Different alleles of a gene encoding leucoanthocyanidin reductase (*PaLAR3*) influence resistance against the fungus *Heterobasidion parviporum* in *Picea abies*

Miguel Nemesio-Gorriz[1], Almuth Hammerbacher[2], Katarina Ihrmark[1], Thomas Källman[3, 4], Åke Olson[1], Martin Lascoux[3], Jan Stenlid[1], Jonathan Gershenzon[5] and Malin Elfstrand[1]

Author affiliation

[1] Department of Forest Mycology and Plant Pathology, Swedish University of Agricultural Sciences, Uppsala, Sweden

[2] Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa

[3] Department of Ecology and Genetics, Evolutionary Biology Centre, Science for Life Laboratory, Uppsala University, 75236 Uppsala, Sweden

[4] BILS, Science for Life Laboratory, Uppsala, Sweden

[5]Department of Biochemistry, Max Planck Institute for Chemical Ecology, Jena, Germany

One Sentence Summary:

The gene *PaLAR3* involved in phenolic secondary metabolism is validated as a marker for pathogen resistance in Norway spruce through the study of two *PaLAR3* alleles and their differential effect on resistance

Author contributions

Miguel Nemesio-Gorriz (MNG), Malin Elfstrand (ME) and Jan Stenlid (JS) conceived the study. MNG designed, preformed and analyzed: the phenotyping experiments, the *in vitro* experiment to determine the effect of catechin on *H.*

parviporum growth, the expression analyses and the cloning of *PaLAR3A* and *B* genomic sequences with assistance from Katarina Ihrmark (KI) and ME. Almuth Hammerbacher (AH) performed the chemical analysis of the bark samples and the *in vitro* experiment for protein activity, she wrote the corresponding parts of the manuscript. Thomas Källman and Martin Lascoux genotyped the populations used in this study on a custom made SNP chip. MNG drafted the manuscript and coordinated the writing. ME, AH, Åke Olson and Jonathan Gershenzon contributed with advice on data analyses and advice during the process of writing the manuscript. MNG wrote the final manuscript and all co-authors read and approved the final version of the manuscript

Financial support of the project

Financial support was received from SSF (grant nr R8b08-0011), and by the Swedish Research Council FORMAS, grant nrs 2012-1276 and 217-2007-433, the European Community's Sixth Framework Programme, under the Network of Excellence Evoltree (www.evoltree.eu/) and by the Seventh Framework Programme (FP7/2007-2013) under the grant agreement nu 211868 (Project Noveltree, www.noveltree.eu/), Support also came from the Biodiversa projects Linktree (www.igv.fi.cnr.it/linktree/) and TipTree, and The Max Planck Society

Keywords

Norway spruce, *Picea abies, Heterobasidion,* catechin, leucoanthocyanidin reductase (LAR), flavonoid, resistance, allele, QTL

Abstract

Despite the fact that fungal diseases are a growing menace for conifers in modern silviculture, only a very limited number of molecular markers for pathogen resistance have been validated in conifer species. A previous genetic study indicated that the resistance of Norway spruce (*Picea abies*) to *Heterobasidion annosum s.l.*, a pathogenic basidiomycete species complex, is linked to a QTL that associates with differences in fungal growth in sapwood (FGS) that includes a gene, *PaLAR3*, which encodes a leucoanthocyanidin reductase. In this study, gene sequences showed the presence of two *PaLAR3* allelic lineages in *P. abies.* Higher resistance was associated with the novel allele, which was found in low frequency in the four *P. abies* populations that we studied. Norway spruce plants carrying at least one copy of the novel allele showed a significant reduction in FGS after inoculation with *Heterobasidion parviporum* compared to their half-siblings carrying no copies, indicating dominance of this allele. The amount of (+) catechin, the enzymatic product of PaLAR3, was significantly higher in bark of trees homozygous for the novel allele. Although we observed that the *in vitro* activities of the enzymes encoded by the two alleles were similar, we could show that allele-specific transcript levels were significantly higher for the novel allele, indicating that regulation of gene expression is responsible for the observed effects in resistance, possibly caused

by differences in *cis*-acting elements that we observe in the promoter region of the two alleles.

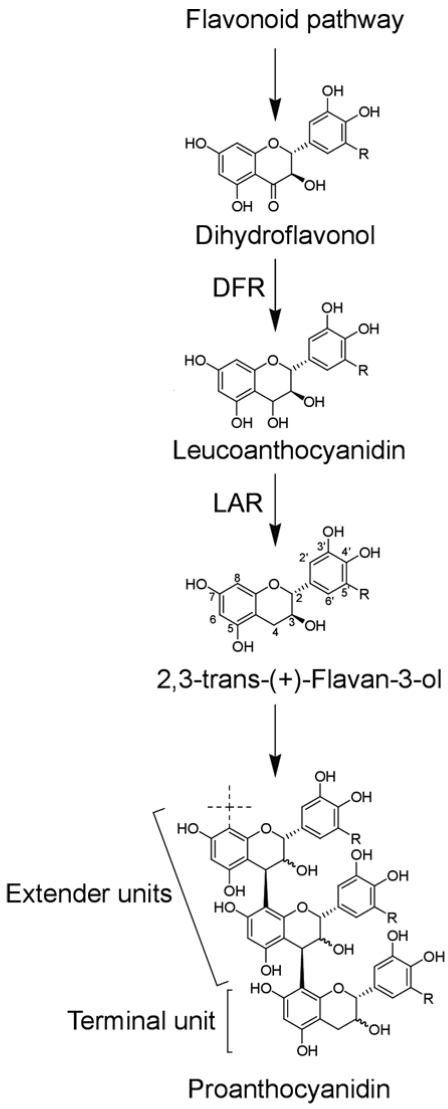
Introduction

Conifers are a major component of boreal and alpine forests in the Northern Hemisphere and are of great importance both ecologically and economically. Norway spruce [*Picea abies* (L.) Karst], in particular, is one of the most important conifer species in Europe and among the most widely planted spruce species, both in and outside of its native range. Infection of *P. abies* by the stem and root-rot fungi from the *Heterobasidion annosum sensu lato* (s.l.) species complex, one of the most significant forest pathogens in the Northern Hemisphere, has a strong negative impact on wood production and quality (Woodward, 1998). The wood decay associated with the spread of *H. annosum* s.l. in the tree reduces high quality saw logs to fire wood (Karlsson and Swedjemark 2006). Following spread from roots or direct infection via wound at stem base, *H. annosum* s.l. can establish a decay column in the wood that may reach up to 10 m in the stem. To achieve this the pathogen must penetrate the bark and necrotize phloem, cambium and sapwood (Oliva et al., 2015).

Resistance to *H. annosum* s.l. in *P. abies* is a quantitative trait (Swedjemark et al., 1997; Arnerup et al., 2010) and it is likely to involve an active defense response against colonization and mechanisms associated with protection of heart- and sapwood. In the first Quantitative Trait Loci (QTL) study on *P. abies* resistance to *H. parviporum,* a member of the *H. annosum s.l.* species complex, Lind et al. (2014) identified 13 QTLs that were associated with four traits related to host resistance. Four of the QTL regions were associated with the control of fungal

growth in sapwood (FGS), which is a trait that reflects the trees capacity to protect its sapwood. In theory trees allowing shorter FGS after inoculation with *H. annosum* s.l. in the sapwood would also display shorter decay columns after natural infections (Woodward, 1998). One of the four QTL associated with FGS is defined by the SNP GQ03204_B13.1.1304 and is based on the *Picea glauca* ortholog of *PaLAR3 (Hammerbacher et al., 2014).*

PaLAR3 encodes a leucoanthocyanidin reductase (LAR), which is an enzyme of the flavonoid biosynthetic pathway (Figure 1) that belongs to the short-chain dehydrogenase/reductase protein family and catalyzes the synthesis of (+) catechin, a flavan-3-ol that is a precursor of the proanthcyanidins from leucoanthocyanidin (Stafford and Lester, 1984; Mauge et al., 2010). *P. abies* saplings overexpressing *PaLAR3* showed higher (+) catechin levels than wild type *P. abies* saplings (Hammerbacher, 2011; Hammerbacher et al., 2014). The enzymatic product, (+) catechin, has been correlated with the level of resistance against various pathogens and pests in different *Picea* species (Brignolas et al., 1995; Brignolas et al., 1998; Lieutier et al., 2003; Danielsson et al., 2011; Fossdal et al., 2012; Porth et al., 2012). Inoculation of *P. abies* with the bark beetle-associated fungus *Endoconidiophora polonica* increases the transcript levels of *PaLAR3* as well as the monomeric and polymeric flavan-3-ols, including (+) catechin *(Hammerbacher et al., 2014).* Inoculation studies of selected *P. abies* genotypes with known resistance to *H. annosum s.l.* provided some support for the direct involvement of *PaLAR3* and (+) catechin in defense, as genotypes with higher levels of resistance against *H. annosum s.l.* showed increased *PaLAR3 (LAR1)* expression and an accumulation of (+) catechin 5 days after inoculation, which could not be found in the susceptible trees (Danielsson et al., 2011).



	R = H	R = OH
Dihydroflavonol	Dihydroquercetin	Dihydromyricetin
Leucoanthocyanidin	Leucocyanidin	Leucodelphinidin
2,3-trans-(+)-Flavan-3-ol	Catechin	Gallocatechin

Figure 1: Biosynthesis of monomeric and polymeric flavan-3-ols via the falvonoid pathway. Abbreviations: DFR, dihydroflavonol reductase; LAR, leucoanthocyanidin reductase

Taken together GQ03204_B13.1.1304 on LG 6 (Lind et al., 2014) appeared to be associated not only with reduced FGS in the QTL study but also to the previously described *PaLAR3* gene and (+) catechin, both associated with biotic stress responses in *Picea spp*. Thus we selected locus GQ03204_B13.1.1304 and its associated gene *PaLAR3* for analysis. We tested the following hypotheses: i)

GQ03204_B13.1.1304 associates with reduced FGS also outside the QTL mapping population; ii) resistance conferred by GQ03204_B13.1.1304 depends on the *PaLAR3* gene and iii) allelic variation in the *PaLAR3* gene can explain the variation in resistance associated with GQ03204_B13.1.1304. To test the first hypothesis we carefully phenotyped FGS in selected progenies, with defined genotypes at the GQ03204_B13.1.1304 locus, from several open-pollinated families across multiple populations. To address our second and third hypotheses we resequenced *PaLAR3* in individuals homozygous at the GQ03204_B13.1 locus to determine the allelic structure of the gene, to identify non-synonymous substitutions, and to analyze motifs in the promoter. We tested the fungistatic effect of (+) catechin on the *H. parviporum* strain that was used for inoculation, we investigated differences in the protein activity between the enzymes produced by the different *PaLAR3* alleles through heterologous expression in *Nicotiana benthamiana*, and we quantified the allele-specific gene expression of the *PaLAR3* alleles by qRT-PCR (quantitative Reverse Transcribed-PCR).

Materials and Methods

Plant Material

The genotype at locus GQ03204_B13.1 was determined by three SNPs (GQ03204_B13.1.64, GQ03204_B13.1.1304 and GQ03204_B13.1.1406) in 773 *P. abies* individuals, which were genotyped for a total of 1733 SNP sites, employing the Illumina GoldenGate® assay used by Lind et al. (2014). The 773 trees originated from 102 half-sib families collected from four populations in Sweden,

Finland and Russia. The plants were grown in a common garden, Lugnet, in Håbo (Sweden) (Supplemental material1).

Based on their genotype at the locus GQ03204_B13.1, 42 individuals from 14 half-sib families were selected for the experiments. From each family three individuals were selected, one individual homozygous for each allele and one heterozygote (Supplemental material 2).

Plant inoculation and material harvest

Rb175, a well-defined strain of *Heterobasidion parviporum* (Fr.) Niemelä & Korhonen (Stenlid, 1987), was grown on Hagem's medium (Stenlid, 1985) and inoculum was prepared as described by Lind et al. (2007).

