Genetica (2005) 123: 63–74 © Springer 2005

Studying genetics of adaptive variation in model organisms: flowering time variation in *Arabidopsis lyrata*

Mona Riihimäki[1], Robert Podolsky[1,2], Helmi Kuittinen[1], Hans Koelewijn[1,3] & Outi Savolainen[1,*]

[1]*Department of Biology, University of Oulu, FIN-90014 University of Oulu, Finland;* [2] *Present address: Office of Biostatistics and Bioinformatics, Medical College of Georgia, Augusta, GA 30912-4900, USA;* [3]*Present address: ALTERRA, Center for Ecosystem Studies, P.O. Box 47, 6700 AA Wageningen, The Netherlands; *Author for correspondence (Phone +358-8-553 1782; Fax +358-8-5531061; E-mail: outi.savolainen@oulu.fi)*

Received 3 March 2003 Accepted 26 November 2003

Key words: adaptation, *Arabidopsis lyrata, A. thaliana,* flowering time, genetic differentiation

Abstract

Arabidopsis thaliana has emerged as a model organism for plant developmental genetics, but it is also now being widely used for population genetic studies. Outcrossing relatives of *A. thaliana* are likely to provide suitable additional or alternative species for studies of evolutionary and population genetics. We have examined patterns of adaptive flowering time variation in the outcrossing, perennial *A. lyrata*. In addition, we examine the distribution of variation at marker genes in populations form North America and Europe. The probability of flowering in this species differs between southern and northern populations. Northern populations are much less likely to flower in short than in long days. A significant daylength by region interaction shows that the northern and southern populations respond differently to the daylength. The timing of flowering also differs between populations, and is made shorter by long days, and in some populations, by vernalization. North American and European populations show consistent genetic differentiation over microsatellite and isozyme loci and alcohol dehydrogenase sequences. Thus, the patterns of variation are quite different from those in *A. thaliana*, where flowering time differences show little relationship to latitude of origin and the genealogical trees of accessions vary depending on the genomic region studied. The genetic architecture of adaptation can be compared in these species with different life histories.

Introduction

Arabidopsis thaliana is the best known plant species in terms of its genome and molecular biology (Arabidopsis Genome Initiative, 2000). Its small genome and readily available mutants have made it a favorite organism for developmental and molecular genetic studies. Recently, the interest in the population genetics of *A. thaliana* has increased (Hanfstingl et al., 1994; Innan et al., 1996; Mitchell-Olds, 2001). At the same time, related species have begun to be seen also as potential model organisms. These relatives offer possibilities to study species with different life histories and the molecular genetic tools of *A. thaliana* can be often readily applied in the relatives (e.g., Kuittinen et al., 2002a). *A. lyrata* is a self-incompatible outcrossing species (Schierup, 1998; Kärkkäinen et al., 1999), to which the extensive population genetics theory of random mating populations can

Table 1. Comparison of *A. lyrata* and *A. thaliana* features

Trait	*A. lyrata*	*A. thaliana*	Reference
Outcrossing rate	1.0	0.02	Abbot and Gomes (1989) Kärkkäinen et al. (1999) and Schierup (1998)
Life cycle	Perennial	Annual	
Diploid genome size	0.46–0.51 pg	0.23–0.29	*Arabidopsis* Genome Initiative (2000), Earle (unpublished)
Chromosome #	8	5	Jones (1963)
Distribution	Palearctic, nearctic	Worldwide	

be applied. In outcrossing species, the different genes evolve more independently than in selfing species, where extensive linkage disequilibrium (LD) of genomes is maintained (Nordborg et al., 2002). The more independent variation of genes may make it easier to examine the evolution and its causes of individual genes. Further, *A. thaliana* is a weedy species, and outcrossing relatives may offer a possibility of studying populations where the effects of recent population expansions are not as much confounding in analyses of sequence variation. Third, for studies of local adaption, it may well be profitable to also use species that are not global generalist weeds.

In this paper, we examine the patterns of variation in one potentially adaptive trait, flowering time. Based on the life history differences between *A. thaliana* and *A. lyrata,* we can ask several questions. First, do the more stable, less weedy populations of *A. lyrata* show signs of local adaptation e.g. in flowering time, related to the environmental conditions. Do the populations of the outcrossing species have much variation within the populations, in comparison to the selfing *A. thaliana*. (e.g. Charlesworth & Charlesworth, 1995). Third, is the current distribution reflected in the genetic structure of *A. lyrata* populations? Do we find consistent patterns of genetic relationships between populations, using data from different parts of the genome. We address these questions with new data on the variation of flowering time, and with some new data and new analysis of earlier genetic markers and sequences. We discuss the implications of the differences between the species for the study of genetics of adaptation.

Materials and methods

Natural history of Arabidopsis lyrata

Arabidopsis lyrata is among the closest relatives to *A. thaliana* based on restriction fragment length polymorphism (RFLP) studies of cpDNA, and sequences of rbcL {Price, Palmer & Al-Shehbaz, 1994). Until recently, the two subspecies of *A. lyrata* (ssp. *lyrata* and ssp. *petraea)* were called *Arabis lyrata* and *Cardaminopsis petreaea,* but O'Kane and Al-Shebaz (1997) placed the species (and several others) in the genus *Arabidopsis*. This view of the systematics has been confirmed in many later studies of the Brassicaceae, using both cpDNA and nuclear sequences (Koch, Bishop & Mitchell-Olds, 1999; Koch, Haubold & Mitchell-Olds, 2000, 2001). The proportion of synonymous substitutions between the two species ranges between 10 and 15%, and for aminoacid changing nonsynonymous substitutions the divergence level is about 1–2%. Koch, Haubold and Mitchell-Olds (2000) have estimated a divergence time of about 5 MY for these two species based on *Adh* and *Chs* sequences.

The diploid genome size of *A. lyrata* (Swedish Mjällom and US Michigan populations) measured with flow cytometry is 0.46–0.51 pg, compared with the estimates for *A. thaliana* of 0.23–0.29 pg in the same set of measurements (Earle, pers. comm.). *A. lyrata* and other close relatives have eight chromosomes, against the five of *A. thaliana* (Jones, 1963). The two species can be crossed (Mesicek, 1967; Redei, 1974). Nasrallah et al. (2000) produced viable vigorous offspring from the

hybrid seeds after embryo rescue. In the backcross offspring of the hybrids, there was no evidence of crossing over between homeologous segments of the genomes of the two species (Nasrallah et al., 2000).

