5-2010

THE GENETICS OF CHILLING REQUIREMENT IN APRICOT (*PRUNUS ARMENIACA* L.)

Bode Olukolu
Clemson University, bolukol@clemson.edu

THE GENETICS OF CHILLING REQUIREMENT IN APRICOT (*PRUNUS ARMENIACA* L.)

A Dissertation
Presented to
The Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Genetics

---

by
Bode Adebowale Olukolu
May 2010

Accepted by:
Dr. Albert G. Abbott, Committee Chair
Dr. Julia Frugoli
Dr. Chin-Fu Chen
Dr. Amy Lawton-Rauh

ABSTRACT

Commercial production of apricot is severely affected by sensitivity to climatic conditions, an adaptive feature essential for cycling between vegetative or floral growth and dormancy. Yield losses are due to either late winter or early spring frosts or inhibited vegetative or floral growth caused by unfulfilled chilling requirement (CR). Studies in this dissertation developed the first high-density apricot linkage map; followed by a comparative mapping strategy to validate conservation of synteny, genome collinearity and stable quantitative trait loci (QTLs) controlling CR and bud break between apricot and peach; and ultimately attempt to identify key candidate genes following a linkage disequilibrium-based association mapping approach to fine map the major CR QTL genomic regions.

Following a two-way pseudotestcross mapping strategy, two high-density apricot maps were constructed using a total of 43 SSR (Simple Sequence Repeats) and 994 AFLP (Amplified Fragment Length Polymorphism) markers that span an average of 502.6 cM with an average marker interval of 0.81 cM. Twelve putative CR QTLs were detected using composite interval mapping, a simultaneous multiple regression fit and an additive-by-additive epistatic interaction model without dominance. An average of 62.3% ± 6.3% of the total phenotypic variance was explained. We report QTLs corresponding to map positions of differentially expressed transcripts and suggest candidate genes controlling CR.

A majority of the QTLs were shown to be stable between both *Prunus* species, as well as similar trends in their QTL effects, with the allele for increasing the trait value mostly originating from the high chill parents. The denser apricot maps, due to more AFLP marker polymorphisms, provide a higher resolution to delineate QTLs to smaller genomic intervals, as well as splitting each of some of the peach QTLs into two. The comparative QTL mapping strategy presented here reveals the transferability of genetic information between two *Prunus* species, the characterization of stable QTLs, the utility of the maps to consolidate each other and to further validate previously identified CR QTL loci as a major controlling factor driving floral bud break.

The LD-based association mapping was limited to marker dense genomic regions within and around previously detected major QTLs on linkage group (LG) 1 and 7. LD decayed below the centimorgan scale, indicating insufficient marker density averaged at 0.44 and 1.58 cM on LG1 and 7, respectively. Denser marker regions averaged at 0.1 and 0.7 cM on LG1 and 7, respectively, revealed significant LD estimates above the baseline threshold. We report significant marker-trait associations and underlying genes the markers were derived from. Our results demonstrate that an LD-based association mapping can be used for validating QTLs, fine mapping and detecting CGs in *Prunus*.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Albert Abbott, for giving me the opportunity to work with him and for always been there as a mentor and guide. I also extend my gratitude to my committee members: Dr. Julia Frugoli, Dr. Chin-Fu Chen and Dr. Amy Lawton-Rauh, for all their support and advice throughout the program.

Thanks to Dr. Doron Holland and Taly Trainin for providing, maintaining and scoring the phentoypes on the apricot mapping population, Dr. William Okie for providing the peach mapping population, Dr. Gregory L. Reighard and Dr. Douglas G. Bielenberg for maintaining the peach mapping population an Dr. Valentina Gorina for maintaining and scoring the phenotypic data on the apricot germplasm used for the association mapping.

Special thanks to members of the Abbott's lab. Thanks to Dr. Tatyana N. Zhebentyayeva, for establishing the collaboration on the association mapping, providing apricot germplasm DNA samples, designing SSR markers and for her help on several occasions as a senior colleague and friend. Thanks to Dr. Laura L. Georgi, who made life in the lab comfortable and who was always available as a resource person. Thanks to Dr. Renate Horn for all her unrelenting help, Dr. Chittaranjan Kole whose help I can't list exhaustively and to graduate students who showed me the ropes: Renea Hardwick, Fan Shenghua and Samuel Forrest. I would also like to thank Sherri Hughes-Murphree, Megan Mosanto and Phullara Kole for their support in the lab and Jeanice Troutman of CUGI for her help with DNA sequencing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

LITERATURE REVIEW

Introduction

An understanding of complex triggers and molecular players regulating induction, maintenance and release of dormancy in plants is crucial to resolving problems associated with crop production. Although mostly motivated by the economic importance of crops, it has been the subject of several research studies concurrently leading to numerous fundamental findings in plant science. For over a decade now, the rate of yield increase for staple food crops has been dwindling and causing widespread concern that we might be approaching the sustainable yield barriers once overcome by the green revolution (Huang *et al.* 2002, Jung and Müller 2009). The need for new technologies is not only necessitated by this challenge for food security, but it is also necessitated by rising and competing demands for plant biomass as a source of renewable energy. For both generative and vegetative crops, the phenological developmental processes are critical to increasing crop yield. This is evident from effects of photoperiod (plant response to short or long day lengths) and vernalization (exposure to cold as a requirement for flowering) on timing and vigor of flowering. A recent study demonstrates that genes linked to the control of circadian-mediated physiological and metabolic pathways have a major influence on growth vigor and accumulation of plant biomass (Ni *et al.* 2009). This is of great significance to productivity as reflected in seed crops where floral transition is a key developmental switch that determines dry matter yield, in vegetative crops like fodder

grasses where early bolting limits potentials for high yields. This problem is mirrored in perennial fruit trees where fruit production is vulnerable to inconsistent late onset of flowering and loss of flowers and/or immature fruits to spring frost.

Although this review centers on dormancy in floral and vegetative buds of perennial fruit trees, other studies on dormancy among several other plant taxa will also be mentioned as they relate to perennial fruit trees. Furthermore, after about a century of research on dormancy, the field of study has branched into studies of the different dormancy aspects, such as site of dormancy (seeds, floral and vegetative buds, tubers, bulbs and stolons; Wareing 1969), photoperiod and other environmental effects causing the induction of dormancy (Wareing 1956), differences in vegetative versus floral bud dormancy (Romberger 1963), modification of CR by environmental factors and cultural practices (Nooden and Weber 1978, Lang 1987), dormancy breaking chemicals and/or stress treatments, chilling requirement (CR) for dormancy release and its effective temperature range and models for calculating CR (Doorenbos 1953, Samish 1954, Vegis 1964, Leike 1965, Perry 1971, Erez and Lavee 1974, Saunders 1978, Saure 1985, Lang 1994, Champagnat 1989, Rowland and Arora 1997, Arora *et al.* 2003, Horvath *et al.* 2003).

Recent advances absent in previous reviews will be highlighted to keep pace with the developments made in this discipline. These will include: mechanisms underlying bud dormancy from induction to release; gene pathways and signals; cell-to-cell crosstalk; physiological delineation of different stages of dormancy; separation of dormancy from other related biological processes like freezing and dehydration tolerance, hormonal

physiology and the genetics of dormancy in woody plants (including identification of associated quantitative trait loci, mapping of dormancy-related genes and gene action of dormancy-related genes and expression profiles of these genes. Dormancy, an adaptive feature, is required to synchronize development with the cyclic climatic conditions and involves a gradual and progressive process until state of rest is reached (Hill *et al* 1998, Lang *et al* 1987). This transition into dormancy is triggered by environmental cues and has been delineated (Fig. 1.1) into 3 stages (induction, maintenance and release), as well as component pathways and processes Arora *et al*. 2003).

Regulation of growth cycles and dormancy in woody perennials

While evidence suggests that angiosperms originated in humid tropical climates where temperature, day length and availability of water were fairly stable all year round, one of the key evolutionary forces differentiating plant species and ultimately temperate species was environmental change (Okubo 2000). This has a profound impact on plant growth habit and life cycle. In order to synchronize timing of flowering with ambient temperatures that are optimal for fertilization and seed/fruit development, perception and transduction pathways (vernalization) that sense prolonged cold winter temperatures evolved that translate environmental cues into increased competence for flowering in spring or summer.

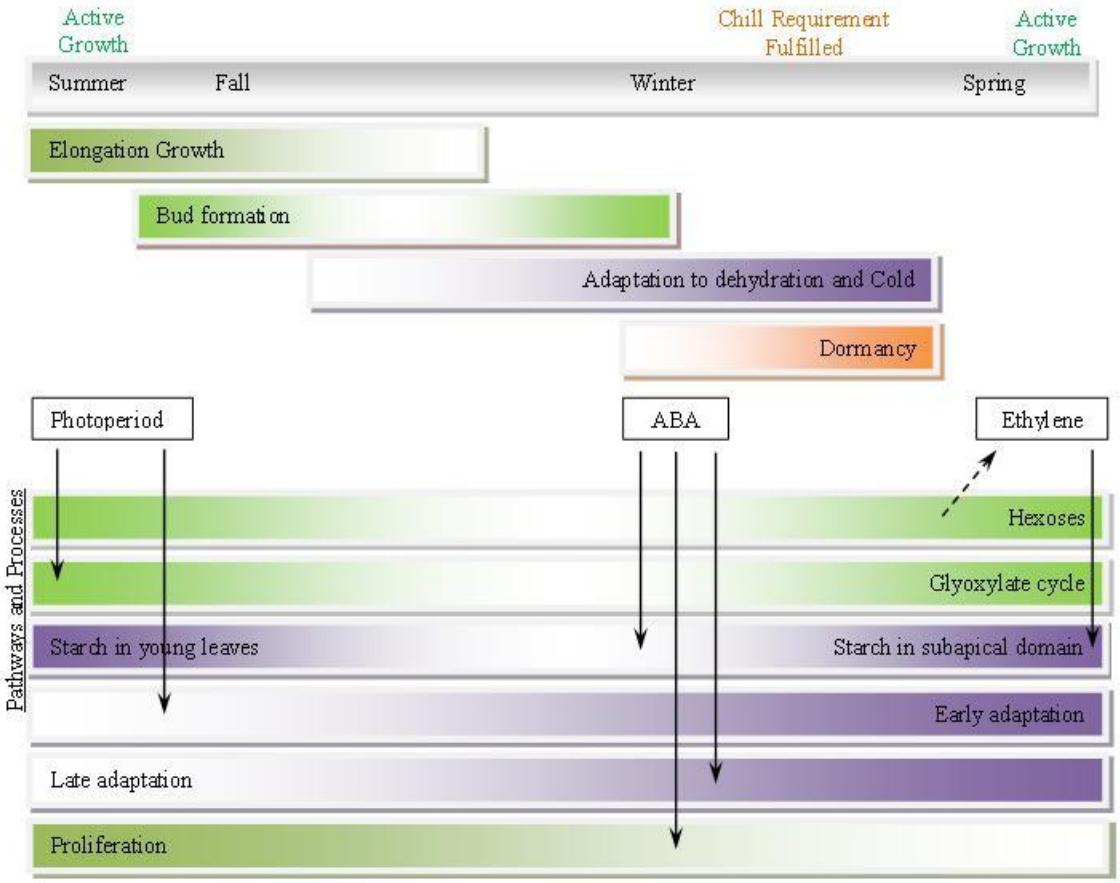


Figure 1.1: Timetable of bud development showing integration of environmental cues and endogenous processes underlying dormancy induction, maintenance and release in woody perennials.

Even though early studies showed that shortened growth period of shoots caused by water stress promoted the early induction of bud dormancy leading to reduced CR, there still remains a poor understanding of the molecular processes involved (Muller-Thurgau 1885, Arora *et al.* 2003). This correlation of the length of shoot growth period and the timing of bud break was confirmed by work of Chandler and Tufts (1934) that showed that an extended growth period of shoots delayed bud break during the following spring when chill accumulation was not sufficient.

Plants initially undergo a juvenility period of vegetative development prior to floral bud induction, and in woody perennials, this vegetative/juvenile stage can last for several years before the switch to a flowering developmental state. In this regard, woody perennials can exhibit significant variation in life cycle with reference to the transition to flowering. Raspberries, having a biennial or perennial growth habit, mostly produce juvenile vegetative primocanes in the first year and adult fruiting laterals in the following year(s). Some varieties exhibit the primocane fruiting phenotype that is a desired growth habit because it allows for some berry production in the first year, although yields and fruit size are low (Keep 1988). *Prunus* species on the other hand don't flower and fruit until two to three years of a juvenile phase is fulfilled. Bernier and Périlleux (2005) provide an extensive review of the major factors that influence this flowering habit. Even though woody perennials require a juvenility period, the size of the plant rather than its age has been confirmed to be specifically more important (Lacey 1986). In nature, some plants that don't flower until the third to fifth vegetative phase are known to flower during the second year under cultivation in resource rich conditions (Lacey 1986,

Klinkhamer *et al.* 1987). Thus, it seems the best predictor of flowering time in perennials is a threshold size that varies across different species and ecotypes (Lacey 1986, Wesselingh *et al.* 1993). Size in turn is directly related to the amount of resources accumulated, which in turn depends on ambient temperature, irradiance, water/mineral availability and presence/absence of resource competing neighboring plants (Bernier and Périlleux 2005).

Considering the difficulty of identifying QTLs in perennials that co-segregate with mapped photo-receptor genes (Frewen *et al.* 2000), there is speculation that there are other molecular players downstream of photo-receptors that regulate flowering and dormancy by transducing the light signal. Perhaps what the plant measures during the vegetative phase of development or before onset of flowering is biomass accumulation, which is a function of light, rather than light signal itself. From analysis of the phloem sap exported by leaves in response to floral induction, Bernier and Périlleux (2005) postulated that sucrose and cytokinin are potential long-distance signaling molecules. The increased export of sucrose in *Arabidopsis* in response to long-day induction might be partially due to increased efficiency of sucrose loading (Corbesier *et al.* 1998). After loading sucrose into the shoot apical meristem, a number of cellular and molecular events are initiated (Bernier 1988) as well as the hydrolysis of sucrose by local invertases i.e. vacuolar (Koch 2004) and cell wall (Heyer *et al.* 2004) invertases. Cytokinins activate invertase and increase the rate of cell division, while hexoses are known to participate with Giberellic acids in the up regulation of *LEAFY* (*LFY*) gene expression (Bernier and Périlleux 2005), see below.

In *Arabidopsis*, the flowering response to environmental cues involves several signaling pathways but they all converge towards the regulation of floral meristem identity genes (Mouradov *et al.* 2002). Downstream of this convergence are the *LEAFY* (*LFY*) and *APETALA 1*(*AP1*) genes that control flower morphogenesis. Genes acting upstream of this are the considered integrator genes and mutants of these show delay in flowering under different growing conditions. The integrator genes include *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1*). The primary environmental factors that initiate the pathways include photoperiod and temperature (Martínez-Zapater *et al.* 1994). Mutants of genes that caused late flowering and flowering delay in long-days were termed genes in the "long day (LD) pathway", while mutants of genes in plants responsive to photoperiod but impaired in their response to cold were considered genes in the "vernalization pathway" (Fig. 1.2). Mutants that were sensitive to both photoperiod and cold temperatures were classified as part of the "autonomous flowering pathway" (Bernier and Périlleux 2005).

The response to vernalization is facilitated by a cascade of gene regulatory networks, that are initiated during prolonged cold exposure by the induction and up-regulation of the homeodomain finger gene *VIN3* (*VERNALIZATION INSENSITIVE 3*) and results in the chromatin-based and mitotically stable repression of the *FLC* (*FLOWERING LOCUS C*) gene (Sung *et al.* 2004). *FLC*, a MADS-box transcription factor, in turn acts as a repressor of floral transition. In the following generation, *FLC* expression is reset around the time of early embryogenesis (Sheldon *et al.* 2008, Choi *et al.* 2009), thus ensuring a renewed requirement for vernalization.

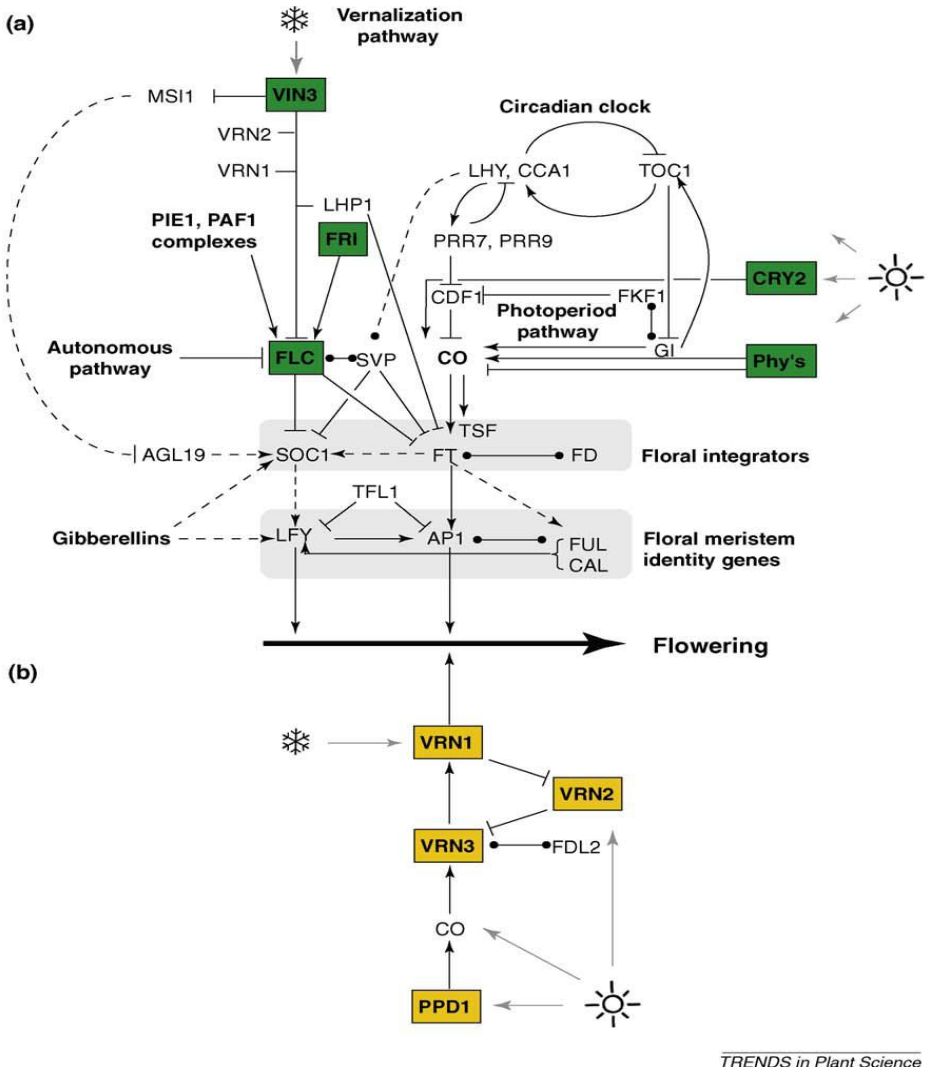


Figure 1.2: Flowering time control in *Arabidopsis* (a) and cereals (b). Exogenous cold (❄) and light (☼) signals are indicated by symbols. Positive and negative regulatory actions are indicated by arrows and lines with bars, respectively. Dashed lines designate more speculative interactions. The dashed line with a single filled circle at the end indicates a regulatory but yet little understood effect of *LHY* and *CCA1* on *SVP* protein accumulation (Fujiwara *et al.* 2008). Lines with filled circles at either end indicate protein–protein interactions. The green and yellow boxes designate genes shown to affect natural variation in flowering time in *Arabidopsis* and cereal accessions, respectively. The figure incorporates aspects from various previously published models (He and Amasino 2005, Trevaskis *et al.* 2007, Alonso-Blanco *et al* 2009, Distelfeld *et al.* 2009) (Jung and Müler 2009).

Although the *FLC* orthologs in other Brassica species are functionally related to the *Arabidopsis FLC* (Tadege *et al.* 2001, Kim *et al.* 2007), the extent of conservation outside the Brassicaceae family is still contentious (Jung and Muller 2009). Expressed sequence tags of the gene in rosids, asterids and caryophyllids have been identified (Reeves *et al.* 2007) but proof for the functional conservation remains inadequate (Jung and Muller 2009). Additionally, in temperate cereals, identification of key regulators of vernalization requirement and response (wheat *VRN1*, *VRN2* and *VRN3*, which are homologs of *Arabidopsis VRN* gene) does not include an *FLC-like* gene and reveal a regulatory pathway whose components differ from the *FLC*-dependent vernalization pathway.

Currently, regulatory mechanisms underlying floral induction in perennial plants remain poorly characterized, although attempts are being made to test pathways (Figure 1.2) already characterized in model plants like *Arabidopsis*. Floral induction in woody perennials differ from that of annual and biennials plants in that they comprise a morphogenetic transition of cells in apical meristems as well as in lateral meristems. In perennials, above ground meristems are not induced by strong floral promoters and therefore remain vegetative, thus guaranteeing a long life span. Little is known about how perennials achieve this but silencing of genes via DNA methylation, inaccessibility of floral promoters, and RNAi are potential players in this process. Additionally, repression floral specification may be achieved by the long time transcription during vegetative/juvenile phase of floral-repressing genes like *FLC* or similar homologs in perennials (Chen and Coleman 2006, Bangerth 2009).

Complexity of bud dormancy and its overlap with related biological processes

Despite extensive general progress made in understanding dormancy, knowledge gaps are still prevalent at all stages of the dormancy process. This is due to the complexity and nature of the molecular pathways that also overlap with other distinct biological processes, some of which are difficult to dissociate from dormancy. An example of such biological processes includes freezing and dehydration tolerance and the complexity of distinguishing cause and effects between them and dormancy. The capability of temperate perennials to survive freezing winter temperatures depends heavily on their adaptation, which involves mechanisms for transitioning into a dormant state, as well as the acquisition of cold hardiness, a measure of freezing and dehydration (Powell 1987). The same environmental cues (photoperiods and colder temperatures) that cause a shift from summer dormancy or correlative dormancy (paradormancy) to winter dormancy (endodormancy) concurrently induces cold acclimatization (ability of temperate plants to tolerate and survive freezing temperatures), while plant tissues become more cold-hardy during winter dormancy (Nissila and Fuchigami 1978). Consequently, the induction and release from dormancy in the annual growth cycles of woody perennials is superimposed on the acquisition and loss of cold hardiness, respectively (Fuchigami *et al.* 1982). To resolve the physiological and molecular events associated with the regulation of bud dormancy and that of cold hardiness, several strategies have been employed to study them independently of each other.

The first effort at resolving dormancy and related traits involved the use of genetically related peach (*Prunus persica*) genotypes that segregated for deciduous (completely lose their foliage at fall and regrowing it at spring) and evergreen (having leaves all year round) habits. The lack of endodormancy in one genotype and cold hardiness in both facilitated their use for a comparative study of changes in protein content as it relates to seasonal changes and the degree of cold hardiness (Arora *et al.* 1992, Arora and Wisniewski 1994, Artlip *et al.* 1997). Another species, *Vitis labruscana*, was originally explored by Fennell and Hoover (1991) since it was able to completely transition into an endodormant state in response to short photoperiod but without cold acclimation. Salzman *et al* (1996) eventually utilized this species to characterize differential expression of proteins in buds exposed to only short photoperiods. A study in blueberry (*Vaccinium* section *Cyanococcus*) cultivars attempted to resolve the problem by observing changes in bud proteins specifically associated with dehardening based on the premise that only temperatures between 0^{o} and 7^{o} C are effective towards contributing to chilling unit accumulation (Erez *et al.* 1979, Erez and Couvillon 1987). Cold acclimated buds (50% CR acquired) were exposed to controlled-temperature regimes warm enough to cause dehardening (reversal of the process of cold acclimation or hardening) without negating accumulated chill units of cold weather (i.e. duration of chilling temperatures measured as a requirement for bud break) or releasing them from winter dormancy (i.e. not affecting the dormancy status of the buds, Arora *et al.* 1997). Based on the studies mentioned above that show differential response as relates to dormancy and cold hardiness, there was a consensus that the metabolism of certain dehydrins, a subgroup of

late embryogenesis abundant (LEA) proteins referred to as the D-11 family (Close 1997), was more closely associated with cold hardiness transitions than bud dormancy (Rowland and Arora 1997, Rowland *et al.* 2003).

Even though dehydrins are ubiquitous hydrophilic proteins considered to guard cells against cellular dehydration (in this case, freeze-induced dessication) and are therefore expected to build up in cold hardened tissues, Faust *et al.* (1997) postulated that they might not be exclusive to cold hardiness but also involved in bud endodormancy. This was based on MRI studies where the bound to free water ratio of buds increased during late fall or early winter (Faust *et al.* 1991), hence, it was proposed that dehydrins bind water after being induced by low temperatures and abscisic acid (ABA), leading to freeze protection and a simultaneous deepening of dormancy (Faust *et al*., 1997).

Others have attempted to dissociate bud dormancy and cold hardiness by exploring the effects of endogenous ABA levels in each process. Numerous studies have implicated ABA as a stress-inducible hormone and growth inhibitor, as well as a mediator of short-day-induced growth cessation and dormancy induction in buds (Lenton *et al.* 1972, Iwasaki and Weaver 1977, Dumbroff *et al.* 1979, Barros and Neill 1989). Manipulating the endogenous ABA content of buds and using an ABA-deficient mutant of birch (*Betula pubescens*), the involvement of ABA in dormancy induction was examined (Welling *et al.* 1997, Rinne *et al.* 1998). Wild-type plants expressed elevated levels of ABA before onset of cold acclimation under short-day regimes, followed by tissue desiccation and accumulation of dehydrin proteins, while the ABA-deficient mutant had lower water loss, lower tolerance to low-temperature stress and lacked accumulation of

dehydrins. Nevertheless, the mutant was still capable of entering dormancy (Rinne *et al.* 1998), suggesting that ABA was not necessary for dormancy induction or that there were other pathways that augment the ABA-induced dormancy response. Increasing ABA content by spraying ABA on long-day exposed and water stressed plants also led to an increase in cold hardiness in the absence of dormancy induction in the wild type (Welling *et al.* 1997). These findings support ABA as being more directly involved with photoperiodic control of cold acclimation rather than in bud dormancy induction; although the influence of ABA in other developmental timed aspects of dormancy (maintenance and release) were not investigated.

Inferences made from ABA experiments are difficult to validate when one considers that the promotion of flowering by primary factors (day length and vernalization) can be delayed or even eliminated by other less predictable factors. This has been demonstrated in studies where flowering was suppressed in favorable photoperiodic conditions by water stress in both long-day (*Lolium temulentum*) and short-day plants (*Xanthium strumarium* and *Pharbitis nil*) or by excess nitrogen input (Bernier and Périlleux 2005). In certain instances, these primary factors can be conditional like in *Calceolaria* which requires vernalization at low irradiance even though vernalization is not required at high irradiance (Bernier 1988). These interactions corroborate the fact that plants are sessile opportunists that must optimize the timing of the commitment to flowering in an environment that displays significant annually fluctuating physical conditions.

Dormancy induction

Prior to growth cessation during winter, paradormancy (apical dominance or correlative inhibition) marks the first stage towards bud dormancy that allows for the plant to: allocate resources for reproduction, control plant architecture, and maximize light harvesting while allowing for regeneration should individual shoots become damaged. Historically, hormones were proposed as the major culprits for the induction of bud dormancy and were implicated as transducers of environmental cues (Hermberg 1949). In fact, the term dormin was proposed as a label for endogenous dormancy inducers (Eagles and Wareing 1963).

Although it's quite alluring to think of dormancy on the basis of hormonal control alone, dormancy is controlled by several integrated plant structures and functions; and even its path is a continuum that begins as early as bud break in spring (Simpson 1990, Crabbe 1994). ABA has been implicated in both short-day and water stress-induced dormancy in *Betula pubescens* (Rinne *et al* 1994a, 1994b, Welling *et al.* 1997) and *Vitis vinifera* 'Merlot Noir' (Koussa *et al.* 1998) where evidence supports a relationship between ABA and bud water content. Additionally, endogenous ABA levels appear to relate to the depth of bud dormancy (Tamura *et al.* 1993). Faust *et al.* (1991) demonstrated that endodormant buds have less free water than ecodormant buds, implying that CR satisfaction is related to the conversion of water from a bound to a free state. *Viccinium* cultivars with the deepest dormancy and highest CR reportedly possess the most bound water (Parmentier *et al.* 1998), while bound water is also shown to increase in

endodormant and freeze tolerant peach buds in response to induction by either photoperiod or cold temperatures (Erez *et al.* 1998). Although the studies above concluded that bound water status was associated with cold temperature stress tolerance rather than directly to dormancy itself, Fennell *et al.* (1996) revealed an increasing amount of bound water after 2 weeks of short-day photoperiod exposure in *Vitis riparia*. Similarly, Fernell and Line (2001) demonstrated an increasing amount of bound water with endodormancy in both grape buds and the cortex/gap tissue adjacent to the bud. Thus, increased bound water and the endodormancy state are potentially more directly connected.

Several studies initially monitored endogenous levels of hormones in whole buds, leaves, stems, cambium and root tissue under fall and dormancy-inducing controlled-environment conditions (Samish 1954, Wareing 1956, Nitsch 1957, Phillips and Wareing 1958, Dennis and Edgerton 1961) but the interpretation of the experimental results (measuring responses and application of hormones) suggested several pitfalls. These problems include: degradation and differential responses between commercial (±)-ABA and natural (+)-ABA (Wilen *et al.* 1996); reduced root uptake of ABA by casparian strip formation in the hypodermis (Freundl *et al.* 2000); loss of ABA to the medium when it is more alkaline than the root cortex; and finally, the pH of root zone and ABA concentration may modify root-to-shoot signaling as they affect apoplastic transport of ABA (Arora *et al* 2003). Strauss *et al.* (2001) also demonstrated experimentally that exogenously applied ABA was distributed differently from compartmentalized endogenous ABA within the cell. Proteins and other molecules that bind and/or modify

ABA might exist in the cytosol and/or endoplasmic reticulum and prevent ABA distribution based on a cellular pH gradient alone.

The problems in the studies of hormone action are further complicated by findings showing that their levels vary from basal to apical parts of the plants (Arora *et al.* 2003). Some examples of factors to consider for hormone studies include the use of lateral buds against terminal buds, distinguishing determinate and indeterminate growth patterns, use of whole buds against partitioned bud tissues, sampling buds at quantitatively established stages of dormancy and differential photoperiodic response of young and mature leaves. In the case of ABA, other more recent studies have further complicated the importance of ABA in dormancy due to many other processes mediated by ABA particularly auxin- and ethylene-triggered ABA induction (Grossmann and Hansen 2000, Hansen and Grossmann 2000, Sharp *et al.* 2000).

While the implication of basipetal transport of auxin as the primary signal regulating paradormancy is well documented (Horvath 2003), based on grafting studies, other signals have been proposed to significantly influence shoot outgrowth (Cline 1994, Beveridge *et al.* 2000). Although growth inhibition via basipetal transport of auxin is slightly complicated by concurrent production of auxin in growing buds and by the plant's requirement for auxin, the effect of auxin produced from the distal meristem seems to be different from that inside the buds once dormancy is broken, implicating different effectors, pathways or interacting partners. Several studies confirm that auxin signaling alters cell cycle directly or through crosstalk in concert with other plant

hormones. It has been shown to inhibit the production or sensing of cytokinin (Francis and Sorrell 2001, Ferguson and Beveridge 2009).

Other plant hormones acting alongside auxin in paradormancy include ABA and GA, which inhibit and promote growth, respectively. ABA induces expression of an inhibitor (*ICK1*) of *CDK* action at the G1-S-phase transition (Wang *et al.* 1997), while GA induces S-phase progression (Sauter 1997). Auxin signaling pathways target degradation of specific proteins and regulation of cytokinin production in the stem segments adjacent to the axillary buds (Shimizu-Sato and Mori 2001, Stirnberg *et al.* 2002, Xiangdong and Harberd 2003). It has also been proposed that auxin might regulate ABA content through expression of a *P450 mono-oxygenase* gene (Shimizu-Sato and Mori 2001). Details of a pathway or an auxin controlled complex remain elusive.

Besides hormones, sugars also play a complex role in paradormancy (Healy *et al.* 2001, Oakenfull *et al.* 2002). The role of sugars in determining the competence of a perennial plant for flowering during the vegetative juvenile stage and just before bud set was mentioned earlier (Bernier and Périlleux 2005). In a peach study, cell wall invertase activity and imported hexose content in the meristematic tissues had positive correlations with the bud break capacity (Maurel *et al.* 2004).

Bud dormancy maintenance and release

At the onset of winter, endodormancy results from physiological changes in woody perennials that follow paradormancy in the growth cycle. This response is internal to the bud and prevents untimely growth during seasonal transitions when temperatures often fluctuate between favorable warm and inhibitory cold temperatures. This stage of dormancy reflects the plants adaptive mechanisms to maintain buds in a physiologically dormant state until a return of persistent favorable conditions. Compared to maintenance of the paradormant state the molecular components of endodormancy maintenance are much less well understood and seem to overlap and share similar aspects with cold acclimation, making it more intractable to investigation than other stages. In dicots, endodormancy has been studied in buds of poplar (*Populus deltoids*) and grape (*Vitis vinifera*) and in potato tuber buds.

Endodormancy occurs concurrently with plant senescence during the fall in several plant systems (Fedoroff 2002), with ethylene and ABA been implicated in both processes. In potato microtubers, ethylene directly induces endodormancy (Suttle 1998); while the role of ABA includes growth cessation in potato tubers and inhibitory effects on seed germination in several plant species (Leung and Giraudat 1998). Cases of phytochromes acting synergistically with both ethylene and ABA have been reported (Finlayson *et al.* 1998, Weatherwax *et al.* 1998). The signaling pathways for this molecular mechanism in woody perennials are still been deciphered, especially for ABA action; however no concrete connections have yet been identified.

One economically important major challenge in the horticultural industry is the importance of adequate chilling to temperate fruit tree production in regions with varying cold and warmer winter temperatures. Warmer climates often lack sufficient chilling required to overcome floral and vegetative bud dormancy, while trees with fulfilled CRs in colder regions are prone to spring frost damage during unusually early warm temperature spells. Varietal breeding programs need to incorporate high CR into varieties destined for cold climates, while cultivars with variable (for early and relatively late blooms) low CR are desirable in warm climates. Several studies have been published on the regulation of bud break (Erez *et al* 1971, Saure 1985, Iwahori *et al.* 2002), as well as, the use of chemicals to break dormancy (Erez *et al.* 1971, Erez 1987, Fernandez-Escobar and Martin 1987, Siller-Cepeda *et al.* 1992, Wood 1993). A proper understanding of pathways, signal molecules and target genes underlying bud dormancy maintenance and release may aid development of markers to assist in the breeding of varieties that match environmental conditions to the proper timing of bud break (Tamura *et al.* 1998).

Several strategies have been utilized to elucidate the process of dormancy and bud break. These include approaches based on regulation within the apical meristem by changes in cell-to-cell communication and plasmodesmatal connections (van der Schoot 1996, Jian *et al.* 1997, Rinne *et al.* 2001), control of the cell cycle (Rohde *et al.* 1997, MacDonald 2000), regulation of water with initial findings based on supercooling after examining the vascular connections into the bud (Sakai 1979, Ashworth 1984, Quamme *et al.* 1995), the sequence and regulation of water uptake into the bud (De Fay *et al.* 2000), water stress and availability during dormancy (Faust *et al.* 1997), studying molecular events involved

in the reception and transduction of dormancy-breaking signals during chemical-induced dormancy release (Or *et al.* 2000, 2002) and the mechanism of dormancy induction and release via a metabolic and communication block or permeability barrier between the bud and adjacent tissues (Champagnat 1989, Crabbe and Barnola 1996, Faust *et al.* 1997).

To reproduce the effect of CR on dormancy release, horticulturalists have successfully used chemicals such as hydrogen cyanamide (HC) for controlled dormancy release in grape and kiwifruit (Henzell 1991, Pérez *et al.* 2008). Transcript populations from HC-treated and control buds have been used to identify a *SUCROSE NON-FERMENTING* (*SNF*)-*like* protein kinase that is upregulated during initial stages of dormancy release (Or *et al.* 2000, 2002). Although the mechanisms underlying dormancy release using the chemicals are unknown, there is mounting evidence that an *SNF-like* protein kinase plays a role in the signaling cascade. Since *SNF-like* protein kinases are known to be transcriptionally regulated by stress stimuli in plants (Anderberg and Walker-Simmons 1992, Hardie 1994), Or *et al.* (2002) suggests that they might be involved in perception of stress signal induced by HC and similar chemicals (e.g. azide, cyanide, thidiazuron) in grape. These chemicals are theorized to transiently disrupt respiratory metabolism by inducing H_2O_2 via oxidative stress, an explanation supported by reduced catalase activity (a free radical scavenger) soon after HC application (Nir *et al.* 1986, Wang *et al.* 1991, Faust and Wang 1993, Pérez and Lira 2005**).** The inhibition of catalase by HC could be as a result of H_2O_2 production or H_2O_2 acting as a chemical signaling molecule inducing the up-regulation of genes related to endodormancy release (Desikan *et al* 2000, Neill *et al.* 2002).

Studies in dormant apple buds indicate that dormancy release in buds coincides with the up-regulation of the antioxidant system, reflected by increased levels of peroxide scavenging enzymes (Wang and Faust 1994, Rowland and Arora 1997). The antioxidant machinery is also known to be up-regulated for protection against freezing stress (Guy 1990). More recently, the *MITOGEN-ACTIVATED PROTEIN KINASE* (*MAPK*) cascade has been implicated to play a role in transducing signals involving reactive oxygen species (ROS) like H_2O_2 (Dóczi *et al.* 2007; Pitzschke and Hirt 2009) to a corresponding H_2O_2-induced dormancy release in grape (Pérez and Lira 2005) and raspberry (Mazzitelli *et al.* 2007). Several studies now show that the *MAPK* cascade is not only induced by ROS but can also regulate production of ROS (Pitzschke and Hirt 2009). *MAPK*, sometimes referred to as extracellular-signal-regulated kinases (*ERKs*), are some of the best studied signal transduction pathways that play central roles in signaling cells to progress past the G1/S boundary (Meskiene and Hirt 2000; Roberts *et al.* 2000). These growth factor signaling pathways are implicated in the up-regulation of *CYCLIN D1* and *CKIs* (Cook *et al.* 2000) and in activation of *CAK* (Chiariello *et al.* 2000). Recently, components of the *MAPK* signal cascade have also been associated with oxidative stress-induced cell cycle arrest at G2/M (Chien *et al.* 2000; Kurata 2000).

As mentioned earlier, along with changes in gene expression, there is also evidence for more general epigenetic changes associated with endodormancy induction and release. Major changes in DNA methylation have been observed during bud set, dormancy induction and release in potato (Law and Suttle 2003) and azalea floral buds (Meijón *et al.* 2010). Increased DNA methylation and histone deacetylation act simultaneously and

coordinately following dormancy induction, suggesting that chromatin remodeling plays an important role in restructuring chromatin and regulating gene expression during this process. Interestingly, the previously mentioned *SNF1-like* protein kinase, activated in grape by HC, is similar to a known component of a DNA modifying protein complex *SW1-SNF* from yeast and animals (Fan *et al.* 2003). Other components of this complex interact with *RB-E2F* (Figure 1.2) in both plants and animals (Shen 2002).

Genetic control of endodormancy-related traits in woody perennials

In the past decades, little effort was made to understand regulation of dormancy from a global regulation and genetic perspective because dormancy-related traits like many other polygenic traits were considered too complex. This was partly due to limited genomic resources and the lack of analytical tools (Tanksley and Hewitt 1988, Tanksley *et al.* 1989). Other obstacles that prevented performing genetic studies in woody perennials include a long generation time, high ploidy levels in economically important crops, inbreeding depression, self- and cross-incompatibility (Janick and Moore 1975, Moore and Janick 1983). Early genetic studies on bud dormancy estimated the heritability and classic Mendelian genetic behavior of a few traits, followed by genetic studies of evergrowing mutants in hazelnut (Thompson *et al.* 1985) and peach (Rodriguez *et al.* 1994) which suggested that their lack of dormancy induction was due to a single recessive gene. Hansche (1990) reported high heritability estimates for leaf abscission during fall and spring bloom date in peach implying a strong genetic component for these

traits. Studies in apple (*Malus x domestica* Borkh) also confirmed a strong genetic component for CR and provided evidence that a major dominant gene controls low CR and minor genes modulate its effect (Hauagge and Cummins 1991).

Most dormancy-related traits are inherited in a quantitative manner and display a continuous distribution in phenotype values among progenies of crosses segregating for these traits. This strongly indicates a polygenic mode of inheritance (Farmer and Reinholt 1986, Billington and Pelham 1991, Bradshaw and Stettler 1995, Lawson *et al.* 1995, Howe *et al.* 1999, 2000). The first QTL analyses on dormancy-related traits in woody perennials were performed in an F1 population (double pseudo-testcross) of apple (Lawson *et al.* 1995) and in an F2 population of poplar (Bradshaw and Stettler 1995). Two QTLs for bud flush were placed on an apple genetic map; while five QTLs explaining 85% of the phenotypic variance were detected in the poplar map. In apple, another study for vegetative bud flush with a larger F1 population size (172 individuals) detected 8 QTLs on 6 linkage groups that explained 42 % of the phenotypic variance (Conner *et al.* 1998); however, none of these linkage groups was homologous to the linkage group with the QTLs from the initial study. In poplar the population size was also increased to 346 in an F2 population segregating for fall bud set and spring bud flush (Frewen *et al.* 2000). With the intent of mapping possible candidate genes, 3 QTLs distributed over 3 linkage groups were associated with bud set and 6 QTLs were distributed over 6 linkage groups for bud flush. The 3 bud set QTLs co-localized with 3 of the QTLs for bud flush implying that a single QTL could have pleiotropic effects on both traits as a result of shared components in their biochemical pathways. After

comparing the 2 poplar maps, 3 QTLs were found to be common in both studies and all 3 contained bud flush QTLs. Following the mapping of candidate genes involved in perception of photoperiod, *PHYB1* and *PHYB2*, and genes involved in the signal transduction of ABA response signals, *ABI1B*, *ABI1D* and *ABI*, only *PHYB2* and *ABI1B* were found to map near but not inside QTLs affecting both bud set and bud flush. The lack of co-localization of the sensors of photoperiod with QTLs that control dormancy, suggest that light may not be the direct regulator of the system. These results would be consistent with a model that light may act indirectly through production of sugars that may more directly regulate the system.

Several other maps have been constructed for detection of QTLs controlling bud dormancy and related traits but with little success at identifying candidate genes. Some of these studies include studies in apple (Liebhard *et al.* 2003a,b, Segura *et al.* 2006), sour cherry (Wang *et al.* 2000), raspberry (Graham *et al.* 2009) and Douglas fir (*Pseudotsuga menziesii* Franco var. *menziesii*) (Jermstad *et al.* 2001). Besides bud set and bud flush QTLs, efforts were directed towards identifying QTLs and mode of potential gene actions underlying CR in blueberry (*Vaccinium* section *Cyanococcus*) (Rowland *et al.* 1999). CR was chosen as a study phenotype because of interest in developing low-CR cultivars for warmer winter conditions (Hancock and Draper 1989, Hancock *et al.* 1995). The CR of a cultivar is known to broadly impact the timing of bud flush, preventing growth during transitory periods, synchronize plant growth with exposure to stable favorable conditions and select for cold hardiness. It is the major factor determining bud break (Ruiz *et al.*,

2007; Alburquerque *et al.*, 2008), which is an important agronomic trait affecting production in temperate fruit tree species.

Vast amounts of information has been revealed from transcriptome analysis and expression studies that propose a plethora of plausible candidate genes in grape (Mathiason *et al.* 2009, Ophir *et al.* 2009), raspberry (Mazzitelli *et al.* 2007) and poplar (Rohde *et al.* 2007), but the short-coming of such studies lies in their inability to identify cause and effect genes from the differential expressions. The expression study on dormancy release by Mathiason *et al.* (2009) reported differential expression of several genes already characterized in vernalization pathways of model annual plants in relation to flowering time, indicating that some components of these pathways are conserved in woody perennials. These genes include *FLOWERING TIME LOCUS T* (*FT*), *SUPPRESSOR OF OVER-EXPRESSION OF CONSTANS 1* (*SOC1*), *LEAFY* (*LFY*), *FRIGIDA* (*FRI*), *FLOWERING LOCUS C* (*FLC*), *GIGANTEA* (*GI*), *CONSTANS* (*CO*), *VERNALIZATION INDEPENDENT 3* (*VIN3*), *VERNALIZATION 1* (*VRN1*) and *VERNALIZATION 2* (*VRN2*).

In poplar, the most prominent genes revealed by differential gene expression after exposure to 24 short-days were identified using an amplified fragment length polymorphism-based (cDNA-AFLP) transcript profiling (Rohde *et al.* 2007). These were three regulatory genes, *AP2/EREBP* (*APETALA 2*/ *ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 13*), *ERF4* (*ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 4*), and *WRKY11* (Calmodulin binding/ transcription factor). These

genes were linked with critical steps in dormancy induction (Rohde *et al.* 2007). Inference from the gene functional analyses of the respective *Arabidopsis* homologues showed that they act downstream of the ethylene and/or abscisic acid (ABA) signaling cascade that is successively initiated during bud development (Ruttink *et al.*, 2007). The closest homologue of the *AP2/EREBP* transcription factor in *Arabidopsis* (RAP2.6L), is an *AP2-like* ABA repressor 1 gene (McGrath *et al.* 2005, Nakano *et al.* 2006) that acts in a network regulating shoot regeneration from root explants (Che *et al.* 2006). The *Arabidopsis* homologue for *ERF4* (McGrath *et al.* 2005) is induced transcriptionally by ethylene, ABA and jasmonate and has been identified independently during short-day-induced bud set in poplar (Ruttink *et al.*, 2007). It acts as a transcriptional repressor that modulates ethylene and ABA responses in *Arabidopsis*, while overexpression of this gene causes ethylene insensitivity and reduced ABA sensitivity (Yang *et al.*, 2005). The role of sugars was verified by Mazzitelli *et al.* (2007) where a putative raspberry plasma membrane H±ATPase gene was significantly up-regulated during dormancy release. Sugar influx has been suggested to occur through H^+/sugar symports based on the pH gradient produced by a plasma membrane H±ATPase (Alves *et al.* 2001). Gevaudant *et al.* (2001) also confirmed in dormancy release of peach buds that carbohydrate uptake capacity of buds increases concurrently with the up-regulation and increased activity of the plasma membrane H±ATPase.

Summary of chapters

CR and dormancy in Prunus species

Besides the mapping (Wang *et al.* 2002), annotation (Bielenberg *et al.* 2008) and expression studies (Li *et al.* 2009, Jiménez *et al.* 2010) of the peach *DAM* genes, there remains a lack of insight into genes associated specifically with this complex trait and generally with the molecular mechanisms underlying the constituent pathways. Several studies in *Prunus* species have attempted to elucidate genetic factors controlling only blooming date using QTL analysis and do not reflect the comprehensive biological processes involved in dormancy ranging from induction to release (Dirlewanger *et al.* 1999, Verde *et al.* 2002, Silva *et al.* 2005). Although QTLs controlling the blooming date trait were detected, defining the genomic regions produced intervals spanning several cM (mostly > 20 cM) due to inadequate marker saturation and limited mapping population size. With adequate resources now available, the *Prunus* system provides the most tractable genetic system in the Rosaceae family and woody perennials in general. This is due to their relatively small genome size (approx. 0.6 pg/2C), which is only twice as much as Arabidopsis (0.3 pg/2C) and the diploid nature of their genome (including their cultivars), unlike the larger genome size and polyploidy observed in several other Rosaceae genera like *Malus* (1.57 pg/2C), *Pyrus* (1.11 pg/2C) and other systems like poplar (1.1 pg/2C).

Project rationale and current state of prunus genomics.

Subsequent studies aim to generate a high-density map of the apricot genome using genetic linkage and LD-based association mapping approaches towards defining genomic regions (QTLs) controlling CR and bud break. The apricot genome serves as an ideal system amenable for genetic studies compared to peach. This is mostly due to the high level of heterozygosity in the genome, enabling greater ease at generating numerous polymorphic loci for linkage mapping. This is also reflected in the broad genetic base of the apricot germplasm and its outcrossing nature, making it an ideal system for resolving QTL regions in greater detail using the LD-based association mapping approach. Although, genetic resources for *Prunus* are mainly based on the peach genome, the highly collinear genomes of peach and apricot allow for easy transferability of marker and genetic information. Complementary studies in both species along with other genomic resources (BAC libraries, peach genome sequence, QTL maps and expression study data) available from the Rosaceae community will also be used to identify candidate genes within the QTLs. Additionally, comparative mapping between peach and apricot will provide us with stable QTLs, as well as differences, between the apricot and peach phenotypes, enabling the study of the evolutionary events underlying the trait. The study provides the genetic substrate for preliminary gene expression and methylation studies of buds sampled during developmental stages spanning dormancy induction, maintenance and release.

The current state of *Prunus* genomics will facilitate dissection of genetic and molecular mechanisms driving natural variations observed in this trait in an unprecedented manner. These resources include *Prunus* BAC and EST libraries, expression studies within *Prunus* and related woody perennials, and an extensive annotation of candidate genes in several systems. The complete sequencing and assembly of a dihaploid-derived peach genome also provides an unprecedented ability to mine candidate genes for QTLs as well as marker design for fine mapping. Reported in this dissertation are:

1) Construction of high-density apricot linkage maps: Two parental maps comprising corresponding to the apricot 8 chromosomes were aligned against the *Prunus* reference map using *Prunus* anchor SSR marker sets to establish conservation of synteny.
2) QTL mapping of CR for vegetative bud break in apricot: A total of 20 putative CR QTLs were detected on the apricot 8 linkage groups after applying a model based on an additive-by-additive epistatic interaction with and without dominance. Four of the 12 QTLs detected for each of the two models were stable, while majority of the alleles that increase trait value were contributed by the high chill parent.
3) Comparative analysis of QTLs underlying CR and bud break in peach and apricot: A majority of the QTLs were shown to be stable between peach and apricot, as well as similar QTL effects that explain the parental origin of the allele that increases the trait value. The study reveals transferability of genetic information between these 2 *Prunus* species and validation of previously identified QTLs.

4) Linkage disequilibrium (LD)-based mapping of CR for floral bud break in an apricot germplasm: Fine mapping of 2 of the major QTLs for gene discovery revealed candidate genes that were strongly and siginificantly associated with the CR trait.

The literature reviewed above is relevant to the next three chapters that comprise research studies attempting to uncover the genetic mechanisms underlying CR and bud break in *Prunus*.

References

Alburquerque, N., Garcia-Montiel, F., Carrillo, A, Burgos, L. 2008. Chilling and heat requirements of sweet cherry cultivars and the relationship between altitude and the probability of satisfying the chill requirements, Environ. Exp. Bot. 64(2): 162-170. doi:10.1016/j.envexpbot.2008.01.003.

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

Alves, G., Sauter, J.J., Julien, J.L., Fleurat-Lessard, P., Ameglio, T., Guillot, A., Petel, G., Lacointe, A., 2001. Plasma membrane H^+-ATPase, succinate and isocitrate dehydrogenases activities of vessel-associated cells in walnut trees, J. Plant Physiol. 158(10): 1263-1271.

Anderberg, R.J., Walker-Simmons, M.K. 1992. Isolation of a wheat cDNA clone for an abscisic acid-inducible transcript with homology to protein kinases. Proc. Natl. Acad. Sci. U.S.A. 89: 10183-10187.

Arora, R., Rowland L.J., Panta, G.R. 1997. Cold hardiness and dormancy transitions in blueberry and their association with accumulation of dehydrin-like proteins. Physiol. Plant. 101: 8-16.

Arora, R., Rowland, L.J., Tanino, K. 2003. Induction and release of bud dormancy in woody perennials: A science comes of age. HortScience. 38 (5): 911-921.

Arora, R., Wisniewski, M.E. 1994. Cold acclimation in genetically related (sibling) deciduous and evergreen peach (*Prunus persica* [L.] Batsch). II. A 60-kilodalton bark protein in cold-acclimated tissues of peach is heat stable and related to the dehydrin family of proteins. Plant Physiol. 105: 95-101.

Arora, R., Wisniewski, M.E., Scorza, R. 1992. Cold acclimation in genetically related (sibling) deciduous and evergreen peach (*Prunus persica* [L.] Batsch) I : seasonal changes in cold hardiness and polypeptides of bark and xylem tissues. Plant Physiol. 99: 1562–1568.

Artlip, T.S., Callahan, A.M., Bassett, C.L., Wisniewski, M.E. 1997. Seasonal expression of a dehydrin gene in sibling deciduous and evergreen genotypes of peach (*Prunus persica* [L.] Batsch). Plant Mol. Biol. 33: 61–70.

Ashworth, E.N. 1984. Xylem development in *Prunus* flower buds and the relationship to deep supercooling. Plant Physiol. 74: 862-865.

Bangerth, K.F. 2009. Floral induction in mature, perennial angiosperm fruit trees: Similarities and discrepancies with annual/biennial plants and the involvement of plant hormones. Sci. Hortic. 122(2): 153-163. doi:10.1016/j.scienta.2009.06.014.

Barros, R.S., Neill, S.J. 1989. The status of abscisic acid in willow as related to the induction of bud dormancy. Acta Physiol. Plant. 11(2): 117-123.

Bernier, G. 1988. The control of floral evocation and morphogenesis. Annu. Rev. Plant. Physiol. Plant Mol. Biol. 39: 175–219.

Bernier, G., Périlleux, C. 2005. A physiological overview of the genetics of flowering time control. Plant Biotech. J. 3: 3–16.

Beveridge, C.A., Symons, G.M., Turnbull, C.G. 2000. Auxin inhibition of decapitation-induced branching is dependent on graft-transmissible signals regulated by genes Rms1 and Rms2. Plant Physiol. 123(2): 689-98.

Bielenberg, D.G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G.L., Scorza, R., Abbott, A.G. 2008. Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet. Genomes 4: 495–507.

Billington, H. L., Pelham, J. 1991. Genetic variation in the date of budburst in Scottish birch populations: implications for climate change. Funct. Ecol. 5: 403–409.

Bradshaw, H. D., Stettler, R. F. 1995. Molecular genetics of growth and development in Populus. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics 139: 963– 973.

Champagnat, P. 1989. Rest and activity in vegetative buds of trees. Annales des Sciences Forestieres 46 (suppl.): 9-26.

Chandler, W.H., Tufts, W.P. 1934. Influence of the rest period on opening of buds of fruit trees in spring and on development of flower buds of peach trees. Proc. Amer. Soc. Hort. Sci. 30: 180-186.

Che, P., Lall, S., Nettleton, D., Howell, S.H. 2006. Gene expression programs during shoot, root, and callus development in *Arabidopsis* tissue culture. Plant Physiol 141: 620–637.

Chen, K.Y., Coleman, G. 2006. Type II MADS-box genes associated with poplar apical bud development and dormancy. Ann. Meeting Amer. Soc. Plant Biologists. Abstr. pp. 118.

Chiariello, M., Gomez, E., Gutkind, J. 2000. Regulation of cyclin-dependent (Cdk) 2 Thr-160 phosphorylation and activity by MAP kinase in late G1 phase. Biochem. J. 349: 869-876.

Chien, M., Rinker-Schaeffer, C., Stadler, W.M. 2000. A G2/M growth arrest response to low-dose intermittent H2O2 in normal uroepithelial cells. Int. J. Oncol. 17: 425–432.

Choi, J., Hyun, Y., Kang, M.J., In Yun, H., Yun, J.Y., Lister, C., Dean, C., Amasino, R.M., Noh, B., Noh, Y.S., Choi, Y. 2009. Resetting and regulation of FLOWERING LOCUS C expression during *Arabidopsis* reproductive development. Plant J. 57: 918–931.

Cline, M. G. 1994. The role of hormones in apical dominance. New approaches to an old problem in plant development. Physiol Plant 90: 230-237.

Close, T.J. 1997. Dehydrins: a commonalty in the response of plants to dehydration and low temperature. Physiol Plant 100: 291–296.

Conner, P.J., Brown, S.K., Weeden, N.F. 1998. Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. Theor. Appl. Genet. 96: 1027–1035.

Cook, S. J., Balmanno, K., Garner, A., Millar, T., Taverner, C., Todd, D. 2000. Regulation of cell cycle re-entry by growth survival and stress signaling pathways. Biochem. Soc. Trans. 28: 233–240.

Corbesier, L., Lejeune, P., Bernier, G. 1998. The role of carbohydrates in the induction of flowering in *Arabidopsis thaliana*: comparison between the wild type and a starchless mutant. Planta 206: 131–137.

Crabbe, J. 1994. Dormancy. In: Arntzen, C.J. and Ritter, E.M. (eds). Encyclopedia of agricultural sciences, Academic Press, New York. vol. 1: 597-611.

