Contributions of Flowering Time Genes to Sunflower Domestication and Improvement

John M. Burke‡, Steven J. Knapp§**, Scott D. Michaels*, and Loren H. Rieseberg*††

*Department of Biology, Indiana University, Bloomington, Indiana 47405
†Department of Biology, Washington University, St. Louis, Missouri 63130
‡Department of Plant Biology, University of Georgia, Athens, Georgia 30602
§Department of Crop and Soil Science, University of Georgia, Athens, Georgia 30602
**Center for Applied Genetic Technologies, University of Georgia, Athens, Georgia 30602
††Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

1Present Address: Department of Biology, Duke University, Durham, North Carolina 27708
2Present Address: Department of Biological Sciences, University of Illinois, Chicago, Illinois 60607

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. GU985570-GU987022, HQ110110-HQ110505.
Running Head: Flowering Genes and Sunflower Domestication

Keywords: flowering time, FLOWERING LOCUS T, domestication, sunflower, Helianthus annuus

Corresponding Author: Benjamin K. Blackman, Department of Biology, Duke University, Box 90338, Durham, NC 27708; Phone: (919) 660-7223; Fax (919) 660-7293; e-mail: bkb7@duke.edu

bkb7@duke.edu
ABSTRACT

Determining the identity and distribution of molecular changes leading to the evolution of modern crop species provides major insights into the timing and nature of historical forces involved in rapid phenotypic evolution. In this study, we employed an integrated candidate gene strategy to identify loci involved in the evolution of flowering time during early domestication and modern improvement of sunflower (Helianthus annuus). Sunflower homologs of many genes with known functions in flowering time were isolated and catalogued. Then, co-localization with previously mapped quantitative trait loci (QTL), expression or protein sequence differences between wild and domesticated sunflower, and molecular evolutionary signatures of selective sweeps were applied as step-wise criteria for narrowing down an original pool of 30 candidates. This process led to the discovery that five paralogs in the FLOWERING LOCUS T / TERMINAL FLOWER 1 gene family experienced selective sweeps during the evolution of cultivated sunflower and may be the causal loci underlying flowering time QTL. Our findings suggest that gene duplication fosters evolutionary innovation and that natural variation in both coding and regulatory sequences of these paralogs responded to a complex history of artificial selection on flowering time during the evolution of cultivated sunflower.
INTRODUCTION

Domestication by early farmers and improvement by modern breeders have dramatically transformed wild plants into today’s crops, and these human-driven phenotypic changes are excellent models for studying the genetics of rapid evolutionary responses to natural selection. Investigations seeking the genetic basis of domestication and crop improvement traits generally fall into two categories: top-down and bottom-up approaches (BURGER et al. 2008; DOEBLEY et al. 2006; ROSS-IBARRA et al. 2007; WRIGHT and GAUT 2005). Top-down studies begin with phenotypic variation and use forward genetic methods to positionally clone genetic variants underlying quantitative trait loci (QTL). Alternatively, association analyses, which exploit the fine mapping resolution provided by the recombination history of natural populations or complex crosses, are performed for candidate genes with known involvement in traits of interest. Top-down approaches have identified genes contributing to domestication and improvement traits in several important crop species (e.g. CONG et al. 2008; DOEBLEY et al. 1997; FRARY et al. 2000; JIN et al. 2008; LI et al. 2006; SIMONS et al. 2006; SUGIMOTO et al. 2010; TAN et al. 2008; THORNSBERRY et al. 2001; WANG et al. 2008; WANG et al. 2005; XIAO et al. 2008) but have several limitations. Positional cloning is costly and labor-intensive, as QTL detection power and fine mapping require large numbers of recombinant individuals and genetic markers, and it may not be feasible for species with long generation times or that are vegetatively propagated. Population structure and bias or error in candidate gene choice can confound association studies.

Bottom-up studies search, on a genomic scale, for molecular evolutionary signatures of selective sweeps with the expectation that the identities of genes under selection and their sequence variants will eventually lead back to phenotypic variation. Genetic targets of selective sweeps exhibit reduced sequence variation relative to interspecific divergence when compared to
neutrally evolving loci (Hudson et al. 1987; Wright and Charlesworth 2004). A localized signature of selection is evident because selection, unlike other evolutionary forces such as genetic drift and inbreeding, acts in a locus-specific manner. The timing of selection can also be examined by comparing diversity levels in wild progenitors, traditional landraces, and elite-bred cultivars (Burke et al. 2005). This approach has successfully identified genes that experienced selection during the domestication or improvement of maize (Hufford et al. 2007; Vielle-Calzada et al. 2009; Vigouroux et al. 2002; Wright et al. 2005; Yamazaki et al. 2005; Yamazaki et al. 2007) and sunflower (Chapman et al. 2008), but was less effective for sorghum (Casa et al. 2005; Hamblin et al. 2006). Though these screens are unbiased because target loci are randomly selected with respect to function, homologs of many genes with known effects on traits of interest are often omitted by chance or, in screens that assay simple sequence repeat (SSR) diversity, because they lack SSRs. The functions of many included genes may be unknown; consequently, connecting genes that exhibit signatures of selective sweeps back to domestication or improvement traits is rarely straightforward.

An alternative strategy is to synthesize information from both top-down and bottom-up methods (Figure 1). First, the genomic locations of genes homologous to those involved in a trait of interest in model species can be examined. The subset co-localizing with QTL intervals constitute an excellent group of candidates, and additional criteria (e.g. coding sequence or expression differences between the cross parents) can be applied to build evidence supporting the candidacy of these genes. Molecular evolution analyses can then be applied to test whether these genes exhibit signatures of selection during a stage of crop evolution. While still beholden to preexisting knowledge, this strategy is not agnostic with respect to phenotype, integrates known details of genetic architecture and mechanistic context, and directs attention to evolutionarily
relevant genes. This provides a sharpened focus in terms of genomic location, tissue, phenotype, and stage of crop evolution for subsequent functional and evolutionary confirmation.

Here, we present the results of such an integrated candidate gene approach in identifying genes involved in the evolution of flowering time during domestication and improvement of cultivated sunflower (*Helianthus annuus* L.). Flowering time is a critical agronomic trait, and its evolution was crucial for the domestication and spread of many crop species into new climatic regions (Colledge and Conolly 2007; Fuller 2007; Izawa 2007). The gene regulatory network controlling flowering time is exceptionally well described, making it an excellent trait for candidate gene analysis. Many genes involved in flowering time regulation are known, and the molecular mechanisms through which environmental and endogenous cues are integrated to trigger the floral transition have been elucidated in many cases (Farrona *et al.* 2008; Kobayashi and Weigel 2007; Michaels 2008).

Plants assess photoperiod cues by integrating information received from the circadian clock and light cues. These signals jointly regulate the abundance of *CONSTANS* (*CO*) such that its transcript and protein only accumulate and activate transcription of the floral inducer *FLOWERING LOCUS T* (*FT*) under inductive conditions (Figure 2; Adrian *et al.* 2010; Fornara *et al.* 2009; Imaizumi *et al.* 2005; Imaizumi *et al.* 2003; Jang *et al.* 2008; Laubinger *et al.* 2006; Liu *et al.* 2008b; Sawa *et al.* 2007; Suárez-López *et al.* 2001; Tiwari *et al.* 2010; Valverde *et al.* 2004; Wenkel *et al.* 2006; Yanovsky and Kay 2002). *FT* expression is also promoted by circadian signals through a *CO*-independent pathway that represses *FT* repressors (Jung *et al.* 2007; Mathieu *et al.* 2009; Wu *et al.* 2009). *FT* protein travels from the leaf through the phloem to the shoot apical meristem (Corbesier *et al.* 2007; Jaeger and Wigge 2007; Lin *et al.* 2007; Mathieu *et al.* 2007; Shalit *et al.* 2009; Tamaki *et al.* 2007). There, it
induces the meristem integrators \textit{SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1} ( \textit{SOC1} ) and \textit{APETALAI} ( \textit{API} ; \textit{Abe et al. 2005}; \textit{Wigge et al. 2005}; \textit{Yoo et al. 2005} ), which in turn promote expression of the meristem identity protein \textit{LEAFY} ( \textit{LFY} ) and initiate a signaling cascade of floral homeotic genes that pattern the floral meristem.

