The responses of *Arabidopsis lyrata* ssp. *petraea* to environmental stimuli

Chloë Elizabeth Thompson

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
Faculty of Biological Sciences
Institute of Integrative and Comparative Biology
Centre for Plant Sciences

September 2010

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

Firstly I would like to thank my supervisor Professor Peter Urwin for his invaluable guidance and support throughout my Ph.D., and for introducing me to the world of molecular biology. Thank you also to my second supervisor Professor William Kunin for his ecological enthusiasm and advice. In addition, I would like to thank past and present members of the HJA/PEU lab and the WEK ecology group for their friendship and support.

The following people deserve special thanks; Dr Philippine Vergeer for sharing her wealth of *A. l. petraea* knowledge; Jennie Hibbard for her invaluable technical support; Professor Howard Atkinson for his statistical words of wisdom; Siân Giddy for her IT wizardry; Dr Nathan Watson-Haigh for assistance with analysis of the microarray data; Professor Roger Butlin of the University of Sheffield for advice; Lance Penketh for assistance in the Plant Growth Suite; Martin Lappage for his friendship and horticultural support; Laura Davies for patiently answering many of my stupid questions and for her dedication to morning tea and biscuits; Nicky Atkinson for tea, cake, and being cheerful; Mark Goddard for his editorial skills and an unfailing commitment to the 14.56 tea break, and; my friends Lucy, Jane, and Steph of the James Farrell Concert Band for their weekly encouragement and word count targets. I am also very grateful to Dr Laura Jones for her assistance with printing and binding of the final thesis. Particular thanks must go to Dr Catherine Lilley for her continued support, encouragement, advice, and skills in panic prevention and alleviation. Finally I would like to thank Shaun Dowman for his support and faith in my abilities, and for always looking after me.

This Ph.D. project was funded by the Biotechnology and Biological Sciences Research Council.

Abstract

Arabidopsis lyrata ssp. *petraea* has a fragmented distribution across Northern Eurasia, and its geographically isolated populations experience varying environmental conditions throughout the species' range. *A. l. petraea* is consequently an appropriate test system to study the effects of abiotic stress on species' distributions and for investigation of local adaptation.

The main objective of this study was to investigate whether or not geographically isolated populations of *A. l. petraea* differ in their responses to environmental stimuli and thus demonstrate potential for local adaptation. Under controlled environmental conditions, populations from the British Isles and Norway demonstrated the highest average growth rates whilst the lowest rates were observed in Swedish populations. Populations also exhibited significantly altered growth rates under imposed hot and cold stress conditions than in their corresponding no-stress controls. The percentage of plants that flowered was significantly higher in Norwegian populations than for those from Sweden, Iceland, and the British Isles

Molecular techniques were employed to investigate differential gene expression in populations of *A. l. petraea* from Norway, Sweden, Iceland, and the British Isles, both prior to and following exposure to cold temperature stress. Extent of up-regulation of expression following exposure to chilling differed for a number of genes including CCA1, NAC2, and SKIP16. Microarray analysis identified 207 genes that were differentially expressed between two focal populations, Leitrim (Ireland) and Helin (Norway), irrespective of temperature treatment ($P < 0.05$, fold change >2). Approximately 6% of the *A. l. petraea* transcriptome was responsive to acclimation whilst expression of 75 genes, many annotated as unknown function, changed following exposure to freezing.

The identified phenotypic variation is indicative of local adaptation, and populations from throughout the species' range may thus harbour unique genetic variation. Further investigation of genes identified as important in plant responses to cold temperatures may facilitate crop improvement through GM technologies.

Contents

List of Figures

List of Tables

Abbreviations

2-iP	2-isopentenyladenine
4-MU	4-methylumbelliferone
ABA	abscisic acid
ANOVA	analysis of variance
AS	artificial selection
BLAST	basic local alignment search tool
bp	base pair
$CaCl_2$	calcium carbonate
CDM	culture dilution medium
cDNA	complementary DNA
CIM	callus-inducing medium
C_t	threshold cycle
CV	coefficient of variation
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DRE	dehydration responsive
EDTA	ethylenediaminetetraacetic acid
ELP	expression level polymorphism
F1	first filial generation
GIS	geographic information systems
GUS	ß-Glucoronidase
h^2	heritability
IAA	indole-3-acetic acid
kb	kilobase
LB	Luria-Bertani (growth medium)
LGM	last glacial maximum
LNS	laboratory natural selection
MIAME	minimum information about a microarray experiment
MS	Murashige and Skoog
MUG	4-methylumbelliferyl ß-D-glucuronide
NaCl	sodium chloride
NaOH	sodium hydroxide

PCA	principal component analysis
PCR	polymerase chain reaction
qRT-PCR	quantitative reverse-transcriptase PCR
R	response to selection
RMA	robust multi-array average
RNA	ribonucleic acid
RNase	ribonuclease
S	selection differential
SD	standard deviation
SDS	sodium dodecyl sulphate
SE	standard error
SOB	super optimal broth
SOM	shoot overlay medium
TAIR	the *Arabidopsis* information resource
TB	transformation buffer
TBE	tris-borate-EDTA
UV	ultraviolet
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
X-Gluc	5-bromo-4-chloro-3-indolyl-glucuronide

1. General Introduction

1.1 Species ranges

Each species has its own individual distributional limits in space, referred to as its geographic range. Limits to the distributional range of a species must arise as a result of the interaction between processes of birth, death, and dispersal (Holt and Keitt, 2005; Gaston, 2009). The boundaries of species' ranges can remain relatively constant when set by major barriers to dispersal such as coastlines, whilst others can alter frequently due to constant changes in competing ecological forces that act to create these borders (MacArthur, 1972; Brown *et al.*, 1996). The sizes of species' ranges are extremely varied, with *Homo sapiens* and some associated bacterial species occupying some of the largest ranges, whilst a large number of narrow endemic species occupy tiny ranges (Gaston, 1996). Examples of narrow endemic species include the perennial plant *Cochlearia bavarica* that only occurs in two regions of Bavaria, Germany (Paschke *et al.*, 2002) and the caddisfly *Wormaldia tagananana* that is restricted to eight catchment areas in the Canary Islands (Kelly *et al.*, 2002). Although range size varies considerably between species, more closely related species tend to have increasingly similar range sizes (Jablonski, 1987; Brown *et al.*, 1995). There are correlations between extent of range occupied and aspects of species' ecology such as feeding habits; for example Rapoport (1982) illustrates that Central and North American insectivorous bats are more widespread than those that feed on fruits and nectar. It has been proposed that the wider range of climatic conditions experienced by high-latitude species has resulted in more northerly species occupying larger ranges. These species may have evolved broader climatic tolerances thus enabling their persistence over a greater geographical area (Stevens, 1989). This hypothesis has, despite some supporting studies (Letcher and Harvey, 1994), been largely discredited as latitudinal patterns in range size have not been realised in a range of species including New World woodpeckers (Husak and Husak, 2003), marine teleost fish (Rohde *et al.*, 1993), and marine molluscs (Roy *et al.*, 1994).

In recent years there has been an increased interest in the study of species ranges and their borders or margins. One reason for this is the development of technology to aid such research including geographic information systems (GIS), remote sensing, and databases that enable information to be compiled and more readily compared (Brown *et al.*, 1996; Holt, 2003; Holt and Keitt, 2005). Holt and Keitt (2005) also suggest that

researchers have an increased appreciation of space and spatial processes, and a requirement to be able to respond to rapid changes that are taking place throughout world environments. Research into species ranges can be applied to a range of ecological issues including responses of species to climatic change (Parmesan and Yohe, 2003; Root *et al.*, 2003) and the effect of land use change on endangered species (Channell and Lomolino, 2000).

1.2 The edge of the range

A wide variety of forces act to determine species ranges and the position of their borders. These include features of the geographic environment, both biotic and abiotic ecological factors, and evolutionary forces. It is essential that these forces are not viewed as mutually exclusive alternatives (Gaston, 2009).

1.2.1 Ecological factors

Species ranges are thought to be constrained by ecological factors including climatic variables and interspecific interactions such as competition and dispersal (Brown *et al.*, 1996; Kirkpatrick and Barton, 1997; Case and Taper, 2000; Holt, 2003; Baack *et al.*, 2006). An understanding of the ecological factors influencing distributions provides the context for developing evolutionary studies (Baack *et al.*, 2006).

Plants in particular, due to their sessile nature, may be restricted in terms of distribution by their physiological responses to climatic variables. For instance, significant climatic differences (e.g. frequency of precipitation and average temperature) were reported between regions where the loblolly pine (*Pinus taeda* L.) persists and regions outside of its range where establishment appears unfeasible (Hocker, 1956). Experimental transplantations of the annual cocklebur (*Xanthium strumarium)* have demonstrated its ability to grow successfully in regions beyond the northern border of its natural range without manipulation of reproductive timing, however, transplanted seedlings were unable to reproduce prior to the onset of frost (Griffith and Watson, 2006). Declines in seed production in populations at range peripheries in three species of the perennial *Circium* are thought to be due to enhanced environmental stresses at the range margins that subsequently result in limited energy availability for reproduction. The founding of new range-expanding populations will be dependent on dispersal of viable seed, and levels of seed production may be too low to enable new populations to establish (Jump

and Woodward, 2003). These examples suggest that climatic variables are fundamental to distribution, and that peripheral populations would need to evolve in order to expand their distributional ranges.

Climatic factors are not the only variables that influence the distribution of plant species. Abrupt distributional boundaries in the native Californian annual *Gilia tricolour* appear to be maintained by the prevention of seedling emergence through accumulation of litter; increased litter accumulation can be attributable to large differences in soil chemistry between core and marginal zones (Baack *et al.*, 2006). It is unlikely, however, that only abiotic factors will differ as a species expands its range. Differing biotic factors, such as pathogen species, will be present, and selection for adaptation to these may counteract adaptive responses that may enable further range extension. Inhibition of seedling emergence at population range margins in *Gilia tricolour* is also thought to be attributable to the presence of competition from other more dominant species that are better suited to conditions in the marginal habitats of this species (Baack *et al.*, 2006).

1.2.2 Evolutionary perspective

Identification of environmental factors that are responsible for limitations to species distributions leads to the expectation that natural selection should result in evolutionary adaptations to these factors, primarily in marginal populations. Such adaptations will result in the ability of the species to persist under new environmental conditions, and consequently over a wider geographic area (Holt, 2003; Griffith and Watson, 2005, 2006). The lack of evidence for continual expansion of species ranges has resulted in the development of a variety of theories as to why natural selection appears unable to improve adaptation in peripheral populations and subsequently enable current distributional limits to be overcome. Two of the main hypotheses are discussed below.

1.2.2.1 Lack of genetic variation in marginal populations

It is hypothesised that marginal populations may be genetically depauperate and that this may result in the inability of such populations to evolve adaptations to locally prevailing environmental conditions. Genetic drift is the effect resulting from chance survival of alleles; allelic changes in a small number of individuals will have a proportionally larger overall effect in populations consisting of fewer individuals than it would in larger populations. Many marginal populations have small effective population sizes due to their fragmented nature and their fluctuating sizes, and these

characteristics consequently put such populations at a greater risk of losing allelic diversity (Glemin *et al.*, 2003; Whitlock, 2004). Alleles that may have been beneficial for adaptation may therefore be lost before they have the opportunity to be selected, and thus local adaptation is prevented. Populations at range boundaries may have, in the past, experienced persistent natural selection due to exposure to stressful environments, and subsequently such populations may have depleted their genetic variation and have reached the limits of their adaptive potential (Hoffmann and Parsons, 1991; Case and Taper, 2000; Eckert *et al.*, 2008). Genetic variation can be replenished through the process of gene flow between populations. Populations at the centre of a range can receive gene flow in all directions from surrounding populations, while gene flow into peripheral populations can only occur in one direction. Problems of habitat fragmentation and resultant periods of small population size are therefore likely to have a greater impact on genetic diversity in peripheral populations, isolating them from gene flow, and causing further reductions in effective population size (Hoffmann and Parsons, 1991; Lonn and Prentice, 2002).

In support of this hypothesis, peripheral populations of the perennial herb *Gypsophila fastigiata* were found to have significantly lower levels of within-population genetic diversity than central populations (Lonn and Prentice, 2002). Core populations of the south east Asian mangrove *Avicennia marina* were found to possess higher allelic richness and a lower extent of inbreeding than populations at the northern and southern distributional limits (Arnaud-Haond *et al.*, 2006). These patterns of genetically depauperate marginal populations in plants have not been observed in all species. For example, within-population genetic diversity in a number of small fragmented populations of the orchid *Orchis purpurea* was reported to be relatively high. There are relatively few studies that compare genetic variation for quantitative traits between core and marginal populations. One such study investigated desiccation resistance in the fly *Drosophila serrata*; populations from the southern margins of the range were able to respond to artificial selection for desiccation resistance yet they displayed lower realised heritability values for this trait than core populations (Blows and Hoffmann, 1993). Traits involved in limiting the distribution of species are expected to have lower realised heritability values than more neutral traits as genetic variation will have been selected against in favour of the optimal trait value. A review of 134 empirical studies found that, on average, within-population diversity decreased and among-population differentiation increased from the centre of the range to the periphery (Eckert *et al.*,

2008). The authors, however, do warn against forming generalisations from these results, highlighting that in many cases differences in genetic diversity were not large and there were both biogeographical and taxonomic biases in the data available (Eckert *et al.*, 2008).

1.2.2.2 Gene flow counteracts local adaptation

Gene flow is a vital process involved in maintaining both genetic and phenotypic homogeneity within a species and can occur in plants via both seed and pollen flow (Hu and He, 2006). This homogeneity occurs to the detriment of local adaptation and may negatively affect a species' ability to persist or expand its geographic range (Haldane, 1956; Lenormand, 2002; Alleaume-Benharira *et al.*, 2006). In this hypothesis, marginal populations are thought of as sinks, receiving immigrants from core populations in a source-sink system (Arnaud-Haond *et al.*, 2006). The effects of natural selection on the gene pool of marginal populations will, in effect, be counteracted by the introduction of alleles representative of core populations. The swamping of marginal populations with maladapted genotypes may thus prevent the divergence of central and marginal populations (Butlin *et al.*, 2003; Arnaud-Haond *et al.*, 2006). Immigration from marginal to core habitats can also occur but, due to the tendency for core populations to be larger, the impact of such gene flow on local adaptation is likely to be very small (Kawecki, 2008).

This theory has been modelled quantitatively, illustrating how dispersal, natural selection, and adaptive trade-offs interact to produce stable range limits, and also how the isolation of peripheral populations from gene flow results in these populations attaining their ecological optimum (Garcia-Ramos and Kirkpatrick, 1997; Kirkpatrick and Barton, 1997). The model by Kirkpatrick and Barton (1997) suggests that the effect of gene flow can be great enough to prevent the spread of a species into a suitable habitat. The biological relevance of this model, however, cannot be verified until data for all of the model parameters are available for a single species (Kirkpatrick and Barton, 1997). A long-term study of blue tit (*Cyanistes caeruleus*) populations in southern France provides biological evidence for the maladaptive effects of gene flow (Blondel *et al.*, 2006). Populations persist in two different habitats, deciduous and evergreen forests, deemed as core and marginal respectively. Breeding success is dependent on synchronisation of chick emergence with the peak in food supply and, despite there being a difference of one month in this peak between habitats, birds from both habitats

lay their eggs at the same time. The evergreen habitat has been demonstrated as a sink, receiving net gene flow from the deciduous habitat, and thus the genetically determined breeding date is unable to evolve to match food supply in this habitat (Blondel *et al.*, 2006). Conversely, there is some evidence for local adaptation in marginal populations despite continuous gene flow; small populations of *Agrostis capillaris* were able to adapt to zinc-polluted habitats despite the close proximity of large core populations persisting on unpolluted land (Al-Hiyaly *et al.*, 1993).

A striking contrast between the two hypotheses outlined above is that the first requires very little or no gene flow into peripheral populations whilst the second is based on an assumption that significant levels of gene flow occur from core to marginal populations. Adaptation to prevailing conditions may actually be most feasible in the zone between central and peripheral populations. In this zone there may be adequate gene flow to maintain sufficient genetic variation yet not enough to overwhelm populations and counteract selection for the most beneficial genotypes. Ecotones are zones of ecological transition and may potentially function to generate and maintain biodiversity (Smith *et al.*, 1997); the persistence of gene flow and divergent pressures of natural selection across these zones may result in high levels of genetic diversity within a species. For example, the ecotone between desert and Mediterranean regions is located towards the margins of the Chukar partridge's (*Alectoris chukar*) continuous range in Israel. Although geographical distances between the partridge populations under investigation were small, within-population diversity was found to be highest within this ecotone thus making it a priority for conservation (Kark *et al.*, 1999). Theoretical work by Barton (2001) provides support for the role of gene flow in facilitating adaptation, in contrast to Kirkpatrick and Barton's (1997) model in which gene flow was found to limit the species' range. In the later model gene flow was found to increase genetic variance and as a result the species was able to spread indefinitely in the absence of other barriers to dispersal (Barton, 2001).

1.3 Evidence for local adaption in plants

Studies of adaptation in natural populations can provide supporting evidence for the hypotheses outlined for the persistence of species range margins. Controlled evolutionary experiments, however, have the potential to directly test factors affecting adaptation to marginal habitats (Kawecki, 2008). A number of reciprocal transplant experiments have been performed involving populations of a given species from widely

differing habitats. For example, populations of *Boechera holboellii*, a perennial relative of *Arabidopsis thaliana*, originating from sites with contrasting hydrology were transplanted to field sites at which either wet or dry soil conditions prevailed; each population demonstrated significantly greater survival in its native habitat indicating that populations were locally adapted in terms of this trait (Knight *et al.*, 2006). Similarly, two closely related *Mimulus* (monkeyflower) species that naturally persist at differing altitudes exhibited greatest survival and reproduction in environments most like those at their range centre in a reciprocal transplant experiment (Angert *et al.*, 2008). Hybrids, produced through crossing of the two sister species, demonstrated intermediate reproduction and survival in both altitudinal environments. A further common-garden experiment involving these hybrids enabled a more detailed investigation into the role of adaptive trade-offs in the limitation of geographic ranges. Hybrids that had been selected in either high or low altitude environments were outperformed by unselected controls when grown in a common environment for which they were not selected. This observation indicates that adaptation to a particular environment entails a cost to adaptation in other environments (Angert *et al.*, 2008).

1.4 The impact of climatic change on species ranges

Recent ecological research has focused on developing an understanding of how species, and also ecological interactions, respond to climatic change (Thomas *et al.*, 2004). It has been confirmed that global climates are warming (IPCC, 2001; 2007) and this will have numerous effects on the world's species, with evidence gathering for both expansion and contraction of species ranges (Butlin *et al.*, 2003). The effects of climatic change on species' ranges has been naturally difficult to study because such changes occur over an extended time period and detailed distribution records for pre-1800 are rarely available (Thomas, 2010). Scientists have utilised the fossil record for investigations into changes in species' distributions and such paleoecological studies have revealed that both insect and plant species were able to track past changes in climate through alterations to their ranges (Coope, 1995; Pitelka, 1997). Mathematical modelling has also been used to predict the consequences of global climatic change on species distributions. Although such studies avoid the problems of time (e.g. time over which climatic changes occur in the field), the lack of field-collected data limits the extent to which they truly represent responses of real flora and fauna. The consequences of climatic change on Mexican fauna have been investigated through the use of models, and varying the dispersal abilities of species produced a range of

different predicted outcomes including extinction, range reduction, and range fragmentation (Townsend-Peterson *et al.*, 2002). A broader-based meta-analysis revealed that in species that have demonstrated shifts in their ranges, a significant proportion of these have shifted in the direction expected in the face of climatic change (Parmesan and Yohe, 2003).

Recent field studies of the changing distributions of a wide range of flora and fauna provide substantial evidence for the movement of species ranges northwards and to areas of higher elevation; examples include marine fish (Perry *et al.*, 2005), amphibians (Pounds *et al.*, 2006), birds (Thomas and Lennon, 1999), butterflies (Parmesan *et al.*, 1999; Hill *et al.*, 2002; Wilson *et al.*, 2005) and the invasion of *Betula pubescens* ssp.. *tortuosa* (mountain birch) above the former level of the treeline in Sweden (Truong *et al.*, 2007). Evidence for retraction of ranges at species' warmer southern margins, however, is less apparent (Thomas *et al.*, 2006). This could be due to the temporary persistence of species in isolated climatic refugia or may reflect a concentration of survey effort at northern range margins. For example, thirty years of regional warming has resulted in upwards shifts of over 200 metres in elevational range for many butterfly species in Sierra de Guadarrama, central Spain and, as this effect is systematic throughout the study species, it is indicative of a consistent widely-acting cause such as climate (Wilson *et al.*, 2005). Recent rapid anthropogenic climate change (the past thirty years) has been treated by Thomas (2010) as a large-scale experiment; using both distributional data and the published literature, the range boundaries for the majority of included terrestrial and freshwater species were found to have shifted in the direction expected if climate change were to be the cause (Thomas, 2010).

1.5 Responses of plants to abiotic stress

Plants are exposed to a wide range of abiotic stresses including temperature, drought, and soil salinity that they are unable to avoid due to their sessile nature. Although plants have evolved a wide range of tolerance mechanisms, such stresses are a primary cause of crop losses worldwide as they can place severe limitations on plant growth (Kant *et al.*, 2008; Agarwal and Jha, 2010). Following exposure to abiotic stress plants undergo a wide range of biochemical and physiological changes as well as alterations to molecular and cellular processes. The responses of plants to abiotic stresses are controlled by signalling networks and these interconnected networks regulate the expression of specific sets of genes (Papdi *et al.*, 2009). Tolerance to abiotic stress is

governed by multiple loci and responses to stress are, as a result, complex (Angarwal and Jha, 2010). The response of plants to abiotic stresses has received much attention in the scientific community not only due to the desire to understand the tolerance mechanisms that have evolved in plants but also due to the potential wider applications of such knowledge; this could be of huge importance in the development of strategies to improve the environmental stress tolerance of commercially important crop plants (Thomashow, 1998; Fowler and Thomashow, 2002; Papdi *et al.*, 2009). Knowledge of responses to plant abiotic stress will be of increasing importance in the face of predicted global climatic change, and the growing pressures for increased food production. The model plant *A. thaliana* has been central to advances in plant stress biology despite its own limited ability to tolerate stress. This can be attributed to the availability of the *A. thaliana* genome sequence and the development of a wide range of genomics tools for this species (Papdi *et al.*, 2009).

1.5.1 The responses of plants to cold stress

Temperature perception in plants is of particular importance as temperature influences the regulation of development events and can also affect plant survival (Penfield and Hall, 2009). As well as perception of absolute temperatures in an environment, plants are also responsive to temperature fluctuations and the accumulation of temperature values over time (Hua, 2009). Global climate change predicts that mean temperatures will rise, yet this will also influence plant responses to cold temperatures as plants evolve sensitivity to cold according to the temperature regime within their environment (Penfield and Hall, 2009). To ensure the appropriate response to temperature, plants must perceive changes in temperature, resulting in the transduction of signals that subsequently evoke appropriate responses. Although much progress has been made in understanding the mechanisms by which plants tolerate temperature, the sensors involved in temperature perception remain largely unknown (Penfield and Hall, 2009).

Cold stress in plants has been studied extensively over the past two decades, following the discovery that exposure to cold temperatures results in changes in the expression of genes within the plant (Guy *et al.*, 1985). There are three stages in the response of plants to low temperatures. Cold acclimation, the first stage, is the process by which plants enhance their tolerance to sub-zero temperatures through prior exposure to low, yet above zero, temperatures (Hughes and Dunn, 1996; Thomashow, 1999). The second stage is referred to as hardening and it is during this phase that the plant's full

tolerance is achieved, for which exposure to sub-zero temperatures is required (Janska *et al.*, 2010). The final stage relates to recovery of the plant following cessation of the cold temperatures and involves tissue thawing, restoration of cell structure, and the resumption of cellular processes (Li *et al.*, 2008). Membrane fluidity, stability of RNA and DNA, and enzyme activity are all affected by low temperatures, and could all potentially have involvement in plant perception of cold (Smallwood and Bowles, 2001). Levels of cytosolic calcium are known to change rapidly upon exposure to cold temperatures (Knight *et al.*, 1991; Knight *et al.*, 1996) and there is growing speculation that this change may be involved in plant cold stress signalling (Xin and Browse, 2000; Chinnusamy *et al.*, 2006; Penfield and Hall, 2009). The plant hormone abscisic acid (ABA) may also act as a signalling molecule during cold temperature exposure (Chinnusamy *et al.*, 2006; Penfield and Hall, 2009), although induction of many genes by low temperature is ABA-independent (Gilmour and Thomashow, 1991); the importance of ABA in cold-acclimation of plants has thus been questioned (Shinozaki and Yamaguchi-Shinozaki, 2000). Protein kinases and phosphatases have also been advocated in cold temperature signal transduction (Chinnusamy *et al.*, 2006).

Although mechanisms for cold temperature perception remain poorly understood, the advancement of genomic techniques has enabled the molecular mechanisms of cold temperature responses to be investigated in detail. The transcriptome in *A. thaliana* undergoes significant changes following exposure to acclimation temperatures. Work in this field has led to the discovery of many cold responsive (COR) genes and regulatory pathways, and it has been estimated that up to 4% of the genome may be affected by cold temperatures (Fowler and Thomashow, 2002). As yet there is no simple model that encompasses the complexity of the signalling pathways involved in cold acclimation (Janska *et al.*, 2010). Genes induced under conditions of stress can be classified into two groups: genes that have functions in stress tolerance and genes that are involved in further regulation of signal transduction and gene expression (Fowler and Thomashow, 2002; Kreps *et al.*, 2002; Seki *et al.*, 2002). Genes with roles in stress tolerance encompass a wide range of functions: these include i) stabilisation of cell membranes to minimise freezing-induced damage; ii) transportation of proline and sugars that may function as osmolytes and thus protect cells from dehydration; iii) cryoprotectant proteins such as heat shock proteins; iv) molecular chaperons that protect enzymes from denaturation; v) prevention of intracellular ice formation; vi) protection of cells from oxidative damage, and; vii) production of primary and secondary

metabolites. Despite considerable research interest in the cold response in plants, the precise function of many cold-responsive genes remains uncertain (Thomashow, 1999).

To date, the most documented cold response pathway is the CBF (C-Repeat Binding Factor) pathway. The promoters of several cold-responsive genes in *A. thaliana* were found to contain a DNA regulatory element known as the CRT (C-Repeat) or DRE (dehydration responsive) element (Yamaguchi-Shinozaki and Shinozaki, 1994). This led to the discovery of three transcriptional activator genes, *CBF1*, *CBF2*, and *CBF3* (also referred to as *DREB1b*, *DREB1c*, and *DREB1a* respectively) that are able to bind to the CRT/DRE element in gene promoters and thus activate their transcription (Stockinger *et al.*, 1997; Liu *et al.*, 1998). Further analysis of the transcriptome has led to recognition of the CBF regulon which refers to genes that are induced by both overexpression of the CBF genes and by cold temperatures (Seki *et al.*, 2001; Fowler and Thomashow, 2002; Kreps *et al.*, 2002; Seki *et al.*, 2002; Vogel *et al.*, 2005). Genes within this regulon span a wide range of functions including transcription factors, cryoprotectant proteins, and genes of unknown function (Fowler and Thomashow, 2002; Vogel *et al.*, 2005). A more recent insight into the CBF pathway has involved the discovery of *ICE1* (*INDUCTION OF CBF EXPRESSION*), a constitutively expressed transcriptional activator that controls expression of *CBF3* (Chinnusamy *et al.*, 2003). ICE1 has also been shown to influence both the basal expression levels and induction of many cold responsive genes (Lee *et al.*, 2005). *HOS1* (*HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE*), a negative regulator of the CBF regulon, functions in the degradation of the ICE1 protein following cold temperature exposure (Lee *et al.*, 2001). Most recently, *SIZ1*, a SUMO (small ubiquitin-related modifier) E3 ligase has been putatively linked to ICE1 stability and activity (Miura *et al.*, 2007; Jan *et al.*, 2009). Components of the *Arabidopsis* CBF cold response pathway are conserved in other plant species including *Brassica napus*, as well as the more distantly related species wheat (*Triticum aestivum*), rye (*Secale cereale*), and tomato (*Lycopersicon esculentum*) (Jaglo *et al.*, 2001; Zhang *et al.*, 2004b).

A number of other pathways are involved in temperature perception and signal transduction during cold acclimation. One of these is the *ZAT12* regulatory pathway. ZAT12, a zinc finger protein, is induced in parallel with the CBF genes and regulates expression of over twenty cold responsive genes. This pathway does appear to interact with the CBF pathway as some genes are regulated by both pathways and induction of

CBF expression is dampened by constitutive *ZAT12* expression (Vogel *et al.*, 2005). The bZIP-mediated ABA-dependent pathway also induces the expression of cold responsive genes (Xin and Browse, 2000; Agarwal and Jha, 2010), and the *ESK1* (*ESKIMO1*) gene has been identified as a novel regulator of freezing tolerance (Xin *et al.*, 2007). The response of plants to cold temperature exposure is clearly regulated by a complex gene network and continued research in this field will no doubt lead to the discovery of many more regulatory genes.

1.6 The study species: *Arabidopsis lyrata* ssp. *petraea*

Arabidopsis lyrata (formerly *Cardaminopsis* and *Arabis*) is a close relative of the extensively studied model plant species *Arabidopsis thaliana*. *Arabidopsis lyrata* persists in cold temperate to mild climatic regions of the Northern hemisphere (Clauss and Koch, 2006). Several subspecies of *A. lyrata* have been described, and each occurs within a different geographic range. *A. lyrata* ssp. *petraea* is found in central and northern Eurasia, *A. lyrata* ssp. *kamchatica* is found in eastern Asia and northwest North America, and *A. lyrata* ssp. *lyrata* is found in central and eastern North America (Clauss and Koch, 2006). This study has focused on populations of *A. l. petraea* (Figure 1.1). Unlike the human commensal *A. thaliana*, for which the natural range margins can be difficult to determine, *A. l. petraea* naturally persists in stable geographically isolated populations. This crucifer species forms basal rosettes and occupies a range of habitats including gravel and scree slopes, rock faces, and rocky coastal regions (Clauss and Koch, 2006). These low-competition environments enable the effects of abiotic stressors to be investigated without the influence of competition, a major biotic variable invoked in explaining species distributions. *A. l. petraea* is predominantly self-incompatible, in contrast to *A. thaliana*, and this facilitates population genomic analysis. Reproduction generally requires insect pollination, and the most common pollinators are small bees and flies (Clauss and Koch, 2006). *A. l. petraea* also demonstrates vegetative propagation through rhizomes (Clauss and Mitchell-Olds, 2006).

A. l. petraea has a fragmented distribution in Europe that is in direct contrast to the model species *A. thaliana* (Clauss and Mitchell-Olds, 2006). Much work has been performed aiming to uncover the historical biogeography of this species. Genetic studies support the hypothesis for a central European refugial population during the last glacial maximum (LGM; 20 000 years ago) with subsequent recolonisation of northern

Europe during the post-glacial period (Clauss and Mitchell-Olds, 2006; Muller *et al.*, 2007). A more recent study, encompassing a broader spectrum of *A. l. petraea* populations, indicates that eastern Austria is the centre of genetic diversity within central Europe and is likely to have functioned as a long-term glacial refugium (Ansell *et al.*, 2010). Muller *et al.* (2007) observed a diversity gradient from diverse central European populations to northern Scandinavian and Icelandic populations, suggesting recolonisation of northern Europe from central Europe, but the study by Ansell *et al.* (2010) did not provide support for this. Estimated divergence times for central and northern European populations predate the LGM (Ross-Ibarra et al, 2008); this has led to the speculation that northern populations are derived from an alternative source population or that they represent genetic diversity that is no longer extant in central Europe (Ansell *et al.*, 2010).

A high degree of population genetic structuring exists within *A. l. petraea* and has been detected at continental, regional, and local scales (Clauss and Mitchell-Olds, 2006; Muller *et al.*, 2007; Ansell *et al.*, 2010). This differentiation, the persistence of large ecologically stable populations (Ross-Ibarra *et al.*, 2008), and the recent release of the genome sequence of closely related *A. l. lyrata* (http://genome.jgi-psf.org/Araly1/Araly1.home.html) will further enhance the value of *A. l. petraea* as a model species for studying the ecology and evolution of local adaptation (Mitchell-Olds, 2001). There is growing evidence for divergent selection in distinct populations of *A. l. petraea*. For example, a high degree of among-population variation in trichome production has been observed in Swedish populations and has been attributed to genetic variation in the regulatory gene *GLABROUS1* (Karkkainen *et al.*, 2004; Kivimaki *et al.*, 2007). Maintenance of the polymorphism for trichome production in *A. l. petraea* is thought to be due to spatial variation in plant-herbivore interactions (Løe *et al.*, 2007). Differences in flowering characteristics have also been observed in *A. l. petraea* populations, including considerable variation in time to flowering both within and between populations (Clauss and Koch, 2006; Vergeer *et al.*, 2008) and differential daylength requirements between northern and southern populations (Riihimaki *et al.*, 2005; Sandring *et al.*, 2007).

A. l. petraea is a perennial, is predominantly self-incompatible, and is found in naturally disturbed sites of low competition. Extensive field data (*P. Vergeer*, unpublished) demonstrates the considerable variation in climatic conditions that exists throughout the

north west European range of this species; this has been illustrated through principal component analysis (Figure 1.2). Of the populations studied, Icelandic populations are exposed to the coldest winter conditions, the British populations experience the mildest winter conditions, and Swedish and low altitude Norwegian populations have the hottest summers (Figure 1.2). Iceland, where *A. l. petraea* is widespread and common, forms the northernmost range margin of this subspecies. The species is locally common in Norway and Sweden yet populations are more fragmented than in Iceland. Welsh and Irish populations constitute the southern range margin of *A. l. petraea*'s north west European distribution and populations in these countries are scarce and isolated.

1.7 Project overview

This work set out to establish insights into the extent to which *A. l. petraea* populations are adapted to their local environmental conditions. This was performed using both ecological and molecular approaches, enabling differences in plant growth parameters and levels of gene expression to be identified between populations.

Key objectives of this work are to:

a) Examine responses to temperature stress in *A. l. petraea* and to determine whether or not geographically distinct populations differ in their abilities to tolerate extremes of temperature;
b) Identify genes with a potential role in the response of plants to both chilling and freezing temperatures;
c) Identify genes with potential roles in the adaptation of *A. l. petraea* to local environmental conditions.

Chapter 2 details those methods that relate to more than one section of the thesis, and should therefore be used as a reference when reading the individual results chapters, whilst methods specific to only one chapter are located within the relevant chapter. Chapter 3 details variability in life history traits between populations of *A. l. petraea* that were exposed to a range of temperature conditions and investigates the potential for differential responses to selection in these populations. In Chapter 4, differential expression of a pre-selected set of genes between thirteen *A. l. petraea* populations in both control conditions and in response to chilling is investigated. Microarray analysis, presented in Chapter 5, involved two *A. l. petraea* populations: Helin in Norway and Leitrim in Ireland. Genes responsive to cold acclimation and a freezing shock were

identified. In addition a set of genes that have potential roles in local adaptation were identified; these genes were responsive to acclimation and also differentially regulated between the populations under investigation. A subset of these genes were characterised further and results are presented in Chapter 6. This involved the generation of transgenic promoter:reporter plants that were used in quantitative GUS assays, and enabled the role of population-specific promoters in regulating gene expression to be explored.

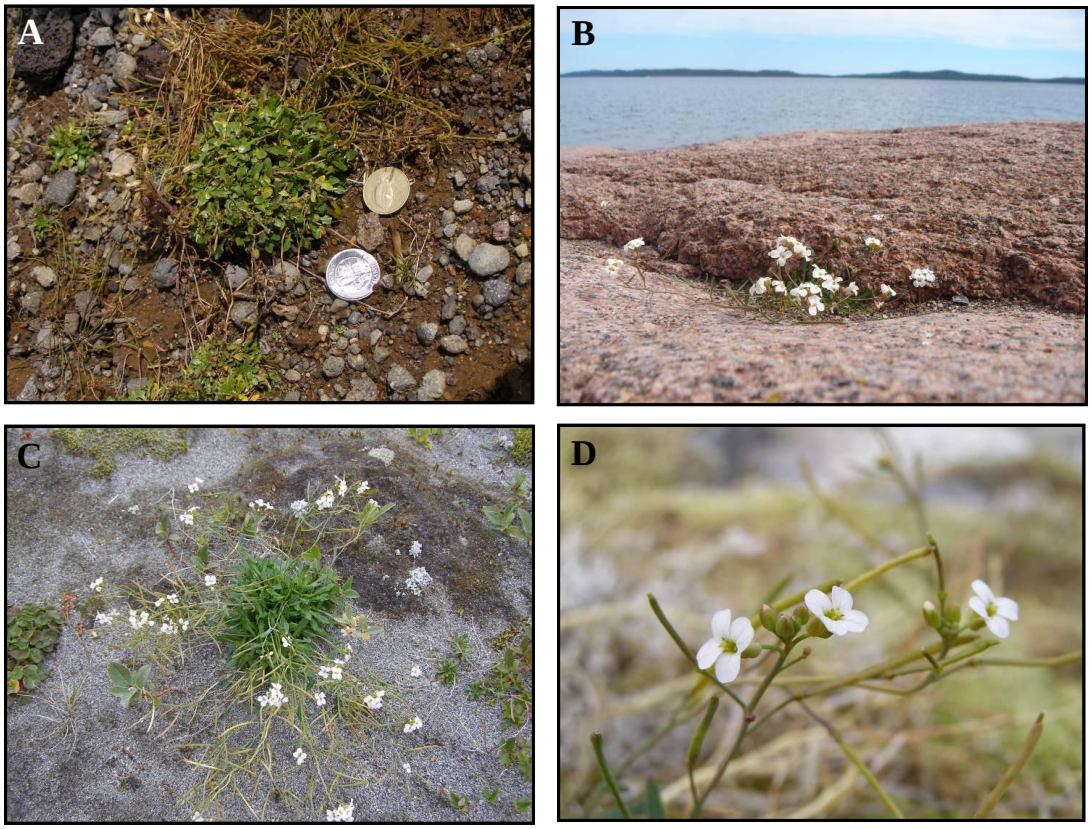


Figure 1.1 ***Arabidopsis lyrata*** **ssp.** ***petraea*****.** A) Near Gullfoss, eastern Iceland; B) coastal habitat along the Gulf of Bothnia, Sweden (Photograph: Dr. P. Vergeer); C) Spiterstulen, southern central Norway: the population persists beside the river Visa on fluvial and glacial depositions; D) Plant in flower at Spiterstulen, Norway.

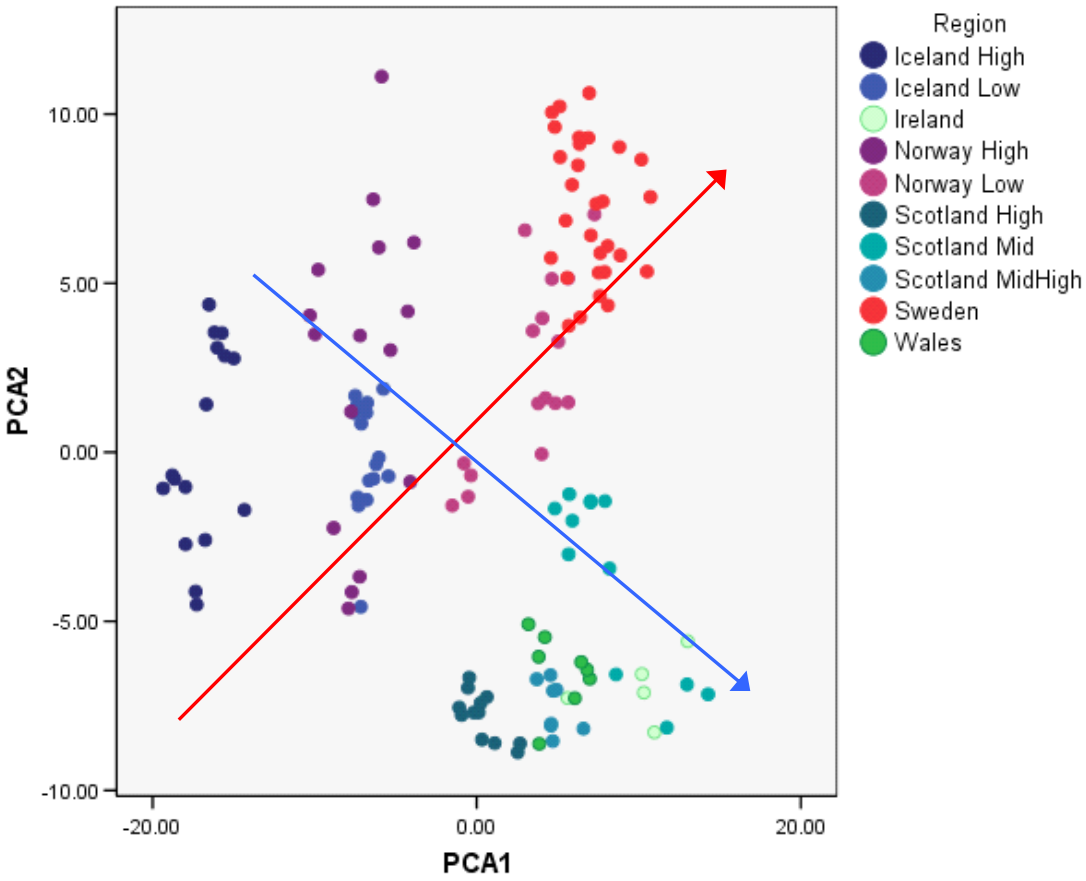


Figure 1.2 Climatic variation between *A. l. petraea* populations. A principal component analysis (PCA) was performed using extensive field collected climatic data. Coloured circles represent individual populations and are colour-coded according to their country of origin (see key). This figures shows axis PCA1 and PCA2 that together represent 75% of the total variation. The red arrow represents increasing summer temperatures (approximate) and the blue arrow represents increasing winter temperatures (approximate). Climates differ considerably between geographic regions (Data and figure from Dr. P. Vergeer).

2. General Materials and Methods

This chapter provides details of general materials, methods, and information that relate to one or more sections of work. Methods specific to particular aspects of work are given in the appropriate chapter.

2.1 Plant species and bacterial strains used

2.1.1 Plant species

- *Arabidopsis thaliana* ecotype Col-0
- *Arabidopsis lyrata* ssp. *petraea* (original seed collected from field sites across Europe). Populations used within each chapter are detailed in Table 2.1.

2.1.2 *E. coli* strains

- *E. coli* DH5α

2.1.3 *Agrobacterium tumefaciens* strains

- *Agrobacterium tumefaciens* GV3101

2.2 Antibiotic stock solutions

Antibiotics were prepared as concentrated stock solutions and stored at -20^{o}C prior to use. Table 2.2 details the preparation of all stock solutions and the working concentrations.

Antibiotic	Solvent	Stock solution	Working concentration
Ampicillin	Distilled water	50 mg ml^{-1}	50 µg ml^{-1}
Cefotaxime	Distilled water	250 mg ml^{-1}	250 µg ml^{-1}
Kanamycin	Distilled water	50 mg ml^{-1}	50 µg ml^{-1}
Rifampicin	Methanol	10 mg ml^{-1}	50 µg ml^{-1}

Table 2.2 Preparation of antibiotic stock solutions.

					Chapter 3			Chapter 4		Chapter 5	Chapter 6
Population Code	**Population Name**	**Latitude**	**Longitude**	**Altitude (m)**	**Selection Exp Gen1**	**Selection Exp Gen 2**	**Selection Exp Gen 3**	**Acclimation Experiment**	**Acclimation experiment additional replicates**	**Microarrays**	**Promoter region analysis**
Iceland1	Raudholar	64.05N	21.44E	90							
Iceland2	Nyidalur North	64.43N	18.07E	875							
Iceland3	Fagrifoss	61.18N	18.58E	635							
Sweden1	Berghamn	62.49N	18.16E	2							
Sweden2	Sjoviken	62.35N	18.01E	2							
Sweden3	Stubbsand	63.12N	18.57E	1.5							
Sweden4	Notsand	62.36N	18.03E	3							
Norway1	Hjelle	61.22N	7.55E	193							
Norway2	Saebo	61.06N	7.25E	4							
Norway3	Spiterstulen	61.39N	8.25E	953							
Norway4	Bovra	61.46N	8.25E	14							
Norway5	Laerdal Island	61.05N	7.31E	493							
Norway6	Smadalen	61.02N	8.33E	1061							
Norway7	Helin	61.02N	8.39E	1130							
Scotland1	River Dee	57N	3.24E	336							
Wales1	Snowdon West	53.05N	4.52W	672							
Ireland1	Leitrim	54.22N	8.22W	355							

Table 2.1 The *A. l. petraea* populations. This table details the geographical location of all populations used in this investigation and the code assigned to each population. Coloured regions indicate inclusion of a population within a results chapter.

2.3 General molecular biology methods used in DNA manipulation

2.3.1 Tissue Collection: *A. l. petraea* and *A. thaliana*

Leaf tissue was collected by snap freezing in liquid nitrogen. Samples were stored in microcentrifuge tubes at -80^{o}C until required.

2.3.2 Genomic DNA Extraction

Genomic DNA was extracted from both *A. thaliana* and *A. l. petraea* leaves using fresh young leaf tissue to a maximum of 100 mg per sample. Tissue was ground in Tris buffer (100 mM Tris pH 8; 50 mM EDTA pH 8; 500 mM NaCl) containing 0.2% (v/v) ß-mercaptoethanol. Cell lysis was achieved by adding 100 µl of a 10% SDS solution, mixing by inversion, and incubating at 65^{o}C for ten minutes. Ice cold 3 M potassium/5 M acetate solution was added (250 µl) and mixed immediately by inversion before incubation on ice for 20 minutes. Cell debris was pelleted by centrifugation in a microcentrifuge at full speed for ten minutes. Supernatant was transferred to a fresh microcentrifuge tube, 0.5 ml of isopropanol added, and samples incubated at -20^{o}C for 30 minutes to precipitate the DNA. DNA was pelleted in a microcentrifuge at full speed for 15 minutes and then washed with 70% (v/v) ethanol. Pellets were allowed to air dry before slow rehydration in 50 µl of TE buffer (10 mM Tris; 1 mM EDTA). RNase A (Qiagen) was added to samples to a concentration of 200 µg ml^{-1}.

2.3.3 Purification of DNA by phenol : chloroform extraction

Solutions of DNA were purified by the addition of an equal volume of phenol/chloroform (1:1, v/v). The DNA phenol/chloroform mixture was then mixed by vortexing. The organic layer of phenol/chloroform and the aqueous layer were then separated by centrifugation at full speed in a microcentrifuge for three minutes. The purified DNA, contained within the aqueous layer, was then removed to a fresh tube.

2.3.4 Extraction of Total RNA

Total RNA was prepared from *A. l. petraea* leaf tissue using the Qiagen RNeasy Plant Mini Kit, according to the manufacturer's protocol. Instead of grinding plant material in liquid nitrogen (as per the protocol), 450 µl buffer RLT (supplied in the kit) was added directly to whole plant material in a microcentrifuge tube with RNase-free sterile sand, and this was ground using a plastic pestle. Buffer RLT provides denaturing conditions required for cell lysis. A maximum of 100 mg leaf tissue was used in each

reaction. Disrupted cells were microcentrifuged at full speed through a QIAshredder spin column to remove cell debris and homogenise the lysate. Ethanol was added to the supernatant to create column-binding conditions and transferred to an RNeasy Spin Column. Following microcentrifugation the optional on-column DNase digestion was performed. The column was then washed with buffers RW1 and RPE to remove contaminants, and 30-50 µl of RNase-free water was used to elute total RNA. This was then stored at -80^{o}C. An additional in-solution DNase digestion was performed according to manufacturer's instructions (Qiagen) if RNA was to be used for applications sensitive to the presence of small quantities of DNA such as quantitative PCR.

2.3.5 Estimation of nucleic acid concentration

A NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies) was used to estimate nucleic acid concentration in 1 µl samples. Standard deviations of <2% or 1.5 ng µl^{-1} are typical (NanoDrop ND-1000 Spectrophotometer V3.3 User's Manual, 2005).

2.3.6 Measurement of RNA quality

The quality of RNA samples was inspected on an Agilent Technologies 2100 Bioanalyser (Agilent Technologies) using an RNA 6000 Nano Assay kit (Agilent Technologies) according to the manufacturer's instructions.

2.3.7 Reverse transcription of RNA

Reverse transcription was performed with 1 µg of total RNA, denatured at 65^{o}C for five minutes in the presence of 0.5 mM each dNTP and 500 nM oligo(dT) primer, followed by ice quenching. The 20 µl reaction was reversed transcribed using SuperScript™ II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. The reaction was heat inactivated for five minutes at 70^{o}C. Negative controls were performed for each sample, replacing the enzyme with water.

2.3.8 PCR Amplification

A standard PCR reaction contained 1 µl template DNA and concentrations of all other reagents as detailed in Table 2.3 to a final volume of 50 µl per reaction. Reactions were then subjected to a PCR cycle using a TProfessional thermocycler (Biometra). The

Reaction component	Final concentration	Volume (µl)
10x reaction buffer	1x	5
dNTP mix	0.2 mM	0.5
$MgSO_4$	1.5 mM	1.5
Forward primer	1 µM	0.5
Reverse primer	1 µM	0.5
Template DNA	As required	1
Enzyme (BIOTAQ™ Red DNA Polymerase)	1 unit	1.5
Sterile, distilled water	-	To 50

Table 2.3 Components of the PCR reaction. This table details all components used in a standard PCR reaction of 50 µl.

	Temperature ^{o}C	Time	Number of Cycles
Activation	94	4 m	1
Denaturation	94	40 s	Variable
Annealing	55 - 63	1 m	Variable
Extension	72	1 m per kb	Variable
Final Extension	72	10 m	1

Table 2.4 The PCR cycle. This table details the standard PCR cycle that was used. The extension time was variable depending on the desired product size.

PCR cycle used is outlined in Table 2.4. Negative control PCR reactions containing water instead of template were included to ensure specificity of any amplified products. Following PCR amplification, 10 µl of product combined with 2 µl of loading buffer (Fermentas) were electrophoresed on agarose gels to ascertain whether or not amplicons of the expected size had been produced.

2.3.9 Agarose gel electrophoresis

DNA samples were electrophoresed in Tris borate EDTA (TBE)-agarose gels of 1-1.5% (w/v) in 1x TBE buffer (0.1 M Tris; 83 mM Boric acid; and 1 mM EDTA) at 60-80 volts for up to one hour. Ethidium bromide was added to the molten gel prior to pouring into the tray to enable visualisation of DNA under UV light. Loading buffer (Fermentas) was added to the DNA samples before loading into the gel (5:1 ratio DNA:buffer). Loading buffer was bromophenol blue-based. Size of the DNA molecules was estimated by comparing to λ DNA/*Pst* I marker (Fermentas) or 100 bp DNA ladder (Invitrogen).

2.3.10 Gel Extraction

Gel extraction was performed using the QIAquick Gel Extraction kit (Qiagen). DNA fragments were excised from agarose gels using a sharp scalpel. A maximum of 400 mg of gel was used per reaction. Three volumes of buffer QG were added to one volume of gel in a microcentrifuge tube to ensure optimal pH (≤ 7.5) for DNA binding. Following incubation at 50^{o}C for ten minutes, one gel volume of isopropanol was added to the sample to increase DNA yield and transferred to a QIAquick spin column and centrifuged for one minute at 11 300 x *g*. Buffers QG and PE were used to wash the spin column membrane; the column was left to stand for five minutes following addition of buffer PE to ensure removal of salts. DNA was eluted in 50 µl of elution buffer (supplied in kit) and stored at -20^{o}C.

2.3.11 PCR Purification

PCR purification was performed using the QIAquick PCR Purification kit (Qiagen). Five volumes of buffer PBI were added to one volume of the PCR sample to ensure optimal pH (≤ 7.5) for DNA binding, and then transferred to a QIAquick spin column and microcentrifuged for one minute at 11 300 x *g*. Buffer PE was used to wash the

spin column membrane to remove unwanted primers and impurities, and to adjust to low-salt conditions. DNA was eluted using 30 µl elution buffer (supplied in the kit).

2.3.12 Precipitation of DNA

Two volumes of ethanol and $\frac{1}{10}$ volume 3 M sodium acetate (pH 4.8 – 5.2) were added to DNA in elution buffer, vortexed briefly and placed into -80^{o}C for at least 30 minutes to increase the efficiency of the precipitation. Following centrifugation at 11 300 x g for ten minutes the supernatant was removed. 200 µl 70% ethanol was added to the pellet and mixed by inversion, then centrifuged at 11 300 x g for two minutes. All ethanol was removed and the pellet left to dry. The DNA was resuspended in an appropriate volume of water.

2.3.13 DNA Sequence Analysis

Samples were prepared at 50 ng µl^{-1} and primers were supplied at 1.6 pmol µl^{-1} (see Table 2.5 for primer sequences). Sequencing was performed by the University of Leeds Genetic Analysis Facility using BioDye® Terminator v 3.1 Cycle Sequencing (ABI) technology using the dye-labelled, dideoxy terminator method. Dye-labelled dideoxy terminators were incorporated in a reaction mix with the DNA sample and a 25 cycle PCR performed. Following precipitation of unincorporated dye products, samples were run on an ABI 3103x1 Capillary Sequencer to generate sequence data output as a chromatogram file.

Primer	Primer Sequence 5' – 3'
M13F	CGT-TGT-AAA-ACG-GCC-AGT
M13R	TCA-CAC-AGG-AAA-CAG-CTA-TGA-C
GUS	GAA-TGC-CCA-CAG-GCC-GTC-GAG

Table 2.5 Primers used for DNA sequencing. M13 primers were used for the sequencing of inserts in the pSTBlue-1 and pBluescript vectors. The GUS primer was used for sequencing of inserts in the pBI101 vector.

2.3.14 Cloning of *A. l. petraea* DNA

2.3.14.1 pSTBlue 1 Vector

Amplified DNA fragments were cloned into pSTBlue 1 (Figure 2.1) using the AccepTor Vector Kit (Novagen) according to the manufacturer's instructions. Four microlitres of precipitated DNA and 5 µl Clonables™ 2x ligation premix (which includes ligase and buffer) were added to 1 µl of pSTBlue-1 vector (50 ng μl^{-1}, Figure 2.1), mixed gently and incubated at 16^{o}C for between 30 minutes and two hours in a waterbath. A greater incubation time results in a greater number of recombinants.

2.3.14.2 Vectors pBluescript and pBI101

Purified PCR products and sequenced pBluescript (Figure 2.2) clones were digested with the appropriate restriction enzymes. Vectors pBluescript (4 µl) and pBI101 (5 µl) were also digested with these same restriction enzymes in final reaction volumes of 30 µl and 20 µl respectively. Following incubation at 37^{o}C for three hours, reactions were made up to a final volume of 100 µl with sterile water. In the case of pBI101 (Figure 2.3), vector and insert were combined prior to this step. Samples were purified using phenol:chloroform (as 2.3.3), precipitated (as 2.3.12), and resuspended in water. Digested pBluescript vector and PCR products were resuspended in 15 µl and 4 µl sterile water respectively, whilst the combined digested pBI101 and insert were together resuspended in 7 µl sterile water. Only 3 µl of the digested pBluescript was used for each subsequent ligation. The ligation of vector and insert was performed following the addition of 1 µl T4 Ligase (Invitrogen) and 2 µl 5x ligase buffer (Invitrogen). Reactions were mixed gently, and incubated at 16^{o}C for at least four hours in a waterbath.

2.3.15 Restriction digests

Restriction digests were performed to determine whether or not amplified DNA fragments had been successfully cloned into the desired vector. The appropriate restriction enzymes (1 µl of each) for each construct were used in a reaction with 3 µl (pSTBlue-1 or pBluescript) or 6 µl (pBI101) plasmid DNA from transformed *E. coli*, 2 µl of the appropriate buffer, and water to a final volume of 20 µl. The restriction enzyme *Eco* R1 was used in all pSTBlue-1 digests. Enzymes and buffers for all pBluescript and pBI101 constructs are detailed in Table 2.6. All digests were incubated at 37^{o}C for three hours and the products separated on a 1% (w/v) agarose gel.

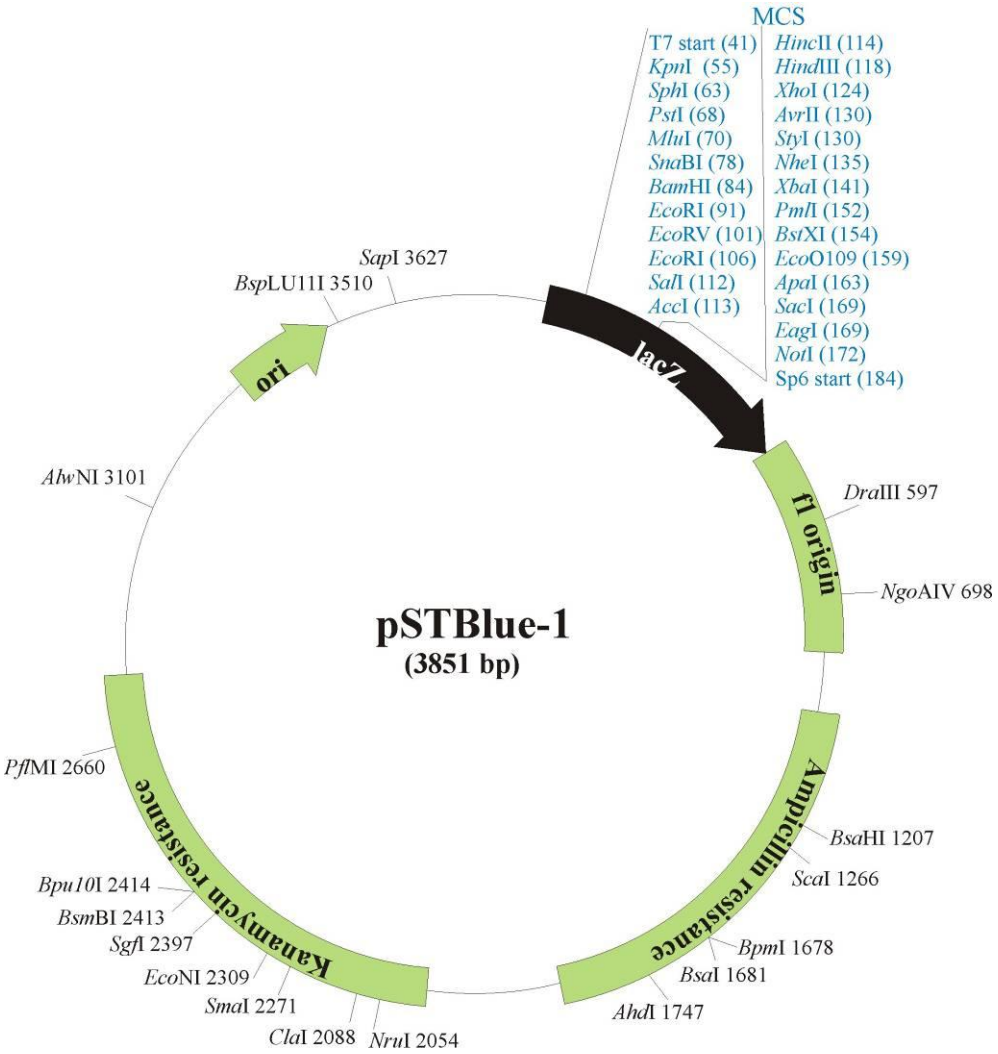


Figure 2.1 Map of the pSTBlue-1 vector. This vector was used for the cloning of *A. l. petraea* DNA. Sequences of interest were cloned between the two *Eco* RI restriction sites. (not to scale).

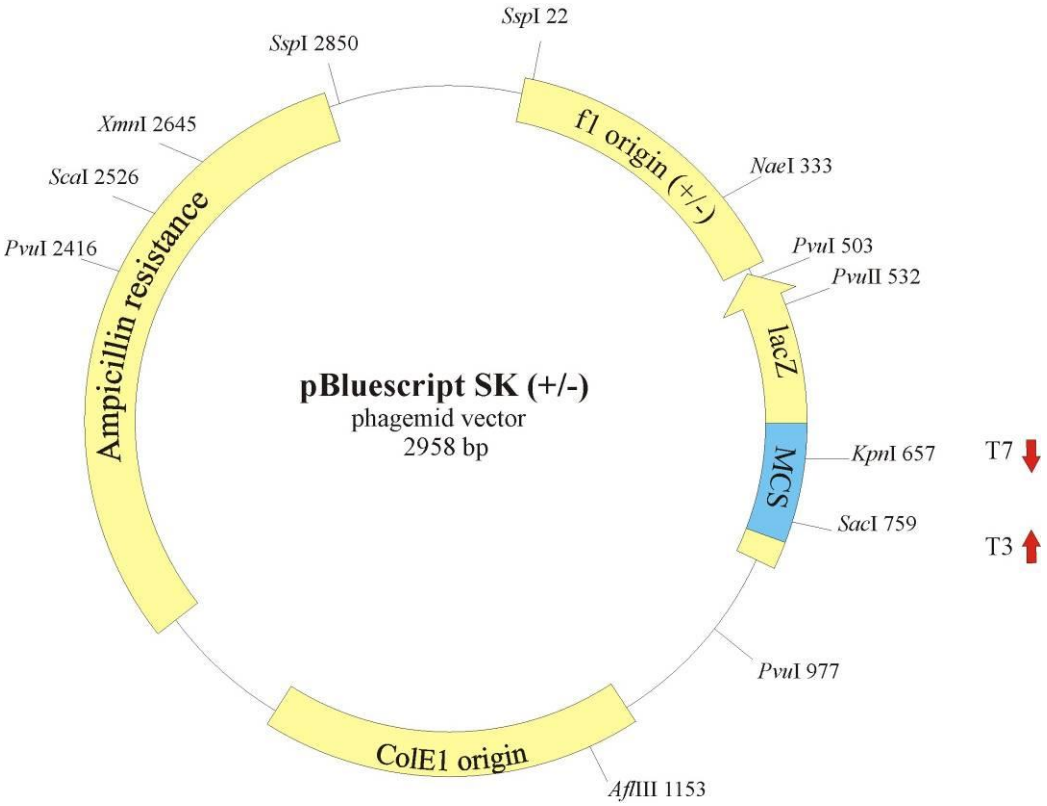


Figure 2.2 Map of the pBluescript vector. This vector (Agilent Technologies) was used for the cloning of *A. l. petraea* promoter regions (Chapter 6), prior to cloning into vector pBI101. Restriction enzymes used were specific to each sequence of interest (see Table 2.6). (not to scale).

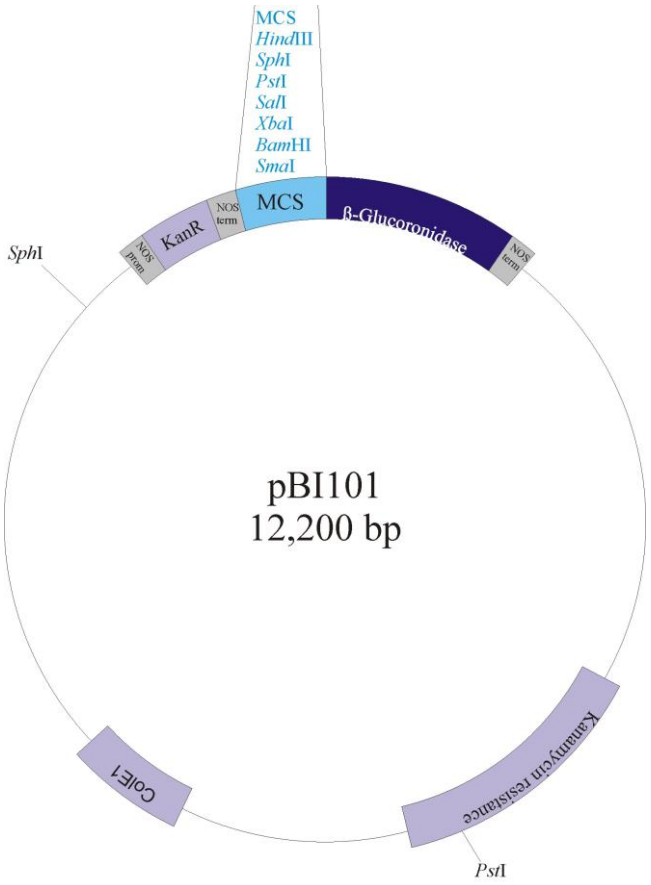


Figure 2.3 Map of the pBI101 vector. This vector was used in the generation of promoter:reporter constructs for plant transformation (Chapter 6). *A. l. petraea* promoter regions were inserted upstream of the ß-glucurounidase gene. Restriction enzymes used were specific to each sequence of interest (Table 2.6). (not to scale)

Gene AGI code	**Restriction site (Forward primer)**	**Restriction site (Reverse primer)**	**Buffer used in digests**
At1g09070	*Hin*d III	*Xba* I	REact 2
At5g42050	*Sal* I	*Bam* H1	BamH1
At5g11740	*Hin*d III	*Bam* H1	Buffer B
At4g31800	*Hin*d III	*Xba* I	REact 2
At1g18210	*Hin*d III	*Bam* H1	Buffer B

Table 2.6 Restriction enzymes and appropriate buffers for pBluescript and pBI101 constructs. Restriction sites were added to forward and reverse primer sequences, ensuring that the chosen restriction sites were not present within the sequence of the *A. l. petraea* region to be cloned.

2.3.16 Preparation of DH5α *E. coli* competent cells

Luria-Bertani (LB; 10 g l^{-1} tryptone; 5 g l^{-1} yeast extract; 10 g l^{-1} NaCl; and 1% (w/v) agar) agar plates were inoculated with *E. coli* DH5α cells from a glycerol stock, and cultured overnight at 37°C. Twelve large colonies were used to inoculate 250 ml SOB (0.5% (w/v) yeast extract; 2% (w/v) tryptone; 10 mM NaCl, 2.5 mM KCl; 10 mM $MgCl_2$; and 10 mM $MgSO_4$) in a one litre flask. This was cultured at 19°C in a shaking incubator (New Brunswick Scientific) to OD_{600} = 0.5. The flask was placed on ice for ten minutes, prior to pelleting of cells at 5000 x g for ten minutes at 4°C. Pelleted cells were resuspended in 80 ml ice cold TB (10 mM PIPES; 15 mM $CaCl_2$; and 250 mM KCl) and stored on ice for ten minutes. Cells were again pelleted at 5000 x g for ten minutes at 4°C, and resuspended in 20 ml ice cold TB and 1.4 ml dimethylsulphoxide (DMSO, Fisher Scientific, previously stored at -20°C overnight). The cell suspension was divided into 200 µl aliquots, flash frozen in liquid nitrogen, and stored at -80°C until required.

2.3.17 Transformation of *E. coli* competent cells

DH5α *E. coli* competent cells (200 µl) were thawed on ice and 4 µl of ligation reaction added. Cells were left on ice for five minutes following gentle stirring. Cells were plated out on Petri dishes of LB growth media, pre-warmed at 37°C. Selection for the plasmid was imposed by addition of 50 µg ml^{-1} kanamycin (pSTBlue1 and pBI101 vectors) or 50 µg ml^{-1} ampicillin (pBluescript) to LB growth media. X-gal, at 70µg ml^{-1}, was included in media to allow for blue/white screening of recombinants. Colonies containing a vector with an insert were white, those without an insert were blue.

2.3.18 Preparation of competent *Agrobacterium tumefaciens* cells

A single colony of *A. tumefaciens* GV3101 was inoculated into 5 ml liquid LB (as LB media but minus agar) containing 50 µg ml^{-1} rifampicin and grown at 28°C for approximately 16 hours. Two millilitres of this culture was added to 50 ml fresh LB containing 50 µg ml^{-1} rifampicin and grown at 28°C for a further three to five hours until the culture reached an OD_{600} of 0.5. The culture was chilled on ice, and cells harvested by centrifugation at 5000 x g for five minutes at 4°C. The supernatant was removed and cells resuspended in 1 ml ice cold 20 mM $CaCl_2$ by gentle shaking. The cell suspension was divided into 200 µl aliquots, flash frozen in liquid nitrogen and stored at -80°C until use.

2.3.19 Transformation of competent *A. tumefaciens* cells

Plant transformation vector pBI101 constructs at a concentration of between 10-100 ng μl^{-1} were added to aliquots of competent *A. tumefaciens* GV3101 (prepared as 2.3.18) that had been thawed on ice. The mixture was incubated at 37^{o}C for five minutes. One millilitre of liquid LB was immediately added to the mixture, and the cells transferred to a sterile 15 ml plastic tube. Cells were incubated at 28^{o}C in a shaking incubator at 200 rpm (Sanyo GallenKamp) for at least two hours. Cells were then plated onto LB agar plates containing 50 µg ml^{-1} kanamycin to select for the pBI101 vectors and 50 µg ml^{-1} rifampicin to select for the *A. tumefaciens*. Plates were incubated at 28^{o}C for 48 hours until colonies were visible. Single colonies were re-streaked onto fresh plates containing 50 µg ml^{-1} kanamycin and 50 µg ml^{-1}-rifampicin and incubated at 28 ^{o}C for 24 h.

2.3.20 PCR screening of transformant colonies

A single colony of a selected clone was transferred to a sterile PCR tube and suspended in 10 µl sterile distilled water. Additional components of a standard PCR reaction were added and a standard PCR cycle used to amplify the gene of interest (2.3.8).

2.3.21 Isolation of plasmid DNA from *E. coli* and *A. tumefaciens*

2.3.21.1 Alkaline lysis method

Between 1.5 and 3 ml of overnight *E. coli* cell cultures in liquid LB medium was pelleted in a microcentrifuge at full speed and the supernatant discarded. The cell pellet was resuspended in 100 µl of iso-osmotic solution comprising 1% (w/v) glucose, 10 mM EDTA (pH 8.0) and 25 mM Tris (pH 8) through vortexing. Cell lysis was achieved by adding 200 µl of a 1% SDS (w/v) and 0.2 M NaOH solution and mixing by inversion. 150 µl of ice cold 3 M potassium/5 M acetate solution was added and mixed immediately by brief vortexing before incubation on ice for five minutes. Cell debris was pelleted in a microcentrifuge at full speed for five minutes. Supernatant was transferred to a fresh microcentrifuge tube, two volumes of 100% ethanol were added and samples left to stand for five minutes to precipitate DNA. The solution was microcentrifuged for five minutes at full speed, ethanol removed and 0.5 ml 70% (v/v) ethanol added to wash the DNA pellet. After recentrifugation, all ethanol was removed with a pipette and the pellet allowed to air dry. The pellet was resuspended in 30 µl 10 mM Tris (pH 8.5), and 1 µl of RNase (10 mg ml^{-1}) was added.

2.3.21.2 Commercially available kit

The QIAprep Spin Miniprep kit (Qiagen) was also used to isolate plasmid DNA, following the manufacturer's protocol. Bacterial cells were lysed under alkaline conditions, then neutralised and adjusted to high-salt binding conditions. DNA was then adsorbed to a silica membrane in a QIAprep spin column. Buffers PB and PE were used to wash the membrane to remove contaminants and adjust to low-salt conditions. DNA was then eluted at a pH optimum of 7.0-8.5. DNA was stored at -20^{o}C.

2.3.22 Transformation of *Arabidopsis thaliana* using *Agrobacterium tumefaciens*

Cultures of *A. tumefaciens* GV3101 carrying appropriate pBI101 constructs were prepared by inoculating a single colony into 5 ml liquid LB media containing 50 µg ml^{-1} rifampicin and 50 µg ml^{-1} kanamycin. Cultures were grown at 28^{o}C in a shaking incubator (200 rpm) for approximately 16 hours. One millilitre of this culture was then added to 200 ml of fresh liquid LB containing 50 µg ml^{-1} rifampicin and 50 µg ml^{-1} kanamycin and grown at 28 ^{o}C in a shaking incubator (200 rpm) for approximately 16 hours. Cells were pelleted by centrifugation at 5000 x *g* at room temperature for ten minutes. Supernatant was removed and cells were resuspended in 5% (w/v) sucrose with the addition of 0.05% (v/v) Silwet L-77 surfactant (Lehle Seeds). Solutions of *A. tumefaciens* were used to inoculate *A. thaliana* Col-0 plants that had been grown from seed in temperate glasshouse conditions until plants produced inflorescences. All flowers were immersed in the solution containing the *Agrobacterium* and incubated with gentle agitation for approximately 30 seconds, a process termed floral dipping. After inoculation plants were placed under a propagator lid or sealed in florist bags for 48 hours to maintain high humidity. The floral dipping process was repeated after five days to maximise the number of inflorescences inoculated. Plants were maintained under temperate glasshouse conditions and seed collected in paper seed bags.

2.3.23 Selection of transformed lines of *A. thaliana* in tissue culture

Aliquots of *A. thaliana* seed were sterilised for twenty minutes in a 20% (v/v) bleach solution. These aliquots were mixed on a rotating platform to ensure all seed made contact with the bleach. The bleach was removed by washing the seed five times with sterile distilled water. Seed was stored in sterile distilled water at 4^{o}C overnight prior to sowing. Seeds were plated on to ½ MS10 media (0.2% (w/v) Murashige and Skoog basal medium including vitamins (Melford); 1% (w/v) sucrose; and 1% (w/v) plant

agar) containing 50 µg ml^{-1} kanamycin to select for transformed plants and 250 µg ml^{-1} cefotaxime to prevent growth of *A. tumefaciens*. Plates were maintained at 20°C, 16 h day length, for approximately three weeks to allow transformed plants to develop (Figure 2.4). These lines were then transferred to potting compost (SHL) in 24-cell tray inserts and grown under temperate glasshouse conditions. Seed was collected from individual plants using Aracons (Lehle Seeds).

2.3.24 Growth of *A. l. petraea* in tissue culture

Aliquots of *A. l. petraea* seed were sterilised for twenty minutes in a 20% (v/v) bleach solution, and were mixed, washed, and stored as described for *A. thaliana* seed above (2.3.23). Seeds were plated on to ½ MS10 media containing 5 mg l^{-1} silver thiosulphate (STS), known to enhance *Arabidopsis* root development in tissue culture grown plants (Clarke and Lindsay, 1992). Four seeds were plated per square plate (10x10 cm, Sterilin), and incubated at 20°C, 16 hour day length, for approximately four weeks (Figure 2.5A) with the plates held vertically.

2.3.25 Transformation of *A. l. petraea* using *A. tumefaciens*

A. l. petraea root systems were removed from plants grown in tissue culture (as 2.3.24) by cutting just below the hypocotyls. Root systems were placed intact on to callus-inducing medium (CIM, pH 5.7; 0.3% (w/v) Gamborg B5 medium (Melford); 0.05% (w/v) MES; 2% (w/v) glucose; and 5.5 g l^{-1} plant agar) containing 5 mg l^{-1} STS, 0.5 mg l^{-1} 2,4-dichlorophenoxyacetic acid (2,4-D) and 0.05 mg l^{-1} kinetin, and incubated at 23°C for three days. Plates were wrapped in a layer of tissue to reduce the light levels. Roots were transferred to a sterile Petri dish and cut into sections of approximately 0.5 cm. These root segments were then co-cultivated with a 20 x dilution, in culture dilution medium (CDM; as CIM but minus plant agar), of a 16 h old culture of *A. tumefaciens* carrying the appropriate pBI101 construct. Following two minutes of co-cultivation, explants were filtered on sterile nylon mesh to remove excess inoculum, and blotted dry on sterile filter paper. The cut root sections were formed into small bundles (approximately 5x5x2 mm), placed on fresh CIM plates and incubated at 20°C, 16 hour day length, for three days (Figure 2.5B). Bundles were then washed

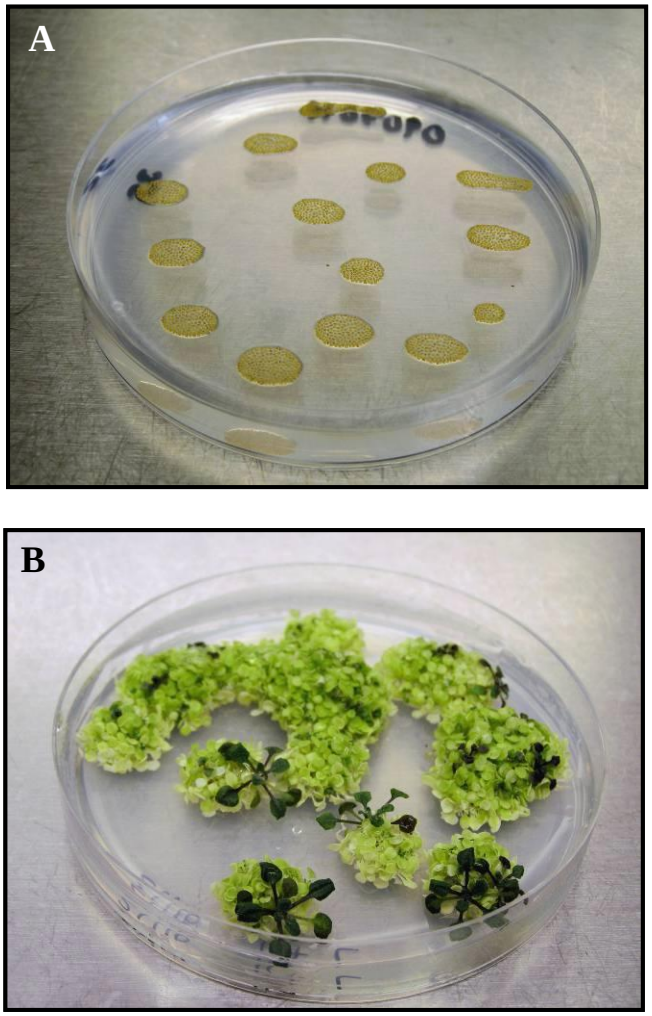


Figure 2.4 Screening of *A. thaliana* seed for transformants. Seed collected from transformed *A. thaliana* plants (section 2.3.22) was rigorously sterilised in a 20% bleach solution prior to plating onto media containing kanamycin (A). The pBI101 T-DNA confers resistance to kanamycin thus only seedlings harbouring the desired construct were able to develop true leaves (B).

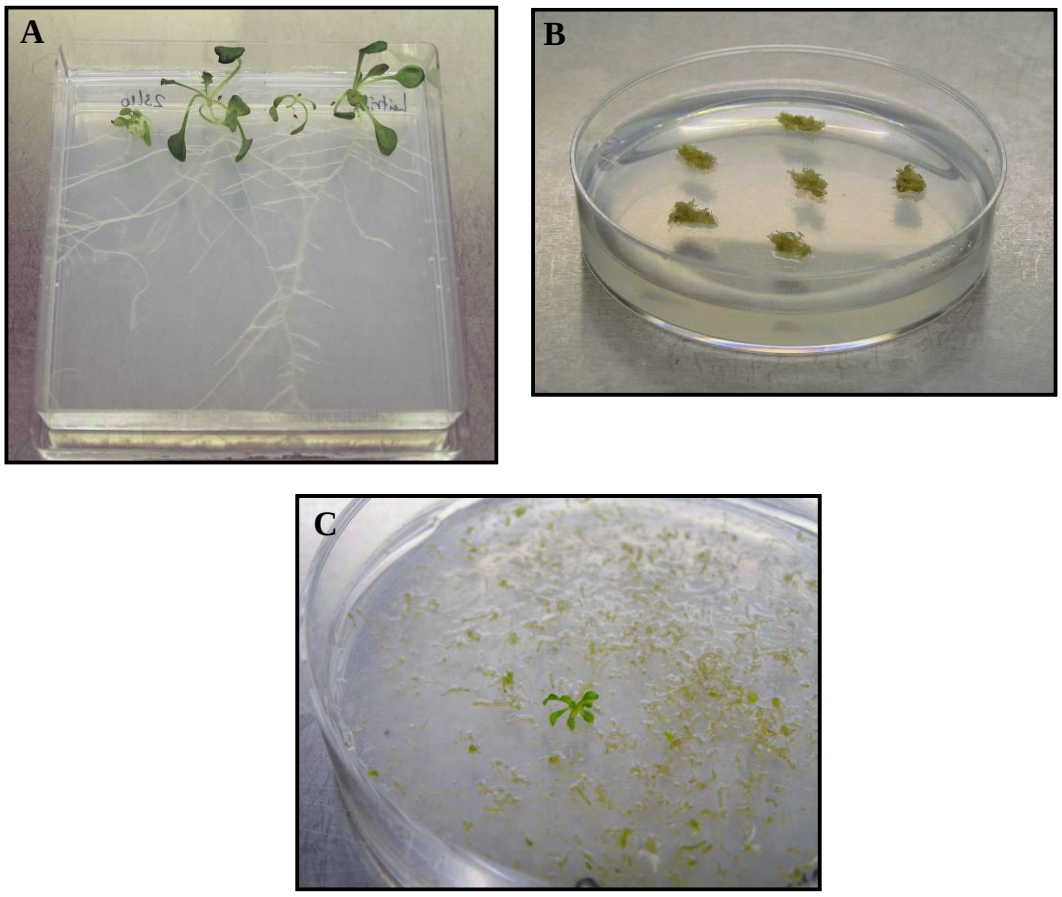


Figure 2.5 Transformation of *A. l. petraea* using root explants. Root systems were removed from plants grown in tissue culture (A) and, following three days on callus-inducing medium, were cut into sections of 0.5 cm. Cut root segments were co-cultivated with *A. tumefaciens* harbouring the desired pBI101 construct. Roots were blotted dry and formed into small bundles (B) and placed on callus-inducing medium. After thorough washing, root sections were re-suspended in agarose and poured onto plates containing shoot-inducing medium. Plates were incubated at 20^{o}C to allow shoots to form from the calli (C). Shoot-inducing media contained kanamycin to select for shoots carrying the appropriate pBI101 construct. Individual transformants were transferred to plant growth media and left to root.

twice in CDM to remove excess *A. tumefaciens* and thoroughly blotted dry on sterile filter paper. Root sections were re-suspended in 10 ml molten shoot overlay medium (SOM, pH 5.7; 0.3% (w/v) Gamborg B5 medium (Melford); 0.05% (w/v) MES; 2% (w/v) glucose; and 8 g l^{-1} low-melting point agarose, 30°C), containing 5 mg l^{-1} 2-isopentenyladenine (2-iP), 0.15 mg l^{-1} indole-3-acetic acid (IAA), 250 µg ml^{-1} cefotaxime and 50 µg ml^{-1} kanamycin, by vigorous shaking, and poured on to shoot-inducing medium (SIM; as SOM, but with 5.5 g l^{-1} plant agar replacing low-melting point agarose). Plates were left to dry in the tissue culture flow hoods for thirty minutes before being sealed with micropore tape and incubated at 23°C for one week. Plates were then incubated at 20°C for approximately six weeks to allow shoots to form from the calli (Figure 2.5C). These shoots were removed to ½ MS10 media containing 50 µg ml^{-1} kanamycin and 250 µg ml^{-1} cefotaxime, and left to root.

2.3.26 Quantitative analysis of gene expression

Primers obtained from MWG or Eurogentec were diluted to 100 µmol l^{-1} in sterile water. Stocks for use in qRT-PCR reactions were prepared by combining forward and reverse primers for the same gene at a concentration of 7.5 pmol $µl^{-1}$ each. Stocks were stored at -20°C.

All qRT-PCR experiments were run on a Agilent Technologies Mx3005P and results analysed using MxPro software (v 4.01, Agilent Technologies). Standard curves were generated for all target genes using serial dilutions (10 fold) of cDNA. Only those primer pairs with efficiencies > 90% and < 110% were used. The specificity of each primer pair was analysed using a dissociation curve. A reaction volume of 25 µl was used for all assays; this contained 1X final concentration of Brilliant® II SYBR® Green qPCR master mix (Agilent Technologies), 500 nM each primer, and 5 µl template cDNA. All standards and samples were run in triplicate, and controls without template were included for all primer pairs. The qRT-PCR cycle protocol was as follows: 95°C for ten minutes, annealing (60°C) for 30 seconds, 72°C for 30 seconds for 40 cycles, followed by a gradual temperature increase from 55°C to 95°C to produce the dissociation curve.

2.4 List of suppliers addresses

1. Affymetrix, 3420 Central Expressway, Santa Clare, CA 95051, USA. www.affymetrix.com
2. Bayer Garden, Bayer Crop Science Ltd, 230 Cambridge Science Park, Milton Road, Cambridge, CB4 0WB. www.bayergarden.co.uk
3. Biometra GmbH, Rudolf-Wissell-Str 30, D-37079, Goettingen, Germany. www.biometra.de
4. Bio-Rad Laboratories Limited, Bio-Rad House, Maxted Road, Hemel Hempstead, Hertfordshire, HP2 7DX. www.bio-rad.com
5. BMG LABTECH, PO Box 73, Aylesbury, HP20 2QJ. www.bmglabtech.com
6. Desch Plantpak LTD, Burnham Road, Mundon, Maldon, Essex, CM9 6NT. www.plantpak.co.uk
7. Eppendorf UK Limited, Endurance House, Chivers Way, Histon, Cambridge, CB24 9ZR. www.eppendorf.com
8. Epson (UK) Ltd, Westide, London Road, Hemel Hempstead, Herts, HP3 9TD.
9. Eurofins MWG Operon, Anzingerstr . 7a, 85560 Ebersberg, Germany. www.eurofinsdna.com
10. Eurogentec S.A. Liège Science Park, Rue Bois Saint-Jean, 5 4102 Seraing, Belgium. www.eurogentec.com
11. Fermentas UK, Sheriff House, Sheriff Hutton Industrial Park, York, YO6 0RZ. www.fermentas.com
12. Invitrogen Ltd, 3 Fountain Driver, Inchinnan Business Park, Paisley, PA4 9RF. www.invitrogen.com
13. Lehle Seeds, 1102-D South Industrial Boulevard, Round Rock, TX, 78681, USA. www.lehleseeds.com
14. Melford Laboratories Ltd, Bildeston Road, Chelsworth, Ipswich, Suffolk, IP7 7LE. www.melford.co.uk
15. NanoDrop Technologies, now a division of Thermo Fisher Scientific. 3411 Siverside Road, Bancroft Building, Wilmington, DE 19810, USA. www.nanodrop.com
16. New Brunswick Scientific, 17 Alban Park, Hatfield Road, St. Albans, Hertfordshire, AL4 OJJ. www.nbsc.com
17. Novagen, Merck Chemicals Ltd, Padge Road, Beeston, Nottingham, NG9 9JR. http://www.merck-chemicals.com/life-science-research
18. Qiagen, Qiagen House, Fleming Way, Crawley, West Sussex, RH10 9NQ. www.qiagen.com
19. Sanyo E+E Europe, 9 The Office Village, North Road, Loughborough, LE11 1QJ. www.sanyo-biomedical.co.uk
20. Snijders Scientific, PO Box 5099, 5004 EB Tilburg, The Netherlands. www.snijder.com
21. Stratagene, now a division of Agilent Technologies UK Ltd, South Queensferry, West Lothian, EH30 9TG. www.home.agilent.com/agilent/home
22. WEISS Gallenkamp, Units 37-38, The Technology Centre, Epinal Way, Loughborough, LE11 3GE. www.weiss-gallenkamp.com
23. William Sinclair Horticulture Ltd, Firth Road, Lincoln, Lincolnshire, LN6 7AH. www.william-sinclair.co.uk

3. Response to selection in populations of *A. l. petraea.*

AIMS

- Design a response-to-selection experiment incorporating populations from throughout the north west European geographic range of *A. l. petraea*.
- Measure the chosen traits following exposure to different temperature regimes, and select an appropriate proportion of each population as parents for the subsequent generation.
- Produce seed for successive generations of the experiment using the selected individuals, and complete as many generations of the experiment as is feasible.
- Calculate response to selection in the populations under investigation and identify any differences between populations of *A. l. petraea*.

3.1 Introduction

3.1.1 Selection experiments: historical context

Since the 1950s, selection experiments have become a broadly used tool in the field of biology (Hill and Caballero, 1992). Artificial selection was initially steered towards improving animals and plants for agricultural purposes but is now being more widely used to study aspects of ecology and evolution. Such experiments have the advantage of a known selection pressure, compared to analysing natural evolution where the selective pressure experienced has to be obtained indirectly from the data (Hill and Caballero, 1992). Superior methods for analysis of results from selection experiments have been developed, including the advancement of genetic techniques that can be use to study the populations involved (Hill and Caballero, 1992).

3.1.2 Types of selection experiment design

It has been argued that there are only two basic types of selection experiment design: (i) laboratory natural selection, and (ii) artificial selection, and that all other variations can be included in this dichotomy (Scheiner, 2002; Fuller *et al.*, 2005).

3.1.2.1 Laboratory natural selection (LNS)

Such experiments involve replicate lines of the species concerned being divided over two or more experimental treatments, and their responses to the treatments being recorded over time (Fuller *et al.*, 2005). In this type of experiment, the investigator is

able to control which forces can act as selective agents but the actual strength of selection as a proportion of the population is not controlled. Strength of selection is therefore not consistent across generations, especially if the population adapts. A wide range of characters may respond to the selection treatment, but the targets of selection can be unclear (Rose, 1990). LNS experiments are used in two ways: firstly to test fundamental predictions of evolutionary theory with little reference to patterns that have been observed in nature, and secondly to test hypotheses related to patterns of variation observed in nature and the selective agents that may be responsible for maintaining these patterns (Fuller *et al.*, 2005).

3.1.2.2 Artificial selection (AS)

In experiments of this design, individuals are selected to be the parents for the next generation according to direct phenotypic measures i.e. if the individual possesses a specified value of a particular phenotypic trait then they will be selected. Individuals are selected if they have a trait value above or below a set value, or if they fall within a certain specified percentage of the population (Bell, 1997; Fuller *et al.*, 2005). In these experiments, the investigator has control over the strength of selection (e.g. proportion that are selected) and its consistency. The trait for selection must be decided in advance. There are many reasons why AS experiments are used. They can be used to investigate the symmetry of responses in a certain trait when the trait is selected in opposing directions, and to look for correlated responses in other traits when selecting for a specific trait (Fuller *et al.*, 2005). One disadvantage of AS over LNS is that it is substantially more labour intensive and therefore population sizes and experiment durations tend to be smaller.

3.1.3 Factors to consider in experimental design

Experimental design is highly dependent on the specifics of the questions being asked (Scheiner, 2002). With finite resources it is impossible to have the ideal values for all design parameters and the investigator must decide which factors to prioritise and which must be given lower importance. These factors include initial population sizes, selection intensity, levels of inbreeding, number of replicate lines, inclusion of controls and the time over which the experiment is run. Design of selection experiments can have profound effects on the outcome, and therefore it is important to produce the most appropriate experimental protocol possible (Fuller *et al.*, 2005).

3.1.3.1 Strength of selection and the impact of mutations

Two key issues involved in designing an experiment that is both useful and powerful are the strength of selection and the role of mutations during the experiment (Fuller *et al.*, 2005). The strength of selection in an experiment can determine the extent of the response. If selection is too strong then beneficial alleles may be lost and the effective population size of the next generation may be reduced significantly; this can enhance drift. If selection is very weak then replicate lines have more chance of diverging through drift, thus confusing interpretation of results. In terms of logistics, an experiment with weak selection (gradual environmental change) will take longer to produce measurable changes than one in which the selection pressure is strong, but there is risk of bias towards fixation of alleles of large effect in experiments with strong selective forces (Fuller *et al.*, 2005). Mutations can arise during the course of an experiment and are generally deleterious; they can, however, contribute to the evolutionary response. Using an outbred population will reduce the impact of newly arising mutations, whereas the response of a highly inbred population will be due to new mutations alone (Fuller *et al.*, 2005).

3.1.3.2 Replication within artificial selection experiments

Replication is an important factor in selection experiments and must be considered carefully. Experiments are naturally constrained by logistics with space for the organisms involved and time investment in trait measurement being key factors. Many published experiments involve replicate lines for different selection regimes, plus control lines that do not undergo selection (Bisazza *et al.*, 2007). Replicates are essential for ensuring that any observed changes are as a result of the selective forces and not due to drift. Selection on a single line represents a unique historical process, thus, such experiments are difficult to interpret, whereas replicate lines, if they eventually converge on a common phenotype that represents the limit to selection, can be more informative (Bell, 1997). Increasing the number of lines results in a trade-off with the number of different lines used and the number of individuals per line. Reducing population sizes can reduce the genetic diversity of the population at the start of the experiment and also affect the fixation probabilities of alleles under selection. In some cases it may therefore be more beneficial to conduct a larger experiment without replicates, rather than two smaller replicated experiments that are, as a result of being smaller, less powerful (Fuller *et al.*, 2005).