Crabbe, J., Barnola, P. 1996. A new conceptual approach to bud dormancy in woody plants. In: ed. G.A. Lang. Plant dormancy, physiology, biochemistry and molecular biology. Wallingford, UK: CAB International.

De Fay, E., Vacher, V., Humbert, F. 2000. Water-related phenomena in winter buds and twigs of *Picea abies* L. (Karst.) until bud-burst: a biological, histological and NMR study. Ann. Bot. 86: 1097–1100.

Dennis, F.G., Edgerton L.J. 1961. Understanding Abscisic AcidThe relationship between an inhibitor and rest in peach flower buds. Amer. Soc. Hort. Sci. 77: 107-116.

Desikan, R., Neill, S.J., Hancock, J.T. 2000. Hydrogen peroxide induced gene expression in *Arabidopsis thaliana*. Free Radic. Biol. Med. 28: 773–778.

Dirlewanger, E., Moing, A., Rothan, C., Svanella, L., Pronier, V., Guye, A., Plomion, C., Monet, R. 1999. Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batch). Theor. Appl. Genet. 98: 18–31.

Distelfeld, A., Li, C., Dubcovsky, J. 2009. Regulation of flowering in temperate cereals. Curr. Opin. Plant Biol. 12(2): 178-84.

Dóczi, R., Brader, G., Pettkó-Szandtner, A., Rajh, I., Djamei, A., Pitzschke, A., Teige, M., Hirt, H. 2007. The *Arabidopsis* mitogen-activated protein kinase kinase MKK3 is upstream of group C mitogen-activated protein kinases and participates in pathogen signaling. Plant Cell, 19(10): 3266–3279.

Doorenbos, J. 1953. Review of the literature on dormancy in buds of woody plants. Meded. van de Landbouwhogesch. Wageningen 53: 1-23.

Dumbroff, E.B., Cohen, D.B., Webb, D.P. 1979. Seasonal levels of abscisic acid in buds and stems of *Acer saccharum*. Physiol Plant 45: 211-214.

Eagles, C. F., Wareing, P. F. 1963. Experimental induction of dormancy in *Betula pubescens*. Nature 199: 874–875.

Erez, A. 1987. Chemical control of bud break, Hortic. Sci. 22: 1240-1243.

Erez, A., Couvillon, G.A. 1987. Characterization of the influence of moderate temperatures on rest completion in peach, J. Am. Soc. Hortic. Sci. 112: 677-680.

Erez, A., Couvillon, G.A., Henderschott, C.H. 1979. Quantitative chilling enhancement and negation in peach buds by high temperatures in a daily cycle. J. Am. Soc. Hortic. Sci. 104: 536-540.

Erez, A., Faust, M., Line, M.J. 1998. Changes in water status in peach buds on induction, development and release from dormancy. Sci. Hortic. 73: 111-123.

Erez, A., Lavee, S. 1974. Recent advances in breaking the dormancy of deciduous fruit trees. 19th Int. Hortic. Congr., Warszawa, pp 69–78.

Erez, A., Lavee, S., Samish, R.M. 1971. Improved methods for breaking rest in the peach and other deciduous fruit species. J. Am. Soc. Hort. Sci. 96: 519-522.

Fan, H.Y., He, X., Kingston, R.E., Narlikar, G.J. 2003. Distinct strategies to make nucleosomal DNA accessible. Mol. Cell 11: 1311–1322.

Farmer Jr., R.E., Reinholt, R.W. 1986. Genetic variation in dormancy relations of balsam poplar along a latitudinal transect in northwestern Ontario. Silvae Genet. 35:38–42.

Faust, M., Erez, A., Rowland, L.J., Wang, S.Y., Norman, H.A. 1997. Bud dormancy in perennial fruit trees: Physiological basis for dormancy induction, maintenance and release. HortScience 32(4): 623-629.

Faust, M., Liu, D., Millard, M.M., Stutte, G.W. 1991. Bound versus free water in dormant apple buds–a theory for endodormancy. HortScience 26(7): 887-890.

Faust, M., Wang, S.Y. 1993. Polyamines in horticuturally important plants. Hortic. Rev. 14: 333-356.

Fedoroff, N. 2002. Cross-talk in abscisic acid signaling. Sci. STKE 140: re10.

Fennell, A., Hoover, E. 1991. Photoperiod influences growth, bud dormancy, and cold acclimation in *Vitis labruscana* and *V. riparia*. J. Am. Soc. Hortic. Sci. 116: 270 273.

Fennell, A., Line, M.J. 2001. Identifying differential tissue response in grape (*Vitis riparia*) during induction of endodormancy using nuclear magnetic resonance imaging. J. Am. Soc. Hortic. Sci. 126(6): 681-688.

Fennell, A., Wake, C., Molitor, P. 1996. Use of ^{1}H-NMR to determine grape bud water state during the photoperiodic induction of dormancy. J. Am. Soc. Hortic. Sci. 121: 1112-1116.

Ferguson, B.J., Beveridge, C.A. 2009. Dependent and independent roles of plant hormones in apical dominance, correlative inhibition and the RMS/Strigolactone pathway. Plant Physiol 149: 1929-1944.

Fernandez-Escobar, R., Martin, R., 1987. Chemical treatments for breaking rest in peach in relation to accumulated chilling. J. Hort. Sci. 62: 457–461.

Finlayson, S.A., Lee, I.J., Morgan, P.W. 1998. Phytochrome B and the regulation of circadian ethylene production in sorghum. Plant Physiol. 116: 17–25.

Francis, D., Sorrell, D. 2001. The interface between the cell cycle and plant growth regulators: a mini review [J]. Plant Growth Reg. 33: 1–12.

Freundl, E., Steudle, E., Hartung, W. 2000. Apoplastic transport of abscisic acid through roots of maize: effect of the exodermis. Planta 210: 222-231.

Frewen, B.E., Chen, T.H., Howe, G.T., Davis, J., Rohde, A., Boerjan, W., Bradshaw Jr., H.D. 2000. Quantitative trait loci and candidate gene mapping of bud set and bud flush in populus. Genetics. 154(2): 837-45.

Fuchigami, L.H., Weiser, C.J., Kobayashi, K., Timmis, R., Gusta, L.V. 1982. A degree growth stage (oGS) model and cold acclimation in temperate woody plants. In: P.H. Li and A. Sakai, eds. Plant cold hardiness and freezing stress. New York: Academic Press.

Fujiwara, S., Oda, A., Yoshida, R., Niinuma, K., Miyata, K., Tomozoe, Y., Tajima, T., Nakagawa, M., Hayashi, K., Coupland, G., Mizoguchi, T. 2008. Circadian clock proteins LHY and CCA1 regulate SVP protein accumulation to control flowering in *Arabidopsis*. Plant Cell. 20(11): 2960-71.

Gevaudant, F., Petel, G., Guilliot, A. 2001. Differential expression of four members of the H + -ATPase gene family during dormancy of vegetative buds of peach trees. Planta 212: 619–626.

Graham, J., Woodhead, M., Smith, K., Russell, J., Marshall, B., Ramsay, G., Squire, G. 2009. A decade of change; new insight into wild raspberry populations using SSR markers. J. Am. Soc. Hortic. Sci. 134 (1): 1-11.

Grossmann, K., Hansen, H. 2000. Ethylene-triggered abscisic acid: a principle in plant growth regulation? Physiol. Plant. 113: 9-14.

Guy, C.L. 1990. Cold acclimation and freezing stress tolerance: role of protein metabolism. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41: 187-223.

Hancock, J. F., Draper, A. D. 1989. Blueberry culture in North America. HortScience 24: 55 1-556.

Hancock, J.F., Erb, W.A., Goulart, B.L., Scheerens, J.C. 1995. Utilization of wild blueberry germplasm: The legacy of Arlen Draper. J. Small Fruit Viticult. 3: 1-16.

Hansche, P.E., 1990. Heritability of spring bloom and fall leaf abscission dates in *Prunus persica*. HortScience 25: 1639–1641.

Hansen, H., Grossmann, K. 2000. Auxin-induced ethylene triggers abscisic acid biosynthesis and growth inhibition. Plant Physiol. 124: 1437-1448.

Hardie, D.G. 1994. Ways of coping with stress. Nature 370: 599–600.

Hauagge, R., Cummins, J. N. 1991. Genetics of length of dormancy period in *Malus* vegetative buds. J. Am. Soc. Hortic. Sci. 116: 121–126.

He, Y., Amasino, R.M. 2005. Role of chromatin modification in flowering-time control. Trends Plant Sci. 10(1): 30-5.

Healy, J.M., Menges, M., Doonan, J.H., Murray, J.A.H. 2001. The *Arabidopsis* D-type cyclins CycD2 and CycD3 both interact in vivo with the PSTAIRE cyclin-dependent kinase Cdc2a but are differentially controlled. J. Biol. Chem. 276: 7041–7047.

Henzell, R.F., Briscoe, M.R., Gravett, I. 1991. Improving kiwifruit vine productivity with plant growth regulators. Acta Hortic. 297: 345–350.

Hermberg, T. 1949. Growth inhibitors in terminal buds of *Fraxinus*. Physiol. Plant. 2: 37.

Heyer, A.G., Raap, M., Schroeer, B., Marty, B., Willmitzer, L. 2004. Cell wall invertase expression at the apical meristem alters floral, architectural, and reproductive traits in *Arabidopsis thaliana*. Plant J. 39: 161-169.

Hill, J., Becker, H.C., Tigerstedt, P.M.A. 1998. Quantitative and ecological aspects of plant breeding. St. Edmundsbury Press, Suffolk, U.K.

Horvath, D.P., Anderson, J.V., Chao, W.S., Foley, M.E. 2003. Knowing when to grow: signals regulating bud dormancy. Trends Plant Sci. 8(11): 534-40.

Howe, G. T., Davis, J., Frewen, B., Saruul, P., Jeknić, Z., Bradshaw Jr., H. D., Chen, T.H.H. 1999. Physiological and genetic approaches to studying endodormancy-related traits in Populus. Hortscience 34: 1174–1184.

Howe, G.T., Saruul, P., Davis, J., Chen, T.H.H. 2000. Quantitative genetics of bud phenology, frost damage, and winter survival in an F2 family of hybrid poplars. Theor. Appl. Genet. 101: 632-642.

Huang, J., Pray, C., Rozelle, S. 2002. Enhancing the crops to feed the poor. Nature 418: 678-684.

Iwahori, S., Gemma, H., Tanabe, K., Webster, A.D., White, A.G. 2002. Proceedings of the international symposium on asian pears commemorating the 100th anniversary of 'Nijisseiki' pear. Acta Hort. (587).

Iwasaki, K., Weaver, R.J. 1977. Effects of chilling, calcium cynamide, and bud scale removal on bud break, rooting, and inhibitor content of buds of 'Zinfandel' grape (*Vitis vinifera* L.). J. Am. Soc. Hortic. Sci. 102(5): 584-587.

Janick, J., Moore, J.N. eds. 1975.Advances in fruit breeding. Purdue Univ. Press, West Lafayette, Ind.

Jermstad, K.D., Bassoni, D.L., Jech, K.S., Wheeler, N.C., Neale, D.B. 2001. Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir. I. Spring bud flush. Theor. Appl. Genet. 102:1142–1151.

Jian, L., Li, P.H., Sun, L., Chen, T.H.H. 1997. Alterations in ultrastructure and subcellular localization of Ca2+ in poplar apical bud cells during the induction of dormancy. J. Exp. Bot. 48: 1195–1207.

Jiménez, S., Li, Z., Reighard, G.L., Bielenberg, D.G. 2010. Identification of genes associated with growth cessation and bud dormancy entrance using a dormancy-incapable tree mutant. BMC Plant Biol. doi:10.1186/1471-2229-10-25.

Jung, C., Müller, A.E. 2009. Flowering time control and applications in plant breeding. Trends Plant Sci. 14(10): 563-73.

Keep, E., 1988. Primocane (autumn)-fruiting raspberries: A review with particular reference to progress in breeding. J. Hort. Sci. 63: 1–18.

Kim, S.Y., Park, B.S., Kwon, S.J., Kim, J., Lim, M.H., Park, Y.D., Kim, D.Y., Suh, S.C., Jin, Y.M., Ahn, J.H., Lee, Y.H. 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 Rep. 26: 327–336.

Klinkhamer, P.G.L., De Jong, T.J., Meelis, E. 1987. Life-history variation and the control of flowering in short-lived monocarps. Oikos, 49: 309–314.

Koch, K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. Curr. Opin. Plant Biol. 7: 235–246.

Koussa, T., Zaoui, D., Broquedis, M. 1998. Relationship between the levels of abscisic acid in latent buds, in leaves and in internodes of *Vitis vinifera* L. cv. Merlot during the dormancy phase. Journal International des Sciences de la Vigne et du Vin 32(4): 203-210.

Kurata, S.I. 2000. Selective activation of p38 MAPK cascade and mitotic arrest caused by low level oxidative stress. J. Biol. Chem. 275(31): 23413–23416.

Lacey, E.P. 1986. Onset of reproduction in plants: size- versus age-dependency. Trends Ecol. Evol. 1: 72–75.

Lang, G.A. 1994. Dormancy - the missing links: molecular studies and integration of regulatory plant and environmental interactions. HortScience: 29: 1255-1263.

Lang, G.A., Early, J.D., Martin, G.C., Darnell, R.L. 1987. Endo-, para- and ecodormancy: physiological terminology and classification for dormancy research. HortScience 22: 371-377.

Law, R.D., Suttle, J.C. 2003. Transient decreases in methylation at 5'-CCGG-3' sequences in potato (*Solanum tuberosum* L.) meristem DNA during progression of tubers through dormancy precede the resumption of sprout growth. Plant Mol. Biol. 51: 437–447.

Lawson, D. M., Hemmat, M., Weeden, N. F. 1995. The use of molecular markers to analyze the inheritance of morphological and developmental traits in apple. J. Am. Soc. Hortic. Sci. 120: 532–537.

Leike, H. 1965. Neuere Ergebnisse ü ber die Ruheperiode (Dormancy) der Gehölzknospen. Wiss. Z. Univ. Rostock, Math.-naturwiss. R. 14: 475-492.

Lenton, J.R., Perry, V.M., Saunders, P.F. 1972. Endogenous abscisic acid in relation to photoperiodically induced bud dormancy. Planta 106: 13-22.

Leung, J., Giraudat, J. 1998. Abscisic acid signal transduction. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49: 199-222.

Li, Z., Reighard, G.L., Abbott, A.G., Bielenberg, D.G. 2009. Dormancy-associated MADS genes from the EVG locus of peach [*Prunus persica* (L.) Batsch] have distinct seasonal and photoperiodic expression patterns. J. Exp. Bot. 60: 3521–3530

Liebhard, R., Kellerhals, M., Pfammatter, W., Jermini, M., Gessler, C. 2003b. Mapping quantitative physiological traits in apple (*Malus* × *domestica* Borkh.). Plant Mol. Biol. 52: 511-526.

Liebhard, R., Koller, B., Patocchi, A., Kellerhals, M., Pfammatter, W., Jermini, M., Gessler, C. 2003a. Mapping quantitative field resistance against apple scab in a 'Fiesta' × 'Discovery' progeny. Phytopathology 93:493-501.

MacDonald, J.E. 2000. The developmental basis of bud dormancy in 1-year-old Picea and Pseudostuga seedlings. In: eds., J.-D. Viemont and J. Crabbe. Dormancy in plants. New York, USA: CABI Publishing.

Martinez-Zapater, J. M., Coupland, G., Dean, C., Koornneef, M. 1994. The transition to flowering in *Arabidopsis*, pp. 403–434 in *Arabidopsis*, edited by E. M. Meyerowitz and C. R. Somerville. Cold Spring Harbor Laboratory Press, New York.

Mathiason, K., He, D., Grimplet, J., Venkateswari, J., Galbraith, D.W., Or, E., Fennell, A. 2009. Transcript profiling in Vitis riparia during CR fulfillment reveals coordination of gene expression patterns with optimized bud break. Funct. Integr. Genomics. 9(1): 81–96

Maurel, K., Leite, G.B., Bonhomme, M., Guilliot, A., Rageau, R., Pétel, G., Sakr, S. 2004. Trophic control of bud break in peach (Prunus persica) trees: a possible role of hexoses. Tree Physiol. 24: 579-588.

Mazzitelli, L., Hancock, R.D., Haupt, S., Walker, P.G., Pont, S.D., McNicol, J., Cardle, L., Morris, J., Viola, R., Brennan, R., Hedley, P.E., Taylor, M.A. 2007. Co-ordinated gene expression during phases of dormancy release in raspberry (*Rubus idaeus* L.) buds. J. Exp. Bot. 58(5): 1035–1045.

McGrath, K.C., Dombrecht, B., Manners, J.M., Schenk, P.M., Edgar, C.I., Maclean, D.J., Scheible, W.-R., Udvardi, M.K., Kazan, K. 2005. Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of *Arabidopsis* transcription factor gene expression. Plant Physiol. 139: 949-959.

Meijón, M., Feito, I., Valledor, L., Rodríguez, R., Cañal, M.J. 2010. Dynamics of DNA methylation and Histone H4 acetylation during floral bud differentiation in azalea. BMC Plant Biol. doi:10.1186/1471-2229-10-10.

Meskiene, I., Hirt, H. 2000. MAP kinase pathways: molecular plug-and-play chips for the cell. Plant Mol. Biol. 42: 791-806.

Moore, J.N., Janick, J. 1983. Methods in fruit breeding. Purdue Univ. Press, West Lafayette, Ind.

Mouradov, A., Cremer, F., Coupland, G. 2002. Control of flowering time: interacting pathways as a basis for diversity. Plant Cell 14: S111–S130.

Muller-Thurgau, H. 1885. Beitrag zur Erklärung der Ruheperioden der Pflanzen. Landw. Jahrb. 14: 851-907.

Nakano, T., Suzuki, K., Fujimura, T., Shinshi, H. 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol, 140: 411-432.

Neill, S.J., Desikan, R., Clarke, A., Hurst, R.D., Hancock, J.T. 2002. Hydrogen peroxide and nitric oxide as signalling molecules in plants. J. Exp Botany. 53: 1237-1247.

Ni, Z., Kim, E.D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q., Chen, Z.J. 2009. Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. Nature. 457(7227): 327-31.

Nir, G., Shulman, Y., Fanberstein, L., Lavee, S. 1986. Changes in the activity of catalase in relation to the dormancy of grapevine (*Vitis vinifera* L.) buds. Plant Physiol. 81: 1140–1142.

Nissila, P.C., Fuchigami, L.H. 1978. The relationship between vegetative maturity and the first stage of cold acclimation. J. Am. Soc. Hortic. Sci. 103: 710–711.

Nitsch, J.P. 1957. Growth responses of woody plants to photoperiodic stimuli. American Society for Horticultural Science Proceedings 70: 512-525.

Nooden, L. D., Weber, J. A. 1978. Environmental and hormonal control of dormancy in terminal buds of plants. Pages 221–268 in M. E. Cutter, ed. Dormancy and Developmental Arrest. New York: Academic Press.

Oakenfull, E.A., Riou-Khamlichi C., Murray, J.A.H. 2002. Plant D-type cyclins (CycDs) and the control of G1 progression. Philos. Trans. R. Soc. London B 357: 749-760.

Okubo, H. 2000. Growth cycle and dormancy in plants. In: eds., J.-D. Viemont and J. Crabbe. Dormancy in plants. New York, USA: CABI Publishing.

Ophir, R., Pang, X., Halaly, T., Venkateswari, J., Lavee, S., Galbraith D., Or E. 2009. Gene-expression profiling of grape bud response to two alternative dormancy-release stimuli expose possible links between impaired mitochondrial activity, hypoxia, ethylene-ABA interplay and cell enlargement. Plant Mol Biol. 71(4-5): 403-23.

Or, E., Vilozny, I., Eyal, Y., Ogrodovitch, A. 2000. The transduction of the signal for grape bud dormancy breaking induced by hydrogen cyanamide may involve the SNF-like protein kinase GDBRPK. Plant Mol. Biol. 43: 483–494.

Or, E., Vilozny, I., Fennell, A., Eyal, Y., Ogrodovitch, A. 2002. Dormancy in grape buds: Isolation and characterisation of catalase cDNA and analysis of its expression following chemical induction of bud dormancy release. Plant Sci. 162: 121–130.

Parmentier, C.M., Rowland, L.J., Line, M.J. 1998. Water status in relation to maintenance and release from dormancy in blueberry flower buds. J. Am. Soc. Hortic. Sci. 123(5): 762-769.

Pérez, F.J., Lira, W. 2005. Possible role of catalase in post-dormancy bud-break in grapevines. J. Plant Phisiol. 162: 301–308.

Pérez, F.J., Vergara, R., Rubio, S. 2008. H_2O_2 is involved in the dormancy-breaking effect of hydrogen cyanamide in grapevine buds. Plant Growth Regul. 55(2): 149-155. doi:10.1007/s10725-008-9269-4.

Perry, T.O. 1971. Dormancy of trees in winter. Science 171: 29-36.

Phillips, I.D., Wareing, P.F. 1958. Studies in dormancy of sycamore. I. Seasonal changes in the growth-substance content of the shoot. J. Exp. Bot. 9(27): 350-364.

Pitzschke, A., Hirt, H. 2009. Disentangling the Complexity of Mitogen-Activated Protein Kinases and Reactive Oxygen Species Signaling. Plant Physiol. 149: 606-615.

Powell, L.E. 1987. The hormonal control of bud and seed dormancy in woody plants. In: ed. P.J. Davies. Plant hormones and their role in plant growth and development. Boston: Martinus Nijhoff Publishers.

Quamme, H.A., Su, W.A., Veto, L.J. 1995. Anatomical features facilitating supercooling of the flower within the dormant peach flower bud. J. Am. Soc. Hortic. Sci. 120(5): 814-822

Reeves, P.A., He, Y., Schmitz, R.J., Amasino, R.M., Panella, L.W., Richards, C.M. 2007. Evolutionary conservation of the FLOWERING LOCUS C-mediated vernalization response: evidence from the sugar beet (*Beta vulgaris*). Genetics. 176:295–307

Rinne, P., Saarelainen, A., Junttila, O. 1994a. Growth cessation and bud dormancy in relation to ABA level in seedlings and coppice shoots of *Betula pubescens* as affected by a short photoperiod, water stress and chilling. Physiol Plant 90(3): 451-458.

Rinne, P., Tuominen, H., Junttila, O. 1994b. Seasonal changes in bud dormancy in relation to bud morphology, water and starch content, and abscisic acid concentration in adult trees of *Betula pubescens*. Tree Physiol. 14(6): 549-561.

Rinne, P., Welling, A., Kaikuranta, P. 1998. Onset of freezing tolerance in birch (*Betula pubescens* Ehrh.) involves LEA proteins and osmoregulation and is impaired in an ABA-deficient genotype. Plant Cell Environ. 21: 601-611.

Rinne, P.L.H., Kaikuranta, P., Van der Schoot, C. 2001. The shoot apical meristem restores its symplastic organization during chilling-induced release from dormancy. Plant J. 26: 249–264.

Roberts, C.J., Nelson, B., Marton, M.J, Stoughton, R., Meyer, M.R., Bennett, H.A., He, Y.D., Dai, H., Walker, W.L., Hughes, T.R., Tyers, M., Boone, C., Friend, S.H. 2000. Signaling and circuitry of multiple MAPK pathways revealed by matrix of global gene expression profiles. Science 287:873–880.

Rodriguez, A.J., Sherman, W.B., Scorza, R., Wisniewski, M., Okie, W.R. 1994. 'Evergreen' peach, its inheritance and dormant behaviour. J. Am. Soc. Hortic. Sci. 119(4): 789-792.

Rohde, A., Ruttink, T., Hostyn, V., Sterck, L., Van Driessche, K., Boerjan, W. 2007. Gene expression during the induction, maintenance, and release of dormancy in apical buds of poplar. J. Exp. Bot. 58(15-16): 4047–4060.

Rohde, A., Van Montagu, M., Inze, D., Boerjan, W. 1997. Factors regulating the expression of cell cycle genes in individual buds of Populus. Planta 201: 43-52.

Romberger, J.A. 1963. Meristems, growth, and development in woody plants. Forest Serv. Tech. Bul. 1293. USDA.

Rowland, L.J., Arora, R. 1997. Proteins related to endodormancy (rest) in woody perennials. Plant Science 126: 119–144.

Rowland, L.J., Mehra, S., Dhanaraj, A., Ogden, E.L., Arora, R. 2003. Identification of molecular markers associated with cold tolerance in blueberry. Acta Hortculturae 625: 59-69.

Rowland, L.J., Ogden, E.L., Arora, R., Lim, C.C., Lehman, J.S., Levi, A., Panta, G.R. 1999. Use of blueberry to study genetic control of CR and cold hardiness in woody perennials. HortScience 34: 1185-1191.

Ruiz, D., Campoy, J.A., Egea, J. 2007. Chilling and heat requirements of apricot cultivars for flowering. Environ. Exp. Bot. 61(3): 254–263.

Ruttink ,T., Arend, M., Morreel, K., Storme, V., Rombauts, S., Fromm, J., Bhalerao, R.P., Boerjan, W., Rohde, A. 2007. A molecular timetable for apical bud formation and dormancy induction in poplar. Plant Cell 19: 2370–2390.

Sakai, A. 1979. Freeezing avoidance mechanism in primordial shoots of conifer buds. Plant Cell Physiol. 20: 1381-1390.

Salzman, R.A., Bressan, R.A., Hasegawa, P.M., Ashworth, E.N., Bordelon, B.P. 1996. Programmed accumulation of LEA-like proteins during desiccation and cold acclimation of overwintering grape buds. Plant Cell Environ. 19: 713 720.

Samish, R.M. 1954. Dormancy in woody plants. Annu Rev Plant Physiol 5: 183-204.

Saunders, P. 1978. Phytohormones and bud dormancy. In: Phytohormones and Related Compounds. A Comprehensive Treatise. (Letham D.S., Goodwin P.B. & Higgins T.J., eds.), Vol. II. Elsevier, Amsterdam, pp. 423-445.

Saure, M.C. 1985. Dormancy release in deciduous fruit trees. In: ed. J. Janick. Horticultural Reviews, volume 7. Westport, Conneticut: AVI Publishing Company

Sauter, M. 1997. Differential expression of a CAK (cdc2-activating kinase)-like protein kinase, cyclins and cdc2 genes from rice during the cell cycle and in response to gibberellin. Plant J. 11: 181–190.

Segura, V., Cilas, C., Laurens, F., Costes, E. 2006. Phenotyping progenies for complex architectural traits: a strategy for 1-year-old apple trees (*Malus* × *domestica* Borkh.). Tree Genet. Genom. 2: 140–151.

Sharp, R.E., LeNoble, M.E., Else, M.A., Thorne, E.T., and Gherardi, F. 2000. Endogenous ABA maintains shoot growth in tomato independently of effects on plant water balance: Evidence for an interaction with ethylene. J. Exp. Bot. 51: 1575–1584.

Sheldon, C.C., Hills, M.J., Lister, C., Dean, C., Dennis, E.S., Peacock, W.J. 2008. Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. Proc. Natl Acad. Sci. U.S.A. 105: 2214–2219.

Shen, W.H. 2002. The plant E2F-Rb pathway and epigenetic control. Trends Plant Sci. 7: 505–511.

Shimizu-Sato, S., Mori, H. 2001. Control of outgrowth and dormancy in axillary buds. Plant Physiol. 127: 1405–1413.

Siller-Cepeda, J.H., Fuchigami, L.H., Chen, T.H.H. 1992. Glutathione content in peach buds in relation to development and release of rest. Plant Cell Physiol. 33: 867–872.

Silva, C., Garcia-Mas, J., Sánchez, A.M., Arús, P., Oliveira, M.M. 2005. Looking into flowering time in almond (*Prunus dulcis* (Mill) D. A. Webb): the candidate gene approach. Theor. Appl. Genet. 110(5): 959-68.

Simpson, G.M. 1990. Seed dormancy in grasses. Cambridge University Press: Cambridge.

Stirnberg, P., Van De Sande, K., Leyser, H.M. 2002. MAX1 and MAX2 control shoot branching in *Arabidopsis*. Development 129: 1131–1141.

Strauss, M., Kauder, F., Peisker, M., Sonnewald, U., Conrad, U., Heineke, D. 2001. Expression of an abscisic acid-binding single-chain antibody influences the subcellular distribution of abscisic acid and leads to developmental changes in transgenic potato plants. Planta 213: 361-369.

Sung, S., Amasino, R.M. 2004. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. Nature 427: 159-164.

Suttle, J. C. 1998. Involvement of ethylene in potato microtuber dormancy. Plant Physiol. 118: 843-848.

Tadege, M., Sheldon, C. C., Helliwell, C. A., Stoutjesdijk, P., Dennis, E.S., Peacock, W.J. 2001 Control of flowering time by FLC orthologues in *Brassica napus*. Plant J. 28: 545-553.

Tamura, F., Tanabe, K., Ikeda, T. 1993. Relationship between intensity of bud dormancy and level of ABA in Japanese pear 'Nijisseiki'. Journal of the Japanese Society for Horticultural Science 62(1): 75-81.

Tamura, F., Tanabe, K., Itai, A. 1998. Characteristics of endodormancy in *Pyrus* species. J. Japan. Soc. Hort. Sci. 67 (Suppl. 1): 97.

Tanksley, S.D., Hewitt, D. 1988 Use of molecular markers in breeding for soluble solids in tomato-a re-examination. Theor. Appl. Genet. 75: 811-823.

Tanksley, S.D., Young, N.D., Paterson, A.H., Bonierbale, M.W. 1989. RFLP mapping in plant breeding-new tools for an old science. Biotechnology 7: 257-264.

Thompson, M.M., Smith, D.C., Burgess, J.E. 1985. Nondormant mutants in a temperate tree species, *Corylus avellana* L. Theor. Appl. Genet. 70: 687–692.

Trevaskis, B., Hemming, M.N., Dennis, E.S., Peacock, W.J. 2007. The molecular basis of vernalization-induced flowering in cereals. Trends Plant Sci. 12(8): 352-357.

Van der Schoot, C. 1996. Dormancy and symplasmic networking at the shoot apical meristem. In: ed., G.A. Lang. Plant dormancy, physiology, biochemistry and molecular biology. Wallingford, UK: CAB International.

Vegis A. 1964. Dormancy in higher plants. Annu. Rev. Plant Physiol. 15: 185-224.

Verde, I., Quarta, R., Cedrola, C., Dettori, M.T. 2002. QTL analysis of agronomic traits in a BC1 peach population. Acta Hortic 592: 291–297.

Wang, D., Karle, R., Iezzoni, A. 2000. QTL analysis of flower and fruit traits in sour cherry. Theor. Appl. Genet. 100:535-544.

Wang, H., Fowke, L.C., Crosby, W.L. 1997. A plant cyclin-dependent kinase inhibitor gene. Nature 386: 451-452.

Wang, S.Y., Faust M. 1994. Changes in the antioxidant system associated with bud break in ‘Anna’ apple (*Malus* x *domestica* Borkh.) buds, J. Am. Soc. Hort. Sci. 119: 735–741.

Wang, S.Y., Jiao, H.J., Faust, M. 1991. Changes in metabolic enzyme activities during thidiazuron-induced lateral bud break of apple. HortScience 82: 231–236.

Wang, Y., Georgi, L.L., Reighard, G.L., Scorza, R., Abbott, A.G. 2002. Genetic mapping of the evergrowing gene in peach [*Prunus persica* (L.) Batsch]. J Hered. 93(5): 352-8.

Wareing, P.F. 1956. Photoperiodism in woody plants. Annu. Rev. Plant Physiol. 7: 191-214.

Wareing, P.F. 1969. Germination and dormancy. In Physiology of Plant Growth and Development. Ed. M.B. Wilkins. McGrawhill, London, pp 605–644.

Weatherwax, S.C., Ong, M.S., Degenhardt, J., Bray, E.A., Tobin, E.M.1996. The interaction of light and abscisic acid in the regulation of plant gene expression. Plant Physiol. 111: 363–370.

Welling, A., Kaikuranta, P., Rinne, P. 1997. Photoperiodic induction of dormancy and freezing tolerance in *Betula pubescens*. Involvement of ABA and dehydrins. Physiol Plant 100(1): 119-125.

Wesselingh, R. A., De Jong, T. J., Klinkhamer, P. G. L., Vandijk, M. J., Schlatmann, E. G. M. 1993. Geographical variation in threshold size for flowering in Cynoglossum officinale. Acta Botanica Neerlandica 42: 81–91.

Wilen, R.W., Fu, P., Robertson, A.J., Abrams, S.R., Low, N.H., Gusata, L.V. 1996. An abscisic acid analog inhibits abscisic acid-induced freezing tolerance and protein accumulation, but not abscisic acid-induced sucrose uptake in a bromegrass (*Bromus inermis Leyss*) cell culture. Planta 200: 138-143.

Wood, B.W. 1993. Hydrogen cyanamide advances pecan budbreak and harvesting. J. Amer. Soc. Hort. Sci. 118: 690-693.

Xiangdong, F., Harberd, N.P. 2003. Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. Nature 421: 740–743.

Yang, Z., Tian, L., Latoszek-Green, M., Brown, D., Wu, K. 2005. *Arabidopsis* ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. Plant Mol. Biol. 58(4): 585-96.

CHAPTER TWO

GENETIC LINKAGE MAPPING FOR MOLECULAR DISSECTION OF CHILLING REQUIREMENT AND BUD BREAK IN APRICOT (*PRUNUS ARMENIACA* L.)

Bode A. Olukolu,[a] Taly Trainin,[d] Shenghua Fan,[a] Chittaranjan Kole,[a] Douglas G. Bielenberg,[b, c] Gregory L. Reighard,[b] Albert G. Abbott,[a] and Doron Holland[d]

[a]Genetics and Biochemistry, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA.
[b]Horticulture, Clemson University, 170 Poole Agriculture Center, Clemson, SC 29634, USA.
[c]Biological Sciences, Clemson University, 132 Long Hall, Clemson, SC 29634, USA.
[d]Fruit Sciences, Newe Ya'ar Research Center, Agricultural Research Organization, P.O. Box 1021, Ramat Yishay 30095, Israel.

Corresponding authors: A.G. Abbott (e-mail: aalbert@clemson.edu) and D. Holland (e-mail: vhhollan@volcani.agri.gov.il).

As it appears in Genome 2009 Oct,52(10):819-828

Received 14 April 2009. Accepted 3 July 2009. Published on the NRC Research Press Web site at http://genome.nrc.ca on 16 September 2009.

Corresponding Editor: P. Gulick.

The maintenance and phenotyping of the mapping population was conducted by Dr. Doron Holland and Taly Trainin, while genotyping of the mapping population, marker and data analysis, map construction and QTL mapping was performed by Bode A. Olukolu.

Abstract

Commercial production of apricot is severely affected by sensitivity to climatic conditions, an adaptive feature essential for cycling between vegetative or floral growth and dormancy. Yield losses are due to late winter or early spring frosts and inhibited vegetative or floral growth caused by unfulfilled chilling requirement (CR). Two apricot cultivars, Perfection and A.1740, were selected for high and low CR, respectively, to develop a mapping population of F_1 individuals using a two-way pseudo-testcross mapping strategy. High-density male and female maps were constructed using, respectively, 655 and 592 markers (SSR and AFLP) spanning 550.6 and 454.9 cM with average marker intervals of 0.84 and 0.77 cM. CR was evaluated in two seasons on potted trees forced to break buds after cold treatments ranging from 100 to 900 h. A total of 12 putative CR quantitative trait loci (QTLs) were detected on six linkage groups using composite interval mapping and a simultaneous multiple regression fit. QTL main effects of additive and additive × additive interactions accounted for 58.5% ± 6.7% and 66.1% ± 5.8% of the total phenotypic variance in the Perfection and A.1740 maps, respectively. We report two apricot high-density maps and QTLs corresponding to map positions of differentially expressed transcripts and suggested candidate genes controlling CR.

Key words: dormancy, bud break, peach, QTL.

Introduction

Apricots, like other temperate stone fruit crops, are grown in climates with well-differentiated seasons where species have adapted mechanisms to survive low winter temperatures and frost damage (Ruiz *et al.* 2007). Survival and reproduction of perennial fruit trees require adaptation to the environment by synchronizing development with the cyclic climatic conditions (Dietrichson 1964, Hill *et al.* 1998).

A prominent adaptive feature, dormancy, involves a gradual and progressive process through autumn until a deep rest state is reached (Hatch and Walker 1969, Lang *et al.* 1987). Transition into dormancy is triggered by environmental cues such as photoperiod, cold, or drought (Arora *et al.* 2003, Rohde *et al.* 2007). Cultivars introduced into a climate where low winter temperatures are not sufficient for breaking dormancy exhibit adverse effects with regard to vegetative or floral growth and fruit-bearing capacity (Coville 1920). On the contrary, cultivars with low chilling requirement (CR) that are grown in cold-winter climates quickly complete CR and bloom too early, leading to yield losses due to late winter or early spring frosts (Scorza and Okie 1990).

The poor understanding of the genetics controlling CR in fruiting trees is most likely due to their long generation time and the complex mode of inheritance of characters related to plant growth and habit. However, the state of the art in fruit tree genetics and genomics affords a unique opportunity to characterize CR. Currently we know that there is a high level of conservation of synteny among the genomes of different *Prunus* species (Dirlewanger *et al.* 2004), facilitating transferability of genomic and genetic resources.

Capitalizing on this fact, a *Prunus* reference genetic and physical map populated with numerous marker types, including those based on expressed sequence tags (ESTs), was constructed (Aranzana *et al.* 2003, Zhebentyayeva *et al.* 2008). With a number of important genes and quantitative trait loci (QTLs) already mapped in *Prunus* species, marker-assisted selection is now becoming a reality in some breeding programs (Dirlewanger *et al.* 2004). Numerous genetic linkage maps marking important characters have been constructed for several *Prunus* species including apricot, peach and related wild species, almond, plum, and cherry (Genome Database for Rosaceae, Jung *et al.* 2008; available at http://www.bioinfo.wsu.edu/gdr/); however, maps identifying genes or QTLs controlling CR in *Prunus* have not been published. The *Prunus* reference map could in principle facilitate discovery of CR-related genes if appropriate mapping populations exist in *Prunus* species.

In this communication, we report the underlying genetic basis of CR in apricot through the development and analysis of a mapping population segregating for CR. Using this mapping population and a 2-way pseudo-testcross mapping strategy, 2 high-density genetic linkage maps with locations of putative QTLs for CR were developed. Twelve QTLs for CR were located on 2 maps generated from high and low CR parents. Because of the utilization of core *Prunus* map markers that are integrated on the peach physical map, the physical map location of these QTL intervals is available and potential candidate gene ESTs have been identified.

Materials and methods

Mapping population

A 2-way pseudo-testcross population consisting of 100 F_1 individuals was developed from crosses between 2 cultivars (Perfection as male and A.1740 as female) with contrasting differences for CR and other traits. The parents and progenies were maintained at the Newe Ya'ar Research Center of the Agricultural Research Organization in Israel. Routine methods of bagging and pollination were followed (Zeaser 2001).

DNA extraction

Genomic DNA was extracted from fresh young leaves using a hexadecyltrimethylammonium bromide-polyvinyl pyrrolidone (CTAB-PVP) method as described in Porebski *et al.* (1997). This procedure is a modification of the CTAB protocol of Eldredge *et al.* (1992) for plants containing high amounts of polysaccharides and polyphenolic compounds. DNA concentrations were quantified by a minifluorimeter (TKO100, Hoefer Scientific).

SSR markers

Most of the simple sequence repeat (SSR) markers used were from the *Prunus* anchor marker set originally developed for peach (Aranzana *et al.* 2003). These markers were selected based on uniform distribution across the linkage groups of the *Prunus* reference map (Dirlewanger *et al.* 2004) to establish a framework map for apricot for studies of genome homology. These SSR markers were screened for polymorphism between the 2

parents and segregation among 6 randomly selected individuals from the pseudo-testcross mapping population. The specific primer pairs, amplification conditions, radioactive labeling, and polyacrylamide gel electrophoresis were employed as described in Combes *et al.* (2000).

PCR reactions of 10 µL contained 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 0.2 mmol/L of each dNTP, 10 pmol of labeled forward primer and 10 pmol of unlabelled reverse primer, 10 ng of genomic DNA, and 0.5 U of *Taq* polymerase (Life Technologies, Inc.). Amplification was conducted with initial denaturing at 94 °C for 3 min followed by 35 cycles of denaturing at 94 °C for 1 min, annealing at 49–56 °C for 1 min, and primer extension at 72 °C for 1 min, and a final extension at 72 °C for 5 min. Forward primers were end-radiolabeled with [γ-^{33}P]ATP (PerkinElmer) and T4 polynucleotide kinase (Promega). PCR products were separated on 6% denaturing polyacrylamide gels which were vacuum-dried, exposed to X-Omat blue XB-1 films (Kodak), and developed after 1–7 days. A DNA standard ladder (Promega, fmol DNA cycle sequencing system) was loaded alongside the samples to determine the sizes of the amplified fragments.

AFLP protocol and markers

Amplified fragment length polymorphism (AFLP) marker analysis was performed as described in Vos *et al.* (1995). An aliquot of 200 ng of genomic DNA was digested with the restriction enzymes *Eco*RI and *Mse*I. Restriction fragments from the digest were ligated to *Eco*RI and *Mse*I adapters and diluted 10-fold for pre-amplification. The pre-

amplification reaction was set up using standard E and M primers corresponding to the *Eco*RI and *Mse*I adapters, respectively, and containing one selective nucleotide (E+A and M+C) at the 3′ end. The pre-amplification reaction mixture was diluted 10-fold and used for selective amplification using various combinations of E primers with 1 additional selective nucleotide and M primers with 2 additional selective nucleotides. The 256 primer combinations initially screened among the parents and 6 progenies include all combinations from EAA to ETT and MCAA to MCTT. Following screening, primer combinations were chosen based on the polymorphism information content (PIC) and used for genotyping the mapping population.

Pre-amplification PCR conditions included 20 cycles of denaturing at 94 °C for 30 s, annealing at 56 °C for 1 min, and primer extension at 72 °C for 1 min, followed by a final extension at 72 °C for 5 min. Selective PCR conditions included 13 cycles of denaturing at 94 °C for 30 s, annealing at 65 °C (decreasing by 0.7 °C per cycle) for 1 min, and primer extension at 72 °C for 1 min; 24 cycles of denaturing at 94 °C for 30 s, annealing at 56 °C for 30 s, and primer extension at 72 °C for 1 min; and a final extension at 72 °C for 5 min. End-radiolabeling of E primers and separation and detection of PCR products were conducted as described under SSR markers.

Genotyping and nomenclature of markers

Genotyped marker data were obtained from visual scoring of the banding patterns. Alleles detected with primers that produced multiple loci were labeled with an alphabetic suffix for SSR markers and a numeric suffix for AFLP markers. The segregating loci

obtained from AFLP and SSR analysis were categorized into 6 classes. These included loci that are heterozygous in both parents and segregate in a 1:1:1:1 ratio, involving 4 alleles (ab × cd) and 3 non-null alleles (ef × eg); loci that are heterozygous in both parents and segregate in a 1:2:1 ratio, with both parents having the same genotype of codominant alleles (hk × hk); loci that are heterozygous in both parents and segregate in a 3:1 ratio (some mapped as dominant bridge markers); and loci that are in a testcross configuration between the parents and segregate in a 1:1 ratio (dominant markers), comprising loci that are heterozygous in the female parent and homozygous in the male parent (lm × ll) and those that are heterozygous in the male parent and homozygous in the female parent (nn × np). Only 39 AFLP markers were scored as codominant for bands showing polymorphism and intensity differences between heterozygous and homozygous allelic states (Castiglioni *et al.* 1999).

Genotypic data for each parental map comprised markers segregating specifically in a parent as well as the bridge markers, which served as anchors to align linkage groups between the 2 parental maps. Raw genotypic data were recorded without any previous knowledge of phase relationship and inheritance, as typical of 2-way pseudo-testcross mapping populations.

Linkage analysis and map construction

Linkage map construction was performed according to the procedures described in Lodhi *et al.* (1995) for 2-way pseudo-testcross populations (Grattapaglia and Sederoff 1994, Maliepaard *et al.* 1997, Lambert *et al.* 2004) using JoinMap version 3.0 (Van Ooijen and

Voorrips 2001). Genotypic data were prepared based on the JoinMap CP (cross-pollinating) function for 2 separate parental maps. JoinMap data analysis tools were used to screen for missing data points, segregation distortion, and similarity between loci and individuals. Segregation distortion was determined by χ^2 analysis. Linkage analysis was performed using a maximum recombination fraction of 0.40 and minimum critical logarithm of odds (LOD) scores of 6.0 and 7.3 for the Perfection and A.1740 maps, respectively. Marker distances were calculated based on the Kosambi mapping function (Kosambi 1944).

Since the linkage phases in a 2-way pseudo-testcross are not known beforehand, a first round of linkage analysis was done to determine loci out of linkage phase, followed by a second round of analysis that included dummy variables (alternative linkage phase) of loci that were not in linkage phase. Map files of both parental maps were used to draw the linkage map in MapChart 2.2 (Voorrips 2002).

Algorithms of the mapping software used do not take into account the 2-way pseudo-testcross population structure for the estimation of actual centimorgan marker intervals (Wu *et al.* 2002, 2007). For instance, the CP function in JoinMap (Van Ooijen and Voorrips 2001) and the F_2 pseudo-testcross population mapping options in MAPMAKER (Lander *et al.* 1987) do not yield exactly the same centimorgan distances. In fact, they do not compute an actual centimorgan distance unit, and JoinMap has been shown to produce a shorter map than MAPMAKER (Van Ooijen *et al.* 1994, Qi *et al.* 1996). Therefore, they cannot be compared directly for the distance estimates obtained from the

Prunus maps based on F_2 and backcross populations. To resolve this conundrum, we utilized IRILmap version 1.1 (Falque 2005), which is capable of computing actual centimorgan distances from recombination fraction per meiosis (r_n). It takes into account the *n* generations of inter-mating by reversing Winkler's formula (Winkler *et al.* 2003) through iteration and finally reapplies a distance function. Since the *n* generations cannot be determined for the highly heterozygous apricot parents, the generation of inter-mating ($n = 4$) that produced expected map lengths estimated from existing *Prunus* maps was utilized.

Evaluation of chilling requirement

Chilling requirement evaluation was performed at the Newe Ya'ar Research Center in Israel under controlled conditions on the Perfection × A.1740 mapping population. Seeds of matured fruits from the cross were washed with water and treated with an antifungal solution of 0.25% Merpan 48 containing 480 g/kg Captan. Seeds were then placed in wet vermiculite at 4–5 °C until the beginning of germination. Upon germination, the seedlings were placed in 1 L plastic pots containing garden soil inside growth chambers. The growth chambers were set at 24 °C and diurnal cycles of 8 h dark and 16 h light in artificial light. Young plants of about 0.5 m length were planted in the field at Newe Ya'ar.

Three-year-old flowering trees developed from the planted seedlings were used as the source material for stem cuttings with buds. Cuttings harvested from individual trees were top-grafted on 1-year-old plum rootstock (Mariana 2624) with low CR. Leaves were

stripped in late November following the onset of sufficient low temperatures in mid-November, marked by growth cessation and leaf senescence. Following defoliation, pots were placed in walk-in cold chambers at 6 °C, and 2 replicates of each individual were removed at 200 h intervals ranging from 200 to 600 h in 2007 and 100 to 900 h in 2008, with the exception of the 700 chilling hour treatment. Chilled potted trees were transferred to a naturally lit greenhouse at day and night temperatures of 25 °C and 13 °C, respectively, to force bud break under normal photoperiod and irradiance.

Due to limitation of space in the controlled cold chambers, it was not feasible to sample large numbers of apricot trees with multiple replicates in the same year. Subsequently, intervals and the range in chilling hours in each year do not adequately represent the phenotypic classes segregating in the population. Pooling data from both years was required to reveal an all-inclusive spectrum of recombinant genotypes. The emphasis of the study was on vegetative bud break and the data were expressed as the time in forcing conditions ("days in forcing"). From analysis of the data in the greenhouse and comparison with similar data from the field, we concluded that the best variable expressing the differences in CR in the population is days to first vegetative bud opening followed by bud flush. This variable was used as a basis to determine CR for each progeny in the following manner: the chilling regime in which bud opening occurred within 15 days in forcing was set as the minimum amount of chilling hours required for bud break. Quantification of CR (chill accumulated in cold chamber) was expressed in chilling hours (Weinberger 1950).

Quantative trait loci analysis

Quantitative trait loci for CR were detected using composite interval mapping (CIM, Zeng 1994) and by integrating genetic information and accumulated chilling hour data using the PC version of PlabQTL version 1.2 (Utz and Melchinger 1996). The PlabQTL algorithm was proposed to improve precision of QTL mapping (Zeng 1994). The CIM employed the cov SELECT option of PlabQTL, which uses a forward stepwise multiple regression to select cofactors automatically. The LOD curve of the PlabQTL multiple regression is similar to that of maximum likelihood-based programs (i.e., MapQTL and MAPMAKER/QTL), though multiple regression-based interval mapping is more robust against non-normality, is statistically well known, and underestimates R^2 (Haley and Knott 1992).

Different genetic models were compared based on Akaike's information criterion and Bayesian information criterion values (Hjorth 1994) to determine the best regression fit. The LOD curves were created by scanning at 1 cM intervals, while a permutation test (1000 resamplings) was performed to determine the critical LOD score appropriate to empirically identify a putative QTL with a genome-wide error at a 0.05 confidence level (Churchill and Doerge 1994). Subsequently, the detected QTLs and their estimated map positions were verified using a simultaneous multiple regression, which accounts for effects of other linked QTLs on a chromosome (Zeng 1993). The phenotypic variance explained by each QTL (R^2) was calculated as the square of the correlation coefficient from the final multiple regression model (Utz and Melchinger 1995, 1996). To examine

the interaction between detected QTLs, the general linear model of variance analysis was implemented.

Results

Molecular marker analysis

Out of a total of 275 SSR primer pairs screened, 225 produced amplification products and 53 were selected based on their uniform distribution over the *Prunus* reference map (Dirlewanger *et al.* 2004) for genotyping 94 pseudo-testcross progenies. After genotyping with the 53 SSR markers, only 43 showed linkage and were mapped (Table 2.1). Segregation patterns of 256 AFLP primer combinations among 6 randomly selected progenies generated 2253 segregating loci (Appendix A). The highest number of loci per primer combination (30 loci) was observed using primers with AT-rich selective nucleotides, while the lowest frequencies of loci per primer combination (2 loci) were observed using primers with GC-rich selective nucleotides. For the parent Perfection, 900 AFLP loci were analyzed, while for A.1740, 716 AFLP loci were analyzed (Table 2.2). The proportion of segregation distortion observed in Perfection was estimated at 10.07% at $P < 0.01$ and 24.60% at $P < 0.05$, while for A.1740 it was 4.97% at $P < 0.01$ and 13.98% at $P < 0.05$.

Table 2.1: SSR markers mapped on Perfection and A.1740 parental maps.

Locus	Segregation type	Classes	Position (cM)		χ^2
			Perfection	A.1740	
CPSCT042†	<abxcd>	[ac:ad:bc:bd]	LG7:43.3	LG7:33.1	10.2**
Taly	<abxcd>	[ac:ad:bc:bd]	-	LG2:36.7	1 ns
BPPCT007†	<efxeg>	[ee:ef:eg:fg]	LG3:0.0	LG3:0.0	1.6 ns
BPPCT025†	<efxeg>	[ee:ef:eg:fg]	LG6:43.2	LG6:43.2	3.6 ns
CPDCT045†	<efxeg>	[ee:ef:eg:fg]	LG4:14.8	LG4:21.4	0.5 ns
CPPCT026†	<efxeg>	[ee:ef:eg:fg]	LG1:53.0	LG1:51.3	4 ns
PceGA025†	<efxeg>	[ee:ef:eg:fg]	LG5:33.8	LG5:28.8	11.8***
PceGA034†	<efxeg>	[ee:ef:eg:fg]	LG2:57.1	LG2:42.1	1.5 ns
Pchmgs001†	<efxeg>	[ee:ef:eg:fg]	LG2:40.9	LG2:31.3	2.5 ns
BPPCT013a†	<hkxhk>	[hh:hk:kk]	LG2:29.9	LG2:22.3	0.2 ns
CPPCT034a†	<hkxhk>	[hh+hk+h-:kk]	LG1:38.7	LG1:45.9	2.7 ns
CPPCT034b	<hkxhk>	[hh+hk+h-:kk]	-	LG1:47.7	0.6 ns
EPDCU2862†	<hkxhk>	[hh:hk:kk]	LG1:63.3	LG1:64.0	3.1 ns
UDP96-005a†	<hkxhk>	[hh+hk+h-:kk]	LG1:38.9	LG1:37.4	1.1 ns
BPPCT004	<lmxll>	[ll:lm]	LG2:23.9	-	1.7 ns
BPPCT028	<lmxll>	[ll:lm]	LG1:76.1	-	0.7 ns
BPPCT030	<lmxll>	[ll:lm]	LG2: 42.2	-	0.0 ns
BPPCT040	<lmxll>	[ll:lm]	LG4:4.7	-	0 ns
CPDCT025	<lmxll>	[ll:lm]	LG3:54.3	-	1.1 ns
CPDCT034	<lmxll>	[ll:lm]	LG8:22.0	-	0.0 ns
CPSCT044	<lmxll>	[ll:lm]	LG2: 28.5	-	1.2 ns
EPDCU3083b	<lmxll>	[ll:lm]	LG3:44.7	-	0.0 ns
EPDCU3454	<lmxll>	[ll:lm]	LG8:42.2	-	0.7 ns
Pchcms002	<lmxll>	[ll:lm]	LG7:43.6	-	0.8 ns
Pchgms044a	<lmxll>	[ll:lm]	LG3:60.1	-	0.0 ns
Pchmgs005	<lmxll>	[ll:lm]	LG4:16.4	-	2.3 ns
SSRM6a	<lmxll>	[ll:lm]	LG8:30.9	-	0.4 ns
UDA002	<lmxll>	[ll:lm]	LG3:45.6	-	0.1 ns
UDA011a	<lmxll>	[ll:lm]	LG8:0.0	-	1.1 ns
UDA011b	<lmxll>	[ll:lm]	LG3:28.9	-	0.0 ns
UDP97-403	<lmxll>	[ll:lm]	LG3:17.2	-	0.2 ns
UDP98-024	<lmxll>	[ll:lm]	LG4:8.9	-	0.3 ns
UDP98-406	<lmxll>	[ll:lm]	LG2: 55.6	-	0.1 ns
UDP98-409a	<lmxll>	[ll:lm]	LG8:38.3	-	0.1 ns
UDP98-409b	<lmxll>	[ll:lm]	LG2: 37.0	-	0.9 ns
UDP98-412	<lmxll>	[ll:lm]	LG6:58.8	-	0.2 ns
BPPCT039	<nnxnp>	[nn:np]	-	LG3:17.6	0.0 ns
EPDCU3083a	<nnxnp>	[nn:np]	-	LG3:21.2	0.0 ns
Pchgms044b	<nnxnp>	[nn:np]	-	LG7:30.5	0.0 ns
SSRM2b	<nnxnp>	[nn:np]	-	LG7:41.0	1.5 ns
UDP96-001	<nnxnp>	[nn:np]	-	LG6:20.9	1.7 ns
UDP96-005b	<nnxnp>	[nn:np]	-	LG1:38.8	0.7 ns
UDP97-401	<nnxnp>	[nn:np]	-	LG5:21.1	0.4 ns

Note: LG, linkage group. χ^2: Chi-square values for expected Mendelian segregation ratio (** and *** denote significance level at 0.01 and P < 0.001, respectively). Segregation types abxcd, efxeg and hkxhk depict segregation in both parents, while lmxll and nnxnp depict segregation in Perfection and A.1740, respectively.

†Bridge-markers have corresponding map positons on both parental maps.

Table 2.2: AFLP and SSR marker analysis.

Features	Perfection		A.1740		
	Dominant	Codominant	Dominant	Codominant	Total
Segregating AFLP loci	879[a]	21	695[a]	21	1252
Segregating SSR loci	25	18	10	18	53
AFLP loci after χ^2-test	794[b]	11	668[b]	11	1136
SSR loci after χ^2-test	25	18	10	18	53
AFLP loci mapped	(422) 610[c]	11	(292) 560[d]	11	994
SSR loci mapped	22	12	7	14	43
Total loci in linkage map	632	23	567	25	

Note: Dominant AFLP markers also include bands heterozygous and segregating in both parents, with 165 of them serving as brigde markers: [a](343), [b](337), [c](188) and [d](268). Numbers in paratheses indicate AFLP markers unique to each parental map (i.e., not bridge markers). AFLP and SSR loci that did not map were due to lack of linkage, and unequal numbers bridge markers were not mapped between the 2 parental maps. [e](176 served as bridge-markers). Total AFLP markers: 422 + (188 – 165) + 292 + (268 – 165) + 165 +11 = 1016.

