OULU 2014
A 643

ACTA UNIVERSITATIS OULUENSIS

Tuomas Toivainen

GENETIC CONSEQUENCES OF DIRECTIONAL SELECTION IN *ARABIDOPSIS LYRATA*

UNIVERSITY OF OULU GRADUATE SCHOOL;
UNIVERSITY OF OULU,
FACULTY OF SCIENCE,
DEPARTMENT OF BIOLOGY;
BIOCENTER OULU

ACTA UNIVERSITATIS OULUENSIS
A Scientiae Rerum Naturalium 643

TUOMAS TOIVAINEN

GENETIC CONSEQUENCES OF DIRECTIONAL SELECTION IN *ARABIDOPSIS LYRATA*

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Health and Biosciences of the University of Oulu for public defence in Kuusamonsali (YB210), Linnanmaa, on 11 December 2014, at 12 noon

UNIVERSITY OF OULU, OULU 2014

Acta Univ. Oul. A 643, 2014

Supervised by
Professor Outi Savolainen
Docent Helmi Kuittinen
Docent Tanja Pyhäjärvi

Reviewed by
Professor Johanna Vilkki
Docent Irma Saloniemi

Opponent
Doctor Thomas Källman

ISBN 978-952-62-0689-9 (Paperback)
ISBN 978-952-62-0690-5 (PDF)

ISSN 0355-3191 (Printed)
ISSN 1796-220X (Online)

Cover Design
Raimo Ahonen

JUVENES PRINT
TAMPERE 2014

Toivainen, Tuomas, Genetic consequences of directional selection in *Arabidopsis lyrata*.
University of Oulu Graduate School; University of Oulu, Faculty of Science, Department of Biology; Biocenter Oulu
Acta Univ. Oul. A 643, 2014
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract

Plants and animals colonized Northern Europe after the last Ice Age from different refugia, not covered by the ice sheet. Many plants, such as the northern rock cress (*Arabidopsis lyrata* ssp. *petraea*) adapted to the short growing season in the North. We thus expect that colonization of the new environment was accompanied by directional selection for traits conferring this adaptation. In this thesis I studied whether recent directional selection can be detected in two important genes, *PHYTOCHROME A* (*PHYA*) and *FLOWERING LOCUS C1* (*FLC1*), related to the flowering time pathway. To detect directional selection, I compared DNA sequence variation from the samples of a southern (Plech, Germany) and a northern (Spiterstulen, Norway) population. I also studied the current response potential to changing conditions in the marginal Spiterstulen population. Adaptation potential was characterized by assessing plasticity and amount of additive genetic variation, focusing on flowering traits. In addition, associations of 21 flowering time candidate genes for phenological and fitness traits were studied.

There were several lines of evidence for recent directional selection in both candidate genes, *PHYA* and *FLC1*, in the northern Spiterstulen population Variation was strongly reduced around both genes and in addition they were highly differentiated between populations. In the Spiterstulen population there was a remarkable reduction in additive genetic variation for flowering traits, for instance when compared with morphological traits. On the other hand, phenological traits showed high plasticity. Some of the photoperiodic pathway genes showed association to flowering or reproductive fitness.

The results suggest that directional selection during the colonization of the northern areas has impacted the two studied genes. Genetic changes were likely involved in altered photoperiodic and vernalization responses which might be adaptive for a short growing season. Further, directional selection was probably responsible for reducing additive genetic variation in flowering traits. Because there was only little genetic variation, adaptation to future environmental change of the marginal Spiterstulen population is likely to rely largely on plastic reactions to environmental signals, or tracking the environment by dispersal.

Keywords: Arabidopsis lyrata, association mapping, FLC, flowering time, phenotypic plasticity, PHYA, response potential, selective sweep

Toivainen, Tuomas, Suuntaavan valinnan geneettiset seuraukset *Arabidopsis lyratalla*.
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Luonnontieteellinen tiedekunta, Biologian laitos; Biocenter Oulu
Acta Univ. Oul. A 643, 2014
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä

Kasvit ja eläimet levittäytyivät Pohjois-Eurooppaan viimeisen jääkauden jälkeen mannerjäätikön ulkopuolella jääneistä refugioista. Useat kasvit, kuten idänpitkäpalko (*Arabidopsis lyrata* ssp. *petraea*) sopeutuivat pohjoisen lyhyeen kasvukauteen. On syytä olettaa, että suuntaava valinta vaikutti sopeutumisessa tärkeisiin ominaisuuksiin. Tässä väitöskirjassa tutkin voidaanko suuntaavan valinnan aiheuttamia jalanjälkiä löytää kahdesta tärkeästä kukkimisaikageenistä, *FYTOKROMI A* (*PHYA*) ja *FLOWERING LOCUS C1* (*FLC1*) geeneistä. Tätä varten vertasin DNA sekvenssimuuntelua pohjoisessa (Norja) ja eteläisessä (Saksa) populaatiossa, kiinnittäen erityisesti huomiota geneettisen muuntelun määrään ja erilaistumiseen. Lisäksi tutkin miten Spiterstulenin reunapopulaatio voi vastata tulevaisuudessa muuttuvaan ympäristöön. Sopeutumispotentiaalia arvioitiin sekä fenotyyppisen plastisuuden että additiivisen geneettisen muuntelun määrällä. Lisäksi tutkin miten vaihtelu 21 kukkimisaikageenissä liittyy fenologisiin ja kelpoisuusominaisuuksiin.

Useat merkit viittasivat siihen, että suuntaava valinta oli vaikuttanut kummassakin tutkitussa geenissä. Muuntelu oli vähentynyt voimakkaasti kumpaakin geeniä ympäröiviltä kromosomialueilta, jotka olivat myös selkeästi erilaistuneet. Additiivinen geneettinen muuntelu oli selvästi vähentynyt kukkimisominaisuuksissa verrattuna morfologisiin ominaisuuksiin, mahdollisesti suuntaavan valinnan johdosta. Kukkimisominaisuudet olivat kuitenkin plastisia. Jotkin valojaksoreitin geenit vaikuttivat sekä kukkimiseen että lisääntymiskykyyn.

Nämä tulokset osoittavat että suuntaava valinta vaikutti kahteen tutkittuun geeniin pohjoiseen levittäytymisen aikana. Geneettiset muutokset liittyivät todennäköisesti muuttuneisiin valojakso-, ja vernalisaatiovasteisiin, jotka saattoivat edistää sopeutumista lyhyeen kasvukauteen. Koska geneettistä muuntelua oli vain hyvin vähän, fenotyyppisellä plastisuudella on todennäköisesti tärkeä rooli sopeutumisessa muuttuvaan ympäristöön Spiterstulenin reunapopulaatiossa.

Asiasanat: adaptaatiopotentiaali, Arabidopsis lyrata, assosiaatiokartoitus, fenotyyppinen plastisuus, FLC, kukkimisaika, PHYA, valinnan pyyhkäisy

Acknowledgements

First I want to thank my principal supervisor, Outi Savolainen, who introduced me the fascinating world of evolutionary genetics. Without Outi's enthusiastic attitude and professional skills, the field would not have been so attractive. I want to thank also my other supervisors, Helmi Kuittinen and Tanja Pyhäjärvi. I am very grateful for your collaboration in working on the original papers and for commenting on the thesis manuscripts.

I want to thank my other co-authors in papers, Anne Niittyvuopio, Ulla Kemi, Timo Vesimäki, Saana Remula and David Remington. Your contributions were essential for this thesis. It has been a pleasure working with you also. I wish to thank my thesis reviewers prof. Johanna Vilkki and Doc. Irma Saloniemi for good advices and constructive feedback. I am also grateful to my follow-up group, professor Taina Pihlajaniemi, Dr. Reetta Vuolteenaho, Dr. Tuire Salonurmi and Dr. Heli Ruotsalainen who supported me and gave good advice concerning my studies and career. I acknowledge financial support by the Biocenter Oulu Doctoral program, Department of Biology of University of Oulu, the Biosciences and Environment Research council (to OS) and the Faculty of Natural Sciences.

Next I wish to thank Soile Alatalo, Hannele Parkkinen, Marja Nousiainen who helped me in genotyping and phenotyping and Matti Rauman who designed the growing conditions with me and Tuomas Kauppila and other staff from the botanical garden. Several people have helped me at the Spiterstulen and Oulu experimental field sites, including Heidi Aisala, Anu Pasanen, Jatta Jääskeläinen, Milla Koramo, Eevi Suleva. Thank you for your help. I want to thank also Lars and Marit Bakkom for their hospitality and kindness and the Sulheim family at Spiterstulen Turisthytte.

I thank Mikko Sillanpää for discussions on natural selection and Jaakko Lumme for interesting discussions on a broad spectrum of topics.

The Plant genetics group has been a good working place. I want to thank former and current Arabidopsis team members, Esa Aalto, Antti Virtanen, Johanna Kinnunen, Kirsi Järvi, Jaro Guzinski, Johanna Leppälä, Päivi Leinonen, Tiina Mattila and Tuomas Hämälä for your help and good company. The pine team, including Komlan Avia, Aleksia Vaattovaara, Yongfeng Zhou, Jaakko Tyrmi and Sonja Kujala have also been easily approachable and helpful, thanks for that.

Finally I want to thank my family and my friends for their support during these years. Mirva, your encouragement and support was crucial for finalizing the thesis.

Oulu, November 2014 Tuomas Toivainen

Abbreviations

CO	CONSTANS
CVA	coefficient of additive genetic variation
FLC1	FLOWERING LOCUS C1
FRI	FRIGIDA
FT	FLOWERING LOCUS T
GWAS	genome-wide association studies
h^2	heritability
LD	linkage disequilibrium
LGM	last glacial maximum
N_e	effective population size
PHYA	PHYTOCHROME A
PHYB	PHYTOCHROME B
QTL	quantitative trait locus
sd	standard deviation
SNM	standard neutral model
SNP	single nucleotide polymorphism
TOC1	TIMING OF CAB EXPRESSION 1
V_A	additive genetic variation

List of original papers

This thesis is based on the following publications, which are referred to throughout the text by their Roman numerals:

I Toivainen T, Pyhäjärvi T, Niittyvuopio A, Savolainen O (2014) A recent local sweep at the *PHYA* locus in the northern European Spiterstulen population of *Arabidopsis lyrata*. Molecular Ecology 23: 1040–1052.

II Kemi U, Niittyvuopio A, Toivainen T, Pasanen A, Quilot-Turion B, Holm K, Lagercrantz U, Savolainen O, Kuittinen H (2013) Role of vernalization and of duplicated *FLOWERING LOCUS C* in the perennial *Arabidopsis lyrata*. New Phytologist 197: 323–335.

III Toivainen T, Vesimäki T, Remula S, Remington D, Kuittinen H, Savolainen O (2014) A marginal *Arabidopsis lyrata* population has low genetic variation but is phenotypically plastic in flowering traits. Manuscript.

Author contributions

Paper	Study design	Data collection	Data analyses	Manuscript preparation
I	OS, **TT**	**TT**, AN	**TT**	**TT**, OS, TP
II	UL, HK, OS	UK and others	UK, **TT** and others	UK, HK, **TT** and others
III	OS, HK, **TT**	**TT**, TV, SR	**TT**, TV, SR, DR	**TT**, HK, OS

Ulla Kemi (UK), Helmi Kuittinen (HK), Ulf Lagercrantz (UL), Anne Niittyvuopio (AN), Tanja Pyhäjärvi (TP), David Remington (DR), Saana Remula (SR), Outi Savolainen (OS), Timo Vesimäki (TV), Tuomas Toivainen (**TT**)

Table of contents

Abstract
Tiivistelmä
Acknowledgements **7**
Abbreviations **9**
List of original papers **11**
Table of contents **13**
1 Introduction **15**
1.1 Genetic adaptation to local conditions requires evolution by natural selection ... 15
1.2 Evolution by random processes and natural selection ... 15
1.3 Genetic architecture of adaptation ... 17
1.4 Detecting natural selection ... 18
1.4.1 Coalescent theory ... 18
1.4.2 A hard sweep as a footprint of natural selection ... 19
1.4.3 Controlling random effects and demography when inferring selection at a single locus ... 20
1.5 Characterizing the response potential for environmental change ... 21
1.5.1 Phenotypic plasticity ... 21
1.5.2 Evolutionary responses ... 22
1.5.3 Association mapping ... 23
1.6 Flowering time genes as targets of selection ... 24
1.7 *Arabidosis lyrata* as an evolutionary genetic model species ... 26
1.8 Aims of the study ... 27
2 Material and methods **29**
2.1 Material for sequence analyses ... 29
2.1.1 PHYA ... 29
2.1.2 FLC ... 30
2.2 Characterizing potential to respond to environmental change in flowering traits ... 31
2.2.1 Study material ... 31
2.2.2 Response potential ... 32
2.2.3 Association mapping ... 32
3 Results and discussion **33**
3.1 Genetic signals of adaptation to the northern conditions in *A. lyrata* ... 33

3.1.1 Photoperiodic pathway - PHYA 33
3.1.2 Vernalization pathway - FLC 34
3.2 Response potential of marginal *A. lyrata* population 36
3.2.1 Plasticity 36
3.2.2 Potential for genetic responses – genetic variation in quantitative traits 37
3.2.3 Photoperiodic pathway genes have small effects on fitness 38
3.2.4 Can Spiterstulen population respond to changing environment in respect of flowering time? 39
4 Conclusions **41**
References **43**
Original articles **55**

1 Introduction

Species inhabiting a heterogeneous environment deal with the variable environment either by phenotypic plasticity or by local adaptation and the associated genetic differentiation. Phenotypic plasticity and genetic differentiation are not mutually exclusive. Often phenotypic plasticity is the first buffer against environmental change and precedes adaptation (Bradshaw 1965). Colonization of new areas is often accompanied by genetic changes that confer adaptation in the new environment. In the long term, genetic changes resulting in adaptation can be the first step towards the evolution of new species (Darwin 1859).

