Independent FLC mutations as causes of flowering time variation in Arabidopsis thaliana and Capsella rubella

Ya-Long Guo*, Marco Todesco*, Jörg Hagmann*, Sandip Das*1 and Detlef Weigel*2

*Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany and §State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

1Present address: Department of Biotechnology, University of Delhi, Delhi-110007, India

1Corresponding author: MPI for Developmental Biology, Spemannstraße 35, D-72076 Tübingen, Germany. E-mail weigel@weigelworld.org

DNA sequences have been deposited in GenBank (JQ992992-JQ993051, JX003246-JX003248); short read data have been deposited in European Nucleotide Archive (ENA) (ERP001651). Sequence variants have been deposited with Dryad (http://datadryad.org), doi:10.5061/dryad.4rq51.
Running Title: Flowering time in Capsella rubella

Keywords: Flowering time, mapping-by-sequencing, Arabidopsis thaliana, Capsella rubella, FLC

Corresponding author:

Detlef Weigel

Department of Molecular Biology
Max Planck Institute for Developmental Biology
Spemannstraße 35
72076 Tübingen
Germany

Ph +49 7071 601 1411
Fx +49 7071 601 1412
Em weigel@weigelworld.org
ABSTRACT

*Capsella rubella* is an inbreeding annual forb closely related to *Arabidopsis thaliana*, a model species widely used for studying natural variation in adaptive traits such as flowering time. Although mutations in dozens of genes can affect flowering of *A. thaliana* in the laboratory, only a handful of such genes vary in natural populations. Chief among these are *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*). Common and rare *FRI* mutations along with rare *FLC* mutations explain a large fraction of flowering time variation in *A. thaliana*. Here we document flowering time under different conditions in 20 *C. rubella* accessions from across the species’ range. Similar to *A. thaliana*, vernalization, long photoperiods and elevated ambient temperature generally promote flowering. In this collection of *C. rubella* accessions, we did not find any obvious loss-of-function *FRI* alleles. Using mapping-by-sequencing with two strains that have contrasting flowering behaviors, we identified a splice-site mutation in *FLC* as the likely cause of early flowering in accession 1408. However, other similarly early *C. rubella* accessions did not share this mutation. We conclude that the genetic basis of flowering time variation in *C. rubella* is complex, despite this very young species having undergone an extreme genetic bottleneck when it split from *C. grandiflora* a few tens of thousands of years ago.
INTRODUCTION

For two decades, Arabidopsis thaliana has been the preeminent model for mechanistic studies of many aspects of plant development and physiology. In addition, it is extensively used to investigate the genetic basis of natural variation for adaptive traits such as the onset of flowering (Bergelson and Roux 2010; Koornneef and Meinke 2010; Weigel 2012). As in other species, flowering of A. thaliana is influenced by prolonged exposure to cold, which signals winter, by ambient temperature and long days, both of which indicate spring, as well as the age of the plant. An indication of the complexity of its regulation is that well over a hundred genes affecting flowering time in the laboratory have been identified through mutant analysis (Srikanth and Schmid 2011).

Arabidopsis thaliana is found throughout much of the Northern hemisphere, with a native range that extends from North Africa to the Arctic Circle, and from the Atlantic coast of Western Europe to Central Asia. Accordingly, A. thaliana accessions show a wide range of flowering time behaviors when grown in climate chambers or in common gardens (Brachi et al. 2010; Gazzani et al. 2003; Hancock et al. 2011; Lempe et al. 2005; Li et al. 2010; Mendoza-Vigo et al. 2011; Strange et al. 2011; Werner et al. 2005). In contrast to the multitude of laboratory-induced mutations that change flowering behavior, genetic analysis has identified only a handful of loci responsible for flowering variation among natural accessions. Chief among these are FRIGIDA (FRI) and FLOWERING LOCUS C (FLC), with FRI repressing flowering by allowing expression
of FLC. Vernalization stably reduces the FLC transcription, thus reversing the action of FRI. Together, allelic differences at FRI and FLC can account for more than half of flowering time variation when A. thaliana accessions are grown in constant, strongly flower-promoting conditions and without vernalization (Atwell et al. 2010; Brachi et al. 2010; Gazzani et al. 2003; Johanson et al. 2000; Le Corre et al. 2002; Lempe et al. 2005; Li et al. 2010; Mendoza-Vigo et al. 2011; Michaels et al. 2003; Salomé et al. 2011; Shindo et al. 2005; Strange et al. 2011; Werner et al. 2005).

The ancestors of the two extant genera Arabidopsis and Capsella separated from each other about twice as long ago as A. thaliana split from other species in the genus Arabidopsis (Koch and Kiefer 2005). Reduced genetic diversity and the pattern of allele sharing suggest that Capsella rubella was founded about 30,000 to 50,000 years ago when a single C. grandiflora individual became self-compatible. The general picture is that most C. rubella alleles can be easily found in the C. grandiflora population, and that loci typically have only one or two different major alleles (Fuxe et al. 2009; Guo et al. 2009). Compared with C. grandiflora, C. rubella has a much wider distribution (Hurka and Neuffer 1997), which raises the question of how C. rubella could adapt so seemingly rapidly to a much wider range of environmental conditions than encountered by the genetically much more diverse C. grandiflora. In addition to new mutations, recombination could expose new epistatic
interactions, allowing selection to act on standing variation (BARRETT and SCHLUTER 2008).