Linkage analysis and map construction

Of the 1059 markers (1016 AFLP and 43 SSR markers) mapped on the linkage maps, 176 AFLP (165 dominant and 11 codominant) and 12 SSR markers (18.1% of mapped markers) were bridge markers mapped in both parents (Table 2.2). These bridge markers allowed for transferability of markers and alignment of homologous linkage groups (Figs. 2.1 and 2.2). Totals of 632 and 567 dominant markers were mapped on the Perfection and A.1740 parental maps, respectively (Table 2.2). Some SSR primers detected multiple loci that mapped within close genomic vicinity. The average marker intervals were 0.84 and 0.77 cM for the Perfection and A.1740 maps, respectively, while the largest gap was only 7.5 cM. The map lengths were 550.6 cM and 454.9 cM for the Perfection and A.1740 maps, respectively (Table 2.3).

To further validate marker order on our maps, the 2 parental maps were compared with the published *Prunus* reference map (Dirlewanger *et al.* 2004) to assess colinearity. All 8 homologous linkage groups (LGs) of our maps correspond to the homeologous linkage groups of the *Prunus* reference map based on alignments with anchor SSR markers (Appendix B1, B2, B3 and B4). All 32 anchor SSR markers mapped in our maps showed perfect conservation of synteny with the *Prunus* reference map (Appendix B) except for 2 markers on LG1 and LG2 (CPPCT034 and BPPCT040, respectively). Comparison with other apricot maps showed that the incongruence at the CPPCT034 loci was consistent between our map and the map by Dondini *et al.* (2007), thus showing a slight variation within a small genetic distance between the peach and apricot genomes.

Fig. 2.1: Genetic linkage maps (linkage groups 1, 2, 3, and 4) derived from a Perfection × A.1740 cross oriented with the *Prunus* reference map using SSR markers (gray text; purple in the Web version). AFLP bridge markers (bold black text) confirm colinearity between parental maps. Detected QTLs are indicated by solid gray (purple in the Web version) fills and bars, with common QTLs in crosshatch fill. The asterisks show the χ^2 *p* levels of significance (*, 0.05; **, 0.01; ***, 0.001).

Fig. 2.2: Genetic linkage maps (linkage groups 5, 6, 7, and 8) derived from a Perfection × A.1740 cross oriented with the *Prunus* reference map using SSR markers (gray text; purple in the Web version). AFLP bridge markers (bold black text) confirm colinearity between parental maps. Detected QTLs are indicated by solid gray (purple in the Web version) fills and bars, with common QTLs in crosshatch fill. The asterisks show the χ^2 *p* levels of significance (*, 0.05; **, 0.01; ***, 0.001).

Table 2.3: Data on Perfection and A.1740 maps: number of mapped markers, linkage group and map lengths, marker density, and marker intervals (gaps).

LG	Perfection				A.1740			
	No. of loci	Map length, cM	Mean Distance, cM	Gaps > 2 cM (*n*)	No. of loci	Map Length, cM	Mean Distance, cM	Gaps > 2 cM (*n*)
1	161	96.1	0.60	2.4, 3.7	149	88.2	0.59	2.4 - 6.9(4)
2	76	59.6	0.78	2.1 - 4.3 (3)	52	44.4	0.85	2.1 -2.9 (4)
3	97	87	0.90	2.1 - 5.9 (7)	87	58.8	0.68	2.8
4	68	44.6	0.66	2.2, 2.5	73	58.1	0.80	2.0 - 4.3 (6)
5	47	58.2	1.24	2.0 - 7.5 (10)	85	44	0.52	2.8, 4.1
6	86	76.4	0.89	2.0 - 5.1 (8)	53	63.8	1.20	2.1 - 4.5 (9)
7	64	74	1.16	2.1 - 6.2 (8)	50	43.5	0.87	2.1 - 3.7 (4)
8	56	54.7	0.98	2.0 - 4.4 (6)	43	54.1	1.26	3.7 - 6.3 (3)
Total	655	550.6	0.84		592	454.9	0.77	

Note: LG, linkage group.

Phenotypic evaluation and mapping of CR QTLs

CR of the parents obtained from the phenotypic assay was in agreement with previous studies establishing CR at 600 and 300 chilling hours for Perfection and A.1740, respectively. CR segregated in the mapping population, with 900, 600, 500, 400, 300, and 200 chilling hours required for 49, 4, 14, 9, 2, and 12 individuals, respectively (Fig. 2.3). Progenies with the parental phenotypes were the least frequent, with 4 and 2 individuals for Perfection and A.1740, respectively (Fig. 2.3). Transgressive segregants were observed for very low and very high CR i.e. 200 and 900 chilling hours, respectively.

Following QTL analysis for co-segregation between phenotypic and genotypic markers using the additive and additive × additive interaction regression model, a total of 12 unique QTLs were detected at initial LOD thresholds of 9.44 and 8.46 (as determined by a permutation test) for the Perfection and A.1740 maps, respectively. LOD scores of the QTL peaks ranged from 10.52 to 64.61, while the QTL support intervals were established at 2-LOD support interval (Table 2.4). The positive additive effects indicate that the female parent with high CR (Perfection) contributed the increasing allele, while the negative additive effects indicate that the male parent with low CR (A.1740) contributed the increasing allele. Most of the increasing QTL alleles came from the high CR parent (Perfection), while 4 of the increasing alleles are from the low CR parent (A.1740).

In addition to the one-dimensional genome-wide scan, the main-effect QTLs of digenic interactions were estimated as shown in Table 2.5. Six main-effect QTLs detected on the Perfection map explained a total of 58.5% ± 6.7% of the phenotypic variance, while 8

main-effect QTLs detected on the A.1740 map accounted for 66.1% ± 5.8% of the phenotypic variance (Table 2.5). Two of the QTLs detected were common to both maps on LGs 1 and 8. Distorted markers were fairly evenly distributed across all linkage groups except for the region around the QTL on LG6 with a peak at 29 cM. The overrepresentation of alleles from A.1740 for all distorted markers in this genomic region (Fig. 2.2) corresponds with results from QTL analysis showing that the increasing allele is from the A.1740 parent (Table 2.4).

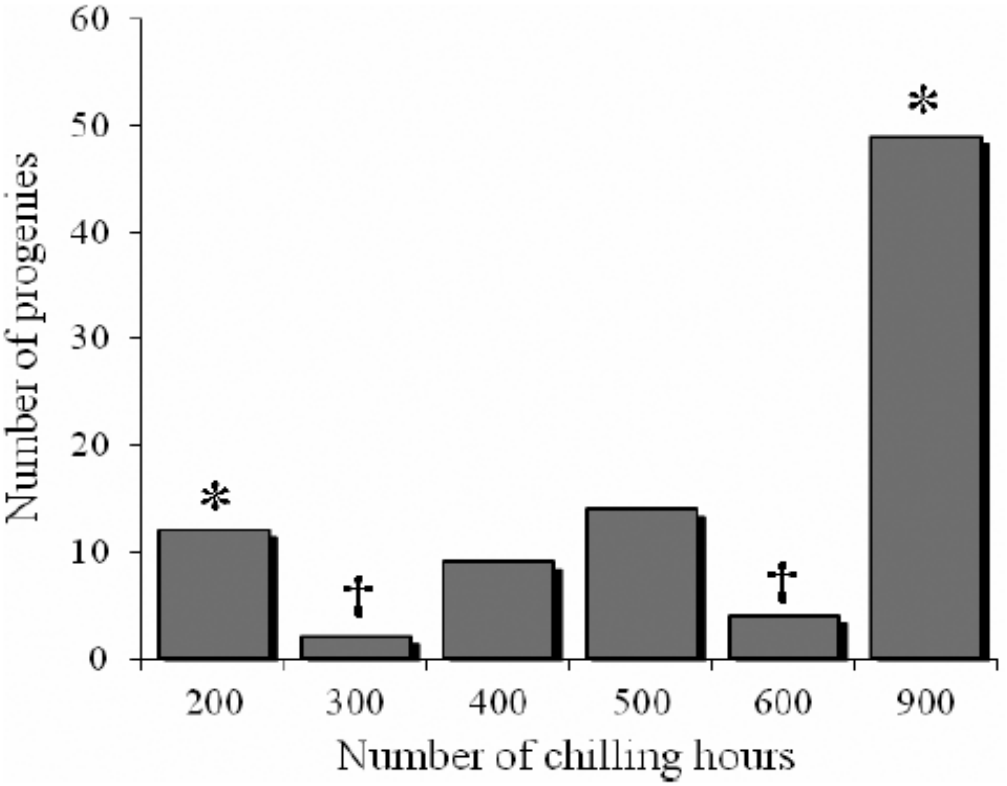


Fig. 2.3: Frequency distribution of chilling requirement (CR) phenotypes in the Perfection × A.1740 mapping population. A.1740 and Perfection had CRs of 300 and 600 chilling hours (†), respectively. Transgressive segregants are indicated by asterisks (*).

Table 2.4: QTLs detected for chilling requirement in parental maps using composite interval mapping.

Map	LG	Position (cM)	SI[a]	Additive effect	SSR Markers within or close to SI
Perfection	1	82	80-84	117.19	BPPCT28-76.1
	2	15	13-17	72.42	BPPCT04-23.9
	6	29	27-31	602.62	BPPCT25-43.2
	6	66	64-68	165.93	UDP98-412-58.8
	7	8	6-10	-205.37	-
	8	44	42-46	-221.19	EPDCU3454-42.2
A.1740	1	69	67-71	400.50	EPDCU2862-64
	1	86	81-89	265.75	-
	2	4	2-6	-301.04	-
	2	36	34-38	247.04	Taly-36.7
	5	1	0-3	-595.14	BPPCT7-0.0
	5	27	25-29	112.82	EPDCU3083-21.2
	7	26	26-34	180.82	Pchgms044b-30.5, CPSCT042-33.1
	8	35	33-38	75.63	-

Note: LG, linkage group. Position, location of the maximum LOD score of the QTL on LG; A, additive QTL effect. Only QTLs above empirical threshold of 9.44 and 8.46 (for the Perfection and A.1740 maps, respectively) are listed (LOD threshold computed by 1,000 permutations). The positive and negative additive effects indicate that the allele which increases the trait values is in the Perfection and A.1740 parent, respectively.

[a]2-LOD support interval in the fit.

Table 2.5: Digenic interactions of QTL controlling CR.

		A		A x A	
QTL 1	QTL 2	R^2 (%)	[d]Std Eff	R^2(%)	[d]Std Eff
Perfection (LOD= 17.19, R =0.765, R^2% = 58.5 ± 6.7)					
LG1:82	LG7:8	-	-	8.9	1.29*
LG2:15	LG8:44	-	-	9.0	1.05*
LG6:29	LG6:66	-	-	12.4	-1.33**
A.1740 (LOD= 21.15,R = 0.813, [b]R^2%= 66.1 ± 5.8)					
LG5:27		6.2	0.56 *	-	-
LG1:69	LG7:26	-	-	16.7	0.95**
LG1:86	LG2:36	-	-	15.5	-0.89**
LG2:4	LG2:36	-	-	16.7	0.76**
LG2:36	LG5:27	-	-	14.1	1.16 **
LG5:1	LG8:35	-	-	10.1	-1.06 **

Note: A and A x A values are the additive and additive x additive interaction effects of QTLs; R, Multiple correlation coefficient; R^2, percentage of phenotypic variance explained by all the QTL interactions; Std eff., Standardized QTL effects. * and ** denote significance level at 0.05 and 0.01 respectively.

Discussion

Linkage analysis and map construction

Comparative mapping within *Prunus*, as well as among related genera (*Malus*, *Pyrus*, *Rosa*, *Rubus*, and *Fragaria*), has emerged as a potential strategy for genetically exploring economically important traits. These maps facilitate the detection of genomic intervals that underlie economic traits, some of which have already been shown to segregate in our mapping population. The high map saturation with AFLP markers reduces detection of spurious QTLs and could potentially provide (with conversion to sequence tagged sites) more tightly linked flanking markers for marker-assisted selection.

The two high-density parental maps we constructed in apricot correspond well to the *Prunus* reference map (Dirlewanger *et al.* 2004), since all except 2 SSR markers were syntenic to those in the reference map in all 8 homeologous chromosomes. The incongruence of the 2 SSR markers (CPPCT034 and BPPCT040) occurred within a small genetic distance, while the position and incongruence of CPPCT034 were validated by another apricot map (Dondini *et al.* 2007) which was mostly syntenic with our maps. Because of the inclusion of the *Prunus* anchor SSR loci in our maps, all functional genomic resources (ESTs, BAC sequences, whole genome sequences, and cDNAs) from the *Prunus* database will provide more information for the genomic intervals in which detected QTLs exist. Our maps provide better genome coverage than the previously published apricot maps. The greater genome coverage observed predominantly in the Perfection parental map can be attributed to the availability of more markers segregating

in this parent. Aligning the linkage groups with bridge markers in similar order provided evidence for regions on either end of the linkage groups that are underrepresented in the other parental map (9.2 to 25.8 cM). At $P < 0.05$, the percentages of marker distortion in this study (24.60% and 13.98%) were comparable with other published works (10.2% to 17%) in apricot (Hurtado *et al.* 2002, Vilanova *et al.* 2003, Lambert *et al.* 2004, Lalli *et al.* 2007).

Phenotypic evaluation and QTL mapping

In this study, 12 unique QTLs for chilling requirement were mapped in the Perfection × A.1740 population. Owing to limited population size, only QTLs with large effects were statistically significant; hence, the number of QTLs detected in this study should be considered a minimal estimate (Dirlewanger *et al.* 1999). Our sampling in 2 years spans 100 to 900 chilling hours, but only the central portion of this range (200–600 h) overlaps between years. A more robust sampling with replicates could allow for better definition or elimination of minor QTLs.

As mentioned earlier, Perfection has high CR, whereas A.1740 has low CR. However, some progenies exhibited lower CR than A.1740. This could be attributed to epistatic interactions of different QTL alleles. Some individuals superseded the high CR parent. This may be due to contribution of some QTL alleles from the low CR parent to these progenies. The transgressive segregants, both positive and negative, may serve as useful materials for future breeding of high and low CR apricot as required for specific agro-climates.

QTLs detected in this study on LGs 1, 5, 6, 7, and 8 were localized to similar map positions in a CR study in peach (Fan *et al.* 2008). In addition, the SSR markers in apricot mapped within or close to the support intervals of the corresponding QTLs in peach. The similarity in QTL positions between peach and apricot deserves special attention because bud flush was studied in peach, while dormancy release of vegetative bud was analyzed in apricot. This suggests that similar genes might be involved in these two different physiological processes.

Following detection of QTLs controlling CR and bud break, the next step is to associate these loci with known genes using functional genomics and transcriptome resources. The QTL on LG1 was shown to map to a region corresponding to the known location of the *EVERGROWING* (*EVG*) locus in peach, characterized as comprising MADS-box transcription factors (Wang *et al.* 2002, Bielenberg *et al.* 2004, Bielenberg *et al.* 2008). The *evg* mutant is known to lack responsiveness to winter temperatures (i.e., *evergrowing* trees keep growing and are killed by low winter temperatures). The presence of the wild-type cold-responsive *EVG* locus within close proximity of our QTL provides us with a potential gene candidate in this region that has been extensively characterized in peach (Bielenberg *et al.* 2008).

Genomic sequence-based and EST-derived SSR markers that mapped directly within QTLs on LGs 7 and 8 included CPSCT042 and EPDCU3454, respectively. The CPSCT042 genomic and EPDCU3454 EST sequences showed homology to the *MITOGEN-ACTIVATED PROTEIN KINASE7 (MPK7*) and *ABSCISIC ACID*

INSENSITIVE3 (*ABI3*) genes, respectively. *ABI3* has been confirmed in several studies to be primarily responsible for imposition and maintenance of seed dormancy (Rohde *et al.* 2002). The processes associated with late seed development, such as reserve accumulation, dormancy imposition, and acquisition of tolerance of seed tissues to desiccation, seem to be controlled by *ABI3* (Bonetta and McCourt 1998). The gene promoter activity of *ABI3* has also been detected in vegetative meristems (Ng *et al.* 2004). MPK7 is a downstream substrate of MKK3 (a *MAPK* kinase), and in a few recent studies it has been shown to have a role in transducing signals involving reactive oxygen species (ROS) (Dóczi *et al.* 2007, Pitzschke and Hirt 2009) and in turn a corresponding H_2O_2-induced dormancy release in grapevine (Pérez and Lira 2005) and raspberry (Mazzitelli *et al.* 2007). Several studies now show that the *MAPK* cascade is not only induced by ROS but can also regulate production of ROS (Pitzschke and Hirt 2009).

Utilizing the available *Prunus* genomic resources (Horn *et al.* 2005, Zhebentyayeva *et al.* 2008), we are currently expanding our search for candidate CR genes in these major QTL intervals.

Conclusion

In this study, we report on 2 high-density parental maps in apricot constructed by using *Prunus* SSR anchor markers and saturated by using AFLP markers. A densely populated map is required for map-based cloning of economically important genes (Zhang 2008)

and dissection of complex traits to understand their genetic basis (Frewen *et al.* 2000). Based on our current map resolution and the degree of microsynteny between peach and apricot (Jung *et al.* 2006), a candidate gene approach should be possible for discovery of genes involved in CR and bud break, particularly as the whole genome sequence of peach is currently been assembled (B. Sosinski, personal communication, 2008). However, to assist in this candidate gene endeavor, it is necessary to detect and locate QTLs more precisely by fine-mapping and other approaches such as association mapping.

Acknowledgements

This research was supported by research grant No. US-3746-05 R from the United States – Israel Binational Agricultural Research and Development Fund (BARD). The authors wish to thank Tatyana N. Zhebentyayeva, Laura L. Georgi, Renate Horn, Iñaki Hormaza, Wu Rongling, and Aldrich Preston for expert technical help.

References

Aranzana, M.J., Pineda, A., Cosson, P., Dirlewanger, E., Ascasibar, J., Cipriani, G., Ryder, C.D., Testolin, R., Abbott, A., King, G.J., Iezzoni, A.F., Arús, P. 2003. A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theor. Appl. Genet. 106(5): 819–825.

Arora, R., Rowland, L.J., Tanino, K. 2003. Induction and release of bud dormancy in woody perennials: a science comes of age. HortScience. 38: 911–921.

Bielenberg, D.G., Wang, Y., Fan, S., Reighard, G.L., Scorza, R., Abbott, A.G. 2004. A deletion affecting several gene candidates is present in the *Evergrowing* peach mutant. J. Hered. 95(5): 436–444.

Bielenberg, D.G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G.L., Scorza, R., Abbott, A.G. 2008. Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet. Genomes. 4: 495-507.

Bonetta, D., McCourt, P. 1998. Genetic analysis of ABA signal transduction pathways. Trends Plant Sci. 3(6): 231–235.

Castiglioni, P., Ajmone-Marsan, P., van Wijk, R., Motto, M. 1999. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor. Appl. Genet. 99(3–4): 425–431.

Churchill, G.A., Doerge, R.W. 1994. Empirical threshold values for quantitative trait mapping. Genetics. 138(3): 963–971.

Combes, M.C., Andrzejewski, S., Anthony, F., Bertrand, B., Rovelli, P., Graziosi, G., Lashermes, P. 2000. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. Mol. Ecol. 9(8): 1178–1180.

Coville, F.V. 1920. The influence of cold in stimulating the growth of plants. J. Agric. Res. 20: 151–192.

Dietrichson, J. 1964. The selection problem and growth rhythm. Silvae Genet. 13: 178–184.

Dirlewanger, E., Moing, A., Rothan, C., Svanella, L., Pronier, V., Guye, A., Plomion, C., Monet, R. 1999. Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batch). Theor. Appl. Genet. 98: 18–31.

Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W., Arús, P. 2004. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc. Natl. Acad. Sci. U.S.A. 101(26): 9891–9896.

Dóczi, R., Brader, G., Pettkó-Szandtner, A., Rajh, I., Djamei, A., Pitzschke, A., Teige, M., Hirt, H. 2007. The *Arabidopsis* mitogen-activated protein kinase kinase MKK3 is upstream of group C mitogen-activated protein kinases and participates in pathogen signaling. Plant Cell. 19(10): 3266–3279.

Dondini, L., Lain, O., Geuna, F., Banfi, R., Gaiotti, F., Tartarini, S., Bassi, D., Testolin, R. 2007. Development of a new SSR-based linkage map in apricot and analysis of synteny with existing *Prunus* maps. Tree Genet. Genomes. 3(3): 239–249.

Eldredge, L., Ballard, R., Baird, W.V., Abbott, A., Morgens, P., Callahan, A., Scorza, R., Monet, R. 1992. Application of RFLP analysis to genetic-linkage mapping in peaches. HortScience. 27(2): 160–163.

Falque, M. 2005. IRILmap: linkage map distance correction for intermated recombinant inbred lines/advanced recombinant inbred strains. Bioinformatics. 21(16): 3441–3442.

Fan, S., Bielenberg, D., Zhebentyayeva, T., Reighard, G., Abbott, A. 2008. QTL analysis of chilling requirement. *In* 4th International Rosaceae Genomics Conference, Pucón, Chile, 16–19 March 2008. [Abstr.]

Frewen, B.E., Chen, T.H.H., Howe, G.T., Davis, J., Rohde, A., Boerjan, W., Bradshaw, H.D., Jr. 2000. Quantitative trait loci and candidate gene mapping of bud set and bud flush in populus. Genetics. 154(2): 837–845.

Grattapaglia, D., Sederoff, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics. 137(4): 1121–1137.

Haley, C.S., Knott, S.A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity. 69(4): 315–324.

Hatch, A.H., Walker, D.R. 1969. Rest intensity of dormant peach and apricot leaf buds as influenced by temperature, cold hardiness and respiration. J. Am. Soc. Hortic. Sci. 94(3): 304–307.

Hill, J., Becker, H.C., Tigerstedt, P.M.A. 1998. Quantitative and ecological aspects of plant breeding. St. Edmundsbury Press, Suffolk, U.K.

Hjorth, J.S.U. 1994. Computer intensive statistical methods. Chapman & Hall, New York.

Horn, R., Lecouls, A.C., Callahan, A., Dandekar, A., Garay, L., McCord, P., Howad, W., Chan, H., Verde, I., Main, D., Jung, S., Georgi, L., Forrest, S., Mook, J., Zhebentyayeva, T., Yu, Y., Kim, H.R., Jesudurai, C., Sosinski, B., Arús, P., Baird, V., Parfitt, D., Reighard, G., Scorza, R., Tomkins, J., Wing, R., Abbott, A.G. 2005. Candidate gene database and transcript map for peach, a model species for fruit trees. Theor. Appl. Genet. 110(8): 1419–1428.

Hurtado, M.A., Westman, A., Beck, E., Abbott, G.A., Llacer, G., Badenes, M.L. 2002. Genetic diversity in apricot cultivars based on AFLP markers. Euphytica. 127(2): 297–301.

Jung, S., Main, D., Staton, M., Cho, I., Zhebentyayeva, T., Arús, P., Abbott, A. 2006. Synteny conservation between the *Prunus* genome and both the present and ancestral *Arabidopsis* genomes. BMC Genomics. 7(1): 81.

Jung, S., Staton, M., Lee, T., Blenda, A., Svancara, R., Abbott, A., Main, D. 2008. GDR (Genome Database for Rosaceae): integrated web-database for Rosaceae genomics and genetics data. Nucleic Acids Res. 36(Database issue): D1034–D1040.

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugen. 12: 172–175.

Lalli, D.A., Abbott, A.G., Zhebentyayeva, T.N., Badenes, M.L., Damsteegt, V., Polák, J., Krška, B., Salava, J. 2007. A genetic linkage map for an apricot (*Prunus armeniaca* L.) BC_1 population mapping *plum pox virus* resistance. Tree Genet. Genomes. 4(3): 481–493.

Lambert, P., Hagen, L.S., Arus, P., Audergon, J.M. 2004. Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas × peach Earlygold reference map for *Prunus*. Theor. Appl. Genet. 108(6): 1120–1130.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., Newburg, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics. 1(2): 174–181.

Lang, G.A., Early, J.D., Martin, G.C., Darnell, R.L. 1987. Endodormancy, paradormancy, and ecodormancy- physiological terminology and classification for dormancy research. HortScience. 22(3): 371–377.

Lodhi, M.A., Daly, M.J., Ye, G.-N., Weeden, N.F., Reisch, B.I. 1995. A molecular marker based linkage map of *Vitis*. Genome. 38(4): 786–794.

Maliepaard, C., Jansen, J., Van Ooijen, J.W. 1997. Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for application. Genet. Res. 70(3): 237–250.

Mathiason, K., He, D., Grimplet, J., Venkateswari, J., Galbraith, D.W., Or, E., Fennell, A. 2009. Transcript profiling in *Vitis riparia* during chilling requirement fulfillment reveals coordination of gene expression patterns with optimized bud break. Funct. Integr. Genomics. 9(1): 81–96.

Mazzitelli, L., Hancock, R.D., Haupt, S., Walker, P.G., Pont, S.D., McNicol, J., Cardle, L., Morris, J., Viola, R., Brennan, R., Hedley, P.E., Taylor, M.A. 2007. Co-ordinated gene expression during phases of dormancy release in raspberry (*Rubus idaeus* L.) buds. J. Exp. Bot. 58(5): 1035–1045.

Ng, D.W., Chandrasekharan, M.B., Hall, T.C. 2004. The 5′ UTR negatively regulates quantitative and spatial expression from the *ABI3* promoter. Plant Mol. Biol. 54(1): 25–38.

Pérez, F.J., Lira, W. 2005. Possible role of catalase in post-dormancy bud break in grapevines. J. Plant Physiol. 162(3): 301–308.

Pitzschke, A., Hirt, H. 2009. Disentangling the complexity of mitogen-activated protein kinases and reactive oxygen species signaling. Plant Physiol. 149(2): 606–615.

Porebski, S., Bailey, L.G., Baum, B.R. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol. Biol. Rep. 15(1): 8–15.

Qi, X.Q., Stam, P., Lindhout, P. 1996. Comparison and integration of four barley genetic maps. Genome. 39(2): 379–394.

Rohde, A., Prinsen, E., De Rycke, R., Engler, G., Van Montagu, M., Boerjan, W. 2002. *PtABI3* impinges on the growth and differentiation of embryonic leaves during bud set in poplar. Plant Cell. 14(8): 1885–1901.

Rohde, A., Ruttink, T., Hostyn, V., Sterck, L., Van Driessche, K., and Boerjan, W. 2007. Gene expression during the induction, maintenance, and release of dormancy in apical buds of poplar. J. Exp. Bot. 58(15-16): 4047–4060.

Ruiz, D., Campoy, J.A., Egea, J. 2007. Chilling and heat requirements of apricot cultivars for flowering. Environ. Exp. Bot. 61(3): 254–263.

Scorza, R., Okie, W.R. 1990. Peaches (*Prunus persica* L. Batsch). Acta Hortic. 290: 177–231.

Utz, H.F., Melchinger, A.E. 1995. PLABQTL: a computer program to map QTL. Version 1.0. University of Hohenheim.

Utz, H.F., Melchinger, A.E. 1996. PLABQTL: a program for composite interval mapping of QTL. JQTL 2(1).

Van Ooijen, J.W., Voorrips, R.E. 2001. JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.

Van Ooijen, J.W., Sandbrink, J.M., Vrielink, M., Verkerk, R., Zabel, P., Lindhout, P. 1994. An RFLP linkage map of *Lycopersicon peruvianum*. Theor. Appl. Genet. 89(7–8): 1007–1013.

Vilanova, S., Romero, C., Abbott, A.G., Llácer, G., Badenes, M.L. 2003. An apricot (*Prunus armeniaca* L.) F2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. Theor. Appl. Genet. 107(2): 239–247.

Voorrips, R.E. 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. J. Hered. 93(1): 77–78.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23(21): 4407–4414.

Wang, Y., Georgi, L.L., Reighard, G.L., Scorza, R., Abbott, A.G. 2002. Genetic mapping of the *evergrowing* gene in peach [*Prunus persica* (L.) Batsch]. J. Hered. 93(5): 352–358.

Weinberger, J.H. 1950. Chilling requirements of peach varieties. Proc. Am. Soc. Hortic. Sci. 56: 122–128.

Winkler, C.R., Jensen, N.M., Cooper, M., Podlich, D.W., Smith, O.S. 2003. On the determination of recombination rates in intermated recombinant inbred populations. Genetics. 164(2): 741–745.

Wu, R.L., Ma, C.X., Painter, I., Zeng, Z.B. 2002. Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. Theor. Popul. Biol. 61(3): 349–363.

Wu, R.L., Ma, C.X., Casella, G. 2007. Statistical genetics of quantitative traits: linkage, maps, and QTL. Springer-Verlag, New York.

Zeaser D. 2001. Vegetative propagation of Gmelina arborea Roxb. In: CAMCORE. International Tree Breeding Short Course Book. North Carolina State University, Raleigh, North Carolina, USA, pp. 3-15.

Zeng, Z.B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc. Natl. Acad. Sci. U.S.A. 90(23): 10972–10976.

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genetics. 136(4): 1457–1468.

Zhang, H.B. 2008. Map-based cloning of genes and quantitative trait loci. *In* Principles and practices of plant genomics. Vol. 1. Genome mapping. *Edited by* C. Kole and A.G. Abbott. Science Publishers, Enfield, N.H. pp. 229–267.

Zhebentyayeva, T.N., Swire-Clark, G., Georgi, L.L., Garay, L., Jung, S., Forrest, S., *et al.* 2008. A framework physical map for peach, a model Rosaceae species. Tree Genet. Genomes. 4(4): 745–756.

CHAPTER THREE

COMPARATIVE ANALYSIS OF QTLS UNDERLYING CHILLING REQUIREMENT AND BUD BREAK IN PEACH (*PRUNUS PERSICA* L.) AND APRICOT (*P. ARMENIACA* L.)

Bode A. Olukolu,[a] Fan Shenghua,[a] Taly Trainin,[d] Tatyana N. Zhebentyayeva,[a] Douglas G. Bielenberg,[b, c] Chittanranjan Kole,[a] Gregory L. Reighard,[b] William R. Okie,[e] Doron Holland,[d] and Albert G. Abbott[a]

[a]Genetics and Biochemistry, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA.
[b]Horticulture, Clemson University, 170 Poole Agriculture Center, Clemson, SC 29634, USA.
[c]Biological Sciences, Clemson University, 132 Long Hall, Clemson, SC 29634, USA.
[d]Fruit Sciences, Newe Ya'ar Research Center, Agricultural Research Organization, P.O. Box 1021, Ramat Yishay 30095, Israel.
[e]ARS-USDA, Southeastern Fruit and Tree Nut Research Laboratory, Byron, GA 31008, USA

Corresponding author: A.G. Abbott (e-mail: aalbert@clemson.edu)

The peach mapping population was provided by Dr. William Okie, maintained by Dr. Douglas G. Bielenberg and phenotyped by Dr. Douglas G. Bielenberg, Fan Shenghua and Bode A. Olukolu. The apricot mapping population was maintained and phenotyped by Dr. Doron Holland and Taly Trainin. Dr. Tatyana N. Zhebentyayeva assisted with developing peach SSR markers used in both maps and linkage analysisfor the peach map. Genotyping, linkage analysis, map construction and QTL analysis for peach and apricot were performed by Fan Shenghua and Bode A. Olukolu, respectively, while the comparative analysis of QTLs was conducted by Bode A. Olukolu.

Abstract

Chilling requirement (CR), a major factor that determines bud break, is a limiting factor for temperate fruit production, hence, a key trait breeders select for to avoid frost damage and to ensure uniform bud burst. In this study, we characterized common quantitative trait loci (QTLs) underlying the trait in mapping populations of two *Prunus* species (peach and apricot). In peach, a total of 8 QTLs detected for CR and 10 QTLs for bloom date mapped to 12 genomic regions, with 6 common QTLs indicating a common underlying genetic factor. Altogether, 20 QTLs were detected in apricot under additive-by-additive epistatic models with and without dominance. Four of the 12 QTLs detected in each of the two models were consistent on linkage groups 1, 2, 6 and 8. A majority of the QTLs were stable between both *Prunus* species, as well as, similar trends in their QTL effects, with the allele for increasing the trait value mostly originating from the high chill parents. The denser apricot map provided a higher resolution to delineate QTLs to smaller genomic intervals, as well as, splitting each of the peach QTLs on linkage groups 2, 4, 6 and 7 into 2 QTLs with smaller genomic intervals. The comparative QTL mapping strategy presented here reveals the transferability of genetic information between two *Prunus* species, characterization of stable QTLs, utility of the maps to consolidate each other and to further validate previously identified CR QTL loci as a major controlling factor driving floral bud break.

Keywords: QTL effects, conservation of synteny, collinearity and marker transferability.

Introduction

The need for transferability of genetic information from one organism, usually a model organism, to a non-model organism is crucial especially for organisms that are intractable or that lack genetic resources for similar studies (Hall *et al.* 2002, Schmidt 2002). The core of comparative mapping for this purpose encompasses syntenic relationships between closely related species within the same taxonomic genus or family, consolidation of genetic maps, verifying quantitative trait loci (QTL), identification of candidate genes underlying QTLs and a better understanding of genome evolution (Sankoff and Nadeau 2000, Kliebenstein *et al.* 2001, Murphy *et al.* 2001, Zhang *et al.* 2001, Schmidt *et al.* 2002). This not only facilitates transfer of genetic information across different species but also allows for the taxonomic family to be viewed as a single genetic system (Freeling 2001). Without sequenced and assembled plant genomes; comparative analysis frequently relies on molecular markers common among species (Cabrera *et al.* 2009). The application of this approach within grasses has proven successful for the positional cloning of important genes such as VRN1 in wheat, a species that map-based cloning was considered impracticable due to the large genome size and several repetitive elements in the genome that would impede chromosome walking (Yan *et al.* 2003).

Genome colinearity, conservation of synteny and marker transferability among member species of *Prunus* has been demonstrated by several studies. These were mainly based on comparisons of 13 maps from different *Prunus* populations (Dirlewanger *et al.* 2004). In the *Rosaceae,* this genus provides the most detailed genetic map which was derived from

an interspecific almond (*P. dulcis*) cv. Texas × peach (*P. persica*) cv. Earlygold (TxE *Prunus* reference map) F_2 mapping population (Joobeur 1998, Aranzana *et al.* 2003, Dirlewanger 2004).

Recently two studies identifying QTL for chilling requirement (CR) and bloom date (BD) in peach and apricot were reported (Olukolu *et al.* 2009, Fan *et al.* 2010). CR is an agronomic trait that specifies floral and vegetative bud break. CR is the major factor determining bud break (Ruiz *et al.*, 2007; Alburquerque *et al.*, 2008), which is an important agronomic trait affecting production in temperate fruit tree species. CR refers to the duration of low temperatures necessary for the release of temperate trees from endo-dormancy so that initiation of growth in response to transient warm temperature is prevented before fulfillment of chill requirement. Consequently, frost damage during later winter or early spring is avoided. Unfortunately this trait which protects fruit trees also limits the climatic distributions of temperature fruit tree genotypes (Coville 1920, Scorza and Okie 1990, Sherman and Beckman 2003). Because the previously mentioned CR QTL studies utilized the framework SSR marker set from the general *Prunus* genetic map (Aranzana *et al.* 2003), we were in the unique position to compare the across species nature of major CR QTLs in these two closely related *Prunus* species each grown in very different environments and conditions. Additionally, in the case of peach CR determinations, floral bud break was scored and in the case of apricot, vegetative bud break was scored, thus, we could potentially identify QTL that were common for CR in these different tissue types as well. Overlapping QTLs in this latter case would provide candidate genes that play a pivotal role to the pathway of tissue dormancy.

The main objective of this study was to integrate common marker data to generate a comparative map of the QTLs for CR and bud break. Most of the QTLs were shown to be stable in both *Prunus* species as well as possess similar trends in their QTL effects i.e. the high chill cultivars contribute the allele that is associated with increase in the trait value.

Material and methods

Mapping populations and linkage maps

The peach map (Fan *et al.* 2010) was constructed using an F_2 population with 378 individuals developed at ARS-USDA, Southern Fruit and Tree Nut Research Laboratory (Byron, GA, USA) by crossing two peach genotypes with high (1050 chilling hr) and low (300 chilling hr) CR values. The female parent (Contender) is a commercial peach cultivar in the southeastern USA developed by the North Carolina Agricultural Service (Raleigh, NC, USA), while the male parent (Fla.92-2C) is a selection from the University of Florida's (Gainesville, FL, USA). The apricot maps (Olukolu *et al.* 2009) were constructed using a two-way pseudo-testcross population consisting of 100 F_1 individuals. The female parent (A.1740) is a North African cultivar with 300 chilling hr, while the male parent (Perfection) is a commercial hybrid cultivar with 600 chilling hr. The parents and F_1 progeny were maintained at the Newe Ya'ar Research Center of the Agricultural Research Organization in Israel. Routine methods of bagging and pollination were followed (Zeaser 2001). The TxE reference map, which was used as a reference

between the peach and apricot maps, was developed from an almond (cv. Texas) x peach (cv. Earlygold) F_2 progeny (Joobeur *et al.* 1998).

Marker analysis

A set of 370 SSR markers isolated from different *Prunus* species were tested for polymorphism in the peach F_2 mapping population using the female grandparent 'Contender' and the F1 tree 'BY01p6245'. Most of the SSR markers used for the apricot maps were mostly *Prunus* anchor marker set originally developed for peach (Aranzana *et al.* 2003) and were selected based on uniform distribution across the linkage groups of the *Prunus* reference map (Dirlewanger *et al.* 2004) to establish a framework map for genome homology studies. The peach linkage map comprised a total of 96 SSR and 30 AFLP loci (Fan *et al.* 2010), while the apricot linkage maps consists a total of 43 SSR and 1016 AFLP loci (Olukolu *et al.* 2009). Alleles detected with primers that produced multiple loci were labeled with an alphabetic suffix for SSR markers and a numeric suffix for AFLP markers.

Linkage analysis and map construction

Genetic linkage mapping was performed using JoinMap version 3.0 software (Van Ooijen and Voorrips 2001). Parameters used for linkage analysis and map construction in peach and apricot are described by Fan *et al.* 2010 and Olukolu *et al.* 2009, respectively. The orientation of the linkage groups was based on that of the TxE *Prunus* reference map and the *Prunus* bin map.

Evaluation of chilling requirement and bud break

The trait evaluation in peach and apricot were performed as described by Fan *et al.* 2010 and Olukolu *et al.* 2009, respectively. The peach population segregated and was scored for chilling requirement (CR) and blooming date (BD), while the apricot population was only scored for CR. For CR determinations, floral bud break and vegetative bud break were scored for peach and apricot, respectively. Starting at 300 chilling hr when air temperature drops below 7.2^0C, the branches of each genotype were harvested approximately every 100 chilling hr interval until 1100 chilling hr and forced to bud break in a glasshouse at 25^0C under a 16 hr photoperiod. CR data for the peach F_2 population were scored over winter 2007/ spring 2008 and winter 2008/ spring 2009, while BD was evaluated in spring of 2006, 2007, 2008 and 2009 on 3 replicates for each progeny. For each genotype, the trees were observed every 1 or 2 days in the spring to determine BD.

CR evaluation on the apricot mapping population was performed with 2 replicates of each progeny. Whole potted trees were subjected to chilling at 6 °C and removed at 200 hr intervals ranging from 200 to 600 hr in 2007 and 100 to 900 hr in 2008, with the exception of the 700 chilling hr treatment. Chilled trees were transferred to a naturally lit greenhouse at day and night temperatures of 25 °C and 13 °C, respectively, to force bud break under natural photoperiod and irradiance.

QTL analysis

Quantitative trait loci were detected using composite interval mapping (CIM, Zeng 1994) using the PC version of PlabQTL version 1.2 (Utz and Melchinger 1996). Parameters used for QTL analysis are described for peach and apricot in Fan *et al.* 2010 and Olukolu *et al.* 2009, respectively. Different genetic models were compared based on Akaike's information criterion and Bayesian information criterion values (Hjorth 1994) to determine the best regression fit in both peach and apricot maps. The model with the minimal value was selected as the best fit. A pure additive model without dominance was applied to the peach QTL analysis, while two loci additive-by-additive epistatic effects were applied to the apricot QTL analysis with and without dominance. CR or BD trait with multiple overlapping QTLs in more than one year were pooled together and represented within the same QTL interval.

The LOD curves were created by scanning at 1 cM intervals, while a permutation test (1000 resamplings) was performed to determine the critical LOD score appropriate to empirically identify a putative QTL with a genome-wide error at a 0.05 confidence level (Churchill and Doerge 1994). Subsequently, the detected QTLs and their estimated map positions were verified using a simultaneous multiple regression, which accounts for effects of other linked QTLs on a chromosome (Zeng 1993). The phenotypic variance explained by each QTL (R2) was calculated as the square of the correlation coefficient from the final multiple regression model (Utz and Melchinger 1995, 1996). To examine the interaction between detected QTLs, the general linear model of variance analysis was

implemented. One- or two-LOD intervals (c. 95% or 99% confidence interval) (Lynch & Walsh, 1998) for QTL detection were reported.

Result

Common QTLs between peach and apricot

In peach, a total of 8 QTLs were detected for CR and 10 QTLs for BD, with all 18 QTLs mapped to 12 genomic regions (Table 3.2). Six of the CR QTLs overlap considerably with 6 of the BD QTLs indicating a common underlying genetic factor (Table 3.1 and Fig. 3.1). Altogether, 20 QTLs were detected on the apricot maps for both models used i.e. two loci additive×additive epistatic effects with (12 QTLs) and without (12 QTLs) dominance. Four of the 12 QTLs detected for each model were consistent on linkage groups 1, 2, 6 and 8 (Fig. 3.1 and Table 3.3). On linkage group 1 (LG1), 2 of the 4 peach QTL intervals map into the same genomic region as 2 of the 3 apricot QTL intervals. On LG2, the single QTL in peach spanned across a similar genomic region in apricot, although the region in apricot represents two distinct QTLs refined to a significantly smaller genomic interval. Likewise on LG4, two distinct apricot QTLs were localized inside a larger genomic interval spanned by 1 of the 2 QTLs in peach. On LG5, 1 of the 2 apricot QTLs map in the same genomic region as the single QTL in peach. Similar to the trend observed on LG4 and 2, a single peach QTL on LG 6 mapped to a genomic region corresponding to the location of 2 of the 3 apricot QTLs. On LG7, 2 of the peach QTLs

have corresponding QTLs within the same genomic region in apricot. The peach QTL on the lower arm of the LG7 also spans a genomic region in apricot containing two distinct QTLs. The single QTLs on LG8 of peach and apricot didn't seem have overlapping genomic regions, although they were in close proximity to each other. The marker density in the region of LG8 on both maps was relatively low especially in the peach map which could result in the incongruence of the QTLs between the species.

The larger QTL intervals observed in peach were better defined in apricot due to the higher map resolution in apricot. This is as a result of high levels of polymorphisms observed using the AFLP markers (Olukolu *et al.* 2009) in apricot as compared to peach where polymorphism was extremely limited. This resulted in the apricot QTLs been defined into smaller genomic regions with the exception of the evergrowing (EVG) locus (Wang *et al.* 2002) on the bottom part of the peach LG1, which also had a high SSR marker saturation in peach. Some of the peach QTLs on LG2, 4, 6 and 7, were defined as representing 2 QTLs on the apricot maps. The splitting of these QTL in apricot could result from insufficient marker density in these regions in the peach map to resolve multiple QTLs within close genomic proximity or alternatively, the additional QTLs could result from additional gene activities associated with CR for vegetative bud break in apricot contrasting to CR for floral bud break in peach.

Table 3.1: Common QTLs and associated LOD scores between peach and apricot.

Linkage groups	[a]QTL support interval (cM)		LOD	
	Peach	Apricot	Peach	Apricot
1	43-56[BD]	52-55[A]	12.7[BD]	13.76[A]
	86-88[BD/CR]	67-71[A]	31.4[BD]/21.7[CR]	14.6[A]
2	20-37[BD]	13-17[P]	5.1[BD]	63.6[P]
	-	18-20[P]	-	45.7[P]
4	4-19[CR]	4-6[A]	9.8[CR]	21.9[A]
	4-33[BD]	24-26[A]	12.4[BD]	17.8[A]
5	24-38[BD/CR]	25-29[A]	4.14[BD]/3.88[CR]	15.5[A]
6	35-43[CR]	54-56[P]	3.3[CR]	64.5[P]
	34-42[BD]	64-68[P]	3.4[BD]	50.8[P]
7	13-22[BD]	6-10[P]	3.9[BD]	46.2[P]
	40-47[BD]/43-59[CR]	41-43[P]/26-34[A]	33.7[BD]/21.3[CR]	11.6[A]/76.3[P]
	-	34-36[A]	-	11.8[A]
8	36-54[CR]	42-46[P]/33-38[A]	3.6[CR]	56.5[P]/75.6[A]

[BD]Blooming date QTLs; [CR]Chilling requirement QTLs; [P]Perfection map; [A]A.1740 map; LOD, logarithm of the odds, threshold computed by 1,000 permutations; [a]2-LOD support interval; Only QTLs above empirical threshold of 9.44 and 8.46 for the Perfection and A.1740 maps, respectively, and 2.85 for the peach map are listed.

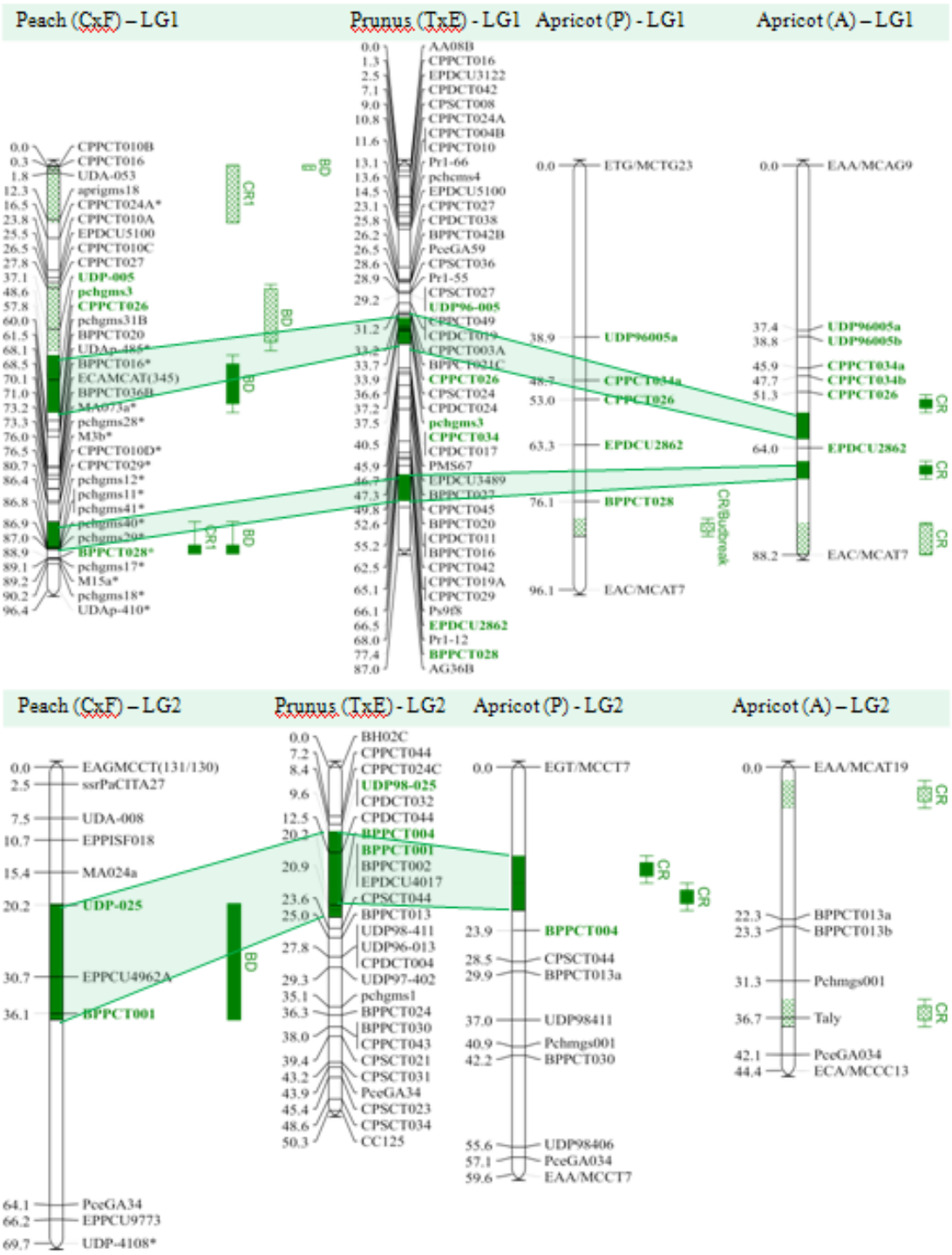


Figure 3.1: Comparative mapping of QTLs between peach (F_2 mapping from Contender x FLa.92-2C) and apricot (F_1 mapping population from Perfection x A.1740). The asterisks show the χ^2 *p* levels of significance (*, 0.05; **, 0.01; ***, 0.001). In peach and apricot, only the the SSR framework markers are shown (i.e. AFLP loci excluded).

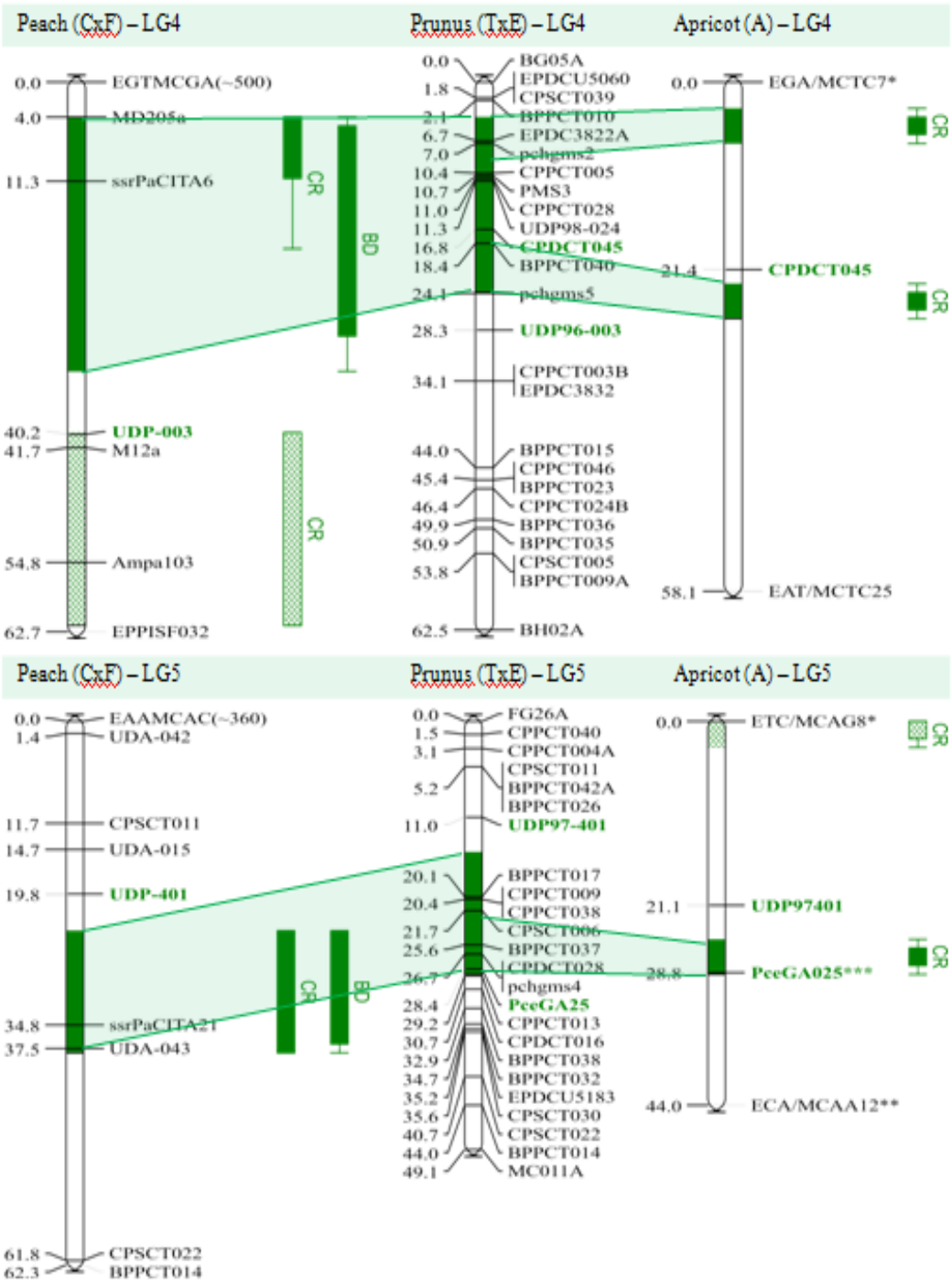


Figure 3.1: Continued

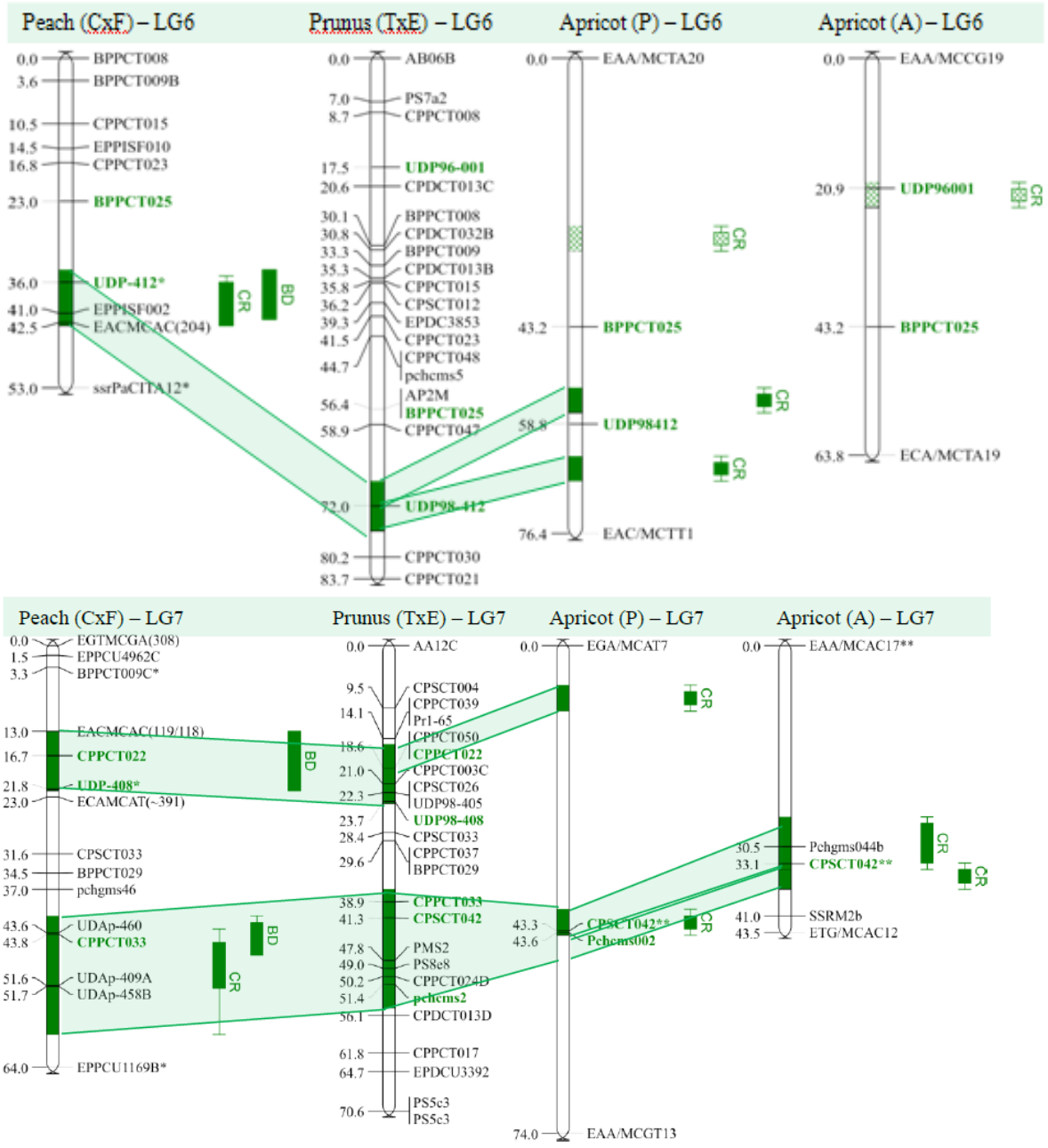


Figure 3.1: continued

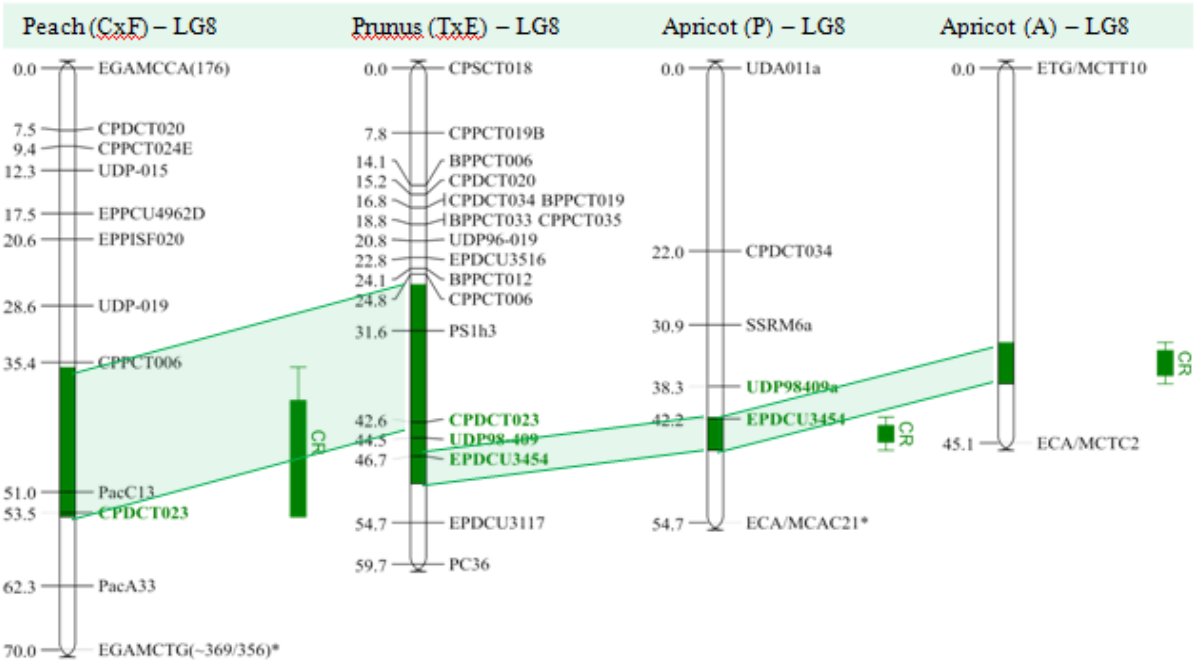


Figure 3.1: continued

QTL effects in peach and apricot

The peach QTLs were detected using only a pure additive model following a statistical test for the model with the best fit (Table 3.2), while the apricot QTL analysis was performed using an additive-by-additive gene action model to show epistatic interactions with dominance and without dominance (Table 3.3). The + or - effects value of the QTL (peach) and QTL interaction (apricot) indicates the parental origin of the allele responsible for increasing the trait value. The positive effect values indicate that the allele for increasing the trait value is from the male and vice versa. In peach, the positive and negative effects values correspond to the low chill FLa.92-2C male parent and the high chill Contender female parent, respectively, while in apricot, it corresponds to the high chill Perfection male parent and the low chill female A.1740 parent, respectively.