Inoculations were done in August 2014 on 11-year old trees growing in the common garden at Lugnet by inserting inoculum wood plugs aseptically into wounds, made using a cork borer with a diameter of 0.7 cm in the bark of one-year-old twigs. Inoculations were covered with Parafilm and left under conditions where daytime temperatures varied between 15 and 25°C during the experimental period. Every plant was inoculated eight times on separate twigs. One twig per individual was harvested at seven days post inoculation (dpi) for RNA extraction. A 20 x 5 mm bark piece around the inoculation point was cut and immediately put in liquid nitrogen. All samples were stored at -80°C until further use. At 21 dpi the remaining seven inoculated twigs were harvested for phenotyping of FGS.

For analysis of constitutive and induced (+) catechin contents, bark (including the phloem) was harvested from intact or inoculated one-year-old twigs, respectively. From each twig, three 20 x 5 mm bark samples including bark,

phloem and parts of the cambium were collected. These bark samples were immediately put in liquid nitrogen and used for biochemical analysis. All samples were stored at -80°C until further use.

Phenotyping of fungal growth in sapwood

Needles were removed from the twigs and five-centimeter cross-sections upward and downward from the inoculation point were cut. Each of these sections was cut into ten five-millimeter pieces that were placed on moist filter paper in a Petri plate, together with the plug that was used for the inoculation. Plates were then left under moist conditions at 21°C and in darkness for one week. After that, a stereomicroscope (Zeiss Stemi 2000C) was used to determine the presence or absence of *H. parviporum* conidia on each one of the five-millimeter plugs under 50x magnification. For each twig, we annotated the sum of the 5 mm-sections upward and downwards from the inoculation point where fungal growth was observed. Plates where no conidia could be observed on the inoculation point or on the inoculation plug were treated as inoculation failures and were discarded. The FGS data was analyzed using a one-way ANOVA with repeated measures in JMP®, Version 10. SAS Institute Inc., Cary, NC, 1989-2007.

Primer design

In order to sequence the 2.8 kb genomic region of *PaLAR3* and create a haplotype network, five primer pairs were designed (a, b, c, d and e) covering the genomic region. For sequencing of the promoter region of the gene three primer pairs spanning 1.5 kb into the 5' upstream region of *PaLAR3* were designed. Primer3 (biotools.umassmed.edu/bioapps/primer3_www.cgi) was used to design all

primers pairs. In order to design these primers, the *P. abies* scaffold MA_176417 sequence from the *P. abies* v.1.0 genome (www.congenie.org) was used as a template. Primers for qRT-PCR were designed manually based on DNA genomic sequences of *PaLAR3* to obtain a 150 bp product and a Tm of 60°C. Primer were checked at www.bioinformatics.org/sms2/pcr_primer_stats for control of quality and properties before primers were synthesized at TAG Copenhagen (Copenhagen, Denmark) (Supplemental material 3).

Sequencing of the *PaLAR3* genomic sequence

Genomic DNA was extracted from 28 individuals homozygous for the GQ03204_B13.1 locus that were included in the inoculation experiment. Firstly, in order to create a haplotype network for *PaLAR3,* 2.8 kb of the *PaLAR3* genomic sequence that spanned from the 5'UTR region to the 3'UTR region of the gene were sequenced. To do this, five overlapping amplicons (a, b, c, d and e) were used to amplify *PaLAR3* (Figure 2b). Secondly, in order to study differences between the promoter regions of the two main *PaLAR3* variants, DNA from two homozygous individuals was selected to sequence 1.5 kb of the 5' upstream region of *PaLAR3*. To do this, three primer pairs positioned at -500, -1000 and -1500 bp from the ORF were used.

For each reaction 50µl PCR-mix with 1x Dream-Taq green buffer, 0.2µM of each of the primers, 0.2mM dNTPs, 1.25U Dream-Taq Polymerase (Fermentas), a final concentration of $MgCl_2$ of 3.25mM, and 0.5-5ng/µl reaction volume of genomic DNA was prepared. The PCR conditions were as follows: initial denaturation at 95°C for 5 minutes, followed by 35 cycles of: 30 s at 95°C, 30 s at 57°C and 2 minutes at 72°C and a final elongation step of 7 minutes at 72°C.

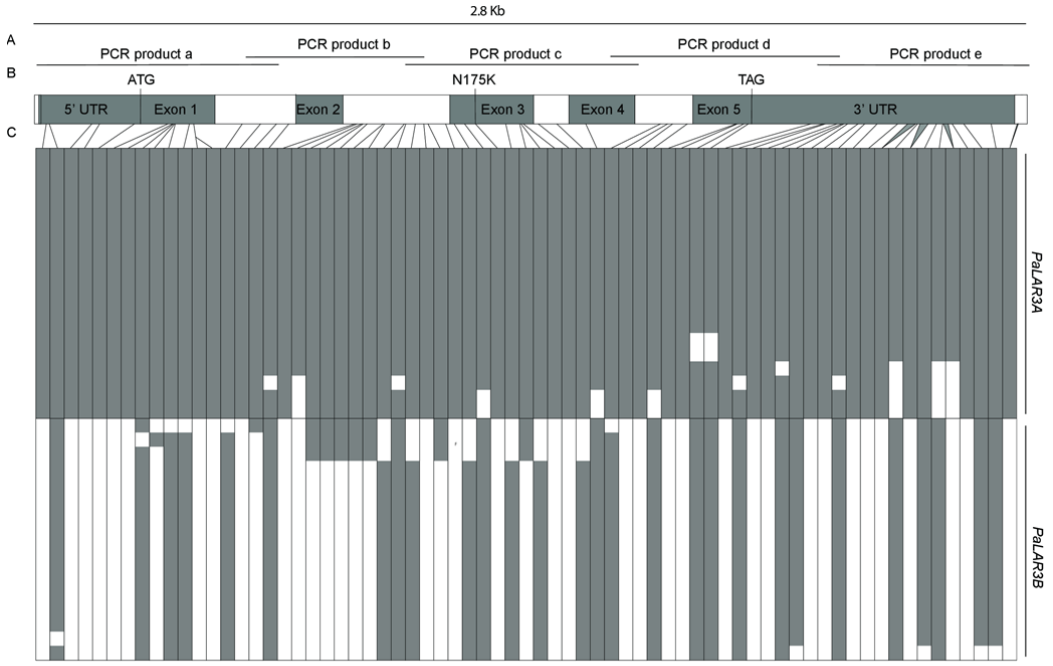


Figure 2: *PaLAR3* structure; A: PCR amplicons used for re-sequencing. B: Gene structure with mutations and their positions. C: Genotype for each mutation in the 36 sequences. Major alleles are represented in gray and minor alleles in white.

PCR products were cloned into pCR2.1-TOPO (Clontech) according to the manufacturer's instructions. Colony PCR was run on selected colonies with M13 primers and the PCR products were sent to Macrogen (Amsterdam, The Netherlands) for Sanger sequencing.

Phylogenetic analysis and haplotype network

Sequence assembly was done with Seqman (DNAstar). Sequences were then imported into MEGA6 (Tamura et al., 2013) and aligned by the ClustalW algorithm with gap opening penalty 15, gap extension penalty 6.66, IUB DNA weight matrix and transition weight 0.5. A Maximum Likelihood tree was created with a bootstrap phylogeny test of 1000 replications and pairwise deletion of missing data.

Genetic distances between haplotypes were imported into HapStar (http://fo.am/hapstar/) to create a Minimum Spanning Network using default

software settings. The network was edited in Illustrator (Adobe) to incorporate allele sizes for each specific subgroup and genotype provenances.

SNP comparison between *Picea glauca* and *P. abies LAR3*

We compared the SNP variation that we observed after re-sequencing *PaLAR3* with the SNP variation that was published by Pavy et al. (2013) for the unigene GQ03204_B13 in *P. glauca*, corresponding to the *P. glauca* full-length sequence BT109050.1 in Genbank. BT109050.1 was used as a reference and the *PaLAR3* genomic sequences were trimmed to fit BT109050 by removing the introns and the 3' and 5' UTR regions that were not covered by BT109050.1. SNPs were compared to our allelic structure by observing the presence or absence of *P. glauca* SNPs in *P. abies* and *vice versa*.

Comparative analysis of the *PaLAR3* promoter region

Cis-regulatory motifs in the isolated 1.5 kb 5' upstream region of *PaLAR3* from two individuals (51: PaLAR3AA and 51:PaLAR3BB) representing the two main *PaLAR3* genotypes were identified using PLACE (Higo et al., 1999). The output files were manually compared to identify motifs differing in frequency or position between the two promoter sequences. The promoter sequences were also imported into MEGA 6.0 to identify any larger structural differences between them.

RNA extraction and cDNA synthesis

Bark samples (split from sapwood at the cambium thus including bark, phloem and some of the cambium tissue) from the inoculated plants that were harvested

one week after the inoculation with *H. parviporum* were pulverized in a mortar with liquid nitrogen and RNA was extracted following the protocol described by Chang et al. (1993). RNA samples were treated with DNaseI (Sigma Aldrich, USA) and RNA concentration was determined with a NanoDrop (Spectrophotometer ND 1000, Saveen Werner). Five hundred ng of total RNA were reverse transcribed with the iScript™ cDNA Synthesis Kit (Bio-Rad) according to the manufacturers' instructions.

Quantitative Polymerase Chain Reaction

A PCR with allele-specific primers (Supplemental material 3) was run on cDNA samples and the two products were purified by cloning the products into TOPO® vectors (Clontech) following the manufacturer's instructions. Plasmids were purified using The PlasmidPrep minikit® and dilution series were then prepared from 10^8 to 10^3 copies/µl to create regression lines that allowed to estimate the transcript levels in the samples.

qRT-PCR was performed with the SsoFast™ EvaGreen® Supermix (Bio-Rad) according to the instructions in the manual, using 0.3µM of each primer and one µl cDNA as template. The quantitative PCR was carried out in an iQ5™ Multicolor Real-Time PCR Detection System thermo cycler (Bio-Rad) using a program with a 30 seconds initial denaturation step at 95°C, followed by 40 cycles of 5s denaturation at 95°C and 10s at 60°C. Three repetitions per standard, sample and negative control were run. Melt curve analyses were used to validate the amplicon. Primer efficiency was confirmed to be similar for the two allele-specific primer pairs used for qPCR. Transcript levels were estimated using the regression line obtained with the Ct values for the standards. A pairwise t test

(t=6.444 df=13) was run to detect differences between transcript levels of the two *PaLAR3* alleles in the heterozygous individuals.

Chemical analysis

For the 42 individuals used for the inoculation study, bark samples (including bark, phloem and some of the cambium) of intact twigs and inoculated twigs were harvested at day 0 and seven days after inoculation, respectively. Bark samples were ground in a mortar with liquid nitrogen. Once pulverized, the samples were freeze-dried and flavan-3-ol content was measured using the method described by Hammerbacher et al. (2014).

Variances in the Flavan-3-ol content were analyzed using the Friedman test, a non-parametric repeated measures one-way ANOVA, together with a Dunn's multiple comparison test to detect variance differences.

Protein activity study

PaLAR3A and *PaLAR3B* were cloned into the expression vector pEAQ Dest 1 (Sainsbury et al., 2009) using LR clonase (Invitrogen) following the manufacturer's protocol. Chemically competent *Agrobacterium tumefaciens* GV3101 were transformed with the constructs. LAR genes were transiently expressed by infiltration of *A. tumefaciens* into the leaves of *Nicotiana benthamiana* using the protocols from Sainsbury et al. (2009). Tobacco leaf tissue expressing *PaLAR3A* and *PaLAR3B* constructs was harvested 5 days after transformation. Proteins were extracted from approximately 300 mg of fresh homogenized tissue with 3 ml extraction buffer (50 mM Tris, pH 7.8, 5 mM ascorbic acid, 5 mM dithiothreitol, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 10 mM $MnCl_2$, 0.5

M NaCl, 10% glycerol (v v^{-1}), 1% (w v^{-1}) polyvinylpyrrolidone (MW 360,000), 4% polyvinylpolypyrrolidone (w v^{-1}) and 0.1% Tween 20 (v v^{-1})) at 4 °C for 30 min. shaking. Crude protein content was quantified using the Bradford reaction. Approximately 200 µg crude protein was mixed with 2 µg bovine serum albumin (BSA) as an internal standard and subjected to tryptic digest following the protocols from Stergachis et al. (2011).