1.1 Genetic adaptation to local conditions requires evolution by natural selection

Locally adapted populations have higher fitness in their home site than any other population introduced to the site (Kawecki & Ebert 2004). During local adaptation, divergent selection pressures in different environments results in genetic changes in adaptive traits conferring fitness advantage in the native site of each population. For instance in trees, genetic differentiation in annual growth periods allows survival and reproduction in different latitudes (Savolainen *et al*. 2007).

Divergent selection does not always result in local adaptation, as random genetic drift or gene flow can prevent differentiation. Reciprocal transplant experiments have shown that some 50-70% of populations are locally adapted (Leimu & Fisher 2008, Hereford 2010). Local adaptation has evolved, for example, in *Arabidopsis thaliana* (Ågren & Schemske 2012), *Mimulus guttatus* (Hall & Willis 2006) and *Arabidopsis lyrata* (Leinonen *et al*. 2009, 2011) and many tree species (Savolainen *et al*. 2007, Alberto *et al*. 2013). When local adaptation has been demonstrated, its genetic basis can be studied. If reciprocal transplant experiments are not possible, then local adaptation can be inferred from other experimental work, or from patterns of genetic variation (Kawecki & Ebert 2004).

1.2 Evolution by random processes and natural selection

Evolution is a consequence of interplay between mutation (and recombination), random genetic drift, migration and natural selection. The effective population size (N_e), the size of an ideal population experiencing the same level of drift as the actual population, is a key factor affecting both the selective and neutral processes. Before

natural selection can be studied, random processes have to be well understood. At the level of an individual neutral locus genetic drift results in random fluctuations of allele frequencies in each generation (binomial variance of allele frequency change/generation $\sigma_p^2 = p(1-p)/2N$, where p is a frequency of allele 1 at a biallelic locus and N is population size) due to random sampling of gametes (Wright 1931). In a small population or when a population goes through a bottleneck, genetic drift can result in large changes in allele frequencies. In large populations the effect of genetic drift is much smaller.

Mutations are the source of new variation. The neutral theory assumes that deleterious mutations are eliminated usually very rapidly and beneficial mutations displace existing alleles in an evolutionarily short time. It then follows that most variants in polymorphic sites within populations (or species) are neutral ($Ns < 1$, (s is the selection coefficient) or nearly neutral mutations on their way to fixation (Kimura 1968, Ohta 1973). The expected level of polymorphism (population mutation rate θ) at equilibrium for the infinite site model (mutations always occur at a new site) is the result of the balance between mutation and drift. It is the product of the effective population size (N_e) and mutation rate (μ), $\theta=4N_e\mu$. In large populations nucleotide polymorphism is expected to be higher because drift reduces variation more slowly than in small populations.

The rate of neutral molecular evolution is not dependent on population size because there is an inverse relationship between the supply of new mutations ($\mu*2N$) and their fixation probability (1/2N), (Kimura 1968). Thus the mutation rate alone determines the rate of neutral molecular evolution.

The neutral theory emphasizes the major role of drift in the molecular evolution (in the short term), but it does not deny the importance of natural selection in adaptation. Natural selection changes allele frequencies to result in higher fitness in the present environment. In a constant size population, an advantageous mutation has fixation probability of 2s (Haldane 1927, Kimura 1962). The supply of new beneficial mutations is $\mu*2N$. Thus the rate of adaptive evolution via new mutations is $4N_e s\mu$. Further, because the effective recombination rate ($4N_e r$) is higher in species with large N_e, selection can influence the genome at a finer resolution and more efficiently (Hill & Robertson 1966, Barton 1995, Neher 2009, Presgraves 2005, Haddrill *et al*. 2007). These effects on adaptive evolution have been demonstrated by experimental data. When related species pairs have been compared, e.g. in fruit flies (Jensen & Bachtrog 2011), mice (Phifer-Rixey *et al*. 2012) and sunflowers (Strasburg *et al*. 2011), species with larger effective population sizes have shown more rapid adaptive evolution than species with smaller populations.

Furthermore, population structure can have an influence on the scale of adaptive evolution. If populations have a fragmented distribution with restricted gene flow, as in many plants, adaptive evolution occurs at a local scale, as has been shown in *A. thaliana* (Horton *et al*. 2012, Fournier-Level *et al*. 2011, Hancock *et al*. 2011, Long *et al*. 2013, Huber *et al*. 2014). Similar findings have also been made in humans (Barreiro *et al*. 2008, Keinan & Reich 2010). Species-wide evolution can then be rare (Cao *et al*. 2011, Hernandez *et al*. 2011). Consistently, rapid adaptive evolution can take place in tree species high migration rate (Ingvarsson *et al*. 2010, Zhou *et al*. 2014).

1.3 Genetic architecture of adaptation

One important question in evolutionary genetics is what kind of genetic changes underlie adaptation: Citing Charles Darwin: "natural selection can act only by taking advantage of slight successive variations, she can never take a leap, but must advance by the shortest and slowest steps" (Darwin 1859). R.A. Fisher also thought that adaptation proceeds via several small effect mutations because large effect mutations are almost always deleterious (Fisher 1930). This gradualistic view was challenged 50 years later by Motoo Kimura in 1980's (Kimura 1983). He included the fixation probability of mutation (2s) in the model and concluded that mutations of intermediate sizes are the most likely genetic source of adaptation. Currently, the prevailing view is based on H.A. Orr's theory (Orr 1998). According to the theory, directional selection in a single population should result in fixation of adaptive mutations, the effect sizes of which should follow the exponential distribution. Further, the first mutation and the largest effect mutation account for a majority of the total fitness increase (Orr 2002). In a heterogeneous environment, Yeaman & Whitlock (2011) suggest that selection for local adaptation with migration-selection balance will result in large effect mutations underlying the differentiation. Empirical findings have demonstrated cases with both large and small effects. Major genes have been shown to govern adaptation in several species, mouse pigmentation (Hoekstra *et al*. 2006), vernalization response in *Arabidopsis thaliana* (Le Corre *et al*. 2002, Johanson *et al*. 2000) and armor plates in Stickleback (Cresko *et al*. 2004). Even taking into account the bias that the large phenotypic effects may have attracted researchers' attention and that major genes underlying adaptation can be more easily detected, this kind of mutations clearly can be important.

1.4 Detecting natural selection

Adaptation can be based on new beneficial mutations that increase in frequency. These new large effect mutations often result in a characteristic sequence pattern including loss of variation (Maynard-Smith & Haigh 1974). These are called "hard" selective sweeps. Adaptation could also result from existing (standing) variation at individual loci. Such a "soft" sweeps results in a less clear signal of past selection (Pennings & Hermisson 2006.). Methods based on coalescent theory (Kingman 1982, Tajima 1983, Hudson 1991, reviewed by Wakeley 2008) have an essential role when recent selection at a single locus is studied. Polygenic adaptation from standing genetic variation by several additive small effect mutations may also be quite frequent, even if it is more difficult to pinpoint the underlying loci. Examples are e.g. human height (Turchin *et al*. 2012) and rabbit domestication (Carneiro *et al*. 2014). Methods for detecting polygenic adaptation are just beginning to be developed (Berg & Coop 2014).

1.4.1 Coalescent theory

Coalescent theory (Kingman 1982) has a root in the standard neutral model (SNM, Wright-Fisher model (Wakeley 2008) of evolution. In this model, a population of constant size reproduces by random mating, with discrete generations. All individuals have an equal probability of survival and reproduction. The coalescence tree depicts the historical genealogical relationships of n sequences (or individuals) backward in time until the most common recent ancestor of all the individuals or sequences in the tree has been found after n-1 coalescence events.

The characteristics of a standard coalescence tree are governed by the size of the population (N). Coalescent times are independent and exponentially distributed. Most coalescences occur rapidly in recent history. Coalescence time for the two last lineages is roughly a half (2N generations) of the total height of the coalescent tree (4N(1-1/n)), where n is the number of individuals sampled. The expected number of mutations in each branch of the tree is approximately Poisson distributed, governed by the parameter θ ($4N_e\mu$) (Wakeley 2008).

The SNM results in a standard coalescence tree, where the expected distribution of mutations occurring in internal and external branches is known, even though the random processes can result in highly diverse individual genealogies of trees. Watterson (1975) and Tajima (1983) derived estimators of θ based on the expected number of segregating sites (θ_W) and pairwise differences (θ_π), respectively. Given the

standard neutral model, the two estimates, θ_W and θ_π are expected to be equal. This result has been used extensively to detect deviations from neutral evolution (Tajima 1989, Fu & Li 1993, and Fay & Wu 2000).

Coalescent theory serves as a computationally efficient tool to model evolution. It is well suited for examination of current data. Coalescent theory starts with a sample of sequences from the current populations, corresponding to the observations. Coalescent theory can be applied to many different questions within population genetics, inferring demography or speciation events. Further, coalescent simulations with selection are useful when studying selective sweeps (Kim & Stephan 2002, Nielsen *et al*. 2005, Pavlidis *et al*. 2013).

1.4.2 A hard sweep as a footprint of natural selection

If a new beneficial mutation is not lost by drift in the early stages when it is very rare, natural selection starts to increase its frequency rapidly, and because of linkage, the chromosomal haplotype carrying it increases in frequency. The linked region hitchhikes with the selected mutation and is finally fixed within a population. This process is called genetic hitchhiking (Maynard-Smith & Haigh 1974), subsequently termed also a selective sweep. This region with the new haplotype is highly differentiated compared to the ancestral haplotype (Sabeti *et al*. 2002, Voight *et al*. 2006). Directional selection eliminates variation around the selected locus. Recombination is critical in the early stage of sweep, because it can shuffle the beneficial mutation to high fitness backgrounds and remove negative associations in the same chromosome (Barton 1995, Neher *et al*. 2009). Independent recombination events on both sides of the selected site limit the length of the swept region. The extent of the swept region is determined by the ratio of the selection coefficient s and the recombination rate r, s/r, (Kim & Stephan 2002). The scaled selection coefficient ($4N_es$) of a sweep can be calculated with this information (Stephan *et al*. 1992). Note that the effective population size does not have an influence on the length of swept region because even if selection is stronger in large populations ($4N_es$), there are more recombination events ($4N_er$) in same time interval. A chromosomal fragment carrying a new beneficial mutation will become fixed rapidly and carry little variation, but after the fixation, it starts to accumulate new mutations, which first occur at low frequency (Tajima 1989). The flanking regions surrounding the swept area harbour an excess of derived high frequency alleles (Fay & Wu 2000). These areas are not completely fixed for the haplotype carrying the selected allele, because they escaped the sweep by recombination. These skews in allele frequency spectra are not expected

in the standard neutral model (SNM, Wright-Fisher), and are a characteristic signal of a selective sweep (Braverman *et al*. 1995). After a sweep, linkage disequilibrium (LD), the non-random association of alleles at two loci, between flanking regions is absent or low due to the independent recombination events, which occurred in different times during a sweep (Kim & Nielsen 2004). In contrast, LD is expected to be high within both of the flanking regions. The most informative signs of a sweep (e.g. LD patterns) do not remain detectable for much longer than 0.1 N_e generations after a sweep (Kim & Stephan 2002, Pfaffelhuber *et al*. 2008).

1.4.3 Controlling random effects and demography when inferring selection at a single locus

Detecting selection is difficult because it occurs concurrently with random processes and demographic events, such as bottlenecks, population expansions, admixture, or population isolation. Such demographic events can result in nucleotide variation patterns resembling footprints of positive selection (Jensen *et al*. 2005, Pavlidis *et al*. 2010, Thornton & Jensen 2007). Thus the effects of demography should be controlled statistically. Coalescence simulations play an important role in this. The likelihoods for the observed data can be calculated using simulations assuming different neutral demographic models, with parameters estimated from the data (Hudson 2002, Csilléry *et al*. 2010). Further, spatial genomic data (along the chromosome) can be utilized to detect selective sweeps by calculating likelihoods for the observed data given the observed parameters and the neutral (Kimura 1971) or hitchhiking model (Fay & Wu 2000) (Kim & Stephan 2002, Jensen *et al*. 2005, Nielsen *et al*. 2005). The current extensive genome-wide sequence data allow more informative comparisons between the genome-wide level variation, influenced mainly by neutral processes, and the level of variation at individual candidate loci that may be influenced by selection (Wright & Charlesworth 2004).

