Jonathan Lamb

St Hugh's College

University of Oxford

Uncovering the genetic basis of natural variation of leaf form in *Cardamine hirsuta*

Hilary term, 2015

Thesis submitted for the degree of Doctor of Philosophy

Jonathan Lamb **St Hugh's College**

Uncovering the genetic basis of natural variation of leaf form in *Cardamine hirsuta*

Thesis submitted for the degree of Doctor of Philosophy, Hilary term 2015

Abstract

A major goal in biology is to understand the genetic basis of morphological variation at different evolutionary scales, for example between and within species. Here I investigate this issue by using plant leaves as an example. Previously comparative studies between the simple leaf model plant *Arabidopsis thaliana* and its dissected leaf relative *Cardamine hirsuta* have shown that inter-specific differences in leaf shape mostly result from variation in local tissue growth and patterning (Vlad et al., 2014; Hay et al., 2006; Barkoulas et al., 2008). Here, I aim to elucidate the genetic basis of natural variation in leaf form within species, by using divergent strains of *C. hirsuta*. I present evidence that variation in six strains collected from geographically diverse locations results from different rates of progression of an age-dependent leaf development programme in a phenomenon known as heteroblasty. By using Quantitative trait loci (QTL) mapping with a recombinant inbred line (RIL) population derived from a cross between the Oxford and Azores strains, I detected six QTL that influence leaflet production on multiple leaves. A QTL located on the 4th linkage group was validated and selected for further analysis. Characterisation of QTL effect indicated that the QTL influences leaf form by altering the rate of heteroblastic development. Subsequently I fine mapped this QTL to a DNA segment of 48 kb containing the gene *SQUAMOSA PROMOTER PROTEIN BINDING LIKE 9* (*ChSPL9*), a previously characterised regulator of age dependent development. The parental alleles of *ChSPL9* show variation in their sequence and were transformed into *A. thaliana* to evaluate whether they contribute to the QTL effect. Resultant phenotypes mirrored the QTL effect suggesting that *ChSPL9* does indeed contribute to this QTL effect. These results indicate that age-dependent leaf shape progression underlies variation in leaflet number within species and more broadly suggest that in the case of plant leaves different processes might underlie morphological variation between and within species.

Acknowledgements

Finished at last. I am exhausted and relieved in equal measure – writing up has been more of a challenge than I ever imagined it could be and submitting this thesis is a testament to the many who encouraged me to persevere when I lost sight of the end.

Firstly, I must thank my supervisor Professor Miltos Tsiantis, without whom this thesis would not exist and my life in Oxford would never have been. Thank you also to Maria Cartolano and Bjorn Pieper for their not inconsiderable guidance, patience and generosity in the lab.

Technical help from Mohsen Hajheidari, Michiel Kwantes and Dana Vlad helped me get over many hurdles, and special thanks to Rita Pabst for guiding me through the plant transformations. In addition, Janne Lempe was invaluable in helping me continue my experiments in Cologne. The members of the Tsiantis group in Oxford and Cologne all helped me at one stage or another, and I am lucky to have been a part of the team.

I owe a particular debt of gratitude to Professor Lee Sweetlove, who was the voice of reason when I needed it most.

As always, I am immensely grateful to my parents for everything. Throughout my childhood, my parents encouraged and nurtured my love of science, in the same way that they encouraged and supported me whilst I wrote this thesis. They have both gone out of their way to help with my DPhil in every way possible, and I can only hope they realise how much I appreciate their love and kindness.

Yn olaf, hoffem ddiolch i'm gwraig, Angharad, sydd siwr o fod wedi dioddef hyd yn oed yn fwy na fi yn ystod cyfnod ysgrifennu's traethawd yma (sydd yn gweud rhywbeth). Bydde ni ar goll hebddo ti yn fy mywyd. Diolch yn fawr cariad.

Mae'r traethawd yma, fel popeth arall, i Angharad.

List of abbreviations

A. thaliana	*Arabidopsis thaliana*
amiRMA	artificial micro RNA
ARF	Auxin Response Factor
bHLH	basic Helix Loop Helix
BP	Brevipedicellus
bp	base pairs
BROMO	DNA-binding bromodomain containing protein
C. hirsuta	*Cardamine hirsuta*
CAPS	Cleaved Amplified Polymorphic Sequence
C-CAP	Cofactor C-like domain containing protein
cDNA	Complementary DNA
cM	centi Morgan
Col-0	Columbia
CUC2	CUP SHAPED COTYLEDON 2
DOPS	Department of Plant Sciences, University of Oxford, UK
dCAPS	derived Cleaved Amplified Polymorphic Sequence
FBH4	FLOWERING BASIC HELIX-LOOP-HELIX 4
HD-ZIPIII	Class III Homeodomain Leucine Zipper
HIF	Heterogeneous Inbred Family
IL	Introgression Line
kb	kilo base
KNOX	class I KNOTTED1-like homeobox
LA	LANCEOLATE
Ler	Landsberg *erecta*
LBD	Lateral Organ Boundary Domain
LG	Linkage Group
LRR	Leucine Rich Repeat
miRNA	micro RNA
MPIPZ	Max Planck Institute for Plant Breeding Research in Cologne, Germany
NIL	Near Isogenic Line
NZ	New Zealand
PCR	Polymerase Chain Reaction
PIN1	PIN-FORMED1
qRT-PCR	quantitative real time Polymerase Chain Reaction
QTL	Quantitative Trait Loci/Locus
RIL	Recombinant Inbred Line
RRM	RNA recognition motif protein
SEM	Standard Error of the Mean
SNP	Single Nucleotide Polymorphisms
SPL9	SQUAMOSA PROMOTER PROTEIN BINDING LIKE 9
ta-siRNAs	Trans-acting small interfering RNAs
Wash	Washington state

Table of Contents

Chapter 1:
Introduction

The leaves of plants exist in an enormous range of forms and thus present a promising system in which to address a key question in biological research: how does diversity arise? In this thesis plant leaves are used to investigate morphological variation within species in an attempt to tackle the problem of whether phenotypic diversity within species and between species (i.e. at two different evolutionary scales) in the same trait has arisen through equivalent or divergent morphogenetic avenues. Previously, comparative developmental studies have shown that the divergent leaf forms displayed by the closely related species *Arabidopsis thaliana* and *Cardamine hirsuta* have arisen mostly as a result of variation in local tissue growth and patterning (Vlad et al., 2014; Hay et al., 2006; Barkoulas et al., 2008). In this thesis I investigate the genetic basis of variation in leaf form between strains of *C. hirsuta* to better understand the processes governing intraspecific morphological variation.

1. 1. Leaf development

Leaves are the principal lateral appendages of the stems of vascular plants. They are typically thin, dorsiventrally flattened organs specialised for light capture to facilitate photosynthesis. They are derived from the shoot apical meristem, a population of pluripotent cells located at the tip of the shoot axis. As growth proceeds, leaves are formed at the flanks of the shoot apical meristem from founder cells that lose their pluripotency and start to become specialised leaf cells. Directional cell division drives outward growth from the edges of the shoot apical meristem to create the initial leaf primordia. Soon after leaf primordia have become established, leaf cell differentiation begins to promote the different layers of specialised cell types that allow important physiological functions to begin. An important cell differentiation event is that which gives rise to the specification of distinct adaxial (upper surface) and abaxial lower surface, once this axis is established the development of the flat lamina with distinct and specialised surfaces can proceed. Beyond this initial cell division and specialisation to establish the basic leaf shape, cell division slows down and leaf growth is driven predominantly by cell

expansion to give rise to the final size and shape (Donnely et al., 1999). There is no defined temporal separation of the separate cell proliferation, differentiation and expansion phases in leaf development and there is, to a certain extent, overlap (Reddy et al. 2004). The coordination of these phases at different stages and areas of the leaf is responsible for determining final leaf morphology.

1. 1. 1. Leaf initiation from the flanks of the shoot apical meristem

Leaves develop from the flanks of the apical meristem as the shoot grows upwards. The very first specification of shoot apical meristem cells as leaf initials is mediated by a mechanism involving down regulation of *KNOX* (class I KNOTTED1-like homeobox) genes whose proteins normally act to maintain meristem identity at the shoot apical meristem. *KNOX* expression is repressed by ARP MYB transcription factors (ASYMMETRIC LEAVES1 (AS1), ROUGH SHEATH2 (RS2) and PHANTASTICA (PHAN)), the LATERAL ORGAN BOUNDARY DOMAIN (LBD) protein AS2 and the chromatin remodelling factor HIRA to promote specification of the organ initiation domain (Hay and Tsiantis, 2010; Barkoulas et al., 2007). Chromatin remodelling is involved in silencing *KNOX* expression in leaf initials; AS1 interacts with the histone deacetylase HAD6 (Luo et al 2012). Additionally a complex created by AS1 and AS2 recruits POLYCOMB-REPRESSIVE COMPLEX 2 (PRC2) which alters chromatin structure around the promoters of two *KNOX* genes ensuring their repression is maintained through multiple rounds of cell division as the leaf primordium develops (Lodha et al., 2013).

A second pathway in leaf initial delimitation involves activity maxima of the indolic plant hormone auxin which work alongside AS1 to promote leaf development, an action that is partly dependent on *BREVIPEDICELLUS* (*BP*) repression (Hay et al., 2006). The auxin activity maxima are set up by auxin efflux facilitator PIN-FORMED1 (PIN1) proteins which direct auxin flux firstly towards the summit of the meristem and secondly down the shoot resulting in the establishment of auxin maxima flanking the shoot apical meristem. A positive feedback loop

ensues in which auxin gradients and or flow direct PIN1 polarization towards areas of leaf initiation. The response to auxin is then mediated by auxin response factors.

The KNOX and the auxin pathways are highly integrated; expression domains of *PIN1* and *KNOX* genes lie in complementary patterns (Heisler et al., 2005), expression domain of *BP* is affected when auxin signalling or PIN1 activity is altered and ectopic expression of *KNOX* genes alters auxin transport activity. Models have been produced that suggest KNOX and auxin activities sit within a feedback loop that produces and reinforces a defined delimitation between the meristem and leaf initials (Barkoulas et al. 2007). Other early markers of leaf initiation domains include the transcription factors from the AINTEGUMENTA (ANT) like (AIL)/PLT family and genes of YABBY (YAB) family of High-Mobility-Group-like proteins (Horstman et al., 2014; Sarojam et al., 2010).

Leaf initial positioning is also partly determined by signals and tissue properties created by mechanistic forces (reviewed in Robinson et al., 2013). Cell walls of different areas of the meristem have been found to have different levels of stiffness (Milani et al., 2011) and variable mechanical properties of cells correlates with increased growth (Kierzkowski et al., 2012). Again auxin has a role in driving these mechanical forces required for leaf initiation. Organ initiation requires that the cell walls are loosened; auxin induces the cell wall modifying enzymes that mediate this (Braybrook and Peaucelle, 2013). Furthermore, mechanical forces and the cell wall have an influence in the polarization of PIN1 in cells (Feraru et al., 2011; Nakayama et al., 2012). Thus organ initiation is directed by variable mechanistic properties of different areas of the shoot apical meristem. This in turn influences auxin distribution leading to changes in cell properties and auxin interactions with additional plant hormones and other leaf development pathways (Bar and Ori, 2014).

1. 1. 2. Initiation and growth of the lamina

To attain the typical flat shape of a leaf the growth of a lamina is initiated during primary morphogenesis. The process is promoted by YAB and AIL/PLT genes in Arabidopsis (Mizukami and Fischer, 2000; Sarojam et al., 2010) and leaf blade growth is redundantly promoted by JAGGED (JAG) and NUBBIN (Dinneny et al., 2006; Ohno et al., 2004). JAG also has a role in promoting differentiation by repressing meristematic and cell cycle genes (Scheissel et al., 2014) suggesting that lamina initiation and expansion require concurrent repression of meristematic properties. Further research is required to see how the different regulators of lamina initiation and growth are coordinated. Lamina growth requires communication between the epidermis and mesophyll layers and this is achieved via the transcriptional co-activator ANGUSTIFOLIA3 (AN3). AN3 is produced in the mesophyll cells but diffuses into the epidermis to promote growth in both layers (Kawade et al., 2013) by interacting with chromatin remodelling factors (Vercruyssen et al., 2014).

Other influences of leaf laminar growth include genes involved in basic cellular functions. These include ribosomal proteins (Horiguchi et al., 2012), the E3 ubiquitin-ligase BIG BROTHER (Disch et al., 2006), poly(A) polymerases (Vi et al., 2013) and the ribosome associated protein SIMPLE LEAF3 (Kougioumoutzi et al., 2013). However, it is unclear whether these have specific roles in regulating laminar development or if the phenotypes observed in their mutants simply reflect a general effect on growth.

1. 1. 3. Axial polarity and patterning

The process that must follow for further development is the establishment of abaxial and adaxial leaf surface identities. This cell specification is essential for leaf function because of the divergent roles of the adaxial (light capture) and abaxial (gas exchange) surfaces.

The adaxial identity is promoted by *Class III Homeodomain Leucine Zipper (HD-ZIP III)* genes such as *PHABULOSA (PHB).* These are repressed in the abaxial side by the microRNAs 165 and 166 (Mallory et al., 2004) and also by the KANADI (KAN) family of proteins (Emery et al., 2003). Abaxial identity is promoted by *YABBY* and *KAN* genes and these are negatively regulated by AS1 and AS2 in adaxial cells. Chromatin modification is also involved; the histone deacetylases HDT1/HD2A and HDT2/HD2B act in concurrence with AS1 and AS2 to regulate the micro RNAs 165 and 166 which repress *HD-ZIP III* expression. Auxin activity also has a role in establishing axial patterning; KAN action in the abaxial side of the leaf is facilitated by auxin response factors (ARFs). In turn these ARFs are down regulated by trans-acting small interfering RNAs (ta-siRNAs).

1. 1. 4. Differentiation versus morphogenesis

Following the initial cell division and specialisation within the leaf primordia, further expansion to the final shape and size is achieved by cell expansion rather than proliferation (Donnelly et al., 1999). Leaf shape modifications are contingent on the organogenic potential held in the marginal blastozones (Hagemann and Gleissberg, 1996). Depending on the species these are regions located either at the tip or the base of a leaf (Tsukaya, 2014) in which transient indeterminate growth is maintained. Initially marginal blastozones are responsible for lamina initiation but they also maintain organogenic potential so that marginal structures such as leaflets can develop. Compound leaf development is contingent on prolonged activity of the marginal blastozones during primary morphogenesis and further elongation of this activity correlates with increased leaf complexity (Hagemann and Gleissberg, 1996). In accordance with the evolutionary origin of the leaf as a modified shoot; maintenance of marginal blastozone activity and maintenance of shoot apical meristem activity have multiple genetic and hormonal regulators in common (Brand et al., 2007; Floyd and Bowman, 2010).

Marginal complexity is determined by a window of morphogenetic activity within the marginal blastozones and the spatial and temporal activity of the window is determined by counteracting factors that either delay or promote differentiation. Differentiation spreads across the leaf gradually from the tip to the base in a moving cell cycle arrest front (Donnelly et al., 1999). The progression of the arrest front is variable; in *Arabidopsis* it goes through two sharp transitions (Andriankaja et al., 2012). The wave is in part initiated by TCP transcription factors (TEOSINTE BRANCHED1/CYCLOIDEA/PCF) which negatively regulate growth. Mutations in these genes result in aberrant leaf shape, curvature, serrated margins and wrinkles (Palatnik et al., 2003; Nath et al., 2003). TCPs also act to restrict the expression domains of locally acting margin outgrowth factors (Koyama et al., 2007) demonstrating the link between whole organ cell-cycle arrest and margin outgrowth.

CIN-TCP transcription factors are key promoters of differentiation. In *Antirrhinum,* CIN promotes tissue differentiation and arrest of lamina growth inducing the onset of secondary morphogenesis (Nath et al., 2003). In *Arabidopsis* sequential TCP activities drives transition from primary morphogenesis into cell expansion and secondary morphogenesis (Efroni et al., 2008). These factors are regulated by miR319 and when the target site of this miR319 is altered in the CIN-TCP member LANCEOLATE (LA) in tomato, *LA* expression is precocious and simple leaves rather than compound leaves are produced (Ori et al., 2007). Overexpression of *miR139* results in down regulation of CIN-TCP genes and delayed onset of leaf maturation and an extended phase of indeterminate growth at the leaf margin (Ori et al., 2007; Shleizer-Burko et al., 2011). These results suggest that the spatio-temporal properties of the morphogenetic window are dependent on TCP activity in the early stages of leaf growth.

In contrast to CIN-TCP transcription factors, KNOX proteins play an important role in maintenance of the transient morphogenetic window during early leaf development in some species in addition to its role in the maintenance of shoot apical meristem properties. *KNOX* over expression in simple leaved species leads to knot like structures for the leaves, curled or lobed leaves, and ectopic meristems on the leaves (Hay and Tsiantis, 2010). In maize, mis-expression of

KNOX results in proximal tissues being placed in more distal locations indicating that KNOX has a role in establishment of leaf polarity (Bolduc et al., 2012). In species with compound leaves, *KNOX* expression is not excluded from developing leaf primordia (Bharathan et al., 2002). This expression drives formation of complex leaves; down regulation of *KNOX* in leaf primordia results in faster maturation and reduced complexity (Hay and Tsiantis, 2006; Shani et al., 2009). Further details of KNOX expression in compound leaves are discussed later in this chapter.

Plant hormones also affect the balance between morphogenesis and differentiation and this effect is often mediated by interactions with the transcription factors considered above. A good example is gibberellin which influences rates of cell proliferation and expansion in *Arabidopsis* (Achard et al., 2009) and promotes simpler and faster maturing leaves in tomato (Bassel et al., 2008; Fleishon et al., 2011; Hay et al., 2002; Jasinski et al., 2008). There is interplay between KNOX and TCP proteins with gibberellin processes. KNOX reduces gibberellin levels by repressing the gibberellin biosynthesis gene *GA20ox* and activation of the gibberellin inactivation gene *GA2ox* (Bolduc and Hake, 2009; Hay et al., 2002; Jasinski et al., 2005; Sakamoto et al., 2001). In tomato, the TCP protein LA regulates gibberellin levels (Yanai et al., 2011). Here we have a clear mechanism by which transcription factors regulate the rate of cell differentiation and marginal elaboration in leaves.

Another common plant hormone besides auxin and gibberellin is cytokinin which influences the transition between morphogenesis and differentiation. Elevated cytokinin degradation in leaf primordia accelerates cell expansion and terminates cell proliferation in *Arabidopsis* (Holst et al., 2011; Werner et al., 2001) meaning that cytokinin is a positive regulator of morphogenesis. In tomato it has been demonstrated that cytokinin maintains morphogenetic activity in the leaf margin (Shani et al., 2010). Like gibberellin and auxin, cytokinin also interacts with the transcription factors KNOX and CIN-TCPs. It acts downstream of KNOX to delay differentiation and CIN-TCPs promote maturation partly by reducing leaf sensitivity to cytokinin (Efroni et al., 2013). Thus, cytokinin represents a key regulator of the balance between differentiation and morphogenesis.

Further elaborations in leaf growth are underpinned by complex space and time dependent inter tissue layer relationships. For example, signals that derive from epidermal cell layers feed into the brassino steroid signalling pathways that regulate cell expansion in other underlying tissue layers (Savaldi-Goldstein et al. 2007).

1. 1. 5. Marginal patterning

Aside from size and overall shape, a major constituent of leaf form variability is that of marginal patterning which involves the production of serrations, lobes and leaflets along the leaf margin during both primary and secondary morphogenesis. These structures are generated by local promotion and restriction of growth in adjacent regions of the leaf (Kawamura et al., 2010; Malinowski et al., 2011; Vlad et al., 2014).

A key interaction in the development of marginal patterning is that which occurs between auxin and *NO APICAL MERISTEM* (*NAM*)/*CUP SHAPED COTYLEDON* (*CUC*) transcription factors. In simple leaves NAM/CUC promote production of serrations (Bilsborough et al., 2011; Hasson et al., 2011; Nikovics et al., 2006), and regulate leaflet separations and specification in compound leaves (Brand et al., 2007). The boundaries between the leaf margin and new leaflets are marked by the expression of *NAM/CUC* in an array of compound leaf bearing species, and silencing expression results in simplified leaves (Berger et al., 2009; Blein et al., 2008; Cheng et al., 2012). *NAM/CUC* expression is negatively regulated by miR164 and when ectopic expression is initiated in transgenic experiments, simple deeply lobed leaves where leaflets have fused are produced (Berger at al., 2009) suggesting that distinct domains of expression are required to create a compound leaf with separated leaflets.

The production of marginal structures such as serrations, leaflets and lobes is another part of leaf development that is dependent on auxin (Bilsborough et al., 2011; Hay et al., 2006). The mechanism by which it operates has similarities to the generation of leaf initials at flanks of the shoot apical meristem; both leaf initiation and lateral leaf organ development require discrete

auxin maxima to promote initials and growth. As in the generation of leaf initials, PIN1 subcellular localisation creates auxin concentration maxima at sites of future lateral leaf structures and the external application of auxin leads to ectopic production of serrations/lobes and/or lamina growth (Barkoulas et al., 2008; Ben-Gera et al., 2012; DeMason and Polowick, 2009; Koenig et al., 2009). Computational modelling was used to propose that CUC2 promotes PIN1 localisation and then in a feedback loop, auxin represses *CUC2* expression to create patterned serration along the leaf margin (Bilsborough et al., 2011). Auxin maxima also drive the formation of major veins (Scarpella et al. 2006).

Additional control over marginal patterning comes in the form of tasiRNAs; mutations in the tomato tasiRNA pathway, which regulate ARF 2, 3 and 4 negatively, produce leaves that are simpler and narrower (Yifhar et al., 2012).

Multiple studies show that common genes can affect both leaf maturation and marginal patterning. In tomato the JAG ortholog, LYRATE (LYR) promotes organ growth at the leaf margin in a similar way to that which JAG promotes the growth of the leaf lamina in *Arabidopsis* (David-Schwartz et al., 2009). CUC genes, *AS1* and auxin responsive genes have been identified as targets of CIN-TCP regulation in *Arabidopsis* (Koyama et al., 2007). The combination of CIN-TCPs down regulation of CUC and STIMPY/WOX9 up regulation results in *Arabidopsis* leaves with greatly increased margin complexity and a leaf shape which is close to that of a compound leaf (Blein et al., 2013).

In summary, leaf complexity by marginal patterning is generated from regional differences in promotion and inhibition of laminal growth. Interaction between plant hormones, transcription factors and growth regulators are responsible for the positioning of these regions and fine tuning the interactions between these creates an indefinite number of leaf margin structures.

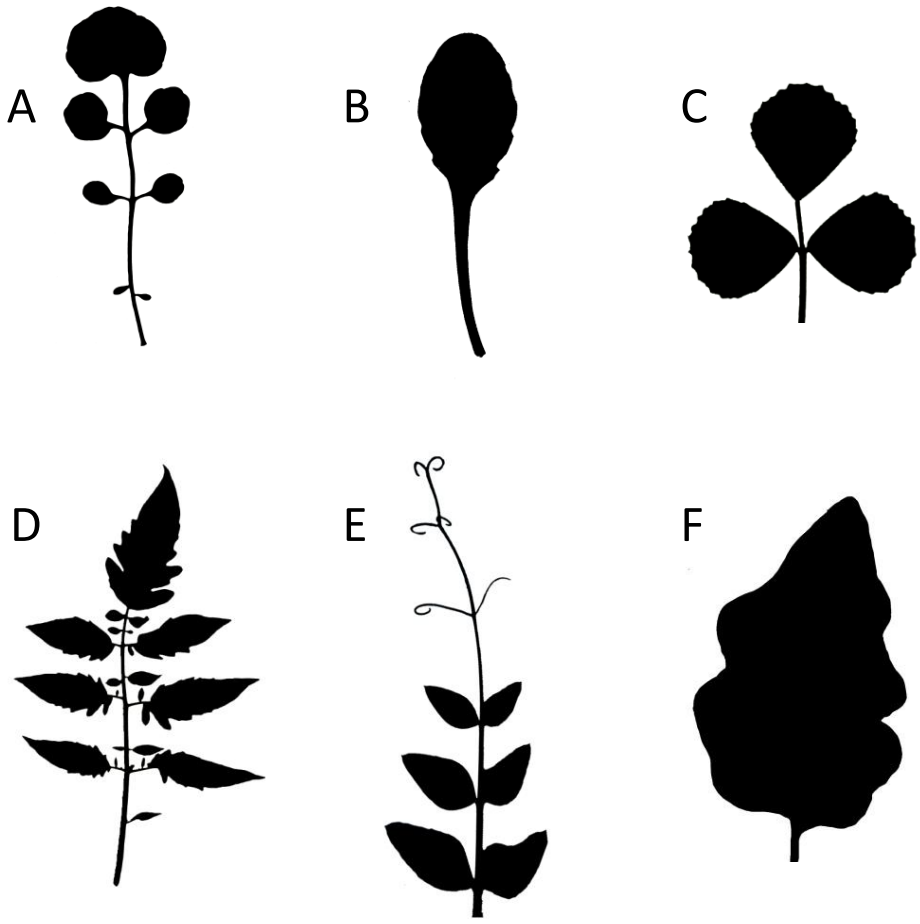
1. 1. 6. Summary

At the broad level leaf development can be seen as a sequence of developmental programs that are performed by different combinations of interacting factors. Each stage is comprised of factors that are unique to that phase but there are also factors that operate across multiple phases and contribute to different pathway or act as the convergence point between different pathways. The artificial division of leaf development into initiation, morphogenetic balance and marginal patterning can be misleading because there are many regulators that act in multiple stages. These regulators and the pathways in which they operate are numerous and their interactions are complex and dependent on spatial and temporal context, which means there is a great deal of potential for modification to generate diversity. Identification of these modifications amongst different species has revealed how leaf shape diversity has evolved.

1. 2. Diversity between species

Of all plant features, leaf shape represents one of the most morphologically diverse characters (Figure 1.1). The diversity is striking and is often used as a marker for species identification. In the following section I summarise some of the identified diversity in leaf development programmes that has produced some of this variation. Much of the discussion will centre on the development of the compound leaf which has a lamina dissected into subunits termed leaflets, each of which resembles a simple leaf. A simple leaf has an entire and continuous lamina. Developmentally, simple leaves differentiate relatively fast and compound leaves have delayed differentiation and, to a certain extent, have intermediate forms somewhere between simple leaves and lateral branches.

Figure 1.1 Morphological diversity of leaves Compound leaf of *Cardamine hirsuta* (A), simple leaf of *Arabidopsis thaliana* with marginal serrations (B), trifoliate compound leaf of *Medicago truncatula* (C), compound leaf of Tomato (*Solanum lycopersicum*) with primary and secondary dissection (D), compound leaf of Pea (*Pisum sativum*) with tendrils (E) and the simple lobed leaf of Aubergine (*Solanum melongena*) (F). Images are not to scale.


A
B
C
D
E
F

The extensive diversity in plant leaf forms is underpinned by corresponding extensive variability in leaf ontogeny which is created by adjustments to partially common factors in leaf development pathways. Combinatorial and spatio-temporal variability allows for an almost infinite number of final leaf shapes and sizes. New interactions between hormones and transcription factors that generate diversity are discovered regularly.

The study of the genetic pathways that underpin leaf development in different model organisms has helped reveal the evolutionary re-configurements that have resulted in diversification of leaf morphology; there are multiple ways in which biochemically equivalent processes have diverged substantially in different taxa with major developmental consequences. From early on in the development of leaf primordia the intricate patterning pathways that work to establish axial polarity have substantial species-specific differences. An example is the biogenesis of tasiRNAs in maize by *LEAFBLADELESS1* (*LBL1*); mutations in this gene create defects in axial patterning but mutations in the *Arabidopsis* orthologue *SUPPRESSOR OF GENE SILENCING3* (*SGS3*) do not (Nogueria et al., 2007; Peragine et al., 2004).

Model species which have compound leaves include *C. hirsuta*, tomato, pea and *M. trunculata* (Figure 1.1) and research into their leaf development programmes have revealed interesting cases where regulatory genes that serve important functions in other developmental areas (in particular the shoot apical meristem) have been shown to operate in leaf development to promote the generation of laminar outgrowths. Many of the model compound leafed species employ many, but not all, of the same developmental pathways to produce a dissected leaf and the resulting variation between these species results in reconfigurement of these pathways which often act to alter the length of the morphogenetic window. Reconfigurements to the developmental pathways are spatially and temporally context dependent and the known alterations together cannot account for the immense variation between compound leaves found in nature. A deeper understanding of specific 'tools' and the fine tuning of their activity in different species will be invaluable in the elucidation of leaf development as a whole.

1. 2. 1. Species diversity in control of differentiation and morphogenesis

As discussed above, production of a non-simple leaf shape requires the maintenance of transient morphogentic activity at the leaf margin in regions called the marginal blastozones (Hagemann and Gleissberg, 1996). There is a temporal window of morphogenetic activity at the blastozones and the extent of this window correlates with increased leaf complexity and is defined by the antagonistic activities of differentiation promoting and differentiation delaying factors. This antagonism is common amongst species with complex leaves but the nature and control of the factors involved has diversified considerably.

KNOX, a delayer of tissue differentiation, is not expressed in simple leaves with an uninterrupted complete leaf margin but *KNOX* expression has been observed in complex dissected leaves that possess leaflets (Hareven et al., 1996; Bharathan et al., 2002). In tomato down regulation of KNOX results in accelerated leaf differentiation and reduced leaf complexity (Shani et al., 2009) and up regulation of KNOX genes prolongs the morphogenetic window and delays differentiation producing leaves that have increased complexity (Chen et al., 1997; Hareven et al., 1996). Regulators of KNOX expression include the CLAUSA (CLAU) and TRIPINNATE (TP) transcription factors (Hay and Tsiantis, 2010; Avivi et al., 2000; Jasinski at al., 2007). Additionally, the BEL LIKE HOMEODOMAIN (BELL) protein BIPINNATE (BIP) interacts directly with KNOX in tomato and increased dissection in the *bip* mutant suggests BIP represses KNOX activity (Smith et al., 2002; Kimura et al., 2008). When KNOX proteins are released from the interaction with BELL, as in the *PETROSELINUM* (*PTS)* gain of function mutation, leaves increase in complexity (Kimura et al., 2008). *PTS* and its negative regulator *BLADE ON PETIOLE* (*BOP*) have been shown to influence the morphogenetic window in tomato and their expression among tomato and related wild species correlated with diversity of leaf complexity (Ichihashi et al., 2014).

The importance of KNOX in promoting compound leaf development is highlighted by contrasting the compound leaved *C. hirsuta* to its simple leaved relative *Arabidopsis*. KNOX is

present in leaves of *C. hirsuta* but not in the leaves of the latter (Hay and Tsiantis, 2006). Knock down of the *C. hirsuta KNOX* gene *SHOOT MERISTEMLESS (STM)* leads to leaves which are less complex, have larger cells and have less cell division. In accordance with this, ectopic KNOX expression leads to increased complexity. The divergent expression resulted from differences in the *cis* regulatory elements and there is evidence that the simple leaves of *Arabidopsis* derived from complex structures through the loss of *KNOX* (*STM*) expression in leaves (Piazza et al., 2010).

In summation; KNOX activity in developing leaf primordia serves to delay differentiation and promote leaf complexity but its effect on leaf morphogenesis is dependent on the timing and location of expression rather than actual expression levels (demonstrated by silencing and overexpression experiments, Shani et al., 2009).

KNOX is affected by ARP (ASYMETRIC LEAVES1 (AS1), PHANTASTICA, CRISPA) transcription factors in some species (Hay and Tsiantis, 2010; Harrison et al., 2005) but their exact effect on KNOX expression and leaf development is highly variable between species. Reducing ARP/AS1 in tomato and *C. hirsuta* affects leaflet number and shape (Kim et al., 2003), but the exact mechanism of interaction with KNOX is unclear. The pea AS1 ortholog, CRISPA, has been reported to suppress KNOX and UNI expression with *crispa* mutants having various leaf abnormalities (Demason et al., 2014; Tattersal et al., 2005). Mutations in *MtPHAN* also produce abnormal leaves in *Medicago* (Ge et al., 2014) and in Medicago the UNI ortholog SINGLE LEAFLET1 (SGL) is uncoupled from ARP regulation (Zhou et al., 2013). It is clear that control of compound leaf development and KNOX expression by ARPs is highly variable and has diverged significantly in different species.

KNOX gene expression in *C. hirsuta* is linked to proteins that have a role in ribosome recycling and translation termination that are conserved in eukaryotes and Archaea. The ribosome-associated protein SIMPLE LEAF3 is required for leaf growth and leaflet development,

it is required for *KNOX* gene expression and localised auxin concentration maxima in leaves (Kougioumoutzi et al., 2013).

However, not all compound leaves rely on the expression of KNOX to promote morphogenetic growth. KNOX proteins are not wholly essential in the elaboration of pea and *Medicago* leaves (Di Giacomo et al., 2008). In pea *UNIFOLIATA* (*UNI*) and in *Medicago SGL1* promote the transient intermediate growth of the leaf and these genes are orthologs to the *Arabidopsis* floral meristem identity gene *LEAFY* (*LFY*) (Wang et al., 2008; Gourlay et al., 2000). *UNI* is up regulated during pea leaf initiation and repressed as leaves mature, *uni* mutant leaves have reduced complexity (Gourlay et al., 2000; Wang et al., 2008) and extended *UNI* expression results in increased marginal blastozone activity (Champagne et al., 2007; DeMason et al., 2013; DeMason and Chetty, 2011; Gourlay et al., 2000). This demonstrates that alterations to different developmental pathways produce the same phenotypic effect.

Transcription factors that promote differentiation include those from the CIN-TCP (CINNCINATA TEOSINTE BRANCHED1-CYCLOIDEA PCF) family (Martin-Trilo and Cubas, 2010). The tomato LANCEOLATE (LA) gene, a CIN-TCP gene, is negatively regulated by miR139. Overexpression of *miR139* leads to extended morphogenesis at the leaf margin because differentiation is delayed and loss of miR139 repression results in small simplified leaves because differentiation is promoted (Ori et al., 2007; Shleizer-Burko et al., 2011). Maintenance of marginal blastozone activity requires low CIN-TCP activity in the early stages of leaf development. LA's effect comes about because it represses APETALA1/FRUITFULL (AP1/FUL) MADS box genes. Less complex leaves are produced when AP1/FUL activity is reduced (Burko et al., 2013) suggesting AP1/FUL promote morphogenetic activity at the leaf margin. The role of CIN-TCP transcription factors in *C. hirsuta* is yet to be elucidated.

Leaf maturation is also regulated by the plant hormones and key examples include auxin, gibberellin and cytokinin. Differences exist amongst species in how they regulate and respond to differing levels of these hormones. In tomato, increased gibberellin levels result in

faster maturing simpler leaves, suggesting that gibberellin is a promoter of leaf maturation (Bassel et al., 2008; Fleishon et al., 2011; Hay et al., 2002; Jasinski et al., 2008). This response is bought about by interactions with TCP proteins (Yanai et al. 2011). In contrast, gibberellin in pea elongates the window of leaf morphogenesis and promotes leaf complexity (Goliber et al., 1999). Similarly auxin response is also altered in pea; auxin in *Cardamine* and tomato promotes leaf simplification but in pea it promotes indeterminate growth by regulating UNI (Demason et al., 2013). Cytokinin contributes to the maintenance of morphogenetic activity in the tomato leaf margin (Shani et al., 2010) and it acts in opposition to gibberellin (Fleishon et al., 2011). In pea, cytokinin promotes shoot growth and helps promote the transport of auxin (Demason et al., 2005). The way in which plant hormones regulate leaf development is highly variable and is a very flexible regulatory system in which changes to relative hormone levels can produce changes to leaf shape and size.

Other factors regulating morphogenetic potential in tomato include TRIFOLIATE (TF), a relative of the *Arabidopsis* MYB transcription factor LATERAL ORGAN FUSION1. Loss of function *tf* mutants have faster growth and earlier differentiation of leaf primordia resulting in simplified leaves lacking marginal elaborations (Naz et al., 2013) suggesting TF is required for the maintenance of morphogenetic potential. Leaf maturation is also influenced by SINGLE FLOWER TRUSS (SFT), and SELF-PRUNING (SP) which are better known for their roles in flowering induction. A high SFT/SP ratio leads to earlier maturation and a simpler leaf form. This effect is also influenced by TF and miR139 (Burko et al., 2013; Shalit et al., 2009). In pea AFILA (AF) and TENDRIL LESS (TL) depress expression of UNI meaning they are promoters of differentiation (Demason et al., 2013; Mishra et al., 2009). In *Medicago* this role is fulfilled by the *PALM* gene which promotes differentiation by repressing *SGL* genes (Chen et al., 2010).

In summation, deviations from the simple leaf shape are facilitated by the characteristics of the marginal blastozones which define a window of morphogenetic activity. These characteristics are influenced by the balance in activity of antagonistically acting

transcription factors that either promote or repress differentiation. The nature and fine tuning of these factors differs among species resulting in extensive leaf shape diversity.

1. 2. 2. Species diversity in control of marginal patterning

Marginal structures such as leaflets, serrations and lobes are a very striking and easily quantifiable source of interspecies variation in leaf shape. Such marginal patterning is created by differential growth in adjacent regions (Malinowski et al., 2011; Vlad et al., 2014). Many of the processes that produce leaflets have much in common with the processes that underlie the formation of serration and lobes. The prevalent marginal patterning mechanism is mediated by the interaction between auxin and the NO APICAL MERISTEM (NAM)/CUP-SHAPED COTYLEDON (CUC) transcription factors, where growth is promoted by auxin and repressed by NAM/CUC proteins. This mechanism of leaflet separation is broadly conserved in tomato (Berger et al., 2009), *Cardamine*, pea (Blein et al., 2008) and *Medicago* (Cheng et al., 2012; Zhou at el, 2013). However in these species there are key modifications which have resulted in phenotypic diversity.

Bilsborough et al., (2011) showed that CUC2, PIN1 and auxin operate in a feedback loop that promotes the production of margins serrations on the leaves of *Arabidopsis*; CUC2 promotes PIN1 localisation and auxin represses CUC2 expression. In *Arabidopsis, NAM/CUC* genes are also regulated by auxin in the shoot apical meristem (Aida and Tasaka, 2006). However in tomato, auxin affects *GOB* (*NAM/CUC* ortholog) in the shoot apices but not in leaf primordia and auxin appears to act downstream of GOB in leaf development (Ben-Gera et al., 2012) suggesting that NAM/CUC and auxin interactions are important in both simple and compound leaves but the relationship is reconfigured to pattern diverse leaf forms. Upstream in this pathway, TCP transcription factors shape the expression domains of *CUC* genes (Koyama et al., 2007) resulting in control over marginal outgrowths and they have been implicated in the

formation of leaflets. In the dissected leaf of tomato, mutations that prevent miRNA regulation of the *TCP* gene *LANCEOLATE* impedes the formation of leaflets (Ori et al., 2007).

As discussed above, auxin maxima are required to establish the formation of marginal outgrowths such as serrations and leaflets, however in *M. truncatula*, mutants of the *MtPIN10/SLM1*, PIN1 ortholog, have increased complexity and decreased marginal patterning. This suggests a complex relationship between auxin activity and leaf shape variation. However, Peng and Chen (2011) and Zhou et al., (2011) suggest that the increased complexity may be a result of fusion of multiple leaflets.

Returning to tomato; tasiRNAs are important in promoting leaf complexity by negatively regulating auxin response factors (Yifhar et al., 2012). However, when the same pathway is compromised in *M. trunculata* there is no effect on the number of leaflets, just increased leaf lobes (Zhou et al., 2013). Disruption of tasiRNA regulation has no effect on leaf development (Hunter et al., 2006). This is an excellent example of how some mechanisms for the control of leaf shape are conserved among species and others have diverged substantially.

An additional controller of marginal patterning in tomato is the ortholog of JAG; LYRATE (LYR). LYR promotes organ growth at the leaf margin, in a role analogous to how JAG promotes growth of the main leaf lamina in *Arabidopsis.* The leaves of *lyr* mutants have more leaflets; *LYR* overexpression leads to leaflet fusion and it is expected the LYR affects auxin response or distribution (David- Schwartz et al., 2009).

The *Potato leaf* (C) gene is an orthologue of the *Arabidopsis* branching regulator *REGULATOR OF AXILLARY MERISTEMS1* (*RAX1*) which also regulates tomato leaf complexity. Mutants of this gene generate leaves with less complexity and less uneven leaf margins than the wild type and experiments with double mutations show that the gene acts partially redundantly with *NAM/CUC* genes (Busch et al., 2011). Like KNOX, this is another instance in which a regulator of shoot growth has been lost or gained in leaf development and has resulted in the generation of diversity.

In *C. hirsuta,* leaflet development requires the input of the REDUCED COMPLEXITY (RCO) homeodomain protein (Vlad et al., 2014). RCO operates by inhibiting growth between leaflets which it does independently of auxin distribution. It has evolved by duplication in the *Brassicaceae* family, but has been lost in *Arabidopsis,* leading to leaf simplification. It was found that *cis* regulatory diversification of RCO paralogs account for variation in leaf shape amongst species of the genus *Capsella* (Sicard et al., 2014). In *C. hirsuta,* differential expression was also the reason for diversified function in leaf complexity. The RCO ortholog in pea is TL which specifies tendril fate (Hofer et al., 2009). When TL is absent tendrils are converted into leaflets and loss of RCO in *C. hirsuta* results in loss of leaflets therefore RCO/TL is clearly important in the generation of lateral organs in compound leaves.

The array of variation visible in the marginal structures of leaves is reflected by the diversity in the regulatory systems that govern marginal patterning. Different species have modified, employed or lost different regulatory elements that dictate the distribution of regions in which growth is promoted or repressed and the result is an array of different leaf shapes.

1. 2. 3. Summary

In the development of leaves, similar mechanisms have been finely re-tuned in context-specific ways to produce enormous variations in form. Orthologous mechanisms, primarily involving hormones and transcription factors, have influence in different contexts to control the balance between indeterminate and determinate growth to determine the fate and location of lateral leaf organs. Comparing and contrasting the key components in the development of leaves in different model species displays clearly how the same components are used in differing ways to generate diversity. Currently many of the key evolutionary modifications that have been identified have led to leaf shape diversification, but there is more still to be discovered that will improve our understanding. There is an indefinite way in which components of leaf shape regulation could be modified and further study of and comparisons of development in

model species is required to see which modifications are prevalent. This inter-species analysis can display how variation has arisen on a large scale over evolutionary time scales, but there is also potential in studying the great deal of intraspecific variation that exists between strains of the same species. This will show how leaf development programmes have been re-configured over different evolutionary scales to provide subtler changes in leaf shape.

1. 3. Natural variation within species

Leaf shape diversity between species is enormous and wide ranging, comparably diversity within species is much smaller and the variability is less striking and much more subtle. Studies exploring the developmental basis of intraspecific leaf shape variation are rarer than those which look at diversity in leaf shape across taxonomic groups. Naturally occurring intraspecific variation often occurs when species spread to and create isolated populations as demonstrated in this thesis. Another major driver of intraspecific variation is domestication.

1. 3. 1. Natural variation as a resource to study gene function

As many of the studies mentioned above demonstrate, induced mutagenesis has been a powerful tool for the functional analysis of genes in developmental regulation. However, the definition of gene functions with this approach has its limitations. For example, it is limited by the small number of genetic backgrounds analysed. The range of mutant phenotypes that can be identified is dependent on the wild type genotype. If the wild type accession carries a natural null allele (silenced or mutated) of a gene, mutant phenotypes for that gene will not be detected and even a weak allele might not be detected. This is a significant limitation as loss of function and null alleles exist in laboratory strains of *Arabidopsis* at a high frequency (Borevitz et al., 2003; Clark et al., 2007). Additionally gene function is often context dependent, this may be because of

gene-by-gene or/and gene-by-environment interactions (Tonsor et al., 2005; Chou et al., 2011). Epistatic interactions sometimes ensure that phenotypes only appear within a certain genetic background. The flowering time genes *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*) could not have been identified using the standard laboratory strains of *Arabidopsis.* The late flowering phenotype is only initiated when active alleles are carried at both loci, and in the common laboratory strains: L*er*, Col, and Ws, one or both of the genes is defective (Koornneef et al., 1998; Koornneef et al., 1994; Lee et al., 1994).

Thus, investigation of the genetic basis underlying diversity within species presents a valuable asset for the functional analysis of genes that regulate morphology, free from the limitations of mutagenesis approaches. This natural variation has been produced by random spontaneous mutations that have been selected for within different genetic backgrounds and environments. Therefore the study of natural variation will also facilitate the study of evolution and adaptation (Benfey and Mitchell-Olds, 2008).

1. 4. QTL analysis

Some leaf trait variation is qualitative in nature, such as the absence or presence of trichomes, but the vast majority of intraspecific variation in leaf shapes is of a quantitative nature. This is usually due to the effects of allelic variation at multiple loci, which when combined with environmental effects determine a continuous (i.e. quantitative) phenotypic distribution of trait values in segregating populations. The genotype at these various quantitative trait loci (QTL) cannot be directly known from the single phenotypic value; they can however be inferred from linked marker loci using a methodology referred to as QTL mapping. This approach is used to estimate the number and genomic location of segregating QTL in a specially designed

experimental mapping population. The population is first genotyped at markers throughout the genome and the trait or traits of interest are phenotyped and then specific software is available to perform tests of associations between phenotypic values of the trait and genotypic classes of the polymorphic markers. Where statistically significant associations arise a QTL is identified and located. From the analysis it is possible to deduce the number and genetic position of loci controlling the trait variation in that population, the additive effect of each locus, the contribution of epistasis between loci and depending on the mapping population used, any dominance effects can be evaluated. The mapping population used has a great influence over the results obtained. Any population of offspring in which the alleles of the ancestor have segregated will yield good results regarding the genetic nature of traits with QTL mapping. The simplest is an F2 population but there are more advanced mapping populations that introduce advantages over an F2. In particular, inbred lines are popular for QTL mapping. For example, recombinant inbred line (RIL) populations produced through self-fertilisation and single seed descent from an F2 population that was the product of a cross between two parental lines. After each generation of selfing, genomic heterozygosity reduces by 50% with the result that all individuals of the population are practically homozygous after about six generations. This lack of allelic segregation allows for the same genetic material to be analysed multiple times and phenotypic values can be based on multiple replicates to minimise environmental effects. In addition, the increased number of generations means that there are more meiotic recombination events that mix up the parental genomes further; this increases the power and resolution of QTL mapping without the need for greater population sizes. QTL analysis software has been available for over two decades that use the classical methods of QTL analysis to fit the effects of a single putative QTL (Lander and Bostein, 1989; Jansen and Stam, 1994). These have now been superseded by advances in mixed model (Cooper et al., 2009) and Bayesian methodology (Yi and Shriner, 2007) meaning that it is now possible to asses genetic effects with more complex models that continue to develop (Alonso-Blanco and Mendez-Vigo, 2014). The application of QTL mapping methodologies has continued to broaden with the availability of high throughput “omics” data (Jansen et al., 2009; Chitwood and Sinha, 2013).

Identification of either the gene(s) or the functional nucleotide polymorphism(s) responsible for creating the QTL effect is not straight forward. Depending on the quality of the QTL mapping assay and its relative effect, a QTL is assigned to a genetic interval of 5 to 50 cM. In *Arabidopsis* this can correspond to on average 1.2 to 12 Mb of DNA which would cover 240 to 2400 predicted genes (Koornneef et al., 2004; Price, 2006). Therefore, the QTL region must be reduced by fine mapping and this requires populations in which the multigenic nature of quantitative traits is reduced to a situation where the only source of genetic variation is around the QTL of interest. Commonly used for this purpose are near isogenic lines (NILs) and heterogenic inbred families (HIFs). NILs are created through repeated backcrossing and marker assisted selection so that there is a single genomic introgression fragment from one parental strain into an otherwise homogenous genetic background from the other parental strain. HIFs are derived from RILs that despite the repeated generations of self-fertilisation retain small regions of residual heterozygosity. Only these regions segregate in the following generation, if the genotype and phenotype co-segregate then the QTL gene is confirmed to be in the that region. Fine mapping is the validation of the QTL in ever decreasing (due to meiotic recombination) and overlapping segregating regions. If the region is small enough following this, candidate genes for the QTL can be selected. Due to the ubiquity of nucleotide polymorphisms between natural strains proof to positively identify a gene as causal to the QTL effects is usually required. Multiple strategies have been employed for this purpose (Alonso-Blanco et al., 2005) and a common strategy is that of complementation of the QTL effect by transformation with the alternative QTL allele. Many *Arabidopsis* QTL have had their genetic determinants identified (Alonso-Blanco et al., 2009) but as more plant genomes are sequenced and genomic resources are developed for other species an increasing number of QTL will have their underlying genes and functional polymorphisms identified too.

One of the major drivers of development in QTL mapping methodologies has been the agricultural industry. Information about loci for important agronomic traits in crop species can be used to implement marker assisted selection in plant breeding programmes (Alonso-Blanco et al.,

2009; Hospital, 2009). However improvements to QTL mapping methodology have also been of benefit in the field of evolutionary biology. Cases of this application include relating genetic variation in leaf shape and size (Langlade et al., 2005) and petal shape and size (Feng et al., 2009) to evolutionary change in *Antirrhinum majus,* the regressive evolution of eye morphology in cave dwelling fish (Protas et al., 2007) and the genetics of plant pollinator interactions (Martin et al., 2008). These studies did not identify the molecular basis underlying these QTL, but there is interest in identification of the exact genes and functional nucleotide polymorphisms creating QTL effects to discover novel gene function and to study natural variation in the context of ecology and adaptation (Alonso-Blanco et al., 2005; Tonsor et al., 2005; Mitchell-Olds et al., 2007; Benfey and Mitchell-Olds, 2008).

1. 4. 1. QTL mapping to study diversity in leaf shape

A QTL mapping approach is an ideal way to exploit natural genetic and phenotypic variation to uncover novel gene functions in the regulation and development of morphology, and thus to reveal some of the key steps that led to the evolution of diversity. As a trait, leaf shape is very amenable to having its genetic architecture dissected by QTL approaches; however, few studies have explicitly studied QTL leaf traits. This is surprising because such phenotypes have been associated with water use efficiency and thermoregulation which are traits important to yield (Nicotra et al., 2011; Chitwood et al., 2012). The QTL studies that have examined leaf morphology are often limited to analyses of size, length, width, and complexity (Pérez-Pérez et al., 2002; Holtan and Hake, 2003; Frary et al., 2004).

Arabidopsis has served as a model for many of these studies. Pérez-Pérez et al. (2002) used a RIL population derived from a Ler (Landsberg *erecta*) x Col-4 (Columbia) cross to locate 16 and 13 QTL that affected highly correlated leaf morphological traits in juvenile and adult leaves respectively. Two of the major QTL were able to predict nearly 45% of the total phenotypic variance in the leaf size of juvenile leaves. Three QTL were identified in a Ler x Cvi

(Cape Verde Islands) (Juenger et al., 2005) two of which co-located with QTL identified in the Ler x Col-4 RIL population. Interestingly these two QTL also mapped to regions that were identified as QTL for the size and shape of floral organs (Juenger et al., 2005), so it would seem that these two QTL regulate the size of lateral organs. Co-location of QTL for leaf and floral traits was also discovered using an F2 population of tomato (Frary et al., 2004). The *ERECTA*(*ER*) gene has been proposed to be linked to the effect of the major QTL identified in Pérez-Pérez et al. (2002). The Ler parental strain carries the *er-1* mutation and complementation studies using the wild-type *ER* allele in a L*er* background and vice versa partially replicates the QTL effect (Pérez-Pérez et al., 2010). The *ER* gene expresses a leucine rich repeat receptor kinase (Torii et al., 1996) that acts in organ primordia to coordinate cell growth and differentiation (Shpak et al., 2005).

Leaf growth can be assessed by image analysis of the flat rosette of *Arabidopsis* and in an experiment designed to investigate the genetic basis of growth related traits F9 RILs from a cross between L*er* and Shakdara (Sha, from Tajikistan) were used to locate 4, 5 and 4 QTL for total leaf area at 10, 15 and 20 days post germination respectively (El-Lithy et al. 2004). Many of these QTL co-located with QTL for traits associated with growth of the entire shoot, thus leaf size is clearly linked to growth rates of the shoot. One of the QTL identified here that affected leaf area mapped to an area shared by major QTL identified in the Juenger et al. (2005) and Pérez-Pérez et al (2002) studies. Such studies that identify the same major QTL in different strains for leaf traits can highlight which QTL have been important in the diversification of leaf form.

Tisné et al., (2008) measured the size of the sixth leaf and the average number and area of epidermal pavement cells in a RIL population derived from a L*er* x An-1 (Antwerp) cross. Co-locating QTL demonstrated that genetic effects on rosette and leaf level are directly correlated to variation at the cellular level, thus underlining the dependency of leaf variables on both leaf and whole plant control mechanisms. In this example a large proportion of the trait variance was attributed to epistatic interactions between QTL.

QTL mapping of morphology in *Arabidopsis* has benefited from new technologies in mapping populations. The F1s from multiple parental strains can be crossed together before multiple generations of selfing and single seed descent to create *Arabidopsis* multiparent RIL (AMPRIL) populations. This novel population has the advantage of introducing additional sources of variation into QTL mapping procedures relative to traditional mapping populations derived from only two parental strains. Huang et al. (2011) used an AMPRIL population to identify QTL associated with ratios of length and width, petiole length and leaf area. Another advanced mapping population is the Multiparent Advanced Generation Inter-Cross (MAGIC) lines derived from 19 inter-crossed *Arabidopsis* accessions (Kover et al., 2009).

QTL mapping studies of leaf shape have also been successful in identifying loci that have contributed to diversification in species other than *A. thaliana.* In *Brassica oleracea* an F2 mapping population was used to locate five QTL responsible for lamina length and four QTL affected petiole length (Lan and Paterson, 2001). In *B. rapa,* F2 and double haploid populations were used to map 2 QTL that are involved in leaf margin serration (Lou et al., 2007). Outside of the *Brassica* family, two QTL were found to affect the length and width of the middle leaflet in *Lotus japonicus* using a population of RILs (Gondo et al., 2007).

Tomato has been the subject of many QTL studies because of its economic importance. Many of these focus on fruit phenotypes but there have been studies that use QTL mapping to analyse variation in its compound leaf shape. Frary et al., (2004) used an F2 population from a cross between the modern cultivar *S. lycopersicum* and its small leaved, drought tolerant, wild relative *S. pennelli* and they were able to locate three QTL that influenced leaflet size and three QTL that affected leaflet shape. A similar methodology, but with aubergine (*Solanum melongena*) instead of tomato identified leaf size and shape QTL that located within the same syntenic regions of the tomato genome that contributed to leaf size and shape (Frary et al., 2003). This suggests that the genes responsible for leaf and leaflet morphology can be shared amongst species within the same family.

Another common QTL mapping population is introgression lines (ILs); each line contains a small introgressed region of one parental genome in the genomic background of the other parental strain or species. ILs derived from *S. lycopersicum* and *S. pennelli* have been used to map QTL marginal patterning variation; 22 QTL for variable levels of dissection and six QTL that affected marginal serration were found by Holtan and Hake (2003). One of the major effect QTL here mapped to the tomato *KNOX* gene *LeT6*, which has been proven to participate in regulation of dissection (Chen et al., 1997). In these studies there was limited overlap between QTL affecting leaf size and leaflet production (Holtan and Hake 2003; Frary et al., 2003, 2004) suggesting that the genetic control of size is independent of control of dissection.

A recent investigation has again used a population of ILs derived from *S. lycopersicum* and *S. pennelli* to map a remarkable 1035 QTL associated with leaf morphology (Chitwood et al., 2013). The authors genotyped the ILs at an ultra-high density using state of the art RNA-seq and reduced representation genomic sequencing to provide the exact gene content of each line. They then performed comprehensive phenotyping of over 11,000 leaves for many different leaf traits including mathematical descriptions of leaf shapes. The data were then reduced into fewer dimensions using principal component analysis and QTL mapping was performed on these. Interestingly, the analysis revealed genetic correlations between fruit glucose levels and genetic control of leaf morphology. Modern tomato varieties have been bred for improved fruit and not leaf shape, but it seems that leaf shape influences photosynthetic efficiency and is correlated with fruit sweetness. It would appear that leaf shape as been inadvertently bred alongside improved fruit during domestication (Chitwood et al., 2013). This is an exemplary case of where a QTL study has helped understand the genetic basis and diversification of leaf shape.

Another study that quantified leaf shape was performed by Langlade et al. (2005); geometrical morphometrics were employed to define a genetically controlled space to capture variation between closely related species of *Antirrhinum.* With these data, QTL were mapped in an F2 population derived from a cross between *A. majus* and its wild relative *A. charidemi.* These

QTL were then used to identify the possible paths by which divergence in leaf shape could have occurred.

Thus, quantitative genetics approaches hold considerable potential to reveal the genetic basis for evolutionary diversification of leaf shape. QTL mapping experiments are able to locate QTL that contribute to variation in leaf size, length, width, marginal complexity, dissection and quantitative measures of leaf shape in a multitude of taxonomic groups.

1. 5. *Cardamine hirsuta* as a model species for studying diversification in leaf shape

Cardamine hirsuta has emerged recently as a genetic system in which to study development. It is a small crucifer closely related to *A.thaliana* and it has many of the desirable features that make for a good model species. It is self-compatible and produces abundant seeds within an eight week generation time. It has a small rosette growth habit that can be cultivated efficiently on a large scale. The genome of *C. hirsuta* is approximately 1.5 times the size of that of *A. thaliana* (Johnston et al., 2005) and it is amenable to transformation by floral dip using *Agrobacterium tumefaciens.* This experimental tractability has allowed for molecular genetic approaches that use transgenic tools available in *A. thaliana* to manipulate and visualise gene expression to understand developmental regulation in *C. hirsuta* (Hay and Tsiantis, 2006; Barkoulas *et al.*, 2008; Canales *et al*., 2010; Vlad *et al*., 2014).

One feature that is not shared between these species is leaf shape: *A. thaliana* bears a simple leaves, whereas *C. hirsuta* produces compound leaves dissected into leaflets. Study of development in these two species that have a broadly comparable genotype to phenotype landscape has allowed identification of the key developmental changes that have resulted in the generation of alternative morphology (Hay et al., 2014; Hay and Tsiantis, 2006; Barkoulas *et al.*, 2008; Canales *et al*., 2010; Vlad *et al*., 2014).

Commonly called Hairy Bittercress or Popping Cress on account of its explosive seed dispersal, *C. hirsuta* is a cosmopolitan weed that thrives on recently disturbed ground. It has spread worldwide but is thought to be native of Europe (Lihova et al. 2006). Concurrent with this distribution was the generation of intraspecific morphological variation which forms the basis of the present study into the genetic basis of leaf shape divergence amongst *C. hirsuta* strains using a QTL mapping approach.

1. 6. Scope of this thesis

The aim of the work presented in this thesis is to uncover the genetic basis of natural variation in leaf form of *Cardamine hirsuta.* QTL mapping on variable leaf shape forms the basis by which loci controlling variation in this trait will be identified. This will be done with a RIL population derived from parental strains from Oxford, UK and the Azores Islands, Portugal. The principal trait to be mapped is variable leaflet number; which is a highly variable trait both between strains of *C. hirsuta* and between leaves of the same plant from different phases of development. Following QTL analysis, one of the mapped QTL was chosen to be validated, fine mapped and cloned.

In Chapter 3, I investigated the intraspecific analysis that exists between six naturally occurring *C. hirsuta* strains originating from different locations around the World. I analysed features associated with variable leaf shape for all rosette leaves. A common feature for all strains was that leaf shape varied with heteroblastic development. A major determinant of leaf shape variability was the rate at which the strain progressed through heteroblastic development. Also phenotypically analysed was the population of Oxford x Azores RILs which revealed that leaflet number is a highly heritable trait, that leaflet number on different leaves were highly correlated with each other and also with rosette leaf number, suggesting that there is overlap in their genetic control. QTL mapping was performed and confirmed this to be the case; in all multi trait QTL

analysis was able to identify six QTL, five of which affect both leaflet production and rosette leaf number.

In chapter 4, I tried to validate four of the detected QTL using heterogeneous inbred families (HIFs), three of which were successful. The QTL discovered on linkage group 4 (QTL-LG4) was selected for further analysis. Two successive rounds of fine mapping were performed and the QTL-LG4 interval was narrowed down to a region of 48.6 kb containing 15 genes. Further to this I make use of the lack of background genetic variance in the HIFs to comprehensively characterise the effect of this QTL on leaf shape in both long day and short day conditions. The QTL effect was largely unaffected by differences in photoperiod.