The amounts of both PaLAR3A and B proteins were quantified by LCMS analysis using an API 5000 mass spectrometer coupled to an Agilent 1260 Infinity HPLC and a DXB-C18 (4.6 X 50 mm 1.8 µm particle size) column. The total mobile phase flow rate for chromatographic separation was 1 ml min^{-1}. The column temperature was maintained at 25°C. Peptides from the tryptic digest were separated using 0.05% (v v^{-1}) formic acid and 5% (v v^{-1}) acetonitrile in water and 0.05% (v v^{-1}) formic acid and 5% (v v^{-1}) water in acetonitrile as mobile phases A and B respectively with the following elution profile: 0-0.5 min, 100% A; 0.5-15 min, 0-35% B in A; 15-17.5 min 35-100% B; 17.5-22.5 min 100% B and 22.5-30 min 100% A.

Multiple reaction monitoring (MRM) transitions for peptides were predicted using Skyline (http://sciex.com/products/software/skyline-software). The ion spray voltage in the MS was maintained at -4500 V. The turbo gas temperature was set at 700 °C. Nebulizing gas was set at 70 psi, curtain gas at 25 psi, heating gas at 60 psi and collision gas at 10 psi. The analyte precursor ion → product ion transitions that were monitored are described in supplemental material 4. The amounts of *Pa*LAR3A and *Pa*LAR3B were calculated by relating the peak areas of the LAR peptides to the peak areas of the internal standard, BSA, and are expressed as BSA equivalents.

LAR enzyme activity was assayed at 28 °C for 15 min using approximately 0.2-1 µg LAR protein, 5 µg dihydroflavonol reductase (*Md*DFR) from *Malus domestica* (Pfeiffer et al., 2006) expressed in *N. benthamiana* as above, 1 mM NADPH and 2.5 mM taxifolin. Products were analyzed and quantified using the protocols described in Hammerbacher et al. (2014). Assays using leucodelphinidin were initiated as above, but without the addition of *Md*DFR.

Specific activities of the enzymes encoded by the two alleles were calculated by taking into account the amount of LAR protein in the assay and the rate of product formation (µmol product formed * mg LAR protein^{-1} * minute^{-1}).

(+) catechin effect on fungal growth of *H. parviporum*

The fungistatic effect of biologically relevant (+) catechin levels was tested by growing *H. parviporum* Rb175 on 90 mm Petri plates with standard Hagems medium as a control and Hagems medium supplemented with (+) catechin levels corresponding to the average (+) catechin levels in bark of *PaLARA* and *PaLARB* homozygotes, 5.4 mg/mL and 7.6 mg/mL, respectively. On each plate, two perpendicular axes having as a center the inoculation point were drawn. Fungal growth was measured every 24 hours and was assessed by marking the furthest point reached by the fungal hyphae on each of the four axes and calculating the average of the four marks. Three repetitions per treatment were done and measurements were done until day 9, when the first hyphae reached the edge of the Petri plates. Data were analysed using a one-way ANOVA in order to detect statistical differences between treatments.

Results

Genotyping shows co-segregation of three SNPs in the *PaLAR3* locus

In addition to the SNP GQ03204-B13.1.1304 that was linked to the FGS QTL, we genotyped 773 individuals for two additional SNPs in the same locus (GQ03204-B13.1.1406 and GQ03204-B13.1.64) using the Illumina GoldenGate® assay. These three SNPs co-segregated in all 773 individuals and were used to define two alleles of *PaLAR3* using the genotype at the locus GQ03204_B13. The major allele in the locus (p=0.78) was called *PaLAR3A* and the minor allele *PaLAR3B,* with an allele frequency that varied between 0.16 and 0.29 among four Northern European populations (Table 1).

PaLAR3 has two allelic lineages

The isolated *PaLAR3* 2.8 kb genomic sequence spanned the whole coding sequence together with the 5' UTR region (284 bp) and the 3' UTR region (706-713 bp). The open reading frame of *PaLAR3* is made up of five exons (Figure 2c). The *PaLAR3* gene was fully sequenced in 28 individuals homozygous for either GQ03204_B13 allele returning in total 36 sequences (NCBI accession numbers KT592308- KT592342). Out of the 36 sequences 15 were unique variants. Phylogenetic analysis of the sequences suggested that two allelic lineages of *PaLAR3*, *PaLAR3A* and *PaLAR3B*, were present in the population (Figure 3). In total 64 SNPs and five indels were detected in the 15 unique sequences that we studied. Twenty-seven SNPs and three indels co-segregated with the original SNPs defining the GQ03204-B13.1.1304 locus, delineating two allelic lineages encoding two variants of *PaLAR3*. The two allelic lineages of *PaLAR3* differed in a a single non-synonymous substitution leading to an amino acid change (N175K) in the PaLAR3 protein. This sequence change substitutes a polar asparagine

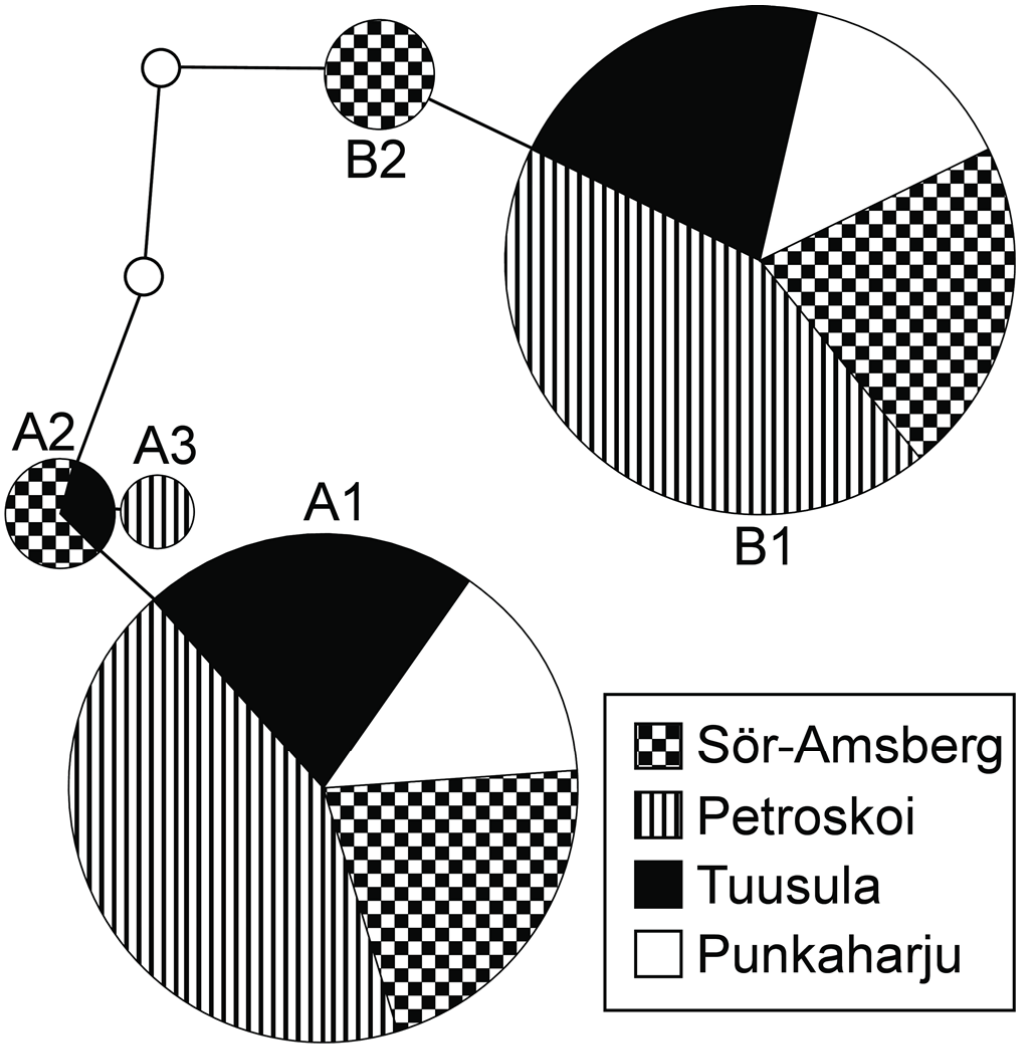


Figure 3: Haplotype network of *PaLAR3* sequences in Picea abies (n=36). Circle diameter is proportional to the number of sequences in a specific haplotype. Patterns indicate the geographical origin of the recorded haplotypes and open circles indicate inferred missing haplotypes.

residue with a positively charged lysine residue and occurs between the tyrosine in position 174 and the lysine in position 176, both members of the catalytic triad of the enzyme (Mauge et al., 2010).

The remaining 39 mutations (37 SNPs and two indels) were found to vary within allelic lineages. Eleven SNPs were mutations that mostly differed between the *PaLAR3A* and *PaLAR3B* allelic lineages except for particular subgroups within the allelic lineages that carried the alternative nucleotide (Figure 2c). Twelve of these mutations were found within the *PaLAR3A* allelic lineage, and 16 in the *PaLAR3B* allelic lineage.

We compared the SNP variation of *PaLAR3* with the SNP variation of the *LAR3* gene reported from *P. glauca* EST sequences (Pavy et al., 2013) to study the possibility that the allelic lineages are conserved between the two species. We found eight SNPs that were shared between the two species, 24 that are *P. glauca*-specific and 26 that are *P. abies*-specific. From the eight SNPs that were shared between the two species, six differed between the *PaLAR3A* and *PaLAR3B* allelic lineages in *P. abies* while the other two SNPs belong to the B2 subclass of the *PaLAR3B* allelic lineage.

The main alleles within the respective lineage *PaLAR3A* A1 and *PaLAR3B* B1 were present in the whole geographic region analysed. Of the minor variants it was only *PaLAR3A* A2 that was present in more than one population (Figure 3).

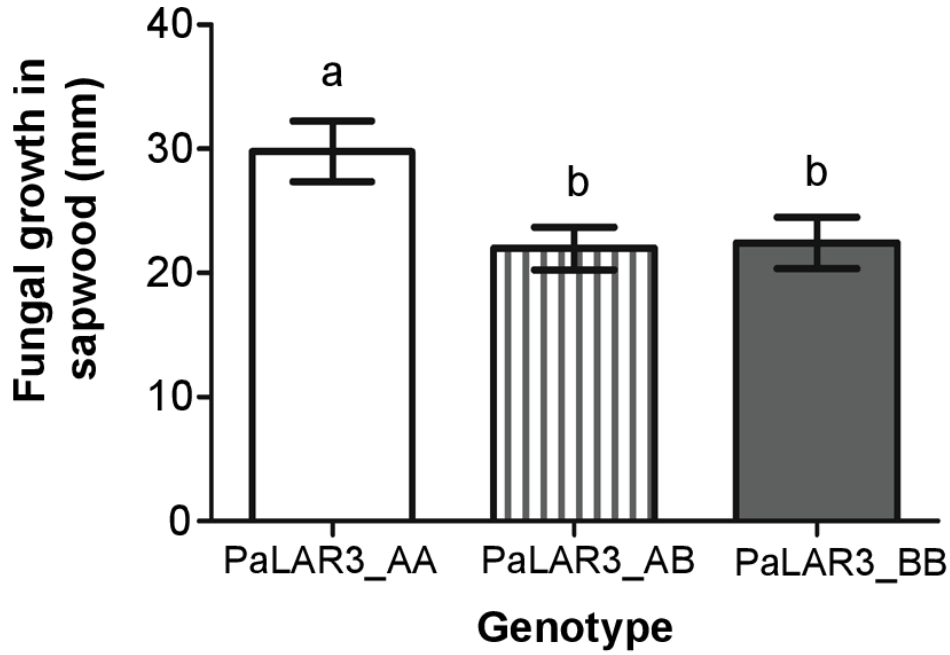


Figure 4: Average *H. parviporum* fungal growth in sapwood (mm) in half-sibs with different genotypes at the *PaLAR3* locus; *PaLAR3A* homozygotes (AA), *PaLAR3B* homozygotes (BB) and heterozygotes (AB) sampled from 14 families. The error bar indicate standard deviation and letters above the columns indicate statistically significant differences (Repeated measures one-way ANOVA, $P < 0.001$).