In addition to demographic events, the recurrent removal of deleterious alleles (background selection), can mimic the traces of a selective sweep (Charlesworth *et al*. 1993, Cai *et al*. 2009). Background selection also removes variation within populations, and gives rise to patterns of nucleotide variation that might be due to positive selection. However, strong background selection is not expected to skew allele frequency spectra as strongly as a selective sweep (Stephan 2010).

Finally, support for the role of selection can be obtained by approaches using statistical tests that are based on different genetic aspects of the data, because individual signs of selection do not always allow making robust conclusions about

directional selection. The amount of nucleotide diversity and divergence (Wright & Charlesworth 2004), genetic differentiation (Foll & Gaggiotti 2008), LD (Voight *et al*. 2006) and allele frequency spectra (Tajima 1989, Fay & Wu 2000) together comprise a powerful tool to detect selection.

1.5 Characterizing the response potential for environmental change

Populations can respond to altered conditions by genetic changes (adaptation) or by tolerating new environmental challenges by phenotypic plasticity. Plants cannot avoid the conditions by migrating. The experimental evidence in plants (Franks *et al*. 2014) and in e.g. corals (Palumbi *et al*. 2014) suggests that both phenotypic and evolutionary responses have been important in responding to rapid environmental changes, but initial responses to a rapid environmental change are likely to be phenotypic (Anderson *et al*. 2012). In the long term, evolutionary responses have been important (e.g. Davis & Shaw 2001). The probability of genetic responses varies depending on the characteristics of the population or species. There are only few documented cases of genetic change in response to climate warming (Gienapp *et al*. 2008). These include rapid responses to drought in flowering time in Brassica (Franks *et al*. 2007), and a change in the critical day length for diapause in the pitcher plant mosquito *Wyeomyia smithii* (Bradshaw & Holzapfel 2001).

1.5.1 Phenotypic plasticity

Phenotypic plasticity means that the same genotype expresses different phenotypes in different environments (Bradshaw 1965). For example, plants growing taller in shaded environment, or flowering earlier in warmer conditions are plastic responses. Phenotypic plasticity is a widespread phenomenon across organisms, even though it is thought to be more common in sessile organisms, such as plants (Bradshaw 1965, Nicotra *et al*. 2010). Phenotypic plasticity can be adaptive, maladaptive or neutral with regard to an individual's fitness. Species inhabiting more heterogeneous environments usually show more plasticity (Sultan 2001, Matesanz *et al*. 2012). In particular, adaptive phenotypic plasticity can be crucial for tolerating new environmental conditions if adaptive genetic variation is low (Bradshaw 1965, Anderson *et al*. 2012). However, phenotypic plasticity can first reduce the efficiency of natural selection and slow down evolutionary responses (Chevin *et al*. 2010).

1.5.2 Evolutionary responses

Adaptation by genetic changes requires available genetic variation. The question of the maintenance of polygenic variation is still poorly understood. Directional selection will deplete additive genetic variation, V_A. Mutations are the ultimate source of V_A. They increase V_A, even relatively rapidly if a trait has a polygenic architecture (Lynch & Walsh 1998). Spatially or temporally varying selection can maintain variation within populations under some conditions (Levene 1953, Via & Lande 1987). Some authors have suggested that antagonistic genetic correlation between fitness components can maintain additive genetic variation in fitness traits (Rose 1985, Charlesworth & Hughes 1996), but the conditions for it can be quire restricted (Hedrick 1999). Further, genotype x environment interactions can maintain variation within a population (Gillespie & Turelli 1989).

Low heritabilities (h^2, the proportion of genetic determination of phenotype) have been found in the traits connected to fitness (Crnokrak & Roff, 1995, Falconer & MacKay 1996), which suggests that natural selection results in low additive genetic variation in those traits (Fisher 1930, Robertson 1955). However, when Houle (1992) scaled additive genetic variation with mean of a trait (CVA), he showed that the amount of additive genetic variation is not lower in fitness traits, but that they harbour a large amount of environmental variance, accounting for the lower heritabilities. This variation can be especially valuable during sudden environmental change. Finally, heritability is a population and environment specific measure for a trait. New conditions can result in different patterns of phenotypic variation with an increased genetic component of variation (e.g. Goodnight 1988). This increase in additive genetic variation may also concern fitness itself (Shaw & Shaw 2014).

The response to selection requires that the trait correlates genetically with fitness (Robertson 1966, Price 1970). The amount of additive genetic variation in a trait and the importance of a trait for fitness (selection differential) determine the expected response. However, because an organism is the product of thousands of traits, among which several are genetically correlated, the selection response is affected by the correlation structure and selection on other traits (Lande & Arnold 1983). Empirical studies have emphasized the importance of antagonistic genetic correlations between traits in reducing selection responses to a warming climate (Etterson & Shaw 2001).

The genetic architecture of a trait and the genetic interactions among loci also have an influence on the selection response. Larger effect loci are expected to be fixed rapidly but additive genetic variance is then rapidly reduced. However, quantitative (polygenic) traits are usually affected by several small effect loci which act additively

(Visscher 2008), and selection does not deplete additive genetic variation as rapidly (Falconer and Mackay 1996). Further, because variation is reduced slowly, new mutations can produce substantial genetic variation in parallel. The fitness of genotypes at one locus can be influenced by the genetic background at other loci. Such epistatic (non-additive) interactions may be common, even if difficult to detect (MacKay 2014). Huang *et al*. (2012) suggested that epistasis is an important part of the genetic architecture of quantitative traits in Drosophila.

1.5.3 Association mapping

Finding quantitative trait loci (QTL) is of central importance in quantitative genetics. Association mapping is a powerful method for mapping loci affecting phenotypic variation at high resolution (Balding 2006). In population samples of especially random mating organisms, linkage disequilibrium is lower than in progeny of QTL crosses because of the historical recombination. Mapping resolution depends essentially on the extent of linkage disequilibrium (LD). In random mating large populations, LD decays more rapidly due to numerous historical recombination events. For example, in outbreeding species such as in maize, rye or *A*. *lyrata*, LD decays more rapidly (Remington *et al*. 2001, Li *et al*. 2011, Wright *et al*. 2006) due to higher $4N_e r$ compared to the inbreeding species, such as *A*. *thaliana*, rice or wheat (Nordborg *et al*. 2002, Garris *et al*. 2003, Somers *et al*. 2007). Association mapping can be conducted by selecting candidate genes for targets or by conducting the analysis genome-wide. The former will miss genes not included in the study. Combing QTL mapping and association mapping is the most powerful tool to find associations (Yu *et al*. 2008). Association mapping can be conducted within populations, as has often been done in human studies (The Welcome consortium 2007), or by combining populations. Because the samples are often genetically structured (e.g. several populations), an important task is to control the confounding effects of heterogeneous genetic backgrounds (Yu *et al*. 2006). This can be also a caveat, because disregarding SNPs associated with population structure can weaken a power to find adaptively important loci correlated with population structure. When an association study is conducted within a population with no significant population structure, the number of spurious associations is strongly diminished and the effect sizes of alleles can be estimated with higher accuracy. Then only those QTLs segregating within the population are detected.

Association mapping has been used successfully to characterize important genes e.g. for several diseases (starting with Wellcome trust 2007) and flowering time in *A*.

thaliana (Atwell *et al*. 2010), maize (Buckler *et al*. 2009) and cold tolerance in forest trees (e.g. Eckert *et al*. 2009).

1.6 Flowering time genes as targets of selection

All plants in natural populations need to adapt to surrounding environmental conditions with respect to flowering time. Flowering time is regulated by environmental cues, most importantly by temperature and light (Thomas &Vince-Prue 1997). Day length and temperature conditions differ between latitudes, giving rise to selection for phenotypic differences for day length and temperature requirements between populations.

The genetic signaling pathways, such as photoperiodic, temperature (or vernalization), and autonomous pathways involved in flowering time are well characterized in several species (Fig. 1) (review by Andres & Coupland 2012). Light is captured by photoreceptors, which respond to different wave-lengths. Phytochromes (e.g. *PHYA*, *PHYB* in angiosperms) are specialized for red and far-red light (reviewed in Sharrock 2008) and cryptochromes (*CRY1*, *CRY2*) for blue and ultra-violet wavelengths (Yu *et al*. 2010). They regulate multiple responses throughout the plant life cycle. Several phytochrome loci such as *PHYB2* in *Populus tremula* (Ingvarsson *et al*. 2006, 2008), *PHYC* in *A*. *thaliana* (Balasubramanian *et al*. 2006) or *PHYE* in *Cardamine nipponica* (Ikeda *et al*. 2009) have been suggested to be differentiated across latitudes due to local adaptation. The clock genes, (e.g. *TOC1*, *LHY*, *ELF3*, *CCA1*, *FKF1*, *GI*, *ZTL*) are regulated mainly by photoreceptors (Somers *et al*. 1998, Devlin & Kay 2000). Clock genes and several downstream targets show adaptive genetic differentiation across latitudes in *Populus balsamifera* (Keller *et al*. 2012) and Norway spruce (Källman *et al*. 2014). The clock genes might have been targeted frequently by recurrent selection also over the long term, as was demonstrated in *Populus tremula* (Hall *et al*. 2011). The last of the downstream genes in the photoperiodic pathway, *CONSTANS*, has a key role in photoperiodically regulated flowering (reviewed by Valverde 2011). *CONSTANS* regulates the *FT* gene which finally triggers flowering and growth cessation, as in *A*. *thaliana* and Populus Suarez-Lopez *et al*. 2001, Böhlenius *et al*. 2006, Hsu *et al*. 2011. Genes with a CCT domain (*CONSTANS*, *CO*-like, and *TOC1*) have been shown to govern adaptive photoperiodic flowering time variation in rice (Xue *et al*. 2008), maize (Hung *et al*. 2012), wheat (Beales *et al*. 2007), barley (Turner *et al*. 2005) and likely in Capsella (Slotte *et al*. 2007). The *CONSTANS* gene family is rapidly evolving (Lagergranz 2000), perhaps due to the important role role in adaptive evolution.

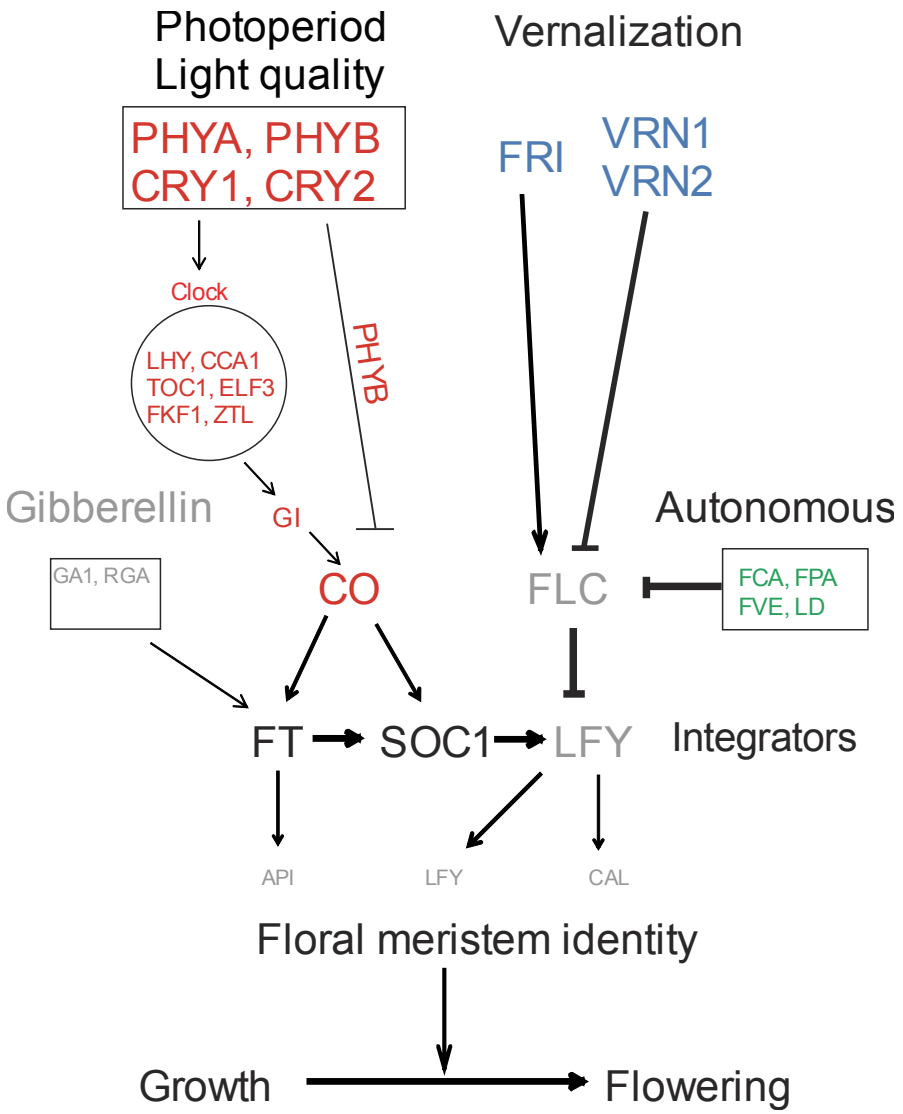

Fig. 1. The main genetic signaling pathways resulting in flowering in *A. thaliana*. The genetic signaling pathways are marked with different colours. Red colour depicts photoperiod/clock pathway, whereas blue and green colour depict vernalization and autonomous pathways, respectively. The full names for genes can be found from the Supplementary Table1S in III. Genes in grey colour were not studied in III. Arrows promote flowering and lines terminated with a bar denote repressive effects. Adapted from Corbesier & Coupland (2006), Mouradov et al. (2002), Blázquez (2000).

