per 10 pollinated flowers. The other response was the production of siliques on 50% of the flowers, and an average of 1.6 seeds per flower. Random mating among the same plants produced siliques on 88% of pollinated flowers and 4.8 seeds per pollinated flower. Darmency and Fleury (2000) regarded that the second class of response indicated that it is likely that some *H. incana* plants are facultative inbreeders, as they are able to produce seeds from selfing (albeit at a low frequency). However, it is apparent that the main breeding strategy is one of outbreeding.

Table 2.2 Insects observed on *H. incana*. *H. incana* is visited by a variety of bees, hoverflies, butterflies and beetles that are attracted to its flowers (species list compiled by Charles Hipkin, data not published).

Pollinators	Herbivores	Species predating on insects
Apis mellifera	*Amblytylus nasutus*	*Adalia bipunctata*
Bombus lapidaries	*Brevicoryne brassica*	
Bombus lucorum	*Cernuella virgata*	
Bombus muscorum	*Meligethes aeneus*	
Bombus pascuorum	*Pieris brassicae*	
Bombus terrestris	*Pieris rapae*	
Eristalis arbustorum		
Eristalis pertinax		
Eristalis tenax		
Helophilus pendulus		
Meligethes aeneus		
Pieris brassicae		
Pieris napi		
Pieris rapae		
Scaeva pyrastri		
Syrphus ribesii		
Syrphus vitripennis		

In addition to breeding experiments, it is possible to estimate breeding strategy by the ratio of pollen to ovules produced by a flower (Cruden 1977). Pollen / ovule ratios of *H. incana* were within the range of 3.3-3.5 (Pers obs. data not published, details given in Appendix), these values lie between the boundaries of facultative and obligate outcrossing as defined by Cruden (1977). Cruden's scale to estimate breeding strategy can be approximated using boundaries of pollen/ ovule ratios as indicators to a species strategy: ratios around 1.8 indicate obligate inbreeding, ratios around 2.4 indicate facultative inbreeding, ratios around

3.2 indicate facultative outcrossing and ratios around 3.8 indicate obligate outcrossing. A study of pollen/ ovule ratios from 66 species of the *Brassicaceae* by Preston (1986) provided an estimate for allogamous and autogamous crucifers of 3500 and 1000 respectively (the range for *H. incana* was 2138-3022 Pers. Obs data not published). These pollen/ ovule ratios agree with the classification by Darmency & Fleury (2000) that *H. incana* is mainly an outbreeding species.

Populations of *H. incana* in Israel have been described as gynodioecious and to contain between 2-10 % male sterile plants (Horovitz & Galil 1972). Flowers of male sterile plants were observed to have smaller corollas and shorter stamens than those of hermaphrodite plants. Their sterile anthers maintain their original introrse orientation during anthesis, their endothecium lacking a dehiscence strip, and their tapetum persisting longer then that in fertile anthers.

2.8.2 ***Hybrids.*** Hybridization of *H. incana* with *B. napus* has been reported by Lefol *et al.* (1995), Lefol *et al.* (1996), and Darmency & Fleury (2000). The hybrid is reported to be triploid with 26 chromosomes (Lefol *et al.* 1996) and has low fertility. There are no reports of *H. incana* hybrids occurring in the wild.

2.8.3 ***Seed production and dispersal.*** Seed counts taken from plants at a site in Elounda (Crete) were used to calculate an estimated seed output of 80,000 seeds m^{-2}. This was for plants that were observed growing at a density of 40 m^{-2} (see chapter 3). At this site, the plant density was approximately uniform across the field of dimensions approximately 50 x 100m which, could be subject to an estimated output of 4 x 10^8 seeds per year. This can be compared to an estimated seed production in Swansea of 127,000 seeds m^{-2} where plants were observed at densities of 30 m^{-2} (please see Chapter 3. for details on field data).

H. incana has been described by Matsinos & Troumbis (2002) as having a long distance dispersal strategy (random distribution of its seeds throughout the landscape). Seed of *H. incana* is most likely dispersed by wind and aided by vehicular traffic vorticies. The dehiscent siliques release seeds readily once dry. The dried out stems of *H. incana* are prone

to break and may disperse seed with a tumble weed effect, given sufficient wind force. Imbert (2002) reported that the seed in the distal part of the silique (beak) does not disperse. Due to the frequency that *H. incana* is found as a roadside species, it is likely that dispersal is assisted by vehicles (see Von Der Lippe & Kowarik 2007 for information on vehicles as dispersal vectors). It is also possible that *H. incana* is also dispersed as a product of ground works associated with development projects where quantities of topsoil are transported between locations.

Interestingly, *H. incana* seed is available to purchase from B & T world seeds, at a cost of 1358 Euros per kilo; although its purpose is not described by B & T it can be assumed that the seed is sold for agricultural purposes. The leaves of the *H. incana* are palatable to grazing animals and humans, and could be planted as a crop is so desired. In addition, due to the tolerance of *H. incana* to high levels of copper in soil the species may have soil remediable values in agricultural land where copper accumulates e.g. viticulture. Kavak *et al.* (2007) report that *H. incana* is cultivated to increase honey production in bees.

2.8.4 ***Viability of seeds: Germination.*** Seed viability of *H. incana* is high and retains viability when stored under optimum conditions. Up to 95% germination success was observed from seed supplied from seed bank that had been collected 40 years prior to germination during this investigation (Pers. obs. see chapter 4 for details). Mean seed mass was also found to vary considerably between populations, between the range of 0.19 and 0.53 µg (Pers. obs. see chapter 4 for details).

The effect of temperature on germination response of *H. incana* (described as a winter annual in California) has been assessed in experiments conducted by Steinmaus et al (2000). They observed a mean value of 75% germinability for *H. incana* seed and described a T_{base} of between 6.5 and 13.5 °C (T_{base} is defined as the temperature below which phenological development ceases). Germination rate was also observed to increase up to 30 °C, after which it began to decrease. Further aspects of germination of *H. incana* have been discussed by Negbi *et al.* (1966). Seed germination of *H. incana* was found sensitive to substrate hydration, being progressively inhibited in increasingly high levels of hydration. Water-sensitivity was unaffected by light conditions, though germination was promoted by a short

irradiation and to a lesser extent by continuous irradiation and inhibited by darkness. In addition, it was observed that gibberellic acid hastened the processes of germination and reduced water-sensitivity.

2.9 Herbivory and disease; *Animal feeders or parasites*

Insects. The polyphagous agromyzid *Liriomyza strigata* and the drosophilid *Scaptomyza flava* have been recorded to mine *H. incana* (leaf and stem mines of British flies and other insects http://www.ukflymines.co.uk/Keys/HIRSCHFELDIA.html). Caterpillars of cabbage white butterflies (*Pieris spp.*) have been observed to cause serious leaf damage to plants growing in high densities (pers. obs.) Aphids may be found at high levels on *H. incana* especially in cultivation in greenhouse conditions, causing damage to leaves and flowering stems (pers. obs.). Butterfly species from Israel, whose larvae predate on *H. incana* include *Euchloe belemia and E. falloui* (Kark *et al.* 2004). White blister disease has been observed on *H. incana* in Turkey, caused by the pathogen *Albugo canida* (Kavak *et al.* 2007). In addition, see table 2.2 for herbivores observed on *H. incana* in South Wales.

Birds. No reported studies.

Mammals. *H. incana* has medium palatability to goats (Simmonds *et al.* 2000). Leaves and flowers are edible and may be consumed by humans.

2.10 History

It has been reported that *H. incana* was cultivated in the UK in the 18th century and was first recorded in the wild in 1837 (Preston *et al.* 2002). The reason for its cultivation in the UK was not discussed, but it is likely to have been used as a crop either for animal or human consumption. Colonisation histories have not been produced for *H. incana*, but with greater knowledge based on molecular markers it may be possible to develop a greater understanding of the dynamics of the spread of this species.

2.11 Conservation

H. incana is an invasive alien species in the UK and has no conservation value. The species is mainly found in areas of high disturbance where it can temporarily dominate the habitat. Its presence is usually linked to anthropogenic activities and in this case the habitat where *H. incana* is found does not normally contain species of conservation value that may be impacted upon by its presence. *H. incana* can be found as an occasional species of sand dunes, but in this habitat it occurs at low densities and is unlikely to cause problems to native species.

2.12 Concluding remarks

Hirschfeldia incana is a species that is undergoing a rapid expansion in its range. It is associated with areas of recent habitat disturbance and is commonly found on waste ground, docks, railways, landfill sites, sand dunes and areas that may be described as urban gaps. The species is considered native to East Asia and Mediterranean regions and can be found as an alien species in an increasing number of countries with temperate and Mediterranean climates.

Populations of *H. incana* tend to be ephemeral and shifting, dependent on areas of recent disturbance as they compete best as pioneer species. *H. incana* has the ability to produce and disperse large quantities of seed and may become a temporary dominant species of recently disturbed ground. *H. incana* shows variation between populations in terms of its life history e.g. annual and perennial behaviour and phenotypic characteristics.

Variation in life history of *H. incana* has been observed in plants between its native and alien range. Plants from its native range such as Crete display a strict annual life history. Plants from its alien range such as Swansea (UK) frequently do not flower during their first year of growth. These plants frequently overwinter and display the ability to perennate, this promotes the potential for greater seed production in plants of ages greater than one year. Due to the successful establishment of this species in the UK it is expected that its frequency will increase in other locations with similar habitats and climate.

2.13 References

Al Gharbi A, Hipkin CR (1984) Studies on nitrate reductase in British angiosperms I. A comparison of nitrate reductase activity in ruderal, woodland-edge and woody species. New Phytologist, 97, 629-639.

Belkhiri A, Lockwood GB (1994) Investigation of mustard oil glucosinolates in cell cultures of three species of Cruciferae. Flavour Fragrance J 9, 1-6.

Bukhov NG, Drozdova IS, Bondar VV (1995) Light response curves of photosynthesis in leaves of sun-type and shade-type plants grown in blue or red light. Journal of Photochemistry and Photobiology B: Biology, 30, 39-41.

Chronopoulos G, Theocharopoulos M, Christodoulakis D (2005) Phytosociological study of Hirschfeldia incana (L.) Lagraze-Fossat (Cruciferae) communities in mainland Greece. Acta Botanica Croatica, 64, 75-114.

Cruden RW (1977) Pollen-ovule ratios: a conservative indicator of breeding systems in flowering plants. Evolution, 31, 32-46.

Darmency H, Fleury A (2000) Mating system in *Hirschfeldia incana* and hybridization to oilseed rape. Weed Research, 40, 231-238.

Ghosh S, Kolipinski M, Quinn LD, Alas M (2008) Poster: Seed biology; Invasive plant seeds germinate after digestion by horses in California. Plant Biology Symposium 2008, Merida, Mexico.

Halkier BA ,Gershenzon J (2006) Biology and biochemistry of Glucosinolates. Annual Review of Plant Biology, 57, 303-333.

Harley JL, Harley EL (1987) A check-list of mycorrhiza in the British Flora. New Phytologist, 105 (Suppl.), 1-102.

Horovitz A, Galil J (1972) Gynodioecism in East Mediterranean *Hirschfeldia incana* Cruciferae. Botanical Gazette, 133, 127-131.

Imbert E (2002) Ecological consequences and ontogeny of seed heteromorphism. Perspectives in Plant Ecology Evolution and Systematics, 5, 13-36.

Jalas J, Suominen J (1996) Atlas Florae Europaeae, Distribution of vascular plants in Europe 11 Cruciferae, Ricotia to Raphanus. The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Kark S, Lens L, Van Dongen S, and Schmidt E (2004) Asymmetry patterns across the distribution range: does the species matter? Botanical Journal of the Linnean Society, 81, 313-324.

Kavak H, Katircioglu Z, Bukun B (2007) *Hirschfeldia incana*, a new host report for white blister caused by *Albugo canida* in Turkey. Australasian Plant Disease Notes, 2, 149.

Kehl H (1995) First published in: Sukopp H, Numata M, Huber A (1995) Urban ecology as the basis of urban planning. pp 85-150, SPB Academic Publishing, Amsterdam, the Netherlands.

Lefol E, Danielou V, Darmency H, Boucher F, Maillet J, Rennard M (1995) Gene dispersal from transgenic crops .1. Growth of interspecific hybrids between oilseed rape and the wild hoary mustard. Journal of Applied Ecology, 32, 803-808.

Lefol E, Fleury A, Darmency H (1996) Gene dispersal from transgenic crops .2. Hybridization between oilseed rape and the wild heavy mustard. Sexual Plant Reproduction, 9, 189-196.

Matsinos YG, Troumbis AY (2002) Modeling competition, dispersal and effects of disturbance in the dynamics of a grassland community using a cellular automaton model. Ecological Modelling, 149, 71-83.

Negbi M, Rushkin E, and Koller D (1966). Dynamic aspects of water-relations in germination of Hirschfeldia incana seeds. Plant and Cell Physiology, 7, 363-376.

Orre KJ, Hufft RA, Parker IM (2005) The effects of grazing on native and exotic seed banks at Elkhorn Slough National Estuarine Research Reserve. Elkhorn Slough Technical Report Series, 2005:2.

Patel R (2004) Ecology and population genetics of the neophyte alien, Hoary mustard (*Hirschfeldia incana*) (L.)Lagreze-Fossat. (Thesis M.Phil) University of Wales, Swansea.

Poschenrieder C, Bech J, Llugany M, Llugany M, Pace A, Fenes E, Barcelo J (2001) Copper in plant species in a copper gradient in Catalonia (North East Spain) and their potential for phytoremediation. Plant and Soil, 230, 247-256.

Preston RE (1986) Pollen-ovule ratios in the Cruciferae. American Journal of Botany, 73, 1732-1740.

Preston CD, Pearman DA, Dines TD (2002) New atlas of the British & Irish flora : an atlas of the vascular plants of Britain, Ireland, the Isle of Man and the Channel Islands. Botanical Society of the British Isles, Oxford University Press.

Radoglou KM, Aphalo P, Jarvis PG (1992) Responses of photosynthesis, stomatal conductamce and water use efficiency to elevated CO2 and nutrient supply in acclimated seedlings of Phaseolus vulgaris L. Annals of Botany, 70, 257-264.

Rodwell JS (2000) British Plant Communities vols. 1-5, Cambridge University Press, Cambridge.

Simmonds H, Holst P, Bourke C (2000) The palatability and potential toxicity of Australian weeds to Goats. Rural Industries Research and Development Corporation: Canberra. 166pp.

Steinmaus SJ, Prather TS, Holt JS (2000) Estimation of base temperatures for nine weed species. Journal of Experimental Botany, 51, 275-286.

Travers-Martin N, Kuhlmann F, Muller C (2008) Revised determination of free and complexed myrosinase activities in plant extracts. Plant Physiology and Biochemistry, 46, 506-516.

USDA: http://plants.usda.gov/java/profile?symbol=HIIN3

Von der Lippe M, Kowarik I (2007) Long-distance dispersal of plants by vehicles as a driver of plant invasions. Conservation Biology, 21, 986-996.

Chapter 3

Field studies of *Hirschfeldia incana* in Swansea and Crete

3.1 Introduction

Geographical differentiation of populations has been interpreted as the result of adaptive responses to environmental gradients and biotic interactions. The forces of natural selection would be expected to drive the evolution of genotypes that have maximum fitness in their local habitat. The Darwinian view is that responses to natural selection are slow processes. However, observations based on species differentiation as a result of recent introductions would suggest that some evolutionary processes may take place over short timescales (Phillips & Shine 2004, Kollmann & Banuelos 2004).

Studying successful recent colonisers may give insight into evolutionary processes that take place over relatively short periods of time. Invasive species, by definition, are species that colonise new areas at high rates (Richardson *et al.* 2000). Therefore, they may offer model conditions to study contemporary evolution. Invasive species are commonly associated with human induced habitat modification or destruction. Species that are able to take advantage of anthropogenic disturbance or changes to habitats have been identified to show a combination of directional selection with a short term opportunity for population growth (Reznick & Ghalambor 2001).

Observations of local adaptations are useful for testing hypotheses about specific traits that are more suited to particular environments (Kawecki and Ebert 2004). Evolutionary adaptations have been described as part of recent colonisations and have been associated with different environmental conditions or stresses (e.g. Berglund *et al.* 2004, Blair & Wolfe 2004, Maron *et al.* 2004, Muller-Sharer *et al.* 2004). This highlights the importance of field studies taken with populations in the wild in heterogeneous environmental conditions. Information gained in this manner, combined with common garden experiments and population genetic information may lead to a fuller understanding of the role of evolutionary processes in the success of invasive species success.

Little data exists on the performance of *H. incana* in the different countries and habitats where it occurs in the wild. *H. incana* has a native and alien range within Europe and has shown considerable range expansion in the UK since 1970 (Jalas & Suominen 1996, Preston *et al.* 2002). The species favors areas of recent disturbance for colonization and fits the criteria described above for investigation of contemporary evolution. In this study, populations were surveyed in order to investigate whether differences in growth habit were apparent between plants in their native compared with alien range. Although *H. incana* is now widespread and well established in a number of geographical areas in Europe, Asia, Australasia and America, financial restrictions limited the survey to two areas. Populations were studied in Swansea (UK) and in Crete (overseas fieldwork was funded by a small grant from the Botanical Society of the British Isles). Crete was selected as a suitable location to study native populations due to the availability of known population locations (R. John pers. comm.). To further understand the processes of local adaptation with regard to deme and habitat interaction, reciprocal transplant and common garden experiments may identify particular ecological and genetic factors that assist or restrict local adaptation (Kawecki and Ebert 2004). However, this methodology is beyond the scope of this study as caution must be exercised when translocating potentially invasive species. This pilot study is intended as a starting point to document field observations from populations of *H. incana* in one native and one alien setting. The preliminary work may act as a resource if future studies are to be planned to cover a greater proportion of its geographical range.

3.2 Method

3.2.1 ***Botanical surveys at sites in Swansea.*** *H. incana* is commonly found in South Wales in a variety of (usually ruderal) habitats. It can be found at high density in sites of recent disturbance, although localized populations are often short lived due to successional changes or habitat management. Plant species diversity surveys were carried out at five sites in the Swansea area where *H. incana* was present (Figure 3.1, Table 3.1). The locations were selected to cover as wide a range of habitat types as possible in the local area.

Table 3.1 Survey site information for locations visited in the Swansea area.

Site	Lat (N) & Long (W)	Altitude (m)	Substrate	Population
Pennard (1)	51°34'33" 04°05'19"	80	Fixed sand dune (high proportion of sand)	~100
Fabian Way (2)	51°37'12" 03°53'25"	15	High proportion of gravel and builders rubble	~200
Swansea SA1 (3)	51°37'37" 03°55'51"	10	Topsoil & gravel	~1000
Swansea Bay Marina (4)	51°36'52" 03°56'03"	5	Sand	~50
Blackpill (5)	51°35'55" 3°59'35"	5	Sand & Boulders	~50

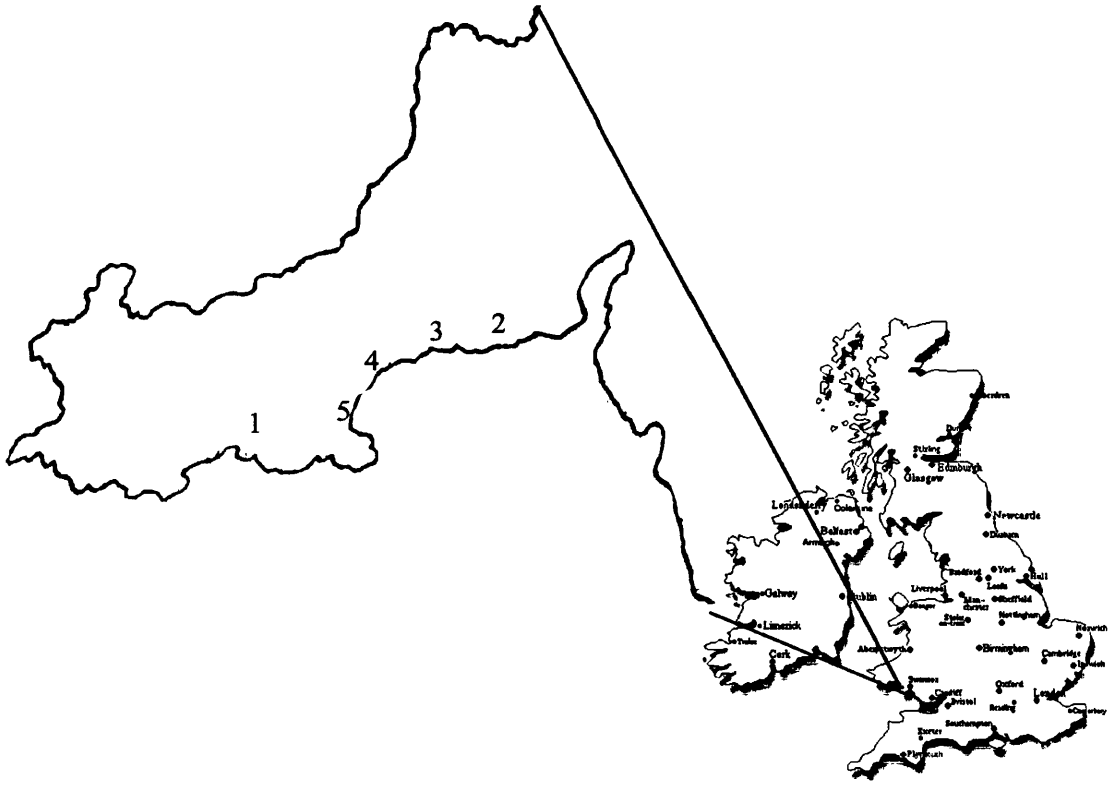


Figure 3.1 Field data collection and survey sites in Swansea, South Wales.

3.2.2 ***Phenotypic characteristics measurements.*** Fieldwork was carried out in Swansea (UK) and Crete to obtain information about the phenotypic characteristics of *H. incana* in its native and alien range. Plants were studied particularly in terms of life history (e.g. annual or perennial strategy) and physical output (plant size and seed production). Different sites were studied to observe whether substantial phenotypic variation occurred between populations from isolated locations across Crete. These characteristics (see below for details) were compared with plants from alien populations located in Swansea UK to identify if life histories varied between populations growing in contrasting environments.

Six populations of *H. incana* were studied in Crete during May 2006. These were selected to cover as large a geographical range across the island as possible during the visit. This was restricted to the Eastern half of Crete to allow outward and return travel from each site to be kept within a day.

Table 3.2 Sample sites for fieldwork carried out in Crete.

Site	Lat (N) & Long (E)	Altitude (m)	Substrate	Population
Wind Farm (1)	35°19'46 25°45'30	250	Clay , coarse gravel	1000+
Elounda (2)	35°15'33 25°43'26	20	Clay	1000+
Exo Lakonia (3)	35°12'21 25°39'13	160	Clay, gravel	80-100
Malia (4)	35°17'46 25°28'25	5	Sand, gravel	75
Irapetra (5)	35°00'13 25°45'56	5	Sand, fine gravel	50
Panorama (6)	Not recorded	NR	Clay, Gravel	70

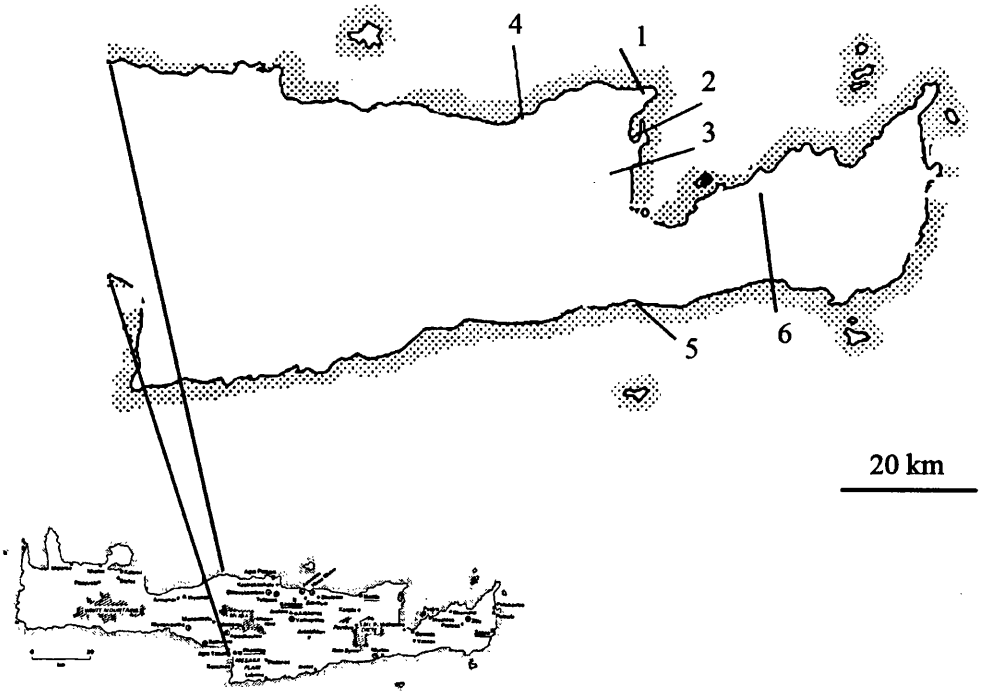


Figure 3.2 Field data collection sites in Crete.

The following morphometric parameters were measured for 10 plants per location:

- Number of basal leaves
- Leaf length, width, lobe length
- Number of basal branches
- Stem length
- Stem width
- Silique number, length and seed content

In 2008, this data was also collected from a further two populations from the Swansea area (Blackpill, Swansea n=10 and Pennard, Gower n=20) (sites 5, and 1 respectively from figure 3.1). Since there is evidence that *H. incana* frequently displays a perennial life history in the UK (Hipkin, unpublished observations), estimations of plant age were attempted by viewing

root sections under a dissecting microscope for plants from Pennard. However, due to the difficulty in resolving annual growth rings, it was only possible to distinguish between plants younger than one year and plants older than one year by the presence of secondary woody growth. This was sufficient to determine differences in seed output between annual and perennial forms of *H. incana*. For the same population in Pennard, counts were made of plants that had flowered and had remained as rosettes. This was performed in October 2008 and served to estimate the proportion of plants that did not flower within their first season of growth. Flowering stems for *H. incana* are typically produced in the UK between may-august, hence if a plant has not bolted by October it is unlikely to flower within that season. Root sections were not taken in Crete. In this location *H. incana* clearly displays an annual strategy; the plants become fully desiccated during the summer dry season.

3.2.3 ***Statistics.*** Means and standard deviations were calculated for each characteristic for each population. Differences in means between populations from Crete were investigated using the Kruskal Wallis test. Data collected from all Cretan populations were pooled as were data from South Wales. These data sets were used to provide a comparison of characteristics from native versus alien individuals and for seed production of one year old plants versus plants of greater than one year. The mean values for each characteristic for alien and native individuals were compared using the Mann Whitney test. Corrections were made for multiple comparisons by applying a sequential Bonferroni correction.

3.3 Results

3.3.1 ***Species diversity surveys*** (species names in bold grey text indicate non-native status in the UK).

Blackpill sea defence 07/05 (located at site 5, on figure 3.1)

Substrate; large boulders in-filled with sandy substrate.

Hirschfeldia incana
Diplotaxis tenuifolia
Hordeum murinum
Rumex crispus
Tripleurospermum maritimum
Achillea millefolium
Plantago lancelota
Hypochaeris radicata
Lotus corniculatus
Plantago coronopus
Lolium perenne
Festuca sp.
Leucanthemum vulgare
Crithmum maritimum
Cerastium glomeratum
Sonchus asper
Buddleja davidii
Elytrigia repens
Erodium cicutarium
Taraxacum officinale
Convolvulus arvensis
Clematis vitalba
Crepis capillaries
Oenanthe crocata
Reseda lutea
Cirsium vulgare
Trifolium repens
Cardaria draba
Senecio squalidus

Blackpill (located at site 5 on figure 3.1)

On beach, sand substrate + cockle shells, near strandline

Polygonum aviculare
Hirschfeldia incana
Chenopodium album
Tripleurospermum maritimum
Atriplex prostrata
Persicaria maculosa
Buddleja davidii
Carex arenaria
Poa annua
Cackile maritime
Medicago lupulina
Trifolium scabrum

Epilobium hirsutum	*Rumex crispus*	*Plantago major*
Elytrigia juncea	*Salsola kali*	*Honkenya peploides*
Coronopus didymus	*Cirsium vulgare*	*Eryngium maritimum*
Melilotus sp.	*Plantago coronopus*	*Juncus bufonius*

Swansea Bay marina; observatory, Dune by Building development. (Located at site 4 on figure 3.1)

Substrate: Sand

Cover: not 100% closed community. 30-40% bare sand.

Diplotaxis tenuifolia	*Eryngium maritimum*	*Aira praecox*
Lolium perenne	*Daucus carota*	*Arenaria serpyllifolia*
Plantago lanceolata	*Rubus caesius*	*Vulpia fasciculate*
Melilotus officinalis	*Anisantha madritensis*	*Hordeum murinum*
Melilotus albus	*Hypochoeris radicata*	*Artemisia vulgaris*
Anthoxanthum odoratum	*Crepis capillaries*	*Elytrigia sp. (repens)*
Coincya monensis ssp cheiranthos	*Fallopia japonica*	*Equisetum arvense*
Plantago coronopus	*Oenothera cambrica (parviflora)*	*Hirschfeldia incana*

SA1 species list (Located at site 3 on figure 3.1)

High proportion of gravel in soil – construction development site (disused waste ground)

Crepis vesicaria

Crepis capillaries

Agrostis stolonifera

Senecio vulgaris

Plantago major

Tripleurospermum maritimum

Medicago lupulina

Tussilago farfara

Plantago lancelota

Trifolium pratense

Trifolium repens

Taraxacum officinale

Sonchus oleraceus

Eupatorium cannabinum

Rumex crispus

Daucus carota

Sonchus asper

Melilotus altissima

Lotus corniculatus

Cirsium vulgare

Senecio squalidus

Holcus lanatus

Melilotus albus

Rubus caesius

Poa annua

Matricaria discoidea

Foeniculum vulgare

Artemisia vulgaris

Lolium perenne

Reseda luteolea

Hypericum perforatum

Rumex obtusifolius

Fabian Way species list (located at site 2 on figure 3.1)

Bare Ground = approx 30%

Agrostis stolonifera

Anagallis arvensis

Anthoxanthum odoratum

Anthyllis vulneraria

Arabidopsis thaliana

Arrhenatherum elatius

Artemesia vulgaris

Barbarea vulgaris

Bellis perennis

Buddleja davidii

Centaurea nigra

Cerastium fontanum

Cerastium glomeratum

Cirsium arvense

Cirsium palustre

Cirsium vulgare
Coincya monensis ssp cleiranthos
Conyza canadensis
Crepis vesicaria
Daucus carrota
Epilobium ciliatum
Epilobium tetragonum
Equisetum arvense
Festuca ovina
Festuca rubra
Foeniculum vulgare
Galium aparine
Geranium dissectum
Hirchfeldia incana
Hypericum perforatum
Leucanthemum vulgare
Linaria repens
Lolium perenne
Lotus corniculatus
Malva sylvestris
Matricaria discoidea
Matricaria recutita
Medicago lupulina
Melilotus officinalis
Oenothera cambrica
Papaver rhoeas
Pastinaca sativa
Picris echioides
Picris hieracioides
Plantago coronopus
Plantago laceolata
Plantago major
Poa annua
Poa pratensis
Poa trivialis
Potentilla reptans
Prunella vulgaris
Pulicaria dysenterica
Quercus ilex
Ranunculus acris
Ranunculus repens
Reseda lutea
Reseda luteola
Rubus caesius
Rubus fruticosus
Rumex crispus
Rumex obtusifolius
Sagina ciliata
Salix alba
Senecio jacobaea
Senecio squalidus
Sinapis arvensis
Sonchus asper
Sonchus oleraceus
Stachys sylvatica
Taraxacum officinale agg.
Trifolium pratense
Trifolium campestre
Trifolium dubium
Trifolium hybridum
Trifolium repens
Tripleurospermum maritimum
Triticum aestivum
Tussilago farfara
Verbascum thapsus
Vicia cracca
Vicia sativa

3.3.2 ***Phenotypic characteristics.*** Some plant characteristics were found to be variable between isolated populations of *H. incana* in Crete, where the species is considered to be native. For example rosette diameter and stem length varied between different locations (P=0.038, 0.018), but when Bonferroni corrections were applied for multiple comparisons, these results were not significant. These indicators of overall plant size are likely to be highly plastic and dependent on resources such as nutrients, water availability and linked to substrate (e.g. root constraint), and also period of growth between germination and the time the measurements were taken (data not available).

Mean seed number per fruit was significantly different between populations (P=0.002) in Crete, but overall mean seed output per plant was not significantly different between these same populations (P=0.122), (Table 3.3 & 3.4). When comparing native plants from Crete to alien plants from Swansea (UK) the most noticeable and ecologically important observation was the difference in seed output, Kruskal Wallis $P = 0.007$, however a sequential Bonferroni correction for multiple comparisons indicated that this result was not significantly different (Table 3.4). This observation was expanded on by comparing seed output from plants (growing in Pennard, UK) in their first year of growth and seed output from plants of age greater than one year (Table 3.5). A significant difference was observed in seed output between plants of these age groups (Table 3.6). A greater seed output was observed in plants of age +1 years than produced by plants in their first season in their alien range. However, seed output from plants in their first year of growth did not vary between populations from Crete and the UK. Out of 77 plants inspected from the population in Pennard, 28 had flowered and 49 had remained as rosettes. It is likely that the plants that were still rosettes in October would not flower until the following spring.

Table 3.3 Summary data for phenotypic characteristics measured from populations of *H. incana* in (A) Crete (native populations) and (B) Swansea (alien populations). For each character, the mean, ±SD and (N) are reported. Characters measured include; rosette diameter (rosette D), basal leaf number (Basal N), leaf length (Leaf L), stem width (Stem W), number of basal branches (Basal B), stem length (Stem L), number of siliques (Silique N), silique length (Silique L), number of seeds in siliques (Seed M) and estimated plant seed production (Seed Est). For certain fields it was not possible to collect data, this has been represented in the table as DNC.

(A)

Population	Rosette D	Basal N	Leaf L	Stem W	Basal B	Stem L	Silique N	Silique L	Seed M	Seed est.
Overall M	271±140 (76)	11.0±6.2 (76)	148±75 (58)	7.4±6.1 (88)	3.9±2.7 (88)	810±385 (88)	385±410 (87)	10.9±1.5 (68)	5.8±1.8 (85)	2205±2091 (85)
Wind farm	268±145 (9)	15.9±12.3 (9)	132±73.1 (9)	9.3±11.7 (9)	5.6±3.9 (9)	609±206 (9)	192±153 (9)	10.3±1.3 (9)	4.5±1.3 (8)	851±775 (9)
Elounda	276±128 (10)	10.6±4.7 (10)	160±73.5 (10)	6.1±2.3 (10)	5.5±4.1 (10)	870±175 (10)	386±345 (10)	11.1±1.7 (10)	5.9±1.8 (10)	2030±1332 (10)
Exo Lakonia	380±168 (10)	8.7±4.2 (10)	212±83.8 (10)	8.2±2.9 (10)	4.3±1.8 (10)	809±206 (10)	205±177 (10)	11.0±2.4 (10)	5.6±2.6 (10)	1326±1235 (10)
Malia	230±107 (10)	11.1±6.5 (10)	125±58.1 (10)	5.2±2.1 (10)	3.9±2.7 (10)	543±178 (10)	230±141 (10)	10.7±0.9 (10)	5.7±1.2 (10)	1385±1053 (10)
Ierapetra	190±82.1 (10)	7.6±3.4 (10)	101±28.3 (10)	5.0±1.9 (10)	3.3±3.1 (10)	624±137 (10)	231±164 (10)	10.1±0.8 (10)	5.5±1.2 (10)	1326±1083 (10)
Panorama	395±166 (10)	12.0±2.5 (10)	163±77.0 (10)	5.0±1.4 (10)	3.2±1.2 (10)	791±204 (10)	268±135 (10)	10.8±1.1 (10)	8.7±1,7 (10)	2280±974 (10)

Table 3.3 continued

(B)

Population	Rosette D	Basal N	Leaf L	Stem W	Basal B	Stem L	Silique N	Silique L	Seed M	Seed est.
Swansea	DNC	DNC	DNC	12.1±8.8 (10)	3.9±2.3 (10)	1022±368 (10)	1002±738 (10)	12.4±1.0 (10)	4.6±1.4 (10)	4224±2544 (10)
Pennard	219±85 (18)	11.6±4.8 (18)	DNC	8.1±6.8 (20)	2.7±1.5 (20)	856±311 (20)	459±359 (19)	DNC	5.6±1.6 (19)	2943±3028 (19)

Table 3.4 Kruskal Wallis *P* values for comparisons of characteristics between populations of *H. incana*. Comparisons have been made between populations from within Crete to observe if character differences are apparent within a relatively small native area. Also, comparisons have been made based on measurements taken and pooled from Cretan populations (native) versus measurements taken and pooled from South Wales populations. This has been performed to identify if differences occur in phenotypic characteristics from sample sites from the species native range and between native and alien populations. *P* values lower than 0.05 are highlighted in bold.

Character	Rosette D	Basal N	Leaf L	Stem W	Basal B	Stem L	Silique N	Silique L	Seed M	Seed est.
Crete	**0.038**	0.123	**0.023**	0.133	0.486	**0.018**	0.658	0.620	**0.002**	0.122
Native vs alien	0.094	0.249	DNC	0.097	0.10	**0.006**	**0.006**	**0.001**	0.068	**0.007**

Table 3.5 Estimation of seed production for plants of different age surveyed in Pennard, (nr Swansea, UK) and Crete. Plant age estimated by root section analysis.

Population	Plant age (years)	Mean (SD) seed estimate
Crete	< 1	1535 (1159), n=59
Pennard	< 1	1302 (992), n=11
Pennard	> 1	5200 (3481), n=8

Table 3.6 Mann – Whitney pairwise comparison of seed output from plants from Crete and Pennard (UK) of two age groups.

	Crete < 1	Pennard < 1
Crete < 1		
Pennard < 1	0.441	
Pennard > 1	4.8×10^{-4}	0.002

Tables 3.5 and 3.6 indicate the relative difference in seed output from *H. incana* plants of ages of less than one year and of greater than one year (i.e. plants that have flowered and produced seed in their first year of growth compared to plants that have either produced flowers and seed for a successive year, or have not flowered until their second year of growth). Data from plants studied from various locations from Crete were pooled together as seed production was not found to significantly differ between localities. Differences in seed production were identified between plants of differing age class within the same location P=0.002 (Pennard, UK). In this population, plants of an age greater than one year produced significantly more seeds than those of an age less than one year. In addition, the same pattern was observed when comparing plants of an age greater than one year from Pennard to plants of an age less than one year from Crete $P=4.8 \times 10^{-4}$. When comparing plants from Crete versus Pennard of age less than one year, no significant difference was observed in seed output. This result indicates a fitness advantage of a perennial life history where environmental conditions permit survival of greater than one year.

3.4 Discussion

Of the sites surveyed species numbers varied between 29-85 and the proportion of non-native species was between 14-29%. The greatest number of species was found at site 2 (Fabian Way). This was a disturbed development site adjacent to a main road and would be expected to contain a high number of early colonising species, potentially including alien species (Hobbs & Huenneke 1992). Two sites contained only 24 species (Blackpill sand dune and Swansea Marina sand dune). The sand dune at Swansea Marine was directly adjacent to a new development and contained the highest proportion of non-native species (7/24). Although no species abundance data was recorded, high densities of *H. incana* were observed on the urban gap sites with the species remaining less frequent on sand dune systems.

Phenotypic traits measured from isolated populations of *H. incana* across Crete were not found to vary greatly between populations. The variation that was observed was as great within populations as between populations. This result is not particularly surprising as the environmental conditions between sites were very similar. The only trait that was identified as significantly different between populations was seed number per fruit. This difference was probably a result of the higher seed count observed at site six (panorama). However, this did not make an overall difference in total plant seed production between populations. An implication of the lack of trait variation in *H. incana* across Crete is that if seeds have been supplied from a single locality, it is possible that they will be representative of plants from the whole region.

In Crete, *H. incana* was exclusively found in sites that had evidence for recent disturbance. It was observed as an agricultural weed, growing on roadsides (of recently constructed roads), and in areas of development and industrial activity. *H. incana* would appear to grow exclusively in areas of recent disturbance in both its native and alien range. Casual field observations in Crete and Swansea indicated that *H. incana* may endure for a longer period of time at individual sites in Crete than in Swansea. This may be explained by slower successional progress in Crete resulting in the presence of bare

soil for a greater period of time. In Swansea, successional changes appear to occur at a greater rate and *H. incana*, although able to rapidly colonise and dominate a newly opened habitat, may be outcompeted by other species within a few years.

When performing a comparison of characteristic traits between alien and native populations of *H. incana*, it may be expected that greater differences would be observed than for example between populations within a native region such as Crete. Initially, results indicated that indeed significant differences had been identified between alien and native individuals in the following traits: stem length, silique number, fruit seed number and plant seed output. However, performing a sequential Bonferroni correction for multiple comparisons indicated that in this instance these differences observed were not statistically significant. A contributing factor for this result could be the small sample sizes taken from each population (10-20) and the number of traits compared (10). Larger sample sizes would be beneficial in this instance and add greater reliability to statistical tests. In addition, data taken from Crete were all from annual plants, data taken from Swansea were taken from plants of both greater than one year's growth and less than one year's growth.