There are important life history differences. *A. thaliana* is annual, *A. lyrata* perennial (See Table 1). There is a well developed self-incompatibility system in *A. lyrata* (Kusaba et al., 2001), which gives rise to a fully outcrossing mating system (Schierup, 1998; Kärkkäinen et al., 1999). This difference is reflected also in the relatively large, pollinator-attracting petals of *A. lyrata* (see figures in Nasrallah et al., 2000). The species *A. lyrata* has a fragmented distribution in Europe, Japan and North America, with largely unknown distribution in Russia (Figure 1), references in Savolainen et al. (2000), whereas *A. thaliana* is a

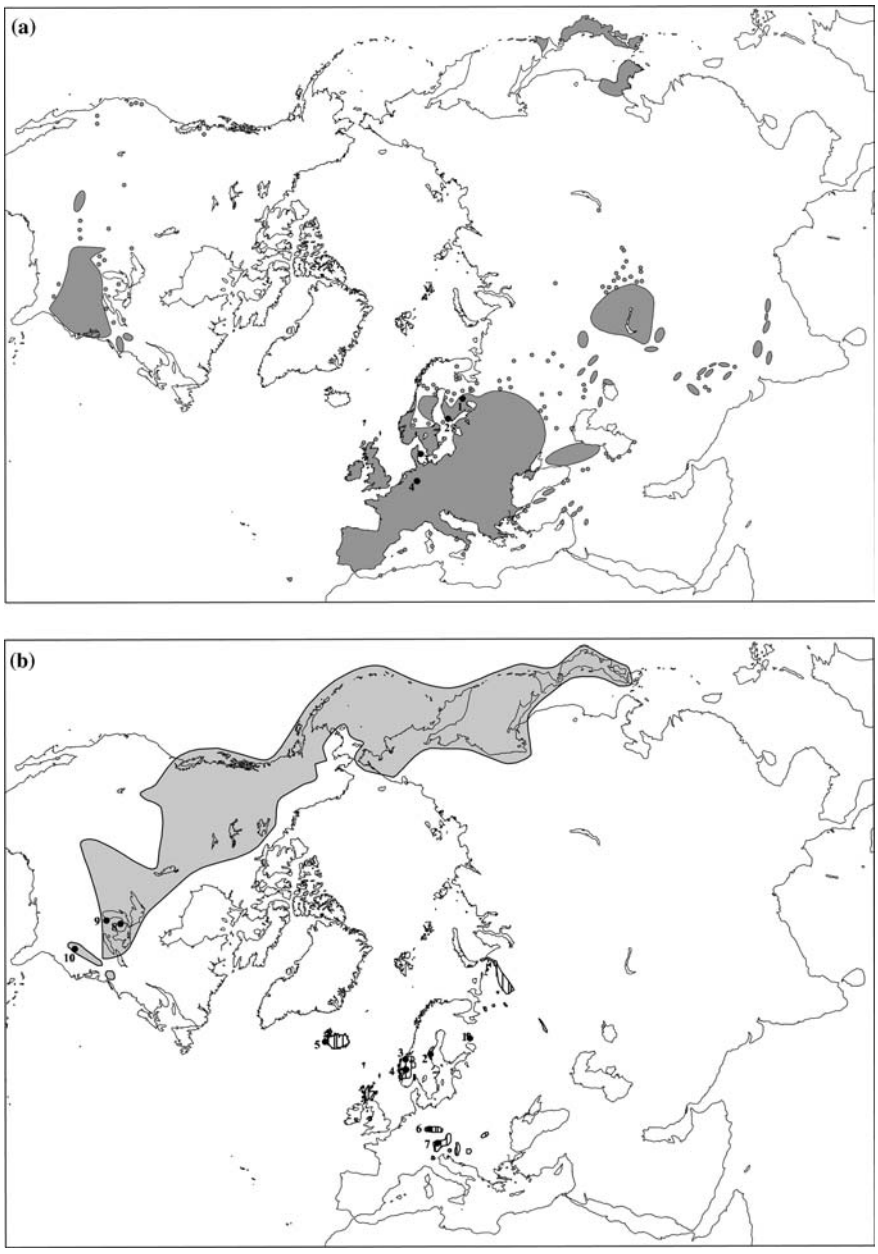

Figure 1. Distribution of (a) *A. thaliana* and (b) *A. lyrata.*

widespread weed. It has its origins in Asia and has spread to Europe (Price, Palmer & Al-Shehbaz, 1994), and has been introduced to other parts of the world, such as to the USA.

Measuring flowering time variation

We have examined the flowering variation in the perennial *Arabidopsis lyrata*. Six populations were chosen for the study (Plech, Germany 49°39′N. Bohemia, The Czech Republic 50°03′N, Spiterstulen, Norway 61°38′N, alt. 1100 m, Litldalen, Norway 62°32′N, Karhumäki, Russia 62°55′N, and Mjällom, Sweden 62°55′N). The seed samples were germinated and grown in long (LD, 20 h) and short (SD, 14 h) daylengths (+20°C). After 6 weeks of growth half of the plants from both daylengths were vernalized in +4°C, for 4 weeks. The nonvernalized plants were kept in +15°C to reduce growth. Both sets of plants received 8 h of light. After vernalization the plants were moved back to LD and SD at +20°C. In each of the four treatments, each population was represented by 12 plants.

We also grew a small set of crosses (12 females crossed each with four males) from the population of Karhumäki, Russia. The plants were not vernalized. They were grown under natural light conditions in the spring time in a greenhouse. The date when the first plant flowered was designated 1.

