Genetic Dissection of Quantitative Traits in Scots Pine

Sara Abrahamsson
Faculty of Forestry
Department of Forest Genetics and Plant Physiology
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2011

Acta Universitatis agriculturae Sueciae
2011:77

Cover: From seedling to full-grown; a Swedish Scots pine
(photo: S.Ranade)

ISSN 1652-6880
ISBN 978-91-576-7621-4

Print: Arkitektkopia, Umeå 2011

Genetic dissection of quantitative traits in Scots pine

Abstract

The phenotypic expressions of Scots pine, have been extensively studied over the years, but the gene regulation behind the traits has only just begun to be elucidated. The overall aim of this thesis was to start dissecting the genetics behind a number of adaptive traits in Scots pine and examine how they are influenced by relatedness using different molecular tools.

In a full-sib family of Scots pine the genetic variation in autumn frost hardiness and height growth was revelead by measuring open-pollinated offspring. A significant genetic variation for both traits was found, a prerequisite to identify quantitative trait loci (QTL). A conceptual statistical model in the Bayesian framework for identifying QTLs in dynamic traits i.e. traits that vary over time, was developed and applied in a QTL study based on a full-sib family of 250 trees. A set of 160 AFLPs were utilized. QTLs for three latent traits were identified: one for the slope (growth) and two for the quadratic term (growth cessation).

As a tool to identify candidate genes for the study of quantitative adaptive traits, gene regulation under continuous red (cR) and far-red (cFR) light was studied in hypocotyls from open-pollinated seeds from a natural population in northern Sweden using microarray technology. The gene expression patterns for the light response pathway in Scots pine under cFR show clear differences from those of angiosperms, wherein we observed up-regulation of cryptochrome1. This gene has, therefore, become a strong candidate gene that deserves further studies to elucidate the genetics behind Scots pine adaptation.

Not knowing the genetic relationship and inbreeding of trees, and how it influences the phenotypic expression, can lead to over- or underestimation of additive genetic values resulting in biased heritability estimates. A natural population of Scots pine, earlier identified as being highly inbreed, was used to investigate the influence of inbreeding by the correlation between heterozygosity and propotion of sound seed (PSS), average seed weight and proportion of rare alleles (PRA). Heterozygosity fitness correlation (HFC) was found positive for PSS and negative for PRA most likely due to recessive deleterious alleles purged in homozygotes. The study provided evidence that, as predicted by theory, inbreeding enhances HFC in a species with high outbreeding rate and high number of lethal equivalents as Scots pine.

Keywords: heterozygosity fitness correlation, inbreeding, molecular markers, microsatellite, QTL, cold hardiness, height, AFLP, microarray, light quality

Author's address: Sara Abrahamsson, SLU, Department of Forest Genetics and Plant physiology, SE-901 83 Umeå, Sweden *E-mail:* Sara.Abrahamsson@ slu.se

Dedication

To my children and husband.

A scientist never really understands his own theories unless he can satisfactorily explain them to the average person.

Albert Einstein

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I **Abrahamsson S**, Nilsson J-E, Wu H, García-Gil MR and Andersson B (2011) Inheritance of height growth and autumn cold hardiness based on two generations of full-sib and half-sib families of *Pinus sylvestris*. *Scandinavian Journal of Forestry* (Submitted manuscript)

II Sillanpää MJ, Pikkuhookana P, **Abrahamsson S**, Knürr T, Fries A, Lerceteau E, Waldmann P and García-Gil MR (2011) Simultaneous estimation of multiple quantitative trait loci and growth curve parameters through hierarchical Bayesian modeling. *Heredity* doi:10.1038/hdy.2011.56

III Ranade S, **Abrahamsson S**, Niemi J, and García-Gil MR. *Pinus taeda* microarray as a tool for candidate gene identification for local red/far-red light adaptive response in *Pinus sylvestris* (Submitted manuscript)

IV **Abrahamsson S**, Hallander J and García-Gil MR The general effect model explains the weak heterozygosity-fitness correlation (HFC) in an inbred Scots pine population (Submitted manuscript)

.

The contribution of Sara Abrahamsson to the papers included in this thesis was as follows:

I Heavily involved in the design of the experiment and responsible for preparing the database, conducting most of the statistical analyses, and writing most of the manuscript.

II Participated in the design of the experiment and carried out all the laboratory work to develop AFLP markers, database maintenance and assisted in the writing of the paper.

III Took part in the design and preparation of the experiment and carried out the greenhouse measurements. She also had an active role in the writing process.

IV Involved in the design of the experiment and active in the statistical analysis of the study. Responsible for the database management. Wrote the paper jointly with co-authors.

Abbreviations

AFLP	Amplified fragment length polymorphism
ASW	Average seed weight
cFR	Continuous far-red light
cR	Continuous red light
CRY1	Cryptochrome 1
CRY2	Cryptochrome 2
CT	Critical temperature
cW	Continuous white light
ESTP	Expression sequence tag polymorphism
F	Inbreeding coefficient
HFC	Heterozygosity fitness correlation
HIR	High irradiation response
HL	Homozygosity by loci
IR	Internal relatedness
LD	Linkage disequilibrium
MLH	Multilocus homozygosity
PSS	Proportion of sound seeds
QTL	Quantitative trait locus
RAPD	Random amplified DNA polymorphism
RFLP	Fragment length polymorphism
RT-PCR	Real time PCR
SE	Standard error
SNP	Single nucleotide polymorphism
SSR	Short sequence repeats

1 Introduction

In order to improve plants productivity breeders select individuals exhibiting superior phenotypes for traits of interest. These traits generally have a continuous distribution of variation. Artificial selection drives the distribution towards one extreme, in crop breeding, for example, this has led to the differentiation of wild cabbage into a wide range of cultivars, including cabbage, broccoli, cauliflower, and so on. The knowledge of how to best select has evolved over the years and mathematical framework of quantitative genetics has made it possible to elucidate the underlying genetic variability (Fisher, 1918). Compared to crop breeding forest trees are just at the beginning of domestication. One reason for this is the long generation time typical in most tree species. Recently, molecular tools have been incorporated in traditional breeding of dairy cattle by using statistical associations between molecular markers and phenotypes leading to an increased genetic gain per breeding cycle. The genomic knowledge in trees, having a huge genome size, is still being developed and molecular markers have so far not been routinely incorporated in forest breeding.

Scots pine is one of the most important economical tree species in Sweden and is therefore being breed to increase productivity. Productivity in northern Sweden does not only depend on growth capacity such as height but also on the trees ability to adapt to the local environment both by tolerating extreme weathers and different light regimes. By selecting trees the natural variation becomes manipulated, which can lead to reduced genetic variation giving rise to inbreeding depression (reduced performance compared to outbreed individuals). This should be avoided in the breeding population where genetic gain and diversity are balanced. In order to make correct estimates of genetic parameters the relatedness within the population must be known and also to which extent inbreeding influences the phenotypic variation. In this thesis we investigated if heterozygosity fitness correlation (HFC) is influenced by

inbreeding. We also explored the genetic variability of three important traits, height, cold tolerance and response to light quality in Scots pine with different molecular tools. A new statistical model has been developed in the Bayesian framework in order to fill some of the knowledge gaps.

1.1 Scots pine (Pinus sylvestris L.)

Scots pine belongs to the genus *Pinus* in the family *Pinaceae*. The genomes of Pines are large (22-30 pg), and are organized on 12 pairs of chromosomes, features that are shared with several divergent subgenera (Wakamiya *et al.*, 1993). High synteny and co-linearity of the chromosomes among different pine species have also been described (Komulainen *et al.*, 2003). Scots pine is a long-lived perennial species with some trees known to be over 700 years old (Engelmark, 1984), although the average longevity among freely growing specimens more commonly ranges between 250 and 400 years. Although it is a wind-pollinated and highly out-crossing species (Muona & Harju, 1989), inbreeding can be substantial depending on the density of stand, environmental conditions and silviculture regimes (García-Gil *et al.*, 2009; Rudin *et al.*, 1974). Inbreeding usually brings with it reduced performance, i.e. inbreeding depression (Charlesworth & Charlesworth, 1979). In conifers, high inbreeding depression is commonly observed early in development, expressed as a high abortion rate of embryos during the development of seeds (Koski, 1971; Sorensen, 1969; Sarvas, 1962). Scots pine inbreeding depression has mainly been documented from its early developmental stages (i.e Kärkkäinen & Savolainen, 1993; Muona *et al.*, 1987; Koski, 1971). However, little effort has been devoted to the estimation of the effect of inbreeding depression on later developmental stages (reviewed by Williams & Savolainen, 1996)

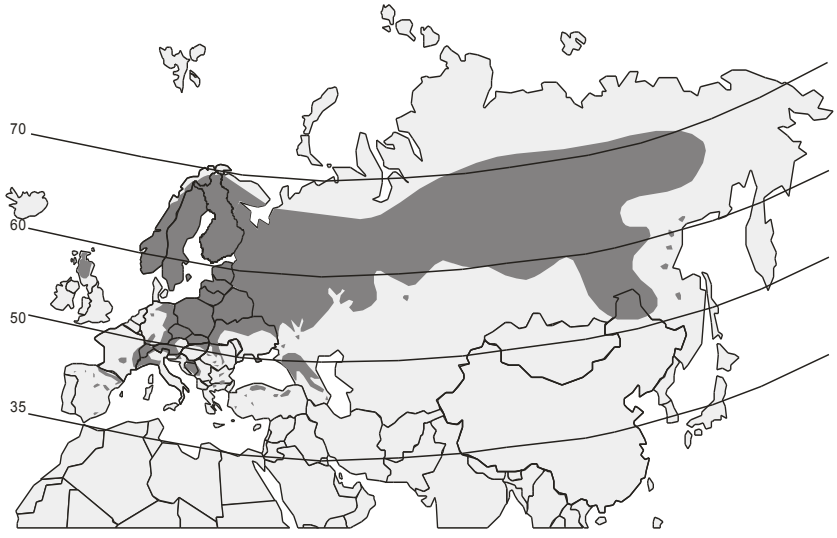

Figure 1. The natural distribution of *Pinus sylvestris* L. extending from Scotland in the east to Russia in the west and from Finland in the north to Spain in the south (from Giertych & Mátyás, 1991).

Scots pine has the widest distribution of all members of the genus *Pinus*. The natural range extends from the Spanish Sierra Nevada in the south (37ºS) to northern Scandinavia (70ºN), and from Scotland and Spain in the west (8ºW) to Siberia in Russia in the east (141ºE) (Giertych & Mátyás, 1991)(Figure 1).

The climate within its natural geographical range has created both a latitudinal and a longitudinal cline across regions of varying temperatures and rainfall (Figure 2). The clines are further enhanced by other environmental cues like photoperiod, light intensity, and light composition (Figure 2). A clear clinal pattern from north to south has been shown for several different adaptive traits such as growth, timing of bud-set and cold adaptation (Notivol *et al.*, 2007; Garcia-Gíl *et al.*, 2003; Nilsson, 2001). A longitudinal cline in frost tolerance has also been identified in populations of Scots pine in Russia and Scandinavia (Andersson & Fedorkov, 2004). Although the broad geographical range of Scots pine has forced the species to adapt to its environment wherever it is located, several studies have shown that there is little subdivision of the populations when neutral markers are assayed (Pyhajarvi *et al.*, 2007; Dvornyk *et al.*, 2002; Karhu *et al.*, 1996).

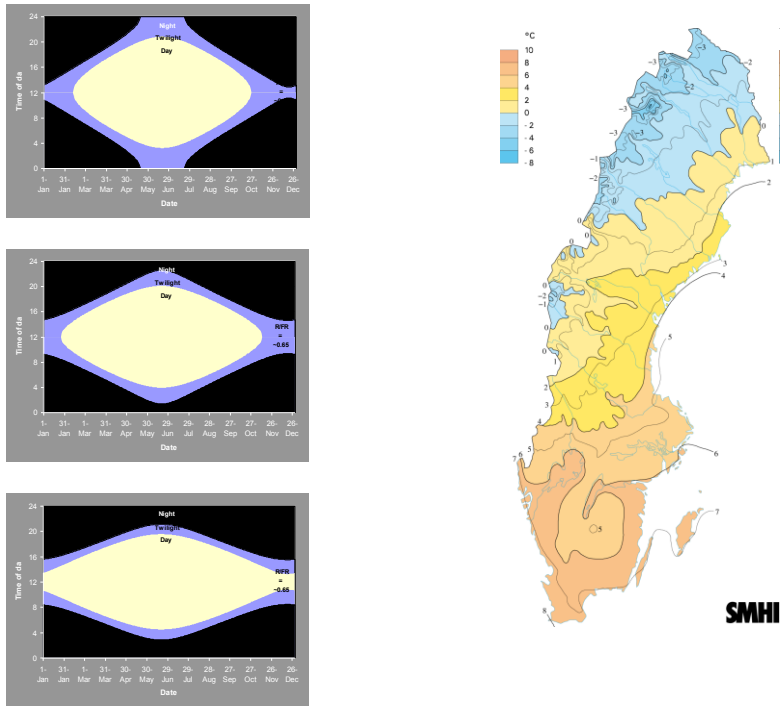

Figure 2. Average temperature throughout Sweden between 1961-1990 as reported by SMHI, on the right. The pictures to the left (adopted from Linkosalo & Lechowicz, 2006) describes how the light quality differs from south (bottom picture) to the north (top picture) throughout the year in Sweden. The black areas are darkness (night), blue area is twilight primarily containing far-red light and the yellow area is day light. The x-axel is a 24h cycle and the y-axel is divided into the months of the year beginning January.

Due to rapid climate change (Houghton *et al.*, 2001) large shifts in species composition have been predicted (Iverson & Prasad, 2002). Provenance trials indicate that the productivity of Scots pine would increase if the climate were to warm (Beuker, 1994), but any rapid shift in climate would not only alter the mean conditions but also change the frequency and level of climatic extremes (IPCC, 2007), such as frost regimes, drought, flooding and so on. The effect of these extremes could counteract any positive effect a warming climate might have on productivity. Cannell and Smith (1986) have proposed a theory in which they state that the earlier budburst in apple trees that would occur due to a warmer climate might make them more vulnerable to late frost damage in spring and so suffer a consequent decrease in growth, rather than an increase of growth, which might normally be expected in a warmer climate. This is in agreement with the results of a simulation study which included several factors such as climate, soil, nutrients, competition and demography as predictors, and

which came to the conclusion that climate change could lead to a drastic decline in the proportion of Scots pine in the north from 63% to 40% (Kellomaki *et al.*, 2001). Savolainen *et al.* (2004) have concluded that although there is a high level of genetic variation for the adaptive traits of frost tolerance and bud-set within extant populations of Scots pine, they predict a very slow adaptation to any increase in temperature. Clearly, predicting the response of Scots pine populations to climate change is not an easy task since there is a very large number of contributory factors. In this respect, it is highly relevant to dissect those genetic parameters that control adaptive traits in Scots pine.

