Norwegian University of Life Sciences
Faculty of Veterinary Medicine and Biosciences
Department of Plant Sciences

Master Thesis 2014
60 credits

Phenotypic and genotypic variation in the allopolyploid *Arabidopsis suecica* and its parent species *A. thaliana* and *A. arenosa*

An investigation of a possible model for studies of polyploidy and invasiveness

Torbjørn Horsberg Kornstad

Preface

The intersection between ecology and genetics is an extremely interesting research field, where many connections and mechanisms are yet to be revealed. If we were able to understand more of the genetic forces that drive adaptation and evolution, we would be better fit to meet ecological challenges like habitat destruction, invasive and alien species, climate change, and so on. Current research has come a long way, and only the last ten years molecular techniques have shown a rapid development, allowing a much better approach to the concept. At the same time, the more we get to know, the more we realize we do not understand. This thesis constitutes a very small piece of the complete picture, but hopefully this piece will fit in somewhere and give further information to the topic that future work can build upon.

Many people have helped me along the path towards completing this thesis. First of all I would like to thank my main supervisor, Associate Professor Siri Fjellheim at the Department of Plant Sciences. She has made a huge effort in supporting me, answering my questions and providing funds, she has joined in fieldwork and she has filled in for me in experimental work when I have needed it. I would also like to thank my co-supervisor Professor Mikael Ohlson at the Department of Ecology and Natural Resource Management for good advice; PhD research fellow Siri Lie Olsen at the Department of Ecology and Natural Resource Management for statistical advice and good writing tips; researcher Marte Holten Jørgensen at the Department of Plant Sciences for statistical advice; engineer Jørn Medlien at the Centre for Plant Research in Controlled Climate and engineers Øyvind Jørgensen, Sylvia Sagen Johnsen and Anne Guri Marøy at the Department of Plant Sciences for help with experiments and lab work; and research technician Leidulf Lund at the University of Tromsø for providing both advice and seeds of diploid *Arabidopsis thaliana*. Lastly I would like to thank my family for support through the educational course towards a Master of Sciences degree, and my friends for a great time at the Norwegian University of Life Sciences.

Ås, 9/5-2014

_____

Torbjørn H. Kornstad

Abstract

Species that possess more than two sets of chromosomes are denoted as polyploids. It is hypothesized that polyploids show high gene redundancy, hybrid vigour and masking of deleterious alleles, and that this make them better at adapting to novel environments because of wider phenotypic response range. It is also speculated that adaptive advantage of polyploidy contributes to invasiveness as there is a trend that polyploids are overrepresented within invasive species. The allopolyploid *Arabidopsis suecica* and its parent species *A. arenosa* and *A. thaliana* were chosen as a model system to investigate relationships between phenotypic plasticity, fitness and genetic variation. In this thesis I try to uncover genetic structures in the study species, and I investigate if *A. suecica* show higher plasticity and/or fitness than its parent species, if the different species show different levels of genetic diversity and whether *A. suecica* could work as a model for studying polyploidy and invasiveness.

Three to four wild Norwegian populations of each species were analyzed for phenotypic responses to differences in availability of nutrient and light, while population structure and genetic diversity was assessed through analysis of AFLP markers. The species were separated into genetic and phenotypic clusters with *A. suecica* being intermediate between its parent species. Clear population structure was inferred in *A. thaliana* and *A. arenosa*, while no structure was inferred in *A. suecica*.

The species exhibited similar phenotypic responses. *A. arenosa* seemed to have higher phenotypic plasticity and higher genetic diversity than the two other species, probably related to its outbreeding reproduction strategy. Furthermore, a general positive relationship between genetic diversity and phenotypic plasticity was found. Low genetic diversity and more population structure were found in the indigenous, inbreeding *A. thaliana*. Population spacing might explain the clear genetic structure in *A. arenosa*, while the lack of structure in *A. suecica* could be due to coherent populations. When it came to fitness measured as allocation of resources to reproduction, the trend pointed towards *A. arenosa* having lower fitness under poor environmental conditions. *A. suecica*, on the other hand, showed the ability to keep up fitness under different environmental conditions, which makes it a promising model for investigating invasiveness and polyploidy. Still, further studies are needed to confirm this.

Keywords: Polyploidy, invasive species, phenotypic plasticity, fitness, genetic variation

Samandrag

Artar som innehar meir enn eitt kromosomsett kallast polyploidar. Ein trur at polyploidar har høg grad av duplikerte gen i genomet, høg heterosis og maskerer skadelege allel, og at det gjer dei betre til å tilpasse seg til nye miljø fordi dei har eit vidare fenotypisk responsområde. Det er òg mogleg at adaptive fordelar ved polyploidi kan bidra til høgare invasibilitet, sia det er ein trend at polyploidar er overrepresentert blant invasive artar. Den allopolyploide arten *Arabidopsis suecica* og foreldreartane *A. arenosa* og *A. suecica* vart vald som eit modellsystem for å undersøke samanhengar mellom fenotypisk plastisitet, fitness og genetisk variasjon. I denne gradsoppgåva prøver eg å avdekke genetiske strukturar i studieartane, og eg undersøker om *A. suecica* har høgare plastisitet og/eller fitness enn foreldreartane, om dei ulike artane har ulike nivå av genetisk diversitet og korvidt *A. suecica* kan fungere som ein modell for studium av polyploidi og invasibilitet.

Tre til fire ville norske populasjonar av kvar art vart analysert for fenotypiske responsar til ulik tilgjengelegheit på næring og lys, medan populasjonsstruktur og genetisk diversitet vart undersøkt gjennom analyse av AFLP-markørar. Artane delte seg i genetiske og fenotypiske klyngar, og *A. suecica* plasserte seg mellom foreldreartane. Det vart finni ein klar populasjonsstruktur i *A. arenosa* og *A. thaliana*, men ikkje i *A. suecica*.

Artane viste liknande fenotypiske responsar. *A. arenosa* verka å ha høgare fenotypisk plastisitet og høgare genetisk diversitet enn dei to andre artane, truleg grunna ein utkryssande reproduksjonsstrategi. Vidare vart ein generell positiv samanheng mellom genetisk diversitet og fenotypisk plastisitet finni. Låg genetisk diversitet og meir populasjonsstruktur vart finni i *A. thaliana* som er stadeigen og innkryssande. Adskilte populasjonar kan kanskje forklara den klare genetiske strukturen i *A. arenosa*, medan manglande strukturar i *A. suecica* kan vera grunna samanhengande populasjonar. Når det kom til fitness målt som allokering av ressursar til reproduksjon, pekte trenden mot at *A. arenosa* kan ha lågare fitness under dårlege miljøtilhøve. *A. suecica* viste derimot evne til å halde oppe fitness under ulike miljøtilhøve, noko som gjer arten til ein lovande modell for å undersøke invasibilitet og polyploidi. Likevel trengs det ytterlegare forsking for å stadfeste dette.

Nøkkelord: Polyploidi, invasive artar, fenotypisk plastisitet, fitness, genetisk variasjon

Table of contents

1. Introduction

Polyploidization, i.e. mutations leading to organisms that possess more than two sets of chromosomes, is recognized as a driving force for adaptation and ecology (Lynch 2007; Sobel et al. 2010). Polyploidy can be observed in numerous taxonomic groups, but is thought to be especially frequent in angiosperms (Wendel 2000; Otto 2007; Song et al. 2012). In fact, it is often assumed that all angiosperms have undergone polyploidization at some point during their evolution (De Bodt et al. 2005; Soltis & Soltis 2009). There are two main ways of gaining polyploidy, namely *autopolyploidy* where the genome is duplicated within a species, and *allopolyploidy* where a new species is formed from hybridization between two parent species combined with whole genome duplication (Soltis & Soltis 2000). Successful allopolyploidization results in rapid speciation in an evolutionary context. The overall polyploidization rate is about $1/10^{th}$ of the overall speciation rate (Meyers & Levin 2006; Otto 2007), meaning that over a longer time span it will not constitute the most important speciation force. Also, polyploids themselves show reduced speciation rates, partly due to the fact that their possibilities of undergoing new polyploidizations are lower than in diploids (Mayrose et al. 2011; Arrigo & Barker 2012). However, in a world with large ecological changes within short time spans, it is reasonable to believe that speciation as a result of polyploidization could have ecological consequences, and these consequences should be investigated.

When a species is polyploid and possesses more than two sets of chromosomes, genetic forces act differently from what they do in diploids. A newly formed allopolyploid combines genes from two unrelated individuals, opening up for hybrid vigour and masking of deleterious alleles (te Beest et al. 2012). The combination of homeologous genes from two parent species often results in one of the genes being silenced, but it is proposed that subfunctionalization could work as a mechanism for retaining homeologous genes in the genome (Lynch & Force 2000; Hegarty & Hiscock 2008). A high gene redundancy due to the presence of homeologous loci suggests that allopolyploids could withstand inbreeding and population bottlenecks better than their diploid counterparts (Song et al. 2012; te Beest et al. 2012). Following this logic, allopolyploids could be better at adapting to new environments and sudden environmental changes, due to the underlying gene redundancy. The generation of new expressional patterns and novel epigenetic variation could also contribute to this (Comai 2005; Chen 2007). At the same time, there are genetic forces associated with polyploidization

that could be detrimental. The genome is notoriously unstable, and polyploidization is a process that changes the genome abruptly in just one generation. This can lead to problems in the mitosis and meiosis giving aneuploid cells, and problems with gene expression due to development of uneven relationships between genes and regulatory factors (Comai 2005). Epigenetic re-modelling could also cause instability in newly formed polyploids (Comai et al. 2003a). Though polyploidization is believed to imply both advantages and disadvantages, the view that polyploidization opens up for wider ecological and phenotypical variation and thus enable species to adapt quickly is widely accepted (Comai 2005; Otto 2007; Flagel & Wendel 2009; te Beest et al. 2012; but see Meyers & Levin 2006; Mayrose et al. 2011; Arrigo & Barker 2012).

Summed up, a theoretical framework for a possible positive relationship between polyploidy and abilities to adapt is established (Flagel & Wendel 2009). An important task now is to find out whether causal relationships exist, and eventually uncover how they work. With this as a background, it is highly interesting to carry out an experiment where an allopolyploid species is compared with its parent species with regard to performance under different environmental conditions. If the proposed ideas on the benefits of being polyploid hold true, the allopolyploid should show better performance and keep up fitness across a range of environmental conditions. A study system with only three species will not provide results that can be directly generalized to all allopolyploids, but it is a good way of building up a model that can be expanded in later experiments. Further on, it has been proposed that polyploids tend to have higher probability of being invasive than diploids (Lee 2002; Pandit et al. 2011; te Beest et al. 2012). In this perspective, a model system for comparing an allopolyploid species with its parent species could also act as a model system for understanding some of the underlying mechanisms that lead to a species becoming invasive.

The species complex chosen to assess these propositions consists of the allopolyploid species *Arabidopsis suecica* (Fr.) Norrl. ex O.E.Schulz and its two parent species, *A. thaliana* (L.) Heynh. and *A. arenosa* (L.) Lawalrée. Within the species complex, the model species *A. thaliana* is well investigated. Further on, the species are simple to grow and have a relatively short lifespan. This provides a good background for the thesis work. In Norway *A. thaliana* is regarded as an indigenous species, although it has the ability to behave like a weed (Elven 2005). The other two are regarded as alien species and classified in the risk category Potentially High risk (PH). This means that they show low or no impact on the Norwegian

nature per now, but it is believed that this could possibly change in the future (Gederaas et al. 2012). With this information as a background, it is investigated how the species complex could work as a model system for studying the genetic and phenotypic effects of allopolyploidization, also in an invasive species perspective. *A. suecica* is not viewed as a threat to Norwegian nature as of today, but it is here proposed that it could have the potential to work as a model species.

Koch and Matschinger (2007) call for genetic research on nonmodel species in the *Arabidopsis* genus. Further on, the chosen species complex has been proposed as a model system for studying what effects polyploidization has on the genome itself (Chen et al. 2004). It is assumed that higher genetic diversity constitutes a foundation for higher fitness (Reed & Frankham 2003). Thus, a study on allopolyploidy and fitness in the *Arabidopsis* genus should include genetic investigations. While population structure and genetic diversity is well investigated in *A. thaliana* (e.g. Beck et al. 2008; Lewandowska-Sabat et al. 2010), there is still a long way to go when it comes to *A. suecica* and *A. arenosa*. A study conducted by Lind-Hallden et al. (2002) compared genetic diversity in the three species, but otherwise little knowlegde is available. This thesis aims to contribute in filling the knowledge gaps by assessing and comparing population structure and genetic diversity between the three species based on genotyping from Amplified Fragment Length Polymorphism (AFLP) markers.

Two terms are assessed specifically in the thesis: *Phenotypic plasticity* and *fitness homeostasis*. The first is the ability to exhibit a wide range of phenotypes across varying environmental conditions (Bradshaw 1965; Schlichting 1986), the second is the ability to keep fitness as equal as possible between varying environmental conditions (Richards et al. 2006; Hulme 2008). It is proposed that high phenotypic plasticity provides wider possibilities to adapt to new environments (Sultan 2000; Davidson et al. 2011), while high fitness homeostasis could imply better abilities at coping with and adapting to stressful environments (Richards et al. 2006; Hulme 2008). The terms are assessed through analysis of phenotypic variation as response to environmental conditions exhibiting different levels of stress. An attempt is done to compare the two terms in light of the results.

Based on the theoretical framework, the following research questions were formulated:

1. Do the allopolyploid *A. suecica* show higher phenotypic plasticity and/or fitness homeostasis than its parent species, and does this reflect a higher ability to adapt to different environments?
2. Do the study species show different levels of genetic diversity, and if so, is this related to phenotypic plasticity?
3. Is *A. suecica* suitable as a model species for studying relationships between polyploidy and invasiveness, even though the species currently does not behave in an invasive way?

2. Materials and methods

2.1. Study area

Seeds from 10 wild populations of *A. thaliana*, *A. suecica* and *A. arenosa* were sampled. The number of sampled populations per species was three *A. thaliana* populations, three *A. suecica* populations and four *A. arenosa* populations. The seeds were sampled from three different geographic areas within Southern Norway, namely Drammen, Eidskog and Gudbrandsdal, so that seeds from at least one population of each species were sampled from each geographical area (Table 1, Fig. 1).

