Natural Variation in *Arabidopsis thaliana* Growth in Response to Ambient Temperatures

Asif Khan Khattak

M.Sc. with Distinction, Roskilde University, Denmark

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

School of Environmental Sciences,

University of East Anglia, UK

March 2014

Abstract

Temperature is one of the most important abiotic environmental regulators of plant growth and development. The temperature-dependent elongation of the *Arabidopsis thaliana* hypocotyls (seedling stems) is a well-characterised environmental response. The aim of this study was to identify allelic variants underlying Quantitative Trait Loci (QTL) responsible for the natural genetic variation of hypocotyl length in response to ambient temperatures.

The *Arabidopsis thaliana* accessions were phenotyped for hypocotyl length at 12°C, 17°C, 22°C and 27°C ambient temperature environments and substantial genetic variation was established. This facilitated a forward genetic approach by performing a QTL analysis to identify the genetic basis of thermal sensitivity. *Firstly*, fine-mapped QTL were identified for hypocotyl length in response to different temperatures. S*MALL AUXIN UPREGULATED RNA 38* (*SAUR38*) is a novel candidate gene for a QTL. Another major-effect QTL 'Temp22.2' was also identified which harbours the candidate gene *PHYTOCHROME B* (*PHYB*). *Secondly*, fine-mapped 'Environmental QTL' were also discovered for the genotype by environment (G x E) interaction. *PHYTOCHROME D* (*PHYD*) is a candidate for a temperature-responsiveness QTL.

For QTL cloning, functional characterisation of *SAUR38* and *PHYB* was carried out by knockout analysis and transgenic allelic complementation. The results showed that *SAUR38* controlled natural variation in the Tsu-0 accession by increasing elongation. The *PHYB* alleles of Ct-1, Sf-2 and Col-0 accessions explain the Temp22.2 QTL. Ct-1 and Sf-2 alleles are positive regulators increasing elongation, whereas Col-0 allele is a negative regulator. Temp22.2 QTL was cloned and novel alleles were discovered revealing the molecular basis of quantitative variation in hypocotyl length in response to temperature.

Acknowledgements

First, I would like to thank Almighty God for all the courage and determination that He has given to me for the completion of this study.

I owe my supervisors, Professor Alastair Grant (School of Environmental Sciences, University of East Anglia) and Dr Philip Wigge (Sainsbury Laboratory, University of Cambridge), a debt of profound gratitude for all their continued advice, support and encouragement throughout my PhD. They have been a constant source of intellectual inspiration and I would like to thank them for giving me the opportunity to do this PhD. Many thanks go to them for reading my thesis during its preparation and for their helpful comments and suggestions.

At the University of East Anglia, Norwich, UK, I am thankful to John Turner, Colwyn Thomas and Richard Bowater for fruitful discussions; and Judith Mayne who helped in the laboratory.

At the John Innes Centre, Norwich, UK, I would like to thank all the members of the Wigge Lab for interesting group meetings and discussions. I am also grateful to Andreas Magusin who helped with the computational analysis for the QTL mapping. Further thanks go to Andrew Davis for very nice photography, and to Barry Robertson and his horticultural team for help in the controlled environmental chambers and glasshouses. Thanks to Caroline Dean, John Snape, Clare Lister, Judith Irwin, Peijin Li, Colin Morgan, Faridoon Yousafzai, Meluleki Zikhali and Susan Duncan for helpful discussions.

I thank my sponsors, The University of Peshawar, Pakistan and The University of East Anglia, UK who have financially supported this PhD.

On my personal front, I would like to thank all my family members for their love, prayers and extraordinary support during my PhD. I owe a lot to my parents, who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. I would like to especially thank my wife who was always there for me and supported me in every possible way to see the completion of this work. I would also like to especially express thanks to my beloved daughter, Hoorain Khattak, who always was an excellent girl and cheered me up, and who helped and enjoyed colouring in the bars in the plots in SPSS.

Table of Contents

Chapter 1

Introduction

1.1 Introduction:

Temperature is one of the most important environmental abiotic factors that affect plant growth and development. Changes in ambient temperature affect plant growth throughout its life cycle at all different stages of development from seed germination, seedling establishment, vegetative growth and flowering time. The problem of global warming is a key concern for plants for they are sessile in nature; and to survive and reproduce they need to rapidly adapt their growth and development to their changing warmer environments. The threats of climate change on plant species and ecosystem dynamics are obvious and increases in mean global temperature poses a significant environmental challenge to them. In such an ever changing climate, it is important to develop our understanding of how plant growth responds to ambient temperatures and identify the underlying genetic basis of this growth regulation.

Plants have developed sophisticated signal transduction mechanisms that allow them to show remarkable growth responses to even small changes in key environmental abiotic factors such as light and ambient temperature. Temperature is a major environmental signal (Franklin, 2010) and plants can sense temperature changes of as slight as 1°C (Argyris *et al*., 2005; Penfield, 2008; Kumar and Wigge, 2010) and alter their growth and development accordingly. The rise in mean global temperatures due to climate change will be a great challenge for plants to adapt to changing warmer climates. This is because the effects of temperature variation on plants can not only be seen in physiological processes such as photosynthesis and metabolism, but also on their growth and development related phenotypes. Minute changes in ambient temperature affect growth and the various stages of development (Penfield, 2008).

The threats of our changing climate on biodiversity and ecosystems are profound (Barnosky *et al*., 2012). The effects of climate change on various plant species have been

well studied, e.g., in the temperate zones, with no changes in photoperiod, elevated spring temperatures have resulted in the accelerated onset of flowering in a number of plant species (Sparks *et al*., 2000; Menzel *et al*., 2001; Fitter and Fitter, 2002) representing altered plant phenology. As a consequence of climate change, plants are already displaying rapid adaptive evolutionary responses in their flowering time (Franks *et al*., 2006). Climate change is also responsible for altered plant distributions (Willis *et al*., 2008). Furthermore, it has been estimated that one-third of all plant species in Europe will go extinct solely due to the effects of increased temperature as a consequence of climate change (Thuiller *et al*., 2005). Crop plants also face serious threats from a warmer world by affecting yields. The significant harmful effects of elevated temperatures on crop yield have already been highlighted (Battisti and Naylor, 2009). It is predicted that as little as 1 – 2°C rise in temperature will have noticeably unfavourable effects on crop yields (Tubiello *et al*., 2007).

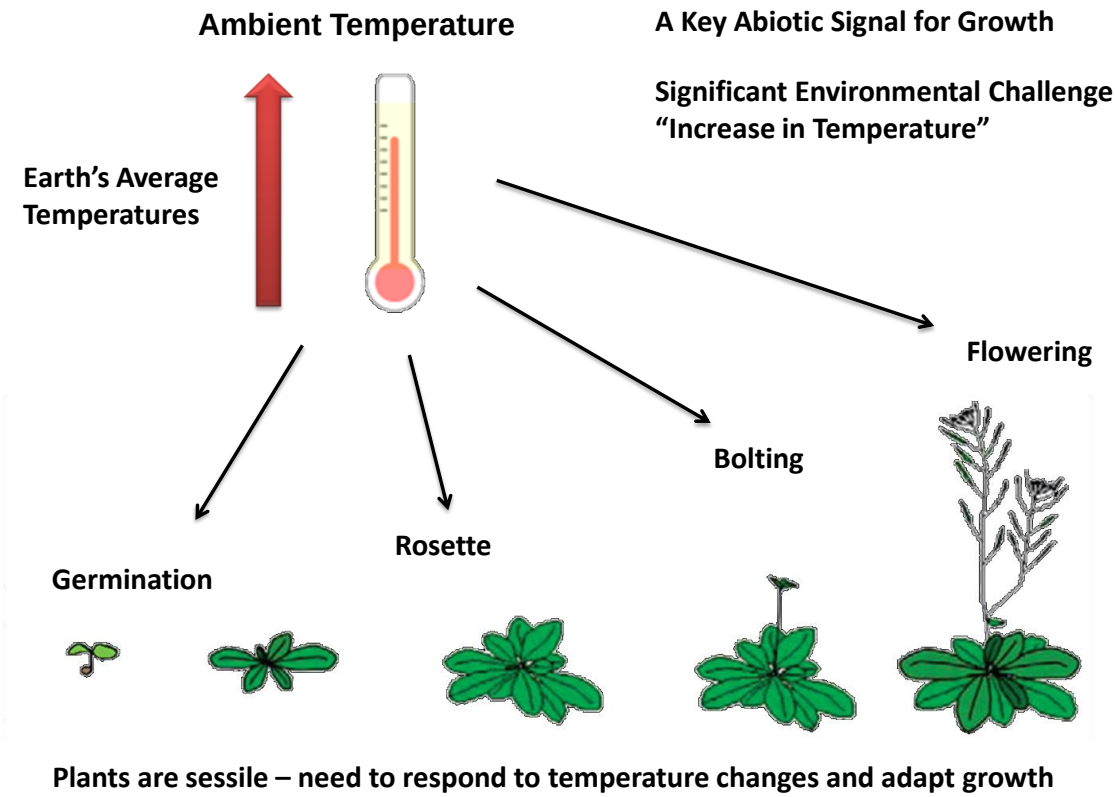

Figure 1.1: Ambient temperature is a key abiotic signal that controls plant growth across all the various developmental stages from juvenile to adult.

Recent climate change trends have shown a steady increase in mean global temperatures. Throughout the last millennium, the period covering the last several decades of the 20th century witnessed an unprecedented rapid global warming (Stott *et al*., 2000). Since the late 19th century, our Earth has experienced an increase in the mean global surface temperature and the decade of 2000's has been the warmest (IPCC, 2013). From 1880 to 2012 our planet underwent a 0.85°C rise in the mean global temperature and it is expected from projections that the mean global surface temperature increase will be in the range of 0.3°C to 0.7°C for the period 2016–2035 relative to 1986–2005 (IPCC, 2013). The rising temperatures will affect local habitats and ecosystems; and this is leading to increased biological responses to global warming (Parmesan, 2006). Global warming is an environmental problem due to gradual increase in average temperatures, but at the same time it is also an evolutionary one, particularly for plants as they can potentially evolve in response to higher temperatures. This depends on the amount of genetic variation present in a particular species, known as intraspecific variation, which is the most primary level of biodiversity. Climate change not only threatens species extinction but also loss of intra-specific diversity termed as cryptic diversity (Balint *et al*., 2011). Therefore, identifying functional natural allelic variation is of high significance.

For survival it is necessary that plants endure changes in environmental temperature. This can be achieved in two ways. To cope with the changing environment plants over the course of their lifetime may undergo phenotypic plasticity, which means that their features can be influenced to a certain degree by the surrounding environment. This includes any kind of change that is not a consequence of changes in its genetic makeup. Hence, plasticity is likely to be an essential feature of persistence of species in a rapidly changing climate (Nicotra *et al*., 2010). Another way of coping with environmental change in longer time scales is to adapt to the environment. Plant populations can adapt if they evolve in response to climate change by improving the function of a particular trait in a specific environment. Many species that lack both the phenotypic plasticity and natural variation may go extinct in the face of global climate change.

Whether plants will be able to adapt to higher temperatures, due to increase in global warming, depends primarily on the genetics of the plant species. Genetic variation in natural plant populations for a particular species is necessary for adaptation to changing

conditions. Naturally occurring variation is likely to be instrumental in allowing adaptation to new temperature environments. Therefore, the whole issue of such responses is addressed by origin and maintenance of natural variation in plant populations. Due to environmental change and related selection pressures, plants will experience growth and development related challenges. Under such circumstances, genetic variation in plants will allow the maintenance of its evolutionary potential and this could lead to their persistence (Hoffmann and Willi, 2008). Polymorphisms having functional effects on phenotypic variation must be identified to comprehend how ecologically important natural variation is evolved (Mitchell-Olds and Schmitt, 2006).

1.1.1 Study of Intraspecific Variation in Response to Temperature:

Arabidopsis thaliana (Arabidopsis) is a small annual flowering plant and it is a model experimental system for molecular and genetic studies (Meyerowitz, 1987).

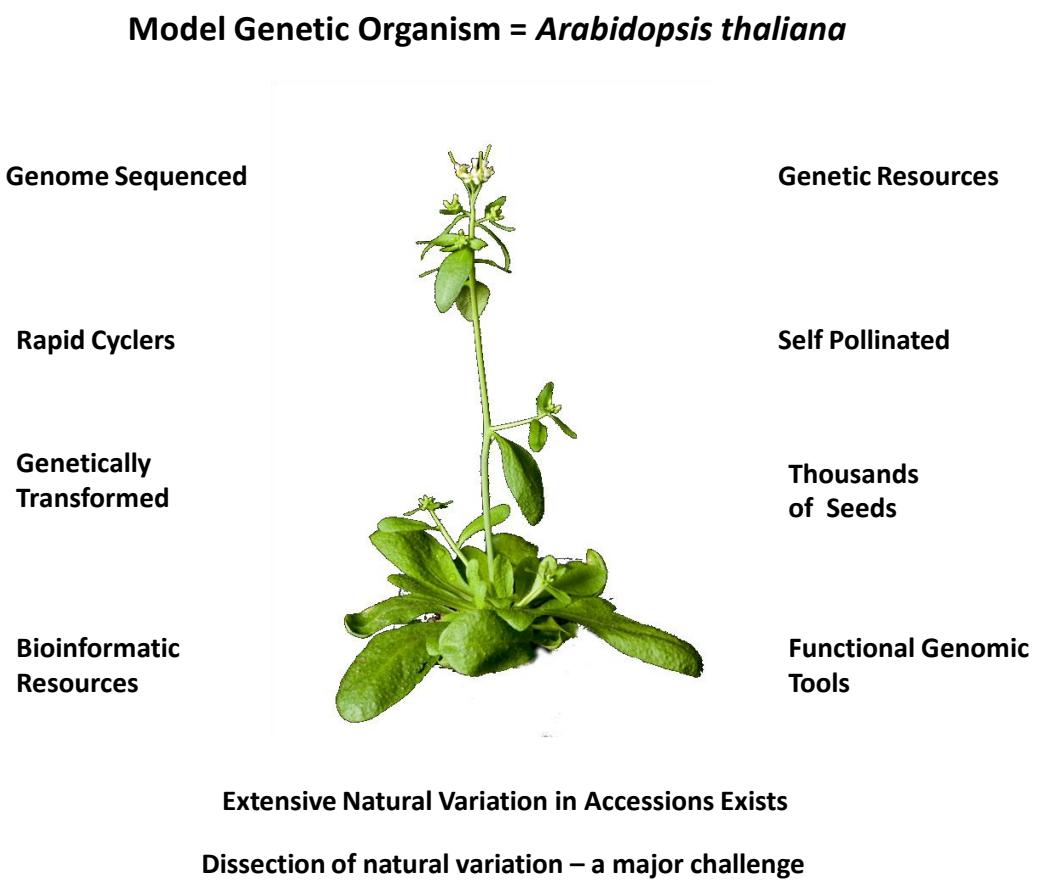

Figure 1.2: Image of Arabidopsis showing characteristics that help make it a useful model organism.

Arabidopsis is broadly distributed in the northern hemisphere (Al-Shebaz and O Kane, 2002; Hoffmann, 2002). It grows in various global climates and habitats. They are not only found in natural landscapes such as fields, rocky slopes and Mediterranean forests, but also in anthropogenic areas such as roadsides and pavements (Hoffmann, 2002; Pico *et al*., 2008).

Arabidopsis is an efficient model plant due to several biological characteristics (Somerville and Koornneef, 2002). It has a short life cycle (6-7 Weeks), a small genome (~150 Mb) and can be easily genetically transformed (Meinke *et al*., 1998). Moreover, it is a selfing plant and natural accessions are homozygous which produce large quantity of seeds (~10,000). Though its genome is relatively small, it still shares the functional genomic characteristics of other higher plants. Arabidopsis was proposed as a model plant for genome analysis (Meinke *et al*., 1998) and due to an international effort the accession Columbia (Col-0) genome was sequenced (AGI, 2000), the first in the plant kingdom. Moreover, various experimental tools are available which include large genetic resources; and bioinformatic resources are becoming increasingly available. Therefore, because of all of its genetic characteristics, it is a model organism for plant genetics and development. Additionally, Arabidopsis is an ideal experimental system for integrating genomics with the study of ecology and evolution (Mitchell-Olds, 2001). The study of ecological and evolutionary processes are not well known from a climate change point of view, including the genetic basis of local adaptation (Anderson *et al*., 2012), and in this regard, the synthesis of functional genomics with the study of ecology and evolution is an interesting emerging interdisciplinary field (Jackson *et al*., 2002; Mitchell-Olds and Schmitt, 2006).

One of the key sources of genetic variation in natural populations is mutation. Over the course of evolutionary history, mutations have occurred in nature, which are now reflected in the genetic diversity seen in Arabidopsis accessions. Understanding the nature of this variation at the molecular level for important plant architectural traits can help us to understand how plant species adapt to their surrounding environment. Arabidopsis is a largely selfing species (Abbott & Gomes, 1989) and a lot of natural variation is distributed among instead of within wild populations. The Arabidopsis accessions collected worldwide vary from one another at thousands of loci and this represents naturally occurring genetic variation that cannot be achieved in the laboratory

(Alonso-Blanco and Koornneef, 2000). This makes these accessions a powerful natural resource to dissect the molecular basis of natural variation for traits of interest. Genetic analyses can be used to investigate responses to environmental factors in these accessions. Differences in accessions are suitable to unravel the molecular basis of their responses to environmental factors for important morphological traits.

The identification of mutant alleles, either dominant or recessive, related with phenotypic variation in Arabidopsis accessions will be a challenge (Meinke, 2013). Furthermore, to understand the molecular and genetic basis of naturally occurring variation in plant developmental traits is one of the major challenges of current biology (Alonso-Blanco *et al*., 2005). It becomes more interesting when we aim to study this variation in response to essential environmental factors such as temperature. This addresses the main notion of whether Arabidopsis accessions respond differentially to temperature changes based on their natural variation. Greater the variation of growth responses to temperature, greater would be the ability of the species to adapt to changing warmer climates.

Generally, plant growth is a trait that is quantitative in nature, showing a continuous distribution, and is controlled by multiple loci and the environment. Quantitative traits are difficult to study as they are under the control of multiple genes and are also influenced by the environment. Undertaking research on quantitative traits related to plant growth and development, in response to temperature, can potentially improve our basic knowledge of the underlying genetic components that contribute to the variation of these traits.

It is quite interesting that those environmental factors such as temperature and light which affect the growth of adult plants also influence the hypocotyl, which is the seedling stem. Therefore, the hypocotyl of Arabidopsis is an important seedling organ and is an excellent proxy of elongation growth in plants. It is a model organ for investigating the effects of environmental factors on its growth. Seedling development is very plastic and is modulated by ambient temperature responses. Plants need to cope with their ever changing environment and successful survival is dependent upon their ability to integrate environmental cues to regulate their growth and development. This ability is most important at the seedling stage and hypocotyl length plays a pivotal role in seedling establishment, which is an ecologically critical developmental process. Hypocotyl length is

one of the important life history traits of plants and is necessary for plant survival in the early juvenile stage. Hypocotyl elongation is a well characterised seedling process (Quail *et al*., 1995; Quail, 2002) which allows the seedlings to initiate photoautotrophic growth. The hypocotyl, in the early photomorphogenesis stage, enables the seedling to cut through the soil and reach sunlight so that the first leaves can appear and start growing in order to initiate photosynthesis for the entire plant. It is of significance both from an ecological and molecular point of view because in addition to the environmental factors it is also influenced by endogenous phytohormone regulators. Multiple cues integrate to control the developmental process of hypocotyl elongation (Alabadí and Blázquez, 2009) and therefore, it is a good system to study the Genotype x Environment (G x E) interactions. In light of all this, the hypocotyl is ideal for ecological, evolutionary and molecular genetic examination.

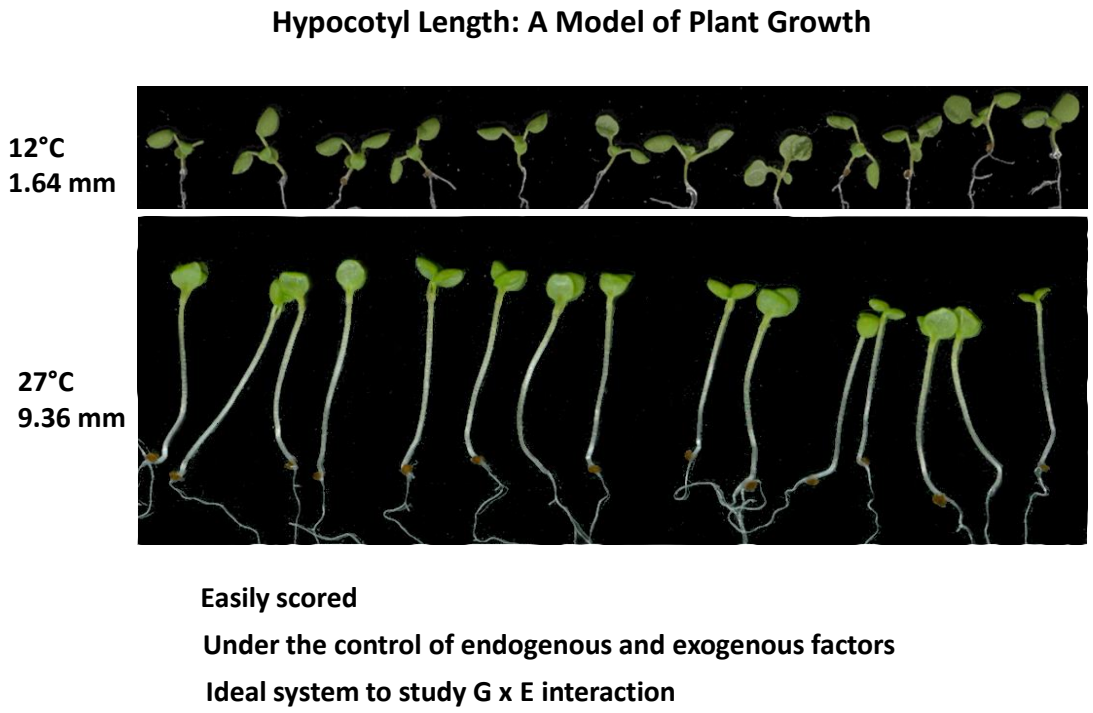

Figure 1.3: 1 week old hypocotyl lengths revealing response to temperature; the hypocotyl lengths are elongated at higher ambient temperature as compared to lower ambient temperature. A recombinant inbred line was phenotyped in short days (8 hours light/16 hours dark) at 170 µE/m^2/sec.

Arabidopsis has a worldwide distribution (Redei, 1970) which reflects variable plant growth environments. Because of diverse geographic origins, the accessions have retained the best alleles for their survival. Hence the study of variation in response to temperature can to help identify the genes and alleles in specific accessions that have been used to survive in the native environments. It is expected that phenotypic variation in hypocotyl length among natural accessions represents the genetic variation which is necessary for adaptation to specific environments. It is anticipated that analysis of the natural genetic variation in hypocotyl length will provide distinctive knowledge from functional perspectives of candidate genes involved and potentially from an ecological perspective too.

1.1.2 Hypocotyl Biology and Factors affecting its Growth:

The hypocotyl is a relatively less complex organ for genetic studies because during germination of the seedling it does not grow as a result of cell division but rather due to cell elongation. It was shown that the hypocotyl cells, after forming in the seed embryo, do not undergo any further divisions after germination; rather the cells elongate during seedling emergence (Gendreau *et al*., 1997). However, the developmental process of hypocotyl elongation is complex, as it is controlled by both the environment and hormones (Jensen *et al*., 1998). Hence, in the hypocotyls, all growth takes place by cell expansion and elongation until saturation is reached. The natural variation affects hypocotyl elongation by modulating cell expansion.

Although hypocotyl length is a manifestation of the genotype, its phenotypic expression is strongly influenced by light, temperature, and hormones. The two most important exogenous factors which crosstalk to regulate general plant development are light and temperature (Franklin, 2009). These two environmental factors affect hypocotyl elongation. *Firstly*, hypocotyl elongation is controlled by light and various phytochromes (*PHYA – PHYE*), which encode for plant photoreceptors (Mathews and Sharrock, 1997) that mediate this response. Phytochrome mutants were identified with long hypocotyls initially through mutagenesis screens (Somers *et al*., 1991; Reed *et al*., 1993; Parks *et al*., 1993) that revealed genetic control of the hypocotyl. Light represses hypocotyl elongation and mutations in the phytochromes give rise to seedlings with long hypocotyls (Shin *et al*., 2009).

Another family of photoreceptors called cryptochromes also control hypocotyl elongation in blue light. Cryptochrome 1 (cry1) and cryptochrome 2 (cry2) are two receptors that mediate several light responses in plants. The functions of cryptochromes in hypocotyl elongation were identified from mutagenesis screens. The isolation of mutants deficient in cry1 and cry2 identified roles for these photoreceptors. Cry1 was identified due to its long hypocotyl phenotype in blue light and with no differences from wild type seedlings grown in R or FR (Koornneef *et al*., 1980; Lin *et al*., 1996). In Arabidopsis, Cry1 mediates inhibition of hypocotyl elongation (Lin *et al*., 1996). Cry2 is also a negative regulator of hypocotyl length. When compared with wild type seedlings, *cry2* null mutants displayed longer hypocotyls (Lin *et al*., 1998).

Light intensity and hypocotyl length are inversely proportional to each other, the lower the light intensity, the bigger the hypocotyls and vice versa. By default, when a seedling germinates it undergoes skotomorphogenesis, i.e. growth in the dark when buried in the soil. This type of growth allows the hypocotyl to elongate and enables the seedling to emerge from the soil. Upon perception of light, plants undergo a developmental change and photomorphogenesis starts i.e. growth in light, and inhibition of hypocotyl elongation takes place; the opening of the apical hook and cotyledons, increase in root growth, and the pigment synthesis is initiated (Chen *et al*., 2004, Franklin *et al*., 2005).

Upon germination, if the seedling senses reduced light quantity and low red/far red ratio due to vegetation canopy in the wild, it undergoes a shade avoidance syndrome which allows the hypocotyl to elongate enabling the seedling to reach higher to receive more light, a response which is mainly mediated by phyB (Smith and Whitelam, 1997). Plants have sophisticated photoreceptors which control these two different types of growth. In Arabidopsis, hypocotyl elongation has been widely studied as a model for light responses, which has revealed various genes coding for photoreceptors and transcription factors (Quail, 2002; Kevei and Nagy, 2003). Plants have evolved various photoreceptor gene families that sense various wavelengths of light.

Secondly, temperature promotes hypocotyl elongation and has an antagonistic effect to light. Temperature is directly proportional to hypocotyl length; the higher the temperature, the bigger the hypocotyls and vice versa. The growth effects of temperature on hypocotyl length both in light and darkness have been previously studied. In

Arabidopsis, temperature-dependent hypocotyl elongation is associated to an increase in auxin levels at higher temperatures in light (Gray *et al*., 1998). It was shown that hypocotyls displayed a dramatic elongation at 29°C as compared to 20°C and it was suggested that growth at higher temperature promoted elevated levels of auxin which resulted in increased elongation of the hypocotyl. In another study, in a temperature-dependent manner, Arabidopsis hypocotyls elongated differentially at various temperatures when grown in darkness; the hypocotyl length at 23°C was bigger than at 16°C (Blázquez *et al*., 2003). Hence in darkness, higher temperatures also increased hypocotyl elongation.

PhyB is one of the important photoreceptors controlling hypocotyl elongation in response to light. In the presence of high light, the Red:Far Red (R:FR) ratio is high. The red-absorbing (Pr) cytosolic form of phyB, after excitation, photoconverts to the far-red-absorbing form (Pfr) where it translocates to the nucleus (Sakamoto and Nagatani, 1996; Kircher *et al*., 1999) and directly interacts with transcription factors called Phytochrome-Interacting Factors (PIFs). This leads to the quick degradation of PIFs (Lorrain *et al*., 2008; Leivar and Quail, 2011) which causes a reduction of growth rate in the seedling stem. Conversely, under low R:FR ratio, the Pfr changes to Pr, that is the inactive form. This conversion allows the PIFs to build up and bind promoters of genes that allow the hypocotyl cells to elongate (Franklin, 2009).

PhyB has been shown to be a major determinant of hypocotyl elongation in white and red light (Reed *et al*., 1993). This has been demonstrated by mutational studies which showed that compared with wild type (WT) seedlings, *phyB* null mutants displayed a long hypocotyl (Reed *et al*., 1993). Moreover, in another study, the seedlings that overexpressed phyB exhibited exaggerated photomorphogenic development with shorter hypocotyls (Wagner *et al*., 1991). PhyB and higher ambient temperature have a close relationship. PhyB has shown to regulate plant architecture at high ambient temperatures (Foreman *et al*., 2011). The results showed that when grown at 28°C, *phyB* mutants have lower biomass and are less viable. This revealed an important role for phyB at higher ambient temperatures. Hence, phyB is vital to temperature response in the hypocotyl growth.

The importance of phyB and temperature interaction can yet be seen in the flowering time trait. In a study, it was shown that a small change from 16°C to 22°C dramatically affected flowering time in the null *phyB* mutants (Halliday *et al*., 2003). The *phyB* mutants flowered earlier at higher temperature but flowered at the same time as WT plants at lower temperature.

Phytochrome-Interacting Factor 4 (*PIF4*) is a transcription factor which incorporates various environmental signals in plant growth and development (Lucyshyn and Wigge, 2009). *PIF4* regulates hypocotyl elongation responses under low light conditions (Lorrain *et al*., 2008). Molecular genetic studies have shown that functional copies of *PIF4* are required for hypocotyl elongation (Huq & Quail, 2002). It is a major target that promotes hypocotyl growth at higher temperatures. Several studies have shown that PIF4 is essential for the elongation of the hypocotyl in response to high temperature (Stavang *et al*., 2009; Koini *et al*., 2009; Foreman *et al*., 2011). *pif4* mutants had significantly smaller hypocotyls at 29°C when compared to Col-0 WT; the hypocotyl length of *pif4* mutants grown at 29°C were similar to Col-0 WT hypocotyls grown at 20°C (Stavang *et al*., 2009).

In another study, *pif4* mutants displayed no hypocotyl elongation when transferred to high temperature (Koini *et al*., 2009); hence PIF4 function is vital for hypocotyl elongation in response to temperature. Recently, it was demonstrated that PIF4 controls synthesis of auxin at high temperature, leading to increased hypocotyl elongation (Franklin *et al*., 2011). Under warm conditions, PIF4 controls hypocotyl elongation by binding to the promoters and regulates the expression of *Tryptophan Aminotransferase of Arabidopsis 1* (*TAA1*) and *CYP79B2*, two genes that encode auxin enzymes (Franklin *et al*., 2011). More recently, it has been demonstrated that PIF4 regulates hypocotyl elongation at higher ambient temperature by directly activating *YUCCA8*, which is another auxin biosynthetic gene (Sun *et al*., 2012).

In another study a direct link between *PIF4* and Phytochrome-Interacting Factor 5 (*PIF5*) and auxin signaling was shown in which *PIF4* and *PIF5* bind to the promoter region of an auxin gene *IAA29*. The results showed that this gene had reduced transcript levels in *pif4*, *pif5* and *pif4pif5* mutants (Hornitschek *et al*., 2012). PIF4 and PIF5 have been implicated in the control of diurnal hypocotyl growth by promoting growth during the dark period at a

time when their proteins are in high quantities, whereas at dawn hypocotyl elongation ceases because the proteins are degraded in the light (Nozue *et al*., 2007).

The hypocotyl elongation in response to temperature has been examined in other species too. For example, in *Abutilon theophrasti*, another annual weed, it was shown that temperature significantly affected hypocotyl elongation (Weinig, 2000). Hypocotyls were bigger at higher temperature than lower temperature under both conditions of simulated foliar shade and full spectrum light.

In addition to light and temperature, which are exogenous environmental cues, endogenous plant growth hormones have been also involved in controlling hypocotyl elongation (Vandenbussche *et al*., 2005). Auxin, gibberellins (GAs) and brassinosteroids increase hypocotyl elongation (Jacobsen and Olszewski, 1993; Romano *et al*., 1995; Clouse, 1996; Cowling and Harberd, 1999), while cytokinins and abscisic acid (ABA) inhibit hypocotyl elongation (Chaudhury *et al*., 1993; Stavang *et al*., 2009). Under skotomorphogenic conditions, ethylene inhibits hypocotyl elongation, whereas, under some photomorphogenic conditions, it promotes hypocotyl elongation (Smalle *et al*., 1997). Hence depending on the light condition, it positively or negatively regulates elongation.

DELLA proteins, like phyB, negatively regulate PIF4. It has been shown that DELLAs are repressors of PIF transcription activity as result of which the PIFs are inactivated (de Lucas *et al*., 2008). In the presence of light due to active phyB, PIF4 is destabilised and DELLAs stop PIF4 transcription which leads to suppression of hypocotyl elongation. The presence of GAs terminates this repression activity leading to the destabilisation of the DELLA proteins. This allows the PIF4 activity to be increased which promotes hypocotyl elongation.

It was also demonstrated that when GAs are not present, the DELLA proteins build up in higher levels and interact with the transcription factor Phytochrome-Interacting Factor 3 (PIF3) (Feng *et al*., 2008). This prevents the binding of PIF3 to target gene promoters and down regulates their gene expression. In this way the PIF3 mediated hypocotyl elongation is terminated. Conversely, in the presence of GAs, DELLA proteins are degraded and this allows the PIF3 to be released from their negative control.

In general, these studies show that, of the external abiotic factors, light responses of hypocotyl length have been well characterised, whereas the ambient temperature responses are not as well studied.

1.1.3 Hypocotyl Length is Quantitative in Nature:

Many Arabidopsis accessions have been surveyed for hypocotyl length when grown under different wavelengths of light and this revealed quantitative variation (Maloof *et al*., 2001; Botto and Smith, 2002). In light of this observation, it is very likely that the hypocotyl length trait shows a continuous distribution when inherited; and therefore it is an excellent growth model to dissect and determine the genetic basis of quantitative variation. Studies on natural variation in hypocotyl growth in Arabidopsis accessions have mainly been carried out in light responses (Maloof *et al*., 2001; Pepper *et al*., 2002; Borevitz *et al*., 2002; Botto *et al*., 2003). This provides a strong indication that accessions can be a valuable natural resource to discover allelic genetic variation in hypocotyl length in response to temperature. The effects of a range of ambient temperatures on accessions and the alleles controlling hypocotyl length natural variation in response to temperature largely remain unknown.

The first review of mapping in natural populations proposed to dissect quantitative traits in natural populations by Quantitative Trait Loci (QTL) mapping and to look into the ecological and evolutionary significance of such variation (Mitchell-Olds, 1995). Due to its importance, this review was regarded as a 'call-to-arms' (Slate, 2005). Complex traits in natural plant populations are inherited quantitatively due to their multifaceted genetic basis (Holland, 2007). Natural variation has been used for the discovery of genes and their functions underlying various traits in Arabidopsis (Koornneef *et al*., 2004). In a similar way, in this study, natural variation is exploited in Arabidopsis to improve our understanding of the gene functions in hypocotyl length in response to ambient temperature. The analysis of accessions found in nature provides a vital source of natural variation that can be utilised for gaining insights into the control of hypocotyl length. Specific allelic variants can be discovered which are present in nature that may have a selective advantage under specific temperature conditions.

The identification of the genetic basis of such traits has been challenging due to the fact that they are under the control of multiple loci and are also affected by environmental factors. The gap between genotype and phenotype can be bridged by identifying the causal genes for natural variation in hypocotyl length in response to temperature. For this purpose there is a need to integrate the disciplines of genetics and molecular biology. Arabidopsis integrates both, due to the widely available methods and resources available and hence is a model system to investigate natural variation in hypocotyl length. QTL mapping, which is the statistical association of the phenotype with the genotype, can be used to dissect the natural variation by identification of the candidate genes that affect hypocotyl length during thermomorphogenesis. It can be pursued down to the molecular level with the help of molecular biology and functional genomic tools in order to clone the QTL, i.e., to find the underlying genes or alleles. QTL cloning is a major objective of quantitative genetics in Arabidopsis for key growth related traits.

Hence, hypocotyl length, which is a plant developmental and architectural trait, can be quantitatively assayed in response to temperature because it is a complex trait in nature; its elongation exhibits a direct proportionality to ambient temperature signals received; it is involved in a straightforward process of seedling emergence; it is a trait that can be easily quantified; it is an important parameter of plant growth; and it may have an ecological adaptive significance.

1.1.4 Plant Growth in Response to Temperature:

Photoreceptors are key regulators controlling various aspects of plant growth and development in response to the environment. Striking temperature-dependent modifications in growth and development related traits have been observed in previous findings, as a consequence of photoreceptor inactivity (Mazzella *et al*., 2000; Blazquez *et al*., 2003; Halliday *et al*., 2003; Dechaine *et al*., 2009). This identifies roles for photoreceptors being involved in growth and development in a temperature-dependent manner. It also shows that temperature and light signalling are linked in nature due to variations in temperature and light in the environment. Interestingly, the photoreceptors also have central functions in the acclimation of plants to their surrounding growth environment (Smith 1995, 2000).

Most of the studies conducted in plants in response to temperature have been focused on freezing and high stress extremes. On one end of the temperature spectrum, cold acclimation, a process where plants are tolerant to sub-zero temperatures, has been extensively investigated in Arabidopsis (Chinnusamy, 2007). It has been shown that transcription factors called C-Repeat Binding Factors (CBFs) are functionally involved in cold acclimation (Stockinger *et al*., 1997). At the other end of the temperature spectrum, plants show responses to extreme high temperatures which are well characterised. Interesting genes have been identified that contribute to the overall thermotolerance in Arabidopsis (e.g. Chen *et al*., 2006; Perez *et al*., 2009). Plants also show acclimation to high temperature stresses, which depends on the expression of heat shock proteins (HSPs) (Queitsch, 2000). The family of HSPs, an important factor in the heat shock response in plants, is activated by Heat Shock Factors (HSFs) (Schramm *et al*., 2008; Liu *et al*., 2008) and are essential for plant thermotolerance reactions that prevent denaturation of target proteins (Sarkar *et al*., 2009) and maintain cell homeostasis (Wang *et al*., 2004). In short, plants are capable of tolerating low and high temperatures. Cold acclimation and thermotolerance allow plants to cope at freezing temperatures and high stress temperatures respectively (Penfield, 2008). Combined, these studies on freezing and high heat demonstrate that extreme temperatures are well studied on both ends of the temperature spectrum. However, the effects of moderate temperatures on plants are less well studied than might be expected (Samach and Wigge, 2005).

Amongst the abiotic environmental factors that affect plant growth and development in Arabidopsis, there is genetic variation present for freezing temperatures (Hannah *et al*., 2006), high stress temperatures (Tonsor *et al*., 2008), drought (Meyer *et al*., 2001; McKay *et al*., 2003; Bouchabke *et al*., 2008), salinity (Katori *et al*., 2010; DeRose-Wilson and Gaut, 2011), UV light (Torabinejad and Caldwell, 2000; Cooley *et al*., 2001), CO_2 (Zhang and Lechowicz, 1995; Tonsor and Scheiner, 2007), and metals (Kobayashi and Koyama, 2002; Hoekenga *et al*., 2003; Baxter *et al*., 2008). Ironically, of all the studies performed on Arabidopsis natural variation in responses to climate change related abiotic factors (Lefebvre, 2009; Assmann, 2013), ambient temperature responses on plant growth are lacking. Studies on ambient temperature responses are greatly required in the face of climate change (Assmann, 2013).

From a climate change perspective, the importance of conducting genetic analyses across a range of environments has been previously emphasised (Nicotra *et al*., 2010). From an ecological point of view, G x E interactions are necessary to understand how ecologically relevant phenotypes interact with the environment. The G x E phenotypic interaction can be observed in Arabidopsis and the genetic basis of this can be detected with the help of QTL analysis by quantifying the QTL x E effects. Several studies have investigated the G x E interactions of the abiotic environmental factors on plant growth. The G x E effects were observed for two growth related traits, shoot biomass and fruit production, in a study on drought response (Hausmann *et al*., 2005). Furthermore, significant G x E effects have also been reported in nutrient responses (Rauh *et al*., 2002; Loudet *et al*., 2003; Prinzenberg *et al*., 2010; Ghandilyan *et al*., 2009). In hypocotyl length, significant G x E interactions has been observed in light and hormonal response (Borevitz *et al*., 2002). In a similar way, this provides a strong indication that significant G x E effects of ambient temperature may also be prevalent for hypocotyl length, which has not been studied in Arabidopsis to date.

As can be observed from the above literature review, previous studies have focussed on either extreme heat or cold to study the effects of temperature on plant growth. However, a recently emerging field of interest is the response of plants to more subtle changes in ambient temperatures (Franklin and Knight, 2010). Our knowledge of the effects of ambient plant growth temperatures is relatively less known amongst the environmental factors that control hypocotyl length. Furthermore, the molecular basis of natural variation in hypocotyl length in response to ambient temperatures is unknown.

1.2 Research Questions:

- Which genes control natural variation in hypocotyl length in response to ambient temperature? And what are the underlying functional alleles?
- Which QTL are responsible for differential hypocotyl length responses to increased temperatures?

To address these questions, a genetic analysis of natural variation in seedling stem length architecture in response to temperature has been performed. In the first question, I am exploring the molecular basis of quantitative variation in hypocotyl length trait at various temperatures; and in the second one, I am discovering the G x E loci which could be involved in the natural variation of a 'temperature-responsiveness trait'; i.e. which 'environmental QTL' are associated with increased temperature responsiveness? Before addressing these questions, initially, it is logical to genetically determine the phenotypic variation in accessions at a range of ambient temperatures; and also investigate how the hypocotyl elongation responds to these temperature differences. This study, in general, aims to explore the phenotype-genotype interaction in an environmental context; and provide a mechanistic understanding, at the molecular level, of hypocotyl length variation and its response to an important ecological signal – temperature.

1.3 Aims and Objectives:

- To explore natural variation in response to a range of ambient temperatures
- To map QTL for hypocotyl length in response to different ambient temperatures
- To map 'environmental QTL' for a temperature-responsiveness trait
- To identify alleles underlying phenotypic variation of the hypocotyl length trait in response to temperature
- To clone novel QTL by functional complementation of QTL candidate genes

1.4 Methodology:

This section provides an overview of the overall methodology adopted to carry out the scientific investigation. Comprehensive research designs for all the different experiments and analyses conducted have been presented in detail in the methods and materials sections in each of the respective chapters.

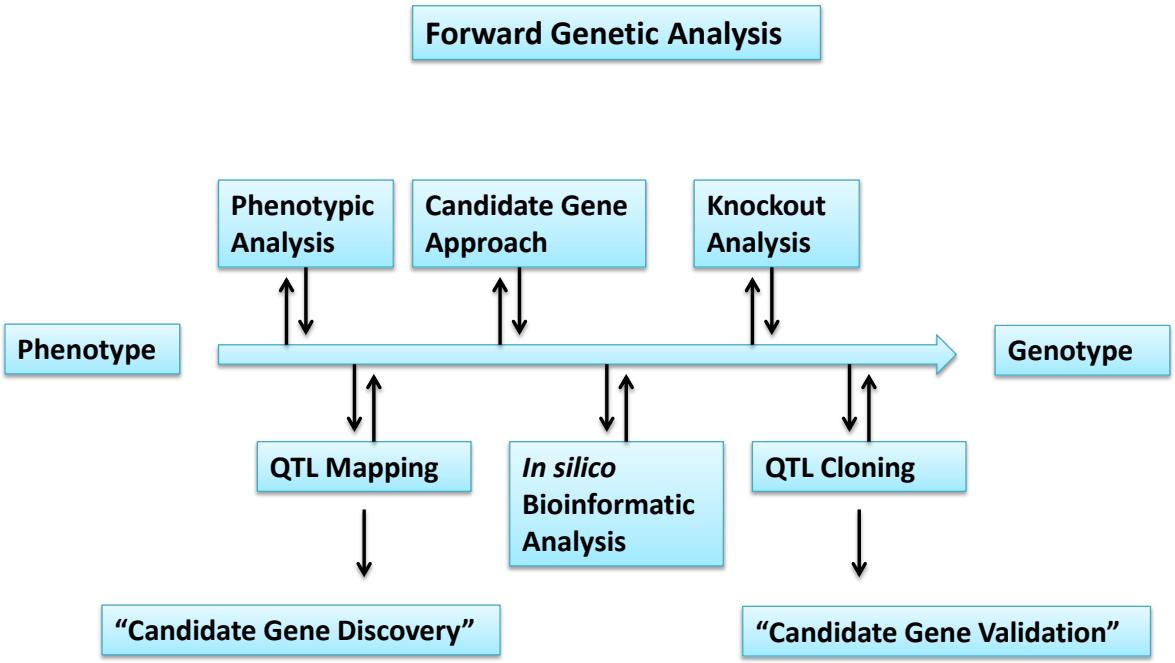

Figure 1.4: Schematic diagram showing a top (phenotype) to bottom (genotype) approach including all milestones for the systematic exploitation of natural variation in order to functionally analyse allelic variation for hypocotyl length. Throughout the course from the phenotype to the genotype, each step feeds back into the mainstream flow facilitating subsequent steps until gene validation.

This methodology aims at dissecting the Arabidopsis intraspecific genetic variation in the hypocotyl length trait in response to temperature. To identify the genetic determinants of phenotypic variation, a forward genetic analysis approach (i.e. from phenotype to genotype) has been adopted for this research study, as seen in figure 1.4. This approach aims to phenotype the trait under investigation at various ambient temperatures, leading to discovering the genotypic factors that contribute to the variation of the trait. This methodology reflects the aims of this study and is described stepwise as follows:

1.4.1 Phenotypic Characterisation:

In order to understand how the hypocotyl elongation responds to various ambient temperatures, a thorough phenotypic analysis was performed and this characterisation

determined the genetically controlled variation. For this purpose, a next-generation QTL experimental population in Arabidopsis known as Multiparent Advanced Generation Inter-Cross (MAGIC) was used which comprises of 19 diverse founding parental accessions and 527 recombinant inbred lines (RILs) (Kover *et al*., 2009). The genetic variation present in the genomes of MAGIC parents has been described after they were sequenced (Gan *et al*., 2011). A detailed description of these lines and their usefulness has been given in Chapter 3. The 19 MAGIC parents were phenotyped for hypocotyl length at 12°C, 17°C, 22°C, and 27°C, which fall within the ambient temperature range of 12°C – 27°C for Arabidopsis (Wigge, 2013). The phenotypic variation was genetically determined by keeping temperature and other environmental variables constant. Since Arabidopsis is well suited for studying the naturally existing phenotypic variation, the hypocotyl length phenotype was measured as an indicator of response to ambient temperature in the accessions. This phenotyping also revealed whether the MAGIC accessions possessed natural variation in their responses to a 5°C or more increase in temperature.

1.4.2 QTL Mapping:

The initial step towards dissecting the molecular basis of complex traits, such as hypocotyl length in Arabidopsis natural accessions, is QTL mapping (Salvi and Tuberosa, 2005). Once genetic variation for hypocotyl length in the parental accessions was established at 12°C, 17°C, 22°C, and 27°C, the next step was phenotyping of the MAGIC RILs at the same four different temperatures. MAGIC RILs are immortal homozygous lines and can be phenotyped under different environmental conditions. This phenotypic quantitative data of the RILs, obtained at these four temperatures, was used for the QTL mapping of two different traits, viz., the 'hypocotyl length trait' and the 'temperature-responsiveness trait'. The quantitative data of the RILs was statistically associated with allelic variation of molecular markers and QTL were identified.

For hypocotyl length, the phenotypic data of the RILs obtained at these respective four temperatures was used in the analysis. This identified QTL at individual temperatures, revealing that QTL are mainly temperature specific. For the temperature-responsiveness trait, in order to map 'environmental QTL' for detection of G x E interactions, all possible

ratios of the four individual temperatures were calculated giving a total of six combinations. This method identified novel QTL for the ratio temperatures; i.e., the genetic basis of the observed phenotypic QTL × E interactions was detected. These QTL have been isolated in response to a change of 5°C or more increase in temperature rather than at an individual temperature. It is these QTL that are contributing to the variation in the temperature responses of the accessions. The 'environmental QTL' for temperature-responsiveness trait were fine mapped and examined for QTL x E interactions. These interactions confirmed unique temperature effects when QTL were only identified in a subset of temperature environments. These QTL display a temperature-dependent correlation with hypocotyl length.

In summary, the interest was to determine how many loci account for the natural variation of these two traits and where they were located. This was performed by QTL mapping which estimated the number and location of QTL segregating in the MAGIC RILs at the four individual temperatures and the six combination ratio temperatures for the two traits.

1.4.3 Candidate Gene Identification:

The post QTL analysis started with a search for strong candidate genes in each of the QTL. A candidate gene approach was used which identified specific genes in a QTL with nucleotide polymorphisms that could possibly determine the phenotypic differences observed in hypocotyl length. Within each identified QTL interval, candidate genes were searched for in the Col-0 accession by looking at the literature, gene ontology and annotations. The next important task was to find the alleles of these candidate genes in the remaining 18 parental accessions for the detection of polymorphisms. For this purpose, several functional genomic tools were used such as the whole genome sequences of the 19 MAGIC parents which were available and proved to be a very useful resource. This opened up for a more realistic sequence polymorphism analysis in the candidate genes and their alleles. With the help of high-throughput bioinformatic approaches, the genome sequence of Col-0, which is well annotated, was used to identify and locate the alleles of these candidate genes in the genomes of the other 18 accessions.

Following their identification, an *in silico* analysis was performed which comprised of multiple sequence alignment and sequence homology methods; and these were used to detect polymorphisms in the protein coding regions of the 19 alleles. Interesting polymorphisms were identified in the major-effect parental accessions that were contributing to the variation of the trait in a particular QTL. These accessions had extreme hypocotyl length QTL phenotypes.

Following the fine mapping of QTL and candidate gene identification, functional strategies were used to test the candidate genes. This was carried out in two steps.

1.4.4 Knockout Analysis:

T-DNA mutant analysis, which is one of the most important functional genomic tools, was the first step allowing functional testing of candidate genes. Since the identification of a T-DNA mutant showing phenotypic effect in hypocotyl length provides further functional evidence for candidate genes, therefore, a knockout analysis was undertaken. This was an efficient strategy to study knockout phenotypes for candidate genes and those which showed interesting altered hypocotyl length phenotypes were further tested for functional complementation in order to clone the QTL.

1.4.5 QTL Cloning by Transgenic Complementation:

The use of molecular genetic tools aims at identifying the genes and their alleles that govern hypocotyl length trait variation. QTL cloning can be performed due to availability of molecular markers, full genome sequences and gene knockouts (Borevitz and Nordborg, 2003; Maloof, 2003); and due to molecular methods of transgenesis. The ultimate proof for a QTL candidate gene comes from by using brute force genetics; and a very powerful method for this is transgenic complementation (Maloof, 2003; Paran and Zamir, 2003) Borevitz and Chory, 2004; Weigel and Nordborg, 2005). Therefore, in order to identify the molecular variation underlying the QTL, which is a major challenge in evolutionary biology (Bergelson and Roux, 2010), a transgenic approach was adopted. Transgenic plants expressing the transgene allele from one accession in the other

accession is a definitive way of QTL cloning which determines the allelic variation for the trait. In Arabidopsis, transgenic complementation has been previously shown, e.g. for the *CRY2* gene (El-Assal *et al*., 2001) and APR2 gene (Loudet *et al*., 2007); and this approach has been applied for QTL cloning in this study. Moreover, the QTL candidate alleles were also transferred into a null background to rescue the hypocotyl length phenotype in the T-DNA knockout mutants.

In summary, for QTL gene identification and to dissect the variation in hypocotyl length, firstly, QTL analysis was performed which has indeed been very successful in Arabidopsis by using RILs. Secondly, a combination of functional genomic and brute force genetics was applied to identify the specific alleles that underlie the QTL in response to temperature.

1.5 References:

Abbott R, Gomes M. (1989) Population Genetic Structure and Outcrossing Rate of *Arabidopsis thaliana* (L.) heynh. Heredity, 62, 411–418.

Alabadí D, Blázquez, M. (2009) Molecular Interactions between Light and Hormone Signaling to Control Plant Growth. Plant Mol. Biol. 69, 409–417.

Al-Shehbaz I, O'Kane S. (2002) Taxonomy and Phylogeny of Arabidopsis (Brassicaceae) in: Somerville CR, Meyerowitz EM, (Ed) The Arabidopsis Book, Rockville, MD: Am.Soc.Plant Biol.

Alonso-Blanco C, Koornneef M. (2000) Naturally Occurring Variation in Arabidopsis: An Underexploited Resource for Plant Genetics. Trends Plant Sci. 5, 1, 22-29.

Anderson J, Panetta A and Mitchell-Olds T. (2012) Evolutionary and Ecological Responses to Anthropogenic Climate Change. Plant Physiology, 160, 4, 1728-1740.

Arabidopsis Genome Initiative (AGI). (2000) Analysis of the Genome Sequence of the Flowering Plant *Arabidopsis thaliana*. Nature, 408, 796–815.

Argyris J, Truco M, Ochoa O, Knapp S, Still D. (2005) Quantitative Trait Loci Associated with Seed and Seedling Traits in Lactuca. Theor Appl Genet. 111, 7, 1365-76.

Assmann S. (2013) Natural Variation in Abiotic Stress and Climate Change Responses in Arabidopsis: Implications for Twenty-First-Century Agriculture, International Journal Of Plant Sciences, 174, Pages: 3-26.

Balint M, Domisch S, Engelhardt C, Haase P, Lehrian S, *et al*. (2011) Cryptic Biodiversity Loss Linked to Global Climate Change. Nat Cli Change 1, 313–318.

Barnosky A, Hadly E, Bascompte J, Berlow E, Brown J, *et al*. (2012) Approaching a State Shift in Earth's Biosphere. Nature 486, 52–58.

Battisti D, Naylor R. (2009) Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. Science 323, 240–244.

Baxter I, Muthukumar B, Park H, Buchner P, Lahner B, *et al.* (2008) Variation in Molybdenum Content across Broadly Distributed Populations of *Arabidopsis thaliana* is controlled by a Mitochondrial Molybdenum Transporter (*MOT1*). PLoS Gene, 4,e1000004.

Bergelson J, Roux F. (2010) Towards Identifying Genes underlying Ecologically Relevant Traits in *Arabidopsis thaliana*. Nature Reviews Genetics, 11, 867-879.

Blazquez M, Ahn J, Weigel D. (2003) A Thermosensory Pathway Controlling Flowering Time in *Arabidopsis thaliana*. Nat. Genet. 33, 168–171.

Borevitz J, Chory J. (2004) Genomics Tools for QTL Analysis and Gene Discovery. Curr. Opin. Plant Biol., 7, 132–136.

Borevitz J, Maloof J, Lutes J, Dabi T, Redfern J, *et al*. (2002) Quantitative Trait Loci Controlling Light and Hormone Response in Two Accessions of *Arabidopsis thaliana*, Genetics, 160, 683–696.

Borevitz J, Nordborg M. (2003) The Impact of Genomics on the Study of Natural Variation in Arabidopsis. Plant Physiol, 132, 718-25.

Botto J, Smith H, (2002) Differential Genetic Variation in Adaptive Strategies to a Common Environmental Signal in Arabidopsis Accessions: Phytochrome-Mediated Shade Avoidance. Plant Cell Environ. 25, 53–63.

Bouchabke O, Chang F, Simon M, Voisin R, Pelletier G, *et al*. (2008) Natural Variation in *Arabidopsis thaliana* as a Tool for Highlighting Differential Drought Responses. PLoS ONE, 3,e1705.

Chaudhury A, Letham S, Craig S, Dennis E. (1993) amp1 - A Mutant with High Cytokinin Levels and Altered Embryonic Pattern, Faster Vegetative Growth, Constitutive Photomorphogenesis and Precocious Flowering. The Plant Journal, 4, 907–916.

Chen M, Chory J, Fankhauser C. (2004) Light Signal Transduction in Higher Plants. Annu Rev Genet, 38, 81-117.

Chen J, Burke J, Velten J, Xin Z. (2006) FtsH11 Protease Plays a Critical Role in Arabidopsis Thermotolerance. The Plant Journal, 48, 73–84.

Chinnusamy V, Zhu J, Zhu J. (2007) Cold Stress Regulation of Gene Expression in Plants. Trends in Plant Science, 12, 444–451.

Clouse S. (1996) Molecular Genetic Studies Confirm the Role of Brassinosteroids in Plant Growth and Development. Plant J., 10, 1, 1-8.

Cooley N, Higgins J, Holmes M, Attridge T. (2001) Ecotypic Differences in Responses of *Arabidopsis thaliana* L. to Elevated Polychromatic UV-A and UV-B+A Inhibition and Growth Rate. J. Photochem. Photobiol., 60,143–50.

Cowling R, Harberd N. (1999) Gibberellins Control Arabidopsis Hypocotyl Growth via Regulation of Cellular Elongation. Journal of Experimental Botany, 50, 337, 1351-1357.

de Lucas M, Davière J, Rodríguez-Falcón M, Pontin M, Iglesias-Pedraz J *et al*., (2008) A Molecular Framework for Light and Gibberellin Control of Cell Elongation. Nature, 24, 451, 7177, 480-4.

Dechaine, J, Gardner G, Weinig C. (2009) Phytochromes Differentially Regulate Seed Germination Responses to Light Quality and Temperature Cues during Seed Maturation. Plant Cell Environ. 32, 1297–1309.

DeRose-Wilson L, Gaut B S. (2011) Mapping Salinity Tolerance during *Arabidopsis thaliana* Germination and Seedling Growth. PLoS ONE 6,e22832.

Doris L, Wigge P. (2009) Plant Development: PIF4 Integrates Diverse Environmental Signals. Current Biology, 19, 6, 265–266.

El-Assal S, Alonso-Blanco C, Peeters A, Raz V, Koornneef M. (2001) A QTL for Flowering Time in Arabidopsis Reveals a Novel Allele of *CRY2*. Nat Genet. 29, 4, 435-440.

Fitter A, Fitter R. (2002) Rapid Changes in Flowering Time in British Plants. Science 296, 1689–1691.

Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, *et al*. (2008) Coordinated Regulation of Arabidopsis thaliana Development by Light and Gibberellins. Nature, 24, 451, 7177, 475-9.

Foreman J, Johansson H, Hornitschek P, Josse E, Fankhauser C, *et al*. (2011), Light Receptor Action is Critical for Maintaining Plant Biomass at Warm Ambient Temperatures. The Plant Journal, 65, 3, 441-452.

Fournier-Level A, Korte A, Cooper M, Nordborg M, Schmitt J, *et al*. (2011) A Map of Local Adaptation in *Arabidopsis thaliana*. Science, 334, 86–89.

Franklin K. (2009) Light and Temperature Signal Crosstalk in Plant Development. Current Opinion in Plant Biology, 12, 1, 63–68.

Franklin K. (2010) Plant Chromatin Feels the Heat. Cell, 140, 1, 26-28.

Franklin K, Knight H. (2010) Unravelling Plant Temperature Signalling Networks. New Phytologist, 185, 8–10.

Franklin K, Larner V, Whitelam G. (2005) The Signal Transducing Photoreceptors of Plants. Int J Dev Biolo, 49, 653-664.

Franklin K, Lee S, Patel D, Kumar S, Spartz A, *et al*. (2011) Phytochrome-Interacting Factor 4 (PIF4) Regulates Auxin Biosynthesis at High Temperature. Proc Natl Acad Sci USA. 108, 50, 20231–20235.

Franks S, Sim S, Weis A. (2006) Rapid Evolution of Flowering Time by an Annual Plant in Response to a Climate Fluctuation. PNAS USA, 104, 4, 1278–1282.

Gan X, Stegle O, Behr J, Steffen J, Drewe P, *et al*. (2011) Multiple Reference Genomes and Transcriptomes for *Arabidopsis thaliana*. Nature, 477, 419–423.

Gendreau E, Traas J, Desnos T, Grandjean O, Caboche M, *et al*. (1997) Cellular Basis of Hypocotyl Growth in *Arabidopsis thaliana*. Plant Physiol. 114, 1, 295–305.

Ghandilyan A, Ilk N, Hanhart C, Mbengue M, Barboza L, *et al*. (2009) A Strong Effect of Growth Medium and Organ Type on the Identification of QTLs for Phytate and Mineral Concentrations in three *Arabidopsis thaliana* RIL Populations. J Exp Bot, 60, 1409–1425.

Halliday K, Salter M, Thingnaes E, Whitelam G. (2003) Phytochrome Control of Flowering is Temperature Sensitive and Correlates with Expression of the Floral Integrator FT. Plant J. 33, 875–85.

Halliday K, Whitelam G. (2003) Changes in Photoperiod or Temperature alter the Functional Relationships between Phytochromes and reveal Roles for phyD and phyE. Plant Physiol. 131, 1913–1920.

Hancock A, Brachi B, Faure N, Horton M, Jarymowycz L, *et al*. (2011) Adaptation to Climate Across the *Arabidopsis thaliana* Genome. Science, 334, 83–86.

Hannah M, Wiese D, Freund S, Fiehn O, Heyer A, *et al*. (2006) Natural Genetic Variation of Freezing Tolerance in Arabidopsis. Plant Physiology, 142, 1, 98-112.

Hausmann N, Juenger T, Sen S, Stowe S, Dawson T, *et al*. (2005) Quantitative Trait Loci affecting δ13C and Response to Differential Water Availability in *Arabidopsis thaliana*. Evolution, 59, 81–96.

Hoekenga O, Vision T, Shaff J, Monforte A, Lee G, *et al*. (2003) Identification and Characterization of Aluminium Tolerance Loci in Arabidopsis (Landsberg erecta × Columbia) by Quantitative Trait Locus Mapping, A Physiologically Simple but Genetically Complex Trait. Plant Physiol. 132, 936–48.

Hoffmann M. (2002) Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). J Biogeogr, 29, 125–134.

Hoffmann A, Willi Y. (2008) Detecting Genetic Responses to Environmental Change. Nat Rev Genet. 9, 6, 421-32.

Holland J. (2007) Genetic Architecture of Complex Traits in Plants. Curr Opin Plant Biol 10, 156–161.

Hornitschek P, Kohnen M, Lorrain S, Rougemont J, Ljung K, *et al*. (2012) Phytochrome Interacting Factors 4 and 5 Control Seedling Growth in Changing Light Conditions by Directly Controlling Auxin Signaling. The Plant Journal, 71, 699–711.

Huq E, Quail P. (2002) PIF4, a Phytochrome-Interacting bHLH Factor, Functions as a Negative Regulator of Phytochrome B Signaling in Arabidopsis. EMBO, 21, 2441–2450.

IPCC, in Climate Change 2013: The Physical Science Basis, (www.ipcc.ch).

Jacobsen S, Olszewski N. (1993) Mutations at the SPINDLY Locus of Arabidopsis alter Gibberellin Signal Transduction. Plant Cell, 5, 8, 887–896.

Jensen P, Hangarter R, and Estelle M. (1998) Auxin Transport Is Required for Hypocotyl Elongation in Light-Grown but Not Dark-Grown Arabidopsis. Plant Physiology, 116, 2, 455-462.

Katori T, Ikeda A, Iuchi S, Kobayashi M, Shinozaki K, *et al*. (2010) Dissecting the Genetic Control of Natural Variation in Salt Tolerance of *Arabidopsis thaliana* Accessions. J Exp Bot 61, 1125–1138.

Kevei E, Nagy F. (2003) Phytochrome Controlled Signalling Cascades in Higher Plants. Physiol. Plant, 117, 305–313.

Kircher S, Kozma-Bognar L, Kim L, Adam E, Harter K, *et al*. (1999) Light Quality-Dependent Nuclear Import of the Plant Photoreceptors Phytochrome A and B. The Plant Cell, 11, 1445–1456.

Kobayashi Y, Koyama H. (2002) QTL Analysis of Al Tolerance in Recombinant Inbred Lines of *Arabidopsis thaliana*. Plant Cell Physiol. 43, 1526–33.

Koini M, Alvey L, Allen T, Tilley C, Harberd N *et al*. (2009) High Temperature-Mediated Adaptations in Plant Architecture require the bHLH Transcription Factor PIF4. Curr Biol. 2009, 19, 408–413.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. (2004) Naturally Occurring Genetic Variation in *Arabidopsis thaliana*, Annu. Rev. Plant Biol., 55, 141–172.

Koornneef M, Rolff E, Spruit C. (1980) Genetic Control of Light-Inhibited Hypocotyl Elongation in Arabidopsis thaliana L. Heynh. Z. Pflanzenphysiol. 100, 147–160.

Kover P, Valdar W, Trakalo J, Scarcelli N, Ehrenreich I, *et al*. (2009) A Multiparent Advanced Generation Inter-Cross to Fine-Map Quantitative Traits in *Arabidopsis thaliana*. PLOS Genetics, 5, 7.

Kumar S, Wigge P. (2010) H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in Arabidopsis. Cell, 8, 140, 1, 136-47.

Lefebvre V, Kiani S, Durand-Tardif M. (2009) A Focus on Natural Variation for Abiotic Constraints Response in the Model Species *Arabidopsis thaliana*. Int J Mol Sci. 10, 3547–3582.

Leivar P, Quail PH. (2011) PIFs: Pivotal Components in a Cellular Signaling Hub, Trends Plant Sci. 16, 1, 19-28.

Lin C, Ahmad M, Cashmore A. (1996) Arabidopsis Cryptochrome 1 is a Soluble Protein Mediating Blue Light-Dependent Regulation of Plant Growth and Development. The Plant Journal, 10, 893–902.

Lin C, Yang H, Guo H, Mockler T, Chen J, *et al*. (1998) Enhancement of Blue Light Sensitivity of Arabidopsis Seedlings by a Blue Light Receptor Cryptochrome 2. Proc. Natl. Acad. Sci. USA, 95, 7686-7699.

Lorrain S, Allen T, Duek P, Whitelam G, Fankhauser C. (2008) Phytochrome-Mediated Inhibition of Shade Avoidance involves Degradation of Growth-Promoting bHLH Transcription Factors. Plant Journal, 53, 312–323.

Loudet O, Saliba-Colombani V, Camilleri C, Calenge F, Gaudon V, *et al*. (2007) Natural Variation for Sulfate Content in *Arabidopsis thaliana* is Highly Controlled by APR2. Nature Genetics, 39, 896 – 900.

Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F. (2003) Quantitative Trait Loci Analysis of Nitrogen Use Efficiency in Arabidopsis. Plant Physiol, 131, 345–358.

Liu H, Gao F, Li G, Han J, Liu D, *et al*. (2008) The Calmodulin-binding Protein Kinase 3 is part of Heat-Shock Signal Transduction in *Arabidopsis thaliana.* Plant J., 55, 760–773.

Maloof J. (2003) Genomic Approaches to Analyzing Natural Variation in *Arabidopsis thaliana.* Curr. Opin. Genet. Dev., 13, 576–582.

Maloof J, Borevitz J, Dabi T, Lutes J, Nehring R, *et al*. (2001) Natural Variation in Light Sensitivity of Arabidopsis. Nat. Genet. 29, 441–46.

Mathews S, Sharrock R. (1997) Phytochrome Gene Diversity, Plant Cell Env. 20, 666–671.

Mazzella M, Bertero D, Casal J. (2000) Temperature-Dependent Internode Elongation in Vegetative Plants of *Arabidopsis thaliana* lacking Phytochrome B and Cryptochrome 1. Planta, 210, 497–501.

McKay J, Richards J, Mitchell-Olds T. (2003) Genetics of Drought Adaptation in *Arabidopsis thaliana*: I. Pleiotropy Contributes to Genetic Correlations among Ecological Traits. Mol. Ecol. 12, 1137–51.

Menzel A, Estrella N, Fabian P. (2001) Spatial and Temporal Variability of the Phenological Seasons in Germany from 1951 to 1996, Global Change Biol. 7, 657–666.

Meyer D, Leonardi A, Brisson G, Vartanian N. (2001) Drought-Adaptive Mechanisms involved in the Escape/Tolerance Strategies of Arabidopsis Landsberg erecta and Columbia Ecotypes and their F1 Reciprocal Progeny. J. Plant Physiol. 158, 1145–52.

Meyerowitz E. (1987) *Arabidopsis thaliana*. Annu. Rev. Genet. 21, 93–111.

Meinke D. (2013) A Survey of Dominant Mutations in *Arabidopsis thaliana.* Trends in Plant Science, 18, 2, 84–91.

Meinke D, Cherry J, Dean C, Rounsley S, Koornneef M. (1998) *Arabidopsis thaliana*: A Model Plant for Genome Analysis. Science, 282, 662-682.

Mitchell-Olds T. (1995) The Molecular Basis of Quantitative Genetic Variation in Natural Populations. Trends in Ecology and Evolution, 10, 324–328.

Mitchell-Olds T, Schmitt J. (2006) Genetic Mechanisms and Evolutionary Significance of Natural Variation in Arabidopsis. Nature, 441, 947–952.

Nicotra A, Atkin O, Bonser S, Davidson A, Finnegan E, *et al*. (2010) Plant Phenotypic Plasticity in a Changing Climate. Trends Plant Sci, 15, 684–692.

Nozue K, Covington M, Duek P, Lorrain S, Fankhauser C, *et al*. (2007) Rhythmic Growth explained by Coincidence between Internal and External Cues. Nature, 448, 358–361.

Paran I, Zamir D. (2003) Quantitative Traits in Plants: Beyond the QTL, Trends in Genetics, 19, 6, 303–306.

Parmesan C. (2006) Ecological and Evolutionary Responses to Recent Climate Change. Annual Review of Ecology Evolution and Systematics, 37, 637–669.

Penfield S. (2008) Temperature Perception and Signal Transduction in Plants, New Phytol. 179, 615–628.

Perez D, Hoyer J, Johnson A, Moody Z, Lopez J, *et al*. (2009) BOBBER1 is a Noncanonical Arabidopsis Small Heat Shock Protein required for both Development and Thermotolerance, Plant Physiol., 151, 241–252.

Picó F, Mendez-Vigo B, Martinez-Zapater J, Alonso-Blanco C. (2008) Natural Genetic Variation of *Arabidopsis thaliana* is Geographically Structured in the Iberian Peninsula. Genetics, 14, 1009–1021.

Prinzenberg A, Barbier H, Salt D, Stich B, Reymond M. (2010) Relationships between Growth, Growth Response to Nutrient Supply, and Ion Content using a Recombinant Inbred Line Population in Arabidopsis. Plant Physiol, 154, 1361–1371.

Quail P. (2002) Phytochrome Photosensory Signalling Networks, Nat. Rev. Mol. Cell Biol. 3, 85–93.

Queitsch C, Hong S, Vierling E, Lindquist S. (2000) Heat Shock Protein 101 Plays a Crucial Role in Thermotolerance in Arabidopsis. The Plant Cell, 12, 479–492.

Rauh B, Basten C, Buckler E. (2002) Quantitative Trait Loci Analysis of Growth Response to Varying Nitrogen Sources in *Arabidopsis thaliana*. Theor Appl Genet, 104, 743–750.

Redei G. (1970) *Arabidopsis thaliana* (L.) Heynh. A Review of the Genetics and Biology. Bibliogr. Genet. 20, 1-151.

Reed J, Nagpal P, Poole D, Furuya M, Chory J. (1993) Mutations in the Gene for the Red/Far-Red Light Receptor Phytochrome B alter Cell Elongation and Physiological Responses throughout Arabidopsis Development. Plant Cell, 5, 147–157.

Romano C, Robson P, Smith H, Estelle M, Klee H. (1995) Transgene-Mediated Auxin Overproduction in Arabidopsis: Hypocotyl Elongation Phenotype and Interactions with

the hy6-1 Hypocotyl Elongation and axr1 Auxin-Resistant Mutants. Plant Mol Biol. 27, 6, 1071-83.

Sakamoto K, Nagatani A. (1996) Nuclear Localization Activity of Phytochrome B. Plant Journal, 10, 859–868.

Salvi S, Tuberosa R. (2005) To Clone or not to Clone Plant QTLs: Present and Future Challenges. Trends Plant Sci, 10, 297–304.

Samach A, Wigge P. (2005) Ambient Temperature Perception in Plants. Current Opinion in Plant Biology, 8, 5, 483–486.

Sarkar N, Kim Y, Grover A. (2009) Rice sHsp Genes: Genomic Organization and Expression Profiling under Stress and Development. BMC Genomics, 10, 393.

Schramm F, Larkindale J, Kiehlmann E, Ganguli A, Englich G, *et al*. (2008) A Cascade of Transcription Factor DREB2A and Heat Stress Transcription Factor HsfA3 Regulates the Heat Stress Response of Arabidopsis. Plant J., 53, 264–274.

Shin J, Kim K, Kang H, Zulfugarov I, Bae G, *et al*. (2009) Phytochromes Promote Seedling Light Responses by Inhibiting Four Negatively-Acting Phytochrome-Interacting Factors. Proc Natl Acad Sci USA, 106, 7660–7665.

Slate J, (2005) Quantitative Trait Locus Mapping in Natural Populations: Progress, Caveats and Future Directions. Molecular Ecology, 14, 2, 363–379.

Smalle J, Haegman M, Kurepa J, Van Montagu M, Straeten D. (1997) Ethylene can Stimulate Arabidopsis Hypocotyl Elongation in the light. PNAS USA, 94, 2756–2761.

Smith H. (1995) Physiological and Ecological Function within the Phytochrome Family. Ann. Rev. Pl. Physiol. Plant Mol. Biol. 46, 289–315.

Smith H. (2000) Phytochromes and Light Signal Perception by Plants – An Emerging Synthesis. Nature, 407, 585–591.

Smith H, Whitelam G. (1997) The Shade Avoidance Syndrome: Multiple Responses Mediated by Multiple Phytochromes. Plant Cell Env. 20, 840–844.

Somerville C, Koornneef M. (2002) Timeline: A Fortunate Choice: The History of Arabidopsis as a Model Plant. Nat Rev Genet, 3, 883–889.

Sparks T, Jeffree E, Jeffree C. (2000) An Examination of the Relationship between Flowering Times and Temperature at the National Scale using Long-Term Phenological Records from the UK. Int. J. Biometeorol. 44, 82–87.

Stavang J, Gallego-Bartolomé J, Gómez M, Yoshida S, Asami T, *et al*. (2009) Hormonal Regulation of Temperature-Induced Growth in Arabidopsis. The Plant Journal, 60, 589–601.

Stockinger E, Gilmour S, Thomashow M. (1997) *Arabidopsis thaliana* CBF1 Encodes an AP2 Domain-Containing Transcriptional Activator that Binds to the C-repeat/DRE, a cis-acting DNA Regulatory Element that Stimulates Transcription in Response to Low Temperature and Water Deficit. PNAS USA, 94, 1035–1040.

Stott P, Tett S, Jones G, Allen M, Mitchell J, *et al*. (2000) External Control of 20th Century Temperature by Natural and Anthropogenic Forcings. Science, 290, 2133–37.

Sun J, Qi L, Li Y, Chu J, Li C. (2012) PIF4-Mediated Activation of *YUCCA8* Expression Integrates Temperature into the Auxin Pathway in Regulating Arabidopsis Hypocotyl growth. PLoS Genet. 8,e1002594.

Thuiller W, Lavorel S, Araujo M, Sykes M, Prentice I. (2005) Climate Change Threats to Plant Diversity in Europe. PNAS USA, 102, 8245–8250.

Tonsor S, Scheiner S. (2007) Plastic Trait Integration across a CO2 Gradient in *Arabidopsis thaliana*. Am Nat, 169, 119–140.