1.1.1 Breeding program

Forest covers 67% of Sweden, Finland being the only land in Europe with a greater coverage (FAO, 2006). The forest industry represents 11.6% of the total value of Swedish exports. Of this, 38% is Scots pine, making it one of the most important commercially grown species in Sweden (Skogsstyrelsen, 2008). It is used for solid wood, paper products and biofuel.

To improve the genetics of Scots pine the Swedish breeding program was initiated in the 1950s when trees with superior phenotypes were selected in the field and grafted onto trees in seed orchards. This first round of seed orchards gave, on average, a 6% increase of production at full rotation (Rosvall *et al.*, 2001). During the 1980s a second round of orchards was established with both genetically tested and untested plus-trees. The second round gave a 10% genetic gain in production and also improved wood quality, stem straightness and branch angle, but did not improve survival rates in the north (Andersson *et al.*, 2007). The third round that is currently being established will ideally give a gain of up to 23% - 27% in volume production, and an additional gain of 5% - 13% in survival rates in harsh climates (Rosvall *et al.*, 2001). The effort in making these improvements has so far focused on three main areas: adaptation to biotic and abiotic stress, growth, and stem and wood quality in both Scots pine and Norway spruce (Karlsson & Rosvall, 1993).

1.2 Molecular Markers: Application in tree breeding

In order to be able to incorporate the new molecular tools into the traditional breeding programs of Scots pine, markers specific for Scots pine most be identified and extensively tested. The next step is to associate the markers to traits of interest for the breeders, for example growth traits and survival. The identified markers can also be used to infer relatedness in unknown pedigrees. Below is some of the more important markers described along with various applications in trees.

1.2.1 Markers

The phenotype of a tree is determined from its genetics and the environment. The genetic information of each individual is stored in the DNA as genes, some of which are being transcribed into RNA, which is translated into proteins that then regulate the building of the tree and partly determine its phenotype. This process is called "the central dogma of biology" and is illustrated in figure 3. The information from thousands of genes can be read

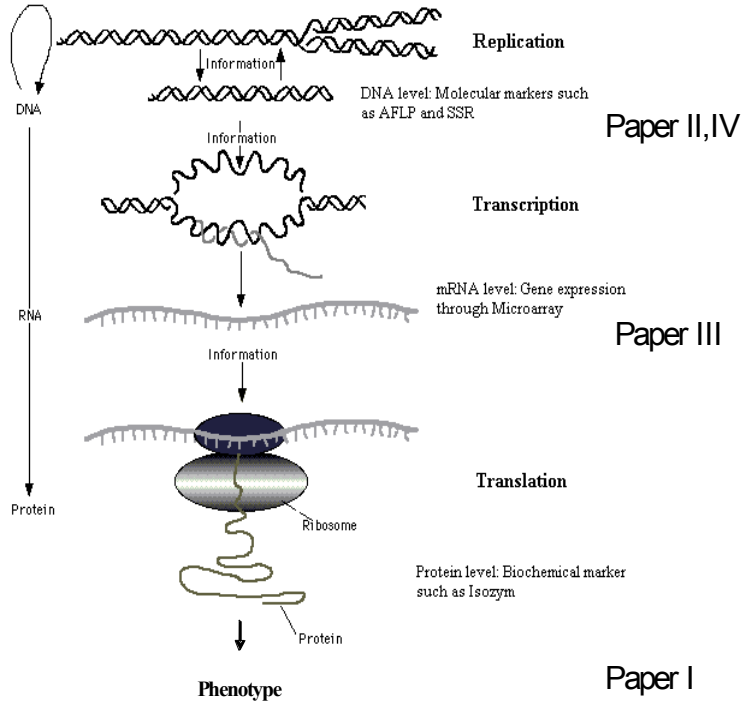

Figure 3. The information stored in the DNA is replicated from cell to cell. Within each cell the DNA is transcribed into RNA which is the template for protein. This whole process is called the central dogma of biology. In this thesis the studies has been executed on different levels within this process indicated on the right side of the figure.

from the DNA. It is possible to study genes and their expression by targeting different levels in the central dogma with help of markers (Figure 3), each of which can represent a region of possible genetic differences between individuals. Markers are generally not considered to be normal genes, as they usually do not have a known biological effect; instead they can be thought of as landmarks in the genome (FAO, 2003). The ideal marker, if existing, would be highly polymorphic, inherited in a co-dominant fashion (in order to see

which are heterozygous and which are homozygous), occur frequently in the genome, be selectively neutral in their behavior, be easy to access, easy and fast to assay, and easy to exchange between laboratories. So far no such ideal marker exists; instead the type of marker that best meets the criteria for each purpose has to be selected.

Markers can be classified as morphological, biochemical or molecular. A morphological marker is a character that can be visually characterized, such as leaf form. A biochemical marker includes allelic variants of enzymes known as isoenzymes. Isoenzymes are defined as enzymes with different isoforms but which have the same function in an individual (Markert & Moller, 1959). One of the disadvantages of using isoenzymes as markers is that they can be differently expressed in different tissues and are therefore sensitive to the environment (rewieved by Scandali, 1969). They also exist at relatively low abundances and are not highly polymorphic, and their selective neutrality is questionable (Krieger & Ross, 2002; Hudson *et al.*, 1994; Berry & Kreitman, 1993). Isozymes, being the first type of markers developed, have been extensively used in Scots pine. In natural populations they have been applied to study population dynamics (Korshikov *et al.*, 2005; Kinloch *et al.*, 1986) and fine-scale genetic structure (Wehenkel *et al.*, 2007; Yazdani *et al.*, 1985; Rudin & Ekberg, 1978) including heterozygosity fitness correlations (Savolainen & Hedrick, 1995). They have also been applied to seed orchards to investigate pollen contamination and reproductive success (Harju & Nikkanen, 1996; Shen *et al.*, 1981)

Molecular markers are DNA based markers, which therefore segregate as single genes and are not affected by the environment. DNA markers are usually divided into two groups: hybridization based markers and PCR based markers. Hybridization based markers are essentially DNA that has been digested with different enzymes and selectively identified with a labeled probe. This method is used for restriction fragment length polymorphism (RFLP). The advantages of this method are that it is fast, and many markers can be produced relatively cheaply. A great disadvantage is the large amount of DNA needed to be able to visualize them without any amplification. RFLP was mainly used in the early 1990s to build genetic maps (Devey *et al.*, 1996; Neale & Williams, 1991) but today they are hardly used due to the low throughput and the need for a priori genomic knowledge that does not exist for non-model species.

Several PCR based markers have been developed. Random amplified DNA polymorphism (RAPD) uses PCR to randomly amplify stretches of DNA. This marker requires less DNA then RFLP but it is dominant, time consuming and gives low reproducibility. RAPD was the first marker technique used to build a genetic map of Scots pine (Yazdani *et al.*, 1995) and has lately been applied in

genetic diversity studies (Fournier *et al.*, 2006; Naugzemys *et al.*, 2006). RAPD and RFLP were further developed to a technique known as amplified fragment length polymorphism (AFLP) - a marker that is a mixture of digestion and PCR amplification. AFLPs are found in abundance and they are relatively cheap to produce because there is no need to have genomic data since the technique uses non-specific primers. The disadvantage is that they are dominant. In Scots pine, AFLP has been used to build genetic maps, and to identify several quantitative trait loci (Komulainen *et al.*, 2003; Yin *et al.*, 2003; Lerceteau *et al.*, 2001; Lerceteau *et al.*, 2000)

Short sequence repeat (SSRs or microsatellites) have become the most common neutral marker of choice, especially for inferring population genetic parameters (Fisher *et al.*, 1998; Powell *et al.*, 1996). SSRs are repeats of short sequence motifs, which are highly polymorphic and have been found in most eukaryotes (Powell *et al.*, 1996). They are evenly distributed throughout the genome but are less common in exons (Goldstein & Schlötterer, 1999). Another advantage is that they are co-dominant and therefore fully informative. The disadvantage is the need for specific primers for each SSR. The high level of polymorphism among SSRs has made them the preferred marker for estimating genetic relationships and in population studies (García-Gil *et al.*, 2009; Torimaru *et al.*, 2009; Smouse & Robledo-Arnuncio, 2005; Floran *et al.*). SSRs have also been applied in gene mapping and in comparative mapping studies (Acheré *et al.*, 2004; Chen *et al.*).

A single nucleotide polymorphism (SNP) is a single base change in the DNA sequence (FAO, 2003). In species where the whole genome has been sequenced (e.g. humans and *Arabidopsis*) substantial SNP data are available, while for non-model species SNP data are still scarce. Although a single SNP is not as informative as SSRs, they are fast becoming the marker of choice because of their abundance and because new high throughput sequencing methods have been developed (Mardis, 2008). SNPs within candidate genes have been used in association mapping and genetic map construction (Eckert *et al.*, 2009; Brown *et al.*, 2003). However, as more SNPs for pines are being identified, population association mapping based on dense SNP arrays will be conducted (Neale, 2007; Neale & Savolainen, 2004). In animal breeding, SNPs are now extensively applied in genomic breeding programs where genetic gain can be inferred by markers instead of by traditional breeding methods (Goddard & Hayes, 2009).

Expressed level polymorphism (ELP) is based on differences in the expression level of genes spotted on an oligo array. Commonly, the genes in the array represent the transcriptome of an organism, therefore, this type of marker is considered to be a functional genomics approach which takes into

consideration the cellular biochemical processes (Schadt *et al.*, 2003). Microarrays were first described by Scheena *et al.* (1995). Thousands of probes for different genes are immobilized on a solid surface, most often a glass-slide. Total RNA extraction is done on the biological sample of interest. The RNA is reversibly transcribed into cDNA and fluorescently labeled. These are then hybridized to the probes and laser-scanned to detect the fluorescence. Gene expression is quantified by hybridizing two samples marked with different fluorescent dyes on the same slide, thereby enabling up and down regulation to be detected. This method has been used in several earlier studies in conifers in regard to wood formation (reviewed by Demura & Fukuda, 2007) and pathogen response (reviewed by Richardson et al., 2010).

1.2.2 Association mapping

The key objective for breeders is to improve economically important traits. In traditional breeding programs, this is done by measuring the trait in question in a known pedigree, and the underlying genetic effects are then determined statistically. Molecular methods are opening the way for a new strategy where markers, associated with specific phenotypes, can be incorporated into the breeding programs. However, before this can be achieved, the association must be identified between the marker/markers and the locus or loci behind the quantitative trait, the so-called quantitative trait locus (QTL). The concepts for detecting QTL were developed more than 75 years ago (Sax, 1923). The objective is to find a region of the genome that is associated with a specific phenotype. The phenotype of interest should be a quantitative trait, meaning that it has a continuous distribution. The theory is that these traits are controlled by an infinite number of genes each having a small effect (Falconer & Mackay, 1996). Most economically important traits in forestry are considered to be quantitative (e.g. growth, hardiness and survival). Mapping of QTL does not mean that a single gene is found; instead it may be the effect of many genes each with a small influence within the identified region (Flint & Mott, 2001). In order for them to be used in the tool kit for breeding, markers explaining a significant portion of the phenotypic (genotypic) variation have to be identified and verified.

Several different methods can be used when doing association studies; which to choose depends on the objective of the study (as for markers) and it is necessary to consider the number of individuals in the study as a factor, which has an enormous effect on the efficiency with which reproducible associations are found. To be able to detect low frequency genes, and genes of modest to low effect, large population sizes are required. Selecting the type of population is part of the association study design. The population can be highly structured

(full-sib), moderately structured (breeding population) or fully unstructured (e.g. a natural, undomesticated forest).

In conifers, which are usually characterized by a low population structure, the short extent of linkage disequilibrium (LD) makes them suitable for fine mapping alleles (Neale & Savolainen, 2004). The disadvantage is that a large number of markers is required to cover the whole genome. On the other hand, a fully structured population, such as a full-sib, allows the use of far fewer markers, but identified QTLs may include several potential genes due to high LD.

Most QTL studies in pines have been conducted on full-sib families, although a few of them have been done in several sib-ships (Kumar *et al.*, 2004) or by sampling unrelated trees in natural populations (Gonzalez-Martinez *et al.*, 2007). The QTL studies undertaken in different pines has been done on two groups of traits: in wood traits for *P. radiata* (Devey *et al.*, 2004; Kumar *et al.*, 2000)*, P. pinaster* (Pot *et al.*, 2006; Markussen *et al.*, 2003) *P. taeda,* (Neale *et al.*, 2002; Sewell *et al.*, 2002; Sewell *et al.*, 2000; Groover *et al.*, 1994) and *P. sylvestris* (Lerceteau *et al.*, 2001; Lerceteau *et al.*, 2000); and in adaptive traits in *P. taeda* (Williams *et al.*, 2007; Gwaze *et al.*, 2003; Kaya *et al.*, 1999), *P. radiata* (Devey *et al.*, 2004) and *P. sylvestris* (Yazdani *et al.*, 2003; Lerceteau *et al.*, 2001; Lerceteau *et al.*, 2000). None of these QTLs has, to my knowledge, been successfully validated and incorporated into a commercial plant-breeding program.

1.2.3 Heterozygosity fitness correlation (HFC)

In order to estimate inbreeding depression, pedigrees have to be known over the course of several generations. This is not an easy task in a species with a long generation time. An alternative method is to score neutral markers and measure the heterozygosity of the individuals leading to a separation of inbred homozygous individuals from outbred heterozygous individuals (Coltman & Slate, 2003). The individual heterozygosity is then correlated to different fitness traits, i.e. heterozygosity fitness correlation (HFC). The reason for there being an HFC has been the subject of much consideration. Three main hypotheses have been proposed. The first of these is a direct effect hypothesis (reviewed in David, 1998) in which the marker itself has a functional influence on the fitness trait by direct over-dominance. By choosing a set of markers that are selectively neutral, this first hypothesis cannot be considered as an explanation for HFC. Microsatellites are considered to be selectively neutral (Jarne & Lagoda, 1996; Queller et al., 1993), and it is therefore reasonable to neglect any direct, functional effects in studies done with these markers (although some authors have contradicted neutrality Dermitzakis *et al.*, 1998).

The second hypothesis suggests that HFC can be explained in terms of a local effect caused by an indirect association between the fitness trait and the microsatellites used. The third hypothesis under consideration is that the observed heterozygosity at a local locus estimates genome-wide heterozygosity that in turn correlates with *f*, the inbreeding coefficient of an individual. These two last models each have shortcomings. The local effects hypothesis is somewhat unlikely because studies usually use only 5 to 10 markers, and any chance linkage with such a small number of markers would indicate a much larger number of loci under balancing selection than is commonly thought to be the case. Since several studies exist where HFC is observed with microsatellites (Chapman *et al.,* 2009; Coltman & Slate, 2003) a more general explanation is called for. The third model is also questionable because, generally, inbreeding in a large, randomly mating population is too low to be detectable in most species (Balloux et al., 2004; Slate et al., 2004). However if there is a high inbreeding in the population through small population size, non-random mating, population ad-mixtures or bottlenecks heterozygosity at neutral markers can reflect heterozygosity at linked and unlinked loci (Szulkin *et al.*, 2010).