To address the extent and genetic basis of variation in a model adaptive trait, we have examined flowering time variation in a set of C. rubella accessions. Using mapping-by-sequencing, we have identified a rare mutation in an FLC homolog as being responsible for rapid flowering in one particularly early accession. We demonstrate by transformation in A. thaliana that a splice site mutation greatly reduces the flowering-repressing activity of this allele.

MATERIALS AND METHODS

**Plant material:** The C. rubella accessions studied (Table 1) have been described (Guo et al. 2009); they included the reference accessions MTE, for which a draft genome sequence has been produced by the US Department of Energy Joint Genome Institute (DOE-JGI). The A. thaliana strain in the Col-0 background with the FRI^{fs2} and flc-3 alleles has been described (LEE et al. 1993; MICHAELS and AMASINO 1999).

For flowering time measurements, siblings from the same maternal family were used. Seeds were stratified for seven days at 4°C in 0.1% agar. For each treatment, we sowed six plants, most of which survived. Trays were moved to random positions in the growth rooms every two days to reduce positional effects. Flowering time was measured as total leaf number (TLN) and days to flowering (DTF).
**Mapping-by-sequencing:** Early-flowering F$_2$ plants from a cross of accession 1408 and the reference accession MTE were selected, and pooled genomic DNA was sequenced to 30-fold coverage on the Illumina GenomeAnalyzer IIx platform with 101 bp paired-end reads (59 million alignable reads, 11 Gb of sequence). SHORE (Ossowski et al. 2008) was used to discover single-nucleotide polymorphisms (SNPs) in the pool by alignment to a preliminary assembly of the MTE genome (CR_stitch_Feb15th.fasta) kindly provided by Jeremy Schmutz (DOE-JGI and HudsonAlpha), Stephen Wright and Khaled Hazzouri (University of Toronto), and Adrian Platts (McGill University). For annotation of variants, we used the Capsella v1.0 gene set released with Phytozome v8.0 (http://www.phytozome.net/capsella/). Using SHORE's scoring matrix approach optimized for heterozygous SNP detection and stringent filtering (requiring uniqueness of reads, minimum 10x coverage, minimum 20% allele frequency, SHORE SNP quality score > 25), 354,080 SNPs could be identified (Table S2). These SNPs were used as markers in SHOREmap (Schneeberger et al. 2009) to identify genomic regions with an excess of homozygous 1408 alleles. The final mapping interval was defined as having an allele frequency of at least 90% on both sides for 40 contiguous SNPs, and fewer than 10 contiguous SNPs with allele frequency below 50% within this region.

**Targeted DNA sequence analysis:** Individual young leaves were collected for DNA extraction using the cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987), followed by polymerase chain reaction (PCR)
amplification with Pfu polymerase (Fermentas, St. Leon-Rot, Germany). Products from two independent reactions were mixed and directly Sanger sequenced on an ABI3730XL (Applied Biosystems, Foster City, CA). Primers for FRI were designed based on sequence of a C. rubella BAC (GenBank accession number JX003248), and a 2.3 kb fragment covering the entire transcription unit was sequenced. Primers for FLC were designed based on A. thaliana genome sequence, and a 5.8 kb fragment encompassing most of the first exon through the 3’ UTR was sequenced. Sequences were assembled and inspected with Lasergene Seqman (DNASTAR, Madison, WI). See Table S1 for oligonucleotide primers.

Sequences were aligned using ClustalX version 1.81 (Thompson et al. 1997), and alignments were refined manually. All polymorphic sites were individually confirmed again based on the original traces. DnaSP version 4.10.9 (Rozas et al. 2003) was used to perform population analyses: levels of nucleotide diversity per site ($\pi$) (Nei 1987).

Expression analysis: RNA was extracted from two-week old plants with the Plant RNeasy kit (Qiagen, Hilden, Germany) or with the TRizol reagent (Invitrogen, Carlsbad, CA, USA) and treated with RNase-free DNase I (Fermentas). Two to four $\mu$g RNA were reverse transcribed using RevertAid First-strand cDNA Synthesis kit (Fermentas). PCRs were carried on in the presence of SYBR Green (Invitrogen, Carlsbad, CA, USA) in 20 $\mu$l reactions. Amplification was monitored in real time with the Opticon Continuous Fluorescence Detection
System (MJ Research, Reno, NV). Threshold cycles (cT) were based on a reaction reaching a specific fluorescence intensity in the log-linear phase of the amplification curve.