The major difference observed between *H. incana* plants from Swansea and Crete was life histories. In Crete all individuals acted as annuals, where in contrast plants in Swansea showed the ability to perennate. *H. incana* observed growing in Crete (one of its native regions) must flower and produce seed within one season. The summers are too arid for the survival of this species, a strict annual life history is observed in combination with rapid flowering and an advanced season. In contrast, in Swansea the species has the capacity to survive for successive seasons and does not always flower within its first year of growth. Gradients in flowering time with changing latitude have been documented for other species, e.g. *Arabidopsis thaliana* (Lempe *et al.* 2005) *A. lyrata* (Riihimaki & Savolainen 2004). Clinal variation in floweting time has also been documented for alien neophyte populations of *Lythrum salicaria* in orth America (Montague *et al.* 2008). A 100 m^2 stand of *H. incana* in Pennard (nr Swansea, UK) was surveyed during October 2008 to identify the proportion of plants that had flowered during the current growing

season and those that would over-winter as rosettes. Out of the 77 plants in this area, 28 had flowered and 49 remained as rosettes. *H. incana* provides another example of a species whose life history is dependent on habitat and/or environment. Similar observations have been reported for *Beta vulgaris* by Hautekeete *et al.* (2002), where a gradient in life span was observed across a latitudinal range.

Fecundity and propagule pressure are important components of an individuals or a population's fitness (Woods *et al.* 2009). Increased levels of propagule pressure increase the possibility of invasion events (Lockwood *et al.* 2005). Seed production was found to be greater in plants of greater than one year old compared to one year old plants. This capacity for the alien *H. incana* to over-winter in the UK greatly enhances its potential to produce a higher seed load. This may be an important factor in the successful colonisation of *H. incana* in the UK.

3.5 References

Blair AC, Wolfe LM (2004) The evolution of an invasive plant: an experimental study with Silene latifolia. Ecology, 85, 3035-3042.

Berglund ABN, Dahlgren S, Westerbergh A (2004) Evidence for parallel evolution and site-specific selection of serpentine tolerance in Cerastium alpinum during the colonisation of Scandinavia. New Phytologist, 161, 199-209.

Hautekeete N-C, Piquot Y, Van Dijk H (2002) Life span in *Beta vulgaris* ssp. *Maritime*: the effects of age at first reproduction and disturbance. Journal of Ecology, 90, 508-516.

Hobbs RJ, Huenneke LF (1992) Disturbance, diversity and invasion: Implications for conservation. Conservation Biology, 6, 324-337.

Jalas J, Suominen J (1996) Atlas Florae Europaeae, Distribution of vascular plants in Europe 11 Cruciferae, Ricotia to Raphanus. The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. Ecology Letters, 7, 1225-1241.

Kollmann J, Banuelos MJ (2004) Latitudinal trends in growth and phenology of the invasive alien plant Impatiens glandulifera (Balsaminaceae). Diversity and Distributions, 10, 377-385.

Lockwood JL. Cassey P, Blackburn T (1990) The role of propagule pressure in explaining species invasions. Trends in Ecology and Evolution, 20, 223-228.

Maron JL, Vila M, Bommarco R, Elmendorf S, Beardsley P (2004) Rapid evolution of an invasive plant. Ecological Monographs, 74, 261-280.

Montague JL, Barrett SCH, Eckert CG (2008) Re-establishment of clinal variation in flowering time among introduced populations of purple loosestrife (*Lythrum salicaria*, Lythraceae). Journal of Evolutionary Biology, 21, 234-245.

Muller-Scharer H, Schaffner U, Steinger T (2004) Evolution in invasive plants: implications for biological control. Trends in Ecology and Evolution, 19, 417-422.

Phillips BL, Shine R (2004) Adapting to an invasive species: Toxic cane toads induce morphological change in Australian snakes. Proceedings of the National Academy of Sciences of the United States of America, 101, 17150-17155.

Preston CD, Pearman DA, Dines TD (2002) New atlas of the British & Irish flora : an atlas of the vascular plants of Britain, Ireland, the Isle of Man and the Channel Islands. Botanical Society of the British Isles, Oxford University Press.

Reznick D, Ghalambor CK (2001) The population ecology of contemporary adaptation: what empirical studies reveal about the conditions that promote adaptive evolution. Getetica, 112-113, 183-198.

Richardson DM, Pyšek P, Rejmánek M, Barbour MG, Panetta F, West C (2000) Naturalisation and invasion of alien plants: concepts and definitions. Diversity and Distributions, 6, 93-107.

Riihimaki M, Savolainen O (2004) Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany, 91, 1036-1045.

Woods TM, Hartnett DC, Fergusos CJ (2009) High propagule pressure and reproductive fitness homeostatis contribute to the invasiveness of *Lespedeza cuneata* (Fabaceae). Biological Invasions, 11, 1913-1927.

Chapter 4

Variation of phenotypic characteristics between populations of *H. incana*

4.1 Introduction

Specialisations in the life history patterns of plants may evolve to match the local environmental conditions of their geographical area (Pigliucci & Marlow 2001). When this happens there may be trade-offs between phenotypic traits. The observed phenotypic variation may be a result of genetic differentiation (e.g. Osborn *et al.* 1997, Schranz and Osborn 2000, Koornneef *et al.* 2004) or be a component of phenotypic plasticity (e.g. Schlichting 1986, Callahan 2005, Lacaze *et al.* 2009). Both plastic and non-plastic ecotypes can evolve specialisations to increase potential fitness within specific habitats (Pigliucci & Marlow 2001). This ability to adapt to a range of environmental conditions is relevant to understanding the adaptive potential of alien and invasive species and may aid the formulation of management policies for the early control of invasive species.

Establishment of species into new regions is a colonisation process and understanding this process is important for understanding invasive biology (Hurka *et al.* 2003). Several themes are commonly studied with respect to non-native invasive species; phenotypic origin, colonisation history, genetic diversity and adaptation. All have roles that can contribute to invasion success. These factors may prompt the question of the origin of invasiveness, whether acquired during the colonisation process after the introduction event, or if the species in question was pre-adapted to the environment of introduction, or indeed if certain traits promote invasiveness. For example, *Capsella bursa-pastoris*, a successful invader with a near world-wide distribution, has had its success attributed to polyploidy, high genetic diversity, a predominantly selfing strategy (Hurka *et al.* 2003), and ecotypic adaptation (Linde *et al.* 2001) resulting in the introduction of pre-adapted or "general purpose" genotypes (Neuffer & Hurka 1988, Neuffer & Hurka 1999). Ecotypic adaptation giving rise to colonisation success for *Capsella bursa-pastoris* may be attributed to differences observed in the expression of genes that promote the change from a vegetative to a flowering state (Linde *et al.* 2001). Other traits associated with invasive potential include a short juvenile period, a large seed crop and small seed mass (Rejmanek & Richardson 1996). Adaptive traits linked to invasive success may include

variation in flowering time (Linde *et al.* 2001) and specific leaf area (Smith & Knapp 2001).

A common observation with alien invasions is that, after initial colonisation by the species, there is a "time lag" prior to rapid growth of the population followed by range expansion (Sakai *et al.* 2001). This time delay to population growth may simply be an ecological function relating to the commonly observed lag phase of an exponential growth curve. However, it is hypothesised that adaptation to local environmental conditions during this time may be important in developing invasive potential (Parker *et al.* 2003). Evolutionary processes or evolutionary potential may complicate the predictability of this species invasive potential in novel habitats. In the case of widely successful invaders, high levels of genetic diversity (Stepien *et al.* 2005, Kolbe *et al.* 2004), large founder populations (Lockwood *et al.* 2005) and multiple introductions (Frankham 2005) are thought to promote a more rapid response to selection.

A number of studies have provided evidence for local adaptation among populations for a range of phenotypic traits among invasive species (Kollmann & Banuelos 2004, Maron *et al.* 2004). Common garden studies can determine if observed phenotypic differences have a genetic basis (e.g. Reinartz 1984 a,b,c, Siemann & Rogers 2001) while reciprocal transplant experiments can indicate relative fitness for a range of habitats (Kawecki & Ebert 2004). With common garden experiments, it may be possible to distinguish between 'general purpose genotypes' – as described by Baker (1965), that successfully colonise new areas by high levels of phenotypic plasticity – and invaders that have spread by rapid adaptation involving changes in the genetic profile of the population.

H. incana, the study species here, is native to Mediterranean and Middle Eastern countries, but has undergone a huge expansion in its range and is currently found in a diverse range of environments. Within Europe, the species has extended its range as far north as Sweden (Jalas and Suominen 1996). Across its current range there is considerable geographical variation in both climate and day lengths. Previous observations have indicated that it displays differences in its life history in the wild

between southern native populations and northern alien populations (data not published, details provided in chapter 3) and that flowering time has differed considerably between native and alien individuals when grown under greenhouse conditions (Hipkin pers. comm.). Populations of *H. incana* have previously been studied in the wild in Crete (native) and in South Wales (alien) by the author. Cretan individuals were observed follow a strict annual life history in contrast to a strategy of short lived perennials in South Wales. The occurrence of perennation was observed with an increase of plant size and in seed production, which may be key attributes to its success as an alien colonizer. Thus, even though *H. incana* at the moment is not known to have a significant negative impact in the UK, it may still serve as a useful study model for an introduced, outcrossing and diploid plant of Mediterranean origin, that has successfully colonised the UK landscape, is recently naturalised and is currently expanding its range. Understanding the basic mechanisms of colonisation could be important for understanding how such types of plants could invade the UK, which could then aid in formulating methods for controlling undesirable invasions. Also, as it is likely to continue to increase its range in the UK *H. incana* may still have the potential to impact on the native flora if begins to encroach on more natural environments.

This study aims to explore the variation in phenotypic traits from a variety of populations of *H. incana* from its native and alien range under common garden conditions to assist with explaining its colonisation success. Comparisons were made to determine whether differences in life history traits are observable in plants representative of its native and alien range. Furthermore, investigations using common garden experiments aimed to establish if this variation is a product of phenotypic plasticity or if local adaptation is evident. Levels of phenotypic plasticity and the influence of local environmental conditions will be examined. Focus will be placed on the life history strategies of plants from native and alien populations grown under greenhouse conditions and under the local environmental conditions of South Wales. Observations of life history strategies of *H. incana* grown under these conditions may give an indication of the role of adaptation on the success of the alien populations. If phenotypic differences or differences in life

history strategy are observed between populations under common garden conditions, it may highlight the areas where adaptive change has occurred.

4.2 Method

4.2.1 ***Seed selection.*** Seed from 10 geographically distinct populations of *H. incana* were selected to be grown for phenotypic comparison (Table 4.1). Seeds were collected locally and sourced from seed collections. Included in the selection were populations from both the native and introduced ranges. From the native range continental populations were sourced from both Europe and North Africa and two island populations were included from the Mediterranean. The alien range was represented by seed sourced from Swansea (UK) – 3 localities within a 5 km radius, and Los Angeles (USA). The seed collected from the three locations in Swansea (see figure 3.1, Table 3.1) have been given different identities to allow for separate analysis, but it is possible that they are subsets of one population. Among the samples selected there are coastal and inland populations represented. Populations were also selected where possible to cover as large a range of latitude and altitude as possible. Where possible, climate data was sourced from nearest available weather stations for each population studied (Table 7.2, http://www.climate-charts.com/).

Table 4.1 Source of seed used for phenotypic comparison (listed in order of increasing altitude).

Name	Location	Country	Latitude	Altitude	Status
Mor1	Mohammedia	Morocco	33°43	10	Native
USA	L.A.	USA	33°47	25	Alien
Alg2	Saida	Algeria	34°49	800	Native
Cyp	Larnaka	Cyprus	34°55	6	Native
Alg1	El Bier	Algeria	35°	100	Native
Crete	Elounda	Crete	35°17	5	Native
Mor2	Taza	Morocco	36°41	50	Native
Turk	Usak	Turkey	38°33	850	Native
Spain	Bejar	Spain	40°23	900	Native
Swan2	Swansea	Wales	51°37'12	5	Alien
Swan3	Swansea	Wales	51°37'37	5	Alien
Swan4	Swansea	Wales	51°36'52	5	Alien

4.2.2 ***Seed germination.*** Germination time and success were estimated for *H. incana* under greenhouse conditions. Seeds were collected from Swansea and sourced from seed collections (Table 4.1 and further details in table 8.9). Mean seed mass was calculated for each population. Seeds from each population were batch weighed and the number of seeds then counted, from this method a mean could be calculated, but no range of population variation could be determined. The collection strategy of supplied seeds was unknown, so it was not possible to determine if all the seeds from distinct populations originated from one or more mother plants. Two germination strategies were used; germination within multi-purpose compost in 5" pots (February 2005) and germination on moist filter paper within a Petri dish (May 2005). The second set of seeds was germinated on Petri dishes rather than in compost to save space and to potentially provide more accurate germination results. When seeds germinate in compost some inaccuracy may me introduced if there are any differences in seed depth within the compost, the use of Petri dishes for germination allows the germinating seed to be viewed directly.

Due to limited stock of seed, 20 seeds from each population were initially used to estimate germination rate per population per germination strategy. Those seeds germinating in compost were lightly covered with compost and watered daily. Seeds on filter paper were kept moist with daily watering. Lids were kept on the Petri dishes to minimise evaporation. Seeds were kept under greenhouse conditions at 18°C under ambient light with an additional artificial light source of 12hr light/dark cycles. Germination was recorded daily over a 30 day period. Mean germination rate and germination success was calculated for each population for both treatments (Table 4.5).

4.2.3 ***Measurement of phenotypic traits.*** Plants resulting from the germination observations were used for comparison of phenotypic traits. The seeds germinated in multi-purpose compost remained in 5" pots and were grown in the greenhouse. The seeds germinated in Petri dishes were transferred temporarily to 1" pots with multi-purpose compost prior to being planted outdoors in an experimental plot at the beginning of July.

Growing conditions were kept as consistent as possible within the greenhouse. Growth lamps were used to supplement the low February light levels, and were set to a 12 hour light/ dark cycle. The lights were to remain on this setting for the duration of the experiment, although the duration of ambient daylight would have increased to a maximum of approximately 17 hours in June. The greenhouse temperature was set at 18°C and was kept at or above this minimum. Regulating the maximum temperature was not possible and during sunny and windless days, temperatures frequently exceeded 30° and reached an observed maximum of 39°C. Plants in the greenhouse were watered daily in some instances it was necessary to water twice daily in hot weather.

Plants were grown over two benches with up to 10 individuals from each population in adjacent rows along the benches. This arrangement was decided upon to allow recognition of patterns to be observed along rows of the same population, which may not have been noticed and recorded had they been grown in a random block design.

To determine if there were observed differences in the form and life history between geographically distinct populations of *H. incana* the following characteristics were recorded:

- Germination rate and success
- Time taken from germination to first open flower (time t)
- Height of flowering stem at the time t
- Number of basal rosette leaves (all individuals were measured on 01/04/05- this was the commencement of flowering stem development in some individuals)
- Leaf dimensions; including length (rosette radius), width and terminal lobe length (measured on 07/04/05)
- Mean number of lobes on leaves from each plant (measured on 08/04/05)

Data collected from these characteristics were initially analysed with a principle component analysis to identify where differing characteristics had the greatest influence and how these characters spatially separated the populations when these components were plotted (SPSS version 13). The characters were then compared for within and between population variation using a Kruskal Wallis test. A Mann Whitney U-test pairwise comparison was used to provide a *post hoc* analysis of the Kruskal Wallis test. This was used to determine where population differences occurred. This was performed for flowering time and basal leaf number data. Sequential Bonferroni corrections were applied to correct for multiple comparisons. A Pearson correlation was performed between leaf number and time (between germination and data collection), this was performed to identify if a correction factor could be applied for leaf number between plants that had germinated at different times.

In addition to the growth experiment in the greenhouse, a plot was prepared in the University Botanic Garden to observe patterns of growth between populations in local environmental conditions. Ideally, a minimum of 10 plants from each population would have been planted, but due to some mortality on transfer to pots this was not achieved for all populations.

Due to the position of a greenhouse directly south of the plot, shading was greater in the area closest to the greenhouse. To account for this gradient in available light and soil moisture five individuals from each population were arranged in a random block, the remaining plants were placed in rows according to population. Plants were only watered for the period of acclimation to outdoor conditions, after which precipitation was the only water source. Natural light was the only light source and at the start of the trial was approximately 15/9hr light/dark cycle, during midsummer this would change to a ratio of 17/5 and by late August to a ratio of 14/10 (by which time all those plants that were to flower in that season had done so).

The following measurements were taken from the plants grown outdoors.

- Time taken from germination to first open flower (time t)
- Height of flowering stem at time t
- Dimensions of largest leaf at time t (rosette radius); including length, width and terminal lobe length
- Number of basal leaves at time t

The characters measured were compared for within and between population variation using a Kruskal Wallis test.

Some changes were made to the way the measurements were taken from the plants grown in the greenhouse. Leaf number and dimensions were measured at time that each plant produced its first open flower to determine the relative allocation of investment in early flowering versus vegetative growth. It is common practice in the literature, that characters are measured at a specific life stage of development e.g. plant bolting or flowering. So it was decided that although the data collected would not be the same as that taken in the greenhouse, it would be useful to use this method in order to be comparable with similar published studies of other species. Furthermore, this approach also provides additional vital information. For example, it was expected that a plant flowering early would produce less vegetation than a plant flowering later in the season. In addition if there were individuals that had an annual life history they would be expected to have a more

rapid flowering strategy than a perennial plant that may invest in increased vegetative growth. Measurements taken at this stage give an indication of total vegetative growth at time of flowering, this compares to the leaf measurements taken at an individual point in time in the greenhouse, which may indicate rate of vegetative growth.

As there were differences in time taken from germination to flowering within populations between greenhouse and outdoor treatments, paired t-tests within populations were performed to identify where differences had occurred. The *P* values were subjected to sequential Bonferroni corrections. Explaining the variation within the flowering times between the two treatments was not possible due to the number of environmental differences in their cultivation conditions. A Wilcoxon signed rank test was performed on the difference observed in flowering time within populations. This test was done to identify if observed changes were all in the same direction, or if different populations reacted in the change in environment in different manners.

4.2.4 ***Self-pollination and pollination within populations.*** Although self incompatibility of *H. incana* has been previously described by Darmency and Fleury (2000), an attempt was made to observe the effects of self pollination of *H. incana* for the populations grown in the greenhouse. Flowers from plants in each population were covered with muslin bags to exclude pollinators cross pollinating between populations. In addition two $1m^3$ frames covered with fine mesh were used to contain 3 plants from Swansea and 3 plants from Morocco. These populations were selected to represent a population from both the native and alien range, a greater number of individuals and populations were not used due to space constraints. Comparisons in flowering strategy were observed between these populations. The aim was to collect seed that had been produced from within population crosses for a repeat of the character measurements to check that there were no observed maternal effects from the original source seed influencing the results previously seen. These plants were frequently agitated in an attempt to achieve pollination. Hand pollination was also attempted with a small paint brush. In both cases (self pollination and cross pollination) seed formation was low, but seed was collected to attempt germination. However, due to low numbers of seed germination and an outbreak of pests

the trial did not yield any usable results. This trial was not repeated as the source plants that would have potentially provided seed were also affected by pest attack.

4.2.5 ***Analysis of petal and leaf shape.*** When *H. incana* was growing during the phenotypic trait trials, different populations had characteristic features. As the individuals from each population were grown in rows, it was easy to observe subtle differences in appearance between each population. As no significant differences in characters were apparent between individuals grown outdoors from the same population, whether grown in rows of in a random block design, it was decided that there were no row effects inducing these differences. The plants' leaves, and in some cases petals, appeared to be more similar within populations than between populations. These similarities were not easy to describe with simple measurements such as length, width or lobe(s) dimensions. One additional problem with simple measurements would be differences in leaf size due to age or leaf or variations in growth rate. Verbal descriptors of shape difference would be difficult to analyse in a statistical manner, and may be biased due to observer interpretation. One method of mathematically describing shape outlines uses elliptic fourier descriptors (Kuhl & Giardina 1982). Different shapes can then be discriminated by principle component analysis. The elliptic fourier developed by Kuhl allows the analysis of complex shapes, where radii emanating from the centre of the shape may cross the outline more than once (Le Minor & Schmittbuhl 1999). Shape version 1.3 (Iwata & Ukai 2002) is a package specifically designed for the quantitative evaluation of shapes, primarily aimed at biologists.

For this analysis one leaf and one flower was removed from each of 5 individuals from the following five populations (Swansea 2, Swansea 3, Elounda - Crete, Irapetra - Crete, and Bejar - Spain. These populations were chosen as leaves from Crete and Swansea plants are visually different, and leaves from individuals from Bejar, Spain appeared to be most similar to those from Swansea (pers. obs). This selection was made as a pilot study to identify if the method would associate populations that were more similar in appearance (e.g. it was anticipated that the two populations from Swansea and Spain were more similar in appearance to the two populations from Crete). If the technique

proves useful, this could be repeated using a greater number of individuals from the whole range of populations to perform a more comprehensive study in the future.

Leaves and flowers from each plant were scanned into Photoshop version 7 using a Canon Canoscan N650U flatbed scanner at 300 dpi. Images were manipulated to maximize contrast between leaves and background. This was achieved using the channel selector, turning the image to monochrome, and altering red, green, blue and contrast until the desired effect was achieved. This would lead to an easier contour identification in the Shape package. Images were saved as full colour bitmap files for import into Shape. A 15mm black square was also scanned to provide a scale. Default settings were used in the Shape package with exception to the Ero Dil Filter. This was set to the value 3 which removes imperfections introduced in scanning, e.g. pixel noise (see Shape User manual, Iwata and Ukai 2002). Principle component analysis was performed and principle components 1 and 2 were plotted against one another to see if any clusters were apparent.

4.2.6 *Environmental responses*

4.2.6.1 Effect of vernalization. To investigate if vernalization (cold treatment) affects flowering time in *H. incana* two populations were selected for comparison (Swansea and Crete, Table 3.2). These populations were selected due to the observed differences in flowering time based on the observations from other growth experiments detailed in this chapter (short flowering times for individuals from Crete versus long flowering times for Individuals from Swansea). Other considerations were seed availability and the time frame that a growth chamber was available to perform the experiments (seeds were available from Crete and Swansea). Seed from Elounda (Crete), which was previously used in the greenhouse, was not available, so an alternative population (Irapetra, Crete) was used in its place. Individuals from both populations were grown in a plant growth chamber. This was programmable and allowed the temperature, light intensity and light duration to be set. Plants grown for these trials were split into two groups. Half of the individuals from each population were split for cold treatment (removed to a constant temperature room) and half for no cold treatment (remaining in the growth chamber). For the trial, day length was set at 16 hours light 8 hours dark. Light intensity was set to

maximum (approximately 7,500 lux). Temperature was constant within the chamber at 20°C. Cold treatment at 4°C in a constant temperature room was given to seedlings after 21 days from germination. At this time approximately 4-6 leaves had developed on each plant. These plants were subject to this cold treatment for 30 days and then returned to the growth chamber at 20°C. Light duration was 16 hours for both groups. Ideally, light intensity would have been the same for both groups. However, equipment available for the cold room produced 2,500 lux. Time was measured from germination until flowering for all individuals.

Table 4.2 Samples compared for the effects of vernalization.

	Sample size	
Population	Vernalized	Non-vernalized
Swansea, Wales	20	20
Irapetra, Crete	6	7

4.2.6.2 Survival of *H. incana* through British winter. *H.* incana is frequently described as either an annual or perennial, and may over winter in some localities. *H. incana* grows over a large geographical range and may or may not be subjected to cold temperatures dependant on location. It is possible that variation in cold tolerance in *H. incana* occurs between populations. To investigate this further, five populations of *H. incana* were selected to determine the effects of frost/ winter conditions on survival (Swansea2, Elounda – Crete, Larnaka – Cyprus, Mohammedia – Morocco and El Bier - Algeria). Seeds were sown in the greenhouse in 1" pots in autumn 2006 and when large enough to handle were transplanted outdoors into the experimental beds in the botanic garden. Plants were grown in rows according to population for ease of observation. The plants were subjected to several frosts during the winter and were covered with snow for approximately 3 days. Plants were checked at least weekly (often more frequently) for damage. The surviving plants were observed for time taken to flower. These data were

combined with the flowering time data from the other growth trials to produce a chart displaying the differences in flowering time between the populations under different environmental conditions.

4.2.7 ***Climatic data.*** Climatic data were collected for all of the locations where *H. incana* seed had been sourced for this study. Using Google Earth satellite mapping facility and www.climate-charts.com/ it was possible to identify the closest weather stations for each of the populations. For each site (where available) mean monthly maximum and minimum temperatures and mean monthly precipitation were reported.

Relationships between phenotypic characters and climatic conditions were investigated by performing Pearson correlations. Each character (from plants grown in the greenhouse and grown outdoors) was investigated for a relationship with mean temperatures and rainfall.

4.3 Results

4.3.1 ***Phenotypic Variables.*** Phenotypic characteristics of *H. incana* were measured for comparative purposes between populations. Within the raw data it was possible to see that variation was apparent among phenotypic traits observed across a range of populations of *H. incana* when grown under two contrasting common garden conditions (see table 7.3, 7.4). These results were then used for successive statistical analysis to determine if populations could be clustered based on these differences and whether significant differences among these traits between populations were apparent.

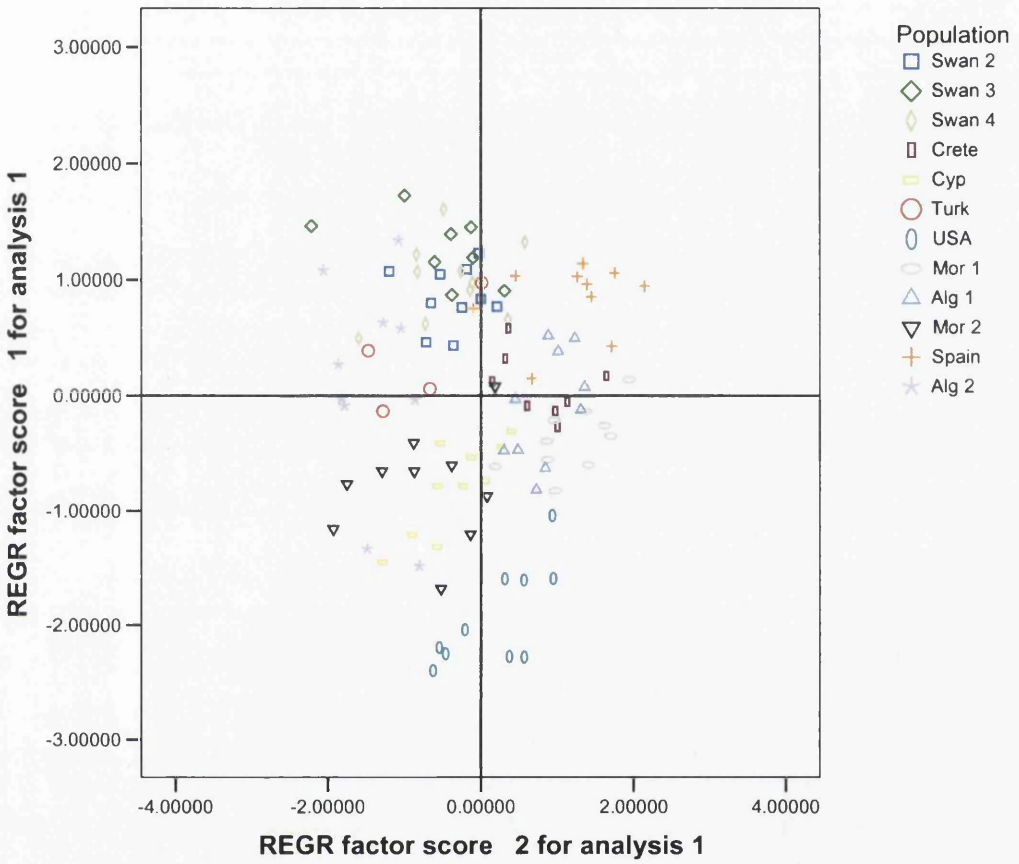


Figure 4.1 Plot of principle components 1 and 2 of a data reduction of characters measured in the greenhouse. The characters included; germination time, leaf number, rosette radius, stem length and flowering time. The differences in principle components produced separation of individuals when they were plotted, these were seen to cluster according to populations. Individuals from Swansea were positioned closest to individuals from populations Alg 2, Spain and Turk. Individuals from populations USA, Mor 1, Mor 2, and Cyp were positioned furthest from Swansea.

Table 4.3 Total variance explained by the principle component analysis for plant characters from the greenhouse.

	Initial Eigenvalues	
Component	Total	% of Variance
1	2.23	44.66
2	1.32	26.32
3	0.67	13.48
4	0.51	10.27
5	0.26	5.28

Table 4.4 Component matrix (factor loadings) produced by the PCA for plant characters from the greenhouse.

	Component	
	1	2
Germination time	-0.782	-0.028
Leaf number	0.763	-0.525
Rosette radius	0.577	0.531
Stem length	0.309	0.837
Flowering time	0.782	-0.238

The output from the principle component analysis (figure 4.1, Tables 4.3, 4.4) indicates that the traits measured were sufficiently different so as to produce a non-random distribution of individuals, based on a plot of the principle components. Individuals from the three localaties in Swansea were clustered together in the upper left of the chart. Other populations that shared similar principal component values and were plotted closest to (or overlapping) the individuals from Swansea included Alg 2, Spain and Turk The populations that had values least similar to Swansea, and were plotted at the greatest distance from Swansea, include USA, Mor 1, Mor 2 and Cyp. Table 4.3 lists the Eigenvalues calculated by the principle component analysis. This displays the proportions of the variation explained by the components. The first two components

explain 71% of the observed variation. Table 4.4 gives the magnitude and direction of the associations of the characters and can be used to indicate which characters are driving the variation. The values in table 4.4 are Pearson correlations of the individual's character measurement with the calculated value of the principle components for each individual. Component one (explaining the greatest proportion of the variation) has a high negative correlation value with germination time (-0.782), in contrast to positive values with all of the other traits. This can be interpreted as, plants that germinated more rapidly accumulated a greater leaf number and rosette radius within the time period, developed longer flowering stems and took a greater number of days to come into flower. In addition, high leaf number or large rosette radius was associated with longer time from germination to flowering. The character 'stem length' had the lowest value for component one in the matrix, indicating that this character explains the least of the observed variation.

Table 4.5 Variation of character traits between 12 populations of *H. incana* grown under two common garden conditions, greenhouse and outdoors. (variation tested using Kruskal-Wallis). This table displays the between population variation of the characters measured. The data without parenthesis are for the plants grown in the greenhouse, the data within parenthesis are for the plants grown outdoors. Data were compared only within treatments to determine whether between population differences occurred for the characters; it is not intended that a comparison is made between the greenhouse and outdoor data in this instance.

Trait	mean	d.f.	Chi-Square	P
Germination time (days)	5.6 (6.0)	11 (10)	81.2 (49.7)	<0.001 (<0.001)
Rosette leaf number	13.7 (24.6)	11 (10)	115.2 (43.9)	<0.001 (<0.001)
Rosette radius (mm)	223 (234)	10 (10)	32.1 (35.0)	<0.001 (<0.001)
Stem length (mm)	558 (362)	11 (10)	102.6 (21.2)	<0.001 (0.019)
Flowering time (days)	100 (88.8)	11 (10)	118.0 (37.7)	<0.001 (<0.001)

High phenotypic variation was observed between populations of *H. incana* grown under common garden conditions. Table 4.5 shows the results of Kruskal Wallis analysis of

ranks of various characteristics measured from plants grown under common conditions within the greenhouse. ANOVA was not used in this case as parametric conditions were not satisfied. As can be seen from the results of the data analysis, significant differences were observed for all traits measured, indicating population differences for these traits.

For other characteristics measured (leaf width and leaf lobe length), significant differences were also observed between populations, but as these characters were both found to correlate with leaf length (or rosette radius) when grown in outdoor conditions (Pearson correlation 0.612, 0.516 respectively, $P < 0.001$ and n = 77 in both cases), and leaf length and width were found to correlate under greenhouse conditions (Pearson correlation 0.388, $P < 0.001$, n = 113) they were not included in table 4.5.

Table 4.6 Mann Whitney U-test pairwise comparison for between population variation of characters for greenhouse grown plants. In this table the characters compared are flowering time (below the diagonal) and Basal leaf number (above the diagonal). This test was performed to provide a *post hoc* analysis of the Kruskal-Wallis analysis to highlight where population differences occurred. Bold type indicates significant differences after sequential bonferroni correction.

	Swan2	Swan3	Swan4	Spain	Turk	Mor2	Crete	Alg1	Cyp	Alg2	USA	Mor1
Swan2		5.12E-01	1.17E-01	**5.20E-07**	8.37E-01	**5.10E-05**	**1.90E-05**	**2.30E-08**	**8.80E-10**	4.84E-01	**9.70E-09**	**2.30E-10**
Swan3	1.60E-02		4.73E-01	**4.10E-05**	6.45E-01	**4.00E-04**	**2.40E-04**	**2.10E-07**	**3.00E-08**	8.22E-01	**9.97E-08**	**9.77E-09**
Swan4	7.20E-02	3.01E-01		**3.10E-04**	3.40E-01	2.00E-03	2.00E-03	**1.80E-05**	**2.70E-06**	7.80E-01	**8.80E-07**	**1.30E-06**
Spain	2.50E-02	**3.93E-04**	**1.20E-03**		1.40E-02	3.94E-01	1.23E-01	**5.80E-05**	**7.50E-05**	8.00E-03	**5.30E-08**	**1.20E-05**
Turk	1.24E-01	4.00E-03	1.50E-02	8.71E-01		4.00E-03	1.10E-02	**1.00E-03**	**1.00E-03**	6.82E-01	**1.00E-03**	**4.50E-04**
Mor2	**3.11E-06**	**2.65E-05**	**3.20E-06**	1.00E-02	6.00E-02		4.12E-01	1.10E-02	2.20E-02	1.10E-02	**8.00E-07**	2.00E-03
Crete	**2.86E-04**	**1.55E-04**	**3.93E-05**	**6.57E-05**	4.00E-03	9.68E-01		1.69E-01	3.96E-01	4.00E-03	**1.60E-05**	8.50E-02
Alg1	7.00E-02	**1.59E-05**	**1.94E-04**	9.36E-01	8.62E-01	**9.90E-05**	**6.30E-05**		3.25E-01	**1.80E-05**	**3.70E-05**	7.04E-01
Cyp	**1.33E-08**	**2.72E-06**	**5.89E-08**	**5.60E-07**	**4.12E-04**	2.94E-01	2.14E-01	**6.70E-08**		**1.20E-05**	**5.00E-07**	1.14E-01
Alg2	**1.70E-04**	**2.72E-06**	**2.01E-05**	1.70E-01	1.22E-01	1.00E-02	**1.00E-03**	5.90E-02	**3.20E-07**		**2.90E-08**	**1.00E-06**
USA	**5.34E-08**	**9.80E-06**	**1.92E-07**	**3.80E-07**	1.00E-02	**7.70E-05**	**4.40E-04**	**3.80E-07**	**9.90E-05**	**5.90E-08**		**9.00E-06**
Mor1	**1.10E-09**	**9.00E-07**	**5.76E-09**	**5.80E-09**	**2.30E-04**	**6.95E-07**	**3.60E-06**	**1.40E-08**	**1.80E-07**	**4.90E-10**	4.72E-01	

To test if planting method outdoors (random block design or populations in rows) gave rise to any within population differences in magnitude of measured characters a Mann Whitney U test was performed for each population for each trait, (Table 8.7). There was no significant difference for these characters observed between the two designs. This

indicates that at this small scale, where it was more convenient to view populations in rows, no discernable differences were observed when plants were grown in random blocks or in rows.

4.3.2 ***Germination time.*** Germination occurred rapidly and was within the range of 2-17 days. While there appears to be significant differences between populations, there may be factors such as seed age and mass that influence germination time. There was a trend of decreasing seed mass with increasing seed age from the seeds listed in Table 4.7 (Pearson correlation = -0.679, $P = 0.015$, n = 12). A partial correlation of germination time and seed age, with seed mass as a co-factor indicated a relationship ($P = 0.077$) between age of seed and germination time (although not significant to $P = 0.05$), which may be a factor influencing the observed population differences. Variation in germination time may be influenced by temperature (Steinmaus *et al.* 2000), seed dormancy (Bewley 1997), germination inhibitors (Evenari 1949), damaged or poor quality seed. Germination success was high (>60%) for all populations other than Elounda, Crete and Usak, Turkey (50% and 25% respectively).

Table 4.7 Seed parameters that may influence germination time or success, including seed mass (sample sizes in parenthesis) and age. Germination % refers to the final percentage of seed that succeeded to germinate from 20 seeds for each population for both treatments. Two values are given for germination success and germination time, seeds germinated in compost (no parenthesis) and Petri dishes respectively (data in parenthesis).

Population	Mean seed mass (mg)	Seed age (Years)	Germination (%)	Germination time (days)
Mor1	0.19 (119)	40 (1965)	95 (80)	4.4 (6.4)
USA	0.37 (192)	40 (1965)	70 (25)	11.0 (11.2)
Alg2	0.30 (106)	35 (1970)	80 (65)	5.4 (5.5)
Cyp	0.29 (244)	5 (2000)	90 (90)	5.8 (6.1)
Alg1	0.23 (145)	40 (1965)	60 (70)	6.0 (8.0)
Crete	0.37 (281)	3 (2002)	50 (30)	3.3 (5.2)
Mor2	0.24 (304)	37 (1968)	55 (65)	7.2 (5.5)
Turk	0.27 (333)	22 (1983)	25 (20)	3.8 (N/A)
Spain	0.29 (282)	34 (1971)	75 (85)	3.3 (4.6)
Swan2	0.53 (372)	<1 (2005)	90 (90)	4.4 (4.3)
Swan3	0.33 (286)	<1 (2005)	65 (60)	4.4 (7.6)
Swan4	0.51 (97)	<1 (2005)	75 (60)	4.1 (4.8)

4.3.3 ***Basal rosette leaf number and size.*** Number of basal rosette leaves recorded from the greenhouse when the first plant had bolted could be split into three categories, high (>15), medium (10-15), and low (<10) leaf production. Individuals from Swansea were among the populations that produced the most leaf growth (Figure 4.2a, Table 7.3). Individuals from LA, USA were observed to produce the fewest leaves (Table 7.3). A correlation was not identified between leaf number and time (time taken between germination and measurement of leaf number at day 62, Pearson $P = 0.49$). This result indicated that leaf production was not linear with time across the range of populations. It was therefore not possible to apply a correction factor to take into account the difference

in germination times between individuals. However, this does indicate that variation in leaf numbers may be related to population differences rather than period of growth.

There was considerable variation observed in the radius of basal rosettes from different populations. Interestingly, with the exception of individuals from Swansea, plants from populations identified as producing large quantities of leaves (Turk & Alg2) had a lower than average rosette radius. Conversely, individuals from Swansea, Wales invested heavily in leaf production, both in terms of leaf number and size. This can be seen clearly from the greenhouse data displayed in figure 4.2a. Plants from Swansea had both high leaf number and large rosette size. In comparison, individuals from southern populations (i.e. those from lower latitudes) that had high leaf numbers typically had smaller rosette radii and those with larger radii typically had fewer leaves. This indicates that at the given point in time when measurements were taken, (62 days) plants from Swansea had produced the most vegetative growth.

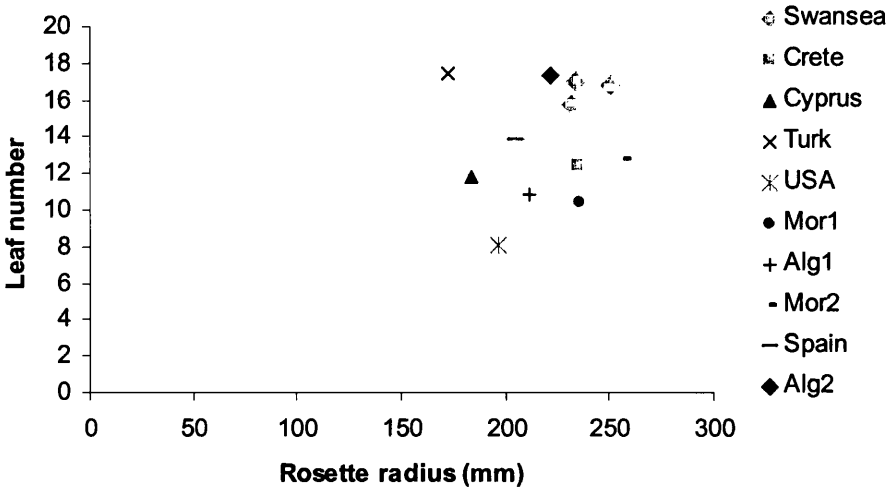


Figure 4.2 a Plot of rosette radius against leaf number for plants grown in the greenhouse. The figure indicates the relative resources invested in leaf production for the different populations investigated. The three populations from Swansea area produced among the highest leaf number and rosette radius. All measurements were taken when the first plant was observed to bolt at day 62.

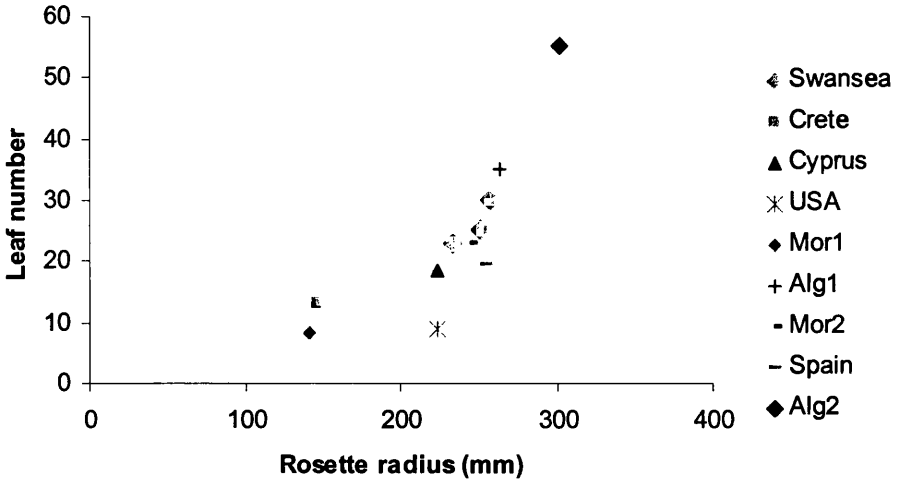


Figure 4.2 b Plot of rosette radius against leaf number. Measurements were taken as each plant was observed to bolt (plants grown outdoors in Swansea University Botanic Garden). This figure indicates the investment in vegetative growth at the developmental stage of flower production. In this instance, although plants from Swansea have both a high leaf number and large rosette radius, the greatest production of vegetation was from the two populations from Algeria.