The alleles for increasing trait values in peach and apricot QTLs are mostly from the high chill parents i.e. Contender and Perfection, respectively. Only 1 of the 8 CR and 2 of the 10 BD peach QTLs had alleles increasing trait value that originated from the FLa.92-2C low chill parent (Table 3.2). These QTLs are localized on LG2 for BD (20-37 cM) and on LG6 for CR and BD (35-43 cM and 34-42 cM, respectively). Similarly, 1 of the 2 apricot QTLs mapping within the corresponding peach QTLs on LG 2 and 6 above had a negative value (i.e. allele increasing trait value in low chill A.1740; Table 3.3). Conversely, the QTL on the upper arm of the LG7 had alleles increasing the trait value originating from the low chill apricot cultivar parent instead of from the high chill parent as indicated in peach. Predominantly, the common QTLs between peach and apricot

showed similar trends in their QTL effects i.e. most of alleles that increase trait value are contributed by the high chill parents. Other QTL effects unique to the peach and apricot maps are shown in Table 3.2 and 3.3.

Proportions of phenotypic variance in peach and apricot

Most of the phenotypic variance in peach was contributed by 2 QTLs on LG1 (86-88 cM) and LG7 (43-57 cM) at about 40%. In the apricot genetic background, the proportions of partial phenotypic variance explained are presented for a QTL interaction when other QTL interactions effects were fixed (Table 3.4 and 3.5). The corresponding QTL genomic regions were consistently associated with QTL interactions that had high contributions to the phenotypic variance (Table 3.4 and 3.5), although these contributions were not as high as those observed in peach.

Table 3.2: The proportion of phenotypic variance contributed by each peach QTL, the additive QTL effect and the source of allelic dominance from either of the grandparents.

Linkage group	[a]QTL SI (cM)	Part R^2 (%)	A
1	0-1^{BD*}	4.0	-0.14
	0-13^{CR*}	7.6	-0.29
	27-42^{BD}	10.0	-0.26
	43-56^{BD}	12.8	-0.42
	86-88^{CR}	40.1	-0.68
	86-88^{BD}	42.65	-0.74
2	20-37^{BD}	6.15	0.228
4	4-19^{CR}	9.7	-0.32
	4-33^{BD}	14.4	-0.34
	40-62^{CR}	5.0	-0.23
5	24-38^{BD}	4.6	-0.20
	24-38^{CR}	5.3	-0.20
6	35-43^{CR}	4.2	0.19
	34-42^{BD}	4.0	0.14
7	13-22^{BD}	3.5	-0.17
	40-47^{BD}	21.4	-0.52
	43-59^{CR}	39.9	-0.71
8	36-54^{CR}	4.4	-0.20

Note: Part R2 (%), percentage of phenotypic variance explained by one QTL when other QTL effects are fixed; A, additive QTL effect. The positive (+) and negative (-) additive effects indicate that the allele which increases the trait values is in the FLa.92-2C and Contender parent, respectively.
[a]2-LOD support interval in the fit.

Table 3.3: QTLs detected for chilling requirement in parental maps using composite interval mapping.

Map	LG	A x A(no dominance)			A x A (dominance)			
		Position (cM)	[a]SI	Add	Position (cM)	[a]SI	A	D
		82	81-83	117.2	82	81-83	-37.5	-242.2
	2	15	14-16	72.4	-	-	-	-
	2	-	-	-	19	18-20	-660.1	-561.3
	3	-	-	-	44	43-45	-	-
	6	29	28-30	602.6	-	-	-	-
	6	-	-	-	55	54-56	-8401	-8491.4
	6	66	65-67	165.9	-	-	-	
	7	8	7-9	-205.4	-	-	-	
	7	-	-	-	42	41-43	-	-
	8	44	43-45	-221.2	44	43-45	576.6	800
A.1740	1	-	-	-	54	52-55	735.8	-
	1	69	68-70	400.5	-	-	-	-
	1	86	82-88	265.8	84	81-88	-10.8	80.3
	2	4	3-5	-301	-	-	-	-
	2	36	35-37	247	36	35-37	225.9	46.3
	3	-	-	-	12	11-13	839.3	-
	4	-	-	-	5	4-6	-	-
	4	-	-	-	25	24-26	-	95469
	5	1	0-2	-595.1	-	-	-	-
	5	27	26-28	112.8	-	-	-	-
	6	-	-	-	22	21-23	80.0	-118
	7	26	25-28	180.8	-	-	-	-
	7	-	-	-	35	34-36	-	397
	8	35	34-37	75.6	-	-	-	-

Note: LG, linkage group; Position, location of the maximum LOD score of the QTL on LG; A, additive QTL effect; D, dominance QTL effect; A x A, additive x additive interaction model. Only QTLs above empirical threshold of 9.44 and 8.46 (for the Perfection and A.1740 maps, respectively) are listed (LOD threshold computed by 1,000 permutations). QTLs consistent between both models, i.e. additive-by-additive epistatic effects model with and without dominance, are listed in the same row. The positive and negative additive effects indicate that the allele which increases the trait values is in the Perfection and A.1740 parent, respectively.

[a]1-LOD support interval in the fit.

Table 3.4: Digenic interactions of apricot QTLs controlling CR using the additive and additive-by-additive epistatic interaction model.

QTL1	QTL2	A		A x A	
		[c]Part R^2%	[d]Std Eff	[c]Part R^2%	[d]Std Eff
Perfection (LOD= 17.19, [a]R =0.765, [b]R^2% = 58.5 ± 6.7)					
LG1:82	LG7:8	-	-	8.9	1.29*
LG2:15	LG8:44	-	-	9.0	1.05*
LG6:29	LG6:66	-	-	12.4	-1.33**
A.1740 (LOD= 21.15, [a]R = 0.813, [b]R^2%= 66.1 ± 5.8)					
LG5:27		6.2	0.56 *	-	-
LG1:69	LG7:26	-	-	16.7	0.95**
LG1:86	LG2:36	-	-	15.5	-0.89**
LG2:4	LG2:36	-	-	16.7	0.76**
LG2:36	LG5:27	-	-	14.1	1.16 **
LG5:1	LG8:35	-	-	10.1	-1.06 **

[a]Multiple correlation coefficient; [b]percentage of phenotypic variance explained by all the QTL interactions; [c]percentage of phenotypic variance explained by a QTL interaction when other QTL interaction effects were fixed; [d]Standardized QTL effects; A and A x A values are the additive effect and additive x additive QTL interaction effects; * and ** denote significance level at 0.05 and 0.01, respectively. The positive and negative additive effects indicate that the allele which increases the trait values is in the Perfection and A.1740 parent, respectively.

Table 3.5: Digenic interactions of apricot QTLs controlling CR using the additive, dominance and additive x additive model.

QTL1	QTL2	A		D		A A	
		[c]Part R^2%	[d]Std Eff	[c]Part R^2%	[d]Std Eff	[c]Part R^2%	[d]Std Eff
Perfection (LOD= 12.71, [a]R =0.692, [b]R^2% = 47.8 ± 7.6)							
LG7:42	-	11.7	1.29**	-	-	-	-
LG1:82	LG3:44	-	-	-	-	6.9	-1.20*
LG2:19	LG6:55	-	-	-	-	9.0	1.66*
LG7:42	LG8:44	-	-	-	-	17.3	1.66**
A.1740 (LOD= 21.15, [a]R = 0.813, [b]R^2%= 66.1 ± 5.8)							
LG1:54	-	11.7	234.69**	11.8	236.19**	-	-
LG2:36	-			9.7	0.60*	-	-
LG1:54	LG3:12			-	-	9.2	1.67*
LG1:54	LG4:5			-	-	15.5	1.27**
LG1:84	LG2:36			-	-	18.6	-0.86**
LG4:25	LG7:35			-	-	11.8	1.50 **
LG6:22	LG7:35			-	-	16.3	1.13 **

[a]Multiple correlation coefficient; [b]percentage of phenotypic variance explained by all the QTL interactions; [c]percentage of phenotypic variance explained by a QTL interaction when other QTL interaction effects were fixed; [d]Standardized QTL effects; A and A x A values are the additive effect and additive x additive QTL interaction effects; * and ** denote significance level at 0.05 and 0.01, respectively. The positive and negative additive effects indicate that the allele which increases the trait values is in the Perfection and A.1740 parent, respectively.

Discussion

In this study, we present a comparative analysis of QTLs that provides a more comprehensive strategy for cataloging QTLs that cannot be achieved from one single population as well as an independent verification of significant QTLs that are stable between two populations. The stability of QTLs in turn is of utmost important since the utility associated markers across different genetic backgrounds are required for marker-assisted breeding. Some of these QTLs were also been reported for bloom date (BD) in previous studies and in other mapping populations (Dirlewanger *et al.* 1999, Verde *et al.* 2002, Silva *et al.* 2005). Although, the study aims to validate common QTLs identified in two different *Prunus* species, identifying QTLs unique to each mapping population is useful for elucidating inter-species trait differences and overcoming some of the limitations of single mapping populations. These limitations include the difficulty to generate a single large mapping population in trees segregating for a complex trait and the incidence of QTLs with large effects that obscure those with smaller effects (Paterson *et al.* 1988, 1990, Lander and Botstein 1989).

Establishing QTLs that are stable within a genus or taxonomic-family is crucial for transferability and optimal use of genetic information and resources in breeding programs. The results above support the conservation of QTLs controlling chilling requirement (CR) and bud break between two *Prunus* species. The maps were also shown to consolidate each other especially in regions where marker density is inadequate on one map, leading to the inability to detect or refine a QTL to a smaller genomic region. This is evident in the marker sparse genomic regions in peach that were potentially further

refined to 1 or 2 smaller genomic regions in the apricot maps that possess a much higher marker saturation with AFLP markers (Olukolu *et al.* 2009). In turn, the high density of SSR markers on the peach map provide informative marker information to anchor the two species maps and provide sequence-based markers that have utility for chromosome walking, isolating BACs within QTLs and defining genomic regions in the physical map and on the complete genome sequence of peach. This allows for transferability of genetic information between both species and for further enriching the genomic resources based on the peach genome. Some of the QTLs were also shown to be common between the CR and BD phenotypes after comparing the CR and BD QTLs of the peach map with the CR QTLs in apricot. This supports claims that CR is a major factor determining bud break (Ruiz *et al.* 2007, Alburquerque *et al.*, 2008).

The alleles for increasing trait value were consistent between both studies, with most of the alleles increasing the trait value originating from the high chill parents. This further implies that the underlying genetic factor or genes within the QTL regions between peach and apricot are also similar, thus, pointing to similar gene pathways driving the trait variation within the genus. The contributions to the phenotypic variance by QTLs seem to follow a similar trend, although the use of different effect models (best model fit determined statistically) prevents a more precise comparison. Most conspicuous is the contribution of peach QTL on LG1 (86-88 cM) and LG7 (43-57 cM) that agrees with the contributions by the corresponding QTLs in apricot, although the contributions by these peach QTLs are relatively higher compared to other detected QTLs. This probably points to a difference in the expression of the trait between peach and apricot or differences of

the response of floral buds vs. vegetative buds to CR and chemicals that induce dormancy release (Arora *et al.* 2003).

In this report, we presented stable and potentially key CR QTL regions based on a comparative QTL mapping approach. These results provide markers for: testing in marker-assisted breeding of CR, increasing our understanding of the genetics of this complex trait, and identifying and characterizing candidate genes whose study will establish the fundamental pathways controlling this important life history trait. The correspondence of the candidate genes in stable QTLs between these two taxa is crucial to validation of the comparative mapping approach and is the current focus of our continuing research in peach and apricot.

Aknowledgements

This research was supported by research grant No. US-3746-05 R from the United States – Israel Binational Agricultural Research and Development Fund (BARD) and the Robert and Lois Coker Trustees Chair in Molecular Genetics.

References

Alburquerque, N., Garcia-Montiel, F., Carrillo, A, Burgos, L. 2008. Chilling and heat requirements of sweet cherry cultivars and the relationship between altitude and the probability of satisfying the chill requirements, Environ. Exp. Bot. 64(2): 162-170. doi:10.1016/j.envexpbot.2008.01.003.

Aranzana, M.J., Pineda, A., Cosson, P., Dirlewanger, E., Ascasibar, J., Cipriani, G., *et al.* 2003. A set of simple-sequence repeat (SSR) markers covering the Prunus genome. Theor. Appl. Genet. 106(5): 819–825.

Arora, R., Rowland, L.J., Tanino, K. 2003. Induction and release of bud dormancy in woody perennials: A science comes of age. HortScience. 38 (5): 911-921.

Cabrera A, Kozik A, Howad W, Arus P, Iezzoni AF, van der Knaap E. 2009. Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. BMC Genomics. 10: 562.

Churchill, G.A., and Doerge, R.W. 1994. Empirical threshold values for quantitative trait mapping. Genetics, 138(3): 963–971.

Coville, F.V. 1920. The influence of cold in stimulating the growth of plants. J. Agric. Res. 20: 151–192.

Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W., and Arús, P. 2004. Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proc. Natl. Acad. Sci. U.S.A. 101(26): 9891–9896.

Dirlewanger, E., Moing, A., Rothan, C., Svanella, L., Pronier, V., Guye, A., *et al.* 1999. Mapping QTLs controlling fruit quality in peach [Prunus persica (L.) Batsch.]. Theor. Appl. Genet. 98(1): 18–31.

Fan, S., Bielenberg, D.G., Zhebentyayeva, T.N., Reighard, G.L., Okie, W.R., Holland, D., Abbott, A.G. 2010. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). New Phytol. 185: 917-930.

Freeling, M. 2001. Grasses as a single genetic system: reassessment 2001. Plant Physiol. 125(3): 1191-7.

Hall, A.E., Fiebig, A., Preuss, D. 2002. Beyond the Arabidopsis genome: opportunities for comparative genomics. Plant Physiol. 129(4): 1439-1447.

Hjorth, J.S.U. 1994. Computer intensive statistical methods. Chapman & Hall, New York.

Joobeur, T., Viruel, M.A., de Vicente, M.C., Jauregui, M.B., Ballester, J., Dettori, M.T., Verde, I., Truco, M.J., Messeguer, R., Batlle, I., Quarta, R., Dirlewanger, E.,

Arús, P. 1998. Construction of a saturated linkage map for *Prunus* using an almond × peach F2 progeny. Theor. Appl. Genet. 97: 1034–1041.

Kliebenstein, D.J., Gershenzon, J., Mitchell-Olds, T. 2001. Comparative quantitative trait loci mapping of aliphatic, indolic and benzylic glucosinolate production in Arabidopsis thaliana leaves and seeds. Genetics. 159(1): 359-70.

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugen. 12: 172–175.

Lander ES, Botstein D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics. 121(1): 185-99.

Lynch, M., Walsh, B. 1998. Genetics and analysis of quantitative traits. Sunderland, MA, USA: Sinauer Associates.

Murphy, W.J., Stanyon, R., O'Brien, S.J. 2001. Evolution of mammalian genome organization inferred from comparative gene mapping. Genome Biol. 2(6): REVIEWS0005.

Olukolu, B.A., Trainin, T., Fan, S., Kole, C., Bielenberg, D.G., Reighard, G.L., Abbott, A.G., Holland, D. 2009. Genetic linkage mapping for molecular dissection of chilling requirement and bud break in apricot (*Prunus armeniaca* L.). Genome 52(10): 819-28.

Paterson, A.H., DeVerna, J.W., Lanini, B., Tanksley, S.D. 1990. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics. 124(3): 735-742.

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln, S.E., Tanksley, S.D. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature. 335(6192): 721-726.

Ruiz, D., Campoy, J.A., and Egea, J. 2007. Chilling and heat requirements of apricot cultivars for flowering. Environ. Exp. Bot. 61(3): 254–263.

Sankoff, D., Nadeau J.H. (Editors). 2000. Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment and the Evolution of Gene Families (Computational Biology Series, Vol. 1). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Schmidt, R. 2002. Plant genome evolution: lessons from comparative genomics at the DNA level. Plant Mol Biol. 48(1-2): 21-37.

Scorza, R., and Okie, W.R. 1990. Peaches (Prunus persica L. Batsch). Acta Hortic. 290: 177–231.

Sherman, W.G., Beckman, T.G. 2003. The climatic adaptation in fruit crops. Acta Horticulturae. 622: 411-428.

Silva, C., Garcia-Mas, J., Sánchez, A.M., Arús, P., Oliveira, M.M. 2005. Looking into flowering time in almond (Prunus dulcis (Mill) D. A. Webb): the candidate gene approach. Theor. Appl. Genet. 110(5): 959-68.

Utz, H.F., and Melchinger, A.E. 1995. PLABQTL: a computer program to map QTL. Version 1.0. University of Hohenheim.

Utz, H.F., and Melchinger, A.E. 1996. PLABQTL: a program for composite interval mapping of QTL. JQTL 2(1).

Van Ooijen, J.W., and Voorrips, R.E. 2001. JoinMap® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.

Verde, I., Quarta, R., Cedrola, C., Dettori, M.T. 2002. QTL analysis of agronomic traits in a BC1 peach population. Acta Hortic 592: 291–297.

Voorrips, R.E. 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. J. Hered. 93(1): 77–78.

Wang, Y., Georgi, L.L., Reighard, G.L., Scorza, R., Abbott, A.G. 2002. Genetic mapping of the *evergrowing* gene in peach [*Prunus persica* (L.) Batsch]. J. Hered. 93(5): 352–358.

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J. 2003. Positional cloning of the wheat vernalization gene VRN1. Proc. Natl. Acad. Sci. U.S.A. 100(10): 6263-6268.

Zeaser D. 2001. Vegetative propagation of Gmelina arborea Roxb. In: CAMCORE. International Tree Breeding Short Course Book. North Carolina State University, Raleigh, North Carolina, USA, pp. 3-15.

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genetics, 136(4): 1457–1468.

Zhang, J., Zheng, H.G., Aarti, A., Pantuwan, G., Nguyen, T.T., Tripathy, J.N., Sarial, A.K., Robin, S., Babu, R.C., Nguyen, B.D., Sarkarung, S., Blum, A., Nguyen, H.T. 2001 Locating genomic regions associated with components of drought resistance in rice: comparative mapping within and across species. Theor. Appl. Genet. 103: 19-29.

CHAPTER FOUR

ASSOCIATION MAPPING FOR CANDIDATE GENES UNDERLYING FLORAL BUD BREAK IN APRICOT (*PRUNUS ARMENIACA* L.)

Bode A. Olukolu[1], Tatyana Zhebentyayeva, [1] Valentina Gorina,[2] and Albert G. Abbott[1]

[1]Genetics and Biochemistry, Clemson University, 100 Jordan Hall, Clemson SC 29634 USA
[2]Department of Pomology and Subtropical Crops, Nikita Botanical Garden, Yalta, Crimea, 98648, Ukraine

Corresponding author: A.G. Abbott (e-mail: aalbert@clemson.edu)

Maintenance and phenotyping of the apricot germplasm was performed by Dr. Valentina Gorina. DNA extraction of the individuals in the germplasm and the AFLP (Amplified Fragment Length Polymorphism) marker analysis was conducted by Dr. Tatyana Zhebentyayeva. Genotyping of individuals in the germplasm and data analysis (population structure analysis, linkage disequilibrium analysis and LD-based association mapping) was performed by Bode A. Olukolu.

Abstract

Despite accomplishments in defining genomic regions controlling dormancy-related traits in *Prunus* (peach and apricot) and a few other woody perennials, limitations imposed by sample size and generation cycle still persist for candidate gene (CG) discovery. We present a linkage disequilibrium-based (LD) mapping strategy aimed towards identifying CGs underlying dormancy and floral bud break (blooming date) in apricot (*Prunus armeniaca* L.). Using 100 individuals of diverse geographical origins, 240 AFLP and 8 SSR markers were used to estimate the population structure. The LD mapping was limited to marker dense genomic regions within and around previously detected major QTLs. These QTLs positioned on linkage groups (LG) 1 and 7 are consistent between the apricot and peach (*Prunus persica* L.) QTL maps. Pair-wise marker association estimates were computed based on 32 LG1 and 23 LG7 SSR markers. Marker-trait associations were revealed using a generalized linear model to eliminate false-positives due to population structure. LD decayed below the centimorgan scale, indicating insufficient marker density averaged at 0.44 and 1.58 cM on LG1 and 7, respectively. Denser marker regions were averaged at 0.1 and 0.7 cM on LG1 and 7, respectively, and revealed significant LD estimates above the LD baseline threshold. We report significant marker-trait associations and the genes these markers were derived from. The CGs on the LG1 QTLs include a transposon (*HARBINGER-LIKE*) involved in RNAi-mediated DNA methylation-induced silencing; *DAM2* (*DORMANCY ASSOCIATED MADS-box 2*); VIN3 (*VERNALIZATION INSENSITIVE 3*) or *VEL1* (*VIN3-LIKE/VERNALIZATION 5*) and *SUT1* (a *SUCROSE TRANSPORTER 1* gene). CGs on the LG 7 QTLs includes *CLF*

(*CURLY LEAF*), a component of a complex required for gene silencing; and *SQUINT* (*SQN*), which is involved in miRNA-mediated vegetative to reproductive phase regulatory pathway. Our results demonstrate that LD-based association mapping can be used for validating QTLs, fine mapping and for the detection of CGs in the *Prunus* genus.

Keywords: Linkage disequilibrium, population structure, general linear model, candidate gene

Introduction

Detection of quantitative trait loci (QTLs) and high resolution mapping to define candidate genes controlling dormancy and bud break on a fine genomic scale, like many other complex traits especially in tree species, is not possible using the conventional bi-parental crosses and linkage analysis. This is mostly due to the need for recombination events required to break up polymorphisms and resolve very small genomic regions. This in turn requires several thousand progenies, an expensive and difficult task. To compound the problem, the statistical power to detect minor QTLs especially in complex traits that lack major QTLs or that comprise several additional minor QTLs, is simply lacking in any existing mapping populations of long-lived species. To combat the problem, numerous recombination events maintained in natural populations have been routinely exploited in animal systems especially for the genetics of complex diseases in humans (Risch 2000). More studies are now using natural plant populations and their allelic variations as an alternative to standard single family mapping approaches. This approach, linkage disequilibrium-based association mapping, is not only becoming an alternative and powerful tool for identifying loci controlling complex traits but also a more realistic and cost-effective way for high-resolution mapping of QTLs compared to single family linkage mapping. Basically, association mapping explores several meiotic events accumulated over numerous generations, hence, establishing population genealogy, marker-marker association and eventually marker-trait association (Remington *et al.* 2001, Thornsberry *et al.* 2001). According to a review by Hirschhorn and Daly (2005),

association mapping and linkage analysis are considered complimentary to each other because they incorporate prior knowledge, cross-validation and statistical power.

The resolution captured by linkage disequilibrium-based (LD) mapping is only a function of the landscape of linkage disequilibrium across the genome. Linkage disequilibrium refers to the non-random association of alleles between loci. The landscape of the LD is in turn influenced by genetic and non-genetic factors like recombination, drift, selection, mating pattern and population admixture (Flint-Garcia *et al.* 2003). The structure of LD is reflected in the decay of LD over genetic distances in a population, which then determines the marker coverage required within a genomic region to perform association analysis. Studies have shown that LD decay varies considerably within and between species as well as across genomic regions. Until recently, one of the major drawbacks of LD mapping has been spurious association due to population structure leading to elevated levels of false positives. The complex evolutionary history and breeding in plants is a causative factor for this. Regardless of this drawback, populations considered for association mapping still include breeding and gene bank collections of accessions, breeding lines and germplasm collections. These collections are most ideal for association studies because curators routinely evaluate them, thus providing a wealth of valuable phenotypic data readily available for QTL/gene discovery.

Recent linkage mapping studies in peach (Fan *et al.* 2009) and apricot (Olukolu *et al.* 2009) for bud break (floral and vegetative) and chilling requirement have identified several QTLs with major contribution on linkage group 1 and 7. Although most of this

QTLs are stable between the two species and explain most of the phenotypic variation (about 70%), the QTL intervals lack sufficient map resolution to determine which of the candidate genes within the intervals control the trait. This study aims to validate some of the QTLs detected in these previous linkage maps based on the variation in the blooming date (floral bud break) in a diverse and structured population of apricot (*Prunus armeniaca* L.).

This study evaluates the extent of population structure in the apricot population based on a genome wide scan using AFLP markers (Falush *et al.* 2007), estimating pair-wise marker association, extent of LD decay within genomic regions spanning QTLs on LG1 and 7, as well as detection of marker-trait association. SSR markers were designed mostly from genic regions to span and saturate these QTL intervals. The wide genetic diversity and the outcrossing nature of apricot (Kostina 1946, Layne *et al.* 1996, Faust *et al.* 1998, Zhebentyayeva *et al.* 2010) make it ideal for association mapping according to criteria designated by (Nordborg and Donnelly 1997). In this study, we report the first LD-based association mapping in *Prunus* for identification of candidate genes underlying blooming date (floral bud break). These results demonstrate the power of coupling single family QTL analysis with association approaches to further refine QTL intervals and to identify potential candidate genes for traits influenced by these QTL loci. It also serves as a potential model for parallel studies in other *Prunus* species, with possible extension to other Rosaceae species.

Materials and Methods

Plant materials

Sampled individuals are native to diverse geographical regions (Table 4.1) and comprised of 100 accessions. Only accessions with documented introduction from respective geographical region were considered as indigenous germplasm. The core set of cultivars comprised of 83 apricot accessions previously described in a genetic diversity study by Zhebentyayeva *et al.* (2003). These accessions represent Chinese, Central Asian (Fergana, Zeravshan, Khorezm and Kopet-Dag subs), European and hybrid cultivars of a known pedigree. Also included are cultivars associated with PPV resistance breeding and non-domesticated germplasm i.e. Dzhungar-Zailij population of *P. armeniaca* and wild species: *Prunus mandshurica* (Maxim), *Prunus sibirica* L., *Prunus sibirica* var. *davidiana* (Carrière), *Prunus armeniaca* var *ansu* (Maxim.) Kost., *Prunus mume* (Sieb. et Zucc.) and alpine plum *Prunus brigantina* Vill. (see Zhebentyayeva *et al.*, 2008). The list of indigenous cultivars was enriched with two North African (Tunisia) cultivars Quardi and Sayed.

Phenotypic data

All but 5 accessions (i.e. Goldrich, LE2904, LE3276, Sayed and Quardi) were maintained at the State Nikita Botanical Garden (Crimea, Ukraine). Apricot cultivars were grown under standard agrotechnique in two orchards on the northern coast of the Black Sea in Yalta characterized by warm Mediterranean climate; one orchard (44°30' 50.81" S 34°

13' 52. 22" E) planted in 1973 and another one (44° 30' 52.44"S 34°13' 59.06" E) planted in 1990-1991. Most of cultivars were replicated in 2 locations in Gvardejskoe (steppe zone of Crimean peninsula) characterized by temperate continental climate with fluctuating winter temperatures; one orchard (45°09' 23. 82" S 33°58' 37, 93" E) planted in 1973 (eradicated in 1996) and another one (45°09'01.38" S 33°59' 17.37" E) planted in 1991 through 1994. Eastern-European cultivar Krasnoshchekii and its later blooming clone Jubileinyi were randomly planted in all orchards and served as control for normalization of bloom date across the locations.

Phenotype evaluation was done according to certified protocol for selection and evaluation of new stone fruit cultivars in the South of USSR by Ryabov (1969). Yearly during the entire period of blooming, tree phenology has been recorded every other day and included 5 records on date of: 1) fully swollen buds, 2) flower bud emergence, 3) emergence of petals, 4) start of the bloom, and 5) full bloom (at least 50% flower buds on tree are open). In this association study we used a data for full bloom as blooming date and categorized apricots into 5 classes according to their blooming date: 1) March 1st - 10th, 2) March 11th – 20th, 3) March 21st – 31st, 4) April 1st – 10th and 5) April 10th – 20th. Average date of blooming was calculated for entire period of observation. Cultivar assignments to phenotypic classes were iteratively verified for entire period of data collection spanning from the first year of flowering (1976) till 2009. Due to consistency of the ensuing phenotypic classes, several breeding programs have consistently and successfully used these accessions for production of hybrids with specific desirable times for blooming.

Table 4.1: List of accessions, their geographical origin and serial numbers corresponding to number assigned in Structure analysis.

Central Asian	***sno***	***Central Asian***	***sno***	***European***	***sno***
Arzami	3	Kzyl Khorezmskii	44	Alberge de Tur	1
B-1-11	4	Kzyl Khurmai	45	Ananasnyi Tsurupinskii	2
B-4-5	5	Kzyl Uryuk	46	Bergeron	8
B-5-3	6	Lyuchak Sumbarskii	52	De Compot	13
Badami	7	Maftobi	53	Early Gold	21
Gulyunghi	23	Mamuri	55	Jubileinyi	27
Iskadari	26	Mirsandzhali	59	Kantsler	30
Kandak-10	28	Nukul Citronnyi	64	Krasnoshchekii	40
Kandak-12	29	Oranzhevo-krasnyi	66	Luizet Krupnoplodnyi	50
KB-12	32	Paivandi Bukharskii	74	Precoce d'Italia	76
KB-9	33	Rukhi Dzhuvanon	81	Real d'Imola	80
Khurmai	34	Rukhi Dzhuvanon Surkh	82	Tilton	93
Khurmai Rannii	35	Samyi Rannii	83	Velkopavlovichka	95
KK(2) N1	37	Supkhani	91	Vengerskii Krypnyi	96
Kok-pshar	38	Tadzhabai	92	Vynoslivyi	98
Kolon Boboi	39	Zard	100		
Kunduzi	42				
Hybrid cultivars	**sno**	**Iran-Caucasian**	***sno***	***Chinese***	***sno***
Dionis	14	Daradzhi ek Shabistr	12	Da-bei	9
Krimskii Amur	41	Katuni	31	Da-chuan-che N1	10
Lunnik	51	Kurbane Marache	43	Da-chuan-che N2	11
Medunets Krimskii	57	Mascat	56	In-ben-sin	25
Naryadnyi	61	Nakhichevanskii	60	Kitaiskii	36
Naslazhdenije	63	Nasera Tabris	62	Lao-yech-lian	47
Olimp	65	Ordubad	67	Mai-che-sin	54
Parnas	75	Shalakh	87	Mi-bada	58
Pruisadebnyi Rannii	77	Shekarpara de Semnan	90	Pui-sha-sin	78
Satser	84	Vaagas Vardaguin	94	Shantunski	88
Shedevr	89			Yuan-sin	99
Non-domesticated	***sno***	***PPV resistant cultivars***	***sno***	***P. armeniaca, wild***	***sno***
P. ansu	68	Goldrich	22	Dzhungarskii 8/55	15
P. manshurica	69	Harlayne	24	Dzhungarskii 18/63	16
P. mume, N 15	70	LE2904	48	Dzhungarskii 18/64	17
P. mume, N18	71	LE3276	49	Dzhungarskii 18/68	18
P. sibirica var *davidiana*	72	Stark Early Orange	86	Dzhungarskii 18/75	19
P.sibirica	73	Vestar	97	Dzhungarskii 18/78	20
Northern African	***sno***				
Sayed	79				
Quardi	85				

Note: accession sno correspond to numbering on bar plot of population structure.

DNA extraction

Genomic DNA was extracted from fresh young leaves using a modified CTAB/PVP method (Sivolap *et al.* 1998) as described by Zhebebntyayeva *et al.* (2003). Working solutions of genomic DNA at 10 ng/µl in 0.1 × TE buffer were prepared for SSR analysis and kept at 4 °C.

Genotypic data

In this study we obtained data for 10 SSR loci (generated from 8 primer combinations) and 240 AFLP loci (generated from 8 primer combinations) from previous studies, in which several geographical regions and hybrid cultivars were excluded from analysis for reasons described in original papers (Zhebebntyaeva *et al.* 2003, 2008). The Quardi and Sayed cultivars absent in the previous studies were treated as missing data when these preexisting genotypic data were used in the analysis. New set of SSR data generated for all 100 accessions included 45 SSR loci (generated from 44 primer combinations) anchored on the *Prunus* reference map (Aranzana *et al.* 2003) and the *Prunus* bin map (Howad *et al.* 2005). Most of these additional *Prunus*-based SSR markers were designed for further saturation within the QTL regions on LG1 and 7 detected in both peach (Fan *et al.* 2009) and apricot (Olukolu *et al* 2009) maps. The specific primer pairs, amplification conditions, radioactive labeling, and polyacrylamide gel electrophoresis were employed as described in Combes *et al.* (2000). For primer design we used genomic sequence derived from the annotated EVG region in peach (Bielenberg *et al.* 2008) and the whole peach genome assembly available at http://peachzome.phytozome.net. Using

QTL flanking markers we delineated genomic regions inside of QTL interval and selected di-nucleotide repeats close to predicted genes or inside of 5` UTR, 3` URT or introns, thereafter, the markers represented mostly genic regions. Homology searches for predicted proteins were done against the Arabidopsis reference protein database using a BLASTp program available on the NCBI website (http://www.ncbi.nlm.nih.gov/). The list of primers for SSR markers and corresponding predicted genes are available in Appendix D and E, respectively.

We used amplified fragment length polymorphism (AFLP) marker analysis as an inexpensive and fast alternative for a genome wide scan. The AFLP analysis was performed according to Vos *et al.* (1995) as described in Zhebentyaeva *et al.* (2008).

Statistical analysis

Population structure: Eight unlinked SSR loci (UDAp485, UDA410, UPD98-4069, Pchgms144, Pchgms137, Pchgms106, Pchgms20b and EPDCU3454), one on each linkage group and 240 AFLP loci were used for evaluation of population structure. Since the model that STRUCTURE software is based on assumes that loci are independent within populations (i.e. not in LD); AFLP loci, which are supposedly unlinked or mostly distantly linked markers due to random distribution over the genome were used to avoid the structure analysis being skewed by using linked SSR markers (Falush *et al.* 2007). Studies by Falush *et al.* (2003) and Conrad *et al.* (2006) have shown that structure performs reasonably well despite the data not completely fitting the model provided there is enough independence across regions that LD within regions does not dominate the data

(Pritchard *et al.* 2000a). This was further verified in our data set by using only 8 unlinked SSR marker sets to assess the confidence with which accessions are assigned to subpopulations for which P > 0.5. All program parameters were kept uniform for each run. The program STRUCTURE version 2.3.2 (Pritchard *et al.* 2000b) was used to test the hypotheses of K = 1 to K = 12 i.e. 1 to 12 populations using an admixture model, a burn-in phase of 10^5 and a sampling phase of 5 x 10^5 replicates. The optimal division of the population into subpopulations was determined when the probability of K was very small for K less than the appropriate value (effectively zero) and then more-or-less plateaus for larger K (documentation for STRUCTURE software, Pritchard *et al.* 2009).

Within subpopulations, it is assumed that the loci are at Hardy-Weinberg equilibrium and are in linkage equilibrium i.e. individuals are assigned to populations in such a way as to achieve this (Pritchard *et al.* 2009). The overall and subpopulation Fst and heterozygosity (Falush *et al.* 2003) parameters estimated using the STRUCTURE version 3.2 (Pritchard *et al.* 2000b) were as a measure of genetic diversity.

Linkage disequilibrium: The LD parameters for pair-wise significance between markers were estimated using the Tassel software (http://www.maizegenetics.net/). The LD parameters were computed by 10^5 permutations and without the rapid permutation test. Prior to marker-trait association, LD using the square of the coefficient of correlation (r^2, Pritchard and Przeworski 2001), was estimated separately for linked loci on the same QTL genomic region on LG1 and 7. The 95^{th} percentile of the square root transformation of the estimates was used to establish a population-specific threshold for evidence of

linkage for each genomic interval (Breseghello and Sorrels, 2006). The r^2 values were plotted against the map distance (cM) on the linkage map and a regression line was drawn by a second-degree loess curve (Cleveland 1979) using the SAS version 9.1 (SAS Institute). The intersection of the loess curve (a type of regression curve) to the baseline (95^{th} percentile critical value of r^2) was regarded as the estimate of the level of LD within each genomic region. The map position of anchored peach SSR markers were obtained from genetic distances from a peach F_2 map (Fan *et al.* 2009), while newly designed peach SSR markers were mapped in the same peach F2 population (Zhebebntyeva *et al.* unpublished).

Association analysis: Associations between SSR markers and the blooming date trait were tested using the General linear model (GLM), where the tested loci were considered as fixed-effects factor and the subpopulations (covariates) were considered as a random-effects factor (Kennedy *et al.* 1992). The GLM model introduces population structure (Q) to reduce false positives due to population stratification i.e.

$$\text{GLM statistical model: } Y = X_{\alpha} + Q_{\beta} + \varepsilon,$$

where Y is the vector of phenotypes, X is the vector of single locus genotypes, α is the vector of fixed effects of the *n*–1 genotype classes, Q is the matrix of the $K - 1$ subpopulation ancestry estimates for each individual from STRUCTURE, β is the vector of the fixed effects for each of the subpopulations, and ε is the vector of residual errors. Significance of associations was based on F-test at *p*-adj, which corresponds to *p*

corrected for multiple testing. Corrected significance levels *p*-adj was computed by 10^5 permutations within a QTL genomic region.

RESULTS

Marker polymorphism

A total of 240 AFLP loci were detected by 8 AFLP primer combinations, while a total of 55 SSR loci were detected by 52 SSR markers after screening and running 102 SSR markers. Thirty-two SSR loci were genotyped for the LG1 QTL genomic regions, while 23 SSR loci were genotyped for the LG7 QTL genomic regions. The total number alleles varied between 2 and 23 alleles for the SSR loci, while the AFLP markers were scored for only 2 alleles per locus i.e. dominant markers scored as absent or present.

Population Stratification and Genetic Structure

Based on 100 genotyped individuals, the sample was stratified into 7 sub-populations (Fig. 4.1). The posterior probability of the data set peaked at 7 subpopulations and subsequently plateaus from 7 to 12 subpopulations (Fig. 4.2). The increase in posterior probability was not significant after 7 populations, while more accessions were split between 2 or more subpopulations from 8 to 12 subpopulations. Based on the contribution of a subpopulation for each accession, the subpopulations comprised 11, 17, 25, 24, 10, 2 and 11 accessions (Appendix C).

The subpopulations were allocated to geographical sampling locations by discretely assigning accessions with P > 0.5 and their corresponding origin to each subpopulation (Table 4.2). Following the allele-frequency divergence (net nucleotide distance) among subpopulations, relationships between subpopulations were established (Table 4.3). The European accessions split between subpopulation 1 and 7, were shown to be more closely related than other subpopulations, while the subpopulation 2 (Chinese), 3 (Central Asia, Fergana) and 4 (Central Asia, Zeravshan) were all closely related although the two Central Asia subpopulations were more similar. Subpopulation 5 is most related to the subpopulation 3 (Central Asia, Fergana) and consist mostly of hybrid cultivars, some of which were also observed in subpopulation 4 (Central Asian, Zeravshan). The subpopulation 6 comprised strictly of the two *P. mume* species that were the most distantly related from any of the subpopulations. They are most related to the Central Asian and Chinese subpopulations, which also contain other non-domesticated species. The Iran-Caucasian accessions were dispersed across the 2 Central Asian subpopulations and the European subpopulation 7. The northern African accessions on the other hand grouped together with the European subpopulation 1. The cultivars involved associated with PPV resistance breeding (Goldrich, Harlayne, SEO and Vestar) appear to be hybridizations resulting from the Chinese and European subpopulation 1.

The confidence of assigning accessions to subpopulation was assessed between 2 population structure analyses based on 8 unlinked SSR loci and 248 loci (240 AFLP and 8 SSR markers) within only 3 subdivisions. At k = 3; 83.1, 77.8 and 79.2 % of the accessions where consistently assigned in 3 subdivisions comprising Central Asian,

Table 4.2: Assigning subpopulations to geographical regions based on member accessions and inferring accession ancestry.

Subpopulation	Predominant group	Other groups
1	European	Northern African, PPV resistant cultivars
2	Chinese	Non-domesticated, PPV resistant cultivars
3	Central Asian, Fergana	*P. armeniaca*-wild, Iran-Caucasian
4	Central Asia, Zeravshan	*P. armeniaca*-wild, Iran-Caucasian, hybrid, Chinese
5	Hybrid cultivars	Iran-Caucasian, Europe
6	*P. mume*	-
7	European group	Iran-Caucasian

Note: Accessions are grouped based on geographical origin, pedigree (PPV and hybrid cultivars) and taxonomic classification (*P. mume* and *P. armeniaca*).

Table 4.3: Allele-frequency divergence between subpopulations (net nucleotide distance) computed using point estimates of probability values.

	1	2	3	4	5	6	7
1	-						
2	48.8793	-					
3	46.2574	35.8506	-				
4	42.9624	43.5332	31.5479	-			
5	51.3395	62.0901	42.0904	47.3089	-		
6	87.9173	84.8668	77.4058	72.7285	91.4628	-	
7	42.1615	74.8547	66.9399	62.7159	62.7465	112.5728	-

Note: Net nucleotide distance, average amount of pairwise difference between alleles from *different* populations (similar populations have distances near 0), beyond the amount of variation found within each population.

Table 4.4: Fst estimates for each sub-population and average distances between individuals in same cluster (expected heterozygosity).

Subpopulation	Fst	Expected heterozygosity
1	0.5711	-228.469
2	0.4758	-215.768
3	0.3421	-205.535
4	0.408	-212.039
5	0.7606	-249.524
6	0.8391	-253.362
7	0.9837	-278.023

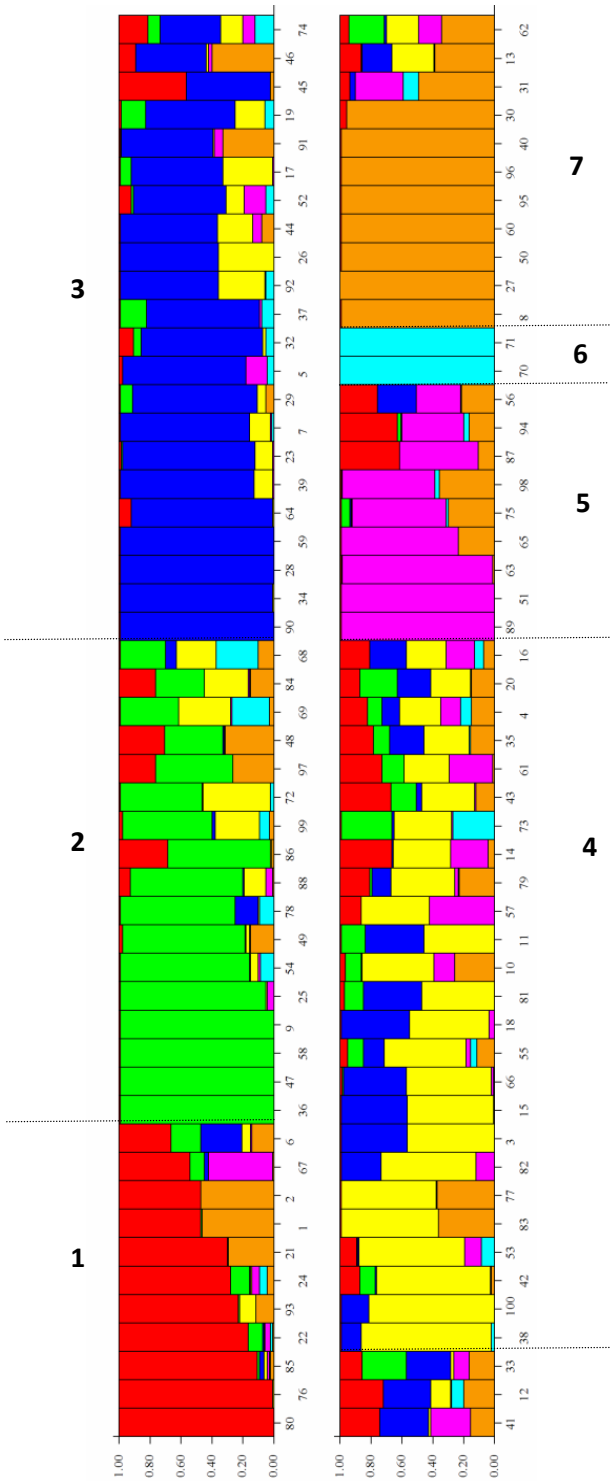


Figure 4.1: Bar plots showing population stratification and ancestry of 100 apricot accessions under an admixture model. Each accession is assigned to subpopulations based of proportional membership (vertical bars expressed as %) at k = 7. Each subpopulation is coded by a different color; Red (European), Green (Chinese), Blue (Central Asian, Fergana), Yellow (Central Asia, Zeravshan), Purple (Hybrid cultivars), Turquoise (*P. mume*), Brown (European group).

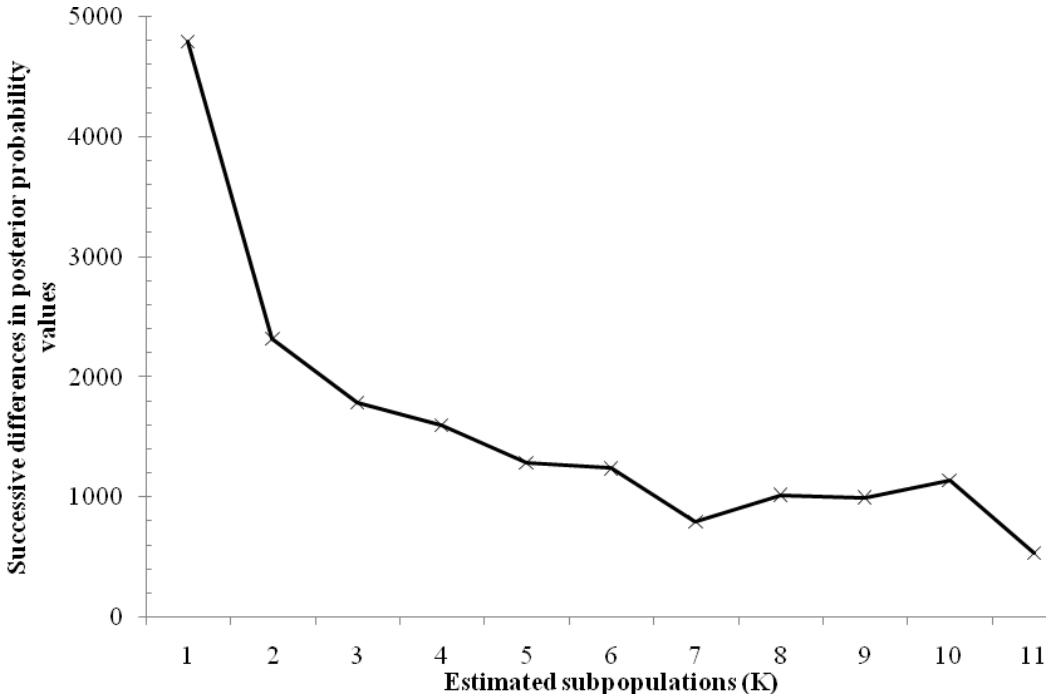


Figure 4.2: Successive differences between posterior probability values from K=1 to k12 (subpopulations) plotted against K values to establish the appropriate number of subpopulation.

Chinese, and European populations. Fst values across the subpopulations was 0.0010, indicating moderate differentiation and high diversity within the apricot germplasm (Table 4.4). Conversely, individual subpopulations had relatively high Fst values, especially in the European subpopulation 7 and *P. mume*, which had 0.984 and 0.839, respectively. The Central Asian and Chinese subpopulations had the least Fst estimate and the most heterozygosity, indicating highest diversity, hence, suggesting these geographical origins as the center of domestication for apricot as reported by Zhebentyayeva *et al.* (2008).

Linkage disequilibrium

Linkage disequilibrium was estimated within a genomic region spanning about 14 and 25 cM on LG1 and 7, respectively. Since the study focused mainly on identifying candidate genes within detected QTLs in apricot and peach crosses (Fan *et al.* 2009, Olukolu *et al.* 2009), marker saturation was emphasized for these regions rather than flanking genomic regions. The average value of marker density within QTL intervals was 0.1 and 0.7 cM on LG1 and 7, respectively. The average marker density across the scanned genomic regions is 0.44 and 1.58 cM on LG1 and 7, respectively.

The pair-wise estimates among 32 markers on the LG1 genomic region (248 estimates) varied between 0.0037 and 0.3006, with an average of 0.0276. On the LG7 genomic region, pair-wise estimates among 23 markers (127) varied from 0.0056 to 0.0681, with an average of 0.0198. The pair-wise LD estimates are shown in the LD plot (Appendix F). The 95th percentile of the distribution of these estimates was implemented as a

population-specific threshold and as an evidence of linkage. The 95th percentile threshold for the r^2 was estimated at 0.0562 and 0.0381 for the LG1 and 7 genomics regions. On the LD vs. genetic distance (cM) plots (Fig. 4.3), 14 and 8 of the r^2 estimates were above the baseline of 0.0562 and 0.0381 for LG1 and 7 genomic regions, respectively, but the loess curve fitted on the distribution of the r^2 estimates did not reach baseline at any point. This implies that marker density was not enough to detect consistent LD and the LD decayed at an average distance of 0.44 and 1.58 cM across the intervals scanned on LG1 (14 cM) and 7 (36.4 cM), respectively (Fig. 4.3). Most of the estimates above the baseline were observed within the marker dense QTL regions and at an average marker density of 0.1 and 0.7 cM on LG1 and 7, respectively. Although there was stronger LD on the LG1 genomic region, it decayed faster than the LD observed on that of LG7.

Association mapping

The association of SSR markers with the blooming date data was performed in the presence of population structure and tested using a general linear model (GLM). More than one significant marker was detected for each QTL interval indicating the presence of multiple candidate genes that exist in a cluster of genes underlying similar molecular pathways (Fig. 4.4). Four Significant markers (Fig 4.4) were detected on the LG1 QTL around the *EVG* locus (Bielenberg *et al.* 2008), while only one marker showed association on the LG1 QTL downstream of the QTL spanning *EVG* locus.

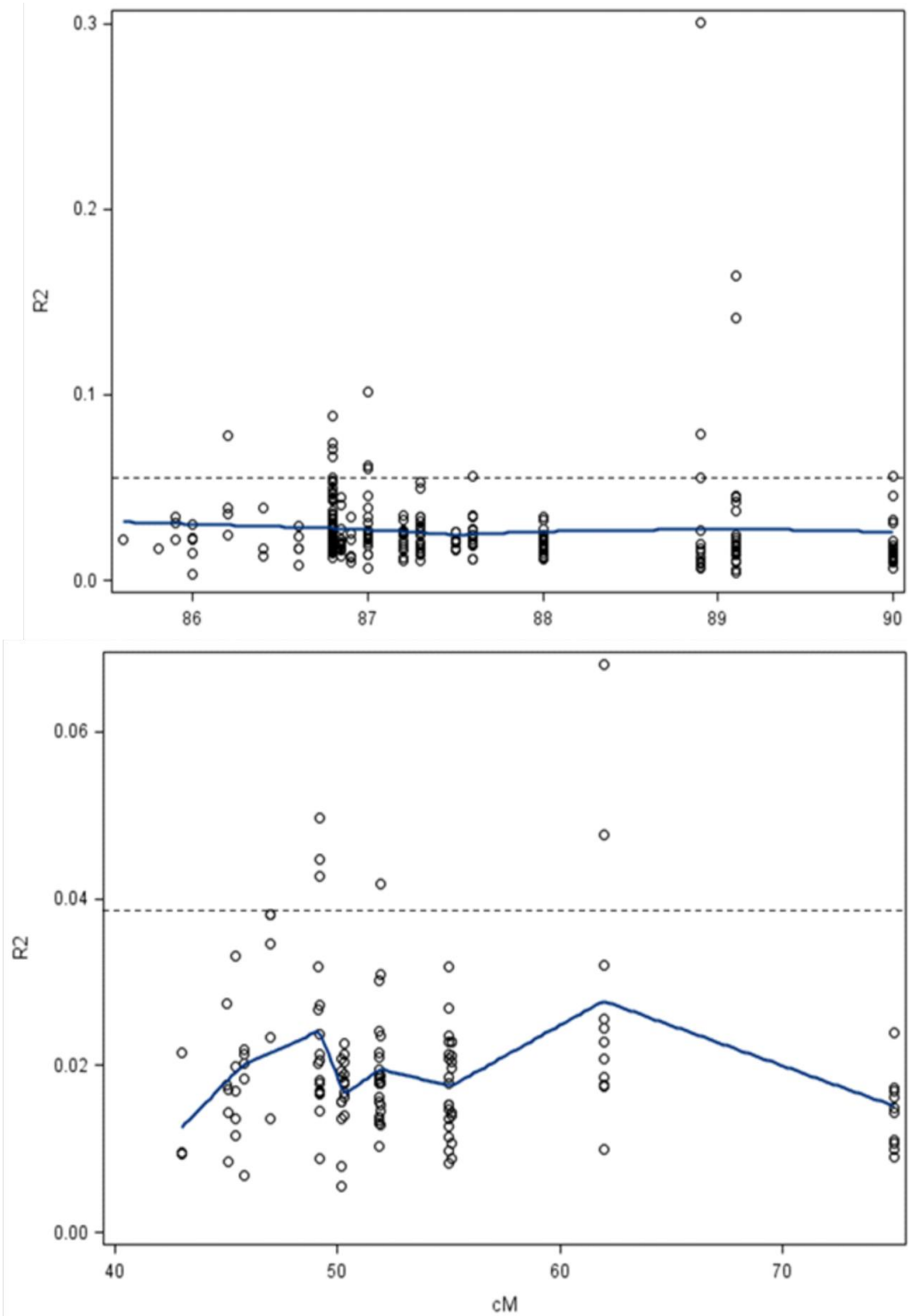


Figure 4.3: LD estimates of r^2 plotted against genetic linkage distance on LG1 (above) and 7 (below). Horizontal dotted lines show the 95th percentile of the distribution of unlinked r^2. Curves were fitted by second-degree loess. Axis scales vary for each plot.