Combinatorial regulation by internal and environmental signals occurs elsewhere in the flowering time gene regulatory network as well. Hormonal signals from gibberellic acid promote flowering by targeting the DELLA family proteins, repressors of \textit{SOC1} and \textit{LFY}, for degradation while signals induced by abiotic stresses oppose these effects (Figure 2; \textit{Achard 2004}; \textit{Achard et al. 2007}; \textit{Achard et al. 2006}; \textit{Dill et al. 2004}; \textit{Murase et al. 2008}; \textit{Schwechheimer and Willige 2009}; \textit{Strader et al. 2004}; \textit{Ueguchi-Tanaka et al. 2005}; \textit{Willige et al. 2007}; \textit{Yamaguchi 2008}). Likewise, the \textit{FRIGIDA} ( \textit{FRI} ) pathway and endogenous chromatin-modifying complexes promote expression of the floral repressor \textit{FLOWERING LOCUS C} ( \textit{FLC} ) while other autonomous signals and external cues from the duration of overwintering, or vernalization, repress \textit{FLC} expression. (\textit{Farrona et al. 2008}; \textit{Lee et al. 2008}; \textit{Levy et al. 2002}; \textit{Li et al. 2008}; \textit{Liu et al. 2008a}; \textit{Michaelis et al. 2003}; \textit{Michaelis 2008}; \textit{Michaelis and Amasino 1999}; \textit{Schönrock et al. 2006}; \textit{Yu et al. 2004}).

Changes in flowering time coincide with major transitions in the evolution of cultivated sunflower. Sunflower was initially domesticated more than 4000 years ago from wild \textit{H. annuus} populations in eastern North America (\textit{Harter et al. 2004}; \textit{Heiser 1951}; \textit{Rieseberg and Seiler 1990}; \textit{Smith 2006}). Over its history as a crop, sunflower experienced several periods of intense selection and population bottlenecks (\textit{Putt 1997}; \textit{Tang and Knapp 2003}), including its transformation in the mid-20\textsuperscript{th} century by breeders into a globally important oilseed crop. While wild \textit{H. annuus} populations range from early to late flowering (\textit{Heiser 1954}), Native American
landraces are primarily late or very late flowering (HEISER 1951). In contrast, most elite-inbred modern cultivated lines are early flowering as a consequence of selection for shorter growing seasons during improvement (GOYNE and SCHNEITER 1988; GOYNE et al. 1989).

QTL studies performed on a landrace x wild *H. annuus* cross (WILLS and BURKE 2007) and an elite x wild *H. annuus* cross (BAACK et al. 2008; BURKE et al. 2002) have found that the genetic architecture of flowering time differences between wild and domesticated sunflower is oligogenic. In each case, one major and several minor QTLs controlling flowering time were detected, and QTLs concordant between these crosses were detected in two regions. No sunflower domestication or improvement locus for flowering time has been positionally cloned. Though a recent bottom-up screen of 492 loci identified two selected genes that map to flowering QTL and belong to gene families with flowering time regulators (CHAPMAN et al. 2008), functional studies confirming a role of either gene in flowering have yet to be completed.

Here, we report our findings using an integrated strategy to study 30 sunflower homologs of flowering time regulators (Figure 1, Table 1). Specifically, we asked 1) whether any of these genes met multiple successive criteria highlighting them as strong candidate genes for domestication or improvement and 2) what the identity and distribution of molecular changes in these genes revealed about the timing and nature of selection on flowering time. This work extends and contextualizes findings reported in our recent study of the *FT/Terminal Flower 1* (*TFL1*) family, which found functional and evolutionary support for a homolog of *FT, HaFT1*, as the gene underlying one of the two concordant QTLs (BLACKMAN et al. 2010).

MATERIALS AND METHODS
Ortholog identification: A list of genes with demonstrated involvement in regulation of flowering time in Arabidopsis was compiled from the literature (Figure 2), and the Helianthus expressed sequence tag (EST) collections generated by the Compositae Genome Project were screened for homologs of these genes. Initial EST library content, as well as methods of library construction, sequencing, and assembly, are described in Heesacker et al. 2008. Subsequent library construction and sequencing has produced new EST collections for *H. ciliaris*, *H. exilis*, and *H. tuberosus*. All results reported are based on the EST assemblies available at http://cgpdb.ucdavis.edu/cgpdb2/asteraceae_assembly. Homologs of flowering time genes, MADS-box genes, and CONSTANS-like (COL) genes were identified by searching a report of top BLASTx hits of the *H. annuus* EST assembly to *Arabidopsis thaliana* proteins by The Arabidopsis Information Resource locus ID number. When searching the *H. annuus* report did not yield any homologs, reports for additional Helianthus species were searched.

The sunflower EST collection has grown incrementally since the beginning of this project, and early releases did not contain homologs of many key flowering time genes. Therefore, PCR and hybridization-based methods were employed to obtain these genes. Acquisition of the four *HaFT* homologs has been described (Blackman et al. 2010). Partial sequences of *HaGI* and *HaTFL1* were obtained with degenerate primers designed to alignments of homologs from other species present in GenBank (Table S1). Previously published degenerate primers successfully amplified partial sequences of *HaLFY* (Aagaard et al. 2006) and *HaCOL2* (Hecht et al. 2005). Partial sequence of *HaSOC1* was obtained with primers designed to a *Chrysanthemum x morifolium* *SOC1*-like sequence (GenBank accession number AY173065). The *H. annuus* cultivated line HA383 BAC library generated by the Clemson University Genome Institute was then screened with radioactively labeled overgo probes (Ross...
et al. 2001) designed to HaCOL2, HaSOC1, and HaTFL1 partial sequences (Table S1). These screens resulted in identification of genomic clones containing full length HaCOL1, HaSOC1, and HaTFL1. Full-length sequences of HaCOL2, HaGI, and HaLFY were acquired by assembly with sequences from subsequently available ESTs, TAIL PCR, and 5’ and 3’ RACE.

**Genetic mapping:** Map positions were obtained for 30 candidate genes by genotyping markers on subsets of one of six mapping panels: 94 of 214 NMS373 x (NMS373 x Ann1811) BC1 individuals (GANDHI et al. 2005); 96 of 378 Hopi x Ann1238 F2 individuals (WILLS and BURKE 2007); 96 of 374 CMSHA89 x Ann1238 F3 individuals (BURKE et al. 2002); 48 of 94 RHA280 x RHA801 RILs (LAI et al. 2005; TANG et al. 2002); 94 of 262 PHC x PHD RILs (S. Knapp, unpublished); and 94 of 94 NMS801 x Arg1805 F1s (HEESACKER et al. 2009). Portions of the genes were amplified by PCR with gene-specific primers (Table S1). Parental DNAs or a subset of progeny of each mapping panel were initially screened for polymorphisms by sequencing. For most of the polymorphic candidate genes, single nucleotide polymorphisms (SNPs) were then scored on a mapping panel by DHPLC analysis carried out on a WAVE nucleic fragment analyzer (Transgenomic Inc.) using a DNAsep HT Column as described previously (LAI et al. 2005). A SSR in HaGID1B was mapped by assaying length polymorphism of fluorescently labeled PCR products on an Applied Biosystems 3730xl sequencer as previously described (BURKE et al. 2002). HaFT3, HaSVP, HaLFY, and HaGA2ox PCR products were directly sequenced. Linkage mapping was performed with MAPMAKER 3.0/EXP (LANDER et al. 1987).

**Gene expression analysis:** To survey candidate gene expression in leaf tissue, wild accession Ann1238 and elite line CMSHA89 were grown at 25.5ºC in growth chambers under short days
(8 hours light, 16 hours dark) or long days (16 hours light, 8 hours dark). At 30 days after sowing, leaf tissue was collected every 4 hours from dawn to 20 hours after dawn. To survey expression in shoot apices, Ann1238 and CMSHA89 plants were grown in long day conditions and shoot apices were collected from germinated seedlings before sowing and seedlings at 10, 20, 30, and 40 days after sowing. Shoot apices were collected from Ann1238 only at 60 days after sowing as well. Some Ann1238 plants were also transferred from long days into short days at 20 days after sowing, and shoot apices were collected from these plants 10 days after transfer. In both the leaf and shoot apex collections, three biological replicates were taken at each time point or developmental age, and samples from three plants were pooled within in each replicate.