Figure 4.2b indicates the relative investment in vegetative production of *H. incana* plants at the developmental stage of flower production, when grown outdoors in Swansea. These results are not directly comparable to the measurements taken from the greenhouse as they represent resource allocation to vegetative growth at a developmental stage in contrast to measurements taken in the greenhouse, which give an indication of vegetative production after a certain period of time (62 days). That said, although not the highest vegetative producer (both populations from Algeria produced the greatest quantity), plants from Swansea were observed to produce both a greater number of basal leaves and a larger basal rosette than many of the southern populations (Figure 4.2 a, b, table 7.3, 7.4).

4.3.4 ***Stem length.*** The majority of plants across the populations were seen to have a stem length of between 400-800 mm when they first came into flower. Two populations had shorter stems relative to the other populations (Taza, Morocco and Saida, Algeria) Mann-Whitney U 168.0, P=<0.001. There were no associations with leaf size or number, or relationships with altitude that may have provided an explanation for this peculiarity (the data was examined to check if there were any associations between; altitude and stem length, latitude and stem length, leaf number and stem length, rosette size and stem length, flowering time and stem length. These comparisons were performed by plotting scatter graphs and calculating Pearson correlations. As there were no relationships, the Pearson coefficients are not shown, the raw data are taken from Tables 4.1, 8.4, 8.5 and 8.6). When these plants were grown outdoors a more uniform stem length was observed on flowering. In this case, only Taza (Morocco) had noticeably shorter stems on flowering than the other populations (Mann-Whitney U 108.5, $P = 0.001$).

4.3.5 ***Variation in Flowering time.*** Time from germination to first open flower varied greatly between populations and in some cases within populations (Tables 4.5, 8.5, 8.6). Flowering was first observed in an individual from the L.A. population after 55 days and took in excess of 160 days in some individuals from Swansea. The range of flowering time within populations varied considerably between populations. Individuals from Mohammedia, Morocco, all flowered over a period of two weeks. In contrast, individuals from Bejar, Spain and Swansea flowered over an eight and 10 week period, respectively. Swansea and Spain were the only populations where some individuals failed to produce flowers in the season's growth in the greenhouse 12 from 50 and 1 from 16 respectively.

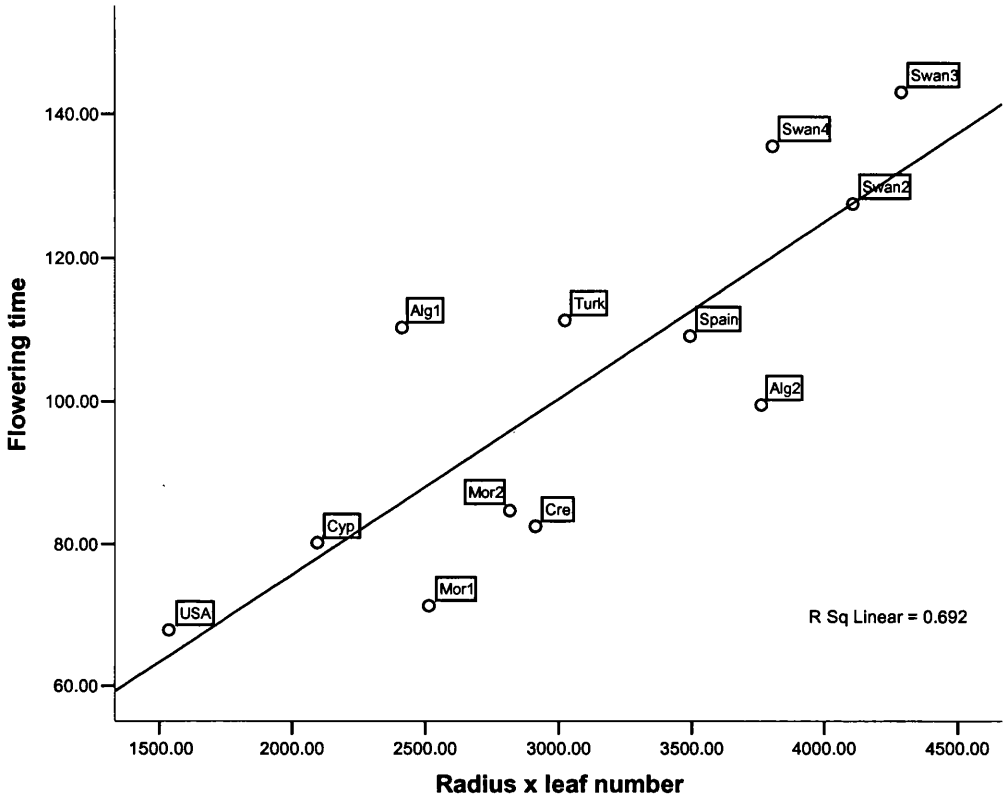


Figure 4.3a Relationship between leaf production and flowering time, for individuals grown in the greenhouse. A linear relationship was observed between increasing flowering time (days) and the estimator of leaf area (rosette radius x leaf number).

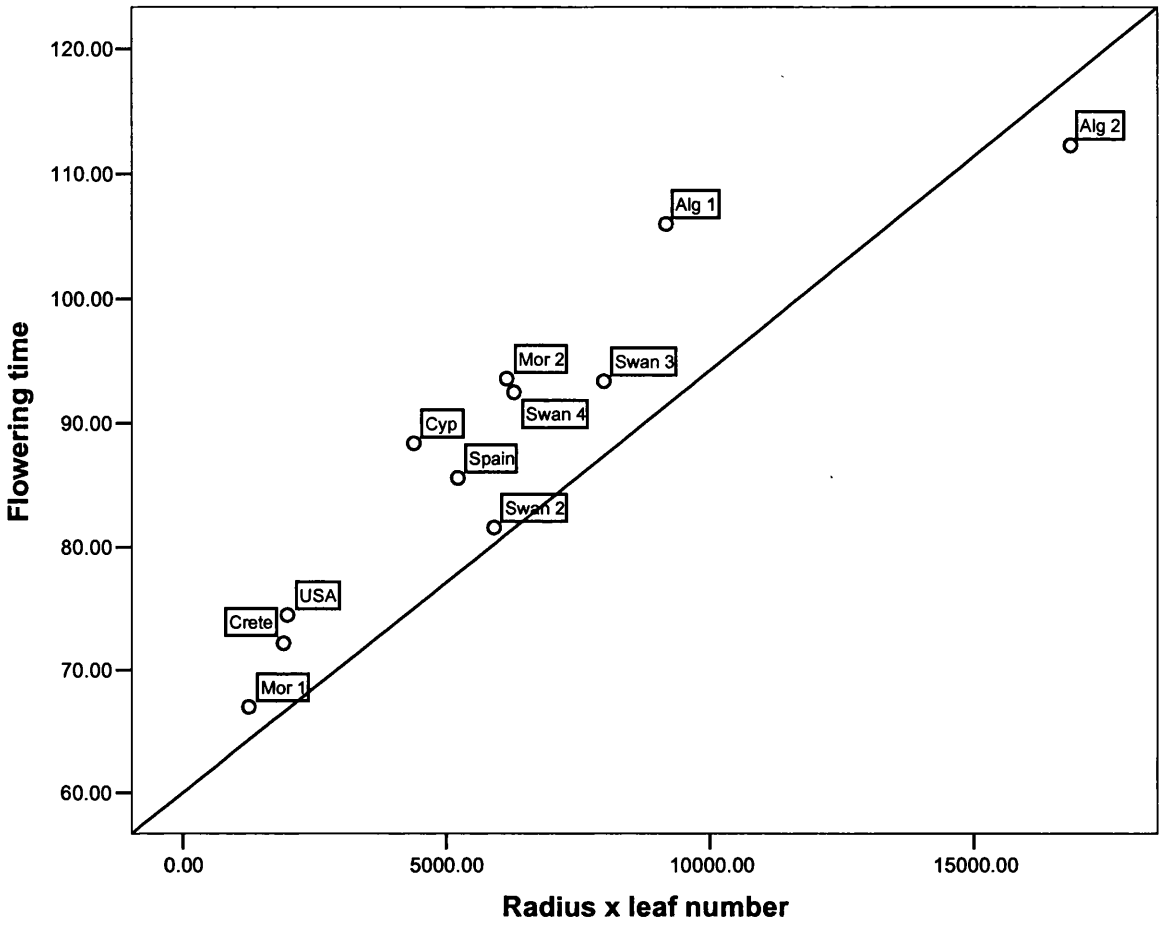


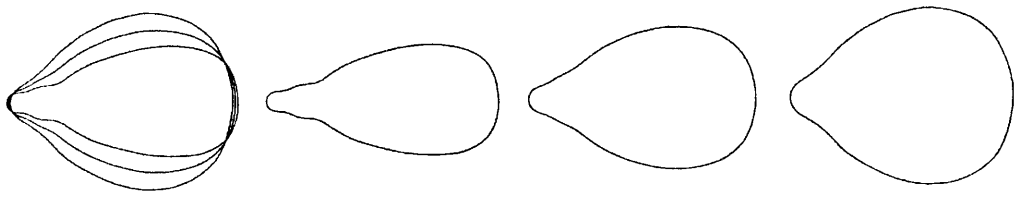
Figure 4.3b Relationship between leaf production and flowering time, for individuals grown outdoors (b). A linear relationship was observed between increasing flowering time (days) and the estimator of leaf area (rosette radius x leaf number).

Time taken to flower from germination was found to correlate with leaf production, with plants grown in the greenhouse, as measured at this discrete time of 62 days (Pearson correlation = 0.832, P = 0.00079, n = 12) indicating a possible difference in resource allocation between populations. This result was also observed in those plants grown outdoors (Pearson correlation = 0.768, $P < 0.001$, n = 79) although the results are not directly comparable as in this case leaf dimensions (rosette radius) and number of leaves were measured at the time of flowering. This is as you would expect, as plants flowering later have had a greater time to produce more leaves through increased acquisition of

resources (sunlight and nutrients). The greatest difference between the two sets of results is influenced by the variation in flowering time of some populations. A lack of physical constraint also lead to greater leaf growth, both in terms of leaf size and leaf number. In the greenhouse, individuals from Swansea flowered latest and produced the greatest quantity of vegetation (Figure 4.3a). In conditions outdoors, individuals from the Algerian populations produced the greatest quantity of vegetation and flowered last (Figure 4.3b).

4.3.6 ***Shape analysis of H. incana leaves and petals.*** Patterns of distinctive leaf shape between populations were not found using the Shape analysis package. When principle components were plotted, there were no distinct clusters based on population identity (data not shown). The graphical representation to describe variation in leaf shape was based on width versus length and did not appear to be able to incorporate variation in lobes. It is possible that the variable location of the leaves lateral lobes was the cause of the lack of differentiation observed by the program. A similar result was observed with the petals in that there were no distinct clusters formed between contrasting populations. However, the graphical output visualised the differences in petal shapes more realistically (figure 4.4). The main differences in shape would appear to be described by the overall roundness of the petal (PC1) and the difference of angles in tapering towards the ends of the petals (PC2). It is possible that a greater sample size may have given greater statistical power to differentiate between populations, but since this was not available, the technique was abandoned at this stage.

PC1



PC2

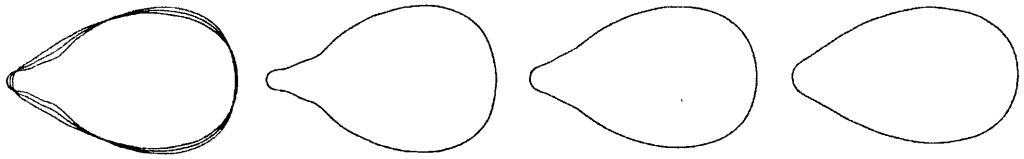


Figure 4.4 Graphical representation of principle components of petal variation drawn by Shape. The two images on the far left display the variation, the third from left images indicate the mean petal shape and the image on either side of these represent 2 standard deviations difference.

4.3.7 ***Responses to Environment.*** There were no correlations between altitude or latitude with any of the measured parameters other than between flowering time (within the greenhouse), and latitude (Pearson correlation = 0.821, $P = 0.004$, n = 10).

A correlation of flowering time was observed with annual rainfall if Taza was excluded from the dataset (Pearson $P = 0.009$), but due to the abnormally high annual rainfall this was not significant when Taza was included (the data point acted as an outlier affecting the overall results). Flowering time was observed to correlate with length of growing season (based on the availability of the arbitrary selected > 20 mm monthly precipitation), Pearson $P = 0.011$. Basal leaf number was also correlated with length of growing season, Pearson $P = 0.014$ (based on >20 mm monthly precipitation). Twenty millimeters of monthly precipitation was selected as a known minimum quantity of precipitation that can support *H. incana*. This assumption was made as 20 mm is the lowest (available) recorded monthly precipitation from an area where populations of *H.*

incana are found (L.A. airport) (Table 7.2). In actual fact this may be an overestimation, as for the majority of the year monthly precipitation is lower than this figure.

Table 4.8 Differences observed in flowering time in days under greenhouse and outdoor growth conditions in Swansea. Shaded boxes indicate significant differences in flowering time between the two growth conditions, based on t-tests with sequential Bonferroni corrections.

Population	Greenhouse ft	Outdoor ft	Difference	P-values
Swan3	143	93.4	-49.6	$9.6x10^{-7}$
Swan2	127.5	81.6	-45.9	$1.6x10^{-5}$
Swan4	135.5	92.5	-43	$2.5x10^{-4}$
Spain	109.1	85.6	-23.5	$1.0x10^{-3}$
Cre	82.5	72.2	-10.3	0.03
Alg1	109	106.3	-2.7	0.06
Mor1	69.7	67	-2.7	0.177
USA	67.9	74.5	6.6	0.195
Cyp	80.2	88.4	8.2	0.301
Mor2	84.3	93.6	9.3	0.466
Alg2	99.5	112.3	12.8	0.756

One significant difference observed in flowering time between *H. incana* grown in the greenhouse and outdoors was the reduced time taken for individuals from Swansea and Bejar to flower when they were grown outdoors (table 4.8). These were the only two populations to display significant differences in the time taken from germination until flowering between the two treatments. It is possible that the availability of more resources allowed more rapid growth, and that a threshold size was reached that allowed flowering to commence in a shorter period of time. Greater resources available may include light (longer daylight hours) and nutrients (unrestricted root growth compared to the constraints of a plant pot). The differences may also be linked to the timing of the planting. The start of *H. incana*'s main growing season in the UK would be around April. Wilcoxon signed ranks test indicated no significant directional difference between

flowering times from plants grown in the greenhouse or outdoors. Z = -0.415, $P = 0.678$ (based on positive ranks). Figure 4.5 displays the time taken to flower for each population when they were grown under three different conditions in Swansea. Although the time taken from germination until flowering varied between populations and growth conditions, individuals from Swansea were among the last to flower in all cases.

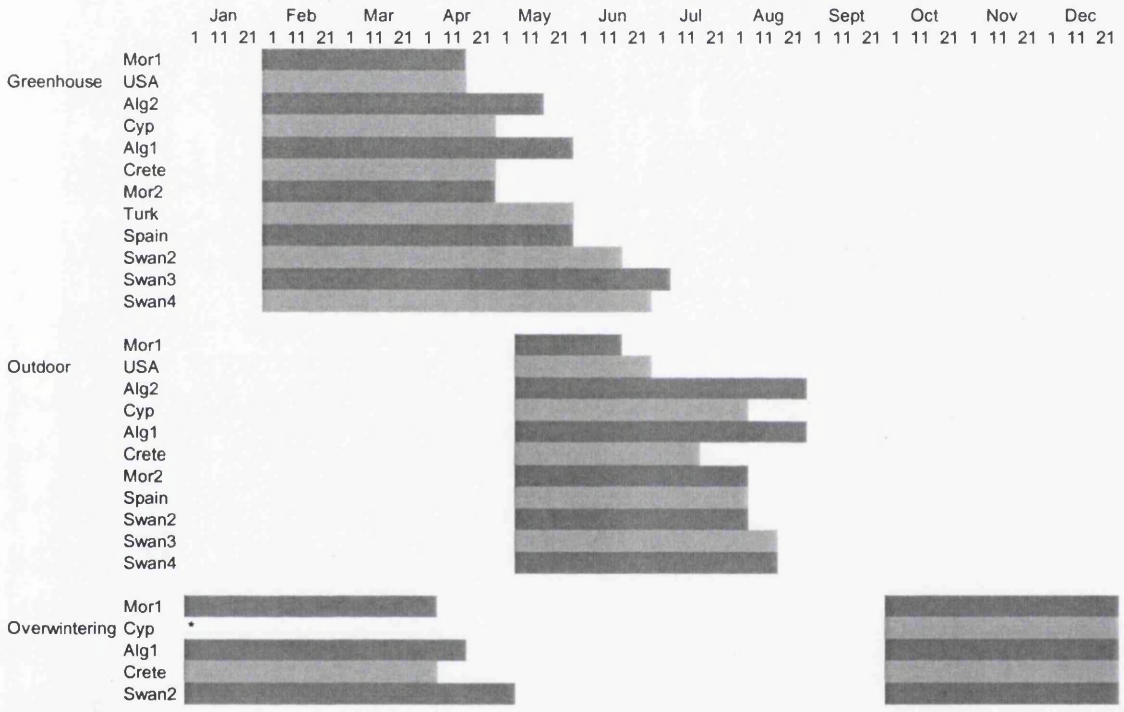


Figure 4.5 Mean time taken between germination and flowering for *H. incana* from a variety of populations germinated at three distinct times. Plants germinated at the beginning of February were grown under greenhouse conditions. The other two groups were germinated in the greenhouse, but transferred outdoors into the University Botanic Garden. * No overwintering survival was observed for any plants from Larnaka, Cyprus.

4.3.7.1 Frost tolerance / survival during winter (Swansea). All plants from Swansea, Crete, Morocco and Algeria survived throughout the winter and proceeded to flower in the spring (Figure 4.4). All plants from Cyprus failed to survive the winter, indicating either a possible lack of frost/cold tolerance. After flowering all individuals from Crete, Morocco and Algeria produced seed and died soon afterwards. All individuals from Swansea survived for the duration of the year and the following winter. These plants successfully flowered during the following spring. Only plants from Swansea displayed the ability to perennate. Individuals from Swansea were typically later to flower than those from southern latitudes, with a difference of approximately 20 days between mean flowering times.

Table 4.9 Survival of *H. incana* through winter in Swansea, and associated flowering strategy (plants germinated 25/09/05). The proportion of plants that survived the winter period and the proportion of plants that persisted after flowering are reported. Only plants from Swansea populations persisted after flowering and they survived until the following season when flowering was observed to commence again. All individuals from Cyprus perished during the first winter. All individuals from Crete, Morocco and Algeria survived the winter, flowered and died within a month after seed has been set.

Population	Winter survival	Mean flowering time (days)	Survival after flowering	Perenniation
Swansea	9/10	02/05/2006 (122)	9/9	9/9
Crete	10/10	08/04/2006 (98)	0/10	0/10
Cyprus	0/10			0/10
Morocco	10/10	10/04/2006 (100)	0/10	0/10
Algeria	10/10	11/04/2006 (101)	0/10	0/10

4.3.7.2. Vernalization. All plants from Crete in the non-vernalized group flowered within 56 and 68 days from germination (Figure 4.5). All plants from Crete in the vernalized group flowered within 72 and 76 days. This time period included the cold treatment of 35 days at 4°C, during which little growth was observed due to the cold temperature. Vernalized plants from Swansea began to flower after 85 days, again this time includes the same number of days of cold treatment. Twelve from twenty plants from the vernalized group of Swansea individuals flowered within the experimental period. Plants from Swansea in the non-vernalized group did not flower within the experimental period of 150 days and as a result are not present in Figure 4.5. If the growth chamber had been available for a greater period of time the experimental period could have been extended. It would have also been interesting to investigate a greater diversity of populations, and in a different experiment investigate if day length also had an effect on flowering times.

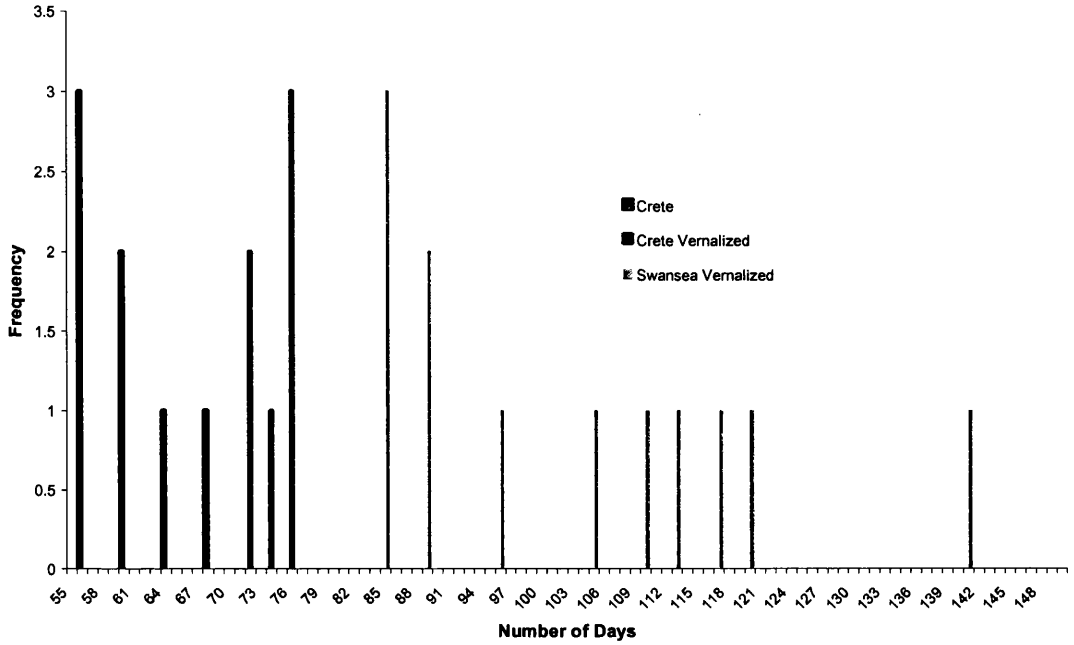


Figure 4.6 Profile of observed flowering time in days from germination for vernalized (35 days of cold treatment at 4 °C) and non-vernalized individuals of *H. incana*. Non-vernalized individuals from Swansea did not flower within the 150 days of the experimental period, and therefore are not included in the plot. The number of days taken to flower includes in this chart includes the period of cold treatment – where growth for these individuals during this period was minimal.

4.3.7.3 Life Strategy. (Indication of annual or perennial growth strategy see table 4.9 including winter survival). During all plant growth experiments, individuals from Crete and Cyprus were observed to die after flowering, and always behaved as annuals. Individuals from Morocco and Algeria indicated an intermediate strategy, the majority of these plants died following flowering. For example, all of the Moroccan and Algerian individuals that were planted in September and flowered in spring died soon after flowering (Table 4.9), but some of the individuals that had been planted outdoors in April survived and flowered the following spring (Pers. obs, data not shown). All individuals

from Swansea and Spain were observed to flower, survive the winter and flower again in the following spring. Some individuals from Swansea were observed to survive for three years, and come into flower during the growing season each year (Pers. obs).

4.4 Discussion

4.4.1 ***Resource allocation.*** The life history of an organism can be described by the combination of resource allocations to maintenance, growth and reproduction (Ehlers & Olesen 2004). An increase in investment in one area of resource allocation may be at the expense of another. This assumption implies that resource allocation is optimised in a way that maximises fitness (Cody 1966; Harper 1967). Benefits resulting from investment of resources to specific functions should outweigh lost competitive ability in those functions where investment is low. Patterns of resource allocation are thought to evolve in a way that maximises reproduction and survival; these can be influenced by external environmental pressures (Schaffer 1974).

H. incana has expanded its range to such an extent that it is now distributed across a wide latitudinal gradient in regions with contrasting environmental conditions (Jalas & Suominen 1996, Preston *et al.* 2002) – for example, in one part of its range it is subject to Mediterranean climates, and in another part of its range, temperate climates instead. The different parts of this range differ considerably in climatic conditions, including timing of the growing season, mean temperatures, patterns of rainfall and day length. Traits specific to native regions (most southern European populations experience a Mediterranean type climate) will not be optimally adapted for northern temperate climates. It is therefore possible that the phenotypic differences observed between these populations, in particular those observed between UK (alien) populations and continental European populations (considered native), may be a result of new adaptations to the novel habitat (northern temperate climate).

Principle component analysis was employed to graphically visualise variations between measured morphometric traits in order to give some indication of similarities and differences between populations. The results of the PCA (Table 4.4) indicated that plants that had germinated earlier gained vegetation more rapidly – as can be reasonably expected. In addition, the PCA suggested that plants that were quicker to gain vegetation were in fact slower to produce flowers.

Differences observed in vegetative growth and duration before flowering may reflect adaptations to local environmental conditions. During the initial investigation when populations of *H. incana* were grown in the greenhouse, individuals from Swansea were observed to produce among the highest quantity of vegetative growth in the given time period. Conversely, individuals from L.A. (USA) (also a non-native population, but in a region with a climate similar to the Mediterranean) were among the lowest producers of vegetation within this time period (Table 8.5). *H. incana* is a successful invader in both California and South Wales, its success in both locations may be explained by the different growth patterns it utilizes in each environment to maximize its fitness. Comparing time taken from germination until onset of flowering between these two populations, two clearly different strategies were apparent under these growth conditions. Individuals from L.A., USA invested less resources into vegetative growth but developed into the flowering stage rapidly (mean basal leaf number at 62 days = 7.8, mean basal rosette radius at 62 days = 196 mm, mean time take for onset of flowering = 67.9 days) (Table 8.5). The strategy adopted by individuals from Swansea (for the sake of this example using data from Swan2) under these same greenhouse conditions were a greater investment in vegetative growth and a delay in flowering time (mean basal leaf number at 62 days = 17.6, mean basal rosette radius at 62 days = 233 mm, mean time taken for onset of flowering = 127.5 days) (Table 8.5). Other populations observed to flower in significantly less time than all groups from Swansea included Mohammedia and Taza (Morocco), Larnaka (Cyprus), Elounda (Crete) and Saida (Algeria). In each of these cases with the exception of Saida (Algeria) the variable 'leaf number x basal rosette diameter' was lower than the overall mean, providing an indication of less than average investment in vegetative growth.

Multiple pairwise comparisons between populations (Table 4.5) test revealed significant differences in terms of flowering time and basal leaf number. Clusters of populations could be identified on the basis of presence or absence of significant differences. Clusters based on differences in flowering time separated into three groups. L.A. (USA) and Mohammedia (Morocco) were grouped together as the two fastest flowering populations. These two sites represent the most arid locations where plants originated in this growth

experiment, including longer dry seasons and lower annual rainfall. The other extreme of slower initiation to flower included Swansea (Wales), Bejar (Spain) and El Bier (Algeria). Climate data for these localities indicates that rainfall is more continuous throughout the year, and where at least 10 months are subject to > 20 mm of precipitation. This contrasts with LA where > 20 mm precipitation only occurs for 3 months. An intermediate cluster including Elounda (Crete), Larnaka (Cyprus), and Taza (Morocco) was apparent. Climate data for these populations indicated seasonal rainfall of between 6-10 months, however, Elounda and Larnaka had low annual mean precipitation while Taza was subject to high levels of rainfall. A correlation of flowering time was observed with annual rainfall if Taza was excluded from the dataset as an outlier (Pearson $P = 0.009$). Length of growing season (based on the availability of the arbitrary selected > 20mm monthly precipitation was observed to correlate with flowering time, (Pearson $P = 0.011$), including all populations. This strongly suggests that for *H. incana* that a major influence in the differences observed between populations phenotypic characters is based on their locations water availability throughout the year. Combining this observation with the data generated by the PCA (Figure 4.1) we could predict the likelihood of *H. incana*'s survival in a given environment and the likely growth strategy that would be required for successful further propagation.

The pairwise comparison for number of basal leaves observed from individuals between populations produced 2-3 main groups, but greater variation resulted in an overlap in clusters and a lack of ability to differentiate clearly between groups. In comparison to the correlation of flowering time with water availability, no relationship was observed between production of leaves and water availability.

Some plants are known to respond to drought conditions by an early onset of flowering (Aronson *et al.* 1992). For example, evolutionary change has been observed for flowering time in *Brassica napus* over short time periods as a response to changing water availability (less than 10 years). Populations subjected to abbreviated growing seasons caused by drought flowered earlier than their ancestors that had been exposed to longer growing seasons. Flowering time in this case was found to be a heritable trait (Franks *et al.* 2007). In contrast to this pattern, plastic responses to water availability have been

observed in some *Brassica* annuals. *Diplotaxis horra*, a common annual species in the desert of Egypt, was observed to be able to switch from an annual to perennial form in wet years with abundant water. A shift was observed from an 'r' to a 'k' selected lifecycle, where perennial plants produced seeds of a higher energy content, increased germinability, viability and longevity (Hegazy 2001). Plasticity may assist evolutionary divergence in certain circumstances. It may improve the chances of survival of a species in a novel habitat, allowing the pressures of other forces to promote genetic differentiation in the gene pool (Agrawal 2001).

Flowering time can be accelerated by a number of stresses in addition to drought conditions. These may include overcrowding (a change in light quality), nutrient deficiency and heat (Simpson & Dean 2002). Additionally, changes in day length are known to regulate the onset of flowering (Koornneef *et al.* 1991, Suarez-Lopez *et al.* 2001, Simpson & Dean 2002).

In the case of the populations grown outdoors in Swansea, different patterns of flowering time and vegetation production were observed between the populations when they were grown in the greenhouse. Individuals from L.A. (USA), Mohammedia (Morocco), Elounda (Crete), and Larnaka (Cyprus) all flowered rapidly with little investment in vegetative growth, as was observed under greenhouse conditions. Individuals from Swansea were also observed to flower in less time outdoors than under greenhouse conditions, albeit with a greater production of vegetative growth compared to the individuals from the quickest to flower populations. Under outdoor growth conditions, individuals from Algerian populations were the slowest to commence flowering and produced the greatest quantity of vegetation at the time of flowering. Perhaps it is not surprising that plants from Swansea took less time to flower under outdoor growth conditions compared to greenhouse conditions. The time of the outdoor experiment corresponded with the main flowering season for the local *H. incana* of this area, compared to the greenhouse experiment which was prior to the local flowering season. In addition, a combination of longer days, increased light and a lack of root constraint from plant pots may have allowed for a greater and more rapid assimilation of vegetative growth, and rapid commencement of flowering when plants had attained a suitable size.

The "home site advantage" hypothesis predicts that fitness or relative success may decrease with increased environmental distance to the local population (Montalvo & Ellstrand 2000). However, in this study, seed output was not recorded as a measure of fitness and there is not enough data to estimate "environmental distances" to be able to test the home site hypothesis. Certainly, local populations demonstrated rapid growth rates (high basal leaf numbers x leaf size) and rapid maturity (commencement of flowering) under local conditions.

When grown under greenhouse conditions, a trend was observed that duration prior to flowering increased with increasing latitude. This difference in flowering time may correspond with a gradient of local environmental conditions across the latitudinal range of these populations. The growing season for *H. incana* may be controlled by factors such as temperature and availability of water. Within Mediterranean areas (e.g. Crete) water availability is greatest between autumn and spring. Droughts combined with high temperatures are usual weather conditions during the summer (table 8.4). Winter temperatures in the Mediterranean are generally not low enough to be limiting to growth. Under these conditions, *H. incana* acts as a winter annual. If survival is unlikely through periods of drought through these predictable Mediterranean summers, a strategy of rapid flowering and seed maturation prior to the dry season would appear logical. Towards the northerly limit of the species range (in my samples, individuals from Swansea, although the species is found further north) winter temperatures are not favorable for growth, but certainly in Swansea average temperatures do not drop so low as to limit survival. Generally, water is more available throughout the year and in this case temperature and light availability may be the main factors regulating the timing of the growing season. In this environment, the species acts as a summer perennial. Survival is not generally restricted by periods of drought and a strategy of delayed flowering and allocating more resources into vegetative growth is apparent. This approach produces larger plants that are able to survive for successive seasons, therefore increasing potential total seed production for that individual.

4.4.2 ***Environmental cues for flowering.*** Vernalization treatment resulted in observable differences in the flowering response between populations of *H. incana* from Crete and Swansea. Vernalisation had little effect on the flowering response of plants from Crete that were observed to flower rapidly under both vernalized and non-vernalized conditions (within 50-70 days - not including the 30 days of cold treatment). A proportion of plants from Swansea were observed to respond to vernalization treatment, resulting in 12/20 plants flowering in the growth cabinet conditions. Those plants from the same Swansea populations grown under constant conditions with no cold treatment were not observed to flower within the time period of the growth experiment (150 days). This suggests that plants from Swansea require an environmental stimulus to commence flowering, although since not all vernalized plants were observed to flower, it is likely that other factors in addition to cold treatment are involved in stimulating plants from Swansea to initiate flowering. It is possible that the remaining plants from Swansea would have flowered if the experiment was continued for a longer period of time. However, the experiment did indicate that vernalisation did prompt flowering in a proportion of the Swansea plants. It also served to show the quite remarkable difference in time taken to flower between plants from Crete and Swansea.

Individuals from Swansea grown in the greenhouse were not subject to cold treatment, but the majority of plants did commence flowering after an extended period of time (mean of 138 days) and individuals grown outdoors were observed to flower in less time than this (mean of 88 days). Therefore, individuals from Swansea do not have an obligatory requirement for vernalization, but do appear to have a facultative response to cold treatment. Additional factors that have not yet been investigated for this species such as changes in day length and differences in night and day temperatures may play an important role in regulating flowering time in these plants. This has been reported for *Arabidopsis lyrata* from northerly latitudes, where changes in day length were identified as controlling factors for flowering time (Riihimaki *et al.* 2005). In this example, plants from northerly populations could take more than 150 days to flower in short day conditions. This was in contrast to individuals from populations from lower latitudes that flowered rapidly regardless of day length (less than 100 days).

Considerable research effort has been placed on understanding the mechanisms that control the timing of flowering (a key adaptive component of the life cycle of a plant) particularly within the model plant *Arabidopsis thaliana*. Environmental stimuli are major determinants of the onset of flowering, which ensures that flowering coincides with the optimal season for the species. Key stimuli for flowering time in *Brassica* species include photoperiod and vernalization controls. The genetic mechanisms have been described in detail in the literature for *A. thaliana* (e.g. Lagercrantz *et al.* 1996, Lagercrantz 1998, Nordborg & Bergelson 1999, Johanson *et al.* 2000, Simpson *et al.* 2002, Caicedo *et al.* 2004, Komeda 2004, Michaels *et al.* 2005, Shindo *et al.* 2005) and comparisons have been made to *Brassica* species (Lagercrantz *et al.* 1996, Osborn *et al.* 1997, Lagercrantz 1998, Osterberg *et al.* 2002). Key observations include photoperiod as the main environmental stimuli in rapid cycling accessions of *A. thaliana* while winter annual *A. thaliana* has both a photoperiod and vernalization requirement for rapid flowering (Michaels *et al.* 2005). As has been done for some *Brassica* species, it may be possible to transfer some of this knowledge to *H. incana* in future studies to help with understanding the observed variation in flowering time with these stimuli in addition to climatic differences.

4.4.3 ***Life history characteristics.*** All populations of *H. incana* that were investigated for survival over a British winter (or frost/snow exposure) were observed to have high survival rates (table 4.9), with the exception of those plants from Larnaka (Cyprus), of which all individuals failed to survive. These results indicate the ability of certain populations of this species to withstand temperatures down to 0 °C, at least for short periods of time. The difference in survival between individuals from Larnaka, Cyprus and all others investigated may relate to local climatic differences experienced by these populations, although the climatic data in table 8.4 does not indicate any obvious patterns. Frost or winter survival has been described by Stone *et al.* 1993 as being a complex trait with polygenic inheritance. It is therefore possible that different populations of *H. incana* are genetically different with regard to this trait – giving rise to the survival to individuals of some, but not all populations. Differences in life history were also

observed among the surviving populations. All individuals from Elounda (Crete), Mohammedia (Morocco) and El Bier (Algeria) flowered in less time than those plants from Swansea and they displayed an annual life form strategy; no survival was observed after flowering and seed set, during this discrete investigation. In contrast, all individuals from Swansea survived and were observed to flower in the following growing season the year after. Similar variation in perennation habit and flowering time for *Mimulus guttatus* has been reported at the population level by Hall *et al.* 2006. Their observations of less vegetative production and quicker arrival to flowering for annual individuals (compared to their perennial counterparts) are consistent with the results observed in this study. *A. lyrata* is reported to show a similar trend of delayed flowering in northern populations compared to those located further south (within the northern hemisphere) (Riihimaki & Salvolainen 2004, Riihimaki *et al.*2005). This difference in flowering time has been attributed to genetic differences between the populations where regulatory control over flowering is influenced by daylength (Kuittinen *et al.*2008). Observations of *H. incana* flowering time under different daylength regimes were beyond the scope of this study. If these were performed in the future, it would be interesting to compare such results to studies such as those by Riihimaki *et al.* (2005) and Kuittinen *et al.* (2008) as it may provide a greater understanding as to the mechanisms influencing flowering time variation in *H. incana.*

The most important, ecologically important variations in phenotypic traits of *H. incana,* (including flowering time, vegetative production and annual or perennial strategies) observed in these growth experiments would appear to be linked to the climatic gradient across the distribution of the species range. Further investigations could focus on the heritability of these traits as this may give an insight into timescales of evolutionary processes in invasive species such as *H. incana.* Additional investigations into the mechanisms of self incompatability for *H. incana* would support the current literature by Darmency and Fleury 2000. Unfortunately the attempts at self fertilization and between population crosses did not yield any results due to an outbreak of pests.

It is possible that the effects of climate change may give rise to opportunities for rapidly evolving species to colonise new areas. In northern temperate climates, temperatures are predicted to increase in combination with an increase in precipitation (e.g. Dukes & Mooney 1999). These conditions may promote evolutionary adaptation in species such as *H. incana* including life history changes from annual to perennial strategies. In turn, this may result in higher propagule pressure from increased fecundity – a result of survival and reproduction for multiple years and the growth of potentially larger plants. This could result in the increased invasibility of opportunistic aliens that may pose a threat to native biodiversity.

4.5 References

Agrawal AA (2001) Phenotypic plasticity in the interactions and evolutions of species. Science, 294, 321-326.

Aronson J, Kigel J, Shmida A, Klein J (1992) Adaptive phenology of desert and Mediterranean populations of annual plants grown with and without water stress. Oecologia, 89, 17-26.

Baker HG (1965) Characteristics and modes of origin of weeds. In: Baker HG, Stebbins GL (eds) The genetics of colonizing species. Academic Press, New York, pp 147-168.

Bewley DJ (1997) Seed germination and dormancy. The Plant Cell, 9, 1055-1066.

Caicedo AL, Stinchcombe JR, Olsen KM, Schmitt J, Purugganan MD (2004) Epistatic interaction between Arabidopsis FRI and FLC flowering time genes generates a latitudinal cline in a life history trait. Proceedings of the National Academy of Sciences of the United States of America, 101, 15670-15675.

Callahan H (2005) Using artificial selection to understand plastic plant phenotypes. Integrative and Comparative Biology, 45, 475-485.

Cody ML (1966) A general theory of clutch size. Evolution, 20, 174-184.

Darmency H, Fleury A (2000) Mating system in Hirschfeldia incana and hybridization to oilseed rape. Weed Research, 40, 231-238.

Dukes JS, Mooney HA (1999) Does global change increase the success of biological invaders? TREE, 14, 135-139.

Ehlers BK, Olesen JM (2004) Flower production in relation to individual plant age and leaf production among different patches of Corydalis intermedia. Plant Ecology, 174, 71-78.

Evenari M (1949) Germination inhibitors. The Botanical Review, 15, 153-194.

Frankham R (2005) Invasion biology - Resolving the genetic paradox in invasive species. Heredity, 94, 385-385.

Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. Proceedings of the National Academy of Sciences of the United States of America, 104, 1278-1282.

Hall MC, Willis JH (2006) Divergent selection on flowering time contributes to local adaptation in Mimulus guttatus populations. Evolution, 60, 2466-2477.

Harper JL, (1967) A Darwinian approach to plant ecology. Journal of Ecology, 58, 681-698.

Hegazy AK (2001) Reproductive diversity and survival of the potential annual Diplotaxis harra (Forssk.) Boiss (Brassicaceae) in Egypt. Ecography, 24, 403-412.

Hurka H, Bleeker W, Neuffer B (2003) Evolutionary process associated with biological invasions in the Brassicaceae, Biological Invasions, 5, 281–292.

Iwata H, Ukai Y (2002) SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. Journal of Heredity, 93, 384-385.

Jalas J, Suominen J (1996) Atlas Florae Europaeae, Distribution of vascular plants in Europe 11 Cruciferae, Ricotia to Raphanus. The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. Science, 290, 344-347.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. Ecology Letters, 7, 1225-1241.

Kuittinen H, Niittyvuopio A, Rinne P, Savolainen O (2008) Natural variation in *Arabidopsis lyrata* vernalization requirement conferred by a FRIGIDA indel polymorphism. Molecular Biology and Evolution, 25, 319-329.

Kolbe JJ, Glor RE, Schettino LRG, et al. (2004) Genetic variation increases during biological invasion by a Cuban lizard. Nature, 431, 177-181.