The ideal selection experiment would involve multiple lines that are selected in both directions for the specified trait value (i.e. for an increase and for a decrease in trait value), plus multiple control lines (Hill and Caballero, 1992; Callahan, 2005; Fuller *et al.*, 2005). As this will be very rarely possible, it may be more efficient to proceed with fewer treatments and a greater population size for each. Choosing the correct opposing responses (i.e. increase and decrease in trait values) can also cause problems; it is often not clear what the corresponding decreasing response is to a trait such as drought tolerance. Control lines are those maintained under the same environmental conditions as the selected lines but with no artificial selective forces acting on them. By maintaining control lines in this way, resources that could be used to carry out more replicates of selected lines are being occupied; it may be more efficient to maintain ancestral stock populations in a different form such as seeds. Directional changes in selected lines that are not related to the selective forces, such as inbreeding depression, will however not be able to be separated from the response to selection (Fuller *et al.*, 2005). A range of experimental designs have been used in selection experiments involving plants according to availability of resources and the chosen study system. Siemens *et al*. (2002) included lines of *Brassica rapa* selected for both high and low glucosinolate content in their artificial selection experiment but, due to the labour intensive nature of the assays involved in scoring phenotypes, did not include an unselected control line. Conversely, Stanton *et al.* (2000) were able to include both replicate lines and unselected controls in their investigation into stress responses in *Sinapsis arvensis*.

3.1.3.3 The optimum number of generations

Decisions must also be made about the number of generations over which to run an experiment (Scheiner, 2002). This depends on resources, but also on the biology of the species involved; if it is a species that has a short generation time then it will be easier to continue the selection for a higher number of generations than if the species has a long generation time. It is difficult to study sexual populations for a prolonged period of time as they involve much more intensive work than asexual populations (Bell, 1997). Collins and Bell (2004) used a microbial model system and they were able to complete 1000 generations of selection due to short generation times and asexual populations, in comparison to work using a teleost fish system that has longer generation times and sexual reproduction for which eight generations of selection were completed (Baer and Travis, 2000). Moreover, a selection experiment using *Senecio jacobaea* investigating

ability to flower following vernalisation was carried out for only one generation, with significant decreases in the fraction of non-flowering plants being recorded between the parental and the F1 generations (Wesselingh and Klinkhamer, 1996). Selection experiments involving plants tend not to span large numbers of generations; this may relate to their increased generation times and larger sizes compared to more commonly used insect species (Callahan, 2005). Callahan (2005) reviews a number of artificial selection experiments that use plant species, and the longest involving *Brassica juncea* (Potvin and Tousignant, 1996) is maintained for just six generations. There are, however, some longer running experiments involving plants; one of these is the Illinois corn experiment. This commenced in 1896 and selection for increased and decreased levels of oil was still producing a detectable response after over 100 generations (Dudley and Lambert, 1992; Moose *et al.*, 2004).

3.1.4 Abiotic factors are a selective force

Changes in environmental parameters are often the main forces for natural selection in wild populations. A change in an abiotic factor, such as temperature, will result in a change in the relative fitness of genotypes and resultant phenotypes within a population. As a result, different genotypes will be favoured in the next episode of selection and, over time, the genetic composition of the population and the mean values for particular character traits will change. Subsequently, the population will become adapted to the prevailing new environmental conditions. If a species is unable to adapt to a particular change in environmental conditions then the population experiencing these changed conditions may suffer a decline; for this reason, environmental parameters play a large role in determining the distribution of species.

Many previously conducted selection experiments have investigated the response of an individual species to some form of environmental stress, be it water supply (Hoffmann and Parsons, 1991; Stanton *et al.*, 2000), temperature (Baer and Travis, 2000), or levels of carbon dioxide (Ward *et al.*, 2000; Collins and Bell, 2004). Here, an artificial selection (AS) experiment involving multiple populations of *A. l. petraea* from across the north-west European geographic range, including both core and marginal populations, has been designed. This species occurs in geographically isolated populations that experience a wide range of environmental conditions. Populations that experience extreme conditions, such as persistently cold temperatures, may have become adapted to these locally prevailing conditions and, as a result, may lack the

ability to respond to imposed environmental stresses. The main objective of this study was to determine whether or not populations from the species' north-west European range exhibit differential response to selection on traits relevant to distribution. It has been hypothesised that, in addition to the environmental limitations to species distributions, marginal populations may have a key evolutionary role in limiting the expansion of a species range through either a lack of available genetic variation or the swamping of local adaptation by gene flow (discussed in section 1.2.2). This study has enabled the differential responses of a broad range of populations to both hot and cold temperature stresses to be investigated. Key factors, including strength of selection, the level of replication, and the range of treatments to investigate, were carefully considered in relation to characteristics of the study species and logistical constraints, and these are discussed in section 3.2. In addition, differences in life history characteristics, such as growth rate and seed weight, between the populations could also be indentified and discussed in relation to climatic conditions experienced by populations in the field. Initially, five treatments were chosen for investigation using seventeen populations of *A. l. petraea*, and it was planned for the experiment to continue for three or four generations.

3.2 Rationale for Experimental Design

3.2.1 Experimental setup

The number of plants was limited by availability of growth cabinet space. In total 5040 plants could be included per generation across five treatments. These plants were from seventeen populations of *A. l. petraea* and are detailed in Table 3.1.

3.2.2 Considering selection pressure and inbreeding risk

3.2.2.1 Selection pressure

The estimated response to selection using a variety of percentage selections was calculated according to the breeder's equation:

$$R = h^2 S \qquad \text{Equation 3.1}$$

where R = response to selection in units of standard deviation (SD)

h^2 = heritability of the trait. This is likely to be between 0.05 to 0.5, with 0.5 being the typical value for more neutral traits (Falconer and Mackay, 1996). Previous selection experiments with plants have reported heritabilities of between 0.09 and 0.62 (Rose, 1990) and between 0.23 and 0.31 (Burgess *et al.*, 2007) for the traits studied.

S = selection differential. This is the difference between the mean value of the current generation (those selected to reproduce) and the previous generation. This was measured in units of standard deviation.

Values of S for different selection intensities were obtained from Falconer and Mackay (1996).

3.2.2.2 Inbreeding

Selecting only a proportion of individuals within a population to act as parents for the next generation brings about a risk of inbreeding within the population. Inbreeding was calculated according to the following equation:

$$\text{Inbreeding} = (1 - \left(\frac{1}{2Ne}\right))^{t} \qquad \text{Equation 3.2}$$

where Ne = Effective population size i.e. the number of individuals selected to breed

t = time i.e. the number of generations of the experiment.

3.2.3 Combining selection and inbreeding to optimise experimental design

Considering both a suitable selection regime and the risk of inbreeding, it was calculated that a final selection of 25% should be used. For traits with heritability of 0.05, response was calculated as 0.1878 and 0.2504 standard deviations (SD) respectively for three and four generations of selection in which 25% of individuals are retained using Equation 3.1 (Table 3.2). This should be possible to detect at the end of the experiment whereas a higher heritability of, for example, 0.1 would result in a readily observable response.

Based on the calculations above, and the availability of growth cabinet space, it was deemed feasible to carry out the experiment using seventeen populations, with a total of 280 plants per population per generation. This equates to 56 plants per population per treatment. With these maximum numbers of plants, 25% initial selection results in the availability of fourteen plants per population per treatment as parents for the subsequent generation. The initial 25% selection allows for a fraction of the plants to fail to flower as *A. l. petraea* flowering is relatively unpredictable, particularly in growth cabinet conditions. Allowing for only 75% of the fourteen selected plants to flower, ten or more plants per population per treatment would be available to provide the seed for the next generation. Performing controlled crossings between limited numbers of individual plants introduces the potential for populations to become inbred. With an effective population size of ten, inbreeding risk was calculated to be minimal; for three generations of selection, Equation 3.2 gave the outcrossing value of 0.85, and for four generations this value was 0.81. Inbreeding is highest with a value of zero, and not apparent with a value of one.

3.2.4 Choice of populations for inclusion

There were several important factors to consider when deciding upon the populations to be included in the selection experiment. Ideally all countries within the north west European geographic range of *A. l. petraea* needed to be represented. Several factors are known about the different populations within each country; a population size estimate (P. Vergeer, *pers. comm.*), the estimated level of heterozygosity within each population based on SNP (single nucleotide polymorphism) data (K. Tanaka, *pers. comm.*), and the altitude at which the population persists. As far as possible, populations within countries were chosen to maximise variation in these three characteristics. Table 3.1

details the populations chosen and their characteristics. It was not possible to obtain an approximate population size for Icelandic populations; populations were effectively infinite, as *A. l. petraea* is very widely distributed in Iceland. All Swedish populations are located within ten metres of sea level therefore it was not possible to include different altitudes for this country. Population choice was also limited to a lesser extent by the availability of seeds. In particular, no laboratory produced seed and very little field seed was available for most populations from Scotland, Ireland and Wales.

3.2.5 Number of plants per population

The total number of seeds required per population was carefully considered in order to ensure that the greatest range of populations could be included in the investigation. The available laboratory-produced seed from all populations was counted in terms of numbers of seed per unique cross. Field collected seed was included in the experiment to maximise the overall number of populations. This field collected seed was counted in terms of seed per mother plant as the exact crosses were unknown.

Ideally all populations were to start the first generation of the experiment with the same founder population size. This was quantified/measured as the numbers of unique crosses (laboratory produced seed) or mother plants (field collected seed) and all populations were comprised of the same number of unique crosses or mothers (now referred to as 'families'), and the same number of individuals per family. The highest possible number of families was used, but as numbers needed to be the same for each population, the lowest common denominator for family number dictated the final number of families used. Proportions of both laboratory-produced and field-collected seed were kept equal between the different populations.

Each population in the first generation comprised 250 plants in total; 150 from field-produced seed (thirteen families) and 100 from lab-produced seed (eight families). Numbers of families used, and quantities of seed within each family, are detailed in Table 3.3. These 250 plants were divided between the five treatments, with each treatment containing exactly the same numbers of plants and families where possible. This ideal sowing plan was still not possible for all chosen populations due to seed shortages. Rather than compromise the size of all populations in the experiment, some populations without ideal numbers of seeds were included because they were the

Country	*Population*	*Population Code*	*Heterozygosity estimate*	*Estimated population size (individuals)*	*Altitude (m)*	*January average day temp. (^{o}C)*	*July average day temp. (^{o}C)*
Iceland	Raudholar	Iceland1	0.202	>100,000	80	-0.8	15.6
	Nyidalur North	Iceland2	0.191	>100,000	880	-1.6	14.2
	Fagrifoss	Iceland3	0.205	>100,000	635	-2.4	15.8
Sweden	Berghamn	Sweden1	0.255	8000	4	-1.7	24.3
	Sjoviken	Sweden2	0.253	31,000	3	-1.9	21.0
	Stubbsand	Sweden3	0.210	5000	1	-3.0	22.5
	Notsand	Sweden4	0.237	16,000	3	-1.7	20.9
Norway	Hjelle	Norway1	0.211	7,000	170	-1.6	14.7
	Saebo	Norway2	0.155	56,000	25	-0.9	19.3
	Spiterstulen	Norway3	0.204	29,000	950	-1.0	14.2
	Bovra	Norway4	0.236	17,000	495	-3.2	16.3
	Laerdal Island	Norway5	0.211	48,000	20	-1.7	19.2
	Smadalen	Norway6	0.219	6,000	1050	-3.3	16.0
	Helin	Norway7	0.211	22,000	1130	-3.6	17.4
Scotland	River Dee	Scotland1	0.209	3,000	340	2.2	19.2
Wales	Snowdon West	Wales1	0.132	4,000	670	2.9	13.8
Ireland	Leitrim	Ireland1	0.203	6,000	355	4.3	16.3

Table 3.1 Characteristics of the *A. l. petraea* populations chosen for inclusion in the selection experiment. Heterozygosity estimates are based on SNP data (Dr K. Tanaka). Estimation of population size, to the nearest thousand, was carried out by Dr P. Vergeer. The countries included in the investigation experience a wide range of winter and summer temperature conditions; data for January and July day temperatures are shown (iButton data, Dr P. Vergeer).

Heritability (H^2)	*Selection differential (S)*	*Response (R) after one generation (in SD units)*	*Response (R) after three generations (in SD units)*	*Response (R) after four generations (in SD units)*
0.5	1.252	0.626	1.878	2.504
0.25	1.252	0.313	0.939	1.252
0.1	1.252	0.1252	0.3756	0.5008
0.05	1.252	0.0626	0.1878	0.2504

Table 3.2 Estimating response to selection. The expected response to selection (R) for a range of heritabilities (h^2) was calculated in units of standard deviation (SD). The value for the selection differential was obtained from Falconer and Mackay (1997), and relates to selection of fifteen individuals from a population of 60 (25% selection). These 25% are retained as parents for the next generation. Response (R) was calculated using the Breeder's equation as detailed in section 3.2.2.

important marginal populations and consequently of biological interest. It was hoped that seed numbers could be bulked up by the second generation, and therefore these populations will have just experienced an unavoidable and relatively minor bottleneck in generation one.

3.2.6 The treatment conditions

Five different selection treatments were used in this experiment. Two of the treatments involved temperature stress (a heat stress and a cold stress) to address the differential responses of populations to abiotic stress. A third treatment involving no selection was included as a baseline for comparison. The remaining two treatments involved selection of an easily measurable trait not known to be relevant to the distribution of *A. l. petraea* in the field; leaf shape was chosen for this purpose. Leaf shape is very variable both within and between different populations of *A. l. petraea* (Jonsell *et al.*, 1995). Some plants have leaves that are very indented with many lobes along the edges, whilst other plants have leaves that are much more rounded (*pers. obs*.). These two final treatments were designed to oppose each other (i.e. selection took place in opposite directions); one treatment favoured lobed leaves whilst the other favoured rounded leaves.

No selection – plants were chosen at random as parents for the next generation. They were grown at 'optimum' conditions throughout the experiment.

Cold stress – plant experienced colder than 'optimum' temperatures for the treatment period. Those with the highest growth rate in these conditions were selected as parents for the next generation.

Heat stress – plants experienced warmer than 'optimum' temperatures for the treatment period. Those with the highest growth rate in these conditions were selected as parents for the next generation.

Arbitrary trait (leaf shape) 'upwards' – plants were selected according to their measurements for leaf shape. Those with values that equate to most lobed leaves were selected as parents for the next generation. Plants were grown at optimum conditions throughout the experiment.

Arbitrary trait (leaf shape) 'downwards' – plants were selected according to their measurements for leaf shape. Those with values that equate to most rounded leaves were selected as parents for the next generation. Plants were grown at optimum conditions throughout the experiment.

Growth conditions for all treatments are detailed in Table 3.4. Day length was consistent across all treatment conditions. The temperature stresses were defined by arbitrary values, primarily due to the differences that exist between growth cabinet and field conditions. It was essential that these arbitrary temperature treatments were non-lethal stresses; plants had to be available post-stress for production of seed for the subsequent generation.

Total plants per population	*Field-collected seed*	*Laboratory-produced seed*
250	150 seed from 13 families Quantity of seed was not equal per family but distributed as follows: Family 1: 20 Family 2: 20 Family 3: 15 Family 4: 10 Family 5: 10 Family 6: 10 Family 7: 10 Family 8: 10 Family 9: 10 Family 10: 10 Family 11: 10 Family 12: 10 Family 13: 5	100 seed from 8 families Quantity of seed was not equal per family but distributed as follows: Family 1: 20 Family 2: 20 Family 3: 15 Family 4: 10 Family 5: 10 Family 6: 10 Family 7: 10 Family 8: 5

Table 3.3 The ideal sowing plan per population for the first generation of the selection experiment. In total 250 plants were used per population, grown from a combination of laboratory-produced and field-collected seed. Genetic diversity was quantified as the number of unique crosses or mother plants for laboratory and field seed respectively, and this was kept equal across populations as far as possible. The 250 plants were divided between the five treatment conditions (detailed in Table 3.4).

Growth conditions	*Day length (hr)*	*Day temperature (°C)*	*Night temperature (°C)*	*Humidity (%)*	*Light µmol m^{-2} s^{-1}*	*Treatments experiencing these conditions*
Optimum (Control)	16	20	15	50	125	1. No selection 2. Arbitrary trait upwards 3. Arbitrary trait downwards
Hot	16	25	20	50	125	1. Heat stress
Cold	16	15	10	50	125	1. Cold stress

Table 3.4 The treatment conditions used in the artificial selection experiment. Five different treatment conditions were used. Two sets of plants were exposed to temperature stress (hot or cold) for eight weeks. Three sets of plants remained in optimal or control temperature conditions throughout the experiment. These included the set of plants for which no selection was performed, and the two arbitrary trait treatments. The arbitrary trait chosen was leaf shape and was selected in two directions: upwards (lobed leaves) and downwards (rounded leaves). Following the treatment period, selected plants from all five treatments were all exposed to the cold conditions for two weeks to promote flowering, and then were maintained in the optimal temperature conditions throughout seed production.

3.3 Methods for plant growth and measurement

3.3.1 Sowing of seeds

Seed was weighed prior to sowing, with ten seeds per family being weighed together to provide an average seed weight for each individual family. Trays containing 60-cell tray inserts were filled with sterilised 3:1 grit sand: SHL compost mix (both from Sinclair Horticulture). This mix was chosen because the low nutrient content and good drainage are representative of the conditions experienced by *A. l. petraea* in the field. Four corner cells were removed to enable watering from the base. Trays were surface watered using a watering can with rose prior to seed sowing. Three seeds from the same family were sown per cell to minimise numbers of trays required for germination (Figure 3.1); all required seedlings were transplanted to individual cells following germination. A greater number of seed than required was sown as far as was possible to allow for low germination success. Additional seed from spare families were also sown as spares to avoid problems such as seed from a particular family having a very low germination success rate. Trays were surface watered using a watering can with rose following sowing, watered from the base by adding water to the tray, covered with a propagator lid and placed in a growth cabinet (16 h day, 8 h night, 20^{o}C : 15^{o}C day : night temperature, 50% humidity regime). Propagator lids remained on trays until the majority of seedlings had emerged (between one and two weeks). All plant growth, either under control conditions or in the temperature stresses, was performed in controlled environment chambers (Sanyo E+E Europe, Snijders Scientific, and WEISS Gallenkamp). Populations known to germinate more slowly (e.g. Icelandic populations, P. Vergeer, *pers. comm.*) were sown first as an attempt to synchronise germination dates across the different populations. Laboratory produced seed and field collected seed were labelled separately to enable them to be distinguished throughout the experiment.

3.3.2 Finalising the planting plan

Germination success was recorded, and compared to the ideal sowing plan. Amendments were made in those populations where there were not enough seedlings from a particular family, and final planting plans were compiled. When doing this, the first priority was to match the founder population size in terms of number of families per population in the five treatments across one particular population. The second priority was to ensure that the numbers of individuals per family per treatment within one population were the same. Finally, as far as possible, these measures were kept

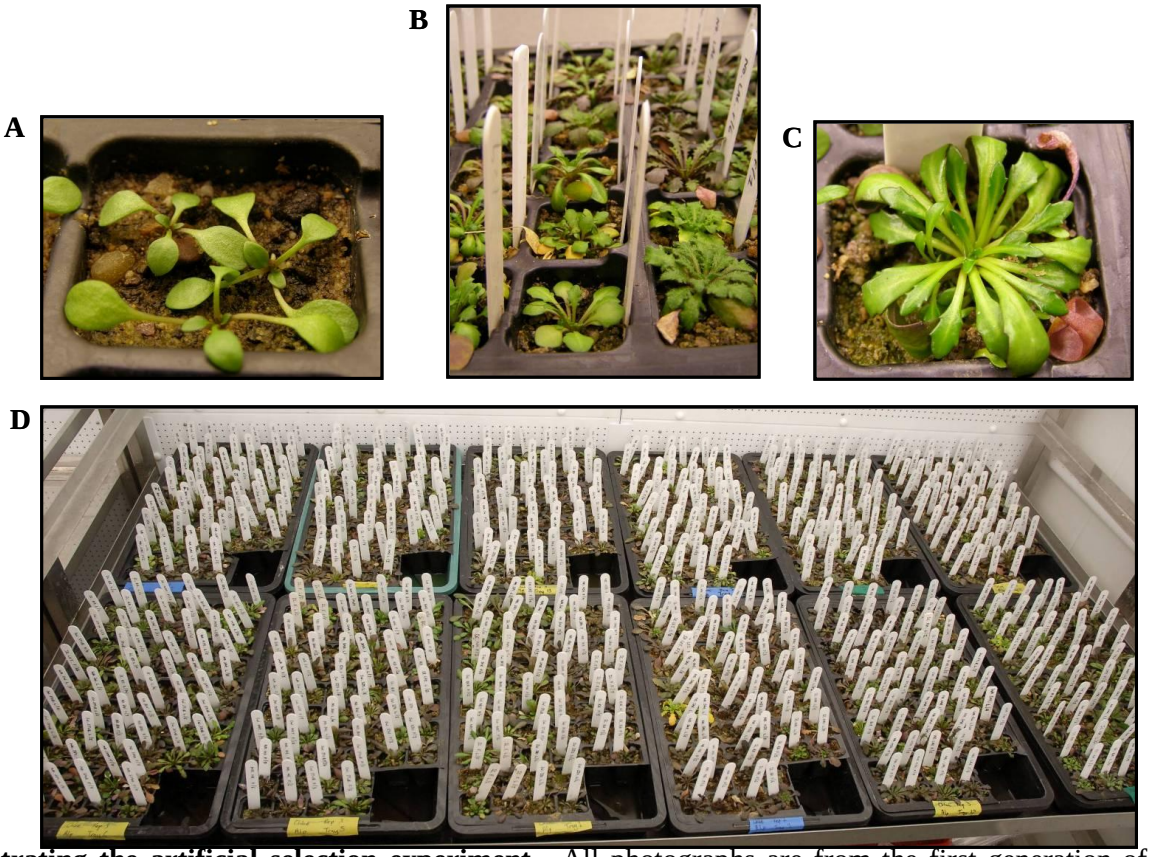


Figure 3.1 Photographs illustrating the artificial selection experiment. All photographs are from the first generation of the experiment. A: *A. l. petraea* seedlings, three per cell, prior to pricking out into the five treatments; B: *A. l. petraea* plants once pricked out, one plant per cell; C: Close up of one plant per cell and; D: one shelf in a Sanyo walk in growth cabinet containing twelve trays of plants. All plants were randomised across trays within each treatment, and each plant was individually labelled.

constant between the seventeen populations. The final planting plan for each of the five treatments, containing all plants required from all populations, was randomised to correct for tray effect. As a result all trays contained a random arrangement of plants from the different populations and families.

3.3.3 Pricking out of seedlings into sets

Trays containing 60-cell tray inserts (Plantpak) were filled with sterilised 3:1 sand: SHL compost mix. Trays were surface watered using a watering can with rose. Seedlings were pricked out into trays, one seeding per cell, according to the planting plan and each individual cell labelled to ease monitoring of the plants throughout the experiment (Figure 3.1). All trays for one treatment set were completed within a period of 48 h. Trays were covered with propagator lids for 24 h post-transplanting to minimise stress experienced by the seedlings. Trays were returned to the growth cabinet (16 h day, 8 h night, 20^{o}C: 15^{o}C day : night temperature, 50% humidity) and seedlings allowed to establish for two weeks.

3.3.4 Exposure of plants to the temperature stress treatments

Following establishment of seedlings and growth measurements (as described in sections 3.3.3 and 3.3.5 respectively), the treatment period commenced. Plants experienced the hot and cold stress conditions for a total of eight weeks (see section 3.2.6 for details of the treatment regimes). At the end of the eight week period, plants were returned to the growth cabinet set at optimum (control) conditions.

3.3.5 Plant growth measurements

Total number of leaves and length of the median and longest leaf were measured for all plants for the two temperature stress treatments, and for the no selection treatment prior to commencement of treatments. This was not required for the two arbitrary trait treatments as they were not being selected on growth rate. It was also noted whether or not plants had trichomes. These measurements were repeated after eight weeks of the treatment conditions.

The biomass index for each individual plant, both before (a) and after (b) the eight week treatment period, was calculated by multiplying the number of leaves by the length of the median leaf at each stage. The change in the biomass index (c) was then calculated

by subtracting the pre-treatment biomass index value from the post-treatment value (a-b). The relative growth rate (d) could then be obtained by dividing the change in biomass index (c) by the pre-treatment biomass index value, and multiplying this by 100 $\left((c-a)/a\right)*100$. Plants within a population were ranked based on this percentage change in biomass index value (a measure of relative growth rate) prior to selection.

3.3.6 Leaf shape measurements

Leaf shape was measured for all plants in the two arbitrary treatments. Three average sized leaves from each plant were removed using tweezers and left on the bench for approximately ten minutes to allow a loss of turgidity. Leaves were then placed on the surface of a flatbed scanner (Epson Perfection V500) and a black and white image of the leaves created (Epson Smart Panel). Images were given a transparent background using Adobe Photoshop and saved as tiff files. These tiff files were imported into Scion Image (Scion Corporation, Beta 4.0.3, 2005) where the area and perimeter of each individual leaf was measured in pixels. Pixel values were converted into area (cm^2) in Excel. The ratio of area : perimeter (area divided by perimeter) was calculated to determine leaf shape; lower values indicated leaves with more lobes or serrated edges, and higher values indicated more rounded leaves.

3.3.7 Performing selection

Selection of the top 25% of each population in each treatment as parents for the subsequent generation took place after eight weeks of treatment. For the hot and cold stress treatments, those plants that exhibited the most rapid growth rate (i.e. greatest percentage change in biomass index) over the treatment period were selected. For the treatment where no selection was to take place, 25% of the plants from each population were selected at random. For the two arbitrary trait treatments, plants were selected according to their lobed or rounded leaf shape values. All selected plants were re-potted into larger cells (40-cell tray inserts) using sterilised 3:1 grit sand: SHL compost mix, transferred to new trays, and randomised within these trays. All trays were placed in a Sanyo Walk-In cabinet under cold conditions (16 h day, 8 h night, 15^oC : 12^oC day : night temperature, 50% humidity) for two weeks to promote flowering before being maintained at optimum conditions (16 h day, 8 h night, 20^oC : 15^oC day : night temperature, 50% humidity). Plants were fed every four weeks with Phostrogen (NPK: 14:10:27; Bayer Garden, Bayer CropScience Ltd, Cambridge).

3.3.8 Controlled crossings of *A. l. petraea*

Following commencement of flowering, plants were organised according to population and treatment to aid the logistics of performing controlled crossings. Plants were placed in aracons (Arasystem, Lehle Seeds) to prevent pollen transfer through contact. Crossings were only performed between plants from the same population and the same treatment. Due to the limited number of plants that possessed ripe flowers at any one time, as many crossings as possible were performed between all available plants within each population. A mature flower was removed from the donor plant and its stamens brushed against the stigma of a mature recipient flower to enable pollen transfer. The recipient flower was then tagged with coloured cotton to ensure correct identification of the ensuing silique. Reciprocal crossings were carried out where possible. Mature siliques were collected and stored at room temperature in labelled paper bags. The number of seeds produced per cross were recorded.

3.3.9 Continuation with further generations of selection

Following seed collection, a sowing plan was produced for the next generation (generation two). This was based on the quantities of seed collected in the first generation, and the number of unique crosses that had been performed within each population and treatment. The main priorities, in order of precedence, were: (i) to reach the required number of 50 plants per population for each treatment; (ii) to include seed from the same number of unique crosses for each population in each treatment; and (iii) to ideally have equal numbers of seed from each unique cross across the populations. Greater quantities of seed than the number of plants required needed to be sown to allow for low germination success and the final planting plan was produced based on germination counts.

Seed counts from the first generation plants revealed that a large number of populations had not produced the required 50 or more seed per treatment (Figure 3.2). Many populations exhibited low levels of flowering, and crosses that were performed often produced low numbers of seed. All Swedish and Icelandic populations produced very little seed across the five treatments, as did the Scottish population Scotland1. The Welsh population (Wales1) did not produce any seed at all. Seed production within the seven Norwegian populations was more successful; most populations produced at least a small quantity of seed for each treatment. The original experimental plan of including

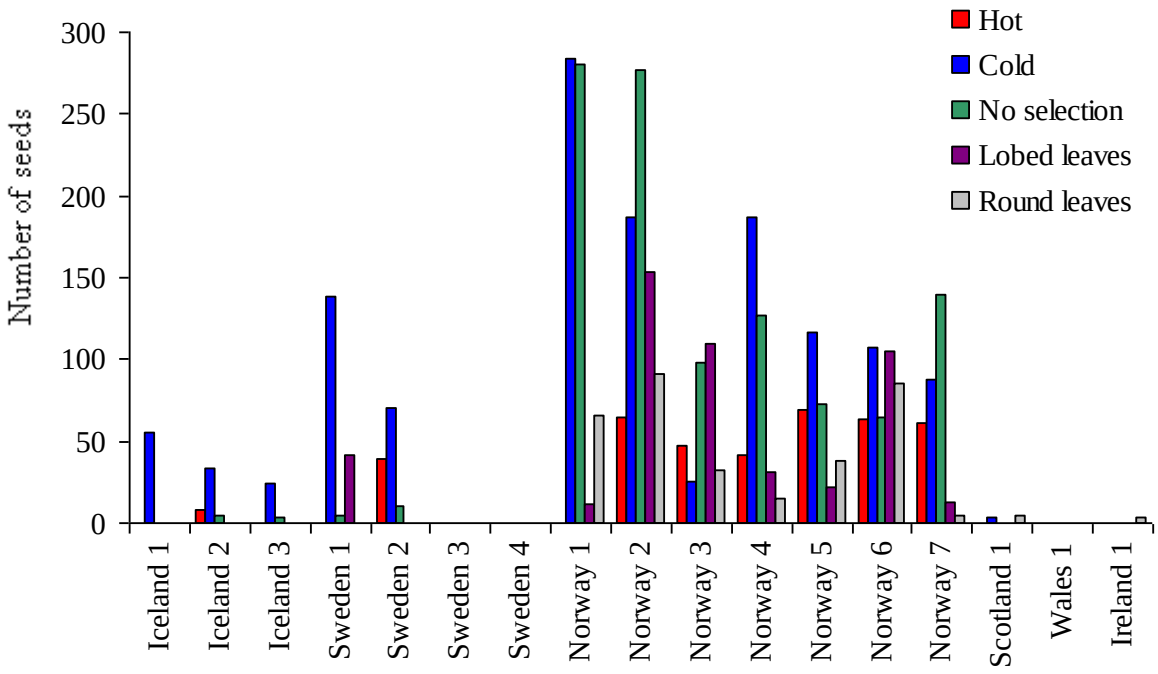


Figure 3.2 Seed production from first generation plants. This graph illustrates the number of seeds produced for each population across the five treatment conditions. At least 50 seed were required per population per treatment to permit inclusion in the second generation of the experiment. This was only possible with the Norwegian populations thus only these were involved in subsequent generations of the selection experiment.

seventeen populations across five treatments was consequently not possible for the second generation. It was decided to proceed with the seven Norwegian populations only; although from the same country, they cover a range of altitudes, population sizes, and estimated heterozygosity values. A second generation enables information about the response to selection across the populations to be obtained; one generation alone prevents this. Even with this smaller scale experiment, it was not possible to include all seven populations for all the treatments. The populations included in the second generation are detailed in Table 2.1. The process of treating, measuring, and selecting plants was then repeated as detailed in previous sections (3.3.4-3.3.7). Selected plants were re-potted and maintained as detailed in section 3.3.7. Controlled crossings were performed between plants from the same population and treatment as described in section 3.3.8.

A sowing plan was produced for the third generation based on the quantities of seed collected from the second generation plants. Priorities for this plan were the same as those for the second generation sowing plan (detailed above). The plan was adjusted based on germination success. Seed production from the second generation plants was not uniform across the populations or treatments, and a number of populations did not produce the required quantity of seed to perform a third generation. With the exception of Norway2, no populations produced sufficient seed for the two leaf shape treatments; many did not produce any seed at all. Consequently it was necessary to remove these treatments from the third generation experiment. Four of the Norwegian populations produced sufficient seed (> 50) for the cold and no selection treatments (Norway1, Norway2, Norway4, and Norway5), and just two of these also produced sufficient seed for the hot treatment (Norway4 and Norway5). The process of treating, measuring, and selecting plants was then repeated as detailed in previous sections (3.3.4-3.3.7).

3.3.10 Calculating response to selection

It was only possible to calculate response to selection in those populations that were able to complete at least a second generation of selection. These were the seven Norwegian populations. A measure of relative growth rate, the percentage change in biomass index, was used as the selection criteria for the hot and cold stress treatments. Mean values for the whole population in each generation were calculated using the change in biomass index values (pre- and post-treatment, in units of biomass index) for each plant rather than the percentage change values. The mean of the selected

individuals from the first generation was calculated in units of standard deviation to give the selection differential (S). The mean of the whole second generation population was calculated in units of standard deviation to give the response to selection value (R). After obtaining both S and R for each population, the heritability (h^2) could be calculated using the breeder's equation (section 3.2.2.1).

3.3.11 Statistical Analysis

Statistical analyses for all plant traits (seed weight, plant growth rate, leaf shape, and flowering success) were performed in SPSS (v 15.0). Univariate general linear models (glm) used population as the explanatory variable and the appropriate plant trait as the response variable. The glm for plant growth rate was also used to test for an interaction between the two independent variables, plant growth rate and seed source. Post-hoc testing (S-N-K) was performed where appropriate.

Biomass index change values for the hot and cold datasets were standardised across populations, using the overall mean and standard deviation for all plants in the first generation. This standardisation allowed for initial differences between the populations to be preserved, rather than performing standardisation within populations. Analysis of variance was performed in SPSS (v 15.0) on the standardised data. Population and generation were used as explanatory variables and the standardised datasets as response variables. In a second ANOVA, the population-specific selection differential (S) was incorporated into the model as a covariate, replacing the fixed factor generation.

3.4 Results from the artificial selection experiment

3.4.1 Seed weights and germination success

Figure 3.3 illustrates the mean seed weight in each population for both laboratory-produced and field-collected seed. Seed weight varied significantly according to source; analysis of variance indicated that laboratory-produced seed was heavier than field-collected seed (F_1 = 44.589, P <0.001). There was also significant variation in seed weight between the study populations (F_{16} = 20.70, P <0.001). Post hoc testing divided populations into five overlapping groups as indicated by the lowercase letters in Figure 3.3. The four Swedish populations, Ireland1, and Norway2 had the lightest mean seed weights, whilst Scotland1, Wales1, Norway6, and Iceland1 and 2 had the heaviest mean seed weights. Germination success was higher for laboratory-produced seed than for field-collected seed in all study populations and this was significantly different (Paired t-test: T_{16} = 2.12, P <0.001); average germination rates were 85.39% and 71.39% respectively. Individual average germination rates for all populations are detailed in Table 3.5, based the first generation of the experiment only.

3.4.2 Plant growth measurements

As illustrated in Figure 3.4, plants exposed to the heat stress grew faster than the no-selection controls in all but two of the populations under investigation. Conversely, the majority of populations exhibited a reduced rate of growth in plants exposed to cold stress conditions in the no-selection controls. Percentage change in biomass index differed significantly between the three temperature conditions (F_2 = 28.69, P <0.001), and post hoc testing divided the data into three distinct subsets. The rate of plant growth differed between the study populations; the four Swedish populations exhibited the lowest average growth rates, whilst Scotland1, Wales1, Norway7, Norway4, and Ireland1 had the highest average growth rates. Analysis of variance indicated a significant effect of population (F_{16} = 4.66, P <0.001). The interaction between temperature treatment and population was significant for plant growth measurements (F = 32, P <0.001), indicating that the populations responded differently to the three temperature conditions. Further analysis revealed a significant effect of temperature treatment on plant growth rate in eleven of the populations. The most common effect was increased rate of growth in the hot temperature treatment compared to both control and cold stress conditions, as demonstrated by eight of the eleven populations (P <0.05 in all populations). Scotland1 exhibited a significantly lower growth rate in the control

conditions (no selection treatment) than for both the hot and cold stress treatments (F_2 = 6.56, P = 0.002). Norway5 exhibited a significantly higher growth rate in the hot stress conditions compared to the cold (F_2 = 4.17, P = 0.017), but neither condition differed from the control treatment growth rate. Plant growth rate was compared between those plants grown from laboratory-produced seed and those grown from field-collected seed; there was no significant effect of seed source on growth rate (F_1 = 3.47, P = 0.063). The population x seed source interaction was, however, significant (F_{16} = 1.79, P = 0.029). Further analysis found no significant differences for sixteen of the population under investigation. The rate of plant growth for Sweden3, however, was significantly higher for plants grown from laboratory-produced seed (F_{49} = 5.83, P = 0.020).

3.4.3 Leaf shape measurements

Mean leaf shape for each population was not significantly different between the two sets of plants that were to be selected based on leaf shape. Leaf shape differed significantly between the populations (Figure 3.5, F_{16} = 41.05, P <0.001). Norway7 and Wales1 had the most lobed leaves, and the most rounded leaves were recorded in Norway3. Figure 3.6 illustrates the range of leaf shapes recorded within and between populations. Representative leaves are presented from four of the populations including those with the most lobed and the most rounded leaves according to the perimeter : area ratio values. For each population, leaves from across the range of observed values are shown. Leaf shape values were compared between those plants grown from laboratory-produced seed and those grown from field-collected seed; no significant differences were observed in the seventeen populations under investigation.

3.4.4 Selection affects variation within populations

For each treatment, the coefficient of variation (standard deviation divided by the mean, CV) was calculated for all plants within a population, and for the selected proportion of plants for that population. Figure 3.7 illustrates these values for the heat stress, cold stress, and no selection treatments. In both of the temperature stress treatments the CV was reduced for the selected proportion of plants compared to the whole population. This was observed in all populations, indicating removal of variation. In the no-selection treatment, plants were chosen at random as parents for the next generation. For these populations there was no pattern of reduction in CV values following

selection; for many populations CV values were similar pre- and post-selection although there were some exceptions, such as Sweden2 and Wales 1 (Figure 3.7). For the majority of populations in the no-selection treatment, variation was maintained in the selected proportion of plants and there was no selection towards the mean. The CV value was also reduced following selection in the majority of populations in the two leaf shape sets of plants, indicating removal of variation in these populations (Figure 3.8).

3.4.5 Variation in percentage flowering between populations and treatments

Figure 3.9 illustrates the percentage of each population (selected proportion only) that had flowered eight weeks after the end of the treatment period. Analysis of variance was performed on arcsine transformed data. Flowering percentages were significantly higher for the set of plants that experienced cold conditions than for the remaining four sets of plants (F_4 = 15.70, P <0.001). Flowering percentages also varied significantly between populations (F_{16} = 3.01, P <0.001). In general Norwegian and Swedish populations demonstrated the highest proportions of flowering plants, whilst British populations had the lowest proportions of flowering plants. Ireland1 was devoid of flowering plants in any of the five treatments at this stage.

3.4.6 Response to selection in the Norwegian populations

Response to selection was calculated for the Norwegian populations following the second generation of selection. Heritability (h^2) ranged from 0.18 to 0.72 for populations experiencing heat stress. There was no relationship between heritability values and i) altitude; ii) summer temperature; iii) original field population size; or iv) the heterozygosity estimate of each population. There was, however, a non-significant positive association between h^2 and degree days above 20^{o}C (r^2 = 0.56) that was opposite to the trend predicted. Populations experiencing cold stress had heritability values ranging from 0.13 to 0.63. Pearson product-moment correlation indicates a significant negative association between h^2 and degree days below -5^{o}C (r = 0.73, d.f. = 6, P <0.01). There was no relationship between heritability values and i) altitude; ii) winter temperature; iii) original field population size; or iv) the heterozygosity estimate of each population. The lowest heritability values for both treatments were observed in Norway4. Heritability values for each population are detailed in Table 3.6.

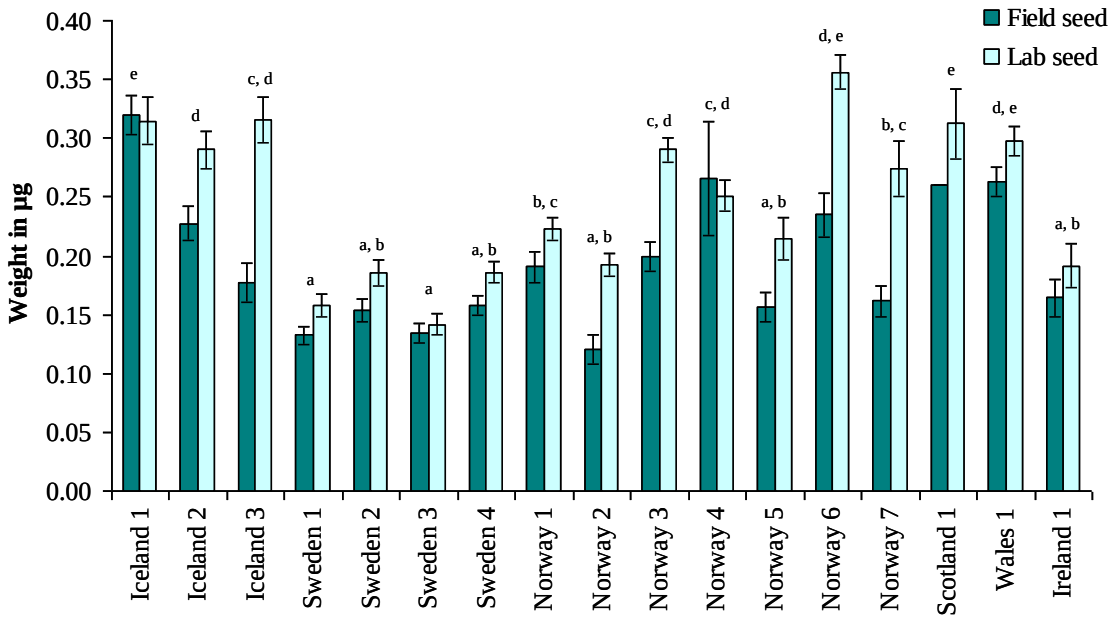


Figure 3.3 Variation in weight of seed across *A. l. petraea* populations and seed sources. Mean individual seed weights for each population, for both field-collected and laboratory-produced seed (± SE), are shown. Laboratory-produced seed was significantly heavier than field-collected seed ($F_1 = 44.589$, $P < 0.001$). Seed weight, based on combined values for laboratory-produced and field-collected seeds, differed between populations (univariate general linear model; $F_{16} = 20.702$, $P < 0.001$) with populations grouping into five overlapping groups (S-N-K) as represented by lower case letters.

	Percentage germination (%)	
Population code	***Field-collected seed***	***Laboratory-produced seed***
Iceland1	87.18	96.02
Iceland2	61.36	86.88
Iceland3	44.56	63.44
Sweden1	78.69	80.32
Sweden2	82.02	95.66
Sweden3	81.90	89.16
Sweden4	79.52	98.11
Norway1	85.64	90.00
Norway2	78.45	80.31
Norway3	85.48	89.82
Norway4	82.56	94.29
Norway5	70.83	92.22
Norway6	81.90	92.08
Norway7	63.10	80.98
Scotland1	55.97	64.56
Wales1	39.64	71.98
Ireland1	54.76	85.81

Table 3.5 Germination success of *A. l. petraea* seed from different populations. This table details the average percentage germination of both laboratory-produced and field-collected seed for all populations included in generation one of the artificial selection experiment. Germination success was higher for laboratory-produced seed in all populations.

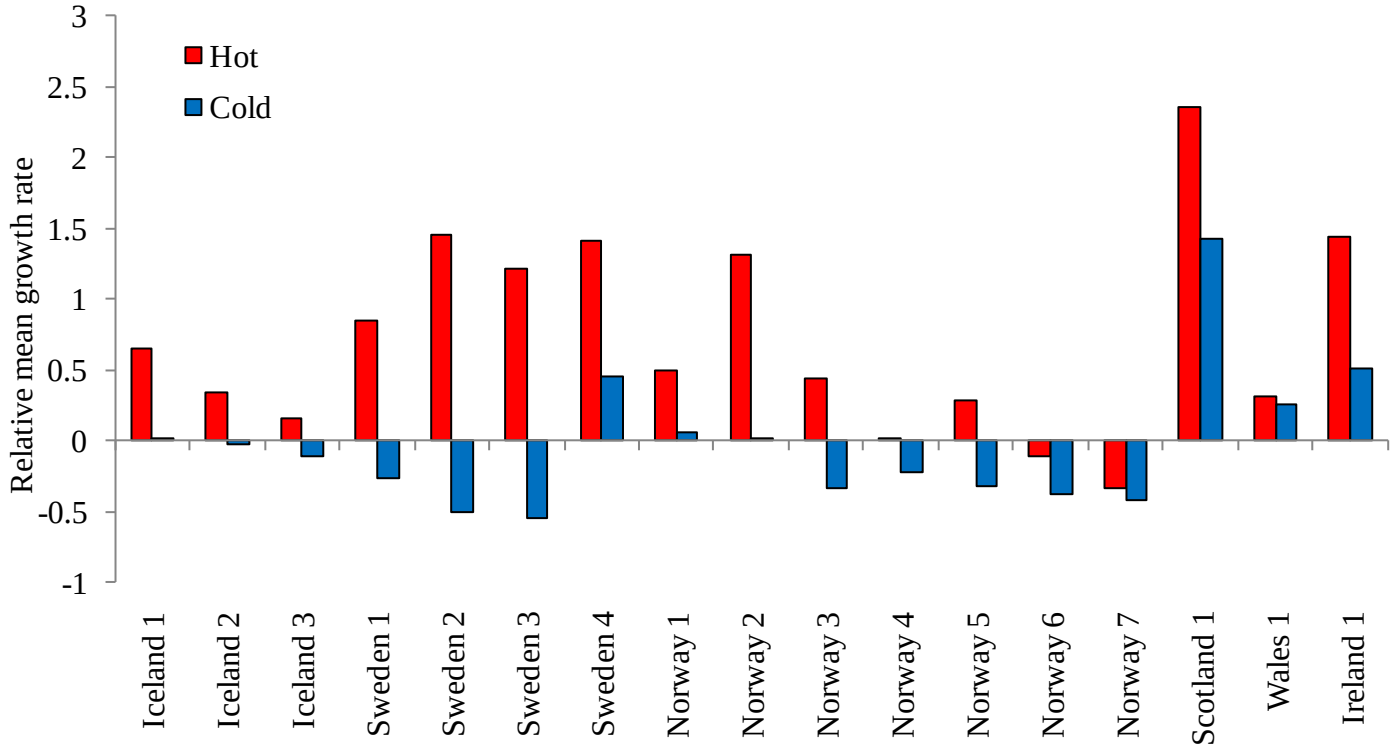


Figure 3.4 Mean growth rate of *A. l. petraea* populations relative to the no-selection controls. For each population, mean percentage change in biomass index for hot and cold treated plants is presented in proportion to the mean percentage change in biomass index of the no-selection plants. Within each population, a positive value indicates that plants grew faster than in control conditions, and a negative value indicates that plants grew more slowly than in control conditions. Temperature regimes for the hot and cold treatments are detailed in Table 3.4.

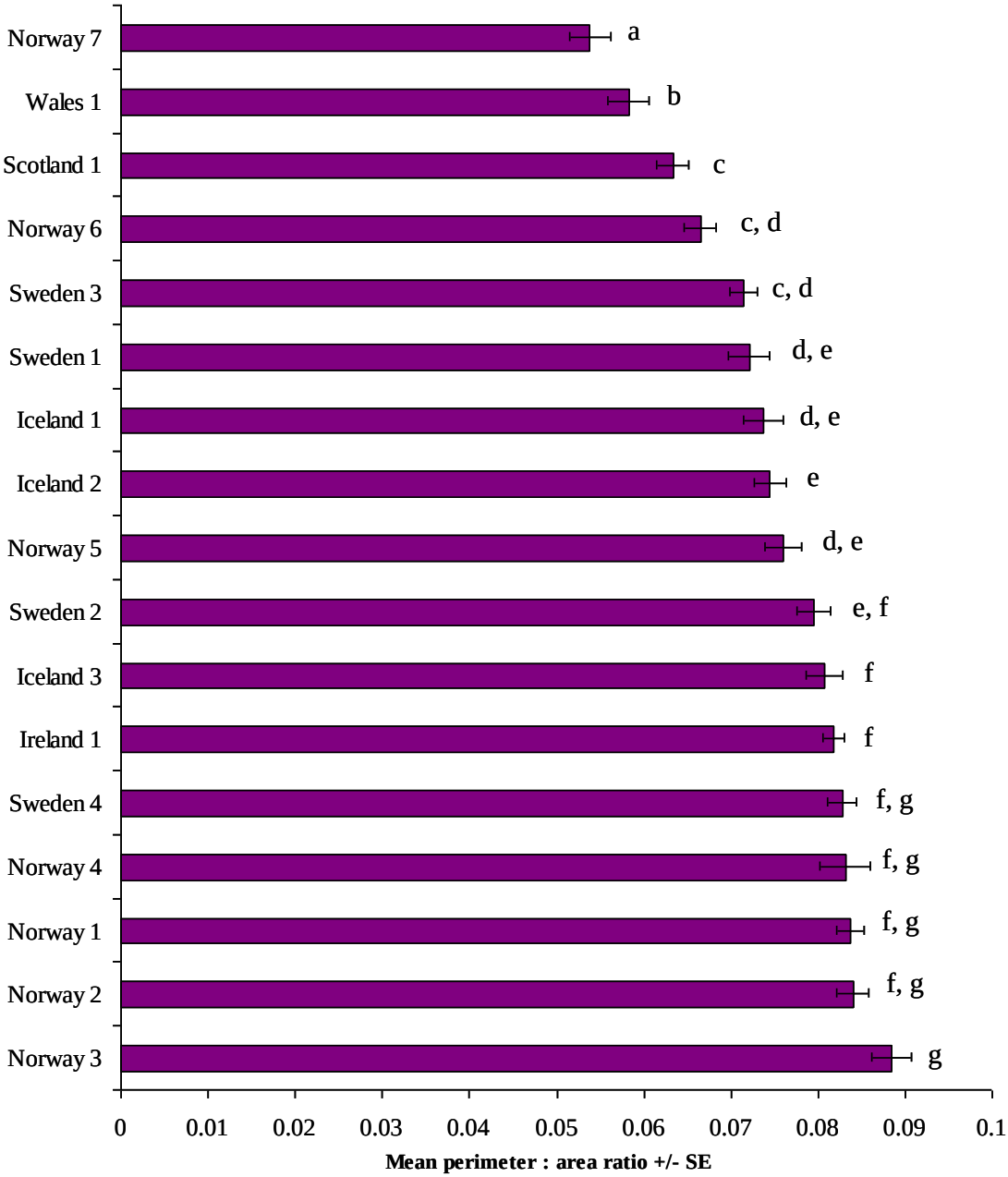


Figure 3.5 Variation in leaf shape across the study populations. Bars show mean leaf shape values (perimeter : area ratio) for plants from each study population prior to selection. Lower perimeter : area ratios indicate lobed leaves and greater perimeter : area ratios indicate rounded leaves. Populations are presented in order from the most lobed (Norway7) to the most rounded (Norway3). Leaf shape was significantly different between populations (F_{16} = 41.05, P <0.001), and lowercase letters indicate allocation of populations to distinct groups.

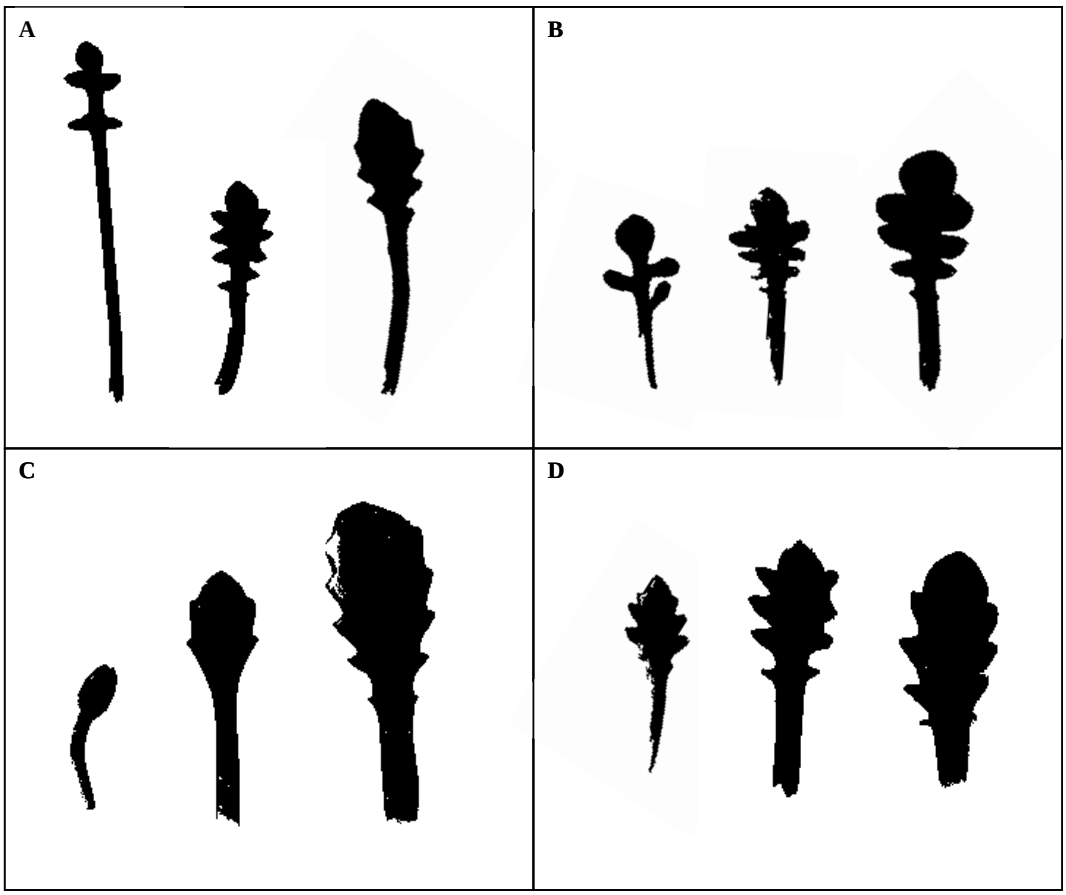


Figure 3.6 Variation in leaf shape within and between populations of *A. l. petraea*. A: Norway7; B: Wales1; C: Norway3; and D: Norway2. For each population three leaves are shown: the most lobed, a leaf in the middle of the range, and the most rounded (left to right) as calculated by the perimeter : area ratio. Norway7 (A) was the population with the most lobed leaves overall, and Norway2 (D) was the population with the most rounded leaves overall.

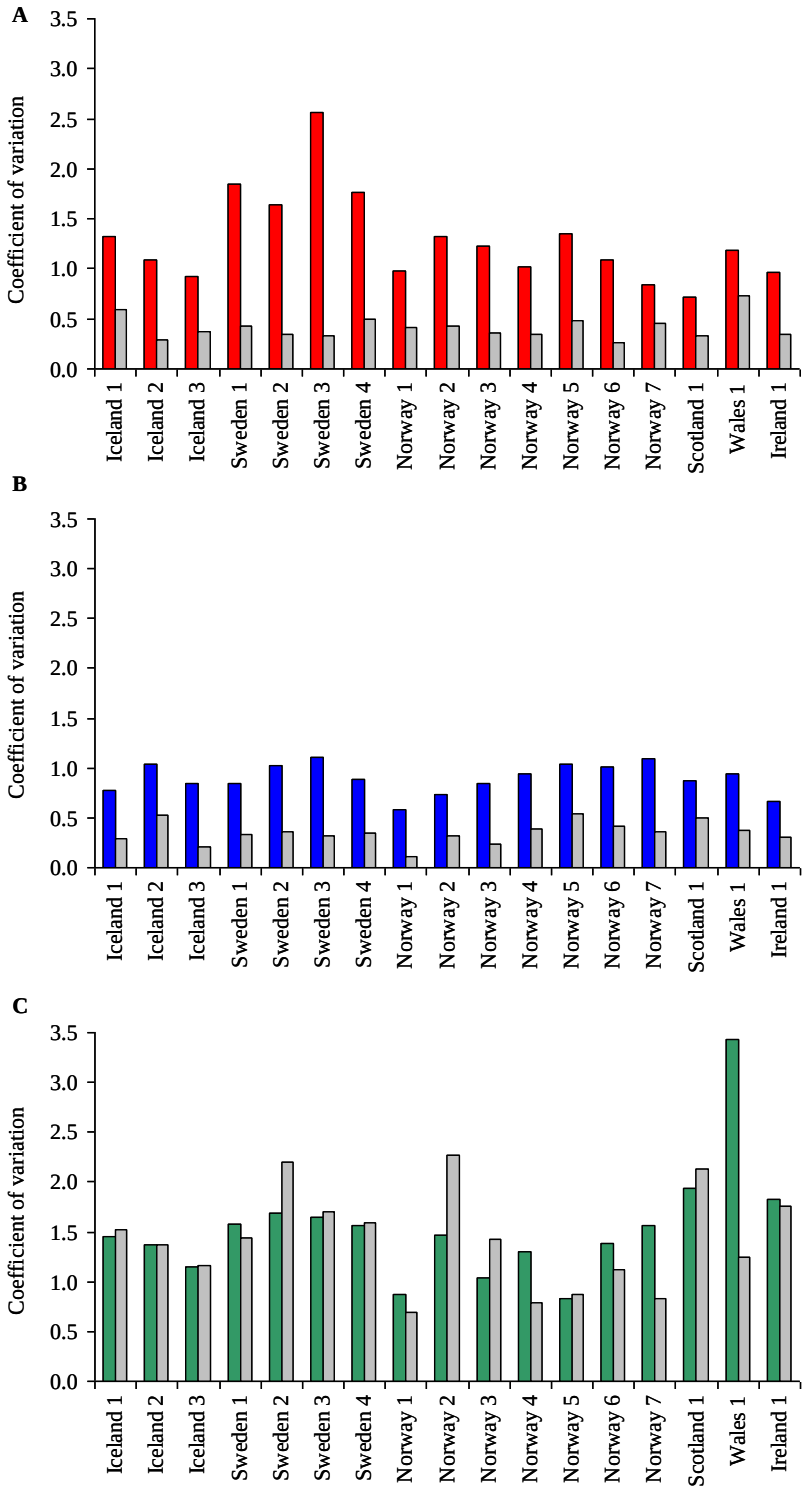


Figure 3.7 Change in coefficient of variation (CV) following selection for growth rate. CV is a measure of the variation relative to the mean. Graphs A, B, and C correspond to plants subjected to hot, cold, and no-selection treatments respectively. Coloured bars represent the CV for the whole population, and grey bars represent the selected proportion only. The sets of plants selected for response to hot and cold temperature stresses illustrate reduced variation following selection (reduced CV following selection). The set of plants where a random proportion were selected (no-selection) do not show this pattern.

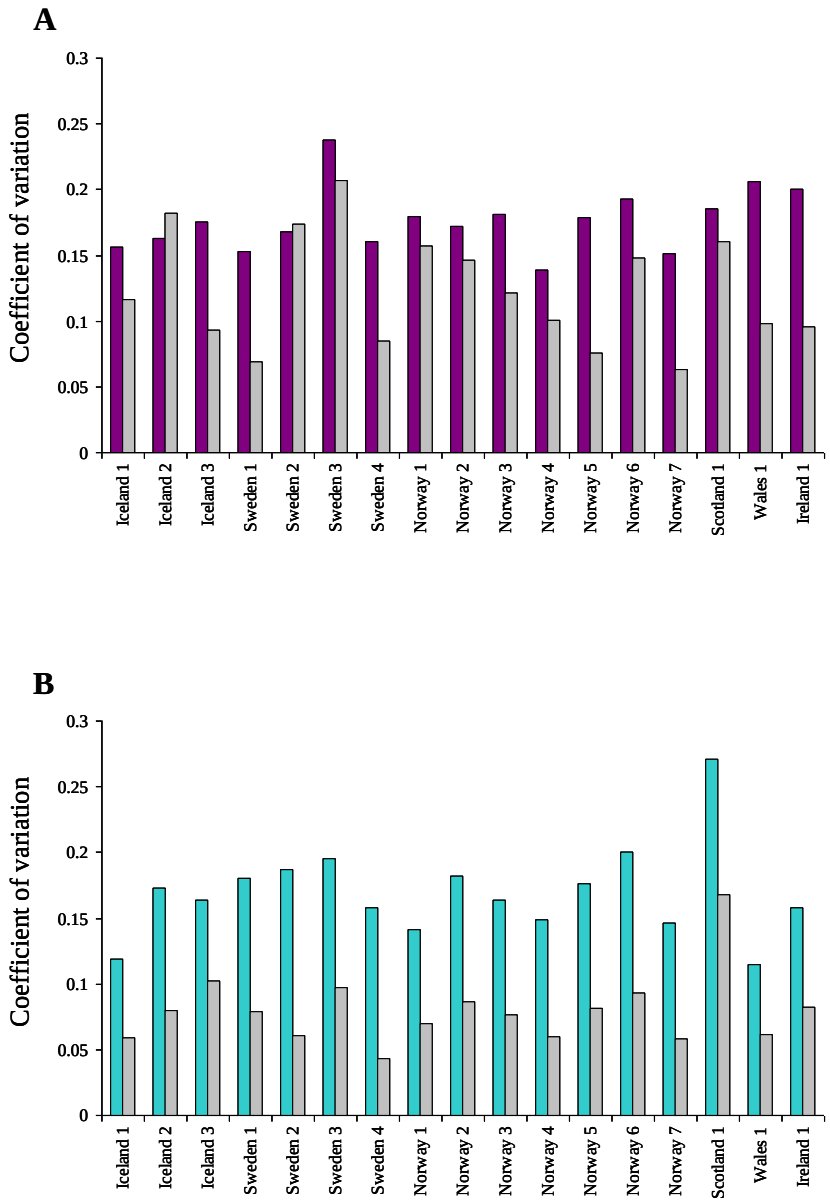


Figure 3.8 Change in coefficient of variation (CV) following selection for leaf shape. CV is a measure of the variation relative to the mean. Graph A corresponds to plants selected for lobed leaves; Graph B corresponds to plants selected for rounded leaves. Coloured bars represent the CV for the whole population, and grey bars represent the selected proportion only. CV values are reduced following selection for the majority of populations; this represents a reduction in variation within the study populations.

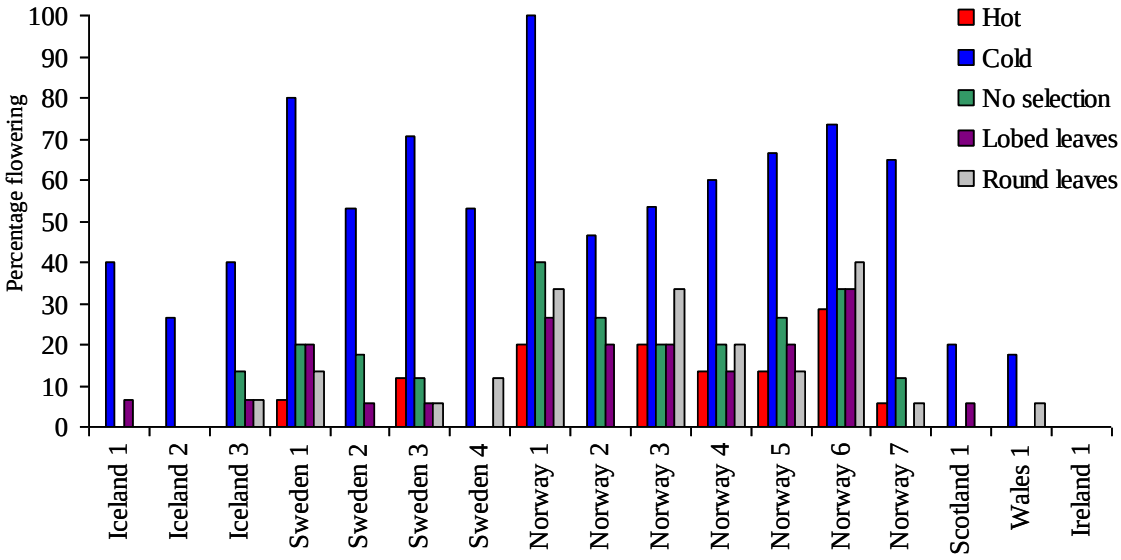


Figure 3.9 Percentage of selected plants that flowered across treatments. This graph illustrates the percentage of plants (selected proportion only) within each population and treatment that were flowering eight weeks post-selection. Percentage of flowering plants was significantly different between the populations ($F_{16} = 3.01$, $P < 0.001$), and the percentage flowering was significantly higher in the cold treated plants than all other treatments ($F_4 = 15.70$, $P < 0.001$).

Population had a significant effect on biomass index change values in cold-treated plants ($F_6 = 15.5$, $P < 0.001$), whilst population effects were not significant in heat-treated plants ($F_5 = 0.901$, $P = 0.480$). Generation had a highly significant effect on growth for both cold- and hot-treated plants ($P < 0.001$), with populations exhibiting higher growth rates in the second generation. This may indicate domestication of the plants towards conditions in the growth chambers. The population x generation interaction was not significant in either dataset (Cold: $F_6 = 0.841$, $P = 0.539$; Hot: $F_5 = 0.411$. $P = 0.841$), suggesting that there is no significant difference in heritability values between the study populations. The selection differential (S) was incorporated into the model as a covariate to account for differences in this value between populations; the population x S interaction was not significant for either cold- or hot-treated plants (Cold: $F_6 = 1.16$, $P = 0.326$; Hot: $F_5 = 1.108$, $P = 0.355$) further indicating the lack of evidence for variation in heritability among populations.

The values for the selection differential (S) were all very close to zero for the no-selection plants (Table 3.6); any deviation from zero will be due to chance and caused by the random selection of plants as parents for the next generation. The majority of response (R) values for the no-selection plants were high: this indicates differences between the growth rates for the first and second generations that, in the absence of selection in these plants, are unexpected. As all R values were positive, they can be presumed to be the result of an unknown environmental effect.

Data from all three generations can be used to investigate changes in the trait mean (mean percentage change in biomass index) over time as illustrated in Figure 3.10. Plants exposed to the cold treatment demonstrate an increase in the trait mean between the first and second generations of the experiment, and either an increase or very little change between the second and third generation, whilst trait means increase and then decrease across the three generations for those plants in the no-selection treatment. Plants exposed to the hot treatment demonstrate an increase in the trait mean between generations one and two, but drastic reductions in the trait mean are apparent for the two populations involved in the third generation of the experiment. These plants did suffer from unintentional drought during the treatment period, and this is likely to have contributed to significant reductions in the recorded growth rates of these plants. These plants will therefore be removed from the analysis of response to selection.

		Generation two			*Generation three*		
		S	**R**	**h^2**	**S**	**R**	**h^2**
Norway1	Hot	1.10					
Norway2	Hot	0.80	0.46	0.58			
Norway3	Hot	1.07	0.30	0.28			
Norway4	Hot	1.25	0.22	0.18	0.69	-1.26	-1.82
Norway5	Hot	0.94	0.45	0.48	0.80	-1.30	-1.63
Norway6	Hot	0.96	0.49	0.51			
Norway7	Hot	0.66	0.48	0.72			
Norway1	Cold	1.10	0.27	0.24	0.81	1.11	1.37
Norway2	Cold	0.74	0.37	0.50	1.01	1.68	1.66
Norway3	Cold	1.14	0.71	0.63			
Norway4	Cold	1.27	0.17	0.13	0.81	1.08	1.32
Norway5	Cold	1.13	0.53	0.47	1.02	1.12	1.10
Norway6	Cold	1.03	0.39	0.38			
Norway7	Cold	0.85	0.43	0.50			
Norway1	No selection	-0.11	0.35	-3.31	-0.18		
Norway2	No selection	-0.23	0.58	-2.49	-0.29		
Norway3	No selection	-0.10	0.49	-4.72			
Norway4	No selection	0.13	0.26	1.98	0.10		
Norway5	No selection	-0.30	0.24	-0.80	0.06		
Norway6	No selection	0.08	0.01	0.07			
Norway7	No selection	0.02	0.30	12.73			

Table 3.6 Calculating parameters of the Breeder's equation following selection. A maximum of three generations of selection were performed in the Norwegian study populations of *A. l. petraea*. Parameters of the Breeder's equation could only be calculated if populations were involved in at least the second generation of the experiment. Parameters are as follows: S: the selection differential; R: response; and h^2: heritability. The h^2 values calculated for the populations in the no-selection treatment are not biologically feasible; this is due to the absence of selection in these populations. Dotted regions indicate lack of data due to removal of populations from the investigation as a result of seed shortages.

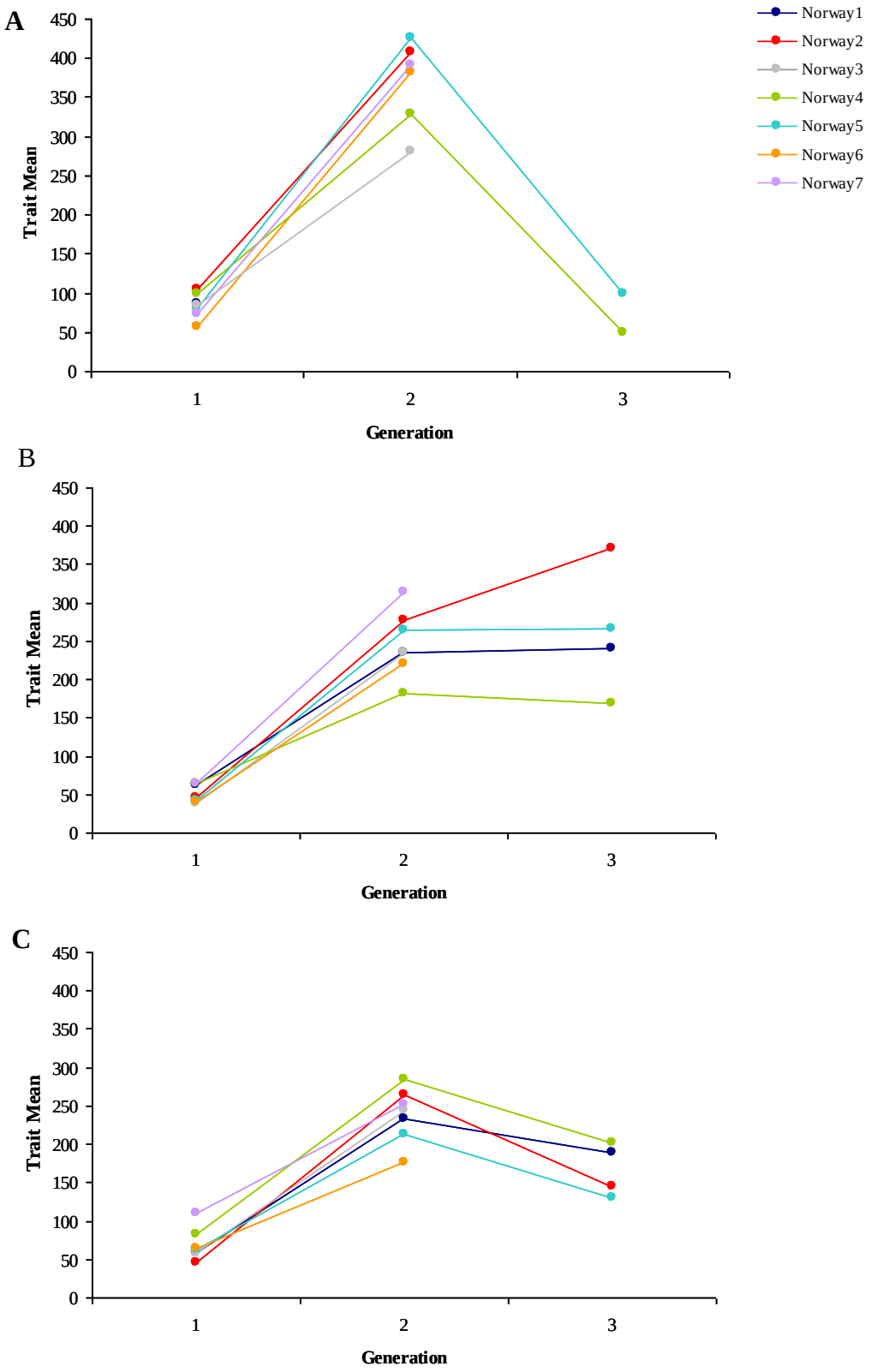


Figure 3.10 Change in mean growth rate over three generations of artificial selection. These graphs illustrate the mean change in biomass index values for each population across the three generations of the artificial selection experiment. Each graph represents plants from a different temperature treatment: A: hot treated plants; B: cold treated plants; C: the no-selection treatment.

3.5 Discussion of the artificial selection experiment

3.5.1 Variations in seed weights and germination success

Laboratory produced seed was found to weigh significantly more than field collected seed (Figure 3.3). Reproductive output is strongly correlated with plant size, with larger plants producing larger seeds (Solbrig and Solbrig, 1984). The plants grown in growth cabinet conditions were larger than at the field sites despite the use of small cell-tray inserts and a low-nutrient compost mix (*pers. obs*.), and this may have resulted in heavier seed being produced from laboratory crossings. A degree of phenotypic plasticity is evident for the seed size trait in response to the conditions under which they develop in a wide range of plant species (Fenner, 1992). Maternal environment effects have been reported to influence early stages of plant development, including seed mass and germination success (Baskin and Baskin, 1998). The more stable conditions within which the laboratory-produced seed developed may have contributed to its higher average weight whilst in the field environmental conditions will have been considerably more variable; factors such as periods of drought may have had an impact upon development of seeds. Plants reared in growth cabinets, and thus grown in a compost mix, are also likely to have had a higher availability of nutrients than those in the field as many *A. l. petraea* populations persist in areas where there is very little topsoil; a greater nutrient availability may have resulted in greater allocation of nutrients to the process of seed production. Maternal nutrient levels are known to affect seed size, and higher nutrient availability generally results in the production of larger seeds (Fenner and Thompson, 2005). Seed filling is significantly more expensive to the plant than production of flowers, and the fraction of reproductive output given to seeds is often reduced in situations of resource limitation (Bazzaz *et al.*, 2000).

Considerable variation in mean seed weight exists between populations in the majority of species that have been studied (Mendez, 1997; Vaughton and Ramsey, 1998; Bazzaz *et al.*, 2000; Jacquemyn *et al.*, 2007); this may be the result of natural selection or genetic drift in the different localities, or a combination of both environmental and genetic factors (Fenner, 1992). Average seed weights differed significantly between *A. l. petraea* study populations, even when considering only laboratory-produced seed, indicating that some populations either allocate a larger percentage of available nutrients to seed production, or that they have an overall greater nutrient availability that enables heavier seeds to be produced. Evidence for variation in the allocation of

resources to reproduction in response to a range of environmental factors is widespread for perennial species (Bazzaz *et al.*, 2000). Consequently, as perennials have more than one opportunity to reproduce, production of viable seeds in less favourable conditions is not essential. The four Swedish populations had some of the lightest average seed weights; these populations persist mainly on rocky substrates along the coast of Sweden that are likely to be relatively nutrient poor environments. Scotland1 and Wales1 are among the populations with the greatest average seed weights; they persist in some of the least extreme climatic conditions experienced by the *A. l. petraea* populations in this study, and may thus have more energy available for seed production. The number of flowers produced per plant in relation to its number of leaves may also affect seed weight. Scandinavian populations produce more inflorescences with greater numbers of flowers than the British populations (*pers. obs.*) and this may enable British populations to invest more resources per seed.

Seed produced in laboratory conditions had a higher average percentage germination rate than that of seed collected from the field (Table 3.5). The greater weight of laboratory seed indicates that this seed contains a greater reserve of energy for use in the germination process, and this increased seed quality is likely to contribute to germination success. Additionally, field seed collection was limited by time constraints; seed could only be collected during scheduled field work visits and therefore may not have been fully developed at the time of harvesting. This is likely to have impacted the percentage germination rates, and may also have affected the observed weights of field-collected seed. In contrast, Riihimäki and Savolainen (2004) used both field-collected and laboratory produced seed in their investigation of flowering time in *A. l. petraea*, yet they observed no consistent differences in germination between the two types of seed.

3.5.2 Growth rate variations

Change in biomass index, when averaged across all study populations, differed significantly between the three temperature treatments. All populations, with the exception of Norway6 and Norway7, exhibited higher growth rates in the hot treatment than in their no-selection controls (Figure 3.4), despite the hypothesis that temperatures of 25^{o}C would pose stress on at least some of the *A. l. petraea* populations. This increased growth rate at higher temperatures is likely to relate to an increased rate of plant metabolic processes; processes such as photosynthesis and respiration may have

occurred at more rapid rates and thus resulted in the more rapid production of plant leaf tissue. As growth rate measurements were based on the actual numbers of leaves per plant, this will have significantly impacted observed growth rate values. These metabolic processes are likely to have been less rapid at the cooler temperatures of the no selection and cold treatments, and thus may explain the lower rates of growth observed in these plants.

Differences in growth rates were also observed between the study populations; the three British populations and two of the Norwegian populations demonstrated the highest average growth rates in terms of change in biomass index over the eight week treatment period. The British populations experience some of the more stable climatic conditions in the field, with a narrower range of temperatures prevailing throughout the year. The basal growth rate of these British populations may therefore be higher than that of populations that experience a broader range of temperatures and thus more extreme climatic conditions in the field. In agreement with these observations, British plants were consistently found to have a significantly greater fresh leaf biomass than Norwegian or Icelandic plants in an investigation concerning a range of nitrogen-deposition rates (Vergeer *et al.*, 2008). It is possible that plants from the different regions have adopted different life history strategies, with British plants investing more in growth and less in reproduction whilst Scandinavian plants invest more in reproduction and less in growth (P. Vergeer, *pers. comm.*). The harsher climatic conditions in Scandinavia may prevent long-term persistence of individuals thus increasing the importance of early reproduction, whilst British plants are likely to survive for longer and consequently can delay reproduction and invest more in growth instead. This hypothesis also correlates with data obtained for flowering and seed production; Scandinavian populations had both higher flowering percentages and greater seed production than British populations (section 3.4.5). The populations exhibiting the lowest average growth rates were all Swedish. For the Swedish populations, there is a large difference between growth rates measured under hot conditions and those measured for the no selection and cold treatments. This may suggest that growth rate in these populations is a plastic trait and thus can be increased or decreased according to the prevailing conditions, or that Swedish populations may be adapted to warm summer conditions and as a result can grow faster at higher temperatures but not at the cooler temperatures. The interaction between temperature treatment and population was found to be significant, indicating that the different study

populations responded differentially, in terms of changes in biomass index, to the three different temperature treatments. The most common interaction was that of significantly higher growth rates for the hot treatment than for the no selection and cold treatments; this was the case in eight of eleven populations for which there was a significant interaction. If the hot temperature parameters used in this investigation did impose stress upon the *A. l. petraea* plants, it seems plausible that observed growth rates for the hot treatment would be lower than for the no selection treatment ('optimal' conditions). The enhanced growth of plants in the heat treatment suggests that many populations are perhaps adapted to warmer climates, and thus were able to take advantage of conditions that were evidently favourable for growth.

3.5.3 Leaf shape differences between populations

Studies of leaf shape differences within a species have mostly been carried out in woody dicot species; correlations between mean annual temperature and number of teeth (lobes) per leaf have frequently been reported with more teeth found on those trees in colder climates (Royer *et al.*, 2005; Royer *et al.*, 2008). Significant differences in leaf shape, measured using the perimeter : area ratio, were observed between populations of *A. l. petraea* under investigation. The trend in leaf shape did not correlate with geographic location of the populations; Norwegian populations possessed both the most lobed and the most rounded leaves (Figure 3.6). Observed leaf shapes also did not correlate with temperature or altitude data. Shape of rosette leaves in populations of *A. l. petraea* has been reported to show distinct patterns of variation between populations, unlike other morphological characteristics that are more irregular and appear to vary according to environmental conditions and life stage of the plant (Jonsell *et al.*, 1995). The study by Jonsell *et al.* (1995) involved populations of *A. l. petraea* from four countries and found characteristic differences in leaf shape between populations, both in the field and when grown in culture. Differences were also observed between populations that grow in close proximity to one another, and it is suggested that vegetative propagation may contribute to the maintenance of genetic variation (Jonsell *et al.*, 1995). Differences in trichome coverage between populations were also observed, with the majority of Swedish populations having glabrous leaves (Jonsell *et al.*, 1995). Populations that had the greatest proportion of pubescent plants in this study were Iceland1, Ireland1, Norway1 and Norway 2, and not all Swedish plants were glabrous (*pers. obs.*).

3.5.4 Proportion of flowering plants

The proportion of *A. l. petraea* plants that were flowering eight weeks post selection differed significantly between the populations under investigation. Norwegian and Swedish populations tended to have higher proportions of flowering plants than British populations at the same time point in the first generation of this experiment. Variation between populations in flowering characteristics, including time to flowering, have been reported in a number of plant species including *Beta vulgaris* (Boudry *et al.*, 2002) and the perennial herb *Lythrum saliceria* (Olsson and Ågren, 2002). Such variations have previously been reported for *A. l. petraea*. For example, Riihimäki and Savolainen (2004) found highly significant differences in the probability to flower between Norwegian and German populations of *A. l. petraea*, as well as clinal variations in time to flowering under growth cabinet conditions. In opposition to results presented here they found that southern populations had both a greater probability of flowering, and flowered more rapidly, than northern populations, with more pronounced differences in short day rather than long day conditions (Riihimäki and Savolainen, 2004). Only four populations were included in their study (Plech from Germany; Bohemia from the Czech Republic; Spiterstulen from Norway, and; Stubbsand from Sweden) whilst this investigation involved seventeen populations, two of which overlap with the published study (Spiterstulen - Norway3, and Stubbsand – Sweden3). The daylength used also differed between the studies, and these factors may have contributed to the observed disparity in findings. Clinal variation in flowering time is not evident in *A. thaliana* (Stenøien et al., 2002) and thus such variations in the proportion of flowering plants between *A. l. petraea* populations are indicative of adaptation to local conditions through natural selection, perhaps due to the isolated nature of the populations.

In this investigation the most pronounced difference in flowering characteristics was between those plants that had experienced the cold stress conditions and the remaining four sets of plants. Riihimäki and Savolainen (2004) investigated the effect of vernalisation on flowering in a range of *A. l. petraea* populations under growth cabinet conditions, and were surprised to find no effect of this cold treatment on flowering probability; this is in contrast to data presented here that demonstrate significantly higher flowering percentages in the set of plants that experienced the cold treatment, although plant age and treatment lengths are not directly comparable. The cold treatment used in Riihimäki and Savolainen's study (2004) was also substantially colder (4^{o}C) than that used in this study (15^{o}C day: 12^{o}C night) The published study also

investigated flowering characteristics of a number of populations in field conditions; when plants experienced a natural winter in a common garden experiment in Finland, all populations flowered the following year. It is thought that this effect may be due to the length and intensity of the winter, but it could also be related to the age of the plants (Riihimäki and Savolainen, 2004). A number of British populations of *A. l. petraea*, those that exhibited the lowest flowering percentages in growth cabinet conditions, were maintained at the university gardens in Leeds over the winter of 2008 – 2009; these plants flowered prolifically the following spring (*pers. obs.*). Average minimum temperatures in the greenhouse used were below 3^{o}C for both December and January with absolute minimums reaching -3 and -4^{o}C respectively, and thus it appears, in concurrence with results reported by Riihimäki and Savolainen (2004), that prolonged exposure to cold temperatures promotes flowering in *A. l. petraea*. Light intensity under greenhouse conditions will have been greater than in the growth cabinets, and this may also have influenced flowering.

3.5.5 Response to selection

The selection pressure imposed within this investigation successfully reduced variation within populations, as shown by the reduction in coefficient of variation (CV) values in the selected proportion of plants as compared to the population as a whole (Figures 3.7 and 3.8). Selection of the top 25% of plants in terms of their percentage change in biomass index values resulted in higher population trait means and reduced within-population variation; as expected, this pattern was not evident for those plants involved in the no-selection treatment. Estimations of response to selection could only be made once data from second generation plants had been collected. In this investigation it was only possible to collect data from a second generation of plants for the seven Norwegian populations (see 3.3.9). Using this data the selection differential (S), response (R), and a heritability value (h^2) could be calculated for each population in each of the treatments; these are detailed in Table 3.6. It was hypothesised that populations from the coldest climates may have been pushed to their limits of adaptation to cold thus resulting in lower levels of heritability, and also that populations from the hottest climates may have reached their limits of adaptation to hot temperatures thus resulting in these populations having lower h^2 values. There is some correlation between h^2 values and temperature conditions in the field; number of degree days of <-5^{o}C negatively correlated with h^2 values for plants selected under cold conditions. This indicates a possible reduction in adaptive potential in climatically marginal populations.

For plants selected under hot conditions, h^2 values correlated positively (but not significantly) with number of degree days >20^oC; this is contrary to the predicted relationship. Despite the significant correlation in cold selected plants, described above, a general linear model incorporating standardised data and S values for each population, found no significance for the population x generation interaction term or the population x S interaction term thus indicating a lack of evidence for significant variation in heritability between populations. Although the Norwegian populations experience a range of climatic conditions in the field, with those at high altitudes experiencing cold winters and those at low altitudes experiencing milder winters, a far greater range of conditions occur across the north west European range of *A. l. petraea* as illustrated in Figure 1.2. Ideally, a larger extent of this range of conditions would have been included in the experiment, and incorporation of this greater variation in field conditions may have proved more successful in identifying variations in heritability across the populations. Unfortunately, the low flowering percentages and consequential low seed production in many of the populations resulted in their elimination from the experiment at the end of the first generation. It was not possible to incorporate a greater initial number of populations as the first generation of this large scale experiment was performed at the maximum capacity for growth cabinet availability. The time that elapsed waiting for the majority of populations to flower had to be carefully considered against the need to perform more than one generation of this experiment; the most practical option was to perform the second generation of the experiment using only a subset of populations once it appeared highly unlikely that the remaining populations would produce sufficient seed to continue. A similar trade-off was performed with regard to the third generation and, as a result, a maximum of four populations per treatment were included in this final generation. All heritability (h^2) values exceeded one in plants experiencing the cold treatment; R values in this generation were particularly high indicating large differences in growth rate between plants in the second and third generations. This is likely to have been caused by an environmental difference between the second and third generations and, as this is an unknown entity, its effect cannot be separated out. The calculated values for heritability, however, are similar between the study populations indicating that even if the true heritability could be calculated, values would not differ significantly between populations.

Those populations exposed to the cold treatment are the only ones for which the increase in the trait mean between the first and second generation is either maintained or,

in the case of Norway2, increased further in the third generation. This suggests that these populations are responding to the imposed selection, although the lack of continued trend for increased trait mean in the majority of populations implies that the selection limits may have been reached. The continued increase in Norway2 indicates that this population may have greater variation for the trait under selection, and can thus continue to respond to the imposed selection. Additional generations of this experiment would be required to test both of these hypotheses, and this was not possible due to logistical limitations. Populations in the no-selection treatment do not exhibit this upward trend in trait mean across all three generations, although the trait means do fluctuate (see Figure 3.10). This indicates that a larger number of generations of selection would be required in order to distinguish true changes in the trait mean from random fluctuations. It is not possible to discuss trait mean trends for plants exposed to the hot treatment as only two generations of reliable data were collected.

3.5.6 Conclusions

Using an artificial selection experiment to investigate differential responses to selection in populations from throughout the range of *A. l. petraea* has proved problematic. The transition from vegetative growth to flowering under growth cabinet conditions occurred in an unexpectedly low proportion of plants resulting in very limited seed production. As a consequence, many populations could not be included in the second and third generations of this investigation. Data collected during the first generation of the experiment, involving over 3000 plants, revealed some interesting observations including differences in weight and germination success between laboratory-produced and field-collected seeds, distinct differences in leaf shape between study populations, and differences in growth rates between populations and temperature treatments. The artificial selection experiment continued to a second and third generation with a subset of the initial populations, all of which were from Norway. Aside from the correlation of heritability with some climatic variables, the results obtained for these populations did not provide evidence for significant variation in heritability between populations, perhaps due to the limited number of generations of selection or to the limited ecological range of the populations involved. The geographically isolated populations of this species, and its self incompatibility, made it a desirable model species for this experiment. Unfortunately, despite the success of *A. l. petraea* as a model species for investigation of local adaptation (Davey *et al.*, 2008; Vergeer *et al.*, 2008; Leinønen *et al.*, 2009), results presented here suggest that it is not particularly suitable for a multi-

generational experiment involving a requirement for seed production under growth cabinet conditions. With a better understanding of how to reliably trigger flowering, future experiments of this type may prove increasingly successful.

CHAPTER 3: SUMMARY

- An artificial selection experiment was designed incorporating seventeen populations of *A. l. petraea* and five different treatments. The level of replication, risk of inbreeding, and required selection pressure were carefully considered.
- Trait measurements were obtained for all populations in the first generation; significant differences in growth rate and leaf shape were observed between the populations under investigation.
- *A. l. petraea* plants exhibited particularly low flowering percentages under growth cabinet conditions. This resulted in seed shortages that prevented continuation of the experiment for the majority of populations.
- Three generations of the experiment were carried out for a small subset of Norwegian populations only; limited evidence was found for differences between populations in heritability (h^2) for the growth rate trait in either hot or cold treated plants.

4. Differential regulation of gene expression between populations of *A. l. petraea* in response to chilling temperatures.

AIMS

- Collect leaf tissue from chilling treated and untreated plants of *A. l. petraea* from a range of study populations.
- Obtain high quality RNA from collected tissue samples for cDNA synthesis.
- Obtain *A. l. petraea* coding region sequences of genes of interest to enable design of species-specific qRT-PCR primers.
- Perform qRT-PCR expression analysis for a range of genes previously identified as chilling responsive and differentially regulated between *A. l. petraea* populations.

4.1 Introduction

4.1.1 Cold temperatures and local adaptation in *Arabidopsis*

Temperature is one of the most important climatic factors determining the distribution of species, particularly those of plants (Thomashow, 1999). As a consequence of their sessile nature, plants that are unable to function successfully under a particular climatic regime will not be able to persist at that location. Populations across the range of a species are exposed to a variety of environmental conditions, and the larger the range the more diverse the conditions experienced. This can result in significant variation in morphology, physiology, and development (Zhen and Ungerer, 2008a). Exposure to extreme climatic conditions places a selective pressure on individuals, and the forces of natural selection may consequently result in adaptation to local environments (Endler, 1977). Low spring and autumnal temperatures limit the geographical distribution of *A. thaliana* (Koornneef et al., 2004) and consequently Hannah *et al.* (2006) suggest that temperature may be a habitat parameter that could influence the genetic composition of local populations through natural selection.

Adaptation to local climatic conditions is evident for *A. thaliana* ecotypes. Hannah *et al.* (2006) reported that the freezing tolerance of *A. thaliana* accessions correlated with the winter temperatures experienced by these nine accessions; those experiencing colder temperatures demonstrated enhanced acclimated freezing tolerance. A wider study of

over 70 accessions of *A. thaliana* from across its geographic range further demonstrates clinal variation in freezing tolerance. This variation was observed in both acclimated and non-acclimated plants, also indicating differences in acclimation capacity between accessions (Zhen and Ungerer, 2008a). *A. thaliana* ecotypes have also been reported to show differential responses to chilling (14^{o}C), cold (6^{o}C), and freezing temperatures across a range of traits including growth characteristics, pigmentation patterns, and chlorophyll content (Hasdai *et al.*, 2006).

Differences in expression of a gene within a species are termed expression level polymorphisms (ELPs) (Doerge, 2002). A study of seven *A. thaliana* accessions revealed ELPs in all pairwise comparisons with, on average, at least 2000 genes identified as differentially expressed between any pair of accessions (Kliebenstein *et al.*, 2006). Gene ontology (GO) annotations of these genes revealed many have involvement in biotic and abiotic responses, stress responses, and signal transduction. These ELPs were found to correlate with sequence divergences between the accessions. Kleibenstein *et al.* (2006) suggest that these gene sequence differences, and their consequential ELPs, have evolved in response to differences in natural environments between the accessions. He *et al.* (2008) also report transcriptional variation between Chinese ecotypes and *Col-0.* Differences in the genetic sequences of *A. thaliana* ecotypes have been reported both across the genome (Hardtke *et al.*, 1996; Innan *et al.*, 1997) and at specific loci (Le Corre *et al.*, 2002), which suggests that selection pressures may differ between geographic localities.

