Revealing the genetic components responsible for the unique photosynthetic stem capability of the wild almond Prunus arabica (Olivier) Meikle

Hillel Brukental[1, 2], Adi Doron-faigenboim[3], Irit Bar-Ya'Akov[2], Rotem Harel-Beja[2], Ziv Attia[2], Tamar Azoulay-Shemer[2], Doron Holland[2*]

[1]Hebrew University of Jerusalem, Israel, [2]Unit of Deciduous Fruit Tree Sciences, Newe Ya'ar Research Center, Agricultural Research Organization, Volcani Center, Israel, [3]Department of Vegetable and Field Crops, Agricultural Research Organization (ARO), Israel

Submitted to Journal:
Frontiers in Plant Science

Specialty Section:
Plant Breeding

Article type:
Original Research Article

Manuscript ID:
779970

Received on:
20 Sep 2021

Revised on:
13 Oct 2021

Journal website link:
www.frontiersin.org

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contribution statement

This study supported by the "LEONA M. AND HARRY B. HELMSLEY CHARITABLE TRUST"

Keywords

almond, prunus arabica, Wild almond, stem photosynthesis, QTL, genetic mapping, decidouos fruit trees

Abstract

Word count: 228

Almond (Prunus dulcis (Mill.) D. A. Webb) is a major deciduous fruit tree crop worldwide. During dormancy, under warmer temperatures and inadequate chilling hours, the plant metabolic activity increases and may lead to carbohydrate deficiency. Prunus arabica (Olivier) Meikle is a bushy wild almond species known for its green, un-barked stem, which stays green even during the dormancy period. Our study revealed that P. arabica green stems assimilate significantly high rates of CO2 during the winter as compared to P. dulcis cv. Um el Fahem (U.E.F), and may improve carbohydrate status throughout dormancy. To uncover the genetic inheritance and mechanism behind the P. arabica Stem Photosynthetic Capability (SPC), a segregated F1 population was generated by crossing P. arabica to U.E.F. Both parent's whole genome was sequenced, and a single nucleotide polymorphism (SNP) calling identified 4,887 informative SNPs for genotyping. A robust genetic map for U.E.F and P. arabica was constructed (971 and 571 markers, respectively). QTL mapping and association study for the SPC phenotype revealed major QTL (log of odd (LOD)=20.8) on chromosome 7, and another minor but significant QTL on chromosome 1 (LOD=3.9). Finally, a list of 71 candidate genes was generated. This work sets the stage for future research to investigate the mechanism regulating the SPC trait, how it affects the tree's physiology, and its importance for breeding new cultivars better adapted to high winter temperatures.

Contribution to the field

The manuscript describes an exceptional discovery showing that the wild almond P. arabica can photosynthesize through its stems. This capability lasts all-year round and even in winter when the plant is dormant and there are no leaves. The level of stem photosynthesis is very similar to that of leaves. Stem photosynthesis was shown in desert plants in California, but never on fruit trees and its biological significance was not shown. The ability to fix CO2 in stems during winter can allow fruit trees to maintain a stable energy balance without exhausting their energy reserves through respiration. This is highly important in deciduous fruit trees that need this energy to support the developing trees in spring when the leaves are not yet fully active. To our best knowledge, this manuscript shows for the first time that the unique stem photosynthesis trait could potentially be used as an important mechanism to increase yield and better cope with climate change particularly in hot winters. We developed for the first time in almond a unique system of SNP markers spaced more or less evenly along the almond genome. These unique SNPs allowed us to find a strong QTL for the trait, suggesting that stem photosynthesis is controlled by a major gene. Mapping of the stem photosynthesis trait was never reported before in any plant.

Ethics statements

Studies involving animal subjects

Generated Statement: No animal studies are presented in this manuscript.

Studies involving human subjects

Generated Statement: No human studies are presented in this manuscript.

Inclusion of identifiable human data

Generated Statement: No potentially identifiable human images or data is presented in this study.

Data availability statement

Generated Statement: The authors acknowledge that the data presented in this study must be deposited and made publicly available in an acceptable repository, prior to publication. Frontiers cannot accept a manuscript that does not adhere to our open data policies.

Revealing the genetic components responsible for the unique photosynthetic stem capability of the wild almond *Prunus arabica* (Olivier) Meikle

Hillel Brukental[1,2], Adi Doron-Faigenboim[3] , Irit Bar-Ya'akov[1], Rotem Harel-Beja[1], Ziv Attia[1] Tamar Azoulay-Shemer[1] Doron Holland[1*]

[1]Unit of Fruit Tree Sciences, Institute of Plant Sciences, Agricultural Research Organization, Newe Ya'ar Research Center, Ramat Yishay, Israel

[2] The Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agriculture, Hebrew University of Jerusalem, Rehovot, Israel

[3]Department of Vegetable and Field Crops, Institute of Plant Sciences, Agricultural Research Organization, Volcani Center, Rishon LeZion, Israel

Correspondence:
Doron Holland

vhhollan@volcani.agri.gov.il

Keywords: almond, *prunus arabica*, wild almond, stem photosynthesis, QTL, genetic mapping, decidouos fruit trees.

Abstract

Almond (Prunus dulcis (Mill.) D. A. Webb) is a major deciduous fruit tree crop worldwide. During dormancy, under warmer temperatures and inadequate chilling hours, the plant metabolic activity increases and may lead to carbohydrate deficiency. Prunus arabica (Olivier) Meikle is a bushy wild almond species known for its green, un-barked stem, which stays green even during the dormancy period. Our study revealed that P. arabica green stems assimilate significantly high rates of CO2 during the winter as compared to P. dulcis cv. Um el Fahem (U.E.F), and may improve carbohydrate status throughout dormancy. To uncover the genetic inheritance and mechanism behind the P. arabica Stem Photosynthetic Capability (SPC), a segregated F1 population was generated by crossing P. arabica to U.E.F. Both

parent's whole genome was sequenced, and a single nucleotide polymorphism (SNP) calling identified 4,887 informative SNPs for genotyping. A robust genetic map for U.E.F and P. arabica was constructed (971 and 571 markers, respectively). QTL mapping and association study for the SPC phenotype revealed major QTL (log of odd (LOD)=20.8) on chromosome 7, and another minor but significant QTL on chromosome 1 (LOD=3.9). As expected, the *P. arabica* allele in the current loci significantly increased the SPC phenotype. Finally, a list of 64 candidate genes was generated. This work sets the stage for future research to investigate the mechanism regulating the SPC trait, how it affects the tree's physiology, and its importance for breeding new cultivars better adapted to high winter temperatures.

Introduction

Almond, *Prunus dulcis* (Mill.) D. A. Webb, is a major fruit tree crop worldwide. As a deciduous fruit tree, it enters dormancy during early winter and renews growth following the fulfillment of a variety-specific period of exposure to low temperatures, known as chilling requirements (CR) and adequate heat requirements. Exposure to a sufficient number of low winter-temperatures is essential for synchronized flowering in the early spring followed by efficient pollination, fruit set and fruit development (Sánchez-Pérez et al. 2014). CR limit growing areas of deciduous fruit trees and dramatically influence the yield and quality of fruit (Atkinson, Brennan, and Jones 2013). When winter temperature increases, CR are not sufficiently provided, and the metabolic activity increases (Sperling et al. 2017). As a result, carbohydrates are consumed, and intense starch synthesis occur. These changes lead to soluble carbohydrate (SC) deficiency in the buds during the period of flowering and fruit set, which results in disruptive flowering that may reduce yield (Tixier et al. 2017; Fernandez et

al. 2018; Guo et al. 2021). The ability of the dormant almond to respond this energy depletion is restricted, mainly due to the shortage in photosynthetic leaves during dormancy. Climate changing trends emphasize the urgent need for deciduous fruit crops to gain more plasticity (Gradziel et al. 2001), particularly for maintaining their nonstructural carbohydrate (NSC) reserves in warmer winters (Zwieniecki, Tixier, and Sperling 2015; Atkinson, Brennan, and Jones 2013).

Prunus arabica (Olivier) Meikle, also known as *Amygdalus arabica Olivier*, defined as a different species from the domesticated almond *P. dulcis*. However, both belong to the *Prunus* genus and are a part of the Rosacea family. The species "arabica" was named after the geographical region where it was first described. This taxon is native to the temperate-Asia zone. It covers the Fertile Crescent Mountains, Turkey, Iran and Iraq. In the Middle East it can be found in Lebanon, Syria, Israel (Judean Desert) and Jordan (Roskov 2019). *P. arabica* can be found in altitudes between 150-1,200 m and rarely up to 2,700 m. It is a bush, rather than a tree, with a very long root system and is considered resistant to drought (Rajabpoor et al. 2014; Sorkheh et al. 2011). As a deciduous tree, *P. arabica* drops its leaves at the end of the summer, turns meristems into buds and stops growing. However, unlike other almond species, its young branches remain green and are not covered with bark (i.e., no cork layer deposition) throughout the dormancy phase (Fig. 1 a-d). In fact, *P. arabica* stems remain moist and green during the whole year. *P. arabica* green stems were previously suggested to photosynthesize (Sorkheh et al. 2009), yet no physiological evidence was published regarding their ability to assimilate external CO_2.

Stem photosynthesis was previously shown in other desert species (Aschan and Pfanz 2003). In these species, which do not belong to the Rosacea family, high efficiency of CO_2 assimilation comparable with that of the leaf was demonstrated. Because stem photosynthesis

was found in desert plants, it was suggested that it might play a role in carbon gain under stress conditions such as heat and drought (Nilsen 1995). The contribution of stem photosynthesis to tree adaptation under drought was further supported by evidence showing that stem photosynthesis assists to embolism repair (De Baerdemaeker et al. 2017; Bloemen et al. 2016). Ability to photosynthesize through stems could prove to be highly beneficial for deciduous fruit trees such as almonds in maintaining the energy balance of the tree, particularly during hot winters and springs when respiration is enhanced, and tree energy is limited due to leaf drop and the lack of photosynthetic organs (Zwieniecki, Tixier, and Sperling 2015; Sperling et al. 2019). Previous genetic studies of *P. arabica* were limited to phylogenetic studies and encompass a small number of markers (few to dozens) (Delplancke et al. 2016; Yazbek and Oh 2013). To the best of our knowledge, no genetic approach detected to uncover the mechanism of the stem photosynthesis phenomena.

Recent important advancement in Rosacea genetics and genomics enables the application of a genetic approach in the study of important physiological processes. Such advancements include the development of genetic maps based on F1 and F2 populations (Sánchez-Pérez et al. 2014), and their usage for mapping QTLs affecting CR in apple (*Malus domestica*)(Miotto et al. 2019), pear (*Pyrus communis*)(Gabay et al. 2018), apricot (*Prunus armeniaca)*(Olukolu et al. 2009), and peach (*Prunus persica)* (Wang et al. 2002). In almonds, genetic mapping of hybrid populations based on distinctive CR, demonstrated that a major gene, *LATE BLOOMING* (*LB*) was associated with blooming date and dormancy release (Ballester et al. 2001). In addition, complete genomes and various transcriptomic datasets of Rosacea: apple, cherry *(Prunus avium)* and peach (Jung et al. 2019) were published. Recently, two almond genomes were published: *P. dulcis cv.* Texas (https://www.ncbi.nlm.nih.gov/bioproject/572860) and *P. dulcis cv.* Lauranne

*(*https://www.ncbi.nlm.nih.gov/bioproject/553424*)*. This trend of new genomic and transcriptomic data of deciduous fruit trees, including almonds, sets the stage for intense genetic research and the development of novel marker-assisted breeding approaches. In this report, we combined physiological and genetic approaches to study the green stems of *P. arabica* throughout the year. By direct measurements of gas exchange, we demonstrate that *P. arabica* stems transpire and assimilate significant levels of CO_2 all year round, including during the dormancy period. Moreover, we undertook a forward genetic approach, and used an F1 population segregating for the stem photosynthesis trait for high-resolution mapping of major QTLs for this trait. The genetic markers and candidate genes that our study underlines pave the way to undermine the physiological role of stem photosynthesis and its utilization for genetic improvement.

