Natural variation in circadian rhythms and photoperiodic flowering in tomato and its wild relatives

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Niels Müller
aus Heidelberg

Köln, 2015

Diese Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Arbeitsgruppe von Dr. José Jiménez-Gómez in der Abteilung für Pflanzenzüchtung und Genetik von Prof. Dr. Maarten Koornneef durchgeführt.

Berichterstatter:	Prof. Dr. Maarten Koornneef Prof. Dr. Ute Höcker
Prüfungsvorsitzender:	Prof. Dr. Siegfried Roth
Tag der mündlichen Prüfung:	28.05.2014

Abstract

The circadian clock controls many important aspects of plant physiology and development. By allowing anticipation of predictable daily and annual seasonal changes it helps plants to appropriately align their daily cycles and to correctly time their development. In this thesis, natural variation in daily and seasonal outputs of the circadian clock was studied in cultivated tomato (*Solanum lycopersicum*) and its wild relatives.

A key seasonal output of the circadian clock is photoperiodic flowering, which is crucial for many plants, either to synchronize for pollination or to avoid unfavorable conditions. While all cultivated tomato varieties flower irrespective of daylength, some of its wild relatives, as *S. galapagense* and *S. habrochaites*, accelerate flowering under short days. Under the long day greenhouse conditions used in this thesis, *S. galapagense* does not flower at all unless grafted onto cultivated tomato. This indicates that in this species a mobile flowering signal is active only under short days. Bulk segregant analysis combined with RNA sequencing using an *S. galapagense* F2 population and quantitative trait locus (QTL) analysis using an *S. habrochaites* introgression line population revealed loci underlying photoperiodic flowering.

Besides photoperiodic flowering, the circadian clock controls many other agronomically significant traits. In addition, natural variation in circadian rhythms appears to be important for adaptation to specific environments. Still, quantitative variation in circadian rhythms due to artificial selection has not yet been reported. Monitoring of two daily outputs of the circadian clock, leaf movements and transcript rhythms, revealed differences between cultivated tomato and its wild relatives, indicating that domestication or early breeding has had an effect on the tomato circadian clock. QTL analysis identified two loci that appear to be responsible for these differences in circadian rhythms. Fine mapping of one of these loci determined a region containing 13 genes. One of them, the homolog of the *Arabidopsis thaliana* gene *EMPFINDLICHER IM DUNKELROTEN LICHT 1* (*EID1*), was cloned and transformed into tomato. Strikingly, near isogenic lines (NILs) differing only in *EID1* and 12 adjacent genes exhibit differences in seed-

ling height under greenhouse conditions. This indicates that the observed variation in circadian rhythms may have an effect on plant growth under natural conditions. In conclusion, this study suggests that humans may have selected for altered circadian rhythms during tomato domestication or early breeding to adapt the species to its agricultural environments.

Zusammenfassung

Die circadiane Uhr kontrolliert viele wichtige pflanzenphysiologische und entwicklungsbiologische Prozesse. Indem sie Antizipation von vorhersehbaren täglich oder jährlich wiederkehrenden Ereignissen ermöglicht, trägt sie dazu bei, dass Tagesrhythmen zur richtigen Tageszeit und Entwicklungsschritte zur richtigen Jahreszeit stattfinden. In dieser Doktorarbeit, wurde natürliche Variation circadianer Rhythmen und photoperiodischen Blühens in Tomate (*Solanum lycopersicum*) und ihren wilden Verwandten untersucht.

Ein Beispiel für ein durch die circadiane Uhr ermöglichtes jahreszeitliches Ereignis ist photoperiodisches Blühen. Dies ist für viele Pflanzen wichtig, um sich für die Bestäubung zeitlich abzustimmen oder um ungünstigen Wetterbedingungen zu entgehen. Während kultivierte Tomatenvarietäten unabhängig von der Tageslänge blühen, zeigen einige wilde Tomatenarten photoperiodisches Blühen. Die enge Verwandte *S. galapagense* zum Beispiel blüht nur unter Kurztagbedingungen oder wenn sie auf eine kultivierte Tomate gepfropft wird. Dies deutet darauf hin, dass in dieser Art ein mobiles Blühsignal nur unter Kurztagbedingungen aktiv ist. Auch in der entfernten Verwandten *S. habrochaites* wird Blühen durch lange Tage inhibiert. Quantitative Trait Locus (QTL) Analyse einer *S. habrochaites* Introgressionslinienpopulation und Bulk Segregant Analyse durch RNA Sequenzierung einer *S. galapagense* F2-Population zeigte Loci auf, die dem photoperiodischen Blühen zugrunde liegen.

Neben dem photoperiodischen Blühen, kontrolliert die circadiane Uhr auch viele andere Prozesse, welche in Nutzpflanzen agronomisch wichtige Eigenschaften darstellen. Außerdem ist natürliche Variation circadianer Rhythmen bei der Anpassung an spezifische Umweltbedingungen von Bedeutung. Trotzdem ist nicht bekannt, ob quantitative Unterschiede circadianer Rhythmen während der Pflanzenzüchtung künstlich selektiert wurden. Die Untersuchung circadianer Blattbewegungen und Transkriptrhythmen zeigte, dass die kultivierte Tomate andere Rhythmen aufweist als ihre wilden Verwandten. Dies deutet darauf hin, dass Domestizierung oder frühe Züchtung einen Einfluss auf die circadiane Uhr der kultivierten Tomate hatten. QTL Analyse identifizierte zwei Loci, die für die Unterschiede in circadianen Rhythmen verantwortlich zu sein scheinen. Feinkartierung eines dieser Loci bestimmte eine Region mit 13 Genen. Eines dieser Gene wurde kloniert und in Tomate transformiert – das Homolog des *Arabidopsis thaliana* Gens *EMPFINDLICHER IM*

DUNKELROTEN LICHT 1 (*EID1*). Bemerkenswerterweise offenbarten zwei Linien die sich nur für die 13 Gene unterschieden Größenunterschiede im Gewächshaus. Dies weist darauf hin, dass die Unterschiede der circadianen Rhythmen möglicherweise einen Effekt auf das Wachstum unter natürlichen Bedingungen haben. Abschließend lässt sich sagen, dass diese Arbeit darauf hindeutet, dass veränderte circadiane Rhythmen vom Menschen während der Domestizierung von Tomate oder der frühen Züchtung selektiert wurden, um die Art an die agrarwirtschaftliche Umwelt anzupassen.

Table of Contents

List of Abbreviations

A	adenine
Arabidopsis	*Arabidopsis thaliana*
bp	base pairs
BC1F2	first backcross F2
BIL	backcross inbred line
BRASS	biological rhythms analysis software system
BSA	bulked segregant analysis
CAPS	cleaved amplified polymorphic sequence
CCA1	CIRCADIAN CLOCK ASSOCIATED 1
CCD	charge-coupled device
CO	CONSTANS
cDNA	complementary DNA
cM	centi Morgan
Col-0	Columbia (USA) wild type
CT	circadian time
cv.	cultivated variety
DNA	deoxyribonucleic acid
EID1	EMPFINDLICHER IM DUNKELROTEN LICHT 1
ELF3	EARLY FLOWERING 3
eQTL	expression QTL
F1	first filial generation
F2	second filial generation
FKF1	FLAVIN BINDING, KELCH REPEAT, F-BOX 1
FT	FLOWERING LOCUS T
GATK	Genome Analysis Toolkit
GI	GIGANTEA
H^2	broad sense heritability
Hd1	Heading date 1
IL	introgression line
indel	insertion or deletion
kb	kilo base pairs
LD	long day (16h light / 8h dark)
LHY	LATE ELONGATED HYPOCOTYL
LOD	logarithm of the odds
Mb	Mega base pairs
MP-GC	Max Planck-Genome-centre Cologne

MPIPZ	Max Planck Institute for Plant Breeding Research
NaClO	sodium hypochlorite
Na_3PO_4	tri-sodium phosphate
NASC	Nottingham Arabidopsis Stock Centre
NIL	near isogenic line
NB	night break
PCR	polymerase chain reaction
PRR	pseudo response regulator
QTL	quantitative trait locus/loci
QTN	quantitative trait nucleotide
RAE	relative amplitude error
RH	relative humidity
RIL	recombinant inbred line
RNA	ribonucleic acid
RNAseq	RNA sequencing
rpm	revolutions per minute
S.	*Solanum*
S. l. cerasiforme	*Solanum lycopersicum* var. *cerasiforme*
SD	short day (8h light / 16h dark)
SEM	standard error of the mean
SNP	single nucleotide polymorphism
Taq	*Thermus aquaticus*
TGRC	tomato genetics resource center
TILLING	Targeting Induced Lesions IN Genomes
UTR	untranslated region
VCF	Variant Call Format
VIGS	virus induced gene silencing
ZT	Zeitgeber time

List of Figures

List of Tables

1 Introduction

1.1 Natural variation in tomato and its wild relatives – domestication and adaptive differentiation

The tomato clade (*Solanum* sect. *Lycopersicon* (Mill.) Wettst.) is part of the large Solanaceae family, which includes several major crops of great economic importance and a number of model plants used for scientific research (e.g. tomato – *Solanum lycopersicum*, potato – *Solanum tuberosum*, tobacco – *Nicotiana* spp. and sweet and chili peppers – *Capsicum* spp.) (Sarkinen et al., 2013). The clade consists of 13 closely related species including cultivated tomato *S. lycopersicum* (Figure 1). In 2012 it was estimated that cultivated tomato was the 9th most important crop plant in the world based on production quantity (161,793,834 tons) and the 4th most important crop plant based on production value ($59,108,521,000) (http://www.fao.org).

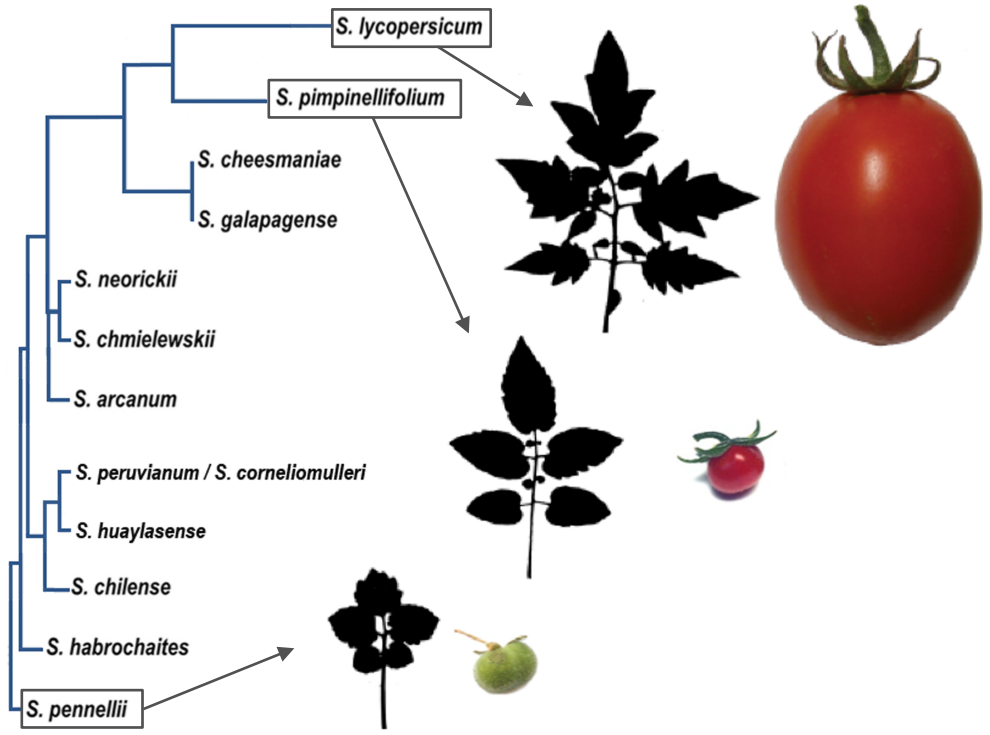

Figure 1: The tomato clade. Phylogenetic tree based on 4072 SNPs (modified from Viquez-Zamora et al. 2013; *S. peruvianum*, *S. corneliomulleri* and *S. huaylasense* were added according to Moyle (2008); their position in the tree is tentative). On the right, leaves and fruits of three representatives of the tomato clade are shown: cultivated tomato - *S. lycopersicum* cv. M82, its closest wild ancestor *S. pimpinellifolium* LA1589 and the distant relative *S. pennellii* LA0716.

The wild tomato species can be used to study naturally occurring variation that can be due to adaptive differentiation and of evolutionary significance. But more importantly, differences between all the wild species and the domesticated species can be analyzed. This allows the identification of the genetic and molecular bases underlying differences caused by domestication and crop evolution. This may lead to a better understanding of the domestication and breeding processes and might ultimately result in improved tomato varieties.

1.1.1 Tomato domestication and breeding

The place where tomato domestication took place has been debated for a long time and has still not been clarified. For more than 60 years now, there have been two alternative hypotheses: The Peruvian hypothesis assumes that tomato domestication occurred in South America in Peru (de Candolle, 1886), while the Mexican theory holds that wild tomato was transferred to Mexico first and subsequently domesticated there (Jenkins, 1948). Another open question regarding tomato domestication is whether the weedy wild cherry tomato, often classified as *S. lycopersicum* var. *cerasiforme* (*S. l. cerasiforme*), represents a transitional step from the wild species *S. pimpinellifolium* to the domesticated species *S. lycopersicum* and is therefore the closest wild ancestor of cultivated tomato, as was widely accepted 40 years ago (Simmonds, 1976), or whether it is an admixture of *S. lycopersicum* with *S. pimpinellifolium*. Several recent studies indicate that the latter is the case (Nesbitt and Tanksley, 2002; Ranc et al., 2008). Nevertheless, the other possibility cannot be ruled out, as some of the wild 'cerasiforme' accessions seem to be ancestral cultivars rather than the product of an admixture (Blanca et al., 2012).

Recent single nucleotide polymorphism (SNP) genotyping data from 272 tomato accessions belonging to the three closely related phylogenetic entities *S. pimpinellifolium*, *S. l. cerasiforme* and *S. lycopersicum*, from different regions of South America and Central America, suggest that tomato domestication took place in two steps, supporting both the Peruvian and the Mexican hypothesis. It is proposed that tomato domestication started with a pre-domestication process in the Andean region of Ecuador. Afterwards, this material was carried to Central America where the true domestication took place. The domesticated tomato was

then taken from Mesoamerica to Spain and Italy in the sixteenth century, and from there to the rest of the world (Blanca et al., 2012). Cultivated tomato therefore faced two major bottlenecks - first during domestication and then during the transfer to Europe and subsequent breeding. This resulted in substantial genetic erosion among cultivated tomato varieties (Miller and Tanksley, 1990; Ranc et al., 2008). Genetic erosion is typical for domesticated species due to domestication bottlenecks (Doebley et al., 2006). However, in cultivated tomato it appears to be even more severe than in other autogamous crops (Mazzucato et al., 2008). Especially vintage cultivars from the late 19th and early 20th century exhibit very low levels of genetic variation (Sim et al., 2012). As tomato can be hybridized with all of its relatives from the tomato clade, wild species germplasm provides a rich reservoir for desirable alleles that were lost during domestication or breeding (Bai and Lindhout, 2007; Simmonds, 1976; Zamir, 2001). In the last decades, this reservoir was exploited by introgression breeding, where 'exotic' genetic material from the wild species was introgressed into cultivated tomato (Rick and Yoder, 1988; Sim et al., 2012). Especially the recent advent of marker-assisted selection has made introgression breeding feasible. Marker-assisted selection, also known as precision breeding, has become possible thanks to numerous natural variation studies identifying molecular markers linked to several quantitative traits of agronomical significance (Bai and Lindhout, 2007; Foolad, 2007). The use of wild species germplasm for breeding purposes has focused mainly on the introduction of disease resistance genes into cultivated tomato (Foolad, 2007; Viquez-Zamora et al., 2013). Nevertheless, complex quantitative traits, such as yield, can also be improved by introgression of defined genomic regions from wild species (Gur and Zamir, 2004).

1.1.2 The wild tomato relatives

Including the domesticated species *S. lycopersicum*, the tomato clade (*Solanum* sect. *Lycopersicon*) currently consists of 13 species, as mentioned above (Figure 1). The 12 wild species only occur in South America where they grow primarily in the Andean region of Colombia, Ecuador, Peru and Chile. Additionally, two species are endemic to the Galapagos Islands – *S. cheesmaniae* and *S. galapagense* (Grandillo et al., 2011). Most of the wild species produce small, round, green

fruits. Only the three closest relatives of cultivated tomato, *S. pimpinellifolium* and *S. cheesmaniae* / *S. galapagense*, bear little red or orange fruits, respectively. These three species unambiguously form a monophyletic group together with *S. lycopersicum* (Figure 1). Despite some ambiguity, *S. pimpinellifolium* and not *S. galapagense* / *S. cheesmaniae* is very likely the closest relative of cultivated tomato (Viquez-Zamora et al., 2013). The wild species *S. pennellii* and *S. habrochaites*, on the other hand, are the two most distant relatives of cultivated tomato within *Solanum* sect. *Lycopersicon*.

There is a very high degree of morphological and genetic variation among the wild species (Grandillo et al., 2011; Nakazato et al., 2008). Part of this variation is likely due to adaptive differentiation. The wild tomatoes evolved according to the conditions they encountered in their diverse habitats. These habitats differ markedly, in altitude, precipitation and temperature. They range from desertlike environments that are inhabited by *S. pennellii* accessions to very high altitude alpine environments in the Andes Mountains where some *S. chmielewskii* and *S. habrochaites* accessions grow above 3000 m (Grandillo et al., 2011; Rick, 1973). *S. galapagense* inhabits another very extreme environment. At the edge of its range it actually grows in the littoral zone, sometimes only a few meters above the high-tide line, where it is faced with high salt concentrations (Rick, 1973). The ecologically most important environmental variables, explaining up to 70 % of the species geographic distribution, are annual temperature and precipitation (Nakazato et al., 2010). In conclusion, adaptive evolution in the wild species has led to a wealth of variation for various physiological and developmental traits. This variation can be used for breeding purposes but also for deciphering basic biological processes in a vegetable crop.

1.1.3 Tomato as a model system for quantitative genetics

There are a number of reasons that make tomato a suitable model plant for genetic studies: First, as mentioned above, there is tremendous phenotypic diversity among the wild species. Second, the vast majority of tomato accessions are diploid. Third, all cultivars and many of the wild accessions are self-compatible, allowing the generation of completely homozygous immortal lines by repetitive self pollination. Further, all of the wild tomato relatives can be hybridized with *S.*

lycopersicum. This not only allows them to be used for breeding purposes, but also enables the generation of mapping populations. Indeed, in the last two decades, several immortal mapping populations have been developed from crosses between various wild tomato species and cultivated varieties to study the genetic basis for differences between wild and domesticated tomato. For example, introgression line (IL) populations have been generated by crossing the wild species *S. pennellii*, *S. habrochaites* and *S. chmielewskii* with different cultivars (Eshed and Zamir, 1995; Monforte and Tanksley, 2000; Prudent et al., 2009). In these populations the genome of the wild species is introgressed into cultivated tomato in defined chromosomal segments. Every line ideally contains only one introgression and can be considered a near isogenic line (NIL). The introgressions of some ILs overlap, thereby partitioning the genome into defined bins to which traits can be mapped (Figure 2).

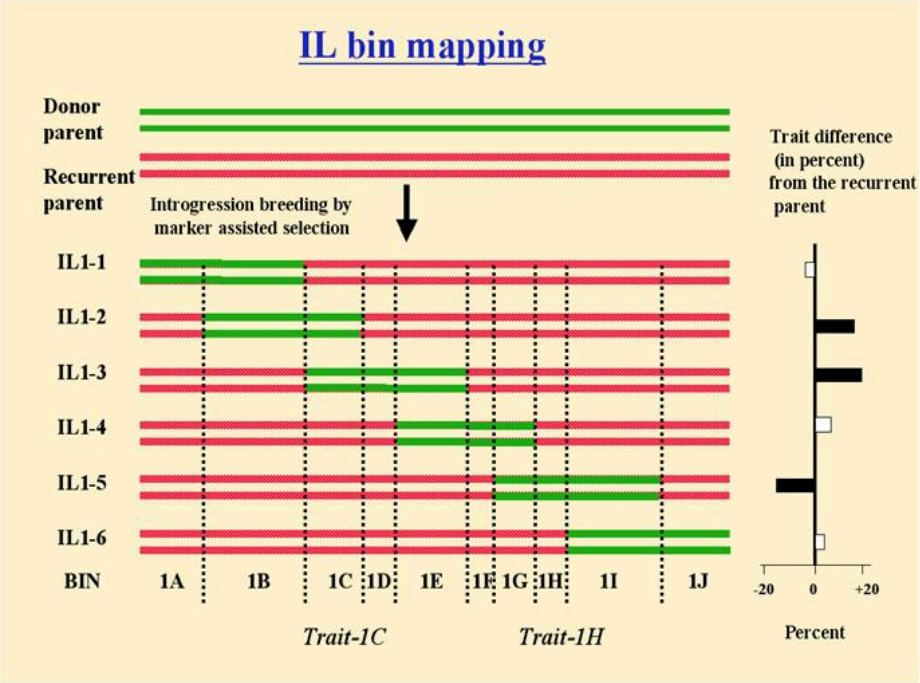

Figure 2: Schematic representation of six ILs for chromosome 1. The overlap of the introgressions allows mapping of traits to specific bins – in the example, 10 such bins can be distinguished: 1A to 1J. The six ILs displayed are homozygous for the recurrent parent for the other 11 chromosomes (http://zamir.sgn.cornell.edu/Qtl/il_story.htm).

Another frequently used type of immortal population consists of recombinant inbred lines (RIL). They are developed by repetitively selfing plants originating from a second filial generation (F2) population (Alonso-Blanco et al., 1998). RIL populations have been generated from crosses between the two close tomato

relatives *S. pimpinellifolium* and *S. galapagense* with cultivated tomato (Paran et al., 1995; Voorrips et al., 2000). One advantage of using RIL populations is that epistatic genetic interactions can be identified as there are several chromosomal segments of each of the parents. This is not possible using ILs. On the other hand, IL populations have increased power for the detection of small effect quantitative trait loci (QTL) since their effect is not masked by other co-segregating loci. The choice of which population is used depends on the parents of interest. RIL populations are often preferred when the parents are closely related, while with genetically more distant parents IL populations are advantageous. Because of the high sterility in the offspring of crosses between distant wild tomato relatives, like *S. pennellii* and *S. habrochaites*, and cultivated tomato, genome-wide coverage can only be obtained by several rounds of backcrossing (Eshed and Zamir, 1995; Keurentjes et al., 2007), resulting in either backcross inbred line (BIL) or IL populations.

In addition to the genetic tools that allow identification of loci involved in phenotypic variation, there are a number of molecular tools available that enable further characterization of these loci. Important examples are *Agrobacterium* mediated transformation using explants culture (McCormick et al., 1986), mutant collections generated by using the 'Targeting Induced Local Lesions IN Genomes' (TILLING) technology (Minoia et al., 2010; Okabe et al., 2013), virus-induced gene silencing (VIGS) (Liu et al., 2002b) and protoplast isolation and transformation (Jongsma et al., 1987; Yoo et al., 2007). These methods can be used to perform complementation tests to prove the genetic basis of a QTL and they allow functional characterization of the corresponding alleles.

Finally, next generation sequencing technologies together with the recently published genome sequence of cultivated tomato (The Tomato Genome Consortium, 2012) give rise to many new and exciting opportunities (Jimenez-Gomez, 2011). One example is the very precise high-density genotyping of mapping populations (Chitwood et al., 2013). Another example is ribonucleic acid (RNA) sequencing (RNAseq), which allows the analysis of sequence variants and expression differences for every expressed gene (Koenig et al., 2013). This kind of data can also greatly facilitate the identification of candidate genes for a QTL. The expected release of the genome sequences of several of the wild tomato relatives will cre-

ate even more opportunities in the near future (Finkers et al., 2013) (Fernie, personal communication).

1.1.4 QTL analyses in tomato

The populations described above and other populations (e.g. F2 and first backcross (BC1) F2 populations that are not immortal) have been used extensively to map QTL for various traits. Thousands of QTL have been published to date for practically every trait that can be measured. Six years ago, almost 3,000 QTL were reported using a single *S. pennellii* IL population (Lippman et al., 2007). Traits analyzed include biotic and abiotic stress responses, flowering time, metabolite content, and morphological characteristics of leaves, flowers, seeds and fruits, many of which are yield-related (for a summary see Foolad 2007). Despite the huge number of QTL, the underlying genes have been identified only in very few cases. The traits for which most QTL have been cloned are fruit weight and fruit shape. The fruit weight QTL fw2.2 was one of the first QTL to be cloned in tomato (Frary et al., 2000). It is responsible for 30 % of the differences in fruit weight between cultivated and wild tomato. It is considered a domestication QTL because it distinguishes wild from cultivated tomato in a dichotomous way (Frary et al., 2000). At about the same time the sugar content QTL Brix9-2-5 was cloned. It was delimited to the underlying nucleotide variant by means of high-resolution mapping (Fridman et al., 2000). Since then, two fruit shape and three additional fruit size QTL have been cloned (Chakrabarti et al., 2013; Cong et al., 2008; Liu et al., 2002a; Munos et al., 2011; Xiao et al., 2008). Further, the genes underlying one leaf shape and one seed size QTL were identified (Kimura et al., 2008; Orsi and Tanksley, 2009). Finally, from the many disease resistance QTL, allelic variation in an RNA-dependent RNA polymerase has been shown to be responsible for differences in virus resistance (Verlaan et al., 2013). The genomic region containing this gene had been introgressed for breeding purposes from *S. chilense* 20 years before (Zamir et al., 1994).

The molecular bases for the QTL cloned in tomato are very diverse, ranging from retrotransposon-mediated gene duplication leading to expression differences responsible for the elongated fruit shape locus *SUN* (Xiao et al., 2008) to simple SNPs causing a premature stop codon and thereby disrupting protein function,

as is the case for the pear shaped fruit locus *OVATE* (Liu et al., 2002a). It is interesting to note that of the six fruit weight and shape QTL that are due to domestication or breeding, five are caused by differences in expression and thus present expression QTL (eQTL). This agrees with the observation that most domestication traits are due to regulatory changes in the underlying genes (Doebley et al., 2006). Although Doebley et al. (2006) also hypothesized that most domestication genes would be transcription factors, this only holds true for one of the domestication genes cloned in tomato so far – the YABBY-like transcription factor underlying the fruit size QTL *fasciated* (Cong et al., 2008).

A major advancement in QTL analyses in tomato was brought about by the genome sequence of cultivated tomato (The Tomato Genome Consortium, 2012). It provides information about all genes located within a defined chromosomal segment of a QTL. The available functional annotation of each of those genes can lead to the identification of candidate genes even in relatively large regions. The positional cloning of a QTL can therefore be stopped at a much earlier stage and the actual cloning of the underlying gene can be based on a candidate gene analysis. For example, a QTL controlling acyl sugar composition in tomato trichomes was recently cloned after mapping the locus to a region that still contained 352 annotated genes. Three of them were very good candidates based on the functional annotation. Through further analysis of those three genes, one of them was demonstrated to underlie the QTL (Schilmiller et al., 2012).

In view of all the resources available today, many QTL are expected to be cloned in tomato in the near future. Comparing their genetic and molecular bases may contribute to a better understanding of tomato domestication and crop evolution. Additionally, the information can be used by breeders to develop new varieties that meet the changing demands of the environment and society.

1.2 The circadian clock

The circadian clock controls many traits of agronomic significance and it appears to be important for adaptation to specific environments. Such adaptation is crucial to agriculture to extend the range of crop plants and to adapt them to climate change. Yet, the circadian clock has never been studied before in tomato or any related species of the Solanaceae family. In general, virtually nothing is known about the significance of naturally occurring variation in circadian rhythms in crop plants.

Circadian clocks are endogenous timekeepers that allow organisms to anticipate predictable daily and annual seasonal changes caused by the rotation of the Earth. They thereby help organisms to appropriately phase their daily cycles and to correctly time their development. The most important characteristic of circadian clocks is that they keep running under constant conditions without any external cues (Sulzman et al., 1984). At the same time, they are able to integrate environmental signals such as light and temperature (Harmer, 2009). These external signals entrain the circadian clock, synchronizing it with the day-night cycles of Earth.

It is widely accepted that the circadian clock in plants coordinates diverse physiological outputs to optimize growth and development (Hsu and Harmer, 2013). Circadian clocks have been proposed as central modulators of overall individual performance.

1.2.1 Evolutionary significance of the circadian clock and its importance for performance

Circadian clocks can be found in all three domains of life (Edgar et al., 2012). Thus, it seems that there was evolutionary pressure to develop circadian timekeeping mechanisms, suggesting that they confer an adaptive advantage. However, providing direct experimental evidence for this assumption is not straightforward. In *Arabidopsis thaliana* (Arabidopsis) the circadian clock can be turned off by overexpression of the core clock gene *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) (Wang and Tobin, 1998). In an experiment using Arabidopsis plants overexpressing *CCA1*, differences in plant performance (seedling survival) were detected under extremely short day lengths of four hours (Green et al., 2002).

Under normal long day or short day laboratory conditions (16:8 or 8:16 hours light/dark cycles, respectively) no differences could be observed (Green et al., 2002). The lack of differences under standard conditions can probably be explained by the fact that even arrhythmic Arabidopsis plants retain the ability to respond to diurnal light and temperature changes. In fact, it seems that in arrhythmic plants only the ability to anticipate such changes and to gate the according responses is lost (Green et al., 2002; Nozue et al., 2007). Therefore, the differences resulting from a nonfunctional circadian clock can be quite subtle and may depend a lot on the growth conditions. Nevertheless, significant differences in chlorophyll content, carbon fixation and biomass between wild type and *CCA1*ox plants were reported in a different study that used 12:12 hours light dark cycles (Dodd et al., 2005). It should be noted, though, that these differences are concluded from a single experiment containing only five to six biological replicates per genotype. For cyanobacteria, on the other hand, the fitness advantage of a functional circadian clock under diurnal conditions was clearly demonstrated by competition experiments (Woelfle et al., 2004). Interestingly, this advantage disappeared or even reversed under constant conditions. A similar observation has been made for arrhythmic Arabidopsis plants (Dodd, personal communication), suggesting that a circadian clock is beneficial when living in a rhythmically changing environment, which is the case for the vast majority of organisms.

1.2.2 Naturally occurring variation in circadian rhythms and adaptive differentiation

As central regulators of multiple fundamental biological processes, an interesting question is how much phenotypic variation in circadian rhythms is present in nature and whether this variation is due to adaptive differentiation. Working with the crop plant tomato, it is most interesting whether variation in circadian rhythms exists due to domestication and breeding.

As mentioned above, the circadian clock is synchronized with the environmental cycles in a process called entrainment. For this reason, clock outputs under diurnal conditions always have a period of exactly 24 hours. To identify differences in circadian periods between different genotypes, the clock output that is used to

monitor the state of the clock has to be observed under artificial constant conditions. The circadian or free-running period can only be assessed under these conditions (Figure 3).

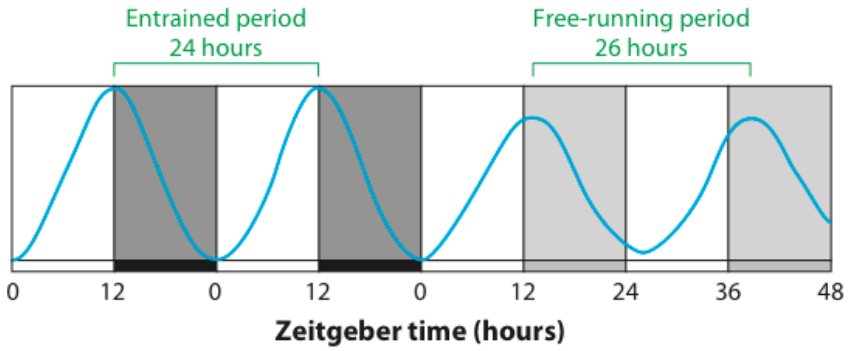

Figure 3: Idealized output of a circadian clock. The left half shows the output under entraining light/dark cycles, the right half under constant conditions. Under natural diurnal conditions the period of clock outputs is always exactly 24 hours. Circadian or free-running periods deviating from 24 hours can only be observed under constant conditions. Modified from Harmer (2009).

For Arabidopsis and Cyanobacteria, it was proposed that circadian resonance, i.e. matching of the circadian clock period with that of the 24 h environmental cycle, enhances growth and fitness (Dodd et al., 2005; Ouyang et al., 1998). Consequently, one would expect that there is evolutionary pressure for circadian clocks to have rhythms that match a circadian period of 24 hours. However, various studies have demonstrated that there is marked natural variation in circadian rhythms, for example in Arabidopsis (Michael et al., 2003b), among cultivars of the related species *Brassica rapa* (Lou et al., 2011), but also among different *Drosophila* strains (Emery et al., 1994). Latitudinal clines of such naturally occurring variation in circadian rhythms suggest that modulation of the circadian system is important for adaptation to specific environments (Joshi, 1999; Michael et al., 2003b; Pittendrigh et al., 1991; Simunovic and Jaenike, 2006). Altered circadian rhythms can affect the timing of a plethora of clock outputs and consequently could be selected for (Yakir et al., 2007). The underlying force driving that selection is still elusive, however. Given the circadian clocks pervasive control, it is likely that different factors are causing the variation in circadian rhythms in diverse species. Whereas in Arabidopsis altered circadian rhythms may lead to differences in flowering time that could be advantageous, as suggested by Michael et al. (2003b), in Drosophila differences in desiccation stress might be responsible for the latitudinal cline in circadian rhythms (Simunovic and Jaenike, 2006).

In conclusion, it seems very likely that having a circadian system that best fits the specific environment of an organism is of adaptive significance (Yerushalmi et al., 2011).

1.2.3 The circadian clock in crop plants

Considering the adaptive significance of the circadian clock and its central regulatory role for many of the traits that are desirable in agriculture, it appears plausible that the circadian clock could have been a target of artificial selection during domestication or breeding. Indeed, several cases have been reported in which homologs of Arabidopsis clock genes have been changed during domestication or breeding. For example, in barley the major determinant of photoperiod response *Ppd-H1* was identified as a pseudo response regulator (*PRR*) by positional cloning, most similar to the Arabidopsis clock gene *PRR7* (Turner et al., 2005). A mutation in *Ppd-H1* results in a reduced response to photoperiod, thereby delaying flowering under long day conditions. This was an important factor in adapting spring barley to the longer growing season in Western Europe and most of North America. Even though a clock gene was targeted to reduce photoperiod sensitivity, the circadian clock itself was not affected by this mutation (Campoli et al., 2012; Turner et al., 2005). In sorghum several independent mutation events in the homolog of barley *Ppd-H1* (*PRR3/7*) also led to reduced photoperiod sensitivity, again without changing expression of other clock genes (Murphy et al., 2011). As in barley, the reduced photoperiod sensitivity was crucial for the range extension of sorghum to the north. However, because sorghum is a short day plant, mutant alleles accelerated flowering under long days, contrary to barley. A third example of a *PRR3/7* gene being targeted by artificial selection is the bolting locus B in beet. A partial loss-of-function allele was selected during domestication, conferring bienniality and thus changing the life cycle of the plant. As in the other two cases, photoperiod sensitivity is reduced without affecting the circadian clock (Pin et al., 2012). It is interesting that mutations in the same gene have resulted in reduced responses to photoperiod in three very diverse plant species.

A second Arabidopsis clock gene, whose homologs were targeted by artificial selection during breeding in various crop plants, is *EARLY FLOWERING 3* (*ELF3*). In

Arabidopsis *ELF3* is a core component of the circadian clock, as its loss leads to arhythmicity (Thines and Harmon, 2010). In contrast, conditional circadian dysfunction only apparent in light has also been reported (Hicks et al., 1996). Recently it has been demonstrated that the gene underlying the early maturity locus *EAM8*/*Mat-a* in barley is homologous to Arabidopsis *ELF3* (Faure et al., 2012; Zakhrabekova et al., 2012). Many different mutant alleles led to early flowering phenotypes, supposedly adapting barley to the short growing seasons with long day lengths at high latitudes in Europe, thereby enabling further range extension. Barley plants with mutant *elf3* alleles exhibit arrhythmic expression of several clock genes and clock output genes (Faure et al., 2012). Loss of function alleles of *elf3* were also responsible for the photoperiod adaptation of pea and lentil (Weller et al., 2012). These alleles confer reduced photoperiod sensitivity and early flowering. Expression of clock genes was severely disrupted in pea genotypes harboring *elf3* mutant alleles. This disruption was only apparent in light, though, consistent with the results reported by Hicks et al. (1996) in Arabidopsis (Weller et al., 2012).

Although the examples described above illustrate that clock genes have been targeted during domestication or breeding to adapt crop species to new environments, they do not represent cases of quantitative variation in the circadian system as reported for natural accessions of Arabidopsis (Michael et al., 2003b). In the case of *Ppd-H1* / *PRR3/7* circadian rhythms are not affected at all. In the case of *ELF3* circadian rhythms are completely abolished. Thus, the question about the role of quantitative variation in circadian rhythms for crop plant adaptation remains unanswered. To my knowledge, there is only one report specifically studying variation in circadian rhythms in a crop plant. Circadian period variation among 50 *Brassica rapa* cultivars is reported (Lou et al., 2011). This study concludes that there are no differences between cultivated varieties and wild populations, implying that a change in the circadian clock was not essential for *Brassica rapa* domestication. It will be necessary to compare circadian rhythms of various crop plants and their wild ancestors to answer the question regarding the significance of quantitative circadian clock variation for domestication and crop evolution.

1.3 Photoperiodism

As mentioned above, the circadian clock not only helps to synchronize daily biological rhythms, but also allows seasonal timing of development. The most important seasonal output of the circadian clock in plants is photoperiodic flowering. For many plant species appropriate timing of this life history trait is essential (Andres and Coupland, 2012). Flowering at the right time can be important for many reasons such as avoiding unfavorable conditions or synchronizing with pollinators or with related individuals for cross pollination. Correct timing of such a critical life history event can be essential for survival and is often under strong natural selection (Wilczek et al., 2009). In order to assess the time of the year, plants measure photoperiod, which is the most reliable seasonal indicator (Jackson, 2009). Changes in photoperiod are exactly the same every year, being independent of inter-annual variation in climate. Plants can perceive differences in daylength as small as 5 minutes, as demonstrated by strong effects on flowering time in rice (Dore, 1959). Even directly at the equator, where photoperiod does not change throughout the year, plants are able to time their development, possibly by measuring variation in sunrise and sunset times (Borchert et al., 2005).

1.3.1 Photoperiodic flowering in Arabidopsis and the external coincidence model

The circadian clock is an essential part of seasonal timekeeping (Thomas et al., 1997). Photoperiodism is lost in arrhythmic plants (Doyle et al., 2002; Hicks et al., 1996) and reduced in clock mutants with altered period lengths (Mizoguchi et al., 2002; Somers et al., 2000). A conceptual framework in which the circadian clock together with light serves as the basis for photoperiodism was already proposed 80 years ago (Bünning, 1936). In this so-called external coincidence model the circadian clock regulates the phasing of light responsiveness. This phasing is day length independent. The responsive phase will only coincide with light under certain photoperiods, thereby making the response day length specific. In Arabidopsis it was shown that the gene *CONSTANS* (*CO*) mediates between the circadian clock and photoperiodic flowering (Suarez-Lopez et al., 2001). The demonstration that *CO* requires light, perceived by the photoreceptors crypto-

chrome 2 and phytochrome A, to activate flowering, firmly established that seasonal time measurement in Arabidopsis is indeed achieved by external coincidence (Yanovsky and Kay, 2002). There are two molecular mechanisms explaining how coincidence of light with circadian timing regulates CO function to achieve photoperiodic flowering. First, daytime *CO* expression is induced only at the end of long days by a protein complex formed by FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1) and GIGANTEA (GI). This long day specific activation is achieved via the clock-controlled expression of FKF1 and GI proteins, leading to peak expression late in the day together with the induction of their interaction by light (Sawa et al., 2007). Peak expression and light-induced interaction only coincide under long day conditions. Additionally, posttranscriptional regulation of the CO protein is crucial for the photoperiodic flowering response. *CO* expression is itself regulated by the circadian clock and peaks late in the day irrespective of photoperiod. But only when this peak coincides with light, which is the case only in long days, the CO protein is stable and can induce flowering by activating transcription of *FLOWERING LOCUS T* (*FT*) (Valverde et al., 2004).