It is well documented that a large number of *Arabidopsis* genes are responsive to cold acclimation or chilling temperatures (Thomashow, 1999; Seki *et al.*, 2001; Fowler and Thomashow, 2002; Kreps *et al.*, 2002; Seki *et al.*, 2002). Genes can be transcriptionally up-regulated or down-regulated, and a large number of cold-responsive genes have now been characterised (section 1.5.1). The extent of these changes in expression in response to cold stress has been reported to differ between different *A. thaliana* ecotypes (He *et al.*, 2008). In that study, the majority of genes exhibiting transcriptional changes were shared between the different ecotypes indicating a conserved response to cold, although each population did display some uniquely changing genes. The extent of these transcriptional changes, however, differed between ecotypes suggesting adaptation to local climatic conditions. Differences in expression between accessions, or ecotypes, have also been reported for several specific genes known to have

involvement in the cold response of *Arabidopsis*, including both *CBF* (*C-REPEAT BINDING FACTOR*) and *COR* (*COLD-REGULATED*) genes (Le *et al.*, 2008; McKhann *et al.*, 2008). Reduced freezing tolerance in southern populations of *A. thaliana* has been partly attributed to relaxed selection on *CBF* genes; both sequence differences and variation in expression levels have been observed between northern and southern populations (Zhen and Ungerer, 2008b). These findings suggest that both gene sequence and expression differences are the result of differential natural selection across climatic environments.

Cold acclimation or chilling temperatures have also been shown to cause extensive changes in the metabolome of *Arabidopsis* and, additionally, the extent of such metabolomic changes has been shown to differ between accessions (Cook *et al.*, 2004). Davey *et al.* (2008) detected significant natural variation in the metabolome between geographically distinct populations of *A. l. petraea*, indicating genetic differentiation and local adaptation. In addition, population-specific metabolic phenotypes associated with cold temperature exposure have been identified in *A. l. petraea* (Davey *et al.*, 2009). Transcriptomic differences between *A. l. petraea* populations have been identified, both under control conditions and in response to chilling through the use of microarrays (C. J. Lilley, *pers. comm.*). As these differences were detected using just two study populations, it was decided to expand this work to incorporate populations from a broader geographical spectrum using quantitative reverse-transcriptase PCR (qRT-PCR) to detect gene expressional differences.

4.1.2 qRT-PCR methodologies

Quantitative reverse transcriptase PCR (qRT-PCR) involves measurement of the amplified PCR product at each cycle throughout the reaction (Gachon *et al.*, 2004). A fluorescent reporter molecule is used and the amount of amplified product is linked to the fluorescent intensity observed. This enables either absolute or relative quantification of the starting material. It has become a very widely used method due to its rapidity, specificity, and reliability (Gachon *et al.*, 2004). A common use of qRT-PCR experiments is to verify microarray results (Wurmbach *et al.*, 2003; Gachon *et al.*, 2004), but a wide range of other uses have been reported. These include the quantification of genetically modified organisms in food products (Hubner *et al.*, 2001) and the detection and identification of bacterial infections on plants (Cubero and Graham, 2005).

The first qRT-PCR experiments used ethidium bromide, and the fluorescence of reactions after an increasing numbers of cycles was recorded using a video camera (Higuchi *et al.*, 1993). SYBR Green I has since been developed as a fluorescent dye for use in qRT-PCR experiments due to its high affinity for double-stranded DNA (Schneeberger *et al.*, 1995). Fluorescence of SYBR Green I increases by over 1000-fold once bound to double-stranded DNA (Stratagene, 2007). An increase in the quantity of double-stranded DNA provides more opportunities for the binding of SYBR Green I, and thus fluorescence increases proportionally to the DNA concentration (Stratagene, 2007). Due to the simplicity of assay design, along with cost considerations and high accuracy, SYBR Green I has become a frequently used reporter molecule (Wurmbach *et al.*, 2003). One inherent limitation of SYBR Green I is that, because the dye will bind to any double stranded DNA, specificity of the reaction is determined solely by the primers used.

Various other strategies can be used to increase the specificity of qRT-PCR experiments, including the use of a range of probe-based chemistries. An internal probe, specific to the amplicon of interest, is added to the reaction. This will only fluoresce when the probe hybridises to the target sequence of interest; prior to this the reporter dye is quenched and thus does not fluoresce. The most commonly used are linear probes (e.g. TaqMan), but more structured probes containing stem-loop structure regions to further enhance specificity can also be used (e.g. Molecular Beacon) (Stratagene, 2007). A significant advantage of probe-based experiments is that a number of targets can be quantified in one reaction; each specific probe can be attached to a different reporter dye (Stratagene, 2007). The use of probes in plant studies will only be valuable if many experiments are to be carried out on the same target due to the costs involved in probe design (Gachon *et al.*, 2004).

The threshold cycle (C_t) is central to the quantification of DNA through qRT-PCR. It refers to the point at which the fluorescent signal emitted is deemed significantly greater than the background signal, and this will always occur in the exponential phase of amplification (Bustin, 2000). A greater quantity of DNA in the starting material will result in the C_t value being reached more quickly (Gibson *et al.*, 1996). C_t values can be compared between samples (e.g. treated and untreated) to ascertain the relative quantity of template (Livak and Schmittgen, 2001), or can be used to accurately quantify the absolute concentration of DNA in the original template with the aid of a standard curve.

A standard curve involves amplification of a range of dilutions of a template with known concentration. The C_t values obtained for each dilution are plotted against the log of the initial template concentration to give the standard curve; this should be a linear regression line if reaction efficiency is maintained across all concentrations. C_t values obtained for experimental samples can be compared to the standard curve and consequently concentration of the unknown samples can be obtained (Bustin, 2000; Stratagene, 2007). Standard curves are also used to determine the efficiency of primer pairs; the optimum efficiency of 100% equates to a doubling of the target amplicon with each PCR cycle.

4.1.3 Interpretation of qRT-PCR results

It is important to minimise any technical variability between samples such as differences in the starting quantity of RNA, and thus normalisation of qRT-PCR data is essential (Nolan *et al.*, 2006). The most common approach is the use of a reference gene or normaliser that is invariantly expressed across experimental conditions and tissue types (Thellin *et al.*, 1999). One of the first genes to be used as a reference was ß-actin (Kreuzer *et al.*, 1999) but many others are commonly used. Recent work by Czechowski *et al.* (2005) identified hundreds of genes that were more stably expressed than the traditionally used reference genes in *Arabidopsis* for a range of experiment conditions and tissue types. Reference genes must therefore be carefully selected based on their suitability for the specific study in question (Pfaffl, 2004). Expression levels of the chosen reference gene are constant, and thus any variation in the C_t of the reference gene in experimental samples can be attributed to other sources of variation (Stratagene, 2007). C_t values for the genes of interest can therefore be corrected for such variation and any remaining differences between treated and untreated samples can be attributed to changes in expression.

The specificity of qRT-PCR reactions can be monitored using a dissociation curve (or melt curve). The melting temperature of DNA depends on its size and nucleotide composition and thus different PCR products will have different melting temperatures (T_m). At the end of a qRT-PCR cycle, samples are melted, annealed at 55^{o}C, and then fluorescence data is collected whilst products are subjected to gradual increases in temperature. Fluorescent dye is released when a product melts, and a melt curve specific to each product is consequently produced by plotting fluorescence as a function of temperature (Ririe *et al.*, 1997; Nolan *et al.*, 2006; Stratagene, 2007). These melt

curve peaks are analogous to bands on an electrophoresis gel (Nolan *et al.*, 2006). A reaction containing a number of peaks is indicative of non-specific amplification, perhaps due to primer dimer formation.

The main objective of this work was to ascertain whether or not *A. l. petraea* populations from differing geographic locations exhibit differential levels of expression of specific genes. SYBR Green I qRT-PCR assays are used to determine the relative expression levels of nine genes in thirteen different populations of *A. l. petraea*. Expression levels are compared between these populations in untreated samples and, additionally, changes in expression following exposure to chilling temperatures are studied. Genes were selected for study based on results from previous microarray analysis of two *A. l. petraea* populations, Helin in Norway and Leitrim in Ireland (Dr. C. J. Lilley, unpublished). All genes chosen were found to be responsive to chilling and were differentially regulated between the two initial populations. Relative quantification of gene expression was calculated following normalisation to a reference gene. Efficiency of all primer pairs was determined using a standard curve, and specificity of reactions was subject to dissociation curve analysis.

4.2 Materials and Methods

4.2.1 Population choice and plant growth

Thirteen populations of *A. l. petraea* were chosen for study based on country of origin and seed availability; populations are detailed in Table 2.1. Trays containing 40-cell tray inserts were filled with sterilised Nottingham compost mix (containing three parts Sinclair potting growing medium: three parts J Arthur Bower's John Innes No.2 Compost: one part Sinclair Horticulture standard grade perlite (2.0 – 5.0 mm), all William Sinclair Horticulture Ltd). A corner cell was removed to allow for watering from the base. Trays were surface watered using a watering can with rose prior to, and following, seed sowing. Two seed were sown per cell. Trays were watered from the base by adding water to the tray, covered with a propagator lid and placed in a Sanyo growth cabinet (16 h day, 8 h night, 20^{o}C:20^{o}C day : night temperature, 50% humidity). Propagator lids remained on trays until the majority of seedlings had emerged (between one and two weeks). Individual seedlings were pricked out into three inch pots of Nottingham compost mix and maintained in a Sanyo growth cabinet (16 h day, 8 h night, 20^{o}C:20^{o}C day : night temperature, 50% humidity) until plants were ten weeks old. Ten plants were grown for each study population; five as control plants and five as treated plants.

4.2.2 Chilling treatment and harvesting of leaf material

At ten weeks, five plants from each population experienced an 8 h chilling treatment at 4^{o}C. A further five plants from each population remained in the Sanyo cabinet at 20^{o}C during this 8 h period. Following the 8 h treatment, two leaves of 3 - 4 cm in length were harvested from each individual plant and snap-frozen in liquid nitrogen.

4.2.3 Processing of leaf tissue

RNA was extracted from individual leaves and quantified as described in sections 2.3.4 and 2.3.5 respectively. If the nucleic acid concentration was less than 100 ng μl^{-1}, RNA extraction was repeated using a second leaf sample. For each population and treatment, the qRT-PCR analysis was carried out using pooled RNA from five individual plants. Four micrograms of RNA from five individual plants for each treatment was combined to create a pooled sample. In-solution DNase treatment (as 2.3.4) was carried out on these pooled samples, in addition to the on-column treatment performed during RNA extraction, to ensure effective removal of any genomic DNA contamination. The

integrity of the pooled samples was assessed using a Bioanalyser (as 2.3.6) prior to synthesis of cDNA, using 1 µg of each pooled sample, for use in qRT-PCR (as 2.3.7).

4.2.4 Selection of genes for study

Microarray analysis had been performed by Dr. Catherine J Lilley using *Arabidopsis* ATH1 GeneChip® Arrays (Affymetrix). This microarray analysis enabled gene expression in two populations, Helin in Norway and Leitrim in Ireland, to be investigated following an 8 h chilling period at 4°C. Results from these microarray experiments have been validated using qRT-PCR (Dr C. J. Lilley). A set of genes was identified that exhibited significantly higher levels of expression following the chilling treatment in both Helin and Leitrim plants; these have been termed 'general cold up-regulated' genes. A set of genes was also identified that exhibited significantly lower levels of expression following the chilling treatment in both Helin and Leitrim plants; these have been termed 'general cold down-regulated' genes. These gene sets, identified through Dr. C. J. Lilley's experiments, were used to select two general cold up-regulated genes (At1g08920 and At2g46830) and one general cold down-regulated gene (At3g48360) for study here (Table 4.1).

Additionally, the microarray analysis (Dr. C. J. Lilley) identified a set of genes responsive to the 8 h chilling treatment, but differentially regulated between the two populations; these were termed 'interaction genes'. Included within the interaction genes are: (i) genes that were up-regulated in response to chilling but to a significantly greater extent in one population than another; (ii) genes that were down-regulated in response to chilling but to a significantly greater extent in one population than another; (iii) genes that were significantly up- or down-regulated in one population but not responsive to temperature in the other; and (iv) genes that were up-regulated in one population and down-regulated in the other. A subset of six of these interaction genes was selected for study in thirteen additional populations. Figure 4.1 illustrates the expression levels (from the microarray analysis) of these six genes in untreated and cold treated plants for the Helin and Leitrim populations. The genes chosen for study in this chapter, and their TAIR annotations, are detailed in Table 4.1. In this chapter, qRT-PCR assays were not performed for the Helin and Leitrim populations as the microarray results had already been validated using this technique (Dr C. J. Lilley).

Gene (AGI Code)	Gene Name	TAIR Annotation	Gene Type
At1g13320	*PDF2*		Normaliser
At4g05320	*UBQ10*		Normaliser
At1g08920		Sugar transporter, putative	General cold up
At2g46830	*CCA1*	CCA1; Circadian Clock Associated 1; transcription factor.	General cold up
At3g48360	*BT2*	BT2; BTB and TAZ domain protein 2; protein binding / transcription factor / transcription regulator.	General cold down
At2g40460		Proton-dependent oligopeptide transporter (POT) family protein. Transporter activity.	Interaction: Up more in Helin
At5g04410	*NAC2*	NAC2; *Arabidopsis* NAC domain containing protein 78; transcription factor.	Interaction: Up more in Helin
At5g05440		Similar to unknown protein.	Interaction: Up more in Leitrim
At1g64890		Integral membrane transporter family protein; transporter activity.	Interaction: Up more in Leitrim
At1g06110	*SKIP16*	SKIP16; SKP1/ASK-Interacting Protein 16; protein binding.	Interaction: Up in Helin, not Leitrim
At2g04040	*AtDXTX1*	AtDTX1; antiporter / multidrug efflux pump / multidrug transporter / transporter.	Interaction: Up in Leitrim, not Helin

Table 4.1 Summary of the genes studied in *A. l. petraea* following chilling treatment. Genes have been assigned to various types as follows: Normaliser; used for normalising of qRT-PCR results, General cold up and general cold down; these genes were found to be up- or down-regulated in both Helin and Leitrim populations of *A. l. petraea* following chilling temperatures through initial microarray analysis, Interaction; these genes were differentially regulated between the two populations following the chilling treatment as reported in initial microarray analysis.

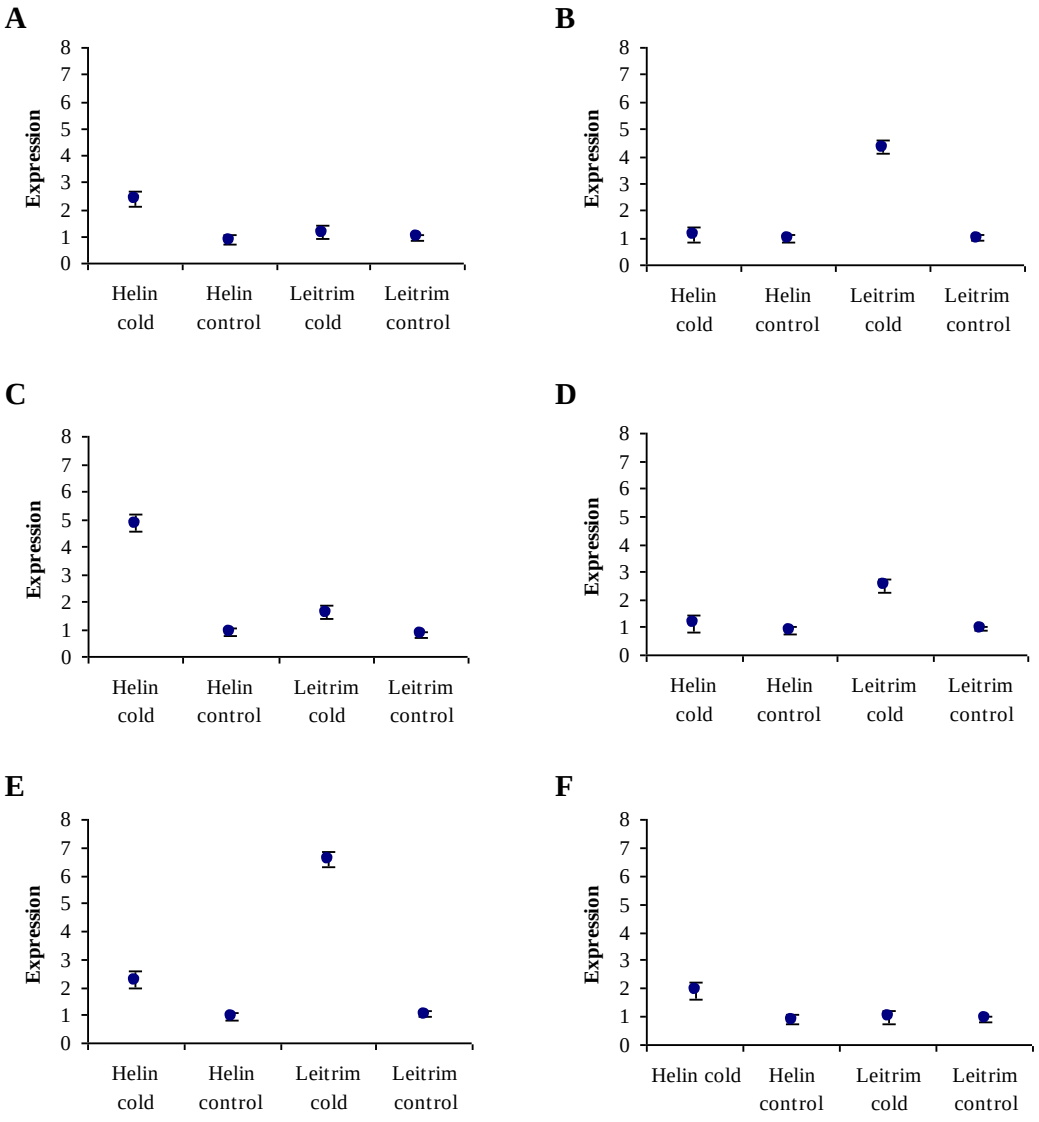


Figure 4.1 Signal values of differentially regulated genes from initial microarray analysis. Microarray analysis was performed by Dr. C. J. Lilley, as described in section 4.2.4, to investigate gene expression in two populations of *A. l. petraea*, Helin in Norway and Leitrim in Ireland, following an 8 h chilling period at 4°C. A number of genes were found to be responsive to the chilling treatment, but were differentially regulated between the two populations. These genes have been termed 'interaction genes' and six (shown here) were chosen for further study in a number of additional populations. Data points refer to the mean signal values for each category across the three arrays (± SE). A) At5g04410 and C) At2g40460: up-regulated in both Helin and Leitrim cold-treated plants but to a greater extent in Helin; B) At2g04040: up-regulated in Leitrim but not Helin; D) At1g64890 and E) At5g05440: up-regulated in both Helin and Leitrim cold-treated plants but to a greater extent in Leitrim; and F) At1g06110: up-regulated in Helin but not Leitrim.

4.2.5 Cloning and sequencing genes of interest

Sequences of *Arabidopsis thaliana* genes were obtained from sequence viewer (SeqViewer) at TAIR (http://www.arabidopsis.org). Primers were designed using Primer3 software (http://frodo.wi.mit.edu/) to amplify regions of approximately 500 bp from the 3' end of the coding sequence for each gene of interest (Table 4.2). *A. l. petraea* cDNA was PCR amplified (as 2.3.8), products analysed (as 2.3.9), PCR purified (as 2.3.11), and precipitated (as 2.3.12). These DNA fragments were then cloned into the pSTBlue-1 vector as detailed in section 2.3.14.1 and transformed into competent DH5α *E. coli* (as 2.3.17). Plasmid DNA was extracted from *E. coli*, digested, and analysed by agarose gel electrophoresis (sections 2.3.21; 2.3.15; and 2.3.9, respectively). Colonies containing an insert of the expected size were sequenced (as 2.3.13) and consequently *A. l. petraea* coding region sequence was obtained for each gene of interest. The *A. l. petraea* sequences for the housekeeping genes *PDF2*, At1g13320, and *UBQ10*, At4g05320, were also obtained using this method.

4.2.6 Design of primers for qRT-PCR

Primers for the two 'general cold up-regulated' genes and the 'general cold down-regulated' gene were designed using the *A. thaliana* gene sequences obtained from sequence viewer (SeqViewer) at TAIR (http://www.arabidopsis.org). The *A. l. petraea* sequences for the six interaction genes and the two housekeeping genes were obtained as detailed in section 4.2.5. Primers for qRT-PCR were designed to these sequences using Primer3 software (http://frodo.wi.mit.edu/) to the following criteria: T_M of 60°C ± 1°C, to amplify fragments 60 – 190 bp in length, primer sequences of ~20 nucleotides and guanine-cytosine content of ~50%. Primers were designed to amplify close to the 3' end of the coding sequence and, where possible, spanned an intron. A full list of genes and primer sequences are detailed in Table 4.3**.**

4.2.7 Measurement of gene expression through qRT-PCR

qRT-PCR was performed on cDNA synthesised from pooled RNA samples as detailed in section 4.2.3. The expression levels of target genes were normalised to the housekeeping gene *PDF2,* At1g13320; the qRT-PCR primers for this gene were based on those described by Czechowski *et al.* (2005) but were altered to ensure that they were specific to the *A. l. petraea* sequence obtained here. Log_2 fold change in gene expression for all target genes was obtained for each population by comparing the

Gene (AGI Code)	Gene Name	Forward primer 5' – 3'	Reverse primer 5' – 3'	Expected product size (bp)
At1g13320	*PDF2*	CGT-TAC-TGC-CAG-CCA-TTG-TA	TGC-GAA-ATA-CCG-AAC-ATC-AA	530
At4g05320	*UBQ10*	AAC-AAT-TGG-AGG-ATG-GTC-GT	CAG-ATA-ATA-AAA-ACG-AAG-CGA-TGA	506
At5g04410	*NAC2*	ATC-TCA-GCT-GTG-CTC-AAC-AAG	TTA-CCG-AGC-AGA-CAC-CAT-GG	646
At2g04040	*AtDTX1*	CAT-ACT-CTG-CTC-AGG-TCT-TC	GAG-ACA-ATT-CTC-TTC-CTT-GC	591 (848 genomic DNA)
At2g40460		GAT-TGC-TTC-GGC-TGT-GGA-GGT-C	AGA-TAC-TAA-GAG-GAG-ATG-TGT-C	500
At1g64890		TTT-CGC-ATG-TGT-TGA-GGA-AG	ATC-AAC-GGA-ATC-CCA-ACG-CAG-C	585
At5g05440		AAC-TCC-AAC-ACG-GCT-CAG-AC	TTA-TTG-CCG-GTT-GGT-ACT-TC	596
At1g06110	*SKIP16*	CTT-GTG-TTC-TGT-CTC-CGT-AAC	CTC-ATG-AAC-TTG-TAA-ACC-AG	565

Table 4.2 Primers used for cloning of *A. l. petraea* genes. Sequences are given for the primer pairs of the normalising genes At1g13320 and At4g05320, and the six genes of interest that were cloned. All primers were designed to the *A. thaliana* sequence obtained from TAIR. All genes were cloned and sequenced from *A. l petraea* to enable specific qRT-PCR primers to be designed.

Gene (AGI Code)	Gene Name	Forward primer 5' – 3'	Reverse primer 5' – 3'	Product size (bp)	Efficiency (%)
At1g13320	*PDF*	TAA-CGT-CGC-CAA-AAT-GAT-GC	GTT-CTC-CAC-AAC-CGC-TTG-GT	61	103
At4g05320	*UBQ10*	GTC-GAC-CCT-TCA-CTT-GGT-GT	GTG-TCG-GAG-CTC-TCC-ACT-TC	102	98.9
At4g05320	*UBQ10 #2*	CAC-CAT-TGA-CAA-CGT-GAA-GG	AAA-GTG-CGA-CCA-TCC-TCA-AG	105	104
At4g05320	*UBQ10 #3*	GAC-TTT-GAC-CGG-AAA-GAC-CA	CGG-AGG-GAT-ACC-TTC-CTT-GT	94	94
At1g08920		CCT-TTG-GTT-GGA-TTG-TTG-CT	TCG-GGT-ACC-ATC-GCA-TAA-AT	124	91.5
At2g46830	*CCA1*	GGT-TCA-GAT-GGA-GCA-GGA-GA	CCG-CCT-CAA-CAT-CAT-CAC-TA	88	98.9
At3g48360	*BT2*	GCA-ATG-GAG-TGC-ATA-GAG-CA	AAC-GCC-TTA-CAA-GGC-TCT-GA	119	97.5
At5g04410	*NAC2*	AAA-GGA-AAG-TGG-CGA-AGG-AT	CAG-AAG-ATT-GTG-CTG-CGT-GT	186	97.4
At2g04040	*AtDTX1*	CCC-AAA-ACT-CGA-GAC-CTC-TG	CGG-AAA-CCC-TAG-CAA-CTT-GA	146	91.6
At2g40460		AGG-CAT-TGG-GGA-TGT-GTT-TA	CCC-TCC-ACC-TTT-ACT-CGT-GA	175	92.6
At1g64890		TCC-TGG-ATG-TGA-AGG-GTC-TC	CAC-GAG-TAA-ACC-CCG-AGA-AA	137	101
At5g05440		TCA-GAC-AGT-GCC-GTA-TCG-TC	CCC-CAC-CAA-CGA-CACPTAAPAG	156	93.3
At1g06110	*SKIP16*	AAT-CCG-GAT-GTC-TCT-CAT-GC	CCA-GCT-TGT-AAG-AGC-GGG-TA	162	98.7

Table 4.3 Primer sequences used for qRT-PCR analysis of *A. l. petraea* samples. Sequences are given for primer pairs of the normalising genes At1g13320 and At4g05320, and for the nine genes of interest. Expected product size from cDNA template is given in base pairs (bp). Efficiency (%) of each primer pair is given based on results from a 10x dilution series. All primer sequences are specific to the *A. l. petraea* sequences obtained.

treated sample to the untreated sample within each population. The significance of these $\log_2$ fold changes within each population was calculated by analysis of variance (SPSS 15.0) using 'treated' and 'untreated' as the grouping variables and the $\log_2$ fold change data as the dependent variable for each population. The dataset containing $\log_2$ fold change values for each gene for all thirteen populations studied was subject to hierarchical cluster analysis using Ward's method (SPSS 15.0). This was to enable populations to be grouped based on the similarity of their transcriptomic responses to chilling across the nine genes investigated. Each gene constituted a variable, and cases were labelled by population.

4.2.8 Additional biological replicates for a subset of populations

Following analysis of data for the thirteen populations, a subset of four populations was selected; this was done using the expression data for all genes and all populations to identify those populations with the greatest differential levels of expression. Two more full biological replicates (with five plants each for treated and control conditions) were performed for these populations to test whether or not gene expression changes following chilling differed significantly between populations. Sample preparation and expression analysis was performed as described in sections 4.2.3 and 4.2.7 respectively. The populations included in the additional replicates are detailed in Table 2.1. These results were analysed using a nested analysis of variance (SPSS 15.0) with the $\log_2$ fold change data as the dependent variable and 'biological replicate' (1, 2, or 3) as a random factor nested within the fixed factor 'population'.

4.3 Results obtained: cloning, sequencing, and qRT-PCR

4.3.1 Cloning and sequencing of the housekeeping genes *PDF2* (At1g13320) and *UBQ10* (At4g05320).

Expected PCR product sizes were based on the *A. thaliana* sequence. For *PDF2*, products were successfully amplified from both Helin and Leitrim *A. l. petraea* cDNA. A PCR product of the expected size (506 bp) could not be obtained using the *UBQ10* primer pair. PCR was instead performed using the *UQB10* forward primer with the CLON PCR primer (AAG-CAG-TGG-TAA-CAA-CGC-AGA-GT) that hybridised to the 3' end of the synthesised cDNA. A specific product, slightly larger than the size expected if the original *UBQ10* reverse primer had been used (506 bp), was obtained from Helin cDNA template whilst a product of the expected size was not obtained from Leitrim cDNA template. PCR products for both genes were cloned directly into the pSTBlue-1 vector according to section 2.3.14.1. Isolated clones that were shown to harbour fragments of the expected size after restriction digestion with *Eco* R1 (section 2.3.15) for *PDF2* and *UBQ10* are shown in Figures 4.2(A + B) and 4.3 respectively.

Plasmid DNA from up to three transformed *E. coli* colonies containing recombinant plasmid was sequenced for each gene. For *PDF2*, a consensus nucleotide sequence was created for Helin and for Leitrim using DNASTAR® Lasergene 4.05 software (DNASTAR, Inc). These sequences are presented in Figure 4.4, aligned to the *A. thaliana* coding sequence for this gene. The *A. l. petraea* consensus sequences for Helin and Leitrim show approximately 98% similarity with the *A. thaliana* sequence. The *A. l. petraea UBQ10* nucleotide sequences are presented in Figure 4.5 with the *A. thaliana* coding sequence for this gene. The two Helin sequences for this gene show approximately 92% shared identity with the *A. thaliana* sequence

4.3.2 Cloning and sequencing of the genes of interest: At2g04040, At2g40460, At1g64890, At5g05440, and At1g06110.

Expected PCR product sizes were based on the *A. thaliana* sequence. Products were amplified from Leitrim *A. l. petraea* cDNA for all genes except At2g04040 for which amplification of cDNA was not successful. For At2g04040 genomic DNA was extracted from Leitrim leaf material (section 2.3.2) and used as template for the PCR reaction. The genomic DNA sequence for the *A. thaliana* At2g04040 gene contains three introns, thus the expected product size from the PCR reaction using genomic DNA

as template was greater than for cDNA template, assuming introns are conserved in *A. l. petraea*. For all genes, PCR products were analysed by agarose gel electrophoresis as detailed in section 2.3.9. Where PCR reactions produced one single product, this DNA was obtained by purifying the reaction (section 2.3.11) and cloned directly into the pSTBlue-1 vector according to section 2.3.14.1. For the gene At1g64890 two products were produced in the PCR reaction and consequently the product of the correct expected size (585 bp) was obtained by gel extraction (as 2.3.10) rather than PCR purification. Following transformation of *E. coli* competent cells and restriction digestion of plasmid DNA preparations with *Eco* R1 (2.3.17 and 2.3.15 respectively), plasmid DNA from up to three transformed *E. coli* colonies containing recombinant plasmid was sequenced for each gene of interest (section 2.3.13). Figure 4.6 illustrates successful amplification of genomic DNA (A) and *Eco* R1 restriction digests of transformed *E. coli* colonies (B) for At2g04040. Figure 4.7 illustrates *Eco* R1 restriction digests of transformed *E. coli* colonies for At2g40460, At1g64890, At5g05440, and At1g06110.

A. l. petraea nucleotide sequences for the genes of interest (At2g04040, At2g40460, At1g64890, At5g05440, and At1g06110) are presented in Figures 4.8 – 4.12 respectively. All coding region sequences had between 89 and 95% identity with the *A. thaliana* homolog sequences (Table 4.4). For At2g04040, genomic DNA was sequenced and thus intron sequences could also be compared; within these regions the similarity between *A. thaliana* and *A. l. petraea* sequences is only approximately 66%.

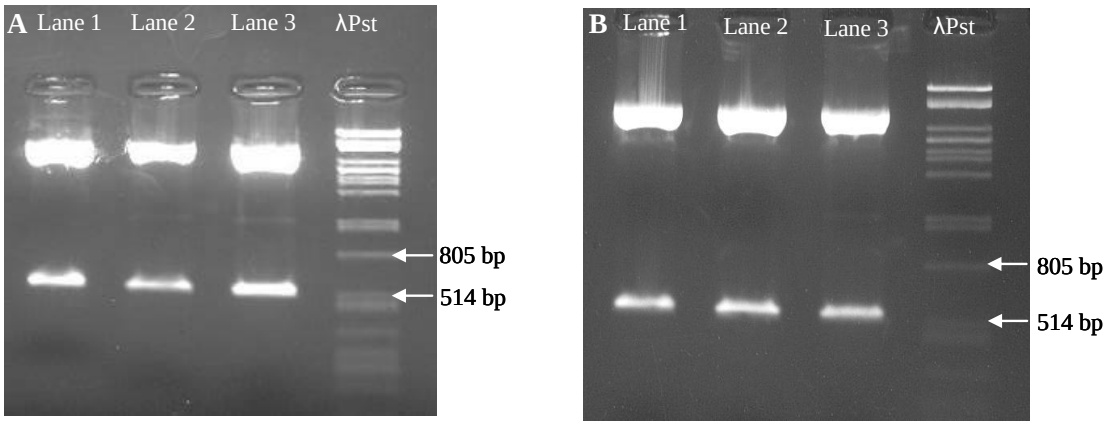


Figure 4.2. Restriction digests of putatively recombinant plasmid harbouring *PDF2* coding sequence of *A. l. petraea*. Plasmid DNA preparations from transformed *E. coli* colonies digested with restriction enzyme *Eco* R1. Expected insert size was 530 bp. A) Lanes 1-3 indicate recombinant plasmid harbouring inserts amplified from Helin cDNA. B) Lanes 1-3 indicate recombinant plasmid harbouring inserts amplified from Leitrim cDNA. Inserts from these six clones were sequenced to obtain the Helin and Leitrim *PDF2* coding sequence for *A. l. petraea*.

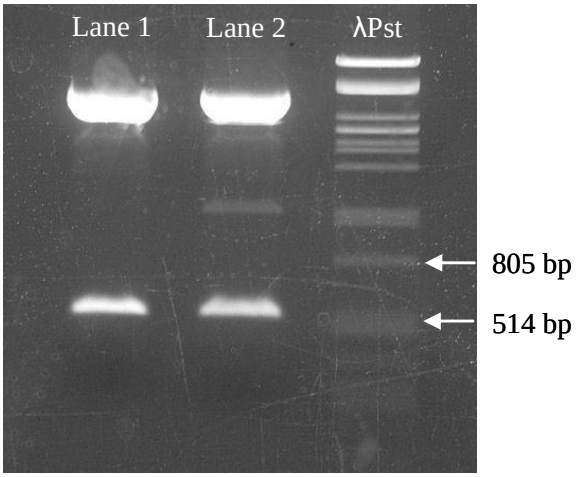


Figure 4.3. Restriction digests of putatively recombinant plasmid harbouring *UBQ10* coding sequence of *A. l. petraea*. Plasmid DNA preparations from transformed *E. coli* colonies digested with restriction enzyme *Eco* R1. Lanes 1 and 2 indicate recombinant plasmid harbouring inserts amplified from Helin cDNA. Inserts from these two clones were used to obtain the *UBQ10* coding sequence for *A. l. petraea*.

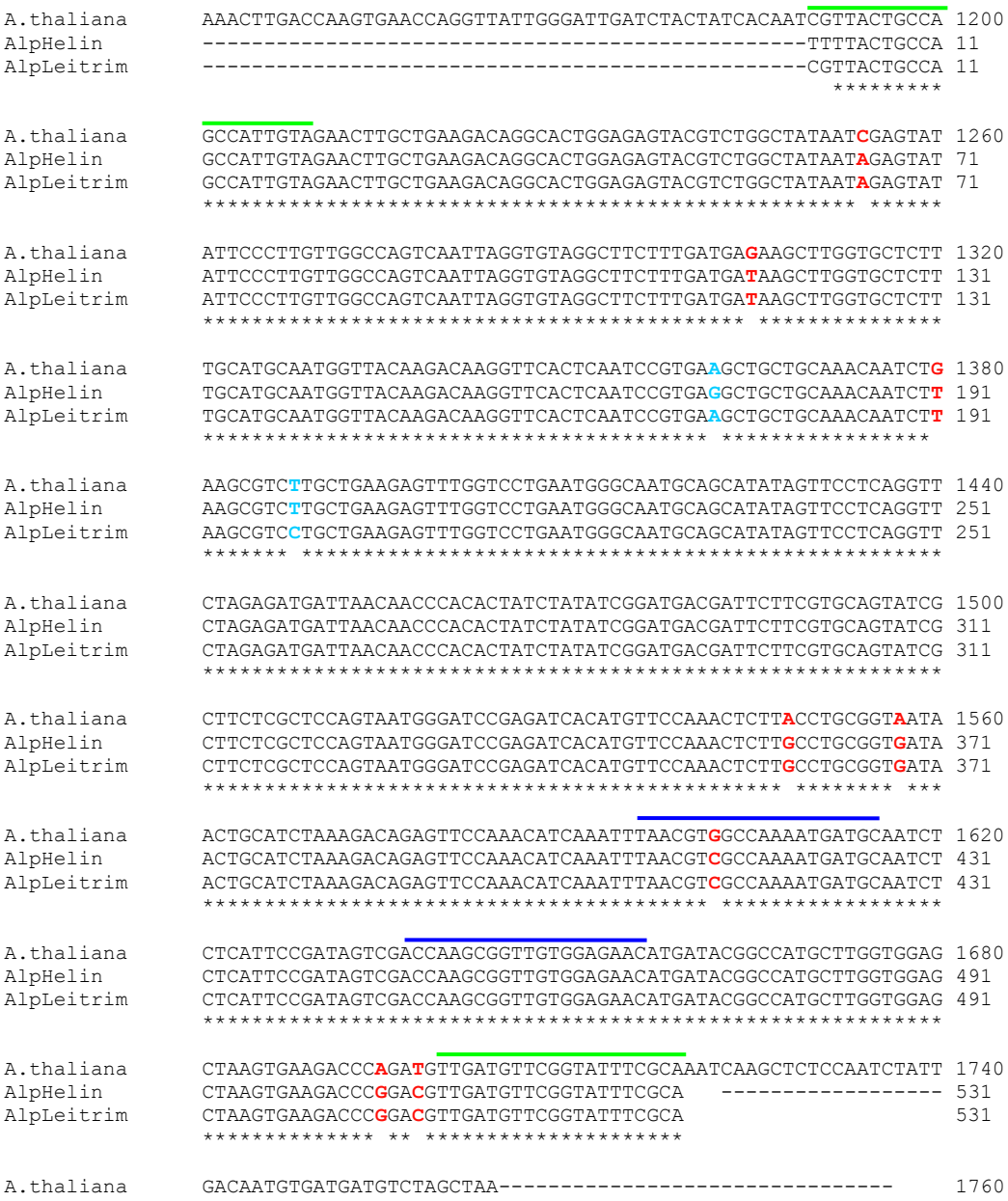


Figure 4.4 The *A. l. petraea* sequence for *PDF2* (At1g13320). The top sequence is the *A. thaliana* coding DNA sequence for the gene At1g13320. AlpHelin and AlpLeitrim refer to the consensus sequences created from sequencing of several individual recombinant plasmid DNA inserts. Asterisk (*) indicate shared identity between all three sequences. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l. petraea* sequences that are confirmed by both *A. l. petraea* traces. Bases coloured in turquoise illustrate differences between the two *A. l. petraea* populations. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments.

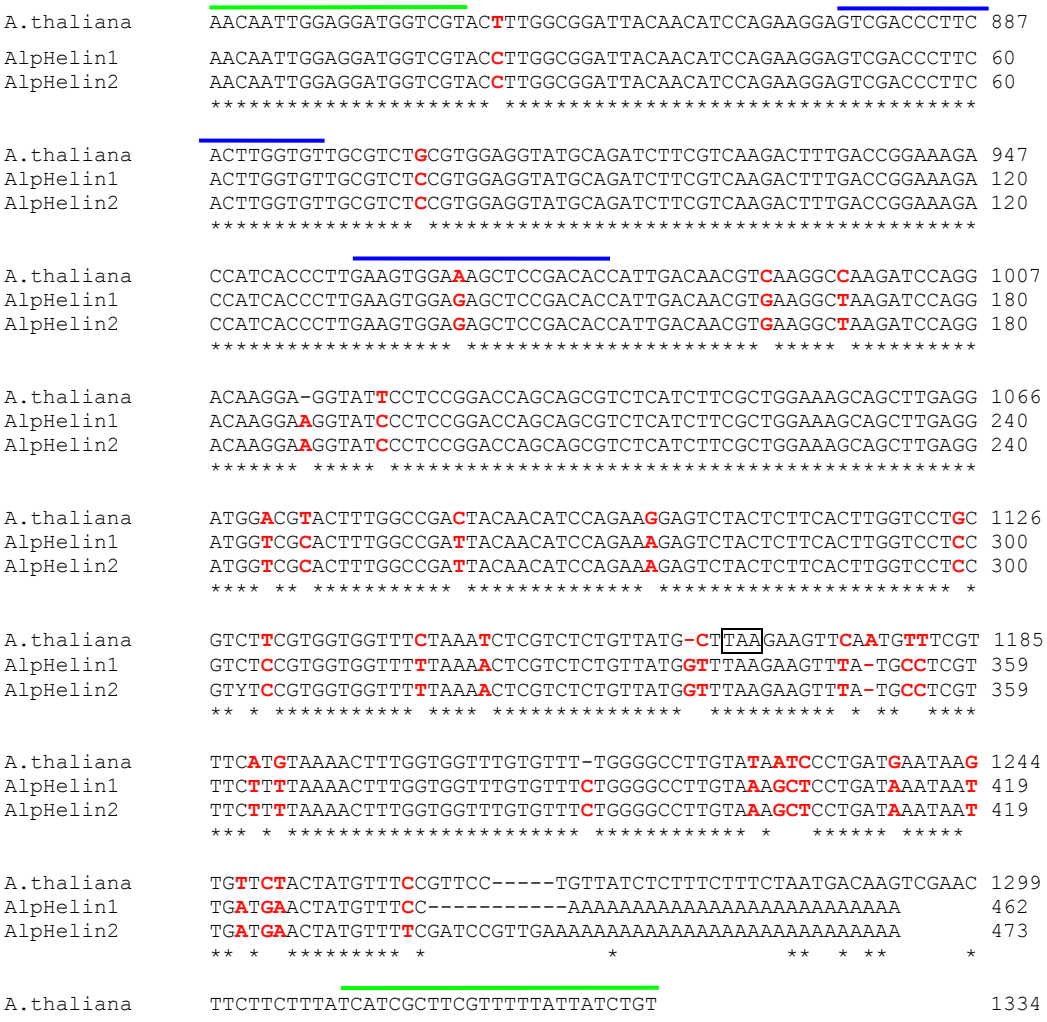


Figure 4.5 The *A. l. petraea* sequence for *UBQ10* (At4g05320). The top sequence is the *A. thaliana* coding DNA sequence for the gene At4g05320. AlpHelin1 and AlpHelin2 refer to the consensus sequences created from sequencing two recombinant plasmid DNA inserts in both directions. Asterisk (*) indicate shared identity between all three sequences. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l. petraea* sequences that are confirmed by both *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments. Boxed region indicates location of the stop codon.

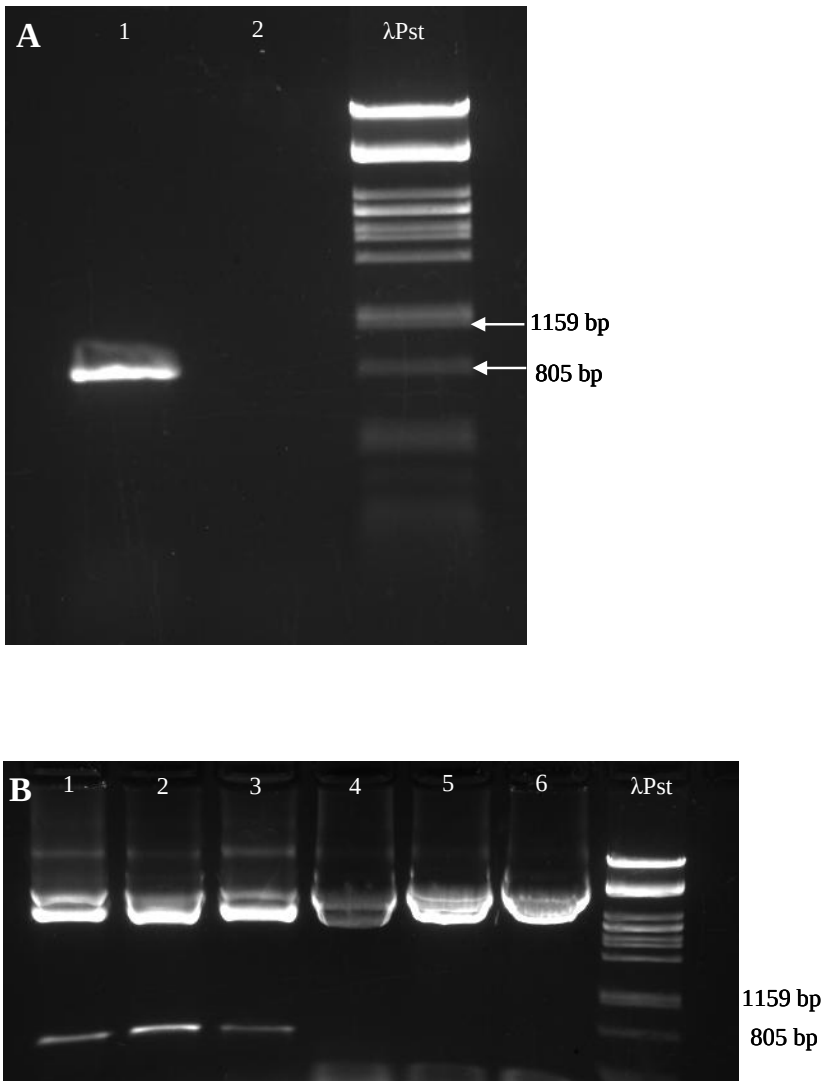


Figure 4.6. Cloning of At2g04040 coding sequence from *A. l. petraea*. A) PCR of *A. l. petraea* genomic DNA from the Leitrim population resulted in a larger product than expected with cDNA template (591 bp). B) Restriction digests with *Eco* R1 of putatively recombinant plasmid harbouring At2g04040 coding sequence of *A. l. petraea*. Lanes 1, 2, and 3 indicate recombinant plasmid harbouring inserts amplified from Leitrim genomic DNA. Lanes 4, 5, and 6 indicate plasmid without desired insert. Inserts from the clones in Lanes 1-3 were used to obtain the At2g04040 coding sequence for *A. l. petraea*.

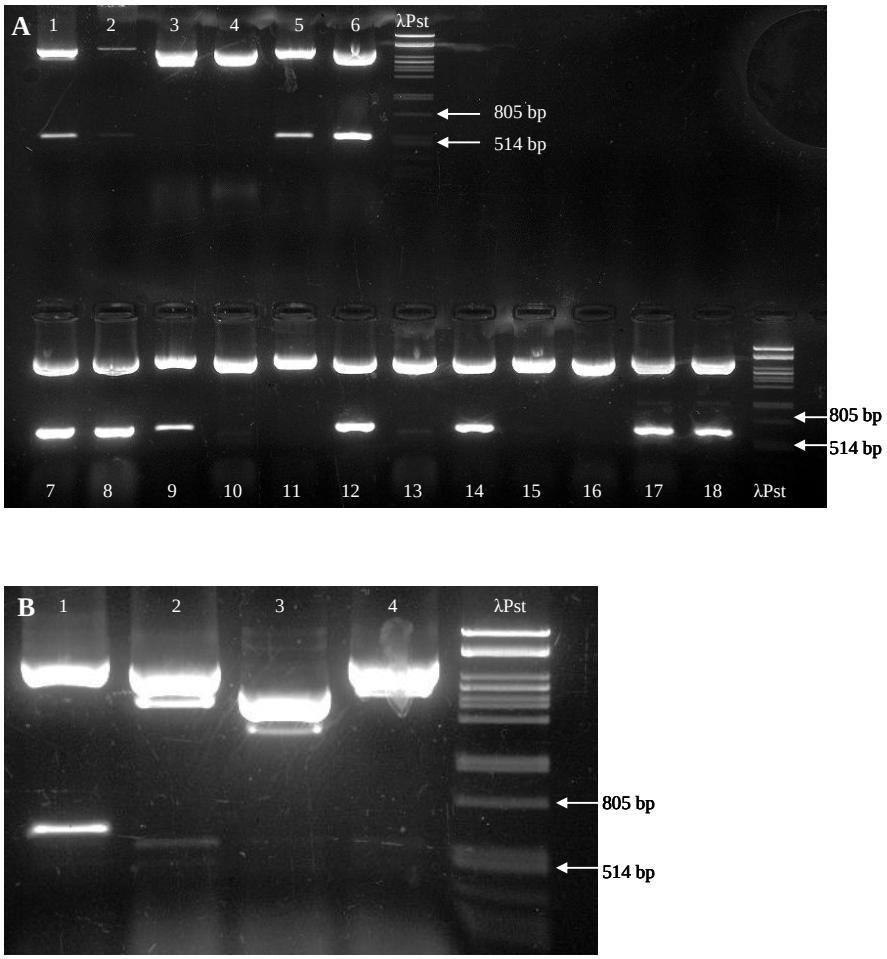


Figure 4.7 Restriction digests of putatively recombinant plasmid harbouring coding sequence of four genes of interest from *A. l. petraea*. Plasmid DNA preparations from transformed *E. coli* colonies digested with restriction enzyme *Eco* R1. A) Lanes 1-6: At2g40460. Lanes 1, 2, 5, and 6 indicate recombinant plasmid harbouring inserts amplified from Leitrim cDNA. Plasmid in lanes 3 and 4 does not contain desired insert. Inserts from the clones in lanes 2, 5, and 6 were used to obtain the *A. l. petraea* sequence for this gene. Lanes 7-12: At4g05440. Lanes 7, 8, 9, and 12 indicate recombinant plasmid harbouring inserts amplified from Leitrim cDNA. Plasmids in lanes 10 and 11 do not contain desired insert. Inserts from the clones in lanes 7, 8, and 12 were used to obtain the *A. l. petraea* sequence for this gene. Lanes 13-18: At1g06110. Lanes 14, 17, and 18 indicate recombinant plasmid harbouring inserts amplified from Leitrim cDNA that were used to obtain the *A. l. petraea* sequence for this gene. Plasmid in lanes 13, 15, and 16 does not contain the desired insert. B) Lanes 1-4: At1g64890. Lane 1 indicates recombinant plasmid harbouring an insert amplified from Leitrim cDNA that was used to obtain the *A. l. petraea* sequence for this gene. Plasmids in lanes 2, 3, and 4 do not contain the desired insert.

```
A.thaliana      ------------------------------------CATACTCTGCTCAGGTCTTCTCC 1020
Alp1            ------------------------------------CATACTCTGCTCAGGTCTTCTAC 23
Alp2            ------------------------------------CATACTCTGCTCAGGTCTTCTAC 23
Alp3            ------------------------------------CATACTCTGCTCAGGTCTTCTAC 23
                                                     ******************** *

A.thaliana      CTAACCCGAAACTCGAGACCTCTGTTCTTTCAATTTGgtaagtaccatagaaggacttct 1080
Alp1            CGAACCCAAAACTCGAGACCTCTGTTCTATCTATTTGgtaagtatcatatca---tttct 80
Alp2            CGAACCCAAAACTCGAGACCTCTGTTCTATCTATTTGgtaagtatcatatca---tttct 80
Alp3            CGAACCCAAAACTCGAGACCTCTGTTCTATCTATTTGgtaagtatcatatca---tttct 80
                * ***** ******************** ** ************ ****  *    ****

A.thaliana      ttttatcttttgtggtctctttactaaaaactaatgaatcctgtttccttcttttcagTC 1140
Alp1            ttttgccctatgtgttctcttgaccaaagtgta-taaatcctatttctttcatgtcagTC 139
Alp2            ttttaccctatgtgttctcttgaccaaagtgta-taaatcctatttctttcatgtcagTC 139
Alp3            ttttaccctatgtgttctcttgaccaaagtgta-taaatcctatttctttcatgtcagTC 139
                ****  * * **** ****** ** ***   ** * ****** **** *** * ******

A.thaliana      TTACAATAGAAACTTTGCACTATGTGATTTCAGCTGGAGTCGCCGCGGCTGTGAGgtatg 1200
Alp1            TCACAATAGGAAATTTGCACTATGTGATTCCCTCTGGAGTCGCGGCGGCTGTGAGgtat- 198
Alp2            TCACAATAGGAAATTTGCACTATGTGATTCCCTCTGGAGTCGCGGCGGCTGTGAGgtat- 198
Alp3            TCACAATAGGAAATTTGCACTATGTGATTCCCTCTGGAGTCGCGGCGGCTGTGAGgtat- 198
                * ******* ** ***************** *  ********** ***************

A.thaliana      aatatgatctcagttttatcattcagaacatcttcatcttattatagacaattctttcct 1260
Alp1            -acatgatctcaattct-caaaccagaacatgttca-----------acaattcctttcc 245
Alp2            -acatgatctcaattct-caaaccagaacatgttca-----------acaattcctttcc 245
Alp3            -acatgatctcaattct-caaaccagaacatgttca-----------acaattcctttcc 245
                 * ********* ** *   *  ******** ****           ******* ** *

A.thaliana      aaacatttagtgattggtttcttgcagCACACGTGTATCAAACAATTTGGGAGCTGGGAA 1320
Alp1            caccagtcatt---tggtttcttgcagCACACGTGTGTCAAACAAATTGGGAGCTGGCAA 302
Alp2            caccagtcatt---tggtttcttgcagCACACGTGTGTCAAACAAATTGGGAGCTGGCAA 302
Alp3            caccagtcatt---tggtttcttgcagCACACGTGTGTCAAACAAATTGGGAGCTGGCAA 302
                 * ** * * *   ********************** ******** *********** **

A.thaliana      TCCTCAAGTTGCTAGGGTTTCAGTATTAGCAGGGCTTTGTCTCTGGATAGTAGAGTCAGC 1380
Alp1            TCCTCAAGTTGCTAGGGTTTCCGTAATGGCAGGGCTTTGTCTCTGGATAGTAGAATCAAC 362
Alp2            TCCTCAAGTTGCTAGGGTTTCCGTAATGGCAGGGCTTTGTCTCTGGATAGTAGAATCAAC 362
Alp3            TCCTCAAGTTGCTAGGGTTTCCGTAATGGCAGGGCTTTGTCTCTGGATAGTAGAATCAAC 362
                ********************* *** * ************************** *** *

A.thaliana      TTTCTTTAGCATACTTCTATTCACCTGCAGGAACATCATAGGATACGCATTCAGCAACAG 1440
Alp1            TTTCTTTAGCACACTTCTCTTCACCTGCCGAAACATCGTAGGCTACGCATTCAGCAACAG 422
Alp2            TTTCTTTAGCACACTTCTCTTCACCTGCCGAAACATCGTAGGCTACGCATTCAGCAACAG 422
Alp3            TTTCTTTAGCACACTTCTCTTCACCTGCCGAAACATCGTAGGCTACGCATTCAGCAACAG 422
                *********** ****** ********* * ****** **** *****************

A.thaliana      CAAAGAAGTTTTGGACTATGTAGCTGACCTAACTCCTTTGCTCTGCCTCTCCTTTATCCT 1500
Alp1            CAAAGAAGTTGTCGACTATGTGGCTGACATATCATATTTGCTCTGTCTCTCCTTCATCCT 482
Alp2            CAAAGAAGTTGTTGACTATGTGGCTGACATATCATATTTGCTCTGTCTCTCCTTCATCCT 482
Alp3            CAAAGAAGTTGTTGACTATGTGGCTGACATATCATATTTGCTCTGTCTCTCCTTCATCCT 482
                ********** * ******** ****** ** *   ********* ******** *****

A.thaliana      CGACGGCTTCACAGCAGTTCTAAATGGAGTTGCTAGGGGAAGTGGTTGGCAACACATTGG 1560
Alp1            AGACGGATTTACCGCAGTTCTAAATGGGGTTGCTAGGGGATGTGGGTGGCAACACATTGG 542
Alp2            AGACGGATTTACCGCAGTTCTAAATGGGGTTGCTAGGGGATGTGGGTGGCAACACATTGG 542
Alp3            AGACGGATTTACCGCAGTTCTAAATGGGGTTGCTAGGGGATGTGGGTGGCAACACATTGG 542
                 ***** ** ** ************** ************ **** **************

A.thaliana      AGCTTGGAACAATACTGTTTCTTATTATCTTGTAGGAGCTCCTGTTGGAATTTACTTAGC 1620
Alp1            AGCTTTGATCAGCGTTGTGGCTTATTATCTCGTAGGAGCTCCTGTTGGAGTTTACTTAGC 602
Alp2            AGCTTTGATCAGCGTTGTGGCTTATTATCTCGTAGGAGCTCCTGTTGGAGTTTACTTAGC 602
Alp3            AGCTTTGATCAGCGTTGTGGCTTATTATCTCGTAGGAGCTCCTGTTGGAGTTTACTTAGC 602
                ***** ** **    ***  ********** ****************** **********

A.thaliana      TTTCAGTCGTGAGTTGAACGGAAAAGGACTGTGGTGCGGTGTTGTGGTTGGATCTACTGT 1680
Alp1            CTTCAGTCGTGAGTGGAACGGAAAAGGACTGTGGTGTGGTGTAGTGGTTGGATCCGCTGT 662
Alp2            CTTCAGTCGTGAGTGGAACGGAAAAGGACTGTGGTGTGGTGTAGTGGTTGGATCCGCTGT 662
Alp3            CTTCAGTCGTGAGTGGAACGGAAAAGGACTGTGGTGTGGTGTAGTGGTTGGATCCGCTGT 662
                 ************* ******************** ***** ***********  ****

A.thaliana      GCAAGCCACTATACTGGCTATTGTCACAGCTTCCATAAATTGGAAGGAACAGgtaaata 1739
Alp1            GCAAGCCATTATATTGACTTTCGTCACAGCGTCCATAAATTGGAAGGAACAGgttag-- 720
Alp2            GCAAGCCATTATATTGGCTTTCGTCACAGCGTCCATAAATTGGAAGGAACAGgttag-- 720
Alp3            GCAAGCCATTATATTGGCTTTCGTCACAGCGTCCATAAATTGGAAGGAACAGgttag-- 720
                ******** **** ** ** * ******** ********************** **
```

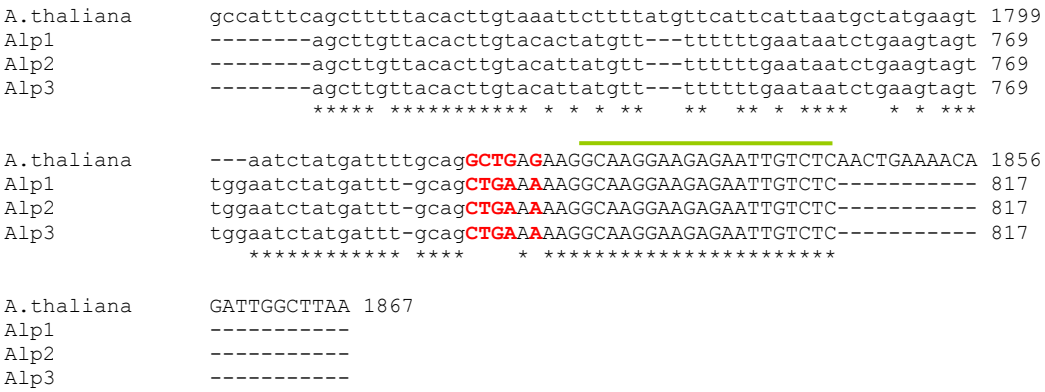


Figure 4.8 The *A. l. petraea* sequence for At2g04040. The top sequence is the *A. thaliana* genomic DNA sequence for the gene At2g04040. Uppercase letters refer to coding sequence and lower case letters refer to introns. Alp1, Alp2, and Alp3 refer individual clones harbouring *A. l. petraea* genomic DNA. Asterisk (*) indicate shared identity between all four sequences. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l. petraea* sequences that are confirmed by all three *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments. qRT-PCR primers were designed to span a region containing introns to enable any genomic DNA contamination of samples to be detected.

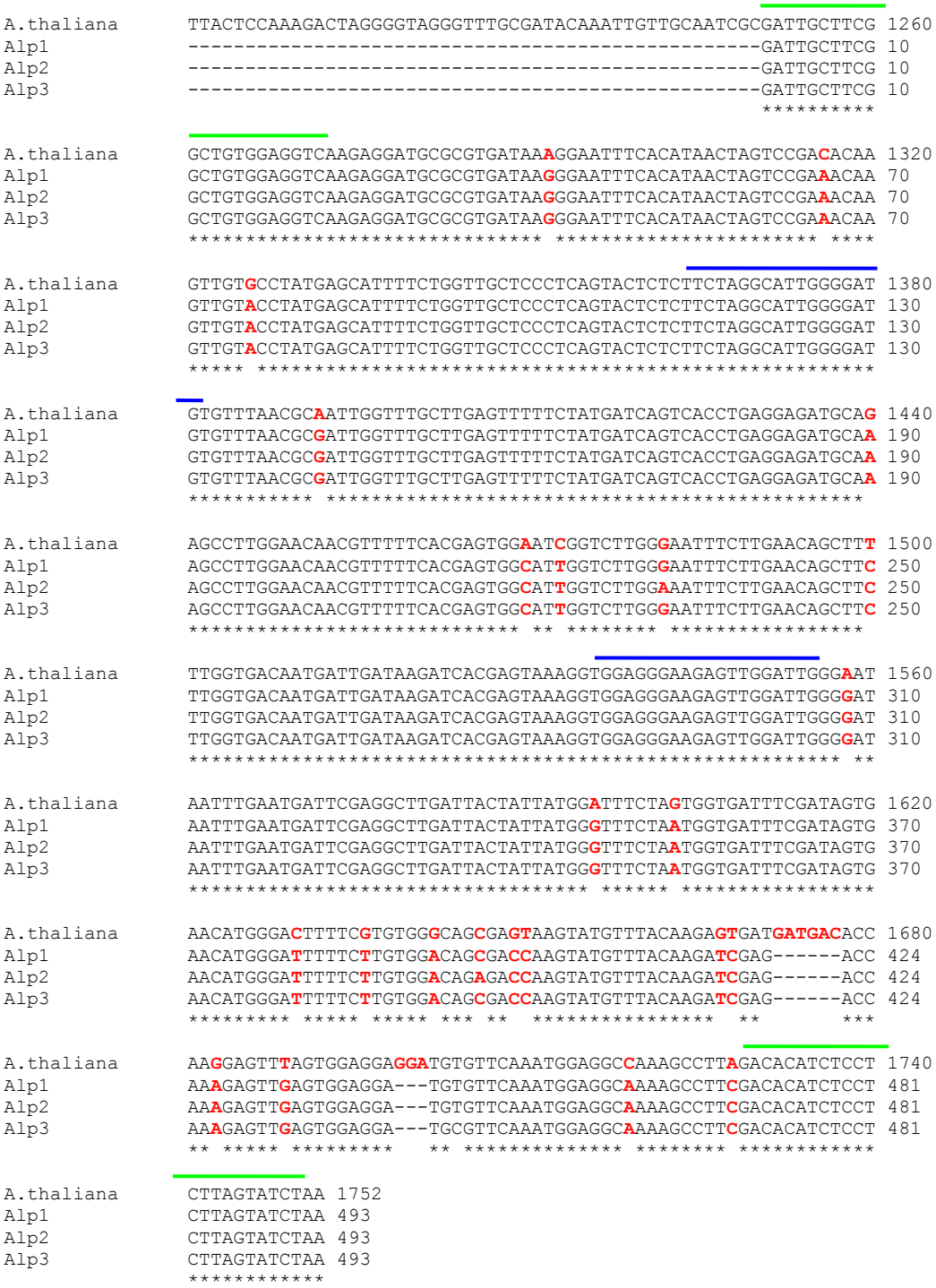


Figure 4.9 The *A. l. petraea* sequence for At2g40460. The top sequence is the *A. thaliana* coding DNA sequence for the gene At2g40460. Alp1, Alp2, and Alp3 refer to the sequences of individual clones harbouring *A. l.* petraea cDNA. Asterisk (*) indicate shared identity between all four sequences. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l. petraea* sequences that are confirmed by all three *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments.

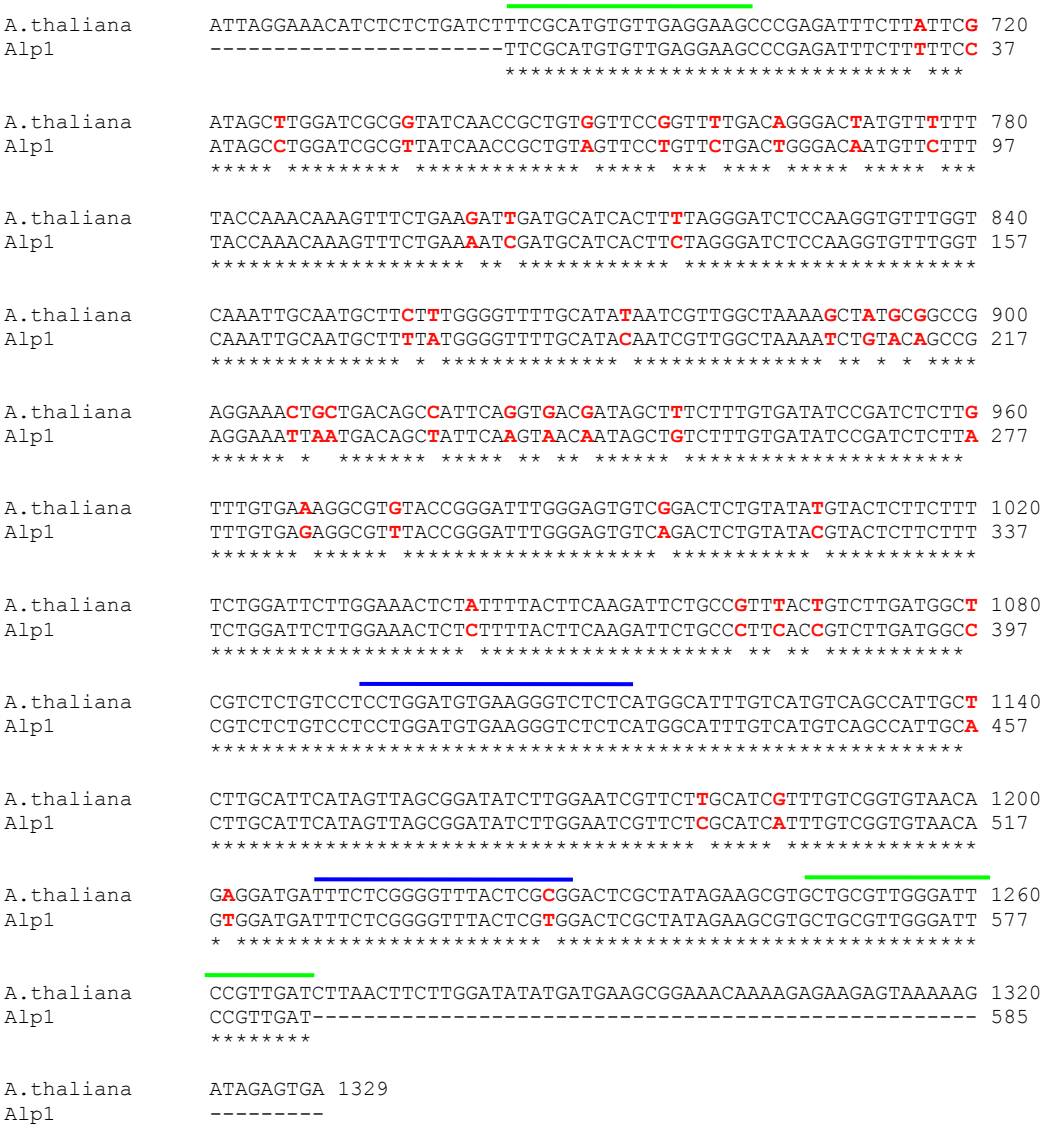


Figure 4.10 The *A. l. petraea* sequence for At1g64890. The top sequence is the *A. thaliana* coding DNA sequence for the gene At1g64890. Alp1 refers to the sequence of an individual clone harbouring *A. l. petraea* cDNA. Asterisk (*) indicate shared identity between both sequences. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l petraea* sequences. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments.

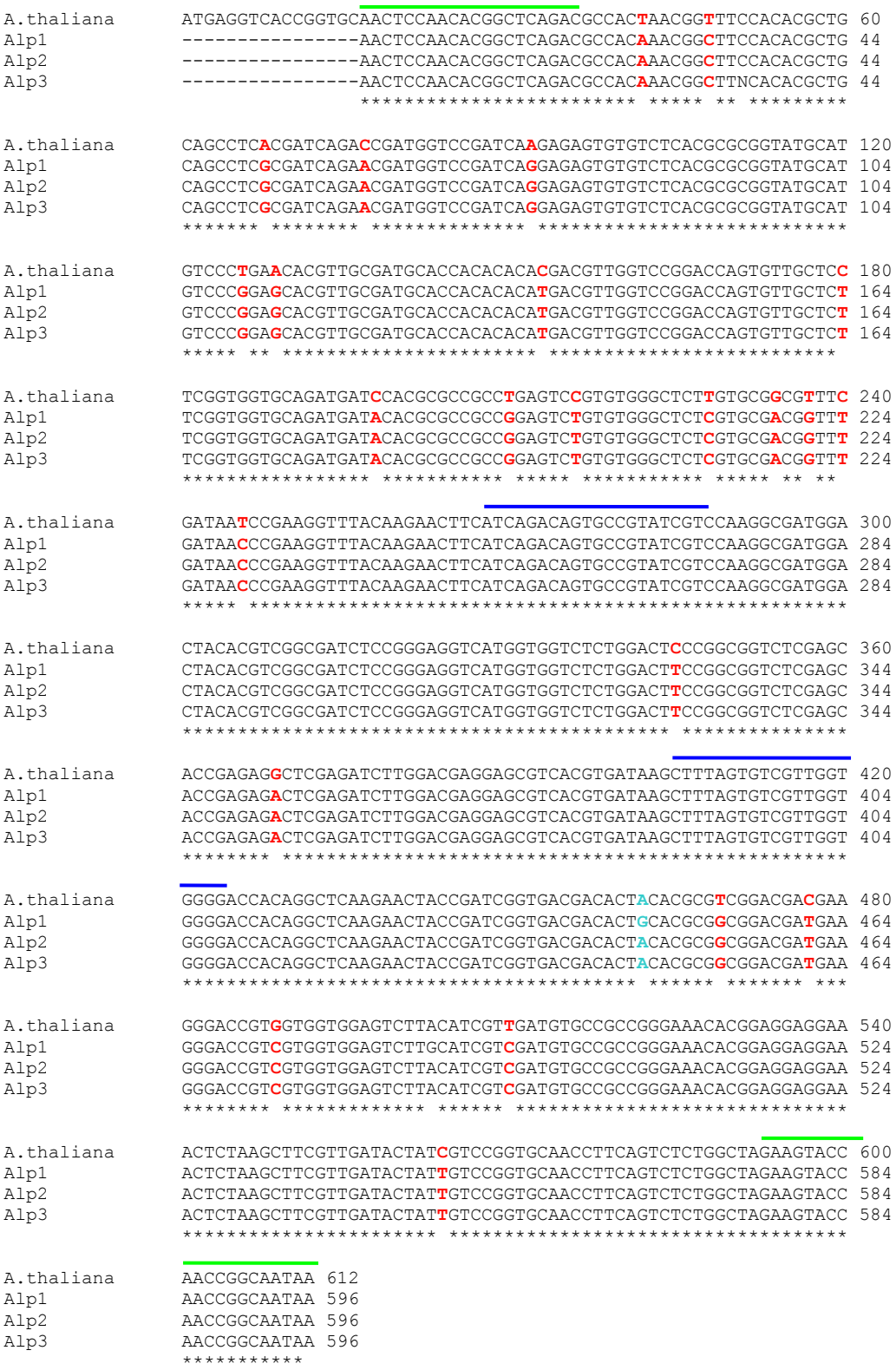


Figure 4.11 The *A. l. petraea* sequence for At5g05440. The top sequence is the *A. thaliana* coding DNA sequence for the gene At5g05440. Alp1, Alp2, and Alp3 refer to sequences of individual clones harbouring *A. l. petraea* cDNA. Asterisk (*) indicate shared identity between sequences. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l. petraea* sequences. Bases coloured in turquoise illustrate sequence differences between *A. l. petraea* clones. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments.

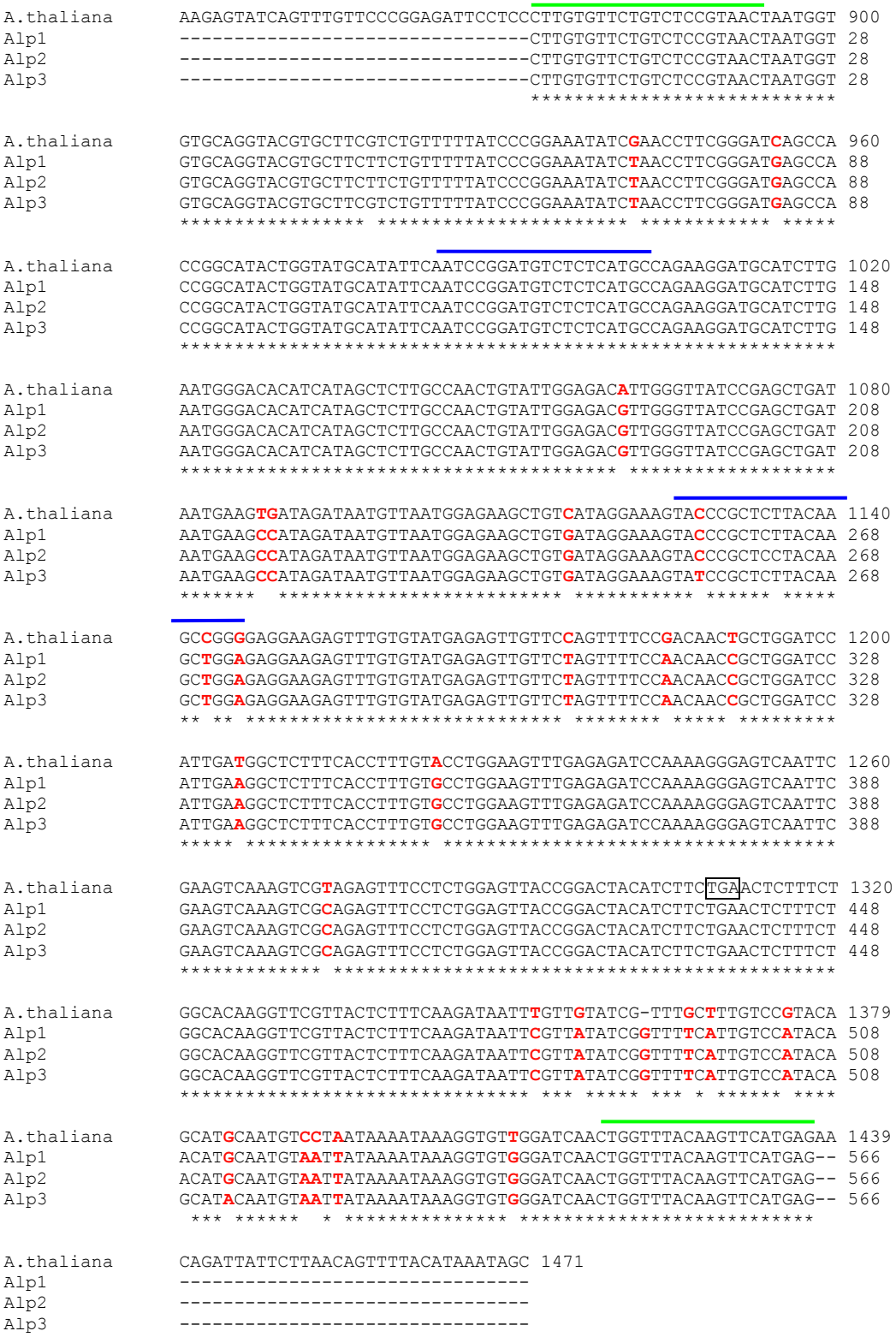


Figure 4.12 The *A. l. petraea* sequence for At1g06110. The top sequence is the *A. thaliana* coding DNA sequence for the gene At1g06110. Alp1, Alp2, and Alp3 refer to sequences of individual clones harbouring *A. l. petraea* cDNA. Asterisk (*) indicate agreement between both traces. Bases coloured in red illustrate differences between the *A. thaliana* and *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Solid blue lines indicate the location of primers used in qRT-PCR experiments. Boxed region indicates location of the stop codon.

Gene (AGI code)	Expected amplicon length (from *A. thaliana* cDNA sequence) (bp)	**Number of recombinant plasmids sequenced and their source population**	Percentage (%) identity between *A. thaliana* and *A. l. petraea*
At1g13320	530	3 Helin / 3 Leitrim	98
At4g05320	506	2 Helin	92
At2g04040	591 (848 from genomic DNA)	3 Leitrim	89 (66 within introns)
At2g40460	500	3 Leitrim	94
At1g64890	585	1 Leitrim	93
At5g05440	596	3 Leitrim	95
At1g06110	565	3 Leitrim	95

Table 4.4 Percentage identity between *A. thaliana* and *A. l. petraea* sequences. Following PCR amplification of *A. l. petraea* DNA using primers designed to the *A. thaliana* sequence, and subsequent cloning of products, individual clones harbouring *A. l. petraea* DNA were sequenced. These sequences were aligned with the *A. thaliana* coding sequence for each gene to obtain percentage identity values. Genomic DNA was sequenced for the gene At2g04040 thus a percentage identity value for the introns was also obtained.

4.3.3 Quality of RNA and cDNA synthesis

All RNA extracted from individual leaves was ≥100 ng μl^{-1} when quantified (section 2.3.5). Following the pooling of samples, as described in section 4.2.3, all samples were found to be of a suitable quality when checked on the Agilent 2100 Bioanalyser (Agilent Technologies) (section 2.3.6). The output from the Agilent 2100 Bioanalyser run is shown in Figure 4.13 for a subset of the samples. All samples have clearly defined 18S and 28S peaks with little noise between these peaks. The additional peaks, before the 18S peak, are due to the presence of chloroplasts in the leaf tissue from which RNA was extracted.

cDNA was synthesised from each pooled sample (section 2.3.7) and PCR with 18S primers was performed to confirm presence and quality of cDNA (forward primer: TCG-GCT-TGC-TCT-GAT-GAT-TC; reverse primer: CCG-ACC-AAT-GCA-CAC-CAA-AG). PCR products were analysed by gel electrophoresis as detailed in section 2.3.9. A specific product of the expected size (452 bp) was observed for all samples tested as shown in Figure 4.14.

4.3.4 Efficiency and specificity of qRT-PCR primers

4.3.4.1 Housekeeping genes PDF2 (At1g13320) and UBQ10 (At4g05320)

The qRT-PCR primers for the housekeeping gene *PDF2* (At1g13320) were based on primers designed to the *A. thaliana* sequence as described by Czechowski *et al.* (2005). Following sequencing of *A. l. petraea* clones (section 4.2.5 and Figure 4.4), primers specific to the *A. l. petraea* sequence were ordered. One change was made to the primers described by Czechowski *et al*. (2005); nucleotide seven in the forward primer (5'-3') was changed from a 'G' as in the *A. thaliana* sequence to a 'C' as in the *A. l. petraea* sequence. This primer pair was found to have an efficiency of 103% (Figure 4.15A), and dissociation curve analysis confirmed specificity of the primer pair (Figure 4.16A).

The qRT-PCR primers for the housekeeping gene *UBQ10* (At4g05320) were designed to the *A. l. petraea* sequences that were obtained for this gene (section 4.2.5 and Figure 4.5). The initial set of primers (*UBQ10*#1) had an efficiency of 98.9% and produced one single product when analysed using a dissociation curve where cDNA from one population was used as template (data not shown). Analysis of the dissociation curve

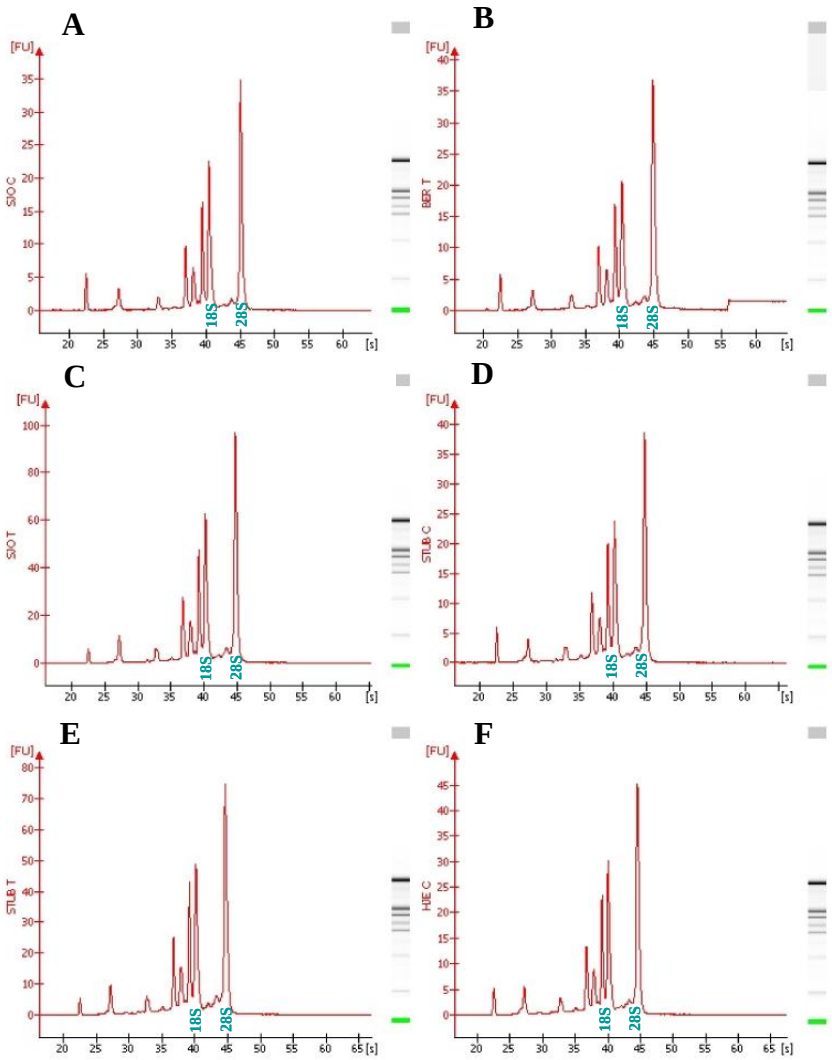


Figure 4.13 Representative Agilent Bioanalyser electropherograms for pooled RNA samples. All RNA samples extracted from individual leaves were pooled as described in section 4.2.3. These pooled samples were analysed on the Agilent 2100 Bioanalyser (Agilent Technologies). Output for a subset of the samples is shown here. A) Sweden2 Untreated; B) Sweden1 Untreated; C) Sweden2 Treated; D) Sweden3 Untreated; E) Sweden3 Treated; and F) Norway1 Untreated. All samples had clearly defined 18S and 28S peaks with little noise between these peaks. The additional peaks, before the 18S peak, are due to the presence of chloroplasts in the leaf tissue from which RNA was extracted.

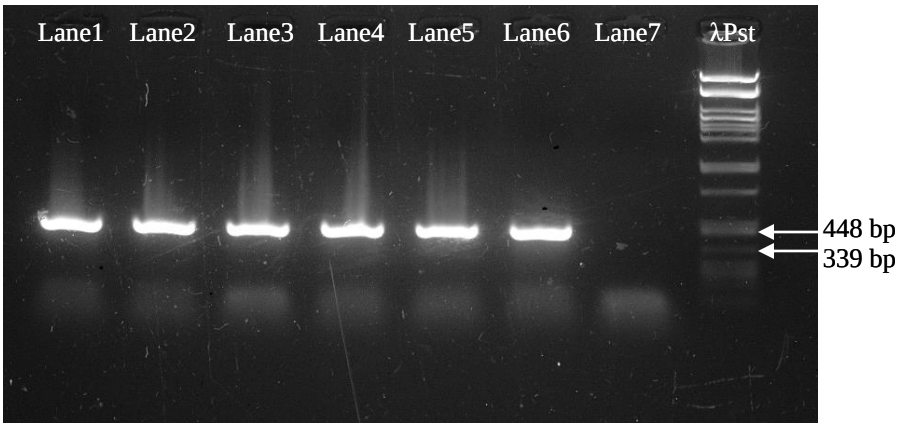


Figure 4.14 PCR of cDNA synthesised from pooled RNA samples using 18S primers. RNA extracted from individual plants was pooled prior to cDNA synthesis (section 4.2.3). PCR was performed on a selection of cDNA (from untreated and treated samples) to confirm successful reverse transcription prior to performing qRT-PCR experiments. A product of 452 bp indicates presence and quality of cDNA. Lane 1:Iceland1 Untreated; Lane 2: Iceland1 Treated; Lane3: Sweden1 Untreated; Lane 4: Norway1 Untreated; Lane 5: Norway1 Treated; Lane 6: Positive control, Leitrim cDNA; Lane 7: Negative control, Water.

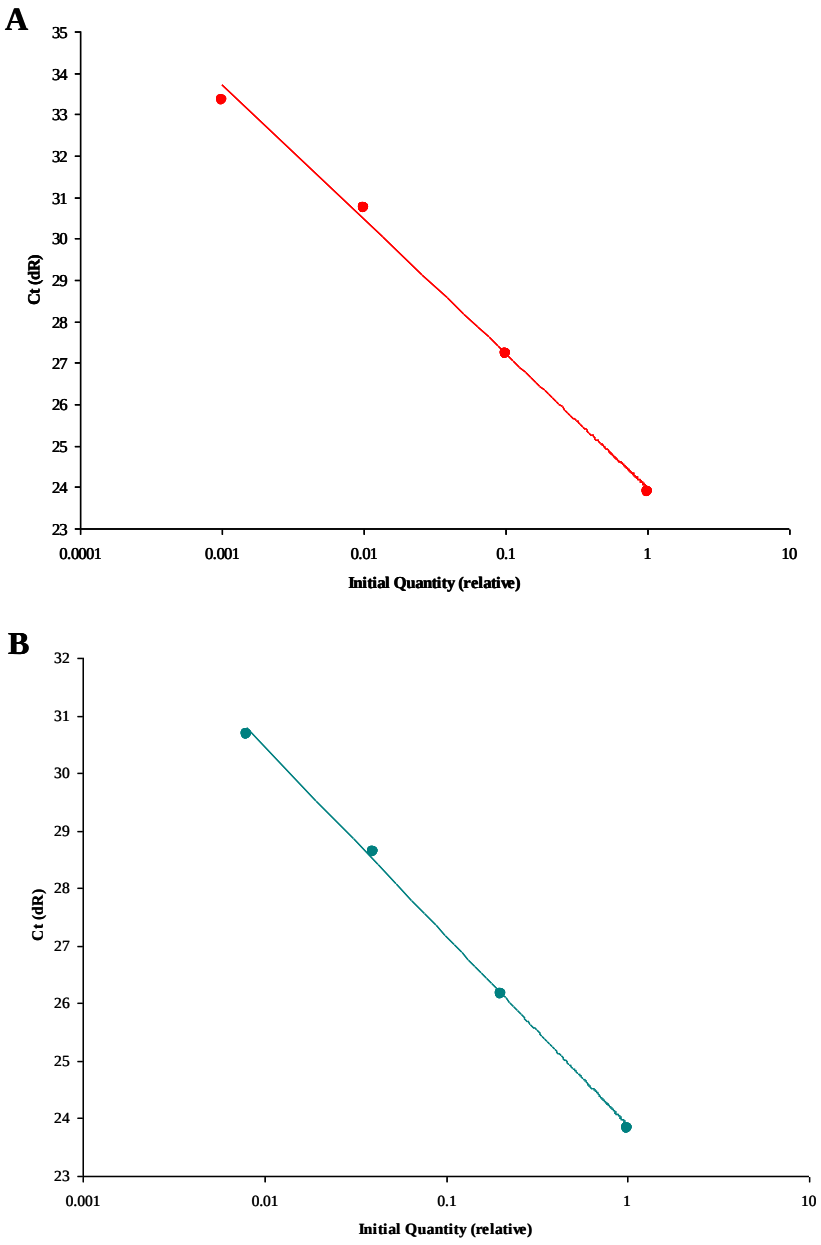


Figure 4.15 Efficiency of *PDF2* and At1g64890 qRT-PCR primers. qRT-PCR primers were designed to the *A. l. petraea* sequences obtained through cloning (section 4.2.5) for both *PDF2* and At1g64890. A standard curve was produced for each primer pair using a 10x dilution series. *PDF2* primers (A) had an efficiency of 103% (RSq = 0.996). At1g64890 primers (B) had an efficiency of 101 % (RSq = 0.999).

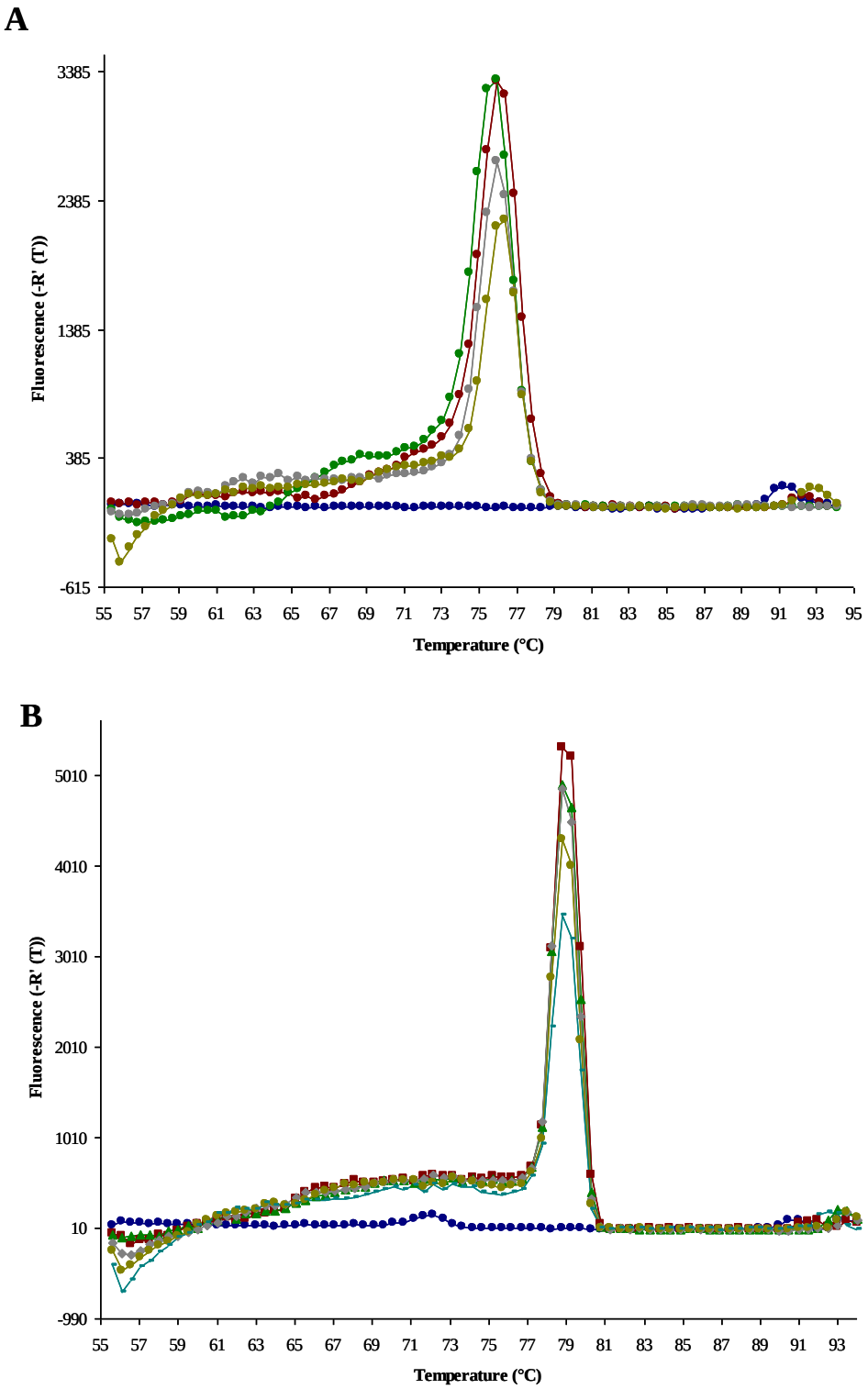


Figure 4.16 Dissociation curves for *PDF2* and At1g64890 qRT-PCR primers. Dissociation curve analysis illustrates the specificity of qRT-PCR primers. The temperature at which the product melts is indicative of its size. One single product is produced for both the *PDF2* (A) and At1g64890 (B) *A. l. petraea* specific qRT-PCR primers.

for these primers revealed that products of different sizes were amplified depending on the source population of the cDNA; Leitrim samples produced a double peak whilst Helin samples produced one single peak. Consequently, two additional pairs of primers (*UBQ10*#2 and *UBQ10*#3) were designed to the *A. l. petraea* sequence. These primer pairs had efficiencies of 104% and 94% respectively. Both primer pairs produced non-specific products for Leitrim samples, but *UBQ10*#3 primers did this to a greater degree (data not shown). All primer sequences are detailed in Table 4.3. Due to the specificity problems of the three separate primer pairs for the *UBQ10* gene, it was decided to use only *PDF2* as the normalising gene in this work. All reported expression levels are subsequent to normalisation with *PDF2*.