Finally, in chapter 5 I describe identification of the gene likely to be underlying the QTL-LG4 effect. I describe the selection process for candidate genes. Potential candidates were selected based on their proposed function from the literature, DNA polymorphisms and amino acid sequence polymorphisms within functional and conserved protein domains. To determine whether the QTL effect is mediated by differential expression of the gene product quantitative real time assays were used to assay relative levels of expression in developing leaves. Following this genomic fragments of both alleles of *SQUAMOSA PROMOTER PROTEIN BINDING LIKE 9* (*ChSPL9*) and *FLOWERING BASIC HELIX-LOOP-HELIX 4* (*ChFBH4*) were transformed into *A. thaliana* to see if there was a differential allelic effect on leaf shape. Leaf shape was affected differently by the alternative alleles of Ch*SPL9* suggesting that variation in this gene probably underlies the QTL-LG4 effect.

By characterizing a QTL for *C. hirsuta* leaf shape and complexity I have been able to show that the causes of diversity within species can be divergent from those that generate diversity between species. Previous work has shown that interspecific differences between *A. thaliana* and *C. hirsuta* leaf growth result mostly from variation in local tissue growth and patterning (Vlad et al., 2014; Hay et al., 2006; Barkoulas et al., 2008). Here, I have identified that an age-dependent leaf shape progression underlies variation in leaf complexity within species.

Chapter 2: Materials and methods

2. 1. Plant growth conditions

Experiments shown in Figures 3.2 – 3.6 and 4.1 - 4-8 used plants grown at Department of Plant Sciences, University of Oxford, UK (DOPS). Controlled environment rooms were set to long day conditions (days: 24°C, 16h; nights: 22°C, 8h, 113 µmol m^{-2} s^{-1} light and 70% relative humidity). The Ox x Azores RILs grown for the QTL mapping experiment (Figures 3.3 – 3.11) were sown in a separate controlled environment room also set to long day conditions but at a lower temperature (days: 18°C, 16h; nights: 16°C, 8h, 113 µmol m^{-2} s^{-1} light and 70% relative humidity). The plants for Figures 4.9 – 4.14 and 5.8 were grown at the Max Planck Institute for Plant Breeding Research in Cologne, Germany (MPIPZ). A greenhouse was used for the long day conditions (days: 22°C, 16h; nights: 20°C, 8h) and a growth cabinet was used for short day conditions (days: 20°C, 8h; nights: 18°C, 16h, 60% Light and 70% relative humidity).

Dry seeds were sown on a well watered soil mixture of 2 parts peat to 1 part vermiculite in 7x7x8 cm pots, then covered and placed at 4°C to stratify for 7 – 10 days before being transferred to their growing conditions. Covers were removed once all seeds had germinated.

In the case of growing the recombinant inbred line (RIL) population for QTL mapping, to ensure enough individuals were propagated for each line, over 10 seeds per line were sown per 7x7x8 cm pot. Of all the successful germinators the seedling nearest the middle was allowed to grow to represent that line, the rest were removed and discarded. This protocol was used so that no selective bias was inadvertently introduced. The RILs and their replicates were randomised within trays and the location of each tray within the controlled environment room was shuffled every three days.

When growing plants for QTL validation, four seeds were grown per pot. Following germination, a single cotyledon was removed for DNA extraction and the appropriate PCR markers were used to determine the genotype of each seedling at the QTL interval. All heterozygotes were removed and homozygotes were retained and allowed to mature. If more than one homozygote was left per pot seedlings were transplanted to new pots before there was any

danger of significant damage to root systems. This was done to ensure that crowded growth conditions did not affect leaf growth.

2. 2. Plant material

2. 2. 1. *Cardamine hirsuta* Strains

The *C. hirsuta* Oxford strain has been described previously (Hay et al., 2014) and all other strains are listed in Table 2.1

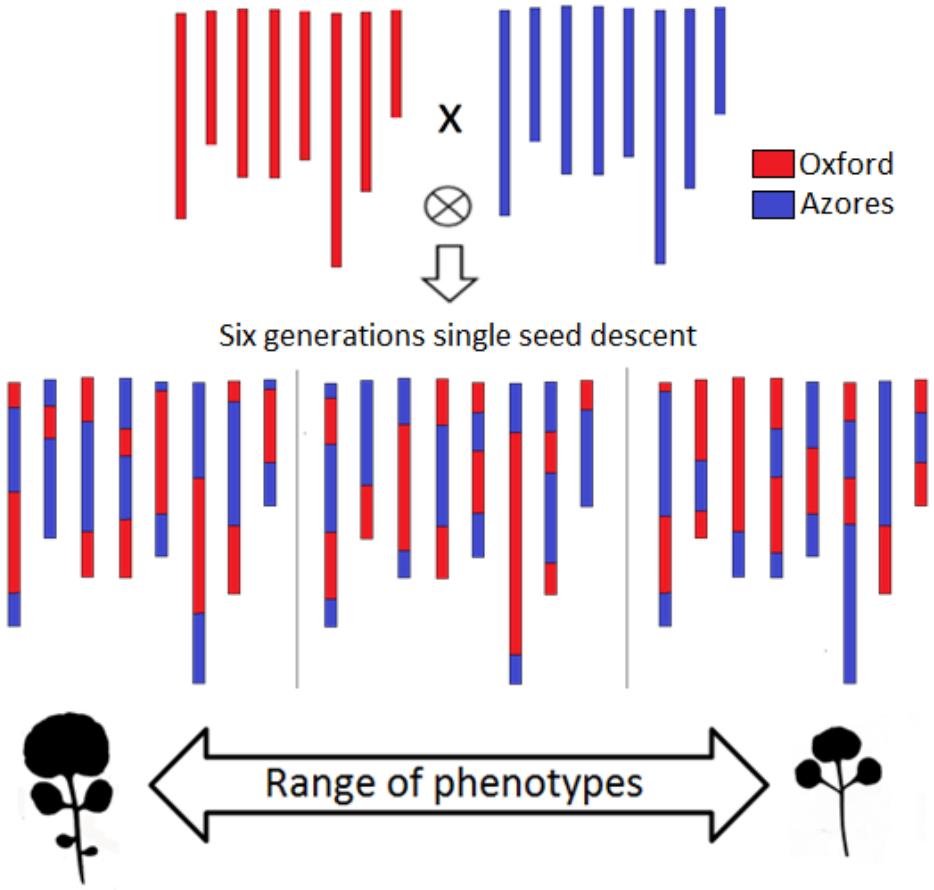
Table 2.1 List of *C. hirsuta* strains used in this thesis.

Strain	Collector	Country	Location collected	Latitude	Longitude
Oxford	Tsiantis, M.	UK	Oxfordshire, Oxford	N 51° 45'	W 1° 15'
Azores	Tattersall, A.	Portugal	Azores, Faial, Flamengos	N 38° 33'	W 28° 39'
Cologne	-	Germany	Nordrhein-Westfalen, Cologne	-	-
Japan	Lihova, J.	Japan	Niigata, Nishi K	N 37° 54'	E 139° 02'
New Zealand (Nz)	Hay, A.	New Zealand	Wellington	-	-
Washington (Wash)	Zika, P.F.	USA	Washington, Clark county	N 45° 44'	W 122° 31'

2. 2. 2. *Arabidopsis thaliana*

All experiments described in this thesis that used *Arabidopsis thaliana* were of the ecotype Columbia and were obtained from the stock grown at MPIPZ.

Figure 2.1 Construction of Oxford x Azores F7 Recombinant Inbred Line (RIL) population. The Oxford and Azores parental strains were crossed to create an F1 and this was followed by six generations of single seed descent. Mixing of the parental genomes creates a range of phenotypes.


x
Oxford
Azores
Six generations single seed descent
Range of phenotypes

2. 2. 3. Recombinant inbred line (RIL) population

I did not have a role in the generation of the Oxford x Azores RIL population used in this thesis; the work was performed by previous members of the Tsiantis group at Oxford. Following the collection of the Azores accession from Faial Island, Azores (Table 2.1) was selfed multiple times then crossed to the Oxford strain which itself had been selfed eight times. A single F1 progeny was selected and allowed to self fertilize to generate an F2 population (Figure 2.1). A total of 195 F2 progeny were propagated through selfing and single seed descent to generate the F7 RIL population. In the QTL mapping experiment 149 of the RILs were genotyped and phenotyped.

The genome of the RILs was assumed to be practically homozygous (later analysis revealed that the RILs were homozygous at an average of 97.5% of the markers used) which meant that it was possible to grow three replicates per line. This helped minimise non-genetic environmental variation. Material for DNA extraction was freeze dried and material from all three replicates was pooled prior to DNA extraction and therefore genotyping information generated represents an F6 RIL population rather than the individual F7 replicates because of the small amount of segregating loci.

2. 3. Materials used

2. 3. 1. Bacterial strains

For plasmid amplification the DH5α strain of *Escherichia coli* was used. For transformation of *A. thaliana* the GV3101 strain of *Agrobacterium tumefaciens* was used.

2. 3. 2. Buffers and media

Media, solutions and buffers were prepared following the protocols described by Sambrook and Russell (2001).

2. 3. 3. Oligonuleotides

Oligonucleotides used for genotyping of HIFs for QTL validation and fine mapping, quantitative real time PCR, amplification of candidate genes and sequencing of clonal inserts are listed in Table 2.2. The design of primers for cleaved amplified polymorphic sequence (CAPS) markers and for derived cleaved amplified polymorphic sequence (dCAPS) markers was guided by the web based software dCAPS Finder 2.0 (http://helix.wustl.edu/dcaps/dcaps.html; Neff et al. 2002). Primers for amplifying fragments to be cloned, for sequencing insertions and for quantitative RT-PCR analyses were designed using the web based Primer3 software (http://primer3.ut.ee/; Untergrasser et al., 2012). All primers were constructed by Sigma-Aldrich.

Single nucleotide and insertion/deletion polymorphism for which genotyping markers were designed around were selected from preliminary whole genome resequencing alignments (unpublished).

Table 2.2 Primers used for PCR, genotyping quantitative RT-PCR, sequencing and genotyping

Primer name	Primer sequence	Use
JLc4_10629-BamHI-F	TGACCAATTCCTGAAAACCC	LG4 QTL validation
JLc4_10629-BamHI-R	GAAGCTATTGATATTTCTCGTGGATC	LG4 QTL validation
JLc4_10747-NlaIII-F	GTAAGTGATTGCAAACCCC	LG4 QTL validation
JLc4_10747-NlaIII-R	GGATCAAGATGTAAGCAACTTGG	LG4 QTL validation
JLc2_5287-NlaIII-F	CCGAAGAAAACAACTGTCAGAAAACCTCAT	LG2 QTL validation

JLc2_5287-NlaIII-R	GTCTGTACTGAGGCACTTTC	LG2 QTL validation
JLc3_7369-MboI-F	CAAGTCCATACCACTGGTCTTTACAATTGC	LG3 QTL Validation
JLc3_7369-MboI-R	GTGATAGCTGTCAAAGATCGGACTATG	LG3 QTL Validation
JLc8_24987-MseI-F	GCTTCATCTACAAAGCTCCCCTCGG	LG8 QTL Validation
JLc8_24987-MseI-R	GTGTCAACAAAAGTGGCGCG	LG8 QTL Validation
JLc8_24829-RsaI-F	CATTCCTGTGCAAAGACATGTCGGAC	LG8 QTL Validation
JLc8_24829-RsaI-R	TCTCATCTATTGCTGGTGGTAATAGGTA	LG8 QTL Validation
JLc4_10462-MseI-F	GATTAATTTCAACTTCCTGATGAGTTTCAG G	Delimiting LG4 Validated region
JLc4_10462-MseI-R	ACAGTTGATTCAGGCTCTAAAGCTCTT	Delimiting LG4 Validated region
JLc4_10607-RsaI-F	CAGAGAGATTCCTCATCTCCGATGG	Delimiting LG4 Validated region
JLc4_10607-RsaI-R	GCAGATGGTTGCGGAAACGGAGTGT	Delimiting LG4 Validated region
JLc4_10763-HpaII-F	GGGAGTCATGGACAATAAGCAAGAGC	Delimiting LG4 Validated region
JLc4_10763-HpaII-R	GCTGGAGAAATATCTGAATCCTCTC	Delimiting LG4 Validated region
JLc4_10770-MseI-F	GCATCGCATACTCATCCATACTGG	Delimiting LG4 Validated region
JLc4_10770-MseI-R	CCATTATTGGGAAAAATGATGAGGCCTT	Delimiting LG4 Validated region
JLc4_10862-HhaI-F	CCGCTGCTCTCGCAGCTTCTTCCATTTTCC G	Delimiting LG4 Validated region
JLc4_10862-HhaI-R	GGTGAAACCACAGGTTCTTGATGAACAGG	Delimiting LG4 Validated region
JLc4_10999-HhaI-F	GGGTTTGAGACAAACAAAGCAG	Delimiting LG4 Validated region
JLc4_10999-HhaI-R	GCAAACTTCATCTCCAACATCTGC	Delimiting LG4 Validated region
JLc4_9935-AluI-F	GGTTAATAAAGTTGAAGACTTTAGCAG	Delimiting LG4 Validated region
JLc4_9935-AluI-R	GGTACCATCAGTTGCAGTAACAG	Delimiting LG4 Validated region
JLc4_9996-AluI-F	GGTTATAGAGGAAGATAAGAGAGCCTTC	Delimiting LG4 Validated region

JLc4_9996-AluI-R	CAGTTGGTAACTCAAGGTTCTGGTTTCTAAGC	Delimiting LG4 Validated region
JLc4_10151-HhaI-F	AGTTATAGATGGCAAAGCTTTGCATTG	Fine mapping LG4 QTL
JLc4_10151-HhaI-R	GGATGGGGTGGTCCTGGTGG	Fine mapping LG4 QTL
JLc4_10218-DdeI-F	CTTTGCCTCGTCAACTAATTTAGAAGCC	Fine mapping LG4 QTL
JLc4_10218-DdeI-R	CAAGTAGATTAGTCAGATCAGCACAAAC	Fine mapping LG4 QTL
JLc4_10240-MboI-F	GGAGCTAATGAGAGCAGTTCTGG	Fine mapping LG4 QTL
JLc4_10240-MboI-R	TCTTTAGTGGATTCCGAGCCCGAGTCAAGA	Fine mapping LG4 QTL
JLc4_10330-AluI-F	GGCATCAGCACTTATCTTCGAAAAG	Fine mapping LG4 QTL
JLc4_10330-AluI-R	GAATGTCGGATATCCTTGCGGTTTAAGC	Fine mapping LG4 QTL
JLc4_10570-RsaI-F	GGCATACACATGGTTGAGTCCAAGG	Fine mapping LG4 QTL
JLc4_10570-RsaI-R	GAACCCCCTTCTGATGAAAACTGG	Fine mapping LG4 QTL
JLc4_10701-HhaI-F	ACCAGTTTGCTTGAGCTTTTCTAGATGGC	Fine mapping LG4 QTL
JLc4_10701-HhaI-R	GGCGTCTCCGCCGCAACCGTCGG	Fine mapping LG4 QTL
JLc4_10974-MboI-F	GACGGTGGCGGCGGCGAGAGAG	Fine mapping LG4 QTL
JLc4_10974-MboI-R	GTGGACACCGAATCGATCAACGG	Fine mapping LG4 QTL
JLc4_11091-AluI-F	CCTTCAACTTCAAGGAGCATTCC	Fine mapping LG4 QTL
JLc4_11091-AluI-R	GTCGGGTGGCTGCATTTTTGG	Fine mapping LG4 QTL
JLc4_9924-MboI-F	GCATCCCTTTCTTCTTTTGTCAGCACACG	Fine mapping LG4 QTL
JLc4_9924-MboI-R	CCAGATCCAGTCACGTCGCATC	Fine mapping LG4 QTL
JLc4_10004-HpaII-F	CCGTTCTCCACCGCTCTCGCCGCC	Fine mapping LG4 QTL
JLc4_10004-HpaII-R	CCGGAATGGCGTCTCGGAGTTC	Fine mapping LG4 QTL
JLc4_10363-DdeI-F	GGTTTGCTAGGTACATCGGCGG	Fine mapping LG4 QTL
JLc4_10363-DdeI-R	CCCTAGTGGTTCAAACTAGGACC	Fine mapping LG4 QTL
JLc4_10589-NlaIV-F	GGCATCCCAGTGTGTGAAGGAGCTGTAGGCAC	Fine mapping LG4 QTL
JLc4_10589-NlaIV-R	CCATTCTCTAATGTTTCCCCCTC	Fine mapping LG4 QTL
JLc4_10791-HhaI-F	GTCACGGGGATTCACTTTTCGATCG	Fine mapping LG4 QTL

JLc4_10791-HhaI-R	GGCACTTGGACATCAAGAGAAGGGCG	Fine mapping LG4 QTL
JLc4_9919-HpaII-F	GAGCGAGCATAGAAGAGTAGAG	Fine mapping LG4 QTL
JLc4_9919-HpaII-R	CGTAAGAGAGCTTCTTCAATCTCAATTCCC	Fine mapping LG4 QTL
FMLG4_6671-DdeI-F	GCAAATGATTGTAGCTAACTCCATCAACG	Fine mapping LG4 QTL
FMLG4_6671-DdeI-R	ATTATGGAAGCATTGATTGATGAAAACAAC	Fine mapping LG4 QTL
FMLG4_9215-RsaI-F	GTAATGATCACACCCAATAATTAGACGTA	Fine mapping LG4 QTL
FMLG4_9215-RsaI-R	CTGCTTGAGGTTTGACCAACAAAGTCGTC	Fine mapping LG4 QTL
FMLG4_1118-RsaI-F	CTCCTGCAGCTATCTAAAAGCAATCCAGTA	Fine mapping LG4 QTL
FMLG4_1118-RsaI-R	CCATCAAAGAGCTCGATATCATGTTAAACC	Fine mapping LG4 QTL
FMLG4_0403-HhaI-F	CCGTCCGAATGACACGTGGAATCC	Fine mapping LG4 QTL
FMLG4_0403-HhaI-R	GAAGTGGAGCATGAAGCTAGGG	Fine mapping LG4 QTL
FMLG4_8309-AluI-F	CAGATTCATAGCCTAAGTTAGATTCGG	Fine mapping LG4 QTL
FMLG4_8309-AluI-R	GATATTTTGTGTTTGCCCATACCAAATCAGC	Fine mapping LG4 QTL
FMLG4_6811-RsaI-F	GGCTGTTGCTACGAACAAGCTGCAAG	Fine mapping LG4 QTL
FMLG4_6811-RsaI-R	CCCATGTTTTAGTGATGATTTTTCTCAGT	Fine mapping LG4 QTL
FMLG4_6864-AluI-F	CTCTAAAAGGAACTTCCAAATTAAGCAG	Fine mapping LG4 QTL
FMLG4_6864-AluI-R	GGCAAAACATATCTTGCCCCAGCATG	Fine mapping LG4 QTL
FMLG4_1491-MboI-F	CATGTAAACGGCAAATCATCCAAGCCC	Fine mapping LG4 QTL
FMLG4_1491-MboI-R	GACACAGAAAAATATCCATTATGGGGG	Fine mapping LG4 QTL
FMLG4_9035-AluI-F	GCTAAGAGAAGTAGCAGCACTCGAGATTTA	Fine mapping LG4 QTL
FMLG4_9035-AluI-R	GCATGGATGTAGATTCTAGAAATACGACGAG	Fine mapping LG4 QTL
FMLG4_0561-HhaI-F	CGCAAGATTCGTCCTCCAGC	Fine mapping LG4 QTL
FMLG4_0561-HhaI-R	GCGATGACTAGCGTGTCGAGGCCACG	Fine mapping LG4 QTL
FMLG4_0101-MseI-F	CGTCAACATGGAATAAAATACTTTAATATGC	Fine mapping LG4 QTL

FMLG4_0101-MseI-R	CAAACGTAGTAACCATGTTGTTGCC	Fine mapping LG4 QTL
fFMLG4_1358-AluI-F	GGGATGTTAAATCTTATTACAAATAAACT CAG	Fine mapping LG4 QTL
FMLG4_1358-AluI-R	CTTCAAGCATCTCTCGTTTACCTTGC	Fine mapping LG4 QTL
FMLG4_0063-MseI-F	CCGCCGGAGATCTAAGAAACATCTCC	Fine mapping LG4 QTL
FMLG4_0063-MseI-R	GGATGATGGTGGTTATAGGGTGTTGGG	Fine mapping LG4 QTL
FMLG4_8857-RsaI-F	GGACACGAGCCGCATCAGCATCATG	Fine mapping LG4 QTL
FMLG4_8857-RsaI-R	GGCTTACGGCATGTTAATTTATCTAATGTA	Fine mapping LG4 QTL
FMLG4_0595-MboII-F	CTTGATCTCTACGTCAAGCAACAGCC	Fine mapping LG4 QTL
FMLG4_0595-MboII-R	GATGCTTTCAGATCTCAAATGAACTGGAA	Fine mapping LG4 QTL
FMLG4_3404-length-F	GTTGAGGGACACGTGGACGTG	Fine mapping LG4 QTL
FMLG4_3404-length-R	CAATTTAGGTTAGTAGATTTTATGC	Fine mapping LG4 QTL
FMLG4_6436-length-F	GGTTAACATGTTTGACCCCAAAAAG	Fine mapping LG4 QTL
FMLG4_6436-length-R	GTGTTATTAGCATGTTTACCTAATTGTATT G	Fine mapping LG4 QTL
FMLG4_0571-length-F	GTTCAGTCTTCAGATCAGACAAGG	Fine mapping LG4 QTL
FMLG4_0571-length-R	CCACACAAGATTCAAACTCATATTTCC	Fine mapping LG4 QTL
LG4-SPL9-1F	GATACAGGGTGACAAATTAACCTAAGG	Amplification of candidate gene
LG4-SPL9-1R	CTACGGTTCTACCTAATTGAAAAA	Amplification of candidate gene
LG4-SPL9-2F	AAATTAACCTAAGGCTTCTTCTTC	Amplification of candidate gene
LG4-SPL9-2R	TTCTTCTACGGTTCTACCTAATTG	Amplification of candidate gene
LG4-RRM-1F	GATAAATTTGTCATTTGAGAAAGAAAGAG	Amplification of candidate gene
LG4-RRM-1R	CCTGTTTCTATAAGATACCCAAGAAG	Amplification of candidate gene
LG4-RRM-2F	CAGAGATGGTAGAGAGAGAGAGAG	Amplification of candidate gene
LG4-RRM-2R	ATGGTTCTGTATCATTTCAAAGTT	Amplification of candidate gene

LG4-FBH4-1F	GCAACAAGACTTCTTCAGGCAATGG	Amplification of candidate gene
LG4-FBH4-1R	GGCGAGTTTAATTGAAGGGTTTAGGTTT	Amplification of candidate gene
LG4-FBH4-2F	AACAAGACTTCTTCAGGCAATG	Amplification of candidate gene
LG4-FBH4-2R	GTTTTGTAGTTTATATGGCGAGTT	Amplification of candidate gene
LG4-Bromo.qRT-2F	GTCAGCTTCTACTTCTTCTACACG	qRT-PCR of *ChBROMO* transcript
LG4-Bromo.qRT-2R	CATCCTATACCCTCTAACCATTGGCC	qRT-PCR of *ChBROMO* transcript
LG4-C-CAP.qRT-1F	GCCAAGGATTCATTCCATTTGTTGATAGC	qRT-PCR of *ChC-CAP* transcript
LG4-C-CAP.qRT-1R	CCACAAACTACAGGTGTAGCATCTCTCT	qRT-PCR of *ChC-CAP* transcript
LG4-RRM.qRT-1F	GAATCTGGTAAGATCATTGGTACAAG	qRT-PCR of *ChRRM* transcript
LG4-RRM.qRT-1R	CCAATTGTGTTGTAAGCAGAACTACTC	qRT-PCR of *ChRRM* transcript
LG4-FBH4.qRT-1F	GTAGATGATGACTTCGGTAAACTCAAG	qRT-PCR of *ChFBH4* transcript
LG4-FBH4.qRT-1R	TATCATCCCTTAAAAGAGAAACAG	qRT-PCR of *ChFBH4* transcript
LG4-LRR.qRT-1F	GTATATACATCAAACCTCGAGCAG	qRT-PCR of *ChLRR* transcript
LG4-LRR.qRT-1R	GGCAGCTTATACTCTGCTTTACATGGT	qRT-PCR of *ChLRR* transcript
LG4-bHLH.qRT-2F	CATCTCAACGACTTGTTTCTTATTC	qRT-PCR of *ChbHLH* transcript
LG4-bHLH.qRT-2R	GTATGGTTATCAGTAGCTTGACCACG	qRT-PCR of *ChbHLH* transcript
LG5-SPL9-1F	AGATCATAGCATACTACAGTGATGAGG	qRT-PCR of *ChSPL9* transcript
LG6-SPL9-1R	GTTTGTGTCGCCAATTCCCTTGTAGG	qRT-PCR of *ChSPL9* transcript
FBH4seq1-F	GGATCAGCGTGAGGTCCTCTTGTCTG	Sequencing inserts
FBH4seq2-F	CTGTTTCTCTGTTGATGTCAAGGATTC	Sequencing inserts
FBH4seq3-F	CCTCATTCGCTCGCTTATCCGC	Sequencing inserts

FBH4seq4-F	CTCTCTTCATCTTCGTCATAATTCATC	Sequencing inserts
FBH4seq5-F	CTAAAACGAAGAAGACCCGACCC	Sequencing inserts
FBH4seq6-F	CAGTAGATTTGAGATGAAAATACACAC	Sequencing inserts
FBH4seq7-F	CCGATTTAAGGGTTGGTTCATGGAACC	Sequencing inserts
FBH4seq8-F	GGAGACATATAGACTGGTGAGAGGG	Sequencing inserts
FBH4seq9-F	CTGTAGCATGTAGAAAGAGTTTGC	Sequencing inserts
FBH4seq10-F	GCACCAGAACCAATTTCCACATAGAG	Sequencing inserts
FBH4seq11-F	GATTACATTGATAGAAGTGGGG	Sequencing inserts
FBH4seq12-F	CCCACTTGGTCATATTAGATAGAAATC	Sequencing inserts
FBH4seq1-R	GTACCACAACAAAAAGAGAAAGAGGCAA	Sequencing inserts
FBH4seq2-R	CATCTTTCCTAACAAATTCTTTTGTCAC	Sequencing inserts
FBH4seq3-R	GCACTTGGACTCAAGACATGAAGTG	Sequencing inserts
FBH4seq4-R	CCTGAAACGTCACGAAACAGAGC	Sequencing inserts
FBH4seq5-R	CACGGGCCCATTATAGTACCTCGCC	Sequencing inserts
FBH4seq6-R	CGTCTTTTCTAATGCAGGTGTACC	Sequencing inserts
FBH4seq7-R	GTATGTGCATGCACGTGCTCGTG	Sequencing inserts
FBH4seq8-R	GTCCCCTTTTGTTATTAATTTGTCTTTG	Sequencing inserts
FBH4seq9-R	GCGATTAAAATCTAAAGAAACTTATAAGG	Sequencing inserts
FBH4seq10-R	GCTTAATTTGGAAGTTCCTTTTAGAG	Sequencing inserts
FBH4seq11-R	CAGTCCTTCCTTACATCATTAC	Sequencing inserts
FBH4seq12-R	GTATGTTATTAATCTCAGACAGCATGATG	Sequencing inserts
RRMseq1-F	GCCTCAAACCAGAGATGGTAGAG	Sequencing inserts
RRMseq2-F	GCCGTTCGCATTTGCTGTTTTCTCCG	Sequencing inserts
RRMseq3-F	GATCCCCACGCTCACGCTCTTTAG	Sequencing inserts
RRMseq4-F	CAAGGTAAGAATACTCTGTTCTGCAC	Sequencing inserts
RRMseq1-R	CACCGATTTAACCGTCTACAC	Sequencing inserts

RRMseq2-R	CCAGGAGTGTGAGAGCTGCCAAAACC	Sequencing inserts
RRMseq3-R	CAACAGGAGTCTGATATGTGCCC	Sequencing inserts
RRMseq4-R	GTGTTGTAATGCTACATGTTTACCTG	Sequencing inserts
SPL9seq1-F	GTGATGAAACAAACAACTTCAAG	Sequencing inserts
SPL9seq2-F	CACATGGAAACTACTTTCAC	Sequencing inserts
SPL9seq3-F	GATTATGAAGATGTGTAAAGTTGGATG	Sequencing inserts
SPL9seq4-F	GCCTAAGTTAGATTCGGTTTG	Sequencing inserts
SPL9seq5-F	CTGTCTGAAGAGGAGGAAACC	Sequencing inserts
SPL9seq6-F	GTTGGGCTTATTTTGGGGAATATG	Sequencing inserts
SPL9seq7-F	GTTCAAGAACGTTGGATACAAGAG	Sequencing inserts
SPL9seq8-F	CCAAACTATTGTTGCCAAAGATTGGC	Sequencing inserts
SPL9seq1-R	CTTTGTCATAGACATCCATCAGAGTC	Sequencing inserts
SPL9seq2-R	GAGTATCTATTTGACACCAGGG	Sequencing inserts
SPL9seq3-R	GAGGTTGGCAATAAATGAAGATG	Sequencing inserts
SPL9seq4-R	GGCCAAACATGAGACCTCCACTG	Sequencing inserts
SPL9seq5-R	GGCTAAGGGTTTCTTATTGTCCATTCAT	Sequencing inserts
SPL9seq6-R	GTCCATAGCCTCTGGAGATGAAAAG	Sequencing inserts
SPL9seq7-R	GTTTCAAGTGTTGATAAATGGGGGAG	Sequencing inserts
SPL9seq8-R	CCTAATTGAAAAATATTCACCAGAAAACAC	Sequencing inserts

2. 4. Genetic methods

2. 4. 1. Bacterial transformations

For transformation of chemically competent DH5α *E. coli* cells; 100ng of plasmid was added to 100µl of cells and these were incubated at 4°C for 30 minutes. The cells were then heat

shocked at 42°C for precisely 30 seconds, before being incubated at 4°C again for 2 minutes. These cells were then added to 200µl LB liquid medium and incubated at 37°C while being agitated at 225rpm for 45 minutes. Cells were then plated on LB agar that had the appropriate antibiotics added, to be incubated at 37°C overnight. For plasmids with ampicillin resistance, cells were grown on 100µg/ml ampicillin (Sigma) LB agar plates and for plasmids with spectinomycin resistance, cells were grown on 70µg/ml spectinomycin (Sigma) LB agar plates.

When transforming electrocompetent GV301 *A. tumefaciens*, 100ng plasmid was added to 100µl cells which were electroporated at 2kV, 200Ω and 25µF. 500µl LB liquid medium was then added to the cells which were then incubated at 28°C and agitated at 225rpm for 2 hours. Cells were transferred to 25µg/ml gentamycin and 70µg/ml spectinomycin (Sigma) LB agar plates and incubated at 28°C for 2 days.

2. 4. 2. Plant transformation of *A. thaliana* and *C. hirsuta*

Vectors containing the fragment of interest were introduced into the GV3101 strain of *A. tumefaciens* by electroporation and a single positive colony was used to inoculate 5ml LB liquid medium with the appropriate antibiotics and incubated at 28°C for 2 days. This culture was used to inoculate 1 Litre YEB with the appropriate antibiotics added and incubated at 28°C until an optical density of OD_{600nm} = 0.7 - 1 was reached. These cells were pelleted using a centrifuge and re-suspended in 500ml infiltration medium (5% sucrose (Fisher), 10mM $MgCl_2$ (Sigma), 1x Gamborg's vitamins (Sigma), 100ng/ml 6-benzylaminopurine (Sigma), 0.03% silwet-77 (GE Silicones). The inflorescences of *A. thaliana* plants with unopened buds were dipped into the infiltration medium for 12 minutes and the whole plants were wrapped in cling film then stored at 4°C overnight before being transferred to the greenhouse. The resultant seeds were harvested once dried.

2. 4. 3. Selection of transgenic plants

Selecting successfully transformed seedlings was done by spraying them with 400µM BASTA (commercial name "Kaspar"; Aventis), as soon as the cotyledons were fully expanded.

2. 5. Molecular biology methods

2. 5. 1. Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) is a molecular biology technique that exponentially amplifies a target sequence of DNA. For standard PCR conditions 1µg of genomic DNA was added to a PCR mix of 1x Mango*Taq* buffer (Bioline), 2.5mM $MgCl_2$ (Bioline), 0.1mM dNTPs (Fermentas), 8µM forward primer (Sigma-Aldrich) 8µM reverse primer (Sigma-Aldrich) and 0.3U *MangoTaq* DNA polymerase (Bioline), in a final volume of 25µl in a 0.2ml PCR tube (Starlab). The reaction itself was performed in an Applied Biosystems 2720 Thermal Cycler using the protocol: 1 cycle at 94°C for 3 minutes, 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 72°C for 15 – 30 seconds per kbp extension, followed by 10 minutes at 72°C.

PCRs for amplification of genomic fragments containing the candidate genes were performed using the high fidelity DNA polymerase Platinum *Taq* (Invitrogen). The PCR mix consisted of the following; 60ng template DNA, 2µl each of 10µM forward and reverse primer (Sigma-Aldrich) and 45µl Platinum supermix (Invitrogen). The reaction was performed with an Applied Biosystems 2720 Thermal Cycler using the protocol; 1 cycle at 94°C for 3 minutes; and 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds and 68°C for 1 minute per Kb extension.

2. 5. 2. Sequencing reactions

Sequencing was performed on the genomic fragment inserts to ensure that no sequence errors had been introduced during PCR amplification. All sequencing was performed by the Max Planck Genome Centre, Cologne (http://mpgc.mpipz.mpg.de/). Any clones that had sequence polymorphisms were discarded.

2. 5. 3. Genomic DNA extraction

In order to extract small amounts of plant genomic DNA for analysis, young leaf tissue (at least 0.1g) was collected in either collection microtubules (racked 8x12, Qiagen), if larger segregating populations were being analysed or in 2ml Eppendorf tubes if fewer samples were being analysed. In either case the same DNA extraction method was used. Included in the tubes along with the plant material were two 3mm tungsten carbide beads (Qiagen). Shortly following collection the material was frozen and stored at -80°C until required. After removal from the -80°C freezer, 400µl of extraction buffer (250mM Tris-HCl pH 7.5, 250mM NaCl, 25mM EDTA, 0.5% SDS) was added. The samples were transferred to a Retsch MM 300 macerating machine and shaken at 30Hz in 1 minute intervals until the samples appeared fully homogenised. Plates were spun at 6000rpm in a centrifuge (Sigma 4-15C) for 20 minutes. 350µl of the supernatant was transferred to a new rack of collection microtubules (or 2ml Eppendorf tube) and to each well 300µl of propan-2-ol was added. The samples was mixed by tipping and then incubated at room temperature for 60 minutes to precipitate DNA. The samples were spun at 6000rpm in a centrifuge (Sigma 4-15C) for 20 minutes and the supernatant was removed. The pellets were washed with 300µl of 70% ethanol before being air dried. To dissolve the DNA 100µl ddH_2O was added. For PCR analysis the DNA was dilutes 1:10 in ddH_2O and 10µl was used per 25µl reaction.

2. 5. 4. Restriction endonuclease digestion of DNA

For genotyping the main markers used were cleaved amplified polymorphic sequence (CAPS) markers or derived cleaved amplified polymorphic sequence (dCAPS) markers which were designed through analysis of sequence information (for primers see Table 2.2). For digestion 10µl of PCR product was added to a digestion mix of 0.5µl of restriction enzyme (New England Biolabs, NEB)), 1.5µl of the corresponding buffer and if suggested by the NEB protocol 0.15µl of BSA. ddH_2O was added to bring the final volume to 15µl total volume. The reactions were incubated as instructed by NEB protocols and the resultant products were separated by agarose gel electrophoresis, as described below.

2. 5. 5. Agarose gel electrophoresis of DNA

PCR products and digested PCR products were separated in gels by electrophoresis. The gels were made from 3% agarose in 1x TBE buffer (0.089 M Tris base, 0.089 M boric acid and 0.022 M EDTA) and ethidium bromide was added to a concentration of 1µg/ml. If the PCR was performed using mangotaq (Bioline) 5X buffer the PCR product was loaded directly into a well in the gel. If the PCR was not performed using a buffer which already contained a loading dye 0.1 volume of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol and 30% glycerol in water) was used. Typically 10-15µl was loaded into 4mm deep wells. So that the sizes of DNA fragments could be estimated, 125ng of DNA size ladder was loaded into each row of wells in a gel. Depending on the size of the DNA fragment being analysed, gels were run at 60-120A/cm for 30–90 minutes. The ethidium bromide allowed for visualisation of the DNA when illuminated using a UV trans illuminator and photographed for records.

2. 5. 6. Quantitative RT-PCR analysis

For expression analysis of potential QTL genes, whole seedlings bearing five visible leaves were collected and stored at -80°C. Total RNA was extracted from these using the Qiagen RNeasy kit. 1ug of RNA was added to a mix of 1 unit DNaseI and 1x DNaseI buffer (both Invitrogen) to final volume of 10 µl. This was incubated at 65°C in 2.27 mM EDTA (Invitrogen) for 10 minutes then 70°C for 10 minutes with 0.5 µg oligo(dT) (Invitrogen). To this a mix of 1x First Strand buffer (Invitrogen) and 0.5mM dNTPs (Fermentas) was added and then two further incubation steps; 42°C for 60 minutes followed by 7°C for 15 minutes. The cDNA amplifications were performed using the Power SYBR green PCR master mix (Applied Biosystems) using 5 µl of cDNA (diluted 1:100) and 0.8 µM primers in 13 µl reactions. The ViiA 7 Real Time PCR system (Applied Biosystems) was used for amplification. Each genotype was analysed in 3 biological replicates and 3 technical replicates. The efficiency for each set of primer was tested by setting up a serial dilution of a single cDNA sample and efficiency was calculated using the Pfaffl method (Pfaffl, 2001). Only primers with an amplification efficiency of over 90% were used. The Pfaffl method was also employed to determine the relative expression levels of the genes and were normalised against values obtained for the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH). All primers used are listed in Table 2.1

2. 5. 7. Cloning candidate genes

Genomic fragments of the genes of interest were ligated into the pGEM plasmid using the pGEM-T Easy Vector System I (Promega) according to the associated protocol. This vector was cloned with *E. coli* and then isolated using the Qiagen Miniprep plasmid purification kit, with no amendments to manufacturer's protocol. The candidate gene genomic fragment was excised using the NotI restriction enzyme (New England Biolabs) and separated using gel electrophoresis. The fragment was extracted from the gel using the Invitrogen Purelink Quick Gel Extraction kit without any amendments to the manufacturer's protocol. The resultant DNA was sequenced and clones without sequence polymorphisms were ligated into the vector pMLBART that had

undergone NotI digestion de-phosphorylation with Antarctic phosphatase (New England Biolabs) using the manufacturer's protocol. The ligation mix consisted of 2µl plasmid (20ng), 6µlfragment DNA (50ng), 1µl 1 x T4 DNA ligase buffer (Promega) and 1µl T4 Ligase (Promega). This vector was then used to transform *A. tumefaciens* before transforming *A. thaliana.*

2. 6. Phenotype analysis

2. 6. 1. Silhouettes

Phenotyping was always performed after the plants had fully flowered, unless otherwise stated. For leaf shape analysis silhouttes of all rosette leaves were captured. Fully developed rosette leaves were adhered to white paper using adhesive spray (Spraymount - 3M) and digitally scanned (Canon scanner). Photoshop (Adobe) was used to remove any bits of dirt of shadow lines from the images that would interrupt the taking of measurements and image analysis software ImageJ (http://imagej.nih.gov/ij/) was used to measure the various leaf shape traits.

2. 6. 2. Cell count analysis

The same plants used to characterise QTL effect on leaf shape were used to analyse cell size. Prior to sticking the leaves to paper for scanning, the adaxial surface of individual leaflets were pressed on to molten 2% agarose and removed once the agarose had set leaving a detailed imprint of the leaf surface. This was photographed under a dissection microscope (Olympus). Five photographs were taken per leaflet and these were constrained to the middle portion of the leaf with the margins and any areas near leaf damage avoided. The resultant images were analysed with Photoshop (Adobe) and ImageJ (http://imagej.nih.gov/ij/). It is

impractical to measure the area of each cell captured so our methodology involved isolating segments of complete cells and determining the area of these segments. Cell size was calculated by division of segment area over cell number in the segment and from this cell number for the entire leaflet was determined from the area of that particular leaflet

2. 7. Statistical methods

For comparison of alternative genotypes, Student t tests were used to compare traits at each leaf node along the heteroblastic series. To assess significant phenotypic correlations between traits in the RIL population (Figure 3.8), Pearson's r correlation test was used. These were performed using either Excel (Microsoft office) or Minitab 17 (http://www.minitab.com/).

2. 8. QTL mapping

2. 8. 1. Linkage map construction

To establish a genetic map of the RIL population, we generated 199 markers by identifying sequence polymorphisms between the two parental strains segregating in the RIL population. The sequence polymorphisms from which markers were designed were identified using existing EST sequence libraries (Hay et al., 2014). Polymorphism selection was in part guided by the known relationships between blocks of the *Arabidopsis* genome and the *Cardamine* genome (as in Schranz et al. 2006). This meant that all the markers were located within predicted protein coding regions genes, which facilitates comparisons of synteny between the resulting map of the RIL population and the Arabidopsis genome. This selection method allowed us to select single nucleotide polymorphisms (SNPs) that would most likely cover all parts of the genome

allowing for any QTL that were detected to be identified effectively. The resulting panel of SNPs were used to develop Sequenom assays for genotyping which were performed by the Wellcome trust, Oxford (www.wellcome.ac.uk). Initially the RIL population was genotyped with SNPs designed for an Oxford x Washington RIL population (see Hay et al., 2014) that were specifically selected for being polymorphic for Oxford/Azores also. The resulting map did not have full coverage and consisted of 12 linkage groups and 5 singletons. A second set of 103 SNPs were specifically selected to fill in the gaps, again by using synteny between *A. thaliana* and the ancestral Brassicaceae genome. The additional SNPs were enough to complete the linkage map with the expected eight linkage groups (Hay et al., 2014). The final map had fewer markers (173) because those with poor genotyping results or where the mapping indicated technical problems were excluded from the analysis.

Linkage analysis and map construction were completed using JoinMap version 4 (Van Ooijen, 2006) with the Kosambi mapping function (Kosambi, 1943) using allelic information from all 149 genotyped RILs. A minimum logarithm of odds (LOD) score of 3.5 or higher was used to establish the linkage map.

2. 8. 2. Single trait QTL mapping

For single trait QTL mapping the statistical package *R* (R Development Core Team, 2011) with the add on program *R/qtl* (Broman et al. 2003; www.rqtl.org) was used. Significance threshold values for Log of odds (LOD) were derived according to the method outlined by Churchill and Doerge (1994); 1000 permutations of the original data set for each trait were used to determine a threshold value to a genome wide type I error rates of 5% and 1%. At first, each trait was mapped with simple marker regression with interval mapping. And then composite interval mapping with those markers that were identified as significantly associated with the trait were used as co-factors. Additive genetic effects and the percentage variance explained by each

QTL were estimated using the composite interval mapping analysis. This analysis was applied to leaflet number on leaves 2 to 10 and on rosette leaf number.

2. 8. 3. Multi trait QTL mapping

The results of the single trait QTL mapping suggested that many of the traits shared QTL which proves genetic correlation and therefore the use of multi trait QTL mapping was deemed appropriate and valid. Multi-trait QTL analysis is based on mixed model methodology and here it was applied to leaflet number and rosette leaf number in the Oxford x Azores RIL population. The analysis was performed using the software Genstat 12th edition (Payne et al, 2009). The mean trait values per RIL were used for leaflet numbers on leaves 2 to 10 and rosette leaf number. The method involved performing a single genome-wide QTL mapping scan for all 10 traits simultaneously. The QTL effects were considered fixed effects while the genetic covariances were modelled explicitly as random effects. Genetic predictors were calculated that were no more than 2 cM apart in distance. The threshold of significance was calculated (P = 0.05) using the method described by Li and Ji (2005). After an initial simple interval mapping scan multiple rounds of composite interval mapping followed in which cofactors were either added or removed until no additional improvements could be made. From each of the detected QTL the most significant genetic predictors were used for the final QTL model to determine an estimate of allelic effects on each trait and the proportion of variances explained by them.

Chapter 3:
Analysis and QTL mapping of intraspecific natural variation in *Cardamine hirsuta* leaf shape

3. 1. Chapter Introduction

Living organisms exhibit striking variations in their morphology. A major goal in biology is to understand how biological forms develop and to discover the genetic basis for their diversity. Plant leaves make a very attractive model with which to address these goals. They are important organs for plant function and they show considerable variation at different levels; between species, within species and at different stages of plant development. The majority of studies into leaf development have focused on inter species differences, this study focuses on intra species differences in leaf morphology.

Our model, *Cardamine hirsuta,* is a plant species that displays intraspecific variation in leaf morphology. *C. hirsuta* is a cosmopolitan weed that thrives on sandy, rocky or disturbed ground. It is thought to be a native to Europe (Lihova et al. 2006), but can be found worldwide thanks to both natural forces and anthropogenic dispersal. This global spread has created separate strains with clear geographic structuring of genetic variation (Hay et al. 2014). The aim of the present study is to characterize the morphological variation in leaf shape within *C. hirsuta* and to uncover the genetic basis of this variation.

Cardamine hirsuta has only recently been developed into a model system for studying morphological diversification (Hay et al, 2014). As a consequence the extent of naturally occurring genetic variation for its key morphological characters within the species, in particular leaf geometry, has not been determined yet. The available collection of accessions in the lab consisted of samples from a broad geographic range. By analysing in detail a small sample of six accessions from locations widely distributed across this range I aim to describe some of the natural variation that has evolved in leaf shape during the dispersion of *C. hirsuta* strains globally.

Previous studies of the genetics of leaf shape in *C. hirsuta* in the context of comparative development have adopted an accession from Oxford, United Kingdom as a common reference (Hay and Tsiantis, 2006; Barkoulas *et al.*, 2008; Canales *et al*., 2010; Vlad *et al*., 2014). Given the considerable genetic and molecular resources previous work have made available, it stands to

reason that this accession is adopted as a reference for the ensuing work on natural variation also. Particular attention is therefore given to contrasting leaf morphology of the analysed accessions against that of the Oxford accession.

My results revealed that there was a great deal of variation amongst the different strains. However despite this, within all strains there was a predictable pattern of leaf shape change at progressive nodes as the shoot developed. This developmentary progression provides the context in which we should be analysing variation in leaf traits. It is clear that the rate at which development progresses is highly variable and that any investigation into variation in leaf traits is highly likely to uncover components that influence or are influenced by this transition.

In advance of attempts to discover the genomic regions that contribute to leaf variation using an F6 recombinant inbred line (RIL) experimental mapping population derived from a cross between the Oxford and Azores strains. There is further focus on the degree of variation between these strains. Analysis of the RIL population allows for predictions to be made about the nature of the genetic control behind leaf shape traits. Importantly, many of the traits are highly heritable indicating they have a strong genetic basis. Significant correlations between traits suggest determinants and phenotype values beyond the range of the parental phenotype indicate transgressive segregation. Thus we can predict that the leaf traits are controlled by multiple common loci.

To discover the identity of these loci, Quantitative trait loci (QTL) mapping was employed. As the results here demonstrate, the traits associated with leaf shape are quantitative in nature; there is continuous variation in a segregating population. The genetic basis of such traits is characterised by multiple loci, mostly of small or intermediate effect referred to as QTL. QTL mapping is the identification of statistically significant associations between phenotypic variation and specific alleles at, and in between marker loci. QTL are located at the position in the genome where these significant associations are discovered. The power to detect QTL and the resolution of QTL location is provided by the mixing of parental genomes that occurs via recombination in

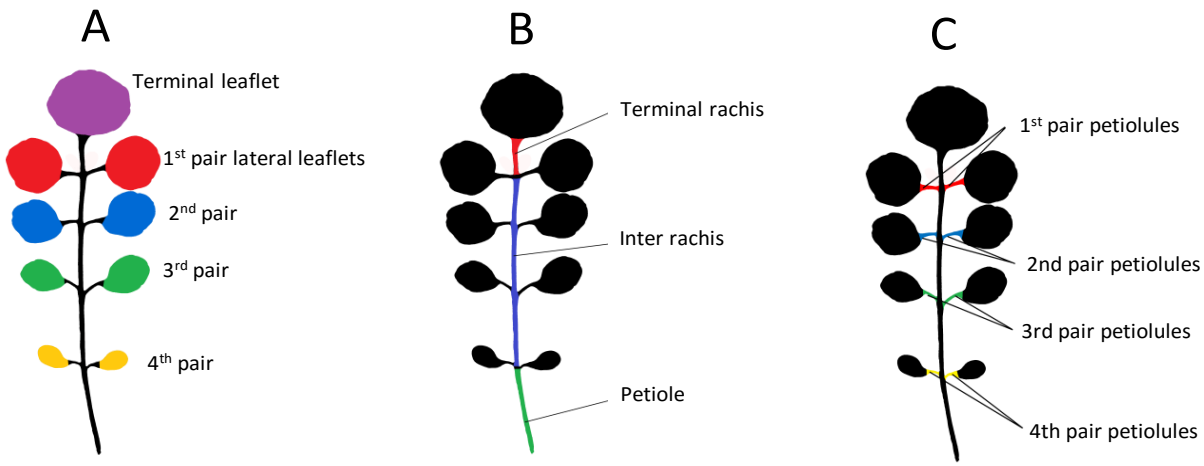
meiosis during the construction of an experimental mapping population (in this study an F6 RIL population). Each member of the population is genotyped at markers located across the genome, and then the population is scored for the quantitative trait of interest. Significant associations are then sought between markers and the quantitative trait using specialised software (In this study R/QTL and Genstat were used). Recombination frequencies between linked markers allow for the construction of a genetic linkage map that describes the location of each marker on the genome making it possible to identify the location of QTL.

In this study QTL mapping is performed using two different statistical methodologies. Initially QTL were identified using single trait analysis which examines association between phenotype and genetic intervals one trait at a time. A more dynamic and realistic model for leaf shape traits was achieved by multi trait QTL mapping (van Eeuwijk et al., 2010) which showed how different traits were influenced by common QTL whose effect was variable across the heteroblastic series. Following these analyses I was able to detect six QTL responsible for variation in leaflet number and rosette leaf number.

3. 1. 1. Anatomy of the *Cardamine hirsuta* leaf

In order to analyse the intraspecific diversification of leaf shape in *C. hirsuta* it is important to clarify the features for which I am measuring variation. *C. hirsuta* bears compound leaves which have a lamina dissected into subunits termed leaflets. The distally situated leaflet at the end of the rachis is the terminal leaflet and the leaflets produced along the rachis are the lateral leaflets (Figure 3.1 A). Lateral leaflets are usually organised in pairs and for the sake of comparison the most distal, and closest to the terminal leaflet, lateral leaflets are termed the 1st pair, the next pair are termed the 2nd pair and so on (Figure 3.1 A). The rachis is then split into the terminal rachis, the inter rachis and the petiole (Figure 3.1 B). The terminal rachis extends from the terminal leaflet to the 1st pair of leaflets, the inter rachis describes the length of rachis

Figure 3.1 Anatomy of the *C. hirsuta* leaf. The terminal leaflet and the lateral leaflets are shown in A, the different parts of the rachis are labelled in B and the petiolules in C There is no scale to this figure as it is for illustrative purposes only.


A
Terminal leaflet
1st pair lateral leaflets
2nd pair
3rd pair
4th pair
B
Terminal rachis
Inter rachis
Petiole
C
1st pair petiolules
2nd pair petiolules
3rd pair petiolules
4th pair petiolules

from the upper most to the lower most lateral leaflet and the petiole is the rachis from the lowermost leaflet to the shoot (Figure 3.1 B). Finally the petiolules attach the lateral leaflets to the rachis and these are termed 1st, second…*etc*. in the same way the lateral leaflets are (Figure 3.1 C).

3. 2 Quantification of heteroblasty in the reference strain Oxford

Most of the developmental and genetic work on the evolution of *C. hirsuta* leaf shape has used the strain form Oxford as the wild type control (Hay and Tsiantis, 2006; Barkoulas *et al.*, 2008; Canales *et al*., 2010; Vlad *et al*., 2014) therefore I have used the same strain in the present study to compare with other strains. Figure 3.2 shows the leaf shape of all leaves produced prior to bolting in the main basal part of the plant, these leaves are referred to as rosette leaves. Bolting (the rapid elongation of the shoot) marks the plant's developmental transition from vegetative to reproductive growth and these accompany marked changes in the character of the shoot such as the production of flowers instead of leaves (Poethig, 2013). The leaves produced away from the main rosette structure (cauline leaves) are not examined in this study, they have a very different shape to the rosette leaves and their development is highly influenced by the switch to reproductive growth therefore including them in this study could complicate the exploration of the genetic basis underlying leaf shape variation during vegetative growth. The results below are based on a sample of 10 observations.

Number of rosette leaves - The mean number of leaves produced prior to bolting by the Oxford strain is 11.2 (SEM: +/- 0.2).

Leaf dissection – The number of leaflets on a leaf was used as a measure of dissection. The mean number of leaflets produced increases at successive nodal positions; the 1st leaf only ever has a single terminal leaflet and the number of lateral leaflets increases steadily by 1-2 leaflets up until the final rosette leaf (Figure 3.3 A).

Figure 3.2 Silhouettes of rosette leaves of *Cardamine hirsuta* - Oxford strain, produced prior to bolting. Scale bar, 1cm. This is a typical example.

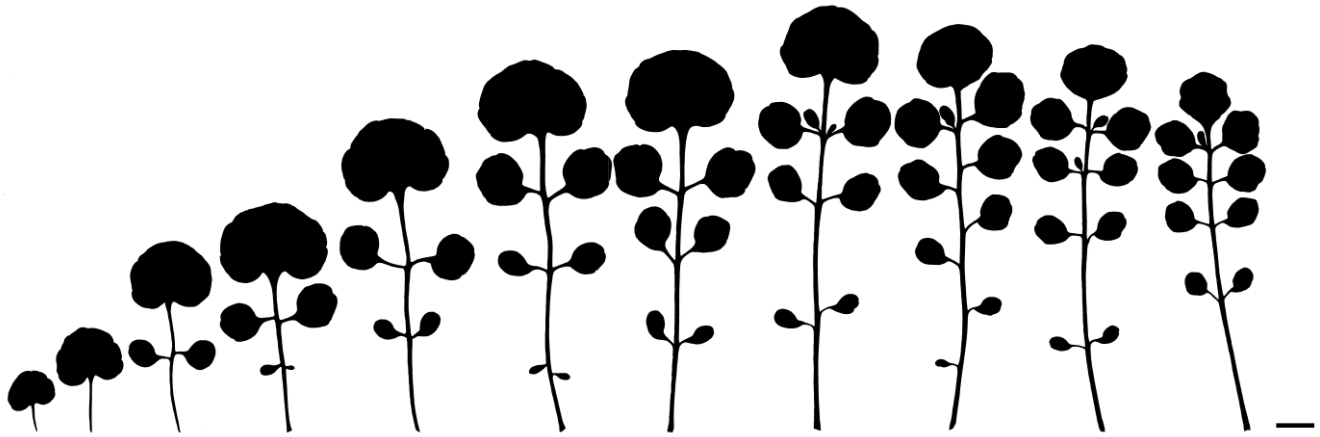
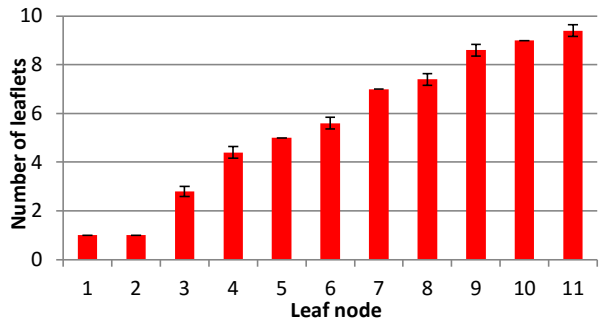
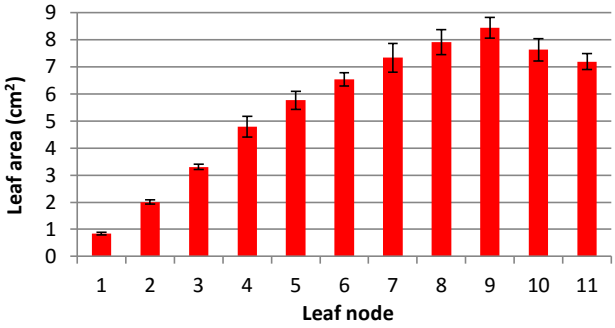


Figure 3. 3. Heteroblastic variation in the leaflets of the Oxford strain of *C. hirsuta.* (A) Mean number of leaflets per leaf, (B) Mean leaf area (C) mean terminal leaflet area, (D) mean area of leaflets. All error bars represent standard error of the mean.

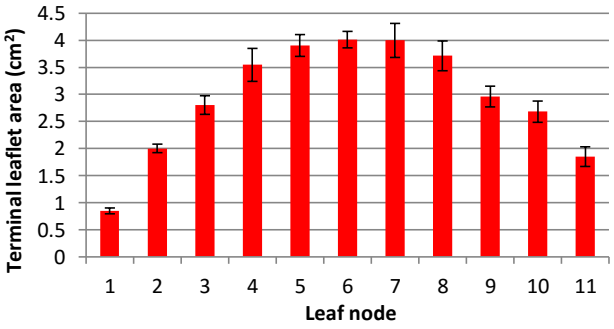
A


Number of leaflets
0
2
4
6
8
10
1 2 3 4 5 6 7 8 9 10 11
Leaf node

B


Leaf area (cm²)
0
1
2
3
4
5
6
7
8
9
1 2 3 4 5 6 7 8 9 10 11
Leaf node

C


Terminal leaflet area (cm²)
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
1 2 3 4 5 6 7 8 9 10 11
Leaf node

D

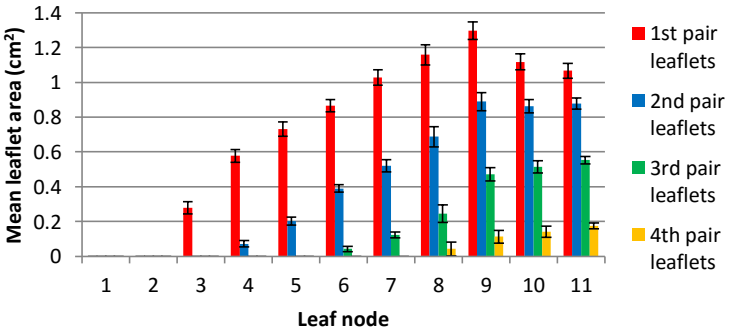
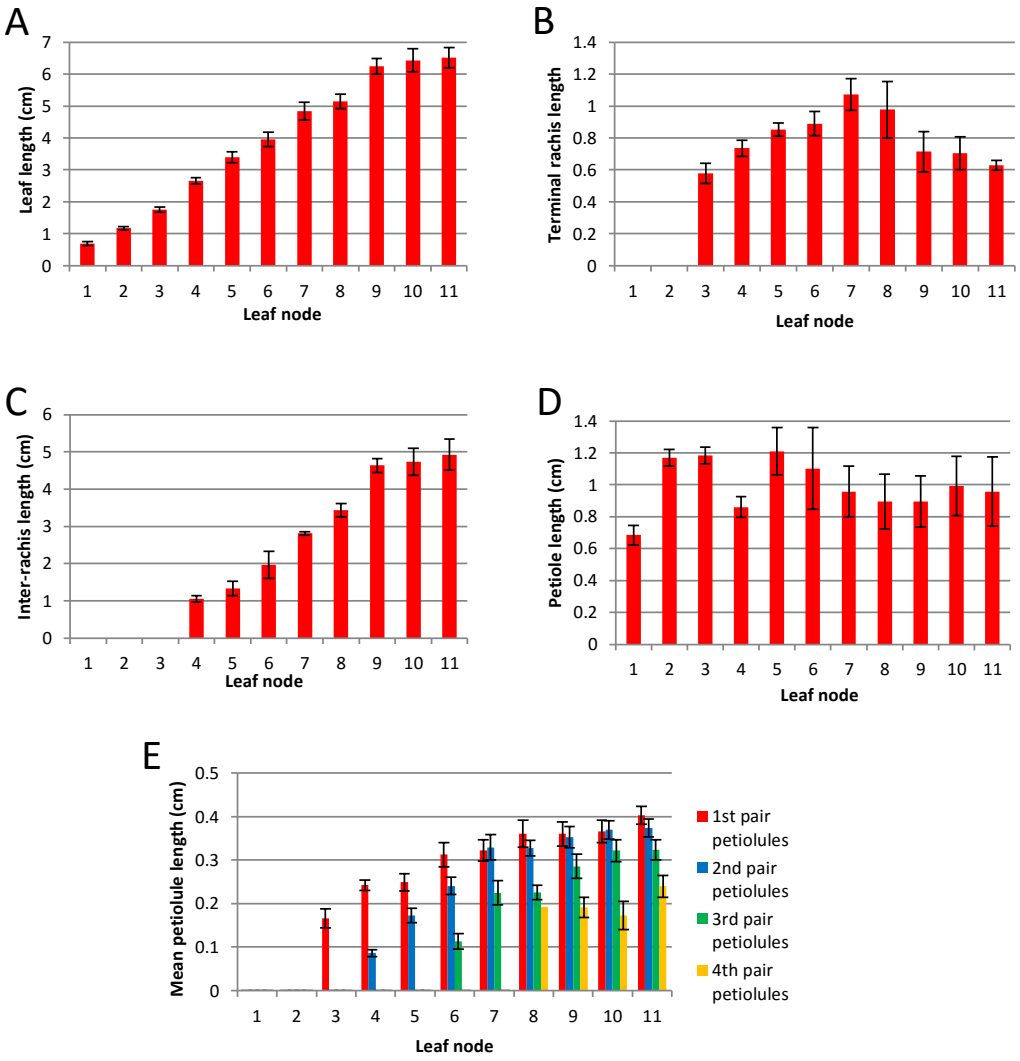

Mean leaflet area (cm²)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1 2 3 4 5 6 7 8 9 10 11
Leaf node
1st pair leaflets
2nd pair leaflets
3rd pair leaflets
4th pair leaflets

Figure 3. 4. Heteroblastic variation of rachis dimensions in the Oxford strain of *C. hirsuta.*

Mean (A) leaf length, (B) terminal rachis length, (C) Inter-rachis length, (D) petiole length and (E) petiolule length. All error bars represent standard error of the mean.