FT has been shown to encode a systemic mobile signal. It is expressed in the leaves but the encoded protein moves through the phloem to the shoot apical meristem where it induces the transition from vegetative to reproductive development (Corbesier et al., 2007). The existence of such a systemic signal, dubbed florigen, was demonstrated in several experiments conducted as early as the 1930s, but the identity of the signal has remained elusive (Zeevaart, 1976). In the last decade it became clear that the FT protein is an important component of florigen not only in Arabidopsis, but also in tomato, rice and potato, where homologs of Arabidopsis *FT* were shown to encode systemic mobile signals that can move across graft junctions to induce flowering (Lifschitz et al., 2006; Navarro et al., 2011; Tamaki et al., 2007).

1.3.2 Photoperiodic flowering outside Arabidopsis

The *CO* / *FT* module regulating photoperiodic flowering in Arabidopsis is conserved across species, as it also controls photoperiodic flowering in many other plant species, e.g. rice and barley (Hayama et al., 2003; Turner et al., 2005). An interesting difference between the short day plant rice and Arabidopsis is the bi-

functionality of the *CO* ortholog *Heading date 1* (*Hd1*). It promotes flowering specifically under short day conditions by activating expression of *Heading date 3a* (*Hd3a*), the homolog of Arabidopsis *FT*. Under long day conditions, however, *Hd1* inhibits flowering by repressing *Hd3a* (Izawa et al., 2002). The molecular basis for this bi-functionality is currently not known (Shrestha et al., 2014).

A third group, besides long day and short day plants, comprises the day neutral plants, which flower irrespective of photoperiod. While the flowering responses of the long day plant Arabidopsis and the short day plant rice are well studied, the day neutral response is only poorly characterized (Mizoguchi et al., 2007). Cultivated tomato is a prominent example of a day neutral plant. Despite the conserved role of *FT* in tomato, homologs of Arabidopsis *CO* do not appear to be involved in the regulation of flowering (Ben-Naim et al., 2006). The molecular basis for the photoperiodic flowering response present in some of the wild tomato species that are short day plants (e.g. *S. habrochaites* and *S. galapagense*) is not yet known.

1.3.3 Other photoperiodic responses

Interestingly, the *CO* / *FT* module is not only conserved across species in controlling photoperiodic flowering, but also in regulating other photoperiodic responses like tuberization in potato and bud cessation in poplar (Bohlenius et al., 2006; Navarro et al., 2011). Cultivated tomato does not possess any of those common photoperiodic responses. Nevertheless, it is strikingly affected by photoperiod. Long days cause light-induced injury in tomato (Arthur et al., 1930; Hillman, 1956; Withrow and Withrow, 1949). This phenomenon can lead to decreased growth, indicated by reduced plant weight, under photoperiods of 17 hours as compared to 12 hours (Arthur et al., 1930). It has been shown that the circadian clock is involved in light-induced injury in tomato (Highkin and Hanson, 1954; Hillman, 1956; Velez-Ramirez et al., 2011). These data fit well with a model proposing that plants go through a photophil and a scotophil phase each day (Bünning, 1950). According to this model, light has different effects depending on the endogenously defined phase. Light during the scotophil phase, which is set by the circadian clock, would be damaging. Since tomato originates from South America relatively close to the equator, it is adapted to photoperiods of

approximately 12 hours. With the transfer of tomato to Mexico during domestication and later to Europe it faced much longer photoperiods. In Rome, for instance, day lengths are up to 15 hours in summer.

Adaptation to new agro-ecological and cultural environments due to geographical radiation is important for crop diversification (Meyer and Purugganan, 2013). Especially reduced photoperiod sensitivity was crucial as discussed above for barley, sorghum, beet and pea. Loss of the photoperiodic tuberization response was also essential for potato cultivation in northern latitudes (Kloosterman et al., 2013). It is an open question whether any adaptation to changed photoperiods occurred in tomato during the domestication process or subsequent crop evolution.

2 Aim of the project

The circadian clock regulates many important aspects of plant physiology and development. Additionally, naturally occurring variation in the circadian system appears to be of adaptive significance. In spite of that, nothing is known about natural variation in circadian rhythms in tomato and its wild relatives. The role of quantitative changes in circadian rhythms during domestication and breeding also remains to be studied. The primary aim of this thesis was to assess the extent of natural variation in circadian rhythms in cultivated tomato and its wild relatives and to decipher the molecular basis of this variation. To do so, a camera system was established to monitor circadian leaf movement rhythms. Using this system, patterns of variation were identified and an effect of domestication on the tomato circadian clock was revealed. Employing mapping populations, QTL responsible for differences between cultivated and wild tomato were detected. One of those loci was fine-mapped and a candidate gene was cloned.

The most important seasonal output of the circadian clock in plants is photoperiodic flowering. Domestication and breeding have led to early flowering and photoperiod-insensitive tomato varieties. In contrast, some of the wild tomato relatives exhibit a pronounced delay in flowering under long day conditions. The molecular basis for photoperiodic flowering is very well understood in model species like Arabidopsis and rice. However, in tomato the genes underlying natural variation in photoperiodic flowering are unknown. Therefore, the second aim of this thesis was to identify the loci responsible for the differences in photoperiodic flowering between cultivated tomato and two of its wild relatives. This was approached by QTL analysis and bulk segregant analysis combined with next-generation RNA sequencing in two different mapping populations. Loci underlying the variation were identified and candidate genes are proposed.

3 Material and Methods

3.1 Plant material

3.1.1 Tomato accessions

Wild and cultivated tomato accessions used to study patterns of variation for circadian rhythms and flowering time were obtained from four different sources: (i) the University of California Davis / C.M. Rick tomato genetics resource center (TGRC), (ii) Enza Zaden (Enkhuizen, Netherlands), (iii) Dr. Richard Finkers from the University of Wageningen and (iv) Prof. Dr. Dani Zamir from the Hebrew University of Jerusalem. For a detailed list of all the accessions used see Table 4 in the Appendix.

3.1.2 Tomato populations

Mapping populations used to map differences in circadian rhythms and flowering time between species were obtained from various sources.

The *S. pennellii* IL population (Eshed and Zamir, 1995) was kindly provided by Prof. Dr. Julin Maloof and Prof. Dr. Neelima Sinha from the University of California Davis. Genotype information for each line was obtained from TGRC (http://tgrc.ucdavis.edu/pennellii_ils.aspx) and from recently published fine-scale genotyping (Chitwood et al., 2013).

Seeds for the *S. habrochaites* IL population (Monforte and Tanksley, 2000) and corresponding genotype information were obtained from TGRC (http://tgrc.ucdavis.edu/hirsutum_ils.aspx).

Seeds for the *S. pimpinellifolium* RIL population (Voorrips et al., 2000) together with the corresponding genotype information for 1966 markers were kindly provided by Dr.ir. Sjaak van Heusden from the University of Wageningen.

Seeds for the *S. pennellii* BIL population were received from Prof. Dr. Dani Zamir from the Hebrew University of Jerusalem. The according genotypes consisting of 1231 markers were obtained from Prof. Dr. Neelima Sinha from the University of California, Davis.

The *S. galapagense* F2 population was generated by pollinating emasculated flowers of *S. lycopersicum* cv. Moneymaker with pollen of the accession LA0530 and subsequent self-pollination of the resulting first filial (F1) hybrid.

3.1.3 Arabidopsis lines

For a heterologous complementation approach with the purpose of testing a QTL candidate gene, mutant and wild type Arabidopsis plants were employed.
Seeds for the Arabidopsis SALK T-DNA insertion mutants (Alonso et al., 2003) (SALK_021058C and SALK_027403C with insertions in the gene *EID1*, At4g02440) were obtained from the Nottingham Arabidopsis Stock Centre (NASC). Seeds for the Columbia (Col-0) accession, which is the genetic background of the SALK lines, were provided by Dr. Inga Schmalenbach from the Max Planck Institute for Plant Breeding Research (MPIPZ).

3.2 Growth conditions

All tomato seeds were treated with either saturated tri-sodium phosphate (Na_3PO_4) or 1 % sodium hypochlorite (NaClO) for 15 minutes to kill viruses on the seed coat (Ling, 2010) and to enhance germination efficiency. After this treatment, seeds were kept in water for three days in the dark and on the third day they were sown on soil. *S. galapagense* seeds have a special seed coat conferring strong seed dormancy (http://tgrc.ucdavis.edu/seed_germ.aspx). Seeds of this species were therefore treated with 2.7 % NaClO for 30 minutes once a week as recommended by TGRC. Enough seeds sprouted one day after the third treatment, and were then sown on soil. If not, the treatment was continued for another week, keeping seeds in water between treatments. Arabidopsis seeds were stratified on a moist filter paper in petri dishes at 4 °C in the dark for three days before sowing.
For leaf movement analyses the soil consisted of a mixture of ten parts Stender-Erde (Typ A240, Stender AG, Schermbeck) and one part sand supplemented with Osmocote® Start (Everris International, Heerlen, Holland) and BioMÜKK® WDG (Biofa AG, Münsingen). For all other experiments the soil consisted of a mixture of 7.2 parts Einheits-Erde® Classic (Typ ED73), 2.4 parts Einheits-Erde® Special (Typ Mini-Tray, both from Einheitserde- und Humuswerke, Sinntal-Jossa) and 0.4 parts sand supplemented with Schwefelsaures Ammoniak 21 (RWZ Raif-

feisen, Raiffeisen Waren-Zentrale Rhein-Main eG, Köln), EPSO Microtop® (K+S KALI GmbH, Kassel) and Grün-Kalk (Deutsche Cuxing Marketing GmbH, Ottendorf).

3.2.1 Leaf movement analysis

Genotypes were positioned based on a completely randomized design. After germination, defined as plants with fully spread cotyledons, seedlings were entrained for two to four days under cool white fluorescent tubes (Philips F32T8/TL741, ~100 µmol m^{-2} s^{-1}) in 12:12 light/dark and 20:18 °C temperature cycles in a growth chamber (Elbanton, Kerkdriel). A polystyrene pebble was glued to the tip of one cotyledon of each seedling using an acid and solvent free Herma glue stick (HERMA GmbH, Filderstadt) (Salathia et al., 2007) (Figure 5B). The next day at the dark:light transition (Zeitgeber (ZT) 0), seedlings were transferred to an identical growth chamber with constant light and temperature (~100 µmol m^{-2} s^{-1} and 25 °C) for circadian leaf movement analysis.

3.2.2 Transcript abundance time course

The three genotypes used in this analysis were distributed evenly in seven blocks. After germination, seedlings were entrained for six to seven days under cool white fluorescent tubes (Philips F32T8/TL741) + incandescent bulbs (Riva 25 W) with a light intensity of about 100 µmol m^{-2} s^{-1} in 12:12 light/dark and 25:18 °C temperature cycles in an Elbanton growth chamber (Elbanton, Kerkdriel, Holland). The growth chamber was then switched to constant conditions (~100 µmol m^{-2} s^{-1} and 25 °C). Samples from green tissue were taken every four hours for three days starting at ZT 0 with three biological replications per genotype and time-point. Biological replicates were taken from three different trays at different positions in the growth chamber to minimize possible positional effects.

3.2.3 Flowering time analysis

3.2.3.1 *Solanum galapagense*

Plants were distributed in three climate chambers (CLF Plant Climatics Gro-Banks) equipped with cool white fluorescent tubes (Philips F40T8/TL841, ~100 µmol m^{-2} s^{-1}) and set to 28 °C during the day and 18 °C during the night.

For short day (SD) conditions plants were grown in 8:16 light/dark cycles, for long day (LD) conditions in 16:8 light/dark cycles and for night break (NB) conditions in 8:16 light/dark cycles with an additional 30 minutes of light in the middle of the dark period. Trays were shuffled between chambers twice a week and the conditions of the chambers were changed accordingly to avoid any effect of the climate chamber.

3.2.3.2 *Solanum lycopersicum* cv. Moneymaker x *Solanum galapagense* F2 population

480 F2 plants were grown in two climate chambers (Bronson, Nieuwkuijk) under cool white fluorescent tubes (Sylvania Luxline Plus F58W/840 T8) + incandescent bulbs (SLU 30W 2100K, Shine Lighting Technology Co., Ltd.) with a light intensity of about 250 µmol m^{-2} s^{-1} in 16:8 light/dark and 24:18 °C temperature cycles. Trays containing six plants were shuffled three times a week within each climate chamber to reduce positional effects.

3.2.3.3 *Solanum habrochaites* introgression line population

Plants were grown in two Bronson climate chambers (Bronson, Nieuwkuijk) under cool white fluorescent tubes (Sylvania Luxline Plus F58W/840 T8) + incandescent bulbs (SLU 30W 2100K, Shine Lighting Technology Co., Ltd.) with a light intensity of about 250 µmol m^{-2} s^{-1} in 8:16 light/dark and 24:18 °C temperature cycles for short days and in 8:16 light/dark cycles with an additional half hour of light and 24 °C in the middle of the dark period for night break conditions. The genotypes were positioned according to a randomized complete block design with each shelf (each chamber has four shelves) being one block. Each genotype was grown once in each block resulting in 8 biological replicates per genotype. Trays containing six plants were shuffled three times a week within each shelf.

3.2.4 Greenhouse experiments

Phenotyping for pleiotropy of circadian rhythms and grafting experiments were carried out under controlled greenhouse conditions at the Max Planck Institute for Plant Breeding Research in Cologne. Photoperiod is maintained throughout the year at 16 hours by providing artificial light from mercury vapor lamps (Philips Deutschland GmbH, Hamburg) in the morning and evening or whenever

the light intensity is below 20 klx between 6 a.m. and 10 p.m. Temperature is kept above a minimum of 18 °C.

Grafting was performed with four weeks old seedlings using a V-cut. The scion was roughly 5 cm in length. Scions and rootstocks were joined using adhesive tape. Plants were covered with plastic bags for one week after grafting for increased humidity.

The light intensity values for the different climate chambers in µmol m^{-2} s^{-1} were determined using a LI-COR light meter (LI-250A) equipped with a Quantum sensor (LI-190), which measures the photosynthetically active radiation (both from LI-COR Biosciences GmbH, Bad Homburg). For all leaf movement experiments light intensity, temperature and humidity were measured every five minutes using a HOBO® data logger (Onset Computer Corporation, Bourne, MA). The data was used to make sure that conditions were constant during the time of the experiments (as an example see Figure 4) since environmental fluctuation can have a strong impact on circadian rhythms.

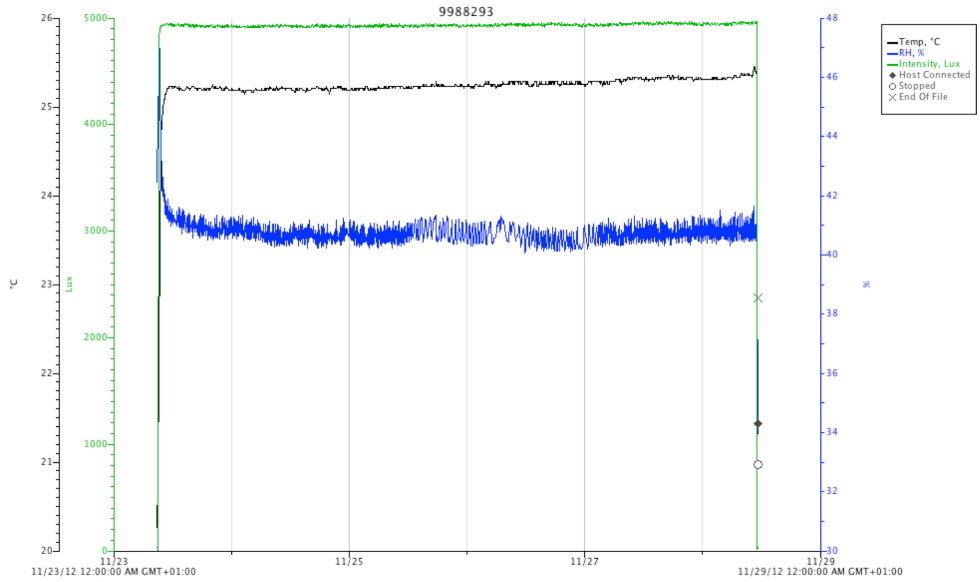

Figure 4: Light intensity, temperature and relative humidity in the imaging chamber. Light intensity in lux (green), temperature in °C (black) and RH in % (blue) as measured with a HOBO® data logger. Data are shown for 5 days corresponding to one leaf movement experiment.

3.3 Phenotyping

3.3.1 Circadian leaf movements

Circadian leaf movements were monitored in seedlings grown in constant light at a constant temperature. Pictures of these seedlings were taken every 20 minutes for four to five days using Pentax Optio WG-1 charge-coupled device (CCD) cameras controlled by their internal intervalometer. Cameras were used in program mode in order to be able to manually adjust settings like white balance and exposure compensation to optimize images for the subsequent analysis. The resolution was set to 2048 x 1536 pixels. Six such cameras were used each taking pictures of 27 tomato seedlings (Figure 5).

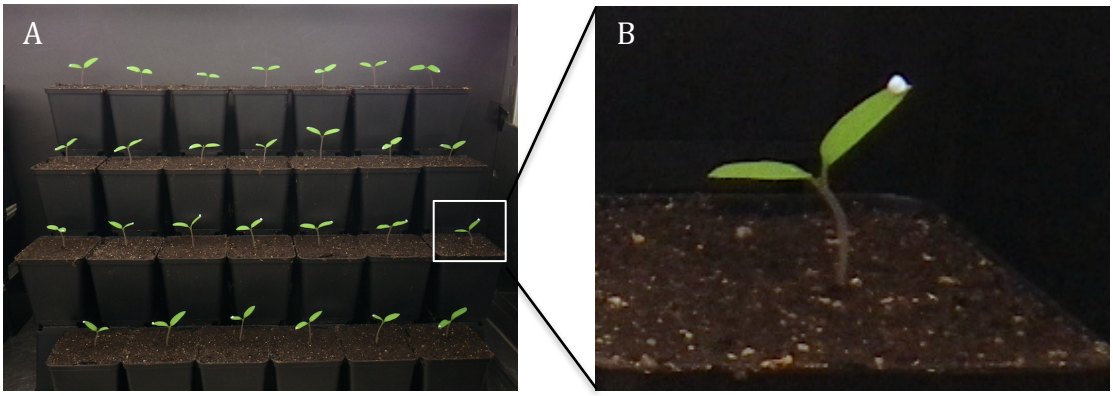

Figure 5: Tomato seedlings in the imaging chamber at ZT24. (A) 27 tomato seedlings are being imaged by one camera. (B) A polystyrene pebble is glued to the tip of one cotyledon of each seedling.

The pictures of the seedlings were analyzed using the freely available software ImageJ (http://rsbweb.nih.gov/ij/). They were imported into ImageJ as a stack, which was then split into the three color channels. Only the blue channel was used. The threshold was adjusted in such a way that only the polystyrene pebbles remained visible as black points. The centroids of these points were measured with the analyze particles command. The Y-values of these measurements represent the vertical positions of the cotyledon tips. These time course data (vertical cotyledon positions) were then analyzed with the biological rhythms analysis software system (BRASS) (available on http://millar.bio.ed.ac.uk) to obtain estimates for the circadian period, phase, amplitude and relative amplitude error (RAE). BRASS estimates these variables via fast Fourier transform nonlinear least-squares analysis (Plautz et al., 1997). BRASS was run using the default settings. The first 24 hours were excluded from the analysis to remove initial

noise that could potentially be caused by the transfer from the entrainment chamber to the imaging chamber. The RAE, estimated by BRASS, is a measure of the robustness of a rhythm and can theoretically have values between 0 and 1. A value of 0 indicates a perfect rhythm and a value of 1 a rhythm that is not statistically significant (Plautz et al., 1997). All seedlings with RAE values below 0.25 were taken for the further analysis. This corresponds to the 95 % quantile as determined for the *S. pennellii* IL population (Figure 6).

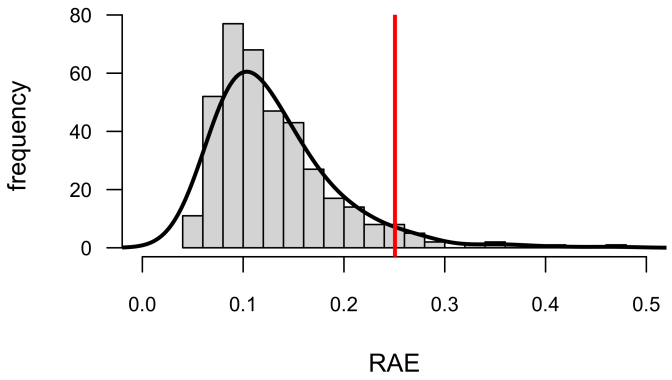

Figure 6: RAE distribution of the *S. pennellii* IL population. The distribution of the RAE estimates of all individuals (n = 385) of the *S. pennellii* IL population is shown. The black line indicates the density, the red line the 95 % quantile, which is 0.2506. Data are from four independent experiments.

Additionally, individuals that were obviously outliers were removed manually before further analysis (see Figure 7 as an example).

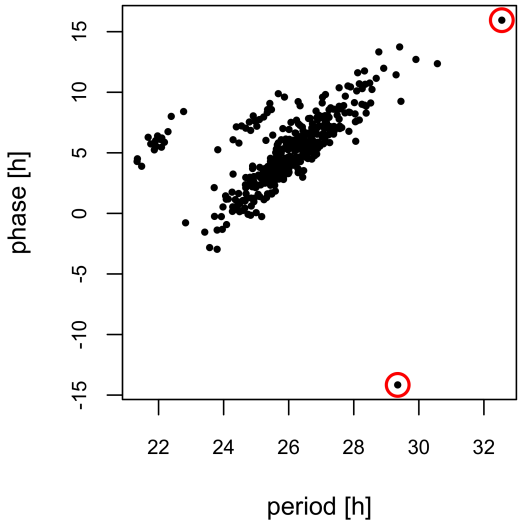

Figure 7: Period and phase estimates of the *S. pennellii* IL population. The period estimates are plotted against the phase estimates of the 365 individuals of the *S. pennellii* IL population with RAE values smaller than 0.25. Two obvious outliers are marked with a red circle. Data are from four independent experiments.

Phase values are outputted by BRASS as circadian time (CT) phases meaning they are normalized by period. The sidereal phases for every seedling were calculated (sidereal phase = period / 24 * CT phase) (Michael et al., 2003a). From these values, the time of the second peak of the theoretical trace (the sine wave fitted by BRASS), termed peak time throughout this thesis, was determined by adding twice the period value (Figure 8).

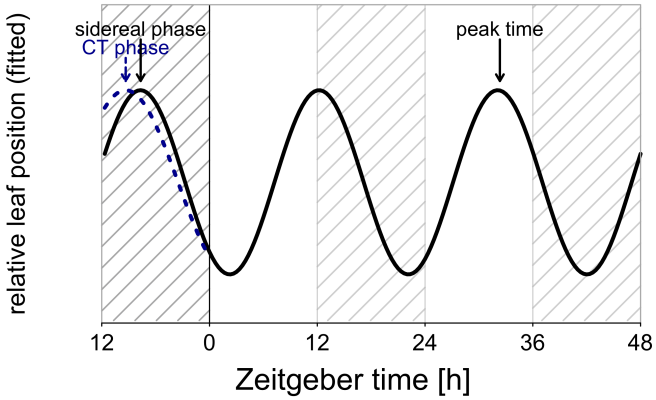

Figure 8: Theoretical leaf movement trace. The actual leaf movement trace starts at Zeitgeber time 0. BRASS outputs the CT phase obtained from the period corrected extension of the fitted curve (shown as dashed blue line). Sidereal phase values are obtained if the fitted sine wave is continued to the left without correcting for period. The time of the second peak (peak time) is obtained by adding two period values to the sidereal phase value.

3.3.2 Flowering time

Days to flowering were counted from germination, defined as seedlings with open cotyledons. Plants were scored as flowering when the first flower of the first inflorescence had completely opened. Germination and flowering were scored once a day at approximately the same time in the afternoon. To determine flower number, all flowers and buds of a terminated inflorescence were counted.

3.3.3 Seedling height

Seedling height was measured manually using a ruler. The distance from the cotyledons to the shoot apical meristem was determined. Values were rounded to half centimeters.

3.4 Designing markers for genotyping

Polymorphic markers were used: (i) to confirm and precisely define the genotypes of introgression lines from the *S. pennellii* and *S. habrochaites* populations; (ii) to identify recombinants in F2 populations from ILs backcrossed to their re-

current parent and to define their recombination breakpoints and (iii) to determine the alleles at specific QTL for an *S. galapagense* F2 population. For this, cleaved amplified polymorphic sequence (CAPS) markers and insertions and deletions (indel) markers were used. CAPS markers are based on SNPs or very short indels that create or remove a restriction site. By digestion with the specific enzyme recognizing the altered motif, the polymerase chain reaction (PCR) fragment of one genotype, but not the other, will be digested. Alleles can then be distinguished by separation of the resulting deoxyribonucleic acid (DNA) fragments through gel electrophoresis on agarose gels. The amplified fragments for indel markers differ in size without any additional enzymatic steps. Because the size differences are often less than 10 base pairs (bp), DNA fragments need to be separated on polyacrylamide gels, which have a higher resolution than agarose gels (for gel electrophoresis methods see section 3.5.6). Markers were designed using custom written Perl scripts. The scripts use Variant Call Format (VCF) files with sequence variants obtained using Genome Analysis Toolkit (GATK) (http://www.broadinstitute.org/gatk/index.php) on RNA-seq data from *S. pennellii*, *S. habrochaites*, *S. chmielewskii*, *S. galapagense*, *S. lycopersicum* cv. M82 (Koenig et al., 2013) and *S. lycopersicum* cv. Moneymaker (unpublished). The script to design CAPS markers in silico generates both alleles of a SNP or indel and compares the sequences with a database of the most common restriction enzymes. When a SNP suitable of a CAPS or indel marker is recognized, the scripts use Primer 3 (http://primer3.sourceforge.net/) to design primers, taking care that primers are not placed in regions with other variants or without coverage of RNAseq reads. A list of all markers used in this thesis can be found in the Appendix (Table 5 + Table 6).

3.5 Molecular analyses

3.5.1 DNA extraction

For large-scale genotyping, two DNA extraction methods were used: (i) 'magic buffer' and (ii) the BioSprint 96 workstation with the BioSprint 96 DNA Plant Kit (both from QIAGEN GmbH, Hilden). For DNA extraction with the 'magic buffer' 200 µl of buffer (see recipe below) and a steel bead (Ø 2 mm) were added to approximately 2 mg of fresh plant material in each well of a 96-well PCR-plate and

closed with 8-cap strips (both from Bio-Budget Technologies GmbH, Krefeld). Tissue was then ground in a bead mill for one minute at 28.5 Hz (RETSCH GmbH, Haan) and afterwards centrifuged for five minutes at 6000 revolutions per minute (rpm). The supernatant was used as DNA template for subsequent PCR reactions. This extraction method only works reliably with amplicons smaller than 500 bp. DNA extraction with the BioSprint 96 DNA Plant Kit was carried out according to the manufacturer protocol. The first step was modified: disruption of the plant tissue was conducted after and not before the addition of 300 µl buffer RLT.

Magic buffer:

10 g Sucrose

5 ml Tris-HCl (pH 8.0)

6 ml 5M NaCl

to 100 ml with distilled water

The buffer can be stored at 4 °C for several months.

For limited numbers of samples and for increased purity, needed for cloning (see section 3.5.9), the DNeasy Plant Mini Kit (QIAGEN GmbH, Hilden) was used following the manufacturer protocol.

DNA concentrations were determined with a spectrophotometer (NanoDrop 1000, peqlab Biotechnologies GmbH, Erlangen) if needed.

3.5.2 RNA extraction

RNA designated for quantitative real-time PCR for the transcript abundance time course experiment, was extracted using the TRIzol® reagent (Life Technologies GmbH, Darmstadt) following a standard protocol. RNA needed for library preparation and subsequent RNA sequencing was extracted using the RNeasy Plant Mini Kit (QIAGEN GmbH, Hilden) following the manufacturer protocol including the optional step of DNA digestion using DNAse I (Roche, Indianapolis, IN). Total RNA is obtained by both techniques.

For RNAseq of the cultivar *S. lycopersicum* cv. Moneymaker, total RNA was extracted from 100 mg tissue from one individual. For bulk segregant analysis (BSA), total RNA of all individuals within each of the bulks was pooled. This was

done by grinding frozen tissue of every individual (approximately 100 mg) on liquid nitrogen. 500 µl of Buffer RLT were added and the samples were vortexed. Then 100 µl of each of the 96 samples of the first bulk were transferred to a 15 ml falcon tube; the same was done for the 90 samples of the second bulk. The falcon tubes were vortexed thoroughly. Finally, 450 µl of each of theses mixtures were taken for RNA extraction with the RNeasy Plant Mini Kit starting with step 4 of the Handbook.

RNA concentrations were determined with a spectrophotometer (NanoDrop 1000, peqlab Biotechnologies GmbH, Erlangen).

3.5.3 Complementary DNA (cDNA) synthesis

cDNA was synthesized from 1 µg of total RNA with the SuperScript™ II Reverse Transcriptase (Life Technologies, Darmstadt) following the manufacturers instructions using Oligo(dT)$_{18}$ primers (Fisher Scientific – Germany GmbH, Schwerte).

3.5.4 Polymerase chain reaction (PCR)

PCRs were run in an Eppendorf Mastercycler® pro (Eppendorf AG, Hamburg).

For genotyping the Ampliqon *Thermus aquaticus* (*Taq*) DNA polymerase (Ampliqon A/S, Odense) with the following mixture was used:

1 µl	template DNA
1 µl	10x buffer
0.2 µl	dNTPs (10 mM each)
0.2 µl	forward primer (10 µM)
0.2 µl	reverse primer (10 µM)
0.1 µl	*Taq* DNA polymerase
7.3 µl	distilled water

The PCR program consisted of the following steps:

1.	95 °C	2 min
2.	95 °C	20 sec
3.	54 °C	30 sec
4.	72 °C	1 min
5.	72 °C	5 min

Steps 2 to 4 were repeated 35 times.

Colony PCRs to test for positive bacteria clones after transformation were performed following the PCR protocol for genotyping with the following modifications: instead of 1 µl of template DNA, a bacteria colony dipped into 2 µl of distilled water served as a template. Further, the PCR program only consisted of 27 cycles instead of 35 and step 4 was 2:30 minutes.

For candidate gene cloning the Phusion® High-Fidelity DNA polymerase (New England Biolabs GmbH, Frankfurt am Main) with the following recipe was used:

100 ng	template DNA / cDNA
10 µl	5x Phusion HF buffer
1 µl	dNTPs (10 mM each)
1.5 µl	forward primer (10 µM)
1.5 µl	reverse primer (10 µM)
0.5 µl	Phusion HF DNA polymerase
to 50 µl	distilled water

The PCR program was as follows:

1.	98 °C	30 sec
2.	95 °C	10 sec
3.	58 °C	30 sec
4.	72 °C	1:30 min
5.	72 °C	7 min

Steps 2 to 4 were repeated 35 times.

3.5.5 Restriction enzyme digestion

For CAPS markers the PCR fragments were digested with the required restriction enzyme to discriminate the two alleles, Every restriction enzyme recognizes a specific palindromic DNA sequence of 4 to 6 base pairs and cuts the DNA at this sequence motif. The enzymes used for the different markers are listed in the Tables A2 + A3 in the Appendix.

EcoRI, HindIII, PstI and Hpy188III (all from New England Biolabs GmbH, Frankfurt am Main) were used in the following mixture:

5 µl	PCR product
1.5 µl	10x buffer (according to the enzyme)
0.5 µl	enzyme

8 µl	H_2O

The mixture was incubated at 37 °C for 30 minutes.

TaqI (Fisher Scientific – Germany GmbH, Schwerte) was used with the following recipe:

5 µl	PCR product
1.5 µl	10x buffer (according to the enzyme)
1 µl	enzyme
7.5 µl	H_2O

The mixture was incubated at 65 °C for 60 minutes.

3.5.6 Gel electrophoresis

Gel electrophoresis is used to separate DNA fragments based on their length, either as a diagnostic tool, e.g. genotyping, or to isolate a specific DNA fragment, for example for molecular cloning. The DNA fragments from the CAPS genotyping markers were separated on agarose gels. Also PCR products for cloning candidate genes and all the colony PCR products were run on agarose gels. For the indel markers, polyacrylamide gels were used.

Agarose gels were prepared by dissolving 1.5 g agarose (my-budget Universal Agarose, Bio-Budget Technologies GmbH, Krefeld) in 150 ml 1x Tris-Acetate-EDTA (TAE) buffer (see recipe below). 4.5 µl ethidium bromide were added (10 mg/ml, Carl Roth GmbH & Co. KG, Karlsruhe). A horizontal gel system was used (PEQLAB Biotechnologie GmbH, Erlangen). 1 µl of loading dye was added to 10 µl of PCR product before loading the samples on the gel. To visualize the DNA fragments, pictures of the gels were taken on an UV-screen using a digital gel documentation system (Intas Science Imaging Instruments GmbH, Göttingen).

<u>50x TAE-buffer:</u>

242 g	Tris
57.1 g	Glacial acetic acid
100 ml	0.5 M EDTA (pH 8.0)

For the polyacrylamide gels a LI-COR DNA analysis system was used (4300 DNA Analyzer, LI-COR Biosciences GmbH, Bad Homburg). Gels were prepared by add-

ing 150 µl 10 % APS and 15 µl TEMED to 20 ml KB Plus 6.5 % Gel Matrix (LI-COR Biosciences GmbH, Bad Homburg). TBE (pH 8.3) was used as running buffer. 10 µl Stop-Mix (see recipe below) were added to 10 µl PCR reactions. Samples were then denatured for two minutes at 94 °C before loading them on the gel.

Stop-Mix:

95 ml	Formamid
0.05 g	Fuchsin
100 µl	NaOH (10 M)

Fill to 100 ml with pure water.

3.5.7 Quantitative real-time PCR

To determine the cycling expression behavior of the tomato homologs of the Arabidopsis clock genes *LATE ELONGATED HYPOCOTYL* (*LHY*) and *GIGANTEA* (*GI*), Solyc10g005080 and Solyc12g05660 respectively, their relative abundance was measured via qRT-PCR. The standard curve method was used for quantification and the AP-2 complex subunit mu (Solyc08g006960) as an internal control to normalize transcript abundance. One of the primers for each of the genes was spanning an exon-exon junction therefore not binding genomic DNA but only cDNA synthesized from spliced mRNA. The primer sequences can be found in the Appendix in Table 7.

Twenty µl reactions were set up in Eppendorf twin.tec PCR plates (Eppendorf AG, Hamburg) using IQ™ SYBR Green Master Mix (Bio-Rad Laboratories GmbH, München) and 500 ng cDNA in triplicates. PCR was performed in an Eppendorf Realplex cycler (Eppendorf AG, Hamburg). The PCR-program consisted of 50 cycles. The CT-values were calculated with the Eppendorf realplex2.2 software with the default method (Noiseband), which specifies a threshold of 10 standard deviations above the noise of the baseline.

3.5.8 RNA sequencing

RNA sequencing was carried out at the Max Planck-Genome-centre Cologne (MP-GC) located at the MPIPZ. Total RNA was provided for library preparation, which was performed by the MP-GC using an Illumina TrueSeq RNA sample preparation kit (Illumina, Inc., San Diego). Approximately 50 million reads were sequenced for each library with 100 bp single reads on the Illumina HiSeq2500

platform (Illumina, Inc., San Diego). The sequencing results were provided by the MP-GC as fastq-files.

The first step of analyzing the sequencing results was a quality check using FastQC. The sequencing quality of our reads was very high (average per base sequence quality = 34.5). Therefore no filtering or trimming steps were carried out based on the FastQC results. Next, the Bowtie2 software was used to generate an index file of the tomato genome, which is needed for subsequent mapping of the reads. The tomato reference genome sequence from the cultivar 'Heinz 1706' was used to do this (The Tomato Genome Consortium, 2012). Version 2.40 of this reference sequence was obtained as a fasta-files from the sol genomics network website (ftp://ftp.solgenomics.net/tomato_genome/wgs/assembly/build_2.40/). For mapping the reads to the reference genome, the software tophat2 was used with the following parameters: --max-insertion-length 12 --max-deletion-length 12 -g 1 -m 1 --read-gap-length 12 --read-edit-dist 12 --read-mismatches 12 --read-realign-edit-dist 0 --no-coverage-search --segment-mismatches 3. With those settings, an average 94.8 % of the reads could be mapped to the reference genome. After mapping, GATK was used to realign the reads around indels. This step is necessary to remove artifactual mismatches caused by misaligned reads due to indels. Finally, SNPs and indels were called from the final BAM-files with the UnifiedGenotyper module from GATK using the default parameters.

3.5.9 Candidate gene cloning

To test whether the candidate gene *EMPFINDLICHER IM DUNKELROTEN LICHT 1* (*EID1*) is underlying one of the identified QTL for circadian rhythms, the *S. pennellii* and the *S. lycopersicum* cv. M82 alleles were each cloned into two destination vectors: pGWB1 (Nakagawa et al., 2007) and pFAST-R01 (Shimada et al., 2010), used to transform tomato and Arabidopsis, respectively. The pFAST vectors have the advantage of using a fluorescent reporter visible in the seed coat, which allows the selection of transgenic T1 seeds without germination. Transgenic seeds can therefore directly be sown on soil without the need of antibiotics or other chemicals.

Because either expression differences or a protein change could be responsible for the QTL, both cDNA alleles were cloned behind both native promoters using the MultiSite Gateway® Pro 2.0 Kit (Life Technologies, Darmstadt). In *S. lycopersicum* the sequence between the end of the upstream gene and the start codon of *EID1* is approximately 2500 bp. The upstream 2465 base pairs starting from the start codon were cloned as the promoter of *S. lycopersicum*. *S. pennellii* has a 2100 bp insertion compared to *S. lycopersicum* starting 250 bp upstream of the start codon. The upstream 2484 base pairs starting from the start codon were cloned as the promoter of *S. pennellii*. The two alleles of the gene were cloned from cDNA and included 45 bp of downstream sequence representing the 3' untranslated region (UTR). The primers used can be found in the Appendix (Table 8).

The PCR for the amplification of the promoters also yielded some unspecific products. Therefore the promoters were first cloned into the pGEM®-T easy vector (Promega GmbH, Mannheim) before subcloning them into the appropriate Gateway pDONR™ vector via the BP recombination reaction. The pGEM®-T easy vector system requires an unpaired adenine (A) at the end of the DNA fragments to be cloned. Such an A-tail is added by *Taq* DNA polymerases but not by the Phusion® High-Fidelity DNA polymerase used for the PCR reaction. Therefore the PCR product had to be A-tailed.

- First the PCR product was cleaned using the QIAquick PCR purification Kit (QIAGEN GmbH, Hilden) following the manufacturer protocol using 32 µl Buffer EB.
- 30 µl cleaned PCR product were used for A-tailing by adding 0.4 µl ampliqon *Taq* DNA polymerase (Ampliqon A/S, Odense), 4 µl 10x PCR buffer, 0.32 µl dATP (25 mM) and 5.28 µl water and incubating the mixture at 72 °C for 30 minutes.
- The 40 µl reaction was then loaded on an agarose gel, the appropriate band was excised and the DNA extracted from the gel slice using the QIAquick gel extraction Kit (QIAGEN GmbH, Hilden) according to the manufacturer protocol.
- For cloning the DNA fragment, the pGEM®-T easy vector system was used following the manufacturers protocol. Step 2 of the ligation was modified:

0.5 µl of vector and 3.5 µl of DNA were used for the ligation reaction because the DNA concentrations were very low with 2-5 µg/µl. The reaction was incubated at 4 °C overnight. The LB plates (see recipe below) for transformation were prepared with carbenicillin (100 µg/ml) instead of ampicillin. Step 2 of the transformation was also modified: 4 µl of ligation reaction were used. White colonies were selected and tested for the presence of the promoter via colony PCR.

- Positive clones were cultured in 3 ml of LB medium containing 3 µl of carbenicillin (100 mg/ml) overnight at 37 °C. The plasmids containing the promoter were isolated from those cultures using the QIAprep Spin Miniprep Kit (QIAGEN GmbH, Hilden) following the manufacturers protocol.