4.3.4.2 Genes of Interest: At1g08920, At2g46830, At3g48360, At5g04410, At2g04040, At2g40460, At1g64890, At5g05440, and At1g06110.

The qRT-PCR primers for genes At1g08920, At2g46830, and At3g48360 were designed to the *A. thaliana* coding sequence available from TAIR (http://www.arabidopsis.org). The qRT-PCR primers for the gene *NAC2* (At5g04410) were designed to *A. l. petraea* sequence previously obtained for this gene (Lilley, C.J, *pers. comm*.), and primers for genes At2g04040, At2g40460, At1g64890, At5g05440, and At1g06110 were designed to the *A. l. petraea* sequences obtained in this work (Section 4.2.5, Figures 4.8 – 4.12). The standard and dissociation curves for At1g64890 are shown in Figures 4.15 and 4.16, respectively. All primer pair efficiencies were > 90% and < 110% and are detailed in Table 4.3. All dissociation curves confirmed the amplification of a single product and the absence of primer dimers (data not shown).

4.3.5 Measurement of gene expression through qRT-PCR

4.3.5.1 Relative quantity of gene expression in untreated samples

In Figures 4.17-4.21 inclusive, the level of gene expression in the untreated samples is shown for each population relative to the level of gene expression recorded for the population Iceland1. Any differences in the expression levels of the genes of interest prior to any cold temperature exposure can consequently be assessed.

For gene At1g08920 (Figure 4.17A), the three Swedish populations exhibit the highest levels of expression in untreated samples whilst Scotland1, Wales1, and Norway3 – Norway6 have the lowest levels of expression. Expression in Sweden3 is almost six times that of in Iceland1 prior to any cold temperature exposure. Expression of the gene

At2g46830 varies considerably across the different populations in the untreated samples (Figure 4.17B). Iceland2, Sweden2 and Norway1 all exhibit at least 3 times the level of expression measured for Iceland1, whilst other populations including Sweden3 and Norway2 exhibit expression levels that are very comparable to that of Iceland1. Norway3 – Norway5 inclusive all exhibit less than half of the expression levels recorded for Iceland1 prior to cold temperature exposure. The cold down-regulated gene At3g48360 has higher levels of expression in the three Swedish populations, Norway1 and Norway2, Scotland1 and Wales1 than in the remaining populations prior to cold temperature exposure. These remaining populations all exhibit relative expression levels very comparable to that of Iceland1 (Figure 4.18).

Results for the two genes that were up-regulated to a significantly greater extent in Helin than in Leitrim are presented in Figure 4.19. For both genes, the levels of expression in these untreated samples are consistent across the different populations, and therefore all populations have expression levels comparable to those of Iceland1. The one exception to this is for the population Norway6; this has almost three times the level of expression of At2g40460 than Iceland1 (Figure 4.19A). Figure 4.20 shows the relative quantity of gene expression for the two genes that were up-regulated to a significantly greater extent in Leitrim than in Helin. At5g05440 is expressed to similar extents across all the populations in the untreated samples, with Iceland2 and Wales1 exhibiting the highest expression. For At1g64890, several populations have higher levels of expression than Iceland1, including Sweden2 and Wales1 which each have over double the level of expression.

The expression of At1g06110 in untreated samples varies considerably across the populations (Figure 4.21A). A number of populations have higher relative levels of expression than Iceland1, notably the three Swedish populations, Norway6 and Wales1. Of these, Norway6 has the highest expression; a value over nine times that of Iceland1. Many populations have higher expression of the gene At2g04040 relative to Iceland1, with many exhibiting greater than double the value for Iceland1 (Figure 4.21B). Expression values for this gene in untreated plants are the highest for Wales1, almost five times that of Iceland1.

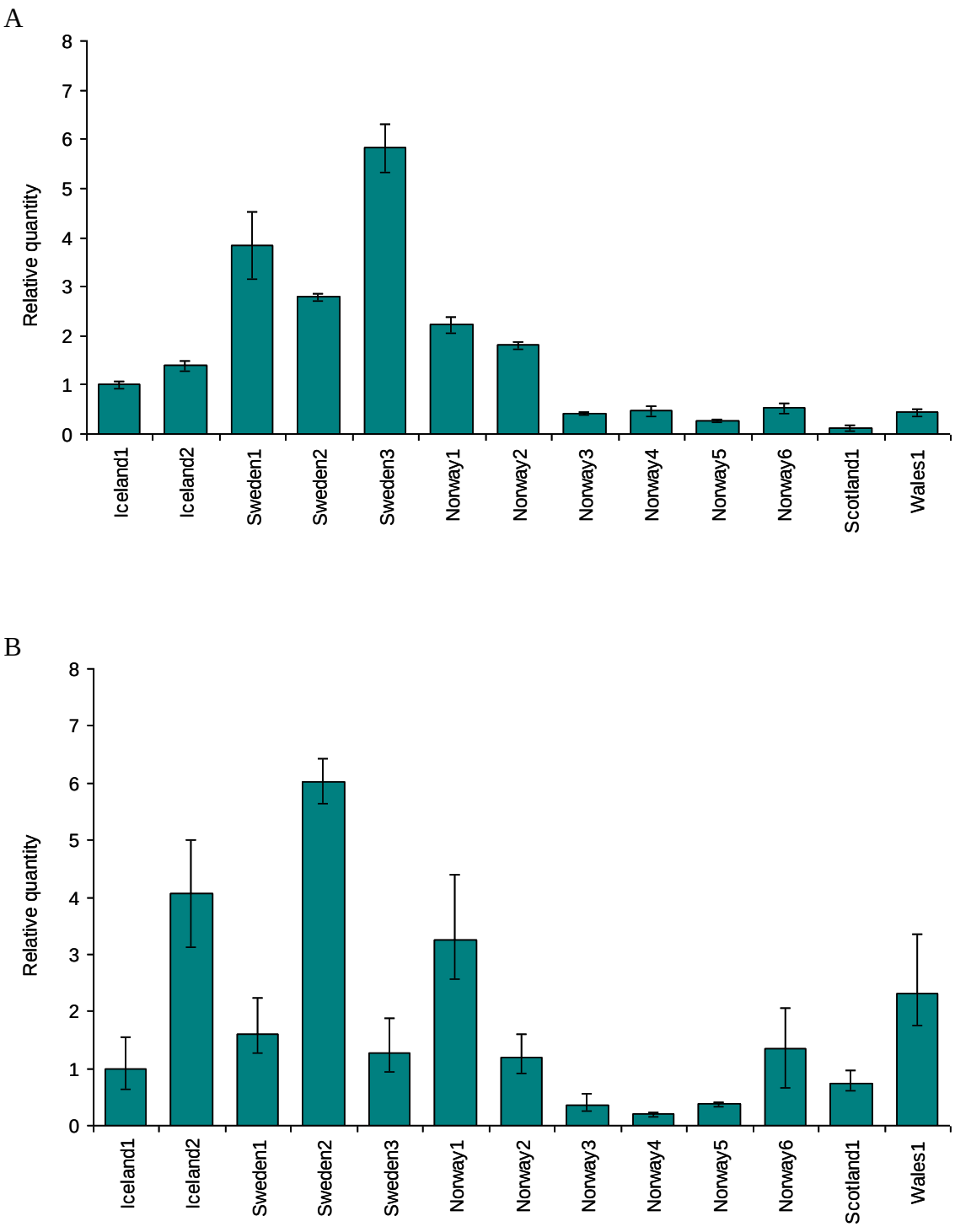


Figure 4.17 Relative quantity of expression of genes At1g08920 and At2g46830 in non-treated samples. Graph showing the relative quantity of gene expression of genes At1g08920 (A) and At2g46830 (B) in *A. l. petraea* populations prior to any temperature treatment. The quantity of gene expression is shown for all populations relative to the level of expression in population Iceland1. Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. These two genes are termed the 'general cold up-regulated' genes (section 4.2.4).

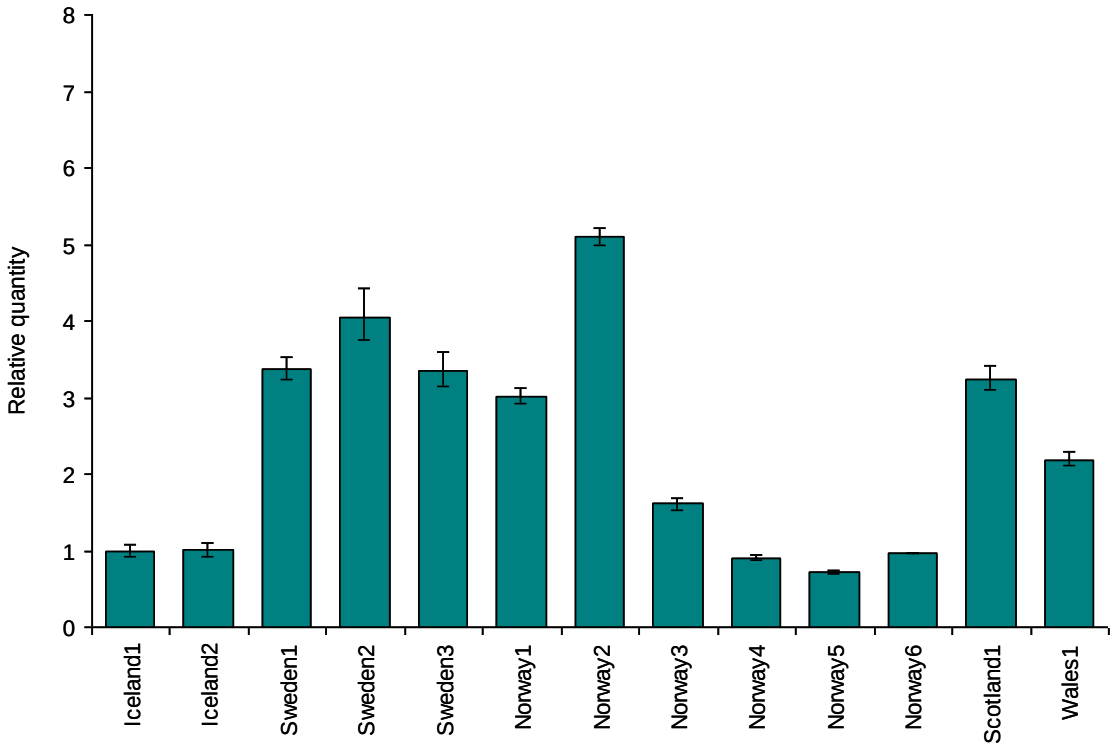


Figure 4.18 Relative quantity of expression of the gene At3g48360 in non-treated samples. Graph showing the relative quantity of gene expression of the gene At3g48360 in *A. l. petraea* populations prior to any temperature treatment. The quantity of gene expression is shown for all populations relative to the level of expression in population Iceland1. Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. This gene is termed the 'general cold down-regulated' gene (section 4.2.4).

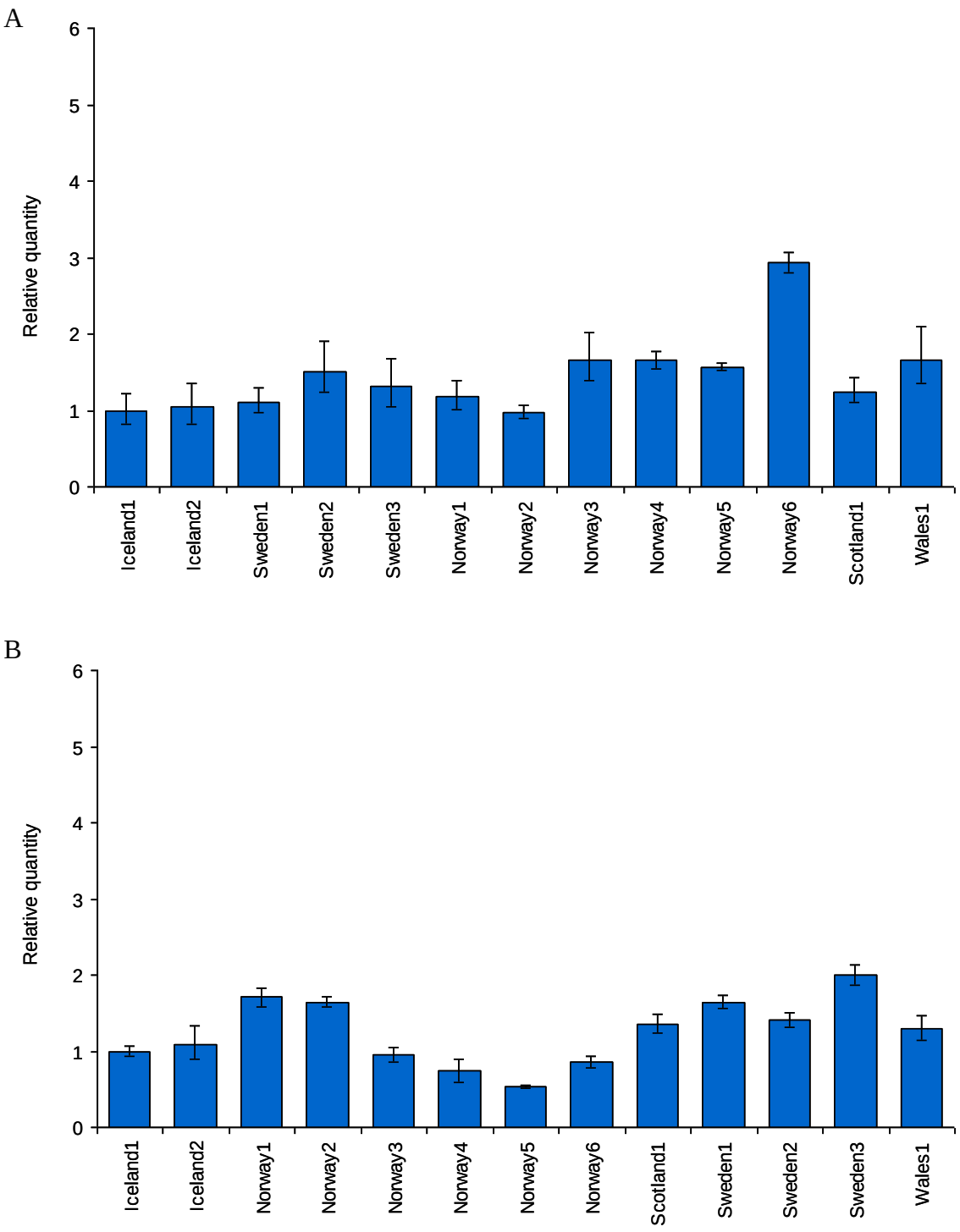


Figure 4.19 Relative quantity of expression of genes At2g40460 and At5g04410 in non-treated samples. Graph showing the relative quantity of gene expression of genes At2g40460 (A) and At5g04410 (B) in *A. l. petraea* populations prior to any temperature treatment. The quantity of gene expression is shown for all populations relative to the level of expression in population Iceland1. Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. These genes are termed 'interaction' genes and were both up-regulated to a greater extent in Helin than Leitrim in original microarray studies (section 4.2.4)

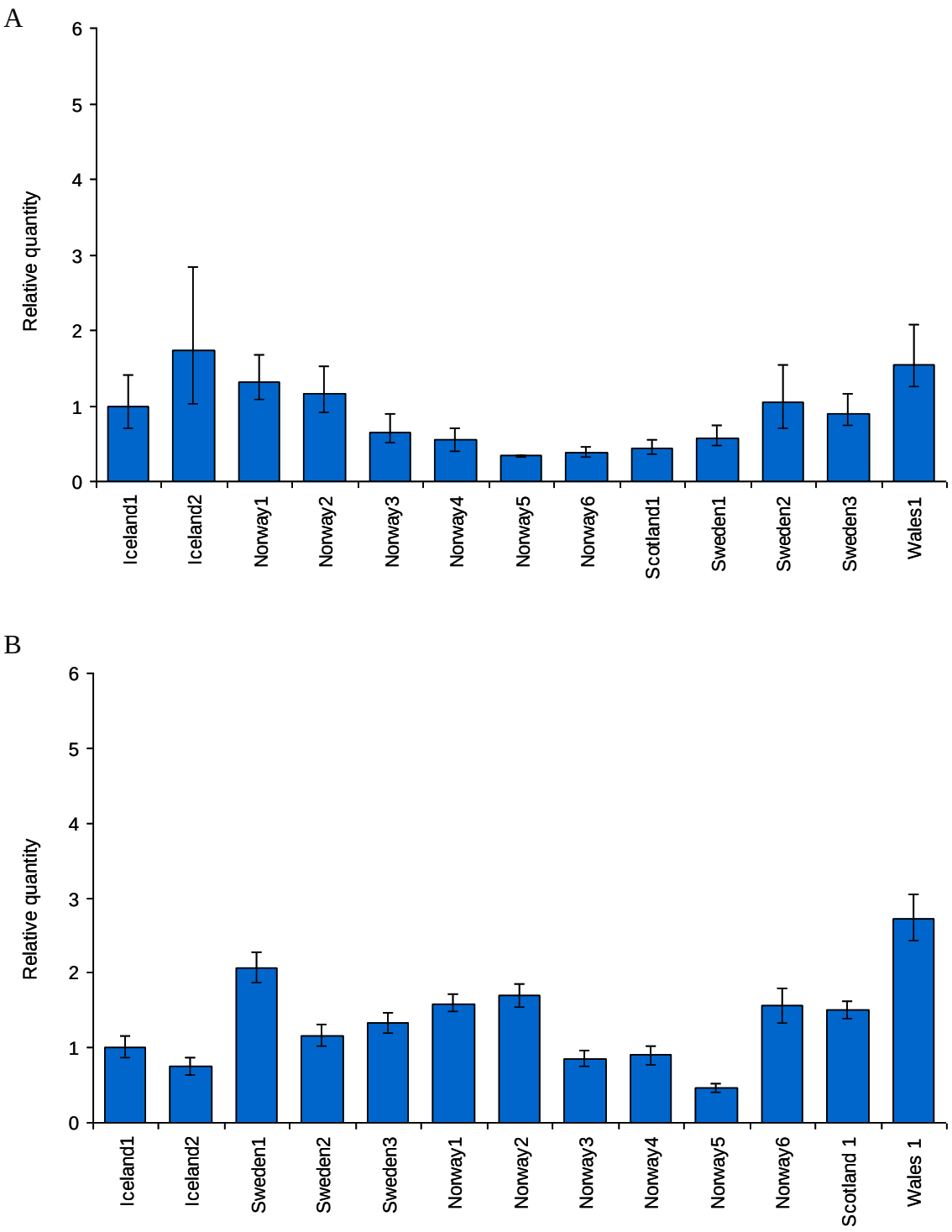


Figure 4.20 Relative quantity of expression of genes At5g05440 and At1g64890 in non-treated samples. Graph showing the relative quantity of gene expression of genes At5g05440 (A) and At1g64890 (B) in *A. l. petraea* populations prior to any temperature treatment. The quantity of gene expression is shown for all populations relative to the level of expression in population Iceland1. Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. These genes are termed 'interaction' genes and were both up-regulated to a greater extent in Leitrim than Helin in original microarray studies (section 4.2.4)

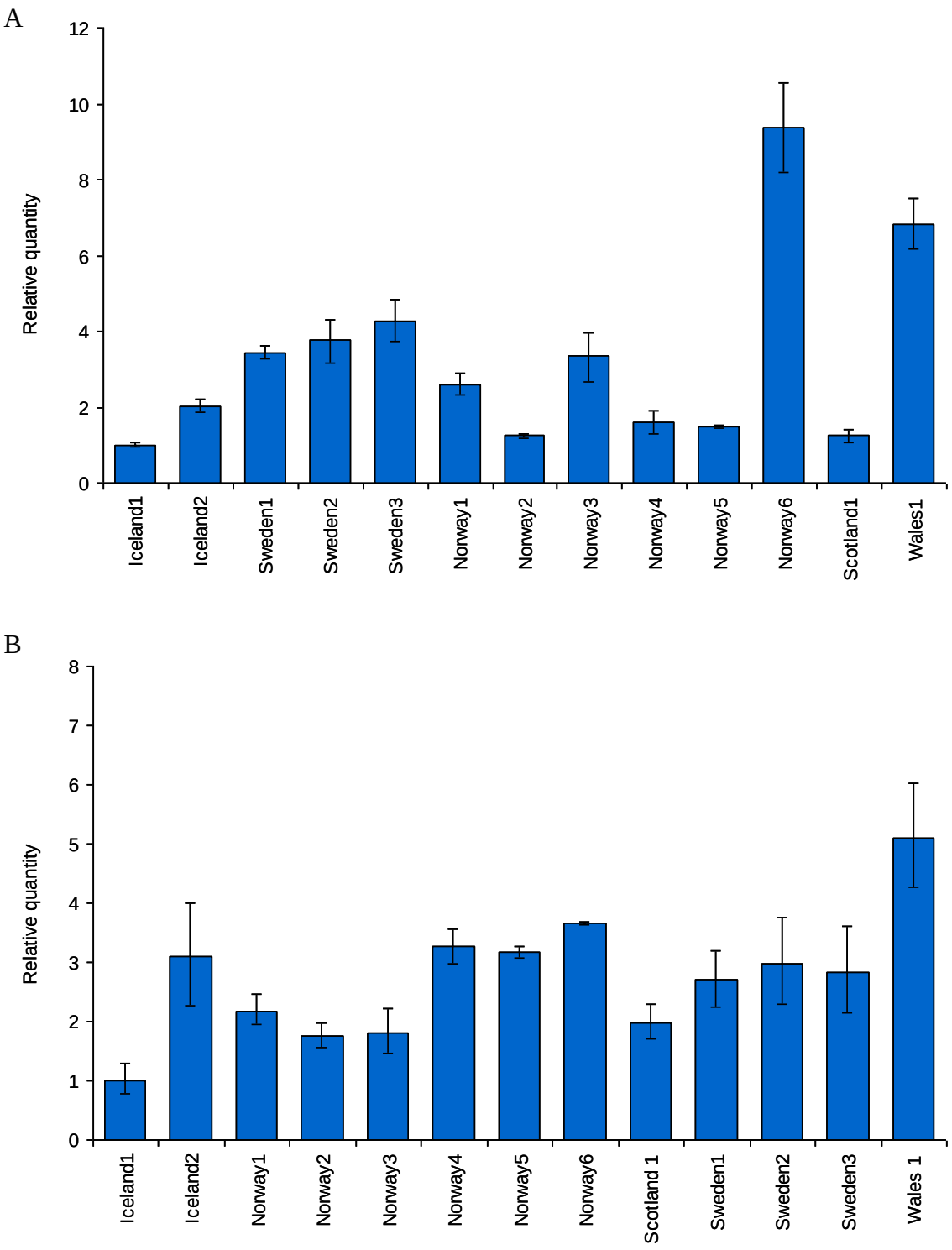


Figure 4.21 Relative quantity of expression of genes At1g06110 and At2g04040 in non-treated samples. Graph showing the relative quantity of gene expression of genes At1g06110 (A) and At2g04040 (B) in *A. l. petraea* populations prior to any temperature treatment. The quantity of gene expression is shown for all populations relative to the level of expression in population Iceland1. Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. These genes are termed 'interaction' genes; At1g06110 was up-regulated in Helin but not Leitrim, and At2g04040 was up-regulated in Leitrim but not Helin, in original microarray studies (section 4.2.4).

4.3.1.1 Fold changes in gene expression following exposure to chilling temperatures

qRT-PCR analysis was used to investigate the transcriptomic responses of nine genes in thirteen *A. l. petraea* populations following an 8 h chilling treatment at 4^{o}C. All genes studied were previously found to be cold responsive in the populations Helin, or Leitrim, or both, as detailed in section 4.2.4.

Genes At1g08920, and At2g46830, both 'general cold up-regulated' genes, were found to be significantly up-regulated in relation to the untreated samples in all thirteen populations following the chilling treatment ($P < 0.005$ for all populations). At1g08920 exhibited $\log_2$ fold changes of between 2.7 and 6.7 across the thirteen populations, and At2g46830 exhibited $\log_2$ fold changes of between 6.2 and 10.5 across the study populations (Figure 4.22). The 'general cold down-regulated' gene (At3g48360) was shown to be significantly down-regulated in all thirteen study populations following the chilling treatment (Figure 4.23). The extent of the $\log_2$ fold change and its significance varied across the populations, with the greatest down-regulation in population Norway6 ($P < 0.005$) and the smallest down-regulation in Norway1 ($P < 0.05$).

Genes At2g40460 and At5g04410 were both found to be up-regulated following chilling treatment in the initial microarray studies. Both genes were up-regulated in the populations Helin and Leitrim, but to a significantly greater extent in Helin. For At2g40460, qRT-PCR results show significant up-regulation of this gene in all thirteen populations ranging from a $\log_2$ fold change of 1.8 for Wales1 to a $\log_2$ fold change of 4 for Sweden1 (Figure 4.24A). For At5g04410, only eleven populations exhibit significant up-regulation following the chilling treatment; Norway2 and Norway6 do not show significant up-regulation of this gene. The highest $\log_2$ fold change for this gene of 2.5 is shown in the population Norway1 (Figure 4.24B). Genes At5g05440 and At1g64890 were both found to be up-regulated following chilling treatment in the initial microarray studies. Both genes were up-regulated in the populations Helin and Leitrim, but to a significantly greater extent in Leitrim. For At5g05440, qRT-PCR results show significant up-regulation of this gene in all thirteen populations; all $\log_2$ fold changes are greater than 4, with the greatest $\log_2$ fold change of 7.8 exhibited in the population Scotland1 (Figure 4.25A). For At1g64890, twelve of the thirteen study populations exhibit significant up-regulation of this gene following the chilling treatment; Norway3 does not show significant up-regulation of this gene. Only two populations, Iceland2 and Scotland1 exhibit $\log_2$ fold changes of greater than 2 (Figure 4.25B).

In the initial microarray studies the gene At1g06110 was shown to be up-regulated in Helin following the chilling treatment, but it was not shown to be temperature responsive in Leitrim. For this gene, qRT-PCR results show significant up-regulation of this gene across all thirteen populations following the chilling treatment, with $\log_2$ fold changes ranging from 0.8 for Norway2 to 4.1 for Sweden1 (Figure 4.26A). In the initial microarray studies the gene At2g04040 was shown to be up-regulated in Leitrim following the chilling treatment, but it was not shown to be temperature responsive in Helin. For this gene, qRT-PCR results show significant up-regulation of this gene across all thirteen populations, with $\log_2$ fold changes ranging from 1.4 for Iceland2 to 4.3 for Scotland1 (Figure 4.26B).

4.3.6 Clustering of fold change results across all study populations

Hierarchical cluster analysis was performed using data for all thirteen populations and fold change data for all nine genes studied. These results as presented as a dendrogram in Figure 4.27. The two most similar populations are Iceland1 and Iceland2. Norway2, 3, 4, and 5 cluster fairly closely indicating similarities in their fold change responses following chilling. The most distinct population is Norway6 as it forms a separate group on the dendrogram.

4.3.7 Additional replicates of qRT-PCR data

Biological replicates of the qRT-PCR data were performed, for genes differentially responsive to chilling temperatures, in a subset of the populations; this was to enable differences between populations to be compared using analysis of variance. Populations were chosen for this based on the qRT-PCR results outlined above and were selected to capture the greatest variation in gene expression across all genes of interest, both prior to and following chilling treatment. Four populations were chosen: Iceland1, Norway2, Norway6, and Wales1, and two further biological replicates were carried out as detailed in section 4.2.8. All thirteen populations could not be included in this due to time constraints.

The six interaction genes (Table 4.1) were studied in these additional biological replicates. Considerable variation was observed between the biological replicates within each population, but no significant differences in gene expression were observed between the populations for the individual genes (nested ANOVA, Table 4.5). Overall,

the patterns of expression obtained in the first replicate were confirmed in these additional replicates, although the extent of expressional changes varied between the replicates (data not shown).

Gene (AGI code)	Variation among biological replicates (qRT-PCR data)	Variation between study populations
At1g06110	$F_{8,21} = 112.4$, $P < 0.005$	$F_{3,8} = 0.64$, $P = 0.609$
At1g64890	$F_{8,23} = 58.5$, $P < 0.005$	$F_{3,8} = 0.73$, $P = 0.561$
At2g04040	$F_{8,23} = 29.43$, $P < 0.005$	$F_{3,8} = 2.69$, $P = 0.117$
At2g40460	$F_{8,24} = 34.31$, $P < 0.005$	$F_{3,8} = 0.25$, $P = 0.859$
At5g04410	$F_{8,24} = 13.16$, $P < 0.005$	$F_{3,8} = 0.43$, $P = 0.737$
At5g05440	$F_{8,24} = 40.17$, $P < 0.005$	$F_{3,8} = 2.79$, $P = 0.110$

Table 4.5 Statistical analysis of qRT-PCR data. Nested ANOVAs reveal significant variation among biological replicates of qRT-PCR data for all genes studied, whilst variation between populations was not found to be significant for any gene under investigation.

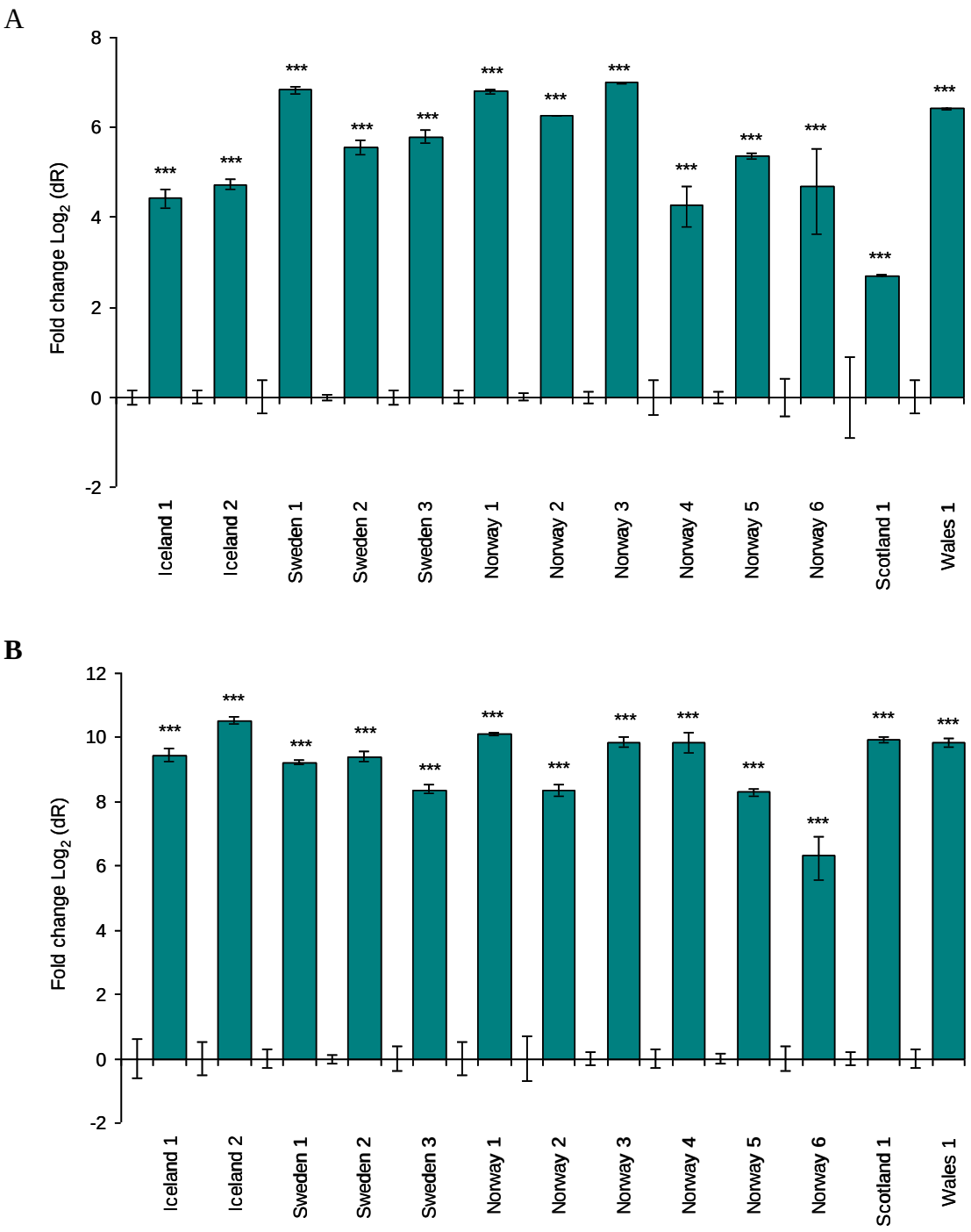


Figure 4.22 Expression of genes At1g08920 and At2g46830 following chilling treatment. Graph showing the $\log_2$ fold change in gene expression of genes At1g08920 (A) and At2g46830 (B) in *A. l. petraea* following an 8 h chilling treatment at 4°C. These genes are termed the 'general cold up-regulated' genes (section 4.2.4). Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. For each population, the first error bars represent the untreated sample and the coloured bars represent the chilling treated sample. Asterisks (*) indicate the significance of the fold change within each population between treated and untreated plants (* $P<0.05$; ** $P<0.01$; *** $P<0.005$) as tested through analysis of variance (Appendix 1).

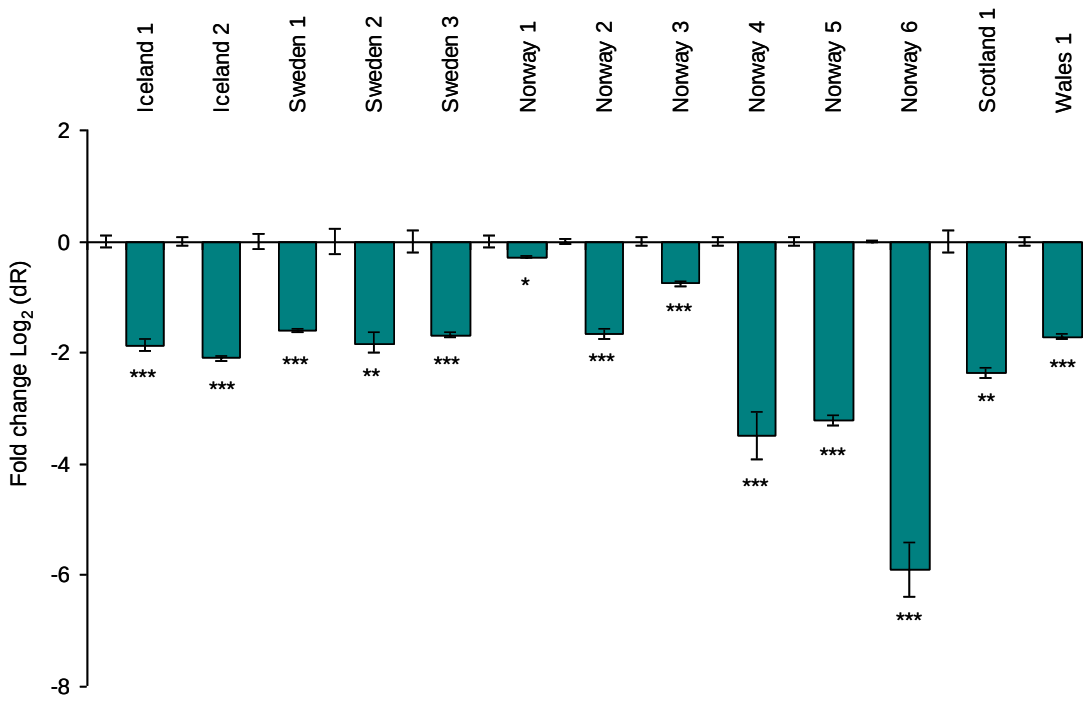


Figure 4.23 Expression of the gene At3g48360 following chilling treatment. Graph showing the $\log_2$ fold change in gene expression of the gene At3g48360 in *A. l. petraea* following an 8 h chilling treatment at 4°C. This gene is termed the 'general cold down-regulated' gene (section 4.2.4). Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. For each population, the first error bars represent the untreated sample and the coloured bars represent the chilling treated sample. Asterisks (*) indicate the significance of the fold change within each population between treated and untreated plants (* P<0.05; ** P<0.01; *** P<0.005) as tested through analysis of variance (Appendix 1).

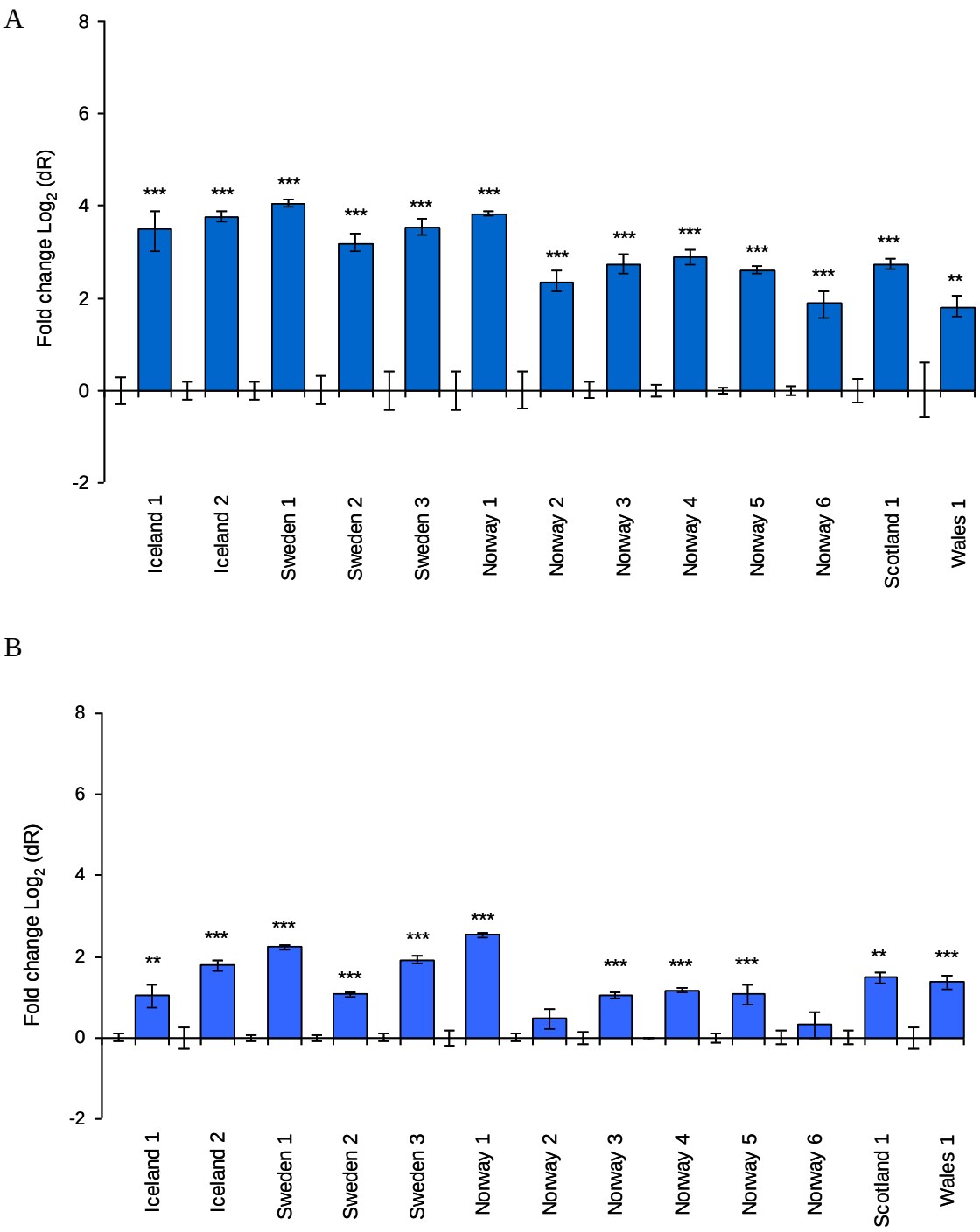


Figure 4.24 Expression of genes At2g40460 and At5g04410 following chilling treatment. Graph showing the $\log_2$ fold change in gene expression of genes At2g40460 (A) and At5g04410 (B) in *A. l. petraea* following an 8 h chilling treatment at 4°C. These genes are termed 'interaction' genes and were both up-regulated to a greater extent in Helin than Leitrim in original microarray studies (section 4.2.4) Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. For each population, the first error bars represent the untreated sample and the coloured bars represent the chilling treated sample. Asterisks (*) indicate the significance of the fold change within each population between treated and untreated plants (* P<0.05; ** P<0.01; *** P<0.005) as tested through analysis of variance (Appendix 1).

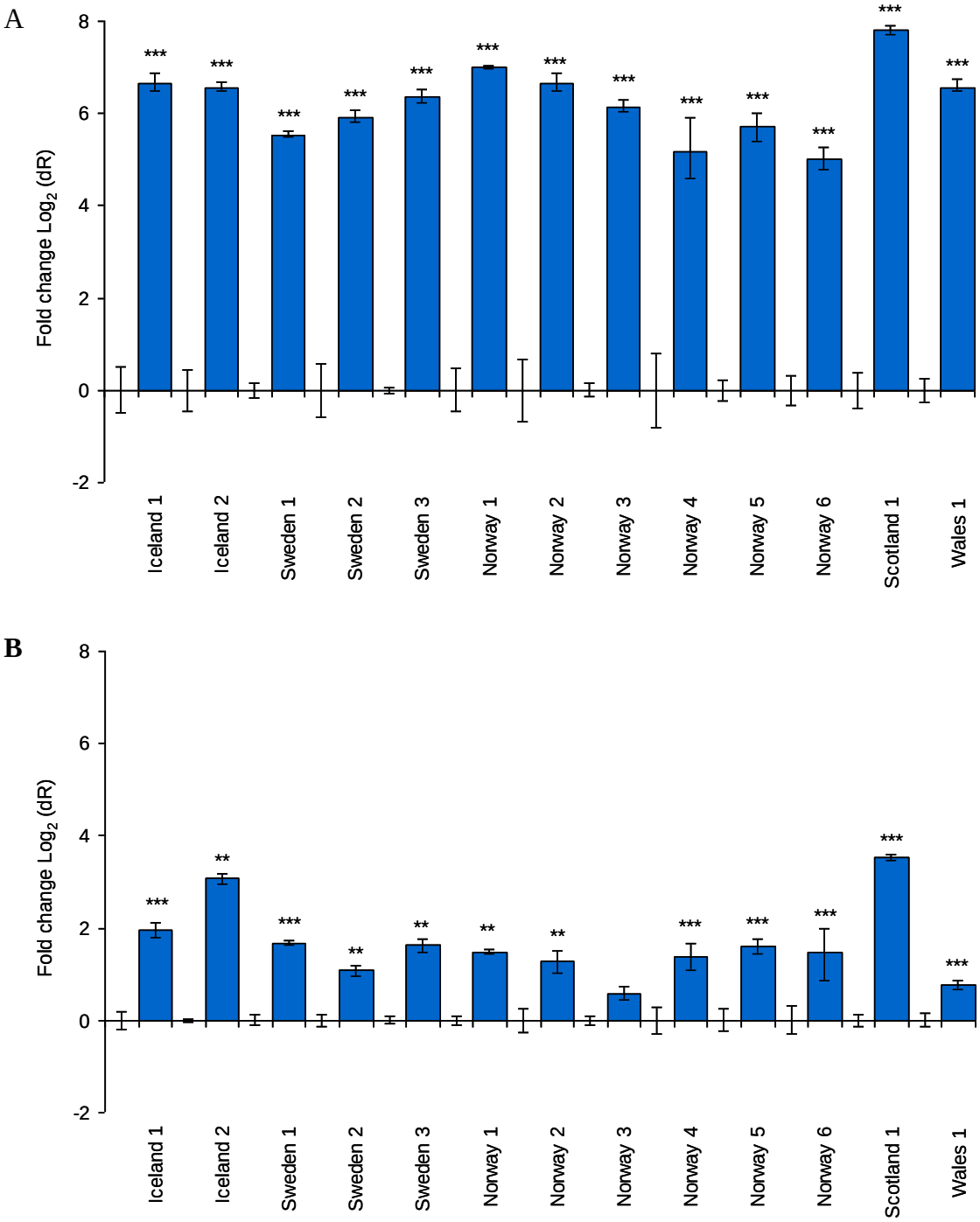


Figure 4.25 Expression of genes At5g05440 and At1g64890 following chilling treatment. Graph showing the $\log_2$ fold change in gene expression of genes At5g05440 (A) and At1g64890 (B) in *A. l. petraea* following an 8 h chilling treatment at 4°C. These genes are termed 'interaction' genes and were both up-regulated to a greater extent in Leitrim than Helin in original microarray studies (section 4.2.4) Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. For each population, the first error bars represent the untreated sample and the coloured bars represent the chilling treated sample. Asterisks (*) indicate the significance of the fold change within each population between treated and untreated plants (* $P<0.05$; ** $P<0.01$; *** $P<0.005$) as tested through analysis of variance (Appendix 1).

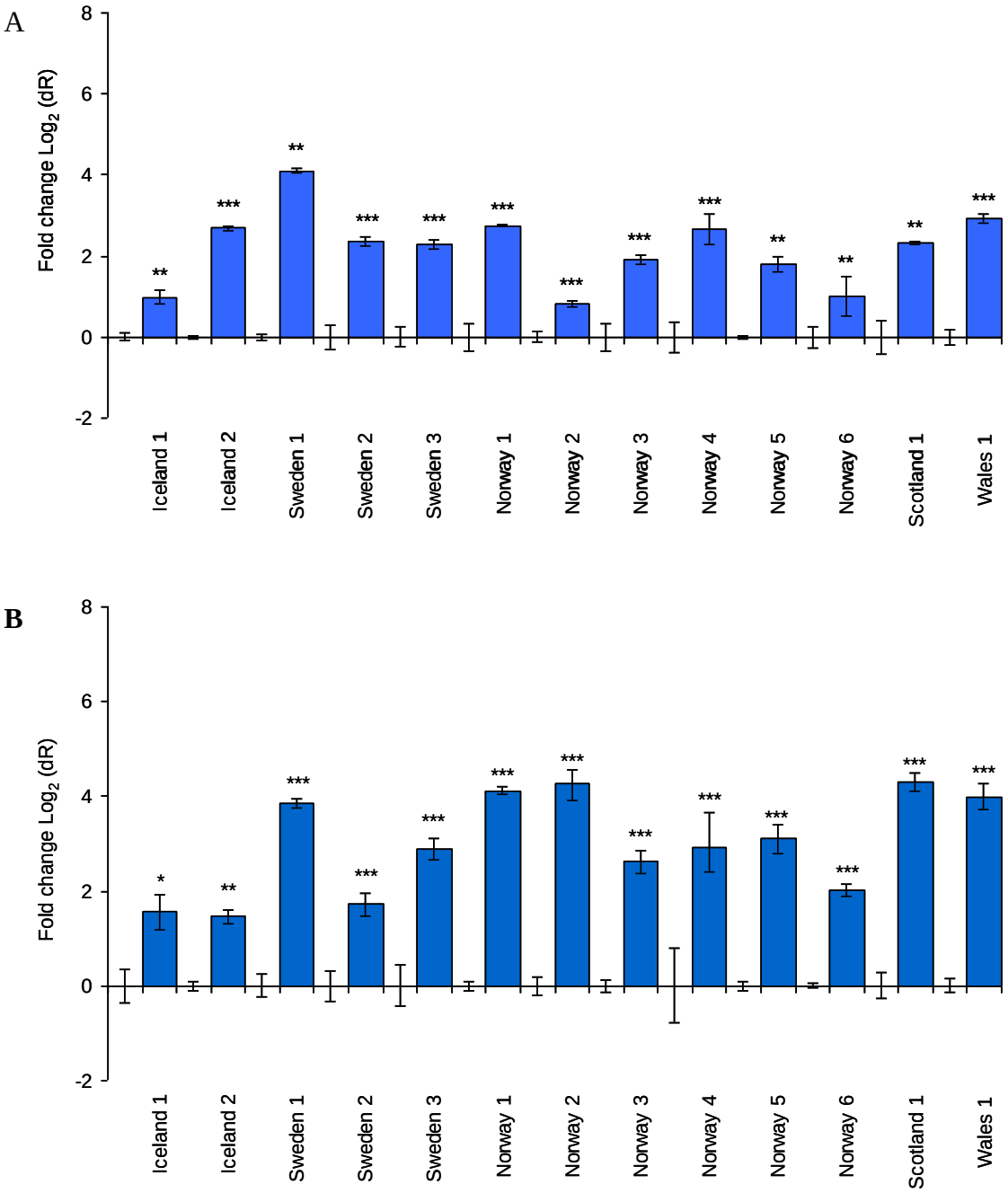


Figure 4.26 Expression of genes At1g06110 and At2g04040 following chilling treatment. Graph showing the $\log_2$ fold change in gene expression of genes At1g06110 (A) and At2g04040 (B) in *A. l. petraea* following an 8 h chilling treatment at 4°C. These genes are termed 'interaction' genes; At1g06110 was up-regulated in Helin but not Leitrim, and At2g04040 was up-regulated in Leitrim but not Helin, in original microarray studies (section 4.2.4). Expression was quantified through qRT-PCR in thirteen populations from across the species' range. Populations are identified in Table 2.1. For each population, the first error bars represent the untreated sample and the coloured bars represent the chilling treated sample. Asterisks (*) indicate the significance of the fold change within each population between treated and untreated plants (* $P<0.05$; ** $P<0.01$; *** $P<0.005$) as tested through analysis of variance (Appendix 1).

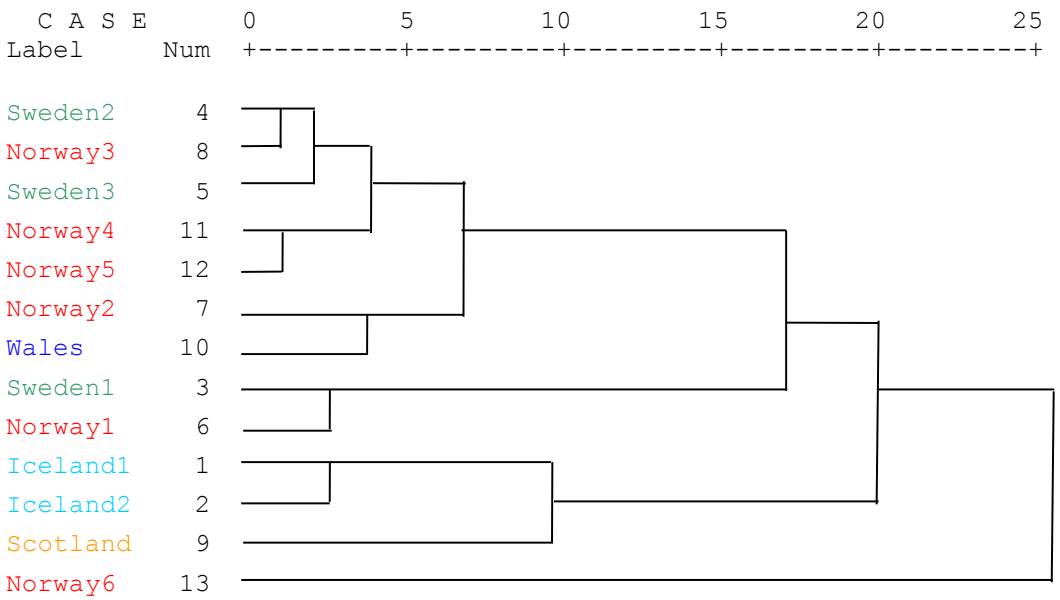


Figure 4.27 Cluster dendrogram of all populations based on gene expression results. Cluster analysis was performed using the qRT-PCR data for all genes and populations studied. Num refers to the order in which the populations clustered. Iceland1 and Iceland2 cluster first and are therefore the most similar populations in terms of their transcriptomic responses to chilling treatment. Norway6 is the most distinct population as this is the last to cluster.

4.4 Discussion of gene expressional differences between populations of *A. l. petraea* in response to chilling temperatures

4.4.1 Normalisation and relative quantification by standard curve

Normalisation of qRT-PCR data is required to correct for issues such as inter-reaction variation and differences in the starting quantities of RNA. Reference genes such as ubiquitin are commonly used for normalisation due to their invariant expression levels across conditions, yet this approach is not without limitations (Thellin *et al.*, 1999). A genome-wide survey of *Arabidopsis* identified hundreds of genes that were more stably expressed than traditionally used reference genes across a range of growth stages, organs, and environmental conditions, including the gene At1g13320 (*PDF2*) (Czechowski *et al.*, 2005). *PDF2* primers used in this work were based on primers designed by Czechowski *et al.* (2005) but were specific to the *A. l. petraea* sequence rather than to *A. thaliana*. In the microarrays performed by Dr. C. J. Lilley, expression of this gene was not significantly different between untreated and cold-treated *A. l. petraea* plants; this further indicates its suitability for use as a reference gene in this study.

Performance of qRT-PCR reactions can be monitored through primer efficiency values. The optimum efficiency is 100% and this equates to a doubling of the target amplicon with each PCR cycle. Efficiencies of between 90% and 110% are considered acceptable for most applications (Stratagene, 2007). Reactions with the reference gene *PDF2* had an efficiency of 103% indicating that this assay was highly optimised and reliable. Efficiencies for all target gene assays were also within the acceptable range thus quantity of expression will have been reliably estimated from C_t values.

4.4.2 Sequence similarities between *A. thaliana* and *A. l. petraea*

A. l. petraea has rapidly developed as a model species, in part due to the genetic similarities with *A. thaliana* that have enabled application of tools developed for *A. thaliana* to *A. l. petraea* studies. The *A. l. petraea* sequences for coding regions obtained in this study demonstrated at least 89% similarity with *A. thaliana* coding sequences thus confirming genetic similarities between the two species. Greater sequence differentiation was observed for introns; there was only 66% similarity between the two species for the introns of At2g04040. This is in agreement with studies of the *FRI* locus that revealed a greater divergence between intronic regions than for

exons between *A. thaliana* and *A. l. petraea*, suggesting that the level of genetic variation in coding regions is subject to selective constraints (Le Corre *et al.*, 2002).

It is known that genetic differentiation exists between different *A. thaliana* ecotypes. Comparison of DNA sequences between a number of ecotypes revealed small clusters of related ecotypes, with extensive differentiation between the clusters (Hardtke *et al.*, 1996). A link between the sequences and the geographic origin of ecotypes was not observed. Microsatellites have also been used to quantify genetic diversity between ecotypes; a wide range of geographic locations were reported within each distinct cluster of ecotypes e.g. plants from Norway and Morocco were placed into the same cluster (Innan *et al.*, 1997). Sequence differentiation between *A. thaliana* ecotypes has been reported for a number of specific genes including floral development genes (Le Corre et al 2002) and plant defence genes (Kuittinen and Aguade, 2000; Hauser *et al.*, 2001). Hauser *et al.* (2001) report high sequence diversity between 26 accessions at the locus *GLABROUS1 (GL1)*, and were able to assign ecotypes to two diverged sequence clades. Le Corre *et al.* (2002) studied the sequences of 25 ecotypes at the locus *FRIGIDA (FRI)*; they reported a range of mutations between the ecotypes including individual nucleotide polymorphisms and indels.

As DNA sequences differ between *A. thaliana* ecotypes, it can be expected that genetic differentiation will be apparent between distinct populations of *A. l. petraea*. Ramos-Onsins *et al.* (2004) reported significant genetic differentiation between populations of *A. l. petraea* from diverse locations including Germany, Russia, and Scotland (eight genes studied). Another close relative, *A. halleri*, also demonstrated genetic differentiation between populations but to a lesser extent than for *A. l. petraea*, thus indicating a higher level of population structure within *A. l. petraea* (Onsins *et al.*, 2004). Variation among *A. l. petraea* populations was reported as three-fold higher than for the *A. thaliana* populations studied (three genes) (Ramos-Onsins *et al.*, 2004). In this study, one housekeeping gene (At1g13320) was sequenced (within the coding region) for both Helin and Leitrim populations, and sequence differences were observed between the populations. Although this data refers to a single gene, these observed population-specific sequences potentially suggest that very little, if any, gene flow occurs between these two populations and that they may therefore be genetically distinct. Primers designed to amplify the coding region of At4g05320 (*UBQ10*) worked well for Helin cDNA template but did not successfully amplify Leitrim cDNA (section 4.3.1).

Additionally, all three pairs of qRT-PCR primers designed to the Helin *A. l. petraea* sequence produced non-specific products for Leitrim template (section 4.3.4.1). These primer specificity issues further suggest sequence differences between populations. Such genetic differences could be responsible for differential expression of genes between populations, and indicate adaptation to local environmental conditions.

4.4.3 Cold up-regulated and down-regulated genes

The two 'general cold up-regulated' genes chosen for study (section 4.2.5) demonstrated significant up-regulation of expression in response to the chilling treatment in each of the additional thirteen study populations. It is therefore likely that these genes are required in the general response of *A. l. petraea* to cold temperature conditions.

The gene At1g08920 was significantly up-regulated in all thirteen study populations following exposure to 4^{o}C, with all populations except Scotland1 exhibiting a greater than 4 fold $\log_2$ change in expression (Figure 4.22). This suggests that up-regulation of this gene is important in the response of *A. l. petraea* to cold temperature. This gene is annotated as a putative sugar transporter and has been named *ERD SIX-LIKE 1* (*ESL1*) due to its close homology with *ERD6* (*EARLY RESPONSE TO DEHYDRATION 6*) (Yamada *et al.*, 2009). *ESL1* has been shown to respond to abiotic stresses in previous microarray studies demonstrating up-regulation in response to cold, drought, osmotic, and salt stresses, and down-regulation in response to heat stress in *A. thaliana* (GENEVESTIGATOR). Yamada *et al.* (2009) monitored *ESL1* expression in *A. thaliana* and found that ESL1 protein was localised at tonoplasts, and characterised it as a facilitated diffusion transporter. Expression levels of *ESL1* were similar to those of vacuolar invertase genes, and Yamada *et al.* (2009) suggest that increased expression of invertases under abiotic stress may result in accumulation of hexose sugars in the vacuole. They consequently infer that ESL1 may function in the efflux of hexoses from the vacuole to the cytoplasm, thus regulating osmotic pressure and aiding sugar remobilisation under abiotic stress conditions. The qRT-PCR results here suggest that *ESL1* performs the same function in *A. l. petraea* as in *A. thaliana*, and it therefore plays an important role in transportation of sugars during abiotic stress. Sugars, and enzymes involved in their synthesis, are documented to increase under low temperature stress (Strand *et al.*, 1999; Kaplan *et al.*, 2004) so observed increases in transporter gene expression correlate well with this. Expression levels of this gene prior to temperature

stress are higher for the three Swedish populations than for all other populations suggesting that perhaps this gene is a more important component of the Swedish plants' cold response, and that it may be of benefit for these populations to be primed ready for exposure to cold temperature.

Expression of the gene At2g46830 (*CIRCADIAN CLOCK ASSOCIATED1*, *CCA1*) was very significantly up-regulated in all populations following exposure to cold temperatures (Figure 4.22). Considering all populations, this gene demonstrated the greatest extent of up-regulation of all genes studied with every population exhibiting a $\log_2$ fold change of at least six. This large degree of up-regulation is indicative of importance in the cold response of *A. l. petraea*. *CCA1* is a core component of the *Arabidopsis* circadian clock, and functions with two other genes (*LATE ELONGATED HYPERCOTYL*, *LHY;* and *TIMING OF CAB EXPRESSION 1, TOC1*) in the primary circadian feedback loop (Fukushima *et al.*, 2009). Expression of *CCA1* and *LHY* rises before dawn to regulate expression of *TOC* (Alabadi *et al.*, 2001). It has been shown that *CCA1* transcript levels degrade more rapidly in light than dark, with a half life of approximately 1.5 h, (Yakir *et al.*, 2007), indicating that they are one of the unstable to moderately unstable *Arabidopsis* genes that comprise about 4% of the genome (Gutierrez *et al.*, 2002). Interactions between cold temperatures and the circadian clock can be assumed as there is evidence of both circadian regulation of the CBF genes (Edwards *et al.*, 2006), and that particular circadian-regulated genes are cold responsive (Kreps *et al.*, 2002). Bieniawska *et al*. (2008) demonstrate through a principal component analysis (PCA) that circadian- and diurnal- regulated genes make the largest contribution to differences between independent investigations into the cold response in *Arabidopsis*. They also reveal that many circadian clock related genes either demonstrate reduced amplitude in cycles of expression, or that oscillations stop completely in response to cold. During four cold treatment regimes *CCA1* increased in expression during the dark phase but was only reduced slightly after the start of the light phase, whilst dramatic reductions in expression were illustrated in the control (20^{o}C) sample after onset of the light phase (Bieniawska *et al*., 2008). Here, qRT-PCR results demonstrate significantly higher expression of *CCA1* in cold treated samples from all *A. l. petraea* study populations. As leaf samples were harvested during the light period, it is possible that *CCA1* cycling was disrupted in these plants in response to the cold temperatures resulting in increased transcript abundance in the treated plants.

The gene *BT2* (At2g48360) is one of five members of the family BTB AND TAZ DOMAIN PROTEINS. Expression of this gene was significantly reduced in all study populations following exposure to cold temperatures, but the extent of down-regulation varied considerably between populations (Figure 4.23). Interestingly, the three populations that exhibited the greatest down-regulation of this gene had the lowest relative expression levels prior to cold temperature exposure (Norway4, Norway5, and Norway6). Functional analysis of the BT family has demonstrated that BT proteins are land-plant specific, and that it is likely that they act in, or perhaps interconnect, multiple cellular pathways (Robert *et al.*, 2009). *BT2* has been identified as an essential component of the TAC1-telomerase pathway, and is induced by the expression of *TAC1* (Ren *et al.*, 2007). Telomerase synthesises and maintains telomeric repeats at chromosome ends. Mandadi *et al.* (2009) demonstrate that *BT2* has broader functions relating to responses to changes in light signals, nutrient status, hormones, and specific stresses. They report that *BT2* is highly expressed in the dark and almost undetectable under light conditions, with high presence of sugars during the light period thought to be responsible. Abscisic acid (ABA) was also shown to decrease *BT2* expression and, as cold levels are known to increase levels of ABA (Chen *et al.*, 1983), this may explain the significant decreases in expression of *BT2* in populations of *A. l. petraea* following exposure to cold. *BT2* has exhibited down-regulation to a range of abiotic stresses in previous microarray experiments including salt and light and, to a lesser extent, cold, with very limited up-regulation in response to any stress treatment (GENEVESTIGATOR). The varying extents of this down-regulation across study populations may indicate limited gene flow between populations, and that a reduction in expression may be of greater adaptive significance for some populations than others. Of the Norwegian populations, Norway4, 5, and 6 experience far more degree days where temperatures are $< -5^{o}C$ (between 48 and 88 degree days) than the remainder of the Norwegian populations (between 0 and 28 degree days). It is possible that this frequent exposure to particularly cold conditions maintains higher levels of ABA in these plants, perhaps explaining the reduced relative expression prior to cold exposure and the more significant response observed to the chilling stress.

4.4.4 Interaction genes

As with the ‘general cold up-regulated’ and ‘general cold down-regulated’ genes, qRT-PCR results demonstrate that the expression levels of the six interaction genes varied considerably between *A. l. petraea* populations, both prior to cold temperature exposure

and in response to the chilling treatment. This was not unexpected as the interaction genes were selected based on the evidence that they were differentially regulated in Helin and Leitrim, the two initial study populations, in microarray studies (section 4.2.4). Taking into account the expression values for all of the interaction genes across the thirteen populations, no obvious patterns of expression in response to chilling emerge. Specific populations do not consistently exhibit the greatest or least fold changes in expression across all genes. Likewise, there are no obvious patterns for the relative quantity expression values prior to cold temperature exposure. The absence of a consistent pattern indicates local adaptation of populations, with certain genes potentially being of greater importance in the cold response of specific populations.

The gene At5g04410 (*NAC2*) was significantly up-regulated in eleven of the thirteen study populations, yet $\log_2$ fold change values for this gene were some of the lowest observed for all genes (Figure 4.24). Data available from GENEVESTIGATOR present both up- and down-regulation of *NAC2* in response to cold temperatures, but to small extents. Together, this suggests that *NAC2* may not be particularly responsive to cold stress in *Arabidopsis*. *NAC2* is one of a large family of over 100 plant-specific NAC transcription factors in *A. thaliana* that appear to be involved in various plant developmental systems (Ooka et al 2003). Some individual NAC genes have been linked to abiotic stresses including dehydration and salinity (Hegedus *et al.*, 2003; Fujita *et al.*, 2004; Tran *et al.*, 2004). NAC2 has recently been characterised; proteins are localised to the nucleus and cytoplasm, and transcription is induced under high light conditions. *Arabidopsis* plants over-expressing this gene exhibited up-regulation of genes related to flavonoid biosynthesis under high light conditions, suggesting that *NAC2* is associated with the induction of flavonoid biosynthesis genes (Morishita *et al.*, 2009). Flavonoids are known to increase under both abiotic and biotic stresses (Dixon and Paiva, 1995; Winkel-Shirley, 2002), and have potential roles in protection against oxidative stress (Hernández *et al.*, 2009), thus increased expression of *NAC2* in *A. l. petraea* populations following cold temperature stress may enhance flavonoid biosynthesis and consequently infer protection against the adverse effects of cold exposure. The fairly low levels of up-regulation indicate the limited importance of this gene in the cold response of *A. l. petraea.* Log_2 fold change values were similar across all populations, and there were also limited differences in relative expression levels prior to cold temperature exposure.

Relative expression of the gene At2g40460 was similar across the study populations prior to cold temperature exposure, with the exception of Norway6 (Figure 4.24). This may reflect the small size of this population and its location at high altitude; these factors may contribute to isolation of Norway6 and prevent potential gene flow. Significant up-regulation of this gene in response to cold in all population indicates that this gene may be of importance in the cold response of *A. l. petraea*. Log_2 fold change values are higher for this gene than for *NAC2* and a range of expression values are observed across the populations. At2g40460 is annotated as a peptide transporter, and is a member of the NRT(PTR) family that contains both low-affinity nitrate transporters and small peptide transporters. This family contains over 50 genes in *Arabidopsis*, far more than the same family in humans, Drosophila, and yeast, and they exhibit a range of different tissue expression patterns that may indicate unique functions (Tsay *et al.*, 2007). Tsay *et al.* (2007) suggest that the presence of such a large family in *Arabidopsis* may indicate involvement in the transport of other substrates (similar to nitrate and dipeptides), including those involved in stress adaptation. If this were the case, it could explain the up-regulation of this peptide transporter in response to temperature stress. In previous microarray experiments this locus has been shown to be up-regulated in response to cold and drought, and down-regulated in response to heat (GENEVESTIGATOR) which may further indicate a role in the abiotic stress response of *Arabidopsis*.

The qRT-PCR results for At5g05440 illustrate significant up-regulation of this gene in all populations following exposure to cold temperatures (Figure 4.25). Data available from GENEVESTIGATOR detail a large extent of both up- and down-regulation of this gene in response to cold temperature stress, and also in response to drought conditions. This gene is shown to be down-regulated to the greatest extent for both heat and glucose treatments. The function of this gene is currently unknown hence it is not possible to explain why significant up-regulation is observed following exposure to 4^oC in *A. l. petraea*. The log_2 fold change values are greater than four in all study populations indicating that this gene is of significant importance in the cold temperature response. The greatest difference in extent of up-regulation is between Norway6 (smallest) and Scotland1 (largest), yet these two population have a very similar relative quantity of expression in untreated plants. This suggests that up-regulation of this gene confers a different selective advantage in individual study populations, with populations demonstrating differing levels of expression in response to the same degree of stress.

The gene At1g64890 is annotated as an integral membrane transporter protein, implying roles in the transport of substances through the membrane. This gene has not been further characterised so the nature of possible substrates is not known. Previous microarray studies report this gene to increase expression levels following exposure to cold stress, and also under osmotic stress, whilst exhibiting reduced expression following exposure to heat (GENEVESTIGATOR). Over-representation of transport proteins was reported in genes deemed responsive to cold temperatures in *Arabidopsis*, indicating the importance of transport in and out of the cell in response to low temperatures (Hannah *et al.*, 2005). Further characterisation of the function of this gene will be required to confirm its specific contribution to the cold response. The qRT-PCR results imply that this gene does exhibit differential expression between the study populations, both pre- and post-temperature stress (Figures 4.20 and 4.25). Wales1 exhibited the highest relative expression in the untreated samples but had one of the lowest amounts of up-regulation following temperature stress. It is possible that there is an optimal level of expression for this gene that enables a sufficient response to be realised, and that other populations require a greater extent of up-regulation upon cold temperature exposure to reach this level. Alternatively, particular populations (in this case, Wales1) may be better primed for potential cold temperature exposure.

The qRT-PCR results for At1g06110 (*SKP1/ASK INTERACTING-PROTEIN 16, SKIP16*) reveal considerable variation in expression between the populations in the untreated samples, with Norway6 and Wales1 having particularly high expression values (Figure 4.21). Following cold temperature exposure, all populations exhibited significant up-regulation but many populations had small $\log_2$ fold change values (Figure 4.26). This gene was up-regulated in Helin but not Leitrim in previous microarray analysis where the cut-off for designating a gene as up-regulated was a fold change of at least two. If the fold change cut-off for the qRT-PCR results is also set at two, then a number of the populations do not exhibit up-regulation of this gene in response to cold temperatures; *SKIP16* is therefore differentially regulated between populations of *A. l. petraea.* *SKIP16* functions in protein binding, and is part of the SCF ubiquitin ligase complex. This complex has a role in the ubiquitination and consequent degradation of proteins, but the precise role of this gene has not been characterised. Data available from GENEVESTIGATOR suggest that this gene is not particularly responsive to environmental stresses in *Arabidopsis*, as only small changes in expression are reported across a range of conditions.

The gene At2g04040, (*Arabidopsis thaliana* *DETOXIFICATION 1*, *AtDTX1*), is a member of a large multigene family in *Arabidopsis* (over 50 genes), is ubiquitously expressed in all organs, and is localised to the cell membrane. It has been characterised as an efflux carrier for toxic compounds including antibiotics and the heavy metal cadmium (Li *et al.*, 2002). The qRT-PCR results illustrate differences in expression between the populations prior to cold temperature exposure indicating adaptation to local environmental conditions (Figure 4.21). Soil properties may vary significantly between sites, and perhaps contain different levels of toxic compounds, hence leading to differential requirements for detoxification. Turner *et al.* (2010) recently demonstrated significant sequence differences between populations of *A. l. petraea* found on serpentine and non-serpentine soils, and these may relate to differences in expression levels of the genes concerned. The extent of fold change for *AtDTX1* in response to cold treatment, determined through qRT-PCR, also varies between the populations (Figure 4.26). This gene has been shown to respond to cold treatment in previous microarrays, with high levels of up-regulation reported, and is also responsive to a range of other stresses including drought (GENEVESTIGATOR). Li *et al.* (2002) speculate that AtDTX1 proteins may be involved in export of compounds into the xylem for long distance transport, and that such compounds could include secondary metabolites. Metabolites are known to increase during exposure to chilling temperatures in both *A. thaliana* and *A. l. petraea* (Hannah *et al.*, 2006; Davey *et al.*, 2009). Population specific differences in metabolites have been demonstrated for *A. l. petraea*, both under control conditions and in response to cold acclimation (Davey *et al.*, 2008; Davey *et al.*, 2009) and these may explain the differential expression of the *AtDTX1* transporter between *A. l. petraea* populations.

4.4.5 Hierarchical Cluster Analysis

Hierarchical cluster analysis was performed to identify populations that responded similarly to the chilling treatment across all genes studied. Iceland1 and Iceland2 were the most similar populations in terms of gene expression following cold exposure; this is likely to reflect the large sizes of the Icelandic populations and the subsequent high probability of gene flow between them. Norway6 is the most distinct population in terms of its transcriptomic responses to chilling for the genes studied. This population is the one of the smallest Norwegian populations and is located at high altitude; these factors may contribute to isolation of the population from gene flow and thus result in its distinct responses.

4.4.6 Additional qRT-PCR replicates

Two additional replicates were performed for four of the thirteen populations, and gene expression quantified through qRT-PCR. Log_2 fold change values obtained for each population differed between the three replicates, although overall trends were confirmed. This illustrates the limitations of qRT-PCR analysis; results are reproducible but exact values fluctuate between replicates. It is therefore important to perform replicates where possible, and not to over-interpret exact fold change values.

4.4.7 Conclusions

In general, results obtained for Helin and Leitrim in the initial microarray studies are confirmed in the additional thirteen *A. l. petraea* population studied through qRT-PCR. Patterns of either up-regulation or down-regulation in Helin and Leitrim are mirrored in the additional populations. For the majority of genes studied, there is variation in the level of expression between the untreated samples of all populations indicating differences in the basal levels of expression. These differences in expression suggest that populations are adapted to the conditions that prevail locally, and are likely to be linked to genetic differences between the populations. Significant metabolomic differences have previously been reported between populations of *A. l. petraea* (Davey *et al.*, 2008; Davey *et al.*, 2009), and this work demonstrates that transcriptomic differences are also evident. These differences may go some way in explaining the different extremes of environmental conditions experienced by *A. l. petraea*, as differences in basal expression levels of genes may confer differential adaptation. In this way, some populations may be better prepared for any potential cold temperature exposure than others. These populations also exhibit different extents of up- or down-regulation of gene expression following cold temperature exposure; differences may reflect the environmental conditions experienced by the populations and the consequential requirement for a greater or lesser transcriptomic change. The transcriptomic differences between study populations also emphasise the limited amounts of gene flow thought to occur between sites. Lack of gene flow has enabled expression differences to prevail and local adaptation to occur.

CHAPTER 4: SUMMARY

- Previous microarray analysis, involving populations of *A. l. petraea* from Helin in Norway and Leitrim in Ireland, identified genes that were responsive to a chilling treatment but differentially regulated between the two populations under investigation (Dr. C. J. Lilley). A number of these genes were chosen for further study in thirteen additional populations of *A. l. petraea*.
- The *A. l. petraea* sequences were obtained for the chosen genes; all coding region sequences had high percentage identities with the *A. thaliana* homolog sequences.
- Quantitative RT-PCR was used to measure the relative expression of nine genes in thirteen populations of *A. l. petraea*, in both untreated plants and in plants that had experienced a chilling treatment at 4^{o}C. Variation in expression between study populations was observed in the untreated samples, and differing extents of up- or down-regulation were evident across the populations following exposure to cold temperatures.
- The observed differences in basal levels of gene expression between the populations are indicative of adaptation to local climatic conditions. Similarly, the differing environmental conditions experienced in the field by the populations under investigation may confer differential requirements for transcriptomic change upon exposure to chilling temperatures.

5. Microarray analysis elucidates population specific transcriptomic responses of *A. l. petraea* to acclimation and freezing stresses.

AIMS

- Obtain high quality RNA from acclimation treated, acclimation and freezing treated, and untreated plants of two distinct populations of *A. l. petraea*.
- Perform microarray analysis on RNA samples through the Nottingham Arabidopsis Stock Centre (NASC) transcriptomic service using Affymetrix ATH1 GeneChip® whole genome arrays.
- Process microarray data to identify genes that are differentially expressed in response to temperature treatments and between study populations.
- Verify microarray results using qRT-PCR expression analysis for a subset of genes.

5.1 Introduction

5.1.1 Microarray technologies: the underlying principles

Microarray technology is a high throughput transcriptional profiling technique which analyses the expression of thousands of genes simultaneously. Data obtained through microarray analysis provide a 'snapshot' of gene expression under a particular set of conditions in the specific tissue or cell type under investigation (Kennedy and Wilson, 2004; Busch and Lohmann, 2007). Use of the technology has increased rapidly over the last decade, and is reflected in the substantial increase in numbers of publications involving microarray data (Barrett and Kawasaki, 2003). The popularity of microarrays is due to their relative simplicity, availability, and affordability (Brown and Botstein, 1999; Kennedy and Wilson, 2004). As whole genomes can be investigated on a single array, genes that are differentially regulated under chosen conditions can be identified in an unbiased way with no requirement for pre-selection. Schena *et al*. (1995) first demonstrated the use of microarray technology by investigating the differential expression of 45 *Arabidopsis* genes between wild type plants and transgenic lines. Since this pioneering experiment, microarrays have been widely used in plant biology to investigate, amongst others: i) the function of genes currently annotated as unknown; ii) expressional changes over a time course or developmental process; iii) responses to

environmental stimuli; and iv) conservation of fundamental processes between cultivars or closely related species (Kennedy and Wilson, 2004).

Although several different microarray platforms are commonly used, that differ in their specific technical approaches, the underlying principle for the technology is common to all. DNA probes, complementary to specific genes in the chosen study system, are immobilised on a substrate surface. Fluorescently labelled cDNA or RNA is then hybridised to these probes or array elements, resulting in a fluorescent signal for each gene represented on the array. The signal values obtained for each probe can then be used to determine the expression of the corresponding genes in the cDNA or RNA under investigation (Busch and Lohmann, 2007). Expression ratios are calculated by comparing cDNA or RNA from different samples, such as 'untreated' and 'treated', and individual genes can be subsequently described as up-regulated, down-regulated, or unchanged (Barrett and Kawasaki, 2003).

5.1.2 Types of microarrays

The two main array types are spotted arrays and photolithographic oligonucleotide arrays. Spotted arrays consist of probes, commonly PCR products or pre-synthesised oligonucleotides of 50 to 70 bases, that are attached (or spotted) directly onto a glass slide using a robotic arrayer (Rensink and Buell, 2005). A distinct advantage of these arrays is that they permit the comparison of two different RNA samples using just one array; two RNA populations are labelled with distinct fluorescent dyes prior to hybridisation to the array and, following scanning at the two appropriate wavelengths, two intensity values are obtained for each probe. These values relate to the starting proportions of the transcript in each of the RNA populations. Photolithographic arrays are produced by synthesising oligonucleotide probes on a solid substrate, using similar methods as for the production of silicon computer chips (Lockhart et al., 1996; Zhang et al., 2004). This method is highly reproducible and has consequently resulted in the commercial production of arrays, or gene chips, for a number of model organisms. The Affymetrix ATH1 GeneChip® is a whole genome array that has been developed for *Arabidopsis*, consisting of 22 747 probe sets that represent approximately 24 000 genes (Redman *et al.*, 2004). Each probe set is designed to query one transcript, but contains 22 probes of 25 bases each. Eleven of these are perfect match (PM) probes, exactly matching the *Arabidopsis* sequence for the desired transcript, whilst the remaining eleven are mismatch (MM) probes that contain a mismatch nucleotide at the central

position (Knudsen, 2004). These mismatch probes are designed to detect non-specific hybridisation (http://www.affymetrix.com). Prior to hybridisation, cDNA is synthesised from the desired RNA population and then amplified in the presence of biotin, producing biotin-labelled cRNA for hybridisation to the GeneChip®. Hybridised biotinylated cRNA is then stained with a fluorescent molecule that binds to biotin before chips are scanned to obtain signal values. The efficiency of hybridisation is determined using probe sets for prokaryotic genes; Affymetrix GeneChips® use *bio*B, *bio*C, *bio*D, and *Cre* (Redman *et al.*, 2004). The differing concentrations of each of the probe sets that are incorporated into the cRNA sample should be reflected in the signal values obtained from the arrays; if this is not the case, this is indicative of hybridisation problems. Probe sets for the control or 'housekeeping' genes, *UBIQUITIN*, *GAPDH* and *ß-ACTIN*, are also included on the Affymetrix ATH1 GeneChip® array; these contain probes specific to the 3', middle, and 5' ends of these genes. Ratios between the intensities of these probes are analysed to determine whether or not RNA has suffered degradation and to confirm that cDNA has been efficiently synthesised. One Affymetrix ATH1 GeneChip® array must be used for each RNA population under investigation; signal values can then be compared between chips to determine differentially regulated genes. It is standard practice to perform more than one replicate array for each experimental condition under investigation although there is currently no consensus on the optimum level of replication that should be used; it has been suggested that using ten replicates produces the most stable results although this is dependent on investigation-specific aims (Pavlidis *et al.*, 2003). Biological variation has been accepted to outweigh technical variation, and thus biological variation should be prioritised in the design of microarray experiments (Zhu and Wang, 2000; Meyers *et al.*, 2004). The number of replicates within a microarray experiment is commonly limited by factors such as availability of biological material and the monetary costs involved (Clarke and Zhu, 2006; Jeffery *et al.*, 2006).

5.1.3 Processing of microarray data

Data generated from microarray images must undergo several transformations and normalisations to ensure that the resultant dataset is robust, and that subsequent comparisons between individual arrays are valid (Clarke and Zhu, 2006). Arrays must be corrected for background noise; for Affymetrix chips, background noise is commonly estimated by averaging the signal values of the lowest two percent of probe sets within a specified region of the image. This value is then subtracted from all other

probe sets in that region to give background corrected signal values (Knudsen, 2004). Normalisation is the process of removing systematic variability such as differences in the starting quantities of RNA, differences in labelling or detection efficiencies, and variability in hybridisation conditions between arrays (Yang *et al.*, 2002; Hatfield *et al.*, 2003; Zhang *et al.*, 2004a). There are a range of approaches that can be used to normalise microarray data. Global normalisation involves the subtraction of a constant value, for example the mean or median signal value, from all signal intensities (Ye and Day, 2003). Housekeeping genes, designated as such because their expression levels remain constant under varying experimental conditions, can also be used for global normalisation; signal intensities for a given gene are multiplied by a constant until intensities are equal across the arrays that are to be compared (Knudsen, 2004). The development of spike controls has added another possible approach for global normalisation. These involve the addition of a known concentration of a known transcript to the experimental mRNA sample; this is assumed to be amplified and labelled in the same way as all other transcripts and a unique probe is used for its detection on the array. The spike signals can consequently be used for normalisation as their expression can be assumed to be invariant (Kreil and Russell, 2005). The field of microarray data normalisation is constantly evolving, and the chosen method must be appropriate to the experiment in question (Kriel and Russell, 2005). Normalisation of data from experiments performed using GeneChips requires particular consideration; each array must be normalised against all others in the experiment as only one biological sample is hybridised to each array, in contrast to spotted arrays where a control and treated sample are hybridised to the same array substrate (Irizarry *et al.*, 2003). Quantile normalisation aims to ensure that all arrays within a set of arrays exhibit the same distribution of probe intensities, and this has been cited as the most favourable method of normalisation for GeneChip datasets (Bolstad *et al.*, 2003). The most effective way to analyse such datasets is to use a method that incorporates background correction, normalisation, and summarisation, and several packages for this have been developed. MAS 5.0, a statistical algorithm by Affymetrix, calculates a robust average of PM probe values minus MM probe values to give single expression values for each probe set; this technique does introduce noise into the dataset at low signal intensity levels, caused when PM probes have lower signal values than MM probes. The RMA method (Robust Multi-array Average) succeeds in decreasing the variation in low abundance transcripts by ignoring MM values completely (Irizarry *et al.*, 2003). During RMA normalisation, background correction is performed for each

array separately followed by quantile normalisation across all arrays to correct for array biases. Finally, summarisation is performed where single signal values are created for each probe set on each array using a technique called median polishing. This method has since been further developed to incorporate the MM values in GCRMA normalisation; MM data are adjusted according to probe affinity prior to subtraction from PM data (Wu and Irizarry, 2004). GCRMA normalisation results in less variation at lower signal values and consequently gives more consistent fold change estimates. Normalisation software packages are available at Bioconductor, an open source software project for bioinformatics (http://www.bioconductor.org). Following global normalisation it is common for microarray datasets to be converted to a log scale to minimise the range of signal intensities (Kreil and Russell, 2005). These datasets can then be used to detect differentially regulated genes; such genes are those that display significant differences in expression levels between the conditions under investigation. Gene lists, or sets, are created using criteria such as fold change, statistical significance, and expression values; it is advised to use a combination of measures to identify differentially regulated genes (Clarke and Zhu, 2006).

The popularity of microarray experiments has consequently produced a wealth of data concerning the transcriptomic responses of a range of organisms in a relatively short period of time. This, however, does not quickly translate to an understanding of gene function; the context within which expression values were measured is critical to interpretation of results (Brazma *et al.*, 2000). Ideally, to use microarray datasets to their full potential, integration of data gathered in different laboratories should be possible; a lack of standardisation for methods such as normalisation and even expression units across the field does not aid such comparisons (Brazma *et al.*, 2001). Attempts have been made to establish standards for reporting microarray experiments; MIAME (Minimum information about a microarray experiment) specifies that arrays should be comprehensively annotated to place experiments in context (Brazma *et al.*, 2001). A number of sections for annotation are suggested, along with controlled vocabularies; these include experiment design as a whole, procedures and parameters for hybridisation, and the normalisation type and specification used (Brazma *et al.*, 2001). Compliance with the MIAME standards is now required by many scientific journals and public databases and, subsequently, comparisons across individual array experiments are more viable. All data generated through NASC's (Nottingham Arabidopsis Stock Centre) transcriptomic service are publically available in the

NASCArrays database (accessed at http://affymetrix.arabidopsis.info), and this database aims to be fully MIAME supportive (Craigon *et al.*, 2004). GENEVESTIGATOR is an additional database specifically concerning *Arabidopsis* GeneChip data; gene expression profiles can be obtained for over 22 000 genes in relation to environmental conditions, growth stages, or particular organs (http://www.genevestigator.ethz.ch). Data mining tools can also be used to identify genes responsive to specific conditions or expressed in particular tissue types, thus directing future functional studies (Zimmermann *et al.*, 2004; Hruz *et al.*, 2008). The development of standards for the reporting of microarray data, and of publically available databases for such data, have dramatically improved the field of microarrays, and this increased connectivity has aided identification and annotation of genes involved in key biological processes.

5.1.4 Use of microarrays for the study of *A. thaliana* close relatives

The original Affymetrix AG GeneChip, containing over 8000 probes, has been utilised in the study of a wide range of processes in *A. thaliana* (Hennig *et al.*, 2003), including stress responses (Kreps *et al.*, 2002). Following development of the Affymetrix ATH1 GeneChips®, with probes representing a far greater extent of the genome, the stress response in *A. thaliana* has been studied in much more detail. More recently, the use of Affymetrix ATH1 GeneChips® has been extended to investigation of the transcriptome in species closely related to *A. thaliana*. Published studies involve the species *A. halleri* (Becher *et al.*, 2004; Weber *et al.*, 2004) and *Thlaspi caerulescens* (Hammond *et al.*, 2006) amongst others, and have investigated transcriptomic responses to a range of environmental stresses. Often these study species have been chosen due to their adaptive traits; interestingly, orthologues of genes induced by stress in *A. thaliana* have regularly exhibited higher expression levels in the adapted species, particularly in the absence of the stressor (Van De Mortel and Aarts, 2006) Such studies will be of value in identifying genes generally involved in stress responses, and genes with specific roles in adaptation to environment (Mitchell-Olds, 2001).

5.1.5 Effects of acclimation, freezing, and population on the transcriptome

Cold temperatures stimulate a wide range of physiological and biochemical changes within plants including accumulation of sugars and starch, and altered membrane behaviour (Ristic and Ashworth, 1993), and calcium ion fluxes (Polisensky and Braam, 1996). Cold acclimation is the process by which plants are able to increase their ability

to withstand freezing temperatures through exposure to low, non-freezing temperatures (Hughes and Dunn, 1996; Thomashow, 1998, 1999). Cold acclimation and resultant freezing tolerance have been studied extensively in *A. thaliana*, and many cold responsive genes and regulatory pathways have been identified as discussed in section 1.5.1. Although acclimation-induced freezing tolerance and associated transcriptomic changes have been studied in detail, the study of transcriptomic responses upon exposure to freezing temperatures is limited (Livingston III *et al.*, 2007). Exposure to freezing temperatures results in freezing damage in plants; this is primarily caused by cellular dehydration, although disruption to tissue and organ structure by ice crystal formation can also contribute (Pearce, 1999). It is feasible to expect additional transcriptomic changes to occur upon exposure to freezing temperatures as a wide range of plant processes, such as photosynthesis, will be adversely affected by this stress. Herman *et al.* (2006) demonstrated that exposure of cold-acclimated wheat plants to subzero temperatures (-3^{o}C) resulted in additional freezing tolerance, coupled with significant transcriptomic and proteomic changes. Large scale transcriptomic changes included up-regulation of genes related to cell wall modification and senescence, whilst photosynthetic genes were down-regulated (Herman *et al.*, 2006). Subsequently, gene expression studies in wheat plants exposed to potentially damaging freezing temperatures (-10^{o}C and -12^{o}C) have revealed extensive transcriptomic changes, thus suggesting that a response mechanism exists to further modulate the transcriptome at such temperatures (Skinner, 2009).

Evidence for local adaptation in different ecotypes of *A. thaliana* is outlined in Chapter 4; ecotypes experience differing climatic conditions in the field and subsequently exhibit differences in ability to acclimate (Hannah *et al.*, 2006), a latitudinal cline in freezing tolerance (Hannah *et al*., 2006; Zhen and Ungerer, 2008a), and differential expression of genes in response to cold stress (He *et al.*, 2008; McKhann *et al.*, 2008). *A. l. petraea* persists in isolated populations throughout Europe, in direct contrast to the continuous distribution of *A. thaliana* (Clauss and Mitchell-Olds, 2006; Davey *et al.*, 2008). The differing climatic conditions experienced by *A. l. petraea* populations may exert selective pressures that result in local adaptation, and the discontinuous nature of these populations may act to enhance this effect by preventing gene flow between them. The ability to acclimate and to survive exposure to freezing temperatures may consequently differ between populations, and, as this is genetically controlled, differences are likely to be reflected in expression levels of genes.

Here, Affymetrix ATH1 GeneChip® arrays are used to identify the transcriptomic responses in two distinct populations of *A. l. petraea* to acclimation, and in plants that have experienced an additional freezing stress following acclimation. A key objective of this work was to identify a set of genes with potentially important roles in the acclimation response of *A. l. petraea*, and to ascertain whether or not these genes are also relevant to the cold response of *A. thaliana*. Also, investigation of the transcriptome following exposure of plants to sub-zero temperatures has enabled genes with potential roles in the freezing response to be identified. As two independent populations were included in this investigation, differences in gene expression between the two populations, either in control conditions or in response to the temperature treatments, were also identified. Differential gene expression under the same environmental conditions may be indicative of adaptation to environment within this species. This gene set therefore provides insight into the biological function of genes involved in adaptation to environment, and serves to assist in the selection of genes for future studies concerning local adaptation.

5.2 Materials and Methods

5.2.1 Plant growth

Work was carried out in parallel with Dr. Matt Davey. Plants were grown at the University of Sheffield by Dr. Matt Davey for metabolomic studies; material from these same plants was used in these microarray experiments. Plants were established under a 20^{o}C day, 15^{o}C night regime, with a 12 h day, 12 h night cycle. Plants from both Helin (Norway7) and Leitrim (Ireland1) populations were grown; these populations were chosen due to their contrasting environmental conditions. Twenty plants per population per treatment were grown, with a total of three treatments. RNA from 18 plants per population per treatment was used in the microarray analysis; only those plants with the highest quantity and quality of extracted RNA were included.

5.2.2 Treatment conditions

Davey (*pers. comm.*) established that acclimation time, duration of cold shock, and freezing temperature all affect chlorophyll fluorescence measurements, plant survival, and electrolyte leakage, and found that a cold shock of -9^{o}C was the temperature that provided the largest difference in fluorescence and survival between the two study populations. A freezing temperature of between -7^{o}C and -9^{o}C provided the largest difference in electrolyte leakage between the two populations thus a freezing temperature of -9^{o}C was chosen for this experiment (Dr. M. Davey, unpublished data). These parameters were selected for and used in the investigation of both metabolomic (Dr. M Davey) and transcriptomic responses in *A. l. petraea*.