After trees were inoculated with *H. parviporum* fungal growth in sapwood was measured. Our data showed a significant effect on FGS for family ($P<0.0001$), genotype ($P<0.0001$) and their interaction ($P<0.0001$) (Supplemental material 5). Homozygous individuals for the *PaLAR3B* allele and heterozygous individuals showed on the average a 27% reduction in FGS compared to their half-siblings homozygous for the *PaLAR3A* allele, suggesting that *PaLAR3B* alleles are dominant over the *PaLAR3A* alleles (Figure 4).

Constitutive bark (+) catechin content is higher in *PaLAR3B* homozygotes

A significant difference in (+) catechin content in bark was observed between individuals with different genotypes. *PaLAR3B* homozygotes showed a higher constitutive (+) catechin content (P = 0.0171, Friedman test) compared with the PaLAR3A homozygotes (Figure 5a). No differences were observed for any other flavan-3-ol or proanthocyanidins that we analyzed. However, analysis of the (+) catechin content in *P. abies* bark after *H. parviporum* infection did not show any significant differences between genotypes (Figure 5b).

The *H. parviporum* colony area was significantly different (P<0.0001) when grown on medium amended with (+) catechin levels corresponding to the estimated constitutive levels in the *PaLAR3A* or *PaLAR3B* homozygotes. After nine-days growth on medium amended with 5.4 mg/mL (+) catechin the colony area was, on average, 48.4% of the area for the control treatment (no (+) catechin supplemented). At the same time, the colony area on medium containing 7.6 mg/mL (+) catechin was, on average 48% of the area in the presence of 5.4 mg/mL (+) catechin and 23.3% of the area of control plates (Figure 6).

***PaLAR3A* and *PaLAR3B* show similar enzymatic activity**

The substitution of a polar amino acid in *PaLAR3A* for a positively charged amino acid in *PaLAR3B* is immediately adjacent to the conserved Tyr174, a member of the predicted LAR catalytic triad, Ser…Tyr-X-X-X-Lys (Mauge et al., 2010) , and so may alter the hydrogen bonds around the catalytic triad, which could influence the enzyme´s catalytic properties. To compare the catalytic properties of the two protein variants the specific activity of heterologously expressed PaLAR3A and PaLAR3B proteins was determined *in vitro*. No significant difference in specific

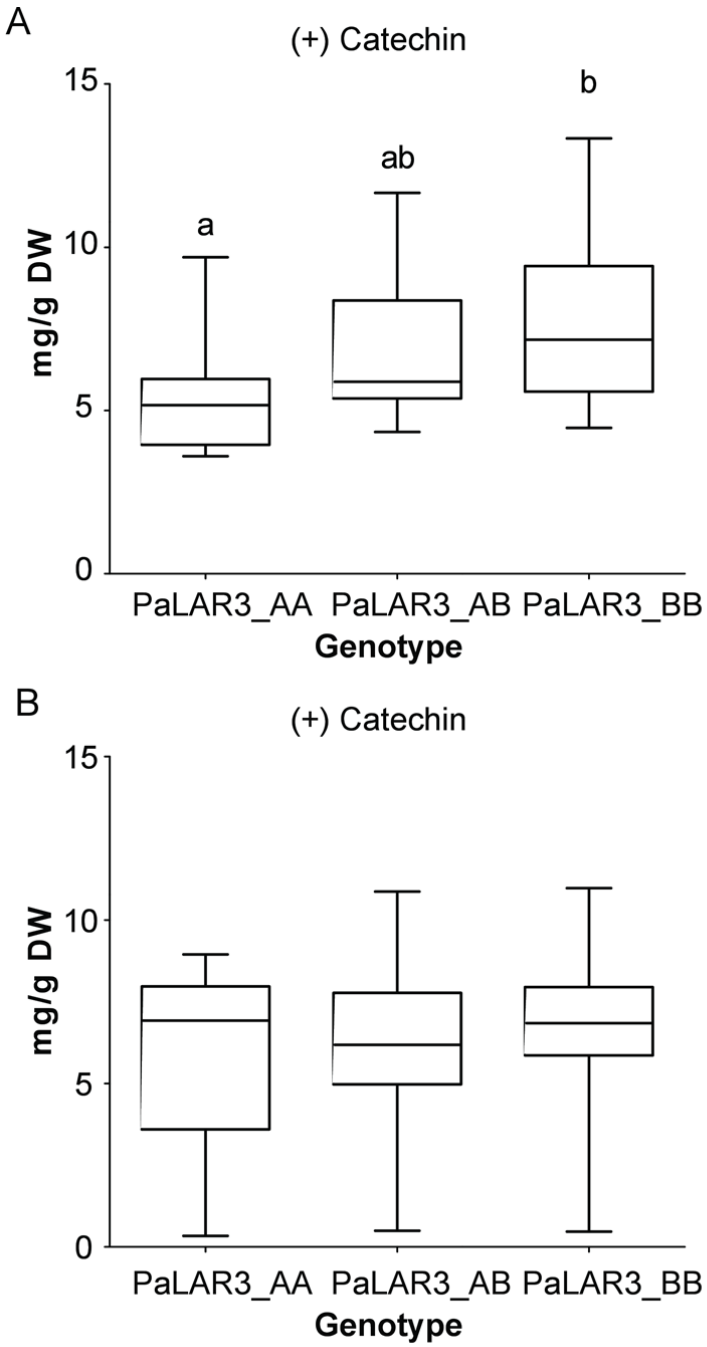


Figure 5: (+) catechin content in bark in mg per gram of dried material for plants with different *PaLAR3* genotype untreated (A) and one week after inoculation (B). Letters indicate statistical differences between genotypes.

activity for the transformation of leucocyanidin into (+) catechin (Figure 7a) or leucodelphinidin into (+) gallocatechin (Figure 7b) was found between PaLAR3A

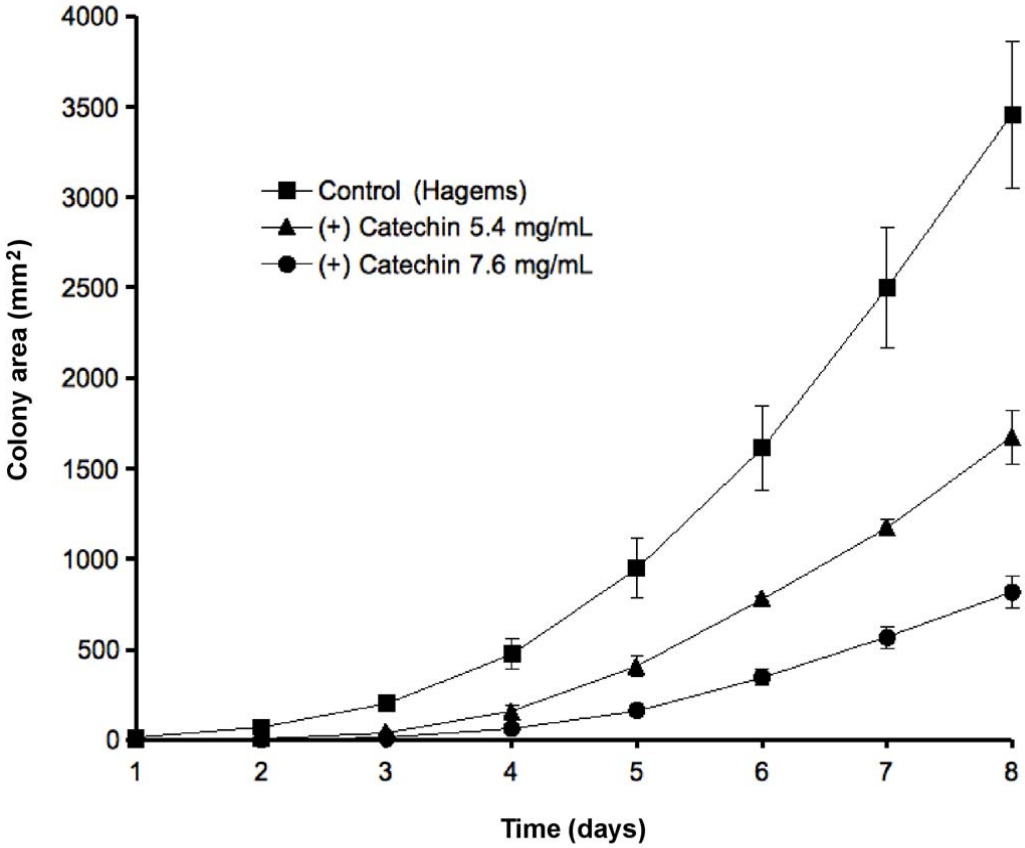


Figure 6: Average colony area for *H. parviporum* Rb175 on Hagems medium and after the addition of 5.4 mg/mL and 7.6 mg/mL (+) catechin. (n=3) horizontal bars show standard deviation.

and PaLAR3B, suggesting that the amino acid substitution does not interfere with the proteins' catalytic properties.

***PaLAR3B* is shows higher transcript levels than *PaLAR3A* after *H. parviporum* challenge**

After comparing the transcript levels of the *PaLAR3A* and *PaLAR3B* alleles in 14 heterozygotes by RT-qPCR with allele-specific primers we observed that *PaLAR3B* transcript levels were higher than the transcript levels of the *PaLAR3A* allele ($P<0.0001$, Paired t-test). The average *PaLAR3B*/*PaLAR3A* ratio was 6.9 suggesting that *PaLAR3B* transcription is higher after *H. parviporum* infection than *PaLAR3A* (Figure 8).

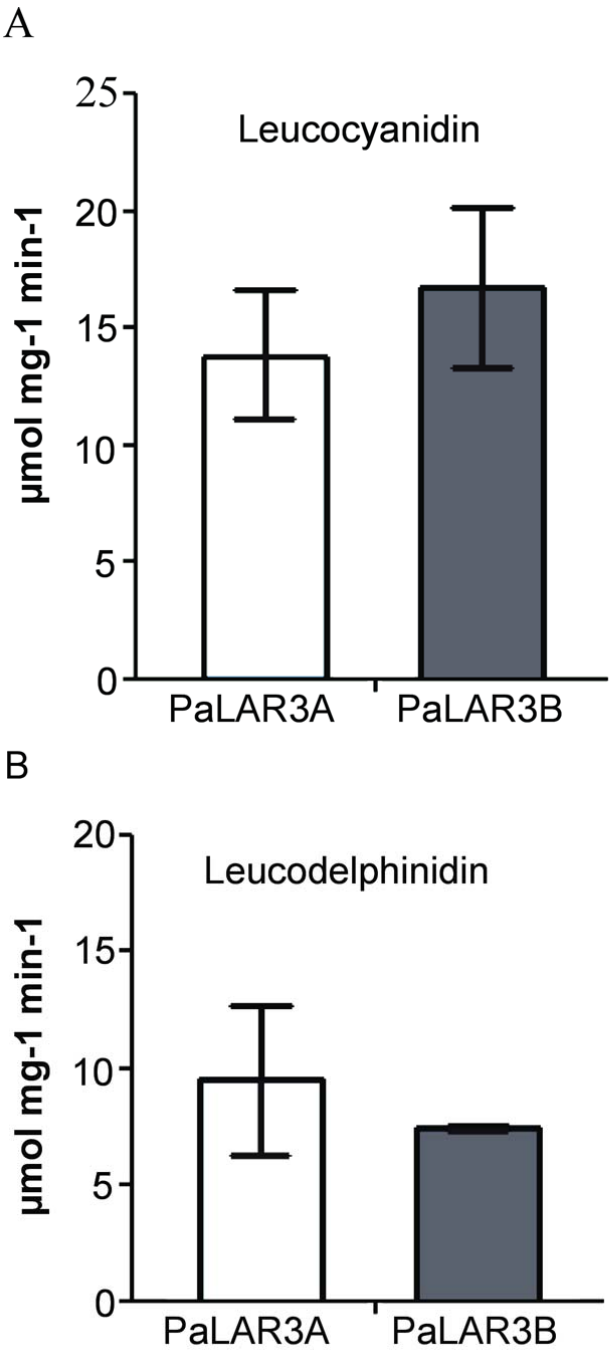


Figure 7: Specific activity of heterologously expressed PaLAR3A and PaLAR3B for leucocyanidin (A) and leucodelphinidin (B). The error bars indicate standard deviation.