Statistical analyses

The flowering time data in the different environments were analyzed using the linear mixed effects model of R, after logarithmic transformation (Pinheiro & Bates, 2000; Team R Development Core Group, 2002). For the purposes of the analysis, the data from the four northern populations were combined to form a northern region, and the two southern populations were likewise combined to form a southern region. Region, daylength and vernalization were treated as fixed factors. The plants were randomized within daylengths on six trays. The tray was regarded as a random factor. The within population family data were also analyzed with ANOVA in R. Mothers and fathers were both treated as fixed effects.

The proportions of flowering could not be transformed to have normal distributions. Hence, we used a Bayesian generalized linear mixed model (GLMM) analysis for this kind of data. The analysis is implemented in the program WINBUGS. Rather than testing significance, the method results in an estimate of the probability that the factor in question has an effect (Clayton, 1996; Spiegelhalter, Thomas & Best, 2000).

Genetic markers and sequencing of A. lyrata

The methods for sequencing the alcoholdehydrogenase gene (*Adh*) of *A. lyrata* have been described by Savolainen et al. (2000). We obtained additional sequences from plants from Mayodan, North Carolina (seeds kindly provided by C.H. Langley) and from Mjällom, Sweden (see Van Treuren et al., 1997 for description of the locations). The earlier data of nine polymorphic enzyme and five microsatellite loci of Van Treuren et al. (Saitou & Nei, 1987) were also used for making genealogical trees of the populations. Neighbor-joining trees (Saitou & Nei, 1987) were constructed with the MEGA program version 1.3 (Kumar et al., 2001).

Results

Probability of flowering

We characterize the flowering of the populations in two ways, first the probability of flowering, and second, the time to flower formation. The measurements were made in four different environmental conditions, long and short days with and without vernalization. The Bayesian analysis of the probabilities showed that the northern and southern (regions) populations differed (Figure 2, Table 2). In short days, the southern populations of Plech and Bohemia were more likely to flower than any of the four northern populations. The northern populations were more likely to flower in long days than in short days. These different reactions to the conditions showed up as a significant interaction between region and daylength. Vernalization effects varied across daylengths and regions. It increased the probability of flowering in the northern populations in both short and long days, but did not have a consistent effect in the southern populations. This resulted in a significant interaction between vernalization and daylength. It should be noted that the results are based on

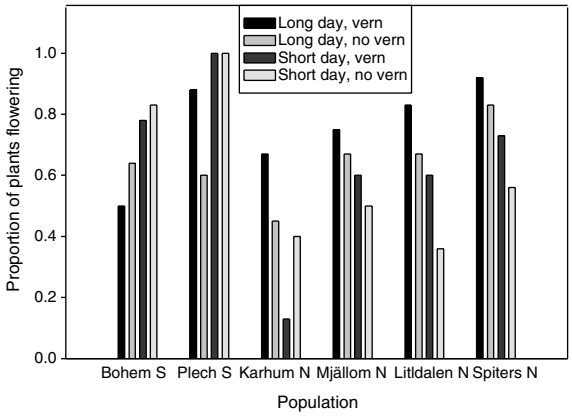

Figure 2. Proportion of plants from six different populations (N – northern, S – southern) flowering in the different day length – vernalization treatments. Long days (LD 20:4*)*, short days (LD 14:10), vernalization – rosette cold treatment during 4 weeks.

Table 2. Bayesian generalized linear mixed model analysis of flowering probability of *A. lyrata* using WinBUGS 3.1

Node	Probability
Region	0.97
Daylength	0.75
Vernalization	0.89
Reg × Dayl	0.98
Reg × Vern	0.67
Dayl × Vern	0.99
Reg × Dayl × Vern	0.79

Names of factors and the probability that the factor has an effect on probability of flowering.

rather small samples, and need to be confirmed in later studies.

Timing of flowering

The shortest flowering times were for southern populations in long days, less than a hundred days, while northern populations in short days could take more than 150 days to flower (Figure 3). In all environmental conditions, the two southern populations (Plech, Bohemia) flowered earlier than the northern populations (for region, $p < 0.001$). All populations also flowered more rapidly in the long days than in the short days. This effect was similar in all populations, with no interaction for region and daylength. Vernalization had an overall effect of speeding up flowering (Table 3), but this

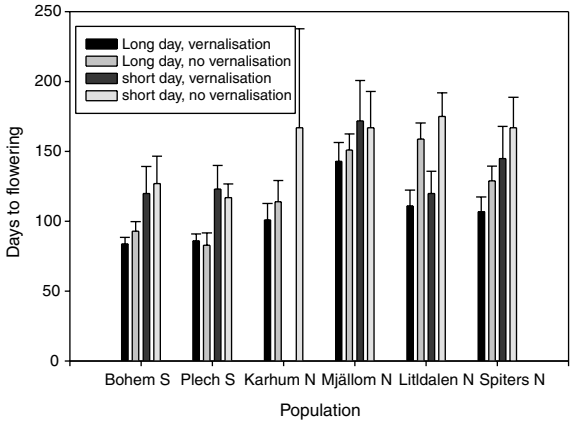

Figure 3. Flowering time of six populations of *A. lyrata* in different environmental conditions. Long days (LD 20:4), short days (LD 14:10), vernalization – rosette cold treatment during 4 weeks. Days to flowering (means and standard errors of the mean). (Too few plants flowered in Karhumäki, short days, vernalization – no result presented).

Table 3. Analysis of variance of flowering time of *A. lyrata* of the grouped northern and southern populations in four different environments

Effect	df	*F*	*p*
Region	1	25.78	0.001
Daylength	1	18.49	0.002
Vernalization	1	10.84	0.013
Reg × Dayl	1	0.98	0.320
Reg × Vern	1	3.66	0.050
Dayl × Vern	1	0.058	0.810
Reg × Dayl × Vern	1	0.002	0.966

effect was strongest in the northern populations of Spiterstulen and Litldalen, resulting in a region by vernalization interaction.