RNA was isolated with the Spectrum™ Plant Total RNA Kit (Sigma) and treated with On-Column DNase (Sigma) during extraction. Shoot apex RNA samples were further cleaned and concentrated with an RNeasy® MinElute™ Cleanup Kit (Qiagen). RNA was converted to cDNA with SuperScript™ III Reverse Transcriptase (Invitrogen), and gene-specific primers (Table S1) were used to amplify candidates by PCR. PCR on leaf cDNA was carried out in 20uL reactions with Platinum® Taq DNA Polymerase (Invitrogen) run for 30 cycles for all candidate genes and for 26 cycles for a control gene, Ha60S rRNA. PCR of shoot apex cDNA was carried out for 32 cycles for HaFT1 and HaSOC1; 30 cycles for HaTFL1, HAM75, HaLFY, HaDELLA2, and HaLD; and 28 cycles for Ha60S rRNA. PCR product concentrations were visualized on 1% agarose gels stained with ethidium bromide.

cDNA sequencing: Gene-specific primers were designed to the 5’ and 3’ ends of the candidate genes mapping to linkage groups (LGs) with flowering time QTL, and full length coding sequences were amplified by PCR from Ann1238 and CMSHA89 leaf or shoot apex cDNAs
(Table S1). For candidate genes where alignments of ESTs or otherwise-isolated sequences did not cover the entire coding sequence, 5’ and 3’ RACE was performed to try to isolate the remaining sequence. If this was unsuccessful, primers were designed to the largest possible portion of the gene. In most cases, full-length sequence could be obtained from first-strand cDNA generated by reverse transcription (RT) reactions primed with an oligo dT primer. However, particularly for long or low abundance transcripts, RT reactions primed with gene-specific primers targeted to the 3’-UTR were required to acquire a cDNA substrate amenable to full length cDNA amplification by PCR. The number of PCR cycles and duration of extension time were increased for genes of large size or low transcript abundance.

**Molecular evolution analyses:** Portions of candidate genes and seven putatively neutral control loci were amplified by PCR and sequenced using gene-specific primers (Table S1) on a diversity panel of wild and domesticated *H. annuus*. This panel included 18 individuals from elite inbred lines, 19 individuals from Native American landraces, 23 individuals from a geographically diverse sample of wild *H. annuus* populations, and 6 individuals from *H. argophyllus* (Table S2). Although *H. argophyllus* has been isolated for ~1.1 million generations from *H. annuus*, due to *H. annuus*’ large effective population size, incomplete lineage sorting may affect divergence estimates because some shared polymorphisms may be confounded with fixed differences (Strasburg *et al.* 2009). PCR fragments from heterozygous individuals were cloned (TOPO TA cloning, Invitrogen) and sequenced with T7 and T3 primers. Multiple clones were sequenced per individual, compared with each other, and compared to the original direct sequencing reads to detect and eliminate errors introduced during PCR and cloning. Seven putative neutral reference loci (Table S3) were chosen because they were shown to be evolving
neutrally (LIU and BURKE 2006) or because they had the most complete sequence information available from an ongoing study of sequence diversity by the Compositae Genome Project. A multi-locus Hudson-Kreitman-Aguadé (HKA; HUDSON et al. 1987) test demonstrated that sequence diversity of the seven reference loci did not deviate from a strictly neutral evolutionary model (http://lifesci.rutgers.edu/~heylab/heylabsoftware.htm#HKA).

Diversity parameters – number of segregating sites (S), number of haplotypes (h), pairwise nucleotide diversity (\( \pi \)), and Watterson’s estimator of diversity (\( \theta_w \)) – were calculated with DnaSP (ROZAS et al. 2003). DnaSP was also used for synonymous substitution rate calculations between paralogs. MLHKA (WRIGHT and CHARLESWORTH 2004) was used to conduct maximum-likelihood HKA tests. For each candidate gene, likelihoods for a strictly neutral model and a model where the candidate was under selection were calculated and compared with a likelihood ratio test. To examine the timing of selection, separate tests were conducted for elite, landrace, and wild samples.

RESULTS

Flowering time gene homologs in sunflower: Our analysis of the Helianthus EST collections revealed that most flowering time genes identified in model plant species have homologs in sunflower (Figure 2). For the 30 genes pursued as potential candidates, GenBank accession numbers for contributing ESTs, percent amino-acid identity to A. thaliana homolog, and percent of the total protein sequence obtained are reported in Table 1. Information for all genes not involved in subsequent analyses is summarized in Table S4. Phylogenetic analyses conducted to more specifically identify homologs in two large gene families—CONSTANS LIKE
(COL) genes and Type II MADS-box transcription factors—are described in the Supporting Information (Extended Materials and Methods, Figures S1 and S2, Tables S5 and S6).

In total, the *H. annuus* EST collection (93,428 sequences) contained just over half the genes in our search (82 of 155 total). Expanding our search to the EST collections of related sunflower species (189,585 additional sequences) identified homologs of 31 additional genes. Sunflower homologs were identified for genes throughout the flowering time gene regulatory network, including genes that function in the circadian clock, light reception, *CO*-dependent and *CO*-independent photoperiod pathways, gibberellin biosynthesis and signaling, the autonomous pathway, vernalization response, chromatin modification, and meristem integration (Figure 2). The absence of ESTs homologous to any *FRI* pathway genes was a notable exception.

Several causes may have contributed to the absence of sunflower EST homologs of certain genes. Homologs were not expected for genes like *PHYTOCHROME D* that arose by duplication events in lineages not ancestral to sunflower (MATHEWS and SHARROCK 1997). Insufficient library sequencing depth or incomplete sampling of particular cell types, developmental stages, circadian time points or environmental conditions may also have produced gaps in coverage. For instance, absence of several critical flowering genes expressed in the shoot apical meristem during floral induction (*MYB33, MYB65, LFY, TFL1, and FD*) signals that samples used for EST library construction were impoverished for this tissue.

In several cases, multiple Helianthus copies were found of genes present as a single copy in Arabidopsis, e.g. *GIGANTEA (GI), FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1), and ZEITLUPE (ZTL)*. Recent gene family expansion has also been described for the *FT/TFL1* family (BLACKMAN *et al.* 2010). Some of these expansions likely result from persistence of duplications that arose during polyploidy events at the base of the Compositae or more recently
within the Heliantheae (Barker et al. 2008). For instance, values of $K_s$, the synonymous substitution rate between two sequences, for sunflower duplicate pairs of ZTL and FKF1 were 0.71 and 0.66, respectively, consistent with the former event; $K_s$ was 0.45 for one of the duplication events in the sunflower $FT$ clade, consistent with the latter event.

Although describing sunflower Type II MADS-box gene diversity was not our principal aim, the involvement of many of these genes in flowering required more rigorous evaluation of orthology with a phylogeny. Though we limited our analysis to $H. annuus$ ESTs and excluded several incomplete sequences, sunflower orthologs clustered in all major MADS-box clades except AGL12 and GGM13 (Figure S2; Becker and Theissen 2003), including sequences related to *SHORT VEGETATIVE PHASE (SVP)*, *SOC1*, *FRUITFULL (FUL)*, *AP1*, and *FLC*. *FLC* and the *MADS AFFECTING FLOWERING (MAF)* genes were once suspected to be a Brassicaceae-specific clade (De Bodt et al. 2003), but recent work has determined that *FLC* homologs are present in diverse eudicots and can act as a cold-repressed regulators of flowering (Hileman et al. 2006; Reeves et al. 2007). While temperature’s influence on flowering in sunflower is well documented (Goyne and Schneiter 1988; Goyne et al. 1989), to our knowledge there is no evidence of vernalization sensitivity. *FLC* also regulates seed germination in *A. thaliana* (Chiang et al. 2009), and this function could be conserved. Therefore, the functional importance of this *FLC* homolog and its regulators requires further investigation.

**Defining candidates by mapping:** As a first criterion for choosing flowering time homologs as candidate domestication or improvement genes, we tested whether genes co-localized with previously identified flowering time QTL (Baack et al. 2008; Burke et al. 2002; Wills and Burke 2007). Markers polymorphic in one of the six mapping panels available to us were
identified (Table 1, Figure 2); and of the 30 genes mapped, 18 mapped to LGs containing flowering time QTL (Figure 3). These genes included representatives from throughout the gene regulatory network. QTLs for flowering time have been detected in just over half (9 of 17) of the LGs in the *H. annuus* genome in previous crosses between wild and domesticated sunflower, and these QTLs, when projected onto the same genetic map, cover ~15% of the genome (148.2 cM / 972.6 cM map from Burke *et al.* 2002). Nine of our 30 candidates (30%) unambiguously co-localized with QTLs, suggesting these regions may be enriched for flowering time genes.