Table 1: List of populations where seeds were sampled, specifying locality codes, locality names, what geographical areas the different localities belong to, species, collection date, latitude in degrees north (Lat (°N)) and longitude in degrees east (Long (°E)).

Code	Locality name	Geographical area	Species	Collection date	Lat (°N)	Long (°E)
T-EID1	Bakkeberget	Eidskog	*A. thaliana*	11.06.2012	60.111	12.123
S-EID3	Åbogen stasjon	Eidskog	*A. suecica*	11.06.2012	60.109	12.116
A-EID4	Pramhus	Eidskog	*A. arenosa*	11.06.2012	60.090	12.149
A-DRA1	Berskog	Drammen	*A. arenosa*	21.06.2012	59.755	10.120
S-DRA2	Drammen stasjon	Drammen	*A. suecica*	17.06.2012	59.741	10.202
T-DRA3	Åslyveien	Drammen	*A. thaliana*	21.06.2012	59.756	10.154
T-SFRO3	Kjorstad	Gudbrandsdal	*A. thaliana*	05.07.2012	61.579	9.894
S-NFRO3	Kvam stasjon	Gudbrandsdal	*A. suecica*	05.07.2012	61.665	9.702
A-NFRO4	Nymoen	Gudbrandsdal	*A. arenosa*	05.07.2012	61.663	9.676
A-GAU1	Steinslia	Gudbrandsdal	*A. arenosa*	07.07.2012	61.220	10.228

For each population, 20 randomly chosen individuals were sampled. If a population consisted of less than 20 individuals, as many individuals as possible were sampled. The plants were dried, and the seeds extracted and transferred to 2 mL tubes (Eppendorf, Hamburg, Germany).

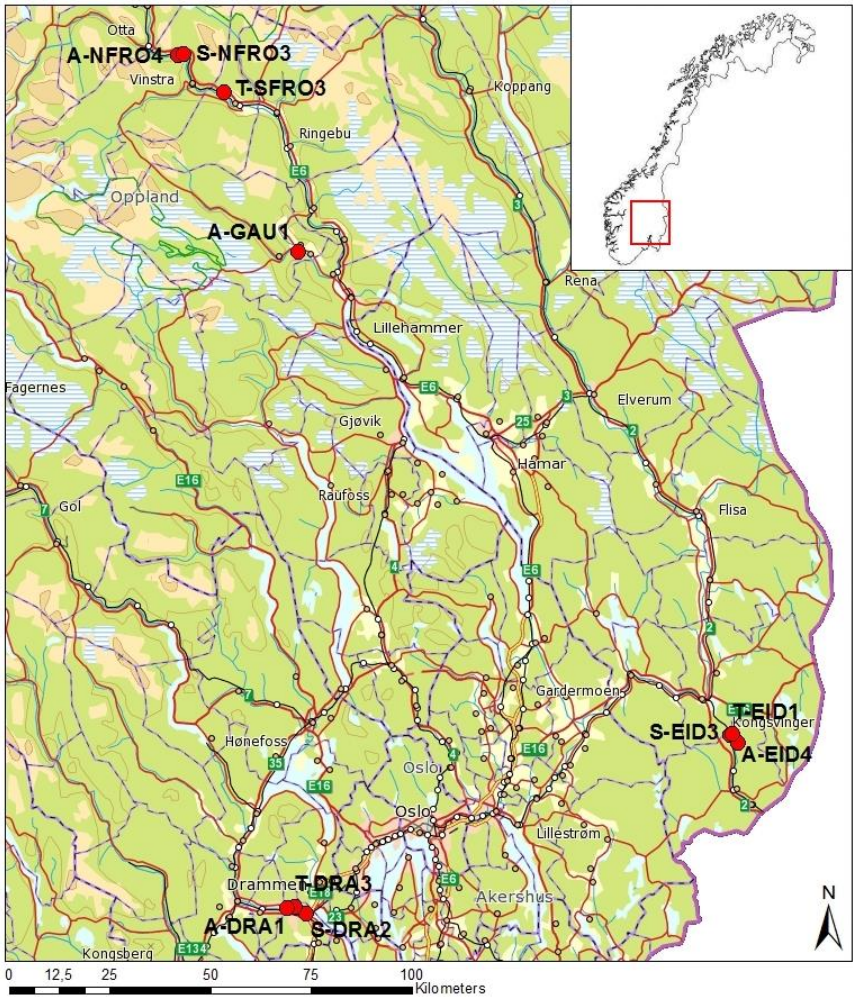


Fig. 1: Map showing localities of populations where seeds were sampled for the experiment.

2.2. Study species

The species collected all belong to the genus *Arabidopsis*, and they form a hybrid complex. *A. suecica* originates from an allypolyploid hybridization between the mostly diploid *A. thaliana* and the mostly autotetraploid *A. arenosa* (O'Kane et al. 1996; Jakobsson et al. 2006), possibly within the eastern parts of *A. thaliana*'s native range (Beck et al. 2008). The formation of the species probably happened through the fertilization of a female, unreduced *A. thaliana* gamete with a normal, male *A. arenosa* gamete (Säll et al. 2003). It is believed to have risen in a single event between 12 000 and 300 000 years ago, somewhere south of its present native distribution in Sweden and Finland (Säll et al. 2003; Jakobsson et al. 2006). It has been shown that out of *A. suecica*'s 26 chromosomal pairs, 16 derive from *A. arenosa* and 10 derive from *A. thaliana* (Comai et al. 2003b). *A. suecica* exhibits bivalent, homologous pairing of its chromosomes in the meiosis (Comai et al. 2003b; Pecinka et al. 2011). Studies indicate that *A.*

suecica expresses more homeologous genes descending from *A. arenosa* than from *A. thaliana* (Chang et al. 2010).

The study species can all be found in the wild in Norway, but *A. thaliana* is the only indigenous species (Elven 2005). *A. thaliana* is distributed in inland valleys and along the coast up to Lofoten and Vesterålen, while both *A. arenosa* and *A. suecica* have easterly distributions. *A. arenosa* is found all the way up to the Russian border (Elven 2005).

All three species are winter annuals, forming an overwintering basal rosette of leaves in the autumn and a flowering stem in the following spring (Baskin & Baskin 1983). *A. thaliana* has small, off-white, inconspicuous flowers, rosette leaves almost without serration, and is quite slender, while *A. arenosa* has bigger, white flowers, heavily serrated leaves and is coarser. *A. suecica* is morphologically intermediate between the two parent species. While *A. arenosa* is a strictly outcrossing species, *A. thaliana* and *A. suecica* are self-fertilizing species that set seeds regardless of whether they are pollinated or not (Säll et al. 2004).

When it comes to habitat preference, all three species are found mostly on dry, sandy soil. *A. thaliana* often grows in rock crevices and on ledges, while the other two mostly are found on road verges and railways (Elven 2005). The last applies especially to *A. suecica*, seeing as all populations that were visited when sampling seeds for this experiment were growing along railways.

2.3. Analysis of phenotypic responses

2.3.1. Experimental design

Seeds from the sampled populations were grown under controlled environmental conditions in a growth chamber. In order to assess whether different species reacts differently to varying environmental conditions, eight different treatments were applied in a 2^3 factorial design. These treatments consisted of all different combinations of wet and dry water conditions, rich and poor nutrient conditions and high and low light conditions. Water as a treatment was not regarded in the statistical analysis as the effects turned out to be small and insignificant, and is not further described in this chapter. Thus, the experiment was reduced to a 2^2 factorial design. The final four treatment levels were as follows: Low light, poor nutrients (LP); low light, rich nutrients (LR); high light, poor nutrients (HP); and high light, rich nutrients (HR).

A number of 8 trolleys with a size of 100 times 60 cm were covered first with plastic and then with felt mats having the ability to transport the water evenly over the whole trolley. 50 circular 8C-101 flowerpots with a diameter of 8 cm (Billund Potter, Billund, Denmark) were placed on each trolley. 400 flowerpots were prepared overall. Each flowerpot was filled with Gartnerjord soil (Tjerbo Torvfabrikk, Rakkestad, Norway) consisting of 86 % Sphagnum peat, 10 % sand and 4 % granule clay. Two trolleys were assigned to each treatment level (Appendix 1). The number of replicates per population was 10 for each treatment level. For each population, seeds from all sampled individuals were mixed on a white paper sheet, then several seeds were drawn randomly and sown in each pot. The different populations were distributed randomly within each trolley. 9 L of water were applied to each trolley after sowing.

The seeds were stratified for four days in 4°C and 24 hours darkness. Then, conditions were changed to 20°C/17°C day/night temperature and an 8/16 hours light/dark cycle. Light was provided by OSRAM 400W Powerstar HQ1®-BT 400W/d Pro Daylight E40 (OSRAM Licht AG, Munich, Germany) light bulbs in GAVITA GAN 400 AL lamps (GAVITA AS, Andebu, Norway). The amount of light in the chamber was measured to 210-250 µmol m^{-2} s^{-1} with a LI-189 quantum/radiometer/photometer (LI-COR Biosciences, Lincoln, Nebraska, USA). The seeds were allowed to germinate at similar conditions for all trolleys, and water was applied regularly in order to avoid desiccation. Almost all seeds from the T-DRA3 population failed to germinate, and the population was excluded from the experiment. The T-DRA3 seeds that did germinate were allowed to grow. A total number of 360 plants distributed on the remaining populations were grown for the experiment. Among these, four died during the experiment and were not included in the analyses.

When the seedlings had reached the stadium where primary leaves started to become visible, they were thinned so that one plant remained in each flowerpot. For some populations, transplantations between pots were done. The plants were allowed one week of optimal growth conditions before treatments were applied. Nutrient treatment was applied by giving nutrient solution made from 1.25 mL Superba NPK 14-4-21 + mikro (Nordic Garden AS, Stokke, Norway) and 1 L water to each of the rich nutrient trolleys once per week, while no nutrients were applied to the poor nutrient trolleys. Light treatment was applied by covering the low light treatment trolleys with XLS 17 Revolux light-reducing fabric (AB Ludvig

Svensson, Kinna, Sweden). The fabric is partly made from aluminium, and it does not change the spectral composition of the light that passes through. The amount of light below the fabric was measured to be 80-90 µmol m^{-2} s^{-1}, equivalent to a reduction of 60-70 %. The light-shading fabric was found to heavily reduce evaporation from the low-light trolleys, so these trolleys were watered more rarely in order to obtain similarity in water conditions between the low-light and the high-light trolleys.

Vernalization was initiated 39 days after sowing (35 days after germination conditions were initiated). Growth conditions were changed to 4°C constant temperature and an 8/16 hours light/dark cycle. Since growth was low during vernalization, nutrients were applied on average every third week, in the same doses as described above. The amount of light in the growth chamber was reduced in order to avoid the plants dying from light stress (J. Medlien, pers. comm.). The amount of light was measured to be 125-135 µmol m^{-2} s^{-1} and 27-32 µmol m^{-2} s^{-1} for the high light and low light treatments, respectively.

Based on findings in Lewandowska-Sabat et al. (2012), vernalization conditions were kept for 9 weeks. At the end of vernalization, 102 days after sowing, growth conditions were changed to 23°C/20°C day/night temperature, 16/8 hours light/dark cycle and full amount of light to allow flowering. Nutrients were given once a week in the same dose as described above. These conditions were kept for 33 days, when the growth experiment was ended. During the whole experiment, the trolleys were moved around within the chamber and pots were moved around on the trolleys periodically to avoid edge effects.

2.3.2. Measurements of phenotypic variables

Phenotypic variables were measured at different times. At the initiation of vernalization, three different variables were measured: Number of rosette leaves per plant, length of the longest rosette leaf for each plant (including both petiole and lamina) and length of the lamina on the longest rosette leaf. In cases where leaves were serrated, the length from the innermost serration to the leaf tip was measured and recorded as lamina length.

At the end of the vernalization period, the days it took for each individual plant to bolt and to open the first flower were counted with the last day of vernalization set as day zero. In addition, the number of rosette leaves was measured at bolting. For plants that bolted, but did

not flower, the time to flowering was set to the ending day of the experiment plus five days. For plants that neither did flower nor bolt, rosette leaves were counted at the ending day of the experiment, time to bolting was set to the ending day plus five days and time to flowering was set to the ending day plus ten days.

At the end of the experiment, five different variables were measured for each plant: Plant height measured as the longest stem from root to tip for each plant, number of branches on the stem, total number of flowers and siliques (denoted as "number of flowers" – buds were not counted) and dry weight of the aboveground biomass. To measure the dry weight, the plants were harvested and dried at 60°C for 24 hours in a TS8136 drying oven (Termaks, Bergen, Norway) before weighing them with AG ED224S scales (Sartorius AG, Groettingen, Germany).

2.4. Measurements of ploidy level and chromosomal numbers

In order to ensure that all populations of the study species had the expected chromosomal numbers and ploidy levels, DNA content was measured with flow cytometry. Seeds from each of the populations grown in the experiment were sown in pots and grown to a certain size. For each population, three individuals were selected for harvesting. Leaves corresponding to a total area of 1-2 cm^2 were harvested. Flow cytometry was performed and DNA ratios were obtained by G. Geenen, Plant Cytometry Services (Schjindel, The Netherlands). Diploid *A. thaliana* from the "Columbia" line was aquired from the University of Tromsø, and provided as a control sample along with the experimental samples. For internal control *Ilex crenata* 'Fastigiata' was used.

