2019-2020

Master Agronomie, Environnement, Territoires, Paysage, Forêt (AETPF)

Major

Forests and their Environment

Genetics of quantitative resistance and its reaction norms in the natural pathosystem of *Populus nigra* and *Melampsora larici-populina* using Genome-Wide Association Study (GWAS)

Firza RIANY

Master thesis, defended at Nancy on 29 September 2020

Supervisor in host institution: Véronique JORGE
Chargeé de Recherche, Directrice Adjointe

Academic referent : Yves ERHART, Lecturer, AgroParisTech

Acknowledgement

The author wishes to thank Véronique Jorge, Yves Erhart, and Leopoldo Sanchez for their support and assistance throughout the production of this report, Celine Ridel, Patrick Poursat, Anita Rinfray, and Remy Gobin for providing the support and care in the laboratory during the experiment, and BioForA unit of INRAE Orléans for their hospitality during the internship.

Contents

1. Introduction

1.1. Management of *Populus* spp.

Populus spp. are a deciduous tree species with a wide natural distribution covering the northern hemisphere until the tropics. The mature tree can reach 20 – 40 m in heights and 60 – 100 cm in diameter with straight, well-formed stems (Dickmann and Kuzovkina, 2014). These characteristics of *Populus* spp. made it an interesting choice for wood production. In Europe, *Populus nigra* is the species that is naturally occurring in the nature. It is a riparian tree species, occurring predominantly on floodplains in mixed forests, especially with *Larix* spp. (Gérard *et al.*, 2006).

Populus spp. has been an attractive subject to breeding programs to produce superior hybrid cultivars, such as a breeding program in Europe that crossed *Populus deltoides* (native to North America) with *Populus trichocarpa* (native to West America) creating *Populus x interamericana* cultivars with high level of heterosis for growth. Another common cultivar is *Populus x euramericana,* created by crossing *Populus deltoides* and *P. nigra*. Besides its ability to be crossed intra-species, *Populus* spp. can also be easily cloned, thus protecting the genetic diversity and the superiority of the parents in cloned off-springs. The cultivars are usually planted in a monoclonal stand for its wood and resistance, a typical practice in northern France (Lefèvre *et al*. 1998; Dowkiw and Bastien 2004; Dowkiw *et al.,* 2003; Dickmann and Kuzovkina, 2014).

1.2. Interaction dynamic between *Populus* spp. and *Melampsora laricii-populina* Kleb.

Populus spp. is frequently attacked by pest and diseases, disorienting its potential growth. *Melampsora laricii-populina* Kleb. (*Mlp*) is the pathogen that reduces the growing stock of *Populus* spp. through decreased photosynthesis efficiency and early defoliation (Dowkiw *et al.* 2003; Gérard *et al.*, 2006; Aylott *et al.*, 2008; Benetka *et al.*, 2011; Polle *et al.*, 2013; Eberl *et al*., 2018). *Mlp* has an efficient reproduction cycle, alternating between sexual reproduction in *Larix* spp. producing the aeciospores, and asexual multiplication in *Populus* spp. producing rust spots consisted of urediniospores. The sexual phase provides *Mlp* with a fresh pool of gene diversity through constant genes recombination, while asexual multiplication amplifies the strength of the most adaptive individuals. These abilities are the fuel for *Mlp* adaptive evolution, providing it with a great advantage to overcome the challenges from poplar resistance (Becheler *et al.,* 2016; Gérard *et al.*, 2006; Persoons *et al.,* 2017). Moreover, gene and genotypic diversity of *Mlp* can always be enriched through urediniospores migration from other *Populus* spp. stands, thus increasing the chances for the *Mlp* to rapidly evolve and adapt to *Populus* spp. (Barrès *et. al.,* 2012).

Intensive human management of *Populus* spp. has opened the way for *Mlp* to evolve vigorously, breaking down *Populus* spp. resistance in less than 20 years. In 1994, the cultivar *P. x interamericana* 'Beaupré' that carries resistance gene R7 was planted widely

in northern France. The R7 gene was putting a strong selection pressure on *Mlp* populations, leading them to evolve by mutating their avirulence gene (Avr7) to alternative virulence (vir7) (Persoons *et al.,* 2017).