In some cases, QTL chromosomal blocks and gene locations clearly overlapped, whereas in other cases candidate genes and QTL did not overlap or the coincidence was ambiguous (Figure 3). For instance, *HaPHYB* mapped to the opposite end of LG1 from the QTL region. The locations of *HaFT1*, *HaFT2*, and *HaFT3* coincided with the LG6 chromosomal block in which a QTL was detected in all three previous QTL studies; *HaSOC1* and *HaSPY* also map to LG6 but fall outside the concordant region. *HaCDFL1* (also mapped as c1921 in Chapman *et al.* 2008) and *HaTFL1* mapped to positions similar to QTL on LG7. Relative positioning of QTLs and genes on the remaining linkage groups was less certain. LG15 was the only group with a flowering time QTL to which no flowering time gene homologs mapped. Notably, this is also the only QTL found in the cross between wild *H. annuus* and a Native American landrace but not in the cross of wild *H. annuus* and an elite-bred line.

Though we could confidently assign genes to LGs, ambiguity in determining whether a gene and QTL were truly co-localized arose from several sources. Since different markers were used in the construction of each of the six mapping panels and only two of the panels were used for QTL detection, approximate map locations of the candidate genes and QTLs were often inferred from the relatively few markers common to more than one mapping panel. Inconsistent
ordering of markers on different panels compounded this issue. The map with the most QTLs (Ann1238 x CMSHA89 F3s; Burke et al. 2002) partly consisted of dominant markers, leading to uncertainties in marker ordering, potential false splitting of single QTL into multiple QTLs, and some large confidence intervals around QTL peaks that often spanned half to all of a LG. Finally, variation in marker density among LGs and panels may have affected the breadth of QTL confidence intervals. Due to these sources of uncertainty, we conservatively designated all 18 genes located on LGs with flowering time QTLs as preliminary candidates.

Expression and sequence surveys of candidates: Next, we winnowed down our preliminary candidate gene pool by surveying for notable differences in expression or protein sequence between parents of the elite x wild cross used to develop two of the three QTL panels. The domesticated parent was inbred line CMSHA89 (Figures 4 and 5, Table 2). The wild parent of this cross, Ann1238, came from a population near Cedar Point Biological Station, Nebraska. Individuals of this population have a short day-sensitive photoperiod response in flowering, unlike CMSHA89, which is long-day sensitive and flowers earlier than Ann1238 in both inductive and non-inductive photoperiods (Blackman et al. 2010). Since no genes mapped to the LG containing the single QTL unique to the landrace x wild panel (Wills and Burke 2007) we did not include the landrace parent of that panel in our analysis.

Expression of candidates was assayed by RT-PCR in leaf and/or shoot apex tissue, depending on the gene’s expression domain(s) in other species. Since many photoperiod pathway genes are controlled by the circadian clock and since the parents differ in photoperiod response, we surveyed gene expression in leaf tissue in plants raised in short days and long days, taking samples every four hours over the course of a single day (Figure 4). For the survey of
shoot apex expression, we observed transcript levels over a developmental time course in long
days, allowing us to look at the relative timing of upregulation of candidates during growth
(Figure 5). Expression in the shoot apices was also examined for a set of Ann1238 individuals
transferred to short days, this population’s inductive photoperiod, for 10 days.

We assayed expression differences between the parents as a proxy for functionally
important cis-regulatory sequence variants. While we expect this served as a reliable indicator,
two situations may have arisen where this was not the case. First, expression differences may
have been too subtle to detect by RT-PCR or occurred in conditions not surveyed. Second, the
parents may have identical expression patterns yet differ by compensatory sequence changes.
Consequently, transgressive variation in gene expression, and potentially in flowering time,
could have segregated in hybrids. Though missing such genes may have led us to discard
candidates underlying particular QTL, attrition of our primary search targets – genes responsible
for phenotypic differences that experienced directional selection during the evolution of
cultivated sunflower – is unlikely.

The entire coding region or the largest partial sequence possible (Table 1) was amplified
from Ann1238 and CMSHA89 leaf or shoot apex cDNA for 17 of the 18 preliminary candidates.
Previous work found that HaFT3 is not expressed, but previously characterized loss of function
mutations were genotyped from genomic DNA (BLACKMAN et al. 2010).

Both expression and sequence differences: Of the 18 preliminary candidates, two exhibited
expression and protein sequence differences. As previously reported (BLACKMAN et al. 2010),
the domesticated allele of HaFT1 contains a 1bp deletion in the third exon that causes a
frameshift and extension of the protein by 17 amino acids. HaFT1 is expressed at very low
levels in Ann1238 plants raised in long days but is upregulated on transfer to short days. In contrast, *HaFT1* is upregulated in long days in CMSHA89 as soon as 10 days after sowing.

In both leaf and shoot apex tissue, *HaSOC1* expression differed. In Ann1238, *HaSOC1* was upregulated at most times of day in leaves of plants raised in short days relative to leaves of plants raised in long days, but the opposite pattern was observed in CMSHA89. Shoot apex expression of *HaSOC1* also started earlier in development in CMSHA89 than Ann1238. A SNP causing a histidine to a glutamine substitution in the K-box domain of *HaSOC1* was also found in CMSHA89. Due to these differences, *HaFT1* and *HaSOC1* retained candidate gene status.

*Sequence differences only:* Five genes showed no notable differences in the timing or abundance of gene expression consistent across all biological replicates, but they did have replacement substitutions or insertion/deletion mutations between Ann1238 and CMSHA89.

*HaCOL1* had three in-frame insertion/deletion differences: a 6bp insertion polymorphic in Ann1238, a 24bp insertion in CMSHA89, and a SSR, which is polymorphic in Ann1238 but always contains more repeats than present in CMSHA89. The two lines also differed by two replacement substitutions in *HaCOL1* that both cause replacement of a histidine in Ann1238 with a glutamine in CMSHA89. Two non-conservative amino acid changes – a secondary structure-altering serine to proline substitution and a charge-changing lysine to phenylalanine substitution – distinguished the *HaCOL2* sequences of Ann1238 and CMSHA89. None of the mutations in *HaCOL1* or *HaCOL2* were in either of the two B-box domains or the CCT domain.

Though we opted to keep *HaCOL1* and *HaCOL2* designated as candidate genes based on these notable mutations, changes in the remaining three genes were not dramatic enough to merit continued study. Two substitutions were found in *HaFKF1* (serine to tyrosine; aspartic acid to
valine), one in an unconserved residue of the PAS domain, the other in the fifth KELCH domain. Ann1238 and CMSHA89 *HaSLY1* proteins sequences differed only by a glutamic acid to aspartic acid substitution in the unconserved 5’ region. Finally, the two lines differed by a serine to glutamine substitution in the unconserved 3’ end of *HaSPY*. Though these substitutions appear conservative, we note that the possibility they have phenotypic effects cannot be excluded.

Of the four previously discovered non-functionalizing mutations in *HaFT3* (Blackman et al. 2010), none were present in Ann1238 and one was present in CMSHA89, a 7bp frameshift mutation in the fourth exon. In that study, lack of expression of *HaFT3* in Ann1238 was also demonstrated for all tissues tested. Even so, we retained *HaFT3* in our candidate list.

**Expression Differences Only:**

Eight genes exhibited differences in expression but had no differences in protein sequence. In the leaf, *HaFT2* and *HaFT4* showed expression patterns consistent with the divergence in photoperiod response between the two parents. Both were expressed only in short days in Ann1238 and only in long days in CMSHA89 (Figure 4), though one CMSHA89 replicate (shown) had a small peak of *HaFT2* expression in short days. In CMSHA89, transcript abundances of both genes appeared higher in the inductive photoperiod at all times of day as well, a result quantitatively confirmed for *HaFT2* previously (Blackman et al. 2010).

For *HaCDFL1*, *HaPHYB*, and *HaDELLA2*, timing of expression in each photoperiod was similar, but expression levels appeared higher at most or all time points in CMSHA89 than in Ann1238 (Figure 4). Similar to *HaFT1*, *HAM75* and *HaTFL1* are expressed at very early developmental stages in CMSHA89 shoot apices in long days; expression of both genes only increases later in development in Ann1238 plants under long days or upon transfer to short days.
(Figure 5). *HaLFY* is also expressed earlier in development in long days in CMSHA89 than in Ann1238, though expression was shifted earlier when Ann1238 was transferred to short days.

Though some of the expression differences observed may be partially or wholly caused by *trans*-acting changes in upstream factors, we cannot rule out contributions of *cis*-regulatory changes. Consequently, all eight genes were kept as candidates for further study.

**No expression or sequence differences:**

*HaCPS* and *HaLD* showed no differences in expression (Figures 4 and 5) between Ann1238 and CMSHA89 consistent across all biological replicates, and no cDNA sequence differences were found (Table 2). Therefore, candidate status was revoked for these genes.