2.5. Analysis of genetic markers

For genetic analyses, seeds from each plant harvested during the fieldwork were sown in individual pots for the populations A-GAU1, A-NFRO4, A-DRA1, S-DRA2, S-NFRO3, S-DRA2, S-EID3, T-SFRO3 and T-EID1. For T-DRA3, seeds harvested from the plants grown in the growth chamber experiment were sown. A-EID4 was not available for analysis, since there were very few viable seeds left. The plants were grown until they were eligible for harvesting. One individual was harvested from each pot. During the harvest, ~100 µg of fresh tissue per plant was cut in pieces with scissors and put into 2 mL tubes (Eppendorf). The

tubes were stored at -80°C. Before isolation of DNA, two 3 mm crushing beads were applied to each tube. The tubes were dipped into liquid nitrogen, before the tissue was crushed with a TissueLyser II (QIAGEN, Hilden, Germany) for 1 minute at 20 r/s. DNA was extracted from the crushed tissue using a DNeasy Plant Mini Kit (QIAGEN). The quantity of DNA in each isolation was checked using a NanoDrop 8000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). A modified AFLP protocol after Hayashi et al. (2005) and Vos et al. (1995) was run on the genomic DNA. One E+1 primer (E01) and one M+1 primer (M01) was used for preamplification. Six different combinations of E+3 and M+3 primers were tested for selective amplification (E33xM36, E33xM37, E33xM38, E42xM36, E42xM37, E42xM38). The three underlined combinations yielded the best testing results, and were used for further runnings (see Appendix 2 for details on AFLP protocol and primers).

The AFLP results were scored using GeneMapper® ver. 5.0 (Life Technologies, Carlsbad, CA, USA). As a general rule, peaks had to have a height of at least 2000 in order to be scored as a band. Manual corrections were run on all samples. Individuals showing anomalous peak patterns on at least one of the three primer combinations were removed completely from the dataset. The number of replicates was 31 (22.7 %) for primer combination E33xM37, 30 (22.1 %) for primer combination E33xM38 and 23 (16.9 %) for primer combination E42xM38. The genotyping error for each primer combination was calculated using the formula (total number of scoring errors)*100/(number of replicates)*(number of markers) (Bonin et al. 2004), then a final genotyping error was calculated by computing a weighted mean between the primer combinations. Alleles showing a high level of inconsistency were removed before calculating the genotyping error, and not included in the analyses. All individuals were assessed for number of bands within each primer combination, and individuals showing an extraordinary high or low number of bands within at least one primer combination were removed from further analysis. The numbers of assessed individuals per population were 6 individuals from A-DRA1, 18 individuals from A-GAU1, 17 individuals from A-NFRO4, 16 individuals from S-DRA2, 16 individuals from S-EID3, 20 individuals from S-NFRO3, 6 individuals from T-DRA3, 18 individuals from T-EID1 and 19 individuals from T-SFRO3.

2.6. Data analysis

All data analyses were done with RStudio version 0.97.551 (RStudio 2013), based on R version 3.0.1 (R Core Team 2013), unless anything else is specified in the text.

2.6.1. Phenotypic responses

Descriptive multivariate analysis using non-metric multidimensional scaling (NMDS) from the R package vegan (Oksanen et al. 2007) was run on all measured response variables, in order to obtain a crude picture of how the different species reacted to the different combinations of treatment. A non-metric approach was chosen since several of the response variables were non-linear and/or non-normal. Some of the variables were discarded from further analysis for different reasons: Days to bolting (closely correlated with days to flowering, $r = 0.84$), number of leaves at start of vernalization (closely correlated with number of leaves at bolting, $r = 0.90$), number of branches (zero inflated and thus hard to analyze properly), and length of lamina on longest leaf and its percentage of total leaf length (irrelevant variables in an ecological perspective). The remaining variables were superimposed onto a biplot of the first two NMDS axes.

To assess the effect of treatment and species on the different variables, linear mixed effects models or generalized linear mixed effects models were run. Table 2 gives an overview of transformation of variables and what type of models that were run for each response variable. For the final models a single factor was constructed, where each level corresponded to a specific combination of light, nutrients and species for a total of 12 levels. Population was added as a random effect. Linear mixed effects models were fit using restricted maximum likelihood, generalized linear models with poisson family were fit using maximum likelihood, and generalized linear models with quasipoisson family were fit using penalized quasi-likelihood. All models were checked for assumptions of normality and equality of variance between groups by conferring Q-Q and residual plots.

Table 2: Overview of type of models run for the phenotypic response variables, including eventual transformation or GLMM family. LMM = Linear mixed model, GLMM = Generalized Linear Mixed Model.

Response	Type of model	Transformation	GLMM family
Biomass	LMM	Natural logarithm	-
Days to flowering	GLMM	-	Poisson
Height	LMM	-	-
Number of leaves at bolting	GLMM	-	Poisson
Number of flowers	GLMM	-	Quasipoisson
Longest leaf at start of vernalization	LMM	-	-

Post-hoc testing of the models was done by applying general linear hypothesis methods from the R package multcomp (Hothorn et al. 2008). These methods give a generalization of the Tukey post-hoc test that can be used on unbalanced designs. To model reaction norms for each species to the applied treatments, common letter displays based on multiple comparisons between all pairs were constructed.

To check whether the species reacted differently to environmental stress and showed differences in phenotypic plasticity, 95 % confidence intervals for estimated differences in phenotypic responses between high and low levels of treatments were constructed. Effects of light were assessed separately within rich and poor level of nutrients, and effects of nutrients were assessed separately within high and low level of light. This gave a total of four assessed treatment combinations for each species. The constructed confidence intervals were compared between species in order to investigate whether the difference in phenotypical responses would vary from one species to another. Since no corrections for multiple comparisons were done, the confidence intervals were interpreted with care.

Coefficients of variation were calculated for each response variable. The measurement gives an indication on the amount of phenotypic plasticity (Schlichting & Levin 1984; Sultan 2001). Variables were not transformed for this calculation. The formula used for calculation was $100 * sd(\bar{X}_i)/mean(\bar{X}_i)$, where i denotes the different treatment levels. This was done both on the population and on the species level. Estimates of uncertainty were unavailable, meaning that it was not possible to evaluate whether significant differences could be found.

To assess fitness homeostasis in the different species, a comparison variable called *C* comparing experimental variables connected to fitness by Davidson et al. (2011) with other experimental variables connected to phenotypic plasticity was constructed. Variables connected to fitness included number of flowers and total biomass, while variables not connected to fitness included height of plants, number of leaves at the end of the experiment and the length of the longest leaf at the start of vernalization. Some of the variables were transformed to make them more linear: Biomass (natural logarithm), number of flowers (natural logarithm of (number of flowers + 1)) and number of leaves at the end of the experiment (natural logarithm). To make the variables comparable, they were standardized to occupy an interval between 0 and 1. This was achieved by 1) adding/subtracting the lowest number in the variable to all observations in the variable so that the lowest number in the variable would be 0 and 2) dividing all observations by the highest number in the variable. From the transformed and scaled variables the formula $C = \frac{Flowers+Biomass}{Nleaves+Leaf\ length+Height}$ was used to construct the comparison variable.

A linear mixed model and general linear hypothesis post-hoc methods as described above were applied to the comparison model. The theory is that a higher value of *C* means relatively more allocation of resources to fitness, and vice versa. A smaller difference in *C* between good and poor environmental conditions can be interpreted as a higher degree of fitness homeostasis.

2.6.2. Population structure and genetic diversity

The dataset was examined for population structure using the software Structure, a software that can allocate individuals to genetic groups based on AFLP data (Pritchard et al. 2000). Analyses were run using Structure ver. 2.3.4 at the Lifeportal, University of Oslo (http://lifeportal.uio.no), with 10^6 iterations and a burn-in of 10^5 iterations. An admixture model was used; meaning that for each individual different parts of the genome is allowed to descend from different groups. Linkage between markers was not considered. A minimum of one population (K = 1) and a maximum of 9 populations (K = 9) was allowed per analysis. For each value of K, 10 independent runs were done. The results were assessed using the R functions in Structure-sum (Ehrich 2011). The number of clusters was chosen after an evaluation based on the following criteria: 1) all runs gave similar results, 2) similarity coefficient close to 1.0, 3) highest possible ln P (data) and 4) highest possible ΔK (Pritchard

et al. 2000; Evanno et al. 2005). Structure analysis was run for each species. In addition, an analysis incorporating all individuals was run in order to see whether the different species clustered separately.

To visualize the clusters in a multidimensional space, Principal Coordinate Analysis (PCO) was run on a distance matrix calculated with Dice's coefficient of similarity (Dice 1945). The PCO analyses were run in PAST ver. 2.17c (Hammer et al. 2001), and scores for the two first components were extracted and plotted in R. PCO analysis was run for all species together, and separately for each species.

To assess and compare the diversity of the sampled populations and species, 95 % confidence intervals for Nei's Genetic Diversity (Nei 1987) was constructed using bootstrapping over 1000 replicates with the R functions in AFLPdat (Ehrich 2006). Analyses of molecular variance (AMOVA) (Excoffier et al. 1992) were performed in Arlequin ver. 3.5 (Excoffier & Lischer 2010). This was done for each species based on groups inferred from the original populations. If the number of clusters inferred from Structure came out differently from the original populations, an additional AMOVA was run based on the inferred clusters (unless the inferred number of clusters was one).

2.6.3. Comparison of genetic diversity and phenotypic plasticity

To assess whether there is a positive relationship between genetic diversity and phenotypical plasticity, a Mantel test was run to compare euclidean distance matrices calculated from 1) Nei's Genetic Diversity and 2) Coefficients of variation for all phenotypic variables. The test was run on the eight populations where results from both growth experiments and genetic analyses were available. A corresponding test was also done with a phenotypic distance matrix calculated from coefficients of variation where each variable was scaled to unity. The scaling was done by dividing all values in the variables by the highest value in the variable.

3. Results

3.1. Growth experiment

3.1.1. Multivariate analysis of phenotypic responses

The ordination in Fig. 2 shows that all species clustered loosely. *A. thaliana* was separated from the other two species. There was a trend that *A. suecica* occupied the space between *A. arenosa* and *A. thaliana*. Within each treatment, the clustering of species was clearer. There was not a clear clustering between treatments, although rich nutrients and high light tended to cluster on the top left side of the plot. This indicates that rich nutrients and high light were associated with taller plants, higher biomass and more flowers. For all variables included in the NMDS, R^2 were > 0.50 and p-values were < 0.001.

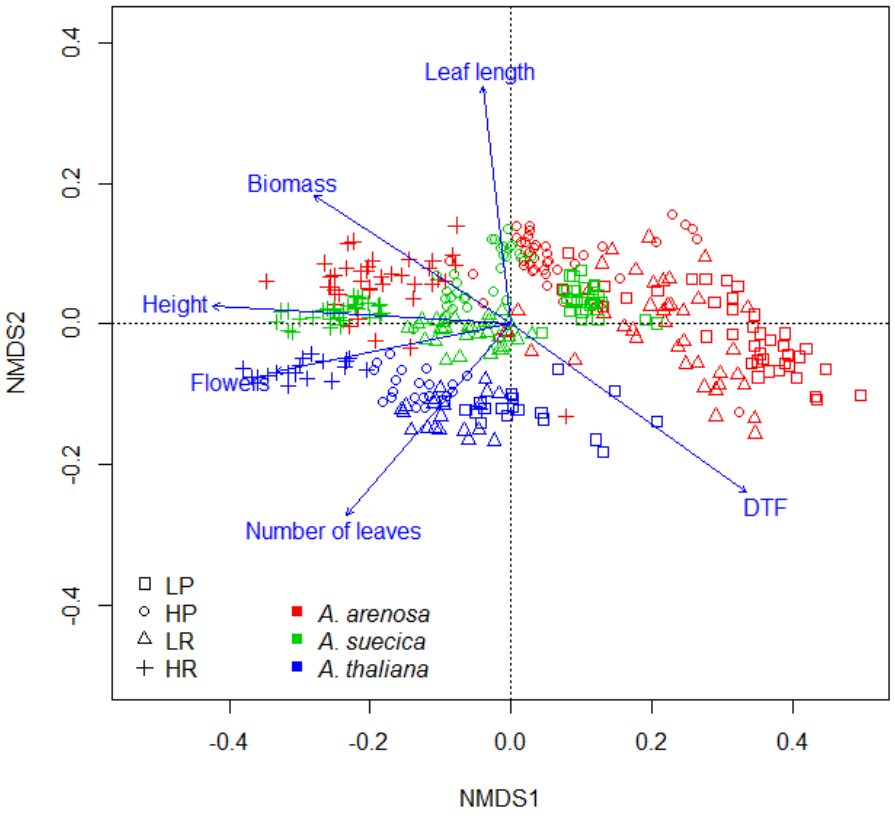


Fig. 2: Biplot of the two first NMDS axes, showing all observations grouped after species and treatment. LP = low light, poor nutrients, HP = high light, poor nutrients, LR = low light, rich nutrients, HR = high light, rich nutrients. The arrows show the phenotypic response variables and what trends they exhibited.

3.1.2. Analyses of phenotypical responses

The general trend was an increase in the phenotypic response variables from low light, poor nutrients via low light, rich nutrients / high light, poor nutrients to high light, rich nutrients (Fig. 3). The exception from this was days to flowering (DTF), where the trend was the opposite. This is expected, since plants are anticipated to flower faster when conditions are better. In that sense, a negative trend in DTF should be treated equally to a positive trend in the other variables.

Reaction norms differed between species in some of the response variables (Fig. 3). Biomass reactions were very similar between all three species, and no significant differences in absolute values between species were found on any treatment levels (Fig. 3a). There was a tendency that *A. arenosa* produced fewer flowers than the other two species, but this was significant only for the low light, rich nutrients treatment (Fig. 3b). Under low light conditions, *A. arenosa* used longer time to flower and the plants were shorter than in the two other species (Fig. 3c and e). *A. thaliana* plants had more leaves than the other species at the time of bolting across all treatments (Fig. 3d). *A. suecica* placed between the parent species when it came to number of flowers, number of leaves at bolting and partially in longest leaf at start of vernalization.

The differences in phenotypic responses between high and low level treatments were in most cases different from 0 (Fig. 4), meaning that the applied treatments provoked responses in the measured variables in general. Nutrient reduction did not seem to have any effect on DTF, and insignificant nutrient responses were also found in number of flowers, number of leaves at bolting and longest leaf at vernalization. Still the general picture is that all species showed phenotypic plasticity as a response to the applied treatments.

In general, the differences in phenotypic responses were quite similar between the species (Fig. 4). *A. arenosa* showed higher reduction in number of flowers and plant height as response to light reduction under rich nutrient conditions (Fig. 4b and e). There was also a trend that *A. arenosa* showed less reduction in number of leaves at bolting as response to light reduction under poor nutrient conditions, but this trend was less clear (Fig. 4d). For *A. thaliana* and *A. suecica*, no treatment responses were observed that were significantly different from all other species.