Variation in flowering time within the population

The average flowering time of individual families of Karhumäki, in long days, with no vernalization had a range of 25 days (the date when the first plant flowered was designated 1). In this pilot study, there were significant differences in flowering time both between the maternal ($F_{11,316} = 7.29$, $p < 0.001$) and paternal ($F_{3,\ 324} = 3.45$, $p < 0.02$) families. Figure 4 shows large maternal family influences, probably partly due to maternal effects,

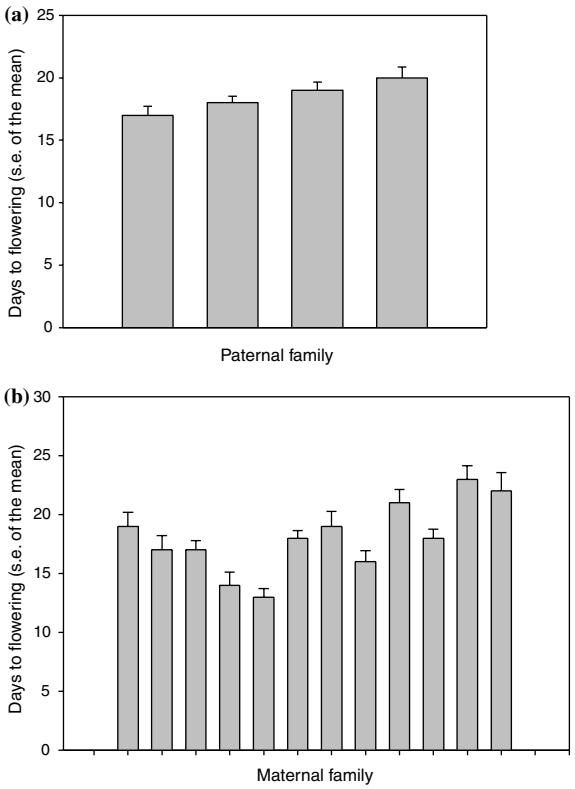

Figure 4. Flowering times for families of *A. lyrata* from Karhumäki, Russia in greenhouse conditions. (a) Means of maternal families (in days after first plant to flower) (± standard error of the mean); (b) paternal families (± standard error of the mean).

Table 4. ANOVA for flowering time variation within *A. lyrata* population of Karhumäki

Factor	df	Mean square	*F*	*p*
Mothers	11	229.3	7.29	0.001
Residuals	316	31.4		
Fathers	3	128.6	3.45	0.017
Residuals	324	37.2		

but the paternal family differences are evidence for genetic variation within the population.

Phylogeographic relationships between populations

Isozyme and microsatellite allele frequencies were available from four different populations (Van Treuren et al., 1997). In addition, we used the

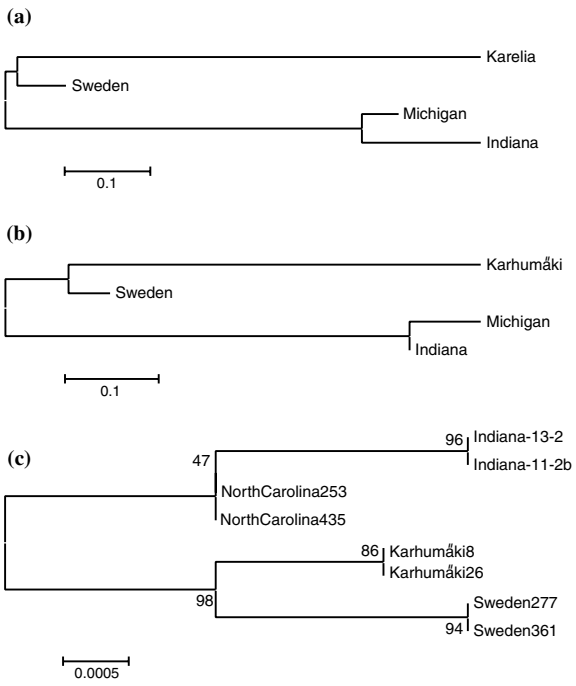

Figure 5. Neighbor-joining trees of *A. lyrata* based on (a) isozyme loci and (b) microsatellites (data of Van Treuren et al., 1997) and (c) *ADH* sequences from Savolainen et al. (2000) and additional sequences from North Carolina and Sweden (with bootstrap support).

alcoholdehydrogenase (*Adh*) sequences of (Savolainen et al., 2000), and some additional sequences obtained for the current purpose from Sweden and North Carolina. From these data, we constructed neighbor-joining trees shown in Figure 5. All data sets give a similar picture of the grouping of the North American and European populations. There is very high bootstrap support for this with the Adh sequences. The two north American populations Michigan and Indiana are rather close to each other based on microsatellites and allozymes, and the Adh sequences show that North Carolina also is not much diverged from Indiana.

Discussion

We have above described patterns of mainly between population variation in the outcrossing *Arabidopsis lyrata*. In comparing the patterns to *Arabidopsis thaliana,* we can test for effects of the outcrossing mating system, but these are

counfounded with the effects of demographic differences between the species.

Patterns of flowering time variation between populations

The set of six populations of *Arabidopsis lyrata* showed consistent differences between populations for both the probability to flower in different conditions and the time to flowering. Southern populations were more likely to flower and flowered more rapidly than the northern ones. Latitudinal clines in timing of reproduction or growth are common in many plant species (Mikola, 1982; Thomas & Vince-Prue, 1999). These patterns are interpreted as adaptations due to natural selection by climatic factors. The flowering time of *Arabidopsis thaliana* accessions has also been extensively studied. In these studies, it is rare that the plants would not flower at all, rather the lack of flowering of *A. lyrata* may correspond to very late flowering in *A. thaliana*. Based on the data of Karlson, Sills and Nienhuis (1993) we have plotted the flowering time (recorded as leaf number at flowering) of accessions against the latitude of origin (Figure 6), which shows that there is no clinal variation. The data of Nordborg and Bergelson (1999) showed a similar lack of clinal variation. Johanson et al. (2000) also did not find a strong relationship of flowering time with latitude. Stenøien et al. (2002) also failed to find clinal variation in populations collected in a south–north transect along the Norwegian coast, even if the same populations did show a cline in hypocotyl responses to red and far-red light. Thus, the early and later flowering of *A. thaliana* seems to be a reflection of whether the plants are winter annuals requiring vernalization or summer annuals without such a requirement. The quantitative variation among the genotypes requiring vernalization does not seem to be directly related to the length of the growing season (latitude of origin).