These associated markers correspond to CGs that include *HARBINGER-LIKE* transposon, *DAM2* (*DORMANCY ASSOCIATED MADS-box 2*), *VIN3* (*VERNALIZATION INSENSITIVE 3*) or *VEL1* (*VIN3-LIKE/VERNALIZATION 5*) and *SUT1* (a *SUCROSE TRANSPORTER 1*).

The respective adjusted p-values are 0.0232, 2×10^{-5}, 1.7×10^{-4}, 0.0186 and 0.0011. The genomic contexts of 2 markers within the LG7 QTLs, Pcghms107 and Pchgms115 (Fig. 4.4) were fully characterized and the corresponding CGs includes *CLF* (*CURLY LEAF*) and *SQN* (*SQUINT*) genes, while the adjusted p-values are 0.0274 and 0.028, respectively. The probable candidate genes corresponding to the other associated markers on LG7 were not resolved although likely candidate genes lie within close proximity. These genes include *DEFICIENT IN DNA METHYLATION 1* (*DDM1*), a chromatin remodeling factor required for maintaining DNA methylation, which is positioned 19 kb downstream of Pchgms90 and Pchgms9; and *VERNALIZATION INSENSITIVE 3-LIKE 1* (*VRN5/VIL1*), which lies between Pchgms100 and Pchcms2.

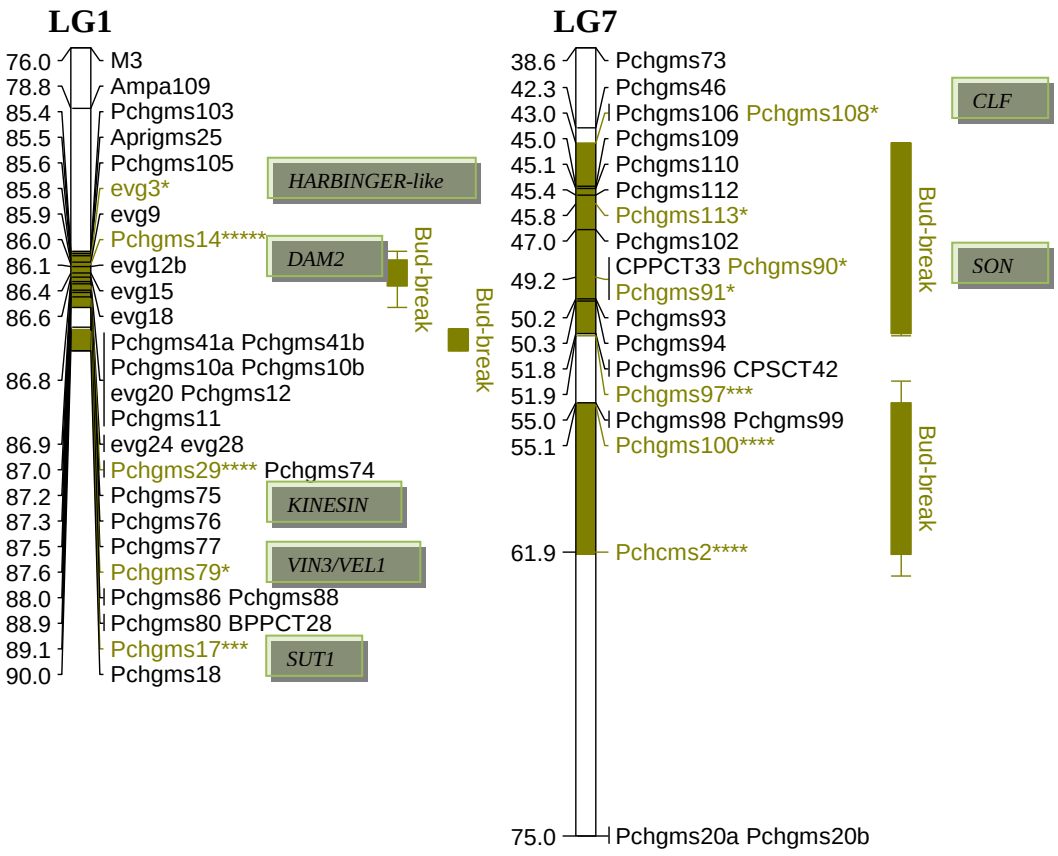


Figure 4.4: Linkage groups 1 and 7 showing loci that are associated with blooming date. The asterisks show the $\chi 2$ p levels of significance (*, 0.05; **, 0.01; ***, 0.001)

Discussion

We report the first LD-based association mapping of a complex trait in *Prunus* towards evaluating the utility of the fine mapping approach in diverse and well curated apricot germplasm and other related Rosaceae species. This will provide preliminary LD parameters that will be useful in improving the approach and designing tools for association mapping. The germplasm developed for this study represents a broad genetic base of apricot from diverse geographical origins where the crop has been domesticated as well as collected from the wild.

Population structure

Although the population analysis identified 7 subpopulations, 3 subdivisions can be delineated based on geographical origin and genetic similarity of 2 European and 2 Central Asian subpopulations. This is in agreement with previous genetic studies by Zhebentyayeva *et al.* (2003) that produced 3 main clusters each with smaller subdivisions. The wide genetic diversity observed in the Chinese and Central Asian subpopulations in this study further support that China and Central Asia are the primary centers of apricot domestication (Vavilov 1951, Kryukova 1989, Zhebentyayeva *et al.* 2010). Occurrence of non-domesticated wild species in the Chinese and Central Asian subpopulations also supports the claim by Mehlenbacher *et al.* (1990) that apricot domestication occurred in two separate regions (i.e. in north and northeastern China where the wild *P. armeniaca* range overlaps with that of *P. mandshurica* and *P. sibirica*).

Our study also support inferences made from a studies about the origin of the North American PPV (Plum Pox virus) resistant cultivars, which suggest contribution from European and Chinese parental genotypes (Zhebentyaeva *et al*. 2008, Pedryc *et al*. 2009).

Moreover, our results on population structure presented here allow direct comparison of model-based clustering method using a Bayesian approach (implemented in STRUCTURE) and distance-based Neighbor Joining (NJ) method used in a previous publication (Zhebentyayeva *et al.* 2008). The structure analysis in this study produced accurate assignment of Northern American apricots as admixed individuals from a European-Chinese genetic background while NJ failed to detect genetic relatedness of Northern American and Chinese cultivars.

Linkage disequilibrium

Even though linkage maps derived from small progeny numbers are inherently inaccurate with reference to marker order and map distances, the pattern of LD observed in this study was relatively consistent with the map positions (Fan *et al.* 2009, Olukolu *et al.* 2009). Although previous studies in other plants have proposed the level of LD vs. the map distance required for LD decay as an arbitrary threshold value of $r^2 = 0.1$ (Remington *et al.* 2001, Nordborg *et al.* 2002, Palaisa *et al.* 2003), we adopted a LD threshold value defined in comparison with the LD observed among unlinked loci in a sample (Luo *et al.* 2000, Breseghello and Sorrels 2006, Laurie *et al.* 2007). The 95^{th} percentile of the distribution of r^2 estimates defines the sample-specific critical threshold value (baseline LD). The point at which regression curve (drawn with the second-degree

loess) intersects the baseline defines the extent of LD due to linkage. The stronger LD on LG1, with an average value of 0.0276 and maximum value of 0.3, confirms the paradox observed between high LD and marker-marker association. The higher LD on the LG1 QTL genomic region results in a lower marker density requirement and more potential of detecting markers strongly associated with the target gene polymorphism even if distantly physically linked. Conversely, the lower LD (average of 0.0198 and maximum value of 0.0681) on the LG7 QTL genomic region requires more marker density, although the resolution of diagnostic markers will be higher.

Across the genomic regions scanned, LD was not sufficient at average genetic distance of 0.44 and 1.58 cM on LG1 and 7, respectively. The extent of LD is similar to that estimated in apple (*Malus x domestica* Borkh), which decayed at distance greater than 1 cM (Micheletti *et al.* 2010). The average map distance of markers within the marker dense QTL regions that demonstrated LD were estimated at 0.1 and 0.7 cM on LG1 and 7, respectively. This suggests the required marker density for sufficient LD in future studies. The physical map distances in the marker dense QTL region of the LG1 correspond to approximately 130 kb of a sequenced BAC containing the *EVERGROWING* locus (Bielenberg *et al.* 2008). Based on estimates from contigs of the peach physical map, the marker dense QTL region on LG7 corresponds to approximately 500 kb. The LD estimates on the marker dense region of LG1 QTL are also comparable to that observed in *Arabidopsis* where $r^2 = 0.1$ within about 250 kb in the genomic region of the FRI gene (Nordborg *et al.* 2002) and in the genomic region of the rice resistance gene, *xa5*, where r2 > 0.10 at a physical distance > 100 kb (Garris *et al.* 2003).

Association mapping and candidate genes

This study focused on estimating LD and detecting marker-trait association within genomic regions of previously identified major QTLs on LG1 and 7. Since some markers were designed within genic regions, we were able to implicate candidate genes playing a role in the phenotypic variation observed for this trait. The *DAM2* gene is one of the six dormancy related MADS-box genes within the peach *EVERGROWING* locus that are now being scrutinized as major candidates for dormancy, chilling and bud break. Characterization of this locus include several linkage mapping studies that have defined a QTL around their genomic location (Fan *et al.* 2009, Olukolu *et al.* 2009); sequencing and annotation the locus (Bielenberg *et al.* 2008), comparative analysis with transcription factor homologs in Arabidodpsis (Jiménez *et al.* 2009); their seasonal and photoperiodic expression patterns (Li *et al.* 2009) and validation of the role they play in leafy spurge and peach dormancy (Horvath *et al.* 2010, Jiménez *et al.* 2010). The peach DAM genes were shown to be homologous to the *Arabidopsis SVP/StMADS11* lineage of type II $MIKC^C$ MADS-box genes and were suggested to have expanded through serial tandem gene duplications (Jiménez *et al.* 2010). These genes in *Arabidopsis* have been associated with vegetative to reproductive meristem transition, with the *AGL22/SVP* genes acting as a flowering repressor (Hartmann *et al.* 2000); while it's close homolog *AGL24* has a reverse effect (Michaels *et al.* 2003, Yu *et al.* 2004). The *DAM2* gene was the only *DAM* gene consistently detected in this study. A closer look at the genomic context in which the 6 genes exist, current expression profiles and their functional homology with the two *SVP* gene homologs mentioned above suggests their function in controlling dormancy.

Sequence analysis between a Nemared BAC clone and the dihaploid Lovell-derived assembled peach genome reveal that the copy number of the *HARBINGER-like* transposon is variable between the two cultivars. The Nemared allele sequenced contains two copies while the Lovell allele contains only one copy of the transposon, suggesting a transposon copy number-dependent locus underlying the trait. The role the associated *HARBINGER-LIKE* candidate gene detected upstream of the *DAM* genes plays in RNAi-mediated DNA methylation-induced gene silencing (Numa *et al* 2010) correlates with the seasonal expression patterns of the DAM genes (Li *et al.* 2009) and the seasonal methylation status as reflected in a woody perennial plant, azalea (Meijón *et al.* 2010). Overlaying the global methylation status in azalea and the expression of the DAM genes suggest that the upregulation of *DAM1 to DAM4* are due to demethylation mediated by the *HARBINGER-LIKE* transposon. It appears the distal distance of the *DAM5* and *DAM6* probably excludes them from this transposon-mediated regulation since their expression does not change during this sharp decrease in global methylation. The *DAM5* and *DAM6* gene expression are proposed to be controlled by decreasing day length before winter and repressed by successive chilling accumulation (Jiménez *et al.* 2010) though this has not been unequivocably detemined. In addition, a *Ty1-copia* element on the minus strand and sharing the same 3` UTR with the *DAM2* candidate gene might explain the difference in its expression pattern, since it was observed to be downregulated earlier than the *DAM1*, *DAM3* and *DAM4* genes, while its upregulation coincides with the onset of bud break (Bielenberg *et al.* 2008). *Ty1-copia* retrotransposons have been described to be also involved in DNA methylation (Zhong *et al.* 2009). These evidences implicating the

HARBINGER-like and *Copia*-like elements indicate that these mobile elements have been co-opted by their host genome and are now a major driving force in plant adaptation (Zhong *et al.* 2009).

Downstream of the *evg* locus are 3 other candidate genes, *KINESIN*, *VIN3/VEL1* and *SUT1* that are well characterized in other species. *KINESINs* have been described extensively in plants and are ATP driven *KINESINs* possess a motor domain for binding to microtubules, acting as motor proteins that play a critical role in cell division (Sharp *et al.*, 2000, Wittmann *et al.* 2001). Perhaps, *KINESIN* plays a role in cell cycle regulation, since following dormancy floral bud break, active floral bud growth is associated with the actions of hormones and increased cell division. During this dormancy release, gene expression changes are tightly coordinated with the cell cycle (Devitt and Stafstrom 1995, Cambell *et al.* 1996, Horvath *et al.* 2002, Freeman *et al.* 2003).

The *VIN3/VEL1* gene family is involved in flowering within the vernalization pathway (Sung and Amasino 2004). The response to vernalization is facilitated by a cascade of gene regulatory networks, that are initiated during prolonged cold exposure by the induction and up-regulation of the homeodomain finger gene *VIN3* (*VERNALIZATION INSENSITIVE 3*) and results in the chromatin-based and mitotically stable repression of the *FLC* (*FLOWERING LOCUS C*) gene (Sung and Amasino 2004), which in turn suppresses flowering. This implicates *VIN3/VEL1* as a positive regulator of blooming date i.e. induction of *VIN3* leads to a progressive downregulation of *FLC* and or in the case of *Prunus*, a flowering inhibiting *AGL22/SVP-like* (*SHORT VEGETATIVE PHASE*)

gene leading to floral bud break. *DAM5* and *DAM6* are possible functional homologs of *AGL22/SVP* since they are upregulated prior to dormancy and downregulated during dormancy before floralbud break (Jiménez *et al.* 2010). If the *DAM5* and *DAM6* locus play a similar role as the *Arabidopsis FLC* or *AGL22/SVP*, then the *VIN3/VEL1* locus might be the *DAM5* and *DAM6* repression factor during vernilzation, leading to a progressive reduction to their transcript accumulation as reflected in expression studies (Li *et al.* 2009, Jiménez *et al.* 2010).

The role played by *SUCROSE TRANSPORTER 1* (*SUT1*) is supported by studies implicating sucrose and cytokinin as signaling molecules (Bernier and Périlleux 2005). Elevated export of sucrose in *Arabidopsis* in response to long-day induction is suggested to be partially due to increased efficiency of sucrose loading (Corbesier *et al.* 1998). Following the loading of sucrose into the shoot apical meristem, a number of cellular and molecular events are initiated (Bernier 1988) as well as the hydrolysis of sucrose by local invertases i.e. vacuolar (Koch 2004) and cell wall (Heyer *et al.* 2004) invertases. Cytokinins activate invertase and increase the rate of cell division, while the ensuing hexoses participate with Giberellic acids in the up regulation of *LEAFY* (*LFY*) gene expression, which is a central regulator of floral meristem identity along with *APETALA1* that acts downstream of it (Bernier and Périlleux 2005, Karim *et al.* 2009).

Candidates genes defined within the LG7 QTL include *CLF* and *SQN* and are well characterized to be functionally conserved across plant and animal systems. The *Arabidopsis CLF* is an ortholog of the Drosophila *PRC2* component of *Enhancer of*

Zeste, a methyl transferase with specificity for H3K27 (Muller *et al.* 2002). Recent studies indicate a role in the repression of *FLC*, a repressor of flowering, by vernalization (Wood *et al.* 2006, Jiang *et al.* 2007). During the course of cold exposure, *FLC* expression is suppressed after several signature features of silenced chromatin accumulate at the *FLC* locus (Sheldon *et al.* 1999, Bastow *et al.* 2004, Sung and Amasino 2004, Mylne *et al.* 2006, Schubert *et al.* 2006, Finnegan and Dennis 2007, Schmitz *et al.* 2008). Changes in the amount and distribution of one of such signature, trimethylation on Lys-27 of H3 (H3K27me3), has been described during and after cold treatment (Schubert *et al.* 2006, Sung *et al.* 2006, Finnegan and Dennis 2007). H3K27me3 is carried out by *Polycomb-Group* (*PcG*) complexes that contain orthologs of Drosophila *Polycomb Repressive Complex 2* (*PRC2*), which include *CLF* and *VIN3* identified as candidates in this study. The regulation of *FLC* by this complex probably mirrors a possibly accumulation around the *AGL22/SVG* homologs in the *Prunus* (*DAM5* and *6*) in order to eliminate their repressive effect on flowering.

SQN, an ortholog of the Drosophila *CYCLOPHILIN 40* (*CYP40,* Galat 1993), is one of the first genes identified to be involved in vegetative phase change (transition from juvenile-adult-reproductive phase) via small RNAs that play a key regulatory role. This is crucial in life cycle of perennials plants that require several years of juvenile phase before they acquire competence to flower or in fruit trees that lose or have reduced competence to flower after a vigorous flowering and fruit season. Studies in *Arabidopsis* and maize show the constitutive expression of microRNA (miRNA) *miRNA156* prolongs the expression of the juvenile phase of vegetative development (Wu and Poethig. 2006,

Chuck *et al.* 2007, Gandikota *et al.*, 2007, Schwarz *et al.* 2008, Wang *et al.* 2008). Recent results suggest that *SQN* promotes *miRNA156* activity by promoting the activity of *ARGONAUTE* (Smith *et al.* 2009), which is responsible for miRNA-directed posttranscriptional silencing (Baumberger and Baulcombe 2005, Qi *et al.* 2005). This cascade acts on *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) transcription factors in *Arabidopsis* that converge on an overlapping set of targets downstream of an *FT* (*FLOWERING LOCUS T*)/bZIP transcription factor *FD* complex. The *FT* protein and the bZIP transcription factor *FD* play a central role in activating genes that execute the switch from vegetative to reproductive development, while the SPLs not only acts downstream of the *FT/FD* complex but also directly activate flower-promoting MADS-box genes by a separate endogenous flowering pathway (Wang *et al.* 2009).

Conclusion

Since the first observation by Muller-Thurgau in 1885, several studies have noted that early inception of bud dormancy leads to a shortened duration of dormancy i.e. reduces chilling requirement (Chandler and Tufts 1934). This implies that the major gene controlling blooming date must have been playing a role between bud set and the onset of dormancy and eventually determining the duration of dormancy and blooming date. Current evidence around *DAM2* implicates it as a major player. *DAM2* is a candidate for *AGL24*, which induces flowering and like *DAM2*, it's expressed after cold treatment

(Horvath *et al.* 2010). *AGL24* (also *AGL19*) are known to function in a vernalization pathway that is independent of *FLC* and mediated by *VIN3* (Alexandre and Hennig 2008), which acts in concert with *CLF* in the polycomb repressive complex to downregulate *DAM5* and *DAM6*. Although *DAM5* and *DAM6* are required to be repressed by vernalization before bud break, it appears they function only to suppress floral growth, while additional factors are required to promote flowering. *DAM2* seems a logical candidate for blooming date considering its expression before and after dormancy. Another genetic factor besides this pathway is the miRNA regulatory pathway induced by *SQN*, which probably serves as an independent floral induction pathway and an additional layer for the complex trait. These results presents functional homologs of genes with the vernalization, flowering and cell division regulation pathway models already characterized in model plants (e.g. *Arabidopsis thialana*), as well as mirroring a comparable gene pathway in *Prunus*.

Acknowledgement

This research was supported by the Robert and Lois Coker Trustees Chair in Molecular Genetics. We thank Dr. Marisa Badenes for providing us with the apricot accessions (Sayeb and Qardi) from Tunisia and Fan Shenghua for designing the evg SSR marker series.

References

Alexandre, C., Hennig, L. 2008. FLC or not FLC – the other side of vernalization. J. Exp. Bot. 59: 1127-35.

Aranzana, M.J., Pineda, A., Cosson, P., Dirlewanger, E., Ascasibar, J., Cipriani, G., Ryder, C.D., Testolin, R., Abbott, A., King, G.J., Iezzoni, A.F., Arús, P. 2003. A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theor. Appl. Genet. 106(5): 819-825.

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A., Dean, C. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. Nature 427: 164-167.

Baumberger, N., Baulcombe, D.C. 2005. *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc. Natl. Acad. Sci. U.S.A.* 102: 11928-11933.

Bernier, G. 1988. The control of floral evocation and morphogenesis. Annu. Rev. Plant. Physiol. Plant Mol Biol 39: 175-219.

Bernier, G., Périlleux, C. 2005. A physiological overview of the genetics of flowering time control. Plant Biotech. J 3: 3-16.

Bielenberg, D.G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G.L., Scorza, R., Abbott, A.G. 2008. Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. Tree Genet. Genomes 4: 495-507.

Breseghello, F., Sorrells, M. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172: 1165–1177.

Campbell, M.A., Suttle, J.C., Sell, T.W. 1996. Changes in cell cycle status and expression of p34 (cdc2) kinase during potato tuber meristem dormancy. *Physiol Plant* 98: 743-752.

Chandler, W.H., Tufts, W.P. 1934. Influence of the rest period on opening of buds of fruit trees in spring and on development of flower buds of peach trees. Proc. Amer. Soc. Hort. Sci. 30: 180-186.

Chuck, G., Cigan, A.M., Saeteurn, K., Hake, S. 2007. The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA. *Nat. Genet.* 39*:* 544-549.

Cleveland, W.S. 1979. "Robust Locally Weighted Regression and Smoothing Scatterplots". Journal of the American Statistical Association 74 (368): 829–836.

Combes, M.C., Andrzejewski, S., Anthony, F., Bertrand, B., Rovelli, P., Graziosi, G., Lashermes, P. 2000. Characterisation of microsatellite loci in *Coffea arabica* and related coffee species. Mol. Ecol. 9: 1178-1180.

Conrad, D. F., Jakobsson, M., Coop, G., Wen, X., Wall, J. D., Rosenberg, N. A., Pritchard, J. K. 2006. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat. Genet.* 38: 1251-1260.

Corbesier, L., Lejeune, P., Bernier, G. 1998. The role of carbohydrates in the induction of flowering in *Arabidopsis thaliana*: comparison between the wild type and a starchless mutant. Planta 206: 131-137.

Devitt, M.L., Stafstrom, J.P. 1995. Cell cycle regulation during growth-dormancy cycles in pea axillary buds. Plant Mol. Biol. 29: 255-265.

Falush, D., Stephens, M., Pritchard J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Falush, D., Stephens, M., Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes 7: 574-578.

Fan, S., Bielenberg, D.G., Zhebentyayeva, T.N., Reighard, G.L., Okie, W.R., Holland, D., Abbott, A.G. 2009. Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). New Phytol. 185: 917-930.

Faust, M., Surány, D., Nyujtó, F . 1998. Origin and dissemination of apricot, p. 225-266. In: J. Janick (ed.), Horticultural Reviews, vol. 22. John Wiley & Sons, Inc., New York, Chich-ester, Weinheim, Brisbane, Singapore, Toronto

Finnegan, E.J., Dennis, E.S. 2007. Vernalization-induced trimethylation of histone H3 lysine 27 at FLC is not maintained in mitotically quiescent cells. Curr. Biol. 17: 1978-1983.

Flint-Garcia, S.A., Thornsberry, J.M., Buckler, E.S. 4th. 2003. Structure of linkage disequilibrium in plants. Annu Rev Plant Biol. 54: 357-74.

Freeman, D., Riou-Khamlichi, C. Oakenfull, E.A., Murray, J.A. 2003. Isolation, characterization and expression of cyclin and cyclindependent kinase genes in Jerusalem artichoke *(Helianthustuberosus* L.). J. Exp. Bot. 54: 303-308.

Galat, A. 1993. Peptidylproline cis-trans-isomerases: Immunophilins. Eur. J. Biochem. 216: 689-707.

Gandikota, M., Birkenbihl, R.P., Höhmann, S., Cardon, G.H., Saedler, H., Huijser, P. 2007. The miRNA156/157 recognition element in the 3′ UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. Plant J. 49: 683-693.

Garris, A., Mccouch, S., Kresovich, S. 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (*Oryza sativa* L.). Genetics 165: 759-769.

Hartmann, U., Hohmann, S., Nettesheim, K., Wisman, E., Saedler, H., Huijser, P. 2000. Molecular cloning of *SVP*: a negative regulator of the floral transition in *Arabidopsis*. Plant J. 21: 351-360.

Heyer, A.G., Raap, M., Schroeer, B., Marty, B., Willmitzer, L. 2004. Cell wall invertase expression at the apical meristem alters floral, architectural, and reproductive traits in *Arabidopsis thaliana*. Plant J. 39: 161-169.

Hirschhorn, J.N., Daly, M.J. 2005. Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. 6(2): 95-108.

Horvath, D.P., Chao, W.S., Anderson, J.V. 2002. Molecular analysis of signals controlling dormancy and growth in underground adventitious buds of leafy spurge. Plant Physiol. 128: 1439-1446.

Horvath, D.P., Sung, S., Kim, D., Chao, W., Anderson, J. 2010. Characterization, expression and function of DORMANCY ASSOCIATED MADS-BOX genes from leafy spurge. Plant Mol Biol. doi:10.1007/s11103-009-9596-5.

Jiang, D., Yang, W., He, Y., Amasino, R.M. 2007. *Arabidopsis* relatives of the human lysine-specific demethylase1 repress the expression of FWA and FLOWERING LOCUS C and thus promote the floral transition, Plant Cell 19: 2975-2987.

Jiménez, S., Lawton-Rauh, A.L., Reighard, G.L., Abbott, A.G., Bielenberg, D.G. 2009. Phylogenetic analysis and molecular evolution of the dormancy associated MADS-box genes from peach. BMC Plant Biol. doi:10.1186/1471-2229-9-81.

Jiménez, S., Li, Z., Reighard, G.L., Bielenberg, D.G. 2010. Identification of genes associated with growth cessation and bud dormancy entrance using a dormancy-incapable tree mutant. BMC Plant Biol. doi:10.1186/1471-2229-10-25.

Karim, M.R., Hirota, A., Kwiatkowska, D., Tasaka, M., Aida, M. 2009. A role for Arabidopsis PUCHI in floral meristem identity and bract suppression. Plant Cell. 21(5): 1360-1372.

Kennedy, B.W., Quinton, M., Vanarendonk, J.A.M. 1992 .Estimation of effects of single genes on quantitative traits. J. Anim. Sci. 70: 2000-2012.

Koch, K. 2004. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. Curr. Opin. Plant Biol. 7:235-246.

Kostina, K.F .1946. The origin and evolution of cultivated apricot. Proceedings (Trudi) of the Nikita Botanical Garden 24: 25-31 (in Russian)

Krška, B., Vachun, Z., Necas.T. 2005. The apricot breeding program at the Horticulture faculty in Lednice. Acta Horticulturae 717:145–148

Kryukova I.V. 1989. Botanical classification and geographical distribution,p 9-23. In: VK Smykov (ed), Apricot, Agropromizdat, Moscow, USSR (in Russian)

Laurie, C.C., Nickerson, D.A., Anderson, A.D., Weir, B.S., Livingston, R.J., Dean, M.D., Smith, K.L., Schadt, E.E., Nachman, M.W. 2007. Linkage disequilibrium in wild mice. PLoS Genet. 8: e144.

Layne, R.E.C., Bailey, C.H., Hough, L.F. 1996. Apricots. In: J. Janick and J.N. Moore (eds.), *Fruit Breeding: Tree and Tropical Fruits.* Vol. I. John Wiley and Sons Inc., New York, pp. 79–111.

Li, Z., Reighard, G.L., Abbott, A.G., Bielenberg, D.G. 2009. Dormancy-associated MADS genes from the *EVG* locus of peach [*Prunus persica* (L.) Batsch] have distinct seasonal and photoperiodic expression patterns. J. Exp. Bot. 60: 3521-3530.

Luo, Z.W., Tao, S.H., Zeng, Z.B. 2000. Inferring linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. Genetics 151: 457-467.

Mehlenbacher, S.A., Cociu, V., Hough, L.F. 1990. Apricots (*Prunus*). In: Moore JN, Ballington JR (eds) Genetic resources of temperate fruit and nut crops. International Society for Horticultural Science, Wageningen, The Netherlands, pp 63-108.

Meijón, M., Feito, I., Valledor, L., Rodríguez, R., Cañal, M.J. 2010. Dynamics of DNA methylation and Histone H4 acetylation during floral bud differentiation in azalea. BMC Plant Biol. doi:10.1186/1471-2229-10-10.

Michaels, S.D., Ditta, G., Gustafson-Brown, C., Pelaz, S., Yanofsky, M., Amasino, R.M. 2003. AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. Plant J. 33(5): 867-74.

Michaels, S.D., Ditta,, G., Gustafson-Brown., C., Pelaz, S., Yanofsky, M., Amasino, R.M. 2003. AGL24 acts as a promoter of flowering in *Arabidopsis* and is positively regulated by vernalization. Plant J. 33: 867-874.

Micheletti, D., Troggio, M., Baldi, P., Costa, F., Pindo, M., Komjanc, M., Malnoy, M., Zarkikh, A., Magnago, P., Velasco, R., Salvi S. 2010. LD estimation and analysis of diversity in apple. In: Plant & Animal Genome 29th Conference, San Diego, California, U.S.A.

Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., Simon, J.A. 2002. Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. Cell 111: 197-208.

Muller-Thurgau, H. 1885. Beitrag zur Erklärung der Ruheperioden der Pflanzen. Landw. Jahrb. 14: 851-907.

Mylne, J.S., Barrett, L., Tessadori, F., Mesnage, S., Johnson, L., Bernatavichute, Y.V., Jacobsen, S.E., Fransz, P., Dean, C. 2006. LHP1, the *Arabidopsis* homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc. Natl. Acad. Sci. U.S.A. 103: 5012-5017.

Nordborg, M., Borevitz, J.O., Bergelson, J., Berry, C.C., Chory, J., Hagenblad, J., Kreitman, M., Maloof, J.N., Noyes, T., Oefner, P.J., Stahl, E.A., Weigel, D. 2002. The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat. Genet. 30: 190-193.

Nordborg, M., Donnelly, P. 1997. The coalescent process with selfing. Genetics 146: 1185-1195.

Numa, H., Kim, J.M., Matsui, A., Kurihara, Y., Morosawa, T., Ishida, J., Mochizuki, Y., Kimura, H., Shinozaki, K., Toyoda, T., Seki, M., Yoshikawa, M., Habu, Y. 2010. Transduction of RNA-directed DNA methylation signals to repressive histone marks in *Arabidopsis thaliana*. EMBO J. 29(2): 352-62.

Olukolu, B.A., Trainin, T., Fan, S., Kole, C., Bielenberg, D.G., Reighard, G.L., Abbott, A.G., Holland, D. 2009. Genetic linkage mapping for molecular dissection of

chilling requirement and budbreak in apricot (*Prunus armeniaca* L.). Genome 52(10): 819-28.

Palaisa, K.A., Morgante, M., Williams, M., Rafalski, A. 2003. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15: 1795-1806.

Pedryc, A., Ruthnera, S., Hermána, R., Krska, B., Hegedűsa, A., Halásza, J. 2009. Genetic diversity of apricot revealed by a set of SSR markers from linkage group G1. Sci. Hortic. 121(1): 19-26.

Pritchard, J.K., Przeworski, M. 2001. Linkage disequilibrium in humans: models and data. Am. J. Human Genet. 69: 1–14.

Pritchard, J.K., Stephens M., Donnelly P. 2000a. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Pritchard, J.K., Stephens, M., Rosenberg, N.A., Donnelly P. 2000b. Association mapping in structured populations. Am. J. Hum. Genet. 67: 170–181.

Pritchard, J.K., Wen, X., Falush D. 2009. Documentation for Structure Software, Version 3.2 Department of Human Genetics, University of Chicago, Chicago.

Qi, Y., Denli, A.M., Hannon, G.J. 2005. Biochemical specialization within *Arabidopsis* RNA silencing pathways. Mol. Cell 19: 421-428.

Remington, D.L., Thornsberry, J.M., Matsuola, Y., Wilson, L.M., Whitt, S.R., Doebley, J., Kresovich, S., Goodman, M.M., Buckler, E.S. 4th. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. U.S.A. 98: 11479-11484.

Risch, N.J. 2000. Searching for genetic determination for the new millennium. Nature 405: 847-856.

Ryabov I.N. 1969. Sortoizuchenie i pervichnoe sortoispyitanie kostochkovyih plodovyih kultur v Gosudarstvennom Nikitskom botanicheskom sadu"/Sortoizuchenie kostochkovyih plodovyih kultur na yuge SSSR. Proceedings (Trudy) of State Nikita Botanical Garden 41: 3-83

SAS Institute Inc., SAS 9.1.3 Help and Documentation, Cary, NC: SAS Institute Inc., 2000-2004.

Schmitz, R.J., Sung, S., Amasino, R.M. 2008. Histone arginine methylation is required for vernalization-induced epigenetic silencing of FLC in winter-annual *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U.S.A. 105: 411-416.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., Goodrich, J. 2006. Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. EMBO J. 25: 4638-4649.

Schwarz, S., Grande, A.V., Bujdoso, N., Saedler, H., Huijser, P. 2008. The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in *Arabidopsis*. Plant Mol. Biol. 67: 183-195.

Sharp, D.J., Rogers, G.C., Scholey, J.M. 2000. Microtubule motors in mitosis. Nature 407: 41-47.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J., Dennis, E.S. 1999. The FLF MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. Plant Cell 11: 445-458

Sivolap, YuM (ed). 1998. Application of PCR analysis to genetic research and breeding (in Russian). Agrarna nauka, Kiev.

Smith, M.R., Willmann, M.R., Wu, G., Berardini, T.Z., Möller, B., Weijers, D., Poethig, R.S. 2009. Cyclophilin 40 is required for microRNA activity in *Arabidopsis*. Proc. Natl. Acad. Sci. U.S.A. 106(13): 5424-5429. doi:10.1073/pnas.0812729106.

Sung, S., Amasino, R.M. 2004. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. Nature 427: 159-164.

Sung, S., He, Y., Eshoo, T.W., Tamada, Y., Johnson, L., Nakahigashi, K., Goto, K., Jacobsen, S.E., Amasino, R.M. 2006. Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1. Nat. Genet. 38: 706-710.

Thornsberry, J.M., Goodman, M.M., Doebley, J., Kresovich, S., Nielsen, D., Buckler E.S. 4th. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. Nat. Genet. 28: 286-289.

Vavilov, N.I. 1951. Phytogeographical basis of plant breeding. In: Chester KS (trans) The origin, variation, immunity and breeding of cultivated plants. Chron. Bot. 13: 13-54.

Vos, P., Hogers R., Bleeker, M., Reijans, M., Van der Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. 1995. AFLP: a new technique of DNA fingerprinting. Nucleic Acids Res. 23: 4407-4414

Wang J.W., Czech B., Weigel D. 2009. miR156-Regulated SPL Transcription Factors Define an Endogenous Flowering Pathway in *Arabidopsis thaliana*. Cell 138(4): 738-749.

Wang, J.W., Schwab, R., Czech, B., Mica, E., Weigel, D. 2008. Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in *Arabidopsis thaliana*. Plant Cell 20: 1231-1243.

Wittmann, T., Hyman, A., Desai, A. 2001. The spindle: A dynamic assembly of microtubules and motors. Nat. Cell Biol. 3: E28-E34.

Wood, C.C., Robertson, M., Tanner, G., Peacock, W.J., Dennis, E.S., Helliwell, C.A. 2006. The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proc. Natl. Acad. Sci. U. S. A. 103: 14631-14636.

Wu, G., Poethig, R.S. 2006. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development* 133*:* 3539-3547.

Yu, H., Ito, T., Wellmer, F., Meyerowitz, E.M. 2004. Repression of AGAMOUS-LIKE 24 is a crucial step in promoting flower development. Nat Genet. 36: 157-161. doi:10.1038/ng1286.

Zhebentyayeva, T.N., Reighard, G.L., Gorina, V.M., Abbott, A.G. 2003. Simple sequence repeat (SSR) analysis for assessment of genetic variability in apricot germplasm, Theor. Appl. Genet. 106: 435-444.

Zhebentyayeva, T.N., Reighard, G.L., Lalli, D., Gorina, V.M., Krška, B., Abbott, A.G. 2008. Origin of resistance to plum pox virus in apricot: what new AFLP and targeted SSR data analyses tell. Tree Genet. Genomes 4: 403-417.

Zhebentyayeva, T.N., Ledbetter, C., Burgos, L., Llácer, G. 2010. Apricots. In: Handbook of Plant Breeding, Vol. 6, Fruit Breeding. Badenes M.L and Byrne D.H. (Eds.). Springer NY, 450 pp (in press)

Zhonga, X., Liub, X., Qia, B., Liu, B. 2010. Characterization of copia retrotransposons in *Zizania latifolia* shows atypical cytosine methylation patterns and differential occurrence from other species of the grass family. Aquat. Bot. 90(3): 213-221.

CHAPTER FIVE

CONCLUSION

Molecular biology is advancing and changing rapidly and with the aid of new high-throughput technologies. Holistic studies on an –omics platform can be conducted on large scale and at a fast pace for the generation of huge amounts of data. This has changed the playing field and made approaches that were not feasible decades ago possible. Gene-by-gene expression analysis and functional genomics approaches identifying gene interactions have opened up new frontiers in understanding regulation of bud dormancy. Ongoing research is shedding more light on differentiating sensing and signaling genes from those that maybe be regulatory or target genes. Genetic mapping studies are providing information about key genes containing polymorphisms that are crucial for regulating bud dormancy. These approaches in concert will help to define dormancy gene cause and effect relationships since by itself, high throughput transcriptome analyses of differential gene expression can implicate genes as part of a network; however, their candidacy for driving phenotype must be examined through genotype-specific responses in populations that segregate for the character under study. This will lead to remarkable practical value for breeding programs and judicious utilization of existing germplasm. A reasonable next step beyond identification of QTLs controlling bud dormancy will be the use of tightly linked markers for marker-assisted breeding and map-based cloning of causative/candidate genes from large insert libraries using chromosome walking approaches in unsequenced genomes or mining the sequences

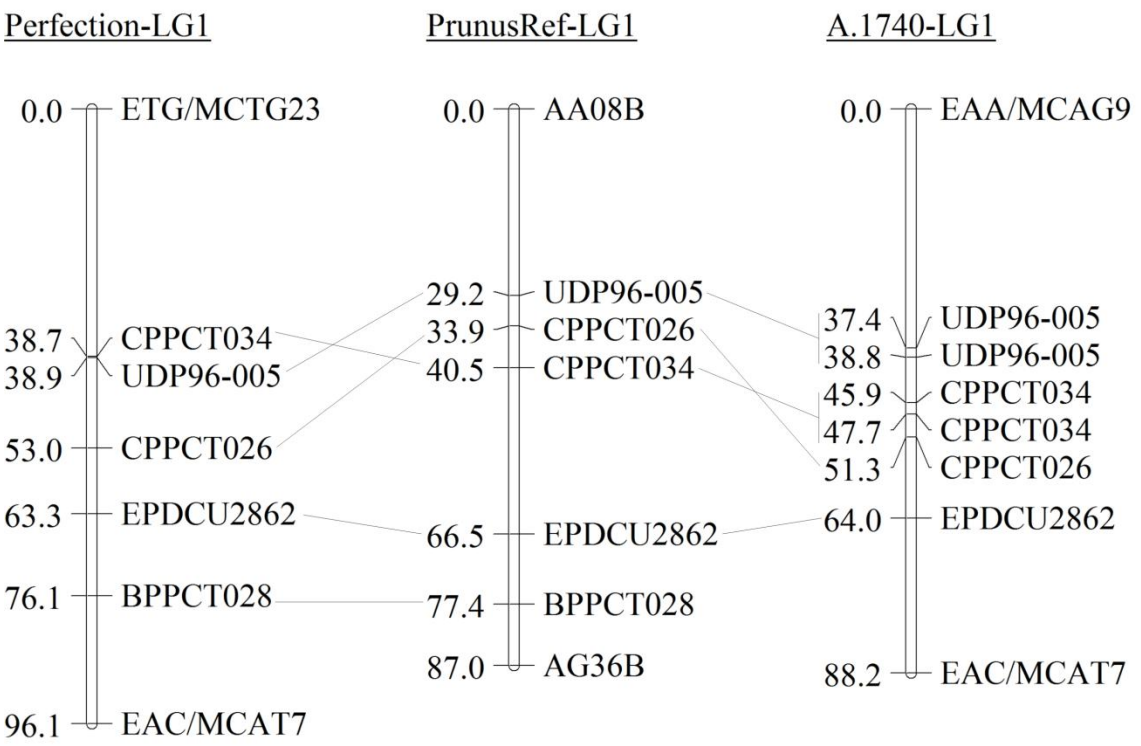
in genomes that are already sequenced. This will provide a targeted approach for identifying and characterizing genes and their causative polymorphisms as well as their roles in signaling pathways. A detailed understanding of the molecular components and gene networks will be indispensable for targeted manipulations of bud dormancy in crops of economic importance.

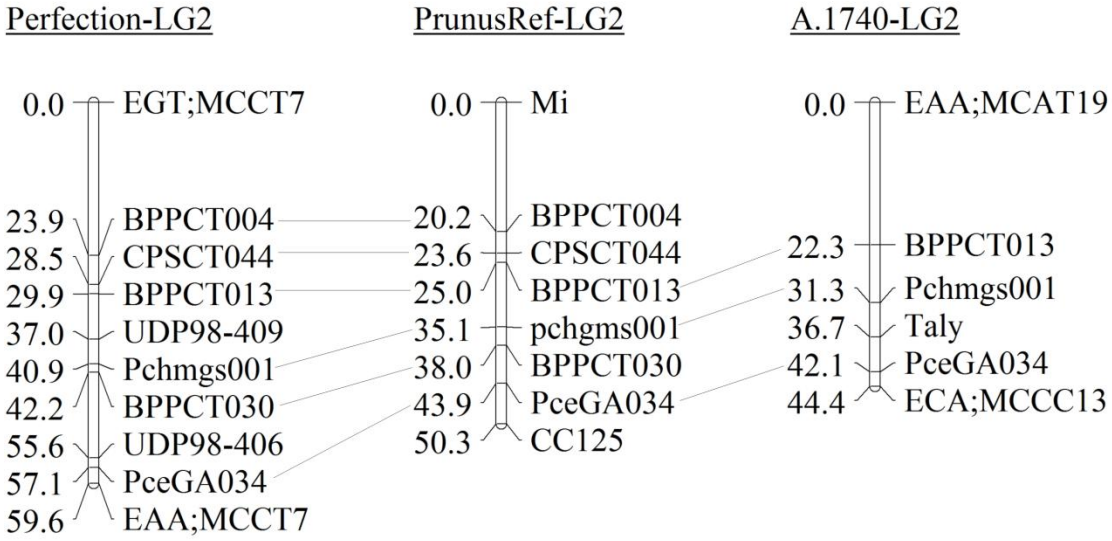
We report a concerted effort at elucidating the genetic and molecular mechanism that underlie chilling requirement (CR) and bud break. With the aid of studies comprising genetic linkage and QTL analysis; linkage disequilibrium-based association mapping and comparative genomics, we were able to define genomic regions and candidate genes controlling these dormancy-related traits. Following a functional genomics approach, these results in concert with other published gene expression and functional studies in woody perennials and other model organism, provide us with resources to define these candidate genes as functional homologs as well as sequence homologs of annotated and characterized genes.

APPENDICES

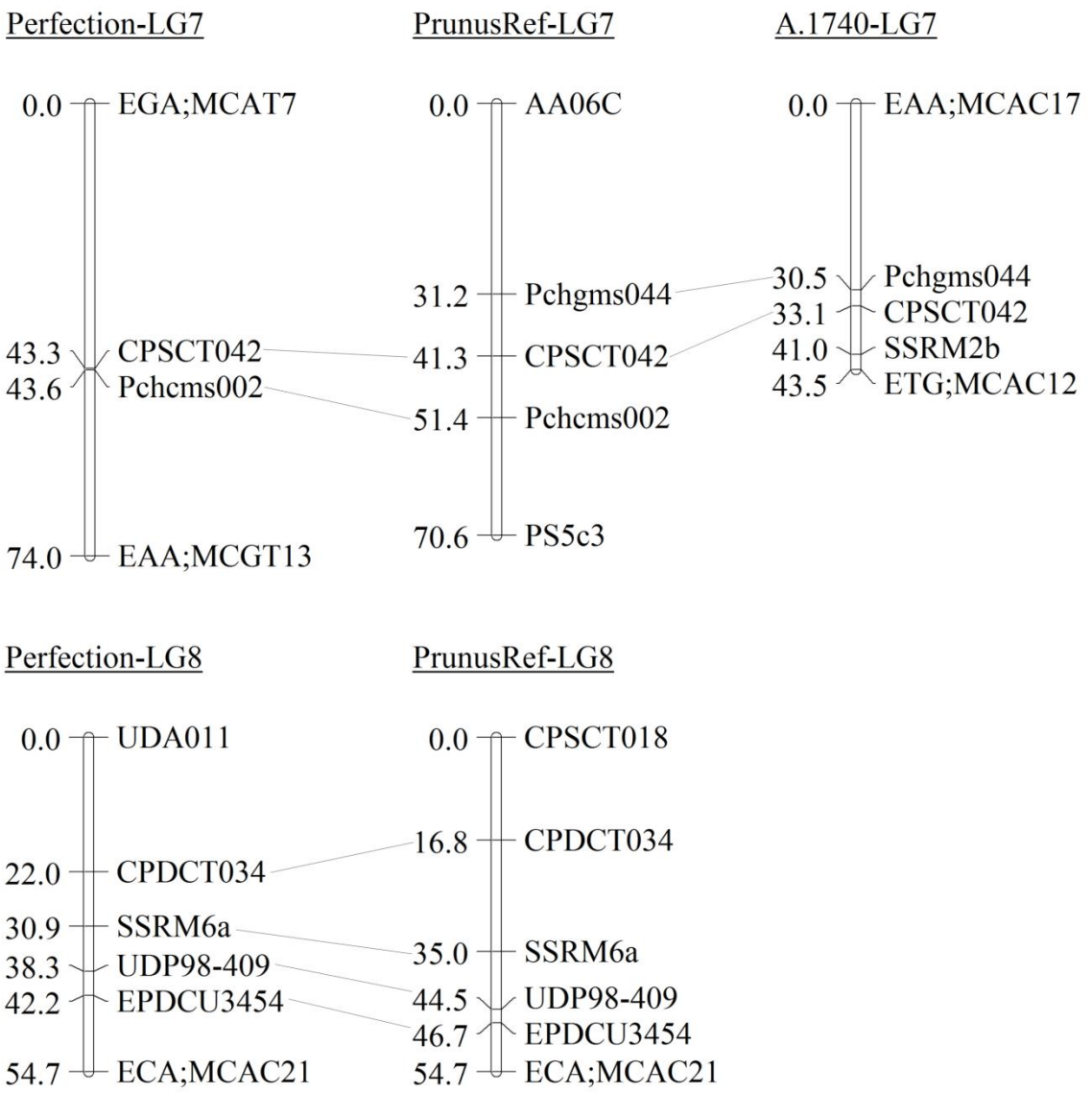
Appendix A: AFLP E (*EcoRI*) and M (*MseI*) primer combinations (PC) and the number of segregating loci.

Primer Combinations	segregating loci/PC	# of PCs	Total # of loci
ACCGC, AGCGT, CACGC, CTCGA, GCCCT, GGCGC, GTCTC, TGCCG, TTCTC	2	9	18
AGCCT, CCCGC, CCCGG, CCCGT, CGCAC, CGCAT, CGCCA, CGCCC, CGCCG, CGCTA, CTCAA, CTCGC, CTCTT, GACGC, GACGT, GCCAA, GCCCA, GCCCC, GCCCG, GCCGA, GCCGC, GCCGG, GCCGT, GCCTA, GCCTC, GCCTG, GCCTT, GGCAG, GGCGG, GGCGT, GGCTG, GTCTG, GTCTT, TCCCG, TGCCC, TTCAA, TTCAC, TTCAG, TTCAT, TTCCA, TTCGC, TTCGG, TTCGT, TTCTG	3	44	132
ACCCC, ACCCT, ACCGA, ATCCA, ATCGT, CACGT, CCCAA, CCCAC, CCCCA, CCCGA, CCCTG, CGCAG, CGCGC, CGCGT, CGCTT, CTCAC, CTCGG, CTCGT, GACGA, GCCAC, GCCAT, GGCCC, GGCCG, GGCGA, GGCTT, GTCAC, GTCGG, TCCCT, TGCCT, TGCGT	4	30	120
AACTG, AGCAA, AGCGG, AGCTC, ATCGC, CCCAT, CCCCC, CCCTC, CGCCT, CGCTC, CTCAT, CTCCG, CTCTC, CTCTG, GACCG, GGCAA, GGCCA, GGCTC, GTCCG, GTCGA, GTCGC, TACAA, TACCA, TACGC, TACGG, TCCCA, TCCCC, TCCGA, TCCGG, TGCAA, TGCAT, TGCGC, TGCTA, TTCCC, TTCGA, TTCTT, TCCTT	5	37	185
ATCCG, ATCCT, CACTT, CCCCG, CCCTA, CCCTT, CGCAA, CGCTG, GACGG, GACTG, GGCCT, GGCTA, TACGT, TCCGC, TGCCA	6	15	90
AACGG, ACCAC, AGCAG, AGCGC, AGCTA, ATCCC, ATCTG, CACCT, CACGG, CACTG, CGCGA, CTCAG, CTCCA, CTCTA, GACAA, GACAC, GCCAG, GTCAT, TACCG, TACTC, TGCAG	7	21	147
ACCCA, AGCAC, AGCCG, AGCTT, ATCAT, ATCGA, ATCGG, CACCG, CTCCC, GACCA, GGCAT, TACAG, TACAT, TACCC, TACGA, TCCAC, TCCGT, TGCGA, TTCCG, TTCCT, TTCTA	8	21	168
AACGC, ACCCG, AGCCC, AGCTG, ATCAG, CCCCT, CGCGG, GACTT, GGCAC, TACAC, TACTA, TACTT, TCCTC, TGCGG, AACAA, GTCTA	9	16	144
GTCGT, GTCCC	10	2	20
TCCTA, ACCGT, GACCC, GACCT, TACCT	11	5	55
ACCTG	12	1	12
ACCAG, CTCCT, GTCAG	13	3	39
ACCGG, ACCTT, TCCTG	14	3	42
AACAG, ATCTA, CACGA, CACTC, CCCAG, GTCCT	16	6	96
GACAT, GTCCA	17	2	34
AGCAT, AGCCA, AGCGA, CACCC, TCCAG	18	5	90
ACCTA, ATCTT, TACTG	19	3	57
AACCA, AACCG, GACAG	20	3	60
AACGA, ACCAT, ATCAA, CACAC, CACCA	21	5	105
AACAT, AACCT, ATCAC, CACAG, GACTC, TGCAC	22	6	132
ACCTC, CACAT, GTCAA, TGCTC	23	4	92
ACCAA, TCCAT	24	2	48
AACTT, CACAA	26	2	52
TGCTT	27	1	27
AACAC, AACCC, AACGT, AACTC, CACTA	28	5	140
GACTA, TCCAA	29	2	58
AACTA, ATCTC, TGCTG	30	3	90
Total	424	256	2253

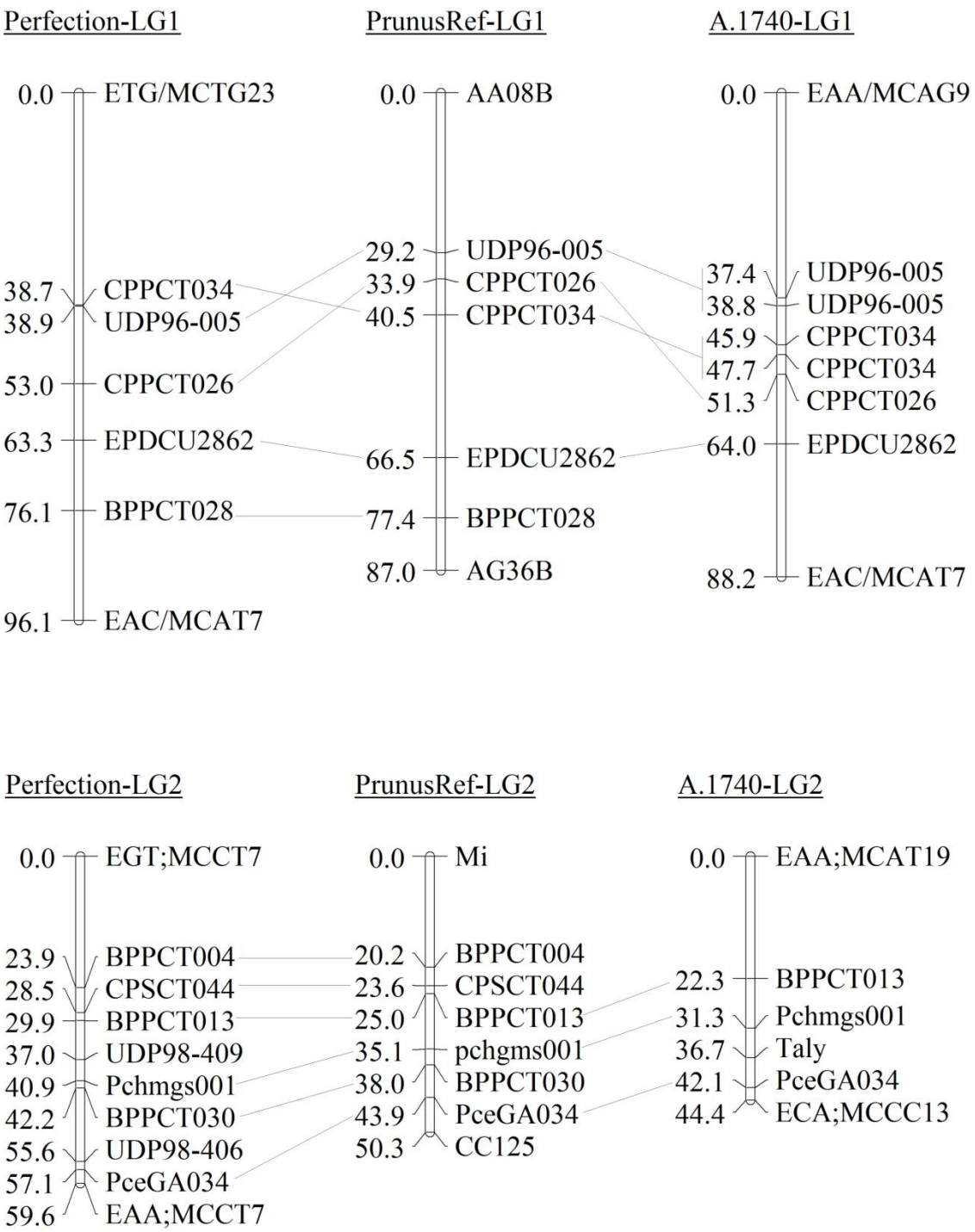




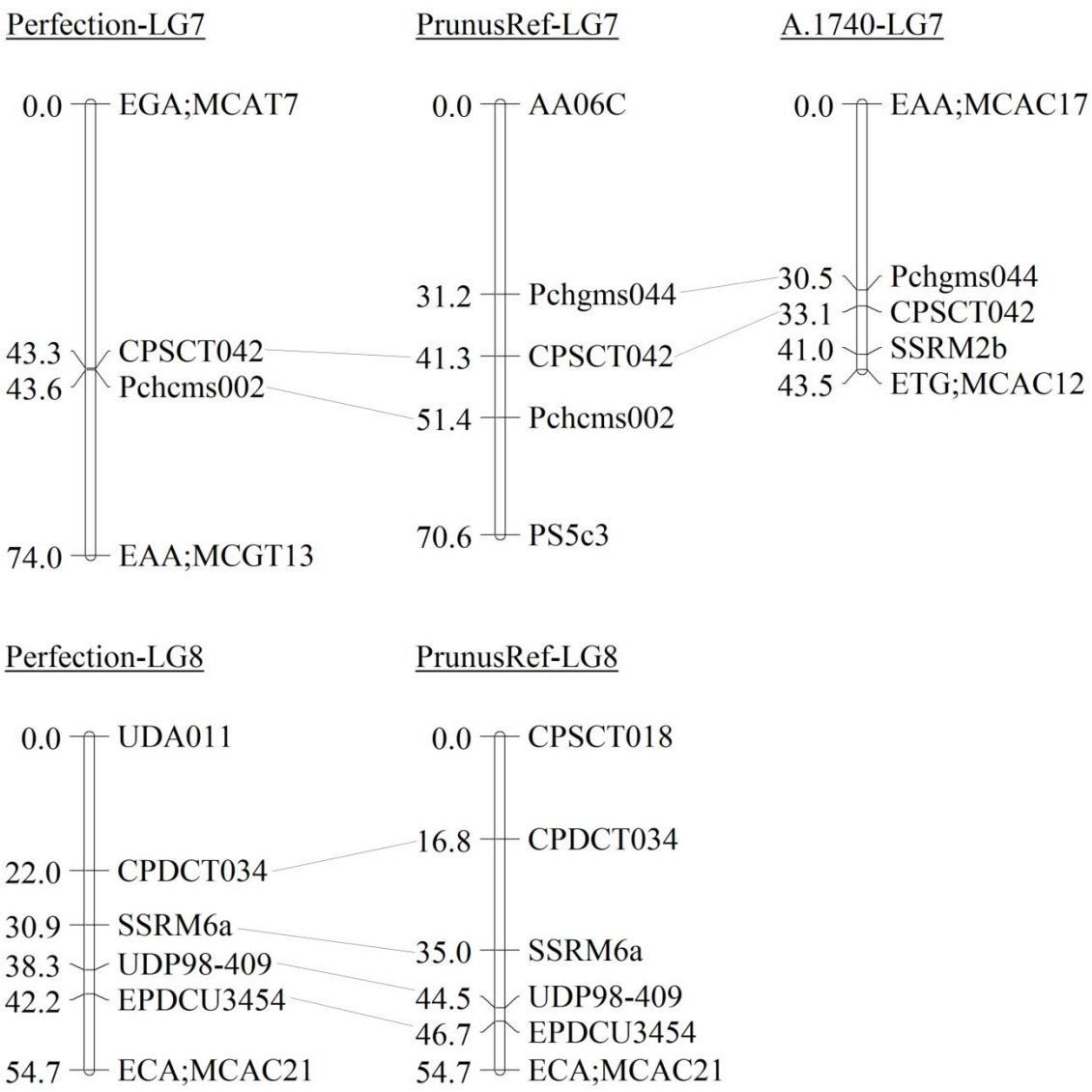
Appendix B1: Comparative analysis of common SSR markers on linkage groups 1 and 2 among the apricot parental (Perfection and A.1740) maps and *Prunus* ref map.