The PCR for the amplification of the two alleles of the gene produced one single band when visualized on an agarose gel. Therefore PCR products were purified as described in the MultiSite Gateway® Pro manual without the use of a gel. These purified PCR products and the isolated plasmids containing the promoters were used for the BP recombination reaction inserting the DNA fragments into the appropriate Gateway pDONR™ vectors. This was done according to product manual, but with half the reaction volume. To ensure that the sequences did not have any artifactual polymorphisms introduced by the PCR amplification, all entry clones were sequenced via Sanger sequencing at the MP-GC and aligned to the corresponding reference sequences using the Lasergene® software SeqMan (DNASTAR, Inc., Madison). The subsequent LR recombination reactions, which combined the promoters with the genes, were also conducted with half the reaction volume. All further steps were carried out according to the manufacturers manual.

LB-plates / medium:

10 g	NaCl
5 g	Yeast extract
10 g	Tryptone/Peptone
12.5 g	Agar

Fill to one liter with pure water. For LB medium no Agar is added.

3.5.10 Transformation of *Agrobacterium thumefaciens*

The *Agrobacterium thumefaciens* strain PM90, harboring rifampicin and gentamycin resistances, was used as a vector to transfer the different alleles of the candidate gene into Arabidopsis and tomato. *Agrobacterium* was transformed with the different expression clones using electroporation. For this 5 ng of plasmid DNA were mixed with 50 µl of electro-competent *Agrobacterium* cells in a 1 mm cuvette. The electroporation was performed with the following parameters:

Capacitance:	25 µFD
Voltage:	2.2 kV
Resistance:	200 Ω
Pulse length:	4-5 seconds

After applying the electric pulse, 950 µl YEP medium (see recipe below) were added. The mixture was shaken for two hours at 28 °C. 10 µl were then plated on YEP plates containing rifampicin, gentamycin and the antibiotics corresponding to the resistances of the according destination vectors (spectinomycin and hygromycin for pFAST-R01 and kanamycin and hygromycin for pGWB1). The final concentrations of the antibiotics were 10 µg/ml for gentamycin and 50 µg/ml for the others. Transformation was confirmed via colony PCR with the according specific primers.

YEP-plates / medium:

5 g	NaCl	
10 g	Yeast extract	
10 g	Tryptone/Peptone	
5 g	Sucrose	→ pH 7.5 (NaOH)
15 g	Agar	

Fill to one liter with pure water. For YEP medium no Agar is added.

3.5.11 Arabidopsis transformation

10 ml YEP medium in small flasks containing the appropriate antibiotics was inoculated with transformed Agrobacteria and incubated at 28 °C for two days in a shaker (220 rpm). On the third day 400 ml YEP medium in big flasks containing the appropriate antibiotics was inoculated with the 10 ml starter culture and in-

cubated overnight at 28 °C in a shaker (200 rpm). The next day the 400 ml culture was centrifuged at 5000 rpm for 20 minutes at room temperature. The pellet was resuspended in 400 ml sucrose solution (20 g sucrose + 400 ml water) with 100 µl Silwet L-77 (LEHLE SEEDS, Round Rock). This suspension has an OD_{600} of approximately 0.8. Pots with flowering Arabidopsis plants are inverted and dipped into the solution for one minute. Afterwards, plants are kept under a cover, which increases the humidity, for one to two days and then grown in the greenhouse until harvest.

3.5.12 Tomato transformation

Sterile seeds are sown into glass jars containing MS medium (see recipe below) and put into a light room (16 h light / 8 h dark) at 25 °C. Glasses must not be sealed airtight. To sterilize the tomato seeds, they are first put into saturated trisodium phosphate (Na_3PO_4) for ten minutes. Then they are rinsed with sterile water and put into 70 % ethanol for ten seconds. Afterwards, the seeds are put into 5 % sodium hypochlorite (NaClO) for ten minutes and then rinsed with sterile water again.

Agrobacteria containing the vector of interest are plated on YEP-plates containing the appropriate antibiotics and incubated at 28 °C for three days. On the third day 10 ml LB medium in small flasks containing the appropriate antibiotics are inoculated with these bacteria and cultured at 28 °C in a shaker with 220 rpm. On the same day, leaves of four to six weeks old sterile tomato plants are cut into pieces (explants) of approximately 0.5 to 1 cm^2 in a petri dish with a bit of sterile water. These explants are placed with their abaxial side on MS-plates (see recipe below) containing Zeatin, Acetosyringon and IAA (1 mg, 36 mg and 2 mg respectively in 500 ml). Plates are closed with parafilm. The next day, the *Agrobacterium* cultures are diluted 1:20 with LB medium. Explants are taken from the plates, on which they were placed the day before, and put into the *Agrobacterium* solution for 15 minutes. Afterwards, the explants are very shortly placed on sterile filter paper and put back onto the MS-plates and kept in the dark at 25 °C for two days. Plates are sealed with Leucopore tape (Duchefa Biochemie B.V, Haarlem), which is air-permeable. After the two days in the dark, explants are shortly placed on sterile filter paper again and then transferred to MS-plates containing

Zeatin (1 mg/ml), Carbenicillin (250 mg/ml) and the appropriate antibiotic (Kanamycin (50 mg/ml) for pGWB1) (1 ml each in 500 ml). After about three weeks, explants that remained green are transferred to fresh MS-plates with the same ingredients. After approximately six weeks, shoots of these explants are cut and transferred to MS-glasses containing Carbenicillin (250 mg/ml) and the appropriate antibiotic (Kanamycin (50 mg/ml) for pGWB1) (1 ml and 0.5 ml in 500 ml, respectively). Once these shoots have developed roots they are transferred to soil and put to the greenhouse. The existence of the transgene in the plants is validated via PCR with specific primers.

MS-medium

4.3 g	MS + vitamins
100 mg	myo-Inositol
30 g	sucrose
7 g	Phyto-Agar

Fill to one liter with pure water and adjust to pH 5.9 with a few drops of 3 M KOH.

3.6 Data analysis

All data analyses, i.e. statistical tests and plotting, were done using R (http://www.r-project.org). For QTL mapping the R-package 'qtl' was employed (Broman and Sen, 2009).

4 Results

4.1 Variation in circadian rhythms in tomato and its wild relatives

As a first step in studying the circadian clock of tomato, leaf movement rhythms of seedlings were analyzed in constant light. Rhythmic leaf movements have been shown to be an output of the circadian clock in several plant species and have been used to study natural variation in circadian rhythms (Bünning, 1932; Lou et al., 2011; Salathia et al., 2007; Swarup et al., 1999). The leaf movements of the cultivated tomato variety *S. lycopersicum* cv. M82 were compared with those of the closest wild relative of cultivated tomato, the red-fruited wild tomato species *S. pimpinellifolium* (LA1589), and with the distantly related green-fruited wild tomato species *S. pennellii* (LA0716). The three species showed very clear differences in circadian leaf movement rhythms (Figure 9). The rhythms were characterized by fitting sinusoid curves to the time course data of every individual seedling to estimate circadian parameters such as period, phase and amplitude. The three species tested exhibit pronounced differences in period. The wild tomato species *S. pennellii* shows periods of only around 22 hours while the cultivated variety exhibits much longer periods of about 26 to 28 hours (Figure 9B). The wild species *S. pimpinellifolium* has periods in between cultivated tomato and *S. pennellii*. As can be seen in Figure 9A, the first circadian cycle of the cultivated tomato variety exhibits pronounced asymmetry. The leaves of all three species move up during the first subjective light to dark transition and start moving down again in the middle of the first subjective night. After that, however, the cultivated variety continues the downward movement for more than 16 hours before moving up again. This asymmetric waveform, together with the period differences, leads to a very late second peak of *S. lycopersicum* compared to the wild species. The parameter “peak time” was defined as the time of the second peak of the idealized rhythm calculated from the period and phase estimates (for details see section 3.3.1).

The peak time estimates clearly distinguish the cultivated tomato variety from its wild relatives (Figure 9B). All M82 seedlings exhibit later peak times than the seedlings of the wild species. Whereas there is some overlap between *S. lycopersicum* cv. M82 and *S. pimpinellifolium* for period there is none for peak time.

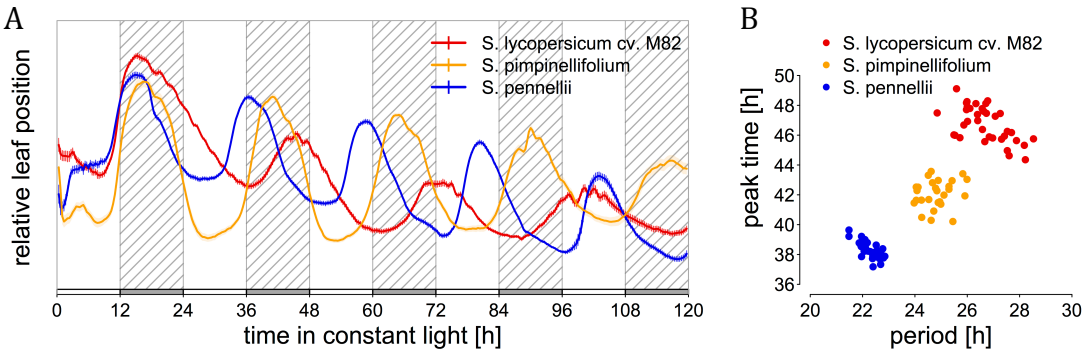

Figure 9: Variation in circadian leaf movements of cultivated tomato and two of its wild relatives. (A) Mean relative position of cotyledon tip ± standard error of the mean (SEM) from between two and three independent experiments for *S. lycopersicum* cv. M82 (n = 36), *S. pimpinellifolium* (n = 27) and *S. pennellii* (n = 36). Hatched areas indicate subjective nights. (B) Circadian period and peak time estimates of the same individuals shown in A.

Because leaf movements are an output of the circadian clock that is several levels away from the core oscillator, it was tested whether the observed differences are due to differences in the molecular circadian clock. To do this, transcript abundance of the tomato homologs of the *Arabidopsis thaliana* core clock genes *LATE ELONGATED HYPOCOTYL* (*LHY*) and *GIGANTEA* (*GI*), Solyc10g005080 and Solyc12g05660, respectively, was examined. Samples were taken from seedlings of the same three genotypes shown in Figure 9 every four hours for three days. Their RNA was extracted and qRT-PCR performed with primers for *LHY* and *GI*. For LHY the transcript abundance was analyzed for all three species, for *GI* for *S. lycopersicum* and *S. pennellii*. The circadian transcript rhythms of *LHY* (Figure 10) and *GI* (Figure 11) resemble the leaf movement rhythms. These results demonstrate that rhythmic leaf movements follow the molecular clock and are indeed a suitable marker to study the circadian clock of tomato. In addition, it was confirmed that there is variation in circadian rhythms between *S. lycopersicum*, *S. pimpinellifolium* and *S. pennellii* at the molecular level.

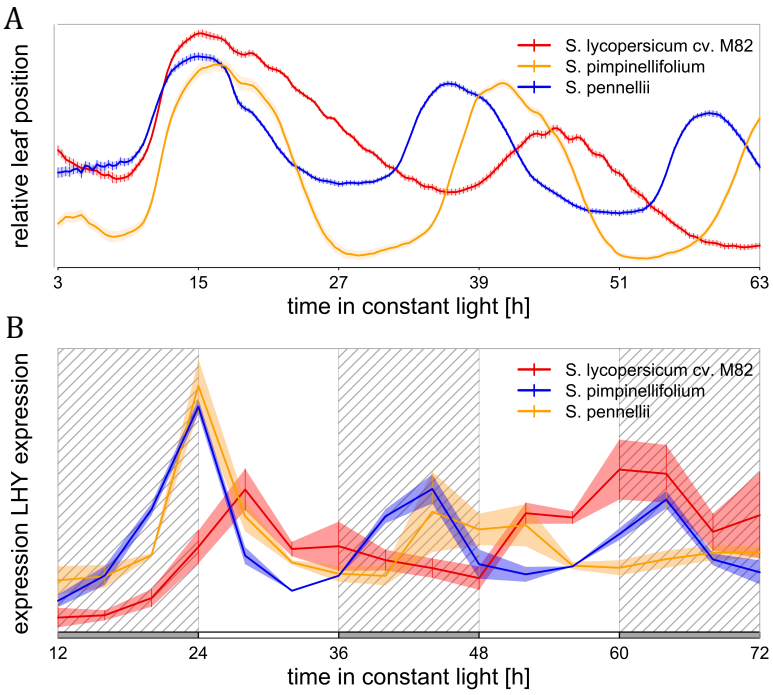

Figure 10: Comparison of circadian leaf movements and *LHY* transcript rhythms. (A) Mean relative position of cotyledon tip ± SEM from between two and three independent experiments for *S. lycopersicum* cv. M82 (n = 36), *S. pimpinellifolium* (n = 27) and *S. pennellii* (n = 36). (B) Mean transcript abundance of the tomato *LHY* homolog (Solyc10g005080) from two biological replicates (shaded area ± SEM); data are normalized against AP-2 complex subunit mu (Solyc08g006960). Hatched areas indicate subjective nights.

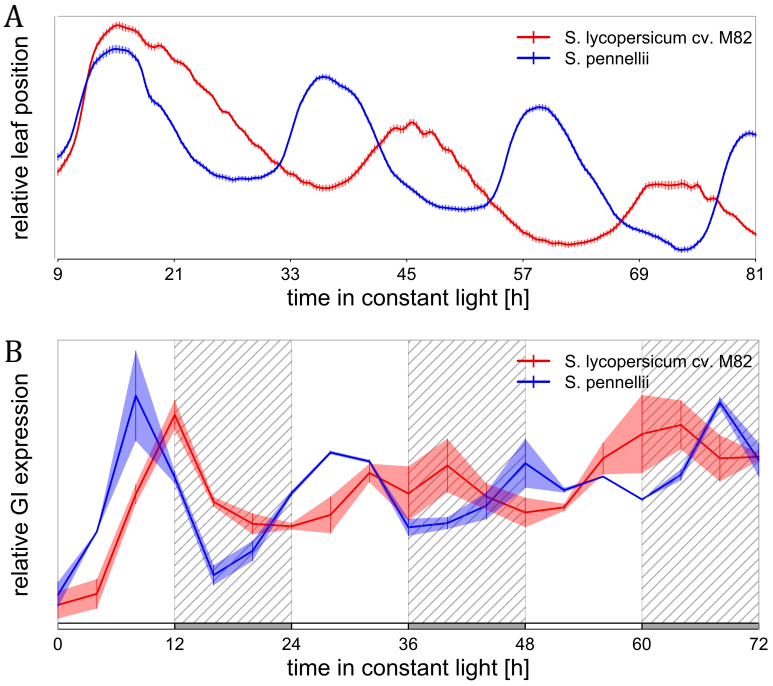

Figure 11: Comparison of circadian leaf movements and *GI* transcript rhythms. (A) Mean relative position of cotyledon tip ± SEM from three independent experiments for *S. lycopersicum* cv. M82 and *S. pennellii* (n = 36 each). (B) Mean transcript abundance of the tomato *GI* homolog (Solyc12g05660) from two biological replicates (shaded area ± SEM); data are normalized against AP-2 complex subunit mu (Solyc08g006960). Hatched areas indicate subjective nights.

4.2 Patterns of variation in circadian leaf movements suggest an effect of domestication on the tomato circadian clock

To determine whether the observed differences in circadian rhythms between the cultivated tomato variety *S. lycopersicum* cv. M82 and the two wild tomato species *S. pennellii* (LA0716) and *S. pimpinellifolium* (LA1589) are specific to these genotypes, 81 accessions of 11 different tomato species were analyzed. Among those, 47 accessions belong to *S. lycopersicum*, including eight wild cherry tomatoes, six Latin-American cultivars and 33 'modern' varieties (referring to varieties developed after the transfer of tomato to Europe in the 16th century). The other 34 accessions belong to 10 of the 11 wild tomato species (a detailed list can be found in the Appendix Table 4).

The circadian period and peak time estimates of 67 of these genotypes (all but the wild cherries and Latin-American cultivars) differentiate three conceptual groups: green-fruited wild species, red-fruited wild species and cultivated varieties (Figure 12A).

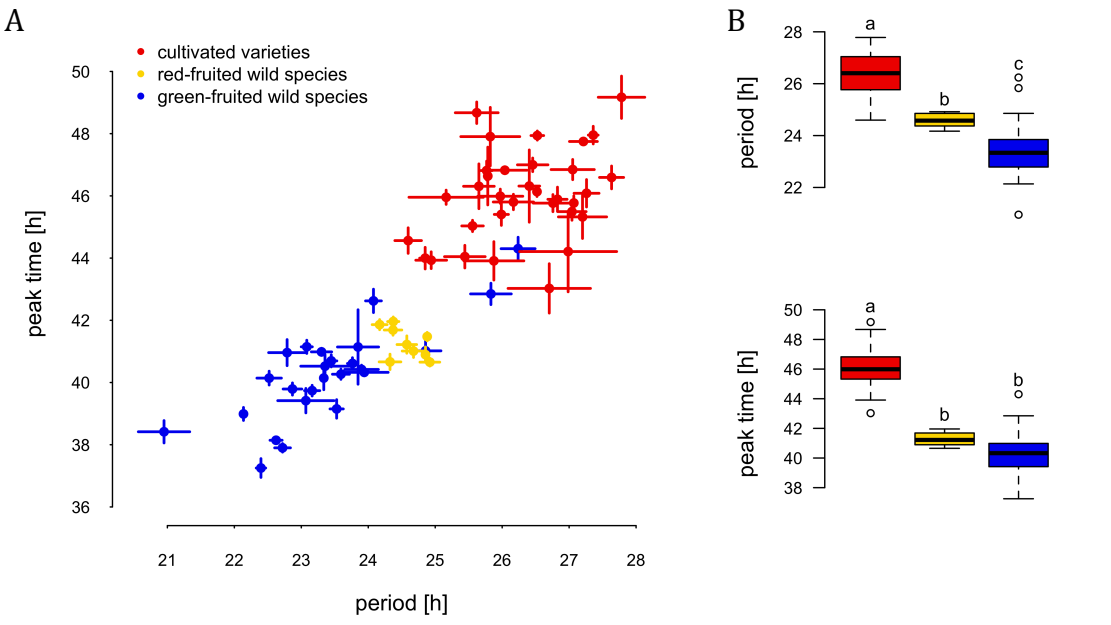

Figure 12: Patterns of variation for three groups of the tomato clade: cultivars, red-fruited and green-fruited wild species. (A) Mean period and peak time estimates ± SEM of 33 cultivars and 34 wild species accessions (n= 2-5). (B) Boxplots of the two traits by group; different letters indicate significantly different means (Tukey's HSD, $p < 0.05$).

All the modern cultivars tested cluster together with much longer periods and later peak times (mean = 26.3 and 46.0, respectively) than the green-fruited wild

species accessions (mean = 23.4 and 40.3). The red-fruited wild species accessions exhibit periods and peak times in between but closer to the green-fruited wild species (mean = 24.6 and 41.3). The mean period and peak time of the cultivated varieties are significantly different from both the mean period and peak time of the green- and the red-fruited wild species (Figure 12B and Table 1, Tukey's HSD < 0.01). For the two groups of wild species only the mean period is significantly different (Tukey's HSD < 0.01) while the mean peak time is not (Tukey's HSD = 0.179) (Figure 12B and Table 1).

Table 1: Mean periods and peak times and statistical comparisons of the three groups – cultivars, red-fruited and green-fruited wild species. Data are based on the 67 genotypes shown in Figure 12. Groups: 1 = modern cultivars, 2 = red-fruited wild species, 3 = green-fruited wild species

period	mean	SD	min	max	p-value (Tukey's HSD)		
group					1	2	3
1	26.3	24.6	24.6	27.8		< 0.001	< 0.001
2	24.6	0.27	24.2	24.9			0.004
3	23.4	1.10	21.0	26.2			
peak time	mean	SD	min	max	p-value (Tukey's HSD)		
group					1	2	3
1	46.0	1.46	43.0	49.2		< 0.001	< 0.001
2	41.3	0.50	40.7	42.0			0.179
3	40.3	1.56	37.3	44.3			

These results demonstrate that the differences observed between the three species shown in Figure 9 are representative of three conceptual groups of tomato species. Faster to slower rhythms in these species differentiate green-fruited wild species, red-fruited wild species and cultivated varieties. The differences between the cultivated varieties and their ancestors, the red-fruited wild species, strongly suggest that domestication or early breeding has had an effect on the circadian clock of tomato.

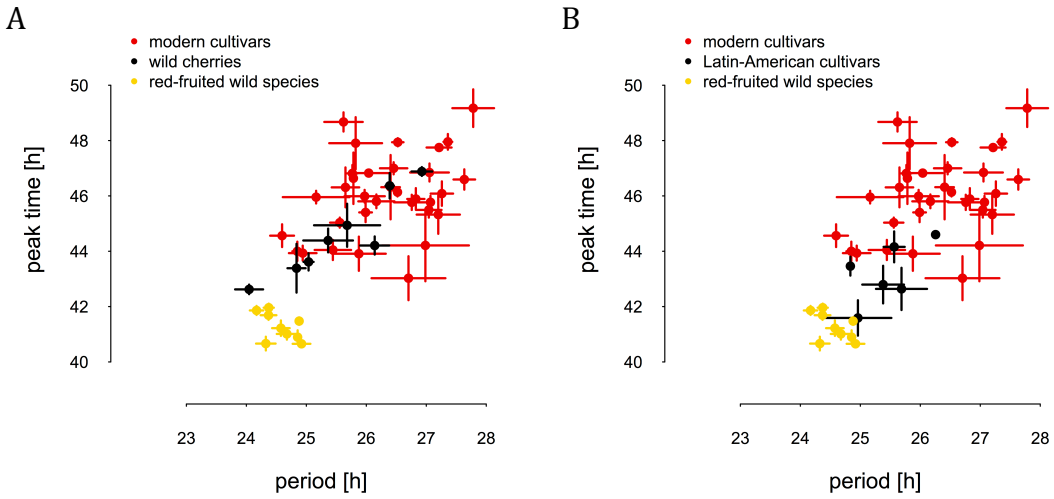

Figure 13: Period and peak time estimates for wild cherries and Latin-American cultivars. (A) Mean period and peak time estimates ± SEM (n= 2-5) for 8 wild cherries compared to cultivars and red-fruited wild species. (B) Mean period and peak time estimates ± SEM (n= 2-5) for 6 Latin-American cultivars compared to the other cultivars and red-fruited wild species.

To further investigate the role of circadian rhythms during domestication, *S. lycopersicum* accessions classified as ancestral to the modern varieties were employed. One of these groups comprises the wild cherry tomatoes, which are often classified as *S. lycopersicum* var. *cerasiforme*. It is important to note that while this group mainly consists of plants that are an admixture of cultivated tomato and *S. pimpinellifolium* and ancestral cultivars, some of them seem to be regular cultivars that were wrongly classified (Blanca et al., 2012; Ranc et al., 2008). The period and peak time estimates for most of the tested wild cherry tomatoes overlap with those of the cultivars (Figure 13A). However, some of the wild cherries have shorter periods and earlier peak times closer to the red-fruited wild species. Further experiments are necessary to test whether these short period accessions are ancestral cultivars or an admixture between cultivated tomato and *S. pimpinellifolium*. Six of the tested tomato genotypes are classified as South or Central American cultivars and, like some of the wild cherries, have shorter periods and especially earlier peak times than the modern cultivars (Figure 5B). It must be noted that several studies agree on the difficulty of unambiguously classifying these genotypes, and these results should be interpreted with care (Blanca et al., 2012; Rick, 1958). Nevertheless, if the accessions with short period and early peak time are indeed ancestral cultivars, the results suggest that the circadian clock of cultivated tomato has been gradually changed during domesti-

cation and/or breeding, with the early cultivars from South and Central America still more closely resembling their wild relatives. Tomato sequentially moved away from the equator during domestication. It is possible that modulation of circadian rhythms drove adaptation to higher latitudes and was artificially selected for by humans.

4.3 QTL mapping for circadian leaf movements

To identify the molecular basis for the differences in circadian rhythms observed between the wild species and the cultivated varieties, QTL mapping was performed in two populations: One generated from a cross between *S. lycopersicum* cv. M82 and the green-fruited wild species *S. pennellii* (Eshed and Zamir, 1995); the other one generated from a cross between *S. lycopersicum* cv. Moneymaker and the red-fruited wild species *S. pimpinellifolium* (Voorrips et al., 2000). QTL present in both populations could potentially represent the loci responsible for the general differences between cultivated and wild tomato and could have played a role during domestication.

4.3.1 QTL mapping in an *S. pennellii* introgression line (IL) population

The *S. pennellii* population consists of 76 introgression lines generated from a cross between the cultivar *S. lycopersicum* cv. M82 and the wild tomato species *S. pennellii* (LA0716) (Eshed and Zamir, 1995). This population is widely used and has led to the identification of many QTL controlling various traits and even to the cloning of genes underlying QTL for fruit size and sugar content (Frary et al., 2000; Fridman et al., 2000). Circadian leaf movements of 69 of the 76 lines of the population were analyzed, providing almost complete coverage of the genome. Parts of chromosomes 1, 3 and 5, present in lines 1-1, 3-4 and 5-3, were missing, because no seeds were available for those lines. In addition, parts of chromosomes 5 and 8 are not covered by any IL of the population (Chitwood et al., 2013).

4.3.1.1 Phenotypic variation

Each line of the population was grown in one of five independent experiments, the recurrent parent *S. lycopersicum* cv. M82 in all five. Figure 14 shows the mean period and phase estimates of all the lines with at least two biological rep-

licates. The period and phase means of the ILs are normally distributed (Shapiro-Wilk test, p = 0.822 and p = 0.508, respectively). Notably, transgressive segregation is present for period and for phase. The normal distribution and wide range of trait values compared to the parental lines suggest that circadian period and circadian phase are quantitatively controlled by several genes.

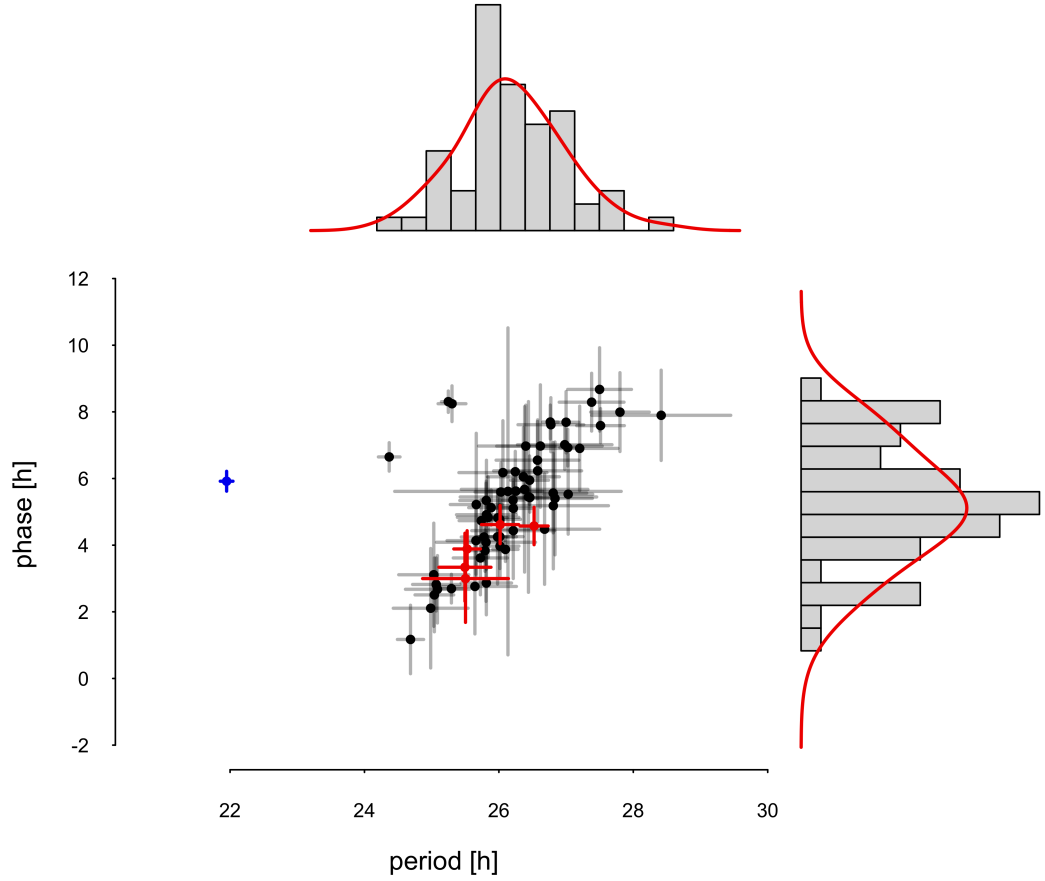

Figure 14: Variation in circadian leaf movements for the *S. pennellii* introgression line population. Mean period and phase estimates ± SEM of the 62 ILs with at least two biological replicates in black and *S. pennellii* in blue (n = 2-8), each grown in one of five independent experiments, and of the recurrent parent *S. lycopersicum* cv. M82 in red (n = 6-8) grown in all five experiments. Histograms on the sides show the distributions of the means of the ILs; the red lines indicate the density.

The five independent measurements for M82 show that there is variation across experiments, so even minor differences in the controlled environment can have an influence on the rhythmic properties of the seedlings (Figure 14). However, the high broad sense heritability (H^2), which is 61 % for period and 59 % for phase, indicates that the differences observed between the lines are mostly determined by their genotype.

Figure 14 also shows a correlation between period and phase (Pearson's ρ = 0.604, p = 2.01e-08). Only three lines show a combination of higher phase and shorter period means, clearly distinguishing them from all the other genotypes (see top-left three points in Figure 14). Those three lines, IL9-2-6, IL9-3 and IL7-4-1, also showed the most distinct leaf movement traces when compared to M82 or other random ILs (Figure 15).

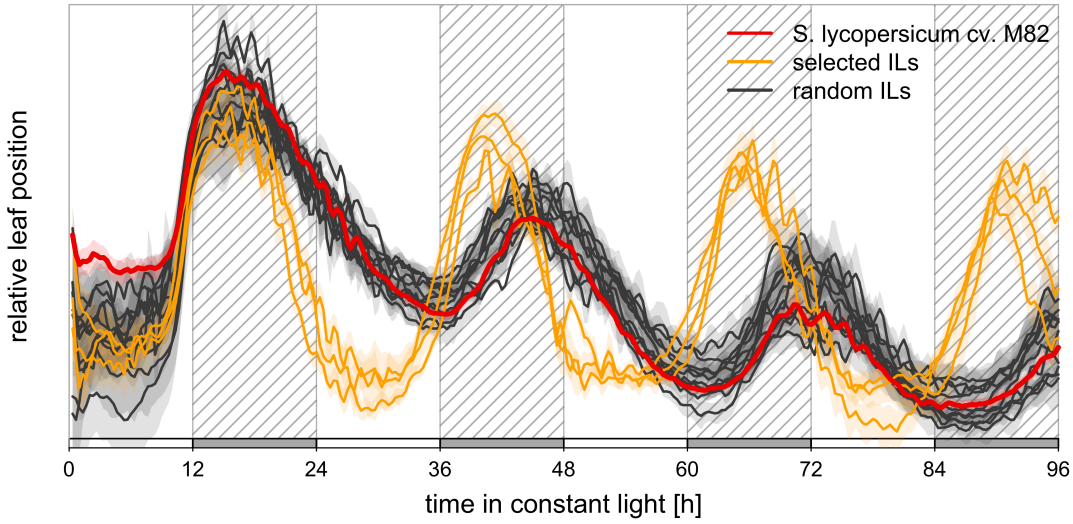

Figure 15: Circadian leaf movements of M82, three selected and 10 random ILs. Mean relative position of cotyledon tip (shaded area: ± SEM) (n = 3-8 for the ILs; n = 30 for M82). IL9-2-6, IL9-3 and IL7-4-1 are shown in orange. A set of ILs randomly selected from the population is shown in black. Each IL was grown in one of five independent experiments, while M82 was grown in all. Hatched areas indicate subjective nights.

Figure 15 again illustrates an asymmetry of the first circadian cycle for M82 and the random ILs. This asymmetry was already noted when comparing cultivated tomato to its wild relatives (Figure 9). The rhythms of the three outlier ILs do not exhibit this asymmetric first waveform, as their leaves move up again at the end of the second subjective day, leading to an earlier second peak in comparison with M82 and the other randomly chosen lines (Figure 15). As a matter of fact, when calculating the time of the second peak of all the ILs, referred to as peak time, IL9-2-6, IL9-3 and IL7-4-1 differ more from the rest of the population than they do when using the phase values (Figure 16). Also the heritability for peak time is higher than for the other traits (H^2 = 69.4 %). For this reason further analysis focused on period and peak time.

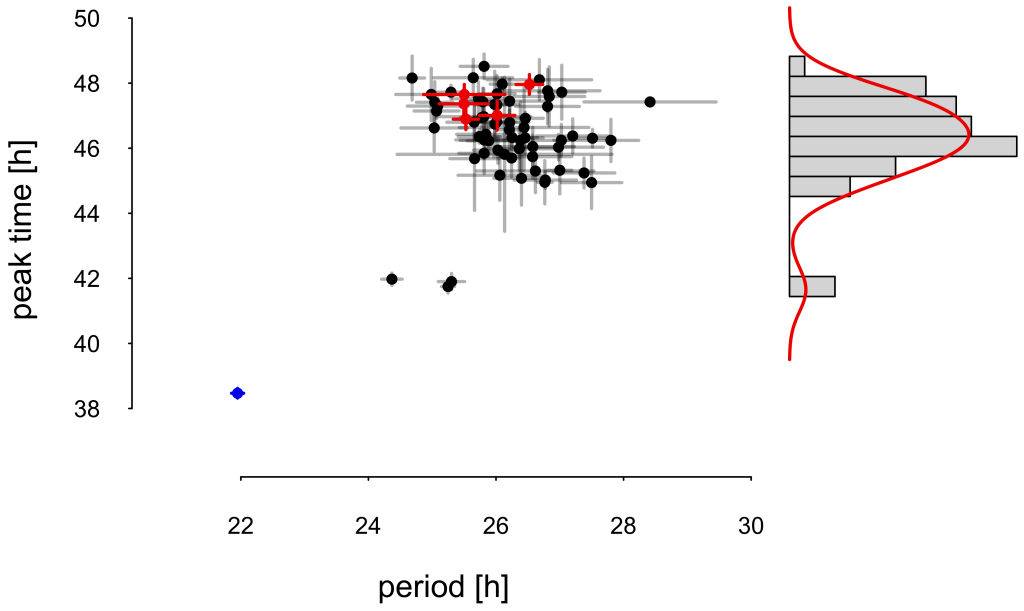

Figure 16: Variation in peak time for the *S. pennellii* introgression line population. Mean period and peak time estimates ± SEM for 62 ILs in black and *S. pennellii* in blue (n = 2-8), each grown in one of five independent experiments, and for *S. lycopersicum* cv. M82 in red (n = 6-8) grown in all five experiments. Histogram on the side shows the distribution of the mean peak times of the ILs; the red line indicates the density.

Nevertheless, two other important parameters for rhythmic time-course data, amplitude and relative amplitude error (RAE), were also explored. For leaf movement analyses, differences in amplitude should be interpreted with caution, because differences in leaf morphology can have a direct influence on this variable. For example, genotypes with smaller cotyledons, like *S. pennellii*, will have lower amplitudes than seedlings with longer cotyledons even though their circadian clock may be running with the same strength. There is only one line in the population, IL7-4-1, that exhibits a significantly higher amplitude than M82 (Figure 17A, linear regression, $p < 0.01$). Interestingly, this is one of the three lines with an early peak time and symmetric waveform already discussed. Additionally, there are 13 lines with significantly lower amplitudes than M82 (linear regression, $p < 0.01$).

Relative amplitude error (RAE) is a measure of the goodness of fit of a sinusoid to the time-course data, and is often used in circadian research to estimate the robustness of a rhythm. RAE can theoretically have values between zero and one. If a rhythm perfectly resembled a sine wave the RAE of this rhythm would be zero. The less robust a rhythm gets, the greater its RAE will be (Plautz et al., 1997).

There are two lines (IL3-2 and IL10-1-1) with significantly higher and one line (IL7-1) with significantly lower relative amplitude errors than M82 (linear regression, $p < 0.01$). However, it would go beyond the scope of this thesis to further analyze those three QTL for RAE.

The data reveal a positive correlation between period and RAE (Pearson's ρ = 0.611, p = 3.224e-08) (Figure 17B). This means that lines with longer periods exhibit less robust rhythms. Interestingly, this correlation is only true for the lines of the *S. pennellii* IL population. When looking at the period and RAE values for the 83 accessions tested, there is no correlation between the two traits (Pearson's ρ = 0.105, p = 0.352).

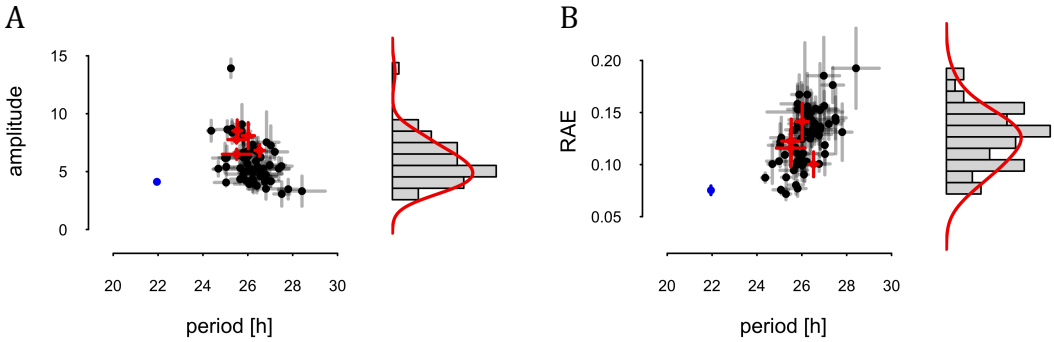

Figure 17: Variation in amplitude and RAE for the *S. pennellii* introgression line population. Mean amplitude (A) and relative amplitude error (RAE) (B) estimates ± SEM of 62 ILs in black and *S. pennellii* in blue (n = 2-8), each grown in one of five independent experiments, and of *S. lycopersicum* cv. M82 in red (n = 6-8) grown in all five experiments. Histograms on the side show the distribution of the means of the ILs; the red lines indicate the density.

4.3.1.2 QTL detection

Seven of the 69 lines tested of the *S. pennellii* population exhibited a significantly different period and six a significantly different peak time (linear regression, $p < 0.01$) compared to the recurrent parent M82 within the same experiment (Table 2).

Table 2: ILs showing significant (linear regression, $p < 0.01$) differences in period or peak time compared to M82. Comparisons are made between M82 and all the genotypes with at least two biological replicates within one experiment. Data are from five independent experiments.

trait	genotype (IL)	difference (IL –M82)	p-value
period	1-1-2	2.00	0.0084
period	1-1-3	2.40	0.0003
period	4-3-2	2.30	0.0026
period	9-2-6	-1.65	0.0007
period	12-1	1.49	0.0036
period	12-3	1.36	0.0046
period	12-4	1.68	0.0016
peak time	1-1-2	-2.71	0.0042
peak time	4-3	-1.72	0.0078
peak time	7-4-1	-4.87	2.23e-13
peak time	9-1	-2.05	0.0029
peak time	9-2-6	-5.03	1.01e-09
peak time	9-3	-5.75	3.04e-07

Since every IL only contains a single genomic fragment from *S. pennellii* in the M82 background, the genomic fragment of an IL significantly differing from M82 can be considered a QTL. Lines with overlapping introgressions can be used to narrow down the QTL. Based on recently published genotype data (Chitwood et al., 2013), five putative period and five putative peak time QTL, with 2 QTL overlapping between the two traits (on chromosomes 1 and 9), were identified (Figure 18).