1.3 The traits studied

Plants, being sessile organisms, have to adapt to their surrounding environment. The appropriate timing of the active growth period of trees, with the seasonal changes in weather, temperature and photoperiod, is crucial for their vitality and growth, this is especially true in the northern countries were survival is depending on the trees ability to withstand harsh winter. We have therefore focused on traits that influence the productivity and adaptability of Scots pine in northern Sweden such as growth, cold tolerance, light perception and seed production.

1.3.1 Height growth

Growth in Scots pine is initiated in spring by accumulated heat sum, which stimulates vegetative bud-break and shoot elongation. First-year seedlings have a free growth pattern with new needles being formed and elongated during the whole growth season until the environment becomes unfavorable. Photoperiod is the main cue which induce growth cessation (Koski & Sievanen, 1985). After the first year, seedlings have a predetermined growth pattern due to all the needle primordia were formed in the previous year. Shoot elongation is already over in July, the rest of the growing season being used for needle elongation, diameter growth, and root extension. A rapid growth from

the beginning of a seedling's life is very important in a pioneer species such as Scots pine, to be able to compete with the surrounding vegetation.

Traits that vary over time, such as height or diameter, are defined as dynamic traits. When studying dynamic traits the time trend can be taken into account in several different ways (reviewed by Wu & Lin, 2006). The classical way, when studying markers for height, is to consider each time point as a separate measurement, and to do individual QTLs. This may be valid if it is done in a multi-trait framework so that co-variances between the years can be evaluated (Lund *et al.*, 2008; Macgregor *et al.*, 2005). Another option is to describe temporal variation by fitting a curve, and then to map those parameters that describe the curve. In conifers, QTL studies for height have been done by analyzing individual time points (Ukrainetz *et al.*, 2008; Brown *et al.*, 2001; Lerceteau *et al.*, 2001; Sewell *et al.*, 2000; Kaya *et al.*, 1999; Emebiri *et al.*, 1998) but none has been done using a curve. There is, however, one study by Ma *et al.* (2004) of a woody species where a growth trajectory was applied.

1.3.2 Autumn cold hardiness

In northern latitudes the ability of perennials to withstand freezing and other stresses during the annual seasonal changes is essential for growth and survival. Cold acclimation in late summer and fall is triggered through shorter day-length (Notivol *et al.*, 2007), quality changes of the light (Clapham *et al.*, 2002) and lower temperature (Beck *et al.*, 2004). When cold hardening is fully achieved, trees are able to withstand temperatures far below what they might normally experience, but during the growing season they are highly vulnerable to frost. Several studies have demonstrated the existence of a clear cline from north to south in the rate of cold acclimation in the fall, the northern populations being more frost tolerant (Lindgren & Nilsson, 1992). This latitudinal difference can be seen while hardiness is building up in the fall, but it is not as detectable during the winter or during de-hardening in spring (Nilsson, 2001; Hurme *et al.*, 1997; Nilsson & Walfridsson, 1995). In a study comparing Swedish and Russian populations, a clear longitudinal cline was also identified (Andersson & Fedorkov, 2004) but such an altitudinal or longitudinal dependence has not been found within Sweden (Sundblad & Andersson, 1995; Aho, 1994). In young seedlings, frost damage can lead to death or reduced growth resulting in a lowered ability to compete with the surrounding vegetation.

Strong correlations have been found between the injuries suffered by one-year-old seedlings during artificial freezing tests in the autumn, and field mortality at ages of 10 to 18 years (Nilsson *et al.*, 1991; Nilsson & Andersson,

1987; Nilsson & Eriksson, 1986). Artificial progeny freeze testing has therefore been used as a selection criteria for seed orchard material that is intended for harsh environments where survival is crucial (Andersson, 1985). Correlations between frost tolerance and height have, in some studies, been shown to be negative (Nilsson *et al.*, 1991; Nilsson & Walfridsson, 1990) indicating that selection for frost tolerance can lead to smaller trees. However, Persson *et al.* (2010) found such correlations in milder sites to be weak and even positive in harsh northern localities. They concluded that height growth expressed in a harsh environment reflects hardiness instead of growth potential leading to a negative correlation between height growth at harsh and mild sites.

1.3.3 Light perception

The perception of light is of most important for a plant, as it is the source of energy. The plant can sense several different aspects of the light such as; light quantity, light quality, direction and duration. Light play a central role in photosynthesis, phototropism and photomorphogenesis. Several different physiological mechanisms are tightly connected to light perception such as germination, hypocotyl elongation, flowering, shade avoidance and so on (Chen *et al* 2004; Kami *et al* 2010; Kendrick and Kronenberg 1994). The plants sense the light changes through different photoreceptors that are mainly divided into four families; phytochromes, cryptochromes, phototropins and the ultraviolet B photoreceptor. Phytocromes (phy) reacts to red and far-red light as a switch between an active and inactive form (Franklin and Quail 2010). Two forms of phy have been identified in conifers; phyO and phyP (Garcia-Gil *et al* 2003). Cryptochromes (Cry), phototropins and ultraviolet B photoreceptor all monitors blue and ultraviolet light in the solar spectrum.

The response to light quality has been studied extensively in conifers under dichromatic light with red and far-red ratios. It is known that red light enhances germination, especially in pine, (Kvaalen and Appelgren 1999) and far-red inhibits (Durzan *et al.* 1979). Red and far-red also induces fall-hardiness (Beck *et al*. 2004)and bud settings in conifers (Clapham *et al.* 1998). In shade avoiding trees such as pine the removal of blue light and high amount of FR compared to R, induces internode elongation and enhance apical dominance (Asakawa *et al.* 1974; Hoddinott and Scott 1996; Sarala *et al.* 2009; Warrington *et al.* 1989).

1.3.4 Seed production

Scots pine, in common with most conifers, has polyembryonic development, meaning that several embryos are developed in each seed, but only one survives to maturation. When no embryo survives an empty seed is produced.

Unpollinated embryos do thus not produce an empty seed instead they do not form a seed at all. On average 12% of empty seeds are produced after open pollination (Koski, 1971). Although Scots pine is a monoecious species, self-pollination is avoided through spatial or temporal separation of male and female strobili. If selfing does occur, most selfed zygotes are killed as embryos by a lethal inbreeding depression system attributed to deleterious lethal genes. It is therefore very difficult to get a correct estimation of the selfing rate, but those that exist estimate it to fall within a range between 10% and 25% (Koski, 1971; Sarvas, 1962). Severe inbreeding causes reduced seed yield (Franklin, 1970; Sarvas, 1962) although a mild inbreeding may not influence the seed yield drastically (Griffin & Lindgren, 1985).

Seed weight of Scots pine varies from year to year. Prescher *et al*. (2005) presented an overview from data gathered from seeds collected over a period of 16 years in natural stands in southern Sweden. The average seed weight over all years was 3.82 mg/seed with a range between 2.89 mg/seed and 4.64 mg/seed. Weight also varies with latitude, the heavier seeds being found in southern Sweden rather than the north.

2 Objectives

In order to understand the genetic behind economically important traits in one of the main tree species in Sweden, the aim this thesis was to start dissecting some adaptive traits in Scots pine species utilizing different molecular tools.

1. QTL studies are usually conducted on association between phenotypes of interest and different markers within a full-sib family, presuming that part of the observed variation of the trait is due to genetics. The aim of this study was to investigate the heritability and additive genetic variation within a full-sib family by testing their half-sib progenies. By utilizing the additive genetic variation instead of the phenotypes when performing a QTL study the environmental influence can be reduced and the identified associations would be more accurate.

2. To dissect traits that evolve over time, such as growth, the different measurements taken over time are considered as one quantitative trait instead of individual time points. The aim of this study was to develop a model in the Bayesian framework in which a curve fitted to growth simulated and real data could be associated to molecular markers in order to discover different QTLs. This approach aims to identify time independent QTLs for a more accurate description of the biological phenomena behind growth.

3. Elucidating what genes are responsible of adaptation to the local conditions would contribute to the understanding of the response of Scots pine populations to climate change. The aim was here to identify genes involved in the regulation of Scots pine hypocotyl response to two essential light wavelengths, which are known to be involved in Scots pine development and growth (e.i. red and far-red light).

4. Wether heterozygosity fitness correlation (HFC) can explain the variation in different fitness related traits, has been controversial since the beginning of the use of markers. The two main hypotheses are the local effect hypothesis and a general effect hypothesis under which a few markers heterozygosity would reflect the overall level in the plant. If for example a population is inbreed the extent of linkage disequilibrium would increase and thereby give merit to the general effect hypothesis. The aim was to study the effect of inbreeding on the heterozygosity-fitness correlation (HFC) for seed production in a Scots pine population.

3 Methodological overview

3.1 Materials and field trials

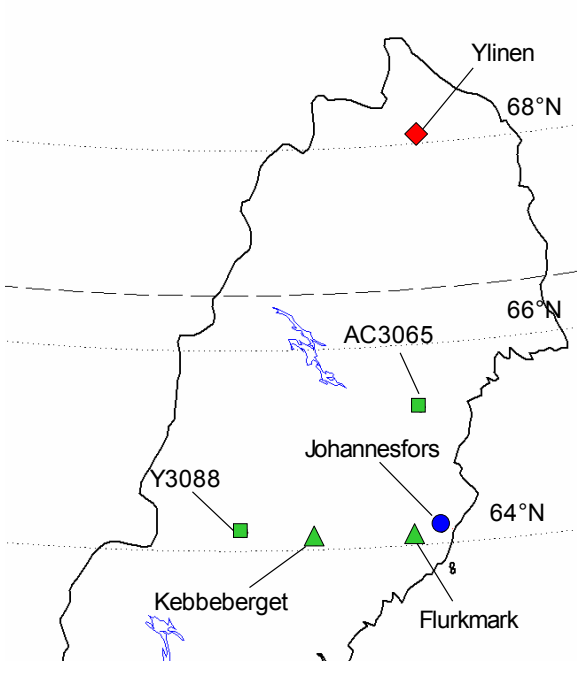

Figure 4. Locations of the studied material: locations of the two parental trees (green squares) crossed to produce a large full-sib family (green triangle Flurkmark); field site where their half sib progenies grow (green triangle Kebbeberget); the managed stand (blue circle); and the natural stand (red diamond).

Within the Swedish breeding program of Scots pine two plus-trees, AC3065 and Y3088 (Figure 4, green squares), were phenotypically selected in 1954 for

superior growth and high vitality in two mature natural stands in northern Sweden.In 1986 AC3065 was artificially pollinated by Y3088, after their superiority had been validated in progeny trials. In1988 Skogforsk established a field trial at Flurkmark near Umeå with a 1000 seedlings from this controlled cross, making it one of the largest existing full-sib families planted in a continuous plot in Sweden (Figure 4 green triangle Flurkmark). Cones from 360 trees (those trees that produced cones) were harvested in winter 2006. Due to the low age off the trees no pollen production was evident so we assume that all seeds were from external pollen sources. The open-pollinated seeds were sawn in a greenhouse producing 360 half-sib families which were planted in four progeny field trials on different localities in Sweden in spring 2008. The field trial with the highest survival rate was chosen for the study in paper I. This field trail was located at Kebbeberget close to Åsele and included a total of 4140 seedlings of 358 of the half-sib families (Figure 4 green triangle Kebbeberget). The unique genetic design, including a large set of full-sib individuals and their half-sib offspring, has given us the opportunity to estimate heritabilities of autumn cold hardiness and height growth and additive genetic variances within the full-sib family (Paper I). By utilizing breeding values instead of phenotypic values for the individuals in the full-sib family in later QTL studies the environmental influence can be excluded and the associations made will be more accurate.

In paper II, 500 individuals of the full-sib family in paper I were used to validate a new model developed in the Bayesian framework with which QTL for dynamic traits could be inferred.

In paper III cones were collected from unrelated trees in a natural stand of Scots pine located at Ylinen in northernmost Sweden (Figure 4 red diamond). Seeds from the open pollinated cones were extracted and grown in a growth chamber at Umeå Plant Science Center, Sweden under different light regimes and the gene expression was studied. A natural stand was chosen in order to be able to investigate how trees react in nature and thereby identify potential candidate genes which are involved in the sensing of different light quality.

A managed stand (Figure 4 blue circle) was established 1965 in Johannesfors, close to Sävar, with wind-pollinated seed trees. It was thinned in 1985 resulting in a stand relatively homogenous in height and age. In 2004 seeds were collected from 96 trees in the stand (Paper IV). A high inbreeding within this population has earlier been identified making it ideal to use in a heterozygosity-fitness correlation (HFC) study. Seed properties were measured at Skogforsk and used to study the influence of inbreeding on the ability to reproduce.

All of the materials are also described in table 1.

Table 1. *The locations of the studied material, the genetic design of the regeneration of the different trials, the year of establishment and the age (years) at which time the different materials where measured.*

Place	Lat, °N	Long,°E	Alt, m	Genetic design	Year planted	Age at measurement
Paper I,II						
Storklinta, AC3065	65°08'N	20°14'E	300	Plustree		
Stortjärn, Hoting, Y3088	64°09'N	16°04'E	250	Plustree		
485 Flurkmark[+]	64°02'N	20°30'E	150	Fullsib family	1988	9-19
726 Kebbeberget[£]	64° 6'N	17°41'E	400	OP*	2008	3
Paper III						
Ylinen, Karhakkamaa	68°N			Natural stand		
Growth chamber	-	-	-	OP*	2005	0
Paper IV						
Johannesfors	63°57'N	20°36'E	45	Seed trees	1965	~50

* OP = Open pollinated families, [+] prefix S23F881, [£] prefix S23F0810

3.2 Traits

3.2.1 Height growth

Seedling height in the three-year half-sib progeny trial, was measured from ground to apex, using a folding ruler (Paper I). The height (Ht) and annual growth (AG) (Papers I and II) of older trees in the full-sib trial was measured annually from the ground to the top of the terminal bud using a telescopic measuring rod.

3.2.2 Autumn cold hardiness

Autumn cold hardiness of studied individuals (Paper I), was assessed by artificial freeze testing in early autumn of current-year needles detached from the uppermost shoots according to Nilsson and Walfridsson (1995). The method is non-destructive for the tested individuals and allows freeze testing of each individual at multiple temperatures. Needle-pairs were collected from each individual and put in three or four pre-moistened zipper bags. The various bags from each individual tree were gradually frozen to preset temperatures for one to two hours, after which they were allowed to thaw slowly until they reached the starting temperature. The bags were then transferred to a climate chamber for a period of 10 days. The level of induced frost injury was then visually assessed on a scale from zero (no injury) to nine (fully discolored needles).

The critical temperature for an individual tree was calculated by linear interpolation between test temperatures as the temperature at which 50% of the needles showed at least 20% discolored tissue.