Appendix B2: Comparative analysis of common SSR markers on linkage groups 3 and 4 among the apricot parental (Perfection and A.1740) maps and *Prunus* ref map.



Appendix B3: Comparative analysis of common SSR markers on linkage groups 5 and 6 among the apricot parental (Perfection and A.1740) maps and *Prunus* ref map.



Appendix B4: Comparative analysis of common SSR markers on linkage groups 7 and 8 among the apricot parental (Perfection and A.1740) maps and *Prunus* ref map.

Appendix C: Assignment of apricot accessions to subpopulations, their ancestry and proportions of contributions from each subpopulation.

sno	Accessions	Subpopulati	Geographical origin	1	2	3	4	5	6	7
1	Alberge de Tur	1	European	0.52	0.00	0	0.00	0.00	0	0.46
2	Ananasnyi	1	European	0.52	0.00	0	0.00	0.00	0	0.47
6	B-5-3	1	Central Asian	0.32	0.19	0.26	0.06	0.00	0	0.14
21	Early Gold	1	European	0.69	0.00	0.00	0.00	0.00	0.00	0.29
22	Goldrich	1	PPV resistant cultivars	0.83	0.08	0.01	0.00	0.03	0.01	0.00
24	Harlayne	1	PPV resistant cultivars	0.71	0.12	0.00	0.00	0.05	0.05	0.04
67	Ordubad	1	Iran-caucassian	0.45	0.09	0.03	0.00	0.41	0	0.01
76	Precoce d'Italia	1	European	0.99	0.00	0.00	0.00	0.00	0.00	0.00
80	Real d'Imola	1	European	0.99	0.00	0.00	0.00	0.00	0.00	0.00
85	Sayed (Tunisia)	1	Northern African	0.89	0.00	0.03	0.02	0.01	0.00	0.02
93	Tilton	1	European	0.77	0.00	0.00	0.10	0.00	0	0.11
9	Da-bei	2	Chinese	0.00	0.99	0.00	0.00	0.00	0.00	0.00
25	In-ben-sin	2	Chinese	0.00	0.94	0.00	0.00	0.04	0.00	0.00
36	Kitaiskii	2	Chinese	0.00	0.99	0.00	0	0	0	0
47	Lao-yech-lian	2	Chinese	0.00	0.99	0.00	0	0.00	0	0.00
48	LE2904	2	PPV resistant cultivars	0.28	0.38	0.00	0.00	0.00	0.00	0.32
49	LE3276	2	PPV resistant cultivars	0.01	0.79	0.00	0.02	0.00	0.00	0.15
54	Mai-ch10-sin	2	Chinese	0.00	0.83	0.00	0.05	0.01	0.08	0.00
58	Mi-bada	2	Chinese	0.00	0.99	0.00	0.00	0.00	0.00	0.00
68	P. ansu	2	Non-domesticated sp.	0.00	0.29	0.06	0.25	0.00	0.27	0.10
69	P. manshurica	2	Non-domesticated sp.	0.00	0.37	0.00	0.33	0.00	0.24	0.02
72	P. sibirica var	2	Non-domesticated sp.	0.00	0.52	0.01	0.43	0.00	0.01	0.00
78	Pui-sha-sin	2	Chinese	0.00	0.74	0.15	0.00	0.00	0.09	0.00
84	Satser	2	Hybrid cultivars	0.23	0.31	0.00	0.28	0.00	0.00	0.15
86	SEO	2	PPV resistant cultivars	0.31	0.65	0.00	0.00	0.00	0	0.01
88	Shantunski	2	Chinese	0.06	0.73	0.00	0.14	0.04	0.00	0.01
97	Vestar	2	PPV resistant cultivars	0.23	0.49	0.00	0.00	0.00	0	0.26
99	Yuan-sin	2	Chinese	0.01	0.57	0.02	0.28	0.00	0.06	0.03
5	B-4-5	3	Central Asian	0.01	0.00	0.79	0.00	0.13	0.04	0.00
7	Badami	3	Central Asian	0.00	0.00	0.83	0.13	0.00	0.01	0.00
12	Daradzhi ek Shabistr	3	Iran-caucassian	0.27	0.00	0.31	0.12	0.01	0.07	0.20
17	Dzhungarskii 18/64	3	P. armerniaca wild	0.00	0.07	0.59	0.32	0.00	0.00	0.00
19	Dzhungarskii 18/75	3	P. armerniaca wild	0.01	0.16	0.57	0.19	0.00	0.05	0.00
23	Gulyunghi	3	Central Asian	0.00	0.01	0.85	0.11	0.00	0.00	0.00
26	Iskadari	3	Central Asian	0.00	0.00	0.63	0.36	0.00	0.00	0.00
28	Kandak-10	3	Central Asian	0.00	0.00	0.98	0.00	0.00	0.00	0.00
29	Kandak-12	3	Central Asian	0.00	0.07	0.80	0.05	0.00	0.00	0.05
32	KB-12	3	Central Asian	0.09	0.04	0.78	0.01	0.00	0.04	0.00
33	KB-9	3	Central Asian	0.14	0.28	0.28	0.02	0.09	0.00	0.16
34	Khurmai	3	Central Asian	0.00	0.00	0.98	0.00	0.00	0.00	0.00
37	KK(2) N1	3	Central Asian	0.00	0.17	0.72	0.00	0.00	0.08	0.00
39	Kolon Boboi	3	Central Asian	0.00	0.00	0.86	0.12	0.01	0.00	0.00
41	Krimskii Amur	3	Hybrid cultivars	0.25	0.00	0.31	0.01	0.25	0.00	0.15
44	Kzyl Khorezmskii	3	Central Asian	0.00	0.00	0.62	0.22	0.05	0.00	0.07
45	Kzyl Khurmai	3	Central Asian	0.43	0.00	0.54	0.00	0.00	0.00	0.02
46	Kzyl Uryuk	3	Central Asian	0.10	0.00	0.45	0.01	0.01	0.00	0.40
52	Lyuchak Sumbarskii	3	Central Asian	0.07	0.01	0.60	0.11	0.14	0.04	0.00
59	Mirsandzhali	3	Central Asian	0.00	0.00	0.98	0.00	0.00	0.00	0.00
64	Nukul Citronnyi	3	Central Asian	0.07	0.00	0.91	0.00	0.00	0.00	0.00
74	Paivandi Bukharskii	3	Central Asian	0.18	0.07	0.39	0.14	0.07	0.12	0.00
90	Shekarpara de Semnan	3	Iran-caucassian	0.00	0.00	0.99	0.00	0.00	0.00	0.00

Appendix C: continued

sno	Accessions	Subpopulati	Geographical origin	1	2	3	4	5	6	7
91	Supkhani	3	Central Asian	0.00	0.00	0.59	0.00	0.05	0.00	0.33
92	Tadzhabai	3	Central Asian	0.00	0.00	0.63	0.29	0.01	0.05	0.00
3	Arzami	4	Central Asian	0.00	0.00	0.43	0.56	0.00	0.00	0.00
4	B-1-11	4	Central Asian	0.17	0.09	0.11	0.26	0.13	0.07	0.15
10	Da-chuan-che N1	4	Chinese	0.03	0.10	0.00	0.46	0.13	0.00	0.26
11	Da-chuan-che N2	4	Chinese	0.00	0.16	0.37	0.45	0.00	0.00	0.00
14	Dionis	4	Hybrid cultivars	0.33	0.00	0.00	0.37	0.23	0.00	0.04
15	Dzhungarskii 18/55	4	P. armerniaca wild	0.00	0.00	0.42	0.55	0.00	0.01	0.00
16	Dzhungarskii 18/63	4	P. armerniaca wild	0.18	0.00	0.23	0.25	0.18	0.06	0.07
18	Dzhungarskii 18/68	4	P. armerniaca wild	0.00	0.00	0.44	0.51	0.04	0.00	0.00
20	Dzhungarskii 18/78	4	P. armerniaca wild	0.12	0.24	0.21	0.26	0.00	0.00	0.15
35	Khurmai Rannii	4	Central Asian	0.21	0.10	0.22	0.29	0.00	0.00	0.16
38	Kok-pshar	4	Central Asian	0.00	0.00	0.12	0.84	0.00	0.02	0.00
42	Kunduzi	4	Central Asian	0.12	0.10	0.00	0.74	0.00	0.00	0.02
43	Kurbane Marache	4	Iran-caucassian	0.32	0.16	0.04	0.34	0.00	0.00	0.12
55	Mamuri	4	Central Asian	0.04	0.1	0.13	0.52	0.02	0.04	0.12
57	Medunets Krimskii	4	Hybrid cultivars	0.13	0.00	0.00	0.44	0.41	0.00	0.00
61	Naryadnyi	4	Hybrid cultivars	0.26	0.13	0.00	0.29	0.27	0.00	0.01
66	Oranzhevo-krasnyi	4	Central Asian	0.00	0.01	0.40	0.55	0.01	0.00	0.00
73	P.sibirica	4	Non-domesticated sp.	0.00	0.32	0.01	0.36	0.00	0.27	0.00
77	Pruisadebnyi Rannii	4	Hybrid cultivars	0.00	0.00	0.00	0.61	0.00	0.00	0.37
79	Quardi (Tunisia)	4	Northern African	0.18	0.02	0.12	0.41	0.02	0.00	0.23
81	Rukhi Dzhuvanon	4	Central Asian	0.02	0.12	0.37	0.47	0.00	0.00	0.00
82	Rukhi Dzhuvanon	4	Central Asian	0.00	0.00	0.25	0.61	0.12	0.00	0.00
83	Samyi Rannii	4	Central Asian	0.00	0.00	0.00	0.62	0.00	0.00	0.36
100	Zard	4	Central Asian	0.00	0.00	0.17	0.81	0.00	0.00	0.00
51	Lunnik	5	Hybrid cultivars	0.00	0	0	0	0.99	0	0.00
53	Maftobi	5	Central Asian	0.10	0.00	0.00	0.68	0.10	0.08	0.00
56	Mascat	5	Iran-caucassian	0.24	0.00	0.24	0.00	0.28	0.00	0.21
63	Naslazhdenije	5	Hybrid cultivars	0.00	0.00	0.00	0.00	0.97	0	0.01
65	Olimp	5	Hybrid cultivars	0.00	0.00	0.00	0.00	0.75	0	0.23
75	Parnas	5	Hybrid cultivars	0.00	0.06	0.00	0.00	0.60	0.01	0.30
87	Shalakh	5	Iran-caucassian	0.37	0.00	0.00	0.00	0.50	0.00	0.11
89	Shedevr	5	Hybrid cultivars	0.00	0	0	0	0.99	0	0.00
94	Vaagas Vardaguin	5	Iran-caucassian	0.36	0.02	0.00	0.00	0.40	0.03	0.16
98	Vynoslivyi	5	European	0.00	0.00	0.00	0.00	0.60	0.02	0.36
70	P. mume, N 15	6	Non-domesticated sp.	0	0	0	0	0	0.99	0
71	P. mume, N18	6	Non-domesticated sp.	0	0	0	0	0	0.99	0
8	Bergeron	7	European	0.00	0	0	0	0	0	0.99
13	De Compot	7	European	0.13	0.00	0.19	0.27	0.00	0.00	0.38
27	Jubileinyi	7	European	0	0	0	0	0	0	0.99
30	Kantsler	7	European	0.04	0	0	0	0.00	0	0.95
31	Katuni	7	Iran-caucassian	0.05	0.00	0.03	0.00	0.30	0.10	0.49
40	Krasnoshchekii	7	European	0.00	0.00	0	0	0.00	0	0.99
50	Luizet Krupnoplodnyi	7	European	0.00	0	0	0	0	0	0.99
60	Nakhichevanskii	7	Iran-caucassian	0.00	0	0	0	0	0	0.99
62	Nasera Tabris	7	Iran-caucassian	0.05	0.23	0.01	0.20	0.15	0.00	0.34
95	Velkopavlovichka	7	European	0.00	0	0	0	0	0	0.99
96	Vengerskii Krypnyi	7	European	0.00	0	0	0	0	0	0.99

Appendix D: List of primer sequence for SSR markers used for linkage disequilibrium-based association mapping on LG1 and 7.

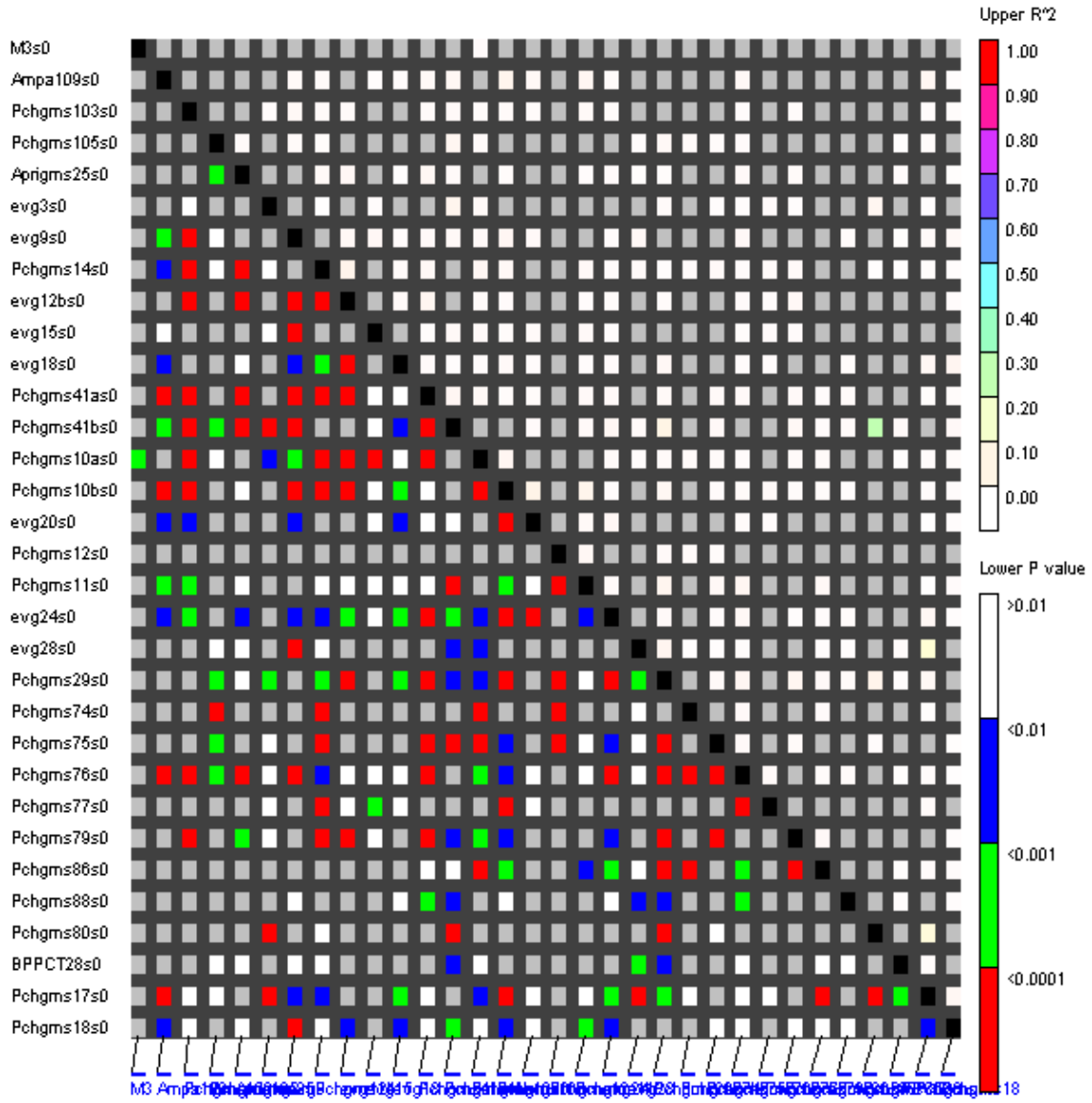
Marker	Forward primer sequence	Reverse primer sequence
M3	F-CGAGAAACTCTGCACAGAGA	R-GTAGCCGATTCAAAGCCTCC
Ampa109	F-GTGTCCCGAATTCCAATATCC	R- TTTGTCTCAACACTTTCCCTCTC
Pchgms103	F-GCCGATGACTACGATTGG	R-ATGTGTTTCGGAGATGTCG
Aprigms25	F-ACACACACGGCTCTTCGT	R-CCTTGTCGCCAGTGAATAATTTGACA
Pchgms105	F-CCGTAGTTGAGAGTAAGTGAAAGC	R-GGGTTATCTCTCTGCTCATCC
evg3	F- TCATTGACGACCCATTTGAA	R- ATTTGGGCGATCATATCAGG
evg9	F- AGAGGGAGAGAAGGGTGGAG	R- GGACGGGGACAAGGATATTT
Pchgms14	F-GCAAAGAGTACAACAATATCTACCG	R-GGATGGTGAAGACGATGAGG
evg12b	F- GGAGGAGAAGGGTGCCTTTA	R- GGACACCCACCTAGACGTTG
evg15	F-GGTGTCCCCAGTTTGAGAAA	R- GAAGGACACCTTTTCGGACA
evg18	F-GCCATAAAGTCCACCAAGGA	R- CAGCCCCGTAAAGAAACAAA
Pchgms41	F-TCAAGCTCAAGGTACCAGCA	R-AAGGCACTCTCCCTCTCCTC
Pchgms10	F-GGTCACGCATCCTTTCATTT	R-GACACCTCCATTTGTATCAAAGC
evg20	F- GTTGCTTGTTGGGTTGGTCT	R- GCTTTATCGCATGGTTTGGT
Pchgms12	F-CGACACTTAGCTAGAAGTTGCCTTA	R-TCAAGCTCAAGGTACCAGCA
Pchgms11	F-AAGCAATAAAACCAGCAGCAA	R-TCAATCAATTGGCATGTTCG
evg24	F-AGGCTTGAACCAGCAGAAAA	R- TCAAGGTGTGGAACAACCAA
evg28	F- GGGCATGGTTGATTGATTCT	R- TCGTTCTGCTGCTTTAATTGTT
Pchgms29	F-CCTGAAGAAGGTGGACCAGA	R-CCTCCCAATTCAAATTCCCT
Pchgms74	F-TGTTCTATTTAGCTTCTTCCTCCAG	R-TCCTTTTCTCTGTATGCCACTTC
Pchgms75	F-CAGCCATTTTGACTGCTTTG	R-TTCTTCTGGCTTGCATTGTG
Pchgms76	F-CAAACAACTCTCCCCATACCA	R-AGAGAGAGAGGGCTTTGACC
Pchgms77	F-GGAGTGTGTTCCTTCTGGATG	R-CCCAGCTACTGTTATGACTTTTCTC
Pchgms79	F-CTCGCTTCTGGGTTGTTCAT	R-TACTCGGATGCCACTTTTCC
Pchgms86	F-CGAAACCCTAACCCTAAACC	R-TGAGCTGTCCTTCACTCTGG
Pchgms88	F-CTGTCGCAGGAACAGTAAGC	R-GATTTTCCGGTACAGTCTGG
Pchgms80	F-GCAGTGCCGTAATGGATTTTC	R-TGTAGGCCCCTTAGTGGATG
BPPCT28	F-TCAAGTTAGCTGAGGATCGC	R-GAGCTTGCCTATGAGAAGACC
Pchgms17	F-ATGCACTCAAGTGGCAAGC	R-GGTTTTTGAGCAAAGATGCAC
Pchgms18	F-TTAAGTGGCGCACGTAAGG	R-TTTTGTGGGTATCTGAGCAAA
Pchgms73	F-TGCACCTTGTTTTTGTTTGC	R-GCCCGGAATATCTGTCACAA
Pchgms46	F-ACACCAAAAGCCACTCAAGTCTC	R-CGTCTTGGCTATTGGCTATTGCT
Pchgms106	F-GCAACAAATCGTCAAACTCC	R-GGAGAAACTGACGCAGAAGG
Pchgms108	F-CTGCAGCCAGTTCAATCC	R-CCCTTCCTGATTAGAGAGTCG
Pchgms109	F-ACAAGTATGGAATGAAACAAGG	R-CACGATTGTTTGTTGGTACG
Pchgms110	F-TGTTACCATGCTTGATGAGC	R-TTGGAAGGGCTTTGTATTCC
Pchgms112	F-GCTGTTATCAGGTGGTCAGG	R-TTGAGCTCTGATTGCTAGGG
Pchgms113	F-GATAATGCGGTGGAAACAACC	R-ACTGGTGAGATGGGTATTGG
Pchgms102	F-AACATCTTCGGGTCCATCG	R-AGGCACTGAAGTTAGAGAAGTCC
CPPCT33	F-TCAGCAAACTAGAAACAAA	R- CCTTGCAATCTGGTTGATGTT
Pchgms90	F-CATTGCGAACTATTCTCAGC	R-AGACTACCGGAGATCAAACG
Pchgms91	F-AGAGAAATGTCTGGCACACC	R-CCCTTCCAATGTCACTCC
Pchgms93	F-TCTTTGCTGAACTTGAATGG	R-TCGAACATTTAGGCTTACCC
Pchgms94	F-TGTTGAAGAAATGCACTTGG	TGTTTCGAAAACTTGAGATGG
Pchgms96	F-GCTTAAGTTAGCAAAGGCAACC	R-CCAGCTGGGAATGTAAAGC
CPSCT42	F- TGGCTCAAAAGCTCGTAGTG	R- CCAACCTTTCGTTTCGTCTC
Pchgms97	F-GCACGGCTATACTTATTTTCC	R-TAAGGCTCTTGGGCTATGG
Pchgms98	F-GGTCTGCCTTTGTTTTCTCG	R-ACACAGTGAACCCCAACTCC
Pchgms99	F-AATTAAAGATGATAACTTATTGGAACG	R-CTGAGTAGAAGACGCTAACACG
Pchgms100	F-AACAACAACACATCCTATTCTCC	R-ACCACCTTGTTCACTTCAGC
Pchcms2	F-AGGGTCGTCTCTTTGAC	R-CTTCGTTTCAAGGCCTG
Pchgms20	F-AATTGCATCACAGCAAGAGC	R-GGGGGTTTGGTTAAGATCG

Appendix E1: List of SSR marker, map position and corresponding predicted genes used for linkage disequilibrium-based association mapping on LG1.

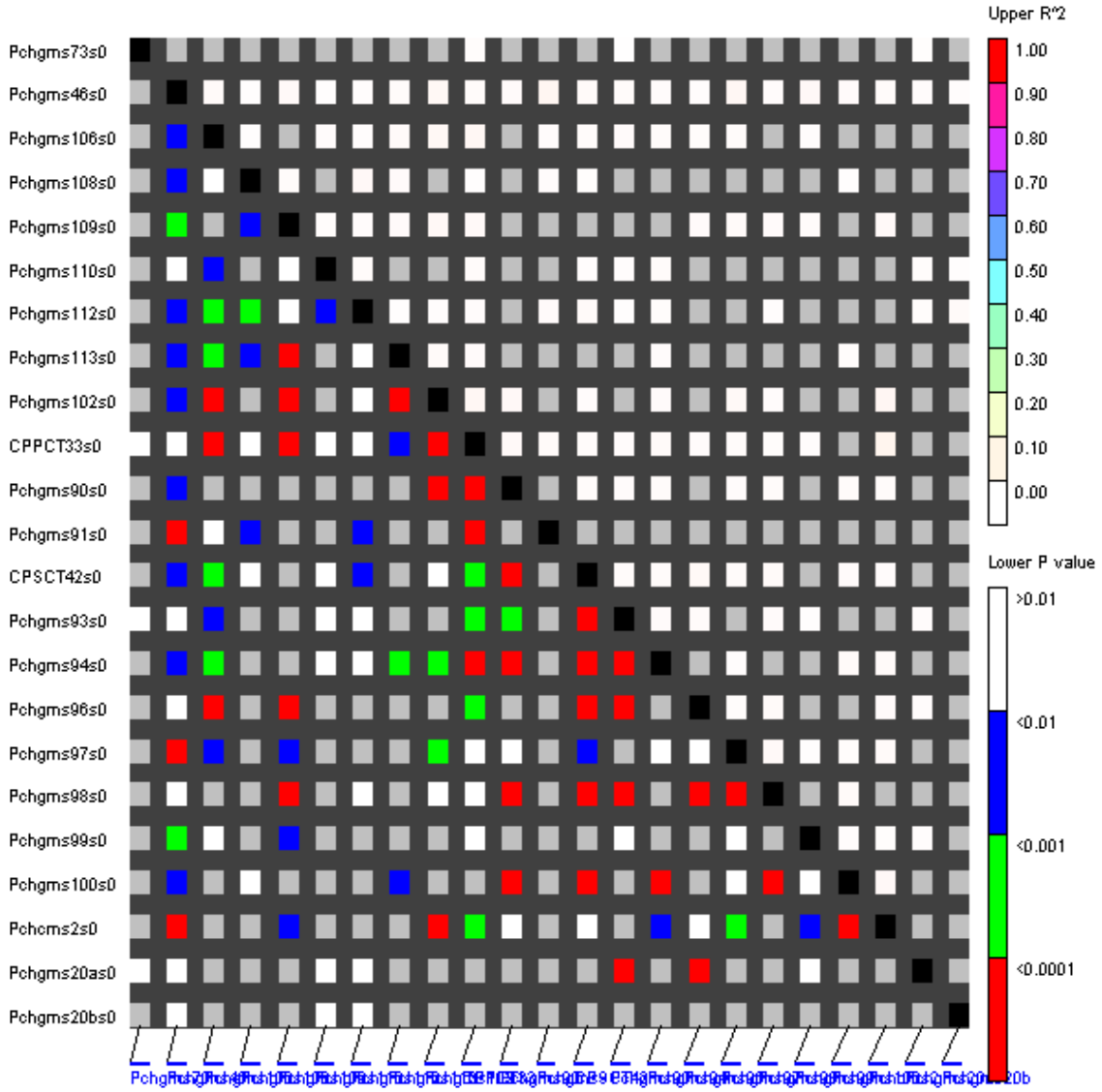
Marker	Position (cM)	Candidate genes
M3	76.00	-
Ampa109	78.8	-
Pchgms103	85.4	-
Aprigms25	85.50	-
Pchgms105	85.6	-
evg3	85.8	-
evg9	85.9	*DAM1*
Pchgms14	86.0	*DAM2*
evg12b	86.1	*DAM2*
evg15	86.4	*DAM3*
evg18	86.6	*DAM4*
Pchgms41a	86.8	*DAM5*
Pchgms41b	86.8	*DAM5*
Pchgms10a	86.80	*DAM5*
Pchgms10b	86.80	*DAM5*
evg20	86.80	*DAM5*
Pchgms12	86.80	*DAM6*
Pchgms11	86.80	*DAM6*
evg24	86.9	-
evg28	86.90	-
Pchgms29	87.0	*kinesin*
Pchgms74	87.0	*kinesin*
Pchgms75	87.2	Kinase interacting family protein (TAIR:AT1G09720.1)
Pchgms76	87.3	Myb family transcription factor (TAIR:AT1G58220.1)
Pchgms77	87.5	ATR2 (*ARABIDOPSIS P450 REDUCTASE 2*); NADPH-hemoprotein reductase (TAIR:AT4G30210.2)
Pchgms79	87.6	*VIN3/VEL1*
Pchgms86	88.0	UDP-glucoronosyl/UDP-glucosyl transferase family protein (TAIR:AT5G65550.1)
Pchgms88	88.0	*CTR1* (*CONSTITUTIVE TRIPLE RESPONSE 1*); kinase/ protein binding / protein serine/threonine kinase/ protein serine/threonine/tyrosine kinase (TAIR:AT5G03730.2)
Pchgms80	88.9	*IRX9* (*IRREGULAR XYLEM 9*); transferase, transferring glycosyl groups / xylosyltransferase (TAIR:AT2G37090.1)
BPPCT28	88.9	oxidoreductase/ zinc ion binding; basic helix-loop-helix (bHLH) family protein (TAIR:AT5G10570.1)
Pchgms17	89.1	*SUT1*
Pchgms18	90.0	*SUT1*

Appendix E2: List of primers for SSR markers, map position and corresponding predicted genes used for linkage disequilibrium-based association mapping on LG7.

Marker	Position (cM)	Candidate genes
Pchgms73	38.6	*TFL*
Pchgms46	42.3	*PHOSPHATIDYL INOSITOL MONOPHOSPHATE 5 KINASE*: 1-phosphatidylinositol-4-phosphate 5-kinase/ ATP binding / phosphatidylinositol phosphate kinase; PIP5K9 interacts with CINV1 to negatively regulate sugar-mediated root cell elongation.
Pchgms106	43.0	*CURLY LEAF*
Pchgms108	43.0	*CURLY LEAF*
Pchgms109	45.0	6kb downstream from AT1G49620: *ICK5*; cyclin binding / cyclin-dependent protein kinase inhibitor (*KIP*-related protein (*KRP*) gene, negative regulator of cell division. 7 kb upstream of AT4G18750 : DOT4 (*DEFECTIVELY ORGANIZED TRIBUTARIES 4*)
Pchgms110	45.1	AT5G66810: hypothetical protein
Pchgms112	45.4	AT5G66820: hypothetical protein
Pchgms113	45.8	*SQN* (*SQUINT*): peptidyl-prolyl cis-trans isomerase; encodes the Arabidopsis homolog of *CYCLOPHILIN 40* (*CyP40*).
Pchgms102	47.0	AT5G49530 - *SIN-like* family protein: DNA-directed RNA polymerase activity
CPPCT33	49.2	2kb downstream of AT3G50690: leucine-rich repeat family protein; 5 kb of AT4G36650: ATPBRP (PLANT-SPECIFIC TFIIB-RELATED PROTEIN); RNA polymerase II transcription factor/ rDNA binding
Pchgms90	49.2	Between 2 copies of AT5G07610: F-box family protein
Pchgms91	49.2	AT3G50670 - U1-70K (U1 SMALL NUCLEAR RIBONUCLEOPROTEIN-70K): RNA binding / nucleic acid binding / nucleotide binding; mRNA splicing
Pchgms93	50.2	AT3G50930 which is *BCS1* (CYTOCHROME BC1 SYNTHESIS): ATP binding / ATPase/ nucleoside-triphosphatase/ nucleotide binding
Pchgms94	50.3	Between duplicated AT2G18190 gene: which is AAA-type ATPase family
Pchgms96	51.8	AT2G18180: SEC14 cytosolic factor, putative / phosphoglyceride transfer protein
CPSCT42	51.8	*ATMPK7*: *MITOGEN ACTIVATED PROTEIN KINASE*; involved in circadian rhythm, signal transduction, response to hydrogen peroxide
Pchgms97	51.9	AT4G18905: transducin family protein / WD-40 repeat family protein
Pchgms98	55.0	AT3G51000 epoxide hydrolase (putative).
Pchgms99	55.0	AT3G51000 epoxide hydrolase (putative); several copies. Close to AT4G36390: radical SAM domain-containing protein / TRAM domain-containing protein
Pchgms100	55.1	AT4G36380: ROT3 ROT3 (*ROTUNDIFOLIA 3*): encodes a cytochrome P-450 gene that is involved in leaf blade expansion by controlling polar cell expansion in the leaf length direction. Member of the *CYP90C CYP450* family
Pchcms2	61.9	Close to AT3G18670-AT3G18670 ankyrin repeat family protein; close to AT2G06025: GCN5-related N-acetyltransferase (*GNAT*) family protein
Pchgms20	75.0	AT2G26580: *YAB5* (*YABBY5*); transcription factor; Nozzle protein bound filamentous flower and yabby3. AT3G15790: MBD11; DNA binding / methyl-CpG binding. Close to AT5G20200 which is nucleoporin-related



Appendix F1: Strength and extent of LD within a 14 cM genomic region of LG1 estimating pair-wise parameter estimates (r^2). Each square in the LD matrix represents a comparison between a pair of polymorphic sites, with the r^2 values displayed above the diagonal and the *P*-values for Fisher's exact test below. Points on the diagonal correspond to comparisons of each site with itself. Color codes for r^2 and *P*-values are given.

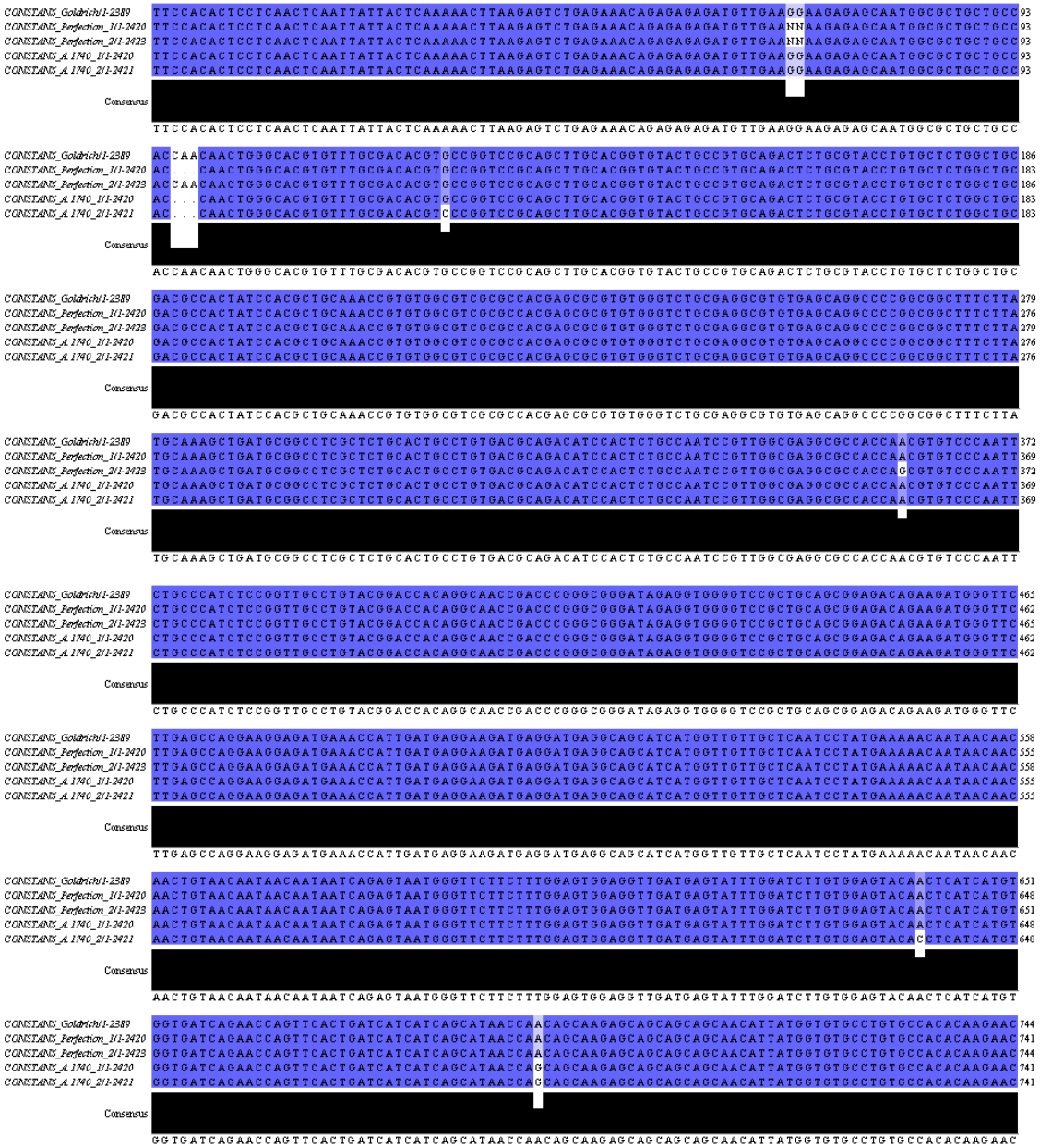


Appendix F2: Strength and extent of LD within a 25 cM genomic region of LG7 estimating pair-wise parameter estimates (r^2). Each square in the LD matrix represents a comparison between a pair of polymorphic sites, with the r^2 values displayed above the diagonal and the *P*-values for Fisher's exact test below. Points on the diagonal correspond to comparisons of each site with itself. Color codes for r^2 and *P*-values are given.

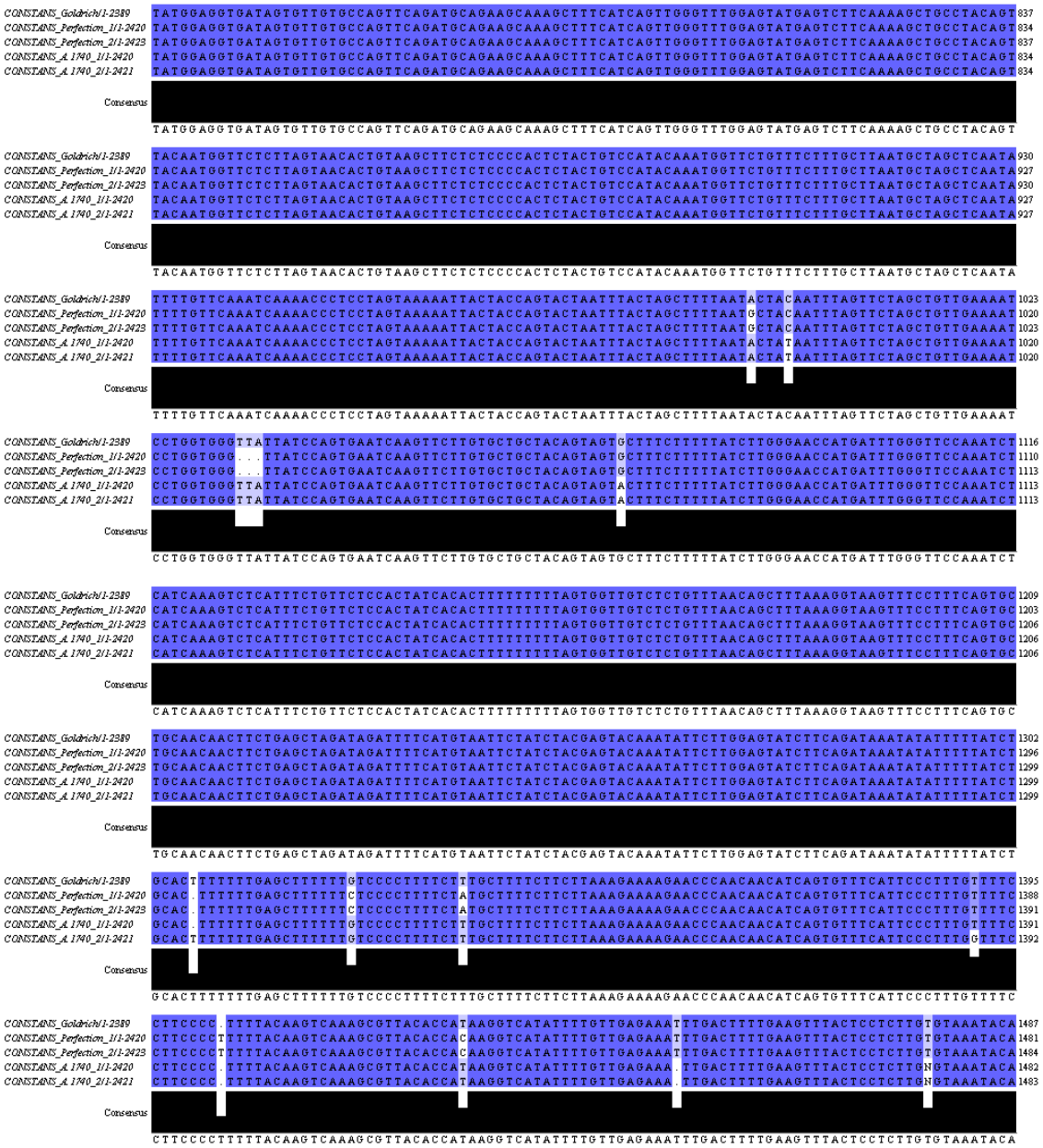
Appendix G: Goldrich apricot BAC clones positive (following southern hybridization) for probes designed from putative candidate gene sequences controlling chilling requirement and bud break.

Candidate genes/SSR marker	Linkage group	Goldrich BAC clones
MPK7/CPSCT 42	7	G086H19,G015N20,G155O12,G176O01,**G212G03**,G062D17,G192 P01,G038F11,G136M05,G137D02
AP2/AP2M	6	G004A12,G112H11,**G224J17**,G104G11,G118B01,G011A24,G073 O19,G187J11,G194I04
CONSTANS	1	G203N03,G101N02,**G179I08**,G160P04,G195C23
SUT1	1	**G158G23**
FT	6	G049L23,G081J14,G226K05,G149C13,G174C04,G158C15,G019J 07,G218M10,G063B08,G043N08,G166E13,G197A12,**G226D18**
TFL1	7	G013O10,G039O12,G061G15,G073G15,G107O03,G135E20,G141 J23,G159M19,G170H21,G170I22,**G210L03**,G215C09,G225B18,G 241F17
LEAFY	5	G195M04,G180G08,G010E10,G011F24,**G023F21**,G216D06
AGL2	3	G023M19,G007M17,G097E07,G178P16,G007O19,G099K05,G148 M17,G067P18,G119P20,G186H19,G048M04
ABI3/EPDCU3 454	8	**G059G19**

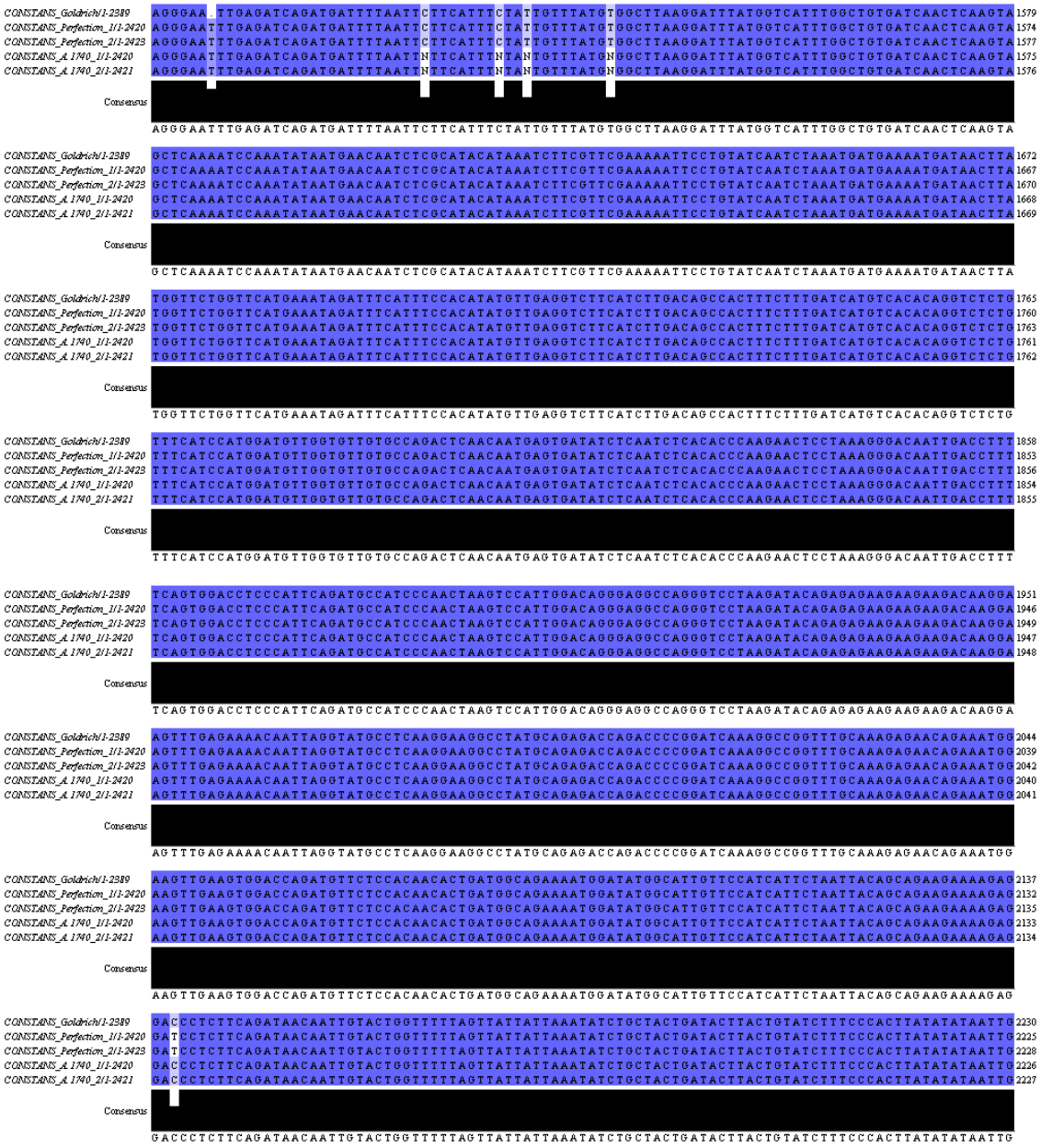
Note: BAC clones in bold were used as template for sequencing candidate genes.



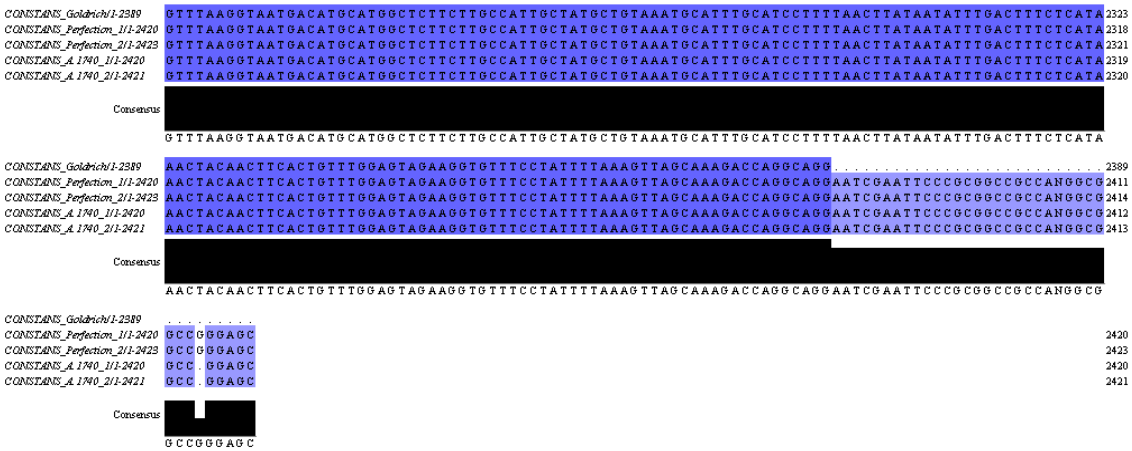
Appendix H1: Map-based cloning of *CONSTANS* from the Goldrich cultivar apricot BAC library and the 2 allelic variants of Perfection and A.1740 apricot cultivars. Sequence alignment performed with MUSCLE v3.7.



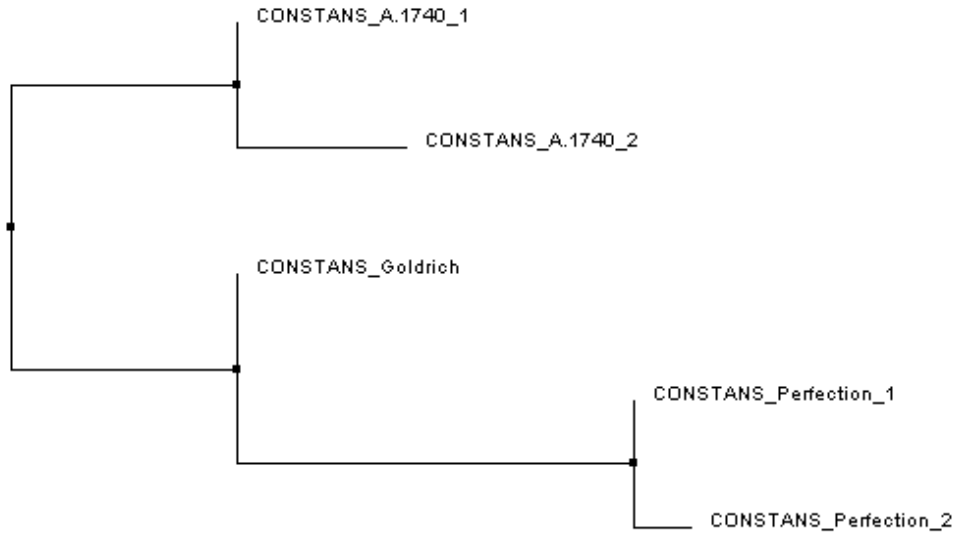
Appendix H1: continued



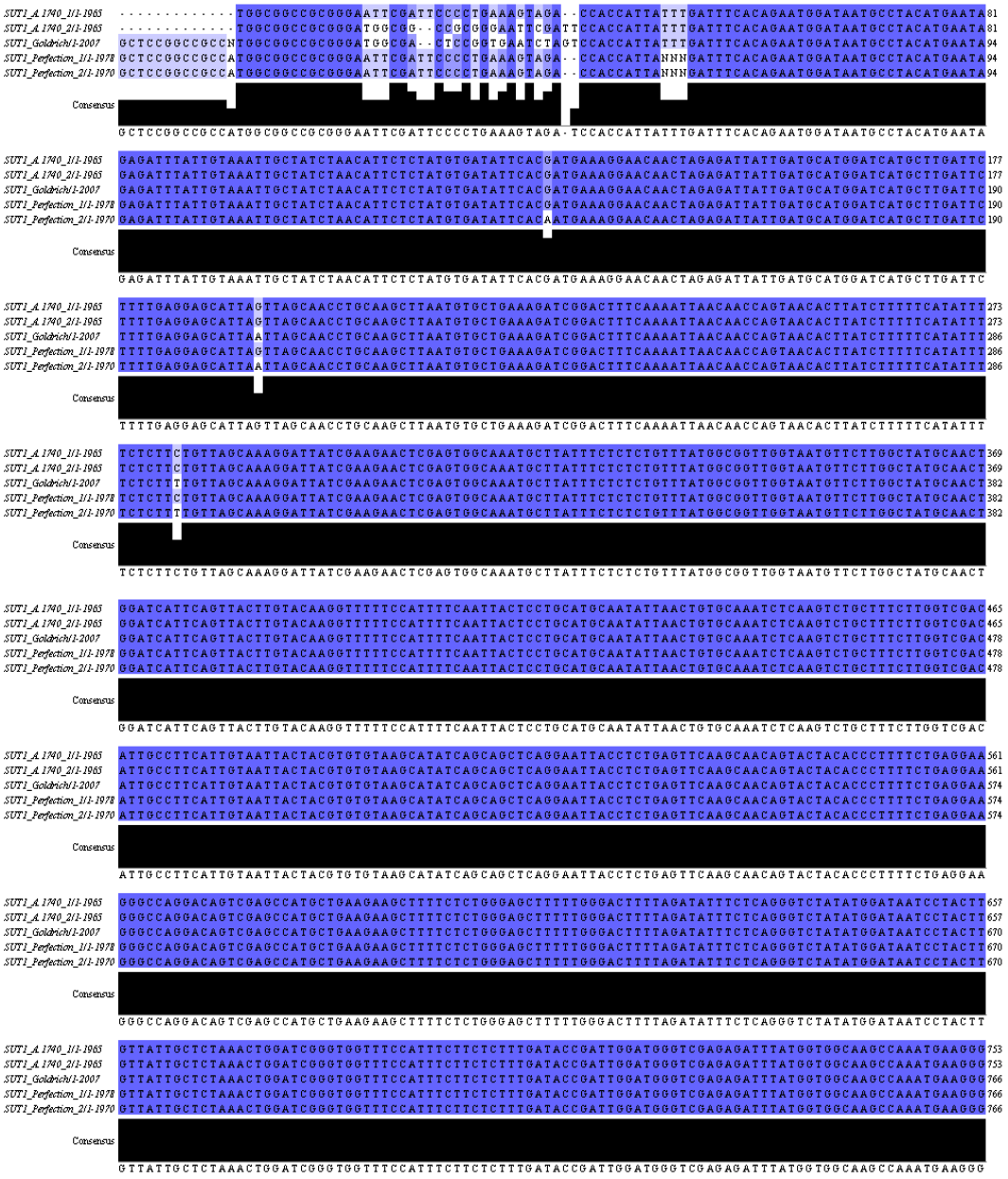
Appendix H1: continued



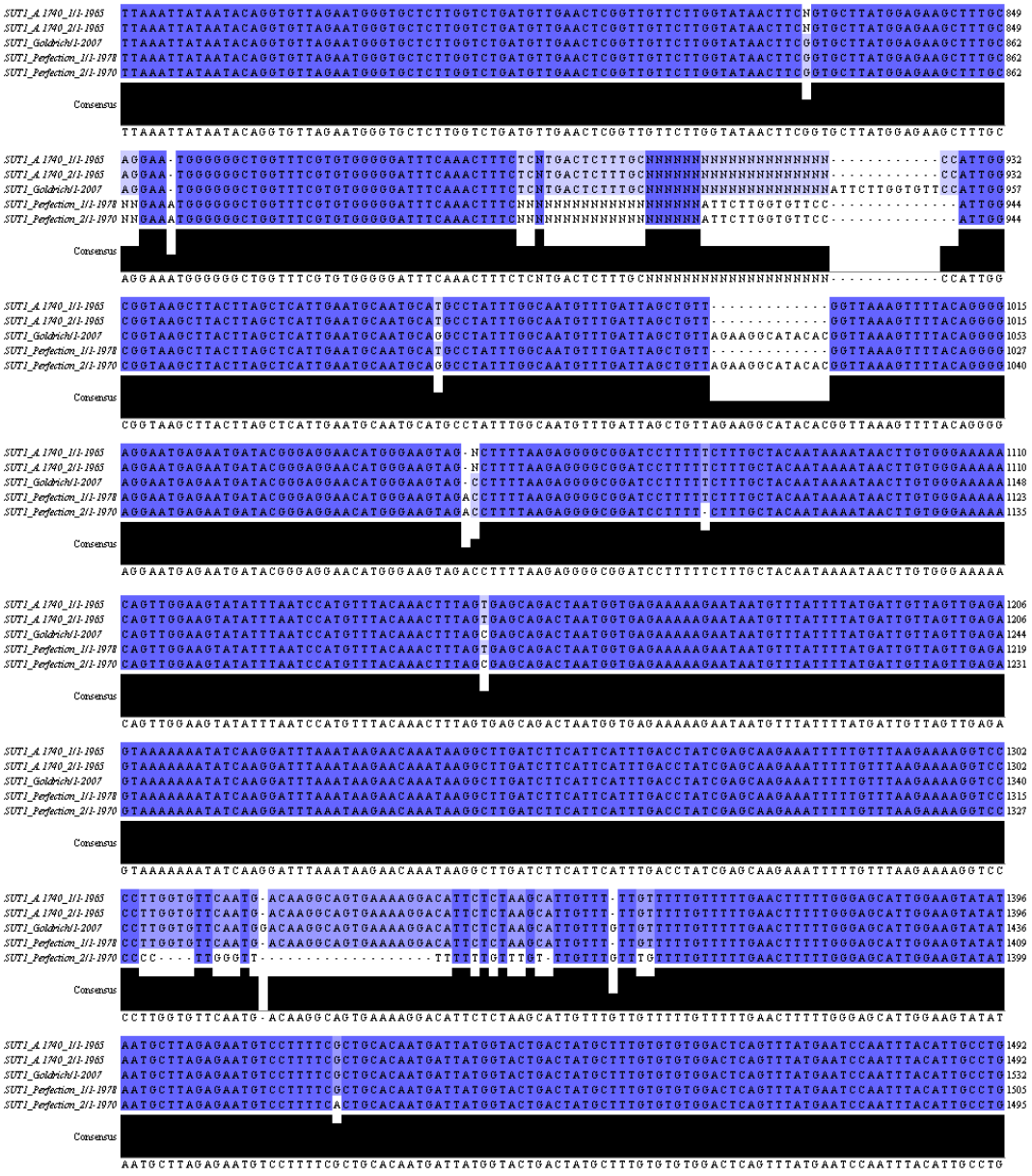
Appendix H1: continued



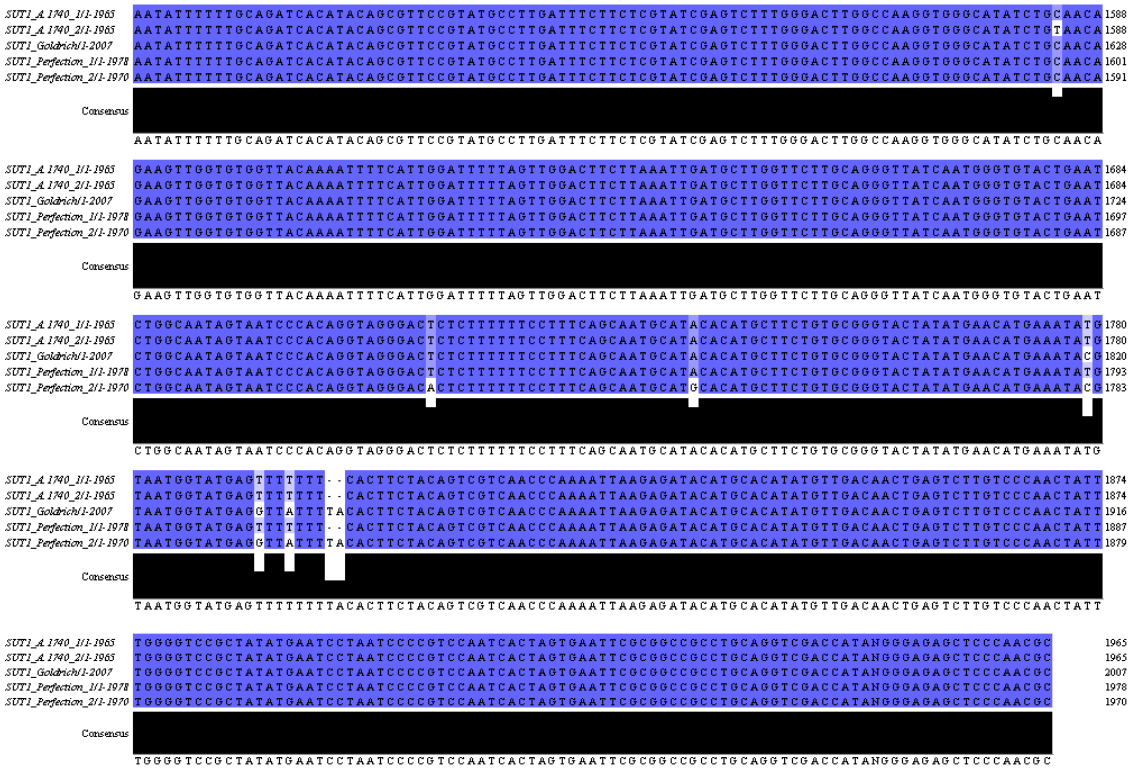
Appendix H2: Gene Phylogeny of *CONSTANS* alleles from Goldrich cultivar apricot BAC library, 2 allelic variants of Perfection apricot cultivar and 2 allelic variants of A.1740 apricot cultivar. Phylogenetic analysis was based on Neighbour joining tree using PID, following alignment using MUSCLE v3.7.