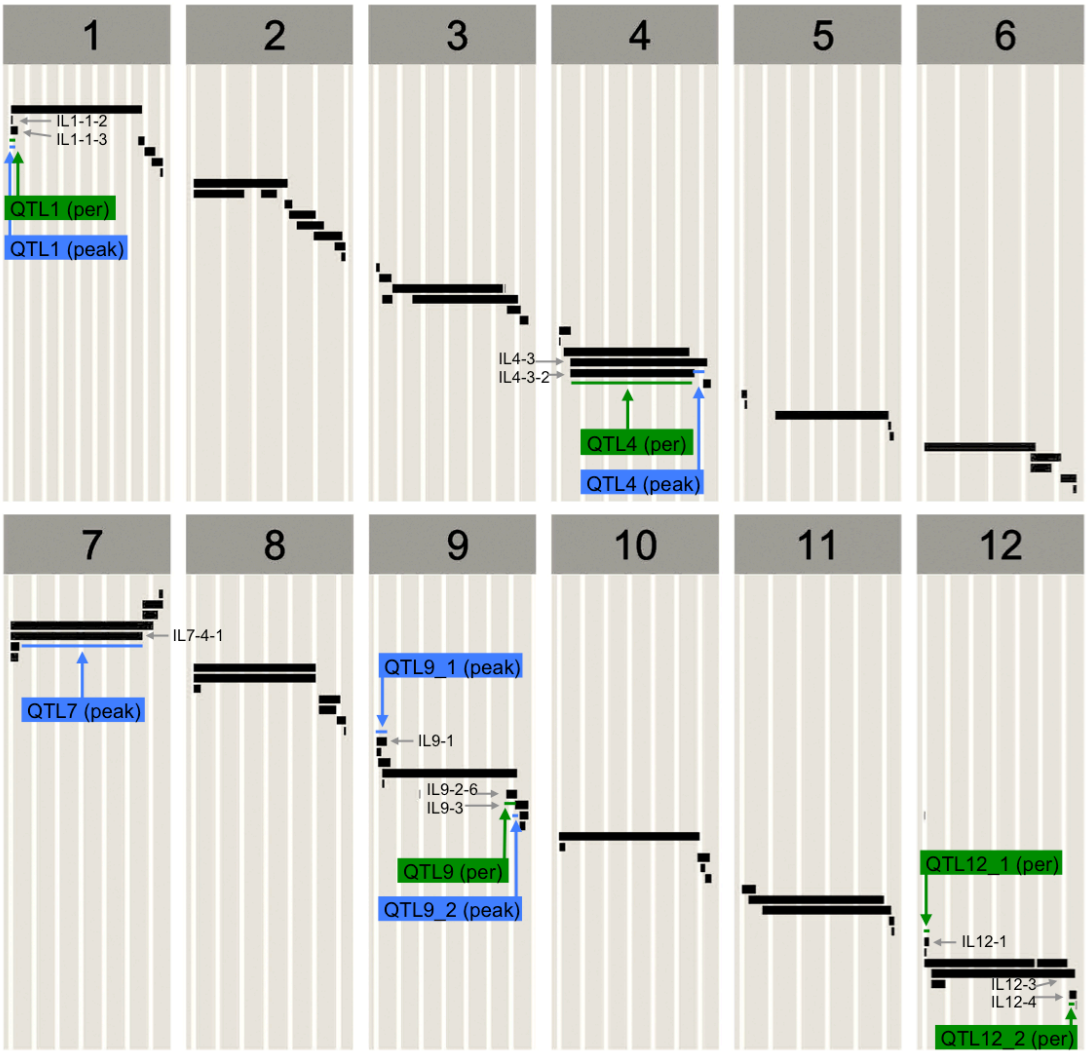

Figure 18: Putative QTL identified in the *S. pennellii* IL population. The 12 boxes represent the 12 chromosomes, the black rectangles inside the boxes the introgressions of the ILs for the respective chromosome. ILs significantly different from M82 (linear regression, $p < 0.01$) are marked with their name. QTL based on those ILs and ILs with overlapping introgressions are indicated with an arrow and their name. The colored horizontal lines show the putative chromosomal region of the QTL. Period QTL are shown in green, peak time QTL in blue. Modified from Chitwood et al. 2013.

4.3.2 QTL mapping in an *S. pimpinellifolium* recombinant inbred line (RIL) population

As already mentioned, *S. pimpinellifolium* is the closest wild relative of cultivated tomato. Thus, mapping of QTL responsible for differences in circadian rhythms between *S. pimpinellifolium* and *S. lycopersicum* may lead to the identification of circadian clock genes that were changed during domestication or breeding. A QTL analysis for circadian leaf movements was performed using a RIL population generated at Wageningen University from a cross between the cultivated tomato variety *S. lycopersicum* cv. Moneymaker (MM) and the wild tomato spe-

cies *S. pimpinellifolium* (CGN 15528) (Voorrips et al., 2000). 90 individuals of this population that have been genotyped with 1969 markers were used (Kazmi et al., 2012).

4.3.2.1 Phenotypic variation and QTL detection

As in the *S. pennellii* IL population, there is transgressive segregation for period but not for peak time (Figure 19). The broad sense heritability (H^2) is very high for both traits: 86.0 % for period and 85.7 % for peak time. This means that most of the observed phenotypic variance can be explained by the genotype. The mean period and peak time estimates are not normally distributed (Shapiro-Wilk test, p = 0.005 and p = 9.442e-05, respectively). Notably, the distribution of peak time is bimodal. Approximately 75 % of the individuals have peak times similar to *S. pimpinellifolium* while the others have peak times similar to Moneymaker (Figure 19). This suggests that there are one or two genes with a very strong effect on this trait.

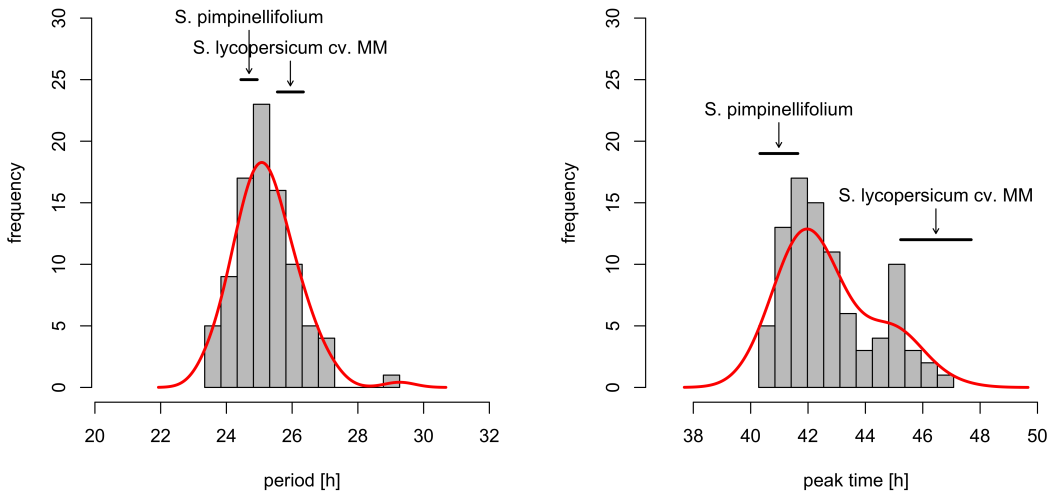

Figure 19: Period and peak time distribution of the *S. pimpinellifolium* RILs. Histograms of the mean period and peak time estimates of 90 lines of the *S. pimpinellifolium* RIL population (n = 3-9). Each RIL was grown in two of five independent experiments. Red lines show the density. Arrows indicate the mean value for the parents of the population, the horizontal line underneath their 99 % confidence interval.

The QTL analysis revealed three highly significant QTL (logarithm of the odds (LOD) score > 5), one for period and two for peak time (Figure 20). The peak time QTL are located on chromosomes 9 and 1 and account for 42.2 % and 23.6 % of the explained variance of the model, respectively. For period, one strong QTL on chromosome 1 can be observed that accounts for 40.7 % of the

variance in the population. This period QTL overlaps with the peak time QTL, as both loci show maximal association with the same marker: 69683843-1. For period and peak time there are several other LOD score peaks that could represent small effect QTL, although they don't pass the 5 % significance threshold as determined by 10,000 permutations (Figure 20).

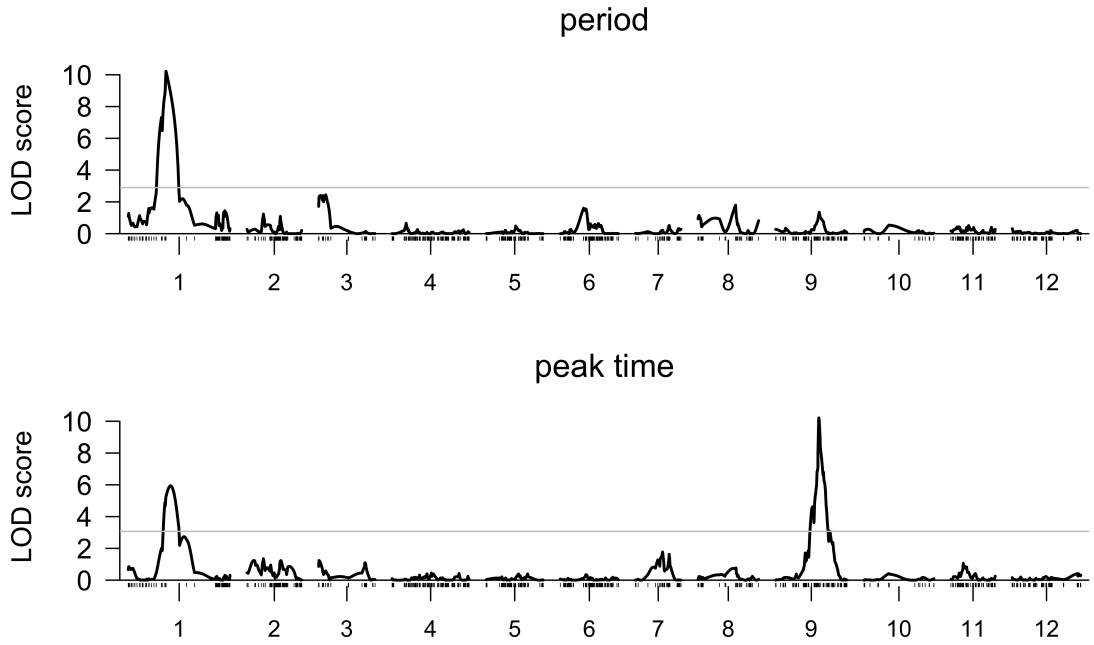

Figure 20: QTL for period and peak time in the *S. pimpinellifolium* RIL population. LOD scores were calculated with a single-QTL model using standard interval mapping with a 1 centi Morgan (cM) grid. The genetic map was estimated based on the genotypic data using the Kosambi map function. Numbers on the x-axis refer to the chromosomes and tick marks represent markers. Gray horizontal lines indicate the 5 % significance threshold as determined by 10,000 permutations.

When separating the population by the maximally linked markers of the two peak time QTL and looking at the peak time distribution of the groups, it becomes obvious that the bimodal distribution is indeed due to the strong effect of these two loci (Figure 21).

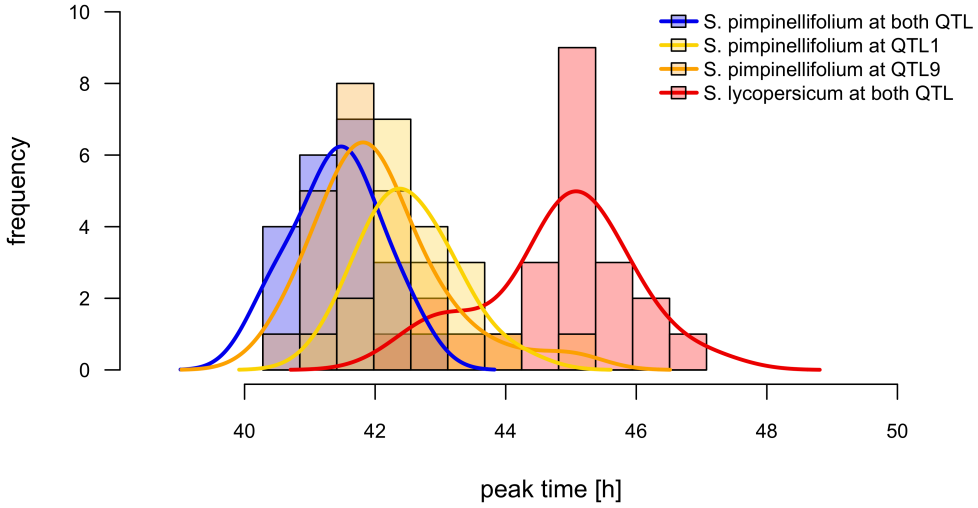

Figure 21: Bimodal peak time distribution due to two strong QTL. Histograms of the mean peak time estimates of 90 RILs (n = 3-9) separated by their genotype at the maximally linked markers for the two peak time QTL. Each RIL was grown in two of five independent experiments. Solid lines indicate the density.

4.3.2.2 Two-QTL interactions

One advantage of RIL populations as compared to IL populations is the possibility of analyzing the interaction between pairs of loci, due to the presence of more than just one genomic fragment of either parent in every line.

As already indicated in Figure 21, the additivity of the two peak time QTL on chromosomes 1 and 9 is confirmed by the two-dimensional two-QTL scan. The two loci together reach a LOD score of more than 20 in the two-QTL model (Figure 22). Also, a strong epistatic interaction between the QTL region on chromosome 1 and a region on chromosome 10, with a LOD score of around seven, can be observed. Additionally, there are three interactions reaching LOD scores higher than three. These are between two regions of chromosome 12, between chromosomes 3 and 10 and between chromosomes 3 and 11 (Figure 22).

Many weak epistatic interactions are present for period, but only one between chromosomes 4 and 12 reaches LOD scores higher than three and should therefore be considered significant (Figure 23).

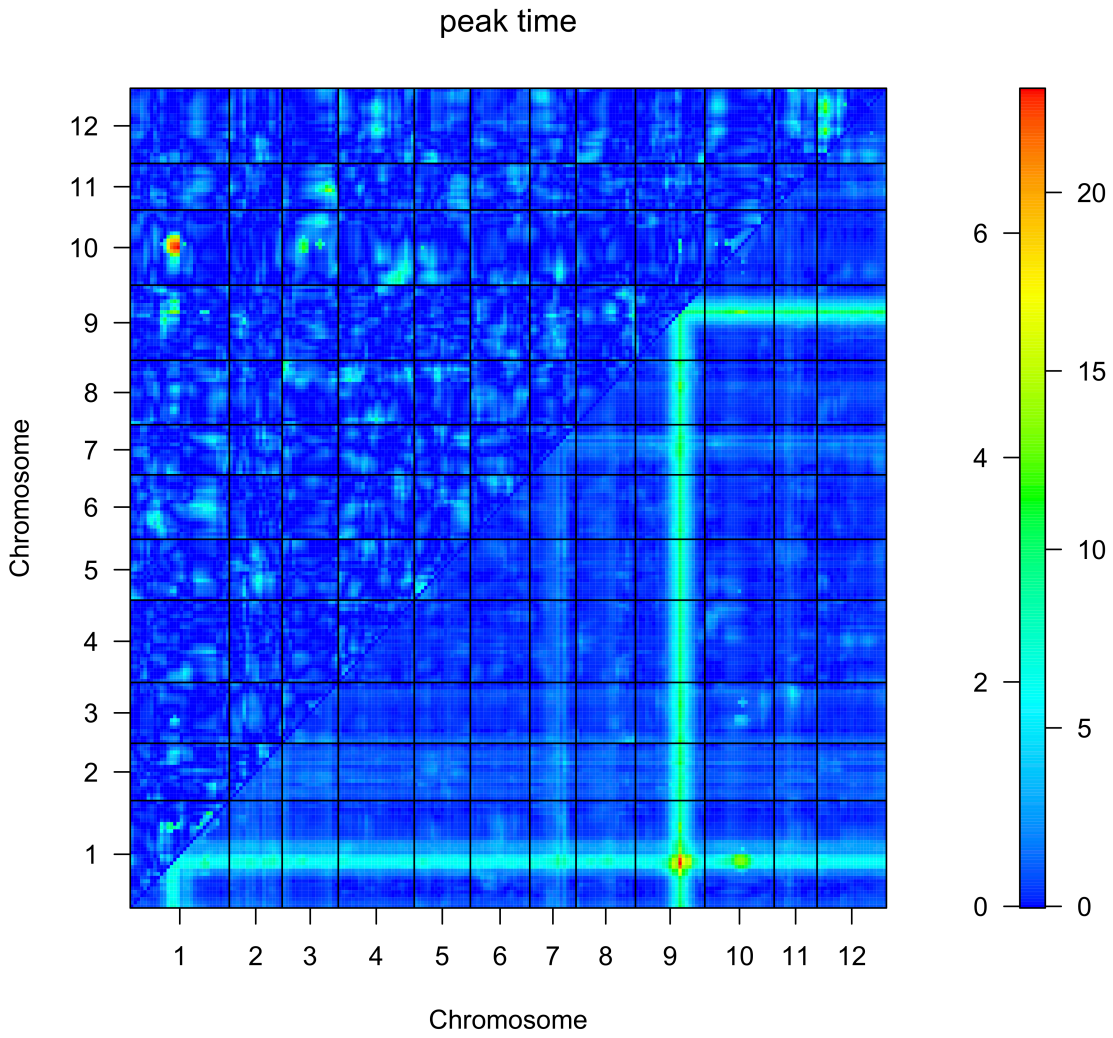

Figure 22: Two-dimensional scan showing QTL and epistatic interactions for peak time. LOD scores were calculated with a two-QTL model using standard interval mapping with a 5 cM grid. The LOD scores for the full two-QTL model (LOD_f) are displayed in the lower right triangle, the LOD scores for epistatic interactions (LOD_i) are displayed in the upper left triangle. Numbers on the left of the color scale correspond to LOD_i, numbers on the right to LOD_f.

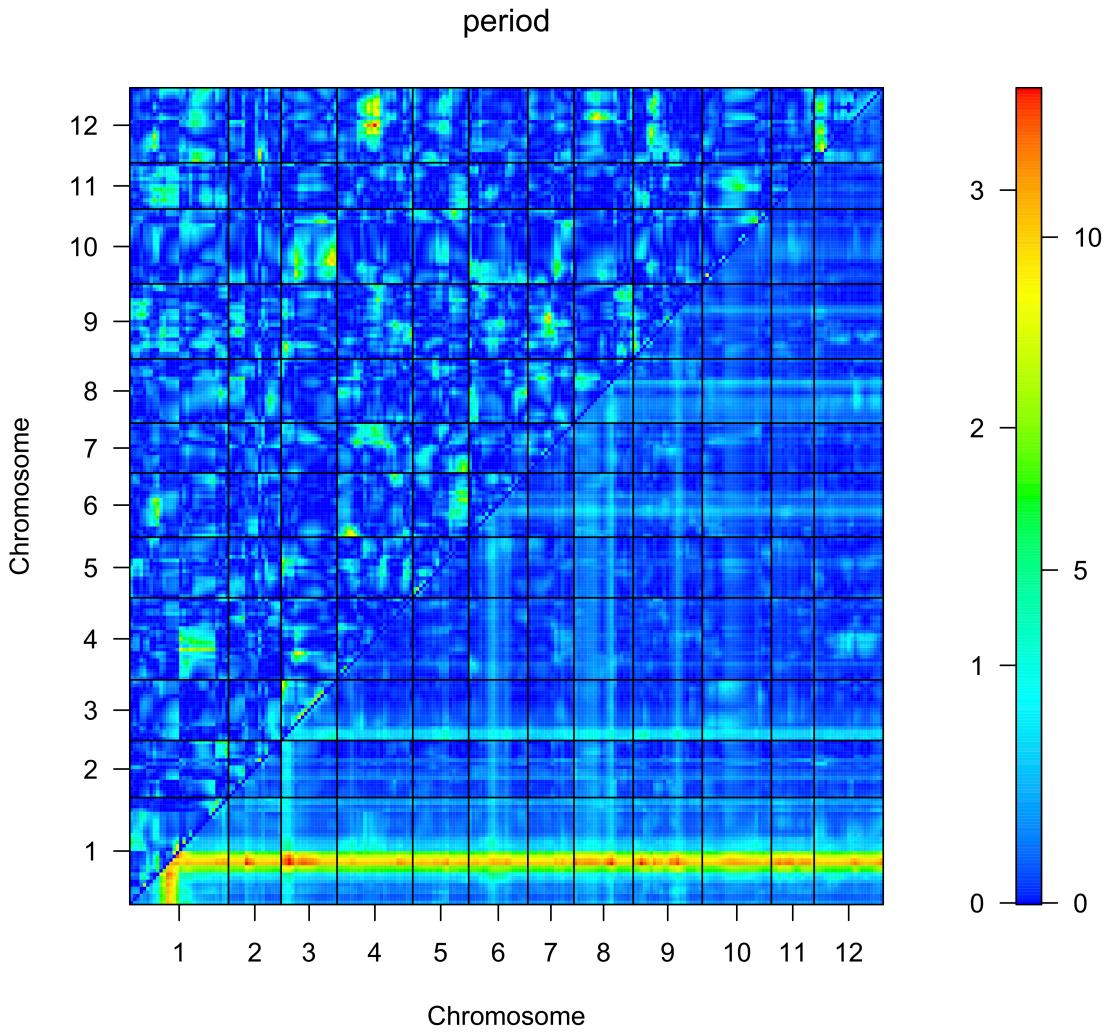

Figure 23: Two-dimensional scan showing QTL and epistatic interactions for period. LOD scores were calculated with a two-QTL model using standard interval mapping with a 5 cM grid. The LOD scores for the full two-QTL model (LOD_f) are displayed in the lower right triangle, the LOD scores for epistatic interactions (LOD_i) are displayed in the upper left triangle. Numbers on the left of the color scale correspond to LOD_i, numbers on the right to LOD_f.

4.4 QTL confirmation and characterization

QTL mapping for circadian leaf movements in two different populations revealed several loci possibly involved in period and peak time differences between wild and cultivated tomato. Five putative period QTL in the *S. pennellii* population and one in the *S. pimpinellifolium* population were identified. Additionally, five putative peak time QTL were detected with the *S. pennellii* ILs and two with the *S. pimpinellifolium* RILs. Most interestingly, the QTL identified in the *S. pimpinellifolium* population are present in both populations as will be shown in more detail later.

4.4.1 QTL only present in the *S. pennellii* IL population

The 11 lines of the *S. pennellii* population that showed significant differences in circadian rhythms compared to M82 (linear regression, $p < 0.01$, Table 2) were grown again in two independent experiments to confirm their phenotypes. Another five lines containing introgressions overlapping with introgressions of these 11 lines were also included in the experiments, as they can help to delimit the chromosomal regions containing the causal gene or genes (Figure 24).

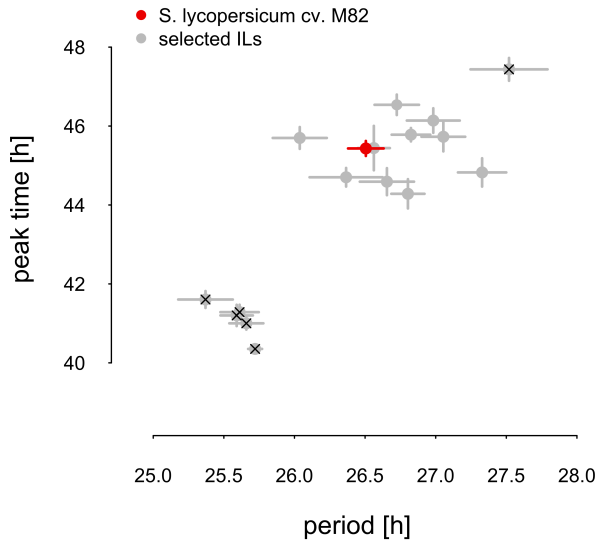

Figure 24: Period and peak time estimates of 16 selected ILs and *S. lycopersicum* cv. M82. Shown are the mean period and peak time estimates ± SEM of 16 selected introgression lines in gray and of M82 in red each from two independent experiments (n = 12-16). Lines that were significantly different from M82 in both experiments (linear regression, $p < 0.05$) are indicated by a black cross.

Of the eleven lines that showed significant differences in the first experiment, six maintained significant differences compared to M82 in both confirmation experiments (linear regression, $p < 0.05$). These lines define and confirm the presence of at least three QTL in the *S. pennellii* population. IL1-1-3 has a longer period and later peak time than M82. IL7-4-1, IL9-2, IL9-2-6, IL9-3 and IL9-3-1 have shorter periods and earlier peak times than M82 (Figure 24). The QTL on chromosomes 1 and 7, defined by IL1-1-3 and IL7-4-1, are specific to the *S. pennellii* population.

4.4.1.1 Period and peak time QTL on chromosome 1

Only one of the introgression lines that had a significantly longer period than M82 in the initial screen could be confirmed in both of the confirmation experiments. This locus therefore represents the only transgressive period QTL found in the population. Leaf movement traces for the confirmation experiment as well as mean period estimates for lines M82 and IL1-1-3 are presented in Figure 25 (A+B).

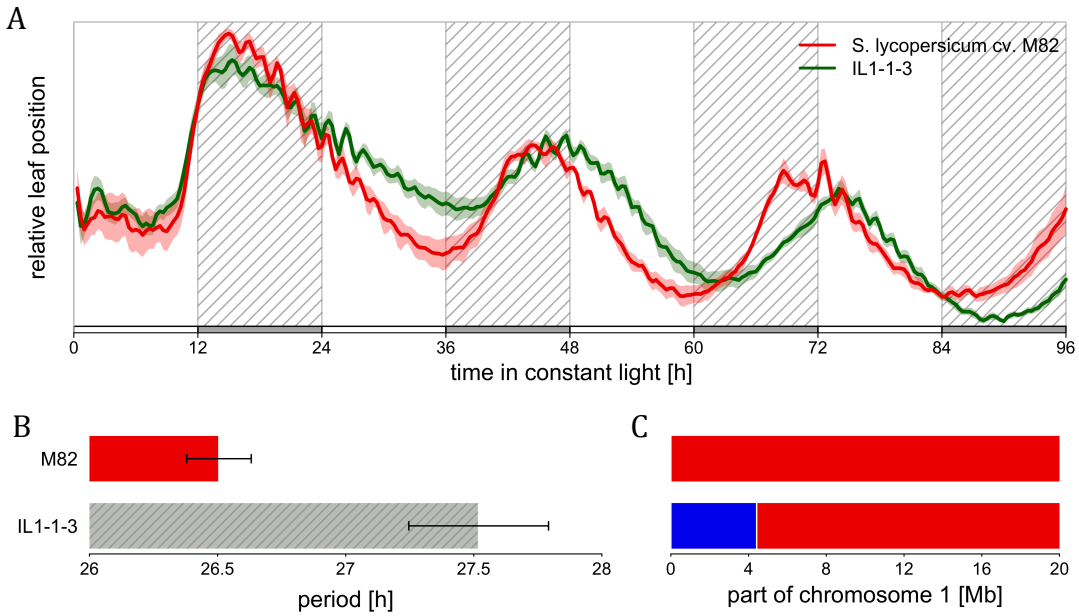

Figure 25: Transgressive QTL on chromosome 1 in the *S. pennellii* population. (A) Mean relative position of cotyledon tip of IL1-1-3 and *S. lycopersicum* M82 from one representative experiment (shaded area: ± SEM) (n = 7 + 6). Hatched areas indicate subjective nights. (B) Mean period estimates of M82 and IL1-1-3 ± SEM (n = 13 + 16) each from two independent experiments. (C) Genotype representation of part of chromosome 1: M82 = red, *S. pennellii* = blue.

IL1-1-3 contains a 4.4 Mega base pairs (Mb) introgression of the top of chromosome 1 (Figure 25C). There are 487 annotated genes in that region, many of which are possible candidate genes based on their functional annotation. Also, most genes have polymorphisms between *S. pennellii* and cultivated tomato that could be responsible for the differences in circadian phenotypes observed. Nevertheless, in this QTL region there is one very obvious candidate gene: a Flavin-binding Kelch-domain F-box protein (Solyc01g005300.2.1), which is the ortholog of the *Arabidopsis thaliana* clock gene *FKF1*. In Arabidopsis the *fkf1* mutant shows longer period than the wild type (Baudry et al., 2010). Additionally, this gene has four amino acid changes between *S. pennellii* and *S. lycopersicum* cv. M82 that could be altering the function of the alleles. Still, as the QTL region contains many other genes, it would be necessary to reduce the size of the introgression by backcrossing IL1-1-3 to M82 and to screen recombinant lines for circadian rhythms to strengthen this gene as a candidate for the QTL.

4.4.1.2 Period, peak time and amplitude QTL on chromosome 7

IL7-4-1 showed a much earlier peak time than M82 in the initial screen. In both confirmation experiments, this phenotype was confirmed and an additional short period phenotype was found. Further, this IL was the only line of the population with a significantly higher amplitude than M82 (linear regression, $p < 0.01$). This amplitude phenotype is transgressive and was again observed in both confirmation experiments (Figure 26).

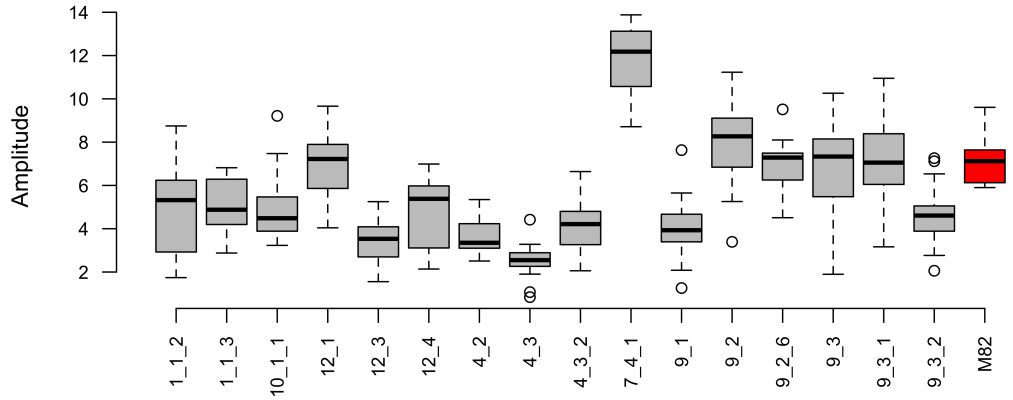

Figure 26: Amplitude estimates of 16 selected ILs and *S. lycopersicum* cv. M82. Boxplot of the amplitude estimates of 16 selected introgression lines in gray and of M82 in red (n = 12-16). Data are from two independent experiments

IL7-4-1 carries an introgression of roughly 57 Mb, representing almost the entire chromosome 7 except for the rightmost 9 Mb. There are 1440 annotated genes in that region. About 400 of those can be excluded because they are shared by an overlapping line, IL7-5, that does not show a peak time, period or amplitude phenotype (Figure 27).

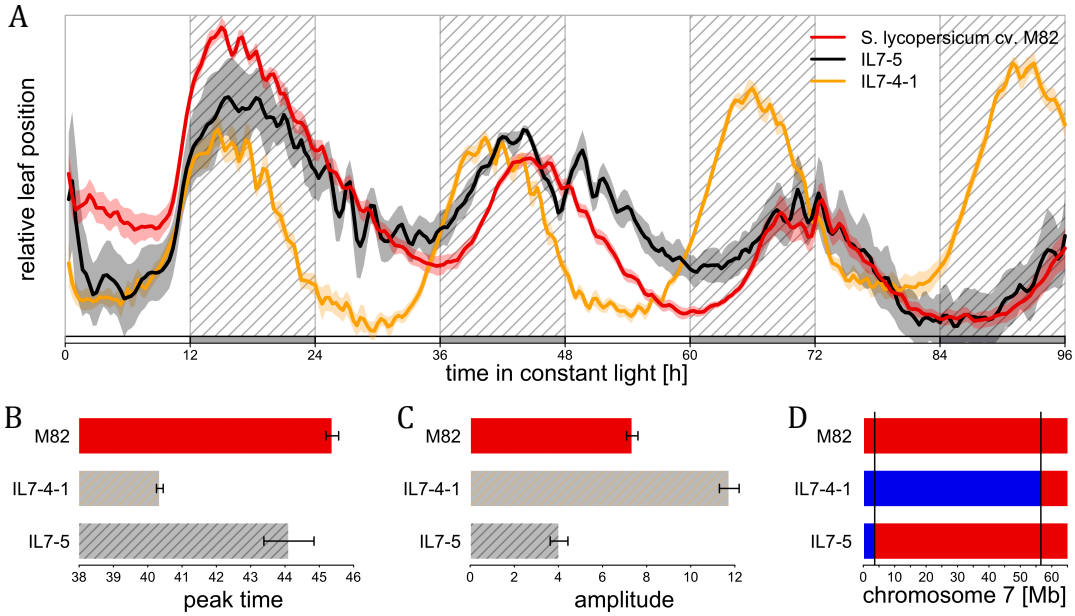

Figure 27: Peak time and amplitude QTL on chromosome 7 in the *S. pennellii* population. (A) Mean relative position of cotyledon tip of the QTL defining ILs and *S. lycopersicum* cv. M82 from one representative experiment each (shaded area: ± SEM, n = 5-10). Hatched areas indicate subjective nights. Mean peak time (B) and amplitude (C) estimates of M82, IL7-4-1 and IL7-5 ± SEM (n = 19, 15 + 5). IL7-4-1 was grown in two, IL7-5 in one experiment, M82 in all three. (D) Genotype representation of chromosome 7: M82 = red, *S. pennellii* = blue. Vertical lines indicate the QTL.

As for the long period QTL on chromosome 1, there is one very obvious candidate gene in the QTL region based on the annotation: the homolog of the *Arabidopsis thaliana* clock gene *ZEITLUPE* (*ZTL*) annotated as Flavin-binding kelch repeat F-box protein (Solyc07g017750.2.1). Interestingly, *ZTL* is the paralog of *FKF1* discussed as a candidate for the long period QTL on chromosome 1. Again, in order to substantiate this gene as a candidate, the QTL should be narrowed down by screening the F2 population of a M82 x IL7-4-1 backcross for recombinants and phenotyping them.

It is worth mentioning that the *S. pennellii* allele shows a high degree of dominance for this QTL (Figure 28). Therefore, transformation of M82 with the wild species allele of the causative gene should significantly alter the circadian phenotype. The transgenics are expected to be similar to the F1 hybrids (M82xIL7-4-1).

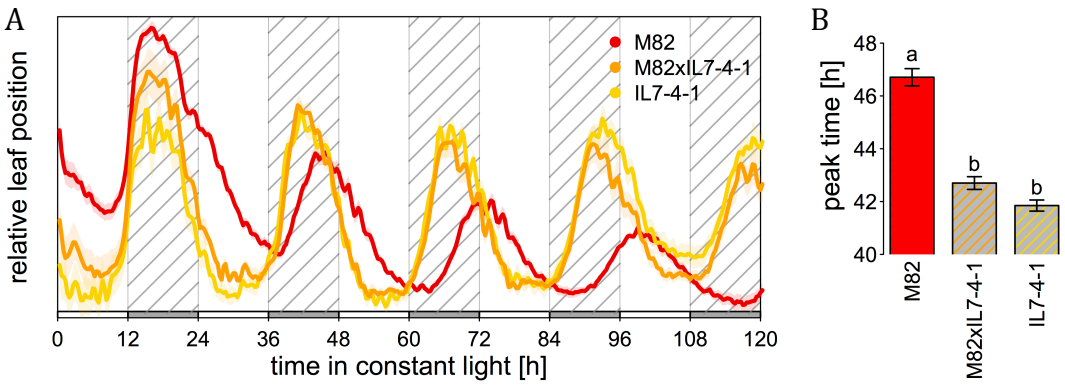

Figure 28: The wild species allele of QTL7 shows a high degree of dominance. (A) Mean relative position of cotyledon tip of IL7-4-1, *S. lycopersicum* cv. M82 and the respective F1 hybrid; data are from one representative experiment (shaded area: ± SEM, n = 9). Hatched areas indicate subjective nights. (B) Mean peak time estimates of the same plants ± SEM (n = 9). Different letters indicate significantly different means (Tukey's HSD, $p < 0.05$).

4.4.2 QTL present in both populations

Clock genes that were changed during tomato domestication or breeding would underlie QTL that differentiate any wild species from any cultivar. Since the main goal of this thesis was to understand possible effects of human selection on the circadian clock, these are the most interesting loci. Both QTL identified in the *S. pimpinellifolium* RIL population (QTL1 and QTL9) also differentiate *S. pennellii* from *S. lycopersicum* cv. M82. Therefore, both loci could potentially be responsible for the general differences in circadian rhythms observed between wild and cultivated tomato.

4.4.2.1 Period and peak time QTL on chromosome 1 (QTL1)

The period QTL on chromosome 1 identified in the *S. pimpinellifolium* population was originally not found in the *S. pennellii* population. However, a closer look at the *S. pennellii* ILs showed that the part of chromosome 1 containing the QTL is only covered by IL1-1, one of the lines that had not been phenotyped because no seeds were available. Therefore, a population of around 500 backcross inbred lines (BILs) was employed. These lines were generated in the lab of Dani Zamir at the Hebrew University of Jerusalem from a cross between the cultivated tomato variety *S. lycopersicum* cv. M82 and the wild tomato species *S. pennellii* (LA0716) just as the introgression line population. Four BILs were identified that have introgressions containing at least part of the QTL on chromosome 1. Phenotyping of those lines for circadian leaf movements revealed that two lines,

BIL497 and BIL484, exhibit significantly shorter periods than M82 (t-test, $p < 0.05$) (Figure 29A). These two lines have a shared segment of *S. pennellii* ranging from 69.7 to 74.5 Mb on chromosome 1 (Figure 29B). Notably, unlike most other BILs, BIL497 has only one introgression and thus can be considered a near isogenic line confirming the QTL (it had a significantly shorter period (t-test, $p < 0.05$) in two independent experiments). Additionally, it can be used to fine-map the QTL by backcrossing it to M82.

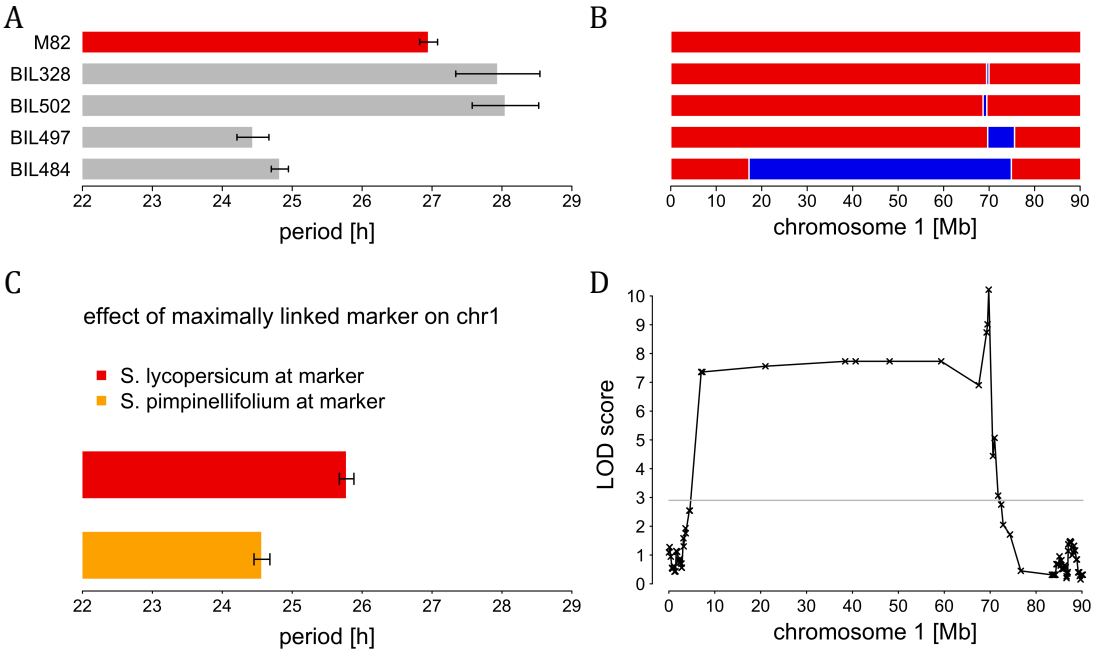

Figure 29: Across population period QTL. (A and B) *S. pennellii* BILs; (C and D) *S. pimpinellifolium* RIL population. (A) Mean period estimates of selected BILs and *S. lycopersicum* cv. M82 ± SEM, BIL328, BIL502 and BIL484 (n = 2 - 4) were grown in one experiment, M82 and BIL497 (n = 12 - 13) in two. (B) Genotype representation of chromosome 1 of the lines shown in A: M82 = red, *S. pennellii* = blue. (C) Allelic mean effect ± SEM of the maximally linked marker on chromosome 1 (69683843-1). (D) LOD scores of a one-QTL model for peak time calculated by marker regression. The gray horizontal line indicates the 5% significance threshold as determined by 10,000 permutations.