A
Leaf length (cm)
0
1
2
3
4
5
6
7
1 2 3 4 5 6 7 8 9 10 11
Leaf node
B
Terminal rachis length
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1 2 3 4 5 6 7 8 9 10 11
Leaf node
C
Inter-rachis length (cm)
0
1
2
3
4
5
6
1 2 3 4 5 6 7 8 9 10 11
Leaf node
D
Petiole length (cm)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1 2 3 4 5 6 7 8 9 10 11
Leaf node
E
Mean petiolule length (cm)
0
0.1
0.2
0.3
0.4
0.5
1 2 3 4 5 6 7 8 9 10 11
Leaf node
1st pair petiolules
2nd pair petiolules
3rd pair petiolules
4th pair petiolules

Leaf area –Total leaf area increases steadily from 0.84cm^2 (SEM: +/- 0.05) at leaf 1 to a peak of 7.91cm^2 (SEM: +/- 0.46) at leaf nine. Subsequent leaves decrease in size to a mean area of 7.20cm^2 at the final leaf (SEM: +/- 0.29, Figure 3.3 B).

Leaflet areas - Clearly leaf area is likely to increase as leaflets increase in number, but the size of the leaflets are also changing during ontogeny (Figure 3.2 C-D). Terminal leaflet area increases steadily from 0.85cm^2 (SEM: +/- 0.05) at the 1st leaf to just under 4cm^2 at leaves 5, 6 and 7 and then decreases steadily hereafter until the final leaf before bolting (Figure 3.3 C). With regards to lateral leaflet area there is variation both within leaves and between leaves at different nodes (Figure 3.3 D). The largest leaflets are always the closest pair to the terminal leaflet and furthest away from the main shoot (1st pair leaflets, Figure 3.3 D). The leaflets decrease in area towards the shoot, although these differences become smaller in later leaves (Figure 3.3 D). There is also variation in lateral leaflet area between leaves; leaflet area with all leaflet pairs increases in successive leaves reaching their peak size at leaf 9, in the subsequent leaves the area of the 1st pair of leaflets decreases but the area of the second, 3rd and 4th pair of leaflets does not vary significantly after the peak size is reached (Figure 3.3 D). Some leaves on some plants possessed a fifth pair of lateral leaflets, but this occurred too rarely to provide sufficient power for analysis.

Leaf length - The length of the rachis also increases as the plant grows; starting at 0.68cm (SEM: +/- 0.06) and increases to 6.51cm (SEM: +/- 0 .31, Figure 3.4 A). The successive increase in rachis length is due to expansion of the terminal rachis, the inter-rachis and the petiole (Figure 3.4 B-D). Leaf length increase in leaves 3 -7 is due to expansion in both the inter-rachis and terminal rachis (Figure 3.4 B-C) whereas leaf extension between leaves 8 and 9 is from expansion in the inter-rachis alone (Figure 3.4 C) because following leaf 7 the terminal rachis length shrinks (Figure 3.4 B). Leaf length does not increase significantly following leaf nine. While the terminal rachis length decreases, total leaf length is maintained by expansion within the inter-rachis. Petiole length increases in the first two leaves but then remains steady with some fluctuation (Figure 3.4 D).

Petiolule length - Attaching each leaflet to the main rachis of the leaf are the petiolules. These determine the width of the leaves and like the lateral leaflets themselves there is variation within leaves and between leaves. The leaflets closest to the terminal leaflet and furthest away from the shoot have the longest petiolules decreasing away from the terminal leaflet; however these differences decrease in later developing leaves (Figure 3.4 E). Petiolule length for the 1st, 2nd and 3rd pair of leaflets increases throughout development until bolting, but petiolule length on the 4th pair of leaflets show less predictable variation in the four leaves they appear on (Figure 3.4 E).

3. 3. Heteroblastic change

The results above demonstrate there is a lot of quantitative change in the shape and size of leaves during the development of the Oxford strain of *C. hirsuta* (Figures 3.2, 3.3 and 3.4). This asymmetry in form between organs that develop at different stages of growth is a widespread phenomenon amongst plants and has been extensively studied since the late 19th century in a multitude of different plant groups (reviewed in Poethig, 2013).

Plants are made up of repeating units and each new unit is initiated within a changing internal (and external) environment that is contingent on the previous developmental history of the plant which in turn is determined by the genetic architecture of the plant. As a result throughout postembryonic development the plant shoot undergoes many quantitative and qualitative changes that are associated with many life-history traits. Of these the most obvious and best understood is the transition a plant undergoes when it switches from vegetative to reproductive development. This is marked by the production of new organs responsible for sexual reproduction such as flowers or cones (Amasino, 2010; Andres and Coupland, 2012; Huijser & Schmid, 2011; Wilkie et al., 2008)

In the present study I am analysing the change that precedes the onset of reproductive growth which is vegetative phase change. The changes are associated with preparing the shoot to respond to the stimuli that induce reproductive development (Poethig 2013). Hildebrand and Goebel in the 19th century (reviewed in Poethig, 2013) were the first to recognise that shoot development could be separated into different phases on the basis of vegetative traits such as leaf shape, leaf size, phyllotaxy, internode length, anthocyanin pigmentation, rooting ability or wood structure. The term "heteroblastic" was coined by Goebel in 1900 to describe those plants that had substantial and abrupt changes in these characters.

Heteroblasty has since been used to encompass any morphological differences between individual metamers, irrespective of the nature and extent of the variation (Zotz et al. 2011; Jones 1999). The abrupt changes that Goebel was referring to in his definition are the points at which vegetative growth switches between juvenile and adult phases. However these terms are now often used to distinguish between vegetative and reproductive growth (Jones 1999). There are in fact juvenile and adult phases within both vegetative development and reproductive development (Poethig 1990). Juvenile and adult phases often have distinguishing morphological markers such as leaf shape and trichome distribution (Poethig, 1997). However based on the above data on the Oxford strain of *C. hirsuta* there are no features that can be used to clearly differentiate leaves developed in the juvenile or adult growth phases. Jones (1999) argues that the change in vegetative phases should be defined by the point at which the shoot gains reproductive competence, but since no attempts were made to determine when this occurs in any of the strains of *C. hirsuta* in this study we cannot define different phases of growth with this criterion either. To avoid the danger of misappropriating different leaf shapes to juvenile and adult phases of growth the terms are avoided in this study and leaves will be described as being part of a heteroblastic series rather than assigning them to any particular phase of vegetative growth. Any attempt to dissect the genetic architecture underlying leaf shape variation needs to consider variation within the dynamic model of heteroblastic development. A study that isolates variation

at one node will fail to capture the underlying process determining leaf shape; thus this study will consider leaf variation within the context of heteroblastic change.

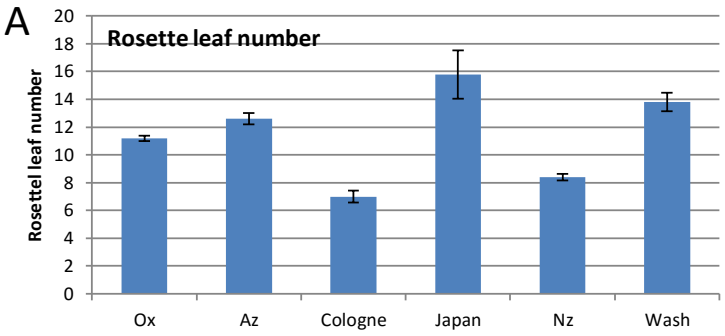
3. 4. Intraspecific variation in the heteroblastic progression of rosette leaf morphology

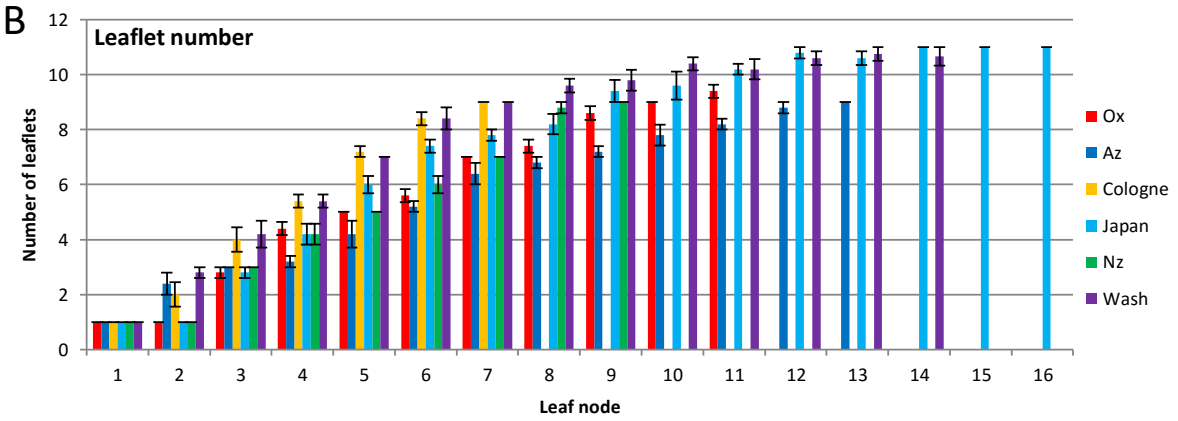
C. hirsuta is thought to be native to Europe but has spread worldwide (Lihova et al. 2006). Many different strains from around the world have been collected by members of the Tsiantis group or provided by collaborators. These have been verified as being of the same species through morphology and sequence analyses and have been propagated for several generations. *C. hirsuta* has a dominant selfing habit so it is assumed that the majority of the positions in the genome are fixed and very little allelic segregation is occurring when seeds are propagated. Additionally, once collected the wild ecotypes were selfed for at least six generations and propagated through single seed descent to ensure that no heterozygosity remained. It is therefore legitimate to analyse the leaf shape of multiple individuals of the same strain with the assumption they are genetically identical. In order to explore what variation might exist amongst different strains of *C. hirsuta* I have selected geographically divergent strains to analyse and compare with the reference strain - Oxford. The aim of this study is to explore the phenotypic variation that exists amongst a few strains of *C. hirsuta*. I made no attempt to extrapolate any trends or meaning from these data, the data is only used to illustrate the intraspecific variation that exists within *C. hirsuta*.

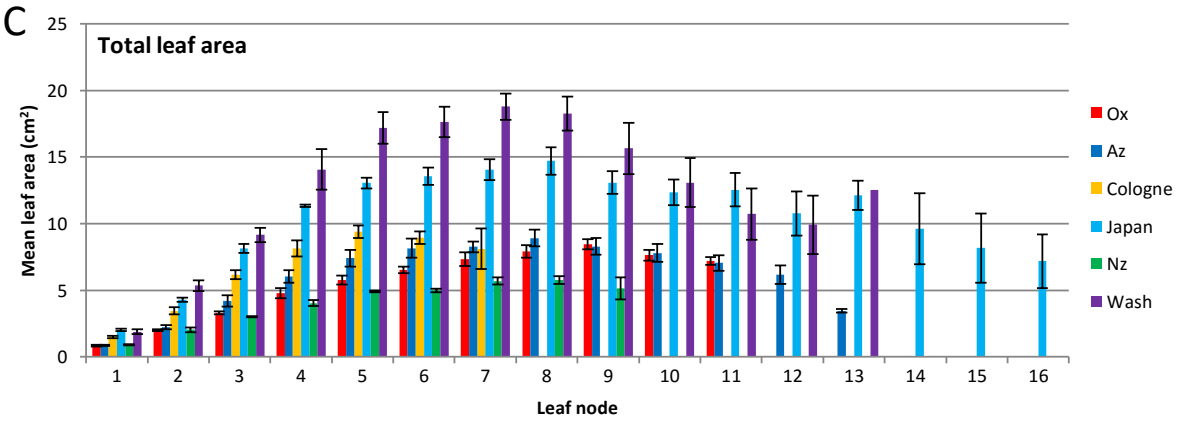
3. 4. 1. Intraspecific variation – rosette leaf production

One of the most notable aspects of variation between the strains studied was the number of rosette leaves produced prior to bolting (Figure 3.5 A). Strains such as those from Cologne and New Zealand produced fewer leaves than the Oxford Strain before bolting with a mean of 7 (SEM: +/- 0.44) and 8.4 (SEM: +/- 0.24) leaves respectively. Conversely the Japanese strain

Figure 3.5 Intraspecific variation in leaf shape along the heteroblastic series. *C. hirsuta* strains were measured for rosette leaf number (A), leaflet number (B), leaf area (C), terminal leaflet area (D), 1st to 4th pair of lateral leaflets (E-H), leaf length (I),terminal rachis length (J), lnter rachis length (K), petiole length (L) and the 1st to 4th pair of petiolules (M-P). All error bars represent standard error of the mean.


A
Rosette leaf number
Rosettel leaf number
20
18
16
14
12
10
8
6
4
2
0
Ox
Az
Cologne
Japan
Nz
Wash


B
Leaflet number
Number of leaflets
12
10
8
6
4
2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Leaf node
Ox
Az
Cologne
Japan
Nz
Wash


C
Total leaf area
Mean leaf area (cm²)
25
20
15
10
5
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Leaf node
Ox
Az
Cologne
Japan
Nz
Wash

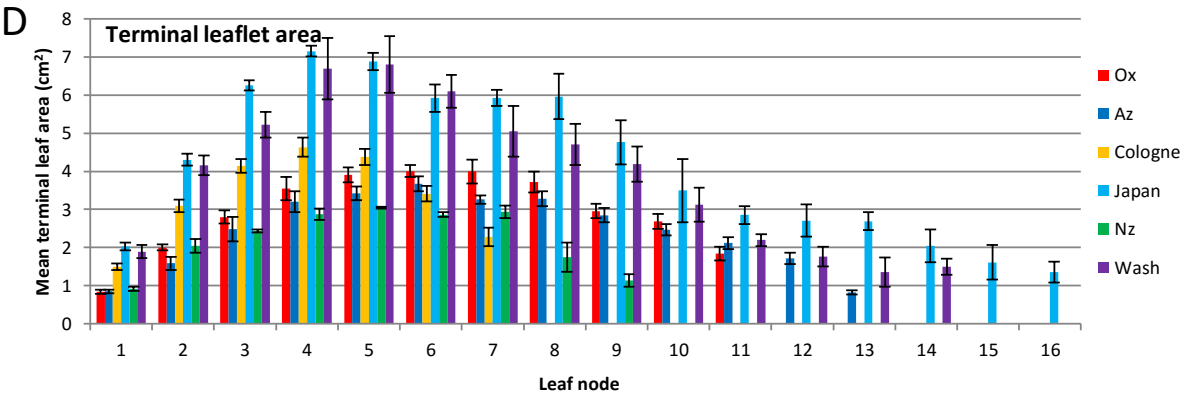

D
Terminal leaflet area
Mean terminal leaf area (cm²)
8
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Leaf node
Ox
Az
Cologne
Japan
Nz
Wash

Figure 3.5 continued

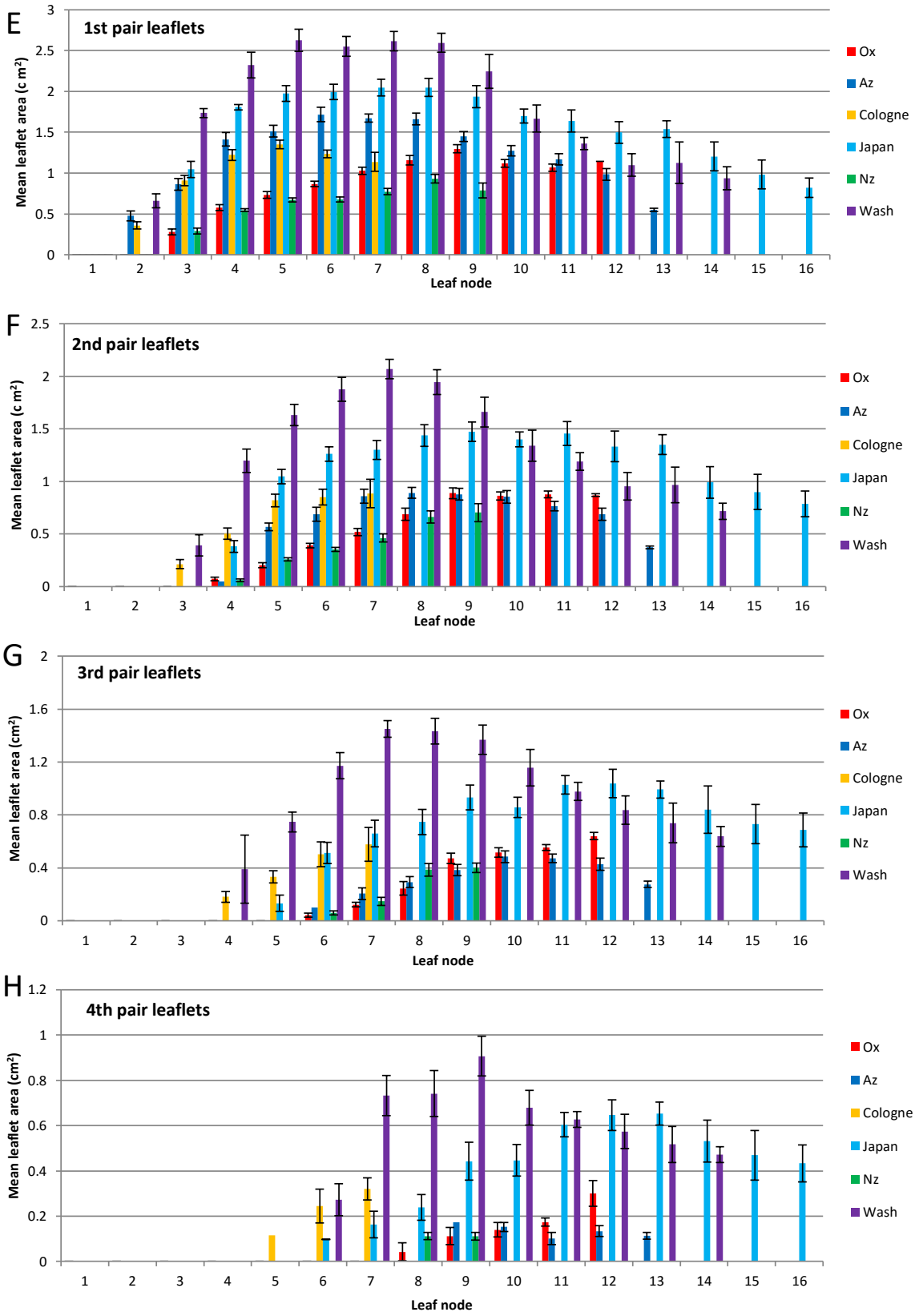


Figure 3.5 continued

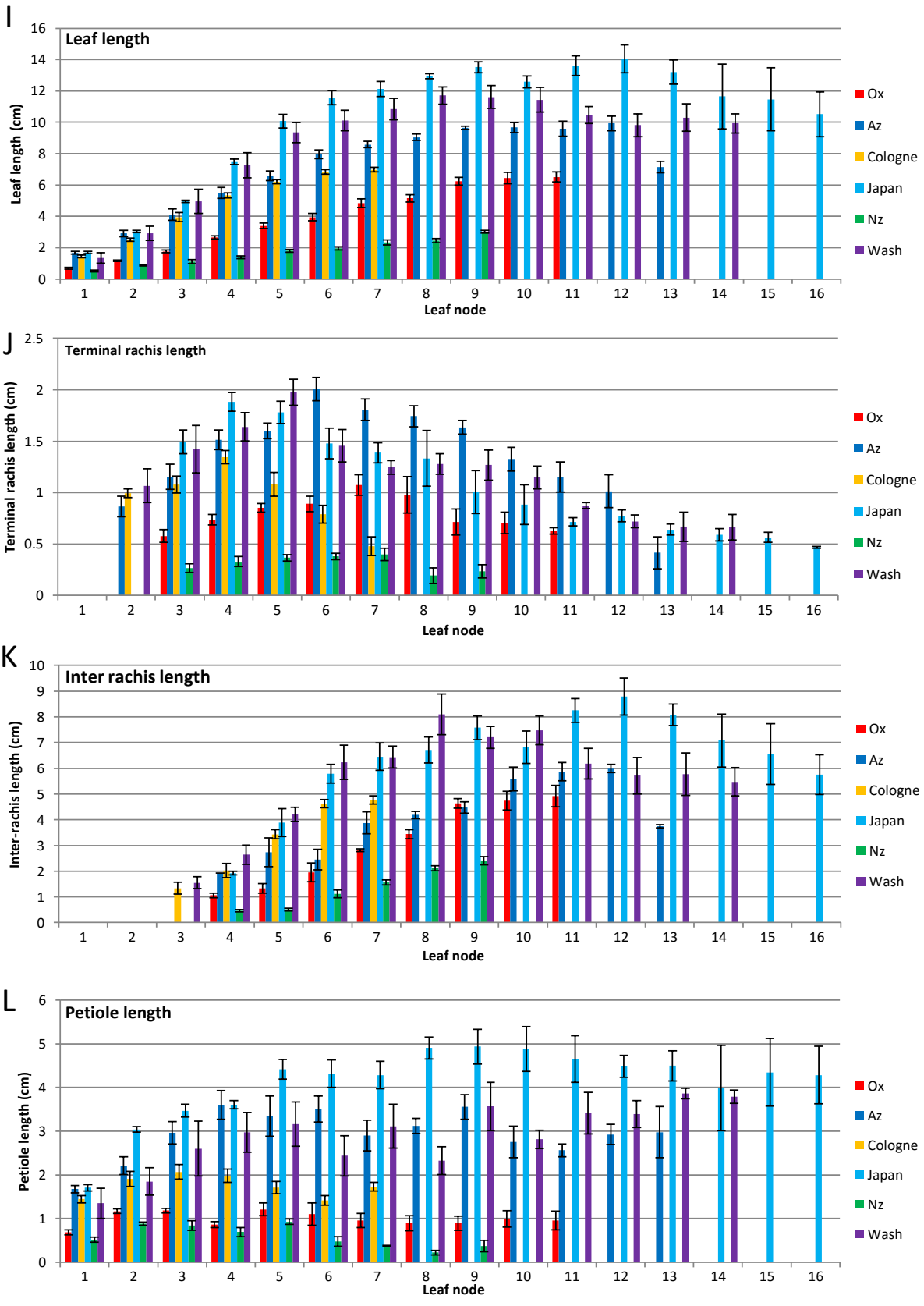
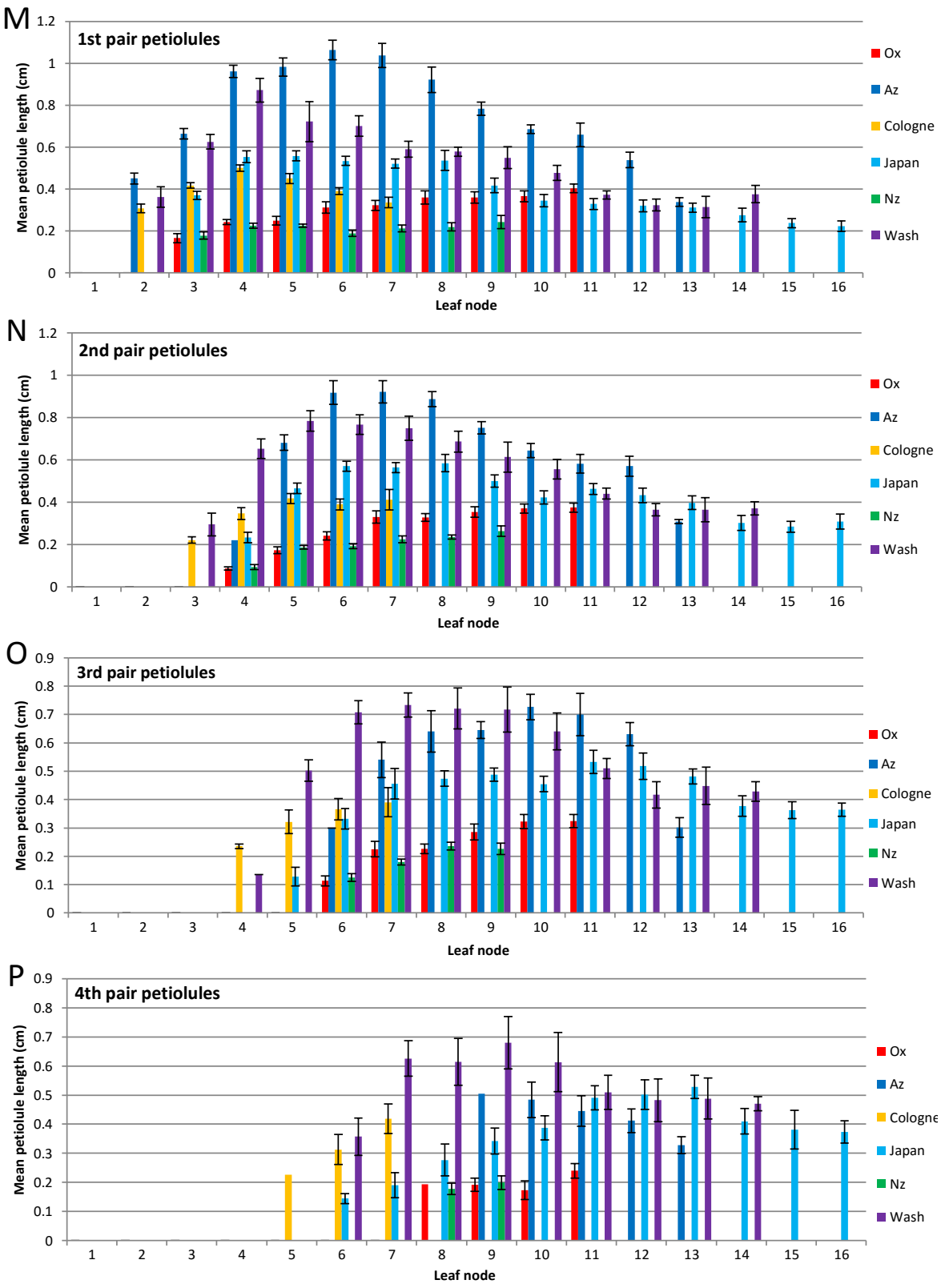


Figure 3.5 continued



produces more rosette leaves at 15.8 (SEM: +/- 1.74) as does the strain from Washington state, USA at 13.8 (SEM: +/- 0.66). In *A. thaliana* the number of rosette leaves produced by the plant and the time taken to flower are highly correlated (Salome et al., 2011), therefore it is reasonable the differences observed here reflect variation amongst accessions in how long vegetative development proceeds before reproductive development is initiated.

3. 4. 2. Intraspecific variation – Leaf shape

Leaf dissection - There is variation in the number of leaflets at all nodes except at the 1st leaf node where there is only ever one leaflet (Figure 3.5 B). At any node the largest difference in leaflet number exists at Leaf node six, where there is a range of 3.2 leaflets between Azores that has a mean of 5.2 leaflets (SEM: 0.2) to Wash and Cologne that both have a mean of 8.4 leaflets (SEM: 0.4 and 0.2 respectively, Figure 3.5 B). In all the strains flowering is never achieved until the plant produces a leaf with at least nine leaflets. This may indicate that the shoot has reached its adult vegetative phase.

Leaf area - In terms of total leaf area the largest difference is captured at leaf node seven between New Zealand (5.7cm^2, SEM: 0.3) and Wash (18.8cm^2, SEM: 0.9) with a range of 13.1cm^2 (Figure 3.5 C)

Leaflet areas - When the area of individual leaflets are looked at independently of the entire leaf; the largest variation for terminal leaflet area exists at leaf node 4 between New Zealand and Japan (2.9cm^2, SEM: 0.15 and 7.1cm^2, SEM: 0.14 respectively) (Figure 3.5 D); at leaf node 5 for the 1st pair of leaflets between New Zealand and Wash (0.7cm^2, SEM: 0.02 and 2.6cm^2, SEM: 0.16 respectively) (Figure 3.5 E); at leaf node seven for the second pair of leaflets between New Zealand and Wash(0.5cm^2, SEM: 0.04 and 2.1cm^2, SEM: 0.09 respectively) (Figure 3.5 F); at leaf node seven for the 3rd pair of leaflets between Oxford and Wash (0.12cm^2, SEM: 0.02 and 1.45cm^2, SEM: 0.06 respectively) (Figure 3.5 G); and, leaf node nine for the 4th pair of leaflets

between New Zealand and Wash (0.11cm^2, SEM: 0.02 and 0.9cm^2, SEM: 0.09 respectively) (Figure 3.5 H).

Leaf length - The largest difference in leaf length exists at leaf node 9; Japan has the longest leaves at 13.5cm (SEM: 0.34) and New Zealand has the shortest leaves at 3.0cm (SEM: 0.1) (Figure 3.5 I). The largest difference in terminal rachis length exists at leaf node six between the strains Azores (2.0cm, SEM: 0.11) and New Zealand (0.4cm, SEM: 0.03) (Figure 3.5 J). Inter rachis is most variable at leaf node eight where Wash has the longest (8.1cm, SEM: 0.79) and New Zealand the smallest (2.1cm, SEM: 0.1, Figure 3.5 K). Petiole length is also most variable at leaf node eight where Japan has the longest (4.9cm, SEM: 0.79) and New Zealand the smallest (0.2cm, SEM: 0.05) (Figure 3.5 L).

Petiolule lengths - Petiolule length determines how wide the leaves are and this is highly variable between different strains. The 1st pair of petiolules are most variable at leaf node 6, where Azores has the longest (1.06cm, SEM: 0.05) and New Zealand the smallest (0.19cm, SEM: 0.01) (Figure 3.5 M). Following on from this; the second pair of petiolules are most variable at leaf node 6 where Azores has the longest petioles and New Zealand, the shortest (0.92cm, SEM: 0.06 and 0.19cm, SEM: 0.01 respectively) (Figure 3.5 N); the 3rd pair of petiolules are most variable al leaf node 6- Wash has the longest and Oxford the shortest (0.71cm^2, SEM: 0.04 and 0.11cm^2, SEM: 0.02 respectively) (Figure 3.5 O); and the 4th pair of petiolules are most variable at leaf node 9 – Wash has the longest and Oxford the shortest petiolules (0.68cm^2, SEM: 0.09 and 0.19cm^2, SEM: 0.02 respectively) (Figure 3.5 P).

The results here demonstrate that despite the variation in the number of rosette leaves and in the length of the vegetative phase all strains share a similar pattern of growth along the heteroblastic series. All the strains produce more leaflets in successive nodes (Figure 3.5 B), leaf area increases in older leaves, as does individual leaflet area (Figure 3.5 C - H). Leaf length shows a similar pattern in all strains to that of Oxford and the way that the petiole, inter-rachis and terminal rachis contribute to this variation at different nodes holds true for all strains as well

(Figure 3.5 I - L). Petiolule length also changes as plant proceeds along the heteroblastic series in a manner that conserved across the strains (Figure 3.5 M - P). However there is still considerable variation in all of these traits at different nodes.

3. 4. 3. Intraspecific variation in rate of change along the heteroblastic series

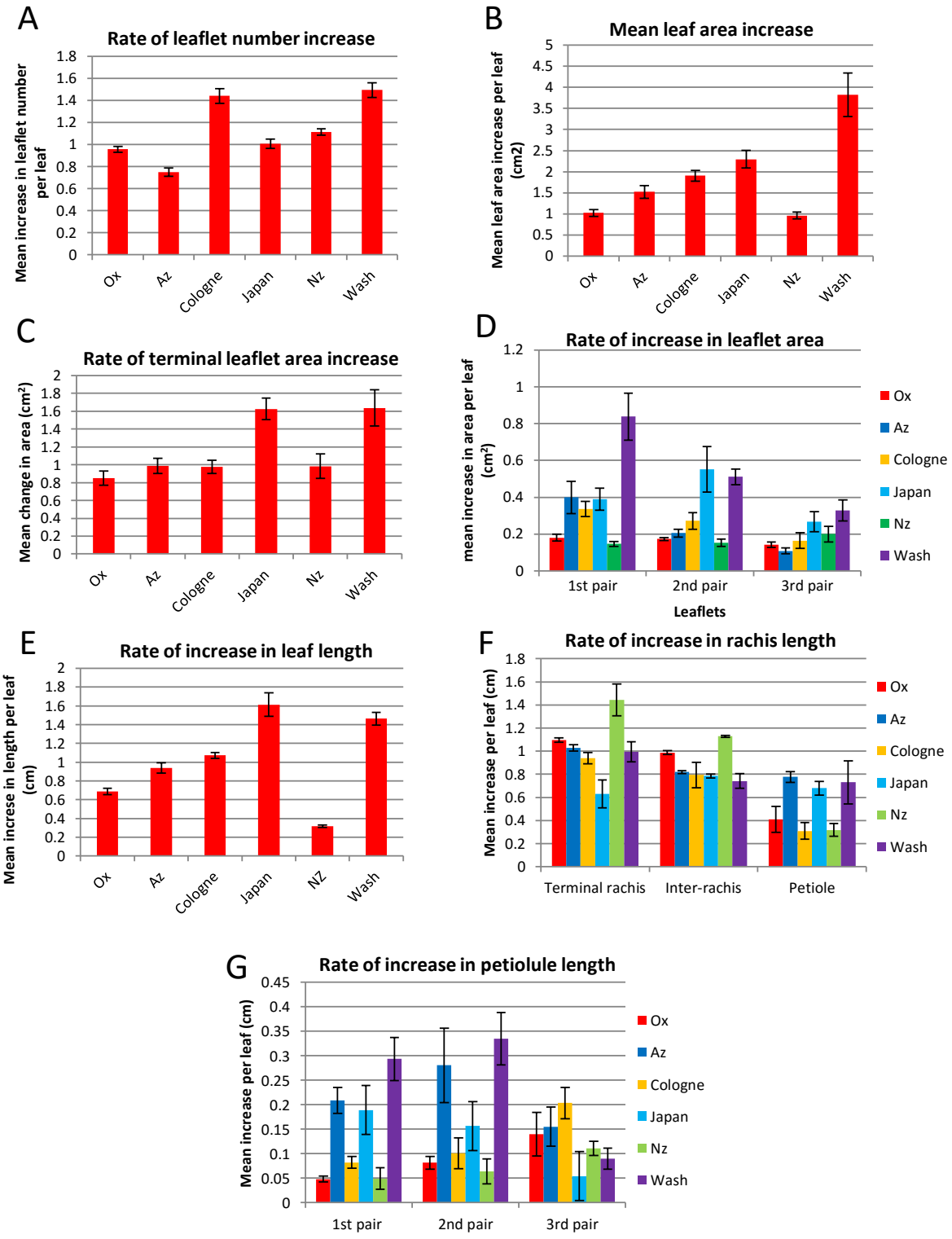
All the traits analysed here show variation at individual nodes, but they do show the same pattern of heteroblastic development. With all characters there is an initial increase in trait value in the early developing leaves and then these values plateau and remain relatively stable in the subsequent leaves. With some strains there is a decrease in the last few leaves before bolting (Figures 3.5) However, the rate at which the leaf traits change during early stage of development is variable. To provide a rough measurement for this rate of change I have determined for each plant the leaf node at which each trait reaches its peak and does not significantly change for the rest of the vegetative phase, I have then calculated the rate of change in the trait from the 1st leaf to this point, which was done by dividing the range in phenotype values by the number of leaves included in this range. This provides only a basic estimate of rate of trait change and it assumes that the change is linear; the aim of the analysis is to provide a basic measurement by which rate of early heteroblastic changes can be compared between strains.

For all traits there was significant variation between strains in the rate of change. This initial change in leaf shape is the most dominant feature in the heteroblastic profile of *C. hirsuta* and could represent the transition between juvenile and adult vegetative leaves. It is where the character of the vegetative shoot is changing the most and therefore it is interesting to compare the speeds with which different strains pass through this transition.

Leaf dissection - The rate of increase in leaflet number over successive leaves is highest in the accessions from Cologne and Wash while it is lowest in the accession from the Azores (Figure 3.6 A).

Leaf area - The Wash strain also gains leaf area at the fastest rate and the strains that have the slowest rate of leaf area increase are Oxford and New Zealand (Figure 3.6 B).

Figure 3.6 Intraspecific variation in rate of leaf shape trait change during initial development. Mean change per leaf in (A) leaflet number, (B) Leaf area, (C) Terminal leaflet area, (D) lateral leaf let area, (E) Leaf length, (F) Terminal rachis, Inter rachis, Petiole and (G) petiolule length. Rates of change calculated up to the leaf where trait value reaches a peak. Error bars indicate standard error of the mean.


A
Rate of leaflet number increase
Mean increase in leaflet number per leaf
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Ox
Az
Cologne
Japan
Nz
Wash
B
Mean leaf area increase
Mean leaf area increase per leaf (cm2)
5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Ox
Az
Cologne
Japan
Nz
Wash
C
Rate of terminal leaflet area increase
Mean change in area (cm²)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Ox
Az
Cologne
Japan
Nz
Wash
D
Rate of increase in leaflet area
mean increase in area per leaf (cm²)
1.2
1
0.8
0.6
0.4
0.2
0
1st pair
2nd pair
3rd pair
Leaflets
Ox
Az
Cologne
Japan
Nz
Wash
E
Rate of increase in leaf length
Mean increse in length per leaf (cm)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Ox
Az
Cologne
Japan
NZ
Wash
F
Rate of increase in rachis length
Mean increase per leaf (cm)
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Terminal rachis
Inter-rachis
Petiole
Ox
Az
Cologne
Japan
Nz
Wash
G
Rate of increase in petiolule length
Mean increase per leaf (cm)
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1st pair
2nd pair
3rd pair
Ox
Az
Cologne
Japan
Nz
Wash

Leaflet areas - Terminal leaflet area increases fastest in the Japan and Wash strains but there is no significant change in the rate of change between any of the other strains (Figure 3.6 C). With the lateral leaflets, Wash consistently has one of the highest rates of increase for each pair of leaflets (Figure 3.6 D). However different pairs of leaflets change in size at different rates in the Wash, Japan, Azores and Cologne strains; the leaflets nearest the terminal leaflet increase in size at a faster rate than the leaflets near the shoot in Wash, Azores and Cologne but the middle leaflets on Japan's leaves increase in area faster than those near the terminal leaflet and faster than those near the shoot (Figure 3.6 D). Conversely the rate of leaflet area increase is relatively uniform across the leaves of Oxford and New Zealand (Figure 3.6 D).

Leaf length - Leaf length increases most rapidly in the Japan and Wash strains and New Zealand shows the slowest increase in length (Figure 3.6 E). When this change in length is split into the different parts of the rachis, expansion in the inter-rachis and terminal rachis contributes more to leaf lengthening than the petiole in Oxford, Cologne and New Zealand (Figure 3.6 F). In the Azores, Japan and Wash strains this difference is much smaller or does not exist (Figure 3.6 F).

Petiolule length - The rate at which petiolules expand on successive leaves determines how fast the leaves widen during development. There is variation in petiolule expansion between strains; there is no strain in which petiolule expansion is stable at each pair of leaflets and there is no consistent pattern of expansion at each pair of leaflets along the longitudinal axis of the leaf (Figure 3.6 G). From the terminal leaflet towards the shoot of the plant; Oxford and Cologne show a slight steady increase in rate of petiolule expansion; Azores has a faster rate of expansion in the 1st and 2nd pairs of leaflets but this rate decreases in the 3rd pair; Japan has a relatively high rate of expansion in the 1st and 2nd pair of leaflets but this is much reduced in the 3rd pair of leaflets; New Zealand has the lowest rate of petiolule expansion in the two pairs of leaflets closest to the terminal leaflet and this increases slightly in the 3rd pair; Wash has the fastest rate of petiolule expansion in the 1st and 2nd pairs of leaflets but this is much reduced in the 3rd pair (Figure 3.6 G). Rate of expansion in the 4th pair of leaflets and petiolules was not analysed

because many of the strains begin bolting very soon after leaves with four pairs of lateral leaflets develop, providing insufficient leaflets to analyse rate of change.

3. 5. Oxford x Azores F6 Recombinant Inbred Line (RIL) mapping population

The remainder of this thesis focuses on the dissection of the genetic basis for leaf shape variation through QTL mapping. This was performed by analysing the phenotype and genotype of a recombinant inbred line (RIL) population descended from the F1 population created by crossing the Oxford strain with the Azores strain. It is important to look at the variation between these two strains to get a better idea of the variation that will be partly explained by the results of the QTL mapping.

3. 5. 1. Variation between Oxford and Azores strains

Rosette leaves - The two strains produce different number of vegetative leaves (Figure 3.5 A; Table 3.1), Oxford produces a mean of 11.2 leaves (SEM: +/- 0.2 leaves) before bolting and Azores produces 12.6 leaves (SEM: +/- 0.4) and the difference is significantly different (T test p = 0.01). This suggests that the vegetative phase of growth lasts longer in Azores than it does in Oxford. This is reflected in the fact that Oxford flowers earlier than Azores.

Leaf dissection - Over the whole course of the vegetative phase of development Oxford has leaves that bear more leaflets than the leaves of Azores, however this does not hold true for every leaf on the heteroblastic series (Figure 3.5 B; Table 3.1). As with all strains the 1st leaf has only the single leaflet, at the second leaf node Azores has a mean number of leaflets greater than two but Oxford never has more than one leaflet in the plants we analysed. There is no significant difference in leaflet number at leaf node 3, Oxford has significantly more leaflets at leaf 4, and

then there is no significant difference until leaf 9 to leaf 11 where Oxford has significantly more leaflets than Azores (Figure 3.5 B; Table 3.1). When the numbers of leaflets are totalled for the 1st 11 leaves, Oxford has significantly more leaflets than Azores (Table 3.1).

Leaf area - The Azores strain has significantly larger leaves at leaf nodes five and six but at no other point along the heteroblastic series do the strains vary significantly in leaf area (Figure 3.5 C; Table 3.1)

Leaflet area - The area of the terminal leaflets do not vary significantly between the two strains at any of the leaf nodes tested but the lateral leaflets do (Figure 3.5 D; Table 3.1). The 1st pair of lateral leaflets are significantly bigger on Azores at leaf nodes 3 to 8 (Figure 3.5 E; Table 3.1). The same holds true for the second pair of leaflets at leaf nodes 5 to 8 (Figure 3.5 F; Table 3.1). The 3rd pair of leaflets do not vary significantly and the 4th pair of leaflets vary significantly at leaf node 11 where those of Oxford are significantly larger (Figure 3.5 G; Table 3.1). There is a much greater difference in leaflet area in the 1st pair of leaflets than there is in the second pair of leaflets – thus there is a gradient of decreasing variation from the terminal leaflet towards the shoot of the plant.

Leaf length - The leaves of Azores are significantly longer than those of Oxford at all leaves along the heteroblastic series (Figure 3.5 I; Table 3.1). The difference is due more to expansion in the terminal rachis and petiole of the Azores leaf as these are both significantly longer in the Azores strain at all leaf nodes whereas Inter rachis length is only significantly longer at two leaf nodes (Figure 3.5 J - L; Table 3.1).

Petiolule length - As well as being longer the leaves of Azores are also much wider than those of Oxford. The petioles on all leaflets are significantly longer on the Azores leaves than they are on the leaves of Oxford (Figure 3.5 M - P; Table 3.1).

Variation in rate of trait change - To gauge how fast the two strains are transitioning through vegetative growth is interesting to compare the rate of change in leaf shape traits (Figure 3.6; Table 3.2). Neither Oxford nor Azores shows consistent faster development of traits, so it is

difficult to decide which makes the faster transition between growth phases. If rate of leaflet number increase is used as an indicator of transition speed then Oxford can be said to develop faster than Azores, but if leaf length and width increase is used; Azores can be said to progress faster through the heteroblastic series (Figure 3.6; Table 3.2). As mentioned before this is only a rough measurement for rate of trait change designed to give an indication of how fast strains might be progressing heteroblastic change.

Table 3.1 Phenotypic differences between Oxford and Azores leaf shape traits. Statistically significant differences are highlighted yellow ($P < 0.05$),

	Oxford trait value	SEM (+/-)	Azores trait value	SEM (+/-)	T test, P value
Rosette leaf number					
	11.20	0.20	12.60	0.40	0.01
Leaflet Number					
- *Leaf 1*	1.00	0.00	1.00	0.00	-
- *Leaf 2*	1.00	0.00	2.40	0.40	0.02
- *Leaf 3*	2.80	0.20	3.00	0.00	0.37
- *Leaf 4*	4.40	0.24	3.20	0.20	0.01
- *Leaf 5*	5.00	0.00	4.20	0.49	0.18
- *Leaf 6*	5.60	0.24	5.20	0.20	0.24
- *Leaf 7*	7.00	0.00	6.40	0.40	0.21
- *Leaf 8*	7.40	0.24	6.80	0.20	0.09
- *Leaf 9*	8.60	0.24	7.20	0.20	< 0.01
- *Leaf 10*	9.00	0.00	7.80	0.37	0.03
- *Leaf 11*	9.40	0.24	8.20	0.20	0.01
Total leaflets on fist 11 leaves					
	61.20	0.66	55.40	1.63	0.01
Leaf Area (cm^2)					
- *Leaf 1*	0.85	0.05	0.85	0.05	0.94
- *Leaf 2*	2.00	0.08	2.25	0.16	0.20
- *Leaf 3*	3.30	0.10	4.21	0.43	0.10
- *Leaf 4*	4.80	0.38	6.04	0.49	0.08
- *Leaf 5*	5.77	0.33	7.41	0.64	0.05
- *Leaf 6*	6.55	0.24	8.15	0.71	0.05
- *Leaf 7*	7.34	0.53	8.27	0.40	0.19
- *Leaf 8*	7.91	0.46	8.92	0.62	0.23

	Oxford trait value	SEM (+/-)	Azores trait value	SEM (+/-)	T test, P value
Leaf length (cm)					
- *Leaf 1*	0.68	0.06	1.67	0.09	< 0.01
- *Leaf 2*	1.17	0.05	2.91	0.22	< 0.01
- *Leaf 3*	1.76	0.08	4.12	0.37	< 0.01
- *Leaf 4*	2.66	0.09	5.50	0.34	< 0.01
- *Leaf 5*	3.40	0.18	6.59	0.31	< 0.01
- *Leaf 6*	3.95	0.23	7.96	0.27	< 0.01
- *Leaf 7*	4.84	0.28	8.59	0.22	< 0.01
- *Leaf 8*	5.15	0.23	9.05	0.20	< 0.01
- *Leaf 9*	6.25	0.24	9.66	0.10	< 0.01
- *Leaf 10*	6.43	0.36	9.67	0.33	< 0.01
- *Leaf 11*	6.51	0.31	9.58	0.48	< 0.01
Terminal rachis length (cm)					
- *Leaf 3*	0.58	0.06	1.15	0.12	< 0.01
- *Leaf 4*	0.74	0.05	1.51	0.10	< 0.01
- *Leaf 5*	0.85	0.04	1.60	0.07	< 0.01
- *Leaf 6*	0.89	0.08	2.01	0.11	< 0.01
- *Leaf 7*	1.07	0.10	1.81	0.10	< 0.01
- *Leaf 8*	0.98	0.18	1.74	0.10	0.01
- *Leaf 9*	0.71	0.13	1.63	0.07	< 0.01
- *Leaf 10*	0.70	0.10	1.33	0.11	< 0.01
- *Leaf 11*	0.63	0.03	1.15	0.15	0.02
Inter Rachis length					
- *Leaf 5*	1.33	0.19	2.73	0.56	0.03
- *Leaf 6*	1.96	0.36	2.45	0.40	0.40

- *Leaf 9*	8.45	0.39	8.30	0.64	0.84
- *Leaf 10*	7.64	0.41	7.80	0.66	0.85
- *Leaf 11*	7.20	0.29	7.05	0.60	0.83
Terminal leaflet area (cm^2)					
- *Leaf 1*	0.85	0.05	0.85	0.05	0.88
- *Leaf 2*	2.00	0.08	1.58	0.17	0.18
- *Leaf 3*	2.80	0.17	2.49	0.32	0.27
- *Leaf 4*	3.55	0.31	3.20	0.28	0.88
- *Leaf 5*	3.91	0.20	3.42	0.18	0.80
- *Leaf 6*	4.01	0.15	3.67	0.19	0.66
- *Leaf 7*	4.00	0.32	3.26	0.10	0.05
- *Leaf 8*	3.72	0.28	3.28	0.20	0.54
- *Leaf 9*	2.96	0.19	2.85	0.19	0.99
- *Leaf 10*	2.68	0.20	2.46	0.15	0.61
- *Leaf 11*	1.85	0.18	2.12	0.16	0.78
1st pair Lateral leaflet area (cm^2)					
- *Leaf 2*	-	-	0.48	0.06	-
- *Leaf 3*	0.28	0.04	0.86	0.07	< 0.01
- *Leaf 4*	0.58	0.04	1.41	0.08	< 0.01
- *Leaf 5*	0.73	0.04	1.51	0.07	< 0.01
- *Leaf 6*	0.87	0.04	1.72	0.09	< 0.01
- *Leaf 7*	1.03	0.04	1.67	0.05	< 0.01
- *Leaf 8*	1.16	0.06	1.66	0.07	< 0.01
- *Leaf 9*	1.30	0.05	1.45	0.06	0.07
- *Leaf 10*	1.12	0.05	1.27	0.06	0.06
- *Leaf 11*	1.07	0.04	1.17	0.07	0.22
2nd pair Lateral leaflet area (cm^2)					
- *Leaf 4*	0.07	0.02	0.05	-	-
- *Leaf 5*	0.20	0.02	0.57	0.03	< 0.01
- *Leaf 6*	0.39	0.02	0.69	0.07	< 0.01
- *Leaf 7*	0.52	0.03	0.86	0.07	< 0.01
- *Leaf 8*	0.69	0.06	0.89	0.05	0.02
- *Leaf 9*	0.89	0.05	0.88	0.06	0.88
- *Leaf 10*	0.86	0.04	0.85	0.06	0.90
- *Leaf 11*	0.88	0.03	0.76	0.05	0.06
3rd pair lateral leaflet area (cm^2)					
- *Leaf 6*	0.04	0.01	0.10	-	-
- *Leaf 7*	0.12	0.02	0.20	0.04	0.12
- *Leaf 8*	0.25	0.05	0.29	0.04	0.48
- *Leaf 9*	0.47	0.04	0.38	0.04	0.15
- *Leaf 10*	0.52	0.04	0.48	0.04	0.59
- *Leaf 11*	0.55	0.02	0.47	0.03	0.06
4th pair lateral leaflet area (cm^2)					
- *Leaf 9*	0.11	0.04	0.17	-	-
- *Leaf 10*	0.14	0.03	0.15	0.02	0.81
- *Leaf 11*	0.17	0.02	0.10	0.03	0.03

- *Leaf 7*	2.81	0.05	3.88	0.43	0.07
- *Leaf 8*	3.44	0.17	4.18	0.15	0.01
- *Leaf 9*	4.64	0.19	4.47	0.22	0.59
- *Leaf 10*	4.74	0.36	5.59	0.46	0.18
- *Leaf 11*	4.92	0.42	5.87	0.36	0.13
Petiole length (cm)					
- *Leaf 1*	0.68	0.06	1.67	0.09	< 0.01
- *Leaf 2*	1.17	0.05	2.21	0.20	0.01
- *Leaf 3*	1.18	0.05	2.96	0.25	< 0.01
- *Leaf 4*	0.86	0.07	3.60	0.33	< 0.01
- *Leaf 5*	1.21	0.15	3.35	0.46	0.01
- *Leaf 6*	1.10	0.26	3.51	0.30	< 0.01
- *Leaf 7*	0.96	0.16	2.90	0.35	< 0.01
- *Leaf 8*	0.90	0.17	3.12	0.17	< 0.01
- *Leaf 9*	0.90	0.16	3.56	0.28	< 0.01
- *Leaf 10*	0.99	0.19	2.75	0.36	< 0.01
- *Leaf 11*	0.96	0.21	2.56	0.15	< 0.01
1st pair petiolule length (cm)					
- *Leaf 2*	-	-	0.45	0.03	-
- *Leaf 3*	0.17	0.02	0.66	0.02	< 0.01
- *Leaf 4*	0.24	0.01	0.96	0.03	< 0.01
- *Leaf 5*	0.25	0.02	0.98	0.04	< 0.01
- *Leaf 6*	0.31	0.03	1.06	0.05	< 0.01
- *Leaf 7*	0.32	0.02	1.04	0.06	< 0.01
- *Leaf 8*	0.36	0.03	0.92	0.06	< 0.01
- *Leaf 9*	0.36	0.03	0.78	0.03	< 0.01
- *Leaf 10*	0.37	0.03	0.69	0.02	< 0.01
- *Leaf 11*	0.40	0.02	0.66	0.06	0.02
2nd pair petiolule length (cm)					
- *Leaf 4*	0.09	0.01	0.22	-	-
- *Leaf 5*	0.17	0.02	0.68	0.04	< 0.01
- *Leaf 6*	0.24	0.02	0.92	0.06	< 0.01
- *Leaf 7*	0.33	0.03	0.92	0.05	< 0.01
- *Leaf 8*	0.33	0.02	0.89	0.04	< 0.01
- *Leaf 9*	0.35	0.02	0.75	0.03	< 0.01
- *Leaf 10*	0.37	0.02	0.64	0.03	< 0.01
- *Leaf 11*	0.37	0.02	0.58	0.04	0.01
3rd pair petiolule length (cm)					
- *Leaf 6*	0.11	0.02	0.30	-	-
- *Leaf 7*	0.23	0.03	0.54	0.06	< 0.01
- *Leaf 8*	0.23	0.02	0.64	0.07	< 0.01
- *Leaf 9*	0.29	0.03	0.65	0.03	< 0.01
- *Leaf 10*	0.32	0.02	0.73	0.05	< 0.01
- *Leaf 11*	0.32	0.02	0.70	0.07	0.02
4th pair petiolule length (cm)					
- *Leaf 9*	0.19	0.02	0.50	-	-
- *Leaf 10*	0.17	0.03	0.48	0.06	0.01
- *Leaf 11*	0.24	0.02	0.45	0.05	0.02

Table 3.2 Variation in rate of trait change between Oxford and Azores. Rates calculated from the first leaf to the peak in trait value. Yellow highlighting indicates significant differences between Oxford and Azores (P<0.05).

Trait	Oxford - Mean trait change per node	Azores - Mean trait change per node	T test, P Value
Leaflet number	0.96	0.75	<0.01
Whole leaf area (cm^2)	1.02	0.96	0.60
Terminal leaflet area (cm^2)	0.85	0.99	0.29
Mean lateral leaflet area (cm^2)	0.17	0.24	0.12
- *1st pair lateral leaflet area (cm^2)*	0.18	0.40	0.07
- *2nd pair lateral leaflet area (cm^2)*	0.17	0.21	0.19
- *3rd pair lateral leaflet area (cm^2)*	0.14	0.11	0.16
Leaf length (cm)	0.69	0.94	<0.01
- *Terminal rachis length (cm)*	1.10	1.03	0.07
- *Inter-rachis length (cm)*	0.99	0.82	<0.01
- *Petiole length (cm)*	0.41	0.78	<0.01
Mean petiolule length (cm)	0.06	0.21	<0.01
- *1st pair petiolule length (cm)*	0.05	0.21	<0.01
- *2nd pair petiolule length (cm)*	0.08	0.31	0.02
- *3rd pair petiolule length (cm)*	0.06	0.14	0.16

3. 5. 2. - Phenotypic analysis of the parental strains in alternative growing conditions

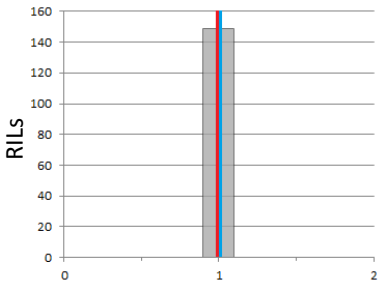
Amongst the RIL population 15 representative of each of the parental strains, Oxford and Azores, were grown for the sake of comparison. As described in the methods chapter of this thesis the conditions were different to those in which the strains were analysed initially. The only difference in growth conditions was that the plants from Figures 3.2 - 3.6 were grown in conditions with a day time temperature of 24°C and night time temperature of 22°C and the plants grown alongside the RILs were exposed to temperatures of 18°C in the day time and 16°C in the night time, the length of photoperiod was identical (16 hours). The resultant plants had differing phenotypes.

The most notable difference in phenotype was the number of leaves produced before bolting; in the first set of conditions Oxford produces a mean of 11.2 rosette leaves (SEM: 0.2) and Azores produced a mean of 12.6 rosette leaves (SEM: 0.4, Figure 3.5 A and Table 3.1). When grown in colder daytime temperatures Oxford produced more rosette leaves at 14.4 (SEM: 0.3) and Azores produced fewer at 11.8 (SEM: 0.2) (Table 3.3). The two strains have reacted in opposite directions and to varying degrees in different growing conditions.

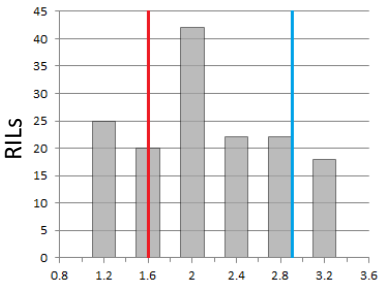
In both strains there were slightly more leaflets at 18°C than at 24°C, but the difference between the two strains are still preserved along the heteroblastic series (Figure 3.5 B, Table 3.1 and Table 3.2). In both conditions leaf two of Azores has more leaflets, at leaf 3 there is very little difference between the strains, and from leaf 4 onwards the Oxford leaves have the greater number of leaflets. The only point at which this difference between the strains is not maintained is at leaf 10, where there is negligible difference in 18°C conditions (Table 3.3).

This experiment has demonstrated that rosette leaf number and leaflet number are plastic traits that are sensitive to deviations in their environment. Any reference to the Oxford or Azores phenotype in the following section refers to the phenotype observed when the strains were grown alongside the RIL population with day time temperature at 18°C.

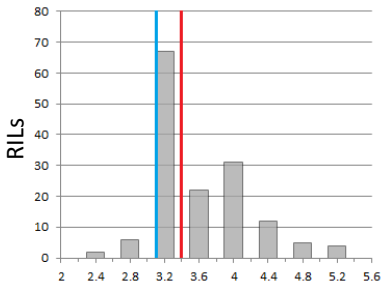
Figure 3.7 Distribution of phenotypes in the Oxford x Azores RIL population. Frequency histograms for the indicated phenotypic parameters in the RIL population. The red and blue lines indicate the mean values of the parental Oxford and Azores strains, respectively.