Tonsor S, Scott C, Boumaza I, Liss T, Brodsky J, *et al*. (2008) Heat Shock Protein 101 Effects in *A. thaliana*: Genetic Variation, Fitness and Pleiotropy in Controlled Temperature Conditions. Mol Ecol 17, 1614–1626.

Torabinejad J, Caldwell M. (2000) Inheritance of UV-B Tolerance in Seven Ecotypes of *Arabidopsis thaliana* (L.) Heynh. and their F1 Hybrids. J. Hered. 91, 228–33.

Tubiello F, Soussana J, Howden S. (2007) Crop and Pasture Response to Climate Change. Proc Natl Acad Sci USA. 104, 50, 19686-90.

Vandenbussche F, Verbelen J, Van Der Straeten D (2005) Of Light and Length: Regulation of Hypocotyl Growth in Arabidopsis. BioEssays 27, 275–284.

Wagner D, Tepperman J, Quail P. (1991) Overexpression of Phytochrome B induces a Short Hypocotyl Phenotype in Transgenic Arabidopsis. Plant Cell. 3, 1275–1288.

Wang W, Vinocur B, Shoseyov O, Altman A. (2004) Role of Plant Heat Shock Proteins and Molecular Chaperones in the Abiotic Stress Response. Trends in Plant Science, 5, 244–252.

Weinig C. (2000) Limits to Adaptive Plasticity: Temperature and Photoperiod Influence Shade-Avoidance Responses. American Journal of Botany, 87, 1660–1668.

Wigge P. (2013) Ambient Temperature Signalling in Plants. Current Opinion in Plant Biology, 16, 5, 661–666.

Willis C, Ruhfel B, Primack R, Miller-Rushing J, Davis C. (2008) Phylogenetic Patterns of Species Loss in Thoreau's Woods are Driven by Climate Change. Proc. Natl Acad. Sci. USA, 105, 17029–17033.

Zhang J, Lechowicz M. (1995) Responses to CO_2 Enrichment by Two Genotypes of *Arabidopsis thaliana* differing in their Sensitivity to Nutrient Availability. Ann. Bot. 75, 491–99.

Chapter 2

Phenotypic Characterisation of Thermo-Morphogenic Responses in Arabidopsis Accessions

2.1 Introduction:

The underlying genetic variation in Arabidopsis accessions for hypocotyl length can only be detected if quantitative differences in hypocotyl length exist. Therefore, the scientific objective of this chapter is to carry out a detailed phenotypic analysis and identify natural variation in Arabidopsis accessions that is genetically determined. For this purpose, the MAGIC parental accessions have been phenotyped because as a first step in the forward genetic analysis shown in figure 1.4, the phenotypic characterisation of the parental accessions is essential. This paves the way for the phenotyping of the RILs as it will reveal the extent of natural variation present in the parents of this mapping population for hypocotyl length. Moreover, as shown in chapter 1, these parental accessions have been collected from diverse geographical habitats and represent genetic variation to those native local environments; hence we expect a greater extent of phenotypic variation in hypocotyl length in response to temperature.

Analysis of genetic variants provides a crucial resource of genetic variation which has been used in this chapter to develop a phenotypic understanding into the control of hypocotyl length in Arabidopsis. The hypocotyl length phenotype is a plant growth parameter and it is a quantitative measurement of growth. Naturally existing variation in Arabidopsis has been exploited to discover the thermo-morphogenic responses. The terms ‘accession’ and ‘genotype’ are used synonymously throughout this chapter.

The initial task, as always with non-Mendelian traits, is to identify whether there is any genetic component in the causation of the variation of hypocotyl length phenotype. This task can be achieved by phenotyping the hypocotyl length of different genotypes under the same environmental conditions. The environmental variation, in this way, is controlled and any variation in the phenotype, if present, is determined genetically.

The genetic elements, the environmental factors and the interaction of both affect hypocotyl length and hence it is a complex trait. One such important abiotic environmental factor is ambient temperature. This study investigates its role in the contribution to the variation of hypocotyl length. The effect of temperature can be observed by changing the growing environment of the seedlings by different ambient temperature treatments. The patterns of phenotypic variation across variable temperatures can be examined. This is important because it not only allows us to comprehend natural variation but also responsiveness to varying temperatures. The interest is to dissect these various components and determine the effect sizes of each of these contributing components. This dissection can be feasibly performed across the temperature range and to also identify the G x E interaction.

Additionally, due to the interest in temperature on which this project is based, this chapter also explains the temperature responsiveness of the 19 parental accessions. How does the hypocotyl length in Arabidopsis respond to different ambient temperatures? This is carried out by phenotyping the accessions across a range of ambient temperatures, 12°C, 17°C, 22°C and 27°C, followed by calculating their responses to temperature. These four temperatures chosen for the study fall within the ambient temperature range and a difference of 5 degrees or more will show the responses of the genotypes. For Arabidopsis, temperatures less than 12°C are considered as sub-ambient; whereas temperatures greater than 27°C fall beyond the higher ambient range. For example, a temperature of 30°C has been shown to be the basal thermotolerance for Col-0 accession (Ludwig-Muller *et al.*, 2000). The different temperature treatments have been used to test for the functional involvement of separate temperature responses of genotypes, controlling thermo-morphogenic hypocotyl growth.

Arabidopsis is an appropriate model to study natural phenotypic variation that is present in various accessions. The hypothesis for the phenotyping experiment is that phenotypic variation between Arabidopsis accessions is a consequence of genetic variation between them. And it is predicted that any given accession, which has a unique genotype, will possess a constant phenotype at any given temperature; and this phenotype will remain distinct from other MAGIC accessions. The other prediction is that phenotypes of the MAGIC parents will change with temperature treatments showing differential phenotypic

plasticity. Phenotypic variation is a pre-requisite for QTL mapping and the phenotypic identification of genetic variation in hypocotyl length will be established. Before initiating analysis of natural variation for hypocotyl length, the overall amount of variation existing for hypocotyl length has been explored by a main quantitative genetic study with the 19 MAGIC parental accessions. The results of this phenotypic assessment would determine if a QTL mapping could prove successful.

Phenotypic variation among the MAGIC parental accessions represents genetic variation. This chapter provides a detailed phenotypic analysis of the MAGIC accessions primarily in the trait of hypocotyl length followed by the plant size trait as estimated from the total leaf area and attempts to find any correlation between the juvenile and vegetative stages of development. Hypocotyl length is also proposed as an ecologically relevant trait by investigating the hypocotyl length trait and plant size trait, of this heterogeneous set of accessions, in relation to their native habitats and local temperature conditions.

Objectives of this chapter:

1. To establish the presence of natural genetic variation underlying phenotypic variation for hypocotyl length trait in response to temperature in Arabidopsis
2. To compare the morphological responses of growth in the accessions to a range of different temperatures
3. To study the phenotypic plastic responses of the various accessions to a range of temperature treatments
4. To identify possible extreme variations among the accessions in their responses to various temperatures
5. To analyse temperature-induced hypocotyl elongation and gain a mechanistic understanding of the connection between ambient temperature and hypocotyl length

2.2 Methods and Materials:

2.2.1 Methods and Materials for Hypocotyl Length Experiments:

The phenotypic characterisation of the 19 parental accessions for the identification of natural variation among accessions and of the RILs for QTL mapping study were carried out in controlled growth conditions. This maintains the environmental conditions constant and any variation that is observed in the hypocotyl length is genetically determined.

2.2.1.1 Sourcing of Seeds:

The seeds of the MAGIC accessions and RILs were ordered from The European Arabidopsis Stock Centre (NASC) in Nottingham, UK. The list of accessions is provided in table 2.4.

2.2.1.2 Surface Sterilisation of Seeds:

This process was carried out in a laminar flow hood to prevent any contamination (especially in the Petri dishes). The seeds were first washed with 70% ethanol with 0.01% Triton X-100 (v/v) in a microcentrifuge tube for 3 minutes. The supernatant ethanol was discarded and seeds were given a second wash with 95 % ethanol for 1 minute. With a pipette, the seeds along with the ethanol were put on sterile filter paper (circular in shape; which provides a means for easily sprinkling the seeds over the Petri dishes). The seeds were allowed to completely air dry in the flow hood for a few minutes before sowing.

2.2.1.3 Plating Procedure:

The seeds were sowed in Petri dishes (plates) containing ½ strength Murashige-Skoog growth media (Murashige and Skoog, 1962) with no sucrose; for sucrose has been shown to interfere with light signalling and decreases light responsiveness (Smeekens, 2000) and hence it has strong effects on growth rates and de-etiolation responses. Once completely dried in the flow hood, the seeds were sprinkled evenly over the Petri dishes to create an equidistant distribution of the seeds.

2.2.1.4 Seed Germination:

Seeds were allowed to imbibe after placing them evenly on agar plates. To account for potential variation caused by germination differences and consequently seedling age differences in the hypocotyls, the seeds were stratified in the dark for 4 days at 4°C to ensure uniform seed germination. Stratification not only improves the germination rate but also the synchronisation. After stratification, in order to induce germination and achieve maximum uniformity in germination the plates were placed at a standard temperature of 22°C in a growth incubator in short days. Germination of the seedlings was carefully monitored for 24-36 hours. It should be noted that since the seedlings undergo de-etiolated germination in the presence of light (short days), the cotyledons emerge very shortly after the radicle protrudes from the testa, due to the photomorphogenic state of growth and 22°C temperature. However, when the seedlings rupture the testa and emerge from the seed and if they are still kept at 22°C for longer than required, then the hypocotyls elongate at 22°C and this will confound the hypocotyl elongation results. So the plates were carefully monitored. Once the seedlings emerged from the testa and the cotyledons could be seen, the plates were then placed in the respective testing temperature treatment incubators at 12°C, 17°C, 22°C and 27°C. The photoperiod/light level etc. conditions of germinating the seeds at 22°C were exactly the same under which the hypocotyls were tested and allowed to elongate.

2.2.1.5 Growth Conditions:

Temperature, photoperiod and light intensity play roles in identifying phenotypic variation in hypocotyl length and hence environmental variables have to be carefully selected while phenotyping hypocotyl lengths for laboratory experiments under controlled conditions to avoid any confounding effects. Therefore, precise environmental conditions for phenotyping are crucial for determining differences in Arabidopsis accessions. Before embarking on a detailed parental phenotyping experiment, pilot experiments were run and these environmental conditions were tested which provided valuable insights into the conditions of growth that showed maximum variability in the accessions.

Climate control growth incubators were used for growing the seedlings. After the stratification, the plates were put in light in the incubators because seed germination is induced by light. Plants were grown in short days (8 hours of daylight) because growth differences at different temperatures are more pronounced in short day grown plants. When the seedlings are grown in moderate light intensity (100 microeinsteins per second per square meter (µE/m^2/sec) or less), they exhibit mild etiolation-like effects where the hypocotyls are elongated. Under higher light intensity (200 µE/m^2/sec or more), hypocotyls display stronger de-etiolation phenotypes and they do not elongate and are shorter. Therefore, a light intensity of 150-170 µE/m^2/sec enables the seedlings to respond in such a manner that the effects of etiolation and extreme de-etiolation are avoided. Therefore, a light intensity of 170 µE/m^2/sec was used in the phenotyping of the hypocotyl lengths. Moreover, the duration of light intensity, i.e. whether seedlings are grown under short day (SD) or long day (LD) also affects hypocotyl length. Both of these conditions of light intensity and light duration have been experimentally tested and the results of photoperiod are analysed in detail in section 2.3.7. The humidity in the incubators does not need any regulation because the micro-climate of the agar plates keeps the seedlings humid. No other objects were placed in the incubators that could reduce the light or cause shade and potentially cause the hypocotyls to undergo the shade avoidance syndrome.

The above optimal conditions for seedling growth were selected so that maximum phenotypic variation could be observed in the 19 accessions in response to temperature.

2.2.1.6 Measurement of Hypocotyl Lengths:

Up to 60 individuals per genotype were measured. The seedlings were measured after one week of growth. They were removed from the incubators and phenotyped shortly after the time when the dark cycle had ended in the incubators, representing early morning because maximum hypocotyl growth occurs at dawn under diurnal conditions.

To ensure the hypocotyl length phenotype was measured precisely, an appropriate quantitative assay was set up and specific environmental, temporal and spatial conditions were taken into account to reduce any variation that may be caused by the environment. To obtain unbiased means and variances, the seedlings of the parental accessions were

grown in Petri dishes in a random design for 7 days and then measured for hypocotyl lengths. The error of the means is reduced by measuring many individual seedlings per accession. The hypocotyls were measured on a flat-bed scanner between two transparency sheets and scanned with a black background. The images were saved in TIFF format at a resolution of 600 dpi, so at this resolution, 1 mm = 24 pixels. The software ImageJ was used for image analysis and the scale was set to 24 pixels per mm. The accuracy of ImageJ scaling was tested by measuring a standard size. To set the scale, a known distance of 1mm was measured and the number of pixels was visually counted on the computer screen. The measurements were taken using the segmented line selection tool of ImageJ. The segmented line followed the central axis of the hypocotyl starting at the base of the two cotyledons which is recognised by a dark shaded 'V', to the region of the shoot-root collet junction. This measuring method was repeated for all the seedlings. A total of 3100 individual seedlings were measured for the hypocotyl length analysis.

2.2.1.7 Data Analysis:

All data were analysed using Excel; and the ANOVA and GLM procedures in SPSS. Since the parental accessions have been selected non-randomly based on the extreme genetic variation present between them, therefore, accession has been considered as a fixed effect. Similarly, ambient temperatures have been considered as a fixed effect because they are chosen for this study in a non-random manner. The correlation analyses were determined using Pearson's correlation between the mean values of the trait for each accession.

2.2.2 Methods and Materials for Plant Size Experiments:

The methods for surface sterilisation of seeds, plating procedure, seed germination and growth conditions were all the same as explained in section 2.2.1. However, there was an additional step of transplanting the seedlings to the pots and a different method for quantitative measurements of plant size which are explained as follows.

2.2.2.1 Transplantation:

After one week of growth at the respective 17°C and 22°C temperatures, the seedlings were transplanted on to soil pots. The same soil was used for all the plants in this

experiment. 10 replicates per accession were grown in a randomised design under the same set of environmental conditions. The seedlings were transplanted with precaution to avoid any damage which could affect its growth. This ensured that the seedlings were healthy and grew uniformly under controlled conditions. All the pots were watered equally periodically. The plants were allowed to grow for three more weeks, so in total the plants had grown for four weeks.

2.2.2.2 Plant Size Measurements:

Images were taken after four weeks of growth from above with the help of a tripod stand to ensure the clarity of the pictures. The images were analysed in MATLAB software. The plant size was estimated by quantifying the rosette leaves during the vegetative stage of development. The estimation was based on the total leaf area in the vegetative stage of development. The green pixels of the total leaf area were counted for all of the individuals phenotyped. After the image analysis plant size was estimated by converting the number of pixels into square centimetres.

2.3 Results:

2.3.1 Arabidopsis Accessions Exhibit Natural Variation in Hypocotyl Length:

To investigate natural variation in hypocotyl length in the MAGIC parents, a phenotyping experiment was conducted in which all the environmental variables were kept constant. Arabidopsis MAGIC parents may differ in their hypocotyl lengths when phenotyped at individual ambient temperatures of 12°C, 17°C, 22°C and 27°C. To test this, a one-way ANOVA for between accessions and the 'tests of between-subjects effects' were performed on each of the individual temperatures separately. The hypocotyl length data was either normally distributed or the sample sizes were > 30. The ANOVA results demonstrated that the mean hypocotyl lengths of MAGIC parents were significantly different ($p < 0.001$) at each of the 12°C, 17°C, 22°C and 27°C temperatures. There are differences between accessions at all temperatures. 12°C = $F(18,875) = 332.36$, $p < 0.001$; 17°C = $F(18,829) = 527.7$, $p < 0.001$; 22°C = $F(18,716) = 638.2$, $p < 0.001$; 27°C = $F(18,604) = 490.01$, $p < 0.001$. These results show that the hypocotyl lengths are responding to the individual temperature treatments. The variation in length between the accessions can be seen in figures 2.1 – 2.4.

At all the four temperatures, Ct-1 always has the longest length (1.58 mm, 2.48 mm, 7.05 mm, 11.01 mm) whereas, the shortest accessions are different at each temperature. Hi-0 (1.03 mm), Ws-0 (1.23 mm), Col-0 (1.84 mm) and Wil-2 (4.59 mm) had the shortest mean hypocotyl lengths at 12°C, 17°C, 22°C and 27°C respectively.

As a consequence to this finding, to determine if the effect of the genotypes on the hypocotyl lengths is big or small at each of the individual temperatures, the effect sizes were estimated. The estimates of the effect size were carried out in the 'tests of between-subjects effects'. Partial Eta squared (η^2) provided the measure of effect size and the genotypes explained 87% - 94% of the variance in hypocotyl length. The genotypes account for a very large percentage of variability in the hypocotyl lengths when grown at these four temperatures. Remarkable natural variation in temperature-dependent hypocotyl length responses to a variety of temperature conditions was observed.

A gradual increase in the pattern for effect size at the different temperatures was observed in this analysis. The effect size is lower at lower temperatures and increases with temperature; 12°C < 17°C < 22°C = 27°C. Therefore, the phenotypic variation in the hypocotyl length that the genotypes produce is lower at lower temperatures and higher at higher temperatures. Based on these overall natural variation results in hypocotyl length, a detailed G x E interaction analysis across the range of temperatures, explained by a two-way ANOVA, is provided in section 2.3.5.

To compare the mean hypocotyl lengths of each of the 19 accessions individually between any two given ambient temperature treatments, the independent-samples t-tests were carried out. Surprisingly, all of the accessions had significant ($p < 0.001$) differences in hypocotyl length means between any two given temperature combinations. Even at the lower temperature comparison of 12°C - 17°C, where temperature effect is lower on hypocotyl length than higher ambient temperatures, all of the MAGIC accessions showed significant differences in hypocotyl length between 12°C and 17°C temperatures. This natural variation analysis of hypocotyl length in response to temperature provides a basis for the detailed genetic analyses performed in the subsequent chapters.

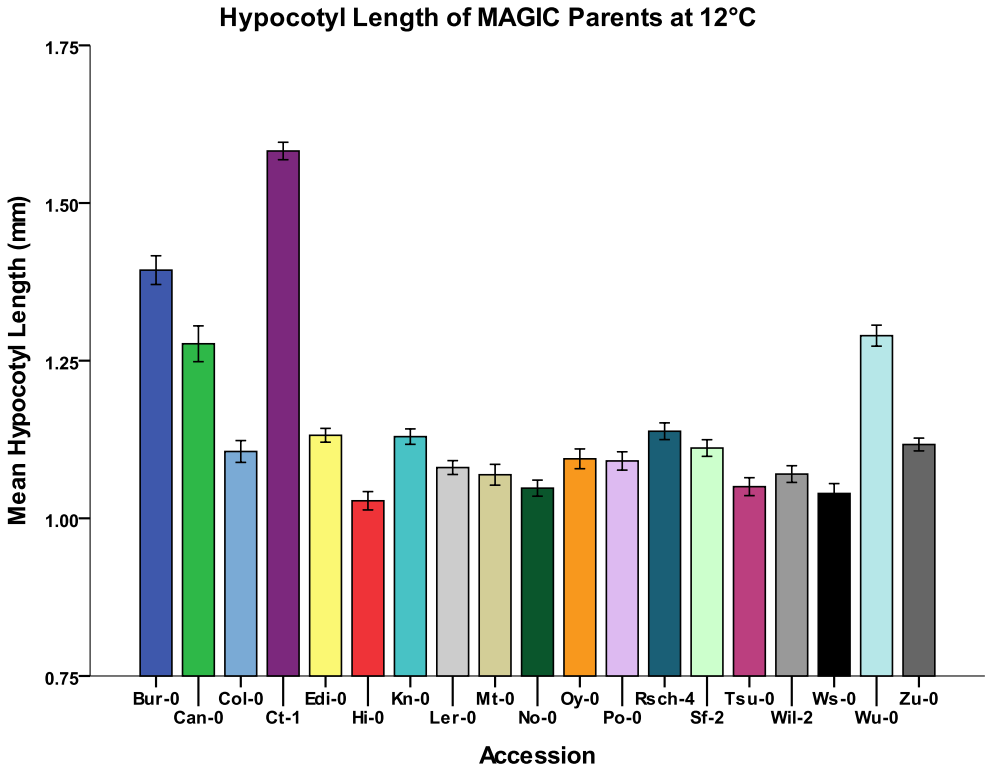

Figure 2.1: Natural variation in 1 week old hypocotyl length in thermal response at 12°C and at 170 µE/m^2/sec short day light. Error Bars: +/- 2 SE.

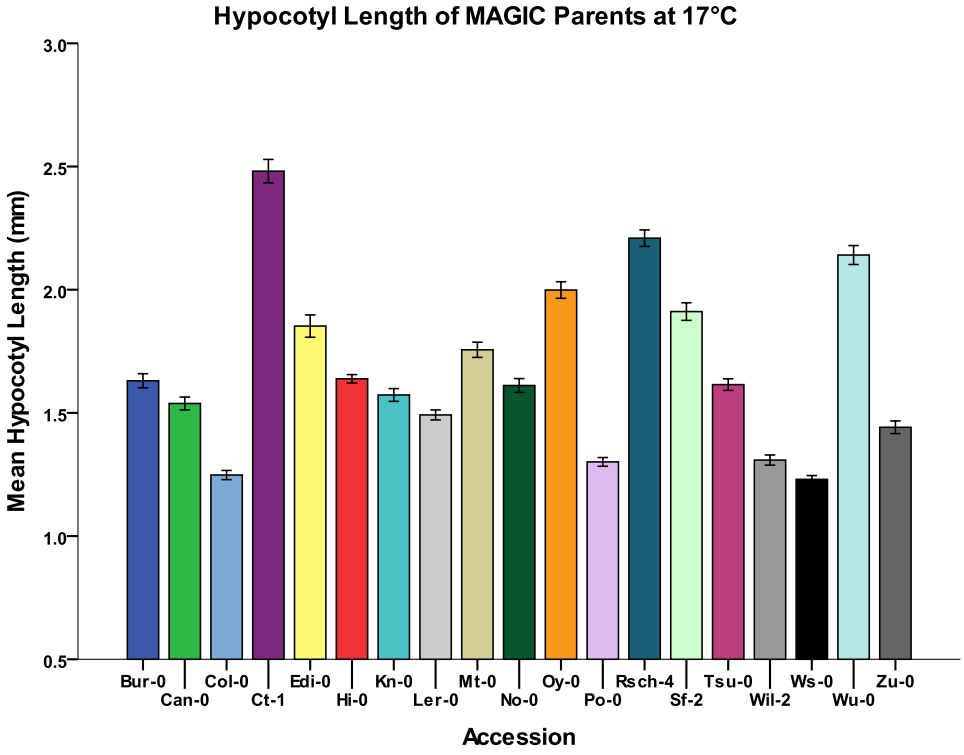

Figure 2.2: Natural variation in 1 week old hypocotyl length in thermal response at 17°C and at 170 µE/m^2/sec short day light. Error Bars: +/- 2 SE.

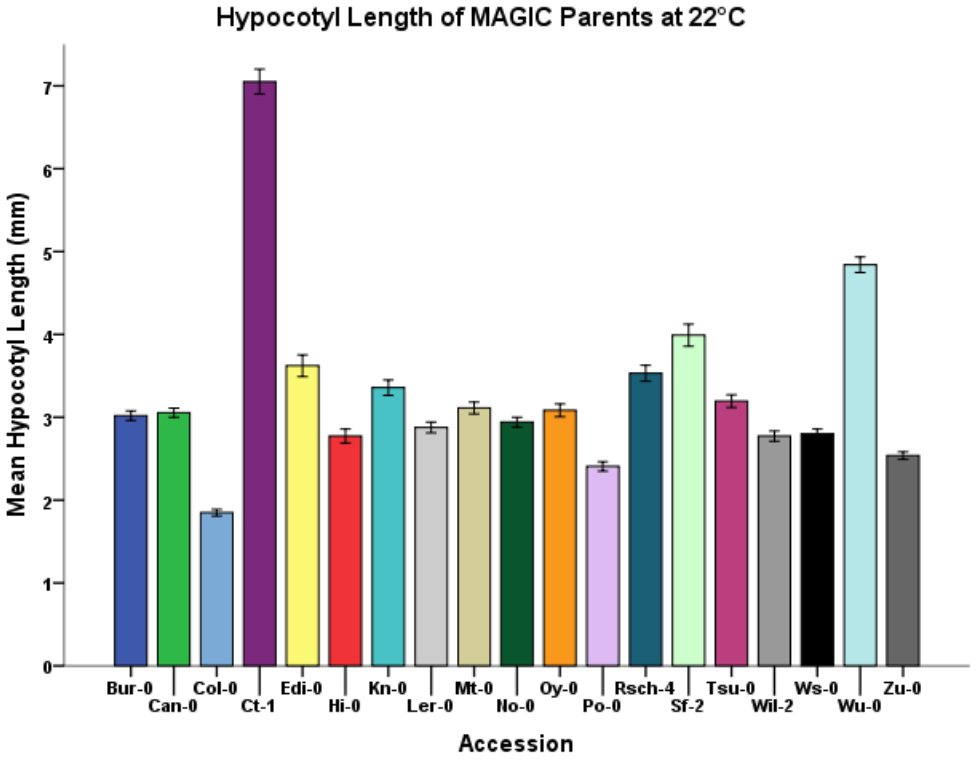

Figure 2.3: Natural variation in 1 week old hypocotyl length in thermal response at 22°C and at 170 µE/m^2/sec short day light. Error Bars: +/- 2 SE.

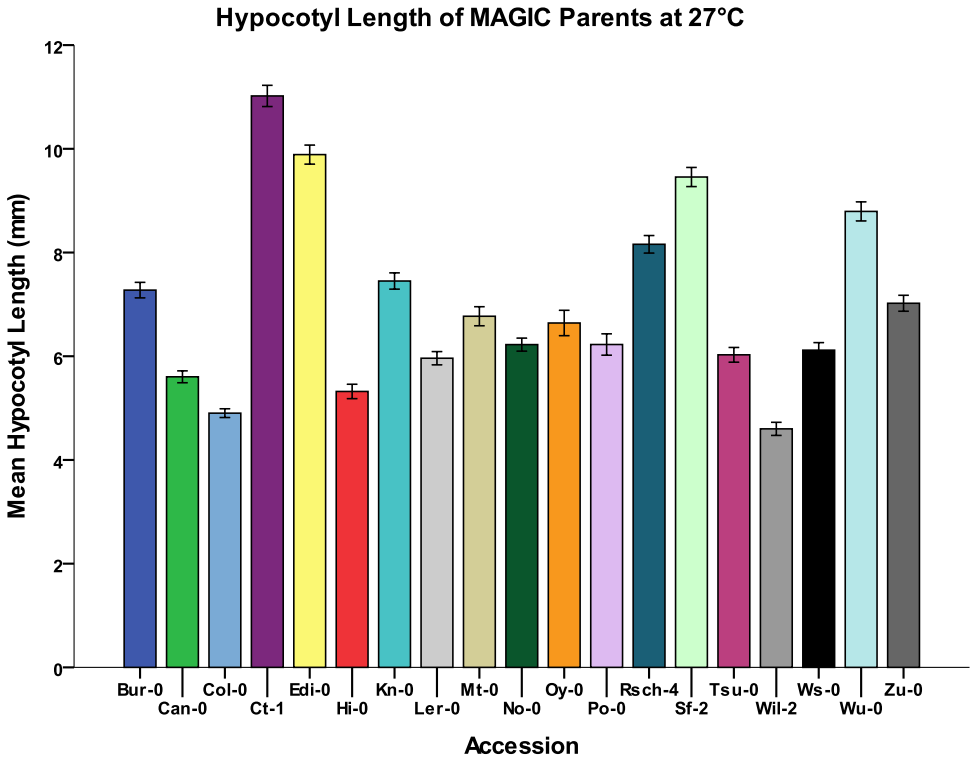

Figure 2.4: Natural variation in 1 week old hypocotyl length in thermal response at 27°C and at 170 µE/m^2/sec short day light. Error Bars: +/- 2 SE.

2.3.2 Skotomorphogenic Natural Variation in Hypocotyl Length:

To test whether there was any genetic variation present when grown in the dark, the MAGIC parental accessions were phenotyped under etiolated conditions and the skotomorphogenic phenotyping results in figure 2.5 reveal the extent to which natural variation is present.

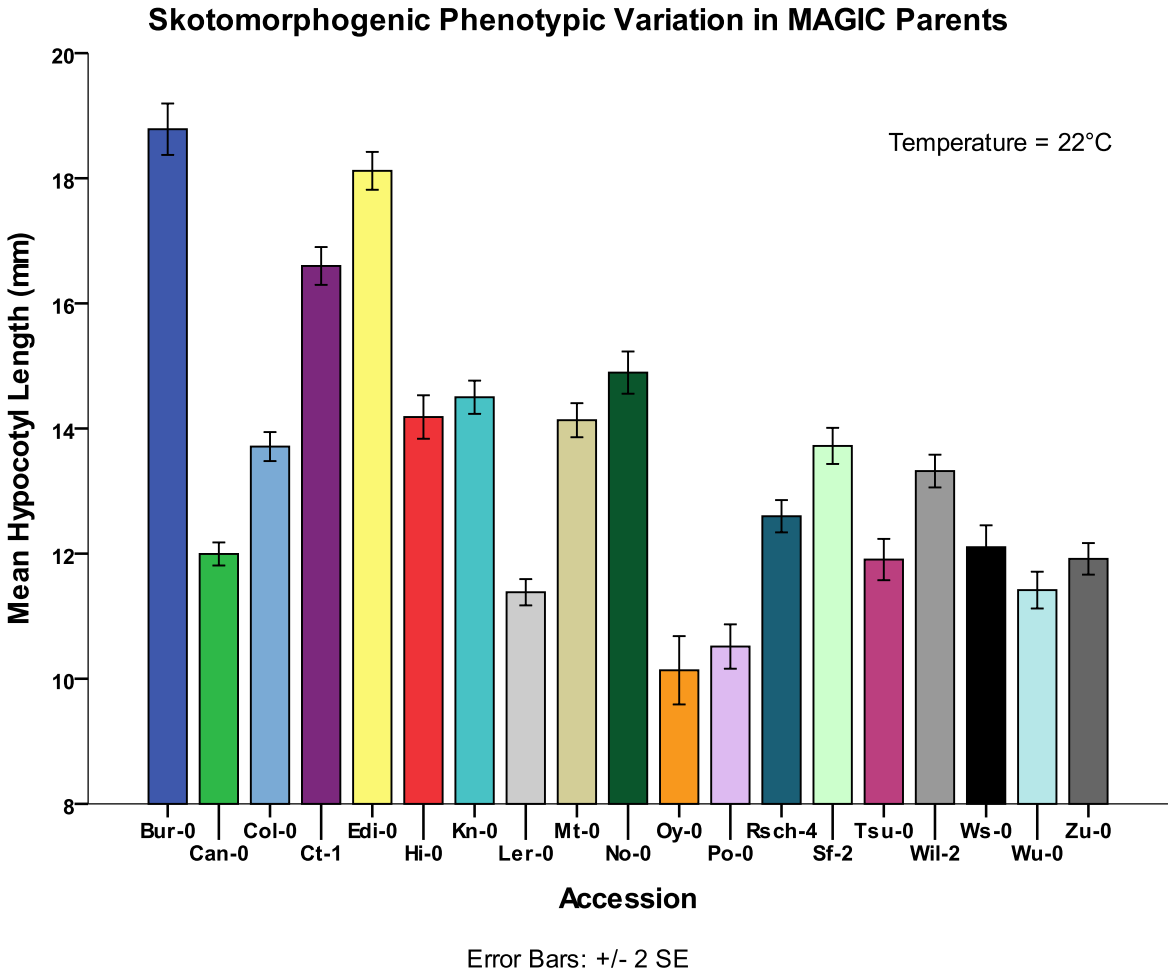

Figure 2.5: Natural variation in 1 week dark-grown hypocotyl length in thermal response at 22°C. Plot showing genetically determined extensive etiolated phenotypic variation in Arabidopsis accessions.

The Arabidopsis MAGIC parents may differ in the hypocotyl lengths when grown in the dark at 22°C ambient temperature. To test this, a one-way ANOVA for between accessions and the 'tests of between-subjects effects' were performed on the quantitative data obtained from phenotyping hypocotyl length in the dark. The hypocotyl

length data was either normally distributed or the sample sizes were > 30. The ANOVA results showed that the mean hypocotyl lengths were statistically significant; 22°C = $F(18,601) = 256.02$, $p < 0.001$. This shows that extensive natural variation is present in hypocotyl lengths when grown in the dark (figure 2.5). The effect sizes of the genotypes on etiolated hypocotyl length were estimated using 'tests of between-subjects effects'. The genotypes explained 89% of the variability in hypocotyl length and substantial variation in etiolation-dependent hypocotyl elongation was observed.

2.3.3 Correlation between Etiolated and De-etiolated Hypocotyl Length:

Furthermore, to investigate if the natural variation in hypocotyl length at 22°C grown in light is correlated to hypocotyl length at 22°C grown in dark, a correlation analysis was performed as can be seen in the scatter plot in figure 2.6. This analysis identified $r = 0.3$ which is showing a weak positive correlation between the dark and light grown hypocotyl lengths. The coefficient of determination, R^2 is near to zero (0.09) and therefore, it is less likely that a linear relationship exists between the two conditions.

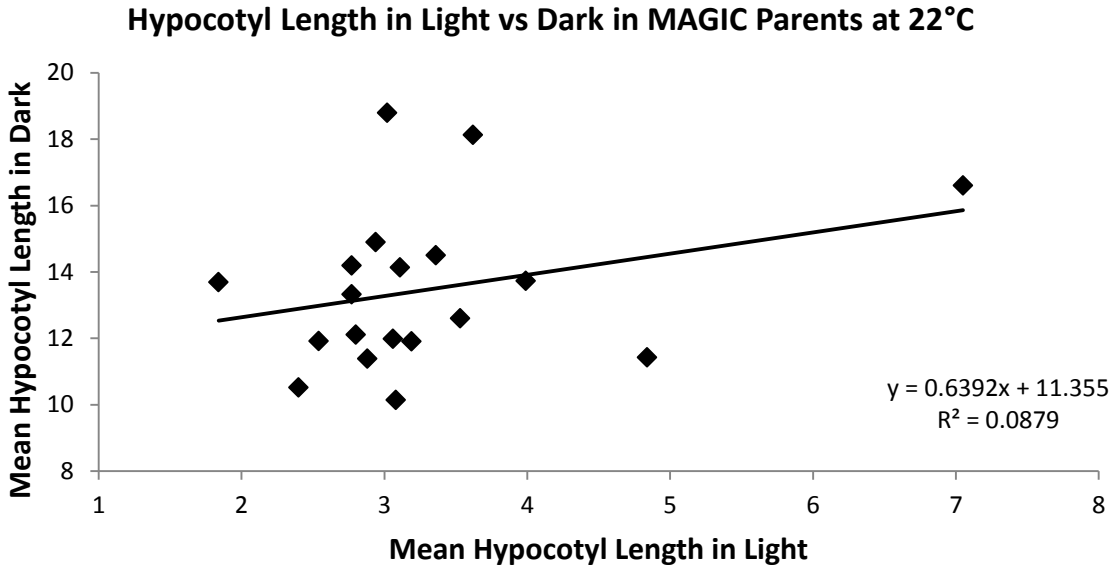

Figure 2.6: Scatter plot of hypocotyl length grown in the light vs. dark at 22°C. The Coefficient of Correlation $r = 0.3$, measures the medium strength and positive direction of the linear relationship. This identifies accessions with strong etiolation responses.

To determine if this identified linear relationship exists in the population from which the samples of accessions were drawn, a significance test showed that there is not a

statistically significant relationship between the two states of growth, r (17) = 0.3, p > 0.05. Additionally, the correlation between hypocotyl length at 27°C grown in light and hypocotyl length at 22°C grown in dark was performed as can be seen in the scatter plot in figure 2.7. This analysis identified r = 0.42 which is also showing a weak to medium positive correlation. The coefficient of determination, R^2 is near to zero (0.1) and therefore, it is less likely that a linear relationship exists between the two conditions. A significance test showed that there is not a statistically significant relationship, r (17) = 0.42, p > 0.05.

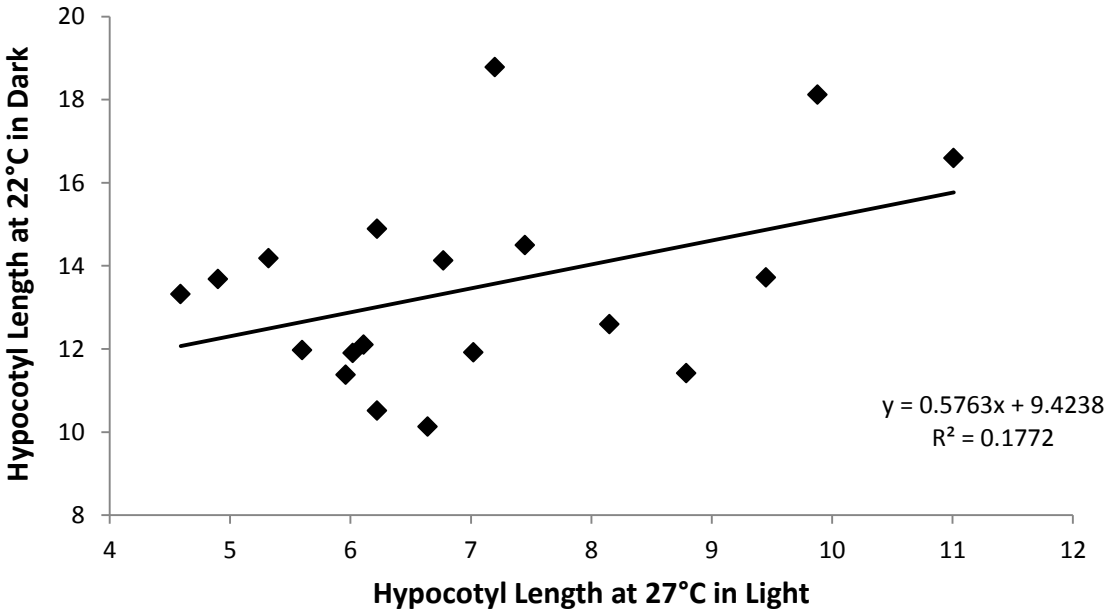

Figure 2.7: Scatter plot of hypocotyl length grown in light at 27°C vs. dark at 22°C. The Coefficient of Correlation r = 0.42, measures the weak to medium strength and positive direction of the linear relationship.

2.3.4 Temperature Responsiveness in Hypocotyl Length:

The variation in hypocotyl length that was seen in figures 2.1 – 2.4 was temperature specific. To study if temperature interacts with the genotype to determine hypocotyl length, the temperature responsiveness of the accessions was examined. This analysis also aims at studying the degree to which MAGIC parents respond to increased changes in

ambient temperature. The phenotyping at multiple temperatures provided a better understanding of the temperature responsiveness concept. This allowed the calculation of the ratios between the means of the phenotypic data for every parental accession at any two temperature treatments, revealing detailed interactions of temperature with the accession.

The temperature responsiveness reaction norms of the 19 MAGIC accessions have been plotted in figure 2.8. These reaction norms define the level of plasticity of hypocotyl length and portray its responses to temperature changes. This has also been quantified in the table 2.1. It illustrates the nature of differential responses to temperature treatments amongst the genotypes; and aids in interpreting the genotype by environment (G x E) interaction that has been statistically explained in detail by a two-way ANOVA in section 2.3.3.

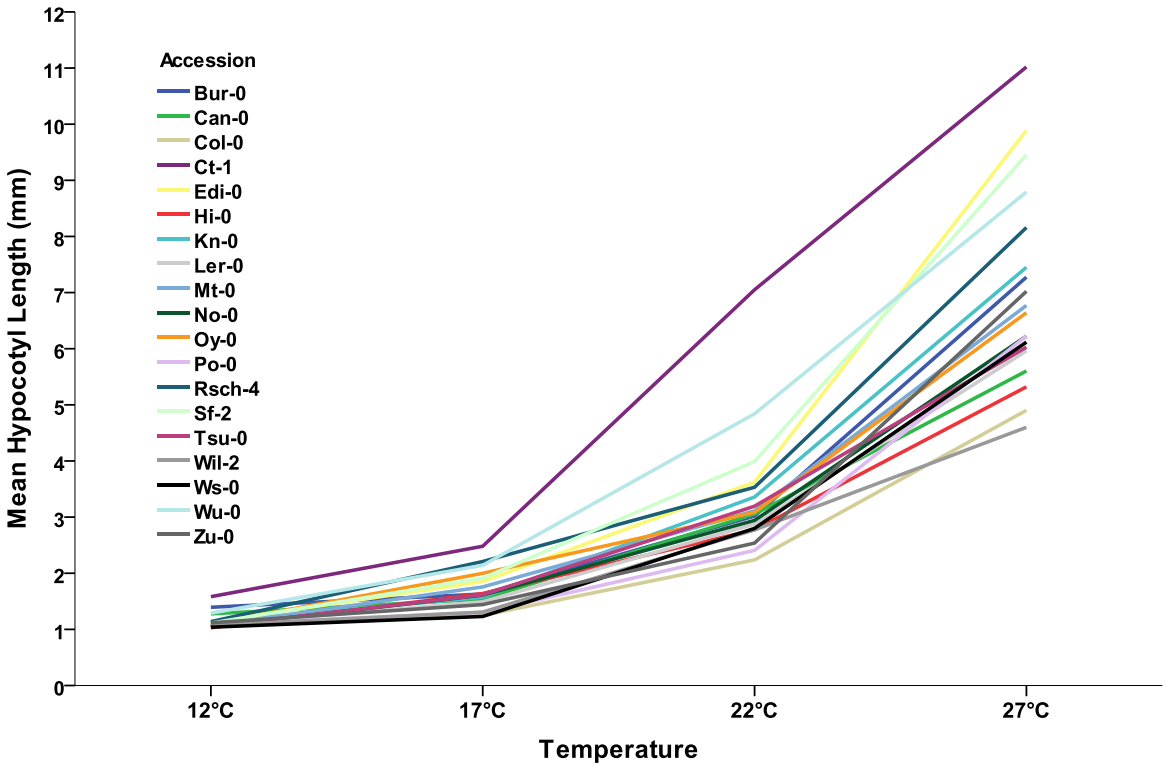

Figure 2.8: Mean hypocotyl lengths of the accessions plotted against four temperatures highlighting their specific temperature response reaction norms. Each line represents a different accession.

Figure 2.8 hints at two major areas of Arabidopsis responses to temperature. One is the interaction of the various genotypes with the temperature environments. This is

explained by the individual response lines for the accessions across the temperatures. The differential responses elucidate the variable interactions of the genotypes with temperature.

The second response is obtained from the changes in the ranking of the genotypes with temperature treatments. Accessions change their ranking across the temperature gradient and show a considerable amount of variation revealing that they are highly responsive to ambient temperature changes. These differential responses are genetically determined. The non-parallel nature of the response lines shows that temperature is interacting with the genotypes in determining the differential responses. The differences in their responses across the temperatures determine their individual ranking at each of the temperatures. Some accessions are more responsive at one temperature than others. However this hierarchy differs between the various temperature ranges. Each of the genotypes has its distinct reaction norm which describes its pattern of phenotypic expression across the temperature gradient. The relative performance of the various accessions varies from one temperature environment to the other.

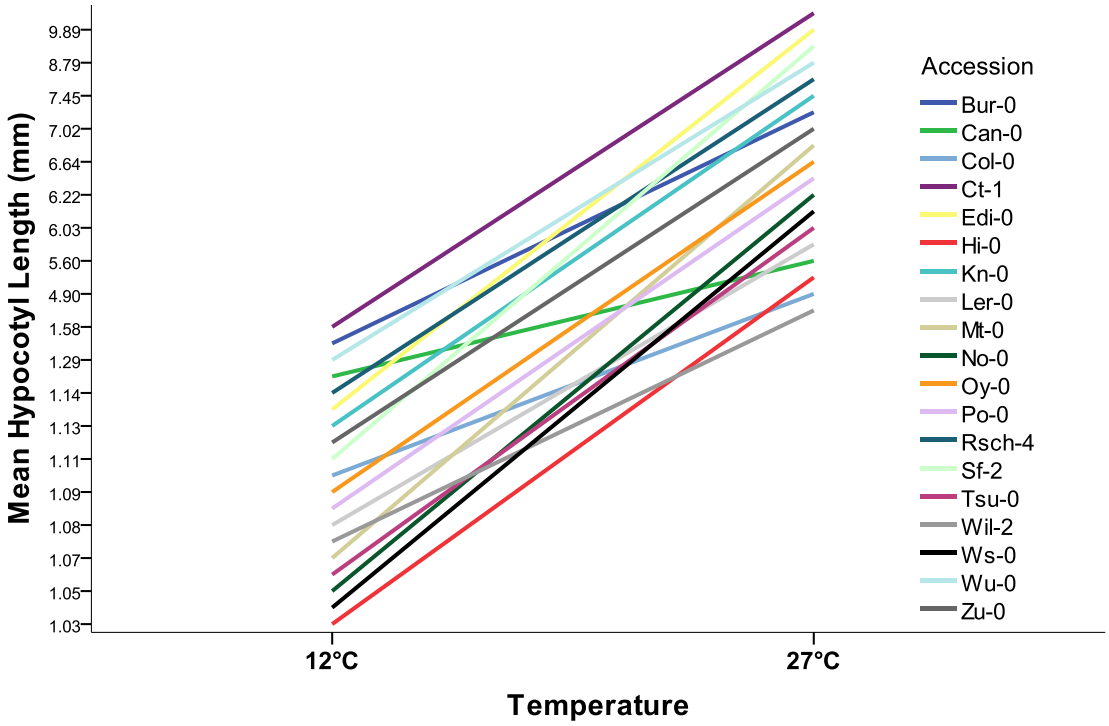

Figure 2.9: Responsiveness reaction norms of the MAGIC parents to a change of 15°C.

The MAGIC parents show an impressive degree of variation in responsiveness to the two extremes of ambient temperatures which has been plotted in figure 2.9. The temperature responses of the accessions are compared to a change in 15°C. The accessions are not only interacting with temperature but are also changing their ranking. This change in ranking pinpoints to accessions being either more or less responsive.

In order to determine the temperature responsiveness in the accessions across the range of temperatures, the relative responsiveness index (RRI) was calculated which is defined as the ratio of hypocotyl length at any one given temperature to hypocotyl length at any other given temperature. For the four temperature treatments, there are six possible RRIs. Table 2.1 quantifies phenotypic plasticity and temperature responsiveness; and identifies the accessions that show high and low phenotypic plasticity.

Accession	**Relative Responsiveness Index (RRI)**					
	27°C/22°C	**27°C/17°C**	**27°C/12°C**	**22°C/17°C**	**22°C/12°C**	**17°C/12°C**
Bur-0	2.41	4.46	5.22	1.85	2.17	1.17
Can-0	1.83	3.64	4.39	1.99	2.39	1.20
Col-0	2.66	3.93	4.43	1.47	1.66	1.13
Ct-1	1.56	4.44	6.96	2.84	4.46	1.57
Edi-0	2.73	5.34	8.74	1.96	3.2	1.64
Hi-0	1.92	3.25	5.18	1.69	2.7	1.59
Kn-0	2.22	4.74	6.6	2.14	2.97	1.39
Ler-0	2.07	4	5.52	1.93	2.66	1.38
Mt-0	2.18	3.86	6.33	1.77	2.91	1.64
No-0	2.12	3.86	5.94	1.83	2.81	1.54
Oy-0	2.15	3.32	6.07	1.54	2.82	1.83
Po-0	2.59	4.79	5.71	1.85	2.21	1.19
Rsch-4	2.31	3.69	7.17	1.6	3.1	1.94
Sf-2	2.37	4.95	8.51	2.09	3.59	1.72
Tsu-0	1.89	3.73	5.74	1.98	3.04	1.54
Wil-2	1.66	3.51	4.3	2.12	2.59	1.22
Ws-0	2.19	4.97	5.89	2.28	2.69	1.18
Wu-0	1.82	4.11	6.82	2.26	3.75	1.66
Zu-0	2.77	4.87	6.29	1.76	2.27	1.29

Table 2.1: Table of RRI for MAGIC parents identifying temperature responses quantitatively.

From table 2.1, we can see that Ct-1 was the least responsive and Zu-0 was the most at RRI of 27°C/22°C. Hi-0 was the least responsive and Edi-0 was the most at RRI of 27°C/17°C. Wil-2 was the least responsive and Edi-0 was the most at RRI of 27°C/12°C. Oy-0 was the least responsive and Ct-1 was the most at RRI of 22°C/17°C. Col-0 was the least responsive and Ct-1 was the most at RRI of 22°C/12°C. Col-0 was the least responsive and Rsch-4 was the most at RRI of 17°C/12°C. The range of variation across all the RRIs reveals the presence of substantial genetic variation in plasticity in MAGIC accessions.

By calculating the means of each of the RRIs across all of the accessions, the overall responsiveness of the accessions can be calculated across the different temperature ranges. This analysis shows that the overall increased responsiveness of the accessions is in the following order: 17°C/12°C = 1.46 < 17°C/22°C = 1.94 < 27°C/22°C = 2.18 < 22°C/12°C = 2.84 < 27°C/17°C = 4.18 < 27°C/12°C = 6.09. The accessions collectively show least responsiveness to a temperature change from 12°C to 17°C, whereas most responsiveness from 12°C to 27°C.

2.3.5 Natural Variation in Hypocotyl Length Temperature Responsiveness:

To determine the components that explained the variation in the hypocotyl length responsiveness and whether Arabidopsis accessions responded to a range of temperature treatments, the G x E interactions of hypocotyl length across the temperature range were evaluated. Ambient temperature may contribute to the hypocotyl length variation in Arabidopsis accessions but this effect may differ across the various accessions. The null hypothesis is that the treatment of temperature has no effect on the hypocotyl length across the temperature range. A two-way ANOVA tested for natural variation in hypocotyl lengths of the MAGIC accessions in response to the ambient temperatures of 12°C, 17°C, 22°C, and 27°C. It tested the main effects of temperature, the main effects of accession and the interaction between the two. Ambient temperature showed a statistically significant effect on hypocotyl length $F(3,3024) = 77233.9$, $p < 0.001$, $\eta^2 = 0.987$. The 19 parental accessions also showed statistically significant effects on the hypocotyl elongation $F(18,3024) = 1443.8$, $p < 0.001$, $\eta^2 = 0.896$. The interaction effects of

accessions and temperature (G x E) were also significant $F(54,3024) = 391.2$, $p < 0.001$, $\eta^2 = 0.875$). Remarkable natural variation in temperature-dependent responsiveness was observed.

From figure 2.8, it is expected that a positive association exists between temperature and hypocotyl length; when temperature is increased, hypocotyl length also increases. The Pearson's correlation r = 0.97 shows a very strong positive correlation between temperature and hypocotyl length and it is highly significant $F(1,3099) = 24566.42$, $p < 0.001$, N = 3100 and adjusted R^2 value = 0.88. Hence 88% of the variability in hypocotyl length can be accounted for by temperature.

To predict the hypocotyl length elongation with an increase in 1°C, a linear regression analysis was performed. The results demonstrated that with an increase of 1°C, hypocotyl length increased by 12%. This was a substantial change in the hypocotyl elongation in response to temperature. It identified temperature as a very meaningful predictor accounting for a large amount of variation in hypocotyl length in Arabidopsis accessions.

2.3.6 Higher Temperatures Determine Higher Variation in Hypocotyl Length:

To determine the most discriminating temperature where maximum phenotypic variation is observed in the MAGIC parents, the phenotypic data across the four temperatures were analysed comparatively. This analysis showed that the total amount of phenotypic variation is directly proportional to temperature; the higher the temperature, the greater the variation. Figure 2.10 shows the four temperatures hosting the phenotypic variation of the MAGIC parents. The descriptive statistical data in table 2.2 illustrate that the, standard deviation, standard error and variances increase with increase in temperature. This reveals that temperature is a major determinant in identifying phenotypic variation in MAGIC parents.

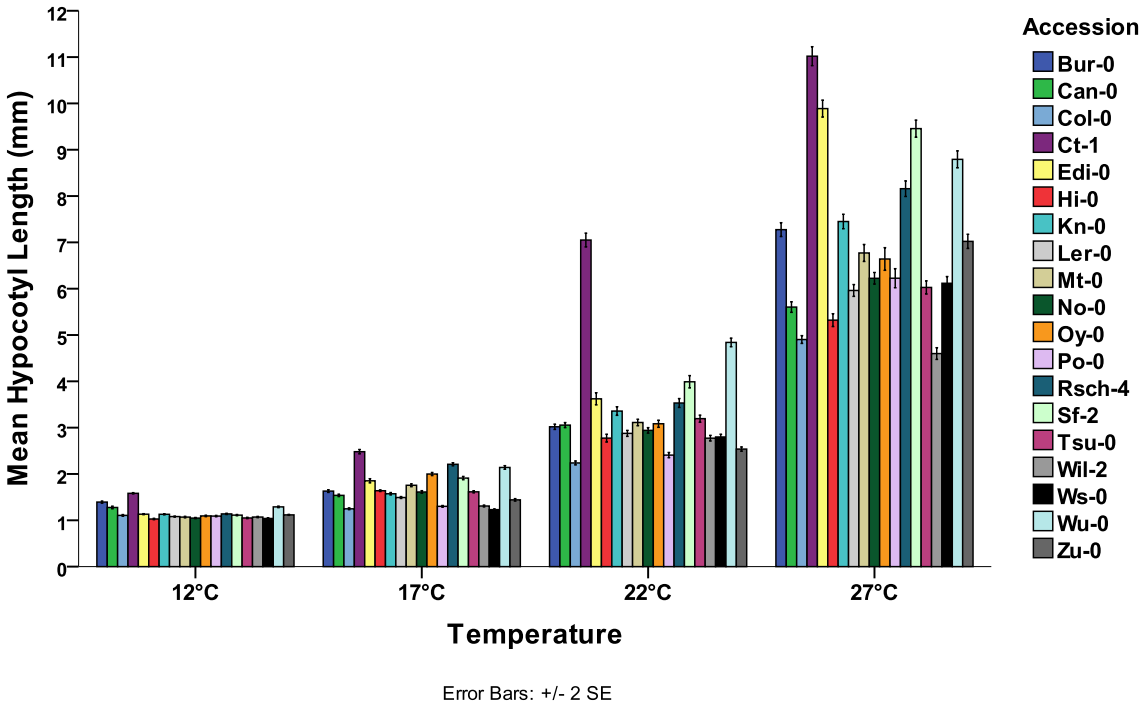

Figure 2.10: Phenotypic variation in the 19 MAGIC parents. Variation in accessions is temperature dependent and it increases with increase in temperature. This figure graphically corresponds to table 2.2 for the four different temperature groups.

Statistic	12°C	17°C	22°C	27°C
Coefficient of Variation	0.12	0.20	0.33	0.24
Standard Deviation	0.14	0.34	1.10	1.74
Standard Error	0.03	0.07	0.25	0.40
Variance	0.01	0.11	1.21	3.02

Table 2.2: Coefficient of variation, standard deviation, standard error and variance calculated from the means of hypocotyl lengths of the accessions at 12°C, 17°C, 22°C and 27°C. The greatest variation between the accessions is observed at 27°C.

2.3.7 Arabidopsis Hypocotyls are Longer when Phenotyped in Short Days:

To determine if there were any significant differences in hypocotyl length between short days and long days, Arabidopsis MAGIC parents were phenotyped separately in these conditions at 22°C. 19 independent-samples t-tests were performed for the 19 parents

separately between long days and short days. For each genotype, the means of the hypocotyl length grown under long days were compared to hypocotyl length grown under short days. All the tests showed that there were significant differences ($p < 0.001$) in hypocotyl length. The photoperiod duration affects hypocotyl length and as can be seen in figure 2.11, hypocotyls grown under short days are overall bigger than long days.

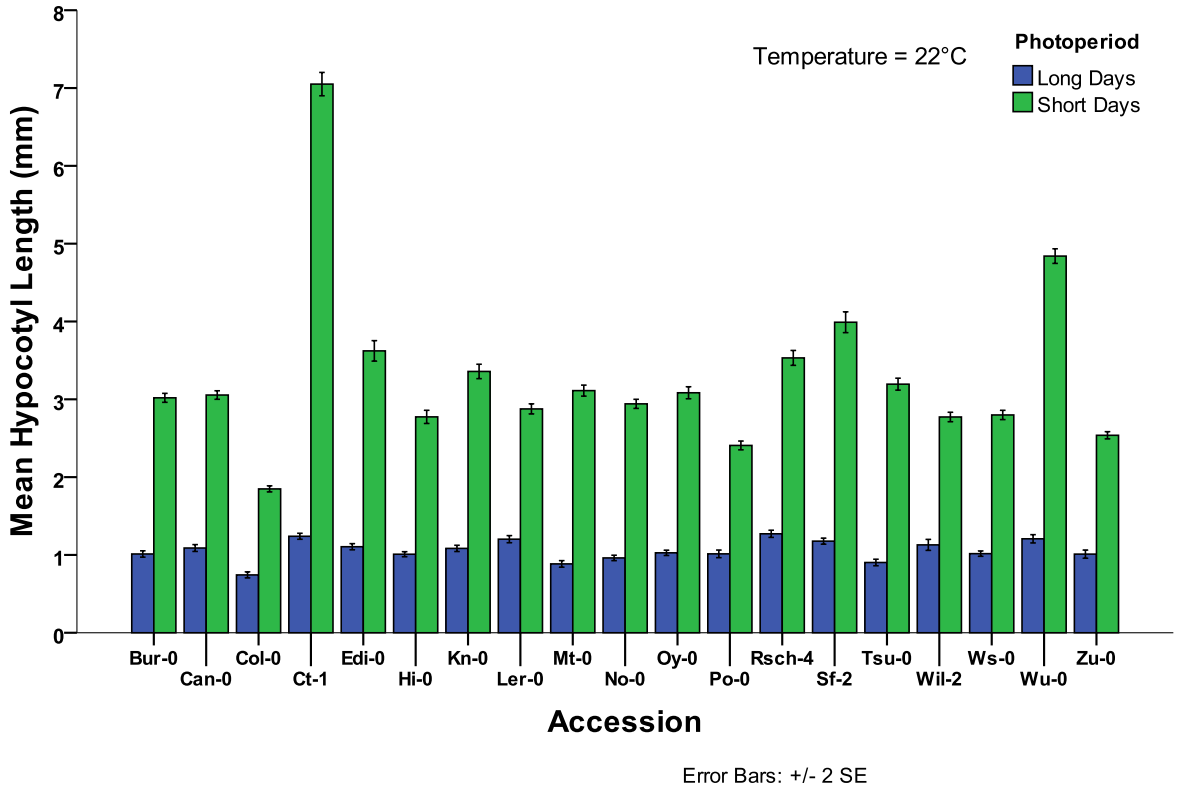

Figure 2.11: Hypocotyl lengths are bigger when phenotyped in short days. In long days they are smaller due to the greater de-etiolation effects. Arabidopsis hypocotyls exhibit higher natural variation when grown for 1 week in 170 $\mu E/m^2/sec$ short days (8 hours daylight) as compared to long days (16 hours daylight).

2.3.8 Natural Variation in Hypocotyl Length is Present in Short Days and Long Days:

Arabidopsis MAGIC parents may differ in their hypocotyl lengths when grown in short day and long day photoperiods at an ambient temperature of 22°C. To test this, two one-way

ANOVAs for 'tests of between-subjects effects' were performed separately for short day phenotype data and long day phenotype data for all the 19 accessions. The hypocotyl length data was either normally distributed or the sample sizes were > 30. The results for both conditions demonstrated that the mean hypocotyl lengths of Arabidopsis accessions are statistically significant ($p < 0.001$). Short days = $p < 0.001$; 22°C = $F(18,716) = 638.2$. Long days = $p < 0.001$; 22°C = $F(18,438) = 36.92$. Natural variation is present in the hypocotyl length when grown in short days and long days.

2.3.9 Natural Variation in Hypocotyl Length is Greater in Short Days:

To determine which photoperiod hosts more of the variation, and if the effect of the genotype on hypocotyl length is big or small in short days or long days, the effect sizes were estimated. ANOVA provided the Partial Eta squared (η^2) values which estimate effect size; and the genotypes account for 94% of the variance in hypocotyl length in short days, whereas in long days the genotypes account for 60% of the variance in hypocotyl length.

To establish which of the two photoperiod conditions enables the Arabidopsis seedlings to show maximum variation in the hypocotyls, the phenotypic data were analysed comparatively. The descriptive statistical data**,** shown in table 2.3, was calculated for both of the photoperiods. The findings demonstrate that in long days, the coefficient of variation, standard deviation, standard error and variances are less as compared to short days. Therefore, the variation between the accessions is greater in short days than long days. This illustrates the effect of photoperiod on the presence of natural variation in Arabidopsis accessions and is a key determinant in identifying phenotypic variation in seedlings.

Statistic	Long Days	Short Days
Coefficient of Variation	0.125	0.333
Standard Deviation	0.133	1.104
Standard Error	0.031	0.254
Variance	0.018	1.218

Table 2.3: Standard deviation, standard error and variance calculated from the means of hypocotyl lengths of the accessions grown in short days and long days. The greatest variation between the accessions is observed in short days.

Accession	NASC ID	Geographical Location	Altitude in Meters	Latitude	Habitat	Average Air Temperature
Bur-0	N6643	Burren, Ireland	1-100	N 54	Temperate, Roadside	Spring: 4°C Summer: 15°C
Can-0	N6660	Las Palmas, Canary Islands, Spain	1200-1300	N 28	Sub-Tropical	Spring: 6°C Summer: 25°C
Col-0	N6673	Gorzow Wielkopolski, Poland	1-100	N 52	Unknown	Spring: 1-7°C Summer: 20-22°C Autumn: 9-19°C
Ct-1	N6674	Catania, Italy	1-100	N 37	Unknown	Spring: 13-14°C Summer: 25°C Autumn: 21-22°C
Edi-0	N6688	Edinburgh, UK	1-100	N 56	Botanic Gardens	Spring: 3-4°C Summer: 16°C Autumn: 9-10°C
Hi-0	N6736	Hilversum, Netherlands	1-100	N 52	Unknown	Spring: 5-6°C Summer: 16°C Autumn: 9-10°C
Kn-0	N6762	Kaunas, Lithuania	1-100	N 54	Unknown	Spring: 3-4°C Summer: 16°C Autumn: 9-10°C
Ler-0	NW20	Poland	1-100	N 53	Unknown	Spring: 1-7°C
Mt-0	N1380	Martuba/Cyrenaika, Libya	100-200	N 33	Unknown	Spring: 15-16°C Summer: 25°C Autumn: 15-16°C
No-0	N6805	Nossen, Germany	200-300	N 51	Unknown	Spring: 2°C Summer: 17°C
Oy-0	N6824	Oystese, Norway	1-100	N 60	Unknown	Spring: 3-4°C Summer: 11°C Autumn: <5-6°C
Po-0	N6839	Poppelsdorf, Germany	1-100	N 50	Dry, Sandy Soil	Spring: 7-8°C Summer: 17°C Autumn: 11-12°C
Rsch-4	N6850	Rschew, Russia	100-200	N 56	Unknown	Spring: <0-2°C Summer: 16°C Autumn: 7-8°C
Sf-2	N6857	San Feliu, Spain	1-100	N 41	Unknown	Spring: 11-12°C Autumn: 19-20°C
Tsu-0	N6874	Tsu, Japan	1-100	N 34	Unknown	Spring: 9-10°C Summer: 22°C Autumn: 19-20°C
Wil-2	N6889	Wilna/litvanian Ssr, Russia	100-200	N 55	Near Towniskaya	Spring): <0-2°C Autumn: 9-10°C
Ws-0	N6891	Wassilewskija, Belarus	100-200	N 52	Sandy Ryefield	Spring: 3-4°C Summer: 17°C Autumn: <5-6°C
Wu-0	N6897	Wurzburg, Germany	100-200	N 49	Sandy Soil	Spring: 5-6°C Summer: 16°C Autumn: 11-12°C
Zu-0	N6902	Zurich, Switzerland	500-600	N 47	Botanic Gardens	Spring: 5-6°C Summer: 17°C Autumn: 9-10°C

Table 2.4: Arabidopsis MAGIC parental accessions used for analysis of natural variation in hypocotyl length in response to temperature. Climatological and Geographical data were obtained from the websites of NASC (http://arabidopsis.info/EcoForm) and (http://arabidopsis.info/info/annualreports/ais24.pdf); and Versailles Biological Resource Centre (http://dbsgap.versailles.inra.fr/vnat/).

2.3.10 Correlation between Hypocotyl Length and Habitat Temperature Clines:

As seen in table 2.4, the different geographic habitats represent varying ambient temperatures. Therefore, it was interesting to analyse hypocotyl length of the MAGIC parents to their respective habitat temperatures which could provide some basic insights, from an ecological perspective, to the study of adaptations of these accessions to their native habitats. To examine if there was any association between the hypocotyl length and the native habitat temperature, a correlation analysis was performed. The coefficient of correlation, r = 0.7, measures a medium to high strength. In figure 2.12 the line of best fit numerically describes the linear relationship between the hypocotyl length and the habitat temperature; where temperature is the explanatory variable and hypocotyl length is the response variable. The coefficient of determination, R^2 is 0.46 and therefore, it is more likely that a linear relationship exists between the two variables. It is assumed that we can model hypocotyl length on habitat temperature with a linear relationship.

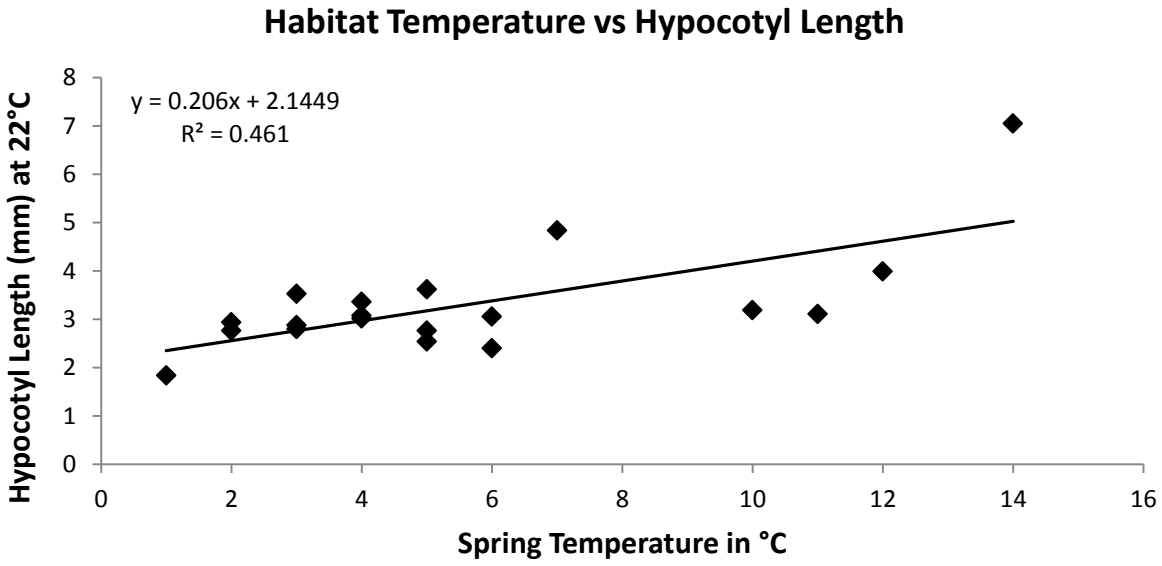

Figure 2.12: Scatter plot of hypocotyl length vs. habitat spring temperature.

To determine if this identified linear relationship exists in the population from which the samples of accessions were drawn, a significance test showed that and there is a

significant relationship between hypocotyl length and habitat temperature, r (17) = 0.7, p < 0.05.

The R^2 in this analysis shows that 46% of the variance in hypocotyl length can be predicted and accounted for by the temperature clines in the native habitats. The gradual change in the inherited trait of hypocotyl length across the geographic range is positively correlated with temperature in the native habitats. The hypocotyl lengths tend to be bigger at native geographic regions where mean temperatures are higher, and this correlation is consistent with the finding in figure 2.10 that at higher temperatures, the hypocotyls are longer.

2.3.11 Correlation between Hypocotyl Length and Latitudinal Clines:

A correlation analysis was performed to study the linear relationship between hypocotyl length and the native latitudes of the MAGIC accessions. In figure 2.13 the results show that r = - 0.3 and there is a weak negative correlation between hypocotyl length and native latitude. The significance test showed that there is not a statistically significant relationship between the hypocotyl length and native latitude, r (17) = - 0.3, p > 0.05.

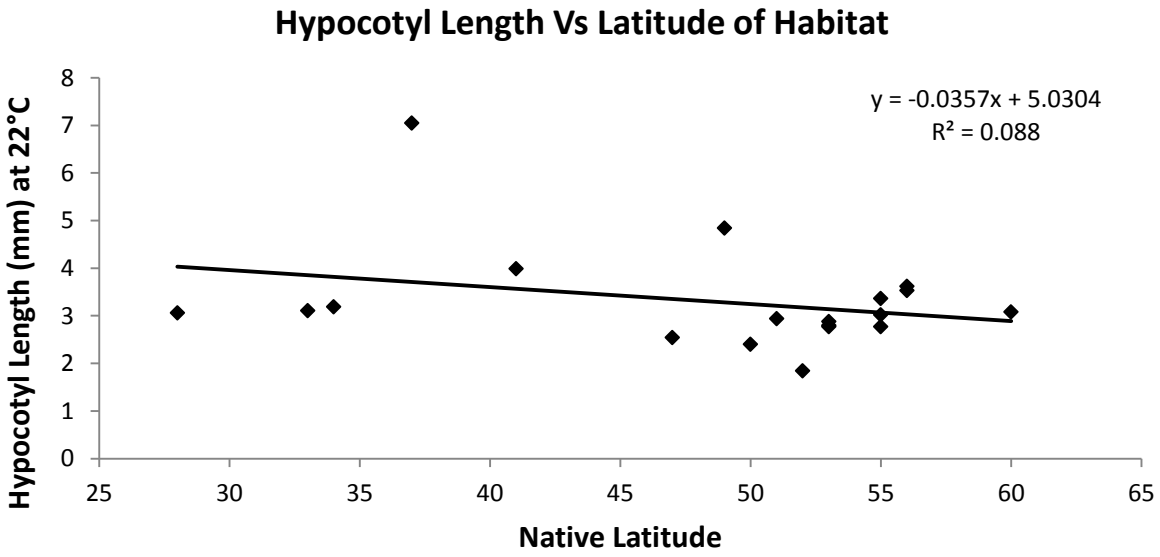

Figure 2.13: Scatter plot of hypocotyl length vs. native latitude.

2.3.12 Correlation between Etiolated Hypocotyl Length and Latitudinal Clines:

It was also of interest to figure out if there is any correlation between hypocotyl lengths grown in the dark with latitudinal clines. From the correlation analysis, the value $r = 0.05$ shows almost no correlation at all between etiolated hypocotyl length and habitat temperature. The coefficient of determination, R^2 is 0.0023 and therefore, it is more likely that no linear relationship exists between the two and no variation in the hypocotyl length can be accounted for by latitude (figure 2.14). The significance test showed that there is not a statistically significant relationship between the etiolated hypocotyl length and native latitude, $r\ (17) = 0.05$, $p > 0.05$.

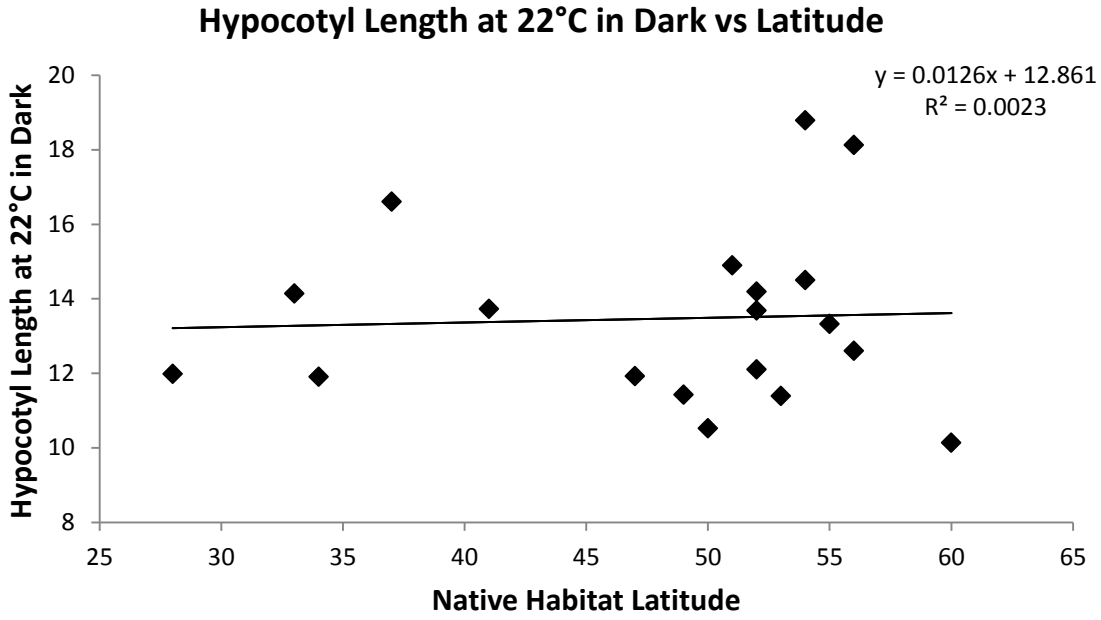

Figure 2.14: Scatter plot of etiolated hypocotyl length vs. native habitat latitude. The coefficient of correlation, $r = 0.05$, measures almost no strength of the linear relationship.

2.3.13 Natural Variation and Phenotypic Analysis of Plant Size:

To investigate if natural variation in the plant size trait is present in the MAGIC accessions and whether this trait showed any responsiveness to changes in ambient temperature, the MAGIC parents were phenotyped at 17°C and 22°C.

Arabidopsis MAGIC parents may differ in their plant sizes when phenotyped at ambient temperatures of 17°C and 22°C. To test this, a one-way ANOVA was performed at each of the individual temperatures separately. The ANOVA results revealed that the mean plant sizes of Arabidopsis accessions are significantly different ($p < 0.001$) at each of the 17°C and 22°C temperatures. 17°C = $F(18,169) = 24.44$, $p < 0.001$; 22°C = $F(18,156) = 28.95$, $p < 0.001$. The phenotypic variation in plant sizes between the accessions at 17°C and 22°C can be seen in figures 2.15 and 2.16 respectively. These data show that there is greater phenotypic variation in the mean plant area amongst the accessions at 22°C (SD = 0.81) as compared to 17°C (SD = 0.64).