Vernalization (Latin: *vernus*, of the spring), a prolonged cold exposure promotes flowering in spring after winter in several plant species. Even if the flowering time regulatory gene network involves dozens of genes, only relatively few, such as *FRIGIDA* (*FRI*) (Johanson *et al*. 2000, Salomé *et al*. 2011, Stinchcombe *et al*. 2004) and *FLOWERING LOCUS C* (*FLC*) have been shown to underlie natural variation in the annual *A*. *thaliana* and in cultivated varieties in annual oilseed rape (*Brassica napus* L.) (Wang *et al*. 2011, Tadege *et al*. 2001). Thus the vernalization pathway seems to be important for flowering time variation in many brassicaceous species. Phytochromes have had a minor role in governing flowering time in *A. thaliana* (Stinchcombe *et al*. 2004, Mendez-Vigo *et al*. 2011).

1.7 *Arabidosis lyrata* as an evolutionary genetic model species

A. *lyrata* is a close relative of *A*. *thaliana*. The species diverged 10 million years ago (Beilstein *et al*. 2010, Ossowski *et al*. 2010) and 15% of the synonymous sites are diverged between species (Yang & Gaut 2011). Despite the close relatedness, there are some fundamental biological differences between species. In contrast to *A*. *thaliana*, *A*. *lyrata* is self-incompatible and perennial.

A. lyrata ssp. *petraea* has a fragmented distribution across central and northern Europe. It prefers low competition habitats, is pollinated by insects. It can also propagate clonally. The northern European *A*. *l*. *petraea* populations have colonized their current areas after the last glacial maximum (LGM) but the exact routes are unknown (Schmickl *et al*. 2010). Overall nucleotide variation in northern European populations is reduced to less than half compared to the central European *A*. *lyrata* populations, possibly due to bottleneck associated with colonization (Wright *et al*. 2003, Muller *et al*. 2008, Pyhäjärvi *et al*. 2012). The high altitude Norwegian (Spiterstulen) population (1100 m.a.s.l.), has been shown to be locally adapted in a comparison with a set of European populations (Leinonen *et al*. 2009).

Photoperiodic responses differ between Central European (Plech) and northern populations (Riihimäki & Savolainen 2004, Leinonen *et al*. 2013, II). The Spiterstulen population requires longer days to start flowering, whereas plants from Plech flower extensively already in 14 h light conditions (Quilot-Turion *et al*. 2013). The northern Spiterstulen population responds more to vernalization in long days (20h) (Kuittinen *et al*. 2008, Riihimäki *et al*. 2005) but not in short days (Quilot-Turion *et al*. 2013), which suggests that vernalization and subsequent long days are strong signals of spring in the northern Spitertulen populations (Leinonen *et al*. 2011, II).

1.8 Aims of the study

Plants and animals colonized the Northern Europe after the last Ice Age. When organisms migrated from Central Europe to the North, adaptation to the short summer and long winter was required. Many plants, such as the northern rock cress (*Arabidopsis lyrata* ssp. *petraea*) adapted to the short growing season in the North. Molecular and developmental biologists have identified several genes which influence the timing of flowering and growth (e.g. Mouradov *et al*. 2002). However, it is not known which of those genes have been important when plants adapted to the northern conditions. The aim of the first part of the thesis (I and II) is to examine directional selection (selective sweeps) at individual flowering time genes. Specifically, we examine two loci known to be potentially functionally important, *PHYA* and *FLC*: Do they show clear signals of directional selection?

The second part of thesis (III) studies the current response potential for changing environmental conditions within a northern *A*. *lyrata* population (Spiterstulen), located at species range margin. Isolated populations located at the species range margin may be vulnerable to extinction (Krajick 2004). To survive a population can respond to environmental change by phenotypic plasticity, adapting by genetic changes, or a combination of the response mechanisms (Franks *et al*. 2007). The information considering the relative importance of the responding mechanisms is still scarce. We focused on flowering time, which is a major adaptive trait in plants. We wanted to know how much plasticity, and additive genetic variation exists for flowering traits. We also evaluated the importance of the trait for fitness within the natural Spiterstulen environment and studied which of the studied flowering time genes govern fitness variation, and are potential targets of selection within the current Spiterstulen population

2 Material and methods

Materials and methods are described shortly. For more detailed information see original articles (I and II) and manuscript (III).

2.1 Material for sequence analyses

Populations that were studied for DNA sequence variation in the *PHYA* and *FLC* genes were Spiterstulen, Norway (61° 38′ N, 8° 24′ E), and Plech, Germany (49° 39′ N, 11° 29′ E). Plants in the Plech population (approx. 400 m.a.s.l.) grow on rock boulders in the forest. The growing season extends from March to October (6 months, Clauss & Koch 2006). In Spiterstulen, plants grow in a mountain valley (1100 m.a.s.l.) on the mossy and rocky bank of the River Visa. The growing season is short: it lasts from the end of May to the beginning of September. Twenty unrelated plants were used for sequence analysis from each population. They were also crossed in ten within-population pairs to obtain progeny for haplotype inference in the case of *PHYA*. DNA was extracted from fresh and frozen leaves from all plants using FastPrep Kit (Qbiogene). The gene regions were sequenced with the Sanger method.

2.1.1 PHYA

To detect a possible selective sweep in the *PHYA* locus we sequenced 9 short gene fragments around the *PHYA* locus from 20 individuals of both Plech and Spiterstulen populations, including parts of the 5'UTR and 3'UTR regions. Amplified loci (300-900 bps) were located in a region of total length of 57 kb (Fig. 1A in I). Parental haplotypes across the 57-kb region were inferred based on progeny genotypes in each locus.

Population genetic summary statistics were calculated with DnaSP 5.10 software (Rozas 2009). We studied selection by examining the level of silent variation (Tajima 1983, Watterson 1975) Selection was tested for by comparing silent nucleotide diversity to neutral divergence at the *PHYA* locus with the MLHKA software (Wright & Charlesworth 2004) using 19 reference loci (Pyhäjärvi *et al*. 2012). LD patterns were characterized for each fragment separately (ZnS, Kelly (1997) and r^2 (Hill & Robertson 1968) and across all fragments variable in both populations (r^2 and p) across the studied 57 kb region. We calculated allele frequency spectra (Tajima 1989, Fay & Wu 2000) for each fragment and tested the fit to the expected based on the standard neutral model by 5000 coalescence simulations (Hudson 1991) without

recombination in both populations. The level of genetic differentiation was characterized as F_{ST} (Hudson 1992). The fit of the data to a model with a selective sweep was tested by coalescence simulations with the ssw and clsw - softwares (Kim & Stephan 2002, Jensen *et al*. 2005). At the same time the location of the selected site and the strength of selection (2Nes) were estimated. Goodness of fit statistics (GOF) was used to exclude some bottleneck and other demographic scenarios.

The ratio of nonsynonymous to synonymous divergences (K_a/K_s) (Nei & Gojobori 1986) gives an estimate of the selective constraint of the locus. K_a/K_s =1 is expected for neutral evolution of a gene, while a ratio higher than 1 is a signal of positive selection. The K_a/K_s ratio (Nei & Gojobori 1986) was used to characterize long term selection at the *PHYA* locus.

2.1.2 FLC

A. lyrata has two tandemly duplicated genes, *FLC1* and *FLC2*. Both of them were studied using two sequence sets (table 3 in II). Two random individuals from each population were used to study sequence variation in the promoters and in the whole *FLC1* (9 kb) and *FLC2* (6.9 kb) genes. The regulatory regions of the *FLC1* gene (3529 bp) were studied in more depth in 7 Spiterstulen and 13 Plech individuals.

Nucleotide diversity in both populations was estimated based on the number of segregating sites and on the average pairwise differences at silent sites, θ_π (Tajima, 1983). Genetic differentiation between populations was estimated as F_{ST} (Hudson *et al*. 1992) and the number of fixed differences.

K_a/K_s - ratio is also useful when studying evolution of gene duplicates. After duplication a duplicate might become nonfunctional and will start to evolve neutrally. High values can suggest non-functional pseudogenes. The gene copies (*FLC1* and *FLC2*) were compared with each other and with *A. thaliana FLC* with K_a/K_s - ratio (Nei & Gojobori 1986).

2.2 Characterizing potential to respond to environmental change in flowering traits

2.2.1 Study material

We carried out an experiment to estimate the Spiterstulen population´s current response potential to environmental change. In 2008, (April-May) in Oulu, ca.108 parental plants (collected as seeds from the natural population in 2002) were crossed according to the North Carolina II design (see Fig. 1 in III) (Lynch & Walsh 1998, 598-602). The design consisted of 27 crossing blocks each with 4 plants. The crosses within blocks resulted in reciprocal full- and half-sib families to allow estimation of additive, dominance and maternal components of variance. In total c.a. 1100 plants from the same families were planted to two environments, Oulu and Spiterstulen (Fig. 2). The Oulu plants were first grown in the growth chamber (onwards from sowing in June 2008) and then planted to the experimental field site in May 2009. Plants for Spiterstulen were first grown in the greenhouse until they were planted to the experimental site in July 2008. Phenotypes were recorded in both environments in both years 2009-2010 and in the growth chamber in 2008 (see Table 1 in III)

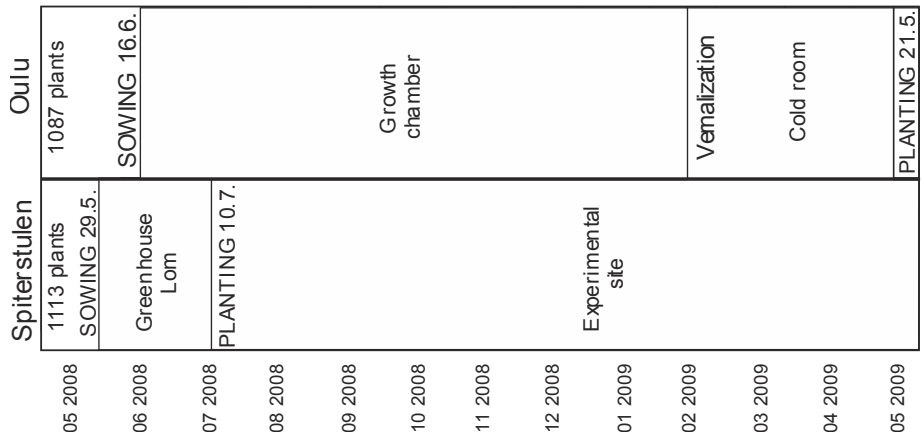

Fig. 2. Description of growing conditions of seedlings eventually transplanted in Spiterstulen and Oulu.

2.2.2 Response potential

The phenotypic response potential was characterized by recording phenotypes in different environments and in different years. The influence of the site on the trait variation reflects phenotypic plasticity. The differences between years are due to both differences in the environment between years and differences due to age. Heritability (h^2) and additive genetic variation (V_A) were calculated for each trait. Paternal families were used in calculations to exclude any maternal effects. Phenotypic and genetic correlations in both environments, and in both years were calculated for each pair of traits to characterize possible selective constraints due to negative genetic correlations. In addition, evolvability, (CVA, additive genetic standard deviation divided by the mean) was calculated to scale additive genetic variance to same scale in all traits.

2.2.3 Association mapping

To study if flowering time candidate genes contribute to variance in flowering time or fitness components, association mapping was conducted in 1077 plants grown in Oulu. 70 SNPs from 21 flowering time genes and 16 reference loci (Supplementary TableS1 in III) were used as markers. Relatedness was taken into account by calculating the kinship matrix with the SPAGeDI-software (Hardy & Vekemans 2002). The kinship matrix was used to avoid spurious associations that can arise due to genetic relatedness between individuals. The TASSEL software 2.1 (Bradbury *et al*. 2007) was used for mixed linear model analyses (Yu *et al*. 2006).