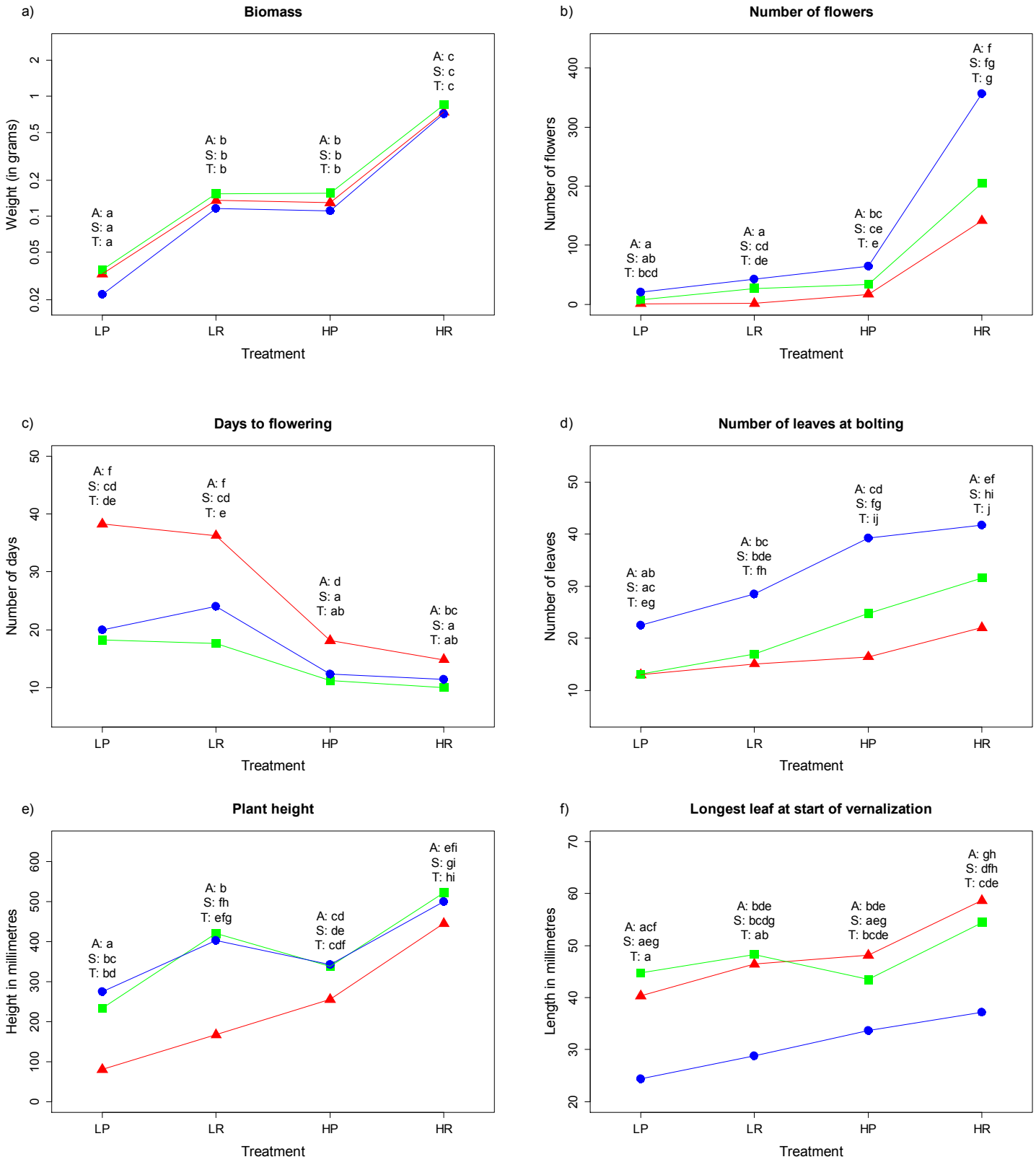


Fig. 3: Reaction norms for the phenotypic response variables within the different species. Common letters denote no significant difference. A (red) = *A. arenosa*, S (green) = *A. suecica*, T (blue) = *A. thaliana*. LP = low light, poor nutrients, LR = low light, rich nutrients, HP = high light, poor nutrients, HR = high light, rich nutrients. a) Biomass (dry weight in g), b) Number of flowers, c) Days to flowering after vernalization, d) Number of leaves at bolting, e) Plant height at harvest (in mm), f) Longest leaf at start of vernalization (in mm).

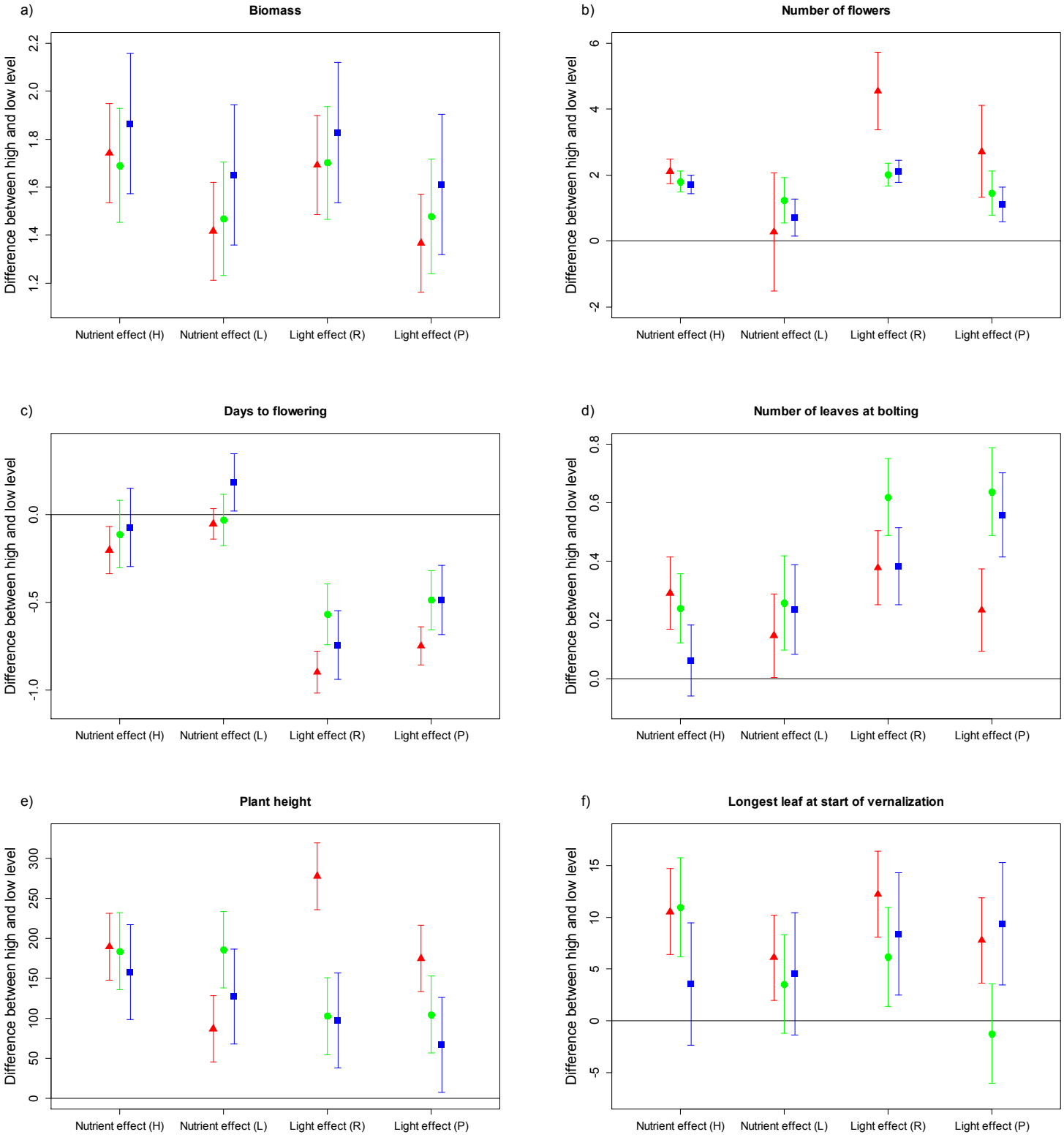


Fig. 4: 95 % confidence intervals for estimated differences in phenotypic responses between high and low levels of treatments. The symbols show the estimated differences, while the error bars show the confidence intervals. Vertical axes denote differences in the models that were run, and thus do show not real values for all variables. Nutrient (H): effect of nutrients within high level of light, Nutrient (L): effect of nutrients within low level of light, Light (R): effect of light within high level of nutrients, Light (P): effect of light within low level of nutrients. Red triangle = *A. arenosa*, green circle = *A. suecica*, blue square = *A. thaliana*. a) Biomass (dry weight in g), b) Number of flowers, c) Days to flowering after vernalization, d) Number of leaves at bolting, e) Plant height at harvest (in mm), f) Longest leaf at start of vernalization (in mm).

The coefficients of variation for the measured response variables are shown in Table 3. *A. arenosa* seemed to exhibit higher phenotypic plasticity when it comes to height, number of flowers and DTF, while *A. suecica* differed from the other species when it comes to leaves at bolting. The variation was large on the population level, but the general trends from the species level were reflected in the populations.

Table 3: Estimated coefficients of variation for the phenotypic response variables, measured across species and populations.

Species	**Biomass**	**Flowers**	**DTF**	**Leaves at bolting**	**Height**	**Longest leaf**
A. arenosa	122.16	168.47	45.12	22.89	65.86	15.74
A. suecica	124.65	133.99	29.93	37.99	32.22	10.19
A. thaliana	130.13	129.96	35.96	27.43	25.11	17.95
Population						
A-DRA1	122.13	171.76	43.52	29.34	53.28	11.71
A-EID4	138.97	171.08	51.41	26.22	77.34	24.01
A-GAU1	117.14	162.21	42.44	19.35	71.33	15.86
A-NFRO4	116.24	163.79	43.47	20.10	71.70	18.31
S-DRA2	125.80	125.15	23.06	33.20	29.26	10.67
S-EID3	126.39	138.98	31.01	41.02	33.38	13.14
S-NFRO3	121.44	140.28	35.21	40.02	36.34	13.37
T-EID1	141.49	149.43	40.82	31.91	33.49	15.20
T-SFRO3	122.46	113.32	31.00	25.43	17.03	22.06

In the analysis of the comparison variable *C*, we see that the species had similar values within the high light treatment (Fig. 5). Within the low light treatment, *A. arenosa* seemed to exhibit lower values than the other two species, although this trend was not significant. However, in a model without population as a random effect *A. arenosa* came out as significantly lower than the other two species within the low light, poor nutrients treatment. This could indicate that *A. arenosa* allocates fewer resources to keep up fitness under low light treatments than the other two species, but these results should be interpreted with caution. Responses in the *C* variable between high and low treatment levels were not significantly different between species within any of the treatments (data not shown).

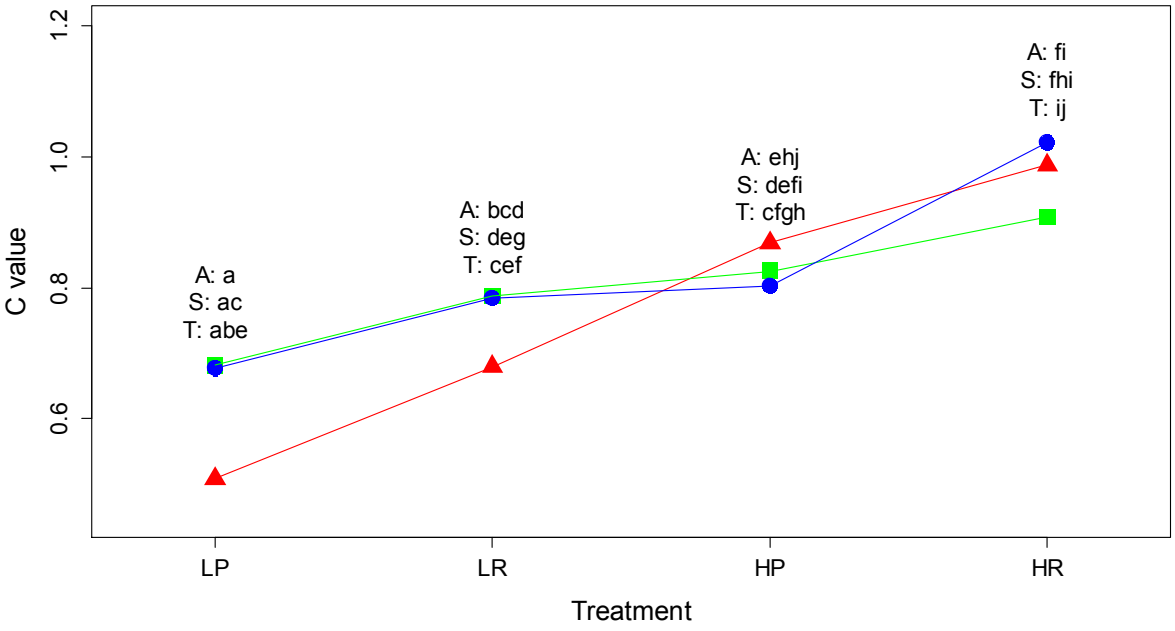


Fig. 5: Reaction norm for the comparison variable *C* within the different species. Common letters denote no significant difference. A (red) = *A. arenosa*, S (green) = *A. suecica*, T (blue) = *A. thaliana*. LP = low light, poor nutrients, LR = low light, rich nutrients, HP = high light, poor nutrients, HR = high light, rich nutrients.

3.2. Measurements of ploidy level and chromosomal numbers

Results from flow cytometry are shown in Table 4. The populations used in the experiment mainly showed the expected chromosomal numbers and ploidy levels: 10 chromosomes/diploid for *A. thaliana*, 32 chromosomes/tetraploid for *A. arenosa* and 26 chromosomes/tetraploid for *A. suecica*. There were two exceptions: Observation number 10 showed a lower chromosomal number than expected in *A. arenosa*. This might be due to aneuploidy, but it might also be due to errors in the measurement. Observation number 30 showed a measurement in *A. thaliana* that one would expect for *A. arenosa*. This is probably due to a confusion of samples.