**Signatures of selection on refined candidate set:** We next conducted MLHKA tests on the remaining 13 candidate genes to determine whether these genes were under selection during early domestication or crop improvement (WRIGHT and CHARLESWORTH 2004). Portions of the candidate genes and seven putative neutral reference genes were sequenced on a panel of elite-bred cultivars, Native American landraces, and wild *H. annuus* (Table S2). Several individuals of *H. argophyllus*, *H. annuus*’ sister species, were sequenced to obtain divergence measures.

Levels of diversity in the reference genes in wild, landrace, and elite lines were comparable to levels found in previous work (Table 3; CHAPMAN et al. 2008; KOLKMAN et al. 2007; LIU and BURKE 2006). Also as in previous studies, average pairwise nucleotide diversity ($\pi$) and average Watterson’s estimator of diversity ($\theta_w$) were highest in the wild populations, intermediate in the landraces, and lowest in the elite lines, indicative of genetic bottlenecks at both the domestication and improvement stages (Table 3; BURKE et al. 2005; CHAPMAN et al.
A similar progressive drop in diversity with each stage of crop evolution was observed for the candidate genes.

In an MLHKA test, the likelihood of a strictly neutral model is compared to the likelihood of a model where a candidate gene is under selection. For each gene, separate tests were performed for elite, landrace, and wild sequence datasets to determine the timing of selection. Five of the 13 candidate genes tested had significant MLHKA tests indicative of selection occurring during domestication or improvement (Table 4). As previously reported (BLACKMAN et al. 2010), sequence variation in HaFT1 was consistent with neutral evolution in wild H. annuus, but consistent with selection in landraces and elite lines. Notably, all additional members of the FT/TFL1 family tested had significant signatures of selection in elite lines but not for wild or landrace datasets. HaFT2, HaFT3, and HaFT4 sequences were identical in all elite lines. A 1bp indel polymorphism in a homopolymer run of adenines present in just a single line was the only SNP segregating in HaTFL1 in elite lines.

DISCUSSION

**Multiple FT/TFL1 family members emerge as strong candidates:** By integrating information gained from both top-down and bottom-up analyses, we identified several strong candidate domestication and improvement genes affecting flowering time (Figure 1). Of the genes pursued, 17% (5/30) exhibited signatures of selection during a stage of cultivated sunflower evolution. Notably, all of these candidates are FT/TFL1 homologs.

HaFT1, HaFT2, and HaFT3 mapped to a large QTL on LG6 that is concordant between elite x wild and landrace x wild crosses and explains 7.6 to 36% of the variance in flowering time depending on the cross (BAACK et al. 2008; BURKE et al. 2002; WILLS and BURKE 2007).
Additional analyses published elsewhere (BLACKMAN et al. 2010) provide strong support for *HaFT1* as the genetic cause of this QTL. Involvement of *HaFT3* is unlikely, as it is non-functionalized. The domesticated allele delays flowering in near isogenic lines (NILs) segregating for this QTL region, and though *cis*-mapping differences affecting *HaFT2* expression were found, they were inconsistent with this result as higher expression of the CMSHA89 allele would be predicted to promote earlier flowering. In contrast, the domesticated allele of *HaFT1* contains a frameshift mutation consistent with the NIL phenotypes provided the mutation has dominant negative action. Transgenic analyses in *A. thaliana* confirmed this by demonstrating that the domesticated allele of *HaFT1* suppresses complementation of late flowering *ft* mutants by *HaFT4* while the wild allele does not (BLACKMAN et al. 2010).

*HaTFL1* mapped to a minor QTL on LG7 concordant between both landrace x wild and elite x wild crosses (BURKE et al. 2002; WILLS and BURKE 2007); however, *HaTFL1* exhibited a signature of selection during improvement but not early domestication. This discrepancy may be explainable if the domesticated parent in the landrace x wild panel carried a functionally equivalent haplotype as the domesticated parent in the elite cross. Alternatively, the QTL may have complex genetic underpinnings.

A recent SSR-based screen identified two other putative flowering time gene homologs in this interval with signatures of selection during improvement (CHAPMAN et al. 2008); however, we consider *HaTFL1* to be the best-supported individual candidate. One gene, c1921, is the same gene as *HaCDFL1* from our analysis, but in our diversity panel no signature of selection in elite lines was apparent. The larger size of our study (24 more elite haplotypes) likely allowed better sampling of low frequency alleles, explaining this discrepancy.

Both \textit{HaFT4} and \textit{HaCOL2} map to LG14, where a flowering time QTL was detected only in the elite x wild cross and only in field-raised plants (Baack \textit{et al.} 2008). While \textit{HaCOL2} falls within the QTL, \textit{HaFT4}’s relative position is ambiguous based on shared markers, and it may fall outside the QTL region. Nonetheless, \textit{HaFT4} exhibits a signature of selection with improvement and \textit{HaCOL2} does not, making \textit{HaFT4} the better candidate improvement gene.

In terms of the genes potentially identifiable through our strategy, we have identified candidates under selection mapping to flowering time QTL on 33\% (3/9) of LGs with QTL. Not all of the genes underlying detected QTL are expected to have experienced the same selective pressures during either domestication or improvement, however. Genetic variation for flowering has been detected in crosses between elite lines (Bert \textit{et al.} 2003; Leon \textit{et al.} 2000; Leon \textit{et al.} 2001), and different lines may have different recent histories of selection on flowering time by farmers in different geographic regions. This scenario would increase diversity and obscure a signature of selection. Since QTL analysis of flowering time has been conducted only in a single elite x wild cross, improvement-specific QTLs and line-specific QTLs cannot be distinguished.
As a result, some CMSHA89 x Ann1238 cross-specific QTLs may be due to functional variants in candidates we identified by mapping, expression, and sequence analyses that did not show signatures of selection (e.g. \textit{HaLFY} on LG9). Though these line-specific variants may be of great value for plant breeders or aid description of the molecular basis of particular phenotypes, they may offer little evolutionary value toward improving knowledge of the early domestication process. Since we examined genes homologous to well-characterized regulators of flowering, there are clear directions for follow up studies.

Our initial concerns about ambiguity in QTL and gene relative positions may have been unfounded. No genes on LGs with flowering time QTL but clearly mapping outside QTL regions (e.g. \textit{HaSOC1} and \textit{HAM75}) exhibited signatures of selection, suggesting our efforts could have been more efficient. In systems where more QTL studies have been conducted, more formal meta-analyses and computation of a composite map combining several densely genotyped panels may further improve the confidence of QTL / candidate gene co-localization and thus also improve the efficiency of this approach (\textsc{Arcade et al.} 2004; \textsc{Chardon et al.} 2004). Several observations suggest genetic hitchhiking is a concern that impacts how effectively improvement loci can be identified, however. Signatures of selective sweeps during improvement were found for all three closely linked \textit{HaFT} genes on LG6, but variation in at least one of these copies, \textit{HaFT3}, is unlikely to have been under selection directly as this gene is non-functional (\textsc{Blackman et al.} 2010). Hitchhiking may also explain why multiple closely linked genes on LG7 exhibit evidence for selection during improvement (\textsc{Chapman et al.} 2008).

Linkage disequilibrium (LD) decayed relatively rapidly, $r^2 < 0.1$ within ~2kb for a sample of elite and landrace lines (\textsc{Liu and Burke} 2006), but LD persisted for much greater distances in a sample of solely elite lines ($r^2 \sim 0.32$ at 5.5kb and $r^2 < 0.1$ at 150kb; \textsc{Fusari et al.}
LD may persist over even longer distances in regions containing targets of selection. For instance, reductions in diversity around targets of selection in maize and rice have been shown to persist over large regions (~250kb to ~1.1Mb) containing multiple additional genes (Olsen et al. 2006; Palaisa et al. 2004; Tian et al. 2009).

Even if LD persists for such large distances in elite sunflower lines, genes would have to be linked by less than 1 cM in sunflower for hitchhiking to be relevant. The sunflower genome is ~3.5 Gb large (Baack et al. 2005), and based on a map length of 1400 cM (Tang et al. 2002), there are ~2.5 Mb/cM. HaFT2 and HaFT3 are indeed that closely linked, though HaFT2 is not present in the 10kb and 100kb of BAC sequence to either side of HaFT3. There is greater ambiguity on LG7 since HaTFL1 was not mapped on the same panel as the other selected genes, but several of those genes map within 1 cM of each other (Chapman et al. 2008). None of these genes are present in the ~60kb and ~38kb BAC of sequence flanking HaTFL1 though.