For the analysis of splice forms, a fragment of about 200 bp was amplified by PCR from pooled complementary DNA (cDNA) obtained from four primary transformants for each transgenic line, and digested for 2 hours with 1 µl of Fast Digest BstXI (Fermentas). See Table S1 for oligonucleotide primers.

Transgenic lines: FLC genomic fragments covering the entire open reading frame and introns were amplified using Phusion polymerase (Thermo Fisher Scientific, Waltham, MA, USA). These fragments were placed under the control of the constitutive CaMV 35S promoter in the pGREEN-derived (HELLENS et al. 2000) vector pFK210, and introduced into different A. thaliana genotypes by Agrobacterium tumefaciens-mediated transformation (WEIGEL and GLAZEBROOK 2002).

Statistical analyses: Mean, median, and standard deviation were calculated using Microsoft Excel (Microsoft, Redmond, WA). Differences between groups in Fig. 7B were tested by ANOVA using Tukey’s test, as implemented in Excel. Spearman’s rank-order correlation coefficient $r_s$ was determined in R (http://www.r-project.org/).
RESULTS

Flowering time variation in *C. rubella*: We analyzed the response of *C. rubella* accessions to long day (16 hour light) and short day (8 hour light) photoperiods, and to vernalization for six weeks at 4°C in short day photoperiods, after seeds had been germinated on soil in 16°C. Twenty accessions were grown in 16°C long days without vernalization (16LD), and in 16°C long days with vernalization (16LDv) (Table 1). A subset of eight accessions was grown in two 23°C environments with contrasting photoperiods, long days (23LD), and short days (23SD) (Table 1). As in *A. thaliana* (Gazzani et al. 2003; Koornneef et al. 1991; Lempe et al. 2005), days to flowering (DTF) and total rosette leaf number (TLN) were correlated (16LD, \( r_s = 0.62, p = 0.003 \); 16LDv, \( r_s = 0.80, p \ll 0.001 \)). Below, we report only DTF data.

Accessions flowered on average much earlier after vernalization (Figure 1). In 16LD, onset of flowering ranged from 51 to 103 days (mean, 69; median, 71). In 16 LDv, it ranged from 28 to 51 days (mean, 36; median, 33). While all accessions responded to vernalization, there was considerable variation in the extent of the response, as seen most clearly in a reaction norm representation (Figure 2A). Late-flowering accessions tended to respond more strongly to vernalization than early-flowering ones, but DTF variation in 16LD explained about 17% of DTF variation in 16LDv (Figure 2B). In addition to vernalization, both longer photoperiods and increased ambient temperature tended to accelerate flowering (Table 1).
**FLC underlies a major effect locus for flowering time:** The reference accession MTE, from Italy, was the last accession to flower in 16LD, after 103 days on average, which was twice as late as accession 1408, from Greece (Table 1). Following vernalization in 16LDv, MTE was still late, flowering after 43 days, but it was no longer the last accession to flower. A robust vernalization response was also seen in accession 1408, which flowered in 16LDv after 29 days, as did several other accessions.

To begin to investigate the genetic basis of flowering time variation in *C. rubella*, we crossed the late-flowering accession MTE with the early-flowering accession 1408. In an F₂ population, about a quarter of plants flowered early, similar to 1408, and three quarters flowered late, similar to MTE (Figure 3), suggesting that a major, recessively acting gene controlled early flowering of accession 1408.

MTE is the reference accession for the ongoing *C. rubella* genome sequencing effort (collaboration with Jeremy SCHMUTZ from the Joint Genome Institute of the US Department of Energy and the HudsonAlpha Institute for Biotechnology, Stephen WRIGHT and Khaled HAZZOURI of the University of Toronto, and Adrian PLATTS of McGill University). The MTE draft genome sequence in turn enables the discovery of polymorphisms between MTE and 1408, including polymorphisms that are potentially causal for the flowering time differences between the two accessions. With this rationale in mind, we selected from 305 F₂ plants of the MTE x 1408 cross 61 individuals that had flowered after 65 days in 16
LD (Figure 3; Table S3). We extracted a single pool of genomic DNA from these plants, and sequenced the pooled DNA on the Illumina GAIIx platform.

Sequencing reads were aligned to a preliminary genome assembly of C. rubella MTE. The proportion of sequence reads that could not be mapped on the MTE genome was 7.3%, which is similar to what we have found with A. thaliana accessions (CAO et al. 2011). As described (SCHNEEBERGER et al. 2009), polymorphisms that can be used for genetic mapping, including single nucleotide polymorphisms (SNPs), can be identified directly from F2 data, if one allows for both homo- and heterozygous SNPs. Accordingly, SNPs characteristic for 1408 were discovered from the bulked segregant data. These SNPs were then used to identify regions enriched for either 1408 or MTE SNPs (SCHNEEBERGER et al. 2009). Rearrangements and indels were not considered in this step. With this strategy, we identified on chromosome 6 of the genome assembly a region of 333 kb that contained over 600 positions where over 90% of reads supported a non-reference polymorphism (Figure 4; Table S4). The overall density of 2.0 SNPs/kb was close to the genome-wide average of 2.3 SNPs/kb.