3.2.3 Hypocotyl length

Hypocotyls have been shown to be an excellent model to study plant development due to high responsiveness to growth regulators such as light quality, temperatures and hormones (Vandenbussche *et al.*, 2005). Another advantage is that it is possible to grow a large amount of individuals (genotypes) in a relative short time which is especially important when working with a large and long-lived species such as Scots pine.

In order to capture the hypocotyls in the same phase the hypocotyls were harvested when they were full-grown, as indicated by the opening of the cotyledons. This was done in order to avoid measuring growth rate and germination, instead of the total height, which would be the case if they all were collected the same day. The length was measured from the crown of the root to the needle base.

3.2.4 Seed characteristics

Two traits were scored, average seed weight (ASW) and the proportion of sound seeds (PSS). PSS were defined as the number of sound seeds per cone divided by the total number of seeds per cone. Development of non-pollinated cones with fully formed but empty seeds (parthenocarpy) is common in *Abies*, *Larix*, *Picea, Tsuga* and *Pseudotsuga* (Orr-Ewing, 1957), but it rarely occurs in *Pinus* (Sarvas, 1962), this property allows for the estimation of PSS parameters as an indicator of the level of abortion.

3.3 Light treatment

To study the global expression of Scots pine under continuous red (cR) and continuous far-red (cFR) light, seedlings were germinated and grown under continuous monochromatic light of the respective treatments. Two treatments were used: cR (620 nm) and cFR (720 nm). To imitate the sunlight, a higher intensity was used for R light compared to FR.

3.4 Genetic markers

3.4.1 AFLP (Paper II)

DNA was extracted from the vegetative buds of 250 individuals from the full-sib grown at Flurkmark using the CTAB method (Doyle & Doyle, 1990). A panel of 160 AFLPs were produced according to Vos *et al*. (1995). To produce these AFLPs, 15 primer enzyme combinations were used as follows: E-act/M-cctg, E-act/M-cccg, E-act/M-ccgc, E-act/M-ccgg, E-act/M-ccag, E-acg/M-cctg, E-acg/M-cccg, E-acg/M-ccgc, E-acg/M-ccgg, E-acg/M-ccag, E-aca/M-cctg, E-aca/M-cccg, E-aca/M-ccgc, E-aca/M-ccgg, and E-aca/M-ccag.

3.4.2 Microarray (Paper III)

The microarray described in this thesis (Paper III) was built from *P. taeda* cDNA by the J. Craig Venter Institute, USA (formerly The Institute of Genomic Research, USA). It represents 12,523 genes isolated from seedlings. The total RNA was isolated from the seedlings grown in each light treatment, cDNA was synthesized and labeled with Cyanine dyes: either Cy3 (green) or Cy5 (red). Dye-swap was used to eliminate any bias from any differential in the take-up of the dyes, i.e. both light treatments were labeled with both dyes.

3.4.3 Microsatellites (Paper IV)

Total DNA was extracted from needles. Fourteen nuclear microsatellites developed for *P. taeda* (Chagne *et al.*, 2004; Liewlaksaneeyanawin *et al.*, 2004; Auckland *et al.*, 2002; Elsik *et al.*, 2000) and *P. sylvestris* (Soranzo *et al.*, 1998) were selected and all individuals were genotyped. The microsatellites used were tested for null alleles.

3.5 Statistical analyses

3.5.1 Paper I

Genetic and phenotypic variances and co-variances were estimated using linear mixed models. Pearson correlations of measurements over several years in the Flurkmark trial were done using the SAS statistical package. Heritability was calculated using the individual model adopted from animal breeders and incorporated in ASReml (Gilmour *et al.*, 2006).

3.5.2 Paper II

A new statistical model using a Bayesian framework was developed for dynamic traits. This involved evaluating a hierarchical model and inferring the

posterior probability of QTLs for a dynamic trait. Individual functional curves were fitted and the curve parameters were treated as latent traits. The latent traits were then mapped in a multiple QTL model. The strength of this framework is that all parameters can be updated throughout the iterative procedures so avoiding being locked into suboptimal positions, which can be the case when a model is developed in several separate layers. A full model description can be found in Paper II. Based on these models, breeding values and heritabilities were calculated for the latent traits and compared to the results from Paper I by linearly regressing the breeding values obtained in Paper II against the predictive family performance obtained from the linear mixed model in Paper I. All regressions were performed in SAS.

3.5.3 Paper III

In order to eliminate systematic sources of error in the microarray study, the data produced were normalized. Several different methods can be used: in Paper III the different scan levels of each slide were merged together with restricted linear scaling (RLS) (Ryden *et al.*, 2006), followed by print-tip loess normalization (Smyth, 2005; Smyth & Speed, 2003) implemented in UPSC-BASE (Sjödin *et al.*, 2006). Computational annotation of the differentially expressed genes was done using Blast2GO (Conesa *et al.*, 2005).

3.5.4 Paper IV

In order to test if any of the microsatellites suffered from null-alleles and thereby an overestimation of homozygotes, null allele frequencies were computed using GENEPOP 4.0 (Rousset, 2008). Two SSRs with high probability of null alleles were identified and therefore exclude in all other calculations. Individual heterozygosity was calculated using the GENHET (Coulon, 2010) function within R (R Development Core Team, 2006). The aim of heterozygosity measurements is to correlate them to the individual inbreeding coefficient, *F*. Several different multilocus homozygosity (MLH) estimates have therefore been developed. Internal relatedness (IR, Amos et al., 2001) is weighted both by the expected heterozygosity and by individuals with a rare allele, these being weighted more than those with a common allele. The distribution of IR is asymmetric, which can lead to an over-estimation of homozygosity in individuals with rare alleles. This asymmetry is most problematic in cases of immigration. Homozygosity by loci (HL) (HL, Aparicio *et al.*, 2006) was designed to solve the asymmetry problem found in IR. It is a measurement that takes each locus into consideration instead of alleles, and gives a locus with more alleles a higher weight than a less informative locus. Simulations have shown that IR is more effective in inbred

lines than HL (Aparicio *et al.*, 2006). We have chosen to include both HL and IR in our study in order to compare their respective performances. The linear models and generalized linear models for traits PSS and ASW were evaluated in R (R Development Core Team, 2006).

4 Results and Discussion

This thesis describes an investigation focusing on the genetic basis of complex traits, such as growth, cold tolerance and response to light. It also includes a study of the effect of inbreeding on the correlation between heterozygosity of molecular markers and two seed related quantitative traits. The studies employed a variety of genomic and functional molecular tools. Some of the results described in the different papers arising from the work are summarized in the following chapters.

4.1 Genetic analysis of growth and autumn cold hardiness (Paper I)

In most QTL studies on trees, it is presumed that the phenotype reflects the genetic variation of the trait of interest to some extent. Full-sib families are preferable due to strong associations, although the markers need to be validated in other genetic backgrounds to be considered general. No study in pine trees has empirically shown how much of the phenotypic variation is due to genetics and how much is the environmental influence in a full-sib family. In Paper I we investigated how the phenotypic variance of autumn frost hardiness and height growth, found in the full-sib family growing at Flurkmark, was due to genetic and how much to environmental influences. We measured the traits in open-pollinated progenies from the full-sib family individuals and found significant genetic variation among the individuals in the full-sib-cross (Paper I, Table 4). The heritability within a single family is usually non-estimable without clonal replication. In this study we used a single full-sib family to estimate genetic parameters by integrating with its offspring half-sib families as a unique population different from conventional population used for traditional genetic evaluation. This gave us the opportunity to infer additive genetic variation and narrow sense heritability of the traits measured.

The heritability of autumn cold hardiness in Scots pine has been investigated in several studies. Andersson and Federkov (2004) found that the heritability for a Russian provenance was 0.22, which is in good agreement with Nilsson (1990) findings in Scandinavia. Although our heritability, h^2 = 0.37, was higher, it is in agreement with the results of a recent study by Persson *et al.* (2010) who found a narrow sense heritabilities between 0.3 and 0.54.

The estimated mean heritability for height growth was 0.19 (SE< 0.05), which is in close agreement with what is found in conventional field trials. Kroon *et al*. (2011) compiled data, collected in Sweden between 1992 and 2006, to describe the genetic variation and heritability from field tests of height increments. They found a mean narrow sense heritability of 0.22 in Scots pine trials. Some studies have found no age trend (Haapanen, 2001) while several others have shown heritability to increase with age (Jansson *et al.*, 2003; Zhelev *et al.*, 2003) including the most recent study done by Kroon *et al.* A general trend with increasing heritabilities for height growth with age, would indicate that the heritability in our population may also increase at higher ages.

The genetic correlation between height and autumn cold hardiness in our investigation was r_a = 0.39 indicating that taller trees acquires autumn cold hardiness later in the fall. This was expected because better height growth is partly an effect of later growth cessation and, consequently, later initiation of autumn cold acclimation. In harsh environments in northern Sweden where survival is important it is relevant to breed for more hardy trees. This will increase the survival, and consequently, will increase productivity per hectare despite the expected decrease in height growth. At milder sites, where survival is a minor problem, breeding is focusing on growth and wood quality instead.

The high narrow sense heritability found for autumn cold hardiness indicates that it has an excellent potential for breeding and is also a candidate for QTL studies. Through the calculation of additive genetic variability in the full-sib family breeding values for the individuals can be inferred. By using individual breeding values instead of measured phenotypic values the non-genetic part, such as environmental influence, can be reduced. The relatively low narrow sense heritability found in height demonstrates that this trait is highly influenced by environment. The accuracy of association between breeding values and markers would therefore be higher than when executing a QTL study on phenotypic values.

4.2 QTL study of dynamic traits (Paper II)

Although the Bayesian model, which was developed to identify QTLs for dynamic traits (i.e. traits that vary over time), was constructed on a conceptual basis, it was successfully applied to 250 individuals from the full-sib cross in Flurkmark. A curve fitted to height measurements across several years was used to derive parameters that could be used as latent traits. Three QTLs were found: one for the slope (speed of growth), and two for the quadratic term (growth cessation) (Paper II, Table 5). Even though the QTL probabilities and effects were generally low, the signals were stronger as compared to the general level in other positions. Slope and height are comparable because they are highly correlated in this cross (average 0.8). In a time point study by Lercetaue *et al.* (2001) in the same population, three QTLs for height were found to be stable over four consecutive years. Although this may seem to reveal that this method is just as successful in mapping QTLs, the QTLs identified by Lerceteau *et al.* are time dependent, furthermore, the trees used in Lerceteau *et al* study were still at an early juvenile age. The advantage of using the parameters of a growth curve compared to a time point study is that by avoiding any time dependency, as in our study, the results can be considered as being more reliable over the full rotation time of a tree. The marker panel used by Lercetaue *et al.* is unfortunately not the same as we used in our paper II so the positions of the markers in the two studies cannot be compared.

By using the marker information, genomic breeding values could also be inferred together with posterior heritability. The heritabilities for the latent traits were small, which indicates that genetic variation was generally low, and therefore, difficult to detect with our sample size. The weak signals of genetic variance were also found in a simulated data set also based on a small sample size. These two data sets were compared towards a simulated data set of much larger sample size where higher estimates were found; indicating that the method used is dependent on sample size. Even though the breeding values thus determined, seemed to be estimated from a population with too few individuals, a comparison of the breeding values for the slope with the predicted family performance for height of the offspring was made, no significant correlation was identified (Fig.5). This also supports the conclusion that the breeding values would benefit from the utilization of a larger sample size.

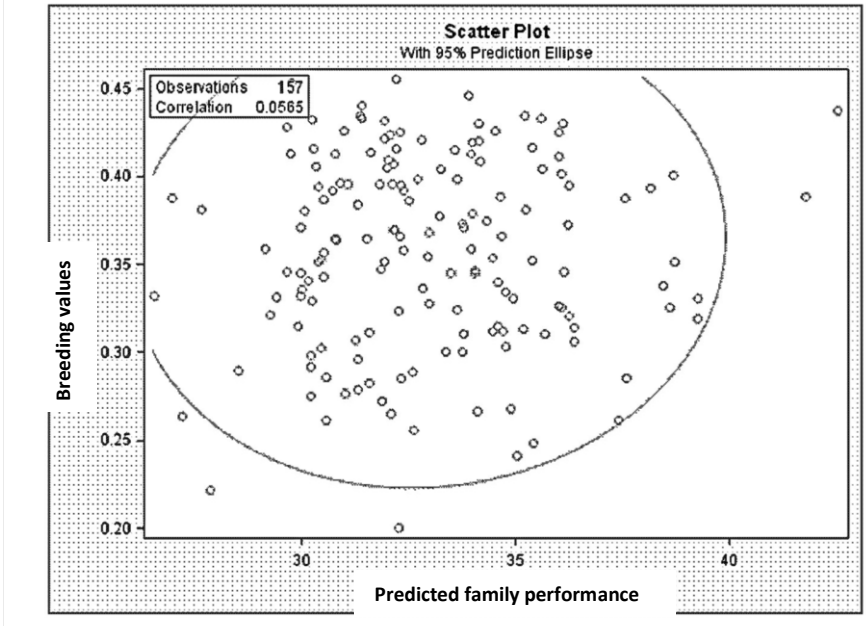

Figure 5: Scatterplot between the predicted family performance of the open-pollinated offspring from the Flurkmark cross and the genomic breeding values of the latent trait slope from the dynamic QTL study. No correlation was found between these two measurements.

4.3 Light treatment induced candidate genes (Paper III)

Plant development is tightly regulated by light, not only the amount (i.e. day-length and intensity) but also the light quality. How plants react to different light qualities such as level of red or far-red follows a cline from north to south (Clapham *et al* 2002,1998). Northern Sweden has longer days during the growing season then southern Sweden and the light also contains more far-red light in the form of twilight. It is therefore highly interesting to study a population from far north in order to capture genes that respond to different light qualities.

Hypocotyls from a northern population in Sweden, were measured after being grown under continuous red (cR) and far-red (cFR) light, and used to create cDNA libraries, which were hybridized against a microarray chip. Microarray has earlier been extensively applied to study wood formation (reviewed by Demura & Fukuda, 2007) and pathogen response (reviewed by Richardson et al., 2010) in different pine species but has not been utilized to identify gene regulation under different light treatments. Although response to light quality is not a classical phenotypic trait, it will become more important

in the light of the changing climate. Scots pine, being a long-lived perennial, will have to adapt to the change in climate. Even if the temperature increase and trees from the south can migrate north, the day length and light quality will not change at the new site, because they are primarily dependent on earth angel, leading to populations that are adjusted either to the new temperature or the light regime. This study was therefore an important step towards identifying the genes involved in the adaptation of Scots pine to different light qualities.