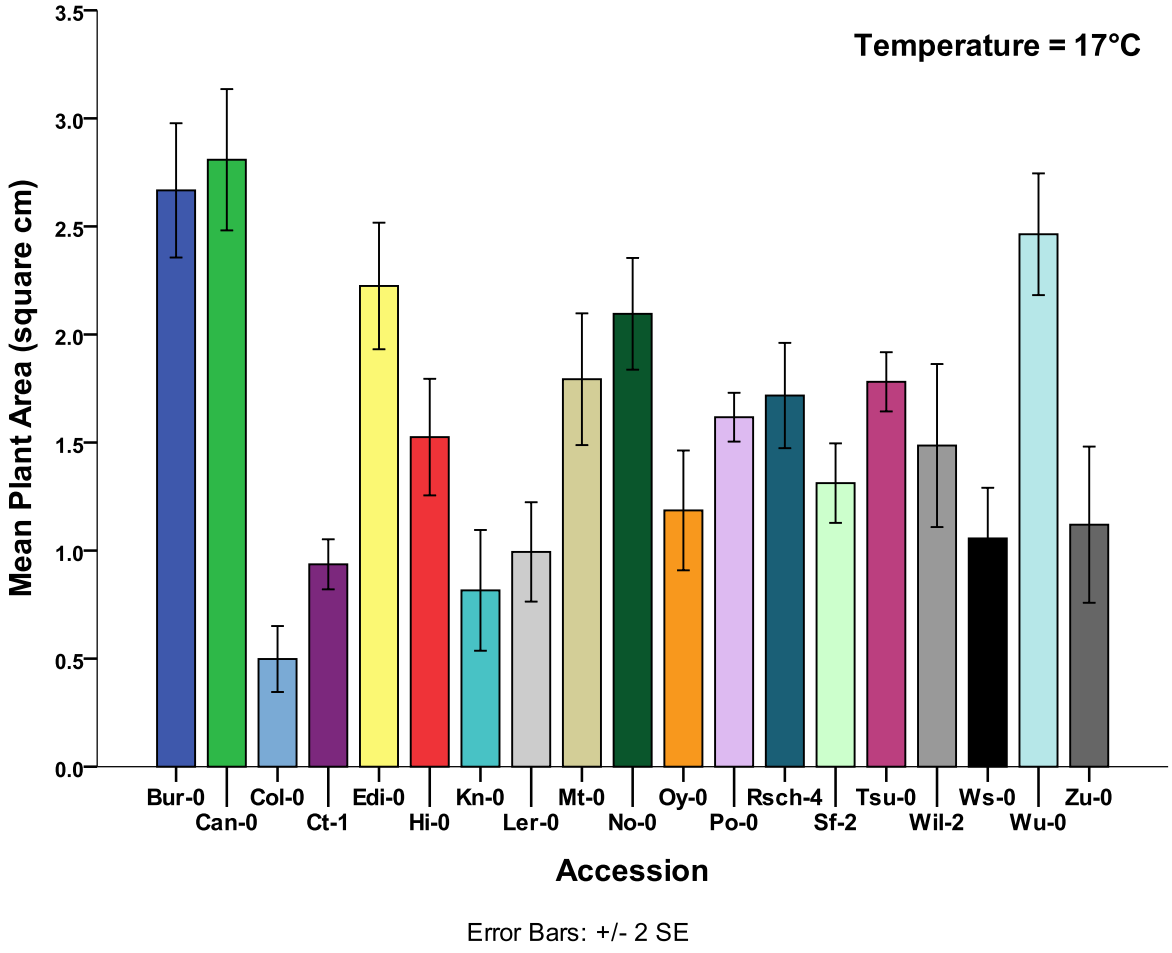

Figure 2.15: Natural phenotypic variation in plant size in MAGIC accessions at 17°C after four weeks of growth in 170 µE/m^2/sec short days (8 hours light/16 hours dark).

Further to this finding, in order to establish if the effect of the genotypes on the plant sizes is big or small at each of the two temperatures, the effect sizes were estimated by Partial Eta squared (η^2). The genotypes account for 72% and 77% of the variance in plant size at 17°C and 22°C temperatures respectively. Notable natural variation in temperature-dependent plant size responses was observed; and similar to the hypocotyl length trait, the variation increases with increase in temperature. These results provided a basis to further explore the G x E interaction by a two-way ANOVA for temperature responsiveness.

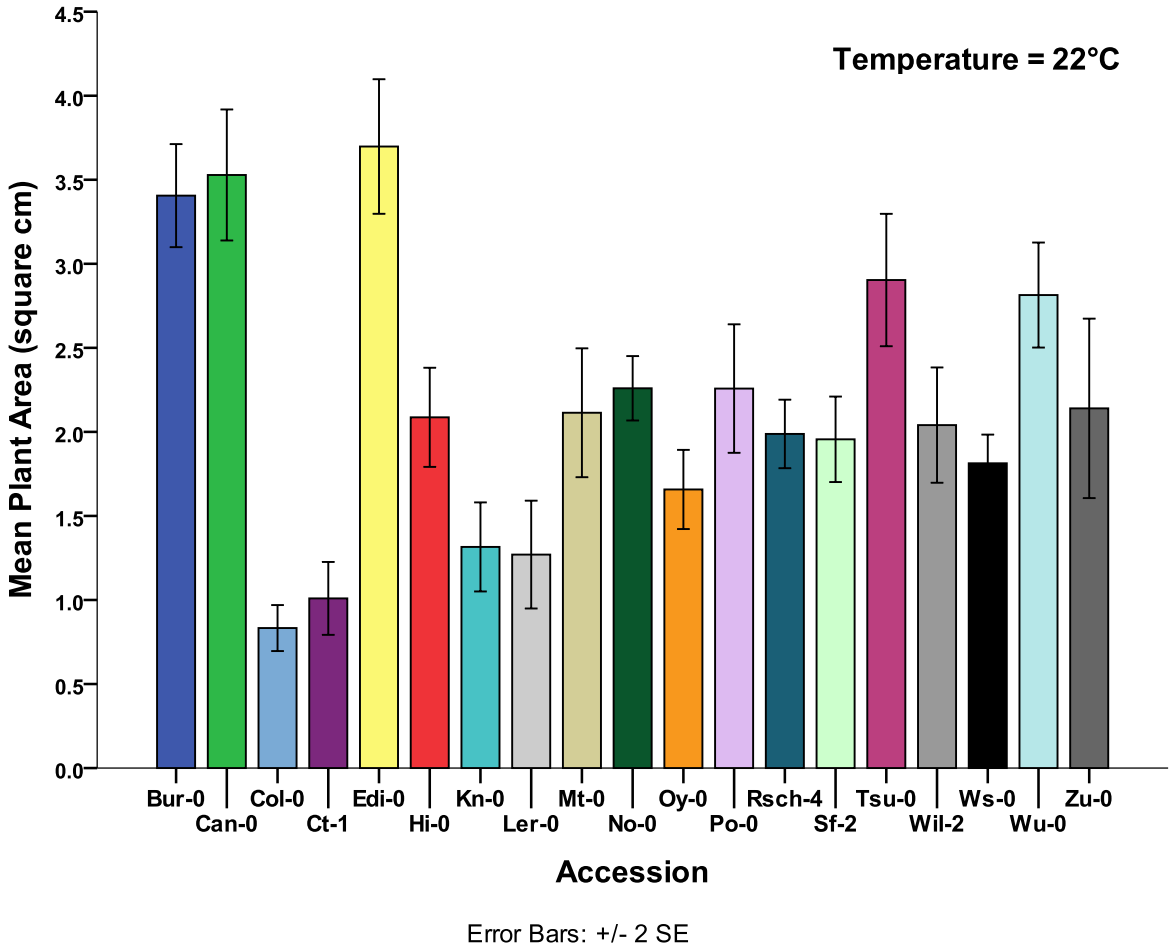

Figure 2.16: Natural phenotypic variation in plant size in MAGIC accessions at 22°C after four weeks of growth in 170 $\mu E/m^2/sec$ short days (8 hours light/16 hours dark).

To determine if there were any significant differences between the mean plant size at 17°C and 22°C for any given accession, 19 independent-samples t-tests were performed on the 19 parents separately. For each accession, the means of the plant size at 17°C were compared to plant size at 22°C. Accessions Ct-1, Ler-0, Mt-0, No-0, Rsch-4, and Wu-0 did not show any significant differences in plant size between the two temperature treatments (figure 2.17). However, it is interesting that all of these accessions showed significant differences in hypocotyl length between the same two temperature treatments.

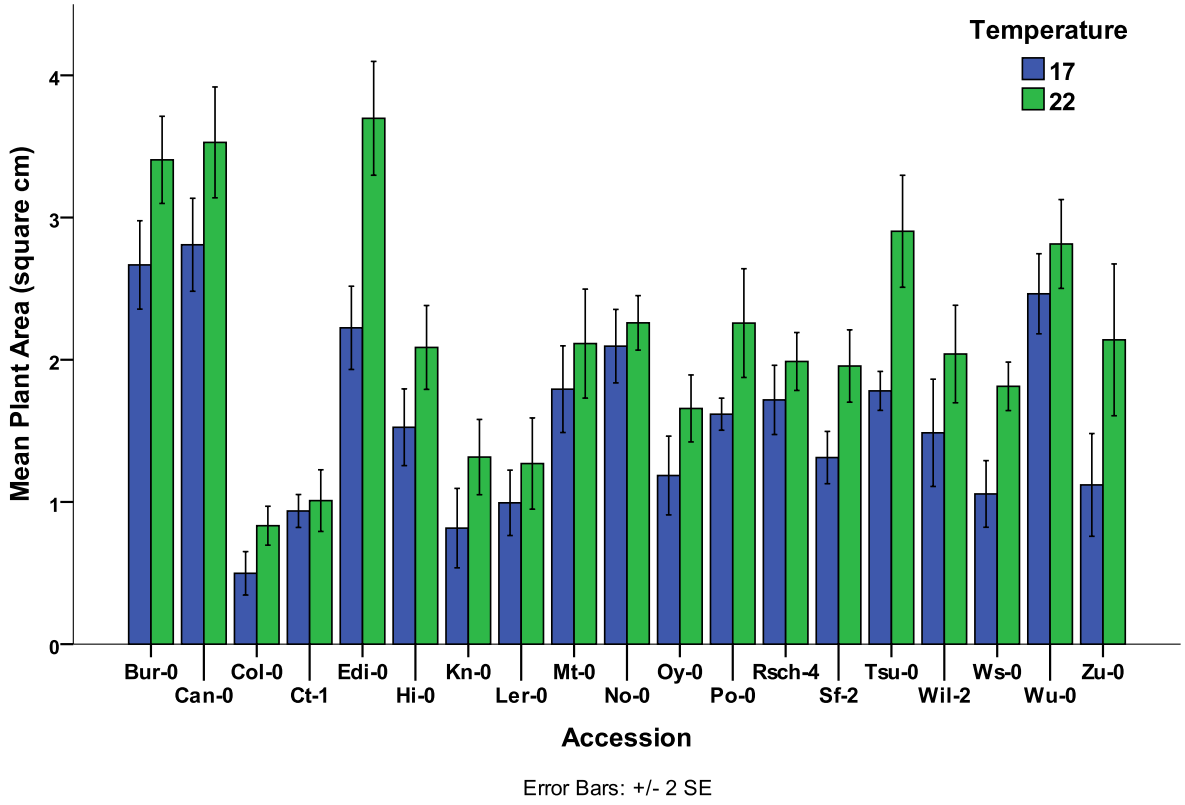

Figure 2.17: Comparison of the 19 MAGIC accessions for plant size at 17°C and 22°C.

2.3.14 Temperature Responsiveness in Plant Size:

In order to study the temperature responsiveness of the plant size trait, the mean phenotypic values of the accessions at 17°C were compared to 22°C. Ambient temperature, may contribute to the plant size variation in Arabidopsis accessions but this

effect may differ across the various accessions. A two-way ANOVA tested for natural variation in plant size temperature responsiveness and compared the plastic responses of the 19 genotypes to 17°C and 22°C. Ambient temperature showed a statistically significant effect on plant size $F(1,325) = 160.5$, $p < 0.001$, $\eta^2 = 0.331$. The genotypes also showed statistically significant effects on plant size $F(18,325) = 51.64$, $p < 0.001$, $\eta^2 = 0.741$. The G x E interaction effects were also significant $F(18,325) = 2.99$, $p < 0.001$, $\eta^2 = 0.142$).

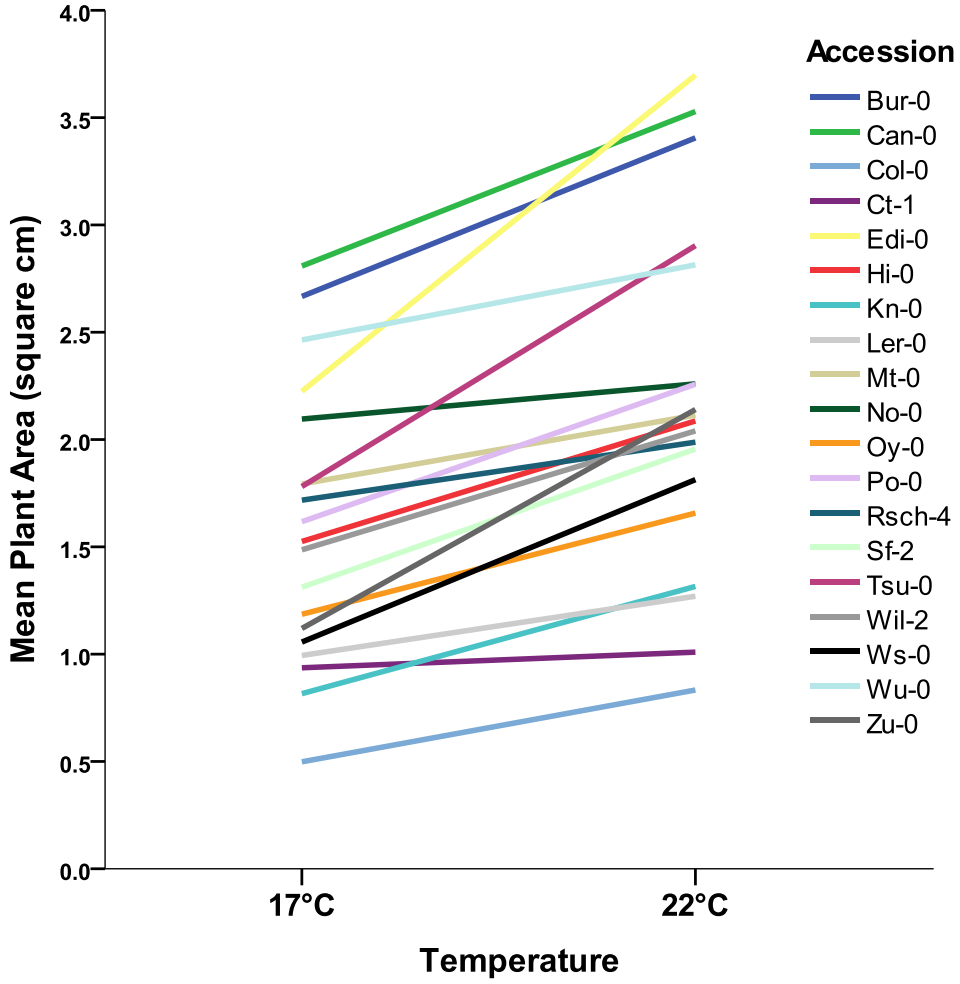

Figure 2.18: Reaction norms of MAGIC accessions for phenotypic plasticity of plant size across the two temperatures of 17°C and 22°C. Accessions show temperature responsiveness in their plant size.

The results of a significant environment term reveal that the genotypes express plasticity by responding to the temperature treatments. The 17°C means for the accessions are lower than 22°C (figure 2.15 and 2.16); hence there is a main effect for temperature. There is a main effect for genotype as variation in the mean plant size is observed. Additionally, a significant G x E term was also found indicating that the Arabidopsis genotypes and ambient temperatures are interacting and affecting the plant size. 14% of the variability in plant size is accounted for by the G x E interaction and can be attributed to the interaction. A distinct reaction norm is seen for each of the genotypes which describe the pattern of phenotypic expression of plant size between the two temperature treatments. The differences in the responses of genotypes across the two temperatures determine their individual ranking at each of the temperatures. The ranking of the mean phenotypic values change between the two temperatures (figure 2.18), which identifies the responsiveness of the accessions resulting in G x E interactions.

Accession	RRI 22°C / 17°C
Ct-1	1.08
No-0	1.08
Wu-0	1.14
Rsch-4	1.16
Mt-0	1.18
Can-0	1.26
Bur-0	1.28
Ler-0	1.28
Hi-0	1.37
Wil-2	1.37
Po-0	1.40
Oy-0	1.40
Sf-2	1.49
Kn-0	1.61
Tsu-0	1.63
Edi-0	1.66
Col-0	1.67
Ws-0	1.72
Zu-0	1.91

Table 2.5: RRI of the 19 MAGIC parents showing hierarchy of least to most temperature responsiveness for plant size.

2.3.15 Comparative Analysis of Hypocotyl Length with Plant Size:

The analyses of the plant size trait paved the way for a comparative natural variation analysis between hypocotyl length and plant size traits at 17°C and 22°C. The results, in the case of plant size, show that the effect size of the genotype is greater than the effect size of temperature (genotype $\eta^2 = 0.741$ > environment $\eta^2 = 0.331$). Conversely, in the case of hypocotyl length trait, the effect size of temperature is greater than the effect size of genotype (environment $\eta^2 = 0.948$ > genotype $\eta^2 = 0.917$). The genetic differences between the genotypes account for greater variation in plant size when compared to the effect of temperature; on the contrary, temperature has a bigger effect on the hypocotyl length trait when compared to the genotypic effect.

The comparative analyses show that hypocotyl length, in response to temperature, displays greater phenotypic plasticity than plant size. Moreover, G x E interaction was greater for the hypocotyl length trait than the plant size trait (the hypocotyl length G x E $\eta^2 = 0.793$ > plant size G x E $\eta^2 = 0.142$). This indicates that the hypocotyl length trait is highly responsive to the G x E interaction (> 5 X) than the plant size trait. It is interesting to establish that the effects of temperature, genotype and their interaction are greater on hypocotyl length than plant size. This makes hypocotyl length a more interesting trait to investigate in response to temperature and to find the underlying genes involved in the natural variation of the trait.

The hypocotyl length and plant size are two different growth traits of two different developmental stages in plants. The hypocotyl elongation takes place during seedling development in the juvenile stage, whereas the plant size is determined during the vegetative stage of development. An interesting biological question relating to the two different stages of development in Arabidopsis is to determine if there is any linear relationship between these two independent variables which show a continuous distribution in their phenotypes. In other words; is hypocotyl length correlated to plant size? Does the plant size increase with increase in hypocotyl length and vice versa? To address these questions, the correlation analyses have been carried out separately for 17°C and 22°C; and for the temperature responsiveness between the two traits.

2.3.16 Correlation between Hypocotyl Length and Plant Size at 17°C:

At 17°C the value *r* = 0.16 showed a very weak positive correlation between hypocotyl length and plant size (figure 2.19). The coefficient of determination, R^2 is very near to zero (0.026) and therefore, it is less likely that a linear relationship exists between hypocotyl length and plant size. There is not a statistically significant relationship between hypocotyl length and plant size at 17°C, r (17) = 0.16, $p > 0.05$.

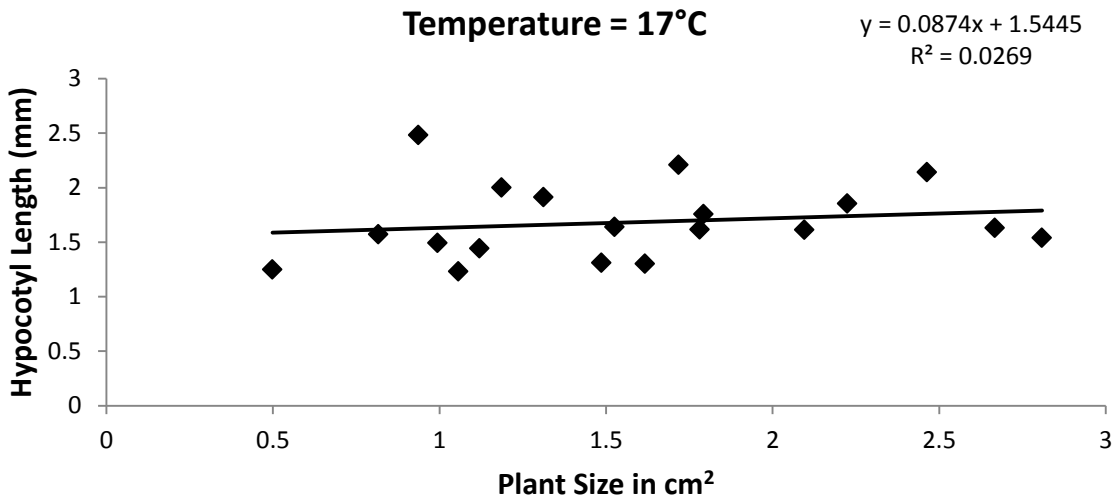

Figure 2.19: Scatter plot of hypocotyl length vs. plant size at 17°C.

2.3.17 Correlation between Hypocotyl Length and Plant Size at 22°C:

At 22°C the value *r* = -0.11 showed a very weak negative correlation between hypocotyl length and plant size. The coefficient of determination, R^2 is almost zero (0.012) and therefore, in a similar way, it is less likely that a linear relationship exists between hypocotyl length and plant size at this temperature (figure 2.20). There is not a statistically significant relationship between hypocotyl length and plant size at 22°C, *r* (17) = -0.11, *p* > 0.05.

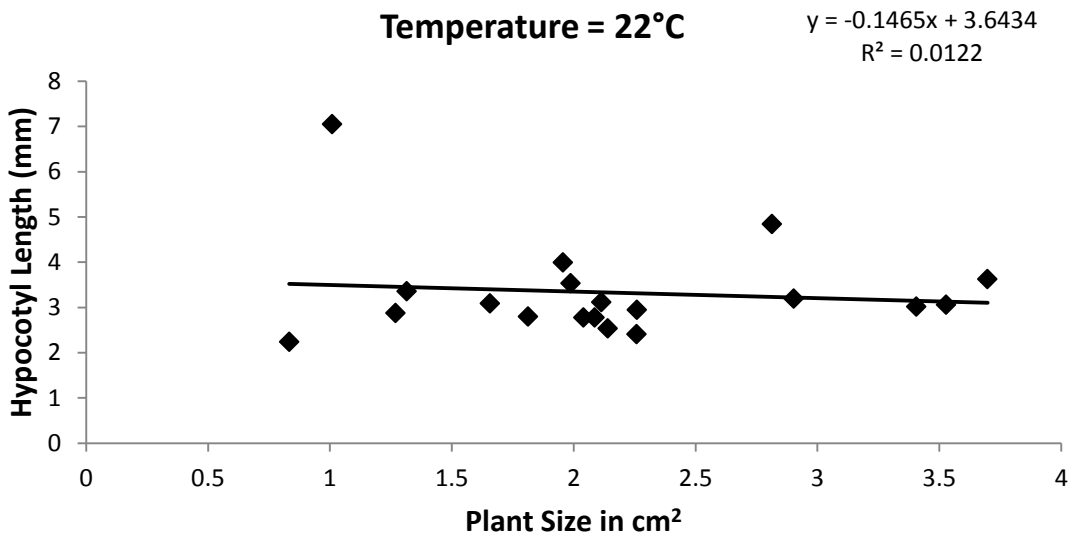

Figure 2.20: Scatter plot of hypocotyl length vs. plant size at 22°C.

2.3.18 Correlation between Hypocotyl Length and Plant Size Responsiveness:

Is there a correlation between temperature responsiveness of hypocotyl length and of plant size? The r = -0.13 showed a very weak negative correlation between hypocotyl length and plant size RRI 22°C/17°C. The coefficient of determination, R^2 is almost zero (0.016) and therefore, it is less likely that a linear relationship exists between hypocotyl length and plant size RRI 22°C/17°C (figure 2.21). There is not a statistically significant relationship,

r (17) = -0.11, *p* > 0.05.

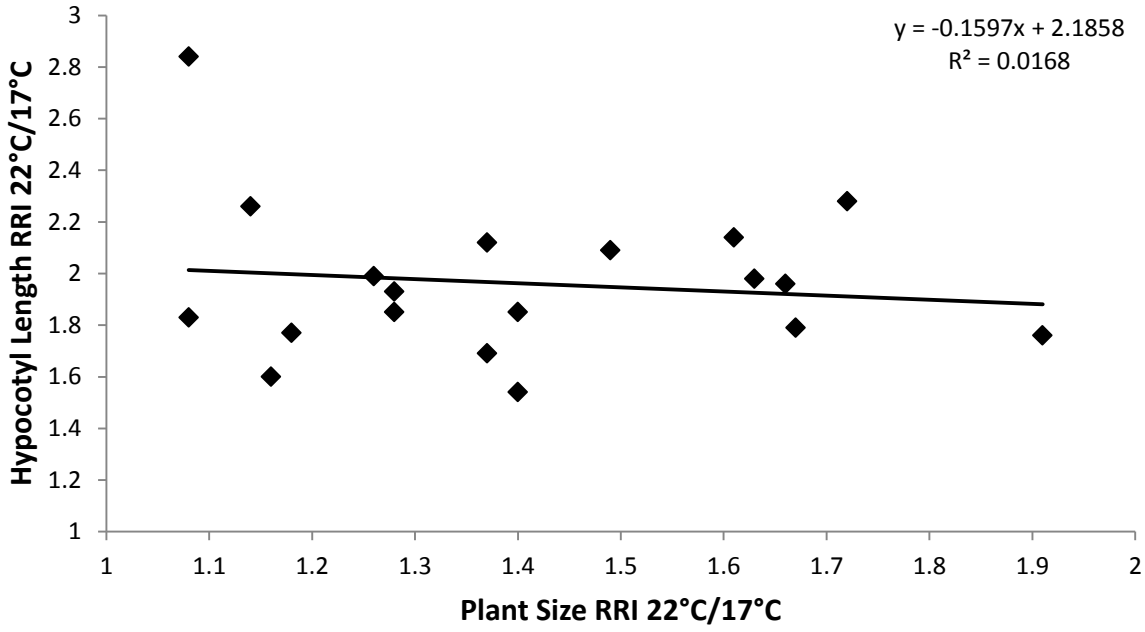

Figure 2.21: Scatter plot of hypocotyl length vs. plant size RRI 22°C/17°C.

2.3.19 Inheritance of Hypocotyl Length as a Quantitative Trait:

To genetically analyse the inheritance pattern of hypocotyl length two genetically distinct accessions were crossed which possess extreme hypocotyl lengths. The MAGIC parental phenotypic analysis showed that Col-0 had shorter and Sf-2 had bigger hypocotyl lengths. The genetic cross between these two accessions allowed the examination of hypocotyl length inheritance both in the F1 and F2 generations. The analysis revealed that the F1 progeny had hypocotyl lengths that were intermediate between those of the Col-0 and Sf-2 parents. The F2 experimental segregating population did not give rise to hypocotyl length distributions that were discrete and hence could not be fitted to monogenic inheritance. The F2 individuals showed a continuous distribution of phenotype as seen in figure 2.23.

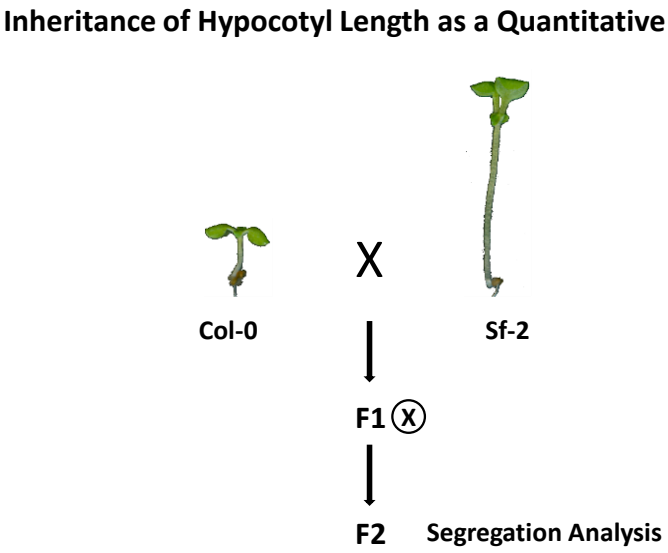

Figure 2.22: The F2 segregation analysis of a cross between two genetically distinct accessions explains the nature of hypocotyl length as a complex trait.

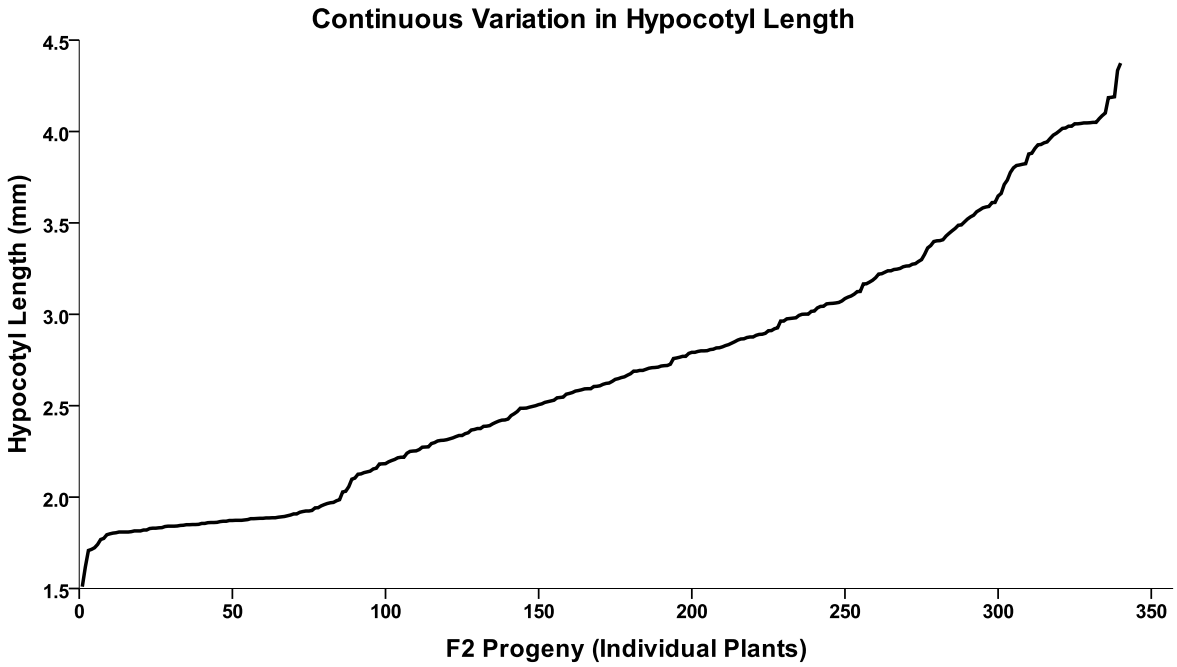

Figure 2.23: Distribution of hypocotyl length in an F2 population of 340 individuals derived from a Col-0 × Sf-2 cross at 22°C.

2.4 Discussion:

2.4.1 Discussion on Hypocotyl Length Natural Variation:

Arabidopsis natural variation in hypocotyl length in response to ambient temperature has not been described previously. Chapter 2 has provided a comprehensive phenotypic analysis of the intraspecific natural variation in an ecologically relevant trait - hypocotyl length in response to temperature. The phenotypic analysis shows that extensive naturally existing variation is present in the 19 MAGIC accessions in the developmental stage of hypocotyl elongation in response to a range of ambient temperatures. This variation at any given temperature shows that there is an underlying genetic basis for this variation. Analysing the phenotypic variation aids in identifying the functional polymorphisms of the causal genes. Natural variation is observed at each of the individual temperatures and Arabidopsis accessions respond to different ambient treatments.

The MAGIC accessions encompass a wide geographic and phenotypic varied sample across the Arabidopsis species (Kover, 2009). These accessions have been locally adapted to their specific native environments over long evolutionary periods of times and this is reflected in the variation of hypocotyl length. The genetic variation that we observe in hypocotyl length in response to temperature has presumably arisen in nature which implies that natural selection has permitted the survival of the underlying alleles that cause this variation.

In general, extensive naturally occurring phenotypic variation is prevalent in Arabidopsis accessions (Alonso-Blanco and Koornneef, 2000). When various genotypes are grown together and are compared under the same environmental conditions, genetic variation has been observed for hypocotyl length trait (figure 2.1 – 2.4). The variation in the hypocotyl length in response to temperature is due to the underlying genetic differences present in the accessions.

In darkness, seedlings of higher plants experience a different and specialised type of development termed as skotomorphogenesis, also known as etiolation. The seeds undertake this developmental program when they are buried under the soil in the dark. Under such conditions of growth, the seedlings have elongated hypocotyls, the cotyledons are folded and an apical hook is present. It was proposed that the apical hook

protects the shoot meristem when the seedling protrudes through the soil (Goeschl *et al*., 1967). Such morphologies of these juvenile organs collectively enable the seedlings to emerge from the soil layer in a swift fashion. This developmental programme ensures the seedlings to utilise the limited stored energy for hypocotyl elongation.

Upon reaching the surface, the seedlings detect light signals and this skotomorphogenic development is changed into a photomorphogenic development in the presence of light, ultimately leading the plant from a heteroautotrophic state to an autotrophic state (Chen *et al.*, 2004; Franklin *et al.*, 2005). The skotomorphogenic elongation growth of seedlings is a default pathway in etiolated development. Under dark growth conditions, this pathway is expressed due to the non-availability of light, as the seeds are buried in the soil and/or dense canopies of vegetation, and plants undergo skotomorphogenic growth.

Skotomorphogenesis has been suggested as an evolutionary adaptation in higher plants that gives them survival advantage in etiolated or etiolated-like terrestrial growth conditions (McNellis and Deng, 1995). There are a few studies where natural variation in etiolated hypocotyls has been studied. In a previous study, a smaller sample of 11 Arabidopsis accessions had shown subtle and indistinguishable differences in hypocotyl lengths when grown in the dark (Pepper *et al.*, 2002). The phenotypic analysis of the heterogeneous set of a small sample population of 19 MAGIC accessions possesses greater natural variation in hypocotyl length when phenotyped in the dark as seen in figure 2.5. These results are complementary to a previous study (Maloof *et al.*, 2001) which used a large number of 141 accessions grown under dark conditions. This further highlights the fact that they are genetically diverse and represent different native habitats. This study has shown that the MAGIC parents exhibit natural variation under both photomorphogenic and skotomorphogenic conditions. The high phenotypic variability demonstrated in photomorphogenesis and skotomorphogenesis in juvenile and vegetative states of growth in Arabidopsis make it an ideal system for ecological, genetic and functional studies.

2.4.2 Discussion on Hypocotyl Length Phenotypic Plasticity:

Phenotypic plasticity is the environmentally induced variation in phenotypes. There is a genetic basis for this response. Bradshaw (1965) defined plasticity as "..... shown by a genotype when its expression is able to be altered by environmental influences". Phenotypic plasticity has been defined as the ability of organisms to change their phenotypes in response to environmental changes (Schlichting, 1986). From a genomic point of view, phenotypic plasticity has been defined as "re-programming of the genome in response to the environment" (Aubin-Horth and Renn, 2009). Phenotypic plasticity can be also considered as an environmental acclimation. When grown under different temperature environments, the seedlings respond and the hypocotyls elongate which is indicative of the seedling to being accustomed to changing temperatures. This response pattern of the hypocotyls describes the whole process of adjusting to an environmental change.

Phenotypic plasticity takes place when the phenotypic responses of plants to exogenous environmental changes are mediated genetically. Plants at the molecular level receive and process environmental signals as a result of which plastic responses are originated. Therefore, the expression of phenotypic plasticity for any trait is mediated at the cellular level (Schlichting and Smith, 2002).

Natural variation in plant species, other than Arabidopsis, for phenotypic plasticity has been well identified (Khan *et al*., 1976 and Jain, 1978). Previous studies on phenotypic plasticity in Arabidopsis have identified significant G x E interactions. For example, a study on plasticity showed significant G x E interactions by quantifying vegetative traits in 26 natural accessions across low and high environments of mineral nutrients (Pigliucci and Schlichting, 1995).

Phenotypic plasticity has been termed as an ecological strategy (Anderson *et al.*, 2012) for species to continue in the face of climate change. Phenotypic plasticity is of great importance for various developmental processes in plants ultimately leading to their survival in the wild. It facilitates various genotypes to thrive in varying environmental conditions by changing phenotypic trait values in order to suite precise conditions (Moczek *et al.*, 2011). In the context of rapidly changing climates, phenotypic plasticity, in

a recent review, has been argued to be a vital determinant of plant responses; and for exploring phenotypic plasticity, plant functional traits should have priority (Nicotra *et al.*, 2010). The hypocotyl length is a plant functional trait involved in seedling shoot architecture; is also responsive to environmental cues and hence is ideal to studying phenotypic plasticity.

The interactions of a genotype with various temperatures can be represented in the form of reaction norms which is an essential subject in the field of molecular ecology. Schmalhausen (1949) introduced the concept of reaction norms of a genotype. A reaction norm is depiction of trait values relative to the environment. In ecological research, reaction norms are generally used to depict phenotypic plasticity (Aubin-Horth and Renn, 2009).

Reaction norms can only be plotted for homozygous individuals so that the same identical genotypes can be replicated and tested in different environments. Given the fact that the Arabidopsis accessions are naturally inbred lines and are homozygous due to the selfing nature of pollination (Lawrence, 1976, and Abbot and Gomes, 1989), their reaction norms for hypocotyl length can be plotted across a range of temperatures. This could well be an ideal example of differentiated reaction norms of a quantitative trait showing responsiveness to temperature.

Little information exists on reaction norms of quantitative traits in any species in general. In Arabidopsis, reaction norms for hypocotyl length in 180 accessions have been plotted for R:FR light responses (Filiault and Maloof, 2012). However, reaction norms in response to temperature have not been reported in Arabidopsis, despite its importance from a climate change point of view. This chapter addressed the basic question whether there is intraspecific natural variation for reaction norms in response to temperature in hypocotyl length. This study reports a novel finding of the prevalence of extensive intraspecific variation for plasticity in response to temperature in Arabidopsis. The ANOVA and the reaction norms strongly identify natural variation, plasticity, and natural variation for plasticity in hypocotyl length. In figure 2.8 the slopes of the intersecting lines between any given two temperature shifts exhibits a vast complexity between the inter-relationship of the genotypes and temperature in determining the hypocotyl length. A unique reaction norm is present in every parental accession for hypocotyl length; the

basis of which is the respective genotype and the temperature condition. This suggests that Arabidopsis has proven as a useful species for the investigation of temperature-dependent plasticity studies. These empirical findings vividly show the effectiveness of Arabidopsis for ecological genetic studies, particularly in response to temperature as the global mean temperatures continue to rise.

A trait can become plastic in nature when a genotype is subjected to an environmental change resulting in different phenotypes (Pigliucci, 2001). The plasticity experiment conducted in this study consists of the 19 parental genotypes grown in a series of temperature treatments in controlled incubators. Environmentally induced variability has been avoided to the maximum possible extent. The parental phenotyping results illustrate that the Arabidopsis genome reacts to the temperature treatment resulting in different hypocotyl phenotypes at different temperatures.

Changes in an organism's phenotype that are induced due to different environments are a result of phenotypic plasticity (Price *et al.*, 2003). Phenotypic Plasticity is observed during the developmental stage of hypocotyl elongation when accessions are grown under different temperature treatments. When the same genotypes are phenotyped across varying temperatures, we observe an environmentally induced variation that could be considered as a functional response. Changes in temperature influence the hypocotyl length and cause changes in phenotype. This functional response may ensure plants to maximise their survival and fitness in changing temperature conditions.

The variation seen in any particular accession across a range of temperatures represents phenotypic plasticity. Interestingly, all the genotypes not only showed notable variation in hypocotyl length at any given temperature, but also in their plastic responses to a range of temperatures. The different accessions reacted differentially to varying temperatures which is pointed out by the very high G x E interaction. Remarkably, none of the genotypes showed a non-significant interaction between any given two temperature treatment comparison (12°C - 17°C, 17°C - 22°C, 22°C - 27°C). The differences between the genotypes may be a result of genetic variation, environmental variation and/or the interaction of both and it is essential to separate the sources of phenotypic variation. Hence the analysis of phenotypic data of hypocotyl length illustrates that this is a trait which is under the control of endogenous genetic and exogenous environmental factors.

Arabidopsis changes its hypocotyl length phenotype in response to increased changes in temperature, hence at different temperatures, each of the accessions have different hypocotyl lengths. From figure 2.8, highly morphological distinct phenotypes can be identified. These results show that hypocotyl length is a trait that displays extensive plasticity due to its sensitivity to temperature as an environmental stimulus. In fact, the various accessions do not develop the same hypocotyl phenotype at any given temperature treatment, and therefore the genetic variation cannot be measured as null. The high degrees of differential plastic responses that are seen in the Arabidopsis accessions across a range of temperatures are due to the genetic differences present between them.

The comparisons of the hypocotyl lengths across the various genotypes are made at a common point in time after one week of growth. This comparison is based on two major standards, viz. chronology and growth, because hypocotyls during elongation show size dependent changes. This also shows that different genotypes growing under different temperature treatments grow at different rates. The hypocotyl elongation increases with increase in temperature and therefore the genotypes have different sizes across the range as can be seen in figure 2.8. As a result of multi-environment phenotyping analysis of the MAGIC accessions, it has been possible to show the phenotypic plasticity present in hypocotyl length.

2.4.3 Discussion on Hypocotyl Length Temperature Responsiveness:

The two-way ANOVA compared the plastic responses of the 19 genotypes for hypocotyl length grown at a range of temperatures. From this analysis, three main parts of variation were obtained; variance resulting from genotypes, variance resulting from temperature, and variance resulting from the interaction of the genotypes and temperature (G x E). The results of a significant environment term reveal that the genotypes express plasticity by responding to the temperature treatment. The 12°C hypocotyl length means for the accessions are the lowest followed by 17°C, 22°C and the highest at 27°C, as seen in the figure 2.8, hence there is a main effect for temperature. The results of a significant genotype term elucidate that across each of the temperatures, variation in the mean

hypocotyl length of the accessions is also seen, and thus there is a main effect for genotype. Moreover, the results of a significant G x E term show that there are differences present between the genotypes for their respective plastic responses. The Arabidopsis genotypes and ambient temperatures are interacting and affecting the hypocotyl length. 88% of the variability in hypocotyl length is accounted for by the G x E interaction and can be attributed to the accession and temperature interaction. How much difference there is in the hypocotyl length between accessions depends on genotype and temperature.

Figure 2.8 shows that Arabidopsis accessions change their ranking across a temperature range. The intersection of response lines at specific points in the figure explains the change in ranking. For example, an accession with a lower rank at one temperature has a higher rank at the other temperature when compared to other accessions. Such nature of the reaction norms for the different genotypes exhibits G x E interaction. This phenomenon explains that certain genotypes of Arabidopsis differ as a function of temperature response. The accessions are dynamic in their responses to temperature treatments which demonstrate their overall stability as observed in figure 2.8. This reveals that the accession performance is affected by the individual temperature environments, but the relative performance is steady across the various environments. This stability may well be a result of the underlying unique genetic makeup that the accessions possess and the effect of ambient temperature on hypocotyl length.

All the accessions show a progressive increase in hypocotyl length starting from 12°C up to 27°C. This highlights the fact that higher temperatures influence hypocotyl elongation as has been previously shown to be associated with higher auxin biosynthesis (Gray *et al.*, 1998). These interesting reaction norm findings reveal that various genotypes are not only highly responsive to temperature treatments but there is also considerable natural variation present between their responsiveness. These results explain that different accessions respond differentially to a range of temperature environments. Interestingly, this analysis reveals that the hypocotyl length trait is much more sensitive to temperature differences at higher temperatures than lower temperatures.

In summary, the greatest effects of the genotype were observed in the temperature combinations of 17°C - 22°C, whereas the greatest effects of temperature were seen in

the combinations 12°C - 27°C and 17°C - 27°C. The greatest effects of the interaction of both the main effects were identified in 12°C - 22°C and 12°C - 27°C combinations. Interestingly, a greater change in temperature results in greater variation in hypocotyl length due to G x E interaction.

In all the possible six combinations of temperature comparisons, the environmental effect is the greatest (table 2.1) which suggests hypocotyl length as a highly responsive trait to changes in ambient temperature conditions. The comparative analysis of hypocotyl length with plant size also supports this notion where in section 2.3.15 it has been shown that hypocotyl length trait is more responsive to changes in temperature than the plant size trait. Perhaps this is because in the Arabidopsis post-germination juvenile developmental stage, the embryonic shoot has to perceive the environmental signals from its surrounding and enable the seedling to being more responsive to these signals for its early on survival.

These interesting G x E interaction findings provide a strong basis for QTL mapping for a 'temperature responsiveness trait' which aims at identifying loci in the Arabidopsis genome associated with increased temperature responses, described in detail in Chapter 3. The temperature responsiveness findings showed that a main part of phenotypic variation in hypocotyl length amongst the MAGIC accessions was strongly temperature-dependent.

2.4.4 Discussion on Hypocotyl Length Photoperiod:

An important environmental factor affecting the hypocotyl length and the variation between the accessions is photoperiod. Therefore, it requires regulation when Arabidopsis accessions are phenotyped under controlled conditions in climate growth chambers. The comparison to classify which of the photoperiod conditions reveals greater natural genetic variation was essential for the identification of the ideal experimental conditions for the phenotyping of the MAGIC parents and their RILs for QTL mapping in the subsequent chapter. This analysis exemplifies the effect of photoperiod on hypocotyl length. Notable natural variation in hypocotyl length responses to short day photoperiods was observed. The estimated measures of effect size suggest that the genotypes grown in

short days account for a very large percentage of variability in the hypocotyl lengths when compared to long days.

Previously a complex association was found between day length and hypocotyl elongation (unpublished results in Maloof *et al.,* 2001). It was also reported that there is more variation in hypocotyl length between Arabidopsis accessions when they are phenotyped in short days (Detlef Weigel, Joanne Chory, Justin Borevitz, Mark Estelle; personal meetings) and the results of this comparative study on the MAGIC parental accessions complement this. They highlight the fact that the hypocotyl lengths are not only bigger when seedlings are grown in short days but they also show greater variation and the differences in hypocotyl lengths are bigger between the accessions. The duration of the daily dark period is an essential feature controlling elongation of hypocotyl length (Niwa *et al*., 2009). Hypocotyl length is very sensitive to photoperiod duration. Under extreme de-etiolation effects usually arising in long days, the hypocotyl ceases to elongate and the phenotypic differences between accessions, if any, are subtle. Based on this result, it is important to notice that photoperiod is an important environmental element determining the extent to which natural accessions show genetic variation in hypocotyl length. Hence, in order to maximise the phenotypic variation, the MAGIC parents and RILs were phenotyped in short days.

2.4.5 Discussion on Hypocotyl Length Adaptive Significance:

Upon discovering that Arabidopsis MAGIC accessions harbour substantial natural variation in hypocotyl length in response to temperature, an interesting question from an ecological point of view is whether hypocotyl length is an ecologically relevant trait and if natural variation for this trait would allow specific natural accessions to being advantageous in certain habitats. From an ecological point of view, the hypocotyl length is a trait that plays a vital role in the seedling establishment leading to its survival. In nature, this organ allows the seedling to emerge from the soil and establish itself in response to the various environmental signals that it receives. In this regard, hypocotyl length has been regarded as an ecologically relevant trait (Thomas Mitchell-Olds; personal meeting). Hypocotyl length has also been listed as one of the ecologically

important developmental phenotypes in a recent genome-wide association study (Atwell *et al.*, 2010).

Arabidopsis is a biogeographic species (Hoffmann, 2002) and table 2.4, comprising of the MAGIC parental accessions, reflects a wide geographic distribution, though being a smaller sample collected worldwide. Due to their collection from and potential adaptation to different geographic locations, the MAGIC accessions can be easily distinguished from one another on the basis of natural differences present in morphological traits, both in the vegetative and flowering stages.

One of the several factors that will determine the degree to which plants will adapt to changing climates is phenotypic variation (Aitken *et al.*, 2008). The temperature responsiveness study in the MAGIC parents suggests adaptive significance for hypocotyl length. The described phenotypic variation in hypocotyl length and their plastic responses to temperature may provide insights in to the degree to which natural populations of Arabidopsis may adapt to changing climates. The reaction norms in section 2.3.4 illustrate that Arabidopsis accessions display variable responses and this suggests that they are probably more suited to adapt to changing temperatures. The amount of plasticity in response to temperature may have evolved differently in various Arabidopsis accessions based on their natural habitats and the local environmental conditions of growth. These high levels of natural variation observed in hypocotyl length in the Arabidopsis accessions may increase their potential, in the early juvenile stage of development, to endure abiotic temperature changes resulting from climate change.

The natural variation observed in plastic responses is important because different genotypes can respond differentially to changing temperatures that may have important consequences for their adaptation to their local habitats. The consequence of plastic responses in the hypocotyl length in response to temperature is a morphological modification of the trait. This modification, in the natural habitat, may allow the different accessions to meet the challenges of the changing environments due to climate change.

Though a smaller sample of MAGIC parents was used for the several correlation analyses performed in this chapter, it provides a basic insight into the study of their adaptations to native habitats. The result of the correlation analysis between hypocotyl length and

habitat spring temperature shows that they have a significant medium to high positive correlation. However, no significant correlation was observed between etiolated and de-etiolated hypocotyl length and latitudinal clines. This result of no significant positive correlation between etiolated hypocotyl length and latitudinal clines is consistent with a previous finding on a larger set of Arabidopsis accessions (Maloof *et al.*, 2001). Several other studies in light responses also found a negative correlation between hypocotyl length and latitude of origin (Stenøien *et al*., 2002; Balasubramanian *et al*., 2006; Filiault and Maloof, 2012).

Arabidopsis latitudinal distribution range covers 0° up to 68°N (Koornneef *et al*., 2004). Latitudinal natural variation in plant size has been well studied in forty Arabidopsis accessions ranging from 16°N to 63°N (Li *et al.*, 1998). However, latitudinal variation in hypocotyl length in response to temperature is less well studied. The MAGIC parental accessions represent considerable variation in the latitudes of their habitats ranging from 28°N to 60°N, suggesting variable growth temperature environments. In general, the tropical climate which is near to the equator of the Earth has a pattern of warm temperatures throughout the year with little variation. As the latitude distance from the equator increases, particularly towards the north where most of the Arabidopsis accessions come from, the average temperatures gradually tend to decrease. Therefore, the varying latitudes, which are generally a good predictor of average temperatures, affect the average local temperature in the native regions of the accessions, in addition to other contributing factors such as altitude. Nonetheless, it may be reasonable to hypothesise that plants may have bigger hypocotyls in habitats nearer to the equator where ambient temperatures are higher; and with the latitudinal cline towards the North the hypocotyls may tend to be smaller. In nature, the Arabidopsis accessions may have adapted to the varied temperatures and latitudes of their native habitats. Results from this study reveal that correlation of hypocotyl length is stronger with native habitat temperature than latitudes.

It has been previously determined that hypocotyl length does not have any correlation with flowering time across latitudinal clines (Stenøien *et al.*, 2002). The comparative analysis of hypocotyl length and plant size in this study has shown no correlation between these two traits across ambient temperature. Hence, information collected on hypocotyl

length cannot be used to infer the phenotypic values of plant size and vice versa. These traits need to be assayed independently.

The presence of phenotypic variation in the MAGIC parental accessions is central for dissecting the genetic architecture of hypocotyl length in response to temperature. Gaining a profound understanding of the plastic responses is critical for not only predicting but also managing the effects of changing climates on plant species and may have broader applications for managing crop plants in response to higher temperatures. The findings of the parental phenotyping experiments highlight the value of such analyses in relation to climate change and are an important milestone in paving the way for undertaking genetic analyses across a range of ambient temperatures. The extensive phenotypic variation discovered in the MAGIC accessions for hypocotyl length in response to temperature provides a strong basis for exploring the underlying genetic basis of this variation, which begins with QTL mapping in Chapter 3.

2.5 References:

Abbot R, Gomes M. (1989) Population Genetic Structure and Outcrossing Rate of *Arabidopsis thaliana* (L.) Heynh. Heredity, 62, 411-418.

Aitken S, Yeaman S, Holliday J, Wang T, Curtis-McLane S. (2008) Adaptation, Migration or Extirpation: Climate Change Outcomes for Tree Populations. Evol Appl. 1, 1, 95–111.

Alonso-Blanco C, Koornneef M. (2000) Naturally Occurring Variation in Arabidopsis: An Underexploited Resource for Plant Genetics. Trends Plant Sci. 5, 1, 22-9.

Anderson J, Panetta A, Mitchell-Olds T. (2012) Evolutionary and Ecological Responses to Anthropogenic Climate Change. Pl Physio, 160, 4, 1728-1740.

Atwell S, Huang Y, Vilhjalmsson B, *et al.* (2010) Genome-wide Association Study of 107 Phenotypes in *Arabidopsis thaliana* Inbred Lines. Nature, 465, 7298, 627-631.

Aubin-Horth N, Renn S. (2009) Genomic Reaction Norms: Using Integrative Biology to Understand Molecular Mechanisms of Phenotypic Plasticity. Molecular Ecology, 18, 18, 3763–3780.

Balasubramanian S, Sureshkumar S, Agrawal M, Michael T, Wessinger C, *et al*. (2006) The PHYTOCHROME C Photoreceptor Gene Mediates Natural Variation in Flowering and Growth Responses of *Arabidopsis thaliana*. Nat Genet, 38, 711–715.

Bradshaw A. (1965) Evolutionary Significance of Phenotypic Plasticity in Plants. Advan in Genetics, 13, 115-155.

Chen M, Chory J, Fankhauser C. (2004) Light Signal Transduction in Higher Plants, Annu. Rev. Genet. 38, 87–117.

Filiault D, Maloof J. (2012) A Genome-Wide Association Study Identifies Variants Underlying the *Arabidopsis thaliana* Shade Avoidance Response. PLoS Genet 8, 3,e1002589.

Franklin K, Larner V, Whitelam G. (2005) The Signal Transducing Photoreceptors of Plants. Int. J. Dev. Biol. 49, 653 – 664.

Goeschl J, Pratt H, Bonner B. (1967) An Effect of Light on the Production of Ethylene and the Growth of the Plumular Portion of Etiolated Pea Seedlings. Plant Physiol. 42, 1077–1080.

Gray W, Östin A, Sandberg G, Romano C, Estelle M. (1998) High Temperature Promotes Auxin-mediated Hypocotyl Elongation in Arabidopsis, Proc. Natl. Acad. Sci. USA, 95, 7197–7202.

Hoffmann M. (2002), Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae), Journal of Biogeography, 29, 1, 125–134.

Jain S. (1978) Inheritance of Phenotypic Plasticity in Soft Chess, *Bromus mollis* L. (Gramineae). Experientia 34, 835-36.

Khan M, Antonovics J, Bradshaw A. (1976) Adaptation to Heterogeneous Environments. III. The Inheritance of Response to Spacing in Flax and Linseed (*Linum usitatissimum*). Aust. J. Agric. Res. 27, 649-59.

Koorneef M, Alonso-Blanco C, Vreugdenhil D. (2004) Naturally Occurring Genetic Variation in *Arabidopsis thaliana*. Annual Review of Plant Biology, 55, 141-172.

Lawrence M. (1976), Variations in Natural Populations of *Arabidopsis thaliana* (L.) Heynh. In 'The Biology and Chemistry of the Cruciferae', ed. Vaughan, Macleod, Jones, 167-190.

Li B, Suzuki J, Hara T. (1998) Latitudinal Variation in Plant Size and Relative Growth Rate in *Arabidopsis thaliana.* Oecologia, 115, 293–301.

Ludwig-Muller J, Krishna P, Forreiter C. (2000) A Glucosinolate Mutant of Arabidopsis is Thermosensitive and Defective in Cytosolic Hsp90 Expression after Heat Stress. Plant Physiology, 123, 949–958.

Maloof J, Borevitz J, Dabi T, Lutes J, Nehring R, *et al*. (2001) Natural Variation in Light Sensitivity of Arabidopsis. Nat. Genet. 29, 441–46.

McNellis T, Deng X. (1995) Light Control of Seedling Morphogenetic Pattern. The Plant Cell, 7, 1749-1761.

Moczek A, Sultan S, Foster S, Ledón-Rettig C, Dworkin I, *et al*. (2011) The Role of Developmental Plasticity in Evolutionary Innovation. Proc Biol Sci, 278, 2705–2713.

Murashige T, Skoog F. (1962) A Revised Medium for Rapid Growth and Bioassays with Tobacco Tissue Cultures. Physiol Plant, 15, 473–497.

Nicotra B, Atkin K, Bonser P, Davidson M, Finnegan J, *et al.* (2010) Plant Phenotypic Plasticity in a Changing Climate. Trends Plant Sci, 15, 684–92.

Niwa Y, Yamashino T, Mizuno T. (2009) The Circadian Clock Regulates the Photoperiodic Response of Hypocotyl Elongation through a Coincidence Mechanism in *Arabidopsis thaliana*. Plant Cell Physiol, 50, 4, 838-854.

Pepper A, Corbett R, Kang N. (2002) Natural Variation in Arabidopsis Seedling Photomorphogenesis reveals a likely Role for TED1 in Phytochrome Signalling. Plant, Cell & Environment, 25, 4, 591–600.

Pigliucci M. (2001) Phenotypic Plasticity: Beyond Nature and Nurture. John Hopkins University Press, Baltimore, MD.

Pigliucci M, Schlichting C. (1995) Reaction Norms of Arabidopsis (Brassicaceae). III. Response to Nutrients in 26 Populations from a Worldwide Collection. Am J Bot, 82, 1117–25.

Pigliucci M, Schlichting C. (1996) Reaction Norms of Arabidopsis. IV. Relationships between Plasticity and Fitness. Heredity, 76, 427–436.

Pigliucci M, Whitton J, Schlichting C. (1995) Reaction Norms of Arabidopsis. I. Plasticity of Characters and Correlations across Water, Nutrient and Light Gradients. J Evol Biol, 8, 421-438.

Price T, Qvarnström A, Irwin D. (2003) The Role of Phenotypic Plasticity in Driving Genetic Evolution. Proc. Biol. Sci. 270, 1523, 1433–40.

Schlichting C. (1986) The Evolution of Phenotypic Plasticity In Plants. Ann. Rev. Ecol. Syst., 17, 667-93.

Schlichting C, Smith H. (2002) Phenotypic Plasticity: Linking Molecular Mechanisms with Evolutionary Outcomes. Evol. Ecol., 16, 189–211.

Schmalhausen I. (1949) A Study of Reaction Norms in Natural Populations of *Drosophila pseudoobscura*. Factors of Evolution, Blakeston, Philadelphia.

Stenøien H, Fenster C, Kuittinen H, Savolainen O. (2002) Quantifying Latitudinal Clines to Light Responses in Natural Populations of *Arabidopsis thaliana* (Brassicaceae). Am J Bot, 89, 1604–1608.

Smeekens S. (2000) Sugar-Induced Signal Transduction in Plants. Annual Review of Plant Physiology and Plant Molecular Biology, 51, 49-81.

Chapter 3

Dissecting the Genetic Basis of Hypocotyl Length in Response to Temperature by QTL Mapping

3.1 Introduction:

Finding key genes or novel alleles of candidate genes for plant growth traits is a major goal for experimental studies in natural variation. Natural variation in Arabidopsis accessions can be exploited for novel allelic discovery. The plant growth trait of hypocotyl length is controlled by multiple loci that contribute to the variation to varying degrees. QTL mapping is an essential tool for dissecting such natural variation. Because continuous variation is present in hypocotyl length within Arabidopsis populations, it is a quantitative trait for which methods of QTL mapping can be deployed (Falconer and Mackay, 1996). QTL mapping utilises intraspecific natural variation (Alonso-Blanco and Koornneef, 2000; Maloof, 2003) and helps in identifying allelic variation for known candidate genes and/or isolating novel genes that may contribute to phenotypic variation in hypocotyl length. For any given trait under investigation, the principle of QTL mapping is based on the genotyping of progenitor lines obtained from intercrossing dissimilar accessions (Slate, 2005). Thus, QTL mapping has an ultimate goal of determining the genes controlling natural variation in the trait (Mackay, 2001). This is also called QTL cloning (Borevitz and Chory, 2004) which is followed by detection of causal sequence polymorphisms. Achieving this goal is a daunting task and QTL studies for quantitative traits usually only identify the approximate map position of QTL. The search for QTL genes is rapidly developing (Barton and Keightley, 2002); however, the identification of genes that affect intraspecific phenotypic variation still remains a major challenge (Christians and Keightley, 2002).

Natural variation in Arabidopsis accessions has been shown to be an essential asset for investigating photomorphogenic responses of seedlings (Maloof *et al*., 2000), and this has paved the way for hypocotyl length analysis to temperature responses. This chapter examines the dissection of thermomorphogenic quantitative responses of seedlings.

Therefore, Arabidopsis is particularly suited to study natural variation and the effects of temperature on hypocotyl length. Also with the available methods, novel allelic discovery is possible.

In chapter 2, phenotypic variation of hypocotyl length in the 19 MAGIC parental accessions in response to temperature was established. This chapter is about the dissection and determination of the genetic basis of hypocotyl length and is the second step in the forward genetic analysis (figure 1.4). This step is accomplished with the help of QTL mapping; a statistical method that looks for significant associations between the molecular markers and the phenotypic trait values (phenotype-genotype association). QTL mapping is for the identification of the genomic regions that affect variation in hypocotyl length. In chapter 3, a QTL mapping study has been applied to examine the basis of natural variation in hypocotyl length and temperature responses. It will look at the phenotypic analysis of the RILs and describe the discovery of the QTL. For this purpose, reliable phenotypic screen of the RILs with precise measurements is essential.

Mather (1949) introduced the term 'polygenes' for genes that underlie quantitative traits. Later, Gelderman (1975) first coined the acronym 'QTL' to describe polymorphic regions of the genome that contribute to the variation of continuous traits. This region could be a single gene or a cluster of linked genes affecting the trait. QTL mapping is performed when the phenotypic data of the trait of interest and genotypic data of molecular markers of experimental populations are statistically compared to look for associations (Alonso-Blanco *et al*. 2006). In a QTL mapping, one is essentially looking at all the genome to find the genes that influence the trait of interest. This process of discovery allows us to identify loci in response to specific environmental conditions of growth.

There are four pre-requisites for QTL mapping:

- An experimental population of lines that is genetically variable
- The presence of molecular markers that allow the genotyping of this experimental population
- Quantitative phenotypic data for a trait
- Appropriate statistical methods for the isolation of significant QTL

In this research study, the general goal of QTL mapping is to correlate the hypocotyl length phenotype with the specific parental accessions by using RILs where the molecular markers segregate due to recombination events. One of the principles of QTL mapping is that the detection of QTL can only take place in genomic regions where the parents of a cross vary.

What are the QTL that control hypocotyl length trait in response to temperature? To address this question the phenotyping of the RILs needs to be performed in multiple temperature environments. This will not only allow the discovery of QTL at individual temperatures but will also facilitate the discovery of environmental QTL for a 'temperature responsiveness trait'.

The goals of this chapter are:

1. Phenotypic analysis of the RILs
2. QTL mapping and discovery of QTL for hypocotyl length trait to temperature responses
3. Discovery of environmental QTL for 'temperature responsiveness trait'
4. Determination of underlying sequence variation in candidate genes
5. Shortlisting of candidate genes for QTL cloning

3.2 Methods and Materials:

The methods and conditions for seedling growth for the calibration experiment (for comparison of seedling hypocotyl length at 12°C and 17°C) and the phenotyping of the RILs were the same as explained in section 2.2.

3.2.1 Growth Conditions for Calibration Experiment:

Phenotyping of the parental accessions showed that seedlings grown at higher ambient temperatures of 22°C and 27°C were bigger and easier to handle while scanning them, with no risk of damage to the hypocotyls. However, seedlings grown at lower ambient temperatures of 12°C and 17°C were very small (1 mm) and difficult to handle, when placing them on to the transparency sheets at the time of scanning. Hence, there was an increased risk of damage to the seedlings, especially the hypocotyls. In addition, at lower ambient temperatures the hypocotyls reached their maximum length after 1 week of growth and the first true leaves started to appear and grow in the presence of light facilitating the start of photosynthesis. Therefore, RILs at 12°C and 17°C were grown for 2 weeks and RILs at 22°C and 27°C were grown for 1 week. This method allowed the roots and first true leaves to develop, making them easier to handle and consequently eliminate the risk of damage to the hypocotyls and without compromising with the hypocotyl length due to the longer time given for growth.

To test if there are any confounding effects due to the difference in time scales resulting in any differences between 1 and 2 weeks of hypocotyl growth at 12°C and 17°C, a calibration experiment was performed which showed that the hypocotyls had equal lengths when grown for 1 and 2 weeks. The seedlings of the 19 parental accessions were grown in Petri dishes for one week at 12°C and 17°C in incubators. After 1 week of growth, digital images were taken with a Nikon D300S camera. The same seedlings were placed back in the respective incubators and allowed to grow for another week. After 2 weeks of growth, images were taken again. The hypocotyls were measured in ImageJ. Up to 66 individuals per genotype were measured. At lower ambient temperatures 12°C and 17°C the cessation of the hypocotyl elongation occurs after 1 week of growth upon reaching saturation. This is can be observed in figures 3.1 and 3.2.

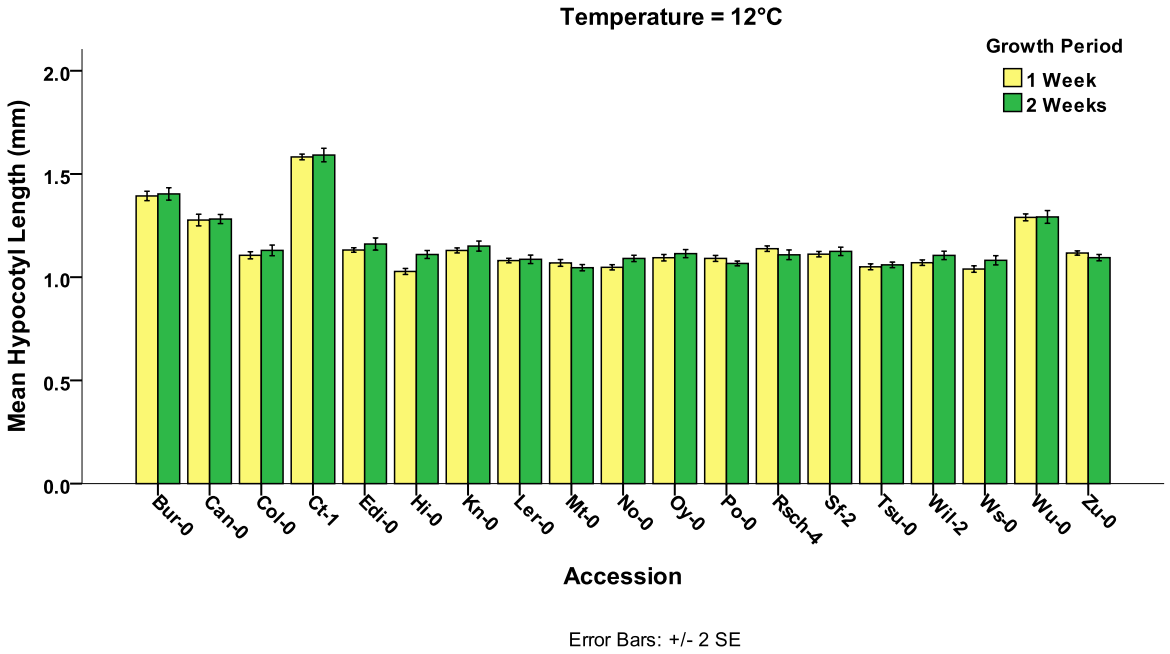

Figure 3.1: 19 Parental accessions phenotyped at 1 week and 2 weeks of growth at 12°C showing similar hypocotyl lengths

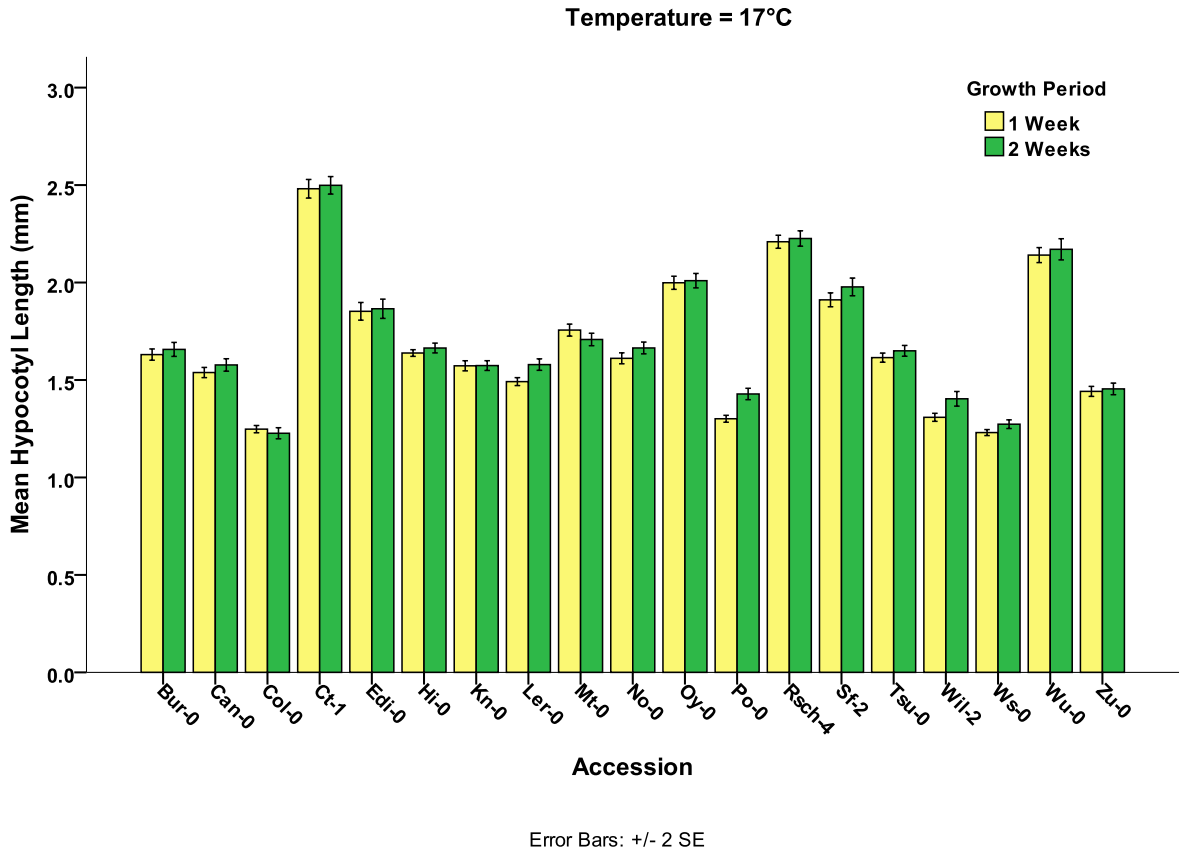

Figure 3.2: 19 Parental accessions phenotyped at 1 week and 2 weeks of growth at 17°C showing similar hypocotyl lengths

These results showed that in any of the given genotypes the hypocotyl lengths were the same and this indicated that the time factor, as a control, for the phenotyping of the RILs at 12°C and 17°C was not compromised. Therefore, since the hypocotyl lengths were similar, the seedlings grown for 2 weeks were used for phenotyping.

3.2.2 Experimental Population for QTL Mapping:

The genetically diverse resource of the Multiparent Advanced Generation Inter-Cross (MAGIC) Recombinant Inbred Lines (RILs) of Arabidopsis was used for the phenotyping of the hypocotyl length. This is a synthetic mapping population and is a family based method. The molecular markers used in MAGIC lines are single nucleotide polymorphisms (SNPs). This segregating RIL population is an alternative resource to genetically dissecting complex traits and is derived from the intercrossing of the 19 parental accessions, which represent more of the genetic and phenotypic variation present.

These accessions come from a wide range of geographical locations and their genotypes represent local adaptations to their original habitats and climates. The advantages of using this mapping population are manifold. The RILs are highly recombinant which improves the mapping resolution. There is greater allelic and phenotypic diversity present in the RILs due to the intermating of 19 heterogeneous parents. Moreover, this immortal population that consists of homozygous individuals allows phenotyping in multiple environments for the detection of G x E interactions.

These lines were constructed by intermating each of the 19 founder accessions as both maternal and paternal parents, 19 X 18 = 342 F1 progeny. The F1 were intermated randomly for four generations producing 342 outbred F4 families. From each F4 family, 3 inbred MAGIC lines were derived by selfing an F4 plant for six generations producing a total of 1026 RILs. The first batch of 527 RILs was obtained from NASC which were used for phenotyping.

3.2.3 Phenotyping of RILs:

One of the pre-requisites for mapping QTL is accurate phenotyping of the RILs for the hypocotyl length trait. By using growth incubators for phenotyping, all environmental variables of growth such as temperature, light and humidity, were controlled and kept

constant to the greatest extent possible. Phenotyping of all the RILs under multiple temperature environments is a bottleneck in QTL analyses. It is a time consuming and laborious task. Under such circumstances, high throughput phenotyping is essential.

Phenotyping of the hypocotyl length under controlled environments should involve high experimental reproducibility and better levels of precision in phenotyping. The accurate phenotyping allows the dissection of the hypocotyl length trait for the large population of the RILs to enable QTL mapping to be undertaken, elucidating the molecular basis of this complex trait.

For RILs, N = 12 and in total > 19,000 plants were measured for QTL mapping. The replication of genetically identical individuals reduced the environmental error and increased the statistical power of analyses and this provided a powerful system to map QTL.

3.2.4 Image Analysis for Quantitation:

The large amount of hypocotyl length quantitative data of the RILs was analysed with the help of using Image J for image analysis. However, it was a laborious task due to the manual method of measuring the hypocotyls of each of the individual seedlings in the images.

3.2.5 QTL Mapping Procedure/Statistical Analysis:

Associations of the hypocotyl length trait differences with specific SNP markers were used for the identification of chromosomal intervals that harbour genes accountable for the natural variation segregating in the MAGIC lines. The QTL mapping was performed in the HAPPY package in the software/programme R which has been developed by Richard Mott (http://www.well.ox.ac.uk/happy/happyR.shtml). The genome-wide molecular genetic map of the RILs is based on SNPs. Hypocotyl length phenotypic values of the RILs were compared with the SNP marker genotypes of the RILs to search for specific genomic intervals (QTL) that showed statistically significant associations with hypocotyl length trait variation.

Associations at the four temperatures was performed and a QTL was declared on the basis of $p = 0.05$ significance threshold. If $p = 0.01$ then there is a chance to have false

negatives as the actual SNPs that may be involved in the QTL may be missed out. If p = 0.05 then there is a chance to have false positives but at the same time we may not miss out actual causal genes that could be involved in the QTL.

This was followed by a candidate gene approach. In order to detect QTL, the QTL mapping software performed a genome scan to determine QTL thresholds for statistical significance. It then found all QTL where the logP of genetic association to the hypocotyl length phenotype data was significant at p = 0.05.

The QTL mapping also provided the mean estimates of the parental accession effects at each QTL peak marker SNP and all the SNP markers within a particular QTL. The QTL mapping was performed as described by Kover *et al*. (2009).

Two different types of QTL mapping were performed; one type was mapping at individual temperatures, the same ambient temperatures at which the parents were phenotyped (12°C, 17°C, 22°C, 27°C). In this type, phenotypic data of all the RILs was used separately at each of the temperatures.

The second kind was mapping by calculating the ratio between any two temperatures to measure the response of the RILs to temperature changes. This type of mapping was carried out for the temperature responsiveness trait. In essence, this mapping should isolate QTL that are implicated in increased temperature responses.