Using all reads from the bulked segregant pools that mapped to this region, we annotated 1408 variants relative to the Capsella v1.0 gene set (http://www.phytozome.net/capsella/), which contained 90 protein coding genes in the final mapping interval. The algorithms implemented in SHORE v0.8 (OSSOWSKI et al. 2008) and pindel v0.2.4s (YE et al. 2009) detected 652 SNPs, 184 small indels of length 1 to 3 bp, 25 longer structural variants (SVs) of length 7
to 679 bp, and 109 highly variable regions, recognizable as simultaneous insertions and deletions (Table S4). Copy number variants (CNVs) and inversions were not detected. Fifty-one SNPs were predicted to cause nonsynonymous amino acid substitutions in 35 genes, with one SNP altering a stop codon. Five SVs affected the coding regions of four genes, and fourteen highly variable regions affected the coding regions of 11 genes. Close to the middle of the final mapping interval from position 2.968 to 3.301 Mb, at position 3.112 to 3.120 Mb, was the only C. rubella homolog of A. thaliana FLC, a well-known flowering repressor (MICHAELS et al. 2003; SHELDON et al. 1999). None of the other genes in this interval were closely related to homologs of known flowering time genes.

In A. thaliana, there is large variation in FLC levels, and accessions with weak FLC alleles often have very little FLC expression (LEMPE et al. 2005; SHINDO et al. 2005). Variation in FLC expression was more moderate in C. rubella accessions. MTE had the highest FLC levels, consistent with it being the latest accession, but accession 1408 did not have the least FLC expression (Figure 5A). Across all accessions, there was no clear correlation of FLC expression and flowering in unvernalized plants (Figure 5A), but a modest correlation of FLC and vernalization response (Figure 5B). As expected, expression of FLC was reduced after vernalization, although to quite different extent in different accessions, which is clearly seen in a reaction norm representation (Figure 5C).

Because it was unclear whether differences of FLC expression were likely to be the cause for earlier flowering of accession 1408 compared to MTE, we
compared the FLC DNA sequence of the two accessions. There is an ACAG-to-AGAG substitution at the very end of intron 6 that shifts the splice acceptor site so that two bases are inserted into the FLC mRNA of accession 1408, as revealed by sequencing of cDNA. This leads to a frame shift and truncation of the open reading frame, removing the last 35 of 198 amino acids (Figure 6).

To test whether this polymorphism could indeed be responsible for the flowering time difference between MTE and 1408, we introduced of both FLC alleles into A. thaliana FRI1-2 flc-3 plants. Because we did not know the extent of regulatory sequences in C. rubella, we expressed the FLC alleles under the control of the constitutive cauliflower mosaic virus 35S promoter. The MTE allele strongly delayed flowering, as did overexpression of a wild-type FLC copy from A. thaliana, and both were similarly effective in repressing the flowering activators FT and SOC1 (Figure 7A, B). In contrast, the 1408 allele delayed flowering by only a few days, from 25.2 to 29.6 days, and it had a much smaller effect on FT and SOC1 expression (Figure 7A, B). FLC transcript levels were similar in all three transgenic lines (Figure 7B). An analysis of cDNA obtained from plants expressing either the MTE or 1408 allele showed that the 1408 allele only produced the misspliced version of the FLC transcript (Figure 7C).

**FLC and variation in flowering time and vernalization response:** Several other accessions flowered as early as 1408 both with and without vernalization (Table 1, Figure 2). A survey of FLC sequences, however, did not reveal any additional accessions with the 1408 mutation. Compared to C. grandiflora,
there were very few FLC polymorphisms in C. rubella. All FLC sequences from C. rubella clustered together and formed a group that was distinct from C. grandiflora FLC alleles (Guo et al. 2009). This strongly suggested that the 1408 mutation arose very recently, after the split of C. rubella and C. grandiflora.

In A. thaliana, large insertions into the first intron are a common cause for reduced FLC activity (Gazzani et al. 2003; Lempe et al. 2005; Liu et al. 2004; Michaels et al. 2003; Sánchez-Bermejo et al. 2012; Shindo et al. 2005; Strange et al. 2011). Three C. rubella accessions had over 1 kb long insertions into the first intron, but none of them stood out as being particularly early flowering or having a particularly weak vernalization response (Table 1, Figure 6). We also surveyed the accessions for changes in the open reading frame of FRI, which is located in a region syntenic to chromosome 5 of A. thaliana (Figure S1) (Irwin et al. 2012). The open reading frames were intact in all 20 accessions. However, there were several non-synonymous substitutions (Figure 6), and π was much higher for FRI than for FLC (0.004 compared with 0.0004).