Kollmann J, Banuelos MJ (2004) Latitudinal trends in growth and phenology of the invasive alien plant Impatiens glandulifera (Balsaminaceae). Diversity and Distributions, 10, 377-385.

Komeda Y (2004) Genetic regulation of time to flower in Arabidopsis thaliana. Annual Review of Plant Biology, 55, 521-535.

Koornneef M, Hanhart CJ, Vanderveen JH (1991) A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. Molecular & General Genetics, 229, 57-66.

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in *Arabidopsis thaliana*. Annual Review of Plant Biology, 55, 141-172.

Kuhl FP, Giardina CR (1982) Elliptic Fourier features of a closed contour. Computer Graphics and Image Processing, 18, 236–258.

Lacaze X, Hayes, PM, Korol A (2009) Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. Heredity, 102, 163-173.

Lagercrantz U, Putterill J, Coupland G, Lydiate D (1996) Comparative mapping in Arabidopsis and Brassica, fine scale genome collinearity and congruence of genes controlling flowering time. Plant Journal, 9, 13-20.

Lagercrantz U (1998) Comparative mapping between Arabidopsis thaliana and Brassica nigra indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics, 150, 1217-1228.

Le Minor JM, Schmittbuhl M (1999) Importance of elliptic Fourier methods for morphometry of complex outlines: application to the distal human femur. Surgical and Radiologic Anatomy, 21, 387-391.

Linde M, Diel S, Neuffer B (2001) Flowering ecotypes of Capsella bursa-pastoris (L.) Medik. (Brassicaceae) analysed by a cosegregation of phenotypic characters (QTL) and molecular markers. Annals of Botany, 87, 91-99.

Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. Trends in Ecology and Evolution, 20, 223-228.

Maron JL, Vila M, Bommarco R, Elmendorf S, Beardsley P (2004) Rapid evolution of an invasive plant. Ecological Monographs, 74, 261-280.

Michaels SD, Himelblau E, Kim SY, Schomburg FM, Amasino RM (2005) Integration of flowering signals in winter-annual Arabidopsis. Plant Physiology, 137, 149-156.

Montalvo AM, Ellstrand NC (2000) Transplantation of the subshrub Lotus scoparius: Testing the home-site advantage hypothesis. Conservation Biology, 14, 1034-1045.

Neuffer B, Hurka H (1988) Germination behavior in populations of Capsella bursa-pastoris (Cruciferae). Plant Systematics and Evolution, 161, 35-47.

Neuffer B, Hurka H (1999) Colonization history and introduction dynamics of Capsella bursa-pastoris (Brassicaceae) in North America: isozymes and quantitative traits. Molecular Ecology, 8, 1667-1681.

Nordborg M, Bergelson J (1999) The effect of seed and rosette cold treatment on germination and flowering time in some Arabidopsis thaliana (Brassicaceae) ecotypes. American Journal of Botany, 86, 470-475.

Osborn TC, Kole C, Parkin IAP, Sharpe AG, Kuiper M, Lydiate DJ, Trick M (1997) Comparison of flowering time genes in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*. Genetics, 146, 1123-1129.

Osterberg MK, Shavorskaya O, Lascoux M, Lagercrantz U (2002) Naturally occurring indel variation in the Brassica nigra COL1 gene is associated with variation in flowering time. Genetics, 161, 299-306.

Parker IM, Rodriguez J, Loik ME (2003) An evolutionary approach to understanding the biology of invasions: Local adaptation and general-purpose genotypes in the weed Verbascum thapsus. Conservation Biology, 17, 59-72.

Pigliucci M, Marlow ET (2001) Differentiation for flowering time and phenotypic integration in Arabidopsis thaliana in response to season length and vernalization. Oecologia, 127, 501-508.

Preston CD, Pearman DA, Dines TD (2002) New atlas of the British & Irish flora : an atlas of the vascular plants of Britain, Ireland, the Isle of Man and the Channel Islands. Botanical Society of the British Isles, Oxford University Press.

Reinartz JA (1984a) Life-History Variation of common mullein (Verbascum thapsus) .1. Latitudinal differences in population dynamics and timing of reproduction. Journal of Ecology, 72, 897-912.

Reinartz JA (1984b) Life-History Variation of common mullein (Verbascum thapsus).3. Differences among sequential cohorts. Journal of Ecology, 72, 927-936.

Reinartz JA (1984c) Life-History Variation of common mullein (Verbascum thapsus). 2. Plant size, biomass partitioning and morphology. Journal of Ecology, 72, 913-925.

Rejmanek M, Richardson DM (1996) What attributes make some plant species more invasive? Ecology 77, 1655-1661.

Riihimaki M, Savolainen O (2004) Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany, 91, 1036-1045.

Riihimaki M, Podolsky R, Kuittinen H, Koelewijn H, Savolainen O (2005) Studying genetics of adaptive variation in model organisms: flowering time variation in Arabidopsis lyrata. Genetica, 123, 63-74.

Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Syndallas B, Cabin RJ, Cohen JE, Ellstrand NC, McCauley DE, O'Neil P, Parker IM, Thompson JN, Weller SG (2001) The population biology of invasive species. Annual Review of Ecology and Systematics, 32, 305-332.

Schaffer W (1974) Optimal reproductive effort in fluctuating environments. American Naturalist, 108, 783-790.

Schlichting CD (1986) The evolution of phenotypic plasticity in plants. Annual Review of Ecology and Systematics, 17, 667-693.

Schranz ME, Osborn TC (2000) Novel flowering time variation in the resynthesized polyploid *Brassica napus*. The Journal of Heredity, 91, 242-246.

Shindo C, Aranzana MJ, Lister C, et al. (2005) Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of Arabidopsis. Plant Physiology, 138, 1163-1173.

Siemann E, Rogers WE (2001) Genetic differences in growth of an invasive tree species. Ecology Letters, 4, 514-518.

Simpson GG, Dean C (2002) Flowering - Arabidopsis, the rosetta stone of flowering time? Science, 296, 285-289.

Smith MD, Knapp AK (2001) Physiological and morphological traits of exotic, invasive exotic, and native plant species in tallgrass prairie. International Journal of Plant Sciences, 162, 785-792.

Steinmaus SJ, Prather TS, Holt JS (2000) Estimation of base temperatures for nine weed species. Journal of Experimental Botany, 51, 275-286.

Stepien CA, Brown JE, Neilson ME, Tumeo MA (2005) Genetic diversity of invasive species in the Great Lakes versus their Eurasian source populations: Insights for risk analysis. Risk Analysis, 25, 1043-1060.

Stone JM, Palta JP, Bamberg JB, Weiss LS, Harbage JF (1993) Inheritance of freezing resistance in tuber-bearing Solanum species: evidence for independent genetic control of nonacclimated freezing tolerance and cold acclimation capacity. Proceedings of the National Academy of Sciences of the United States of America, 90, 7869-7873.

Suarez-Lopez P, Wheatley K, Robson F, et al. (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature, 410, 1116-1120.

Chapter 5

Cross species amplification of Brassica SSR primers in *Hirschfeldia incana*

5.1 Introduction

Invasive species provide good model systems to study rapid evolutionary change (Durka *et al.* 2005). Conversely, the evolutionary genetics of invasive species may offer insight into the mechanisms of invasion (Lee 2002). In this context *H. incana*, a recent invader of Northern Europe, the Americas, Japan and Australasia (Morales & Aizen 2006, USDA, Weeds Australia), is potentially an excellent species for investigations of contemporary evolution. Molecular methods can provide genetic information to determine patterns of genetic diversity. This information can define population genetic structure and be used to infer routes of introduction and provide a measure of diversity of alien individuals compared to the source population. Simple sequence repeats (SSRs) or microsatellites (hereafter referred to as SSRs) are often used for estimating genetic diversity within and between populations. Although SSR's are not without analytical problems such as homoplasy (same product length but different DNA sequence), they are very good for addressing fine scale ecological questions (Selkoe & Toonen 2006). This study aimed to identify a number of SSR markers that could be used for future research programmes to study patterns of genetic diversity between populations of *H. incana*.

SSRs ('Simple Sequence Repeats', also known as 'microsatellites') consist of tandem repeated sets of DNA sequence of 1-6 bp within nuclear or organellar DNA. It is believed that errors easily occur here during the DNA replication process where extra units of DNA repeats are added to or removed from the SSR sequence by mistakes such as replication slippage (Ellegren 2000). SSR rates of mutation are higher than point mutations in non-repeating DNA sequence, as a result, SSRs are very common in the genome (e.g., Dib *et al.* 1996). The lengths of SSR loci can be highly variable between individuals within a population e.g. average expected SSR loci heterozygosities often well above 50%, (Jarne & Lagoda 1996). Since SSRs are thought to be largely neutral and co-dominant, in addition to being "hypervariable", they have been very successfully used as population genetic markers in many ecological studies (Selkoe & Toonen 2006).

It is essential that the primers do not amplify non-specific PCR products; they must therefore be unique to the flanking region of the SSR and not amplify artefacts such as primer dimers. Analysis of SSRs involves the comparison of lengths of PCR products (Goldstein & Clark 1995, Slatkin 1995). The PCR products will vary in length because SSR alleles will have different numbers of repeats. For example the length of an SSR PCR product consisting of CAA repeated three times (indicated as $(CAA)_3$) will be the sum of both flanking regions of the SSR within the targeted region of the primers and 9 bp, but the length of an SSR PCR product consisting of CAA repeated four times (or $(CAA)_4$) will be the sum of both flanking regions of the SSR and 12 bp (e.g. figure 5.1).

TTACTATCCTCCGCACAACAGAAG CAA CAA CAA AAACGAGGGCGGATTAGACGCAG
TTACTATCCTCCGCACAACAGAAGCAA CAA CAA CAA AAACGAGGGCGGATTAGACGCAG

Figure 5.1 Representation of a variable SSR locus, in both alleles, the priming and flanking regions remain the same, however variation is observed in the number of CAA repeats.

The high variability of SSRs in populations may be explained by higher mutation rates as a result of slipped strand mispairing during DNA replication (Kruglyak *et al.* 1998, Calabrese *et al.* 2001). However, if one assumes that the SSR genotype of an individual is largely due to inheritance rather than mutation, then the patterns of frequencies of the SSR alleles in populations can be used to infer the degree of genetic divergence between populations, and identify founder events and provide estimates of genetic diversity and population dynamics (Goldstein & Clark 1995, Tarr *et al.* 1998). Due to the amplification of two products of different lengths in heterozygote samples, measures of heterozygosity can be directly calculated. This is an important advantage over other methods such as RAPDs and AFLPs since heterozygosity can only be estimated using these techniques.
Various methods used to genotype SSRs by differentiating the PCR products based on size, include separation on polyacrylamide gels (Creste *et al.* 2001, White *et al.* 2002, Ceplitis *et al.* 2005), dHPLC (Pan *et al.* 2003), SSCP (Sunnucks *et al.* 2000), capillary electrophoresis (Mansfield *et al.* 1997), and automated sequencing instruments (e.g. Cawkwell *et al.* 1993, Weckworth *et al.* 2005 Woodhead *et al.* 2005).

SSR loci are therefore useful markers for studying population genetics, and the PCR-based genotyping means that the technique is amenable to high-throughput analysis. However, development of new SSR primers can be a time consuming and therefore costly process (Parker *et al.* 1998). This may represent their main disadvantage as a molecular method for low-throughput and cash-poor ecological investigations. Other limitations include the occurrence of null alleles due to mutations in the PCR priming region in some individuals (Callen *et al.* 1993, Lehmann *et al.* 1996) and mutational processes that do not follow the assumptions of models of step-wise mutation. Occurrence of null alleles means that heterozygosity is underestimated and frequencies of the different alleles in a population may not be accurately assessed (Dakin & Avise 2004). PCR can also display preferential amplification of particular alleles and this may lead to misidentification of heterozygote samples as homozygotes (Gagneux *et al.* 1997). Nevertheless due to the increased availability of sequence data and PCR primers for a variety of species, SSR analysis has become a popular technique for ecology and may be one of the most frequently used molecular marker for population studies at present.

SSR regions can be identified by searching DNA sequence databases, either manually or using automated scanning software such as Repeat masker (http://www.repeatmasker.org/). Once SSR sites have been identified, primers can then be designed to target PCR amplification of these regions. If DNA sequence data are lacking, randomly located SSRs can be discovered using hybridisation and cloning techniques. Fluorescently labelled probes designed to hybridise with certain tandem repeats can be used to identify the presence of SSR regions. Positively labelled clones are sequenced and microsatellite primers designed to complement the sequences of the flanking regions. This approach can be time consuming, as not each SSR region will necessarily show polymorphisms and therefore a large proportion of candidate SSR loci may prove not to be useful.

However, cross-species SSR primer use has been applied (e.g. Edh *et al.* 2007) and may offer a low cost alternative to developing SSRs for each species. Previously identified primers may or may not amplify the same locus in related species, depending on the

conservation of flanking region sequences. Thus, as more primers to amplify polymorphic SSR loci become available, greater resources are available for the ecologist wishing to perform population studies. When developing SSR primers for target species, researchers are now regularly checking for cross-species amplification (e.g., Primmer 1996, Gauffre 2006). Closely related species have a greater likelihood of amplifying the same locus successfully. For example, Hoshino (2006) describes the development of SSR primers for species from the genus *Arachis* using genomic libraries. Fifteen primers were tested in 76 accessions of 34 species from nine sections of the genus *Arachis*. High transferability of primers across the species range was observed (primers amplified a product in 91% of the species), amplification of polymorphic loci was also observed. In another example, Hansen (2005) reported amplification of five SSR primers designed for *Abies nordmanniana* across 19 species of the *Abies* genus; cross-species amplification was found to be high (successful amplification in over 85% of species) throughout the genus.

A substantial volume of genetic information is available for species in the Brassicaceae, of which *H. incana* is a member. Many of the genetic studies have involved the model plant *Arabidopsis thaliana* and more recently there has been a greater understanding of the genetics of *Brassica* species. Large amounts of sequenced data generated by high-throughput methods such as shotgun sequencing have allowed the development and characterization of SSR markers by digitally searching the genome data for SSRs (e.g. Burgess et al 2006). In addition, SSR regions can be digitally isolated from expressed sequence data. These SSRs are shown to give reliable amplification, polymorphisms and transferability across species (e.g. Batley *et al.* 2007). Database searches of expressed sequence tags (EST)-SSRs using Blastx (Basic Local Alignment Tool) (Altschul *et al.* 1990) can identify possible homologous sequences and may assist in the association of molecular markers with candidate genes (Hopkins *et al.* 2007).

Given the close phylogenetic association of *Hirschfeldia incana* with *Brassica* (Facey 2006), a logical starting point to find suitable primers to amplify microsatellites in *Hirschfeldia incana* would be to test primers designed for *Brassica* species. With a

greater interest in *Brassica* genomics and greater availability of sequence data numerous microsatellite primers have been designed for species in this genus. Approximately 800 SSR primers for the use with *Brassica* species are listed by the *Brassica* Geneflow Consortium project (http://www.brassica.info/ssr/SSRinfo.htm). Major contributors include: Lagercrantz *et al.* (1993), Bell and Ecker (1994), Kresovich *et al.* (1995), Szewc-Mc Fadden *et al.* (1996), Westman and Kresovich (1998), Raybould *et al.* (1999), Uzunova and Ecke (1999), Saal *et al.* (2000), Plieske and Struss (2001), Lowe *et al.* (2002), Suwabe *et al.* (2002), Tommasini *et al.* (2003) and Lowe *et al.* (2004). Additional primer information is compiled within the Molecular Ecology Notes Database (http://tomato.bio.trinity.edu/MENotes/home.html) including a large number of primers for SSR amplification for many species within the Brassicaceae. When available, cross-species amplification information is included in the *Brassica* DB. In addition, for some primers, some information or comment is included on interpretability of the PCR products.

Due to the lack of existing sequence data for *H. incana* and the time and cost involved in designing SSR primers using hybridization techniques, the approach I took was to select a proportion of the primers described in the *Brassica* database (selection criteria described in materials and methods) and test their applicability for use with *H. incana.* I tested these primers by PCR amplification with DNA extracted from *H. incana.* Primers found to amplify products of expected size were then investigated further to identify polymorphic loci. PCR products were also sequenced to confirm presence or absence of the expected SSR. If these conditions are met these SSRs may be suitable for future population studies of *H. incana.*

5.2 **Methods and materials**

5.2.1 ***Primer Selection.*** The Brassica Microsatellite Information Exchange (http://www.brassica.info/ssr/SSRinfo.htm) was used as the source for potential primer pairs to amplify SSRs in *H. incana*. Primers were initially ranked according to possible suitability for amplification in *H. incana*. Selection criteria for potentially suitable SSRs included reports of cross-species amplification, reports of polymorphic amplification within multiple species and interpretability of amplification. Forty two primer pairs were short-listed from the database that met these conditions (table 8.8) from these, 10 were selected for testing (Table 5.1). The final 10 were selected out of the 42 based on reported interpretability of data (reported as a score of 1-5). In addition, 3 base-pair repeats were given preference over 2 base-pair repeats because the 3-base difference provides a greater degree of separation that would be easier to detect using electrophoresis methods. No repeats were listed in the database containing more that 3 base-pairs (e.g. tetramers).

Table 5.1 *Brassica* SSR primer pairs identified to test with *H. incana*. * indicates that the repeat number was not listed in the database.

Primer Nuclear	Forward	Reverse	Repeat (n)
Na14 E08	TTACTATCCCCTCTCCGCAC	GCGGATTATGATGACGCAG	GCC (6)
OI10 H02	AACAGGAAGAAACGACGAGG	AGAGAGCCATGAGAAGCACC	GGC (8)
OI11 G11	GTTGCGGCGAAACAGAGAAG	GAGTAGGCGATCAAACCGAG	CCG (5)
BRMS 001	GGTGGCTCTAATTCCTCTGA	ATCTTTCTCTCACCAACCCC	GA (*)
BRMS 003	ACGAATTGAATTGGACAGAG	CAGATGGGAGTCAAGTCAAC	CT (*)
BRMS 019	CCCAAACGCTTTTGACACAT	GGCACAATCCACTCAGCTTT	GT (*)
Na10 D09	AAGAACGTCAAGATCCTCTGC	ACCACCACGGTAGTAGAGCG	GT (11)
Na12 A01	GCATGCTCTTGATGAACGAA	GCTTCAACCTCTCAATCGCT	CT (34)
Na12 A02	AGCCTTGTTGCTTTTCAACG	AGTGAATCGATGATCTCGCC	CT (16)
Na10 F06	CTCTTCGGTTCGATCCTCG	TTTTTAACAGGAACGGTGGC	CCG (6)

In addition to those primer pairs identified in the BBSRC database, an additional 13 primer pairs were provided by Lund University (table 5.2). These primers had been used

previously with *Brassica cretica*, and included loci on the chloroplast genome (typically mononucleotide repeats e.g. Provan *et al.* 2001).

Table 5.2 *Brassica* SSR primers supplied by Lund University.

Primer	Forward	Reverse	Repeat
Nuclear			
BN12A	GCCGTTCTAGGGTTTGTGGGA	GAGGAAGTGAGAGCGGGAAATCA	(GA)(AAG)
MB4	TGTTTTGATGTTTCCTACTG	GAACCTGTGGCTTTTATTAC	TG
Na12-A07	TCAAAGCCATAAAGCAGGTG	CATCTTCAACACGCATACCG	GT
Ni4-B10	GTCCTTGAGAAACTCCACCG	CCGATCCCATTTCTAATCCC	CT
Ol10-F11	TTTGGAACGTCCGTAGAAGG	CAGCTGACTTCGAAAGGTCC	GGC
Ol12-F02	GGCCCATTGATATGGAGATG	CATTTCTCAATGATGAATAGT	TC
sORA26	TGTTTACCTGTTGGAGAT	AACCCTAAGCATCTGCGA	GA
Ol09-A06	TGTGTGAAAGCTTGAAACAG	TAGGATTTTTTTGTTCACCG	GA
nga111	TGTTTTTTAGGACAAATGGCG	CTCCAGTTGGAAGCTAAAGGG	CT
Chloroplast			
ccmp6	CGATGCATATGTAGAAAGCC	CATTACGTGCGACTATCTCC	(T)C(T)
ccmp10	TTTTTTTTTAGTGAACGTGTCA	TTCGTCGDCGTAGTAAATAG	T
ATCP 28673	GCGTTCCTTTCATTTAAGACG	TGCACTCTTCATTCTCGTTCC	T
ATCP 70189	CGGGTTGATGGATCATTACC	GCAATGCACAAAAAAAGCCT	A

5.2.2 ***DNA Extraction.*** DNA was extracted from *H. incana* using the Gentra Puregene mouse tail kit. Extractions were performed from fresh, frozen or dried leaf tissue following the manufacturer's recommended protocol for plant material. Extracted DNA was visualised by electrophoresis using 1% agarose gels stained with 0.5 μg/ml ethidium bromide and viewed under a U.V. transilluminator. DNA extractions were checked for purity and concentration in a spectrophotometer (Nanodrop) at A260/280 nm. DNA extractions were diluted to 100 ng/μl for storage and diluted into aliquots at concentrations of 10 ng/μl for PCR reactions. One hundred and twenty nine DNA extractions were prepared from 9 discrete populations of *H. incana*, including: Swan2 (27), Mor2 (2), Spain (22), Alg1 (4), Mor1 (27), USA (2), Turk (1), Crete (20), Cyp (1), Alg2 (23), population codes are described in table 4.1. All samples were taken from plants grown from seed planted in Swansea. Seeds were either collected personally (Swansea by the author, Crete by Dr. Rosemary John of Swansea University) or supplied

from seed collections from the following institutions; Tohoku University (Turk, USA, Mor1, Alg1), University of Hamburg (Cyp), University of Madrid (Mor2, Spain, Alg2). DNA was also extracted from *Brassica napus*, *Brassica oleracea* and *Brassica nigra* samples to act as controls when used with primers designed for these species.

5.2.3 ***SSR PCR amplification.*** Primer pairs were initially tested for amplification using a reaction profile described in Lowe 2004. PCR amplification was performed using initially using a DNA Engine (MJ Research), using the following thermocycle profile: 94°C for 1min 30 s, followed by 35 cycles of 94°C for 30 s, 56°C for 1 min, 72°C for 1 min, and ending with a final extension at 72°C for 10 min. Reaction volumes of 10 µl were used, with the following reagents: 10 ng template, 1x buffer, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 500 µM of each primer, 0.4 U taq DNA polymerase and H_20 to make up 10 µl. Optimisation of PCR conditions was carried out. $MgCl_2$ concentration was varied between 1.5 and 2.5 mM and annealing temperature was varied from 45 - 60°C.

To check first for the presence of PCR products, these were separated by electrophoresis in 1.5% agarose gels containing ethidium bromide (0.5 µg/ml), visualised under a U.V. trans-illuminator and photographed in a BioRad Gel Doc with Bio Rad Quantity 1 software version 4.4.0 Basic.

Primers resulting in clean bands of PCR products with sizes that approximated those of published data were considered for further testing. Further analyses of PCR products involved polyacrylamide gel electrophoresis, Real-Time PCR and finally, sequencing. PAGE gels were inspected to determine if SSR variation could be visualised and whether this would be a suitable technique for identifying polymorphic loci and scoring population variation. Real-time PCR and DNA melting curves were also used to differentiate between products (see Chapter 6) and to assist in selecting different products for sequencing. A range of different PCR products were selected for sequencing to confirm if the desired SSR was present and if there was variation at the allele. Different products were identified either by size differences observed on PAGE gels, or by different melting temperatures observed on the Roche lightCycler (DNA melting analysis

described in Chapter 6). Only homozygous products were useful for sequencing. Homozygosity was determined by either the observation of single bands on PAGE gels or by the generation of a single peak by using the melt analysis feature of the Roche LightCycler. The most frequently used method for analysis of microsatellites is currently fluorescently labeled primers and automated sequencing. The outputs from automated sequencing include the product lengths of individual alleles. It can identify heterozygotes and can be used for genotyping. However, due to the expense of this technique and lack of freeware software for data analysis (only recently available – too late for this project), it was outside of the scope of this study and less expensive techniques were attempted.

PCR products were run on 6% polyacrylamide gels 15x12 cm for increased resolution, particularly in differentiating between homozygous and heterozygous samples, and increased sensitivity in revealing non-specific products. These gels were run at 300 V for approximately three hours. PCR products were visualised on the PAGE gels using the following silver staining procedure:

Polyacrylamide gels were placed individually into suitably sized trays to which 1/3 of the prepared fixing solution (A) was added (Table 5.3). The gels were then agitated in this solution for 2-3 minutes. This solution was then discarded and replaced with a further 1/3 of the original volume of solution A and then agitated again for an additional 2-3 minutes. The solution was again discarded and the whole volume of the staining solution (B) was added. Gels were agitated in solution B for 10 minutes and then thoroughly rinsed in dH_2O. To develop the bands, all of solution C was added to the gels and they were agitated for 10-15 minutes or until the bands became clearly visible. Finally, the remaining 1/3 of solution A was added to fix the gel.

Table 5.3 Solutions required in the following proportions for silver staining procedure.

Solution A: Fixing	Solution B: Staining	Solution C: Developer
dH_20 180 ml	dH_20 100 ml	dH_20 150 ml
Absolute Ethanol 20 ml	Silver nitrate 0.1 g	NaOH 2.25 g
Glacial Acetic Acid 1 ml		NaBHydride 0.015 g
		Formaldehyde 600 µl

Primers identified as producing a single clean product of approximately the expected size were used in real-time PCR with the Roche LightCycler PCR thermocycler. The following thermocycle profile was used with the Roche LightCycler: 95°C for 10 minutes, followed by 50 cycles of; 95°C for 5 seconds, 58°C for 5 seconds and 72°C for 10 seconds. The advantages of the Roche LightCycler over the MJ Research DNA Engine are faster PCR cycles due to rapid temperature ramping rates, very consistent temperatures between all samples during the reaction (as it uses a rotary carousel system rather than a heated block design), and the ability to perform post PCR DNA melt analysis to assist in PCR product identification (This is focused on in greater detail in Chapter 6, see 6.1.2 for background information and 6.2.1 for more detailed methodology).

5.2.4 ***PCR product sequencing.*** PCR products were treated with the GenElute® PCR Clean-up kit from Sigma-Aldrich prior to performing sequencing reactions. This removes excess primers, nucleotides, DNA polymerase, oils and salt, most primer-dimers and other small DNA fragments <40bp. Sequencing reactions were performed using Applied Biosystems Big Dye Terminator v3.1 cycle sequencing kit. Sequencing reactions were performed with the following reagents and protocol: reaction volume 10 µl, Big Dye 0.5 µl, 5x buffer 1.75 µl, H_20 5.75 µl, template DNA 1 µl (1-3 ng). The following profile was used for the sequencing reactions: denaturing at 96°C for 1 minute followed by 25 cycles of; 96°C for 10 seconds and 50°C for 5 seconds. This was then terminated by holding the temperature at 60°C for 4 minutes.

Sequencing reactions were then cleaned up and precipitated using the following procedure (protocol from Zoology Department, Oxford University): 15 µl H_20 was added to each sample. 7 ml of absolute ethanol was mixed with 280 µl NaOAc pH 5, from this mixture 52 µl was dispensed into each sample. The samples were sealed, vortexed and incubated at room temperature for 15 minutes. The samples were then placed in a centrifuge for 1 hour at 2750 x g at 4°C. The seals were then removed and samples were

inverted onto absorbent paper. The inverted samples were spun in the centrifuge briefly at 300 rpm. To each sample 150 µl 70% ethanol was added and returned to the centrifuge for 10 minutes at 2750 x g. Samples were inverted again onto absorbent paper and spun in the centrifuge briefly at 300 rpm. Finally samples were placed into a vacuum centrifuge for 10 minutes to fully remove traces of ethanol.

Sequenced products were electrophoretically separated on ABI automated sequencing machines by an external sequencing facility (Zoology Department, Oxford University). Sequence traces were viewed using the software FinchTV version 1.4. and analysed with *Staden* programs (Staden 1996), as provided within the *Bio-Linux* bioinformatics package (NERC Environmental Bioinformatics Centre, http://envgen.nox.ac.uk/biolinux.html). PCR sequences were compared with published data to determine if the correct microsatellite region was amplified. Sequences were investigated to determine if variation was present between samples and also to identify if SSRs consisted of perfect (e.g. CT_x) or imperfect repeats (e.g. GCC_xGTC_y).

5.3 Results

5.3.1 ***Details on the amplification of Brassica SSR primers with H. incana DNA*** **(selected from the BBSRC Brassica microsatellite database).** The success of PCR amplifications of *H. incana* DNA with *Brassica* primers was mixed. Of the 10 primers selected for potential suitability, seven yielded products of expected sizes when viewed on agarose gels. Five of these PCR products gave single bands on agarose gels indicating good potential for use with this species. The resolution of agarose gel electrophoresis is not sufficient to identify heterozygote SSRs, therefore if multiple bands were present it would indicate the generation of non-specific products. Amplification and specificity of the PCR reaction were found to be the most successful at 1.5-2 mM $MgCl_2$ with annealing temperatures of 56°C. These conditions are the same as those reported for the original target species. Lower annealing temperatures and higher magnesium concentrations both resulted in the formation of additional non-specific products for all primers. Annealing temperatures greater than 56°C resulted in a decrease in intensity of the amplified product when visualised on an agarose gel.

Table 5.4 Primers tested from BBSRC Brassica microsatellite database and those supplied by Lund University. The table indicates the expected PCR product size when amplified with the original species that the primer was designed for, the approximate product size amplified with *H. incana* DNA, whether a single or multiple bands were produced from *H. incana* DNA (agarose gel electrophoresis), the published SSR, and the observed SSR sequenced from *H. incana* ("Not visible" indicates that a SSR was not observed in the sequence trace, "not determined" indicates that sequencing was not undertaken).

Primer	Expected band size	Approximate Product size	Single band on agarose	Published SSR repeat	Observed SSR repeat
Nuclear					
Na14 E08	116	110	Y	GCC	GCC
Ol10 H02	220	220	N	GGC	Not visible
Ol11 G11	132	130	N	CCG	Not determined
BRMS 001	139	600	N	GA	Not determined
BRMS 003	192	300	N	CT	Not determined
BRMS 019	220	220	Y	GT	Not determined
Na10 D09	280	260	Y	GT	Not determined
Na12 A01	176	150	Y	CT	CT
Na12 A02	190	200	Y	CT	CT
Na10 F06	110	220	N	CCG	Not determined
Primers supplied by Lund University					
Nuclear					
OL10F11	180	180	Weak band	GGC	Not determined
OL12F02	120	No amplification	N/A	TC	Not determined
NA12A07	180	180	50% samples	GT	GT
NI4B10	220	200	Y	CT	CT
OL9A06	130	No amplification	N/A	GA	Not determined
Nga111	120	No amplification	N/A	CT	Not determined
MB4	100	No amplification	N/A	TG	Not determined
Chloroplast					
CCMP6	100	100	2 products	(T)C(T)	(T)C(T)
CCMP10	100	100	Y	T	T
ATCP70189	120	120	Y	A	A
ATCP28673	130	130	Y	T	T

The details of amplification of PCR products from the primers listed in Table 5.4 are individually summarised below with the exception of OL10F11, OL12F02, OL9A06, NGA111 and MB4 which did not amplify a product and were not investigated further.

5.3.1.1 NA14E08. Polyacrylamide gels showed that NA14E08 produced a single product of size close to what is expected, indicating good specificity of the primer. Polymorphisms were also visible between some individuals which gave an early indication that a variable SSR product may be present (Figure 5.1).

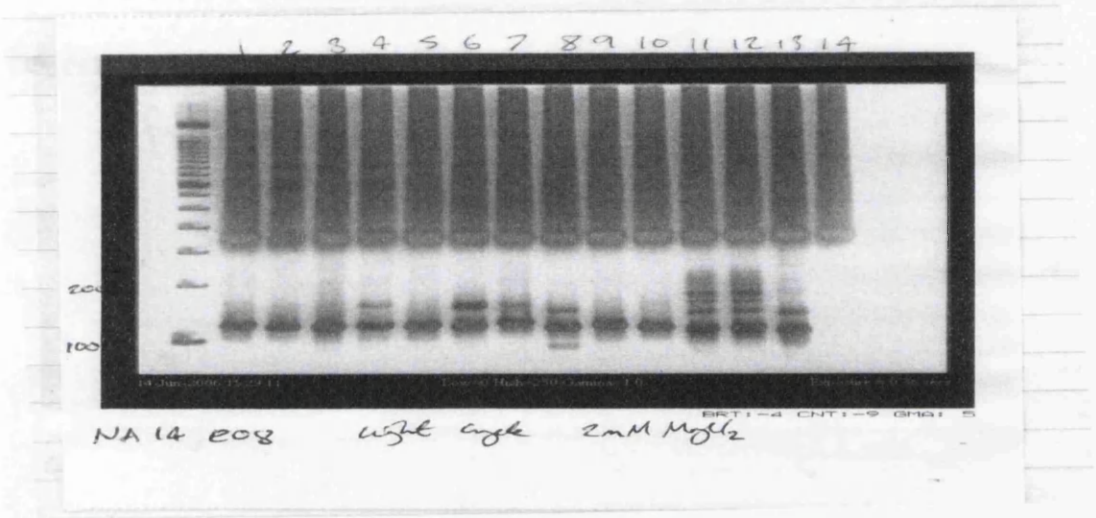


Figure 5.2 Replication of a polyacrylamide gel produced using primer NA14E08. Wells 1-10 contain DNA template from *H. incana* 11-13 from *B. napus* and 14 H_20. Close inspection reveals slight variation in product size between some of the *H. incana* samples. Smearing on the gel was due to the presence of dye included in this particular PCR master mix.

SSR variation observed by sequence data:

A total of 15 PCR products for NA14E08 were sequenced in both directions. The published SSR repeat in *Brassica napus* is $(GCC)_6$, the variation in the SSR region in *H. incana* was observed as follows:

$(GCC)_3(GTC)_1$ (5 sequences), all 5 individuals from Swan2

$(GCC)_3(GTC)_0$ (3 sequences), all 3 individuals from Mor1

$(GCC)_4(GTC)_0$ (4 sequences), all 4 individuals from Crete

$(GCC)_5(GTC)_0$ (2 sequences), both 2 individuals from Crete

$(GCC)_5(GTC)_1$ (1 sequence), individual from Spain

TTACTATCCC CTCTCCGCAC NNTNGNGATC GCAGCCGCCG CCGTCGGTGA TGATAACGAT GAAGAGTACC TCGCCGACGA ACAAGCCAGC GCTGCAGACG NCNGCTGACG TCATCATAAT CCGC

Figure 5.3 Consensus sequence for NA14E08, variant $(GCC)_3(GTC)_1$ (samples from Swan2). This was produced with GAP4 software that assembles sequences by use of a Bayesian consensus alogarithm. The sequence listed here is based on the assembly of four sequences.

5.3.1.2 OL10H02. This SSR locus initially looked promising, but repeated PCR reactions, visualization with PAGE gels (production of multiple bands, Figure 5.3) and real-time PCR (production of multiple melt peaks) confirmed that the primers repeatedly produced non-specific products. Sequencing was attempted with two samples, but the sequence traces were not of good quality, the SSR was not present and this primer was not investigated further.

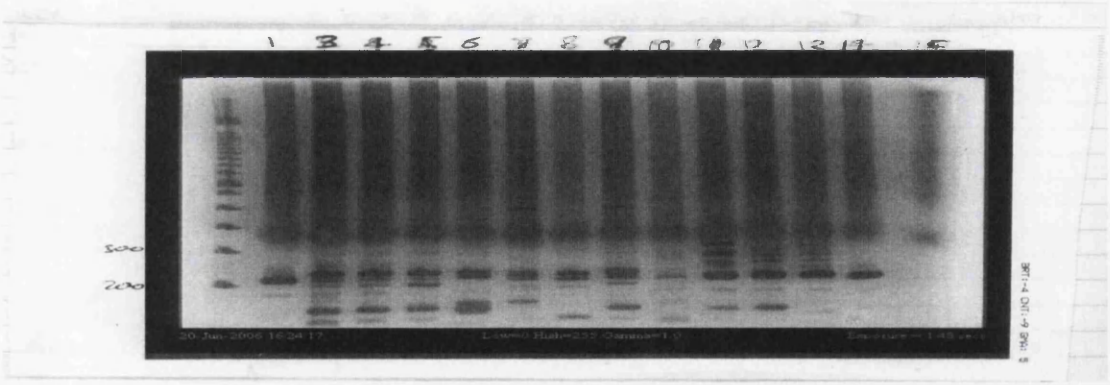


Figure 5.4 Replication of a polyacrylamide gel produced using primer OL10H02. Wells 1-10 contain DNA template from *H. incana* 11-13 from *B. napus* and 14 H_2O. Multiple products were produced for *H. incana* samples indicating that the primer was not suitable for use with this species.

5.3.1.3 OL11G11. Primers for OL11G11 amplified a product of approximately the correct size in *H. incana* albeit with rather wide, non distinct bands, when visualised on agarose gels. When these products were visualised on PAGE gels with silver staining a large number of different sized products were visible, in addition to inconsistent amplification of the desired product between different individuals. This primer was not investigated further.

5.3.1.4 BRMS001 & BRMS003. Primers for BRMS001 amplified 2 PCR products. Stronger bands were seen on agarose gels for the band of 'non-expected' size (600 bp) than 'expected' size (140 bp) and this was not resolved by altering reaction conditions. This was also the case for BRMS003 where bands of 300 bp and 190 bp were produced. The difference in size for the heterozygote products from these primers was too great to be from variation in SSR length and both primers were rejected at this stage.

5.3.1.5 BRMS019. Primers for BRMS019 amplified single PCR products of the expected size, but inconsistently as the primer did not amplify the desired product with all individuals i.e. null alleles were encountered. Since other primers performed more reliably, and financial resources were limiting, those other primers were prioritized. PCR products of BRMS019 were not taken further than visualization on agarose gel and were not sequenced.

5.3.1.6 NA10D09. Primers for NA10D09 amplified bands of approximately expected sizes (280 bp), but inconsistently. For some samples a smaller product was produced (150 bp), this unknown product was also observed with *B. napus* template. In some cases, NA10D09 did not amplify any product with *H. incana* template DNA. Products were not sequenced for this primer.

5.3.1.7 NA12A01. Primers for NA10A01 appeared to amplify a single product of slightly shorter length than expected when viewed on agarose gels (approximately 150bp instead of an expected 176bp). When visualised on PAGE gels some additional products were visible. Sequencing confirmed the presence of the expected SSR region. However, sequences were not easy to interpret due to poorly differentiated peaks in the sequence trace. The SSR region was shorter than reported for the original species (*Brassica napus*). Sequence quality was not good enough to produce a consensus using GAP4, however the SSR region was observed as variable.

SSR variation observed by sequence data:
A total of 8 PCR products for NA12A01 were sequenced in both directions. The published SSR repeat in *Brassica napus* is $(CT)_{34}$, the variation in the SSR region in *H. incana* was observed as follows:

$(CT)_4$ (2 sequences) 1 individual from each of Alg1 and Cyp
$(CT)_8$ (4 sequences) 3 individuals from Swansea, 1 individual from Spain

5.3.1.8 NA12A02. Primer NA10A02 amplified a single product with *H. incana* template when visualised on an agarose gel. Silver staining of PAGE gels indicated signs of some non-specific product. The PCR product was slightly shorter than expected (approximately 170bp compared to the expected 190bp). Sequencing confirmed the presence of the expected SSR region and that the loci were variable. The sequencing traces included some areas of poorly differentiated peaks which may have resulted from heterozygote PCR products. Sequence quality was not good enough to produce a consensus using GAP4.

SSR variation observed by sequence data:
A total of 8 PCR products for NA12A02 were sequenced in both directions. Two sequences traces were not clear enough to identify the number of repeats. The published SSR repeat in *Brassica napus* is $(CT)_{16}$, the variation in the SSR region in *H. incana* was observed as follows:

$(CT)_7$ (3 sequences) 2 individuals from Crete, 1 individual from Mor1

$(CT)_8$ (1 sequence) 1 individual from Swan2

$(CT)_{14}$ (2 sequences) 2 individuals from Mor1

5.1.3.9 NA10F06. Primers for NA10F06 did not amplify a PCR product in *H. incana*; products were only generated using DNA template from the target species of *Brassica napus*.

5.3.2 ***Details on the amplification of SSR primers supplied from Lund University with H. incana DNA.*** Of the 11 primers supplied by Kristina Edh of Lund University (previously used with *Brassica cretica* at that institution), 6 amplified products (viewed on agarose gels) of the expected size in *H. incana*. Two primers that successfully amplified products were those targeting SSR loci in nuclear DNA while four were for chloroplast DNA SSRs. The two nuclear primers amplified well with an annealing temperature of 56°C and an $MgCl_2$ concentration of 2 mM and the chloroplast primers amplified desired sized products with a lower annealing temperature of 50°C (as recommended by Edh pers. comm.). The PCR products produced by the primers supplied by Lund University were only viewed using agarose gels as they arrived at a late stage during the project.

5.3.2.1 NA12A07. Primers for NA12A07 amplified single products of expected size in *H. incana*. However, in some cases poor amplification was observed. Two samples amplified with NA12A07 primers were sequenced. Both of these products contained the repeat $(GT)_4$. The SSR was only visible on the reverse sequences, which gave clearer results than forward sequences. The published repeat in *Brassica napus* is $(GT)_{11}$.

Primers for NI4B10 amplified well and produced clean PCR products when visualised on agarose gels. Two samples amplified with NI4B10 primers were sequenced. Both of these products contained the SSR repeat CT. Variation was identified between the two samples. One sample produced $(CT)_9$, the other $(CT)_{11}$. The SSR repeat was visible in both forward and reverse sequences; however, sequences derived from the forward primer had greater clarity. The published repeat in *Brassica nigra* is $(CT)_{20}$.