The next task was to find polymorphisms in candidate genes that may be involved in temperature responsiveness.

3.2.6 Bioinformatic Methods:

The *in silico* analysis of genomic sequences was performed for the candidate genes in the different QTL. The software 'BioEdit Sequence Alignment Editor' was used for genomic sequence analysis, which is a very effective programme to performing this data analysis.

For each of the genes, the predicted protein sequences for all the 19 parents were aligned using Multiple Sequence Comparison by Log- Expectation (MUSCLE) software available at http://www.ebi.ac.uk/Tools/msa/muscle/. This multiple sequence alignment software

achieves both better speed and average accuracy as compared to ClustalW2 or T-Coffee softwares.

At the time of this analysis, the sequences of the MAGIC parents were not available online on bioinformatic search websites but were available to download from Richard Mott's webpage in the raw form.

In order to do a multiple sequence alignment for the detection of polymorphisms, it was necessary to *in silico* extract the alleles of the candidate genes in all the founder accessions of the MAGIC lines. The identification of alleles in parental accessions other than Col-0 was a difficult step because they have different physical positions on the respective chromosomes. Therefore, the coding sequences of the DNA of all the candidate genes were manually extracted for each of the 18 parental accessions by using the reference sequences from Col-0 accession which were available on TAIR's website.

The DNA sequences were translated to protein coding sequences and multiple sequence homology was used for the detection of amino acid changes or other indels in the candidate genes.

3.3 Results:

3.3.1 Phenotypic Characterisation of MAGIC RILs:

As mentioned before, in a QTL analysis, the phenotypic characterisation of RILs is a bottle neck and requires high throughput phenotyping methods that need to be accurate and precise. This accuracy of the quantitative trait data is essential for the detection of QTL.

To study the amount of variation in the RILs, a one-way ANOVA was performed at each of the four treatment temperatures. The results show that there are highly significant differences between the means of hypocotyl lengths of the RILs.

12°C = $F(411, 4532) = 680.77$, $p < 0.001$)
17°C = $F(339, 3740) = 710.32$, $p < 0.001$)
22°C = $F(425, 4686) = 1420.5$, $p < 0.001$)
27°C = $F(408, 4499) = 4707.7$, $p < 0.001$)

RILs show significant differences in hypocotyl length because they were constructed by the intermating of the 19 founder accessions and hence each RIL carries a unique fine mosaic of the parental genomes. The phenotypic analysis of the RILs provided valuable insights into the large range of quantitative data which the hypocotyl length trait exhibits. The range of hypocotyl lengths at each of the temperatures shows the extent of genetic variation in the RILs;

12°C = 0.79 – 3.29 mm,
17°C = 0.96 – 4.39 mm,
22°C = 1.29 – 7.07 mm,
27°C = 3.10 – 11.64 mm

3.3.2 Hypocotyl Length Trait is Continuously Distributed:

The phenotypic characterisation of the RILs reveals that hypocotyl length is quantitatively inherited and therefore shows a continuous phenotypic distribution. Figure 3.3 shows the continuous range of hypocotyl lengths in the RILs at different temperatures. The inheritance of hypocotyl length in an F2 progeny of a cross between Col-0 and Sf-2 accessions has been studied thoroughly in chapter 2.

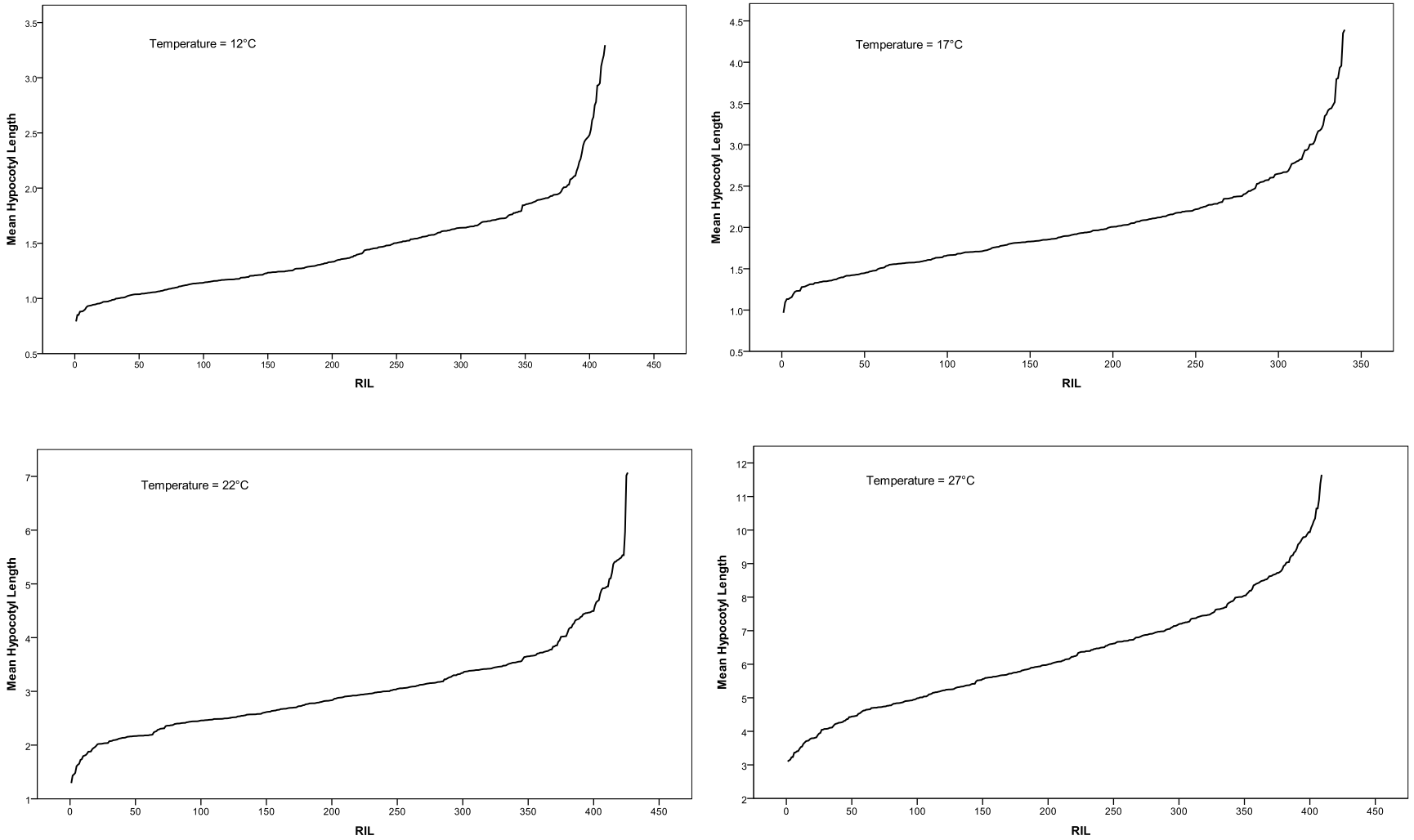

Figure 3.3: Plots revealing the continuous distribution of RILs at different temperatures; a typical feature of quantitatively inherited traits.

3.3.3 Discovery of QTL at Individual Temperatures 12°C, 22°C, 27°C:

QTL were mapped which were responsible for hypocotyl length variation in response to temperature in the MAGIC RILs. Temperature-dependent QTL have been isolated. At different temperatures, different sets of QTL appear suggesting that particular QTL affect natural variation in hypocotyl length trait under a specific set of environmental conditions. The QTL are temperature sensitive as seen in figure 3.4, revealing the importance of temperature.

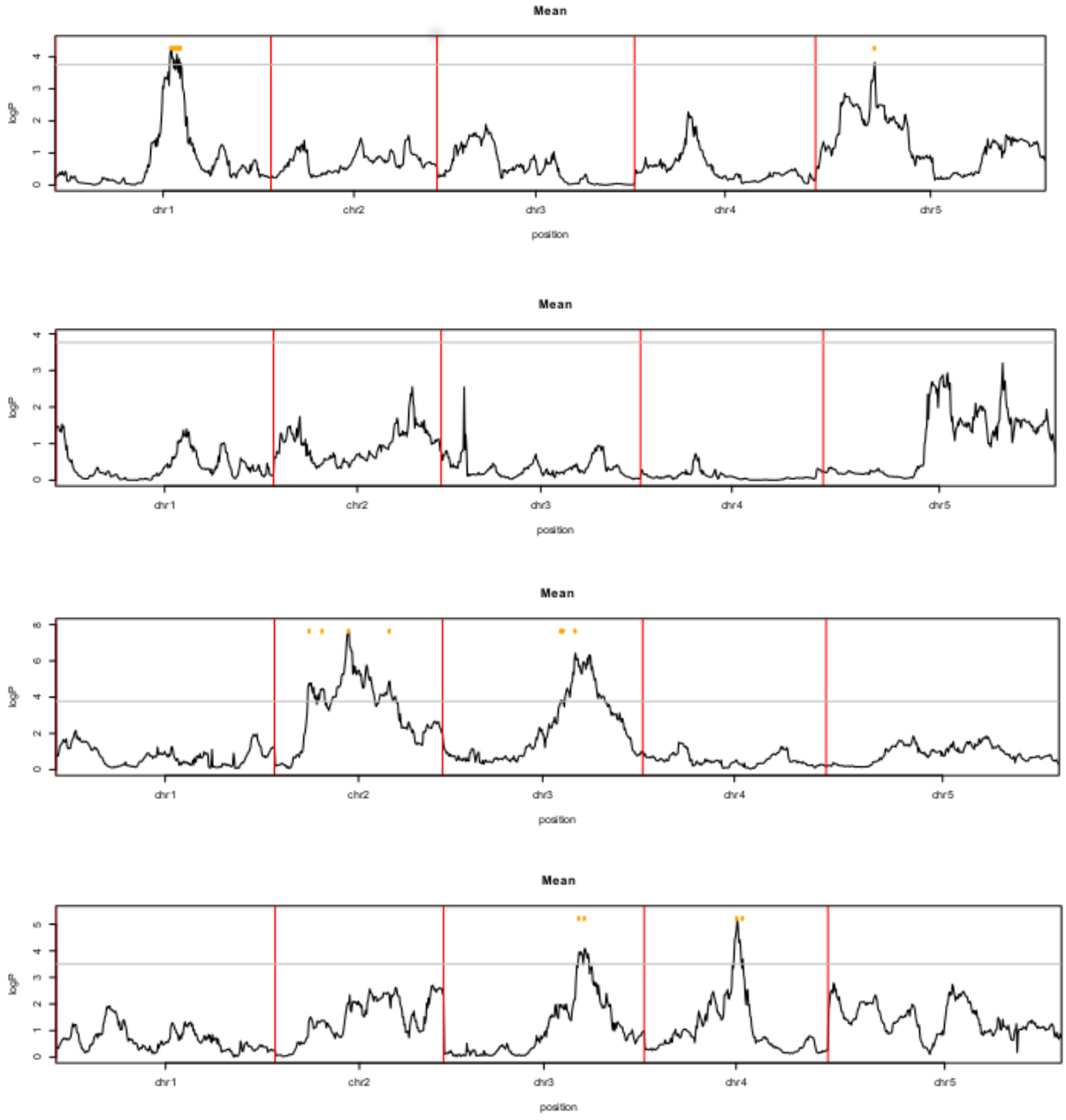

Figure 3.4: QTL scanning maps at 12°C, 17°C, 22°C and 27°C (top to bottom) showing the isolated QTL with peak marker positions as orange dots. The red lines are boundaries of chromosomes.

12°C							
Chromosome	QTL	Island.From.bp	Island.To.bp	Peak.bp	Peak.SNP	logP	Genomewide.pvalue
Chr1	QTL1	13832180	15765615	14922486	PERL0126101	4.255643273	0.021
Chr1	QTL2	15928121	15974746	15974746	NMSNP1_15977186	3.894408886	0.0385
Chr1	QTL3	16134927	16644538	16251782	PERL0133385	4.08008612	0.0282
Chr1	QTL4	16871886	16872100	16872100	MN1_16874540	3.928843916	0.0359
Chr1	QTL5	17148697	17179544	17179544	MN1_17181992	3.797107195	0.0461
Chr5	QTL1	5005747	5145069	5145069	MASC01136	3.82599005	0.044
22°C							
Chromosome	QTL	Island.From.bp	Island.To.bp	Peak.bp	Peak.SNP	logP	Genomewide.pvalue
Chr2	QTL1	6154830	7722797	7065654	MN2_7072735	4.786055984	0.008
Chr2	QTL2	7722958	8144073	8141710	PHYB_2850	4.482622015	0.0137
Chr2	QTL3	8742966	13319028	10933252	MASC02949	7.628476898	0.0002
Chr2	QTL4	13782937	15244621	14399791	MN2_14406873	4.896008743	0.0062
Chr3	QTL1	15317766	15668809	15668414	MN3_15679400	3.845717159	0.0439
Chr3	QTL2	15760957	15869095	15869095	MN3_15880080	3.793127808	0.0482
Chr3	QTL3	15966634	19913837	17155377	MN3_17166362	6.437050678	0.0013
27°C							
Chromosome	QTL	Island.From.bp	Island.To.bp	Peak.bp	Peak.SNP	logP	Genomewide.pvalue
Chr3	QTL1	17370484	17639552	17438018	MASC07256	3.960640517	0.0197
Chr3	QTL2	17755915	18428879	17871357	NMSNP3_17882342	4.094330979	0.0142
Chr4	QTL1	10042015	10977564	10302880	MASC01526	5.207123235	0.0012
Chr4	QTL2	11001770	11003558	11003558	TSF_606	3.691393187	0.0325

Table 3.1: Different QTL isolated at 12°C, 22°C and 27°C for hypocotyl length. At 12°C, a total of 6 QTL; at 22°C 7 QTL; and at 27°C 4 QTL were identified.

The use of MAGIC RILs has enabled fine-scale and refined QTL mapping to isolate majority of the intervals that contain a manageable number of genes for the candidate gene approach to follow. Upon identification of obvious candidate genes involved in hypocotyl elongation, a comprehensive methodology was undertaken to shortlist the candidate genes. This has been described in detail in section 3.3.5.

When compared to traditional biparental QTL mapping, the isolated QTL intervals have narrow boundaries. The mapping resolution of MAGIC RILs is very high and several QTL have been mapped in the sub-centimorgan (cM) range as shown in table 3.2. This fine mapping of QTL allows direct investigation of the candidate genes for further shortlisting and functional analysis. The QTL represent a mapping resolution of the order of few to several centimorgans.

12°C					
QTL Name	**QTL Interval in bp**	**QTL Interval in Kb**	**QTL Interval in Mb**	**QTL Interval in cM**	**No. of Genes**
Chr1, QTL1	1933435	1933.435	1.933435	9.66	42
Chr1, QTL2	46625	46.625	0.046625	0.23	4
Chr1, QTL3,	509611	509.611	0.509611	2.54	51
Chr1, QTL4	214	0.214	0.000214	0.001	1
Chr1, QTL5,	30847	30.847	0.030847	0.15	9
Chr5, QTL1	139322	139.322	0.139322	0.69	39
22°C					
Chr2, QTL1	1567967	1567.967	1.567967	7.83	288
Chr2, QTL2	421115	421.115	0.421115	2.10	110
Chr2, QTL3	4576062	4576.062	4.576062	22.88	1160
Chr2, QTL4	1461684	1461.684	1.461684	7.30	426
Chr3, QTL1	351043	351.043	0.351043	1.75	38
Chr3, QTL2	108138	108.138	0.108138	0.54	19
Chr3, QTL3	3947203	3947.203	3.947203	19.73	1009
27°C					
Chr3, QTL1	269068	269.068	0.269068	1.34	61
Chr3, QTL2	672964	672.964	0.672964	3.36	177
Chr4, QTL1	935549	935.549	0.935549	4.67	243
Chr4, QTL2	1788	1.788	0.001788	0.008	1

Table 3.2: QTL intervals in bp, Kb and Mb and total number of protein-coding genes in each of the QTL.

3.3.4 Genetic Architecture of Hypocotyl Length as a Quantitative Trait:

In plants, the number of genes, their location on the genome and QTL effect sizes, describes the genetic architecture of complex traits (Holland, 2007). An interesting feature of MAGIC RILs is that for each SNP marker, the QTL mapping reconstructs the mean estimates of hypocotyl lengths for all the 19 parental accessions. This data calculates the effect sizes of the parental accessions and identifies the parental contribution to the variation of the hypocotyl length trait. The effect sizes of those SNPs that fall within a QTL interval were all available and that provided insights to the genetic architecture of hypocotyl length.

From the QTL mapping, it is interesting to know that at each of the temperatures, few large effect QTL determine the quantitative genetic variation of hypocotyl length. This in turn means that the few QTL explain a greater proportion of the variation. As the number of QTL is inversely proportional to the distribution of variance, fewer QTL isolated account for the variance and hence they are large effect QTL and vice versa. The underlying genetic variance is attributable to few QTL. Therefore, due to the fact that multiple parents are used in MAGIC QTL mapping, instead of having an overall percentage of variation of hypocotyl length for each of the QTL, it is of great interest to know which of the accessions are major effect parents and show more differences in the hypocotyl length phenotypes at each particular isolated QTL. This is explained by the mean accessions estimates, shown in section 3.3.5.2, which provides a meaningful way to analyse the contribution of the parents to the variation of the QTL, at the peak individual SNP molecular marker.

There are two important quantitative analyses that are informative about the genetic architecture of hypocotyl length. The first is to calculate how much of the phenotypic variance between accessions is attributable to genetic variation. This is the simplest way of analysing genetic and phenotypic variance. This has already been carried out in chapter 2 that provides the estimates of heritability. In this analysis, since the sample is very big, therefore, it is a very good estimate of heritability.

The second analysis is to determine what proportion of the phenotypic variation in hypocotyl length is explained by each of the QTL. In a QTL if there is a single gene, it is

interesting to analyse the proportion of the contributing variance of this locus to hypocotyl length phenotypic variation. This proportion of variance has been estimated by the QTL mapping procedure by the reconstruction of the parental accessions explaining their effect sizes at this particular locus. Hence this quantitative data is more useful in estimating the proportion of variation in hypocotyl length by each of the loci showing a significant effect.

The major effect size parental accessions and their hypocotyl length phenotypes correlate with each other. This information is available from the box plots in section 3.3.5.2 which show that e.g. Sf-2 has a bigger phenotype and Col-0 has a smaller phenotype. This reconstruction data from the QTL mapping is congruent to the actual hypocotyl lengths when phenotyped.

3.3.5 Post QTL Mapping Analyses:

The discovery of QTL for the hypocotyl length trait in response to temperature, in the pre-QTL cloning step, is one of the most noteworthy results and an achievement in itself, however, to enhance our molecular understanding of the genomic regions involved in the natural variation of hypocotyl length in response to temperature, the post QTL mapping analysis aimed to identify candidate genes. Due to the fine mapping of QTL, the approach adopted for this purpose was to conduct a detailed analysis of the isolated fine-mapped QTL and subsequent QTL cloning. This allowed the detailed examination of small genomic intervals that may be involved in the natural variation of hypocotyl length.

3.3.5.1 A Three-Dimensional Strategy for Candidate Gene Identification:

The post QTL mapping approach, for the preliminary shortlisting of candidate genes, was based on a three-dimensional strategy. This effective method, which culminated in an exhaustive list of candidates, included:

1. A candidate gene approach by identifying genes functionally involved in the hypocotyl length phenotype
2. Identification of molecular allelic polymorphisms that may be responsible for the functional variation of hypocotyl length phenotype in candidate genes

3. Identification of the major effect size parental accessions for these candidate genes in the QTL showing allelic polymorphism

These three criteria form the basis of an integrated approach to data analysis as follows. The identification of candidate genes was carried out in a systematic way once the QTL were isolated to a relatively narrow interval by fine-QTL mapping. The candidate gene screening comprised of two parallel and complementary methods for each of the QTL at different temperatures. *First* was to retrieve online gene descriptions and orthology from TAIR's website (http://www.arabidopsis.org) and screen for the significant ones according to their known or predicted function in hypocotyl elongation. The lists of genes in each of the QTL were compared against exhaustive lists of genes which have been already implicated in growth and development related phenotypes or characteristics; such as hypocotyl length, temperature response, light response, seedling germination, auxin response, shoot development and developmental regulation. These comprehensive lists, based on phenotypes, were prepared from a study (http://walnut.usc.edu/2010) related to the 1001 genomes project (http://www.1001genomes.org/index.html). Subsequently, more than 130 protein-coding genes, associated to the above mentioned characteristics or phenotypes, were identified in all of the QTL. *Second* was to look in the literature and spot all those pre-identified genes that have functional relationships to the trait; and observe if they are present in the isolated QTL. This method ensured that no significant gene was left out in the screening process. And as it is usually the case, the candidate gene approach was an informed first step in the shortlisting of genes in the QTL.

The *in silico* bioinformatic analysis of genome sequences for identifying allelic sequence polymorphisms in the protein coding regions of the candidate genes in the QTL was a second step of gene shortlisting. The genome sequences of the 19 parental accessions were available and this greatly benefitted the QTL cloning as it reduced the total number of candidate genes in a QTL. All those genes were disregarded and excluded from any further analysis that did not show any allelic polymorphism in any of the 19 accessions in their protein coding regions. This is because these genes are unlikely to be causative of the QTL based on their coding sequence (CDS). This was an important step in the shortlisting of candidate genes. Conversely, those genes and their splice variants were selected for further investigation which had sequence polymorphisms, such as indels and

SNPs in the coding sequences of the genes. The DNA sequences of the candidate genes and their splice variants, if any, in all the 19 parental accessions, were analysed for identification of sequence polymorphisms. However, SNP mutations were not studied in the DNA sequences, because these point mutations could lead to synonymous changes. These types of polymorphisms were studied in greater depth in the protein sequences that were obtained by *in silico* translation of the DNA sequences. The bioinformatic protocol created and used for this analysis is provided in the Appendix (section 3.9).

The various kinds of polymorphisms that were analysed are insertions, deletions, (indels), nonsynonymous mutations such as nonsense and missense mutations, which could potentially lead to a different amino acid substitution in the polymer or a pre-mature truncation stop codon, eventually forming a different protein structure and hence altering its function (loss of function mutations).

The first and second steps of shortlisting identified candidate genes with potentially significant sequence polymorphisms. In the third step it was essential to pinpoint which parents were major contributors to a QTL. To distinguish between minor and major effect size parents, the corresponding box plots for the peak SNP marker and for the flanking markers of the genes in the respective QTL were analysed. The box plots estimate the effects of the parental accession at each locus. The box plots represent parental reconstruction data from the QTL mapping showing hypocotyl length effect sizes. This data revealed candidate genes which showed sequence polymorphisms in the major effect size parents. These are parents which show extreme phenotypes (bigger or smaller) in the box plots. These accessions show the most significant association with the peak SNP markers or the flanking SNP markers in the QTL. Therefore, candidate genes that had polymorphisms in the main effect size parental accessions in a QTL were investigated further for genetic analyses. The shortlist comprised of a total of ~ 30 candidate genes in various QTL.

3.3.5.2 Implementing an Integrated Approach for Gene Discovery:

Following the shortlisting of genes according to the three-tier strategy as explained in section 3.3.5.1, candidate genes were identified in all of the different QTL at all the

temperatures. This section provides the results of the QTL mapping and details of the candidate genes in each of the QTL. For each of the QTL and their candidate genes the molecular polymorphisms have also been identified.

Since all the candidates underwent a rigorous shortlisting methodology, this allowed for very few strong candidate genes to be taken to the next level of functional analysis. In the following sections, the figures with the box plot data show effect sizes of the parental accessions contributing to the variation of the trait at a particular peak SNP marker; and the tables show types of mutations in accessions compared to the reference Col-0 genome.

3.3.5.2.1 QTL Data Analysis at 12°C:

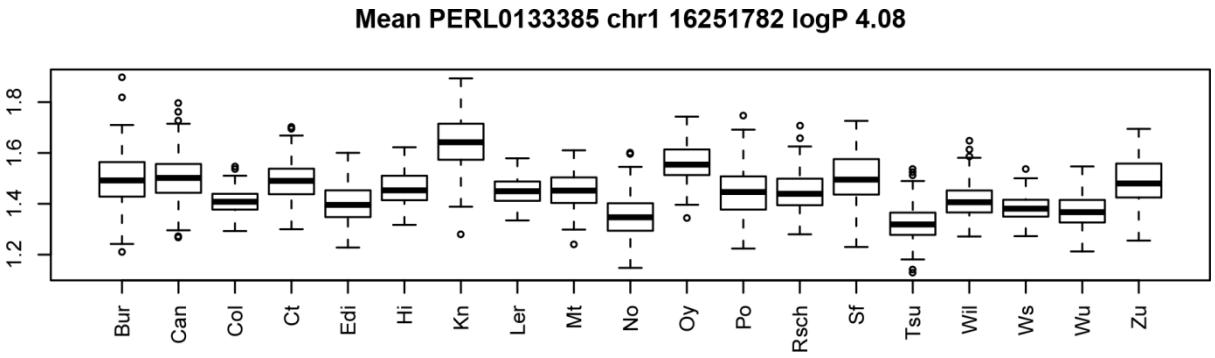

Figure 3.5: Effect sizes of parental accessions at peak SNP marker at Chr1, QTL3

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr1, QTL3	AT1G43040	*SAUR58*, auxin-responsive protein, putative	Kn-0, Wil-2, Zu-0	Stop Codon

Table 3.4: Table of candidate genes at Chr1, QTL3

The *SAUR58* gene belongs to the Small Auxin Up-RNA (SAUR)-like auxin responsive protein family. The total length of the DNA sequence is 315bp. In the accessions Kn-0, Wil-2 and Zu-0 at 70 bp, C => T, a SNP mutation which changes the amino acid Arginine (R) into a truncation stop codon TGA and hence terminates the predicted protein prematurely. This gene is an interesting candidate for the QTL and a T-DNA knockout 'GK-441A08-018217' is available on TAIR's website. However, this gene could not be further functionally characterised as no seeds for knockouts were received from the stock centre.

3.3.5.2.2 QTL Data Analysis at 22°C:

The following data are for Chr2 QTL1:

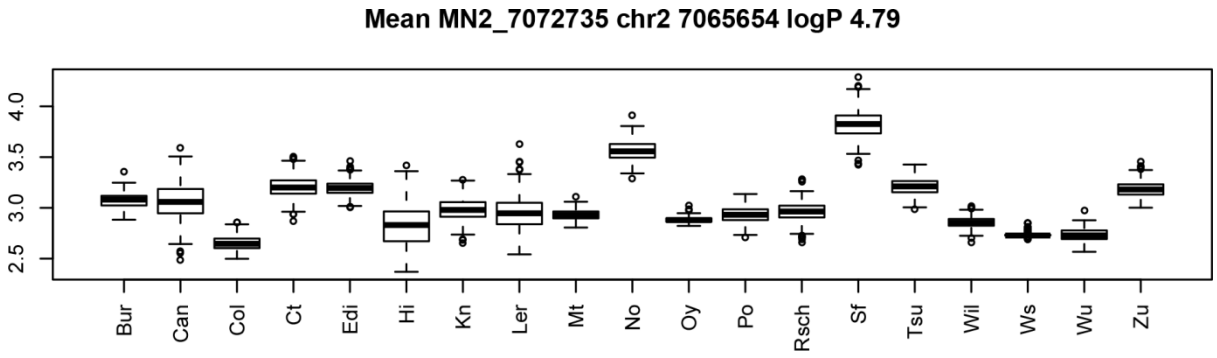

Figure 3.6: Effect sizes of parental accessions at peak SNP marker at Chr2, QTL1

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr2, QTL1	AT2G14820	*NPY2*, phototropic-responsive NPH3 family protein	Ct-1, Kn-0, No-0, Sf-2, Tsu-0, Zu-0	non-synonymous
Chr2, QTL1	AT2G16580	*SAUR8*, auxin-responsive protein, putative	No-0, Sf-2, Tsu-0	non-synonymous

Table 3.5: Table of candidate genes at Chr2, QTL1

Following are the non-synonymous mutations in the candidate gene *NPY2*:

In Ct-1, No-0, Sf-2, Tsu-0 at residue 177 Histidine => Glutamine and at residue 270 Isoleucine => Valine. In Ct-1 at residue 376 Leucine => Phenylalanine and at residue 416 Serine => Isoleucine. In Kn-0 at residue 341 Aspartic acid => Glutamic acid. In Sf-2 and Tsu-0 at residue 569 Isoleucine => Methionine. In Kn-0 at residue 444 Lysine => Asparagine and at residue 598 Glycine => Glutamic acid.

Following are the non-synonymous mutations in the candidate gene *SAUR8*:

In No-0, Sf-2, Tsu-0 at residue 38 Valine => Aspartic acid. In Tsu-0 at residue 50 Proline => Histidine.

The following data are for Chr2 QTL2:

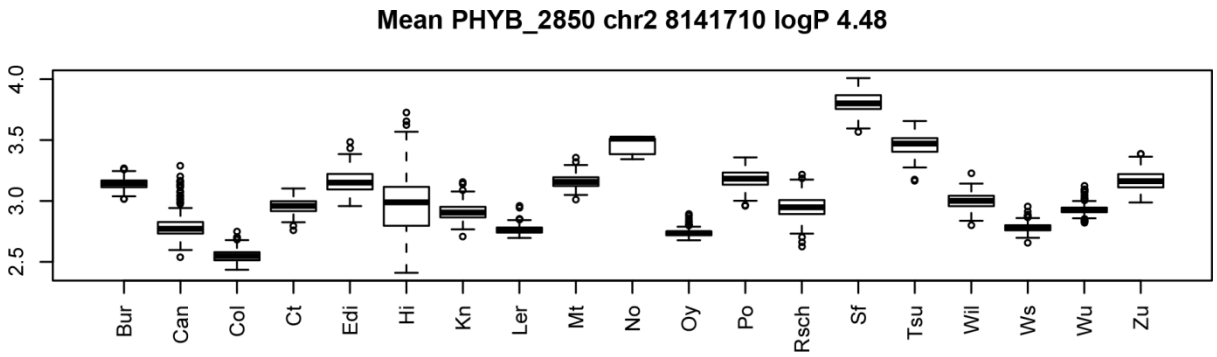

Figure 3.7: Effect sizes of parental accessions at peak SNP marker at Chr2, QTL2

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr2, QTL2	AT2G18790	*PHYB* (PHYTOCHROME B); G-protein coupled photoreceptor/ signal transducer	Ct-1, Kn-0, No-0, Sf-2, Wil-2, Zu-0	non-synonymous

Table 3.6: Table of candidate genes at Chr2, QTL2

Following are the non-synonymous mutations in the candidate gene *PHYB*:

In Kn-0, No-0, Sf-2, Wil-2, Zu-0 there is a 4 residue (GGGR) deletion starting at residue 9. In No-0 at residue 19 Glutamic acid => Lysine and at residue 980 Valine => Isoleucine. In Ct-1 at residue 93 Glycine => Aspartic acid. In Kn-0, No-0, Sf-2, Wil-2, Zu-0 at residue 143 Isoleucine => Leucine. In Kn-0, Wil-2, Sf-2 at residue 247 Alanine => Serine. In Zu-0 at residue 709 Glutamic acid => Lysine. In Sf-2, Wil-2, Kn-0 at residue 736 Serine => Threonine. In Kn-0 and Wil-2 at residue 949 Alanine => Threonine and at residue 989 Glutamic acid => Aspartic acid. In Ct-1, Kn-0, No-0, Sf-2, Wil-2, Zu-0 at residue 1072 Leucine => Valine.

PHYB was an obvious candidate and this QTL has been named as 'Temp22.2' QTL. The QTL has been named according to the temperature environment in which it has been isolated and the chromosome number.

The following data are for Chr2 QTL3:

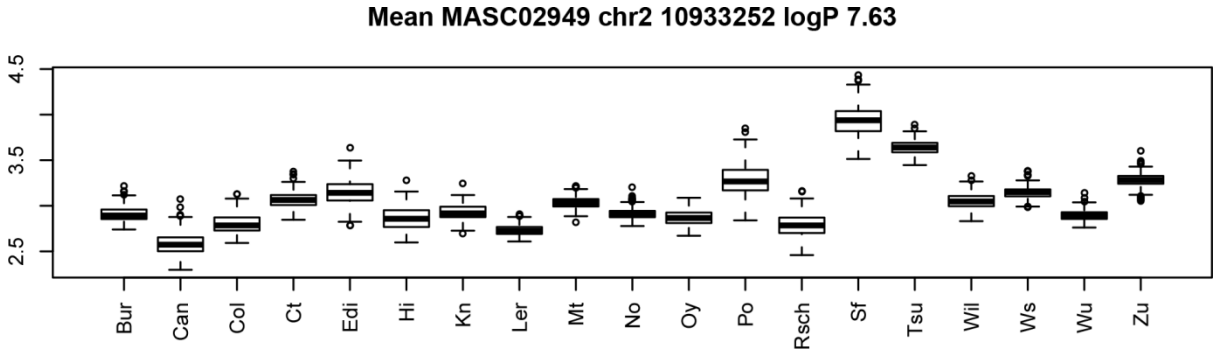

Figure 3.8: Effect sizes of parental accessions at peak SNP marker at Chr2, QTL3

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr2, QTL3	AT2G23050	*NPY4*, phototropic-responsive NPH3 family protein	Can-0, Kn-0, No-0, Wil-2, Ct-1, Sf-2	non-synonymous
Chr2, QTL3	AT2G23380	*CLF* (CURLY LEAF); transcription factor	Ct-1, Sf-2, No-0, Tsu-0	non-synonymous
Chr2, QTL3	AT2G24400	*SAUR38*, auxin-responsive protein, putative / small auxin up RNA (SAUR_D)	Can-0, Kn-0, Sf-2, Mt-0, Tsu-0, Wil-2, Ws-0	non-synonymous
Chr2, QTL3	AT2G24790	*COL3* (CONSTANS-LIKE 3); protein binding / transcription factor/ zinc ion binding	Edi-0, Bur-0, No-0, Tsu-0, Wil-2, Ct-1	non-synonymous
Chr2, QTL3	AT2G27380	*ATEPR1 Arabidopsis thaliana* extensin proline-rich 1	Sf-2, Tsu-0, Zu-0	non-synonymous
Chr2, QTL3	AT2G28350	*ARF10* (AUXIN RESPONSE FACTOR 10); miRNA binding / transcription factor	Sf-2, Tsu-0, Ct-1	non-synonymous
Chr2, QTL3	AT2G28890	*PLL4* (POLTERGEIST LIKE 4); protein serine/threonine phosphatase	Tsu-0, Bur-0, Zu-0, Ct-1, Sf-2, Can-0	non-synonymous
Chr2, QTL3	AT2G30520	*RPT2* (ROOT PHOTOTROPISM 2)	Sf-2, Po-0, Zu-0, Can-0	non-synonymous
Chr2, QTL3	AT2G30950	*VAR2* (VARIEGATED 2); ATP-dependent peptidase/ ATPase/ metallopeptidase/ zinc ion binding	Tsu-0, Edi-0, Oy-0, Bur-0	non-synonymous

Table 3.7: Table of candidate genes at Chr2, QTL3

Following are the non-synonymous mutations in the candidate gene *NPY4*:

In Can-0, Kn-0, No-0, Wil-2, an insertion of T at residue 61 causes a frame shift mutation.

In the same accessions at residue 245 Alanine => Threonine; at residue 257 Glutamic acid

=> Glycine; at residue 268 Asparagine => Aspartic acid; at residue 305 Lysine => Arginine; at residue 321 Aspartic acid => Histidine; and at residue 339 Asparagine => Aspartic acid. In Ct-1 at residue 14 Leucine => Glutamine, at residue 268 Asparagine => Aspartic acid. In Sf-2 at residue 14 Leucine => Glutamine.

Following are the non-synonymous mutations in the candidate gene *CLF*:

In Ct-1, No-0 at residue 204 Tyrosine => Serine, at residue 408 Arginine => Glutamine, at residue 453 Histidine => Asparagine, at residue 510 Serine => Asparagine, at residue 542 Leucine => Phenylalanine. In Sf-2, Tsu-0 at residue 225 Serine = Asparagine, there is an insertion of two residues at 337 Threonine and 338 Glycine, and at residue 415 Lysine => Asparagine.

Following are the non-synonymous mutations in the candidate gene *SAUR38*:

In Can-0 at residue 48 Serine => Glycine. In Kn-0, Sf-2 at residue 121 Glutamic acid => Aspartic acid. In Mt-0, Tsu-0, Wil-2, Ws-0 at residue 156 Serine => Arginine.

Following are the non-synonymous mutations in the candidate gene *COL3*:

In Edi-0, Bur-0, No-0 at residue 30 Aspartic acid => Asparagine, at residue 286 Glycine => Cysteine. In Tsu-0, Wil-2 at residue 64 Lysine => Glutamic acid. In Tsu-0 at residue 182 Isoleucine => Valine. In Ct-1 at residue 273 Aspartic acid => Glutamic acid.

Following are the non-synonymous mutations in the candidate gene *EPR1*:

In Zu-0 at residue 274 Valine => Leucine, at residue 723 Isoleucine => Threonine. In Tsu-0 at residue 280 Isoleucine => Threonine, at residue 623 Proline => Serine. In Sf-2 at residue 307 Valine => Isoleucine, at resuide 422 Leucine => Proline, at residue 509 Glutamine => Lysine, at residue 558 Isoleucine => Valine, at residue 710 Proline => Glutamine.

Following are the non-synonymous mutations in the candidate gene *ARF10*:

In Sf-2, Tsu-0, Ct-1 at residue 411 Histidine => Aspartic acid.

Following are the non-synonymous mutations in the candidate gene *PLL4*:

In Tsu-0 at residue 436 Alanine => Proline. In Bur-0, Tsu-0 at residue 327 Aspartic acid => Histidine. In Zu-0, Ct-1, Sf-2, Tsu-0 at residue 363 Leucine => Serine. In Zu-0 at residue 315 Alanine => Glycine. In Can-0 at residue 208 Proline => Serine.

Following are the non-synonymous mutations in the candidate gene *RPT2*:

In Po-0, Sf-2, Zu-0 at residue 244 Tyrosine => Phenylalanine. In Can-0 at residue 458 Isoleucine => Valine. In Sf-2 at residue 555 Lysine => Arginine.

Following are the non-synonymous mutations in the candidate gene *VAR2*:

In Edi-0, Tsu-0, Oy-0, Bur-0 at residue 14 Serine => Leucine.

The following data are for Chr3 QTL3:

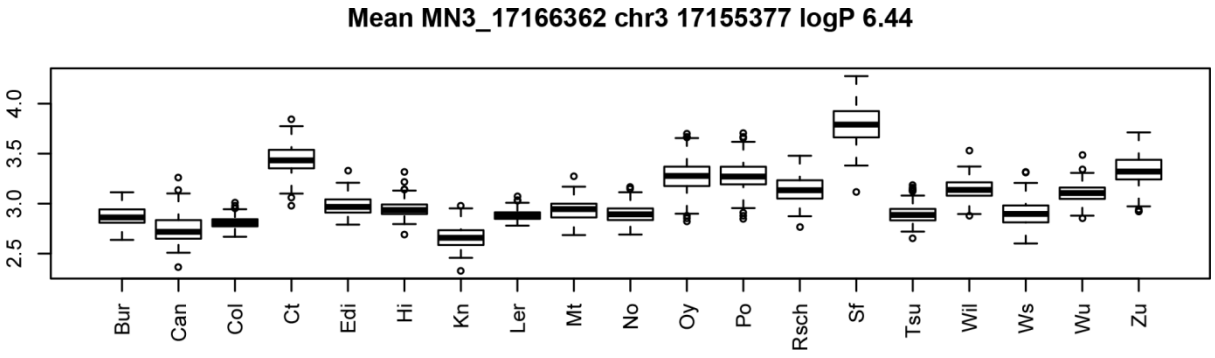

Figure 3.9: Effect sizes of parental accessions at peak SNP marker at Chr3, QTL3

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr3, QTL3	AT3G44600	*CYP71* (CYCLOPHILIN71); chromatin binding / histone binding / peptidyl-prolyl cis-trans isomerase	Ct-1, Sf-2, Tsu-0, Zu-0, Oy-0, Po-0, Bur-0	non-synonymous
Chr3, QTL3	AT3G45780	*PHOT1* (phototropin 1); kinase	Zu-0, Can-0, Oy-0, Po-0, Sf-2	non-synonymous
Chr3, QTL3	AT3G49120	*PCB*, defense response to bacterium, defense	Sf-2, Zu-0, Mt-0, Ct-1, Oy-0,	non-synonymous

		response to fungus, response to light stimulus,	Po-0	
Chr3, QTL3	AT3G49670	*BAM2* (big apical meristem 2); ATP binding / protein serine/threonine kinase	Sf-2, Can-0, Oy-0, Po-0	non-synonymous
Chr3, QTL3	AT3G51200	*SAUR18*, auxin-responsive family protein	Sf-2, Mt-0, Can-0, Bur-0, Ler-0	non-synonymous
Chr3, QTL3	AT3G51240	*TT6*, F3H (TRANSPARENT TESTA 6); naringenin 3-dioxygenase	Sf-2, Zu-0, Edi-0, Ct-1, Oy-0, Po-0	non-synonymous
Chr3, QTL3	AT3G52910	*ATGRF4* (GROWTH-REGULATING FACTOR 4)	Zu-0, Can-0, Oy-0, P-0, Ct-1, Sf-2	non-synonymous

Table 3.8: Table of candidate genes at Chr3, QTL3

Following are the non-synonymous mutations in the candidate gene *CYP71*:

In Ct-1, Oy-0, Po-0 at residue 9 Glycine => Glutamic acid, at residue 357 Phenylalanine => Leucine. Bur-0, Tsu-0 at residue 9 Glycine => Arginine. In Zu-0 at residue 181 Leucine => Serine, at residue 278 Alanine => Serine, at residue 357 Phenylalanine => Leucine. In Sf-2 at residue 357 Phenylalanine => Leucine.

Following are the non-synonymous mutations in the candidate gene *PHOT1*:

In Can-0 at residue 41 Phenylalanine => Tyrosine, at residue 365 Aspartic acid => Asparagine. In Oy-0, Po-0 at residue 67 Proline => Leucine. In Ct-1 at residue 79 Isoleucine => Serine. In Zu-0 at residue 268 Lysine => Asparagine. In Bur-0 at residue 836 Glutamine => Glutamic acid.

Following are the non-synonymous mutations in the candidate gene *PCB*:

In Mt-0 at residue 53 Valine => Glutamic acid. In Ct-1, Oy-0, Po-0, Sf-2, Zu-0 at residue 270 Arginine => Glutamine.

Following are the non-synonymous mutations in the candidate gene *BAM2*:

In Can-0 at residue 39 Threonine => Alanine, at residue 77 Serine => Proline, at residue 184 Alanine => Proline, at residue 353 Glutamine => Lysine, at residue 983 Alanine =>

Serine. In Sf-2 at residue 130 Asparagine => Aspartic acid. In Oy-0, Po-0 at residue 353 Glutamine => Histidine.

Following are the non-synonymous mutations in the candidate gene *SAUR18*:

Sf-2, Mt-0, Can-0, Bur-0 at residue 44 Tyrosine => Histidine. In Ler-0 at residue 21 Glutamine => Lysine.

Following are the non-synonymous mutations in the candidate gene *TT6*:

In Edi-0 at residue 11 Glycine => Arginine. In Ct-1, Sf-2, Oy-0, Po-0 at residue 148 Aspartic acid => Asparagine, at residue 350 Aspartic acid => Alanine. In Zu-0 at residue 151 Glutamic acid => Lysine. In Oy-0, Po-0 at residue 158 Glutamic acid => Lysine.

Following are the non-synonymous mutations in the candidate gene *GRF4*:

In Can-0, Zu-0 at residue 204 Alanine => Threonine, at residue 380 Histidine => Glutamine. In Oy-0, Po-0 at residue 380 Histidine => Glutamine. In Ct-1, Sf-2 at residue 120 Leucine => Isoleucine.

3.3.5.2.3 Data Analysis at 27°C:

The following data are for Chr4 QTL1:

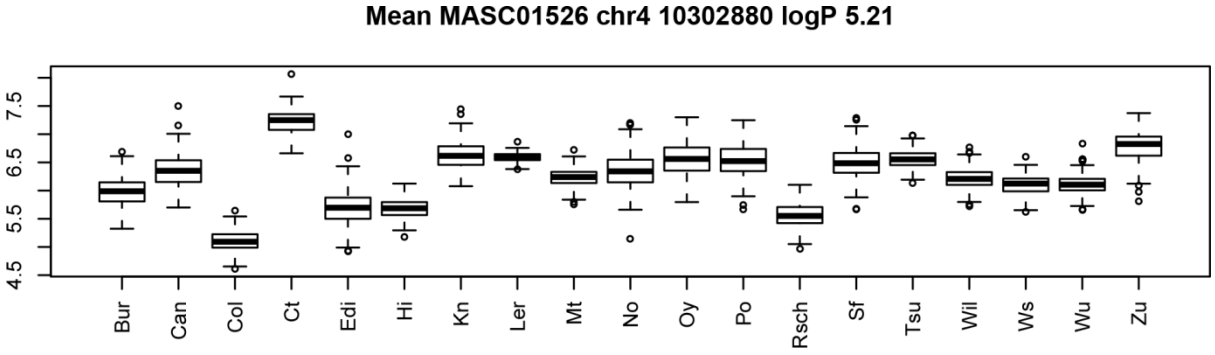

Figure 3.10: Effect sizes of parental accessions at peak SNP marker at Chr4, QTL1

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr4, QTL1	AT4G18130	*PHYE*, peptidyl-histidine phosphorylation, protein-chromophore linkage, red, far-red phototransduction,	Ct-1, Can-0, Sf-2, Zu-0	non-synonymous
Chr4, QTL1	AT4G18610	*LSH9* flower, hypocotyl, inflorescence meristem, leaf apex, root, shoot apex	Wil-2, Ct-1, No-0, Oy-0, Sf-2, Tsu-0, Zu-0	non-synonymous
Chr4, QTL1	AT4G18710	*BIN2* (BRASSINOSTEROID-INSENSITIVE 2); kinase	Tsu-0,	non-synonymous
Chr4, QTL1	AT4G18780	*CESA8* (CELLULOSE SYNTHASE 8); cellulose synthase/ transferase, transferring glycosyl groups	Ct-1, Bur-0, Sf-2, Wu-0	non-synonymous
Chr4, QTL1	AT4G19600	*CYCT1*;4; cyclin-dependent protein kinase	Zu-0, Can-0, Mt-0, Tsu-0, Bur-0	non-synonymous

Table 3.9: Table of candidate genes at Chr4, QTL1

Following are the non-synonymous mutations in the candidate gene *PHYE*:

In Can-0 at residue 261 Glutamic acid => Glycine. In Ct-1, Sf-2, Zu-0 at residue 499 Phenylalanine => Isoleucine. In Ct-1 at residue 686 Tyrosine => Serine, at residue 878 Valine => Isoleucine, at residue 900 Glutamic acid => Aspartic acid, at residue 910 Glutamic acid => Aspartic acid, at residue 925 Serine => Glycine, at residue 959 Isoleucine => Methionine, at residue 977 Threonine => Alanine, at residue 1016 Glycine => Arginine, at residue 1104 Glycine => Valine. In Mt-0 977 Threonine => Alanine, at residue 1016 Glycine => Arginine.

Following are the non-synonymous mutations in the candidate gene *LSH9*:

In Wil-2 at residue 186 Proline => Leucine. In Ct-1, No-0, Oy-0, Sf-2, Tsu-0, Zu-0 at residue 173 Valine => Methionine.

Following are the non-synonymous mutations in the candidate gene *BIN2*:

In Tsu-0 at residue 53 Isoleucine => Valine.

Following are the non-synonymous mutations in the candidate gene *CESA8*:

In Sf-2 at residue 24 Phenylalanine => Valine, at residue 85 Serine => Proline. In Ct-1 at residue 145 Glutamic acid => Aspartic acid, at residue 157 Threonine => Methionine. In Bur-0, Wu-0 at residue 144 Histidine = Glutamine.

Following are the non-synonymous mutations in the candidate gene *CYCT1*:

In Zu-0 at residue 14 Serine => Isoleucine, at residue 364 Alanine => Valine, at residue 501 Leucine => Phenylalanine. In Can-0 at residue 313 Serine => Leucine, at residue 529 Tyrosine => Histidine. In Mt-0 at residue 386 Valine => Alanine. In Tsu-0 at residue 400 Isoleucine => Leucine. In Bur-0 at residue 501 Leucine => Phenylalanine.

The following data are for Chr4 QTL2:

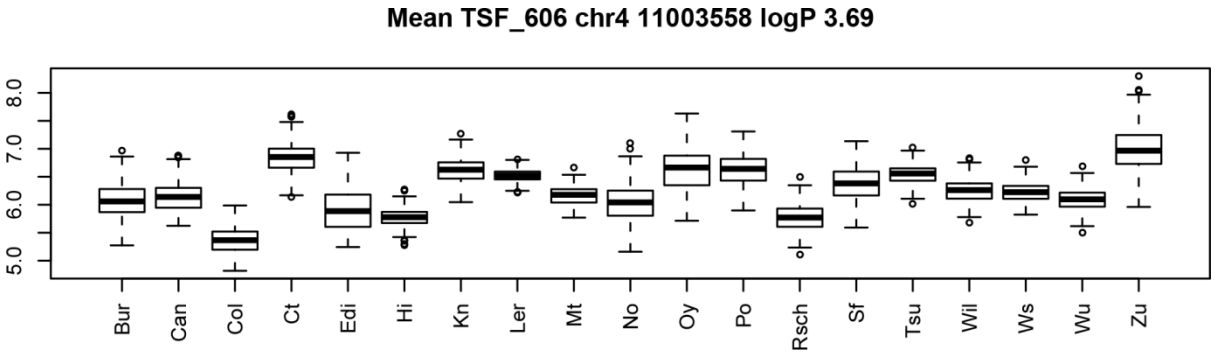

Figure 3.11: Effect sizes of parental accessions at peak SNP marker at Chr4, QTL2

QTL	Gene Locus	Gene Name and Description from TAIR	Major Effect Parents	Type of Mutation
Chr4, QTL2	AT4G20370	*TSF* (TWIN SISTER OF FT); phosphatidylethanolamine binding	Can-0, Kn-0	Non-synonymous

Table 3.10: Table of candidate genes at Chr4, QTL2

Following are the non-synonymous mutations in the candidate gene *TSF*:

In Can-0 at residue 12 Glycine => Serine, at residue 46 Serine => Phenylalanine. In Kn-0 at residue 24 Arginine => Serine.

3.3.6 Discovery of Environmental QTL (G x E Interaction):

As shown in the previous chapter, hypocotyl length as a complex trait is also influenced by the environment. Temperature has a significant effect on hypocotyl length. Due to the fact that the phenotyping of the RILs was carried out in a range of temperature environments, QTL for increased temperature responses could be identified. This analysis aims at the following questions: What are the genomic regions (QTL) that show significant effects to changes in temperature? What are the loci that are associated with the temperature responsiveness of the seedlings? The genetic basis of the phenotypic G x E interactions in hypocotyl length to increased temperatures has been investigated and analysed in this section. As mentioned in section 3.1, the interest is to map QTL for a 'temperature responsiveness trait'. These QTL can also be referred to as 'environmental QTL'.

We can treat the ratio of hypocotyl length at two different temperatures as a quantitative trait and carry out a QTL analysis to identify candidate genes for temperature responsiveness. This trait can be quantitatively calculated when the phenotype is a product of the ratio of two temperature environments. The temperature responsiveness trait shows the genetic interaction between the loci and the temperature environments.

To explore the overall amount of variation present in the MAGIC RILs for hypocotyl length in response to temperature, a two-way ANOVA was performed. The results reveal that RILs are significantly different $F(433, 17457) = 3692.36$, $p < 0.001$), the effect of temperature on the RILs is significant $F(3, 17457) = 4034818.84$, $p < 0.001$) and the G x E interaction is also highly significant $F(1150, 17457) = 1705.32$, $p < 0.001$). (ANOVA table 3.1 in appendix).

For RILs, the partial eta squared showed the effect sizes which are: for genotype = 0.98, for temperature = 0.99 and for G x E = 0.99. This shows that hypocotyl length in the MAGIC RILs is a highly heritable trait. Temperature has a very high effect on hypocotyl length and it is interesting to find genes at individual temperature treatments and also in temperature response, as the G x E both are interacting and affecting the hypocotyl length. The range of differences in hypocotyl length of the RILs at various temperature ratios is shown in table 3.11.

Relative Responsiveness Index (RRI)	Range of hypocotyl length
27°C/22°C	1.11 – 4.51 mm
27°C/17°C	1.17 – 6.07 mm
27°C/12°C	1.45 – 8.52 mm
22°C/17°C	1.07 – 3.97 mm
22°C/12°C	1.12 – 5.26 mm
17°C/12°C	1.05 – 2.82 mm

Table 3.11: Range of hypocotyl length of the RILs at each of the RRI combinations showing very high responsiveness to temperature

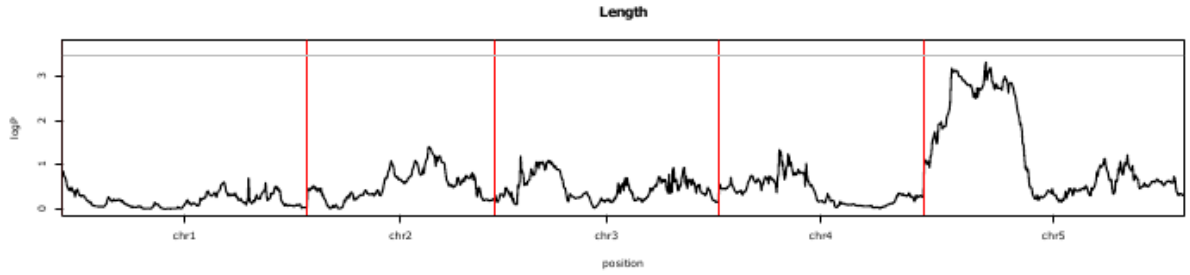

Figure 3.12: QTL scan results for temperature ratio 12°C/17°C

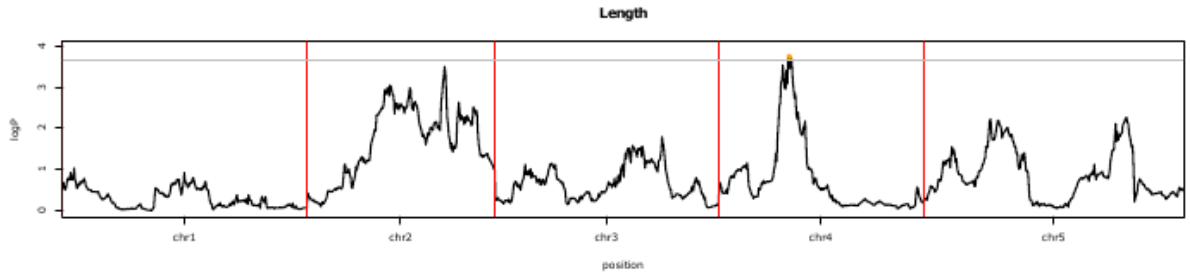

Figure 3.13: QTL scan results for temperature ratio 12°C/22°C

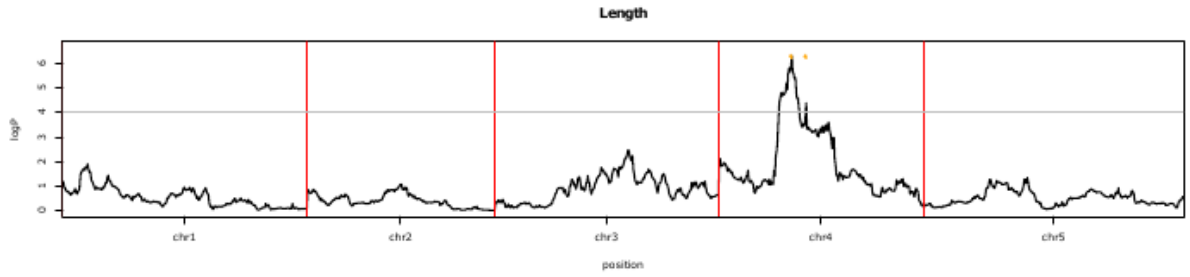

Figure 3.14: QTL scan results for temperature ratio 12°C/27°C

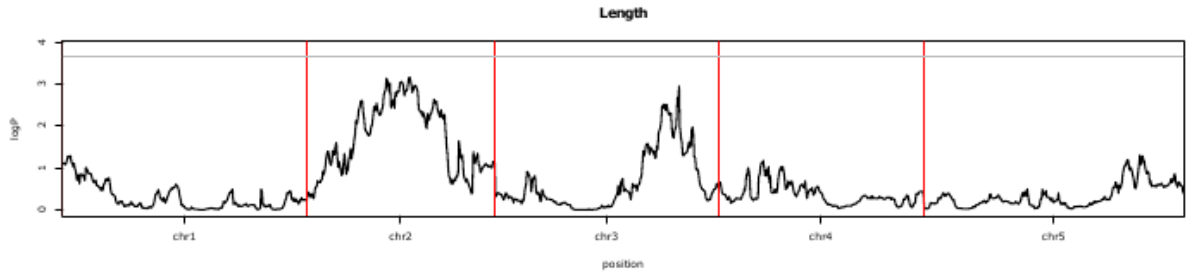

Figure 3.15: QTL scan results for temperature ratio 22°C/17°C

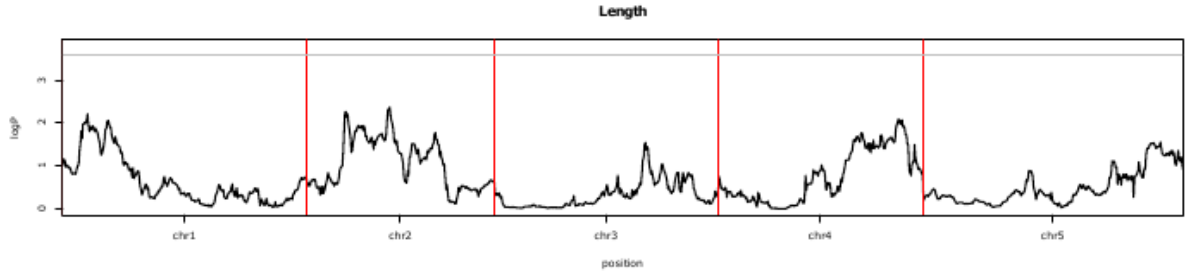

Figure 3.16: QTL scan results for temperature ratio 22°C/27°C

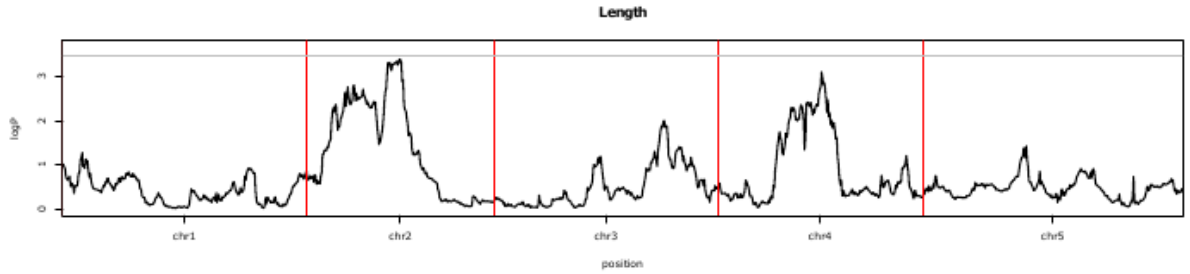

Figure 3.17: QTL scan results for temperature ratio 27°C/17°C

phenotype	chromosome	island.from.bp	island.to.bp	peak.bp	peak.SNP	logP	genomewide.pvalue
Length.12.22	chr4	8048343	8078653	8078653	MN4_8078648	3.766076634	0.0442
Length.12.22	chr4	8177688	8482332	8482332	NMSNP4_8482327	3.746734574	0.0463
phenotype	chromosome	island.from.bp	island.to.bp	peak.bp	peak.SNP	logP	genomewide.pvalue
Length.12.27	chr4	6523087	9121201	8789187	MN4_8789187	6.332249203	0.0017
Length.12.27	chr4	9198009	9198484	9198484	PHYD_1815	4.408593757	0.0278

Table 3.12: Four environmental QTL isolated at 22°C/12°C and 27°C/12°C for the temperature responsiveness trait

RRI	QTL Name	QTL Interval in bp	QTL Interval in Kb	QTL Interval in Mb	QTL Interval in cM	No. of Genes
12°C/22°C	Chr4, QTL1	30310	30.31	0.03031	0.15	9
12°C/22°C	Chr4, QTL2	304644	304.644	0.304644	1.52	82
12°C/27°C	Chr4, QTL1	2598114	2598.114	2.598114	12.99	640
12°C/27°C	Chr4, QTL2	475	0.475	0.000475	0.0023	1

Table 3.13: QTL intervals in bp, Kb and Mb and total number of protein-coding genes in each of the environmental QTL

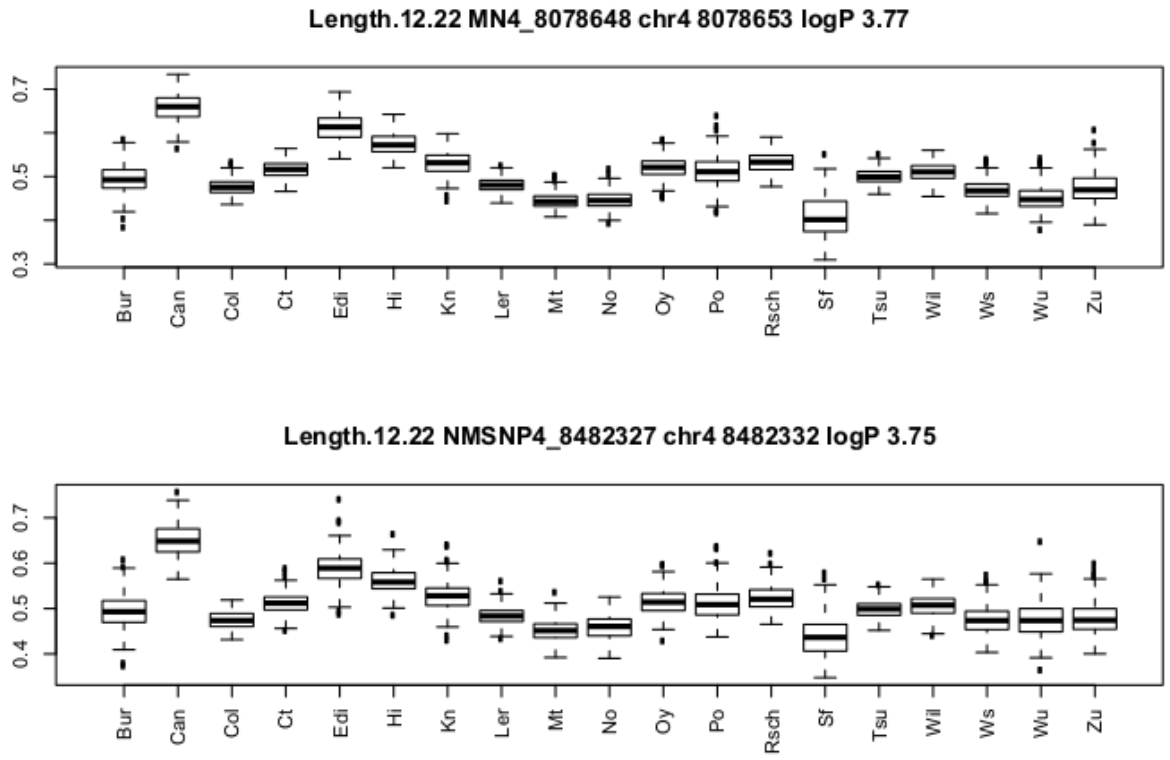

Figure 3.12: Effect sizes of parental accessions at peak SNP markers at ratio 22°C/12°C for QTL1 (top) and QTL2 (bottom)

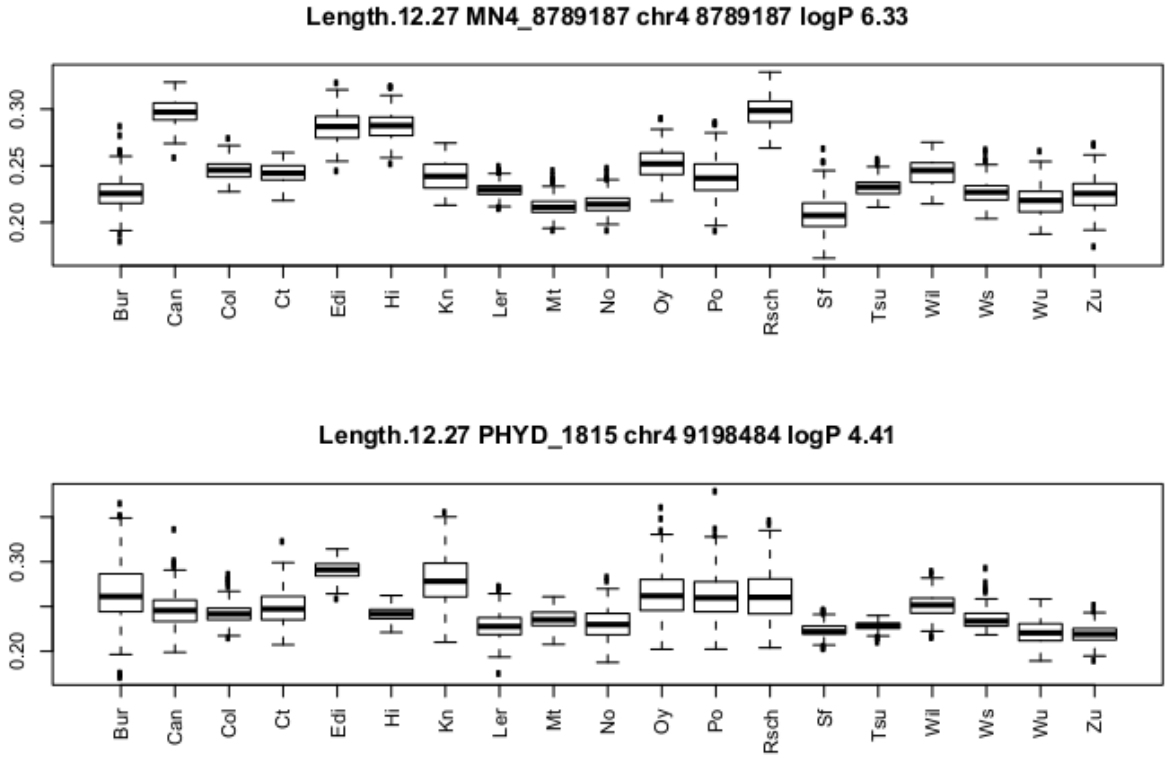

Figure 3.13: Effect sizes of parental accessions at peak SNP markers at 27°C/12°C for QTL1 (top) and QTL2 (bottom)

3.6 Discussion:

The approach of using naturally occurring variation has been exploited by using the MAGIC homozygous RILs population for the dissection of hypocotyl length in response to temperature. The genetic basis of natural variation in hypocotyl length in response to temperature was determined by QTL analysis. The principle of QTL mapping in the MAGIC lines is based on the genotyping of the SNP markers of the progeny obtained from the inter-crossing of distinct accessions for the trait. It requires an experimental population (MAGIC) where genetic variation has segregated between the lines.

As a useful resource, the whole genome sequences of the 19 parental accessions were utilised effectively with the help of a high-throughput bioinformatics approach. *In silico* screens were performed for non-synonymous mutations that allowed the quick identification of candidate genes with polymorphisms. Such remarkable advances in technology and resources facilitated the search for underlying genes. Candidate genes were selected based on sequence polymorphism data, such as indels and SNPs in the coding sequences of the genes.

Two essential features for detection of QTL are power and precision. Power means the probability of isolating a QTL which is segregating and precision is the location error connected to the mapped QTL and the actual QTL.

The main method to identify the genetic basis of ecological and evolutionary complex traits is through genotype-phenotype association. For the identification of the loci that underpin natural variation in complex traits, this approach statistically associates genotypic and phenotypic variation. These loci can be identified by association mapping which uses historical recombination events (Flint-Garcia *et al*., 2003) in natural populations or by QTL mapping which exploits recent recombination events in RILs.

QTL mapping is a well established method in the field of genetics dealing with quantitative traits in plants (Hackett, 2002), whereas the association mapping method has been more recently introduced in plant genetic analyses and is still being significantly developed and improved. The first genome-wide association (GWA) mapping study in Arabidopsis has been recently reported (Atwell *et al*., 2010).

Hence the analysis of natural variation can be performed on natural populations as well as synthetic populations. In the approach that uses natural populations to dissect QTL using association mapping, the analysis looks for associations between SNPs and trait variation across natural accessions (Atwell *et al*., 2010). It exploits linkage disequilibrium (LD), which is the non-random association of alleles, to detect QTL in natural populations. The LD for association mapping is much shorter than in RILs and in principle it isolates fine mapped intervals.

One of the benefits is that it tackles wider natural variation as large numbers of accessions are used in the analysis. Another benefit is that it provides increased precision as compared to family based linkage mapping (Mackay and Powell, 2007) as QTL detection is based on correlation between a trait and a molecular marker.

A major drawback is presence of increased rate of false positives due to population structure that has to be controlled for (Yu *et al*., 2006). In association studies the population structure may be controlled but with a trade-off of reducing the power to detect associations between the genotype and the phenotype (Atwell *et al*., 2010).

Another problem with association mapping is that it usually detects the variants that are common in populations of wild accessions, and therefore may not pick out rare alleles. In general, association mapping provides increased precision but has lower power for isolation of QTL.

On the other hand, artificial populations are used to dissect QTL. These experimental segregating populations are derived from natural populations by controlled crosses. The detection of QTL can only take place in genomic regions for which polymorphisms in the parental accessions exist. The classical method of generating an experimental population (RILs or F2s) is by crossing two parents. These are called biparental populations. Even though inbred lines with large phenotypic differences have been used in classical biparental QTL mappings (Doerge, 2002), yet the genetic variation is limited to two lines. And it isolates QTL intervals at a low mapping resolution of 10–30 cM (Cavanagh *et al*., 2008) or ranging from 5–50 cM (Alonso-Blanco *et al*., 2005) with hundreds to thousands of protein-coding genes. Consequently, there are relatively few traditional QTL studies that have determined the genes and nucleotide polymorphisms underlying QTL (Flint and

Mott, 2001). Examples of biparental mapping populations are Col-0 x Ler (Lister and Dean, 1993) and Bay-0 x Shahdara (Loudet *et al*., 2002).

A higher mapping resolution is directly proportional to the number of recombination events in these types of mapping populations. The higher the number of recombination events, the higher will be the mapping resolution. That is why in QTL mapping where biparental populations are used, there are lower number of recombinations leading to lower mapping resolution which is a drawback. Conversely, in association mapping since natural accessions are used, the historical recombination events are greater leading to higher mapping resolution. However, as described above this method has other drawbacks.

In QTL mapping, in order to overcome this limitation of low resolution, greater numbers of recombinations are required. This can be accomplished by increasing the number of crosses in a population. An extension of the classical RILs populations derived from biparental crosses were proposed by Darvasi and Soller (1995) known as the advanced intercross (AIC) population aiming at improving QTL location estimates.

The advanced intercross lines can be constructed from two parents (Balasubramanian *et al*., 2009 and Balint-Kurti *et al*., 2010). In this method, the F2 progeny arising from the biparental cross are intercrossed to increase recombination events. This is followed by selfing, which produce RILs. With this approach, in these lines, although the mapping resolution is increased, yet the issue of reduced genetic diversity is not addressed as it is still biparental. These methods have shortcomings in identifying candidate genes and subsequently in the post QTL mapping process of QTL cloning.

Therefore, in QTL mapping, a next-generation approach that maintains a high mapping resolution, and at the same time overcomes the issues of reduced genetic variation in biparentals and population structure problems in natural accessions, is to use multiparental populations with advanced intercrosses. This allows the inclusion of a bigger fraction of the genetic variation present in natural populations. This type of experimental population is a further extension to the advanced intercross lines. The multiparent intercrossing lines undergo additional recombination events, which allow isolation of QTL with smaller intervals. In plants, Cavanagh *et al*. (2008) proposed

multiparent advanced generation intercross (MAGIC) populations obtained from crossing multiple parents and have described the usefulness of MAGIC populations for QTL mapping.

The MAGIC population was initially used in mapping genes in mice also called the heterogeneous stock which isolated QTL with small confidence intervals (Yalcin *et al*., 2005). With the help of this RIL population, derived from multiparents, the power to detect precise QTL has been demonstrated and is applicable to plant genetics (Valder *et al*., 2006). In Arabidopsis examples of populations obtained from multiple parents are Arabidopsis Multi Parent RIL (AMPRIL) (Huang *et al*., 2011) and MAGIC which has been used for fine mapping of quantitative traits (Kover *et al*., 2009).

MAGIC lines have been shown to be a better mapping population to identify QTL intervals, underlying quantitative traits, due to its higher mapping accuracy and detection in comparison to biparental mapping populations (Kover, 2009). The 527 RILs of this population and the 19 parents have been genotyped with 1260 single nucleotide polymorphisms (SNPs) across the whole genome. With the presence of the 1260 SNPs the MAGIC lines offer better resolution to localise candidate loci.

In the MAGIC lines the added rounds of progeny crosses result in greater recombination events leading to decay in LD. This in turn increases the precision of a QTL location (Richard Mott, personal meeting). The QTL mapping in this chapter using MAGIC lines has isolated fewer QTL of large effects than traditional RILs which is in congruence to a previous study that has used the multiparent AMPRIL population (Huang *et al*., 2011).

MAGIC lines have benefits of no population structure and increased genetic diversity. Due to the prevalence of greater number of alleles, from many parents, segregating in the population, QTL have been mapped to significantly smaller QTL intervals as has been demonstrated in this chapter (table 3.2). These RILs possess an increased genetic heterogeneity, by being a fine genetic mosaic, leading to a higher mapping resolution property.

The MAGIC lines are a new generation of multiparent genetic mapping population which could be considered as next-generation experimental populations. The 19 MAGIC accessions have been shown to exhibit extensive natural variation for disease resistance

(Kover and Schaal, 2002). MAGIC lines have also been used to validate previously identified QTL and novel QTL have been identified for the traits of glaborous and erecta (Kover *et al*., 2009). They offer exploitation of more natural variation by using multiple parents with large phenotypic differences. In terms of the genetic variation present in MAGIC lines, the 19 parental accessions represent the amount of commonly available molecular variation present in Arabidopsis (Kover and Mott, 2012). This suggests that although 19 accessions were used to make the MAGIC RILs, yet they do capture a significant proportion of the common molecular variation present in the species.

In comparison to a large set of natural accessions used for association mapping, although MAGIC RILs experience less recombination and thus their mapping resolution is lower, yet there are less false positive QTL, which could lead to identifying the actual loci involved in the variation of the trait.

In summary, for dissecting the genetic basis of hypocotyl length trait, the MAGIC population is ideal for it represents diversity, has no population structure and allows fine mapping of QTL intervals. In this study the MAGIC lines have proved effective in the discovery and subsequent characterisation of the underlying allelic variation for hypocotyl length, a complex trait. It has also been proposed to use the MAGIC approach in all crop plants (Mackay and Powell, 2007).