**Molecular signature of changing selection pressure**: A surprising result from our work was that all the members of the FT/TFL1 gene family tested were under selection at some stage of crop evolution. While HaFT3 is likely a false positive and the contributions of HaTFL1 and HaFT4 to particular QTL remain ambiguous, the timing of selection on these genes and their protein sequence or expression differences between wild and domesticated sunflower suggest an appealing evolutionary scenario. They provide a molecular signature of a reversal in the direction of selection on flowering time occurring over the history of sunflower cultivation.

The frameshift mutation in HaFT1 experienced selection during the initial stages of domestication, either due to direct selection for later flowering or selection on indirect developmental effects on other traits. Alternatively, selection may have favored direct
pleiotropic effects of the frameshift allele on other domestication traits, and later flowering was a byproduct of this selection. QTL for additional vegetative and reproductive traits map to this locus in multiple crosses (BAACK et al. 2008; BURKE et al. 2002; WILLS and BURKE 2007), and NILs segregating for the locus show genotype-dependent differences in some of these traits (BLACKMAN et al. 2010). Genetic evidence supports roles for FT in plant architecture, meristem size, leaf development, and abscission zone formation in Arabidopsis and tomato (JEONG and CLARK 2005; JEONG et al. 2008; KRIEGER et al. 2010; MELZER et al. 2008; SHALIT et al. 2009).

Modern sunflower breeders have imposed selection for early flowering. In-frame HaFT1 alleles may have been excluded by the genetic bottleneck at this stage of crop evolution, or selection may have continued to favor the frameshift mutation due to its effects on other traits. It is also possible that the fixation of additional mutations in HaFT1 or epistatic loci during early domestication also thwarted selection for a simple reversal of the frameshift in the domesticated background (BRIDGHAM et al. 2009). Consequently, variation in other genes responded to selection. Genetic redundancy, their similar positions to HaFT1 in the flowering time regulatory network, and their molecular mode of action may all have primed HaFT2, HaFT4, and HaTFL1 to respond most appropriately. Both FT and TFL1 homologs interact with FD in other organisms (ABE et al. 2005; PNUELI et al. 2001; WIGGE et al. 2005), and it has been hypothesized that phenotypic effects of these genes on flowering and growth depend on the outcome of competitive interactions between members of this family (AHN et al. 2006; KRIEGER et al. 2010; SHALIT et al. 2009). Increased expression of HaFT2 and HaFT4 are consistent with this scenario, particularly since there is evidence that HaFT2’s elevated expression level is controlled by differences in cis-regulatory elements (BLACKMAN et al. 2010). Expression of HaTFL1, a repressor of flowering, rose to higher levels earlier in development in CMSHA89 relative to
Ann1238 though, an observation inconsistent with this scenario unless *HaTFL1* expression is evolving to alleviate particular deleterious effects of increased *HaFT* expression.

Further functional work is required to support this scenario. Isolation of recombinants between *HaFT1* and *HaFT2* will be necessary to tease apart the contributions of the wild and domesticated alleles in these two genes to flowering. In addition, allele-specific expression analysis in additional elite x wild crosses will help determine how widespread putative *cis*-regulatory differences affecting *HaFT2*, *HaFT4*, and *HaTFL1* are.

Functional variants in *FT* orthologs contribute to natural variation in flowering time in cereal varieties and natural populations of *A. thaliana*. In cultivated rice, the promoter type of *Hd3a*, an *FT* ortholog, is a significant predictor of expression level and flowering time (Kojima et al. 2002; Takahashi et al. 2009). Alleles of *FT* ortholog, *VRN3*, in wheat and barley have non-coding variants that bypass a vernalization requirement for expression (Yan et al. 2006). Finally, allelic variation in *FT* and *TWIN SISTER OF FT* has been associated with natural variation in flowering among *A. thaliana* accessions (Atwell et al. 2010; Brachi et al. 2010; Schwartz et al. 2009). Notably, all of these previous findings involve *cis*-regulatory changes while our data suggest that both changes in expression and coding sequence of *FT* orthologs were involved in evolution of cultivated sunflower. Thus, our findings challenge the prediction that downstream genes like *FT* are more likely to exhibit regulatory variation (Schwartz et al. 2009). The redundancy afforded by a recent history of duplication, the novel derived expression domain of *HaFT1*, or multifarious selection acting on multiple pleiotropic effects may have uniquely fostered a selective sweep on a structural change in this case (Blackman et al. 2010); however given the expansion of *FT/TFL1* homolog numbers observed in other species (Carmel-
Goren et al. 2003; Danilevskaya et al. 2008; Faure et al. 2007; Igasaki et al. 2008; Mimida et al. 2009; Nishikawa et al. 2007), similar results may soon be observed in additional systems.

Acknowledgements: The authors thank A. Kozik and the Compositae Genome Project for providing BLAST summaries of EST library assemblies; D. Wills and A. Heesacker for providing genetic map data for various mapping panels prior to publication; Z. Lai, H. Luton, A. Posto, and K. Turner for technical assistance; the Indiana University Greenhouse staff; L. Washington and the Indiana Molecular Biology Institute for sequencing; and M. Hahn, L. Moyle, and past and present members of the Rieseberg, Michaels, and Moyle lab groups for helpful discussions. The research was supported by NSF (DBI0421630) and NIH (GM059065) grants to LHR; NSF (IOB0447583) and NIH (GM075060) grants to SDM; a NIH Ruth L. Kirschstein Postdoctoral Fellowship (5F32GM072409-02) to JLS; and an NSF Doctoral Dissertation Improvement Grant (DEB0608118) to BKB.
REFERENCES


CHARDON, F., B. VIRLON, L. MOREAU, M. FALQUE, J. JOETS et al., 2004 Genetic architecture of flowering time in maize as inferred from quantitative trait loci meta-analysis and Synteny conservation with the rice genome. Genetics 168: 2169-2185.


COLASANTI, J., R. TREMBLAY, A. Y. WONG, V. CONEVA, A. KOZAKI et al., 2006 The maize INDETERMINATE1 flowering time regulator defines a highly conserved zinc finger protein family in higher plants. BMC Genomics 7: 158.


COLLEDGE, S., and J. E. CONOLLY, 2007 The Origins and Spread of Domestic Plants in Southwest Asia and Europe. Left Coast Press, Walnut Creek, CA.


MIMIDA, N., N. KOTODA, T. UEDA, M. IGARASHI, Y. HATSUYAMA et al., 2009 Four TFL1/CEN-like genes on distinct linkage groups show different expression patterns to regulate vegetative and reproductive development in apple (Malus x domestica Borkh.). Plant Cell Physiol. 50: 394-412.


VALVERDE, F., A. MOURADOV, W. SOPPE, D. RAVENSCROFT, A. SAMACH et al., 2004

VIELLE-CALZADA, J.-P., O. MARTINEZ DE LA VEGA, G. HERNANDEZ-GUZMAN, E. IBARRA-LACLETTE, C. ALVAREZ-MEJIA et al., 2009 The Palomero genome suggests metal effects on domestication. Science 326: 1078-.

VIGOUROUX, Y., M. McMULLEN, C. T. HITTINGER, K. HOCHINS, L. SCHULZ et al., 2002


WILLIGE, B. C., S. GHOSH, C. NILL, M. ZOURELIDOU, E. M. DOHMANN et al., 2007 The DELLA domain of GA INSENSITIVE mediates the interaction with the GA INSENSITIVE DWARF1A gibberellin receptor of Arabidopsis. Plant Cell 19: 1209-1220.


Yamasaki, M., M. I. Tenaillon, I. V. Bi, S. G. Schroeder, H. Sanchez-Villeda et al., 2005 A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement. Plant Cell 17: 2859-2872.