Results

P. arabica assimilate CO_2 through green stems

P. arabica stems remain green during winter while cultivated almonds develop an outer grey cork layer (Fig.1b, d). To study if these stems are actively assimilating CO_2, we undertook gas exchange measurements of tree stems in the orchard with the Licor 6800 Portable Photosynthesis System. Two different almond species were compared, the wild almond *P. arabica* and the cultivated almond *P. dulcis* (U.E.F), throughout the entire year (Fig. 1e). The data indicate that *P. arabica* ~~assimilates~~ assimilated CO_2 through its green stems during all year (annual average of 8±0.19 µmol CO_2 m^{-2} sec^{-1}), while similar one-year old stems of U.E.F assimilation capacity is almost ~~nil~~null (annual average of 0.5±0.05 µmol CO_2 m^{-2} sec^{-1}). The significantly high CO_2 assimilation rates of *P. arabica* stems were found

comparable with assimilation rates of *P. arabica* leaves (11.2±0.8 CO_2 m^{-2} sec^{-1}, July average, data not shown). Although some fluctuations ~~were~~ observed between the different seasons, pronounced high CO_2 assimilation rates were found in *P. arabica* stem during the whole year (Fig. 1e). Finally, *P. arabica* stem transpiration rate is relatively low in the dormancy phase and gradually increases until it peaks in October (1.2±0.18 in January to 5±0.7 mmol H_2O m^{-2} sec^{-1} in October). In contrast, transpiration from U.E.F stems is relatively constant and low throughout the year (0.46±0.11 mmol H_2O m^{-2} sec^{-1}; Fig. 1e). Transpiration rate fluctuation also attribute to high instantaneous water use efficiency (iWUE) of *P. arabica* during the dormancy phase (two-fold higher than U.E.F in December; Fig. 1e).

Previous studies on temperate fruit trees showed a positive correlation between tissue temperature and SC consumption through respiration (Sperling et al. 2019). To find out how stem respiration of *P. arabica* and U.E.F are influenced by temperature, we measured the respiration rate of one-year old stems while exposing them to three different temperatures (17°, 28° and 34° C) (Fig. 1f). Increased respiration rate in response to elevated temperature was observed in both almond species (0.5±0.11, 2.2±0.27, 3.5±0.44 and 0.33±0.12, 1.4±0.35, 3±0.68 µmol CO_2 m^{-2} sec^{-1} for *P. arabica* and U.E.F respectively for each temperature), while no significant differences were observed between species (for each measured temperature).

Stem assimilation is genetically inherited

To elucidate the genetic nature of the assimilating stem trait of *P. arabica*, an F1 hybrid population (n = 92) was established by crossing *P. arabica* (male) with U.E.F (female) (Fig. 2c, d). The same approach of gas exchange measurements in the field was used for phenotyping the SPC trait among the three year-old F1 population during dormancy. Twelve

offspring assimilated CO_2 via their stems in a similar level as *P. arabica* (offspring 24H27 is the highest; 8.3±0.14 µmol CO_2 m^{-2} sec^{-1}), and Thirty- seven individuals assimilated as U.E.F or less, all others offspring phenotypes fluctuated between the parents (Fig. 2a). Analysis of distribution demonstrated two prominent peaks within the histogram (Fig. 2b). Although the '3 Normal Mixture' is the most accurate model to describe the the current phenotype distribution (achieved the lowest AICc and the -2Log Likelihood values). Broad-sense heritability (h^2) was found to be high (0.91).

Sequence comparisons, SNPs identification, and genotyping of the F1 population

Segregation of the SPC trait among the F1 population rendered it the F1 population as suitable infrastructure for genetic mapping. For this purpose, we sequenced the *P. arabica* and the U.E.F genomic DNA, targeting for high coverage, to ensure reliable (SNP) calling. The reads were aligned against the reference genome of *P. dulcis cv.* Lauranne, becausedue to awhich better mapping results related towere obtained with the lLauranne genome as compared to *P. dulcis cv.* Texas genome was found as the closest (> 97%, ~85% respectively) mapped reads) of the two published almond genomes (Table 1).

A total of 3,750,363 and 2,407,787 variants (i.e., SNPs or short InDels) were detected for *P. arabica* and U.E.F respectively, against the cv. Lauranne reference genome. Analyzing the variants showed that 71.5% and 72.6% (*P. arabica* and U.E.F, respectively) are in the intergenic region (Table S1). Furthermore, a higher variant number was detected in the

intronic regions in relation to the exons (Table S1). The initial number of identified SNPs (Total variant sites in Table 1) were filtered by several types of criteria as specified in materials and methods.

Overall, 4,887 SNPs that are heterozygous for one of the parents and homozygous for the second were selected for F1 genotyping screening. The SNPs are spread at intervals of about 40K along the almond genome.

The F1 population was successfully genotyped with 4,6125 SNPs. The resulting genotyping quality data (Table S2) represent high coverage (152X), and low number of missing data (5.5%). Further analysis of the genotyped F1 population with the SNP panel described above show the allelic frequency within the F1 population is 50%, as expected from an F1 population (Fig. 3). Moreover, this expected ratio suggest there is no plant contamination in our F1 plant material. However, since the allelic composition in this bi-parental population is AA x Aa, we can also refer this ratio as the allelic frequency of the heterozygous genotype. Therefore, data presented (Fig. 3) also indicates exceptional chromosomal regions (hot spot) with a unique pattern of inheritance that deviates from the 1:1 ratio, for example, in chromosome 3 (see black arrow in Fig. 3).

Construction of genetic maps for the F1 population

To establish a genetic map of the F1 population, Join Map 4.1 software was used (J. W. & J. J. Van Ooijen 2013). CP (cross pollination) population type was performed with the lmxll code for markers that were homozygous for the male parents (*P. arabica*) and heterozygous for the female parent (U.E.F). The code nnxnp used for the opposite case. A significant

190 portion of the markers was filtered, most of them due to complete similarity (~50%). Overall,
191 1,533 SNPs were used for mapping (Table 2). Because there were no common markers for
192 both parents, the hkxhk code was not applied. Using the pseudo test cross method (Olukolu et
193 al. 2009), we separated the markers for two different maps: one map for the U.E.F (where *P.*
194 *arabica* is homozygous, lmxll code), and the second map for the *P. arabica* parent (where
195 U.E.F is homozygous, nnxnp code). Applying this strategy, we obtained two maps with
196 robust numbers of markers and good density. The U.E.F map was found to be denser than the
197 *P. arabica* map and includes 971 markers with an average distance of 0.533 centiMorgen
198 (cM), while *P. arabica* map contains 572 SNPs with an average distance of 1.093 (cM)
199 (Table 2). It can be clearly seen that the distribution of the SNP markers is well spread (the
200 biggest gap is 6.173 cM in the U.E.F and 18.85 cM in *P. arabica* map) over the eight almond
201 linkage groups (LG) (Fig. 4).

Formatted: Font: Italic, Complex Script Font: Italic

202 To assess the validity of the genetic map, the order of SNP markers as determined by the
203 U.E.F genetic map was compared with the deduced order from the physical map as
204 determined by cv. Lauranne reference genome. The analysis (Fig. 5) demonstrates a good co-
205 linearity between the genetic and the physical map. Remarkably, most of the markers from
206 the genetic map were highly correlated with the physical order (Fig. 5), yet, few markers did
207 not correlate (chromosome 6, Fig. 5). The genetic map divided the markers into eight LGs
208 parallel to the previously published chromosome organization order. Moreover, the slopes
209 generated between the physical orders to the genetic order represent recombination frequency
210 (cM / Mb). Thus, one can see that around the centromere, the slope is more horizontally,
211 meaning the cM / Mb ratio is relatively low. Thirty-eight SNP markers representing un-
212 scaffold contigs (i.e., chromosome 0) in the reference genome project were assembled into six
213 linkage groups based on the genetic maps (marked by yellow dots in Fig. 5).

214

215 *QTL analysis and genome wide association study (GWAS) of the SPC trait*

216 Two main approaches were initiated for detecting genomic regions regulating the SPC. QTL
217 mapping, computed with Map QTL by interval mapping (IM) analysis (J. W. Van Ooijen
218 2006) and Genome wide association (GWAS) by TASSEL software (Bradbury et al. 2007).
219 QTL mapping generated two significant QTLs. Each QTL was discovered only in one of the
220 two genetic maps. Thus, one major QTL (LOD=20.8) was mapped on LG 7 spanning a region
221 of 2.4 cM detected on the U.E.F map. The second, minor but significant QTL (LOD=3.9) was
222 detected at the end LG 1, spanning a region of 4.4 cM on the *P. arabica* map (Table S3, Fig.
223 6).

224 Applying GWAS approach with TASSEL enabled us to simultaneously detect two genomic
225 sites that regulate the SPC on chromosomes 1 and 7 at positions similar to those detected by
226 QTL mapping. The major region on chromosome 7 spanning only 400kb, and the minor on
227 chromosome 1 containing 700kb. Moreover, GWAS analysis showed significant associations
228 with markers aligned to chromosome 0. Interestingly, two of these markers assembled into
229 the major QTL in locus 7 (Table S3, marked with gray background). The major QTL
230 explained 67% of the phenotypic variance, while the minor QTL explained 19.3% (Table S3).

231

232 *QTL's interaction*

233 As presented, two significant loci were discovered as regulating the SPC (Fig. 6, Table S3).
234 Full factorial test shows a significant additive effect between those two associated loci
235 (<0.0001; Fig. 7a, b). Yet, no epistatic effect was found (p-value= 0.676; Fig. 7c). As

expected, in both QTLs the *P. arabica* alleles were the increasing alleles regarding the SPC trait.

Generating a list of candidate genes

Combining the data on ~~GWAS and QTLs data~~the QTLs bordering markers (which was wider than the one obtained by the ~~-~~associated region detected in GWAS (Table S3)) ~~described above~~ with that of the cv. Lauranne reference sequence allowed us to delineate a list of genes within the regions that are predicted as responsible for the SPC trait ~~(Table S5)~~. In total~~,~~ ~~350~~, 350 genes were detected under the QTLs (Table S5)~~.~~ Tt~~T~~he region at chromosome 1 includes 82 annotated genes ~~(GCA_008632915.2)~~ with SNPs between *P. arabica* and U.E.F ~~(Table 3)~~. Among those, 15 include non-synonymous SNPs in the genes coding region. The associated region at chromosome 7 consists 218 genes with SNPs, of which only 49 have non-synonymous SNPs in their coding region. The data is based on the usage of ~~as detected by~~ the SnpEff software (Cingolani et al. 2012) (Table S1)~~,~~ with the NCBI genome database (GCA_008632915.2)~~-~~.