Table 4: Results from flow cytometry, showing observation numbers, species, populations, sample numbers, measured DNA ratio, chromosomal number and ploidy level.

#	Species	Population	Sample	DNA ratio	Chromosomal number and ploidy level
0	*A. thaliana*	Control		0.16	2n = 2x = 10
1	*A. arenosa*	A-EID4	1	0.39	2n = 4x = 32
2			2	0.40	2n = 4x = 32
3			3	0.40	2n = 4x = 32
4		A-DRA1	1	0.40	2n = 4x = 32
5			2	0.40	2n = 4x = 32
6			3	0.40	2n = 4x = 32
7		A-GAU1	1	0.40	2n = 4x = 32
8			2	0.42	2n = 4x = 32
9			3	0.41	2n = 4x = 32
10		A-NFRO4	1	0.38	2n = 4x = 30
11			2	0.40	2n = 4x = 32
12			3	0.40	2n = 4x = 32
13	*A. suecica*	S-EID3	1	0.37	2n = 4x = 26
14			2	0.36	2n = 4x = 26
15			3	0.36	2n = 4x = 26
16		S-DRA2	1	0.36	2n = 4x = 26
17			2	0.36	2n = 4x = 26
18			3	0.36	2n = 4x = 26
19		S-NFRO3	1	0.36	2n = 4x = 26
20			2	0.36	2n = 4x = 26
21			3	0.36	2n = 4x = 26
22	*A. thaliana*	T-EID1	1	0.16	2n = 2x = 10
23			2	0.16	2n = 2x = 10
24			3	0.16	2n = 2x = 10
25		T-DRA3	1	0.16	2n = 2x = 10
26			2	0.16	2n = 2x = 10
27			3	0.16	2n = 2x = 10
28		T-SFRO3	1	0.16	2n = 2x = 10
29			2	0.16	2n = 2x = 10
30			3	0.41	2n = 4x = 32

3.3. Analyses of population structure and genetic diversity

A total number of 136 individuals were analyzed for variation in 274 AFLP markers (100 E33xM37 markers, 97 E33xM38 markers and 77 E42xM38 markers). 63 markers were only present in *A. arenosa*, 27 were only present in *A. suecica*, 16 were only present in *A. thaliana*, 67 were present in *A. arenosa* and *A. suecica*, 45 were present in *A. suecica* and *A. thaliana*, 12 were present in *A. arenosa* and *A. thaliana* and 44 were present in all three species. The percentage of polymorphic markers was 95.2 % in *A. arenosa*, 82.5 % in *A. suecica* and 77.8 % in *A. thaliana*. The genotyping error was calculated to be 3.30 %.

3.3.1. Population structure

The results from Structure showed a clear clustering of the different species (Fig. 6a). This was confirmed by the PCO (Fig. 7a), where we also see that *A. suecica* was placed in the middle of the first axis between its parent species. On the population level, *A. arenosa* showed a clear population clustering both in Structure and PCO (Fig. 6b, Fig. 7b). In *A. suecica* no clear population structure was found (Fig. 6c), but the PCO indicated a clustering of the different populations (Fig. 7c). In *A. thaliana*, Structure identified one cluster consisting of T-SFRO3 and one cluster consisting of T-DRA3 and T-EID1 (Fig. 6d). One individual in T-DRA3 clustered with T-SFRO3, and this was reflected in the PCO plot (Fig. 7d). This individual was removed before analyzing genetic diversity and running AMOVA. In a structure analysis run only on T-DRA3 and T-EID1 without the misplaced individual, all individuals clustered to their respective populations (data not shown). No individuals showed mixed descent within any of the Structure analyses. Appendix 3 shows the graphs that underlie the decisions on optimal numbers of clusters.

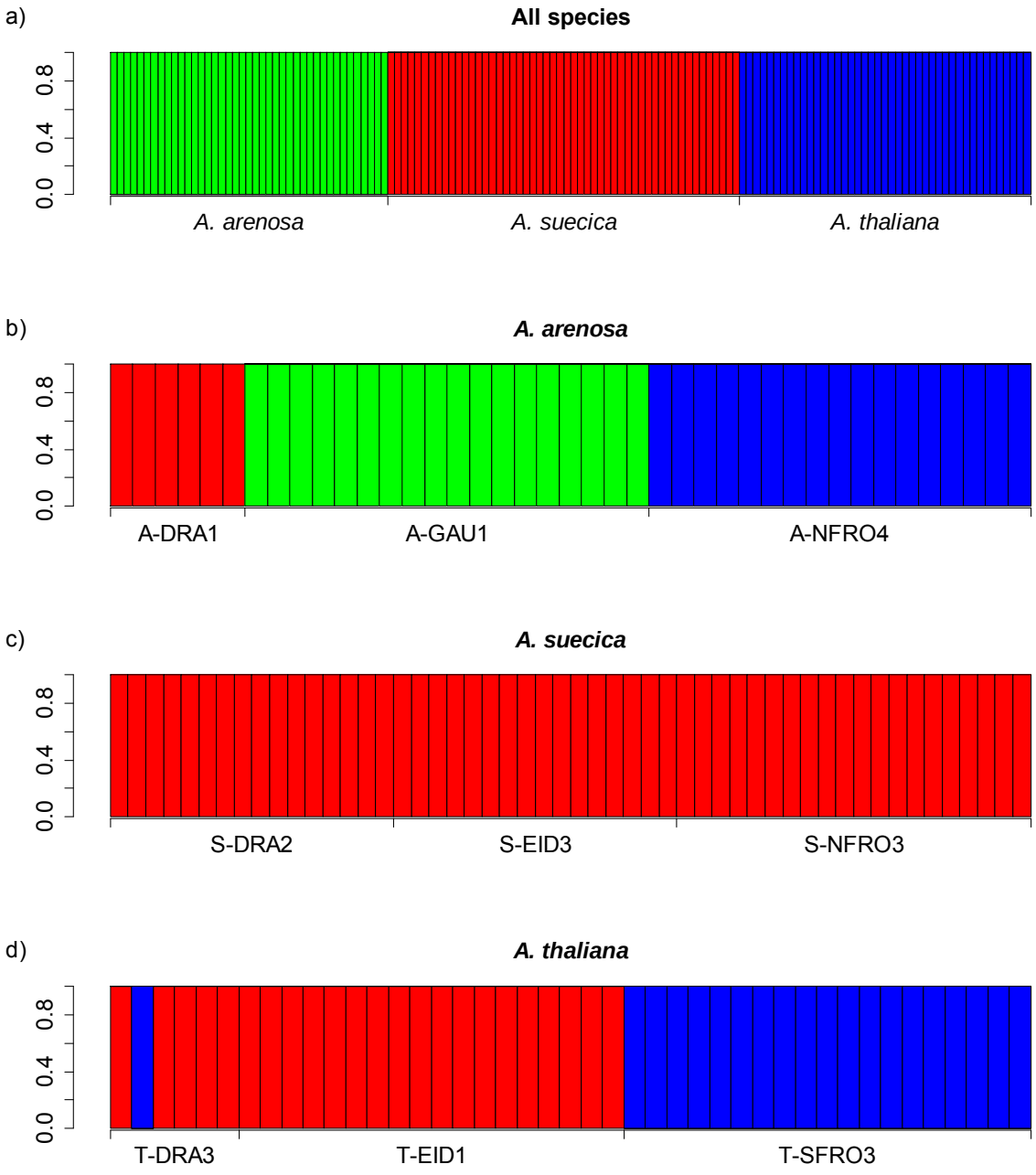


Fig. 6: Bar plots showing allocations to clusters from Structure. The vertical axis denote probability of allocation to a cluster. a) Analysis of all individuals (K = 3), b) Analysis of *A. arenosa* (K = 3), c) Analysis of *A. suecica* (K = 1), d) Analysis of *A. thaliana* (K = 2).

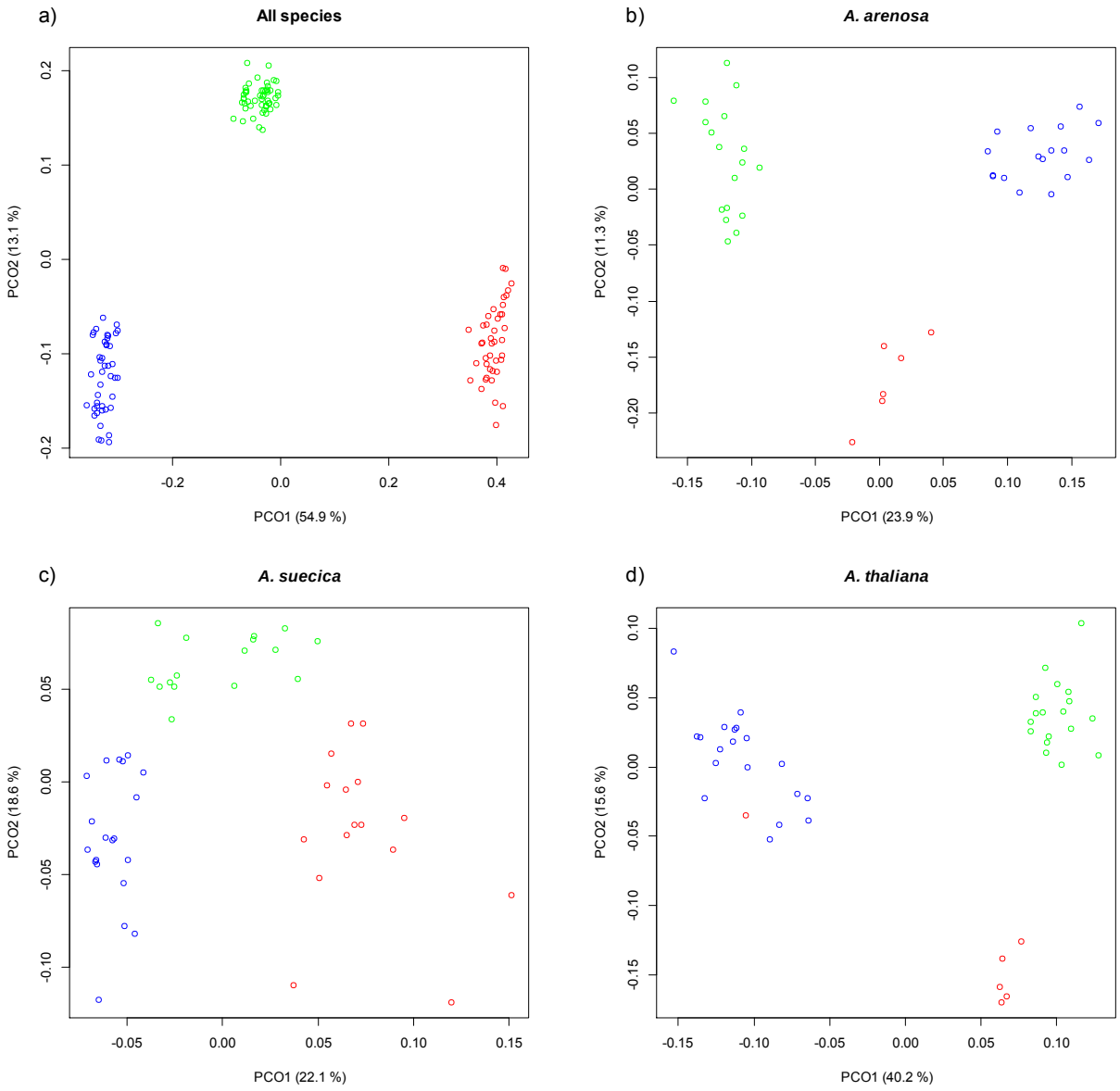


Fig. 7: Plots showing scores on the first and second PCO components from PCO analyses on Dice distances between AFLP markers. a) All species (red = *A. arenosa*, green = *A. suecica*, blue = *A. thaliana*), b) *A. arenosa* (red = A-DRA1, green = A-GAU1, blue = A-NFRO4), c) *A. suecica* (red = S-DRA2, green = S-EID3, blue = S-NFRO3), d) *A. thaliana* (red = T-DRA3, green = T-EID1, blue = T-SFRO3).

3.3.2. Genetic diversity

The *A. arenosa* populations exhibited significantly higher genetic diversity than the *A. suecica* and *A. thaliana* populations (Fig. 8), and this was confirmed on the species level (Fig. 9). Two of the *A. thaliana* populations (T-DRA3 and T-EID1) exhibited the lowest genetic diversity. For T-DRA3, the number of sampled individuals was so low that the total sample did not necessarily reflect the population diversity. More diversity was observed within *A. suecica* populations than within *A. thaliana* populations (Fig. 8). No significant difference could be found between *A. suecica* and *A. thaliana* on the species level (Fig. 9).

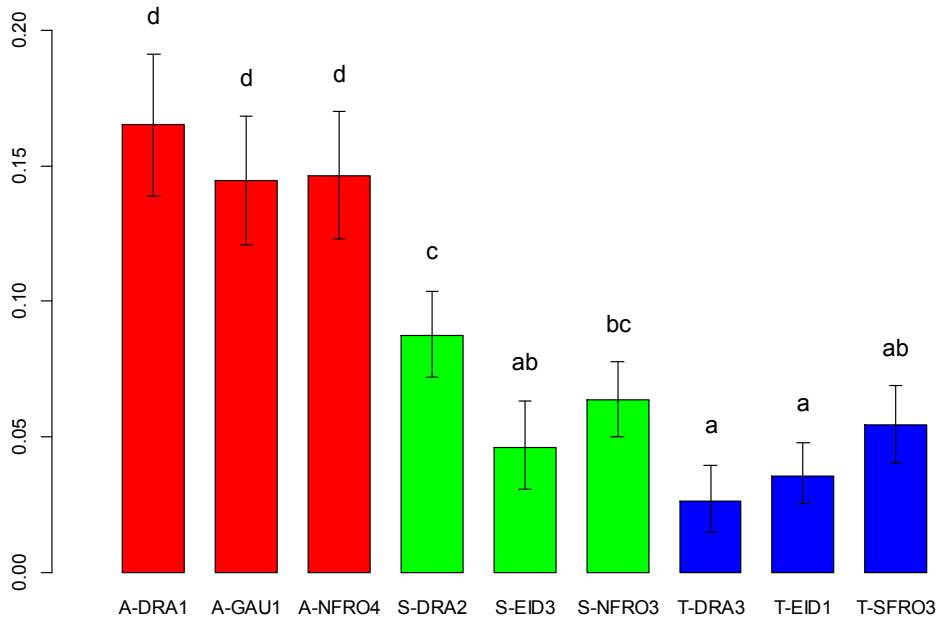


Fig. 8: Barplot of Nei's Genetic Diversity within the investigated populations. The vertical axis shows the diversity measure. Error bars denote 95 % confidence intervals, calculated using bootstrapping over 1000 replicates and all AFLP markers. Common letters denote populations that are not significantly different from each other.