### TABLE 1

**Sunflower Homologs of Flowering Time Genes**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Arabidopsis Homolog</th>
<th>Arabidopsis Locus ID</th>
<th>GenBank No.</th>
<th>% ID</th>
<th>% ID</th>
<th>Linkage Group</th>
<th>Panel</th>
<th>Closest Marker</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Photoperiod Pathway</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaFT1</td>
<td>FT</td>
<td>AT1G65480</td>
<td>DY917234</td>
<td>72.6</td>
<td>100</td>
<td>6</td>
<td>1</td>
<td>ORS483</td>
</tr>
<tr>
<td>HaFT2</td>
<td>FT</td>
<td>AT1G65480</td>
<td>EL485572</td>
<td>73.7</td>
<td>100</td>
<td>6</td>
<td>1</td>
<td>ORS483</td>
</tr>
<tr>
<td>HaFT3</td>
<td>FT</td>
<td>AT1G65480</td>
<td>-</td>
<td>71.4</td>
<td>100</td>
<td>6</td>
<td>2</td>
<td>ORS349</td>
</tr>
<tr>
<td>HaFT4</td>
<td>FT</td>
<td>AT1G65480</td>
<td>EL482916</td>
<td>73.7</td>
<td>100</td>
<td>14</td>
<td>2</td>
<td>HT765</td>
</tr>
<tr>
<td>HaCOL1</td>
<td>CO</td>
<td>AT5G15840</td>
<td>-</td>
<td>41.1</td>
<td>100</td>
<td>9</td>
<td>2</td>
<td>ORS1155</td>
</tr>
<tr>
<td>HaCOL2</td>
<td>CO</td>
<td>AT5G15840</td>
<td>DY912615</td>
<td>41.5</td>
<td>100</td>
<td>14</td>
<td>2</td>
<td>HT842</td>
</tr>
<tr>
<td>HaGI1</td>
<td>Gl</td>
<td>AT1G22770</td>
<td>DY913818</td>
<td>71.3</td>
<td>100</td>
<td>11</td>
<td>3</td>
<td>ORS228</td>
</tr>
<tr>
<td>HaGI2</td>
<td>Gl</td>
<td>AT1G22770</td>
<td>EL438742</td>
<td>57.8</td>
<td>27</td>
<td>10</td>
<td>1</td>
<td>CRT278</td>
</tr>
<tr>
<td>HaCDFL1</td>
<td>CDF1</td>
<td>AT5G62430</td>
<td>BU025202</td>
<td>37.0</td>
<td>100</td>
<td>7</td>
<td>1</td>
<td>ORS331</td>
</tr>
<tr>
<td>HaPHYB</td>
<td>PHYB</td>
<td>AT2G18790</td>
<td>DY908939</td>
<td>78.2</td>
<td>60</td>
<td>1</td>
<td>2</td>
<td>HT636</td>
</tr>
<tr>
<td>HaPHYC</td>
<td>PHYC</td>
<td>AT5G35840</td>
<td>EE622685</td>
<td>51.7</td>
<td>29</td>
<td>11</td>
<td>1</td>
<td>HT555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EE622732</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaCRY1</td>
<td>CRY1</td>
<td>AT4G08920</td>
<td>CF081828</td>
<td>30.5</td>
<td>22</td>
<td>5</td>
<td>3</td>
<td>ORS1153</td>
</tr>
<tr>
<td>HaCRY2</td>
<td>CRY2</td>
<td>AT1G04400</td>
<td>BQ970293</td>
<td>41.2</td>
<td>52</td>
<td>3</td>
<td>6</td>
<td>ORS1114</td>
</tr>
<tr>
<td>HaZTL</td>
<td>ZTL</td>
<td>AT5G57360</td>
<td>BU034705</td>
<td>91.1</td>
<td>31</td>
<td>3</td>
<td>2</td>
<td>HT745</td>
</tr>
<tr>
<td>HaFKF1</td>
<td>FKF1</td>
<td>AT1G68050</td>
<td>DY912573</td>
<td>74.9</td>
<td>100</td>
<td>17</td>
<td>2</td>
<td>ZVG80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Meristem Integrators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaTFL1</td>
<td>TFL1</td>
<td>AT5G03840</td>
<td>-</td>
<td>68.0</td>
<td>100</td>
<td>7</td>
<td>1</td>
<td>ORS331</td>
</tr>
<tr>
<td>HaSOC1</td>
<td>SOC1</td>
<td>AT2G45660</td>
<td>DY916215</td>
<td>60.7</td>
<td>100</td>
<td>6</td>
<td>1</td>
<td>ORS1229</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Protein</td>
<td>Accession Number</td>
<td>Identity (%)</td>
<td>Similarity (%)</td>
<td>Accession Number</td>
<td>Identity (%)</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>
</tr>
<tr>
<td>HAM75</td>
<td>AP1</td>
<td>AT1G69120</td>
<td>55.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100</td>
<td>8</td>
<td>2</td>
<td>ORS744</td>
<td></td>
</tr>
<tr>
<td>HaLFY</td>
<td>LFY</td>
<td>AT5G61850</td>
<td>61.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100</td>
<td>9</td>
<td>1</td>
<td>ORS844</td>
<td></td>
</tr>
<tr>
<td>HaSVP</td>
<td>SVP</td>
<td>AT2G22540</td>
<td>54.9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>100</td>
<td>5</td>
<td>1</td>
<td>HT440</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD848755</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DY916321</td>
<td></td>
</tr>
</tbody>
</table>

**Gibberellin Pathway**

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Protein</th>
<th>Accession Number</th>
<th>Identity (%)</th>
<th>Similarity (%)</th>
<th>Accession Number</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HaCPS</td>
<td>CPS</td>
<td>AT4G02780</td>
<td>42.6&lt;sup&gt;c&lt;/sup&gt;</td>
<td>47</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>HaKO</td>
<td>KO</td>
<td>AT5G25900</td>
<td>54.3&lt;sup&gt;a&lt;/sup&gt;</td>
<td>48</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>HaGA2ox</td>
<td>GA2ox</td>
<td>AT1G78440</td>
<td>55.3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD849186</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaGID1B</td>
<td>GID1B</td>
<td>AT3G63010</td>
<td>75.6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>37</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BU022119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaSLY1</td>
<td>SLY1</td>
<td>AT4G24210</td>
<td>50.0&lt;sup&gt;c&lt;/sup&gt;</td>
<td>100</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>HaDELLA1</td>
<td>RGA</td>
<td>AT2G01570</td>
<td>59.6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>100</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BU028290</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BQ912776</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BQ913059</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DY948837</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>DY916832</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaDELLA2</td>
<td>GAI</td>
<td>AT1G14920</td>
<td>64.1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>29</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD849186</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CD850340</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HaSPY</td>
<td>SPY</td>
<td>AT3G11540</td>
<td>67.3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>37</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

**Autonomous Pathway**

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Protein</th>
<th>Accession Number</th>
<th>Identity (%)</th>
<th>Similarity (%)</th>
<th>Accession Number</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HaFCA</td>
<td>FCA</td>
<td>AT4G16280</td>
<td>13.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>92</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>HaLD</td>
<td>LD</td>
<td>AT4G02560</td>
<td>39.8&lt;sup&gt;c&lt;/sup&gt;</td>
<td>39</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

---

49
GenBank accession numbers for sunflower EST sequences used for design of mapping primers.

Percent amino acid identity to *A. thaliana* protein sequence

Calculated using cDNA sequence obtained during this study from Ann1238

Calculated using EST sequence

Percent of the full *A. thaliana* cDNA covered by cDNA sequences from Ann1238 or EST sequences

Numeric code for mapping panels used to place genes on genetic map is as follows: 1) Hopi x Ann1238 F2, 2) NMS373 x Ann1811 BC1, 3) RHA280 x RHA801 RIL, 4) CMSHA89 x Ann1238 F3, 5) PHC x PHD RIL, 6) NMS801 x Arg1803 F1
# TABLE 2

Nucleotide Changes Present In Cross Parent Accession Coding Sequences

<table>
<thead>
<tr>
<th>Gene</th>
<th>Synonymous</th>
<th>Replacement</th>
<th>Deletion</th>
<th>Synonymous</th>
<th>Replacement</th>
<th>Deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>HaFT1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>HaFT2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HaFT4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HaCOL1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>HaCOL2</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>HaCDFL1</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>HaPHYB</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>HaFKF1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>HaTFL1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HaSOC1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>HAM75</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>HaLFY</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>HaCPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>HaSLY1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HaDELLA2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HaSPY</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>HaLD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
**TABLE 3**

Average sequence diversity values for candidate and reference genes

<table>
<thead>
<tr>
<th>Diversity Statistics</th>
<th>Elites</th>
<th>Landraces</th>
<th>Wilds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Candidate Genes (13 loci)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>35.7 (0.3)</td>
<td>37.9 (0.2)</td>
<td>44 (0.9)</td>
</tr>
<tr>
<td>L</td>
<td>686.3 (30.9)</td>
<td>683.9 (30.9)</td>
<td>671.5 (32.2)</td>
</tr>
<tr>
<td>S</td>
<td>6.6 (2.1)</td>
<td>14.9 (3.3)</td>
<td>43.5 (6.4)</td>
</tr>
<tr>
<td>h</td>
<td>2.4 (0.3)</td>
<td>7.5 (1.1)</td>
<td>23.6 (1.8)</td>
</tr>
<tr>
<td>(\pi)</td>
<td>0.0026 (0.0010)</td>
<td>0.0056 (0.0013)</td>
<td>0.0101 (0.0013)</td>
</tr>
<tr>
<td>(\theta_w)</td>
<td>0.0024 (0.0007)</td>
<td>0.0053 (0.0012)</td>
<td>0.0149 (0.0021)</td>
</tr>
</tbody>
</table>

| Reference Genes (7 loci) |              |              |              |
| n                       | 36 (0.0)     | 38 (0.0)     | 46 (0.0)     |
| L                       | 601.1 (38.6) | 593.6 (40.4) | 608 (41.9)   |
| S                       | 12.43 (1.2)  | 24.71 (4.3)  | 51.17 (6.5)  |
| h                       | 4.714 (0.9)  | 12 (2.4)     | 31.5 (2.9)   |
| \(\pi\)                | 0.0056 (0.0010) | 0.0095 (0.0013) | 0.0129 (0.0018) |
| \(\theta_w\)           | 0.005 (0.0006) | 0.01 (0.0016) | 0.02 (0.0028) |