Before the appearance of vir7, there were already six virulences (vir1 – vir6) in *Mlp* populations. These virulences are commonly working together to generate the pathogenicity in *Populus* – *Mlp* interaction, creating a pathotype (i.e. combination of virulences). There was a complex pathotype pre-vir7 consisted of 1 – 3 – 4 – 5 virulences. This pathotype was then amplified by the addition of vir7 when it appeared, creating a more complex pathotype that can create a massive destruction on *Populus* spp.. However, this is particularly the case in *Populus* cultivated stands while neither the complex pathotype nor vir7 individuals were observed in natural *P. nigra* stands so far (Persoons *et al.,* 2017; Gérard *et al.*, 2006).

There is a difference in co-evolution progress between the two stands that make the complex pathotype is present in one stand and not in the other. Cultivated *Populus* exerts strong selection pressure on rust as previously described, whereas natural *P. nigra* is generating a stable *Populus* – *Mlp* interaction at its co-evolving loci. Studies showed that *P. nigra* is perennial to the oscillation of resistance alleles, thus retaining the possible selection pressure that can cause the *Mlp* to evolve, and furthermore maintaining the long-term co-evolution between the two. Moreover, these differences in *Populus* characteristics have led to different evolution in the genetics of *Mlp* pathogenicity in these two stands (Persoons, *et al.,* 2017). Sexual recombination often occurs between individuals with the same pathotypes. In cultivated stands, *Mlp* individuals with high genetic diversity in complex pathotypes are favored, thus the genetic diversity for complex pathotypes keeps increasing through sexual recombination. In natural *P. nigra* stands, while the *Mlp* has only vir2 and vir4, stable *Populus* – *Mlp* interaction does not create the urgency for them to evolve these virulences (Gerard *et al.*, 2006).

1.3. The search of durable resistance in *Populus* spp.

Since the resistance exerted by current *Populus* spp. cultivars are rapidly being overcame by *Mlp,* the search for durable resistance is a challenge. Natural population of *P. nigra* provides a reservoir of genetic diversity that can be explored in the search for durable resistance. Beside that *P. nigra* is co-evolving with *Mlp* in their natural population, Bastien *et. al.,* (2015) indicated that *P. x euramericana* (a breed of *P. deltoides* and *P. nigra)* showed better qualitative resistance than *P. x interamericana*. It indicates that *P. nigra* may support or improve the resistance exerted by *P. deltoides*. Moreover, it also provides a starting ground to study *P. nigra* as a potential parent in *Populus* spp. breeding strategies for durable resistance.

There are methods to carry on the search of durable resistance in *Populus* spp. A study of *P. nigra* genetic diversity should give the information for its potential. Faivre-Rampant *et. al.,* (2016) had re-sequenced *P. nigra* individuals that were collected from their natural population in Western Europe. The sequence resulted in 10k SNPs that relate to rust resistance, wood properties, water-use efficiency, bud phenology and natural genetic

diversity across the genome. This set of sequence was used to genotype 880 *P. nigra* from the same natural populations. The genotyping showed *P. nigra* genetic diversity for the above-mentioned traits. The diversity also reflected the geographical diversity of *P. nigra.* The result provides the potential of *P. nigra* genetic diversity to be explored for rust resistance.

The evaluation of *P. nigra* resistance in the field and laboratory experiments had been conducted. Field resistance was evaluated by scoring the susceptibility of 1 141 *P. nigra* clones collected from the same natural populations as in Faivre-Rampant *et. al.,* (2016). Bastien *et. al.,* (2015) showed that *P. nigra* in Ticino, Basento, and Paglia from Italy as well as in Durance, Ramieres, Nohedes, and Adour from France had high field resistance. The clones from Adour and Nohedes showed the highest field resistance.

Field resistance can be complemented by evaluating the development of epidemiological traits that explain *P. nigra* quantitative resistance to specific strains in laboratory experiments. Quantitative resistance in laboratory is studied by breaking down epidemiological traits into latent period (LP), uredinia number (UN), and uredinia size (US). Combining field resistance with laboratory experiment, Bastien *et. al.,* (2015) showed that field resistance was genetically controlled and was explained by latent period and uredinia size. Furthermore, clones from Ramieres, Val Allier and Loire (France) showed higher variation in latent period and uredinia size than the clones from Adour and Nohedes.