Discussion

P. arabica can assimilate CO_2 through its green stems all year round

This study investigated the ~~unique~~ wild almond *P. arabica*, a member of the Rosacea family. We demonstrated for the first time that the wild almond *P. arabica* is assimilating external CO_2 via its green stems all year round, including during the dormancy period when the tree sheds its foliage. In contrast, the cultivated almond U.E.F is unable to do so (Fig. 1e). Interestingly, even the young green stems of U.E.F (Fig. 1c), which are not covered with bark layer, in spring, are assimilating CO_2 in a negligible levels (Fig. 1e). *P. arabica* is a desert plant, and its native growing ecosystems encompass the desert margins of the Fertile Crescent (Roskov 2019). Previous studies have indicated that perennial desert shrubs, such as *Ambrosia salsola* and *Bebbia juncea*, are able to photosynthesize through their stems (Ávila-Lovera et al. 2019). This capability was attributed to better adaptation of the plant to arid climate conditions (Aschan and Pfanz 2003). The geographical distribution of *P. arabica* suggests that stem photosynthesis might be important for its survival in dry areas(Nilsen 1995). Indeed, several studies have suggested that stem CO_2 assimilation may help prevent embolism damage in response to drought (De Baerdemaeker et al. 2017; Bloemen et al. 2016). Others have claimed that SPC enables more efficient carbon gain with respect to water loss (transpiration) due to smaller surface area when compared to leaves (Ávila-Lovera, Zerpa, and Santiago 2017). Interestingly, fluctuations were observed in transpiration rates of *P. arabica* stem during the year, which resulted in maximal iWUE in winter time since CO_2 assimilation was not highly affected (Fig.1e). Further physiological and anatomical analyses are needed to understand better this unique SPC trait and its contribution to almonds in harsh climates.

Among the wild and cultivated species of almond, only *P. arabica* and its very close relative, *Prunus scoparia* (Spach) Schneider, are known to produce all- year green stems (Khadivi-Khub and Anjam 2014). Yet, stem CO_2 assimilation in *P. scoparia* was not demonstrated. To

our knowledge, *P. arabica* is the only known deciduous fruit tree that possesses the SPC trait. As such, *P. arabica* is unique among all deciduous fruit trees in its capability to assimilate external CO_2 from stems in winter. In addition to the reasonable assumption that SPC contributes to plant survival in dry conditions by preventing embolism damage and by improving carbon gain, it is anticipated that SPC in winter could influence other critical physiological processes that are dependent on carbohydrate management, including dormancy break and ability to support heavy yield (Granot, David-Schwartz, and Kelly 2013; Guo et al. 2021). This is particularly true for deciduous fruit trees under the pressure of energy shortage when the trees have to support both leaf and flower development (Fernandez et al. 2018). Deciduous fruit trees rely on their energy reserves during this growth phase since their leaves are not yet fully developed (Sperling et al. 2019). Recent studies on deciduous fruit trees demonstrate how higher temperatures are positively correlated with tree NSC consumption. It was shown that during higher winter temperatures, NSC reserves depleted rapidly (Zwieniecki, Tixier, and Sperling 2015). For setting these results in *P. arabica* and U.E.F, we measured respiration rate in response to elevated temperatures. As expected, environmental temperature indeed raised the respiration rate in both species (Fig. 1f). This result emphasizes the possible link between winter temperatures and trees' NSC status during winter. Importantly, it highlights the advantage of a functional photosynthetic organ in warm winters during dormancy phase. Nevertheless, to establish the link between tissue temperature to tree NSC and their contribution to yield, we are currently conducting additional research.

The SPC trait depict a high heritability value and segregates in an F1 population

As a means of investigating the functional role of SPC in almonds, we undertook a forward genetics approach. Identification of the genetic components underlying SPC is crucial for deciphering the role of SPC in respect to dormancy break and adaptation to upcoming climate changes. Moreover, molecular markers for the genes in question could provide an efficient tool for breeding and genetic manipulation for better almond trees. For this purpose, we established an F1 population from a cross between *P. arabica* and U.E.F. Remarkably, SPC measurements of individual progenies within the F1 population indicate that the SPC trait is inherited and segregated already in the F1 population (Fig. 2a). Interestingly, the distribution of the trait is not normal but displays two main peaks, suggesting that this trait is probably controlled by a small number of genes or QTLs (Fig. 2b). Altogether, the data indicated that the F1 hybrid population could be used as a mapping population for the SPC trait. Furthermore, the high heritability value (0.91) suggesting the SPC trait can be integrated into cultivars through classical genetic crosses. Yet, current data is not sufficient to determine the heritability control of the trait (dominant/recessive) due to the continuous nature of the SPC phenotype.

Genotyping

In this study, we report the feasibility of using a novel methodology to obtain genetic markers through QTL mapping of almond. The method of 'targeted SNP seq' that was already described for other plants (Zhang et al. 2020), takes advantage of the availability of almond chromosomal organization and genome sequencing data. Only markers that are evenly spread along the chromosomes at an average of 40K were chosen in order to target the desired marker density and genomic distribution. High coverage for parent's WGS, and the targeted

genotyping, yielded a good SNPs panel with almost no missing data or out-filtered SNPs for low quality.

The expected segregation of the selected markers (i.e., the parent's alleles) is 1:1, meaning the allelic frequency should be 0.5 on average. Indeed, our results, which established the genotyping quality of the SNPs in the F1 population, match this prospect (Fig. 3). The allelic frequency by physical position presentation shows a small number of "hot spots" that display genomic regions with an unexpected segregation ratio (see black arrow on chromosome 3; Fig. 3), where the frequency of the parent's allele is significantly deviating from 1:1 ratio. This may indicate that the SNPs in that region could be placed in or near a lethal allele, which its presence in the progeny is not favorable. Such an allele, for example, could be an incompatibility *S* allele (Gómez et al. 2019).

Genetic maps construction

Two maps were constructed, each for a different parent (Fig. 5). The map for the U.E.F was denser, with 971 markers divided for all the linkage groups (Table 2, Fig. 5). Both maps were fitted to the eight chromosomes published in the two almond reference genomes (https://www.ncbi.nlm.nih.gov/bioproject/572860; https://www.ncbi.nlm.nih.gov/bioproject/553424). The current U.E.F map density (one marker per ~0.5 cM) is almost equal to the average recombination frequency (0.475 cM per Mb) (Table 2). Considering this, together with the small population size (n=92), we assume that the recombination rate (i.e., population size) and not the number of markers is the bottleneck for obtaining better mapping resolution.

High co-linearity between the physical and the genetic maps (Fig. 5) demonstrates the reliability of the constructed map. Furthermore, it gives an overview of the genomic distance between the wild almond species, *P. arabica*, and *P. dulcis* cultivar cv. Lauranne, which, remarkably, seems to be quite similar. Moreover, 98% of the sequenced reads of *P. arabica* were mapped to cv. Lauranne reference genome. Although two reference genomes are available for almond, a genetic map is essential for two main reasons. Firstly, genetic map, which relies on recombination frequency of the current population, should represent the most accurate result for marker arrangement of the population compared to the reference genome (Oren et al. 2020). In this respect, both genetic maps conform well to the physical data of the cv. Lauranne reference genome. Nonetheless, the few markers which are not correlating could emphasize some chromosomal aberrations as translocation. Secondly, using the genetic map, 38 markers that were delineated as chromosome 0, were linked to LGs by recombination (see yellow dots in Fig. 5).

The F1 population and the character of markers selected, resulted in construction of two maps. SNP markers are mainly di alleles, by choosing SNPs that are heterozygous for one parent and homozygous for the other, we can assure that the markers will segregate, and also to identify the parental origin of each allele. In order to intersect the maps, shared markers are needed. Joining the F1 maps could be achieved by SSR markers that have more than di alleles, as was done in apricot F1 population (Hurtado et al. 2002), or by SNPs that are heterozygous for both parents, as was done in pear (Gabay et al. 2018).

SPC genetic mapping

Linkage analysis (i.e., QTL mapping) and GWAS were two approaches used in this study for linking the SPC phenotype and genetic markers (Fig. 6 and Table S3). While GWAS only associates between genomic markers without any data on their specific genomic location, the linkage analysis, which is based on a genetic map connects phenotype to a specific genomic locus / loci (Oren et al. 2020).

QTL mapping discovered two significant loci. One major locus with an exceptional LOD score of ~20 on LG/chromosome 7 was found in the U.E.F map, and another minor (LOD score of ~4) was detected at the end of LG/chromosome 1 in *P. arabica* map. GWAS revealed both loci. Moreover, the loci identified by the GWAS overlapped those found by QTL analysis. Six m~~M~~arkers on chromosome 0 also demonstrate significant LOD score (~20) (Fig. 6, Table S3). Using the genetic map and the QTL analysis enabled the positioning of two of these markers to the~~locus 7~~ QTL of locus 7 ~~(Table S3)~~, while the other markers were positioned in other LGs, or discarded during the genetic map construction~~building~~ procedure (Table S3). ~~.~~ Here we demonstrated the importance of integrating the association data, which is based on the physical reference genome and the QTL mapping approach based on the F1 recombination frequency. The GWAS enabled us to run the analysis with higher number of SNPs~~more saturated map when~~ comparing to the genetic map (3,700 versus 970 in the U.E.F genetic map). The overlapping results between the two methods emphasize the unbiased genetic infrastructure and corroborate the mapping data. The data presented identified new genetic loci on the almond genome. To our knowledge, such a mapping effort on the SPC trait has never been done before for any plant, above all, in trees. This strong mapping data is important in order to fully comprehend the physiological role of SPC and identify new genetic components that control photosynthesis in plants. The availability of

segregating population and genetic markers highly associated with SPC provides a powerful means to explore this trait.

Candidate genes

The high-resolution QTL mapping and the robust annotations data sets available, enabled the establishment of a putative list of candidate genes. The list is based primarily on the data from the QTLs boundaries~~limits~~ (Table S3)~~genetic maps and association studies~~. The list was filtered for non-synonymous polymorphism in the coding region of the genes between the F1 parents, (Table S4). The final list ~~contains six~~contains six genes with the "HIGH" impact variant score as analyzed by the SnpEff software. Based on annotation, the list contains, among others, several genes that are involved in sugar transport (gene ID: Prudu_004403, Prudu_004404, Prudu_004408; Fig. S3). Preliminary results demonstrated a high negative correlation in the F1 population between cork layer (i.e., periderm) development (qualitative 1-5 scale measurements) and the SPC (data not shown). This data suggest that periderm development genes may be involved. For example, *HXXXD-type acyl-transferase* and *MYB-family transcription factor* (gene ID: Prudu_018862, Prudu_018912; Fig. S3) genes, which were suggested to be involved in cork synthesis (Vulavala et al. 2019; Soler et al. 2007). Interestingly, the *EPIDERMAL PATTERNING FACTOR-like protein 2* (EPF2) gene is a member of this list (Table S3; Prudu_018883). The EPF2 gene is a direct regulator of epidermis cell development and stomatal density (Hara et al. 2009). Although the EPF2 gene looks as a promising candidate to control SPC through its role in controlling stomatal density, advanced genetic research should done for establishing the link between the SPC phenotype to this gene. However, this study indicates that the variation of SPC trait in the F1 population

is controlled by a small number of genes localized to only two loci in the almond genome. This work sets the stage for further studies aimed to delineate the genetic nature of the SPC trait, define its importance to the tree, and understand how it could be utilized for tree improvement targeted to produce fruit trees adapted to extreme climate. A recent study demonstrated that the cultivated almond breeding lines are highly conserved and are founded only on the cvs. Tuono, Cristomorto, and Nonpareil (Pérez de los Cobos et al. 2021). This study emphasizes the importance of introduction and utilization of genetic material originating from wild sources. Our current study sets the way for utilizing the wild almond P. arabica as a new source for widening and enriching the current narrow base of almond breeding material.