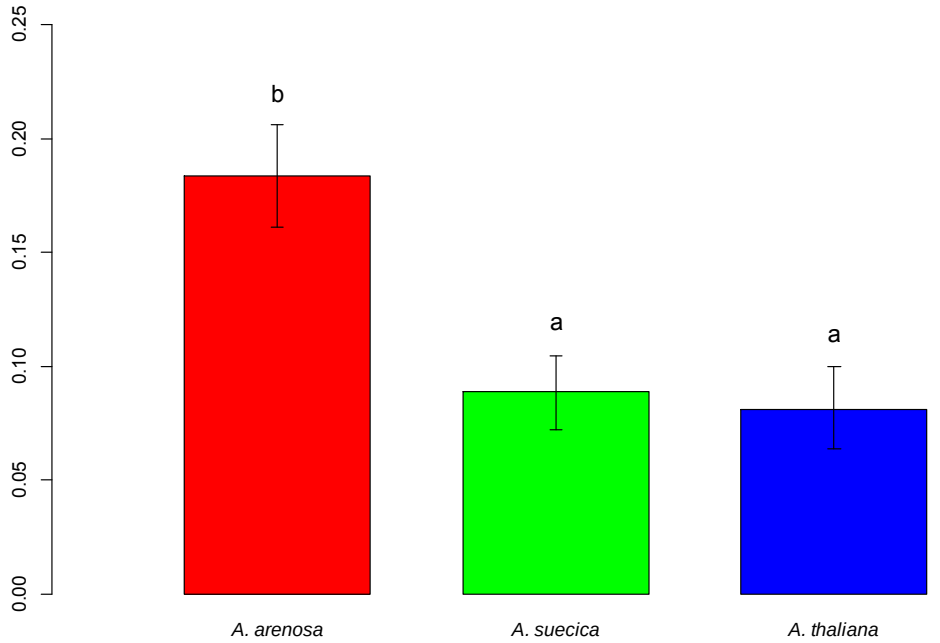


Fig. 9: Barplot of Nei's Genetic Diversity within the investigated species. The vertical axis shows the diversity measure. Error bars denote 95 % confidence intervals, calculated using bootstrapping over 1000 replicates and all AFLP markers. Common letters denote species that are not significantly different from each other.

3.3.3. Analysis of molecular variance (AMOVA)

The AMOVA showed that the between-populations percentage of variation in the AFLP markers was 27.5 % in *A. arenosa*, 34.5 % in *A. suecica* and 58.8 % in *A. thaliana* when considering the original populations (Table 5). When considering K=2 clusters in *A. thaliana*, the between-population percentage of variation was still quite high (48.6 %).

Table 5: Analysis of molecular variance (AMOVA) on a) Populations within *A. arenosa*, b) Populations within *A. suecica*, c) populations within *A. thaliana* and d) clusters within *A. thaliana* inferred from the K=2 Structure analysis. P-values for all estimations are < 0.001.

Species	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
a) *A. arenosa*	Among populations	2	234.480	7.704	27.52
	Within populations	38	770.935	20.288	72.48
b) *A. suecica*	Among populations	2	180.178	4.709	34.47
	Within populations	49	438.688	8.953	65.53
c) *A. thaliana*	Among populations	2	224.049	8.456	58.82
on populations	Within populations	39	230.856	5.919	41.18
d) *A. thaliana*	Among clusters	1	155.028	7.090	48.60
on K=2 clusters	Within clusters	40	299.876	7.497	51.40

3.4. Comparison of genetic diversity and phenotypic plasticity

The Mantel test showed a significant positive correlation between phenotypic plasticity measured as coefficients of variation and the measurements of Nei's Genetic Diversity (Table 6). This indicates that there is a relationship between higher genetic diversity and higher phenotypic plasticity on the population level among the study species. The corresponding test done with a distance matrix created from coefficients of variation scaled to unity also yielded a significant positive correlation (Table 6).

Table 6: Results from Mantel tests comparing distance matrices constructed from unscaled and scaled coefficients of variation from Table 3 with the distance matrix constructed from the estimates of Nei's Genetic Diversity presented in Fig. 8. The table shows estimates, lower and upper bound for 95 % confidence intervals and p-values for correlation.

	Estimate	Lower bound 95 % CI	Upper bound 95 % CI	P-value
Unscaled	0.648	0.543	0.750	0.019
Scaled	0.518	0.344	0.741	0.022

4. Discussion

4.1. *Arabidopsis suecica* places between its parent species in both pheno- and genotype

When considering the genetic structure analyses of all three species together, they show clear clusters corresponding to species (Fig. 6a, Fig. 7a). This implies that there is no interspecific hybridization. *A. suecica* falls between its parent species on the first PCO axis (Fig. 7a), confirming its status as an allopolyploid offspring species. This is also reflected in the NMDS created from phenotypic responses (Fig. 2), although the tendency is not as clear for the phenotypic analysis as it is for the genotypic.

Many of the phenotypic variables show common trends among the species. However, *A. arenosa* differed from the other species in several response variables when it came to response to light treatment. The reason for this might be found in the species' life histories. Since *A. arenosa* requires insect pollination (Säll et al. 2004), it could have incentives for allocating less resources to flowers and grow relatively less tall when conditions are shady and thus less attractive for pollinators (Kilkenny & Galloway 2008). The two other species are selfers (Säll et al. 2004), which means that the light gradient could have less impact on their ability to reproduce successfully. On the other hand, all three species are semelparous. Thus *A. arenosa*, as the two other species, has nothing to lose by allocating resources to reproduction in the longer run. The experiment ran for 33 days after vernalization conditions were ended, not a very long time considering that the growing season lasts 150-200 days within the sampling area (Skaugen & Tveito 2004). It is possible that other results had been observed if the experiment had lasted longer.

Nutrient availability gives similar responses in all three species. In the wild, the species tend to grow in sandy, nutrient-poor soil (Elven 2005). The similar response patterns suggest that they are able to thrive under poor conditions, but have the capability to behave opportunistically when nutrient availability improves. When it comes to water, it is hard to tell why no response was observed. One explanation might be that the applied treatments did not concur with what could be classified as high and low levels of water for *Arabidopsis* species.

4.2. No clear population structure could be identified in *Arabidopsis suecica*

From the findings in Structure, it can be inferred that the populations within *A. arenosa* and *A. thaliana* are separated (Fig. 6b and d). In *A. suecica* no population clusters could be inferred. However, the PCO plot (Fig. 7c) shows that the populations form loose clusters. When comparing the differences in population structure between the species, their histories of immigration and physical spacing of populations could offer some explanations. *A. thaliana* is at least partly indigenous in Norway, and the investigated populations are well separated. In addition, the population in Gudbrandsdal seems to be isolated from the Drammen and Eidskog populations. It is surprising that no clear population structure could be found in *A. suecica*, since inbreeding species are expected to exhibit more genetic structure among populations than outcrossing species like *A. arenosa* (Loveless & Hamrick 1984).

Both *A. arenosa* and *A. suecica* are immigrants in the Norwegian flora, but *A. arenosa* has probably been here for a longer time. There are numerous Norwegian herbarium records of *A. arenosa* from the late 19th century, while the earliest herbarium records of *A. suecica* in Norway are from 1934 (Artsdatabanken 2014). It was noted during fieldwork that *A. arenosa* tended to grow in small, isolated populations, meaning that the gene flow between populations could be limited. Meanwhile, all populations of *A. suecica* were found along railway lines. This might mean that most of the Norwegian *A. suecica* consists of a large, coherent population. Even though *A. suecica* is self-fertilizing, it has larger flowers than *A. thaliana* and might thus be more appealing to pollinators. It could be hypothesized that there is a certain degree of opportunistic outcrossing in *A. suecica* that contributes to the lack of population structure. Further on, it is plausible that the railway populations of *A. suecica* in Norway have a very recent common ancestor, and that there has not been enough time for a clear population structure to develop. It would be interesting to include "non-railway" populations in analyzes of *A. suecica*, and see whether they differ from populations growing along railways.

The AMOVA results show that the between-population percentage of variation was much higher in *A. thaliana* than in the two other species (Table 5). This is to expect, both because *A. thaliana* has grown here for much longer and thus have had better time to develop genetic

isolation between the different populations, and because *A. thaliana* is an inbreeding species (Loveless & Hamrick 1984).

4.3. *Arabidopsis arenosa* is most plastic and show the highest level of genetic diversity

I hypothesized that *A. suecica* due to its allopolyploid nature would show the highest level of phenotypic plasticity. This was not the case, and it rather seems that *A. arenosa* shows higher phenotypic plasticity than the two other study species in this experiment. This is inferred from significant differences in phenotypic responses, and from coefficients of variation. When it comes to genetic diversity, it is expected that the outcrossing *A. arenosa* would show a higher level than the inbreeding *A. suecica* and *A. thaliana*, as was found in this thesis. Lind-Hallden et al. (2002) found *A. suecica* to possess the lowest genetic diversity among the three species. However, this was not the case in my findings. They rather indicate that some of the *A. thaliana* populations were the least genetically diverse. If *A. suecica* commit opportunistic outcrossing as discussed above, it could also offer some explanation for the higher genetic diversity observed in some of the populations in the species.

It should be mentioned that the two least genetically diverse populations, T-EID1 and T-DRA3, seemed to be unstable in the wild. When visited in 2013, the year after initial collection of seeds, I was unable to locate any plants in either population. This might be a trend in *A. thaliana*. During fieldwork in 2012 we were unable to locate populations of the species on localities where it was present ten years earlier (S. Fjellheim pers. comm.). It is plausible that there still is a seed bank in the soil on localities where no plants were found. Such seed banks can contribute to increase the genetic diversity (Lundemo et al. 2009), and explain why local populations seem to be extinct in certain years. For T-DRA3 the number of sampled individuals was so low (5) that it is possible that not all variation was detected. On the other hand, much higher genetic diversity was found in A-DRA1 with a similar sample size (6).

4.4. There is a positive relationship between genetic diversity and plasticity

There is evidence for a positive relationship between genetic diversity and phenotypic plasticity (Table 6). This is an interesting result, although it is must be interpreted with caution. The general trend is a weak positive relationship between variation in molecular

markers and in morphological traits (Hufbauer 2004). AFLP markers are often considered to be neutral and mostly within non-coding regions, and thus high diversity in AFLP markers should not necessarily confer a higher expressional diversity. However, Caballero et al. (2013) investigated distribution of AFLP markers in the genome, and found that for the *Eco*RI/*Mse*I system up to 87 % of the markers were within coding regions depending on species. This means that the view of AFLP markers as exclusively neutral should be nuanced.

The positive relationship between phenotypic plasticity and genetic diversity that was found might be explained by the species' life histories. Both higher plasticity, mainly as the result of more extreme responses to the light treatment, and higher genetic diversity could be due to *A. arenosa*'s outcrossing, insect pollinated nature (Schoen & Brown 1991; Kilkenny & Galloway 2008).

4.5. Phenotypic plasticity and high genetic diversity does not imply higher fitness

Even though *A. arenosa* show higher phenotypic plasticity and higher genetic diversity than *A. suecica* and *A. thaliana*, it does not show higher fitness as assessed by the *C* variable (Fig. 5). My findings rather lean towards the conclusion that *A. arenosa* shows the lowest fitness among the study species when conditions get poor. This illustrates the importance of separating between plasticity and fitness. It is implied that phenotypic plasticity plays an important role in a species' ability to adapt to novel environments (Via et al. 1995; Davidson et al. 2011). Plasticity could also increase a species' tolerance to herbivore attacks (Agrawal 2000). This indicates that species with higher plasticity should show higher fitness, but it has been hard to establish a relationship between those two (Hulme 2008; Davidson et al. 2011). Further on, fitness homeostasis (the ability to keep up reproduction when conditions get worse) is not necessarily favoured by a high degree of plasticity in traits directly connected to fitness, and the terms "phenotypic plasticity" and "fitness homeostasis" cannot be used interchangeably (Hulme 2008). My findings support this postulation. The comparison variable *C* constructed in this thesis could constitute a way of assessing fitness homeostasis, even though it is not an established measurement.

Focus on the relationship between phenotypic plasticity, evolution and adaptation has long been present in research (Bradshaw 1965; Schlichting 1986). A postulate is that "Plasticity is favorable if the environment is variable" (Callaway et al. 2003). With this experiment taken

into consideration, it could be questioned whether said quote is true on a general basis. It could hardly be seen as favorable to lower the number of flowers in short-lived species, as was observed in *A. arenosa*. This rather promotes the view that *A. thaliana* and *A. suecica* have a wider capacity for adapting to variable environments, since they were more stable in number of flowers. Fitness is also a question of viability. Seed production and viability was not assessed in the experiment, but during harvesting and weighing of the plants it was observed that *A. suecica* seemed to produce relatively larger siliques with more seeds than *A. thaliana* under low light conditions. This observation was not assessed statistically, but it should be included in future experiments.

Fitness is difficult to measure directly. The best way would be to run an experiment over several generations and quantify fitness from that, but this was not possible within the timeframe of this project. Hence, this experiment was restricted to measuring certain variables that could be considered more or less connected to fitness. A qualitative approach to this was chosen, classifying variables as either connected to fitness or not connected to fitness. This is a crude approach, but still useful as an entrance to the concept. The choice of variables connected to fitness in this experiment (flowers and biomass) was chosen based on methods used by Davidson et al. (2011). More flowers confer possibilities for higher offspring production. When it comes to biomass, Weiner et al. (2009) advocates an allometric relationship between biomass and reproduction. In that perspective, total biomass could be viewed as a good fitness proxy.