Parameters listed include the average number of sequences (n), average sequence length (L), average number of segregating sites (S), average number of haplotypes (h), average pairwise nucleotide diversity (\(\pi\)), and average Watterson’s estimator of diversity (\(\theta_w\)). Values in parentheses are standard errors.
TABLE 4

Results of MLHKA Tests for Selection on Candidate Genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>MLHKA P Values&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Timing of Selection&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elites</td>
<td>Landraces</td>
</tr>
<tr>
<td>HaFT1</td>
<td>0.014*</td>
<td>0.045*</td>
</tr>
<tr>
<td>HaFT2</td>
<td>0.001***</td>
<td>0.123</td>
</tr>
<tr>
<td>HaFT3</td>
<td>0.001***</td>
<td>0.086</td>
</tr>
<tr>
<td>HaFT4</td>
<td>0.001**</td>
<td>0.122</td>
</tr>
<tr>
<td>HaCOL1</td>
<td>0.906</td>
<td>0.839</td>
</tr>
<tr>
<td>HaCOL2</td>
<td>0.330</td>
<td>0.533</td>
</tr>
<tr>
<td>HaCDFL1</td>
<td>0.909</td>
<td>0.922</td>
</tr>
<tr>
<td>HaPHYB</td>
<td>0.075</td>
<td>0.876</td>
</tr>
<tr>
<td>HaTFL1</td>
<td>0.001***</td>
<td>0.103</td>
</tr>
<tr>
<td>HaSOC1</td>
<td>0.841</td>
<td>0.914</td>
</tr>
<tr>
<td>HAM75</td>
<td>0.813</td>
<td>0.548</td>
</tr>
<tr>
<td>HaLFY</td>
<td>0.495</td>
<td>0.480</td>
</tr>
<tr>
<td>HaDELLA2</td>
<td>0.909</td>
<td>0.968</td>
</tr>
<tr>
<td>7 Reference</td>
<td>0.517</td>
<td>0.600</td>
</tr>
</tbody>
</table>

Loci<sup>c</sup>

<sup>a</sup>*, P < 0.05; **, P < 0.01; ***, P < 0.001

<sup>b</sup>Defined based on interpretation of MLHKA P values.

<sup>c</sup>P values for reference loci are reported for HKA results (see Materials and Methods)
FIGURE LEGENDS

Figure 1. Flowchart illustrating the criteria applied by integrated candidate gene approach and the serial refinement of the candidate gene pool.

Figure 2. Sunflower contains homologs of many genes in the A. thaliana flowering time gene regulatory network. Regulatory relationships between genes as determined in A. thaliana are depicted. Junctions indicate protein complex formation, dashed arrows indicate biosynthetic action, and a wavy arrow indicates protein transport between tissues. Genes without identifiable homologs in sunflower are italicized; genes with homologs genetically mapped in sunflower as part of this study are shown in bold.

Figure 3. Relative positions of candidate genes and flowering time QTL on the sunflower genetic map (BURKE et al. 2002). QTL positions and candidate gene positions were approximated by homothetic projection from markers concordant between multiple mapping populations as performed in BioMercator (Table S7; ARCADE et al. 2004). The BURKE et al. 2002 map was chosen as a framework since the most flowering QTLs and several candidate genes were mapped in this population. Only the nine LGs containing flowering time QTL of the 17 LGs in sunflower are shown. QTL regions, shown as 1-LOD intervals, detected in an elite x wild population in the greenhouse (BURKE et al. 2002) and the wild (BAACK et al. 2008) are shown in black and gray bars, respectively. QTL regions detected in a landrace x wild population grown in the greenhouse (WILLS and BURKE 2007) are shown as white bars.
Figure 4. Comparison of candidate gene expression patterns in leaf tissue between parent populations of an elite x wild QTL cross. Expression was assayed by RT-PCR in leaves collected every four hours starting at dawn over the course of a day from wild (Ann1238) and elite (CMSHA89) plants raised in long days and short days for 30 days after sowing. Data from one of three biological replicates are shown.

Figure 5. Comparison of candidate gene expression patterns in shoot apex tissue between parent populations of an elite x wild QTL cross. Expression was assayed by RT-PCR in shoot apices collected over a developmental time course for wild (Ann1238) and elite (CMSHA89) plants raised in long days. Plants were sown at day 0 and collection occurred every ten days until day 40 for both lines. A collection at 60 days was also conducted for Ann1238. In addition, some Ann1238 plants (T30) were transferred from long days to short days at 20 days, and shoot apices were collected from these plants on day 30. Data from one of three biological replicates are shown. HaFT1 data are from BLACKMAN et al. 2010.
FIGURE 1.

30 Flowering Time Gene Homologs

Maps to Linkage Group with Flowering Time QTL?

no

12 genes
HaGl1 HaCRY2 HaGA2ox
HaGl2 HaZTL HaGID1B
HaPHYC HaSVP HaDELLA1
HaCRY1 HaKO HaFCA

yes

18 genes

Notable Expression or cDNA Sequence Differences Between Cross Parents?

no

5 genes
HaFKF1 HaCPS HaSLY1 HaLD Ha Spy

yes

13 genes

Signature of Selection During Early Domestication or Improvement?

no

8 genes
HaCOL1 HaSOC1 HaCOL2 HAM75 HaCDF1 HaLFY HaPHYB HaDELLA2

yes

5 genes
HaFT1 HaFT2 HaFT3 HaFT4 HaTFL1
FIGURE 2.
FIGURE 4.

<table>
<thead>
<tr>
<th>h after dawn</th>
<th>Ann1238</th>
<th>CMSHA89</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short Days</td>
<td>Long Days</td>
</tr>
<tr>
<td>HaFT2</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaFT4</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaCOL1</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaCOL2</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaCDFL1</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaPHYB</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaFKF1</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaSOC1</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaCPS</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaDELLA2</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaSLY1</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>HaSPY</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>Ha60s rRNA</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
</tbody>
</table>
FIGURE 5.

<table>
<thead>
<tr>
<th>age</th>
<th>Ann1238</th>
<th>CMSHA89</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 10 20 30 40 60 T30</td>
<td>0 10 20 30 40</td>
</tr>
<tr>
<td>HaFT1</td>
<td><img src="image" alt="HaFT1 Image" /></td>
<td><img src="image" alt="HaFT1 Image" /></td>
</tr>
<tr>
<td>HaTFL1</td>
<td><img src="image" alt="HaTFL1 Image" /></td>
<td><img src="image" alt="HaTFL1 Image" /></td>
</tr>
<tr>
<td>HaSOC1</td>
<td><img src="image" alt="HaSOC1 Image" /></td>
<td><img src="image" alt="HaSOC1 Image" /></td>
</tr>
<tr>
<td>HAM75</td>
<td><img src="image" alt="HAM75 Image" /></td>
<td><img src="image" alt="HAM75 Image" /></td>
</tr>
<tr>
<td>HaLFY</td>
<td><img src="image" alt="HaLFY Image" /></td>
<td><img src="image" alt="HaLFY Image" /></td>
</tr>
<tr>
<td>HaDELLA2</td>
<td><img src="image" alt="HaDELLA2 Image" /></td>
<td><img src="image" alt="HaDELLA2 Image" /></td>
</tr>
<tr>
<td>HaLD</td>
<td><img src="image" alt="HaLD Image" /></td>
<td><img src="image" alt="HaLD Image" /></td>
</tr>
<tr>
<td>Ha60S rRNA</td>
<td><img src="image" alt="Ha60S rRNA Image" /></td>
<td><img src="image" alt="Ha60S rRNA Image" /></td>
</tr>
</tbody>
</table>