All four chloroplast DNA primers (CCMP6, CCMP10, ATCP70189 and ATCP28673) amplified well with *H. incana* template DNA, producing single products of expected size. When visualised on an agarose gel, bands were strong and clean.

5.3.2.2 CCMP6. Two samples amplified with CCMP6 primers were amplified. Reverse sequence traces were found to have poorly differentiated peaks. Variation was identified between the two samples, including repeats of $(T)_2C(T)_7$ and $(T)_3C(T)_9$ that were identified in the forward sequences.

5.3.2.3 CCMP10. Two samples amplified with CCMP10 primers were sequenced. The peaks of the sequence traces were poorly differentiated resulting in useful information from only one sequence. Only the reverse sequence from one sample gave clear results. This sequence contained two repeats; $(T)_7$ and $(T)_{12}$, that potentially could be the targeted SSR.

5.3.2.4 28673. Two samples amplified with 28673 primers were sequenced. The peaks of the sequence traces were poorly differentiated. The expected repeat of T was observed in the reverse sequence of one sample, $(T)_3$, and the forward sequence of the other sample, $(T)_4$. However, such short single base repeats may not be the expected SSR or may represent only a remnant of an ancestral SSR. In any case this is unlikely to be useful for population studies as there will not be much variation at such a locus.

5.3.2.5 70189. Two samples amplified with 70189 primers were sequenced. Both these products contained an $(A)_{10}$ repeat which was visible in both forward and reverse sequences.

Table 5.5 Summary table of observed SSR amplification in *H. incana* using *Brassica* primers. Published SSR indicates the repeat observed from the original target species. Observed SSR indicates the observed repeat from this loci in *H. incana*. SSR variation indicates the variability of the loci based on the number of repeats observed from sequenced samples.

Primer	Published SSR	Observed SSR	SSR variation
Nuclear			
NA14E08	$GCC_{(6)}$	$(GCC)_x(GTC)_y$	(3,0), (4,0), (5,0), (3,1), (5,1)
NA12A01	$CT_{(34)}$	$CT_{(x)}$	(4), (8)
NA12A02	$CT_{(16)}$	$CT_{(x)}$	(7), (8), (14)
NA12A07	$GT_{(11)}$	$GT_{(x)}$	(4)
NI4B10	$CT_{(20)}$	$CT_{(x)}$	(9), (11)
Chloroplast			
CCMP6	(T)C(T)	$(T)_xC(T)_y$	(2,7), (3,9)
CCMP10	T	$T_{(x)}$	(7), (12)
ATCP70189	T	$T_{(x)}$	(3), (4)
ATCP28673	A	$A_{(x)}$	(10)

5.4 Discussion

For an SSR locus to be useful for an ecological population study it must fit certain criteria. The alleles of the locus should be consistently amplified for all individuals in all populations of a species. To be useful as a population marker, SSRs must be variable (polymorphic). This variation must be detectable by an approach that can accurately quantify it. It is preferable that the variation is restricted to the repeat number of the SSR itself, rather than caused by base substitutions, insertions or deletions occurring within the flanking region of the PCR product, including the primer site. Variation within the flanking site will complicate differentiation between alleles. Variation within the primer sites may lead to poor quality amplification, mispriming or no formation of a PCR product (null allele) due to too much mismatch between the primer sequence and the sequence of the target template.

Cross species amplification of nuclear and chloroplast SSRs using conserved primer regions have been reported frequently (e.g. Cairney *et al.* 2000, Soliva *et al.* 2000, Pleiske & Struss 2001, Suwabe *et al.* 2002, Sawabe *et al.* 2004). However, confirmation of cross species amplification is often limited to the presence of a band when visualised on an agarose gel. In some cases further investigations are performed to determine if polymorphisms exist (e.g. Clauss *et al.* 2002, Bon *et al.* 2005).

In contrast to previously published work relating to cross species amplification of primers, e.g. within the *Brassicaceae* (Suwabe *et al.* 2004), the genus *Abies* (Hansen 2005), and the genus *Arachis* (Hoshino 2006), successful cross amplification with *H. incana* appeared low. Five out of 17 nuclear *Brassica* primers (29.4% success) and 3 out of 4 chloroplast *Brassica* primers (75%) successfully amplified in *H. incana* compared to a set of primers that have had reported success rates of >90% among *Brassica* species, 78.5% in *Raphanus sativus*, 68% in *Sinapis alba* and 39.8% in *Arabidopsis thaliana* (Suwabe *et al.* 2004). However the evaluation of primer performance in this study was also backed up by sequencing the PCR products to confirm the presence of the SSR repeat motif, not just based on the presence of a band on an agarose gel. If only the

presence of bands on agarose gels were relied upon to access the success of cross amplification, the impression would have been that 10 out of 17 of the nuclear primers and all of the chloroplast primers successfully amplify a product (58.8% and 100% success rate in both cases respectively compared to 29.4% and 75% confirmed by sequencing).

Amplification of SSR products varied between each individual primer pair. This was demonstrated by the variation in amplification success and clarity of PCR products when visualised on PAGE or agarose gels e.g. the production of a single band. Sequence data also gave an indication of the conservation of primer sites based on the overall clarity of the sequences. Clean sequences with little interference or scatter would imply good conservation of the primer site and that the primers have successfully amplified a single product. Conversely, multiple bands on electrophoresis gels or sequences with heavy background noise indicate poor amplification of the PCR product, contamination with non-specific products (Figure 5.4) or the presence of heterozygotes. PCR products from primers NA14E08, NA12A02, NA12A07, and NI4B10 gave electrophoresis gels with strong clean bands and readable sequence data. Primer NA12A01 produced apparently clean PCR products when visualised on agarose gels, but the sequence data was of poor quality. Primer OL10H02 suffered from generation of additional non specific products and consequently very poor sequence data.

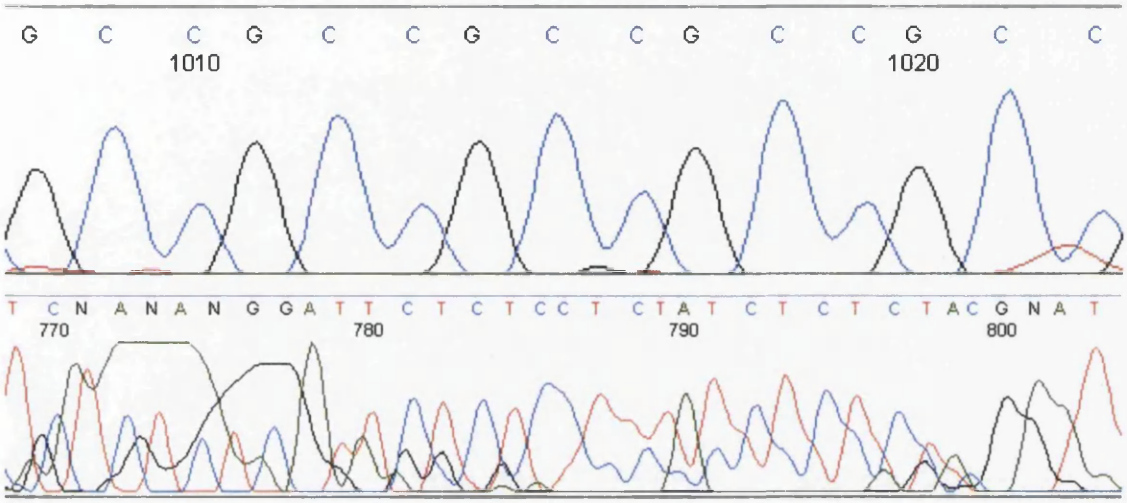


Figure 5.5 Variation in sequence trace quality experienced with different primers used with *H. incana*. Well differentiated peaks from NA14E08 (above) are easily distinguished and interpreted. Poorly differentiated peaks, this example from NA10A02, are not as clear cut and the inferred sequence may be subject to an individuals (or software) interpretation, possible sources of multiple peaks include the amplification of more than one product.

From the results of provided by electrophoresis and sequence data, primers NA14E08, NA12A02, NA12A07, and NI4B10 have shown the greatest potential for use with *H. incana*. These primers would make a good starting point in any investigation using SSRs for *H. incana*. Primers NA14E08 and NA12A02 displayed polymorphic products in *H. incana*, NA12A07, and NI4B10 have not as yet been tested for polymorphic products due to financial limitations, as stated earlier in the chapter. Amplification using individuals from a range of populations should give an indication of any polymorphism. It may be advisable to reject all other primers tried in this study for use in *H. incana* and test other potentially suitable primers from the Brassica database or the literature.

The low observed success rate (29.4%) of transferable nuclear SSR primers between *Brassica* species and *H. incana* when compared to other *Brassica* species implies that they have low value for use with this species. Ideally, SSR loci would be located from

DNA libraries created directly from *H. incana* sequence traces. These would still require further investigation to identify amplifiable polymorphic loci, but there would be greater control over primer design for the investigator. However, in the absence of such data, and the funds and time to perform such a task, selected SSR primers from the *Brassica* DB may be functional and used with *H. incana*. Due to the low success rate observed it is recommended that a greater number of primers are investigated. If possible primers could be tested in groups of 25-50, working through the database in the order they are listed, rather than selecting primers based on certain criteria. PCR amplification and visualisation on polyacrylamide gels using the protocols listed above should quickly identify potential candidates. Using an estimated success rate of 20% (based on the observed rate of 29.4%), each set of 50 primers tested would yield 10 additional SSR primers for potential use with *H. incana*.

5.5 References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology, 215, 403–410.

Bately J, Jewell E, Edwards D (2007) Automated discovery of single nucleotide polymorphism and simple sequence repeat molecular genetic markers. Methods in Molecular Biology, 406, 473-494.

Bell CJ, Ecker JR (1994) Assignment of 30 Microsatellite Loci to the Linkage Map of Arabidopsis. Genomics, 19, 137-144.

Burgess SL, Fleischer RC (2006) Isolation and characterization of polymorphic microsatellite loci in the Hawaiin flycatcher, the elepaio (Chasiempis sandwichensis). Molecular Ecology Notes, 6, 14-16.

Cairney M, Taggart JB, Hoyheim B (2000) Characterization of microsatellite and minisatellite loci in Atlantic salmon (Salmo salar L.) and cross-species amplification in other salmonids. Molecular Ecology, 9, 2175-2178.

Calabrase P, Durret R, Aquadro C (2001) Dynamics of Microsatellite Divergence Under Stepwise Mutation and Proportional Slippage/Point Mutation Models. Genetics, 159, 839-852.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR (1993) Incidence and origin of null alleles in the (AC)N microsatellite markers. American Journal of Human Genetics, 52, 922-927.

Cawkwell L, Bell SM, Lewis FA, Dixon MF, Taylor GR, Quirke P (1993) Rapid detection of allele loss in colorectal tumours using microsatellites and fluorescent DNA technology. British Journal of Cancer, 67, 1262-1267.

Ceplitis A, Su YT, Lascoux M (2005) Bayesian inference of evolutionary history from chloroplast microsatellites in the cosmopolitan weed Capsella bursa-pastoris (Brassicaceae). Molecular Ecology, 14, 4221-4233.

Creste S, Neto AT, Figuerira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Molecular Biology Reporter, 19, 299-306.

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. Heredity, 93, 504-509.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature, 380, 152-154.

Durka W, Bossdorf O, Prati D, Auge H (2005) Molecular evidence for multiple introductions of invasive garlic mustard (Alliaria petiolata, Brassicaceae) to North America. Molecular Ecology, 14: 1697-1706.

Edh K, Widen B, Ceplitis A (2007) Nuclear and chloroplast microsatellites reveal extreme population differentiation and limited gene flow in the Aegean endemic Brassica cretica (Brassicaceae). Molecular Ecology, 16, 4972-4983.

Ellegren H (2000) Heterogeneous mutation processes in human microsatellite DNA sequences. Nature Genetics, 24, 400-402.

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. Molecular Ecology, 6, 861-868.

Gauffre B, Coeur d'Acier A (2006) New polymorphic microsatellite loci, cross-species amplification and PCR multiplexing in the black aphid, Aphis fabae Scopoli. Molecular Ecology Notes, 6, 440-442.

Goldstein DB Clark AG (1995) Microsatellite variation in North American populations of Drosophila melanogaster. Nucleic Acids Research, 23, 3882-3886.

Hansen OK, Vendramin GG, Sebastiani F, Edwards KJ (2005) Development of microsatellite markers in Abies nordmanniana (Stev.) Spach and cross-species amplification in the Abies genus. Molecular Ecology Notes, 5, 784-787.

Hopkins CJ, Cogan NOI, Hand M, Jewell E, Kaur J, Li X, Lim GAC, Ling AE, Love C, Mountford H, Todorovic M, Vardy M, Spangenberg GC, Edwards D, Batley J (2007) Sixteen new simple sequence repeat markers from Brassica juncea expressed sequences and their cross-species amplification. Molecular Ecology Notes, 7, 697-700.

Hoshino AA, Bravo JP, Angelici C, et al. (2006) Heterologous microsatellite primer pairs informative for the whole genus Arachis. Genetics and Molecular Biology 29, 665-675.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends in Ecology and Evolution, 11, 424–429.

Kresovich S, Szewcmcfadden AK, Bliek SM, McFerson JR (1995) Abundance and Characterization of Simple-Sequence Repeats (Ssrs) Isolated from a Size-Fractionated Genomic Library of Brassica-Napus L (Rapeseed). Theoretical and Applied Genetics, 91, 206-211.

Kruglyak S, Durrett RT, Schug MD, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc. Natl. Acad. Sci. USA 95, 10774–10778.

Lagercrantz U, Ellegren H, Andersson L (1993) The Abundance of Various Polymorphic Microsatellite Motifs Differs between Plants and Vertebrates. Nucleic Acids Research, 21, 1111-1115.

Lee CE (2002) Evolutionary genetics of invasive species. Trends in Ecology and Evolution, 17, 386-391.

Lehmann T, Hawley WA, Collins FH (1996) An evaluation of evolutionary constraints on microsatellite loci using null alleles. Genetics, 144, 1155-1163.

Lowe AJ, Jones AE, Raybould AF, Tricks M, Moule CL, Edwards KJ (2002) Transferability and genome specificity of a new set of microsatellite primers among Brassica species of the U triangle. Molecular Ecology Notes, 2, 7-11.

Lowe AJ, Moule C, Trick M, Edwards KJ (2004) Efficient large-scale development for marker and mapping applications in Brassica crop species. Theor Appl Genet, 108, 1103-1112.

Luikart G, England PR (1999) Statistical analysis of microsatellite DNA data. Trends in Ecology and Evolution, 14, 253-256.

Mansfield ES, Vainer M, Harris DW, Gasparini P, Estivill X, Surrey S, Fortina P (1997) Rapid sizing of polymorphic microsatellite markers by capillary array electrophoresis. Journal of Chromatography A, 781, 295-305.

Morales CL, Aizen MA (2006) Invasive mutualisms and the structure of plant-pollinator interactions in the temperate forests of north-west Patagonia, Argentina, Journal of Ecology, 94, 171–180.

Pan K, Lui W, Lu YY, Zhang L, Li Z, Lu WL, Thibodeau SN, You WC (2003) High throughput detection of microsatellite instability by denaturing high-performance liquid chromatography. Human Mutation, 22, 388-394.

Parker PG, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: Choosing and using a molecular marker. Ecology, 79, 361-382.

Plieske J, Struss D (2001) Microsatellite markers for genome analysis in Brassica. I. development in Brassica napus and abundance in Brassicaceae species. Theoretical and Applied Genetics, 102, 689-694.

Primmer CR, Moller AP, Ellegren H (1996) A wide-range survey of cross-species microsatellite amplification in birds. Molecular Ecology, 5, 365-378.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trends in Ecology and Evolution, 16, 142-147.

Raybould AF, Mogg RJ, Clarke RT, Gliddon CJ, Gray AJ (1999) Variation and population structure at microsatellite and isozyme loci in wild cabbage (Brassica oleracea L.) in Dorset (UK). Genetic Resources and Crop Evolution, 46, 351-360.

Saal B, Plieske J, Hu J, Quiros CF, Struss D (2001) Microsatellite markers for genome analysis in Brassica. II. Assignment of rapeseed microsatellites to the A and C genomes and genetic mapping in Brassica oleracea L. Theoretical and Applied Genetics, 102, 695-699.

Selkoe KA, and RJ Toonen (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecolological Letters, 9, 615-629.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139, 457-462.

Soliva M, Gautschi B, Salzmann C, Tenzer I, Widmer A (2000) Isolation and characterization of microsatellite loci in the orchid Ophrys araneola (Orchidaceae) and a test of cross-species amplification. Molecular Ecology, 9, 2178-2179.

Staden R (1996) The Staden sequence analysis package. Molecular Biotechnology, 5, 233-241.

Suwabe K, Iketani H, Nunome T, Kage T, Hirai M (2002) Isolation and characterization of microsatellites in Brassica rapa L: Theoretical and Applied Genetics 104, 1092-1098.

Suwabe K, Iketani H, Nunome T, Ohyama A, Hirai M, Fukuoka H (2004) Characteristics of microsatellites in Brassica rapa genome and their potential utilization for comparative genomics in Cruciferae. Breeding Science, 54, 85-90.

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in Ecology and Evolution, 15, 199-203.

SzewcMcFadden AK, Kresovich S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated Brassica species. Theoretical and Applied Genetics 93, 534-538.

Tarr CL, Conant S, Fleischer RC (1998) Founder events and variation at microsatellite loci in an insular passerine bird, the Laysan finch (Telespiza cantans). Molecular Ecology, 7, 719-731.

Tommasini L, Batley J, Arnold GM, Cooke R, Donini P, Lee D, Law J, Lowe C, Moule C, Edwards K (2003) The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (Brassica napus L.) varieties. Theoretical and Applied Genetics 106, 1091-1101.

USDA: http://plants.usda.gov/java/profile?symbol=HIIN3

Uzunova MI, Ecke W (1999) Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (Brassica napus L.). Plant Breeding, 118, 323-326.

Weckworth BV, Talbot S, Sage GK, Person DK, Cook J (2005) A signal for independent coastal and continental histories among North American wolves. Molecular Ecology, 14, 917-931.

Westman AL, Kresovich S (1998) The potential for cross-taxa simple-sequence repeat (SSR) amplification between Arabidopsis thaliana L. and crop brassicas. Theoretical and Applied Genetics 96, 272-281.

White E, Sahota R, Edes S (2002) Rapid microsatellite analysis using discontinuous polyacrylamide gel electrophoresis. Genome, 45, 1107-1109.

Woodhead M, Russell J, Squirrel J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W (2005) Comparative analysis of population genetic structure in Athyrium distentifolium (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Molecular Ecology, 14, 1681-1695.

Chapter 6

Use of Real-Time PCR to distinguish SSR (Simple Sequence Repeat) alleles for application in molecular ecology

6.1 Introduction

6.1.1 ***SSR (microsatellite analysis).*** In order to genotype SSR loci, usually these are first isolated and amplified by PCR (polymerase chain reaction) using primers that target areas in the flanking regions of the SSR (i.e. the product will include some non-repeating and conserved sequence on either side of the SSR itself). After SSR alleles are amplified by PCR, they are commonly characterised by their product length. This works because if SSR mutations occur largely by replication slippage of individual units of repeats, then different alleles should differ by the number of repeats, and the overall size of the PCR product would reflect this length variation (Table 6.1). By measuring the sizes of PCR products and assuming that any non-repeating sequences within the PCR product remains constant, it is therefore possible to estimate how many alleles are present at each locus and observe if patterns of allele variation exist between different populations. Generally, this approach works well, but with some exceptions. For example, with compound SSRs, two alleles may have the same PCR product size but in fact have a different sequence of repeats (Table 6.1) (Garza & Freimer 1996). SSR genotyping by PCR product length assessment may therefore underestimate variation. Also, variation in the non-repeating sequences within the PCR product may lead to errors (Table 6.1) (Grimaldi & CrouauRoy 1997). Nevertheless, the approach to genotype SSRs based on PCR product size is more efficient and cost-effective than sequencing every single PCR product, and so most studies would adopt this approach and assume that errors are sufficiently low that analyses are not significantly biased. To ensure the latter, one can take precautions, such as to undertake preliminary sequencing of a subset of alleles (e.g. several alleles of similar length from diverse populations) before deciding to include a locus in the study, and to avoid including extremely divergent groups (e.g. different species) in the study since homoplasy is expected to be more problematic as populations become more distantly related (Selkoe & Toonen 2006).

Table 6.1 Examples of different types of SSR. Any method of differentiation between alleles will need to detect variation of a single repeat unit. PCR product size may not be an entirely accurate way of estimating SSR allele variation: *the same sized product may consist of different repeats and **variation in non-repeat sequences can appear to be SSR variation.

SSR type	Sequence (flanking and non repeating sequence in small caps; SSR in all caps)	Repeat motif	Size (bp)
Di-nucleotide pure repeat	gcttactgaCACACACACACAgggtcagta	$(CA)_n$	30 -
	gcttactgaCACACACACACACAgggtcagt a		32 -
Tri-nucleotide compound repeat	gttcattCAAGAAGAAGAAcctgat	$(CAA)_n(GAA)_n$	25 -
	gttcattCAACAAGAAGAAGAAGAAcctg at*		31 -
	gttcattCAACAACAAGAAGAAGAAcctg at*		31 -
Tetra-nucleotide interrupted repeat	ttgattaGATAGATAttGATAGATAtggggc a	$(GATA)_n$tt(GA TA)	32 -
	ttgattaGATAGATAGATAttGATAGATA tggggca**		36 -
	ttgattaGATAGATAGATAttGATAGATA tgca**		33 -

Following PCR, various methods exist to differentiate the PCR products based on size, including separation on polyacrylamide gels (e.g. Creste *et al.* 2001, White *et al.* 2002, Ceplitis *et al.* 2005), dHPLC (e.g. Pan *et al.* 2003), SSCP (e.g. Sunnucks *et al.* 2000), capillary electrophoresis (e.g. Mansfield *et al.* 1997), and automated sequencing instruments (e.g. Cawkwell *et al.* 1993, Weckworth *et al.* 2005 Woodhead *et al.* 2005). The presence of the PCR product is then detected by various means: staining with ethidium bromide or other fluorescing dyes that bind to DNA, autoradiography (e.g. Tessier & Bernatchez 1999), or by using PCR primers that are labelled with a fluorescing molecule. Comparison of the position of the PCR product on the gel with the positions of DNA size standards enables estimation of the size of the PCR product.

Methods of post-PCR analysis vary in their ease of application and their cost. The cheapest technique is to use polyacrylamide gels with resolutions of at least 2 bp. However, this is a highly manual procedure with problems in resolving very small differences in product size and requires exposure to toxic chemicals (acrylamide – neurotoxin, formaldehyde – carcinogenic). It may be cheap in materials, but is not efficient for analysis of large samples sizes and many loci. Instead, the most popular current technique for genotyping microsatellites uses the fragment analysis capabilities of automated sequencing instruments. This requires the use of fluorescently labelled primers, but enables multiplexed analysis of both loci and samples so that many PCR products can be typed at the same time. However, the cost of running the PCR products through an automated sequencing instrument and purchasing labelled primers is much higher, and a PCR is still required as a separate step in the approach. Regardless of the means of detecting PCR products, all current standard approaches to genotyping SSRs require two steps: PCR and post-PCR analysis.

A possible alternative to this existing approach is 'real-time PCR'. Real-time PCR was developed primarily as a technique for quantification of nucleic acids, but now has other applications. The key idea with real-time PCR is to detect and characterise the PCR product during the course of the PCR itself – this eliminates any need for post-PCR analysis. In theory, this could be a simpler, more efficient and cost-effective approach for

ecological studies using SSR data. Real-time PCR has been used to identify (genotype), or screen wheat cultivars based on microsatellites (Shepherd and Henry 1998) and more recently for grapevine and olive cultivars (Mackay *et al.* 2008), but this technique has not been used to study microsatellite variation between wild populations as has been attempted in this study.

6.1.2 ***Real time PCR and DNA melting analysis.*** Soon after PCR was first developed in the mid-80s (Saiki *et al.* 1985; Saiki *et al.*1988) real-time PCR was already being developed (e.g. Holland *et al.* 1991; Higuchi *et al.* 1993; Gibson *et al.* 1996; Heid *et al.* 1996). The method relies on the detection of fluorescence that is associated with newly amplified DNA. The detection system may be indirect – detecting the accumulation of fluorescence rather than the amplified PCR product. The most commonly used indirect detection system is the "Taqman" system (Heid *et al.* 1996), which requires fluorescently labelled probes synthesized for every application, and is therefore an expensive technique. More generic in application, and cheaper, are direct detection systems, based on the binding of a fluorescent dye to double-stranded DNA (dsDNA) (Wittwer *et al.* 1997). Some commercially available dyes include SYBR green, LCGreen and CYT09. When these types of dye are bound to dsDNA and are excited by a specific frequency they will fluoresce. The strength of the fluorescent signal will be relative to the concentration of dsDNA present, but this approach is less specific than the "Taqman" system because any dsDNA will contribute to increased signal. The dye is included in the PCR reaction mix, and the fluorescence is measured by the real-time PCR instrument at the end of each amplification stage in the PCR reaction. In theory, the amount of DNA present doubles with every round of PCR. This will result in a doubling of fluorescence with each cycle until a plateau is reached as the reaction reaches it's end-stages (e.g. limited by available reagents). How soon the plateau is reached is also determined by the amount of starting DNA– the more copies of template DNA, the fewer PCR cycles are required to reach a specific threshold of detection – it is therefore possible to quantify the starting amount of DNA. This capability for DNA quantification is the chief application for real-time PCR, and most of the early development of the technology has been towards refining and expanding quantitative real-time PCR analysis (Walker 2002).

With the direct system of detection, the fluorescence is measured at regular intervals during the PCR reaction, to show the increasing amount of dsDNA present in the reaction. However, it is not possible to determine what fragments of DNA are being amplified. An increase in fluorescence will be observed in cases of amplification specific to the primers, mispriming and the formation of primer-dimer artefacts. To determine product specificity, DNA melt curve analysis to determine DNA melting temperatures (T_m) was developed (Ririe *et al.*1997; Lyon *et al.* 1998). To perform a melt curve the final PCR product is heated by small increments while fluorescence is monitored. Typically, monitoring of fluorescence during a melt curve is either performed continuously or after each incremental increase in temperature. The rate of change of relative fluorescence with time is plotted against temperature; this will produce a peak at the melting temperature (T_m) of the product(s), where T_m is defined as the temperature at which half of DNA double stranded helical structure is lost (Rajeevan *et al.* 2001). Melting temperature of DNA products is determined by a combination of base composition and length. G-C concentration is important due to greater bond strength (3 H bonds) compared to A-T associations (2 H bonds). Different products of either length or base composition may be identified by different melting peaks (T_m). Primer dimers are frequently identified by low melting temperatures and can thus be distinguished from specific PCR products (Ririe *et al.* 1997).

Early studies in melt curve variation investigated the possibility of identifying differences between homoduplex and heteroduplex single nucleotide polymorphisms (SNP's) (e.g. Lay and Wittwer 1997; Bernard *et al.* 1998). A heteroduplex SNP may contain a single base pair mismatch in contrast to a homoduplex, where all bases are paired in the usual expected A-T and C-G form. To detect this fluorescent probes are designed for the region of interest. One of these probes would anneal to the region that contains the mutation. When both probes hybridize in close proximity a fluorescent emission can be observed. By increasing the temperature during a melt analysis while monitoring fluorescence, where mutations are present a difference in melting temperature of the probes is apparent (e.g. Lay and Wittwer 1997). Genotypes that contain a mutation have a decrease in T_m

compared to the wild type that the probe was specifically designed for. Hybridisation probes are effective for the identification of SNP's, but their disadvantages include the requirement of detailed knowledge of SNP locations for the design of costly probes. This technology would not be suitable for the differentiation of microsatellites.

In contrast, the use of SYBR Green, or other dsDNA fluorescent dyes, for DNA melting analysis does not require the design of specific probes. However, to identify different PCR products there must be a sufficiently large difference in melting temperature between products, which may be due to a sufficient difference in length or GC composition (Aoshima *et al.* 2000). DNA melt analysis has been used for the purposes of crop variety and cultivar identification (Shepherd & Henry 1998). For such application, differences in the melting profile of 5S ribosomal RNA genes for six species allowed the differentiation of 13 out of 15 of the possible pairwise combinations of these species. Identification of wheat varieties was achieved by comparison of melting temperatures and profiles of PCR products amplified with microsatellite primers. However, it was found that fewer discriminations were possible with DNA melt analysis using SYBR Green 1 than with PAGE electrophoresis (Shepherd & Henry 1998). During a study of human inherited metabolic diseases, SYBR Green has been successfully used to identify a 9-bp deletion from a PCR product with a length of 55 bp. A 1.8°C difference in DNA melting temperature between the two products was observed (Aoshima 2000). SYBR Green has been used to distinguish between and also identify new SNPs (Lipsky *et al.* 2001). However this was not achieved in a one step process. After amplification of the product, purification and addition of SYBR green at high concentrations (that would inhibit PCR) was required. Melt analysis was then carried out on an ABI 7700 instrument. Variations in melting temp were observed between homoduplex and heteroduplex products. These ranged from 0.6-4.9° for products of 167 and 15 bp that containing SNPs.

Early methods to detect SNPs using DNA melt analysis relied on the use of restriction enzymes to create large differences between product sizes where SNPs were present or absent (Ye *et al.* 2002). However, improvements in dye chemistry and development of

PCR thermocyclers with "high resolution" melting capacity now allow the detection of SNPs directly with melt profiles and no post-processing of PCR products (e.g. Corbett Rotogene 6000, Idaho HM1, Roche Lightcycler 480). A direct comparison of SYBR green used with the Roche Lightcycler and a single labelled probe used with a "high resolution" melting instrument (HR-1 – Idaho Technology) showed a better detection of heterozygotes with the HR-1 system than with the LightCycler (Gundry *et al.* 2003). Also, cross platform comparison of instruments and dyes by Herrmann *et al.* (2006) indicated that while both SYBR green 1 and the saturating dye LCgreen could be used for genotyping SNPs by amplicon melting and analysis of T_m, only LCgreen proved useful for detecting heterozygotes by changes in the shape of the melt curve.

Therefore, with the combination of development of new dyes and new instruments, SNP detection without the need to synthesize probes is now achievable (e.g. Herrmann *et al.* 2006), and some real-time PCR systems are specifically designed for this application. In theory, the level of resolution to detect SNPs should be sufficient to detect SSR variation. In addition it is possible that other variation within microsatellite PCR products may be detected that may be missed using methods based purely on length separation. For example, an insertion or deletion within the flanking region of a microsatellite PCR product may be incorrectly identified as a change in SSR length using traditional methods (Table 6.1). This may be correctly identified as a different product using DNA melt analysis because the method is able to detect different base composition. The potential advantages of SSR identification by melt curve analysis, based on reduced costs and the rapid turnaround of results make it an interesting technique for investigation.

The aim of this study was to investigate the potential of real-time PCR and DNA melt curve analysis to differentiate SSRs in *H. incana* and if possible characterise differences in microsatellites across a geographical range of populations. To determine the suitability of this technique several areas were investigated to satisfy the requirements for genotyping samples from natural populations. The resolution of DNA melt analysis was investigated to quantify identification of differences between microsatellite alleles. Mixed samples were studied to investigate the success of identification of heterozygotes.

6.2 Method

Brassica SSR primer pairs were selected and tested for amplification with *H. incana* in Chapter 5. Primer pairs that produced products of expected size with clean bands when visualised on agarose gels were considered for use with real-time PCR. The presence of the SSR region was confirmed by sequencing as reported in chapter 5. Although a range of primers were investigated (Table 6.2), the majority of resources (e.g. time, reagents) were invested in the most promising primers NA12A02 and NA14E08. Two real-time systems were tried: the Roche Lightcycler, and the Corbett Rotogene 6000. Sufficient funds were not available for comparisons with ABI genotyping (Total project budget £1000, including the identification of functional primers).

6.2.1 ***Roche Lightcycler real-time PCR system.*** The Roche LightCycler ® Real-Time PCR System is a rotary carousel based thermal cycler platform with fluorescence detection system. PCR occurs in specially designed glass capillaries, which have an optimal surface-to-volume ratio to ensure rapid equilibration between the air and the reaction components. A typical amplification cycle requires only 30 to 60 seconds (Roche http://www.roche-applied-science.com/lightcycler-online/). The Roche Lightcycler can use either SYBR green dye chemistry or it can be used for hybridisation probe mutation screening (e.g. Taqman system).

The PCR master mix used was Lightcycler fast start DNA master plus SYBR green 1 (Roche). PCR thermal cycling protocols were modified from those used with standard PCR (as described in Chapter 5). PCR protocols were changed initially to follow a standard protocol supplied by Roche specifically for the LightCycler. Due to the design of the Lightcycler temperature ramping times are very quick in comparison to ordinary block-based PCR thermal cyclers. This allows the length of each step to be reduced during PCR. PCR was performed in 10 µl volume reactions. The reaction mix consisted of:

1 µl undiluted Roche master mix including SYBR green, $MgCl_2$ (2 mM), 10 ng of DNA, and 0.5 µM of each primer.

The following protocol was found to be successful in most cases, although for some runs where amplification was considered poor, the annealing temperature was adjusted upwards or downwards by 1 degree:

An initial denature period at 95°C was held for 10 minutes to activate the enzymes in the Roche master mix. This was followed by 50 cycles of: 95°C for 5 seconds (denature), 58°C for 5 seconds (annealing), and 72°C for 10 seconds (extension). This was followed by the melt curve analysis, which had the following profile: 95°C for 5 seconds (temperature transition 20°C/s), 65°C for 15 seconds (temperature transition 20°C/s), 95°C for 0 seconds (temperature transition 0.1°C/s) and finally 40°C for 30 seconds (cooling).

Dye fluorescence is measured (referred to as "acquisition") at the end of the extension period to monitor the progress of the PCR amplification and during the melt analysis to determine melting temperatures (T_m) of PCR products.

Primers pairs tested with real-time PCR on the Roche included: NA14E08, NA12A01, NA12A02 and OL10H02 (see Chapter 5 for details of loci and primers).

Success of the PCR reactions was also determined by electrophoresis, in addition to analysis of DNA melt profiles. For electrophoresis, PAGE (silver staining) or agarose (ethidium bromide staining) gels were used to determine product length and the presence of primer dimers. DNA melt profiles were generated for samples from all available populations to give an indication of melting temperature variation among samples. Individual samples were tested repeatedly to determine reproducibility of results from one PCR run to another (see Table 6.2).

Table 6.2a SSR primers tested with *H. incana* DNA samples with the Roche Lightcycler Real-Time PCR system.

Primer	No. of individuals tested	No. of samples repeated	Target genome
BRMS19	7	7 samples repeated 2x	Nuclear
OL10H02	57	10 samples repeated 2x	Nuclear
NA10D09	10	5 samples repeated x3	Nuclear
NA12A01	10	5 samples repeated x2	Nuclear
NA12A02	>100	30 samples repeated 2-10x	Nuclear
NA12A07	8	0 repeats	Nuclear
NA14E08	>100	17 samples repeated 2-12x	Nuclear
NI4B10	8	0 repeats	Nuclear
28673	8	0 repeats	Chloroplast
70189	8	0 repeats	Chloroplast
CCMP6	8	0 repeats	Chloroplast
CCMP10	8	0 repeats	Chloroplast

Table 6.2b Breakdown of populations and number of individuals investigated per primer (Table 6.2a).

Primer	Populations investigated (number of individuals)
BRMS19	Sw002(2), Spain(1), Crete(2), Mor1(2), *B. napus*, *B. oleracea*
OL10H02	Sw002(11), Spain (17), Crete (4), Mor1 (11), Alg2 (13), Cyprus (1), USA (1), *B. napus*, *B. oleracea*
NA10D09	Sw002(10)
NA12A01	Sw002(4), Spain(1), Crete(1), Cyprus(1), Mor1(1), USA(1), *B. napus*
NA12A02	Sw002(20), Spain(21), Crete(17), Mor1(21), Alg2(22), *B. napus*
NA12A07	Sw002(4), Spain(2), Crete(1)
NA14E08	Sw002(28), Spain(23), Mor1(27), Crete(20), Alg2(23), Alg1(4), USA(2), Mor2(2), Cyprus(1), *B. napus*, *B. oleracea*
NI4B10	Sw002(4), Spain(2), Crete(1)
28673	Sw002(4), Spain(2), Crete(1)
70189	Sw002(4), Spain(2), Crete(1)
CCMP6	Sw002(4), Spain(2), Crete(1)
CCMP10	Sw002(4), Spain(2), Crete(1)

A 10X dilution series was performed to determine impact of starting template concentration on melting profile and melting temperature. This was performed by using a quantity of DNA (10 ng) in a PCR reaction in dilutions of 1, 10, 100, 1000 and 10000. This dilution series was only performed using the primer NA14E08 and with two samples. Other primers were not tested with a range of starting conditions as the

amplification with NA14E08 was most efficient at the recommended quantity of 1 ng in a 10 µl reaction volume therefore investigating this further was not considered a priority.

The possibility of including a standard sample in each PCR run was investigated. This was done to determine if product melting temperature varied between different runs and if the affect was the same for all samples included (i.e. a linear shift in melting temperature). To investigate this; two sets of samples were PCR amplified twice. The melting profiles were then checked for each sample to determine if a change in melt temperature had occurred. A standard sample was included in successive PCR amplifications in the event that the melt profiles would be used for genotyping. If temperature shifts were linear it would increase comparability between runs. In the event that a temperature shift was observed for the 'standard' sample, it would be assumed that all samples within this PCR run will be subject to this temperature shift and could be adjusted accordingly to provide estimated standardised melt temperatures between runs. This was performed with primers Na12A02 and NA14E08, for which the greatest number of individuals were tested. Melting temperatures were plotted as frequency histograms for these two primers to visualise the general structure of melting peaks across the populations. Box plots were drawn to indicate the range of melting temperatures observed within and between populations. To analyse the differences in observed melting temperatures between populations Mann-Whitney U pairwise comparison were calculated across all populations tested with NA14EO8. Sequencial bonferroni corrections were applied to allow for multiple comparisons.

DNA melting was performed with continuous and stepwise acquisition to assess the sensitivity and clarity of the melt under both conditions. A stepwise profile measures fluorescence of each sample after each incremental increase in temperature. A continuous profile measures fluorescence continually as the melt temperature increases without the intervals of the stepwise profile.

Where variation was observed between melting profiles of PCR products, these products were sequenced to determine the differences between products (see chapter 5). When

DNA sequences were known for several samples, and SSR repeat number was known for given melt temperatures, real-time PCRs were performed to determine if sequence differences would be identified for both homozygous and heterozygous products. This was achieved by comparing melt profiles of PCRs for DNA templates with homozygous alleles and with mixed DNA templates that simulate heterozygous samples (artificial mixture of two homozygote DNA templates with different SSR alleles).

6.2.2 ***Corbett real time PCR.*** The Corbett Rotor-gene 6000 was also used to analyse SSR variation in *H. incana*. The instrument is a rotary PCR thermal cycler that can also perform DNA melt analysis with SYBR green dye chemistry. However, with more advanced dyes that bind to dsDNA in a more reliable way than SYBR green, the manufacturer claims it has the capacity to perform high resolution melt analysis (HRM) with a detection resolution of 0.02°C. This feature is supposed to be able to identify single nucleotide polymorphisms (SNP), including heterozygotes, and can therefore be used for SNP genotyping without the need for probes. This is possible because of new dyes (LCgreen®, EvaGreen™, or SYTO®9) that are advancements on SYBR®Green (traditionally used in real-time PCR) as they do not inhibit PCR under saturating conditions. Previously, melt curve analysis with real-time PCR and SYBR Green (i.e. as is done with the Lightcycler) were limited in sensitivity because of this problem, and also because real-time instruments were not strictly intended for this application. However, the combination of purpose-developed instrumentation for "high resolution melt analysis" (RotorGene™ 6000 from Corbett Life Science, and HR1™ and 384 well LightScanner™ from Idaho Technology) and the new dyes means that HRM is now a feasible method for accurate DNA fragment typing. It is also cheaper and more generic in application than TaqMan® approaches because bespoke dye-labeled probes/primers are unnecessary. The dye is added to the PCR reaction mix, together with any un-labelled primers of choice.

I used the HRM (high resolution melt) feature of the RotorGene™ 6000 to gain maximum resolution from the DNA melting analysis. Briefly, analysis is based on not just T_m (as is done with the Lightcycler) but also the shape of the melt curve, e.g. a

heteroduplex (present in heterozygous samples) has a different shaped curve compared to a homozygous sample and different homoduplexes (resulting from homozygous samples) show a simple T_m shift rather than alteration in curve shape (see Corbett website http://www.corbettlifescience.com/ for details of the instrument and HRM analysis). This difference in melt curve analysis is supposed to provide the greater sensitivity for detecting heterozygotes. *SensiMix* PCR reagents (Quantace) was used as the PCR master mix. This product includes EvaGreen™ dye for detection of amplification and DNA melting.

Initially, the PCR thermocycling protocol for the Lightcycler was used, but this required optimising due to low amplification as a result of slower temperature ramping speeds. The following PCR protocol was used for amplification using primer pair NA14E08:

An initial denature period at 95°C was held for 10 minutes to activate the enzymes in the Quantace sensi mix. This was followed by 50 cycles of: 95°C for 20 seconds (denature), 60°C for 20 seconds (annealing), and 72°C for 30 seconds (extension). This was followed by the melt curve analysis, which had the following profile: 75-90°C (0.1°C each step and sample fluorescence acquisition on the HRM channel).

One aim for the amplification on the Rotogene was to assess the capacity for identification of heterozygous SSRs (using this more sensitive DNA melt technology to compare polymorphisms) and to determine the potential for this to be used to genotype SSRs. Samples that had been identified as homozygotes from the investigation in Chapter 4 and that had been sequenced were selected for use with the Corbett. Samples with known SSR sequences were run singly and also in different combinations to determine the capability of the Rotogene system to differentiate between the alleles in a heterozygous product (Table 6.3).