4.6. Is *Arabidopsis suecica* a suitable model for studies of polyploidy and invasiveness?

The findings in my thesis build up under *A. suecica* as a model species for studying polyploidy and invasiveness, though they do not provide an unambigious conclusion. In Richards et al. (2006), a set of different hypotheses considering invasive species' abilities to outcompete native species are put forward. They are called "jack-of-all-trades" (an ability to keep up fitness under poor conditions), "master-of-some" (an ability to max out fitness under good conditions) and "jack-and-master" (a combination of jack-of-all-trades and master-of-some), When comparing the analysis of the variable *C* with the jack-of-all-trades scenario (Fig. 10), there is a tendency that *A. thaliana* and *A. suecica* behaves like jack-of-all-trades species compared with *A. arenosa*. However, this trend is weak and should thus be interpreted more like a possible pinpoint than a positive finding. It could also be suggested that *A.*

thaliana behaves like a master-and-jack species compared with the two other species, as it seems to have the highest *C* values overall (though this is not significant). One objection to the scenarios in Richards et al. (2006) is that it is difficult to define stressful and favorable conditions, and whether you move from poor to good or from good to better. It might be that what are reckoned as poor conditions in my experiment in fact are good conditions. However, my results are still interesting as a practical assessment to fitness under different environmental conditions.

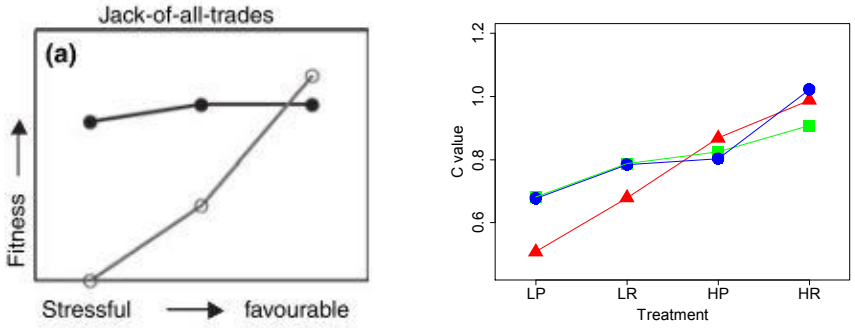


Fig. 10: Comparison of the jack-of-all-trades scenario from Richards et al. (2006) (left) with my results from Fig. 5 (right). Red = *A. arenosa*, green = *A. suecica*, blue = *A. thaliana*.

A. suecica does not behave like an invasive species in Norway at the current point. The question therefore is whether it still could be regarded as a model species for studying how polyploidy could lead to higher invasibility. Van Kleunen et al. (2010) found that invasive species are likely to show higher fitness, size and growth rate than non-invasive species. As discussed above, there does not seem to be differences between my study species when it comes to size (measured in total biomass), while the tendencies when it comes to fitness (measured as the comparison variable *C*) are weak.

When it comes to growth rate, this was not measured explicitly in the experiment. However, it could be argued that days to flowering (DTF) is a good proxy for growth rate after vernalization is ended. In that sense, it could be argued that both *A. suecica* and *A. thaliana* show higher potential for invasibility than *A. arenosa*. The practical interpretation of this is that species using shorter time on flowering and setting seeds have higher reproduction potential, since they could have more generations during one growing season. The study species are principally defined as winter annuals, germinating in the autumn and flowering in the following spring. However, it was observed that during the growing of plants for genetic

analysis, all individuals in one of the *A. suecica* populations (S-EID3) bolted and flowered without vernalization. Occurences of the same was observed in several of the other populations within all three species. This indicates that the study species have the capability to behave as summer annuals without vernalization demands. Further on, the observations discussed above that *A. suecica* might be better at setting seeds under poor conditions could also imply that it has higher capacity for becoming invasive than its parent species.

As discussed above, it seems like *A. suecica* mainly disperse along railways at the current point without venturing into surrounding areas. It is known that species often show a "lag phase" after they are introduced where they remain stable, before they suddenly become invasive (Mooney & Cleland 2001). If *A. suecica* turns invasive in a longer perspective, railway populations might constitute a source for further dispersal into vulnerable habitats. Though *A. suecica* is a small and modest species, it could still have an impact i.e. on indigenous vegetation composed of small, annual herbs that have low competitive abilites.

The growth experiment was confined to stress on only two abiotic variables, whereas in nature there are a wide array of both abiotic and biotic variables that makes up the total amount of stress. Thus, the experiment is not necessarily a good reproduction of natural conditions (Davidson et al. 2011; Drenovsky et al. 2012). In Hegarty and Hiscock (2008) it is called for field-based experiments to assess the adaptability of genetic changes that are associated with polyploids, while Davidson et al. (2011) call for experiment where multiple environmental conditions are assessed. A good way to venture further into the question of whether *A. suecica* could work as a model species for studying polyploidy and invasiveness would be to set up a competition experiment where the three study species compete over several generations under multiple environmental conditions. An experiment like that would need to take into account that *A. arenosa* demands cross-pollination. Further on, it would be interesting to assess how gene expression varies between the study species under different levels of environmental stress.

5. Conclusion

A. suecica did not show higher phenotypic plasticity than its parent species. On the contrary, *A. arenosa* seemed to be the most plastic species. Genetic diversity was also highest in *A. arenosa*, and a positive relationship between phenotypic plasticity and genetic diversity was found. However, higher phenotypic plasticity did not correlate with higher fitness in the experiment. The trend rather pointed towards that *A. thaliana* and *A. suecica* had higher fitness under poor conditions. This might mean that they have higher invasibility than *A. arenosa* in a "jack-of-all-trades"-perspective. Both *A. arenosa* and *A. thaliana* showed clearly defined population structure, while no structure could be identified in *A. suecica*. This might be because all sampled *A. suecica* were railway populations that probably immigrated quite recently from the same source. The habit of *A. suecica* to disperse along railways might constitute a future source for the species to invade vulnerable communities. Combined with the observations that *A. suecica* might have better seed setting under poor conditions and used shorter time to flower, it is conceivable that the species could be used as a future model for studying polyploidy and invasiveness. This should be investigated further through competition experiments. In light of hypotheses put forward by Comai (2005) and Chen (2007) on emerging of new epigenetic and expressional patterns in polyploids, measurements of epigenetic and expressional diversity should also be performed.

This thesis constitutes a thorough investigation of phenotypical responses, population structure and genetic diversity, and it is a piece in the puzzle of answering the call from Koch and Matschinger (2007) to investigate species in the *Arabidopsis* genus. The last word is not said in the debate on whether polyploidy is an important driving force in evolution (Soltis et al. 2014), and hopefully this thesis will contribute to enlightenment both when it comes to the adaptability of polyploidy and the connections between polyploidy and invasiveness. It is implied by Fawcett et al. (2009) that polyploidization is more facilitated in times where environmental stochasticity is high, which might lead to the rapid conclusion that the question is irrelevant for current research. However, the times we are living in now truly can be said to be stochastic for the environment, seeing as we probably have entered the sixth period of mass extinction in the Earth's history (McKinney & Lockwood 1999).

6. Litterature

Agrawal, A. A. (2000). Benefits and costs of induced plant defense for *Lepidium virginicum* (Brassicaceae). *Ecology*, 81 (7): 1804-1813.

Arrigo, N. & Barker, M. S. (2012). Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology*, 15 (2): 140-146.

Artsdatabanken. (2014). Artskart 1.6. Available at: http://artskart.artsdatabanken.no/default.aspx.

Baskin, J. M. & Baskin, C. C. (1983). Seasonal changes in the germination responses of buried seeds of *Arabidopsis thaliana* and ecological interpretation. *Botanical Gazette*, 144 (4): 540-543.

Beck, J. B., Schmuths, H. & Schaal, B. A. (2008). Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Molecular Ecology*, 17 (3): 902-915.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C. & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, 13 (11): 3261-3273.

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. *Advances in Genetics*, 13 (1): 115-155.

Caballero, A., Garcia-Pereira, M. & Quesada, H. (2013). Genomic distribution of AFLP markers relative to gene locations for different eukaryotic species. *BMC Genomics*, 14 (1): 528.

Callaway, R. M., Pennings, S. C. & Richards, C. L. (2003). Phenotypic plasticity and interactions among plants. *Ecology*, 84 (5): 1115-1128.

Chang, P. L., Dilkes, B. P., McMahon, M., Comai, L. & Nuzhdin, S. V. (2010). Homoeolog-specific retention and use in allotetraploid *Arabidopsis suecica* depends on parent of origin and network partners. *Genome Biology*, 11 (12): R125.

Chen, Z. J., Wang, J., Tian, L., Lee, H. S., Wang, J. J., Chen, M., Lee, J. J., Josefsson, C., Madlung, A. & Watson, B. (2004). The development of an *Arabidopsis* model system for genome-wide analysis of polyploidy effects. *Biological Journal of the Linnean Society*, 82 (4): 689-700.

Chen, Z. J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology*, 58: 377-406.

Comai, L., Madlung, A., Josefsson, C. & Tyagi, A. (2003a). Do the different parental 'heteromes' cause genomic shock in newly formed allopolyploids? *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 358 (1434): 1149-1155.

Comai, L., Tyagi, A. P. & Lysak, M. A. (2003b). FISH analysis of meiosis in *Arabidopsis* allopolyploids. *Chromosome Research*, 11 (3): 217-226.

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, 6 (11): 836-846.

Davidson, A. M., Jennions, M. & Nicotra, A. B. (2011). Do invasive species show higher phenotypic plasticity than native species and, if so, is it adaptive? A meta-analysis. *Ecology Letters*, 14 (4): 419-431.

De Bodt, S., Maere, S. & Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. *Trends in Ecology & Evolution*, 20 (11): 591-597.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26 (3): 297-302.

Drenovsky, R. E., Grewell, B. J., D'Antonio, C. M., Funk, J. L., James, J. J., Molinari, N., Parker, I. M. & Richards, C. L. (2012). A functional trait perspective on plant invasion. *Annals of Botany*, 110 (1): 141-153.

Ehrich, D. (2006). AFLPdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes*, 6 (3): 603-604.

Ehrich, D. (2011). Structure-sum v. 2011: A series of R functions for summarizing the outputs of the program Structure ver. 2.3.3 (and some previous versions). *Unpublished, available from the author upon request*.

Elven, R. (2005). *Norsk Flora*. 7 ed. Oslo: Det Norske Samlaget.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14 (8): 2611-2620.

Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131 (2): 479-491.

Excoffier, L. & Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10 (3): 564-567.

Fawcett, J. A., Maere, S. & Van de Peer, Y. (2009). Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (14): 5737-5742.

Flagel, L. E. & Wendel, J. F. (2009). Gene duplication and evolutionary novelty in plants. *New Phytologist*, 183 (3): 557-564.

Gederaas, L., Moen, T. L., Skjelseth, S. & Larsen, L. K. (2012). *Alien species in Norway – with the Norwegian Black List 2012*. Trondheim: The Norwegian Biodiversity Information Centre.

Hammer, Ø., Harper, D. & Ryan, P. (2001). Past: Paleontological Statistics Software Package for education and data analysis. Paleontología Electrónica 4: 1-9.

Hayashi, E., Chi, H., Boyer, S. & Still, D. (2005). Amplified fragment length polymorphism protocol for plant science on CEQ series genetic analysis system. *Beckman Coulter, Fullerton, CA.*

Hegarty, M. J. & Hiscock, S. J. (2008). Genomic clues to the evolutionary success of polyploid plants. *Current Biology*, 18 (10): R435-R444.

Hothorn, T., Bretz, F. & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50 (3): 346-363.

Hufbauer, R. A. (2004). Population genetics of invasions: Can we link neutral markers to management? *Weed Technology*, 18: 1522-1527.

Hulme, P. E. (2008). Phenotypic plasticity and plant invasions: Is it all Jack? *Functional Ecology*, 22 (1): 3-7.

Jakobsson, M., Hagenblad, J., Tavare, S., Säll, T., Hallden, C., Lind-Hallden, C. & Nordborg, M. (2006). A unique recent origin of the allotetraploid species *Arabidopsis suecica*: Evidence from nuclear DNA markers. *Molecular Biology and Evolution*, 23 (6): 1217-1231.

Kilkenny, F. F. & Galloway, L. F. (2008). Reproductive success in varying light environments: Direct and indirect effects of light on plants and pollinators. *Oecologia*, 155 (2): 247-255.

Koch, M. A. & Matschinger, M. (2007). Evolution and genetic differentiation among relatives of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (15): 6272-6277.

Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, 17 (8): 386-391.

Lewandowska-Sabat, A. M., Fjellheim, S. & Rognli, O. A. (2012). The continental-oceanic climatic gradient impose clinal variation in vernalization response in *Arabidopsis thaliana*. *Environmental and Experimental Botany*, 78: 109-116.

Lewandowska-Sabat, A. M., Fjellheim, S. & Rognli, O. A. (2010). Extremely low genetic variability and highly structured local populations of *Arabidopsis thaliana* at higher latitudes. *Molecular Ecology*, 19 (21): 4753-4764.

Lind-Hallden, C., Hallden, C. & Säll, T. (2002). Genetic variation in *Arabidopsis suecica* and its parental species *A. arenosa* and *A. thaliana*. *Hereditas*, 136 (1): 45-50.

Loveless, M. D. & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, 15: 65-95.

Lundemo, S., Falahati-Anbaran, M. & Stenøien, H. K. (2009). Seed banks cause elevated generation times and effective population sizes of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology*, 18 (13): 2798-2811.

Lynch, M. & Force, A. (2000). The probability of duplicate gene preservation by subfunctionalization. *Genetics*, 154 (1): 459-473.

Lynch, M. (2007). *The Origins of Genome Architecture*: Sinauer Associates Inc.

Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H. & Otto, S. P. (2011). Recently formed polyploid plants diversify at lower rates. *Science*, 333 (6047): 1257-1257.

McKinney, M. L. & Lockwood, J. L. (1999). Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology & Evolution*, 14 (11): 450-453.

Meyers, L. A. & Levin, D. A. (2006). On the abundance of polyploids in flowering plants. *Evolution*, 60 (6): 1198-1206.