Following establishment, plants were subjected to control conditions, an acclimation period at 2^{o}C of two days, or an acclimation period at 2^{o}C of two days followed by a cold shock at -9^{o}C. The period used for the freezing shock was a ramp from +2^{o}C to -9^{o}C over five hours 40 minutes. The different treatments are detailed in Table 5.1.

5.2.3 Tissue collection and RNA extraction

Leaf tissue was harvested during the night cycle and flash frozen in liquid nitrogen. Plants receiving a cold shock of -9^{o}C were harvested upon reaching the freezing temperature. Tissue samples were stored at -80^{o}C until required. RNA was extracted from leaf material collected from all 20 plants from each of the treatment conditions

Treatment	Acclimation		Freezing shock	Population
	Temperature	Time	Temperature	
1 – Control 20 °C	X	X	X	Helin
2 – Control 20 °C	X	X	X	Leitrim
3	2 °C	2 days	X	Helin
4	2 °C	2 days	X	Leitrim
5	2 °C	2 days	-9 °C	Helin
6	2 °C	2 days	-9 °C	Leitrim

Table 5.1 Treatment conditions used in the Affymetrix ATH1GeneChip® arrays. This table details the treatment conditions experienced by plants from both Helin and Leitrim populations of *A. l. petraea* prior to RNA extraction and subsequent microarray analysis.

(Table 5.1) as detailed in 2.3.4. Nucleic acid concentration was estimated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies) as 2.3.5. Repeat RNA extraction was carried out if the concentration was less than 100 ng μl^{-1}.

5.2.4 Pooling of samples for microarray analysis

For each treatment, the microarray analysis was carried out in triplicate using pooled RNA from six individual plants for each replicate. Five micrograms of RNA from six individual plants for each treatment was combined to create a pooled sample containing 30 µg of RNA. Pooling of samples was performed to enable a larger number of individual plants to be included in the study. The pooled samples (three per treatment per population) were checked for quality using a Bioanalyser (section 2.3.6), and then precipitated prior to microarray analysis (section 5.2.5).

5.2.5 ATH1 Arrays

Hybridisation of Biotin-labelled cRNA to Affymetrix Arabidopsis ATH1 GeneChip® arrays and array scanning were performed by the Nottingham Arabidopsis Stock Centre transcriptomic service (Craigon *et al.*, 2004) following the standard Affymetrix protocol.

5.2.6 Microarray data analysis

Data analysis was carried out in collaboration with Dr Nathan Watson-Haigh. The Affymetrix CEL files containing the raw probe intensity values were used for analyses performed in R using open source software available at Bioconductor (http://www.bioconductor.org). Pre-normalisation quality control included generation of pseudochip images to analyse chip hybridization quality and assessment of sample quality by examination of 3' to 5' intensity ratios of ß-actin and GAPDH control genes. Data were subjected to GCRMA background correction and quantile normalisation across all arrays. Genefilter was used to remove genes with very low expression across all arrays from the analysis and the limma package was used to create a linear model and extract differentially expressed genes of interest. The Benjamini and Hochberg multiple test correction was employed (Benjamini and Hochberg, 1995). Genes were designated as responsive to a treatment if they had expression changes, either positive or negative, of greater than two fold (>1 fold on a $\log_2$ scale) and had significance (P values) of less than 0.05. Over- and under-representation of GO Slim biological process categories was determined using the TAIR website

(www.arabidopsis.org/tools/bulk/go/index.jsp). Lists of differentially expressed gene locus IDs were uploaded separately to the website, and gene ontology data retrieved. Enrichment or depletion of GO Slim categories was determined by chi-squared tests of comparison in Excel. GeneSpring v7.3 was used to investigate more specific Gene Ontology (GO) classifications of the differentially expressed genes and to identify metabolic pathways that were over-represented in each list of differentially expressed genes.

5.2.7 Verification of microarray results through qRT-PCR

Pooled RNA was used for first-strand cDNA synthesis with reverse transcriptase (SuperScript II, Invitrogen) according to the manufacturer's instructions (section 2.3.7). The three original RNA pools per treatment used for the microarrays (section 5.2.4) were combined to create one pool per treatment containing a total of 1 μg RNA from which cDNA was synthesised. This pooling was performed to minimise costs associated with qRT-PCR.

Seventeen genes present on the arrays were chosen for verification of the microarray data through quantitative reverse-transcriptase PCR (qRT-PCR). Sequences of *Arabidopsis thaliana* genes were obtained from sequence viewer (SeqViewer) at TAIR (http://www.arabidopsis.org). Primers for qRT-PCR were designed using Primer3 software (http://frodo.wi.mit.edu/) to the following criteria: TM of $60^{o}C \pm 1^{o}C$, to amplify fragments 90 – 160 bp in length, primer sequences of ~20 nucleotides and guanine-cytosine content of ~50%. Primers were designed to amplify close to the 3' end of the coding sequence and, where possible, spanned an intron. The efficiency of all primer pairs was determined using a 10x dilution series of cDNA template, and the specificity of primers was determined through dissociation curve analysis The expression levels of target genes were normalised to the housekeeping gene PDF2, At1g13320; the primers for this gene were based on those described by Czechowski *et al*. (2005) but were specific to the *A. l. petraea* sequence (section 4.2.6). A full list of genes and primer sequences is detailed in Table 5.2.

Gene	TAIR Annotation	Forward primer sequence 5'-3'	Reverse primer sequence 5'-3'	Product size (bp)	Efficiency (%)
At1g13320	PDF2	TAA-CGT-CGC-CAA-AAT-GAT-GC	GTT-CTC-CAC-AAC-CGC-TTG-GT	61	103
At1g08920	Sugar transporter, putative	CCT-TTG-GTT-GGA-TTG-TTG-CT	TCG-GGT-ACC-ATC-GCA-TAA-AT	124	91.5
At1g20440	Cor47	ACA-AGC-CTA-GTG-TCA-TCG-AAA-AGC	TCT-TCA-TCG-CTC-GAA-GAG-GAA-G	67	102
At1g23205	Invertase/pectin methylesterase inhibitor family protein	CAG-ATG-AGG-GAG-CTG-GTC-TC	GCG-CAT-ATG-TCA-TCC-TTC-CT	158	101.8
At1g24070	ATCSLA10_CSLA10; transferase, transferring glycosyl groups	TGT-CAT-TGC-GAC-ACC-AAG-AT	TCT-CAG-TGA-CAA-CCC-ATT-CG	140	100.6
At1g49720	ABF1: DNA binding / transcription activator/ transcription factor	TCG-GGG-AAG-AAG-AAG-CAA-TA	CTC-AGC-TTC-CAG-TTC-CAA-GG	133	102.7
At1g54820	Protein kinase family protein	ACT-CTT-AAC-GGG-TCG-CAC-AC	GGT-TGG-ATC-CAC-CAT-TTC-AC	121	100.6
At1g74670	Gibberellin-responsive protein, putative	ATG-CAT-GTT-CTT-CTG-CCA-AA	TGG-ACA-TTT-TGG-TCC-ACC-TT	127	101.7
At1g77120	ADH1 (Alcohol dehydrogenase 1)	GAA-TCG-CTG-GTG-CTT-CTA-GG	CTC-AGC-GAT-CAC-CTG-TTG-AA	134	110
At1g79270	ECT8 (evolutionarily conserved C-terminal region 8)	CAA-TTC-TGC-GGT-GTA-GCT-GA	GTT-GCG-GAT-TTG-GAA-CAT-CT	133	95.5
At2g34810	FAD-binding domain-containing protein	CTT-GGC-GAT-TGC-AAG-AAG-TT	GCT-CCT-TCA-TAC-GTC-GCA-TT	135	106.2
At2g42540	Cor15a	AAC-GAG-GCC-ACA-AAG-AAA-GC	CAG-CTT-CTT-TAC-CCA-ATG-TAT-CTG-C	132	103.3
At3g05500	Rubber elongation factor (REF) family protein	TGT-CGT-TGG-TCC-TGT-CTA-CG	GCA-GAC-ACT-TGC-TTG-ACG-AC	132	107.4
At3g05890	RCI2B (Rare cold-inducible 2)	TGA-GTA-CAG-CCA-CTT-TCG-TAG-AG	TCC-GGG-AAG-ATA-ACC-AAA-CA	131	98.4
At3g48360	BT2; protein binding / transcription factor	GCA-ATG-GAG-TGC-ATA-GAG-CA	AAC-GCC-TTA-CAA-GGC-TCT-GA	119	97.5
At3g50530	CRK (CDPK-related kinase); calcium ion binding	GAT-GAA-CTC-GCT-TCG-GAA-CT	GCG-GCT-AGA-CAC-ACC-ATG-TA	135	90
At3g55610	P5CS2	AAG-TGG-ACA-GTG-CTG-CTG-TTT	CTT-GTG-CTT-ATT-CCC-ACC-TCA	94	101.1
At5g15970	Kin2	CAA-AGC-TGA-GGA-GAA-GAG-CAA	GCC-GCA-TCC-GAT-ATA-CTC-TTT	105	99.3
At5g61160	AACT1; transferase	GTT-TGG-CTT-CAC-AAG-GGG-TA	CCG-AAG-TCC-GTT-TCG-TAC-AT	130	108.5

Table 5.2 Primers used in qRT-PCR verification of microarray results. Primers are designed to the *A. thaliana* sequences obtained from TAIR, with the exception of the PDF2 primers that are designed to the *A. l. petraea* sequence as described in section 4.2.6. All primer pairs had efficiencies of > 90% and < 110%, based on results from a 10x dilution series.

5.3 Results obtained from microarray analysis

5.3.1 Quality of RNA samples

All RNA extracted from individual leaves was ≥100 ng μl^{-1} when quantified as detailed in section 5.2.3. Following the pooling of samples, as described in section 5.2.4, all samples were found to be of a suitable quality when checked on the Agilent 2100 Bioanalyser (Agilent Technologies) (section 2.3.6). This output is shown in Figure 5.1. All samples have clearly defined 18S and 28S peaks with little noise between these peaks. The additional peaks, before the 18S peak, are due to the presence of chloroplasts in the leaf tissue from which RNA was extracted.

5.3.2 Microarray quality control

Quality control measures were performed on the raw data files prior to analysis to ensure no degradation of samples had occurred and that no chips showed sub-optimal hybridisation. Positive and negative border elements (probes located around array edges) were analysed and hybridisation was found to be uniform around the edges of all the arrays. The Affymetrix ATH1GeneChip® arrays contain control spots for *GAPDH*, *ß-ACTIN,* and *UBIQUITIN* with probes that are specific to the 3', middle, and 5' regions of these genes. Ratios between the intensities of these probes were analysed to confirm that RNA samples were not degraded and that cDNA was successfully synthesised. Sample replication is also important; biological replicate arrays for both control and treated samples clustered into distinct groups. Similar quality control measures were performed on the data following GCRMA background correction and normalisation, and all arrays were confirmed to be of sufficient quality.

5.3.3 Identification of differentially expressed genes

The GCRMA package at Bioconductor (http://www.bioconductor.org) was used to identify genes that were differentially regulated following the acclimation, and the acclimation followed by freezing treatments, as well as those differentially regulated between the two study populations. A total of 4663 genes were found to have significantly altered expression within this investigation ($P < 0.05$). A Venn diagram illustrates overlapping and non-overlapping gene expression in the different treatments (Figure 5.2). This shows that 2796 (60%) genes exhibited a significant change in expression in response to the acclimation treatment alone, whilst only 44 (0.94%) genes significantly changed in response to the additional freezing treatment alone. The

regions of overlap, for example between 'acclimation' and 'population', represent genes that both responded to the temperature treatment and were differentially expressed between the two populations (sections 5.3.3.3 and 5.3.3.4). Gene sets were then filtered based on the extent of the fold change exhibited by each gene; Figure 5.3 illustrates the numbers of genes responding to each treatment ($P < 0.05$) that had fold changes of two or more (>1 fold on a $\log_2$ scale).

5.3.3.1 Genes responsive to the acclimation treatment

The acclimation treatment resulted in the largest number of significant changes in gene expression (3742 genes). These genes either exhibited an up-regulation or a down-regulation in expression levels following exposure to the acclimation conditions (2^oC for 2 days). A list was produced containing those genes for which, in addition to a significant change in expression ($P < 0.05$), the extent of the fold change in expression was greater than two (1651 genes as illustrated in Figure 5.3); this list was used in all subsequent analyses. This list was then split into two; the acclimation up-regulated list containing those genes with positive fold changes (799 genes: Appendix 2a), and the acclimation down-regulated list containing those genes with negative fold changes (852 genes: Appendix 2b). The top twenty genes (i.e. those exhibiting the greatest fold change in expression) for both of these lists are detailed in Table 5.3. The acclimation up-regulated list contains several genes annotated as cold responsive including At5g52310 (*COR78*), as well as several transferases and binding proteins. The two most up-regulated genes, with $\log_2$ fold changes greater than seven are At1g22770 (*GIGANTEA*, binding) and At4g14690 (*ELIP2*, chlorophyll binding). Eighty-seven percent of the top 100 acclimation induced genes have been reported as cold responsive (> 2 fold up-regulated) in previous microarray studies (GENEVESTIGATOR). The acclimation down-regulated list includes hydrolases, transferases, and transcription factors. The two most down-regulated genes both have $\log_2$ fold changes greater than minus five; At4g37800 (Xyloglucosyl transferase) and At1g74670 (gibberellin-responsive protein). Seventy-four percent of the top 100 genes down-regulated following acclimation have previously been reported as cold responsive (> 2 fold down-regulated).

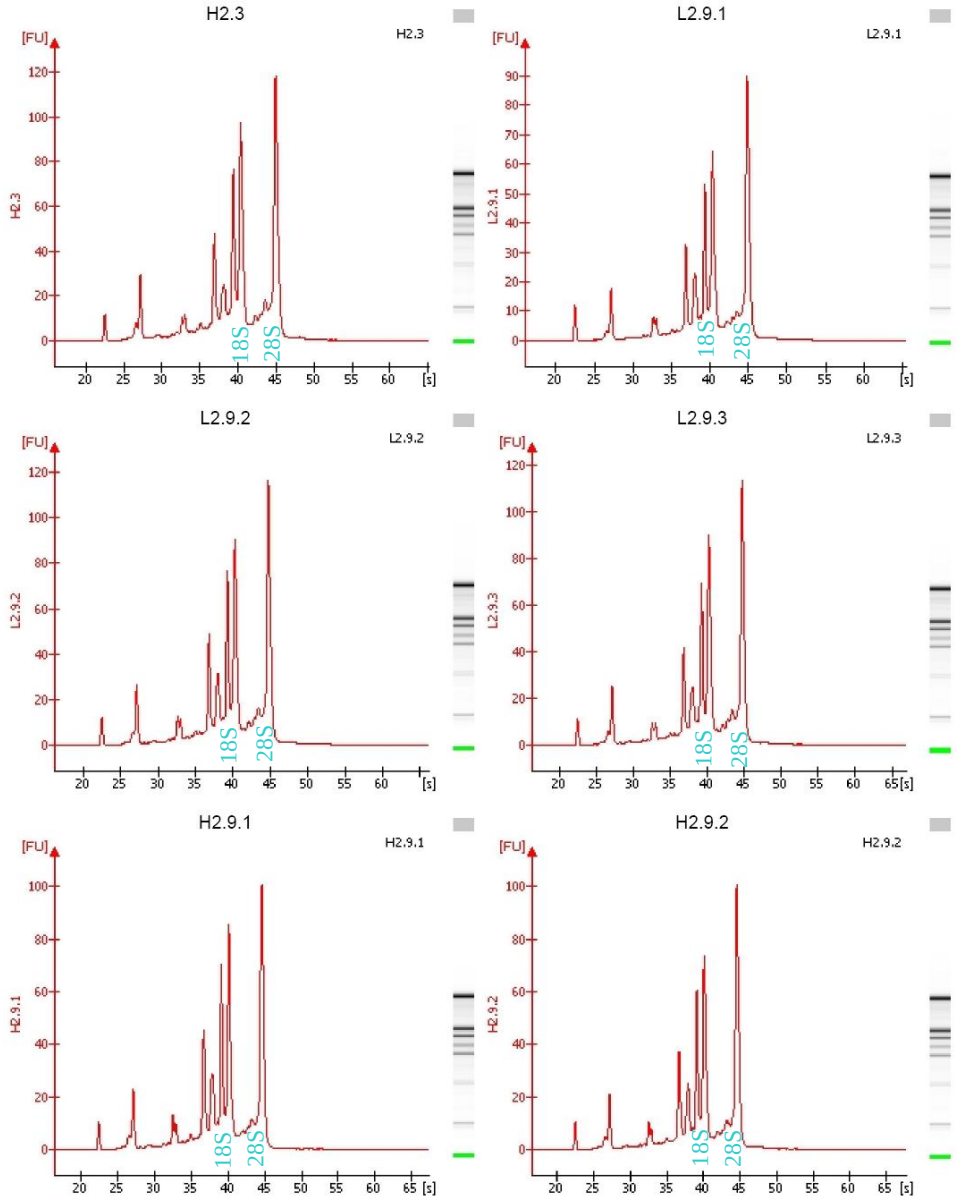


Figure 5.1 Representative Agilent Bioanalyser electropherograms for RNA samples pooled prior to microarray analysis. All RNA samples extracted from individual leaves were pooled as described in section 5.2.4. These pooled samples were analysed on the Agilent 2100 Bioanalyser (Agilent Technologies). Output for a subset of the samples is shown here. A-C: Leitrim two day acclimation samples, D-F: Helin two day acclimation samples. All samples had clearly defined 18S and 28S peaks with little noise between these peaks. The additional peaks, before the 18S peak, are due to the presence of chloroplasts in the leaf tissue from which RNA was extracted.

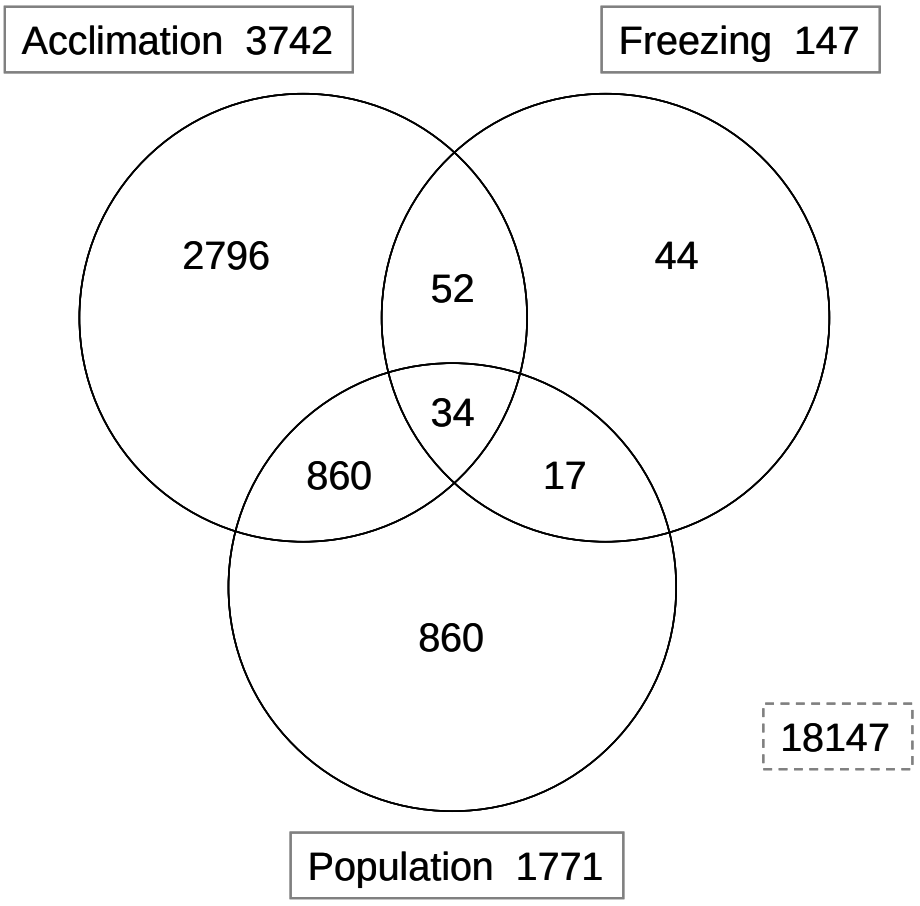


Figure 5.2. Venn diagram summarising numbers of genes with significantly altered expression following each treatment. Results from analysis of Affymetrix ATH1 GeneChip® arrays for *A. l petraea* following acclimation and freezing treatments. Two populations were used in this study: Helin in Norway and Leitrim in Ireland. Values in rectangles indicate the total number of genes affected by each treatment. The value in the hashed rectangle indicates the number of probe sets for which no significant change in expression was observed. Regions of overlap represent those genes that were responsive to more than one treatment.

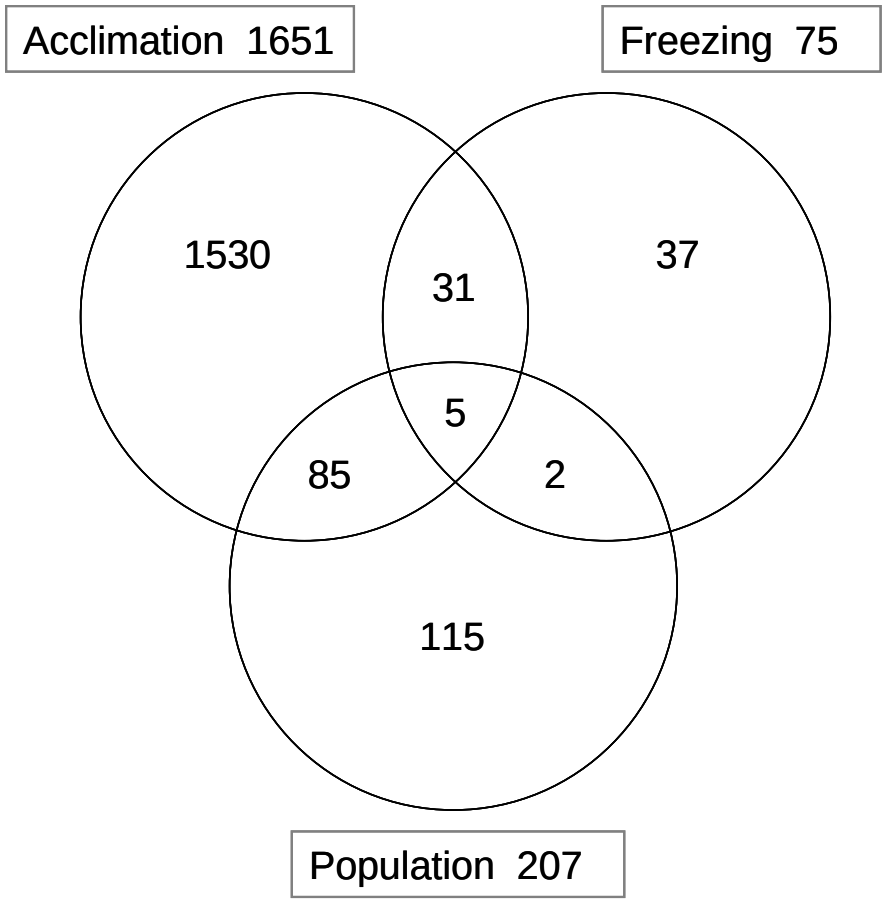


Figure 5.3. Venn diagram summarising numbers of genes exhibiting a significant >2 fold change in expression following each treatment. Results from analysis of Affymetrix ATH1 GeneChip® arrays for *A. l. petraea* following acclimation and freezing treatments; only genes that exhibited a fold change in expression of >2 and had a significance P value of <0.05 are included in this figure. Two populations were used in this study: Helin in Norway and Leitrim in Ireland. Values in rectangles indicate the total number of genes affected by each treatment. Regions of overlap represent those genes that were responsive to more than one treatment.

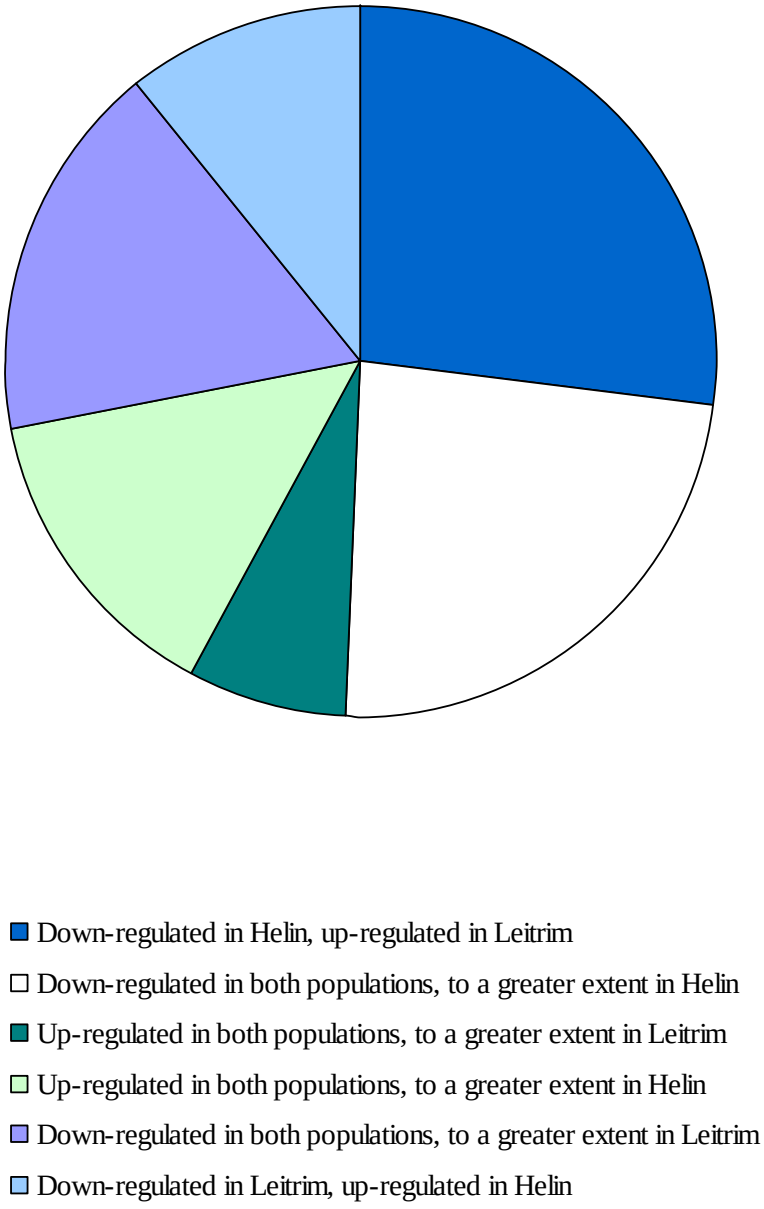


Figure 5.4 The different types of interaction gene response. All genes that were responsive to acclimation and differentially regulated between the two study populations (470 genes) were divided into categories depending upon the nature of the interaction. Average signal values from the arrays were used to determine the interaction type. These categories are detailed in the key. The two study populations were Helin in Norway and Leitrim in Ireland.

Affymetrix Probe ID	Gene ID	Annotation	Fold change (Log_2)
264211_at	At1g22770	Gigantea, binding	7.37
245306_at	At4g14690	ELIP2, chlorophyll binding	7.15
264511_at	At1g09350	ATGOLS3, transferase	6.74
253416_at	At4g33070	Pyruvate decarboxylase, putative	6.57
247095_at	At5g66400	RAB18, responsive to ABA 18	6.35
252102_at	At3g50970	Low temperature induced 30 (LTI30), XERO2	6.13
264612_at	At1g04560	AWPM-19-like membrane family protein	5.85
245627_at	At1g56600	ATGOLS2, transferase	5.84
247851_at	At5g58070	Lipocalin, putative	5.82
258979_at	At3g09440	Heat shock cognate 70 kDA protein	5.74
248337_at	At5g52310	COR78	5.55
255521_at	At4g02280	SUS3, sucrose synthase/transferase	5.53
266462_at	At2g47770	Benzodiazepine receptor-related	5.44
262644_at	At1g62710	BETA-VPE	5.30
258321_at	At3g22840	ELIP1, chlorophyll binding	5.19
266578_at	At2g23910	Cinnamoyl-CoA reductase-related	5.09
256245_at	At3g12580	HSP70, ATP binding	4.97
264147_at	At1g02205	CER1	4.65
254952_at	At4g10955 At4g10960	Lipase class 3 family protein	4.65
253627_at	At4g30650	Putative / low temperature and salt responsive protein	4.64
253040_at	At4g37800	Xyloglucosyl transferase	-5.14
260221_at	At1g74670	Gibberellin-responsive protein	-5.07
266123_at	At2g45180	Protease inhibitor / seed storage / LTP family protein	-4.93
247266_at	At5g64570 At5g64572	XYL4, hydrolase	-4.92
258054_at	At3g16240	Delta tonoplast integral protein	-4.79
265066_at	At1g03870	FLA9	-4.35
259788_at	At1g29670	GDSL-motif lipase / hydrolase family protein	-4.32
248566_s_at	At5g49730 At5g49740	FRO6, ferric-chelate reductase / oxidoreductase	-4.27
266790_at	At2g28950	ATEXPA6	-4.07
259272_at	At3g01290	Band 7 family protein	-4.00
266899_at	At2g34620	mTERF-related	-3.82
248186_at	At5g53880	Similar to unknown (GB:ABK92577.1)	-3.71
266215_at	At2g06850	EXGT-A1. hydrolase, acting on glycosyl bonds	-3.70
256772_at	At3g13750	BGAL1, beta-galactosidase	-3.69
247162_at	At5g65730	Xyloglucosyl transferase	-3.67
246487_at	At5g16030	Similar to unknown protein (At3g02500.1)	-3.63
252367_at	At3g48360	BT2, protein binding / transcription factor	-3.61
265400_at	At2g10940	Protease inhibitor / seed storage / LTP family protein	-3.52
246011_at	At5g08330	TCP family transcription factor, putative	-3.49
255786_at	At1g19670	ATCLH1	-3.39

Table 5.3 Top twenty up- and down-regulated genes in *A. l. petraea* following two day 2°C acclimation treatment. This table contains the genes that exhibited the greatest log_2 fold changes in expression ($P < 0.05$). Positive fold changes refer to up-regulated genes and negative fold changes to down-regulated genes.

Affymetrix Probe ID	Gene ID		Fold change (Log_2)
253643_at	At4g29780	Similar to unknown protein (TAIR: At5g12010.1)	3.52
266545_at	At2g35290	Similar to unnamed protein product (GB:CA063442.1)	3.15
247925_at	At5g57560	TCH4, hydrolase, acting on glycosyl bonds / xyloglucan:xyloglucosyl transferase	3.06
260227_at	At1g74450	Similar to unknown protein (TAIR: At1g18740.1)	3.05
255064_at	At4g08950	Phosphate responsive protein, putative	2.87
261648_at	At1g27730	STZ ZAT10, nucleic acid binding / transcription factor / zinc ion binding	2.77
266834_s_at	At2g05050 At2g30020 At3g27140 At4g08260	Multiple genes represented by probe set.	2.67
261405_at	At1g18740	Similar to unknown protein (TAIR: At1g74450.1)	2.56
256633_at	At3g28340	GATL10, transferase, transferring hexosyl groups	2.48
262307_at	At1g71000	DNAJ heat shock N-terminal domain-containing protein	2.37
254120_at	At4g24570	Mitochondrial substrate carrier family protein	2.33
257654_at	At3g13310	DNAJ heat shock N-terminal domain-containing protein	2.26
252474_at	At3g46620	Zinc finger family protein	2.20
248389_at	At5g51990	CBF4, DREB1D, DNA binding, transcription activator, transcription factor	2.10
253859_at	At4g27657	Similar to unknown protein (TAIR:At4g27652.1)	2.09
253830_at	At4g27652	Similar to unknown protein (TAIR:At4g27657.1)	2.07
265561_s_at	At2g05440 At2g05510	Glycine-rich protein	2.02
252679_at	At3g44260	CCR4_NOT transcription complex protein, putative	1.95
261470_at	At1g28370	ATERF11, ERF11, DNA binding, transcription factor, transcription repressor	1.91
253915_at	At4g27280	Calcium-binding EF hand family protein	1.88
246935_at	At5g25350	EBF2	-2.25
261263_at	At1g26790	Dof-type zinc finger domain-containing protein	-2.24
263549_at	At2g21650	MEE3, DNA binding, transcription factor	-1.81
266363_at	At2g41250	Haloacid dehalogenase-like hydrolase family protein	-1.58
260556_at	At2g43620	Chitinase, putative	-1.47
261958_at	At1g64500	Glutaredoxin family protein	-1.31
252972_at	At4g38840	Auxin-responsive protein, putative	-1.24
267230_at	At2g44080	ARL (ARGOS-LIKE)	-1.09
244973_at	AtCg00690	PSBT, encodes photosystem II 4 kD protein subuni PSII-T	-1.01
265943_at	At2g19570	CDA1, DESZ	-1.00

Table 5.4 Top up- and down-regulated genes in *A. l. petraea* following additional -9°C freezing shock. This table contains the genes that exhibited the greatest log_2 fold changes in expression ($P < 0.05$). Positive fold changes refer to up-regulated genes and negative fold changes to down-regulated genes. The top twenty up-regulated genes are shown. There were only a total of ten genes that were significantly down-regulated ($P < 0.05$) and had fold changes of >2.

Affymetrix Probe ID	Gene ID	Annotation	Fold change (Log_2)
251241_s_at	At3g62460 At3g62530	Similar to PBS lysase HEAT-like repeat-containing protein.	4.98
262399_at	At1g49500	Similar to unknown protein (TAIR:At3g19030.1)	3.10
246603_at	At1g31690	Copper ion binding	2.85
245757_at	At1g35140	PHL-1	2.74
252367_at	At3g48360	BT2, protein binding, transcription factor, transcription regulator	2.56
250228_at	At5g13840	WD-40 repeat family protein	2.46
260283_at	At1g80480	PTAC17	2.26
266142_at	At2g39030	GCN5-related N-acetyltransferase family protein	2.19
266566_at	At2g24040	Hydrophobic protein, putative, low temperature and salt responsive protein	1.96
255479_at	At4g02380	ATLEA5, SAG21	1.93
261410_at	At1g07610	MT1C	1.85
254804_at	At4g13010	Oxireductase, zinc-binding dehydrogenase family protein.	1.82
244966_at	AtCg00600	PETG, cytochrome b6-f complex, subunit V.	1.81
263046_at	At2g05380	GRP3S	1.80
260568_at	At2g43570	Chitinase, putative	1.77
245611_at	At4g14390	Ankyrin repeat family protein	1.65
245442_at	At4g16710	Glycosyltransferasr family protein 28	1.65
248191_at	At5g54130	Calcium-binding EF hand family protein	1.64
252450_s_at	At3g47090 At3g47580	Leucine-rich repeat transmembrane protein kinase, putative	1.57
249125_at	At5g43450	2-oxoglutarate-dependent dioxygenase, putative	1.55
248197_at	At5g54190	PORA, oxidoreductase/protochlorophyllide	-3.93
265111_at	At1g62510	Protease inhibitor/seed storage/lipid transfer protein (LTP)	-3.18
259799_at	At1g72290	Trypsin and protease inhibitor family protein	-2.89
250558_at	At5g07990	TT7, CYP75B1, flavonoid 3'-monooxygenase/oxygen binding	-2.65
259431_at	At1g01620	Plasma membrane intrinsic protein 1;3	-2.64
258224_at	At3g15670	Late embryogenesis abundant protein	-2.34
255080_at	At4g09030	AGP10	-2.30
253879_s_at	At4g27560 At4g27570	Glycosyltransferase family protein	-2.26
245307_at	At4g16770	Oxidoreductase, 2OG_Fe(II) oxygenase family protein	-2.24
264672_at	At1g09750	Chloroplast nucleoid DNA-binding protein-related	-2.09
246038_s_at	At5g19460 At5g19470	ATNUDT20, hydrolase	-2.05
249761_at	At5g23970	Transferase family protein	-2.04
255786_at	At1g19670	ATCLH1 / CORI1	-2.04
254915_s_at	At4g11310 At4g11320	RD21, cysteine proteinase, putative	-2.00
260309_at	At1g70580	AOAT2 / GGT2, alanine transaminase	-2.00
249690_at	At5g36210	Serine-type peptidase	-1.94
255619_s_at	At4g01350 At5g59920	DC1 domain-containing protein	-1.91
245465_at	At4g16590	ATCSLA01, glucosyltransferase/ transferase, transferring glycosyl groups	-1.76
264160_at	At1g65450	Transferase family protein	-1.75
264741_at	At1g62290	Aspartyl protease family protein	-1.73

Table 5.5 Genes differentially regulated between *A. l. petraea* populations. This table contains the genes that exhibited the greatest differential expression, either under control conditions or following a temperature treatment, between the two populations. Positive log_2 fold changes refer genes expressed more highly in Leitrim, and negative log_2 fold changes to genes expressed more highly in Helin.

Probe ID	AGI Code	Helin Difference between acclimation and control	Leitrim Difference between acclimation and control	Interaction	Magnitude of difference between Helin and Leitrim
249208_at	**At5g42650**	-1891.26	609.04		2500.31
267083_at	At2g41100	-2590.78	-473.17		2117.61
261648_at	At1g27730	-1741.76	164.36		1906.11
250358_at	**At5g11740**	2549.53	4453.50		1903.97
249237_at	**At5g42050**	400.49	2211.57		1811.08
254085_at	At4g24960	3055.68	1371.24		1684.43
252126_at	At3g50950	-1446.09	209.97		1656.06
247450_at	At5g62350	1122.97	-425.35		1548.32
261143_at	At1g19770	-1702.34	-208.63		1493.71
256129_at	**At1g18210**	-831.78	658.37		1490.15
246289_at	At3g56880	-1189.20	175.09		1364.29
264000_at	At2g22500	-1419.74	-72.29		1347.45
255064_at	At4g08950	-1242.25	94.40		1336.65
250676_at	At5g06320	-1194.84	94.71		1289.55
247925_at	At5g57560	-1180.43	15.12		1195.55
260876_at	At1g21460	2514.11	1341.42		1172.69
253284_at	At4g34150	-1003.89	109.84		1113.73
252592_at	At3g45640	-916.00	164.89		1080.89
264655_at	**At1g09070**	-309.88	608.19		918.06
255116_at	At4g08850	-730.60	147.37		877.97
261892_at	At1g80840	-776.54	89.30		865.84
254120_at	At4g24570	-773.35	59.21		832.56
254926_at	At4g11280	-821.28	-17.24		804.04
266371_at	At2g41410	-651.83	116.38		768.21
265149_at	At1g51400	-1461.07	-695.93		765.14
255786_at	At1g19670	-807.86	-81.95		725.91
253485_at	**At4g31800**	-508.23	217.32		725.55
252377_at	At3g47960	-879.26	-172.11		707.15

	Down-regulated in Helin, up-regulated in Leitrim
	Down-regulated in both populations, to a greater extent in Helin
	Up-regulated in both populations, to a greater extent in Leitrim
	Up-regulated in both populations, to a greater extent in Helin
	Down-regulated in Leitrim, up-regulated in Helin

Table 5.6 Acclimation interaction genes exhibiting the greatest differential regulation between Helin and Leitrim. Genes in this table were responsive to an acclimation treatment, but were also differentially regulated between the two study populations. The central columns indicate the differences between the average acclimation and control signal values (acclimation minus control) for each population. Red text indicates a reduction in the level of expression; blue text indicates an increase in the level of expression. Types of interactions are colour coded according to the key. The final column refers to the magnitude of difference between the responses to acclimation of the two study populations (absolute difference between the central two columns). AGI codes in bold text refer to genes chosen for study in Chapter 6.

5.3.3.2 Genes responsive to the additional freezing treatment

Following the additional freezing treatment, 147 further significant ($P < 0.05$) changes in gene expression were observed. A total of 75 of these genes, in addition to a significant change in expression, exhibited fold changes that were > 2 following this additional freezing treatment (Figure 5.3); of these, 65 genes were up-regulated (Appendix 2c) and only ten genes were down-regulated. The top twenty up-regulated genes and the ten down-regulated genes are detailed in Table 5.4. The genes that were up-regulated following the additional freezing treatment include transcription factors, binding proteins and a number of unknown proteins. The two most up-regulated genes are both annotated as unknown (At4g29780 and At2g35290), and both have $\log_2$ fold changes of greater than three. Seventy seven percent of these have been previously reported as cold responsive (> 2 fold up-regulated). The number of genes that were down-regulated following the additional freezing treatment is far fewer than for the acclimation treatment; there were only ten genes in total in this list and 50% of them have been previously reported as cold responsive (> 2 fold down-regulated). The two most down-regulated genes are At5g25350 (*EBF2*) and At1g26790 (Dof-type zinc finger domain-containing protein), both with $\log_2$ fold changes of greater than minus four.

5.3.3.3 Genes differentially regulated between Helin and Leitrim study populations

Over 1000 genes were differentially expressed ($P < 0.05$) between populations, irrespective of treatment. Positive and negative fold changes for these genes refer to which population had the higher level of expression; positive fold changes denote genes that were more highly expressed in Leitrim than Helin, and negative fold changes denote genes that were expressed more highly in Helin than Leitrim. Following filtering of genes to include only those for which the significant change in expression was greater than a fold change of two, there were 96 genes that exhibited positive > 2 fold differences in expression (Appendix 2d), and 111 genes that exhibited negative > 2 fold differences in expression (Appendix 2e; Figure 5.3). Genes that were found to be differentially expressed between the populations included several binding proteins and transferases. The gene most differentially expressed between the populations is annotated as similar to PBS lysase HEAT-like repeat-containing protein (expressed more highly in Leitrim than Helin). The top twenty positive fold change and negative fold change genes are detailed in Table 5.5.

5.3.3.4 Genes responsive to cold treatments and also differentially regulated between Helin and Leitrim study populations

In the Venn diagram (Figure 5.2), the region of overlap between acclimation and population denotes those genes that were responsive to the acclimation treatment ($P < 0.05$) but, in addition, were significantly differentially regulated between the two study populations ($P < 0.05$). There are 894 genes in this category, and 34 of these were also responsive to the additional freezing treatment. Similarly, the region of overlap between freezing and population denotes those genes that were responsive to the additional freezing treatment ($P < 0.05$) and were differentially regulated between the two populations ($P < 0.05$). There are 51 genes in this category, with just seventeen that were only responsive to freezing (34 overlap with the acclimation responsive genes). This set of genes that were responsive to cold treatments, but were differentially regulated between the two populations; have been termed the 'interaction genes'. Included within the interaction genes are: (i) genes that were up-regulated in response to acclimation or freezing but to a significantly greater extent in one population than another; (ii) genes that were down-regulated in response to acclimation or freezing but to a significantly greater extent in one population than another; (iii) genes that were significantly up- or down-regulated in one population but not responsive to temperature in the other; and (iv) genes that were up-regulated in one population and down-regulated in the other.

A list of the interaction genes has been produced using just those genes that were significantly differentially regulated ($P < 0.05$) and that exhibited > 2 fold changes in expression following the acclimation treatment in at least one of the populations; there are 469 of these genes in total and this gene set has been used in all subsequent analyses. The average signal values obtained from the arrays enabled the interaction type to be determined for each gene. These genes encompass six different interactions as detailed in Figure 5.4. The two most common interactions found were i) genes that were down-regulated in Helin but up-regulated in Leitrim, and ii) genes that were down-regulated in both populations but to a greater extent in Helin. The interaction that was the least common was for those genes that were down-regulated in Leitrim but up-regulated in Helin. Analysis of the average signal values for the control arrays indicates that, for many genes, the level of expression prior to temperature treatment is higher in Helin than in Leitrim.

The acclimation interaction genes include At4g31800 (*WRKY18*, transcription factor) that was down-regulated in Helin but up-regulated in Leitrim following the acclimation treatment. The expression level of this gene in the control samples was over ten times greater for Helin than for Leitrim. At5g42050 (unknown protein) was up-regulated in both populations following acclimation, but to a greater extent in Leitrim. The expression levels of this gene following acclimation are similar between the two populations, but the Helin population had a higher level of expression prior to cold temperature exposure. Further acclimation interaction genes are detailed in Table 5.6; the twenty genes with the greatest magnitude of difference between the Helin response to acclimation and the Leitrim response to acclimation are shown. This was calculated by determining the difference between the average control and acclimated signal values for each population independently, and then comparing these values between the two populations.

5.3.4 Functional classification of responsive genes

Genes with altered expression levels following the treatments within this study were categorised according to the gene ontology of their associated gene products (GO analysis) using GeneSpring 7.3 software. There are three main categories of gene ontology; biological process, molecular function, and cellular component. All three categories can be split into increasingly specific ontologies.

It was possible to determine the percentage of genes within an individual response set (e.g. those genes up-regulated in both populations following the acclimation treatment) that belonged to particular gene ontologies. These could then be compared to the percentage of genes from the *Arabidopsis* genome as a whole that were also assigned to these categories. GO Slim analysis (TAIR: www.arabidopsis.org/tools/bulk/go/index) was used to give an overview of enrichment or depletion of biological process categories in all sets of genes (Acclimation up- and down-regulated, freezing up- and down-regulated, population up- and down-regulated, and interaction). These data are shown in Figure 5.5. The categories ‘response to abiotic and biotic stimulus’ and ‘response to stress’ were significantly over-represented in all lists.

5.3.4.1 Genes responsive to the acclimation treatment

The set of genes showing significant increases in expression following the acclimation treatment (acclimation up-regulated) was enriched for a large number of ontologies

within the biological process category. The twenty most significantly over-represented ontologies are illustrated in Figure 5.6A. These include genes in the GO categories response to stimulus (GO:50896), response to stress (GO:6950), metabolism (GO:8152), and physiological process (GO:7582). Within the response to stress ontology, three increasingly specific ontologies were over-represented: response to cold (GO:9409), response to water deprivation (GO:9414), and cold acclimation (GO:9631). A number of more specific ontologies were over-represented within the metabolism category including secondary metabolism (GO:19748) and macromolecule metabolism (GO:43170). Within the molecular function category, the most significantly over-represented ontology was that of oxidoreductase activity (GO:16491).

The acclimation down-regulated set of genes was also enriched for a large number of ontologies within the biological process category. The twenty most significantly over-represented ontologies are shown in Figure 5.6B. These include transport (GO:6810) and growth (GO:40007). A large number of ontologies within transport were significantly over-represented including the most significant ontologies oligopeptide transport (GO:6857) and peptide transport (GO:15833). The most significantly over-represented cellular component for this set of genes is that of the membrane (GO:16020), and the some of the most significantly over-represented molecular function ontologies are related to transporter activity (GO:5215).

5.3.4.2 Genes responsive to the freezing treatment

The sets of genes found to be responsive to the additional freezing treatment were much smaller than for the acclimation treatment (see section 5.3.3.2), thus a smaller number of ontologies were found to be over-represented in these gene sets. The most significantly over-represented ontologies for the freezing up-regulated gene set were photosynthetic acclimation (GO:9643) and response to stress (GO:6950). Very few genes were linked to over-represented ontologies from the cellular component and molecular function categories. The most significantly over-represented ontologies for the freezing down-regulated gene set were related to metabolism and catabolism.

5.3.4.3 Genes differentially regulated between Helin and Leitrim study populations

Genes differentially regulated between Helin and Leitrim are those that were expressed more highly in one population than the other, either in control conditions or in response to the treatments. For the set of genes that were more highly expressed in Leitrim than

Helin (Population up-regulated), the most significantly over-represented gene ontologies were within the response to stimulus categories (response to copper ion (GO:46688), metal ion (GO:10038), and inorganic substance (GO:10035)) and the transport categories (disaccharide transport (GO:15766) and sucrose transport (GO:15770)). Molecular function ontologies linked to catalytic activity and binding were significantly over-represented in this gene set. A much wider range of gene ontologies were over-represented for the set of genes that were more highly expressed in Helin than Leitrim (Population down-regulated). The most significantly over-represented ontologies included secondary metabolism (GO:19728), cellular process (GO:9987), and physiological process (GO:7582). A number of more specific ontologies related to metabolism were significantly over-represented, particularly those linked to flavonoids and jasmonic acid. Very few individual genes were linked to over-represented ontologies for cellular component and molecular function categories.

5.3.4.4 Genes responsive to cold treatments and also differentially regulated between Helin and Leitrim study populations: Interaction genes.

Interaction genes are those that were found to be responsive to the temperature treatments, but were also differentially regulated between the two study populations. A more detailed explanation of 'interaction genes' can be found in section 5.3.3.4.

Figure 5.7A illustrates those ontologies from the biological function category to which this gene set are linked. Response to stimulus (GO:50896) is the most significantly over-represented ontology. Figure 5.7B illustrates more specific ontologies within the response to stimulus ontology; the most significantly over-represented ontologies include defence response (GO:6952) and response to stress (GO:6950). For those ontologies within the cellular component category, chloroplast inner membrane (GO:9706) and plastid inner membrane (GO:9528) were the most significantly over-represented. Within the molecular function category, protease inhibitor activity (GO:30414) and several ion binding ontologies were the most significantly over-represented.

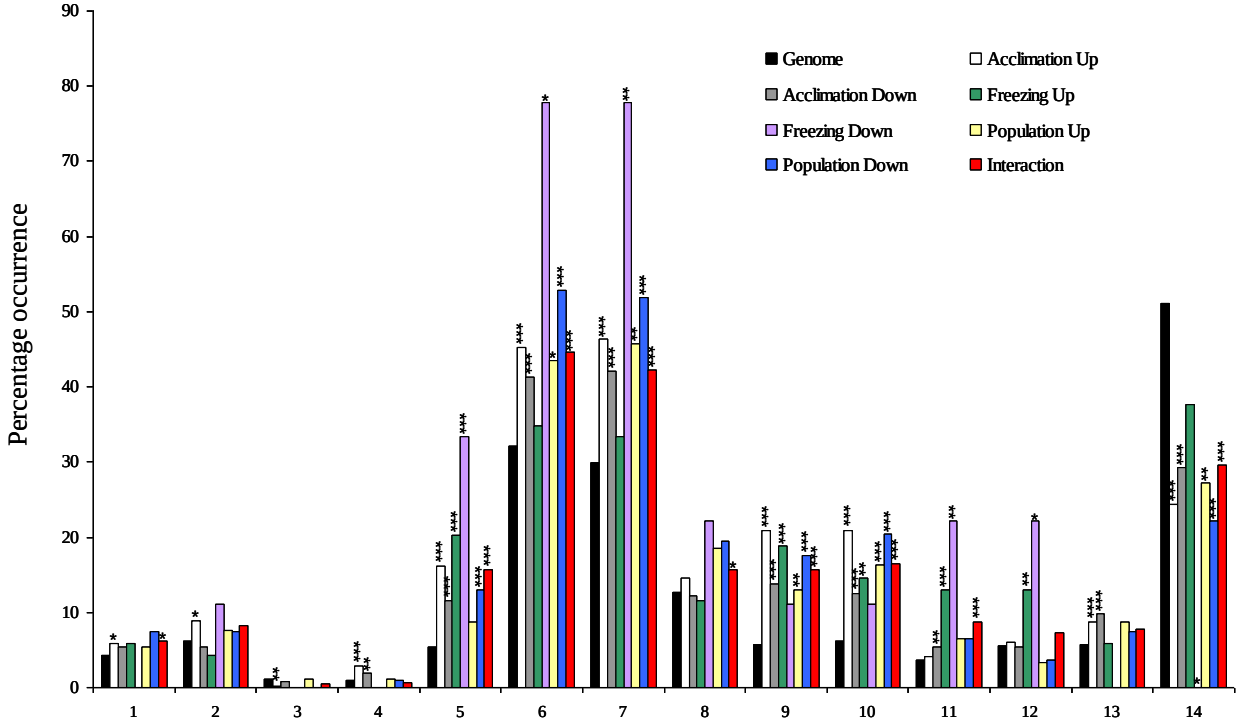
5.3.5 Pathway Analysis

GeneSpring 7.3 was used to investigate if genes responding to the treatments in this study relate to particular metabolic pathways. Firstly, the set of genes that exhibited increased expression following acclimation (acclimation up-regulated) contained ten

genes from the flavonoid synthesis pathway and consequently this pathway was significantly over-represented ($P < 0.0005$). The proteasome pathway was also significantly over-represented in this set of genes ($P < 0.0005$, eight genes). The acclimation up-regulated set of genes was split according to population to give a set of genes responsive to acclimation for Helin, and a separate set for Leitrim, thus enabling pathway analysis to be performed for the populations independently. Pathways that were over-represented in just one of the two populations could therefore be identified providing further insight into population-specific transcriptomic responses.

In the Helin-only set of genes, the starch and sucrose metabolism pathway was significantly over-represented ($P < 0.05$, eight genes), yet this was not the case for the Leitrim-only gene set. The nucleotide sugars metabolism pathway was significantly over-represented ($P < 0.05$, five genes) in the Leitrim-only gene set, but not in the Helin-only gene set. The set of genes that exhibited reduced levels of expression following acclimation in both populations (acclimation down-regulated) contained eleven genes from the photosynthetic pathway resulting in over-representation of this pathway ($P < 0.0001$). Pathway analysis was also performed on the sets of genes that were responsive to the additional freezing treatment (freezing up-regulated and freezing down-regulated); no pathways were significantly over-represented in these gene sets.

For genes that were differentially regulated between the populations, those genes that were expressed more highly in Leitrim than Helin were not significantly linked to any pathways. Those genes expressed more highly in Helin than Leitrim contained three genes from the flavonoid synthesis pathway thus this was significantly over-represented ($P < 0.05$). Pathway analysis was also performed on the sets of genes described as interaction genes; no pathways were significantly over-represented in these gene sets.



1	Cell organisation and biogenesis
2	Developmental processes
3	DNA or RNA metabolism
4	Electron transport or energy pathways
5	Other biological processes
6	Other cellular processes
7	Other metabolic processes
8	Protein metabolism
9	Response to abiotic or biotic stimulus
10	Response to stress
11	Signal transduction
12	Transcription
13	Transport
14	Unknown biological processes

Figure 5.5 GO Slim biological process annotations for all gene sets. All sets of differentially expressed genes were compared to the whole *Arabidopsis* genome using the TAIR website (http://www.arabidopsis.org/tools/bulk/go/index.jsp). Enrichment or depletion of GO Slim categories was determined by chi-squared tests of comparison. Acclimation up and acclimation down refer to those genes significantly up and down regulated following the acclimation treatment. Similarly, freezing up and freezing down refer to those genes significantly up and down regulated following the freezing treatment. Population up refers to genes expressed more highly in Leitrim than Helin, and population down refers to genes expressed more highly in Helin than Leitrim. The interaction gene set includes those genes that were significantly differentially regulated between the populations in response to the acclimation treatment.

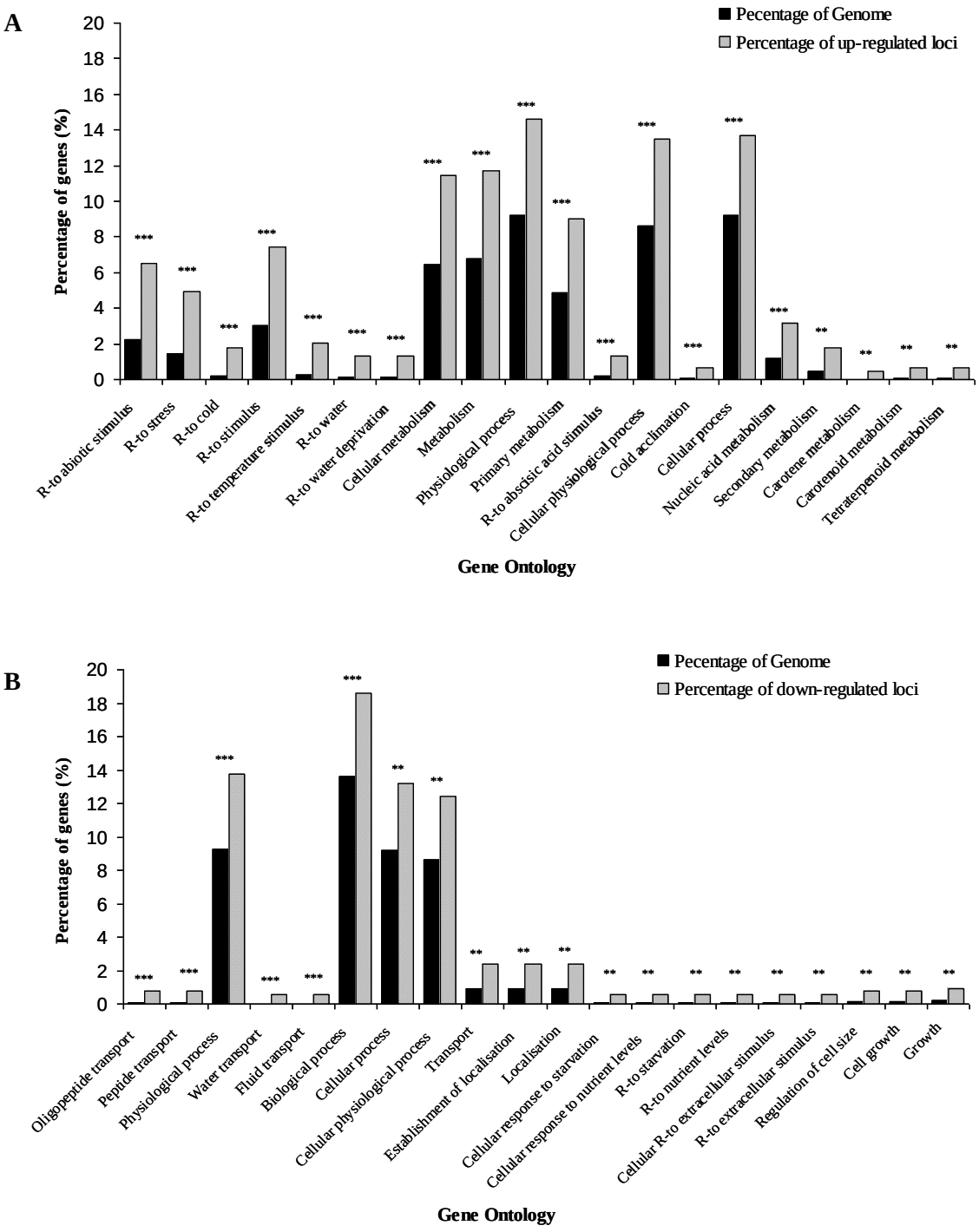


Figure 5.6 Gene ontology (GO) biological annotation for genes responsive to acclimation in *A. l. petraea*. Black bars illustrate the percentage of the *Arabidopsis* genome that is annotated to each of the ontologies (GO category). Shaded bars show the percentage of acclimation responsive genes, as identified in this study, that are annotated to each category: A) acclimation up-regulated; B) acclimation down-regulated. The twenty most significantly over-represented GO categories are shown in this figure. Asterisks indicate significance of over-representation (** P <0.01, *** P <0.001).

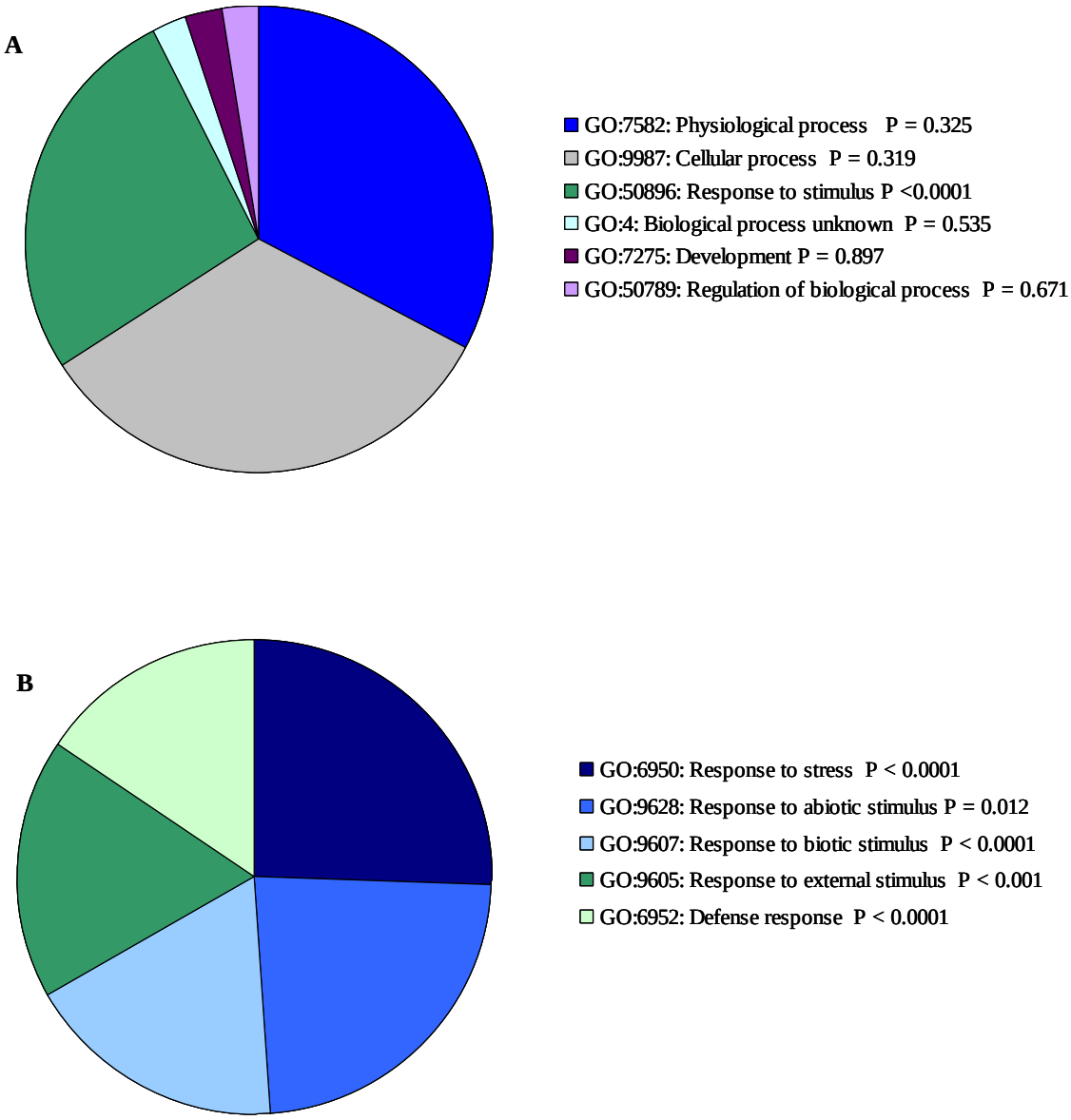


Figure 5.7 Gene ontology (GO) categories the acclimation interaction gene set. Figure A illustrates the percentage of genes in the acclimation interaction gene set (section 5.3.3.4) that relate to gene ontologies within the 'Biological Process' category. P values of less than 0.05 illustrate that the particular ontology is significantly over-represented in this gene set. Figure B refers to more specific gene ontologies within the 'response to stimulus' category. All five ontologies are significantly over-represented in this gene set.

5.3.6 Verification of microarray results through qRT-PCR

The results obtained from the Affymetrix ATH1 GeneChip® arrays were verified using qRT-PCR. Genes used for qRT-PCR were representatives selected to encompass a range of fold changes in response to acclimation, both positive and negative. All primer pairs used had efficiencies > 90% and < 110% (Table 5.2 and Figure 5.8). The specificity of each primer pair was analysed using a dissociation curve; one clean peak was observed for all primer pairs (Figure 5.9). The majority of primer pairs that were designed to *A. thaliana* coding sequences worked effectively for *A. l. petraea* cDNA template.

The $\log_2$ fold changes from the microarrays have been correlated against those for qRT-PCR for each population individually (Figure 5.10). For both Helin and Leitrim, the mean extent of the fold change measured using qRT-PCR was greater than for the microarrays. Pearson product-moment correlation indicated a strong and very significant positive association between microarray $\log_2$ fold change and qRT-PCR $\log_2$ fold change for both Helin ($r = 0.901$, d.f. = 14, $P < 0.001$) and Leitrim ($r = 0.855$, d.f. = 14, $P < 0.001$).

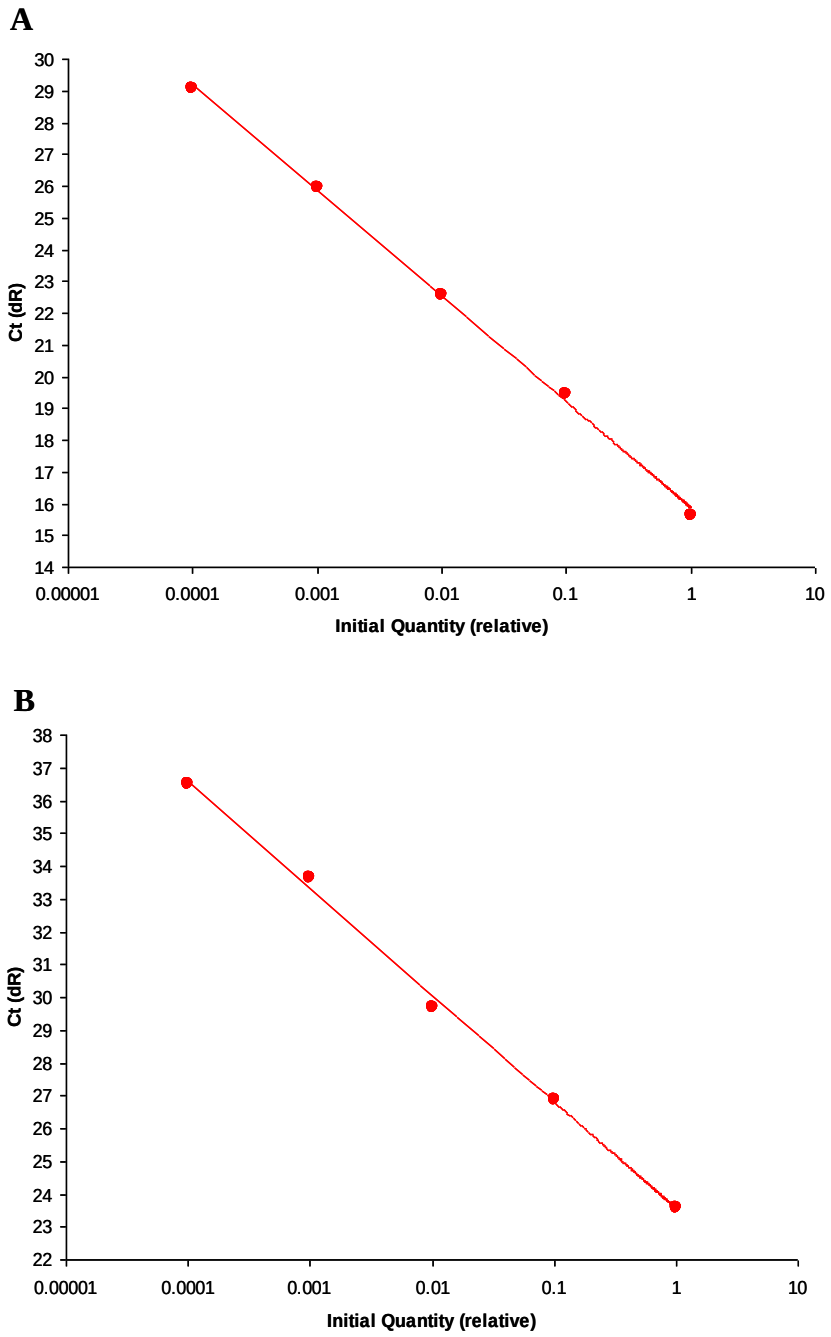


Figure 5.8 Efficiency of qRT-PCR primer pairs used in microarray verification. Primers were designed for use in qRT-PCR to the *A. thaliana* sequence obtained from sequence viewer (SeqView) at TAIR (http://www.arabidopsis.org) for seventeen genes. These genes all exhibited altered levels of expression in both Helin and Leitrim plants following the two day acclimation at 2°C. A standard curve was produced for each primer pair using a 10x dilution series; all pairs had efficiencies of > 90% and < 110%. Two representative standard curves are illustrated in this figure: A) At5g15970 primers had an efficiency of 99.3% (RSq = 0.999), and B) At1g47920 primers had an efficiency of 102.7% (RSq = 0.998).

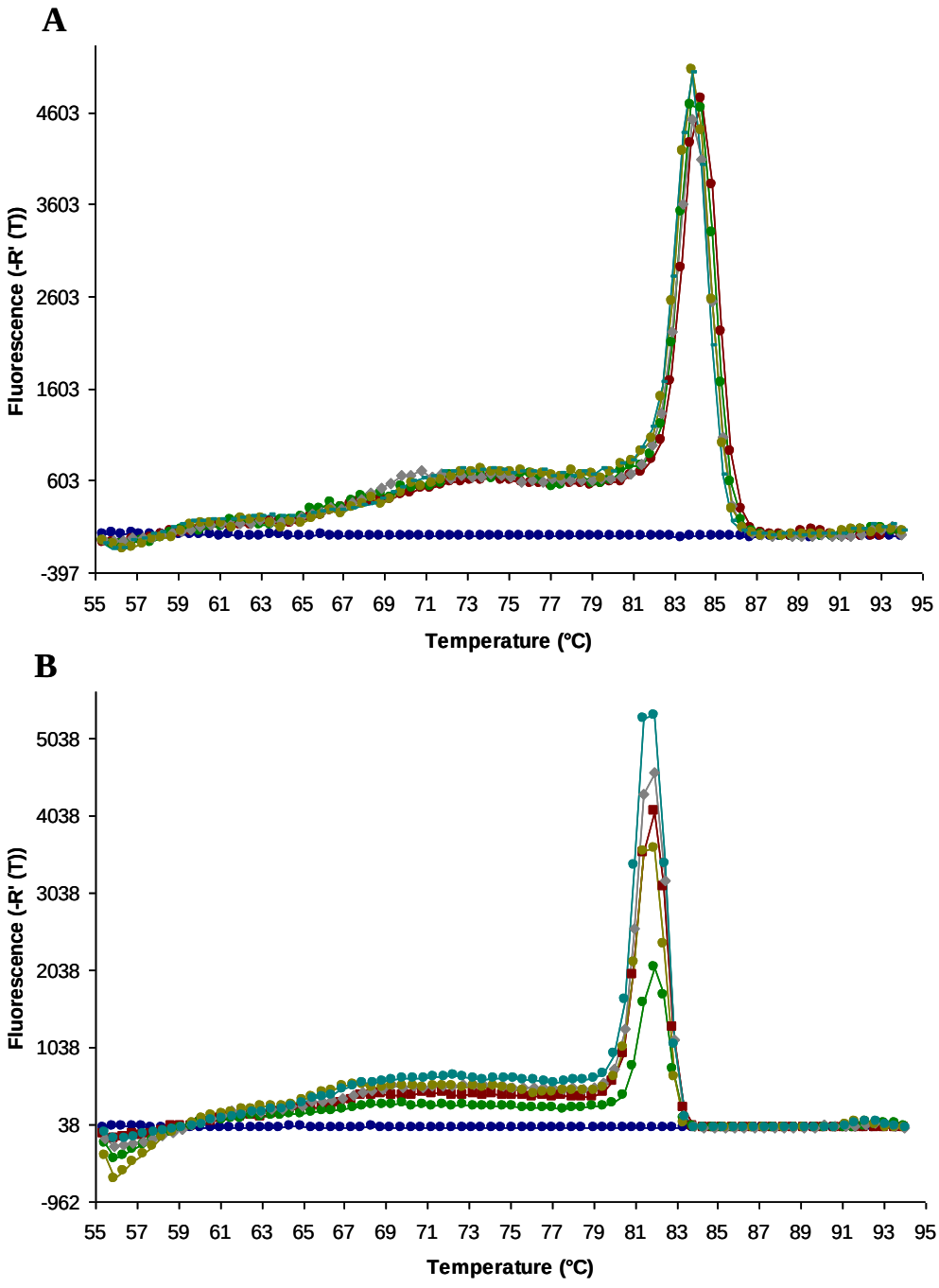


Figure 5.9 Representative dissociation curves for qRT-PCR primers used in verification of microarray results. Dissociation curve analysis illustrates the specificity of qRT-PCR primers. The temperature at which the product melts is indicative of its size. One single product is produced for both the At5g15970 (A) and the At1g47920 (B) qRT-PCR primers. Primers were designed to the *A. thaliana* sequence for amplification of *A. l. petraea* cDNA. One single product was also produced for the remaining fifteen primer pairs (data not shown). Dark blue circles indicate no template control.

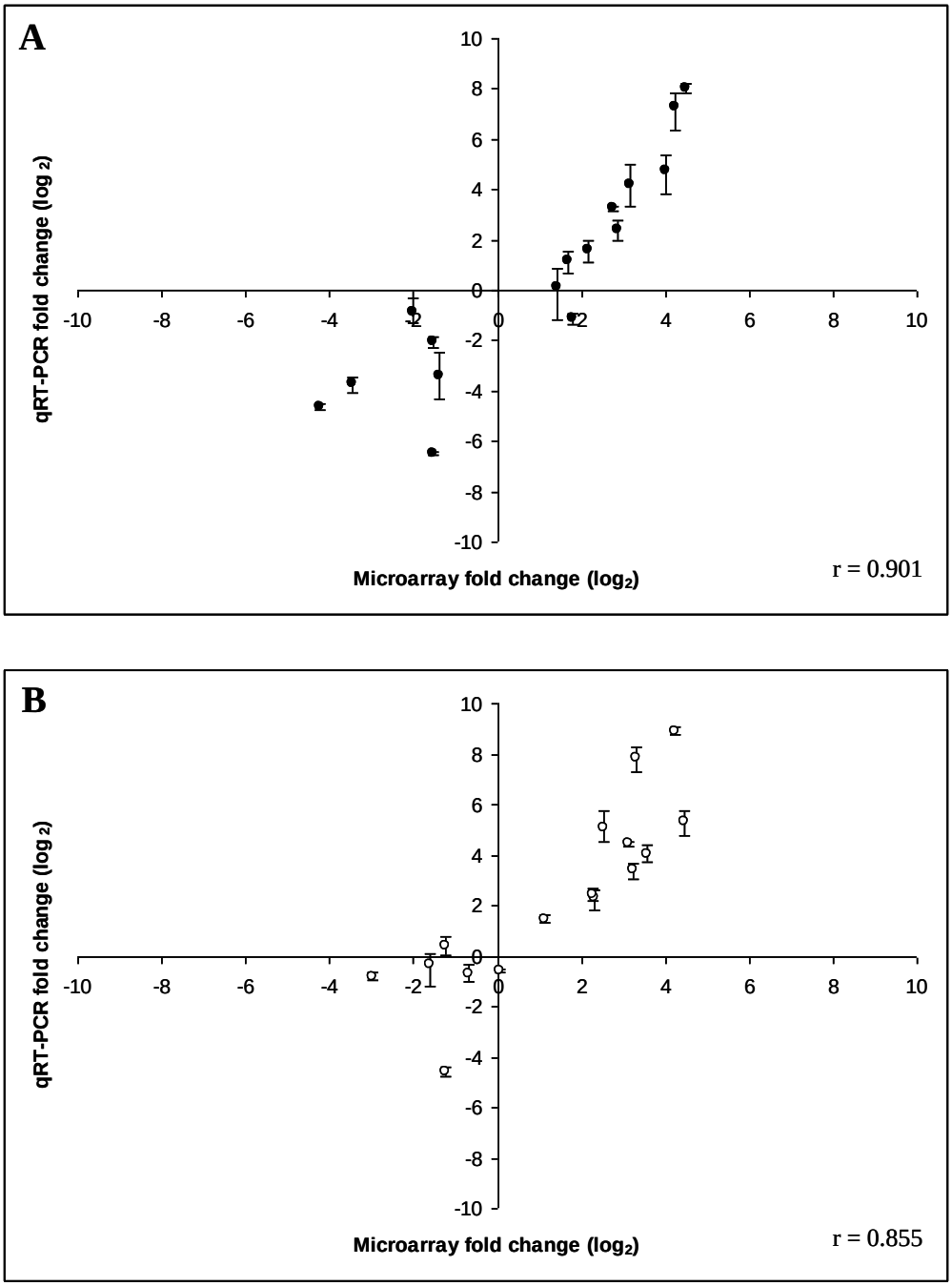


Figure 5.10 Correlation of Affymetrix ATH1 GeneChip® array and qRT-PCR results. Significant positive correlation is shown between fold change (log_2) results from the Affymetrix arrays and qRT-PCR verification for both populations; Helin (A, solid circles) and Leitrim (B, open circles). Seventeen genes, randomly selected from the acclimation-responsive gene set, were used for the verification of microarray results. Significance of the correlations was tested using Pearson product-moment correlations and the r values from these are indicated on each graph ($P < 0.001$ for both correlations).

5.4 Discussion of differentially regulated genes of *A. l. petraea* in response to acclimation, freezing, and between study populations.

5.4.1 Validity of performing cross-species transcriptomics

Affymetrix ATH1 arrays are designed to the *A. thaliana* sequence and contain probe sets for 24,000 *Arabidopsis* genes. As previously reviewed by Clauss and Koch (2006), average species divergence between *A. thaliana* and *A. l. petraea* is estimated at 1-15% indicating that there is a high level of sequence similarity between these species. These sequence similarities enable technologies that have been developed for *A. thaliana* to be easily transferred to the study of *A. l. petraea* and other closely related species. Affymetrix ATH1 GeneChip® arrays have been successfully used to study other relatives of *Arabidopsis thaliana* including two *Thlaspi* species (Hammond *et al.*, 2005) and *Arabidopsis halleri* (Becher *et al.*, 2004; Weber *et al.*, 2004). Hammond *et al.* (2005) report hybridisation for all probes on the ATH1 GeneChip®, and were able to detect over 5000 differentially expressed genes between the two *Thlaspi* species, one of which is a zinc (Zn) hyperaccumlator. Good signal intensities were obtained for a large number of probe sets on the arrays presented here and for previous ATH1 GeneChip® arrays (C. J. Lilley), both of which have used *A. l petraea* material. Quality control measures indicate good hybridisation across the arrays, and genes that were responsive to the different experimental conditions have been successfully identified. These results demonstrate that *Arabidopsis* ATH1 GeneChip® arrays are suitable for and useful in the study of the transcriptome of *A. thaliana* close relatives.

5.4.2 Corroboration of microarray results by qRT-PCR

A wide range of techniques have been used in the validation of published microarray experiments, and there is no consensus on the most applicable method (Rockett and Hellmann, 2004). qRT-PCR is the most commonly used technique for array corroboration, with over two hundred published references using this method (Chuaqui *et al.*, 2002; Dallas *et al.*, 2005). Its popularity is likely to relate to the rapidity and sensitivity of the technique, and the minimal quantities of starting template required (Gachon *et al.*, 2004; Dallas *et al.*, 2005). Here, corroboration of the microarray results was carried out through qRT-PCR. Expression values obtained from the two platforms demonstrate significant correlation for both of the study populations (Helin, r = 0.901; Leitrim, r = 0.855). This provides a high level of confidence in the overall ability of the microarrays to detect differences in gene expression between untreated and temperature

treated *A. l. petraea* plants, and further confirms the validity of performing cross-species transcriptomics. Correlation coefficients are in line with those previously reported in the literature (He *et al.*, 2008; Trzcinska-Danielewicz *et al.*, 2009).

Of the two platforms employed in this study, the more reliable measure of gene expression is considered to be qRT-PCR (Gachon *et al.*, 2004), as there is no strict linear relationship between transcript amount and signal intensity for hybridisation-based technologies; comparisons between untreated and treated samples in microarrays are therefore qualitative rather than quantitative (Czechowski *et al.*, 2005). Here, fold change values obtained through qRT-PCR were higher than those obtained from the Affymetrix ATH1 GeneChip® arrays; this is in agreement with other published studies that have used qRT-PCR for verification of microarray experiments (Yuen *et al.*, 2002; Redman *et al.*, 2004; Dallas *et al.*, 2005; Trzcinska-Danielewicz *et al.*, 2009). Ratios of expression levels obtained from microarray studies are known to be underestimated in both cDNA and oligonucleotide arrays (Yuen *et al.*, 2002), and consequently the magnitude of detected expression changes tends to differ between microarrays and other technologies such as qRT-PCR (Draghici *et al.*, 2006).

5.4.3 Detecting genes of interest from microarray data

Normalisation of microarray data must be performed before comparisons between arrays are valid (Clarke and Zhu, 2006). Global normalisation acts to minimise and standardise non-biological variation, such as unequal starting quantities of RNA, background intensity, and hybridisation conditions (Quackenbush, 2002). Values are normalised across all experimental chips to give the same distribution of signal values, and consequently comparisons between arrays can then be made. The GCRMA method for normalisation was used here; this adjusts for optical noise and non-specific binding. Differences between arrays were successfully minimised, and thus valid comparisons between the arrays could be made.

Variability in microarray experiments can be minimised through replication; this can be performed at the biological or technical level. Biological variation provides a measure of the natural variability in the study system and any random variation that may occur in sample preparation, whilst technical variability identifies experimental reliability by measuring the same biological sample multiple times (Quackenbush, 2002; Clarke and Zhu, 2006). Expression values from replicate experiments are usually averaged

following normalisation to reduce the overall complexity of the dataset. The extent of replication in microarray experiments is often restricted by factors including the limited availability of biological material and the high costs involved (Clarke and Zhu, 2006; Jeffery *et al.*, 2006). The optimal level of replication for microarray experiments has not been clearly defined. It is accepted that biological variation outweighs that of technical variation, and thus biological replication can be prioritised in the design of microarray experiments (Zhu and Wang, 2000; Meyers *et al.*, 2004). Published microarray studies in plant molecular biology typically perform three or four replicates for each experimental condition (Zhang *et al.*, 2004; Ditt *et al.*, 2006; Kant *et al.*, 2008), although some studies report fewer replicates and are consequently able to investigate a wider range of experimental conditions (Swindell, 2006). It has been suggested that at least five biological replicates are required for stable results to be obtained, with the ideal being between ten and fifteen replicates (Pavlidis *et al.*, 2003). The three biological replicates for each of the treatment combinations that were performed in this study clustered together well (data not shown), thus indicating the similarity of expression patterns between biological replicates.

Microarrays generate a wealth of data; this both contributes to their appeal and creates challenges in terms of extracting biologically meaningful data. Differentially regulated genes are those that display significant differences in expression between samples originating from the various conditions under investigation, and they are identified by the creation of gene lists based on selection criteria that are primarily chosen by the researcher. A fold change cut-off is the most common technique used to identify differentially regulated genes; this involves a gene being designated as significantly changed if its expression level varies by a chosen constant between the control and treated conditions (Hatfield *et al.*, 2003). There are limitations to this method, and selection of genes based on a combination of fold change, statistical significance, and expression level is advised (Clarke and Zhu, 2006). In this study, genes were designated as responsive to a particular treatment if they exhibited a fold change of greater than two (or of one on a $\log_2$ scale), and significance P values of less than 0.05. Due to these strict criteria and the inclusion of three biological replicates for each treatment, expressional changes reported are very likely to reflect real changes in the transcriptome of *A. l. petraea*.

5.4.4 Differential regulation of genes in response to acclimation in *A. l. petraea*

Over 1600 genes were deemed responsive to acclimation across both populations ($P < 0.05$, fold change >2), indicating that at least 6% of the 24 000 genes included on the ATH1 GeneChip® respond to cold temperatures in *A. l. petraea*. This is concurrent with findings for *A. thaliana* for which 4 – 8% of the transcriptome has been reported as responsive to cold temperature exposure (Hannah *et al.*, 2005; He *et al.*, 2008). Acclimation temperatures are known to stimulate a large variety of physiological and biochemical responses in *Arabidopsis* including changes in lipid composition, accumulation of cryoprotectants, reduced growth rates, and increased antioxidant production (Thomashow, 1999; Xin and Browse, 2000; Smallwood and Bowles, 2001). Significant changes in the transcriptome are to therefore be expected, as many genes will be involved in the regulation of these responses and will consequently be induced or repressed (Smallwood and Bowles, 2001). Kreps *et al.* (2002) investigated the effects of three abiotic stresses on plants and reported nearly twice as many cold responsive genes as there were to the other two stresses (salt and drought), indicating that cold stress has a large effect on gene expression relative to that of other stresses. Similar numbers of genes were both up- and down-regulated in response to acclimation in *A. l. petraea*, with slightly more genes exhibiting down-regulation; Kreps *et al.* (2002) report a comparable pattern for *A. thaliana*.

5.4.4.1 Genes previously annotated as cold responsive

Many genes previously reported to be cold responsive in *A. thaliana* are shown to respond to cold acclimation in *A. l. petraea* in this study. For a range of plant species, components homologous to the CBF cold response pathway in *Arabidopsis* have been reported: *Brassica napus* (Jaglo *et al.*, 2001); barley (Choi *et al.*, 2002); strawberry and sour cherry (Owens *et al.*, 2002). This indicates that at least some aspects of the cold response are conserved between species. It is therefore sensible to expect many of the genes responsive to cold temperatures in *A. l. petraea* to overlap with those for *A. thaliana*. Vogel *et al*. (2005) describe a core set of cold up- and down-regulated genes for *A. thaliana*; these are genes that were consistently responsive to cold under two different sets of growth conditions. This gene set comprises 512 genes, and 224 of the genes responsive to acclimation in this study overlapped with this gene set. This indicates that transcriptomic responses are conserved between these two closely related species. In addition, Vogel *et al.* (2005) assigned a number of genes to the CBF regulon

based on their expression levels under constitutive *CBF2* expression; there are 85 genes in this regulon and over half of them were deemed cold responsive in *A. l. petraea*.

Following exposure to cold temperatures a cascade of signalling pathways results in the regulation of cold responsive genes. These genes can be classified into two groups; the first containing genes that encode proteins with roles in regulation of gene expression and signal transduction (Thomashow, 1999), and the second containing genes encoding functional proteins likely to have roles in stress tolerance (Seki *et al.*, 2001; Seki *et al.*, 2002). Transcription factor proteins (members of the first group) bind to specific DNA sequences adjacent to the genes that they regulate, and consequently can up-regulate or down-regulate the expression of particular genes. There are 85 transcription factors designated as cold responsive in this investigation; 74 in response to acclimation (32 up-regulated and 42 down-regulated), and eleven in response to the additional freezing treatment (ten up-regulated and one down-regulated). The large number of cold responsive transcription factors indicates that there are many transcriptional regulatory mechanisms involved in responses of *A. l. petraea* to cold stress, and is in concurrence with data obtained by Hannah *et al.* (2005); they report a wide range of transcription factors as responsive to differing periods of cold temperature exposure in *A. thaliana*. Particular families of transcription factors are shown to be responsive to the acclimation treatment including five genes from the WRKY family (At2g30250 and At4g30935 are up-regulated; At3g56400, At2g38470, and At3g04670 are down-regulated). Seki *et al.* (2002) reported up-regulation of four WRKY transcription factors following cold treatment, and this family have also been shown to be responsive to other stresses including heat and pathogen attack (Eulgem *et al.*, 2000; Nolan *et al.*, 2006). Other transcription factors up-regulated in response to acclimation include members of the AP2 and NAC families, both of which have previously been identified as having involvement in the cold response of *A. thaliana* (Seki *et al.*, 2002; Hannah *et al.*, 2005). These families of transcription factors are therefore likely to function in the regulation of some stress-inducible genes.

Within the set of genes up-regulated following the acclimation treatment, a number of gene ontologies relating to stress and, more specifically, to cold temperatures were over-represented. This indicates that many of the genes known to be regulated in response to cold temperatures in *A. thaliana* were also responsive to this acclimation treatment in *A. l. petraea*, and thus there is evidence for a conserved response to cold in the two

species. This is further confirmed by comparison of the top 100 acclimation-responsive genes in this study to available microarray data, with 87% of the genes having previously been shown to respond to cold (GENEVESTIGATOR). A number of designated COR/LEA genes, members of the second group of cold regulated genes, were identified within this set of genes. Many members of this group of genes are cold induced and encode hydrophilic proteins that are thought to have roles in freezing tolerance through the stabilisation of membranes against freezing damage (Thomashow, 1999). These include *RESPONSIVE TO ABA 18* (*RAB18*, At5g66400), *LOW TEMPERATURE INDUCED 30* (*LTI30/XERO2*, At3g50970) and *COLD-RESPONSIVE 78 (COR78;* At5g52310). Hannah *et al.* (2005) proposed the classification of COR/LEA proteins based on hydrophilicity and glycine contents; using their classification system, XERO2 and RAB18 were deemed two of the most hydrophilic COR/LEA proteins. Overexpression of both of these genes has been shown to enhance freezing tolerance in *Arabidopsis* (Puhakainen *et al.*, 2004), and this may explain the significant up-regulation of these genes observed in *A. l. petraea* plants exposed to cold stress.

5.4.4.2 Primary and secondary metabolism

Gene ontologies relating to metabolism, both primary and secondary, were also significantly over-represented in the set of genes up-regulated in response to acclimation. It has been documented that the metabolome of *A. thaliana* is responsive to changes in temperature, with cold shock causing more profound changes than heat shock (Kaplan *et al.*, 2004). Cook *et al.* (2004) report that 75% of metabolites monitored were increased in plants exposed to cold temperatures, and that the majority of these metabolites responded similarly in untreated transgenic lines of *A. thaliana* over-expressing *CBF3*, a major component of the CBF cold response pathway. Significant metabolic changes in *A. thaliana* in response to cold temperature exposure include rapidly increased salicylic acid (SA) levels, implicating SA as an early signalling molecule, and the early accumulation of Suc, a metabolite of primary metabolism (Kaplan *et al*., 2004). Kaplan *et al.* (2004) suggest that, because the promoter regions of many cold responsive genes contain sugar-responsive elements, Suc may be a candidate sugar signalling molecule with an important role in the acquired freezing tolerance of *A. thaliana*. Changes in the metabolome of *A. l. petraea* have also been documented in response to cold temperatures; Davey *et al.* (2009) report increases in free carbohydrates including fructose, raffinose, and mannose, with particularly large

increases observed for sucrose. It is suggested that, as sucrose can be easily mobilised and transported, its increase may be beneficial for the additional respiratory needs of plants during cold acclimation, during which many biochemical changes will occur (Davey *et al*., 2009). Secondary metabolites such as flavonoids are known to accumulate during plant stress (Winkel-Shirley, 2002), and significant over-representation of genes involved in the production of secondary metabolites has been reported in acclimated *A. thaliana* during short, medium, and long-term exposure to cold temperatures (Hannah *et al.*, 2005). In addition to over-representation of the gene ontology for secondary metabolism in the gene set responsive to acclimation in *A. l. petraea*, pathway analysis revealed over-representation of the flavonoid synthesis pathway. Several potential stress-related functions for flavonoids have been reported in the literature, including protection of cells from oxidative stress through the absorbance of UV light (Winkle-Shirley, 2002). Cold temperature specific roles for flavonoids have also been suggested; these include the scavenging of reactive oxygen species (ROS) and direct effects on the stability of cell membranes (Korn *et al.*, 2008).

5.4.4.3 Photosynthesis and growth

For the set of genes that were down-regulated in response to the acclimation treatment, GO categories such as physiological and cellular processes, and growth, were found to be over-represented. Additionally, pathway analysis revealed that the photosynthetic pathway is significantly over-represented in this gene set. The response of plants to cold temperatures involves signalling pathways and cascades of changes in gene expression, and this will prove metabolically costly (Kreps *et al.*, 2002; Hannah *et al.*, 2006). It is therefore feasible that plant processes such as photosynthesis and growth will be halted or significantly reduced, at least during the initial phase of exposure to stress. Reduction in photosynthetic rate has been demonstrated in *A. thaliana* following exposure to chilling temperatures (Strand *et al.*, 1997; Savitch *et al.*, 2001), and is coupled with the accumulation of soluble sugars and the reduction in transcript levels of several photosynthetic genes (Strand *et al*., 1997). Additionally, Seki *et al.* (2002) report the down-regulation of a number of photosynthetic genes in *A. thaliana* following exposure to abiotic stress, a pattern also reported in spinach (Hua, 2009) and more recently in cold-acclimated perennial ryegrass (Zhang *et al.*, 2009). The observed down-regulation of genes involved in photosynthesis and growth in *A. l. petraea* following cold acclimation is thus in concordance with trends for other species, and indicates that these processes are similarly affected in both populations studied.