The lengths of the hypocotyls grown under cR were significantly shorter than those grown under cFR light (Paper III, Fig. 1). Our data concerning the lack of inhibition of hypocotyl elongation under cFR, support current thinking that conifers lack a high irradiation response (HIR) (Burgin *et al.*, 1999; Fernbach & Mohr, 1990). However, the inhibition of hypocotyl elongation has been described in *Picea abies* grown under cFR (Kvaalen & Appelgren, 1999; Scharff, 1962). This may be explained by the shade avoidance mechanism, which is typical in shade intolerant species like pines, but not found in shade tolerant species like spruce (Jarvis & Leverenz, 1983).

In the microarray, 405 genes were found differentially expressed under cR compared to the 239 genes that were differentially expressed under cFR. Thus, there were more genes that were expressed in the metabolic pathways under cR than under cFR. This is in agreement with the results of previous studies on the response of Arabidopsis seedlings to light quality (Ma *et al.*, 2001).

To avoid confounding the results with other stress factors, such as temperature and water potential, we kept temperature and humidity constant. Although we saw differential expression of several different genes, such as those coding for the L-ascorbate peroxidase and gluthathione reductase enzymes that are involved in protection against reactive oxygen species, no necrosis or seedling decay was observed under either of the light treatments. This probably indicates that although the plants were under stress, they did not reach a threshold level of stress. These findings support the conclusion that the observed variation in response was due to the different light regimes, and not to any other induced stress.

With the exception of those involved in photoreception, most of the genes involved in the metabolic pathways observed in our study were regulated in the same way as those in angiosperms such as *Arabidopsis*. The array used in the present study did not include genes for the phytochromes, but we did see indirect regulation of genes involved in mechanisms that interact with the phytochromes in both cR and cFR (Paper III, Table S1 and S2). Cryptochrome1 (CRY1), another photoreceptor, was differentially expressed under cFR compared to cR, and has previously been reported to inhibit

hypocotyl elongation under blue light (Ahmad & Cashmore, 1993; Koornneef *et al.*, 1980). However, nothing has been published about the regulation of CRY1 under other light conditions, although a CRY1-CRY2 mutant has been shown to have a lower level of mRNA expression under both blue and red light than the wild type (Lin *et al.*, 1996; Lin *et al.*, 1995), but the hypocotyls of the CRY1-CRY2 mutants were found to be shorter. CRY1 has also been shown to be induced by blue light suppressing levels of gibberellins and auxin or sensitivity to them (Folta *et al.*, 2003). However, in our study, auxin responsive protein and gibberellin receptor genes were differentially expressed under cFR. Auxin is known to elongate the hypocotyl in Arabidopsis (Romano *et al.*, 1995) and gibberellin has been shown to induce internode elongation, cell division, and cell elongation in beans (Beall *et al.*, 1996). An auxin responsive gene was found to be suppressed under cR light, which suggests a link to the promotion of genes involved in flavonoid biosynthesis under cR light since the flavonoids are known to inhibit auxin transport. Similarly, an ethylene responsive gene was also induced under cR light, and ethylene is considered to be involved in growth inhibition.

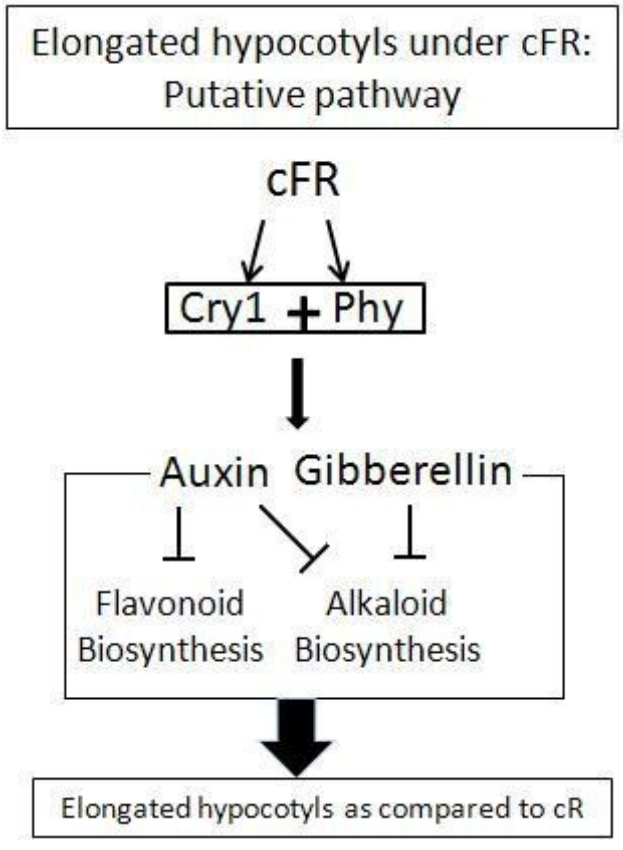

Figure 6. The proposed pathway controlling the elongation of the hypocotyl elongation under far-red (cFR) light compared to red (cR) light. The cFR light acts through Cryptochrome 1 (Cry1) and Phytochrom O (Phy) and up-regulates auxin and gibberellin production. These inhibit the production of flavonoid and alkaloid biosynthesis leading to elongation of the hypocotyls.

The molecular mechanism behind the regulation of CRY1 under cFR has not yet been studied. El-Assal *et al.* (2001) have described variation in cryptochrome that leads to altered biochemical properties, and CRY1 has been shown to be functionally dependent on PhyA and PhyB in Arabidopsis under blue light (Ahmad & Cashmore, 1997). It may therefore also be the case that the CRY1-phytochrome interaction functions differently in pine species then Arabidopsis, and so leads to the elongation of hypocotyls under cFR, thus allowing the production of auxin and gibberellins (Fig 6), although we have not measures the auxin and gibberellin levels.

We propose to that Cry1 is one candidate gene that is involved in the reaction to light quality. Other interesting genes identified are the light dependent genes aspargine synthetase, Ribulose-1,5-bisphosphate carboxylase

oxygenase and Cinnamoyl-CoA. To further study the system behind Scots pine adaptions to different light treatments a real-time PCR study including some of the identified candidate genes would be highly interesting but we foresee substantial problems because of the complexity of the genome and the insufficient genomic resources.

4.4 Heterozygosity-fitness correlation (Paper IV)

Heterozygosity-fitness correlations (HFC) have been studied since the first genetic markers were developed. Three main hypotheses have been proposed; a direct effect, local effect or genome wide hypothesis. The direct effect hypothesis (reviewed in David, 1998) in which the marker itself has a functional influence on the fitness trait. This theory is however no longer of interest when the markers preferred today i.e. microsatellites, are believed to be selectively neutral. The second theory is that HFC is due to a local effect caused by an indirect association between fitness and molecular marker. The third model under consideration is that the observed heterozygosity at a local locus estimates genome-wide heterozygosity that in turn correlates with F, the inbreeding coefficient of an individual.

Several meta-analyses have been done over the years including both plants and animals (Chapman *et al.*, 2009; Coltman & Slate, 2003; Britten, 1996). All these studies conclude that HFC is generally positive although very weak, while non-significant and negative correlations of fitness traits with heterozygosity are less often reported. However, several authors have pointed out the potential bias towards the publication of positive significant results, together with the application of inappropriate statistics and in correct definition of fitness traits (Szulkin *et al.*, 2010; Chapman *et al.*, 2009).

Based on theoretical expectations, HFC should be enhanced in inbred populations (reviewed by Szulkin *et al.*, 2010). Our Scots pine population showed high levels of inbreeding (García-Gil *et al.*, 2009)(mean of inbreeding coefficient =0.25, Paper IV, Fig. I) making it especially suitable for the detection of HFC. We also found that the SSRs, to large extent, departed from Hardy-Weinberg equilibrium and had a deficiency of heterozygotes, which further support non-random mating or mating among relatives.

The traits measured, PSS and ASW, both showed an unusual high variation within the population compared to earlier reported findings. The percentage of sound seed varied between 36% and 96% (Paper IV, Fig 1). Several studies done in seed orchards show a low variation (8-27% of variation) (Almqvist and Pulkkinen 2005; Yazdani *et al.*1995) in which the design is optimized to reduce the co-ancestry between the trees within the orchards. Griffin and

Lindgren (1985) did a comparison of how different levels of inbreeding influenced the level of PSS and found that selfed had as low as 33 % of sound seeds while half-sibs had 82%. The high variation found in our population therefor further supports a mixture of different levels of inbreeding.

Szulkin et al (2010) stated that all models expect linkage to increase the HFC, the question is therefore not whether local effect exist, but if it is great enough to influence the overall HFC effect. Following the statistical method proposed, the global effect could be invoked. In light of this result single-locus heterozygosity was not assayed, instead we carried out regression of fitness on MLH (HL and IR). We observed a positive association between proportion of sound seed and the two measures of homozygosity, IR and HL (Paper IV Fig. 4), meaning that the trees with lower heterozygosity had higher proportion of sound seeds. This enhanced HFC can be explained by the high inbreeding coefficient which has produced an increased variation within the population. There was no significant association between ASW and MLH which could be expected because seed weight has been shown to be highly influenced by the maternal environment (e.g. Castro 1999). In order to see the influence of rare alleles on fitness we correlated proportion of rare alleles on MLH and found a significant negative correlation with HL (Paper IV Fig. 3). The individuals carrying more rare alleles are also more heterozygote in our population. Most of these rare alleles can be expected to be recessive deleterious alleles due to the high rate of lethal equivalents carried by conifers (Sorensen 1969, Franklin 1972). These alleles would therefore been exposed in homozygous individual, leading to abortion of the seed. The heterozygotes that still carry the rare alleles would therefore have a disadvantage, which would explain the positive association between PSS and homozygotes.

The inference of inbreeding coefficient through microsatellites can be influenced by null alleles. A microsatellite allele that consistent fail to amplify, through PCR, to a detectable level is a null-allele. There are mainly three sources of null-alleles in microsatellites. The first one is due to a mutation in the actual primer annealing site leading to poor amplification. The second is due to differential amplification of size-variant alleles (Wattier *et al.*, 1998). Null-alleles can also be due to inconsistent DNA template quality. The second and third source for null-alleles can be kept under control through careful preparation of the PCR amplification conditions i.e. prepare the template carefully and load a higher concentration of template. Apart from these “real” null-alleles, several different population events can give the false impression of null-alleles. We therefore choose to test for null-alleles, even though the test overestimates the null-alleles frequency, and excluded the alleles with the highest probability of null-alleles. Another way to test for null-allele would be

by testing mega-gametophytes although this test would only identify null-alleles specific to the tree from which the seed were picked. Basham et al (personal communication April 11, 2011) has evidence that due to somatic mutations, null-alleles can even be specific to part of a tree. It is therefore sounder to statistically test for null-alleles.

5 Conclusions and future perspectives

Rotation times in the northern countries for soft woods are between 60 and 120 years which means that the trees that are planted today need to withstand future hazards, such as the changing climate or other upcoming events. It is therefore important, as a breeder, to take this into account when choosing trees for the future. The aim of tree breeding is to improve the trees as much as possible under as short as possible time without jeopardizing the trees ability to survive and withstand biotic and abiotic stress. To be able to do this, knowledge about the traits of economic and ecological interest is of great importance. The work presented in this thesis has provided new insights into the molecular basis of the response of Scots pine to light treatments and several candidate genes were identified. It has also shown that there is a genetic basis in the variation of height and frost hardiness within a full-sib family and it also indicates that increased accuracy can be achieved by using the predicted family performance to reduce the environmental influence on the QTL studies. The thesis also has improved the methodology with which functional traits can be mapped. It finally has given an insight into how HFC can be created in pines.

In Paper I we have provided empirical evidence that the phenotypic variation in height and critical temperature within a full-sib family, is linked to significant genetic variation, which is a prerequisite when identifying QTLs for any trait. This investigation also illustrate that by using open pollinated offspring to estimate genetic values for the mothers would improve the ability to get accurate genetic associations between markers and traits of interest. Our work indicates that it is more accurate to estimate variability at later developmental stages, and with additional sets of progeny tests in order to exclude any confounding interactions that may exist between genetic and environmental sources. Our intention is to conduct an association study in the future, in which SNPs, AFLPs and SSRs will be used to map the identified additive genetic effects instead of phenotypes.

In future association studies, more complex pedigrees should be taken into consideration in order to identify the genes that underlay economically important traits such as height and frost tolerance. With the molecular marker developing technology approaching a highly time and cost effective stage, the collection of phenotypic data will become the most critical step in any study. One way to ameliorate this would be to use pedigrees in existing breeding programs, the phenotypes of which have already been extensively investigated. By utilizing such breeding populations, any QTLs found could be directly incorporated into the breeding program. Such QTLs could first be used for pre-screening in order to identify seedlings with a favorable genetic structure. In this way, more seedlings could be tested before being introduced into expensive field trials, thereby increasing the gain from each breeding cycle.

In paper II we provided a new method with which it is possible to estimate QTLs that do not take individual time points of the some trait as being different QTLs, instead it provides a way to more accurately estimate them as varying over time. This way the identified QTLs would be stable over time and also provide a more biological meaning than single data points. This method will be useful in later studies especially as the power of the method increases with larger dataset, which is the most likely scenario in the future.

In paper III, we studied the genetic basis of Scots pine response to two different light treatments. CRY1 was identified as a suitable candidate for elucidating the light response seen in Scots pine grown under cFR. Initially, the amount of differentially expressed of CRY1 under cFR should be studied with RT-PCR, and should include other genes known to be involved in the response to light such as Phytochrome O and Phytochrome P. A study is under way in which we are including the two phytochromes and also three variants of Cryptochrome with RT-PCR although such study in a species with pseudogenes most likely will be problematic. In this study we are also going to include more light qualities such as blue, white, red and far-red treatments on hypocotyls from both south and north Sweden to investigate the cline more extensively. Given that shade avoidance may be part of the different responses of Scots pine and spruce, it would also be interesting to compare the two species directly. Such a study could be done at several different life stages including hypocotyls, seedlings and adult trees.

In paper IV we had the opportunity to study the influence of inbreeding, through an atypical inbreed population of Scots pine, identified by Garcia-Gil et al (2009), on HFC allowing us to further support the hypothesis that inbreeding explains HFC and not the local effect hypothesis. This is most likely also the case in other conifers with similar characteristics in terms of population size, mating system, pollen dispersal and so on. The reduced

production of sound seeds that heterozygotes showed were most likely due to a large proportion of deleterious rare alleles carried by the heterozygotes. It could be argued that, a certain percent of inbreeding to create purging of rare alleles would increase the fitness of a breeding population but we only tested for seed production traits and have not taken growth, survival and so on into consideration. In order to be able to generalize the result more traits must be taken into consideration.

The population, which was used in paper IV, is believed to have an atypical high inbreeding for being a natural regenerated stand, but this has not yet been tested. We intend to test several populations regenerated in different ways, such as natural population, natural regeneration with seed trees and a planted population, in order to investigate if inbreeding generally is high in stands generated by seed trees or if this is specific to the stand investigated in paper IV.