Promoter analysis

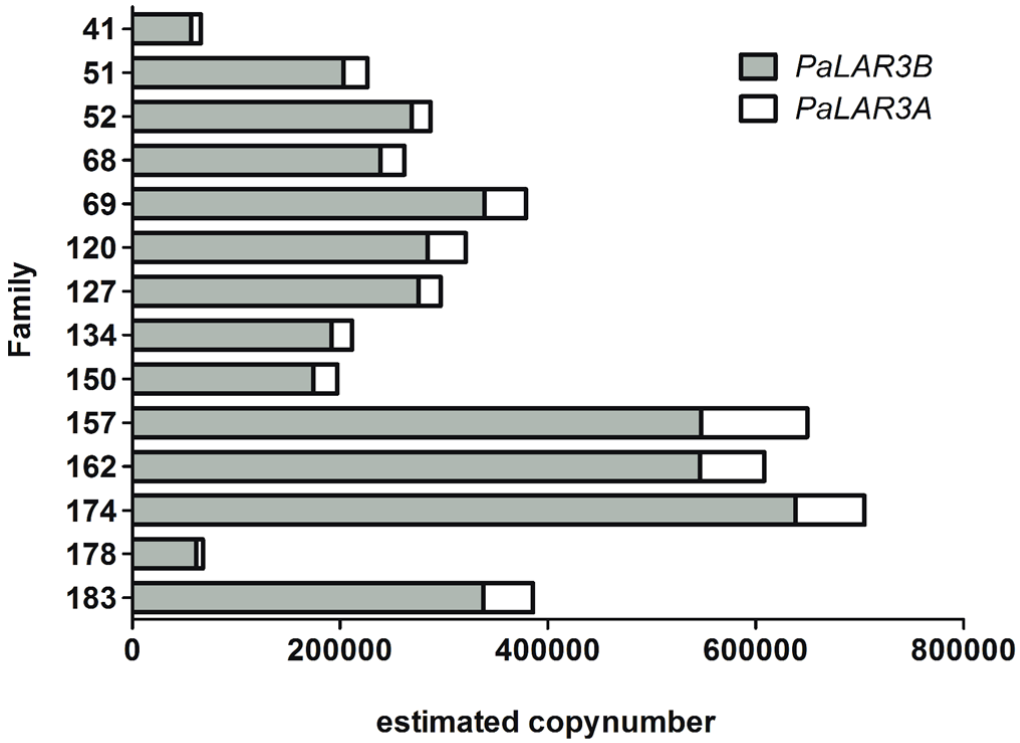


Figure 8: Transcript levels for *PaLAR3* in *P. abies* bark of 14 heterozygous individuals one week after inoculation with *H. parviporum* Rb175. For each individual, white indicates the transcript levels corresponding to *PaLAR3A* and grey for *PaLAR3B.*

The promoter regions of *PaLAR3A* (1733 bp) and *PaLAR3B* (1611 bp) were isolated. Like the *PaLAR3* gene sequence, the promoter region shows consistent differences between the allelic lineages including SNPs and indels, which lead to a divergence in promoter motifs. Analyses in the PLACE database (www.dna.affrc.go.jp/**PLACE**/*)* revealed 35 motifs in *PaLAR3A* that are not present in *PaLAR3B* and 22 motifs are present in *PaLAR3B* but not in *PaLAR3A* (Supplemental material 6). These motifs are related with biotic and abiotic stress, development and hormone signaling.

The most obvious difference between the two promoters is a 129 bp region that is not present in *PaLAR3B*. This 129 bp region is repeated three times in

PaALAR3A and twice in *PaLAR3B*, even though one of these two repeats in PaLAR3B lacks a part of the sequence repeat (Supplemental material 7).

Discussion

This study aimed to investigate the genetic and physiological background behind the higher resistance levels to *H. parviporum* in *P. abies* associated with the locus GQ03204_B13.1. A novel allelic lineage of *PaLAR3, PaLAR3B,* was identified and showed dominance conferring higher resistance to *H. parviporum* growth in sapwood and higher amounts of (+) catechin in bark. We observed that the biologically relevant (+) catechin levels have a fungistatic effect reducing the growth of *H. parviporum in vitro*. The transcript levels of the *PaLAR3B* allele were significantly higher than of *PaLAR3*A while the specific activities and substrate specificity of the enzymes encoded by the two allelic lineages, measured by *in vitro* assays, were similar. Our results indicate that regulation of *PaLAR3* expression is responsible for the observed effects on resistance.

Deducing the association genetics of complex traits in conifers is a demanding task and has often proven unsuccessful, as shown by the scarce number of molecular markers that have been validated so far (Neale and Savolainen, 2004) in this taxon. In cases where resistance is associated with single genes, such as the major QTL *Fr1* conferring resistance to *Cronartium quercuum* in *Pinus* (Wilcox et al., 1996; Kuhlman et al., 1997), genetic effects may be large and detection can be relatively uncomplicated (Quesada et al., 2010; Neale et al., 2014; Quesada et al., 2014). In contrast, quantitative disease resistance, e.g. the resistance to *H. annosum s.l.* in conifers, is typically controlled by many *loci* each with a relatively small effect (Quesada et al., 2010). The validated QTL, GQ03204_B13.1, is estimated to explain just under 5% of the phenotypic variation detected for FGS of *H. parviporum* in *P. abies* (Lind et al., 2014). Thus, a precise disease phenotyping is a critical step towards identifying and validating

specific loci and alleles associated with quantitative disease resistance (Kayihan et al., 2005; Quesada et al., 2010); while careful experimental design and precision in the validation experiment is necessary to allow detection of effects of the QTL outside the original pedigree. The FGS phenotyping method relies on the formation of fungal conidia gives a semi-quantitative measure of the spread of living mycelium from the inoculation site at a given time point. The progression of the mycelium after inoculation is positively correlated to the susceptibility of the host and the incubation time (Swedjemark et al., 1997; Bodles et al., 2007) allowing quantitative assessments of resistance, or fungal virulence, in the conifer-*Heterobasidion* pathosystem (Swedjemark et al., 1997; Bodles et al., 2007; Lehtijärvi et al., 2009; Arnerup et al., 2010). The careful phenotyping of FGS in selected progenies with defined genotypes at the GQ03204_B13.1 locus from several open-pollinated families across multiple populations allowed us to validate the locus GQ03204_B13.1. The locus comprises the *PaLAR3* gene (Hammerbacher et al., 2014) and we identified *PaLAR3B,* a novel *PaLAR3* allele, as a genetic marker for *H. parviporum* resistance in *P. abies*.

When resequencing the *PaLAR3* locus in multiple *P. abies* individuals we observed a dimorphism, where two main alleles could be identified without evidence of haplotype clustering by geographical distribution. Allele dimorphisms have been reported for genes associated with abiotic stress in *Pinus* species (Gonzalez-Martinez et al., 2006; Wachowiak et al., 2009). Patterns of conserved dimorphism can result from admixture of isolated populations, neutral coalescence process, or by balancing selection. Since the same pattern of dimorphism was observed in all four populations and there are no data that

support an admixture of the populations used in our study, this is not the reason in our case. For dimorphism to occur under neutral coalescence process, no or low recombination frequency has to be assumed, which has been reported to maintain two major haplotypes in genes encoding enzymes in the phenylpropanoid pathway in *A. thaliana (Aguade, 2001)*. The breeding system in conifers and the observed differences in FGS and transcript levels combined with the role of the enzymatic product of *PaLAR3*, (+) catechin, in defense (Brignolas et al., 1995; Brignolas et al., 1998; Lieutier et al., 2003; Danielsson et al., 2011; Fossdal et al., 2012; Porth et al., 2012) makes a mechanism involving a long-term balancing selection between two traits a likely explanatory model for the existence of the two allele dimorphism in *PaLAR3*.

Even though the two *PaLAR3* alleles differ in their gene sequences by multiple co-segregating mutations, the difference between the encoded proteins is limited to a single amino acid substitution at position 175. As pointed out before, the active site regions of LAR enzymes are similar between angiosperms and gymnosperms, with conserved His, Ser, and Tyr residues (Hammerbacher et al., 2014). Thus, compared to the available VvLAR1 structure (Mauge et al., 2010), it appears that the substitution at position 175 is flanked by two of the putative members of the catalytic triad (Ser, Tyr, Lys) of the enzyme, namely Tyr174 and Lys177. Amino acid replacements at positions flanking members of the catalytic triad have been shown to alter protein function and stability in a *B. amyloliquefaciens* subtilisin (Estell et al., 1985). Consequently we analyzed whether the variation between a polar amino acid (Asn, *PaLAR3A*) and a positively charged amino acid (Lys, *PaLAR3B*) at position 175 affects the specific activity of the protein. This amino acid substitution did not interfere with the

catalytic properties of the enzymes, as no significant differences in protein activity or specificity were found between *Pa*LAR3A and *Pa*LAR3B. This result is in agreement with previous studies on different short chain dehydrogenases/hydrolases (SDR) proteins suggesting that the amino acids at the positions between the conserved Tyr and Lys in SDRs are not important for the enzymatic activity (Johnson et al., 2001; Petit et al., 2007; Mauge et al., 2010). More importantly, the similar protein activities and substrate specificities of *Pa*LAR3A and *Pa*LAR3B, shows that functional differences between the two proteins are not the reason behind the differences in constitutive bark (+) catechin content and FGS associated with the different genotypes.

The steady-state mRNA levels of *PaLAR3B* were significantly higher than those of *PaLAR3A* in heterozygous plants challenged with *H. parviporum*. There are several possible mechanisms that could result in different steady-state mRNA levels between two alleles. First, the mRNA originating from the different alleles could have different stability (Narsai et al., 2007; Plantegenet et al., 2009). Secondly, polymorphisms in the *cis*-regulatory region of genes may results in considerable differences in expression between alleles (de Meaux et al., 2005; Bentsink et al., 2006; Plantegenet et al., 2009). Analysis of SNPs and indels along the *PaLAR3* 5'-upstream region revealed 57 *cis*-regulatory motifs that differed between the promoters of the two alleles. In *Arabidopsis* drastic structural changes in the promoter, open reading frame or flanking sequences are major causes for expression level polymorphisms between alleles at loci exhibiting simple, single locus inheritance patterns (Plantegenet et al., 2009). The main difference between the promoters of *PaLAR3A* and *PaLAR3B* is a 129bp repeat found three times in *PaLAR3A* but only twice in *PaLAR3B.* This repeat includes

cis-regulatory motifs and shows a certain similarity with a known ATAF2-binding sequence identified in *A. thaliana* (Wang and Culver, 2012). Interestingly, one of the two repeats of *PaLAR3B* includes a deletion that could affect the capacity of a transcription factor to bind at this site as shown by Wang and Culver (2012). This leaves *PaLAR3A* with three, and *PaLAR3B* with one complete repetition of the *cis*-regulatory motif. Taken together, the structural differences that we observe in the promoter regions of the two *PaLAR3* alleles are likely to cause differential expression of the two alleles, but this would need to be confirmed with *in vivo* experiments.