5.4.5 Differential regulation of genes in response to freezing in *A. l. petraea*

The freezing treatment, applied to plants that had already experienced a two day acclimation period, resulted in the additional regulation of a small set of genes (75 genes: $P < 0.05$, fold change >2), the majority of which were up-regulated rather than down-regulated. There is a wealth of literature documenting the transcriptional changes that occur during cold acclimation in *Arabidopsis* (e.g. Kreps *et al*., 2002; Hannah *et al*., 2005; He *et al*., 2008), and the consequential increases in freezing tolerance that are observed (Wanner and Junttila, 1999; Zhen and Ungerer, 2008a); information on the transcriptional changes that occur in plants that have experienced freezing stress, however, appears scarce. There are several practical limitations to the collection of such data that may help to explain the lack of published research into the effects of freezing on the transcriptome including: i) without a prior period of acclimation, the exposure of *Arabidopsis* to freezing temperatures is likely to cause severe damage to leaf tissues and thus prevent extraction of high quality RNA for transcriptomic analysis; and ii) many standard growth cabinets do not have the capacity to reach freezing temperatures, and the use of a general lab freezer for such treatments does not permit control over light conditions. Photosynthesis is required for a full acclimation response to be realised; exposure of *Arabidopsis* to cold temperatures in darkness does not result in enhanced freezing tolerance (Wanner and Junttila, 1999). Within the GENEVESTIGATOR database there are no datasets involving exposure to freezing temperatures so although many of the genes deemed responsive to freezing in *A. l. petraea* have previously been reported as responsive to cold temperatures, their involvement in the response of *Arabidopsis* to freezing stress cannot be confirmed.

Although there is practically no information on the transcriptomic responses of *Arabidopsis* following exposure to sub-zero temperatures, limited studies have been performed in wheat (*Triticum aestivum* L.). Extensive transcriptomic changes have been observed at both -3^{o}C (Herman *et al.*, 2006) and at -10 and -12^{o}C (Skinner, 2009), indicating that the transition to sub-zero temperatures is complex. Expression of over 400 genes is reported to alter in both investigations, with the majority of genes exhibiting up-regulation (Skinner, 2009). Many of the up-regulated genes are involved in signal transduction and transcriptional regulation, whilst down-regulated genes are related to plastids and photosynthesis (Herman *et al*., 2006; Skinner, 2009). Freezing-responsive genes were identified following comparison with untreated rather than acclimated plants, thus many of the genes may, in fact, be responding to the cold-

acclimation period. This may explain the large number of freezing-responsive genes identified in wheat in relation to the 75 genes identified in *A. l. petraea*.

5.4.5.1 Stress-related genes

The set of genes up-regulated in response to freezing in *A. l. petraea* was enriched for the gene ontology response to stress. This indicates that genes known to have involvement in response to other stresses in *Arabidopsis* are also involved in the response to freezing stress. The gene *CBF4* (At5g51990) has been previously described as responsive to a number of stresses in *Arabidopsis* including salt, drought, and abscisic acid, but no up-regulation has been observed under cold stress at 4^{o}C (Haake *et al.*, 2002; Sakuma *et al.*, 2002). A functional genomics study of *CBF4* revealed that the encoded protein is a close homolog to proteins encoded by *CBF1, 2,* and *3,* and that over-expression of *CBF4* resulted in constitutive expression of the cold and drought inducible genes *COR15a* and *COR78a*. Transgenic plants were also found to be more tolerant to both freezing and drought stress, and consequently *CBF4* has been implicated as a member of the *CBF/DREB1* family (Haake *et al.*, 2002). *CBF4* is one of the most up-regulated genes in response to freezing stress in *A. l. petraea,* thus indicating that it does in fact have a role in the cold response of *Arabidopsis*, but, unlike *CBF1, 2,* and *3,* it is only up-regulated under conditions of extreme cold. *TCH4* (*TOUCH 4*) is one of the most up-regulated genes in response to freezing stress in *A. l. petraea* with a $\log_2$ fold change of three. This gene is responsive to a diverse range of stimuli including touch, darkness, cold temperatures, and auxin, and previous studies have revealed transient expression patterns in response to stimuli with rapid accumulation and subsequent reductions in transcript levels (Polisensky and Braam, 1996; Iliev *et al.*, 2002). Mechanically stimulated plants have been shown to exhibit greater freezing tolerance than untreated controls (Jaffe and Biro, 1979), indicating that although TCH genes are regulated by a range of stimuli, they may result in similar physiological responses. TCH4 is known to have activity in cell wall modification (Xu *et al.*, 1995; Campbell and Braam, 1999), and it has subsequently been suggested that properties of the cell wall are sensed and transmitted through a pathway that then affects the expression of *TCH4* (Iliev *et al.*, 2002). If this were the case, the observed up-regulation of this gene in response to freezing stress may be explained by the severe effects that freezing will have on cell wall structure.

5.4.5.2 Unknown proteins

Over 18% of the genes up-regulated in response to freezing in this study are annotated as unknown proteins, indicating that little is known of their functions in *Arabidopsis.* It is possible that many of these genes actually have important roles in response to freezing stress, and that the lack of previous work in this area has prevented the selection of these genes for functional genomics studies. Three of the four most freezing up-regulated genes in this study relate to unknown proteins; all three of these have previously been reported as responsive to cold stress but exhibit greater extents of up-regulation under colder temperature regimes (i.e. at 4^{o}C rather than at 10^{o}C or greater) (GENEVESTIGATOR). This suggests that there may be a set of genes that, although responsive to cold stress, are particularly responsive to below-freezing temperatures and that their main roles may relate to protection of the plant upon exposure to freezing conditions. Le *et al.* (2008) investigated acclimation at both above and below zero temperatures in *A. thaliana*; expression of specific *CBF* and *COR* genes was quantified for both conditions and no correlation was found between gene expression and subsequent changes in freezing tolerance. This has lead to the conclusion that the genetic and molecular basis of acclimation is different to that of subzero acclimation. Although not directly comparable to the data presented here, this does perhaps help to explain the significant number of genes responsive to the freezing treatment that are annotated as unknown proteins; different genes are likely to be involved in the response to freezing temperatures and lack of studies in this area have to date prevented their identification. A preliminary microarray study found that the greatest increase among over 700 genes differentially regulated in response to subzero temperatures (-3^{o}C for at least one day) was for genes of unknown function (unpublished – referred to in Livingston III *et al*., 2007), further supporting presence of a distinct set of as yet uncharacterised freezing responsive genes. Study of the wheat transcriptome following exposure to subzero temperatures resulted in extensive gene expressional changes; 25% of these responsive genes were also of unknown function (Skinner, 2009).

5.4.5.3 Genes down-regulated in response to freezing

There were very few genes deemed to be significantly down-regulated in response to the additional freezing treatment: only ten in total. Over-representation of gene ontologies related to metabolism and catabolism indicate that freezing stress causes further strain on the ability of the plant to perform normal metabolic functions;

particular processes may be sacrificed in order for freezing tolerance mechanisms to be implemented or maintained. *EBF2* is one of the most down-regulated genes; this gene is involved in the down-regulation of ethylene signalling and may be important in the rapid recovery of plants following an abiotic stress that has caused ethylene accumulation (Mineko and Shuichi, 2008). Other down-regulated genes are involved in transcription and photosynthesis suggesting that transcriptional regulation of gene expression occurs in response to freezing stress as well as acclimation, and that key metabolic processes such as photosynthesis are also further affected.

5.4.6 Differential expression of genes between *A. l. petraea* study populations

In concordance with the ecological and metabolic characteristics known to differ between populations of *A. l. petraea* (e.g. Vergeer *et al.*, 2008; Davey *et al.*, 2008; Davey *et al.*, 2009), a set of 207 genes was shown to be differentially expressed between the two study populations, irrespective of treatment ($P < 0.05$, fold change >2). Genes within this set were either expressed more highly in Helin than Leitrim, or vice versa, and the differentially regulated genes cover a wide range of functions. The two study populations are geographically distinct and experience vastly different conditions in the field; Helin, persisting at high altitude, experiences the more extreme climate with an average yearly site temperature of 4.5^{o}C, whilst Leitrim, a low altitude population, experiences a much milder climate with an average yearly site temperature of 8.8^{o}C. Helin experiences 23 degree days where temperatures are below minus five and reaches minimum nightly winter temperatures of below -16^{o}C, whilst at the Leitrim site such cold temperatures do not occur; these extreme climatic differences will encourage adaption to the locally prevalent conditions. Gene ontologies relating to response to stimuli are over-represented among those genes that are expressed more highly in Leitrim than Helin. This suggests that Leitrim is either able to maintain higher levels of expression of these genes, perhaps due to the greater availability of energy in the warmer climate, or that higher levels of expression of such genes are required to confer the same level of response or protection as in Helin. *BT2*, a gene studied in a range of *A. l. petraea* populations both prior to and following a chilling stress (see Chapter 4), is expressed more highly in Leitrim than Helin in these microarrays. This is in concurrence with data presented in Chapter 4; populations that experience colder climatic conditions in the field exhibited lower expression of this gene, possibly due to maintained levels of ABA in these populations as discussed in 4.4.3.

Those genes that are more highly expressed in Helin than Leitrim cover a wider range of gene ontologies, but in particular incorporate secondary metabolism, and cellular and physiological processes. Population specific differences in metabolite production, in both control conditions and following exposure to cold temperatures, have been reported for *A. l. petraea* (Davey *et al*., 2008; Davey *et al.*, 2009), hence, it is not surprising that genes involved in the production of metabolites are differentially regulated between the populations. In addition to the secondary metabolism gene ontology, the flavonoid biosynthesis pathway is also over-represented in this gene set; both of these results are indicative of a greater requirement for secondary metabolites such as flavonoids in the Helin population. This may be caused by the much more frequent exposure to cold temperatures experienced by this population; it may be beneficial for plants to maintain higher levels of metabolites that will consequently confer greater protection against cold temperature stress. Over-representation of genes involved in cellular and physiological processes in this gene set is indicative of differential regulation of core processes between the Helin and Leitrim populations. Exposure to different climatic conditions will require different physiological and cellular adaptations in order for populations to persist. The presence of many different transferases in the set of genes expressed more highly in Helin than Leitrim is suggestive of differential biochemical requirements within each population; Helin may require varying types or proportions of molecules than Leitrim in order to maintain normal physiological function under the different climatic conditions experienced. The higher expression of many physiological and cellular genes in Helin may suggest that either i) this population is better adapted to maintain core processes under the cold climatic conditions experienced, or ii) greater expression levels of these genes are required in Helin to confer comparable levels of core physiological and cellular processes to those for Leitrim.

5.4.7 Differential regulation of genes between *A. l. petraea* study populations in response to cold temperatures.

The set of genes termed the interaction genes (469 genes) were responsive to the acclimation treatment but were also differentially regulated between the two study populations in response to acclimation. Of the six different possible types of interaction (Figure 5.4.), two interaction types accounted for almost half of all the interaction genes. These were i) genes down-regulated in Helin yet up-regulated in Leitrim following acclimation, and ii) genes down-regulated in both populations following acclimation but

to a greater extent in Helin. For many of these genes, expression levels were far higher in Helin than Leitrim prior to the cold temperature treatments (i.e. under control conditions), whilst expression levels of the two populations became much more comparable after the acclimation treatment. It seems sensible to suggest that the higher expression of these cold temperature responsive genes in Helin, prior to cold temperature exposure, may relate to the more frequent exposure of Helin to cold temperatures in the field. Maintaining higher expression of these genes under warmer temperatures may be beneficial to Helin as it is then primed ready for cold exposure, whilst this may not be an efficient use of energy resources for Leitrim as cold exposure is much less frequent. This, however, does not explain the down-regulation of these genes in Helin following an acclimation treatment; if expression of these genes is beneficial under cold conditions then why would expression be reduced upon exposure to cold conditions? The control conditions used in this experiment were chosen based on published cold stress research to ensure that suitable comparisons between results could be made; for published *Arabidopsis* work, a control treatment using approximately 20^{o}C for the day temperature has frequently been used (Hannah *et al.*, 2005; Le *et al.,* 2008). Temperature data collected at the specific field locations of the two study populations (P. Vergeer, *pers. comm.*) reveals that, although both populations do experience temperatures of approximately 20^{o}C during the summer months, the difference between the average daily and nightly temperatures is greater for Helin than for Leitrim during this time. It is therefore possible that Helin is responding to the control treatment used in this investigation, perhaps as reaction to continuously warm temperatures, whilst the control conditions are more comparable to field conditions experienced by Leitrim. This could explain the higher levels of gene expression observed for Helin under the control conditions; Helin may be regulating gene expression in response to the persistently warm temperatures. Following exposure to acclimation temperatures, the levels of expression of many of these genes are similar for Helin and Leitrim indicating that acclimation may involve a specific transcriptomic response with particular levels of gene expression required. It is also possible that Leitrim is responding to the cold temperatures by up-regulating the expression of many genes whilst Helin, which in the field is exposed to cold temperatures much more frequently, is returning to its more basal levels of gene expression. To clarify these observed patterns, it would be beneficial to explore expression of some of the interaction genes in both Helin and Leitrim under a wider range of temperature regimes. One obvious consideration would be that of the day : night temperatures; incorporating

a greater distinction between these two temperatures, for example 20^{o}C day : 10^{o}C night, may help to identify whether or not Helin is responding to the continuously warm temperature conditions in these control arrays.

Within the interaction genes a number of gene ontologies are over-represented. These include response to stimulus, and more specifically response to stress and defence response. As genes within this set are differentially regulated between the two study populations in response to cold temperatures, the over-representation of these categories suggests populations perceive temperature in different ways, and consequently that populations are adapted to local environmental conditions. Differing climatic conditions, for example the much colder temperatures experienced in the field by Helin compared to Leitrim, will require differing levels of protection and response. A range of differentially regulated genes are discussed below.

5.4.7.1 Transcription factors

Many differentially regulated interaction genes are transcription factors, and will thus be involved in the transcriptional regulation of other genes within the plant. A large proportion of these differentially regulated transcription factors are expressed more highly in Helin than Leitrim under the control conditions but demonstrate similar levels of expression in response to acclimation in the two populations. Many of these transcription factors are up-regulated in Leitrim following acclimation, suggesting that they are required in the acclimation response of *A. l. petraea*. Several WRKY transcription factors demonstrate these expression patterns, including *WRKY40* (At1g80840) and *WRKY18* (At4g31800). The WRKY family is a large family of transcription factors exclusive to plants, with up to a hundred genes in the *Arabidopsis* genome; genes are divided into subgroups depending on the structure of their WRKY domains (Eulgem *et al.*, 2000). These genes are known to be induced by pathogens, and to be expressed in the vascular bundles of plants, and thus have been implicated in the defence signalling response (Chen and Chen, 2002; Dong *et al.*, 2003). WRKY mRNA accumulation is rapid indicating that genes may be involved in the regulation of secondary response genes (Eulgem *et al.*, 2000). Some WRKY genes have been implicated in the biosynthesis of secondary metabolites (Eulgem *et al.*, 2000), and in the regulation of expression of other regulatory genes including protein kinases (Singh *et al.*, 2002). A number of other transcription factors also demonstrate similar expression patterns to the WRKY transcription factors in this investigation; these include *ZAT10*

(At1g27730), *ERF4* (At3g15210) and *ERF8* (At1g53170), and *CDF3* (At3g47500). Up-regulation of *ZAT10* has been demonstrated in *Arabidopsis* in response to a range of abiotic stresses including drought, high salt, and cold, and it has been implicated in the regulation of stress down-regulated genes such as those involved in photosynthesis and carbohydrate metabolism (Sakamoto *et al.*, 2004). ERF (ethylene response element binding factor) genes, like WRKY genes, are unique to plants and are defined by their AP2 / ERF domains; most ERF proteins are activators of transcription, although some act to repress transcription (Singh *et al.*, 2002). *ERF4* has been implicated in the negative regulation of ethylene, jasmonate, and ABA responsive genes (Yang *et al.*, 2005), and, as *ERF8* is a member of the same sub-group of ERF genes as *ERF4* (Nakano *et al.*, 2006), it is possible that both genes may be involved in the transcriptional repression of genes during abiotic stress. *CDF3* (CYCLING DOF FACTOR 3) is one of 37 members of the Dof family of transcription factors in *Arabidopsis*. This family function as transcriptional activators and repressors and, whilst many have not been characterised, they are thought to have involvement in a wide range of plant physiological functions (Yanagisawa, 2002). *CDF3* has previously been reported as cold responsive in *Arabidopsis* (Kreps *et al*., 2002; Vogel *et al*., 2005) and thus it seems feasible to suggest that it has involvement in the regulation of secondary gene expression in response to temperature stress.

5.4.7.2 Protein kinases and phosphatases

Many of the acclimation interaction genes are annotated as protein kinases and phosphatases; these function in the modification of specific proteins and consequently result in signal transduction (Schenk and Snaar-Jagalska, 1999). Interestingly, these genes exhibit a range of different interactions or expression patterns in the two populations including, for the protein kinases: down-regulation in Helin with up-regulation in Leitrim (At4g08850 and At2g05940); down-regulation in both populations but by a greater extent in Helin (At3g02880); up-regulation in Helin with down-regulation in Leitrim (At1g51940); and up-regulation in both populations but by a greater extent in Leitrim (At2g43850). This range of interactions, unlike that for the transcription factors detailed above, reflects the diverse signalling pathways that these proteins are involved in and may also indicate the differential requirements of the two study populations for protein modification following cold temperature exposure. The sessile nature of plants requires external stimuli to be recognised and for appropriate intracellular responses to be subsequently initiated; protein kinases and phophatases

play a vital role in the transduction of such signals (Stone and Walker, 1995; Schenk and Snaar-Jagalska, 1999). The differential levels of expression of these genes are indicative of adaptation to environment in the *A. l. petraea* populations; very diverse expression patterns are observed under uniform laboratory conditions implying dissimilar signal transduction pathway requirements.

5.4.7.3 Calcium signalling

Calcium levels are known to increase dramatically in plants following exposure to a range of abiotic stresses including cold temperatures (Knight *et al.*, 1991; Knight *et al.*, 1996). The precise kinetics, magnitude of increase, and cellular source of calcium have all been proposed to contribute to the specificity of the response elicited, thus enabling plants to respond accordingly to specific stimuli (Knight and Knight, 2001). Perception of cold temperatures has long been linked to the fluidity state of the plasma membrane (Levitt, 1980) and, more recently, rigidification of the membrane has been shown to result in calcium influx (Örvar *et al.*, 2000; Sangwan *et al.*, 2001). Cellular calcium is known to regulate many proteins, and is essential for the cold acclimation response and cold-induced gene expression in *Arabidopsis* (Tähtiharju *et al.*, 1997). Many genes involved in calcium binding are present in the acclimation interaction gene set (e.g. At1g18210 and At1g08450), as well as a range of calcium regulated effector proteins (e.g. At2g41410 and At5g66210). Calcium binding proteins have shown up-regulation in *A. thaliana* at time points soon after cold exposure (Lee *et al.*, 2005). Effector proteins include calmodulin; this binds calcium and consequently interacts with enzymes and kinases to mediate cellular responses to calcium fluxes and thus to abiotic stimuli (Braam and Davis, 1990). The TCH gene family in *Arabidopsis* is a calmodulin-related gene family; several TCH genes have been reported to respond to cold temperatures and intracellular calcium has been implicated as the signalling molecule involved (Polisensky and Braam, 1996). In addition to the calcium binding proteins and effector proteins, *TCH3* (At2g41100) and *TCH4* (At5g57560) are also present within the acclimation interaction gene set and are thus differentially expressed between the study populations, Helin and Leitrim. The differential regulation of a wide range of genes related to calcium signalling, both in control conditions and in response to cold temperatures, is indicative of differences in the perception of temperature between the two populations.

5.4.8 Conclusions

Cross species transcriptomics, using Affymetrix ATH1 GeneChip® arrays, have been shown to be a viable method for the detection of real changes in gene expression both within and between populations of *A. l. petraea*. Following a two day acclimation at 4°C many changes in gene expression were observed within the two study populations. The majority of these genes have previously been reported as cold responsive, and thus expression patterns observed in *A. l. petraea* are in concordance with published *A. thaliana* literature; this indicates that cold response pathways are conserved between the two species. The additional freezing shock resulted in further expressional changes of 75 genes; transcriptomic changes following freezing temperatures have not previously been reported in the *Arabidopsis* literature. These data imply that there are a number of genes involved in response to freezing temperatures in *Arabidopsis*; many of these genes are annotated as unknown proteins and thus have not been characterised. This is in concordance with transcriptomic responses of wheat to subzero temperatures; 25% of expressional changes were in genes of unknown function (Skinner, 2009). One of the genes exhibiting the greatest fold change response to freezing in this investigation, *CBF4*, has not previously been demonstrated as cold responsive despite over-expression of this gene resulting in the expression of known cold-inducible genes and in increased freezing tolerance in *Arabidopsis* (Haake *et al.*, 2002; Sakuma *et al.*, 2002). In these arrays, differences in gene expression between the two study populations of *A. l. petraea* were possible to detect, both under control conditions and in response to temperature treatments. Differential gene expression between the populations could be indicative of adaptation to local environments, and these two populations experience vastly differing climatic conditions in the field. Differences in the metabolome between populations of *A. l. petraea* have previously been reported (Davey *et al.*, 2008; Davey *et al.*, 2009), and the transcriptomic data presented here are in accordance with these results. Of the sets of differentially regulated genes identified in this investigation, those termed the interaction genes are possibly of the greatest biological interest. The interaction genes are those genes responsive to the temperature differences, but exhibiting differing expression patterns between the two study populations and are thus more likely to infer local adaptation. Many of the interaction genes are involved in signalling and transcriptional regulation, and consequently point towards adaptation of the populations to locally experienced conditions.

CHAPTER 5: SUMMARY

- Affymetrix ATH1 GeneChip® arrays were successfully employed in the detection of significant changes in gene expression in populations of *A. l. petraea* in response to both acclimation and freezing stresses. Genes were also differentially expressed between the two populations under investigation, both in response to cold temperatures and under control conditions.
- Many of the genes responsive to acclimation were in concordance with the published literature for *A. thaliana*, including genes involved in the CBF cold response pathway.
- The additional freezing stress resulted in changes in the expression of many genes that are currently annotated as having an unknown function, indicating the possibility of a distinct set of as yet uncharacterised freezing response genes in *Arabidopsis*.
- Differential expression of genes between the populations under investigation is indicative of adaptation to local environment; the two study populations experience vastly different climatic conditions in the field.
- The genes termed 'interaction genes' are of the greatest biological interest, as they are responsive to cold temperatures yet exhibit differential expression patterns between the populations under investigation. A subset of these genes has been chosen for further investigation; results from this work are presented in Chapter 6.

6. The role of population-specific gene promoters in differential regulation of gene expression.

AIMS

- Select differentially regulated genes identified by microarray analysis with potential roles in adaptation of *A. l. petraea* populations to local environments.
- Obtain population-specific promoter region sequences for the genes of interest. Generate GUS promoter-reporter fusion constructs for the transformation of *A. thaliana* and *A. l. petraea*.
- Perform fluorometric GUS assays to quantify GUS activity in these transformed plants following an acclimation treatment, and subsequently determine the role of the promoter in the differential regulation of gene expression.

6.1 Introduction

6.1.1 Reporter gene fusions

Reporter genes are used as a versatile tool within molecular biology to investigate gene expression as their gene products possess properties, such as the ability to fluoresce, that aid their detection and quantification. Coding regions of reporter genes are fused to the regulatory sequence of the gene under investigation and expression of the gene can subsequently be readily assayed in the organism harbouring the construct. Several reporter gene systems have been developed for use in plants including ß-galactosidase from *E. coli* (Helmer *et al.*, 1984), octopine synthase from *A. tumefaciens* (DeGreve *et al.*, 1982) and the luciferase gene from firefly (Ow *et al.*, 1986); problems with endogenous expression in plants or assay accuracy have prevented wider use of these systems. The ß-glucuronidase gene (GUS) from *E. coli* has been developed as a reporter gene system in plants (Jefferson *et al.*, 1987; Jefferson, 1989). This gene encodes an enzyme that catalyses the cleavage of a range of commercially available substrates, and its activity can thus be assayed using spectrophotometric, fluorometric, and histochemical techniques. There is no endogenous ß-glucuronidase activity in higher plants (Jefferson *et al.*, 1987), therefore, background expression does not have to be considered when assaying transgenic plants. In addition, expression of this gene in plants does not confer any negative effects on growth or reproduction (Jefferson *et al.*, 1987). Sensitive histochemical assays can be easily and cheaply performed and these

enable GUS activity to be localised to particular cell types (Jefferson *et al.,* 1987). Diffusion of the reaction product has prompted concern as to the specificity of localisations, but alterations to incubation times and the inclusion of K-ferricyanide can improve this (Guivarc'h *et al.,* 1996).

Green fluorescent protein (GFP), isolated from the jellyfish *Aequorea victoria*, is another successful reporter gene system. This protein was first expressed in *E. coli* and the nematode *C. elegans* (Chalfie *et al.,* 1994) and the system has also subsequently been developed for use in plants (Haseloff and Amos, 1995). Advantages of this reporter system stem from the fact that the protein is itself fluorescent. This means that expression can be studied in living tissue as there is no requirement for destructive enzymatic assays, and problems such as reaction product diffusion are eliminated. Drawbacks to this system include slow formation of the fluorophore in relation to the initiation of gene expression. The fluorescing protein may also still be present after gene expression has ceased due to its stable nature (Guivarc'h *et al*., 1996).

6.1.2 Use of the GUS reporter system in plant studies

The GUS reporter system is of enormous value in plant science and this is illustrated by the rapid increase in the number of published studies using this system since 1989 (Guivarc'h *et al*., 1996). Following identification of a differentially regulated gene through molecular methods (e.g. qRT-PCR or microarray analysis), a more complete depiction of how that gene is affected by the experimental conditions can be obtained using GUS reporter analysis. The system can be used to investigate localisation of gene products or to study the activity of specific gene promoters as selected by the researcher. The promoter activity of a wide range of genes has been investigated in *A. thaliana* using this system. The promoter of *RD29A*, a gene responsive to desiccation, cold, and salinity, was shown to drive GUS expression in transgenic *A. thaliana* plants in response to desiccation; histochemical assays revealed promoter activity in almost all organs and tissue (Yamaguchi-Shinozaki and Shinozaki, 1993). The CBF/DREB1 transcription factors (*CBF1*, *2*, and *3*) play an important role in cold acclimation. Following identification of these factors, Novillo *et al.* (2007) used the GUS reporter gene system to determine whether or not these three factors are functionally equivalent. GUS activity, driven by the promoters for the three genes, revealed that expression of *CBF1* and *CBF3* differed considerably from that of *CBF2* during plant development and in response to cold (Novillo *et al.,* 2007). Expression of *Arabidopsis* ACC synthase

genes (ACS), a multi-gene family with roles in ethylene production, has also been investigated using the GUS reporter system. Histochemical assays revealed that members of this family were localised to different regions of the plant during development, and fluorometric assays demonstrated differential regulation in response to a range of hormones and abiotic stresses (Wang *et al.*, 2005).

6.1.3 The role of promoters in the regulation of gene expression

The products of stress inducible genes can be classified into two groups: those with direct roles in stress tolerance and those that are involved in signal transduction and regulation of gene expression. Regulation of gene expression in response to stress involves several components, and the promoter regions of genes are of key importance. Both abscisic acid (ABA) dependent and ABA-independent systems are involved in the abiotic stress response (Thomashow, 1999; Shinozaki and Yamaguchi-Shinozaki, 2000). Elements, known as *cis*-acting elements, are present in the promoter sequences of genes. These elements interact with transcription factors (*trans*-acting factors) to form a transcriptional initiation complex, and subsequently these interactions function as switches and initiate transcription of the stress-inducible gene (Yamaguchi-Shinozaki and Shinozaki, 2005).

A number of *cis*-acting elements have been identified that have roles in the response of plants to both cold and osmotic stress. One of the major regulatory elements functions in ABA-dependent gene expression. Promoter regions of a number of ABA-dependent genes were studied, leading to the identification of a conserved sequence referred to as ABA responsive element (ABRE); this was first identified in wheat (Guiltinan *et al.*, 1990) and rice (Mundy *et al.*, 1990). A single copy of ABRE within the promoter is not sufficient for ABA-responsive gene expression and it has been demonstrated that promoter regions must either contain two ABRE sequences or additional coupling elements (Yamaguchi-Shinozaki and Shinozaki, 2005). DRE/CRT is a major *cis*-acting element that functions in ABA-independent expression. The proteins encoded by CBF/DREB1 and DREB2 transcription factors bind to the DRE/CRT element in the promoter regions of downstream genes, thus activating their transcription (Stockinger *et al.*, 1997; Liu *et al.*, 1998). These downstream genes are commonly referred to as *COR* (cold-responsive) genes (section 1.5.1). CBF/DREB1 genes are cold responsive, and their expression is controlled by the transcription factor *ICE1* (INDUCER OF CBF EXPRESSION1) (Chinnusamy *et al.*, 2003). The CBF/DREB1 genes also have *cis*-

acting elements in their promoters that function to induce gene expression in response to cold (Zarka *et al.*, 2003). A wide range of *cis*-acting elements are known to have critical roles in the regulation of gene expression in response to various abiotic stresses, and these have resulted in the identification of complex regulatory networks. Alterations to these networks, such as the loss of particular *cis*-acting elements, could potentially be important in the adaptation of plants to locally prevailing environmental conditions.

Here, a number of genes identified to be differentially regulated between populations of *A. l. petraea* in response to cold acclimation (chapter five) are investigated further. The key biological objective of this work was to determine whether or not population-specific promoters are responsible for this observed differential regulation of gene expression. In addition, sequence differences between gene promoters from different populations of *A. l. petraea* could be identified and analysed with reference to known *cis*-acting regulatory elements. Following study of the gene promoter sequences for both Helin and Leitrim, population specific promoters were fused to the GUS reporter gene. These constructs were used to transform *A. thaliana* and *A. l. petraea*. Fluorometric GUS assays were used to quantify GUS activity in untreated transformants and in transformants that had been exposed to cold acclimation conditions. GUS activity reports expression of the population-specific gene promoters and thus the potential role of these in regulation of gene expression can be discussed.

6.2 Materials and Methods

6.2.1 Selection of genes for study

Microarray analysis was performed using Affymetrix ATH1 GeneChip® arrays to investigate gene expression in two populations of *A. l. petraea*, Helin in Norway and Leitrim in Ireland. Treated plants were exposed to a 48 h acclimation period at 2°C, or an 48 h acclimation period at 2°C followed by a freezing shock of -9°C. Results from these arrays are presented in chapter five. A set of genes were identified that were responsive to the acclimation treatment but were also differentially regulated between the two populations; these were termed 'interaction genes' (section 5.3.3.4). Table 5.6 details those genes that had the greatest magnitude of difference between the Helin response to acclimation and the Leitrim response to acclimation. Five of these genes were selected for further study: At5g11740, At5g42050, At1g18210, At1g09070, and At4g31800. Biological functions of the chosen genes are summarised in Table 6.1. Figure 6.1 illustrates expression values for these five genes in the untreated plants and in plants that experienced the 48 h acclimation.

6.2.2 Primer design for cloning of promoter regions

Primers to amplify promoter regions of the genes chosen for investigation were designed to the sequence of *Arabidopsis lyrata lyrata* (the subspecies from central and eastern North America, *A. l. lyrata* hereafter) that has recently become publicly available; the sequence data is not currently annotated and is available in the form of contigs and larger assembled scaffolds (sequence available at http://genome.jgi-psf.org). The *A. l. lyrata* sequence was BLAST searched using the *A. thaliana* coding sequence for each gene of interest. Due to high sequence similarity of the coding regions the *A. l. lyrata* orthologues of the *A. thaliana* genes could be identified. The promoter region upstream of the *A. l. lyrata* coding sequence was then used for primer design. The reverse primer was designed immediately upstream of the start codon. Forward primers were designed between 1500 and 2000 base pairs upstream of the start codon where possible as it was expected that elements involved in regulation of gene expression would be located within this region of promoter sequence. Primer pairs for the gene At5g42050 were designed to amplify shorter products as the promoter region for this gene in *A. thaliana* is less than 1000 bp. Two forward primers were designed per gene as percentage identity between *A. l. lyrata* and *A. l. petraea* sequences was unknown and primer sequence specificity could not, therefore, be guaranteed. These

promoter sequences were checked for absence of suitable restriction enzyme sites that were added to the primer sequences. Three random bases (ACA in all primers) were included 5' of the restriction site to ensure efficient digestion of amplified DNA fragments. Primer sequences are detailed in Table 6.2.

6.2.3 Cloning of *A. l. petraea* promoter regions

A. l. petraea genomic DNA, obtained individually from both Helin and Leitrim leaf material, was PCR amplified (section 2.3.8), products analysed (section 2.3.9), PCR purified (section 2.3.11), and precipitated (section 2.3.12). Following digestion with the appropriate restriction enzymes, these population specific DNA fragments were then cloned into the pBluescript vector (pre-digested with the same restriction enzymes as the DNA fragments) as detailed in section 2.3.14.2 and transformed into competent DH5α *E. coli* (section 2.3.17). Plasmid DNA was extracted from *E. coli*, digested using primer specific restriction enzymes, and analysed by agarose gel electrophoresis (sections 2.3.21, 2.3.15, and 2.3.9 respectively). Colonies containing an insert of the correct size were sequenced (section 2.3.13) and consequently the *A. l. petraea* promoter region sequences for both Helin and Leitrim were obtained for each gene of interest.

Recombinant plasmids, harbouring the desired insert for either Helin or Leitrim, were digested with restriction enzymes, and fragments were cloned into a pBI101 GUS reporter vector (pre-digested with the same restriction enzymes), as detailed in section 2.3.14.2. Competent DH5α *E. coli* cells were transformed with these constructs (section 2.3.17). Plasmid DNA was isolated from transformed *E. coli* (section 2.3.21) and recombinant plasmids containing an insert of the correct size were sequenced to confirm cloning of the desired product. Verified plasmids were used to transform competent *A. tumefaciens* GV3101 cells (section 2.3.19). Successfully transformed *A. tumefaciens* colonies, were used for the transformation of both *A. thaliana* and *A. l. petraea* plants.

6.2.4 Transformation of *A. thaliana* and *A. l. petraea*

All constructs, containing either the Helin promoter region or the Leitrim promoter region, for the genes of interest were used to transform flowering *A. thaliana* plants (section 2.3.22). Seeds were collected and screened for transformants in tissue culture (section 2.3.23).

These constructs were also used to transform *A. l. petraea* plants that had been grown in tissue culture (sections 2.3.24 and 2.3.25). Plants from both the Helin and the Leitrim populations were grown. Each construct was used to transform both Helin and Leitrim plants. As transformants developed, they were transferred to ½ MS10 media containing 50 µg ml^{-1} kanamycin and 250 µg ml^{-1} cefotaxime. Individual transformants were divided, or subcultured, regularly to enable a number of clones within one transformed line to be developed.

AGI Code	Interaction	TAIR Annotation and relevant publications
At5g11740	Up-regulated in both populations, to a greater extent in Leitrim	**ARABINOGALACTAN PROTEIN 15 (*AGP15*).** Encodes peptide with small protein backbone. Expression profile has not been characterised (Schultz *et al.*, 2000).
At5g42050	Up-regulated in both populations, to a greater extent in Leitrim	**Unknown protein.** Expression induced by touch (rapid kinetics), wounding, $CaCl_2$ and $MgCl_2$ (Chotikacharoensuk *et al.*, 2006). Drought responsive in *Thellungiella* (Wong *et al.*, 2005).
At1g18210	Down-regulated in Helin, up-regulated in Leitrim	**WRKY DNA BINDING PROTEIN 18 (*WRKY18)*. Response to bacterial and fungal pathogens.** High levels of expression result in stunted growth. Expression enhances resistance to bacterial pathogen in mature plants (Chen and Chen, 2002).
At1g09070	Down-regulated in Helin, up-regulated in Leitrim	**SOYBEAN GENE REGULATED BY COLD-2 (*SRC2*). Protein binding.** Induced by a broad range of stresses. Component of the vesicle transport pathway (Ma and Bohnert, 2007). Cold responsive (Robinson and Parkin, 2008)
At4g31800	Down-regulated in Helin, up-regulated in Leitrim	**Calcium binding, putative.**

Table 6.1 Biological functions of the genes of interest involved in GUS assays. Five of the interaction genes, as identified in Chapter 5, were selected for further investigation. The column entitled 'Interaction' refers to the differential response to acclimation, identified through microarray analysis, in the two study populations of *A. l. petraea*, Helin and Leitrim. TAIR annotations for gene function are given in bold.

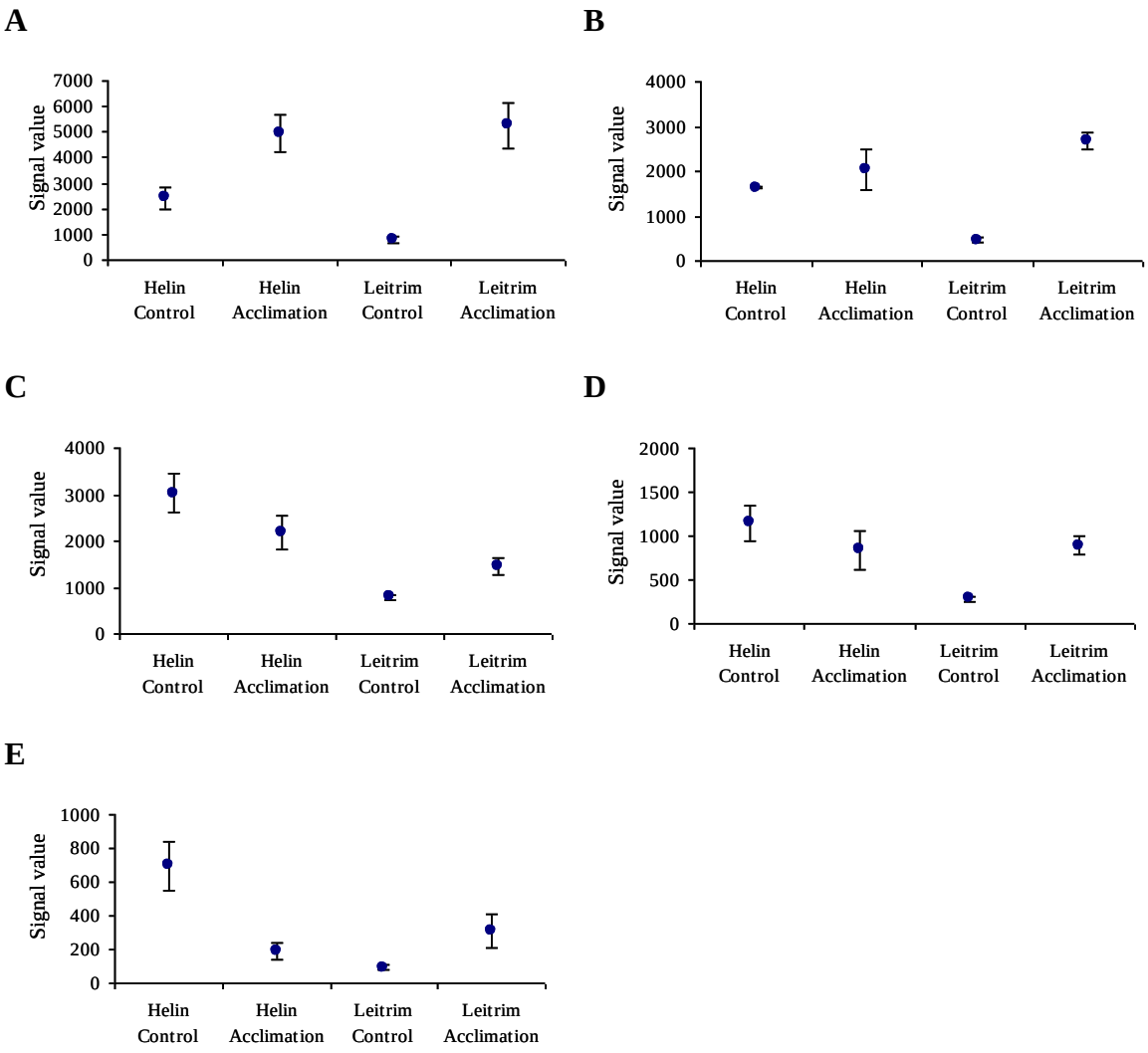


Figure 6.1 Signal values of interaction genes chosen for further investigation. Microarray analysis was performed, as described in Chapter 5, to investigate gene expression in two populations of *A. l. petraea*, Helin in Norway and Leitrim in Ireland. A number of genes were found to be responsive to an acclimation treatment (2^{o}C for 48 h), but were differentially regulated between the two populations. These genes have been termed 'interaction genes'. Five, shown here, were chosen for further investigation; genes exhibiting the greatest magnitude of difference between the Helin and Leitrim response to acclimation were selected (Table 5.6). Data points refer to the mean signal values for each category across the three replicate arrays (±SE). A) At5g11740; B) At5g42050; C) At1g18210; D) At1g09070; and E) At4g31800.

Gene (AGI Code)	Forward primer 5' – 3'	Reverse primer 5' – 3'	Product size (bp)
At5g11740	ACAAAGCTTTCG-TAA-CAA-GTT-AGA-GGA-TA	ACAGGATCCTTC-AAA-TAG-TAA-AGA-GAT-TTG	1960
	ACAAAGCTTCAA-GAG-CCT-GCC-AAC-GTC-TA		1640
At5g42050	ACAGTCGACATC-CAT-CTC-TTT-CTC-TGT-GA	ACAGGATCCCTC-TTC-TCT-GAT-TGA-TCT-GG	927
	ACAGTCGACTCT-CAA-ACT-TAA-TGG-GCT-TA		824
At1g18210	ACAAAGCTTAGT-AAT-TAG-CCA-TGC-GTC-AT	ACATCTAGATGC-CAT-TTT-TTT-TTT-GTT-AAA-GGT	2000
	ACAAAGCTTGAT-GGG-AAA-AGA-GAA-CAC-AT		1600
At1g09070	ACAAAGCTTAGG-AGA-ATA-AAG-TGG-CAA-AG	ACATCTAGAGGT-TGG-TTA-CTC-GGA-TCG-AA	1600
	ACAAAGCTTACC-AAG-AAA-GAT-GCT-TGA-GA		1400
At4g31800	ACAAAGCTTACG-GAT-CCA-CGG-TAT-TGA-AT	ACATCTAGAAAA-ATA-AAT-CTA-TAA-GTT-ATC	2000
	ACAAAGCTTTTA-CGC-AGT-CTC-TTT-TAT-GA		1500

Table 6.2 Primer sequences used in the cloning of *A. l. petraea* interaction gene promoter regions. Two forward primers were designed for each gene to amplify products of differing sizes. Sequences shown in green identify the forward primers with which products were successfully amplified; these products were used in cloning. Restriction sites, highlighted in grey, and three additional random bases were added to the 5' end of each primer.

6.2.5 GUS histochemical analysis to select transformed lines for further work

One leaf from each transformed line of *A. thaliana* or *A. l. petraea* was assayed for GUS activity to confirm successful transformation and to qualitatively assess expression of the GUS gene in each line. Leaves were collected in individual tubes and placed at 4^{o}C for 2 h prior to incubation at 37^{o}C in 5-bromo-4-chloro-3-indolyl-glucuronide (X-Gluc, Melford). The 20 mg ml^{-1} X-Gluc stock was made in *N*,*N*-dimethylformamide and diluted to a working concentration of 1 mg ml^{-1} in 0.1 M sodium phosphate buffer pH 7.0; 10 mM EDTA, 0.1% Triton X-100 and 2 mM each potassium ferricyanide and potassium ferrocyanide. ß-glucuronidase cleaves the substrate X-Gluc to produce glucuronic acid and chloro-bromoindigo. Components of the sodium phosphate buffer catalyse dimerisation of the soluble monomer chloro-bromoindigo, resulting in an insoluble blue dimer that is localised to regions of enzyme production. Stained leaves for all lines of a particular construct were put in order from least blue to most blue, qualitatively representing increasing levels of GUS expression, and three lines from the middle of the range were selected for further study. Lines with very high or very low GUS expression were consequently avoided. The numbers of available lines from the *A. l. petraea* transformations were far fewer than for *A. thaliana*, and thus all available lines for the latter species were used in subsequent experiments.

6.2.6 Cold acclimation treatment of *A. thaliana* and *A. l. petraea* transformed lines

6.2.6.1 A. thaliana growth

For each construct, seed was sown from the three chosen lines (section 6.2.5, Table 6.3). Aliquots of *A. thaliana* seed were sterilised (section 2.3.23) and stored in sterile distilled water at 4^{o}C overnight prior to sowing. Seeds were plated on to ½ MS10 media (0.2% (w/v) Murashige and Skoog basal medium including vitamins (Melford); 1% (w/v) sucrose; and 1% (w/v) plant agar) containing 50 µg ml^{-1} kanamycin and 250 µg ml^{-1} cefotaxime. Two plates of seed were sown per line, with twenty seed per plate; tissue was required from five plants per plate. Plates were maintained at 20^{o}C, 16 h day length, for approximately three weeks.

6.2.6.2 A. l. petraea growth

Single transformed shoots, obtained from SIM plates (section 2.3.25), were maintained on ½ MS10 media containing 50 µg ml^{-1} kanamycin and 250 µg ml^{-1} cefotaxime. Each

shoot was given a line number. When plants were of a sufficient size and had developed root systems, they were split. This process enabled a number of clones to be developed within each line. Two plates, each containing at least three individuals, were required for each line before cold acclimation was performed. Following subculturing, plants were maintained at 20^{o}C, 16 h day length for approximately three weeks. Lines used are detailed in Table 6.3.

6.2.6.3 Cold acclimation and tissue collection

One plate for each *A. thaliana* or *A. l. petraea* transformed line was transferred to 4^{o}C for 48 h under a 16 h day length regime. The second plate for each line remained at 20^{o}C during this time period. Tissue samples were collected at the start of the day cycle (section 2.3.1). Whole seedlings were harvested for *A. thaliana*, and individual leaves from *A. l. petraea*.

6.2.3 Protein extraction and quantification

Plant tissue, from either *A. thaliana* or *A. l. petraea*, was ground in 300 µl GUS extraction buffer (50 mM sodium phosphate pH 7.0; 10 mM EDTA; 0.1% SDS (w/v) and 0.1% Triton X-100 (v/v), containing 10 mM ß-mercaptoethanol); fine sand was added to the sample to aid homogenisation. Cell debris was pelleted by centrifugation in a chilled microcentrifuge at full speed for ten minutes. Supernatant was then transferred to a fresh microcentrifuge tube.

Concentration of protein extract, diluted 1:10 using GUS extraction buffer, was quantified using the BIO-RAD colorimetric RC DC protein assay (RC: reducing agent compatible; DC: detergent compatible) according to the manufacturer's protocol. Twenty five microlitres of diluted extract was incubated with 125 µl RC Reagent I for one minute, and 125 µl RC Reagent II was then added. Following vortexing, samples were centrifuged at 15 000 x *g* for five minutes. Supernatant was discarded by inversion of tubes on clean absorbent tissue paper. Precipitate was then dissolved in 125 µl Reagent A (alkaline copper tartrate) containing 2% (v/v) Reagent S (surfactant solution). One millilitre of Reagent B (dilute Folin reagent) was added to each sample, followed by immediate vortexing. Samples were incubated at room temperature for fifteen minutes before measuring absorbance at 750 nm in an ELX800 plate reader (Bio-Tek Instruments, Inc.). Protein concentration (mg ml^{-1}) was calculated using a BSA (bovine serum albumin) standard curve (0.2 mg ml^{-1} to 1 mg ml^{-1}). Samples were

assayed in duplicate, and absorbance was measured in four replicates of each standard to increase the accuracy of the standard curve.

6.2.4 Quantitative GUS assay

Assay buffer was prepared as GUS extraction buffer (section 6.2.7) with the addition of 1 mM 4-methylumbelliferyl ß-D-glucuronide (MUG; pre-dissolved in N,N-dimethylformamide). The fluorometric substrate MUG is hydrolysed by GUS to produce 4-methylumbelliferone (4-MU). For each protein sample to be assayed, 500 µl assay buffer was aliquotted into a microcentrifuge tube and pre-warmed at 37^{o}C. Reactions were set up at precise 30 second intervals; 50 µl protein extract was added to an assay buffer aliquot and mixed by pipetting. One hundred microlitres of this reaction was immediately transferred to a 900 µl aliquot of stop buffer (0.2 M Na_2CO_3): the zero time point sample. Reactions were incubated at 37^{o}C; further 100 µl aliquots were transferred to additional stop buffer aliquots after twenty, forty, and sixty minutes of incubation for *A. thaliana* samples and after thirty and sixty minutes of incubation for *A. l. petraea* samples. Stop buffer terminates the reaction and, in addition, causes 4-MU to become fully fluorescent. A standard curve was generated for 4-MU using a range of dilutions from 100 pM to 100 nM (Figure 6.2). Fluorescence of all samples and standards was measured using a POLARstar OPTIMA fluorometer (BMG LABTECH) with an excitation wavelength of 365 nm and an emission wavelength of 460 nm. Standards were assayed using four replicates of each dilution and samples were assayed in duplicate. GUS activity in each protein sample was then expressed as pmol 4-MU released min^{-1} mg $protein^{-1}$ using the 4-MU standard curve. The significance of differences between treated and untreated samples for each transformation type (see Table 6.3) was analysed using analysis of variance (SPSS v. 15.0), with temperature treatment as the fixed effect and transgenic line as a random factor.

Gene	Number of unique lines: *A. thaliana*	Number of unique lines: *A. l. petraea*	Numbered lines used in the quantitative GUS assays
At5g11740	AtH: 37 AtL: 50	AlpHH: 4 AlpHL: 8 AlpLL: 1 AlpLH: 4	
At5g42050	AtH: 48 AtL: 48	AlpHH: 4 AlpHL: 7 AlpLL: 5 AlpLH: 18	AtH: 1, 3, 7 AtL: 5, 15, 17 AlpHH: 1, 4 AlpHL: 1, 5 AlpLL: 3, 5 AlpLH: 4, 7, 14
At1g18210	AtH: 9 AtL: 11	AlpHH: 1 AlpHL: 11 AlpLL: 6 AlpLH: 11	
At1g09070	AtH: 23 AtL: 42	AlpHH: 0 AlpHL: 0 AlpLL: 5 AlpLH: 5	
At4g31800	AtH: 35 AtL: 48	AlpHH: 24 AlpHL: 18 AlpLL: 0 AlpLH: 0	AtH: 1, 12, 17 AtL: 10, 20, 21 AlpHH: 1, 3, 22 AlpHL: 1, 2, 3

Table 6.3 *A. thaliana* and *A. l. petraea* transformed lines. The constructs containing the *A. l. petraea* promoter, from either Helin or Leitrim, fused to the GUS gene were used to transform both *A. thaliana* and *A. l. petraea* plants. AtH and AtL refer to *A. thaliana* plants transformed with the Helin promoter construct or the Leitrim promoter construct respectively. The constructs were used to transform *A. l. petraea* plants from both the Helin and Leitrim populations. AlpHH and AlpHL refer to Helin plants transformed with the Helin promoter construct or the Leitrim promoter construct respectively. AlpLL and AlpLH refer to Leitrim plants transformed with the Leitrim promoter construct or the Helin promoter construct respectively. This table details the numbers of independent transformed lines obtained for each transformation type, and the numbered lines used in subsequent quantitative GUS assays.

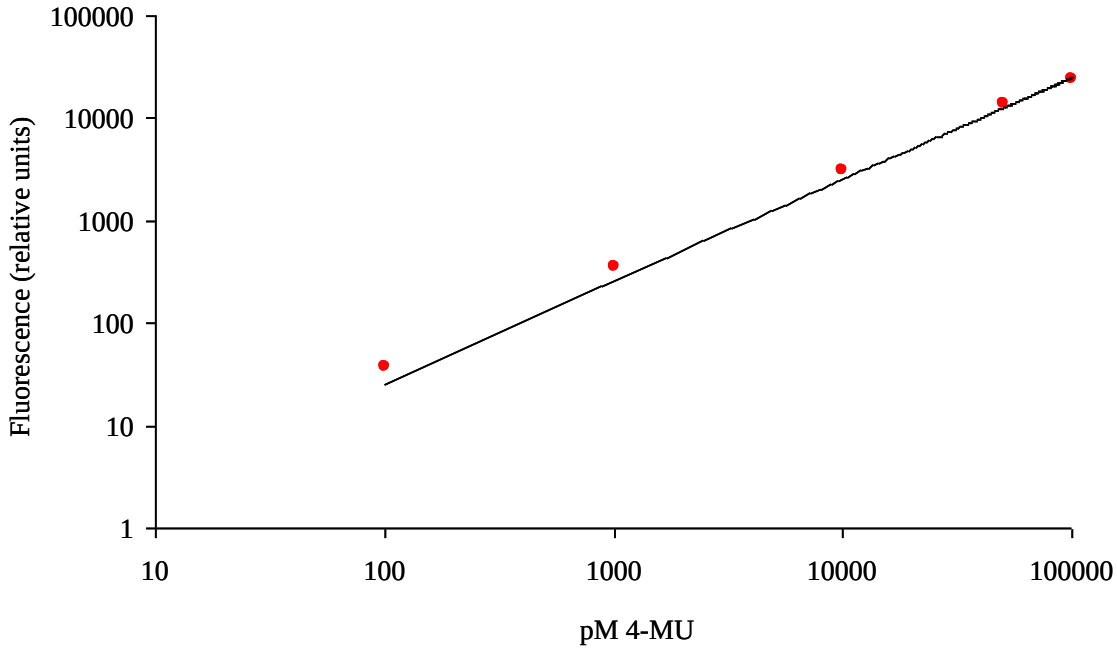


Figure 6.2 Standard curve for fluorescence of 4-MU. A standard curve was generated using a five point dilution series of 4-MU from 100 pM to 100 nM; fluorescence was measured using a PolarSTAR OPTIMA fluorometer (BMG LABTECH). Both axes are shown on a logarithmic scale. Standards were assayed using four replicates of each dilution; standard errors (+/-) fall within the size of the red markers. The equation of the straight line was used to calculate quantity of 4-MU in all unknown samples (RSq = 0.993).

6.3 Results for promoter analysis of interaction genes

6.3.1 Cloning and sequencing of the promoter regions of genes of interest: At5g11740, At5g42050, At1g18210, At1g09070, and At4g31800.

Products of the expected size, based on the *A. l. lyrata* sequence, were successfully amplified from both Helin and Leitrim genomic DNA for at least one primer pair for each gene of interest. Where both primer pairs for a particular gene successfully amplified Helin and Leitrim genomic DNA, the PCR products of the largest size were subsequently used. For At1g18210 the PCR product amplified from *A. l. petraea* genomic DNA was approximately 1000 bp and was therefore smaller than expected. The promoter sequence for this gene fell between two scaffolds of the *A. l. lyrata* sequence and thus the majority of sequence was unavailable. Following transformation, plasmid DNA from up to three transformed *E. coli* colonies containing recombinant plasmid was sequenced for each gene of interest. Figures 6.3 - 6.7 illustrate successful amplification of A. l. petraea genomic DNA and restriction digests of transformed E. coli colonies for the five genes of interest (At5g11740, At5g42050, At1g18210, At1g09070, and At4g31800).

Following confirmation of successful cloning of promoter regions into the pBluescript vector, recombinant plasmid was digested with restriction enzymes and fragments cloned into the pBI101 vector. Constructs were used to transform *E. coli* competent cells and restriction digests of transformed colonies are shown for the five genes of interest in Figures 6.3 - 6.7. Plasmid DNA isolated from these transformed *E. coli* colonies was used to successfully transform *A. tumefaciens*; transformed *A. tumefaciens* colonies were used in the subsequent transformation of *A. thaliana* and *A. l. petraea* plants.

A. l. petraea nucleotide sequences for the promoter regions (up to 2000 bp upstream of the coding sequence) of the five genes of interest are presented in Figures 6.8 - 6.12, aligned to the *A. l. lyrata* sequences for these regions. All promoter sequences had at least 85% identity with the *A. l. lyrata* orthologue sequences with the exception of At1g18210 for which the majority of the *A. l. lyrata* promoter region sequence was unavailable. *A. l. petraea* and *A. thaliana* promoter regions did not have high percentage identities (alignments not shown). The *A. l. petraea* Helin and Leitrim promoter regions for the genes of interest were found to possess sequence differences.

All identified differences were single nucleotide substitutions or indels, except for At5g42050 within which there is one inserted region in the Helin sequence of seventeen bases (at approximately base 670). These Helin and Leitrim promoter sequences were analysed using PLACE, a database of motifs found in *cis*-acting regulatory DNA elements of plants (http://www.dna.affrc.go.jp/PLACE/, Prestridge, 1991; Higo et al, 1999). Aside from the large numbers of motifs that occurred frequently within both the *A. thaliana* and *A. l. petraea* gene promoter sequences, there were a number of less common motifs conserved between all three sequences that could have potential roles in regulation of expression. These are detailed in Table 6.4. Regions of sequence differentiation between Helin and Leitrim corresponded to the presence of different promoter motifs, although the majority of these differences involved additional sites for an already common motif or concerned motifs identified in plant species other than *Arabidopsis*. Two motifs, MYB1AT and ACGTATERD1, were present in the Helin promoter sequence for the genes At5g42050 and At4g31800 respectively yet not in the corresponding Leitrim sequences; both of these motifs are annotated as relevant to genes responsive to dehydration.

6.3.2 *A. thaliana* and *A. l. petraea* transformed lines

For each gene of interest, *A. thaliana* plants were obtained that harboured either the Helin promoter:GUS construct (AtH) or the Leitrim promoter:GUS construct (AtL). Each individual plant, selected in tissue culture, constituted a unique transformed line and at least nine lines were obtained for each construct (Table 6.3). As plants used in the quantitative GUS assays were produced using seed collected from these lines, numbers of available *A. thaliana* individuals for these assays were not limiting. Large numbers of transformed *A. l. petraea* individuals within lines, however, were more difficult to obtain. As with *A. thaliana*, each transformed individual constituted a unique transformed line and the number of lines obtained varied between constructs (Table 6.3). To increase the number of individuals within a line, the original transformant had to be subcultured, or divided, a number of times. This process was dependent on plants reaching a sufficient size and required a greater amount of time than collection of seed from *A. thaliana* transformants. The minimum requirement for completion of the quantitative GUS assay was two plates per line (for the treated and untreated samples) each containing up to four *A. l. petraea* individuals that had been subcultured from the original transformant. For each gene in *A. l. petraea*, there were

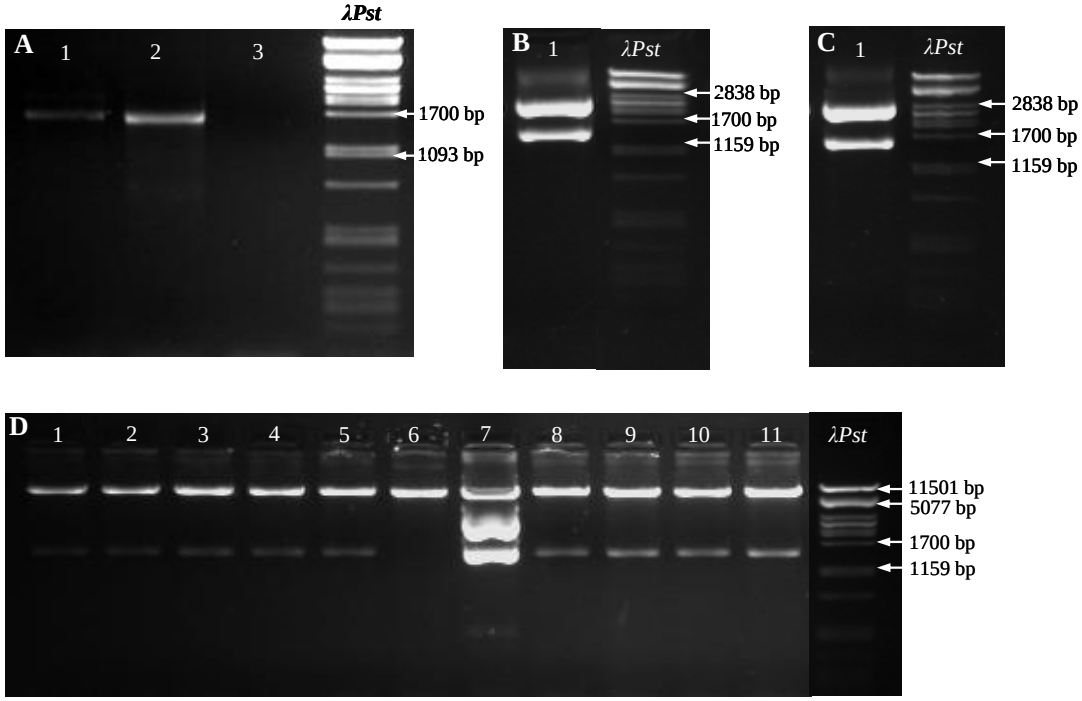


Figure 6.3 Cloning of the promoter region of At5g11740 from *A. l. petraea.* A) PCR of Helin (Lane 1) and Leitrim (Lane 2) *A. l. petraea* genomic DNA, expected size 1640 bp. Lane 3: negative control. B) and C) Restriction digests of putatively recombinant pBluescript (2958 bp) plasmid harbouring the At5g11740 promoter region amplified from Helin (B) and Leitrim (C) genomic DNA. D) Restriction digests of putatively recombinant pBI101 (12 000 bp) harbouring the At5g11740 promoter region amplified from Helin (Lanes 1-5) and Leitrim (Lanes 6-11) genomic DNA. These constructs were used for transformation of *A. tumefaciens*, prior to transformation of both *A. thaliana* and *A. l. petraea*.

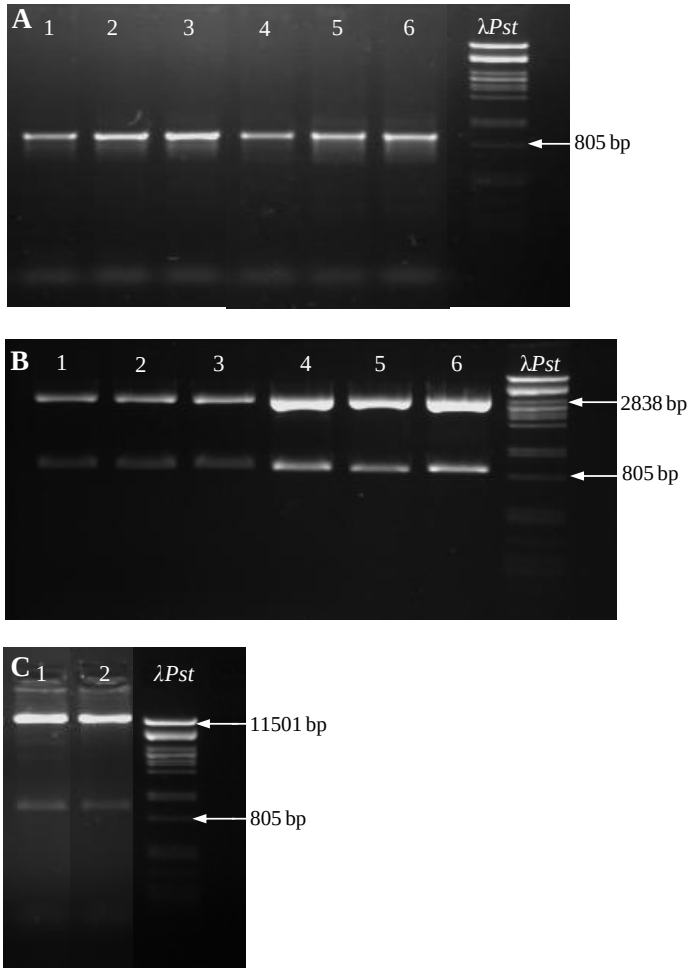


Figure 6.4 Cloning of the promoter region of At5g42050 from *A. l. petraea.* A) PCR of Helin (Lanes 1-3) and Leitrim (Lanes 4-6) *A. l. petraea* genomic DNA, expected size 927 bp. B) Restriction digests of putatively recombinant pBluescript (2958 bp) plasmid harbouring the At5g42050 promoter region amplified from Helin (Lanes 1-3) and Leitrim (Lanes 4-6) genomic DNA. C) Restriction digests of putatively recombinant pBI101 (12 000 bp) harbouring the At5g42050 promoter region amplified from Helin (Lane 1) and Leitrim (Lane 2) genomic DNA. These constructs were used for transformation of *A. tumefaciens*, prior to transformation of both *A. thaliana* and *A. l. petraea*.

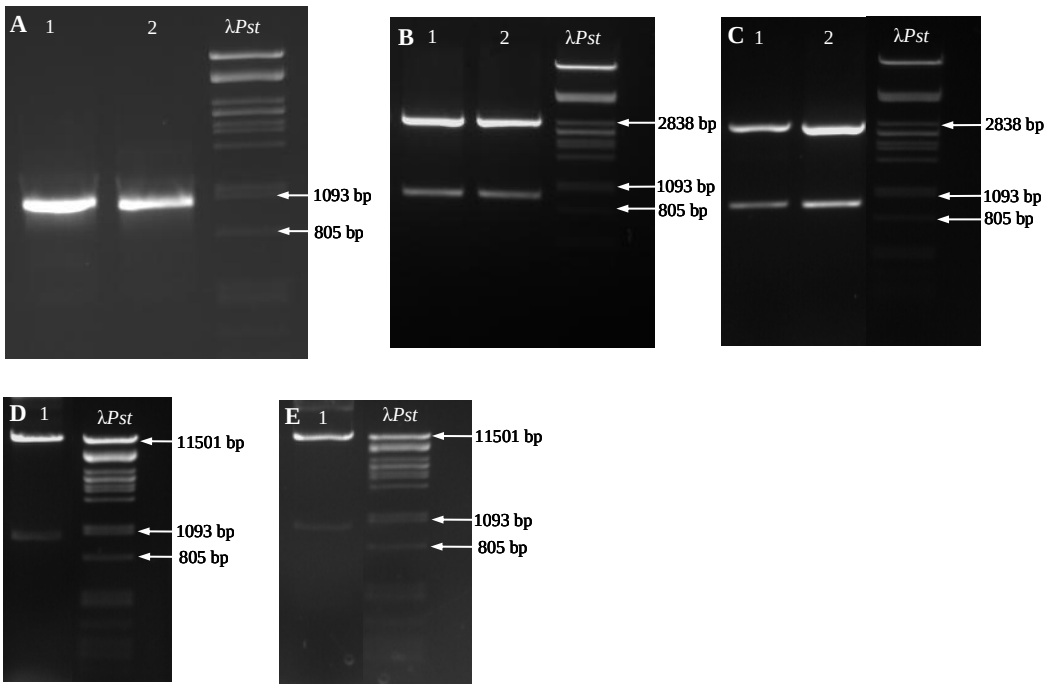


Figure 6.5 Cloning of the promoter region of At1g18210 from *A. l. petraea.* A) PCR of Helin (Lane 1) and Leitrim (Lane 2) *A. l. petraea* genomic DNA, expected size 1600 bp. B) and C) Restriction digests of putatively recombinant pBluescript (2958 bp) plasmid harbouring the At1g18210 promoter region amplified from Helin (B, Lanes 1 and 2) and Leitrim (C, Lanes 1 and 2) genomic DNA. D) and E) Restriction digests of putatively recombinant pBI101 (12 000 bp) harbouring the At1g18210 promoter region amplified from Helin (D, Lane 1) and Leitrim (E, Lane 1) genomic DNA. These constructs were used for transformation of *A. tumefaciens*, prior to transformation of both *A. thaliana* and *A. l. petraea*.

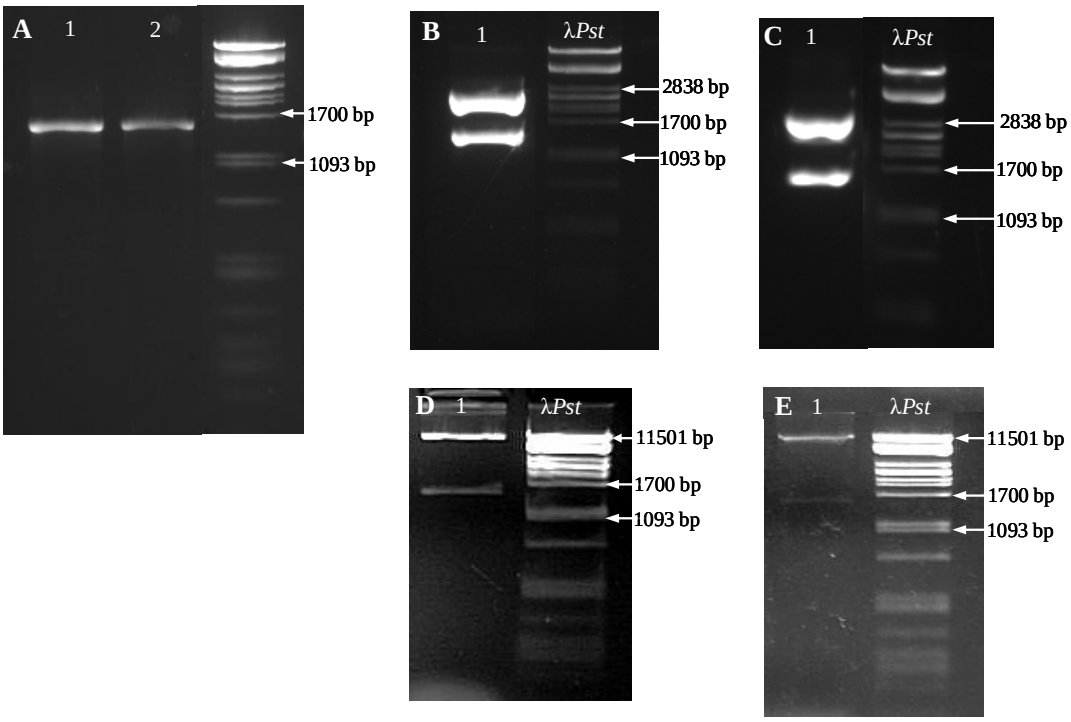


Figure 6.6 Cloning of the promoter region of At1g09070 from *A. l. petraea.* A) PCR of Helin (Lane 1) and Leitrim (Lane 2) *A. l. petraea* genomic DNA, expected size 1600 bp. B) and C) Restriction digests of putatively recombinant pBluescript (2958 bp) plasmid harbouring the At1g09070 promoter region amplified from Helin (B, Lane 1) and Leitrim (C, Lane 1) genomic DNA. D) and E) Restriction digests of putatively recombinant pBI101 (12 000 bp) harbouring the At1g19070 promoter region amplified from Helin (D, Lane 1) and Leitrim (E, Lane 1) genomic DNA. These constructs were used for transformation of *A. tumefaciens*, prior to transformation of both *A. thaliana* and *A. l. petraea*.

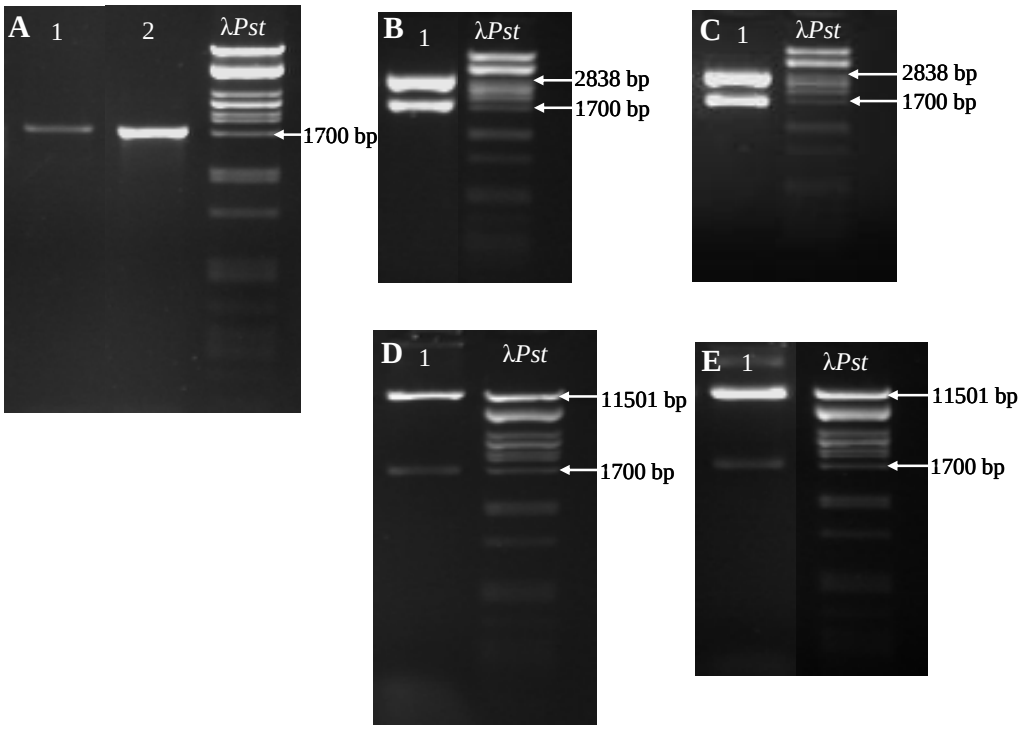
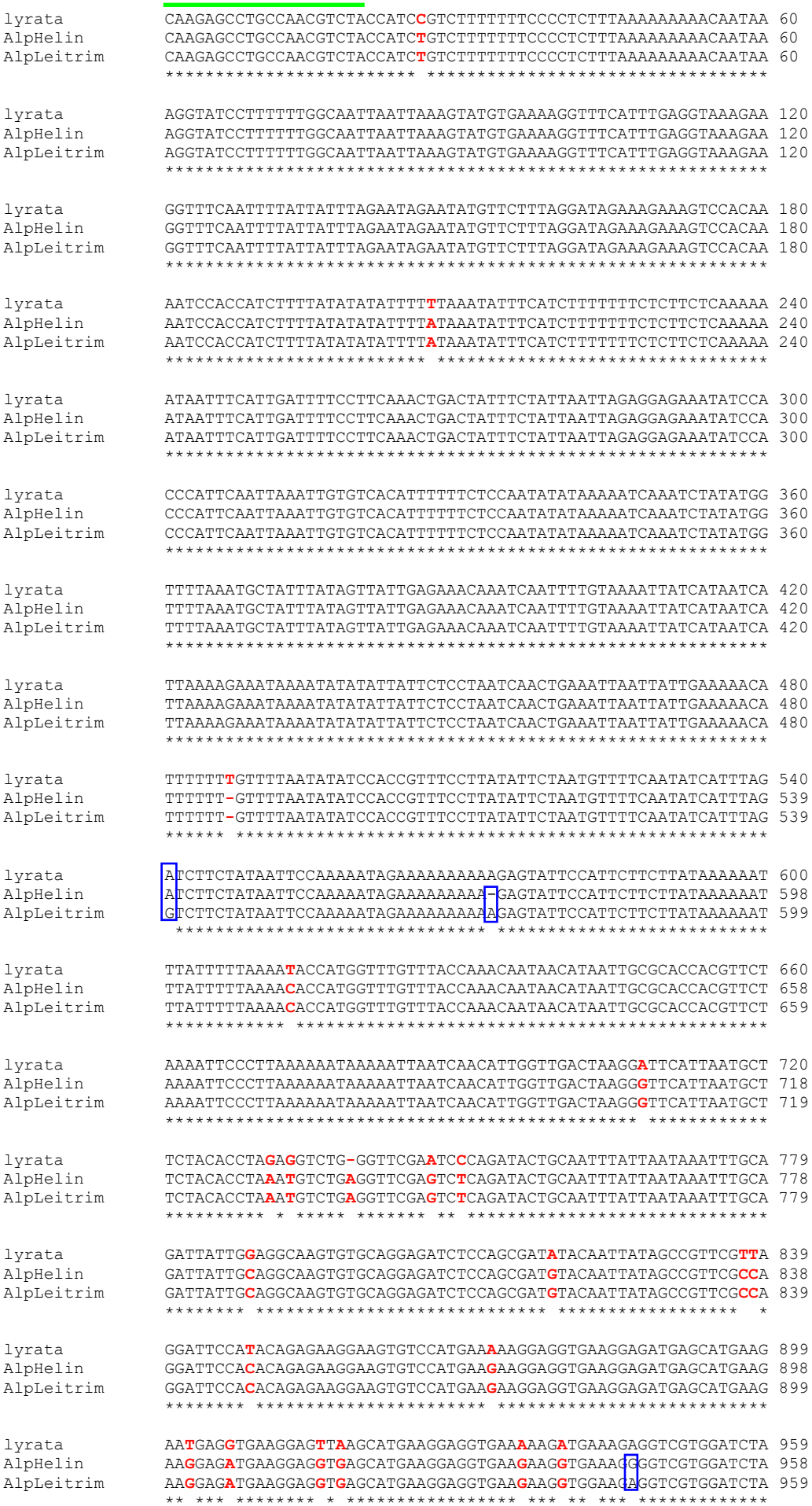


Figure 6.7 Cloning of the promoter region of At4g31800 from *A. l. petraea.* A) PCR of Helin (Lane 1) and Leitrim (Lane 2) *A. l. petraea* genomic DNA, expected size 2000 bp. B) and C) Restriction digests of putatively recombinant pBluescript (2958 bp) plasmid harbouring the At4g31800 promoter region amplified from Helin (B, Lane 1) and Leitrim (C, Lane 1) genomic DNA. D) and E) Restriction digests of putatively recombinant pBI101 (12 000 bp) harbouring the At4g31800 promoter region amplified from Helin (D, Lane 1) and Leitrim (E, Lane 1) genomic DNA. These constructs were used for transformation of *A. tumefaciens*, prior to transformation of both *A. thaliana* and *A. l. petraea*.


lyrata CAAGAGCCTGCCAACGTCTACCATCCGTCTTTTTTTCCCCTCTTTAAAAAAAAACAATAA 60
AlpHelin CAAGAGCCTGCCAACGTCTACCATCTGTCTTTTTTTCCCCTCTTTAAAAAAAAACAATAA 60
AlpLeitrim CAAGAGCCTGCCAACGTCTACCATCTGTCTTTTTTTCCCCTCTTTAAAAAAAAACAATAA 60
lyrata AGGTATCCTTTTTTGGCAATTAATTAAAGTATGTGAAAAGGTTTCATTTGAGGTAAAGAA 120
AlpHelin AGGTATCCTTTTTTGGCAATTAATTAAAGTATGTGAAAAGGTTTCATTTGAGGTAAAGAA 120
AlpLeitrim AGGTATCCTTTTTTGGCAATTAATTAAAGTATGTGAAAAGGTTTCATTTGAGGTAAAGAA 120
lyrata GGTTTCAATTTTATTATTTAGAATAGAATATGTTCTTTAGGATAGAAAGAAAGTCCACAA 180
AlpHelin GGTTTCAATTTTATTATTTAGAATAGAATATGTTCTTTAGGATAGAAAGAAAGTCCACAA 180
AlpLeitrim GGTTTCAATTTTATTATTTAGAATAGAATATGTTCTTTAGGATAGAAAGAAAGTCCACAA 180
lyrata AATCCACCATCTTTTATATATATTTTTTAAATATTTCATCTTTTTTTCTCTTCTCAAAAA 240
AlpHelin AATCCACCATCTTTTATATATATTTTTATAAATATTTCATCTTTTTTTCTCTTCTCAAAAA 240
AlpLeitrim AATCCACCATCTTTTATATATATTTTTATAAATATTTCATCTTTTTTTCTCTTCTCAAAAA 240
lyrata ATAATTTCATTGATTTTCCTTCAAACTGACTATTTCTATTAATTAGAGGAGAAATATCCA 300
AlpHelin ATAATTTCATTGATTTTCCTTCAAACTGACTATTTCTATTAATTAGAGGAGAAATATCCA 300
AlpLeitrim ATAATTTCATTGATTTTCCTTCAAACTGACTATTTCTATTAATTAGAGGAGAAATATCCA 300
lyrata CCCATTCAATTAAATTGTGTCACATTTTTTCTCCAATATATAAAAATCAAATCTATATGG 360
AlpHelin CCCATTCAATTAAATTGTGTCACATTTTTTCTCCAATATATAAAAATCAAATCTATATGG 360
AlpLeitrim CCCATTCAATTAAATTGTGTCACATTTTTTCTCCAATATATAAAAATCAAATCTATATGG 360
lyrata TTTTAAATGCTATTTATAGTTATTGAGAAACAAATCAATTTTGTAAAATTATCATAATCA 420
AlpHelin TTTTAAATGCTATTTATAGTTATTGAGAAACAAATCAATTTTGTAAAATTATCATAATCA 420
AlpLeitrim TTTTAAATGCTATTTATAGTTATTGAGAAACAAATCAATTTTGTAAAATTATCATAATCA 420
lyrata TTAAAAGAAATAAAATATATATTATTCTCCTAATCAACTGAAATTAATTATTGAAAAACA 480
AlpHelin TTAAAAGAAATAAAATATATATTATTCTCCTAATCAACTGAAATTAATTATTGAAAAACA 480
AlpLeitrim TTAAAAGAAATAAAATATATATTATTCTCCTAATCAACTGAAATTAATTATTGAAAAACA 480
lyrata TTTTTTTGTTTTAATATATCCACCGTTTCCTTATATTCTAATGTTTTCAATATCATTTAG 540
AlpHelin TTTTTT-GTTTTAATATATCCACCGTTTCCTTATATTCTAATGTTTTCAATATCATTTAG 539
AlpLeitrim TTTTTT-GTTTTAATATATCCACCGTTTCCTTATATTCTAATGTTTTCAATATCATTTAG 539
lyrata ATCTTCTATAATTCCAAAAATAGAAAAAAAAAAGAGTATTCCATTCTTCTTATAAAAAAT 600
AlpHelin ATCTTCTATAATTCCAAAAATAGAAAAAAAAA-GAGTATTCCATTCTTCTTATAAAAAAT 598
AlpLeitrim GTCTTCTATAATTCCAAAAATAGAAAAAAAAAAGAGTATTCCATTCTTCTTATAAAAAAT 599
lyrata TTATTTTTAAAATACCATGGTTTGTTTACCAAACAATAACATAATTGCGCACCACGTTCT 660
AlpHelin TTATTTTTAAAACACCATGGTTTGTTTACCAAACAATAACATAATTGCGCACCACGTTCT 658
AlpLeitrim TTATTTTTAAAACACCATGGTTTGTTTACCAAACAATAACATAATTGCGCACCACGTTCT 659
lyrata AAAATTCCCTTAAAAAATAAAAATTAATCAACATTGGTTGACTAAGGATTCATTAATGCT 720
AlpHelin AAAATTCCCTTAAAAAATAAAAATTAATCAACATTGGTTGACTAAGGGTTCATTAATGCT 718
AlpLeitrim AAAATTCCCTTAAAAAATAAAAATTAATCAACATTGGTTGACTAAGGGTTCATTAATGCT 719
lyrata TCTACACCTAGAGGTCTG-GGTTCGAATCCCAGATACTGCAATTTATTAATAAATTTGCA 779
AlpHelin TCTACACCTAAATGTCTGAGGTTCGAGTCTCAGATACTGCAATTTATTAATAAATTTGCA 778
AlpLeitrim TCTACACCTAAATGTCTGAGGTTCGAGTCTCAGATACTGCAATTTATTAATAAATTTGCA 779
lyrata GATTATTGGAGGCAAGTGTGCAGGAGATCTCCAGCGATATACAATTATAGCCGTTCGTTA 839
AlpHelin GATTATTGCAGGCAAGTGTGCAGGAGATCTCCAGCGATGTACAATTATAGCCGTTCGCCA 838
AlpLeitrim GATTATTGCAGGCAAGTGTGCAGGAGATCTCCAGCGATGTACAATTATAGCCGTTCGCCA 839
lyrata GGATTCCATACAGAGAAGGAAGTGTCCATGAAAAAGGAGGTGAAGGAGATGAGCATGAAG 899
AlpHelin GGATTCCACACAGAGAAGGAAGTGTCCATGAAGAAGGAGGTGAAGGAGATGAGCATGAAG 898
AlpLeitrim GGATTCCACACAGAGAAGGAAGTGTCCATGAAGAAGGAGGTGAAGGAGATGAGCATGAAG 899
lyrata AATGAGGTGAAGGAGTTAAGCATGAAGGAGGTGAAAAAGATGAAAGAGGTCGTGGATCTA 959
AlpHelin AAGGAGATGAAGGAGGTGAGCATGAAGGAGGTGAAGAAGGTGAAACGGTCGTGGATCTA 958
AlpLeitrim AAGGAGATGAAGGAGGTGAGCATGAAGGAGGTGAAGAAGGTGGAAGAGGTCGTGGATCTA 959

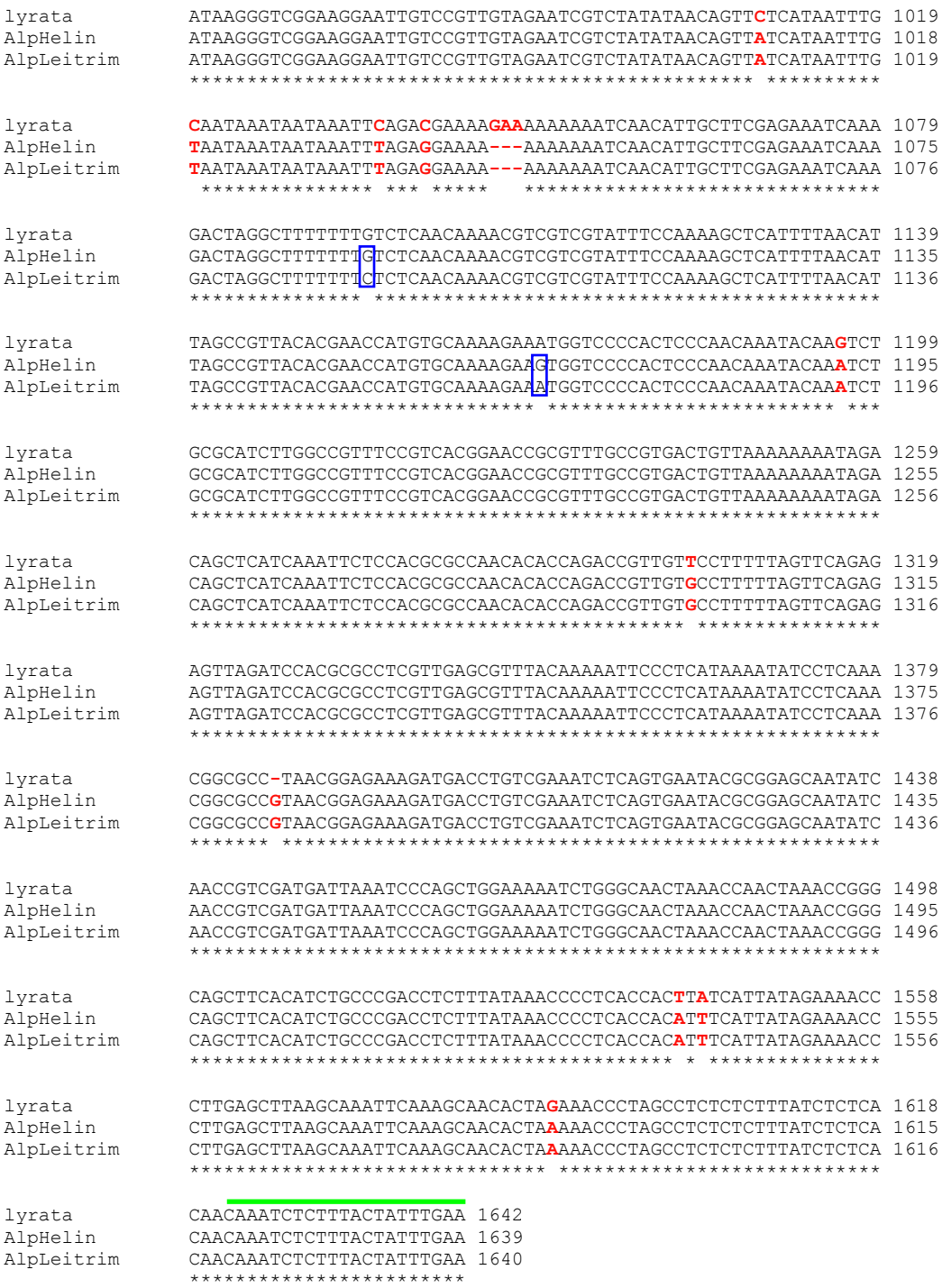
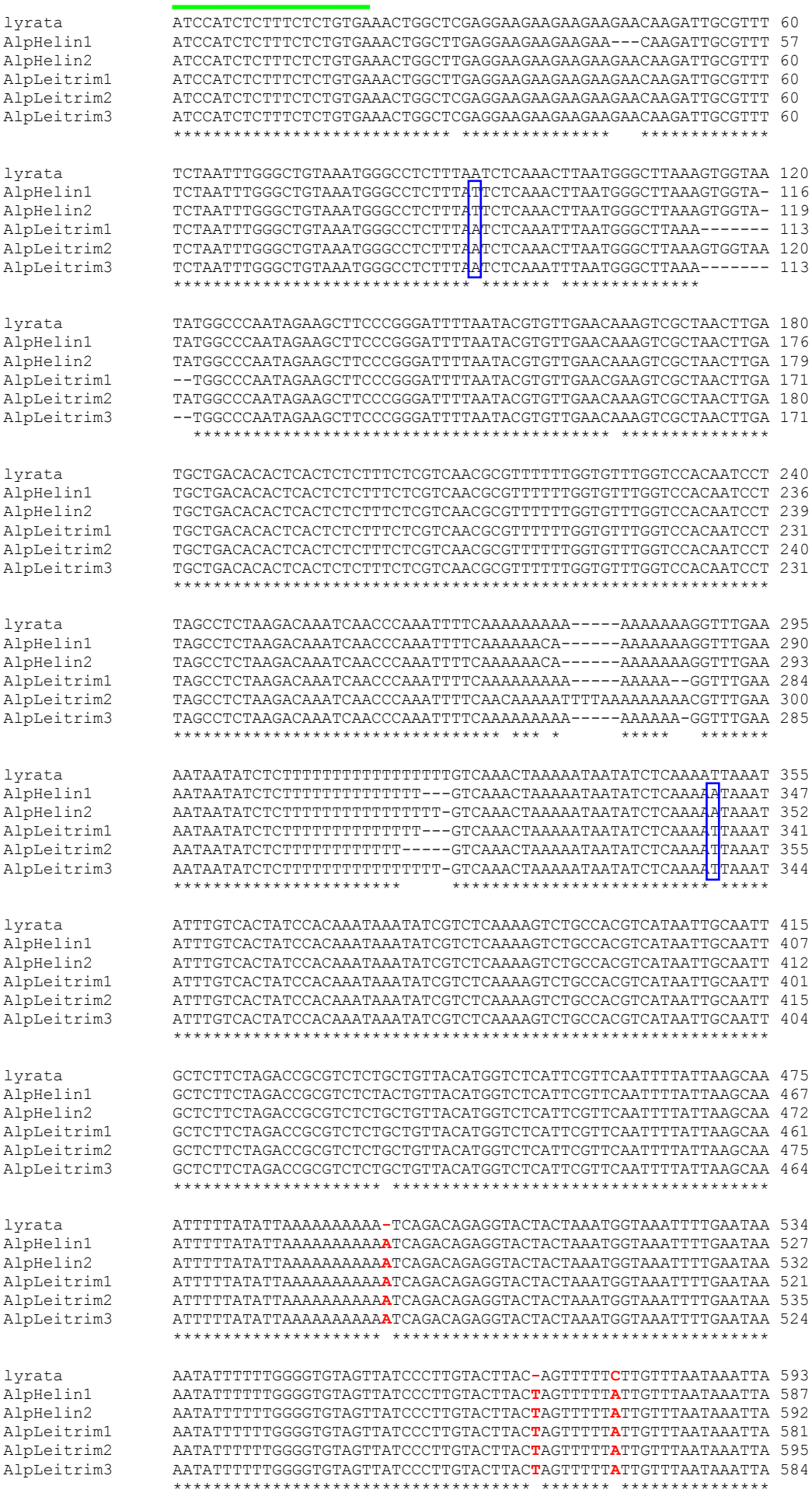