6 Sammanfattning

Tallens genetik har under lång tid studerats indirekt genom att observera och analysera olika egenskaper hos träden och deras avkommor, men den faktiska genregleringen av de flesta egenskaperna är ännu inte klargjorda. Syftet med denna avhandling har varit att ta ett första steg mot att förstå den genreglering som styr egenskaper så som höjdutveckling, köldtålighet, och ljusberoende samt inavel hos tall.

I det traditionella sättet att identifiera lokus för kvantitativa egenskaper (QTL) använder man sig ofta av helsyskonfamiljer. Det antas då att den variation som observeras i en egenskap till viss del beror på genetiska faktorer och inte bara miljön. För att fastställa om så verkligen är fallet studerade vi variationen i köldhärdighet och höjdtillväxt för en helsyskonfamilj och för ett stort antal vindpollinerade halvsyskonfamiler från denna familj. Den genetiska variationen i båda egenskaperna visade sig vara statistiskt signifikant i helsyskonfamiljen vilket stöder antagandet att man kan använda sig av en helsyskonfamilj vid QTL analyser.

Egenskaper som varier över tiden så som till exempel höjd kan statistiskt hanteras på flera olika sätt. Ett sätt är att räkna ut QTLs för varje år för sig och anta att både QTLerna och mätningarna ar helt oberoende från år till år men eftersom ett träd som växt mycket året innan i förhållande till andra träd kan antas växa bättre även nästa år så är inte mätningarna oberoende. Istället kan man inkludera beroendet i modellen. Vi har utvecklat en modell i vilken en funktion eller kurva anpassas till varje individs mätningar och använder kurvparametrarna som latenta egenskaper när man associerar dem mot markörerna. Tre QTL för höjd identifierades med hjälp av denna modell men den statistiska signifikansen var låg. En QTL för lutningen i.e. tillväxten och två för den kvadratiska termen i.e. tillväxtsavslutning. QTLs identifierade med denna metod är stabila över tiden och har en större biologisk mening än QTL identifierade för varje separat tidpunkt.

De flesta aspekter av ett träds liv styrs av ljus, inte bara av dagslängd och ljusintensitet utan även av den ljuskvalitet trädet exponeras för. Ljuskvaliten är olika beroende på var i Sverige man befinner sig, exempelvis har norra Sverige betydligt mycket mer skymningsljus (bestående av långvågigt rött ljus) än södra Sverige. Detta har gjort att det existerar ett klinalt samband mellan ljuskänslighet och latitud. Vi har med hjälp av microarray identifierat vilka gener som utrycks i kontinuerligt rött ljus (kR) i jämförelse med kontinuerligt långvågigt rött ljus på gränsen till vad ögat kan uppfatta (far-red på engelska, förkortat kFR). Genom att använda oss av ett chip bestående av ett objektglas på vilken tusentals olika små DNA avsnitt är fixerade och hybridiserat dem med cDNA från tall hypocotyler (=stamdelen mellan rot och nedersta hjärtbladet hos en groddplanta) som växt under de olika ljusregimerna, har vi identifierat ett antal gener som är involverade i plantans reaktion på dessa ljusregimer. Hypocotylerna blev betydligt längre under kFR jämfört med kR. De flesta gener som identifierades var samma som även har identifierats i organismmodellen backtrav.Vi fann dock att under kFR så uttrycktes ett antal gener kopplade till fotoreceptorerna, avvikande från vad som tidigare rapporterats för backtrav. Microarray chipet som vi använde saknade fytokrom generna men vi såg en uppreglering av ett flertal gener som interagerar med dessa under båda ljusregimerna. Under kFR så uttrycktes även Cryptocrome 1 (Cry1) och identifierades därför som en potentiell kandidatgen med vilken man kan studera latitudrelaterade skillnader hos tall.

Vid beräkning av genetiska parametrar inom skogsträdsförädling så antar man att de valda ursprungsträden var helt obesläktade. Om detta antagande inte stämmer kan det göra att genetiska parametrar, som ärftlighet, över- eller underskattas. Garcia-Gil m.fl. identifierade 2009 en naturligt föryngrad population med hög inavel. Deras modeller visade att populationen med största sannolikhet hade uppkommit genom att besläktade träd korsats och producerat plantorna, som utgjordes av en blandning av helsyskon, halvsyskon och kusiner. För att veta hur stort inflytande inavel kan få på olika konditionsrelaterade egenskaper så som reproduktion har vi tittat på korrelationen mellan heterozygositet och fröproduktion med hjälp av mikrosatelliter. Det har under en längre tid pågått en diskussion om vad korrelationer mellan heterozygositet och konditionrelaterade egenskaper (HFC) beror på och varför man bara i vissa fall kan detektera dem. Två huvudhypoteser har vuxit fram; att det beror på en lokal effekt där markörerna är direkt kopplade till egenskapen eller att markörerna visar hur stor variation det finns i hela genomet. Genom att testa markörerna var för sig mot fenotyperna och även göra en multimarkör analys kunde vi utesluta att det enbart var en lokal effekt, utan att inaveln hade ökat sammankopplingen av

genomet och därigenom skapat en genomomfattande effekt. I träd som normalt har väldigt låg inavel kan man således inte hitta HFC, men i de fall där populationshistoriken medfört ökad inavel så kan HFC detekteras. Vi fann att träd som var mer homozygota producerade större proportion friska frön än träd med högre heterozygositet. Eftersom tallen har en stor andel sällsynta alleler som förväntas vara dödliga eller väldigt viltalitetsnedsättande testade vi om homozygoter hade färre sällsynta gener - och så var fallet! Man kan förklara homozygoternas fördel genom att de till större del sorterat bort de sällsynta generna redan innan fröutvecklingen genom att sällsynta alleler blivit exponerade som homozygoter vid det naturliga urvalet.

References

Acheré, V., Faivre-Rampant, P., Jeandroz, S., Besnard, G., Markussen, T., Argones, A., Fladung, M., Ritter, E. & Favre, J.-M. (2004). A full saturated linkage map of *Picea abies* including AFLP, SSR, ESTP, 5S rDNA and morphological markers. *Theoretical Applied Genetic* 108, 1602-1613.

Ahmad, M. & Cashmore, A.R. (1993). Hy4 gene of A.thaliana encodes a protein with characteristics of a blue-light Photoreceptor. *Nature* 366(6451), 162-166.

Ahmad, M. & Cashmore, A.R. (1997). The blue-light receptor Cryptochrome 1 shows functional dependence on Phytochrome A or Phytochrome B in *Arabidopsis thaliana*. *Plant Journal* 11(3), 421-427.

Aho, M.L. (1994). Autumn frost hardening of one-year-old *Pinus sylvestris* (L.) seedlings - effect of origin and parent trees. *Scandinavian Journal of Forest Research* 9(1), 17-24.

Amos, W., Wilmer, J.W., Fullard, K., Burg, T.M., Croxall, J.P., Bloch, D. & Coulson, T. (2001). The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268(1480), 2021-2027.

Andersson, B. Establishment of Scots pine seed orchards for areas with harsh climate. In: Nilsson, J.E. (Ed.) *Proceedings of Reforestation material for harsh northern sites seminar*, Dept of forest genetics and plant physiology, February 19-20 1985. pp. 75-83: Swedish university of agricultural science.

Andersson, B., Elfving, B., Persson, T., Ericsson, T. & Kroon, J. (2007). Characteristics and development of improved *Pinus sylvestris* in northern Sweden. *Canadian Journal of Forest Research* 37(1), 84-92.

Andersson, B. & Fedorkov, A. (2004). Longitudinal differences in Scots pine frost hardiness. *Silvae Genetica* 53(2), 76-80.

Aparicio, J.M., Ortego, J. & Cordero, P.J. (2006). What should we weigh to estimate heterozygosity, alleles or loci? *Molecular Ecology* 15(14), 4659-4665.

Asakawa S, Sasaki S, Morikawa Y (1974) Growth of tree seedlings under the lights of different spectral compositions. *Journal of the Japanese Forestry Society* 56 (12), 441-447

Auckland, L., Bui, T., Zhou, Y., Shepherd, M. & Williams, C. (2002). Conifer microsatellite handbook. In: *Conifer microsatellite handbook*. p. iv + 57 pp. College Station USA: Texas A & M University System, Texas Agricultural Experiment Station, Department of Agricultural Economics.

Balloux, F., Amos, W. & Coulson, T. (2004). Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology* 13(10), 3021-3031.

Beall, F.D., Yeung, E.C. & Pharis, R.P. (1996). Far-red light stimulates internode elongation, cell division, cell elongation, and gibberellin levels in bean. *Canadian Journal of Botany-Revue Canadienne De Botanique* 74(5), 743-752.

Beck, E.H., Heim, R. & Hansen, J. (2004). Plant resistance to cold stress: Mechanisms and environmental signals triggering frost hardening and dehardening. *Journal of biosciences* 29(4), 449-459.

Berry, A. & Kreitman, M. (1993). Molecular analysis of an allozyme cline-alkohol-dehydrogenase in *Drosophila melanogaster* on the east-coast of North America. *Genetics* 134(3), 869-893.

Beuker, E. (1994). Long-term effects of temperature on the wood production of *Pinus sylvestris* L. and *Picea abies* (L) karst in old provenance experiments. *Scandinavian Journal of Forest Research* 9(1), 34-45.

Britten, H.B. (1996). Meta-analyses of the association between multilocus heterozygosity and fitness. *Evolution* 50(6), 2158-2164.

Brown, G.R., Bassoni, D.L., Gill, G.P., Fontana, J.R., Wheeler, N.C., Megraw, R.A., Davis, M.F., Sewell, M.M., Tuskan, G.A. & Neale, D.B. (2003). Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. *Genetics* 164(4), 1537-1546.

Brown, G.R., Kadel, E.E., 3rd, Bassoni, D.L., Kiehne, K.L., Temesgen, B., van Buijtenen, J.P., Sewell, M.M., Marshall, K.A. & Neale, D.B. (2001). Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. *Genetics* 159(2), 799-809.

Burgin, M.J., Casal, J.J., Whitelam, G.C. & Sanchez, R.A. (1999). A light-regulated pool of phytochrome and rudimentary high-irradiance responses under far-red light in *Pinus elliottii* and *Pseudotsuga menziesii*. *Journal of Experimental Botany* 50(335), 831-836.

Cannell, M.G.R. & Smith, R.I. (1986). Climatic warming, spring budburst and frost damage on trees. *Journal of Applied Ecology* 23(1), 177-191.

Chagne, D., Chaumeil, P., Ramboer, A., Collada, C., Guevara, A., Cervera, M.T., Vendramin, G.G., Garcia, V., Frigerio, J.M.M., Echt, C., Richardson, T. & Plomion, C. (2004). Cross-species transferability and mapping of genomic and cDNA SSRs in pines. *Theoretical and Applied Genetics* 109(6), 1204-1214.

Chapman, J.R., Nakagawa, S., Coltman, D.W., Slate, J. & Sheldon, B.C. (2009). A quantitative review of heterozygosity-fitness correlations in animal populations. *Molecular Ecology* 18(13), 2746-2765.

Charlesworth, D. & Charlesworth, B. (1979). Evolutionary genetics of sexual systems in flowering plants. *Proceedings of the Royal Society of London Series B-Biological Sciences* 205(1161), 513-530.

Chen, M., Chory, J. & Fankhauser, C. (2004) Light signal transduction in higher plants. *Annual Review of Genetics* 38:87-117.

Chen, M.M., Feng, F.J., Sui, X., Li, M.H., Zhao, D. & Han, S.J. Construction of a framework map for *Pinus koraiensis Sieb*. et Zucc. using SRAP, SSR and ISSR markers. *Trees-Structure and Function* 24(4), 685-693.

Clapham, D.H., Ekberg, I., Eriksson, G., Norell, L. & Vince-Prue, D. (2002). Requirement for far-red light to maintain secondary needle extension growth in northern but not southern populations of *Pinus sylvestris* (Scots pine). *Physiologia Plantarum* 114(2), 207-212.

Clapham, D.H., Dormling, I., Ekberg, I., Eriksson, G., Qamaruddin, M. & Vince-Prue, D. (1998) Latitudinal cline of requirement for far-red light for the photoperiodic control of budset and extension growth in *Picea abies* (Norway spruce). *Physiologia Plantarum* 102 (1),71-78

Coltman, D.W. & Slate, J. (2003). Microsatellite measures of inbreeding: A meta-analysis. *Evolution* 57(5), 971-983.

Conesa, A., Gotees, S., García-Gómez, J., Terol, J., Talon, M. & Robles, M. (2005). Blast2GO: A universal annotation and visualization tool in functional genomics research. Application note. *Bioinformatics* 21, 3674-3676.

Coulon, A. (2010). GENHET: an easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources* 10(1), 167-169.

David, P. (1998). Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* 80, 531-537.

Demura, T. & Fukuda, H. (2007). Transcriptional regulation in wood formation. *Trends in Plant Science* 12(2), 64-70.

Dermitzakis, E.T., Clark, A.G., Batargias, C., Magoulas, A. & Zouros, E. (1998). Negative covariance suggests mutation bias in a two-locus microsatellite system in the fish *Sparus aurata*. *Genetics* 150(4), 1567-1575.

Devey, M.E., Bell, J.C., Smith, D.N., Neale, D.B. & Moran, G.F. (1996). A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. *Theoretical and Applied Genetics* 92(6), 673-679.

Devey, M.E., Carson, S.D., Nolan, M.F., Matheson, A.C., Riini, C.T. & Hohepa, J. (2004). QTL associations for density and diameter in *Pinus radiata* and the potential for marker-aided selection. *Theoretical and Applied Genetics* 108(3), 516-524.

Doyle, J.J. & Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus* 12(1), 13-15.

Durzan D., Campbell R. & Wilson A. (1979) Inhibition of female cone production in white spruce by red light treatment during night under field conditions. *Environ Exp Bot* 19 (3), 133-135.

Dvornyk, V., Sirvio, A., Mikkonen, M. & Savolainen, O. (2002). Low nucleotide diversity at the pal1 locus in the widely distributed *Pinus sylvestris*. *Molecular Biology and Evolution* 19(2), 179-188.

Eckert, A.J., Pande, B., Ersoz, E.S., Wright, M.H., Rashbrook, V.K., Nicolet, C.M. & Neale, D.B. (2009). High-throughput genotyping and mapping of single nucleotide polymorphisms in Loblolly pine (P*inus taeda* L.). *Tree Genetics & Genomes* 5(1), 225-234.

El-Assal, S.E.D., Alonso-Blanco, C., Peeters, A.J.M., Raz, V. & Koornneef, M. (2001). A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. *Nature Genetics* 29(4), 435-440.

Elsik, C.G., Minihan, V.T., Hall, S.E., Scarpa, A.M. & Williams, C.G. (2000). Low-copy microsatellite markers for *Pinus taeda* L. *Genome* 43(3), 550-555.

Emebiri, L.C., Devey, M.E., Matheson, A.C. & Slee, M.U. (1998). Age-related changes in the expression of QTLs for growth in Radiata pine seedlings. *Theoretical and Applied Genetics* 97(7), 1053-1061.