Table 6.3 Template DNA of known sequence used to test ability of Rotogene to detect heterozygous SSR products. All possible combinations of each sample were tested.

DNA sample ID	Population	Sequence
2	Swan002	$(GCC)_3(GTC)_1$
14	Mor1	$(GCC)_3(GTC)_0$
119	Crete	$(GCC)_4(GTC)_0$

DNA melt profiles were then examined to assess the capacity to identify different products. Melt profiles generated with this combination of samples were tested with the genotyping software provided by Corbett, although this software is specifically designed for use with SNPs rather than SSRs.

6.3 Results

6.3.1 ***Roche Lightcycler real-time PCR.*** Of the 12 primers tested with the Lightcycler, all amplified a PCR product. Clarity of melt peaks was considered to be good for PCR products from all primers with the exception of OL10HO2. Melt peaks were considered to have good clarity if they had a sufficient change in fluorescence to produce well differentiated peaks. PCR products from the primer OL10H02 suffered from artefacts of non-specific products that resulted in multiple peaks of varying heights masking the signal from the desired product in the melt profile. From the remaining primer sets, four amplified PCR products had little sign of primer dimer amplification, and also did not produce more than 2 melt peaks per individual sample (NA10D09, NA12A02, NA14E08 and NI4B10). These were considered suitable for further analysis and two primer pairs from this set were focused on. NA14E08 was selected on the basis of its reported tri-nucleotide repeat, potentially providing greater differences in product length between SSR alleles. NA12A02 was selected as a di-nucleotide SSR for further investigation of the differentiation capacity of the Lightcycler. With additional resources NA10D09 and NI4B10 would also have been investigated in greater depth, but it was felt sensible to study two primers first in detail.

Table 6.4 Amplification of *H. incana* DNA with a range of *Brassica* primers with the Roche Lightcycler. Reported fields indicate whether a PCR product was observed ('Amplification'), if melt peaks were well differentiated ('Clarity of melt peaks'), and whether at first inspection the melt peaks gave sensible outputs ('Interpretability of melt peaks'). Interpretability was described as 'good' if no greater than two peaks were present and no primer dimers were observed. The presence of primer dimers would reduce the interpretability to 'medium' and additional non-specific products could interfere with the specific melt peaks and would reduce the interpretability to 'poor'. * marks the two primer sets used in further investigation

Primer	Amplification	Clarity of melt peaks	Interpretability of melt peaks
BRMS19	Y	good	poor
OL10H02	Y	poor	poor
NA10D09	Y	good	good
NA12A01	Y	good	medium
NA12A02*	Y	good	good
NA12A07	Y	good	medium
NA14E08*	Y	good	good
NI4B10	Y	good	good
28673	Y	good	medium
70189	Y	good	medium
CCMP6	Y	good	medium
CCMP10	Y	good	medium

6.3.1.1 Amplification of NA14E08. The primer pair for NA14E08 amplified well in the Lightcycler; exponential amplification occurred after approximately 28 cycles (Figure 6.1). Typically, products were identified by a strong single melting peak (Figure 6.2), generally with no sign of primer dimers when template DNA was present. Figure 6.3 indicates the variability of melting peaks and the lower melting peak produced when a primer dimer or non-specific product was formed (75 °C). Melting temperatures of the PCR products for this locus ranged from 85-89°C (Figure 6.3). In some cases two peaks were observed for a single DNA template (data not shown).

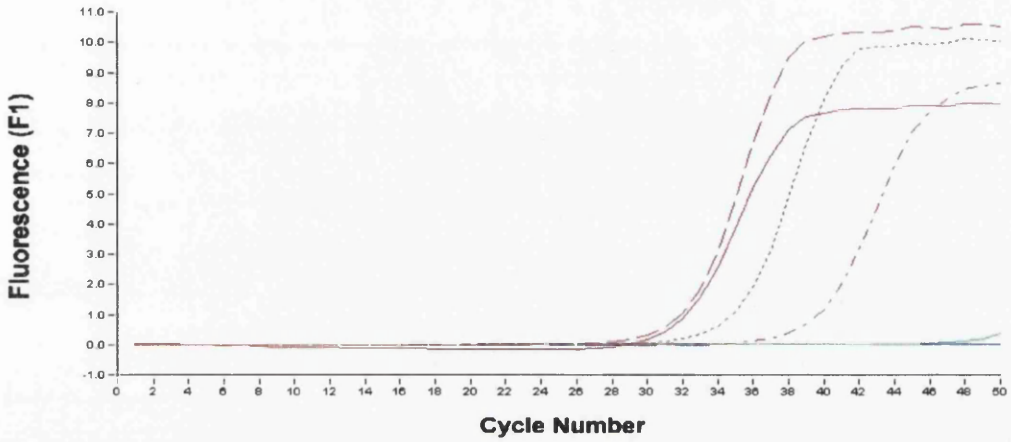


Figure 6.1 Amplification of *H. incana* template DNA dilution series using primers NA14E08. This output shows the progression of the PCR reaction using *H. incana* template from Swansea. In a 10 µl reaction, DNA template quantities were 10 ng (red solid line), 1 ng (red long dash), 0.1 ng (red dotted line), 0.01 ng (red dash and dot line), 0.001 ng (green solid line) and an H_2O blank (blue solid line). Samples containing 10 ng and 1 ng template DNA began amplification after approximately the same number of cycles (28). Each progressive dilution resulted in the delay of the exponential phase of the reaction and the water blank did not produce any amplification product.

The dilution series shows amplification of each 10x dilution of DNA template (Figure 6.1). In theory a 10x dilution would lead to slower amplification in the early stages of the reaction. This would be observed by the delay of the visible exponential phase of the amplification by 2 cycles. In practice this does not always happen, as shown in Figure 6.1. Samples 1 and 2, containing 10 ng and 1 ng template DNA respectively, both show amplification after approximately 28 cycles. Samples 3 and 4, containing 0.1 ng and 0.01 ng template DNA, can be seen to amplify after 30 and 36 cycles respectively. Any amplification that does not start after approx 35 cycles may therefore indicate poor template or reaction conditions, and should be checked carefully that amplification has not occurred as a result of contamination. For this reaction, optimal starting quantity of template DNA was 1 ng in a 10 µl reaction volume. This was the quantity that gave both the earliest initiation of the exponential phase of the reaction and greatest final

concentration of PCR product. A concentration of template DNA greater than 1 ng/µl did not increase the amplification rate and resulted in lower final product concentration of PCR product. This is shown in Figure 6.1 by the lower fluorescence produced (related to DNA concentration) when 10 ng of template DNA was used. 0.1 ng/µl began its amplification slightly later than 1 ng/µl but produced a similar final concentration of product. Lower starting template concentrations than 0.1 ng/µl did produce amplification products, but product quality may be questionable due to late amplification.

Melting peaks are generated after the completion of the melt cycle. The output can then be customized to display the whole melt, or peak(s) of interest. Melt peaks can be estimated manually by positioning a vertical line on the apex of the peak (Figure 6.2) or can be estimated automatically on an individual basis if the number of peaks has been specified. In the example above the melting peak was estimated manually as 87.7°C. The single well defined peak indicated generation of a single PCR product.

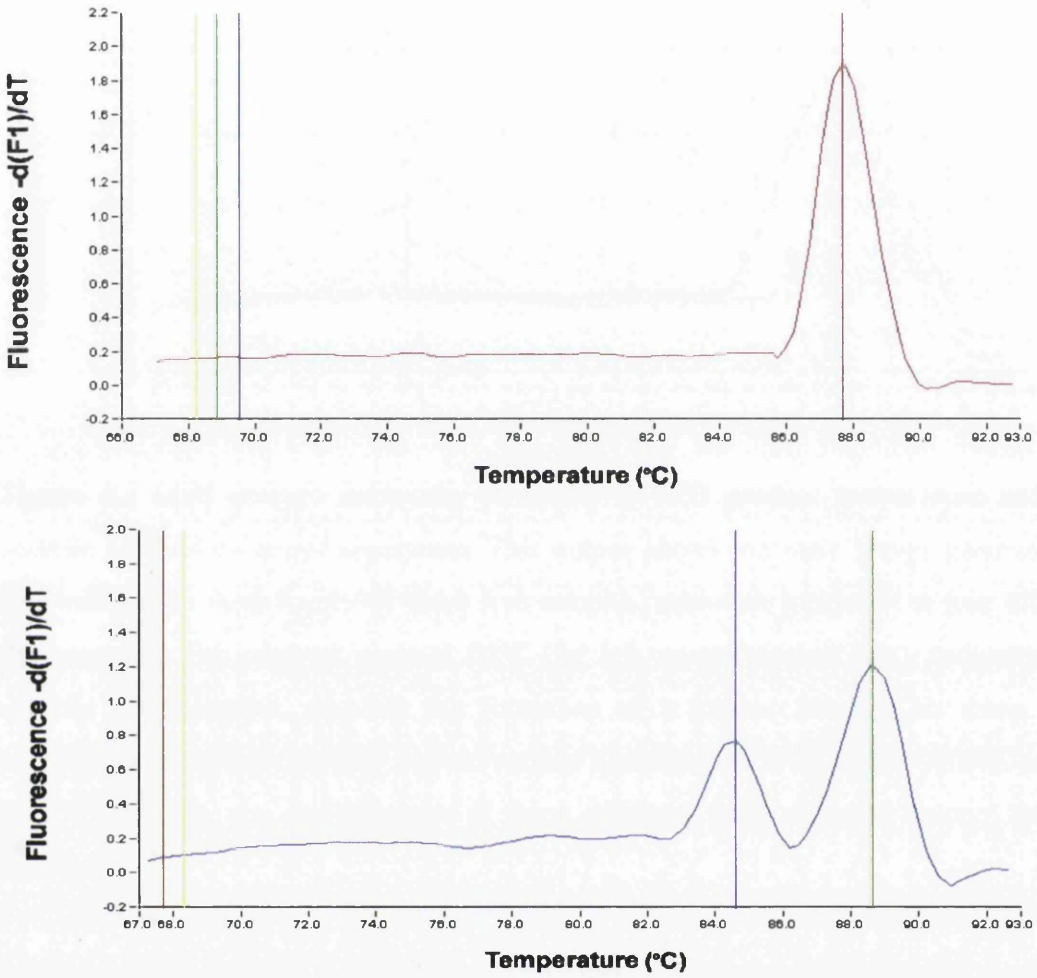


Figure 6.2 a.b Graphical output of Roche Lightcycler showing a single PCR product of a single sample (a) and a multiple product (b). Output (a), of a sample of *H. incana* from Swansea, demonstrates the melt curve generated for a single sample with a single allele. The melt peak at 87.7 °C, (determined by positioning the vertical bar on the apex of the peak) was very well differentiated allowing easy recognition. Output (b), of a sample of *H. incana* from Swansea, demonstrates the melt curve generated for a possible heterozygote.

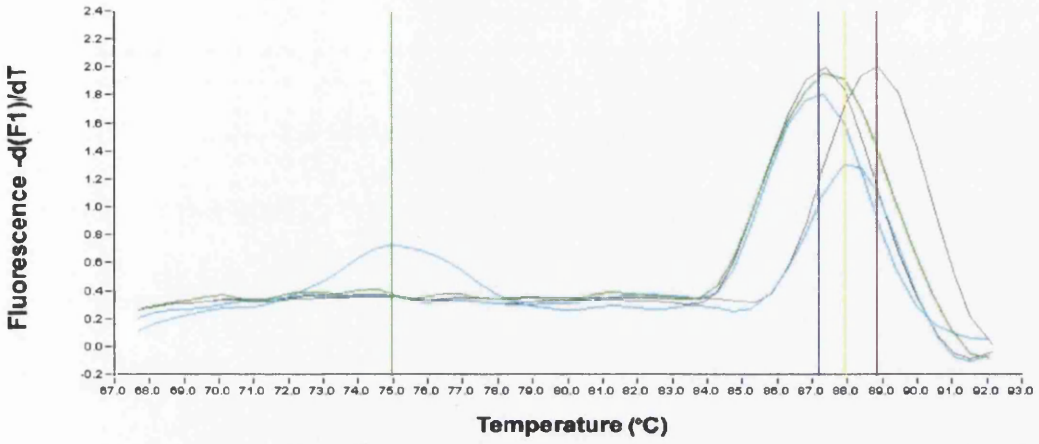
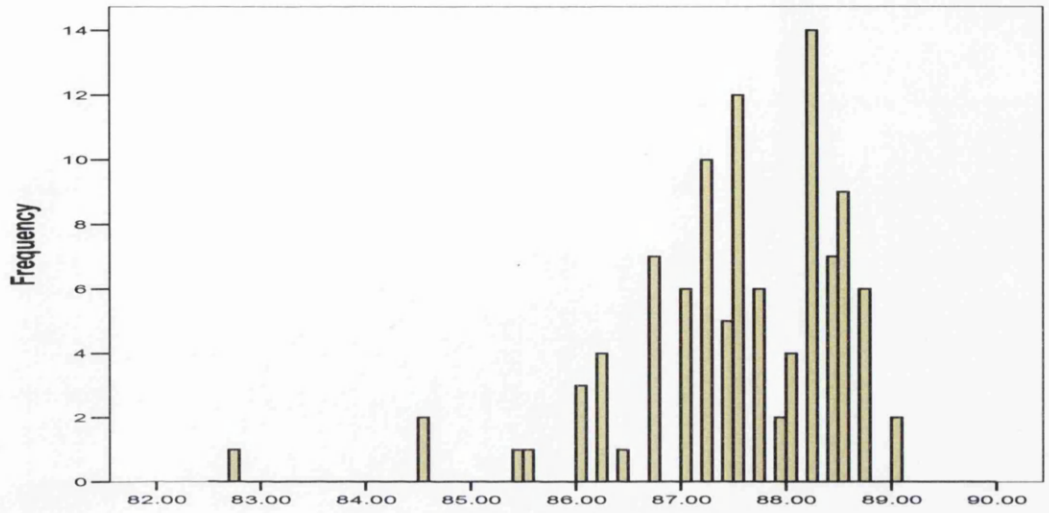
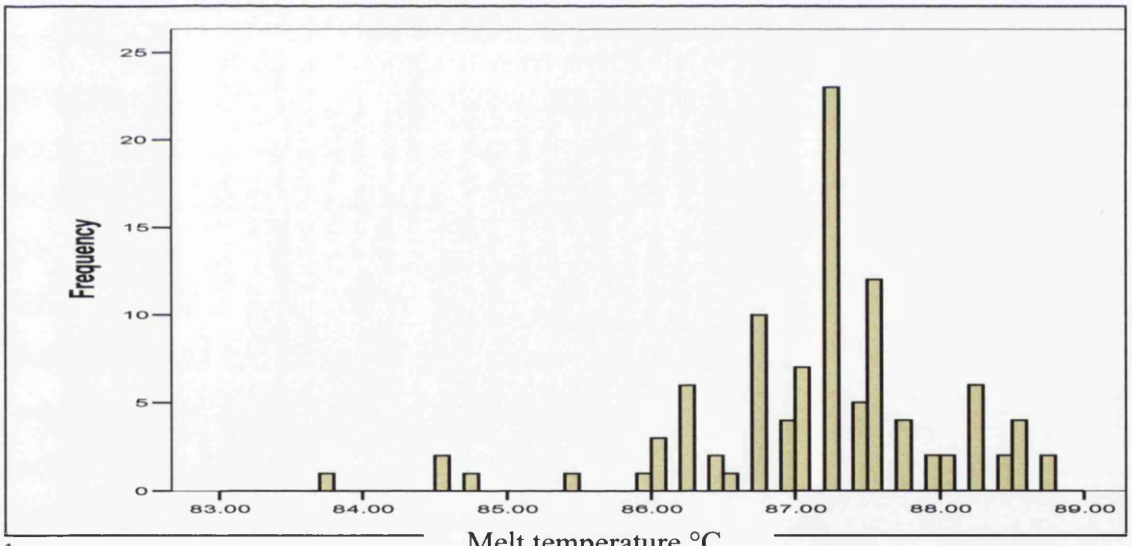


Figure 6.3 Melt analysis indicating variability in PCR product temperature and non-specific product melting temperature. This output shows the melt curves generated for five individuals from Spain. Of these five samples, peaks are produced at four different temperatures. The smallest peak at 75°C (far left, green vertical line), indicates non-specific amplification, possibly the formation of a primer dimer. The three peaks highlighted by the blue, yellow and red vertical lines (blue 87.2°C, yellow 88.0°C and red 88.9°C) indicate the amplification of three different SSR products among the five samples.

Melting peaks are useful in identifying or differentiating between different DNA products and non-specific amplicons. In figure 6.3, four peaks' temperatures are clearly visible among the five samples. The lowest peak at 75°C represents a primer dimer or other non-specific product and can be discounted. The three remaining peaks at 87.2°C, 88.0°C and 88.9°C represent three alleles of the microsatellite locus. Products with melt temperatures 87.2°C and 88.9°C were sequenced and SSR sequences were found to be $(gcc)_3(gtc)_1$ and $(gcc)_5(gtc)_1$, respectively.



a.



b.

Figure 6.4 Frequency of DNA melting temperature across 5 populations of *H. incana* for primer NA14E08 for 106 samples. 6a displays the raw unadjusted temperature data as recorded from the Lightcycler, temperature data in 6b has been modified (i.e. the melt temperature of the standard sample was adjusted, by addition or subtraction, to be the same for all PCR runs. Based on this adjustment, all other samples in the same PCR batch were adjusted by the same magnitude - on the assumption that the temperature shift was linear). The aim of this modification was to attempt to standardize temperatures and remove the fluctuations between successive PCR runs.

The range of DNA melting temperatures for primer NA14E08 is indicated in Figure 6.4. The aim of this graph was to identify if there were distinct boundaries between peaks of melting temperatures – that could possibly be an identification tool for different alleles. As can be seen in the chart there are peaks and troughs in frequency across the melting temperatures, but the differentiation between peaks is not clear enough to state that different peaks are indeed from different alleles. The standardized temperature data covers a narrower range than the raw data (83.7-88.7°C compared to 82.7-89.1), but peaks in frequency follow a similar pattern for both data sets and unfortunately from this chart it is not straightforward to identify if changes in DNA melt temperature relate to different alleles.

The repeatability of the melt peak temperature being produced by the same DNA template was tested by recording the melt peak for samples that had been PCR amplified multiple times (Table 6.5, 6.6). These repeated individuals were usually only included as single samples within each PCR batch, with the exception of the investigation into template concentration (where six samples were included, five from the same template DNA and an H_2O blank). Temperature shifts of up to 1°C were observed between different PCR batches for the same DNA template with the same primer (data not shown). Although not conclusive, due to a small sample size, shifts in temperature would seem to affect all samples in a batch with similar magnitude (Table 6.6) With a standard sample (or more than one standard) included in each batch it may be possible to apply a correction factor to standardize results between batches based on the observed shift in melting temperature for the known standard.

Table 6.5 Summary data for individual DNA samples that were PCR amplified with primer NA14E08 multiple times. This information gives an indication of the variation in melt peaks between separate PCR batches. Temperatures listed for mean melt temperature are from unadjusted (raw) data. Where available, sequence data for SSR repeats have been included.

DNA	Mean melt temp	SD	N	Sequence
1	87.98	0.44	5	(gcc)3(gtc)1
2	87.40	0.42	2	(gcc)3(gtc)1
3	87.55	0.35	2	(gcc)3(gtc)1
4	87.63	0.21	3	(gcc)3(gtc)1
11	87.60	0.20	3	-
14	86.92	0.31	2	(gcc)3(gtc)0
17	87.73	0.25	3	-
18	88.55	0.35	2	-
19	87.70	0.14	2	-
20	86.65	0.35	2	-
23	87.97	0.19	2	-
100	88.53	0.23	11	-
101	88.30	0.00	2	-
102	88.40	0.00	2	-
103	88.60	0.00	2	-
104	88.70	0.00	2	-

Table 6.6 Melt peak temperatures indicating temperature shifts between samples for successive PCR amplification batches (figure in parenthesis indicate temperature shift). For DNA samples 1-19 a consistent temperature shift was observed for both samples ± 0.1°C between batches. Samples 100-104 had consistent melt peaks for all samples between the two batches. The run including a gradient in template dilution shows the variation observed that is influenced by template concentration with the same amplification.

	Melt temperatures variation for different sets of PCRs				
DNA ID	13/06/2006	24/07/2006	28/07/2006	21/08/2006	22/08/2006
1 (10 ng)	-	-	-	-	88.70
1 (1 ng)	-	-	-	-	88.80
1 (0.1ng)	-	-	-	-	88.90
1 (0.01ng)	-	-	-	-	88.40
1 (0.001ng)	-	-	-	-	88.60
H_2O (blank)	-	-	-	-	77.50
1	87.80	87.30 (-0.50)	-	-	-
3	87.80	87.30 (-0.50)	-	-	-
4	87.80	87.40 (-0.40)	-	-	-
11	87.80	87.40 (-0.40)	-	-	-
17	88.00	87.70 (-0.3)	-	-	-
19	86.90	86.40 (-0.5)	-	-	-
100	-	-	88.5	88.5 (0.0)	-
101	-	-	88.3	88.3 (0.0)	-
102	-	-	88.4	88.4 (0.0)	-
103	-	-	88.6	88.6 (0.0)	-
104	-	-	88.7	88.7 (0.0)	-

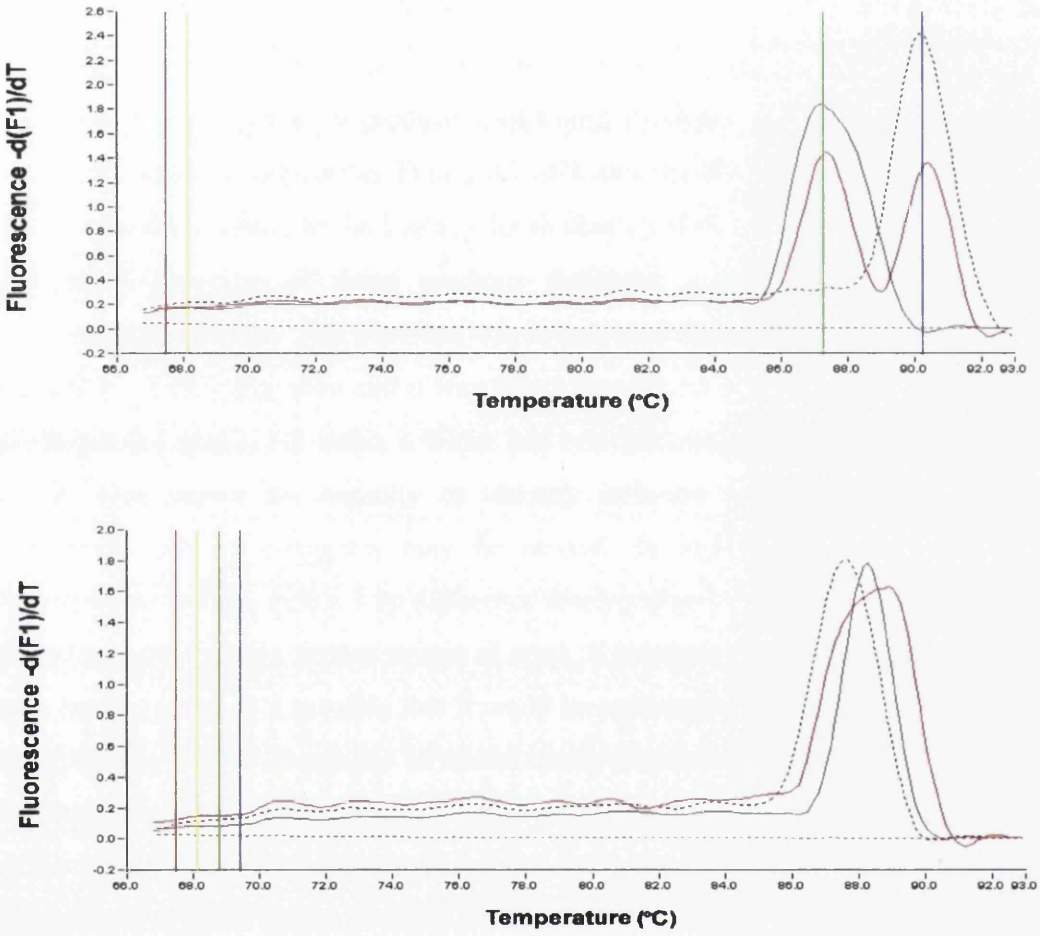
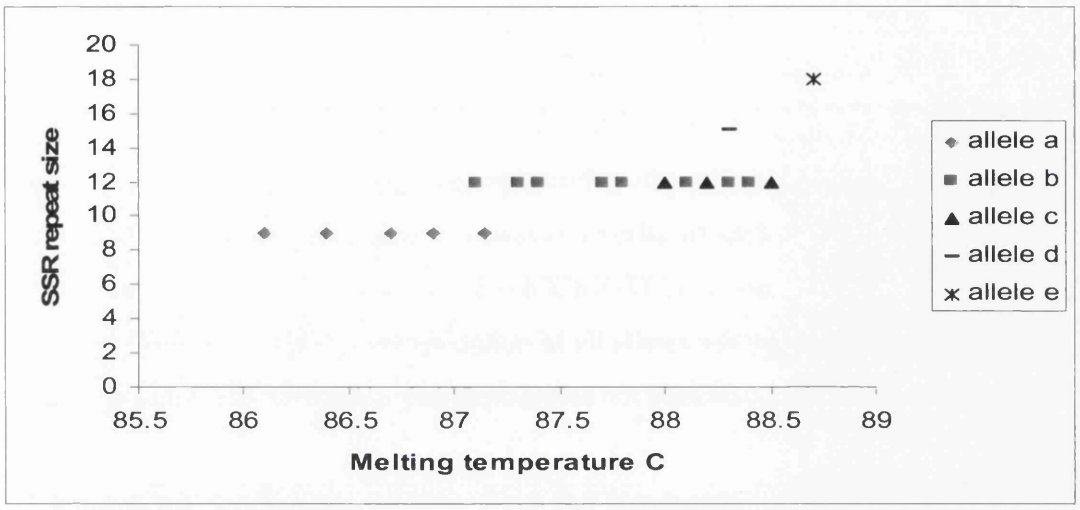


Figure 6.5 a,b (a) Differentiation of heterozygous SSR products with a 9 base pair difference in the alleles. The melt peaks (87.3 °C and 90.2 °C) are for alleles $(GCC)_3(GTC)_0$ and $(GCC)_5(GTC)_1$. Individuals are from Mohammedia, Morocco (purple solid line), and Bajar, Spain (purple dashed line) respectively. The red solid line represents the melt profile of a mixed sample of template DNA from Morocco and Spain (simulating a heterozygote product). **(b)** Differentiation of heterozygous SSR products with a 3 base pair difference in the alleles. The melt peaks (87.5 °C and 88.1 °C) are for alleles $(GCC)_3(GTC)_0$ and $(GCC)_3(GTC)_1$. The individuals are from Mohammedia, Morocco (purple dashed line) and Swansea, UK (purple solid line) respectively. The red solid line represents the melt profile of a mixed sample of template DNA from Morocco and Swansea (simulating a heterozygote product).

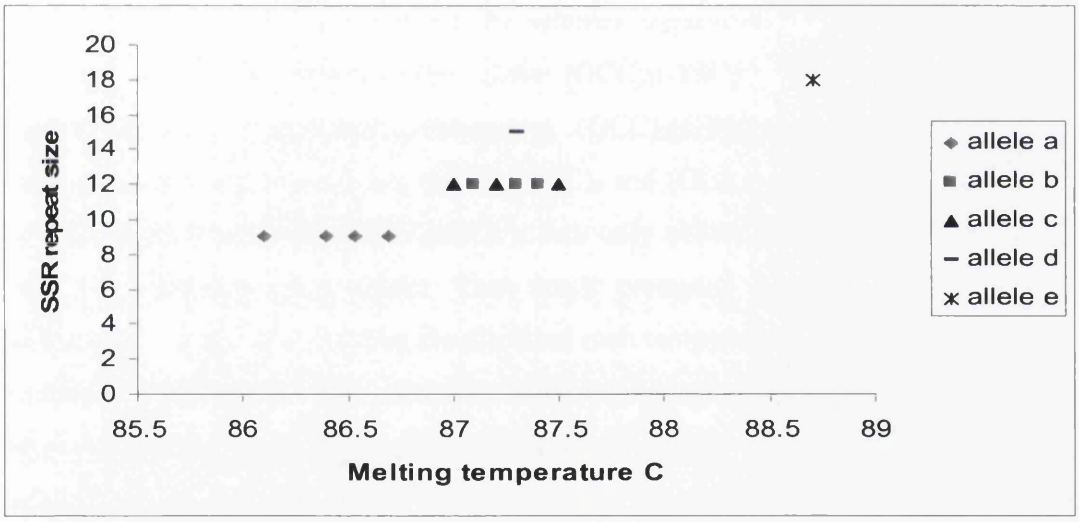
To determine the ability of the Roche Lightcycler to detect heterozygous products, PCR conditions were set up to simulate the generation of a heterozygous product. Two DNA templates were selected to produce contrasting products. These were PCR amplified separately and in combination. Figure 6.5 indicates the DNA melting curve of two such products and the ability of the Lightcycler to identify these two products in the combined sample. Sequencing of these products indicated a nine base difference in the microsatellite product. This protocol was repeated to identify if there was a limit to the capability of differentiation and it was found that a 3 bp difference did not produce two differentiated peaks, but rather a wider and less differentiated single melt peak (Figure 6.5b). This shows the capacity to identify different homozygous alleles of 3 bp difference, but heterozygotes may be missed. In addition, the melt peak for the heterozygote sample with a 3 bp difference displayed an increase in melting temperature, therefore introducing a further source of error. If a sample such as this was not identified as a heterozygote, it is possible that it could be incorrectly labeled as a homozygote of a different allele. The occurrence of stutter bands was not identified using this method as homozygote products produced a single well defined peak. It is possible that a stutter effect would result in a wider melting peak and may also cause a slight shift in melting temperature (similar to that seen for a heterozygote with a small difference in base pairs, as seen in Figure 6.5b).

Table 6.7 Melting peaks matched with sequenced samples using standardized temperatures. Mean melting temperatures were calculated using temperature data for all individuals where their allele had been identified.

Sequence	Mean melt temp	SD	N
$(GCC)_3(GTC)_0$	86.40	0.30	3
$(GCC)_3(GTC)_1$	87.73	0.40	12
$(GCC)_4(GTC)_0$	88.23	0.20	4
$(GCC)_5(GTC)_0$	88.30	0.00	2
$(GCC)_5(GTC)_1$	88.70	0.00	1



a.



b.

Figure 6.6 a, b Distribution of melting temperature for the known alleles of NA14E08 in *H. incana*. These two charts indicate the range of temperatures for the following alleles; a $(GCC)_3(GTC)_0$, b $(GCC)_3(GTC)_1$, c $(GCC)_4(GTC)_0$, d $(GCC)_5(GTC)_0$, e $(GCC)_5(GTC)_1$. Figure 6a is based on the raw melt temperature output of the Lightcycler, figure 6.6b is based on data where a correction factor (based on the inclusion of a standard sample for each PCR run) has been applied to standardize temperatures between runs. A visual comparison between the two charts indicates that the application of the correction factor has reduced the variation in observed melting temperature.

Figure 6.6 was produced to investigate the possibility of matching DNA melt temperatures with alleles for genotyping purposes. From the samples shown, allele a and allele e $(GCC)_3(GTC)_0$, $(GCC)_5(GCT)_1$ had sufficiently different melting temperatures that they did not overlap melting temperatures for other alleles, indicating the potential to identify these alleles by melt peaks. However, overlap of melting temperatures for alleles b $(GCC)_3(GTC)_1$, c $(GCC)_4(GTC)_0$, d $(GCC)_5(GTC)_0$ in the range of 87-87.6 °C, was observed. This prevented the identification of all alleles across the populations using melt peaks and as a result meant that genotyping was not possible.

An interesting observation resulting from the sequencing of a range of samples that covered the range of melt peaks was the distribution of alleles among the populations. Initial observations indicate that for the samples sequenced, little allele variation was observed within populations. The allele $(GCC)_3(GTC)_0$ was only observed from individuals from Mohammedia (Morroco), $(GCC)_3(GTC)_1$ was only observed from individuals from Swansea (UK), $(GCC)_4(GTC)_0$ and $(GCC)_5(GTC)_0$ were only observed in individuals from Crete. $(GCC)_5(GCT)_1$ was only observed in one sequence and was from an individual from Spain. This result prompted the plotting of melt peak temperature by population (using standardized melt temperatures, Figure 6.7). Melt peaks indicated a pattern of little variation within populations; the majority of variation occurred between populations (Kruskal- Wallis $P<0.000$, $N=101$). To demonstrate the two clusters of melt peaks seen from individuals from Saida (Algeria) the temperature data were classified into two groups based on the observed melting temperature of the PCR product. This chart indicates that population differences were present and can be visualized based on the DNA melting temperature of SSRs at NA14E08. Pairwise comparisons indicate significant differences in melt peak temperatures between many of the populations studied. Significant differences were also found between the two clusters within the Algerian population (Table 6.9).

Table 6.8 Results of sequencing for samples that cover the observed range of DNA melting temperatures gives an indication of the clustering of alleles by populations. No sequence data were available for individuals from Algeria.

DNA sample ID	Population	GCC repeat	Standardized temp
14	Morocco 1	(gcc)3(gtc)0	86.7
20	Morocco1	(gcc)3(gtc)0	86.4
60	Morocco 1	(gcc)3(gtc)0	86.1
1	Swansea 2	(gcc)3(gtc)1	87.3
2	Swansea 2	(gcc)3(gtc)1	87.1
3	Swansea 2	(gcc)3(gtc)1	87.3
4	Swansea 2	(gcc)3(gtc)1	87.4
119	Crete	(gcc)4(gtc)0	87
120	Crete	(gcc)4(gtc)0	87.5
125	Crete	(gcc)4(gtc)0	87.2
129	Crete	(gcc)4(gtc)0	87.2
118	Crete	(gcc)5(gtc)0	87.3
121	Crete	(gcc)5(gtc)0	87.3
52	Spain	(gcc)5(gtc)1	88.7

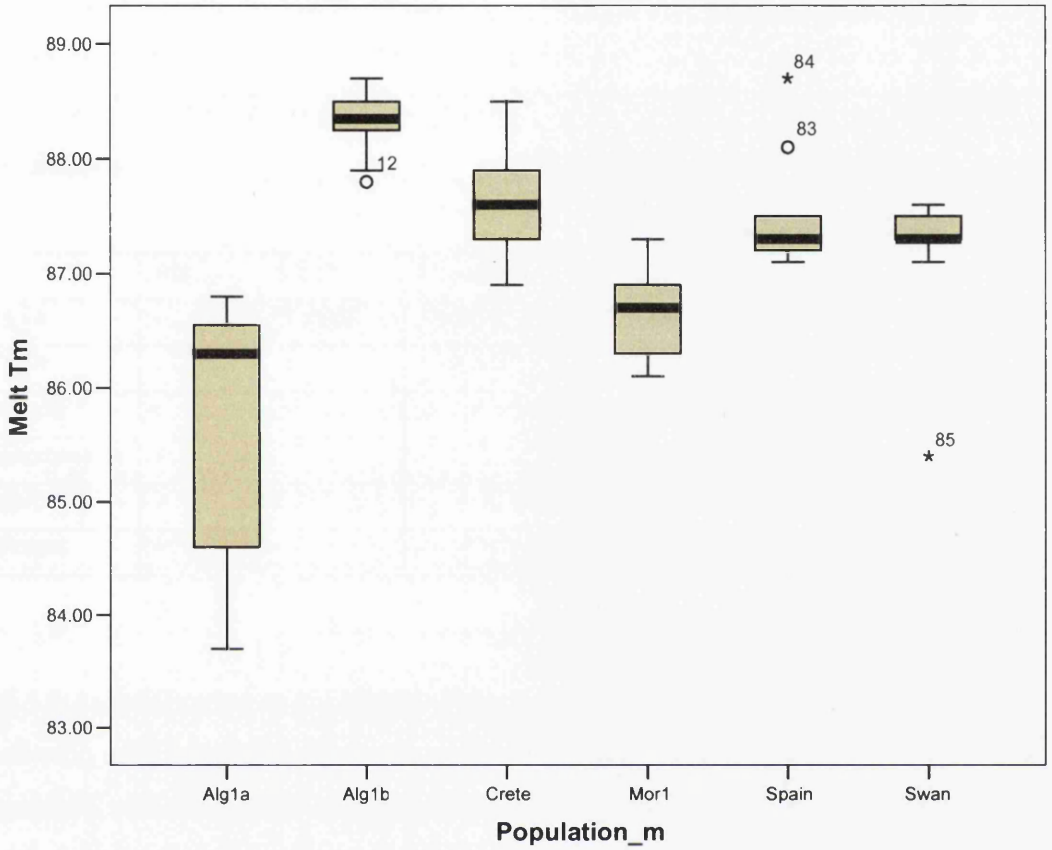


Figure 6.7 Distribution of DNA melt temperatures for NA14E08 across the populations studied. Individuals from Algeria were split into two groups based on the fact that two distinct sets of melting temperatures were observed among these individuals (the products were not identified by sequencing). The populations tested contained the following number of individuals; Saida (Algeria) n=23 (split into a n=11, b n=12), Elounda (Crete) n=17, Mohammedia (Morocco) n=22, Bejar (Spain) n=22, Swansea (UK) n=17. Four individuals from three populations contained individuals that were considered outliers based on their melt temperatures (indicated by; o and *). With the exception of Algeria, all other populations had a narrow, within population, range of melting temperatures.

Table 6.9 *P*-values of Mann-Whitney U pairwise comparison of observed melt peaks across all populations tested with NA14E08. A sequential Bonferroni correction was performed to allow for multiple comparisons. * indicates significant differences between populations.

	Alg 1a	Alg 1b	Elounda	Mohammedia	Bejar	Swansea
Alg 1a	-	$*1.5x10^{-6}$	$*9.3x10^{-8}$	0.021	$*1.0x10^{-8}$	$*4.2x10^{-6}$
Alg 1b	-	-	$*1.0x10^{-4}$	$*3.6x10^{-6}$	$*1.4x10^{-6}$	$*3.9x10^{-8}$
Elounda	-	-	-	$*3.8x10^{-9}$	0.098	0.099
Mohammedia	-	-	-	-	$*5.2x10^{-8}$	$*5.7x10^{-7}$
Bejar	-	-	-	-	-	0.59
Swansea	-	-	-	-	-	-

6.3.1.2 Amplification of NA10A02. Primers for NA10A02 also PCR amplified a product well with the Lightcycler. However, amplification curves often appeared wider than those generated with NA14E08 or alternatively transitions were often observed at some point on the melt curve (Figure 6.8). Variation observed in melting peaks in Figure 6.8 may be a result of allele variation. However, sequence data is only currently available for the sample (bold red line) with the melt peak at 83.4 °C (SSR repeat $(CT)_6$) so it is not possible to determine if melt peaks at 83.4 °C and 84.1 °C represent different or the same SSR alleles. These results suggested that differences between melt profiles of NA10A02 were poorly defined and were perhaps caused by heterozygotes. These results indicated the limitations of using the Roche Lightcycler with SYBR green to identify or differentiate between SSR alleles with very similar sequences. As NA12A02 represents a di-nucleotide SSR the difficulties observed from studying NA14E08 (tri-nucleotide SSR) were expected to be greater for this primer.