When combining the genotype information of BIL497 and the QTL confidence interval from the *S. pimpinellifolium* RIL population, the QTL has a size of approximately 900 kilo base pairs (kb) containing 78 annotated genes. Many of those genes are possible candidates, making it difficult to speculate on the molecular basis for this QTL. Momentarily F1 hybrids from backcrosses of BIL497 with *S. lycopersicum* cv. M82 are growing. These are needed to generate an F2-population that can be screened for recombinants.

Interestingly, the wild species allele for QTL1 is dominant as is shown in Figure 30.

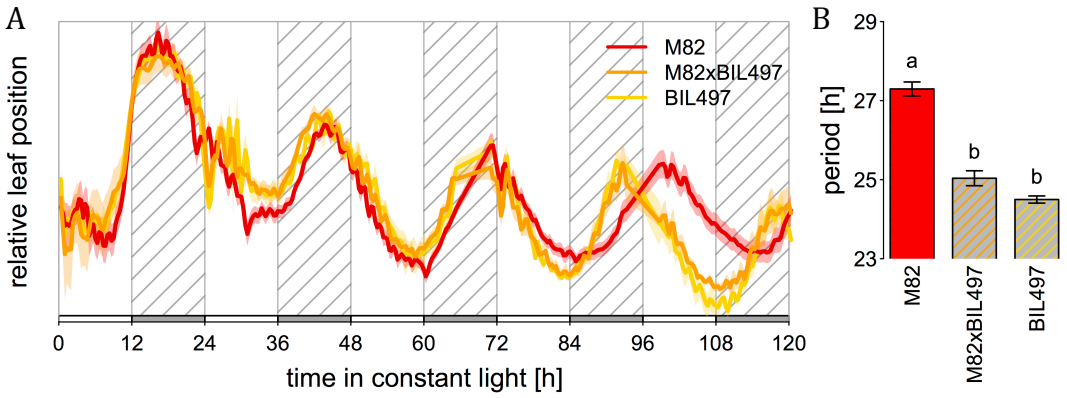

Figure 30: The wild species allele of QTL1 shows a high degree of dominance. (A) Mean relative position of cotyledon tip of BIL497, *S. lycopersicum* cv. M82 and the respective F1 hybrid; data are from one representative experiment (shaded area: ± SEM, n = 5-13). Times 66 to 71 and 94 to 99 are connected by straight lines to remove artifactual drops (probably caused by warm air blown into the chamber to maintain 25 °C). Hatched areas indicate subjective nights. (B) Mean period estimates of the same plants ± SEM (n = 5-13). Different letters indicate significantly different means (Tukey's HSD, $p < 0.05$).

It is likely that the same gene underlies the QTL in both populations. Especially when considering that this is the only period QTL in the *S. pimpinellifolium* population and that it is the period QTL with the strongest effect in the *S. pennellii* population. The *S. pennellii* allele of this QTL shortens the period by almost three hours (see BIL497 vs. M82 in Figure 29A). This effect is considerably stronger than that of the other two period QTL on chromosomes 7 and 9 that shorten the period by about one hour.

4.4.2.2 Period and peak time QTL on chromosome 9 (QTL9)

Four of the five ILs with earlier peak times and shorter periods have an introgression of chromosome 9: IL9-2, IL9-2-6, IL9-3 and IL9-3-1. Genotyping of these four ILs revealed that IL9-3-1 actually has the same genotype as IL9-3. The early peak time and the symmetric waveform of the four lines are associated with a genomic region of about 900 kb containing 126 annotated genes (Figure 31). This suggests that allelic variation for one of those genes is responsible for the observed differences in circadian rhythms.

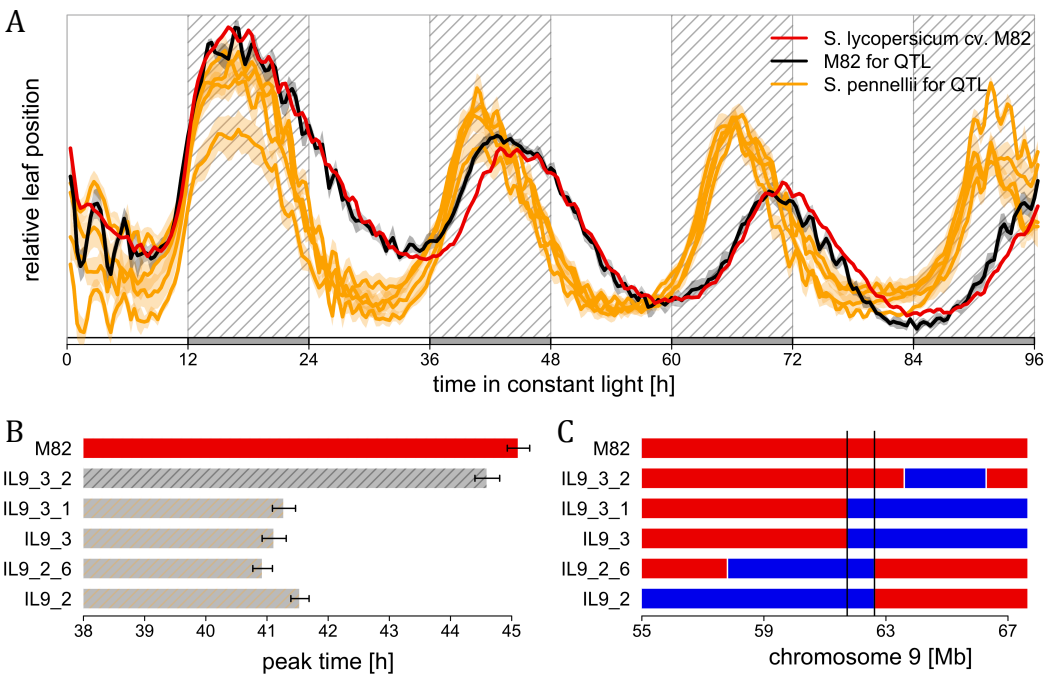

Figure 31: Peak time QTL on chromosome 9. (A) Mean relative position of cotyledon tip of the QTL defining ILs and *S. lycopersicum* M82 from one representative experiments (shaded area: ± SEM) (n = 7-8). Hatched areas indicate subjective nights. (B) Mean peak time estimates of the same lines as in A ± SEM (n = 13-22). Each line was grown in two independent experiments. (C) Genotype representation of part of chromosome 9: M82 = red, *S. pennellii* = blue. Vertical lines indicate the QTL region.

Most interestingly, the peak time QTL on chromosome 9 identified in the *S. pimpinellifolium* RIL population perfectly overlaps the peak time QTL from the *S. pennellii* IL population (Figure 32B and D). Additionally, the cultivated species allele increases peak time in both populations (Figure 32A and C). This supports the hypothesis that the causal genetic change underlying this QTL has occurred only once - in cultivated tomato, possibly already during domestication.

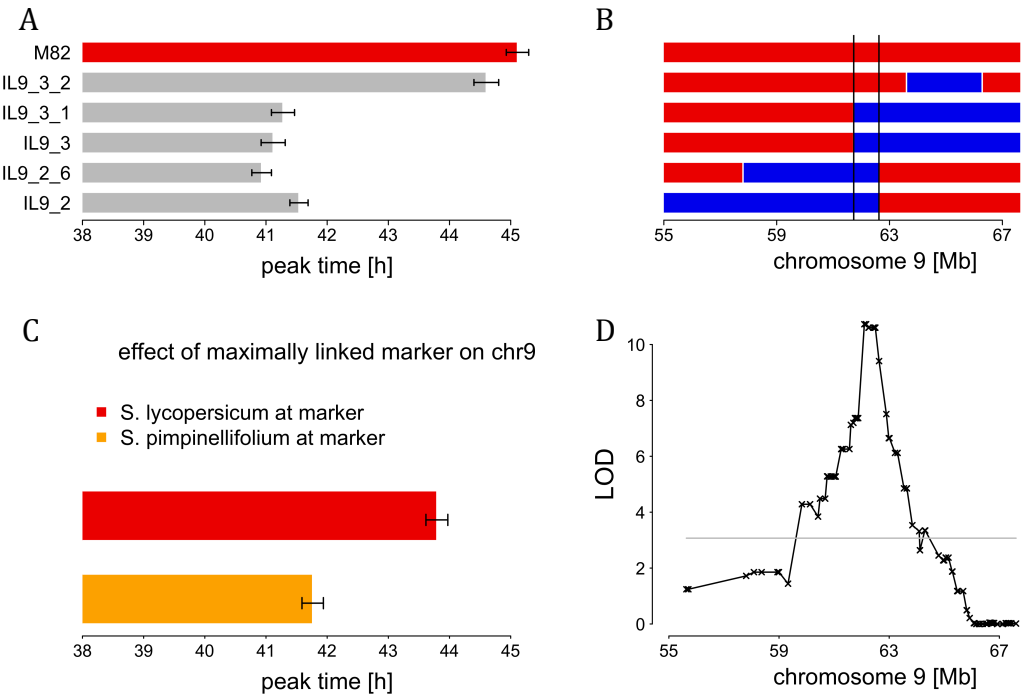

Figure 32: Across population peak time QTL. (A and B) *S. pennellii* IL population; (C and D) *S. pimpinellifolium* RIL population. (A) Mean peak time estimates of the QTL defining ILs and *S. lycopersicum* cv. M82 ± SEM (n = 13-22). Each line was grown in two independent experiments. (B) Genotype representation of part of chromosome 9 of the lines shown in A: M82 = red, *S. pennellii* = blue. Vertical lines indicate the putative QTL. (C) Allelic mean effect ± SEM of the maximally linked marker on chromosome 9 (62161978-9). (D) LOD scores of a one-QTL model for peak time calculated by marker regression. The gray horizontal line indicates the 5% significance threshold as determined by 10,000 permutations.

4.4.3 QTL also present in *S. habrochaites* and *S. galapagense* populations

If the cultivated alleles of the two QTL shared among populations originated during domestication or early breeding, these QTL would be found in crosses of any wild species with any cultivated tomato. To test this, first an already existing *S. habrochaites* IL population covering about 85 % of the *S. habrochaites* genome was used (Monforte and Tanksley, 2000). Unfortunately the regions of QTL1 and QTL9 are missing. With this population it could therefore not be tested whether these loci also differentiate *S. habrochaites* from *S. lycopersicum*. It should be noted that the genotype information for this population available from TGRC (http://tgrc.ucdavis.edu/) is not correct. Of several lines that were genotyped only few had the genotype published by Monforte and Tanksley (2000), even though the seeds were obtained directly from TGRC.

Nevertheless, two lines of the population more closely resemble the wild parent *S. habrochaites*, exhibiting significantly earlier peak times (LA3931 and LA3950)

and a significantly shorter period (LA3950) (linear regression, $p < 0.01$) than the recurrent parent *S. lycopersicum* cv. E6206 (Figure 33).

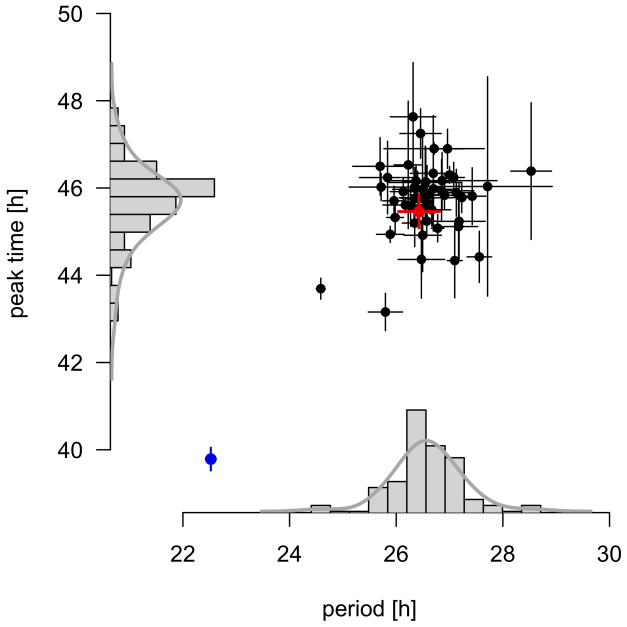

Figure 33: Variation in circadian leaf movements for an *S. habrochaites* introgression line population. Mean period and peak time estimates ± SEM of the 51 ILs with at least two biological replicates in black (n = 2-9) and the wild parent *S. habrochaites* in blue (n =9), each grown in one of four independent experiments, and the recurrent parent *S. lycopersicum* cv. TA209 in red (n = 17) grown in each of the four experiments. Histograms on the axes show the distributions of the means of the ILs; the gray lines indicate the density.

The IL with the short period has an introgression of almost the entire chromosome 7 except the leftmost few Mb. The phenotype of this line and IL7-4-1 from the *S. pennellii* population (discussed in section 4.4.1.2) could theoretically be caused by allelic variation of the same gene or genes. The QTL might be involved in the differences between the green-fruited and red-fruited tomato species.

To test whether QTL1 and QTL9 also differentiate other wild species from cultivated tomato, an F2 population from a cross between the cultivated tomato variety *S. lycopersicum* cv. Moneymaker (MM) and the wild species *S. galapagense* (LA0530) was employed. Circadian leaf movements of 96 individuals were phenotyped. These individuals were genotyped with two markers distinguishing the cultivar Moneymaker from *S. galapagense*. The first marker is a three bp indel closely linked to QTL9 in the *S. pimpinellifolium* and the *S. pennellii* populations and is located in the coding region of the homolog of the Arabidopsis gene

EMPFINDLICHER IM DUNKELROTEN LICHT 1 (*EID1*) (Solyc09g075080), which is the best candidate gene for the QTL on chromosome 9, as described below. This mutation might represent the causal nucleotide variant responsible for the QTL - the quantitative trait nucleotide (QTN). The second marker is a three bp indel as well and is close to the maximally linked marker of QTL1 in the *S. pimpinellifolium* population. It is located in the 3'UTR of the Crossover junction endonuclease MUS81 (Solyc01g068350).

As can be seen in Figure 34A, the peak time QTL on chromosome 9 also exists in the *S. galapagense* population. It shows a high degree of dominance. On the other hand, the marker on chromosome 1, which is supposed to be associated with QTL1, does not significantly differentiate the genotypes (Tukey's HSD, $p > 0.05$) (Figure 34B). Nevertheless, there is a trend of plants with the *S. galapagense* allele showing shorter periods than plants with the Moneymaker allele. It is possible that the marker used is not linked closely enough to the causal polymorphism to show a significant association.

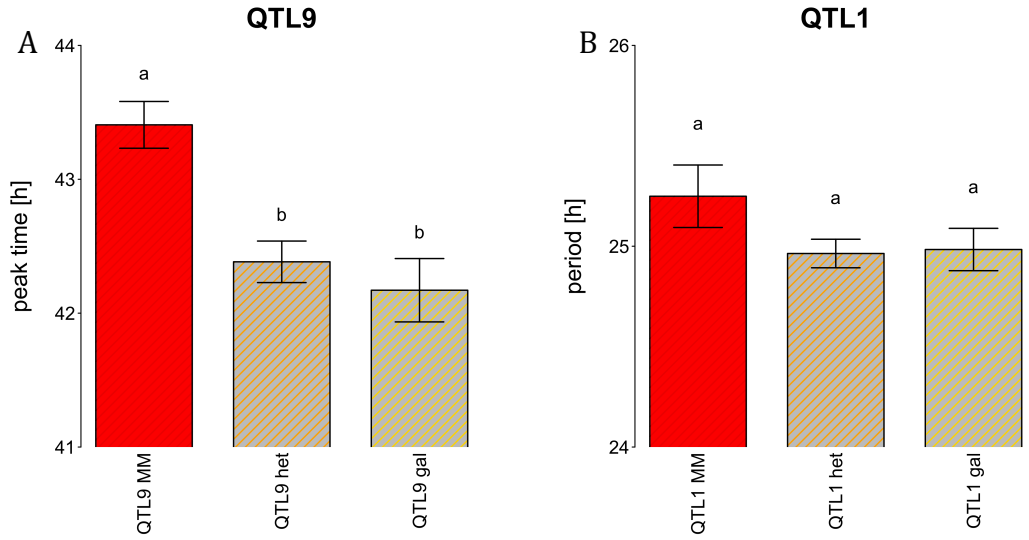

Figure 34: QTL9 and QTL1 in an *S. galapagense* F2-population. (A) Mean peak time estimates ± SEM (n = 15-37) of lines with differing genotypes for a marker closely linked to QTL9. (B) Mean period estimates ± SEM (n = 25-43) of lines with differing genotypes for a marker closely linked to QTL1. Data are from two independent experiments. Different letters indicate significantly different means (Tukey's HSD, $p < 0.05$). MM: *S. lycopersicum* cv. Moneymaker; gal: *S. galapagense*; het: heterozygous.

4.4.4 Conclusions

The presence of at least four QTL for circadian leaf movements was confirmed. One of them, specific to the *S. pennellii* population, is transgressive with the wild species allele lengthening the circadian period. Another QTL is shared by the two green-fruited wild species *S. pennellii* and *S. habrochaites*. Two loci, QTL1 and QTL9, are present in all populations tested (although QTL1 is not significant in the *S. galapagense* population) and could be responsible for the general differences between wild and cultivated tomato. The molecular change leading to the causative allelic variation most likely occurred only once during domestication or crop evolution.

Current and future efforts include fine mapping of the two putative domestication QTL found in this study – QTL1 and QTL9. Of these, QTL 9 has already been narrowed down to less than one megabase and will be detailed below.

4.5 Fine-mapping of QTL9

The *S. pennellii* BIL population (already described in 4.4.2.1) was employed to narrow down the region of QTL9. It is a great resource for fine-mapping QTL identified in the *S. pennellii* IL population. All BILs with an *S. pennellii* introgression overlapping any part of QTL9, as defined by the *S. pennellii* ILs, were selected. Additionally some BILs with introgressions on the left or right of the QTL region were chosen. In total 13 BILs were selected and assayed for circadian leaf movement rhythms. It must be mentioned that most of the BILs contain more than one introgression of *S. pennellii* (on average each line contains three introgressions). In spite of the heterogeneous genetic backgrounds of the BILs, every line carrying an *S. pennellii* introgression containing the rightmost part of QTL9 exhibits an early peak time (Figure 35A and B), a symmetric waveform (Figure 35C) and a short period (data not shown).

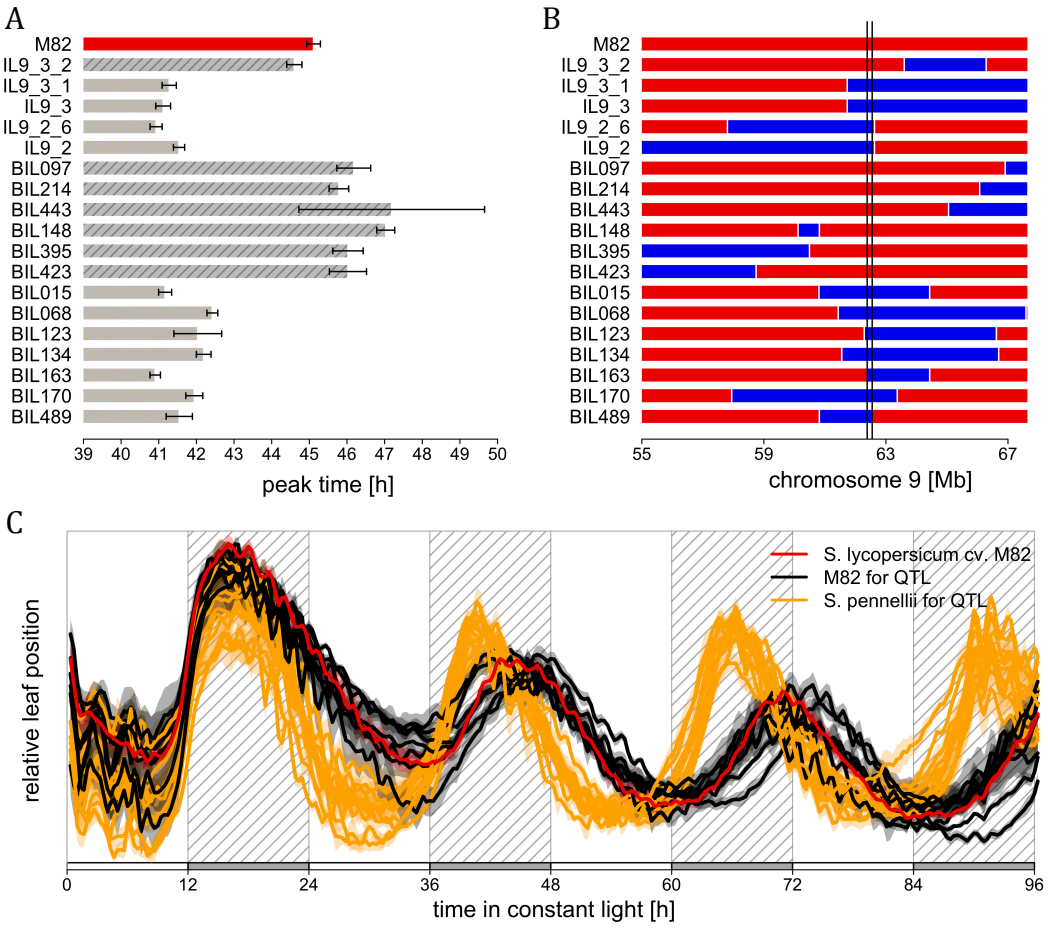

Figure 35: Fine-mapping the peak time QTL on chrmosome 9. (A) Mean peak time estimates of the QTL defining ILs and selected BILs ± SEM (n = 10-22). Each line was grown in two independent experiments. (B) Genotype representation of part of chromosome 9 of the lines shown in A: M82 = red, *S. pennellii* = blue. Vertical lines indicate the putative QTL region. (C) Mean relative position of cotyledon tip of the same lines from one representative experiment (shaded area: ± SEM) (n = 4-8). Hatched areas indicate subjective nights.

This analysis narrowed down the region of the QTL to about 200 kb containing 29 annotated genes. To narrow down the QTL even further, the F2 progeny of a cross between IL9-2-6 and *S. lycopersicum* cv. M82 was used. These plants were screened for recombinants with a crossover between two markers flanking the QTL. Of roughly 800 F2 plants tested, two exhibited crossovers in the QTL region. This means that the QTL region has a size of approximately 0.25 centimorgans (cM). For the first recombinant the crossover occurred in the middle of the QTL region at position 62.47 (rec47). It therefore loses 16 genes from the right side of the QTL. In the second recombinant the crossover took place at the upstream border of the QTL region, at position 62.38 (rec38). This recombinant loses the entire QTL region (Figure 36).

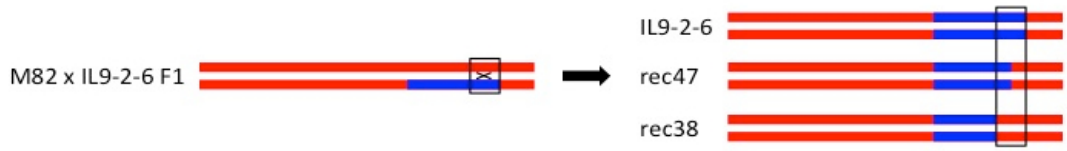

Figure 36: Recombinants from a M82 x IL9-2-6 backcross. Genotypes are represented schematically for chromosome 9 by horizontal bars. Each allele is shown by one bar. M82: red, *S. pennellii*: blue. The box indicates the region of the QTL, the cross symbolizes a cross-over event.

The recombinant rec47 and IL 9-2-6 exhibit the same circadian phenotype, while the recombinant rec38 has the same phenotype as *S. lycopersicum* cv. M82 (Figure 37A). Since the two recombinants only differ in a region of less than 100 kb containing 13 annotated genes, the QTL can be narrowed down to this region. Natural allelic variation for one of those 13 genes must be responsible for the differences in circadian rhythms.

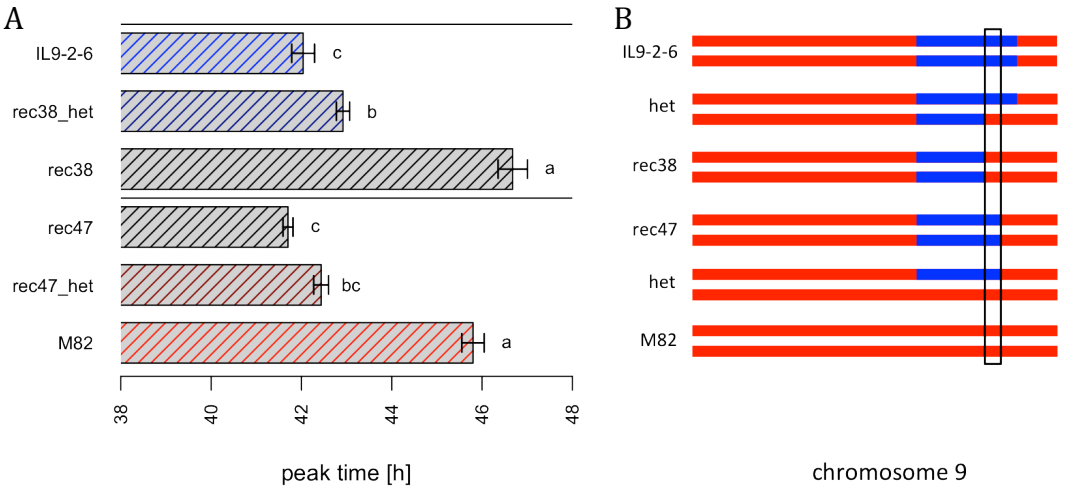

Figure 37: Peak time estimates of the recombinants, parental lines and heterozygotes. (A) Mean peak time estimates of the recombinants in black (middle bars), the parental lines from the backcross, M82 in red and IL9-2-6 in blue, and the heterozygotes ± SEM (n = 10-24). Different letters indicate significantly different means (Tukey's HSD, $p < 0.05$). (B) Genotypes are represented schematically for chromosome 9 by horizontal bars. Each allele is shown by one bar. M82: red, *S. pennellii*: blue. The box indicates the region of the QTL.

Natural alleles underlying quantitative variation are seldom absolutely dominant or recessive. The degree of dominance (d/a) for an allele can reach values between -1 and 1 (Tanksley, 1993). The *S. pennellii* allele shows high degrees of dominance (Figure 37): for "rec38_het" d/a is 0.62; for "rec47_het" it is 0.64. This

is important for choosing the right background for transformation when cloning the QTL. The heterozygous plants show that a single copy of the wild species allele has a significant effect in an otherwise M82 background ($p < 0.001$). For this reason M82 and not IL9-2-6 was later transformed with the different alleles of the candidate gene.

4.6 Candidate genes and cloning of QTL9

QTL9-associated variation in circadian rhythms can be caused either by a DNA sequence polymorphism or by an epigenetic modification. Epigenetic modifications can lead to differences in expression. Also SNPs can affect expression by changing the rate of transcription or mRNA stability. Non-synonymous sequence polymorphisms can lead to changes in protein structure and function. Silent SNPs could change translation efficiency or lead to differential alternative splicing. RNA sequencing data can be used to analyze most of these possibilities: polymorphisms in coding regions, gene expression and alternative splicing. It can thereby help to identify possible candidate genes underlying a QTL. In addition, knowledge about the function of genes in other species, especially in the model plant *Arabidopsis thaliana*, can be helpful for choosing candidate genes for cloning.

4.6.1 Candidate genes based on information from other species

Based on the functional annotation, two genes seem to be very good candidates for QTL9. One is a phytochrome A-associated F-box protein described in Arabidopsis to be a negatively acting component of the phytochrome A signaling pathway called EMPFINDLICHER IM DUNKELROTEN LICHT 1 (EID1) (Buche et al., 2000). The *EID1* mutant (*eid1-3*) seems to have an effect on circadian expression of the *CAB:LUC* reporter (Wenden et al., 2011). In addition, the two closest paralogs of *EID1* in tomato are annotated as circadian clock coupling factors based on a study performed in tobacco (Xu and Johnson, 2001). The other candidate gene is annotated as cryptochrome DASH family protein (*Cry-DASH*). Cryptochromes are core components of the circadian clock in mammals, birds and insects (Bell-Pedersen et al., 2005), and play a role in circadian and diurnal rhythms in Arabidopsis and tomato (Facella et al., 2008; Somers et al., 1998). For this reason, *Cry-DASH* also seems to be a good candidate gene for QTL9.

4.6.2 Candidate genes based on polymorphisms, expression and alternative splicing

Since for QTL9 all wild species show the same differences when compared with cultivated tomato, most probably the same polymorphism, epigenetic modification or alternative splicing event is responsible for this differentiation. Therefore RNAseq data from the cultivars Moneymaker (unpublished data) and M82 and the wild species *S. pimpinellifolium*, *S. pennellii* and *S. galapagense* (Koenig et al., 2013) were screened for polymorphisms differentiating wild from cultivated tomato. Of the 13 genes in the QTL9 region there are only five genes with polymorphisms that fulfill this requirement (Table 3).

Table 3: Candidate genes for QTL9 based on polymorphism data. Polymorphism data were generated from RNA sequencing data as described in Methods (section 3.5.8) and in Koenig, Jimenez-Gomez et al. 2013. nsSNP = non-synonymous SNP, sSNP = synonymous SNP

gene-ID	shared polymorphisms	gene description
Solyc09g075070	5 sSNPs and 3 nsSNP	Beta-glucosidase
Solyc09g075080	3bp insertion and 1 nsSNP	Phytochrome A-associated F-box protein
Solyc09g075150	1 sSNP	60S ribosomal protein L22-2
Solyc09g075170	6 sSNPs and 7 nsSNPs	Pentatricopeptide repeat protein
Solyc09g075190	6 sSNPs	Gamma-tubulin complex component 4

The same RNAseq data were used to analyze differences in expression and alternative splicing differentiating the wild species from cultivated tomato. For this analysis only the samples from *S. pennellii*, *S. pimpinellifolium* and *S. lycopersicum* cv. M82 were compared because they were grown in the same experiment. Of the 13 genes in the QTL interval, none is differentially expressed between the two wild species and *S. lycopersicum* cv. M82. Also none of the 13 genes is differentially alternatively spliced in the wild species compared to the cultivar M82 (Arunkumar Srinivasan, unpublished data).

Combining all the information, the Phytochrome A-associated F-box protein *EID1* (Solyc09g075080) is the most likely candidate gene for QTL9. Either the three bp insertion in the wild tomato species leading to a one amino acid insertion (-269K) or the non-synonymous SNP leading to an amino acid change (F180L) could modify the function of the protein and cause the observed differences in circadian rhythms. Interestingly, the wild alleles at these two polymorphisms are shared with the homologs of *EID1* in potato (PGSC0003DMT400029665, *Solanum tuberosum* group Phureja (Xu et al., 2011)) and pepper (Capang03g003624,

Capsicum annuum cv. Zunla-1 (Qin et al., 2014)). Therefore the allele of cultivated tomato represents a derived state while the wild tomato allele is the ancestral one.

Noteworthy, the tomato gene encoding the Phytochrome A-associated F-box protein unambiguously is the ortholog of the Arabidopsis *EID1* gene, as microsynteny exists between the two species – the upstream gene is an argine/serine-rich splicing factor in both species. Additionally, the protein similarity is 60 %, whereas the similarity between tomato *EID1* and the three Arabidopsis *EID1*-like (*EDL*) genes is only 20 to 30 %.

4.6.3 Cloning the Phytochrome A-associated F-box protein *EID1*

To test whether *EID1* really is the causal gene underlying QTL9, first a heterologous complementation approach was used. The *S. pennellii* and *S. lycopersicum* cDNAs each under the control of both native promoters were cloned into the binary vector pFAST-R01 (Shimada et al., 2010). Those four constructs and the empty vector were transformed into two different T-DNA insertion lines of *Arabidopsis thaliana* (Alonso et al., 2003). Both lines are in the Col-0 background and have a single homozygous T-DNA insertion in the only exon of *EID1*. The T-DNA insertion of the line SALK_021058C (= SALK-1) is at position 152 of the exon. Line SALK_027403C (= SALK-2) contains an insertion at position 947. Circadian leaf movements of the two SALK lines and Col-0 revealed a significant effect (Tukey's HSD, $p < 0.05$) of the T-DNA insertion of the line SALK-2 on period but not on peak time (Figure 38). In contrast, the line SALK-1 did not show any differences compared to the wild type Col-0. This was observed in two independent experiments.

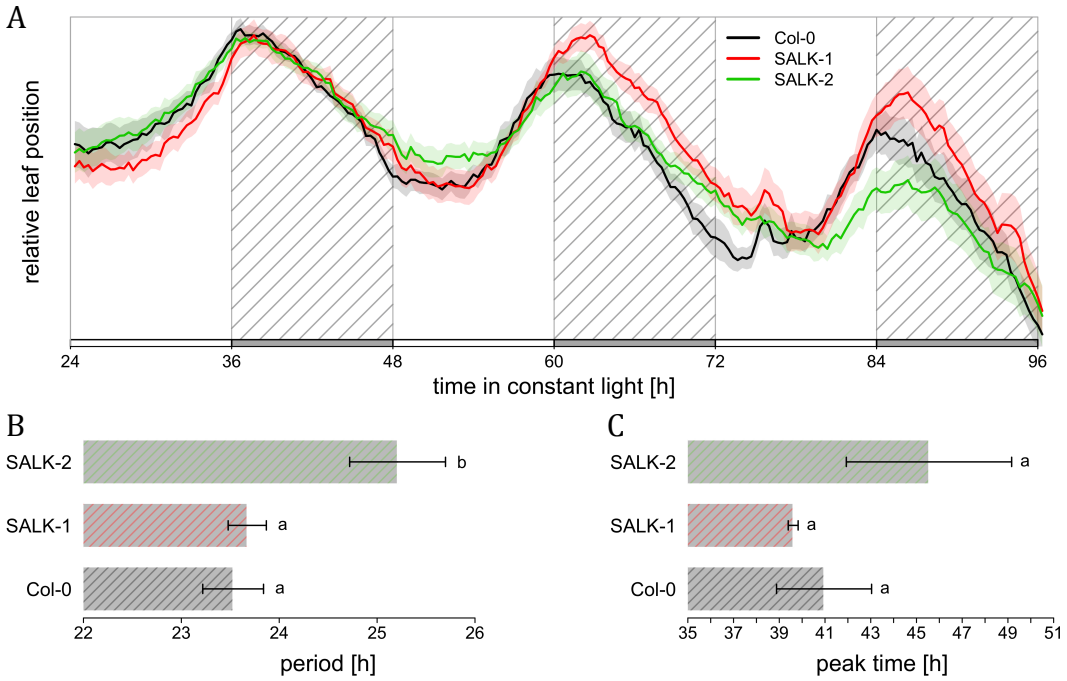

Figure 38: Circadian rhythms of Arabidopsis insertion mutants for *EID1*. (A) Mean relative position of cotyledon tip of two T-DNA insertion lines and Col-0 (shaded area: ± SEM) (n = 12). Hatched areas indicate subjective nights. Mean period (B) and peak time (C) estimates of the same lines as in A ± SEM (n = 11-12) Data are from two independent experiments. Different letters indicate significantly different means (Tukey's HSD, $p < 0.05$).

The phenotypic differences between the two insertion mutants could be due to the different T-DNA insertion sites. The coding sequence of *EID1* has two in frame start codons. In line SALK-1 the second start codon could theoretically be used to initiate translation. This would give rise to a protein representing the right quarter of the full-length protein. If this part of the protein affects circadian rhythms, that would explain the observed differences. In line SALK-2 the second start codon cannot be used because the T-DNA insertion is downstream (Figure 39). Interestingly, the importance of the C-terminus was suggested in a study using different *eid1* mutant alleles (Dieterle et al., 2001).

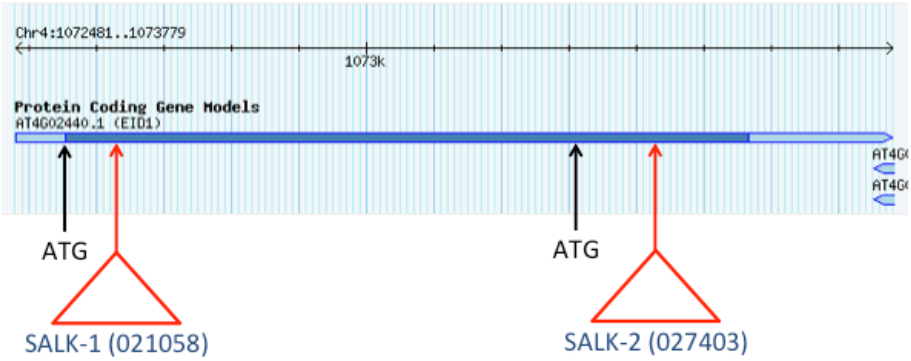

Figure 39: T-DNA insertion mutants of *EID1* in *Arabidopsis thaliana* Col-0. Schematic representation of the exon of *EID1* with the UTR colored in light blue and the coding sequence in dark blue (modified from www.arabidopsis.org). Red triangles and arrows indicate the site of the T-DNA insertion. The positions of the two start codons (ATG) are shown with black arrows.

Unfortunately it could not be tested whether the two tomato alleles have a different effect on circadian rhythms in the Arabidopsis insertion mutants of *EID1* because the transformation efficiency of the constructs was low, except for the empty vector control. It was not possible to obtain transgenic plants with both tomato alleles in either of the two SALK lines.

The same four constructs described above (both alleles driven by both promoters) were generated using the pGWB1 vector (Nakagawa et al., 2007) and transformed into the tomato cultivar *S. lycopersicum* cv. M82. At this moment, stable transgenic lines are being developed and will be used for phenotyping circadian rhythms and possible pleiotropic effects of the alleles. Due to time limitations, characterization of these lines is out of the scope of this thesis.

4.7 Pleiotropic effects of QTL9

The differences in circadian rhythms between cultivated tomato and the wild species could be a consequence of a founder effect and genetic drift caused by a domestication bottleneck. In this case, it is possible that the genes underlying the circadian differences are neutral and there would have been no need to reverse the phenotype by breeding. Another possibility is that the altered circadian rhythms are beneficial for tomato cultivation. In this case, the alleles causing differences in circadian rhythms could have been artificially selected by humans during domestication or breeding, leading to the patterns of variation present today – the differentiation of wild from cultivated tomato. If this hypothesis is

true, differences in circadian rhythms should lead to improved agricultural traits. In other words, the differences in circadian rhythms should affect traits that are visible under natural conditions.

To test this, the two near isogenic lines (described in section 4.5 and outlined in Figure 36) were grown under greenhouse conditions typical for tomato cultivation. The two NILs only differ for the 13 genes of QTL9. After three weeks of seedling growth differences in seedling height were observed and indeed these differences are significant (t-test, $p = 0.0064$). Individuals harboring the cultivated allele for QTL9 are shorter than those with the wild species allele (Figure 40).

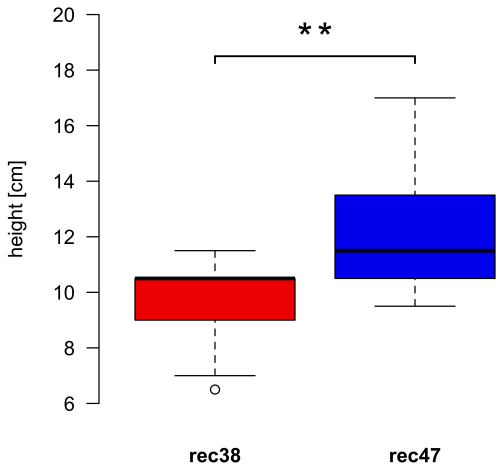

Figure 40: Two near isogenic lines for QTL9 exhibit growth differences. Boxplots of seedling heights (excluding the hypocotyl) of two near isogenic lines (n = 12 each); rec47 contains the QTL9 region from *S. pennellii*, rec38 from *S. lycopersicum* cv. M82. Asterisks indicate significant differences of mean (t-test, $p < 0.01$).

The results suggest that altered circadian rhythms conferred by allelic variation at QTL9 do have an impact on seedling growth under diurnal conditions. This means that there is visible variation under natural conditions possibly due to differences in the circadian clock. Thus, it appears possible that the slower circadian rhythms observed in cultivated tomato were actually selected for by humans during domestication or crop evolution.