Engelmark, O. (1984). Forest fires in the Muddus National Park (northern Sweden) during the past 600 years. *Canadian Journal of Botany* 62(5), 893-898.

Falconer, D.S. & Mackay, T.F.C. (1996). *Introduction to Quantitative Genetics*.4th ed., Malaysia: Prentice-Hall.

FAO (2003). Molecular marker assisted selection as a potential tool for genetic improvement of crops, forest trees, livestock and fish in developing countries. In: *10:th conference of FAO Biotechnology Forum.* Internet.

FAO (2006). Global forest resources assessment 2005 In: *Progress towards sustainable forest management.* Rome: FAO. pp. 1-320.

Fernbach, E. & Mohr, H. (1990). Coaction of blue ultraviolet - Alight and light absorbed by phytochrome in controlling growth of pine (*Pinus sylvestris* L.) seedlings. *Planta* 180(2), 212-216.

Fisher, P., Richardson, T. & Gardner, R. (1998). Characteristics of single- and multi-copy microsatellites from *Pinus radiate*. *Theoretical Applied Genetic* 96, 969-979.

Fisher, R.A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* 52, 399-433.

Flint, J. & Mott, R. (2001). Finding the molecular basis of quantitative traits: successes and pitfalls. *Nat Rev Genet* 2(6), 437-45.

Floran, V., Ganea, S., Sestras, R. & Gil, M.R.G. Genetic variability in populations of Scots pine from Romania and Sweden. *Bulletin of University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca. Horticulture* 67(1), 481.

Folta, K.M., Pontin, M.A., Karlin-Neumann, G., Bottini, R. & Spalding, E.P. (2003). Genomic and physiological studies of early cryptochrome 1 action demonstrate roles for auxin and gibberellin in the control of hypocotyl growth by blue light. *Plant Journal* 36(2), 203-214.

Fournier, N., Rigling, A., Dobbertin, M. & Gugerli, F. (2006). Random amplified polymorphic DNA (RAPD) patterns show weak genetic differentiation between low- and high-elevation types of Scots pine (*Pinus sylvestris* L.) in dry continental valleys in the Alps. *Annals of Forest Science* 63(4), 431-439.

Franklin K.A & Quail P.H. (2010) Phytochrome functions in *Arabidopsis* development. *Journal of experimental Botany* 61, 11-24.

Franklin, E.C. (1970). Survey of mutant forms and inbreeding depression in species of the family Pinaceae. *USDA Forest Service: Research paper*. S.E.-61, 1-21.

García-Gil, M.R., Francois, O., Kamruzzahan, S. & Waldmann, P. (2009). Joint analysis of spatial genetic structure and inbreeding in a managed population of Scots pine. *Heredity* 103(1), 90-96.

Garcia-Gil, M.R., Mikkonen, M. & Savolainen, O. (2003). Nucleotide diversity at two phytochrome loci along a latitudinal cline in Pinus sylvestris. *Molecular Ecology* 12(5), 1195-1206.

Giertych, M. & Mátyás, C. (Eds.) (1991). *Genetics of Scots Pine*. Amsterdam: Elsevier Science Publisher. (Developments in Plant Genetics and Breeding; 3).

Gilmour, A., Gogel, B., Cullis, B. & Thompson, R. (2006). ASReml user guide release 2.0. *VSN International Ltd, Hemel Hempstead, UK*, 320.

Goddard, M.E. & Hayes, B.J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews Genetics* 10(6), 381-391.

Goldstein, D. & Schlötterer, C. (1999). *Microsatellites: evolution and applications (POD)* Oxford: Oxford University Press.

Gonzalez-Martinez, S.C., Wheeler, N.C., Ersoz, E., Nelson, C.D. & Neale, D.B. (2007). Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* 175(1), 399-409.

Griffin, A.R. & Lindgren, D. (1985). Effect of inbreeding on production of filled seed in *Pinus radiata* - experimental results and a model of gene-action. *Theoretical and Applied Genetics* 71(2), 334-343.

Groover, A., Devey, M., Fiddler, T., Lee, J., Megraw, R., Mitchelolds, T., Sherman, B., Vujcic, S., Williams, C. & Neale, D. (1994). Identification of quantitative trait loci influencing wood specific-gravity in an outbred pedigree of Loblolly pine. *Genetics* 138(4), 1293-1300.

Gwaze, D.P., Zhou, Y., Reyes-Valdes, M.H., Al-Rababah, M.A. & Williams, C.G. (2003). Haplotypic QTL mapping in an outbred pedigree. *Genetical Research* 81(1), 43-50.

Haapanen, M. (2001). Time trends in genetic parameter estimates and selection efficiency for Scots pine in relation to field testing method. *Forest genetics* 8(2), 129-144.

Harju, A.M. & Nikkanen, T. (1996). Reproductive success of orchard and nonorchard pollens during different stages of pollen shedding in a Scots pine seed orchard. *Canadian Journal of Forest Research* 26(6), 1096-1102.

Hoddinott J. & Scott R. (1996) The influence of light quality and carbon dioxide enrichment on the growth and physiology of seedlings of three conifer species .1. Growth responses. *Canadian Journal of Botany* 74 (3), 383-390.

Houghton, J.T., Ding, Y., ., D.J.G., Noguer, M., Linden, P.J.v.d. & Xiaosu, D. (2001). *Climate Change 2001: The Scientific Basis*.

Hudson, R.R., Bailey, K., Skarecky, D., Kwiatowski, J. & Ayala, F.J. (1994). Evidence for positive selection in the superoxide-dismutase (SOD) region of *Drosophila melanogaster*. *Genetics* 136(4), 1329-1340.

Hurme, P., Repo, T., Savolainen, O. & Paakkonen, T. (1997). Climatic adaptation of bud set and frost hardiness in Scots pine (*Pinus sylvestris*). *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere* 27(5), 716-723.

IPCC (2007). *Climate change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Iverson, L.R. & Prasad, A.M. (2002). Potential redistribution of tree species habitat under five climate change scenarios in the eastern US. *Forest ecology and managment* 155, 205-222.

Jansson, G., Li, B.L. & Hannrup, B. (2003). Time trends in genetic parameters for height and optimal age for parental selection in Scots pine. *Forest Science* 49(5), 696-705.

Jarne, P. & Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* 11(10), 424-429.

Jarvis, P. & Leverenz, J. (1983). Productivity of temperate, decidous, and evergreen forests. In: Lange, O*., et al.* (Eds.) *Physiological Plant Ecology, IV. Ecosystem Processes: Mineral Cycling, Productivity and Man's Influence*. pp. 223-280. New York: Springer.

Kami C., Lorrain S., Hornitschek P. & Fankhauser C. (2010) Light-regulated plant growth and development *Current Topic in Developmental Biology* 91, 29-66.

Karhu, A., Hurme, P., Karjalainen, M., Karvonen, P., Karkkainen, K., Neale, D. & Savolainen, O. (1996). Do molecular markers reflect patterns of differentiation in adaptive traits of conifers? *Theoretical and Applied Genetics* 93(1-2), 215-221.

Kärkkäinen, K. & Savolainen, O. (1993). The degree of early inbreeding depression determines the selfing rate at the seed stage - model and results from *Pinus sylvestris* (Scots pine). *Heredity* 71, 160-166.

Karlsson, B. & Rosvall, O. Breeding programs in Sweden–Norway spruce. In: Lee S.J. (ed) (1993) Progeny testing and breeding strategies. *Procedings of the Nordic Group of Tree Breeding*. Edinburgh: Forestry Commission .

Kaya, Z., Sewell, M.M. & Neale, D.B. (1999). Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics* 98(3-4), 586-592.

Kellomaki, S., Rouvinen, I., Peltola, H., Strandman, H. & Steinbrecher, R. (2001). Impact of global warming on the tree species composition of boreal forests in Finland and effects on emissions of isoprenoids. *Global Change Biology* 7(5), 531-544.

Kendrick R.E. & Kronenberg G.H.M. (1994). *Photomorphogenesis in plants*, 2nd ed., Dordrecht: Kluwer academic publisher.

Kinloch, B.B., Westfall, R.D. & Forrest, G.I. (1986). Caledonian Scots pine - origins and genetic-structure. *New Phytologist* 104(4), 703-729.

Komulainen, P., Brown, G.R., Mikkonen, M., Karhu, A., Garcia-Gil, M.R., O'Malley, D., Lee, B., Neale, D.B. & Savolainen, O. (2003). Comparing EST-based genetic maps between *Pinus sylvestris* and *Pinus taeda*. *Theoretical and Applied Genetics* 107(4), 667-678.

Koornneef, M., Rolff, E. & Spruit, C.J.P. (1980). Genetic-control of light-inhibited hypocotyl elongation in *Arabidopsis thaliana* (L) Heynh. *Zeitschrift Fur Pflanzenphysiologie* 100(2), 147-160.

Korshikov, II, Kalafat, L.A., Pirko, Y.V. & Velicoridko, T.I. (2005). Population-genetic variation in Scots pine P*inus sylvestlis* L. from the main forest regions of Ukraine. *Russian Journal of Genetics* 41(2), 155-166.

Koski, V. (1971). Embryonic lethals of *Picea abies* and Pinus sylvestris. *Communicationes Instituti Forestalis Fenniae* 75(3), 30 pp.

Koski, V. & Sievanen, R. (1985). Timing of growth cessation in relation to the variations in the growing season. *Crop physiology of forest trees. Proceedings of an international conference on managing forest trees as cultivated plants, held in Finland, 23-28 July 1984*, 167-193.

Krieger, M.J.B. & Ross, K.G. (2002). Identification of a major gene regulating complex social behavior. *Science* 295(5553), 328-332.

Kroon, J., Ericsson, T., Jansson, G. & Andersson, B. (2011). Patterns of genetic parameters for height in field genetic tests of *Picea abies* and *Pinus sylvestris* changes with age. *Tree Genetics and Genome.* DOI: 10.1007/s11295-011-0398-y

Kumar, S., Echt, C., Wilcox, P.L. & Richardson, T.E. (2004). Testing for linkage disequilibrium in the New Zealand Radiata pine breeding population. *Theoretical and applied genetics.* 108(2), 292-8.

Kumar, S., Spelman, R.J., Garrick, D.J., Richardson, T.E., Lausberg, M. & Wilcox, P.L. (2000). Multiple-marker mapping of wood density loci in an outbred pedigree of Radiata pine. *Theoretical and Applied Genetics* 100(6), 926-933.

Kvaalen, H. & Appelgren, M. (1999). Light quality influences germination, root growth and hypocotyl elongation in somatic embryos but not in seedlings of Norway spruce. *In Vitro Cellular & Developmental Biology-Plant* 35(6), 437-441.

Lerceteau, E., Plomion, C. & Andersson, B. (2000). AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. *Molecular Breeding* 6(5), 451-458.

Lerceteau, E., Szmidt, A.E. & Andersson, B. (2001). Detection of quantitative trait loci in Pinus sylvestris L. across years. *Euphytica* 121(2), 117-122.

Liewlaksaneeyanawin, C., Ritland, C.E., El-Kassaby, Y.A. & Ritland, K. (2004). Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theoretical and Applied Genetics* 109(2), 361-369.

Lin, C., Ahmad, M., Gordon, D. & Cashmore, A.R. (1995). Expression of an arabidopsis Cryptochrome gene in transgenic tobacco results in hypersensitivity to blue, UVA, and green light. *Proceedings of the National Academy of Sciences of the United States of America* 92(18), 8423-8427.

Lin, C.T., Ahmad, M. & Cashmore, A.R. (1996). Arabidopsis Cryptochrome 1 is a soluble protein mediating blue light-dependent regulation of plant growth and development. *Plant Journal* 10(5), 893-902.

Lindgren, K. & Nilsson, J.E. (1992). Cold acclimation of *Pinus sylvestris* and *Pinus contorta* provenances as measured by freezing tolerance of detached needles. *Scandinavian Journal of Forest Research* 7(3), 309-315.

Linkosalo, T. & Lechowicz, M.J. (2006). Twilight far-red treatment advances leaf bud burst of silver birch (*Betula penduld*). *Tree Physiology* 26(10), 1249-1256.

Lund, M.S., Sorensen, P., Madsen, P. & Jaffrezic, F. (2008). Detection and modelling of time-dependent QTL in animal populations. *Genetics Selection Evolution* 40(2), 177-194.

Ma, C.X., Lin, M., Littell, R.C., Yin, T. & Wu, R. (2004). A likelihood approach for mapping growth trajectories using dominant markers in a phase-unknown full-sib family. *Theoretical and Applied Genetics* 108(4), 699-705.

Ma, L.G., Li, J.M., Qu, L.J., Hager, J., Chen, Z.L., Zhao, H.Y. & Deng, X.W. (2001). Light control of Arabidopsis development entails coordinated regulation of genome expression and cellular pathways. *Plant Cell* 13(12), 2589-2607.

Macgregor, S., Knott, S.A., White, I. & Visscher, P.M. (2005). Quantitative trait locus analysis of longitudinal quantitative trait data in complex pedigrees. *Genetics* 171(3), 1365-1376.

Mardis, E.R. (2008). Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics* 9, 387-402.

Markert, C.L. & Moller, F. (1959). Multiple forms of enzymes - tissue, ontogenetic, and species specific patterns. *Proceedings of the National Academy of Sciences of the United States of America* 45(5), 753-763.

Markussen, T., Fladung, M., Achere, V., Favre, J.M., Faivre-Rampant, P., Aragones, A., Perez, D.D., Harvengt, L., Espinel, S. & Ritter, E. (2003). Identification of QTLs controlling growth,

chemical and physical wood property traits in Pinus pinaster (Ait.). *Silvae Genetica* 52(1), 8-15.

Muona, O. & Harju, A. (1989). Effective population sizes, genetic-variability, and mating system in natural stands and seed orchards of *Pinus sylvestris*. *Silvae Genetica* 38(5-6), 221-228.

Muona, O., Yazdani, R. & Rudin, D. (1987). Genetic change between life stages in *Pinus sylvestris* - allozyme variation in seeds and planted seedlings. *Silvae Genetica* 36(1), 39-42.

Naugzemys, D., Zvingila, D., Aucina, A. & Rancelis, V. (2006). Comparison of DNA polymorphism in seedlings of *Pinus sylvestris* L. from different populations by RAPD markers. *Biologija* (1), 30-35.

Neale, D.B. (2007). Genomics to tree breeding and forest health. *Current Opinion in Genetics & Development* 17(6), 539-544.

Neale, D.B. & Savolainen, O. (2004). Association genetics of complex traits in conifers. *Trends in Plant Science* 9(7), 325-330.

Neale, D.B., Sewell, M.M. & Brown, G.R. (2002). Molecular dissection of the quantitative inheritance of wood property traits in Loblolly pine. *Annals of Forest Science* 59(5-6), 595-605.

Neale, D.B. & Williams, C.G. (1991). Restriction-fragment-length-polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Canadian Journal of Forest Research* 21(5), 545-554.