There is a common thought that qualitative resistance is strain-specific, thus is easy to be overcame by *Mlp* strain as the genes in *Mlp* is interacting in gene-by-gene way with *P. nigra* (Dowkiw and Bastien 2004; do Vale *et. al.,* 2001). Therefore, some studies also mentioned that quantitative resistance is more durable. However, studies on quantitative resistance (Bastien *et. al.,* 2015) of 26 *P. nigra* clones showed that quantitative resistance of these *P. nigra* clones was highly strain-specific as its qualitative resistance.

A QTL analysis also showed that different strains triggered different loci responsible for resistance in *P. nigra*. ElMalki *et. al.,* (2013) showed that QTL LGI was responsible for *P. nigra* quantitative resistance to one strain among the 12 strains with the code name OSSAS 28.5.6. One year later, a study by Albert (2013) found that the strain OSSAS 28.5.6 had overcome the resistance in QTL LGI. These follow up studies showed different behaviors of *P. nigra* individuals when they are infected by different *Mlp* strains in terms of quantitative resistance.

1.4. Aims of the study

Finally, using the genetic variation of *P. nigra* that has been studied in Faivre-Rampant *et. al.,* (2016), this study was aimed to find the genetic control of quantitative resistance in 154 *P. nigra* genotypes that were collected from 12 natural populations in Western Europe. This study conducted a genome-wide association study (GWAS) to find the candidate genes for quantitative resistance. The quantitative resistance was studied in terms of latent period, uredinia number, and uredinia size in a laboratory experiment after inoculating 154 *P. nigra* genotypes with three *Mlp* strains (09AX27, 93JE3, and P72). This

study was then aimed to evaluate the strain specificity of *P. nigra* genotypes and *Mlp* strains through their reaction norm.

2. Material and method

2.1. Plant materials

Populus nigra material was consisted of 168 cloned genotypes originated from 12 natural populations in Western Europe (Figure 1). These populations were selected because the poplar individuals showed high variability in their resistance to rust infection in the field, meaning that the poplar individuals in these populations have variation that can be selected. Within 168 genotypes, there are seven genotypes that were used as parents in a breeding program in France, one genotype as control ('Robusta' cultivar), and eight genotypes as discriminants.

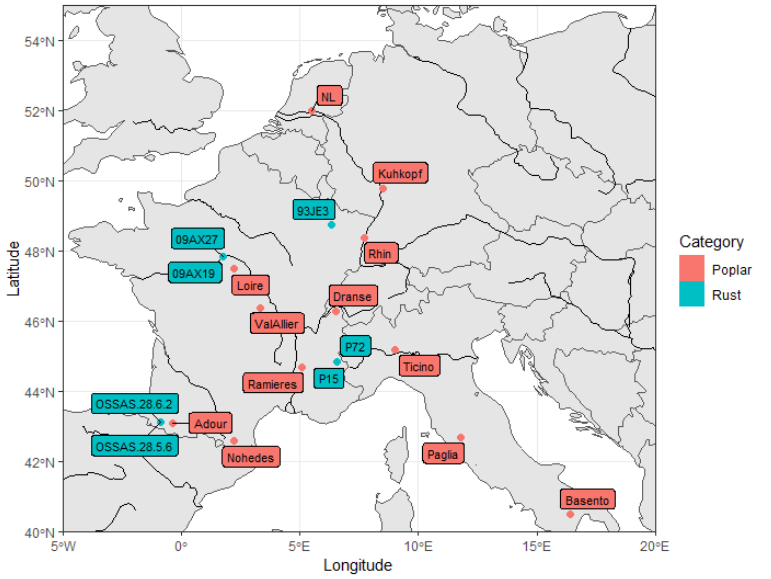


Figure 1 The locations of P. nigra population (red) and M. larici-populina population (green) where the genotypes and the strains where collected, respectively.

The seven genotypes that have been used as parents in the breeding program was included to identify interesting crosses in terms of their reaction to the three strains for future analysis. There were also eight genotypes that were used as discriminants to the *Mlp* strains, and one susceptible genotype 'Robusta' that was used as a control. These discriminants were used to identify the pathotypes of the *Mlp* strains as explained in chapter 2.2. These genotypes were excluded for analysis, and for further description, the study will focus on 158 genotypes.