Conclusion

This paper is the first to establish a genetic study on mapping the unique stem photosynthetic capability originating from the wild *P. arabica* almond. Here we localized the genetic components that regulate the trait and narrowed the whole ~240 Mb almond genome towards only two loci, with one major locus spanning only ~400kb and explaining 67% of the SPC phenotype, which eventually provided a list of 64 candidate genes. Forward genetic approach based on the establishment of a cross-bred population with genetic mapping and GWAS provides a remarkable infrastructure for future introduction of beneficial traits from wild almond origin. This approach is highly efficient for both, study the genetics of important agricultural traits and introducing of new breeding material into highly conserved almond cultivars.

Material and methods

Plant material

All trees are growing in the almond orchard in Newe Ya'ar Research Center in the Yizre'el Valley (latitude $34^0$42'N, longitude $35^0$11'E, Mediterranean temperate to subtropical climate). The parents of the F1 population, *P. arabica and* the Israeli leading commercial cv. Um el Fahem (U.E.F) are grown at two copies for each, grafted on GF.677 rootstock, and planted in winter 2018. The F1 population (*P. arabica* X U.E.F) contains 92 seedlings that were germinated in the nursery in winter 2017 and replanted in the orchard in winter 2018.

Gas exchange measurements

Gas exchange measurements were done in the field on one-year old stems (i.e. current year growth) of three years old *P. arabica* and *P. dulcis* (U.E.F) trees, from October 2019 to October 2020. Each month, two reciprocal days were chosen; in each day, four stems per genotype were analyzed (n=8 per month). All measurements were conducted between the hours 8:30-10:30 a.m. (the latest were in winter). When there were leaves on the stems, they were removed two days prior to measurements to eliminate wounding stress effect. Measurements were carried out with the LI-6800 Portable Photosynthesis System (LI-COR Biosciences, USA), using the 6x6-needle chamber, which is compatible with tree branches of 2.5-4.5 mm diameter. The following conditions were held constant in the chamber: photon flux density of 1,200 µmol m^{-2} sec^{-1} (90% red, 10% blue) and CO_2 reference of 400 PPM was set. Chamber relative humidity, and the temperature held for each month according to the multi-annual average. Gas exchange results were normalized to stem surface area and displayed as net assimilation rates (µmol CO_2 m^{-2} sec^{-1}), transpiration rates (mmol H_2O m^{-2}

461 sec^{-1}), and instantaneous water use efficiency (iWUE; the ratio between net assimilation and
462 transpiration rates).

463 Gas exchange measurements on the F1 population were conducted in February 2020 for two
464 weeks, while the trees were dormant, between the hours 9:30-11:30 a.m. In dormancy there is
465 no interaction with the photosynthetic state of the tree leaves, and the results are more stable.
466 Four stems were measured (n=4) for each genotype. The measurement protocol was the same
467 as mentioned above. The measurements were conducted in~~conducted a year before~~in two
468 successive years, 2019 and 2020. Those of 2019 were conducted with a ~~other~~ portable
469 instrument (CIRAS 3 pp-system U.S.A). Nonetheless, they were significantly correlative with
470 the results obtained with the Licor 6800 instrument in 2020~~ F1~~. (Correlation of 0.67 p value <
471 0.0001 by spearmen). To determine stem respiration rates, stem gas exchange measurements
472 were conducted under the same conditions as described above. Next, the Licor 6800 light
473 source was turned off for ~2 minutes (for stabilization of ΔCO_2), and data were recorded. In
474 dark, the net assimilation value represents respiration. Stem respiration rate was recorded
475 under 17°C, 28°C and 34°C.

476

477 *Whole genome sequencing (WGS) and SNP calling*

478 DNA extracted from young leaves using the plant/fungi DNA isolation kit (NORGEN BIOTEK
479 CORP, Canada). DNA of *P. arabica* and U.E.F, was sent to Macrogen (Macrogen, Korea) for
480 WGS - Illumina Nova Seq 6000, with targeted coverage of X50 on average, read length of 150
481 bp with paired-end sequencing. OmicsBox software (version 1.3.11;
482 https://www.biobam.com/omicsbox/) was used for preprocessing the raw-reads based on
483 Trimmomatic (Bolger, Lohse, and Usadel 2014) for removing adapters and contamination

sequences, trimming low-quality bases, and filtering short and low-quality reads. The cleaned reads were mapped onto the reference genomes: *P.dulcis* cv. Lauranne (https://www.ncbi.nlm.nih.gov/bioproject/553424) and *P. dulcis* cv. Texas (https://www.ncbi.nlm.nih.gov/bioproject/572860), using the Burrows-Wheeler Aligner (BWA) software 0.7.12-r1039, with its default parameters (Li and Durbin 2009). The resulting mapping files were processed using SAMtools/Picard tool (http://broadinstitute.github.io/picard/, version 1.78) (Li et al. 2009); for adding read group information, sorting, marking duplicates, and indexing. Then, the local realignment process for locally realigning reads was performed so that the number of mismatching bases was minimized across all reads using the RealignerTargetCreator and IndelRealigner of the Genome Analysis Toolkit version 3.4-0 (GATK; version http: //www.broadinstitute.org/gatk/) (Depristo et al. 2011). Finally, the variant calling procedure was performed using HaplotypeCaller of the GATK toolkit (https://gatk.broadinstitute.org/hc/en-us) developed by Broad Institute of MIT and Harvard (Cambridge, MA, USA). Only sites with DP (read depth) higher than 20 were further analyzed. SnpEff program (Cingolani et al. 2012) was used to categorize the effects of the variants in the genomes (Table S1, Table S4). The program annotates the variants based on their genomic location (intron, exon, untranslated region, upstream, downstream, splice site, or intergenic regions) including in the Almond GFF file extracted from the NCBI database (GCA_008632915.2). Then it predicts the coding effect such as synonymous or non-synonymous substitution, start or stop codon gains or losses, or frame shifts.

Population genotyping

Based on WGS of *P. arabica* and U.E.F, a SNP calling was performed in order to select SNPs that will detect polymorphism within the F1 population. The following criteria were set: (1)

remove sites with DP lower than 20; (2) an isolated SNP over 100 bp interval; (3) the SNP is unique with no matching on other genomic regions on the reference genome; (4) informative SNPs for the F1 population, that are homozygous for one parent and heterozygous for the other. In addition, SNPs were chosen at intervals of 40 kb along the almond genome (*P. dulcis* cv. Lauranne; https://www.ncbi.nlm.nih.gov/bioproject/553424) to obtain an unbiased representation through the whole chromosomes. Overall, a set of 5,000 markers was selected for genotyping (Fig. 3). The F1 population screening was accomplished by "targeted SNP Seq" by LGC (LGC Genomics, Germany) for SNPs genotyping.

Genetic map construction

For generating the genetic map the JoinMap®4.1 software (J. W. & J. J. Van Ooijen 2013) was used. Cross-pollination population type was used with the code lmxll for markers that were homozygous for the male parents (*P. arabica*) and heterozygous in the female parent (U.E.F), and the code nnxnp for the opposite case. Because there were no common markers (hkxhk) we did not combine the two marker types, and undertook the pseudo test cross method (Swinburne and Lindgren 2013), meaning we separated the markers into two different maps, one map for the U.E.F (where *P. arabica* is homozygous- lmxll code), and one for the *P. arabica* parent (where U.E.F is homozygous nnxnp code). Markers were filtered for three parameters: (1) More than ~11% missing data; (2) Non-Mendelian segregation (X^2>6.5, DF=1) (3) Remove markers in similarity of 1.0. The "Independence LOD" algorithm was used for linkage groups clustering (LOD>8), and the Kosambi's function was chosen for calculating genetic distance.

QTL mapping

In order to conduct the QTL analysis we used the Map QTL®5 software(J. W. Van Ooijen 2006). QTLs and their significance were calculated using interval mapping (IM). A QTL was determined as significant when its LOD score was higher than the calculated threshold (1000 permutation at α=0.05), and the QTL spanning was determined by ±1 LOD from the max LOD marker.

Genome wide association study (GWAS)

Association was calculated by TASSEL 5.2.59(Bradbury et al. 2007). The set of SNPs was filtered; marker discard when missing data was >8.6%, and the allele frequency was set for $0.2<x<0.8$ for preventing overestimated impact of rare alleles. The General linear model (GLM) was applied for the phenotypic and genotypic intersect data set to test the association. Threshold for significance result was assessed by 1000 permutation test α=0.05.

Statistics

All significance tests were done by the statistical software JMP (JMP® PRO 15.0.0 © 2019 SAS Institute Inc.), α=0.05. To test significance when the variance was unequal, a simple T-test was used, and if it was equal, the pooled t- Anova test was performed. Tukey - Kremer's test was used to analyze variance in the population when the distribution was normal, and the variance inside the groups was equal; when it was not equal or normal, Wilcoxon non-parametric test was used. Broad sense heritability of the SPC was calculated on the F1 (full sibs) by the 'Rsquare adj' value, given by a simple Anova test.

553 *Data availability*

554 All data supporting the results are specified in the manuscript or in the supplementary data. *P.*
555 *arabica* leaf net CO_2 assimilation, SPC distribution analysis, and correlation analysis between
556 the SPC and cork development are available from the corresponding author upon reasonable
557 request.

558

559 *Conflict of Interest*

560 The authors declare no conflict of interest.

561

562 *Author Contributions*

563 HB designed and conducted all experiments and wrote the manuscript. ADF processed and
564 assembled the row sequences of the parent's DNA, the SNPs calling, analyzing the SNPs
565 effect and all genes' annotation. IBY and RHB generated the F1 population. TAS, ZA, and
566 HB developed the infrastructure for measuring stem gas exchange. DH is the corresponding
567 author. Designed the experiments, supervised the study, and wrote the manuscript. All authors
568 discussed and commended on the manuscript.

569

570 *Funding*

571 This study supported by the "LEONA M. AND HARRY B. HELMSLEY CHARITABLE
572 TRUST"

Acknowledgment

Thanks to Kamel Hatib for planting, grafting and maintaining the experimental orchard. Thanks to Dr. Elad Oren for the close guidance in designing the SNPs panel and processing the genotyping results. The authors thank the generous support of the Leona M. and Harry B. Helmsley Charitable Trust for their help in funding of the current research.

References

Aschan, Guido, and Hardy Pfanz. 2003. "Non-Foliar Photosynthesis - A Strategy of Additional Carbon Acquisition." *Flora* 198 (2): 81–97. https://doi.org/10.1078/0367-2530-00080.

Atkinson, C. J., R. M. Brennan, and H. G. Jones. 2013. "Declining Chilling and Its Impact on Temperate Perennial Crops." *Environmental and Experimental Botany* 91: 48–62. https://doi.org/10.1016/j.envexpbot.2013.02.004.