In this study, QTL were mapped for each temperature environment independently. The QTL mapping has identified loci that are associated with variation in hypocotyl length. This mapping method was found to be very informative and it revealed that these QTL are temperature sensitive as they appear only at individual temperatures suggesting that different regions of the genome control natural variation in hypocotyl length at different temperatures.

This finding illustrates that different functional genes are involved in controlling natural variation in hypocotyl length at various temperatures. And therefore, it signifies the relevance of understanding the genetic basis of growth at different temperature regimes. It also shows that any given temperature treatment and the QTL are interacting at this specific environmental condition and are affecting the variation in the elongation of the hypocotyl. The individual temperature environment played a role in determining which

parts of the genome show significant associations with the phenotype. From the comparison of associations between genotype and phenotype across different temperature environments, it was seen that very different sets of loci were associated with hypocotyl length.

The importance of carrying out genetic analyses across a range of environments has been previously emphasised (Nicotra *et al*., 2010). G x E interaction studies in Arabidopsis can identify genomic regions involved in responses to various abiotic environmental factors. Most interaction studies in Arabidopsis have focused on the G x E interactions to abiotic nutrient responses (Rauh *et al*., 2002), (Loudet *et al*., 2003), (El-Lithy *et al*., 2006), (Ghandilyan *et al*., 2009), (Prinzenberg *et al*., 2010). A recent study in Arabidopsis also found different loci for flowering time under different laboratory and field conditions (Brachi *et al*., 2010). These studies reveal the importance of multi-environment phenotyping and QTL mapping. The G x E interactions in these studies indicate that such interaction is likely to be present in other environmental abiotic factors including temperature. Moreover, the data on phenotyping of accessions in response to increased temperatures in chapter 2 is a strong indication of possible QTL x Temperature Environment (Q x E_T) interactions for hypocotyl length.

The genetic basis of phenotypic plasticity and the intraspecific variation in plastic responses to temperature present in Arabidopsis accessions in hypocotyl length is poorly understood. One way of developing our understanding of plasticity is to identify genes and QTL involved in plastic responses (Bradshaw, 2006). The hypocotyl length is regulated by environmental factors such as temperature, light quality and quantity, photoperiod, etc, and there is a lack of studies investigating significant Q x E_T interactions.

In this study, significant Q x E_T interactions have been, for the first time, detected between hypocotyl length and temperature responsiveness under different controlled temperature environments. For an understanding of the G x E_T it was of interest to know what the effect of the temperature environment is in determining associations between the genotype and phenotype.

While the QTL cloning was being carried out to identify the molecular basis of variation in hypocotyl length, the data for the environmental QTL mapping were also analysed. In

addition to QTL isolation at individual temperatures, 'environmental QTL' were mapped for increased temperature responses of hypocotyl length looking at G X E_T interaction. The environmental QTL have been mapped for a trait termed as a 'temperature responsiveness trait' that is calculated as the ratio of hypocotyl length at one temperature to the hypocotyl length at another temperature.

This method of QTL analysis proved very useful as it dissects the variation in hypocotyl to temperature responsiveness. And since the phenotyping of the RILs was performed at four different temperatures, a combination of six temperature responses was carried out. The purpose of such QTL mapping was to address the question: what are the genomic regions that control variation in hypocotyl elongation responses to increased temperatures? The resulting analysis, to address this question, aims at comprehending the G x E interaction. It was found that a few high resolution QTL appeared in this analysis suggesting that there are a small number of loci on the genome that are involved in controlling variation in hypocotyl length to increased temperature responses in natural accessions.

One such very high resolution QTL in the sub cM range harbours a single phytochrome gene, *PHYD* and sequence alignment analysis in this gene has revealed important polymorphisms such as indels in some of the MAGIC accessions (Hi-0, Mt-0 and Wil-2) that also show differences in their phenotypic plastic responses to increased temperature. And therefore, *PHYD* is an important candidate gene underlying a temperature responsiveness QTL. The Wassilewskija (Ws) accession has previously been shown to be deficient in *PHYD* due to the presence of an indel resulting in a premature stop codon (Aukerman *et al*., 1997). *phyD* mutants had bigger hypocotyls because they showed a reduced inhibition of elongation compared to WT seedlings, when grown under red or white light (Aukerman *et al*., 1997). The identification of the Ws *PHYD* allele provided evidence that *PHYD* performs a similar role to *PHYB*.

The discovery of this mutation in the Ws-0 accession was further exploited in another study to investigate the role of *PHYD* in the shade avoidance syndrome (Devlin *et al*., 1999) which suggested that *PHYD* is significantly involved in the shade avoidance response. Due to the nature of the mutation, its functional role, and isolation of an environmental QTL harbouring it, *PHYD* gene would seem a reasonable candidate for

underlying differences in temperature responsiveness. The *PHYD* environmental QTL may be a major effect QTL for natural variation in thermo-sensitivity. And this can be tested by functional complementation using a transgenic approach followed by quantifying the differential responses of the major effect size accessions that reveal polymorphisms.

The G x E analysis addresses the question of genetic basis of phenotypic plasticity and the MAGIC RILs have proved to be an effective tool for studying the genetic basis of hypocotyl length under multiple temperature environments. In general Arabidopsis RILS, which are homozygous immortal mapping populations, can be genotyped or sequenced once and can be replicated across different environments (Nordborg and Weigel, 2008). In particular, MAGIC lines are ideal for investigating natural variation in different environments.

One of the benefits of these RILs, similar to others, is that unlimited replicates of each line can be phenotyped across multiple environments for growth and development related traits. In this way, phenotypic data can be accumulated for hypocotyl length, facilitating the investigation of G x E interactions and localising QTL for the environmental responsiveness. This approach has been effectively utilised in this study by phenotyping the RILs across a range of temperatures (12°C, 17°C, 22°C and 27°C) and by identifying loci associated with the temperature responsiveness between any two temperature conditions/treatments.

Due to the complexity of quantitative traits and their interaction with the environment, G x E analyses are usually difficult to comprehend. Studying G x E interactions is an opportunity rather than just a problem (Simmonds, 1991). The G x E_T in this study has also been considered as such an opportunity. This is evident from the fact that a pre-QTL mapping phenotypic analysis to temperature responses in chapter 2, has demonstrated that G x E_T interactions exist between the Arabidopsis accessions (figure 2.8). It is of great interest to identify the molecular and genetic basis of the observed phenotypic G x E_T interactions.

For the temperature responsiveness QTL mapping, this has been accomplished by using the ratios of the phenotypic values and hence the Q x E_T effects have been quantified. The results of this mapping show that QTL reveal a temperature-dependent correlation with

hypocotyl length and hence these QTL are variable-effect QTL. Q x E_T interactions in the hypocotyl length trait in Arabidopsis has not been previously undertaken especially in response to temperature. Studying Q x E_T interactions in response to increased ambient temperatures could identify candidate genes involved in temperature responses of Arabidopsis growth.

This forms the basis of temperature responsiveness, which influences the phenotype of hypocotyl length as shown in chapter 2. The G x E_T interaction confirmed unique temperature effects when QTL were detected in a specific temperature environment. The QTL mapping for the temperature responsiveness trait in various temperature environments has allowed the dissection of responses to temperature of hypocotyl length.

In a post QTL mapping analysis, a QTL causal gene can be isolated on the basis of its chromosomal location, known as positional cloning. This classical method that identifies genes for which no prior knowledge of the protein product or its function is known has been previously used effectively (Tanksley *et al*., 1975). Positional cloning methods were developed for identifying genes with Mendelian effects. Variations of these methods have been used for the identification of QTL (Frary *et al*., 2000; Fridman *et al*., 2000; Johanson *et al*., 2000; El-Assal *et al*., 2001; Takahashi *et al*., 2001; Liu *et al*., 2002).

Another approach is to test candidate genes of known function associated with the phenotypic trait that map to QTL intervals (Doebley *et al*., 1995; Thornsberry *et al*., 2001; Loudet *et al*., 2007). Genes that are functionally related to a trait and that map at a QTL critical region can be selected as candidates (Wayne and McIntyre, 2002).

The candidate gene method, used in conjunction with QTL mapping, is a powerful tool to discover the genes involved in natural variation in a trait. This involves searching for strong candidate genes in a QTL interval that are functionally involved in controlling a trait. It is then possible to look for molecular polymorphisms in these candidate genes that may explain the variation in the trait. This method allows to shortlist and prioritises genes in a QTL. The candidate gene method can be used to proceed directly to functional testing if a strong candidate gene is present within a QTL (Remington *et al*., 2001). If the biological function of a candidate gene is known, this functional information could be

used to test for natural variation in the trait. This approach has been utilised successfully in plant genetics (Byrne and McMullen, 1996) and is facilitated with the availability of genome sequences because the polymorphisms in the sequences of the candidate genes and their alleles are known.

In this study, the candidate gene approach was applied for the characterisation and subsequent cloning of the QTL intervals. The working hypothesis is that a molecular polymorphism in the protein-coding region of the candidate gene is related to the observed phenotypic variation in the hypocotyl length trait. This approach hypothesises that genes of known function could correspond to QTL controlling natural variation in hypocotyl length.

Consequently, the analysis of sequence data, from the sequencing of the whole genomes, combined with QTL mapping parental reconstruction data led to an integrated and informed approach, which has played a fundamental role in the shortlisting of the number of genes identified in each of the QTL.

The ultimate goal of QTL mapping is to identify the causal alleles controlling a quantitative trait (Mackay, 2001) also known as QTL cloning. In this study, the QTL have been fine mapped to a few cM or less and there are a number of interesting candidate genes in the various QTL. Hence, the identification of the responsible genes in the QTL intervals by candidate gene approach should be a less daunting task where the QTL are a few cM or less. The QTL cloning includes approaches of mutational and transgenic analyses for functional testing of candidate genes which comprise of the subsequent two chapters respectively.

3.7 References:

Alonso-Blanco C, Koornneef M. (2000) Naturally Occurring Variation in Arabidopsis: An Underexploited Resource for Plant Genetics. Trends Plant Sci. 5, 1, 22-9.

Alonso-Blanco C, Mendez-Vigo B, Koornneef M. (2005) From Phenotypic to Molecular Polymorphism Involved in Naturally Occuring Variation of Plant Development. Int J Dev Biol, 49, 717–732.

Atwell S, Huang Y, Vilhjalmsson B, Willems G, Horton M, *et al*. (2010) Genome-Wide Association Study of 107 Phenotypes in *Arabidopsis thaliana* Inbred Lines. Nature, 465, 627–631.

Aukerman M, Hirschfeld H, Wester L, Weaver M, Clack T, *et al*. (1997) A Deletion in the *PHYD* Gene of the Arabidopsis Wassilewskija Ecotype Defines a Role for Phytochrome D in red/far-red Light Sensing. The Plant Cell, 9, 8, 1317-1326.

Balasubramanian S, Schwartz C, Singh A, Warthmann N, Kim M, *et al*. (2009) QTL Mapping in New *Arabidopsis thaliana* Advanced Intercross-Recombinant Inbred Lines. PLoS ONE, 4, e4318.

Balint-Kurti P, Yang J, Esbroeck G, Jung J, Smith M. (2010) Use of a Maize Advanced Intercross Line for Mapping of QTL for Northern Leaf Blight Resistance and Multiple Disease Resistance. Crop Sci, 50, 458–466.

Barton N, Keightley P. (2002) Understanding Quantitative Genetic Variation. Nat. Rev. Genet., 3, 11–21.

Bechtold N, Ellis J, Pelletier G. (1993) In Planta Agrobacterium Mediated Gene Transfer by Infiltration of Adult *Arabidopsis thaliana* Plants. Acad. Sci. Paris, 3, 1194–1199.

Bradshaw A. (2006) Unravelling Phenotypic Plasticity – Why Should We Bother? New Phytologist, 170, 4, 644–648.

Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, *et al*. (2010) Linkage and Association Mapping of *Arabidopsis thaliana* Flowering Time in Nature. PLoS Genet, 6.

Byrne P, McMullen M. (1996) Defining Genes for Agricultural Traits: QTL Analysis and the Candidate Gene Approach. Probe, 7, 24–27.

Cavanagh C, Morell M, Mackay I, Powell W. (2008) From Mutations to MAGIC: Resources for Gene Discovery, Validation and Delivery in Crop Plants. Curr Opin Plant Biol, 11, 215–221.

Christians J, Keightley P. (2002) Dissecting the Genetic Basis of Phenotypic Variation. Current Biology, 12, R415-R416.

Darvasi A, Soller M. (1995) Advanced Intercross Lines, an Experimental Population for Fine Genetic Mapping. Genetics, 141, 1199–1207.

Devlin P, Robson P, Patel S, Goosey L, Sharrock R, *et al*. (1999) Phytochrome D Acts in the Shade-Avoidance Syndrome in Arabidopsis by Controlling Elongation Growth and Flowering Time. Plant Physiology, 119, 3, 909-916.

Doebley J, Stec A, Gustus C. (1995) teosinte branched 1 and the Origin of Maize: Evidence for Epistasis and the Evolution of Dominance. Genetics, 141, 333–346.

Doerge R. (2002) Mapping and Analysis of Quantitative Trait Loci in Experimental Populations. Nat Rev Genet, 3, 43–52.

El-Assal S, Alonso-Blanco C, Peeters A, Raz V, Koornneef M. (2001) A QTL for Flowering Time in Arabidopsis Reveals a Novel Allele of CRY2. Nat Genet, 29, 435–440.

El-Lithy M, Bentsink L, Hanhart C, Ruys G, Rovito D, *et al*. (2006) New Arabidopsis Recombinant Inbred Line Populations Genotyped using SNPWave and their use for Mapping Flowering-time Quantitative Trait Loci. Genetics, 172, 1867–1876.

Falconer D, Mackay T. (1996) Introduction to Quantitative Genetics, Edn 4.

Flint J, Mott R. (2001) Finding the Molecular Basis of Quantitative Traits: Successes and Pitfalls. Nat. Rev. Genet., 2, 437–445.

Flint-Garcia S, Thornsberry J, Buckler E. (2003) Structure of Linkage Disequilibrium in Plants. Annu Rev Plant Biol, 54, 357–374.

Frary A, Nesbitt T, Frary A, Grandillo S, van der Knaap E, *et al*. (2000) fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. Science, 289, 85–88.

Fridman E, Pleban T, Zamir D. (2000) A Recombination Hotspot Delimits a Wild-species Quantitative Trait Locus for Tomato Sugar Content to 484 bp within an Invertase Gene. Proc Natl Acad Sci USA, 97, 4718–4723.

Ghandilyan A, Ilk N, Hanhart C, Mbengue M, Barboza L, *et al*. (2009) A Strong Effect of Growth Medium and Organ Type on the Identification of QTL for Phytate and Mineral Concentrations in three *Arabidopsis thaliana* RIL Populations. J Exp Bot, 60, 1409–1425.

Geldermann H. (1975) Investigations on Inheritance of Quantitative Characters in Animals by Gene Markers. I. Methods, Theor Appl Genet, 46, 319–330.

Gray W, Östin A, Sandberg G, Romano C, Estelle M. (1998) High Temperature Promotes Auxin-mediated Hypocotyl Elongation in Arabidopsis. Proc Natl Acad Sci USA, 95, 12, 7197–7202.

Hackett C. (2002) Statistical Methods for QTL Mapping in Cereals. Plant Mol Biol, 48, 585-599.

Holland J. (2007) Genetic Architecture of Complex Traits in Plants. Curr Opin Plant Biol, 10, 156–161.

Huang X, Paulo M, Boer M, Effgen S, Keizer P, *et al*. (2011) Analysis of Natural Allelic Variation in Arabidopsis using a Multiparent Recombinant Inbred Line Population. Proc Natl Acad Sci USA, 108, 4488–4493.

Johanson U, West J, Lister C, Michaels, S, Amasino R, *et al*. (2000) Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. Science, 290, 344–347.

Kover P, Schaal B. (2002) Genetic Variation for Disease Resistance and Tolerance among *A. thaliana* Accessions. Proc Natl Acad Sci. 99, 11270–11274.

Kover P, Mott R. (2012) Mapping the Genetic Basis of Ecologically and Evolutionarily Relevant Traits in *Arabidopsis thaliana.* Current Opinion in Plant Biology, 15, 2, 212–217.

Kover P, Valdar W, Trakalo J, Scarcelli N, Ehrenreich I, *et al*. (2009) A Multiparent Advanced Generation Intercross to Fine-map Quantitative Traits in *Arabidopsis thaliana*. PLoS Genet, 5.

Lister C, Dean C. (1993) Recombinant Inbred Lines for Mapping RFLP and Phenotypic Markers in *Arabidopsis thaliana*. Plant J, 4, 745–750.

Liu J, Van Eck J, Cong B, Tanksley S. (2002) A New Class of Regulatory Genes underlying the Cause of Pear-shaped Tomato Fruit. Proc Natl Acad Sci USA, 99, 13302–13306.

Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F. (2002) Bay-0 × Shahdara Recombinant Inbred Line Population: A Powerful Tool for the Genetic Dissection of Complex Traits in Arabidopsis. Theor Appl Genet, 104, 1173–1184.

Loudet O, Chaillou S, Merigout P, Talbotec J, Daniel-Vedele F. (2003) Quantitative Trait Loci Analysis of Nitrogen Use Efficiency in Arabidopsis. Plant Physiol, 131, 345–358.

Loudet O, Saliba-Colombani V, Camilleri C, Calenge F, Gaudon V, *et al*. (2007) Natural Variation for Sulfate Content in *Arabidopsis thaliana* is highly Controlled by APR2. Nat Genet. 39, 7, 896-900.

Mackay I, Powell W. (2007) Methods for Linkage Disequilibrium Mapping in Crops, Trends in Plant Science, 12, 2, 57–63.

Mackay T. (2001) The Genetic Architecture of Quantitative Traits. Annu. Rev. Genet., 35, 303–339.

Maloof J. (2003) Genomic Approaches to Analyzing Natural Variation in *Arabidopsis thaliana.* Curr Opin Genet Dev. 13, 6, 576-82.

Maloof J, Borevitz J, Weigel D, Chory J. (2000) Natural Variation in Phytochrome Signaling. Cell and Developmental Biology, 11, 523–530.

Nordborg M, Weigel D. (2008) Next-generation Genetics in Plants. Nature, 456, 720-3.

Prinzenberg A, Barbier H, Salt D, Stich B, Reymond M. (2010) Relationships between Growth, Growth Response to Nutrient Supply, and Ion Content using a Recombinant Inbred Line Population in Arabidopsis. Plant Physiol, 154, 1361–1371.

Rauh B, Basten C, Buckler E. (2002) Quantitative Trait Loci Analysis of Growth Response to Varying Nitrogen Sources in *Arabidopsis thaliana*. Theor Appl Genet, 104, 743–750.

Simmonds N. (1991) Selection for Local Adaptation in a Plant Breeding Programme. Theoretical and Applied Genetics, 82, 3, 363-7.

Slate J. (2005) Quantitative Trait Locus Mapping in Natural Populations: Progress, Caveats and Future Directions. Mol Ecol. 14, 2, 363-79.

Takahashi Y, Shomura A, Sasaki T, Yano M. (2001) Hd6, a Rice Quantitative Trait Locus involved in Photoperiod Sensitivity Encodes the α Subunit of Protein Kinase CK2. Proc Natl Acad Sci USA, 98, 7922–7927.

Tanksley S, Ganal M, Martin G. (1995) Chromosomal Landing: A Paradigm for Map-based Cloning in Plants with Large Genomes. Trends Genet. 11, 63–68.

Thornsberry J, Goodman M, Doebley J, Kresovich S, Nielsen D, *et al*. (2001) Dwarf8 Polymorphisms Associate with Variation in Flowering Time. Nat Genet, 28, 286–289.

Valdar W, Flint J, Mott R. (2006) Simulating the Collaborative Cross: Power of QTL Detection and Mapping Resolution in Large Sets of Recombinant Inbred Strains of Mice. Genetics, 172, 1783–1797.

Wayne M, McIntyre L. (2002) Combining Mapping and Arraying: An Approach to Candidate Gene Identification. Proc Natl Acad Sci USA, 99, 14903–14906.

Yalcin B, Flint J, Mott R. (2005) Using Progenitor Strain Information to Identify Quantitative Trait Nucleotides in Outbred Mice. Genetics, 171, 673–681.

Yu J, Pressoir G, Briggs W, Vroh B, *et al.* (2006) A Unified Mixed-model Method for Association Mapping that Accounts for Multiple Levels of Relatedness. Nat Genet, 38 203–208.

3.8 Appendix:

Tests of Between-Subjects Effects

Dependent Variable: Hypocotyl Length

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	86654.055[a]	1586	54.637	10028.098	.000	.999
Intercept	186085.482	1	186085.482	34154297.644	.000	.999
RIL	8710.832	433	20.117	3692.365	.000	.989
Temperature	65949.640	3	21983.213	4034818.847	.000	.999
RIL * Temperature	10684.938	1150	9.291	1705.324	.000	.991
Error	95.112	17457	.005			
Total	286179.130	19044				
Corrected Total	86749.168	19043				

a. R Squared = .999 (Adjusted R Squared = .999)

Table 1: Two way ANOVA showing significance and heritability in the MAGIC RILs

Procedure for the Manual Extraction of Gene Sequences:

I created the following bioinformatic protocol for the software BioEdit for the *in silico* extraction of genes from the completely sequenced MAGIC parental genomes. A local BLAST nucleotide library was constructed by using the genomic sequences of the parental accessions. This library was then used to perform a BLAST search of the genes of interest.

Protocol:

To begin with, the sequence of the gene of interest was obtained from the reference wild type annotated Col-0 accession by using the sequence viewer section on TAIR's website. This sequence contained the whole gene from the ATG start codon to the stop codon including both the exons and introns. The extracted Col-0 sequence was saved as a text document in a fasta file format and then opened in BioEdit. BLAST is included in BioEdit as an accessory application and this function was used to perform a local BLAST by using this file on the specific chromosome of the genomes where the gene of interest lies. The BLAST result provided the coordinates of the gene in each of the parental accessions.

These coordinates were used in BioEdit to extract the gene of interest by pinpointing the start and end base positions of the gene. The gene of interest was extracted in the same manner from all the parental accessions and saved as a text document. This file, containing the gene of interest from the nineteen parental accessions, was opened in BioEdit and the exons were extracted from the whole gene sequence. All the exons were concatenated together after splicing the introns and hence the whole coding sequence was obtained. This coding sequence was then translated in to the amino acid sequence using a function in BioEdit.

Chapter 4

Functional Analysis 1 – Knockout Analysis of Candidate Genes

4.1 Introduction:

In order to initiate a preliminary analysis of the causal relationship between candidate gene sequences and their potential function in hypocotyl length phenotype, gene knockouts have been used which possess a loss of function mutation. One of the primary goals of knockout analysis is the identification of the altered hypocotyl length phenotype that is caused by the insertional mutation in the candidate genes under study. Gene deficiencies in Arabidopsis knockouts may provide insights into the roles of QTL candidate genes involved in hypocotyl length trait. It is interesting to explore the phenotypic consequences of the missing QTL candidate genes. Furthermore, this analysis also provides a mean to further shortlist the QTL candidate genes for further investigation.

This knockout strategy is effective as a first step in functional analysis because in this study it allows focusing experimentation on a smaller number of already shortlisted candidate QTL genes showing interesting non-synonymous sequence polymorphisms. In this knockout study, since the null mutations provide a direct means for the determination of the function of candidate gene sequences *in situ*, hence they are of due importance in post QTL analyses. Insertional mutagenesis allows the disruption of gene function by inserting a foreign sequence into the QTL candidate genes under study. This approach allows the direct monitoring of the effect of the deficiency of the candidate gene sequence in plants.

Agrobacterium tumefaciens (Agrobacterium) is a widespread soil bacterium which occurs naturally, induces crown gall, and has the capability to introduce new genetic material into the plant cell (Gelvin, 2003). Agrobacterium transfer-DNA (T-DNA) has been used for the generation of large collections of Arabidopsis insertion mutants (Alonso *et al*., 2003). The T-DNA insertional mutation is one of the tools of functional genomics and is a simple and direct gene targeting approach for exploring candidate gene function because a

known genomic QTL candidate sequence has been interrupted with a T-DNA. The main objective of this chapter is to use T-DNA insertional mutations in QTL candidate genes to understand the biological function that they may have in controlling the hypocotyl length phenotype.

In the forward genetic approach loss of function analysis of candidate genes is the fifth step as outlined in figure 1.4. For the successful identification of a particular gene that underlies a QTL identified in the previous chapter, direct tests have been carried out on promising candidates for their functional involvement in hypocotyl length variation. This has been feasible due to several reasons; the T-DNA knockouts are easily available from stock centres, homozygous individuals for the T-DNA insertion can be genotyped by PCR, and the phenotype of hypocotyl length can be easily scored.

It is imperative that gene function be explored to determine if any of the QTL candidate genes identified in chapter 3 are functionally involved in the hypocotyl length phenotype. Such biological function of the candidate genes can be initially ascertained in the hypocotyl length trait through knockout mutants obtained by gene disruption. The most direct method of understanding gene function in general is its complete inactivation. To facilitate the study of natural variation, a knockout approach has been applied in a post QTL mapping analysis with the objective to experimentally test if a particular gene shows any function in controlling hypocotyl length in response to temperature.

In chapter 3, although bioinformatic tools were used effectively for the detection of sequence polymorphisms by *in silico* analysis in QTL candidate genes, this analysis is only indicative and insufficient for the definition of gene function. Genomic sequence data is not quite enough to attribute hypocotyl length function to any particular candidate gene. Hence, one way to overcome this problem is to perform a phenotypic analysis of mutated candidate genes by analysing plants that are homozygous for the mutation. The *in silico* analysis aids in the shortlisting of QTL candidate genes and experimental evidences are necessary for functional characterisation purposes.

The results of this preliminary analysis provide a roadmap for the functional shortlisting and subsequent direct testing of candidate sequences by using brute force genetics including transgenic complementation methods described in the next chapter 5.

4.2 Methods and Materials

This section describes the PCR-based genotyping and molecular procedures used for the identification of homozygous knockout lines of candidate genes. It also describes the phenotyping of the knockout lines for hypocotyl length.

4.2.1 Knockout Mutants:

For the selection of knockout lines and the location of the mutation, computational searches were performed using the SALK Institute Genomic Analysis Laboratory (SIGnAL) database called 'T-DNA Express: Arabidopsis Gene Mapping Tool' (Alonso *et al*., 2003) http://signal.salk.edu/cgi-bin/tdnaexpress. The interactive map shows the gene position, its location and T-DNA insertions along with its orientation. This allowed *in silico* searches in the database for knockout lines which had T-DNA inserted in the candidate genes. The choice of mutation, for the purpose of QTL candidate gene functional study, was prioritised to maximise the possibility that the gene was knocked out; and hence the preferred location of the T-DNA insertion sites were exons of candidate genes. The SALK lines are in the Col-0 background and the size of the T-DNA is 4501 bp, which confers resistance to Kanamycin.

The seeds of the knockout lines were obtained from the European Arabidopsis Stock Centre (NASC) (www.arabidopsis.info). These knockout seeds are usually heterozygous and are segregating for the T-DNA (25% homozygous WT: 50% heterozygous: 25% homozygous mutant) thus allowing for homozygous lines to be identified. To determine the hypocotyl length phenotype of the knockouts, homozygous plants were required where no segregation of the T-DNA insertion is taking place. Therefore, the seeds were bulked up for the detection of homozygous individuals. For this purpose, the seeds were germinated as described in Chapter 2. One week old seedlings were transplanted onto the soil pots and were covered with lids for a few days to retain the moisture. The plants were grown in controlled environmental rooms (CERs) under long days at 22°C. For DNA extraction, fresh leaf tissue was taken from 10-12 individual plants for each knockout line progeny and put in separate tubes and in dry ice. The individual plants from which leaf tissues were taken were allowed to self and set seed. The genotyping by PCR and gel electrophoresis determined which of these individual plants for each of the knockout

lines were homozygous and these lines were used for the phenotyping of hypocotyl length.

4.2.2 DNA Extraction:

Plant DNA was extracted using CTAB buffer extraction method. The protocol is as follows:

1. Put floral meristem or leaf into tube (freeze until enough samples)
2. Put metal ball into each well and put on cap
3. Freeze at -70°C
4. Shake for 30 seconds at speed 800 on Geogrinder
5. Add 300µl extraction buffer and mix by inversion
6. Heat at 60°C for 15 minutes in the oven
7. Add 300µl chloroform and mix by inversion
8. Spin for 20 minutes at speed 5000
9. Transfer supernatant (clear upper layer) to a new plate
10. Add 300µl Isopropanol and mix by inversion
11. Leave at room temperature for 10 minutes
12. Spin DNA for 10 minutes at speed 5000 and remove liquid
13. Wash pellet with 500µl of 70% ethanol and spin for 5 minutes
14. Air dry pellet in laminar flow hood
15. Resuspend DNA in 100µl water and leave in fridge overnight or store at -20°C

4.2.3 Genotyping of Knockout Lines:

For the identification of homozygous individuals in each of the candidate genes under study, the segregating T-DNA lines were genotyped. The genotype of the DNA samples obtained from the multiple individuals for each knockout line was determined by PCR and gel electrophoresis. Two separate PCR reactions were set up for each candidate gene. One reaction was performed using a gene specific primer and a T-DNA specific primer to test for the presence of a T-DNA insertion in the gene of interest. The second reaction was performed using a pair of gene specific primers flanking the insertion site to test for the presence of a WT undisrupted allele.

The genomic primers were designed on the SIGnAL website: (http://signal.salk.edu/tdnaprimers.2.html) using 'T-DNA Primer Design'. The default settings for the primer design on the website were changed in order to reduce the total length of the PCR product (<1000 bp). The ranking of the genomic primers was checked in Net Primer (http://www.premierbiosoft.com). BLAST (http://blast.ncbi.nlm.nih.gov) was used to check if the designed primers would virtually anneal to the specific genomic location showing primer specificity. All the primers annealed virtually to the targeted genomic sequence. For the genotyping of the individual plants the reagents of the PCR and their quantities were as follows for a total of 20 µl PCR mix:

Primer 1 = 1 µl, Primer 2 = 1 µl, dNTPs = 1 µl, Buffer = 2 µl, Taq = 0.1 µl, DNA = 1 µl, Water = 13.9 µl

The PCR programme for genotyping was set up as: denaturation at 95°C for 3 min, then 35 cycles of denaturation at 95°C for 15 sec; annealing at 55–60°C for 30 sec (as per recommendation of primer manufacturer SIGMA), and elongation at 72°C for 1 min.

The PCR product was loaded on 1% Agarose Gel. Ethidium bromide added to the gel was 6 µl. 6 µl of 2 Log ladder was used in the gel as a size standard. DNA loading dye used was 2 µl. After gel electrophoresis, the gel was placed under UV light to detect the DNA bands and images were taken. The genotyping data was analysed for all the individuals for each knockout line and homozygous individuals for the T-DNA insertion were selected for phenotyping.

4.2.4 Determination of Homozygous Knockout Lines:

The genotyping by PCR not only identifies the presence of a T-DNA in the mutant plants but it is also a means to identify the homozygous lines. The phenotyping of heterozygous lines for the trait does not provide an accurate phenotyping result due to the segregation of the T-DNA in the background. Therefore, it is imperative to identify homozygous lines.

A PCR product as a result of a gene specific primer pair meant that the plants were either WT or heterozygous. Additionally, a PCR product as a result of a gene specific primer and a T-DNA border primer meant that the plants were either heterozygous or homozygous for the insertion. Both results combined provided complete information on the genotypes

of the plants. To detect an undisrupted wild type allele of a specific candidate gene, in the first PCR, a gene specific primer pair flanking the region of insertion was used. Therefore, individuals that showed only a band for the gene specific primer pair were wild type plants. In the second PCR reaction, to test for the presence of a T-DNA insertion, T-DNA specific and gene specific primers were used. Individuals that showed one band for the gene specific primer and second band for the T-DNA left border primer and gene specific primer were heterozygous plants. Moreover, individuals that only showed a band for the T-DNA left border primer and a gene specific primer were homozygous plants. The evidence for individual plants to be homozygous for the T-DNA insertion was based on the presence of a T-DNA specific product and the absence of a gene specific product. The T-DNA border primer was very effective in identifying insertions as it consistently amplified a known T-DNA insertion in mutant lines.

The knockout lines and their list of PCR primers used are provided in the appendix at the end of this chapter.

4.2.5 Phenotyping Growth Conditions:

The seeds were stratified after plating for 4 days at 4°C. They were then germinated in growth incubators at 22°C and were allowed to grow for 1 week in the respective treatment temperatures. The Petri dishes were placed vertically on the shelves of the incubators. This allowed the seedlings to grow horizontally along the growth media.

4.2.6 Image Analysis:

The Petri dishes were placed horizontally on a black background surface and were photographed from above with a digital SLR Nikon camera. A scale was included in each image for calibration of pixels with length in millimetres, and all the images were taken at the same magnification. The number of pixels was counted between the two lines of a 1 millimetre for scaling. Image J software was used to quantitatively measure the hypocotyl lengths as described in Chapter 2.

4.2.7 Statistical Analysis:

SPSS version 19; Inc.; IBM software was used to perform the Student's t-test to determine if the knockout lines were significantly different from the WT.

4.3 Results and Discussion:

4.3.1 Phenotypic Analysis of Knockout Mutants for Candidates at 22°C and 27°C:

The mutational analysis allows us to predict gene functions of the candidate genes by the loss of function mutations in them. Therefore, the phenotypic analysis of these knockout candidate genes reveals their potential functional involvement in the hypocotyl length trait and their contribution to the variation of the trait. From the post QTL mapping analysis in Chapter 3, a total of 25 candidate genes were shortlisted in various QTL at 22°C and 27°C.

In the gene knockout analysis, the effect of the mutant gene sequences for all the candidate genes on the hypocotyl length phenotype has been studied. The main assumption for this analysis is that if the phenotype of the mutants is different from the WT, then it can be implied that this particular candidate gene may be playing a role in the natural variation of the trait. The mutant plants were screened by phenotyping to look for unusual phenotypes when compared to Col-0 WT. For the mutant plants which showed a different phenotype, it was assumed that the T-DNA insertion has caused the candidate gene, which is related to hypocotyl length, to be inactivated.

Since hypocotyl length is a trait that can be easily scored after one week of growth, a phenotyping experiment was performed on the knockout seeds obtained from the stock centre. The results revealed that there was great variation in the hypocotyl length phenotype between the individuals of any particular knockout line. This implied that some of the lines were not homozygous as the lines were still segregating which could be seen in the phenotype. This hypothesis was later found to be correct when the lines were genotyped by PCR methods to determine the genotype of the knockout lines.

After genotyping there were no heterozygous or homozygous individuals identified for the candidate genes *PCB*, *RPT2*, *CYP71*, *BAM2*, and *SAUR18*, which indicates a problem with knockout lines. And hence their knockout phenotypic analysis could not be performed as all the plants genotyped were WT. However, for *CYP71* a homozygous knockout line *cyp71-1* was available in the Wigge Lab. The knockout phenotyping of the remaining candidate genes was performed and hypocotyl length phenotype data is represented in figures 4.1 and 4.2.

The seeds obtained from the stock centre were bulked up and after determining the genotypes of the individual plants, homozygous individuals for all the knockout lines were identified and phenotyped at the respective QTL temperatures.

4.3.2 From Gene to Phenotype; Analysing Mutants to Identify Gene Function:

The characterisation of the knockout phenotypes were performed under the QTL environmental conditions. This analysis has shown which of the candidate genes are functionally involved in the hypocotyl length phenotype. The first step towards the description of gene function was the detection of knockout mutants. As a next step, the roles of the candidate genes have been examined by studying these mutant lines in which the genes have been changed. These mutant lines have been phenotyped to investigate the consequences of the T-DNA mutation on hypocotyl length relative to the WT. Homozygous individuals were phenotyped to ensure that both copies of the candidate genes were disrupted. The KO lines used for the candidate genes are shown in table 4.1.

Candidate Gene	TAIR Name	KO Line Name	NASC ID	Insertion Location
ARF10	AT2G28350	SALK_087247C	N655696	Exon
BAM2	AT3G49670	SAIL_1053_E09	N878024	Exon
BIN2	AT4G18710	GK-244F08.01	N756025	Exon
CESA8	AT4G18780	SALK_046685C	N679626	Exon
CLF	AT2G23380	SALK_088542C	N672840	Exon
COL3	AT2G24790	SAIL_361_A08	N873490	Exon
CYCT1	AT4G19600	SALK_139324 (BY)	N639324	Exon
EPR1	AT2G27380	GK-710H02.01	N332643	Exon
GRF4	AT3G52910	SALK_037642C	N657589	Exon
LSH9	AT4G18610	SALK_039868C	N660824	Exon
NPY2	AT2G14820	SALK_058416C	N659409	Exon
NPY4	AT2G23050	SALK_151725	N651725	Exon
PCB	AT3G49120	SAIL_143_G09	N873282	Exon
PHOT1	AT3G45780	SALK_088841C	N666670	Exon
PHYB	AT2G18790	SALK_069700C	N675665	Exon
PHYE	AT4G18130	SALK_092529C	N671700	Exon
PLL4	AT2G28890	SALK_047827	N547827	Exon
RPT2	AT2G30520	SAIL_140_D03	N871473	Exon
SAUR38	AT2G24400	SALK_001154C	N669301	Exon
SAUR8	AT2G16580	SALK_058324C	N679733	Exon
SAUR18	AT3G51200	GK-695B04.01	N316614	Exon
TSF	AT4G20370	SALK_087522C	N663213	Exon
TT6	AT3G51240	SALK_113321C	N653439	Exon

Table 4.1: Knockout lines for candidate genes showing location of T-DNA insertion.

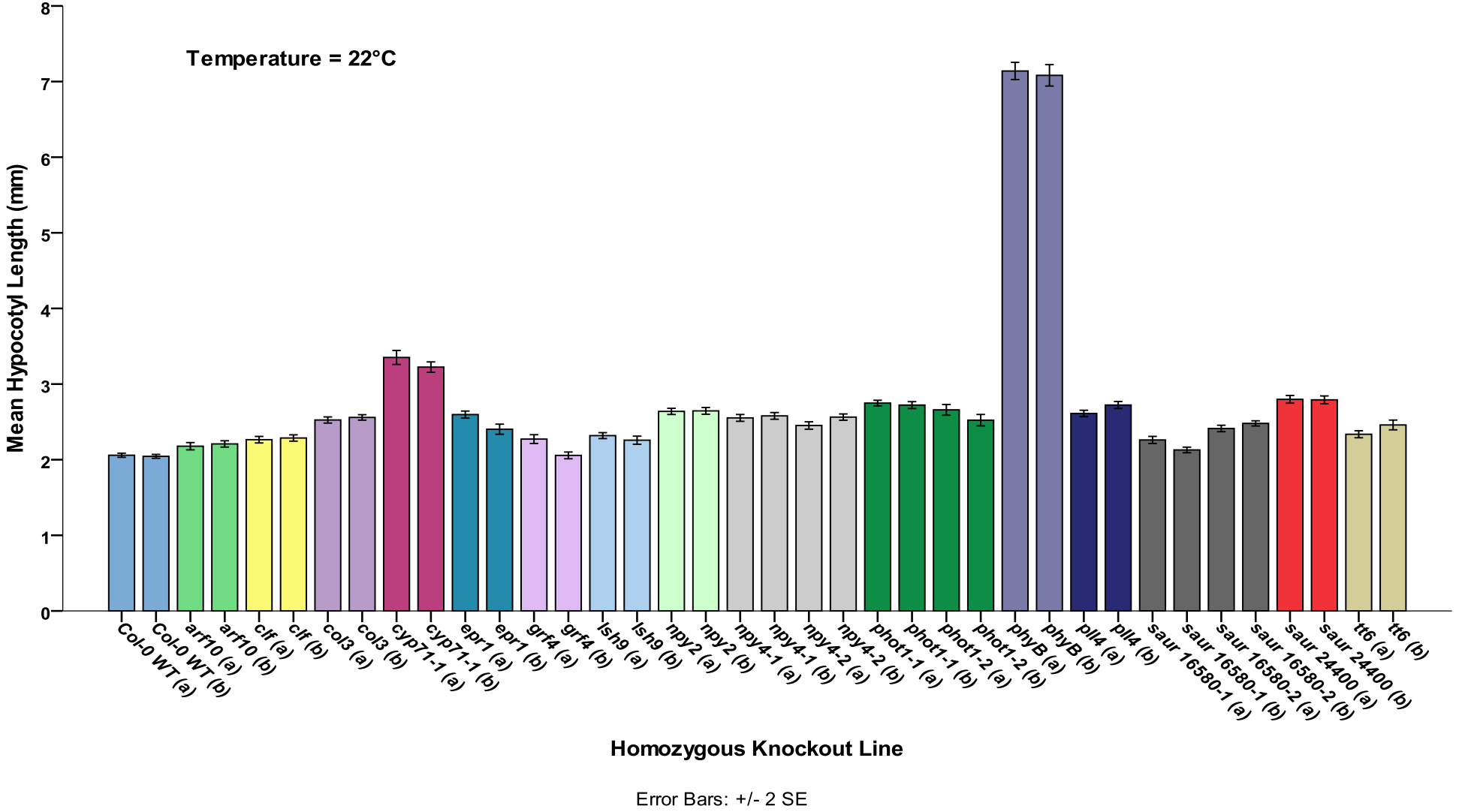

Figure 4.1: 1 week old hypocotyl length of candidate gene knockouts compared against Col-0 WT at 22°C, 170 µE/m^2/sec short days; (a) and (b) are technical replicates.

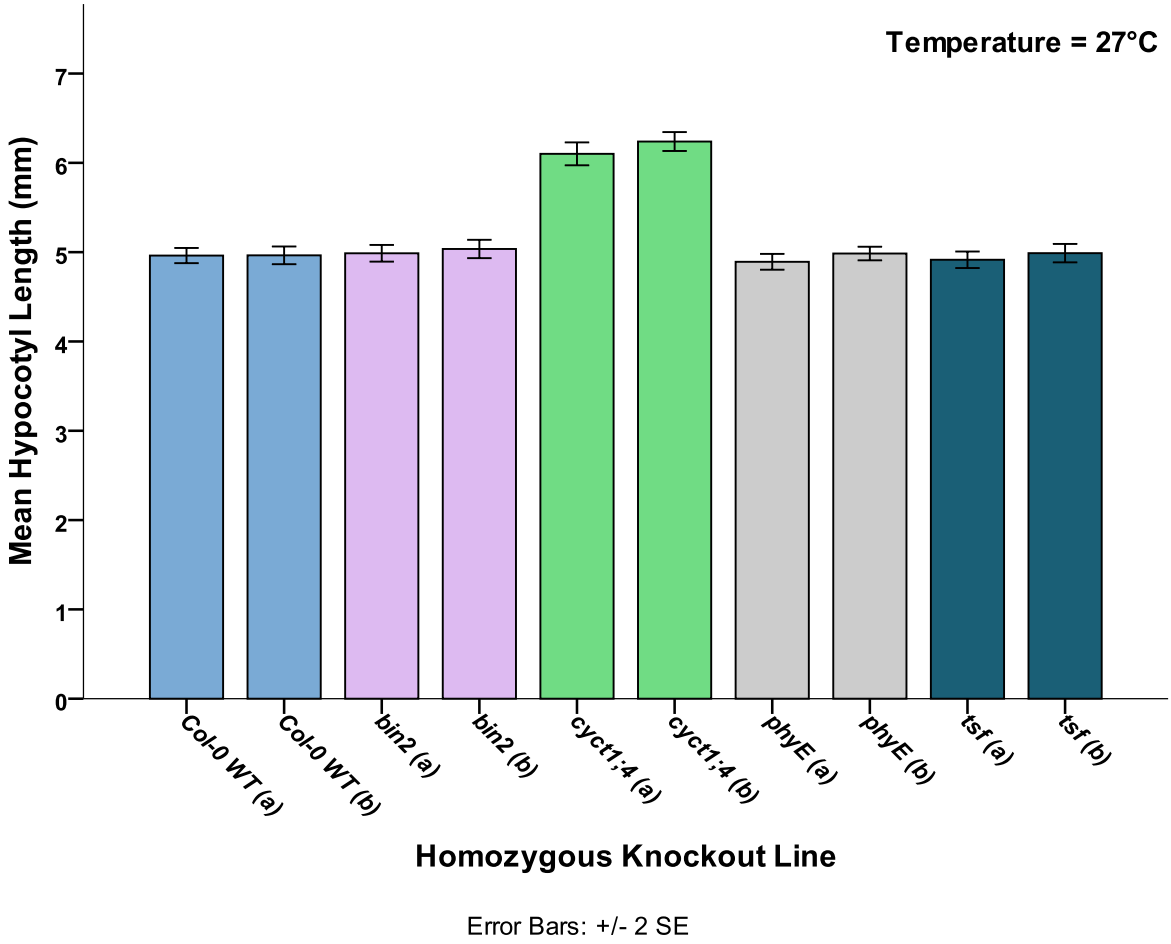

Figure 4.2: 1 week old hypocotyl length of candidate gene knockouts compared against Col-0 WT at 27°C, 170 µE/m^2/sec short days; (a) and (b) are technical replicates.

Figures 4.1 and 4.2 illustrate the hypocotyl length phenotypes of homozygous knockout mutants which have been phenotyped under controlled temperature conditions similar to those in which the QTL were identified. We are interested in identifying mutant phenotypes which are different from the WT under these specific temperature conditions.

An independent-samples t-test was performed to test for the differences in hypocotyl lengths of the knockouts as compared to WT at the respective QTL temperatures. The most significant and distinctive differences in hypocotyl lengths were demonstrated by the mutant lines which are presented in the following table:

Temperature	KO mutant	T-test result	Approximate percent increase in hypocotyl length
22°C	*phyB*	$t(166) = -124.5$, $p < 0.001$	250%
22°C	*cyp71*	$t(158) = -36.54$, $p < 0.001$	75%
22°C	*saur38*	$t(183) = -28.04$, $p < 0.001$	50%
22°C	*npy2*	$t(197) = -24.58$, $p < 0.001$	40%
27°C	*cyct1;4*	$t(59) = -9.91$, $p < 0.001$	20%

A loss of function mutation in an accession could result in the hypocotyl length either being bigger or smaller than WT Col. This would depend on the function of the gene. If the gene acts as a positive regulator of hypocotyl length, and when altered with T-DNA, we would expect shorter hypocotyls. Conversely, if it acts as a negative regulator of hypocotyl length, and when altered with T-DNA, we would expect longer hypocotyls. For example, in the case of phyB, which is a negative regulator of hypocotyl length, the *phyB* mutants show bigger hypocotyl lengths. Several of the knockout lines show increases but this is not surprising as QTL had small hypocotyls for Col-0.

The phenotypic analysis demonstrates the roles played by the mutated genes in controlling the altered phenotype of hypocotyl length. Under the experimental test conditions, knockouts that have a similar phenotype to WT imply that the gene that has been knocked out is not functionally involved in hypocotyl length. Although several knockouts have been phenotyped, few of them have shown significant informative and visible altered phenotypes relative to WT and hence provide a direct indication to gene function.

The strategy for selecting a few of the shortlisted candidate genes was to observe the phenotype of the homozygous knockout mutants. Those knockout mutants whose hypocotyl phenotypes were the most significantly different (p=0.05) from the WT Col-0 were selected for further investigation. These genes may functionally be playing a role in the hypocotyl elongation, as observed from their phenotypes (figures 4.1 and 4.2). The knockout mutants which did not show a phenotype or showed a lesser significant phenotype were excluded from further investigation.

The knockout phenotyping results were compared with the phenotypes of the MAGIC parents to look for similarities. Quite interestingly, it was found that the *phyB* knockout phenotype, which showed the most pronounced difference in hypocotyl length at 22°C compared to the WT Col-0 (figure 4.1), had a similar phenotype to the Ct-1 accession at 22°C (figure 2.3). These results were obtained under dark growth conditions.

From the skotomorphogenic analysis of the MAGIC parents (figure 2.5), the etiolated phenotype of Ct-1 at 22°C had already been determined. The next question was whether the *phyB* knockout also had a similar phenotype to Ct-1 in the dark. To determine this, the *phyB* knockout was phenotyped in the dark and the results (figure 4.3) revealed that they both had a similar phenotype. This provided detailed evidence for a further functional study of these phenotypes in the transgenic analysis, as described in chapter 5.

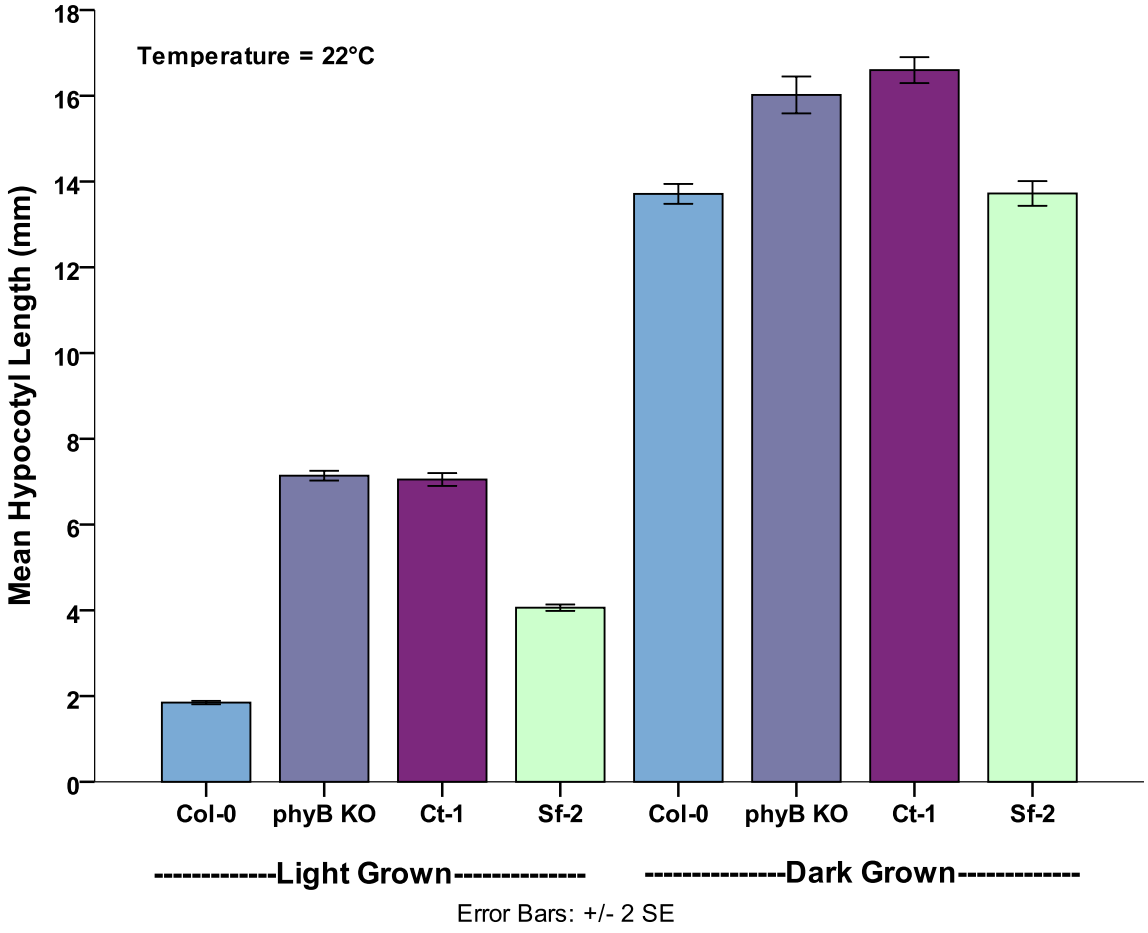

Figure 4.3: Comparison of 1 week old *phyB* knockout phenotype with Ct-1 and Sf-2 accessions in light and dark.

The study of T-DNA mutants with noticeable hypocotyl length phenotypes is undoubtedly a resource that has provided insights into the roles of the candidate genes. There are a few candidate genes at different QTL and temperatures that have been implicated in the hypocotyl length phenotype as a result of the knockout analysis. From the phenotypic analysis of the mutants, those which were identified as demonstrating the most significant differences were selected for the next step: QTL cloning by transgenic complementation of candidates.

4.3.3 Discussion of Candidate Genes used in Knockout Phenotypic Analysis:

To help interpret the results of this knockout analysis, the known mutant phenotypes of the candidate genes in various stages of development including hypocotyl length have been summarised. This is not a comparative study with the phenotypic analysis of the knockout mutants of this study, but it helps to develop a basic understanding of the roles that these candidate genes play in growth and development in general.

AUXIN RESPONSE FACTOR 10 (*ARF10*)

Auxin response factors (ARF) are transcription factors which function in auxin signal transduction in several plant growth phases (Guilfoyle and Hagen, 2007) by controlling the expression of auxin response genes (Tiwari *et al*., 2003). *ARF10* has been identified as one of the key regulators of root cap formation and is involved in cell differentiation. The *arf10* mutants show impaired seedling establishment and their leaves and flowers are defective (Wang *et al*., 2005). It has also been reported that repression of *ARF10* is vital for both seed germination and post-germination phases (Liu *et al*., 2007).

The T-DNA insertion in the *arf10* mutant (SALK T-DNA line: SALK_087247C) is in the first exon. The *arf10* mutant was phenotyped at 22°C.

BRASSINOSTEROID INSENSITIVE 2 (*BIN2*)

Brassinosteroids are ubiquitous plant steroid hormones (Schumacher and Chory, 2000) that regulate various aspects of normal plant growth and development including cell elongation (Clouse and Sasse, 1998). *BIN2* was identified by map based cloning and

reverse genetics as being involved in the cross-talk between auxin and brassinosteroid signalling pathways and *bin2* mutants are insensitive to brassinosteroids (Li *et al*., 2001). Skotomorphogenic hypocotyl elongation was studied in the *bin2* mutants and the seedlings were germinated and grown in darkness. The seedlings showed de-etiolated phenotypes that include short hypocotyls with open cotyledons, a characteristic of plants grown in light (Li *et al*., 1996). The *bin2* mutant showed reduced sensitivity to growth in the dark as compared to WT Col-0. Light signals inhibit brassinosteroids synthesis in the seedlings and hence they contribute to overall photomorphogenesis (Li *et al*., 2001).

The T-DNA insertion in the *bin2* mutant (GABI-Kat T-DNA line: GK-244F08.01) is in the first exon. The *bin2* mutant was phenotyped at 27°C.

CELLULOSE SYNTHASE 8 (*CESA8*)

In Arabidopsis there are 10 CESA proteins (Richmond and Somerville, 2000). The *CESA8* gene belongs to the cellulose synthase family protein and is involved in secondary cell wall biosynthesis (Turner and Somerville, 1997; Taylor *et al*., 2003). The *cesa8* mutants were characterised by abnormal xylem formation and had 70% lesser cellulose content compared to WT (Turner and Somerville, 1997).

The T-DNA insertion in the *cesa8* mutant (SALK T-DNA line: SALK_046685C) is in the eighth exon. The knockout phenotype of *CESA8* could not be measured because the adult homozygous knockout plants were very small and they did not yield any seeds. This is an example of a knockout mutation in these plants leading to a lethal mutation with no seed production.

CURLY LEAF (*CLF*)

The *CLF* gene appears to play a key role in shoot growth. It was initially identified through forward genetics as a suppressor of floral genes (Goodrich *et al*., 1997). The *clf* mutants have altered leaf morphology as the *CLF* gene regulates cell elongation (Kim *et al*., 1998).

The T-DNA insertion in the *clf* mutant (SALK T-DNA line: SALK_088542C) is in the twelfth exon. The *clf* mutant was phenotyped at 22°C.

CONSTANS-LIKE 3 (*COL3*)

COL3 belongs to a family of five CONSTANS (*CO*)-like proteins and it is most closely related to the gene CONSTANS (*CO*) (Robson *et al*., 2001) which plays a role in shoot growth and development. *COL3* has been identified through reverse genetics as an interacting protein with *CONSTITUTIVE PHOTOMORPHOGENIC1* (*COP1*), a gene which represses photomorphogenesis in etiolated growth conditions and hence *COL3* is important in regulating plant development in light (Datta *et al*., 2006). The mutant analysis of *col3* indicated that *COL3* positively regulates development in light by inhibiting hypocotyl elongation under short day conditions. The seedlings studied in white light had a similar hypocotyl phenotype to WT, however, *col3* was found to be less sensitive to red light; as a result had long hypocotyls in short days. These results suggested that *COL3* acted as a positive regulator of phytochrome mediated suppression of hypocotyl elongation (Datta *et al*., 2006).

The T-DNA insertion in the *col3* mutant (SAIL T-DNA line: SAIL_361_A08) is in the first exon. The *col3* mutant was phenotyped at 22°C.

CYCLIN T1;4 (*CYCT1;4*)

In Arabidopsis *CYCT1;4* belongs to a family of five genes that encode cyclin T-like proteins (Wang *et al*., 2004). Although *cyct1;4* loss of function mutants have been previously studied for its roles in infection by Cauliflower Mosaic Virus (CaMV), their roles in vegetative and reproductive growth phenotypes were also examined and it was found that *cyct1;4* mutants had altered leaf and flower growth compared to WT (Cui *et al*., 2007).

The T-DNA insertion in the *cyct1;4* mutant (SALK T-DNA line: SALK_139324 (BY)) is in the sixth exon. The *cyct1;4* mutant was phenotyped at 27°C.

CYCLOPHILIN 71 (*CYP71*)

CYP71 in Arabidopsis encodes a nuclear protein which is necessary for organogenesis (Li *et al*., 2007). The *cyp71* loss of function mutants have been considered as developmental mutants of Arabidopsis because of dramatic defects in multiple shoot and root organs;

and mutants had significantly reduced number of flowers compared to WT (Li *et al*., 2007).

The T-DNA insertion in the *cyp71-1* mutant (SALK accession number JP69.6C06) is in the first intron/exon. The *cyp71-1* mutant was phenotyped at 22°C.

EXTENSIN PROLINE-RICH 1 (*EPR1*)

It has been shown that extensins are expressed in elongating cells and are involved in modification of the cell wall structure (Bernhardt & Tierney, 2000). From a mutant analysis, another study found that *EPR1* encodes a proline-rich extensin-like protein which is involved in positive regulation of seed germination and plays a more specific role in the modification of cell wall structure during germination (Dubreucq *et al*., 2000).

The T-DNA insertion in the *epr1* mutant (GABI-Kat T-DNA line: GK-710H02.01) is in the first exon. The *epr1* mutant was phenotyped at 22°C.

GROWTH-REGULATING FACTOR 4 (*GRF4*)

In Arabidopsis, *GRF4* is one of the nine members of a GRF gene family. *GRF4* encodes a putative transcription factor and mutational analysis revealed that it is involved in growth and development of shoot and leaves (Kim *et al*., 2003). *GRF4* in Arabidopsis is a homologue to *GROWTH-REGULATING FACTOR1* of rice (*Oryza sativa*) (Os*GRF1*) which is involved in the regulation of stem growth (Van der Knaap *et al*., 2000).

The T-DNA insertion in the *grf4* mutant (SALK T-DNA line: SALK_037642C) is in the first exon. The *grf4* mutant was phenotyped at 22°C.

LIGHT SENSITIVE HYPOCOTYLS 9 (*LSH9*)

LSH9 is a hypothetical protein but gets its name by sequence homology with other LSHs which do have a known function. The T-DNA insertion in the *lsh9* mutant (Salk T-DNA line: SALK_039868C) is in the first exon. The *lsh9* mutant was phenotyped at 22°C.

NAKED PINS IN YUC MUTANTS 2 (*NPY2*) and NAKED PINS IN YUC MUTANTS 4 (*NPY4*)

NPY2 and *NPY4* are members of the NPY gene family and have been shown to be involved in auxin-mediated organogenesis (Cheng *et al*., 2008).

The T-DNA insertion in the *npy2* mutant (SALK T-DNA line: SALK_058416C) is in the second exon. The *npy2* mutant was phenotyped at 22°C.

The T-DNA insertion in the *npy4* mutant (SALK T-DNA line: SALK_151725) is in the fourth exon. The *npy4* mutant was phenotyped at 22°C.

PHOTOTROPIN 1 (phot1)

Phot1 is one of the two phototropins which function as blue light receptors in Arabidopsis. They absorb blue light and play an important role in the phototropism of the hypocotyls which allows them to bend towards blue light and are the main photoreceptors mediating this light response (Stowe-Evans *et al*., 2001). Additionally, phototropins also regulate leaf expansion, chloroplast movement and stomatal opening (Sakai *et al*., 2001; Sakamoto and Briggs, 2002). It has been demonstrated that *PHOT1* is the first light receptor which rapidly inhibits hypocotyl elongation under blue light growth conditions (Folta and Spalding, 2001). All such responses support plant growth and development. It has been revealed that under blue light the *phot1* loss of function mutant lacked the rapid inhibition of hypocotyl elongation and hence was blind and continued to elongate straight (Folta and Spalding, 2001).

The T-DNA insertion in the *phot1* mutant (SALK T-DNA line: SALK_088841C) is in the fourth exon. The *phot1* mutant was phenotyped at 22°C.

PHYTOCHROME B (phyB)

In Arabidopsis, phyB belongs to a family of five phytochromes (Sharrock and Quail, 1989) which regulate plant growth and development mainly under red light. PhyB is the main light receptor which predominantly regulates de-etiolation seedling responses in red light; however, other phytochromes are also involved in such responses (Franklin *et al*., 2003).

PhyB is well studied and the dramatic developmental phenotypes, as a result of insensitivity to red light, were first described in *phyB* mutant seedlings that had long hypocotyls and unexpanded cotyledons (Koornneef *et al*., 1980). PhyB functionally controls the inhibition of hypocotyl elongation under red light (Reed *et al*., 1993). It is a negative regulator of hypocotyl elongation and the mutants are less sensitive to red/far

red light as a result of which the hypocotyl elongates. This function of the gene is also well understood from the overexpression of *PHYB* gene. A transgenic *PHYB* overexpression line showed an exaggerated short hypocotyl phenotype that is severely light dependant (Wagner *et al*., 1991). Furthermore, the *phyB* mutant adult plants displayed elongated petioles and flowered early (e.g., Nagatani *et al*., 1991; Halliday *et al*., 1994). Additionally, the *phyB* null mutants which display an array of these noticeable phenotypes are characteristic of the shade-avoidance syndrome seen in WT seedlings exposed to low R/FR ratio light and include elongated hypocotyls and petioles, retarded leaf growth and early flowering (Somers *et al*., 1991; Smith and Whitelam, 1997). The *phyB* mutants that are phenotyped above 21°C display the early flowering phenotype of WT plants that avoid shade and hence phyB is considered as a suppressor of flowering (Simpson *et al*., 1999). This phenotype is independent of the short day or long day photoperiods (Blazquez and Weigel, 1999) but reveals sensitivity to ambient temperature (Blazquez *et al*., 2003; Halliday *et al*., 2003). Hence phyB has been implicated as a major contributor to the shade avoidance syndrome (Whitelam and Devlin, 1997).

It has been shown that the early flowering phenotype in the *phyB* mutant is temperature dependent; which was observed at 22°C but was abolished at 16°C (Halliday *et al*., 2003). It is interesting to note that in this study (chapter 3) *PHYB* has been identified as a strong candidate gene for controlling natural variation in hypocotyl length in a temperature-dependent QTL identified at 22°C. Moreover, in the knockout analysis, *phyB* mutants show the most significantly different hypocotyl length at 22°C.

It has been previously revealed that the *phyB* mutant shows a striking long hypocotyl phenotype when grown in white light (Somers *et al*., 1991; Reed *et al*., 1994). The results of the knockout analysis complement this finding and the *phyB* mutants showed reduced sensitivity to light at 22°C when grown under white light. The T-DNA insertion in the *phyB* mutant (SALK T-DNA line: SALK_069700C) is in the second exon.

PHYTOCHROME E (phyE)

PhyE is another member of the phytochrome family. It has been previously shown that the *phyE* mutants had a similar phenotype to WT seedlings (Devlin *et al*., 1998). PhyE has

been implicated in the light-induction of seed germination in Arabidopsis (Hennig *et al*., 2002).

The T-DNA insertion in the *phyE* mutant (SALK T-DNA line: SALK_092529C) is in the first exon.

POLTERGEIST LIKE 4 (*PLL4*)

PLL4 is known to regulate shoot and leaf development in Arabidopsis. A reverse genetics study showed abnormal leaf morphology in *pll4* mutants (Song *et al*., 2005).

The T-DNA insertion in the *pll4* mutant (SALK T-DNA line: SALK_047827) is in the first exon. The *pll4* mutant was phenotyped at 22°C.

SMALL AUXIN UP RNA 38 (*SAUR38*) and *SMALL AUXIN UP RNA* 8 (*SAUR8*)

SAUR genes belong to a large family that are induced and expressed in the presence of auxin. These are putative auxin-responsive genes and it has been demonstrated that they are generally expressed in elongated hypocotyls under shade in Arabidopsis (Roig-Villanova *et al*., 2007). These genes are also temperature sensitive and in another study it has been shown that *SAUR* genes are highly expressed in hypocotyls at higher ambient temperature conditions (Franklin *et al*., 2011). More recently the role of *SAUR19-24* genes in auxin transport has also been identified (Spartz *et al*., 2012). These genes are rapidly induced in the presence of auxin in elongating tissues and positively regulate cell expansion.

For the *SAUR38* the T-DNA insertion in the *saur* mutant (SALK T-DNA line: SALK_001154C) is in the first exon. The *saur* mutant was phenotyped at 22°C.

For the *SAUR8* the T-DNA insertion in the *saur* mutant (SALK T-DNA line: SALK_058324C) is in the first exon. The *saur* mutant was phenotyped at 22°C.

TWIN SISTER OF FT (*TSF*)

TSF is a homolog of FLOWERING LOCUS T (*FT*) (Kobayashi *et al*., 1999) which is responsible for the induction of floral transition in a redundant manner with *FT* and is under the photoperiod control (Yamaguchi *et al*., 2005). The *tsf* mutants displayed a late

flowering phenotype when grown under short days and *TSF* was expressed in the hypocotyls of long day seedlings (Yamaguchi *et al*., 2005). For flowering time, ambient temperature influences *TSF* expression which is higher at 16°C than at 23°C (Blázquez *et al*., 2003). Recently, a mutant analysis showed that stomatal opening was suppressed in *tsf* mutants in the presence of light and indicated that *TSF* provided a positive effect on it (Ando *et al*., 2013).

T-DNA insertion in the *tsf* mutant (SALK T-DNA line: SALK_087522C) is in the third exon. The *tsf* mutant was phenotyped at 27°C.

TRANSPARENT TESTA 6 (*TT6*)

It was demonstrated that the *tt6* knockout mutants had a yellow seed coat (Wisman *et al*., 1998). Moreover, the *tt6* mutants displayed architectural defects in the seedlings by having long hypocotyls (Buer and Djordjevic 2009). T-DNA insertion in the *tt6* mutant (SALK T-DNA line: SALK_113321C) is in the second exon. The *tsf* mutant was phenotyped at 22°C.

4.3.4 Discussion on Knockout Methods and Results:

In this knockout analysis, the aim was to associate the candidate gene functions with hypocotyl length which is an important seedling organ for plant survival. The stem length of the seedling ensures that the seedlings elongate and reach out to light from the soil, enabling the seedling to pursue photoautotrophic growth and development. The knockout approach has been employed to deduce the functional involvement of the candidate genes. When a specific gene is disrupted with a T-DNA insertion, it usually ends up as a null gene and is not translated into a functional protein. The hypocotyl length may or may not be altered with this null gene and this can be observed by phenotypic analysis of mutants.

T-DNA insertional mutagenesis is an essential tool of functional genomics. The approach of using knockouts to test for the involvement of various QTL candidate genes in hypocotyl length function was applied, which could provide clues to their potential role in contributing to natural variation. This approach has been facilitated due to the availability

of insertion mutants and the complete genome sequence in Arabidopsis which are probably the most important functional genomics tool. The capability of creating loss of function mutations for genes using T-DNA insertional mutagenesis aids in understanding the functional consequences of a missing QTL candidate gene on the hypocotyl length phenotype.

Identification of gene functions in Arabidopsis is a major goal of developmental biology. More than 25,000 genes have been identified in Arabidopsis (Yamada *et al*., 2003; Bevan and Walsh, 2005), however, less than 10% of these genes have been assigned biological functions (Ostergaard and Yanofsky, 2004; Bouche and Bouchez, 2001). Furthermore, various accessions of Arabidopsis possess naturally existing alleles for these genes that gives rise to functional variation.

There are two primary methods to determine gene function. The analysis can work from genotype to phenotype or from phenotype to genotype. In the former method, known as reverse genetics, knowledge of the gene sequence is essential and gene functions can be identified through mutational methods. Any particular gene of known sequence can be selected and mutated. Disruption in the gene sequence by various methods would potentially knockout the gene and the probable function can be inferred by observing the phenotype. Mutations in Arabidopsis can be induced by various agents, which include mutagenic chemicals, being exposed to radiation, and insertional mutagenesis. The induced mutations in Arabidopsis are only limited to a few accessions which are Col-0, Ler-0, Ws-0 and C24. These contain a very small portion of genetic variation found in the accessions.

In the latter, known as forward genetics, one starts from an individual with an interesting phenotype, by selecting either unusual natural accessions or unusual individuals in a mutagenised population, and gene functions can be identified. Natural variation is extensively present and observed in accessions collected from around the world and there is a great potential to explore this. In some accessions, around 9.4% of the protein-coding genes have polymorphisms which result in knockout effects and hence are naturally absent (Clark *et al*., 2007). Therefore, natural variation is a resource which could also be used to identify gene functions and this has been carried out in this study for the hypocotyl length trait. Strong candidate genes have been identified by QTL analysis and

then a knockout analysis has been performed, indicating whether they play a functional role in hypocotyl length. Hence, QTL analysis can be considered as a gene discovery tool because candidate genes are identified and their functional role in natural variation is experimentally tested.

In this study, the knockout analysis was carried out to investigate if the candidate genes are functionally involved in hypocotyl length at the specific temperature conditions under which the QTL were identified. Because a certain set of genes were identified at a particular temperature condition, the phenotypic analysis of knockouts is highly informative, for it explains their potential involvement in hypocotyl length under those conditions. This further provides a strong basis to determine whether the molecular polymorphisms present in the alleles of a specific candidate gene, showing interesting functional association, are responsible for the natural variation of the trait. The establishment of a functional relationship between a candidate gene and the phenotype, under those QTL environmental conditions, paves way for further experimentation and testing by transgenic methods for allele discovery.

Agrobacterium has the natural capability to transfer a specific segment of its DNA into the genomes of plants. This segment of the Ti plasmid, known as Transfer-DNA (T-DNA), transmits the genes that allow the genetic transformation of the recipient plant genome producing crown gall tumours (Chilton *et al*., 1977). Apart from using T-DNA to incorporate foreign genes into plants, early research also revealed that the DNA sequence of plant genes can be interrupted by T-DNA causing direct disruption to its function (Feldmann *et al*., 1989; Koncz-Kalman *et al*., 1990). The T-DNA insert that is several kb in length potentially leads to complete inactivation of genes. And hence T-DNA has been exploited and utilized for insertional mutagenesis.

Agrobacterium has developed a sophisticated mechanism for the transfer of its DNA into plant genomes causing disease. There are a set of genes which the T-DNA carries which include genes for synthesis of auxin and cytokinin that control the neoplastic cell growth (Akiyoshi *et al*., 1984; Schroder *et al*., 1984) and for synthesis of opines which is the growth medium of the bacterium (Schroder *et al*., 1984; Zambryski *et al*., 1989). This T-DNA region, which is flanked by short repeat border sequences (the left and right T-DNA borders), allows the delivery of this DNA segment into the host plant genome. For the

genetic transformation to take place there is another set of virulence genes which the Ti-plasmid carries. Interestingly, this set is not found on the T-DNA region (Veluthambi *et al*., 1989; Zambryski *et al*., 1989) and does not need to be linked physically to the T-DNA.

The position of the T-DNA insertion plays an important role in the disruption of the gene. It is more effective in the exons, lesser effective in the introns and least effective in the promoters. In the promoter, the expression of a gene may be reduced or increased. Insertions in introns could be spliced out whereas insertions in the exons usually lead to null mutations and are therefore the preferred location. The protein product is truncated in this situation and hence it is less likely to retain its biological function. And therefore, knockout lines which possessed T-DNA insertions in the exons were selected for the phenotypic analysis of candidate genes.

T-DNA insertional mutagenesis provides a means of inducing mutations by foreign DNA in Arabidopsis to analyse gene function. With the integration of T-DNA, a critical portion of the coding sequence of a candidate gene is interrupted, representing a mutation and leading to loss of function. This is followed by physically observing the effect of gene inactivation on the phenotype. Due to the fact that an efficient procedure for genetic transformation is available in plants, T-DNA has been the main agent of insertional mutagenesis. For the candidate genes, obtaining T-DNA mutants is an essential step for phenotypic analysis of the trait. It is a preliminary powerful tool required to understand how the function of candidate genes is carried out.

One of the benefits of T-DNA in this study is that the foreign T-DNA disrupts the expression of the candidate genes and its effect can be monitored by the phenotypic analysis of hypocotyl length trait. One of the advantages of using the T-DNA mutagen is that it produces stable insertions in the targeted genomic DNA. Further steps for the stabilisation of the insert are not required and transformed lines can be bred for homozygous lines. As an essential powerful genomic tool for functional characterisation of genes and for linking genotypes to phenotypes, T-DNA has proven effective for genome-wide mutagenesis (Krysan *et al*., 1999). Moreover, random insertional mutagenesis on a large scale is an effective strategy in Arabidopsis because its genome is suitable due to high gene density (Bevan *et al*., 1999). As a result of this, there are several Arabidopsis populations that have been widely mutagenised with T-DNA, thus allowing

isolation of knockouts (Koncz *et al*., 1989; Feldmann 1991; Bechtold *et al*., 1993; Krysan *et al*., 1999; Galbiati *et al*., 2000; Sessions *et al*., 2002; Rosso *et al*., 2003; Alonso *et al*., 2003).

The results of the phenotypic analysis on the mutant plants show whether the hypocotyl lengths are different from WT Col-0 or the same. An altered phenotype would indicate that the missing gene in the mutant plant is functionally involved in the trait. The phenotyping of the knockout lines was carried out at the specific temperature conditions under which the QTL were identified. Therefore, those candidate genes whose mutant plants showed interesting phenotypes were further shortlisted for testing.

Though the knockout analysis is a means of a straightforward way of establishing gene function by loss of function mutations due to the insertion of the T-DNA, it misses out the effects of individual naturally occurring alleles that are found in the different 19 parental accessions. The T-DNA knockout lines focus on extreme null mutations primarily available in Col-0 background, whereas naturally existing alleles in Arabidopsis account for a broader range of genetic variation. The identification of such alleles is important for functional characterisation because there is a possibility that some null mutations could be lethal, or detectable phenotypes may not be observed in the mutants due to genetic redundancy (Bouche and Bouchez, 2001).

For identification of candidate genes and for determining their biological function in the hypocotyl length trait, a global approach has been adopted that includes various experimentation tests. Gene knockout analysis is insufficient to make conclusions and assess the role of candidate genes in natural variation of hypocotyl length but provides a strong basis on which further tests can be carried out on promising candidates. Therefore, it is necessary to enrich these findings with follow up experimentation and test for natural mutations in backgrounds other than Col-0 of specific candidate alleles using transgenesis. The functional validation of candidate genes needs to be taken to a further experimental step of testing by transgenic complementation.