3 Results and discussion

3.1 Genetic signals of adaptation to the northern conditions in *A. lyrata*

3.1.1 Photoperiodic pathway - PHYA

We found strong evidence that directional selection targeted the *phytochrome A* (*PHYA*) locus after the LGM. Variation was reduced strongly at the *PHYA* locus in contrast to the expectation based on the genome-wide level variation and divergence (Table 2 in I). Reduced variation extended in total across a 9.4 kb region which carried a derived haplotype compared to the ancestral Plech population. *PHYA* at the Spiterstulen population was also differentiated at multiple nonsynonymous sites compared to the southern (Plech) population (Fig. 1C in I) and there was no LD across *PHYA* (Fig. 2A and 2D in I). In addition, coalescent based analysis of Kim & Stephan (2002) indicated a selective sweep.

Populations were highly differentiated at *PHYA* (F_{ST} = 0.6-0.8, Fig. 1B in I) compared to the genome-wide average (F_{ST} = 0.35, Pyhäjärvi et al. 2012). Three nonsynoymous fixed differences between populations were observed, which was not expected because low K_a/K_s (0.05) indicated that most non-synonymous mutations at the *PHYA* locus are deleterious and removed rapidly. High differentiation extended at least across the 9.4 kb chromosomal fragment. Variation was almost completely removed from the same region from the Spiterstulen population (Fig. 1A in I), which suggests that the whole haplotype has increased rapidly in frequency due to selection. The selected site was estimated to be in the 3'UTR region of *PHYA*, although the wide area of reduced variation prevented an accurate estimation.

To study the sweep hypothesis more carefully, we inferred haplotypes based on progeny genotypes across the 57 kb studied region. LD pattern in Spiterstulen population fitted well to the expectations of a hard sweep hypothesis (Kim & Nielsen 2004). We also observed skews in allele frequency spectra. A significant excess of low frequency alleles (Tajima 1989) was found in the 3'UTR region of *PHYA* ($D = -1.99$, $P < 0.05$) and (Fig. 3A in I). An excess of derived high frequency alleles was found from the flanking loci (fragment no 9: Fay & Wu $H_n = 3.4$, $P < 0.02$), (fragment no 3: $H_n = -1.74$, $0.05 < P < 0.1$) in Spiterstulen (Fig. 3B in I). These loci were the nearest to the low variation region. This suggested that these flanking loci had escaped a sweep by recombination.

Plants and animals colonized Scandinavia after the last glacial maximum 8 000-10 000 years ago (Björck 1995; Hewitt 1999). Pioneer plants, such as *A. lyrata*, were among the first plants that inhabited the exposed land areas after the ice sheets retreated. We estimated that the new beneficial mutation arose at the (*PHYA*) locus less than 8 200 years ago, given the length of selective phase c.a. 1 800 years (result not included in I), which agrees well with the estimated time of colonization.

The selection coefficient (s = 0.01) estimated for *PHYA* suggested that the new mutation had a large effect. In polygenic adaptation the individual effect sizes are usually smaller (Turchin *et al*. 2012) than observed here. As a comparison, in a genome-wide study of Drosophila, 3% of new nonsynonymous advantageous mutations with largest effect had mean s = 0.005, (Sattath *et al*. 2011).

PHYA, only found in angiosperms, is the most important phytochrome responding to far-red light and it measures the day length (Yanovsky & Kay 2002). Flowering is promoted by *PHYA* in far red enriched long-days for example, in *A. thaliana* (Johnson *et al*. 1994; Mockler *et al*. 2003), pea (Weller *et al*. 1997) and wheat (Carrsmith *et al*. 1994). In *A. lyrata*, Leinonen *et al*. (2013) found a QTL in the genomic region covering *PHYA*, where northern alleles promoted flowering in light conditions resembling early summer in northern Europe.

We found that the C-terminal half of the gene product was highly differentiated (3 nonsynonymous differences) between the northern and southern populations. C-terminal domains (PAS repeat domain and histidine kinase-related domain) mediate light signals to the nucleus and have an important role in transcription regulation and spectral sensitivity (Quail *et al*. 1995, Wang *et al*. 2011). All nonsynonymous mutations were derived compared to *A. thaliana* and it is possible that they have modified the function of *PHYA*, but this would require further study.

3.1.2 Vernalization pathway - FLC

We studied the expression of two duplicated *FLC* genes (*FLC1* and *FLC2*) in the same two populations. The *FLC1* gene was more highly expressed in Spiterstulen compared to the more southern Plech population before vernalization, but there was no difference after vernalization (Fig. 5 and 6 in II). Further, an expression quantitative trait locus (eQTL) covered the *FLC* region in a cross between the same populations (Fig. 7 in II).

We found that the *FLC1* gene was highly differentiated (F_{ST} = 0.62) between populations. The differentiation was almost two times larger compared to the genome-wide average (F_{ST} = 0.35, Pyhäjärvi *et al*. 2012). In Spiterstulen, a 350 bps

deletion was fixed at the promoter region and, in addition, 7 indels and 27 SNPs were fixed between populations, mostly located in the first intron (F_{ST} = 0.85, Fig. 8A in II). These regions are important for the regulation of *FLC* expression by cold temperatures (vernalization) repression (Sheldon *et al*. 2002, Helliwell *et al*. 2011). Fixed differences seemed to cover in total a 3.3 kb region along the promoter and first intron regions. Even if the regulatory regions were highly differentiated, the coding regions were identical.

Neutral diversity also showed unexpected pattern in the Spiterstulen population. Neutral diversity in Spiterstulen was less than 20% of that found in Plech (Fig. 8B, Table 3 in II). The reduction was very large, compared to the genome wide average, as Spiterstulen had on average slightly less than half the diversity of Plech (Pyhäjärvi *et al*. 2012). However, the neutral divergence (K_s =0.11) at the coding regions of *FLC1*, reflecting the mutation rate, was only slightly below the average between *A. thaliana* and *A. lyrata* (0.144, Pyhäjärvi *et al*. 2012) (0.147, Yang & Gaut 2011).

The difference in diversities between populations was largest in the promoter and in the first intron regions (Table 3 in II). Very low variation in the promoter region (Table 3, sequence set 4 in II) was unexpected because it was highly diverged from *A. thaliana* (aligning was impossible), suggesting a high mutation rate or perhaps that several indels have occurred after divergence. In contrast to Spiterstulen, the Plech population showed substantial variation in the same region (Fig. 8B, Table 3 in II). Strong background selection and drift can also remove variation (Charlesworth et al. 1993) but they rarely result in rapid fixation of large deletions and indels between populations, especially, if they are located in important regulatory regions. Alltogether, high genetic differentiation and low variation suggested recent hitchhiking at the *FLC1* regulatory regions (Maynard-Smith & Haigh 1974).

The K_a/K_s ratio between gene duplicates (0.27) and between each gene and the *A. thaliana FLC* (K_a/K_s= 0.28 for *FLC1* and 0.23 for *FLC2*) indicated that both genes are functional. In Spiterstulen population, however, some individuals have a non-functional *FLC2* gene (Fig. 3A in II) whereas in Plech *FLC1* is not functional in all individuals (Kemi 2013, Doctoral Dissertation). Gene duplication is one of the most important sources for adaptive evolution. Gene duplicates may increase expression diversity (Ha *et al*. 2009), which can be important for subfunctionalization (Force *et al*. 1999). For example, in *Brassica napus* the *FLC* homologues are expressed differently in vegetative and reproductive organs (Zou *et al*. 2012). Interestingly, the coding regions between the homologues are conserved but introns and promoter regions are diverged between duplicates. Also in Populus the *FT* paralogs are expressed in different life stages (Hsu *et al*. 2011).

To summarize, the results suggest that recent directional selection targeted the *FLC1* gene. The high altitude Spiterstulen population is facing long winters and short summers (growing season 3 months) and plants have to start flowering rapidly after snow melt in May, when days are already long. It is possible that that the high expression of *FLC1* gene is involved in strong vernalization requirement in Spiterstulen population. The high expression of *FLC1* might ensure that plants do not start flowering before the first winter and that only a long cold period lowers expression to a level adequate for flowering.

3.2 Response potential of marginal *A. lyrata* population

3.2.1 Plasticity

Flowering traits showed differences both between environments and between years in the natural Spiterstulen environment (2009 and 2010). In 2010, plants grown in the Oulu environment flowered 20 days earlier than the plants grown in Spiterstulen (Fig. 2, in III). In 2009, in Spiterstulen, plants flowered 13 days earlier than in 2010. The spring temperature was likely the most important factor determining the differences in flowering date (Vince-Prue & Thomas 1997) (but plants were also a year older). For example, in 2009, the average spring temperature was 1.7 degrees higher compared to the spring temperature in 2010 in Østlandet (same climatic region where Spiterstulen is located). During the last 35 years (1980-2014) the average spring temperature (March-May) has increased 1.5 degrees in the same climatic region (Norwegian Meteorological Institute). Flowering time may have become earlier during 35 years in the high altitude Spiterstulen population. Anderson *et al*. (2012) studied phenological changes in *Boechere stricta* plants growing in Rocky Mountains. In 40 years, there had been a significant increase in minimum temperatures during spring. They found that flowering date had advanced about 14 days during last 40 years. They estimated also that 80% of that shift was covered by plasticity.

Phenotypic plasticity was observed in the new environment, Oulu. Plants grown in Oulu had high reproductive success (Fig. 2 in III), whereas survival was lower than in the native environment, as only 70% plants survived after the first winter. The Oulu environment differs in several ways from the Norwegian environment. Oulu has a short growing season (our main focus here), but Oulu also is close to seashore at sea level, whereas Spiterstulen is a high altitude area, with very different vegetation. The Oulu conditions resulted in a higher reproductive result compared to the natural site,

which may contributed to lower survival over the next winter. Transplantation effects may also have differed. This kind of large distance transplantation studies can still provide interesting information on the effects of large scale climatic differences on the phenotypes.

An earlier study also showed the importance of phenotypic plasticity in the new environments (Vergeer & Kunin 2013). They examined the relative importance of planting site (i.e. phenotypic plasticity) and genetic changes (local adaptation) in reciprocal transplant experiments having *A. l. petraea* populations from Iceland, Sweden, Norway and UK. They found that the effect of planting site exceeded the population effect, which pinpoints the importance of phenotypic plasticity for survival in different environments.

3.2.2 Potential for genetic responses – genetic variation in quantitative traits

We found that additive genetic variation, especially in the timing traits, was low in the field conditions (Table 2 in III). The highest observed heritability for flowering date was only 0.11 in 2009 in the Spiterstulen natural environment (not statistically significant). Timing traits also had low evolvabilities (Table 2 in III). Vernalization and long days in the early summer resulted in that most plants flowered within a short time span within all field conditions (see Fig. 2 in III). The low V_A was thus likely partly due to the favorable conditions for flowering. In such conditions, the delaying effects of some genes are mostly not seen. Earlier studies have also shown that the Spiterstulen population responds more rapidly to long days after vernalization than Plech (Riihimäki & Savolainen 2004, II). This results in faster flowering compared to the southern Plech population (II). This differential response is likely an adaptation to the short growing season, as was also suggested Boudry *et al*. (2002). When snow melts in May and days are already long (16-17 h), plants respond rapidly to environmental cues of beginning summer.

While a population can deal with varying environmental challenges to some extent by phenotypic plasticity, when environmental changes exceed a critical threshold genetic changes are required (Chevin *et al*. 2010). Large populations usually have much standing genetic variation, which allows adaptive changes, whereas small populations can be dependent on new beneficial mutations (Pennings & Hermisson 2006).

Earlier studies have shown that northern populations of *A. lyrata* have lost genetic variation likely due to drift (Wright *et al*. 2003, Ross-Ibarra *et al*. 2008,

Muller *et al*. 2008) because colonization is associated with bottlenecks. Bottlenecks can result in reduced additive genetic variation as was demonstrated in *Mercurialis annua* populations located at the range margins (Pujol & Pannell 2008). Further, peripheral populations of *Chamaecrista fasciculate* harbored less genetic variation than central populations (Etterson 2004). We found that some morphological traits still had considerable V_A (Table 2 in III). This suggests that in addition to drift directional selection is a plausible explanation for a low genetic variation in the timing traits. Further, directional selection favors early flowerers currently, as was shown by Sandring *et al*. (2007) and which was demonstrated also in this study (Fig. 3 in III). Our other studies also show evidence of directional selection in northern populations at the individual loci of the flowering time pathway (Mattila, Aalto, Toivainen et al., manuscript in prep.), and the results in this theses show directional selection specifically in the *FLC* (II) and *PHYA* (I) genes.

To summarize, directional selection accompanying local adaptation after the LGM and the current directional selection for earlier flowering are plausible explanations for low additive genetic variation and evolvability in flowering date.

3.2.3 Photoperiodic pathway genes have small effects on fitness

We conducted association mapping with 21 well-known flowering time candidate genes (Supplementary Table1S in III) and 16 reference loci timing and fitness traits due to a selected set of genes.