Ávila-Lovera, Eleinis, Roxana Haro, Exequiel Ezcurra, and Louis S. Santiago. 2019. "Costs and Benefits of Photosynthetic Stems in Desert Species from Southern California." *Functional Plant Biology* 46 (2): 175–86. https://doi.org/10.1071/FP18203.

Ávila-Lovera, Eleinis, Antonio J. Zerpa, and Louis S. Santiago. 2017. "Stem Photosynthesis and Hydraulics Are Coordinated in Desert Plant Species." *New Phytologist* 216 (4): 1119–29. https://doi.org/10.1111/nph.14737.

Baerdemaeker, Niels J.F. De, Roberto Luis Salomón, Linus De Roo, and Kathy Steppe. 2017. "Sugars from Woody Tissue Photosynthesis Reduce Xylem Vulnerability to Cavitation." *The New Phytologist* 216 (3): 720–27. https://doi.org/10.1111/nph.14787.

595 Ballester, J., R. Socias I Company, P. Arus, and M. C. De Vicente. 2001. “Genetic Mapping of a
596 Major Gene Delaying Blooming Time in Almond.” *Plant Breeding* 120 (3): 268–70.
597 https://doi.org/10.1046/j.1439-0523.2001.00604.x.

598 Bloemen, Jasper, Lidewei L. Vergeynst, Lander Overlaet-Michiels, and Kathy Steppe. 2016. “How
599 Important Is Woody Tissue Photosynthesis in Poplar during Drought Stress?” *Trees - Structure*
600 *and Function* 30 (1): 63–72. https://doi.org/10.1007/s00468-014-1132-9.

601 Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for
602 Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20.
603 https://doi.org/10.1093/bioinformatics/btu170.

604 Bradbury, Peter J., Zhiwu Zhang, Dallas E. Kroon, Terry M. Casstevens, Yogesh Ramdoss, and
605 Edward S. Buckler. 2007. “TASSEL: Software for Association Mapping of Complex Traits in
606 Diverse Samples.” *Bioinformatics* 23 (19): 2633–35.
607 https://doi.org/10.1093/bioinformatics/btm308.

608 Cingolani, Pablo, Adrian Platts, Le Lily Wang, Melissa Coon, Tung Nguyen, Luan Wang, Susan J.
609 Land, Xiangyi Lu, and Douglas M. Ruden. 2012. “A Program for Annotating and Predicting the
610 Effects of Single Nucleotide Polymorphisms, SnpEff: SNPs in the Genome of Drosophila
611 Melanogaster Strain W1118; Iso-2; Iso-3.” *Fly* 6 (2): 80–92. https://doi.org/10.4161/fly.19695.

612 Delplancke, Malou, Mariana Yazbek, Nils Arrigo, Anahí Espíndola, Helene Joly, and Nadir Alvarez.
613 2016. “Combining Conservative and Variable Markers to Infer the Evolutionary History of
614 Prunus Subgen. Amygdalus s.l. under Domestication.” *Genetic Resources and Crop Evolution*
615 63 (2): 221–34. https://doi.org/10.1007/s10722-015-0242-6.

616 Depristo, Mark A., Eric Banks, Ryan Poplin, Kiran V. Garimella, Jared R. Maguire, Christopher
617 Hartl, Anthony A. Philippakis, et al. 2011. “A Framework for Variation Discovery and

618 Genotyping Using Next-Generation DNA Sequencing Data." *Nature Genetics* 43 (5): 491–501.
619 https://doi.org/10.1038/ng.806.

620 Fernandez, Eduardo, Graeme Baird, Daniela Farías, Eduardo Oyanedel, José A. Olaeta, Patrick
621 Brown, Maciej Zwieniecki, Aude Tixier, and Sebastian Saa. 2018. "Fruit Load in Almond Spurs
622 Define Starch and Total Soluble Carbohydrate Concentration and Therefore Their Survival and
623 Bloom Probabilities in the next Season." *Scientia Horticulturae* 237 (April): 269–76.
624 https://doi.org/10.1016/j.scienta.2018.04.030.

625 Gabay, Gilad, Yardena Dahan, Yacov Izhaki, Adi Faigenboim, Giora Ben-Ari, Yonatan Elkind, and
626 Moshe A. Flaishman. 2018. "High-Resolution Genetic Linkage Map of European Pear (Pyrus
627 Communis) and QTL Fine-Mapping of Vegetative Budbreak Time." *BMC Plant Biology* 18 (1):
628 1–13. https://doi.org/10.1186/s12870-018-1386-2.

629 Gómez, Eva María, Federico Dicenta, Ignasi Batlle, Agustí Romero, and Encarnación Ortega. 2019.
630 "Cross-Incompatibility in the Cultivated Almond (Prunus Dulcis): Updating, Revision and
631 Correction." *Scientia Horticulturae* 245 (July 2018): 218–23.
632 https://doi.org/10.1016/j.scienta.2018.09.054.

633 Gradziel, T. M., P. Martínez-Gómez, F. Dicenta, and D. E. Kester. 2001. "The Utilization of Related
634 Prunus Species for Almond Variety Improvement." *Fruit Varieties Journal* 55 (2): 100–108.

635 Granot, David, Rakefet David-Schwartz, and Gilor Kelly. 2013. "Hexose Kinases and Their Role in
636 Sugar-Sensing and Plant Development." *Frontiers in Plant Science* 4 (MAR): 1–17.
637 https://doi.org/10.3389/fpls.2013.00044.

638 Guo, Chunmiao, Yu Wei, Bo Yang, Mubarek Ayup, Ning Li, Jun Liu, Kang Liao, and Huan Wang.
639 2021. "Developmental Transcriptome Profiling Uncovered Carbon Signaling Genes Associated
640 with Almond Fruit Drop." *Scientific Reports* 11 (1): 1–12. https://doi.org/10.1038/s41598-020-

69395-z.

Hara, Kenta, Toshiya Yokoo, Ryoko Kajita, Takaaki Onishi, Saiko Yahata, Kylee M. Peterson, Keiko U. Torii, and Tatsuo Kakimoto. 2009. "Epidermal Cell Density Is Autoregulated via a Secretory Peptide, EPIDERMAL PATTERNING FACTOR 2 in Arabidopsis Leaves." *Plant and Cell Physiology* 50 (6): 1019–31. https://doi.org/10.1093/pcp/pcp068.

Hurtado, M. A., C. Romero, S. Vilanova, A. G. Abbott, G. Llácer, and M. L. Badenes. 2002. "Genetic Linkage Maps of Two Apricot Cultivars (Prunus Armeniaca L.), and Mapping of PPV (Sharka) Resistance." *Theoretical and Applied Genetics* 105 (2–3): 182–91. https://doi.org/10.1007/s00122-002-0936-y.

Jung, Sook, Taein Lee, Chun Huai Cheng, Katheryn Buble, Ping Zheng, Jing Yu, Jodi Humann, et al. 2019. "15 Years of GDR: New Data and Functionality in the Genome Database for Rosaceae." *Nucleic Acids Research* 47 (D1): D1137–45. https://doi.org/10.1093/nar/gky1000.

Khadivi-Khub, Abdollah, and Karim Anjam. 2014. "Morphological Characterization of Prunus Scoparia Using Multivariate Analysis." *Plant Systematics and Evolution* 300 (6): 1361–72. https://doi.org/10.1007/s00606-013-0967-7.

Li, Heng, and Richard Durbin. 2009. "Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform." *Bioinformatics* 25 (14): 1754–60. https://doi.org/10.1093/bioinformatics/btp324.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. 2009. "The Sequence Alignment/Map Format and SAMtools." *Bioinformatics* 25 (16): 2078–79. https://doi.org/10.1093/bioinformatics/btp352.

Miotto, Yohanna Evelyn, Carolina Tessele, Ana Beatriz Costa Czermainski, Diogo Denardi Porto, Vítor da Silveira Falavigna, Tiago Sartor, Amanda Malvessi Cattani, et al. 2019. "Spring Is

664 Coming: Genetic Analyses of the Bud Break Date Locus Reveal Candidate Genes from the Cold
665 Perception Pathway to Dormancy Release in Apple (Malus × Domestica Borkh.).” *Frontiers in*
666 *Plant Science* 10. https://doi.org/10.3389/fpls.2019.00033.

667 Nilsen, ET. 1995. “Stem Photosynthesis: Extent, Patterns and Rol in Plant Carbon Economy.” *Plant*
668 *Stems: Physiology and Functional Morphology.*, 223–40.

669 Olukolu, Bode A., Taly Trainin, Shenghua Fan, Chittaranjan Kole, Douglas G. Bielenberg, Gregory
670 L. Reighard, Albert G. Abbott, and Doron Holland. 2009. “Genetic Linkage Mapping for
671 Molecular Dissection of Chilling Requirement and Budbreak in Apricot (Prunus Armeniaca
672 L.).” *Genome* 52 (10): 819–28. https://doi.org/10.1139/G09-050.

673 Ooijen, J. W. & J. Jansen Van. 2013. “JoinMap®4.1, Software for the Calculation of Genetic
674 Linkage Maps in Experimental Populations of Diploid Species. Kyazma B.V. Wageningen,
675 Netherland.” Wageningen, Netherland: Kyazma B.V.

676 Ooijen, J. W. Van. 2006. “Map QTL 5, Software for the Mapping of Quantitative Trait Loci in
677 Experimantal Population. Kyazma B.V. Wageningen, Netherlands.” Wageningen, Netherland.

678 Oren, Elad, Galil Tzuri, Asaf Dafna, Ayala Meir, Ravindra Kumar, Nurit Katzir, Yonatan Elkind, et
679 al. 2020. “High-Density NGS-Based Map Construction and Genetic Dissection of Fruit Shape
680 and Rind Netting in Cucumis Melo.” *Theoretical and Applied Genetics* 133 (6): 1927–45.
681 https://doi.org/10.1007/s00122-020-03567-3.

682 Pérez de los Cobos, Felipe, Pedro J. Martínez-García, Agustí Romero, Xavier Miarnau, Iban
683 Eduardo, Werner Howad, Mourad Mnejja, et al. 2021. “Pedigree Analysis of 220 Almond
684 Genotypes Reveals Two World Mainstream Breeding Lines Based on Only Three Different
685 Cultivars.” *Horticulture Research* 8 (1). https://doi.org/10.1038/s41438-020-00444-4.

686 Rajabpoor, Shakiba, Soghra Kiani, Karim Sorkheh, and Farahnaz Tavakoli. 2014. “Changes Induced

by Osmotic Stress in the Morphology, Biochemistry, Physiology, Anatomy and Stomatal Parameters of Almond Species (Prunus L. Spp.) Grown in Vitro." *Journal of Forestry Research* 25 (3): 523–34. https://doi.org/10.1007/s11676-014-0491-9.

Roskov, Y. et al. 2019. "Species 2000 & ITIS Catalogue of Life, 2019 Annual Checklist." Digital Resource. 2019. https://doi.org/ISSN 2405-884X.

Sánchez-Pérez, Raquel, Jorge Del Cueto, Federico Dicenta, and Pedro Martínez-Gómez. 2014. "Recent Advancements to Study Flowering Time in Almond and Other Prunus Species." *Frontiers in Plant Science* 5 (JUL): 1–7. https://doi.org/10.3389/fpls.2014.00334.