DISCUSSION

Flowering time variation in the Brassicaceae: Apart from disease resistance, flowering time is perhaps the most comprehensively studied adaptive trait in A. thaliana. In both laboratory and field conditions, there is a wide range in the onset of flowering. Much of this variation can be explained by allelic variation at FRI along with FLC and its homologs. Consistent with the role of FRI
and FLC in repressing flowering in unvernalized plants, much of the flowering variation disappears after vernalization (recently reviewed by Weigel 2012).

Because allelic variation at FRI and FLC is so common in A. thaliana, their homologs have been prime candidates for loci affecting flowering time in other Brassicaceae (Axelsson et al. 2001; Lou et al. 2007; Tadege et al. 2001; Uptmoor et al. 2012; Zhao et al. 2010a). There is evidence for FRI or FLC homologs being responsible for quantitative trait loci (QTL) affecting flowering in A. lyrata, Brassica napus and B. rapa, with the situation being less clear for B. oleracea (Kuittainen et al. 2008; Long et al. 2007; Okazaki et al. 2007; Pires et al. 2004; Razi et al. 2007; Schranz et al. 2002; Wang et al. 2011; Zhao et al. 2010b). While the potential molecular causes for a change in FLC function in these cases remain uncertain, splice site mutations in FLC homologs have been linked to earlier flowering in B. rapa and Capsella bursa-pastoris (Hu et al. 2011; Slotte et al. 2009; Yuan et al. 2009). However, in all these cases there has been little direct proof for a causal role of FLC or FRI alleles in flowering time variation. An induced mutation in the FLC homolog PERPETUAL FLOWERING 1 (PEP1) changes the perennial flowering habit of Arabis alpina. Its response to vernalization is different from the one of A. thaliana FLC, pointing to diversification in the function of FLC homologs in the Brassicaceae (Wang et al. 2009).

**Flowering time variation in C. rubella:** Capsella is more distantly related to A. thaliana than the other Arabidopsis species, but more closely than Brassica and Arabis. Capsella is a small genus with only five species, including one
tetraploid species, C. bursa-pastoris, one of the most frequent and widespread cosmopolitan weeds (CEPLITIS et al. 2005; HURKA et al. 2012; HURKA and NEUFFER 1997). The four diploid species include the self-compatible C. rubella, which originated only a few tens of thousands of years ago from the self-incompatible C. grandiflora (FOXE et al. 2009; GUO et al. 2009). Capsella grandiflora has a restricted distribution, in Western Greece and Albania and, rarely, Northern Italy, while C. rubella has spread across the Mediterranean, and followed European settlers into the New World and Australia (HURKA and NEUFFER 1997). In this work, we have documented flowering time variation in C. rubella under laboratory conditions. Similar to A. thaliana, flowering of many C. rubella accessions is accelerated by vernalization, longer photoperiods and increased ambient temperature. As in A. thaliana, later-flowering C. rubella accessions tend to respond more strongly to vernalization than early-flowering ones (LEMPE et al. 2005; SHINDO et al. 2005). However, the correlation between flowering time and vernalization response, or between flowering time and FLC expression, is less pronounced in C. rubella. This could be due to biased sampling, or it might reflect functional differences between C. rubella and A. thaliana. Given the different ranges of the two species, the second hypothesis is well worth further testing.

FRI alleles with premature stop codons explain a large fraction of flowering time variation in A. thaliana (ATWELL et al. 2010; LE CORRE et al. 2002; LEMPE et al. 2005; SHINDO et al. 2005; WERNER et al. 2005) In C. rubella, we did not
find any potential loss-of-function allele at FRI, even though the C. rubella accessions surveyed flower faster than many A. thaliana accessions. An indication for the functionality of FRI is that all C. rubella accessions analyzed responded to vernalization. In A. thaliana, structural variation at FLC itself appears to be often responsible for differences in the magnitude of the vernalization response (LEMPE et al. 2005; MICHAELS et al. 2003; SÁNCHEZ-BERMEJO et al. 2012; SHINDO et al. 2006). In C. rubella, we did not observe an obvious correlation between structural variation at FLC and vernalization response.

**Flowering time variation through independent FLC mutations:** The reference accession MTE turned out to be one of the latest-flowering C. rubella accessions we surveyed. The segregation pattern in the F2 of a cross to one of the earliest flowering accession in our collection, 1408, suggested the presence of a single major effect QTL responsible for much of the flowering time difference between the two accessions. To rapidly fine-map this QTL, we exploited recent advances in mapping-by-sequencing methods that do not require separate sequence analysis of the parental genomes (SCHNEEBERGER et al. 2009). We found that the gene most likely to be responsible for the QTL is the C. rubella homolog of FLC. While we cannot exclude that other genes in the interval contribute to this QTL, none of the 89 other genes in the mapping interval is closely related to a known flowering time regulator. Moreover, we note that it is difficult to draw firm conclusions about the quantitative differences in the activity of the MTE and 1408 alleles from the A. thaliana experiments, since
there might be additional trans and cis factors that modify FLC activity between the two species.