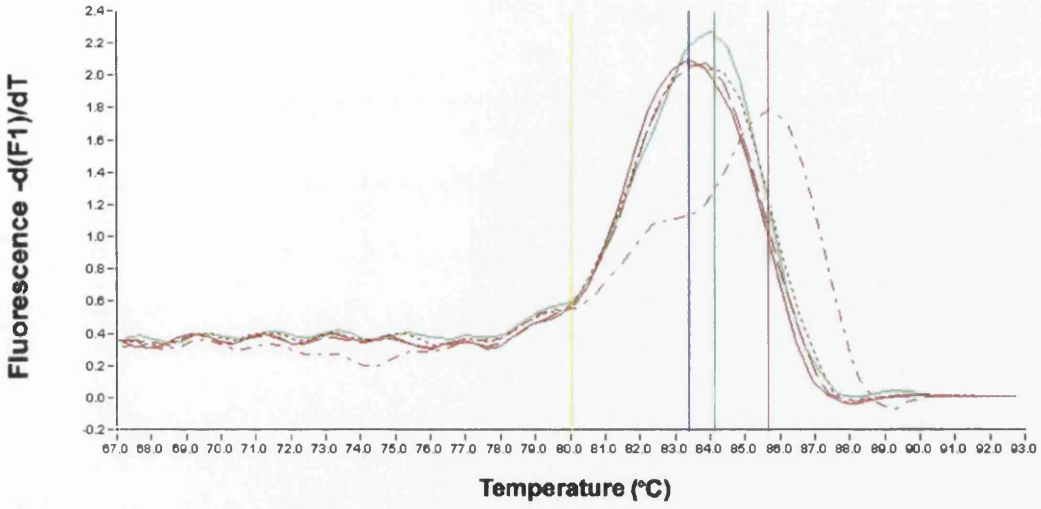
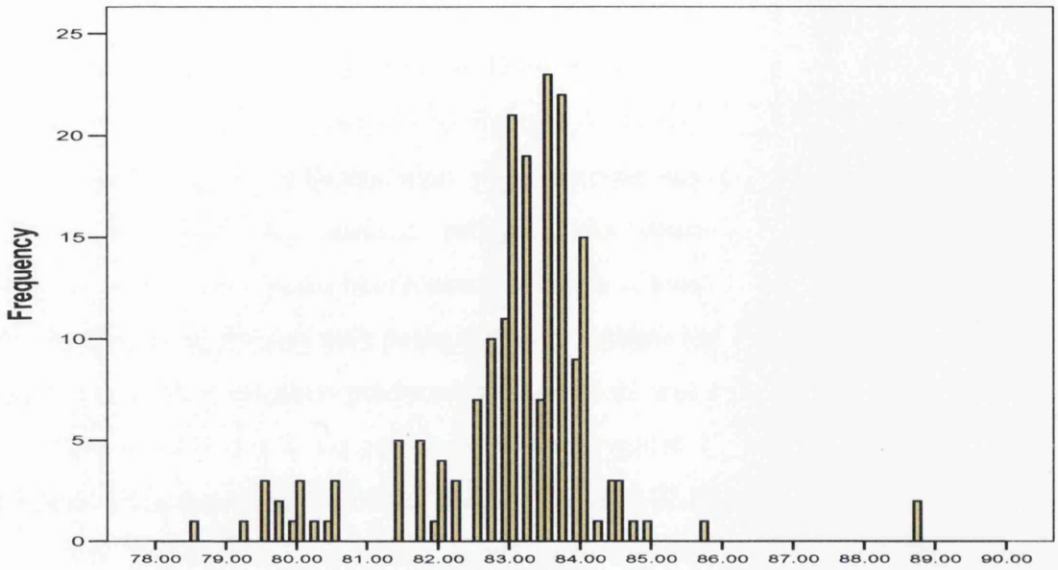
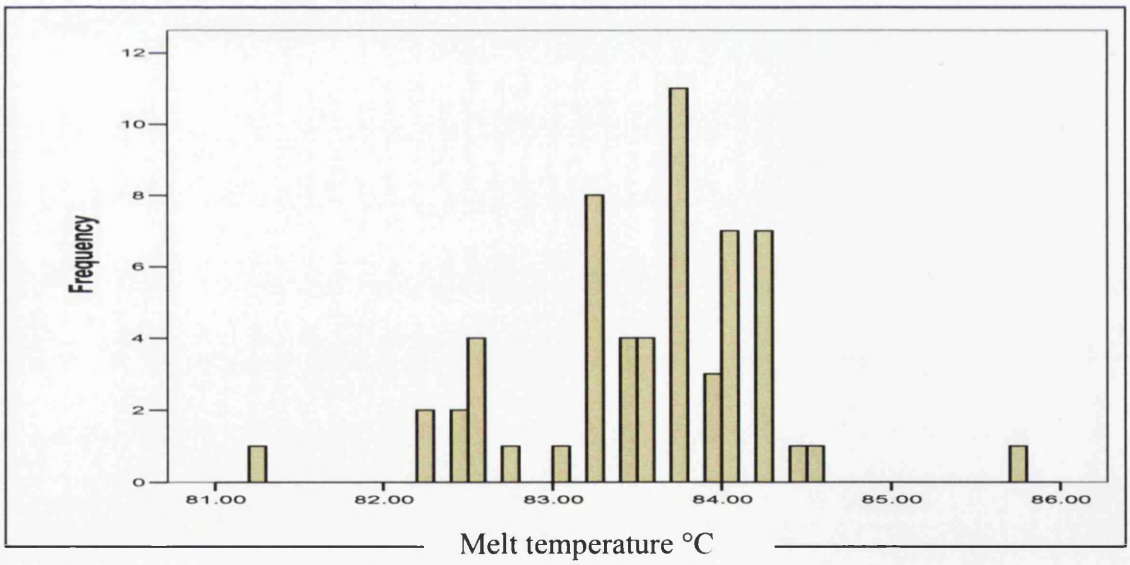


Figure 6.8 Lightcycler output generated from amplification of *H. incana* DNA with NA10A02. This output shows the melt curves for five different samples from Swansea with potentially different products, based on the variation observed of each melt peak. Three main melting peaks were apparent, 83.4 °C, 84.1 °C and 85.7 °C. Two samples produced wide melt peaks intermediate between 83.4 °C and 84.1 °C (long dashed red line and short dashed red line) hinting at a possible heterozygote product. One sample (dot and dash red line) produced a major peak at 85.7 °C with an obvious transition in the melt profile at 83.4 °C strongly suggesting a heterozygous product.



a.



b.

Figure 6.9 a, b Frequency of DNA melting temperature across 5 populations of *H. incana* for primer NA12A02. Figure 6.9a represents raw non-manipulated melting peak data taken directly from the Lightcycler, for 190 samples. Figure 6.9b represents standardised temperatures for 58 individuals where a standard DNA sample was included for each successive PCR batch.

Figure 6.9 indicates the frequency of DNA melt temperatures across the range of populations studied. In comparison to figure 6.4, displaying the same information for primer NA14E08, less differentiation was observed across the range, increasing the difficulty of identifying specific products. As observed for primer NA14E08, standardization of melt peaks has reduced the range of temperatures observed, but has not aided in separating obvious melt peaks that may indicate the presence of different alleles. Sequencing of PCR products produced by NA12A02 was not as successful as those for NA14E08, possibly due to the presence of heterozygotes. It was not possible to produce melt peak temperatures for different alleles of NA12A02. It is also possible there is very little allelic variation for this locus, as indicated by the lack of observed variation in product melt temperatures. Melt temperatures varied by up to 5 °C for different alleles at NA14E08 in contrast to NA12A02 where the majority of melt temperatures were all within 2 °C of each other.

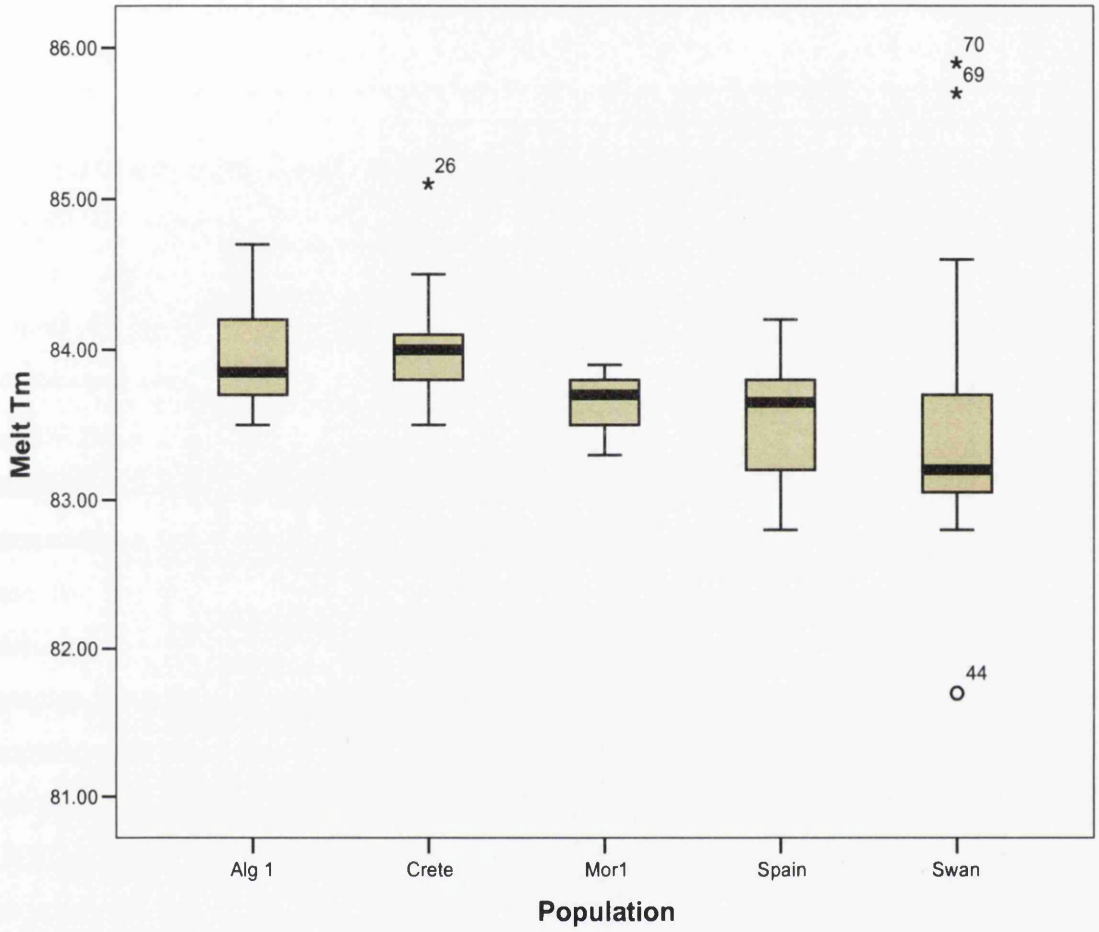


Figure 6.10 Distribution of DNA melt temperatures across the 5 populations studied for primer NA12A02. The populations tested contained the following number of individuals; Saida (Algeria) n=8, Elounda (Crete) n=14, Mohammedia (Morocco) n=9, Bejar (Spain) n=5, Swansea (UK) n=22.

The distribution of melt peaks across the 5 populations for NA12A02 (Figure 10) was not as well differentiated as observed for NA14E08 (Figure 6.7), although a Kruskal-Wallis test indicated that significant differences were present P=0.001. A Mann-Whitney U pairwise comparison with sequential Bonferroni correction applied indicated that the only significant differences occurred between Swansea (UK) and Elounda (Crete) $P= 3.9 \times 10^{-4}$.

6.3.2 ***Corbett real time PCR.*** The Corbett Rotorgene performs high resolution melt analysis using different dye chemistry from the Lightcycler and using plastic reaction vessels compared to the glass capillaries of the LightCycler. Changes were required to the PCR protocols used with the LightCycler to successfully amplify products with the Corbett, an extension in time for each stage of the reaction as detailed in the methodology provided greater success in product amplification. As well as producing graphical outputs similar to the melt curves generated by the Lightcycler, the Rotorgene 6000 uses normalized melt curves (Figure 6.11) to differentiate between products. At first glance this output is not as easy to interpret as the melt peaks, but it is possible to identify subtle differences between x axis temperature shifts and melt curve shape variation. Figure 6.11 demonstrates the output when heterozygous and homozygous products are present; in this case for the loci NA14EO8. In this example PCR reactions and melt analysis were performed with samples that had been previously sequenced (Table 6.3). Therefore, samples were hand picked based on their known SSR repeat length. The purpose of this amplification was to determine if a 3 bp difference, especially for a heterozygous sample, was distinguishable. Three samples were run in duplicate; sample 1 contained the allele $(GCC)_3(GTC)_0$, sample 2 contained the allele $(GCC)_3(GTC)_1$, sample 3 was a mixture of the previous 2 templates – simulating a heterozygous sample.

The three base pair difference is visible by an x axis shift in melt curves (Figure 6.11). The red and yellow curves (far right) represent the $(GCC)_3(GTC)_1$ allele, while the blue and red curves (far left) represent the $(GCC)_3(GTC)_0$ allele. The pink and green curves represent the heterozygous product. The heterozygous curve has a characteristic dip in fluorescence prior to the main decrease in fluorescence. This variation in melt curve shape allows for easy recognition of the heterozygous product in this case.

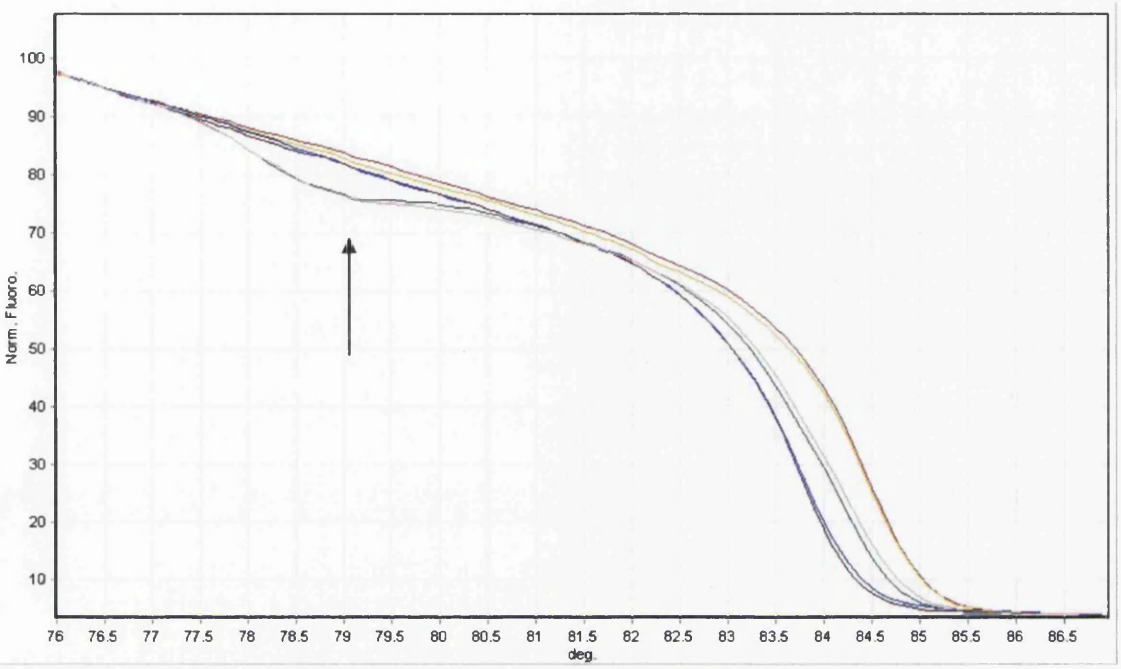


Figure 6.11 Normalized melt curve indicating identification of heterozygote product by the Corbett Rotorgene. In this example three reactions were carried out in duplicate. The Corbett Rotorgene has clearly differentiated between samples containing alleles with a 3 bp difference, and a mixed sample simulating a heterozygote SSR with 3 bp difference (arrow indicates characteristic dip in melt curve indicating a heterozygote). The blue and purple curves have been generated by samples from Mohammedia (Morocco), containing the $(GCC)_3(GTC)_0$ allele. The red and yellow curves have been generated by samples from Swansea (UK) containing the $(GCC)_3(GTC)_1$ allele. The pink and green curves have been generated by samples containing a mixture of both $(GCC)_3(GTC)_0$ and $(GCC)_3(GTC)_1$ alleles.

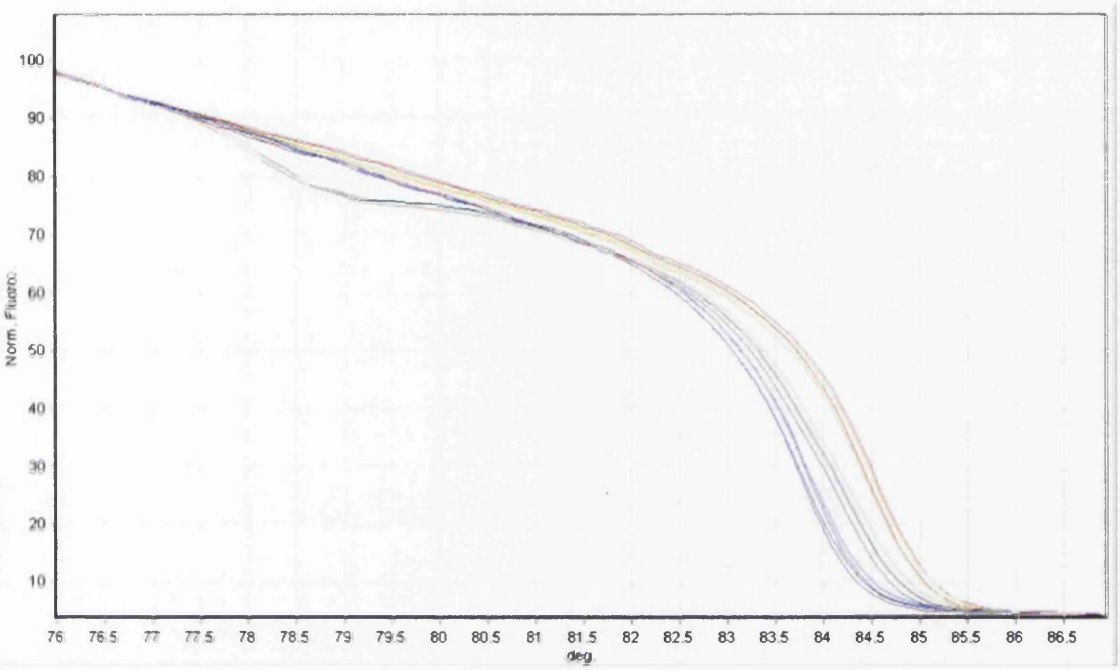


Figure 6.12 Superimposed melt curve of two separate PCR runs, with the same samples as displayed in Figure 6.11. This figure demonstrates the repeatability of DNA melting curve shapes produced by the Rotogene for the same sample between different PCR runs. The output indicates only a very slight x-axis shift in temperature within the same samples between the two PCR reactions. However, the different products can clearly be differentiated between and within the two runs. The colour or the curves in this figure match those from figure 6.11. The blue and purple curves have been generated by samples from Mohammedia (Morocco), containing the $(GCC)_3(GTC)_0$ allele. The red and yellow curves have been generated by samples from Swansea (UK) containing the $(GCC)_3(GTC)_1$ allele. The pink and green curves have been generated by samples containing a mixture of both $(GCC)_3(GTC)_0$ and $(GCC)_3(GTC)_1$ alleles.

The Rotogene allows two methods of genotyping: either single or multiple genotypes can be defined. Genotypes are expressed as a percent similarity to a melt curve of a selected DNA sample from the PCR run. In Table 6.10 genotypes were defined individually and similarity percentages were calculated for each DNA sample. Column 2 displays the percentage similarity of each sample to the melt curve generated for DNA 2. Column 3 displays the percentage similarity to the melt curve generated for DNA 14. Column 4

displays the percentage similarity to the melt curve generated for DNA 119, column 5 displays the percentage similarity to the melt curve generated for the mixed sample containing DNA 2 + 14 and column 6 displays the percentage similarity to the melt curve generated for the mixed sample containing DNA 2 + 119.

Table 6.10 Experimentation with the 'auto genotyping' function supplied with the Corbett Rotorgene software with SSR products. A selection of alleles were amplified; $(GCC)_3(GTC)_1$, $(GCC)_3(GTC)_0$ and $(GCC)_4(GTC)_0$, from DNA sample ID number 2 (Swansea, UK), 14 (Mohammedia, Morocco) and 119 (Crete). 'Synthetic' heterozygotes were also amplified by combining template DNA from samples 2 + 14 and samples 2 + 119, all samples were run in duplicate. The genotyping feature allows you to name one sample from the PCR run with an allele code. From the curve generated by the melt analysis, the software calculates a percentage similarity score for all the other samples in the PCR run (based on the similarity of their melt curve profile to the named sample). This allows the user to ask the question, "of all the samples amplified, how many produce the same allele as sample x". This table shows percentage similarity scores (based on similarity of melt curves) that have been calculated for all combinations of the alleles. The first row of the table indicates the named 'allele' in each case. The first column lists the samples that were PCR amplified. The numbers displayed are the similarity scores of the curve generated for the sample when compared to the curve for the named allele. E.g. in the high-lighted box; when DNA 119 is the named allele, the software calculated that the melt curve for the first DNA 2 sample had 19.2% similarity to that of DNA 119. The reason to produce these comparisons was to investigate the ability of the software to genotype different SSRs and to see if there was a capacity to identify components of a heterozygote when these components had also been amplified in isolation.

Sample	DNA 2	DNA 14	DNA 119	DNA 2 + 14	DNA 2 + 119
DNA 2	95.7	0.05	19.2	2.19	85.6
DNA 2	95.7	0.28	17.2	47.6	75.46
DNA 14	0.16	98.9	0	3.1	0.03
DNA 14	0.09	98.9	0	2.4	0.02
DNA 119	4.27	0.14	19.6	6.8	5.90
DNA 119	3.26	0	19.6	0.1	9.28
DNA 2+14	24.3	2.65	3.49	96.7	16.35
DNA 2+14	44.4	2.62	4.7	96.7	27.15
DNA 2+119	87.6	0.16	25.5	47	83.85
DNA 2+119	56.6	0	39.2	7	83.85
H20	0	0	0	0	0

In column 2 of Table 6.10 similarity scores were calculated for all samples to DNA 2. The first two samples (both consisting only of DNA 2) have been assigned 95.7% similarity (effectively to their own genotype, as here we are asking the software the following question; what similarity does the melt curve generated by DNA 2 have to itself? The expected answer is 100%, but the generated result was 95.7%). The other homozygous samples (DNA 14 and DNA 119) are identified with less than 5% similarity to the genotype of DNA 2. The heterozygous samples that include DNA 2 (DNA 2+14 and DNA 2+119) have between 24.3-87.6% similarities to the genotype defined for DNA 2. Water controls indicated 0% similarity to the genotype. For column 2 all samples containing DNA 2 had higher similarity percentages (to the named genotype of DNA 2) than those that did not contain DNA 2.

In column 3 of Table 6.10, similarity scores were calculated for all samples to DNA 114. DNA 14 showed a high similarity percentage to its own genotype (98.9%), as expected. All other samples show low similarity (<2.65%) including the heterozygous samples that include DNA template 14.

In column 4 of Table 6.10, similarity scores were calculated for all samples to DNA 119. However, the samples that were this genotype only had a similarity percentage of 19.6, even though this was the genotype defined for this PCR product. It was expected that the similarity percentage for the melt curves, when comparing a sample with itself, would be 100%. The first two samples (consisting of DNA 2) had similarity scores of 19.2 and 17.2 to DNA 19. The mixed samples containing DNA 2 + 119 had the highest percentage similarity scores of 25.5 and 39.2.

In column 5 of Table 6.10, similarity scores for melt curves were calculated for all samples to the mixure of DNA 2+14. Genotyping using the heterozygous product of DNA 2&14 gave similarity percentages of 96.7 for the samples containing this template. Also high similarity scores were found for samples containing DNA 2 and the combined sample of DNA 2 + 119.

In column 6 of Table 6.10, similarity scores for melt curves were calculated for all samples to the mixure of DNA 2+119.Genotyping using the heterozygous product of DNA 2&119 had similarity of 83.85% for the samples that specifically contained these templates, but high similarities were also found for samples that contained DNA 2. It may be possible that a dominant signal was picked up from the melt of DNA 2 and that this was producing the higher similarity scores for those samples containing this DNA. This was also supported by the melt profile for this combination of templates in figure 6.13, as the expected 'heterozygote shaped curve' was not produced. It is possible that in this case a 'null allele' effect is being observed and of the heterozygous mixture only the genotype of DNA 2 ($(GCC)_3(GTC)_1$) was being amplified .

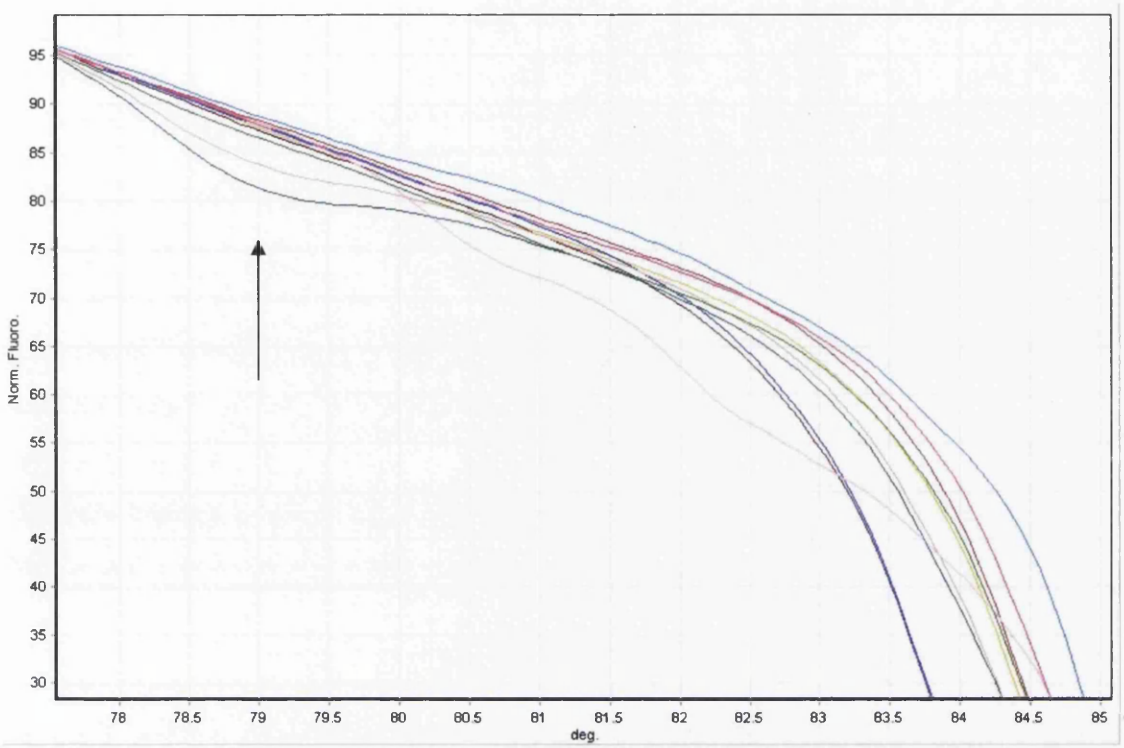


Figure 6.13 This figure is a zoomed in section of the melt curve generated using the DNA templates and combinations described in Table 6.10. In this figure the DNA samples have the following colours and x-axis intersections: DNA 14, purple and dark blue, 83.75°C; DNA 14+2, dark green and peach, 84.3°C; DNA 2, yellow and green, 84.5°C; DNA 2 + 119, pale pink and red, 84.5°C and 84.7°C; DNA 119, dark pink and blue, 84.7°C and 84.8°C. The arrow indicates the characteristic heterozygote dip in the curve for DNA samples 2 +14. For each of the DNA templates run in isolation, DNA 2 $(GCC)_3(GTC)_1$, DNA 14 $(GCC)_3(GTC)_0$ and DNA 119 $(GCC)_4(GTC)_0$, melt curves could be distinguished from each other, the alleles for the curves produced can be identified as they cross the x-axis, in increasing temperature DNA 14, 2 + 14, 2, (2+119) and 119. The combined sample 2 + 119 did not produce the characteristic heterozygote shaped curve: one sample produced a curve similar to DNA 2, while the second sample produced a multiple dipped curve.

The software is not specifically designed to define a mixed product profile as a genotype, which causes difficulty in identifying the components of heterozygous products and in

differentiating between two different heterozygous products. Repeated runs using the Corbett with the samples from Table 6.10 indicated good differentiation between single products, but less successful differentiation of mixed products. PCR products with a 3 bp difference in length could be consistently differentiated. PCR products with no difference in length, (DNA 2 $(GCC)_3(GTC)_1$ and DNA 119 $(GCC)_4(GTC)_0$) could be separated visually by the shape of the melt curve, but the genotype software could not be used to distinguish between them. PCR products amplified from combined samples again produced characteristic shaped melt curves and the melt profile was intermediate to the homozygous samples. However, a greater number of different homozygous alleles and different heterozygous alleles would need to be combined in the same run to understand the variation of melt curves of heterozygous samples in greater detail.

6.4 Discussion

SSR analysis has been achieved by a variety of methods, but it typically performed using either polyacrylamide gels and silver staining or the use of capillary electrophoresis and automated sequencing. Both these methods require the post processing of PCR products and incur additional costs either in terms of labour (polyacrylamide gels) or financially (the use of fluorescently labelled primers and the associated costs of automated sequencing). Real time PCR and DNA melt analysis is a technique that provides the opportunity to distinguish and identify PCR products in a one step process, reducing labour requirements, risk of cross contamination and potentially reducing overall financial costs. SSR polymorphisms have been studied using real time PCR methods such as hybridisation probes (Dietmaier and Hofstädter 2001) and more recently for varietal identification in viticulture (Mackay *et al.* 2008), but the methodology is yet to be developed to address the requirements of population analysis for which fragment analysis on an automated sequencing instrument is the current standard technique. The advances in real time PCR thermocyclers and dye chemistries provide exciting opportunities to study the technique and assess its potential for ecological applications. The sensitivity of DNA melt analysis and its capacity to differentiate between PCR products with small differences in base length or composition may lend itself to applications such as SSR differentiation.

To satisfy the strict requirements for population genotyping certain criteria must be met, and these have been investigated in this study using two real-time PCR instruments capable of DNA melt analysis using two different dye chemistries. Initially, this technique was studied using the Roche Lightcycler and SYBR Green 1; this was then followed up with the Corbett Rotogene using the HRM compatible dye Eva Green.

6.4.1 ***Roche Lightcycler analysis.*** PCR amplification of SSR products using the Roche Lightcycler with well matching primers was found to be excellent. The consistent temperatures between samples due to the rotating drum ensured reliable reaction conditions that can sometimes be an issue with block-based PCR thermocyclers (Linz

1990). Due to rapid temperature ramping rates, 40-50 reaction cycles and melting analysis could be completed within 45 minutes, greatly saving time compared to block-based thermal cyclers. Once initial reactions had been confirmed to contain the product of interest using gel electrophoresis or sequencing, DNA melt profiles could be used to determine the specificity of further reactions without the need for post PCR techniques. The optimal starting DNA template concentration was determined as 1 ng in a 10 µl reaction, which was within the range recommended by the manufacturer. PCR reactions containing this quantity of DNA template entered their logarithmic phase earlier and also resulted in a higher final concentration of PCR product than those starting with other concentrations.

DNA melt analysis of SSRs using the Lightcycler provided consistent patterns of results. With primers that produced clean products, alleles of different lengths could be clearly differentiated to a resolution of 9 BP. Successive reactions containing the same template DNA demonstrated the repeatability of melt profile patterns. However, in some cases melt peaks for the same template were found to vary by up to 1 °C between different PCR runs. This variation is comparable to data reported by Worm *et al.* 2001 in an investigation on 5-methylcytosine in human genomic DNA, who observed variation of up to 0.6 °C for the same samples between successive PCR batches. These may have been due to polymerization of amplicons that may be related to template copy number (Weng *et al.* 2005). This observation indicated the need to include some form of standard sample (i.e. at least one template with known SSR repeat length) between different runs to identify if a shift in melt peaks had occurred. This provided the opportunity to allow a correction factor to be applied to attempt to standardise melt temperatures between successive PCR runs. This technique was used to generate comparable datasets between PCR batches, but caution should be taken when relying on this data, as it is possible that error is introduced at this step. This technique for 'standardisation' between PCR batches has not been previously documented and only serves as an attempt to make a combined dataset more comparable. However, Mackay *et al.* (2008) has advocated the use of including known genotypes in PCR batches to serve as identification aids. Although, in that case the identification of genotypes was based on melt curve shapes rather than

melting peak temperature. I would recommend that greater attention should be given to understanding the variation between melt temperatures from successive PCR batches including performing multiple PCR batch repeats to investigate this area further.

The generation of small quantities of primer dimers did not affect the ability to differentiate between SSR products as their melting temperature was sufficiently low as to not interfere with the main melting peaks (also see Ririe *et al.* 1997). However, the generation of larger non-specific products could seriously compromise the clarity of melt peaks (for example in this study, primer OL10H02), in which case product differentiation was not be possible. This reinforced the need for well designed primers and highlighted the disadvantages of using primers that had been designed for different – albeit closely related species. Combined PCR reactions, in this study using the primer NA14E08, determined that heterozygote products with 9 bp differences could be clearly identified by two distinct melt peaks. This agrees with the 9 bp difference observed using this technique by Aoshima (2000), but their overall product was shorter at only 55 bp compared to 110 bp in this case. When the difference in heterozygous product size was reduced to 3 bp, one wide melting peak was observed. This peak being wider and less differentiated than those containing only one product may alert the observer to a possible heterozygote, but the two individual alleles could not be reliably identified. Therefore, any heterozygote products containing alleles of 3 bp differences or less could be misidentified using the Roche Lightcycler, and this result confirmed its lack of suitability for direct SSR differentiation. Unfortunately, it was not possible to investigate if a 6 bp difference in a heterozygous sample would be identified due to financial constraints.

A plot showing the observed melting temperatures for PCR products of each known allele of NA14EO8 (5 alleles in total) indicated that there was an overlap in melt temperatures for 4 of the alleles with non-standardised melt temperatures. An overlap of 3 alleles was observed for standardised melt temperatures (Figure 6.6 a, b). These 3 alleles had the following repeats; b $(GCC)_3(GTC)_1$, c $(GCC)_4(GTC)_0$, d $(GCC)_5(GTC)_0$. Two of these alleles have only one base pair difference (b & c), and a three bp difference is present between b & d. Differences in melting temperature for the PCR products of

allele a & b could be differentiated as their melting temperature was sufficiently different, even though there was only a 3 bp difference between alleles. This was also the case for alleles d & e, although only one sample was sequenced with this particular allele, so an indication of melting temperature variation is not available. This observation indicated that although it was possible to distinguish between some alleles with a size difference of 3 bp (e.g. allele a and allele b), this was not possible for all alleles (e.g. allele c and d) and some variation would therefore go unnoticed. Therefore, it would not be possible to differentiate between these three alleles using this method. The lack of ability of the Lightcycler to differentiate between alleles b/c and d may be a combination of the small size difference and the influence that base composition has on DNA melting. The lack of sufficient resolution to identify all alleles using primer NA14E08 with a 3 bp repeat difference gives good reason to suggest that alleles would also be missed for shorter 2 bp repeats (e.g. NA12A02) so this was not investigated further at this stage.

The Roche Lightcycler proved to be a useful instrument for rapid identification of suitable SSRs for a population study, and could give an indication of polymorphic loci prior to further analysis using an alternative, more accurate method for genotyping. In this context, real-time PCR may still be a useful application during the developmental phases of SSR investigation, although it may not be used to provide accurate genotyping.

6.4.2 ***Corbett HRM analysis.*** PCR amplification of SSR products using the Corbett Rotor-gene 6000 was found to be less straightforward than the Lightcycler. In some cases, where the Lightcycler had amplified a product with a particular template DNA, no product for this template was produced with the Corbett until the PCR reaction profile was altered to that reported in the methodology. Reaction optimisation was more critical with the Corbett and longer annealing and extension stages were required during the amplification reaction. This may be due to the reaction vessels, which are constructed in plastic, conducting heat less efficiently than the glass capillary tubes used with the Lightcycler (Elenitoba-Johnson *et al.* 2008). However, when reaction conditions were optimised, and samples were standardised to the same concentration, PCR amplification was reliable and the instrument's capacity for high resolution product differentiation was

demonstrated. The Corbett could clearly identify a heterozygous product where the difference in products' length was 3 bp. This exceeded the maximum observed resolution of 9 bp for a heterozygous product with the Lightcycler. In addition, differences were observed in melt curve shape between two products of identical length, even with a single base change in the microsatellite region. However, attempts to differentiate these products in a synthesised heterozygous situation were less successful (Figure 6.13 and Table 6.10), and in at least one instance this may be due to apparent problems of amplification of only one allele ('null alleles' or allele drop-out) (Figure 6.13).

The Corbett Roto-gene had the resolution to differentiate between alleles tested in this study based on their product length or base composition, but the main problem with the technique was found to be with the analysis of the outputs produced by the instrument. For its main application of identification of SNPs, analysis is relatively straight forward because either a 'wild' type allele, a 'mutant' allele or a heterozygote may be produced (e.g. Herrmann *et al.* 2006). In the case of SSRs, any number of alleles and various combinations are possible. This greatly complicates the allocation of allele identities with the HRM melt curve outputs produced by the Corbett. With standard melt curve analysis (i.e. not HRM), the melt curves can be manipulated to produce melt peaks where a specific temperature can be designated as the melting point of each product (as used with the Lightcycler in this study). If sufficient resolution was available to differentiate between all alleles using this format, analysis of SSRs would be simple as each allele could be easily identified – and characterised by its melt temperature. However, at present this is not possible and while it is relatively simple to identify different alleles from homozygote products with the Corbett, it was not possible to identify the alleles that made up heterozygous products using the HRM facility.

An area for future study using the Corbett rotor-gene 6000 for SSR analysis may involve the use of chloroplast SSRs. While the majority of population studies using SSRs focus on nuclear DNA, polymorphisms at chloroplast SSR loci have been reported, and have been utilised for ecological purposes (Powell *et al.* 1995; Provan *et al.* 1997; Provan *et al.*1999). The chloroplast genome shares many of its features with animal mitochondrial

DNA, including a conserved gene order and a lack of heteroplasmy (the presence of a mixture of more than one type of an organellar genome within a cell) and recombination. These characteristics combined with the presence of polymorphic mononucleotide repeats within the chloroplast genome make them potentially useful as tools for ecological and evolutionary biology studies. As chloroplast primers have been reported as cross amplifiable between many closely related species (Provan *et al.* 2001, Wills *et al.* 2005), they may be of particular interest for studies involving species with a lack of sequence information (such as *H. incana*).

During this study, 3/4 previously described chloroplast SSR primers were identified as potentially useful with *H. incana*. Due to the effective haploid nature of the chloroplast genome, it may be possible to use HRM techniques to differentiate between alleles as the problems identified in this study with this technique for heterozygote SSRs would not be an issue. Chloroplast markers may provide aid in identifying historical population bottlenecks, founder effects and genetic drift (refs in Provan *et al.* 2001); the effects of these events may play an important role in the success of alien plant establishment and future studies comparing native and alien populations may provide an increased understanding of the mechanisms involved in these events. Sufficient funds were not available to continue this area of investigation at this time with *H. incana*. Since chloroplast DNA is typically maternally inherited it would be important to consider this when designing the sampling strategy i.e. to ensure that individuals selected were representative of a population.

6.4.3 ***Allele Variation.*** Although reliable genotyping was not possible with the results generated from this study, an interesting pattern of variation was observed when melting temperatures were plotted for the populations tested with NA14E08 (Figure 6.7). With the exception of Saida (Algeria), little variation was observed within populations, but significant variation was observed among populations (Kruskal-Wallis $p<0.000$). This result was also casually observed with the samples that were sequenced to cover the range of observed melting temperatures. A pattern was seen of different alleles occurring (for the main part) only within discrete populations (Table 6.8). This implies that NA14E08 variation is highly structured between populations of *H. incana,* which may be of interest for future population studies.

In contrast, low variation was observed at the locus NA12A02. In cases such as this, where one allele predominates at a higher frequency, but other alleles are observed at low frequency, it is possible that the most frequently observed allele is the oldest (or ancestral) and the other alleles are the result of more recent insertions or deletions. However it was not confirmed whether this was a real feature of this locus, or if the technique used was failing to identify variation. A previous study investigating the variation RAPD markers of *H. incana* revealed a lack of variation between individuals from the species native and alien ranges (Lee *et al.* 2004). In addition, similar levels of genetic diversity were identified between the native and alien populations, indicating comparable levels of outbreeding and therefore a lack of evidence for higher levels of inbreeding within the alien population. Initial results from this study using SSR loci, suggest that there may be more genetic differentiation between isolated populations of *H. incana* than previously identified using RAPDs. However, from the limited number of SSR loci studied during this investigation and the results generated from the novel approach taken, it has not been possible to estimate levels of genetic diversity for comparative purposes.

6.5 References

Aoshima T, Sekido Y, Miyazaki T, Kajita M, Mimura S, Watanabe K, Shimokata K, Niwa T (2000) Rapid detection of deletion mutations in inherited metabolic diseases by melting curve analysis with LightCycler. Clinical Chemistry, 46, 119-122.

Bernard PS, Lay MJ, Wittwer CT (1998) Integrated amplification and detection of the C677T point mutation in the methylenetetrahydrofolate reductase gene by fluorescence resonance energy transfer and probe melting curves. Analytical Biochemistry, 255, 101-107.

Cawkwell L, Bell SM, Lewis FA, Dixon MF, Taylor GR, Quirke P (1993) Rapid detection of allele loss in colorectal tumours using microsatellites and fluorescent DNA technology. British Journal of Cancer, 67, 1262-1267.

Ceplitis A, Su YT, Lascoux M (2005) Bayesian inference of evolutionary history from chloroplast microsatellites in the cosmopolitan weed Capsella bursa-pastoris (Brassicaceae). Molecular Ecology, 14, 4221-4233.

Creste S, Neto AT, Figuerira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Molecular Biology Reporter, 19, 299-306.

Dietmaier W, Hofstädter F (2001) Detection of microsatellite instability by real time PCR and hybridization probe melting point analysis. Laboratory Investigation, 81, 1453–1456.

Elenitoba-Johnson O, David D, Crews N, Wittwer CT (2008) Plastic versus glass capillaries for rapid-cycle PCR. Biotechniques, 44, 487-492.

Garza JC, Freimer NB (1996) Homoplasy for size at microsatellite loci in humans and chimpanzees. Genome Research, 6, 211-217.

Gibson UEM, Heid CA, Williams PM (1996) A novel method for real time quantitative RT PCR. Genome Research, 6, 995-1001.

Grimaldi MC, CrouauRoy B (1997) Microsatellite allelic homoplasy due to variable flanking sequences. Journal of Molecular Evolution, 44, 336-340.

Gundry CN, Vandersteen JG, Reed GH, Pryor RJ, Chen J, Wittwer CT (2003) Amplicon melting analysis with labeled primers: A closed-tube method for differentiating homozygotes and heterozygotes. Clinical Chemistry, 49, 396-406.

Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. Genome Research, 6, 986-994.

Herrmann MG, Durtschi JD, Bromley LK, Wittwer CT, Voelkerding KV (2006) Amplicon DNA melting analysis for mutation scanning and genotyping: Cross-platform comparison of instruments and dyes. Clinical Chemistry, 52, 494-503.

Higuchi R, Fockler C, Dollinger G, Watson R (1993) Kinetic PCR analysis - real-time monitoring of DNA amplification reactions. Bio-Technology, 11, 1026-1030.

Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'-3' exonuclease activity of Thermus aquaticus DNA polymerase. Proceedings of the National Academy of Sciences of the United States of America, 88, 7276-7280.

Lay MJ, Wittwer CT (1997) Real-time fluorescence genotyping of factor V Leiden during rapid-cycle PCR. Clinical Chemistry, 43, 2262-2267.