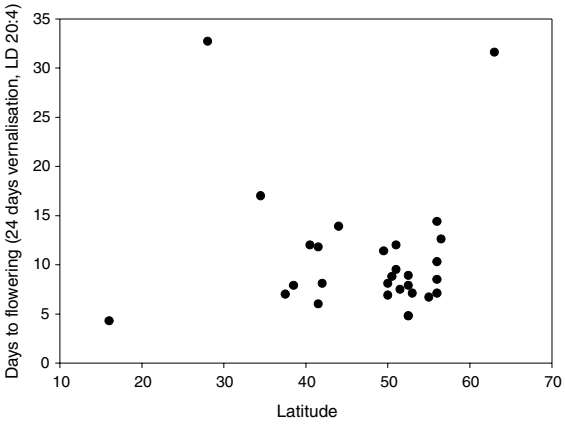

Figure 6. Variation of flowering time (measured as leaf number at flowering) in *Arabidopsis thaliana* in relation to latitude, based on data of Karlsson et al. (1993).

Environmental factors influencing flowering

We also gained some understanding on the factors controlling the probability to flower by growing the plants in several environments. The region by daylength interaction suggests that the southern and northern populations respond differentially to daylength, with northern populations more likely to flower in long days. In this experiment, vernalization had a stronger effect on the time to flowering rather than the probability to flower. In *A. thaliana,* the different accessions or ecotypes differ considerably with respect to vernalization response. It is well known that there are winter annual ecotypes requiring vernalization (e.g. Stockholm), late flowering summer annuals that flower faster after a cold treatment (such as Gr) and early flowering summer annuals which are not influenced by cold treatment (such as Li-5) (Zenker, 1955; Napp-Zinn, 1957). Napp-Zin (1957) already identified the locus *FRI*. This gene has been recently cloned and its role in determining flowering time differences in the wild between winter and summer annual ecotypes has been examined in detail (Johanson et al., 2000).

However, as mentioned, the distribution of these ecotypes is not related to latitudinal climatic variation. All populations do eventually flower even in the absence of vernalization. Interestingly, a third close relative, *A. hirsuta,* seems not to flower at all without a vernalization treatment (Zenker, 1955). Thus, in the related species the relative importance of the different pathways may vary.

The *A. thaliana* ecotypes also have variable responses to photoperiod (Karlsson, Sills & Nienhuis 1993), and in their G × E interactions (Pigliucci, Pollard & Cruzan, 2003). The photoperiodic pathway of *Arabidopsis thaliana* and its relationship to flowering time control has been

well described (e.g. Koornneef, Hanhart & van der Veen, 1991; Suarez-López et al., 2001). Developmental studies of the gene CONSTANS have shown that it has an important role (Putterill et al., 1995; Yanovsky & McKay, 2002). Further, El-Assal et al. (2001) recently demonstrated that the *CRY2* cryptochrome locus is largely responsible for a flowering time difference between two early flowering accessions. QTL studies have identified other loci in crosses between summer annuals (Jansen et al., 1995). In addition, phytochrome *A* has been shown to influence flowering time differences between natural populations of *A. thaliana* (Maloof et al., 2001). The initial results on *A. lyrata*, in combination with the well known pathways of *A. thaliana*, suggest further studies on the genetic mechanisms governing these differences.

Variation within populations

We also demonstrated that there are quantitative genetic differences between families in the Russian Karhumäki population, when plants were grown under long days without vernalization. These findings are consistent with the existence of considerable within population genetic variation, as has been found earlier for marker genes (Van Treuren et al., 1997; Schierup, 1998; Clauss, Cobban & Mitchell-Olds, 2002) and for sequence variation at the *Adh* gene (Savolainen et al., 2000). *Arabidopsis thaliana* populations have been examined only rarely for quantitative genetic variation. Early British studies found evidence of segregating major gene variation (Westerman & Lawrence, 1970; Jones, 1971b, a), presumably due to the FRl gene (Johanson et al., 2000). Kuittinen, Mattila and Savolainen (1997) found that many marginal populations had no variation for flowering time. Likewise, the within population variation in microsatellites or isozymes has been found to be low (Abbott & Gomes, 1989; Todokoro, Terauchi & Kawano, 1996), as well as in restriction fragment length polymorphism (RFLP) studies (Bergelson et al., 1998) and studies of sequence variation (Stahl et al., 1999; Kuittinen, Salguero & Aguadé, 2002b).

The reduced level of genetic variation in flowering time and other traits found in at least some populations could be due to the effects of the mating system and a reduction of effective population size, due to background selection or hitchhiking (Kaplan, Hudson & Langley, 1989; Charlesworth, Morgan & Charlesworth, 1993; Charlesworth & Charlesworth, 1995). In addition to this, the weedy life history of *Arabidopsis thaliana* may also give rise to extinctions and recolonizations. The metapopulation structure is expected to lead to much reduced variation within populations, beyond the mere effects of selfing (Ingvarsson, 2002). *A. lyrata* in turn is a perennial, and is less likely to suffer frequent population extinctions.

Population history in the light of distribution of marker and sequence genetic variation

The small set of populations that was studied in *A. lyrata* demonstrated that isozymes (nine loci), microsatellites (five loci) and sequence variation at the alcoholdehydrogenase (*Adh*) locus (1700 nt) all result in a clear separation of the North American and European populations. The *Adh* trees also show that the variation between populations is high relative to within population variation, as was found earlier for isozymes and microsatellites (Van Treuren et al., 1997). These sets of populations have evidently been isolated for considerable time. When we use all our available *Adh* data (34 sequences from North America, 15 from Europe), we obtain a net divergence d_A of 0.0033 (Nei & Kumar, 2000). If we use the rate of synonymous substitution at the *Adh* locus suggested by Koch, Haubold and Mitchell-Olds (2000) of 1.5×10^{-8}/bp/year, we obtain a rough estimate of separation of at least 100,000 years for the North American and European populations. The European populations are more similar to each other, as are the two North American ones.