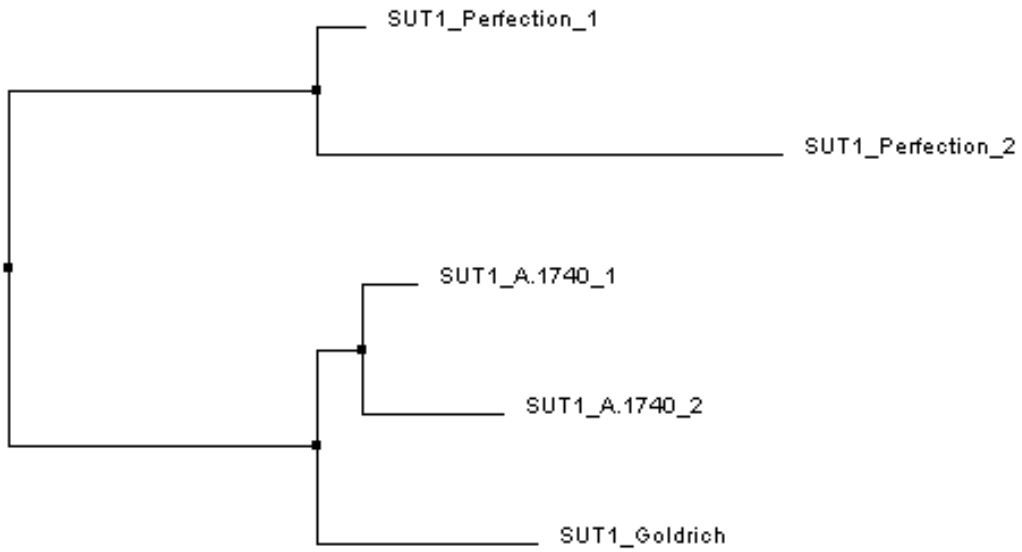
Appendix I1: Map-based cloning of *SUCROSE TRANSPORTER 1* (*SUT1*) from the Goldrich cultivar apricot BAC library and the 2 allelic variants of Perfection and A.1740 apricot cultivars. Sequence alignment performed with MUSCLE v3.7.



Appendix I1: Continued



Appendix I1: Continued



Appendix I2: Gene Phylogeny of *SUCROSE TRANSPORTER 1* (*SUT1*) alleles from Goldrich cultivar apricot BAC library, 2 allelic variants of Perfection apricot cultivar and 2 allelic variants of A.1740 apricot cultivar. Phylogenetic analysis was based on Neighbour joining tree using PID, following alignment using MUSCLE v3.7.

> *LFY* sequence from apricot Goldrich cultivar; 53 bps missing from START CODON. 322 bps after TGA stop codon

GCCCAGNNCCCAGNNCANNGCGTNTNGCNTCCNCATTAGCCCTCCCNNNCAAGNNCCGTCAGTTAACNNNAATATTTT
ACTATTANNTNNGAAAAGAAGATAAAGTCACAAACTTGTGAACCCTCTCGTGTTTTAATCTCCATTTTCTCTCCCNCATC
ACATCACCACTTCATCGTTCACCATTCATCATCAGAACCGTCTGCCTTGGGCACGGCTCGTGGACCACCACCACCGCAC
CAGACCNTCCGCCGTCCTCCGCGCGACACGTGGGTGGGAGCCGCAACAACATCGTATTCGTCAGGGCCACGTGTCGCT
GTATAAACGGGAACACGTAGGTGTTGATGCACCTGATCGTGACACTAGCCTAAGAAGAAGAAGATGGATATGGTTTCG
TTTGTCTTTGAGAAAGTACAGAAAAGCAGCCTTTCAGAACACCTGCCCACCTCTGTACTTCCCTGCCGCTTCTTATATCG
CCAAGGGGAACAGTAAATTTCACTCGTTATTATAACATCGAGTGGAATGACGTGTTTGCCCCTCTTGCATGCTTGGCTT
GTTCGGGTAATTAAAGCTTGAGAATCTCTACAGAGATATGGGTTTTCGTCACGTTGTCCGAGGAAGAGATGATGATGAT
GATGGTGATTGTGAAAATTACCATCTATTACACAAACGCCAATCTTAATCCCCACATTCCATTATGCCAATTTTACTAAG
CTTTGCCCCCATCAATTTCACCGTCAAAGATTTCTTGTTACTACCTACTGCCAACAATAAATAAATAAACCCAAAAATA
AGAGTGTGTGTGCATCTCAAAATCACGTCATAGAGTCAATGGGTTCTCACCATAGTGGTGGTCGTTTCCCCCTCCTCAC
GGCAACCCAGTTCGAACTCGTTCGAACTCCAATGTAACCAACAAAAATCACGTCATAGATTTTGGGCCCTTCAAAATGT
TGATTTTGTGCCATAAACTAGGAGACCACGCAGACAAGTAGACGAGAAGAAAGAAGCAAGAAAGTTTGGAAAACTTC
ACTTCTTATTAGCAGTTTTAGCACTCTGAGGAGCAACATCTACACGCATGAGTTTCATTATTAAAATCATTATGTTTATT
AAAAAATAAAAATCGTAAACTCTAGGGCTAACATAACATAACAAGTTTAAAAGTTCATTGGGTGGTGGGACCAATTAG
AAGAACCCTAAAGAGTGTTTCATTCTGAAATACCGAGTTTGAAACCATCAAACCCCAGTTTGCTTAGTTCCTCGGGAAA
CATCCCATATGTAGGAGGGTAGTTTCGGAATAAACTAAAATGGTCTTGTTTTTATAGGAACAGAAGACTACACACACCA
ACCACAGTGCCAAAGACTAAAANCAAAGCTAGAGAAGAAGCTGTGGTCTTCATCTGTGTGNGGTTTCACACNGCGG**AA**
AAnnnnnnnnnn**GGTTG**TTCCGCCGAGTCGGGCTCAGCTAGAAGCCGCCGTGACGCCTCAAGCTGCCNNNNCGGCTTACG
CTGCCGTGAGGCCCCCGAGAGAGCTCGGAGGGCTTGAGGACTTATTCCAGGCTTATGGGGTCAGATACTACACGGCAG
CGAAGATAGCCGAGCTCGGCTTTACTGTCAACACCCTTTTGGATATGAGGGACGGTGAGCTTGACGACATGATGAGTA
GCCTCTCTCAGATATTCAGGTGGGATTTGCTTGTGGGTGAGAGGTACGGTATCAAAGCCGCCGTCAGAGCAGAGCGTCG
CCGCCTCGATGACGAGGACTCCAGGCGGCGCCACACCGTCTCCGGCGACACCACCACCACCAATGCCCTAGATGCTCT
CTCCCAAGAAGGTTCGTTAGTCACTATTACATGAATTCCTGGAATGAAAATTTACATGTTAGCATAAAATTATACACGC
AATATTTCATCATAAGATGCAAAATTATTTAATCAATTTGTTACAATATTTCATCATAAAACATTTTTTTTATATGAATC
GAACAAAAAACTAATTTTTATTTAATAATATTACATGTGAAACTATTTAATTGTTACCATATTTCATCATAAAACATTTT
TAATATGTATGTATATGACATGGCGTGCATGGGATTGTAGGGTTGTCGGAGGAGCCGGTGCAACAAGAGAAGGAGATG
GTGGGGAGCGGCGGAGGGGCCGTGTGGGAAGTGGTGGCGGCGGCGGGGGNNNNNCGGAAGAAGCAGCGAAGGACGA
GGAAGGGGCAATATAGGAATTTCAATGGTATCGGAGGGGGGCATAATAATGATCATAATGAGGGTGTGGACAACGAG
GACGACAACGACATGGACGACATGAATGGGCACGGGAACGGTGCAGGAGGGCGGTTGCTGAGCGAGAGGCAGAGGG
AGCACCCGTTCATTGTGACTGAGCCTGGGGAGGTGGCACGTGGCAAAAAGAACGGCCTAGATTACCTCTTCCATCTCTA
CGAGCAGTGCCGTGATTTCTTGATCCAGGTCCAAAACATTGCAAAGGAGCGCGGTGAAAAATGTCCAACCAAGGTACG
GAGTTTACCCAACCCCCGTCTTCATAACCTAAATGCATACGCTGATTTATACTGTGGTAAATAGTAAATACTAAAATAG
TAACTTGGCGCATGGACTATCATACCTGGTCAATGTGGTCCCATTTCCGCAAGTACGAACAATTACAATCTAGTGGCCC
GACAGTTTCAATTGGAAAGGCCTGCTGACAGCATCAGTATAATGTTTGAGCTAACTCTGGACCGGGTCTGAATTTTTCA
TTGATCATGTGGCAGACTAAGTTCACAATAATTTTTTTGAAAAAGATCATTGTGAATAAGATCATATTTTAATCCTACTT
TGGTCTATATTTGATCTGCTTTGTTAGAAATATATAGTGATTCCAACAATGATGTTAATATGACGTAGAATTGGATAATT
GGTTAGCTTAATAAAGGTAGAGATTTTCACACTCGCGTTTTGTTAACTCCGCTCTTTGTTATATTTTGATGATCTATCATC
TACTTTTATTTCTCAAGTAGAAAAACAATAGAAAAGTGCAAACATATAAAATCGGAGTGTGAAAGTCGACACCCTTAG
GTATCCGATATTTGTAAATAATGAATGATTTATCTTGTGACCATGACAAGTGAATTGTTGTATTTTGGTTGTGTGAACAT
GAATTATTGTGCAGGTAACAAACCAAGTGTTTAGGTTTGCAAAAAAGGCAGGGGCAAGCTACATCAACAAGCCCAAGA
TGCGACACTACGTGCATTGCTATGCGCTGCATTGCTTGGACGAGGAGGCCTCCAATGCACTGAGGAGAGTTTTTAAGGA
GAGAGGCGAAAATGTGGGGGCCTGGAGACAGGCATGTTACAAGCCTCTTGTGGCCATTGCAGCAGGCCAAGGCTGGGA
CATTGATGCCATCTTCAATTCTCATCCCCGACTCTCCATTTGGTATGTTCCCACCAAGCTCCGTCAGCTTTGTCACACTG
AGCGCAACAATGCCACAGCCTCTAGCTCTGCCTCCGGTGGTGGTGGTGGCGGCGATCACCTACCCTAC**TGA**TCAG
AGTGCATGCCCCTCGATATGAGTTTGAAGAAAGAGAGAATTTGAGGAAATGACCTAGTGGGTTTTATGTACTTTAGGTT
AGAACTTAGAACGGAGACGGTTAAATGGATGAGTCTTATTTNNNATCTTTGTGTTCTGATGTTTCAAAAGTTAGCTATA
TTTAACTGTTTTTACTAATCTACCACAATGATAAAATCTAATGNGCTTGTAAATGGAATGNGNAATTGNNNNNNGTTTT
TTANAGCACTTTCNNTCATGNNNCATGNNANNNNANGGNNNNNNNNNNTNCCTACCTTTTTTGNTTNNTTTNTCCNTNC
NN

Appendix J: Map-based cloning of *LEAFY* (*LFY*) from the Goldrich cultivar apricot BAC library.

> *AP2/ERF* sequence from apricot Goldrich cultivar; 53 bps missing from START CODON. 90 bps after TAG stop codon

CATTTACAAANGGNGTTTTCATGATTTAATTGCTTATAAAACNTGATAGTCATGCACCAACCCACCCTAAACANGGATT
CATAAAAGATGTTAAGAAAACTTAGGTATTTAGGGATGACCACACTACCNTTCTATCAATTCCAATNGCAGCTTAATGT
GAACTACAAGAGTATGCTTATGCATTCTTAAGCTATAAGAACAAAAATAGAAACCCTAACAAACTAGAGTTACTAAAT
TTGTAATTCGGCCTAAAGTATGGTTTCGGTTCTTTTATTACAAGTTTTGATAGTTTAATTATTTTGTGTTGGATCATACTA
CTGAACTAGATTGGTCATATAACAATCTCAATAACAAATCCTTAAGGTATAATAACGATAATGTCGACACTTGATGTTT
AAAAAAAGTCAATTCCCATAAAATCTAGTTTTGACTTGTGACTATATTTTCATAGGGCATATGTAAACTTATAAATTCTC
TTAAATTGTAATACACGTATTATTTACTTAAACCAGAAATTCCAAATTGTGAGACTACTAATCCTACACATGTTACAATT
CTTGAATACTTGTAACACCCCATTAACATAGTCGAACTAAACTATTCAAGTGAAGGAAGGACCACAAATGGTGTGAAA
AGGAATATAGTTTGAATGAAGAAGAAGTAATCTTGTCATTTGTAAAAAATTGTTTAAGCATTAACAAGGGGAAAGATG
AAGTTGAAAGAGAAAAGNNNGGAAGAAGAGAAGAAGTGTAGGAGACGCGGGAAAGGAAGAGAGNGAGGAGAGGnnn
nnnnnnnCACANTACGANGTACAGCAAATAGGAAGGGGCATTGGTCGCTAGAAGAACGAAAGAGGCAAAGNNGCCCTCT
TCCNGTANTCTNGGATNGGACTTTTTNGNAAAGACNTTCCTTTTCCTACCATTATTCCCNTTTCCTTTTTCTGATACCAAA
CCAAACCCAGTAATAACAACCCAGTCACCACTCACCACCTCTCTCTCTCTCTCTGAGACTGAGAGAGAGAGGTCCCTGA
TTTGATAATTACACCCTTTTTTTCTTTCCCAAAATGCCCCGGTTATCCACCCACTATTTACTGCCCTTCTTCCTTATGGTA
AAAACTAAAAACCAGGTGAAACACAAGATACCTGAATATGTGAAAAAAAGTGAAATCAAATAGTTTTTTTATTAATTA
TTTATTTATTTTGTTGACTTGTGATGTGAACGTTGAAATAATTAAGTTGCTAATAAAGGCAATGGAGTAATTAGCCATCT
ATTTTAGTGTTTGGGCTGTCTACTTCCATTAGTGGGTAAATAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCA
GAATGCTTTTCGATTTGGTTAACGTGGGAAAGTTGTGTAAGAAAGAGATGGTTCTTTCTTTTCTCTCCCCATTACCCCAT
CTGATCCATGCTCCTTTGAACCCACCAATAACCCCAAAAAAAGAGAAAAAAAGAAAAAAAAGAAAAAAAGAAACAGA
GAAAGAGAAAGAGAGACAACAGTAAAGTAGAACTGTACACAACAAGTTTAAAGCCCAAAAACTTCAAGTTTAATCTTT
ATTTAATATAGATGCCAAAAATGCATACTACAAAATTTTATAAAATTGAGTGATTGCACAAAATAAAAAGTAAAGAGA
GAGAGAGAGAGAGAGAGAGAGAGAGAATTATGAGTCATTGAAAGGAAAAGGAAGCAAGTGGTTAGAGATGGAGAGA
GAGAGAGAGAGAGGGGTATTAAGGGTGGGGTGAGAAAGCTTGAGGCTAAGGCTTTACTGAAACTGGCCTCGTCTATGC
TGTATTGGGCAGTTTTAGGCGCCATATACAAGCACAAGATCTTGGAGTCCATAGCTCCATAGCTTCATAGCTCCATAGC
CCTCTCTCTCTCTGTCTCTCTATTAACCCTCCTCCTACTCTTGCCTTTTTGCAAAGACACAAAACCTGCAACATTACCCAC
CTCTCTCTTTTCTCTCTTTAACTTTTCCTTTTCGTGTTTCTTGTTGAGGCCTTTTTTTTTGTTTGCCACTCTGAATTCTGGGT
TTGATGTAAAAANTAATACTGAACACTGAGCAAACATCTGGTAGACAAACCCAACAAAGGGACAGAGAGAGAGAAAAAN
AAAnnnnnnnnnnGCGNNGATCGGAGGGNTGCTCCTCCCNGAAGACCTCGGCCGACGGAGACGAGGAAAAGGGCAAACG
GGTCGGATCCGTGTCCAATTCAAGCTCCTCGGCCGTGGTCGTCGGGGACGACGGATCAGACGAGGAAGAAGACGACGA
CGGCCCCGCCAAGCTCGCCAAGACGCGGGGAGGAGGAGGGGGCAAGATATTCGGGTTCTCCATGACCCACGAGGAGT
CCATGGATAGTGATCCGCCCCGTGACCGTGACCCGACAGTTCTTTCCGGTGGAGCTAGACTCCACCTCTGAAATAATGGG
GCCCACACGAGGAGGCGTCCCGCCTCCAGCTCCTCCTTCTTCGTCGTTGTCGTCGTTTCCCAGGGCCCACTGGGTCGGG
GTCAATTTCGGGCAGTCGGATTCTGGCAGCCCCGGGAAGCCGCCGGCCGCGGTGGAGGCAGCCCACCAGCCCATGAAG
AAGAGCCGGCGTGGGCCTCGCTCCAGGAGCTCTCAGTACCGCGGCGTGACGTTTTACCGGAGGACTGGCCGATGGGAG
TCTCATATTTGGTCAGTTATTAGTTAGTTAATTTCACTTCACATAATTAGTAACTAATTTTTTCAGTTACTATTCCTATGT
TTCAAACTATATATATATATATATATATATATATATATTTCTATGTTTAATTTATTTTGTACTTTGGCACAGGGATTGTGG
GAAGCAAGTTTATCTTGGTAAGACTCTGACAAAAATGATGCAATTTTAATTTTATTTTATGGTGTGTTTACTGATGGGAC
TTTTTTTTCCTTTCTCCATAATACGATTAATAGGTGGATTTGACACAGCACATGCGGCTGCACGGTGAGCTCTTCATCTT
GAAAAACCAAGGTTTAATTTTGTGGGGGAAAATGTGGGGGCTGGAAAAAAGATGAGAATTTATTTTGTGTTATGCAGT
GCTTATGATCGAGCGGCCATCAAGTTCCGGGGAGTGGAGGCTGACATAAATTTTAGCATAGAGGATTATGAAGAAGAC
TTGAAACAGGTGAGAGAGAGTGTACGGTCCCCCAAAAAAAGAGAAAGGGCTTTGAACTTATCTTCAACAGTTTTTTTTT
TTTAATGCCTTATCTTCAACAGTTGCAGAAAATAAATAAATAAATTTCTTTTTAAATTTTTTTATTTTATTTCTGAGTTTT
GGTTTGTTATAATTTTCTTGGGATGTTTGCATGAAATGAAACACAGATGACCAATTTAACAAAGGAAGAATTTGTGCAT
GTACTTCGCCGACAAAGCACCGGGTTCCCCAGAGGTAGCTCCNAATATAGAGGGGTCACCTTGCACAAGTGTGGGAGA
TGGGAGGCCAGGATGGGCCAATTTTTAGGCAAAAAGTAATCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTCTCTCTAGCTTTGCCCTTCAATTCTCATTTCACCAGTTTTCTTGTACTACTTCCTTGCTTCATGAATTGCTCAATG
ATTTTCATAGGTACGTGTATTTGGGCTTGTTTGATACCGAGATTGATGCTGCAAGGTACATATATATACATGATATCGTG
CTACGGATCTGCTCTCTCTCTATATATAAATATATATATATATATATATATATCTCGTAATGTTGAGTTCAGTTCCTGAA
GATTTGAGTTTTTTTTGGGGTTAATTTTATGCTACTTCAATTCTTGTAGGGCCTATGACAAAGCAGCAATCAAGTGCAAT
GGCAAGGAAGCTGTCACTAATTTTGATCCCAGCATCTACGAGAACGAGCTGAACCCCTCCTCTGGTAATTCAAATTGTA
ATAACATTATTAATGCATTGTAATTAATCTGATAATTTGGATATATGTATTTGCGTAGAATCATCCGGCGTTAATCCTGC
AGAACACAATCTCGATTTGAGTTTGGGCAATTCAAACTCGAAGAAAAACAATCAAGCTTTTGGGAGTAGTGATCATGG
CCAAAATGCTGCAATGGAAGTTCAACATTCTGCCTCAATGCAACTCGAAGCCGATTGGCGGAATCAAGGGTTTCGACA
AAAGGTTTCGTAACAAATGTTATTTTTATTTAATTACATTAATTCTTGCTTGAACGTGTGATCCATACATATTTGTATCAT
ATATCCGTCTAATTATTTTTGTTTCTTACATTTAATAGCTTAACCTACAAAGGGATCGATCTAGAGAGGAGAGTGATGCG
CACAGAAGAGATGGATACTTAGAGACAGAAGCCATGCAGCTACTACTCAGAACCAACCTTCATTCTCCAGCCCCCACT
GAAATGCATAAATATGGGCAGTTTAGTAGGAGGCCTAACGTTGGAGACACCCAAATGCCTCACACTTTCCCACCACATT
TCAACTCACCAAACAATTACCACCATGTAAGCACAATTAACTCATTTTCGACTCATGATTTGCTACGTTGATGATCAATT
TGCTTAAAATTATTGTGATTATTTTTCCTTTTGTCACAGGTTCAGTTTCCAAGCAGCAGCGAAGGAGGCCGCATCGGCAG
TGATCTTTCACTCTCGATGAGTGACCACCCACACCAACAACAATGGCAATCCGGCATGCCGACTTCCGATATATTTGCA
ACTGCTGCAGCATCATCNGNANTCCCACCTCAAATCANANCGTCCGCGCAAAATTGCTGGCTGCAGAAAAGTGGCTTC
CACTCTCTCACGAGACGC**TAG**CTATAGCTCCTCTCTGACCAATGAGCCATGACACACCCCCCCCCCCCCTACATTTTTCT
GCCTAATTTCCACCATGTCCCATTTCATCA

Appendix K: Map-based cloning of *APETALA 2* (*AP2*) from the Goldrich cultivar apricot BAC library.

> *FT* sequence from apricot Goldrich cultivar

GAANGAAAAAAAANNNNNNAAAAANTAANNAAANTTNNNCCNNNNAGNTNGATNNTCGNGATNTTTTGACCNNTCCA
AANGANCNNGGCNCTNGNNNNNNCGGATCGTNGGGTNTAAGTTNACNCTNTAANGNNGNGAGATTCGGCAGTTGNTA
ATACTAATATANGTANNTNTTATATTTAGGAGANGAGATCAGAGACAGNCAGGACAACCGATGTATAGTGGGGAATGG
AATGGAAGTGCATGCGTAGGTACTAGGTGATGGTGGTGNGGGGTGATATNTCNTNCCCNGAAGATNTCAGTAATTAAG
CGCATGCGTGNTTCGCTGCTCATGATGTATATATNTGACTGAGACTGACTGATCACTGCTCACNNNNACTGACCATATA
TAAGTATGTATATAAACCAGTGACATGATATGATATGATGATACGATGAATATGTAGTTTACGAGTTGGCAGTGTGCAG
TGAAACCTACAGTTTTGAGTTCATGAGCTACCAAATAGTACAAGTACCACCAACTACAGGGCTACAGGGCTACAGGGC
TACAGGGCTACAGGGCTACGGGGCTACAGGGCTACAGATACAAGGACTCCAAGGAGCAAAGATATGTATGTATTCATC
ATCGTCAACTATTTCGAATCTCATAAAATCATGTATGTCCAAACTCAAAATTAACCACATTCACGCGATTCTTCTAAACC
CTTTGGTTCAATTCATCACATCCACACACGTACCCATTCCAGACTAGCTACTTTGTTTGAATCCACCAAGATTCCCTGAG
AAAGCGATCTGCCACTGCCTCGCCTATGCCACACAACCCAATAACAAAACTTAGGCTTTAATTTGCATTTATGTTTGTTT
CTCTTTCTCTTTTTATCCACTCACACTTTTACACTTTCACACTTGAGATAGCATTATCTCAAGTGTAAGAATAATTATTAT
TNTTAACCAATTCAGTTACAAGTCTCTAATCGCTGACTATTTCATATTATATACATATATGTAATGCAAAAAAGAATTTG
AATAATTTGAGATTGCCAGGTAGCTAGAATTTTCTTCTGGAGGGGAAGAGCAGATTTNGACAACATGTACGATGATGG
GGGAGCTTTTTGGACCGACAACGCATAAACTGATGAATTGGAGAATAAAGAGTGGTGGTCCACTACTTGACTTTACAG
GTTGATGAAAGAGAGGAACTTAGTAAGAATCAAAAGAAAATAAAAATATCGAGCAAAAGGCAAAAGCTTTCCACGGA
GCAATCAGAATATCTTTTTTAAGTTACTGGTTGCACAAGAAGAAAAACAAAATGGGCAGCAGAATATAAATAGCCAAT
GGCCCTCTCGAATTGGATCACCAAACCAACTCGATCTTCTCAATATCAAATACAACTAGTTATTTGGATACTTTGTGTTT
GAGTTCGAGTTTGAGCTTCGTGTATATATCTATATATATATATATATTTACTCGTCAAAAGATAAAAAGAAAGAAAGGA
AGAAA**ATG**CCTAGGGACAGGGACCCGCTTGTTGTTGGAAGAGTGGTAGGTGATGTTTTAGACCCGTTTACAAGGTCTG
TTTCTCTCAGGGTCACTTACGGTATGAAGGAGGTTAACAATGGTTGCGAGCTCAAACCTTCCCAAGTTGTCCAACAACC
TAGAGTTGATACTGGTGGGGATGATCTTAGGACTTTCTACACTCTGGTAAGTGAAGTAATTAACGTTTAGTATACTTTTA
CTACTTTATTATAATTAGNTCTCTCTCTCTCTCTACTTTCTTCCACTCCATCTTNTTNNNNNNNATTAAGTGCATACGTAC
GGGGTCAATCCNTCCTCTTGCTTGCATGCTNNNTNNCAAnnnnnnnnnnTACATANGATACATTAATTCATTGTGCAGGTCA
TGGTGGATCNTGATGCNCCCAGCCCAAGTGACCCCANCNTTAAGGAATATTTGCATTGGTGTGTATTATTTGTAGCTTCT
CTCTCCCAAGGAATAAAGACATCTCTCTCTCTCTCTCCCTCCCTCTTTTCCGTAGAAATTTGTCCTATCCCAGCTATAT
AGATTGACTATATCTTCCCCCTTTGGCTTAAGCAATTTGCCATCAAAGTTGAGGTTTCGAGTTCAAATCTTCATCCCTTTT
TTGACCAAATATGTATTGGGGAAAGAGTATATATAGACGAAGAAGGATTTGCTTAGGAAAAAGAAGAATGTTCATTCA
TTTATTCCTGGTTGTTTTCGGAATATCAGTTAAATGTTGGTCTTATACTTCTGATTGTGCCAACTTTTCAGGTTGGTTACG
GATATACCAGCAACAACGGCGGCAAGCTTTGGTGAGTAGTTCCTATTATATTCTAGTTAGGGTAATGGTAGGCTTAATT
ACCACATTTTTATACCACACCGTGTACCACCTGTCAAATAGAGATGGAGCCTACCAATACAATGGGGCCCACATCTATT
AGAGAGGTTGTACATAATGTGGTATAAAAACGTGATAAATCTAGCATTTTCCTTCTAGTTATTGGGCATTTTTTCTTGGT
TTTTTTTGGTGGAGGGGGGAATATTTCACTCATTACCATATAAGTATGAGCTCCTACATTCTTACCAATAGTAATGTGAC
AGGTGTGCATAAAATTTTCAAAGAAATGTTCGATTGGTCTAGTCTACAAAGGCATAATTANNNNNCATTTGATTATATT
TTTGGATGCATGACCAATTTATAANGAAACCCCATATACGTNCATTTTCGCTTTGTTTTGGNGNTATTTCGGCAATTGGG
TAGGCAAACAGTGNNNNNNNNCCGGGGTGGCGCCAGAACTTTAATACTAGAGACTTTGCGGAGCTTTACAATCTTGGAT
TACCGGTATCTGCCGTCTATTTTAACTGCCAAAGGGAGAGCGGCTCTGGAGGGAGGAGAAGA**TAA**TTAAGTTCGATAT
TATTAAGCATGTTATACTTGTAATATTATTGACGAAGCTAGTTAATCTATATATAACGTGGAAACTAATAGTAATAATA
ATTTAGCTAGCTAGTGGTAATCAATCTCCTACTAGCTAGCTAGTCCATGCCATGGCACTAGGGCACTAGGGCACCAGGG
CACAGCATGTTGTATGCTTGTTTGAGACTACATCCGTACGTCTCCACTTTCAAATTAACAAGGCTAGCTATATATATATA
TATATAATAATAATAATTAAGTAAATGTCAAAGTTCTCTTCTTCGGTCGAAGGGGATGTCGACTCCATATAAATACAAA
GTGTCGATCAAATATTCATATTATATTATATTTTCCATATTTATTATAGAGATCAATTCAATTGTCAAATGATGAATTAT
TATGTACTTTCCATTCCATGCATGTAACATGTTGTCGTCTAGTGCTACTCAGACGACATTGCGGTCAGCTCAGGCATCAT
TTCTGCNAACTTCTGCACTATAATTCCCACTAATTCTCACATCCCATTAGTTTGACTAACACCAACAAAATATGTTTTTG
GTCATACCTAGCTAATAATAATAAACACGCAATTACGAATCTGGCAAATTATTCAAATATATATATATATATATATATT
TACACAATAAATATATGTTTTTTCAGTATGGCATTCCGTACAGTTCACCTACCGTCTATACATGTTGTAAAAGAATCCATG
CAAAATAAATAAACTATAACTATATGTAATCTATATATATAAAGTAAAAGACAGAGAATGGTGAAACATTTAAAATAC
CAAAAAATGTCCTTGGTTAATGCAAACATTAAGAATTGAAATTATTAATTAAATGAGGATAATATGGTAAATTCACAAT
TTTTCGTATTAAAAAAATTAAAATTAAAAGAAAATTCAGATAATGGGTCCTATTTTTATGGAACACAAATATCCATTCT
CTTTTTTAATTCTAAAATAAATTTAAATTTTTTTAAAAAAAATCTCTCGCATGCGCGGAAGCGCGTGCAGAGAGGCTAG
TATATATTAATTAAGGATTAATTACAGTTTAGTACTATAGGGTTACACCTTTAAGACATGTTAGTCCCTATCTTTTCAAT
TTTAACAATCACATATCCTGATTTTTTAAATTTGTTATAATGTAGTTTAACCGTTAGGTTTCCGTCAGATTTCTCTTTTTA
GTTGTCTGCCATTATGGGTCTCACATTTTTCAATTTTTATTACTTTTAATAAATGGATTAAAATTAAATATTTAATTTATT
TTTTTTTTAAAAAAAATCTCTCTCTCTCTCTCTCTCTCTCATGCGATTTCTTCTCCCTCCACCGCTTGTGATTTTCTAAATT
TACATATTTTAATCTATTAGTTTATATGACAACACAAAAATTAAAAGATTCTCCATTTTCCTTCACTTTCTCACCAACCA
AACACTGACTTTGCAAAAATGAAGAAACGGAAAAAATTGAGTTCGGGGTTTGGGGCTCTCGATTTCTTTCCCTCGCCCA
AATTTCTTCCTTCCCCCCGAATTTCTCCCCCCACCCAAATTTCTTCTCCCCTTACCCTCCGATTTCTTCTCCTCATCAAAC
AATCAACATCCAAGCCACAATCAAAACAGAGGAAGTGGGGAAGAAATCGGCTACTGTGATTTCAAAACTGGTTATTTG
ATCAGATCTTTATTGGTAATTGGAAGGGGGAAGAAATCGGAGGAATGAGGAGGGGGAAGAAATTGGAAGGAATGAGG
GGGAGGANGTCNAAGTANTTGCGGGNNGNGGTTTGGGTTCGTGTGNTNNNNGTTGNNNGCTCGGG

Appendix L: Map-based cloning of *FLOWERING LOCUS T* (*FT*) from the Goldrich cultivar apricot BAC library.

> *TFL* sequence from apricot Goldrich cultivar

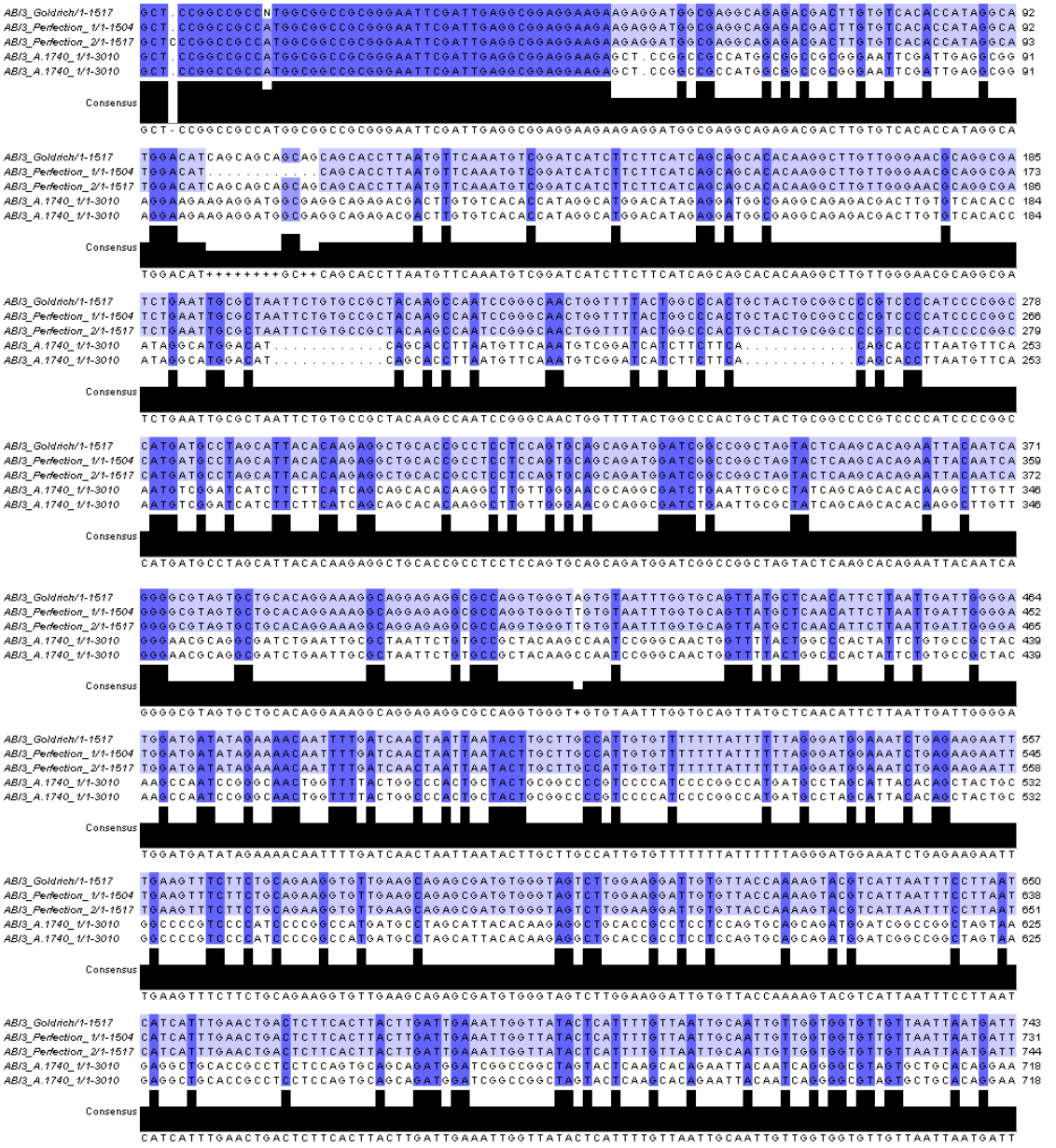
GTCCAGCTTGNANNNNCNGAANNAAAAAAAAGTAGTNGGNTNNTNNTATGGAAANTCNCCNNNTTNNCATNNCGNTN
CNGGAACTNNTNTGAAAATCACTTTATTTNATCCGAGTTATATATATGAGATGNNCTNTCNCANCAAAANCAGTATANC
AGCNNNCCAGGCACTCATAGAGGTGGAGTTCCTTTATGCCCTCAGCCNGTTTTCAAAGATAGAGANTCTCTNTGAAAAT
GTTGAAATTCCCTCTGTTTCNGTATTTATATATGCACNCATTCAACATTTTCTTCTGATTACCCAAACTGCAAAAGATGA
ATTTCCTTTATAAGCTTGTGAATGTTCAGCAAGAACCATAGACATTTTGCTGTTTTTGCTGTGTATTGCTTTTTATATATA
TATACATATATAATAGCATGAATTGACAGCCATGGAAACGTAAGGACCCCACCGTGGCATGCAGGATTGTATATAATA
ATGGTTTTCTGGCCTGACCATGGAAAAACATCCCCTCGTTTTGAATAAAAATGTTCAATAATTAAGTGCACAGCCAATC
AAATCAAGAACCAATTCCAAGGGACCCAATTAATATGAGAGGCACAGATATATAACATAAGTATATTATGCATATTAT
CCTAAATAAGCAAAAATTATTGCATACACCACACAGTTAATCATAGAAAGCGTGTTAATTAGTAAAAATCTCAAAGTC
GTCATAAGAGCACATAATTCAACTTATTATGTTTGATTTTGATATGGCAAATCATTACAAGAAAAATTGGCTAGTGTAA
TGAAATTAAAAGTGACGTTTGTAAATTCGTCACATGTGGTTGGAATATATGAGGAATTTATTTGGTCGTCACTTATAGA
ATGACTAAGAAAACAACAGTTTGTGAAAAAGACCAAAGTCGTCACAAACATGAGCGCGATTCTAGGTGGGAAAGACT
AACAAGTGTTTGTGATGCCAATAGAGTCACACATAAAATTCTAATGCGTCACAAATATTACATGTTGGAATCTCTAGTG
AAAAAAGAGAATACAAAAAGAANGAGAGGAAAANTTCCCTNCCCNAAAGAAAAAATATATATTTCNGGANTTCTTTN
GGTCATCGCATATAATNAATTATGACAAATTATGTGAGGATNGGGCCTCACTTATTCACATCTAATATTGANGAAACCN
NGTTATAGAATTAATACAACCGTNANACATACAGGAATTTTTCTGGAAACGATACCACTAAACTCTGTTGATGTGACAT
AGAAAAGAGAGGGTGGGAGGTGGGGGACTGGCTAACAATATTTTCTAGAGATGTCCCTATTATTCTATGCCTTTATACT
CTTTCTTCTTCTTAATGCCTGTCTTAAAACACATAATAGAGATCTAAAGTAACTCGCTTGAAAACACTATAAATATAACC
TTACAGGAATTCCACTACCAAGCAACATAAGAAGTACATCTCTAATTCCTCCTGAGTTCTTATCTGTTAATCTTTGATCA
ATAGCTACACTTTCTTTCTCACTCCTCATTCTCTCTGTCTCAAAAA**ATG**GCAAGAATGTCTGAGCCTTTGGTTGTTGGGA
GAGTGATAGGAGATGTTCTTGATTGTTTCACCCCAACAACAAAAATGTCTGTCACTTACAACACCAGGCTAGTCTGCAA
TGNNNATGAGCTGTATCCTTCTGCTGTCACCACCAAACCTAGAGTTGAGATTCAAGGAGGAGATATGAGAACTTTCTTT
ACTCTGGTATATATATCTATACAAATTTCCAATATTTCTCCAGACTCCCACTTCAATTAATTAGATTTTGTTTCTCTATGA
GTCTCCATTATCGGATCAATTTGATTTTACTTTGCATCTCTTGTGATATAGTTTTCTTGTTATTGCAGATCATGACAGACC
CTGATGTTCCTGGCCCTAGTGATCCTTATTTAAGGGAGCACCTGCACTGGTATACTTATTAACCCACTATCAAATTAACC
TAGCAACAAACTAATCAATTAAAAGACAAATAAATATTCACTTAATTGATTTCATGAGCATGCTACGCTAGAGCAGATG
TACTCCCCCTTTTGCCCCCCTCCTCTCACCCCCACACGTGCTGGTAGTTTAGATTAGATTAAAGTAGAACATATNNAACC
CATAAAACAATGGTCCAGATACAAGCACCTAACAGTCCTTCCGTACTTGTTAAGCATAAAAAATGACTTCAGTCAAAGC
TTCTAATAGTTTGTAAAACATTCATTGTATACCTCTACATAAACCTTGTTTTTGTTTGCGTTTCTCAATTATTTGTTCATA
AATTATATTCAGAATTTTCAAATTTCATCTAATATATATAAATATATGAACTGATACCTAGCTAGAATAGCANGCTAAA
AGTTGTCTAAGTATAAAGATACCTAGCTAGGAACTGACAAGTTTGCTTTAATTTTGAACTAAACCCAGGATTGTGACAG
ATATTCCAGGCACCACAGATGCCACATTTGGTAAGTCAAATAATGTTTTCTTTTTTGATTATTTTCCCCATTTTAATTAAC
TTAACTGTGGTTAGATTAATTGTAAATTAATTTNACAGGAAGAGAGGTGGTGAGCTATGAGATGCCAAGGCCCAACATT
GGCATCCACAGGTTTGTGTTTGTTCTCTTCAAGCAGAAAAGAAGGCAGTCTGTGAACCCTCCTTCCTCAAGGGATCATT
TCAGTGCTCGAAGCTTCGCAGCTGAAAACGACCTGGATCTTCCTGTCGCTGCCGTTTACTTCAATTGCCAGAGAGAAAC
GGCAGCTAGAAGACGC**TAG**CTAGCTACCCAGAACCCCAAAGCTCCTCCATAATATGTTAATTTTAAATAAAATTATTAT
CAAGTGTGTTTCATCATCCTCCCTTGTCGTTAGAGTTGTATTAGGCTAAAACTACTCACATGTAACCAGAATAATTTCCA
GTCAAGAGAGAGCAAGGTGCGTTCGGCTCTGTCATTTTGCTGCTCTCTGATCGGACTTTATGAATTAATTTGCAGAAAG
TTTGAAGTCTCACCTGACCATCTGTGAAATATCTCTCTATGGGGTTTTTAATTTATATTTTATTATAACGGGACAGGTGG
AACCAAACGTCCCACTTTTTGTCCTTTTTAAGTTAGGGTTTTGGCTTCCATTTCGACTGGAAAAGTAGAAGACCCTCTAT
TATTGACTTCATCTCTTCTTTCCAGTTTAGGGTTTTGGGTTGCACCACGTTTCAAGCTCAGAAATATGTGCGATTTGACA
GCAGTCAGTCACACCACATCCAGCTATGCCGACGCTTATAATAAAGAGGCAACTAGTATTTTTTTGCTTTTGAGAACNA
ANAAACTAAANCCTAAAACNTCTGTCANTTGATAGGAAAACAGCCTCAAATTTTATGCTTATGAACCCTAATATTATAT
GAGTTATTAAATTATTGANCTTTTCNACCNTGAATTTATGGGTTTATTTATTTCCCCTCTTCATTTTNNACATGGGTTTAA
TTTTCCCAATTATGGANNCNTGNCTATTTNAGTNNAACAAAAACGATAGTTAACGGGNTCNAGNNNNNTGAAAAATGA
TNTNAATNNCTA

Appendix M: Map-based cloning of *TERMINAL FLOWER* (*TFL*) from the Goldrich cultivar apricot BAC library.

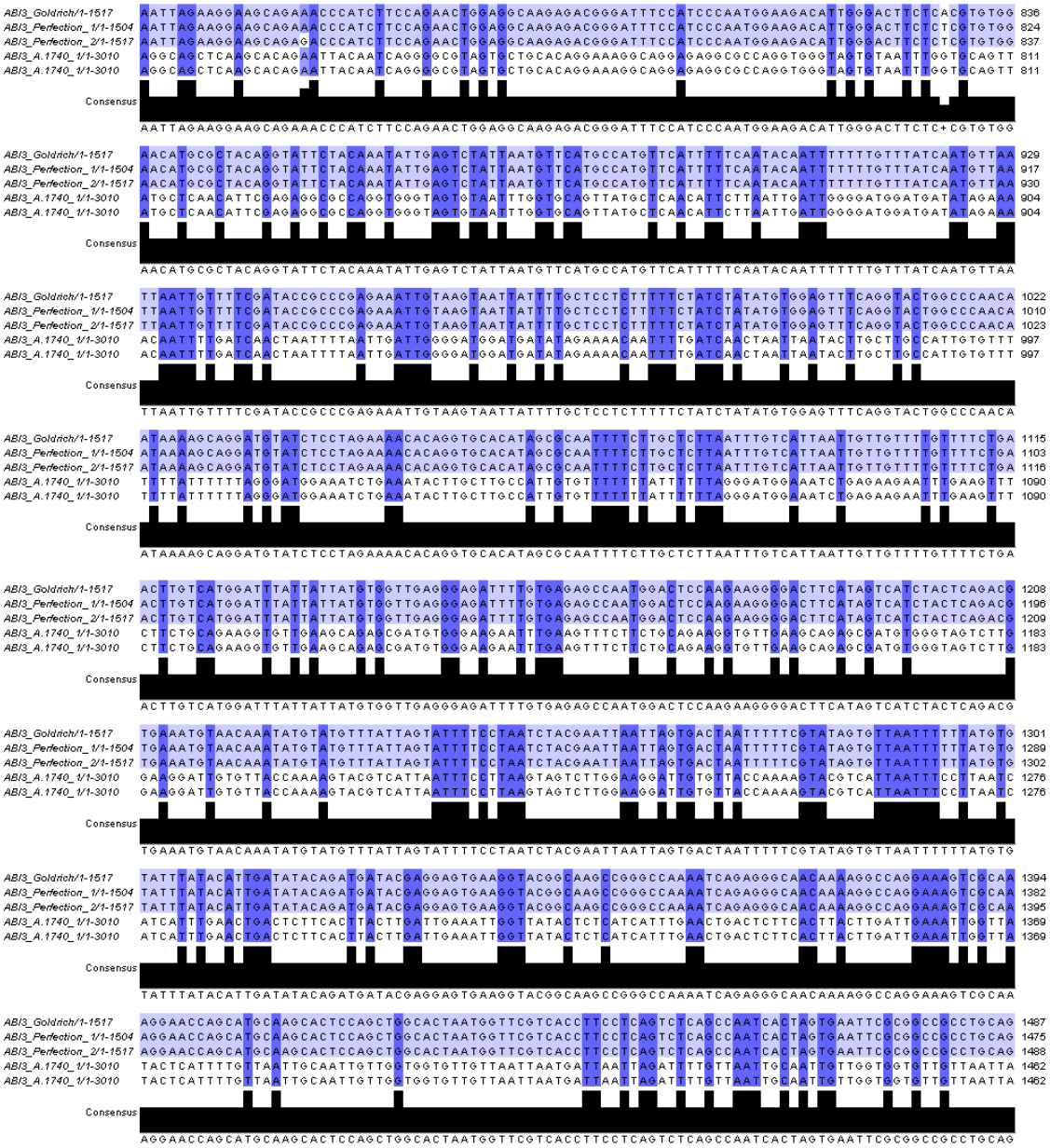
> *MPK7* sequence from apricot Goldrich cultivar; STOP CODON within 2nd gap (971 bps estimated from mRNA).

AAGCTCAAGTGGGACCCCCCACCCACCGAGCAAGCTCAAGCTAGCAAAGGCAACCGGCCTCGGGCAGCTGCCACTTTA
CACGTGGCCCGCCGCTTGCTTAAGGGAGGAGAAGTGAGCATATCATGGGAATGTTTCTTTGAAAGATTCACACGAGAG
AGAGAGAGAGAGAGAGTGAATGAATGAATTGTAGGGAAAGCAAAGGCAAACTTTCAACCTTGCCATATCTTCAAGTTG
GACTGCCACCCCTAGGGTTTTTTAGGCTTTACATTCCCAGCTGGCTAACTTAGAGCTCCCACCCACCATAAACTTTACAC
ATAGTAATATACATCAGCATTATAAAATCACTAAATCATTCCTTTTTTGTTACACATAAAATAACAAAAAGAGTGTGAG
CCATCTAGTTAACCTAATGGTTAATCTAAAGATCCTTAGATTCGGAGAGCCAAATTTTAAAACTTAATAGCAAATGACT
TTATGATTTTTTTTTCTAAGTGTTAAAAGGGTTGTAAAAGAGGTGGGAGCCAAGGCCACCTTGCACCTGGCTCTAGTTAT
GTCCGTGCTAATGGTATTTCTTTTCTGTTCAAGTTCGAAATTCCCGTATACTATCAATGTATAATCCAATATATAATTCAT
AAAAATTAAGCTATAATTACTATTAATATTTCTAAAAATAAATATTTTAGATCTCTTAACTATAATTACCATATTAAATT
TTGGTGATAATACTTTAAAATTAGAAAATTATTGTGCCGTTGAAGACATAAATTTCCAAATAAGTAAGTGTTGAAAAAA
GGTATATACATGAGTGAGAGAGACTTTCCCTGTTCCTTGTGATAAGTGATTGACTGATAAATAGTCAGCCAACCAAACC
AAGCCGAGCTGACAAGGAGCCGCCGTCGGCAGGCTGCGTTGGCGTTGGCGCTGAGCGTGAGCGTGTTTGTTTCTTTCAT
TTCCATGATTTGGTTGTGTTGATTGCACATGAAAGAAGCAAGAAAACGAAGACCGAAAGAGAGATGGATGTTAAGAAG
AGCTGTCGTTTCTGAGCAGTTTTTGTACCGAAATTTGCCGTCGTGTCATGGGTTTGGCTCATTCTCATTTCTCTGCCCCAT
TCCCCCAAACCCAAAACCAACCTTTCGTTTCTTCTCTATATAATATAACCCATGAGCCTCTCTCTCTCTCTCTCTCTCTAT
AGTATTTGCGTGCCATGTAGCGTGTCGCTAAGATCTCTCACTGATCACGACGAGATAACGAATCATAACTCGTGATGAT
CTGATCACTACGAGCTTTTGAGCCATTTGCTTTTGAGCCATTTATATCTATATATATTNATAGATAAAGGAGGGATCTGC
ATAACTACATGCCACTGTCTCCACGACCNNAGAnnnnnnnnnnCTCTTCTGGGTTTTTTTTCTTCTTCTTCAATTTTGTGTAA
AGTCGTAGAATNGAGCTGCGAATTCATGAATTGGGTATTAGGAAGTGAATTTGGGAGCTTCTTGTTGGGAAAAACCTTG
CTTTCATAGCTTAGCTTTGTTCAATTTTGAGTTTGAAGGCATGAAACAGATAGTGTAAAAGGTGGAGAGTTTTTGGTGT
GAATGGGATTTGAGAATTGCTGGGTCGACAAGGAAAAGAAAAAAAAAAAGGTTTAGGAGTGGACAGAGTGGGGTTTT
AAGAAGAAAAAGAGAAAAGCCTAAAGGGGAAGTGAAATTTTTTTATAAATATAAAAAAGGTAGGAAAAAGGCAACGA
GAATAACTCAAACAGGGAAGAGGGTGACCGGAAAAGATCCAAGCCCAGAAGTGGTAAGGGCAAGTAAGGATTGATTC
CATTCTAAAGCCTCCTTCTTCACACACGCAAATACACCAAGACGTATTGTTTTCTAGTTTTGTATTATTTTCATCTAGTGC
ATGTCATGCACTTCAACTTTTGTCACTTGTTTATGTTCCTTGTCTAAATTGTTGTTCCAACCAGCTTCACCAGTAAGTTAA
TATGGTAATTCATGTTGTTCTTATTATATTTCATTGCCTGAAGGTCACTTAAGAAGCATTTGTGGTTTTAAGTTATACTGG
GTATTGATTCATTTAGACCATTTTATGGTGATTTGCANAAGAAAAAAC**ATG**GCAACTTTAGTTGAGCCTCCAGATGGAA
TTAGGCAACGGGGGAAGCATTATTACTCAATGTGGCAAACATTGTTTGANGTTGACACCAAGTACGTTCCGATCAAACC
CATAGGGCGAGGANCATATGGNATAGTGTGCTCATCNNTCAATAGGGCAACNAATGANAAAGTTGNAATCAAGAANA
TCNATANTGTGTTTGANAACCNAATCGANGCNTTGAGGANTCTGANGGANTTGAAGCTTCTNAGGCATATCCNGNACN
ANAATGTGATTGCTTTNAAGNNNNNTATGANGCCNATCCNCNGGANNAGTTTCANGGATGNGNANTTTGNTTATGnnnn
nnnnnnAGTGTACGACTATGGTTCTGTATTATCGATAATGCATGTAAGATATTTCCTAAGAAAAACCAAGGGAATTGAAT
AAAGATGCATCTATTTTTATTATTGTTATTTTCTTCTCTCTTTTTTTCTTTTCTTGATTTCTCTACTCAATCTTAATAGGTA
GATGCATGGGTCACTTACTCAAGCATATTGTATATGTTGAAACGTTTGCTGGTCTGGTTTACCTGAAAGGAAGTTATAC
AAATCTTTCCTAAGTACTTTCACTTTATGATTTGGGCCATATGCTTAATAGTTCAATGCCCATCAGAAGTAAGAAGGCCT
ATCCTAGGGCCAGCCCAAGTGGTTCAAGAGTCATTATATAGTGAGGGCCTTACATATGGCTCTTAGGTAAGGCCGTATC
TCTTCATTGTTAGATCAGATTAGCTAATATAATCTAACGGTTAACCAATTTGATCAACAAAAAAAGAGATTATTATCAA
AGCTGCTACATGGAACGGAATGGAATGCTGCTTTTTAGGAAAGCACGCACCCTATATCCTACTCTTAAAAAGCATGTCA
CTTCCAAAAGTTCTATTAATAAAACAGAGCACACAATTTGGGACCCAAGACCCGATATGCGCCCAACTTTTCTTGACCC
CCAATACCCACTAACAGCCCACCACTGCCCACTGCTCTTGCCTGGTTTCCTTTTGCCTTCCCTCCCTTCTTTTTCGAATAA
CTGGGTCCAAGCCTCCAAGGCAACTTGCAAGACCGTGTGCTTTTTTAAGGCACAGAAAAAAGTTGCTTCAGATTTCATT
GGTTTTCATATATCTGCTTGCTTGTGTTCACTGTTCAATACGCTTGAAAAAGAAC

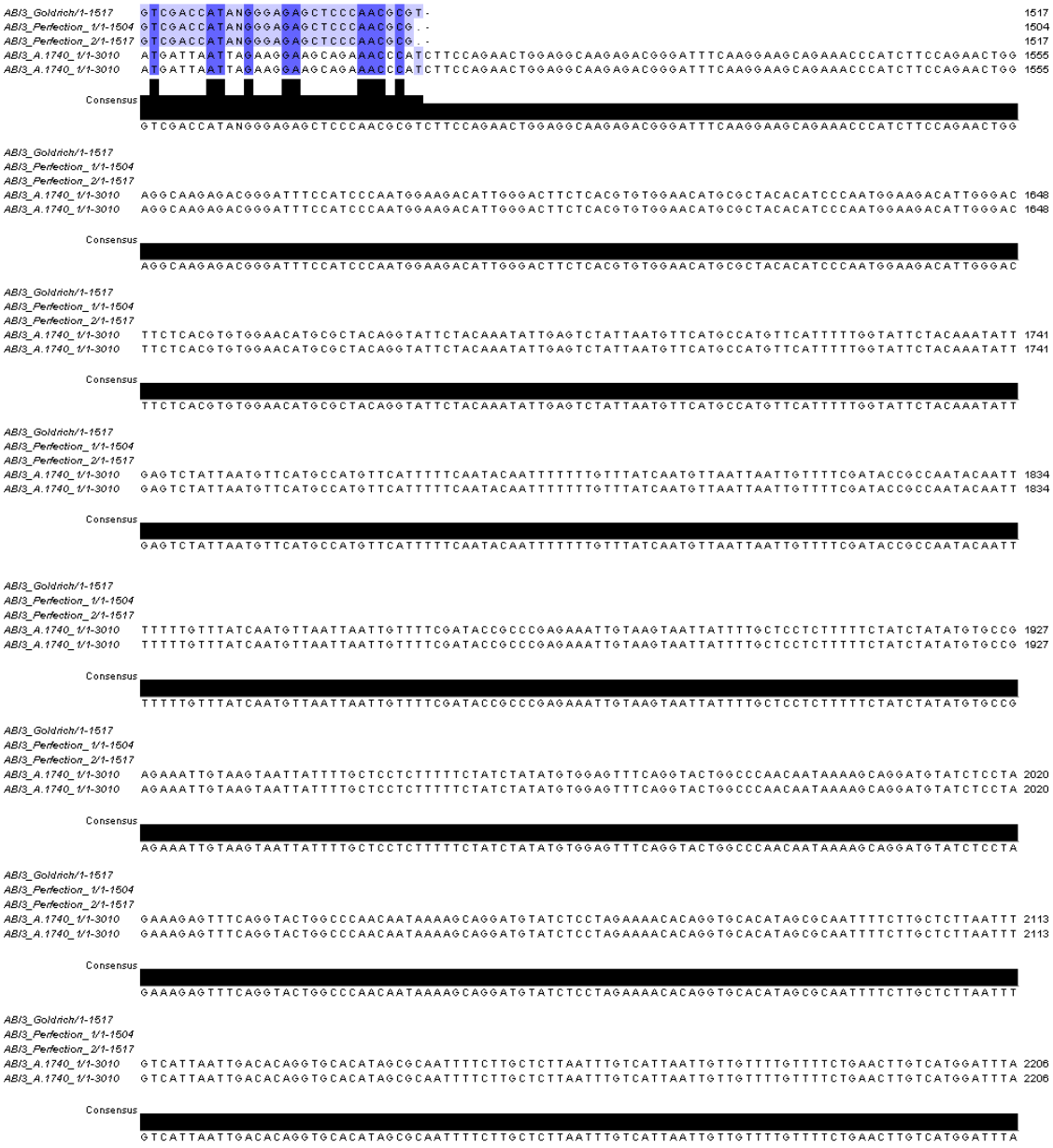
Appendix N: Map-based cloning of candidate genes for *MITOGEN ACTIVATED PROTEIN KINASE 7* (*MPK7*) from the Goldrich cultivar apricot BAC library.