4.8 Natural variation in photoperiodic flowering – short day flowering in the close tomato relative *S. galapagense*

A key seasonal output of the circadian clock in plants is photoperiodic flowering. Photoperiod flowering is crucial for many plants as it allows proper timing of reproductive development. Cultivated tomato, however, is considered a day neutral plant not changing flowering time due to differences in photoperiod (Samach and Lotan, 2007). Nevertheless, some of its wild relatives are short day plants that accelerate flowering under short days (Prat 2011, unpublished data). The genes responsible for this naturally occurring variation in photoperiodic flowering in tomato have not yet been identified.

In this thesis, flowering was monitored in a subset of the domesticated and wild tomato accessions scored for circadian rhythms (see section 4.2). It was observed that *S. galapagense* plants (LA0530) did not produce any flowers after more than one year under standard greenhouse conditions (artificial lighting was provided in the morning and evening to have 16 hours of light every day). When plants derived from cuttings of the non-flowering individuals were transferred to a short days (8 hours light, 16 hours dark) environmental chamber they produced abundant flowers within a month. This indicates that *S. galapagense* is a short day plant. This flowering behavior appears to be obligate as *S. galapagense* does not produce any flowers under long day greenhouse conditions.

4.8.1 Grafting experiments

Previous studies in several plant species have demonstrated that a systemic signal, called florigen, is produced in the leaves under inductive photoperiods and moves through the phloem to the shoot apical meristem where it induces flowering. This signal seems to be mainly encoded by the Arabidopsis *FLOWERING LOCUS T* (*FT*) gene and its homologs in other species. (Corbesier et al., 2007; Lifschitz et al., 2006; Navarro et al., 2011; Tamaki et al., 2007). To test whether the inhibition of flowering in *S. galapagense* observed under greenhouse conditions is due to the lack of such a systemic inductive signal, a grafting experiment was performed. *S. galapagense* scions were grafted onto *S. lycopersicum* cv. Moneymaker rootstocks and onto *S. galapagense* rootstocks as a negative control (Figure 41). The grafting was done in a way to retain four to five leaves on the

rootstock. Grafted plants were kept under greenhouse conditions in which *S. galapagense* plants normally never flower. The grafted shoots, however, formed first flower buds after about one month and started to flower after two months when the rootstock was *S. lycopersicum* (n = 9), whereas the shoots did not even form flower buds when the rootstock was *S. galapagense* (n = 6). This indicates that the lack of a mobile inducing signal is indeed responsible for the inhibition of flowering in *S. galapagense* under greenhouse conditions. It is likely that this signal is encoded by *SINGLE FLOWER TRUSS* (*SFT*), the homolog of Arabidopsis *FT* (Lifschitz et al., 2006).

Figure 41: Grafting of *S. galapagense* onto *S. lycopersicum* induces flowering under non-inductive long days. The left plant shows a *S. galapagense* scion grafted onto a *S. lycopersicum* rootstock; the right plant shows the negative control where *S. galapagense* is grafted onto *S. galapagense*. The white circles indicate the graft junctions. (Picture taken by Maret-Linda Kalda, photographer of the MPIPZ)

4.8.2 Flowering time and flower number under long day, short day and night break conditions

To quantify the flowering behavior of *S galapagense* (LA0530), plants were grown in climate chambers set to long day, short day or night break conditions. Night break conditions are the same as short days except that 30 minutes light are provided in the middle of the night. This treatment ensures that flowering time differences are due to photoperiod sensitivity and not due to other factors such as unequal light quantity or stress caused by the longer light exposure under long days. Figure 42 shows *S. galapagense* plants after two and a half months under night break or short day conditions, demonstrating that half an hour of light in the middle of the dark period has a strong inhibitory effect on flowering.

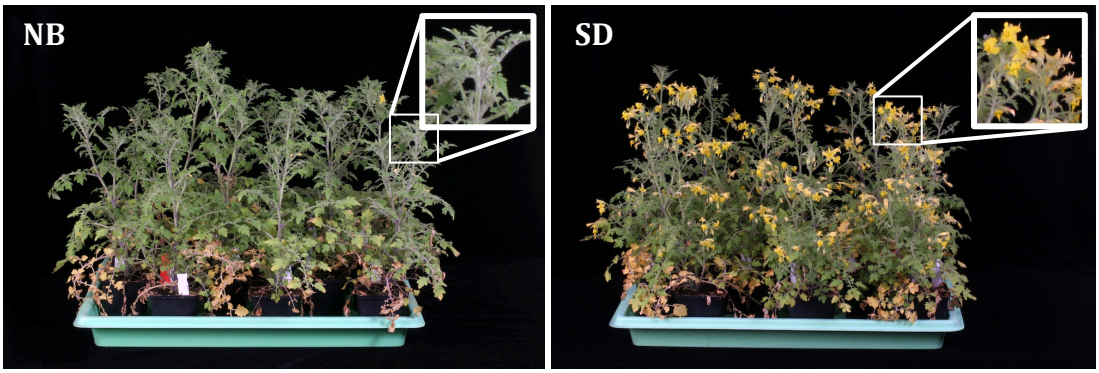

Figure 42: *S. galapagense* plants grown under night break or short day conditions. NB: 8h light / 16h dark + 30 min light in the middle of the dark period. SD: 8h light / 16h dark. For each condition 12 representative plants 75 days after germination are shown. The inserts in the right top corners show close-ups of the indicated regions. (Pictures taken by Maret-Linda Kalda, photographer of the MPIPZ)

As in the greenhouse, *S. galapagense* never flowered under long day conditions in the growth chambers used for this experiment. Six of seventeen plants (35%) did not flower under night break conditions either. The ones that did flower, on average flowered about ten days later than the plants grown under short day conditions (Figure 43A).

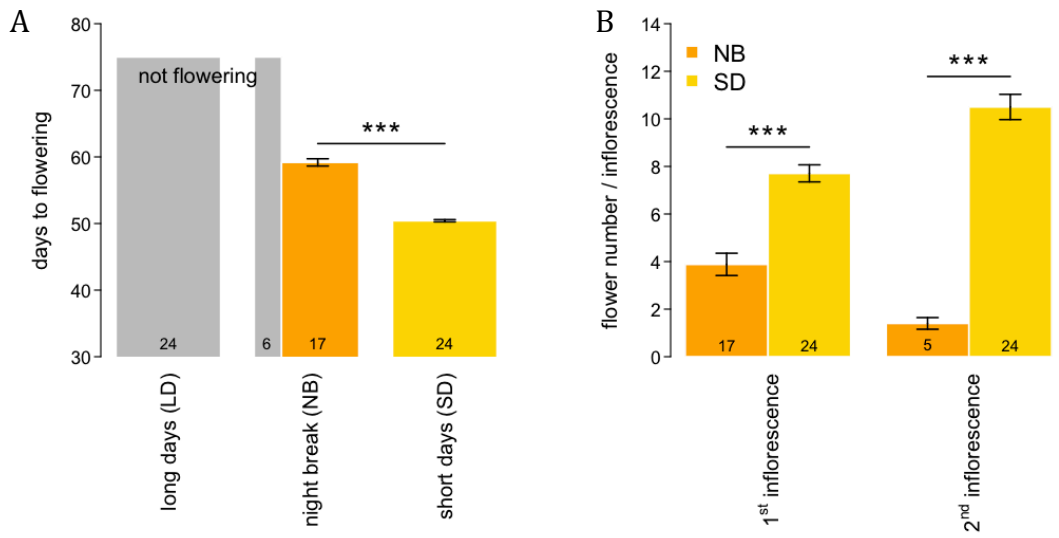

Figure 43: Flowering time and flower number of *S. galapagense* under different light conditions. (A) Days to flowering under LD, NB and SD conditions; data are mean ± SEM (n = 23-24). (B) Flower number per inflorescence of plants under NB and SD conditions counted two weeks after opening of the first flower of the inflorescence; data are mean ± SEM (n = 5-24). Asterisks indicate significant differences of means (t-test, $p < 0.001$). Numbers within the bars specify the number of plants.

There is also a significant effect (t-test, $p < 0.01$) of night breaks on the number of flowers per inflorescence. Under night break conditions the plants produced fewer flowers than under short days. This effect is especially strong for the second inflorescence (Figure 43B). Additionally, in two thirds of the plants (12/17) only the first inflorescence flowered under night break conditions, whereas under short day conditions every plant generated further flowering inflorescences. In conclusion, these data demonstrate that *S. galapagense* LA0530 is a short day plant. Night breaks delay flowering time and decrease flower number per inflorescence and the number of flowering inflorescences per plant. As a matter of fact, long days completely inhibit flowering under certain growth chamber conditions and in the greenhouse. Thus, the short day flowering behavior of *S. galapagense* is obligate under certain conditions.

4.8.3 Bulk segregant analysis in an S. lycopersicum x S. galapagense F2 population using RNA sequencing

To map the region(s) responsible for the photoperiod response of *S. galapagense* (LA0530), bulk segregant analysis was performed using an *S. lycopersicum* cv. Moneymaker x *S. galapagense* F2 population.

The F1 hybrids obtained from the parental lines flowered under long day conditions in the greenhouse, where *S. galapagense* never flowered. Thus, the inhibition of flowering under long day conditions appears to be recessive. Additionally, the grafting experiment demonstrated that the lack of a mobile signal, likely *SFT*, is involved in the inhibition. Taken together, a recessive lack of florigen seems to be responsible for the photoperiodic flowering behavior of *S. galapagense.* This could be due to a protein change of SFT leading to short specific activation. However, the most straightforward explanation would be short day specific activation of expression due to a change in a cis regulatory element in the *S. galapagense SFT* promoter

To test this hypothesis 480 plants of a Moneymaker x *S. galapagense* F2 population were grown under long day conditions (in climate chambers with high light intensities, ~250 µmol m^{-2} s^{-1}). Tissue was collected from the first 96 plants that flowered and from the 90 plants that did not flower at all or only produced a maximum of three flowers per inflorescence. This cutoff was taken because in the climate chambers used for this experiment about 40 % of *S. galapagense* plants flowered under long day conditions. However, those flowering plants produced a maximum of three flowers per inflorescence as compared to at least six under short day conditions. Tissue from each group was pooled and sequenced using RNAseq. RNAseq instead of DNAseq was chosen because less reads are needed to obtain sufficient coverage and because it could allow the detection of changes in expression between pools. The reads were aligned to the tomato genome and exonic SNPs were identified. For these SNPs the allele frequencies in each pool were calculated.

If only one Mendelian recessive gene (supposedly *SFT*) is responsible for the inhibition of flowering under long day conditions, the frequency of the wild species allele should be 100 % in the nonflowering pool and 33 % in the flowering one.

This is not the case, as there is no region in the genome with a wild species allele frequency of 100 % (see green lines in Figure 44).

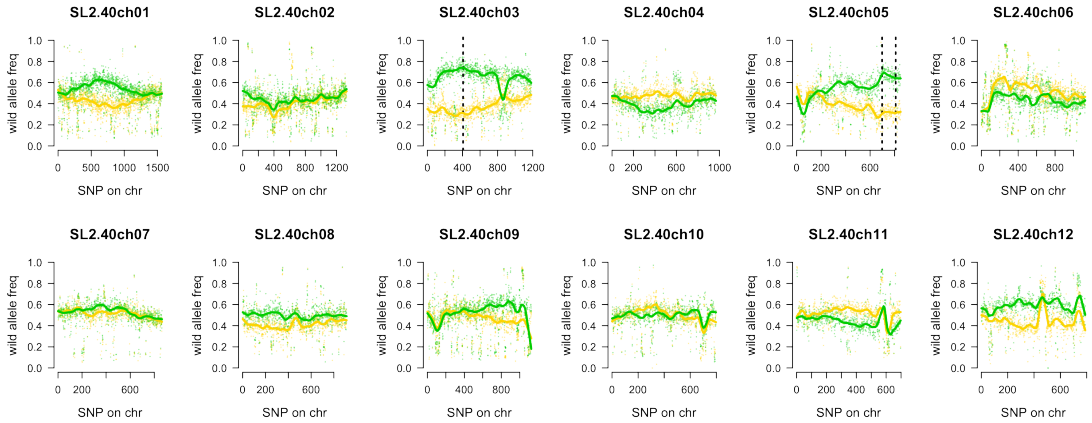

Figure 44: Wild species allele frequency in the two pools for the 12 chromosomes. Points in the graph represent allele frequencies of the wild allele for every SNP with a coverage >100 in both pools. Lines show a locally weighted smoothing regression (LOWESS) of the points in each chromosome. The pool formed by plants flowering under long day conditions is shown in yellow, the pool composed of nonflowering plants is shown in green. Dashed vertical lines indicate the SNP closest to the a priori candidate gene *SFT* (Solyc03g063100) on chromosome 3 and its two close paralogs *SP5G* (Solyc05g053850) and *SP6A* (Solyc05g055660) on chromosome 5.

Nevertheless, two regions in the genome were identified - one on chromosome 3 and one on chromosome 5 - that exhibit allele frequencies that would fit a scenario in which two recessive loci inhibit flowering under long days. In the case of a two genes model, the wild species allele frequency should be 71.4 % at those loci in the nonflowering pool and 33 % in the flowering one. The frequency plot of SNPs in 150 SNP windows showing these allele frequencies generates two peaks, one on chromosome 3 and one on chromosome 5. In these two regions almost half of the SNPs have wild species allele frequencies of 71 ± 5 % in the nonflowering pool and 33 ± 5 % in the flowering one (Figure 45).

These positions coincide with *SFT* (*SP3D*, Solyc03g063100) on chromosome 3, which was the a priori candidate gene, and its two close paralogs, *SP5G* (Solyc05g053850) and *SP6A* (Solyc05g055660), located on chromosome 5 (see dashed vertical lines in Figure 44 and Figure 45).

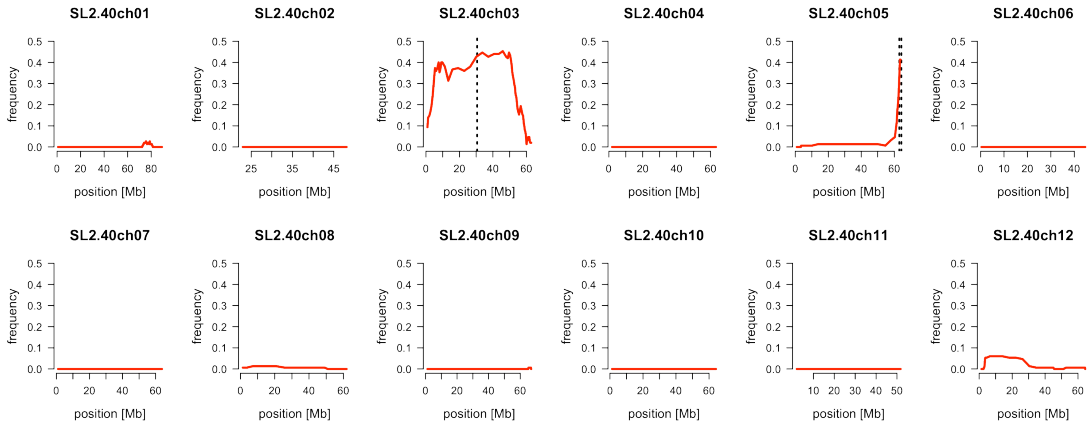

Figure 45: Frequency of SNPs with allele frequencies expected for two recessive genes. Shown is the frequency of SNPs in 150 SNP windows with steps of 15 SNPs that have a wild species allele frequency of 71 ± 5 % in the nonflowering pool and of 33 ± 5 % in the flowering one. Dashed vertical lines indicate the positions of the a priori candidate gene *SFT* (Solyc03g063100) on chromosome 3 and its two close paralogs *SP5G* (Solyc05g053850) and *SP6A* (Solyc05g055660) on chromosome 5.

In conclusion, the results demonstrate that more than one gene is involved in the photoperiodic flowering behavior of *S. galapagense.* Data are provided indicating that two recessive genes could be regulating the inhibition of flowering under long day conditions. The broad peak on chromosome 3 (Figure 45) suggests that the gene underlying this peak is in the centromeric region with a low recombination rate, as is the case for *SFT*. The narrow peak on chromosome 5 is expected for a gene in a region with high recombination rates such as *SP5G* and *SP6A*.

4.9 Photoperiodic flowering in *S. habrochaites*

Another line that showed an interesting flowering time phenotype is *S. habrochaites* (LA1777). Flowering was strongly inhibited under long day conditions in this accession (Figure 46).

Figure 46: *S. habrochaites* plans grown under long day and short day conditions. Plants were grown for three months either under long day (LD) or short day (SD) conditions at light intensities of about 250 µmol m^{-2} s^{-1}. LD: 16h light / 8h dark, SD: 8h light / 16h dark.

Also, flowering of *S. habrochaites* (LA1777) is delayed under night break conditions as compared to short days (Prat, personal communication).

4.9.1 QTL analysis in an *S. habrochaites* IL population

To identify the chromosomal regions responsible for the response to photoperiod, an introgression line population generated from a cross between cultivated tomato and *S. habrochaites* (already discussed in section 4.4.3) was employed (Monforte and Tanksley, 2000). Fifty-six lines and the recurrent parent *S. lycopersicum* cv. E6206 were grown in eight replicates under short day and under night break conditions. Days to flowering of the first inflorescence were scored. Some individuals are affected by senescing inflorescences. This refers to inflorescences of which all buds turn yellow and fall off before flowering (Figure 49A). This characteristic is mainly displayed by first inflorescences and was also recorded.

On average, all the introgression lines showed delayed flowering in the night break experiment. Also the recurrent parent *S. lycopersicum* flowered later under

these conditions. Furthermore, the range of flowering time was larger under night break conditions compared to short days, as can be seen from the wider distribution of the mean days to flowering of the ILs (Figure 47).

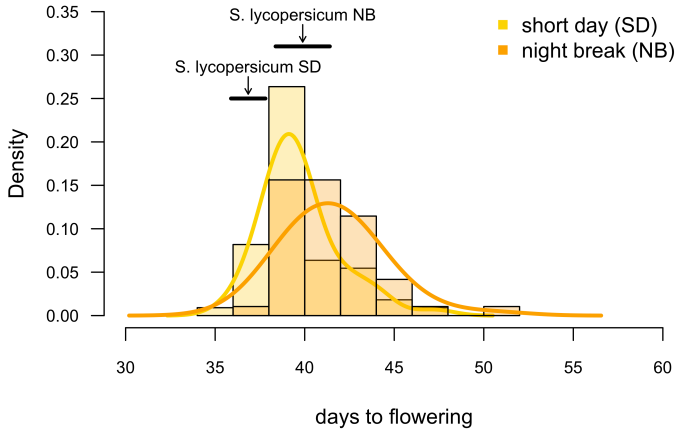

Figure 47: Distribution of days to flowering of the introgression lines under two conditions. Histograms of mean days to flowering of the 48 ILs with at least one measurement in both experiments (n = 1-8) under short day conditions and under night break conditions. Solid colored lines show the density. Arrows indicate the mean values for the recurrent parent *S. lycopersicum* under both conditions; the horizontal lines underneath represent the 99% confidence interval.

To identify lines that differ in their flowering time response to night breaks compared to the recurrent parent of the population *S. lycopersicum* cv. E6206, the relative responses of days to flowering of all the ILs with at least two biological replicates under both conditions were analyzed (Figure 48). One line (LA3968) exhibited a significantly stronger response to night breaks than *S. lycopersicum* (linear regression, $p < 0.05$). An additional line (LA3943) also showed a stronger response to the night break conditions than *S. lycopersicum*, although the difference is not significant on the 5 %-level (linear regression, $p = 0.121$).

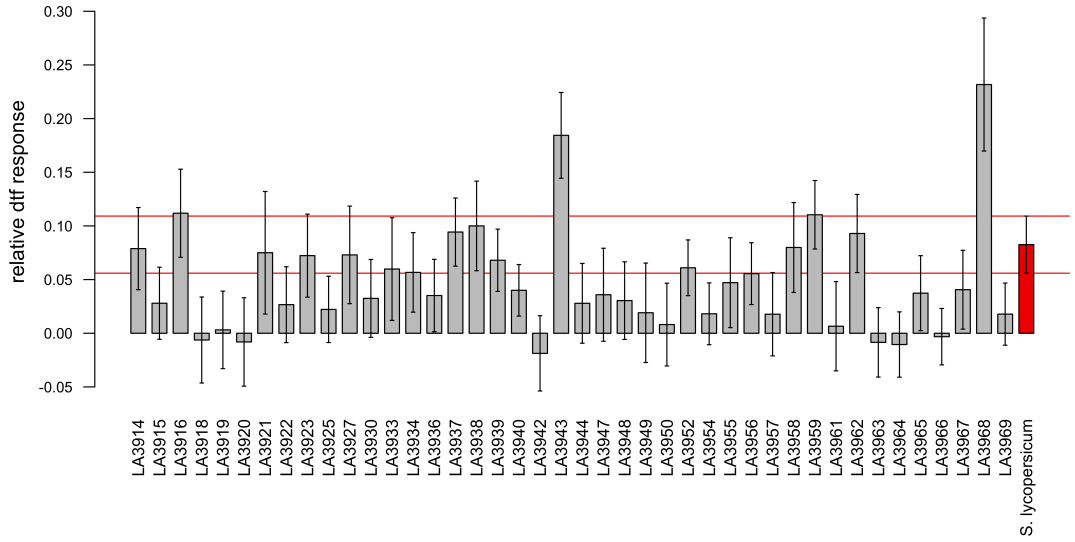

Figure 48: Relative response of days to flowering to night breaks. The relative response of days to flowering to night breaks is shown for all ILs with at least two measurements per condition ± SEM. The recurrent parent *S. lycopersicum* is shown in red, horizontal red lines indicate the range of two standard errors.

Under night break conditions, several lines including *S. lycopersicum* were affected by senescing inflorescences (Figure 49A). Six lines exhibited a significantly higher senescence ratio (ratio of plants with senescing inflorescences to plants with normal inflorescences) under night break conditions compared to short days (fisher test, $p < 0.05$). In two of those lines (LA3941 and LA3924) this response was stronger than in the cultivar E6206 (Figure 49B).

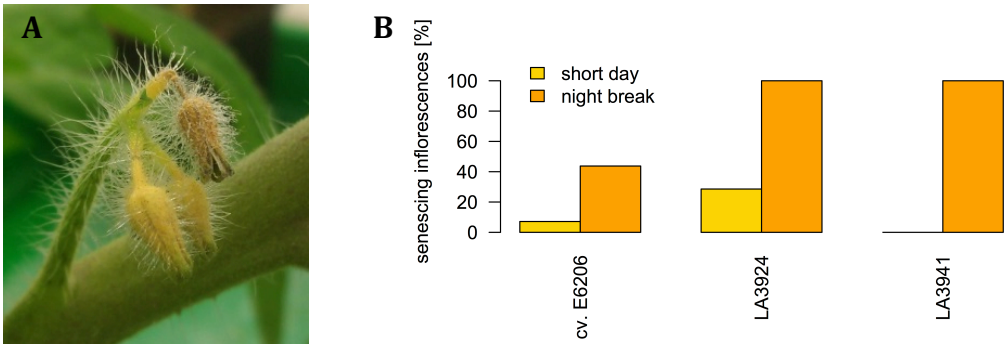

Figure 49: Night breaks lead to senescing inflorescences. (A) Example of a senescing inflorescence (picture taken by Cris Wijnen). (B) Percentage of senescing inflorescences of the cultivar E6206 and two ILs with stronger response than the cultivar under short day and night break conditions.

Thus, inhibition of flowering under non-inductive photoperiods (night breaks) is stronger in these two lines than in the cultivar E6206. In fact, three quarters of

the LA3941 plants did not produce any viable flower under night break conditions, demonstrating a strong inhibitory effect.

LA3941 has an introgression comprising the end of chromosome 5 overlapping with the peak identified in the bulk segregant analysis of *S. galapagense*. This region includes the two FT genes (*SP5G* and *SP6A*) discussed above (section 4.8.3). Genotyping of LA3924 and LA3968 with 97 CAPS markers evenly distributed over the genome did not reveal any *S. habrochaites* introgressions. The introgressions of these two lines are probably relatively small and were missed because of being located in between two markers.

4.9.2 Conclusions

In conclusion, allelic variation for *SFT* or other *FT*-like genes may be responsible for the differences in photoperiodic flowering between cultivated tomato and the wild species *S. galapagense* and *S. habrochaites*. This is in accordance with other studies that demonstrated that the photoperiodic flowering pathway is highly conserved across different plant species (Bohlenius et al., 2006; Hayama et al., 2003; Navarro et al., 2011). The loci underlying natural variation in photoperiodic flowering appear to be independent of the circadian clock, since they do not coincide with circadian rhythm QTL. Sequence and expression analysis of the candidate genes may give further insight into the bases for the differential photoperiod responses between wild and cultivated tomato.

5 Discussion

5.1 Naturally occurring variation in circadian rhythms among cultivated and wild tomato accessions

It is widely accepted that the circadian clock coordinates diverse physiological outputs to optimize growth and development in plants (Hsu and Harmer, 2013). The circadian clock is therefore thought to confer an adaptive advantage and to be of evolutionary significance.

In this thesis it was demonstrated that there is pronounced natural variation in circadian rhythms among wild and cultivated tomato accessions. Notably, this variation differentiates the cultivated varieties from their wild relatives. This differentiation is particularly clear for the time of the second peak, here termed peak time. There is an evident relationship between peak time and period. The two variables are highly correlated (Pearson's ρ = 0.88, p = 2.2e-16) because longer periods lead to increased peak times. Nevertheless, an asymmetry of the waveform during the first circadian cycle (Figure 50) only present in cultivated tomato seems to significantly contribute to the later peak times of cultivated tomato compared to its wild relatives. This notion is supported by the identification of a peak time specific QTL

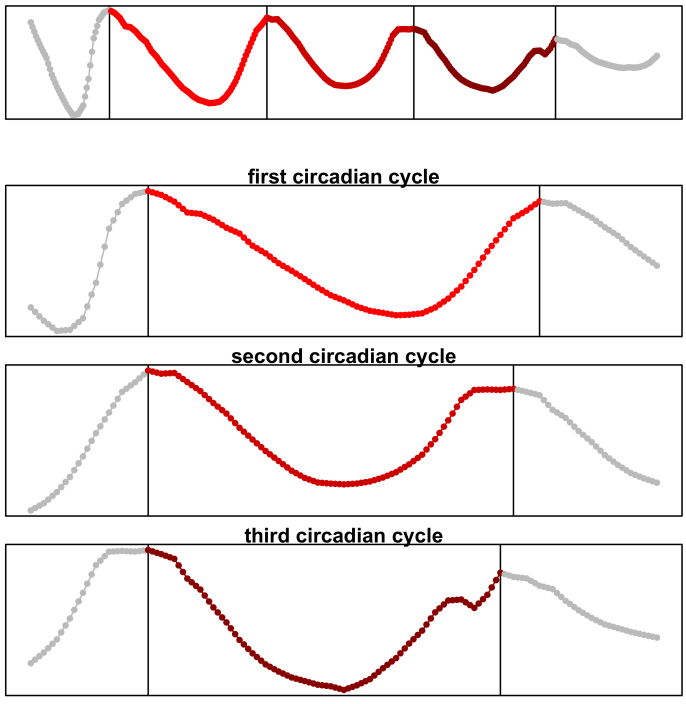

Figure 50: The first circadian cycle of cultivated tomato exhibits pronounced asymmetry. Mean relative cotyledon position of S. lycopersicum cv. M82 (n = 36). Data are detrended by taking the residuals of a polynomial regression. Vertical lines indicate local maxima. Leaf movements are shown over five days in the upper panel and over two days in the three lower panels.

without a co-localizing period QTL.

In a simplified view, period can be defined as the distance between two consecutive peaks. The asymmetric waveform of cultivated tomato leads to a larger distance between the two peaks of the first circadian cycle, compared to the peaks of the second and third cycles. Thus, the period is not constant in cultivated tomato but rather changes over time. The oscillation is slower at the beginning and then speeds up (Figure 50). It should be mentioned that the first 24 hours of the time-courses were excluded before the data analysis with BRASS, a standard software that estimates circadian variables such as period, phase and amplitude. This is a common practice to eliminate initial noise potentially caused by the transfer from entraining to constant conditions. The period estimates of the cultivated tomato varieties should therefore not be affected by the asymmetry of the first circadian cycle.

The fast Fourier transform nonlinear least-squares analysis employed by BRASS is considered the standard method to analyze circadian time course data. Period, phase and amplitude values are estimated from the fit of a perfectly symmetric sinusoid. However, this method does not provide information about the shape of the waveform. By taking into account peak time, at least one aspect of the waveform was captured in this thesis – the asymmetry of the first circadian cycle. It has been shown before that differences in circadian period can have an influence on plant performance under diurnal light dark cycles (Dodd et al., 2005). Here, evidence is provided that differences in the waveform of the first circadian cycle may also affect seedling growth under diurnal conditions.

It would be an interesting work for the future to try to extract other variables from time-course data that represent more detailed properties of the waveforms than period, phase and amplitude.

5.1.1 Effect of domestication on the circadian clock of tomato

For Arabidopsis and cyanobacteria, it has been proposed that matching of the circadian cycle with the environment enhances growth and fitness (Dodd et al., 2005; Ouyang et al., 1998). This implies that every organism should target a circadian period of 24 hours. But as a matter of fact, various studies have shown that there is marked variation in circadian rhythms and that some genotypes de-

viate strongly from a 24-hour period (Emery et al., 1994; Lou et al., 2011; Michael et al., 2003b). Concordantly, this study finds a high degree of naturally occurring variation in circadian rhythms among wild and cultivated tomato accessions. The circadian periods range from 21.0 to 26.2 hours in the wild accessions and from 24.6 to 27.8 hours in the cultivars tested.

Notably, accessions from *S. pimpinellifolium*, the closest wild relative of *S. lycopersicum*, have significantly shorter periods and earlier peak times than cultivated tomato varieties (p = 3.66e-11 in t-test for period, p = 2.2e-16 in t-test for peak time). A combination of the period and peak time values differentiates the two groups in a dichotomous way. Since cultivated tomato *S. lycopersicum* is not a natural species but emerged through human selection, this strongly suggests that the circadian clock of tomato was changed during domestication or early breeding. To my knowledge, this is the first study reporting variation in circadian period and peak time due to domestication or crop evolution.

Several of the cultivars that were used for this thesis are very old, e.g. 'Yellow Pear', which was first described in 1805, 'Rutgers', which was developed in New Jersey in the early 20th century or 'Aisla Craig' and 'Moneymaker', which were bred in the UK at the beginning of last century. All of these cultivars have the long period and late peak time phenotypes (see Table A6 in Appendix), ruling out a recent breeding event as the cause for the change in circadian rhythms.

Interestingly, the cultivars classified as Latin American, which possibly are direct descendants of the first cultivated varieties, have significantly longer periods and later peak times than *S. pimpinellifolium* accessions (t-test, $p < 0.05$). At the same time their periods and peak times are shorter and earlier than those of the 'modern' cultivars ('modern' referring to all varieties that were developed after the transfer of tomato from Mexico to Europe in the 16th century). This observation suggests that the deceleration of circadian rhythms in cultivated tomato may have happened in two steps. In this scenario, the clock would have been modulated during domestication, which was accompanied by the migration of tomato from South America to Mesoamerica. Crop evolution after the transfer of tomato from Mexico to Europe would have further changed the clock, leading to even greater differentiation between cultivated tomato and its wild ancestors.

5.1.2 Identification of the genetic basis for the variation in circadian rhythms

For Arabidopsis an elaborate model of the circadian clock consisting of many interacting genes forming interconnected transcriptional feedback loops has been constructed (Nagel and Kay, 2012). Most genes in this circadian clock model have been identified via mutant screens. Mutagenized populations were monitored directly for circadian rhythms (Millar et al., 1995; Somers et al., 2000), for photoperiod insensitive flowering (Doyle et al., 2002; Hicks et al., 1996; Schaffer et al., 1998) and for hypocotyl elongation (Hazen et al., 2005). While these approaches were of great use in deciphering the circadian clock network of Arabidopsis, they do not provide information about the genes that were targeted by natural selection to modulate circadian rhythms allowing adaption to specific environments. Such information can only be obtained by the analysis of naturally occurring variation. By mapping QTL for circadian rhythms the genes underlying naturally occurring differences between accessions can be identified. Likewise, the genetic basis for the differences between a domesticated species and its wild ancestor can be resolved. QTL mapping employing populations generated from crosses between cultivated and wild accessions will yield loci possibly underlying the variation caused by domestication.

5.1.2.1 Genetic architecture of the variation in circadian rhythms in tomato – two putative domestication QTL

The QTL analysis in a RIL population generated from a cross between cultivated tomato and its wild progenitor *S. pimpinellifolium* revealed a simple genetic architecture of circadian leaf movement variation. One period QTL was identified on chromosome 1. Additionally, two peak time QTL were detected on chromosomes 1 and 9, the first one overlapping the period QTL. This simple genetic architecture supports the notion that altered circadian rhythms may have played a role during tomato domestication, as domestication traits are often mono- or digenic (Abbo et al., 2014). Also, domestication phenotypes are often caused by recessive alleles. Interestingly, the domesticated alleles are recessive for both circadian rhythm QTL identified in the *S. pimpinellifolium* RIL population.

The two QTL are also present in populations generated from crosses between two other wild tomato species (*S. galapagense* and *S. pennellii*) and two cultivars

(Moneymaker and M82, respectively). This strongly suggests that the causal molecular changes occurred only once in cultivated tomato. It is therefore not unlikely that these QTL are responsible for at least part of the general differences observed between wild and cultivated tomato.

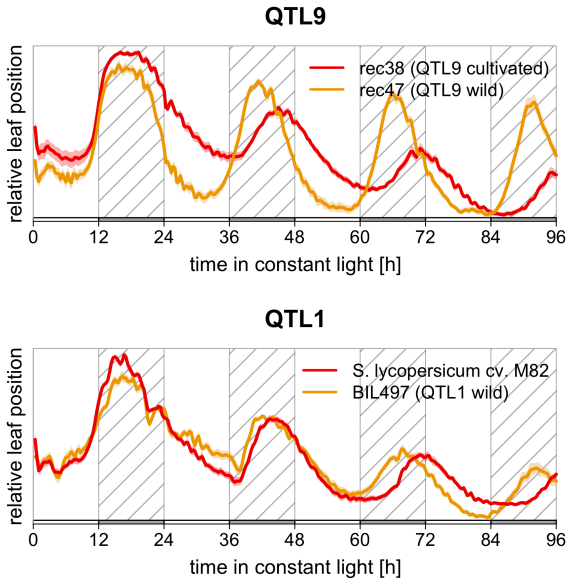

Figure 51: Two putative domestication QTL – one for peak time and one for period. Mean relative position of cotyledon tip (shaded are: ± SEM). In the upper panel two NILs - rec38 and rec47 - differing only for 13 genes (i.e. QTL9) are shown (n = 11-12). In the lower panel the cultivar *S. lycopersicum* cv. M82 and BIL497, which carries an introgression from *S. pennellii* including QTL1, are depicted (n = 9-10). Data are from one representative experiment. Hatched areas indicate subjective nights.

The peak time QTL on chromosome 9 (QTL9) seems to be caused almost entirely by differences in the waveform of the first circadian cycle, discussed above, as the locus does not confer significant period differences in the *S. pimpinellifolium* population. The wild species allele leads to a symmetric waveform, whereas plants with the cultivated allele exhibit the asymmetry typical for cultivated tomato (Figure 51 – QTL9). Period differences of around one hour associated with this locus were detected in the *S. pennellii* ILs. However, those differences were not observed in the *S. pennellii* NILs only differing for 13 genes. Thus, while the peak time differences are obvious, period differences caused by QTL9 are not consistent. The period QTL on chromosome 1 (QTL1), on the other hand, consistently shortens the circadian period by about three hours. However, it does not affect the waveform, and the associated peak time differences are caused by the period alteration. This is visible when comparing the cultivar M82 and an *S. pennellii* BIL carrying an introgression containing QTL1: The waveform of the circadian leaf movements is very similar in both lines, but the shorter period in the BIL leads to decoupling of the rhythms in the final days of the experiment (Figure 51 – QTL1).

In summary, differences between cultivated tomato and its wild progenitor seem to be caused by two loci. Pyramiding of these two QTL in a cultivated variety should restore the *S. pimpinellifolium* circadian phenotype. It would be interesting to cross the according ILs to test this.

5.1.2.2 Molecular basis for the peak time QTL on chromosome 9

Natural variation in circadian rhythms has been investigated in Arabidopsis (Michael et al., 2003b; Swarup et al., 1999) and in two *Brassica* species (Lou et al., 2011; Salathia et al., 2007), resulting in the identification of four to six QTL in each of those studies. However, none of the genes underlying these QTL have been cloned. In this thesis, a candidate gene for a peak time QTL was identified by positional cloning with the help of next generation sequencing data, functional annotation and integration of information from other species. The candidate gene is the homolog of the Arabidopsis Phytochrome A associated F-box protein *EMPFLINDLICHER IM DUNKELROTEN LICHT1* (*EID1*). The complementation experiments by transformation of the allelic variants to prove that *EID1* is the gene underlying the QTL, are currently being carried out.

Since *EID1* appears to be responsible for part of the general differences in circadian rhythms between wild and cultivated tomato, the causal sequence variant should be present in every cultivar but absent in every wild species accession. Two sequence variants in this gene differentiate the three wild parents from the two cultivated parents of the mapping populations used: a 3 bp indel and a non-synonymous SNP. The wild species alleles for these two polymorphisms are also shared with the homologs of *EID1* in cultivated potato and pepper. The alleles for these two polymorphisms in cultivated tomato therefore clearly represent a derived state. Recently, genome sequencing data of 95 tomato accessions have been made available by the 150 tomato genome consortium in collaboration with the user community (http://www.tomatogenome.net/). According to these data the two *EID1* variants (the indel and the non-synonymous SNP) are present in every 'modern' cultivar (n = 40), i.e. all of those cultivars harbor the derived alleles. Interestingly, the derived allele for the non-synonymous SNP is also present in two *S. pimpinellifolium* accessions. These two accessions do not show different rhythms compared to the other *S. pimpinellifolium* accessions that do not contain

the SNP, arguing against the assumption that this polymorphism represents the QTN. On the other hand, none of the wild accessions (n = 31) possess the 3 bp deletion present in all the 'modern' cultivars. This deletion could therefore be the causal sequence variant, as it perfectly differentiates wild from 'modern' cultivated tomato. It is worth mentioning that the 3 bp indel but not the SNP change a region of the protein that is highly conserved among five Solanaceae species (cultivated pepper – *Capsicum annuum*, cultivated potato – *S. tuberosum*, the wild tomato species *S. pennellii* and *S. pimpinellifolium* and cultivated tomato).

Interestingly, of the eight tomato cultivars from South or Central America four have the wild species allele. This suggests that the two alleles might be present at equal frequencies in the Latin American cultivars and that the deletion may have been fixed with the transfer of tomato to Europe in the 16th century. This would also be consistent with the intermediate circadian phenotype observed for the Latin American cultivars. Nevertheless, the difficulty of unambiguously classifying these cultivars has to be taken into account. It cannot be ruled out that European cultivars were reintroduced to Latin America, leading to the presence of 'modern' tomato alleles in the cultivars classified as Latin American (Rick, 1958). On the other hand, it is also possible that hybridization events with *S. pimpinellifolium* led to introgression of the wild species alleles (Rick, 1958).

Notably, the three South American *S. lycopersicum* var. *cerasiforme* accessions collected in Ecuador, have the wild species allele. Since their phylogenetic position is also uncertain - the *S. l. cerasiforme* accessions could represent either ancient cultivars or an admixture of *S. lycopersicum* with *S. pimpinellifolium* - this observation does not allow a clear conclusion either. Nevertheless, it rather supports the scenario of the fixation of the deletion in *EID1* after the initial domestication process, i.e. during crop evolution.

In Arabidopsis, *EID1* is involved in phytochrome A specific light signaling. The loss of function leads to enhanced phyA activity (Dieterle et al., 2001). Since overexpression of phyA in Arabidopsis leads to a shortening of the circadian period (Anderson et al., 1997; Kolmos et al., 2011), the longer period observed for one of the SALK insertion mutants tested in this thesis is unexpected. Nevertheless, because *EID1* is phyA associated in Arabidopsis, changes in circadian rhythms in tomato, supposedly caused by *EID1*, could be phytochrome depend-

ent. It would be interesting to test the two *EID1* alleles in phytochrome deficient backgrounds. Another interesting experiment would be to test circadian leaf movements of the two NILs under different light conditions: blue light and red light. A red light specific difference would strongly support the hypothesis of *EID1* affecting the circadian clock through phytochrome signaling.