Soler, Marçal, Olga Serra, Marisa Molinas, Gemma Huguet, Silvia Fluch, and Mercè Figueras. 2007. "A Genomic Approach to Suberin Biosynthesis and Cork Differentiation." *Plant Physiology* 144 (1): 419–31. https://doi.org/10.1104/pp.106.094227.

Sorkheh, K., B. Shiran, M. Khodambshi, V. Rouhi, and S. Ercisli. 2011. "In Vitro Assay of Native Iranian Almond Species (Prunus L. Spp.) for Drought Tolerance." *Plant Cell, Tissue and Organ Culture* 105 (3): 395–404. https://doi.org/10.1007/s11240-010-9879-1.

Sorkheh, K., B. Shiran, V. Rouhi, E. Asadi, H. Jahanbazi, H. Moradi, T. M. Gradziel, and P. Martínez-Gómez. 2009. "Phenotypic Diversity within Native Iranian Almond (Prunus Spp.) Species and Their Breeding Potential." *Genetic Resources and Crop Evolution* 56 (7): 947–61. https://doi.org/10.1007/s10722-009-9413-7.

Sperling, Or, Tamir Kamai, Aude Tixier, Anna Davidson, Katherine Jarvis-Shean, Eran Raveh, Ted M. DeJong, and Maciej A. Zwieniecki. 2019. "Predicting Bloom Dates by Temperature Mediated Kinetics of Carbohydrate Metabolism in Deciduous Trees." *Agricultural and Forest Meteorology* 276–277 (April): 107643. https://doi.org/10.1016/j.agrformet.2019.107643.

Sperling, Or, Lucas C.R. Silva, Aude Tixier, Guillaume Théroux-Rancourt, and Maciej A.

710 Zwieniecki. 2017. “Temperature Gradients Assist Carbohydrate Allocation within Trees.”
711 *Scientific Reports* 7 (1): 1–10. https://doi.org/10.1038/s41598-017-03608-w.

712 Swinburne, June, and Gabriella Lindgren. 2013. “Genetic Linkage Maps.” *Equine Genomics* 1137:
713 11–47. https://doi.org/10.1002/9781118522158.ch2.

714 Tixier, Aude, Or Sperling, Jessica Orozco, Bruce Lampinen, Adele Amico Roxas, Sebastian Saa, J.
715 Mason Earles, and Maciej A. Zwieniecki. 2017. “Spring Bud Growth Depends on Sugar
716 Delivery by Xylem and Water Recirculation by Phloem Münch Flow in Juglans Regia.” *Planta*
717 246 (3): 495–508. https://doi.org/10.1007/s00425-017-2707-7.

718 Vulavala, Vijaya K.R., Edna Fogelman, Adi Faigenboim, Oded Shoseyov, and Idit Ginzberg. 2019.
719 “The Transcriptome of Potato Tuber Phellogen Reveals Cellular Functions of Cork Cambium
720 and Genes Involved in Periderm Formation and Maturation.” *Scientific Reports* 9 (1): 1–14.
721 https://doi.org/10.1038/s41598-019-46681-z.

722 Wang, Y., L. L. Georgi, G. L. Reighard, R. Scorza, and A. G. Abbott. 2002. “Genetic Mapping of the
723 Evergrowing Gene in Peach [Prunus Persica (L.) Batsch].” *Journal of Heredity* 93 (5): 352–58.
724 https://doi.org/10.1093/jhered/93.5.352.

725 Yazbek, M., and S. H. Oh. 2013. “Peaches and Almonds: Phylogeny of Prunus Subg. Amygdalus
726 (Rosaceae) Based on DNA Sequences and Morphology.” *Plant Systematics and Evolution* 299
727 (8): 1403–18. https://doi.org/10.1007/s00606-013-0802-1.

728 Zhang, Jian, Jingjing Yang, Like Zhang, L Jiang, Hong Zhao, and Jianan Zhang. 2020. “OPEN A
729 New SNP Genotyping Technology Target SNP-Seq and Its Application in Genetic Analysis of
730 Cucumber Varieties,” 1–11. https://doi.org/10.1038/s41598-020-62518-6.

731 Zwieniecki, Maciej A., Aude Tixier, and Or Sperling. 2015. “Temperature-Assisted Redistribution of
732 Carbohydrates in Trees.” *American Journal of Botany* 102 (8): 1216–18.

https://doi.org/10.3732/ajb.1500218.

Tables

Table 1. Quality data from whole genome sequencing of *P. arabica* and U.E.F. Quality parameters (Q20, Q30) from sequencing of each parent is presented with respect to the reference genomes of *P. dulcis* cv. Lauranne and *P. dulcis* cv. Texas. The row sequence data was mapped to each of the reference genomes. The total variant sites (SNPs and InDels) between each species and cv. Lauranne reference genome are also presented.

Species	Average coverage	Q20	Q30	% Mapping VS Lauranne	% Mapping VS Texas	Total variant sites	% Heterozygous variant sites
P. arabica	~ X 57	94.5	87.37	97.93	85.94	3,750,363	35.18
U.E.F *(P.dulcis)*	~ X 55.5	94.7	87.71	98.37	89.82	2,407,787	69.29

Table 2. Characteristics of the established genetic maps. U.E.F map (1) was generated from markers that were homozygous in *P. arabica* (male) and heterozygous in the U.E.F (female). *P. arabica* map (2) was generated from markers that were homozygous for U.E.F and heterozygous for *P. arabica*.

	Map type		number of markers	LGs number	Total length (cM)	marker density (cM)			
	U.E.F	P. arabica				average	median	maximum distance	standard error
1	Aa	aa	971	8	504.64	0.533	0.375	6.173	0.017
2	AA	Aa	572	8	568.76	1.093	0.768	18.854	0.078

~~**Table 3. Summary of candidate genes' list.** The annotated genes within the regions spanning the two QTLs are presented. Only genes that have non- synonymous variants between the two parents are presented. Genes are divided for polymorphism region as determined by the D.N.A sequence.~~

~~Locus~~	~~Coding region~~	~~Non-coding region~~	~~Total~~
~~1~~	~~17~~	~~96~~	~~113~~
~~7~~	~~54~~	~~282~~	~~336~~
~~**Total**~~	~~**71**~~	~~378~~	~~449~~

Figure captions

Figure 1. Gas exchange measurements of *P. arabica* and *Prunus dulcis* cv. Um el Fachem (U.E.F.) stems. Three years old trees of the Israeli cultivar *Prunus dulcis* cv. U.E.F (a), and the wild almond *P. arabica* (b) at spring. Stems of U.E.F. (c), and *P. arabica* (d) annually developed: one-year old (c1, d1), second year (c2, d2) and three years old stems (c3, d3). Gas exchange data of one-year old stems along the year (e) of *P. arabica* (solid gray line), and U.E.F (dashed black line). Each dot denotes the average of two independent days of measuring for each month (n=8). Stem respiration rate in response to three different temperatures (f) of *P. arabica* (grey bar) and U.E.F (black bar). Four stems were measured for each genotype in each temperature (n=4). Different capital letter represents significance (α=0.05) between temperatures, not between species. The error bars represent ± SE.

Figure 2. Stem photosynthetic capability (SPC) in the F1 progeny. Levels of net CO_2 assimilation for each offspring (a). Each box plot presents the average of four stems (n=4). The F1 progeny parents, *P. arabica* and U.E.F are marked by dashed arrow and simple arrow, respectively. Distribution histogram of the same data is presented in (b), while parents' data is highlighted. Measurements were conducted during February 2020 while the trees were dormant. Representative pictures of the F1 population while dormant (c) in February, and during the vegetative phase in April (d).

Figure 3. Allelic frequency among the F1 population. Allelic frequency of the heterozygous allele in the F1 population for each marker. Blue line indicates the *P. arabica* allelic incidence, and red line indicates the U.E.F allelic incidence. X axis is the physical position in Mb, and Y axis represents the allelic ratio (from 0 to 1). Black arrow in chromosome 3 represents an example of un-expected deviated region ("hot spot"). References lines presented the 'tolerance interval' limits.

Figure 4. Graphic presentation of markers density and distribution along the eight linkage groups. Comparison between U.E.F map (left graph) and *P. arabica* map (right graph). Each horizontal line represents a single marker.

Figure 5. Comparison between the genetic and physical order of the markers. SNP markers were placed according to their physical position on the Lauranne reference genome sequence (X axis), and their position on the U.E.F genetic map (Y axes). The yellow dots represent markers from unplaced scaffolds, according to the Lauranne reference genome (chr-0). Those markers were mapped to several chromosomes in this study based on the genetic map data.

Figure 6. QTLs and GWAS analysis for the SPC trait. Major QTL of 2.4 Cm width and LOD score of 20.8, was detected in chromosome 7 by using the U.E.F genetic map (a). Minor QTL of 4.4 Cm width, and LOD score (3.9) was located at the end of chromosome 1 by using the *P. arabica* genetic map (b). Results of GWAS using the whole set of markers (3,800) sorted by their physical position according to the reference genome (c), revealed both loci in chromosome 7 and chromosome 1. Markers in c were sorted by their physical position according to the reference genome. Markers that were placed on chromosome 0 and found as highly associated to the SPC trait are also shown in (c). The horizontal dash line represents significance level according to permutation test (1000 times at α=0.05).

Figure 7. QTLs effect on net CO_2 assimilation and the synergistic effect between QTLs. Least square of means for each allelic combination is presented (a). Each box plot represents the population individual's average phenotype-grouped for their allelic combination. Y axis is the level of net CO_2 assimilation. The X axis represents the allelic combination. On the X axis, the capital letter A, refers to individuals with *P. arabica* allele combination, U refers to

individuals with U.E.F allele combination in each one of the QTL, and superscript numbers (7, 1) present the QTL identity. Numerical presentation of the data is presented in a (b). *P. arabica* allele combination is marked in green, and U.E.F marker combination is marked in red. Different letters indicate significance (α=0.05). Statistical evaluation of each QTL effect, and interaction between the two loci (c). Presented results were analyzed by the "Full factorial test", blue line is equivalent for p-value of α=0.01.