The likely causal polymorphism in FLC affects splicing, which in turn removes the C-terminal 35 amino acids. That most of the open reading frame remains intact may explain why accession 1408 retains a robust vernalization response. In A. thaliana, reduction- or loss-of-function alleles of FLC either have insertions in the first, large intron, or they carry premature stop codons that can lead to alternative splicing (Lempe et al. 2005; Liu et al. 2004; Ménendez-Vigo et al. 2011; Michaels et al. 2003; Sánchez-Bermejo et al. 2012; Shindo et al. 2005; Werner et al. 2005). However, alleles with splice site mutations have, to our knowledge, not been described in A. thaliana. It is striking that FLC alleles with splice site mutations have now been found in three other species, in B. rapa, C. bursa-pastoris and in C. rubella (this work, Hu et al. 2011; Slotte et al. 2009; Yuan et al. 2009).

Although several other accessions show a similar flowering behavior as 1408, the 1408 FLC allele is not shared by any other accession. Moreover, the pattern of polymorphisms in FLC alleles suggests that the 1408 mutation arose only after the split from C. grandiflora (Figure S2). Because of the strong recent bottleneck experienced by C. rubella (Foxe et al. 2009; Guo et al. 2009), there is only limited standing variation in this species, and new mutations likely play an important role in the adaptation to different environments. There is precedence for FLC, and to a lesser extent, FRI, loss-of-function alleles being rare. Although
few individual loss-of-function alleles segregate at appreciable frequency in A. thaliana flowering, collectively, such alleles are quite frequent. There is good evidence that the more common FRI loss-of-function alleles have increased in frequency due to selection for early flowering (Toomajian et al. 2006), a conclusion that is supported by experimental selection experiments (Scarcelli and Kover 2009). In animals, most cases of parallel evolution are from closely related species, but there are also examples of the same gene being responsible for genetic variation in more distantly related taxa, the most notorious one being the melanocortin-1 receptor gene (Mc1r), which underlies changes in pigmentation in reptiles, birds and mammals (Arendt and Reznick 2008; Elmer and Meyer 2011; Gompel and Prud’homme 2009; Hoekstra 2006; Protas and Patel 2008; Wood et al. 2005). Similarly, there are several cases of parallel evolution in flower pigmentation (Des Marais and Rausher 2010; Hoballah et al. 2007; Quattrocchio et al. 1999; Schwinn et al. 2006; Smith and Rausher 2011; Streisfeld and Rausher 2009; Whittall et al. 2006). How broadly FLC contributes to flowering time variation in C. rubella is, however, not known yet. Thus, further studies are needed to determine whether it is appropriate to speak of parallel evolution when comparing the basis of flowering time variation in A. thaliana and C. rubella.

In summary, we have found substantial flowering time variation in C. rubella, although its extent is smaller than in A. thaliana, perhaps reflecting the more restricted geographic range of C. rubella (Hoffmann 2002; Hurka and
In one accession, a major effect QTL affecting flowering maps FLC. In experimental crosses of A. thaliana, a small number of major effect QTL, including FRI, FLC, the FLC homolog FLOWERING LOCUS M (FLM)/MADS AFFECTING FLOWERING 1 (MAF1), the MAF2-5 cluster and FLOWERING LOCUS T (FT), explain much of the variation in flowering time (BRACHI et al. 2010; SALOMÉ et al. 2011; STRANGE et al. 2011). Given the reduced genetic diversity in C. rubella (FOXE et al. 2009; GUO et al. 2009), it will be interesting to determine how many additional loci are responsible for flowering time variation.

ACKNOWLEDGEMENTS

We thank Jeremy Schmutz (DOE-JGI and HudsonAlpha), Stephen Wright and Khaled Hazzouri (University of Toronto), and Adrian Platts (McGill University) for access to a preliminary version of the C. rubella genome assembly. We thank Sureshkumar Balasubramanian, Patrice Salomé, Jia-Wei Wang, Xuan Zhao and Zhong Zhao for advice, Dan Koenig and Patrice Salomé for comments on the manuscript, Korbinian Schneeberger for help with SHOREmap analyses, and Christa Lanz for sequencing support. Supported by Bundesministerium für Bildung und Forschung (BMBF) grant Next Generation Forward Genetics for Crop Plants (NUGGET), a Gottfried Wilhelm Leibniz Award to the Deutsche Forschungsgemeinschaft (DFG) and the Max Planck Society.
REFERENCES


AXELSSON, T., O. SHAVORSKAYA and U. LAGERCRANTZ, 2001 Multiple flowering time QTLs within several Brassica species could be the result of duplicated copies of one ancestral gene. Genome 44: 856-864.