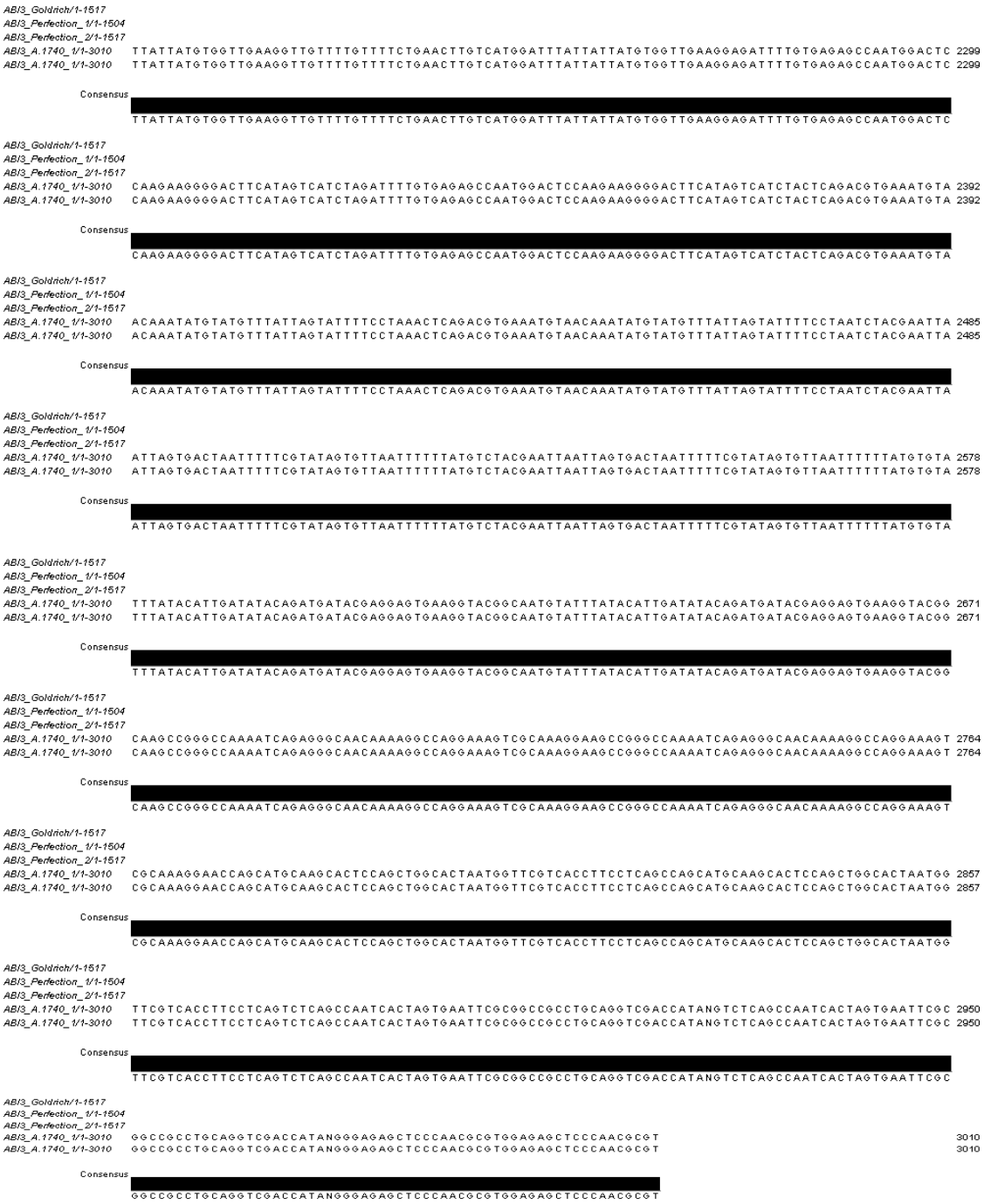
Appendix O1: Map-based cloning *ABSCISIC ACID–INSENSITIVE 3* (*ABI3*) from the Goldrich cultivar apricot BAC library and the 2 allelic variants of Perfection and A.1740 apricot cultivars. Sequence alignment performed with MUSCLE v3.7.



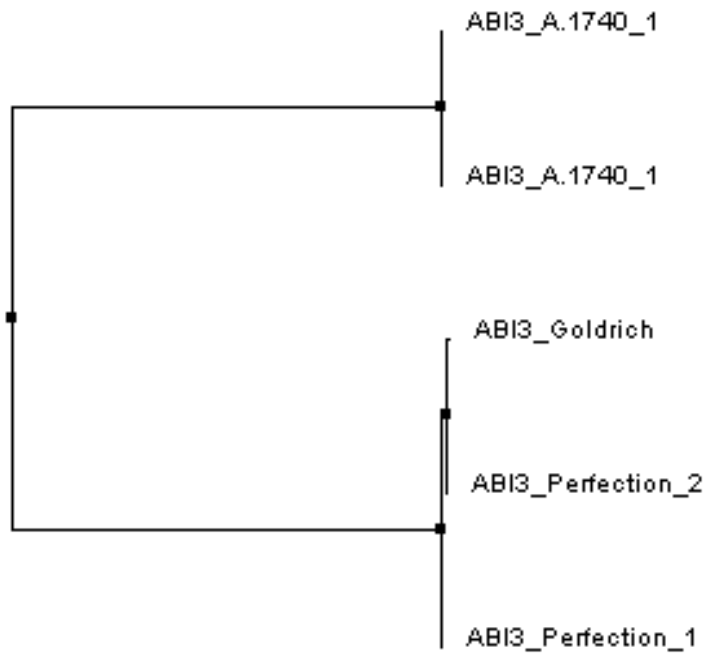
Appendix O1: continued



Appendix O1: continued



Appendix O1: continued



Appendix O2: Gene Phylogeny of *ABSCISIC ACID–INSENSITIVE 3* (*ABI3*) alleles from Goldrich cultivar apricot BAC library, 2 allelic variants of Perfection apricot cultivar and 2 allelic variants of A.1740 apricot cultivar. Phylogenetic analysis was based on Neighbour joining tree using PID, following alignment using MUSCLE v3.7.

APPENDIX P

COMPARATIVE AND FUNCTIONAL GENOMICS OF SYMBIOTIC NITROGEN FIXATION IN RHIZOBIUM-LEGUME AND FRANKIA-ACTINORHIZAL SYSTEMS

Bode A. Olukolu, Julia Frugoli, Amy Lawton-Rauh, Chin-Fu Chen, and Albert G. Abbott

Genetics and Biochemistry, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA.

Corresponding authors: A.G. Abbott (e-mail: aalbert@clemson.edu)

Project Summary

Nitrogen, the most limiting element for the nutritional needs of the ecosystem, is commonly assimilated as nitrate yet it is often a major source of groundwater and surface water contamination in the form of inorganic nitrate fertilizer. To solve the resulting environmental and human health problems estimated to cost many several billion dollars, interest has revolved around biological nitrogen fixation in plant-bacterium interactions. Rhizobium-Legume genes implicated in biological nitrogen fixation have been revealed by means of functional genomics and now provide the long-awaited insights into these complex interactions. Despite the vast amount of knowledge, which has provided a significant amount of our understanding about plant-microbe interactions, using the Rhizobium-Legume model systems still seem remote for utility in the genetic modification of commercial crops lacking the trait. With the advent of recent molecular -omics technologies and tools, organisms which in the past were considered intractable for basic research are now used for cutting edge studies because they now present unique opportunities to broaden and enhance our understanding of the evolution of various specialized biological systems and their divergent functions in different plants.

Here, we propose an approach to bridge the gap between plant-microbe symbiosis systems and commercial crops by implementing and integrating high throughput approaches including sequencing, transcriptomics, comparative genomics, phylogenomics and metabolic profiling. The experimental design exploits the recently evolved nitrogen-fixation symbiosis of the Frankia-Actinorhizal species. In view of the

recent emergence and diverse evolutionary lineages, the seemingly huge divergence from the ancestral Arbuscular Mycorrhizal symbiosis, which pervades most plant families, can now be tractable for in-depth scrutiny on a molecular scale. The project aims to dissect the molecular components of the poorly understood Frankia-Actinorhizal symbiotic association by:

1. Differential expression transcript profiling of Frankia and Cercocarpus based on pyrosequencing and cDNA-AFLP techniques and concurrently generating EST libraries of symbiont and host for further studies.

2. Annotation, Characterization, expression studies of the transcripts and utilizing transgene strategies to understand specific gene pathways.

3. Comparative studies of gene structure and expression patterns to delineate the evolution and biological uniqueness of the trait. This will enable simulation of evolutionary trends across both nodulating systems and their closely related non-nodulating Rosaceae species.

4. Integration of the differentially generated transcripts and metabolites to enhance understanding the roles these genes and associated metabolites within the signal transduction pathways of nodulation.

The feasibility of transferring this trait from a model organism to closely related crops is strongly supported by preliminary studies which vary from *in silico* to experimental studies. These studies range from the heritable vascular uninfected nodules in Apple

(*Malus x domestica)* to high identity matches of nodulation gene sequences lacking in other distantly related species. Ultimately, the novel approach holds promise to elucidate the molecular mechanisms conferring plant-microbe interactions within an evolutionary context of symbiotic nitrogen-fixation

Introduction

Symbiotic associations between microbes and plants have received much attention due to their effects on plant morphogenesis, nutrition, infectious diseases and understanding of basic cell biology. These associations ranging from parasitic to mutual symbiotic interactions are critical to food security and nutritional needs of the ecosystem (Vance *et al*. 2002, Verghese and Misra 2002). These essentials rely heavily on the availability of fixed nitrogen, which is the most limiting element for the synthesis of proteins, amino acids, nucleotides and vitamins. Biological nitrogen fixation is a very cheap and sustainable source of soil nitrates unlike fertilizer application. Furthermore, fertilizer applications create a fully nitrogen-saturated ecosystem, which leads to several harmful consequences to the functioning of the ecosystem. The ensuing negatively charged nitrates (from ammonium build up) carry along with them positively charged alkaline minerals into groundwater, surface water and the atmosphere. Besides leaching of nutrients, the acidified soil leads to the mobilization of Aluminum ions with accumulations reaching toxic concentrations that damage plant roots and eventually top soil structure degradation (Aber 1992). The two major plant-bacterium systems involved in endosymbiotic nitrogen-fixation include Rhizobium-Legume and Frankia-Actinorhizal interactions. Based on phylogenetic analysis using *rbcl* chloroplast gene sequences, the Legumes and Actinorhizal plant families belong to the same Rosid I clade, suggesting that a genetic tendency to form root-nodule symbioses originated in a common ancestor (Soltis *et al.* 1995, Doyle 1998, Hocher *et al*. 2006).

Understanding the basis of symbiotic relationship is crucial to unraveling the persistent environmental and resource conundrums experienced before and even more intensified after the green revolution. We propose a concerted and coordinated genomic analysis of the symbionts and hosts involved in symbiotic nitrogen fixation. Genome analysis, in this context, refers to the structural and functional analysis of the DNA including the genes, expressed proteins, the metabolites involved in the signal transduction pathways, as well as non-coding sequences involved in genome dynamics and function. The study will complement the extensive databases and well studied Rhizobium-Legume symbiosis with the molecular studies of the basal clade of the Actinorhizal genera (Cercocarpus, Chamaebatia, Cowania, Dryas, Purshia) in the Rosaceae family (Potter *et al*. 2002). This study of the Rosaceae Actinorhizal symbiosis will broaden our knowledge of symbiotic nitrogen-fixation beyond the Rhizobium-Legume symbiosis, while providing novel insight to the biology of Actinorhizal species.

Present State of Knowledge

The genus Frankia, formerly considered a filamentous fungus due to its hyphal morphology, was revised and renamed as an Actinobacteria in the family Frankiaceae (Becking 1970). They consist of Gram-positive bacteria, which nodulate 8 plant families (Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Elaeagnaceae, Myricaceae, Rhamnaceae and Rosaceae) comprising a total of 25 genera (woody, dicotyledonous, perennial angiosperms). They are mostly pioneer species in nitrogen-deficient soils,

hence establishing and sustaining forests (Silvester, 1976). Economic importance includes use in regeneration of waste lands, as windbreaks, pulpwood, fuel and timber wood, in human diet, as forage for livestock, ornamental and horticultural applications (Silvester 1976, Verghese and Misra 2002).

Though there are many similarities between the Rhizobium-Legume and Frankia-Actinorhizal systems demonstrated within their application and functional context, questions abound about differences that hold promising insights into novel mechanism in plant-microbial interactions. Some of these questions include the incongruence in range of host plant family, highly variable growth rates and taxonomically diverse symbionts that retain similar quantum of fixed nitrogen (Verghese and Misra 2002). A major advance in Frankia-Actinorhizal molecular biology will require a comprehensive approach to dissecting genome-wide machineries exclusive to each and common to both host-symbiont systems.

Nodules are commonly modified lateral roots and are mostly initiated via root hair infection in most Legumes and some Actinorhizal species (*Alnus, Casuarina, Comptonia, Myrica* (Torrey 1976, Callaham 1979, Berry 1983). Root hairs curl around and entirely encapsulate the bacteria that divide to form a micro-colony and migrate down an infection thread (initiated by plant host) into the inner cortical cells (Geurts *et al.* 2005, Riely *et al.* 2006, Stacey *et al.* 2006). An alternative approach to circumvent inhibition of root hair curling by ethylene in the aquatic Legume *Sesbania rostrata* allows bacterial (*Azorhizobium caulinodans*) invasion on submerged roots and stem via intercellular entry

(D'Haeze *et al*. 2003, Den Herder 2007). Actinorhizal species such *Ceanothus, Elaeagnus and Shepherdia* also utilize this mode of infection via direct intercellular penetration of root epidermis cells and cortex, but unlike in the Rhizobia counterpart (Den Herder 2007); Frankia does not require gaps in root epidermis for entry (Miller and Baker 1985, Racette and Torrey 1989, Liu and Berry 1991, Vessey *et al*. 2005). The relationship culminates with a controlled exchange of fixed nitrogen and the plant carbon source required for the energy demanding process of nitrogen fixation in the symbiont. Unlike Rhizobia that utilizes simple sugars (mostly sucrose) from the photosynthetic plant host as its primary of carbon; Frankia strains are suggested to rather utilize lipids based on studies reporting lipid-rich Alnus nodules (Maudinas *et al.* 1982) and exponential growth rates of Frankia isolates in culture supplemented with long-chain fatty acids (Selim *et al*. 1996, Verghese and Misra 2002). Conversely, a report of high levels of sucrose synthase expression in *Alnus Glutinosa* nodules elicits questions about possible multiple carbon sources and preferences (Van Ghelue *et al*. 1996, Vessey 2005).

Genes involved in nodulation

Rapid increase in our understanding of molecular interactions between Legumes and Rhizobia elaborate a molecular dance in which plant and bacterial partners signal and respond to each other through a complex series of feedback loops resulting in differentiation and initiation of specialized root structures (nodules) to house the bacteria.

According to recent studies by Normand *et al*. (2007), BLAST Searches for nodulation gene homologs in Frankia have been further enhanced with the availability of 3

completely sequenced genomes of Frankia strains to date. So far, evidence for nod gene clusters has not been confirmed in any of the genomes. The *nodABC* genes in Rhizobium code for the nod factors (decorated lipo-chitin-oligosaccharide) elicited by plant flavonoids and which in turn trigger a cascade of signal transduction events in the plant host response for bacterial infection (Long 1984, Verghese and Misra, 2002). The *NodA* ortholog was completely absent from the available Frankia genome sequence though orthologs of the *NodB* and *NodC* were found but at low identity levels (32-48% and 24-43% respectively). Likewise, the *NodD* gene responsible for inducing nod factor genes and *NodO* gene were discovered with a low identity. These low levels of identity do not seem to support that the typical Rhizobium Nod factor genes are responsible for initiating early infection in Frankia since many organisms show similar low identity matches. Furthermore, the homologs lack the synteny observed in Rhizobia nod gene clusters. To further verify similarities in the early stage of nodule development and genetic components, multiple attempts at genetic complementation of Rhizobia nod mutants with Frankia DNA (Chen *et al*. 1991, Reddy *et al*. 1992, Ceremonie *et al*. 1998) and hybridization of both symbionts did not yield satisfactory results to implicate nod genes in Actrinorhizal symbiosis (Simonet *et al*. 1988, Verghese and Misra 2002). An alternative strategy has been suggested implicating the Frankia pel genes, which have a high sequence similarity with the pectate lyase gene from *Erwinia chrysanthemi* (plant pathogen) and is responsible for invasion of host cell by degradation of pectin (Seguin and Lalonde 1989, Verghese and Misra 2002).

Interestingly, Rhizobium strains (photosynthetic Bradyrhizobia strains: *BTail* and *ORS278*) forming functional nodules in *Aeschynomene sp.* have been reported to also utilize alternative signaling strategies to trigger nodule formation (Chaintreuil *et al*. 2001, Giraud and Fleischman 2004). Similar to Frankia, their genome sequences reveal the absence of symbiotic genes (nod, nif or fix) on islands and low levels of synteny. Homologs of *nodA* and nodC from BLAST searches yielded identity scores of 33-36% (Giraud *et al*. 2007). Transposon insertion mutagenesis libraries suggest that a purine derivative may be involved in initiating the symbiosis (Giraud *et al*. 2007).

Downstream of the initial signal, genes with high sequence similarity have been implicated in functional conservation though phylogenetics studies propose a lateral gene transfer from Rhizobia. The uptake hydrogenases (hup) gene conserved across nitrogen-fixing bacteria species are used for recovering loss of energy by scavenging for hydrogen generated during the nitrogen fixation process is an example of a gene suggestive of lateral gene transfer (Leul *et al*. 2007). The hup genes are reported in all 3 sequenced Frankia genomes (Normand *et al*. 2007) and have been confirmed to hybridize with the Bradyrhizobium hup gene (Verghese and Misra 2002).

Regulation of symbiosis

To mutually optimize benefits, interactions between host and symbiont are delicately synchronized to levels where the genetics and physiology of both organism appear to function nearly as one, however, the exact nature of controls involved are not very clear (Verghese and Misa 2002, Vessey 2005). At a glance though, expression patterns evocate

three factors comprising the host, symbiont and environmental influence (Verghese and Misa 2002).

Many investigations suggest a dominant though not exclusive role played by the host regulation in recognition and selection of strains at infection; nodule development; selective suppression of host defense genes leading to growth of symbiont while other strains or pathogens are constrained; controlling leghaemoglobin levels, thickness of vesicle walls and host cells barriers for proper functioning of the bacterial nitrogenase; regulating metabolism in nodules by altering demands for fixed nitrogen, export of available ATP for synthesis and regulating the amount of carbon compounds available to the microsymbiont (Verghese and Misra 2002). These levels of control have been elucidated in variable nitrogenase activities (nitrogen fixation rates) of Frankia in pure culture and in situ conditions for various strains found in different nodule morphologies (Reddell and Bowen 1985, Dawson and Sun 1981, Sougoufara *et al.* 1992, Verghese and Misra 2002)

Evolution of nodulation

Unlike the relatively poorly understood Frankia-Actinorhizal symbiosis; the complex biology, genetic components and regulatory pathways of Rhizobium-Legume symbiosis is one of the best-studied interactions between prokaryotes and eukaryotes. However, little progress has been made towards transferring this knowledge to benefit or induce this trait in non-nodulating species, even though this goal has been a priority. Credit has been given to the puzzling and complex network of feedback regulations of the symbiosis

and the lack of multiple genetic components in non-nitrogen fixing species. Since studies indicate that the nodulation pathway evolved from the more ancient Arbuscular Mycorrhizal symbioses found in most plant species (Duc *et al.* 1989, Kistner & Parniske 2002), it is only rational to anticipate that the evolutionary divergence of nodulating and non-nodulating species can be bridged.

It is striking that the single Legume family contains both perennials and annuals that form symbiotic root nodules, while the more taxonomically diverse Actinorhizals known till date are all perennials. This raises questions about a possible evolutionary significance. Annuals are typically known to have evolved from perennial ancestors (Laroche and Bousquet 1999) and have higher rates of evolution (Eyre-Walker and Gaut 1997, Muse 2000). This trend correlates with phylogenetic inferences from molecular data rooted with characterized Legume fossils (Lavin *et al.* 2005), which reveals the order in which nodulation evolved among Legume subfamilies i.e. Caesalpinoideae (most primitive, mostly perennials), Mimosoideae (mostly perennials) and Papilionoideae (more annuals) are comprised of rare, common and very common nodulating member species respectively (Allen and Allen 1981, Sprent and James 2007). It appears Legume annuals have evolved at a faster rate considering that their Rhizobia symbionts have a very narrow host range, while the tree species studied to date show a wide range in specificity and are generally promiscuous in the Rhizobia with which they nodulate (Batzli *et al*. 1992, Odee *et al*. 1995, Sprent and Parsons 2000). Indeed, most of the Rhizobial symbionts described have been isolated from tropical/sub-tropical Legume trees (Moreiro and Franco 1994, Sprent and Parsons 2000). Simonet *et al.* (1999) suggests an initial

promiscuous state in which any species in the family nodulate indifferently by a variety of symbionts but some associations might have evolved to an intermediate state in which selection favored particular associations for higher efficiency.

Similar to the seemingly slower mutation rates in tree Legumes, some Actinorhizal species are known to have an even broader range of host specificity (Benson & Silvester 1993, Kohls *et al*. 1994, Vanden Heuvel *et al.* 2004) suggestive of a recent incidence of the trait.

Preliminary Research

Rosaceae genomic resources

Central to the preference for Rosaceous Actinorhizal species as taxa of choice for detailed study of Actinorhizal symbiotic nitrogen fixation is the available and extensive genomic resources in the Rosaceae community which is unparalleled by other Actinorhizal families. EST databases for Rosaceae species are highly enriched with genome-wide sequences across the family at the Genome Database for Rosaceae (GDR: www.mainlab.clemson.edu/gdr/). It includes annotated unigene data sets for Malus (82,850), Prunus (23,721), Fragaria (10,012), Rosa (2,963) and Pyrus (271). They represent 3 of the traditional 4 Rosaceae sub-families including Rosoideae (Strawberry, Raspberry, Rose); Maloideae (Apple, Pear); Amygdaloideae (Peach, Almond, Cherry, Apricot) and subfamily spiraeoideae (Schulze-Menz 1964). The classification of the

Actinorhizal genera within Rosaceae has been unclear and they were originally placed in the subfamily Rosoideae along with Strawberry but recently classified (Thorne and Reveal 2007) in the subfamily spiraeoideae and tribe Dryadeae, which includes Cercocarpus, Chamaebatia, Purshia and Dryas (Morgan *et al*. 1994, Takhajan 1997, Potter 1997, Evans *et al*. 2000, Potter *et al*. 2002). Besides these EST resources, extensive large insert BAC libraries are available in the Abbott laboratory and through collaborations. Available BAC libraries include Peach, Apricot, Plum, Cherry, Raspberry (Scottish Crops Research Institute), Strawberry (V. Shuleav, Virginia Tech, under construction in the Abbott lab), Rose (T. Debener, University of Hannover, Germany) and Apple (S. Korban). Furthermore, physical maps are been developed from this libraries for Peach (Abbott laboratory, in completion), Apple (Korban laboratory in progress), Strawberry and raspberry (Abbott *et al*. 2006, USDA grant proposal).

Data mining for Rosaceae and Frankia orthologs of nodulation genes

The extensive genomic resources provide an exceptional opportunity to study the system biology associated with the Actinorhizal nitrogen fixation. In this regard, we mined the Rosaceae EST database for all known plant genes associated with nitrogen fixation in Legumes and have identified many of the major players with significant high similarities (Data available on request). Genes from several Legume species were used for the BLAST search and this include genes at all stages of nodulation (signaling, development and functional stage, Table P2 and P2).

Table P1: Genes Identified by mutation in model systems as being involved in nodulation.

Gene name	Gene product	Sequence orthologs in other species	EST hit in Rosacae with bit score > 60 and E value <-12
NFR1, NFR5, LYK3, LYK4, NFP, SYM2, SYM10	LysM receptor kinases	yes	*Prunus persica, Malus domestica, Pyrus pyrifolia, Pyrus communis and Rosa roxburghi*
DMI1, CASTOR, POLLUX	ion channels	yes	*Prunus persica, Malus domestica*
DMI2, NORK, SYMRK, SYM19	LRR receptor kinase	yes	*Prunus persica, Malus domestica, Pyrus pyrifolia,* Pyrus communis, *and Rosa roxburghi*
DMI3	Ca^{++}/Calmodulin receptor kinase	yes	*Malus domestica, Fragaria x ananassa, Prunus armeniaca and Rose hybrid cultivar Asami*
NSP1,NSP2	GRAS family putative transcription factors	yes	*Malus domestica,* Prunus persica and *Rosa roxburghi*
MtNIN, LjNIN. SYM35	putative transcription factor	yes	*Malus domestica*
MtCRE1, SNF2, HIT1	Cytokinine receptor	yes	?
Mt HAP-2	CCAAT-binding transcription factor	yes	?
SKL	*EIN2* (ethylene signaling) ortholog	yes	?
SUNN, HAR, NARK, SYM 29	*CLV1* like receptor kinase	yes	*Prunus persica, Malus domestica, Pyrus pyrifolia, Pyrus communis and Rosa roxburghi*
ASTRAY	BZIP transcription factor w/ RING finger domain	yes	?
LjnsRING	RING finger protein	yes	?

Note: *M. truncatula* gene names are in black, *Lotus japonicus* in blue, other legumes in green.

Table P2: Genes identified by differential expression in model systems as being involved in nodulation.

Gene symbol	Gene function	EST hit in Rosacae with bit score > 60 and E value <-12
ENOD11	Repetitive proline-rich protein	no significant hit in Rosaceae
ENOD12	Early nodulin (	no significant hit in Rosaceae
ENOD16	Early nodulin	no significant hit in Rosaceae
ENOD20	Early nodulin	no significant hit in Rosaceae
ENOD40	Small untranslated RNA	no significant hit in Rosaceae
GS1	Glutamine synthase	no significant hit in Rosaceae
MtN1,3 5,6,12, 13	Nodulins	significant hit, only *MtN13*
MtSucS1	Sucrose synthase	significant hit *found*
PR-1	Pathogenesis-related gene	Significant hit *found*
PRP4	Proline rich nodulin	no significant hit in Rosaceae
RIP1	Peroxidase	no significant hit in Rosaceae
LOXN2	Lipoxygenase	Significant hit found

Note: "nodulins" are genes up-regulated during nodulation- no function has been assigned to these genes and they occur in species outside legumes, including *Arabidopsis*.

Also, the new and completely sequenced Frankia genomes (*Frankia* strains: CcI3, Ean1pec and ACN14a) were examine for ORFs using gene prediction softwares, as well as annotated genes available on NCBI and TIGR.

Expression studies of Strawberry DMI3 homolog in Medicago

Our database searches identified a highly significant match in Strawberry (*FaCDPK*, Llop-Tous *et al*. 2002) as a potential ortholog for the DM13 gene in Medicago (Abbott *et al*. 2006, USDA grant proposal). Young leaf tissues were obtained from diploid Strawberry (Fragaria vesca; Accession PI: 551572) and total RNA was isolated according to modifications to the RNA miniprep protocol for peach tissue (Dr. Zhigang Li, Clemson University). The cDNA of the transcript was generated using RT-PCR with oligos designed from the Strawberry sequences to give an expected size product (Fig. P1). The cDNA was cloned into pUC19 and verified by sequencing. For expression studies in *Medicgo truncatula* (A17) via hairy root transformation, the gene construct in pCAMBIA2304 was designed to be driven by CaMV 35S promoter. To initiate spontaneous nodules, the *MAX17* gene was truncated by removal of the auto-inhibitory domain and EF hands, using appropriate restriction enzymes and ensuring that the ORF is in frame. The modification of the *DMI3* gene is reported in Gleason *et al*. (2006) to result in a constitutively active kinase that triggers spontaneous nodulation. Our construct also initiated spontaneous nodules in Medicago roots (unpublished).

>*Fragaria vesca* calcium-dependent protein kinase (*MAX17*) mRNA, complete cds
TTGGACTAATACACCGGTTTTGGGTAGATGGGTAATTGCTGTGTCACCCCTCCCCAGACGGGT
TCGCCGTTAAAGAACAAGAAGAATAAGCCAAACCCGTTTGCGATAGACTACGTTGTCGCCAA
TGGCGGCAAGCTCTCCGTTTTGAAGAACCCAACCGGCACTGAAATCGAGCAGACTTACGAGCT
GGGCCGCGAGCTCGGCCGCGGAGAGTTCGGGATTACGTATCTGTGTACTGACAAGGCCACCA
ACGAGAACTACGCTTGCAAATCGATATCGAAGCAGAAACTGAGGACGGCTGTGGATATTGAA
GATGTGAGGAGGGAAGTTGAGATCATGAAGCACTTGCCTAAGCATCCCAATATTGTGAGCTTG
AAAGATACTTACGAGGATGATAATGCTGTCCATCTTGTTATGGAGCTCTGTGAGGGCGGTGAG
CTTTTTGATCGGATCGTGTCTAGGGGACATTACACTGAGCGTGCTGCTGCTGCTGTCACTAAG
ACTATTGTGGAAGTTGTTCAGATGTGCCACAAGCATGGTGTGATGCACCGGGATCTTAAACCT
GAGAACTTTTTGTTTGCAAACAAGAAAGAAACAGCGCCCTTGAAGGCAATTGATTTTGGGTTG
TCAGTGTTCTTTAAGCCTGGTGAAAGATTCAGTGAAATTGTTGGAAGTCCATACTACATGGCT
CCTGAGGTGCTAAGACGCAATTATGGTCCTGAAGTTGATGTGTGGAGTGCTGGAGTTATACTT
TACATCTTACTTTGTGGTGTTCCGCCTTTCTGGGCAGAAACTGAACAGGGAGTTGCACAAGCA
ATTATACGGTCTGTTGTAGATTTTAAGAGGGACCCCTGGCCTAAGGTTTCTGATAATGCAAAA
GACCTTGTGAAAAAGATGCTTGATCCTGACCCGAAGCGGAGGCTTACAGCTCAGCAAGTTCTA
GATCATACTTGGTTGCAAAATGCAAAGAGAGCTCCAAATGTTTCTTTAGGTGAAACAGTGAGA
GCAAGGCTCAAGCAGTTCTCTGTAATGAACAAGCTTAAGAAAAGTGCACTGAAGGTCATAGC
TGAGCATTTGTCACAGGAGGAAGTTGCTGGCATACAAGAGGGATTTAAGATCATGGATACTA
GCAATAAGGGCAAGATTAACATTGATGAGCTAAGAGTTGGGTTACATAAACTAGGCCATCAG
ATTCCTGATGCTGATGTTCATATCCTAATGGAAGCTGGTGATGTAGATAATGATGGGTATCTG
GACTATGGGGAGTTTGTTGCCATTTCTGTTCACCTAAGAAGGATGGGCAATGATGATGAGCAC
CTTCGCAAAGCTTTTGACTTCTTTGATCAAAACAAAAGTGGGTTCATTGAAGTCGAGGAGTTG
CGAACTGCCTTGGCTACTGAAGTTGATGACCACGTTGAAGATGTTATTAGTGCCATTATCAGT
GACGTGGATACAGACAAGGATGGAAAAATAAGTTACGAGGAGTTTGCCACCATGATGAAGGC
CGGCACAGATTGGAGAAAGGCCTCAAGGCAGTATTCACGAGAGCGGTTCAATAGTCTCAGTT
TGAAATTGATGAGGGATGGATCATTGGAAGGTAAAACCGAGAGCAAATGACACATCATACAT
GTTAATGAAAGAATTGTTCATTTTTTGTTTGTGTTTTTGTAATTCTTTCTTTTGTAAGTTTTCTCT
GTTAATTTTACATCCTTTTGTAGACCCTTCTGTGATTATTAGGATATGAGCCAAGGGTTTTCTC
AT

Figure P1: Strawberry *MtDMI3* homolog (*MAX17*) cloned from cDNA reverse transcribed from mRNA obtained from strawberry leaf sample.

Studies reveal the rice *DMI3* ortholog initiates nodules but lack infection within the organ (Godfroy *et al*. 2006), while the Poplar DMI3 ortholog (Pers. Comm., Netherlands) only initiates infected nodules when driven by the Medicago *DMI3* promoter but not with its own native promoter. At present, studies are conducted on the full length *MAX17* gene under the expression of Medicago *DMI3* and its own native promoter to test for functional similarity.

Phylogeny study based on Frankia and Rhizobia genomes and genes

Rhizobia nodulation genes and annotated orthologs in Frankia are currently been used for phylogenetic inferences and sequence data includes homologs in Rhizobium, Bradyrhizobium, Sinorhizobium, Mesorhizobium and Frankia. Our preliminary *in silico* studies support suggestions from literature about lateral gene transfer between different symbiont genera (Bailly *et al.* 2007, Nandasena *et al.* 2007, Normand *et al*. 2007). From our study, Frankia is revealed to be more related to Bradyrhizobium than any other Rhizobia species, indicating that Bradyrhizobium might have played a major role in the evolution of nod genes with reference to bridging the Rhizobia and Frankia clades via lateral gene transfer. The genome comparisons of Bradyrhizobium and Frankia strains support the proposed role of Bradyrhizobium in the evolution and acquisition of nodulation in Frankia (Giraud *et al*. 2007, Normand *et al*. 2007). This deduction is subject to more analysis on acquisition of more sequence data. Whole Genome alignments were performed using GenomeVista (Couronne 2003) with a 611kb genomic region of Mesorhizobium nod gene island (Uchiumi *et al*. 2004) against genomes of

Frankia and Rhizobium (nodulating and non-nodulating species). The Frankia genome had greater sequence similarity with the Mesorhizobium nod gene island than closely related non-nodulating species of Rhizobium, further implicating functional similarity and horizontal gene transfers (Data available on request).

Project Rationale

With the advent of recent molecular –omics technologies and tools, organisms which in the past were considered intractable for basic research are now used for cutting edge studies because they now present unique opportunities to broaden and enhance our understanding of the evolution of specialized biological systems and their divergent functions in various plant families (Abbott *et al*. 2006, USDA grant proposal). The suitability of Rosaceae for this study is mirrored in evidences ranging from the heritable vascular nodulation in *Malus x domestica (*McIvor *et al.* 2001) to the high nodulation gene sequence matches lacking in other distantly related species. Rationales for this study include:

1. Enriching knowledge of Frankia-Actinorhizal biology through dissection of gene regulatory networks and ultimately better understanding of nodulation symbiosis and the transition from the highly pervasive Arbuscular Mycorrhizal symbiosis. The proximity of the Actinorhizal species in the Rosid 1 clade and more recent evolution of nodulation offer a strategic edge for in-depth comparative study

between closely related species across extremes of the nodulation symbiosis spectrum.

2.Generation of a comprehensive Actinorhizal ESTs database for the Actinorhizal and nodulation symbiosis community at large.

3.Rosaceae features extensive genomics resources available for identification and characterization of genes important in nodulation (Abbott *et al*. 2006, USDA grant proposal). The resources, available in 3 of the 4 subfamilies, represent the most diverse family (3,000-4,000 species in 100-120 genera) of the Actinorhizal clades.

4. The genomes of Rosaceae species are among the smallest plant genomes making them exceptionally amenable to large scale high throughput genomics studies (Abbott *et al*. 2006, USDA grant proposal).

5.Closely related genera to the Actinorhizal clades that do not fix nitrogen have significant genomic resources available (ESTs and large insert libraries) and are easily transformed (Strawberry) for functional studies (Abbott *et al*. 2006, USDA grant proposal).

References

Aber J.D. 1992. Nitrogen cycling and nitrogen saturation in temperate forest ecosystems. Trends in Ecology and Evolution 7:220-223.

Allen O.N., Allen E.K. 1981. The Leguminosae: A Source Book of Characteristics, Uses and Nodulation. University of Wisconsin Press, Madison, WI/Macmillan Publishing, London

Bailly X., Olivieri I., Brunel B., Cleyet-Marel J.C., Bena G. 2007. Horizontal gene transfer and homologous recombination drive the evolution of the nitrogen-fixing symbionts of Medicago species. J Bacteriol. 189(14):5223-36

Batzli J.M., Graves W.R., Van Berkum P. 1992. Diversity among Rhizobia effective with *Robinia pseudoacacia*. Appl. Environ. Microbiol. 58, 2137-2143.

Becking J.H. 1970. *Frankiaceae* fam. nov. (Actinomycetales) with one new combination and six new species of the genus *Frankia* Brunchorst 1886, 174. Inter. J. System. Bacteriol. 20:201-220.

Benson D.R., Silvester W.B. 1993. Biology of Frankia Strains, Actinomycete Symbionts of Actinorhizal Plants, Microbiological Reviews, p 293-319.

Berry A.M., Torrey J.G. 1983. Root hair deformation in the infection process of *Alnus rubra*. Can. J. Microbiol. 61:2863-2976.

Chaintreuil C., Boivin C., Dreyfus B., Giraud E. 2001. Characterization of the common nodulation genes of the photosynthetic *Bradyrhizobium sp* ORS285 reveals the presence of a new insertion sequence upstream of nodA FEMS Microbiology Letters 194 (1): 83-86

Couronne O., Poliakov A., Bray N., Ishkhanov T., Ryaboy D., Rubin E., Pachter L., Dubchak I. 2003. Strategies and Tools for Whole-Genome Alignments. Genome Res., 13:7.

Dawson J.O., Sun S.H. 1981. The effect of Frankia isolates from *COMPTONIA-PEREGRINA* and *ALNUS-CRISPA CRISPA* on the growth of *ALNUS-GLUTINOSA*, A *ALNUS-CORDATA*, and A *ALNUS-INCANA* clones Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere 11 (4): 758-762

Den Herder J., Vanhee C., De Rycke R., Corich V., Holsters M., Goormachtig S. 2007. Nod factor perception during infection thread growth fine-tunes nodulation. Molecular Plant-Microbe Interactions 20 (2): 129-137

Desbrosses G.G., Kopka J., Udvardi M.K. 2005. *Lotus japonicus* Metabolic Profiling. Development of Gas Chromatography-Mass Spectrometry Resources for the Study of Plant-Microbe Interactions Plant Physiol. 137(4): 1302–1318.

D'Haeze W., De Rycke R., Mathis R., Goormachtig S., Pagnotta S., Verplancke C., Capoen W., Holsters M. 2003. Reactive oxygen species and ethylene play a positive role in lateral root base nodulation of a semiaquatic Legume. Proceedings of the National Academy of Sciences of the United States of America 100 (20): 11789-11794

Doyle J.J. 1998. Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. Trends Plant Sci. 3: 473–478.

Duc G., Trouvelot A., Gianinazzi-Pearson V., Gianinazzi S. 1989. First report of non-mycorrhizal plant mutants (myc$^-$) obtained in pea (*Pisum sativum*) and fababean (*Vicia faba* L.). Plant Sci 60: 215–222.

Evans R.C., Alice L.A., Campbell C.S., Kellogg E.A., Dickinson T.A. 2000. The granule-bound starch synthase (GBSSI) gene in Rosaceae: multiple loci and phylogenetic utility. Mol. Phylogen. Evol. 17:388-400.

Eyre-Walker A., Gaut B.S. 1997. Correlated rates of synonymous site evolution across plant genomes. Mol. Biol. Evol. 14:455–460.

Geurts R., Fedorova E., Bisseling T. 2005. Nod factor signaling genes and their function in the early stages of Rhizobium infection. Current Opinion in Plant Biology 8: 346-352.

Giraud E., Hannibal L., Fardoux J., Jaubert M., Jourand P., Dreyfus B., Sturgis J.N., Vermeglio A. 2004. Two distinct crt gene clusters for two different functional classes of carotenoid in Bradyrhizobium Journal of Biological Chemistry 279 (15): 15076-15083

Giraud E., Moulin L., Vallenet D., Barbe V., Cytryn E., Avarre J.C., Jaubert M., Simon D., Cartieaux F., Prin Y., Bena G., Hannibal L., Fardoux J., Kojadinovic M., Vuillet L., Lajus A., Cruveiller S., Rouy Z., Mangenot S., Segurens B., Dossat C., Franck W.L., Chang W.S., Saunders E., Bruce D., Richardson P., Normand P., Dreyfus B, Pignol D., Stacey G., Emerich D., Vermeglio A., Medigue C., Sadowsky M. 2007. Legumes symbioses: absence of Nod genes in photosynthetic Bradyrhizobia. Science. 316(5829):1307-12.

Gleason C., Chaudhuri S., Yang T., Munoz A., Poovaiah B.W., Oldroyd G.E.D. 2006. Nodulation independent of Rhizobia induced by a calcium-activated kinase lacking autoinhibition, Nature, 441; 1149-1152.

Godfroy O., Debelle F., Timmers T., Rosenberg C. 2006. A rice calcium- and calmodulin-dependent protein kinase restores nodulation to a Legume mutant. Mol Plant Microbe Interact. 19(5):495-501.

Hirsch A.M., LaRue T.A. 1997. Is the Legume nodule a modified root or stem or an organ sui generis? Critical Reviews in Plant Sciences, 16: p. 361-392

Hocher V., Auguy F., Argout X., Laplaze L., Franche C., Bogusz D. 2006. Expressed sequence-tag analysis in Casuarina glauca actinorhizal nodule and root. New Phytol.; 169(4):681-8.

Kistner C. Parniske M. 2002. Evolution of signal transduction in intracellular symbiosis. Trends Plant Sci. Nov; 7(11):511-8. Review.

Kohls S.J., Thimmapuram J., Buschena C.A., Buschena C.A., Paschke M.W., Dawson J.O. 1994. Nodulation patterns of actinorhizal plants in the family Rosaceae. Plant and Soil 162 (2): 229-239

Lavin M., Herendeen P.S., Wojciechowski M.F. 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. Syst Biol 54: 574–594

Laroche J., Bousquet J. 1999. Evolution of the mitochondrial *rps3* intron in perennial and annual angiosperms and homology to nad5 intron 1. Mol. Biol. Evol. 16:441–452.

Leul M., Normand P., Sellstedt A. 2007. The organization, regulation and phylogeny of uptake hydrogenase genes in Frankia Physiologia Plantarum 130 (3): 464-470

Liu, Q., Berry A.M. 1991. The infection process and nodule initiation in the *Frankia-Ceanothus* root nodule symbiosis. Protoplasma 163:82-92.

Llop-Tous I., Dominguez-Puigjaner E., Vendrell M. 2002 Characterization of a Strawberry cDNA clone homologous to calcium-dependent protein kinases that is expressed during fruit ripening and affected by low temperature. J Exp Bot. 53(378):2283-5.

Long S.R. 1984. in *Plant Microbe Interactions* (eds Kosuge, T. and Nester, E. W.), Macmillan, New York, , vol. I, pp. 265–280.

Maudinas B., Chemardin M., Gadal P. 1982. Fatty-acid composition of roots and root-nodules of Alnus species. Phytochemistry 21 (6): 1271-1273

McIvor I., Lang S. Jameson P.E. 2001. Incidence of vascular nodules in the fruit of Pacific Rose™ and related Apples (*Malus x domestica* borkh.). Acta Hort. (ISHS) 564:123-127

Miller I. M., Baker D.D. 1985. Initiation, development and structure of root nodules in *Elaeagnus angustifolia* L. (Elaeagnaceae). Protoplasma 128:107-119.

Moreiro F.M.S., Franco A.A. 1994. Rhizobia-host interactions in tropical ecosystems in Brazil. In: Sprent J. I., McKey D. (Eds.), Advances in Legume Systematics, Part 5. The Nitrogen Factor. The Royal Botanic Gardens, Kew, pp. 63-74.

Morgan D.R., Soltis D.E., Robertson K.R. 1994. Systematic and evolutionary implications of *rbc*L sequence variation in Rosaceae. American J. Bot. 81:890-903.

Muse S.V. 2000. Examining rates and patterns of nucleotide substitution in plants. Plant Mol. Biol. 42:25-43.

Nandasena K.G., O'hara G.W., Tiwari R.P., Sezmis E., Howieson J.G. 2007. In situ lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of Mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture Legume *Biserrula pelecinus* L. Environ Microbiol.; 9(10):2496-511.

Normand P., Queiroux C., Tisa L.S., Benson D.R., Rouy Z., Cruveiller S., Medigue C. 2007. Exploring the genomes of Frankia Physiologia Plantarum 130 (3): 331-343

Odee D.W., Sutherland J.M., Kimiti J.M., Sprent J.I. 1995. Natural Rhizobial populations and nodulation status of woody Legumes growing in diverse Kenyan conditions. Plant Soil 173, 211-224.

Potter D. 1997. Variation and phylogenetic relationships of nucleotide sequences for putative genes encoding polygalacturonase inhibitor protein (PGIPs) among taxa of Rosaceae. American J. Bot. 84(6 Suppl.):224.

Potter D., Gao F., Botiri P.E., Oh S.H., Baggett S. 2002. Phylogenetic relationships in Rosaceae inferred from chloroplast *matK* and *trnL-F* nucleotide sequence data. Plant Syst. Evol. 231:77-89.

Racette S., Torrey J.G. 1989. Root nodule initiation in *Gymnostoma* (Casuarinaceae) and *Shepherdia* (Elaeagnaceae) induced by *Frankia* strain HFPGpI1. Can. J. Bot. 67:2873-2879.

Reddell P., Bowen G.D. 1985. *Frankia* source affects growth, nodulation and nitrogen fixation in *Casuarina* species. New Phytol. 100:115-122. Dawson, J. and Sun, Soon-Hwa, *Can. J. For. Res*., 1981, 11, 758–762.

Riely B.K., Mun J.H. Ane J.M. 2006. Unravelling the molecular basis for symbiotic signal transduction in Legumes. Molecular Plant Pathology 7: 197-207.

Schulze-Menz G.K. 1964. Rosaceae, p. 209-218. *In* H. Melchior (ed.), Engler's Syllabus der Pflanzenfamilien II 12th ed. Gebruder Boerntraeger, Berlin.

Seguin A., Lalonde M. 1989. Detection of pectolytic activity and *pel* homologous sequences in *Frankia*. Plant and Soil 118:221-229.

Selim S., Delacour S., Schwencke J. 1996. Specific long-chain fatty acids promote optimal growth of Frankia: accumulation and intracellular distribution of palmitic and propionic acid. Arch Microbiol 165:252-7.Maudinas, B., Chemardin, M. and Gadal, P., *Phytochemistry*, 1982, 21, 1271–1273.

Silvester W.B. 1976, Ecological and economic significance of the non-Legume symbiosis, p 489-506 in W.E. Newton and C. J. Nyman eds, Proceedings of the First International Symposium on Nitrogen Fixation, Washington University Press (Pullman, WA)

Simonet P., Navarro E., Rouvier C., Reddell P., Zimpfer J., Dommergues Y., Bardin R., Combarro P., Hamelin J., Domenach A., Gourbiere F., Prin Y., Dawson J.O., Normand P. 1999. Co-evolution between Frankia populations and host plants in the family Casuarinaceae and consequent patterns of global dispersal Environmental Microbiology 1(6), 525-533

Simonet P., Normand P., Bardin R. 1988. Heterologous hybridization of *Frankia* DNA to *Rhizobium meliloti* and *Klebsiella pneumoniae nif* genes. FEMS Microbiol. Lett. 55:141-146. Seguin, A. and Lalonde, M., *Plant Soil*, 1989, 118, 221–229.

Soltis D.E., Soltis P.S., Morgan D.R., Swensen S.M., Mullin B.C., Dowd J.M., Martin P.G. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. Proc. Natl. Acad. Sci. 92: 2647–2651.

Sougoufara B., Maggia L., Duhoux E., Dommergues, Y.R. 1992. Nodulation and N-2 fixation in 9 Casuarina clone-Frankia strain combinations. *Acta Oecol*., 13, 497–503.

Sprent J.I., James E.K. 2007. Legume evolution: Where do nodules and mycorrhizas fit in? Plant Physiology 144 (2): 575-581

Sprent J.I., Parsons R. 2000. Nitrogen fixation in Legume and non-Legume trees, Field Crops Research 65, 183-196.

Stacey G., Libault M., Brechenmacher L., Wan J., May, G.D. 2006. Genetics and functional genomics of Legume nodulation. Current Opinion in Plant Biology 9: 110-121.

Thorne R.F., Reveal J.L. 2007. An updated classification of the class magnoliopsida ("Angiospermae") Botanical Review 73 (2): 67-181

Torrey J.G. 1976. Initiation and development of root nodules of *Casuarina* (Casuarinaceae). Amer. J. Bot. 63:335-344.

Uchiumi T., Ohwada T., Itakura M., Mitsui H., Nukui N., Dawadi P., Kaneko T., Tabata S., Yokoyama T., Tejima K., Saeki K., Omori H., Hayashi M., Maekawa T., Sriprang R., Murooka Y., Tajima S., Simomura K., Nomura M., Suzuki A., Shimoda Y., Sioya K., Abe M., Minamisawa K. 2004. Expression islands clustered on the symbiosis island of the *Mesorhizobium loti* genome. J Bacteriol. 186(8):2439-48.

Ueda T., Suga Y., Yahiro N., Matsuguchi T. 1995. Phylogeny of sym plasmids of Rhizobia by PCR based sequencing of a *nodC* segment. J. Bact. 177: 468–472.

Vance C., Young N., Samac D., VandenBosch K., Gantt S., Retzel, E. 2002. Rationale for Meeting. Program and Abstract Book: International Conference on Legume Genomics and Genetics: Translation to Crop Improvement, U of Minnesota.

Van Ghelue M., Ribeiro A., Solheim B., Akkermans A.D.L., Bisseling T., Pawlowski K. 1996. Sucrose synthase and enolase expression in actinorhizal nodules of *Alnus glutinosa*: Comparison with Legume nodules. Mol. Gen. Genet. 250, 437–446.

Vanden Heuvel, B.D., Benson D.R., Bortiri E., Potter, D. 2004. Low genetic diversity among *Frankia spp*. strains nodulating sympatric populations of Actinorhizal species of Rosaceae, Ceanothus (Rhamnaceae) and *Datisca glomerata* (Datiscaceae) west of the Sierra Nevada (California). Can. J. Microbiol. 50: 989–1000.

Verghese S.K., Misra A.K. 2002. Frankia-Actinorhizal symbiosis, with special reference to host-microsymbiont relationship. Curr. Sci.:404-408.

Vessey J.K., Pawlowski K., Bergman B. 2005. Root-based N-2-fixing symbioses: Legumes, *actinorhizal plants, Parasponia sp* and cycads Plant and Soil 274: 51-78

Appendix Q: License for Figure 1.2

ELSEVIER LICENSE
TERMS AND CONDITIONS

Mar 27, 2010

This is a License Agreement between Bode A Olukolu ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Bode A Olukolu
Customer address	Genetics & Biochemistry Dept. Office Clemson, SC -0318
License Number	2396131195500
License date	Mar 25, 2010
Licensed content publisher	Elsevier
Licensed content publication	Trends in Plant Science
Licensed content title	Flowering time control and applications in plant breeding
Licensed content author	Christian Jung, Andreas E. Müller
Licensed content date	October 2009
Volume number	14
Issue number	10
Pages	11
Type of Use	Thesis / Dissertation
Portion	Figures/tables/illustrations
Number of Figures/tables/illustrations	1
Format	Both print and electronic
You are an author of the Elsevier article	No
Are you translating?	No
Order Reference Number	
Expected publication date	Mar 2010
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
Value added tax 0.0%	0.00 USD
Total	0.00 USD