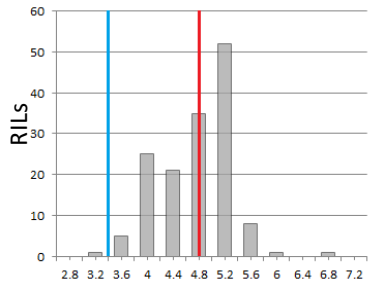
Leaflets Leaf 1



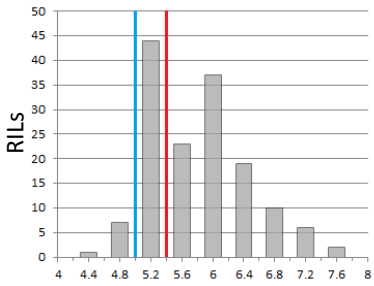
Leaflets Leaf 2



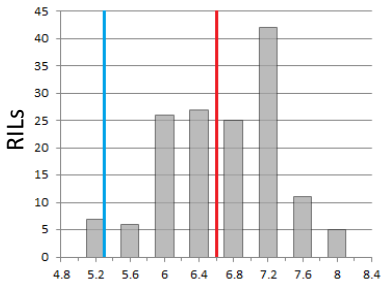
Leaflets Leaf 3



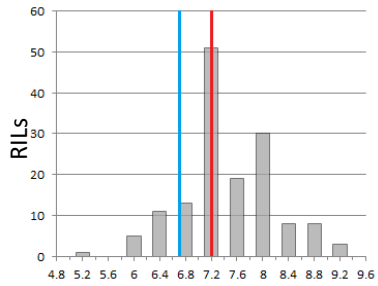
Leaflets Leaf 4



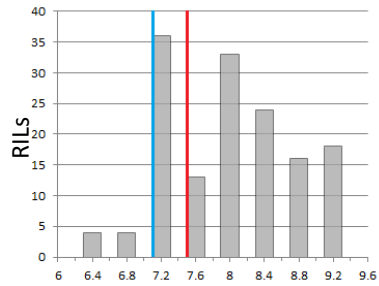
Leaflets Leaf 5



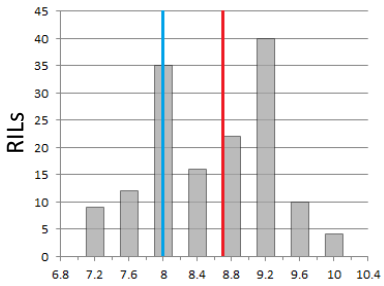
Leaflets Leaf 6



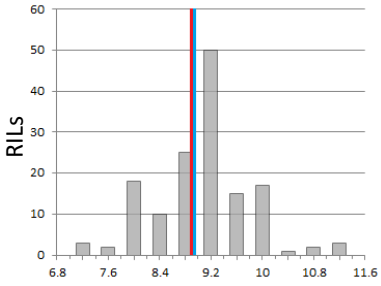
Leaflets Leaf 7



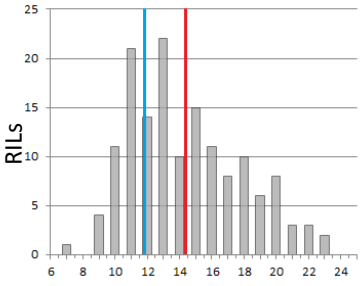
Leaflets Leaf 8



Leaflets Leaf 9



Leaflets Leaf 10

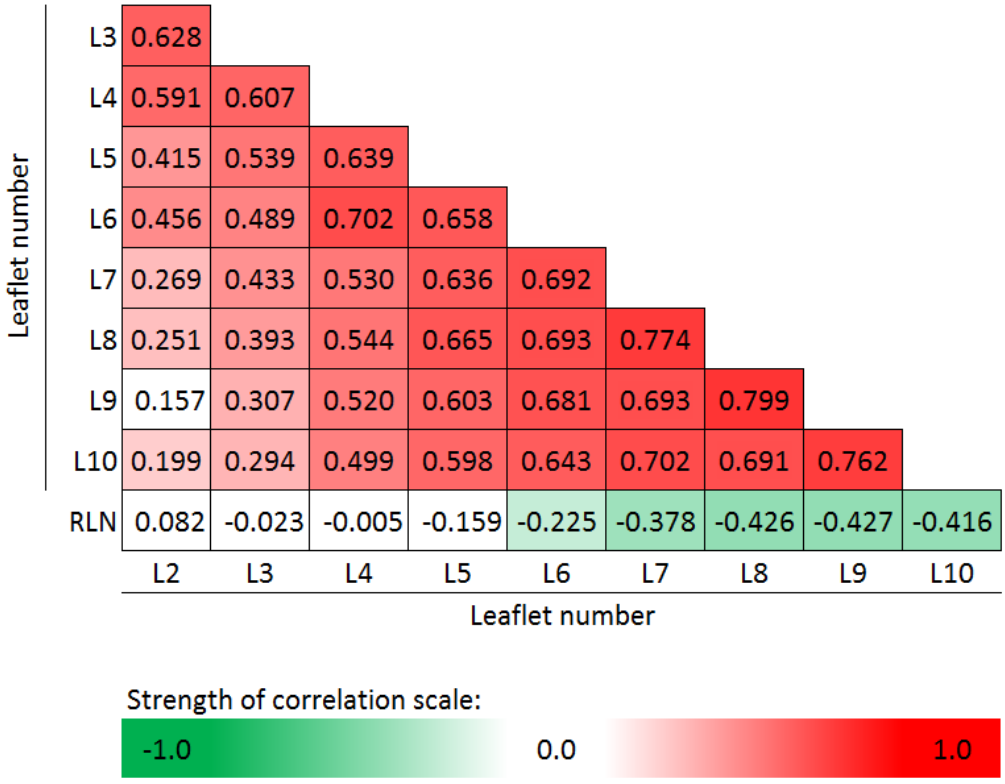


Rosette leaves

Oxford mean value

Azores mean value

Figure 3.8 Genetic Correlations Between Leaf Traits in RIL Population. Values are Pearson's correlation coefficients between phenotypic parameters measured in the RIL population. Significant correlations are highlighted ($P < 0.05$), red highlighting indicates a positive correlation and green highlighting indicates a negative correlation, the intensity of the colour corresponds with the strength of the correlation**.** RLN: rosette leaf number.


Leaflet number
L3 0.628
L4 0.591 0.607
L5 0.415 0.539 0.639
L6 0.456 0.489 0.702 0.658
L7 0.269 0.433 0.530 0.636 0.692
L8 0.251 0.393 0.544 0.665 0.693 0.774
L9 0.157 0.307 0.520 0.603 0.681 0.693 0.799
L10 0.199 0.294 0.499 0.598 0.643 0.702 0.691 0.762
RLN 0.082 -0.023 -0.005 -0.159 -0.225 -0.378 -0.426 -0.427 -0.416
L2 L3 L4 L5 L6 L7 L8 L9 L10
Leaflet number
Strength of correlation scale:
-1.0
0.0
1.0

3. 5. 3. Phenotypic analysis of the Oxford x Azores recombinant inbred line population

After the plants had produced flowers the number of rosette leaves was counted as was the number of leaflets on each of these. Due to the multiple generations of selfing involved in the construction of the RIL population it was assumed the genome was quasi-homozygous (later genotyping revealed that the RILs were homozygous at an average of 97.5% of the markers used) therefore individuals of the same line would be genetically identical. This permitted us to phenotype multiple replicates for each line to minimise variation due to environmental factors. Three replicates were phenotyped for each of the 149 RILs.

Phenotype variation - There was considerable variation in the number of rosette leaves produced before bolting (Table 3.3, Figure 3.7). The least number of leaves produced was 6, and the most number of leaves produced was 27. There was also variation in terms of leaflet dissection; leaflet number varied enormously with the largest range at leaf nodes 7 and 10, at 4 leaflets. The only leaf that did not have variable number of leaflets was the 1st leaf; all 1st leaves never produced any lateral leaflets, only a single terminal leaflet (Table 3.3, Figure 3.7), a characteristic shared with all other strains (Figure 3.5 B). For the sake of efficient analysis only the first 10 leaves of all RILs were analysed extensively in the later sections of this thesis.

Trait Heritability - The broad sense heritability for all traits was calculated. This is the proportion of variation attributable to genotype, which includes epistasis and not just additive effects (it would also include dominance effects if the RIL population was not quasi-homozygous). The heritability for the traits measured ranged from 0.368 for leaflet number on leaf nine and 0.615 for leaflet number on leaf eight (Table 3.3). This indicates that leaflet number has a strong genetic basis.

Transgressive segregation All the traits showed transgressive segregation compared to the parental strains of the RIL population (Table 3.3, Figure 3.7); the parental values in this instance are from those grown alongside the RIL population and not from the plants described in the earlier part of this chapter. Transgressive segregation, the presence of more extreme phenotypes

than in either of the parents, can be due to the parents having both positively and negatively acting alleles at different loci that are rearranged in extreme combinations because of recombination during the creation of the RIL population (Reiseberg et al, 1999). It is less likely that it is due to over-dominance, which is where the phenotype of the heterozygote sits outside the phenotypical range of both homozygote parents, because in this RIL population the entire genome is essentially homozygous. Therefore the presence of transgressive segregation can be taken as evidence that leaflet number is influenced by multiple loci and that both parental strains harbour a mixture of alleles that promote and inhibit leaflet production. It is interesting to note that these RILs captured the full extent of phenotypic variation observed earlier in this chapter amongst the six strains analysed, despite the fact that the two founder strains did not represent the phenotypic extremes.

Trait correlation - It is possible to determine whether the different traits are likely to share genetic determinants or be under the control of independent loci by analysing phenotypic correlations between the traits in the RIL population. If two traits correlate with each other then it is likely that they are influenced by many of the same loci, whereas a lack of correlation suggests the traits are controlled by independent loci.

Leaflet number on all leaves have significant positive correlations with the leaflet number on all other leaves, except for leaves two and nine where there was no significant relationship (Figure 3.8). This suggests that there are probably loci that influence leaflet number for the entirety of vegetative development. However as can be seen by the intensity of highlighting in Figure 3.8, the strength of a correlation between leaflet numbers on any two leaves is influenced by the proximity of those leaves to each other on the heteroblastic series. The strongest correlations in leaflet number exist on leaves that developed immediately before or after one another and the correlations that exist in leaflet number on leaves that developed at very different times of development are much weaker. This suggests that while there are loci that influence the

degree of leaf dissection throughout heteroblastic development there are also additional genetic determinants of leaflet number that act independently of these loci and have influence at particular phases of development. Alternatively, the stronger correlation between leaves in close proximity could be enhanced by environmental effects.

Variation in the number of rosette leaves produced before bolting showed significant correlation with leaflet number on leaves six to ten. The correlation is negative; plants that produce more vegetative leaves produce leaves with fewer leaflets. It appears that the same loci that extend the vegetative phase of development also have a negative effect on leaflet number.

Table 3.3 Quantitative-Genetic Parameters for Morphological Traits. Vg, among-RILs (genetic) variance; Ve, residual (environmental) variance; H^2, broad-sense heritability ; SEM, standard error of the mean; n.a., not applicable.

	Oxford		Azores		RILs						
Trait	Mean	SEM	Mean	SEM	Mean	SEM	Minimum	Maximum	H2	Vg	Ve
L1 leaflets	1.00	0.00	1.00	0.00	1.00	0.00	1.00	1.00	n.a.	0.00	0.00
L2 leaflets	1.60	0.21	2.92	0.08	1.96	0.04	1.00	3.00	0.609	214.83	138.17
L3 leaflets	3.27	0.12	3.08	0.08	3.42	0.04	2.33	5.00	0.558	155.44	123.33
L4 leaflets	4.80	0.11	3.42	0.15	4.59	0.04	3.00	6.50	0.494	131.06	134.33
L5 leaflets	5.40	0.19	5.00	0.00	5.62	0.04	4.33	7.50	0.572	199.41	149.17
L6 leaflets	6.60	0.13	5.25	0.13	6.51	0.04	5.00	8.00	0.549	187.30	153.67
L7 leaflets	7.20	0.11	6.67	0.19	7.27	0.04	5.00	9.00	0.582	220.68	158.67
L8 leaflets	7.53	0.19	7.08	0.08	7.82	0.05	6.33	9.00	0.615	246.49	154.17
L9 leaflets	8.67	0.16	8.00	0.21	8.51	0.15	7.00	10.00	0.368	1486.12	2555.50
L10 leaflets	8.87	0.13	8.92	0.08	8.86	0.05	7.00	11.00	0.609	214.83	138.17
Rosette leaf number	14.40	0.29	11.83	0.17	14.05	0.18	6.00	27.00	0.565	67.59	52.00

3. 6. Intraspecific variation in leaf shape - Discussion

These results have demonstrated that there is considerable variation between different geographically disparate strains. There were significant differences for many traits at different leaves but the basic pattern of heteroblastic change was shared by all strains; an increase in leaflet number, leaflet area, length and width for the initial part of shoot development followed by a plateau of less change before the reproductive phase of growth is initiated. However the speed at which different strains progress through heteroblastic change was variable as demonstrated through the analysis of rate of trait change during the early part of shoot development. This is suggestive of varying lengths of vegetative juvenile and vegetative adult phases (as in Poethig, 1990). Thus the major source of intraspecific variability amongst *C. hirsuta* natural strains comes from the variation in the rate of progression through heteroblastic development.

There is no strain in which all leaf traits change faster or slower than all the other strains, suggesting that different leaf traits alter, to a limited extent, independently of each other. It appears that different traits are responding to heteroblastic change at variable rates, for my purposes this makes it hard to use any of the measured traits to assign different section of the heteroblastic series as belonging to either vegetative juvenile or vegetative adult phases. In *A. thaliana,* trichome distribution can be used as a marker for different phases of vegetative growth (Telfer et al., 1997); we have not found a similar trait for *C. hirsuta.*

These observations highlight the importance of considering leaf trait variation within the context of heteroblastic development. Studying variation at an arbitrary fixed point in the series will not provide a picture of the dynamic interactions acting during development to produce variable morphology. The complexity is two-fold; not only do different strains progress through heteroblastic change at variable rates but different traits change at different rates in different strains.

I looked in more detail at comparing and contrasting the Oxford and Azores strains since these were the parental strains used in constructing the experimental mapping populations later in

this chapter. Relative to the variation observed amongst the other strains the range in the number of rosette leaves differed little between the two strains. On the leaves themselves, Oxford produced a greater number of lateral leaflets than Azores, Oxford had larger lateral leaflets, and the variation was most pronounced on the pair of leaflets closest to the terminal leaflet and furthest away from the shoot. As well as being larger, the leaves of Oxford were also longer and wider due to expansion of the terminal rachis, petiole and petiolules. With regards to analysis of trait rate change, different traits showed variation and no conclusion about which strain, Oxford or Azores, transfers the fastest through heteroblasty.

Data were obtained for these two strains in alternative growing conditions. In colder conditions both strains produced more leaflets than they did in the original warmer conditions. At 24°C Azores produced more rosette leaves but at 18°C Oxford produced more leaves. These results indicate that rosette leaf production is highly plastic and sensitive to external stimuli. This is an unsurprising result since plants need to react to the external environment in order that they allocate their resources towards reproductive growth when conditions are best suited to pollination and seed production.

From phenotypic analysis of the Oxford x Azores RIL population and by calculating heritability values, it has been possible to determine that number of rosette leaves and degree of leaf dissection have strong genetic determinants. All the traits show greater variation in the RIL population than that which exists between the parental strains, which indicates that each trait has many genetic determinants that have been rearranged in separate lines by recombination. Examination of phenotypic correlations show evidence that there are loci affecting leaflet number on all of the first 10 leaves but that there are other loci have temporally localised influences on leaflet production. It is difficult to determine whether genetic correlations between the number of rosette leaves and leaflet production indicate a common genetic control or that the correlation comes about through indirect effects.

Assessing all these results together we can expect that as I dissect the genetic architecture behind these leaf traits, I expect to see; the influence of multiple loci affecting multiple traits and a system that is open to external stimuli.

3. 7. QTL mapping

The identification QTL controlling leaf shape traits in *C. hirsuta* is an important step towards understanding how the production of alternative leaf shapes are genetically controlled. In this study the variation in leaf number was analysed using an F_6 RIL population derived from a cross between the Oxford and Azores strain.

As seen in the above results, the leaves of different strains of *C. hirsuta* vary in many different ways; however one of the most striking variations is leaflet number. This trait can be measured quickly and easily, no subjective judgment is required in its measurement and there is also reduced risk, relative to size and shape traits, that it will be affected by leaf damage. I have therefore placed the emphasis on discovering QTL associated with variable leaflet number, however as will be seen in the following chapter the QTL discovered not only affect leaflet number they also affect the size and shape of leaves and leaflets.

QTL mapping was performed with the aim of identifying the regions of the genome that influence leaflet number and rosette leaf number. Initially composite interval mapping was performed one trait at a time using R/QTL (Broman et al. 2003; www.rqtl.org). This approach successfully identified QTL for leaflet number on most of the leaves and there were numerous co-locations of QTL. This partly explains the strong phenotypic correlations discovered amongst the different traits in the RIL population (Figure 3.8). The confirmation of this genetic correlation made it valid to perform multi trait QTL mapping. Multi-trait QTL models are essential in order to understand the genetic basis of correlations between traits (van Eeuwijk et al, 2010), such as

the strong correlations between leaflet numbers on leaves 2 to 10. Mixed models allow estimating the fixed QTL effects while explicitly modelling the random genetic and non-genetic covariances. This approach improves on the single trait QTL mapping by modelling the genetic correlations while fitting QTL effects, this provides increased power to detect QTL and it means that the data is analysed more realistically relative to a collection of single trait analyses that assume no genetic correlations (Malosetti et al., 2008). By applying multi trait mixed model methodology using Genstat (Payne et al., 2009) six QTL were detected that affect leaflet numbers on leaves 2 to 10 and rosette leaf number. QTL mapping is only capable of identifying QTL that have a major effect on the trait in question but because of the inherent variation and loci linkage it will not be able to identify many of the QTL that have a smaller effect on the trait (Koornneef et al., 2004). This weakness is recognised and we do not expect to be able to uncover loci that explain 100% of the heritable variation.

3. 7. 1. Linkage map

Linkage mapping using the 173 single nucleotide polymorphisms (SNP) markers and the Joinmap 4.0 software was used to obtain a map with a total of eight linkage groups which corresponded to the eight chromosomes expected (Hay et al., 2014). The assembled map had a total length of 868.1 cM, with an average interval size between markers of 5.04 cM and the largest interval between adjacent markers was 24.3 cM (Figure 3.9)

3. 7. 2 QTL mapping – Single trait analysis

For each trait many additive QTL were detected for leaflet number and rosette leaf number using composite interval mapping in *R/QTL*, these are summarised in Table 3.4 and LOD score profiles of the Oxford/Azores genome showing the detection of QTL are presented in Figure 3.10. Percent of the explained phenotypic variation for single QTL effects varied among

Figure 3.9 The Oxford/Azores genetic linkage map. Eight linkage groups are numbered corresponding to the eight chromosomes in *C. hirsuta.* Horizontal bars indicate the position of 173 molecular markers mapped to the Oxford x Azores F6 recombinant inbred line (RIL) population. Marker positions in cM are shown to the left of each linkage group and marker identifiers to the right.

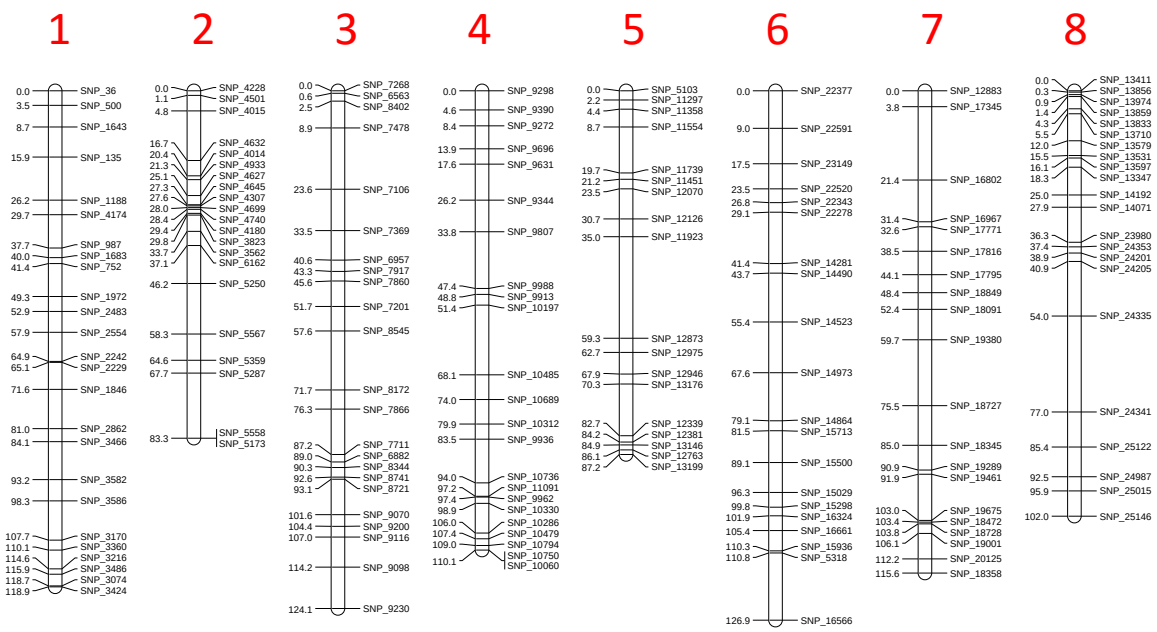
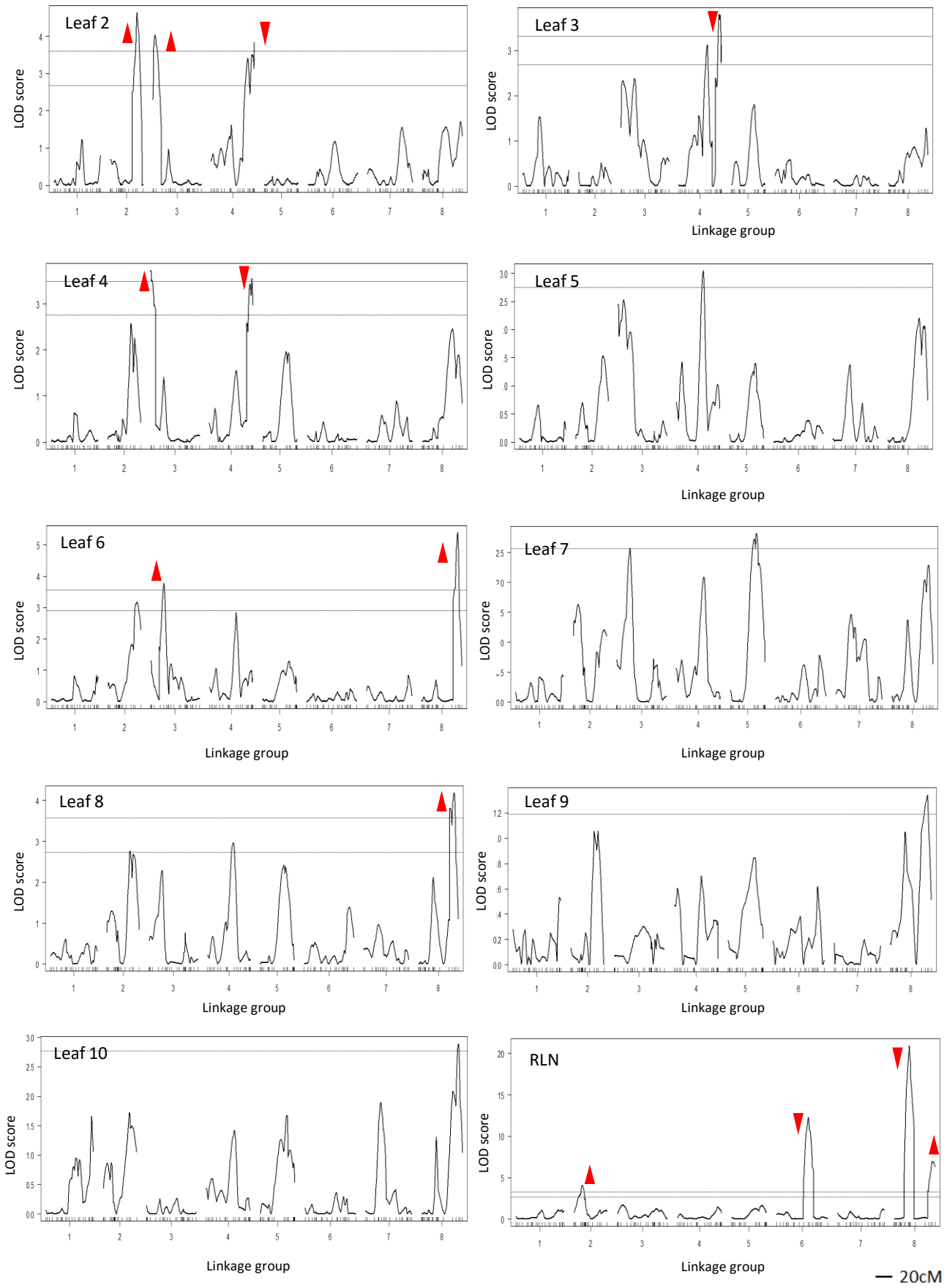

1
0.0 SNP_36
3.5 SNP_500
8.7 SNP_1643
15.9 SNP_135
26.2 SNP_1188
29.7 SNP_4174
37.7 SNP_987
40.0 SNP_1683
41.4 SNP_752
49.3 SNP_1972
52.9 SNP_2483
57.9 SNP_2554
64.9 SNP_2242
65.1 SNP_2229
71.6 SNP_1846
81.0 SNP_2862
84.1 SNP_3466
93.2 SNP_3582
98.3 SNP_3586
107.7 SNP_3170
110.1 SNP_3360
114.6 SNP_3216
115.9 SNP_3486
118.7 SNP_3074
118.9 SNP_3424
2
0.0 SNP_4228
1.1 SNP_4501
4.8 SNP_4015
16.7 SNP_4632
20.4 SNP_4014
21.3 SNP_4933
25.1 SNP_4627
27.3 SNP_4645
27.6 SNP_4307
28.0 SNP_4699
28.4 SNP_4740
29.4 SNP_4180
29.8 SNP_3823
33.7 SNP_3562
37.1 SNP_6162
46.2 SNP_5250
58.3 SNP_5567
64.6 SNP_5359
67.7 SNP_5287
83.3 SNP_5558
SNP_5173
3
0.0 SNP_7268
0.6 SNP_6563
2.5 SNP_8402
8.9 SNP_7478
23.6 SNP_7106
33.5 SNP_7369
40.6 SNP_6957
43.3 SNP_7917
45.6 SNP_7860
51.7 SNP_7201
57.6 SNP_8545
71.7 SNP_8172
76.3 SNP_7866
87.2 SNP_7711
89.0 SNP_6882
90.3 SNP_8344
92.6 SNP_8741
93.1 SNP_8721
101.6 SNP_9070
104.4 SNP_9200
107.0 SNP_9116
114.2 SNP_9098
124.1 SNP_9230
4
0.0 SNP_9298
4.6 SNP_9390
8.4 SNP_9272
13.9 SNP_9696
17.6 SNP_9631
26.2 SNP_9344
33.8 SNP_9807
47.4 SNP_9988
48.8 SNP_9913
51.4 SNP_10197
68.1 SNP_10485
74.0 SNP_10689
79.9 SNP_10312
83.5 SNP_9936
94.0 SNP_10736
97.2 SNP_11091
97.4 SNP_9962
98.9 SNP_10330
106.0 SNP_10286
107.4 SNP_10479
109.0 SNP_10794
SNP_10750
110.1 SNP_10060
5
0.0 SNP_5103
2.2 SNP_11297
4.4 SNP_11358
8.7 SNP_11554
19.7 SNP_11739
21.2 SNP_11451
23.5 SNP_12070
30.7 SNP_12126
35.0 SNP_11923
59.3 SNP_12873
62.7 SNP_12975
67.9 SNP_12946
70.3 SNP_13176
82.7 SNP_12339
84.2 SNP_12381
84.9 SNP_13146
86.1 SNP_12763
87.2 SNP_13199
6
0.0 SNP_22377
9.0 SNP_22591
17.5 SNP_23149
23.5 SNP_22520
26.8 SNP_22343
29.1 SNP_22278
41.4 SNP_14281
43.7 SNP_14490
55.4 SNP_14523
67.6 SNP_14973
79.1 SNP_14864
81.5 SNP_15713
89.1 SNP_15500
96.3 SNP_15029
99.8 SNP_15298
101.9 SNP_16324
105.4 SNP_16661
110.3 SNP_15936
110.8 SNP_5318
126.9 SNP_16566
7
0.0 SNP_12883
3.8 SNP_17345
21.4 SNP_16802
31.4 SNP_16967
32.6 SNP_17771
38.5 SNP_17816
44.1 SNP_17795
48.4 SNP_18849
52.4 SNP_18091
59.7 SNP_19380
75.5 SNP_18727
85.0 SNP_18345
90.9 SNP_19289
91.9 SNP_19461
103.0 SNP_19675
103.4 SNP_18472
103.8 SNP_18728
106.1 SNP_19001
112.2 SNP_20125
115.6 SNP_18358
8
0.0 SNP_13411
0.3 SNP_13856
0.9 SNP_13974
1.4 SNP_13859
4.3 SNP_13833
5.5 SNP_13710
12.0 SNP_13579
15.5 SNP_13531
16.1 SNP_13597
18.3 SNP_13347
25.0 SNP_14192
27.9 SNP_14071
36.3 SNP_23980
37.4 SNP_24353
38.9 SNP_24201
40.9 SNP_24205
54.0 SNP_24335
77.0 SNP_24341
85.4 SNP_25122
92.5 SNP_24987
95.9 SNP_25015
102.0 SNP_25146

Figure 3.10 QTL analysis of leaflet number and rosette leaf number (RLN). Logarithm of the odds (LOD) profiles are shown against genetic linkage map. Dashed lines indicate significance determined by permutation testing where P = 0.05 (lower line or where there is only one line) and P = 0.01 (higher line). Red arrow heads indicate the direction of the QTL effect; upwards: the Azores allele increase trait value relative to the Oxford allele, downwards: the Azores allele decreases trait value relative to the Oxford allele.


Leaf 2
Leaf 3
Leaf 4
Leaf 5
Leaf 6
Leaf 7
Leaf 8
Leaf 9
Leaf 10
RLN
LOD score
Linkage group
20cM

all traits from 7.3% for the QTL detected on Linkage group 3 for leaflet number on leaf two to 39.6% for a QTL affecting rosette leaf number on linkage group 6 (Table 3.4). There were some QTL that influenced leaflet number on different leaves which is not surprising since leaflet number on different leaves correlated with each other (Figure 3.8). These results suggest that the correlation is due in part to common genetic determinants.

3. 7. 3. QTL mapping - Multi-trait analysis

Multi-trait QTL mapping was also performed using Genstat and the results are presented in Figure 3.11. Here the model for QTL detection and effect estimation is based on the correlations between different traits to provide a more realistic and dynamic view of the genetic determinants underlying trait variation.

The results demonstrate that leaflet number and rosette leaf number are influenced by many of the same QTL as there is only one QTL that affects leaflet number independently of rosette leaf number (linkage group 1 QTL). In this instance the multi-trait QTL analysis has been effective in displaying how the QTL affect different parts of the heteroblastic series; the QTL on Linkage groups 3 and 4 only effect the early leaves, the QTL on linkage group 6 and 8a do not affect the earlier leaves and the QTL on linkage groups 2 and 8b affect a larger part of the heteroblastic series (Figure 3.11). From this we can surmise that the effects of these QTL are affected by, or are influencing, heteroblastic development.

In five out of six instances a QTL that affects rosette leaf number also affects leaflet number at some stage of development. However, there is a lack of consistency in the direction of QTL effects on each respective trait. For example QTL on linkage groups 3 and 4 both influenced leaflet number and rosette leaf number in the same directions (increasing and decreasing respectively) but QTL detected on linkage groups 6 and 8 influenced these traits in opposite directions (Figure 3.11).

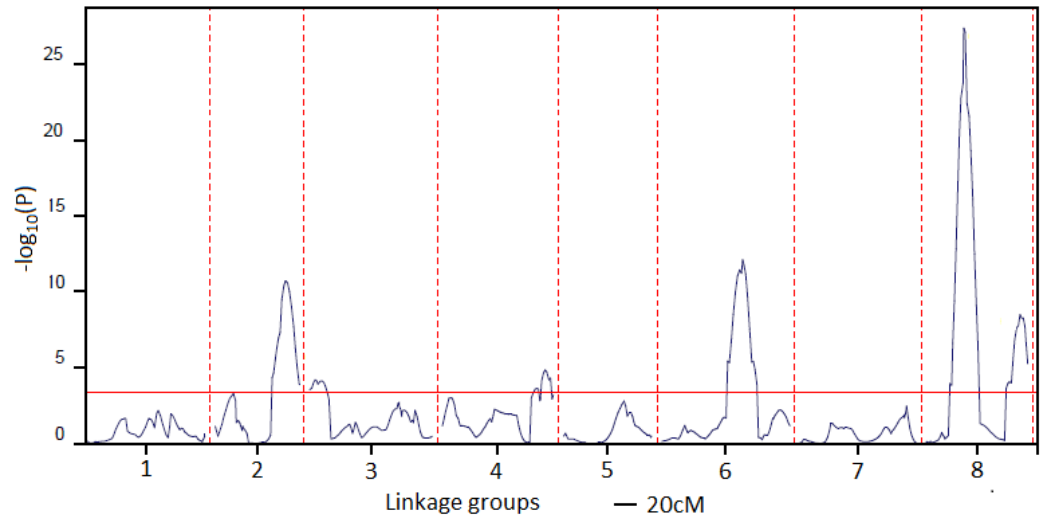
The results also show that both the Oxford and Azores strains possess QTL alleles that both increase and decrease trait values – neither strain has QTL alleles that just increase or decrease leaflet number.

Table 3.4 QTL analysis of leaflet number on leaves 2 -10 and rosette leaf number. QTL where the LOD score exceeded the P = 0.01 threshold are shown with their position on the genetic linkage map, the trait variance explained and predicted effect. These results are from initial single trait analysis.

	Trait	Number of QTL with significance p<0.01 for trait	Linkage group	Position on genetic linkage map (+/- 95 % C.I.) (cM)	Trait variance explained by QTL (%)	Predicted QTL effect expressed as trait value relative to Oxford strain
Leaflet number	*Leaf 2*	3	2	67.7 (+/- 2.1)	12.366	0.248
			3	6 (+/- 3.4)	7.263	0.204
			4	110.1 (+/- 1.9)	8.588	-0.203
	Leaf 3	1	4	105 (+/- 6.0)	15.867	-0.25
	Leaf 4	2	3	0.6 (+/- 2.1)	9.607	0.181
			4	107.4 (+/- 3.7)	8.326	-0.164
	Leaf 5	0	-	-	-	-
	Leaf 6	2	3	33.5 (+/- 6.4)	17.66	0.212
			8	91 (+/- 6.5)	25.44	0.262
	Leaf 7	0	-	-	-	-
	Leaf 8	1	8	91 (+/- 5.0)	12.86	0.293
	Leaf 9	0	-	-	-	-
	Leaf 10	0	-	-	-	-
Rosette leaf number		3	2	20.4 (+/- 1.3)	22.76	0.964
			6	79.1 (+/- 2.9)	39.62	-1.317
			8	38 (+/- 1.5)	10.53	2.118

Figure 3.11 Multi-trait QTL mapping results Genome scan (A)–$\log_{10}$ transformed p-values are plotted against the genetic linkage map, the redline indicates threshold value for significant QTL ($P < 0.05$). QTL effects are shown in table (B) where the values indicate the effect the QTL has on the trait relative to the Oxford parent, the direction and magnitude of effects are also indicated by colour, only significant effects are shown ($p < 0.05$). LG: linkage group

A



B

Leaflet number	Trait	QTL effect on trait					
		LG2	LG3	LG4	LG6	LG8a	LG8b
	Leaf 2	0.26	0.19	-0.22	-	-	0.13
	Leaf 3	-	0.18	-0.24	-	-	0.11
	Leaf 4	0.15	0.18	-0.17	-	-	0.13
	Leaf 5	0.18	0.19	-	-	-	0.15
	Leaf 6	0.27	0.17	-0.10	-	-0.11	0.20
	Leaf 7	0.18	0.12	-	-	-0.19	0.18
	Leaf 8	0.27	0.15	-	-	-0.24	0.28
	Leaf 9	0.24	-	-	-	-0.26	0.26
	Leaf 10	0.23	-	-	0.12	-0.24	0.27
	RLN	-	0.44	-0.45	-1.50	2.21	-1.26

Azores allele decreases trait value

Azores allele increases trait value

3. 8. QTL mapping - Discussion

In this chapter we present a QTL analysis for leaflet number and rosette leaf number in a *C. hirsuta* F6 RIL population derived from a cross between the Oxford and Azores strains. Using composite interval mapping I have detected and located QTLs involved in the production of differing number of lateral leaflets in the first 10 leaves. Multi-trait QTL mapping was then used to improve the analysis by fitting the QTL effects to the genetic correlations amongst the traits. This allowed us to see the variable effect of QTL along the heteroblastic series – the developmental range in which QTL are active was very different.

Five QTL were mapped that influence both leaflet number and rosette leaf number, a single QTL was mapped that affected leaflet number independently of rosette leaf number and no QTL were detected that influenced rosette leaf number independently of leaflet number. All QTL effecting leaflet number had a significant effect at more than one leaf node but the range of influence QTL had across the heteroblastic series was variable.

These results demonstrate that variable leaflet number and rosette leaf number are regulated by multiple pleiotropic loci. This is not unexpected since it was predicted that the genetic determinants of leaf traits would have a role in phase change and that QTL promoting mature leaf morphology (increased leaflets) because they accelerate phase change so that bolting and the onset of reproductive growth come earlier when less rosette leaves had been developed. This prediction holds for the three QTL on linkage groups six and eight (both 8a and 8b) (Figure 3.11). However, in contrast, the QTL on linkage groups 3 and 4 do not conform to this prediction; the alleles associated with adult leaf morphology (increased leaflets) at these QTL do not bring about earlier transitions to reproductive growth, they actually promote production of more rosette leaves (Figure 3.11). A possible explanation for this observation is that early transition through the heteroblastic development is not always coupled to the transition to reproductive growth. More work is required to see how flowering time and length of the vegetative phase correlates to rosette leaf production to get a more comprehensive explanation, unfortunately difficulties

associated with poor germination rates and over-heating problems with the controlled environment rooms hampered my plastochron experiments.

It was interesting to note that alleles from either parent did not affect leaflet number in the same direction at every QTL, in other words neither of the parental strains possess QTL alleles that either increase or decrease leaflet number. This could suggest that strong directional selection on leaflet number has not occurred in the establishment of either of the Oxford or Azores populations in their respective habitats. Had leaflet number been under the influence of such selection then it would be expected that the genome of one strain would only have QTL alleles working to maximise or minimise leaflet production according to the direction of selection.

Previously in this chapter comparison of the parental strains revealed that Oxford had consistently more leaflets than Azores across most of the heteroblastic series. However, for four of the six QTL detected the allele that increases leaflet number has come from the Azores parent. This would suggest that there are additional QTL to be detected that have been missed in this experiment, I would expect that the Oxford allele would increase leaflet production at the majority of these missing QTL in order to account for the observed variation between the parents.

The QTL detected here will provide the basis for future studies aimed at the elucidation of the underlying molecular basis for variable leaflet number.

Chapter 4:

QTL validation, fine mapping and characterisation

4. 1. Introduction

Quantitative trait loci (QTL) are not mapped to a high resolution, so only regions containing QTL can be identified and these can cover tens of cM. In small genomes like that of *Arabidopsis* a single cM can equate to 120kb on average (Alonso-Blanco et al. 2005), but this figure varies depending on whether the QTL is mapped to euchromatic or heterochromatic regions within the genome. These large intervals make identifying the causal gene underlying the QTL difficult. To make this task easier the QTL interval must be reduced by fine mapping. Fine-mapping requires the use of populations whose genomes differ only at the QTL interval.

The most common populations used for this purpose are near isogenic lines (NILs) and heterozygous inbred families (HIFs). NILs are lines containing a single (or a small number of) genomic introgression fragment from one parental strain into an otherwise homogenous genetic background from the other parental strain, they are created through repeated backcrossing and marker assisted selection. HIFs are derived from RILs which, despite the repeated generations of self-fertilisation, have retained some residual heterozygosity around the QTL region. In both cases only the region around the QTL will segregate in the subsequent generation, if the phenotype co-segregates with the genotype then it is confirmed that the QTL lies within that region - the QTL has been validated.

In fine mapping, individuals from the above populations are selected where a recombination event has occurred within the segregating region of a validated QTL. These events allow for the QTL to be validated in smaller genomic regions, thus making it easier to identify the exact genes causal to the QTL. A further generation may be required to obtain the desired recombinants.

Before recombinant lines for fine mapping can be selected it is first required that the QTL is validated, using either NILs or HIFs. Within QTL mapping there is an inherent risk of false positive QTL detection especially with QTL of small effect. Identifying QTL effect in NILs or

HIFs can be made problematic by epistatic interactions (e.g. Keurentjes et al., 2007) and non-genetic variation. For these reasons QTL validation is a pre-requisite of fine mapping.

To avoid the analysis of a large number of replicates the QTL effect being validated needs to be of sufficient size and heritability. QTL of large effect are much easier to validate and fine map than QTL with lesser effects; this is reflected in the number of small and large effect QTL that have been fine mapped to a resolution fine enough to identify the causal gene (Kroymann and Mitchell-Olds, 2005).

QTL validation is an essential procedure before fine mapping can occur. A clear segregation of phenotypes will indicate whether fine mapping is feasible and will provide a more accurate estimate of QTL effect than is provided by QTL mapping results. A better idea of the size of QTL effect will inform the experimental design required during fine mapping in terms of the population size required.

In this chapter I attempt to validate QTL responsible for variable leaflet number with a range of effects. I have used HIFs selected from the mapping population of RILs. On average the RILs retained 2.5% heterozygosity following six generations of self-fertilisation. I was able to select the lines where this residual heterozygosity overlapped with the region containing a QTL. These lines became the HIFs and the QTL validation was performed by comparing the mean leaflet number of the progeny which were homozygous for one parental allele at the QTL to the leaflet number of the progeny which were homozygous for the other parental allele at the QTL.

The QTL detected on linkage groups 2, 3, 4 and 8 could be validated by using HIFs thus confirming the success of the QTL mapping. However, the QTL effects validated were not exactly as predicted (Figure 3.11 B); typically; QTL effects were observed across a wider breadth of the heteroblastic series than was expected. It was also clear from the validation experiments that variability between experiments has a large influence over whether the QTL effect is observed. The success of this validation means that it is now possible to begin fine mapping these

QTL and take the first steps towards identifying the underlying genes. The QTL on linkage group 4 (QTL-LG4) was fine mapped to a DNA segment of 48.6 kb.

I have made use of the lack of genetic variation inherent in HIFs to fully characterise the effect of QTL-LG4 on other aspects of leaf development. The QTL was mapped and validated on the basis of leaflet number and rosette leaf number data but it is likely that the QTL has pleiotropic effects on other leaf shape traits. In order to explore this, the leaves of HIFs with an alternative genotype at the QTL region were analysed. Concurrently with this the effect of variable photoperiod on QTL effect was analysed by growing the same HIFs in short day conditions. Heteroblastic development is influenced by many external stimuli including photoperiod (Jackson, 2009) and an analysis of QTL effect in different photoperiods will show whether this QTL is influenced by day length.

4. 2. QTL validation

QTL to be validated were chosen to capture a range of QTL effects that operated in different directions and had different levels of pleiotropy. HIFs were selected by scanning the genotype of each RIL to assess its suitability as an HIF to validate a particular QTL. I aimed to get HIFs that had residual heterozygosity at the QTL being validated but with minimal heterozygosity elsewhere. Leaflet number was recorded for each genotype and used to assess co-segregation of phenotype and genotype. In addition to leaflet number on each individual leaf, total leaflets on the number of leaves held by the plant that produced the fewest leaves were used to look for QTL effect.

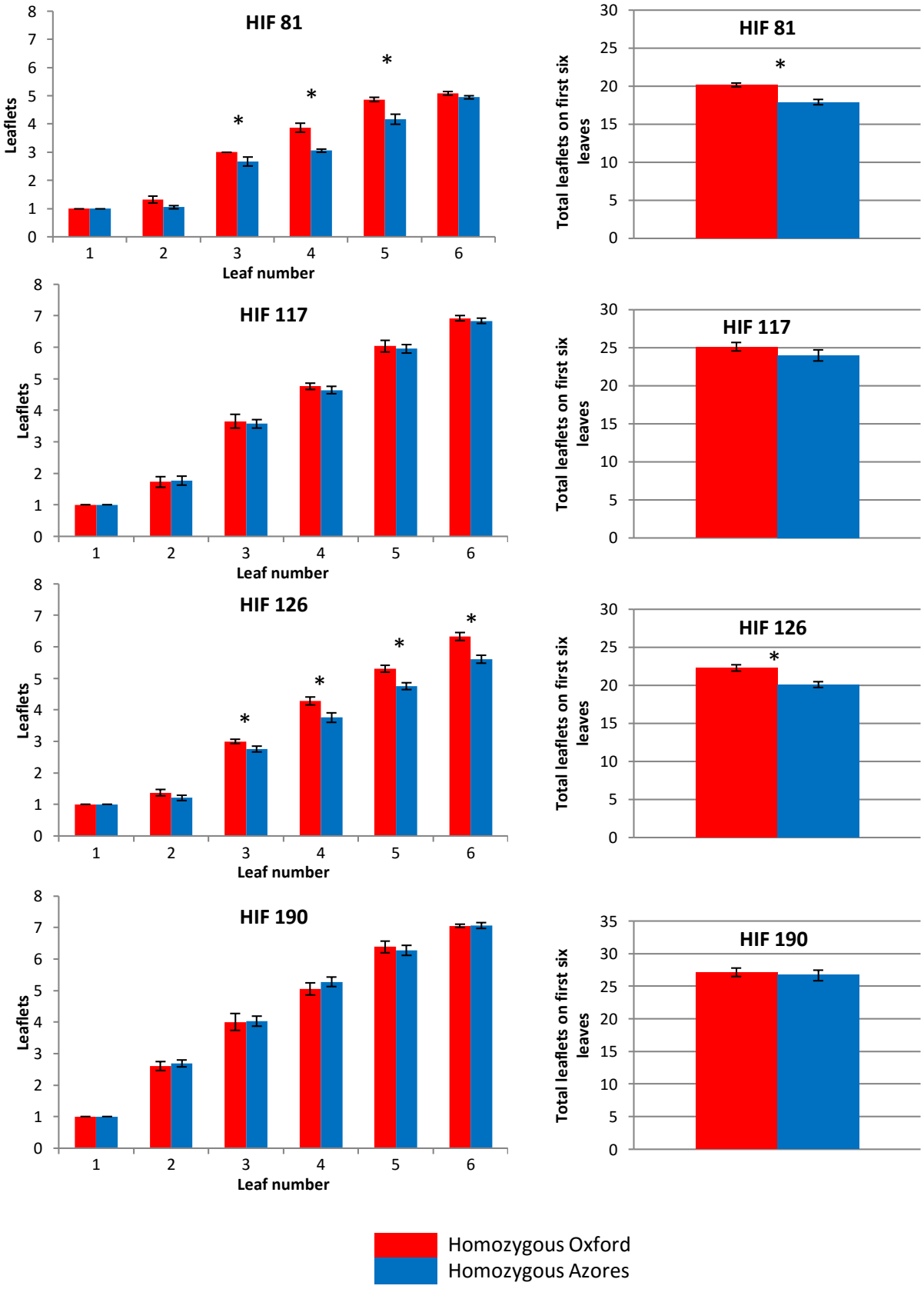
4. 2. 1. Validation of QTL-LG4

A QTL of relatively large effect was mapped to the lower portion of linkage group 4 (QTL-LG4) at about 107.4 cM. It has an effect on leaflet number on leaves 2, 3, 4 and six as well as rosette leaf number (Figure 3.11).

Four RILs to become HIFs were selected to validate this QTL which had heterozygous regions around the QTL location – these were HIFs 81, 117, 126 and 190. None of the HIFs were completely free of other areas of heterozygosity in their genomes: HIF 81 had areas of heterozygosity near the middle of linkage group 1 and the tip of linkage group six; HIF 117 had areas of heterozygosity near the middle of linkage group 3 and the base of linkage group 7; HIF 126 has residual heterozygosity at the top of linkage group 5 and 8; and HIF 190 has heterozygous regions at the top of linkage group 1 and the middle of linkage group 3. All of these additional areas of residual heterozygosity were judged to be far enough away from genomic regions predicted to play a role in variable leaflet number. 45 seeds per HIF were sown and these were genotyped using PCR- based markers designed around SNPs. Due to time constraints, leaflet number was captured only for the first six or seven leaves and at this stage we were not able to see if the QTL had any subsequent effect on leaves or on rosette leaf number.

HIF 81 was able to validate the QTL (Figure 4.1). There were significant differences between plants that were homozygous for the Oxford allele and plants homozygous for the Azores allele at leaves 3, 4 and 5. When the numbers of leaflets were summed for the first six leaves there was also a significant difference (Figure 4.1). The QTL was also validated with the HIF 126 (Figure 4.1). There were significant differences in the number of leaflets between the alternative genotypes at leaves 3, 4, 5 and 6 and in total leaflets for the first six leaves (Figure 4.1). The Oxford and Azores genotypes of HIFs 117 and 190 did not display significant differences in leaflet number at any of the leaves and there was no significant difference in the total number of leaflets on the first six leaves. Thus the QTL was not validated in either of these HIFs.

Figure 4.1 QTL-LG4 validation. Leaflets on individual leaves and total leaflets on first six leaves of HIFs homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).


HIF 81
Leaflets
Leaf number
Total leaflets on first six leaves
HIF 117
HIF 126
HIF 190
Homozygous Oxford
Homozygous Azores

4. 2. 2. Validation of QTL-LG2

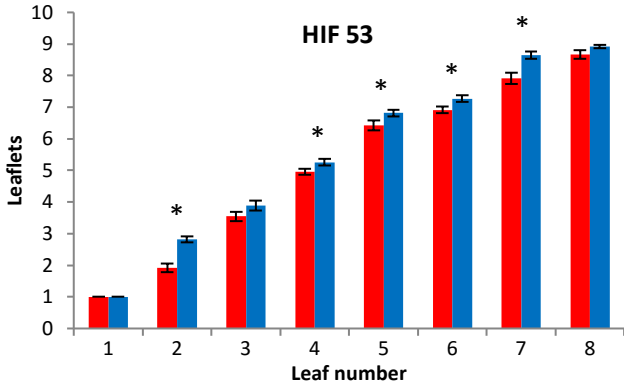
A QTL affecting leaflet number was mapped to the base of linkage group 2 (QTL-LG2). The analysis placed the QTL at around 67 cM; the QTL affects leaflet production on leaves 2 to 8 and rosette leaf number (Figure 3.11)

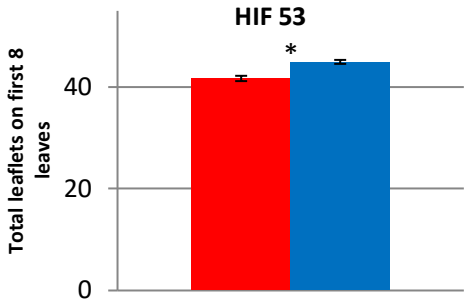
On this occasion three RILs were selected to validate the QTL: these became HIFs 53, 102 and 112. Prior to selecting these I ensured there was minimal background genetic variation created from other area of residual heterozygosity. However for HIF 53 there were heterozygous markers at the base of linkage group 1, the middle of linkage group 3 and the middle of linkage group 5. HIF 102 had a single heterozygous marker in the upper half of linkage group 1. According to the QTL mapping none of these areas are predicted to affect leaflet number. The genome of HIF 112 was free from heterozygous regions away from the QTL under consideration. Leaflet number was recorded on the first eight leaves.

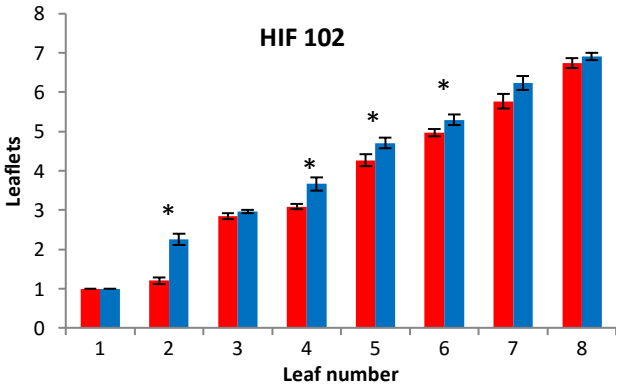
HIF 53 plants that were homozygous for the Azores allele had significantly more leaflets than plants that were homozygous for the Oxford allele at leaves 2, 4, 5 , 6 and 7 as well as greater total number of leaflets on the first eight leaves (Figure 4.2). HIF 102 showed similar results to HIF 53; plants homozygous for the Azores allele had a greater number of leaflets than those homozygous for the Oxford allele at leaves 2, 4, 5 and 6 as well as a greater total number of leaflets on the first eight leaves (Figure 4.2). For HIF 112 the reverse situation appeared to be true; plants homozygous for the Oxford allele had significantly greater leaflets than plants homozygous for the Azores allele on leaves 5, 6 and 8 as well as more leaflets altogether on the first eight leaves (Figure 4.2).

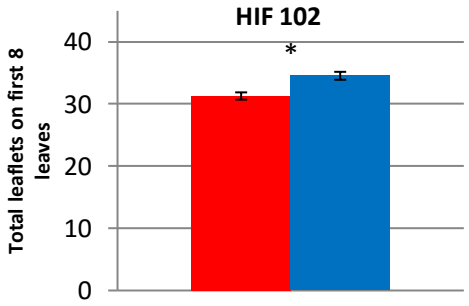
While HIF 53 and HIF 102 validate the QTL as expected per the QTL mapping results (Figure 3.11), the QTL effect in HIF 112 has been reversed. In an attempt to repeat these results, seeds harvested from the homozygous plants were sown. 30 seeds per genotype for all three HIFs were sown and grown in exactly the same conditions as they were previously. On this occasion leaflet number was recorded for every rosette leaf produced before bolting.

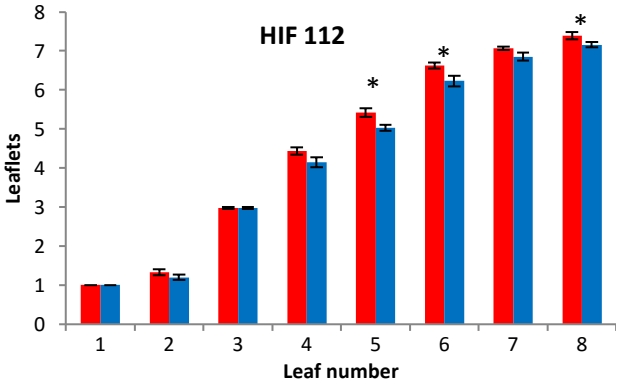
Figure 4.2 Linkage group 2 QTL validation. Leaflets on individual leaves and total leaflets on first eight leaves of HIFs homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).

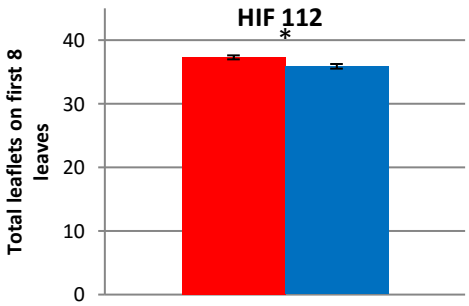

HIF 53
Leaflets
Leaf number
0
1
2
3
4
5
6
7
8
9
10
1
2
3
4
5
6
7
8
*


HIF 53
Total leaflets on first 8 leaves
0
20
40
*


HIF 102
Leaflets
Leaf number
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
*


HIF 102
Total leaflets on first 8 leaves
0
10
20
30
40
*


HIF 112
Leaflets
Leaf number
0
1
2
3
4
5
6
7
8
1
2
3
4
5
6
7
8
*


HIF 112
Total leaflets on first 8 leaves
0
10
20
30
40
*

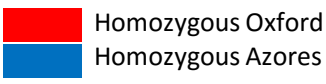
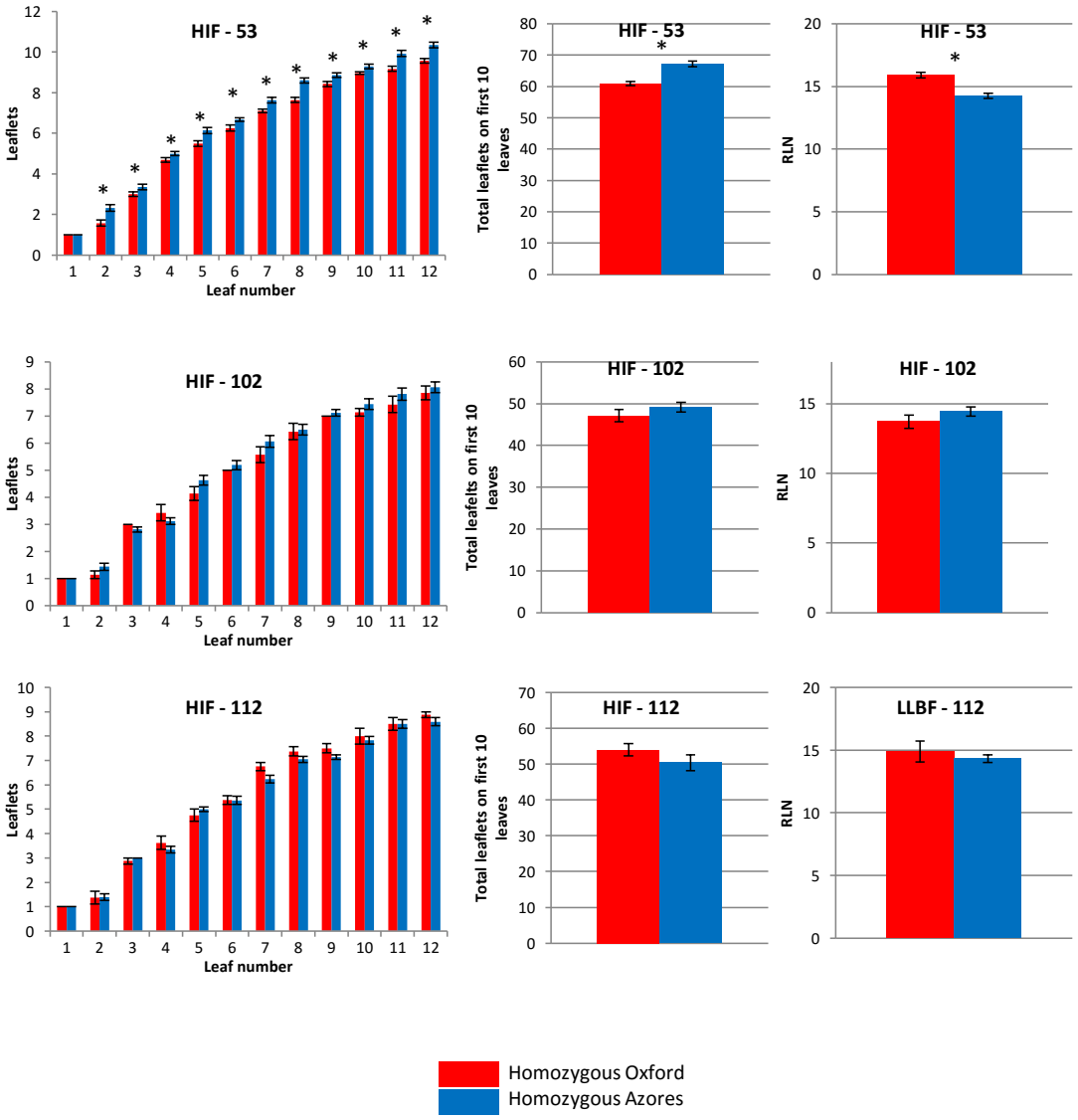

Homozygous Oxford
Homozygous Azores

Figure 4.3 Linkage group 2 QTL validation, second experiment. Leaflets on individual leaves, total leaflets on first 10 leaves and rosette leaf number (RLN) of HIFs homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).