The knockout analysis has provided a solid foundation upon which a complete understanding of the candidate QTL alleles involved in natural variation of hypocotyl length can be built. This has been described in detail in the next chapter.

4.4 References:

Akiyoshi D, Klee H, Amasino R, Nester E. Gordon M. (1984) T-DNA of *Agrobacterium tumefaciens* Encodes an Enzyme of Cytokinin Biosynthesis. Proc. Natl Acad. Sci. USA, 81, 5994–5998.

Alonso J, Stepanova A, Leisse T, Kim C, Chen H, *et al*. (2003) Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana.* Science, 1, 301, 5633, 653-7.

Ando E, Ohnishi M, Wang Y, Matsushita T, Watanabe A, *et al*. (2013) TWIN SISTER OF FT, GIGANTEA, and CONSTANS Have a Positive But Indirect Effect on Blue Light-Induced Stomatal Opening in Arabidopsis. Plant Physiology, 162, 1529–1538.

Bechtold N, Ellis J, Pelletier G (1993) In planta Agrobacterium Mediated Gene Transfer by Infiltration of Adult *Arabidopsis thaliana* plants. C R Acad Sci Paris, 316, 1194 -1199.

Bernhardt C. Tierney M. (2000) Expression of AtPRP3, a Proline-Rich Structural Cell Wall Protein from Arabidopsis, is regulated by Cell-type-specific Developmental Pathways Involved in Root Hair Formation. Plant Physiol., 122, 705 714.

Bevan M, Bancroft I, Mewes H, Martienssen R, McCombie R (1999) Clearing a Path through the Jungle: Progress in Arabidopsis. Genomics BioEssays, 21, 110-120.

Bevan M, Walsh S (2005) The Arabidopsis Genome: A Foundation for Plant Research. Genome Res. 15, 12, 1632-42.

Blázquez M. Weigel D. (1999) Independent Regulation of Flowering by Phytochrome B and Gibberellins in Arabidopsis. Plant Physiol, 120, 1025–1032.

Blázquez M, Ahn J, Weigel D. (2003) A Thermosensory Pathway Controlling Flowering Time in *Arabidopsis thaliana*. Nat. Genet. 33, 168–171.

Bouche N, Bouchez D. (2001) Arabidopsis Gene Knockout: Phenotypes Wanted. Current Opinion in Plant Biology. 4, 2, 568-577.

Buer C, Djordjevic M. (2009) Architectural Phenotypes in the Transparent Testa Mutants of *Arabidopsis thaliana.*, J Exp Bot. 60, 3, 751–763.

Chilton M, Drummond M, Merio D, Sciaky D, Montoya A, *et al*. (1977) Stable Incorporation of Plasmid DNA into Higher Plant Cells: The Molecular Basis of Crown Gall Tumorigenesis. Cell, 11, 263–271.

Clouse S, Sasse J. (1998) Brassinosteroids: Essential Regulators of Plant Growth and Development. Annu Rev Plant Physiol Plant Mol Biol, 49, 427–451.

Cui X, Fan B, Scholz J, Chen Z. (2007) Roles of Arabidopsis Cyclin-Dependent Kinase C Complexes in Cauliflower Mosaic Virus Infection, Plant Growth, and Development. Plant Cell. 19, 4, 1388–1402.

Datta S, Hettiarachchi G, Deng X, Holm M. (2006) Arabidopsis CONSTANS-LIKE3 Is a Positive Regulator of Red Light Signaling and Root Growth. Plant Cell, 18, 1, 70–84.

Devlin P, Patel S, Whitelam G. (1998) Phytochrome E Influences Internode Elongation and Flowering Time in Arabidopsis. Plant Cell, 10, 9, 1479-87.

Dubreucq B, Berger N, Vincent E, Boisson M, Pelletier G, *et al*. (2000), The Arabidopsis AtEPR1 Extensin-like Gene is Specifically Expressed in Endosperm During Seed Germination. The Plant Journal, 23, 643–652.

Feldmann K. (1991) T-DNA Insertion Mutagenesis in Arabidopsis: Mutational Spectrum. Plant J, 1, 71 -82.

Feldmann K, Marks M, Christianson M, Quatrano R. (1989) A Dwarf Mutant of Arabidopsis Generated by T-DNA Insertion Mutagenesis. Science, 243, 1351–1354.

Folta, K. Spalding E. (2001) Unexpected Roles for Cryptochrome 2 and Phototropin Revealed by High-Resolution Analysis of Blue Light-Mediated Hypocotyl Growth Inhibition. Plant J. 26, 471–478.

Franklin K, Davis S, Stoddart W, Vierstra R. Whitelam, G. (2003) Mutant Analyses Define Multiple Roles for Phytochrome C in Arabidopsis Photomorphogenesis. Plant Cell, 15, 1981–1989.

Franklin K, Lee S, Patel D, Kumar S, Spartz A, *et al*. (2011) PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) Regulates Auxin Biosynthesis at High Temperature. 13, 108, 50.

Galbiati M, Moreno M, Nadzan G, Zourelidou M, Dellaporta S. (2000) Large-scale T-DNA Mutagenesis in Arabidopsis for Functional Genomic Analysis. Funct Integr Genomics, 1, 25-34.

Gelvin B. (2003) Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. Microbiology and Molecular Biology Reviews, 67, 1, 16–37.

Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz E, *et al*. (1997) A Polycomb-Group Gene Regulates Homeotic Gene Expression in Arabidopsis. Nature, 386, 44–51.

Guilfoyle T, Hagen G. (2007). Auxin Response Factors. Curr. Opin. Plant Biol. 10, 453–460.

Halliday K, Koornneef M, Whitelam G. (1994) Phytochrome B and at least one other Phytochrome Mediate the Accelerated Flowering Response of *Arabidopsis thaliana* L. to Low red/far-red Ratio. Plant Physiol. 104, 1311–1315.

Halliday K, Salter M, Thingnaes E, Whitelam G. (2003) Phytochrome Control of Flowering is Temperature Sensitive and Correlates with Expression of the Floral Integrator FT. Plant J. 33, 875–885.

Halliday K, Whitelam C. (2003) Changes in Photoperiod or Temperature Alter the Functional Relationships between Phytochromes and Reveal Roles for phyD and phyE. Plant Physiol. 131, 1913–1920.

Hennig L, Stoddart W, Dieterle M, Whitelam G, Schäfer E. (2002) Phytochrome E Controls Light-Induced Germination of Arabidopsis. Plant Physiol. 128, 1, 194–200.

Kim G, Tsukaya H, Uchimiya H. (1998) The CURLY LEAF Gene Controls both Division and Elongation of Cells during the Expansion of the Leaf Blade in *Arabidopsis thaliana*, PLANTA, 206, 2, 175-183.

Kim J, Choi D, Kende H. (2003) The AtGRF Family of Putative Transcription Factors is Involved in Leaf and Cotyledon Growth in Arabidopsis. The Plant Journal, 36, 1, 94–104.

Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T. (1999) A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. Science, 286, 1960–1962.

Koncz C, Martini N, Mayerhofer R, Koncz-Kalman Z, Korber H, *et al*. (1989) High-Frequency T-DNA-Mediated Gene Tagging in Plants. Proc Natl Acad Sci USA, 86, 8467–8471.

Koncz C, Mayerhofer R, Koncz-Kalman Z, Nawrath C, Reiss B, *et al*. (1990) Isolation of Encoding Novel Chloroplast Protein by T-DNA Tagging in *Arabidopsis thaliana*. EMBO J. 9, 1337–1346.

Koornneef M, Rolff E, Spruit C. (1980) Genetic Control of Light-Inhibited Hypocotyl Elongation in *Arabidopsis thaliana* L. Heynh. Z. Pflanzenphysiol. 100, 147–160.

Krysan P, Young J, Sussman M. (1999) T-DNA as an Insertional Mutagen in Arabidopsis. Plant Cell, 11, 2283-2290.

Li H, He Z, Lu G, Lee S, Alonso J, *et al*. (2007) A WD40 Domain Cyclophilin Interacts with Histone H3 and Functions in Gene Repression and Organogenesis in Arabidopsis. Plant Cell. 19, 8, 2403–2416.

Li J, Nagpal P, Vitart V, McMorris T, Chory J. (1996). A Role for Brassinosteroids in Light-Dependent of Arabidopsis. Science, 272, 398–401.

Li J, Nam K, Vafeados D, Chory J. (2001) BIN2, A New Brassinosteroid-Insensitive Locus in Arabidopsis. Plant Physiol. 127, 1, 14-22.

Liu P, Montgomery T, Fahlgren N, Kasschau K, Nonogaki H, Carrington J. (2007) Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is Critical for Seed Germination and Post-Germination Stages. The Plant Journal, 52, 133–146.

Nagatani A, Chory J, Furuya M. (1991) Phytochrome B is Not Detectable in the hy3 Mutant of Arabidopsis, which is Deficient in Responding to end-of-day far-red Light Treatments. Plant Cell Physiol. 32, 1119–112.

Østergaard L, Yanofsky M. (2004) Establishing Gene Function by Mutagenesis in *Arabidopsis thaliana*. The Plant Journal, 39, 5, 682–696.

Reed J, Nagpal P, Poole D, Furuya M, Chory J. (1993) Mutations in the Gene for the red/far-red Light Receptor Phytochrome B alter Cell Elongation and Physiological Responses throughout Arabidopsis Development. Plant Cell, 5, 147–157.

Reed J, Nagatani A, Elich T, Fagan M, Chory J. (1994) Phytochrome A and Phytochrome B have Overlapping but Distinct Functions in Arabidopsis Development. Plant Physiol. 104, 1139-1149.

Richmond T, Somerville C. (2000) The Cellulose Synthase Superfamily. Plant Physiol. 124, 2, 495-8.

Robson F, Costa M, Hepworth S, Vizir I, Pineiro M, *et al*. (2001). Functional Importance of Conserved Domains in the Flowering-Time Gene CONSTANS Demonstrated by Analysis of Mutant Alleles and Transgenic Plants. Plant J. 28, 619–631.

Roig-Villanova I, Bou-Torrent J, Galstyan A, Carretero-Paulet L, Portoles S, *et al*. (2007) Interaction of Shade Avoidance and Auxin Responses: A Role for Two Novel Atypical bHLH Proteins. EMBO J. 26, 4756–4767.

Rosso M, Li Y, Strizhov N, Reis B, Dekker K, *et al*. (2003) An *Arabidopsis thaliana* T-DNA Mutagenized Population (GABI-Kat) for Flanking Sequence Tag-based Reverse Genetics. Plant Mol Biol, 53, 247–259.

Sakai T, Kagawa T, Kasahara M, Swartz T, Christie J, *et al*. (2001) Arabidopsis nph1 and npl1: Blue Light Receptors that Mediate both Phototropism and Chloroplasts Relocation. PNAS USA, 98, 6969–6974.

Sakamoto K, Briggs W. (2002) Cellular and Subcellular Localization of Phototropin 1. Plant Cell, 14, 1723–1735.

Song S, Clark S. (2005) POL and Related Phosphatases are Dosage-Sensitive Regulators of Meristem and Organ development in Arabidopsis. Dev. Biology, 285, 1, 272–284.

Sharrock R, Quail P. (1989) Novel Phytochrome Sequences in *Arabidopsis thaliana*: Structure, Evolution, and Differential Expression of a Plant Regulatory Photoreceptor Family. Genes Dev. 3, 11, 1745-57.

Schroder G, Waffenschmidt S, Weiler E, Schroder J. (1984) The T-region of Ti Plasmids Codes for an Enzyme Synthesizing Indole-3-Acetic Acid. Eur. J. Biochem. 138, 387–391.

Schumacher K, Chory J. (2000) Brassinosteroid Signal Transduction: Still Casting the Actors. Curr. Opin. Plant Biol. 3, 79–84.

Sessions A, Burke E, Presting G, Aux G, McElver J. (2002) A Highthroughput Arabidopsis Reverse Genetics System. Plant Cell, 14, 2985–2994.

Simpson G, Gendall A, Dean C. (1999) When to Switch to Flowering. Annu Rev Cell Dev Biol, 99, 519–550.

Smith H, Whitelam G. (1997) The Shade Avoidance Syndrome: Multiple Responses Mediated by Multiple Phytochromes. Plant Cell Environ. 20, 840–844.

Somers D, Sharrock R, Tepperman J, Quail P. (1991) The hy3 Long Hypocotyl Mutant of Arabidopsis is Deficient in Phytochrome B. Plant Cell 3, 1263–1274.

Spartz A, Lee S, Wenger J, Gonzalez N, Itoh H, *et al*. (2012) The SAUR19 Subfamily of SMALL AUXIN UP RNA Genes Promote Cell Expansion. Plant J, 70, 978–990.

Stowe-Evans E, Luesse D, Liscum E. (2001) The Enhancement of Phototropin-Induced Phototropic Curvature in Arabidopsis occurs via a Photoreversible Phytochrome A-Dependent Modulation of Auxin Responsiveness. Plant Physiol. 126, 826–834.

Taylor N, Howells R, Huttly A, Vickers K, Turner S. (2003) Interactions among Three Distinct CesA Proteins Essential for Cellulose Synthesis. Proc. Natl Acad. Sci. USA, 100, 1450-1455.

Tiwari S, Hagen G, Guilfoyle T. (2003) The Roles of Auxin Response Factor Domains in Auxin-Responsive Transcription. Plant Cell, 15, 533–543.

Turner S, Somerville C. (1997) Collapsed Xylem Phenotype of Arabidopsis identifies Mutants Deficient in Cellulose Deposition in the Secondary Cell Wall. Plant Cell, 9, 5, 689-701.

Van der Knaap E, Kim J, Kende H. (2000) A Novel Gibberellin-Induced Gene from Rice and its Potential Regulatory Role in Stem Growth. Plant Physiol. 122, 695–704.

Veluthambi K, Krishnan M, Gould J, Smith R, Gelvin S. (1989) Opines Stimulate Induction of the Vir Genes of the *Agrobacterium tumefaciens* Ti Plasmid. J. Bacteriol. 171, 3696–3703.

Wagner D, Tepperman J, Quail P. (1991) Overexpression of Phytochrome B Induces a Short Hypocotyl Phenotype in Transgenic Arabidopsis, Plant Cell. 3, 12, 1275-1288.

Wang G, Kong H, Sun Y, Zhang X, Zhang W, *et al*. (2004) Genome-Wide Analysis of the Cyclin Family in Arabidopsis and Comparative Phylogenetic Analysis of Plant Cyclin-like Proteins. Plant Physiol. 135, 2, 1084-99.

Wang J, Wang L, Mao Y, Cai W, Xue H, *et al*. (2005) Control of Root Cap Formation by MicroRNA-Targeted Auxin Response Factors in Arabidopsis, Plant Cell, 17, 2204-2216.

Whitelam G, Devlin P. (1997) Roles of Different Phytochromes in Arabidopsis Photomorphogenesis. Plant Cell Environ. 20, 752–758.

Wisman E, Hartmann U, Sagasser M, Baumann E, Palme K, *et al*. (1998) Knock-out Mutants from an En-1 Mutagenized *Arabidopsis thaliana* Population Generate Phenylpropanoid Biosynthesis Phenotypes, Proc. Natl. Acad. Sci. USA, 95, 12432–12437.

Yamada K, Lim J, Dale J, *et al*. (2003) Empirical Analysis of Transcriptional Activity in the Arabidopsis Genome. Science, 302, 842–846.

Yamaguchi A, Kobayashi Y, Goto K, Abe M, Araki T. (2005) TWIN SISTER OF FT (TSF) Acts as a Floral Pathway Integrator Redundantly with FT. Plant Cell Physiol. 46 8, 1175-89.

Cheng Y, Qin G, Dai X, Zhao Y. (2008) NPY Genes and AGC Kinases Define Two Key Steps in Auxin-Mediated Organogenesis in Arabidopsis, PNAS, 105, 52, 21017-21022.

Zambryski P, Tempe J, Schell J. (1989) Transfer and Function of T-DNA Genes from Agrobacterium Ti and Ri Plasmids in Plants. Cell, 56, 193–201.

4.5 Appendix:

List of candidate genes and primer information for T-DNA knockout lines:

Gene	Primer Information for T-DNA lines
AT2G18790 *PHYB*	SALK_069700.46.10.x PRODUCT_SIZE 668 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.46 LP ATGTACCTGAATTCTGTGCGG Len 21 TM 60.01 GC 47.62 SELF_ANY_COMPL 0.46 3'_COMPL 0.00 RP ATTGGGGCTTTCTGTTTCTTG Len 21 TM 60.47 GC 42.86 SELF_ANY_COMPL 0.46 3'_COMPL 0.00 Insertion chr2 8143030 BP+RP_PRODUCT_SIZE 272-572
AT2G23380 *CLF*	SALK_088542.54.90.x PRODUCT_SIZE 638 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.07 LP TAGGTTCAGGAGAATGGTTGG Len 21 TM 59.04 GC 47.62 SELF_ANY_COMPL 1.07 3'_COMPL 0.00 RP GAATGTGTAGTGGTGGGGATG Len 21 TM 60.10 GC 52.38 SELF_ANY_COMPL 1.07 3'_COMPL 0.00 Insertion chr2 9959378 BP+RP_PRODUCT_SIZE 296-596
AT2G23380 *CLF*	SALK_139371.50.70.x PRODUCT_SIZE 640 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.50 LP TTGCCTGTATAAGATGGTCGG Len 21 TM 59.97 GC 47.62 SELF_ANY_COMPL 0.50 3'_COMPL 0.00 RP AAATGTAAGCCCCTCGCATAG Len 21 TM 60.46 GC 47.62 SELF_ANY_COMPL 0.50 3'_COMPL 0.00 Insertion chr2 9956619 BP+RP_PRODUCT_SIZE 283-583
AT2G24400 *SAUR*	SALK_001154.29.99.f PRODUCT_SIZE 648 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.19 LP AGGAACAGAGGTCCACATCAC Len 21 TM 59.02 GC 52.38 SELF_ANY_COMPL 1.19 3'_COMPL 0.00 RP AAAGCAGCCAACAACAACAAC Len 21 TM 60.20 GC 42.86 SELF_ANY_COMPL 1.19 3'_COMPL 0.00 Insertion chr2 10378295 BP+RP_PRODUCT_SIZE 247-547
AT2G24400 *SAUR*	SALK_001155.48.55.x PRODUCT_SIZE 633 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.22 LP CATCACGGTTTCAATGATCTG Len 21 TM 58.99 GC 42.86 SELF_ANY_COMPL 1.22 3'_COMPL 0.00 RP AAAGCAGCCAACAACAACAAC Len 21 TM 60.20 GC 42.86 SELF_ANY_COMPL 1.22 3'_COMPL 0.00 Insertion chr2 10378297 BP+RP_PRODUCT_SIZE 245-545
AT3G45780 *PHOT1*	SALK_088841.45.35.x PRODUCT_SIZE 626 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.02 LP TCACGATTGCTCCCATTAAAG Len 21 TM 60.08 GC 42.86 SELF_ANY_COMPL 0.02 3'_COMPL 0.00 RP TCTCCGATTTTGTCATGAAGG Len 21 TM 60.06 GC 42.86 SELF_ANY_COMPL 0.02 3'_COMPL 0.00 Insertion chr3 16820402 BP+RP_PRODUCT_SIZE 251-551
AT3G45780 *PHOT1*	SALK_146058.55.25.x PRODUCT_SIZE 566 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.11 LP TCACGGAATATAGCCGTGAAG Len 21 TM 60.10 GC 47.62 SELF_ANY_COMPL 0.11 3'_COMPL 0.00 RP AATGACATTGCGAACTGGTTC Len 21 TM 59.99 GC 42.86 SELF_ANY_COMPL 0.11 3'_COMPL 0.00 Insertion chr3 16821572 BP+RP_PRODUCT_SIZE 244-544

AT3G51240 *TT6*	SALK_113321.21.00.x PRODUCT_SIZE 625 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.28 LP AGTGGATACACTGTGGCATCC Len 21 TM 59.87 GC 52.38 SELF_ANY_COMPL 0.28 3'_COMPL 0.00 RP GCTGTGCAAGATTGGAGAGAG Len 21 TM 60.15 GC 52.38 SELF_ANY_COMPL 0.28 3'_COMPL 0.00 Insertion chr3 19026019 BP+RP_PRODUCT_SIZE 254-554
AT3G51240 *TT6*	SALK_061570.56.00.x PRODUCT_SIZE 579 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.67 LP CACAAAACACACCGAGCCTAG Len 21 TM 60.73 GC 52.38 SELF_ANY_COMPL 0.67 3'_COMPL 0.00 RP CCTTGCTGCTACAAGACCAAG Len 21 TM 60.06 GC 52.38 SELF_ANY_COMPL 0.67 3'_COMPL 0.00 Insertion chr3 19026312 BP+RP_PRODUCT_SIZE 243-543
AT3G52910 *GRF4*	SALK_037642.56.00.x PRODUCT_SIZE 560 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.09 LP GGAAGTGATATTTCCAAGGCAG Len 22 TM 59.97 GC 45.45 SELF_ANY_COMPL 0.09 3'_COMPL 0.00 RP CAGCAGAAGTAGCAGTTTGGG Len 21 TM 60.06 GC 52.38 SELF_ANY_COMPL 0.09 3'_COMPL 0.00 Insertion chr3 19618268 BP+RP_PRODUCT_SIZE 226-526
AT3G52910 *GRF4*	SALK_077829.19.30.x PRODUCT_SIZE 633 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.15 LP ATACATGTTGGCTGGTGCTTC Len 21 TM 60.01 GC 47.62 SELF_ANY_COMPL 0.15 3'_COMPL 0.00 RP AAAGAAGATGCGGTTGTTGTG Len 21 TM 60.16 GC 42.86 SEL_ANY_COMP 0.15 3'_COMPL 0.00 Ins chr3 19617438 BP+RP_PRO_SIZE 295-595
AT4G18130 *PHYE*	SALK_092529.39.10.n PRODUCT_SIZE 585 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.14 LP GGATTTGGAGCCTGCTAAATC Len 21 TM 60.05 GC 47.62 SELF_ANY_COMPL 0.14 3'_COMPL 0.00 RP AACTCACAAGCATACCGCAAC Len 21 TM 60.19 GC 47.62 SELF_ANY_COMPL 0.14 3'_COMPL 0.00 Insertion chr4 10044964 BP+RP_PRODUCT_SIZE 236-536
	Same Primers for both of these lines as insertion site is the same
AT4G18130 *PHYE*	SALK_040131.50.45.x PRODUCT_SIZE 585 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.14 LP GGATTTGGAGCCTGCTAAATC Len 21 TM 60.05 GC 47.62 SELF_ANY_COMPL 0.14 3'_COMPL 0.00 RP AACTCACAAGCATACCGCAAC Len 21 TM 60.19 GC 47.62 SELF_ANY_COMPL 0.14 3'_COMPL 0.00 Insertion chr4 10044962 BP+RP_PRODUCT_SIZE 234-534
AT4G18610 *LSH9*	SALK_039868.20.35.x PRODUCT_SIZE 547 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 2.07 LP TCCTTCTTGACAACAACGACC Len 21 TM 60.14 GC 47.62 SELF_ANY_COMPL 2.07 3'_COMPL 0.00 RP TTCCCCTAATAATGTCTTCGG Len 21 TM 58.08 GC 42.86 SELF_ANY_COMPL 2.07 3'_COMPL 0.00 Insertion chr4 10250903 BP+RP_PRODUCT_SIZE 220-520
AT4G18610 *LSH9*	GABI_419F10 PRODUCT_SIZE 696 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.57 LP TGCATGCAACGTACAGAAATG Len 21 TM 60.71 GC 42.86 SELF_ANY_COMPL 0.57 3'_COMPL 0.00 RP TTGTGTCTTCTTCGGACAACC Len 21 TM 60.14 GC 47.62 SELF_ANY_COMPL 0.57 3'_COMPL 0.00 Insertion chr4 10251258 BP+RP_PRO_SIZE 298-598

AT4G18780 *CESA8*	SALK_046685.56.00.x PRODUCT_SIZE 626 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.02 LP AGAAACAGCATGCTTGCTAGG Len 21 TM 59.67 GC 47.62 SELF_ANY_COMPL 0.02 3'_COMPL 0.00 RP GACTCTGTAAACCCGTCAACG Len 21 TM 59.65 GC 52.38 SELF_ANY_COMPL 0.02 3'_COMPL 0.00 Insertion chr4 10313904 BP+RP_PRODUCT_SIZE 226-526
AT4G18780 *CESA8*	GABI_339E12 PRODUCT_SIZE 602 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.68 LP TTTAGGGTTCTTTGTGATGCG Len 21 TM 60.12 GC 42.86 SELF_ANY_COMPL 0.68 3'_COMPL 0.00 RP AAGATACGACGTCGCTTAACG Len 21 TM 59.44 GC 47.62 SELF_ANY_COMPL 0.68 3'_COMPL 0.00 Insertion chr4 12018722 BP+RP_PRO_SIZE 220-520
	GABI_339E12 PRODUCT_SIZE 656 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.33 LP AATGTTGCGGTTAGCATATCG Len 21 TM 60.00 GC 42.86 SELF_ANY_COMPL 0.33 3'_COMPL 0.00 RP CGTTAACCGATGATTGAATCG Len 21 TM 60.33 GC 42.86 SELF_ANY_COMPL 0.33 3'_COMPL 0.00 Insertion chr4 10314972 BP+RP_PROD_SIZE 265-565
AT4G20370 *TSF*	SALK_087522.53.25.x PRODUCT_SIZE 555 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.72 LP TTATGGTGGATCCAGATGTGC Len 21 TM 60.74 GC 47.62 SELF_ANY_COMPL 0.72 3'_COMPL 0.00 RP AAGATATGGTGCAGTGCAACC Len 21 TM 60.01 GC 47.62 SELF_ANY_COMPL 0.72 3'_COMPL 0.00 Insertion chr4 11001694 BP+RP_PRODUCT_SIZE 230-530
AT4G20370 *TSF*	GABI_585C12 PRODUCT_SIZE 672 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.08 LP TTCTTGACAAGTTGCTGCTCTC Len 22 TM 59.80 GC 45.45 SELF_ANY_COMPL 1.08 3'_COMPL 0.00 RP CGGTTAACTTGATTTTGTTTCG Len 22 TM 58.72 GC 36.36 SELF_ANY_COMPL 1.08 3'_COMPL 0.00 Insertion chr4 11002318 BP+RP_PROD_SIZE 272-572
AT2G16580 *SAUR*	SALK_058324.53.25.x PRODUCT_SIZE 588 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.77 LP TCCGTGGAAATAAAAAGCCTC Len 21 TM 60.43 GC 42.86 SELF_ANY_COMPL 0.77 3'_COMPL 0.00 RP AGTGAGCTCGAGAAACTGTGG Len 21 TM 59.66 GC 52.38 SELF_ANY_COMPL 0.77 3'_COMPL 0.00 Insertion chr1 26151014 BP+RP_PRODUCT_SIZE 249-549
	SALK_058324.50.00.x PRODUCT_SIZE 601 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.60 LP GCTTTCTTGCCAGATCTTAACC Len 22 TM 59.40 GC 45.45 SELF_ANY_COMPL 0.60 3'_COMPL 0.00 RP GCAATGCTTAGGCAGATTCTG Len 21 TM 60.00 GC 47.62 SELF_ANY_COMPL 0.60 3'_COMPL 0.00 Insertion chr2 7186725 BP+RP_PRODUCT_SIZE 264-564
	SALK_058324.50.00.x PRODUCT_SIZE 588 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.77 LP TCCGTGGAAATAAAAAGCCTC Len 21 TM 60.43 GC 42.86 SELF_ANY_COMPL 0.77 3'_COMPL 0.00 RP AGTGAGCTCGAGAAACTGTGG Len 21 TM 59.66 GC 52.38 SELF_ANY_COMPL 0.77 3'_COMPL 0.00 Insertion chr1 26151015 BP+RP_PRODUCT_SIZE 248-548
AT2G16580 *SAUR*	SALK_003272.49.45.x PRODUCT_SIZE 601 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.60 LP GCTTTCTTGCCAGATCTTAACC Len 22 TM 59.40 GC 45.45 SELF_ANY_COMPL 0.60 3'_COMPL 0.00 RP GCAATGCTTAGGCAGATTCTG Len 21 TM 60.00 GC 47.62 SELF_ANY_COMPL 0.60 3'_COMPL 0.00 Insertion chr2 7186743

	BP+RP_PRODUCT_SIZE 246-546
AT2G14820 *NPY2*	SALK_058416.55.50.x PRODUCT_SIZE 656 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.98 LP TAAGAACGGTTTGGTGGTCTG Len 21 TM 60.02 GC 47.62 SELF_ANY_COMPL 0.98 3'_COMPL 0.00 RP ACAAGTGCTTGCACTTTTGTG Len 21 TM 59.04 GC 42.86 SELF_ANY_COMPL 0.98 3'_COMPL 0.00 Insertion chr2 6359297 BP+RP_PRODUCT_SIZE 287-587
AT2G14820 *NPY2*	SALK_142094.45.55.n PRODUCT_SIZE 669 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.73 LP CATTTTGCTTCTCTGATTGGC Len 21 TM 59.84 GC 42.86 SELF_ANY_COMPL 0.73 3'_COMPL 0.00 RP TCTAGGCACAAAAGTGCAAGC Len 21 TM 60.57 GC 47.62 SELF_ANY_COMPL 0.73 3'_COMPL 0.00 Insertion chr2 6358939 BP+RP_PRODUCT_SIZE 297-597
AT2G28350 *ARF10*	SALK_087247.50.00.x PRODUCT_SIZE 606 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.47 LP TCTCTGGTTTGTTTGGTTCTTG Len 22 TM 59.26 GC 40.91 SELF_ANY_COMPL 0.47 3'_COMPL 0.00 RP AGAGGATAAGAGGTGGAACGC Len 21 TM 59.73 GC 52.38 SELF_ANY_COMPL 0.47 3'_COMPL 0.00 Insertion chr2 12114319 BP+RP_PRODUCT_SIZE 280-580
AT2G28350 *ARF10*	GABI_274H01 PRODUCT_SIZE 654 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.18 LP AGGATCAATGGTTCAAATCCC Len 21 TM 60.01 GC 42.86 SELF_ANY_COMPL 0.18 3'_COMPL 0.00 RP AATCCAGGGTAAGGATTGTCG Len 21 TM 60.19 GC 47.62 SELF_ANY_COMPL 0.18 3'_COMPL 0.00 Insertion chr2 12114831 BP+RP_PRODUCT_SIZE 300-600
AT2G28890 *PLL4*	SALK_047827.46.25.x PRODUCT_SIZE 652 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.03 LP CGCCACACCACTCTCTACTTC Len 21 TM 59.92 GC 57.14 SELF_ANY_COMPL 0.03 3'_COMPL 0.00 RP AGAGAGTAAATAGTCCGGCGC Len 21 TM 59.89 GC 52.38 SEL_ANY_COM 0.03 3'_COMPL 0.00 Ins chr2 12407346 BP+RP_PRO_SIZE 262-562
AT2G28890 *PLL4*	SALK_047818.16.90.x PRODUCT_SIZE 652 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.03 LP CGCCACACCACTCTCTACTTC Len 21 TM 59.92 GC 57.14 SELF_ANY_COMPL 0.03 3'_COMPL 0.00 RP AGAGAGTAAATAGTCCGGCGC Len 21 TM 59.89 GC 52.38 SELF_ANY_COMPL 0.03 3'_COMPL 0.00 Insertion chr2 12407346 BP+RP_PRODUCT_SIZE 262-562
AT2G23050 *NPY4*	SALK_151725.55.25.x PRODUCT_SIZE 599 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.10 LP TTGGTGATTGGACTAAATCGC Len 21 TM 59.95 GC 42.86 SELF_ANY_COMPL 0.10 3'_COMPL 0.00 RP AACATGCGGTACAGAACGAAC Len 21 TM 60.05 GC 47.62 SELF_ANY_COMPL 0.10 3'_COMPL 0.00 Insertion chr2 9812467 BP+RP_PRODUCT_SIZE 265-565
AT2G23050 *NPY4*	SALK_046452.51.25.x PRODUCT_SIZE 629 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.15 LP GCGATTCTATTTTTGTTGACCC Len 22 TM 59.85 GC 40.91 SELF_ANY_COMPL 0.15 3'_COMPL 0.00 RP AGAAGAAGAGAGTATGCGGGC Len 21 TM 60.00 GC 52.38 SELF_ANY_COMPL 0.15 3'_COMPL 0.00 Insertion chr2 9812376 BP+RP_PRODUCT_SIZE 231-531

AT2G30950 *VAR2*	SAIL_253_A03 PRODUCT_SIZE 589 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.09 LP GTTGGTAGGCAAAGAGGAACC Len 21 TM 59.99 GC 52.38 SELF_ANY_COMPL 0.09 3'_COMPL 0.00 RP CACTTACCCACACACTGCATG Len 21 TM 60.08 GC 52.38 SELF_ANY_COMPL 0.09 3'_COMPL 0.00 Insertion chr2 13176138 BP+RP_PRO_SIZE 232-532
AT3G49120 *PCB*	SAIL_143_G09 PRODUCT_SIZE 630 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.18 LP ATAGGCAGGAGGTCCTTCTTG Len 21 TM 59.72 GC 52.38 SELF_ANY_COMPL 0.18 3'_COMPL 0.00 RP TTGGTGTAATGTTTCCCATCC Len 21 TM 59.54 GC 42.86 SELF_ANY_COMPL 0.18 3'_COMPL 0.00 Insertion chr3 18209766 BP+RP_PRO_SIZE 271-571
AT3G49120 *PCB*	GABI_728F08 PRODUCT_SIZE 657 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.18 LP GAAAAAGAAAATGCCTCCAGC Len 21 TM 60.20 GC 42.86 SELF_ANY_COMPL 0.18 3'_COMPL 0.00 RP TGTTTTGGCATCATCATTGTG Len 21 TM 60.37 GC 38.10 SELF_ANY_COMPL 0.18 3'_COMPL 0.00 Insertion chr3 18208160 BP+RP_PRO_SIZE 294-594
AT2G24790 *COL3*	SAIL_361_A08 PRODUCT_SIZE 587 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.58 LP TTCCTTTTGTGAAAGTCCGTG Len 21 TM 60.13 GC 42.86 SELF_ANY_COMPL 0.58 3'_COMPL 0.00 RP ACAGCGTCGTAGAAAGGAGTG Len 21 TM 59.55 GC 52.38 SELF_ANY_COMPL 0.58 3'_COMPL 0.00 Insertion chr2 10567096 BP+RP_PRO_SIZE 261-561
AT2G30520 *RPT2*	SAIL_140_D03 PRODUCT_SIZE 624 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.44 LP CTTGCAGGTCAAGGAAGTGTC Len 21 TM 59.90 GC 52.38 SELF_ANY_COMPL 1.44 3'_COMPL 0.00 RP CTGTTGTCAGACTATGATTGCG Len 22 TM 58.46 GC 45.45 SELF_ANY_COMPL 1.44 3'_COMPL 0.00 Insertion chr2 13004284 BP+RP_PRO_SIZE 247-547
AT3G44600 *CYP71*	SALK_024686.20.75.x PRODUCT_SIZE 626 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.19 LP GCTAACCTGCATTGCTGAAAC Len 21 TM 59.90 GC 47.62 SELF_ANY_COMPL 0.19 3'_COMPL 0.00 RP CTTTGGAACCCTCTCTGTTCC Len 21 TM 60.10 GC 52.38 SELF_ANY_COMPL 0.19 3'_COMPL 0.00 Insertion chr3 16168816 BP+RP_PRODUCT_SIZE 284-584
AT3G44600 *CYP71*	SALK_050092.55.00.x PRODUCT_SIZE 640 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.21 LP TGCCACCAACTTAACCTGATC Len 21 TM 59.99 GC 47.62 SELF_ANY_COMPL 0.21 3'_COMPL 0.00 RP ATCCTCTTGGGGATGGTACTG Len 21 TM 60.20 GC 52.38 SEL_AN_COM 0.21 3'_COMPL 0.00 Ins chr3 16165705 BP+RP_PRO_SIZE 251-551
AT3G49670 *BAM2*	SAIL_1053_E09 PRODUCT_SIZE 654 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.98 LP CGACTTTTCAAGTTTTTAATTTGTG Len 25 TM 58.47 GC 28.00 SELF_ANY_COMPL 1.98 3'_COMPL 0.00 RP AGGAGAGTTCATCGGGAAATG Len 21 TM 60.45 GC 47.62 SELF_ANY_COMPL 1.98 3'_COMPL 0.00 Insertion chr3 18417979 BP+RP_PRO_SIZE 276-576
AT3G49670 *BAM2*	GABI_791G02 PRODUCT_SIZE 620 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.47 LP CGTTTCAAAACAACAAGGCTC Len 21 TM 59.78 GC 42.86 SELF_ANY_COMPL 0.47 3'_COMPL 0.00 RP CCGGAGAGGTTAAGACCTGAG Len 21 TM 60.25 GC 57.14 SELF_ANY_COMPL 0.47 3'_COMPL 0.00 Insertion chr3 18417856

	BP+RP_PRO_SIZE 232-532
AT4G19600 *CYCT1*	SALK_139324.42.80.x PRODUCT_SIZE 634 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.43 LP TTTCTCGTCTTTTCACCTTGG Len 21 TM 59.34 GC 42.86 SELF_ANY_COMPL 0.43 3'_COMPL 0.00 RP CACCTCGACAACTAGAGGGTG Len 21 TM 59.77 GC 57.14 SELF_ANY_COMPL 0.43 3'_COMPL 0.00 Insertion chr4 10675741 BP+RP_PRODUCT_SIZE 300-600
AT4G19600 *CYCT1*	SALK_139322.15.30.x PRODUCT_SIZE 618 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.11 LP TGACCGTGACATCAGCTGTAC Len 21 TM 59.77 GC 52.38 SELF_ANY_COMPL 0.11 3'_COMPL 0.00 RP TTGCTGCGAAGTTCCTTAAAG Len 21 TM 59.66 GC 42.86 SELF_ANY_COMPL 0.11 3'_COMPL 0.00 Insertion chr4 10675648 BP+RP_PRODUCT_SIZE 279-579
AT2G27380 *EPR1*	GABI_710H02 PRODUCT_SIZE 643 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.60 LP AAAAGCTGCCATTAAGGGATC Len 21 TM 59.58 GC 42.86 SELF_ANY_COMPL 0.60 3'_COMPL 0.00 RP TGGGAGGACTATAAGTTGGGG Len 21 TM 60.19 GC 52.38 SELF_ANY_COMPL 0.60 3'_COMPL 0.00 Insertion chr2 11713514 BP+RP_PRO_SIZE 291-591
AT4G18710 *BIN2*	GABI_244F08 PRODUCT_SIZE 652 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.07 LP CGAGATTGCTTGAAGTCGAAC Len 21 TM 60.01 GC 47.62 SELF_ANY_COMPL 0.07 3'_COMPL 0.00 RP TACCTTATCATCAGCCATGGC Len 21 TM 59.94 GC 47.62 SELF_ANY_COMPL 0.07 3'_COMPL 0.00 Insertion chr4 10296332 BP+RP_PRO_SIZE 259-559
	GABI_244F08 PRODUCT_SIZE 641 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.11 LP CGTTTGTATCTTCCCTCGAATC Len 22 TM 59.97 GC 45.45 SELF_ANY_COMPL 0.11 3'_COMPL 0.00 RP GTTCGAGTCTCTCACATTCGG Len 21 TM 59.86 GC 52.38 SELF_ANY_COMPL 0.11 3'_COMPL 0.00 Insertion chr4 10035274 BP+RP_PRO_SIZE 280-580
AT3G51200 *SAUR*	GABI_695B04 PRODUCT_SIZE 656 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.66 LP GTTGCCATGTCAAGCAACTAG Len 21 TM 58.45 GC 47.62 SELF_ANY_COMPL 0.66 3'_COMPL 0.00 RP ATCATCATGTGAGGAAGCCTC Len 21 TM 59.11 GC 47.62 SELF_ANY_COMPL 0.66 3'_COMPL 0.00 Insertion chr3 19018977 BP+RP_PRO_SIZE 267-567
	GABI_695B04 PRODUCT_SIZE 577 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 1.39 LP TCGTTAATCTCATCGTCTCCC Len 21 TM 59.14 GC 47.62 SELF_ANY_COMPL 1.39 3'_COMPL 0.00 RP ACTTTGAGTCCGTTAAACCGC Len 21 TM 60.53 GC 47.62 SELF_ANY_COMPL 1.39 3'_COMPL 0.00 Insertion chr1 6136804 BP+RP_PRO_SIZE 225-525
AT2G17950 *WUS*	SAIL_150_G06 PRODUCT_SIZE 623 PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.36 LP CCATTTACACCACTAGCTATTACATG Len 26 TM 58.27 GC 38.46 SELF_ANY_COMPL 0.36 3'_COMPL 0.00 RP ACCATAGATCCATAGACATGGC Len 22 TM 57.90 GC 45.45 SELF_ANY_COMPL 0.36 3'_COMPL 0.00 Insertion chr2 7809606 BP+RP_PROD_SIZE 223-523

Chapter 5

Functional Analysis 2 – QTL Cloning by Transgenic Allelic Complementation

5.1 Introduction:

The ultimate aim of QTL mapping is to reveal the underlying genes that control a quantitative trait. This stage of the project is aimed at fulfilling this ultimate objective. This can be accomplished by direct testing of candidate alleles to explore if they influence natural variation.

Upon discovery and isolation of QTL, the next major step is the identification of the causal genes (Salvi and Tuberosa, 2005), which is regarded as one of the greatest challenges for geneticists this century (Luo *et al*., 2002). To determine the molecular mechanism for functional variation, it is vital to identify the underlying alleles of the QTL and this process is known as QTL cloning (Borevitz and Chory, 2004). QTL cloning still remains an important challenge (Weigel and Nordborg, 2005) and it allows candidate gene verification by functional experimental methods. One of these methods is transgenic complementation and it represents a gold standard of proof that a candidate gene is actually controlling natural variation (Drinkwater and Gould, 2012). This is a definitive step to confirm a QTL by introducing alleles into QTL parental accessions and null knockout mutants (Borevitz and Chory, 2004). Such a transgenic approach should identify alleles that have significantly different effects on the hypocotyl length. This *in vivo* analysis consisting of phenotypic complementation experiments is thus required; which is an essential part of the 'burden of proof' to determine the causal genes responsible for the variation in the trait.

While considerable progress has been made in mapping QTL, explaining the underlying molecular basis of QTL has remained a bottleneck and a challenge. The identification of a QTL allele that causes natural variation requires laborious cloning techniques. There are very few instances where QTL have been cloned particularly in plant growth and

morphology traits (Frary *et al*., 2000; El-Assal *et al.*, 2001). Despite the challenges, Arabidopsis as a model plant offers several advantages for QTL studies. For example, QTL cloning is feasible because its genome is relatively small, genetic resources are available and candidate genes can be easily tested with transgenic transformation to directly examine the effects of their different alleles.

A combination of follow up experimental methods is essential for definitive identification of causal QTL alleles which include molecular genetic, functional and mechanistic studies that unequivocally link the genotype to the phenotype (Weigel and Nordborg 2005). Therefore, a three tier post QTL mapping strategy was adopted in this study. The high level of available molecular genetic information gained from post QTL analyses in the first two tiers (candidate gene identification and knockout analysis) for hypocotyl length in response to temperature can be further exploited by transgenic analysis and potentially this can yield in QTL cloning. This chapter comprises of the third and final tier, which is functional testing of candidates by transgenic complementation experiments. This tier combined with the first tier (candidate gene identification) and second tier (knockout analysis) provides a rigorous understanding of the candidates and is a means of QTL cloning.

The mutational analysis in Chapter 4 has identified several potential candidate genes by ascertaining their biological involvement in the trait. Candidate gene knockouts provided a causal relationship between function of a gene and its sequence, i.e. a direct way of determining function. In this chapter, QTL cloning aims at discovering novel functional alleles of candidate genes for natural variation in hypocotyl length at the QTL specific temperatures. Therefore, a subset of these genes, based on their phenotypes in the knockout analysis, has been selected for direct experimentation by transgenic analysis to test for their functional roles. This analysis tests whether the identified mutations in the candidates could possibly be influencing natural variation in the hypocotyl length phenotype in the corresponding parental accessions. The allelic functional approach targets QTL to measure the effect of allelic variation at that particular locus. Functions of new alleles can be determined by merging the findings from knockout analysis with functional allelic testing. This leads the path from QTL mapping to QTL cloning which is the last step in the forward genetic analysis.

QTL mapping is a statistical association study that leads to the prioritisation of candidate genes. Any such association study needs to be verified by functional experimentation and transgenesis provides an *in vivo* understanding of the candidate alleles. Transgenic plants have been used as a tool for allele discovery and confirmation of hypocotyl length QTL natural variation. Whether these candidate genes play a vital role in natural variation of the hypocotyl length is an interesting question and is addressed by the transgenic complementation analysis. Following the genetic dissection of the hypocotyl length by QTL mapping, the functional analysis of Arabidopsis candidate genes is pursued. A direct method of QTL cloning is to complement phenotypes of the major effect size accessions by using homozygous transgenic lines. The parental accessions selected for allelic complementation are based on the data analysis of their effect sizes that are contributing to the respective QTL as described in Chapter 3.

The aims for transgenic allelic complementation, which is a brute force genetics approach, are as follows:

- To find the molecular basis of quantitative variation in the hypocotyl length phenotype in response to temperature
- To directly investigate the phenotypic effects of the different candidate alleles in major effect parents that may contribute to the variation of the trait in response to temperature
- To establish a link between allelic variation at candidate loci and the phenotype in response to temperature from which gene function can be inferred
- To provide an insight in to the mechanistic control of hypocotyl length by several candidate genes in multiple accessions

5.2 Methods and Materials:

5.2.1 Laboratory Methods:

This section describes the methods and materials used in the laboratory for different experiments. These methods are as follows:

5.2.1.1 Extracting and Amplifying Genomic DNA for Candidate Genes:

The DNA from the plants of each of the accessions was extracted and PCR performed as explained in the methods section in Chapter 4. The extension time was set up according to the total length of the genomic DNA, i.e. 30 seconds per kb.

The genomic region for the candidate gene fragments were obtained and amplified by PCR using gene specific primers listed in table 5.1 in the appendix. These primers were designed to amplify the endogenous promoters and the coding sequence for each of the candidate genes. The forward PCR primers were designed with sequence CACC at its 5´ end to allow directional cloning. To make sure that the PCR product was inserted directionally in the pENTR, the reverse PCR primer was designed in such a manner that it was not complementary to the overhang sequence GTGG at the 5´ end. The composition of the PCR mix was as follows:

DNA = 1 µl, dNTPs = 1 µl, primer 1 = 1 µl, primer 2 = 1 µl, DMSO = 0.3 µl, Phusion = 0.1 µl, water = 13.6 µl, buffer = 2 µl

5.2.1.2 Purification of Genomic DNA:

The 1% Agarose gel was observed under UV light for the detection of the PCR amplified DNA bands. The DNA bands from the gel were cut out under an UV transilluminator. The DNA exposure to UV was minimised to avoid any DNA damage. The PCR amplified genomic DNA was purified from the gel using the QIAquick Gel Extraction Kit (from QIAGEN) with following protocol:

1. Place the gels in 1.5 ml tubes
2. Add 3 volumes of QG buffer to 1 volume gel
3. Heat at 40°C until the gel is completely dissolved

4. Add 1 gel volume of 100% Isopropanol to the sample and mix

5. No need to leave longer and pipette the sample in the columns to bind the DNA

6. Spin for 1 min at 13,000 RPM and discard the flow-through

7. Repeat step 6 if there are more samples

8. To wash, add 750 µl PE buffer and leave the column for 5 minutes

9. Spin for 1 minute at 13,000 RPM and remove the flow-through

10. Spin again for 1 minute at 13,000 RPM to completely remove the flow-through

11. Place the column into a clean 1.5 ml tube

12. To elute DNA, add 30 µl EB buffer to the centre of the column and leave for 2 minutes

13. Centrifuge for 1 min at 13, 000 RPM and collect DNA in the 1.5 ml tube

5.2.1.3 Ligation into pENTR_TOPO Vector:

The PCR amplified and purified genomic DNA of the various candidate genes were ligated into separate pENTR vectors. The Gateway cloning system requires initially the insertion of genomic DNA into pENTR which has two flanking recombination sequences called “att L 1” and “att L 2”, as can be seen in figure 5.1, which develops a ‘Gateway Entry clone’ (Invitrogen nomenclature).

The pENTR transformation took place by adding the following materials and leaving at room temperature for 1 hour:

Materials for pENTR_TOPO reaction:

PCR product = 2 µl,

Salt solution = 0.5 µl,

TOPO Vector = 0.5 µl

The virtual pENTR constructs were prepared in the software ‘Gene Construction Kit’ (GCK). These various constructs for all the candidate genes functionally tested for transgenic complementation experiments can be viewed in Figure 5.1.

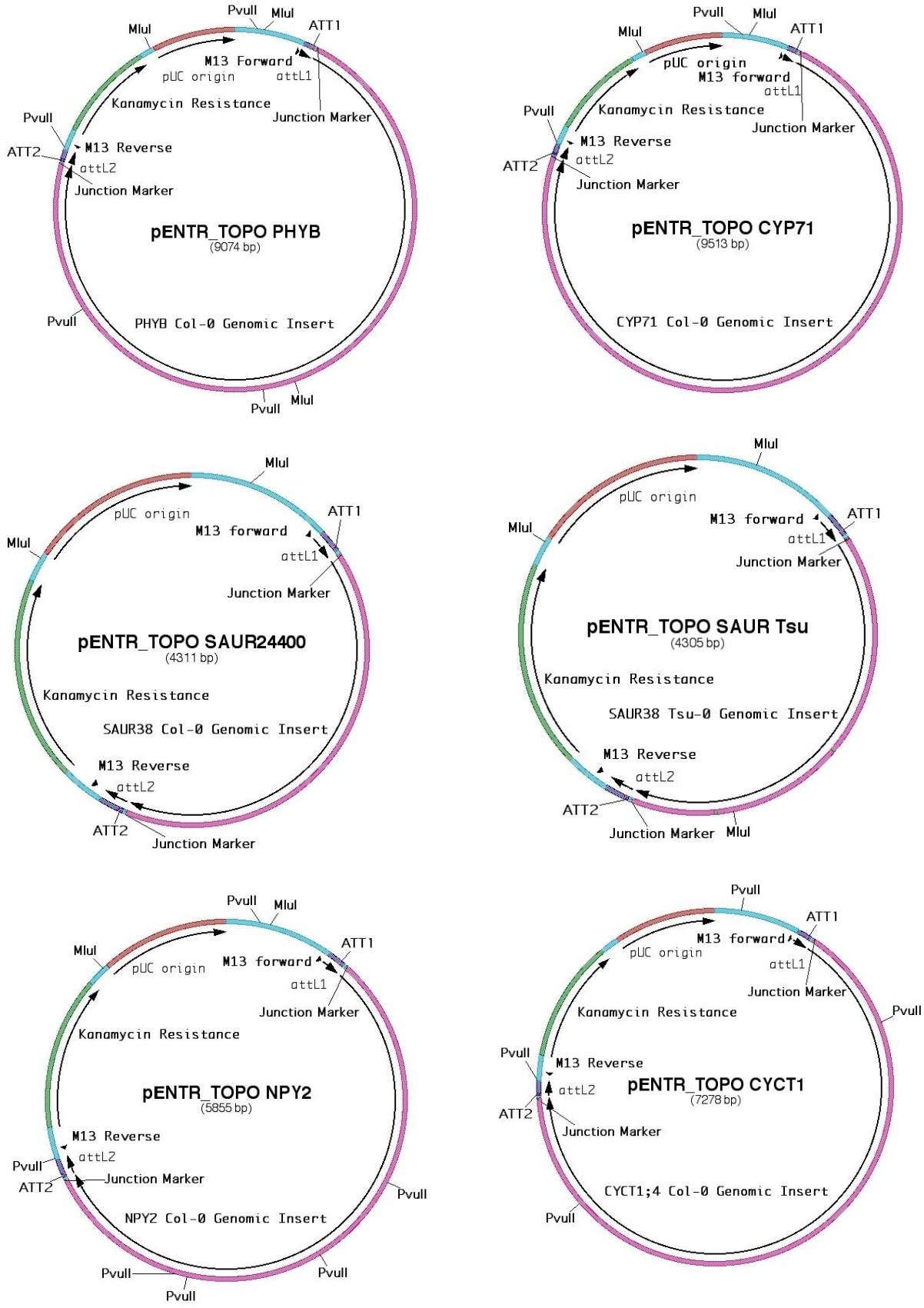

Figure 5.1: pENTR vector constructs of various candidate genes used in gene cloning.

5.2.1.4 Transformation of *E. coli* with pENTR:

TOP10 chemically competent bacterial cells (from Invitrogen) were transformed with the pENTR constructs according to the following protocol:

1. Add 10 µl of TOP10 cells to the tubes which have been incubated at room temperature for 1 hour after the pENTR reaction

2. Leave for 30 minutes on ice

3. Place at 42°C in a water-bath for 45 seconds for the heat shock

4. Place on ice for 5 minutes

5. Add 450 µl of SOC and recover at 37°C

6. Plate transformed cells on LB + Kanamycin selection and incubate at 37°C overnight

7. Pick 6 colonies from each of the plates and put them in the small glass jars containing 10 ml LB and 10 µl Kanamycin

8. Incubate overnight at 37°C in the shaker and then the growth medium is ready for plasmid DNA isolation from *E. coli*

5.2.1.5 Purification of Plasmid DNA from *E. coli*:

For the isolation of plasmid DNA from *E. coli*, the NucleoSpin Plasmid Kit (from Macherey-Nagel) was used with the following protocol:

1. To cultivate and harvest bacterial cells put 2 ml of *E. coli* LB culture in 2 ml tubes and centrifuge at 11,000 x g (rcf) for 1 minute

2. For cell lysis, add 250 µl A1 buffer and re-suspend pellet by vortexing so that no cell clumps should remain

3. Add 250 µl A2 buffer and mix gently by inverting the tube 10 times, do not vortex

4. Incubate at room temperature for 5 minutes or until lysate appears clear

5. Add 300 µl A3 buffer and mix thoroughly by inverting the tube 10 times, do not vortex

6. To clarify the lysate centrifuge for 6 minutes at 11,000 x g, repeat if the supernatant is not clear

7. To bind DNA place a column in a collection tube 2 ml

8. Take 750 µl of the supernatant and pipette into the column

9. Centrifuge for 1 minute at 11,000 x g

10. Discard flow-through and place the column back into the collection tube

11. To wash silica membrane, add 500 µl pre-heated buffer AW to the column

12. Centrifuge at 11,000 x g for 1 min

13. Add 600 µl A4 buffer to the column and centrifuge at 11,000 x g for 1 min

14. Discard the flow-through and place the column back into the empty collection tube

15. To dry silica membrane, centrifuge the column for 2-3 minutes at 11,000 x g and discard the collection tube

16. To elute DNA, place the column in a 1.5 ml tube and add 50 µl AE buffer

17. Incubate at room temperature for 1 minute

18. Centrifuge at 11,000 x g for 1 minute and collect the eluted DNA in 1.5 ml tube

5.2.1.6 Digestion of Plasmid DNA:

The purified plasmid DNA quantity was measured with a Nanodrop spectrophotometer. This helped in identifying colonies which possessed the genomic inserts because such colonies measured a significantly higher quantity of DNA than those colonies which did not possess the genomic inserts. To analyse the plasmid DNA from the selected colonies, restriction analysis was carried out to confirm the presence of the genomic inserts in the plasmids. Virtual gels were prepared in the GCK software for all candidate genes assisting in selecting the restriction enzymes. The reagents used in the digestion were:

Plasmid DNA = 2 µl, Buffer 10x React 3 = 1 µl, water = 7 µl, restriction enzyme = 0.1 µl

The reaction was incubated at 37°C for 3 hours. DNA was loaded into a gel for gel electrophoresis and the DNA bands were visualised under UV light. The DNA bands in the gel were compared against the virtual gels and were found to be correct.

5.2.1.7 Sequencing of pENTR vectors:

After the successful identification of the correct clones which carry the genomic inserts, to confirm the proper orientation of the inserts in the vector, the constructs were sequenced with pENTR M13 F and R Primers. The M13 F Primer is: 5´-GTAAAACGACGGCCAG-3´ and the M13 R Primer is: 5´-CAGGAAACAGCTATGAC-3´.

Once the genomic DNA of the candidate genes was confirmed in the pENTR vectors, the next step was to sequence the whole genomic DNA of candidate genes to identify any polymorphisms that may have arisen in the cloning process. Primers were designed with an interval of about 500 bp and the DNA was sequenced. The pENTR sequenced DNA was compared to the already existing reference sequence of Col-0 on TAIR's website and interestingly no mutations were found, showing successful cloning of the candidate genes. The software DNA Strider 4.2 was used to tally the DNA sequences.

5.2.1.8 Gateway Clonase Reaction for Expression Vector:

Gateway technology (Invitrogen) was used for the cloning of all the plasmid constructs by recombination of PCR products into destination vectors. After the entry clone was obtained, an LR recombination reaction was performed using the Gateway LR Clonase II enzyme mix in order to transfer the candidate gene from the pENTR entry construct into the destination vector pJHA212B to generate an expression clone. This vector carries a gluphosinate resistance marker ($Basta^R$) and confers Basta resistance in plants. The T-DNA lines have transgenes that confer Kanamycin resistance, therefore Basta resistance was used. The materials required for the reaction were:

Purified plasmid DNA of entry clone = 0.4 µl

Destination vector = 0.4 µl

Gateway LR Clonase II enzyme = 0.3 µl

TE Buffer, pH 8.0 (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) = 0.3 µl

2 µg/µL Proteinase K solution = 0.5 µl

TOP10 competent *E. coli* cells = 10 µl

The LR reaction was performed as follows:

1. Add 0.4 µl purified plasmid DNA of entry clone to 0.4 µl of destination vector
2. Incubate at room temperature for 2 hours
3. Terminate the reaction with 0.5 Proteinase K solution
4. Incubate at 37°C for 10 minutes
5. Add 10 µl TOP10 competent cells to the reaction
6. Incubate on ice for 30 minutes

7. Heat shock at 42°C for 45 seconds
8. Add 250 µl SOC and recover in shaker at 37°C for 1 hour
9. Plate transformed cells: LB + Spectinomycin 100 (1 µl/ml) and incubate overnight at 37°C

5.2.1.9 Purification of Expression Vector DNA from *E. coli*:

The plasmid DNA of the expression vector was purified from the transformed *E. coli* cells according to the protocol in section 5.2.1.5

5.2.1.10 Digestion of Expression Vector DNA:

To analyse the plasmid DNA from the selected colonies, restriction analysis was carried out according to the protocol in section 5.2.1.6 to confirm the presence of the genomic inserts in the expression vectors. Virtual gels were prepared in the GCK software for all candidate genes. The DNA bands in the gel were compared against the virtual gels and were found to be correct.

5.2.1.11 Sequencing of Expression Vector pJHA212B:

With the digestion confirmation of the presence of the candidate genes in the colonies of expression vector, the next step was to sequence the DNA of the vector. Since the whole genomic DNA of candidate genes was sequenced in the entry pENTR vector, therefore, in the expression vector sequencing only with the F and R primers was performed to further confirm the presence of the genomic inserts and their orientation. The candidate gene sequences in the vector were matched with the Col-0 genomic sequence by using DNA Strider 1.4 software.

5.2.1.12 Transformation of Agrobacterium:

The Agrobacterium strain GV3101 was transformed with the expression vector, containing the gene of interest, with the following protocol:

1. Put the cuvettes on ice
2. Thaw out an aliquot of Agrobacterium (50-100 µl) on ice

3. Dilute the plasmid (miniprep) 1:40

4. Prepare tubes with 1 ml of SOB or YEP (important without antibiotic)

5. Settings of the Gene Pulser (electroporator) should be: capacitance extender = 25, pulse controller = 400, in Gene Pulser press Set volts to 2.5

6. Take a tube containing the Agro, add 1 µl of the previously diluted plasmid

7. Introduce the cells into a cuvette, dry out the external surface with some paper

8. Put the cuvette the white tray and bring it to the end

9. Press PULSE, then you will hear a beep that indicates that the electroporation is done

10. Press TIME CONST (in the Gene Pulser) a value around 9 is very good

11. Put the YEP and SOB in the cuvette and mix it with the cells, collect the mixture and put it back in the tube

12. If there is a bang during the pulse, throw away this cuvette as it is not useful any more

13. Incubate the cells for at least 2-4 hours

14. For the GV3101 prepare 100 ml LB agar plates with the following antibiotics (100 µl Rifampicin + 100 µl Gentamyicin + 80 µl Spectinomycin)

15. Spin cells down (6000 rpm for 1 minute), remove 700 µl of medium and resuspend the cells. Pour cells on the plates containing the antibiotics and incubate at 30°C for 48 hours

5.2.1.13 Purification of Plasmid DNA from Agrobacterium:

The plasmid DNA of the expression vector was purified from the transformed Agrobacterium cells according to the protocol in section 5.2.1.5

5.2.1.14 Digestion of Plasmid DNA:

To analyse the plasmid DNA from the selected Agrobacterium colonies, restriction analysis was carried out according to the protocol in section 5.2.1.6 to confirm the presence of the genomic inserts in the expression vectors in the Agrobacterium. Virtual gels were prepared in the GCK software for all candidate genes. The DNA bands in the gel were compared against the virtual gels and were found to be correct.

5.2.1.15 Plant Transformation by Floral Dipping:

The technique used for plant transformation of the various accessions was Agrobacterium-mediated transformation using in the direct floral dipping method (Clough and Bent, 1998). The protocol for dipping the plants is as follows:

1. Grow Agrobacterium culture in selective medium from starter culture to about OD600 (0.6 – 0.9) in ~ 200 ml YEP or LB. (Can dilute saturated starter 1:1000 into 250 ml, harvest next day). Use the spectrophotometer to measure the optical density of the growth medium

2. Harvest cells at 4000 rpm for 10 minutes at room temperature

3. Resuspend pellet in 0.5 x MS with 5% sucrose (per litre: 2.2 g MS, 50 g sucrose). Use deionised water, should have pH 5.7. (Resuspend in equivalent volume of media for OD600 to be 0.6 – 0.9)

4. Add Silwet L-77 to 0.02 % i.e. 200 µl per litre

5. Dip plants in suspension. It is best to dip plants with secondary inflorescences. For this purpose, cut primary shoots and grow for 1 week until there are many buds but no siliques

6. Keep overnight covered in bags

7. Place the plants in growth incubators and allow them to form transgenic seed

5.2.2 Glasshouse Methods:

This section describes the methods and materials used in the glasshouses for different experiments. These methods are as follows:

5.2.2.1 Production of Genetically Engineered Plants:

Transgenic plants were produced by the direct floral dipping method as described above and the T1 seeds obtained from the florally dipped T0 plants were harvested. This section describes the raising of T0 plants leading all the way to identification of T3 homozygous transgenic lines. A large collection of transgenic lines for different candidate genes and various Arabidopsis accessions were generated. The screening in the T1 plants for

transformed individuals and in the subsequent T2 and T3 generations for the determination of heterozygous and homozygous individuals was based on glufosinate ammonium (Basta) resistance (2.5 ml/1 litre harvest water of Basta). The precise genotype of large numbers of progeny in each generation was determined by Basta resistance. The T-DNA insertion carries the Basta resistance gene and is expressed in transformed individuals making them resistant. During each of the successive selection steps, three rounds of Basta herbicide were applied to rule out any false positive seedlings.

5.2.2.2 Identification of Independently Transformed T1 Lines:

Numerous T0 plants were transformed for various accessions for the different transgenic experiments. Following the floral dipping, the T-DNA was transferred from the Agrobacterium to the genomic DNA of developing Arabidopsis ovaries in the various accessions. The transformed ovaries were allowed to self-pollinate which yielded T1 seeds. The T1 seeds are the primary transformants in the hemizygous state. These T1 seeds were grown under Basta selection to identify individuals containing a T-DNA carrying the gene of interest. Transgenic lines were selected by screening the T1 seed progeny in soil.

Since the T-DNA insertion into the genomic DNA takes place in a completely random manner, therefore, multiple independently transformed lines were required to assess the effect of the insertion site on the phenotype. For this purpose, the seed from the T-DNA containing T1 primary transformants were individually collected, each establishing an independently transformed T-DNA line. A high transformation efficiency of 2-4% was observed in the T1 transgenic plants. This allowed for the identification of independently transformed lines for all the different constructs that were used, which are essential for understanding the effect of the transgene on the phenotype.

Figure 5.2: T1 transformed plants growing that have survived the Basta spray. These plants carry the Basta resistant transgene as a selectable marker. The seeds that were not transformed during the floral dipping produced seedlings that were susceptible and all those plants died.

Using Basta for the screening for T1 transformed heterozygous plants proved very effective. In figure 5.2 the dead organic material can be seen lying on the soil of those plants which did not survive the herbicide and died. The resistant transformed hemizygous individuals were left behind to grow as independent lines.

5.2.2.3 Identification of Independently Transformed T2 and T3 Lines:

The identification of hemizygous and homozygous lines was determined by the segregation analysis of the transgene in T2 seedlings grown under herbicide selection. In the T2 generation, as mentioned above, all the homozygous untransformed plants died after the Basta spray leaving behind only transformed hemizygous and homozygous plants. For the identification of independently transformed stable homozygous transgenic lines, 15 T2 progeny of a single self-pollinated T1 plant were grown. These 15 T2 plants for each transgenic line that survived the selection were randomly chosen and were selfed to bulk up and set T3 seed. Since the transgene usually segregates in a Mendelian fashion, therefore 25% of these should be homozygous for the TDNA.

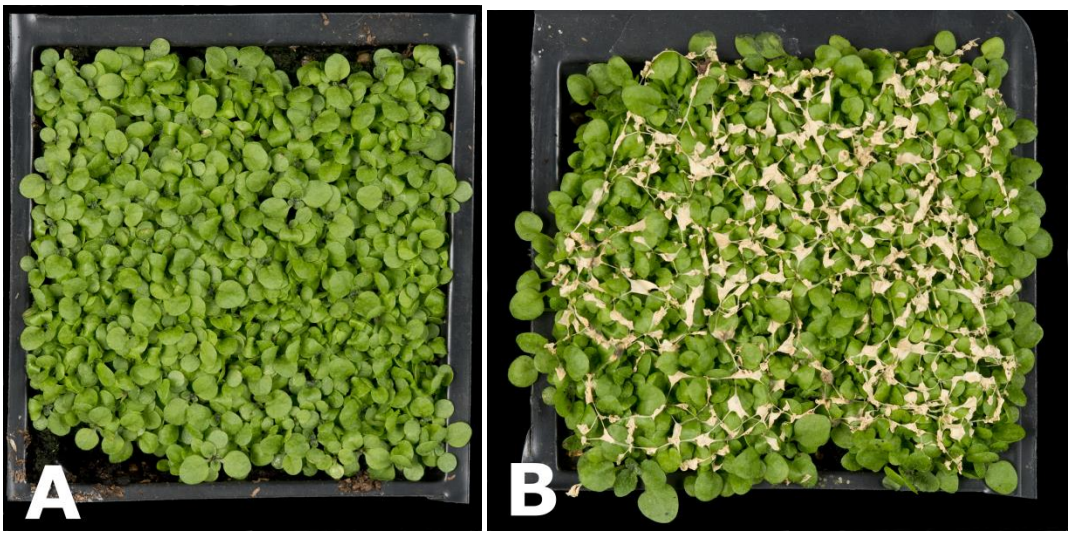

Figure 5.3: Screen for homozygous lines; Image A shows a T2 homozygous line where 100% of the seedlings are resistant to Basta; and Image B shows a T2 hemizygous line where the transgene is still segregating and 25% of the homozygous untransformed seedlings have died after Basta spraying, the dead organic matter can be seen on the top of other seedlings.

T3 generation seeds were harvested from each T2 transformed plant and aliquots of these seeds were tested for herbicide resistance. They were sown on soil and seedlings were grown for one week, as part of the screening experiment. They were sprayed with three rounds of Basta herbicide. Those stable T3 lines in which a 100% of the T3 progeny of a single T2 plant were resistant to Basta indicated that the T2 parent was homozygous for the T-DNA. There was no segregation of the transgene taking place and hence the phenotypes of all the seedlings were uniform. The segregation in the hemizygous lines was also obvious due to non-uniform phenotypes. Therefore, the T3 seeds of homozygous lines were used for the phenotyping of the hypocotyl length. For all of the six constructs, multiple independent homozygous T3 lines were identified.

5.2.2.4 Identification of Transgene Copy Number:

To test whether a particular transgenic line had a single copy of the transgene, a transgene segregation analysis was carried out. For each transgenic line, about 200 T2 seeds were sown in a pot evenly distributed and were allowed to grow for 10 days. Images of these pots were taken from above with a digital camera and the total number of seedlings were counted using Image J. After counting the total seedlings, the pots were

sprayed with three rounds of Basta and the total number of homozygous untransformed seedlings that died after the spray was counted. In this way only the transformed hemizygous and homozygous individuals were left in the pots. For each pot, the total number of seedlings that died was deducted from the total number of seedlings and in this way the total number of resistant seedlings was identified. This was then compared against a 3:1 ratio and statistically tested with a Chi square test. This analysis identified transgenic lines which were likely to carry only a single copy transgene number.

5.2.2.5 Phenotypic Analysis of T3 Homozygous Plants:

The phenotypic analysis was performed in a similar manner as described in the previous chapters.

5.3A Results: Transgenic Complementation with *PHYB*$_{Col-0}$ allele:

The results in this section describe in detail the analysis of functional variation of *PHYB* alleles in the 'Temp22.2' QTL in hypocotyl length phenotype in response to 22°C temperature in multiple Arabidopsis accessions. Multiple major effect parental accessions have been experimentally tested for complementation with the *PHYB*$_{Col-0}$ allele at 22°C. The results of these experiments have been described in detail in individual sub sections for each accession separately. These results provide the definitive experimental evidence for the *PHYB* QTL alleles by individual transgenic allelic complementation studies. Before embarking on a detailed transgenic analysis to determine if *PHYB* plays a role in the natural variation of hypocotyl length in response to temperature in multiple major effect parental accessions, it is necessary to establish whether the long *phyB* KO hypocotyl phenotype is rescued by the re-introduction of *PHYB*$_{Col-0}$ allele. The main objective for this experiment is to validate the functional role of *PHYB*.

5.3A.1 Transgenic Complementation of *phyB* KO with *PHYB*$_{Col-0}$ allele:

The hypothesis for this experiment was that the WT Col-0 phenotype could be restored in the *phyB* knockout mutant with the *PHYB*$_{Col-0}$ allele. This was a positive control; *PHYB* from Col-0 should rescue the WT phenotype in *phyB* mutants. From the KO analysis, it was known that *PHYB* is functionally involved in the hypocotyl length phenotype. This experiment that would validate its functional role was performed for the complementation of *phyB* KO mutant with the *PHYB*$_{Col-0}$ allele. The *phyB* KO plants were transformed with the *PHYB*$_{Col-0}$ construct. This was a gene rescue experiment which is considered as a standard testing procedure to prove that the T-DNA mutated allele is controlling the phenotype. For this purpose, the WT copy *PHYB*$_{Col-0}$ allele, including all introns, exons and the 5' flanking promoter sequence and 3' flanking sequence was transformed directly into the *phyB* KO mutants.

5.3A.1.1 Phenotypic Analysis of Homozygous T3 *phyB* KO Lines:

As seen in figure 5.4, the phenotyping results revealed that the Col-0 WT phenotype was rescued in the *phyB* mutant by the expression of the QTL candidate transgene. The T3

phyB KO transgenic Line 1 carrying the $PHYB_{Col\text{-}0}$ construct had complemented and reverted back to the Col-0 WT phenotype and the seedlings exhibited the WT short hypocotyl phenotype. This experiment validated the functional role of *PHYB* already identified in controlling hypocotyl length from the KO analysis in Chapter 4.

Moreover, in order to test for the effect of *PHYB* gene copy number in Col-0 WT and *phyB* KO plants, a control experiment was performed in which these plants were transformed with the $PHYB_{Col\text{-}0}$ construct. The Col-0 WT only has a single copy number of the *PHYB* gene. As seen in figure 5.4, it is interesting to note that with the introduction of the $PHYB_{Col\text{-}0}$ allele into Col-0 WT plants, the T3 Col-0 lines carrying the $PHYB_{Col\text{-}0}$ construct (Line 1, Line 2 and Line 3) are T3 independently transformed homozygous transgenic lines, and showed an even shorter hypocotyl (exaggerated) because they each have an estimated double copy number of *PHYB*.

Similarly, the double copy number KO transgenic line 2 carrying the $PHYB_{Col\text{-}0}$ construct, which was used as a control to observe for the effect of copy number in the transgenic KO plants, also showed an exaggerated shorter phenotype exactly similar to the T3 Col-0 plants transformed with the $PHYB_{Col\text{-}0}$ construct. This independently transformed homozygous T3 KO line also possessed an estimated double copy number of the *PHYB* allele. These results from the T3 Col-0 transgenic and KO lines indicate that the differential hypocotyl growth responses are sensitive to *PHYB* copy number. The gene copy numbers in the independently transformed transgenic lines were estimated from the segregation analysis in the T2 generation. The ratios of Basta-resistant to Basta-sensitive seedlings were determined. The hypocotyl length phenotypes shown in figure 5.4 inversely correlate with the *PHYB* gene copy number, i.e. fewer gene copy number, bigger hypocotyl phenotype and vice versa.

- Long homozygous KO null mutant phenotype has zero *PHYB* gene,
- Col-0 WT and 'Line1 T3 *phyB* KO + $PHYB_{Col\text{-}0}$' have an estimated single copy number.
- T3 Col-0 + $PHYB_{Col\text{-}0}$ (3 lines) and 'Line 2 T3 *phyB* KO + $PHYB_{Col\text{-}0}$' have an estimated double copy number. These have the shortest exaggerated phenotype.