Constitutively higher levels of secondary metabolites can act as a defense to herbivores and pathogens (Verne et al., 2011; Hall et al., 2011; Mageroy et al., 2015); thus the constitutively higher levels of (+) catechin in bark of *PaLAR3B* homo- or heterozygotes could act as a barrier to fungal growth. PaLAR3 was shown to be the most highly expressed LAR in Norway spruce under normal growth conditions (Hammerbacher et al., 2014) and a more highly transcribed allele of this gene is thus likely to maintain high constitutive catechin levels in the tree. The correlation between bark (+) catechin content and genotype at the *PaLAR3* locus could not be determined after *H. parviporum* challenge, probably because other members of the *PaLAR* family, e. g. *PaLAR1, PaLAR2* and *PaLAR4*, are more strongly induced in response to fungal infection than *PaLAR3 (Danielsson et al., 2011; Arnerup et al., 2013; Hammerbacher et al., 2014)*. Furthermore, the time interval between inoculation and tissue harvest and catechin analysis (7d) in our experiment might have been too short for the shoots to reach their final metabolite concentration.

Our observations regarding the expression levels of the two alleles at the *PaLAR3*

locus and the constitutive (+) catechin content in bark are reminiscent of several other defense associated loci comprising genes in the secondary metabolic pathways in *Picea*; higher transcript levels of the β-glucosidase gene *Pgβglu-1* were recently shown to correlate with higher levels of piceol and pungenol and were associated with resistance to spruce budworm in *P. glauca* (Mageroy et al., 2015), showing that variation in transcript levels can explain conifer resistance phenotypes. Similarly Porth et al. (2012) employed a QTL analysis based on gene expression levels (eQTL analysis) and identified 38 phenylpropanoid-related genes, including two LARs, that co-segregated with weevil resistance in a *P. glauca* x *P. engelmanii* (interior spruce) cross. These studies and our observation on the *PaLAR3* locus underline the role of *cis*- and *trans*-regulatory elements in constitutive as well as induced resistance.

The three SNP-based markers that were used for the first genotyping of the locus associated with GQ03204_B13.1 in our population were originally developed in *P. glauca* (Rigault et al., 2011; Pavy et al., 2012). In addition to the mutation that leads to the amino acid substitution at position 175, we identified seven SNPs shared between *P. abies* and *P. glauca,* which are estimated to have diverged 13-20 million years ago (Nystedt et al., 2013; Warren et al., 2015). Our observations suggest the possibility of a similar allelic structure in the *P. glauca PaLAR3* ortholog. Convergent evolution cannot be discarded as a mechanism creating similar allelic structure in the two sister taxa if there is a pressure favoring the specific mutations. However, the similarity in specific activity of the *PaLAR3* allelic lineages with the substrates leucocyanidin or leucodelphinidin suggests a little selection pressure for the N175K mutation. Moreover, Bouille and Bousquet (2005) and Chen et al. (2010) reported vast numbers of trans-species shared

polymorphisms in the genus *Picea,* indicative of an incomplete lineage sorting at speciation, highlighting the possibility that the *PaLAR3* allelic lineages predate the species-split. Full-length sequencing of the ortholog in multiple *P. glauca* individuals should be helpful in resolving this question.

Taken together, the validation of the *PaLAR3* locus offers a glimpse of the genetics underlying the previously reported (Brignolas et al., 1995; Brignolas et al., 1998; Lieutier et al., 2003; Danielsson et al., 2011; Fossdal et al., 2012) association between flavanol content and reduced susceptibility to pathogenic fungi in *Picea*, the differential transcript levels of the two alleles indicates that the regulation of secondary metabolism in defense is more complex, and possibly even more interesting, than hitherto understood.

Supplemental materials associated with the article

Supplemental material 1: GPS coordinates for the common garden and for the four populations that were used in this study.

Supplemental material 2: List of half-sib families with their provenance and country of origin.

Supplemental material 3: List of primers that were used for the genomic sequencing of *PaLAR3* and allele-specific qRT-PCR.

Supplemental material 4: Multiple reaction monitoring (MRM) transitions for peptides analyzed to quantify LAR and BSA by liquid chromatography-tandem mass spectrometry (DP, declustering potential; CE, collision energy; CXP, collision cell exit potential).

Supplemental material 5: Average fungal growth in sapwood and standard deviation for each one of the genotypes in the 14 families that were studied.

Supplemental material 6: Motif list for the *PaLAR3A* and *PaLAR3B* promoter sequences.

Supplemental material 7: Sequence of 1.5 kb promoter of the promoter region of

PaLAR3A and *PaLAR3B.* Black boxes indicate a 129 bp repetition and red characters indicate the position of the CAAT and TATA boxes, and the ORF of the gene.

Figure list

Figure 1: Biosynthesis of monomeric and polymeric flavan-3-ols via the falvonoid pathway. Abbreviations: DFR, dihydroflavonol reductase; LAR, leucoanthocyanidin reductase.

Figure 2: *PaLAR3* structure; A: PCR amplicons used for re-sequencing. B: Gene structure with mutations and their positions. C: Genotype for each mutation in the 36 sequences. Major alleles are represented in gray and minor alleles in white.

Figure 3: Haplotype network of *PaLAR3* sequences in Picea abies (n=36). Circle diameter is proportional to the number of sequences in a specific haplotype. Patterns indicate the geographical origin of the recorded haplotypes and open circles indicate inferred missing haplotypes.

Figure 4: Average *H. parviporum* fungal growth in sapwood (mm) in half-sibs with different genotypes at the *PaLAR3* locus; *PaLAR3A* homozygotes (AA), *PaLAR3B* homozygotes (BB) and heterozygotes (AB) sampled from 14 families. The error bar indicate standard deviation and letters above the columns indicate statistically significant differences (Repeated measures one-way ANOVA, $P < 0.001$).

Figure 5: (+) catechin content in bark in mg per gram of dried material for plants with different *PaLAR3* genotype untreated (A) and one week after inoculation (B). Letters indicate statistical differences between genotypes.

Figure 6: Average colony area for *H. parviporum* Rb175 on Hagems medium and after the addition of 5.4 mg/mL and 7.6 mg/mL (+) catechin. (n=3) horizontal bars show standard deviation.

Figure 7: Specific activity of heterologously expressed PaLAR3A and PaLAR3B for leucocyanidin (A) and leucodelphinidin (B). The error bars indicate standard deviation.

Figure 8: Transcript levels for *PaLAR3* in *P. abies* bark of 14 heterozygous individuals one week after inoculation with *H. parviporum* Rb175. For each individual, white indicates the transcript levels corresponding to *PaLAR3A* and grey for *PaLAR3B.*

Acknowledgements

We thank Dr. Lawrie Wright for his considerable contribution concerning the protein quantifications by mass spectrometry and Bettina Raguschke for assistance in the

laboratory.

Financial support was received from SSF (grant nr R8b08-0011), and by the Swedish Research Council FORMAS, grant nrs 2012-1276 and 217-2007-433, the European Community's Sixth Framework Programme, under the Network of Excellence Evoltree (www.evoltree.eu/) and by the Seventh Framework Programme (FP7/2007-2013) under the grant agreement nu 211868 (Project Noveltree, www.noveltree.eu/), Support also came from the Biodiversa projects Linktree (www.igv.fi.cnr.it/linktree/) and TipTree, and The Max Planck Society.

Allele	Punkaharju	Tuusula	Sör-Amsberg	Petroskoi	Total
PaLAR3A	0.82	0.84	0.74	0.71	0.78
PaLAR3B	0.18	0.16	0.26	0.29	0.22

Table 1: Major and minor allele frequencies at the *PaLAR3* locus in the four populations that were studied and their average.

References

Aguade M (2001) Nucleotide sequence variation at two genes of the phenylpropanoid pathway, the FAH1 and F3H genes, in Arabidopsis thaliana. Molecular Biology and Evolution 18: 1-9

Arnerup J, Nemesio-Gorriz M, Lunden K, Asiegbu FO, Stenlid J, Elfstrand M (2013) The primary module in Norway spruce defence signalling against H. annosum s.l. seems to be jasmonate-mediated signalling without antagonism of salicylate-mediated signalling. Planta 237: 1037-1045

Arnerup J, Swedjemark G, Elfstrand M, Karlsson B, Stenlid J (2010) Variation in growth of Heterobasidion parviporum in a full-sib family of Picea abies. Scandinavian Journal of Forest Research 25: 106-110

Bentsink L, Jowett J, Hanhart CJ, Koornneef M (2006) Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 103: 17042-17047

Bodles WJA, Beckett E, Woodward S (2007) Responses of sitka spruce of different genetic origins to inoculation with Heterobasidion annosum: lesion lengths, fungal growth and development of the lignosuberized boundary zone. Forest Pathology 37: 174-186

Bouille M, Bousquet J (2005) Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer Picea (Pinaceae): Implications for the long-term maintenance of genetic diversity in trees. American Journal of Botany 92: 63-73

Brignolas F, Lacroix B, Lieutier F, Sauvard D, Drouet A, Claudot AC, Yart A, Berryman AA, Christiansen E (1995) Induced responses in phenolic metabolism in two Norway spruce clones after wounding and inoculations with Ophiostoma polonicum, a bark beetle-associated fungus. Plant Physiology 109: 821-827

Brignolas F, Lieutier F, Sauvard D, Christiansen E, Berryman AA (1998) Phenolic predictors for Norway spruce resistance to the bark beetle Ips typographus (Coleoptera : Scolytidae) and an associated fungus, Ceratocystis polonica. Canadian Journal of Forest Research 28: 720-728

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter 11: 113-116

Chen J, Kallman T, Gyllenstrand N, Lascoux M (2010) New insights on the speciation history and nucleotide diversity of three boreal spruce species and a Tertiary relict. Heredity 104: 3-14

Danielsson M, Lunden K, Elfstrand M, Hu J, Zhao T, Arnerup J, Ihrmark K, Swedjemark G, Borg-Karlson AK, Stenlid J (2011) Chemical and transcriptional responses of Norway spruce genotypes with different susceptibility to Heterobasidion spp. infection. BMC Plant Biology 11

de Meaux J, Goebel U, Pop A, Mitchell-Olds T (2005) Allele-specific assay reveals functional variation in the chalcone synthase promoter of Arabidopsis thaliana that is compatible with neutral evolution. Plant Cell 17: 676-690

Estell DA, Graycar TP, Wells JA (1985) Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. Journal of Biological Chemistry 260: 6518-6521

Filatov DA, Charlesworth D (1999) DNA polymorphism, haplotype structure and balancing selection in the Leavenworthia PgiC locus. Genetics 153: 1423-1434

Fossdal CG, Nagy NE, Hietala AM, Kvaalen H, Slimestad R, Woodward S, Solheim H (2012) Indications of heightened constitutive or primed host response affecting the lignin pathway transcripts and phenolics in mature Norway spruce clones. Tree Physiology 32: 1137-1147

Gonzalez-Martinez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB (2006) DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in Pinus taeda L. Genetics 172: 1915-1926

Hall DE, Robert JA, Keeling CI, Domanski D, Lara Quesada A, Jancsik S, Kuzyk MA, Hamberger B, Borchers CH, Bohlmann J (2011) An integrated genomic, proteomic and biochemical analysis of (+)-3-carene biosynthesis in Sitka spruce (Picea sitchensis) genotypes that are resistant or susceptible to white pine weevil. The Plant Journal 65: 936-948

Hammerbacher A (2011) Biosynthesis of polyphenols in Norway spruce as a defense strategy against attack by the bark beetle associated fungus Ceratocystis polonica. Doctoral Thesis. Max Planck Institute for Chemical Ecology, Jena, Germany

Hammerbacher A, Paetz C, Wright LP, Fischer TC, Bohlmann J, Davis AJ, Fenning TM, Gershenzon J, Schmidt A (2014) Flavan-3-ols in Norway spruce: biosynthesis, accumulation, and function in response to attack by the bark beetle-associated fungus Ceratocystis polonica. Plant Physiology 164: 2107-2122

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Research 27: 297-300

Johnson ET, Ryu S, Yi HK, Shin B, Cheong H, Choi G (2001) Alteration of a single amino acid changes the substrate specificity of dihydroflavonol 4-reductase. Plant Journal 25: 325-333

Kayihan GC, Huber DA, Morse AM, White TL, Davis JM (2005) Genetic dissection of fusiform rust and pitch canker disease traits in loblolly pine. Theoretical and Applied Genetics 110: 948-958

Kuhlman EG, Amerson HV, Jordan AP, Pepper WD (1997) Inoculum density and expression of major gene resistance to fusiform rust disease in loblolly pine. Plant Disease 81: 597-600