This pattern is in strong contrast to the situation found in *A. thaliana*. Most studies of molecular genetic variation in *A. thaliana* have been based on examining a set of accessions collected from around the world (Hanfstingl et al., 1994; Innan, Terauchi & Miyashita, 1997; Miyashita, Kawabe & Innan, 1999; Sharbel, Haubold & Mitchell-Olds, 2000). Several loci have shown a pattern of strong dimorphism, with two divergent haplotypes (e.g. Aguadé, 2001), whereas others show no such pattern. Evidence of recombination has been found in all the genes examined to date (Innan et al., 1996). Gene genealogies of

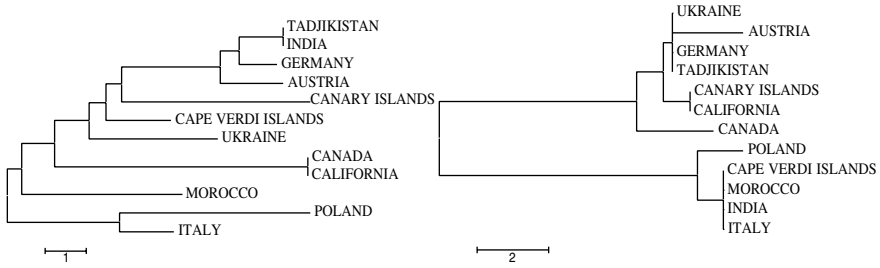

Figure 7. Neighbor-joining trees of *A. thaliana* accessions, on top *CHI,* based on data of Kuittinen et al. (2000) and on the bottom, *FAH,* based on the data of Aguadé (2001). The trees show the geographical areas, the accession names can be found in the original papers. The scale shows numbers of nucleotide substitutions.

accessions from different geographical areas based on variation at the different loci do not show geographic consistency. Figure 7 shows examples of the genealogies for the dimorphic *FAH1* and the nondimorphic *CHI* based on data of Kuittinen and Aguadé (2000) and Aguadé (2001). Several studies (e.g. Sharbel, Haubold & Mitchell-Olds, 2000) suggest that the population has expanded recently. Thus, there seems to be no one genealogical tree of the accessions or populations, an important feature of *A. thaliana*.

Implications for studying the molecular basis of adaptation

Most studies on the genetic basis of quantitative variation in plants have been on cultivated species, such as *Brassicas* (Lagergrantz et al., 1996; Lagercrantz, 1998), where domestication may have influenced patterns of variation. *A. lyrata* and other relatives of *A. thaliana* offer many opportunities to the study of adaptation in natural population, with variable mating systems and life histories.

The mating system is one of the key determinants of plant population genetics (Hamrick & Godt, 1996), and potentially modes of adaptation. Population genetics theory has several specific predictions about the expected levels of neutral variation within and between populations (Charlesworth, Morgan & Charlesworth, 1993). A comparison of the closely related species *A. thaliana* and *A.lyrata* allows investigation of the effects of the mating system of patterns of sequence evolution (Savolainen et al., 2000; Wright, Lauga & Charlesworth, 2002). The mating system effects, however, are also confounded with other life history traits, for instance as perenniality and other demographic aspects, such as the level of migration or the occurrence of extinction/colonization cycles (Pannell & Charlesworth, 1999). Variable selfing and a possible metapopulation structure add complexity to the models (Nordborg & Donnelly, 1997; Pannell & Charlesworth, 2000; Wakeley & Aliacar, 2001). Interpreting the effects of natural selection against a background of other evolutionary forces, such as effects of history, genetic drift, selection at other linked loci may be easier in random mating species as the population genetical theory for random mating populations with reasonable constant size is well developed (e.g. Hudson, 1990).

The mating system also influences patterns of linkage disequilibrium, i.e. statistical association between alleles at different loci or nucleotide sites. LD has become an important tool in genetic mapping of human diseases (Nordborg & Tavaré, 2002; Weiss & Clark, 2002) or loci responsible for quantitative genetic variation in plants (Thornsberry et al., 2001). This technique relies on examining the association of densely situated single nucleotide polymorhisms (SNPs) and phenotypic traits. SNPs close or at the disease/phenotype causing nucleotide site will be in disequilibrium, while those further away will show less association. Selfing species such as *A. thaliana* are expected to have high LD because of little effective recombination in mostly homozygous individuals (Allard, Jain & Workman, 1968). Recently, Nordborg et al. (2002) found that extent in a global sample *A. thaliana,* LD decayed over 250 kb, indicating that recombination has occurred over the long time span represented by this sample. In a local sample, LD extended over whole chromosomes, as there had been little breakdown in disequilibrium over the short time span represented by these collections. Association mapping

cannot be used within local populations, as the linkage disequilibrium will be uniformly high across large parts of chromosomes. Worldwide samples will have the necessary structure of declining disequilibrium, but in such a sample the quantitative traits may be genetically heterogeneous (Nordborg et al., 2002). Disequilibrium within populations of *A. lyrata* will decline much more rapidly than in *A. thaliana*. Then associations of nucleotide variation with the phenotypic variation could be studied at a smaller scale, utilizing also the within population phenotypic variation that has been demonstrated above.

Acknowledgments

We acknowledge the financial support of the Research council for Environment and Biosciences of Finland, and the Graduate School for Forest Sciences for financial support. We thank Mark McNair, Maria Clauss, Katri Kärkkäinen and C. H. Langley for help with seeds.

References

Abbott, R.J. & M.F. Gomes, 1989. Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh. Heredity 62: 411–418.

Aguadé, M., 2001. Nucleotide sequence variation at two genes of the phenylpropanoid pathway, the *FAHl* and *F3H* genes, in *Arabidopsis thaliana*. Mol. BioI. Evol. 18: 1–9.

Allard, R.W., S.K. Jain & P. Workman, 1968. The genetics of inbreeding species. Adv. Genetics 14: 55–131.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815.

Bergelson, J., E. Stahl, S. Dudek & M. Kreitman, 1998. Genetic variation within and among populations of *Arabidopsis thaliana*. Genetics 148: 1311–1323.

Charlesworth, B., M.T. Morgan & D. Charlesworth, 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134: 1289–1303.

Charlesworth, D. & B. Charlesworth, 1995. Quantitative genetics of plants: the effect of the breeding system on genetic variability. Evolution 49: 911–920.