Figure 6.8 The *A. l. petraea* promoter region sequence for At5g11740. The top sequence is the *A. l. lyrata* sequence for the region upstream of the coding sequence for At5g11740. AlpHelin and AlpLeitrim (*A. l. petraea)* refer to the consensus sequences created from sequencing recombinant plasmid DNA inserts in both directions (Helin: one clone; Leitrim: two clones). Asterisk (*) indicate shared identity between all three sequences. Bases coloured in red illustrate differences between the *A. l. lyrata* and *A. l. petraea* sequences that are confirmed by both *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Regions in blue boxes highlight differences between the Helin and Leitrim promoter sequences.


lyrata ATCCATCTCTTTCTCTGTGAAACTGGCTCGAGGAAGAAGAAGAAGAACAAGATTGCGTTT 60
AlpHelin1 ATCCATCTCTTTCTCTGTGAAACTGGCTTGAGGAAGAAGAAGAA---CAAGATTGCGTTT 57
AlpHelin2 ATCCATCTCTTTCTCTGTGAAACTGGCTTGAGGAAGAAGAAGAAGAACAAGATTGCGTTT 60
AlpLeitrim1 ATCCATCTCTTTCTCTGTGAAACTGGCTTGAGGAAGAAGAAGAAGAACAAGATTGCGTTT 60
AlpLeitrim2 ATCCATCTCTTTCTCTGTGAAACTGGCTCGAGGAAGAAGAAGAAGAACAAGATTGCGTTT 60
AlpLeitrim3 ATCCATCTCTTTCTCTGTGAAACTGGCTCGAGGAAGAAGAAGAAGAACAAGATTGCGTTT 60
lyrata TCTAATTTGGGCTGTAAATGGGCCTCTTTAATCTCAAACTTAATGGGCTTAAAGTGGTAA 120
AlpHelin1 TCTAATTTGGGCTGTAAATGGGCCTCTTTATTCTCAAACTTAATGGGCTTAAAGTGGTA- 116
AlpHelin2 TCTAATTTGGGCTGTAAATGGGCCTCTTTATTCTCAAACTTAATGGGCTTAAAGTGGTA- 119
AlpLeitrim1 TCTAATTTGGGCTGTAAATGGGCCTCTTTAATCTCAAATTTAATGGGCTTAAA------- 113
AlpLeitrim2 TCTAATTTGGGCTGTAAATGGGCCTCTTTAATCTCAAACTTAATGGGCTTAAAGTGGTAA 120
AlpLeitrim3 TCTAATTTGGGCTGTAAATGGGCCTCTTTAATCTCAAATTTAATGGGCTTAAA------- 113
lyrata TATGGCCCAATAGAAGCTTCCCGGGATTTTAATACGTGTTGAACAAAGTCGCTAACTTGA 180
AlpHelin1 TATGGCCCAATAGAAGCTTCCCGGGATTTTAATACGTGTTGAACAAAGTCGCTAACTTGA 176
AlpHelin2 TATGGCCCAATAGAAGCTTCCCGGGATTTTAATACGTGTTGAACAAAGTCGCTAACTTGA 179
AlpLeitrim1 --TGGCCCAATAGAAGCTTCCCGGGATTTTAATACGTGTTGAACGAAGTCGCTAACTTGA 171
AlpLeitrim2 TATGGCCCAATAGAAGCTTCCCGGGATTTTAATACGTGTTGAACAAAGTCGCTAACTTGA 180
AlpLeitrim3 --TGGCCCAATAGAAGCTTCCCGGGATTTTAATACGTGTTGAACAAAGTCGCTAACTTGA 171
lyrata TGCTGACACACTCACTCTCTTTCTCGTCAACGCGTTTTTTGGTGTTTGGTCCACAATCCT 240
AlpHelin1 TGCTGACACACTCACTCTCTTTCTCGTCAACGCGTTTTTTGGTGTTTGGTCCACAATCCT 236
AlpHelin2 TGCTGACACACTCACTCTCTTTCTCGTCAACGCGTTTTTTGGTGTTTGGTCCACAATCCT 239
AlpLeitrim1 TGCTGACACACTCACTCTCTTTCTCGTCAACGCGTTTTTTGGTGTTTGGTCCACAATCCT 231
AlpLeitrim2 TGCTGACACACTCACTCTCTTTCTCGTCAACGCGTTTTTTGGTGTTTGGTCCACAATCCT 240
AlpLeitrim3 TGCTGACACACTCACTCTCTTTCTCGTCAACGCGTTTTTTGGTGTTTGGTCCACAATCCT 231
lyrata TAGCCTCTAAGACAAATCAACCCAAATTTTCAAAAAAAAA-----AAAAAAAGGTTTGAA 295
AlpHelin1 TAGCCTCTAAGACAAATCAACCCAAATTTTCAAAAAACA------AAAAAAAGGTTTGAA 290
AlpHelin2 TAGCCTCTAAGACAAATCAACCCAAATTTTCAAAAAACA------AAAAAAAGGTTTGAA 293
AlpLeitrim1 TAGCCTCTAAGACAAATCAACCCAAATTTTCAAAAAAAAA-----AAAAA--GGTTTGAA 284
AlpLeitrim2 TAGCCTCTAAGACAAATCAACCCAAATTTTCAACAAAAATTTTAAAAAAAAACGTTTGAA 300
AlpLeitrim3 TAGCCTCTAAGACAAATCAACCCAAATTTTCAAAAAAAAA-----AAAAAA-GGTTTGAA 285
lyrata AATAATATCTCTTTTTTTTTTTTTTTTGTCAAACTAAAAATAATATCTCAAAATTAAAT 355
AlpHelin1 AATAATATCTCTTTTTTTTTTTTTT---GTCAAACTAAAAATAATATCTCAAAATAAAT 347
AlpHelin2 AATAATATCTCTTTTTTTTTTTTTTT-GTCAAACTAAAAATAATATCTCAAAATAAAT 352
AlpLeitrim1 AATAATATCTCTTTTTTTTTTTTTT---GTCAAACTAAAAATAATATCTCAAAATTAAAT 341
AlpLeitrim2 AATAATATCTCTTTTTTTTTTTT-----GTCAAACTAAAAATAATATCTCAAAATTAAAT 355
AlpLeitrim3 AATAATATCTCTTTTTTTTTTTTTTT-GTCAAACTAAAAATAATATCTCAAAATTAAAT 344
lyrata ATTTGTCACTATCCACAAATAAATATCGTCTCAAAAGTCTGCCACGTCATAATTGCAATT 415
AlpHelin1 ATTTGTCACTATCCACAAATAAATATCGTCTCAAAAGTCTGCCACGTCATAATTGCAATT 407
AlpHelin2 ATTTGTCACTATCCACAAATAAATATCGTCTCAAAAGTCTGCCACGTCATAATTGCAATT 412
AlpLeitrim1 ATTTGTCACTATCCACAAATAAATATCGTCTCAAAAGTCTGCCACGTCATAATTGCAATT 401
AlpLeitrim2 ATTTGTCACTATCCACAAATAAATATCGTCTCAAAAGTCTGCCACGTCATAATTGCAATT 415
AlpLeitrim3 ATTTGTCACTATCCACAAATAAATATCGTCTCAAAAGTCTGCCACGTCATAATTGCAATT 404
lyrata GCTCTTCTAGACCGCGTCTCTGCTGTTACATGGTCTCATTCGTTCAATTTTATTAAGCAA 475
AlpHelin1 GCTCTTCTAGACCGCGTCTCTACTGTTACATGGTCTCATTCGTTCAATTTTATTAAGCAA 467
AlpHelin2 GCTCTTCTAGACCGCGTCTCTGCTGTTACATGGTCTCATTCGTTCAATTTTATTAAGCAA 472
AlpLeitrim1 GCTCTTCTAGACCGCGTCTCTGCTGTTACATGGTCTCATTCGTTCAATTTTATTAAGCAA 461
AlpLeitrim2 GCTCTTCTAGACCGCGTCTCTGCTGTTACATGGTCTCATTCGTTCAATTTTATTAAGCAA 475
AlpLeitrim3 GCTCTTCTAGACCGCGTCTCTGCTGTTACATGGTCTCATTCGTTCAATTTTATTAAGCAA 464
lyrata ATTTTTATATTAAAAAAAAAA-TCAGACAGAGGTACTACTAAATGGTAAATTTTGAATAA 534
AlpHelin1 ATTTTTATATTAAAAAAAAAAATCAGACAGAGGTACTACTAAATGGTAAATTTTGAATAA 527
AlpHelin2 ATTTTTATATTAAAAAAAAAAATCAGACAGAGGTACTACTAAATGGTAAATTTTGAATAA 532
AlpLeitrim1 ATTTTTATATTAAAAAAAAAAATCAGACAGAGGTACTACTAAATGGTAAATTTTGAATAA 521
AlpLeitrim2 ATTTTTATATTAAAAAAAAAAATCAGACAGAGGTACTACTAAATGGTAAATTTTGAATAA 535
AlpLeitrim3 ATTTTTATATTAAAAAAAAAAATCAGACAGAGGTACTACTAAATGGTAAATTTTGAATAA 524
lyrata AATATTTTTTGGGGGTGTAGTTATCCCTTGTACTTAC-AGTTTTTCTTGTTTAATAAATTA 593
AlpHelin1 AATATTTTTTGGGGGTGTAGTTATCCCTTGTACTTACTAGTTTTTATTGTTTAATAAATTA 587
AlpHelin2 AATATTTTTTGGGGGTGTAGTTATCCCTTGTACTTACTAGTTTTTATTGTTTAATAAATTA 592
AlpLeitrim1 AATATTTTTTGGGGGTGTAGTTATCCCTTGTACTTACTAGTTTTTATTGTTTAATAAATTA 581
AlpLeitrim2 AATATTTTTTGGGGGTGTAGTTATCCCTTGTACTTACTAGTTTTTATTGTTTAATAAATTA 595
AlpLeitrim3 AATATTTTTTGGGGGTGTAGTTATCCCTTGTACTTACTAGTTTTTATTGTTTAATAAATTA 584

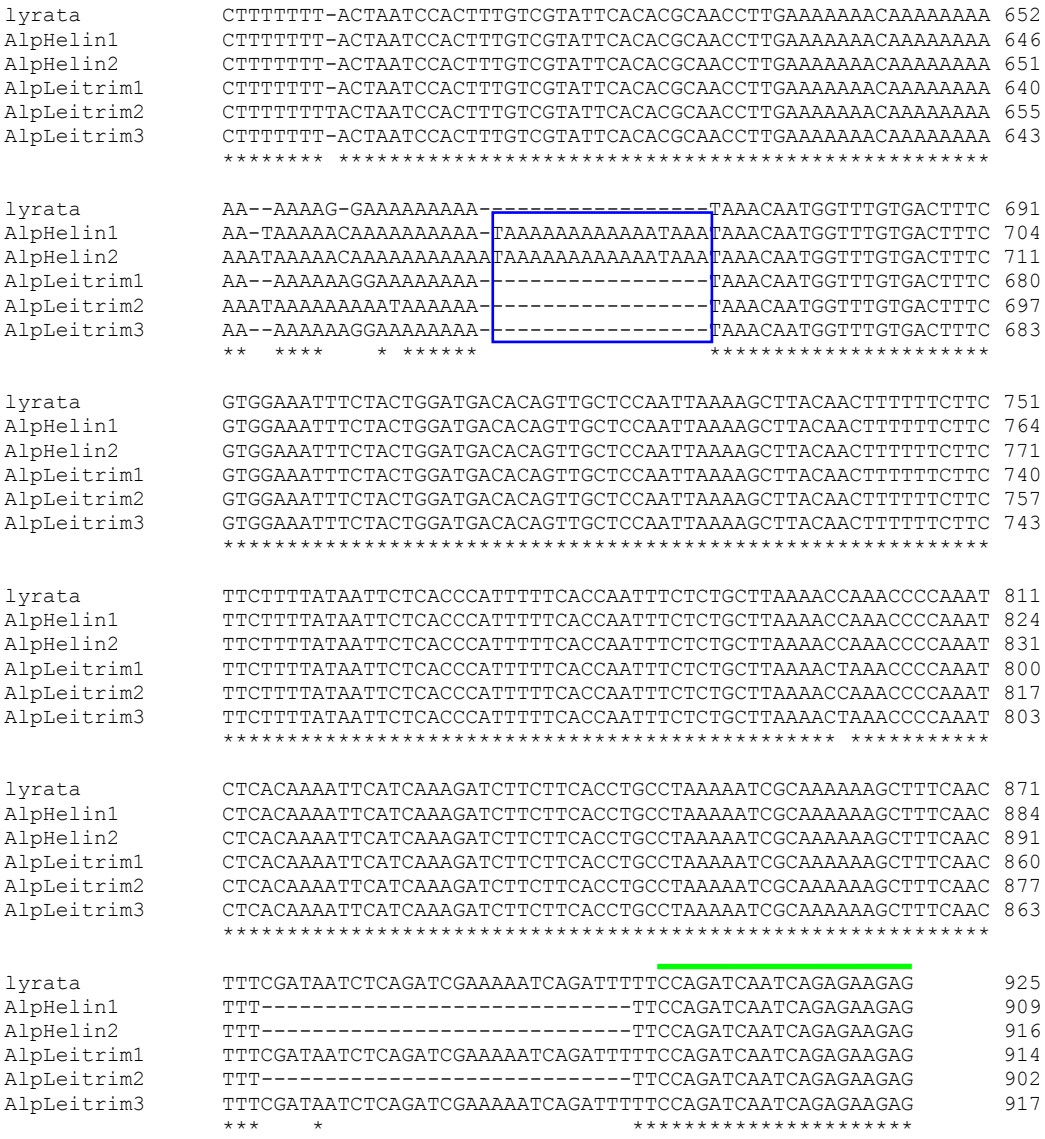
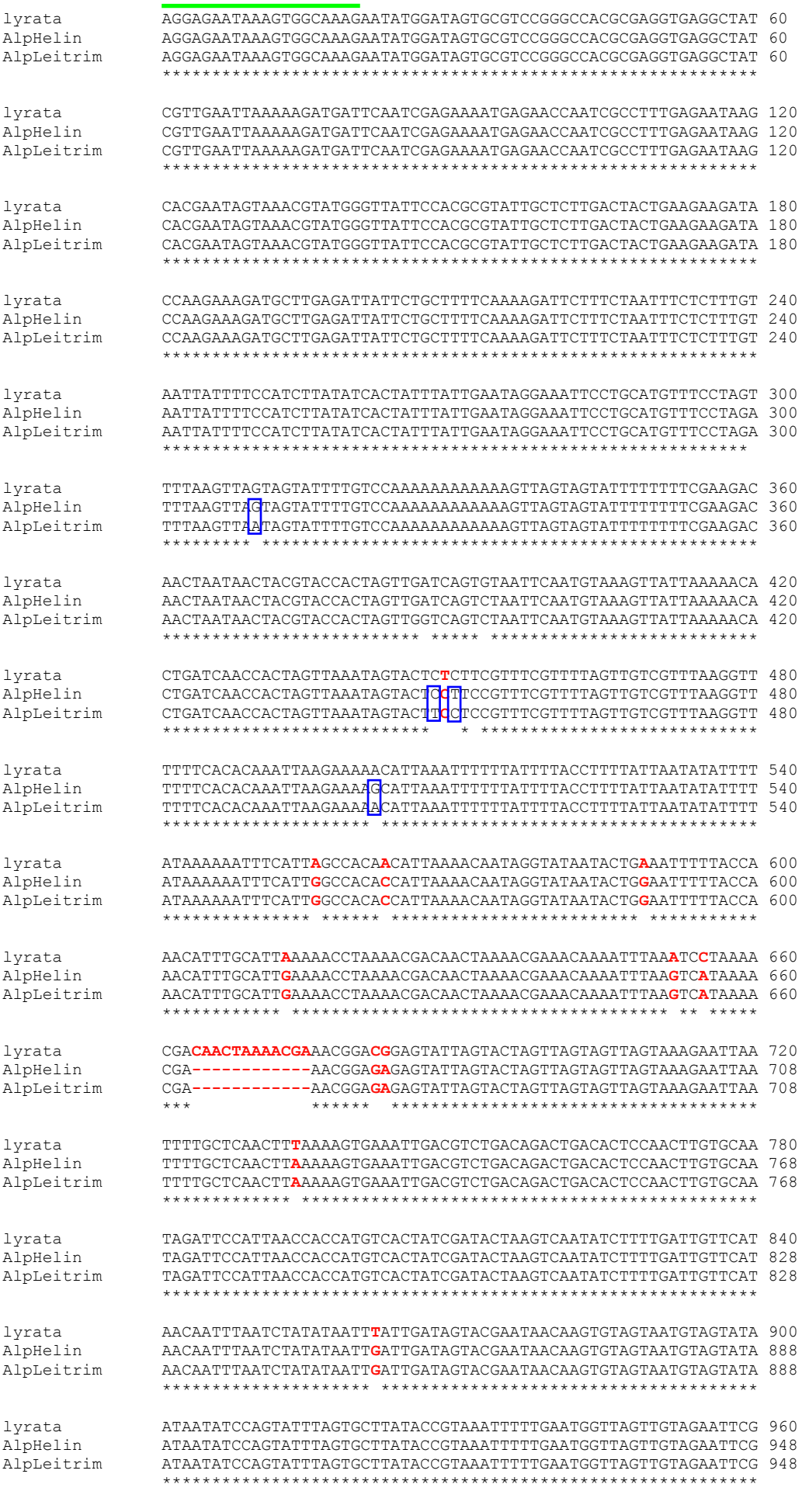


Figure 6.9 The *A. l. petraea* promoter region sequence for At5g42050. The top sequence is the *A. l. lyrata* sequence for the region upstream of the coding sequence for At5g42050. Consensus sequences created from sequencing recombinant plasmid DNA inserts in both directions are presented for two Helin clones (AlpHelin1 and AlpHelin2) and three Leitrim clones (AlpLeitrim1, AlpLeitrim2, and AlpLeitrim3), all from *A. l. petraea.* Asterisk (*) indicate shared identity between all sequences. Bases coloured in red illustrate differences between the *A. l. lyrata* and *A. l. petraea* sequences that are confirmed by all *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Regions in blue boxes highlight consistent differences between the Helin and Leitrim promoter sequences.

```
lyrata          GATGGGAAAAGAGAACACATAAAGGTCCGACAATGGAGATT------------------- 41
AlpHelin        GATGGGAAAAGAGAACACATAAAGGTCCGACAATGGAGATTCGAAAAGAAAATTGTGAAA 60
AlpLeitrim      GATGGGAAAAGAGAACACATAAAGGTCCGACAATGGGGATTCGAAAAGAAAATTGTGAAA 60
                ************************************

lyrata          ------------------------------------------------------------
AlpHelin        TTTTAAAAAATATCGTGAAACATAGAGAAAATACATGAATTTATCATGTAAATGGCATGT 120
AlpLeitrim      TTTTAAAAAATATCGTGAAACATAGAGAAAATACATGAATTTATCATGTAAATGGCATGT 120

lyrata          ------------------------------------------------------------
AlpHelin        AACATGAATAAAAATATGATTGCTTGAAGTAAATAAAATTAGCATGTGATATGCCATTTT 180
AlpLeitrim      AACATGAATAAAAATATGATTGCTTGAAGTAAATAAAATTAGCATGTGATATGCCATTTT 180

lyrata          ------------------------------------------------------------
AlpHelin        TAAACATATCATGACATCTCGAAAAACATAACGTTCGTAGTATAATTGCTTGAAGTAATT 240
AlpLeitrim      TAAACATATCATGACATCTCGAAAAACATAACGTTCGTAGTATAATTGCTTGAAGTAATT 240

lyrata          ------------------------------------------------------------
AlpHelin        AGAAATAGTATGTGATATTATCATTTGTTGGAAAATATATCATGTCGTTTTTAAAAATAG 300
AlpLeitrim      AGAAATAGTATGTGATATTATCATTTGTTGGAAAATATATCATGTCGTTTTTAAAAATAG 300

lyrata          ------------------------------------------------------------
AlpHelin        CATGACATTTCGAAAAATATGTTAATAAAAAATAGCAAGAGGAGGTAATTAAGCAATAAC 360
AlpLeitrim      CATGACATTTCGAAAAATATGTTAATAAAAAATAGCAAGAGGAGGTAATTAAGCAATAAC 360

lyrata          ------------------------------------------------------------
AlpHelin        AGAGGTGTATAAAGGCGGAGAAAATCTGGGATTTCAAAACAGCGGATAATATTTAGAAAT 420
AlpLeitrim      AGAGGTGTATAAAGGCGGAGAAAATCTGGGATTTCAAAACAGCGGATAATATTTAGAAAT 420

lyrata          ------------------------------------------------------------
AlpHelin        AGTGATTCAACATTGGCATTTTGAAACATATGAGCATGATTTGATTCTGCGGAAATGACG 480
AlpLeitrim      AGTGATTCAACATTGGCATTTTGAAACATATGAGCATGATTTGATTCTGCGGAAATGACG 480

lyrata          ------------------------------------------------------------
AlpHelin        TTGTTGTTGGTGGTGAAATTTCCTATTGAACAAAAAATAATGAGATATATTTTAACTTAT 540
AlpLeitrim      TTGTTGTTAGTGGTGAAATTTCCTATTGAACAAAAAATAATGAGATATATTTTAACTTAT 540

lyrata          ------------------------------------------------------------
AlpHelin        CTGAAAAAGATTGATTATGGATTCGAGCATTATTGAAGCAAAGAAGATGAAGGTGAAGAT 600
AlpLeitrim      CTGAAAAAGATTGATTATGGATTCGAGCATTATTGAAGCAAAGAAGATGAAGGTGAAGAT 600

lyrata          ------------------------------------------------------------
AlpHelin        AGACTGGCGACAGAAGAGGGATTTAGAGAGAAGCCAAAATGAAGAGTGACGAATGAAGAA 660
AlpLeitrim      AGACTGGCGACAGAAGAGGGATTTAGAGAGAAGCCAAAATGAAGAGTGACGAATGAAGAA 660

lyrata          ----------------------------------------------------ACTATTTATC 51
AlpHelin        ATAATTTTTAAAGGTTGTATATGTAATTTACTTACCCTTCCCGACAAGAAACTATTTATC 720
AlpLeitrim      ATAATTTTTAAAGGTTGTATATGTAATTTACTTACCCTTCCCGACAAGAAACTATTTATC 720
                                                    **********

lyrata          ATGGTTAGTTTCTTTATTATTTTTGGTTTCATGGCTAGTTAGTGTATATTCCCTAATTTA 111
AlpHelin        ATGATTAATTTCTTTATTTTTTTTGGTTTCATGGCTAGTTAGTGTATATTTCCTAAT--- 777
AlpLeitrim      ATGATTAATTTCTTTATTTTTTTGGTTTCATGGCTAGTTAGTGTATATTTCCTAAT--- 777
                *** *** ********** ****************************** ******

lyrata          TATTTAATAAGAAAGAATTAAATAGTTCAGATTTCAAC-AAATACAGTATTTTCAAACCA 170
AlpHelin        -ATTTAATAAGAAAGAATTAAATAGTTCAGATTTCAACCAAATACAGTATTTTCAAACCA 836
AlpLeitrim      -ATTTAATAGGAAAGAATTAAATAGTTCAGATTTCAACCAAATACAGTATTTTCAAACCA 836
                 ******** **************************** ********************

lyrata          GATACGAATTATAAATGTATTAATCCAAATTATCAAAGTATTATAATATGAAGGGAAATT 230
AlpHelin        GATACGAATTATAAATGTATTAATCCAAATTATCAAAGTATTATAATATGAAGGGAAATT 896
AlpLeitrim      GATACGAATTATAAATGTATTAATCCAAATTATCAAAGTATTATAATATGAAGGGAAATT 896
                ************************************************************

lyrata          AAATACATAAAGATTCGATTCCAAACTTGTAATTAGTGGCATCTATAAAAAAAAATGATA 290
AlpHelin        AAATACATAAAGATTCGATTCCAAACTTGTAATTAGTGGCATCTATAAAAAAAAATGATA 955
AlpLeitrim      AAATACATAAAGATTCGATTCCAAACTTGTAATTAGTGGCATCTATAAAAAAAAATGATA 956
                ************************************************** *****
```

```
lyrata          TCACAAAAAGGTCTATAAAAGGGTCAACTCATCAGCCTCCAATTGATCCAAAACACTTTA 350
AlpHelin        TCACAAAAAGGTCTATAAAAGGGTCAACTCATCAGCCTCCAATTGATCCAAAACACTTTA 1015
AlpLeitrim      TCACAAAAAGGTCTATAAAAGGGTCAACTCATCAGCCTCCAATTGATCCAAAACACTTTA 1016
                ************************************************************

                ____________________
lyrata          CCTTTAACAAAAAAAAAATGGCA 373
AlpHelin        CCTTTAACAAAAAAAAAATGGCA 1038
AlpLeitrim      CCTTTAACAAAAAAAAAATGGCA 1039
                ***********************
```

Figure 6.10 The *A. l. petraea* promoter region sequence for At1g18210. The top sequence is the *A. l. lyrata* sequence for the region upstream of the coding sequence for At1g18210. This region fell between two scaffolds of the available *A. l. lyrata* sequence and thus the majority of the sequence was unavailable. AlpHelin and AlpLeitrim (*A. l. petraea)* refer to the consensus sequences created from sequencing recombinant plasmid DNA inserts in both directions (Helin: three clones; Leitrim: three clones). Asterisk (*) indicate shared identity between all three sequences. Bases coloured in red illustrate differences between the *A. l. lyrata* and *A. l. petraea* sequences that are confirmed by both *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Regions in blue boxes highlight differences between the Helin and Leitrim promoter sequences.


lyrata AGGAGAATAAAGTGGCAAAGAATATGGATAGTGCGTCCGGGCCACGCGAGGTGAGGCTAT 60
AlpHelin AGGAGAATAAAGTGGCAAAGAATATGGATAGTGCGTCCGGGCCACGCGAGGTGAGGCTAT 60
AlpLeitrim AGGAGAATAAAGTGGCAAAGAATATGGATAGTGCGTCCGGGCCACGCGAGGTGAGGCTAT 60
lyrata CGTTGAATTAAAAAGATGATTCAATCGAGAAAATGAGAACCAATCGCCTTTGAGAATAAG 120
AlpHelin CGTTGAATTAAAAAGATGATTCAATCGAGAAAATGAGAACCAATCGCCTTTGAGAATAAG 120
AlpLeitrim CGTTGAATTAAAAAGATGATTCAATCGAGAAAATGAGAACCAATCGCCTTTGAGAATAAG 120
lyrata CACGAATAGTAAACGTATGGGTTATTCCACGCGTATTGCTCTTGACTACTGAAGAAGATA 180
AlpHelin CACGAATAGTAAACGTATGGGTTATTCCACGCGTATTGCTCTTGACTACTGAAGAAGATA 180
AlpLeitrim CACGAATAGTAAACGTATGGGTTATTCCACGCGTATTGCTCTTGACTACTGAAGAAGATA 180
lyrata CCAAGAAAGATGCTTGAGATTATTCTGCTTTTCAAAAGATTCTTTCTAATTTCTCTTTGT 240
AlpHelin CCAAGAAAGATGCTTGAGATTATTCTGCTTTTCAAAAGATTCTTTCTAATTTCTCTTTGT 240
AlpLeitrim CCAAGAAAGATGCTTGAGATTATTCTGCTTTTCAAAAGATTCTTTCTAATTTCTCTTTGT 240
lyrata AATTATTTTCCATCTTATATCACTATTTATTGAATAGGAAATTCCTGCATGTTTCCTAGT 300
AlpHelin AATTATTTTCCATCTTATATCACTATTTATTGAATAGGAAATTCCTGCATGTTTCCTAGA 300
AlpLeitrim AATTATTTTCCATCTTATATCACTATTTATTGAATAGGAAATTCCTGCATGTTTCCTAGA 300
lyrata TTTAAGTTAGTAGTATTTTGTCCAAAAAAAAAAAAAGTTAGTAGTATTTTTTTTCGAAGAC 360
AlpHelin TTTAAGTTAGTAGTATTTTGTCCAAAAAAAAAAAAAGTTAGTAGTATTTTTTTTCGAAGAC 360
AlpLeitrim TTTAAGTTAATAGTATTTTGTCCAAAAAAAAAAAAAGTTAGTAGTATTTTTTTTCGAAGAC 360
lyrata AACTAATAACTACGTACCACTAGTTGATCAGTGTAATTCAATGTAAAGTTATTAAAAACA 420
AlpHelin AACTAATAACTACGTACCACTAGTTGATCAGTCTAATTCAATGTAAAGTTATTAAAAACA 420
AlpLeitrim AACTAATAACTACGTACCACTAGTTGGTCAGTCTAATTCAATGTAAAGTTATTAAAAACA 420
lyrata CTGATCAACCACTAGTTAAATAGTACTCTTCTTCGTTTCGTTTTAGTTGTCGTTTAAGGTT 480
AlpHelin CTGATCAACCACTAGTTAAATAGTACTCCTTCCGTTTCGTTTTAGTTGTCGTTTAAGGTT 480
AlpLeitrim CTGATCAACCACTAGTTAAATAGTACTTCCTCCGTTTCGTTTTAGTTGTCGTTTAAGGTT 480
lyrata TTTTCACACAAATTAAGAAAAACATTAAATTTTTTATTTTACCTTTTATTAATATATTTT 540
AlpHelin TTTTCACACAAATTAAGAAAAGCATTAAATTTTTTATTTTACCTTTTATTAATATATTTT 540
AlpLeitrim TTTTCACACAAATTAAGAAAAACATTAAATTTTTTATTTTACCTTTTATTAATATATTTT 540
lyrata ATAAAAAATTTCATTAGCCACAACATTAAAACAATAGGTATAATACTGAAATTTTTACCA 600
AlpHelin ATAAAAAATTTCATTGGCCACACCATTAAAACAATAGGTATAATACTGGAATTTTTACCA 600
AlpLeitrim ATAAAAAATTTCATTGGCCACACCATTAAAACAATAGGTATAATACTGGAATTTTTACCA 600
lyrata AACATTTGCATTAAAAACCTAAAACGACAACTAAAACGAAACAAAATTTAAATCCTAAAA 660
AlpHelin AACATTTGCATTGAAAACCTAAAACGACAACTAAAACGAAACAAAATTTAAGTCATAAAA 660
AlpLeitrim AACATTTGCATTGAAAACCTAAAACGACAACTAAAACGAAACAAAATTTAAGTCATAAAA 660
lyrata CGACAACTAAAACGAAACGGACGGAGTATTAGTACTAGTTAGTAGTTAGTAAAGAATTAA 720
AlpHelin CGA----------AACGGAGAGAGTATTAGTACTAGTTAGTAGTTAGTAAAGAATTAA 708
AlpLeitrim CGA----------AACGGAGAGAGTATTAGTACTAGTTAGTAGTTAGTAAAGAATTAA 708
lyrata TTTTGCTCAACTTTAAAAGTGAAATTGACGTCTGACAGACTGACACTCCAACTTGTGCAA 780
AlpHelin TTTTGCTCAACTTAAAAAGTGAAATTGACGTCTGACAGACTGACACTCCAACTTGTGCAA 768
AlpLeitrim TTTTGCTCAACTTAAAAAGTGAAATTGACGTCTGACAGACTGACACTCCAACTTGTGCAA 768
lyrata TAGATTCCATTAACCACCATGTCACTATCGATACTAAGTCAATATCTTTTGATTGTTCAT 840
AlpHelin TAGATTCCATTAACCACCATGTCACTATCGATACTAAGTCAATATCTTTTGATTGTTCAT 828
AlpLeitrim TAGATTCCATTAACCACCATGTCACTATCGATACTAAGTCAATATCTTTTGATTGTTCAT 828
lyrata AACAATTTAATCTATATAATTTATTGATAGTACGAATAACAAGTGTAGTAATGTAGTATA 900
AlpHelin AACAATTTAATCTATATAATTGATTGATAGTACGAATAACAAGTGTAGTAATGTAGTATA 888
AlpLeitrim AACAATTTAATCTATATAATTGATTGATAGTACGAATAACAAGTGTAGTAATGTAGTATA 888
lyrata ATAATATCCAGTATTTAGTGCTTATACCGTAAATTTTTGAATGGTTAGTTGTAGAATTCG 960
AlpHelin ATAATATCCAGTATTTAGTGCTTATACCGTAAATTTTTGAATGGTTAGTTGTAGAATTCG 948
AlpLeitrim ATAATATCCAGTATTTAGTGCTTATACCGTAAATTTTTGAATGGTTAGTTGTAGAATTCG 948

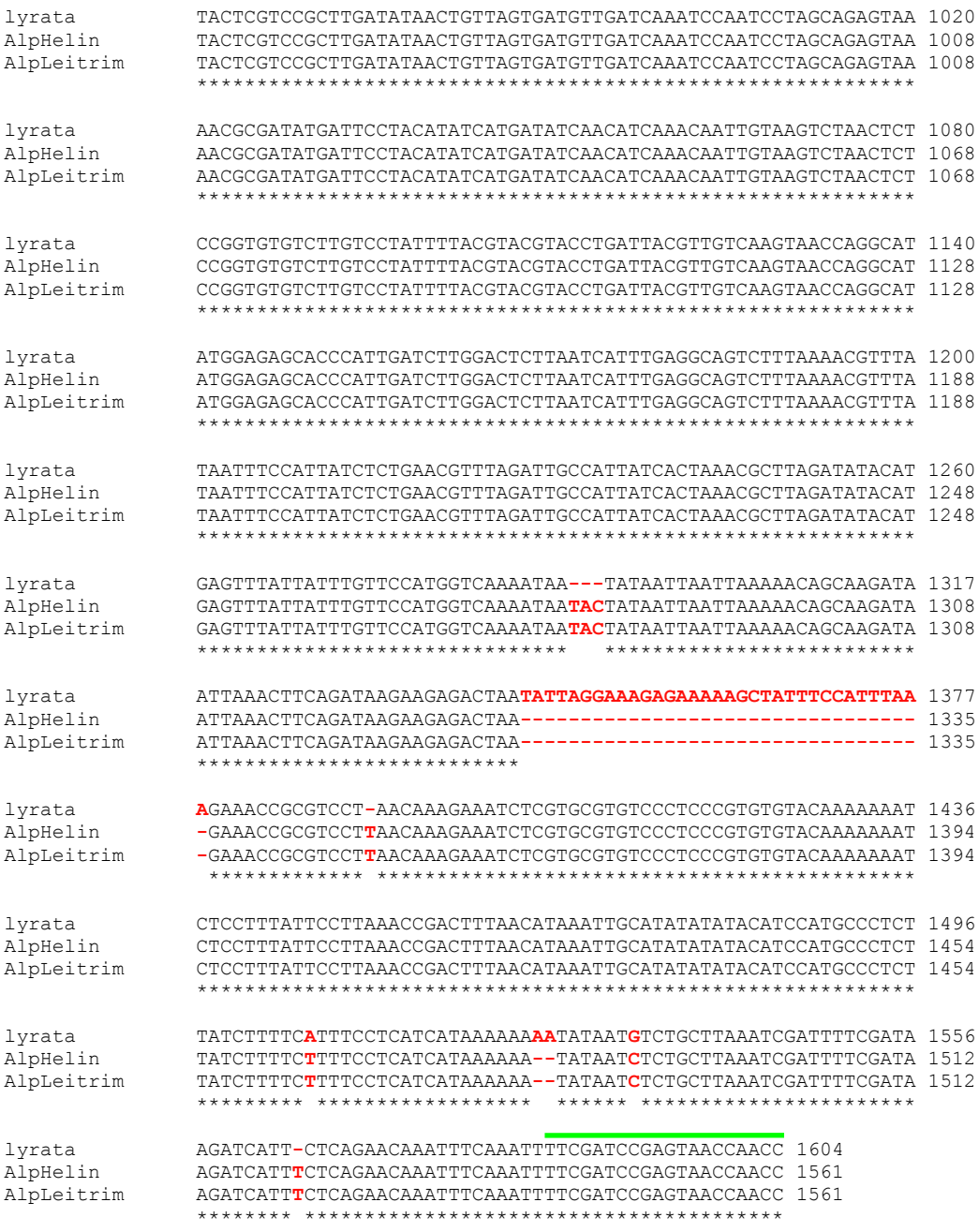
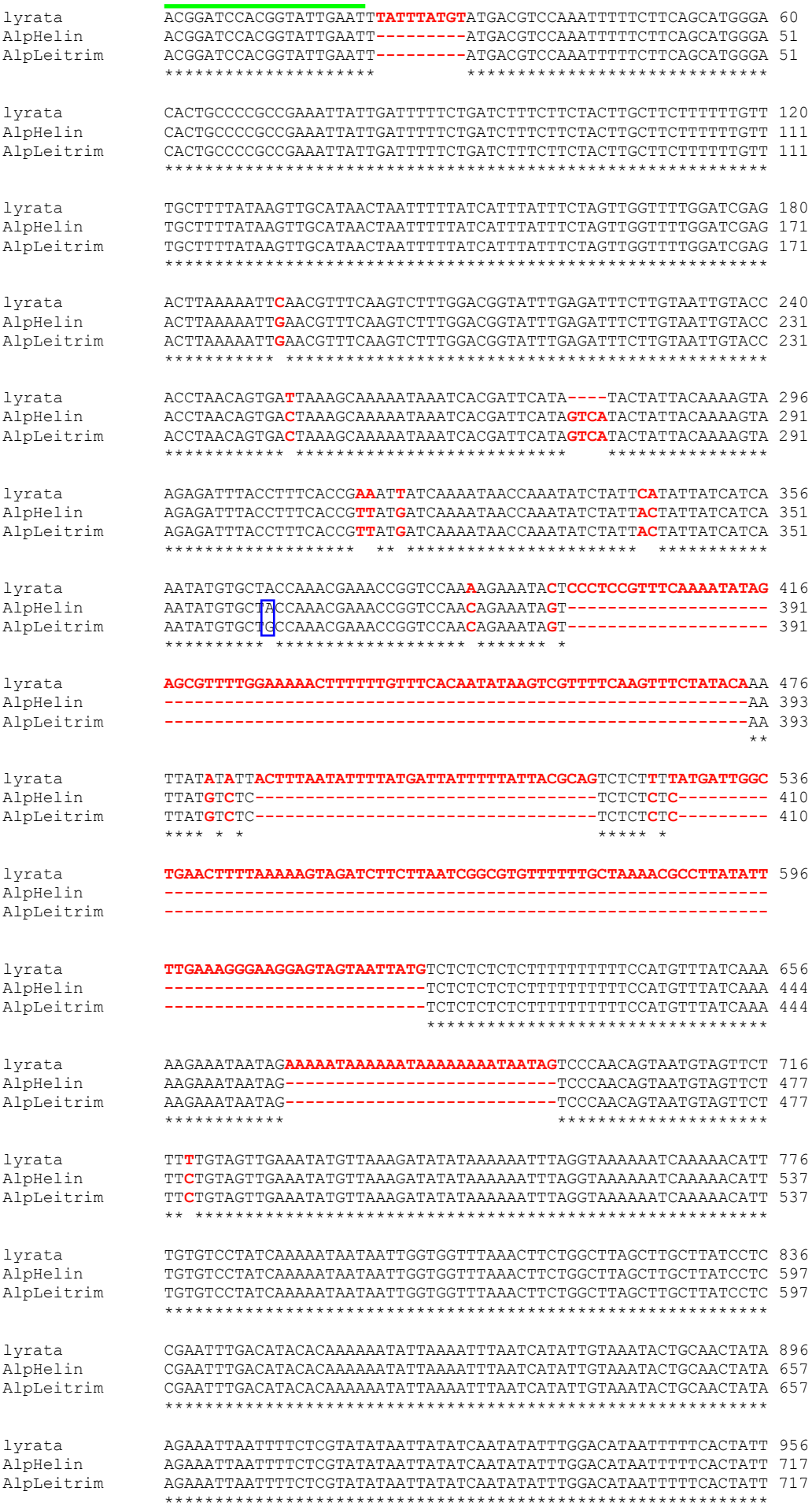


Figure 6.11 The *A. l. petraea* promoter region sequence for At1g09070. The top sequence is the *A. l. lyrata* sequence for the region upstream of the coding sequence for At1g09070. AlpHelin and AlpLeitrim (*A. l. petraea)* refer to the consensus sequences created from sequencing recombinant plasmid DNA inserts in both directions (Helin: three clones; Leitrim: three clones). Asterisk (*) indicate shared identity between all three sequences. Bases coloured in red illustrate differences between the *A. l. lyrata* and *A. l. petraea* sequences that are confirmed by both *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Regions in blue boxes highlight differences between the Helin and Leitrim promoter sequences.


lyrata ACGGATCCACGGTATTGAATTTATTTATGTATGACGTCCAAATTTTTCTTCAGCATGGGA 60
AlpHelin ACGGATCCACGGTATTGAATT---------ATGACGTCCAAATTTTTCTTCAGCATGGGA 51
AlpLeitrim ACGGATCCACGGTATTGAATT---------ATGACGTCCAAATTTTTCTTCAGCATGGGA 51
lyrata CACTGCCCCGCCGAAATTATTGATTTTTCTGATCTTTCTTCTACTTGCTTCTTTTTTGTT 120
AlpHelin CACTGCCCCGCCGAAATTATTGATTTTTCTGATCTTTCTTCTACTTGCTTCTTTTTTGTT 111
AlpLeitrim CACTGCCCCGCCGAAATTATTGATTTTTCTGATCTTTCTTCTACTTGCTTCTTTTTTGTT 111
lyrata TGCTTTTATAAGTTGCATAACTAATTTTTATCATTTATTTCTAGTTGGTTTTGGATCGAG 180
AlpHelin TGCTTTTATAAGTTGCATAACTAATTTTTATCATTTATTTCTAGTTGGTTTTGGATCGAG 171
AlpLeitrim TGCTTTTATAAGTTGCATAACTAATTTTTATCATTTATTTCTAGTTGGTTTTGGATCGAG 171
lyrata ACTTAAAAATTCAACGTTTCAAGTCTTTGGACGGTATTTGAGATTTCTTGTAATTGTACC 240
AlpHelin ACTTAAAAATTGAACGTTTCAAGTCTTTGGACGGTATTTGAGATTTCTTGTAATTGTACC 231
AlpLeitrim ACTTAAAAATTGAACGTTTCAAGTCTTTGGACGGTATTTGAGATTTCTTGTAATTGTACC 231
lyrata ACCTAACAGTGATTAAAGCAAAAATAAATCACGATTCATA----TACTATTACAAAAGTA 296
AlpHelin ACCTAACAGTGACTAAAGCAAAAATAAATCACGATTCATAGTCATACTATTACAAAAGTA 291
AlpLeitrim ACCTAACAGTGACTAAAGCAAAAATAAATCACGATTCATAGTCATACTATTACAAAAGTA 291
lyrata AGAGATTTACCTTTCACCGAAATTATCAAAATAACCAAATATCTATTCATATTATCATCA 356
AlpHelin AGAGATTTACCTTTCACCGTTATGATCAAAATAACCAAATATCTATTACTATTATCATCA 351
AlpLeitrim AGAGATTTACCTTTCACCGTTATGATCAAAATAACCAAATATCTATTACTATTATCATCA 351
lyrata AATATGTGCTACCAAACGAAACCGGTCCAAAAGAAATACTCCCTCCGTTTCAAAATATAG 416
AlpHelin AATATGTGCTACCAAACGAAACCGGTCCAACAGAAATAGT---------- 391
AlpLeitrim AATATGTGCTGCCAAACGAAACCGGTCCAACAGAAATAGT---------- 391
lyrata AGCGTTTTGGAAAAACTTTTTTGTTTCACAATATAAGTCGTTTTCAAGTTTCTATACAAA 476
AlpHelin ----------AA 393
AlpLeitrim ----------AA 393
lyrata TTATATATTACTTTAATATTTTATGATTATTTTTATTACGCAGTCTCTTTTATGATTGGC 536
AlpHelin TTATGTCTC----------TCTCTCTC---------- 410
AlpLeitrim TTATGTCTC----------TCTCTCTC---------- 410
lyrata TGAACTTTTAAAAAGTAGATCTTCTTAATCGGCGTGTTTTTTGCTAAAACGCCTTATATT 596
AlpHelin ----------
AlpLeitrim ----------
lyrata TTGAAAGGGAAGGAGTAGTAATTATGTCTCTCTCTCTTTTTTTTTTCCATGTTTATCAAA 656
AlpHelin ----------TCTCTCTCTCTTTTTTTTTTCCATGTTTATCAAA 444
AlpLeitrim ----------TCTCTCTCTCTTTTTTTTTTCCATGTTTATCAAA 444
lyrata AAGAAATAATAGAAAAATAAAAAATAAAAAAAATAATAGTCCCAACAGTAATGTAGTTCT 716
AlpHelin AAGAAATAATAG----------TCCCAACAGTAATGTAGTTCT 477
AlpLeitrim AAGAAATAATAG----------TCCCAACAGTAATGTAGTTCT 477
lyrata TTTTGTAGTTGAAATATGTTAAAGATATATAAAAAATTTAGGTAAAAAATCAAAAACATT 776
AlpHelin TTCTGTAGTTGAAATATGTTAAAGATATATAAAAAATTTAGGTAAAAAATCAAAAACATT 537
AlpLeitrim TTCTGTAGTTGAAATATGTTAAAGATATATAAAAAATTTAGGTAAAAAATCAAAAACATT 537
lyrata TGTGTCCTATCAAAAATAATAATTGGTGGTTTAAACTTCTGGCTTAGCTTGCTTATCCTC 836
AlpHelin TGTGTCCTATCAAAAATAATAATTGGTGGTTTAAACTTCTGGCTTAGCTTGCTTATCCTC 597
AlpLeitrim TGTGTCCTATCAAAAATAATAATTGGTGGTTTAAACTTCTGGCTTAGCTTGCTTATCCTC 597
lyrata CGAATTTGACATACACAAAAAATATTAAAATTTAATCATATTGTAAATACTGCAACTATA 896
AlpHelin CGAATTTGACATACACAAAAAATATTAAAATTTAATCATATTGTAAATACTGCAACTATA 657
AlpLeitrim CGAATTTGACATACACAAAAAATATTAAAATTTAATCATATTGTAAATACTGCAACTATA 657
lyrata AGAAATTAATTTTCTCGTATATAATTATATCAATATATTTGGACATAATTTTTCACTATT 956
AlpHelin AGAAATTAATTTTCTCGTATATAATTATATCAATATATTTGGACATAATTTTTCACTATT 717
AlpLeitrim AGAAATTAATTTTCTCGTATATAATTATATCAATATATTTGGACATAATTTTTCACTATT 717

```
lyrata          TAAAATTTCTTTATATTAATAGTTTAACTATGTAAACTTTCTTTATATAATTAGTATTTT 1016
AlpHelin        TAAAATTTCTTTATATTAATAGTTTAACTATTTAAACTTTCTTTATATAATTAGTATTTT 777
AlpLeitrim      TAAAATTTCTTTATATTAATAGTTTAACTATTTAAACTTTCTTTATATAATTAGTATTTT 777
                ****************************** ************************** 

lyrata          TTTTTGAAAAATATATTCTTTATATAATTAGTTTAAAATTTGATATATATACAACCTAGC 1076
AlpHelin        TTTTTGAAAAATATATTCTTTATATAATTAGTTTAAAATTTGATATATATACAACCTAGC 837
AlpLeitrim      TTTTTGAAAAATATATTCTTTATATAATTAGTTTAAAATTTGATATATATACAACCTAGC 837
                ************************************************************ 

lyrata          TAATTTCTGTCAAATAGTTGCAACAAACTTTTCAAATTTGTAAATGTGGATATATTATTT 1136
AlpHelin        TAATTTCTGTCAAATAGTTGCAACAAACTTTTCAAATTTGTAAATGTGGATATATTATTT 897
AlpLeitrim      TAATTTCTGTCAAATAGTTGCAACAAACTTTTCAAATTTGTAAATGTGGATATATTATTT 897
                ************************************************************ 

lyrata          TGGGAAACATGTCTATGGTAATCACCTGTATCTTTAATAGTCTACTAGAAACGATATTTG 1196
AlpHelin        TGGGAAACATGTCTATGGTAATCACCTGTATCTTTAATAGTCTACTAGAAACGATATTTG 957
AlpLeitrim      TGGGAAACATGTCTATGGTAATCACCTGTATCTTTAATAGTCTACTAGAAACGATATTTG 957
                ************************************************************ 

lyrata          TCTCCTTCATCGTATAACGTATTTTTATCCATAATGAAAGATCTTGTTTTTAACATGTAC 1256
AlpHelin        TCTCCTTCATCGTATAACGTATTTTTATCCATAATGAAAGATCTTGTTTTTAACATGTAC 1017
AlpLeitrim      TCTCCTTCATCGTATAACGTATTTTTATCCATAATGAAAGATCTTGTTTTTAACATGTAC 1017
                ************************************************************ 

lyrata          TATAGAGATGATTACGACAGTAAATTTAACCCATACATACCCGCCCGCATTCGAAAACTA 1316
AlpHelin        TATAGAGATGATTACGACAGTAAATTTAACCCATACATACCCGCCCGCATTCGAAAACTA 1077
AlpLeitrim      TATAGAGATGATTACGACAGTAAATTTAACCCATACATACCCGCCCGCATTCGAAAACTA 1077
                ************************************************************ 

lyrata          CATACGATAGATATAACTCAATTCAGAATGAT---------CCCATATTTTTAGCCAATC 1367
AlpHelin        CATACGATAGATATAACTCAATTCAGAATGATTTACATGATCCCATATTTTTAGCCAATC 1137
AlpLeitrim      CATACGATAGATATAACTCAATTCAGAATGATTTACATGATCCCATATTTTTAGCCAATC 1137
                ******************************         ******************** 

lyrata          GCTGTGATTAAAATACTTAATAGACAGCTTTCTATGTCTTGTTGTGCTTTTTCTTCTTTT 1427
AlpHelin        GCTGTGATTAAAATACTTAATAGACAGCTTTCTATGTCTTGTTGTGCTTTTTCTTCTTTT 1197
AlpLeitrim      GCTGTGATTAAAATACTTAATAGACAGCTTTCTATGTCTTGTTGTGCTTTTTCTTCTTTT 1197
                ************************************************************ 

lyrata          TTCCCAAAGTCTTTAATGTCTTCTGGACTTTATA-AAGACAGCTTTTTATCTTCACTTTC 1486
AlpHelin        TTCCCAAAGTCTTTAATGTCTTCTGGACCTTATAGAAGACAGCTTTTTATCTTCACTTTC 1257
AlpLeitrim      TTCCCAAAGTCTTTAATGTCTTCTGGACCTTATAGAAGACAGCTTTTTATCTTCACTTTC 1257
                **************************** ***** ************************ 

lyrata          CAACGACTTTCCTTCCTACCATATCCCAAAACCTCTCTCTTTATTCCTTTCGTACTTACA 1546
AlpHelin        CAACGACTTTCCTTCCTACCATATCCCAAAACCTCTCTCTTTATTCCTTTCGTACTTACA 1317
AlpLeitrim      CAACGACTTTCCTTCCTACCATATCCCAAAACCTCTCTCTTTATTCCTTTCGTACTTACA 1317
                ************************************************************ 

lyrata          TCATTTCTTAAAAAACTTTCTTATTTTAATCACTAACTAATTAATAAAATATATGA---- 1602
AlpHelin        TCATTTCTTAAAAAACTTTCTTATTTTAATCACTAACTAATTAATAAAATATACCACTAT 1377
AlpLeitrim      TCATTTCTTAAAAAACTTTCTTATTTTAATCACTAACTAATTAATAAAATATACCACTAT 1377
                ********************************************************  * 

lyrata          --------GAATTTCATTTTGGTTAAAGCTCTCATATAAAAAATTTAATCATGATAAAAA 1654
AlpHelin        TTATTTATGAATTTCATTTTGGTTAAAGCTCTCATATAAAAAATTTAATCATGATAAAAA 1437
AlpLeitrim      TTATTTATGAATTTCATTTTGGTTAAAGCTCTCATATAAAAAATTTAATCATGATAAAAA 1437
                        ************************************************** 

lyrata          CTATGTTATGCCATGATCTGTTGTAAATTGTAACACAATAAAATTTATTAAACTTATATT 1714
AlpHelin        CTATGTTATGCCATGATCTGTTGTAAATTGTAACACAATAAAATTTATTAAACTTATATT 1497
AlpLeitrim      CTATGTTATGCCATGATCTGTTGTAAATTGTAACACAATAAAATTTATTAAACTTATATT 1497
                ************************************************************ 

lyrata          TTGACCTAGGATAAAATTATTTTCTACAAATTTTACTTTTCCCACACTGATAAATTTATA 1774
AlpHelin        TTGACCTAGGATAAAATTATTTTCTACAAATTTTACTTTTCCCACACTGATAAATTTATA 1557
AlpLeitrim      TTGACCTAGGATAAAATTATTTTCTACAAATTTTACTTTTCCCACACTGATAAATTTATA 1557
                ************************************************************ 

lyrata          TATTAATATTATGTCTGTTATAAAGTAAATAAATTATTATTTATTTCTTGTCTCTATCTT 1834
AlpHelin        TATTAATATTATGTCTGTTATAAATTAAATAAATTATTATTTATTTCTTGTCCCTGTCTT 1617
AlpLeitrim      TATTAATATTATGTCTGTTATAAATTAAATAAATTATTATTTATTTCTTGTCCCTGTCTT 1617
                ************************ ************************** ** **** 

lyrata          TGTAGTTTTAGTATAAATAAAGCCCGCTTGCTCTGAGGAACTTCTCATTCACTCGGACCA 1894
AlpHelin        TGTAGTTTTAGTATAAATAAAGCCCGCTTGCTCTGAGGAACTTCTCATTCACTCGGACCA 1677
AlpLeitrim      TGTAGTTTTAGTATAAATAAAGCCCGCTTGCTCTGAGGAACTTCTCATTCACTCGGACCA 1677
                ************************************************************ 
```

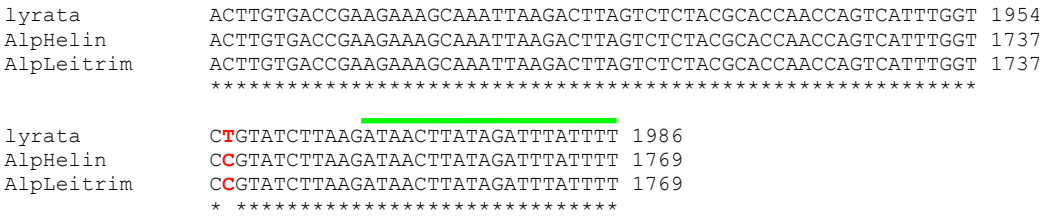


Figure 6.12 The *A. l. petraea* promoter region sequence for At4g31800. The top sequence is the *A. l. lyrata* sequence for the region upstream of the coding sequence for At4g31800. AlpHelin and AlpLeitrim (*A. l. petraea)* refer to the consensus sequences created from sequencing recombinant plasmid DNA inserts in both directions (Helin: two clones; Leitrim: three clones). Asterisk (*) indicate shared identity between all three sequences. Bases coloured in red illustrate differences between the *A. l. lyrata* and *A. l. petraea* sequences that are confirmed by both *A. l. petraea* sequences. Solid green lines indicate the location of primers used for cloning. Regions in blue boxes highlight differences between the Helin and Leitrim promoter sequences.

Motif ID, PLACE identifier, and *SEQUENCE*	**PLACE Database Description**	**Promoter sequence and location (+ or – strand)**
ABRERATCAL (S000507) ***MACGYGB***	'ABRE-related sequence' or 'Repeated sequence motifs' identified in the upstream regions of 162 Ca(2+)-responsive up-regulated genes.	**At5g11740** *A. thaliana*: 1485(+), 1435(-) Helin *A. l. petraea*: 650(-), 1224(+), 1273(-), 1325(+) Leitrim *A. l. petraea*: 651(-), 1225(+), 1274(-), 1326(-). **At5g42050** *A. thaliana*: 223(+), 224(-) Helin *A. l. petraea*: 214(+), 215(-) Leitrim *A. l. petraea*: 206(+), 207(-) **At1g09070** *A. thaliana*: 986(+), 1170(+) *A. l. petraea* Helin: 148(+) *A. l. petraea* Leitrim: 148(+)
CBFHV (S000497) ***RYCGAC***	Binding site of barley CBF1, and also of barley CBF2; CBF= C-repeat (CRT) binding factors; CBFs are also known as dehydration-responsive element (DRE) binding proteins (DREBs).	**At5g11740** *A. thaliana*: 518(+), 554(+), 1315(+), 1315(-), 1652(-) *A. l. petraea* Helin: 1440(-) *A. l. petraea* Leitrim: 1441(-) **At1g09070** *A. thaliana*: 1949(+) *A. l. petraea* Helin: 1411(+) *A. l. petraea* Leitrim: 1411(+)
DRECRTCOREAT (S000418) ***RCCGAC***	Core motif of DRE/CRT (dehydration-responsive element/C-repeat) cis-acting element found in many genes in *Arabidopsis* and in rice.	**At1g09070** *A. thaliana*: 1949(+) *A. l. petraea* Helin: 1411(+) *A. l. petraea* Leitrim: 1411(+)
LTRECOREATCOR15A (S000153) ***CCGAC***	Core of low temperature responsive element (LTRE) of cor15a gene in *Arabidopsis*; A portion of repeat-C (C-repeat),TGGCCGAC, which is repeated twice in *COR15a* promoter (Baker *et al*., 1994); ABA responsiveness;	**At5g11740** *A. thaliana:* 519(+), 555(+), 1723(+) *A. l. petraea* Helin*:* 965(-), 1511(+) *A. l. petraea* Leitrim: 966(-), 1512(+) **At1g09070** *A. thaliana:* 1950(+) *A. l. petraea* Helin: 1412(+) *A. l. petraea* Leitrim: 1412(+) **At4g31800** *A. thaliana:* 509(+), 813(+), 250(-), 256(-), 450(-), 1996(-) *A. l. petraea* Helin: 62(+) *A. l. petraea* Leitrim: 62(+)

Table 6.4 Potential regulatory promoter motifs present in both *A. thaliana* and *A. l. petraea* sequences. The PLACE database (http://www.dna.affrc.go.jp/PLACE/) was used to analyse cis-acting regulatory DNA elements present in the promoter sequences of genes of interest from *A. thaliana* and *A. l. petraea* (both Helin and Leitrim sequences). Based on their annotations, a number were subsequently identified to have potential involvement in regulation of gene expression. Symbols for bases in addition to A/G/C or T: M=A/C, Y=C/T, B=C/G/T, R= A/G.

four transformation types: Helin plants harbouring the Helin promoter construct (HH); Helin plants harbouring the Leitrim promoter construct; Leitrim plants harbouring the Leitrim promoter construct (LL) and; Leitrim plants harbouring the Helin promoter construct (LH). GUS assays were performed only if there were sufficient individuals within at least two unique lines for both transformation types involving Helin plants (HH and HL) or both transformation types involving Leitrim plants (LL and LH). Completion of the quantitative GUS assays was prioritised in those constructs for which there were sufficient individuals for both *A. thaliana* and *A. l. petraea* transformed lines.

6.3.3 Quantification of GUS activity

The quantitative GUS assay was performed for both *A. thaliana* and *A. l. petraea* plants transformed with constructs for the genes At5g42050 and At4g31800. In the time available, insufficient numbers of individuals within successfully transformed lines were generated for *A. l. petraea* transformed with constructs for the three remaining genes of interest (At5g11740, At1g18210, and At1g09070). As a result, quantitative GUS assays could not be carried out for these constructs. For At5g42050 and At4g31800 constructs, GUS activity was quantified for individual plants at more than one time point during the assay. For the majority of samples the rate of GUS activity remained constant over time confirming that samples were obtained during the linear phase of the reaction. The average rate of GUS activity, quantified as production of pmol 4-MU min^{-1} mg $protein^{-1}$, was used in statistical analyses. Lines used in the quantitative GUS assays are detailed in Table 6.3.

6.3.3.1 At5g42050: A. thaliana and A. l. petraea assays

Quantitative GUS assays for transformed *A. thaliana* plants were performed using three lines harbouring the Helin promoter (AtH) and three lines harbouring the Leitrim promoter (AtL) (Figure 6.13A). In the individual lines, effect of the acclimation treatment on GUS activity was not significant for the AtH lines whilst GUS activity was significantly reduced following acclimation in one of the three AtL lines (AtL line 15: $F_{1,9} = 14.42$, $P = 0.005$). In order to ascertain the overall effect of the acclimation treatment on GUS activity for each construct (AtH or AtL) the data for all transformed A. thaliana plants within a construct was put into a general linear model with line and treatment as fixed factors. Mean GUS activity decreased following the acclimation treatment although the effect of treatment was not significant for plants harbouring the Helin promoter ($F_{1,24} = 4.33$, $P = 0.17$) or the Leitrim promoter ($F_{1,24} = 16.48$, $P = 0.06$).

GUS activity level, irrespective of treatment, was variable between lines although the effect of line was only nearing significance for plants harbouring the Helin promoter ($F_{2,24} = 20.66$, $P = 0.046$). Individual plants within a line also demonstrated variable levels of GUS activity as indicated by the standard error bars (Figure 6.13A). The treatment x line interaction was not significant for plants harbouring either the Helin or Leitrim promoter indicating that lines were not responding significantly differently to the acclimation treatment.

For this gene, sufficient *A. l. petraea* individuals were available for at least two lines per transformation type (e.g. AlpHH, AlpHL) and therefore the quantitative GUS assay could be performed for all four transformation types (Figure 6.14). For all the individual lines, the effect of the acclimation treatment on mean GUS activity was not significant. In order to ascertain the overall effect of the acclimation treatment on GUS activity for each transformation type, the data for all plants within a transformation type were put into a general linear model with line and treatment as fixed factors. Mean GUS activity, following the acclimation treatment, decreased in AlpHH plants, increased in both sets of plants harbouring the Leitrim promoter (AlpHL and AlpLL) and did not change in AlpLH plants; these effects of treatment on GUS activity were, however, not significant. The treatment x line interaction was not significant for any of the transformation types indicating that lines within a transformation type were not responding significantly differently to the acclimation treatment.

6.3.3.2 At4g31800: A. thaliana and A. l. petraea assays

GUS assays for transformed *A. thaliana* plants were performed using three lines harbouring the Helin promoter and three lines harbouring the Leitrim promoter (Figure 6.13B). In the individual lines, effect of treatment was not significant the AtH lines whilst GUS activity was significantly reduced in one of the three AtL lines (AtL line 20: $F_{1,9} = 11.28$, $P = 0.01$). In order to ascertain the overall effect of the acclimation treatment on GUS activity for each construct (AtH or AtL) the data for all plants within a construct was put into a general linear model with line and treatment as fixed factors. Mean GUS activity decreased following the acclimation treatment although the effect of treatment was not significant for plants harbouring the Helin promoter ($F_{2,24} = 7.76$, $P = 0.11$) or the Leitrim promoter ($F_{2,24} = 1.05$, $P = 0.41$). Individual plants within a line had variable extents of GUS activity, indicated by the standard error bars (Figure 6.13B). The treatment x line interaction was not significant for plants

harbouring the Helin promoter ($F_{2,24}$ =0.55, P = 0.56) indicating that lines were not responding significantly differently to the acclimation treatment. The treatment x line interaction for plants harbouring the Leitrim promoter was, however, highly significant ($F_{2,24}$ = 10.86, P <0.001) indicating that lines were responding significantly differently to the acclimation treatment; line twenty demonstrates a large decrease in GUS activity following acclimation whilst activity in lines ten and 21 did not change (Figure 6.13B).

For this gene sufficient *A. l. petraea* individuals were available for three lines of the AlpHH and AlpHL transformation types whilst there were not sufficient individuals within lines for the two transformation types that involved plants from Leitrim (AlpLL and AlpLH). The quantitative GUS assay was, therefore, performed using only two of the four possible transformation types (AlpHH and AlpHL, Figure 6.15). The effect of the acclimation treatment on GUS activity was not significant for any of the individual lines. In order to ascertain the overall effect of the acclimation treatment on GUS activity for both of the transformation types, the data for all plants within a transformation type were put into a general linear model with line and treatment as fixed factors. Mean GUS activity, following the acclimation treatment, was unchanged in AlpHH plants and decreased in AlpHL plants; the effect of treatment was not significant for either transformation type (AlpHH: $F_{1,18}$ = 0.001, P = 0.98; AlpHL: $F_{1,12}$ = 2.91, P = 0.17). The treatment x line interaction was not significant for either of the transformation types indicating that lines within a transformation type were not responding significantly differently to the acclimation treatment.

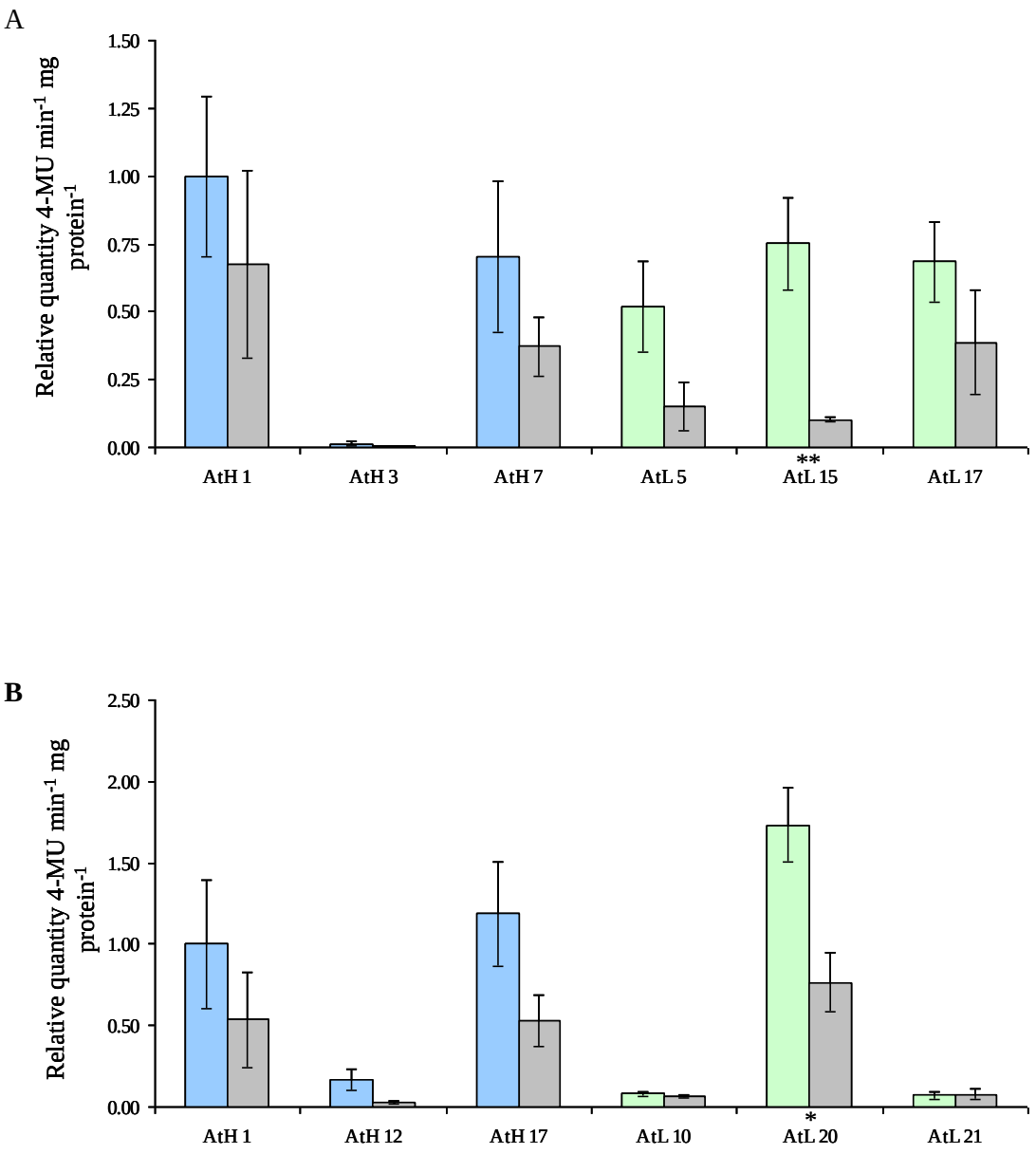


Figure 6.13 Quantitative GUS assay for *A. thaliana* plants harbouring *A. l. petraea* gene promoters fused to GUS. A: *A. thaliana* plants harbouring population specific *A. l. petraea* promoters for At5g42050 fused to the GUS gene. B: *A. thaliana* plants harbouring population specific *A. l. petraea* promoters for At4g31800 fused to the GUS gene. In both A and B, blue bars represent untreated plants harbouring the promoter from the Helin population (AtH) and green bars represent untreated plants harbouring the promoter from the Leitrim population (AtL). Grey bars represent corresponding cold treated samples; treated plants were exposed to a two day acclimation at 4^{o}C. Numbers on the x axis refer to individual transformed lines. All GUS activity is expressed relative to the Helin line one (AtH 1) untreated sample for each gene. Asterisk (*) indicate the level of the significant difference between untreated and treated samples within an individual line. * P <0.05; ** P <0.01.

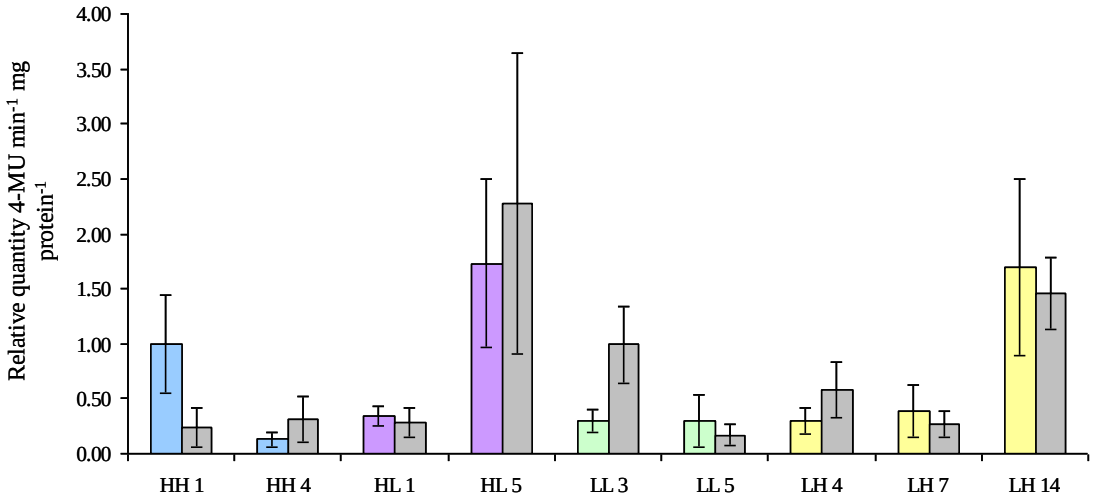


Figure 6.14 Quantitative GUS assay for *A. l. petraea* plants harbouring the At5g42050 gene promoter fused to GUS. The coloured bars represent untreated plants from either Helin or Leitrim harbouring either the Helin or Leitrim promoter fused to the GUS gene; Blue bars: Helin plants with the Helin promoter (HH); purple bars: Helin plants with the Leitrim promoter (HL); green bars: Leitrim plants with the Leitrim promoter (LL); yellow bars: Leitrim plants with the Helin promoter (LH). Grey bars represent corresponding cold treated samples; treated plants were exposed to a two day acclimation at 4^{o}C. Numbers on the x axis refer to individual transformed lines. All GUS activity (±SE) is expressed relative to the HH line one untreated sample.

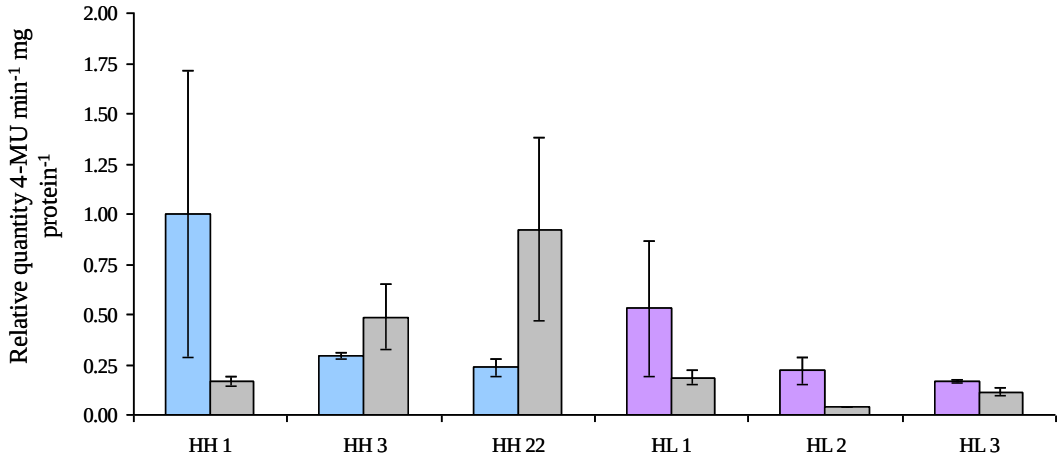


Figure 6.15 Quantitative GUS assay for *A. l. petraea* plants harbouring the At4g31800 gene promoter fused to GUS. Blue bars represent untreated Helin plants harbouring the promoter from the Helin population (HH) fused to the GUS gene, and purple bars represent untreated Helin plants harbouring the promoter from the Leitrim population (HL) fused to the GUS gene. Grey bars represent corresponding cold treated samples; treated plants were exposed to a two day acclimation at 4°C. Numbers on the x axis refer to individual transformed lines. All GUS activity (±SE) is expressed relative to the HH line one untreated sample.

6.4 Discussion of promoter:GUS fusion results

6.4.1 Promoter regions sequences and potential regulatory motifs

The *A. l. petraea* sequences for promoter regions obtained in this study demonstrated at least 85% identity with the available sequences for the North American subspecies *A. l. lyrata*, yet sequences for both subspecies exhibited limited similarity to the corresponding *A. thaliana* sequences. Similarity between *A. thaliana* and *A. l. petraea* coding region sequences for genes studied in chapter four (section 4.3.2) was far greater than that observed between promoter sequences studied in this chapter; this suggests that coding and promoter regions are subject to differing selective constraints. Although there are no published studies concerning differentiation between *A. l. petraea* and *A. thaliana* promoter region sequences, greater sequence divergence has been reported between intragenic regions of DNA than for exonic regions (Le Corre et al, 2002; Chapter 4, section 4.4.2). Sequence differentiation has been reported between distinct populations of *A. l. petraea* (Ramos-Onsins *et al.*, 2004), as well as between ecotypes of *A. thaliana* (Kuittinen and Aguade, 2000; Hauser *et al.*, 2001; Le Corre *et al.*, 2002). Here, promoter region sequences for five genes were found to differ between Helin and Leitrim and, with the exception of At5g42050, these differences involved single nucleotide polymorphisms or indels. Analysis of these sequences using PLACE (section 6.3.1) identified differences in promoter motifs that corresponded with these minor sequence differences; annotations of these motifs, however, do not suggest roles in regulation of gene expression in response to stress. Many of the identified motifs have been reported in other plant species rather than *Arabidopsis*, thus limiting scope for discussion of their potential roles in *A. l. petraea.* The minimal sequence differentiation between promoters of Helin and Leitrim indicates that the promoter may have a limited role in the differential regulation of gene expression between populations. Interestingly a number of motifs with annotations that may indicate roles in regulation in response to stress were identified that were present in both *A. l. petraea* sequences and the *A. thaliana* sequence. As the *A. thaliana* and *A. l. petraea* sequences were relatively dissimilar, the presence of the same motif within both provides stronger support for roles in regulation of gene expression.

The motif ABRERATCAL that has been identified in the upstream region of calcium-responsive genes was present in both *A. l. petraea* sequences and the *A. thaliana* sequence for three of the genes of interest: At5g11740, At5g42050, and At1g09070.

Calcium levels are known to increase dramatically in plants following exposure to abiotic stresses (Knight *et al.*, 1991; Knight *et al.*, 1996), and cellular calcium is essential for the cold acclimation response and cold-induced gene expression in *Arabidopsis* (Tähtiharju *et al.*, 1997). The presence of this motif therefore strongly indicates regulation of these genes in response to enhanced calcium levels during cold stress. The motif is also present in only the *A. thaliana* sequence for At1g18210 and At4g31800.

Components homologous to the *Arabidopsis* CBF cold response pathway have been reported in a range of plant species including barley (Choi *et al.*, 2002). The motif CBFHV, annotated as the binding site for *CBF1* and *CBF2* in barley, is present in the promoter sequences of both species for At5g11740 and At1g09070. This could indicate that these two genes are regulated as part of the *Arabidopsis* CBF cold response. The motif is also present in the *A. thaliana* promoter sequence for At4g31800. CBF genes are also referred to as DREBs (dehydration responsive element binding proteins), and encode proteins that bind to the DRE element motif in the promoter sequence of a gene to activate its transcription (Yamaguchi-Shinozaki and Shinozaki, 2005). The dehydration responsive element (DRE) was originally identified as an essential *cis*-acting element for regulation of the gene *RD29A* (also called *COR78* and *LTI78*) in response to cold and dehydration (Yamaguchi-Shinozaki and Shinozaki, 1994), but has since been reported in a number of cold and drought inducible genes (Thomashow, 1999; Shinozaki and Yamaguchi-Shinozaki, 2000). The motif DRECRTCOREAT is present in the *A. l. petraea* and *A. thaliana* sequences for At1g09070. This motif (CCGACR, where R = G or A) is the core motif for DRE and thus provides support for the hypothesis that the promoter of this gene is involved in regulation of expression in response to abiotic stresses. A number of *Arabidopsis* genes responsive to both cold and drought, identified using a full length cDNA microarray, contained this core motif in their promoters and are reported as *DREB1A/CBF3* target genes (Seki *et al.*, 2001). Two *cis*-acting elements, low-temperature responsive element (LTRE) and C-repeat (CRT) both contain the motif that forms the core of the DRE sequence and regulation of cold-inducible promoters has been reported for both elements (Baker *et al.*, 1994; Thomashow, 1999). The motif LTRECOREATCOR15A is present in both the *A. l. petraea* promoters and the *A. thaliana* promoter for the genes At5g11740 and At1g09070, and may thus be involved in the regulation of this gene in response to cold stress. It is also present in the *A. l. petraea* promoter sequences for At1g18210.

Another LTRE (LTRE1HVBLT49) that has been identified in a barley gene promoter is present in the At4g31800 promoters of both *Arabidopsis* species and may have involvement in cold temperature-induced regulation of this gene. Of the five genes studied, Helin and Leitrim promoters for the gene At5g42050 are the only sequences lacking an LTRE motif; they do, however, contain a CRT element (CRTDREHVCBF2). This core motif has been identified as a binding site for CBF2 proteins in barley, with the highest levels of binding activity reported at cold temperatures (Xue, 2003).

Another motif of interest is CCA1ATLHCB1, the binding site for the transcription factor *CCA1 (*CIRCADIAN CLOCK ASSOCIATED1). This is present in the *A. l. petraea* and *A. thaliana* promoter sequences for At1g09070. Expression of *CCA1* was significantly up-regulated following cold temperature exposure in Helin and Leitrim (C. J. Lilley, unpublished data) and in thirteen additional populations of *A. l. petraea* (chapter four, section 4.3.5.2). This enhanced expression is indicative of importance in the cold temperature response of *A. l. petraea* as discussed in section 4.4.3. The presence of the binding site for *CCA1* in the promoter region of At1g09070 suggests that the cold-responsive *CCA1* transcription factor may regulate expression of this gene upon exposure to cold temperatures. At1g09070 (SRC2: SOYBEAN GENE REGULATED BY COLD2) is down-regulated in Helin and up-regulated in Leitrim following acclimation (Figure 6.1), and has been reported as cold-responsive in *A. thaliana* (Robinson and Parkin, 2008). It has also been recognised as a member of the universal stress response transcriptome that includes genes responsive to a range of stresses (e.g. cold, osmotic, biotic), and is involved in vesicle transport (Ma and Bohnert, 2007).

A potentially interesting motif, present only in the *A. l. petraea* promoter sequences for At1g18210, is TCA1MOTIF. This motif consists of ten nucleotides (TACTCCTCCT), and is conserved among stress inducible genes from a range of plants species including barley, potato, and *Arabidopsis* (Goldsbrough *et al.*, 1993). Little is known of the function of At1g18210 in *A. thaliana* but it is putatively involved in calcium binding. Expression of this gene has been reported to increase in response to potassium, wounding, and nematode infection in *A. thaliana* but expression changes following exposure to cold temperatures were limited (GENEVESTIGATOR). Following acclimation in *A. l. petraea*, this gene was down-regulated in Helin and up-regulated in Leitrim although, interestingly, post-acclimation signal values were very high in both

populations (Figure 6.1). The motif TCA1MOTIF in the promoter of both *A. l. petraea* populations may function in the regulation of this gene in response to cold stress.

6.4.2 Transformation success in *A. thaliana* and *A. l. petraea*

Transformations involving *A. thaliana* plants were generally more successful, in terms of numbers of unique transformed lines obtained, than those involving *A. l. petraea*. The lifecycle differences between *A. thaliana* and *A. l. petraea*, most notably the fact that *A. l. petraea* is self-incompatible, necessitated the use of an alternative method of plant transformation to the frequently used floral dip technique (section 2.3.22). Although the alternative method (section 2.3.25) did yield transformants for the majority of transformations performed, numbers of unique lines obtained were fewer. The requirement for subculturing of *A. l. petraea* individuals to allow for replication within a line also reduced the numbers of lines available for investigation using a quantitative GUS assay. The availability of seed, however, minimised this issue for the *A. thaliana* lines.

6.4.3 Analysis of promoter region function using fluorometric GUS assays

Quantitative GUS assays were performed for both *A. thaliana* and *A. l. petraea* plants transformed with constructs containing either the Helin or the Leitrim promoter fused to the GUS reporter gene for two of the genes of interest, At5g42050 and At4g31800. As illustrated in Figures 6.13 - 6.15, GUS activity varied considerably between replicate lines and between individuals within a line for all transformation types. Consequently, results obtained from these assays did not detect a significant effect of the cold acclimation treatment on GUS activity in the overall results for each transformation type. For *A. thaliana*, variation between individuals within a line may be partly due to the inclusion of a mix of heterozygotes and homozygotes for the transformed gene. Time constraints prevented the additional generation of plant growth that would have ensured only homozygotes were used. The inclusion of kanamycin in plant growth media, however, excluded the use of wild type plants.

In quantitative GUS assays performed for *A. thaliana*, mean GUS activity (pmol 4-MU^{-1} min^{-1} mg protein^{-1}) did decrease following cold acclimation for plants harbouring the Helin and Leitrim promoters for both genes studied. This effect of cold treatment, although not significant, indicates a reduction in the quantity of GUS enzyme within the

protein extract of acclimated plants. As production of GUS was under the control of the Helin and Leitrim promoters for genes At5g42050 and At4g31800, it can be suggested that these promoters initiate higher rates of transcription in untreated plants than in those exposed to cold temperatures. This partially supports results obtained from the microarrays for the gene At4g31800 as expression of this gene was reduced in Helin plants following acclimation. In Leitrim plants the arrays revealed an up-regulation of this gene in response to acclimation that is therefore not supported by these GUS assay results. The results from the *A. thaliana* GUS assays are also in contrast to the array results for the gene At5g4250, as this gene demonstrated increased expression in both populations (to a greater extent in Leitrim) following exposure to cold temperatures. The quantitative GUS assays performed using transformed *A. thaliana* therefore suggest that the population-specific promoters are unlikely to confer the differential regulation of gene expression observed in the microarrays. It is important to interpret these results with caution as the observed decreases in GUS activity following acclimation were not significant (see above). Published studies that have performed fluorometric GUS assays in *A. thaliana* tend to involve between three and five independent lines per construct, and three replicate GUS assays per line (Wang *et al.*, 2005; Li *et al.*, 2010); in this work three independent lines were used with five replicates per line. Incorporation of an increased number of lines per construct would increase confidence in the results obtained and perhaps reveal a significant effect of treatment but this was not possible due to time constraints.

The quantitative GUS assays were also performed using transformed *A. l. petraea* plants from both Helin and Leitrim. It was important to include these assays, despite the reduced ease of performing *A. l. petraea* transformations, as the Helin and Leitrim promoters may not function in the same way in *A. thaliana* plants as they do in *A. l. petraea*. As in the *A. thaliana* assays, the effect of acclimation on GUS activity was not significant. Mean GUS activity (pmol 4-MU^{-1} min^{-1} mg $protein^{-1}$) for plants harbouring the At5g42050 constructs did alter following acclimation for three of the four transformation types (AlpHH, AlpHL, and AlpLL). GUS activity increased in both sets of plants harbouring the Leitrim promoter for this gene, and decreased in AlpHH plants. In the arrays, expression of this gene increased in both populations following acclimation but this increase was significantly higher in Leitrim. The observed post-acclimation increase in mean GUS activity for plants harbouring the Leitrim promoter may therefore indicate that the Leitrim promoter is more effective in increasing

transcription of this gene in response to cold temperatures. *A. l. petraea* GUS assay results for the gene At4g31800 were only performed using Helin tranformants due to lack of available Leitrim transformants. The three AlpHH lines did not exhibit a consistent response to acclimation (Figure 6.15), thus, mean GUS activity did not change in AlpHH plants, whilst mean GUS activity in AlpHL plants decreased following cold temperature exposure. In the arrays, expression of At4g31800 decreased in Helin and increased in Leitrim (Figure 6.1). These assay results therefore suggest that the population specific promoter is not responsible for the observed differential regulation of gene expression in these populations following acclimation. Even if the differential regulation of gene expression observed in the microarrays is not dependent solely on the gene promoter, correct regulation of gene expression might be expected when a promoter is in its correct genetic background. For example, Helin plants harbouring the Helin promoter:GUS construct (AlpHH) might be expected to exhibit the same regulation of gene expression as observed in the arrays as all other required regulatory components should be both present and responding to the acclimation treatment. As this was not the case for the majority of AlpHH or AlpLL plants, other factors may be responsible for the dissimilarity between array and quantitative GUS assay results. It is possible that the length of promoter regions included within the constructs did not contain all the regulatory regions that play a role in expression of the endogenous gene. For At4g31800 the promoter region in *A. thaliana* extends further upstream than 2000 bp and this could also be the case in *A. l. petraea*. It is therefore possible that the region used in the promoter:GUS construct may be missing some important regulatory elements despite the presence of some motifs clearly related to abiotic stress responses. For At5g42050 the *A. thaliana* promoter region for this gene was only 880 bp in length. The promoter region was found to be of a similar length in *A. l. lyrata* as sequence further than approximately 900 bp upstream of the coding sequence corresponded to coding sequence of the previous gene. The entire length of the *A. l. petraea* promoter region is therefore likely to have been included in the At5g42050 promoter:GUS construct, although the possibility of important regulatory elements being present further upstream than this cannot be ruled out. The known stability of the GUS protein (Jefferson *et al.*, 1987; Kavita and Burma, 2008) may have also contributed to the observed dissimilarity between the array and the GUS assay results, particularly in those assays for which a reduction in GUS activity was expected following exposure to acclimation. For example, a reduction in GUS activity was expected in the *A. l. petraea* Helin plants that were harbouring the Helin At4g31800

gene promoter (AlpHH) yet no significant change in GUS activity was observed. Transcript levels of this gene were reduced following exposure to acclimation as detected in the microarray experiments, but the stability of the GUS protein may have prevented detection of this in the quantitative GUS assays. Investigation of expression of the endogenous gene in the transgenic *A. l. petraea* plants that were involved in the quantitative GUS assays could be performed, using qRT-PCR, to confirm whether or not the same overall response was being induced as seen in the microarray experiments. Also, as with the *A. thaliana* GUS assays, incorporation of a greater number of lines per transformation type would be beneficial. This was not possible as sufficient individuals were not yet available within the majority of successfully transformed lines.

6.4.4 Conclusions

The results obtained from these quantitative GUS assays, along with the limited sequence differences that exist between the *A. l. petraea* promoters studied, indicate that other factors aside from the promoters of these genes are likely to be responsible for their observed differential expression. Differences may exist between the populations in genes upstream of these in the stress response pathway, resulting in divergences in signalling cascades and consequently causing differential regulation of gene expression. Although there were some promoter motif differences between the Helin and Leitrim promoters for specific genes, it was not possible to link any of these to response to stress. There were, however, a number of *cis*-acting elements ubiquitous to all sequences for individual genes that have known roles in response to stress and particularly in response to cold. These elements act as binding sites for the protein products of upstream genes and therefore differences in expression of these upstream genes between populations may affect binding to, and the resultant transcription of, the genes in this study. Further investigations would be required to test this hypothesis. For example upstream cold response genes, such as the CBF genes, could be examined in both Helin and Leitrim. Using loss-of-function *A. l. petraea* plants, expression of the five interaction genes could be studied and if differential regulation was no longer observed in response to cold this would suggest that such upstream genes were important in this response. Differences in the promoter regions of these upstream genes may also be involved and this could also be deduced using a transgenic system.

CHAPTER 6: SUMMARY

- Microarray analysis, presented in Chapter 5, identified a set of genes that were responsive to an acclimation treatment but differentially regulated between two populations of *A. l. petraea*. These populations were Helin in Norway and Leitrim in Ireland. A subset of these genes was identified for further investigation.
- Promoter region sequences were obtained for both the Helin and Leitrim populations. These sequences had high homology to the *A. l. lyrata* sequence, but differed considerably to the *A. thaliana* homolog promoter sequences. Limited differences were observed between the Helin and Leitrim sequences for each gene's promoter.
- Promoter sequences were analysed using the PLACE plant *cis*-acting motif database (http://www.dna.affrc.go.jp/PLACE/). A number of elements were identified in both *A. l. petraea* sequences and the *A. thaliana* sequence that could have potential involvement in regulation of gene expression in response to stress.
- Promoter:GUS fusions were constructed for both the Helin and Leitrim promoters of each gene. These constructs were used to transform *A. thaliana* and *A. l. petraea* (Helin and Leitrim plants). Transformants were subject to fluorometric GUS assay analysis.
- GUS activity was not significantly affected by cold acclimation in these transformed plants. This suggests that it is not the promoters of these interaction genes that are responsible for observed population-specific gene expression. For the transformed *A. l. petraea* plants, however, correct regulation of expression was not observed even when promoters were in their correct genetic background. This indicates that other factors may have contributed to the disparity between microarray and quantitative GUS assay results.

7. General Discussion

7.1 Genetic variation and local adaptation

This study has successfully identified a range of phenotypic differentiation between geographically distinct populations of *A. l. petraea* including differences in life history characteristics and in gene expression. Populations from throughout the north west European range of this species experience a wide variety of climatic conditions (Figure 1.2) and field locations also differ in soil characteristics such as pH and organic matter content (Dr P. Vergeer, *pers. comm.*). Populations included within this study were chosen to ensure variation in: (i) altitude (high and low); (ii) position within geographic range (centre and margin); and (iii) estimated population size. Plants both perceive and respond to environmental variables at multiple scales from the level of the genes to that of the phenotypes (Smith and Gallon, 2001) and thus it may be expected that populations that are exposed to differing conditions may become distinct at the genetic and/or the phenotypic level. Cloning and sequencing of gene coding regions and promoter regions (Chapters 4 and 6) revealed genetic differences between study populations. Genetic differentiation between populations has been reported previously for *A. l. petraea* (Ramos-Onsins *et al.*, 2004); this differentiation suggests that limited gene flow occurs between populations and thus local adaptation is a feasible possibility. Significant phenotypic differences were also evident between the populations under investigation. Within the selection experiment (Chapter 3) populations exhibited differential growth rates under the three temperature regimes, perhaps indicating their adaptation to differing temperature optima. Differentiation of leaf shape between populations was also observed in concurrence with differences reported by Jonsell *et al.* (1995). Populations within close geographic proximity to one another displayed varying leaf shapes, both in this study and in work by Jonsell *et al.* (1995), further supporting the hypothesis for limited gene flow between *A. l. petraea* populations. It is possible that vegetative propagation contributes to the maintenance of variation in leaf shape (Jonsell *et al.*, 1995). Phenotypic variation in the proportion of plants that flowered was also observed with Norwegian and Swedish populations demonstrating higher flowering percentages than those from Britain; this pattern is in concurrence with previously collected data from plants reared under growth cabinet conditions (P. Vergeer, unpublished data). As discussed in Chapter 3, this may relate to differences in life history strategy between the populations.