5.1.3 Possible reasons for the variation in circadian rhythms

So far it has been discussed that the tomato clock was changed during domestication or/and early breeding, leading to marked differences in circadian rhythms between wild and cultivated tomato accessions. Also it has been discussed that these differences are mainly caused by two QTL, one of which is probably explained by allelic variation in the tomato homolog of the Arabidopsis *EID1* gene. The most important question that remains is whether humans actively selected the differences in circadian rhythms during domestication or breeding of tomato. This would imply that the differences in circadian rhythms are advantageous for agriculture. Strikingly, NILs only differing in the 13 genes of QTL9 exhibit significant differences (t-test, $p < 0.05$) in seedling height under greenhouse conditions. This suggests that differences in circadian rhythms associated with QTL9 may actually affect plant growth under diurnal conditions. Such an effect of circadian rhythms on plants under natural conditions opens the possibility of direct selection. Alternatively, variation in circadian rhythms could be neutral and have emerged randomly through a founder effect and genetic drift. The third possibility is that the altered circadian clock of tomato is negative for the crop and was caused by hitchhiking of the corresponding alleles, which can happen through close linkage to a gene under positive selection. In the first and in the last case, signatures of selection should be detectable for the regions responsible for altered circadian rhythms. It will be very interesting to do the relevant analyses in the future. Data needed for this has already been generated and is awaiting publication (Huang et al., 2013).

Additionally, in the case of direct selection or indirect selection via hitchhiking, QTL for agronomically important traits should co-localize with the loci underlying the differences in circadian rhythms. No evidence was found for this in data from previous studies (Causse et al., 2004; Eshed and Zamir, 1995), although it is

possible that expression of the phenotypic differences is dependent on particular environmental conditions. Another possibility is that potential effects may have been masked by the heterogeneous backgrounds of the ILs, since they still have relatively big introgressions. To test whether the circadian rhythm QTL are directly associated with agronomic traits, NILs with contrasting circadian phenotypes only differing in very few genes, or ultimately transgenic tomatoes expressing the two alleles of the genes causative for the circadian rhythm QTL, should be thoroughly phenotyped for traits of agronomic significance. Ideally, experiments should be performed under various conditions and using various genetic backgrounds. Because the change in circadian rhythms most probably occurred more than 100 years ago, differences in agronomic traits due to an altered circadian clock may not be visible in the modern elite breeding varieties used today, many of which have a mutation in the *SELF PRUNING* (*SP*) gene, leading to a determinate growth habit. Recently yield heterosis has been reported to be dependent on the presence of *SP* (Jiang et al., 2013). Therefore transgenic plants are currently being generated using the modern determinate cultivar M82, but also the much older indeterminate cultivar Moneymaker.

It does not seem unlikely that the altered circadian clock observed in cultivated tomato was under direct artificial selection conferring an adaptive advantage. In different species latitudinal clines of naturally occurring variation in circadian rhythms have been reported, suggesting that modulation of the circadian system is important for adaptation to specific environments (Michael et al., 2003b; Pittendrigh et al., 1991; Simunovic and Jaenike, 2006). As a matter of fact, during domestication and subsequent crop evolution, tomato has faced considerable environmental changes. One very obvious change was the alteration in photoperiod due to the transfer from the equatorial region to Mesoamerica and later to Europe. In both of those steps tomato was moved northwards by approximately 20° of latitude and thereby photoperiods changed considerably. Whereas the day length is always 12 hours at the equator, during summer it is approximately 13 hours in Mexico City and 15 hours in Rome.

Cultivated tomato and its ancestors are injured by the presence of light at an inappropriate time of the daily cycle, which is set by the circadian clock (Highkin and Hanson, 1954). These injuries can lead to decreased plant performance, as

demonstrated for plants grown under photoperiods of 17 hours compared to 12 hours (Arthur et al., 1930). Therefore, avoidance of light-induced injuries could have been the factor driving selection for slower rhythms in cultivated tomato. Deceleration of the circadian clock in light could have led to a later onset of the so called scotophil phase (Bünning, 1950) in which light is damaging. This could have led to enhanced growth in the long day environments found in temperate regions worldwide. However, considering the extent to which tomato was changed during domestication and breeding, and the pervasive control of the circadian clock, modulation of circadian rhythms could be adaptive by affecting any of the numerous clock outputs.

5.2 Photoperiodic flowering in two wild tomato relatives

For many plants the correct seasonal timing of flowering time is essential for reproductive success. Synchronization via photoperiodic control of flowering has been used in agriculture to maximize seed yields (Andres and Coupland, 2012). However, in most reported cases, loss of photoperiod sensitivity has been favored during crop evolution (Murphy et al., 2011; Pin et al., 2012; Turner et al., 2005; Weller et al., 2012). Cultivated tomato and its wild ancestor *S. pimpinellifolium* are both day neutral plants, i.e. they do not exhibit differences in flowering time in response to photoperiod. However, some accessions from the closely related species *S. galapagense*, which is endemic to the Galapagos Islands, show short day flowering behavior. Also accessions from some other wild tomato species, including *S. habrochaites* and species from the *S. peruvianum* complex, are short day plants. The polyphyletic character of photoperiodic flowering behavior in the tomato clade indicates that it arose several times independently. One of the first scientific reports on effects of day length on flowering was motivated by the occasional sudden occurrence of giant tobacco plants (Allard, 1919; Garner and Allard, 1920). This gigantism is caused by a mutation in a single gene leading to obligatory short day flowering behavior. These examples from tomato and tobacco demonstrate that a seemingly complex trait - photoperiodic flowering - can appear through spontaneous mutations. However, the molecular and genetic bases for such switches in photoperiod response have not yet been identified.

5.2.1 QTL analysis for photoperiodic flowering in *S. habrochaites*

The analysis of photoperiodic flowering in an introgression line population resulted in the identification of three lines differing from the recurrent parent *S. lycopersicum* cv. E6206. Flowering in those lines was inhibited or delayed by night breaks, which are commonly used to assess photoperiod responsive flowering. However, the introgression could be assigned to a chromosomal region only for one of those lines (LA3941). For the other two lines (LA3924 and LA3968) introgressions from the wild species *S. habrochaites* were not detected. This suggests that the two lines may contain relatively small segments that were not covered by any of the genotyping markers. It will be interesting to genotype those two lines with higher resolution. One possibility to do this is to use RNA sequencing. Since almost every gene has a SNP distinguishing *S. lycopersicum* from *S. habrochaites*, RNAseq of the ILs would lead to the identification of thousands of markers, thus allowing the introgressions to be precisely defined. At the same time differentially expressed and alternatively spliced genes could be identified if not only the RNA of the ILs but also of the recurrent parent is sequenced. Since about half of all QTL cloned in diverse species and for various traits appear to be due to differences in expression of the underlying gene and thus represent cis-eQTL (Alonso-Blanco et al., 2009), the ILs and the recurrent parent should both be sequenced.

In line 41 (LA3941) an introgression from *S habrochaites* was found at the end of chromosome 5. In that chromosomal region there are three obvious candidate genes: the two *FT* paralogs, *SP5G* (Solyc05g053850) and *SP6A* (Solyc05g055660), and the photoreceptor *PHYTOCHROME B2* (*PhyB2*, Solyc05g053410), whose Arabidopsis homolog is known to be involved in the control of photoperiodic flowering (Andres and Coupland, 2012). It has been reported that *SP6A* is not expressed in cultivated tomato due to a premature stop codon (Carmel-Goren et al., 2003). Assuming that the wild allele of *SP6A* has a positive effect on flowering time, as has been suggested, and taking into account that the QTL inhibits flowering under night breaks rather than accelerating it under short days, this gene is not a likely candidate. On the other hand, *SP5G* has been proposed to negatively affect flowering (Navarro et al., 2011). This gene therefore is a good candidate. Sequence and expression analysis of the candidate

genes may give further insight into the genetic basis for the QTL. Ultimately, transformation of the recessive genotype with the opposing allele or silencing of the gene in the dominant genotype should revert the photoperiod response and would prove the corresponding gene to underlie the QTL.

5.2.2 Bulk segregant analysis via RNAseq – mapping the long day specific inhibition of flowering in *S. galapagense*

Not only mutations but also QTL can be mapped rapidly by sequencing the DNA of the phenotypically extreme individuals of mapping populations (Schneeberger and Weigel, 2011). Also it has been shown that RNAseq can successfully be used for mutation mapping (Hill et al., 2013; Miller et al., 2013). However, mapping QTL using RNAseq has not yet been reported. To map the loci underlying the photoperiod response of *S. galapagense* - flowering in *S. galapagense* (LA0530) is strongly inhibited by long day photoperiods - bulk segregant analysis (BSA) via RNAseq was performed using an F2 population. Allele frequencies of the two parents were determined for the early flowering and non-flowering bulk. One apparent problem associated with this method is the influence of expression levels on the allele frequencies. Differential expression between the parents will translate into differences in allele frequencies. DNAseq data would not exhibit this problem. Nevertheless, the apparent drawback also provides a benefit: differential expression can be analyzed between the two bulks, allowing the identification of potential eQTL.

Two putative QTL were identified using the BSA / RNAseq approach, indicating that QTL mapping via RNAseq of bulks is feasible. However, the approach is only applicable when the trait of interest exhibits a simple genetic architecture. For traits controlled by more than two QTL, differences in allele frequencies between the two bulks would likely become too small to be reliably distinguished from background noise.

One of the two QTL identified contains the end of chromosome 5. Interestingly, this QTL coincides with the one identified in the *S. habrochaites* IL population. Thus, the same two genes are candidates for the QTL: *SP5G* and *PhyB2*. The other QTL covers almost the complete chromosome 3 and includes *SFT* (Solyc03g063100), the tomato homolog of the major flowering regulator *FT* of

Arabidopsis. Remarkably, the grafting experiments already pointed to *SFT* as a candidate for the differences in photoperiodic flowering. The lack of *SFT* expression under long days in *S. galapagense* would be the most straightforward explanation for the photoperiod response. Surprisingly, *SFT* expression does not show any differences in expression between the two bulks. This indicates that differences in protein function, leading to long-day specific inhibition of SFT, might be responsible for the phenotype. However, the equal expression for *SFT* in flowering and non-flowering *S. galapagense* plants still needs to be confirmed.

One final question that arises is whether the photoperiodic flowering response of *S. galapagense* is of ecological significance, conferring an adaptive advantage to the plants in their natural environment. Since the Galapagos Islands are located directly at the equator where photoperiod is constant throughout the year, it is not evident why *S. galapagense* plants should respond to differences in day length. Nevertheless, it has been suggested that the mechanism that is manifested as a response to day length in growth chambers may exist in nature to measure small differences in sunrise or sunset close to the equator, enabling synchronized flowering (Borchert et al., 2005). This can be important for synchronization with related individuals for cross-pollination or for synchronization with pollinators. Accessions of *S. galapagense* exhibit a high degree of self-pollination (Rick and Fobes, 1975). Still, the vibration caused by pollinating insects can increase fruit set in self-pollinating tomatoes (Rick and Dempsey, 1969). Carpenter bees, which are the main pollinator of the Galapagos Islands, have been observed on *S. galapagense* plants in nature (McMullen, 1989). It is therefore conceivable that the photoperiodic flowering observed in *S. galapagense* leads to seasonal flowering on the Galapagos Islands coinciding with yearly fluctuations in pollinator activity. It will be interesting to study the frequency of photoperiodic flowering among *S. galapagense* accessions. The accession LA0528, for example, does not exhibit photoperiodic flowering, even though it appears phenotypically very similar to LA0530. The analysis of several different accessions would allow the association of flowering behavior with sequence variation of putative candidate genes underlying the differences in photoperiodic flowering to be tested.

5.3 Synopsis

In conclusion, in this thesis naturally occurring variation in circadian rhythms and photoperiodic flowering were analyzed and QTL underlying this variation were identified. A previously unknown photoperiodic flowering response of the close tomato relative *S. galapagense* was discovered and mechanistic and genetic principles of this response were examined. Further, it was revealed that domestication or early breeding very likely had an effect on the circadian clock of cultivated tomato, making it the first study reporting differences in circadian period and peak time due to human selection. One locus most probably underlying part of the general differences in circadian rhythms between wild and cultivated tomato was fine-mapped and a candidate gene was cloned. Additionally, the results obtained give rise to several new questions thereby providing a basis for follow-up studies.

References

Abbo, S., Pinhasi van-Oss, R., Gopher, A., Saranga, Y., Ofner, I., and Peleg, Z. (2014). Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes. Trends Plant Sci.

Allard, H.A. (1919). Gigantism in Nicotiana tabacum and its alternative inheritance. Am Nat *53*, 218-233.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R.*, et al.* (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science *301*, 653-657.

Alonso-Blanco, C., Aarts, M.G., Bentsink, L., Keurentjes, J.J., Reymond, M., Vreugdenhil, D., and Koornneef, M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell *21*, 1877-1896.

Alonso-Blanco, C., Koornneef, M., and Stam, P. (1998). The use of recombinant inbred lines (RILs) for genetic mapping. Methods Mol Biol *82*, 137-146.

Anderson, S.L., Somers, D.E., Millar, A.J., Hanson, K., Chory, J., and Kay, S.A. (1997). Attenuation of phytochrome A and B signaling pathways by the Arabidopsis circadian clock. Plant Cell *9*, 1727-1743.

Andres, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. Nat Rev Genet *13*, 627-639.

Arthur, J.M., Guthrie, J.D., and Newell, J.M. (1930). Some effects of artificial climates on the growth and chemical composition of plants. Am J Bot *17*, 416-482.

Bai, Y., and Lindhout, P. (2007). Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? Ann Bot *100*, 1085-1094.

Baudry, A., Ito, S., Song, Y.H., Strait, A.A., Kiba, T., Lu, S., Henriques, R., Pruneda-Paz, J.L., Chua, N.H., Tobin, E.M.*, et al.* (2010). F-box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control Arabidopsis clock progression. Plant Cell *22*, 606-622.

Bell-Pedersen, D., Cassone, V.M., Earnest, D.J., Golden, S.S., Hardin, P.E., Thomas, T.L., and Zoran, M.J. (2005). Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nat Rev Genet *6*, 544-556.

Ben-Naim, O., Eshed, R., Parnis, A., Teper-Bamnolker, P., Shalit, A., Coupland, G., Samach, A., and Lifschitz, E. (2006). The CCAAT binding factor can mediate interactions between CONSTANS-like proteins and DNA. Plant J *46*, 462-476.

Blanca, J., Canizares, J., Cordero, L., Pascual, L., Diez, M.J., and Nuez, F. (2012). Variation revealed by SNP genotyping and morphology provides insight into the origin of the tomato. PLoS One *7*, e48198.

Bohlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A.M., Jansson, S., Strauss, S.H., and Nilsson, O. (2006). CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. Science *312*, 1040-1043.

Borchert, R., Renner, S.S., Calle, Z., Navarrete, D., Tye, A., Gautier, L., Spichiger, R., and von Hildebrand, P. (2005). Photoperiodic induction of synchronous flowering near the Equator. Nature *433*, 627-629.

Broman, K.W., and Sen, S. (2009). A guide to QTL mapping with R/qtl (Dordrecht: Springer).

Buche, C., Poppe, C., Schafer, E., and Kretsch, T. (2000). eid1: a new Arabidopsis mutant hypersensitive in phytochrome A-dependent high-irradiance responses. Plant Cell *12*, 547-558.

Bünning, E. (1932). Über die Erblichkeit der Tagesperiodizität bei den Phaseolus-Blättern. Jahrbücher für wissenschaftliche Botanik *77*, 283-320.

Bünning, E. (1936). Die endonome Tagesrhythmik als Grundlage der photoperiodischen Reaktion. Berichte der deutschen botanischen Gesellschaft *54*, 590-607.

Bünning, E. (1950). Über die photophile und skotophile phase der endogenen tagesrhythmik. Planta *38*, 521-540.

Campoli, C., Shtaya, M., Davis, S.J., and von Korff, M. (2012). Expression conservation within the circadian clock of a monocot: natural variation at barley Ppd-H1 affects circadian expression of flowering time genes, but not clock orthologs. BMC Plant Biol *12*, 97.

Carmel-Goren, L., Liu, Y.S., Lifschitz, E., and Zamir, D. (2003). The SELF-PRUNING gene family in tomato. Plant Mol Biol *52*, 1215-1222.

Causse, M., Duffe, P., Gomez, M.C., Buret, M., Damidaux, R., Zamir, D., Gur, A., Chevalier, C., Lemaire-Chamley, M., and Rothan, C. (2004). A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. J Exp Bot *55*, 1671-1685.

Chakrabarti, M., Zhang, N., Sauvage, C., Munos, S., Blanca, J., Canizares, J., Diez, M.J., Schneider, R., Mazourek, M., McClead, J.*, et al.* (2013). A cytochrome P450 regulates a domestication trait in cultivated tomato. Proc Natl Acad Sci U S A *110*, 17125-17130.

Chitwood, D.H., Kumar, R., Headland, L.R., Ranjan, A., Covington, M.F., Ichihashi, Y., Fulop, D., Jimenez-Gomez, J.M., Peng, J., Maloof, J.N.*, et al.* (2013). A quantitative genetic basis for leaf morphology in a set of precisely defined tomato introgression lines. Plant Cell *25*, 2465-2481.

Cong, B., Barrero, L.S., and Tanksley, S.D. (2008). Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet *40*, 800-804.

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., Giakountis, A., Farrona, S., Gissot, L., Turnbull, C.*, et al.* (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science *316*, 1030-1033.

de Candolle, A. (1886). Origin of Cultivated Plants.

Dieterle, M., Zhou, Y.C., Schafer, E., Funk, M., and Kretsch, T. (2001). EID1, an F-box protein involved in phytochrome A-specific light signaling. Genes Dev *15*, 939-944.

Dodd, A.N., Salathia, N., Hall, A., Kevei, E., Toth, R., Nagy, F., Hibberd, J.M., Millar, A.J., and Webb, A.A. (2005). Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science *309*, 630-633.

Doebley, J.F., Gaut, B.S., and Smith, B.D. (2006). The molecular genetics of crop domestication. Cell *127*, 1309-1321.

Dore, J. (1959). Response of Rice to Small Differences in Length of Day. Nature *183*, 413-414.

Doyle, M.R., Davis, S.J., Bastow, R.M., McWatters, H.G., Kozma-Bognar, L., Nagy, F., Millar, A.J., and Amasino, R.M. (2002). The ELF4 gene controls circadian rhythms and flowering time in Arabidopsis thaliana. Nature *419*, 74-77.

Edgar, R.S., Green, E.W., Zhao, Y., van Ooijen, G., Olmedo, M., Qin, X., Xu, Y., Pan, M., Valekunja, U.K., Feeney, K.A.*, et al.* (2012). Peroxiredoxins are conserved markers of circadian rhythms. Nature *485*, 459-464.

Emery, P.T., Morgan, E., and Birley, A.J. (1994). An investigation of natural genetic variation in the circadian system of Drosophila melanogaster: rhythm characteristics and methods of quantification. Chronobiol Int *11*, 72-84.

Eshed, Y., and Zamir, D. (1995). An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics *141*, 1147-1162.

Facella, P., Lopez, L., Carbone, F., Galbraith, D.W., Giuliano, G., and Perrotta, G. (2008). Diurnal and circadian rhythms in the tomato transcriptome and their modulation by cryptochrome photoreceptors. PLoS One *3*, e2798.

Faure, S., Turner, A.S., Gruszka, D., Christodoulou, V., Davis, S.J., von Korff, M., and Laurie, D.A. (2012). Mutation at the circadian clock gene EARLY MATURITY 8 adapts domesticated barley (Hordeum vulgare) to short growing seasons. Proc Natl Acad Sci U S A *109*, 8328-8333.

Finkers, R., Smit, S., van Heusden, A.W., Visser, R.G., Sanchez-Perez, G., Aflitos, S., Schijlen, E., Borm, T., de Jong, H., and Peters, S. (2013). Discovery of patterns of structural variation in the tomato clade. 10th Solanaceae Conference Abstract Book, 76.

Foolad, M.R. (2007). Genome mapping and molecular breeding of tomato. Int J Plant Genomics *2007*, 64358.

Frary, A., Nesbitt, T.C., Grandillo, S., Knaap, E., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K.B., and Tanksley, S.D. (2000). fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science *289*, 85-88.

Fridman, E., Pleban, T., and Zamir, D. (2000). A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc Natl Acad Sci U S A *97*, 4718-4723.

Garner, W.W., and Allard, H.A. (1920). Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. J Agric Res *18*, 0553-0606.

Grandillo, S., Chetelat, R.T., Knapp, S., Spooner, D., Peralta, I., Cammareri, M., Perez, O., Termolino, P., Tripodi, P., Chiusano, M.L.*, et al.* (2011). Solanum sect. Lycopersicon. Wild Crop Relatives: Genomics and Breeding Resources, Vegetables, 129-215.

Green, R.M., Tingay, S., Wang, Z.Y., and Tobin, E.M. (2002). Circadian rhythms confer a higher level of fitness to Arabidopsis plants. Plant Physiol *129*, 576-584.

Gur, A., and Zamir, D. (2004). Unused natural variation can lift yield barriers in plant breeding. PLoS Biol *2*, e245.

Harmer, S.L. (2009). The circadian system in higher plants. Annu Rev Plant Biol *60*, 357-377.

Hayama, R., Yokoi, S., Tamaki, S., Yano, M., and Shimamoto, K. (2003). Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature *422*, 719-722.

Hazen, S.P., Schultz, T.F., Pruneda-Paz, J.L., Borevitz, J.O., Ecker, J.R., and Kay, S.A. (2005). LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. Proc Natl Acad Sci U S A *102*, 10387-10392.

Hicks, K.A., Millar, A.J., Carre, I.A., Somers, D.E., Straume, M., Meeks-Wagner, D.R., and Kay, S.A. (1996). Conditional circadian dysfunction of the Arabidopsis early-flowering 3 mutant. Science *274*, 790-792.

Highkin, H.R., and Hanson, J.B. (1954). Possible Interaction between Light-dark Cycles and Endogenous Daily Rhythms on the Growth of Tomato Plants. Plant Physiol *29*, 301-302.

Hill, J.T., Demarest, B.L., Bisgrove, B.W., Gorsi, B., Su, Y.C., and Yost, H.J. (2013). MMAPPR: mutation mapping analysis pipeline for pooled RNA-seq. Genome Res *23*, 687-697.

Hillman, W.S. (1956). Injury of tomato plants by continuous light and unfavorable photoperiodic cycles. Am J Bot *43*, 89-96.

Hsu, P.Y., and Harmer, S.L. (2013). Wheels within wheels: the plant circadian system. Trends Plant Sci.

Huang, S., Du, Y., Ye, Z., and Li, J. (2013). A brief genomic history of tomato breeding. 10th Solanaceae Conference Abstract Book, 74.

Izawa, T., Oikawa, T., Sugiyama, N., Tanisaka, T., Yano, M., and Shimamoto, K. (2002). Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. Genes Dev *16*, 2006-2020.

Jackson, S.D. (2009). Plant responses to photoperiod. New Phytol *181*, 517-531.

Jenkins, J.A. (1948). The Origin of the Cultivated Tomato. Economic Botany *2*, 379-392.

Jiang, K., Liberatore, K.L., Park, S.J., Alvarez, J.P., and Lippman, Z.B. (2013). Tomato yield heterosis is triggered by a dosage sensitivity of the florigen pathway that fine-tunes shoot architecture. PLoS Genet *9*, e1004043.

Jimenez-Gomez, J.M. (2011). Next generation quantitative genetics in plants. Front Plant Sci *2*, 77.

Jongsma, M., Koornneef, M., Zabel, P., and Hille, J. (1987). Tomato protoplast DNA transformation: physical linkage and recombination of exogenous DNA sequences. Plant Mol Biol *8*, 383-394.

Joshi, D.S. (1999). Latitudinal variation in locomotor activity rhythm in adult Drosophila ananassae. Can J Zool *77*, 865-870.

Kazmi, R.H., Khan, N., Willems, L.A., AW, V.A.N.H., Ligterink, W., and Hilhorst, H.W. (2012). Complex genetics controls natural variation among seed quality phenotypes in a recombinant inbred population of an interspecific cross between Solanum lycopersicum x Solanum pimpinellifolium. Plant Cell Environ *35*, 929-951.

Keurentjes, J.J., Bentsink, L., Alonso-Blanco, C., Hanhart, C.J., Blankestijn-De Vries, H., Effgen, S., Vreugdenhil, D., and Koornneef, M. (2007). Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. Genetics *175*, 891-905.

Kimura, S., Koenig, D., Kang, J., Yoong, F.Y., and Sinha, N. (2008). Natural variation in leaf morphology results from mutation of a novel KNOX gene. Curr Biol *18*, 672-677.

Kloosterman, B., Abelenda, J.A., Gomez Mdel, M., Oortwijn, M., de Boer, J.M., Kowitwanich, K., Horvath, B.M., van Eck, H.J., Smaczniak, C., Prat, S.*, et al.*

(2013). Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature *495*, 246-250.

Koenig, D., Jimenez-Gomez, J.M., Kimura, S., Fulop, D., Chitwood, D.H., Headland, L.R., Kumar, R., Covington, M.F., Devisetty, U.K., Tat, A.V.*, et al.* (2013). Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. Proc Natl Acad Sci U S A *110*, E2655-2662.

Kolmos, E., Herrero, E., Bujdoso, N., Millar, A.J., Toth, R., Gyula, P., Nagy, F., and Davis, S.J. (2011). A reduced-function allele reveals that EARLY FLOWERING3 repressive action on the circadian clock is modulated by phytochrome signals in Arabidopsis. Plant Cell *23*, 3230-3246.

Lifschitz, E., Eviatar, T., Rozman, A., Shalit, A., Goldshmidt, A., Amsellem, Z., Alvarez, J.P., and Eshed, Y. (2006). The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. Proc Natl Acad Sci U S A *103*, 6398-6403.

Ling, K.-S. (2010). Effectiveness of Chemo- and Thermotherapeutic Treatments on Pepino mosaic virus in Tomato Seed. Plant Disease *94*, 325-328.

Lippman, Z.B., Semel, Y., and Zamir, D. (2007). An integrated view of quantitative trait variation using tomato interspecific introgression lines. Curr Opin Genet Dev *17*, 545-552.

Liu, J., Van Eck, J., Cong, B., and Tanksley, S.D. (2002a). A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. Proc Natl Acad Sci U S A *99*, 13302-13306.

Liu, Y., Schiff, M., and Dinesh-Kumar, S.P. (2002b). Virus-induced gene silencing in tomato. Plant J *31*, 777-786.

Lou, P., Xie, Q., Xu, X., Edwards, C.E., Brock, M.T., Weinig, C., and McClung, C.R. (2011). Genetic architecture of the circadian clock and flowering time in Brassica rapa. Theor Appl Genet *123*, 397-409.

Mazzucato, A., Papa, R., Bitocchi, E., Mosconi, P., Nanni, L., Negri, V., Picarella, M.E., Siligato, F., Soressi, G.P., Tiranti, B.*, et al.* (2008). Genetic diversity, structure and marker-trait associations in a collection of Italian tomato (Solanum lycopersicum L.) landraces. Theor Appl Genet *116*, 657-669.

McCormick, S., Niedermeyer, J., Fry, J., Barnason, A., Horsch, R., and Fraley, R. (1986). Leaf disc transformation of cultivated tomato (L. esculentum) using Agrobacterium tumefaciens. Plant Cell Rep *5*, 81-84.

McMullen, C.K. (1989). The Galapagos carpenter bee, just how important is it? Noticias de Galapagos *48*, 16-18.

Meyer, R.S., and Purugganan, M.D. (2013). Evolution of crop species: genetics of domestication and diversification. Nat Rev Genet *14*, 840-852.

Michael, T.P., Salome, P.A., and McClung, C.R. (2003a). Two Arabidopsis circadian oscillators can be distinguished by differential temperature sensitivity. Proc Natl Acad Sci U S A *100*, 6878-6883.

Michael, T.P., Salome, P.A., Yu, H.J., Spencer, T.R., Sharp, E.L., McPeek, M.A., Alonso, J.M., Ecker, J.R., and McClung, C.R. (2003b). Enhanced fitness conferred by naturally occurring variation in the circadian clock. Science *302*, 1049-1053.

Millar, A.J., Carre, I.A., Strayer, C.A., Chua, N.H., and Kay, S.A. (1995). Circadian clock mutants in Arabidopsis identified by luciferase imaging. Science *267*, 1161-1163.

Miller, A.C., Obholzer, N.D., Shah, A.N., Megason, S.G., and Moens, C.B. (2013). RNA-seq-based mapping and candidate identification of mutations from forward genetic screens. Genome Res *23*, 679-686.

Miller, J.C., and Tanksley, S.D. (1990). RFLP analysis of phylogenetic relationships and genetic variation in the genus Lycopersicon. Theor Appl Genet *80*, 437-448.

Minoia, S., Petrozza, A., D'Onofrio, O., Piron, F., Mosca, G., Sozio, G., Cellini, F., Bendahmane, A., and Carriero, F. (2010). A new mutant genetic resource for tomato crop improvement by TILLING technology. BMC Res Notes *3*, 69.

Mizoguchi, T., Niinuma, K., and Yoshida, R. (2007). Day-neutral response of photoperiodic flowering in tomatoes: possible implications based on recent molecular genetics of Arabidopsis and rice. Plant Biotechnology *24*, 83-86.

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H.-R., Carre, I.A., and Coupland, G. (2002). LHY and CCA1 Are Partially Redundant Genes Required to Maintain Circadian Rhythms in Arabidopsis. Developmental Cell *2*, 629-641.

Monforte, A.J., and Tanksley, S.D. (2000). Development of a set of near isogenic and backcross recombinant inbred lines containing most of the Lycopersicon hirsutum genome in a L. esculentum genetic background: a tool for gene mapping and gene discovery. Genome *43*, 803-813.

Moyle, L.C. (2008). Ecological and evolutionary genomics in the wild tomatoes (Solanum sect. Lycopersicon). Evolution *62*, 2995-3013.

Munos, S., Ranc, N., Botton, E., Berard, A., Rolland, S., Duffe, P., Carretero, Y., Le Paslier, M.C., Delalande, C., Bouzayen, M.*, et al.* (2011). Increase in tomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol *156*, 2244-2254.

Murphy, R.L., Klein, R.R., Morishige, D.T., Brady, J.A., Rooney, W.L., Miller, F.R., Dugas, D.V., Klein, P.E., and Mullet, J.E. (2011). Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. Proc Natl Acad Sci U S A *108*, 16469-16474.

Nagel, D.H., and Kay, S.A. (2012). Complexity in the wiring and regulation of plant circadian networks. Curr Biol *22*, R648-657.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T., and Kimura, T. (2007). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. J Biosci Bioeng *104*, 34-41.

Nakazato, T., Bogonovich, M., and Moyle, L.C. (2008). Environmental factors predict adaptive phenotypic differentiation within and between two wild Andean tomatoes. Evolution *62*, 774-792.

Nakazato, T., Warren, D.L., and Moyle, L.C. (2010). Ecological and geographic modes of species divergence in wild tomatoes. Am J Bot *97*, 680-693.

Navarro, C., Abelenda, J.A., Cruz-Oro, E., Cuellar, C.A., Tamaki, S., Silva, J., Shimamoto, K., and Prat, S. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature *478*, 119-122.

Nesbitt, T.C., and Tanksley, S.D. (2002). Comparative sequencing in the genus Lycopersicon. Implications for the evolution of fruit size in the domestication of cultivated tomatoes. Genetics *162*, 365-379.

Nozue, K., Covington, M.F., Duek, P.D., Lorrain, S., Fankhauser, C., Harmer, S.L., and Maloof, J.N. (2007). Rhythmic growth explained by coincidence between internal and external cues. Nature *448*, 358-361.

Okabe, Y., Ariizumi, T., and Ezura, H. (2013). Updating the Micro-Tom TILLING platform. Breed Sci *63*, 42-48.

Orsi, C.H., and Tanksley, S.D. (2009). Natural variation in an ABC transporter gene associated with seed size evolution in tomato species. PLoS Genet *5*, e1000347.

Ouyang, Y., Andersson, C.R., Kondo, T., Golden, S.S., and Johnson, C.H. (1998). Resonating circadian clocks enhance fitness in cyanobacteria. Proc Natl Acad Sci U S A *95*, 8660-8664.

Paran, I., Goldman, I., Tanksley, S.D., and Zamir, D. (1995). Recombinant inbred lines for genetic mapping in tomato. Theor Appl Genet *90*, 542-548.

Pin, P.A., Zhang, W., Vogt, S.H., Dally, N., Buttner, B., Schulze-Buxloh, G., Jelly, N.S., Chia, T.Y., Mutasa-Gottgens, E.S., Dohm, J.C.*, et al.* (2012). The role of a pseudo-response regulator gene in life cycle adaptation and domestication of beet. Curr Biol *22*, 1095-1101.

Pittendrigh, C.S., Kyner, W.T., and Takamura, T. (1991). The amplitude of circadian oscillations: temperature dependence, latitudinal clines, and the photoperiodic time measurement. J Biol Rhythms *6*, 299-313.

Plautz, J.D., Straume, M., Stanewsky, R., Jamison, C.F., Brandes, C., Dowse, H.B., Hall, J.C., and Kay, S.A. (1997). Quantitative analysis of Drosophila period gene transcription in living animals. J Biol Rhythms *12*, 204-217.

Prudent, M., Causse, M., Genard, M., Tripodi, P., Grandillo, S., and Bertin, N. (2009). Genetic and physiological analysis of tomato fruit weight and composition: influence of carbon availability on QTL detection. J Exp Bot *60*, 923-937.

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., Cheng, J., Zhao, S., Xu, M., Luo, Y.*, et al.* (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. Proc Natl Acad Sci U S A.

Ranc, N., Munos, S., Santoni, S., and Causse, M. (2008). A clarified position for Solanum lycopersicum var. cerasiforme in the evolutionary history of tomatoes (solanaceae). BMC Plant Biol *8*, 130.

Rick, C.M. (1958). The role of natural hybridization in the derivation of cultivated tomatoes of western South America. Economic Botany *12*, 346-367.

Rick, C.M. (1973). Potential genetic resources in tomato species: clues from observations in native habitats. Basic Life Sci *2*, 255-269.

Rick, C.M., and Dempsey, W.H. (1969). Position of Stigma in Relation to Fruit Setting of Tomato. Bot Gaz *130*, 180-&.

Rick, C.M., and Fobes, J.F. (1975). Allozymes of Galapagos Tomatoes - Polymorphism, Geographic Distribution, and Affinities. Evolution *29*, 443-457.

Rick, C.M., and Yoder, J.I. (1988). Classical and molecular genetics of tomato: highlights and perspectives. Annu Rev Genet *22*, 281-300.

Salathia, N., Lynn, J.R., Millar, A.J., and King, G.J. (2007). Detection and resolution of genetic loci affecting circadian period in Brassica oleracea. Theor Appl Genet *114*, 683-692.

Samach, A., and Lotan, H. (2007). The transition to flowering in tomato. Plant Biotechnology *24*, 71-82.

Sarkinen, T., Bohs, L., Olmstead, R.G., and Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evol Biol *13*, 214.

Sawa, M., Nusinow, D.A., Kay, S.A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science *318*, 261-265.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I.A., and Coupland, G. (1998). The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. Cell *93*, 1219-1229.

Schilmiller, A.L., Charbonneau, A.L., and Last, R.L. (2012). Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. Proc Natl Acad Sci U S A *109*, 16377-16382.

Schneeberger, K., and Weigel, D. (2011). Fast-forward genetics enabled by new sequencing technologies. Trends Plant Sci *16*, 282-288.

Shimada, T.L., Shimada, T., and Hara-Nishimura, I. (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. Plant J *61*, 519-528.

Shrestha, R., Gomez-Ariza, J., Brambilla, V., and Fornara, F. (2014). Molecular control of seasonal flowering in rice, arabidopsis and temperate cereals. Ann Bot.

Sim, S.C., Van Deynze, A., Stoffel, K., Douches, D.S., Zarka, D., Ganal, M.W., Chetelat, R.T., Hutton, S.F., Scott, J.W., Gardner, R.G*., et al.* (2012). High-density SNP genotyping of tomato (Solanum lycopersicum L.) reveals patterns of genetic variation due to breeding. PLoS One *7*, e45520.

Simmonds, N.W. (1976). Evolution of crop plants (London ; New York: Longman).

Simunovic, A., and Jaenike, J. (2006). Adaptive variation among Drosophila species in their circadian rhythms. Evolutionary Ecology Research *8*, 803-811.

Somers, D.E., Devlin, P.F., and Kay, S.A. (1998). Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. Science *282*, 1488-1490.

Somers, D.E., Schultz, T.F., Milnamow, M., and Kay, S.A. (2000). ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. Cell *101*, 319-329.

Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature *410*, 1116-1120.

Sulzman, F.M., Ellman, D., Fuller, C.A., Moore-Ede, M.C., and Wassmer, G. (1984). Neurospora circadian rhythms in space: a reexamination of the endogenous-exogenous question. Science *225*, 232-234.

Swarup, K., Alonso-Blanco, C., Lynn, J.R., Michaels, S.D., Amasino, R.M., Koornneef, M., and Millar, A.J. (1999). Natural allelic variation identifies new genes in the Arabidopsis circadian system. Plant J *20*, 67-77.

Tamaki, S., Matsuo, S., Wong, H.L., Yokoi, S., and Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. Science *316*, 1033-1036.

Tanksley, S.D. (1993). Mapping polygenes. Annu Rev Genet *27*, 205-233.

The Tomato Genome Consortium (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature *485*, 635-641.

Thines, B., and Harmon, F.G. (2010). Ambient temperature response establishes ELF3 as a required component of the core Arabidopsis circadian clock. Proc Natl Acad Sci U S A *107*, 3257-3262.

Thomas, B., Vince-Prue, D., and Vince-Prue, D.P.i.p. (1997). Photoperiodism in plants, 2nd ed. / Brian Thomas and Daphne Vince-Prue. edn (San Diego ; London: Academic Press).

Turner, A., Beales, J., Faure, S., Dunford, R.P., and Laurie, D.A. (2005). The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science *310*, 1031-1034.

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. Science *303*, 1003-1006.

Velez-Ramirez, A.I., van Ieperen, W., Vreugdenhil, D., and Millenaar, F.F. (2011). Plants under continuous light. Trends Plant Sci *16*, 310-318.

Verlaan, M.G., Hutton, S.F., Ibrahem, R.M., Kormelink, R., Visser, R.G., Scott, J.W., Edwards, J.D., and Bai, Y. (2013). The Tomato Yellow Leaf Curl Virus resistance genes Ty-1 and Ty-3 are allelic and code for DFDGD-class RNA-dependent RNA polymerases. PLoS Genet *9*, e1003399.

Viquez-Zamora, M., Vosman, B., van de Geest, H., Bovy, A., Visser, R.G., Finkers, R., and van Heusden, A.W. (2013). Tomato breeding in the genomics era: insights from a SNP array. BMC Genomics *14*, 354.

Voorrips, R.E., Verkerke, W., Finkers, R., Jongerius, R., and Kanne, J. (2000). Inheritance of taste components in tomato. Acta Physiol Plant *22*, 259-261.

Wang, Z.Y., and Tobin, E.M. (1998). Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. Cell *93*, 1207-1217.

Weller, J.L., Liew, L.C., Hecht, V.F., Rajandran, V., Laurie, R.E., Ridge, S., Wenden, B., Vander Schoor, J.K., Jaminon, O., Blassiau, C.*, et al.* (2012). A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proc Natl Acad Sci U S A *109*, 21158-21163.

Wenden, B., Kozma-Bognar, L., Edwards, K.D., Hall, A.J., Locke, J.C., and Millar, A.J. (2011). Light inputs shape the Arabidopsis circadian system. Plant J *66*, 480-491.

Wilczek, A.M., Roe, J.L., Knapp, M.C., Cooper, M.D., Lopez-Gallego, C., Martin, L.J., Muir, C.D., Sim, S., Walker, A., Anderson, J.*, et al.* (2009). Effects of genetic perturbation on seasonal life history plasticity. Science *323*, 930-934.

Withrow, A.P., and Withrow, R.B. (1949). Photoperiodic Chlorosis in Tomato. Plant Physiol *24*, 657-663.

Woelfle, M.A., Ouyang, Y., Phanvijhitsiri, K., and Johnson, C.H. (2004). The adaptive value of circadian clocks: an experimental assessment in cyanobacteria. Curr Biol *14*, 1481-1486.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J., and van der Knaap, E. (2008). A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science *319*, 1527-1530.

Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J.*, et al.* (2011). Genome sequence and analysis of the tuber crop potato. Nature *475*, 189-195.

Xu, Y., and Johnson, C.H. (2001). A clock- and light-regulated gene that links the circadian oscillator to LHCB gene expression. Plant Cell *13*, 1411-1425.

Yakir, E., Hilman, D., Harir, Y., and Green, R.M. (2007). Regulation of output from the plant circadian clock. Febs J *274*, 335-345.

Yanovsky, M.J., and Kay, S.A. (2002). Molecular basis of seasonal time measurement in Arabidopsis. Nature *419*, 308-312.

Yerushalmi, S., Yakir, E., and Green, R.M. (2011). Circadian clocks and adaptation in Arabidopsis. Mol Ecol *20*, 1155-1165.

Yoo, S.D., Cho, Y.H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat Protoc *2*, 1565-1572.

Zakhrabekova, S., Gough, S.P., Braumann, I., Muller, A.H., Lundqvist, J., Ahmann, K., Dockter, C., Matyszczak, I., Kurowska, M., Druka, A.*, et al.* (2012). Induced mutations in circadian clock regulator Mat-a facilitated short-season adaptation and range extension in cultivated barley. Proc Natl Acad Sci U S A *109*, 4326-4331.

Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. Nat Rev Genet *2*, 983-989.

Zamir, D., Ekstein-Michelson, I., Zakay, Y., Navot, N., Zeidan, M., Sarfatti, M., Eshed, Y., Harel, E., Pleban, T., van-Oss, H.*, et al.* (1994). Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, TY-1. Theor Appl Genet *88*, 141-146.

Zeevaart, J.A.D. (1976). Physiology of Flower Formation. Annu Rev Plant Physiol Plant Mol Biol *27*, 321-348.

Appendix

Table 4: Tomato accessions used in this thesis. Source: (i) TGRC, (ii) Enza = Enza Zaden, (iii) RF = Dr. Richard Finkers, (iv) DZ = Prof. Dr. Dani Zamir.

species	accession	cultivar	source
S. arcanum	LA2172		Enza
S. chmielewskii	LA1840		TGRC
S. chmielewskii	LA2663		Enza
S. chmielewskii	LA2695		RF
S. chilense	LA1556		RF
S. chilense	LA1558		RF
S. chilense	LA1932		Enza
S. chilense	LA2884		Enza
S. corneliomulleri	LA1274		TGRC
S. galapagense	LA0528		Enza
S. galapagense	LA0530		TGRC
S. habrochaites	LA1223		TGRC
S. habrochaites	LA1777		TGRC
S. habrochaites	LA2099		Enza
S. habrochaites	LYC4		RF
S. habrochaites f. *glabratum*	CGN1.1561		RF
S. habrochaites f. *glabratum*	CGN15792		RF
S. habrochaites f. *glabratum*	LA1718		RF
S. habrochaites f. *glabratum*	PI134418		RF
S. huaylasense	LA1365		RF
S. lycopersicum	LA2838A	Alisa Craig	DZ
S. lycopersicum		Amish Paste	Enza
S. lycopersicum	LA4451	Black Cherry	RF
S. lycopersicum	LA4449	Black Plum	Enza
S. lycopersicum		Bloody Butcher	RF
S. lycopersicum		Cherokee Purple	Enza
S. lycopersicum		Chih-Mu-Tao-Se	RF
S. lycopersicum		Cross Country	RF
S. lycopersicum		Dana	RF
S. lycopersicum		Dixy Golden Giant	RF
S. lycopersicum	LA4024	E6203	TGRC
S. lycopersicum		Galina	RF
S. lycopersicum		Gardeners Delight	RF
S. lycopersicum		Globe	Enza
S. lycopersicum	LA4355	Gold Nugget	Enza
S. lycopersicum		Heinz 1370	Enza
S. lycopersicum		Howard German	Enza
S. lycopersicum		Iidi	RF
S. lycopersicum		Katinka Cherry	RF
S. lycopersicum		Large Red Cherry	RF
S. lycopersicum	LA2706	M82	TGRC
S. lycopersicum		Marmande VFA	RF
S. lycopersicum	LA2706	Moneymaker	TGRC
S. lycopersicum		Opalka	Enza
S. lycopersicum		Paragon	Enza
S. lycopersicum		Ponderosa	Enza
S. lycopersicum		Porter	RF
S. lycopersicum	LA3229	Prospero	Enza

S. lycopersicum		Rote Beere	RF
S. lycopersicum	LA1090	Rutgers	DZ
S. lycopersicum		Watermelon Beefsteak	RF
S. lycopersicum		Wheatley's Frost Resistant	RF
S. lycopersicum		Yellow Pear	Enza
S. lycopersicum	LA0126	Latin American cultivar	Enza
S. lycopersicum	LA0146	Latin American cultivar	TGRC
S. lycopersicum	LA0468	Latin American cultivar	DZ
S. lycopersicum	LYC00196	Latin American cultivar	Enza
S. lycopersicum	LYC06710	Latin American cultivar	Enza
S. lycopersicum	LYC08804	Latin American cultivar	Enza
S. lycopersicum var. cerasiforme	LA1204		DZ
S. lycopersicum var. cerasiforme	LA1455		Enza
S. lycopersicum var. cerasiforme	LA1542		Enza
S. lycopersicum var. cerasiforme	LA1623		Enza
S. lycopersicum var. cerasiforme	LA2076		Enza
S. lycopersicum var. cerasiforme	LA2078		Enza
S. lycopersicum var. cerasiforme	LA2131		Enza
S. lycopersicum var. cerasiforme	LA2632		Enza
S. neorickii	G1.1601		RF
S. neorickii	LA1326		Enza
S. neorickii	LA2133		Enza
S. pennellii	LA0716		DZ
S. pennellii	LA1926		Enza
S. pennellii	LYC 1831		RF
S. pimpinellifolium	G1.1554		RF
S. pimpinellifolium	LA1269		Enza
S. pimpinellifolium	LA1478		Enza
S. pimpinellifolium	LA1582		Enza
S. pimpinellifolium	LA1584		Enza
S. pimpinellifolium	LA1589		TGRC
S. pimpinellifolium	LYC 2798		RF

Table 5: Genotyping markers. Allele 1 is the allele of genotype 1 and vice versa. *S. pennellii* = LA0716; *S. habrochaites* = LA1777

chromosome	type	position	forward primer	reverse primer	enzyme	fragments allele1	fragments allele2	genotype1	genotype2
7	CAPS	1,171,189	CCAAGAGTAGTTGCAGACTTTCC	AATCTCAGAAACTACCGTGGACA	EcoRI	205 & 541	746	*S. lycopersicum* cv. M82	*S. pennellii*
7	CAPS	53,752,820	GGACCAGCTGTTAAGTCTCAATG	TCCAACTGTCCAAAAAGTTGAAT	EcoRI	101 & 341 & 65	442 & 65	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	61,669,277	TGCTCTTTCATTTTCACTGTTGA	TGGGATTACTAGTGATGCGACTT	EcoRI	643	274 & 369	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,275,735	TATTGGCGCTAAAACATCATACC	ATAAGCTTTCCATATTCCTTCCG	EcoRI	270 & 40	310	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,325,922	GAAGGAACTGTGAGAAGCAACAT	TGCCTCGTCCATAAATAGAAGTC	HindIII	315 & 233	548	*S. lycopersicum* cv. M82	*S. pennellii*
9	indel	62,342,697	CCTGAGTTCATTTACCCGAAAC	TGAGTGTAAACAGCACCACCA		700	780	*S. lycopersicum* cv. M82	*S. pennellii*
9	indel	62,382,209	GTGCCACCTGAATGAGTCAA	GGGAGCTATGACCCGATTTA		470	1200	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,390,560	ACAGCCTAAAATGGCAAACAA	TGTTCACATTTGGCTTCTGG	Hpy188III	324 & 135	459	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,398,757	AACCCACAGTTATTACCAAAGCTC	ATGGCAACCTACAATGATACACC	EcoRI	197 & 126	323	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,427,490	GCACAGATTGTACACAAACCAAA	GTTTGGGGAAAATACGTACCAGT	TaqI	478	423 & 55	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,452,888	CATTTGTCCAAACTTCGCTAATC	TTCGTCTCCATAACGTTTCTGAT	TaqI	757 & 237	483 & 274 & 237	*S. lycopersicum* cv. M82	*S. pennellii*
9	indel	62,468,836	GGGTTGTGAAGTCTAGGTGCTT	GCTCAATGTGGAGAACAGCA		417	347	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,491,315	ATTGTGGACAAACCAAGAAATTG	TTCCCTCACTGCTTACGACTTAC	EcoRV	347	186 & 161	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,512,737	CATCTGCTTCAGATAGGAGAGGA	GGCTTTTAACTGAAGTCCTGGAT	PstI	570	264 & 306	*S. lycopersicum* cv. M82	*S. pennellii*
9	indel	62,582,474	TGCGTGATTGCATCTCTTTT	CTCATTTTATCTCCCACCTAGTTT		500	630	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	62,672,983	TCTTCACACTGCTCTCTACCAAA	GAGTCAAGGATCCAGAGATCAAA	EcoRI	31 & 558	589	*S. lycopersicum* cv. M82	*S. pennellii*
9	CAPS	63,036,812	TCTTGCTGTAGAACCTTCACCTC	TCAAGACATTCAACTCTTGGACA	EcoRI	596	324 & 272	*S. lycopersicum* cv. M82	*S. pennellii*
5	CAPS	62,913,970	CATTTACACATCTGCAAAATGGA	TATTGTGCCTTCTATGCATTCCT	EcoRI	453	100 & 353	*S. lycopersicum* cv. E6203	*S. habrochaites*
5	CAPS	64,355,315	AATGGAGGTTTTTGGAGGAACTA	TCCATATCATGAACTTGTTTCCC	EcoRI	564 & 295	859	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	3,015,306	TGACAAAAATTAATAAAAACAGCAA	TATTTGGGACTAACAAGCGATTC	HindIII	483	54 & 429	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	14,346,657	TCACATGGATTGTTAGATCATGC	CTTGACCGGACAGATAGTAATGC	HindIII	319 & 127 & 260	446 & 260	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	27,576,525	ATTGAGGAAAAGACTGGAGAAGG	TTCACTGGTATGTCAAGGTTTCC	HindIII	85 & 550	635	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	35,137,171	TCAAGCTGATCAAGTATCTGCAA	AAGTGTATCGAAGATGGAAGCTG	HindIII	231 & 79	310	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	42,881,234	AGGCACTTATCATTTTCCTGACA	CTTTACCTGTGAAAGAAGGGGTT	HindIII	450	383 & 67	*S. lycopersicum* cv. E6203	*S. habrochaites*

7	CAPS	48,524,796	CAGAGCGACTAAGAGTGTAGGGA	CGTTGGGATACTATGTGAAAACC	HindIII	451	273 & 178	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	53,300,792	GGCAAATGAAAAAGACAATGAAG	GAGATCATCGTTCATGGAAAGTC	HindIII	282 & 317	599	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	58,229,091	TCAGTCTTCGTCCCTATATTCCA	TGTGAATCTCCTGTTTGACAATG	HindIII	393	182 & 211	*S. lycopersicum* cv. E6203	*S. habrochaites*
7	CAPS	63,286,843	AGCTCTTCTCCTGGTTGAATACC	GAAGGATTGCCTAGATTGCTTTT	HindIII	623	165 & 458	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	50,030,979	GCATAACTCAGAGGAAAAGCAAA	TCTGCAAGCACTGTTTGTAACAT	EcoRI	459	382 & 77	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	58,661,205	CTACTACCACTTGCTTTGCACCT	TGTTCATCCTTTGGTGAAACTCT	EcoRI	450	294 & 156	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	59,816,720	TTCATATTACACTGCTTCAAGTGC	CAGTTTTGACAGATGCTCAAGTG	EcoRI	304	181 & 123	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	61,952,508	GATTTTCGTCTTCTCAACCAATG	CCTGACTTACAAGTTGCTCCCTA	EcoRI	148 & 231	379	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	64,141,413	TGTCCCATGACTTTCTGAAGATT	GATGCGTTATTGCTACTTTGGAA	EcoRI	730	212 & 518	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	66,508,158	AGGTGTTGATGCTTCTTTTTGAG	TTCTAGGGGAGAAAAGATGGAAG	HindIII	644 & 275	919	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	66,921,482	GGAAATATGGGTGTTACTGGTCA	TATGAACACCTTCCTACGCAAAT	HindIII	588	34 & 554	*S. lycopersicum* cv. E6203	*S. habrochaites*
9	CAPS	67,650,306	ATTGAAGCCAACCTAGAGAGCTT	GTCGCCCTACACCATATAACCTA	HindIII	332	241 & 91	*S. lycopersicum* cv. E6203	*S. habrochaites*
1	indel	70,610,399	AGCGTTCTGCTCTTTGGATG	GAACACCCCCTTAATTTATTCCA		153	163	*S. lycopersicum* cv. MM	*S. pimpinellifolium* (G1.1554)
1	indel	70,961,041	GTTGAGGGTAAGAAAGGGGC	CCCTCGAAGACAACAACGAT		192	180	*S. lycopersicum* cv. MM	*S. pimpinellifolium* (G1.1554)
1	indel	71,732,751	AACATCTTCTTCCAACTTACAACA	TGAGAAGTGTGGCTCTGCTG		249	257	*S. lycopersicum* cv. MM	*S. pimpinellifolium* (G1.1554)
1	indel	69,960,920	TGCATCTGGTTCTGTCTTTGTT	ATGATCACACAGCCAAACTTTG		100	103	*S. lycopersicum* cv. MM	*S. galapagense* (LA0530)
9	indel	62,390,823	TTCTGGTGTATGGAATTTGAGCC	AGATCCCACGTCTGTTTCGAT		246	249	*S. lycopersicum* cv. MM	*S. galapagense* (LA0530)

Table 6: Genotyping markers for the *S. habrochaites* ILs. Allele 1 is *S. lycopersicum*, allele 2 is *S. habrochaites*. chr = chromosome

chr	position [bp]	primer_left	primer_right	enzyme	allele1	allele2
1	1399661	TTAAAATGAAGCATAAAATTTCCA	ATGCTGTCAACGAGCAATAAAGT	EcoRI	391	100 & 291
1	11954956	AGTACGATGAATCTTTCCAATCG	TGAATTGCCACGTACAATTAGAG	EcoRI	372	225 & 147
1	35705399	AGACCCAAAAACTTGTGAAATGA	TGGGTTTTATGATGAATGCCTAC	EcoRI	718	224 & 494
1	47598377	ACACTCAGGAGTTGGAAAATGAA	GTCCAAAGACAACAACTACTGGC	EcoRI	545	343 & 202
1	50770687	GGCTGGATGATACTCTGAATTTG	ATTCTTCATATGTTTGGGGAGGT	EcoRI	128 & 493	621
1	61733150	GACCTTTGCCTCTGTGATGTACT	TCCAGTTGCCATTTTCTTACACT	EcoRI	342	219 & 123
1	75449528	CACAGACCATACTTGAGCAAACA	CAATGTCAATAGCAAAATCACCA	EcoRI	569 & 313	882
1	86699267	AGGGACATGAAGAACTCTAAGGG	ATGTGGGTTAAAGAACGAAGGTT	EcoRI	536 & 412	948
2	7826885	CATGCATGCCTATTAGTTTCACA	ACTATTCCCTCCTTACATCTCGG	EcoRI	735	586 & 149
2	11563786	TTGTGACATCAACAGAAATCTCG	ATCTTTCGGACAAGTTGGAGTTT	HindIII	645	252 & 393
2	14981764	TCCATGTCTTGATGAGTTCAATG	GATGATTGGGAGAGAGACCTGAT	EcoRI	788	606 & 182
2	21213173	TTTGATGCCCAATGTCGTT	GGATTTGAGCAGGAGAAGAAAGT	EcoRI	573 & 263	836
2	30629824	TTCCTTCCAATTACAAACAAATGA	GACGCCTGCTCCACTAGTAGTTA	EcoRI	697	301 & 396
2	33988991	TCATTTGAAGACCCTCAACTTGT	CATTCACAGGACTAAATCTTCCG	EcoRI	843	559 & 284
2	40095870	GAGGATGAGACCTATGCTTCAGA	GATAAATCTGTCCCAAACTGCTG	EcoRI	926	650 & 276
2	48088454	TAGAGGATCGTTAAGCGAATGAG	ATTTGGCTTTTGATCACTTCGTA	EcoRI	617 & 386	1003
3	7966724	AGCCCAGAAAATACAGTAAAGGC	ATCGATGAACTCTTGTTTTTGGA	EcoRI	1185	631 & 554
3	15929601	GCAATTAGACCCTGGATCTTTCT	GACACATACTTCAGGTGGGAGAC	EcoRI	455 & 424	879
3	19500566	ATGAATAGCAGTTGCAGGAAATC	AGTTGAGTTTGTGCATGATGTTG	HindIII	707	148 & 559
3	27042823	TTTTTCTACTACTCTTTCCCCGC	CGTCTGTATTGTGAAGCAGATCA	EcoRI	156 & 223	379
3	31763457	TGTGGTTTGCAAGATACTTTACG	TTACTGACATGCAATGAATACCC	EcoRI	319 & 151	470
3	43603720	ATGATGGAGAGATCACTATCGGA	GCATTAGGGATGATGGTGAATTA	EcoRI	507 & 125	409 & 98 & 125
3	52913624	TTGCTTATTTAGCCTTTGTTCCA	AAAGCACTCTGCCTCATAGTCAC	EcoRI	216 & 575	791
3	61668044	ATGCAGTTATCCAGGGGTTATTT	ATTACACGCGAAAGTACATCCAT	EcoRI	421 & 448	869
4	3083430	GCTGACTTGGTTTTTATTCATGG	CAGGGTCATTCAAGACCTCATAG	EcoRI	1068	463 & 605
4	7035506	TACCCATGTACCTTCAGAGAGGA	CTTTTGCATGAGAACTGGCTACT	EcoRI	686	337 & 349
4	23184114	GAGCCTGTAATTTTGATTGATCG	GTACCCATGTTCAGCAGCTATTC	HindIII	753	489 & 264
4	25135478	AATGAGTATGACTGCGGTCTGTT	TGGATGGGTCAATCAGTTCTACT	EcoRI	256 & 222	478
4	31078440	TGATACAGCAGTTGTTTTTCCCT	AACTTGATGACAGTTCGTTTTCG	HindIII	863	256 & 607

4	42790264	ATGAGAATCCTTGACC ACTGTGT	TGAACCATTCAGGGAA GACTCTA	EcoRI	26 & 109 & 419	554
4	48129395	CCTGTTGGTTTGACTC ATTTTTC	TTTTCAACTGCAAGGT TTCGTAT	EcoRI	769	363 & 406
4	56747161	GAGAATCTAGGCAGCT TTGATGA	AAATTTTTGCATGCTT TGTTTGT	EcoRI	497 & 507	1004
5	8814794	ACCCAAAAACACCTCT ACCATTT	AGACGACACATTGCTC CCTATTA	EcoRI	538	234 & 304
5	13143150	TGGATATCATCGCTTA TTGCTCT	GGCATTTGATAAAGAT TCAGTCG	HindIII	649 & 591	1240
5	20382207	GTGAGCTCCAAAACAA CAATTTC	TCCTGTATTTTCACCT GCTTCAT	HindIII	219 & 642	861
5	28878474	GACCCACTTCACCTGT ACTTGTC	TTATTAGGCCGTGTTA CTGTCGT	EcoRI	151 & 198	349
5	36208713	GCACAAAACAGCAAAT ACATCAA	GCCTTGTCCCTGTACT CTATGTC	HindIII	110 & 216	326
5	50849922	GAACCCTGAACTCGAA GATTTTG	AAAGGGTTTGACCTGG TACATTT	EcoRI	468	292 & 176
5	56381938	CGGTAGAACTGGACAA AGAAGAA	GTAAGTGGAAGGACAC AACTTGG	EcoRI	812	325 & 487
5	59879995	ACGAGGGAGATATAGC ATTAGCC	GTGCAAGGAAAATGTT CTCAAAC	EcoRI	1020	570 & 450
5	62913970	CATTTACACATCTGCA AAATGGA	TATTGTGCCTTCTATG CATTCCT	EcoRI	100 & 353	453
6	4696548	TGGACTTAAAGTTGGG TTGAAGA	AGTACAAGATCCAATG CCAGTGT	EcoRI	475 & 186	661
6	12039159	CTTCCTCTGGGGATGT TCTTAGT	CAAAGCCCAGTAATCT TCCTTTT	HindIII	335 & 266	601
6	21357280	AAGAAGGCTAACCATC TGTTTCC	CCAAGATAGTCACTTT GGTGGAG	EcoRI	484	187 & 297
6	25909775	GCTGCATTTGCTTCTA TATGTCC	ATCGTTACTTTTTGCG GATGTAA	EcoRI	402	189 & 213
6	28524434	ATCAACTTGGCAGGTC TTTATCA	GCTATCAATTAGGACA TTGCTGG	EcoRI	1010	545 & 465
6	34675312	ACAGATTGTCTAAAAG CAACCCA	TTGCTTTCTCCGCTGA TATTAAG	EcoRI	894	494 & 400
6	38853889	GGATGAATTAACAGCA TTTCAGC	AATGCTAGGGAAAGGA AGTTTTG	EcoRI	379 & 425	804
6	43821858	TAGTGGTTCAAATCAT CTTGGCT	AGTTTGCTAGCCAATT CTCTCCT	EcoRI	242 & 474	716
7	6517044	TAAGTTACGCAAGTTG GAGGAAG	TATGCAAGTTGGAACA TGTTGAG	HindIII	850	549 & 301
7	11881401	TGGAAAGTTAGTGCAA CAGGAGT	CATCAGAGTATTCAAC ACAGCGA	HindIII	185 & 404	185 & 268 & 136
7	27576525	ATTGAGGAAAAGACTG GAGAAGG	TTCACTGGTATGTCAA GGTTTCC	HindIII	635	85 & 550
7	33218683	TACTTATCCAATACCA CCCATCG	TGCTGGTTCTCATTCC TAATCAT	HindIII	950	322 & 628
7	35825601	CGGAGAGAGAAGTGAC TGAGAAG	GCGTCTAGATCATCAC AAGGTTT	EcoRI	207 & 219	426
7	41547259	TGGAAAGGAACTGTTA AATGAGC	TGAAAATTAATTGTTG GGTTGCT	HindIII	727	303 & 424
7	48524796	CAGAGCGACTAAGAGT GTAGGGA	CGTTGGGATACTATGT GAAAACC	HindIII	273 & 178	451
7	54343101	AGTGTTGTTGAAGAGC TGAAAGG	TTCTGCTTCAATTTCT CCTCAAG	HindIII	963	482 & 481
8	230734	AAGGCTAGGAAAAGTG TGAGCTT	CTTTCCTCACTTTCCTT TTCCAT	EcoRI	1208	603 & 605
8	6901615	CTCATTTCAAAGAGCA AAGAGGA	TTACCGTATGAAACTC TGGGCTA	EcoRI	304 & 571	875
8	12013918	TCAAGGTTCTCTGGTG ATGAAGT	CCACTTCGTCCTCTAA AATGATG	EcoRI	612 & 352	964
8	16846213	CTTTATCCACTCGGAA TCCTCTT	TTGAAATTCTTTTATT GGGCAGA	EcoRI	685	541 & 144

8	39943720	TACAACTTGGCCTTGGAAACTAA	TTCTCTTTTGATGGTTCTTGCTC	HindIII	232 & 641	873
8	46496449	GAAACCCTCTCCATTTCTTTGAT	AGGGAGTAGGAATTTGAGTGAGG	EcoRI	122 & 329	451
8	54003948	TTAAGACTTTTGCATCCCATGTT	AACTATCCATTGGGCTAGAGGAG	EcoRI	490	367 & 123
8	59116675	ACCAATGAAGTTTGAGAGGACAA	CATATGAATGGGAAAACCAAGAA	EcoRI	623 & 340	963
9	636902	TCAGGATGTATATCTTCGGCACT	TGAGTTTCAGGATGACCAATCTT	EcoRI	678	493 & 185
9	6037134	CACATAAATTACAGCAGATGGCA	AGCTTTGTATGAGCTTACGTTGC	EcoRI	491 & 474	965
9	11831070	TGGAGAAGGATCTTTACTTGCAG	GACCCTGTTTTCTGTTTCCTTCT	EcoRI	260 & 265 & 254	525 & 254
9	24425107	CCAGCCCACTATCACTATCAAAC	GTGTGGAAAACACAAAATCCAAT	EcoRI	213 & 217	430
9	30609146	AACGAGTGTTCATGTTACTGCCT	GACATTGACAATTGAAAGCATCA	EcoRI	370	43 & 327
9	40020071	TGAAAATTATGCTTCATGGGTCT	ACTCTATTGATTTCCCCCTCAAG	EcoRI	846	225 & 621
9	56535705	CTGGAATCGATTTGAATTTTCTG	CTTCTTGCTTTGCTCTCTGCTAC	EcoRI	591 & 330	921
9	62437250	CAAACTGACTTGAAAAACTCGCT	TGGACCACCTTGAGTATTAGCAT	EcoRI	702	446 & 256
10	5081837	AAAGTCGATCATACCAGCTCAAA	TTCTACCGCGTGAAATTAAGTGT	EcoRI	200 & 353 & 259	553 & 259
10	17828697	GAAATTGGCAATCAAAGTTGAAG	GATCTTGGAAAGTGTTTGAGGTG	EcoRI	294 & 437	182 & 549
10	23758217	AATGTCAAGGCTTCCATTCATAA	GAAGCAGAATTTGCAGTTCCTAA	EcoRI	331 & 125	456
10	28656702	GCACCCTGTAAATCTCCTTTCTT	GTGATGATCTCACTGATGCATGT	HindIII	236 & 233	469
10	35689747	AAGAACCGTACCTCGATTTATGA	TATGCTGCAACAACTTGAGAAAA	EcoRI	648	219 & 429
10	38709013	CGTGTTTACACAATAACTCAACCA	TGCAATCAAGAAGTTTGAACAAG	HindIII	646	239 & 407
10	45229743	CTTTTCTGGGAACACTTGTTTGT	CTTAGATCACTCCACAAGCCATC	EcoRI	1030	499 & 531
10	57606124	GAAAGATGTTGACCGAAGATGAC	TGCTTGTGTAAGCAGAACAAAAA	EcoRI	955	456 & 499
11	6019288	CTTGTGACCTTAGTCGAAGAGGA	ATTGTGGTCCAAATAGAAGCAGA	HindIII	416	276 & 140
11	10019927	GAATCCAAACTGAAAGAGCTCAA	CATGAGACCATTATAAATGGGGA	HindIII	643 & 345	988
11	15608283	CTATTAGGGGTATAGGCCGTGTC	GAAGCTGTCAAATGAGGCTTCTA	HindIII	532 & 617	1149
11	22978185	TCAGCGGTCTACCTCTAGAAATG	AAGCTTTCCCGATCTTTTGTTAC	EcoRI	51 & 228 & 159	51 & 133 & 95 & 159
11	30326400	CACATTCGGAAGCATTTTCTGTA	TAAGTTCACATTGCCAGATCTCA	HindIII	115 & 185	300
11	36953170	AATTGAATGGATTTACGGCTTTT	AAGTTATTTAGCGCGATTTCCAT	HindIII	375	250 & 125
11	47595305	GTGAAGATGAAGATGAAGATGGA	AGCAAGCTCAGCATCAATTAGTC	HindIII	107 & 240	347
11	52040340	CCATGCCCAGAAGTACTACAGTC	TGCTGGTACGATGATTAGGTCTT	HindIII	373	180 & 193
12	10410907	TCTGTTTGAACTCTTCCCAGGTA	CCAAAGTGTACTTGGTGGTTAGC	HindIII	876	600 & 276
12	24687848	ATACCAATTTCATTTTGGGACCT	TTTCTCTCTCAGCAGAATCCAAC	HindIII	853	579 & 274
12	28841269	AATGCCAATTTTACAGAAGACCA	TGAATAAGCAAATGCCAAATCTT	HindIII	474	64 & 410
12	35935736	ACATTGCTGACAAAACCTCTGTT	TTCTGTTGATGATGGAATAACCC	HindIII	355	217 & 138

12	44076980	CCATTTGCCAAGAACTATGAAAG	TCGAGGTTTCAAGTCCATCTTTA	HindIII	1015	491 & 524
12	51463785	GTCCCTGGGTATGTTGTTGTTTA	ATTTTTGCACCAAACACACTCTT	HindIII	301 & 110	411
12	54321054	AAATCCTCTGAAAATGCAAATCA	TTGGACTTGGAAGGTTTTCACTA	EcoRI	300 & 109	409
12	62380237	TACTCCGACACAGAGGGTAACAT	TCCGCAGATATGGAAGAACTTTA	HindIII	106 & 555	661

Table 7: Primers for qRT-PCR. The primers were tested with *S. lycopersicum* cv. M82, *S. pimpinellifolium* (LA1589) and *S. pennellii* (LA0716)

gene	left primer	right primer
GI (Solyc12g05660)	CATCTGGTCTGCGTCTTCCTA	AAGTGTCGCCACCCAGAAT
LHY (Solyc10g005080)	TTTACAAAGTTAGAAAAGGAGGCTCT	TAAGTTCCTTTCCTCACAGATGG
CAC (Solyc08g006960)	CCTCCGTTGTGATGTAACTGG	ATTGGTGGAAAGTAACATCATCG

Table 8: Primers for cloning of *EID1*. The colored parts of the primers represent the specific *att*-sites needed for gateway cloning

	species	primer	sequence
***EID1* gene**	*S. lycopersicum*	left	GGGGACAACTTTGTATACAAAAGTTGTGAGCAAGGGGTTTGATTCTA
		right	GGGGACCACTTTGTACAAGAAAGCTGGGTCAGTTCAAGTTAACACGATTACCC
	S. pennellii	left	GGGGACAACTTTGTATACAAAAGTTGATTGGATTGAGCAAGGGTTTT
		right	GGGGACCACTTTGTACAAGAAAGCTGGGTTTAACATGATTACCCCATAATTGGT
***EID1* promoter**	*S. lycopersicum*	left	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTAGATGGCAGCTTTGGGTCT
		right	GGGGACAACTTTTGTATACAAAGTTGCCAATGGAATTAAACACCAAA
	S. pennellii	left	GGGGACAAGTTTGTACAAAAAAGCAGGCTGAATTTTTCTTCTGGAAATCGTG
		right	GGGGACAACTTTTGTATACAAAGTTGAAACACCAAATTTATCAGCAAAAA

Table 9: Circadian parameters of all tomato accessions. per = period, peak = peak time, pha = (circadian time) phase, amp = amplitude, RAE = relative amplitude error.

species	accession	cultivar	per	peak	pha	amp	RAE
S. arcanum	LA2172		22.87	39.79	6.23	3.15	0.09
S. chmielewskii	LA1840		22.14	38.99	5.73	3.38	0.06
S. chmielewskii	LA2663		22.40	37.25	8.08	2.30	0.08
S. chmielewskii	LA2695		22.62	38.14	7.53	2.68	0.07
S. chilense	LA1556		23.53	39.15	8.06	3.10	0.08
S. chilense	LA1558		23.34	40.14	6.72	4.93	0.07
S. chilense	LA1932		23.85	41.14	6.56	2.85	0.12
S. chilense	LA2884		20.95	38.42	3.97	3.06	0.11
S. corneliomulleri	LA1274		22.79	40.96	4.85	2.89	0.10
S. galapagense	LA0528		24.17	41.86	6.43	3.47	0.08
S. galapagense	LA0530		24.37	41.69	6.94	3.57	0.07
S. habrochaites	LA1223		24.86	41.02	8.39	3.32	0.08

S. habrochaites	LA1777		22.52	40.14	5.21	2.86	0.11
S. habrochaites	LA2099		23.59	40.27	7.03	2.94	0.10
S. habrochaites	LYC4		25.83	42.85	8.16	2.55	0.11
S. habrochaites f. *glabratum*	CGN1.1561		26.24	44.30	7.45	2.85	0.09
S. habrochaites f. *glabratum*	CGN15792		24.08	42.62	5.51	2.31	0.10
S. habrochaites f. *glabratum*	LA1718		23.08	41.15	5.21	3.10	0.08
S. habrochaites f. *glabratum*	PI134418		23.30	40.99	5.78	3.81	0.09
S. huaylasense	LA1365		23.94	40.33	7.56	3.20	0.09
S. lycopersicum	LA2838A	Alisa Craig	26.53	47.94	4.62	5.66	0.10
S. lycopersicum		Amish Paste	24.85	44.00	5.50	6.19	0.10
S. lycopersicum	LA4451	Black Cherry	26.46	47.00	5.35	6.95	0.07
S. lycopersicum	LA4449	Black Plum	27.78	49.17	5.49	3.70	0.11
S. lycopersicum		Bloody Butcher	25.62	48.67	2.40	3.94	0.07
S. lycopersicum		Cherokee Purple	26.99	44.21	8.62	4.22	0.10
S. lycopersicum		Chih-Mu-Tao-Se	27.04	45.49	7.62	7.71	0.08
S. lycopersicum		Cross Country	26.83	45.89	6.94	6.76	0.07
S. lycopersicum		Dana	25.82	47.91	3.40	4.82	0.07
S. lycopersicum		Dixy Golden Giant	25.44	44.05	6.43	5.67	0.09
S. lycopersicum	LA4024	E6203	25.88	43.91	7.21	5.24	0.13
S. lycopersicum		Galina	26.17	45.80	5.96	4.64	0.10
S. lycopersicum		Gardeners Delight	27.21	47.75	5.88	4.63	0.08
S. lycopersicum		Globe	25.65	46.31	4.65	5.02	0.13
S. lycopersicum	LA4355	Gold Nugget	26.76	45.77	6.93	3.87	0.11
S. lycopersicum		Heinz 1370	25.56	45.03	5.70	8.12	0.06
S. lycopersicum		Howard German	26.70	43.03	9.27	6.09	0.14
S. lycopersicum		Iidi	26.04	46.83	4.84	4.79	0.07
S. lycopersicum		Katinka Cherry	27.63	46.59	7.52	3.21	0.07
S. lycopersicum		Large Red Cherry	25.79	46.64	4.59	7.07	0.07
S. lycopersicum	LA2706	M82	27.26	46.08	7.41	5.06	0.11
S. lycopersicum		Marmande VFA	27.20	45.32	7.99	3.20	0.11
S. lycopersicum	LA2706	Moneymaker	25.97	45.99	5.47	6.75	0.07
S. lycopersicum		Opalka	26.41	46.32	5.91	2.64	0.10
S. lycopersicum		Paragon	24.60	44.56	4.50	4.59	0.12
S. lycopersicum		Ponderosa	27.05	46.85	6.42	4.38	0.09
S. lycopersicum		Porter	24.94	43.93	5.72	5.54	0.10
S. lycopersicum	LA3229	Prospero	26.52	46.14	6.25	5.10	0.09
S. lycopersicum		Rote Beere	25.99	45.40	6.06	8.00	0.05
S. lycopersicum	LA1090	Rutgers	25.77	46.81	4.40	7.31	0.09
S. lycopersicum		Watermelon Beefsteak	27.07	45.77	7.42	6.22	0.12
S. lycopersicum		Wheatley's Frost Resistant	25.16	45.96	4.12	4.17	0.09
S. lycopersicum		Yellow Pear	27.36	47.95	5.93	7.31	0.09
S. lycopersicum	LA0126	Latin American	26.26	44.60	7.23	8.12	0.06
S. lycopersicum	LA0146	Latin American	25.69	42.64	8.15	6.59	0.08
S. lycopersicum	LA0468	Latin American	24.83	43.46	6.00	13.17	0.08
S. lycopersicum	LYC00196	Latin American	25.56	44.16	6.54	7.31	0.10

S. lycopersicum	LYC06710	Latin American	24.96	41.59	7.94	5.28	0.16
S. lycopersicum	LYC08804	Latin American	25.38	42.80	7.53	8.14	0.08
S. lycopersicum var. cerasiforme	LA1204		24.05	42.62	5.44	7.44	0.12
S. lycopersicum var. cerasiforme	LA1455		25.68	44.94	5.98	3.75	0.07
S. lycopersicum var. cerasiforme	LA1542		25.36	44.39	5.96	3.50	0.13
S. lycopersicum var. cerasiforme	LA1623		25.04	43.62	6.19	4.60	0.06
S. lycopersicum var. cerasiforme	LA2076		26.39	46.37	5.83	5.46	0.09
S. lycopersicum var. cerasiforme	LA2078		26.93	46.88	6.21	8.39	0.05
S. lycopersicum var. cerasiforme	LA2131		26.14	44.21	7.39	8.27	0.06
S. lycopersicum var. cerasiforme	LA2632		24.84	43.39	6.08	4.74	0.10
S. neorickii	G1.1601		23.16	39.74	6.82	3.39	0.08
S. neorickii	LA1326		23.45	40.70	6.34	2.65	0.08
S. neorickii	LA2133		23.76	40.61	6.98	5.61	0.06
S. pennellii	LA0716		22.72	37.90	7.96	3.79	0.06
S. pennellii	LA1926		23.90	40.42	7.40	3.27	0.07
S. pennellii	LYC 1831		23.35	40.53	6.33	4.32	0.09
S. pimpinellifolium	G1.1554		24.92	40.65	8.84	6.47	0.07
S. pimpinellifolium	LA1269		24.58	41.22	7.74	7.44	0.08
S. pimpinellifolium	LA1478		24.68	41.01	8.10	6.06	0.08
S. *pimpinellifolium*	LA1582		24.86	40.90	8.51	6.76	0.07
S. pimpinellifolium	LA1584		24.37	41.96	6.68	6.69	0.08
S. pimpinellifolium	LA1589		24.33	40.66	7.87	7.28	0.06
S. pimpinellifolium	LYC 2798		24.88	41.48	7.99	6.90	0.06

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Maarten Koornneef betreut worden.

Köln, März 2014_____
Niels Müller

Lebenslauf

Persönliche Daten

Name:	Niels Müller
Geburtsort:	Heidelberg, Deutschland
Geburtsdatum:	8. Juli 1983

Berufliche Tätigkeiten

03/2010- 12/2010 Wissenschaftlicher Mitarbeiter am Staatlichen Weinbauinstitut, Freiburg im Breisgau
Projekt: Interaktionen zwischen Reblaus und Weinrebe

Ausbildung

seit 01/2011: Doktorarbeit am Max-Planck-Institut für Pflanzenzüchtungsforschung Köln, AG Dr. José Jiménez-Gómez, Abteilung Prof. Dr. Maarten Koornneef
Thema: „Natural variation in circadian rhythms and photoperiodic flowering in tomato and its wild relatives“

02/2010: Diplom in Biologie an der Albert-Ludwigs-Universität Freiburg im Breisgau

Hauptfach:	Botanik
Nebenfächer:	Genetik, Biochemie und Landespflege
Diplomarbeit:	„Verwilderte Reben an den Böschungen des Kaiserstuhls“

09/2006 – 03/2007 Auslandsstudium an der Universidad Autónoma de Madrid, Spanien, im Rahmen des ERASMUS Programms

10/2005: Vordiplom in Biologie an der Universität Kassel

07/2003 Abitur an der Elisabeth von Thadden Schule in Heidelberg

Acknowledgments

First, I would like to express my gratitude to my supervisor Dr. José Jiménez-Gómez for giving me the opportunity to do my PhD in his group in the first place, for his support throughout my thesis, for numerous inspiring discussions and for many helpful comments. He always had an open door and made time for my questions.

Further, my sincere thanks go to Prof. Dr. Maarten Koornneef for his support of my project, for his encouragement and for valuable advice. Also I want to thank him for the great working environment and the motivating atmosphere in his department, which are in large part his credit.

Besides, I would like to thank the rest of my thesis committee, Prof. Dr. Ute Höcker, Prof. Dr. Siegfried Roth and Dr. Wim Soppe.

Also, I would like to express my greatest appreciation to many of my colleagues from the MPIPZ for plenty of though provoking discussions and their readiness to help, but also for the nice time outside of work. Namely I want to thank Ute Tartler and Cris Wijnen for technical support.

I am also deeply grateful to my parents who supported me throughout my life and made possible the great years of my study.

Finally, I want to deeply thank my beloved girlfriend and little daughter for keeping me from overestimating the importance of my work and reminding me every day of what really matters in life.