In general, variation at flowering time genes was associated with fitness traits (rosette size 2008 in growth chamber, fruit production 2009, good seed production 2009) more than expected (Fig. 5 in III) during the first year after sowing (2008-2009). Some individual flowering time candidate genes stood out in results. The *FRIGIDA* gene associated most consistently with the timing traits in different years (and different conditions, Fig. 4 and 6 in III). This was not unexpected based on earlier studies in *A*. *thaliana* (Johanson 2000, Salomé *et al*. 2011) *Brassica napus* (Wang *et al*. 2011) and *A*. *lyrata* (Kuittinen *et al*. 2008). This finding also showed that this set of plants had statistical power to detect genetic variants underlying trait variation.

In 2008, in the growth chamber, rosette size was measured from all plants 8 weeks after the mean flowering date. The photoperiodic pathway loci showed some associations to this trait (Fig. 5 in III). It is possible that rosette size reflects differences in resource allocation strategies because there was a negative phenotypic correlation with flowering probability and rosettes size (Supplementary Fig. 4 in III).

These genes are involved in the pathways that integrate different environmental signals and control the progress to adopt flowering (reviewed in Andrés & Coupland 2012).

In 2009, flowering time was correlated with fruit production in Oulu - early flowerers produced more fruits (Fig. 3 in III). Consistently, we found that flowering time genes, especially in a photoperiodic pathway, were associated more than expected with production of fruits and good seeds (Fig. 5 in III).

In 2010, flowering date and fruit number were not correlated in the Oulu environment (Fig. 3 in III). Rosette size at flowering date (after two year of growth) was a more important determinant of reproductive fitness (Supplementary Fig. 4 in III). In agreement with this the flowering time genes were not associated more than expected with reproductive fitness traits (Fig. 5 in III). Because Oulu was a new environment to the plants, some other traits were perhaps favored. It is also possible that plants had already lost vigor (survival had been much lower than in the native site). In 2010, however, in natural Spiterstulen conditions, flowering time was still correlated with fitness (Fig. 3 in III), as was demonstrated also by Sanding et al. (2007).

Because genetic correlations were low between the same traits in different environments (Supplementary Fig. 4 in III), the relevance of associations in natural conditions is hard to predict. However, some individual genes showed associations in different conditions or with different traits. *FT* and *TOC1* were associated with fruit production in 2009 and survival in 2010, and the *FT* gene with rosette size in 2008 and the *TOC1* gene with number of flowers in 2008 (Fig. 5 and 6 in III).

3.2.4 Can Spiterstulen population respond to changing environment in respect of flowering time?

Flowering traits showed plasticity which has an important role in changing environment, especially if the population size is small and adaptive genetic variation is not available. Additive genetic variation for flowering date was low, which suggests that, new mutations would be required for evolutionary response. However, low frequency alleles may not contribute much to additive genetic variation, (h^2 or V_A), but in a changed environment they might still be important. For example, in 2009, single nucleotide polymorphisms (SNPs) at *PHYB* and *FRI* genes were significantly associated with flowering date (Fig. 4 in III). However, overall there was no significant additive genetic variation. The associated loci had very low minor allele frequencies (0.04, 0.05), and overall there was no sign of heritable variation at the

quantitative genetic level. At the population level there might exist rare variants, which cannot be detected by traditional quantitative genetics methods.

Spiterstulen is located in a valley surrounded by high mountains. Thus, as a response to warming climate, it would also be possible to disperse to more high altitudes to maintain current conditions and as a low competitor, to escape competitors. In Norway, *A. lyrata* occurs at higher elevations than in Spiterstulen (Gaudeul *et al*. 2007), which suggests that seed migration is feasible. Within the Spiterstulen population there is very little spatial structure, between sites located about 1 km for each other (Lundemo *et al*. 2010). This shows that within such short distances, gene flow is possible. This would facilitate cross-pollination of the self-incompatible species during dispersal. Thus, the population might be able to expand to higher elevations.

4 Conclusions

Adaptation to the northern conditions has involved genetic changes e.g. in photoperiodic and temperature signaling pathways. In. *A. lyrata*, selection has targeted individual loci, such as *PHYA* and *FLC1*, which has resulted in selective sweeps across c.a. 10 kb and 3 kb chromosomal regions, respectively. Three nonsynonymous fixed differences at the *PHYA* locus suggest that some structural changes underlie the selective advantage. At the *FLC1* gene regulatory regions were highly differentiated coding regions being identical and *FLC1* gene expression is altered in the northern Spiterstulen population. The functional roles of mutations are not known, but other studies suggest that they could be closely related to the adaptation to the short growing season. Thus, functional studies would be needed to uncover the adaptive physiological mechanisms. Directional selection for adaptation to the Spiterstulen conditions and current directional selection towards earlier flowering has resulted in low genetic variation in flowering traits in the northern Spiterstulen population. Thus the genetic response potential is low. We did not find strong associations even though the studied flowering time genes associated with fitness more than expected during the first of growth in Oulu. As was shown in this study flowering date is highly dependent on environmental signals, especially temperature sum, and this plasticity is an important buffer for changing environment. However, more detailed studies concerning phenotypic plasticity would be needed.

References

Ågren J & Schemske DW (2012) Reciprocal transplants demonstrate strong adaptive differentiation of the model organism *Arabidopsis thaliana* in its native range. New Phytologist 194: 1112–1122.

Alberto FJ, Aitken SN, Alia R, Gonzalez-Martinez SC, Hanninen H, Kremer A, Lefevre F, Lenormand T, Yeaman S, Whetten R & Savolainen O (2013) Potential for evolutionary responses to climate change evidence from tree populations. Global Change Biology 19: 1645–1661.

Anderson JT, Inouye DW, McKinney AM, Colautti RI & Mitchell-Olds T (2012) Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. Proceedings of the Royal Society B-Biological Sciences 279: 3843–3852.

Andrés F & Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics 13: 627–639.

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J & Nordborg M (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465: 627– 631.

Balasubramanian S, Sureshkumar S, Agrawal M, Michael TP, Wessinger C, Maloof JN, Clark R, Warthmann N, Chory J & Weigel D (2006) The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of *Arabidopsis thaliana*. Nature Genetics 38: 711–715.

Balding DJ (2006) A tutorial on statistical methods for population association studies. Nature Review Genetics 7: 781–791.

Barreiro LB, Laval G, Quach H, Patin E & Quintana-Murci L (2008) Natural selection has driven population differentiation in modern humans. Nature Genetics 40: 340–345.

Barton NH (1995) Linkage and the limits to natural selection. Genetics 140: 821–841.

Beales J, Turner A, GriYths S, Snape JW & Laurie DA (2007) A Pseudo-response regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 115: 721–733.

Beilstein MA, Nagalingum NS, Clements MD, Manchester SR & Mathews S (2010) Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America 107: 18724–18728.

Berg JJ & Coop G (2014) A population genetic signal of polygenic adaptation. PLoS Genetics 10: e1004412–e1004412.

Björck S (1995) A review of the history of the Baltic Sea, 13.0-8.0 ka BP. Quaternary International 27: 19–40.

Blázquez M (2000) Flower development pathways. Journal of Cell science 113: 3547-3548.

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

Boudry P, McCombie H & Van Dijk H (2002) Vernalization requirement of wild beet *Beta vulgaris* ssp *maritima*: among population variation and its adaptive significance. Journal of Ecology 90: 693–703.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y & Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633–2635.

Bradshaw AD (1965) Evolutionary significance of phenotypic plasticity in plants. Advances in Genetics: 13: 115–155

Bradshaw WE & Holzapfel CM (2001) Genetic shift in photoperiodic response correlated with global warming. Proceedings of the National Academy of Sciences of the United States of America 98: 14509–14511.

Braverman JM, Hudson RR, Kaplan NL, Langley CH & Stephan W (1995) The hitchhiking effect on the site frequency-spectrum of DNA polymorphisms. Genetics 140: 783–796.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Cinta Romay M, Romero S, Salvo S, Sanchez Villeda H, da Silva HS, Sun Q, Tian F, Upadyayula N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S & McMullen MD (2009) The genetic architecture of maize flowering time. Science 325: 714–718.

Cai JJ, Macpherson JM, Sella G, Petrov DA (2009) Pervasive hitchhiking at coding and regulatory sites in humans. PLoS Genetics 5:e10003.

Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Müller J, Alonso-Blanco C, Borgwardt K, Schmid KJ & Weigel D (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nature Genetics 43: 956–963.

Carneiro M, Rubin C, Di Palma F, Albert FW, Alfoeldi J, Barrio AM, Pielberg G, Rafati N, Sayyab S, Turner-Maier J, Younis S, Afonso S, Aken B, Alves JM, Barrell D, Bolet G, Boucher S, Burbano HA, Campos R, Chang JL, Duranthon V, Fontanesi L, Garreau H, Heiman D, Johnson J, Mage RG, Peng Z, Queney G, Rogel-Gaillard C, Ruffier M, Searle S, Villafuerte R, Xiong A, Young S, Forsberg-Nilsson K, Good JM, Lander ES, Ferrand N, Lindblad-Toh K & Andersson L (2014) Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. Science 345: 1074–1079.

Carrsmith H, Johnson C, Plumpton C, Butcher G & Thomas B (1994) The kinetics of type-1 phytochrome in green, light-grown wheat (*Triticum-Aestivum* L.). Planta 194: 136–142.

Charlesworth, B, Morgan, MT, Charlesworth, D (1993) The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–303

Charlesworth B & Hughes KA (1996) Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence. Proceedings of the National Academy of Sciences of the United States of America 93: 6140–6145.

Chen J, Tsuda Y, Stocks M, Källman T, Xu N, Karkkainen K, Huotari T, Semerikov VL, Vendramin GG & Lascoux M (2014) Clinal variation at phenology-related genes in spruce: parallel evolution in *FTL2* and *Gigantea*? Genetics 197: 1025–1038.

Chevin L, Lande R & Mace GM (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS Biology 8: e1000357.

Corbesier L & Coupland G (2006) The quest for florigen: a review of recent progress. Journal of Experimental Botany 57: 3395–3403.

Cresko WA, Amores A, Wilson C, Murphy J, Currey M, Phillips P, Bell MA, Kimmel CB & Postlethwait JH (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. Proceedings of the National Academy of Sciences of the United States of America 101: 6050–6055.

Crnokrak P & Roff DA (1995) Dominance Variance - Associations with Selection and Fitness. Heredity 75: 530–540.

Csilléry K, Blum MGB, Gaggiotti OE & Francois O Approximate Bayesian Computation (ABC) in practice. Trends in Ecology & Evolution 25: 410–418.

Darwin C (1859) On the Origin of Species by Means of natural Selection or the Preservation of Favoured Races in the Struggle of Life. London, John Murray.

Davis MB & Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. Science 292: 673–679.

Devlin PF & Kay SA (2000) Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. Plant Cell Online 12: 2499–2509.

Eckert AJ, Bower AD, Wegrzyn JL, Pande B, Jermstad KD, Krutovsky KV, St. Clair JB & Neale DB (2009) Association Genetics of Coastal Douglas Fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-Hardiness Related Traits. Genetics 182: 1289–1302.

Etterson JR & Shaw RG (2001) Constraint to adaptive evolution in response to global warming. Science 294: 151–154.

Etterson JR & Willis J (2004) Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. ii. genetic architecture of three populations reciprocally planted along an environmental gradient in the great plains. Evolution 58: 1459–1471.

Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics, Ed 4. Harlow, Essex, Longmans Green.

Fay JC & Wu CI (2000) Hitchhiking under positive Darwinian selection. Genetics 155: 1405– 1413.

Fisher RA (1930) The genetical theory of natural selection. Oxford, Clarendon Press.

Foll M & Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. Genetics 180: 977–993.

Force A, Lynch M, Pickett FB, Amores A, Yan Y & Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151: 1531–1545.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J & Wilczek AM (2011) A map of local adaptation in *Arabidopsis thaliana*. Science 334: 86–89.

Franks SJ, Sim S & Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. Proceedings of the National Academy of Sciences of the United States of America 104: 1278–1282.

Franks SJ, Weber JJ & Aitken SN (2014) Evolutionary and plastic responses to climate change in terrestrial plant populations. Evolutionary Applications 7: 123–139.

Fu YX & Li WH (1993) Statistical Tests of Neutrality of Mutations. Genetics 133: 693–709.

Garris AJ, McCouch SR & Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the *xa5* locus of rice (Oryza sativa L.). Genetics 165: 759–769.

Gaudeul M, Stenøien HK & Ågren J (2007) Landscape structure, clonal propagation, and genetic diversity in Scandinavian populations of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany 94: 1146–1155.

Gienapp P, Teplitsky C, Alho JS, Mills JA & Merila J (2008) Climate change and evolution: disentangling environmental and genetic responses. Molecular Ecology 17: 167–178.

Gillespie JH & Turelli M (1989) Genotype-environment interactions and the maintenance of polygenic variation. Genetics 121: 129–138.