Gene expression, both under control conditions and in response to stress, also differed significantly between populations of *A. l. petraea*. Thirteen populations of *A. l. petraea,* investigated using quantitative RT-PCR (Chapter 4), demonstrated differences in basal expression levels of a number of genes including a putative sugar transporter (At1g08920: *ESL1*) and a gene involved with the circadian clock (At2g46830: *CCA1*). The differing field conditions perceived by the populations under investigation leads to the hypothesis that observed inherent differences in gene expression are due to local adaptation. Microarray analysis (Chapter 5) was also successful in identifying over 200 genes that were inherently differentially expressed ($P < 0.05$, fold change >two) between a central (Helin) and a marginal (Leitrim) population of *A. l. petraea*. Over 180 of these genes overlap with genes that were found to be differentially expressed between these same two populations in a previous microarray study (C. J. Lilley, unpublished data); this is encouraging as it suggests that expression differences between populations are real and not an artefact of the experimental protocol, and thus further supports the hypothesis that *A. l. petraea* populations are locally adapted. A number of published studies further maintain this hypothesis in *A. l. petraea*. Populations of *A. l. petraea* consist of differing proportions of glabrous and trichome-producing individuals and divergent selection on the locus controlling this trait is evident (Karkkainen *et al.*, 2004). These differences in trichome coverage correspond to differences in resistance to herbivory (Kivimaki *et al.*, 2007). As trichomes are likely to be a costly adaptation, they are liable to be selected against in regions of low threat from herbivory. Phenotypic variation in resistance to powdery mildew (*Golovinomyces orontii*) has been reported amongst populations of *A. l. petraea,* and this is reflected in variation in the resistance gene *RPW8* between populations (Jorgensen and Emerson, 2009). A reciprocal transplant experiment involving Norwegian, Swedish, and German populations of *A. l. petraea* found that, after one year, cumulative fitness was higher in the local population than at least one of the non-local populations and this demonstrates adaptive population differentiation (Leinønen *et al.*, 2009). Vergeer *et al.* (2008) investigated the responses of *A. l. petraea* populations to nitrogen deposition and attributed differences in phenological and physiological variables to local adaption within the populations. Metabolite profiling has revealed metabolic differences between populations of *A. l. petraea* and differences in metabolic phenotypes were also evident following exposure to cold temperatures; this may signify adaptation to local climates (Davey *et al.*, 2008; Davey *et al.*, 2009).

There are a number of mechanisms that may be responsible for local adaptation in *A. l. petraea*. The ultimate source of phenotypic variation between populations is variation in the genomic sequence (Siomos, 2009). Sequence variations can arise as the result of mutations and, if these prove to be beneficial within the population, such variations can be selected for through the process of natural selection and thus become locally prevalent. Isolation of *A. l. petraea* populations reduces potential for gene flow between populations and thus may result in population-specific gene sequences. Differences in the promoter region of a gene could affect the gene's own expression while differences in the coding region of a regulatory or signalling gene may potentially affect expression of a downstream gene. Sequence differences between Helin and Leitrim may therefore be responsible for some of the observed inherent differences in gene expression. The genomic DNA sequence is not, however, solely responsible for phenotype and thus the phenotype of an organism can be altered without a requirement for changes in the coding sequence or promoter region of a gene (Grant-Downton and Dickinson, 2005; Rapp and Wendel, 2005). This phenomenon is known as epigenetics and refers to heritable molecular events that do not result in DNA sequence changes (Bender, 2002). The association of DNA with histone proteins enables DNA to be packaged into the nucleus of cells and the resultant DNA-protein complex is referred to as chromatin. Some regions of the genome can be loosely packaged (euchromatin) whilst others are tightly packaged (heterochromatin) and this can affect patterns of gene expression (Bender, 2002). The condensation level of chromatin can be controlled by histone modifications and the methylation of DNA (Rapp and Wendel, 2005; Grant-Downton and Dickinson, 2006). These processes are both epigenetic mechanisms that have the potential to influence gene expression. In plants, stable epigenetic information can be transmitted between generations (Bender, 2002; Grant-Downton and Dickinson, 2006) and could therefore be responsible for observed differences in gene expression between populations of *A. l. petraea*. It has been suggested that environmental stresses, such as exposure to extreme temperatures, could disturb epigenetic systems and that subsequent epigenetic modifications may result in the creation of novel phenotypes (Rapp and Wendel, 2005; Grant-Downton and Dickinson, 2006). It is therefore feasible to suggest that exposure of the Helin population of *A. l. petraea* to frequent cold periods may have resulted in epigenetic changes leading to differing patterns of gene expression to those seen in the Leitrim population.

The biogeographic distributions of many plant species encompass a wide range of environmental variables and this in itself can arguably be viewed as evidence for evolution of local adaptations (Perez De La Vega, 1996). Local adaptation is apparent in other plant species including *Boechera holboellii*, a perennial relative of *A. thaliana* (Knight *et al.*, 2006). Two populations experiencing contrasting hydrology and altitudes were studied; in reciprocal transplant experiments individuals had significantly increased survival in their native habitats. Differential investment in root biomass was evident between populations, and differential regulation of gene expression was observed in response to drought stress (Knight *et al*., 2006). Differences in water use efficiency between populations of *Impatiens capensis* have also been linked to local adaptation (Heschel *et al.*, 2002). Another study identified differential effects of ozone exposure on reproductive development in resistant and sensitive populations of *Plantago* (Black *et al.*, 2010), and sensitivity or resistance to ozone was correlated with extent of previous ozone exposure in the field (Reiling and Davison, 1992; Pearson *et al.*, 1996). The ability of plant species to colonise environments less favourable to growth, such as serpentine soils, strongly supports the hypothesis that plants are able to evolve adaptations that enable persistence in such habitats. Recent work involving populations of *A. l. petraea* from both serpentine and granitic soils successfully identified polymorphisms associated with soil type (Turner *et al.*, 2010). The locations of these polymorphisms are enriched for loci involved in processes such as heavy-metal detoxification and are thus strong candidates for local adaptation to serpentine conditions. In the onset of predicted climatic changes, the ability to adapt to local climatic conditions will be essential if a species is to persist in its current location, and this process will require sufficient genetic variation to be present within a species.

7.2 Response to abiotic stress in *A. l. petraea*

The global expression profiling of *A. l. petraea* by microarray analysis following exposure to cold temperature stresses is in general concordance with the *A. thaliana* cold stress literature. A number of genes annotated as cold responsive and with known functions in cold response regulatory pathways were shown to respond to acclimation in *A. l. petraea* including *COR15A* (At2g42540), *COR47* (At1g20440) and *COR78* (At5g52310). There was also considerable overlap of the *A. l. petraea* acclimation responsive genes with the core set of cold responsive genes described by Vogel *et al.* (2005). This concordance of responses to cold indicates the conservation of cold response mechanisms between the two species. Some key genes in the *A. thaliana* cold

response, such as *ICE1* and CBF/DREB genes, were not found to be differentially regulated in response to cold in the microarray experiments. These transcriptional activators are known to be up-regulated rapidly following exposure to cold temperatures and are involved in the regulation of downstream genes that have functions in stress tolerance (Thomashow, 2001). Following this rapid up-regulation, their expression may have then returned to similar levels as in the untreated plants, hence, at the time of sample collection no change in their expression was detected. The additional freezing stress enabled identification of a set of genes responsive to freezing in *A. l. petraea*, many of which were annotated as being of unknown function. This emphasises current limits in knowledge of the transcriptomic responses of *Arabidopsis* to freezing stress.

Investigation of the response to cold in *A. l. petraea* also revealed further evidence for local adaptation as populations exhibited different responses to cold stress. Microarray analysis identified a set of genes that were regulated following cold temperature exposure but that, in addition, were also differentially regulated between the two populations under investigation (Helin from Norway and Leitrim from Ireland). This implies that the populations require different levels of expression of particular genes to confer the same protection against cold, or that the cold conditions in the investigation were perceived as a different level of stress by the two populations. The qRT-PCR data in Chapter 4 is also supportive of local adaptation in *A. l. petraea* as the genes investigated were found to be differentially regulated following chilling between a much larger set of populations from six countries within the north west European range of *A. l. petraea*. The response-to-selection experiment (Chapter 3) suggested that Norwegian populations may differ in their adaptive potential; heritability (h^2) in cold-treated plants correlated negatively with the number of degree days of $<-5^oC$ meaning that those populations experiencing colder conditions had lower heritability for growth rate in cold conditions. The lack of significant difference between heritability values of the populations, however, limits the conclusions that can be drawn from this.

7.3 Climatic change and its influences on both wild and crop plants

It is known that the world's climate is changing, and changes are expected to continue into the foreseeable future (IPCC, 2007). Temperature is not the only climatic variable expected to change with alterations in precipitation, radiation, and wind speed also predicted (Hulme, 2005). In addition to changes in mean values, the extreme values of abiotic variables are likely to change (Easterling *et al.*, 2000), resulting in greater

climatic variability and increased frequencies of extreme events. Accurate predictions as to the magnitude and extent of climatic change are difficult to make, as the climate is such a complex system, incorporating interactions between the atmosphere, biosphere, and the oceans (Slingo *et al.*, 2005). Current knowledge of the climate system is not sufficient for the development of highly accurate climate models and existing models require significant improvements (Hulme, 2005; Slingo *et al*., 2005). Worldwide socio-economic developments are also likely to influence climatic predictions, such as the potential rate of temperature increase (Visser, 2008). There is demand for climate models to be developed at higher resolutions, especially considering the scale gap that exists between climate models and models concerning potential changes in crop production (Slingo *et al*., 2005; Visser, 2008).

There is growing evidence that climatic change has already had biological consequences (Walther *et al.*, 2002; Parmesan and Yohe, 2003; Root *et al.*, 2003; Traill *et al.*, 2010); these include altered species phenology, range shifts, behavioural changes, and local extinctions (Reusch and Wood, 2007). Ecosystem service providers, such as plant pollinators and seed dispersers that provide benefits to human society, will also be affected by temperature changes (Traill *et al*., 2010). In a meta-analysis study of published literature, 62% of 677 species assessed, from a diverse range of taxa, showed phenological trends towards spring advancement (Parmesan and Yohe, 2003). Data from over 30 years of observations of trees and shrubs in Europe has revealed general trends for advancements in spring events such as unfolding of leaves and also delays in autumnal events such as leaf fall (Menzel and Fabian, 1999). Parmesan and Yohe's (2003) meta-analysis also revealed that in those species exhibiting shifting ranges, significantly more of these have shifted in the direction expected in the face of climatic change than expected by chance. Climate change will not only affect individual species but will also impact upon interactions between species (Walther *et al*., 2010; Traill *et al*., 2010). As discussed above, changes in phenology result from climatic change, but not all species will respond in the same way or at the same rate, thus resulting in differences in the timings of ecological interactions (Ings *et al.*, 2009). Range shifts will also influence species interactions as when species extend into new habitats or regions they will encounter different species than in their native range. This may result in the shifting of dominance between species and in the formation of new community structures (Walther *et al*., 2010). As a consequence of this broad range of effects on

species' biology, climate change is a serious threat to global biodiversity and ecosystem dynamics (McLaughlin *et al.*, 2002; Thomas *et al.*, 2004).

In addition to the threat to biodiversity, crop plants of economic importance are also under threat from climate change. The interaction between climatic variation and successful crops has been a topic of continual scientific interest due to its direct impact upon humans (Porter and Semenov, 2005). Increasing temperatures and erratic rainfall will have direct impacts upon crop yields; crops are likely to suffer from drought, be exposed to higher salinity, experience a greater frequency of detrimental high temperature events, and be exposed to pressures from new pests and diseases (Tester and Langridge, 2010). In addition to the pressures of climate change on crop species, there is an ever growing demand for increased food production in the face of an expanding human population. The current target is an average annual increase in production of 44 million metric tons per year (Tester and Langridge, 2010). Other pressures include a need for increased food quality, the desired reduction in agricultural inputs such as nitrogenous fertilisers, and the requirement for crops as biofuel (Tester and Langridge, 2010). It is clear that there is a need for crop improvement to ensure that crop species will be able to combat global environmental changes.

In general, crop plants appear able to adapt to a wide range of environmental conditions as domestication has involved the spread of such species far from their centres of origin to a range of differing climates (Hawtin *et al.*, 1996; Koski, 1996). The process of domestication, however, has resulted in reductions in the genetic diversity of gene pools of cultivars (Hawkes, 1983; Frankel *et al.*, 1995), particularly for species such as peanut and soybean (Tester and Langridge, 2010). Genetic diversity is the basis of evolution by natural selection and is therefore essential for the improvement of a species; without sufficient variation, crop species will be unable to adapt to changing climatic conditions. Studies that increase understanding of the distribution of genetic diversity within a species and provide insight into the extent of local adaptation are therefore of high importance. *A. l. petraea* has a wide distributional range across Europe and persists in distinct populations that experience disparate climatic conditions. It is therefore a suitable model species for studying the distribution of genetic diversity and the extent of local adaptation. In addition to the work presented here, there is growing evidence for local adaptation in *A. l. petraea* (e.g. Leinønen *et al.*, 2009; Vergeer *et al.*, 2008). Populations of *A. l. petraea* that persist in different geographic locations are likely to

harbour variation that may be crucial to the species' own survival in the face of climatic change but that may also be beneficial to the development of crop cultivars with increased stress tolerance.

7.4 Sources of genetic variation for crop improvement

Future improvements in the adaptability of crop species to changed climatic regimes will rely heavily on the introduction of additional genetic variation. Various techniques are available to utilise naturally occurring variation in a plant's ability to respond to temperature, such as that evident in *A. l. petraea*. Novel variation in traits such as the ability to tolerate high or low temperatures can also be created artificially by using techniques that expand the existing variation within a species. Both of these approaches will be discussed below.

7.4.1 Plant breeding techniques

The established method for crop improvement involves traditional plant breeding techniques, and thus this may have the potential to accelerate local adaptation. The first step in plant breeding is the location of genetic variation that corresponds to the desired trait and it is for this step that knowledge of the natural distribution of genetic variation within a species can be beneficial (Hawtin *et al.*, 1996). Wild relatives or progenitors of crop species can be excellent sources of genetic variation as they will have persisted in a wide range of environments and survived under strong pressures from natural selection. Breeders can then target populations that persist in particularly hot or cold areas to locate potentially useful genetic variation in temperature tolerance for example. Crossings can then be performed to introgress the desired characteristic into the crop cultivar. Wild relatives have been successful sources of novel variation for a range of traits in economically important crops including tomato, potato, and rice. A review of recent developments found that in the past twenty years over 100 beneficial traits have been utilised from 60 wild species in the improvement of thirteen crop species (Hajjar and Hodgkin, 2007). The main characteristics sought from wild relatives are pest and disease resistance, and successful examples include resistance to grassy stunt virus in rice (Hajjar and Hodgkin, 2007) and to both leaf and stem rust and powdery mildew in wheat (Hoisington *et al.*, 1999). All resistance of lettuce to downy mildew and lettuce aphid originates from wild relatives (Crute, 1992). Cultivars have been regularly released since the 1980s that confer resistance to downy mildew and without such

resistance it would be very difficult to successfully cultivate lettuce in much of Europe (Hajjar and Hodgkin, 2007). Resistance is regularly overcome in lettuce and consequently breeders must repeatedly locate further genetic variation from wild relatives (Crute, 1992). Another characteristic derived from wild relatives is tolerance to abiotic stress although few examples of this have reached the stage of cultivar release. A chickpea cultivar has been developed with increased tolerance of high temperatures and drought, and improved tolerance of rice to both acidic soils and drought have enabled the spread of rice production to land previously unsuitable for cultivation (Hajjar and Hodgkin, 2007). Differential responses to cold stress have been observed in *A. l. petraea* within this study, and it is studies such as these that may provide the first steps towards identifying further sources of genetic variation in abiotic stress tolerance. Work with model systems such as *A. l. petraea* enables identification of genes and pathways for which homologues can be characterised in crop wild relatives. Similar microarray studies to those presented here could also be performed using other close relatives of *A. thaliana* such as *A. cebennensis* that persist in warmer climates; this may enable identification of further useful variation in traits relevant to response to temperature. Despite the successes in utilising genetic variation that is naturally occurring within wild relatives, there are problems with this technique including hybrid sterility and the retention of undesirable traits associated with the wild relative following introgression of the desired trait (Hawtin *et al.*, 1996; Hajjar and Hodgkin, 2007).

The application of naturally occurring genetic variation in crop wild relatives to the improvement of crop cultivars emphasises the importance of conserving natural biodiversity. Genetic diversity is known to vary between populations and is also partially correlated with ecological factors (Nevo, 1998). In Spanish populations of slender wild oat (*Avena barbata*) genetic variability both within and between populations is higher in regions of favourable climatic conditions, and in less favourable conditions a reduction in genetic diversity is observed (Perez De La Vega, 1996). In the less favourable conditions, genotypes less suited to the environment are eliminated by selection and consequently result in the observed reductions in genetic diversity. Similar patterns have been reported for both wild emmer (*Triticum dicoccoides*) and wild barley (*Hordeum spontaneum*) where adaptive genetic diversity was high in central populations and reduced in ecologically peripheral and isolated populations (Nevo, 1998). Substantial numbers of alleles were unique to particular peripheral populations

yet were locally common within these populations, strongly indicating their importance in adaptation to prevailing marginal conditions. Variation, particularly in the regulation of gene expression, has been demonstrated between geographically distinct populations of *A. l. petraea*, thus providing strong evidence for local adaptation. The populations investigated using microarrays (Chapter 5) were from the centre and the periphery of *A. l. petraea*'s north west European distributional range and exhibited both inherent and cold-responsive differences in gene expression. Such variation is also likely to be present in other plant species and, as genes for many adaptations such as temperature tolerance will be found in extreme environments at the periphery of species ranges, it is not sufficient to focus on just the core of the range for conservation. The long term storage of seeds and tissues, referred to as *ex situ* conservation, can be beneficial in conserving genetic variation but it is important to consider geographical origin of materials when developing core collections to ensure that the widest possible range of variation is preserved (Hawtin *et al.*, 1996). *In situ* conservation is increasing in importance; originally important to only conservationists, it is now of increasing consequence to agronomists (Hawtin *et al.*, 1996). Such conservation can include nature reserves, national parks, and the implementation of specific management regimes. These will be of increasing importance in the face of global climatic change, in particular to prevent the loss of potentially beneficial adaptations harboured within peripheral populations.

Useful genetic diversity can also be obtained from within agricultural species, and this source is likely to be favoured by many breeders. Traditional varieties of crops or landraces that are still used by farmers in preference to modern cultivars are likely to be adapted to local conditions, and particularly to marginal habitats (Hawtin *et al.*, 1996). It is therefore essential that these landraces are conserved, and not put at risk by habitat destruction. As with wild relatives, transfer of desired variation to new cultivars can be problematic, especially if the trait is under complex genetic control. Existing modern cultivars are also used by plant breeders to aid future crop improvements. Benefits of these include their pre-selected ability to produce high yields and their elite genetic background (Hawtin *et al.*, 1996).

Although plant breeding can successfully enhance crop cultivars and increase their adaptability to environments, there are limitations to its use in the development of crops that will be viable following climate change. The development of new cultivars is a

time consuming task, especially for those species with long generation times, such as fruit trees (Koski, 1996). It is also costly, and there is increasing pressure to demonstrate the commercial returns of breeding programmes (Hawtin *et al*., 1996). The unpredictability of the extent of future climatic changes further complicates the instigation of plant breeding programmes as it is not possible to say with certainty to which environmental variables crop plants need to be further adapted. If plant breeding is to have a role in the alleviation of pressures from climate change, a range of traits should therefore be considered (Koski, 1996).

The question that remains unanswered for the vast majority of species is whether or not the existing levels of genetic variation will be sufficient to confer adaptation to the changing global climate. For crop cultivars, it seems highly unlikely that the current levels of genetic variation will be sufficient to cope with predicted rates of change but that breeding programmes, utilising variation present in wild relatives and landraces, may go some way to enhancing adaptation. The time required for breeding, however, puts this technique at a distinct disadvantage, especially considering the rapid developments occurring in gene technology and genomics. The development of DNA markers in the 1980s has been very beneficial to plant breeders, further enhancing the development of crops with improved characteristics. Such molecular markers enable the identification of regions within genomes that are associated with particular quantitative traits, for example drought or salinity tolerance; these regions are commonly referred to as quantitative trait loci (QTLs) (Collard *et al.*, 2005). Molecular markers represent sites of variation in DNA sequence. Use of such markers in plant breeding, by employing them as a means to select plants containing QTLs linked to traits of agronomic importance, is now referred to as molecular breeding (Rafalski and Tingey, 1993). The potential of this technique to improve the efficiency of conventional breeding (Kasha, 1999) has been realised and it has been incorporated into breeding programs for a wide range of crop plant species, including rice, wheat, barley, and maize (Collard *et al*., 2005). Molecular breeding also has limitations due to lack of precise knowledge concerning genes within QTLs. The introgression of QTLs into crop plants can result in transfer of other less desirable characteristics (Bhatnagar-Mathur *et al.*, 2008).

7.4.2 Artificial creation of genetic variation

Genetic mutations occur spontaneously in nature resulting in genetic variation within a population; the rate of natural mutations is, however, too low for these to be of significant use in plant breeding (Parry *et al.*, 2009). Physical and chemical mutagens can be utilised to artificially create novel genetic variation that would not occur naturally within a plant species. This has the potential to extend genetic variation far beyond its current limits and may therefore be of significant relevance in the enhancement of adaptability of crop species to climatic changes. Chemical mutagens often only affect single nucleotide pairs, and the degree of mutation throughout the genome is dependent upon the length of exposure and concentration of the mutagen (Parry *et al*., 2009). A reverse genetics strategy known as TILLING (Targeting Induced Local Lesions IN Genomes) has been developed that utilises chemical mutagenesis. Briefly, chemically mutagenised populations are developed by treating seeds with a mutagen. Plants are then advanced to a generation (generally the third generation) within which any mutations will be stably inherited and background mutations are of minimal concern (Henikoff *et al.*, 2004; Slade and Knauf, 2005). DNA from a number of individuals is then pooled to increase the throughput of a TILLING assay. A single gene target of interest is PCR amplified from the pools of DNA using dye-labelled primers and products are subsequently denatured and re-annealed, thus enabling the formation of heteroduplexes (mismatched base pairs). These heteroduplexes represent both naturally occurring single nucleotide polymorphisms (SNPs) and SNPs induced through mutagenesis. An endonuclease (*Cel*1), sourced from celery (Oleykowski *et al.*, 1998), is then used to cleave mismatched DNA resulting in the identification of novel SNPs and, as cleavage occurs just 3' of the mismatch, the location of the mutation can also be approximated (Slade and Knauf, 2005; Parry *et al*., 2009). The gene of interest must be identified in advance before TILLING can be performed and thus knowledge of genes involved in response to a specific stress or that confer a desired trait is essential.

TILLING was originally developed for use in functional genomics as the ability to detect the location of novel mutations aids subsequent sequencing (Henikoff *et al.*, 2004). Also, as numerous mutations can be detected within a specifically targeted region of the genome, an allelic series can be developed that encompasses a wide range of phenotypes from subtle to strong. This has obvious advantages when investigating the potential functions of a gene, especially as prior knowledge of the gene product is not required (Slade and Knauf, 2005). Since the development of TILLING, mutations

in hundreds of *A. thaliana* genes have been discovered (Till *et al.*, 2003). It has been established that TILLING assays can be successfully applied to crop improvement. Mutagenised populations have been produced for a range of crop species including rice (*Oryza sativa*; (Suzuki *et al.*, 2008)) and barley (*Hordeum vulgare*; (Caldwell *et al.*, 2004)). Many novel alleles have been generated and identified at the *waxy* loci in wheat through TILLING assays; these alleles represent a valuable source of genetic diversity with potential for use in the modulation of starch quality in wheat (Slade *et al.*, 2005). Here, microarray analysis has identified a large number of genes that have potential roles in the adaptation of *A. l. petraea* to local temperature conditions. As there is a growing need to develop crop plants with increased tolerances to temperature stress, such information could be used to identify genes that, through TILLING, could yield much relevant and potentially beneficial variation. The benefits of TIILING assays for increasing available genetic variation can only be realised if scientists know which genes to target in a species. Orthologues of genes identified as important in the response to temperature stress in *A. l. petraea* could be investigated through TILLING in the mutagenised populations that are now available for important crop species. This may consequently reveal variation that enables these crop species to tolerate further extremes of temperature.

This technique, which enables the creation and detection of novel genetic variation within a species, has several advantages over previously discussed approaches to crop improvement. It is a high throughput process, genetic variation can be introduced into an elite germplasm without the additional transfer of agriculturally undesirable traits and, importantly, mutagenesis is not currently regarded as genetic manipulation, thus removing GM issues (Slade and Knauf, 2005; Parry *et al.*, 2009). The development of allelic series will be of particular importance for traits relevant to adaptation to climate change, as the extent of required adaptation remains uncertain due to the unpredictability of future climates. Natural polymorphisms can also be detected through TILLING, hence further widening the application of this technology; this process is referred to as ecoTILLING (Comai *et al.*, 2004). It is possible that *A. l. petraea* may prove a useful source of natural polymorphisms, especially in light of the growing evidence for local adaptation in this species.

7.4.3 Genetic modification of plants

The genetic engineering of plants enables the introduction or over-expression of specific genes; this technology can be used to enable improvement of stress tolerance in crop species and offers a faster alternative to traditional and molecular breeding (Bhatnagar-Mathur *et al.*, 2008). Genetic modification (GM) facilitates the generation of novel genetic variation that exceeds naturally occurring variation (Tester and Langridge, 2010), and is of particular importance if the gene that confers the desired trait originates from a species unable to cross successfully with the crop plant in question (Bhatnagar-Methur *et al*., 2008). Recent negative economic impacts of abiotic stress on crops have resulted in an increased focus upon genes that can alter a plant's ability to withstand stress amongst the biotechnology industry (Kolodyazhnaya *et al.*, 2009). Research into the responses of plants to abiotic stresses is therefore of great importance, especially considering future expected climatic change. Such studies enable the identification of genes that could have significant potential in the development of transgenics with increased stress tolerance. Knowledge of specific genes that have roles in enhancing a species' tolerance of a particular abiotic stress, for example, a gene that confers an ability to tolerate sub-zero temperatures, is crucial to the success of GM technologies. If a gene of large effect is identified in one plant species, GM technologies enable this gene to then be expressed in another species, with the potential of significantly improving its stress tolerance. Whole genome expression studies, such as microarrays, enable genes that are responsive to particular environmental conditions to be readily identified. In this work many genes that are responsive to cold acclimation and to freezing have been identified in *A. l. petraea* through microarray analysis (Chapter 5). In addition to genes of known function such as *TCH4* (At5g57560), a number of genes annotated as unknown function (e.g. At1g74450 and At1g18740) were up-regulated following the freezing shock and may therefore have roles in the responses of *A. l. petraea* to sub-zero temperatures. Identification of these is the first step in their development as genes that may confer increased tolerance to temperature stress and with further investigation some of these may prove valuable in improvement of crop species through GM technologies. Genes with inherent differences in expression between the two populations under investigation may also be beneficial in this context as such differences are indicative of adaptation to local climatic conditions. Genes identified as being expressed more highly in Helin than in Leitrim (e.g. At5g54190 and Atg62510) are potentially of importance in tolerance to more frequent and prolonged

cold temperature exposure, as such conditions are experienced by this population in the field.

Genetic modification of plants originally involved single action genes of known function including osmoprotectants that confer higher stress tolerance, detoxifying genes, and transporter genes that enable plants to re-establish homeostasis (Bhatnagar-Mathur *et al.*, 2008). An increased understanding of the multigenic nature of plant responses to abiotic stresses has resulted in the use of genes coding for regulatory proteins, such as stress inducible transcription factors, in plant transformation. These proteins function in the regulation of many other genes that are involved in the stress response of the plant and consequently can confer a greater extent of stress tolerance (Kasuga *et al.*, 1999). A wide range of transgenic plants have been developed that are tolerant to abiotic stresses. Examples include cold tolerance in papaya (*Carica papaya:* (Dhekney *et al*., 2007)) and *Brassica napus* (Jaglo-Ottosen *et al.*, 1998; Jaglo *et al.*, 2001) conferred by CBF transcription factors; tolerance of ultraviolet B radiation in rice (*Oryza sativa*: (Wang *et al.*, 2007)) conferred by a WRKY transcription factor; and drought tolerance in tobacco (*Nicotiana tabacum*: (Kasuga *et al.*, 2004) conferred by a DREB transcription factor.

GM technology has been rapidly embraced by both agriculture and food biotechnology (Huang *et al.*, 2002; Halpin, 2005) and many crop cultivars have now been released that harbour transgenes. These include cotton (*Gossypium hirsutum*) cultivars with resistance to a range of herbicides and to Lepidopteron insects, maize with resistance to the European corn borer (*Ostrinia nubilalis*), and potato (*Solanum tuberosum*) with resistance to the Colorado potato beetle (*Leptinotarsa decemlineata*) (Singh *et al.*, 2006). There are, however, a number of potential risks linked with GM crops including the presence of putatively allergenic or harmful proteins in food and possible effects of the transgenic plants on non-target organisms (Singh *et al.*, 2006). The major constraints to the widespread use of GM are the considerable financial costs imposed by the high levels of regulation that are required and the lack of consumer acceptance (Tester and Langridge, 2010). These factors may therefore limit, or at least slow the application of, the potential for the use of GM technologies in alleviating impacts of climate change on food crops.

7.5 Summary

Predicted climatic changes include an increase in climatic variability and in the frequency of extreme weather events, and such changes will place pressure on both natural flora and crop plants. The expanding human population intensifies this pressure as significant increases in yield of crops are required over the coming decades. There is therefore a requirement for genetic variation as this is needed if plants are to be able to adapt to changes in conditions. Variation does exist in natural flora and is generally correlated with the geographic distribution of species. This has been demonstrated in geographically distinct populations of *A. l. petraea*, as well as in a wide range of other plant species. Central populations tend to have high levels of genetic diversity whilst marginal populations are increasingly adapted to extreme conditions and thus possess lower levels of genetic diversity. Conservation of this diversity is essential given the threats to plant populations from climatic change; *in situ* and *ex situ* conservation can be implemented to both protect biodiversity and to maintain a genetic resource that can be utilised for the improvement of crops. Techniques to enhance diversity are of particular relevance to crop plant species, for many of which the process of domestication has resulted in depletion of gene pools. Traditional plant breeding can utilise genetic variation that is naturally present in wild relatives of crops and in landraces for the improvement of crop cultivars. Genetic variation can also be artificially created through mutagenesis and applied to the successful improvement of crops through TILLING; this method also requires knowledge of genes within which additional variation could be beneficial. TILLING has the potential to play a pivotal role in the adaptation of crop species to future climatic changes. Here, the study of distinct populations of *A. l. petraea* has enabled the identification of many genes with potential roles in local adaptation, and that are of particular relevance to temperature tolerance. Similar techniques could be applied to the study of other species as a useful method for detecting existing variation. This will be made easier by the increasing range of species for which sequence data are available. Detailed sequence information simplifies the development of technologies such as GeneChips and the design of primers. GM technologies can further widen the availability of genetic resources as genes from non-related species can also be incorporated into crops; for this it is essential to identify genes that confer the desired trait. Orthologues of the genes identified in this work could also be studied in crop species, with a view to enhancing their ability to tolerate a range of temperature conditions.

8. References

Al-Hiyaly, S.A.K., McNeilly, T., Bradshaw, A.D. & Mortimer, A.M. (1993). The effect of zinc contamination from electricity pylons. Genetic constraints on selection for zinc tolerance. *Heredity* **70**(1): 22-32.

Alabadi, D., Oyama, T., Yanovsky, M.J., Harmon, F.G., Mas, P. & Kay, S.A. (2001). Reciprocal regulation between TOC1 and LHY/CCA1 within the *Arabidopsis* circadian clock. *Science* **293**(5531): 880-883.

Angert, A.L., Bradshaw, H.D. & Schemske, D.W. (2008). Using experimental evolution to investigate geographic range limits in monkeyflowers. *Evolution* **62**(10): 2660-2675.

Ansell, S.W., Stenøien, H.K., Grundmann, M., Schneider, H., Hemp, A., Bauer, N., Russell, S.J. & Vogel, J.C. (2010). Population structure and historical biogeography of European *Arabidopsis lyrata*. *Heredity*.

Arnaud-Haond, S., Teixeira, S., Massa, S.I., Billot, C., Saenger, P., Coupland, G., Duarte, C.M. & Serrao, E.A. (2006). Genetic structure at range edge: low diversity and high inbreeding in Southeast Asian mangrove (*Avicennia marina*) populations. *Molecular Ecology* **15**(12): 3515-3525.

Baack, E.J., Emery, N.C. & Stanton, M.L. (2006). Ecological factors limiting the distribution of *Gilia tricolor* in a California grassland mosaic. *Ecology* **87**(11): 2736-2745.

Baer, C.F. and Travis, J. (2000). Direct and correlated responses to artificial selection on acute thermal stress tolerance in a livebearing fish. *Evolution* **54**(1): 238-244.

Baker, S.S., Wilhelm, K.S. & Thomashow, M.F. (1994). The 5'-region of *Arabidopsis thaliana* COR15a has *cis*-acting elements that confer cold-regulated, drought-regulated and ABA-regulated gene-expression. *Plant Molecular Biology* **24**(5): 701-713.

Barrett, J.C. and Kawasaki, E.S. (2003). Microarrays: the use of oligonucleotides and cDNA for the analysis of gene expression. *Drug Discovery Today* **8**(3): 134-141.

Barton, N.H. (2001). Adaptation at the edge of a species range. In *Integrating ecology and evolution in a spatial context.* Silvertown, J. & Antonovics, J. Blackwell Science, New York**:** 365-392.

Baskin, J.M. and Baskin, C.C. (1998). Seeds. Ecology, biogeography, and evolution of dormancy and germination. . San Diego Academic Press, San Diego.

Bazzaz, F.A., Ackerly, D.D. & Reekie, E.G. (2000). Reproductive allocation in plants. In *Seeds: the ecology of regeneration in plant communities*. Fenner, M. CABI Publishing, Wallingford/New York**:** 1-29.

Becher, M., Talke, I.N., Krall, L. & Kramer, U. (2004). Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *Plant Journal* **37**(2): 251-268.

Bell, G. (1997). *The basics of selection*. Chapman & Hall,, New York ; London.

Bender, J. (2002). Plant epigenetics. *Current Biology* **12**(12): R412-R414.

Bhatnagar-Mathur, P., Vadez, V. & Sharma, K. (2008). Transgenic approaches for abiotic stress tolerance in plants: retrospect and prospects. *Plant Cell Reports* **27**(3): 411-424.

Bieniawska, Z., Espinoza, C., Schlereth, A., Sulpice, R., Hincha, D.K. & Hannah, M.A. (2008). Disruption of the *Arabidopsis* circadian clock is responsible for extensive variation in the cold-responsive transcriptome. *Plant Physiology* **147**(1): 263-279.

Bisazza, A., Dadda, M., Facchin, L. & Vigo, F. (2007). Artificial selection on laterality in the teleost fish *Girardinus falcatus*. *Behavioural Brain Research* **178**(1): 29-38.

Black, V.J., Stewart, C.A., Roberts, J.A. & Black, C.R. (2010). Direct effects of ozone on reproductive development in *Plantago major* L. populations differing in sensitivity. *Environmental and Experimental Botany* **69**(2): 121-128.

Blondel, J., Thomas, D.W., Charmantier, A., Perret, P., Bourgault, P. & Lambrechts, M.M. (2006). A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *BioScience* **56**(8): 661-673.

Blows, M.W. and Hoffmann, A.A. (1993). The genetics of central and marginal populations of *Drosophila serrata*. 1. Genetic variation for stress resistance and species borders. *Evolution* **47**(4): 1255-1270.

Bolstad, B.M., Irizarry, R.A., Astrand, M. & Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**(2): 185-193.

Boudry, P., McCombie, H. & Van Dijk, H. (2002). Vernalization requirement of wild beet *Beta vulgaris* ssp. *maritima*: among population variation and its adaptive significance. *Journal of Ecology* **90**(4): 693-703.

Braam, J. and Davis, R.W. (1990). Rain-, wind-, and touch-induced expression of calmodulin and calmodulin-related genes in *Arabidopsis*. *Cell* **60**(3): 357-364.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C.P., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J. & Vingron, M. (2001). Minimum information about a microarray experiment (MIAME) - toward standards for microarray data. *Nature Genetics* **29**(4): 365-371.

Brazma, A., Robinson, A., Cameron, G. & Ashburner, M. (2000). One-stop shop for microarray data - Is a universal, public DNA-microarray database a realistic goal? *Nature* **403**(6771): 699-700.

Brown, J.H., Mehlman, D.W. & Stevens, G.C. (1995). Spatial variation in abundance. *Ecology* **76**(7): 2028-2043.

Brown, J.H., Stevens, G.C. & Kaufman, D.M. (1996). The geographic range: size, shape, boundaries, and internal structure. *Annual Review of Ecology and Systematics* **27**: 597-623.

Brown, P.O. and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nature Genetics* **21**: 33-37.

Burgess, K.S., Etterson, J.R. & Galloway, L.F. (2007). Artificial selection shifts flowering phenology and other correlated traits in an autotetraploid herb. *Heredity* **99**(6): 641-648.

Busch, W. and Lohmann, J.U. (2007). Profiling a plant: expression analysis in Arabidopsis. *Current Opinion in Plant Biology* **10**(2): 136-141.

Bustin, S.A. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**(2): 169-193.

Butlin, R., Bridle, J. & Kawata, M. (2003). Genetics and the boundaries of speices' distributions. In *Macroecology: concepts and consequences. The 43rd Annual Symposium of the British Ecological Society.* Blackburn, T.a.G., KJ. Blackwell Publications.

Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. & Waugh, R. (2004). A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *Plant Journal* **40**(1): 143-150.

Callahan, H.S. (2005). Using artificial selection to understand plastic plant phenotypes. *Integrative And Comparative Biology* **45**(3): 475-485.

Campbell, P. and Braam, J. (1999). Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions. *Trends in Plant Science* **4**(9): 361-366.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W. & Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. *Science* **263**(5148): 802-805.

Channell, R. and Lomolino, M.V. (2000). Dynamic biogeography and conservation of endangered species. *Nature* **403**(6765): 84-86.

Chen, C. and Chen, Z. (2002). Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced *Arabidopsis* transcription factor. *Plant Physiol.* **129**(2): 706-716.

Chen, H.-H., Li, P.H. & Brenner, M.L. (1983). Involvement of abscisic acid in potato cold acclimation. *Plant Physiol.* **71**(2): 362-365.

Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B.H., Hong, X.H., Agarwal, M. & Zhu, J.K. (2003). ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes & Development* **17**(8): 1043-1054.

Chinnusamy, V., Zhu, J. & Zhu, J.K. (2006). Gene regulation during cold acclimation in plants. *Physiologia Plantarum* **126**(1): 52-61.

Choi, D.-W., Rodriguez, E.M. & Close, T.J. (2002). Barley CBF3 gene identification, expression pattern, and map location. *Plant Physiol.* **129**(4): 1781-1787.

Chotikacharoensuk, T., Arteca, R.N. & Arteca, J.M. (2006). Use of differential display for the identification of touch-induced genes from an ethylene-insensitive Arabidopsis mutant and partial characterization of these genes. *Journal of Plant Physiology* **163**(12): 1305-1320.

Chuaqui, R.F., Bonner, R.F., Best, C.J.M., Gillespie, J.W., Flaig, M.J., Hewitt, S.M., Phillips, J.L., Krizman, D.B., Tangrea, M.A., Ahram, M., Linehan, W.M., Knezevic, V. & Emmert-Buck, M.R. (2002). Post-analysis follow-up and validation of microarray experiments. *Nature Genetics* **32**: 509-514.

Clarke, J.D. and Zhu, T. (2006). Microarray analysis of the transcriptome as a stepping stone towards understanding biological systems: practical considerations and perspectives. *Plant Journal* **45**(4): 630-650.

Clauss, M.J. and Koch, M.A. (2006). Poorly known relatives of *Arabidopsis thaliana*. *Trends In Plant Science* **11**(9): 449-459.

Clauss, M.J. and Mitchell-Olds, T. (2006). Population genetic structure of *Arabidopsis lyrata* in Europe. *Molecular Ecology* **15**(10): 2753-2766.

Collard, B., Jahufer, M., Brouwer, J. & Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* **142**(1): 169-196.

Collins, S. and Bell, G. (2004). Phenotypic consequences of 1,000 generations of selection at elevated CO^2 in a green alga. *Nature* **431**(7008): 566-569.

Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R. & Henikoff, S. (2004). Efficient discovery of DNA polymorphisms in natural populations by ecotilling. *The Plant Journal* **37**(5): 778-786.

Cook, D., Fowler, S., Fiehn, O. & Thomashow, M.F. (2004). A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **101**(42): 15243-15248.

Coope, G. (1995). The effects of Quaternary climatic changes on insect populations: lessons from the past. In *Insects in a Changing Environment*. Harrington, R. & Stork, N. Academic, London**:** 29-48.

Craigon, D.J., James, N., Okyere, J., Higgins, J., Jotham, J. & May, S. (2004). NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. *Nucleic Acids Research* **32**: D575-D577.

Crute, I.R. (1992). From breeding to cloning (and back again) - a case-study with lettuce downy mildew. *Annual Review of Phytopathology* **30**: 485-506.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K. & Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiology* **139**(1): 5-17.

Dallas, P.B., Gottardo, N.G., Firth, M.J., Beesley, A.H., Hoffmann, K., Terry, P.A., Freitas, J.R., Boag, J.M., Cummings, A.J. & Kees, U.R. (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR - how well do they correlate? *Bmc Genomics* **6**.

Davey, M.P., Burrell, M.M., Woodward, F.I. & Quick, W.P. (2008). Population-specific metabolic phenotypes of *Arabidopsis lyrata* ssp *petraea*. *New Phytologist* **177**(2): 380-388.

Davey, M.P., Woodward, F.I. & Quick, W.P. (2009). Intraspecfic variation in cold-temperature metabolic phenotypes of *Arabidopsis lyrata* ssp. *petraea*. *Metabolomics* **5**(1): 138-149.

DeGreve, H., Dhaese, P., Seurinck, J., Lemmers, M., Van Montagu, M. & Schell, J. (1982). Nucleotide sequence and transcript map of the *Agrobacterium tumefaciens* ti plasmid encoded octopine synthase gene. *Journal of Molecular and Applied Genetics* **1**(6): 499-512.

Dhekney, S.A., Litz, R.E., Amador, D.A.M. & Yadav, A.K. (2007). Potential for introducing cold tolerance into papaya by transformation with C-repeat binding factor (CBF) genes. *In Vitro Cellular & Developmental Biology-Plant* **43**(3): 195-202.

Ditt, R.F., Kerr, K.F., de Figueiredo, P., Delrow, J., Comai, L. & Nester, E.W. (2006). The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. *Molecular Plant-Microbe Interactions* **19**(6): 665-681.

Dixon, R.A. and Paiva, N.L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* **7**(7): 1085-1097.

Doerge, R.W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nat Rev Genet* **3**(1): 43-52.

Dong, J., Chen, C. & Chen, Z. (2003). Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. *Plant Molecular Biology* **51**(1): 21-37.

Draghici, S., Khatri, P., Eklund, A.C. & Szallasi, Z. (2006). Reliability and reproducibility issues in DNA microarray measurements. *Trends in Genetics* **22**(2): 101-109.

Dudley, J.W. and Lambert, R.J. (1992). 90 generations of selection for oil and protein in maize. *Maydica* **37**(1): 81-87.

Easterling, D.R., Meehl, G.A., Parmesan, C., Changnon, S.A., Karl, T.R. & Mearns, L.O. (2000). Climate extremes: observations, modeling, and impacts. *Science* **289**(5487): 2068-2074.

Eckert, C.G., Samis, K.E. & Lougheed, S.C. (2008). Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology* **17**(5): 1170-1188.

Edwards, K.D., Anderson, P.E., Hall, A., Salathia, N.S., Locke, J.C.W., Lynn, J.R., Straume, M., Smith, J.Q. & Millar, A.J. (2006). FLOWERING LOCUS C mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock. *Plant Cell* **18**(3): 639-650.

Endler, J.A. (1977). *Geographic variation, speciation, and clines*. Princeton University Press, Princeton.

Eulgem, T., Rushton, P.J., Robatzek, S. & Somssich, I.E. (2000). The WRKY superfamily of plant transcription factors. *Trends in Plant Science* **5**(5): 199-206.

Falconer, D. S. and Mackay, T. F. C. (1996). *Introduction to quantitative genetics.* Edn. 4. Longmans Green, Harlow, Essex.

Fenner, M. (1992). Environmental influences on seed size and composition. *Horticultural Reviews* **13**: 183-213.

Fowler, S. and Thomashow, M.F. (2002). Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* **14**(8): 1675-1690.

Frankel, O.H., Brown, A.H.D. & Burdon, J.J. (1995). *The conservation of plant biodiversity*. Cambridge University Press, Cambridge.

Fukushima, A., Kusano, M., Nakamichi, N., Kobayashi, M., Hayashi, N., Sakakibara, H., Mizuno, T. & Saito, K. (2009). Impact of clock-associated *Arabidopsis* pseudo-response regulators in metabolic coordination. *Proceedings of the National Academy of Sciences of the United States of America* **106**(17): 7251-7256.

Fuller, R.C., Baer, C.F. & Travis, J. (2005). How and when selection experiments might actually be useful. *Integrative and Comparative Biology* **45**(3): 391-404.

Gachon, C., Mingam, A. & Charrier, B. (2004). Real-time PCR: what relevance to plant studies? *Journal of Experimental Botany* **55**(402): 1445-1454.

Garcia-Ramos, G. and Kirkpatrick, M. (1997). Genetic models of adaptation and gene flow in peripheral populations. *Evolution* **51**(1): 21-28.

Gaston, K.J. (1996). Species-range-size distributions: patterns, mechanisms and implications. *Trends in Ecology & Evolution* **11**(5): 197-201.

Gaston, K.J. (2009). Geographic range limits: achieving synthesis. *Proc R Soc B* **276**(1661): 1395-1406.

Gibson, U.E.M., Heid, C.A. & Williams, P.M. (1996). A novel method for real time quantitative RT PCR. *Genome Research* **6**(10): 995-1001.

Gilmour, S.J. and Thomashow, M.F. (1991). Cold acclimation and cold-regulated gene expression in ABA mutants of *Arabidopsis thaliana*. *Plant Molecular Biology* **17**(6): 1233-1240.

Glemin, S., Ronfort, J. & Bataillon, T. (2003). Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics* **165**(4): 2193-2212.

Goldsbrough, A., P. , Albrecht, H. & Stratford, R. (1993). Salicylic acid-inducible binding of a tobacco nuclear protein to a 10 bp sequence which is highly conserved amongst stress-inducible genes. *The Plant Journal* **3**(4): 563-571.

Grant-Downton, R.T. and Dickinson, H.G. (2005). Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Annals of Botany* **96**(7): 1143-1164.

Grant-Downton, R.T. and Dickinson, H.G. (2006). Epigenetics and its implications for plant biology 2. The 'epigenetic epiphany': Epigenetics, evolution and beyond. *Annals of Botany* **97**(1): 11-27.

Griffith, T.M. and Watson, M.A. (2006). Is evolution necessary for range expansion? Manipulating reproductive timing of a weedy annual transplanted beyond its range. *American Naturalist* **167**(2): 153-164.

Guiltinan, M.J., Marcotte, W.R. & Quatrano, R.S. (1990). A plant leucine zipper protein that recognizes an abscisic-acid response element. *Science* **250**(4978): 267-271.

Guivarc'h, A., Caissard, J.C., Azmi, A., Elmayan, T., Chriqui, D. & Tepfer, M. (1996). In situ detection of expression of the GUS reporter gene transgenic plants: Ten years of blue genes. *Transgenic Research* **5**(5): 281-288.

Gutierrez, R.A., Ewing, R.M., Cherry, J.M. & Green, P.J. (2002). Identification of unstable transcripts in *Arabidopsis* by cDNA microarray analysis: Rapid decay is associated with a group of touch- and specific clock-controlled genes. *Proceedings of the National Academy of Sciences of the United States of America* **99**(17): 11513-11518.

Guy, C.L., Niemi, K.J. & Brambl, R. (1985). Altered gene-expression during cold-acclimation of spinach. *Proceedings of the National Academy of Sciences of the United States of America* **82**(11): 3673-3677.

Haake, V., Cook, D., Riechmann, J.L., Pineda, O., Thomashow, M.F. & Zhang, J.Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. *Plant Physiol.* **130**(2): 639-648.

Hajjar, R. and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica* **156**(1): 1-13.

Halpin, C. (2005). Gene stacking in transgenic plants - the challenge for 21st century plant biotechnology. *Plant Biotechnology Journal* **3**(2): 141-155.

Hammond, J.P., Bowen, H.C., White, P.J., Mills, V., Pyke, K.A., Baker, A.J.M., Whiting, S.N., May, S.T. & Broadley, M.R. (2006). A comparison of the *Thlaspi caerulescens* and *Thlaspi arvense* shoot transcriptomes. *New Phytologist* **170**(2): 239-260.

Hannah, M.A., Heyer, A.G. & Hincha, D.K. (2005). A global survey of gene regulation during cold acclimation in *Arabidopsis thaliana*. *Plos Genetics* **1**(2): 179-196.

Hannah, M.A., Wiese, D., Freund, S., Fiehn, O., Heyer, A.G. & Hincha, D.K. (2006). Natural genetic variation of freezing tolerance in *Arabidopsis*. *Plant Physiology* **142**(1): 98-112.

Hardtke, C.S., Muller, J. & Berleth, T. (1996). Genetic similarity among *Arabidopsis thaliana* ecotypes estimated by DNA sequence comparison. *Plant Molecular Biology* **32**(5): 915-922.

Hasdai, M., Weiss, B., Levi, A., Samach, A. & Porat, R. (2006). Differential responses of *Arabidopsis* ecotypes to cold, chilling and freezing temperatures. *Annals of Applied Biology* **148**(2): 113-120.

Haseloff, J. and Amos, B. (1995). GFP in plants. *Trends in Genetics* **11**(8): 328-329.

Hatfield, G.W., She-pin, H. & Pierre, B. (2003). Differential analysis of DNA microarray gene expression data. *Molecular Microbiology* **47**(4): 871-877.

Hauser, M.-T., Harr, B. & Schlatterer, C. (2001). Trichome distribution in *Arabidopsis thaliana* and its close relative *Arabidopsis lyrata*: molecular analysis of the candidate gene GLABROUS1. *Molecular Biology and Evolution* **18**(9): 1754-1763.

Hawkes, J.G. (1983). *The diversity of crop plants*. Harvard University Press, Cambridge.

Hawtin, G., Iwanaga, M. & Hodgkin, T. (1996). Genetic resources in breeding for adaptation. *Euphytica* **92**(1-2): 255-266.

He, F., Kang, J.Q., Zhou, X., Su, Z., Qu, L.J. & Gu, H.Y. (2008). Variation at the transcriptional level among Chinese natural populations of *Arabidopsis thaliana* in response to cold stress. *Chinese Science Bulletin* **53**(19): 2989-2999.

Helmer, G., Casadaban, M., Bevan, M., Kayes, L. & Chilton, M.D. (1984). A new chimeric gene as a marker for plant transformation - the expression of *Escherichia coli* beta-galactosidase in sunflower and tobacco cells. *Bio-Technology* **2**(6): 520-527.

Henikoff, S., Till, B.J. & Comai, L. (2004). TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* **135**(2): 630-636.

Hennig, L., Menges, M., Murray, J.A.H. & Gruissem, W. (2003). *Arabidopsis* transcript profiling on Affymetrix GeneChip arrays. *Plant Molecular Biology* **53**(4): 457-465.

Herman, E.M., Rotter, K., Premakumar, R., Elwinger, G., Bae, R., Ehler-King, L., Chen, S. & Livingston, D.P., III (2006). Additional freeze hardiness in wheat acquired by exposure to -3^{o}C is associated with extensive physiological, morphological, and molecular changes. *J. Exp. Bot.* **57**(14): 3601-3618.

Hernández, I., Alegre, L., Van Breusegem, F. & Munné-Bosch, S. (2009). How relevant are flavonoids as antioxidants in plants? *Trends in Plant Science* **14**(3): 125-132.

Heschel, M.Â S., Donohue, K., Hausmann, N. & Schmitt, J. (2002). Population differentiation and natural selection for water-use efficiency in *Impatiens capensis* (Balsaminaceae). *International Journal of Plant Sciences* **163**(6): 907-912.

Higuchi, R., Fockler, C., Dollinger, G. & Watson, R. (1993). Kinetic PCR analysis - real-time monitoring of DNA amplification reactions. *Bio-Technology* **11**(9): 1026-1030.

Hill, W.G. and Caballero, A. (1992). Artificial selection experiments. *Annual Review of Ecology and Systematics* **23**: 287-310.

Hocker, H.W. (1956). Certain aspects of climate as related to the distribution of Loblolly pine. *Ecology* **37**(4): 824-834.

Hoffmann, A. and Parsons, P. (1991). *Evolutionary genetics and environmental stress*. Oxford University Press, Oxford.

Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J.V., Skovmand, B., Taba, S. & Warburton, M. (1999). Plant genetic resources: What can they contribute toward increased crop productivity? *Proceedings of the National Academy of Sciences of the United States of America* **96**(11): 5937-5943.

Holt, R.D. and Keitt, T.H. (2005). Species' borders: a unifying theme in ecology. *Oikos* **108**(1): 3-6.

Hu, X.S. and He, F.L. (2006). Seed and pollen flow in expanding a species' range. *Journal of Theoretical Biology* **240**(4): 662-672.

Hua, J. (2009). From freezing to scorching, transcriptional responses to temperature variations in plants. *Current Opinion in Plant Biology* **12**(5): 568-573.

Huang, J.K., Pray, C. & Rozelle, S. (2002). Enhancing the crops to feed the poor. *Nature* **418**(6898): 678-684.

Hubner, P., Waiblinger, H.U., Pietsch, K. & Brodmann, P. (2001). Validation of PCR methods for quantitation of genetically modified plants in food. *Journal of Aoac International* **84**(6): 1855-1864.

Hughes, M.A. and Dunn, M.A. (1996). The molecular biology of plant acclimation to low temperature. *Journal of Experimental Botany* **47**(296): 291-305.

Hulme, P.E. (2005). Adapting to climate change: is there scope for ecological management in the face of a global threat? *Journal of Applied Ecology* **42**(5): 784-794.

Husak, M.S. and Husak, A.L. (2003). Latitudinal patterns in range sizes of New World woodpeckers. *The Southwestern Naturalist* **48**(1): 61-69.

Iliev, E.A., Xu, W., Polisensky, D.H., Oh, M.-H., Torisky, R.S., Clouse, S.D. & Braam, J. (2002). Transcriptional and post-transcriptional regulation of *Arabidopsis* TCH4 expression by diverse stimuli. Roles of *cis*-regions and brassinosteroids. *Plant Physiol.* **130**(2): 770-783.

Ings, T.C., Montoya, J.M., Bascompte, J., Bluthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009). Ecological networks - beyond food webs. *Journal of Animal Ecology* **78**(1): 253-269.

Innan, H., Terauchi, R. & Miyashita, N.T. (1997). Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. *Genetics* **146**(4): 1441-1452.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U. & Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat* **4**(2): 249-264.

Jablonski, D. (1987). Heritability at the species level: analysis of geographic ranges of cretaceous mollusks. *Science* **238**: 360-363.

Jaffe, M.J. and Biro, R. (1979). Thigmo-morphogenesis: the effect of mechanical perturbation on the growth of plants with special reference to anatomical changes the role of ethylene and interaction with other environmental stresses. In *Stress Physiology in Crop Plants*. Mussell, H. & Staples, R.C. John Wiley and Sons, Inc, New York, USA: Chichester, England P25-60.

Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O. & Thomashow, M.F. (1998). *Arabidopsis* CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science* **280**(5360): 104-106.

Jaglo, K.R., Kleff, S., Amundsen, K.L., Zhang, X., Haake, V., Zhang, J.Z., Deits, T. & Thomashow, M.F. (2001). Components of the *Arabidopsis* C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. *Plant Physiology* **127**(3): 910-917.

Jefferson, R.A. (1989). The GUS reporter gene system. *Nature* **342**(6251): 837-838.

Jefferson, R.A., Kavanagh, T.A. & Bevan, M.W. (1987). GUS fusions - beta-glucuronidase as a sensitive and versatile gene fusion marker in higher-plants. *Embo Journal* **6**(13): 3901-3907.

Jeffery, I., Higgins, D. & Culhane, A. (2006). Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. *BMC Bioinformatics* **7**(1): 359.

Jonsell, B., Kustås, K. & Nordal, I. (1995). Genetic variation in *Arabis petraea*, a disjunct species in northern Europe. *Ecography* **18**(4): 321-332.

Jorgensen, T.H. and Emerson, B.C. (2009). RPW8 and resistance to powdery mildew pathogens in natural populations of *Arabidopsis lyrata*. *New Phytologist* **182**(4): 984-993.

Jump, A.S. and Woodward, F.I. (2003). Seed production and population density decline approaching the range-edge of *Cirsium* species. *New Phytologist* **160**: 349-358.

Kant, P., Gordon, M., Kant, S., Zolla, G., Davydov, O., Heimer, Y.M., Chalifa-Caspi, V., Shaked, R. & Barak, S. (2008). Functional-genomics-based identification of genes that regulate *Arabidopsis* responses to multiple abiotic stresses. *Plant Cell and Environment* **31**(6): 697-714.

Kaplan, F., Kopka, J., Haskell, D.W., Zhao, W., Schiller, K.C., Gatzke, N., Sung, D.Y. & Guy, C.L. (2004). Exploring the temperature-stress metabolome of *Arabidopsis*. *Plant Physiology* **136**(4): 4159-4168.

Kark, S., Alkon, P.U., Safriel, U.N. & Randi, E. (1999). Conservation priorities for chukar partridge in Israel based on genetic diversity across an ecological gradient. *Conservation Biology* **13**(3): 542-552.

Karkkainen, K., Loe, G. & Agren, J. (2004). Population structure in *Arabidopsis lyrata*: Evidence for divergent selection on trichome production. *Evolution* **58**(12): 2831-2836.

Kasha, K.J. (1999). Biotechnology and world food supply. *Genome* **42**(4): 642-645.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. (1999). Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology* **17**(3): 287-291.

Kasuga, M., Miura, S., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2004). A combination of the *Arabidopsis* DREB1A gene and stress-inducible rd29A promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant and Cell Physiology* **45**(3): 346-350.

Kavita, P. and Burma, P.K. (2008). A comparative analysis of green fluorescent protein and beta-glucuronidase protein-encoding genes as a reporter system for studying the temporal expression profiles of promoters. *Journal of Biosciences* **33**(3): 337-343.

Kawecki, T.J. (2008). Adaptation to marginal habitats. *Annual Review of Ecology, Evolution, and Systematics* **39**(1): 321-342.

Kelly, L.C., Rundle, S.D. & Bilton, D.T. (2002). Genetic population structure and dispersal in Atlantic Island caddisflies. *Freshwater Biology* **47**(9): 1642-1650.

Kennedy, G.C. and Wilson, I.W. (2004). Plant functional genomics: opportunities in microarray databases and data mining. *Functional Plant Biology* **31**(4): 295-314.

Kirkpatrick, M. and Barton, N.H. (1997). Evolution of a species' range. *American Naturalist* **150**(1): 1-23.

Kivimaki, M., Karkkainen, K., Gaudeul, M., Loe, G. & Agren, J. (2007). Gene, phenotype and function: GLABROUS1 and resistance to herbivory in natural populations of *Arabidopsis lyrata*. *Molecular Ecology* **16**(2): 453-462.

Kliebenstein, D.J., West, M.A.L., van Leeuwen, H., Kim, K., Doerge, R.W., Michelmore, R.W. & St Clair, D.A. (2006). Genomic survey of gene expression diversity in *Arabidopsis thaliana*. *Genetics* **172**(2): 1179-1189.

Knight, C.A., Vogel, H., Kroymann, J., Shumate, A., Witsenboer, H. & Mitchell-Olds, T. (2006). Expression profiling and local adaptation of *Boechera holboellii* populations for water use efficiency across a naturally occurring water stress gradient. *Molecular Ecology* **15**(5): 1229-1237.

Knight, H. and Knight, M.R. (2001). Abiotic stress signalling pathways: specificity and cross-talk. *Trends in Plant Science* **6**(6): 262-267.

Knight, H., Trewavas, A.J. & Knight, M.R. (1996). Cold calcium signaling in *Arabidopsis* involves two cellular pools and a change in calcium signature after acclimation. *Plant Cell* **8**(3): 489-503.

Knight, M.R., Campbell, A.K., Smith, S.M. & Trewavas, A.J. (1991). Transgenic plant aequorin reports the effects of touch and cold-shock and elicitors on cytoplasmic calcium. *Nature* **352**(6335): 524-526.

Knudsen, S. (2004). *Guide to analysis of DNA microarray data*. John Wiley and Sons, Inc, Hoboken, New Jersey.

Kolodyazhnaya, Y., Kutsokon, N., Levenko, B., Syutikova, O., Rakhmetov, D. & Kochetov, A. (2009). Transgenic plants tolerant to abiotic stresses. *Cytology and Genetics* **43**(2): 132-149.

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

Korn, M., Peterek, S., Mock, H.P., Heyer, A.G. & Hincha, D.K. (2008). Heterosis in the freezing tolerance, and sugar and flavonoid contents of crosses between *Arabidopsis thaliana* accessions of widely varying freezing tolerance. *Plant Cell and Environment* **31**(6): 813-827.

Koski, V. (1996). Breeding plans in case of global warming. *Euphytica* **92**(1-2): 235-239.

Kreil, D.P. and Russell, R.R. (2005). Tutorial section: There is no silver bullet -- a guide to low-level data transforms and normalisation methods for microarray data. *Brief Bioinform* **6**(1): 86-97.

Kreps, J.A., Wu, Y.J., Chang, H.S., Zhu, T., Wang, X. & Harper, J.F. (2002). Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiology* **130**(4): 2129-2141.

Kreuzer, K.A., Lass, U., Landt, O., Nitsche, A., Laser, J., Ellerbrok, H., Pauli, G., Huhn, D. & Schmidt, C.A. (1999). Highly sensitive and specific fluorescence reverse

transcription PCR assay for the pseudogene-free detection of beta-actin transcripts as quantitative reference. *Clinical Chemistry* **45**(2): 297-300.

Kuittinen, H. and Aguade, M. (2000). Nucleotide variation at the CHALCONE ISOMERASE locus in *Arabidopsis thaliana*. *Genetics* **155**(2): 863-872.

Le Corre, V.r., Roux, F. & Reboud, X. (2002). DNA polymorphism at the FRIGIDA gene in *Arabidopsis thaliana*: Extensive nonsynonymous variation Is consistent with local selection for flowering time. *Molecular Biology and Evolution* **19**(8): 1261-1271.

Le, M.Q., Engelsberger, W.R. & Hincha, D.K. (2008). Natural genetic variation in acclimation capacity at sub-zero temperatures after cold acclimation at 4 degrees C in different *Arabidopsis thaliana* accessions. *Cryobiology* **57**(2): 104-112.

Lee, B.H., Henderson, D.A. & Zhu, J.K. (2005). The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1. *Plant Cell* **17**(11): 3155-3175.

Lee, H.J., Xiong, L.M., Gong, Z.Z., Ishitani, M., Stevenson, B. & Zhu, J.K. (2001). The *Arabidopsis* HOS1 gene negatively regulates cold signal transduction and encodes a RING finger protein that displays cold-regulated nucleo-cytoplasmic partitioning. *Genes & Development* **15**(7): 912-924.

Leinønen, P.H., Sandring, S., Quilot, B., Clauss, M.J., Mitchell-Olds, T., Agren, J. & Savolainen, O. (2009). Local adaptation in European populations of *Arabidopsis lyrata* (brassicaceae). *American Journal of Botany* **96**(6): 1129-1137.

Letcher, A.J. and Harvey, P.H. (1994). Variation in geographic range size among mammals of the palearctic *American Naturalist* **144**(1): 30-42.

Levitt, J. (1980). *Responses of plants to environmental stress. Chilling, freezing, and high temperature stresses.* Academic Press, New York.

Li, L.G., He, Z.Y., Pandey, G.K., Tsuchiya, T. & Luan, S. (2002). Functional cloning and characterization of a plant efflux carrier for multidrug and heavy metal detoxification. *Journal of Biological Chemistry* **277**(7): 5360-5368.

Li, W., Wang, R., Li, M., Li, L., Wang, C., Welti, R. & Wang, X. (2008). Differential degradation of extraplastidic and plastidic lipids during freezing and post-freezing recovery in *Arabidopsis thaliana*. *Journal of Biological Chemistry* **283**(1): 461-468.

Li, Y., Fu, Y., Ji, L., Wu, C. & Zheng, C. (2010). Characterization and expression analysis of the *Arabidopsis* mir169 family. *Plant Science* **178**(3): 271-280.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* **10**(8): 1391-1406.

Livak, K.J. and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**(4): 402-408.

Livingston III, D.P., Van, K., Premakumar, R., Tallury, S.P. & Herman, E.M. (2007). Using *Arabidopsis thaliana* as a model to study subzero acclimation in small grains. *Cryobiology* **54**(2): 154-163.

Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Norton, H. & Brown, E.L. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotech* **14**(13): 1675-1680.

Løe, G., Torang, P., Gaudeul, M. & Ägren, J. (2007). Trichome production and spatiotemporal variation in herbivory in the perennial herb *Arabidopsis lyrata*. *Oikos* **116**(1): 134-142.

Lonn, M. and Prentice, H.C. (2002). Gene diversity and demographic turnover in central and peripheral populations of the perennial herb *Gypsophila fastigiata*. *Oikos* **99**(3): 489-498.

Ma, S. and Bohnert, H. (2007). Integration of *Arabidopsis thaliana* stress-related transcript profiles, promoter structures, and cell-specific expression. *Genome Biology* **8**(4): R49.

MacArthur, R. (1972). *Geographical ecology*. Princeton University Press, Princeton, New Jersey.

Mandadi, K.K., Misra, A., Ren, S.X. & McKnight, T.D. (2009). BT2, a BTB protein, mediates multiple responses to nutrients, stresses, and hormones in *Arabidopsis*. *Plant Physiology* **150**(4): 1930-1939.

McKhann, H., Gery, C., Berard, A., Leveque, S., Zuther, E., Hincha, D., De Mita, S., Brunel, D. & Teoule, E. (2008). Natural variation in CBF gene sequence, gene expression and freezing tolerance in the Versailles core collection of *Arabidopsis thaliana*. *BMC Plant Biology* **8**(1): 105.

McLaughlin, J.F., Hellmann, J.J., Boggs, C.L. & Ehrlich, P.R. (2002). Climate change hastens population extinctions. *Proceedings of the National Academy of Sciences of the United States of America* **99**(9): 6070-6074.

Menzel, A. and Fabian, P. (1999). Growing season extended in Europe. *Nature* **397**(6721): 659-659.

Meyers, B.C., Galbraith, D.W., Nelson, T. & Agrawal, V. (2004). Methods for transcriptional profiling in plants. Be fruitful and replicate. *Plant Physiol.* **135**(2): 637-652.

Mineko, K. and Shuichi, Y. (2008). Ethylene signaling in *Arabidopsis* involves feedback regulation via the elaborate control of *EBF2* expression by *EIN3*. *The Plant Journal* **55**(5): 821-831.

Mitchell-Olds, T. (2001). *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology & Evolution* **16**(12): 693-700.

Moose, S.P., Dudley, J.W. & Rocheford, T.R. (2004). Maize selection passes the century mark: a unique resource for 21st century genomics. *Trends in Plant Science* **9**(7): 358-364.

Morishita, T., Kojima, Y., Maruta, T., Nishizawa-Yokoi, A., Yabuta, Y. & Shigeoka, S. (2009). *Arabidopsis* NAC transcription factor, ANAC078, regulates flavonoid biosynthesis under high light. *Plant and Cell Physiology* **50**(12): 2210-2222.

Muller, M.H., Leppälä, J. & Savolainen, O. (2007). Genome-wide effects of postglacial colonization in Arabidopsis lyrata. *Heredity* **100**(1): 47-58.

Mundy, J., Yamaguchishinozaki, K. & Chua, N.H. (1990). Nuclear proteins bind conserved elements in the abscisic acid-responsive promoter of a rice RAB gene. *Proceedings of the National Academy of Sciences of the United States of America* **87**(4): 1406-1410.

Nakano, T., Suzuki, K., Fujimura, T. & Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* **140**(2): 411-432.

Nevo, E. (1998). Genetic diversity in wild cereals: Regional and local studies and their bearing on conservation *ex situ* and *in situ*. *Genetic Resources and Crop Evolution* **45**(4): 355-370.

Nolan, T., Hands, R.E. & Bustin, S.A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature Protocols* **1**(3): 1559-1582.

Novillo, F., Medina, J. & Salinas, J. (2007). *Arabidopsis* CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. *Proceedings of the National Academy of Sciences* **104**(52): 21002-21007.

Oleykowski, C.A., Mullins, C.R.B., Godwin, A.K. & Yeung, A.T. (1998). Mutation detection using a novel plant endonuclease. *Nucleic Acids Research* **26**(20): 4597-4602.

Olsson, K. and Ågren, J. (2002). Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb *Lythrum salicaria*. *Journal of Evolutionary Biology* **15**(6): 983-996.

Örvar, B.L., Sangwan, V., Omann, F. & Dhindsa, R., S. (2000). Early steps in cold sensing by plant cells: the role of actin cytoskeleton and membrane fluidity. *The Plant Journal* **23**(6): 785-794.

Ow, D.W., Wood, K.V., Deluca, M., Dewet, J.R., Helinski, D.R. & Howell, S.H. (1986). Transient and stable expression of the firefly luciferase gene in plant-cells and transgenic plants. *Science* **234**(4778): 856-859.

Owens, C.L., Thomashow, M.F., Hancock, J.F. & Iezzoni, A.F. (2002). CBF1 orthologs in sour cherry and strawberry and the heterologous expression of CBF1 in strawberry. *Journal of the American Society for Horticultural Science* **127**(4): 489-494.

Papdi, C., Joseph, M.P., Salamo, I.P., Vidal, S. & Szabados, L. (2009). Genetic technologies for the identification of plant genes controlling environmental stress responses. *Functional Plant Biology* **36**(8): 696-720.

Parmesan, C. and Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**(6918): 37-42.

Parry, M.A.J., Madgwick, P.J., Bayon, C., Tearall, K., Hernandez-Lopez, A., Baudo, M., Rakszegi, M., Hamada, W., Al-Yassin, A., Ouabbou, H., Labhilili, M. & Phillips, A.L. (2009). Mutation discovery for crop improvement. *J. Exp. Bot.* **60**(10): 2817-2825.

Paschke, M., Abs, C. & Schmid, B. (2002). Relationship between population size, allozyme variation, and plant performance in the narrow endemic *Cochlearia bavarica*. *Conservation Genetics* **3**(2): 131-144.

Pavlidis, P., Li, Q.H. & Noble, W.S. (2003). The effect of replication on gene expression microarray experiments. *Bioinformatics* **19**(13): 1620-1627.

Pearce, R.S. (1999). Molecular analysis of acclimation to cold. *Plant Growth Regulation* **29**: 47-76.

Pearson, S., Davison, A.W., Reiling, K., Ashenden, T. & Ollerenshaw, J.H. (1996). The effects of different ozone exposures on three contrasting populations of *Plantago major*. *New Phytologist* **132**(3): 493-502.

Penfield, S. and Hall, A. (2009). A role for multiple circadian clock genes in the response to signals that break seed dormancy in *Arabidopsis*
Plant Cell **21**(6): 1722-1732.

Perez De La Vega, M. (1996). Plant genetic adaptedness to climatic and edaphic environment. *Euphytica* **92**(1-2): 27-38.

Perry, A.L., Low, P.J., Ellis, J.R. & Reynolds, J.D. (2005). Climate change and distribution shifts in marine fishes. *Science* **308**(5730): 1912-1915.

Pfaffl, M.W. (2004). Quantification strategies in real-time PCR. In *A-Z of quantitative PCR*. Bustin, S.A. International University Line, La Jolla, CA. **5**.

Pitelka, L.F. (1997). Plant migration and climate change. *American Scientist* **85**(5): 464-472.

Polisensky, D.H. and Braam, J. (1996). Cold-shock regulation of the *Arabidopsis* TCH genes and the effects of modulating intracellular calcium levels. *Plant Physiology* **111**(4): 1271-1279.

Porter, J.R. and Semenov, M.A. (2005). Crop responses to climatic variation. *Philosophical Transactions: Biological Sciences* **360**(1463): 2021-2035.

Potvin, C. and Tousignant, D. (1996). Evolutionary consequences of simulated global change: genetic adaptation or adaptive phenotypic plasticity. *Oecologia* **108**(4): 683-693.

Pounds, J.A., Bustamante, M.R., Coloma, L.A., Consuegra, J.A., Fogden, M.P.L., Foster, P.N., La Marca, E., Masters, K.L., Merino-Viteri, A., Puschendorf, R., Ron, S.R., Sanchez-Azofeifa, G.A., Still, C.J. & Young, B.E. (2006). Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature* **439**(7073): 161-167.

Puhakainen, T., Hess, M.W., Makela, P., Svensson, J., Heino, P. & Palva, E.T. (2004). Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Molecular Biology* **54**(5): 743-753.

Quackenbush, J. (2002). Microarray data normalisation and transformation. *Nature Genetics* **32**(496-501).

Rafalski, J.A. and Tingey, S.V. (1993). Genetic diagnostics in plant-breeding - RAPDs, microsatellites and machines. *Trends in Genetics* **9**(8): 275-280.

Ramos-Onsins, S.E., Stranger, B.E., Mitchell-Olds, T. & Aguade, M. (2004). Multilocus analysis of variation and speciation in the closely related species *Arabidopsis halle*r and *A. lyrata*. *Genetics* **166**(1): 373-388.

Rapoport, E.H. (1982). *Areography: geographical strategies of species*. Pergamon Press, Oxford : New York.

Rapp, R.A. and Wendel, J.F. (2005). Epigenetics and plant evolution. *New Phytologist* **168**(1): 81-91.

Redman, J.C., Haas, B.J., Tanimoto, G. & Town, C.D. (2004). Development and evaluation of an *Arabidopsis* whole genome Affymetrix probe array. *Plant Journal* **38**(3): 545-561.

Reiling, K. and Davison, A.W. (1992). Spatial variation in ozone resistance of British populations of *Plantago major* L. *New Phytologist* **122**(4): 699-708.

Ren, S.X., Mandadi, K.K., Boedeker, A.L., Rathore, K.S. & McKnight, T.D. (2007). Regulation of telomerase in *Arabidopsis* by BT2, an apparent target of TELOMERASE ACTIVATOR1. *Plant Cell* **19**(1): 23-31.

Rensink, W.A. and Buell, C.R. (2005). Microarray expression profiling resources for plant genomics. *Trends in Plant Science* **10**(12): 603-609.

Reusch, T.B.H. and Wood, T.E. (2007). Molecular ecology of global change. *Molecular Ecology* **16**: 3973-3992.

Riihimaki, M. and Savolainen, O. (2004). Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). *Am. J. Bot.* **91**(7): 1036-1045.

Ristic, Z. and Ashworth, E.N. (1993). Changes in leaf ultrastructure and carbohydrates in *Arabidopsis thaliana* L. (heyn) cv. columbia during rapid cold acclimation. *Protoplasma* **172**(2-4): 111-123.

Robert, H.S., Quint, A., Brand, D., Vivian-Smith, A. & Offringa, R. (2009). BTB and TAZ domain scaffold proteins perform a crucial function in *Arabidopsis* development. *Plant Journal* **58**(1): 109-121.

Robinson, S. and Parkin, I. (2008). Differential SAGE analysis in *Arabidopsis* uncovers increased transcriptome complexity in response to low temperature. *Bmc Genomics* **9**(1): 434.

Rockett, J.C. and Hellmann, G.M. (2004). Confirming microarray data - is it really necessary? *Genomics* **83**(4): 541-549.

Rohde, K., Heap, M. & Heap, D. (1993). Rapoport's rule does not apply to marine teleosts and cannot explain latitudinal gradients in species richness. *American Naturalist* **142**(1): 1-16.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds, J.A. (2003). Fingerprints of global warming on wild animals and plants. *Nature* **421**(6918): 57-60.

Rose, M., Graves, JL., and Hutchinson, EW. (1990). The use of selection to probe patterns of pleiotrophy in fitness characters. In *Insect Life Cycles; genetics, evolution and co-ordination.* Springer**:** 29-49.

Ross-Ibarra, J., Wright, S.I., Foxe, J.P., Kawabe, A., DeRose-Wilson, L., Gos, G., Charlesworth, D. & Gaut, B.S. (2008). Patterns of polymorphism and demographic history in natural populations of *Arabidopsis lyrata*. *PLoS ONE* **3**(6): e2411.

Roy, K., Jablonski, D. & Valentine, J.W. (1994). Eastern Pacific molluscan provinces and latitudinal diversity gradient - no evidence for Rapoport's rule. *Proceedings of the National Academy of Sciences of the United States of America* **91**(19): 8871-8874.

Sakamoto, H., Maruyama, K., Sakuma, Y., Meshi, T., Iwabuchi, M., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2004). *Arabidopsis* Cys2/His2-Type Zinc-Finger Proteins Function as Transcription Repressors under Drought, Cold, and High-Salinity Stress Conditions. *Plant Physiol.* **136**(1): 2734-2746.

Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2002). DNA-Binding Specificity of the ERF/AP2 Domain of *Arabidopsis* DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. *Biochemical and Biophysical Research Communications* **290**(3): 998-1009.

Sangwan, V., Foulds, I., Singh, J. & Dhindsa, R.S. (2001). Cold-activation of *Brassica napus* BN115 promoter is mediated by structural changes in membranes and cytoskeleton, and requires Ca^{2+} influx. *The Plant Journal* **27**(1): 1-12.

Savitch, L.V., Barker-Astrom, J., Ivanov, A.G., Hurry, V., Oquist, G., Huner, N.P.A. & Gardestrom, P. (2001). Cold acclimation of *Arabidopsis thaliana* results in incomplete recovery of photosynthetic capacity, associated with an increased reduction of the chloroplast stroma. *Planta* **214**(2): 295-303.

Scheiner, S.M. (2002). Selection experiments and the study of phenotypic plasticity. *Journal of Evolutionary Biology* **15**(6): 889-898.

Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. (1995). Quantitative monitoring of gene-expression patterns with a complementary DNA microarray. *Science* **270**(5235): 467-470.

Schenk, P.W. and Snaar-Jagalska, B.E. (1999). Signal perception and transduction: the role of protein kinases. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1449**(1): 1-24.

Schneeberger, C., Speiser, P., Kury, F. & Zeillinger, R. (1995). Quantitative detection of reverse transcriptase-PCR products by means of a novel and sensitive DNA stain. *Pcr-Methods and Applications* **4**(4): 234-238.

Schultz, C.J., Johnson, K.L., Currie, G. & Bacic, A. (2000). The classical arabinogalactan protein gene family of arabidopsis. *Plant Cell* **12**(9): 1751-1767.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y. & Shinozaki, K. (2001). Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell* **13**(1): 61-72.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A., Sakurai, T., Satou, M., Akiyama, K., Taji, T., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y. & Shinozaki, K. (2002). Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant Journal* **31**(3): 279-292.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000). Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinion in Plant Biology* **3**(3): 217-223.

Siemens, D.H., Garner, S.H., Mitchell-Olds, T. & Callaway, R.M. (2002). Cost of defense in the context of plant competition: *Brassica rapa* may grow and defend. *Ecology* **83**(2): 505-517.

Singh, K.B., Foley, R.C. & Oñate-Sánchez, L. (2002). Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology* **5**(5): 430-436.

Singh, O.V., Ghai, S., Paul, D. & Jain, R.K. (2006). Genetically modified crops: success, safety assessment, and public concern. *Applied Microbiology and Biotechnology* **71**(5): 598-607.

Siomos, M.F. (2009). Shaped by the environment - adaptation in plants. *Febs Journal* **276**(17): 4705-4714.

Skinner, D. (2009). Post-acclimation transcriptome adjustment is a major factor in freezing tolerance of winter wheat. *Functional & Integrative Genomics* **9**(4): 513-523.

Slade, A.J., Fuerstenberg, S.I., Loeffler, D., Steine, M.N. & Facciotti, D. (2005). A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nature Biotechnology* **23**(1): 75-81.

Slade, A.J. and Knauf, V.C. (2005). TILLING moves beyond functional genomics into crop improvement. *Transgenic Research* **14**(2): 109-115.

Slingo, J.M., Challinor, A.J., Hoskins, B.J. & Wheeler, T.R. (2005). Introduction: Food crops in a changing climate. *Philosophical Transactions: Biological Sciences* **360**(1463): 1983-1989.

Smallwood, M. and Bowles, D.J. (2001). Plants in a cold climate. Discussion Meeting on Coping with the Cold, London, England.

Smith, C.J. and Gallon, J.R. (2001). Living in the real world: how plants perceive their environment. *New Phytologist* **151**(1): 1-6.

Smith, T.B., Wayne, R.K., Girman, D.J. & Bruford, M.W. (1997). A role for ecotones in generating rainforest biodiversity. *Science* **276**(5320): 1855-1857.

Solbrig, O.T. and Solbrig, D.J. (1984). Size inequalitities and fitness in plant populations. *Oxford Survey in Evoltionary Biology* **1**: 141-159.

Stanton, M.L., Roy, B.A. & Thiede, D.A. (2000). Evolution in stressful environments. I. phenotypic variability, phenotypic selection, and response to selection in five distinct environmental stresses. *Evolution* **54**(1): 93-111.

Stenøien, H.K., Fenster, C.B., Kuittinen, H. & Savolainen, O. (2002). Quantifying latitudinal clines to light responses in natural populations of Arabidopsis thaliana (Brassicaceae). *American Journal of Botany* **89**(10): 1604-1608.

Stevens, G.C. (1989). The latitudinal gradient in geographic range - how so many species coexist in the tropics. *American Naturalist* **133**(2): 240-256.

Stockinger, E.J., Gilmour, S.J. & Thomashow, M.F. (1997). *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proceedings of the National Academy of Sciences of the United States of America* **94**(3): 1035-1040.

Stone, J.M. and Walker, J.C. (1995). Plant protein kinase families and signal transduction. *Plant Physiol.* **108**(2): 451-457.

Strand, Å., Hurry, V., Gustafsson, P. & Gardeström, P. (1997). Development of *Arabidopsis thaliana* leaves at low temperatures releases the suppression of photosynthesis and photosynthetic gene expression despite the accumulation of soluble carbohydrates. *The Plant Journal* **12**(3): 605-614.

Stratagene (2007) Introduction to quantitative PCR.

Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y. & Kurata, N. (2008). MNU-induced mutant pools and high performance

TILLING enable finding of any gene mutation in rice. *Molecular Genetics and Genomics* **279**(3): 213-223.

Swindell, W.R. (2006). The association among gene expression responses to nine abiotic stress treatments in *Arabidopsis thaliana*. *Genetics* **174**(4): 1811-1824.

Tähtiharju, S., Sangwan, V., Monroy, A.F., Dhindsa, R.S. & Borg, M. (1997). The induction of kin genes in cold-acclimating *Arabidopsis thaliana*. Evidence of a role for calcium. *Planta* **203**(4): 442-447.

Tester, M. and Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science* **327**(5967): 818-822.

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A. & Heinen, E. (1999). Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology* **75**(2-3): 291-295.

Thomas, C.D. (2010). Climate, climate change and range boundaries. *Diversity and Distributions* **16**(3): 488-495.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Townsend Peterson, A., Phillips, O.L. & Williams, S.E. (2004). Extinction risk from climate change. *Nature* **427**(6970): 145-148.

Thomas, C.D., Franco, A.M.A. & Hill, J.K. (2006). Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution* **21**(8): 415-416.

Thomas, C.D. and Lennon, J.J. (1999). Birds extend their ranges northwards. *Nature* **399**(6733): 213-213.

Thomashow, M.F. (1999). Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology* **50**: 571-599.

Thomashow, M.F. (2001). So what's new in the field of plant cold acclimation? Lots! *Plant Physiology* **125**(1): 89-93.

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., Henikoff, J.G., Comai, L. & Henikoff, S. (2003). Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Research* **13**(3): 524-530.

Townsend-Peterson, A., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H. & Stockwell, D.R.B. (2002). Future projections for Mexican faunas under global climate change scenarios. *Nature* **416**(6881): 626-629.

Traill, L.W., Lim, M.L.M., Sodhi, N.S. & Bradshaw, C.J.A. (2010). Mechanisms driving change: altered species interactions and ecosystem function through global warming. *Journal of Animal Ecology* **79**(5): 937-947.

Truong, C., Palme, A.E. & Felber, F. (2007). Recent invasion of the mountain birch *Betula pubescens* ssp *tortuosa* above the treeline due to climate change: genetic and ecological study in northern Sweden. *Journal of Evolutionary Biology* **20**(1): 369-380.

Trzcinska-Danielewicz, J., Bilska, A., Fronk, J., Zielenkiewicz, P., Jarochowska, E., Roszczyk, M., Jonczyk, M., Axentowicz, E., Skoneczny, M. & Sowinski, P. (2009). Global analysis of gene expression in maize leaves treated with low temperature: I. Moderate chilling (14 °C). *Plant Science* **177**(6): 648-658.

Tsay, Y.F., Chiu, C.C., Tsai, C.B., Ho, C.H. & Hsu, P.K. (2007). Nitrate transporters and peptide transporters. *Febs Letters* **581**(12): 2290-2300.

Turner, T.L., Bourne, E.C., Von Wettberg, E.J., Hu, T.T. & Nuzhdin, S.V. (2010). Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics* **42**(3): 260-U42.

Van De Mortel, J.E. and Aarts, M.G.M. (2006). Comparative transcriptomics – model species lead the way. *New Phytologist* **170**(2): 199-201.

Vergeer, P., Van Den Berg, L.L.J., Bulling, M.T., Ashmore, M.R. & Kunin, W.E. (2008). Geographical variation in the response to nitrogen deposition in *Arabidopsis lyrata petraea*. *New Phytologist* **179**(1): 129-141.

Visser, M.E. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proc R Soc B* **275**(1635): 649-659.

Vogel, J.T., Zarka, D.G., Van Buskirk, H.A., Fowler, S.G. & Thomashow, M.F. (2005). Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. *Plant Journal* **41**(2): 195-211.

Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.M., Hoegh-Guldberg, O. & Bairlein, F. (2002). Ecological responses to recent climate change. *Nature* **416**(6879): 389-395.

Wang, H.H., Hao, J.J., Chen, X.J., Hao, Z.N., Wang, X., Lou, Y.G., Peng, Y.L. & Guo, Z.J. (2007). Overexpression of rice WRKY89 enhances ultraviolet B tolerance and disease resistance in rice plants. *Plant Molecular Biology* **65**(6): 799-815.

Wang, N.N., Shih, M.-C. & Li, N. (2005). The GUS reporter-aided analysis of the promoter activities of *Arabidopsis* ACC synthase genes AtACS4, AtACS5, and AtACS7 induced by hormones and stresses. *Journal of Experimental Botany* **56**: 909-920.

Wanner, L.A. and Junttila, O. (1999). Cold-induced freezing tolerance in *Arabidopsis*. *Plant Physiology* **120**(2): 391-399.

Weber, M., Harada, E., Vess, C., von Roepenack-Lahaye, E. & Clemens, S. (2004). Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *Plant Journal* **37**(2): 269-281.

Wesselingh, R.A. and Klinkhamer, P.G.L. (1996). Threshold size for vernalization in *Senecio jacobaea*: Genetic variation and response to artificial selection. *Functional Ecology* **10**(2): 281-288.

Whitlock, M. (2004). Selection and drift in metapopulations. In *Ecology, genetics, and evolution of metapopulations*. Hanski, I. & Gaggiotti, O. Elsevier, Amsterdam : London**:** pp. 153–73.

Wilson, R.J., Gutierrez, D., Gutierrez, J., Martinez, D., Agudo, R. & Monserrat, V.J. (2005). Changes to the elevational limits and extent of species ranges associated with climate change *Ecology Letters* **8**(11): 1138-1146.

Winkel-Shirley, B. (2002). Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology* **5**(3): 218-223.

Wong, C.E., Li, Y., Whitty, B.R., Diaz-Camino, C., Akhter, S.R., Brandle, J.E., Golding, G.B., Weretilnyk, E.A., Moffatt, B.A. & Griffith, M. (2005). Expressed sequence tags from the Yukon ecotype of *Thellungiella* reveal that gene expression in response to cold, drought and salinity shows little overlap. *Plant Molecular Biology* **58**(4): 561-574.

Wu, Z. and Irizarry, R.A. (2004). Preprocessing of oligonucleotide array data. *Nat Biotech* **22**(6): 656-658.

Wurmbach, E., Yuen, T. & Sealfon, S.C. (2003). Focused microarray analysis. *Methods* **31**(4): 306-316.

Xin, Z. and Browse, J. (2000). Cold comfort farm: the acclimation of plants to freezing temperatures. *Plant Cell and Environment* **23**(9): 893-902.

Xin, Z., Mandaokar, A., Chen, J., Last, R.L. & Browse, J. (2007). Arabidopsis ESK1 encodes a novel regulator of freezing tolerance. *The Plant Journal* **49**(5): 786-799.

Xu, W., Purugganan, M.M., Polisensky, D.H., Antosiewicz, D.M., Fry, S.C. & Braam, J. (1995). *Arabidopsis* TCH4, regulated by hormones and the environment, encodes a xyloglucan endotransglycosylase. *Plant Cell* **7**(10): 1555-1567.

Xue, G.-P. (2003). The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature. *The Plant Journal* **33**(2): 373-383.

Yakir, E., Hilman, D., Hassidim, M. & Green, R.M. (2007). CIRCADIAN CLOCK ASSOCIATED1 transcript stability and the entrainment of the circadian clock in *Arabidopsis*. *Plant Physiology* **145**(3): 925-932.

Yamada, K., Osakabe, Y., Mizoi, J., Nakashima, K., Fujita, Y., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2009). Functional analysis of an *Arabidopsis thaliana* abiotic stress-inducible facilitated diffusion transporter for monosaccharides. *Journal of Biological Chemistry* **285**(2): 1138-1146.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (1993). Characterization of the expression of a desiccation-responsive RD29 gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. *Molecular & General Genetics* **236**(2-3): 331-340.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (1994). A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell* **6**(2): 251-264.

Yamaguchi-Shinozaki, K. and Shinozaki, K. (2005). Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science* **10**(2): 88-94.

Yanagisawa, S. (2002). The Dof family of plant transcription factors. *Trends in Plant Science* **7**(12): 555-560.

Yang, Z., Tian, L., Latoszek-Green, M., Brown, D. & Wu, K. (2005). *Arabidopsis* ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. *Plant Molecular Biology* **58**(4): 585-596.

Ye, S. and Day, I.N.M. (2003). *Microarrays & microplates: applications in biomedical sciences*. BIOS, Oxford.

Yuen, T., Wurmbach, E., Pfeffer, R.L., Ebersole, B.J. & Sealfon, S.C. (2002). Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucl. Acids Res.* **30**(10): e48.

Zarka, D.G., Vogel, J.T., Cook, D. & Thomashow, M.F. (2003). Cold induction of *Arabidopsis* CBF genes involves multiple ICE (Inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. *Plant Physiology* **133**(2): 910-918.

Zhang, C.Z., Fei, S.Z., Warnke, S., Li, L.J. & Hannapel, D. (2009). Identification of genes associated with cold acclimation in perennial ryegrass. *Journal of Plant Physiology* **166**(13): 1436-1445.

Zhang, W., Shmulevich, I. & Astola, J. (2004). *Microarray quality control*. John Wiley and Sons Inc., Hoboken, New Jersey.

Zhen, Y. and Ungerer, M.C. (2008a). Clinal variation in freezing tolerance among natural accessions of *Arabidopsis thaliana*. *New Phytologist* **177**(2): 419-427.

Zhen, Y. and Ungerer, M.C. (2008b). Relaxed selection on the CBF/DREB1 regulatory genes and reduced freezing tolerance in the southern range of *Arabidopsis thaliana*. *Molecular Biology and Evolution* **25**(12): 2547-2555.

Zhu, T. and Wang, X. (2000). Large-scale profiling of the *Arabidopsis* transcriptome. *Plant Physiol.* **124**(4): 1472-1476.