HIF - 53
Leaflets
Leaf number
Total leaflets on first 10 leaves
RLN
HIF - 102
HIF - 112
LLBF - 112
Homozygous Oxford
Homozygous Azores

As before the QTL was validated for HIF 53. Plants homozygous for the Azores allele produced significantly more leaflets than those homozygous for the Oxford allele at leaves 2 to 12 and for the total number of leaflets on the first 12 leaves (Figure 4.3). Since rosette leaf number was also recorded we were able to see if the QTL has an effect on this trait. Plants homozygous for the Oxford allele produced significantly more rosette leaves than those homozygous for the Azores allele.

There were no significant differences between the two genotypes in either leaflet number or rosette leaf number for HIFs 102 or 112. However this analysis may be compromised because of poor germination in this instance, especially with the seeds homozygous for the Oxford allele.

The results presented here display that there is a QTL which has been validated with HIF 53, the effect of which were predicted with QTL mapping; this was result was also repeatable. However in the case of HIF 102 the QTL affect was reversed with the Oxford allele conferring more leaflets than the Azores. This result has not been confirmed however. HIF 53 was able to validate the QTL in both experiments, but there was still variation between the experiments. Such variability can also be expected between RIL replicates in the QTL mapping highlighting how the prediction of QTL effects with QTL mapping models is inherently imprecise.

4. 2. 3. Validation of QTL-LG3

A QTL mapped to the top end of linkage group 3 (QTL-LG3) at around 0-6 cM where the Azores allele had the effect of increasing leaflet number on leaves 2 to 8 and decreasing rosette leaf number (Figure 3.11)

Three suitable HIFs were selected from the RILs to validate this QTL – HIFs 70, 74 and 147. All three HIFs had areas of heterozygosity in their genomes away from the QTL; HIF 70 had heterozygous markers at the top of linkage groups 2 and 4; HIF 74 in the middle of linkage

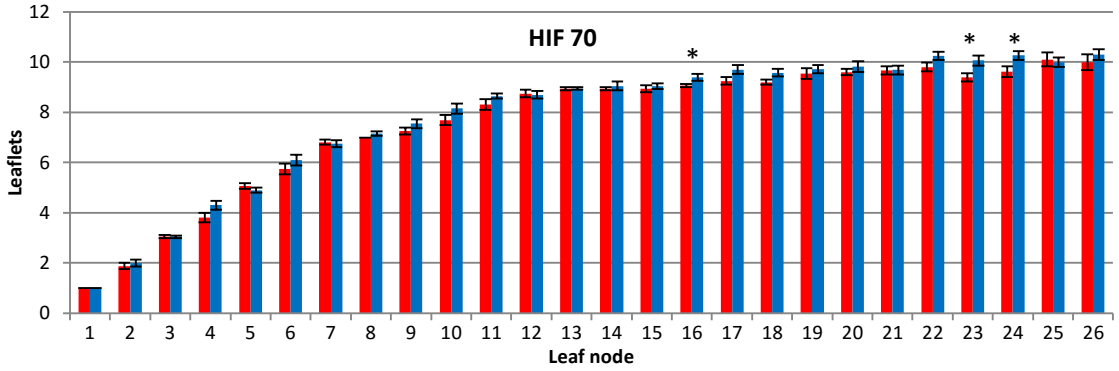
groups 1 and 4 and at the top of linkage group 6 and HIF 147 at the top of linkage groups 4 and 6 as well as at the base of linkage group 6.

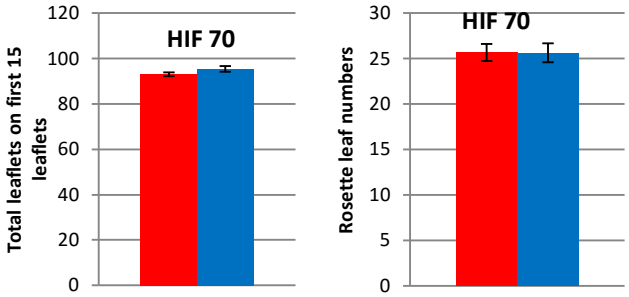
Unfortunately poor germination rate in relation to HIF 74 impeded validation. With HIF 70 the only significant differences in leaflet number between homozygotes for the Oxford allele and homozygotes for Azores allele occurs at leaves 16, 23 and 24 (Figure 4.4). Additionally there was no significant difference in total leaflet number for the first 15 leaves, neither was there a significant difference in the number of rosette leaves produced. These results do not provide sufficient evidence to conclusively validate the QTL in HIF 70.

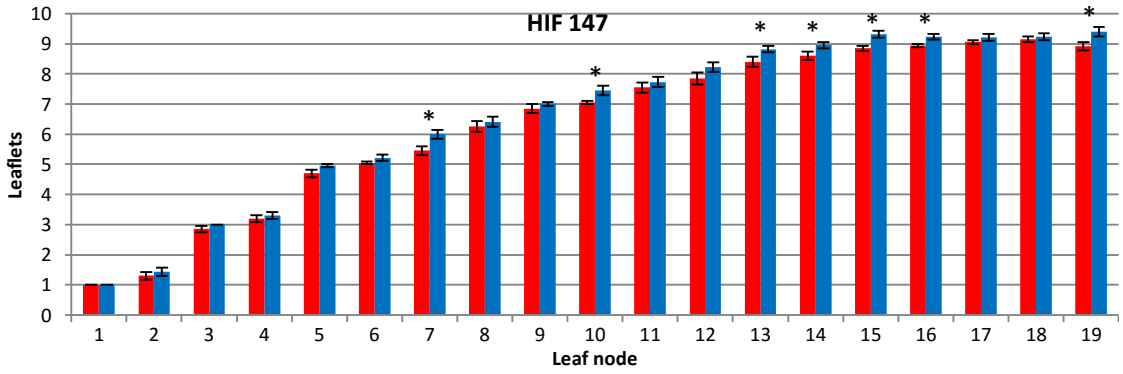
The alternate homozygotes produced by HIF 147 showed some significant differences in leaflet number at leaves 7, 10, 13, 14, 15, 16, 19 and 20 but there was no significant difference in the total number of leaflets produced on the first 15 leaves (Figure 4.4). Rosette leaf number did not vary significantly between the alternate homozygotes (Figure 4.4). These results are suggestive, if not conclusive, evidence that the QTL is validated here. The QTL mapping predicted larger differences in the early growing leaves than that observed here. It may be the case that background variation, from either the environment or other segregating areas of the genome, has obscured the QTL effect which was relatively small (Figure 3.11)

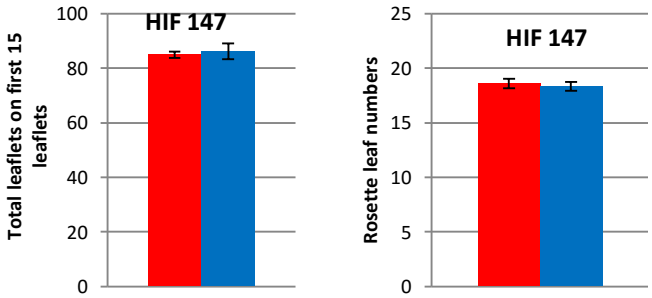
From this evidence, it cannot be conclusively stated that the LG3 QTL is validated. Since fine mapping requires repeated validation in multiple alternate recombinant lines, further fine mapping of this QTL would be problematic.

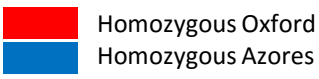
Figure 4.4 Linkage group 3 QTL validation. Leaflets on individual leaves and total leaflets on first 15 leaves and rosette leaf number of HIFs homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. RLN: Rosette leaf number. Significant differences are noted by an asterisk ($P<0.05$).


HIF 70
Leaflets
Leaf node
0
2
4
6
8
10
12
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
*
*
*


HIF 70
Total leaflets on first 15 leaflets
0
20
40
60
80
100
120
HIF 70
Rosette leaf numbers
0
5
10
15
20
25
30


HIF 147
Leaflets
Leaf node
0
1
2
3
4
5
6
7
8
9
10
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
*
*
*
*
*
*
*


HIF 147
Total leaflets on first 15 leaflets
0
20
40
60
80
100
HIF 147
Rosette leaf numbers
0
5
10
15
20
25


Homozygous Oxford
Homozygous Azores

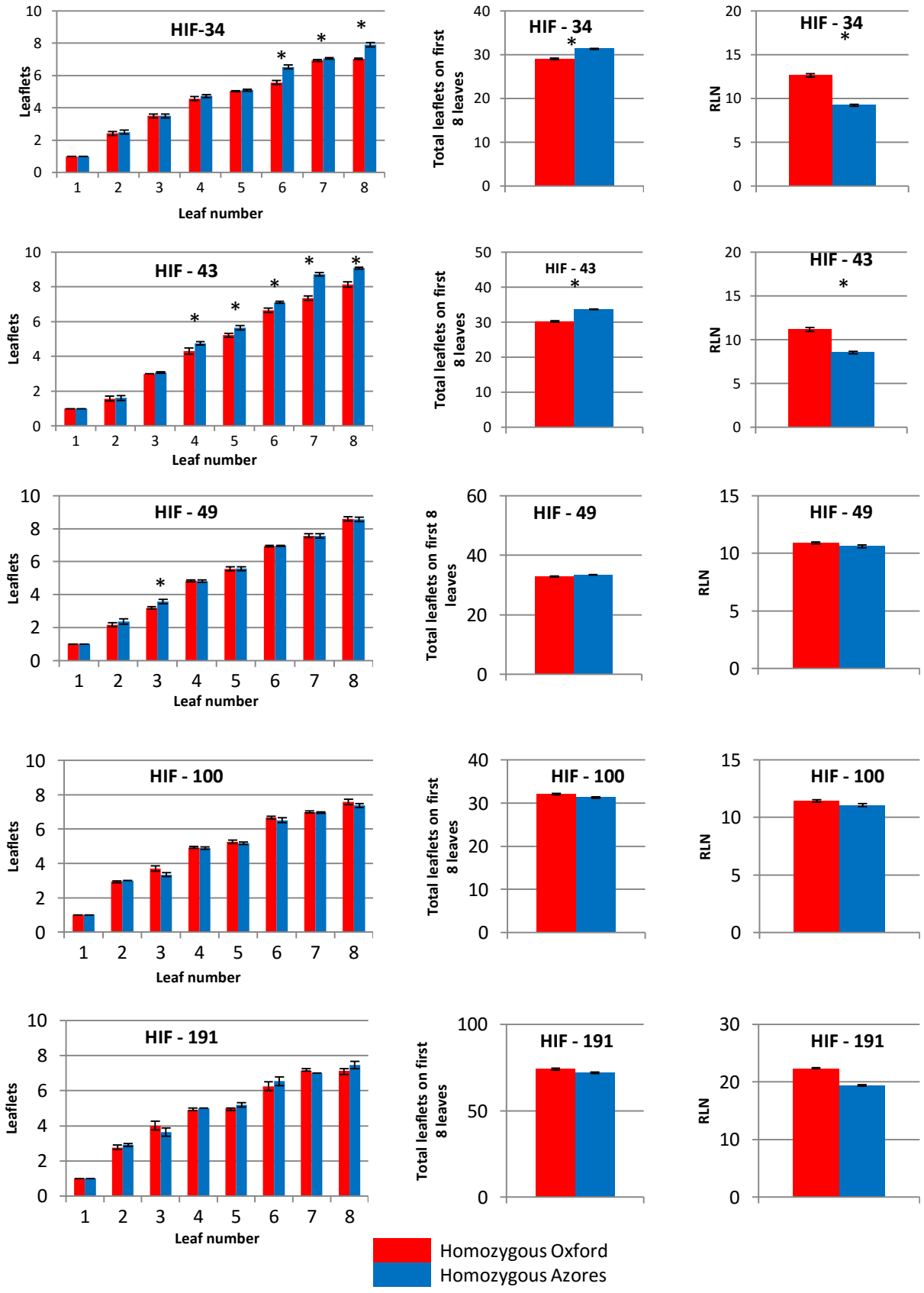
4. 2. 4. Validation of QTL-LG8

There were two QTL affecting leaflet number that mapped to linkage group 8; here I tried to validate the lower most QTL on the chromosome (LG8b from Figure 3.11). The Azores allele is predicted to increase leaflet number on all leaves analysed and decrease rosette leaf number.

Five RILs were used to create HIFs to validate the QTL. Background genetic variation was minimised by selecting HIFs with little no heterozygous regions other than at the QTL region. HIF 34 had heterozygous regions at the top of linkage group 7, HIF 43 had heterozygous regions at the top of linkage group 8, HIF 49 had heterozygous regions at the top of linkage group 1 and the base of linkage group 6, HIF 100 had heterozygous regions at the top of linkage groups 2, and 8 and HIF 191 had heterozygous regions at the top of linkage groups 2, 4 and 6, and at the base of linkage group 5 and 7. The number of leaflets on all rosette leaves was recorded.

The different classes of homozygotes produced by HIF 34 did show significant variation in leaflet number at leaves 6 to 9 (Figure 4.5). The total sum of all leaflets on the first eight leaves also varied significantly as did the number of rosette leaves (Figure 4.5). HIF 43 also showed significant variation between plants of different genotypes: at leaf nodes 4 to 9 there were significantly more leaflets on plants homozygous for the Azores allele (Figure 4.5). This was reinforced when the total number of leaflets on the first seven leaves was examined. There was also a significant difference in rosette leaf number; the Oxford allele conferred a greater number of leaves relative to the Azores allele (Figure 4.5). For HIFs 49, 100 and 191 there was no evidence that phenotype was co-segregating with genotype (Figure 4.5). Thus, there is strong enough evidence that the QTL map on lower portion on linkage group is genuine and not an erroneous result provided by the QTL mapping analysis.

Figure 4.5 Linkage group 8 QTL validation. Leaflets on individual leaves,total leaflets on first eight leaves and rosette leaf number (RLN) of HIFs homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).


HIF-34
Leaflets
Leaf number
HIF - 34
Total leaflets on first 8 leaves
RLN
HIF - 43
HIF - 49
HIF - 100
HIF - 191
Homozygous Oxford
Homozygous Azores

4. 3. Fine mapping

LG4-QTL was chosen for further examination by fine mapping. This QTL was chosen because its effect is strong and easily visible in the early developing leaves which makes validation in multiple alternate recombinant lines more efficient, because fewer individuals are required to provide sufficient statistical power and the phenotype can be observed early on in development.

4. 3. 1. Delimiting QTL region

The advantages of validating the QTL with multiple HIFs is that it is possible to examine the overlapping residual heterozygous regions in the validating and non-validating lines to reduce the region in which the QTL is known to reside. Genotyping of additional markers between those of the linkage map made it possible to locate the recombination events more precisely and following this, the QTL region was reduced to a region of 705 Kb (maximum size), which equates to approximately 170 genes.

4. 3. 2. Searching for recombinant lines

Fine mapping requires the validation of QTL effect in smaller regions. These smaller regions are attained through recombination events within the region known to contain the QTL. In this instance HIF 126 (used to validate QTL-LG4, Figure 4.1) was used to generate lines that had recombination events within the QTL region. Initially a total of 720 seeds from HIF 126 heterozygous at the QTL, were sown in batches of about a 100 seeds. The growing conditions of these seeds were exactly the same as those used for the QTL validation. DNA extraction was performed on material collected from seedlings when they had grown 4-5 leaves and initially these were genotyped with markers from either end of the QTL region. When the genotype at

these markers differed from each other it was taken as an indication that recombination had occurred within that stretch of genome. Following the discovery of recombination events, their location was determined by genotyping with markers located within the validated QTL region.

Unfortunately there was poor germination with this batch of seeds and only 447 seedlings germinated. However, 41 recombinant lines were discovered amongst these. From these specific recombinants were selected to provide a distribution of recombinant break points within the QTL region (Figure 4.6). The recombinant lines 2, 9 and 12 were used to begin fine mapping. Validation was also attempted with recombinants 11 and 37. However poor germination and technical difficulties prevented their analysis.

4. 3. 3. QTL fine mapping first pass

During fine mapping a QTL can be assigned to smaller regions when they are validated in different recombinants. In the same way that the QTL were validated previously, I was looking for co-segregation of genotype and phenotype and thus, the experimental design and analysis was very similar in both circumstances.

With recombinant HIF 126.2 there was no evidence of phenotype and genotype co segregation; leaflet number did not vary significantly at any leaf and neither did the total number of leaflets on the first eight leaves (Figure 4.7). Recombinants HIF 126.9 and HIF 126.12 were able to validate the QTL; there was a significant difference in leaflet number at a number of leaf nodes as well as in total leaflet number for the first eight leaves.

Since the QTL was validated in two recombinant variants the QTL must lie within a genomic region that is heterozygous in both lines. This region equated to 110 Kb, a large reduction on the 705 kb in which the QTL was validated previously.

Figure 4.6 Fine mapping LG4-QTL. Graphical summaries of the recombinants of HIF 126. Each bar represent the genome around the QTL and the genotype is indicated by colour. Markers are indicated by black bands, those in bold are from the Oxford/Azores genetic linkage map (Figure 3.9). To the left are the recombinants used in the first round of fine mapping and to the right are those in the second round of fine mapping. The yellow bar indicates the maximum region in which the QTL must exist. The Purple bar indicates proposed position of putative second QTL. The black bar indicates scale.

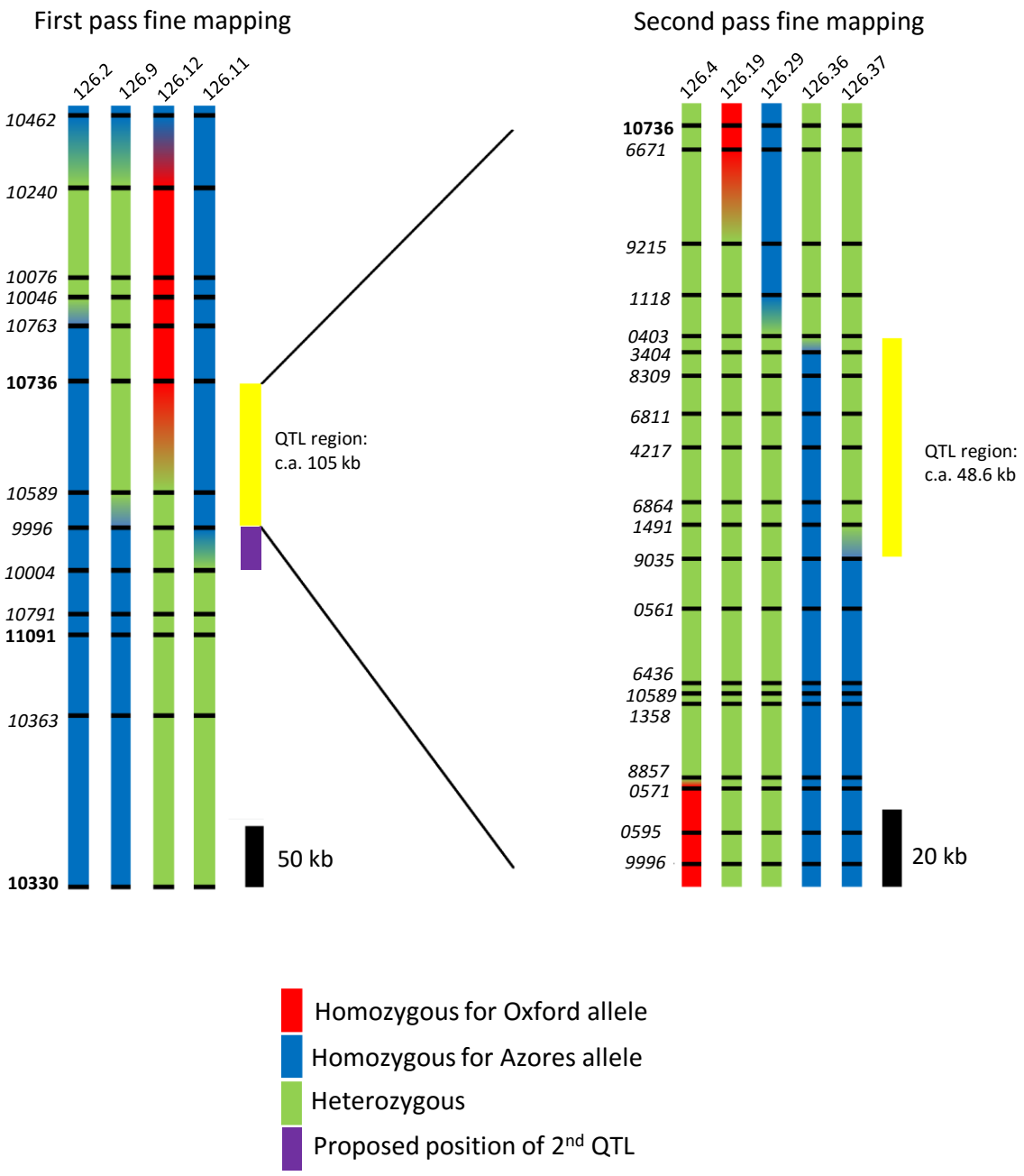
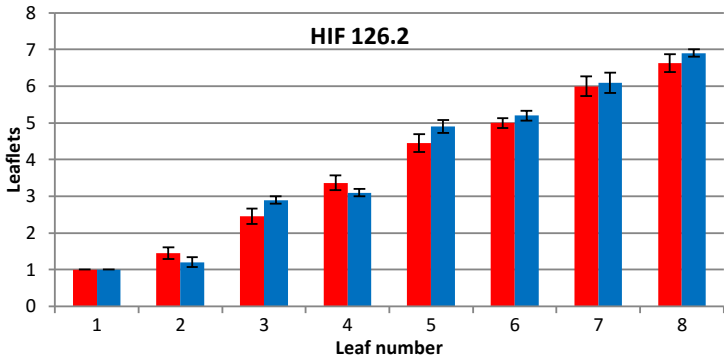
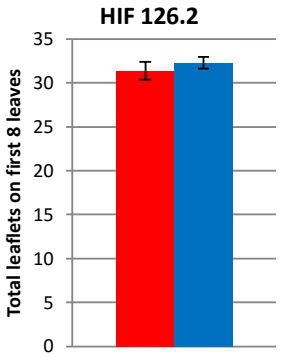
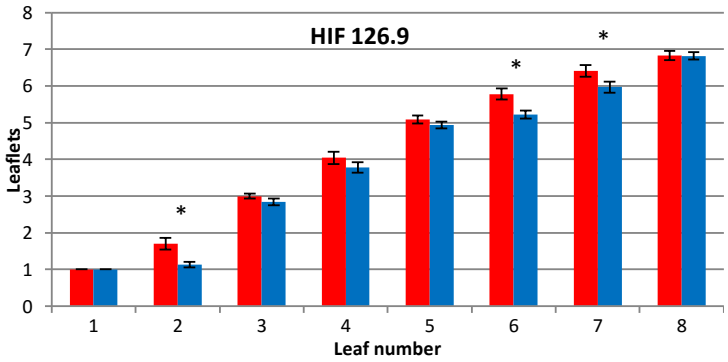
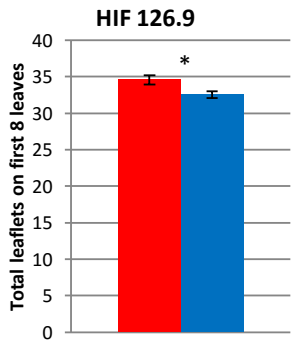

First pass fine mapping
126.2
126.9
126.12
126.11
10462
10240
10076
10046
10763
10736
10589
9996
10004
10791
11091
10363
10330
QTL region:
c.a. 105 kb
50 kb
Second pass fine mapping
126.4
126.19
126.29
126.36
126.37
10736
6671
9215
1118
0403
3404
8309
6811
4217
6864
1491
9035
0561
6436
10589
1358
8857
0571
0595
9996
QTL region:
c.a. 48.6 kb
20 kb
Homozygous for Oxford allele
Homozygous for Azores allele
Heterozygous
Proposed position of 2nd QTL

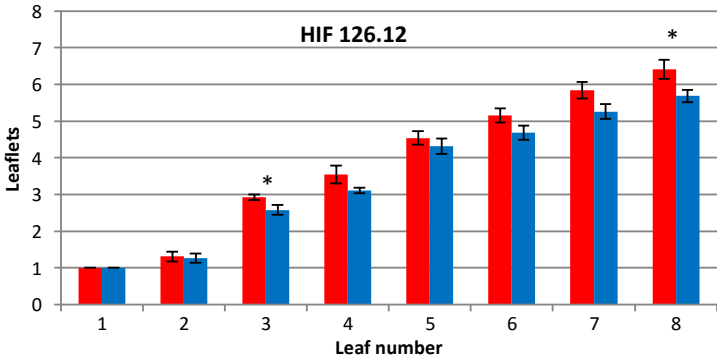
Figure 4.7 QTL-LG4 fine mapping – First pass. Leaflets on individual leaves and total leaflets on first eight leaves of recombinant variants of HIF 126 homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).

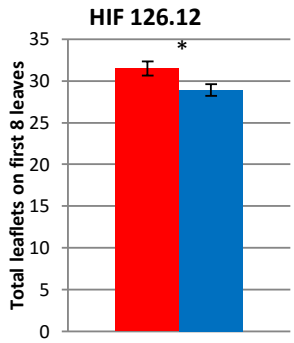

HIF 126.2
Leaflets
Leaf number
0
1
2
3
4
5
6
7
8


HIF 126.2
Total leaflets on first 8 leaves
0
5
10
15
20
25
30
35


HIF 126.9
Leaflets
Leaf number
*
*
*


HIF 126.9
Total leaflets on first 8 leaves
*


HIF 126.12
Leaflets
Leaf number
*
*


HIF 126.12
Total leaflets on first 8 leaves
*

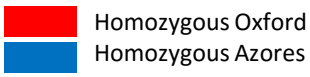

Homozygous Oxford
Homozygous Azores

Figure 4.8 QTL-LG4 fine mapping – Second pass. Leaflets on individual leaves , total leaflets on first 18 leaves and rosette leaf number of recombinant variants of HIF 126 homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).

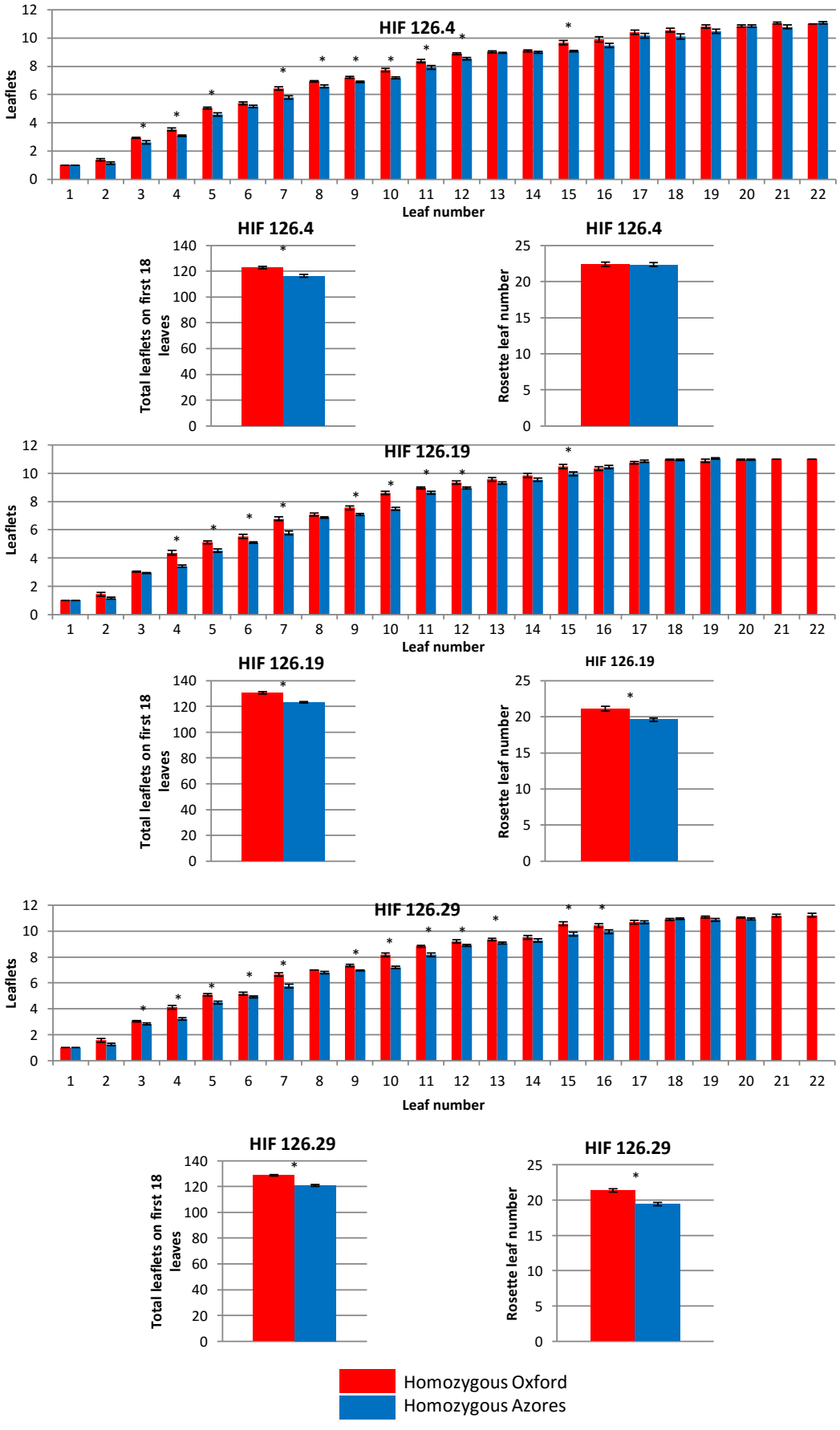
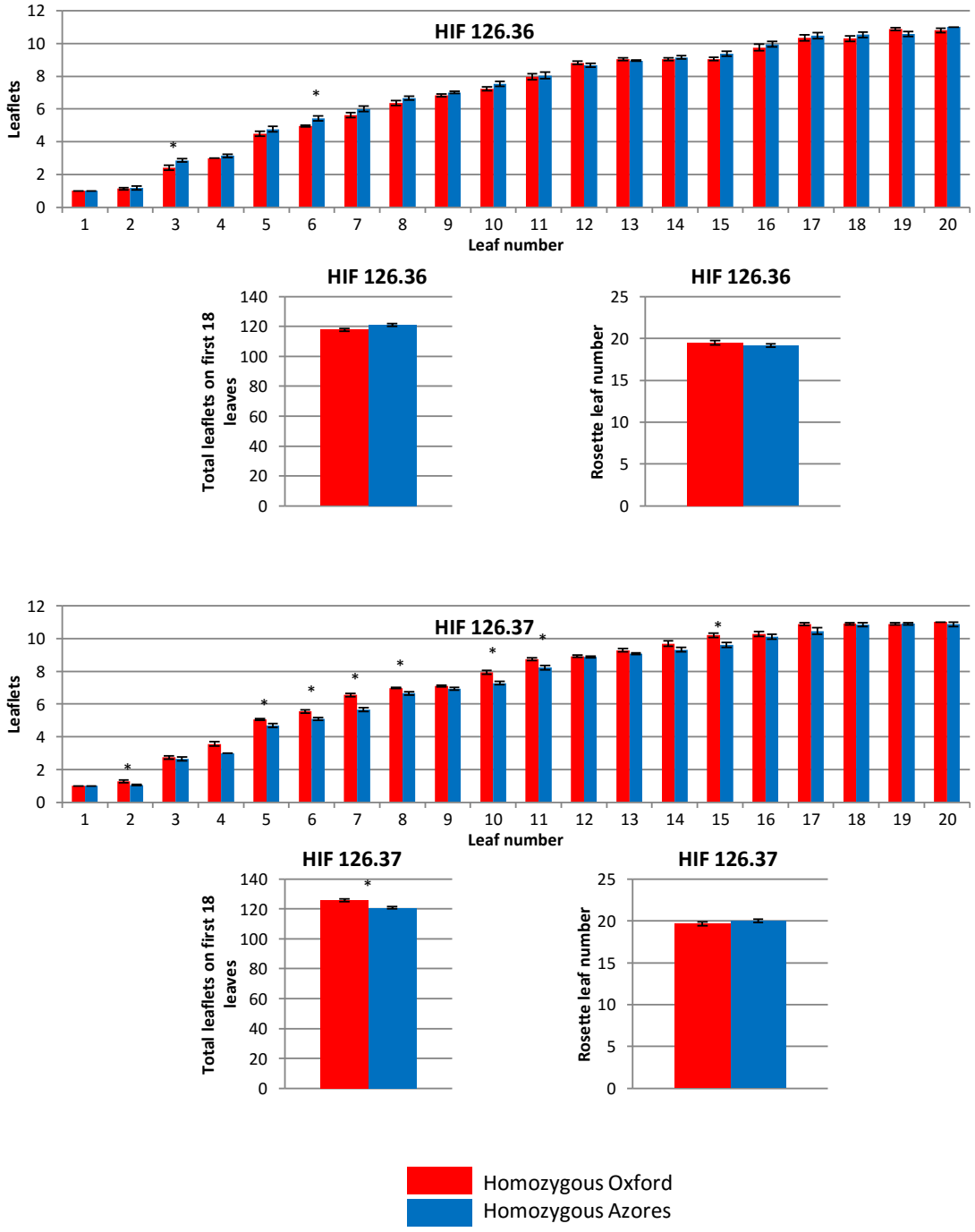

HIF 126.4
Leaflets
Leaf number
HIF 126.4
Total leaflets on first 18 leaves
HIF 126.4
Rosette leaf number
HIF 126.19
Leaflets
Leaf number
HIF 126.19
Total leaflets on first 18 leaves
HIF 126.19
Rosette leaf number
HIF 126.29
Leaflets
Leaf number
HIF 126.29
Total leaflets on first 18 leaves
HIF 126.29
Rosette leaf number
Homozygous Oxford
Homozygous Azores

Figure 4.8 continued



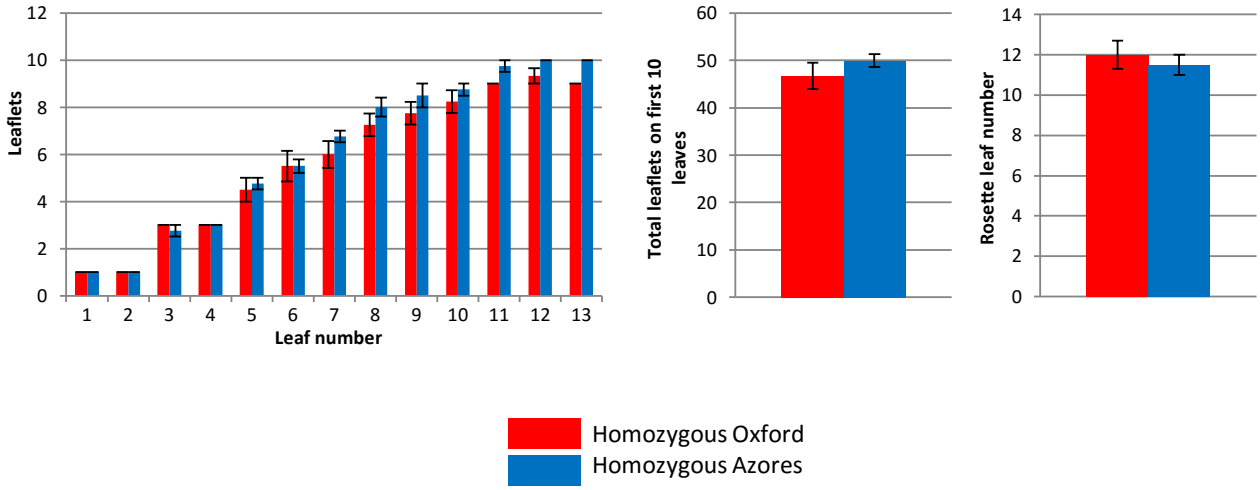
4. 3. 4. QTL fine mapping second pass

Despite a successful first round of fine mapping, at 105 kb the QTL region is still too large and contains too many genes to be able to efficiently identify possible candidates for the gene underlying the QTL, so a further reduction is required. From the same selection of recombinants, five additional lines that showed recombination events between markers 10736 and 10589 were selected for QTL validation (Figure 4.6).

There was co-segregation of phenotype and genotype in the HIF recombinant variants HIFs 126.4, 126.19, 126.29 and 127.37. Plants homozygous for the Oxford allele had significantly greater number of leaflets at multiple leaf nodes and a greater total number of leaflets on the first 18 leaves than plants homozygous for the Azores allele (Figure 4.8). Such co-segregation did not occur with HIF 126.36 (Figure 4.8).

This was the first occasion while fine mapping that rosette leaf number was recorded. HIFs 4, 36 and 37 showed no significant difference between the alternate homozygotes. However, with HIFs 19 and 29, plants homozygous for the Oxford allele had significantly more rosette leaves than plants that were homozygous for the Azores allele (Figure 4.8). Based on the areas segregating in these different lines (Figure 4.8), these results suggest there that may be an additional QTL, affecting rosette leaf production, downstream of this QTL affecting leaflet number alone. If two separate QTL do exist, then their close proximity in the genome could cause the QTL mapping model to identify a single peak on the LOD score profile.

Figure 4.9 QTL Validation with HIF 126.11 Leaflets on individual leaves , total leaflets on first 10 leaves and rosette leaf number of recombinant of HIF 126.11 homozygous for Oxford and Azores allele. Error bars indicate standard error of the mean. Significant differences are noted by an asterisk ($P < 0.05$).


Leaflets
12
10
8
6
4
2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
Leaf number
Total leaflets on first 10 leaves
60
50
40
30
20
10
0
Rosette leaf number
14
12
10
8
6
4
2
0
Homozygous Oxford
Homozygous Azores

4. 3. 5. QTL validation using HIF 126.11

In an effort to validate this potential second QTL affecting rosette leaf number, the progeny of HIF 126.11 (Figure 4.6) was examined to see if there was a phenotypic difference between the alternate homozygotes.

This experiment was performed in the greenhouses of the Max Planck Institute for Plant Breeding Research, Cologne, Germany (MPIPZ) (other experiments have since proved that the QTL can still be validated in these different conditions, see below).

There was no significant difference in rosette leaf number between the alternate genotypes (Figure 4.9). Neither was there any significant differences found when examining the leaflet data (Figure 4.9). This suggests that if there is in fact a second QTL, then it must lie between the markers 0571 and 1004 a region of c.a. 33 kb (Figure 4.6). Unfortunately no recombinant variants of the HIF were discovered whose heterozygous area included this section of genome independently of the genome between markers 0403 and 9035 (Figure 4.6), so we cannot test for the second QTL with this method.

4. 3. 6. Locating recombination break points

With this collection of recombinant variants that did and did not validate the presence of the QTL effect it is possible to identify a much smaller region known to contain the QTL. In order to do this the recombinant break points must be located at a finer resolution which was done with a new set of markers. Genotyping at new markers improved the resolution at which recombinant break points could be identified and resulted in the QTL being located to a genomic region spanning 48.6 Kb (Figure 4.6) that is known to contain 15 genes, some of which are strong candidates for the quantitative trait gene. Identification of potential candidate genes to underlie the QTL on linkage group 4 is outlined in the following chapter.

4. 4. Characterisation of QTL effect

In order to be able to fully describe the contribution a QTL has in generating natural variation it is important to accurately characterise the effect a QTL has on the trait of interest and the organ as a whole. An accurate description of QTL effect can also contribute to understanding how the QTL brings about certain changes. For this purpose HIFs are again a useful resource; the absence of background genetic variation provides comfort that all non-environmental variation is a result of the different QTL alleles present. In order to characterise the QTL-LG4 we have selected HIF recombinant lines 126.4 and 126.29.

The rationale behind selecting two different lines to characterise this QTL lies in the possibility that there are two closely linked QTL. The evidence for two closely linked QTL affecting leaflet number and rosette leaf number is discussed in the fine mapping results. Those results imply that HIF 126.4 is segregating for one QTL and HIF 126.29 is segregating for two QTL. If in fact only one QTL is present there will be no difference between the two lines when QTL effect is characterised, but if two QTL are present there will be different QTL effects observed.

Previously (Chapter 3) it was demonstrated that leaflet number varies predictably along the heteroblastic series as the shoot progresses through different phases within vegetative development. Vegetative phase change is influenced by many environmental factors which include photoperiod (reviewed in Jackson, 2009). For this reason it is reasonable to expect that the QTL may respond differently to different photoperiods. In order to see how the QTL responds to different photoperiods, each of the HIFs have been grown in short day and long day conditions.

The plants used in QTL mapping, QTL validation and QTL fine mapping were grown in controlled environment rooms in the Department of Plant Sciences at the University of Oxford, UK (DOPS). However the QTL characterisation described here took place in the greenhouses (for long day conditions) and controlled environment rooms (for the short day conditions) of the Max Planck Institute for Plant Breeding Research, Cologne, Germany (MPIPZ). This provides further

opportunity to examine the QTL effect in alternative growing conditions. Exact growing conditions are described fully in Chapter 2 of this thesis.

4. 4. 1. HIF 126.4 long day conditions

Leaflets - In accordance to what was observed in the DOPS growing conditions, the Oxford genotype conferred greater number of leaflets relative to the Azores genotype. Significant differences in number of leaves were found at leaves 5, 6 and 9 (Figure 4.10 A). Unsurprisingly, the sum of leaflets on the first nine leaves was significantly greater for the Oxford plants (Figure 4.10 B). In contrast to what was observed in DOPS, here the QTL effect is visible only in the leaves developing further along the heteroblastic series whereas in Oxford the QTL effect was observed in earlier developing leaves also in the same HIF (Figure 4.8 and Figure 4.10 A).

Rosette leaf number - Relative to the Oxford genotype, the Azores genotype produced more rosette leaves before bolting, but this is not a significant difference (Figure 4.10 C). It is worth noting that in the MPIPZ conditions the same HIF has flowered much earlier than it did in DOPS conditions (Figure 4.8).

Leaf area - The Oxford genotype produced leaves of greater size than the Azores genotype (Figure 4.10 D). This difference increases along the heteroblastic series but diminishes greatly in the last leaves before flowering.

Terminal leaflet area - It is apparent that the difference in leaf expansion observed above is not attributable to differences in terminal leaflet area. There was no significant difference in terminal leaflet area between the alternate homozygotes at any leaf node (Figure 4.10 E).

Figure 4.10 QTL-LG4 characterisation with HIF 126.4 in long day conditions. Leaflet number per leaf (A), total leaflets on first 10 leaves (B), whole leaf area (C), terminal leaflet area (D), lateral leaflet areas (F – I), rachis length (J), terminal rachis length (K), inter rachis length (L), petiole length (M) and petiolule lengths (N – Q). Significant differences between genotypes are signified by an asterisk (*) ($P < 0.05$). Error bars represent standard error of the mean.

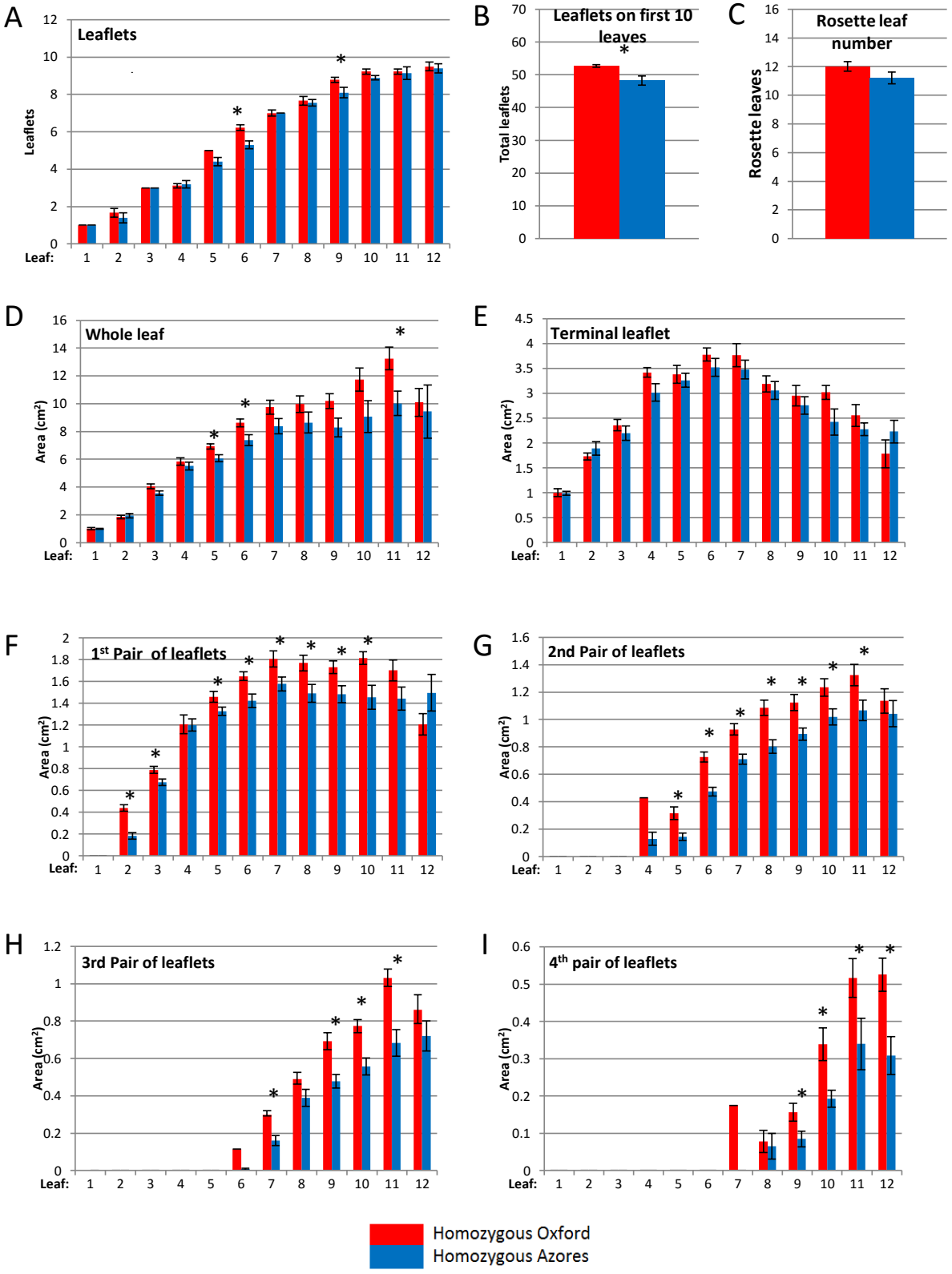
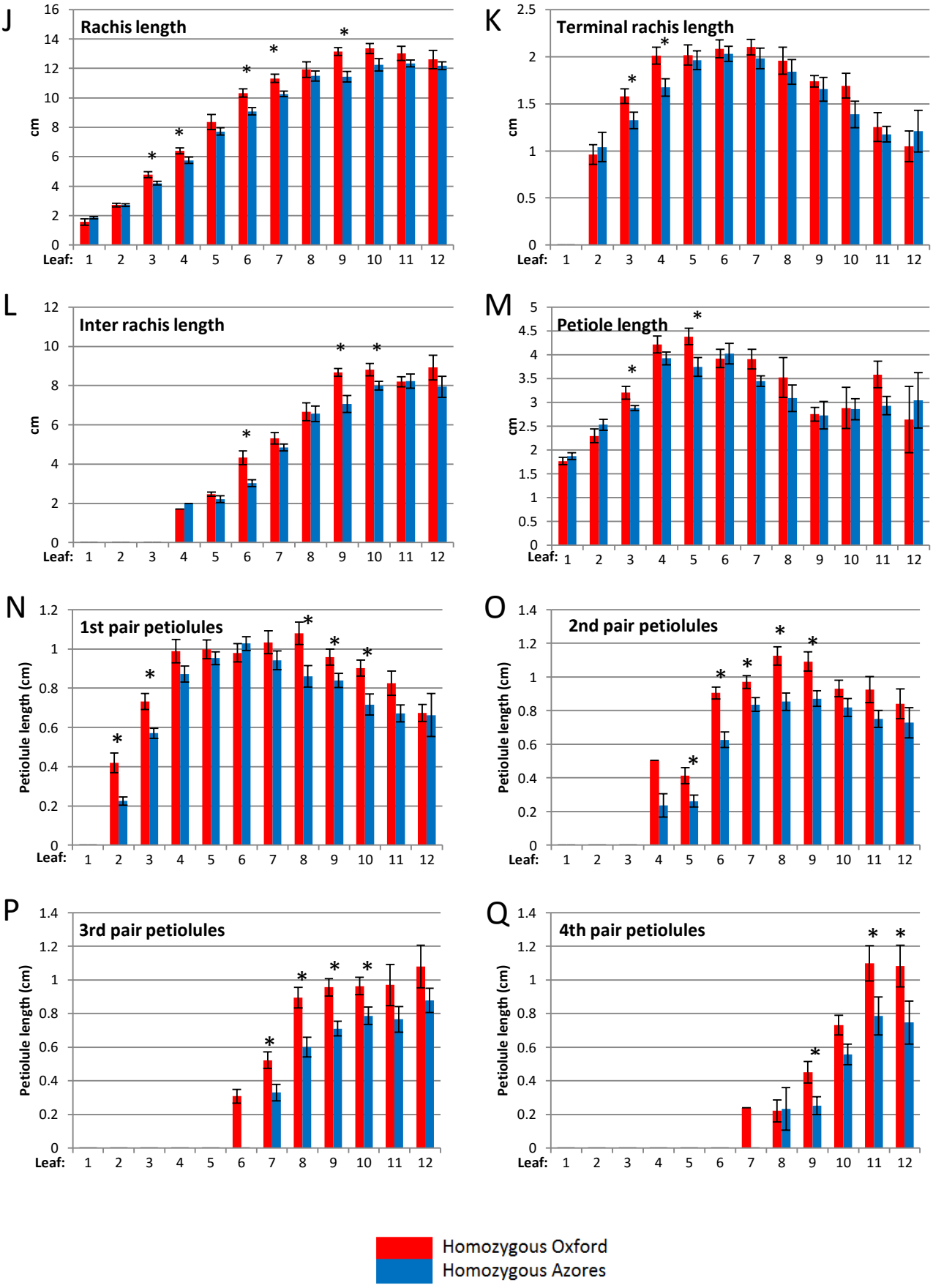

A
Leaflets
Leaflets
Leaf:
B
Leaflets on first 10 leaves
Total leaflets
C
Rosette leaf number
Rosette leaves
D
Whole leaf
Area (cm²)
Leaf:
E
Terminal leaflet
Area (cm²)
Leaf:
F
1st Pair of leaflets
Area (cm²)
Leaf:
G
2nd Pair of leaflets
Area (cm²)
Leaf:
H
3rd Pair of leaflets
Area (cm²)
Leaf:
I
4th pair of leaflets
Area (cm²)
Leaf:
Homozygous Oxford
Homozygous Azores

Figure 4.10 continued



Lateral leaflets - Lateral leaflets were larger on the leaves produced by the Oxford genotype than the Azores genotype. This difference was consistent across many leaves and true for all pairs of leaflets (Figure 4.10 F – H)). This clearly demonstrates that the difference in leaf area between the different genotypes is attributable to expansion in the lateral leaflets and not the terminal leaflet.

Rachis length - Rachis length at leaves 3 – 10 are significantly longer in the Oxford genotype (Figure 4.10 J, exclusive of leaf 5 and 8). Outside this region there is minimal variation between the alternate genotype at the start and end of the heteroblastic series.

Terminal rachis - In leaves 3 and 4 plants with the Oxford genotype have significantly longer terminal rachises (Figure 4.10 K). However this is the only point at which differences are observed and both genotypes display similar variation in this trait across development.

Inter rachis - Leaves 6, 9 and 10 had significantly longer inter rachises in the Oxford genotype relative to the Azores genotype (Figure 4.10 L).

Petiole length - Petioles of the Oxford genotype are significantly longer than those of the Azores genotype at leaves 3 and 5, but other than this there is little difference (Figure 4.10 M)

Petiolule length - The leaves produced by the Oxford genotype have longer petiolules relative to the Azores genotype across all petiolule pairs (Figure 4.10 N - Q). With the 1st pair, petiolules of oxford are longer on leaves 2 -4 and then again on leaves 8 – 11. With the 2nd pair, petiolule length is significantly different at leaves 5 to 9. With the 3rd and 4th pairs oxford has longer leaves across almost all leaf nodes.

4. 4. 2. HIF 126.4 short day conditions

Leaflets - Damage to the leaves of plants grown in short day conditions was extensive, especially amongst the leaves that grew during the earlier stages of growth. An additional 10 plants representing each genotype were sown to capture the QTL effect on leaflet production in the early

stages of growth in short day conditions making it possible to properly quantify the effect of photoperiod on the trait for which the QTL was mapped for. Only leaflet number on the first 10 leaves was recorded. This was important to do as the QTL is predicted to have a greater effect earlier in development. This data was combined with data obtained from the original plants to gain a full picture of variation in leaflet number across the whole heteroblastic series. As predicted by the QTL model, plants of the Oxford genotype produce leaves with more leaflets, which is evident when examining the total leaflets on the first 10 leaves (Figure 4.11 B). However there is only one instance in which the difference is significant at a single leaf node, leaf 5 (Figure 4.11 A).

Rosette leaf number - In this instance each genotype produced a mean number of rosette leaves identical to one another (Figure 4.11 C)

Leaf area - This analysis suffers heavily from leaf damage, however, as observed with this HIF in long day conditions, the Oxford genotype produced leaves with larger surface areas (Figure 4.11 D). This difference is significant at many leaf nodes from leaf 7 to 21. Analysis before and after these nodes is hampered due to leaf damage which meant there were multiple missing observations (Figure 4.11 D).

Terminal leaflet area - Again, analysis is hampered by damage to leaves especially in early developing leaves (hence the missing observations in Figure 4.11 E). There does however appear to be a difference; with the Oxford genotype having significantly larger terminal leaflets at leaf nodes 7 and 9 (Figure 4.11 E). This difference diminishes along the heteroblastic series.

Lateral leaflet area - The lateral leaflets of the Oxford genotype are larger than those of the Azores genotype (Figure 4.11 F - I). For all lateral leaflet pairs, the difference in area is larger in early leaves and diminishes as development proceeds and there is no difference in the leaves leading up to bolting.

Figure 4.11 QTL-LG4 characterisation with HIF 126.4 in short day conditions. Leaflet number per leaf (A), total leaflets on first 10 leaves (B), whole leaf area (C), terminal leaflet area (D), lateral leaflet areas (F – I), rachis length (J), terminal rachis length (K), inter rachis length (L), petiole length (M) and petiolule lengths (N – Q). Significant differences between genotypes are signified by an asterisk (*) ($P < 0.05$). Error bars represent standard error of the mean.