Goodnight CJ (1988) Epistasis and the effect of founder events on the additive genetic variance. Evolution 42: 441–454

Ha M, Kim E & Chen ZJ (2009) Duplicate genes increase expression diversity in closely related species and allopolyploids. Proceedings of the National Academy of Sciences 106: 2295–2300.

Haldane JBS (1927) A mathematical theory of natural and artificial selection, Part V: selection and mutation. Mathematical Proceedings of the Cambridge Philosophical Society 23: 838–844

Haddrill PR, Halligan DL, Tomaras D & Charlesworth B (2007) Reduced efficacy of selection in regions of the Drosophila genome that lack crossing over. Genome Biology 8: R18.

Hall D, Ma X & Ingvarsson PK (2011) Adaptive evolution of the *Populus tremula* photoperiod pathway. Molecular Ecology 20: 1463–1474.

Hall MC & Willis JH (2006) Divergent selection on flowering time contributes to local adaptation in *Mimulus guttatus* populations. Evolution 60: 2466– 2477.

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F & Bergelson J (2011) Adaptation to climate across the *Arabidopsis thaliana* genome. Science 334: 83–86.

Hardy OJ & Vekemans X (2002) SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2: 618–620.

Hedrick PW (1999) Antagonistic pleiotropy and genetic polymorphism: a perspective. Heredity 82: 126–133.

Helliwell CA, Robertson M, Finnegan EJ, Buzas DM & Dennis ES (2011) Vernalization-repression of Arabidopsis *FLC* requires promoter sequences but not antisense transcripts. PLoS One 6(6): e21513.

Hereford J (2010) Does selfing or outcrossing promote local adaptation? American Journal of Botany 97: 298–302.

Hernandez RD, Kelley JL, Elyashiv E, Melton SC, Auton A, McVean G, 1000 Genomes Project, Sella G & Przeworski M (2011) Classic selective sweeps were rare in recent human evolution. Science 331: 920–924.

Hewitt G (1999) Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68: 87–112.

Hill WG & Robertson A (1966) The effect of linkage on limits to artificial selection. Genetical Research 8: 269–294.

Hill WG & Robertson A (1968) Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38: 226–231.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA & Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse colour pattern. Science 313: 101–104.

Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjálmsson BJ, Nordborg M, Borevitz JO & Bergelson J (2012) Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. Nature Genetics 44: 212–216.

Houle D (1992) Comparing evolvability and variability of quantitative traits. Genetics 130: 195–204.

Hsu C, Adams JP, Kim H, No K, Ma C, Strauss SH, Drnevich J, Vandervelde L, Ellis JD, Rice BM, Wickett N, Gunter LE, Tuskan GA, Brunner AM, Page GP, Barakat A, Carlson JE, dePamphilis CW, Luthe DS & Yuceer C (2011) *FLOWERING LOCUS T* duplication coordinates reproductive and vegetative growth in perennial poplar. Proceedings of the National Academy of Sciences of the United States of America 108: 10756–10761.

Huang W, Richards S, Carbone MA, Zhu D, Anholt RRH, Ayroles JF, Duncan L, Jordan KW, Lawrence F, Magwire MM, Warner CB, Blankenburg K, Han Y, Javaid M, Jayaseelan J, Jhangiani SN, Muzny D, Ongeri F, Perales L, Wu Y, Zhang Y, Zou X, Stone EA, Gibbs RA & Mackay TFC (2012) Epistasis dominates the genetic architecture of Drosophila quantitative traits. Proceedings of the National Academy of Sciences of the United States of America 109: 15553–15559.

Huber CD, Nordborg M, Hermisson J & Hellmann I (2014) Keeping it local: evidence for positive selection in Swedish *Arabidopsis thaliana*. Molecular Biology and Evolution 31: 3026–39.

Hudson RR (1991) Gene genealogies and the coalescent process. Oxford Survey in Evolutionary Biology 7: 1–44.

Hudson R, Slatkin M & Maddison W (1992) Estimation of levels of gene flow from DNA-sequence data. Genetics 132: 583–589.

Hudson RR (2002) Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics 18: 337–338.

Hung H, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, McMullen MD, Ware D, Buckler ES, Doebley JF & Holland JB (2012) *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proceedings of the National Academy of Sciences of the United States of America 109: E1913–E1921.

Ikeda H, Fujii N & Setoguchi H (2009) Molecular evolution of phytochromes in *Cardamine nipponica* (Brassicaceae) suggests the involvement of *PHYE* in local adaptation. Genetics 182: 603–614.

Ingvarsson P, Garcia M, Hall D, Luquez V & Jansson S (2006) Clinal variation in *phyB2*, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics 172: 1845–1853.

Ingvarsson PK, Garcia MV, Luquez V, Hall D & Jansson S (2008) Nucleotide polymorphism and phenotypic associations within and around the *phytochrome B2* locus in European aspen (*Populus tremula*, Salicaceae). Genetics 178: 2217–2226.

Ingvarsson PK (2010) Natural selection on synonymous and nonsynonymous mutations shapes patterns of polymorphism in *Populus tremula*. Molecular Biology and Evolution 27: 650–660.

Jensen JD, Kim Y, DuMont VB, Aquadro CF & Bustamante CD (2005) Distinguishing between selective sweeps and demography using DNA polymorphism data. Genetics 170: 1401–1410.

Jensen JD & Bachtrog D (2011) Characterizing the Influence of Effective Population Size on the Rate of Adaptation: Gillespie's Darwin Domain. Genome Biology and Evolution 3: 687–701.

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

Johnson E, Bradley M, Harberd N & Whitelam G (1994) Photoresponses of light-grown *PHYA* mutants of Arabidopsis – Phytochrome A is required for the perception of daylength extensions. Plant Physiology 105: 141–149.

Kawecki TJ & Ebert D (2004) Conceptual issues in local adaptation. Ecology Letters 7: 1225–1241.

Keinan A & Reich D (2010) Human population differentiation is strongly correlated with local recombination rate. PLoS Genetics 6: e1000886.

Keller SR, Levsen N, Olson MS & Tiffin P (2012) Local adaptation in the flowering–time gene network of balsam poplar, *Populus balsamifera* L. Molecular Biology and Evolution 29: 3143–3152.

Kemi U (2013) Adaptation to growing season length in the perennial *Arabidopsis lyrata*. *Acta Universitatis Ouluensis A616. Pub. thesis*.

Kelly JK (1997) A test of neutrality based on interlocus associations. Genetics 146: 1197–1206.

Kim Y & Stephan W (2002) Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics 160: 765–777.

Kim Y & Nielsen R (2004) Linkage disequilibrium as a signature of selective sweeps. Genetics 167: 1513–1524.

Kimura M (1962) On the probability of fixation of mutant genes in a population. Genetics 47: 713–719.

Kimura M (1968) Evolutionary rate at the molecular level. Nature 217: 624– 626.

Kimura M (1971) Theoretical foundation of population genetics at the molecular level. Theoretical Population Biology 2: 174–208.

Kimura, M. (1983). The neutral theory of molecular evolution. Cambridge, Cambridge University Press.

Kingman JFC (1982) On the genealogy of large populations. Journal of Applied Probability 19A: 27–43

Krajick K (2004) Climate change: All downhill from here? Science 303: 1600–1602.

Kuittinen H, Niittyvuopio A, Rinne P & Savolainen O (2008) Natural variation in *Arabidopsis lyrata* vernalization requirement conferred by a *FRIGIDA* indel polymorphism. Molecular Biology and Evolution 25: 319–329.

Källman T, De Mita S, Larsson H, Gyllenstrand N, Heuertz M, Parducci L, Suyama Y, Lagercrantz U & Lascoux M (2014) Patterns of nucleotide diversity at photoperiod related genes in Norway Spruce [*Picea abies* (L.) Karst.]. PLoS One 9: e95306.

Lagercrantz U & Axelsson T (2000) Rapid evolution of the family of *CONSTANS* like genes in plants. Molecular Biology and Evolution 17: 1499–1507.

Lande R, Arnold SJ (1983) The measurement of selection on correlated characters. Evolution 37: 1210–1226.

Le Corre V, Roux F & Reboud X (2002) DNA polymorphism at the *FRIGIDA* gene in Arabidopsis thaliana: Extensive nonsynonymous variation is consistent with local selection for flowering time. Molecular Biology and Evolution 19: 1261–1271.

Leimu R & Fischer M (2008) A meta-analysis of local adaptation in plants. PLoS One 3: e4010.

Leinonen PH, Sandring S, Quilot B, Clauss MJ, Mitchell-Olds T, Ågren J & Savolainen O (2009) Local adaptation in European populations of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany 96: 1129–1137.

Leinonen PH, Remington DL & Savolainen O (2011) Local adaptation, phenotypic differentiation, and hybrid fitness in diverged natural populations of *Arabidopsis lyrata*. Evolution 65: 90–107.

Leinonen PH, Remington DL, Leppälä J & Savolainen O (2013) Genetic basis of local adaptation and flowering time variation in *Arabidopsis lyrata*. Molecular Ecology 22: 709–723.

Levene H (1953) Genetic equilibrium when more than one ecological niche is available. American Naturalist 87: 331–333

Li Y, Haseneyer G, Schon C, Ankerst D, Korzun V, Wilde P & Bauer E (2011) High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (*Secale cereale* L.) genes involved in frost response. BMC Plant Biology 11: 6.

Long Q, Rabanal FA, Meng D, Huber CD, Farlow A, Platzer A, Zhang Q, Vilhjalmsson BJ, Korte A, Nizhynska V, Voronin V, Korte P, Sedman L, Mandakova T, Lysak MA, Seren U, Hellmann I & Nordborg M (2013) Massive genomic variation and strong selection in *Arabidopsis thaliana* lines from Sweden. Nature Genetics 45: 884–U218.

Lundemo S, Stenoien HK & Savolainen O (2010) Investigating the effects of topography and clonality on genetic structuring within a large Norwegian population of *Arabidopsis lyrata*. Annals of Botany 106: 243–254.

Lynch M, and Walsh B (1998) Genetics and analysis of quantitative traits. Sunderland, MA, Sinauer Associates.

Ma X, Hall D, Onge KRS, Jansson S & Ingvarsson PK (November 2010) Genetic differentiation, clinal variation and phenotypic associations with growth cessation across the *Populus tremula* photoperiodic pathway. Genetics 186: 1033–1044.

Mackay TFC (2014) Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nature Review Genetics 15: 22-33.

Matesanz S, Horgan-Kobelski T & Sultan SE (2012) Phenotypic plasticity and population differentiation in an ongoing species invasion. PLoS One 7: e44955.

Maynard Smith J & Haigh J (1974) The hitch-hiking effect of a favorable gene. Genetical Research 23: 23–35.

Mendez-Vigo B, Xavier Pico F, Ramiro M, Martinez-Zapater JM & Alonso-Blanco C (2011) Altitudinal and climatic adaptation is mediated by flowering traits and *FRI*, *FLC*, and *PHYC* genes in Arabidopsis. Plant Physiology 157: 1942–1955.

Mockler T, Yang H, Yu X, Parikh D, Cheng YC, Dolan S & Lin C (2003) Regulation of photoperiodic flowering by Arabidopsis photoreceptors. Proceedings of the National Academy of Sciences of the United States of America 100: 2140–2145.

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: Interacting pathways as a basis for diversity. Plant Cell 14:S111–S130.

Muller MH, Leppälä J & Savolainen O (2008) Genome-wide effects of postglacial colonization in *Arabidopsis lyrata*. Heredity 100: 47–58.

Neher RA, Shraiman BI & Fisher DS (2009) Rate of adaptation in large sexual populations. Genetics 184: 467–481.

Nei M & Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Molecular Biology and Evolution 3: 418–426.

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F & van Kleunen M (2010) Plant phenotypic plasticity in a changing climate. Trends in Plant Science 15: 684–692.

Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG & Bustamante C (2005) Genomic scans for selective sweeps using SNP data. Genome Research 15: 1566–1575.

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA & Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nature Genetics 30: 190–193.

Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246: 96–98.

Orr HA (1998) The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. Evolution 52: 935–949.

Orr HA & Whitlock M (2002) The population genetics of adaptation: the adaptation of DNA sequences. Evolution 56: 1317–1330.

Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D & Lynch M (2010) The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. Science 327: 92–94.

Palumbi SR, Barshis DJ, Traylor-Knowles N & Bay RA (2014) Mechanisms of reef coral resistance to future climate change. Science 344: 895–898.

Pavlidis P, Jensen JD & Stephan W (2010) Searching for footprints of positive selection in whole-genome SNP data from non-equilibrium populations. Genetics 185: 907–922.

Pavlidis P, Zivkovic D, Stamatakis A & Alachiotis N (2013) SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. Molecular Biology and Evolution 30: 2224–2234.

Pennings PS & Hermisson J (2006) Soft sweeps II-molecular population genetics of adaptation from recurrent mutation or migration. Molecular Biology and Evolution 23: 1076–1084.

Pfaffelhuber P, Lehnert A & Stephan W (2008) Linkage disequilibrium under genetic hitchhiking in finite populations. Genetics 179: 527–537.