Lehtijärvi A, Aday A, Dogmus-Lehtijärvi H (2009) Turkish Heterobasidion abietinum is pathogenic to inoculated Abies nordmanniana ssp. nordmanniana and ssp. bornmülleriana. Forest Pathology 39: 200-209

Lieutier F, Brignolas F, Sauvard D, Yart A, Galet C, Brunet M, van de Sype H (2003) Intra- and inter-provenance variability in phloem phenols of Picea abies and relationship to a bark beetle-associated fungus. Tree Physiology 23: 247-256

Lind M, Dalman K, Stenlid J, Karlsson B, Olson A (2007) Identification of quantitative trait loci affecting virulence in the basidiomycete Heterobasidion annosum s.l. Current Genetics 52: 35-44

Lind M, Kallman T, Chen J, Ma XF, Bousquet J, Morgante M, Zaina G, Karlsson B, Elfstrand M, Lascoux M, Stenlid J (2014) A Picea abies linkage map based on SNP markers identifies QTLs for four aspects of resistance to Heterobasidion parviporum infection. PLoS ONE 9

Mageroy MH, Parent G, Germanos G, Giguere I, Delvas N, Maaroufi H, Bauce E, Bohlmann J, Mackay JJ (2015) Expression of the ß-glucosidase gene Pgßglu-1 underpins natural resistance of white spruce against spruce budworm. The Plant Journal 81: 68-80

Mauge C, Granier T, d'Estaintot BL, Gargouri M, Manigand C, Schmitter JM, Chaudiere J, Gallois B (2010) Crystal structure and catalytic mechanism of leucoanthocyanidin reductase from Vitis vinifera. Journal of Molecular Biology 397: 1079-1091

Narsai R, Howell KA, Millar AH, O'Toole N, Small I, Whelan J (2007) Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. Plant Cell 19: 3418-3436

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends in Plant Science 9: 325-330

Neale DB, Wegrzyn JL, Stevens KA, Zimin AV, Puiu D, Crepeau MW, Cardeno C, Koriabine M, Holtz-Morris AE, Liechty JD, Martinez-Garcia PJ, Vasquez-Gross HA, Lin BY, Zieve JJ, Dougherty WM, Fuentes-Soriano S, Wu LS, Gilbert D, Marcais G, Roberts M, Holt C, Yandell M, Davis JM, Smith KE, Dean JFD, Lorenz WW, Whetten RW, Sederoff R, Wheeler N, McGuire PE, Main D, Loopstra CA, Mockaitis K, deJong PJ, Yorke JA, Salzberg SL, Langley CH (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. Genome Biology 15XNystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A, Vicedomini R, Sahlin K, Sherwood E, Elfstrand M, Gramzow L, Holmberg K, Hallman J, Keech O, Klasson L, Koriabine M, Kucukoglu M, Kaller M, Luthman J, Lysholm F, Niittyla T, Olson A, Rilakovic N, Ritland C, Rossello JA, Sena J, Svensson T, Talavera-Lopez C, Theissen G, Tuominen H, Vanneste K, Wu ZQ, Zhang B, Zerbe P, Arvestad L, Bhalerao R, Bohlmann J, Bousquet J, Gil RG, Hvidsten TR, de Jong P, MacKay J, Morgante M, Ritland K, Sundberg B, Thompson SL, Van de Peer Y, Andersson B, Nilsson O, Ingvarsson PK, Lundeberg J, Jansson S (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497: 579-584

Oliva J, Rommel S, Fossdal CG, Hietala AM, Nemesio-Gorriz M, Solheim H, Elfstrand M (2015) Transcriptional responses of Norway spruce (Picea abies) inner sapwood against Heterobasidion parviporum. Tree Physiology 35: 1007-1015

Pavy N, Gagnon F, Rigault P, Blais S, Deschenes A, Boyle B, Pelgas B, Deslauriers M, Clement S, Lavigne P, Lamothe M, Cooke JEK, Jaramillo-Correa JP, Beaulieu J, Isabel N, Mackay J, Bousquet J (2013) Development of high-density SNP genotyping arrays for white spruce (Picea glauca) and transferability to subtropical and nordic congeners. Molecular Ecology Resources 13: 324-336

Pavy N, Pelgas B, Laroche J, Rigault P, Isabel N, Bousquet J (2012) A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. BMC Biology 10

Petit P, Granier T, d'Estaintot BL, Manigand C, Bathany K, Schmitter JM, Lauvergeat V, Hamdi S, Gallois B (2007) Crystal structure of grape dihydroflavonol 4-reductase, a key enzyme in flavonoid biosynthesis. Journal of Molecular Biology 368: 1345-1357

Pfeiffer J, Kuhnel C, Brandt J, Duy D, Punyasiri PA, Forkmann G, Fischer TC (2006) Biosynthesis of flavan 3-ols by leucoanthocyanidin 4-reductases and anthocyanidin reductases in leaves of grape (Vitis vinifera L.), apple (Malus x domestica Borkh.) and other crops. Plant Physiology and Biochemistry 44: 323-334

Plantegenet S, Weber J, Goldstein DR, Zeller G, Nussbaumer C, Thomas J, Weigel D, Harshman K, Hardtke CS (2009) Comprehensive analysis of Arabidopsis expression level polymorphisms with simple inheritance. Mol Syst Biol 5: 242

Porth I, White R, Jaquish B, Alfaro R, Ritland C, Ritland K (2012) Genetical genomics identifies the genetic architecture for growth and weevil resistance in spruce. PloS ONE 7

Quesada T, Gopal V, Cumbie WP, Eckert AJ, Wegrzyn JL, Neale DB, Goldfarb B, Huber DA, Casella G, Davis JM (2010) Association mapping of quantitative disease resistance in a natural population of loblolly pine (Pinus taeda L.). Genetics 186: 677-U336

Quesada T, Resende MFR, Munoz P, Wegrzyn JL, Neale DB, Kirst M, Peter GF, Gezan SA, Nelson CD, Davis JM (2014) Mapping fusiform rust resistance genes within a complex mating design of loblolly pine. Forests 5: 347-362

Rigault P, Boyle B, Lepage P, Cooke JEK, Bousquet J, MacKay JJ (2011) A white spruce gene catalog for conifer genome analyses. Plant Physiology 157: 14-28

Sainsbury F, Thuenemann EC, Lomonossoff GP (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnology Journal 7: 682-693

Stafford HA, Lester HH (1984) Flavan-3-ol biosynthesis: The conversion of (+)-dihdydroquercetin and flavan-3,4-cis-diol) leucocyanidin) to (+)-catechin by reductases extracted from cell suspension cultures of Douglas fir. Plant Physiology 75: 155-155

Stenlid J (1985) Population structure of Heterobasidion annosum as determined by somatic incompatibility, sexual incompatibility and isoenzyme patterns. Canadian Journal of Botany 63: 2268-2273

Stenlid J (1987) Controlling and predicting the spread of Heterobasidion annosum from infected stumps and trees of Picea abies. Scandinavian Journal of Forest Research 2: 187-198

Stergachis AB, MacLean B, Lee K, Stamatoyannopoulos JA, MacCoss MJ (2011) Rapid empirical discovery of optimal peptides for targeted proteomics. Nature Methods 8: 1041-1043

Swedjemark G, Stenlid J, Karlsson B (1997) Genetic variation among clones of Picea abies in resistance to growth of Heterobasidion annosum. Silvae Genetica 46: 369-374

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729

Verne S, Jaquish B, White R, Ritland C, Ritland K (2011) Global transcriptome analysis of constitutive resistance to the white pine weevil in spruce. Genome Biol Evol 3: 851-867

Wachowiak W, Balk PA, Savolainen O (2009) Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (Pinus sylvestris L.). Tree Genetics & Genomes 5: 117-132

Wang X, Culver JN (2012) DNA binding specificity of ATAF2, a NAC domain transcription factor targeted for degradation by Tobacco mosaic virus. BMC Plant Biology 12: 157

Warren RL, Keeling CI, Yuen MM, Raymond A, Taylor GA, Vandervalk BP, Mohamadi H, Paulino D, Chiu R, Jackman SD, Robertson G, Yang C, Boyle B, Hoffmann M, Weigel D, Nelson DR, Ritland C, Isabel N, Jaquish B, Yanchuk A, Bousquet J, Jones SJ, MacKay J, Birol I, Bohlmann J (2015) Improved white spruce (Picea glauca) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. The Plant Journal 83: 189-212

Wilcox PL, Amerson HV, Kuhlman EG, Liu BH, O'Malley DM, Sederoff RR (1996) Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. Proceedings of the National Academy of Sciences 93: 3859-3864

Woodward S (1998) Heterobasidion annosum : biology, ecology, impact, and control. CAB International, Wallingford, Oxon, UK ; New York

	Latitude	Longitude
Lugnet common garden	59°38'	17°31'
Family provenances		
Sör-Amsberg	60°28'	15°23'
Tuusula	60°21'	25°35'
Punkaharju	61°34'	29°23'
Petroskoi	63°22'	34°19'

Supplemental material 1: GPS coordinates for the common garden and for the four populations that were used in this study.

Family	Population	Origin
41	Punkaharju	Finland
51	Punkaharju	Finland
52	Punkaharju	Finland
68	Tuusula	Finland
69	Tuusula	Finland
120	Sör-Amsberg	Sweden
127	Sör-Amsberg	Sweden
134	Sör-Amsberg	Sweden
157	Petroskoi	Russia
160	Petroskoi	Russia
162	Petroskoi	Russia
174	Petroskoi	Russia
178	Petroskoi	Russia
183	Petroskoi	Russia

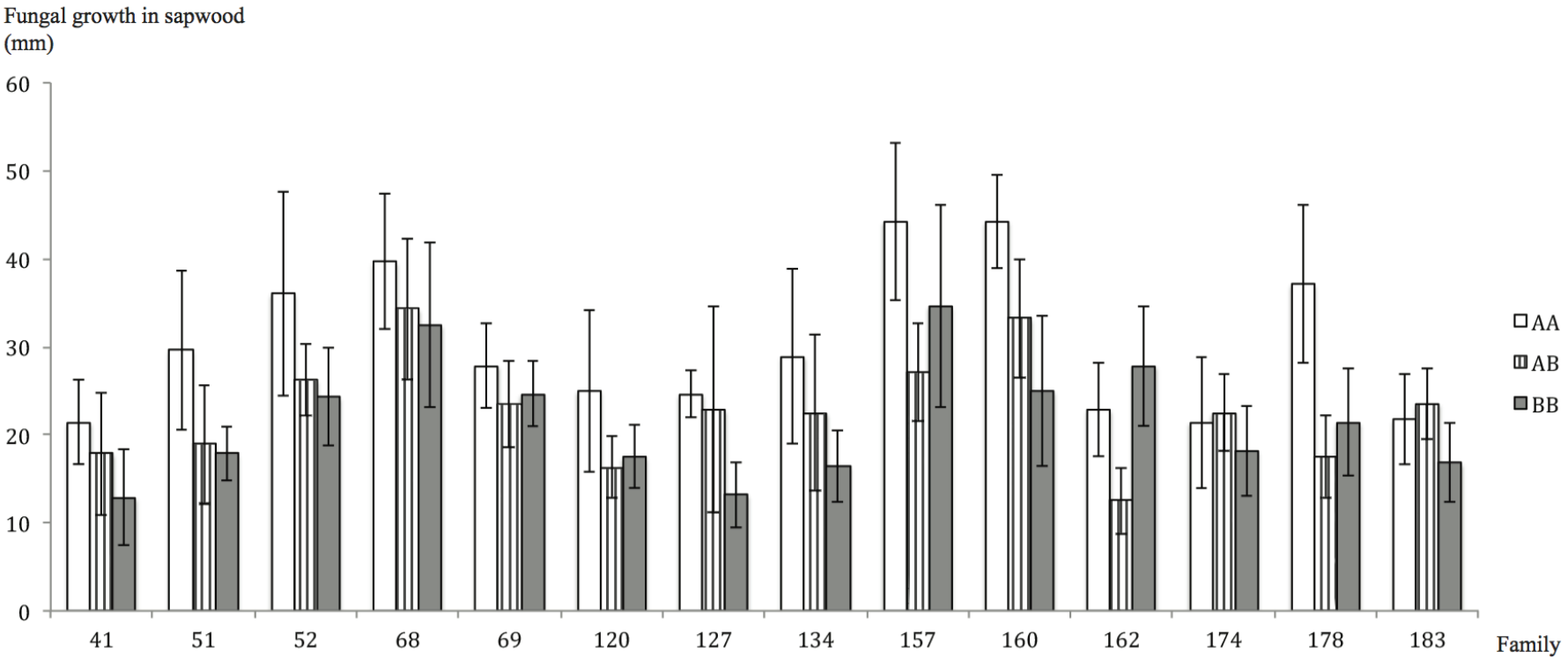
Supplemental material 2: List of half-sib families with their provenance and country of origin.