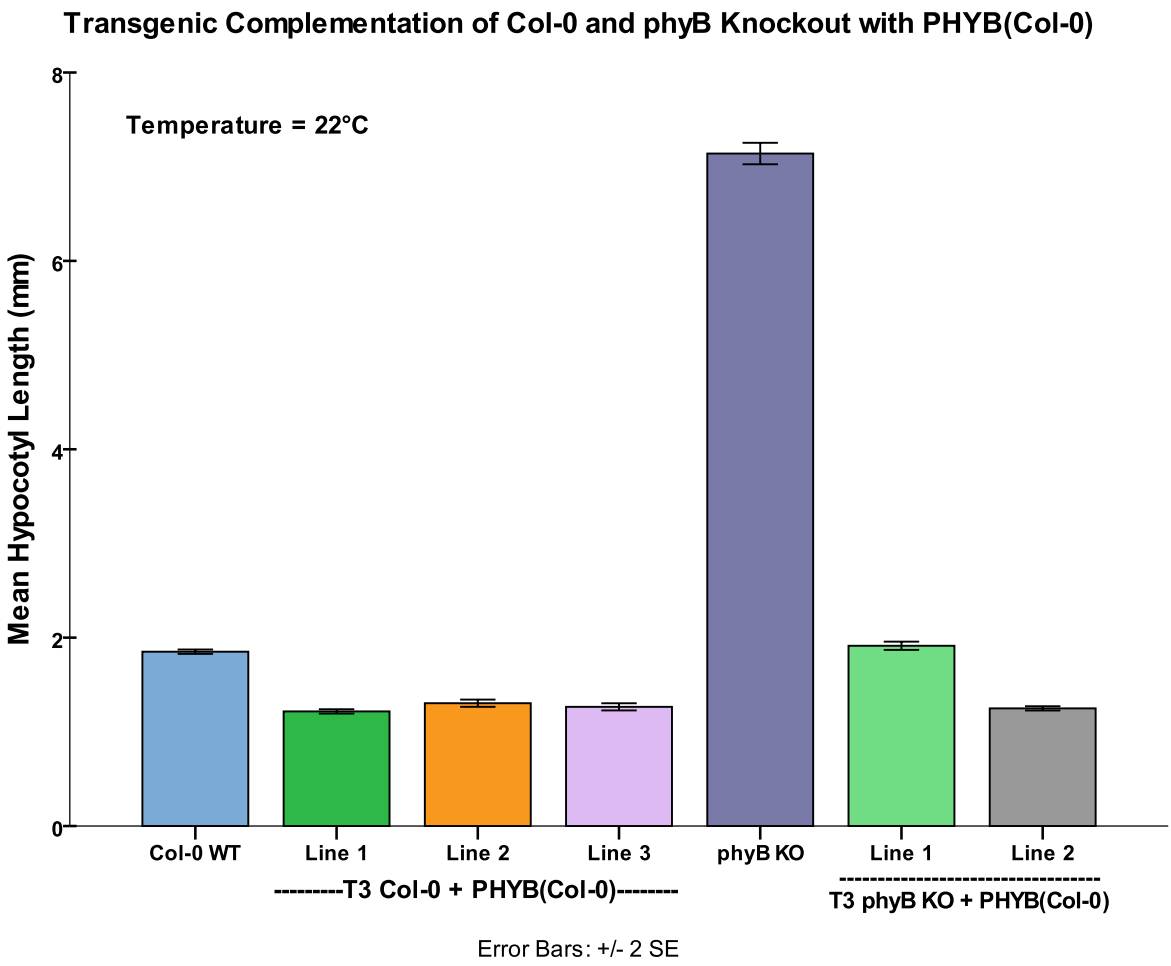

Figure 5.4: Transgenic Col-0 and *phyB* lines showing complementation of phenotypes and effect of gene copy numbers. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

Figure 5.5: A T2 *phyB* line transformed with $PHYB_{Col\text{-}0}$ construct. The tall seedlings are homozygous untransformed which died after Basta spray. The short are transformed and Col-0 WTphenotype has been rescued in these knockouts.

Figure 5.6: A T3 *phyB* homozygous line transformed with $PHYB_{Col\text{-}0}$ construct. All seedlings are small and the phenotype has been complemented.

Since the *phyB* KO phenotype is well characterised and the plant architectural phenotypes of hypocotyl length, plant size and petiole length are known, it was interesting to test for the effect if *PHYB* copy number in the vegetative growth stage and compare it with the effect on hypocotyl length phenotype in the various lines. An experiment was performed to compare the hypocotyl length phenotype with the plant size and petiole length phenotypes of the vegetative stage. The results showed that the hypocotyl length phenotypes of the various lines in figure 5.4 strikingly corresponded to the plant size and petiole length phenotypes of the same lines as seen in figure 5.7. In a similar way to the hypocotyl length phenotype, the three T3 Col-0 lines and Line 2 of T3 *phyB* KO carrying the $PHYB_{Col\text{-}0}$ construct showed an exaggerated smaller plant size and petiole length phenotype as compared to the Col-0 WT plants due to double *PHYB* copy number. Line 1 of T3 *phyB* KO carrying the $PHYB_{Col\text{-}0}$ construct showed a similar phenotype to Col-0 WT rescuing the vegetative phenotypes of plant size and petiole length.

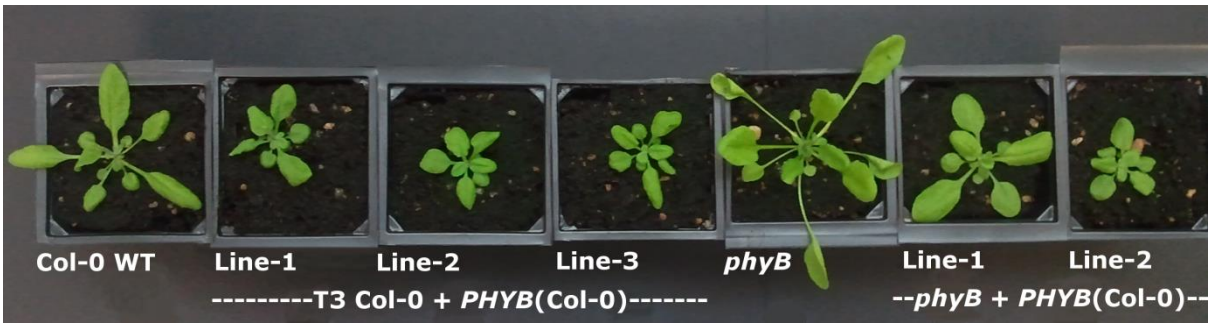

Figure 5.7: Effect of $PHYB_{Col\text{-}0}$ allele on vegetative phenotypes (petiole length and plant size) of transgenic Col-0 and *phyB* plants. These phenotypes are proportionate to hypocotyls in figure 5.4.

5.3A.2 Transgenic Complementation of Ct-1 accession with *PHYB*$_{Col-0}$ allele:

Transgenic complementation experiments were designed and performed for the validation of *PHYB* as a candidate gene in the Ct-1 accession. The aim was to test the hypothesis that allelic variation at *PHYB* candidate locus contributed to the observed phenotypic differences between the Ct-1 and Col-0 accessions. Is the Ct-1 long hypocotyl phenotype due to mutations in Ct-1 allele of *PHYB*? If yes, then with the insertion of *PHYB*$_{Col-0}$ allele into Ct-1, the long hypocotyl length phenotype of Ct-1 should be transformed to the short phenotype of Col-0.

The hypothesis is that the Ct-1 allele has a loss of function mutation in *PHYB*, as has been observed in the phenotype in figure 5.8 which is the same as *phyB* KO. Therefore, if a functional allele from Col-0 is inserted into the genome of Ct-1 we expect to restore the WT phenotype, as has been experimentally demonstrated is the case in transgenic *phyB* lines in figures 5.4 and 5.5.

For the detection of independently transformed single transgene copy Ct-1 lines carrying the *PHYB*$_{Col-0}$ construct, at least 10 independent Basta-resistant primary transformants/regenerants (T1 generation) were selected. These transformants were allowed to self and set T2 generation seed. The T2 segregation data for each independently transformed Ct-1 line was analysed. The empirical data was generated as explained in the methods section. To test if the observed data was significantly different from the expected data, a chi-square test was performed; and the results are shown in table 5.3 in the appendix.

The chi-square results (p = 0.05) show that amongst the 10 independent lines, 7 showed a close to 3:1 segregation for Basta-resistance in the T2 generation, which indicated a single locus insertion of the transgene. The presence of a single locus of the transgene was shown by the monogenic segregation of the Basta resistance marker.

5.3A.2.1 Phenotypic Analysis of Homozygous T3 Ct-1 Lines:

As can be observed in figure 5.8, the Ct-1 accession, which has non-synonymous protein-coding sequence polymorphism in the *PHYB* allele, resembles the *phyB* mutant in having an elongated hypocotyl, which implies that the mutations in the Ct-1 *PHYB* allele could be causative of the long hypocotyl, as outlined in the hypothesis. Similar to *phyB*, Ct-1 exhibits an approximately 250% increase in hypocotyl length when compared to Col-0. The transgenic complementation results can be seen in the phenotypic analysis of the independently transformed single copy T3 homozygous Ct-1 lines 1 – 4 carrying the $PHYB_{Col\text{-}0}$ construct. The Ct-1 accession having a long hypocotyl phenotype was transformed, due to the $PHYB_{Col\text{-}0}$ allele, to the short phenotype of Col-0 WT. This very vividly provides direct allelic evidence for natural variation by restoring the short hypocotyl and shows that the Ct-1 allele is a positive growth regulator of hypocotyl length as opposed to the Col-0 allele which is a negative growth regulator of hypocotyl length.

To test for natural variation, five independently T3 homozygous transformed Ct-1 lines carrying the $PHYB_{Col\text{-}0}$ construct were phenotyped. For each independent line, two different but identical plants (offspring of two different plants) were phenotyped and scored (as a control) which showed similar hypocotyl phenotypes. From the segregation analysis in the T2 generation (explained in section 5.2.2.4), Line 5 is estimated to have double copy number of *PHYB*, hence being shorter than the others and showing an exaggerated short phenotype. The remaining four Ct-1 transgenic lines (Line 1 – 4) have similar short phenotypes and have complemented the Col-0 WT phenotype.

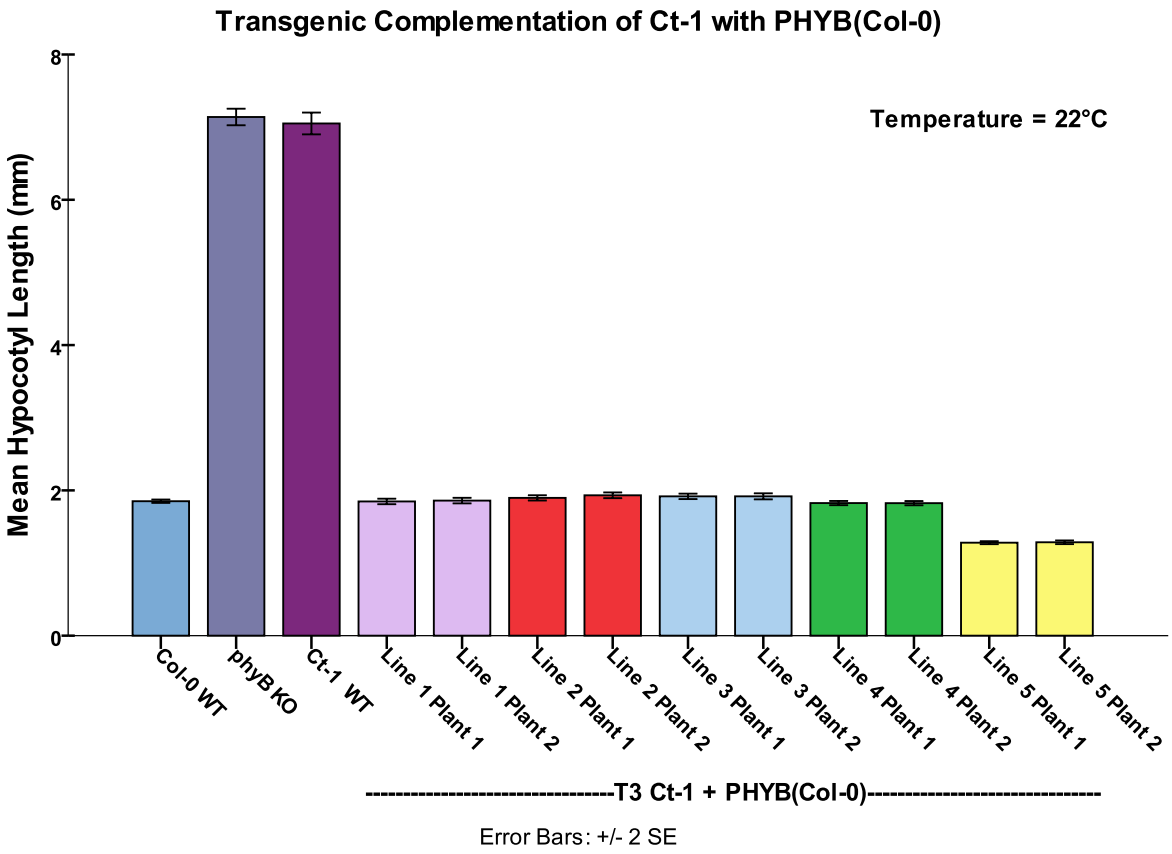

Figure 5.8: Phenotypic analysis of Col-0 and Ct-1 WT, *phyB* and Ct-1 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

Figure 5.9: Transgenic allelic complementation of T2 Ct-1 line with $PHYB_{Col\text{-}0}$ allele where the transgene is segregating and the effect can be seen on the hypocotyl length phenotype. The seedlings with long hypocotyls are the untransformed Ct-1 WT plants which died after the Basta spray.

As can be seen in figure 5.9, there was an unambiguous classification of seedlings into two discrete categories and two clear distinct hypocotyl length phenotypes were observed. The individuals with the tall hypocotyls are homozygous untransformed WT Ct-1 plants, whereas the individuals with the short hypocotyls are transformed Ct-1 plants possessing the *PHYB* allele transgene from Col-0. Ct-1 has complemented Col-0 hypocotyl length phenotype with the introduction of the transgene showing that the *PHYB* allele of the Temp22.2 QTL is responsible for the natural variation of hypocotyl length at 22°C. Since the plants were grown in a uniform environment therefore no environmental variation was observed; all the variation in the hypocotyls was genetic.

Figure 5.10: T3 homozygous Ct-1 transgenic line in which seedlings with no long hypocotyls can be seen. Also no environmental variation can be observed with uniform growth.

Another interesting phenotypic result is the comparison of the hypocotyl length of T3 Ct-1 Line 5 in figure 5.8 with the hypocotyl length of T3 *phyB* Line 2 in figure 5.4. Both of these lines carry double copy number of the transgene and have exactly the same phenotypes. This shows that the effect of the $PHYB_{Col\text{-}0}$ copy number has a similar effect on both the Ct-1 and *phyB* plants supporting further the complementation results with a single gene copy number.

The comparative phenotyping of seedlings and vegetative stages which was done in the *phyB* transgenic phenotyping (figures 5.4 and 5.7) was also carried out for the Ct-1 accession and *PHYB* transgenic plants of other accessions.

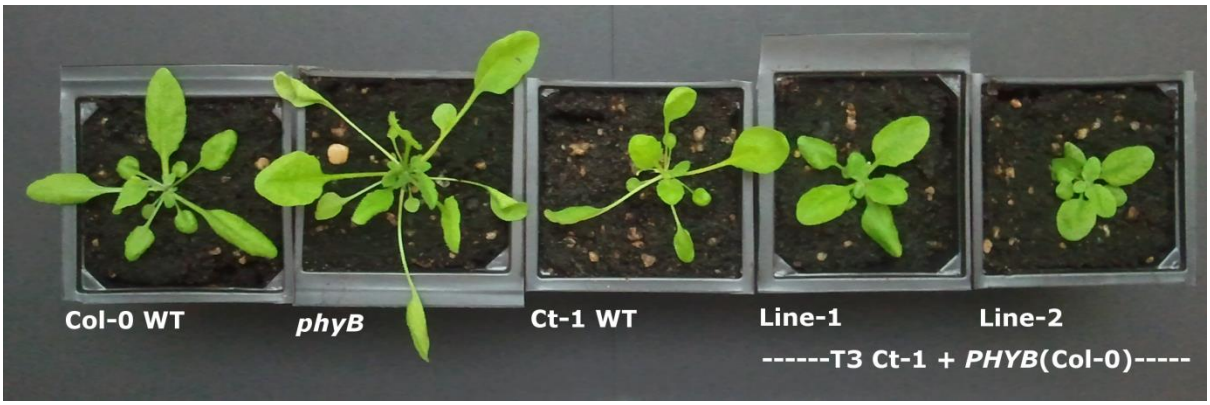

Figure 5.11: Effect of $PHYB_{Col\text{-}0}$ allele on vegetative phenotypes (petiole length and plant size) of transgenic Ct-1 plants. These phenotypes are proportionate to hypocotyls in figure 5.8.

As can be seen in figure 5.11, the vegetative stage phenotypes of the Ct-1 WT plant were similar to the *phyB* KO plant with longer petioles and bigger plant size, typical of the shade avoidance syndrome. This also suggested that *PHYB* allele of Ct-1 could be playing a role in such plant architecture. This observation is further strengthened by the phenotype of T3 Ct-1 Line-1 which carries a single copy of the $PHYB_{Col\text{-}0}$ construct and restores the phenotype to Col-0 WT phenotype. T3 Ct-1 Line 2 carries two copies of this transgene and hence the exaggerated smaller phenotype can be seen. In the case of Ct-1, the hypocotyl length phenotype is directly proportional to the petiole length in the vegetative stage of development. It possesses a relatively bigger hypocotyl and the petioles are also longer.

This comparative study showed that the hypocotyl is an excellent example and proxy of general plant growth because the effect of transgenes and their copy numbers can be observed in the phenotypes in both stages of development.

5.3A.3 Transgenic Complementation of Sf-2 accession with $PHYB_{Col\text{-}0}$ allele:

Transgenic complementation experiments were designed and performed for the validation of *PHYB* as a candidate gene in the Sf-2 accession. The aim was to test the hypothesis that allelic variation at *PHYB* candidate locus contributed to the observed

phenotypic differences between the Sf-2 and Col-0 accessions. Is the Sf-2 long hypocotyl phenotype due to mutations in Sf-2 allele of *PHYB*? If yes, then with the insertion of $PHYB_{Col\text{-}0}$ allele into Sf-2, the long hypocotyl length phenotype of Sf-2 should be transformed to the short phenotype of Col-0.

To identify single copy lines, T2 segregation data for each independently transformed Sf-2 line was analysed. To test if the observed data were significantly different from the expected frequencies, a chi-square test was performed; and the results are shown in table 5.4 in the appendix. The chi-square results (p = 0.05) show that amongst the 10 independent lines, 8 showed a close to 3:1 segregation for Basta-resistance in the T2 generation, which indicated monogenic segregation of a single locus insertion of the transgene. The remaining 2 independent lines showed a 15:1 segregation with an estimated two copies of the transgene, and were not included in the analysis. The T2 segregation analyses of all the lines in all of the subsequent experiments for other candidates were performed in the same manner and single copy lines were identified.

5.3A.3.1 Phenotypic Analysis of Homozygous T3 Sf-2 Lines:

From figure 5.12, it can be seen that Sf-2 WT has bigger hypocotyls than Col-0 WT, a difference which is in the same direction of the *phyB* KO long phenotype. Sf-2 exhibits an approximately 100% increase in hypocotyl length when compared to Col-0. To test for natural variation, eight independently T3 homozygous transformed single copy Sf-2 lines carrying the $PHYB_{Col\text{-}0}$ construct were phenotyped. $PHYB_{Col\text{-}0}$ reverted the long Sf-2 hypocotyl length phenotype to the short phenotype and these transgenic lines have complemented the Col-0 WT phenotype. For each T3 independent line, replicates of two plants were phenotyped and scored. All of the single copy Sf-2 lines 1 – 8 carrying the $PHYB_{Col\text{-}0}$ construct show similar hypocotyl phenotypes. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. The Sf-2 long hypocotyl was transformed to the short phenotype of Col-0 WT. These results also provide direct allelic evidence for natural variation. These experimental results suggest that the long hypocotyl phenotype of the Sf-2 accession is due to the mutations in the *PHYB* allele.

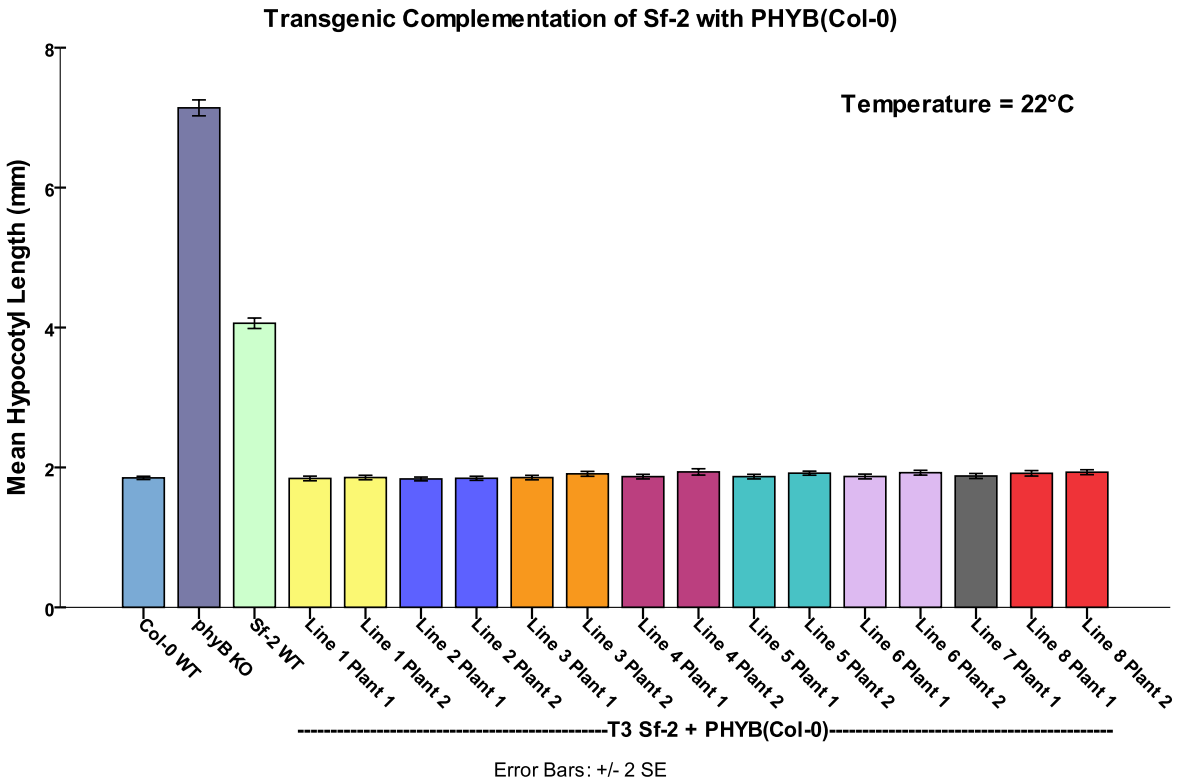

Figure 5.12: Phenotypic analysis of Col-0 and Sf-2 WT, *phyB* and Sf-2 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

Figure 5.13: Transgenic allelic complementation of T2 Sf-2 line with $PHYB_{Col\text{-}0}$ allele where the transgene is segregating and the effect can be seen on the hypocotyl length phenotype. The seedlings with long hypocotyls are the untransformed Sf-2 WT plants which died after Basta spray.

As shown in figure 5.13, the Sf-2 T2 seedlings also had two discrete categories. The individuals with the tall hypocotyls are homozygous untransformed WT Sf-2 plants which

died after the Basta spray, whereas the individuals with the short hypocotyls are transformed Sf-2 plants which have complemented Col-0 phenotype.

Figure 5.14: Effect of $PHYB_{Col\text{-}0}$ allele on vegetative phenotypes (petiole length and plant size) of transgenic Sf-2 plants. These phenotypes are proportionate to hypocotyls in figure 5.12.

As seen in figure 5.14, the vegetative stage phenotypes of the Sf-2 WT plant are similar to the *phyB* KO plant by having long petioles, bigger plant size and similar leaf shape. With the introduction of the single copy $PHYB_{Col\text{-}0}$ construct in the T3 Sf-2 line, the petiole length and plant size is significantly reduced. This elaborates the role of *PHYB* allele of Sf-2 in plant architecture. In the case of Sf-2, the hypocotyl length phenotype is directly proportional to the petiole length in the vegetative stage of development. It possesses a relatively bigger hypocotyl and the petioles are also longer.

5.3A.4 Transgenic Complementation of No-0 accession with $PHYB_{Col\text{-}0}$ allele:

For the validation of *PHYB* as a candidate gene in the No-0 accession, transgenic complementation experiments were designed and carried out. The aim was to test the hypothesis that allelic variation at *PHYB* candidate locus contributed to the observed phenotypic differences between the No-0 and Col-0 accessions. Is the No-0 long hypocotyl phenotype due to mutations in No-0 allele of *PHYB*? If yes, then with the insertion of $PHYB_{Col\text{-}0}$ allele into No-0, the long hypocotyl length phenotype of No-0 should be transformed to the short phenotype of Col-0.

5.3A.4.1 Phenotypic Analysis of Homozygous T3 No-0 Lines:

Figure 5.15 shows that No-0 WT has bigger hypocotyls than Col-0 WT, a difference which is in the same direction of the *phyB* KO long phenotype. No-0 exhibits an approximately 50% increase in hypocotyl length when compared to Col-0. To test for natural variation, five independently T3 homozygous transformed single copy No-0 lines carrying the $PHYB_{Col\text{-}0}$ construct were phenotyped. $PHYB_{Col\text{-}0}$ transformed the long No-0 phenotype to the short phenotype and these transgenic lines have complemented the Col-0 WT phenotype. The phenotypic results can be seen in the single copy No-0 lines 1 –5 carrying the $PHYB_{Col\text{-}0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results provide direct allelic evidence for natural variation in another major effect parental accession. These experimental results suggest that the long hypocotyl phenotype of No-0 is due to the mutations in the *PHYB* allele.

As seen in figure 5.16, the vegetative stage phenotype of the No-0 WT plant is not similar to the Col-0 WT plant. The No-0 plant has bigger and broader leaves. With the introduction of the single copy $PHYB_{Col\text{-}0}$ construct in the T3 No-0 lines 1-5, the petiole length and plant size is significantly reduced showing the effect of $PHYB_{Col\text{-}0}$ in its plant architecture. In the case of No-0, the hypocotyl length phenotype is directly proportional to the petiole length in the vegetative stage of development. It possesses a relatively smaller hypocotyl and the petioles are also shorter.

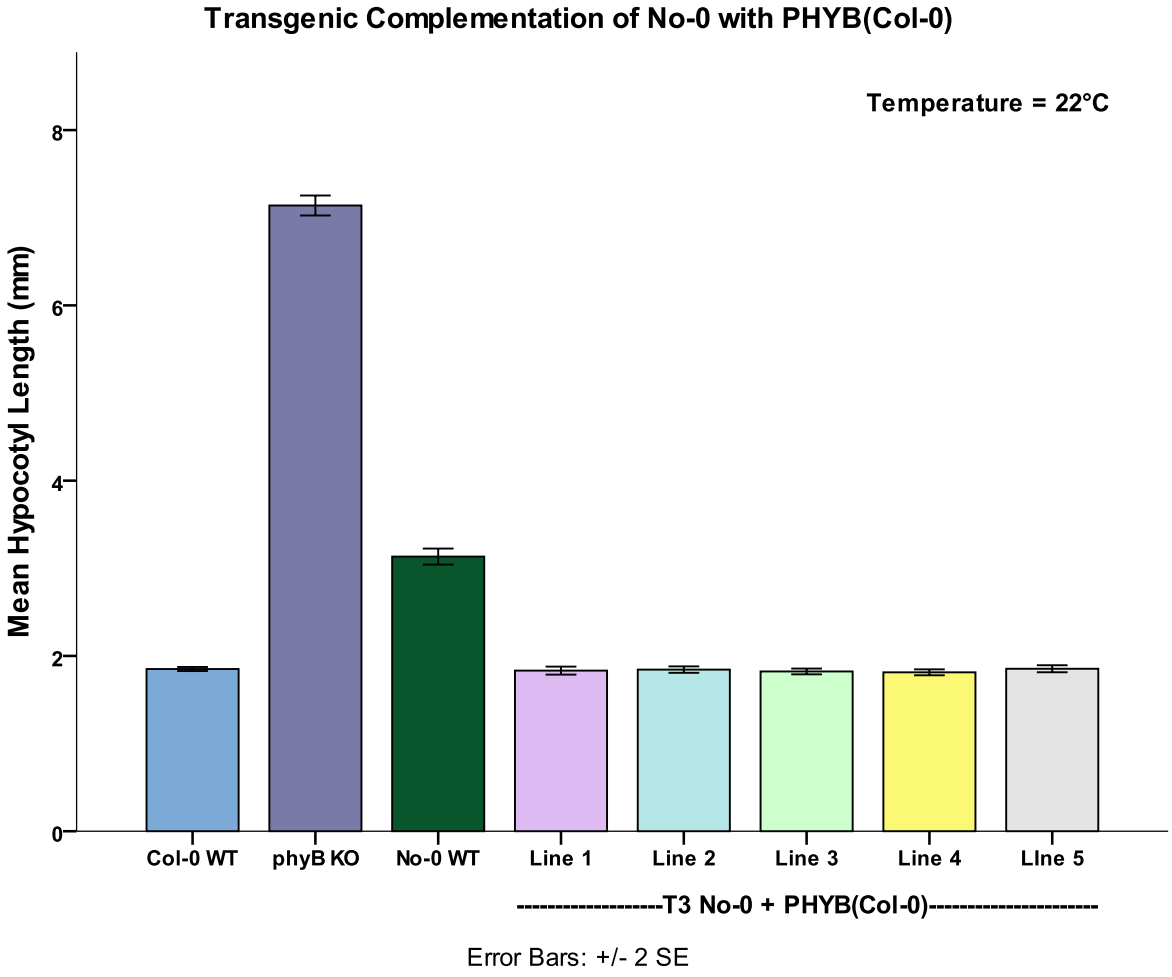

Figure 5.15: Phenotypic analysis of Col-0 and No-0 WT, *phyB* and No-0 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

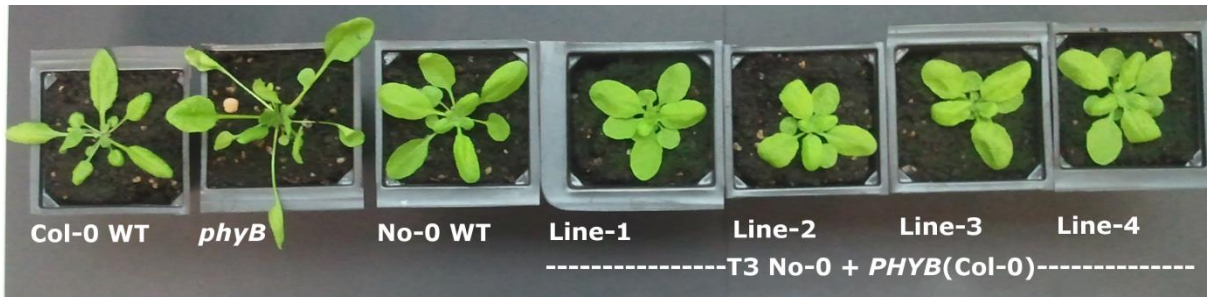

Figure 5.16: Effect of $PHYB_{Col\text{-}0}$ allele on vegetative phenotypes (petiole length and plant size) of transgenic No-0 plants. These phenotypes are proportionate to hypocotyls in figure 5.15.

5.3A.5 Transgenic Complementation of Zu-0 accession with *PHYB*$_{Col-0}$ allele:

For the validation of *PHYB* as a candidate gene in the Zu-0 accession, transgenic complementation experiments were designed and carried out. The aim was to test the hypothesis that allelic variation at *PHYB* candidate locus contributed to the observed phenotypic differences between the Zu-0 and Col-0 accessions. Is the Zu-0 long hypocotyl phenotype due to mutations in Zu-0 allele of *PHYB*? If yes, then with the insertion of *PHYB*$_{Col-0}$ allele into Zu-0, the long hypocotyl length phenotype of Zu-0 should be transformed to the short phenotype of Col-0.

5.3A.5.1 Phenotypic Analysis of Homozygous T3 Zu-0 Lines:

Figure 5.17 reveals that Zu-0 WT has bigger hypocotyls than Col-0 WT, a difference which is in the same direction of the *phyB* KO long phenotype. Zu-0 exhibits an approximately 25% increase in hypocotyl length when compared to Col-0. To test for natural variation, two independently T3 homozygous transformed single copy Zu-0 lines carrying the *PHYB*$_{Col-0}$ construct were phenotyped. *PHYB*$_{Col-0}$ transformed the long Zu-0 phenotype to the short phenotype and these transgenic lines have complemented the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Zu-0 lines 1 –2 carrying the *PHYB*$_{Col-0}$ construct. The phenotypes of these two independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results provide direct allelic evidence for natural variation in yet another major effect parental accession. These experimental results suggest that the long hypocotyl phenotype of Zu-0 is due to the mutations in the *PHYB* allele.

The vegetative stage phenotype of the Zu-0 WT plant is not similar to the Col-0 WT plant and has longer petioles as seen in figure 5.18. With the introduction of the *PHYB*$_{Col-0}$ construct in the T3 Zu-0 lines 1-2, the petiole length and plant size is significantly reduced showing the effect of *PHYB*$_{Col-0}$ in its plant architecture. In the case of Zu-0, the hypocotyl length phenotype is inversely proportional to the petiole length in the vegetative stage of development. It possesses a relatively smaller hypocotyl but the petioles are longer.

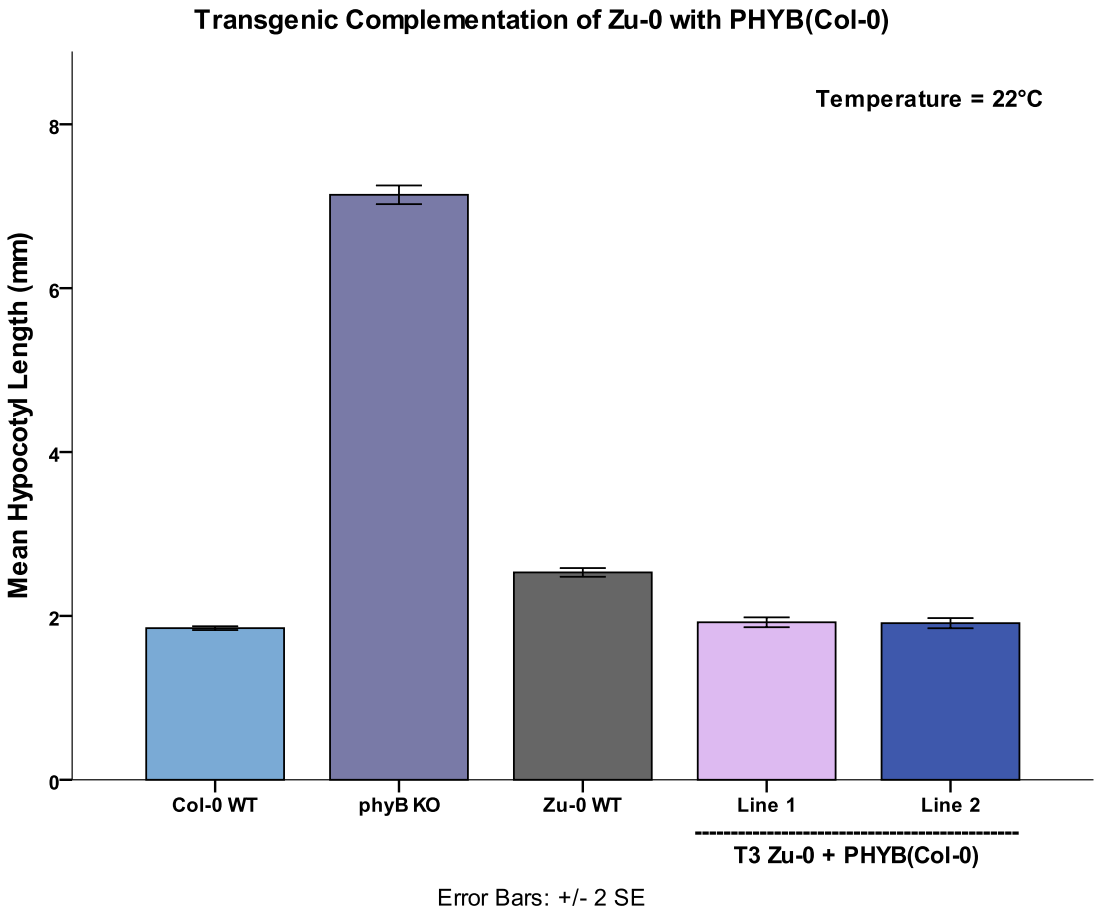

Figure 5.17: Phenotypic analysis of Col-0 and Zu-0 WT, *phyB* and Zu-0 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

Figure 5.18: Effect of $PHYB_{Col-0}$ allele on vegetative phenotypes (petiole length and plant size) of transgenic No-0 plants. These phenotypes are proportionate to hypocotyls in figure 5.17.

5.3B Results: Transgenic Complementation with *CYP71*$_{Col-0}$ allele:

The results in this section describe the analysis of functional variation of *CYP71* alleles from 'Chr3.QTL3' QTL in hypocotyl length phenotype in response to temperature at 22°C in multiple Arabidopsis accessions. Multiple major effect parental accessions have been experimentally tested for complementation with the *CYP71*$_{Col-0}$ allele at 22°C. The results of these experiments have been described in detail in individual sub sections for each accession separately which provide the definitive experimental evidence for the *CYP71* QTL.

5.3B.1 Transgenic Complementation of Ct-1 accession with *CYP71*$_{Col-0}$ allele:

To confirm *CYP71* as a candidate gene in Ct-1, transgenic complementation experiments were carried out. The aim was to test the hypothesis that allelic variation at *CYP71* candidate locus contributed to the observed phenotypic differences between the Ct-1 and Col-0 accessions. Is the Ct-1 long hypocotyl phenotype due to mutations in Ct-1 allele of *CYP71*? If yes, then with the insertion of *CYP71*$_{Col-0}$ allele into Ct-1, the long hypocotyl length phenotype of Ct-1 should be transformed to the short phenotype of Col-0.

5.3B.1.1 Phenotypic Analysis of Homozygous T3 Ct-1 Lines:

Figure 5.19 shows that Ct-1 WT has bigger hypocotyls than Col-0 WT, a difference which is in the same direction of the *cyp71* KO long phenotype. To test for natural variation, five independently T3 homozygous transformed single copy Ct-1 lines carrying the *CYP71*$_{Col-0}$ construct were phenotyped. The results reveal that *CYP71*$_{Col-0}$ did not transform the long Ct-1 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Ct-1 lines 1 –5 carrying the *CYP71*$_{Col-0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results did not provide direct allelic evidence for natural variation between Ct-1 and Col-0 parental accessions and suggest that the long hypocotyl phenotype of Ct-1 is not due to the mutations in the *CYP71* allele.

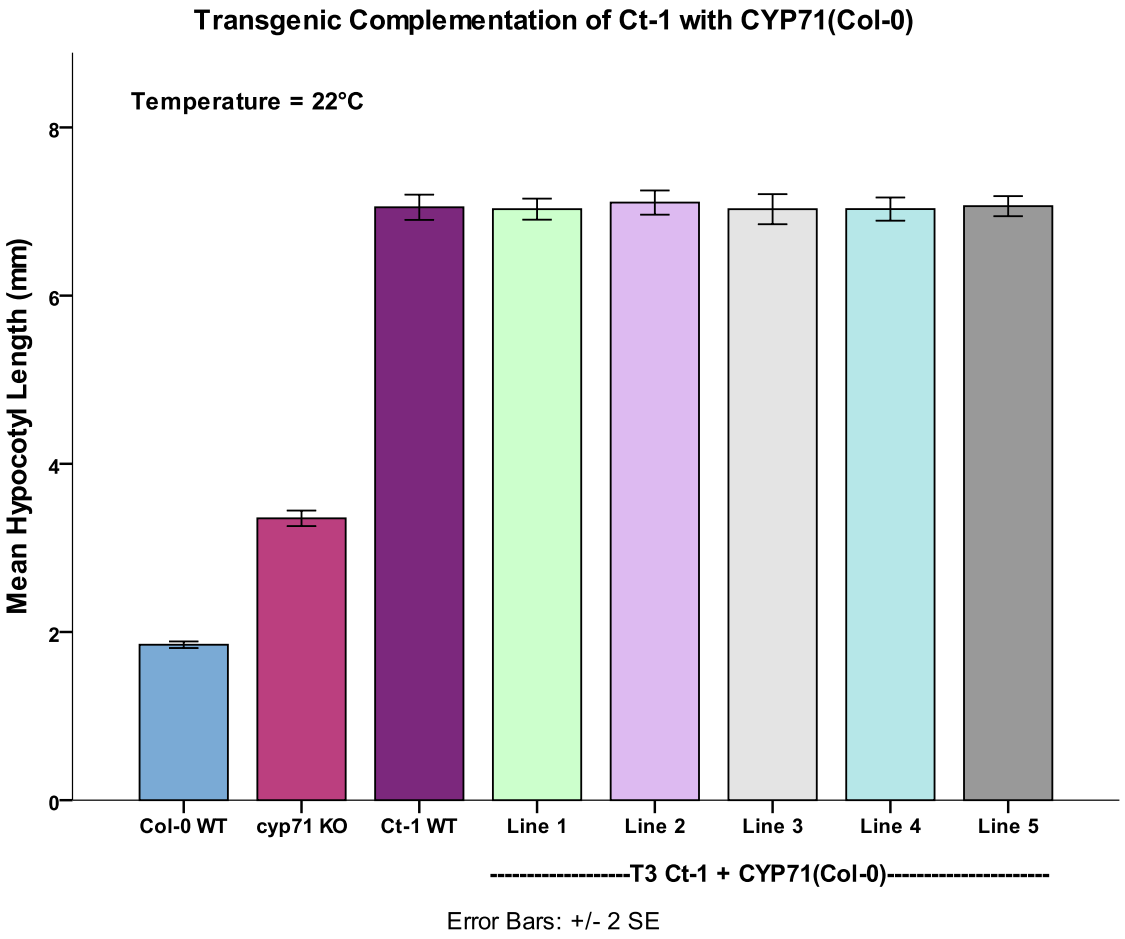

Figure 5.19: Phenotypic analysis of Col-0 and Ct-1 WT, *cyp71* and Ct-1 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

Figure 5.20: T3 Ct-1 line carrying the $CYP71_{Col\text{-}0}$ allele showing no change in the long hypocotyl length phenotype. These transgenic lines did not complement the Col-0 WT phenotype.

The T2 generation individuals displayed only one uniform phenotype of long hypocotyls and the segregation effect of *CYP71*$_{Col-0}$ allele could not be physically seen in the phenotype which suggested that the *CYP71*$_{Col-0}$ allele was not playing a role in the natural variation of the trait. This was confirmed in the phenotyping of the T3 Ct-1homozygous lines where the hypocotyl phenotype was unaffected by the *CYP71*$_{Col-0}$ allele, as can be seen in figure 5.20.

5.3B.2 Transgenic Complementation of Sf-2 accession with *CYP71*$_{Col-0}$ allele:

For the confirmation of *CYP71* as a candidate gene in Sf-2, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *CYP71* candidate locus contributed to the observed phenotypic differences between the Sf-2 and Col-0 accessions. Is the Sf-2 long hypocotyl phenotype due to mutations in Sf-2 allele of *CYP71*? If yes, then with the insertion of *CYP71*$_{Col-0}$ allele into Sf-2, the long hypocotyl phenotype of Sf-2 should be transformed to the short phenotype of Col-0.

5.3B.2.1 Phenotypic Analysis of Homozygous T3 Sf-2 Lines:

Figure 5.21 shows that Sf-2 WT has bigger hypocotyls than Col-0 WT, a difference which is in the same direction of the *cyp71* KO long phenotype. To test for natural variation, four independently T3 homozygous transformed single copy Sf-2 lines carrying the *CYP71*$_{Col-0}$ construct were phenotyped. The results reveal that *CYP71*$_{Col-0}$ did not transform the long Sf-2 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Sf-2 lines 1–4 carrying the *CYP71*$_{Col-0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results did not provide direct allelic evidence for natural variation between Sf-2 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Sf-2 is not due to the mutations in the *CYP71* allele.

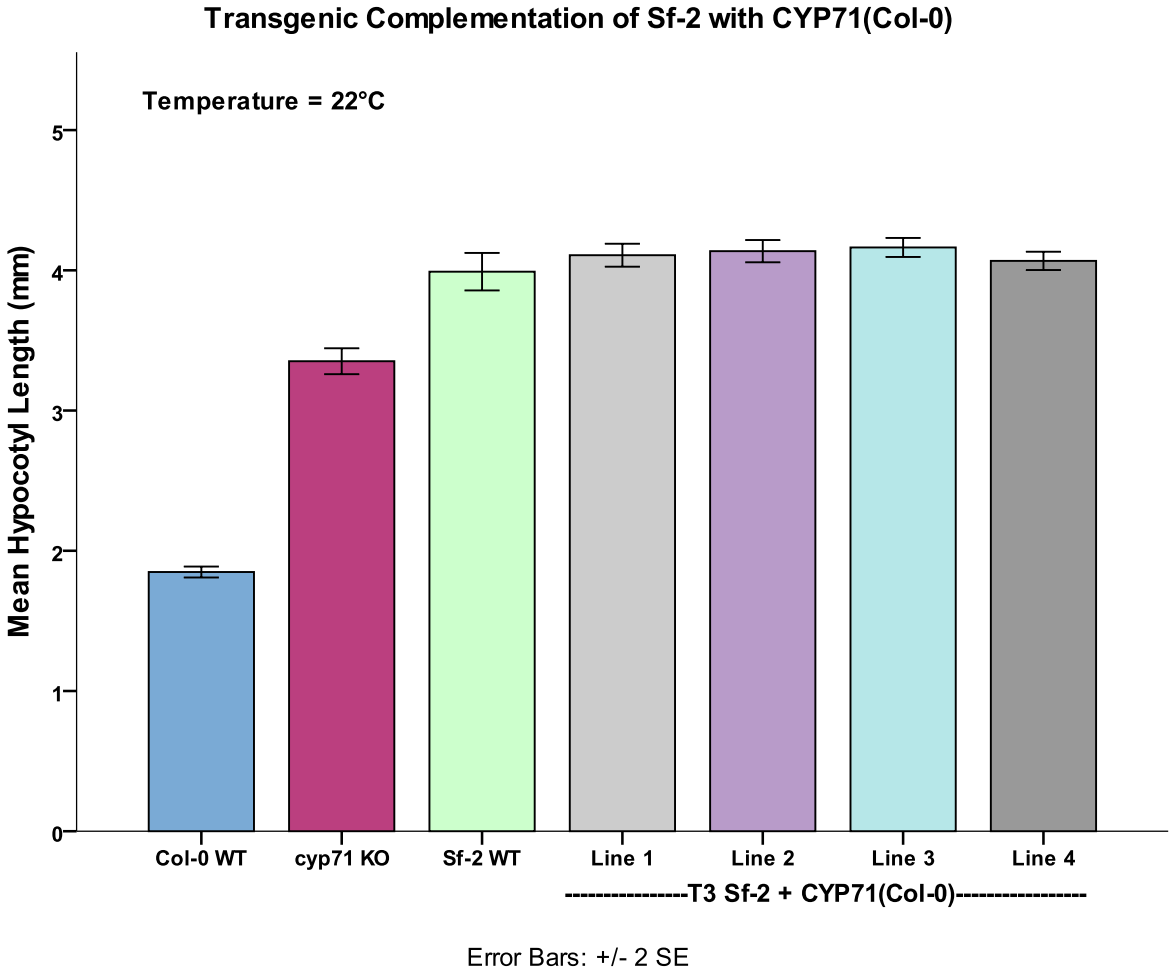

Figure 5.21: Phenotypic analysis of Col-0 and Sf-2 WT, *cyp71* and Sf-2 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C

Figure 5.22: T3 Sf-2 line carrying the $CYP71_{Col\text{-}0}$ allele showing no change in the long hypocotyl length phenotype. These transgenic lines did not complement the Col-0 WT phenotype.

5.3B.3 Transgenic Complementation of Zu-0 accession with *CYP71*$_{Col-0}$ allele:

To verify *CYP71* as a candidate gene in Zu-0, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *CYP71* candidate locus contributed to the observed phenotypic differences between the Zu-0 and Col-0 accessions. Is the Zu-0 long hypocotyl phenotype due to mutations in Zu-0 allele of *CYP71*? If yes, then with the insertion of *CYP71*$_{Col-0}$ allele into Zu-0, the long hypocotyl length phenotype of Zu-0 should be transformed to the short phenotype of Col-0.

5.3B.3.1 Phenotypic Analysis of Homozygous T3 Zu-0 Lines:

It can be observed in figure 5.23 that Zu-0 WT has bigger hypocotyls than Col-0 WT, a difference which is in the same direction of the *cyp71* KO long phenotype. To test for natural variation, three independently T3 homozygous transformed single copy Zu-0 lines carrying the *CYP71*$_{Col-0}$ construct were phenotyped. The results reveal that *CYP71*$_{Col-0}$ did not transform the long Zu-0 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Zu-0 lines 1–3 carrying the *CYP71*$_{Col-0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results did not provide direct allelic evidence for natural variation between Zu-0 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Zu-0 is not due to the mutations in the *CYP71* allele.

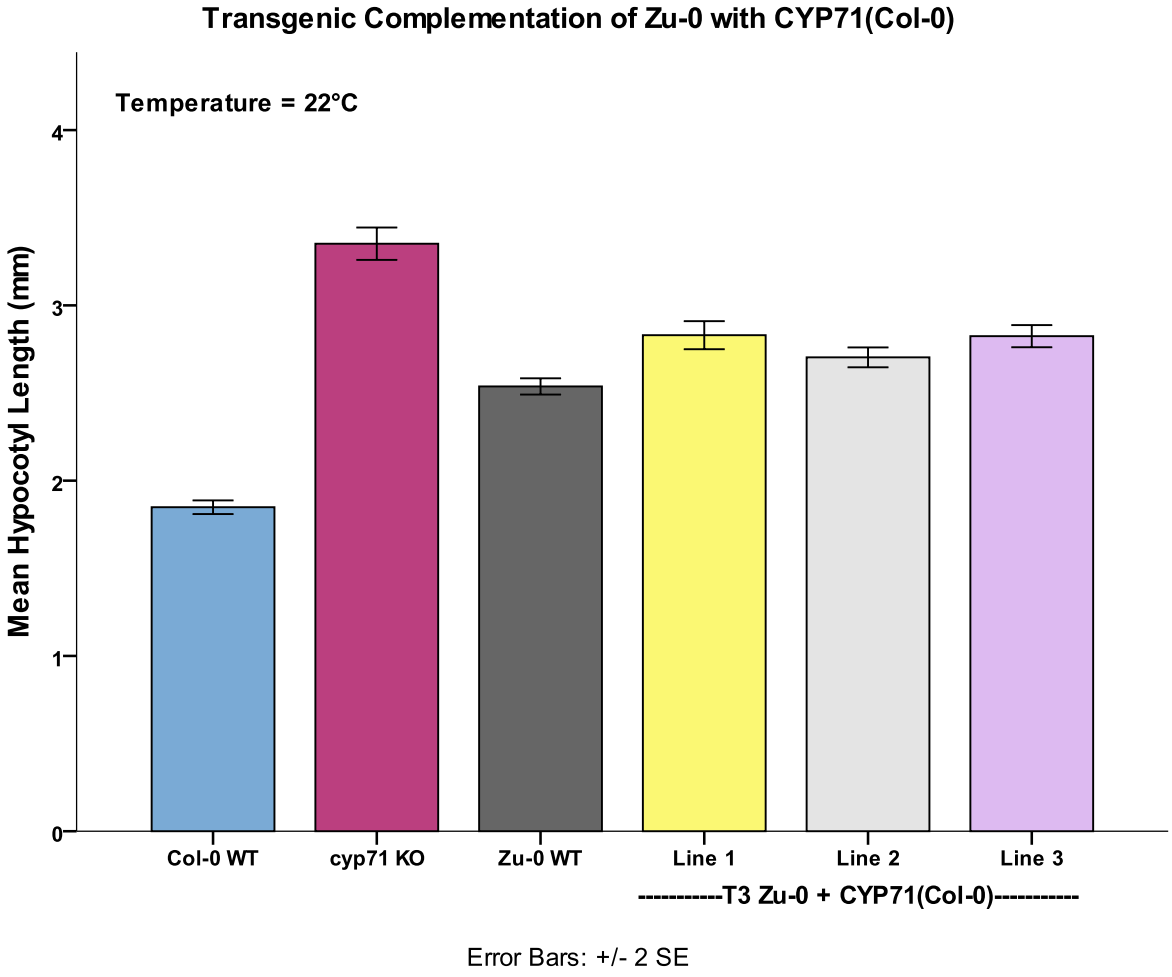

Figure 5.23: Phenotypic analysis of Col-0 and Zu-0 WT, *cyp71* and Zu-0 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

5.3C Results: Transgenic Complementation with *SAUR38*$_{Col-0}$ allele:

The results in this section describe the analysis of functional variation of *SAUR38* alleles from 'Chr2.QTL3' QTL in hypocotyl length phenotype in response to temperature at 22°C in multiple Arabidopsis accessions. Multiple major effect parental accessions have been experimentally tested for complementation with the *SAUR38*$_{Col-0}$ allele at 22°C. The results of these experiments have been described in detail in individual sub sections for each accession separately. These results provide the definitive experimental evidence for the *SAUR38* QTL alleles by individual transgenic allelic complementation studies.

5.3C.1 Transgenic Complementation of Tsu-0 accession with *SAUR38*$_{Col-0}$ allele:

In order to functionally test *SAUR38* as a candidate gene in Tsu-0, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *SAUR38* candidate locus contributed to the observed phenotypic differences between the Tsu-0 and Col-0 accessions. Is the Tsu-0 long hypocotyl phenotype due to mutations in Tsu-0 allele of *SAUR38*? If yes, then with the insertion of *SAUR38*$_{Col-0}$ allele into Tsu-0, the long hypocotyl length phenotype of Tsu-0 should be transformed to the short phenotype of Col-0.

5.3C.1.1 Phenotypic Analysis of Homozygous T3 Tsu-0 Lines:

Figure 5.24 shows that Tsu-0 WT has bigger hypocotyls than Col-0 WT, a difference which is similar to and is in the same direction of the *saur38* KO long phenotype. To test for natural variation, three independently T3 homozygous transformed single copy Tsu-0 lines carrying the *SAUR38*$_{Col-0}$ construct were phenotyped. The results reveal that *SAUR38*$_{Col-0}$ did not transform the long Tsu-0 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Tsu-0 lines 1–3 carrying the *SAUR38*$_{Col-0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of

variability in the hypocotyl length phenotype. These results did not provide direct allelic evidence for natural variation between Tsu-0 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Tsu-0 is not due to the mutations in the *SAUR38* allele.

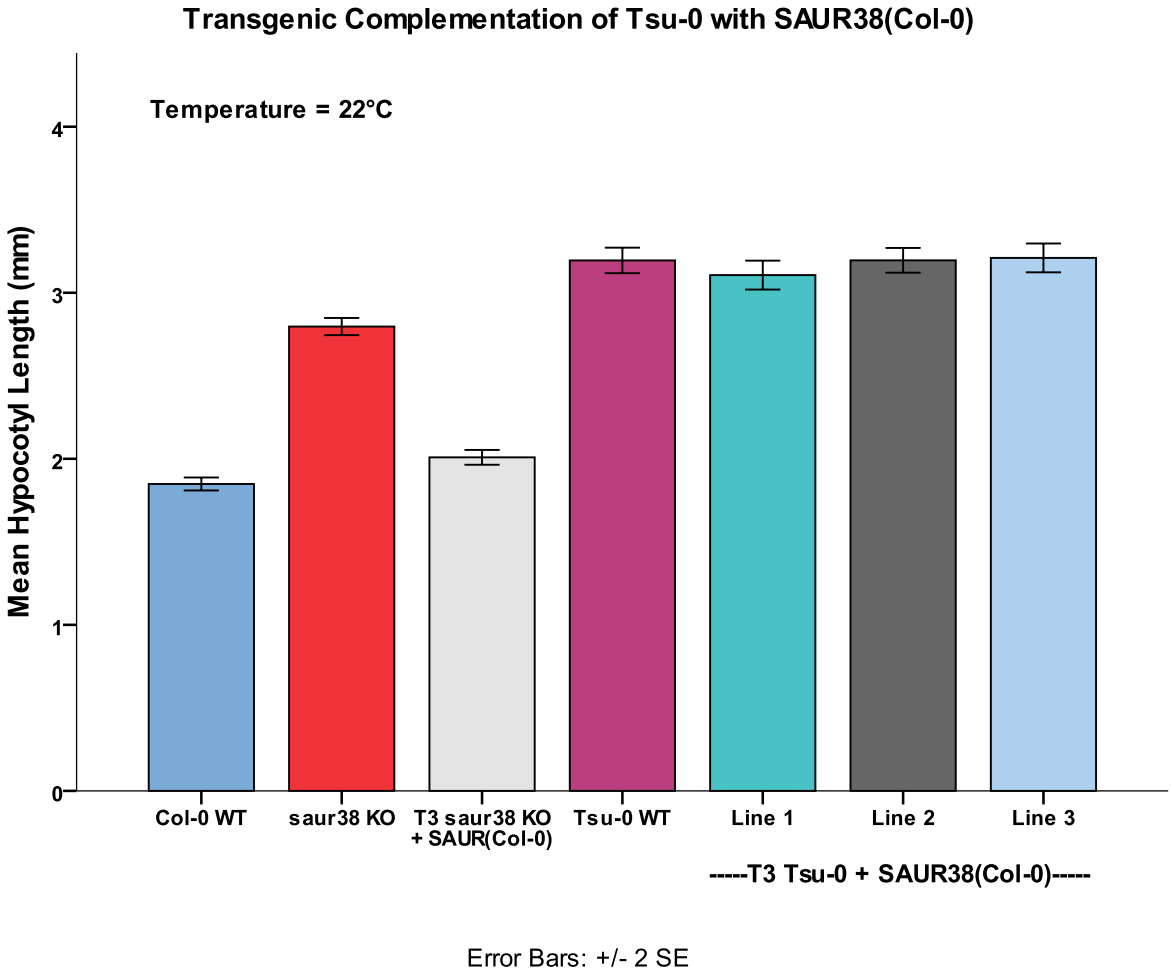

Figure 5.24: Phenotypic analysis of Col-0 and Tsu-0 WT, *saur38* and Zu-0 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

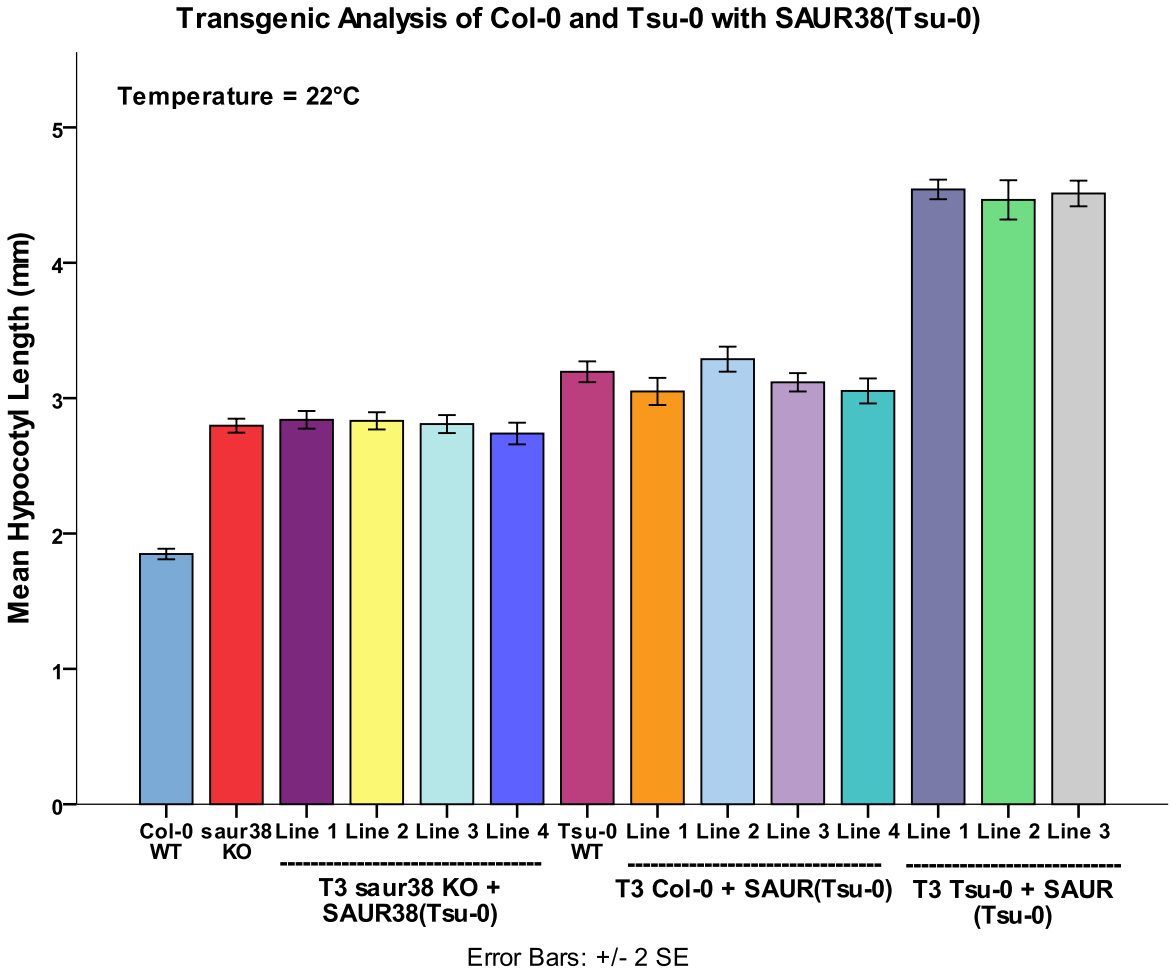

Figure 5.25: Phenotypic analysis of Col-0 and Tsu-0 WT, *saur38*, Col-0 and Tsu-0 transgenic lines. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

Since a novel function has been identified for the *SAUR38* gene, a transgenic analysis (figure 5.25) of the $SAUR38_{Tsu\text{-}0}$ allele was performed to further elucidate the role of this candidate gene in the natural variation of hypocotyl length architecture. The role of the $SAUR38_{Col\text{-}0}$ allele has been established by knockout analysis in Chapter 4 and transgenic analysis in this Chapter. The results respectively reveal that *SAUR38* is functionally involved in hypocotyl length**.** Furthermore, the transgenic Col-0 carrying the $SAUR38_{Tsu\text{-}0}$ allele plants had increased the hypocotyl length and these plants had ~ 90% bigger hypocotyls than Col-0 WT. The gene copy number effects can be seen in transgenic Tsu-0 lines carrying the $SAUR38_{Tsu\text{-}0}$ allele. Due to the presence of an extra copy of the

SAUR38$_{Tsu-0}$ allele, these plants showed an exaggerated phenotype and hypocotyl lengths were ~ 45% bigger than Tsu-0 WT plants.

These are very interesting results. It suggests that the *SAUR38*$_{Tsu-0}$ allele is dominant, and able to promote hypocotyl elongation, since it increases the length of hypocotyls in both Col-0 and even Tsu-0 (which is already long!). By contrast the *SAUR38*$_{Col-0}$ allele does not increase hypocotyl elongation; in the knockout it even rescues the longer hypocotyl. These results are compatible with the *SAUR38*$_{Tsu-0}$ allele being a causative agent in the Tsu-0 phenotype but acting in a dominant way, so that it cannot be rescued by the *SAUR38*$_{Col-0}$ allele (figure 5.24). Overall, these results shed light on the role of *SAUR38* alleles that contribute to the natural variation of hypocotyl length.

5.3C.2 Transgenic Complementation of Wil-2 accession with *SAUR38*$_{Col-0}$ allele:

In order to functionally test *SAUR38* as a candidate gene in Wil-2, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *SAUR38* candidate locus contributed to the observed phenotypic differences between the Wil-2 and Col-0 accessions. Is the Wil-2 long hypocotyl phenotype due to mutations in Wil-2 allele of *SAUR38*? If yes, then with the insertion of *SAUR38*$_{Col-0}$ allele into Wil-2, the long hypocotyl length phenotype of Wil-2 should be transformed to the short phenotype of Col-0.

5.3C.2.1 Phenotypic Analysis of Homozygous T3 Wil-2 Lines:

Figure 5.26 shows that Wil-2 WT has bigger hypocotyls than Col-0 WT, which is exactly a similar phenotype of the *saur38* KO long phenotype. To test for natural variation, four independently T3 homozygous transformed single copy Wil-2 lines carrying the *SAUR38*$_{Col-0}$ construct were phenotyped. The results reveal that *SAUR38*$_{Col-0}$ did not transform the long Wil-2 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Wil-2 lines 1–4 carrying the *SAUR38*$_{Col-0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length

phenotype. These results did not provide direct allelic evidence for natural variation between Wil-2 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Wil-2 is not due to the mutations in the *SAUR38* allele.

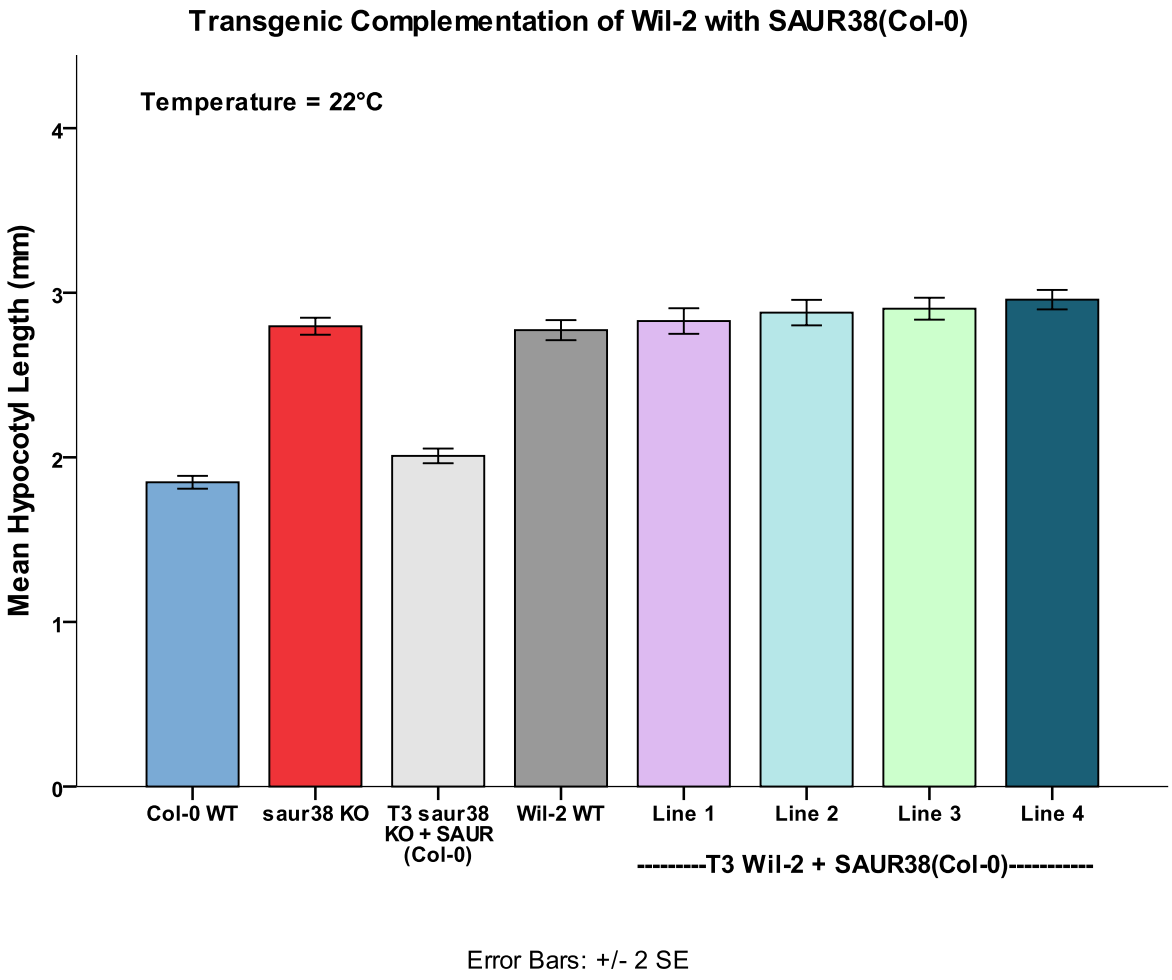

Figure 5.26: Phenotypic analysis of Col-0 and Wil-2 WT, *saur38* and Wil-2 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 22°C.

5.3D Results: Transgenic Complementation with *NPY2*$_{Col-0}$ allele:

The results in this section describe the analysis of functional variation of *NPY2* alleles from 'Chr2.QTL1' QTL in hypocotyl length phenotype in response to temperature at 22°C in multiple Arabidopsis accessions. Multiple major effect parental accessions have been experimentally tested for complementation with the *NPY2*$_{Col-0}$ allele at 22°C. The results of these experiments have been described in individual sub sections for each accession separately. These results provide the definitive experimental evidence for the *NPY2* QTL alleles by individual transgenic allelic complementation studies.

5.3D.1 Transgenic Complementation of Ct-1 accession with *NPY2*$_{Col-0}$ allele:

In order to functionally test *NPY2* as a candidate gene in Ct-1, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *NPY2* candidate locus contributed to the observed phenotypic differences between the Ct-1 and Col-0 accessions. Is the Ct-1 long hypocotyl phenotype due to mutations in Ct-1 allele of *NPY2*? If yes, then with the insertion of *NPY2*$_{Col-0}$ allele into Ct-1, the long hypocotyl length phenotype of Ct-1 should be transformed to the short phenotype of Col-0.

5.3D.1.1 Phenotypic Analysis of T2 Ct-1 Lines:

To test for natural variation, four independently T2 single copy Ct-1 lines carrying the *NPY2*$_{Col-0}$ construct were phenotyped. The results reveal that *NPY2*$_{Col-0}$ did not transform the long Ct-1 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. In figure 5.27, the phenotype of the single copy Ct-1 line carrying the *NPY2*$_{Col-0}$ construct can be seen. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results did not provide direct allelic evidence for natural variation between Ct-1 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Ct-1 is not due to the mutations in the *NPY2* allele.

Figure 5.27: T2 Ct-1 line carrying the *NPY2*$_{Col-0}$ allele showing no change in the long hypocotyl length phenotype. These transgenic lines did not complement the Col-0 WT phenotype.

5.3D.2 Transgenic Complementation of Sf-2 accession with *NPY2*$_{Col-0}$ allele:

In order to functionally test *NPY2* as a candidate gene in Sf-2, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *NPY2* candidate locus contributed to the observed phenotypic differences between the Sf-2 and Col-0 accessions. Is the Sf-2 long hypocotyl phenotype due to mutations in Sf-2 allele of *NPY2*? If yes, then with the insertion of *NPY2*$_{Col-0}$ allele into Sf-2, the long hypocotyl length phenotype of Sf-2 should be transformed to the short phenotype of Col-0.

5.3D.2.1 Phenotypic Analysis of T2 Sf-2 Lines:

To test for natural variation, four independently T2 single copy Sf-2 lines carrying the *NPY2*$_{Col-0}$ construct were phenotyped. The results reveal that *NPY2*$_{Col-0}$ did not transform the long Sf-2 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. In figure 5.28, the phenotype of the single copy Sf-2 line carrying the *NPY2*$_{Col-0}$ construct can be seen. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length

phenotype. These results did not provide direct allelic evidence for natural variation between Sf-2 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Sf-2 is not due to the mutations in the *NPY2* allele.

Figure 5.28: T2 Sf-2 line carrying the $NPY2_{Col\text{-}0}$ allele showing no change in the long hypocotyl length phenotype. These transgenic lines did not complement the Col-0 WT phenotype.

5.3E Results: Transgenic Complementation with *CYCT1;4*$_{Col-0}$ allele:

The results in this section describe the analysis of functional variation of *CYCT1;4* alleles from 'Chr4.QTL1' QTL in hypocotyl length phenotype in response to temperature at 27°C in multiple Arabidopsis accessions. A major effect parental accession Zu-0 has been experimentally tested for complementation with the *CYCT1;4*$_{Col-0}$ allele at 27°C. These results provide the definitive experimental evidence for the *CYCT1;4* QTL alleles by individual transgenic allelic complementation studies.

5.3E.1 Transgenic Complementation of Zu-0 accession with *CYCT1;4*$_{Col-0}$ allele:

In order to functionally test *CYCT1;4* as a candidate gene in Zu-0, transgenic complementation experiments were performed. The aim was to test the hypothesis that allelic variation at *CYCT1;4* candidate locus contributed to the observed phenotypic differences between the Zu-0 and Col-0 accessions. Is the Zu-0 long hypocotyl phenotype due to mutations in Zu-0 allele of *CYCT1;4*? If yes, then with the insertion of *CYCT1;4*$_{Col-0}$ allele into Zu-0, the long hypocotyl length phenotype of Zu-0 should be transformed to the short phenotype of Col-0.

5.3E.1.1 Phenotypic Analysis of Homozygous T3 Zu-0 Lines:

Figure 5.29 shows that Zu-0 WT has bigger hypocotyls than Col-0 WT, a difference in the same direction of the *cyct1;4* KO long phenotype. To test for natural variation, four independently T3 homozygous transformed single copy Zu-0 lines carrying the *CYCT1;4*$_{Col-0}$ construct were phenotyped. The results reveal that *CYCT1;4*$_{Col-0}$ did not transform the long Zu-0 phenotype to the short phenotype and these transgenic lines did not complement the Col-0 WT phenotype. The phenotypic results can be seen in the single copy Zu-0 lines 1–4 carrying the *CYCT1;4*$_{Col-0}$ construct. The phenotypes of the multiple independently transformed lines revealed no degree of variability in the hypocotyl length phenotype. These results did not provide direct allelic evidence for natural variation between Zu-0 and Col-0 parental accessions. These experimental results suggest that the long hypocotyl phenotype of Zu-0 is not due to the mutations in the *CYCT1;4* allele.

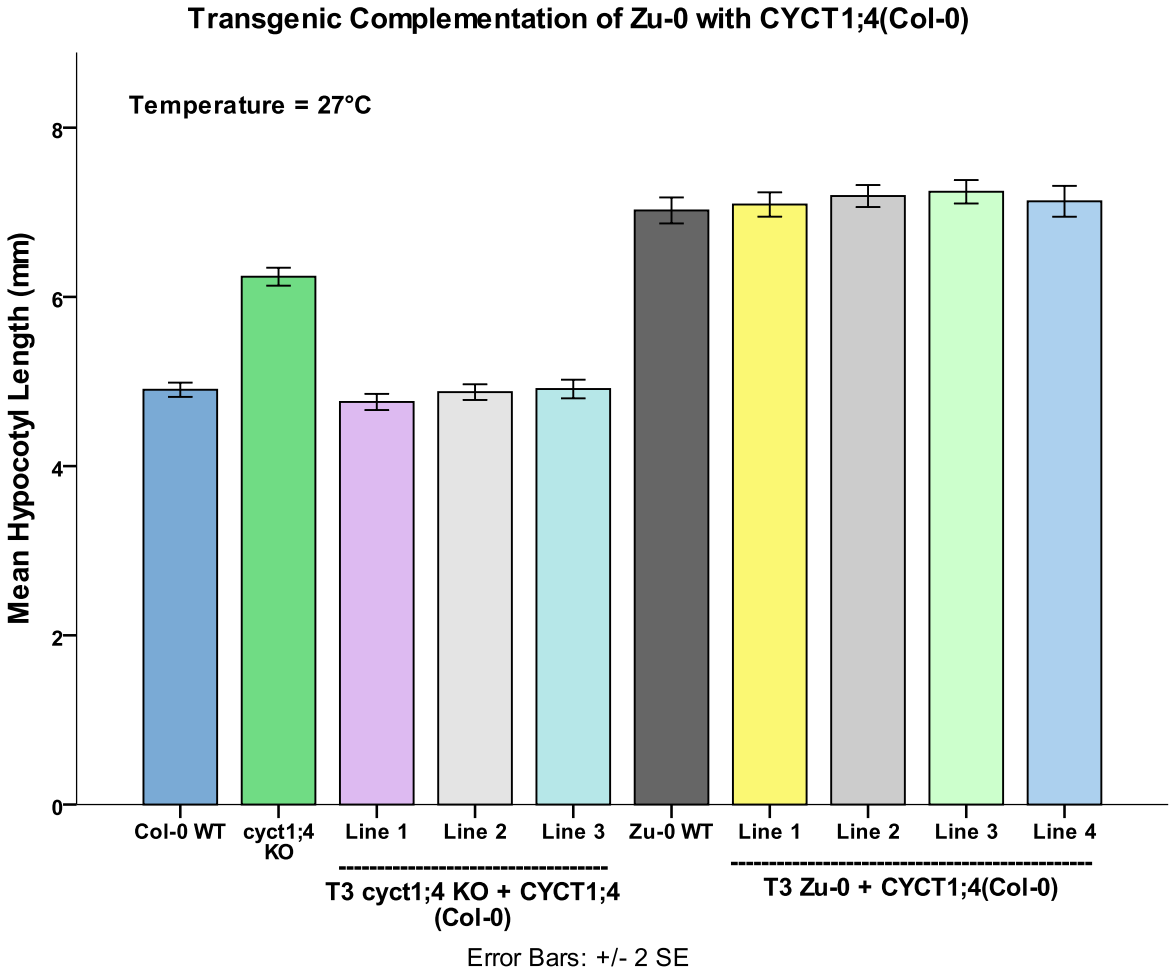

Figure 5.29: Phenotypic analysis of Col-0 and Zu-0 WT, *cyct1;4* and Zu-0 transgenic lines showing complementation results. The different line numbers refer to independent transformants. 1 week old hypocotyls were grown in 170 µE/m2/sec short days at 27°C.