KOCH, M. A., and M. KIEFER, 2005 Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species–


### Table 1. Flowering times of *C. rubella* accessions*

<table>
<thead>
<tr>
<th>Accession</th>
<th>Country of Origin</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
<th>Elevation (m)</th>
<th>16LD</th>
<th></th>
<th></th>
<th>16LDv</th>
<th>23LD</th>
<th></th>
<th>23SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n</td>
<td>Range</td>
<td>Mean</td>
<td>Median</td>
<td>n</td>
<td>Range</td>
<td>Mean</td>
<td>Median</td>
<td>n</td>
<td>Range</td>
</tr>
<tr>
<td>690</td>
<td>ESP</td>
<td>+36.15</td>
<td>-5.58</td>
<td>50</td>
<td>6</td>
<td>65-74</td>
<td>71.8</td>
<td>74.0</td>
<td>6</td>
<td>32-35</td>
<td>34.5</td>
</tr>
<tr>
<td>697</td>
<td>ITA</td>
<td>n/a</td>
<td>950</td>
<td></td>
<td>6</td>
<td>56-56</td>
<td>56.0</td>
<td>56.0</td>
<td>6</td>
<td>35-35</td>
<td>35.0</td>
</tr>
<tr>
<td>698</td>
<td>ITA</td>
<td>n/a</td>
<td>1180</td>
<td></td>
<td>6</td>
<td>70-74</td>
<td>72.0</td>
<td>72.0</td>
<td>6</td>
<td>49-49</td>
<td>49.0</td>
</tr>
<tr>
<td>762</td>
<td>GRE</td>
<td>+37.58</td>
<td>+23.43</td>
<td>n/a</td>
<td>6</td>
<td>72-72</td>
<td>72.0</td>
<td>72.0</td>
<td>6</td>
<td>32-32</td>
<td>32.0</td>
</tr>
<tr>
<td>844</td>
<td>GRE</td>
<td>n/a</td>
<td>200</td>
<td></td>
<td>6</td>
<td>60-64</td>
<td>62.5</td>
<td>63.0</td>
<td>6</td>
<td>37-37</td>
<td>37.0</td>
</tr>
<tr>
<td>879</td>
<td>GRE</td>
<td>+35.29</td>
<td>+24.42</td>
<td>1200</td>
<td>6</td>
<td>51-51</td>
<td>51.0</td>
<td>51.0</td>
<td>6</td>
<td>32-32</td>
<td>32.0</td>
</tr>
<tr>
<td>907</td>
<td>GRE</td>
<td>+39.40</td>
<td>+19.48</td>
<td>40</td>
<td>6</td>
<td>61-61</td>
<td>61.0</td>
<td>61.0</td>
<td>5</td>
<td>51-51</td>
<td>51.0</td>
</tr>
<tr>
<td>925</td>
<td>GRE</td>
<td>+39.40</td>
<td>+20.51</td>
<td>450</td>
<td>6</td>
<td>98-100</td>
<td>98.3</td>
<td>98.0</td>
<td>6</td>
<td>51-51</td>
<td>51.0</td>
</tr>
<tr>
<td>120</td>
<td>ESP</td>
<td>+28.19</td>
<td>-16.34</td>
<td>1000</td>
<td>6</td>
<td>72-74</td>
<td>72.3</td>
<td>72.0</td>
<td>6</td>
<td>37-37</td>
<td>37.0</td>
</tr>
<tr>
<td>120</td>
<td>ESP</td>
<td>+28.19</td>
<td>-16.34</td>
<td>1100</td>
<td>6</td>
<td>56-58</td>
<td>57.7</td>
<td>58.0</td>
<td>6</td>
<td>29-29</td>
<td>29.0</td>
</tr>
<tr>
<td>120</td>
<td>ESP</td>
<td>+28.19</td>
<td>-16.34</td>
<td>100</td>
<td>6</td>
<td>51-56</td>
<td>53.5</td>
<td>53.5</td>
<td>6</td>
<td>29-29</td>
<td>29.0</td>
</tr>
<tr>
<td>120</td>
<td>ESP</td>
<td>+28.19</td>
<td>-16.34</td>
<td>1200</td>
<td>6</td>
<td>70-72</td>
<td>70.3</td>
<td>70.0</td>
<td>6</td>
<td>32-32</td>
<td>32.0</td>
</tr>
<tr>
<td>121</td>
<td>ESP</td>
<td>+28.19</td>
<td>-16.19</td>
<td>0</td>
<td>6</td>
<td>60-65</td>
<td>62.3</td>
<td>63.0</td>
<td>6</td>
<td>27-29</td>
<td>28.7</td>
</tr>
<tr>
<td>131</td>
<td>FRA</td>
<td>+42.53</td>
<td>-0.06</td>
<td>n/a</td>
<td>6</td>
<td>70-81</td>
<td>76.5</td>
<td>77.0</td>
<td>6</td>
<td>27-29</td>
<td>28.0</td>
</tr>
<tr>
<td>137</td>
<td>ARG</td>
<td>-34.40</td>
<td>-58.30</td>
<td>10</td>
<td>6</td>
<td>65-70</td>
<td>66.7</td>
<td>65.0</td>
<td>6</td>
<td>32-32</td>
<td>32.0</td>
</tr>
<tr>
<td></td>
<td>Country</td>
<td>Code</td>
<td>Latitudes</td>
<td>Longitudes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
<td>------</td>
<td>-----------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>GRE</td>
<td>+35.29, +24.42</td>
<td>170</td>
<td>6</td>
<td>51-52</td>
<td>51.3</td>
<td>51.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>ITA</td>
<td>+43.28, +11.02</td>
<td>n/a</td>
<td>6</td>
<td>74-77</td>
<td>75.5</td>
<td>75.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>148</td>
<td>AUS</td>
<td>-31.56, +115.50</td>
<td>n/a</td>
<td>6</td>
<td>77-79</td>
<td>77.7</td>
<td>77.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MTE</td>
<td>n/a</td>
<td>n/a</td>
<td>6</td>
<td>100-106</td>
<td>102.5</td>
<td>103.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RIA</td>
<td>n/a</td>
<td>n/a</td>
<td>6</td>
<td>70-74</td>
<td>71.3</td>
<td>71.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*n/a, information not available; †+/− indicate Northern and Southern latitudes, and Eastern and Western longitudes.