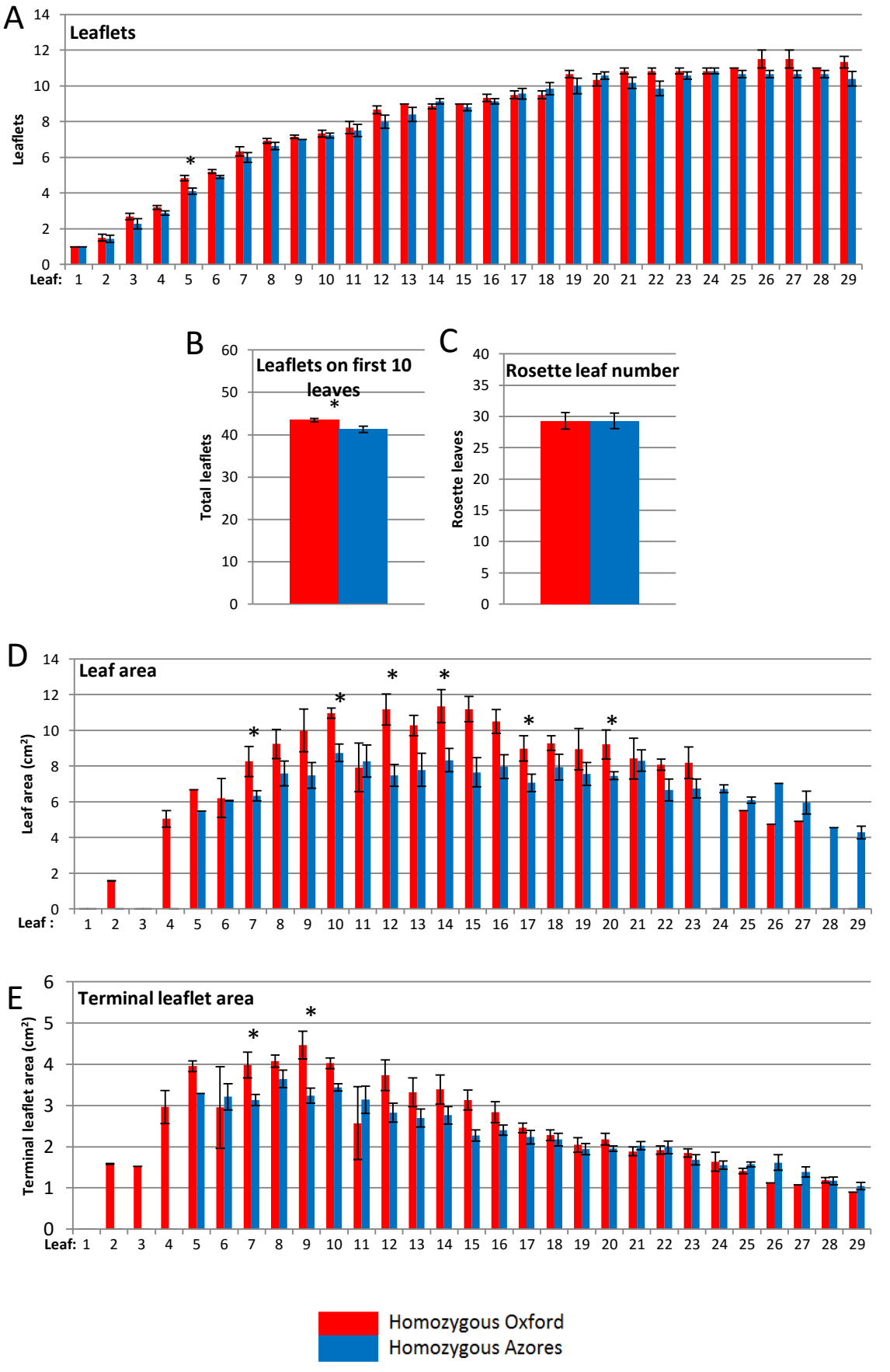

A
Leaflets
Leaflets
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
B
Leaflets on first 10 leaves
Total leaflets
C
Rosette leaf number
Rosette leaves
D
Leaf area
Leaf area (cm²)
Leaf : 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
E
Terminal leaflet area
Terminal leaflet area (cm²)
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
Homozygous Oxford
Homozygous Azores

Figure 4.11 continued

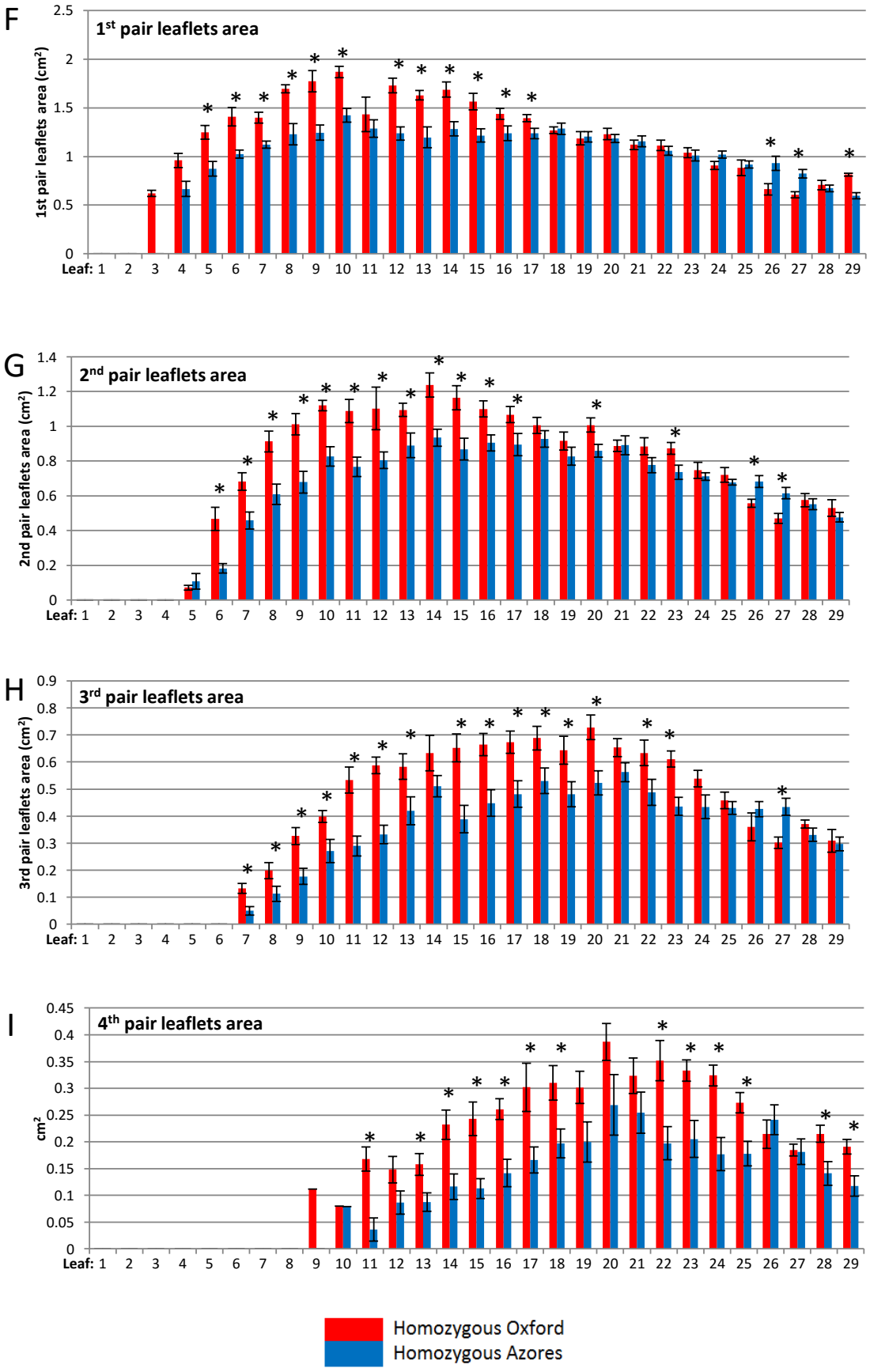


Figure 4.11 continued

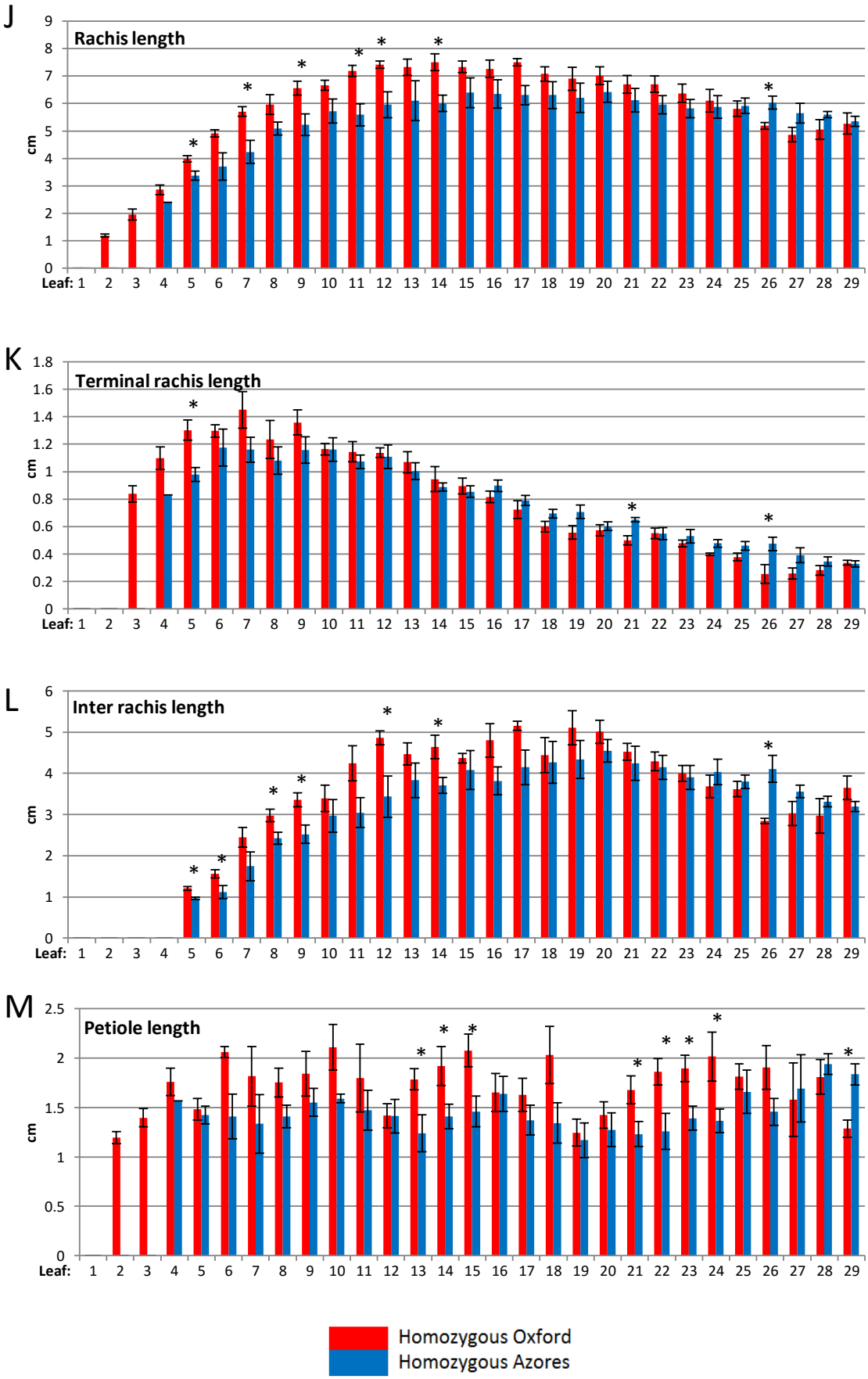
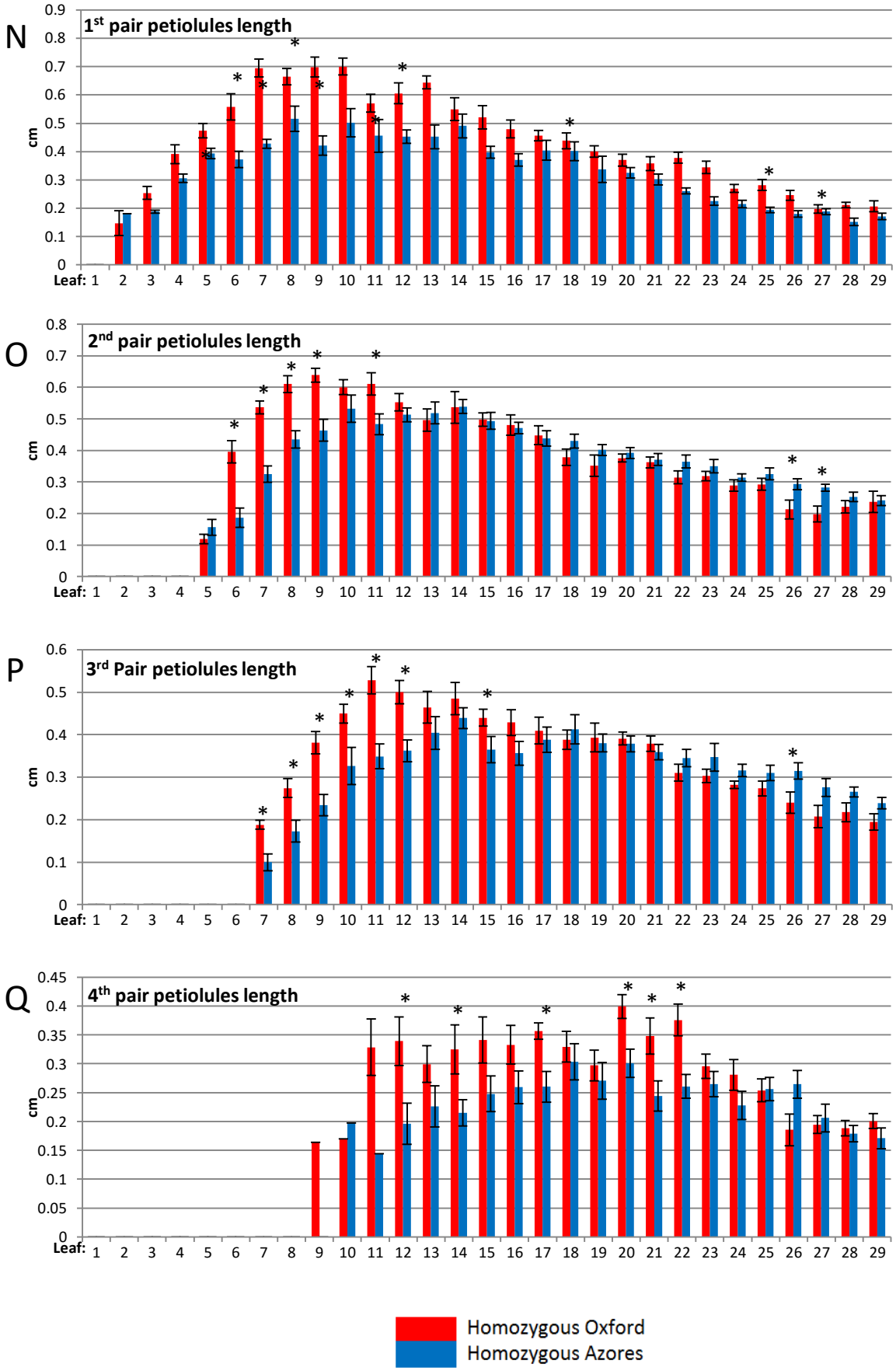


Figure 4.11 continued



Rachis length - The Oxford genotype produced plants with longer rachis than the Azores genotype (Figure 4.11 J). The disparity decreased with development and the latest point at which there is a significant difference is at leaf 17.

Terminal rachis - In the early developing leaves the Oxford genotype produces longer terminal rachises than the Azores genotype, but this is only significant at leaf 5 (Figure 4.11 K). After leaf 10 there is no significant difference between the genotypes, except at leaves 21 and 24 where the inter rachises of the Azores genotype are significantly longer (Figure 4.11 K).

Inter rachis length - From leaf 5 to leaf 17 the Oxford genotype has significantly longer inter rachis than the Azores genotype at the majority of nodes (Figure 4.11 L). After this there is no significant difference between the genotypes until leaf 26 where the Azores has a significantly longer inter rachis (Figure 4.11 L).

Petiole length - The Oxford genotype has significantly longer petioles at leaves 13 to 15 and then again at leaves 21 to 24 (Figure 4.11 M).Otherwise there are is no clear pattern to the QTL effect. At leaf 29, the Azores genotype has significantly longer petioles (Figure 4.11 M).

Petiolule length - For the 1st, 2nd and 3rd pair of petiolules; the Oxford genotype has the longer petiolules during early growth, then there is minimal difference between the genotypes for the majority of development, and in the last stages of growth Azores often has significantly longer petiolules (Figure 4.11 N – P) . At the 4th pair of leaflets, the Oxford genotype has longer petiolules on most leaves (Figure 4.11 Q)

4. 4. 3. HIF 126.29 long day conditions

Leaflets - As seen before in DOPS conditions, the Oxford genotype produces more leaflets than the Azores genotype (Figure 4.12 A and B). The differences were significant at leaves 4 to 7 and for the total number of leaflets on the first 10 leaves.

Figure 4.12 QTL-LG4 characterisation with HIF 126.29 in long day conditions. Leaflet number per leaf (A), total leaflets on first 10 leaves (B), whole leaf area (C), terminal leaflet area (D), lateral leaflet areas (F – I), rachis length (J), terminal rachis length (K), inter rachis length (L), petiole length (M) and petiolule lengths (N – Q). Significant differences between genotypes are signified by an asterisk (*) ($P < 0.05$). Error bars represent standard error of the mean.

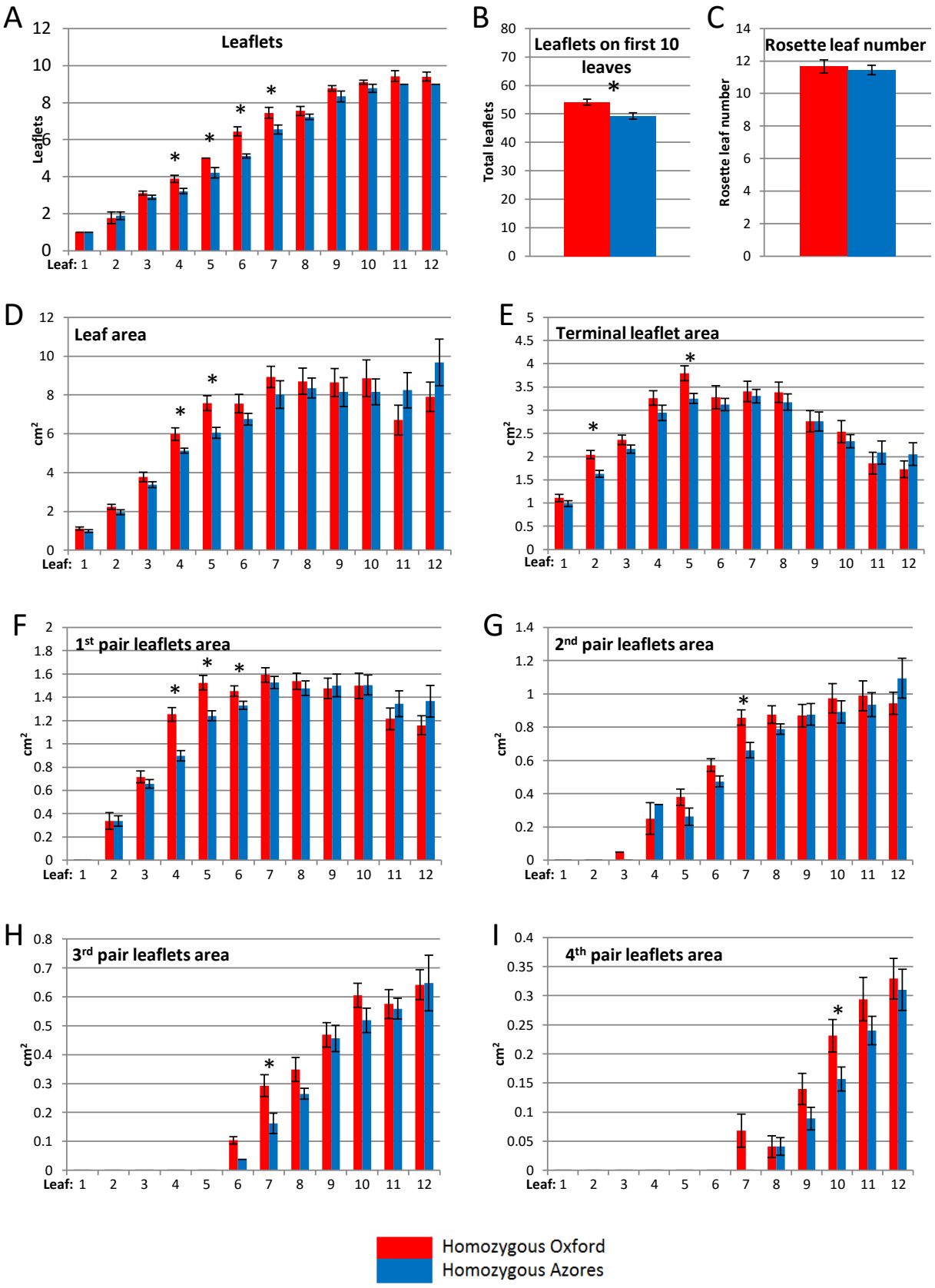
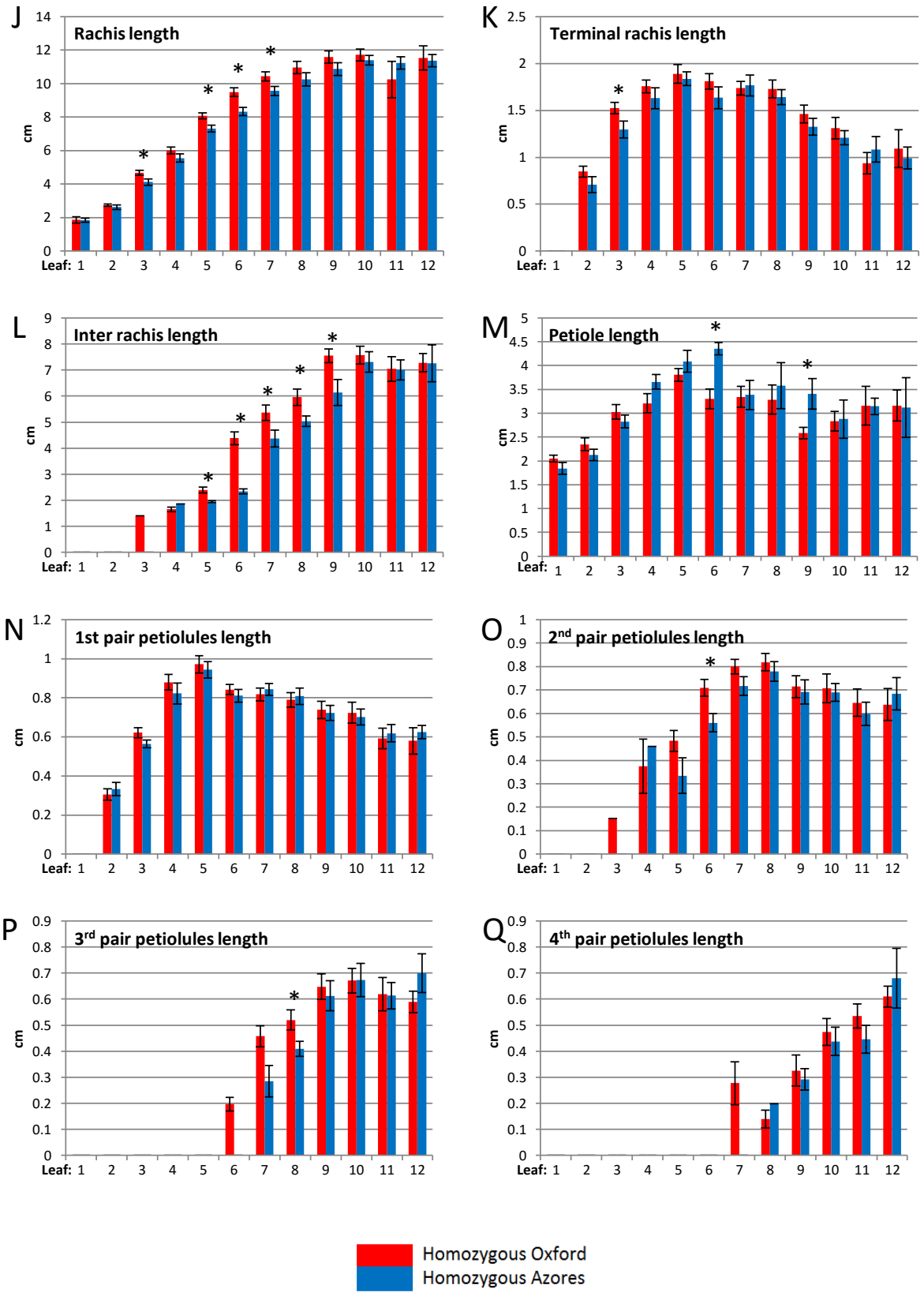

A
Leaflets
Leaflets
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
B
Leaflets on first 10 leaves
Total leaflets
C
Rosette leaf number
Rosette leaf number
D
Leaf area
cm²
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
E
Terminal leaflet area
cm²
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
F
1st pair leaflets area
cm²
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
G
2nd pair leaflets area
cm²
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
H
3rd pair leaflets area
cm²
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
I
4th pair leaflets area
cm²
Leaf: 1 2 3 4 5 6 7 8 9 10 11 12
Homozygous Oxford
Homozygous Azores

Figure 4.12 continued



Rosette leaf number - There was no significant difference between the two genotypes in regard to the number of rosette leaves produced (Figure 4.12 C). This is in contrast to the results obtained in the DOPS conditions where the Oxford genotype had the greater number of rosette leaves (Figure 4.8).

Leaf area - The leaves of the Oxford genotype are significantly larger than those of the Azores genotype at leaves 4 and 5 but there was no significant difference before or after this region in the heteroblastic series (Figure 4.12 D).

Terminal leaflet area - The Oxford genotype has significantly larger terminal leaflets than the Azores genotype on leaves 2 and 5 (Figure 4.1 E2); at all other leaf nodes both genotypes produce terminal leaflets of similar areas.

Lateral leaflets - The lateral leaflets of the Oxford genotype are generally larger than those of the Azores genotype, however the difference is only significant at a few leaf nodes (Figure 4.12 F - I).

Rachis length - The Oxford genotype has significantly longer rachis than Azores on leaves 3, 5, 6 and 7 (Figure 4.12 J). The first two leaves and leaves 8 to 12 do not vary significantly between the genotypes.

Terminal rachis - The Oxford genotype has a significantly longer terminal rachis than the Azores genotype at leaf 3, but otherwise the trait follows a similar developmental trajectory in both genotypes (Figure 4.12 K)

Inter rachis - The Oxford genotype has significantly longer inter rachis on leaves 5 to 9 and at leaves 10-12 there is minimal variation between the genotypes (Figure 4.12 L).

Petiole length - Petiole length follows a similar developmental trajectory in both genotype, except at nodes 6 and 9 where the Azores petiole is significantly longer (Figure 4.12 M).

Petiolule length - There is no significant difference in petiolule length between oxford and Azores genotype for the 1st pair of leaflets (Figure 4.12 N). For the 2nd and 3rd pairs, there is one node each where the oxford genotype has significantly longer petiolules (Figure 4.12 N). There are no significant differences between the genotypes in petiolule length on the 4th pair of leaflets (Figure 4.12 N).

4. 4. 4. HIF 126.29 short day conditions

Leaflets - As with HIF 126.4 grown in short day conditions, there was considerable damage amongst the early developing leaves precluding the analysis of early development. A further 10 plants were analysed only for the leaflets on the first 10 leaves. This data was combined with data obtained from the original plants to gain a full picture of variation in leaflet number across the whole heteroblastic series. The Oxford genotype had significantly more leaflets than the Azores genotype at leaves 3, 6, 9, 10, 11, 16, 18 and 19 (Figure 4.13 A). The sum of leaflets on the first 10 leaves was also significantly greater with the Oxford genotype (Figure 4.13 B).

Rosette leaf number - There was no significant difference between the oxford and Azores genotypes for rosette leaf number (Figure 4.13 C)

Leaf area - Damaged leaves here have hampered the analysis of whole-leaf area, differences in whole leaf area could not be identified with these limited data. As can be seen from Figure 4.13 D, the damage did prevent any observations being for some early leaves

Terminal leaflet area - Again, leaf damage has hindered this analysis as can be seen from Figure 4.13 E where there are multiple missing observations due to the damage. Terminal leaflets are particularly vulnerable to damage when plants are handled. The later developing leaves were less damaged; there does not appear to be any difference between the two genotypes (Figure 4.13 E).

Figure 4.13 QTL-LG4 characterisation with HIF 126.29 in short day conditions. Leaflet number per leaf (A), total leaflets on first 10 leaves (B), whole leaf area (C), terminal leaflet area (D), Lateral leaflet areas (F – I), rachis length (J), Terminal rachis length (K), Inter rachis length (L), petiole length (M) and petiolule lengths (N – Q). Significant differences between genotypes are signified by an asterisk (*) ($P < 0.05$). Error bars represent standard error of the mean.

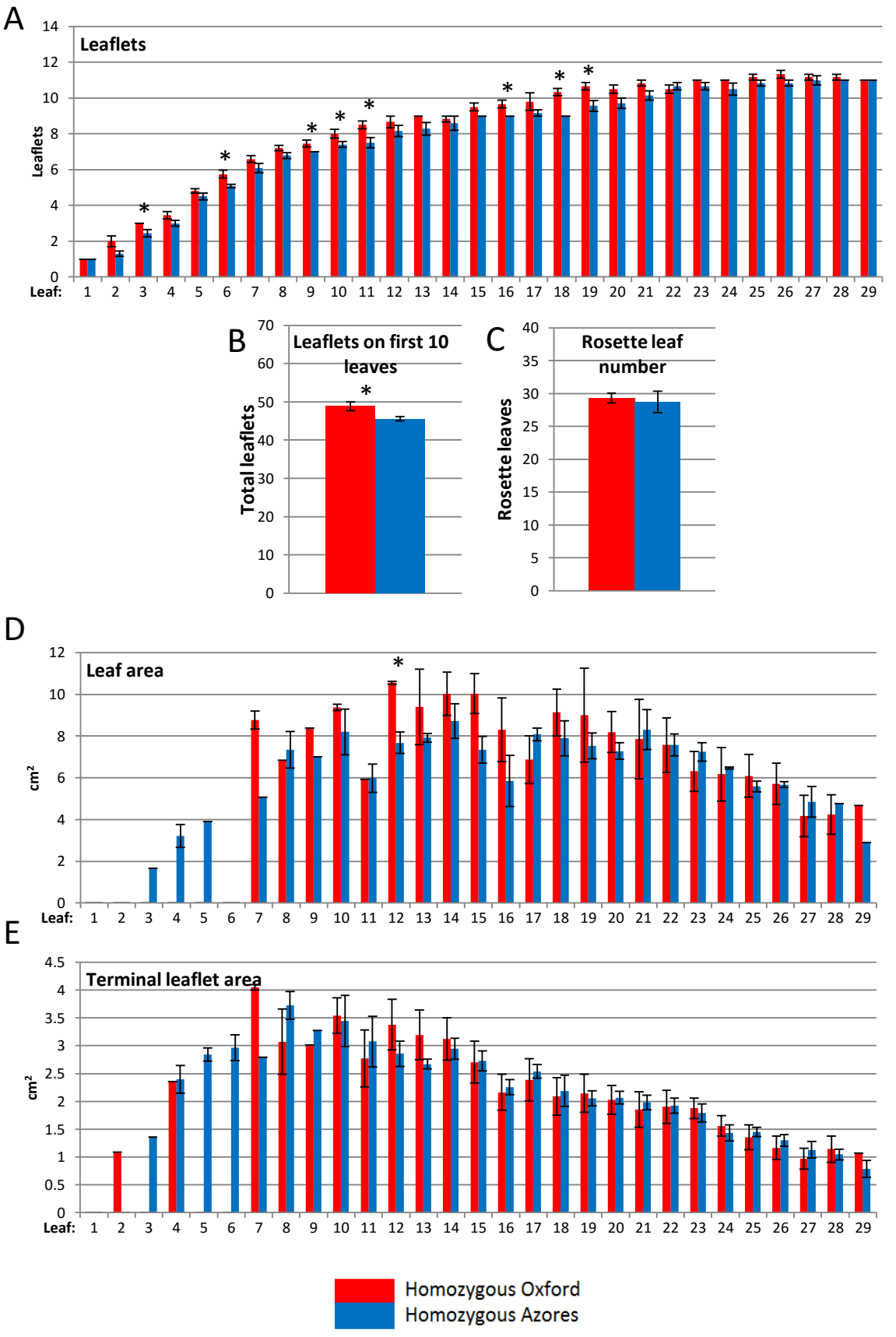

A
Leaflets
Leaflets
Leaf:
B
Leaflets on first 10 leaves
Total leaflets
C
Rosette leaf number
Rosette leaves
D
Leaf area
cm²
Leaf:
E
Terminal leaflet area
cm²
Leaf:
Homozygous Oxford
Homozygous Azores

Figure 4.13 continued

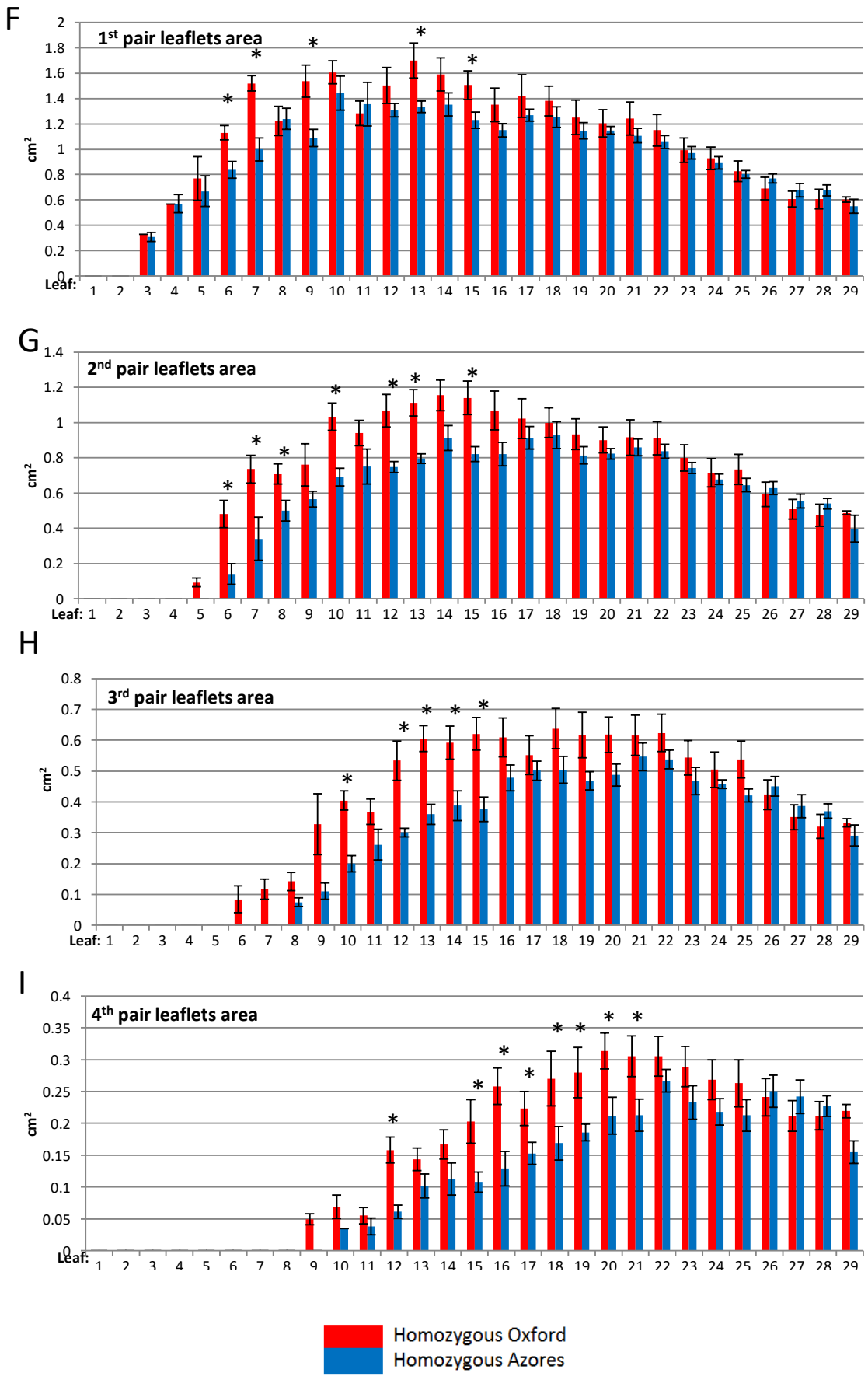


Figure 4.13 continued

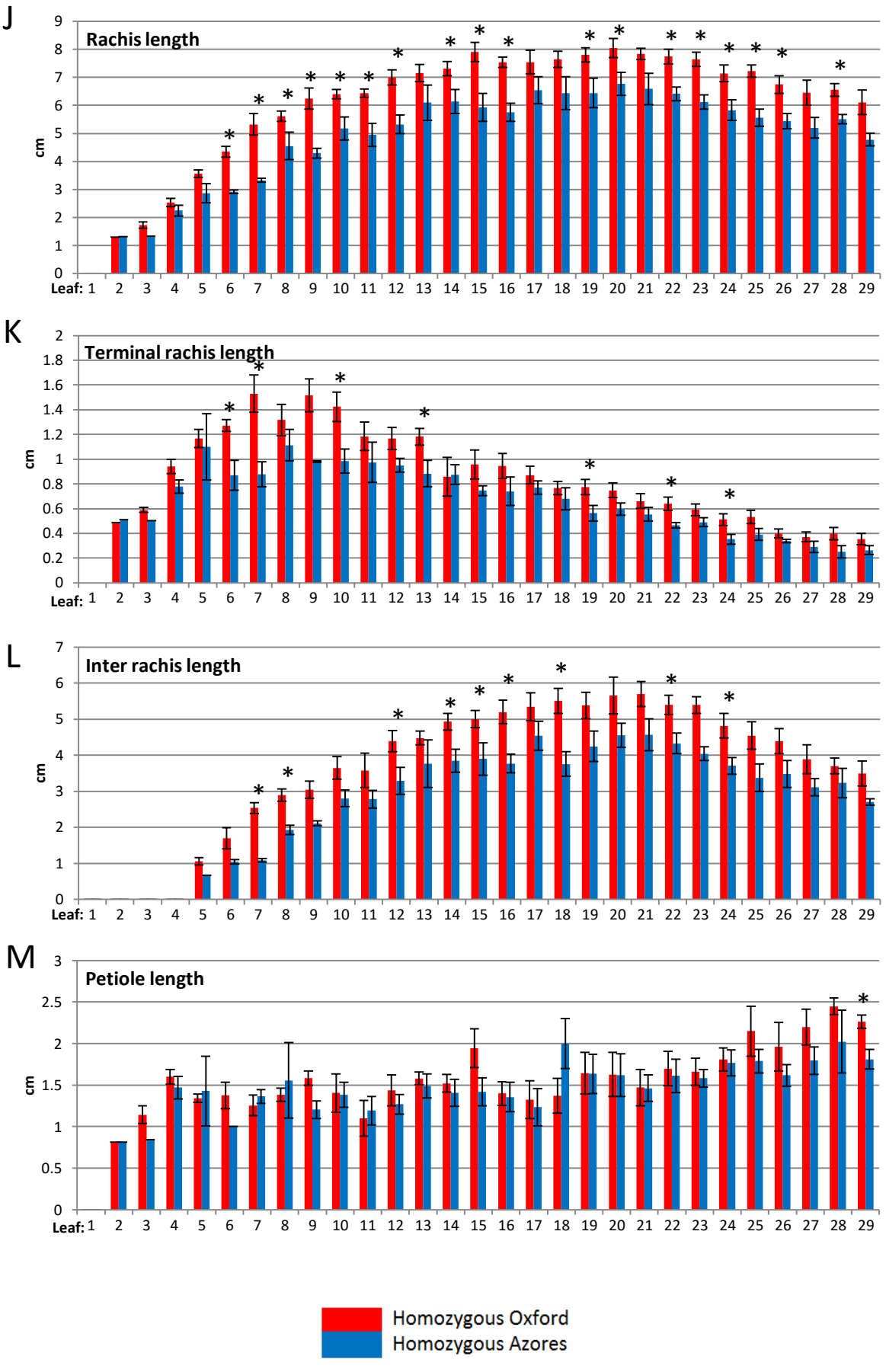
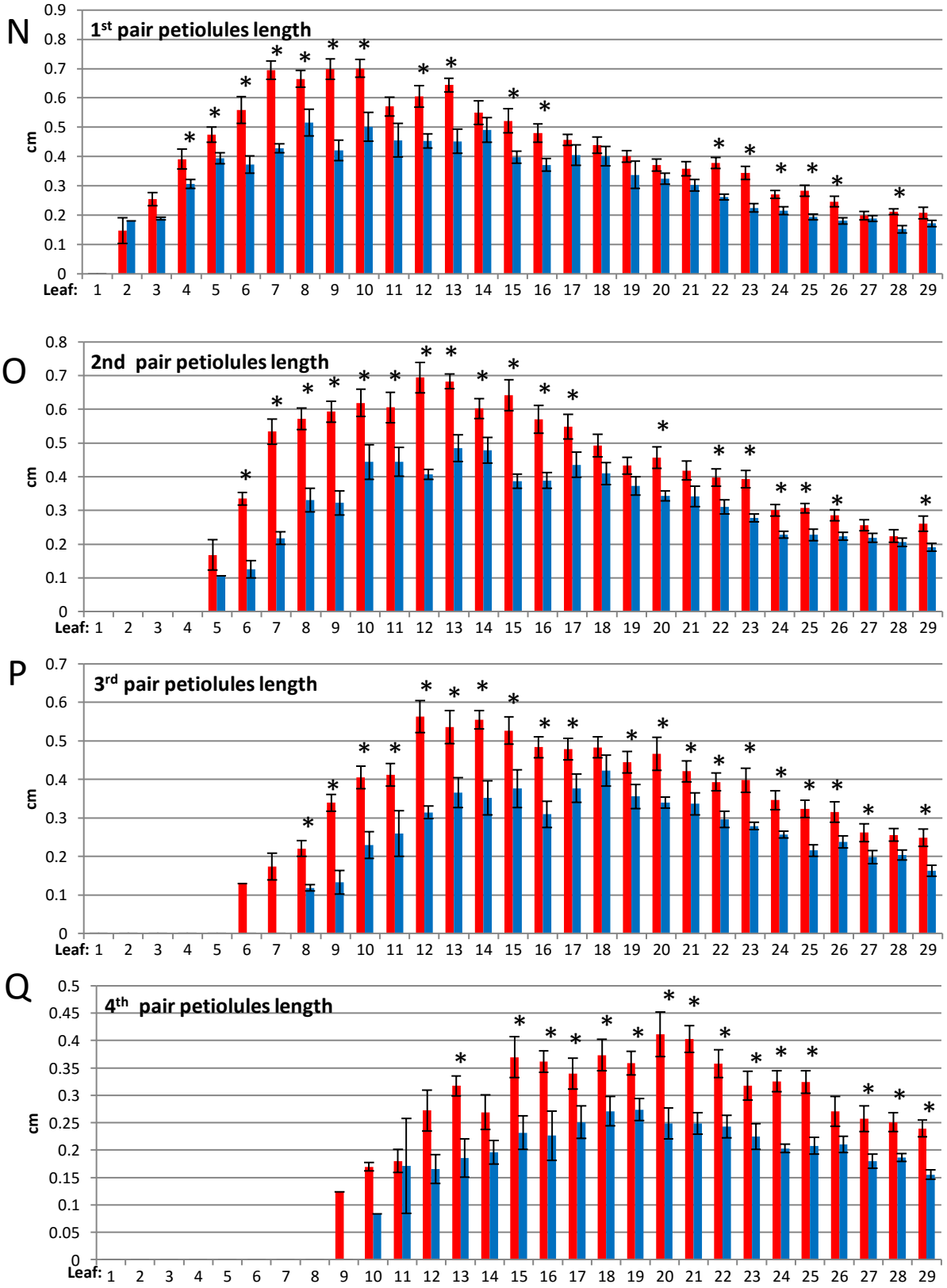


Figure 4.13 continued



Lateral leaflets - Upon their first emergence, the 1st leaflet pairs in each of the genotypes does not significantly differ in area between the two genotypes (Figure 4.13 F). From leaf 5 onwards the Oxford genotype has significantly larger leaflets at many nodes and the difference diminishes as development progresses. The 2nd, 3rd and 4th pair of leaflets all follow a similar pattern to each other; there is initially a large difference between the genotype that diminishes along the heteroblastic series (Figure 4.13 G - I)

Rachis length - The oxford genotype has significantly greater rachis length than the Azores genotype for almost the entire heteroblastic series excluding leaves 2 - 5 (Figure 4.13 J).

Terminal rachis - The terminal rachises on the leaves of the Oxford genotype are significantly longer than those of Azores genotype at leaves 6, 7, 10, 13, 19, 22 and 24 (Figure 4.13 K).

Inter rachis - The Oxford genotype has longer inter rachises than the Azores genotype across the entire heteroblastic series, with significant differences occurring at multiple nodes (Figure 4.13 L)

Petiole length - Petiole length does not appear to differ significantly between the oxford and Azores genotypes at any point during vegetative development (Figure 4.13 M).

Petiolule length - For all leaflet pairs, petiolule length is longer on the leaves of the oxford genotype relative to the Azores genotype (Figure 4.13 N - Q). This dissimilarity is maintained across the entire heteroblastic series

4. 4. 5. Summary of QTL characterisation

Sensitivity of QTL effect to variable conditions - Since we have phenotyped HIF 126.4 at MPIPZ and at DOPS it is possible to compare the two long day conditions. In both instances we were able to observe the QTL effect on leaflet number, however the QTL effect was observed across a broader range of the heteroblastic series in DOPS relative to MPIPZ conditions. In both cases, no significant effect was observed with regard to rosette leaf number, however in MPIPZ

conditions far fewer rosette leaves were produced than were produced in DOPS (Oxford: 10.6 and Azores: 10.7 in MPIPZ and Oxford: 21.4 and Azores: 19.4 in DOPS). Therefore it appears that the MPIPZ conditions have accelerated the progression through vegetative development towards flowering relative to the DOPS conditions but this does not prevent the QTL effect from being observed.

QTL effect in variable photoperiod - When comparing MPIPZ long and short day conditions there is no significant variation in QTL effect observed. With HIF 126.4 the QTL effects observed were not strikingly different in either of the conditions. QTL effects appear to vary in the same way as vegetative development progresses, except that this progression is spread over more leaf nodes in short day conditions. In contrast to 126.4, when comparing QTL effects in 126.29 observed in the two conditions there is a difference. In long day conditions there is little difference in petiolule length between the alternate genotypes, however in short day conditions Oxford has significantly longer petiolules.

Evidence for multiple linked QTL - Two recombinant versions of the same HIF were analysed here in an effort to test the hypothesis that there are two closely linked QTL in the same region; the fine mapping results above suggest that it is HIF 126.29 that has two QTL and that HIF 126.4 only has one (Figure 4.6 and 4.8). If both recombinants only had one QTL we would expect that an identical phenotype would co-segregate with genotype in both instances. The presence of a single QTL is supported by leaflet number, rosette leaf number and rachis length where there is little difference in observed QTL effects between the two HIFs. Additionally, further evidence that there are not two QTL comes from the lack of difference in QTL effects between the two HIFs grown in long day conditions. However the variation between the alternative genotypes in lateral leaflet area and in petiolule length is significant at more leaf nodes within HIF 126.4 compared to HIF 126.29 in long day conditions (Figure 4.10 and 4.12). This suggests that a different profile of genetic determinants influencing leaf shape is segregating in each recombinant and therefore supports the hypothesis that there are 2 QTL closely linked together on linkage group 4. Whilst this is compelling evidence, it should not be taken as proof of the presence of

two QTL – this can be achieved by validating the QTL in an additional HIF where a recombination event has been detected between the two possible QTL.

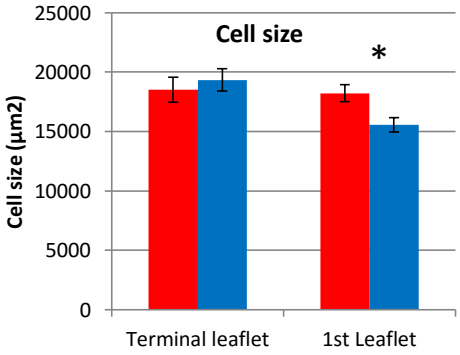
QTL effect characterisation - Up to this point the linkage group 4 QTL was mapped and validated using phenotype data on leaflet production and the number of rosette leaves produced. However, the results here clearly demonstrate that the QTL has pleiotropic effects on lateral leaflet area, rachis length and petiolule length. It would be interesting to retrospectively perform QTL mapping on these phenotypes to see if this QTL is identified. It would provide information on how efficiently QTL mapping procedures identify QTL and how accurately they predict QTL effect.

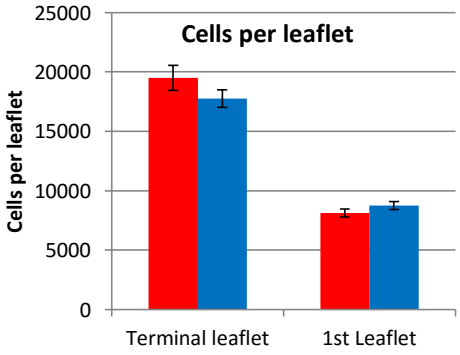
4. 5. Cell size vs. cell proliferation

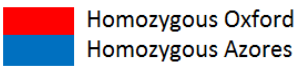
One of the conclusions drawn from the analysis of the alternative homozygotes of HIF 126.4 is that the Oxford QTL allele increases leaf area (Figure 4.10 D). Furthermore, it was demonstrated that this is due to expansion in the lateral leaflets and not in the terminal leaflet (Figure 4.10 E - I). In order to determine whether this is mediated by cell expansion or cell proliferation in the lateral leaflets, mean cell size was determined in the terminal and lateral leaflets.

Cells were analysed using agarose imprints but as this methodology is particularly labour intensive, cell imprints for the entire heteroblastic series were not analysed; leaf five was used as a representative for each allelic variant. So while this explains how the QTL mediates leaf expansion at leaf five it may not be representative of the QTL action throughout the heteroblastic series. For each leaf agarose imprints were produced for the terminal leaflet and for the lateral leaflet closest to the terminal leaflet (hereby referred to as the 1st leaflet). The 1st pair of leaflets showed the biggest difference between the alternative genotypes, which makes it more likely that a difference in cell size or cell proliferation can be detected.

Figure 4.14 Cell size vs cell proliferation in variable leaflet size . Estimated mean cell size and cell number based on agarose imprints is shown in the terminal leaflet and first leaflet for Oxford and Azores genotype of HIF 126.4. Error bars indicate standard error of the mean and asterisks (*) indicate significant differences ($P < 0.05$).


Cell size
*
Cell size (µm2)
25000
20000
15000
10000
5000
0
Terminal leaflet
1st Leaflet


Cells per leaflet
Cells per leaflet
25000
20000
15000
10000
5000
0
Terminal leaflet
1st Leaflet


Homozygous Oxford
Homozygous Azores

This methodology encountered some technical difficulties; many agarose leaf imprints did not provide sufficient resolution to be able to see the cell wall and counting of cells was impaired as a result. The final sample size used in the analysis was 11 and 10 terminal leaflets and 12 and 10 1st leaflets for Oxford and Azores homozygotes respectively.

4. 5. 1. Cell size vs. cell proliferation - Results

Within the terminal leaflets; cell size did not vary significantly between the alternate genotypes and there was no significant difference in estimated number of cells making up the leaflet (Figure 4.14). This is unsurprising as no significant difference was detected between the terminal leaflets of leaf 5 between the Oxford and Azores genotypes.

Cell size in the lateral leaflets did vary significantly; the Oxford genotype had significantly larger cells than the Azores genotype but there was no significant difference in the estimated number of cells in the leaflet (Figure 4.14).

4. 5. 1. Cell size vs. cell proliferation - Summary

These results indicate that the increased lateral leaflet area in the Oxford genotype is mediated by increased cell size. Relative to the leaflets of the Azores genotype the leaflets of the Oxford genotype were 9.8% larger. This experiment indicates that lateral leaflet cells of the Oxford genotype are 17.1% larger than those of the Azores genotype. The variance between these values suggests that cell proliferation is also affected by the QTL however in this case no significant difference was detected. This may be due to the inherent experimental variation associated with this methodology.

As mentioned above, the sample size was compromised by poor quality agarose imprints. But there are also additional factors that increase the experimental variation: cell size is estimated

rather than a precise measurement; isolating areas of complete cells requires a certain amount of subjective judgement; cell wall boundaries are not always particularly clear and more subjectivity is required to count cells. Also, uncertainties exist over whether the agarose imprints warp in changing temperatures as well as when they begin to dry out. This inherent variation might explain why allelic variation in leaflet size does not match allelic variation in cell size.

4. 6. Discussion

In this chapter I have presented the successful validation of the QTL on linkage groups 4, 2 and 8 which will provide additional avenues by which the genetic basis of intraspecific variation in leaflet production can be studied. In addition to validating the QTL-LG4 I was able to fine map it successfully to a genomic region of 48.6 kb that contains 15 genes. The relative likelihood that each of these genes is responsible for the QTL is discussed in the next chapter.

For the time being the results from HIF 126.4 should be used to make conclusions about the overall effect of QTL-LG4 as it is more likely that only one QTL is segregating in this instance. QTL effect can be seen to promote a more mature leaf phenotype. The Oxford allele confers a greater number of leaflets, larger lateral leaflets, longer leaves and wider leaves - all features associated with leaves that develop later in the heteroblastic series. However if the QTL was simply accelerating heteroblastic development we would expect that phenotypes of the Oxford and Azores alleles would converge later in development when the leaf shape fully matures. This is not what we observed; in the case of lateral leaflet area and petiolule length the Oxford allele conferred phenotypic values beyond that ever reached by the Azores genotype anywhere along the heteroblastic series. Thus, an alternative interpretation of these results is that the Oxford allele promotes increased leaflets, greater expansion of lateral leaflets and extension of petiolules relative to the Azores allele in a manner that is to a certain extent independent of heteroblastic development. But it is unlikely that QTL effect is completely uncoupled from heteroblastic development since the QTL effect clearly varies with development; this is

particularly clear when analysing the data from short-day conditions. A further interpretation of QTL effect is that the Oxford allele accelerates phenotypic development relative to the Azores allele during early development and then also confers greater phenotypic values after this initial acceleration in development.

The apparent lack of difference in QTL effects between the alternative growing conditions in different photoperiods suggests that the gene underlying QTL-LG4 has little or no involvement in pathways that are affected by day length, or if it does, the polymorphism responsible does not affect sensitivity to light. However the difference in QTL effects observed between growing conditions in DOPS and MPIPZ suggest that the QTL is influenced by external stimuli. It would be very difficult to pinpoint exactly which stimuli are responsible for the altered QTL effect in the different locations.

Chapter 5: QTL cloning and identifying the quantitative trait gene

5. 1. Chapter introduction

Previously we have described some of the variation that exists between various geographically distinct strains of *Cardamine hirsuta*, and then mapped quantitative trait loci (QTL) that contribute towards this variation. Using heterogeneous inbred families (HIFs), we were able to validate three of these. Next, a QTL on linkage group 4 (QTL-LG4) was selected for fine mapping and using recombinant variants of the HIF used for validation it was possible to fine-map the QTL to a genomic region of 48.6 kb. Controlled phenotypic experiments revealed that the QTL not only affects leaflet number and rosette leaf number but also leaflet size and rachis dimensions. The natural next step in this investigation is to determine the gene underlying the QTL, which is what the present chapter aims to do. Identifying the gene underlying the QTL will allow us to examine the mechanism by which the QTL exerts its influence over the quantitative trait. Identification of the gene can then be followed by identification of the causal sequence polymorphism and it becomes possible to see how leaf development programmes have been reconfigured in the respective parental strains to produce variable morphology.

Initially, identification of the causal gene not very simple because typically (and true in this case) even after fine mapping a QTL region will contain many genes, all of which could potentially be the quantitative trait gene. It is impractical to experimentally test all genes within the region so before any experimental work is performed the candidate gene approach is adopted. This approach involves assessing the likelihood that a gene will be involved in the development of the trait of interest and then experimental work is prioritised for the genes which are deemed to have the highest likelihood. Firstly, this involves identifying all the genes in the mapped region by examining nucleotide sequences. In some cases the genes will have been the subject of previous studies that have identified or have found evidence to suggest the function of that gene. In other cases sequence data will provide information on the expected protein sequence and the functional domains within, so that predictions can be made about the function of that gene. Once the function of a gene is known or predicted it is possible to infer which genes may play a role in trait development and those can be selected as candidates for the quantitative trait gene. Selection

of candidates can also be aided by identification of polymorphisms between the sequences of the two parental strains. By virtue of the very fact a QTL has been identified there must be a polymorphism within the genetic sequence that has affected either the expression or function of a gene. It is therefore possible to search for polymorphisms within predicted upstream promoter regions or polymorphisms within the coding sequence that have significantly altered the amino acid structure of the final protein to guide selection of candidate genes. Following selection of candidate genes there are several strategies that can be adopted to prove that a gene is causal to the QTL effects (Alonso-Blanco et al., 2005). A popular method is complementation of the QTL phenotype by transformation with the alternative allele of the candidate gene.

The fine mapped region of 48.6 kb includes 15 genes; here we present a brief description of each and the rationale behind selection of candidate genes. Of these genes, eight were not looked at further because it is clear from their known or predicted function that they are very unlikely to play a role in leaf development. The seven genes remaining consisted of transcription factors or were predicted to have nucleotide binding capabilities and could therefore have a role to play in regulation of leaf development. Genomic sequences of the Oxford and Azores strains were examined for polymorphisms in the coding sequence or near the expected promoter regions for these genes in an effort to see if expression or function of any of these genes had been disrupted in either of the strains.

The strongest candidate genes were *SQUAMOSA PROMOTER PROTEIN BINDING LIKE 9* (*ChSPL9*) and *FLOWERING BASIC HELIX-LOOP-HELIX 4* (*ChFBH4*) because of their known involvement in growth phase transition (Wu et al., 2009 and Ito et al., 2012). At the time of writing phenotype complementation has not been performed but for both alleles of both genes the complete genomic sequence has been cloned and transformed into *Arabidopsis thaliana.* The leaves of *A. thaliana* transformed with the Oxford allele of *ChSPL9* had more complex margins than those transformed with the Azores allele. Trichome density was used as an indication of vegetative phase transition and it was confirmed that the Oxford allele of *ChSPL9* accelerated phase transitions relative to the Azores allele. Transformation of *A. thaliana* with *ChFBH4*

conferred no such differences between the Oxford and Azores alleles. These results suggest that *ChSPL9* is the gene underlying the QTL.

5. 2. Candidate gene selection by gene function

Analysis of the genomic sequence of the 48.6 kb QTL region revealed there to be 15 genes, the sequence of which were input to a BLAST search to discover the ortholog in *Arabidopsis thaliana* and then The Arabidopsis Information Resource (TAIR http://www.arabidopsis.org/) was used to get the gene ontology which was used to assess the likelihood that the gene could be causal to the QTL effect. This information is listed in Table 5.1 and the colour coding denotes which genes were subject to further analysis; these included all genes with DNA binding capacity and which are therefore more likely to have regulatory functions. Also *ChLRR* was considered for further analysis as Leucine-rich repeat receptor kinases are common within developmental pathways (Torii, 2004). Not considered for further consideration at this stage were those genes that appear to have functions not related to regulation of morphology (Table 5.1)

There are two genes contained within the fine mapped QTL region which stand out as strong candidates for the quantitative trait gene because of their reported functions; these are Ch*SPL9 and ChFBH4.*

Table 5.1 Information on the *A. thaliana* orthologs found in the fine mapped QTL region. Information from The Arabidopsis Information Resource (TAIR) and from research literature is used to infer probable functions of genes and asses how likely it is that these genes are involved in eliciting the QTL effect. The genes marked in red are those that were analysed

A. thaliana locus	Gene name	Given name in text	Gene description according to TAIR	Function
AT3G57990	-	-	Unknown protein	No specific reference in literature
AT2G42150	-	*BROMO*	DNA-binding bromodomain-containing protein	No specific reference in literature
AT2G42190	-	-	Unknown protein	No specific reference in literature
AT2G42200	*SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 9 (SPL9)*	*SPL9*	Transcriptional regulator involved in the vegetative to reproductive phase transition.	Key regulator in the transition from juvenile to adult vegetative growth (Wu et al. 2009)
AT2G42210	*OEP16-3*	-	Homologous to pea OEP16 and barley pPORA (OEP16), a member of Arabidopsis OEP16 family.	Has a role in protein and amino acid import across plastid envelopes (Reinbothe et al. 2004)
AT2G42220	-	-	Rhodanese/Cell cycle control phosphatase superfamily protein	Involved in sulphur metabolism and possibly Cadmium tolerance (Lee et al, 2006)
AT2G42230	-	*C-CAP*	C-CAP/cofactor C-like domain-containing protein	No specific reference in literature
AT2G42240	-	*RRM*	RNA-binding (RRM/RBD/RNP motifs) family protein	No specific reference in literature
AT2G42250	*CYTOCHROME P450*	-	Member of CYP712A	Is an electron transfer protein
AT2G42260	*UV-B-INSENSITIVE 4, UVI4*	-	Encodes a novel plant-specific protein of unknown function	UVI4 is necessary for maintenance of the mitotic state (Hase et al. 2006)
AT2G42270	-	-	U5 small nuclear ribonucleoprotein helicase	This gene is down regulated by *STABILIZED1* a gene associated with stress tolerance (Lee et al. 2006)
AT2G42280	*FLOWERING BHLH 4 (FBH4)*	*FBH4*	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	Associated with stomatal opening (Takahashi et al. 2013) and controls expression of the flowering regulator *CONSTANS* and overexpression results in premature flowering (Ito *et al. 2012)*
AT2G42290	-	*LRR*	Leucine-rich repeat protein kinase family protein	No specific reference in literature
AT2G42300	-	*bHLH*	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	No specific reference in literature
AT2G42320	-	-	Nucleolar protein gar2-related	No specific reference in literature

Based on function Ch*SPL9* is a very strong candidate for the QTL. In *A. thaliana SPL9* acts in a pathway with microRNA156 (miRNA156) and miR172 to regulate the transition from juvenile vegetative growth to adult vegetative growth (Schwarz et al, 2008; Wu et al. 2009). In the shoot development of *Arabidopsis* and maize, miR156 is highly expressed very early on and decreases over time, the expression of miR172 follows the opposite pattern (Aukerman and Sakai, 2003; Chuck et al. 2007; Lauter et al. 2005; Wu and Poethig 2006). Over expression of miR156 prolongs the juvenile phase of growth and delays flowering (Chuck et al. 2007, Wu and Poethig, 2006). In contrast, miR172 overexpression results in accelerated flowering. Thus the actions of these miRNAs are integral in the transition from juvenile to adult vegetative growth and they are functionally linked by the action of SPL9. SPL9 transcript is targeted by miR156 and, redundantly with SPL10, SPL9 directly promotes the expression of *mir172b* (Wu et al, 2009). Thus decreasing expression of *miR156* in a growing shoot drives increased expression of *SPL9*, which in turn promotes expression of *miR172b* and the onset of an adult shoot identity. We have demonstrated in previous chapters that the QTL effect appears to be accelerating transit through the heteroblastic series, therefore *ChSPL9* must be considered a strong candidate for the QTL because it is important to the timing of vegetative phase change (Wu et al, 2009).

ChFBH4 can be considered a candidate for similar reasons; it may also have an effect on changes in growth phases. In *Arabidopsis* over expression of *FBH* genes induces early flowering by elevating levels of *CONSTANS* (Ito *et al*., 2012), an important transcriptional regulator in controlling the timing of flowering (Amasino, 2010). As *ChFBH4* is involved with control of developmental timing, it is possible that it may have effects on heteroblastic development.

5. 3. Candidate gene selection by genome sequence analysis

After looking at likely functions of gene, another factor by which candidate gene selection can be guided by is genetic sequence polymorphisms within genes. In order for the

QTL to be identified during QTL mapping there must be a polymorphism between the strains that is eliciting the variable genetic activity to produce the phenotype. By looking for polymorphisms that might affect a genes activity it may be possible to select candidate genes based on any polymorphisms within promoter regions of the genes and any polymorphisms that affect the final protein structure.

For this analysis the entire genomic sequence of both Azores and Oxford were compared and the position of every polymorphism was recorded and then checked to see if they sat within upstream regions of genes or within the coding sequence of the genes (Table 5.2). Polymorphisms within the coding sequence that do not affect protein sequence are unlikely to alter gene function so only polymorphisms that elicit amino acid substitutions, open reading frame shifts or premature stop codons were considered in the analysis. Amino acid substitutions were classified into three categories; conservative, semi conservative and non-conservative depending on the physio-chemical properties of the amino acids involved (Livingstone and Barton, 1993). A conservative substitution is where one amino acid is replaced by another with similar physio-chemical properties, a non-conservative substitution is where one amino acid has been substituted for another with very different properties and a semi conservative substitution lies between these extremes. Non conservative amino acid substitutions are more likely to alter protein structure and function than conservative and semi conservative substitutions and therefore more likely to bring about phenotype modification and thus present a good candidate for the QTL gene.

Non-coding regions can have regulatory roles, so it is possible that polymorphisms in regions not identified as exons could produce the QTL effect. In this respect it is interesting to note the high density of polymorphisms that lie upstream of *ChFBH4* (Table 5.2).