Primer pair	Forward sequence	Reverse sequence
Genomic fragment amplification		
PaLAR3 _a1	GAAATCGAATTCAGGATGTTCA	CGGGTCGAGCACAATAAGAT
PaLAR3 _a2	GAAATCGAATTCAGGATGTTCT	GGGGTCGAGCACAATAAGAT
PaLAR3 _b1	GCTGAAATGGGCCTATCTTATAC	CCGCCATATGGGAGAAGAA
PaLAR3 _b2	GCTGAAATGGGCCTATCTTATAG	CCGCCATATGGGAGAAGAA
PaLAR3 _c1	TGCCAACAAATCTTGCTCAG	AACGTTCTATTGGCCAGTCC
PaLAR3 _c2	TGCCAACCAATCTTGCTCAG	AACGTTCTATTGGCCAGTCC
PaLAR3 _d	AGCTTCTCATATTTGCCCTGA	CCGGCGTCCATATTCTTTAT
PaLAR3 _e	AACCTAAAACTGGGGGATCG	GAGGTCAACGAGGGTGAAAA
Promoter primers		
PaLAR3 -500	GGAGTATGAGTTAGCTCAACCAAC	CAGTCCGGCCATCCCTATTTC
PaLAR3 -1000	AAGGGAGTAGGTTAAATTGGGA	TTGGTTGAGCTAACTCATGCTC
PaLAR3 -1500	TAATCTGTTTCCATGTGCCTTG	TCCAGTACAGTCCTATTTTGCC
qRT-PCR primers		
PaLAR3_A	CGGACATTGTGACACGAAAC	CGGAGTTTATACCCGTTCCA
PaLAR3_B	CGGACATTGTGACACGAAAC	AACCATTCACGGTCATGATGTA

Supplemental material 3: List of primers that were used for the genomic sequencing of *PaLAR3* and allele-specific qRT-PCR.

Parent ion	Product ion	Dwell time	Peptide	DP	CE	CXP
616,872	794,416	20	LAR_A.ILVIGATGYIGR. +2y8.light	76,1	30,9	13
616,872	737,394	20	LAR_A.ILVIGATGYIGR. +2y7.light	76,1	30,9	13
616,872	666,357	20	LAR_A.ILVIGATGYIGR. +2y6.light	76,1	30,9	13
461,748	722,408	20	BSA.AEFVEVTK.+2y6.light	64,8	22,1	13
461,748	575,34	20	BSA.AEFVEVTK.+2y5.light	64,8	22,1	13
507,813	785,513	20	BSA.QTALVELLK.+2y7.light	68,1	24,7	13
507,813	601,392	20	BSA.QTALVELLK.+2y5.light	68,1	24,7	13

Supplemental material 4: Multiple reaction monitoring (MRM) transitions for peptides analyzed to quantify LAR and BSA by liquid chromatography-tandem mass spectrometry (DP, declustering potential; CE, collision energy; CXP, collision cell exit potential).



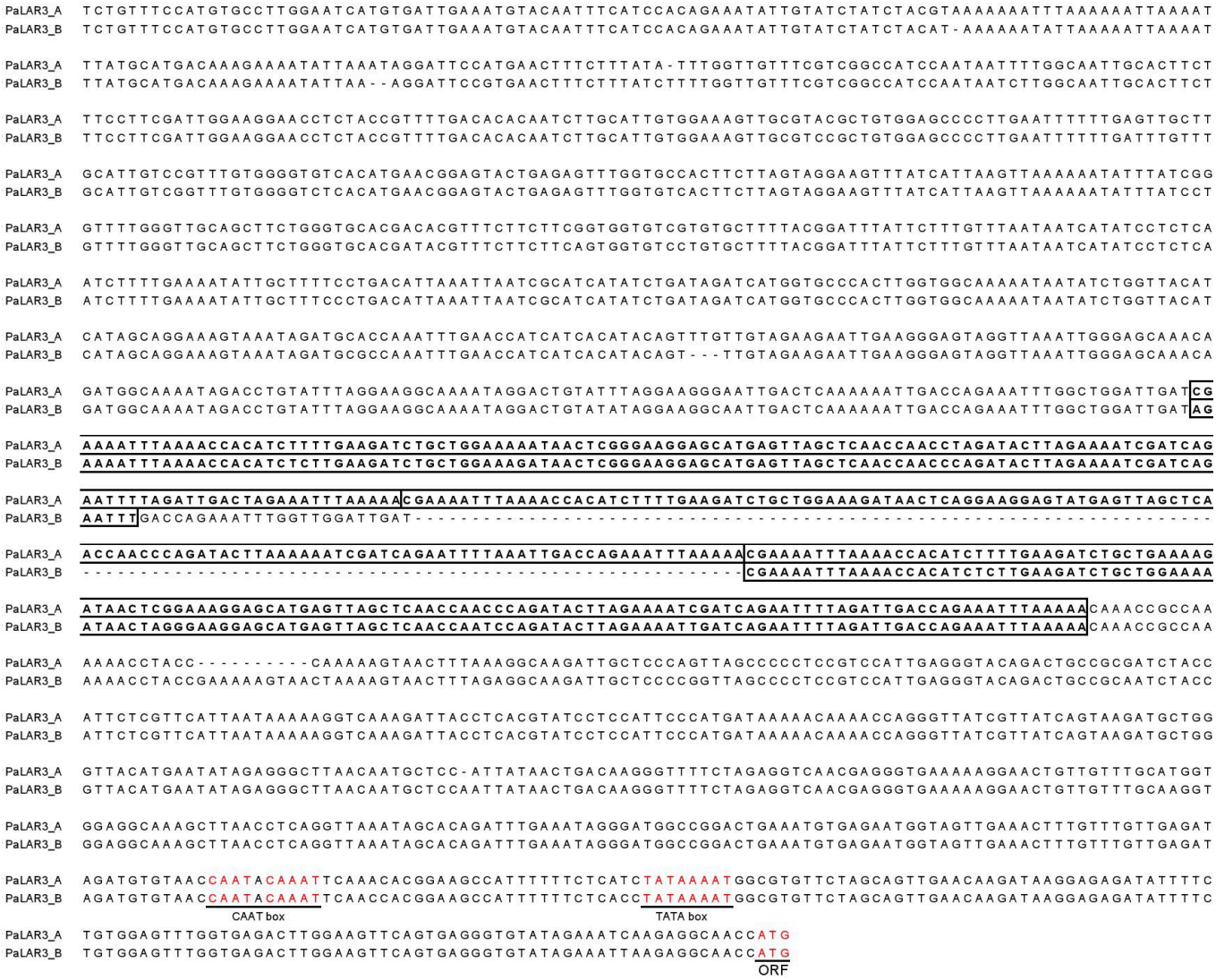
Supplemental material 5: Average fungal growth in sapwood and standard deviation for each one of the genotypes in the 14 families that were studied.

	LAR_A	LAR_B	Function	Comments
ACGTATERD1	3	2	Ethiolation & dehydration	
ABRELATERD1	2	1	Ethiolation & dehydration	
ACGTABREMOTIFA2OSEM	1	0	ABA response	
AACACOREOSGLUB1	3	2	Endosperm-specific	
BOXLCOREDCPAL	4	2	PAL-regulation	
CBFHV	1	2	Low temperature & dehydration	
E2FCONSENSUS	1	2	Cell cycle	
ANAERO1CONSENSUS	5	6	Anaerobic	
ARR1AT	19	21	Cytokinin response	
BOXIINTPATPB	1	2	Plastid expre	The one in LARA al
CAATBOX1	24	28		
CARGNCAT	1	0	Giberellin	
CCAATBOX1	5	6	Increases promoter activity	
CIACADIANLELHC	1	2	Circadian	
EBOXBNNAPA	7	8	Phenylpropanoid R2R3, HZIP and bHLH	
GT1GMSCAM4	3	4	Pathogen	
GTGANTG10	8	11	Pollen	
LTRE1HVBLT49	0	1	Low emperature	
MYBST1	2	3	MYB TF	
MYCCONSENSUSAT	7	8	MYC, low temperature and dehydration	
NODCON2GM	1	3	Infected cells	Same motif as OSE2
PREATPRODH	3	2	Hypoosmolarity and proline dehydroge	
RBCSCONSENSUS	0	1	ribulose-1,5-bisphosphate carboxylase	
ROOTMOTIFTAPOX1	7	8	rolD gene Agrobacterium	
SREATMSD	0	1	stem decapitation sugar repressive mot	
SURECOREATSULTR11	1	2	sulfur-responsive element and auxin	
TL1ATSAR	0	1	NPR1-responsive ER-resident genes	
WBBOXPCWRKY1	1	2	WRKY	
2SSEEDPROTBANAPA	1	0	Storage	
AACACOREOSGLUB1	3	2	Endosperm-specific	
ABRERATCAL	1	0	Ca2+ responsive	
ACGTABOX	1	0	Sugar drop	
BOX1PSGS2	1	0	Glutamine synthase	
BOXLCOREDCPAL	4	2	MYB and PAL	
CANBNNAPA	1	0	endosperm-specific	
CAREOSREP1	1	0	Giberellic acid	
CURECORECR	4	3	Copper-response element	
DOFCOREZM	25	22	Zinc finger binding	
EECCRCAH1	3	1	MYB and CO2 response	
ERELEE4	2	1	Ethylene and stress	

GADOWNAT	1	0	GA and germination
MYB1AT	6	5	MYB, MYC and ABA
MYB2CONSENSUSAT	3	2	MYB, MYC ar Two of LARA also carry the
MYBPLANT	6	4	MYB and phenylpropaniod
MYBPZM	4	3	Floral pigmentation and flavonoid path
NODCON1GM	7	4	Root nodule Same motif as OSE1ROOTI
PREATPRODH	3	2	Hipoosmolarity
QARBNEXTA	1	0	Wounding and tensile stress
RAV1AAT	3	2	Leaf and root development
RYREPEATBNNAPA	2	1	ABA and seed specific
SEF3MOTIFGM	4	3	Embryo development
SORLIP1AT	2	1	Cotyledon and root
T/GBOXATPIN2	1	0	JA, MYC and bHLH
TATABOXOSPAL	2	1	PAL promoter
WBOXATNPR1	8	7	WRKY SA and pathogen-induced
WBOXHVISO1	2	1	WRKY sugar responsive
WBOXNTERF3	7	6	ERF and wounding

	1
	1
	1
	1
	2
	-1
	-1
	-1
	-2
.so encodes a BOX1PSGS2 motif(glutam:	-1
	-4
	1
	-1
	-1
	-1
	-1
	-3
	-1
	-1
ı	-1
?ROOTNODULE	-2
nase	1
	-1
	-1
:if	-1
ı	-1
	-1
	-1
	1
	1
	1
	1
	1
	2
	1
	1
	1
	3
	2
	1

	1
	1
e motif MYBAT2 (MYB and dehydration) and only one	1
	2
way	1
NODULE	3
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1
	1



Supplemental material 6: Sequence of 1.5 kb promoter of the promoter region of *PaLAR3A* and *PaLAR3B.* Black boxes indicate a 129 bp repetition and red characters indicate the position of the CAAT and TATA boxes, and the ORF of the gene.