FIGURE 1.—Distribution of flowering time among 20 C. rubella accessions in 16LD and 16LDv. Notice the differences in scale. Accessions, which are ordered by mean flowering time in each condition, are color coded for comparison. For each accession, the number of individuals that flowered at a given day is indicated.
FIGURE 2.—Relationship of flowering time variation and vernalization response. (A) Reaction norms of flowering time in 16LD without and with vernalization (-vern, +vern). (B) Correlation of mean flowering times with and without vernalization; same data as in Figure 2A.
FIGURE 3.–Flowering time in 16LD among 305 F$_2$ individuals from the MTE x 1408 cross. Average and range of flowering times of grandparents are indicated. Darker shading indicates the 61 individuals used for bulked segregant mapping.
Figure 4.—SHOREmap analysis of early flowering QTL. The homozygosity estimator is 0 at even allele frequencies of both parents, 1 when homozygous for late-flowering accession MTE, and -1 when homozygous for early-flowering accession 1408 (Schneeberger et al. 2009). Sliding windows of 100 kb with step size 10 kb were used. The region of chromosome 6 enriched at over 600 markers indicative of 1408 is indicated in grey. The eight largest scaffolds of a preliminary C. rubella
genome assembly are shown, corresponding to the majority of the eight chromosomes.
**Figure 5.**—Relationship between FLC expression measured by qRT-PCR and flowering time of accessions. (A) Correlation between FLC expression and flowering without vernalization. The lowest value was arbitrarily set to 1. (B) Correlation between FLC expression and vernalization response. (C) Effect of vernalization (vern) on FLC expression. The difference between FLC and BETA-TUBULIN, ΔcT, was used to calculate FLC expression, assuming PCR efficiency was 100% for both genes. Two biological replicates, each with two technical replicates, were analyzed.
Figure 6.—Sequence variation of FLC and FRI in C. rubella. In addition to the FLC point mutation at the end of the last intron in accession 1408, two insertions that are segregating in the population are shown. Thin lines indicate introns, thick lines protein-coding sequences. The C → G substitution creates a new splice acceptor site. PCR amplification and sequencing of cDNA confirmed that two bases are inserted into the FLC mRNA in accession 1408. There was no apparent heterogeneity in the sequence, suggesting that the canonical splice form, if it exists at all, is rare. For FRI, amino acid residues that appear to be ancestral based in several Brassicaceae (Irwin et al. 2012) are shown in bold; numbers indicate position in the peptide sequence. Gray numbers indicate affected accessions.
**Figure 7.**—Functional analysis of *C. rubella* FLC alleles in *A. thaliana*. (A) Flowering time of plants grown in 23LD. The three left strains are controls. *frICol-0 FLCCol-0* is the Col-0 reference strain, which carries a functional FLC allele, but a naturally inactive FRI allele (Johanson et al. 2000). *FRISf-2 flc-3* carries an introgression of the functional FRI allele from the Sf-2 accession and has also an induced mutation at the FLC locus (Michaels and Amasino 1999). The three right strains are transgenic lines, expressing the indicate FLC allele from the constitutive CaMV 35S promoter. The effect of the *C. rubella* MTE allele is similar to that of the fully functional *A. thaliana* Col-0 allele. n ≥ 30. (B) Expression of FLC and of two downstream
flowering regulators in transgenic and non-transgenic $FRJ^{Sf-2}$ $flc$-3 plants, determined by qRT-PCR using the $2^{ΔΔCt}$ method. Expression levels were normalized to those of wild-type Col-0 plants. Averages for four individuals for controls and eight for transgenic lines are shown. Error bars indicate standard errors of the mean. Differences between groups are significant at $p < 0.05$. (C) cDNA analysis of MTE and 1408 alleles expressed in transgenic A. thaliana plants. The splicing variant generated by the SNP in 1408 abolishes a BstXI restriction site in the amplified cDNA fragment.