Clauss, M.J., H. Cobban & T. Mitchell-Olds, 2002. Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae). Mol. Ecol. 11: 591–601.

Clayton, D., 1996. Generalized linear models, pp. 275–301 in Markov Chain Monte Carlo Methods in Practice, edited by W.R. Gilks, S. Richardson & D.J. Spiegelhalter. Chapman and Hall, London.

El-Assal, S.E.D., A.-B.C., A.J.M. Peeters, V. Raz & M. Koorneef, 2001. A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. Nature Genetics 29: 435–440.

Hamrick, J.L. & M.J.W. Godt, 1996. Effects of life history traits on genetic diversity in plant species. Phil. Trans. R. Soc. Lond. B 351: 1291–1298.

Hanfstingl, U., A. Berry, E.A. Kellogg, J.T. Costa, W. Ruediger & F.M. Ausubel, 1994. Haplotypic divergence coupled with lack of diversity at the *Arabidopsis thaliana* alcohol dehydrogenase locus: roles for both balancing and directional selection? Genetics 138: 811–828.

Hudson, R.R., 1990. Gene genealogies and the coalescent process, pp. 1–44 in Oxford Surveys in Evolutionary Biology, edited by D. Futuyma & J. Antonovics. Oxford.

Ingvarsson, P.K., 2002. A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. Evolution 56: 2368–2373.

Innan, H., F. Tajima, R. Terauchi & N.T. Miyashita, 1996. Intragenic recombination in the *Adh* locus of the wild plant *Arabidopsis thaliana*. Genetics 143: 1761–1770.

Innan, H., R. Terauchi & N. Miyashita, 1997. Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146: 1441–1731.

Jansen, R., J. Van Ooijen, P. Starn, C. Lister & C. Dean, 1995. Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. Theor. Appl. Genet. 91: 33–37.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino & C. Dean, 2000. Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290: 344–347.

Jones, B.M.G., 1963 Experimental Taxonomy of the Genus *Arabis*. University of Leicester, UK, Leicester.

Jones, M.E., 1971a. The population genetics of *Arabidopsis thaliana*. II. Population structure. Heredity 27: 51–58.

Jones, M.E., 1971b. The population genetics of *Arabidopsis thaliana*. I. The breeding system. Heredity 27: 39–50.

Kaplan, N.L., R.R. Hudson & C.H. Langley, 1989. The 'hitchhiking' effect revisited. Genetics 123: 887–899.

Kärkkäinen, K., H. Kuittinen, R. van Treuren, C. Vogl, S. Oikarinen & O. Savolainen, 1999. Genetic basis of inbreeding depression in *Arabis petraea*. Evolution 53: 1354–1365.

Karlsson, B.H., G.R. Sills & J. Nienhuis, 1993. Effects of photoperiod and vernalization on the number of leaves at flowering in 32 *Arabidopsis thaliana* (Brassicaceae) ecotypes. Amer. J. Bot. 80: 646–648.

Koch, M., J. Bishop & T. Mitchell-Olds, 1999. Molecular systematics and evolution of *Arabidopsis* and *Arabis*. Plant Biol. 1: 529–537.

Koch, M., R Haubold & T. Mitchell-Olds, 2001. Molecular systematics of the Brassicaceae: evidence from coding platidic *MatK* and nuclear *CHS*. Amer. J. Bot. 88: 534–544.

Koch, M.A., B. Haubold & T. Mitchell-Olds, 2000. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis, Arabis,* and related genera (Brassicaceae). Mol. Biol. Evol. 17: 1483–1498.

Koornneef, M., C.J. Hanhart & J.H. van der Veen, 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol. Gen. Genet. 229: 57–66.

Kuittinen, H. & M. Aguadé, 2000. Nucleotide variation at the *Chalcone Isomerase* locus in *Arabidopsis thaliana*. Genetics 155: 863–872.

Kuittinen, H., M. Aguadé, D. Charlesworth, A. De Haan, B. Lauga, T. Mitchell-Olds, S. Oikarinen, S. Ramos-Onsins,

B. Stranger, P. van Tienderen & O. Savolainen, 2002a. Primers for 22 candidate genes for ecological adaptations in Brassicaceae. Mol. Ecol. Notes 2: 258–262.

Kuittinen, H., A. Mattila & O. Savolainen, 1997. Genetic variation at marker loci and in quantitative traits in natural populations of *Arabidopsis thaliana*. Heredity 79: 144–152.

Kuittinen, H., D. Salguero & M. Aguadé, 2002b. Parallel patterns of sequence variation within and between populations at three loci of *Arabidops*is *thaliana*. Mol. Biol. Evol. 19: 2030–2034.

Kumar, S., K. Tamura, LB. Jakobsen & M. Nei, 2001. MEGA2: Molecular Evolutionary Genetics Analysis Software. Arizona State University, Tempe, Arizona.

Kusaba, M., K. Dwyer, J. Hendershot, J. Vrebalov, J.B. Nasrallah & M.E. Nasrallah, 2001. Self-incompatibility in the genus *Arabidopsis*: characterization of the S locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. Plant Cell 13: 627–643.

Lagercrantz, U., 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150: 1217–1228.

Lagercrantz, U., J. Putterill, G. Coupland & D. Lydiate, 1996. Comparative mapping in *Arabidopsis* and *Brassica,* fine scale genome collinearity and congruence of genes controlling flowering time. Plant J. 19: 13–20.

Maloof, J., J. Borevitz, T. Dabi, J. Lutes, R. Nehring, J. Redfern, G. Trainer, J. Wilson, T. Asami, C. Berry, D. Weigel & J. Chory, 2001. Natural variation in light sensitivity of *Arabidopsis*. Nat. Genet. 29: 441–446.

Mesicek, J., 1967. The chromosome morphology of *Arabidopsis thaliana* (L.) Heynh. and some remarks on the problem of *Hylandra suecica* (Fr.) Love. Folia Geobot. Phytotaxonom. 2: 433–436.

Mikola, J., 1982. Bud-set phenology as an indicator of climatic adaptation of Scots pine in Finland. Silva Fennica 16: 178–184.

Mitchell-Olds, T., 2001. *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. Trends Ecol. Evol. 16: 693–700.