5. 3. 1. Search for polymorphisms within the miR156 target sequence of ChSPL9

From the researching of gene functions, *ChSPL9* emerged as a strong candidate as it is involved in shoot maturation in *Arabidopsis* (Schwarz et al, 2008; Wu et al, 2009). The juvenile

phase of growth is maintained by the degradation of *SPL9* transcript by miR156. If *ChSPL9* proves to be the quantitative trait gene it is possible that the mechanism by which the QTL effect is elicited could be a result of changes in the ChmiR156 – *ChSPL9* interaction. The target site of ChmiR156 was identified in *ChSPL9* by using the corresponding *A. thaliana pre-MIR156* hairpin sequence in a BLAST search. The sequences identified in the Oxford and Azores sequences were identical. This suggests that if *ChSPL9* is shown to be the QTL gene, then the QTL effect is not as a result of the *Ch miR156 – ChSPL9* relationship.

Table 5.2. Distribution of sequence polymorphisms in and around candidate genes and their effect on protein sequence. Upstream polymorphisms include all polymorphisms upstream of the gene up to the final stop codon of the gene immediately upstream. Categories of amino acid substitutions are based on the physio-chemical differences between the amino acids involved. CDS: coding sequence, ORF: open reading frame. (* Last 27 amino acids affected)

Gene	Upstream polymorphisms	Polymorphisms in CDS	Amino acid substitutions			ORF shift	Premature stop codon
			Conservative	Semi-conservative	non-conservative		
ChSPL9	5	2	2	0	0	N	N
ChBROMO	6	6	1	0	1	Y*	N
ChC-CAP	3	1	0	0	0	N	N
ChRRM	3	4	2	0	1	N	N
ChFBH4	83	2	1	1	0	N	N
ChLRR	6	1	0	1	0	N	N
ChbHLH	6	2	1	1	0	N	N

5. 4. Candidate gene selection by protein sequence analysis

Polymorphisms between the Oxford and the Azores were ubiquitous and polymorphisms were discovered upstream and within the coding sequence of all genes (Table 5.2). All of the polymorphisms in Table 5.2 were single nucleotide polymorphisms and no large scale multi nucleotide deletions or insertions were discovered to be associated with any of the candidate genes. No attempt at locating exact promoter regions was made and thus predictions about the impact of upstream polymorphisms are limited. However the high density of polymorphisms upstream of *ChFBH4* may be significant if the promoter region is affected. In terms of large scale changes to protein structure; no premature stop codons were introduced by polymorphisms and in only one case was a gene's open reading frame altered. This occurred in *ChBROMO* but only the final 27 amino acids of the protein were affected, so the basic structure of the protein is unlikely to be affected significantly. From these results it appears that some of the candidates have been affected minimally by polymorphisms within the coding sequence; the protein sequence of ChC-CAP has not been altered at all by polymorphisms and within ChSPL9 there has only been substitutions with conservative physio-chemical differences. ChBROMO, ChRRM, ChFBH4, ChLRR and ChbHLH all have at least a single semi- or non-conservative substitution so it may be more likely that their structures and activity have allelic variation.

However analysis of these polymorphisms within the coding sequence in this way is subject to a degree of subjectivity in regards to whether polymorphisms have a tangible effect on protein activity. More informed predictions can be made if it is known that the amino acid substitutions occur within the functional domains of proteins. An amino acid change within a domain is more likely to alter function so substitutions within domains can inform candidate gene selection. Presence and locations of key functional domains were discovered using the protein BLAST application at www.blast.ncbi.nlm.nih.gov/.

Important domains can also be identified by analysing protein sequence conservation between closely related species. Regions that are unimportant for effective protein activity are

less likely to be subject to selection than those regions important for protein activity. Therefore, analysis of protein conservation between related plant species can provide information regarding the relative importance of certain protein sequences. Following this logic, it may be possible to assess the impact on protein activity by finding polymorphisms within these regions that inform candidate gene selection. Protein sequences for the genes were obtained for 12 species within the Malvidae clade from www.phytozome.net/.

The polymorphisms within ChBROMO lie outside of the known protein domains (Figure 5.1 A) showing that the known domains are unaffected. The non-conservative substitution (glutamic acid in Oxford, glycine in Azores) was at a position which is only conserved in three other species, and the conservative substitution (asparagine in Oxford, lysine in Azores) appears to occur in an area that shows little conservation (Figure 5.1 B). Additionally a nucleotide deletion in the coding sequence of ChBROMO created an open reading frame shift at the carboxy terminus of ChBROMO, such a polymorphism might normally have a big impact on protein structure. However, in this instance the frame shift has not altered a region shared by any of the other species in the analysis (Figure 5.1 B) which suggests that protein activity is unaffected.

ChSPL9 only has one known functional domain, the SBP box, and this is unaffected by the amino acid sequence differences between the Oxford and Azores proteins (Figure 5.2 A). One of the polymorphisms (glutamic acid in Oxford, glutamine in Azores) is at a position in which there is a high degree of conservation; all but two of the Malvidae species analysed here have a glutamic acid in that position (Figure 5.2 B). The two species that do not have a glutamic acid at that position were the two *Citrus species*; *C. sisnensis and C. clementi*, these have a glycine (Figure 5.2 B). Also different is Azores which has a glutamine at that particular position (Figure 5.2 A). The other conservative amino acid substitution (aspartic acid in Oxford, aspargine in Azores) in the protein sequence of ChSPL9 occurred at a non-conserved region in which the only species that has the same amino acid was *Brassica rapa* which has an asparagine in common with Azores (Figure 5.2).

Figure 5.1 Amino acid sequence polymorphisms within ChBROMO. Comparison of Oxford and Azores amino acid sequence for ChBROMO (**A**); polymorphisms are highlighted by coloured boxes and the location of the protein domains are shown by coloured lines. The blue line represents the SANT chromatin remodelling and transcriptional regulator domain and the red line shows the protein interacting Bromodomain. Alignment for amino acid sequences of species within the Malvidae (**B**) around the Oxford/Azores polymorphisms which are highlighted by coloured boxes. Numbers to right of sequences indicate amino acid number for the end of line shown. The black line in both A and B shows where the open reading frame shift has occurred in the Azores protein sequence Amino acid sequences obtained from http://www.phytozome.net/ and alignments performed using online alignment tool www.genome.lbl.gov/vista.

A.


SANT
Bromodomain
Frame shift

B.

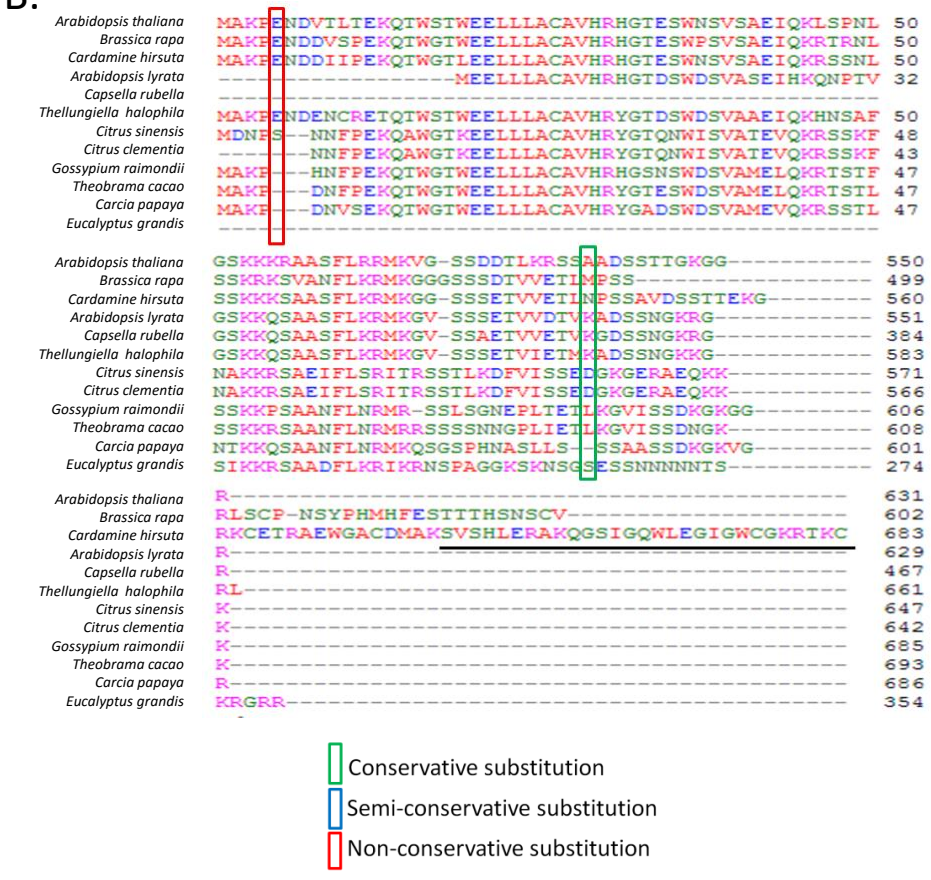
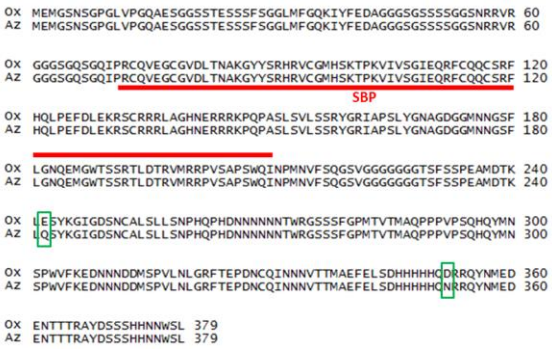
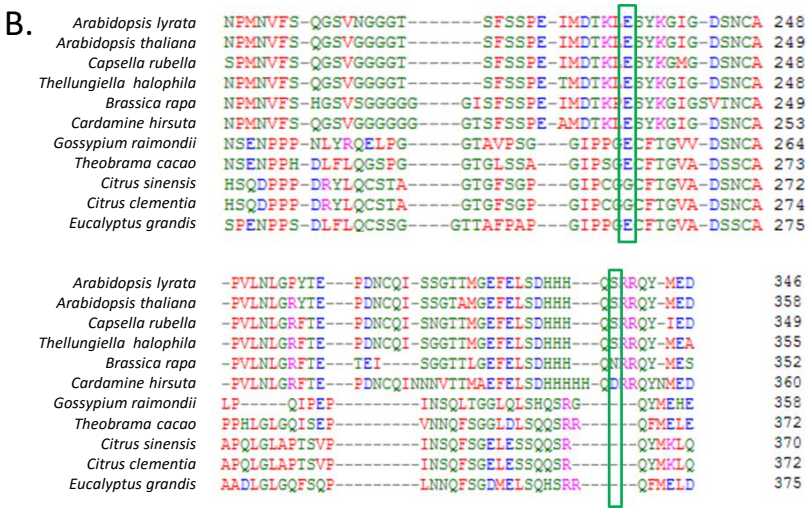

Arabidopsis thaliana
Brassica rapa
Cardamine hirsuta
Arabidopsis lyrata
Capsella rubella
Thellungiella halophila
Citrus sinensis
Citrus clementia
Gossypium raimondii
Theobrama cacao
Carcia papaya
Eucalyptus grandis
Conservative substitution
Semi-conservative substitution
Non-conservative substitution

Figure 5.2 Amino acid sequence polymorphisms within ChSPL9. Comparison of Oxford and Azores amino acid sequence for ChSPL9 (**A**); polymorphisms are highlighted by coloured boxes and the location of *SQUAMOSA* promoter binding domain (SBP) shown by red line. Alignment for amino acid sequences of species within the Malvidae (**B**) around the Oxford/Azores polymorphisms are highlighted by coloured boxes. Numbers to right of sequences indicate amino acid number for the end of line shown. Amino acid sequences obtained from http://www.phytozome.net/ and alignments performed using online alignment tool www.genome.lbl.gov/vista.

A.



B.



Conservative substitution

Semi-conservative substitution

Non-conservative substitution

Figure 5.3 Amino acid sequence polymorphisms within ChRRM. Comparison of Oxford and Azores amino acid sequence for ChRRM (**A**); polymorphisms are highlighted by coloured boxes and the location of the two RNA recognition motifs (RRM) are shown by red lines. Alignment for amino acid sequences of species within the Malvidae (**B**) around the Oxford/Azores polymorphisms are highlighted by coloured boxes. Numbers to right of sequences indicate amino acid number for the end of line shown. Amino acid sequences obtained from http://www.phytozome.net/ and alignments performed using online alignment tool www.genome.lbl.gov/vista.

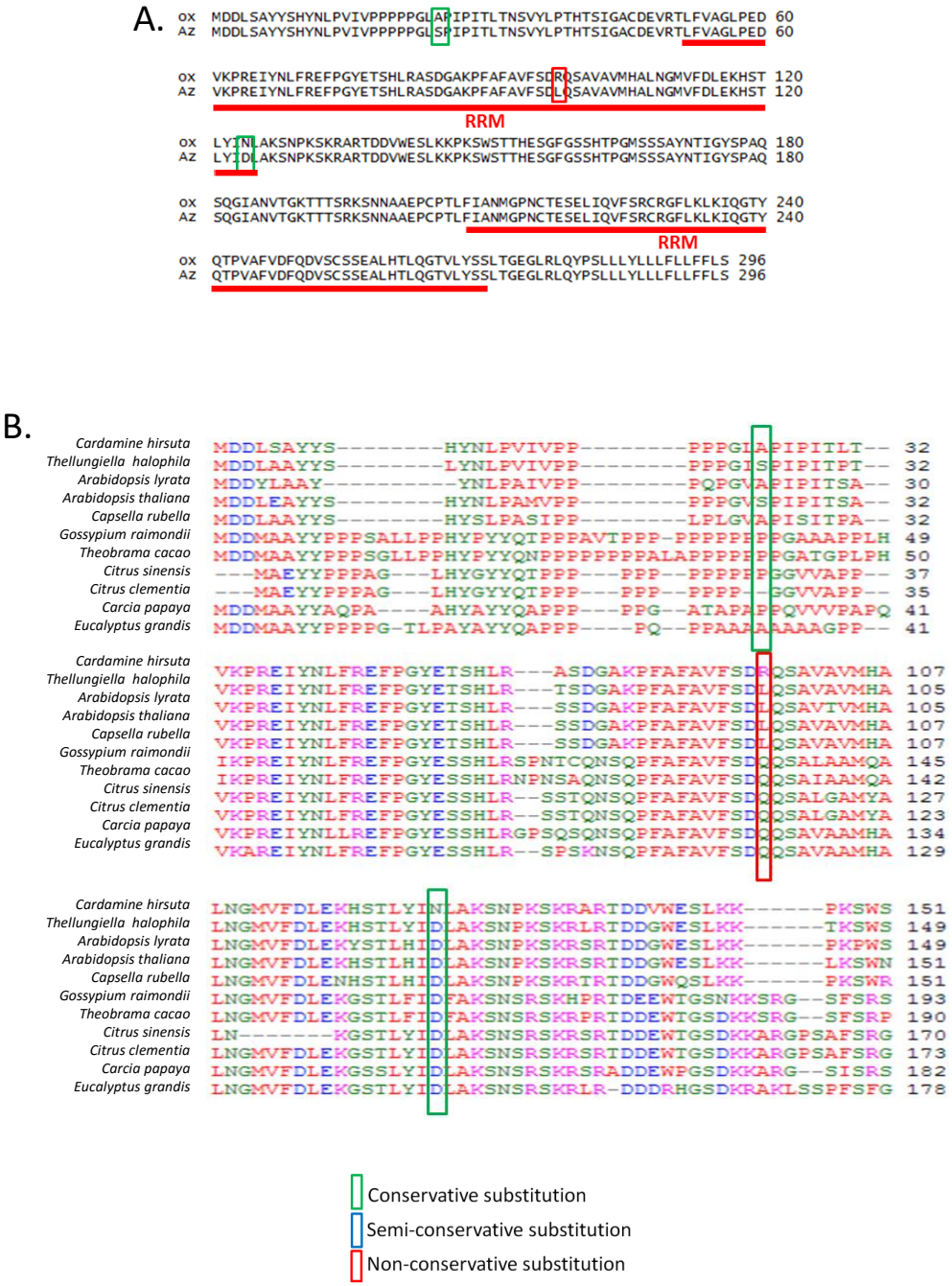
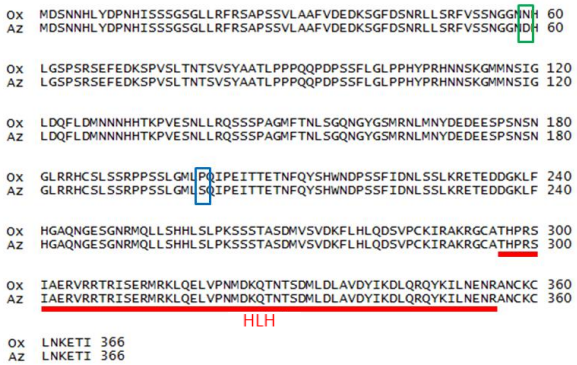

A.
ox MDDLSAYYSHYNLPVIVPPPPPGLAPIPITLTNSVYLPTHTSIGACDEVRTLFVAGLPED 60
Az MDDLSAYYSHYNLPVIVPPPPPGLSPIPITLTNSVYLPTHTSIGACDEVRTLFVAGLPED 60
ox VKPREIYNLFREFPGYETSHLRASDGAKPFAFAVFSDRQSAVAVMHALNGMVFDLEKHST 120
Az VKPREIYNLFREFPGYETSHLRASDGAKPFAFAVFSDLQSAVAVMHALNGMVFDLEKHST 120
RRM
ox LYINLAKSNPKSKRARTDDVWESLKKPKSWSTTHESGFGSSHTPGMSSSAYNTIGYSPAQ 180
Az LYIDLAKSNPKSKRARTDDVWESLKKPKSWSTTHESGFGSSHTPGMSSSAYNTIGYSPAQ 180
ox SQGIANVTGKTTTSRKSNNAAEPCPTLFIANMGPNCTESELIQVFSRCRGFLKLKIQGTY 240
Az SQGIANVTGKTTTSRKSNNAAEPCPTLFIANMGPNCTESELIQVFSRCRGFLKLKIQGTY 240
RRM
ox QTPVAFVDFQDVSCSSEALHTLQGTVLYSSLTGEGLRLQYPSLLLYLLLFLLFFLS 296
Az QTPVAFVDFQDVSCSSEALHTLQGTVLYSSLTGEGLRLQYPSLLLYLLLFLLFFLS 296
B.
Cardamine hirsuta MDDLSAYYS---------HYNLPVIVPP---------PPPGLAPIPITLT-- 32
Thellungiella halophila MDDLAAYYS---------LYNLPVIVPP---------PPPGISPIPITPT-- 32
Arabidopsis lyrata MDDYLAAY----------YNLPAIVPP---------PQPGVAPIPITSA-- 30
Arabidopsis thaliana MDDLEAYYS---------HYNLPAMVPP---------PPPGVSPIPITSA-- 32
Capsella rubella MDDLAAYYS---------HYSLPASIPP---------LPLGVAPISITPA-- 32
Gossypium raimondii MDDMAAYYPPPSALLPPHYPYYQTPPPAVTPPP-PPPPPPPPPGAAAPPLH 49
Theobrama cacao MDDMAAYYPPPSGLLPPHYPYYQNPPPPPPPPALAPPPPPPPPGATGPLPH 50
Citrus sinensis ---MAEYYPPPAG---LHYGYYQTPPP---PPP--PPPPPPGGVVAPP-- 37
Citrus clementia ---MAEYYPPPAG---LHYGYYQTPPP---PPP--PPPP---GGVVAPP-- 35
Carcia papaya MDDMAAYYAQPA----AHYAYYQAPPP---PPG--ATAPAPPPQVVVPAPQ 41
Eucalyptus grandis MDDMAAYYPPPPG-TLPAYAYYQAPPP----PQ--PPAAAAAAAAGPP-- 41
Cardamine hirsuta VKPREIYNLFREFPGYETSHLR---ASDGAKPFAFAVFSDRQSAVAVMHA 107
Thellungiella halophila VKPREIYNLFREFPGYETSHLR---TSDGAKPFAFAVFSDLQSAVAVMHA 105
Arabidopsis lyrata VKPREIYNLFREFPGYETSHLR---SSDGAKPFAFAVFSDLQSAVTVMHA 105
Arabidopsis thaliana VKPREIYNLFREFPGYETSHLR---SSDGAKPFAFAVFSDLQSAVAVMHA 107
Capsella rubella VKPREIYNLFREFPGYETSHLR---SSDGAKPFAFAVFSDLQSAVAVMHA 107
Gossypium raimondii IKPREIYNLFREFPGYESSHLRSPNTCQNSQPFAFAVFSDQQSALAAMQA 145
Theobrama cacao IKPREIYNLFREFPGYESSHLRNPNSAQNSQPFAFAVFSDQQSAIAAMQA 142
Citrus sinensis VKPREIYNLFREFPGYESSHLR--SSTQNSQPFAFAVFSDQQSALGAMYA 127
Citrus clementia VKPREIYNLFREFPGYESSHLR--SSTQNSQPFAFAVFSDQQSALGAMYA 123
Carcia papaya VKPREIYNLLREFPGYESSHLRGPSQSQNSQPFAFAVFSDQQSAVAAMHA 134
Eucalyptus grandis VKAREIYNLFREFPGYESSHLR--SPSKNSQPFAFAVFSDQQSAVAAMHA 129
Cardamine hirsuta LNGMVFDLEKHSTLYINLAKSNPKSKRARTDDVWESLKK------PKSWS 151
Thellungiella halophila LNGMVFDLEKHSTLYIDLAKSNPKSKRLRTDDGWESLKK------TKSWS 149
Arabidopsis lyrata LNGMVFDLEKYSTLHIDLAKSNPKSKRSRTDDGWESLKK------PKPWS 149
Arabidopsis thaliana LNGMVFDLEKHSTLHIDLAKSNPKSKRSRTDDGWESLKK------LKSWN 151
Capsella rubella LNGMVFDLENHSTLHIDLAKSNPKSKRTRTDDGWQSLKK------PKSWR 151
Gossypium raimondii LNGMVFDLEKGSTLFIDFAKSNSRSKHPRTDEEWTGSNKKSRG--SFSRS 193
Theobrama cacao LNGMVFDLEKGSTLFIDFAKSNSRSKRPRTDDEWTGSDKKSRG--SFSRP 190
Citrus sinensis LN--------KGSTLYIDLAKSNSRSKRSRTDDEWTGSDKKARGPSAFSRG 170
Citrus clementia LNGMVFDLEKGSTLYIDLAKSNSRSKRSRTDDEWTGSDKKARGPSAFSRG 173
Carcia papaya LNGMVFDLEKGSSLYIDLAKSNSRSKRSRADDEWPGSDKKARG--SISRS 182
Eucalyptus grandis LNGMVFDLEKGSTLYIDLAKSNSRSKRLR-DDDRHGSDKRAKLSSPFSFG 178
Conservative substitution
Semi-conservative substitution
Non-conservative substitution

Figure 5.4 Amino acid sequence polymorphisms within ChFBH4. Comparison of Oxford and Azores amino acid sequence for ChFBH4 (**A**); polymorphisms are highlighted by coloured boxes and the location of helix loop helix DNA binding domain (HLH) shown by red line. Alignment for amino acid sequences of species within the Malvidae (**B**) around the Oxford/Azores polymorphisms which are highlighted by coloured boxes. Numbers to right of sequences indicate amino acid number for the end of line shown. Amino acid sequences obtained from http://www.phytozome.net/ and alignments performed using online alignment tool www.genome.lbl.gov/vista.

A.


Ox MDSNNHLYDPNHISSSGSGLLRFRSAPSSVLAAFVDEDKSGFDSNRLLSRFVSSNGGNNH 60
Az MDSNNHLYDPNHISSSGSGLLRFRSAPSSVLAAFVDEDKSGFDSNRLLSRFVSSNGGNDH 60
Ox LGSPSRSEFEDKSPVSLTNTSVSYAATLPPPQQPDPSSFLGLPPHYPRHNNSKGMMNSIG 120
Az LGSPSRSEFEDKSPVSLTNTSVSYAATLPPPQQPDPSSFLGLPPHYPRHNNSKGMMNSIG 120
Ox LDQFLDMNNNHHTKPVESNLLRQSSSPAGMFTNLSGQNGYGSMRNLMNYDEDEESPSNSN 180
Az LDQFLDMNNNHHTKPVESNLLRQSSSPAGMFTNLSGQNGYGSMRNLMNYDEDEESPSNSN 180
Ox GLRRHCSLSSRPPSSLGMLPQIPEITTETNFQYSHWNDPSSFIDNLSSLKRETEDDGKLF 240
Az GLRRHCSLSSRPPSSLGMLSQIPEITTETNFQYSHWNDPSSFIDNLSSLKRETEDDGKLF 240
Ox HGAQNGESGNRMQLLSHHLSLPKSSSTASDMVSVDKFLHLQDSVPCKIRAKRGCATHPRS 300
Az HGAQNGESGNRMQLLSHHLSLPKSSSTASDMVSVDKFLHLQDSVPCKIRAKRGCATHPRS 300
Ox IAERVRRTRISERMRKLQELVPNMDKQTNTSDMLDLAVDYIKDLQRQYKILNENRANCKC 360
Az IAERVRRTRISERMRKLQELVPNMDKQTNTSDMLDLAVDYIKDLQRQYKILNENRANCKC 360
HLH
Ox LNKETI 366
Az LNKETI 366

B.

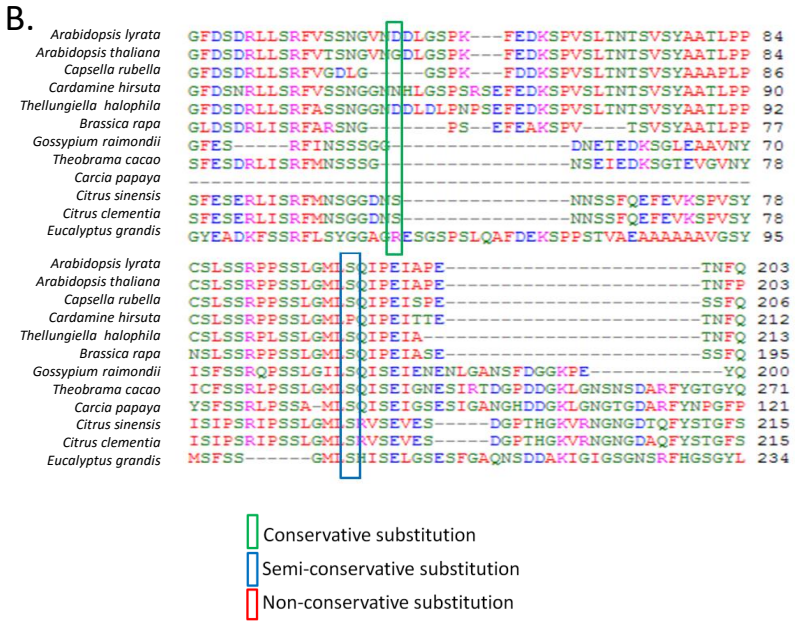
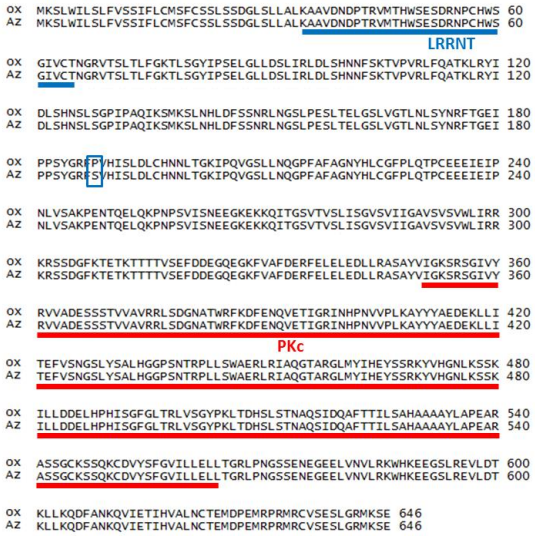
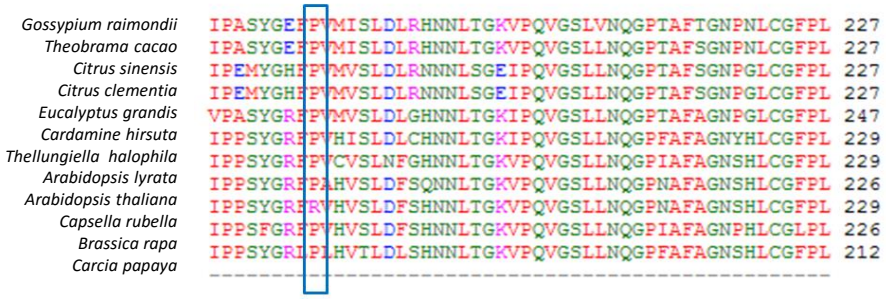

Arabidopsis lyrata
Arabidopsis thaliana
Capsella rubella
Cardamine hirsuta
Thellungiella halophila
Brassica rapa
Gossypium raimondii
Theobrama cacao
Carcia papaya
Citrus sinensis
Citrus clementia
Eucalyptus grandis
Conservative substitution
Semi-conservative substitution
Non-conservative substitution

Figure 5.5 Amino acid sequence polymorphisms within ChLRR. Comparison of Oxford and Azores amino acid sequence for ChLRR (**A**); polymorphisms are highlighted by coloured boxes and the location of protein domains. The blue line presents the leucine rich repeat N terminal (LRRNT) domain and the red line represents the protein kinase C-like domain (PKc). Alignment for amino acid sequences of species within the Malvidae (**B**) around the Oxford/Azores polymorphisms which are highlighted by coloured boxes. Numbers to right of sequences indicate amino acid number for the end of line shown. Amino acid sequences obtained from http://www.phytozome.net/ and alignments performed using online alignment tool www.genome.lbl.gov/vista.

A.



B.



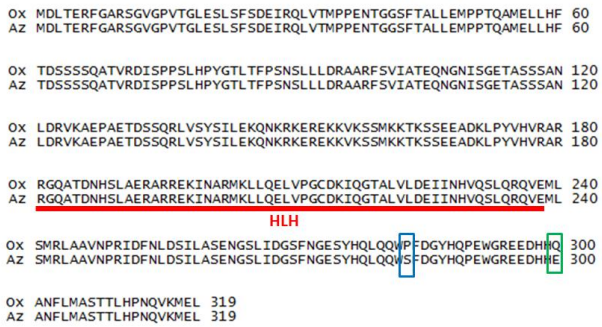
Conservative substitution

Semi-conservative substitution

Non-conservative substitution

Figure 5.6 Amino acid sequence polymorphisms within ChbHLH. Comparison of Oxford and Azores amino acid sequence for ChbHLH (**A**); polymorphisms are highlighted by coloured boxes and the location of helix loop helix DNA binding domain (HLH) shown by red line. Alignment for amino acid sequences of species within the Malvidae (**B**) around the Oxford/Azores polymorphisms which are highlighted by coloured boxes. Numbers to right of sequences indicate amino acid number for the end of line shown. Amino acid sequences obtained from http://www.phytozome.net/ and alignments performed using online alignment tool www.genome.lbl.gov/vista.

A.



B.

Conservative substitution

Semi-conservative substitution

Non-conservative substitution

There are three protein sequence differences between Oxford and Azores in ChRRM; two of these are within one of the RNA recognition motifs (Figure 5.3 A). The conservative polymorphism outside the motif (alanine in Oxford, serine in Azores) does not lie in an area of high conservation. Only three other species have Oxford's alanine (*A. lyrata, C. rubella* and *E. grandis*); however *T. halophila* and *A. thaliana* share a serine with Azores (Figure 5.3 B). The non-conservative substitution within the RNA recognition motif (arginine in Oxford, leucine in Azores) is in an area of high conservation (Figure 5.3 B). The amino acids immediately flanking the substitution site are conserved amongst all species and the leucine of Azores is present in *T. halophila, A. lyrata, A thaliana* and *C. rubella* (Figure 5.3 B). The conservative substitution in the RNA recognition motif (asparagine in Oxford, aspartic acid in Azores) is within a highly conserved region; amino acids immediately flanking the substitution and the aspartic acid of Azores are conserved in all species (Figure 5.3 B); it is the Oxford protein that has diverged from the rest.

The two amino acid substitutions in ChFBH4 have occurred outside of the helix hoop helix DNA binding domain (Figure 5.4 A). The conservative substitution (aspargine in Oxford, aspartic acid in Azores) is within a region which is not highly conserved (Figure 5.4 B). The semi-conservative substitution (proline in Oxford, serine in Azores) is within a conserved region; the serine in Azores is also found in all other species analysed here. (Figure 5.4 B).

The single semi-conservative amino acid substitution (proline in Oxford, serine in Azores) found in ChLRR is in neither the leucine rich repeat N terminus domain or in the protein kinase C-like domain (Figure 5.5 A). This substitution does however affect a region which is highly conserved; the proline found in Oxford is also found in all other species shown here with the exception of *A. thaliana* which has an arginine at that position (Figure 5.5 B).

There are two differences between the protein sequences of Oxford and Azores for ChbHLH but neither result in an alteration to the helix-loop-helix DNA binding domain (Figure 5.6 A). The semi conservative substitution (proline in Oxford, serine in Azores) is in a partially

conserved region; the proline in Oxford is also in *A. lyrata, A. thaliana, C. rubella, T. halophila* and *B. rapa* but not in any of the other species (Figure 5.6 B). The conservative substitution (glutamine in Oxford, glutamic acid in Azores) is also in a partially conserved region; the glutamine found in Oxford is also shared by *A. lyrata, A. thaliana* and *C. rubella.*

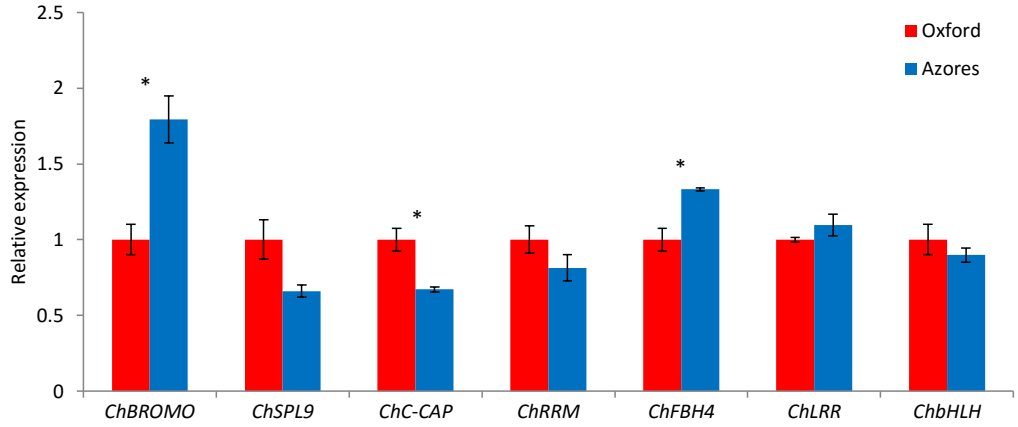
This analysis is not applicable to ChC-CAP as none of the polymorphisms in the coding sequence resulted in an alteration in protein sequence (Table 5.2)

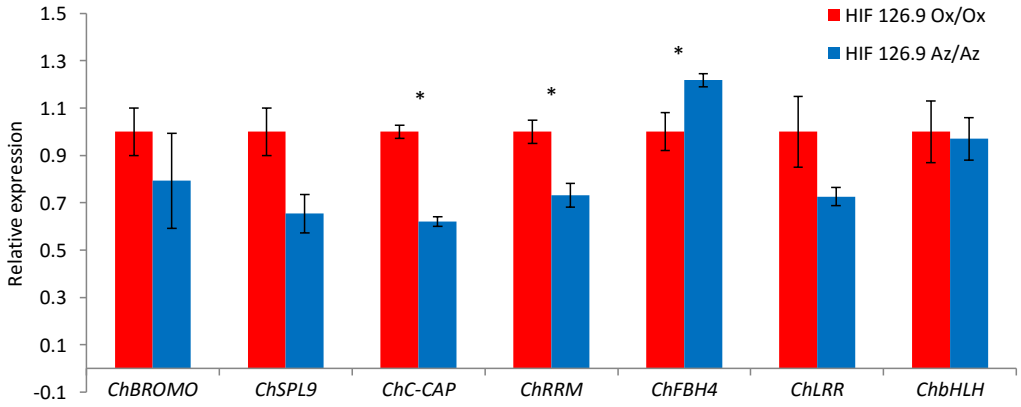
From this analysis of amino acid sequence polymorphisms between the proteins of Oxford and Azores the genes that appear to warrant further investigation are *ChSPL9, ChRRM, ChFBH4 and ChLRR.* The protein products of these genes all have polymorphisms within regions that the protein alignments suggest are important. In the case of ChRRM there is a polymorphism within RNA recognition motif adding further evidence that protein activity could have allelic variance in activity. Protein sequence variation in the ChBROMO, ChbHLH and ChC-CAP proteins did not occur within either known domains or conserved regions and therefore do not support the selection of these genes as candidates for underlying the QTL.

5. 5. Candidate gene selection by variable expression levels

The above analysis on coding sequence polymorphisms looks for candidate genes on the basis that the QTL gene is modified because of a coding sequence alteration. However it is possible that the QTL effect is exerted not by a change in protein structure but by variable expression levels. This is a common mechanism by which variable morphology is produced. The sequence analysis approach taken above is harder to implement in this case because the effect of changes to non-coding regulatory regions around genes is harder to predict. An experimental approach to selecting candidate genes is to quantify the levels of each gene's transcript. A gene in

Figure 5.7 Relative expression levels of the candidate genes. For each gene transcript levels are compared between the parental strains (A) and also between the alternate homozygous genotypes of HIF 126.9 (B). Error bars indicate standard error and significant differences between genotypes (t-test $P < 0.05$) are denoted by an asterisk (*).


2.5
2
1.5
1
0.5
0
Relative expression
Oxford
Azores
*
*
*
ChBROMO
ChSPL9
ChC-CAP
ChRRM
ChFBH4
ChLRR
ChbHLH


1.5
1.3
1.1
0.9
0.7
0.5
0.3
0.1
-0.1
Relative expression
HIF 126.9 Ox/Ox
HIF 126.9 Az/Az
*
*
*
ChBROMO
ChSPL9
ChC-CAP
ChRRM
ChFBH4
ChLRR
ChbHLH

which variable transcript levels correlate with QTL genotype is a strong candidate for being the gene underlying that QTL.

5. 5. 1. Quantitative real time PCR

Variable expression levels in the seven candidate genes (Table 5.2) were assayed using quantitative real time PCR on the whole shoot at the point leaf 5 first emerges. For each gene expression levels were assayed in both Oxford and Azores strains and in each of the alternate homozygous genotypes of HIF 126.9 (Figure 5.7). If transcript level varied significantly in both comparisons in the same direction, then that result would be considered supporting evidence that the gene could be responsible for the QTL.

According to the quantitative real time PCR analysis the expression levels of *ChBROMO* were significantly higher in Oxford than they were in Azores (t test $P < 0.01$) but there was no significant difference between the alternate genotypes of the HIF (Figure 5.7, t test $P=0.35$). The expression levels of *ChSPL9* did not differ significantly between either the parental strains (T test $P=0.07$) nor the different HIF genotypes (Figure 5.7, t test $P = 0.08$). The expression levels of *ChC-CAP* were significantly different in both comparisons; Oxford has significantly more transcript than Azores (t test $P = 0.02$) and the Oxford HIF had significantly more transcript than the Azores HIF (Figure 5.7, t test $P = 0.01$). The expression levels between *ChRRM* displayed no significant differences between Oxford and Azores (t test $P = 0.16$) but the HIF 126.9 homozygous for the Oxford allele had significantly higher levels of transcript than HIF 126.9 homozygous for the Azores allele (Figure 5.7, t test $P = 0.01$). With regard to *ChFBH4*; Oxford had significantly higher expression than Azores (t test $P = 0.04$) and likewise the Oxford genotype of HIF 126.9 had significantly higher expression than the Azores counterpart (Figure 5.7, t test $P = 0.04$). The transcript of *ChLRR* was more abundant in Azores than it was in Oxford but not significantly more (t test $P = 0.11$) and there was no significant difference in transcript levels between the Oxford and Azores genotypes of HIF 126.9 (Figure 5.7, t test $P = 0.07$).

Expression of *ChbHLH* was not significantly different between Oxford and Azores (t test P = 0.49) and not significantly different between either of the alternate HIFs (Figure 5.7, t test P = 0.57)

Of all the genes in which expression levels were assayed with quantitative real time PCR only *ChC-CAP* and *ChFBH4* showed significantly different expression levels that co-segregated with genotype. Expression levels of *ChFBH4* were higher in the presence of the Oxford QTL allele and expression levels of *ChC-CAP* were higher in the presence of the Azores allele whether that was in the background of the parental strain or HIF 126.9 (Figure 5.7). However in both cases, despite being statistically significant, it is difficult to ascertain whether relative expression levels are sufficiently different enough to elicit the QTL effect. For *ChC-CAP* expression in Oxford was only ca. 1.5 times higher than in Azores and expression in the Oxford HIF genotype was only ca. 1.6 times higher than in the Azores genotype (Figure 5.7). Likewise for *ChFBH4,* expression in Azores was only ca. 1.3 times higher than in Azores and expression in the Oxford genotype of the HIF was only ca. 1.2 higher than in the Azores genotype of the HIF. Transcript levels of *ChBROMO* were significantly different between the parental strains but not between the alternate HIF genotypes. The reverse situation was true regarding transcript levels of *ChRRM*; there was a significant difference between the genotypes of the HIF but not between Oxford and Azores. In these cases it may be that expression of the gene is dependent on epistatic relationships within the genome that are created or absent in the HIF or parents. However, similar to *ChC-CAP* and *ChFBH4* differences in expression levels are not big enough to draw conclusions on their contribution to the QTL effect. It is unclear whether these relatively small changes in transcript level could elicit the QTL effect observed but it may be the case that the QTL effect is produced as a result of fine tuning of expression rather than large scale changes in transcript levels. As for, *ChSPL9, ChLRR* and *ChbHLH* differences in transcript level were not significantly different and therefore if it is the case that one of these genes is underlying the QTL then the effect is unlikely to be attributed to polymorphism within regulatory regions of the gene.

5. 6. Initial selection of candidate genes

From initial analysis of the reported function of genes within the fine mapped region there were two genes that stood out as candidates for the QTL effect; *ChSPL9* and *ChFBH4* (Table 5.1). Both these genes are known to have roles associated with transitions through different stages of plant growth. This twinned with the fact that the QTL has an effect on the heteroblastic series, means that *ChFBH4* and *ChSPL9* warrant further investigation by experimental work. Analysis of sequence polymorphisms and their potential effect on the functional domains of protein products also implicated *ChFBH4* and *ChSPL9* (Table 5.2). Both genes had polymorphisms within regions that displayed strong conservation within the Malvidae (Figure 5.2 and Figure 5.4). This analysis also showed that *ChRRM* had protein sequence polymorphisms within one of the RNA recognition motifs; such modifications could alter protein function and therefore should be considered as another strong candidate for the QTL. The analyses of the quantitative real time PCR experiments were mostly inconclusive but do support the selection of *ChFBH4* as a candidate gene.

It was planned that if experimental procedures failed to yield promising results regarding the identification of the quantitative trait gene, additional genes would be selected for experimentation using the results from the analyses above.

Whole genomic fragments of *ChFBH4, ChSPL9* and *ChRRM* were genetically isolated and cloned for transformation of *A. thaliana.* However technical difficulties associated with amplifying an error free genomic fragment for cloning prevented the completion of this with regards to *ChRRM.* The results of these experiments are presented in the next section.

5. 7. Transformation of *A.thaliana* with whole genome fragments of *ChSPL9* and *ChFBH4*

Attempts were made to transform whole genomic fragments of each allele of *ChSPL9* and *ChFBH4* into *C. hirsuta* and *A. thaliana.* The decision to transform *A. thaliana* was made partly

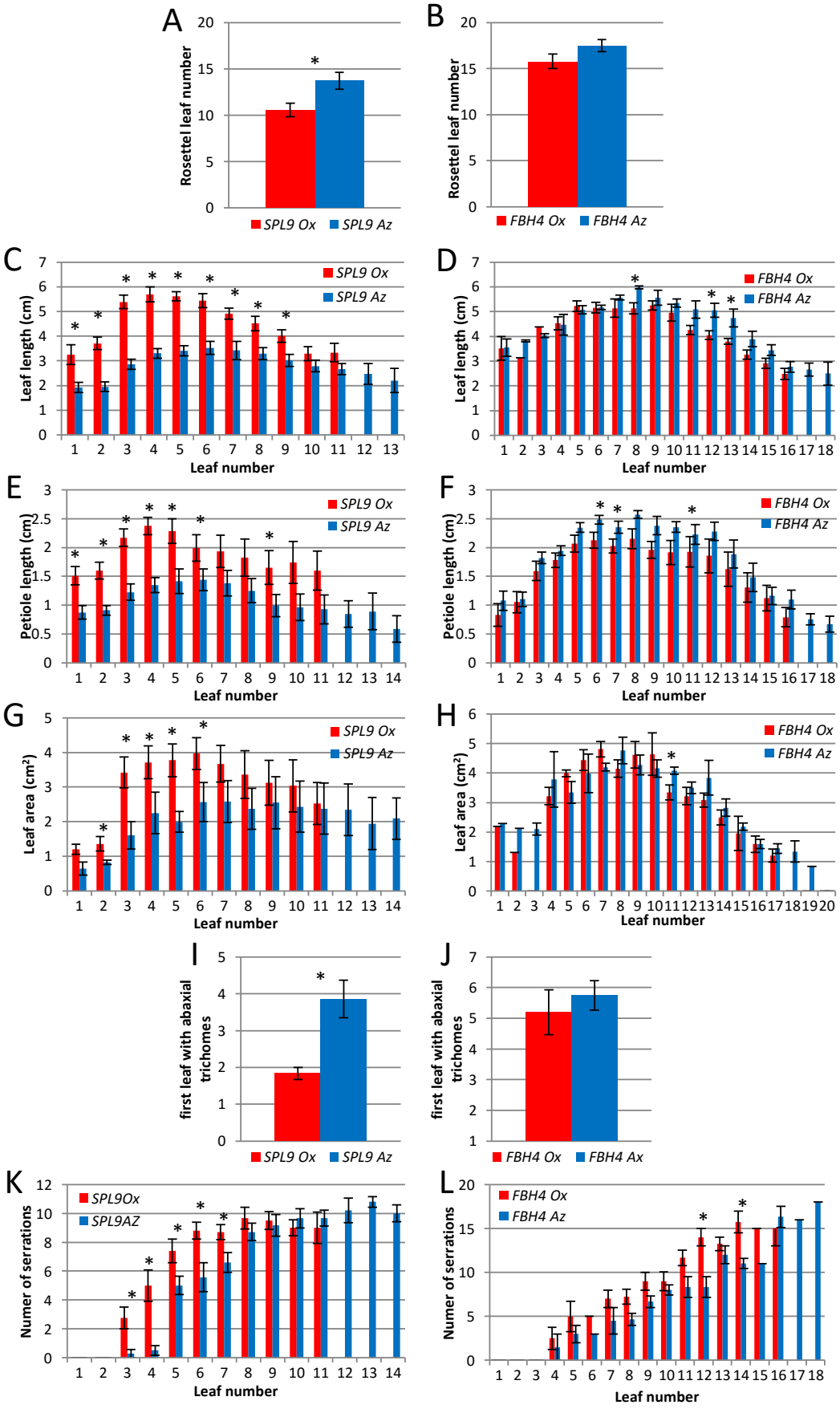
for practical reasons. While both species are amenable to transformation by floral dip with *Agrobacterium tumefaciens*, efficiency is better with *A. thaliana* than it is with *C. hirsuta* (0.3% and 0.1% respectively, Hay et al. 2014). It is also not unreasonable to hypothesize that a QTL affecting leaf complexity in *C. hirsuta* will also affect complexity in *A. thaliana* since the two are closely related and many leaf development programmes operating in one species have been found to operate in the other (Canales et al., 2010). In practice transformation of *C. hirsuta* was unsuccessful due to technical difficulties, however interesting results were obtained from the transformation of *A. thaliana.*

The Col-0 strain of *A. thaliana* was used for this purpose and the whole genome fragments consisted of the whole sequence spanning from the transcription termination site of the gene immediately up stream to the transcription start site of the gene immediately downstream of the candidate gene in question. The cloned fragments for *ChSPL9* were 5.3kb and for *ChFBH4* it was 7.5kb. The successfully transformed T1 lines were analysed for the following traits: rosette leaf number, leaf length, petiole length, leaf area, appearance of abaxial trichomes, and number of leaf margin serrations. Fifteen plants of each genotype for each transcript were analysed, the results are presented below.

Rosette leaf number - *A. thaliana* T1 plants transformed with the Azores allele of *ChSPL9* (*SPL9Az*) produced on average 13.7 rosette leaves which was significantly greater (t test P = 0.02) than the average 10.6 leaves produced by T1 plants transformed with the Oxford allele of *ChSPL9* (*SPL9Ox*) (Figure 5.8 A).

In contrast there was no significant difference (t test P = 0.16) in rosette leaf number between T1 plants transformed with the Oxford allele of *ChFBH4* (*FBH4Ox*, 15.8 leaves) and the Azores allele of *ChFBH4* (*FBH4Az*, 17.5 leaves, Figure 5.8 B).

Figure 5.8 Phenotypic analysis of *ChSPL9* and *ChFBH4 T1* plants. Oxford and Azores alleles of Ch*SPL9* and *ChFBH4* were transformed into *A. thaliana* Col-0 and measured for the following leaf traits: rosette leaf number (A, B), leaf length (C, D), petiole length (E, F), leaf area (G, H), appearance of abaxial trichomes (I, J) and number of leaf margin serrations (K, L). In all cases *ChSPL9* transformation results are on the left column of charts and *ChFBH4* results are on the right column. All error bars indicate standard error of the mean and significant differences (t test, $P < 0.05$) are denoted by asterisks (*).


A
Rosettel leaf number
SPL9 Ox
SPL9 Az
B
Rosettel leaf number
FBH4 Ox
FBH4 Az
C
Leaf length (cm)
Leaf number
SPL9 Ox
SPL9 Az
D
Leaf length (cm)
Leaf number
FBH4 Ox
FBH4 Az
E
Petiole length (cm)
Leaf number
SPL9 Ox
SPL9 Az
F
Petiole length (cm)
Leaf number
FBH4 Ox
FBH4 Az
G
Leaf area (cm²)
Leaf number
SPL9 Ox
SPL9 Az
H
Leaf area (cm²)
Leaf number
FBH4 Ox
FBH4 Az
I
first leaf with abaxial trichomes
SPL9 Ox
SPL9 Az
J
first leaf with abaxial trichomes
FBH4 Ox
FBH4 Ax
K
Numer of serrations
Leaf number
SPL9Ox
SPL9AZ
L
Numer of serrations
Leaf number
FBH4 Ox
FBH4 Az

Leaf length - The leaves of *SPL9Ox* transformants were significantly longer than the leaves of *SPL9Az* transformants at leaf nodes 1 to 9 (Figure 5.8 C). This difference was most pronounced at leaf three where there was a difference in mean length of 2.53cm.

Only at leaves 8, 12 and 13 was there significant differences in leaf length between the leaves of *FBH4Ox* and *FBH4Az* transformed plants; at these leaves the Azores allele conferred longer leaves (Figure 5.8 D)

Petiole length - At leaves 1-6 and at leaf 9 the petioles of *SPL9Ox* plants are significantly longer than those of *SPL9Az* plants (Figure 5.8 E). The largest difference is at leaf 3; 1.22cm.

Significant differences in petiole length between *FBH4Ox* and *FBH4Az* pla1stnts occur at leaves 6, 7 and 11 (Figure 5.8 F). *FBH4Az* plants have longer petioles than *FBH4Az* plants.

Leaf area - The leaves of plants transformed with *SPL9Ox* are significantly larger than the leaves of plants transformed with *SPL9Az* at leaves 2 , 3, 4, 5 and 6 (Figure 5.8 G). The largest difference exists at leaf three where the mean leaf area of *SPL9Ox* is 3.06cm^2 larger than the mean leaf area of *SPL9Az* plants.

The leaf area of leaf 11 is significantly larger in plants transformed with *FBH4Az* than in plants transformed with *FBH4Ox*, but there at no other leaf node is there a significant difference (Figure 5.8 H).

Appearance of abaxial trichomes - In *Arabidopsis*, vegetative phase change is marked by the appearance of trichomes on the abaxial surface of leaves (Telfer et al. 1997). On average the first appearance of trichomes on the abaxial (lower) surface of the leaf occurs after 2.4 leaves in plants transformed with *SPL9Ox* and after 3.5 leaves in plants transformed with *SPL9Az*, this difference is significant (Figure 5.7 I, t test P=0.04).

The average leaf at which abaxial trichomes appeared did not significantly vary between T1 plants transformed with either version of *ChFBH4* (Figure 5.7 J, *FBH4Ox*: 5.2, *FBH4Az*: 5.5, t test P = 0.53).

Number of leaf margin serrations - Serrations along the leaf margin were counted for all T1 plants. The plants transformed with *SPL9Ox* had significantly more serrations than the plants transformed with *SPL9Az* at leaves at nodes 3 – 7 (Figure 5.7 K). The largest difference came at leaf 4 where there was a difference of 4.5 serrations between the mean values of the *SPL9Ox* and *SPL9Az* transformants.

The number of serrations on the leaves of plants transformed with the *FBH4Ox* did not differ significantly from the leaves of the *FBH4Az* transformants except at leaves 12 and 14 (Figure 5.7 8)

5. 8. Discussion

The results here give strong support towards the identification of the gene underlying QTL-LG4 as being *ChSPL9.* Transformation of *A. thaliana* with the different whole genome sequences of the Oxford and Azores alleles of *ChSPL9* resulted in different leaf phenotypes in T1. *SPL9Ox* T1 plants produced more rosette leaves, longer leaves, longer leaf petioles, larger leaves, more leaves with abaxial trichomes and more leaf margin serrations than the T1 *SPL9Az* plants. These results demonstrate that the alternative alleles are functionally different from one another and I can conclude that it is extremely likely that *ChSPL9* is responsible for the QTL.

The results do not suggest that *ChFBH4* is responsible for producing the QTL effect. Significant differences in trait values between *FBH4Ox* and *FBH4Az* transformants were much rarer than they were for *ChSPL9* transformants. When significant differences were observed they did not follow the model predicted by QTL mapping and neither did they follow the observations observed during QTL mapping and QTL characterisation.

The likelihood that the QTL is bought about by the action of *ChFBH4* is further diminished by the findings from the QTL characterisation experiment. *ChFBH4* affects shoot

development by promoting the expression of *CONSTANS* (*CO*) (Ito et al, 2012). However the CO protein can only influence downstream floral inducing genes when a sufficiently long photoperiod is detected (Kobayashi and Weigel 2007; Amasino 2007; Imaizumi 2010). When analysing the alternate genotypes of HIF 126.4 in short and long day conditions showed that QTL effect is independent of photoperiod (Figure 4.10 – 4.11) therefore the QTL gene is unlikely to be a factor in such a photoperiod dependent developmental pathway.

Further evidence that *ChSPL9* is responsible for the QTL is evident from the similarities of these results and the results obtained in Chapter 3 when the QTL was mapped and when the QTL was characterised. The model produced by the multi-trait QTL mapping predicted that QTL effect is largest in the early leaves and then diminishes as the shoot develops (Figure 3.11). This was mirrored in the analysis here; the largest differences in phenotype of *SPL9ox* and *SPL9Az* were observed in the early leaves, and in particular leaf 3 (Figure 5.7).

Phenotypic analysis of HIF 126.4 in long day conditions showed that the Oxford QTL allele increased leaflet area and increased leaf length. This is comparable to the observation that *SPL9Ox* results in larger and longer leaves than *SPL9Az.*

The developmental pathways that lead to the formation of leaflets and the production of leaf margin serrations share many of the same elements (Bar and Ori, 2014) and marginal serration is a trait that varies along the heteroblastic series. This provided the rationale in recording the extent of margin serration on the leaves of *ChSPL9* T1 plants. In both phenotypic analysis of HIFs and in the analysis here of *ChSPL9* T1 plants, the Oxford allele has increased the number of these analogous structures which adds to the evidence that *ChSPL9* underlies the QTL.

Throughout this thesis variation in phase change has been implicated repeatedly in creating variation in leaf traits. Analysis of the intraspecific variation between the different strains showed that the leaves of different strains reach a mature leaf shape earlier in development than others. The QTL effect seems to be the appearance of more adult features in younger leaves; under the influence of the Oxford QTL allele leaves are larger, longer, wider and more dissected

at an earlier stage. However for all these analyses there has been no marker by which to observe phase change; we have had to frame our discussion within the context of heteroblastic change. Analysis of the effect of QTL alleles in *A. thaliana* means we can use the appearance of abaxial trichomes to better measure the onset of the juvenile phase of vegetative growth. In *Arabidopsis*, the switch to adult vegetative phase change is marked by the appearance of trichomes on the abaxial side of a leaf (Telfer et al. 1997). By demonstrating that *ChSPL9ox* induces the earlier appearance of abaxial trichomes I have confirmed that the QTL accelerates the transition through the vegetative phases of growth.

At present we can only hypothesize about how allelic variation in *ChSPL9* has been created; it could be due to any of the seven polymorphisms identified in and around the locus. I have previously demonstrated that it is unlikely that the QTL effect has arisen from changes in the ability of miR156 to target *ChSPL9* transcript as no sequence polymorphisms were located within the target site. It is possible that the ability of ChSPL9 to promote expression of *ChmiR172* has been altered. In this case I would expect SPL9Ox to have a stronger effect on expression relative to SPL9Az because miR172 induces adult identity (Wu et al, 2009) and we have shown here that SPL9Ox accelerates the onset of adult morphology. In addition to control via micro RNAs, *ChSPL9* has also been shown to be under the influence of gibberellin (Yu et al, 2012) therefore this is another mechanism by which the divergent allelic effect may have been bought about.

One polymorphism was identified in this chapter has strong potential for creating the QTL effect. A polymorphism was identifies that created an amino acid substitution at a position in the protein sequence that is conserved among the Malvids except in *C. clementia* and *C. sinensis*. This is potentially of significance as these species have been domesticated, and the polymorphism at this position may have been the subject of selection, therefore there is a possibility that the polymorphism seen in the Azores has implications for selection and adaptation. The amino acid retained by most of the Malvids, including the Oxford strain at this position is glutamic acid, the *Citrus* species have a glycine, and the Azores strain possesses a glutamine (Figure 5.2). A search of the published literature failed to discover more about the

implications of these amino acid substitutions and therefore further experimentation is required to verify the functional significance of the polymorphism.

The results presented here provide strong evidence that *ChSPL9* is the quantitative trait gene; however they are not fully conclusive. The predicted effect produced by multi trait QTL analysis indicates that the Oxford QTL allele induces production of rosette leaves and this was replicated in QTL validation results (Figure 4.8). However in the analysis of *A. thaliana ChSPL9* transformants the Azores allele produced significantly more rosette leaves than the Oxford allele; a reversal of previous results.

5. 8. 1. Further work

The alternative *ChSPL9* alleles should be transformed into *C. hirsuta.* It is possible that the divergent allelic effects observed in *A. thaliana* may not be reproduced in *C. hirsuta* because of interactions with the genetic background. An appropriate approach would be to transform each genotype of a HIF segregating for the QTL effect with the opposite allele, i.e. a HIF homozygous for the Oxford QTL allele would be transformed with *ChSPL9Az* and a HIF homozygous for the Azores QTL allele would be transformed with *ChSPL9Ox*. If the QTL effect is complemented by these transformations then this could provide complete formal proof that QTL is under control of *ChSPL9.*

Chapter 6:
Discussion

6. 1. Analysis of intraspecific variation in leaf shape in *C. hirsuta*

My work on understanding the basis of intraspecific variation of the *C. hirsuta* leaf began with an analysis of the range of phenotypes that exist. The rosette leaves of six strains from geographically disparate locations were analysed, resulting in the immediate observation that variation existed amongst leaves of the same plant at different stages of development. All strains displayed this form of longitudinal asymmetry along the shoot – the earliest leaves had few leaflets, short length and small area; as growth proceeded leaves gained more leaflets, increased in area and got longer and wider. This age-dependent leaf shape progression, or heteroblastic development, is common amongst plants (Poethig, 2013) and forms the context within which leaf shape variation needs to be assessed. Analysis and genetic dissection of variation at one fixed point in development (i.e. if we selected one leaf to analyse as a reference) would not capture the dynamic interactions that act during development to produce variable leaf morphology.