Nilsson, J.E. (1990). Variation in early winter hardening within families of *Pinus sylvestris* L. from northern Sweden. *Scandinavian Journal of Forest Research* 5(1), 17-28.

Nilsson, J.E. (2001). Seasonal changes in phenological traits and cold hardiness of F1-populations from plus-trees of *Pinus sylvestris* and *Pinus contorta* of various geographical origins. *Scandinavian Journal of Forest Research* 16(1), 7-20.

Nilsson, J.E. & Andersson, B. (1987). Performance in freezing tests and field experiments of full-sib families of *Pinus sylvestris* (L.). *Canadian Journal of Forest Research* 17(11), 1340-1347.

Nilsson, J.E., Andersson, B. & Walfridsson, E.A. (1991). Progeny freeze testing progeny field testing and parental phenology of *Pinus sylvestris* L. clones in northern Sweden. *Scandinavian Journal of Forest Research* 6(2), 177-196.

Nilsson, J.E. & Eriksson, G. (1986). Freeze testing and field mortality of *Pinus sylvestris* in northernn Sweden. *Scandinavian Journal of Forest Research* 1(2), 205-218.

Nilsson, J.E. & Walfridsson, E.A. (1990). Autumn freeze testing of one year reciprocal families of *Pinus sylvestris* (L.). *Scandinavian Journal of Forest Research* 5(4), 501-512.

Nilsson, J.E. & Walfridsson, E.A. (1995). Phenology variation among plus-tree clones of *Pinus sylvestris*(L) in northern Sweden. *Silvae Genetica* 44(1), 20-28.

Notivol, E., Garcia-Gil, M.R., Alia, R. & Savolainen, O. (2007). Genetic variation of growth rhythm traits in the limits of a latitudinal cline in Scots pine. *Canadian Journal of Forest Research* 37(3), 540-551.

Persson, T., Andersson, B. & Ericsson, T. (2010). Relationship between autumn cold hardiness and field performance in northern *Pinus sylvestris*. *Silva Fennica* 44(2), 255-266.

Pot, D., Rodrigues, J.C., Rozenberg, P., Chantre, G., Tibbits, J., Cahalan, C., Pichavant, F. & Plomion, C. (2006). QTLs and candidate genes for wood properties in Maritime pine (*Pinus pinaster* Ait.). *Tree Genetics & Genomes* 2(1), 10-24.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. & Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2(3), 225-238.

Prescher, F., Lindgren, D., Wennström, U., Almqvist, C., Ruotsalainen, S. & Kroon, J. Seed production in Scots pine seed orchards. In: Fedorkov, A. (Ed.) *Proceedings of Status, monitoring and targets for breeding programs. Proceedings of the meeting of Nordic forest tree breeders and forest geneticistis*, Syktyvkar2005. pp. 65-72.

Pyhajarvi, T., Garcia-Gil, M.R., Knurr, T., Mikkonen, M., Wachowiak, W. & Savolainen, O. (2007). Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics* 177(3), 1713-1724.

Queller, D.C., Strassmann, J.E. & Hughes, C.R. (1993). Microsatellites and Kinship. *Trends in Ecology & Evolution* 8(8), 285-&.

R Development Core Team, R. (2006). R: A language and environment for statistical computing. *R Foundation forStatistical C om puting, V ienna, A ustria, ISB*, 07-0.

Richardson, B.A., Ekramoddoulah, A.K.M., Liu, J.J., Kim, M.S. & Klopfenstein, N.B. (2010). Current and future molecular approaches to investigate the white pine blister rust pathosystem. *Forest Pathology* 40(3-4), 314-331.

Romano, C.P., Robson, P.R.H., Smith, H., Estelle, M. & Klee, H. (1995). Transgene-mediated auxin overproduction in Arabidopsis - hypocotyl elongation phenotype and interactions with the Hy6-1 hypocotyl elongation and Axr1 auxin-resistant mutants. *Plant Molecular Biology* 27(6), 1071-1083.

Rosvall, O., Jansson, G., Andersson, B., Ericsson, T., Karlsson, B., Sonesson, J. & Stener, L.G. (2001). Genetic gain from present and future seed orchards and clone mixes. *Redogorelse - SkogForsk* (1), 41 pp.

Rousset, F. (2008). GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8(1), 103-106.

Rudin, D. & Ekberg, I. (1978). Linkage studies in *Pinus sylvestris* L. - using macro gametophyte allozymes. *Silvae Genetica* 27(1), 1-12.

Rudin, D., Eriksson, G., Ekberg, I. & Rasmuson, M. (1974). Studies of allele frequencies and inbreeding in Scots pine populations by the aid of the isozyme technique. *Silvae Genetica* 23(1/3), 10-13.

Ryden, P., Andersson, H., Landfors, M., Naslund, L., Hartmanova, B., Noppa, L. & Sjostedt, A. (2006). Evaluation of microarray data normalization procedures using spike-in experiments. *BMC Bioinformatics* 7.

Sarala M., Taulavuori E., Karhu J., Savonen E.M., Laine K., Kubin E. & Taulavuori K. (2009) Improved elongation of Scots pine seedlings under blue light depletion is not dependent on resource acquisition. *Functional Plant Biology* 36 (8), 742-751.

Sarvas, R. (1962). Investigations on the flowering and seed crop of Pinus silvestris. *Commun Inst Forest Fenn* 53((4)), 1-198.

Savolainen, O. & Hedrick, P. (1995). Heterozygosity and fitness - no association in Scots pine. *Genetics* 140(2), 755-766.

Savolainena, O., Bokmaa, F., Garcı´a-Gila, R., Komulainena, P.i. & Repo, T. (2004). Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of *Pinus sylvestris* to climatic changes. *Forest ecology and managment* In press.

Sax, K. (1923). Association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8, 552-560.

Scandali, J. (1969). Genetic control of multiple molecular forms of enzymes in plants - a review. *Biochemical Genetics* 3(1), 37-&.

Schadt, E., Monks, S., Drake, T., Lusis, A., Che, N., Colinayo, V., Ruff, T., Milligan, S., Lamb, J., Cavet, G., Linsley, P., Mao, M., Stoughton, R. & Friend, S. (2003). Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422(6929), 297-302.

Scharff, O. (1962). Effects of red and far-red light on hypocotyl of *Picea abies*. *Physiologia Plantarum* 15(4), 804-&.

Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. (1995). Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. *Science* 270(5235), 467-470.

Sewell, M.M., Bassoni, D.L., Megraw, R.A., Wheeler, N.C. & Neale, D.B. (2000). Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. *Theoretical and Applied Genetics* 101(8), 1273-1281.

Sewell, M.M., Davis, M.F., Tuskan, G.A., Wheeler, N.C., Elam, C.C., Bassoni, D.L. & Neale, D.B. (2002). Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). II. Chemical wood properties. *Theor Appl Genet* 104(2-3), 214-222.

Shen, H.H., Rudin, D. & Lindgren, D. (1981). Study of the pollination pattern in a Scots pine seed orchard by means of isoenzyme analysis. *Silvae Genetica* 30(1), 7-15.

Sjödin, A., Bylesjö, M., Skogström, O., Eriksson, D., Nilsson, P., Rydén, P., Jansson, S. & Karlsson, J. (2006). UPSC-BASE - Populus transcriptomics online. *Plant Journal* 48(5), 806-817.

Skogsstyrelsen (2008). Skogsstatistisk årsbok 2008. In: Skogsstyrelsen (Ed.). Jönköping: Loman, JO. pp. 1-337.

Slate, J., David, P., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E. & McEwan, J.C. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity* 93(3), 255-265.

Smouse, P.E. & Robledo-Arnuncio, J.J. (2005). Measuring the genetic structure of the pollen pool as the probability of paternal identity. *Heredity* 94(6), 640-649.

Smyth, G.K. (2005). Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solution Using R and Bioconductor*, 397-420.

Smyth, G.K. & Speed, T. (2003). Normalization of cDNA microarray data. *Methods* 31(4), 265-273.

Soranzo, N., Provan, J. & Powell, W. (1998). Characterization of microsatellite loci in *Pinus sylvestris* L. *Mol Ecol* 7(9), 1260-1.

Sorensen, F. (1969). Embryonic genetic load in coastal Douglas fir, *Pseudotsuga menziesii* var . Menziesii. *American Naturalist* 103(932), 389-&.

Sundblad, L.G. & Andersson, B. (1995). No difference in frost hardiness between high and low-altitude *Pinus sylvestris* (L) offspring. *Scandinavian Journal of Forest Research* 10(1), 22-26.

Szulkin, M., Bierne, N. & David, P. (2010). Heterozygosity-fitness correlations: a time for reappraisal. *Evolution* 64(5), 1202-1217.

Torimaru, T., Wang, X.R., Fries, A., Andersson, B. & Lindgren, D. (2009). Evaluation of pollen contamination in an advanced Scots pine seed orchard. *Silvae Genetica* 58(5-6), 262-269.

Ukrainetz, N.K., Ritland, K. & Mansfield, S.D. (2008). Identification of quantitative trait loci for wood quality and growth across eight full-sib coastal Douglas-fir families. *Tree Genetics & Genomes* 4(2), 159-170.

Vandenbussche, F., Verbelen, J.P. & Van der Straeten, D. (2005). Of light and length: regulation of hypocotyl growth in Arabidopsis. *Bioessays* 27(3), 275-284.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP - a new technique for DNA-fingerprinting. *Nucleic Acids Research* 23(21), 4407-4414.

Wakamiya, I., Newton, R.J., Johnston, J.S. & Price, H.J. (1993). Genome size and environmental-factors in the genus *pinus*. *American Journal of Botany* 80(11), 1235-1241.

Warrington I.J., Rook D.A., Morgan D.C. & Turnbull H.L. (1989) The influence of simulated shadelight and daylight on growth, development and photosynthesis of *Pinus radiata*, *Agathis australis* and *Dacrydium cupressinum*. *Plant Cell Environ* 12 (4), 343-356.

Wattier, R., Engel, C.R., Saumitou-Laprade, P. & Valero, M. (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Molecular Ecology* 7(11), 1569-1573.

Wehenkel, C., Bergmann, F. & Gregorius, H.R. (2007). Genotype-species interactions in neighbourhoods of forest tree communities. *Silvae Genetica* 56(3-4), 101-110.

Williams, C.G., Reyes-Valdes, M.H. & Huber, D.A. (2007). Validating a QTL region characterized by multiple haplotypes. *Theoretical and Applied Genetics* 116, 87-94.

Williams, C.G. & Savolainen, O. (1996). Inbreeding depression in conifers: Implications for breeding strategy. *Forest Science* 42(1), 102-117.

Wu, R.L. & Lin, M. (2006). Opinion - Functional mapping - how to map and study the genetic architecture of dynamic complex traits. *Nature Reviews Genetics* 7(3), 229-237.

Yazdani, R., Muona, O., Rudin, D. & Szmidt, A.E. (1985). Genetic structure of a *Pinus sylvestris* L. seed-tree stand and naturally regenerated understory. *Forest Science* 31(2), 430-436.

Yazdani, R., Nilsson, J.E., Plomion, C. & Mathur, G. (2003). Marker trait association for autumn cold acclimation and growth rhythm in *Pinus sylvestris*. *Scandinavian Journal of Forest Research* 18(1), 29-38.

Yazdani, R., Yeh, F.C. & Rimsha, J. (1995). Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. *Forest genetics* 2(2), 109-116.

Yin, T.-M., Wang, X.-R., Andersson, B. & Lerceteau-Köhler, E. (2003). Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. *Theoretical and Applied Genetics* 106(6), 1075-1083

Zhelev, P., Ekberg, I., Eriksson, G. & Norell, L. (2003). Genotype environment interactions in four full-sib progeny trials of *Pinus sylvestris* (L.) with varying site indices. *Forest genetics* 10(2), 93-102.

Acknowledgments

Sari, my main supervisor, thanks for leading me into the world of plants. You have given me time and guidance under this long but wonderful journey. The knowledge that you have given me is priceless. I have enjoyed our long talks both about science and families; solving how to keep all balls in the air without losing track of what is important.

Tack till mina andra handledare. Bengt för att du har hållit ett vakande öga på tiden och sedan sett till att jag faktiskt kommer i mål. Jan-Erik för att du tagit dig tid att svara på mina frågor i tid och otid. Rishi…

All the best to all that has passed through Saris group. You all know who you are. It has been a privilege to work and learn from you. Sonali, thanks for reading the thesis over and over again.

Jag vill även passa på att tacka all i forskarskolan. Maria, Elli, Jon, Jenny, Henrik, Mats, David, David, Johan och Lasse det har varit otroligt roligt att ha er som resesällskap, inte bara på de otaliga resorna som vi gjort utan också resan till slutet av avhandlingarna. Även ett stort tack till Jan-Erik, Bengt och Gunnar som har lyckats organisera det hela och ta oss från början och nästan till slutet (Maria och David, tack för att ni tar upp kön, så slipper jag).

Ett rungande tack till all administrativ personal på UPSC. Maria L, Inga-Lis, Kjell, Inga-Britt, Siv och Ingela. Utan er skulle UPSC inte hållas flytande. Tack Stefan och Kjell som utan att knota ryckt in när den där j… datorn inte gör som man tänkt.

I would also like to thank all fellow students that have given me support and a good talk over coffee. All work and no play do not keep a human sane.

Tack Erik N, min företagsmentor på Holmen skog, för att du har gett mig en möjlighet att se den “verkliga” världen utanför universitetet. Det har varit mycket lärorikt att få vara med som en i gänget. Jag har uppskattat all den tid

jag har varit på Holmen skog. Tack även alla trevliga medarbetare från stab skogsvård. Jag kommer sakna er, speciellt Helena, Bror och Yvonne.

Tjejgänget som har hängt med i vått och tort; Jenny, Maria och Kicki. Tack för alla långa kvällar med te och vin. Synd att vi ska bo så långt ifrån varandra men det är skönt att veta att det alltid finns någon där som ställer upp när man som mest behöver det.

Mamma och pappa som fått stå ut genom åren med långa konversationer om både det ena och det tredje. Sedan att ni under de senaste åren ständigt ställt upp och löst både barn, hund och bil problem har varit ovärderligt. Mormor och Hasse, tack för att ni alltid har funnits där. Hasse trots att det varit otroligt irriterande vid tillfällena har det varit ovärderligt att du hela tiden tvingat mig att argumentera för min sak.

Sist men absolut inte minst min underbara man Jonas som har lyckats jonglera med otaliga bollar och tagit största delen av det mesta under tiden som jag har försökt att komma till ett slut. Jag ska försöka att återgälda det om bara… Isak och David jag älskar er.

This thesis was supported by the Knowledge Foundation, the Kempe Foundation and The Swedish Tree Breeding Association through the Research School in Forest Genetics and Breeding at The Swedish University of Agricultural Sciences and also Gunnar and Lillian Nicholson Exchange Fund, which is gratefully acknowledged

September 2011.