Figure 1.TIF

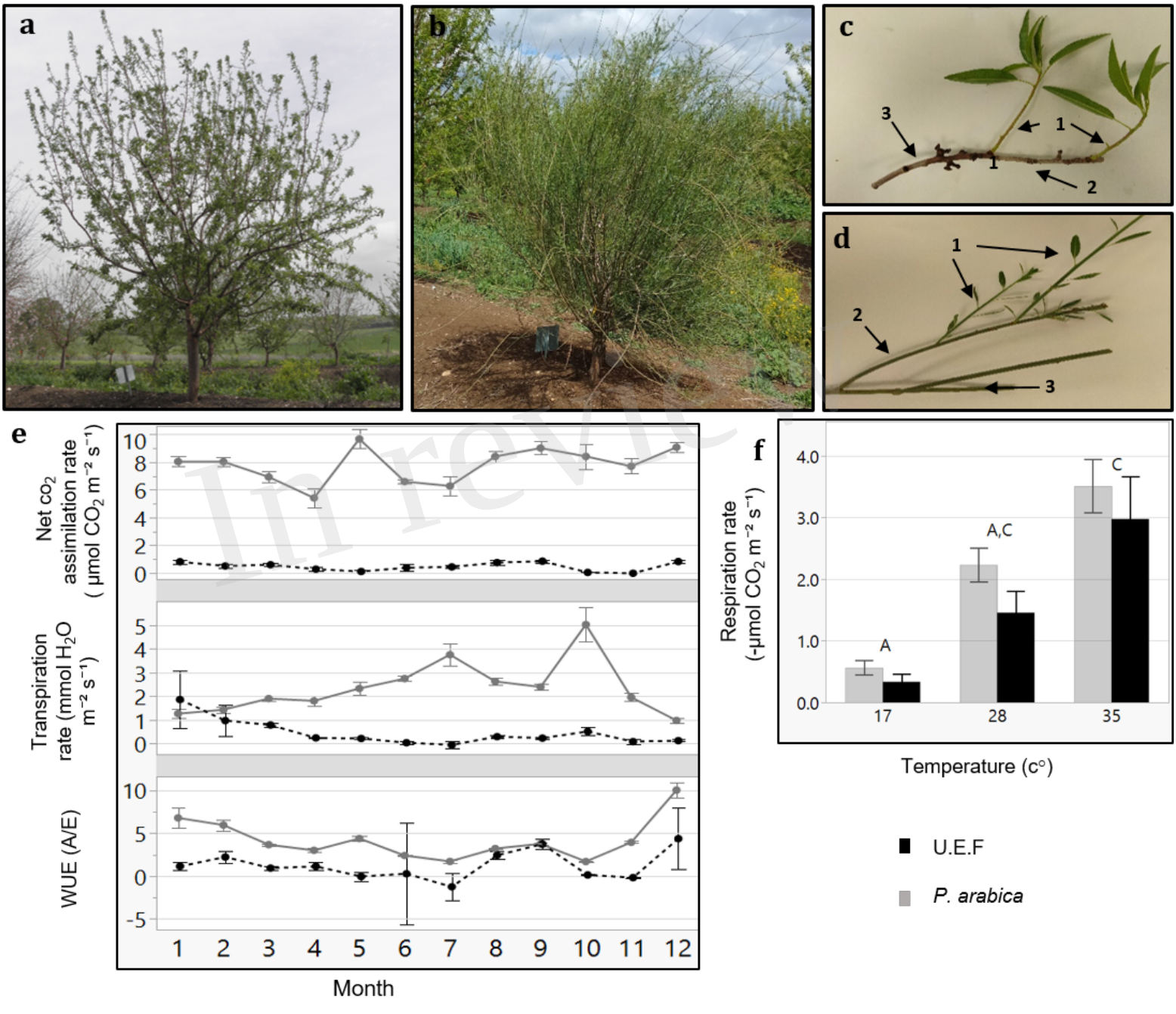

a
b
c
3
1
1
2
d
1
2
3
e
Net co2 assimilation rate (µmol CO2 m−2 s−1)
Transpiration rate (mmol H2O m−2 s−1)
WUE (A/E)
10
8
6
4
2
0
5
4
3
2
1
0
10
5
0
-5
1 2 3 4 5 6 7 8 9 10 11 12
Month
f
Respiration rate (-µmol CO2 m−2 s−1)
4.0
3.0
2.0
1.0
0.0
A
A,C
C
17
28
35
Temperature (c°)
U.E.F
P. arabica

Figure 2.TIF

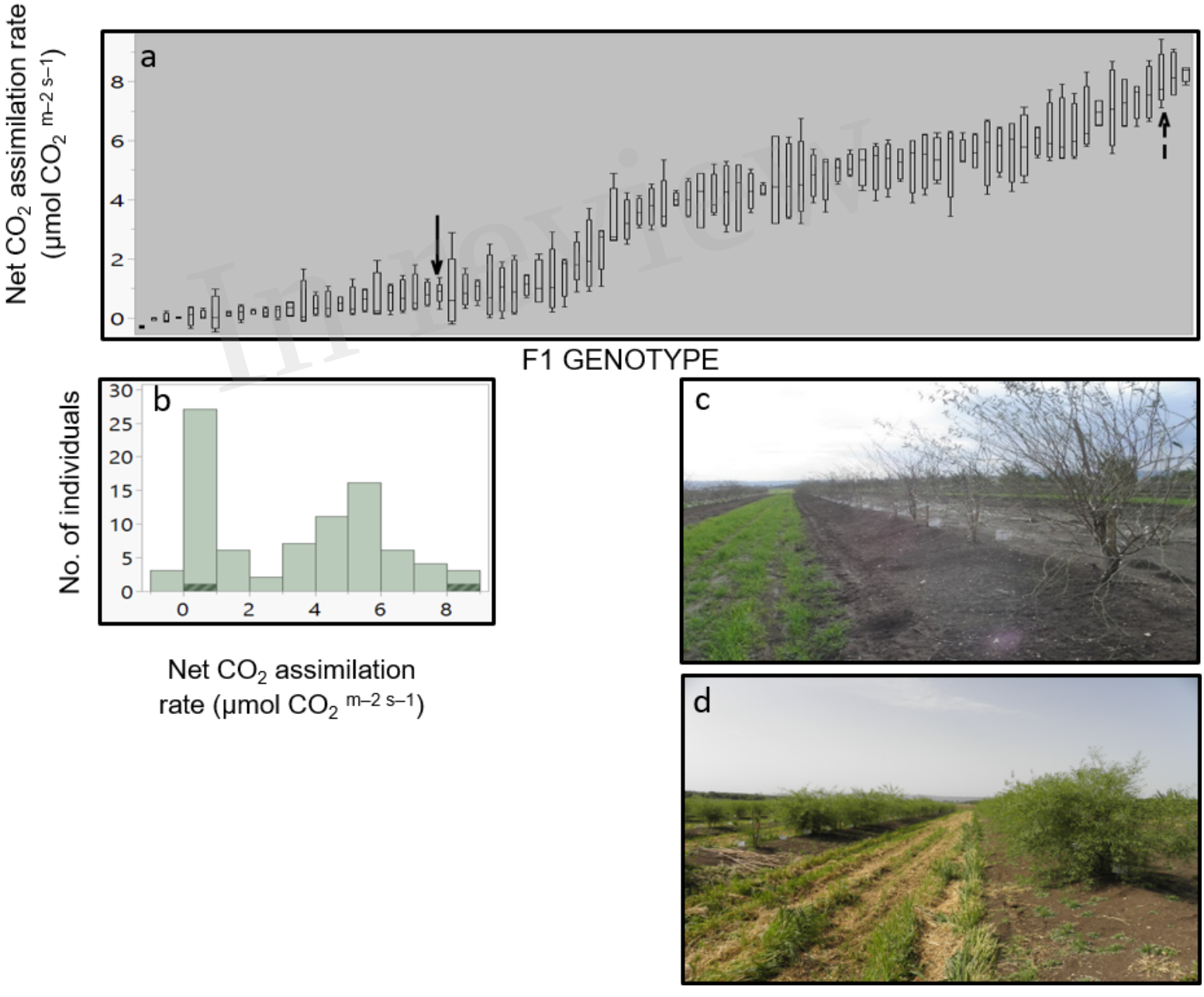

a
Net CO_2 assimilation rate (µmol CO_2 $^{m-2\ s-1}$)
0
2
4
6
8
F1 GENOTYPE
b
No. of individuals
0
5
10
15
20
25
30
0
2
4
6
8
Net CO_2 assimilation rate (µmol CO_2 $^{m-2\ s-1}$)
c
d