Phifer-Rixey M, Bonhomme F, Boursot P, Churchill GA, Piálek J, Tucker P & Nachman M (2012) Adaptive evolution and effective population size in wild house mice. Molecular Biology and Evolution 29: 2949–2955.

Presgraves DC (2005) Recombination enhances protein adaptation in *Drosophila melanogaster*. Current Biology 15: 1651–1656.

Pujol B, Pannell JR (2008) Reduced responses to selection after species range expansion. Science 321:96–96.

Pyhäjärvi T, Aalto E & Savolainen O (2012) Time scales of divergence and speciation among natural populations and subspecies of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany 99: 1314–1322.

Quail PH, Boylan MT, Parks BM, Short TW, Xu Y & Wagner D (1995) Phytochromes: photosensory perception and signal transduction. Science 268: 675–680.

Quilot-Turion B, Leppälä J, Leinonen PH, Waldmann P, Savolainen O & Kuittinen H (2013) Genetic changes in flowering and morphology in response to adaptation to a high-latitude environment in *Arabidopsis lyrata*. Annals of Botany 111: 957–968.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM & Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proceedings of the National Academy of Sciences of the United States of America 98: 11479–11484.

Riihimäki M & Savolainen O (2004) Environmental and genetic effects on flowering differences between northern and southern populations of *Arabidopsis lyrata* (Brassicaceae). American Journal of Botany 91: 1036 –1045.

Riihimäki M, Podolsky R, Kuittinen H, Koelewijn H & Savolainen O (2005) Studying genetics of adaptive variation in model organisms: flowering time variation in *Arabidopsis lyrata*. Genetica 123: 63–74.

Robertson A (1955) Selection in animals: synthesis. Cold Spring Harbor Symposium Quantitative Biology 20: 225–229

Rose MR (1985) Life history evolution with antagonistic pleiotropy and overlapping generations. Theoretical Population Biology 28: 342–358.

Ross-Ibarra J, Wright SI, Foxe JP, Kawabe A, DeRose-Wilson L, Gos G, Charlesworth D, Gaut BS (2008) Patterns of polymorphism and demographic history in natural populations of *Arabidopsis lyrata*. PLoS One 3:e2411.

Rozas J (2009) DNA sequence polymorphism analysis using DnaSP. Methods in Molecular Biology 537: 337–350.

Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R & Lander ES (2002) Detecting recent positive selection in the human genome from haplotype structure. Nature 419: 832–837.

Salomé PA, Bomblies K, Laitinen RAE, Yant L, Mott R & Weigel D (2011) Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. Genetics 188: 421–433.

Sandring S, Riihimäki M, Savolainen O, Ågren J (2007) Selection on flowering time and floral display in an alpine and a lowland population of *Arabidopsis lyrata*. Journal of Evolutionary Biology 20: 558–567.

Sattath S, Elyashiv E, Kolodny O, Rinott Y & Sella G (2011) Pervasive adaptive protein evolution apparent in diversity patterns around amino acid substitutions in *Drosophila simulans*. PLoS Genetics 7: e1001302.

Savolainen O, Pyhäjärvi T & Knürr T (2007) Gene flow and local adaptation in trees. Annual Review of Ecology, Evolution, and Systematics 38: 595–619.

Schmickl R, Jorgensen MH, Brysting AK & Koch MA (2010) The evolutionary history of the *Arabidopsis lyrata* complex: a hybrid in the amphi-Beringian area closes a large distribution gap and builds up a genetic barrier. BMC Evolutionary Biology 10: 98.

Sharrock RA (2008) The phytochrome red/far-red photoreceptor superfamily. Genome Biology 9:230

Shaw RG & Shaw FH (2014) Quantitative genetic study of the adaptive process. Heredity 112: 13–20.

Sheldon CC, Conn AB, Dennis ES & Peacock WJ (2002) Different regulatory regions are required for the vernalization-induced repression of *FLOWERING LOCUS C* and for the epigenetic maintenance of repression. Plant Cell 14: 2527–2537.

Slotte T, Holm K, McIntyre LM, Lagercrantz U & Lascoux M (2007) Differential expression of genes important for adaptation in *Capsella bursa-pastoris* (Brassicaceae). Plant Physiology 145: 160–173.

Somers DE, Devlin P, Kay SA (1998) Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. Science 282: 1488–1490

Somers DJ, Banks T, DePauw R, Fox S, Clarke J, Pozniak C & McCartney C (2007) Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. Genome 50: 557–567.

Stephan W, Wiehe T & Lenz M (1992) The effect of strongly selected substitutions on neutral polymorphism: Analytical results based on diffusion theory. Theoretical Population Biology 41: 237–254.

Stephan W (2010) Genetic hitchhiking versus background selection: the controversy and its implications. Philosophical Transactions of the Royal Society B-Biological Sciences 365: 1245–1253.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD & Schmitt J (2004) A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. Proceedings of the National Academy of Sciences of the United States of America 101: 4712–4717.

Strasburg JL, Kane NC, Raduski AR, Bonin A, Michelmore R & Rieseberg LH (2011) Effective population size is positively correlated with levels of adaptive divergence among annual sunflowers. Molecular Biology and Evolution 28: 1569–1580.

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

Sultan SE (2001) Phenotypic plasticity for fitness components in Polygonum species of contrasting ecological breadth. Ecology 82: 328–343.

Tadege M, Sheldon CC, Helliwell CA, Stoutjesdijk P, Dennis ES & Peacock WJ (2001) Control of flowering time by *FLC* orthologues in *Brassica napus*. The Plant Journal 28: 545–553.

Tajima F (1983) Evolutionary relationship of DNA-Sequences in finite populations. Genetics 105: 437–460.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585–595.

Thomas B, Vince-Prue D (1997) Photoperiodism in plants. London, Academic Press.

Thornton KR & Jensen JD (2007) Controlling the false-positive rate in multilocus genome scans for selection. Genetics 175: 737–750.

Turchin MC, Chiang CWK, Palmer CD, Sankararaman S, Reich D, Hirschhorn JN & Genetic Invest ANthropometric (2012) Evidence of widespread selection on standing variation in Europe at height-associated SNPs. Nature Genetics 44: 1015–1019.

Turner A, Beales J, Faure S, Dunford RP & Laurie DA (2005) The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. Science 310: 1031–1034.

Wakeley J (2008) Coalescent Theory: An Introduction. Greenwood Village, CO. Roberts & Company Publishers.

Valverde F (2011) *CONSTANS* and the evolutionary origin of photoperiodic timing of flowering. Journal of Experimental Botany 62: 2453–2463.

Vergeer P & Kunin WE (2013) Adaptation at range margins: common garden trials and the performance of *Arabidopsis lyrata* across its northwestern European range. New Phytologist 197: 989–1001.

Via S & Lande R (1987) Evolution of genetic variability in a spatially heterogeneous environment: effects of genotype–environment interaction. Genetics Research 49: 147–156.

Visscher PM (2008) Sizing up human height variation. Nature Genetics 40: 489–490.

Voight BF, Kudaravalli S, Wen XQ & Pritchard JK (2006) A map of recent positive selection in the human genome. PLoS Biology 4(4): 659–659.

Wang N, Qian W, Suppanz I, Wei L, Mao B, Long Y, Meng J, Mueller AE & Jung C (2011) Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the *FRIGIDA* homologue *BnaA.FRI.a*. Journal of Experimental Botany 62: 5641–5658.

Wang X, Roig-Villanova I, Khan S, Shanahan H, Quail PH, Martinez-Garcia JF & Devlin PF (2011) A novel high-throughput in vivo molecular screen for shade avoidance mutants identifies a novel *PHYA* mutation. Journal of Experimental Botany 62: 2973–2987.

Watterson G (1975) On the number of segregating sites in genetical models without recombination. Theoretical Population Biology 7: 256–276.

Wellcome Trust Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.

Weller JL, Murfet IC & Reid JB (1997) Pea mutants with reduced sensitivity to far-red light define an important role for *Phytochrome A* in day-length detection. Plant Physiology 114: 1225–1236.

Wright S (1931) Evolution in mendelian populations. Genetics 16: 97–159.

Wright SI, Lauga B & Charlesworth D (2003) Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. Molecular Ecology 12: 1247–1263.

Wright SI & Charlesworth B (2004) The HKA test revisited: a maximum-likelihood-ratio test of the standard neutral model. Genetics 168: 1071–1076.

Wright SI, Foxe JP, DeRose-Wilson L, Kawabe A, Looseley M, Gaut BS & Charlesworth D (2006) Testing for effects of recombination rate on nucleotide diversity in natural populations of *Arabidopsis lyrata*. Genetics 174: 1421–1430.

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X & Zhang Q (2008) Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. Nature Genetics 40: 761–767.

Yang L & Gaut BS (2011) Factors that contribute to variation in evolutionary rate among Arabidopsis Genes. Molecular Biology and Evolution 28: 2359–2369.

Yanovsky M & Kay S (2002) Molecular basis of seasonal time measurement in Arabidopsis. Nature 419: 308–312.

Yeaman S & Whitlock MC (2011) The genetic architecture of adaptation under migration-selection balance. Evolution 65: 1897–1911.

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S & Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38: 203-208.

Yu J, Holland JB, McMullen MD & Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178: 539–551.

Yu X, Liu H, Klejnot J & Lin C (2010) The cryptochrome blue light receptors. The Arabidopsis Book: e0135.

Zhou Y, Zhang L, Liu J, Wu G & Savolainen O (2014) Climatic adaptation and ecological divergence between two closely related pine species in Southeast China. Molecular Ecology 23: 3504–3522.

Zou X, Suppanz I, Raman H, Hou J, Wang J, Long Y, Jung C & Meng J (2012) Comparative analysis of *FLC* homologues in Brassicaceae provides insight into their role in the evolution of oilseed rape. PLoS One 7: e45751.

Original articles

I Toivainen T, Pyhäjärvi T, Niittyvuopio A, Savolainen O (2014) A recent local sweep at the *PHYA* locus in the northern European Spiterstulen population of *Arabidopsis lyrata*. Molecular Ecology 23: 1040–1052.

II Kemi U, Niittyvuopio A, Toivainen T, Pasanen A, Quilot-Turion B, Holm K, Lagercrantz U, Savolainen O, Kuittinen H (2013) Role of vernalization and of duplicated *FLOWERING LOCUS C* in the perennial *Arabidopsis lyrata*. New Phytologist 197: 323–335.

III Toivainen T, Vesimäki T, Remula S, Remington D, Kuittinen H, Savolainen O (2014) A marginal *Arabidopsis lyrata* population has low genetic variation but is phenotypically plastic in flowering traits. Manuscript.

Original publications are not included in the electronic version of the dissertation.

627. Jaakkonen, Tuomo (2014) Intra- and interspecific social information use in nest site selection of a cavity-nesting bird community

628. Päätalo, Heli (2014) Stakeholder interactions in cross-functional productization : the case of mobile software development

629. Koskela, Timo (2014) Interaction in asset-based value creation within innovation networks : the case of software industry

630. Stibe, Agnis (2014) Socially influencing systems : persuading people to engage with publicly displayed Twitter-based systems

631. Sutor, Stephan R. (2014) Large-scale high-performance video surveillance

632. Niskanen, Alina (2014) Selection and genetic diversity in the major histocompatibility complex genes of wolves and dogs

633. Tuomikoski, Sari (2014) Utilisation of gasification carbon residues : activation, characterisation and use as an adsorbent

634. Hyysalo, Jarkko (2014) Supporting collaborative development : cognitive challenges and solutions of developing embedded systems

635. Immonen, Ninna (2014) Glaciations and climate in the Cenozoic Arctic : evidence from microtextures of ice-rafted quartz grains

636. Kekkonen, Päivi (2014) Characterization of thermally modified wood by NMR spectroscopy : microstructure and moisture components

637. Pietilä, Heidi (2014) Development of analytical methods for ultra-trace determination of total mercury and methyl mercury in natural water and peat soil samples for environmental monitoring

638. Kortelainen, Tuomas (2014) On iteration-based security flaws in modern hash functions

639. Holma-Suutari, Anniina (2014) Harmful agents (PCDD/Fs, PCBs, and PBDEs) in Finnish reindeer (*Rangifer tarandus tarandus*) and moose (*Alces alces*)

640. Lankila, Tiina (2014) Residential area and health : a study of the Northern Finland Birth Cohort 1966

641. Zhou, Yongfeng (2014) Demographic history and climatic adaptation in ecological divergence between two closely related parapatric pine species

642. Kraus, Klemens (2014) Security management process in distributed, large scale high performance systems

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/

UNIVERSITY OF OULU P.O. Box 8000 FI-90014 UNIVERSITY OF OULU FINLAND

ACTA UNIVERSITATIS OULUENSIS

ISBN 978-952-62-0689-9 (Paperback)
ISBN 978-952-62-0690-5 (PDF)
ISSN 0355-3191 (Print)
ISSN 1796-220X (Online)

UNIVERSITY of OULU
OULUN YLIOPISTO