When comparing equivalent leaf nodes amongst the different strains there was an enormous amount of variation. However the principal driver of this was the variable rates at which the different strains progress through heteroblastic development. Some strains progressed through heteroblastic development in relatively few leaf nodes, while the leaves of others displayed a more gradual change. A rough measure of rate of change in the different strains was used to show that they did indeed have variable rates of trait change. No conclusion was reached on which strains reached a mature looking leaf phenotype the fastest; this could indicate that different traits respond to age-dependent regulation in variable ways in different traits. Equally, a better measure of rate of trait change may facilitate the identification of strains that develop through heteroblasty the fastest or slowest, giving a better idea of the intraspecific range that exists for this phenotype.

The genetic determinants of variable leaf shape were identified using QTL mapping with an F6 RIL population derived from the Oxford and Azores strains. Phenotypic analysis of this RIL population gave indications about the genetic control of variable leaflet number. Leaflet

production on different leaves is a highly heritability trait, as is rosette leaf number (Table 3.3). Examination of phenotypic correlations between leaflet numbers at different leaf nodes indicated a high degree of shared genetic control (Figure 3.8). Leaflet numbers also correlated with rosette leaf numbers suggesting that the genetic factors determining leaflet production also affect characteristics of the main shoot. This result is not surprising since leaflet production is an age-dependent trait, and also because regulators of shoot apical meristem activity have also been found to be necessary for the development of the compound leaf (e.g. *KNOX* genes; Hay and Tsiantis, 2010).

My analysis of leaf shape variation focused on measures of quantity, length, width and area. Undoubtedly, these are important traits for natural variation but further analysis of natural variation would benefit from methods to quantify shape variance. Langlade et al. (2015) successfully used coordinates at key landmarks of the leaf margin that could be input into principal component analysis to create a quantitative measure that defines variation in leaf shape amongst species of *Antirrhinum*. A particularly innovative example of this was published recently by Chitwood et al. (2014). The authors employed elliptical Fourier descriptor analysis to quantify tomato leaflet shape and the data produced were used to create principal component axes that could describe variation. This was applied to over 33,000 leaflets from a set of tomato introgression lines and used to build a model that quantified the changes in shape across successive leaves in the heteroblastic series. QTL mapping was used to determine the genetic factors underlying variation in this model. The study went beyond measures of size and complexity to demonstrate that shape variation in a single leaf outline can be separated into distinct genetic and heteroblastic components. Similar work with the leaves of *C. hirsuta* would certainly yield an improved understanding of how leaf shape has diverged.

6. 2. The genetic architecture of variable leaf shape

QTL analysis of the Oxford x Azores RIL population successfully identified loci underlying variation in leaflet and rosette leaf number. Initially single trait QTL mapping was able to identify QTL for leaflets on different leaves a co-located to the same genomic regions, suggesting they share common genetic control. Multi trait mixed model QTL mapping was then used to directly model the genetic correlations, providing analysis with greater power that reflects a more realistic model than a collection of single trait analyses (van Eeuwijk et al., 2010). At the end of the analysis, six QTL were detected that had statistically significant effects on leaflet number.

All of the QTL affected leaflet number on multiple leaves at successive nodes. This suggests that the different determinants of leaf shape are active at different stages of heteroblastic development. QTL on linkage group 2 and the lower portion on linkage group 8 affect leaflet number throughout the heteroblastic development whereas the QTL on linkage groups 3 and 4 only affect leaflets early on in growth, and the QTL on linkage group 6 and upper portion of linkage group 8 affect leaflet number later on in development (Figure 3.11). This demonstrates that leaf shape is affected by a dynamic suite of regulators that changes with heteroblastic development.

There was a high degree of commonality in QTL that affected leaflet and rosette leaf number with only one of the five QTL affecting leaflet number independently of rosette leaf number, reflecting the results of the phenotype correlation analysis (Figure 3.8). Pleiotropy in this regard is unsurprising as QTL that produce fewer leaflets are bolting earlier and have therefore transitioned faster through development, thus promoting more mature leaves with a greater number of leaflets. This expectation of QTL alleles that increase rosette leaf number will decrease leaflet number and vice versa is met with the three QTL on linkage groups 6 and 8. However, the converse situation is observed for the QTL on linkage groups 3 and 4; both QTL affect trait value in the same direction (Figure 3.11). This could indicate that heteroblastic development is not

always coupled to the transition to reproductive growth. Further work needs to be done on the correlation between flowering time and rosette leaf production to get a comprehensive explanation.

The QTL mapping results suggest that leaflet production was not under selection during the establishment of these strains in their respective locations. Neither strain possesses QTL alleles that act in one direction. If leaflet number had been subject to selection then it would be expected that the genome of one strain would only have QTL alleles working to maximise or minimise leaflet production according to the direction of selection. Orr (1998) describes a test by which the number of plus and minus alleles for a trait can be used to determine whether a phenotypic difference was caused by natural selection or random genetic drift. However, six QTL do not provide enough power to perform this test.

A factor that was not considered in this analysis is epistasis; interactions between QTL and the genetic background can have a large effect on phenotype. Kroymann and Mitchell-Olds (2005) demonstrated that epistasis significantly influenced the effect of a QTL for *Arabidosis* growth. QTL analysis procedures are available that detect significant interactions between QTL. Further understanding of the genetic architecture of leaf shape could be achieved if epistatic interactions are characterised.

6. 3. Validation and characterisation of QTL effect

Of the six QTL identified during QTL mapping I was able to validate the QTL on linkage groups 2, 4 and 8 using HIFs. This demonstrates that the results yielded by the QTL analysis are reliable and can be used to make inferences on the genetic architecture underlying leaf shape.

HIFs were also used to comprehensively characterise the effect of LG4-QTL and it was revealed that this QTL has pleiotropic effect on other leaf shape traits. Not only does the Oxford

allele confer greater numbers of leaflets, it also increases lateral leaflet area, rachis length and petiolule length relative to the Azores allele. The Oxford allele is promoting earlier development of mature looking leaf phenotype. These additional traits were increased above a trait value that is ever reached by plants of the Azores genotype at any time during vegetative development. This could suggest that LG4-QTL effect operates, to a certain degree, independently of heteroblastic development.

LG4-QTL effect was also analysed in short day conditions to analyse whether the QTL was affected at all by differential photoperiod. The QTL effect was very similar to that observed in long day conditions. This would suggest that the QTL effect is independent of those pathways affected by day length. However, it should be noted that leaf damage and senescence limited the effectiveness of this experiment. This does not mean that the QTL effect is independent of environmental effects. The QTL was validated and fine mapped in a controlled environment room in the Department of Plant Sciences, University of Oxford but characterised in greenhouse of the Max Planck institute for Plant Breeding Research, Cologne, and the QTL effect was different in each condition. There is clearly a gene by environment interaction for this QTL.

Not in any of the QTL validation, fine mapping or characterisation experiments did the observed QTL effect match the effect that was predicted in the QTL mapping analysis. There are numerous possible reasons for this. Despite the use of controlled environment rooms, it is possible that factors such as seed quality, soil quality, small temperature deviations, deterioration of light sources amongst others that could introduce small variations between experiments. It is also possible that additional small QTL effects that were not detected in the QTL analysis are detected when background genetic variation is removed in the HIFs. Another factor could be that epistatic interactions between the detected QTL (and background interactions) may be responsible for small modifications of the QTL effects. These variables could also explain why the attempted validation of LG3-QTL was inconclusive.

Identification of the genetic basis of the other QTL responsible for variable leaflet number detected and validated here will allow further refining of the genetic networks and pathways identified from the analysis of mutants in model species.

6. 4. QTL fine mapping and the identification of the quantitative gene

The QTL on linkage group 4 (QTL-LG4) was chosen for fine mapping and elucidation of its molecular basis. Fine mapping was able to delimit the QTL interval to a genomic region of 48.6 kb spanning 15 genes. While fine mapping, evidence emerged that there may be a two QTL closely linked; one that affects leaflet number and one that affects rosette leaf number (and could also affect leaflet number). Characterising QTL effect in two HIFs, one that segregated only for one QTL and another that segregated for an additional putative QTL, supported the presence of a second QTL. This situation is possible because at the resolution of QTL mapping, no distinction can be made between the pleiotropic effects of a single QTL and the effects of two closely linked QTL. Unfortunately no HIFs, or recombinant derivatives, were available to validate the second putative QTL independently of the leaflet QTL. Until this has happened we cannot confirm the presence of this second QTL.

Of the 15 genes in the fine mapped QTL region the most attractive candidate in terms of function was *ChSPL9*. This is because *A. thaliana* orthologue has been shown to be a key regulator of the transition from juvenile vegetative growth to adult vegetative growth through interaction with the micro RNAs mir156 and mir172 (Wu et al., 2009). Transformation of *A. thaliana* with both alleles of *ChSPL9* revealed allelic effects. Relative to *SPL9Az, SPL9Ox* increased leaf length, petiole length and leaf area which mirrors the QTL effect observed in *C. hirsuta* HIFs. Additionally *SPL9Ox* also increased serrations in *A. thaliana* leaves which to certain extent can be considered homologous to leaflets, thus the QTL effect has been replicated in the *A. thaliana* transformants. Furthermore, I had hypothesised that the QTL effect was bought

about by accelerating the transition through development. This hypothesis was supported by the appearance of abaxial trichomes earlier in *SPL9Ox* transformants than in *SPL9Az* transformants. Abaxial trichomes can be used as a marker for vegetative phase change (Telfer et al., 1997) and therefore from this analysis I can prove that *SPL9Ox* brings about earlier transitions through developmental stages than *SPL9Az.* Thus, there is very strong evidence that *ChSPL9* is causal to the QTL detected on linkage group 4.

Transformation of alternative *C. hirsuta* QTL alleles into *A. thaliana* was able to elicit the QTL effect because the two species share a broadly similar genotype to phenotype model. However for complete formal proof, transgenic complementation of QTL effect should take place in a *C. hirsuta* background. A HIF homozygous for the Oxford QTL allele would be transformed with *SPL9Az* and a HIF homozygous for the Azores QTL allele would be transformed with *SPL9Ox*.

Transgenic complementation could also be used to pinpoint the causal region or nucleotide of a QTL allele. The QTL phenotype in alternative genotypes of HIFs or Near Isogenic lines could be complemented using chimeric genes. A potential complication with this kind of transgenic complementation arises because it is possible that the addition of an extra copy of a gene in the same pathway can affect the phenotype in a way that obscures interpretation of the observed allelic affects. An alternative or complementary approach is to use artificial micro RNAs (amiRNAs; reviewed in Ossowski et al., 2008). Transgenes can be engineered that do not alter the encoded protein but do prevent silencing by a specific amiRNA (Palatnik et al., 2003). Thus, it is possible to use an amiRNA to repress expression of the endogenous gene and contemporaneously introduce a variant copy of the gene that is not affected by the amiRNA, i.e. this would essentially functionally replace one allele with the other (Weigel et al., 2012).

Without identification of the causal polymorphism it is only possible to speculate about how the activity of *ChSPL9* has been modified to bring about allelic variation. At this stage of enquiry all that can be concluded is that *SPL9Ox* promotes the appearance of adult-like

morphology (more leaflets, larger leaflets, longer rachises and wider leaves) and therefore the QTL effect in the Oxford allele could be bought about by reduced sensitivity to miR156, increased expression via *cis* regulatory regions, or an improved ability to promote miR172 expression. The last of these is the most likely, since the target sequence of miR156 was unaltered in either of the *ChSPL9* sequences and a quantitative real time PCR assay revealed no significant allelic variation in *ChSPL9* transcript level. The QTL effect could be due to the interaction of ChSPL9 with DELLA proteins. Auxin-controlled DELLA directly binds to SPL proteins and inactivates mir172 in leaves and MADS box genes at the shoot apex to decelerate floral transition (Yu et al., 2012).

The work presented here complements a recent study that demonstrates mechanism by which *SPL9* operates to elicit an age dependant effect. Rubio-Somoza et al (2014) proved that SPL9 protein interacts with TCP to stop it interfering with CUC2 and CUC3 heterodimers. As SPL9 protein accumulates with age (because of decreasing miR156 levels) it interferes with TCP at a greater rate meaning that the CUC2-CUC3 heterodimer is more readily able to promote more complex morphology in later developing leaves (Rubio-Somoza et al., 2014). It is possible that the QTL effect observed with *ChSPL9* in the Oxford and Azores strains is based on this mechanism.

All the published research on the *SPL* gene family in the Brassicacea has relied on induced mutant alleles and transgenic lines to investigate their role and function in development. Here I present the discovery of a naturally occurring functional variant. Further study into how the QTL effect is elicited promises to reveal more about how *SPL* dependent pathways function and have been re-configured to create divergent morphology within the Brassicacea. There have, however, been natural variants of *SPL* genes in other taxa. For example natural mutations in the maize orthologs of *SPL9/15, ub2/3* have been discovered that affect agronomic traits (Chuck et al., 2014). In rice, orthologs of *SPL* genes have roles in shoot architecture and have been identified as causal to QTL mapped for shoot architecture (Miura et al., 2010; Jiao et al., 2010; Lu et al., 2013). And in tomatoes a spontaneous epigenetic mutation was found at the *Colourless*

non-ripening locus which affects an orthologue of the *SPL* gene family to produce colourless fruit that does not ripen normally (Manning et al., 2006). It will be interesting to compare the results of these studies to future work done on *ChSPL9*, especially if any age dependant effects of *SPL* mutations are discovered in these species.

6. 5. Conclusion

A major focus of this study was to elucidate the molecular basis of naturally variable leaf shape between strains of *C. hirsuta*. This goal has been achieved with the identification of *ChSPL9* as a factor in leaf shape diversification. Future work will focus on how the allelic forms of this gene adjusted leaf shape development.

Of key interest in biological research are the processes that underlie morphological variation, both between species and within species. Here, by using plant leaves as an example, I demonstrate that the causes of diversity at these two evolutionary scales can be divergent. Previous work has shown that interspecific differences between *A. thaliana* and *C. hirsuta* arise mostly from variation in local tissue growth and patterning (Vlad et al., 2014; Hay et al., 2006; Barkoulas et al., 2008). Now, by characterising a QTL for *C. hirsuta* leaf shape and complexity it is demonstrated that a different process, age-dependent leaf progression, is a major contributor to intraspecific variation in this trait. It will be interesting to study the ecological and physiological significance of this observation. It is possible that variation in the heteroblastic pathway allows rapid optimisation of leaf physiology in the diverse habitats and environments faced by weedy and invasive species like *C. hirsuta.*

References

P. Achard, A. Gusti, S. Cheminant, M. Alioua, S. Dhondt, F. Coppens, G.T.S. Beemster, & P. Genschik (2009). Gibberellin signalling controls cell proliferation rate in Arabidopsis. *Current Biology : CB*, *19*(14), 1188–93.

Aida, M., & Tasaka, M. (2006). Genetic control of shoot organ boundaries. *Current Opinion in Plant Biology*, *9*(1), 72–7.

Alonso-Blanco, C., Aarts, M. G. M., Bentsink, L., Keurentjes, J. J. B., Reymond, M., Vreugdenhil, D., & Koornneef, M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? *The Plant Cell*, *21*(7), 1877–96.

Alonso-Blanco, C., & Méndez-Vigo, B. (2014). Genetic architecture of naturally occurring quantitative traits in plants: an updated synthesis. *Current Opinion in Plant Biology*, *18*, 37–43.

Alonso-Blanco, C., Mendez-Vigo, B., & Koornneef, M. (2005). From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. *The International Journal of Developmental Biology*, *49*(5-6), 717–32.

Amasino, R. (2010). Seasonal and developmental timing of flowering. *The Plant Journal : For Cell and Molecular Biology*, *61*(6), 1001–13.

Andrés, F., & Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nature Reviews. Genetics*, *13*(9), 627–39.

Andriankaja, M., Dhondt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., … Inzé, D. (2012). Exit from proliferation during leaf development in Arabidopsis thaliana: a not-so-gradual process. *Developmental Cell*, *22*(1),

Aukerman, M. J., & Sakai, H. (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2 -Like Target Genes The Plant Cell. *The Plant Cell*, *15*, 2730–2741.

Avivi, Y., Lev-Yadun, S., & Morozova, N. (2000). Clausa, a tomato mutant with a wide range of phenotypic perturbations, displays a cell type-dependent expression of the homeobox gene LeT6/TKn2. *Plant Physiology*, *124*, 542.

Bar, M., & Ori, N. (2014). Leaf development and morphogenesis. *Development*, *141*(22), 4219–4230.

Bar, M., & Ori, N. (2015). Compound leaf development in model plant species. *Current Opinion in Plant Biology*, *23*, 61–69.

Barkoulas, M., Galinha, C., Grigg, S. P., & Tsiantis, M. (2007). From genes to shape: regulatory interactions in leaf development. *Current Opinion in Plant Biology*, *10*(6), 660–6.

Barkoulas, M., Hay, A., Kougioumoutzi, E., & Tsiantis, M. (2008). A developmental framework for dissected leaf formation in the Arabidopsis relative Cardamine hirsuta. *Nature Genetics*, *40*(9), 1136–41.

Bartels, A., Mock, H.-P., & Papenbrock, J. (2007). Differential expression of Arabidopsis sulfurtransferases under various growth conditions. *Plant Physiology and Biochemistry 45*(3-4), 178–87.

Bassel, G. W., Mullen, R. T., & Bewley, J. D. (2008). Procera is a putative DELLA mutant in tomato (Solanum lycopersicum): effects on the seed and vegetative plant. *Journal of Experimental Botany*, *59*(3), 585–93.

Benfey, P., & Mitchell-Olds, T. (2008). From genotype to phenotype: systems biology meets natural variation. *Science*, *320*(5875), 495–497.

Ben-Gera, H., Shwartz, I., Shao, M.-R., Shani, E., Estelle, M., & Ori, N. (2012a). ENTIRE and GOBLET promote leaflet development in tomato by modulating auxin response. *The Plant Journal : For Cell and Molecular Biology*, *70*(6), 903–15.

Ben-Gera, H., Shwartz, I., Shao, M.-R., Shani, E., Estelle, M., & Ori, N. (2012b). ENTIRE and GOBLET promote leaflet development in tomato by modulating auxin response. *The Plant Journal : For Cell and Molecular Biology*, *70*(6), 903–15.

Berardini, T., Bollman, K., Sun, H., & Poethig, R. (2001). Regulation of vegetative phase change in Arabidopsis thaliana by cyclophilin 40. *Science*, *291*, 1–4.

Berger Y., Harpaz-Saad S., Brand A., Melnik H., Sirding N., Alvarez J. P., Zinder M., Samach A., Eshed Y., Ori N. (2009). The NAC-domain transcription factor GOBLET specifies leaflet boundaries in compound tomato leaves. *Development*, *136*(5), 823–32.

Bharathan, G., Goliber, T., Moore, C., & Kessler, S. (2002). Homologies in leaf form inferred from KNOXI gene expression during development. *Science*, *296*, 1858.

Bilsborough G. D., Runions, A., Barkoulas, M., Jenkins, H. W., Hasson, A., Galinha, C., Laufs, P., Hay, A., Prusinkiewicz, P. & Tsiantis, M. (2011). Model for the regulation of Arabidopsis thaliana leaf margin development. *PNAS*, *108*(8).

Blein, T., Pautot, V., & Laufs, P. (2013). Combinations of Mutations Sufficient to Alter Arabidopsis Leaf Dissection. *Plants*, *2*(2), 230–247.

Blein, T., Pulido, A., & Vialette-Guiraud, A. (2008). A conserved molecular framework for compound leaf development. *Science*, (December), 1835–1839.

Bolduc, N., & Hake, S. (2009). The maize transcription factor KNOTTED1 directly regulates the gibberellin catabolism gene ga2ox1. *The Plant Cell*, *21*(6), 1647–58.

Bolduc, N., Yilmaz, A., Mejia-Guerra, M. K., Morohashi, K., O'Connor, D., Grotewold, E., & Hake, S. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes & Development*, *26*(15).

Borevitz, J.O., D. Liang, D. Plouff e, H.S. Chang, T. Zhu, D. Weigel, C.C. Berry, E. Winzeler, and J. Chory. (2003). Large-scale identification of single-feature polymorphisms in complex genomes. *Genome Research*, *13*(3), 513–23.

Brand, A., Shirding, N., Shleizer, S., & Ori, N. (2007). Meristem maintenance and compound-leaf patterning utilize common genetic mechanisms in tomato. *Planta*, *226*(4), 941–51.

Braybrook, S. a, & Peaucelle, A. (2013). Mechano-chemical aspects of organ formation in Arabidopsis thaliana: the relationship between auxin and pectin. *PloS One*, *8*(3), e57813.

Broman, K. W., Wu, H., Sen, S., & Churchill, G. a. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, *19*(7), 889–890.

Burko Y, Shleizer-Burko S, Yanai O, Shwartz I, Zelnik ID, Jacob-Hirsch J, Kela I, Eshed-Williams L, Ori N (2013). A role for APETALA1/fruitfull transcription factors in tomato leaf development. *The Plant Cell*, *25*(6), 2070–83.

Busch, B. L., Schmitz, G., Rossmann, S., Piron, F., Ding, J., Bendahmane, A., & Theres, K. (2011). Shoot branching and leaf dissection in tomato are regulated by homologous gene modules. *The Plant Cell*, *23*(10), 3595–609.

Canales, C., Barkoulas, M., Galinha, C., & Tsiantis, M. (2010). Weeds of change: Cardamine hirsuta as a new model system for studying dissected leaf development. *Journal of Plant Research*, *123*(1), 25–33.

Champagne, C. E. M., Goliber, T. E., Wojciechowski, M. F., Mei, R. W., Townsley, B. T., Wang, K., … Sinha, N. R. (2007). Compound leaf development and evolution in the legumes. *The Plant Cell*, *19*(11), 3369–78.

Chen, J., & Janssen, B. (1997). A gene fusion at a homeobox locus: alterations in leaf shape and implications for morphological evolution. *The Plant Cell Online*, *9*, 1289–1304.

Chen, J., Yu, J., Ge, L., & Wang, H. (2010). Control of dissected leaf morphology by a Cys (2) His (2) zinc finger transcription factor in the model legume Medicago truncatula. *PNAS* 107: 10754–10759

Cheng, X., Peng, J., Ma, J., Tang, Y., Chen, R., Mysore, K. S., & Wen, J. (2012). NO APICAL MERISTEM (MtNAM) regulates floral organ identity and lateral organ separation in Medicago truncatula. *The New Phytologist*, *195*(1), 71–84.

Chitwood D. H., Headland L. R., Filiault D. L., Kumar R., Jimenez-Gomez J. M., Schrager A. V., Park D. S., Peng J., Sinha N. R., Maloof J. N. (2012). Native environment modulates leaf size and response to simulated foliar shade across wild tomato species. *PloS One*, *7*(1), e29570.

Chitwood, D.H., Kumar, R., Headland, L.R., Ranjan, A., Covington, M.F., Ichihashi, Y., Fulop, D., Jiménez-Gómez, J.M., Peng, J., Maloof, J.N., Sinha, N.R. (2013). A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. *The Plant Cell*, *25*(7), 2465–81.

Chitwood, D. H., Ranjan, a., Kumar, R., Ichihashi, Y., Zumstein, K., Headland, L. R., Ostria-Gallardo, E., Aguilar-Martínez, J. A., Bush, S., Carriedo, L., Fulop, D., Martinez, C. C., Peng, J., Maloof, J. N., & Sinha, N. R. (2014). Resolving Distinct Genetic Regulators of Tomato Leaf Shape within a Heteroblastic and Ontogenetic Context. *The Plant Cell*, *26(9)*, 3616–3629.

Chitwood, D. H., & Sinha, N. R. (2013). A census of cells in time: quantitative genetics meets developmental biology. *Current Opinion in Plant Biology*, *16*(1), 92–9.

Chou, H., Chiu, H., Delaney, N., Segrè, D., & Marx, C. (2011). Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, *332*, 1190–1192.

Chuck, G., Cigan, a M., Saeteurn, K., & Hake, S. (2007). The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. *Nature Genetics*, *39*(4), 544–9

Chuck, G. S., Brown, P. J., Meeley, R., & Hake, S. (2014). Maize SBP-box transcription factors unbranched2 and unbranched3 affect yield traits by regulating the rate of lateral primordia initiation. PNAS 111(52), 18775–80.

Churchill, G. a, & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics*, *138*(3), 963–71.

Clark R. M., Schweikert G., Toomajian C., Ossowski S., Zeller G., Shinn P., Warthmann N., Hu T. T., Fu G., Hinds D. A., Chen H., Frazer K. A., Huson D. H., Scholkopf B., Nordborg M., Ratsch G., Ecker J. R., & Weigel D. (2007). Common Sequence Polymorphisms Shaping Genetic Diversity in Arabidopsis thaliana. *Science*, *317*(July), 338–343.

Cooper, M., van Eeuwijk, F. a, Hammer, G. L., Podlich, D. W., & Messina, C. (2009). Modeling QTL for complex traits: detection and context for plant breeding. *Current Opinion in Plant Biology*, *12*(2), 231–40.

David-Schwartz, R., Koenig, D., & Sinha, N. R. (2009). LYRATE is a key regulator of leaflet initiation and lamina outgrowth in tomato. *The Plant Cell*, *21*(10), 3093–104.

DeMason, D. A. (2005). Auxin-cytokinin and auxin-gibberellin interactions during morphogenesis of the compound leaves of pea (Pisum sativum). *Planta*, *222*(1), 151–66.

DeMason, D. A, & Chetty, V. (2014). Phenotypic characterization of the CRISPA (ARP gene) mutant of pea (Pisum sativum; Fabaceae): a reevaluation. *American Journal of Botany*, *101*(3), 408–27.

Demason, D. a, Chetty, V., Barkawi, L. S., Liu, X., & Cohen, J. D. (2013). Unifoliata-Afila interactions in pea leaf morphogenesis. *American Journal of Botany*, *100*(3), 478–95.

DeMason, D. a, & Chetty, V. J. (2011). Interactions between GA, auxin, and UNI expression controlling shoot ontogeny, leaf morphogenesis, and auxin response in Pisum sativum (Fabaceae): or how the uni-tac mutant is rescued. *American Journal of Botany*, *98*(5), 775–91.

DeMason, D. A., & Polowick, P. L. (2009). Patterns of DR5::GUS Expression in Organs of Pea (Pisum sativum). *International Journal of Plant Sciences*, *170*(1), 1–11.

Di Giacomo, E., Sestili, F., Iannelli, M. A., Testone, G., Mariotti, D., & Frugis, G. (2008). Characterization of KNOX genes in Medicago truncatula. *Plant Molecular Biology*, *67*(1-2), 135–50. 7

Dinneny, J. R., Weigel, D., Development, M. F. Y., Dinneny, J. R., Weigel, D., & Yanofsky, M. F. (2006). NUBBIN and JAGGED define stamen and carpel shape in Arabidopsis. *Development*, *133*(11), 2285–2285.

Disch, S., Anastasiou, E., Sharma, V. K., Laux, T., Fletcher, J. C., & Lenhard, M. (2006). The E3 ubiquitin ligase BIG BROTHER controls arabidopsis organ size in a dosage-dependent manner. *Current Biology : CB*, *16*(3), 272–9.

Donnelly, P. M., Bonetta, D., Tsukaya, H., Dengler, R. E., & Dengler, N. G. (1999). Cell cycling and cell enlargement in developing leaves of Arabidopsis. *Developmental Biology*, *215*(2), 407–19.

Efroni, I., Blum, E., Goldshmidt, A., & Eshed, Y. (2008). A protracted and dynamic maturation schedule underlies Arabidopsis leaf development. *The Plant Cell*, *20*(9), 2293–306.

Efroni, I., Han, S.-K., Kim, H. J., Wu, M.-F., Steiner, E., Birnbaum, K. D., Efroni I., Han S.-K., Kim, Hye J., Wu M.-F., Steiner E., Birnbaum, Kenneth D., Hong Jong C., Eshed Y., & Wagner D. (2013). Regulation of leaf maturation by chromatin-mediated modulation of cytokinin responses. *Developmental Cell*, *24*(4), 438–45.

El-Lithy, M., Clerkx, E., & Ruys, G. (2004). Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. *Plant Physiology*, *135*(May), 444–458

Emery, J.F., Floyd, S.K., Alvarez, J., Eshed, Y., Hawker, N.P., Izhaki, A., Baum, S.F. and Bowman, J.L. (2003). Radial Patterning of Arabidopsis Shoots by Class III HD-ZIP and KANADI Genes. *Current Biology*, *13*(20), 1768–1774.

Feng X., Wilson Y., Bowers J., Kennaway R., Bangham A., Hannah A., Coen E., & Hudson A. (2009). Evolution of allometry in antirrhinum. *The Plant Cell*, *21*(10), 2999–3007.

Feraru E., Feraru M. I., Kleine-Vehn J., Martinière A., Mouille G., Vanneste S., Vernhettes S., Runions J., & Friml J. (2011). PIN polarity maintenance by the cell wall in Arabidopsis. *Current Biology : CB*, *21*(4), 338–43.

Fleishon, S., Shani, E., Ori, N., & Weiss, D. (2011). Negative reciprocal interactions between gibberellin and cytokinin in tomato. *The New Phytologist*, *190*(3), 609–17.

Floyd, S. K., & Bowman, J. L. (2010). Gene expression patterns in seed plant shoot meristems and leaves: homoplasy or homology? *Journal of Plant Research*, *123*(1), 43–55.

Francki, M., Gondo, T., & Sato, S. (2007). Quantitative trait locus analysis of multiple agronomic traits in the model legume Lotus japonicus. *Genome*, *637*, 627–637.

Frary, A, Doganlar, S., Daunay, M. C., & Tanksley, S. D. (2003). QTL analysis of morphological traits in eggplant and implications for conservation of gene function during evolution of solanaceous species. *Theoretical and Applied Genetics. 107*(2), 359–70.

Frary, A., Nesbitt, T., & Frary, A. (2000). fw2. 2: a quantitative trait locus key to the evolution of tomato fruit size. *Science*, *289*, 85.

Ge, L., Peng, J., Berbel, A., Madueño, F., & Chen, R. (2014). Regulation of compound leaf development by PHANTASTICA in Medicago truncatula. *Plant Physiology*, *164*(1), 216–28.

Gourlay, C., Hofer, J., & Ellis, T. (2000). Pea compound leaf architecture is regulated by interactions among the genes UNIFOLIATA, COCHLEATA, AFILA, and TENDRIL-LESS. *The Plant Cell Online*, *12*, 1279–1294.

Hagemann, W., & Gleissberg, S. (1996). Systematics and Evolution Organogenetic capacity of leaves : the significance of marginal blastozones in angiosperms. *Plant Systematics and Evolution*, *199*, 121–152.

Hareven, D., Gutfinger, T., Parnis, A., Eshed, Y., & Lifschitz, E. (1996). The Making of a Compound Leaf: Genetic Manipulation of Leaf Architecture in Tomato. *Cell*, *84*(5), 735–744.

Hase, Y., Trung, K. H., Matsunaga, T., & Tanaka, A. (2006). A mutation in the uvi4 gene promotes progression of endo-reduplication and confers increased tolerance towards ultraviolet B light. *The Plant Journal : For Cell and Molecular Biology*, *46*(2), 317–26.

Hasson A., Plessis A., Blein T., Adroher B., Grigg S., Tsiantis M., Boudaoud A., Damerval C., & Laufs P. (2011). Evolution and diverse roles of the CUP-SHAPED COTYLEDON genes in Arabidopsis leaf development. *The Plant Cell*, *23*(1), 54–68.

Hay, A., Barkoulas, M., & Tsiantis, M. (2006). ASYMMETRIC LEAVES1 and auxin activities converge to repress BREVIPEDICELLUS expression and promote leaf development in Arabidopsis. *Development*, *133*(20), 3955–61.

Hay, A., Kaur, H., Phillips, A., & Hedden, P. (2002). The gibberellin pathway mediates KNOTTED1-type homeobox function in plants with different body plans. *Current Biology*, *12*(02), 1557–1565.

Hay, A. S., Pieper, B., Cooke, E., Mandáková, T., Cartolano, M., Tattersall, A. D., Ioio R. D., McGowan S. J., Barkoulas, M., Galinha, C., Rast, M. I., Hofhuis, H., Then, C., Plieske, J., Ganal, M., Mott, R., Martinez-Garcia, J. F., Carine, M. A., Scotland, R. W., Gan, X., Filatov, D. A.,. Lysak, M. A., and Tsiantis, M. (2014). Cardamine hirsuta: a versatile genetic system for comparative studies. *The Plant Journal : For Cell and Molecular Biology*, *78*(1), 1–15.

Hay, A., & Tsiantis, M. (2006). The genetic basis for differences in leaf form between Arabidopsis thaliana and its wild relative Cardamine hirsuta. *Nature Genetics*, *38*(8), 942–7.

Hay, A., & Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development 137*(19), 3153–65.

Heisler, M. G., Ohno, C., Das, P., Sieber, P., Reddy, G. V, Long, J. a, & Meyerowitz, E. M. (2005). Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the Arabidopsis inflorescence meristem. *Current Biology : CB*, *15*(21), 1899–911.

Holst, K., Schmülling, T., & Werner, T. (2011). Enhanced cytokinin degradation in leaf primordia of transgenic Arabidopsis plants reduces leaf size and shoot organ primordia formation. *Journal of Plant Physiology*, *168*(12), 1328–34.

Holtan, H. E. E., & Hake, S. (2003). Quantitative trait locus analysis of leaf dissection in tomato using Lycopersicon pennellii segmental introgression lines. *Genetics*, *165*(3), 1541–50.

Horiguchi, G., Van Lijsebettens, M., Candela, H., Micol, J. L., & Tsukaya, H. (2012). Ribosomes and translation in plant developmental control. *Plant Science : An International Journal of Experimental Plant Biology*, *191-192*, 24–34.

Horstman, A., Willemsen, V., Boutilier, K., & Heidstra, R. (2014). AINTEGUMENTA-LIKE proteins: hubs in a plethora of networks. *Trends in Plant Science*, *19*(3), 146–57.

Hospital, F. (2009). Challenges for effective marker-assisted selection in plants. *Genetica*, *136*(2), 303–10.

Huang, X., Paulo, M., & Boer, M. (2011). Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. *PNAS*, *108*(11): 4488-4493

Huijser, P., & Schmid, M. (2011). The control of developmental phase transitions in plants. *Development*, *138*(19), 4117–29.

Hunter, C., Willmann, M. R., Wu, G., Yoshikawa, M., de la Luz Gutiérrez-Nava, M., & Poethig, S. R. (2006). Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in Arabidopsis. *Development*, *133*(15), 2973–81.

Ichihashi, Y., Aguilar-Martínez, J.A, Farhi, M., Chitwood, D.H., Kumar, R., Millon, L.V., Peng, J., Maloof, J.N., Sinha, N.R. (2014). Evolutionary developmental transcriptomics reveals a gene network module regulating interspecific diversity in plant leaf shape. *PNAS 111*(25), E2616–21.

Imaizumi, T. (2010). Arabidopsis circadian clock and photoperiodism: time to think about location. *Current Opinion in Plant Biology*, *13*(1), 1–13.

Ito, S., Song, Y. H., Josephson-Day, A. R., Miller, R. J., Breton, G., Olmstead, R. G., & Imaizumi, T. (2012). FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. *PNAS*, *109*(9), 3582–7.

Jackson, S. (2009). Plant responses to photoperiod. *New Phytologist*, (2008), 517–531.

Jansen, R., & Stam, P. (1994). High resolution of quantitative traits into multiple loci via interval mapping. *Genetics*, *1455*, 1447–1455.

Jasinski, S., Kaur, H., Tattersall, A., & Tsiantis, M. (2007). Negative regulation of KNOX expression in tomato leaves. *Planta*, *226*(5), 1255–63.

Jasinski S., Piazza P., Craft J., Hay A., Woolley L., Rieu I., Phillips A., Hedden P. T., & Tsiantis M. (2005). KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. *Current Biology : CB*, *15*(17), 1560–5.

Jasinski S., Tattersall A., Piazza P., Hay A., Martinez-Garcia J. F., Schmitz G., Theres K., McCormick S., & Tsiantis M (2008). PROCERA encodes a DELLA protein that mediates control of dissected leaf form in tomato. *The Plant Journal*, *56*(4), 603–12.

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., Dong,G., Zeng, D., Lu, Z., Zhu, X., Qian, Q., & Li, J. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nature Genetics*, *42*(6), 541–4.

Johnston J. S., Pepper A. E., Hall A. E., Chen Z. J.,Hodnett G., Drabek J., Lopez R., Price H. J. (2005). Evolution of genome size in Brassicaceae. *Annals of Botany*, *95*(1), 229–35.

Jones, C. (1999). An essay on juvenility, phase change, and heteroblasty in seed plants. *International Journal of Plant Sciences*, *160*. S105-S111

Juenger, T., Pérez-Pérez, J. M., Bernal, S., & Micol, J. L. (2005). Quantitative trait loci mapping of floral and leaf morphology traits in Arabidopsis thaliana: evidence for modular genetic architecture. *Evolution & Development*, *7*(3), 259–71.

Kawade, K., Horiguchi, G., Usami, T., Hirai, M. Y., & Tsukaya, H. (2013). ANGUSTIFOLIA3 signaling coordinates proliferation between clonally distinct cells in leaves. *Current Biology : CB*, *23*(9), 788–92.

Kawamura, E., Horiguchi, G., & Tsukaya, H. (2010). Mechanisms of leaf tooth formation in Arabidopsis. *The Plant Journal : For Cell and Molecular Biology*, *62*(3), 429–41.

Keurentjes, J. J. B., Fu, J., Terpstra, I. R., Garcia, J. M., van den Ackerveken, G., Snoek, L.B., Peeters, A. J. M., Vreugdenhil, D., Koornneef, M., & Jansen, R. C. (2007). Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *PNAS*, *104*(5), 1708–13.

Kierzkowski D., Nakayama N., Routier-Kierzkowska A. L., Weber A., Bayer E., Schorderet M., Reinhardt D., Kuhlemeier C., Smith R. S. (2011). Elastic Domains Regulate Growth and Organogenesis in the Plant Shoot Apical Meristem. *Science*, *1015*(2006), 1096 –1099.

Kim, M. (2003). Reduced leaf complexity in tomato wiry mutants suggests a role for PHAN and KNOX genes in generating compound leaves. *Development*, *130*(18), 4405–4415.

Kimura, S., Koenig, D., Kang, J., Yoong, F. Y., & Sinha, N. (2008). Natural variation in leaf morphology results from mutation of a novel KNOX gene. *Current Biology : CB*, *18*(9), 672–7.

Kobayashi, Y., & Weigel, D. (2007). Move on up, it's time for change - mobile signals controlling photoperiod-dependent flowering. *Genes & Development*, *21*, 2371–2384.

Koenig, D., Bayer, E., Kang, J., Kuhlemeier, C., & Sinha, N. (2009). Auxin patterns Solanum lycopersicum leaf morphogenesis. *Development*, *136*(17), 2997–3006.

Koornneef, M., Alonso-Blanco, C., Peeters, A. J. M., & Soppe, W. (1998). Genetic Control of Flowering Time in Arabidopsis. *Annual Review of Plant Physiology and Plant Molecular Biology*, *49*, 345–370.

Koornneef, M., Alonso-Blanco, C., & Vreugdenhil, D. (2004). Naturally occurring genetic variation in Arabidopsis thaliana. *Annual Review of Plant Biology*, *55*, 141–72.

Koornneef, M., Vries, H., & Hanhart, C. (1994). The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg erecta wild-type. *The Plant Journal*, *6*, 911–919.

Kosambi, D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics*. *12*(1), 172-175

Kougioumoutzi E., Cartolano M., Canales C., Dupré M., Bramsiepe J., Vlad D., Rast M., Dello Ioio R., Tattersall A., Schnittger A., Hay A., Tsiantis M. (2013). SIMPLE LEAF3 encodes a ribosome-associated protein required for leaflet development in Cardamine hirsuta. *The Plant Journal : For Cell and Molecular Biology*, *73*(4), 533–45.

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., Durrant C., & Mott, R. (2009). A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in Arabidopsis thaliana. *PLoS Genetics*, *5*(7), e1000551.

Koyama, T., Furutani, M., Tasaka, M., & Ohme-Takagi, M. (2007). TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in Arabidopsis. *The Plant Cell*, *19*(2), 473–84.

Kroymann, J., & Mitchell-Olds, T. (2005). Epistasis and balanced polymorphism influencing complex trait variation. *Nature*, *435*(7038), 95–8.

Lan, T. H., & Paterson, a. H. (2001). Comparative mapping of QTLs determining the plant size of Brassica oleracea. *TAG Theoretical and Applied Genetics*, *103*(2-3), 383–397.

Lander, E. S., & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, *121*(1), 185–99.

Langlade, N. B., Feng, X., Dransfield, T., Copsey, L., Hanna, A. I., Thébaud, C., Bangham, A. Hudson, A., & Coen, E. (2005). Evolution through genetically controlled allometry space. *PNAS*, *102*(29), 10221–6.

Lauter, N., & Kampani, A. (2005). microRNA172 down-regulates glossy15 to promote vegetative phase change in maize. *PNAS 102*(26), 9412-9417.

Lee, B., Kapoor, A., Zhu, J., & Zhu, J. (2006). STABILIZED1, a stress-upregulated nuclear protein, is required for pre-mRNA splicing, mRNA turnover, and stress tolerance in Arabidopsis. *The Plant Cell Online*, *18*(July), 1736–1749.

Lee, I., & Michaels, S. (1994). The late-flowering phenotype of FRIGIDA and mutations in LUMINIDEPENDENS is suppressed in the Landsberg erecta strain of Arabidopsis. *The Plant Journal*, *6*, 903–909.

Li, J., & Ji, L. (2005). Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, *95*(3), 221–7.

Livingstone, C. D., & Barton, G. J. (1993). Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Bioinformatics*, *9*(6), 745–756.

Lodha, M., Marco, C. F., & Timmermans, M. C. P. (2013). The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes & Development*, *27*(6), 596–601.

Lolle, S. J., Victor, J. L., Young, J. M., & Pruitt, R. E. (2005). Genome-wide non-mendelian inheritance of extra-genomic information in Arabidopsis. *Nature*, *434*(7032), 505–9.

Lou, P., Zhao, J., Kim, S. S., Shen, S., Del Carpio D. P., Song, X., Jin, M., Vreugdenhil, D., Wang, X., Koornneef, M., & Bonnema, G. (2007). Quantitative trait loci for flowering time

and morphological traits in multiple populations of Brassica rapa. *Journal of Experimental Botany*, *58*(14), 4005–16.

Lu, Z., Yu, H., Xiong, G., Wang, J., Jiao, Y., Liu, G., Jing, Y., Meng, X., Hu, X., Qian, Q., Fu, X., Wang, Y., & Li, J. (2013). Genome-wide binding analysis of the transcription activator ideal plant architecture1 reveals a complex network regulating rice plant architecture. *The Plant Cell*, *25*(10), 3743–59.

Luo, M., Yu, C. W., Chen, F. F., Zhao, L., Tian, G., Liu, X., Cui, Y., Yang, J. Y., & Wu, K. (2012). Histone deacetylase HDA6 is functionally associated with AS1 in repression of KNOX genes in arabidopsis. *PLoS Genetics*, *8*(12), e1003114.

Malinowski, R., Kasprzewska, A., & Fleming, A. J. (2011). Targeted manipulation of leaf form via local growth repression. *The Plant Journal : For Cell and Molecular Biology*, *66*(6), 941–52.

Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., Barton, M. K., & Bartel, D. P. (2004). MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *The EMBO Journal*, *23*(16), 3356–64.

Malosetti, M., Ribaut, J. M., Vargas, M., Crossa, J., & van Eeuwijk, F. A. (2007). A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (Zea mays L.). *Euphytica*, *161*(1-2), 241–257

Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., Giovannoni, J. J., & Seymour, G. B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics*, *38*(8), 948–52.

Martin, N. H., Sapir, Y., & Arnold, M. L. (2008). The genetic architecture of reproductive isolation in Louisiana irises: pollination syndromes and pollinator preferences. *Evolution; International Journal of Organic Evolution*, *62*(4), 740–52.

Martín-Trillo, M., & Cubas, P. (2010). TCP genes: a family snapshot ten years later. *Trends in Plant Science*, *15*(1), 31–9.

Milani, P., Gholamirad, M., Traas, J., Arnéodo, A., Boudaoud, A., Argoul, F., & Hamant, O. (2011). In vivo analysis of local wall stiffness at the shoot apical meristem in Arabidopsis using atomic force microscopy. *The Plant Journal : For Cell and Molecular Biology*, *67*(6), 1116–23.

Mishra, R. K., Chaudhary, S., Kumar, A., & Kumar, S. (2009). Effects of MULTIFOLIATE-PINNA, AFILA, TENDRIL-LESS and UNIFOLIATA genes on leaf blade architecture in Pisum sativum. *Planta*, *230*(1), 177–90.

Mitchell-Olds, T., Willis, J. H., & Goldstein, D. B. (2007). Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews. Genetics*, *8*(11), 845–56.

Miura, K., Ikeda, M., Matsubara, A., Song, X.-J., Ito, M., Asano, K Matsuoka M., Kitano, H., & Ashikari, M. (2010). OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nature Genetics*, *42*(6), 545–9.

Mizukami, Y., & Fischer, R. L. (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proceedings of the National Academy of Sciences*, *97*(2), 942–947.

Nath, U., Crawford, B., Carpenter, R., & Coen, E. (2003). Genetic control of surface curvature. *Science*, *299*, 1404-1407.

Naz, A. A., Raman, S., Martinez, C. C., Sinha, N. R., Schmitz, G., & Theres, K. (2013). Trifoliate encodes an MYB transcription factor that modulates leaf and shoot architecture in tomato. *PNAS*, *110*(6), 2401–6.

Neff, M., Turk, E., & Kalishman, M. (2002). Web-based primer design for single nucleotide polymorphism analysis. *TRENDS in Genetics*, *18*(12), 2000–2002

Nicotra, A., Leigh, A., & Boyce, C. (2011). The evolution and functional significance of leaf shape in the angiosperms. *Functional Plant Biology 38*(7), 535–552.

Nikovics, K., Blein, T., Peaucelle, A., Ishida, T., Morin, H., Aida, M., & Laufs, P. (2006). The balance between the MIR164A and CUC2 genes controls leaf margin serration in Arabidopsis. *The Plant Cell*, *18*(11), 2929–45.

Nogueira, F., & Madi, S. (2007). Two small regulatory RNAs establish opposing fates of a developmental axis. *Genes & Devlopment*, *2*, 750–755.

Ohno, C. K., Reddy, G. V., Heisler, M. G. B., & Meyerowitz, E. M. (2004). The Arabidopsis JAGGED gene encodes a zinc finger protein that promotes leaf tissue development. *Development*, *131*(5), 1111–22.

Ori, N., Cohen, A. R., Etzioni, A., Brand, A., Yanai, O., Shleizer, S., Menda, N., Amsellem, Z., Efroni, I., Pekker, I., Alvarez, J. P., Blum, E., Zamir, D., Eshed, Y. (2007). Regulation of LANCEOLATE by miR319 is required for compound-leaf development in tomato. *Nature Genetics*, *39*(6), 787–91

Orr, H. A. (1998). Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics*, *149*(4), 2099-2104.

Ossowski, S., Schwab, R., & Weigel, D. (2008). Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal : For Cell and Molecular Biology*, *53*(4), 674–90.

Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., & Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature*, *425*(6955), 257–63.

Payne, R.W., Murray, D.A., Harding, S.A., Baird, D.B. & Soutar, D.M. (2009). GenStat for Windows (12th Edition) Introduction. VSN International, Hemel Hempstead.

Peng, J., & Chen, R. (2011). Auxin efflux transporter MtPIN10 regulates compound leaf and flower development in Medicago truncatula. *Plant Signaling & Behavior*, *6*(10), 1537–44.

Peragine, A., & Yoshikawa, M. (2004). SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes & Development*, *3*, 2368–2379.

Pérez-Pérez, J. M., Esteve-Bruna, D., & Micol, J. L. (2010). QTL analysis of leaf architecture. *Journal of Plant Research*, *123*(1), 15–23.

Pérez-Pérez, J. M., Serrano-Cartagena, J., & Micol, J. L. (2002). Genetic analysis of natural variations in the architecture of Arabidopsis thaliana vegetative leaves. *Genetics*, *162*(2), 893–915.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*(9), 45e–45.

Piazza, P., Bailey, C. D., Cartolano, M., Krieger, J., Cao, J., Ossowski, S., Schneeberger, K., He, F., de Meaux, J., Hall, N., Macleod, N., Filatov, D., Hay, A., & Tsiantis, M. (2010). Arabidopsis thaliana leaf form evolved via loss of KNOX expression in leaves in association with a selective sweep. *Current Biology*, *20*(24), 2223–8.

Poethig, R. (1990). Phase Change and the Regulation in Plants Morphogenesis Shoot. *Science*, *250*(4983), 923–930.

Poethig, R. (1997). Leaf morphogenesis in flowering plants. *The Plant Cell*, *9*, 1077–1087.

Poethig, R. S. (2013). *Vegetative phase change and shoot maturation in plants. Current topics in developmental biology* (1st ed., Vol. 105, pp. 125–52). Elsevier Inc.

Price, A. H. (2006). Believe it or not, QTLs are accurate! *Trends in Plant Science*, *11*(5), 213–6.

Protas, M., Conrad, M., Gross, J., Tabin, C., & Borowsky, R. (2007). Regressive Evolution in the Mexican Cave Tetra, Astyanax mexicanus. *Current Biology*, *17*(5), 452–454.

Reddy, G. V., Heisler, M. G., Ehrhardt, D. W., & Meyerowitz, E. M. (2004). Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana. *Development*, *131*(17), 4225–37.

Reinbothe, S., Quigley, F., Springer, A., Schemenewitz, A., & Reinbothe, C. (2004). The outer plastid envelope protein Oep16: role as precursor translocase in import of protochlorophyllide oxidoreductase A. *PNAS*, *101*(7), 2203–8

Rieseberg, L. H., Archer, M. A., & Wayne, R. K. (1999). Short Review Transgressive segregation , adaptation and speciation. *Heredity*, *83*, 363.

Robinson, S., Burian, A., Couturier, E., Landrein, B., Louveaux, M., Neumann, E. D., Peaucelle, A., Weber, A., & Nakayama, N. (2013). Mechanical control of morphogenesis at the shoot apex. *Journal of Experimental Botany*, *64*(15), 4729–44.

Rubio-Somoza, I., Zhou, C.-M., Confraria, A., Martinho, C., von Born, P., Baena-Gonzalez, E., Wang, J. W. & Weigel, D. (2014). Temporal control of leaf complexity by miRNA-regulated licensing of protein complexes. *Current Biology : CB*, *24*(22), 2714–9

Sakamoto, T., & Kamiya, N. (2001). KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. *Genes & Development*, *1*, 581–590.

Salomé, P. a, Bomblies, K., Laitinen, R. a E., Yant, L., Mott, R., & Weigel, D. (2011). Genetic architecture of flowering-time variation in Arabidopsis thaliana. *Genetics*, *188*(2), 421–33.

Sambrook, J. and Russell, D.W. (2001) Molecular cloning. A laboratory manual. New York, Cold Spring Harbor Laboratory Press.

Sarojam, R., Sappl, P. G., Goldshmidt, A., Efroni, I., Floyd, S. K., Eshed, Y., & Bowman, J. L. (2010). Differentiating Arabidopsis shoots from leaves by combined YABBY activities. *The Plant Cell*, *22*(7), 2113–30.

Savaldi-Goldstein, S., Peto, C., & Chory, J. (2007). The epidermis both drives and restricts plant shoot growth. *Nature*, *446*(7132), 199–202.

Scarpella, E., & Marcos, D. (2006). Control of leaf vascular patterning by polar auxin transport. *Genes & Development*, 1015–1027.

Schiessl, K., Muiño, J. M., & Sablowski, R. (2014). Arabidopsis JAGGED links floral organ patterning to tissue growth by repressing Kip-related cell cycle inhibitors. *PNAS*, *111*(7), 2830–5.

Schranz, M. E., Lysak, M. A., & Mitchell-olds, T. (2006). The ABC ' s of comparative genomics in the Brassicaceae : building blocks of crucifer genomes. *Trends in Plant Science*, *11*(11).

Schwarz, S., Grande, A. V, Bujdoso, N., Saedler, H., & Huijser, P. (2008). The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. *Plant Molecular Biology*, *67*(1-2), 183–95.

Shalit, A., Rozman, A., Goldshmidt, A., Alvarez, J. P., Bowman, J. L., Eshed, Y., & Lifschitz, E. (2009). The flowering hormone florigen functions as a general systemic regulator of growth and termination. *PNAS*, *106*(20), 8392–7.

Shani, E., Ben-Gera, H., Shleizer-Burko, S., Burko, Y., Weiss, D., & Ori, N. (2010). Cytokinin regulates compound leaf development in tomato. *The Plant Cell*, *22*(10), 3206–17.

Shani, E., Burko, Y., Ben-Yaakov, L., Berger, Y., Amsellem, Z., Goldshmidt, A., Sharon, E., & Ori, N. (2009). Stage-specific regulation of Solanum lycopersicum leaf maturation by class 1 KNOTTED1-LIKE HOMEOBOX proteins. *The Plant Cell*, *21*(10), 3078–92.

Shleizer-Burko, S., Burko, Y., Ben-Herzel, O., & Ori, N. (2011). Dynamic growth program regulated by LANCEOLATE enables flexible leaf patterning. *Development 138*(4), 695–704.

Shpak, E., McAbee, J., Pillitteri, L., & Torii, K. (2005). Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science*, *309*(July), 290–293.

Sicard, A., Thamm, A., Marona, C., Lee, Y. W., Wahl, V., Stinchcombe, J. R., Wright, S., Kappel, C., & Lenhard, M. (2014). Repeated evolutionary changes of leaf morphology caused by mutations to a homeobox gene. *Current Biology : CB*, *24*(16), 1880–6.

Smith, H. M. S., Boschke, I., & Hake, S. (2002). Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9579–84.

Takahashi, Y., Ebisu, Y., Kinoshita, T., Doi, M., Okuma, E., Murata, Y., & Shimazaki, K.-I. (2013). bHLH transcription factors that facilitate K^+ uptake during stomatal opening are repressed by abscisic acid through phosphorylation. *Science Signaling*, *6*(280), ra48.

Tattersall, A., & Turner, L. (2005). The mutant crispa reveals multiple roles for PHANTASTICA in pea compound leaf development. *The Plant Cell*, *17*(April), 1046–1060.

Telfer, A., Bollman, K., & Poethig, R. (1997). Phase change and the regulation of trichome distribution in Arabidopsis thaliana. *Development*, *654*, 645–654.

Tisné, S., Reymond, M., Vile, D., Fabre, J., Dauzat, M., Koornneef, M., & Granier, C. (2008). Combined genetic and modeling approaches reveal that epidermal cell area and number in leaves are controlled by leaf and plant developmental processes in Arabidopsis. *Plant Physiology*, *148*(2),

Tonsor, S. J., Alonso-Blanco, C., & Koornneef, M. (2005). Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in Arabidopsis thaliana. *Plant, Cell and Environment*, *28*(1), 2–20.

Torii, K., Mitsukawa, N., & Oosumi, T. (1996). The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell*, *8*(April), 735–746.

Torii, K. U. (2004). Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathways. *International Review of Cytology*, *234*, 1–46.

Tsukaya, H. (2014). Comparative leaf development in angiosperms. *Current Opinion in Plant Biology*, *17*, 103–9.

Van Eeuwijk, F. a, Bink, M. C. a M., Chenu, K., & Chapman, S. C. (2010). Detection and use of QTL for complex traits in multiple environments. *Current Opinion in Plant Biology*, *13*(2), 193–205.

Van Ooijen, J.W. (2006) Joinmap® 4, software for the calculation of genetic linkage maps in experimental populations. Wageningen, Netherlands.: Kyazma B.V.

Vercruyssen, L., Verkest, A., Gonzalez, N., Heyndrickx, K. S., Eeckhout, D., Han, S.-K., Jégu, T., Archacki, R., Van Leene, J., Andriankaja, M., De Bodt, S., Abeel, T., Coppens, F., Dhondt, S., De Milde, L., Vermeersch, M., Maleux, K., Gevaert, K., Jerzmanowski, A., Benhamed, M., Wagner, D., Vandepoele, K., De Jaeger, G., & Inzé, D. (2014). ANGUSTIFOLIA3 binds to SWI/SNF chromatin remodeling complexes to regulate transcription during Arabidopsis leaf development. *The Plant Cell*, *26*(1), 210–29.

Vi, S. L., Trost, G., Lange, P., Czesnick, H., Rao, N., Lieber, D., Laux, T., Gray, W. M., Manley, J. L., Groth, D., Kappel, C., & Lenhard, M. (2013). Target specificity among canonical nuclear poly(A) polymerases in plants modulates organ growth and pathogen response. *PNAS*, *110*(34), 13994–9.

Vlad, D., Kierzkowski, D., Rast, M. I., Vuolo, F., Dello Ioio, R., Galinha, C., Gan, X., Hajheidari, M., Hay, A., Smith, R. S., Huijser, P., Bailey, C. D., & Miltos Tsiantis, M. (2014). Leaf shape evolution through duplication, regulatory diversification, and loss of a homeobox gene. *Science 343*(6172), 780–3.

Wang, H., Chen, J., Wen, J., Tadege, M., Li, G., Liu, Y., Mysore, K. S., Ratet, P., & Chen, R. (2008). Control of compound leaf development by FLORICAULA/LEAFY ortholog SINGLE LEAFLET1 in Medicago truncatula. *Plant Physiology*, *146*(4), 1759–72.

Wang, Y., Carrie, C., Giraud, E., Elhafez, D., Narsai, R., Duncan, O., Whelan, J., & Murcha, M. W. (2012). Dual location of the mitochondrial preprotein transporters B14.7 and Tim23-2 in complex I and the TIM17:23 complex in Arabidopsis links mitochondrial activity and biogenesis. *The Plant Cell*, *24*(6), 2675–95.

Weigel, D. (2012). Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiology*, *158*(1), 2–22.

Werner, T., & Motyka, V. (2001). Regulation of plant growth by cytokinin. *PNAS*, *98*(18), 10487–10492.

Wilkie, J. D., Sedgley, M., & Olesen, T. (2008). Regulation of floral initiation in horticultural trees. *Journal of Experimental Botany*, *59*(12), 3215–28.

Wu, G., Park, M. Y., Conway, S. R., Wang, J.-W., Weigel, D., & Poethig, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell*, *138*(4), 750–9.

Wu, G., & Poethig, R. (2006). Temporal regulation of shoot development in Arabidopsis thaliana by miR156 and its target SPL3. *Development*, *133*(18), 3539–3547.

Yanai, O., Shani, E., Russ, D., & Ori, N. (2011). Gibberellin partly mediates LANCEOLATE activity in tomato. *The Plant Journal*, *68*(4), 571–82.

Yi, N., & Shriner, D. (2008). Advances in Bayesian multiple quantitative trait loci mapping in experimental crosses. *Heredity*, *100*(3), 240–52.

Yifhar, T., Pekker, I., Peled, D., Friedlander, G., Pistunov, A., Sabban, M., Wachsman, G., Alvarez, J. P., Amsellem, Z., & Eshed, Y. (2012). Failure of the tomato trans-acting short interfering RNA program to regulate AUXIN RESPONSE FACTOR3 and ARF4 underlies the wiry leaf syndrome. *The Plant Cell*, *24*(9), 3575–89.

Yu, S., Galvão, V. C., Zhang, Y.-C., Horrer, D., Zhang, T.-Q., Hao, Y.-H., Feng, Y. Q., Wang, S., Schmid, M., & Wang, J.-W. (2012). Gibberellin regulates the Arabidopsis floral transition through miR156-targeted SQUAMOSA promoter binding-like transcription factors. *The Plant Cell*, *24*(8), 3320–32.

Zhou, C., Han, L., Fu, C., Wen, J., Cheng, X., Nakashima, J., Ma, J., Tang, Y., Tan, Y., Tadege, M., Mysore, K. S., Xia, G., & Wang, Z.-Y. (2013). The trans-acting short interfering RNA3 pathway and no apical meristem antagonistically regulate leaf margin development and lateral organ separation, as revealed by analysis of an argonaute7/lobed leaflet1 mutant in Medicago truncatula. *The Plant Cell*, *25*(12), 4845–62.

Zhou, C., Han, L., Hou, C., Metelli, A., Qi, L., Tadege, M., Mysore, K. S., & Wang, Z.-Y. (2011). Developmental analysis of a Medicago truncatula smooth leaf margin1 mutant reveals context-dependent effects on compound leaf development. *The Plant Cell*, *23*(6), 2106–24.

Zotz, G., Wilhelm, K., & Becker, A. (2011). Heteroblasty - A Review. *The Botanical Review*, *77*(2), 109–151.