Miyashita, N.T., A. Kawabe & H. Innan, 1999. DNA variation in the wild plant *Arabidopsis thaliana* revealed by amplified fragment length polymorphism analysis. Genetics 152: 1723–1731.

Napp-Zinn, K., 1957. Untersuchungen zur Genetik des Kältebedurfnisses bei *Arabidopsis thaliana*. Zeitschrift fur indukt. Abstammungs- und Vererbungslehre 88: 253–285.

Nasrallah, M.E., K. Yogeeswaran, S. Snyder & J.B. Nasrallah, 2000. *Arabidopsis* species hybrids in the study of species differences and evolution of amphiploidy in plants. Plant Physiol. 124: 1605–1614.

Nei, M. & S. Kumar, 2000. Molecular Evolutionary Phylogenetics. Oxford University Press, Oxford.

Nordborg, M. & J. Bergelson, 1999. The effect of seed and rosette cold treatment on germination and flowering time in some *Arabidopsis thaliana* (Brassicaceae) ecotypes. Am. J. Bot. 86: 470–475.

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

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

Nordborg, M. & S. Tavaré, 2002. Linkage disequilibrium: what history has to tell us. Trends Genet. 18: 83–90.

O'Kane, S.L. & LA. AI-Shehbaz, 1997. A synopsis of *Arabidopsis* (Brassicaceae). Novon 7: 323–327.

Pannell, J.R. & B. Charlesworth, 1999. Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. Evolution 53: 664–676.

Pannell, J.R. & B. Charlesworth, 2000. Effects of metapopulation processes on measures of genetic diversity. Phil. Trans. R. Soc. Lond. B 355: 1851–1864.

Pigliucci, M., H. Pollard & M.B. Cruzan, 2003. Comparative studies of evolutionary responses to light environments in *Arabidopsis*. Amer. Natur 161: 68–82.

Pinheiro, J.C. & D.M. Bates, 2000. Mixed Effects Models in Sand S-PLUS. Springer Verlag, New York.

Price, R.A., J.D. Palmer & I.A. Al-Shehbaz, 1994. Systematic relationships of *Arabidopsis*, pp. 7–19 in *Arabidopsis*, edited by E.M. Meyerowitz & C. Somerville. Cold Spring Harbor Press, Cold Spring Harbor, NY.

Putterill, J., F. Robson, K. Lee, R. Simon & G. Coupland, 1995. The CONSTANS gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. Cell 80: 847–857.

R Development Core Team. 2002. R: a language and environment for statistical computing. Vienna, Austria.

Redei, G.P., 1974. Is *Hylandra* an amphidiploid of *Arabidopsis* and *Cardaminopsis arenosa? Arabidopsis* Inf. Servo 11: 5.

Saitou & Nei, 1987. The nieghbor-joining method: a new method for constructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

Savolainen, O., C.H. Langley, B. Lazzaro & H. Freville, 2000. Contrasting patterns of nucleotide variation at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. Mol. Biol. Evol. 17: 645–655.

Schierup, M.H., 1998. The effect of enzyme heterozygosity on growth in a strictly outcrossing species, the self-incompatible *Arabis petraea* (Brassicaceae). Hereditas 128: 21–31.

Sharbel, T.F., B. Haubold & T. Mitchell-Olds, 2000. Genetic isolation by distance in *Arabidopsis thaliana:* biogeography and postglacial colonization of Europe. Mol. Ecol. 9: 2109–2118.

Spiegelhalter, D.J., A. Thomas & N.G. Best, 2000. WinBUGS Version 1.3. User Manual. MRC Biostatistics Unit, Cambridge, UK.

Stahl, E.A., G. Dwyer, R. Mauricio, R. Kreitman & J. Bergelson, 1999. Dynamics of disease resistance polymorphism at the *Rpml* locus of *Arabidopsis*. Nature 400: 667–671.

Stenøien, H., C. Fenster, H. Kuittinen & O. Savolainen, 2002. Quantifying latitudinal clines to light responses in natural populations of *Arabidopsis thaliana* (Brassicaceae). Am. J. Bot.

Suarez-López, P., K. Wheatley, F. Robson, H. Onouchi, F. Valverde & G. Coupland, 2001. CONSTANS mediates

between the cirdacian clock and the control of flowering time in *Arabidopsis*. Nature 410: 1116–1120.

Thomas, B. & D. Vince-Prue, 1999. Photoperiodism in Plants. Academic Press, San Diego.

Thornsberry, J.M., M.M. Goodman, J. Doebley , S. Kresovich, D. Nielsen & E.S. Buckler, 2001. Dwarf8 polymorphisms associate with variation in flowering time. Nature Genet. 28: 286–289.

Todokoro, S., R. Terauchi & S. Kawano, 1996. Microsatellite polymorphism in natural populations of *Arabidopsis thaliana* in Japan. Jpn. J. Genet. 70: 543–554.

Van Treuren, R., H. Kuittinen, K. Kärkkäinen, E. Baena-Gonzalez & O. Savolainen, 1997. Evolution of microsatellites in *Arabis petraea* and *A. lyrata,* outcrossing relatives of *Arabidopsis thaliana*. Mol. BioI. Evol. 14: 220–229.

Wakeley, J. & N. Aliacar, 2001. Gene genealogies in a metapopulation. Genetics 159: 893–905.

Weiss, K.M. & A.G. Clark, 2002. Linkage disequilibrium and the mapping of complex human traits. Trends Genet. 18: 19–24.

Westerman, J.M. & M.J. Lawrence, 1970. Genotype-environment interaction and developmental regulation in *Arabidopsis thaliana*. I. Inbred lines; description. Heredity 25: 609–627.

Wright, S.I., B. Lauga & D. Charlesworth, 2002. Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. Mol. Biol. Evol. 19: 1407–1420.

Yanovsky, M.J. & S.A. McKay, 2002. Molecular basis of seasonal time measurement in *Arabidopsis*. Nature 419: 308–312.

Zenker, A.M., 1955. Jarowisationsuntersuchungen an sommerannuellen *Arabidopsis* Rassen. Beitr. Biol. Pflantz. 32: 135–170.