Figure 3.TIF

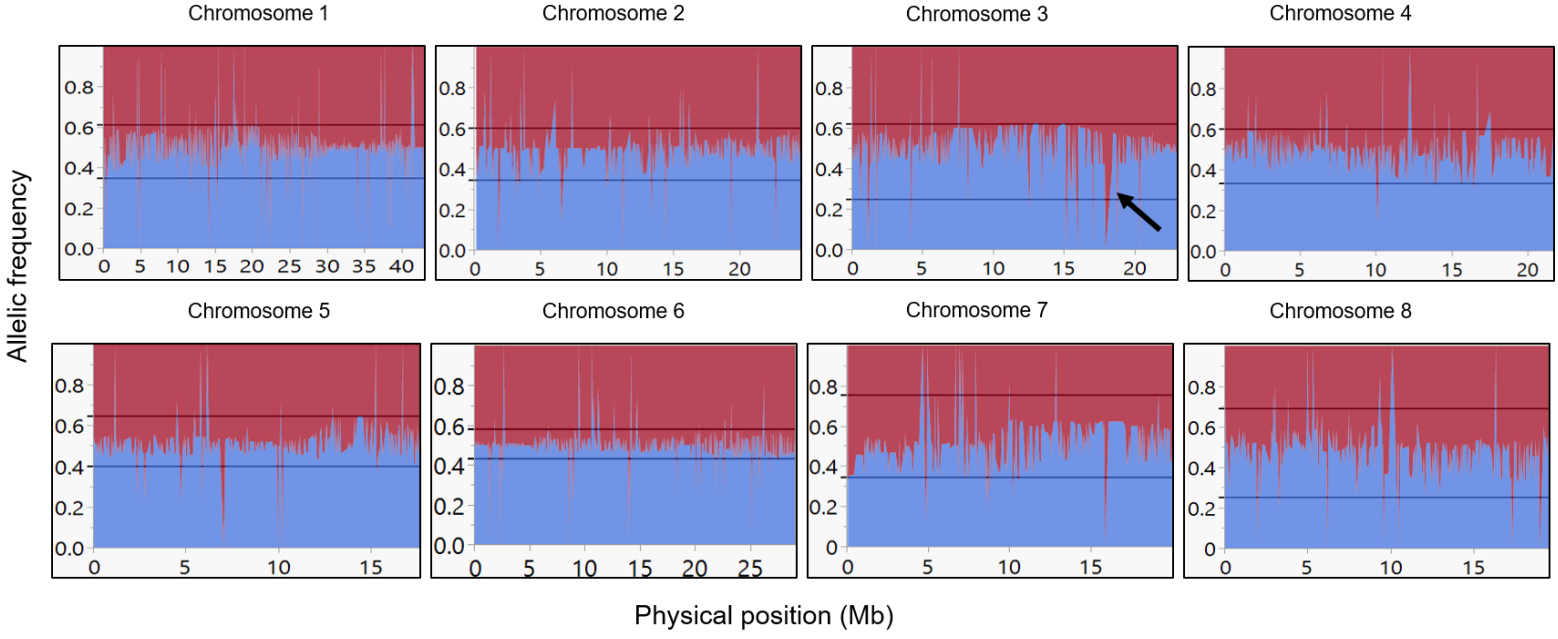


Figure 4.TIF

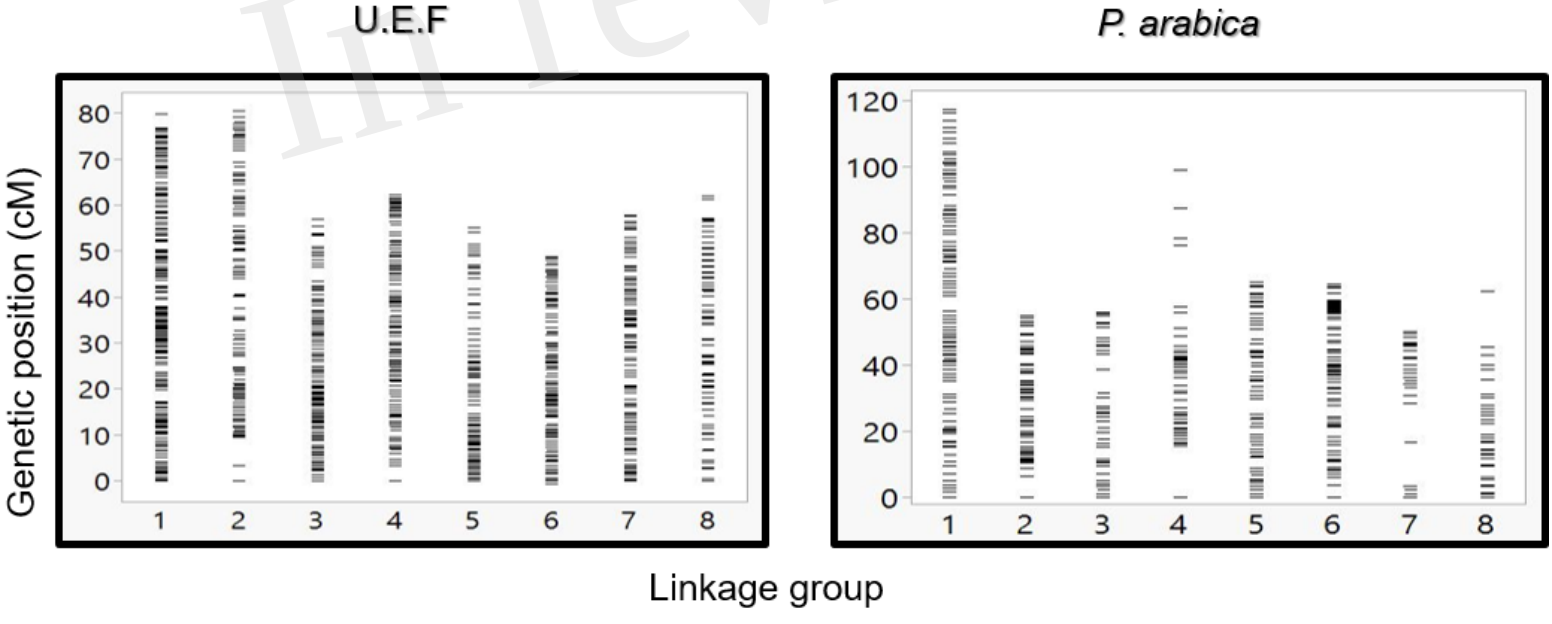


Figure 5.TIF

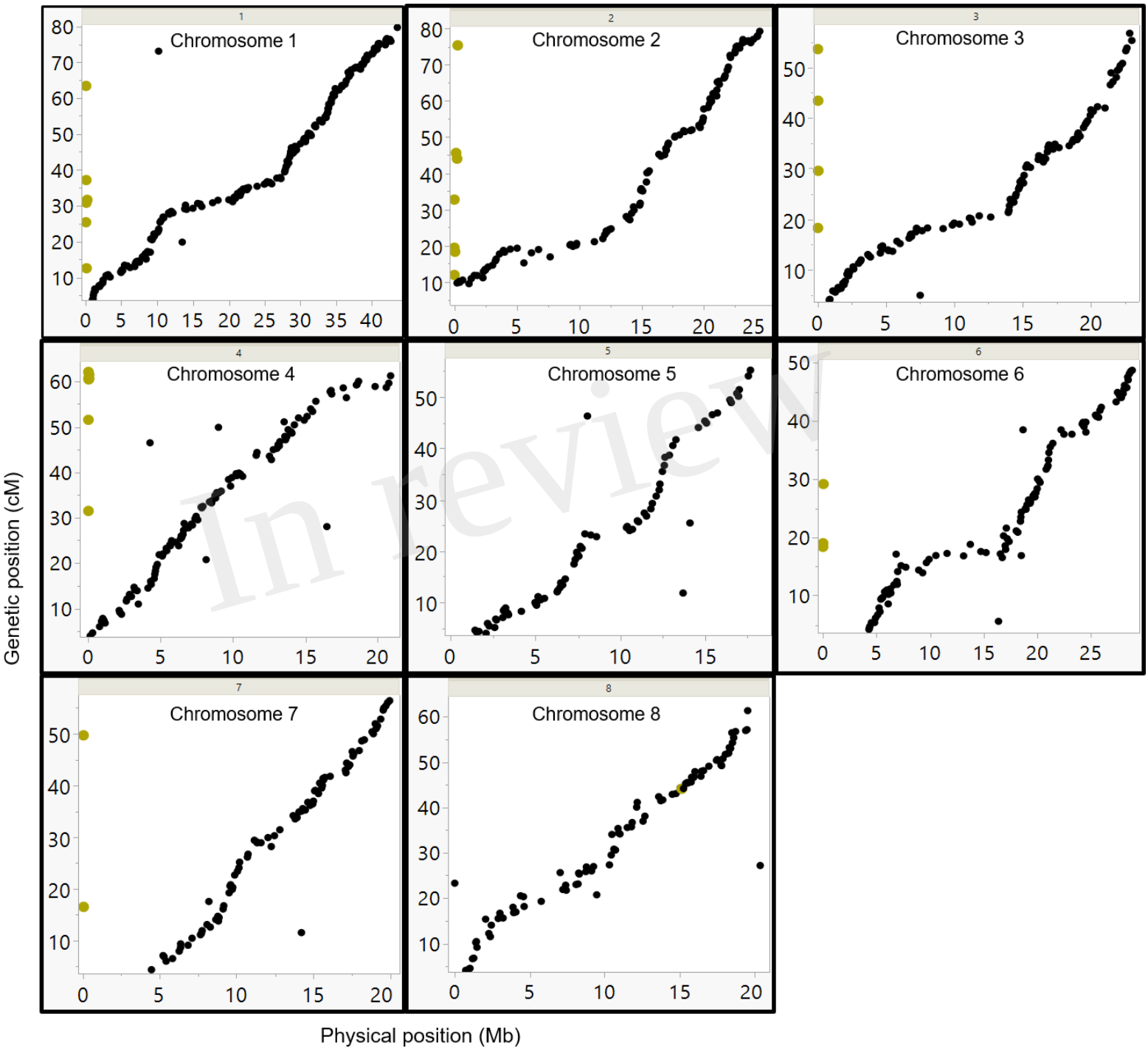


Figure 6.TIF

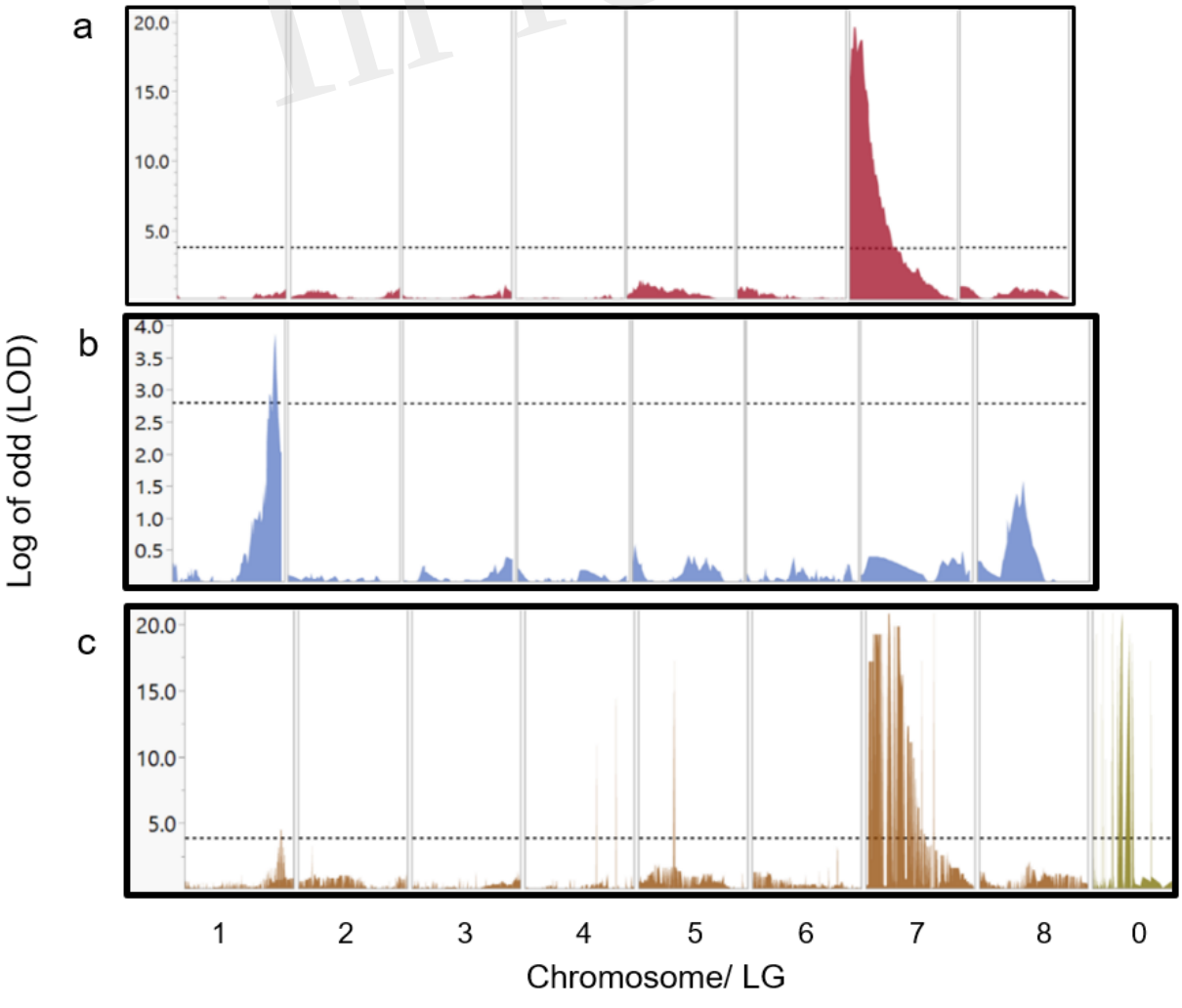
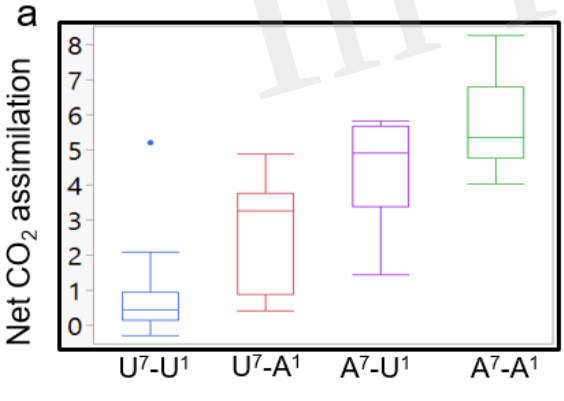


Figure 7.TIF

a



b

Allelic state		Least Sq Mean (net CO_2 assimilation)	SE
Locus 7 (chr 7)	Locus 1 (chr 1)		
A	A	[a]5.7339041	0.250632
A	U	[b]4.4461958	0.389623
U	A	[c]1.827375	0.318126
U	U	[c]0.8019354	0.275505

c

Effect Summary

Source	LogWorth	PValue
LOCUS 7	18.877	0.00000
LOCUS 1	3.395	0.00040
LOCUS 1*LOCUS 7	0.170	0.67633