Lee PLM, Patel RM, Conlan RS, Wainwright SJ, Hipkin CR (2004) Comparison of genetic diversities in native and alien populations of hoary mustard (Hirschfeldia incana [L.] Lagreze-Fossat). International Journal of Plant Sciences, 165, 833-843.

Linz U (1990) Thermocycler temperature variation Invalidates PCR results. Biotechniques, 9, 286-292.

Lipsky RH, Mazzanti CM, Rudolph JG, et al. (2001) DNA melting analysis for detection of single nucleotide polymorphisms. Clinical Chemistry, 47, 635-644.

Lyon E, Millson A, Phan T, Wittwer CT (1998) Detection and identification of base alterations within the region of factor V leiden by fluorescent melting curves. Molecular Diagnosis, 3, 203-210.

Mackay JF, Wright CD, Bontigoli RG (2008) A new approach to varietal identification in plants by microsatellite high resolution melting analysis: application to the verification of grapevine and olive cultivars. Plant Methods, 4-8.

Mansfield ES, Vainer M, Harris DW, Gasparini P, Estivill X, Surrey S, Fortina P (1997) Rapid sizing of polymorphic microsatellite markers by capillary array electrophoresis. Journal of Chromatography A, 781, 295-305.

Pan K, Lui W, Lu YY, Zhang L, Li Z, Lu WL, Thibodeau SN, You WC (2003) High throughput detection of microsatellite instability by denaturing high-performance liquid chromatography. Human Mutation, 22, 388-394.

Powell W, Morgante M, Andre C, McNicol JW, Machray GC, Doyle JJ, Tingey SV, Rafalski JA (1995a) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Current Biology, 5, 1023-1029.

Powell W, Morgante M, McDevitt R, Vendramin GG, Rafalski JA (1995b) Polymorphic simple sequence repeat regions in chloroplast genomes - applications to the population genetics of Pines. Proceedings of the National Academy of Sciences of the United States of America, 92, 7759-7763.

Provan J, Corbett G, McNicol JW, Powell W (1997) Chloroplast DNA variability in wild and cultivated rice (Oryza spp.) revealed by polymorphic chloroplast simple sequence repeats. Genome, 40, 104-110.

Provan, J. et al. (1999) The use of uniparentally inherited simple sequence repeat markers in plant population studies and systematics. In Molecular Systematics and Plant Evolution (Hollingsworth, P.M. et al., eds), pp. 35–50, Taylor and Francis.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trends in Ecology and Evolution, 16, 142-147.

Rajeevan MS, Ranamukhaarachchi DG, Vernon SD, Unger ER (2001) Use of real-time quantitative PCR to validate the results of cDNA array and differential display PCR technologies. Methods, 25, 443-451.

Ririe KM, Rasmussen RP, Wittwer CT (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Analytical Biochemistry, 245, 154-160.

Saiki RK, Scharf S, Faloona F, et al. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle-cell anemia. Science, 230, 1350-1354.

Saiki RK, Gelfand DH, Stoffel S, et al. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science, 233, 1076–1078.

Selkoe KA, and RJ Toonen (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecolological Letters, 9, 615-629.

Shepherd M, Henry R (1998) Monitoring of fluorescence during DNA melting as a method for discrimination and detection of PCR products in variety identification. Molecular Breeding, 4, 509-517.

Sunnucks P (2000) Efficient genetic markers for population biology. Trends in Ecology and Evolution, 15, 199-203.

Tessier N, Bernatchez L (1999) Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic salmon (Salmo salar L.). Molecular Ecology, 8, 169-179.

Walker NJ (2002) A technique whose time has come. Science, 296, 557-559.

Weckworth BV, Talbot S, Sage GK, Person DK, Cook J (2005) A signal for independent coastal and continental histories among North American wolves. Molecular Ecology, 14, 917-931.

Weng TT, Jin NL, Liu L (2005) Differentiation between amplicon polymerization and nonspecific products in SYBR green I real-time polymerase chain reaction. Analytical Biochemistry, 342, 167-169.

White E, Sahota R, Edes S (2002) Rapid microsatellite analysis using discontinuous polyacrylamide gel electrophoresis. Genome, 45, 1107-1109.

Wills DM, Hester ML, Liu AZ, Burke JM (2005) Chloroplast SSR polymorphisms in the Compositae and the mode of organellar inheritance in Helianthus annuus. Theoretical and Applied Genetics, 110, 941-947.

Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques, 22, 130-138.

Woodhead M, Russell J, Squirrel J, Hollingsworth PM, Mackenzie K, Gibby M, Powell W (2005) Comparative analysis of population genetic structure in Athyrium distentifolium (Pteridophyta) using AFLPs and SSRs from anonymous and transcribed gene regions. Molecular Ecology, 14, 1681-1695.

Worm J, Aggerholm A, Guldberg P (2001) In-tube DNA methylation profiling by fluorescence melting curve analysis. Clinical Chemistry, 47, 1183-1189.

Ye J, Parra EJ, Sosnoski DM, et al. (2002) Melting curve SNP (McSNP) genotyping: a useful approach for diallelic genotyping in forensic science. Journal of Forensic Sciences, 47, 593-600.

Chapter 7

General discussion and summary

H. incana is rapidly expanding its range and is likely to continue to do so because it is able to exploit areas of recent disturbance and to dominate early successional communities. It is classified as an introduced invasive species in the USA and Hawaii, South America, New Zealand, Japan, Ireland, Lithuania and as a noxious weed in Australia. In the UK the species does not have invasive status, but it is regarded as an alien.

The capacity for a species to become established and potentially invasive in new locations depends on a combination of the individual species' characteristics, the frequency of introduction and the availability of a suitable habitat where the species may persist or flourish. *H. incana* is not a domesticated species, but it has reportedly been grown as a crop (Preston *et al.* 2002) and may be a contaminant in bird seed. This information indicates that multiple introductions of *H. incana* are likely. As discussed in chapter 1, certain characteristics can increase the potential for a species to become invasive. Species traits may also give an indication of which habitats may be at risk from invaders as their success may be contingent on the availability of suitable habitat types. Below, I consider the characteristics discussed in chapter 1 with respect to what is known for *H. incana*.

Competitiveness. We do not have any evidence that *H. incana* displays additional competitive traits outside of its native range, as has been found for some invasive species. It probably does not have a competitive advantage due to an enemy release phenomenon, as the main species that predate on *H. incana* are generalists. In addition, there are no reported allelopathy interactions between *H. incana* and other species. *H. incana* can dominate open early successional communities, but competes poorly as the community progresses into later stages.

High turnover. In this thesis I showed that individual plants of *H. incana* do indeed have the capacity to produce large quantities of seed. In addition, in areas of disturbed or bare soil, the species can attain high densities. Rapid generation times combined with high fecundity give rise to high levels of propagule pressure, an important characteristic

associated with invasiveness (Kolar & Lodge 2001). In some cases, invasive ecotypes have been observed to display shorter times to first reproduction than their native counterparts (Burnes 2008). This is not the case for *H. incana* as reproduction is observed in the first season for both alien and native populations. However, in the case of alien populations in Swansea, the resulting seed generation from perenniating plants far exceeds that recorded from plants from a native location. This is a key reproductive advantage in the alien habitat that the established, naturalised plants in Britain have over those that originate directly from native areas. The findings that plants from native areas still flowered quickly and behaved as annuals when grown in the British environment, whereas British plants responded to cold treatment and formed rosettes to overwinter, are strong indications that the shift in reproductive strategy has evolved as adaptations to the British environment rather than being merely plastic traits. The fact that plants that overwinter to reproduce a second year enjoy much higher overall fecundity means there is a clear selective advantage to this life history strategy.

Dormancy. Dormancy of *H. incana* seed in the seed bed has not been specifically tested, but under optimal storage conditions seed of over 40 years can be successfully germinated.

Dispersal. *H. incana* does not display specialist dispersal mechanisms. It is likely that the main natural dispersal agents are wind and animal. *H. incana* is also likely to be unintentionally distributed by anthropogenic activities. *H. incana* is often found in recently disturbed sites, frequently isolated from other populations. The exact mechanism of dispersal to these sites is not clear, but may be associated with movement of topsoil, construction vehicles or building materials.

Asexual reproduction. *H. incana* has the capacity to self-fertilise; however this has only been observed at low levels (Darmency & Fleury 2000). The species would be generally described as outcrossing, pollen to ovule ratios for individuals from both the native and alien range were within the boundaries of facultative outbreeding species. Attempts were made to measure rates of self-pollination of *H. incana* for a range of populations,

however in this case useable results were not achieved. A comprehensive study of self-pollination rates for native and invasive populations of *H. incana* would show whether similar breeding strategies occur for all populations, or if some populations show greater ability to self-pollinate. Self pollination is widely regarded as a useful trait for the colonization process.

Rate of growth. *H. incana* displays rapid growth rates and in early successional communities can dominate over other species. Individuals from an alien population in Swansea displayed higher growth rates than most populations from the native range. Additionally, the perennial habit of individuals in Swansea leads to higher levels of biomass accumulation than found in native regions, where an annual habit is the norm.

Adaptability to local climate. *H. incana* displays the capacity to adapt to a wide range of climates. This is indicated by its increasing range from its native region of the Mediterranean and Middle East into temperate and tropical regions (Jalas & Suominen 1996, Preston *et al.* 2002). Currently it has not been confirmed if this successful expansion is due to inherently high levels of plasticity, or if rapid adaptation to novel habitats has occurred. Quite possibly it is a combination of both. The phenotypic differences observed in this study between populations under common garden conditions would suggest that genetic differences are apparent between populations. The differences in reproductive strategy between British and native populations that persists in the common garden further suggests evolution rather than plasticity of these particular traits. Nevertheless, there must have been some plasticity in the original plants that first arrived in order for these to survive and produce the first few generations, prior to evolving the adaptations that resulted in self-sustaining, invasive populations. The source populations of *H. incana* that have made up the current populations in the UK are not known. The possible sources of *H. incana* include contaminated bird seed, a contaminant in wool and a potential historic crop. It is likely that multiple introductions of *H. incana* have occurred from a variety of source locations. It may or may not be possible to determine the origins of the colonization using molecular methods. This may depend on the number of introductions and the level of outcrossing between individuals.

Reproductive timing. *H. incana* is a winter annual in its native range. This coincides with the main growing season for annuals in Mediterranean habitats and is a function of water availability. In the UK, *H. incana* not only behaves as a perennial, but its reproductive timing has changed to coincide with the summer, which provides better growing conditions in the UK and also the availability of pollinators. The common garden experiments suggest these are evolved rather than plastic traits, and perhaps a key adaptation leading to the success of *H. incana* in establishing populations in the UK.

In summary, *H. incana* has many of the characteristics that would be expected when attempting to predict invasive behavior. Indeed with its rapid colonization and expansion rates it would qualify as an invasive species under the criteria speculated by Richardson *et al.* 2000.

Perhaps the most important ecologically significant observation documented in this thesis is the changes in life history observed between the species native range and populations in the UK (as studied in Swansea). The species acts as a winter annual in its native region and behaves as a perennial in the UK. A strategy of rapid growth and completion of seed generation is required in the Mediterranean before the dry season effectively ends the growing season for a species such as this. Under common garden conditions the general pattern was for individuals from the native regions studied to produce flowers more rapidly than individuals from Swansea (an example of a northern alien population). At some stage during the species northern expansion there has been a change in the timing of flowering initiation to coincide with the more favorable season for growth and availability of pollinators for plants in northern latitudes. In addition, investment in plant growth is far higher in its alien range e.g. Swansea when compared to native individuals e.g. Crete. As a strategy of perennation is apparent in Swansea, rapid flowering and production of seed is not so critical in this location when compared to plants from a climate with a pronounced dry season. Seed production in plants from Swansea and Crete flowering in their first year was comparable, but plants of greater than one year flowering

in Swansea produced significantly more seeds, greatly increasing the propagule pressure of the species. This may be a key factor in the species ability to rapidly colonize new regions and dominate early successional habitats within the UK. Propagule pressure is an extremely important factor that may determine the ability of a pioneering species to persist. Increasing the number of populations studied from the alien range would give a better understanding if the observations made in Swansea are common to all northern alien populations, or if each isolated population must be treated as a separate case.

Phenotypic differences observed under common garden conditions would imply genetic differences are present - as external environmental pressures (the other factor determining phenotype) were acting equally on the different populations. Flowering times for *H. incana* were more rapid for plants from Mediterranean type regions such as Crete, Cyprus, Morocco, Algeria and California in comparison to plants from Swansea under common garden conditions, both under greenhouse conditions and when grown outdoors in Swansea in the spring. In addition, greater investment in vegetative growth was observed in plants from Swansea under these same conditions, including the persistence of basal leaves after flowering only in individuals from Swansea. With high levels of phenotypic diversity between populations, one may expect high levels of genetic diversity to be apparent also. One may hypothesize that, greater genetic differentiation would be observed between more isolated populations and between populations displaying greater differences in characteristic traits e.g between plants originating from Mediterranean regions and from temperate regions, or plants displaying rapid flowering times versus delayed flowering.

Unfortunately during this study it was not possible to determine levels of genetic diversity of *H. incana* or confirm similarity or differentiation of genetic structure between the populations investigated during the phenotypic comparisons due to difficulties encountered during the evaluation of real time PCR and DNA melting analysis for the application of SSR population studies. The level of resolution required to differentiate between SSRs of similar repeat number could not be achieved using the Roche Lightcycler. Although it appeared possible to differentiate between homozygous

individuals with a 3 bp difference in PCR product length, this could not be applied to heterozygous individuals, where an observed maximum resolution of 9 bp was recorded. The differential capacity was not sufficient in this case and although the technique initially appeared promising it could not be used for analytical purposes. The ease and speed of the technique would, however, lend itself to pilot studies investigating the suitability of SSR primers for use with new species and as a starting point to identify which primers amplify polymorphic products. In this case the technique is faster and more user friendly than PCR followed by post processing techniques such as polyacrylamide gel electrophoresis. In addition this technology does not require the use of specialist software that is required to analyze data produced by fluorescent genotyping methods using automated sequencers.

The new breeds of high resolution DNA melting apparatus offer significantly greater differentiation power due to advanced dye chemistry and greater sensitivity in terms of measuring fluorescence of double or single stranded DNA molecules. Whilst the Corbett Rotorgene has sufficient resolving capacity to differentiate between heterozygote SSR products, at present it does not appear to be possible to easily identify products based on their individual melt curves with currently available software, which is designed for SNP analysis.

The main purpose of this study was to investigate the characteristics of *H. incana* that make it a successful invader of the UK. By comparing populations of the species from a range of isolated locations it was found that the species displays high levels of phenotypic variation and although a full study was not undertaken to determine levels of genetic diversity, preliminary results suggest that genetic differences are observable between populations. Although sample sizes of sequenced SSRs isolated from *H. incana* were small, differences observed within SSR loci were observed between populations (Table 5.8). These results do not eliminate the possibility of phenotypic plasticity being a component of the observed differences in phenotypic characteristics, but they do demonstrate a difference in genetic structure between populations. It was not possible to identify the source population of *H. incana* found in Swansea in this study. However,

close inspection of loci NA14E08 indicated that only individuals from Swansea and Spain contained the same compound SSR of (GCC)(GTC) whereas individuals from all other populations had a simple SSR of (GCC). One could speculate that this result hints of a closer relationship between populations from Swansea and Spain than to other populations studied here, but a greater number of associations from a larger set of SSRs would be required to draw any conclusions with confidence. It is interesting to note that some phenotypic characteristics also showed greater similarity between Swansea and Spain. Leaf structure had greater visual similarity between Swansea and Spain than between other southern populations. In addition, under greenhouse conditions, Swansea and Spain were the only populations to contain individuals that did not come into flower rapidly in one growing season.

This study provides an example of the advantages of taking a multi-disciplinary approach to understand the processes involved in invasive biology. By comparing the ecology of a species from contrasting habitats it is possible to have a greater understanding of the species potential for successful expansion across a wider range of locations. It is useful to study species in the wild to observe community structure and combine this with common garden studies where differences in external environmental stimuli are minimised. This allows for greater predictive power when considering how a species will behave if introduced into a novel location. This may allow for early recognition of potential problem species and give time for measures to be taken to prevent invasive species from becoming naturalised in new areas.

This study has highlighted some of the characteristics that make *H. incana* a successful invader of the UK, and this information is likely to be relevant to other northern temperate locations where *H. incana* is increasing in its range. Observations made during this study prompt questions about the mechanisms involved in rapid adaptive change to new environments. A future study investigating candidate genes may give a greater understanding of the capacity for genetic change within short time scales. By studying genes linked to the observations documented in this thesis, such as flowering time regulatory genes, it may be possible to determine detailed information of genetic

differentiation between populations that has only been speculated here. In combination with a detailed population study, it may be possible to determine not only routes of introduction of this invasive species, but also (if indeed genetic variation was found among candidate genes) at what stage in the species history genetic differentiation occurred.

References

Burns JH (2008) Demographic performance predicts invasiveness of species in the Commelinaceae under high-nutrient conditions. Ecological Applications, 18:335-346.

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution, 16, 199-204.

Preston CD, Pearman DA, Dines TD (2002) New atlas of the British & Irish flora: an atlas of the vascular plants of Britain, Ireland, the Isle of Man and the Channel Islands. Botanical Society of the British Isles, Oxford University Press.

Chapter 8

Appendix

8.1 ***Fv/Fm Ratios.*** After a period of hot sunny days and cold nights (in Swansea), the production of anthocyanin pigmentation in some individuals of *H. incana* was observed among some of the plants that had been grown to investigate the ability of *H. incana* to survive British winters (see chapter 4). The production of anthocyanin in stems and leaves may be linked to photoinhibition. Individuals from Elounda, Crete appeared to have produced greater quantities of anthocyanin and this would indicate a higher level of photoinhibition. This observation prompted the measurement of Fv/Fm ratios and calculation of photoinhibition scores for the individuals from Swansea, and those from Crete.

Fv/Fm ratios were measured in 10 leaves each from Swansea and Elounda, Crete using the Hansatech Handy PEA. Leaf clips were attached to each leaf over the sample site. These were then left for 30 minutes to exclude light and allow dark adaptation to occur. The Hansatech Handy PEA was then attached to each sample and Fv/Fm ratios were measured.

Table 8.1 Photoinhibition measurements for *H. incana* for individuals from Swansea and Crete, grown outdoors in the Botanic garden at Swansea University. Data presented includes mean and standard deviations for ratios of Fv/Fm, mean and standard deviations for PI, and sample sizes.

	Fv/Fm			PI		
	Mean	SD	N	Mean	SD	N
Swansea	0.81	0.016	8	2.405	0.6775	8
Crete	0.77	0.024	10	1.663	0.9074	10

A significant difference was observed between the Fv/Fm ratios for individuals from Swansea and Crete (independent sample t-test P=0.005). Fv, the variable fluorescence, is given by the difference between the maximum fluorescence (Fm) and the fluorescence under dark conditions (Fo). The depressed Fm/Fv ratio for individuals from Crete in this instance gives an indication that they may have been subjected to higher levels of photoinhibition than individuals from Swansea. This suggests that plants from Crete may be more sensitive to chronic photoinhibition when subjected to high daytime irradiance (PPFD > 400 $\mu molm^{-2}s^{-1}$) following low night time temperatures (minimum temperature < 5 °C), than plants from

Swansea (UK) under the same conditions. A typical range for non-photoinhibited plants would be > 0.80, but values of < 0.75 would indicate severe photoinhibition. PI values are calculated directly by the Handy PEA software, and is described by Strauss *et al.* (2006) and Strasser & Strasser 1995.

8.2 ***Nitrate Reductase.*** Nitrate is an important source of nitrogen for assimilation into amino acids and ultimately proteins. One rate limiting step during this process is the reduction of nitrate to nitrite (catalysed by the enzyme nitrate reductase, NR). Furthermore, the *in vivo* NR activity measured in leaves can provide an indication of nitrate availability in different habitats. Leaves were collected from two populations of *H. incana* from contrasting sites (Sites 1 and 2 from field work chapter 3). From each site leaves were collected from 5 plants. The following methodology follows the *in vivo* nitrate reductase (NR) standard assay described by Al-Gharbi (1984).

Leaves were washed with tap water and blotted dry. Midribs were discarded and 0.2 g of leaf tissue were cut into smaller pieces and placed into McCartney bottles. Into each bottle 5 ml of assay buffer (containing 50 mM potassium phosphate, pH 7.0, 40 mM potassium nitrate and 1.5% (v/v) n-propanol) was added. Bottles were fitted with vaccine stoppers and evacuated to produce a partial vacuum. Assays were initiated by wrapping bottles in silver foil to exclude light, and terminated by placing the bottles into a boiling water bath for 5 min. Nitrite was measured by the Griess-Ilosvay method. Equal volumes of reagent A (10 g sulphanilimide in 800 ml distilled water + 200 ml conc. HCl) and reagent B (1 g N-1-naphthyl ethylene diamine dihydrochloride in 1.0 litre of distilled water) were mixed prior to use. One ml of the mixed reagent was added to each sample. After a period 20 minutes the absorbance of the solutions was measured in a spectrophotometer at 535 nm which were compared to a calibration curve prepared from standards containing known amounts of nitrite.

Table 8.2 Nitrate reductase activity for two populations of *H. incana* in the Swansea area. Data presented includes; mean NR activity, Standard deviation and sample size. The habitats can be described as follows: Pennard - improved grassland, plants growing on the edge of a golf course, sandy substrate. Swansea docks - disused industrial site, substrate containing high proportion of gravel and building materials, e.g. rubble.

Population	Mean NR activity	SD	N
Pennard	1.56	0.4	5
Swansea Docks	1.13	0.3	5

These are typical levels of nitrate reductase activity for a nitrophilous species, i.e. high NR activity in association with high nitrogen content of soils. These NR values are comparable to those observed for *Cardamine hirsuta* (1.68-2.09) and *Coronopus didymus* (1.19-2.85) from the Brassicaceae (Al Gharbi & Hipkin 1984). These species, in common with *H. incana*, are frequently found on disturbed sites.

8.3 ***Calculation of pollen to ovule ratios and the estimation of breeding system.*** Pollen to ovule ratios were calculated for individuals of *H. incana* from Swansea and Crete populations to give an indication of breeding strategy. Flowers were taken from plants grown in the greenhouse. Each population was grown from seed from one maternal source. In addition, pollen to ovule ratios were calculated for the same Swansea population grown outdoors for comparative purposes. For each population, five flowers from each of eight individuals were selected for counting (this was the number of plants in flower available for use at this time). For pollen counts all the anthers from a flower were placed in 5 ml of water and crushed to release the pollen grains. 1ml was transferred to a sedgewick rafter counter, of which 30 squares were counted at x20 magnification for pollen number. Ovules were dissected from the ovary and counted at x6 magnification. Pollen to ovule ratios were then calculated by the total pollen number, as estimated from the sub sample, divided by the ovule count. Pollen to ovule ratios were then compared to Crudens logarithmic scale for estimation of breeding system.

Pollen/ ovule ratios have been used as an indicator of breeding strategy. Pollen / ovule ratios of *H. incana* were within the range of 3.3-3.5 (Table x), these values lie between the boundaries of facultative and obligate outcrossing as defined by Cruden (1977). The mean pollen / ovule ratio was significantly higher for Crete than for Swansea (ANOVA $P < 0.001$), but this difference was not sufficient to indicate that its breeding strategy in Swansea is any different than in Crete.

Table 8.3 Pollen ovule ratios for *H. incana*.

Population	Pollen / ovule ratio	Log P/O ratio
Swansea (greenhouse)	2140	3.3
Swansea (outdoor)	2540	3.4
Crete (greenhouse)	3020	3.5

The amount of pollen produced by a flower is determined not only by breeding system, but by additional factors such as the vector required for pollination and whether pollen or nectar is consumed by the vector. In some cases misleading results can be given if external factors are not taken into account, which is a potential pitfall in the technique. However, a study of 66 species from the Brassicaceae by Preston (1986) indicated that if such factors are given consideration then pollen to ovule ratios do indeed give a good measure of breeding system and an estimate was provided for allogamous and autogamous crucifers of 3500 and 1000 respectively (the range for *H. incana* was 2138-3022).

Table 8.4 Climatic data from (http://www.climate-charts.com/) from the nearest weather stations to each population site indicating mean maximum and minimum monthly temperatures (°C) and mean monthly precipitation (mm) in each row respectively.

Site	Records	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mor 1	Casablan	17	17	18	20	22	23	25	25	25	22	20	17
	ca	8	9	10	12	15	17	20	20	17	15	12	9
		60	50	45	40	20	5	0	0	5	40	80	80
L.A.	La Int	19	19	19	21	21	22	24	26	24	24	21	20
	Airport	10	10	10	12	14	16	18	18	17	16	11	9
		20	20	20	5	2	2	0	0	2	2	13	15
Alg2	N/A												
Cyp	Larnaka	17	17	19	22	25	30	32	32	30	27	22	17
		8	8	9	12	15	17	21	20	18	16	12	8
		60	50	40	15	5	0	0	0	5	20	40	60
Alg1	Bejar	17	17	18	20	22	25	30	30	27	25	20	17
		8	8	9	10	12	16	20	20	18	15	12	8
		100	100	100	80	40	20	0	10	40	80	100	120
Crete	Heraklio	15	15	17	20	22	27	28	27	26	22	20	17
	n	9	9	10	12	15	18	22	22	19	17	14	10
		85	80	60	30	20	5	0	5	20	60	60	70
Mor2	Taza	15	16	17	19	23	25	31	31	26	24	17	16
		6	7	8	8	13	15	18	18	16	14	10	7
		220	180	210	170	110	40	10	5	80	130	150	290
Turk	Usak	6	7	12	17	22	26	29	29	26	20	12	7
		-1	0	2	5	7	12	14	14	11	7	3	2
		75	70	60	50	40	20	15	10	20	40	60	80
Spain	Salamanc	40	40	30	40	40	30	20	20	30	40	40	40
	ar *												
Swan	Cardiff	7	7	10	12	15	17	20	20	17	15	11	8
		2	2	3	4	7	10	12	12	10	7	5	2
		90	70	70	60	60	70	75	80	90	100	100	100

* No temperature data available from this site

Table 8.5 Measurements of *H. incana* traits from plants grown within the university greenhouse (germination time, basal leaf number, rosette radius, leaf number multiplied by rosette radius, leaf width, leaf lobe length, flowering stem height at time of first flower opening and time from germination until flowering). These include population means, standard deviations and sample sizes for each trait measured. Populations are listed in order of increasing latitude.

	Phenotypic traits, Mean values							
	Germ	Leaf n	Radius mm	LxR	Leaf w mm	Lobe l mm	Stem h mm	Flower t days
Overall Mean	5.6,4.1,166	13.7,3.7,164	223,36,113	3051	63,10.0,113	75,21.6,113	558,192,148	100.1,29.9,148
Pop								
Mor1	4.4,1.0,19	10.7,1.8,19	235,27,10	2515	67,5.1,10	87,12.6,10	699,78,19	69.7,4.3,19
USA	11.0,2.5,14	7.8,1.3,14	196,36,10	1534	55,12.0,10	90,16.9,10	453,110,12	67.9,8.3,13
Alg2	5.4,4.5,16	16.9,3.6,16	205,28,10	3763	60,7.1,10	66,12.7,10	302,128,16	99.5,12.9,16
Cyp	5.8,0.9,17	11.4,1.5,17	183,35,10	2095	62,7.3,10	74,10.9,10	506,72,16	79.6,4.0,16
Alg1	6.5,2.7,13	10.8,1.6,12	223,26,10	2414	65,6.9,10	103,17.5,10	753,79,11	108.7,12.0,12
Crete	3.3,1.3,10	12.4,2.6,9	235,25,9	2914	80,13.8,9	97,14.2,9	673,96,8	82.5,5.1,8
Mor2	7.2,2.8,11	13.4,2.4,11	211,33,10	2839	61,8.3,10	89,19.7,10	294,114,11	84.3,10.7,11
Turk	3.5,1.0,4	17.5,2.5,4	172,12,4	3024	57,11.5,4	58,12.4,4	600,203,4	110.5,12.9,4
Spain	5.8,10.0,16	13.9,1.6,15	211,28,10	2945	61,6.9,10	63,10.7,10	793,123,13	109.1,17.6,13
Swan2	4.4,4.5,18	17.6,1.6,18	233,22,10	4108	66,6.7,10	53,9.9,10	586,133,15	127.5,19.8,15
Swan3	4.4,1.0,13	17.1,1.7,14	250,21,10	4290	59,8.0,10	60,10.5,10	523,140,8	154.8,23.0,8
Swan4	4.1,1.1,15	16.4,2.7,16	232,25,10	3806	60,8.1,10	53,11.9,10	556,128,13	141.8,26.1,13

Table 8.6 Measurements of *H. incana* traits from plants grown outdoors within the university botanic garden. These include population means, standard deviations and sample sizes for each trait measured. Populations are listed in order of increasing latitude.

	Phenotypic traits, Mean values							
	Germ	Leaf n	Radius mm	LxR	Leaf w mm	Lobe l mm	Stem h mm	Flower t days
Overall Mean	6.0,2.6,144	24.6,15.6,79	234,15.6,79	5764	69,19.3,79	94,29.0,79	362,180,79	88.8,18.4,79
Pop								
Mor1	6.4,2.1,16	8.3,3.8,8	142,29.3,8	1174	55,12.0,8	84,19.7,8	346,119,8	67,9.5,8
USA	11.2,1.1,5	9.0,4.2,2	224,9.2,2	2012	70,0,2	107,18.4,2	255,78,2	74.5,4.9,2
Alg2	5.5,1.8,13	55.3,13.4,8	302,35.0,8	16605	85,12.0,8	84,22.0,8	340,167,8	112.3,18.4,8
Cyp	6.1,1.6,18	18.4,10.1,7	224,51.9,7	4123	69,16.8,7	78,23.3,7	333,198,7	88.4,16.0,7
Alg1	8.0,3.3,14	35.2,17.0,6	263,75,6	9268	91,17.2,6	115,35.6,6	322,194,6	106.3,19.2,6
Crete	5.2,1.9,6	13.2,8.4,5	146,13,5	1932	51,13.5,5	76,17.1,5	302,89,5	72.2,9.9,5
Mor2	5.5,2.3,13	23.1,11.3,8	245,68,8	5660	72,20.6,8	129,44.6,8	158,102,8	93.6,18.3,8
Spain	4.6,2.6,17	19.4,8.4,10	254,64,10	4933	59,12.0,10	80,14.4,10	409,149,10	85.6,11.3,10
Swan2	4.3,1.6,18	23.1,10.1,10	233,65,10	5387	76,29.4,10	96,22.3,10	401,216,10	81.6,11.4,10
Swan3	7.6,2.4,12	30.0,9.8,5	258,51,5	7728	60,7.8,5	87,21.7,5	450,227,5	93.4,14.4,5
Swan4	4.8,1.8,12	25.3,7.0,10	251,31,10	6350	65,7.1,10	105,22.7,10	521,132,10	92.5,11.2,10

Table 8.7 Comparison of growing conditions (random block design vs rows) and their impact on characteristic traits. For each trait measured, a Mann Whitney U test has been calculated between growth conditions. Bonferroni corrections for multiple comparisons have been calculated. For a *P* value to be considered significant it must be lower than the Bonferroni correction. In this case no significant differences were observed between growing conditions across the traits measured.

	Leaf length	Lobe length	Leaf width	Leaf number	Flowering time	Stem length
P	0.010	0.046	0.088	0.090	0.122	0.300
N	79	79	79	79	79	79
Bonferroni Correction	0.0083	0.0100	0.0125	0.0167	0.0250	0.050

Table 8.8 Selection of primers identified as successfully amplifying a product across a range of species from the Brassicaceae (information from the Brassica DB (http://www.brassica.info/ssr/SSRinfo.htm)).

Name	primer	primer b
BRMS-001	GGTGGCTCTAATTCCTCTGA	ATCTTTCTCTCACCAACCCC
BRMS-003	ACGAATTGAATTGGACAGAG	CAGATGGGAGTCAAGTCAAC
BRMS-008	AGGACACCAGGCACCATATA	CATTGTTGTCTTGGGAGAGC
BRMS-019	CCCAAACGCTTTTGACACAT	GGCACAATCCACTCAGCTTT
BRMS-020	AACAAGAGAAGGAGAGCCACCG	CGCTTATAAAATGGCAGTCGCA
BRMS-027	GCAGGCGTTGCCTTTATGTA	TCGTTGGTCGGTCACTCCTT
BRMS-029	AACAAATGACACACACCACACT	ATTGAAAATCTTAACCGTGAAG
BRMS-033	GCGGAAACGAACACTCCTCCCATGT	CCTCCTTGTGCTTTCCCTGGAGACG
BRMS-037	CTGCTCGCATTTTTTATCATAC	TACGCTTGGGAGAGAAAACTAT
BRMS-040	TCGGATTTGCATGTTCCTGACT	CCGATACACAACCAGCCAACTC
BRMS-044	AGGCGAGGAGAAGACAACACAA	TACGGGTGGTTTGAATCAGCAG
BRMS-049	GATCTTCTCTCCAAAACTCTCT	AAAGTCCAAGCTAAATTACAAA
BRMS-051	GGCCAAGCCACTACTGCTCAGA	GCGGAGAGTGAGGGAGTTATGG
sORA43	GCGCGTGTGGGATCAGAA	CTTCTCCACCGTCGATCG
CALSSR	GTTAAGTGTGGCGTTAGAGG	CCTTGGTACATGCCACTGAA
SSR Na10-D09	AAGAACGTCAAGATCCTCTGC	ACCACCACGGTAGTAGAGCG
SSR Na10-F06	CTCTTCGGTTCGATCCTCG	TTTTTAACAGGAACGGTGGC
SSR Na12-A01	GCATGCTCTTGATGAACGAA	GCTTCAACCTCTCAATCGCT
SSR Na12-A02	AGCCTTGTTGCTTTTCAACG	AGTGAATCGATGATCTCGCC
SSR Na12-A07	TCAAAGCCATAAAGCAGGTG	CATCTTCAACACGCATACCG
SSR Na12-B05	CAAATATCCGTCATCGGAGC	CCTGCGGGATATTGAAGACC
SSR Na12-G05	CCGATCATACCTTTTACTCTAGCC	GATGTTCCTCTCGGTGATGC
SSR Na14-E11	TCATCCTTCTCACACCAAAATC	CCTCGAAATAGCTCCAACCC
SSR Ni2-C12	ACATTCTTGGATCTTGATTCG	AAAGGTCAAGTCCTTCCTTCG
SSR Ni2-F02	TGCAACGAAAAAGGATCAGC	TGCTAATTGAGCAATAGTGATTCC
SSR Ni4-B10	GTCCTTGAGAAACTCCACCG	CCGATCCCATTTCTAATCCC
SSR Ol10-F09	AGAGAGCGAGATTGATTGGC	AAACGACCACGAGTGATTCC
SSR Ol12-A04	TGGGTAAGTAACTGTGGTGGC	AGAGTTCGCATACTCTGGAGC
SSR Ol12-F02	GGCCCATTGATATGGAGATG	CATTTCTCAATGATGAATAGT
SSR Ol12-F11	AAGGACTCATCGTGCAATCC	GTGTCAGTGGCTACAGAGAC
SSR Na12-C08	GCAAACGATTTGTTTACCCG	CGTGTAGGGTGATCTAGATGGG
SSR Na12-E05	CGTATGTTTGTTCCACCTGC	ACTAGCAACCACAACGGACC
SSR Na14-E08	TTACTATCCCCTCTCCGCAC	GCGGATTATGATGACGCAG
SSR Ni4-C11	ATAAGGCCGATGTTTCCTCC	CACAGTTGCTTTGGATCTCG
SSR Ni4-D09	AAAGGACAAAGAGGAAGGGC	TTGAAATCAAATGAGAGTGACG
SSR Ni4-D10	ACATGCGAAAGGGATTTGAC	TGCAAGTGAACTCAAAACAAAAG
SSR Ni4-H04	CAAGAAAGGGTATTGCGTCG	TGTTTTAGAAATGGTATGCCCC
SSR Ol10-H02	AACAGGAAGAAACGACGAGG	AGAGAGCCATGAGAAGCACC
SSR Ol11-B05	TCGCGACGTTGTTTTGTTC	ACCATCTTCCTCGACCCTG
SSR Ol11-G11	GTTGCGGCGAAACAGAGAAG	GAGTAGGCGATCAAACCGAG
SSR Ol11-H08	ATCGAAGGCAAATCACCAAG	CAGTGATTCAGAGGTATAGAG
SSR Ra2-G09	ACAGCAAGGATGTGTTGACG	GATGAGCCTCTGGTTCAAGC
SSR Ra3-E05	TTCTCATGCTCCAACCACAG	GTTTCTTCCAAGCCAAGCTG

Table 8.9 Source of *H. incana* seed.

Population	Supplier	Institute
Mor1	Takeshi Nishio	Tohoku University
USA	Takeshi Nishio	Tohoku University
Alg2	Gomez-Campo	Universidad Politecica de Madrid
Cyp	Unknown	University of Hamburg
Alg1	Takeshi Nishio	Tohoku University
Crete	Rosemary John	N/A
Mor2	Gomez-Campo	Universidad Politecica de Madrid
Turk	Takeshi Nishio	Tohoku University
Spain	Gomez-Campo	Universidad Politecica de Madrid
Swan2	Melvin Smith	Swansea University
Swan3	Melvin Smith	Swansea University
Swan4	Melvin Smith	Swansea University

9. Glossary

Allelopathy; inhibition of plant growth by the release of chemicals (allelochemicals) from neighbouring plants.

Allopolyploid; polyploid arising from the hybridisation of different species.

Annual; a plant that completes its entire growing lifecycle in a single growing season.

Anthesis; the period that a flower is fully open and functioning.

Autecology; the ecological relationship between as organism or species and its environment.

Autogamy; self fertilization in plants.

Capitate Stigma; a stigma with a single head rather than being divided into a number of stigmatic surfaces.

Dehiscent; opening in some definite way, e.g. the capsule of a plant opening to release seeds.

Emarginated petal, margin of petal broken by notch or notches.

Emarginated stigma; margin broken by notch or notches.

Endothecium; the sub-epidermal layer of the anther wall, the dehiscent strip is a strip of cells that become thickened as the anther matures and may aid in the splitting open of the anther before the pollen is released.

Fecundity; reproductive capacity, number of seeds set.

Fv/Fm ratio; the maximum quantum efficiency of photosystem II.

Founder event; introduction of a gene pool that may not represent the parent population due to a low number of low introductory individuals.

Gene flow; transfer of genes between populations of the same species, by migration or dispersal of seeds and pollen.

Genotype; the genetic constitution or allele composition of an individual.

Glabrous; no hairs, projections or pubescence.

Glaucous; pale grey to blue waxy coating that can be removed.

Glucosinolates; secondary metabolitos of plants, particularly common within the Brassicaceae.

Gregariousness; grouping structure of a species, related to density at which plants in a population may grow.

Gynodioecious; having bisexual flowers on some plants and only female flowers on other plants of the same species.

Heteroplasmy; mutations in a fraction of organellar DNA (e.g. mitochondrial DNA), in contrast to homoplasmy where the mutation is present in all copies.

Heterozygous; having different alleles at a particular gene locus on equivalent chromosomes.

Homologous; (of chromosomes) a pair of chromosomes, one inherited from each parent, having corresponding gene sequences. (Of genetic sequences) similarity of genetic sequences (often called conserved) due to related ancestry, as contrasted to similar sequences that occur due to chance.

Homoplasy; similarity due to convergent evolution, but of different ancestral origin.

Homozygous; having the same alleles at a particular gene locus on equivalent chromosomes.

Intron; a segment of a gene between exons that does not function in coding for protein síntesis.

Introrse anthers; anthers facing inward turned toward the axis.

Linear leaf, long slender simple leaf; lancolate, tapering to a point at both ends.

nad4; from nicotinamide adenina dinucleotide, a coenzyme used in glycolysis.

Null alleles; (with regard to molecular markers), a mutation in the annealing site of a PCR primer resulting in no product amplification.

Obovate petal; teardrop-shaped petal, attaches to a tapering point.

Obtuse petal; petal with blunt tip.

Perennial; a plant that under natural conditions lives for several to many years (in contrast to annual or biennal plants).

Petal claw; the narrow stalklike basal part of some petals.

Petiolate leaf; leaf containing short stalk attaching leaf blade to the stem.

Phenology; the relationship between periodic biological phenomenon and the climate.

Phenotype; observable physical and chemical characteristics, determined by genetic and environmental influences.

Photoinhibition; a reduction in photosynthesis capacity due to absorption of excess light energy compared to the photosynthetic capacity.

Phytosociology; the study of the characteristics, classification, relationships and distribution of plant communities.

Pinnatifid leaf; a leaf blade that is deeply lobed with lobes that cut at least half way in from the margin to the mid-rib

PPFD; photosynthetic photon fluence (flux) density, a measure of photosynthetically active radiation.

Primer dimer; product formed during PCR due to interaction between the two primers.

Polyphagous; feeding on many sources of food.

Recombination; crossing over of DNA during meiosis between paired chromosomes, resulting in offspring having different combinations of genes from their parents.

Ruderal; an early coloniser of disturbed land (natural or anthropogenic disturbance).

Segetal; weeds of cereal fields, or more generally applying weeds of arable land.

Shotgun sequencing; method of sequencing long strands of DNA by the process of splitting the DNA into shorter sections which are then sequenced, this is repeated multiple times after which the sequences can be reassembled by computer programmes by way of recognising overlapping ends from the shorter sequences.

Tapetum; a nutritious layer of cells that surrounds the spore mother cells (which give rise to the pollen) in the anthers.

Thin elliptical petal; oval with short or no point

Travertine; porous calcite, CaCo3, deposited from solution in ground or surface waters.

Vernalization; exposure to a period of cold treatment that initiates or hastens a plant to flower.