Mooney, H. A. & Cleland, E. E. (2001). The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences of the United States of America*, 98 (10): 5446-5451.

Nei, M. (1987). *Molecular Evolutionary Genetics*: Columbia University Press.

O'Kane, S. L., Schaal, B. A. & Al-Shehbaz, I. A. (1996). The origins of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences. *Systematic Botany*, 21 (4): 559-566.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B. & Stevens, M. H. H. (2007). *vegan: Community Ecology Package.* 2.0-8 ed.

Otto, S. P. (2007). The evolutionary consequences of polyploidy. *Cell*, 131 (3): 452-462.

Pandit, M. K., Pocock, M. J. O. & Kunin, W. E. (2011). Ploidy influences rarity and invasiveness in plants. *Journal of Ecology*, 99 (5): 1108-1115.

Pecinka, A., Fang, W., Rehmsmeier, M., Levy, A. A. & Scheid, O. M. (2011). Polyploidization increases meiotic recombination frequency in *Arabidopsis*. *BMC Biology*, 9 (1): 24.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155 (2): 945-959.

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reed, D. H. & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, 17 (1): 230-237.

Richards, C. L., Bossdorf, O., Muth, N. Z., Gurevitch, J. & Pigliucci, M. (2006). Jack of all trades, master of some? On the role of phenotypic plasticity in plant invasions. *Ecology Letters*, 9 (8): 981-993.

RStudio. (2013). *RStudio*. 0.97.551 ed. Boston, MA.

Säll, T., Jakobsson, M., Lind-Hallden, C. & Hallden, C. (2003). Chloroplast DNA indicates a single origin of the allotetraploid *Arabidopsis suecica*. *Journal of Evolutionary Biology*, 16 (5): 1019-1029.

Säll, T., Lind-Hallden, C., Jakobsson, M. & Hallden, C. (2004). Mode of reproduction in *Arabidopsis suecica*. *Hereditas*, 141 (3): 313-317.

Schlichting, C. D. & Levin, D. A. (1984). Phenotypic plasticity of annual *Phlox*: Tests of some hypotheses. *American Journal of Botany*, 71 (2): 252-260.

Schlichting, C. D. (1986). The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics*, 17: 667-693.

Schoen, D. J. & Brown, A. (1991). Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 88 (10): 4494-4497.

Skaugen, T. E. & Tveito, O. E. (2004). Growing-season and degree-day scenario in Norway for 2021-2050. *Climate Research*, 26 (3): 221-232.

Sobel, J. M., Chen, G. F., Watt, L. R. & Schemske, D. W. (2010). The biology of speciation. *Evolution*, 64 (2): 295-315.

Soltis, D. E., Segovia-Salcedo, M. C., Jordon-Thaden, I., Majure, L., Miles, N. M., Mavrodiev, E. V., Mei, W., Beatriz Cortez, M., Soltis, P. S. & Gitzendanner, M. A.

(2014). Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). *New Phytologist*.

Soltis, P. S. & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (13): 7051-7057.

Soltis, P. S. & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, 60: 561-588.

Song, C., Liu, S. J., Xiao, J., He, W. G., Zhou, Y., Qin, Q., Zhang, C. & Liu, Y. (2012). Polyploid organisms. *Science China - Life Sciences*, 55 (4): 301-311.

Sultan, S. E. (2001). Phenotypic plasticity for fitness components in *Polygonum* species of contrasting ecological breadth. *Ecology*, 82 (2): 328-343.

Sultan, S. E. (2000). Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science*, 5 (12): 537-542.

te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubesova, M. & Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annals of Botany*, 109 (1): 19-45.

Van Kleunen, M., Weber, E. & Fischer, M. (2010). A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters*, 13 (2): 235-245.

Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S. M., Schlichting, C. D. & Van Tienderen, P. H. (1995). Adaptive phenotypic plasticity: Consensus and controversy. *Trends in Ecology & Evolution*, 10 (5): 212-217.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van De Lee, T., Hornes, M., Friters, A., Pot, J., Paleman, J. & Kuiper, M. (1995). AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, 23 (21): 4407-4414.

Weiner, J., Campbell, L. G., Pino, J. & Echarte, L. (2009). The allometry of reproduction within plant populations. *Journal of Ecology*, 97 (6): 1220-1233.

Wendel, J. F. (2000). Genome evolution in polyploids. *Plant Molecular Biology*, 42 (1): 225-249.

Appendix 1: Pictures from the growth chamber experiment

Fig. A1: Picture showing an overview of the high light treatment trolleys. The low light treatment trolleys are to the right, covered by light-reducing fabric.

Fig. A2: Close-up photo of an *Arabidopsis suecica* replicate from the S-NFRO3 population at the early stages of flowering.

Appendix 2: Protocol for running AFLP

The PCR machines used for incubating and running PCR were a Tetrad 2 Thermal Cycler (Bio-Rad, Hercules, CA, USA) and a Mastercycler ep Gradient Thermal Cycler (Eppendorf, Hamburg, Germany).

Restriction:

Approximately 400 ng Genomic DNA was diluted with MilliQ water to a total volume of 35 µL. 1x RL-buffer (100mM trisHAc, 100 mM MgAc, 500 mM KAc, 50 mM DTT), 0.05 µg Bovine Serum Albumin (BSA) (New England BioLabs, Ipswich, MA, USA), 0.125 units of *Eco*RI enzyme (Invitrogen, Carlsbad, CA, USA) and 0.125 units of *Mse*I enzyme (New England BioLabs) was mixed in a total volume of 40 µL. The mixture was incubated at 37°C for 75 minutes.

Ligation:

Adapters were annealed by mixing F- and R-adapters to a concentration of 10 µM (*Eco*RI-adapters) or 50 µM (*Mse*I-adapters), and incubating the mixtures at 65°C for 10 minutes, 37°C for 10 minutes and 25°C for 10 minutes (see Table A1 for adapter sequences.) MilliQ water, 0.1 µM annealed *Eco*RI-adapter (Invitrogen), 1 µM annealed *Mse*I-adapter (Invitrogen), 0.2 µM ATP (Sigma Aldrich, St. Louis, MO, USA), 1x RL-buffer, 0.05 µg BSA (New England BioLabs) and 0.02 units of T4 DNA ligase (Invitrogen) was added to the restricted DNA in a total volume of 50 µL. The mixture was incubated at 37°C for 3 hours. The ligated DNA was diluted 10x with MilliQ water.

Table A1: Oligonucleotide sequences in adapters used for ligation

Adapter	**Oligonucleotide sequence**
*Eco*RI-F	5'- CTC GTA GAC TGC GTA CC -3'
*Eco*RI-R	5'- AAT TGG TAC GCA GTC TAC -3'
*Mse*I-F	5'- GAC GAT GAG TCC TGA G -3'
*Mse*I-R	5'- TAC TCA GGA CTC AT -3'

Preamplification:

MilliQ water, 1x PCR buffer (Applied Biosystems, Carlsbad, CA, USA), 2 mM $MgCl_2$ (QIAGEN, Hilden, Germany), 0.2 mM dNTP (Invitrogen), 0.3 µM E+1 primer (Invitrogen), 0.3 µM M+1 primer (Invitrogen) and 0.038 units of AmpliTaq DNA polymerase (Applied Biosystems) was mixed to a volume of 10 µL (see Table A2 for primer sequences). 3 µL of

diluted RL-DNA was added so that preamplification was done in a total volume of 13 µL. PCR was run with the following program: 94°C for 2 minutes, then 20 cycles of 94°C for 20 seconds, 56°C for 30 seconds and 72°C for 2 minutes, then 72°C for 2 minutes, then 60°C for 30 minutes. The preamplified DNA was diluted 10x with MilliQ water.

Selective amplification:

MilliQ water, 1 µL 1x PCR buffer (QIAGEN), 0.5 mM $MgCl_2$ (QIAGEN), 0.2 mM dNTP (Invitrogen), 0.625 µM fluorescently labeled E+3 primer (Invitrogen), 0.625 µM M+3 primer (Invitrogen) and 0.025 units of HotStarTaq DNA polymerase (QIAGEN) was mixed to a volume of 7.5 µL. 2.5 µL of diluted preamplified DNA was added so that selective amplification was run in a total volume of 10 µL. PCR was run using the following program: 95°C for 15 minutes, then 10 cycles of 94°C for 20 seconds, 66°C for 30 seconds with a reduction of 1°C per cycle and 72°C for 2 minutes, then 25 cycles of 94°C for 30 seconds, 56°C for 30 seconds and 72°C for 3 minutes, then 60°C for 30 minutes. Six different primer combinations were tested, consisting of all possible E+3/M+3 combinations from the primers listed in Table A2. The three combinations that yielded the best test results were chosen for further running: E33xM37, E33xM38 and E42xM38. The amplified DNA was diluted 20x with MilliQ water before electrophoresis.

Table A2: Oligonucleotide sequences in primers used for preamplification and selective amplification.

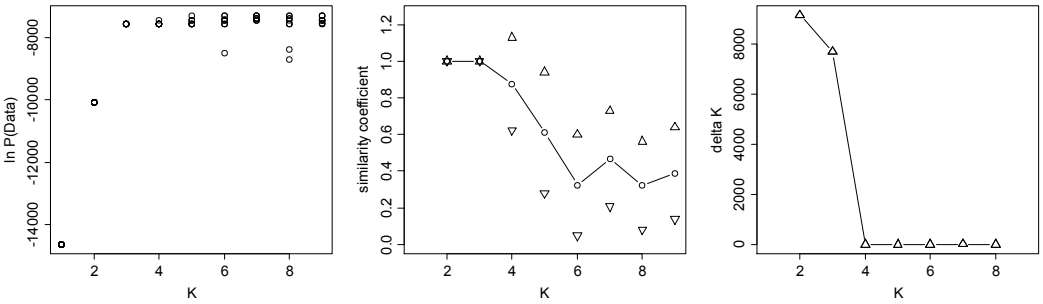
Primers	**Oligonucleotide sequence**
Preamplification primers +1/+1	
E01	5'GAC TGC GTA CCA ATT C**A**3'
M01	5'GAT GAG TCC TGA GTA A**A**3'
Selective amplification primers +3/+3	
E33 (fluorescently labeled)	5'GAC TGC GTA CCA ATT C**AA G**3'
E42 (fluorescently labeled)	5'GAC TGC GTA CCA ATT C**AG T**3'
M36	5'GAT GAG TCC TGA GTA A**AC C**3'
M37	5'GAT GAG TCC TGA GTA A**AC G**3'
M38	5'GAT GAG TCC TGA GTA A**AC T**3'

Electrophoresis:

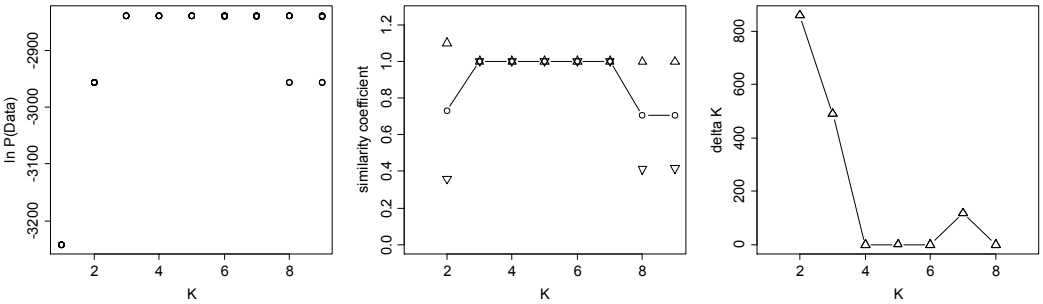
8.95 µL Hi-Di™ formamide (Life Technologies, Carlsbad, CA, USA), 0.05 µL GeneScan™ 500 LIZ® Size Standard (Life Technologies) and 1 µL diluted amplified DNA was mixed and denatured for 3 minutes at 95°C. Electrophoresis was performed with an ABI PRISM 3730 DNA analyzer (Applied Biosystems).

Appendix 3: Figures used for inferring numbers of clusters in Structure

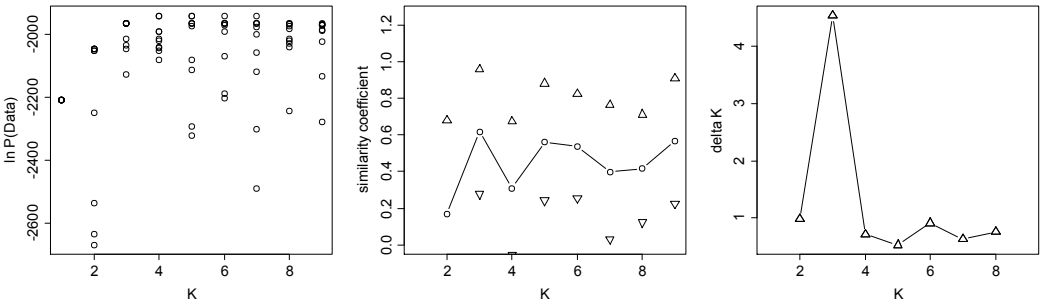
All species



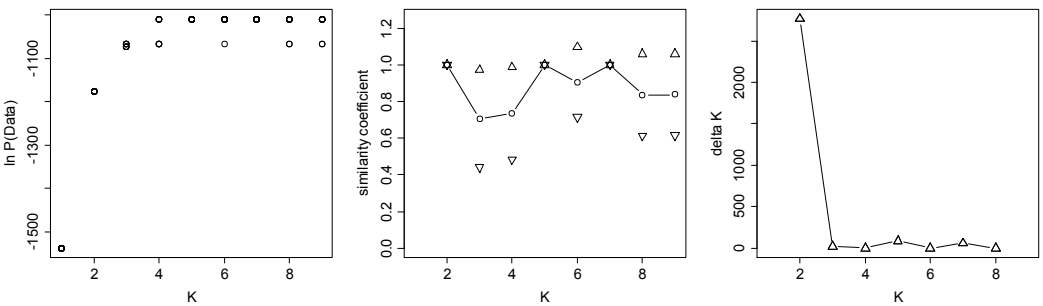
Arabidopsis arenosa

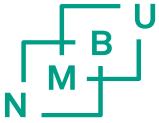


Arabidopsis suecica



Arabidopsis thaliana





Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no