5.4 Discussion:

5.4.1 Discussion on Methods:

This study has conducted QTL cloning using transgenic plants, which is a valuable and powerful tool to help us understand the function of alleles of QTL candidate genes. The use of recombinant DNA technology in the production of transgenic plants for the purpose of studying gene function and expression has been widespread. Transgenic research on the model organism Arabidopsis has helped us to understand how naturally occurring genetic variation in the alleles of various QTL candidate genes control hypocotyl length function at the molecular level at specific temperature responses.

To establish definitively that the natural mutations in the various alleles cause the hypocotyl phenotype, one must complement the mutation by introducing a WT copy of the allele into the mutant accessions for restoration of WT phenotype by using transgenesis. In this regard, one of the benefits of using transgenic complementation, as a direct method of testing for the alleles of the candidate genes, is that it conclusively demonstrates the alleles causing the natural variation in hypocotyl length. What are the consequences of the genetic changes that we have made? The functional analysis allows the expression of transgenes to manipulate the underlying biological processes involved in hypocotyl elongation. This approach allows the activity of an allele to be functionally investigated in the corresponding accession. The phenotype of hypocotyl length can be analysed in the presence of the specific allele which may suggest its functional role.

To generate transgenic Arabidopsis plants, Agrobacterium-mediated transformation was used, which transferred the engineered T-DNA carrying the foreign gene into the genomes of the various accessions. The use of insertional T-DNA provides a swift way of transferring the gene of interest in to various accessions for functional testing. This is the most widely used and successful method to generate transgenic plants in Arabidopsis (Bechtold *et al*., 1993; Clough and Bent, 1998). It has been exploited as an effective tool for transgenic technology.

This method has a number of strengths. To understand gene function, it is essential to stably insert candidate genes in plants. A very useful feature of Arabidopsis is that it can be genetically transformed by the simple method of direct floral dipping. This

transformation method of floral dipping is very effective in generating transgenics. Benefits of this method are that no tissue culture is required which saves a lot of time and it has high transformation efficiency as demonstrated in this study. This method also avoids any unwanted genetic changes such as induction of point mutations which could arise during tissue culture procedures. Therefore, Agrobacterium-mediated transformation has been the method of choice for transforming plants in this study. From a QTL cloning point of view, the capability of Agrobacterium to stably insert candidate genes into various accessions is essential to elucidate and verify their functions.

There are also some weaknesses observed in this transformation method. *Firstly*, the construct is randomly inserted into the host genome which could cause differential expressions of the candidate genes depending on the insertion location. However, this weakness is overcome by using several independently transformed lines which account for any differential expression, if present. As a result, several independent transgenic lines have been analysed to confirm the effect of the transgene.

Secondly, more than one copy of the construct can be inserted into the host genome. Ideally a single copy of the transgene is functionally informative in the QTL cloning process. This problem is overcome by identifying single copy inserts in the various independently transformed lines by traditional genetic analysis to look for monogenic segregation of the transgene in T2 generation. It helps in the identification of homozygous and hemizygous lines. For example, in this study in the *PHYB* candidate gene experiment of Ct-1 transgenic phenotyping, Line 5 (Plants 1 and 2) is a control experiment showing that the copy number of the transgene has an effect on the hypocotyl length. This particular line has a 15:1 segregation ratio due to double copy number whereas all the others have 3:1 ratio and hence the same phenotype as Col-0 WT. The phenotype of transgenic Ct-1 Line 5 has the same phenotype (smaller) as transgenic Col-0 having an extra copy of the gene, indicating the exaggeration of the hypocotyl length phenotype.

Moreover, in the case of transgenic Col-0, a control experiment using a line with double copy transgene number was phenotyped; which showed an exaggerated smaller phenotype. This is in congruence to a previously reported study which showed that an extra copy of the *PHYB* in WT background caused an exaggeration of a shorter hypocotyl phenotype (Wester *et al*., 1994). These results reveal that in Col-0 and other MAGIC

parental accessions, the photomorphogenic hypocotyl length responses are sensitive to *PHYB* copy number.

Due to the presence of the transgene dosage effects, varying phenotypes may be observed in the segregating hemizygous and homozygous plants (Dai *et al*., 1999). The phenotype of hypocotyl length may be affected due to transgene dosage effect in the homozygous or hemizygous states. To avoid any such unwanted effects, T3 homozygous stable individuals were identified after screening with Basta resistance and phenotyped for the transgenic complementation results. The single copy multiple independently transformed lines were tested for any possible between-transformants phenotypic variability. For example, the examination of Ct-1 independently transformed plants, in which the $PHYB_{Col\text{-}0}$ construct has been randomly inserted, revealed no degree of variability in the hypocotyl length phenotype at 22°C. Therefore, these results indicated that there were no position effects of the random insertion of the *PHYB* transgene in the independently transformed lines and the adjacent Ct-1 plant genomic DNA is not influencing any expression changes of the Col-0 allele. However, the control Ct-1 independent Line 5 which had two copies of the Col-0 allele showed variability in the hypocotyl phenotype when compared to the independent lines 1-4 which had single transgene copies but this exaggerated phenotype was associated to an extra copy of the transgene.

Thirdly, low transformation efficiency is observed during the Agrobacterium-mediated transformation process (1 – 2%). This can be addressed by using a selectable marker gene (SMG) in the construct; such as Basta resistance and hence it is a necessity that all transformed plants be identified in this way. In this study the Basta selection has been extraordinarily useful and proved very effective in eliminating all of the untransformed seedlings in all the generations, as explained in section 5.2 by negative selection – a method which kills all the non-transgenic seedlings.

Gateway cloning technology was used to make the different constructs. For functional analysis of the candidate genes, it is a premier cloning technology and is an ideal method, as it is very efficient. It is a recombination based cloning strategy that allows cloning of a gene in a quick and robust way. It also maintains the orientation without using any restriction enzymes. There are two main components of Gateway. First is inserting DNA

into a Gateway vector (plasmid) by generating an 'entry clone' and the second is to transfer this into an expression system. In the first step, PCR is used to amplify the candidate gene with primer sequences that are compatible with the Gateway system. A simple five minute reaction allows the insertion of the candidate gene into the vector. The second step involves a subcloning technique in which a recombination reaction takes place to move the candidate gene into an expression system (bacterial expression). With the help of Gateway technology, the T-DNA in the expression vector was engineered by incorporating the gene of interest in the construct.

QTL cloning by transgenic complementation analysis is a very laborious and time consuming method but is a definitive way of identifying the underlying natural variation in candidate genes. Very few candidate genes until now have been proven in plants by genetic transformation methods. In this study, transgenic methods have provided very important tools to prove whether or not QTL candidate genes actually underlie natural variation in hypocotyl length in response to temperature. This novel contribution to the field in dissecting the natural variation in hypocotyl length, a plant architectural trait, by using transgenic analysis to clone QTL has been very robust and effective.

The transgenic complementation analysis was carried out on both knockout mutants and accessions for any particular candidate gene. An interesting experiment for the knockouts is to complement a deficient phenotype for hypocotyl length with the corresponding missing WT candidate gene. If the QTL candidate transgene rescues the functional phenotype, then it can be assumed that the candidate gene is the gene of interest and can be further tested in the parental accessions which contribute to that particular QTL. This gene rescue experiment in the knockouts is considered as the standard testing procedure to prove that the T-DNA mutated candidate gene is controlling the phenotype.

For all the transgenic lines, the seedlings of the T2 generation were not phenotyped for analysis because the genotype (homozygous and hemizygous) of a line was not determined. In the T2 generation only those lines were selected for further analysis which showed a close to 3:1 segregation for the transgene. As mentioned in the section 5.2, for each of the single copy independently transformed T2 Basta-resistant lines, 15 individuals (homozygous and heterozygous) were selected randomly and allowed to self pollinate forming the T3 seeds. These T3 seeds were sown on pots and sprayed with Basta and

lines in which all the individuals survived were determined as T3 homozygous lines. The stable homozygous T3 lines, identified as a result of the described screening in the section 5.2.2.4, were used for the data collection and empirical analysis of hypocotyl length.

Generally, artificial mapping populations have been fruitful in isolating and cloning QTL in plants (Alonso-Blanco *et al*., 2009). Specifically, analysing the MAGIC mapping population has been clearly successful in identifying and cloning QTL in Arabidopsis for hypocotyl length in response to temperature. The functional testing of QTL candidate genes was performed by transgenic complementation. This approach was carried out in order to establish molecular causality for the natural variation in hypocotyl length in response to temperature. QTL cloning by transgenic complementation experiments is an important approach in establishing allelic functional variation in candidate genes for hypocotyl length in accessions. Though it is labour-intensive, it is a definitive method which provides direct evidence for natural variation (Weigel and Nordborg, 2005). With the help of transgenic complementation, the genetic analyses of hypocotyl length were pursued down to the molecular level facilitating the molecular dissection of natural variation in hypocotyl length.

It has been previously argued that mutagenesis is the way forward for dissecting complex traits and that QTL cloning is an inferior method as compared to it (Nadeau *et al*., 2000); however, this study has demonstrated that QTL cloning has identified novel functional variation and has proved as an effective method in discovering the underlying genetic variation in natural populations. Furthermore, with mutagenesis, quantitative traits may not be properly understood in natural populations as mutagenesis is usually carried out in one genetic background whereas, QTL mapping in the MAGIC population has exploited natural variation in many parental accessions. And as an example, in this study novel alleles have been discovered that control natural variation. A more effective way forward is to combine mutational analysis with QTL mapping and in this study the use of mutagenesis has been exploited in a pre-QTL cloning step to identify functional involvement of the candidate genes in the hypocotyl length trait.

Until recently, the general understanding was that quantitative traits were governed by many QTL and a large number of genes (Barton and Turell, 1989). If this were the case, then QTL cloning of hypocotyl length in temperature response would have not been

feasible because of the prevalence of many small effect QTL. However, since the high density genetic maps of molecular markers were developed, dissection of quantitative traits in many species became possible (Tanksley, 1993; Lander and Schork, 1994). Recently, several QTL studies have identified few major effect QTL for various quantitative traits controlling developmental modifications in Arabidopsis (Strange *et al*., 2011; Salomé *et al*., 2011), in maize (Brown *et al*., 2011) and in rice (Huang *et al*., 2010; Huang *et al*., 2012)**.** The results of the QTL mapping in this study are complementary in this regard showing that hypocotyl length is a trait which is controlled by a few major effect QTL at individual temperatures. This is one of the contributing characteristics due to which QTL cloning has been successfully achieved. A few QTL with large effects represent the fact that few genomic locations are contributing to the variation of the trait. This feature of hypocotyl length trait being under the control of few QTL in response to temperature restricts the search of causal stretches of DNA in these specific few QTL. Furthermore, as described in Chapter 3, the QTL also have narrow boundaries limiting the total number of genes located within them.

Collectively, due to the nature of the identified QTL being few and major effect, and presence of strong candidate genes in the QTL, at various temperature treatments, Mendelisation of the QTL was feasible as part of the QTL cloning process. Although hypocotyl length is a quantitative trait under the control of more than one locus as shown in the segregation analysis of the two parental accessions Col-0 and Sf-2 in Chapter 2, for the identification of the causal alleles it is necessary to Mendelise the major effect QTL by studying individual candidate genes. The candidate genes in this chapter have been Mendelised as part of a process of fine gene mapping and for understanding their effects on the variation of hypocotyl length. The assumption is that individual genes behave in a Mendelian fashion. To give specific examples, in the case of *PHYB* alleles, the hypocotyl length phenotypes of Ct-1 and Sf-2 accessions are drastically different from Col-0 and this phenotypic difference was physically seen in the T2 segregating lines as seen in figures 5.9 and 5.13 respectively. It is quite interesting to note that for both Ct-1 and Sf-2, the phenotypic observation in the T2 segregating population revealed that the independent lines transformed with a single *PHYB* transgene showed simple Mendelian inheritance (3:1). This provided further phenotypic evidence for *PHYB* being a large effect QTL candidate gene which showed monogenic inheritance of the trait and in this way the QTL

had been Mendelised. Whereas, for the other candidate genes tested, no significant and physically observable differences could be seen in the hypocotyls e.g. figures 5.20, 5.22, 5.27 and 5.28, where no segregation was seen. In summary, the *PHYB* phenotyping has shown successful transgenic complementation of the lines. From the experimental testing with this method, the *PHYB* QTL has been cloned and the underlying genetic basis of the natural variation that is observed in hypocotyl length has been dissected.

The functional analysis of multiple alleles is essential in determining natural variation in candidate genes. This approach of analysing multiple alleles before assigning any gene function is a standard principle of genetics (Hirschi, 2003).

5.4.2 Discussion on Results:

The Temp22.2 QTL has been localised to the *PHYB* gene and therefore, *PHYB* was an obvious strong candidate in this QTL. Like the other candidate genes, *PHYB* was selected for further allelic functional investigation based on the mean effect size estimates of the parental accessions at the *PHYB* locus obtained from the parental reconstruction step in the MAGIC QTL mapping. This was combined with the *in silico* analysis which identified *PHYB* allelic sequence polymorphisms in the major effect size parents. The identification of *PHYB* as a candidate gene is twofold. Firstly, *PHYB* is a functional candidate gene involved in hypocotyl length. Secondly, QTL mapping in MAGIC lines identified *PHYB* as a strong positional candidate gene in response to temperature, because positional candidates can be chosen according to their proximity to the segregating QTL. In this case, the co-segregation of a functional candidate gene and a QTL is taking place.

In a previous study, QTL mapping in a biparental population was done in response to light which identified LIGHT2 QTL (Borevitz *et al*., 2002). In that study, although *PHYB* was not within the 8 cM LIGHT2 QTL, it was suggested as a candidate based on its proximity to GPA1, a marker which was 14 cM distal to *PHYB*. At a significance threshold of $p = 0.01$, the 8 cM LIGHT2 QTL interval contained ~200 other genes. Conversely, the QTL mapping, in this study with the MAGIC population at a significance threshold of $p = 0.01$, identified *PHYB* as the only gene in the <0.1 cM Temp22.2 QTL in response to temperature. The SNP marker 'PHYB_2850' is present within the *PHYB* gene and therefore, the gene and the marker are very tightly linked. The position of the segregating marker 'PHYB_2850' is very

convincing statistical evidence that Temp22.2 QTL is a real QTL. At a significance of p = 0.05, the QTL interval is still fine mapped at only 2 cM with ~ 100 genes. Within the Temp22.2 QTL, the peak marker 'PHYB_2850' was statistically the most strongly associated locus with hypocotyl length at 22°C. By using the MAGIC population the QTL has been directly mapped to *PHYB* based on fine mapping and previously it has been shown that QTL can be fine mapped directly to genes (Fridman *et al*., 2000; Kroymann *et al*., 2001).

The transgenic allelic complementation experiments have proved very effective in elucidating the role of *PHYB* in natural variation. *PHYB* has a very clear effect, with several loss of function or reduced function mutants that are complemented by the functional Col-0 allele. In particular, the results reveal naturally occurring variation in *PHYB* in the Ct-1 and Sf-2 alleles which drastically increase seedling stem elongation and plants have longer hypocotyls. Due to the mutations in the Ct-1 and Sf-2 accessions, it seems that *PHYB* is not functional or less functional and therefore with the incorporation of a WT *PHYB* allele from Col-0, the short hypocotyl phenotype is restored. Ct-1 and Sf-2 have complemented the Col-0 short hypocotyl length phenotype. In all of the tested accessions, *PHYB* Col-0 seems to be the dominant allele. These experiments demonstrate the definitive proof that *PHYB* is involved in controlling the natural variation of hypocotyl length phenotype at 22°C. Had the transgenic allelic complementation experiments not been carried out, the underlying causal alleles of the QTL would have not been discovered.

The transgenic transformation experiment in the *phyB* knockout mutant Col-0 background was a control experiment and the insertion of the *PHYB* allele from Col-0 rescued the WT phenotype in the *phyB* knockout. These experimental findings are similar to a previously reported study (Wester *et al*., 1994) which was looking at the effect of gene dosage and its expression. It showed that a mutant allele of *PHYB* Bo64, which had a loss of function mutation due to a premature stop codon, was transformed with a WT allele which restored the phenotype. It complemented all the mutant phenotypes including hypocotyl length.

These results demonstrate that the mutations in the different *PHYB* alleles in all four accessions could cause a partial or complete loss of function affecting the hypocotyl

phenotype in the same direction as the *phyB* KO mutant, i.e. long hypocotyls. With the incorporation of the allele from Col-0, accessions both with drastic and subtle differences in the hypocotyl lengths were transformed to the short phenotype of Col-0. Furthermore, these results show that multiple alleles of *PHYB* gene are involved in controlling the natural variation of the trait. This highlights the possibility of different mutations in the *PHYB* gene having differential effects on the phenotypes of the hypocotyl lengths. Given the fact that the same short hypocotyl length phenotype is linked to the same transgene in several independently transformed individuals provides a strong argument that the mutations in the *PHYB* alleles may well be responsible for the altered phenotypes and are controlling natural variation. Hence, this suggests that in Arabidopsis *PHYB* is involved in producing either long or short hypocotyl lengths at 22°C. These results demonstrate the effect of an individual QTL affecting the quantitative trait; and that the phenotypic variation in hypocotyl length between the various accessions is due to the major effect of allelic variation at *PHYB* locus.

If the candidate alleles are of small or moderate effect, quantitative traits may prove hard to study by transgenic complementation (Tian *et al*., 2003). However, QTL mapping with the MAGIC population has identified *PHYB* as a major effect QTL and experimental evidence from the transgenic complementation of various accessions strongly supports this finding. Additionally, the importance of *PHYB* as a major effect gene can be realised by the fact that several independently transformed lines which have variability in the transgene insertions display similar phenotypes, otherwise very often small to moderate effect QTL display heterogeneous phenotypes.

The experimental results on *PHYB* have revealed that its alleles have very large effects on the highly heritable trait of hypocotyl length. This is because *PHYB* is a major effect gene as observed in the phenotypic analysis of transgenic plants and also it is highly heritable as established in Chapter 2. After Mendelising the trait at the *PHYB* locus, the results of transgenic complementation show that allelic differences at this particular locus at 22°C are unequivocally responsible for the trait variation. Collectively, from previous studies and the findings in this study, it could be hypothesised that *PHYB* is involved in and is key to the evolution of hypocotyl length differences within Arabidopsis accessions.

The photobiology of phytochromes in Arabidopsis is well studied in general and some studies on natural variation in phytochromes in response to light have been carried out. It interesting to note that to date, no study on natural variation in photoreceptors in response to temperature has been performed. Phenotypic variation in hypocotyl length due to a missense substitution in the coding sequence in *PHYA* has been shown (Maloof *et al*., 2001) through a genetic complementation test to determine a *PHYA* naturally occurring allele in the Le Mans (Lm-2) accession. It was shown that this allele had reduced far-red sensitivity due to an amino acid change altering the protein resulting in a substantial loss of *PHYA* function. In another study, a nonsense substitution forming a truncated protein in the coding sequence in *PHYC* was shown to be responsible for phenotypic variation in hypocotyl length in light response (Balasubramanian *et al*., 2006). Due to a natural deletion in *PHYD* in the accession Wassilewskija (Ws), hypocotyl elongation was increased in response to red/far-red light sensing (Aukerman *et al*., 1997). This 14 bp deletion in the coding sequence caused a naturally existing loss of function allele in Ws due to a truncated protein. However, in white light the variation was more subtle are compared to red/far-red light. The loss of function mutation shown in *PHYD* (Devlin *et al*., 1998) was in an induced mutant and is not a natural variation study. In all of these above studies, changes in the different phytochrome genes were limited to individual accessions.

The importance of *PHYB* as a candidate gene in this study can be evaluated from previous QTL studies in Arabidopsis which have suggested *PHYB* as a candidate in response to light treatments. *PHYB* was proposed as a candidate for a LIGHT2 QTL in response to red and white light (Borevitz *et al*., 2002) in a RIL population derived from a biparental cross of Cape Verde Islands (Cvi) and Landsberg erecta (Ler) accessions. In their study, LIGHT2 QTL was confirmed in a NIL on the basis of which they suggested *PHYB* as a candidate gene.

In another study QTL were mapped for light response in RILs obtained from the Columbia (Col-gl1) and Kashmir (Kas-1) accessions (Wolyn *et al*., 2004). A RED2 QTL in response to red light was associated with *PHYB* and the Kas-1 allele contributed to long hypocotyls while those from Col-gl1 contributed to short ones. They suggested *PHYB* as a candidate for RED2 QTL because the Kas-1 allele had a 12-bp deletion and an amino acid substitution as compared to Col allele. However, *PHYB* was only detected in red light

whereas in the LIGHT2 QTL in the Ler × Cvi population (Borevitz *et al*., 2002), it was also detected in white light.

Since Ler and Cvi accessions were previously shown to segregate for the LIGHT2 QTL and *PHYB* was a candidate gene (Borevitz *et al*., 2002), the two alleles of Ler and Cvi *PHYB* were experimentally tested for functional variation in hypocotyl length (Filiault *et al*., 2008). They conducted an overexpression experiment on the two different alleles in a loss of function *phyB* mutant and showed that due to amino acid polymorphisms *PHYB*-Cvi conferred less light responsiveness than *PHYB*-Ler in transgenic *phyB* Col-0 mutants. Hence transgenic plants carrying the *PHYB*-Ler allele were shorter than those carrying *PHYB*-Cvi allele.

Another QTL study (Botto *et al*., 2003), which also searched for hypocotyl length QTL involved in light responses in Arabidopsis*,* used the Ler × Cvi population and identified five QTL but no candidate genes were suggested. In a more recent study (Kasulin *et al*., 2013), *PHYB* was proposed as a candidate gene for hypocotyl length EODRATIO5 QTL in the Ler x Cvi-0 RIL population in shade avoidance and light response which was simulated by end-of-day far-red (EOD) treatment. Natural variation in all of the above mentioned studies were limited to two parental accessions which showed very subtle differences in hypocotyl lengths in various light responses. However, in this study accessions with drastic hypocotyl length differences in temperature responses have been found and the underlying *PHYB* QTL has been successfully cloned in multiple MAGIC parents.

Various photoreceptor mutants were initially identified in light responses, for example functional characterisation of the primary photoreceptor *PHYB* was carried out by studying *phyB* mutants (Reed *et al*., 1993; Somers *et al*., 1993). These mutants displayed altered hypocotyl lengths in light responses and therefore, QTL mapping in hypocotyl length has been previously focused in light responses as described above. It has only been a relatively recent event that the phenotypic variation in hypocotyl length due to polymorphisms in *PHYB* in a QTL approach has been described in response to light by experimentation (Filiault *et al*., 2008). More recently, in Chapter 2, it has been demonstrated that natural variation in hypocotyl length is prevalent in response to temperature in multiple accessions. It is of profound interest to find the underlying genetic elements responsible for natural variation in specific temperature responses.

PHYB is involved in thermal responses in Arabidopsis flowering time (Halliday *et al*., 2003) and hence it is also relevant to identify *PHYB* alleles that control another growth related trait - hypocotyl length - in response to temperature. This could further help to develop our understanding of the role of *PHYB* in an important seedling architectural trait. For an in-depth understanding of natural variation in hypocotyl length, it is essential to explore whether Temp22.2 QTL effects are dependent on the temperature environment and genetic backgrounds in which it is segregating. Since this QTL is temperature dependent, i.e. it has only been detected at 22°C, it shows that *PHYB* action is dependent on the specific environment.

The phytohormone Indole aceticacid (IAA) is the major natural auxin found in plants. It controls almost all aspects of plant growth and development through modulation of gene expression achieved through the degradation of repressor proteins called Aux/IAA proteins. The *SAUR* gene family (McClure *et al*., 1987) is one of the three gene families that are induced by auxin. In Arabidopsis, it consists of over 75 genes that belong to several groups depending on their sequence homology. Due to their importance of being responsive to auxin, and potentially involved in growth promotion, *SAUR* genes have also been studied in model crop plants such as rice (Kant *et al*., 2009). However, their exact functions in plants are not yet properly understood.

As described in the discussion section of chapter 4, the expression of *SAUR* genes is usually connected with elongating tissue in the hypocotyls (Roig-Villanova *et al*., 2007) and also at higher ambient temperatures (Franklin *et al*., 2011). However, it has been reported that no direct role of any *SAUR* gene in growth promotion has been demonstrated (Stavang *et al*., 2009). The candidate gene *SAUR38* has had a putative role in auxin response; however, the combined results of the knockout and transgenic analysis have discovered and suggested a novel role for *SAUR38* by being involved in controlling hypocotyl length in Arabidopsis. This new role has been functionally identified in homozygous *saur38* knockout lines and subsequently validated by transgenic complementation and gene copy number analysis by direct experimentation in different Arabidopsis accessions showing differential hypocotyl length phenotypes. Results from the transgenic analysis reveal that *SAUR38* positively regulates hypocotyl elongation and this finding is congruent to the results of another study on a member of the *SAUR* gene

family, that has shown that the overexpression of Arabidopsis *SMALL AUXIN UP RNA 63* (*SAUR63*) promoted hypocotyl elongation (Chae *et al*., 2012). The results of this study have shown that at 22°C, *SAUR38* not only controls hypocotyl elongation but is also involved in its natural variation.

For the candidate genes *CYP71*, *NPY2*, and *CYCT1;4*, the transgenic results did not show a neat complementation. However, this does not imply that they may not be involved in natural variation. For these genes, there is an effect of the knockout on the hypocotyl phenotype. Additionally, the complementation of the knockout by the candidate genes was also seen. Transgenic complementation works if a functioning transgene is added into a line with a loss of function mutant allele. For example, in the transgenic complementation experiment of Zu-0 accession with $CYCT1;4_{Col\text{-}0}$ allele, if the Zu-0 allele of *CYCT1;4* is an always on mutant or produces a functional protein with a large effect, then adding a Col-0 allele with a smaller or better regulated effect would have little impact. Therefore, in such experiments, the simple effect of complementation cannot be observed because either the candidate is part of a gene network, or is a loss of function or reduced function allele. In such a situation, the Col-0 alleles seem to behave in a recessive manner. This is nicely illustrated in the *SAUR38* example, where the transgenic analysis shows that the $SAUR38_{Tsu\text{-}0}$ allele is acting in a dominant way, so that long hypocotyl phenotype of Tsu-0 cannot be rescued by the $SAUR38_{Col\text{-}0}$ allele (figure 5.24). It is also possible that the candidates participate in gene networks or pathways. There may also be interactions with other genes, e.g. one of these genes may only make a difference in the presence of a functional (or maybe a non-functional) *PHYB* gene. Therefore, transforming them into Col-0 will give a different response to their function in the accession that they come from.

QTL cloning for hypocotyl length trait in response to temperature in Arabidopsis has not been previously performed. Successful QTL cloning of *PHYB* for hypocotyl length trait by transgenic complementation in response to temperature in natural accessions has been carried out for the first time in this study. This study brings a twofold novelty; firstly, it has cloned a QTL for hypocotyl length which is a growth, developmental and plant architectural trait; secondly, it is in response to temperature which is one the most important environmental abiotic factors governing plant growth and development.

5.5 References:

Aukerman M, Hirschfeld M, Wester L, Weaver M, Clack T, *et al*. (1997) A Deletion in the *PHYD* Gene of the Arabidopsis Wassilewskija Ecotype Defines a Role for Phytochrome D in red/far-red Light Sensing. Plant Cell, 9, 1317–1326.

Balasubramanian S, Sureshkumar S, Agrawal M, Michael T, Wessinger C, *et al*. (2006) The PHYTOCHROME C Photoreceptor Gene Mediates Natural Variation in Flowering and Growth Responses of Arabidopsis *thaliana* . Nat Genet, 38, 711–715.

Barton N, Turell M. (1989) Evolutionary Quantitative Genetics. Annu. Rev. Genet. 23, 337 – 370.

Bechtold N, Ellis J, Pelletier G. (1993) In planta *Agrobacterium* Mediated, Gene Transfer by Infiltration of Adult *Arabidopsis thaliana* Plants. C R Acad Sci Paris, 316, 1194 -1199.

Borevitz J, Chory J. (2004) Genomics Tools for QTL Analysis and Gene Discovery. Current Opinion in Plant Biology, 7, 132–136.

Borevitz J, Maloof J, Lutes J, Dabi T, Redfern J, *et al*. (2002) Quantitative Trait Loci Controlling Light and Hormone Response in Two Accessions of *Arabidopsis thaliana*. Genetics, 160, 683–696.

Botto J, Alonso-Blanco C, Gararón I, Sánchez R, Casal J. (2003) The Cape Verde Island Allele of Cryptochrome 2 Enhances Cotyledon Unfolding in the Absence of Blue Light in Arabidopsis. Plant Physiol. 133, 1547–1556.

Brown P, Upadyayula N, Mahone G, Tian F, Bradbury P, *et al*. (2011) Distinct Genetic Architectures for Male and Female Inflorescence Traits of Maize. PLoS Gene. 7,e1002383.

Clough S, Bent A. (1998) Floral Dip: A Simplified Method for *Agrobacterium*-Mediated Transformation of Arabidopsis *thaliana*. The Plant Journal, 16, 735–743.

Devlin P, Patel S, Whitelam G. (1998) Phytochrome E Influences Internode Elongation Flowering Time in Arabidopsis. Plant Cell, 10, 1479–1487.

Drinkwater N, Gould M. (2012) The Long Path from QTL to Gene. PLOS Genetics, 8, 9.

El-Assal S, Alonso-Blanco C, Peeters A, Raz V, Koornneef M. (2001) A QTL for Flowering Time in Arabidopsis Reveals a Novel Allele of CRY2. Nat Genet, 29, 435–440.

Filiault D, Wessinger C, Dinneny J, Lutes J, Borevitz J, *et al*. (2008) Amino Acid Polymorphisms in Arabidopsis Phytochrome B cause Differential Responses to Light. Proc Natl Acad Sci USA, 26, 105, 8, 3157–3162.

Frary A, Nesbitt T, Grandillo S, Knaap E, Cong B, *et al*. (2000) fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. Science, 289, 85–88.

Fridman E, Pleban T, Zamir D, (2000) A Recombination Hotspot Delimits a Wild-Species Quantitative Trait Locus for Tomato Sugar Content to 484 bp within an Invertase Gene. Proc Natl Acad Sci USA, 97, 4718–4723.

Hagen G, Guilfoyle T. (2002) Auxin-Responsive Gene Expression: Genes, Promoters, and Regulatory Factors. Plant Mol. Biol. 49, 373–385.

Hirschi K. (2003) Insertional Mutants: A Foundation for Assessing Gene Function. Trends in Plant Science, 8, 5.

Huang X, Wei X, Sang T, Zhao Q, Feng Q, *et al*. (2010) Genome-Wide Association Studies of 14 Agronomic Traits in Rice Landraces. Nat. Genet. 42, 961–67.

Huang X, Zhao Y, Wei X, Li C, Wang A, *et al*. 2012. Genome-Wide Association Study of Flowering Time and Grain Yield Traits in a Worldwide Collection of Rice Germplasm. Nat. Genet. 44, 32–39.

Kasulin L, Agrofoglio Y, Botto JF. (2013) The Receptor-like Kinase ERECTA Contributes to the Shade-Avoidance Syndrome in a Background-dependent Manner. Ann Bot. 111, 5, 811-9.

Kroymann J, Textor S, Tokuhisa J, Falk K, Bartram S, *et al*. (2001) A Gene Controlling Variation in Arabidopsis Glucosinolate Composition is part of the Methionine Chain Elongation Pathway. Plant Physiol, 127, 1077–1088.

Lander E, Schork J. (1994) Genetic Dissection of Complex Traits. Science, 265, 2037 – 48.

Luo Z, Wu C, Kearsey M. (2002) Precision and High-Resolution Mapping of Quantitative Trait Loci by use of Recurrent Selection Backcross or Intercross Schemes. Genetics, 161, 915–929.

Maloof J, Borevitz J, Dabi T, Lutes J, Nehring R, *et al*. (2001) Natural Variation in Light Sensitivity of Arabidopsis. Nat. Genet. 29, 441–46.

McClure B, Guilfoyle T. (1987) Characterization of a Class of Small Auxin-Inducible Soybean Polyadenylated RNAs. Plant Mol. Biol. 9, 611–623.

Nadeau J, Frankel W. (2000) The Roads from Phenotypic Variation to Gene Discovery: Mutagenesis Versus QTLs. Nature Genetics, 25, 4, 381-384.

Reed J, Nagpal P, Poole D, Furuya M, Chory J. (1993) Mutations in the Gene for the red-far-red Light Receptor Phytochrome B Alter Cell Elongation and Physiological Responses throughout Arabidopsis Development. Plant Cell 5, 147–157.

Roig-Villanova I, Bou-Torrent J, Galstyan A, Carretero-Paulet L, Portoles S, *et al*. (2007) Interaction of Shade Avoidance and Auxin Responses: A Role for Two Novel atypical bHLH Proteins. EMBO J. 26, 4756–4767.

Salomé P, Bomblies K, Laitinen R, Yant L, Mott R, *et al*. (2011) Genetic Architecture of Flowering-Time Variation in Arabidopsis thaliana. Genetics, 188, 421-433.

Salvi S, Tuberosa R. (2005) To Clone or not to Clone Plant QTLs: Present and Future Challenges. Trends Plant Sci. 10, 6, 297-304.

Somers D, Sharrock R, Tepperman J, Quail P. (1993) The hy3 Long Hypocotyl Mutant of Arabidopsis is Deficient in Phytochrome B. Plant Cell 3, 1263–1274.

Stavang J, Gallego-Bartolomé J, Gómez M, Yoshida S, Asami T, *et al*. (2009) Hormonal Regulation of Temperature-Induced Growth in Arabidopsis. The Plant Journal, 60: 589–601.

Strange A, Li P, Lister C, Anderson J, Warthmann N, *et al*. (2011) Major-Effect Alleles at Relatively Few Loci Underlie Distinct Vernalization and Flowering Variation in Arabidopsis Accessions. PLoS ONE 6, 5,e19949.

Tanksley S. (1993) Mapping Polygenes. Annu. Rev. Genet. 27, 205 – 233.

Tian D, Traw M, Chen J, Kreitman M, Bergelson J (2003) Fitness Costs of R-Gene-Mediated Resistance in *Arabidopsis thaliana*. Nature 423, 74–77.

Weigel D, Nordborg M. (2005) Natural Variation in Arabidopsis. How Do We Find the Causal Genes? Plant Physiol, 138, 567–568.

Wester L, Somers D, Clack T, Sharrock R. (1994) Transgenic Complementation of the hy3 Phytochrome B Mutation and Response to PHYB Gene Copy Number in Arabidopsis. Plant J, 5, 2, 261-72.

Wolyn D, Borevitz J, Loudet O, Schwartz C, Maloof J, *et al*. (2004) Light-Response Quantitative Trait Loci Identified with Composite Interval and eXtreme Array Mapping in *Arabidopsis thaliana*. Genetics, 167, 2, 907-917.

5.6 Appendix:

Gene	Primer	Primer Sequence	bp	Tm	Net Primer Rating
PHYB	Left	CACCCTCCTCCGCCACATCCTTTG	20+4	62.65	100
	Right	GGCCTTTACCTCTTGATTGCGTA	23	62.75	83
SAUR	Left	CACCCATGGTGTGCGATGTTTTGAGC	22+4	63.66	90
	Right	TGGGCTGGTGGGTTCTTGTG	20	63.67	100
CYP71	Left	CACCTCTGATATGCAGTGGCAGAACCA	23+4	62.98	85
	Right	GTCCCCACGGAAAGGAGTCG	20	63.86	99
NPY2	Left	CACCATGCTGTGTATCGCCGTGGTT	21+4	62.69	100
	Right	CGCCACCGTGTCGTTTGAGT	20	63.94	100
CYCT1	Left	CACCGTGGAAATGTGGCAAATCAAGCA	23+4	64.67	100
	Right	CATCCTTCGGCAATGAATGACAA	23	64.61	100

Table 5.1: Primers used for PCR amplification of genomic DNA of candidate genes

Gene	Primer	Primer Sequence	bp	Tm	Net Primer Rating
PHYB Col-0	Left 1	tccccatttttcttcttcctcaa	22	60.54	100
	Left 2	GCCGTCGATGAATCCAGTTT	20	59.81	87
	Left 3	TTCCTCAAGCGTCAAGGTTCTT	22	60.8	91
	Left 4	CCGTTTCCGCTAAGGTATGC	20	60.09	100
	Left 5	AGAATCTGAGGCGGCTATGAAC	22	60.34	100
	Left 6	AAGGGAAAGCAGTTTTTGTGGTT	23	61.47	100
	Left 7	TGGGGCTTTCTGTTTCTTGC	20	60.25	100
	Left 8	ATCCGTGACATTCCCGAAGA	20	59.95	91
	Left 9	tcaaattgatgaaaaccagctca	23	60.12	86
SAUR38 Col-0	Left 1	GCAGAAGCAAGCAATGAACAA	21	59.11	100
	Right 1	aaaccaaagcatatgattagcacaa	25	60.57	85
SAUR38 Tsu-0	Left 1	AAACCAAAGCATATGATTAGCACAA	25	60.57	85
	Right 1	CGCGTCTTATCATCTGCTGCT	21	60.68	81
NPY2 Col-0	Left 1	ggtccccaccctcagttctt	20	60.02	100
	Left 2	gcttgcacttttgtgcctagaa	22	60.15	84
	Left 3	GGTATTCCCGGTGGTCCAA	19	60.07	82
	Left 4	TTGGTGGAGAGGTCATTGGAG	21	59.62	100
	Left 5	CAAAGCTGATGGTGGCAAAAG	21	60.85	88
CYCT Col-0	Left 1	AACCAAGCTGATCCCTCTTATGAA	24	61.4	88
	Left 2	AGGTTCTGTAGGATCGAATTTTGAA	25	60.9	87
	Left 3	ttttccttgcgtccgcttat	20	60.21	100
	Left 4	tttgcactgtttctcgcatattct	24	61.21	86
	Left 5	TTTTTCTTGCTGCGAAGTTCC	21	60	91

Table 5.2: Primers used for 500 bp sequencing of pENTR DNA of candidate genes

Independently Transformed Line	No. of Basta Resistant Seedlings	No. of Basta Sensitive Seedlings	Observed Resistant to Sensitive Ratios	Estimated transgene copy number from T2 segregation analysis	Chi square results P=0.05
Line 1	157	44	3:1	1	1.03
Line 2	169	63	3:1	1	0.57
Line 3	166	44	3:1	1	1.83
Line 4	128	35	3:1	1	1.08
Line 5	180	19	15:1	2	3.68
Line 6	110	32	3:1	1	0.46
Line 7	195	17	15:1	2	1.13
Line 8	93	39	3:1	1	1.45
Line 9	134	36	3:1	1	1.32
Line 10	135	15	15:1	2	3.59

Table 5.3: T2 segregation of Ct-1 construct in independently transformed lines determining the ratios of Basta-resistant to Basta-sensitive seedlings

Independently Transformed Line	No. of Basta Resistant Seedlings	No. of Basta Sensitive Seedlings	Observed Resistant to Sensitive Ratios	Estimated transgene copy number from T2 segregation analysis	Chi square results P=0.05
Line 1	110	44	3:1	1	1.04
Line 2	98	44	3:1	1	2.71
Line 3	145	38	3:1	1	1.75
Line 4	90	20	3:1	1	2.72
Line 5	157	68	3:1	1	3.27
Line 6	90	27	3:1	1	0.23
Line 7	102	28	3:1	1	0.83
Line 8	127	47	3:1	1	0.37
Line 9	147	16	15:1	2	3.54
Line 10	133	14	15:1	2	2.69

Table 5.4: T2 segregation of Sf-2 construct in independently transformed lines determining the ratios of Basta-resistant to Basta-sensitive seedlings

Chapter 6

General Discussion, Conclusions and Future Prospects

6.1 General Discussion and Conclusions:

Ambient temperature is an essential environmental regulator of plant growth and development, and hypocotyl length is an ecologically relevant trait that enables the plant to survive in its infancy. Despite the importance of these two, studies investigating natural variation in hypocotyl length in response to ambient temperatures are lacking. Hence, this is a pioneering study in QTL cloning, as to date QTL have not been cloned in Arabidopsis for hypocotyl length in response to ambient temperature. It is the first example of cloning a QTL using the MAGIC experimental population. Employing the MAGIC lines in a forward genetics approach has proven to be a quick way of finding the genes involved in the natural variation of an important plant growth trait.

This study comprises of basic research connecting phenotype to genotype by using the developmental process of hypocotyl elongation. The genetic investigation began with an observed variation in hypocotyl morphology in the MAGIC parents. Natural genetic differences can partly describe the intraspecific variation seen in Arabidopsis due to widespread sequence variation (Nordborg *et al*., 2005; Schmid *et al*., 2003). Therefore, the preliminary question was whether the observed variation in the trait was influenced at all by genetic variation? If yes, then are there alleles segregating in the population that produce a differential effect on the character, or is all the variation simply the result of environmental variation? This study has documented that natural variation, combined with an important environmental factor, is responsible for the observed intraspecific variation in hypocotyl length.

This study has attempted to dissect the genetic variation in the environment-dependent developmental process of hypocotyl elongation in response to temperature. It has focused on the natural variation in thermal sensitivity of hypocotyl length in a diverse mapping set of Arabidopsis. It aimed at discovering new allelic variants that affect

hypocotyl length in natural accessions. A connection between specific allelic variants and phenotypic variation has improved our basic knowledge of plant form and growth, as hypocotyl length is a plant architectural and developmental trait. This study aimed at gaining a mechanistic understanding of the connection between temperature and plant growth. The dramatic elongation of hypocotyl length and their differential responses illustrate temperature control of plant development in the juvenile stage.

With evidence of rising mean global temperatures due to climate change (IPCC, 2013) and many plant species responding to global warming by advancing their flowering time (Parmesan 2006), studies on plant-temperature interactions are necessary. Current changes in climate strongly affect ambient growth temperature, which has direct effects on plant functioning in all developmental stages.

With the current observed effects of climate change and increased temperatures of 2-3°C by the end of the 21st century (IPCC, 2013), plant thermo-morphogenesis is likely to emerge as a model for investigating interactions of the environment with plant development. In this regard, recently a field study on the MAGIC lines was performed for phenological sensitivity and fitness responses to simulated increases in temperature (Springate and Kover, 2014). The study found that plants exposed to increased temperatures flowered earlier than plants grown in ambient temperatures, and significant G x E interaction was found.

One of the post Arabidopsis genome sequencing aims is the rapid discovery of gene and allele functions. Mutation is a major source of genetic variation in natural populations which leads to intraspecific variation. One way of allele discovery is to exploit untapped natural variation to identify polymorphism genes responsible for variation in growth and response to environment. Studying natural variation has proven valuable in the characterisation of hypocotyl length and in identifying the genetic factors controlling it. The genetic variation that is due to the universal prevalence of Arabidopsis in a wide variety of habitats is an ideal resource for examining how organisms have adapted to live under different temperatures. Thus, adaptation to the native environments is likely to have been an important aspect of its evolutionary history. Obtaining knowledge of how plants are capable of adapting to climate is essentially interesting at a scientific level. However, prior to addressing questions of interest to the evolutionary ecology of plants,

it is imperative to make two more immediate problem formulations. Which genes control specific trait variation? What are the allelic variants of these genes functionally affecting a trait? These are undoubtedly hard questions to answer; however, this study has attempted to answer them by exploiting the natural variation in hypocotyl length through a systematic methodology.

Such a methodology consisted of experimental analyses at the organismal and molecular levels; and in a specific order as described in figure 1.4. This study has documented extensive variation in hypocotyl length in accessions in response to ambient temperatures ranging from 12°C to 27°C. The plasticity in the phenotype was caused by the temperature effect. The various accessions reacted differentially to varying temperatures which was explained by the very high G x E interaction. Thus, variable temperature effects on hypocotyl length are vital in understanding plant seedling growth. This finding suggested that hypocotyl length may well be an interesting plant architectural trait to dissect with the help of QTL mapping leading to QTL cloning.

The preliminary understanding of the genetic basis of hypocotyl length was achieved by QTL mapping that evaluated the association between molecular markers and phenotypic variation in MAGIC lines. Since the QTL were fine-mapped, the critical region segregating for the trait was smaller allowing focusing on a smaller number of candidates. A comprehensive analysis included searching for candidate genes in various QTL at different temperatures, instead of limiting search for candidates in any one particular QTL. Candidate genes that are more likely to influence hypocotyl growth were identified from QTL mapping.

Though a high resolution QTL mapping identified several candidate genes, functional strategies were required for QTL cloning. This was carried out by phenotyping homozygous knockout mutants followed by direct testing of the candidate genes through transgenic complementation by importing novel alleles in to the accessions. Functional analysis of natural alleles has enabled the identification of genetic factors that control natural variation in hypocotyl length.

Since hypocotyl length is a continuous trait, the genotypes of the QTL cannot be directly determined from single phenotypic values of a plant, as observed in monogenic

segregation, but it can be indirectly inferred from linked marker loci. The post fine-QTL mapping strategy adopted a single locus approach. In order to clone a large effect QTL, it is important to Mendelise it and work under the assumption that the QTL harbours a single large effect gene. Allelic variants at a QTL have similar molecular basis to the variations at monogenic Mendelian loci. The basis of this is that protein functions or gene expressions vary. The rationale behind focusing on allelic variation is that particular mutations will directly impact how the gene functions, and lead to the phenotype being variable in natural populations.

This study has also attempted to promote synergy between ecology and molecular biology, both which are usually thought of fields of study on the far ends of the spectrum. Based on the fact that climate change will affect future plant growth scenarios, the genetic and molecular bases of plant responses to temperature is becoming the subject of dynamic research in plant environmental response biology. This is becoming to be a new emerging field of environmental sciences which requires an inter-disciplinary approach with other related fields of science.

No single type of experimental evidence may be sufficient to clone a QTL. Nevertheless, several lines of experimental evidence can provide a definitive proof for QTL cloning (Weigel and Nordborg, 2005). There were several types of information required to show that a candidate is the gene for hypocotyl length. First, a candidate was statistically associated with hypocotyl length from QTL mapping. Secondly, in the post QTL mapping, this study carried out several steps of shortlisting that include identification of potential candidates, selecting major effect parental accessions, and finding polymorphisms in the candidate genes of such parents. Thirdly, combinations of different functional strategies were applied to the hypocotyl length natural variation study, that include mutational analysis in which the mutant shows an altered phenotype; and transgenic complementation analysis in which various alleles of candidates have been functionally tested. Thus, various pieces of information have been put together to make a complete picture.

Where could the mutations lie in the genes that give rise to phenotypic variation? Mutations in the regulatory regions, in the introns and in the exons can affect the function of the gene and hence the phenotype that it controls. In the promoter region,

mutations can affect the gene expression levels, e.g. in the case of the fw2.2 growth QTL in tomato (Frary *et al.*, 2000), and they can affect gene silencing as more recently shown in the Arabidopsis *FLOWERING LOCUS C* (*FLC*) gene (Coustham *et al*., 2012). In the introns, they can affect splicing and cause null alleles, e.g. in the *FLC* gene (Michaels and Amasino, 1999; Gazzani *et al*., 2003; Michaels *et al*., 2003). In the exons, mutations can cause non-synonymous amino acid changes leading to truncated DNA sequences and loss of function alleles. They can result in change of function alleles that alter the protein function or stability. Several studies in Arabidopsis photoreceptor genes have shown that natural variation in hypocotyl length is due to amino acid changes. Mutations in the protein coding regions of *PHYA* (Maloof *et al*., 2001), *PHYB* (Filiault *et al*., 2009), *PHYC* (Balasubramanian *et al*., 2006), *PHYD* (Aukerman *et al*., 1997) and *CRY2* (El-Assal *et al*., 2001) affect natural variation in hypocotyl length. Because they can lead to phenotypic variation, non-synonymous mutations in the coding sequences of candidate alleles are a powerful means for explaining their potential role in the natural variation of hypocotyl length. Therefore, the protein coding regions were a reasonable place to look for interesting polymorphisms in the candidates.

6.2 Discussion and Conclusions on *PHYB* Candidate Gene:

Different MAGIC accessions have been experimentally studied which are not only the common laboratory genotypes such as Col-0 and Ler-0 and therefore, novel *PHYB* alleles have been identified that contribute to the natural variation. This supports the general notion that there is greater untapped genetic variation present in natural accessions which needs to be exploited to discover new alleles that play important roles in gene function pertaining to growth and development in response to temperature.

One of the advantages of exploiting natural variation, in comparison to classical induced mutagenesis, is that this approach allows the identification of variant alleles of genes which could be the likely targets for the evolution of hypocotyl length variation. This also helps in the identification of recessive mutations which are not deleterious in nature because nature has allowed the persistence of accessions carrying such mutations.

From the comparative analysis in chapter 4, it has been shown that Ct-1 and the *phyB* knockout mutant in Col-0 background have very similar phenotypes under both light and dark growth conditions. Hence, it seems that the Ct-1 allele of *PHYB* is a loss of function allele due to the natural mutations present in it. Furthermore, transgenic complementation results in chapter 5 showed very explicitly that the short WT hypocotyl phenotype was restored and re-established in single copy transgenic Ct-1 independently transformed lines carrying the Col-0 PHYB allele. Loss of function mutations are usually deleterious in nature and over time they are expected to be cleared from the gene pool. However, as the *PHYB* results in this study reveal, it may be the case that nature has maintained and fixed such null alleles in the population giving it a selective advantage in the wild under the native environmental conditions and improving the fitness of the species. Accessions having such alleles may germinate better in the field when elongated hypocotyls would allow the seedling to reach to light. Therefore, the mutations causing such alleles to be present in populations could be regarded as beneficial mutations. Loss of function or change of function mutations could be beneficial in the case of hypocotyl length which controls seedling plant architecture. The polymorphisms that have arisen in nature in *PHYB*, during the course of evolution, may be seen as a source of variation producing multiple alleles that have contributed to the natural variation of hypocotyl length in the Arabidopsis populations.

Is *PHYB* a genetic driver of adaptive variation in response to temperature with large scale consequences in Arabidopsis populations? It has been previously shown that several Arabidopsis photoreceptor genes are a major driving force of natural variation in plant hypocotyl growth responses to light; these include *PHYA* (Maloof *et al*., 2001), *PHYB* (Filiault *et al*., 2009), *PHYC* (Balasubramanian *et al*., 2006), *PHYD* (Aukerman *et al*., 1997) and *CRY2* (El-Assal *et al*., 2001). The results of this study strongly support this notion and provide a new dimension to the role of *PHYB* as an important QTL, and show that naturally existing alleles in *PHYB* are a major driving force of natural variation in hypocotyl length at 22°C. Furthermore, it is worth mentioning that photoreceptor and temperature interaction has been previously investigated and a close relationship between the two has been found in several studies, as shown in chapter 1. Consistent with this finding, this study demonstrates that the interaction of *PHYB* and ambient temperature control natural variation in hypocotyl length.

Large effect *PHYB* alleles for hypocotyl length in response to temperature were identified, in what is otherwise a typical quantitative trait. The transgenic phenotypic characterisation of the Temp22.2 fine-mapped QTL determines that the hypocotyl length trait variation in natural accessions is due to mutations at this single locus. *PHYB* individually explains a large proportion of the phenotypic variance at 22°C. It could, therefore, be hypothesised that QTL genes that have relatively large effects on phenotypic traits seem to be important in evolution. Important mutations in the QTL alleles are likely to explain a big portion of the natural intraspecific variation in Arabidopsis. This has been previously shown in the case of *FRIGIDA* (*FRI*) which is a major determinant of natural variation in Arabidopsis flowering time; though Arabidopsis accessions show quantitative variation in the vernalization responses, the *FRI* gene acts like a single gene trait (Sanda *et al*., 1997; Johanson *et al*., 2000).

The results of the transgenic complementation studies reveal allelic variation in *PHYB* resulting in phenotypic diversity of hypocotyl length in Arabidopsis. *PHYB* is a gene that is responsible for altering the quantitative aspect of plant growth and morphology as seen in the case of hypocotyl length. This project has provided insights into the genetic basis of functional trait evolution in Arabidopsis. It sheds some light on the extent to which a major phytochrome gene may be involved in the evolution of hypocotyl length trait differences within Arabidopsis.

The fact that very few QTL have been cloned is a manifestation of the long path from phenotype to genotype, despite the wealth of genetic information and resources available. However, the *PHYB* gene has been cloned that explains the Temp22.2 QTL revealing the molecular basis of quantitative variation. This study has identified molecular variants of *PHYB* accounting for natural variation in hypocotyl length and has contributed to the discovery of novel *PHYB* alleles. It is an example of how natural variation, as a resource for biology, has been exploited down to the molecular level in Arabidopsis to advance our knowledge of the genetic basis of seedling architecture.

The genetic analyses have provided a unique understanding of natural variation in the candidate gene *PHYB* from a functional perspective in response to temperature. Moreover, it has helped us to discover novel genetic elements that control hypocotyl

length in response to specific temperatures, e.g. in the case of *SAUR38* alleles that have been shown to control hypocotyl elongation at 22°C in chapter 5.

6.3 Ecological Importance and Adaptation:

Plants are capable of responding and adapting to changes in abiotic factors in nature due to the various sophisticated mechanisms that they have evolved over time. It is important to develop our understanding of the genetic basis of local adaptation due to its relevance to biodiversity and climate change. It is generally accepted that intraspecific natural variation improves the ability of plants to adapt to new changes in the environment as a consequence of climate change (Jump *et al*., 2009). If there is greater genetic diversity, there is greater ability to adapt to changing temperatures as a result of climate change. The considerable variation for temperature responses observed in MAGIC accessions corresponds to the adaptation to specific temperature conditions.

Though many plant populations are locally adapted to native environmental conditions, the genetic basis of their adaptation is not well known (Savolainen, 2011). From an evolutionary point of view, there is interest in identifying the genomic basis of adaptation to local climates and studies on plant-environment interactions in Arabidopsis can facilitate this approach.

Over the years, the analysis of natural variation in Arabidopsis has had great interest from both ecological and evolutionary points of view (Mitchell-Olds, 2001; Remington and Purugganan, 2003). What are the characteristics of a complex trait that determine whether it is an important ecological trait or not? This primarily depends on the trait in question and the role that it plays in the survival and fitness of plants in the wild. It is important to identify the genes that underlie ecologically important complex traits. Is hypocotyl length assumed to be an ecologically important trait for adaptation to various environments? The study of natural variation of hypocotyl length provides a strong evolving link, between the molecular analysis of gene function and ecological analysis of plant adaptation to local habitats. This could aid in providing new insights into the molecular mechanisms that shape complex trait variation in Arabidopsis natural populations. Is hypocotyl length an ecologically important trait for plant growth and

development in response to temperature? Laboratory based experiments cannot address this question entirely due to its limitations; however, as shown in this study, it can develop our basic understanding of hypocotyl responses to a range of ambient temperatures. The natural variation of the candidate genes in this study could provide clues to determine hypocotyl length as an important ecological complex trait. Thus, further experimentation and studies are required under field conditions to establish whether the polymorphisms identified in the candidate genes, under laboratory conditions in this study, are relevant for environmental adaptation. Growing Arabidopsis accessions in the native sites under field conditions could provide information about the adaptation of hypocotyl length to the local climates. Such analysis may be used in due course to predict and manage responses to climate change.

There are several ways of how this could be investigated. One way of studying this in the field could be to grow RILs with extreme hypocotyl phenotypes. Another way could be to conduct reciprocal transplant experiments of accessions that can verify the existence of natural variation in hypocotyl length in the native sites. Moreover, such experiments can confirm that the fitness of the native accession is higher than any introduced non-local accession. In such studies, the correlation of hypocotyl length with fitness can also be examined.

An interesting ecological question is whether differences in hypocotyl length have any correlation to the fitness of the accessions? It would be interesting to compare the hypocotyl lengths of the accessions with total number of seeds produced and determine the correlation. The results of a previous reciprocal transplant study provide evidence of strong adaptive differentiation between Arabidopsis natural populations in their native environments (Agren and Schemske, 2012). Together data from laboratory based experiments with these field based experiments would provide a strong basis for determining the ecological significance of hypocotyl length.

QTL cloning results from this study show that molecular variation in *PHYB* between the accessions controls hypocotyl length variation. What are the consequences of molecular variation in *PHYB* to fitness? Do these alleles have any selective advantage under specific temperature conditions? These questions are difficult to address, however, such a study could provide insights into the adaptive significance of hypocotyl length in the accessions.

It has been demonstrated that Arabidopsis natural accessions contain allelic variants that confer individual fitness to various environmental factors and climates (Fournier-Level *et al*., 2011; Hancock *et al*., 2011; Agren and Schemske, 2012; Horton *et al*., 2012). Therefore, the allelic variants of *PHYB* for hypocotyl length in response to temperature may have a potential to confer fitness in Arabidopsis to changing climates in natural habitats.

Is *PHYB* a key gene necessary for adapting to different climates? The *PHYB* alleles have been maintained in nature and it is likely that this is a photoreceptor that has changed in response to the environment, in different accessions from different geographical areas allowing them to adapt. In this regard, natural variation in *PHYB* function may play a role in allowing plants to adapt to unique future temperature environments. This could be a step forward towards understanding the genetic architecture and basis of hypocotyl natural variation involved in climatic adaptation, as the basis of quantitative traits associated with climatic adaptation remains unresolved (Aitken, 2008).

The phenotypic variation observed in hypocotyl length to temperature responses explained in chapter 2 seems to reflect adaptation to local climate and habitats from where they originally belong. This variation in wild Arabidopsis populations is thought to have been shaped by the driving force of natural selection. There may be an increasing interest to understand the adaptive nature of temperature responses in plants especially in growth and development related traits due to the effects of climate change. Has hypocotyl length adapted to local temperature conditions in the native habitats from where these accessions were collected? Has the function of the hypocotyl length improved (i.e. become bigger or shorter) in response to temperature in the native environments? The understanding of intraspecific phenotypic variation present in Arabidopsis hypocotyl length may provide clues to how these accessions adapt to their environment. Findings of this study show that ambient temperature is one such important abiotic environmental factor regulating plant growth and development. Arabidopsis hypocotyl length shows unique developmental plasticity to changing temperatures.

6.4 Application and Translational Biology:

Developing our understanding of the genetic basis of such important quantitative traits is important for enhancing basic knowledge of plant form and growth. To exploit the intraspecific natural variation, the discovery of allele variants that affect quantitative traits in Arabidopsis is essential. The investigation of variation that is present within a species in response to climate factors is a preliminary step, which is currently in its infancy and once a profound understanding of this has been developed, one can then look into comparing complex variation in plant growth between related species. Once the underpinning genetic components of the QTL variation are identified, our understanding of complex phenotypes could be improved by linking these genetic components to genomic databases of other related species. There is a great potential for comparative genomics between Arabidopsis genome and related crop genomes for candidate genes. In comparative genomics, the sequences of candidates can be compared to orthologues in crops and their functions may be predicted in response to environmental change.

The QTL that contribute to quantitative variation in natural populations may provide prospects for translational biology to other species. One such example is the *FLC* gene that regulates flowering time and confers a requirement for vernalization. This gene was initially identified in Arabidopsis but has been involved in controlling flowering time in related species such as *Brassica napus* (Tadege *et al.*, 2001), *Thellungiella halophila* (Fang *et al.*, 2006), *Brassica rapa* (Kim *et al.*, 2007), *Brassica oleracea* (Okazaki *et al.*, 2007); and unrelated species such as *Beta vulgaris* (Reeves *et al.*, 2007). This example demonstrates how genes discovered in Arabidopsis can be used for understanding variation in other species. In light of this, the QTL cloning results from this study in Arabidopsis could facilitate similar studies in other related lineages because orthologues of phytochromes with similar functions may be exploited for plant growth. The naturally occurring alleles of *PHYB* may also be used to better comprehend the control of plant growth and development.

There is considerable interest in identifying this natural variation to improve crop species and to dissect the genetic basis of adaptation. There are two potential ways of improving related crop species with useful candidate genes from Arabidopsis. Genes controlling plant morphology and development can be either directly transferred into crops or useful

orthologues can be identified by homology (Spannagl *et al*., 2011). How can the knowledge gained in Arabidopsis be translated to other species? Potentially, related species in the Brassica family and others such as rice could benefit (Rensink and Buell, 2004); as Arabidopsis and rice share a considerable number of orthologous genes (Izawa *et al*., 2003).

The empirical results in this study on the reference species Arabidopsis may provide resources or knowledge for translational biology methods to improve stem architecture leading to yield increases in related crop species. There is great potential in the future to apply and use the approaches described for the natural variation in Arabidopsis to other crop species. This will depend on the availability of genetic resources being developed for crop plants.

6.5 Future Prospects:

At the beginning of this century, it was put forward that the future should lie in obtaining a clearer understanding of QTL at the molecular level that would lead to a better understanding of the genes affecting important traits (Kearsey, 2000). In this study, the QTL mapping provided the first bridge between hypocotyl length variation and genes that control them in response to temperature. The motivation for examining natural variation in wild accessions is to find new alleles or new genes involved in the early development of the seedling in response to temperature, and to try to establish a foundation for understanding the molecular basis of adaptations to the local environment. Potentially, this study provides a basis for further analysis to exploit this variation in hypocotyl length to unravel the molecular basis of plant adaptation to different environments.

In general, much progress has been made in mapping QTL; however, finding the causal molecular polymorphisms also known as Quantitative Trait Nucleotide (QTN) remains a bottleneck. Understanding how natural variation in DNA allelic sequences is responsible for phenotypic variation in quantitative traits is a major challenge of contemporary biology (Mackay, 2014). The description of the genetic architecture of hypocotyl length will be more complete by specifying the polymorphic sites, such as QTN, in the alleles that cause differences in the hypocotyl length. This will reveal molecular genetic basis of the

putative adaptive trait of hypocotyl length in Arabidopsis. The goal would be to determine the QTN responsible for ecologically and evolutionarily relevant hypocotyl phenotypic diversity.

Analysis of naturally existing genetic variation comprises of two major challenging steps. These include the determination of the underlying genetic basis of variation, and finding the molecular nature of the allelic differences that are responsible for genetic differences leading to phenotypic variation (Reymond *et al.*, 2007). In this study, the first step in determining the molecular basis of natural variation in hypocotyl length was to determine the causal gene by cloning the QTL. Following QTL cloning, the next logical step is to test the identified functional polymorphisms in the *PHYB* alleles in the Temp22.2 QTL that explain natural variation in hypocotyl length in response to temperature. This will help in understanding in depth the patterns of phenotypic and molecular variation with an aim to infer how allelic variation is maintained. This may lead to clues for identifying which *PHYB* allelic variants could be adaptive under specific temperature conditions.

What are the mechanisms through which causal polymorphisms affect important growth phenotypes? To obtain a complete understanding of the molecular and genetic architecture of hypocotyl length, in a post QTL cloning strategy, the investigation of the molecular polymorphisms in *PHYB* that functionally define the Temp22.2 QTL will be required. This could be a key to understanding the evolutionary significance of Temp22.2 QTL and may provide valuable information based on which inferences could be made whether selection has operated on this QTL. Furthermore, the native habitats provide a wealth of environmental knowledge that can be utilised to hypothesise whether the allelic variation at Temp22.2 QTL has been adaptive. The findings in this study may also pave way for new insights into the networks of gene regulation for hypocotyl length in response to temperature.

The transgenic analysis of *PHYB* appreciates that natural mutations in its alleles can have widespread effects on hypocotyl length. Therefore, the functional analysis of natural variation in mutant alleles at a single locus of *PHYB* has shown to be an effective starting point to unravel molecular mechanisms that control the natural variation in the complex trait of hypocotyl length. This is because in addition to the immediate mechanistic understanding of how *PHYB* influences hypocotyl length, the nature of inherited variation

in hypocotyl length could provide clues to the evolution of the trait in response to temperature.

The QTL underlying the hypocotyl length sensitivity to temperature were detected by examining the G x E interaction. The interest was to identify QTL that contribute to the variation in temperature responsiveness. For the 'environmental QTL', identified for the 'temperature responsiveness trait' in chapter 3, functional characterisation of the important candidate gene, *PHYD*, could provide interesting results by cloning this thermal sensitive QTL. As a way forward, a transgenic allelic approach followed by phenotypic analysis of differential responses to temperature may well explain natural variation in the temperature responsiveness of the major effect MAGIC accessions. Identification of the thermal sensitive QTL, contributing to natural variation in responsiveness, is highly relevant for understanding plant-temperature interactions from the climate change point of view. The alleles controlling such loci could facilitate a population to adapt to changing climates (Via and Lande, 1985). The thermal sensitive loci identified from the G x E mapping may provide plasticity in hypocotyl length which could help Arabidopsis to adapt to warmer climates. Arabidopsis populations that possess allelic variation at such loci could be able to adapt better than those that lack it.

With a thorough understanding of the candidate gene functions, we will be able to profoundly broaden our understanding of plant-temperature interactions and the important genetic elements controlling their natural variation. The field of evolutionary and developmental biology is expanding and by comparing the alleles responsible for variation in hypocotyl length, it may be possible to understand how this developmental trait became modified during evolution in response to abiotic environmental factors. With the help of such information, future plant adaptation to warmer climates may be envisaged.

This natural variation study in response to temperature has identified new QTL candidate genes and novel alleles of known candidate genes, controlling hypocotyl length, that have evolved in specific genetic backgrounds under specific local environmental conditions. This study and future related research can potentially give important insights into plant growth responses to increased temperature connected to global climate change.

6.6 References:

Agren J, Schemske D. (2012) Reciprocal Transplants Demonstrate Strong Adaptive Differentiation of the Model Organism *Arabidopsis thaliana* in its native range. New Phytol. 194, 4, 1112-22.

Aitken S, Yeaman S, Holliday J, Wang T, Curtis-McLane S. (2008) Adaptation, Migration or Extirpation: Climate Change outcomes for Tree Populations. Evolutionary Applications 1, 95–111.

Aukerman M, Hirschfeld M, Wester L, Weaver M, Clack T, *et al*. (1997) A Deletion in the *PHYD* Gene of the Arabidopsis Wassilewskija Ecotype Defines a Role for Phytochrome D in Red/Far-Red Light Sensing. Plant Cell 9, 1317–1326.

Balasubramanian S, Sureshkumar S, Agrawal M, Michael T, Wessinger C, *et al*. (2006) The PHYTOCHROME C Photoreceptor Gene Mediates Natural Variation in Flowering and Growth Responses of *Arabidopsis thaliana*. Nat Genet 38, 711–715.

Borevitz J, Maloof J, Lutes J, Dabi T, Redfern J, *et al*. (2002) Quantitative Trait Loci Controlling Light and Hormone Response in two Accessions of *Arabidopsis thaliana*. Genetics, 160, 683–96.

Coustham V, Li P, Strange A, Lister C, Song J, *et al*. (2012) Quantitative Modulation of Polycomb Silencing Underlies Natural Variation in Vernalization. Science, 337, 584-587.

El-Assal S, Alonso-Blanco C, Peeters A, Raz V, Koornneef M. (2001) A QTL for Flowering Time in Arabidopsis Reveals a Novel Allele of *CRY2*. Nat Genet, 29, 435–440.

Hancock A, Brachi B, Faure N, Horton M, Jarymowycz L, *et al*. (2011) Adaptation to Climate across the *Arabidopsis thaliana* Genome. Science, 334, 83–86.

Horton M, Hancock A, Huang Y, Toomajian C, Atwell S, *et al*. (2012) Genome-Wide Patterns of Genetic Variation in Worldwide *Arabidopsis thaliana* Accessions from the RegMap Panel. Nat Genet, 44, 212–216.

Fang Q, Xu Z, Song R. (2006) Cloning, Characterization and Genetic Engineering of *FLC* Homolog in *Thellungiella halophila*. Biochemical and Biophysical Research Communications, 347, 707–714.

Filiault D, Wessinger C, Dinneny J, Lutes J, Borevitz J, *et al*. (2008) Amino Acid Polymorphisms in Arabidopsis Phytochrome B cause Differential Responses to Light. Proc Natl Acad Sci USA, 105, 3157–3162.

Fournier-Level A, Korte A, Cooper M, Nordborg M, Schmitt J, *et al*. (2011) A Map of Local Adaptation in *Arabidopsis thaliana*. Science, 334, 86–89.

Frary A, Nesbitt T, Grandillo S, Knaap E, Cong B, *et al*. (2000) fw2.2: A Quantitative Trait Locus Key to the Evolution of Tomato Fruit Size. Science, 289, 5476, 85-8.

Gazzani S, Gendall A, Lister C, Dean C. (2003) Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions. Plant Physiol. 132, 1107–1114.

IPCC, Climate Change 2013 (www.ipcc.ch).

Izawa T, Takahashi Y, Yano M. (2003) Comparative Biology comes into Bloom: Genomics and Genetic Comparison of Flowering Pathways in Rice and Arabidopsis. Curr. Opin. Plant Biol. 6, 113–20.

Johanson U, West J, Lister C, Michaels S, Amasino R, *et al*. (2000), Molecular Analysis of FRIGIDA, a Major Determinant of Natural Variation in Arabidopsis Flowering Time. Science, 290, 5490, 344-347.

Jump A, Marchant R, Peñuelas J. (2009) Environmental Change and the Option Value of Genetic Diversity. Trends Plant Sci., 14, 51–58.

Kearsey M. (2000) Quantitative Genetics and Breeding Methods: The Way Ahead. Paris (France), August 30-31-September 1, 2000, Ed. INRA Paris, 2001 in Eucarpia. Biometrics in Plant Breeding. Meeting, André Gallais).

Kim S, Park B, Kwon S, *et al.* (2007) Delayed Flowering Time in Arabidopsis and *Brassica rapa* by the Overexpression of *FLOWERING LOCUS C (FLC)* Homologs isolated from Chinese Cabbage (*Brassica rapa* L. Ssp pekinensis). Plant Cell Reports, 26, 327–336.

Mackay T. (2014) Epistasis and Quantitative Traits: Using Model Organisms to Study Gene–Gene Interactions. Nature Reviews Genetics, 15, 22–33.

Maloof J, Borevitz J, Dabi T, Lutes J, Nehring R, *et al*. (2001) Natural Variation in Light Sensitivity of Arabidopsis. Nat. Genet. 29, 441–46.

Michaels S, Amasino R. (1999) *FLOWERING LOCUS C* Encodes a Novel MADS Domain Protein that acts as a Repressor of Flowering. The Plant Cell, 11, 949–956.

Michaels S, He Y, Scortecci K, Amasino R. (2003) Attenuation of *FLOWERING LOCUS C* Activity as a Mechanism for the Evolution of Summer-Annual Flowering Behavior in Arabidopsis. PNAS USA, 100, 10102–10107.

Mitchell-Olds T. (2001). *Arabidopsis thaliana* and its Wild Relatives: A Model System for Ecology and Evolution. Trends Ecol. Evol. 16, 693–700.

Nordborg M, Hu T, Ishino Y, *et al*., (2005) The Pattern of Polymorphism in *Arabidopsis thaliana.* PLoS Biol, 3, e,196.

Okazaki K, Sakamoto K, Kikuchi R *et al.* (2007) Mapping and Characterization of *FLC* Homologs and QTL Analysis of Flowering Time in *Brassica oleracea*. Theoretical and Applied Genetics, 114, 595–608.

Parmesan C. (2006) Ecological and Evolutionary Responses to Recent Climate Change. Annual Review of Ecology Evolution and Systematics. 6, 37, 637–669.

Reeves P, He Y, Schmitz R, Amasino R, Panella L, *et al*. (2007) Evolutionary Conservation of the *FLOWERING LOCUS C*-Mediated Vernalization Response: Evidence from the Sugar Beet (*Beta vulgaris*). Genetics, 176, 295–307.

Remington D, Purugganan M. (2003) Candidate Genes, Quantitative Trait Loci, and Functional Trait Evolution in Plants, Int. J. Plant Sci. 164.

Rensink W and Buell C. (2004) Arabidopsis to Rice. Applying Knowledge from a Weed to enhance our Understanding of a Crop Species. Plant Physiology, 135, 622–629.

Reymond M, Pieper B, Barbier H, Ihnatowicz A, El-Lithy M, *et al*. (2007) Genetic and Molecular Analysis of Growth Responses to Environmental Factors using *Arabidopsis*

thaliana Natural Variation. In: Volume 21 Scale and Complexity in Plant Systems Research: Gene-Plant-Crop Relations, 2007, X, 332 p., ISBN: 978-1-4020-5905-6.

Sanda S, John M, Amasino R. (1997) Analysis of Flowering Time in Ecotypes of *Arabidopsis thaliana*. J Hered, 88, 1, 69-72.

Savolainen O. (2011) The Genomic Basis of Local Climatic Adaptation. Science, 49-50.

Schmid K, Sorensen K, Stracke R, Torjek O, Altmann T, *et al*. (2003) Large-Scale Identification and Analysis of Genome-Wide Single-Nucleotide Polymorphisms for Mapping in *Arabidopsis thaliana*. Genome Res, 13, 1250–1257.

Spannagl M, Mayer K, Durner J, Haberer G, Fröhlich A. (2011) Exploring the Genomes: From Arabidopsis to Crops. J Plant Physiol, 168, 3–8.

Springate D, Kover P. (2014) Plant Responses to Elevated Temperatures: A Field Study on Phenological Sensitivity and Fitness Responses to Simulated Climate Warming. Global Change Biology, 20, 2, 456-465.

Tadege M, Sheldon C, Helliwell C, Stoutjesdijk P, Dennis E, *et al*. (2001) Control of Flowering Time by FLC Orthologues in *Brassica napus*. Plant Journal, 28, 545–553.

Via S, Lande R. (1985) Genotype-Environment Interaction and the Evolution of Phenotypic Plasticity, Evolution 39, 505–522.

Weigel D, Nordborg M. (2005) Natural Variation in Arabidopsis. How Do We Find the Causal Genes? Plant Physiol. 138, 567-568.