The cloned genotypes were grown in a greenhouse maintained by the institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE) in Orléans, France from May to August 2020. They were grown from cuttings in 4-liter pots containing 40% white sphagnum peat, 50% frozen black sphagnum peat, 10% wood fiber, 20 kg/m^3 milled clay, 1.2 kg/m^3 PHMix 14.16.18, and 0.15/m^3 humectant agent in rust-free greenhouse conditions. For the tending, the plants were watered and fertilized daily with the ratio of 15:10:15 solution (1 g/liter). The clones were planted in 4 blocks in which each genotype is represented by one plant.

2.2. Fungal material

To assess the quantitative resistance of 158 genotypes, three *Mlp* strains were used as the inoculum (Table 1). Strains 09AX27, 93JE3, and P72 were used because they are

originated from different geographical places in France (Prelles, Champenoux, and Saint Ay; Figure 1), two are present in *P. nigra* natural populations and one is present in a plantation of *Populus* cultivars. Their variability in interaction pattern with *P. nigra* genotypes is another reason of choosing them and this has been described in another study (ElMalki *et. al.,* 2013; Bastien *et. al.,* 2015). Four other strains were initially included in the study representing a larger diversity of the pathogen (represented in Figure 1), but were not used in the present study.

The variability within the *Mlp* strains was described in terms of virulences. A strain can have only one virulence or a combination of virulences. For the latter term, the combination will create a pathotype. The identification of virulences was carried out following the method in Gérard, *et. al.,* (2006). The eight genotypes that were used as discriminants in this study were used to identify the virulences of the *Mlp* strains. This is due to the reason that these discriminants carry their own resistance factor that is interacting with the virulence gene in *Mlp*. The discriminants are *Populus x euramericana* 'Ogy' (vir 1), *Populus x jackii* 'Aurora' (vir 2), *Populus x euramericana* 'Brabantica' (vir 3), *Populus x interamericana* 'Unal' (vir 4), *Populus x interamericana* 'Rap' (vir 5), *P. deltoides* '87B12' (vir 6), *Populus interamericana* 'Beaupré' (vir 7), and *Populus x interamericana* 'Hoogvorst' (vir 8) (Table 1).

The identification of *Mlp* virulences was carried out by first inoculating the *Mlp* strains on the leaves of the discriminants. When a discriminant Z showed a symptom (such as sporulating uredinia) within 14 days after being inoculated with strain A, its associated virulence (i.e. vir Z) explains the virulence carried by the strain A. In this study, inoculating the discriminants with strain 09AX27 created symptoms on the discriminants associated with vir 3 and 4, thus creating pathotype 3-4. Strain 93JE3 has pathotype 2-3-4 and strain P72 has pathotype 3-4.

Table 1 Geographical origin and the pathotypes of the selected Melampsora-larici populina strains

Strains	Origin		Isolation	Pathotype
	City	Department		
09AX27	Saint Ay	Loiret	P. nigra natural population	3-4
93JE3[a]	Champenoux	Grand-Est	P. deltoides x P. nigra cv 'Blanc du Poitou' [b]	2-3-4
P72	Prelles	Top of Alpes	P. nigra natural population	3-4

[a]the strain has been sequenced.
[b]the cultivars have been developed since 1993.

2.3. Genomic data

The genomic data (158 genotypes) used in this study was a subset from 706 wild *P. nigra* genotypes that have been genotyped using 7 896 high-quality Single Nucleotide Polymorphism (SNPs) (Faivre-Rampant *et. al.,* 2016).

The procedure in Faivre-Rampant *et. al.,* (2016) covered the creation of template for *P. nigra* reference genome mapping, detection of SNPs, and selection of SNPs. The selection of SNPs was done using the approach "candidate-genomic regions" which means SNPs were selected in the regions where quantitative trait loci (QTL) for rust resistance and bud phenology in *P. nigra,* as well as QTL for water-use efficiency and

wood properties in other *Populus* species were found. In addition to this set of SNPs, other SNPs that were spreading across the poplar genome were selected to retain *P. nigra* neutral genetic diversity. The final selection was resulting in 10 331 SNPs which were used to genotype 888 wild *P. nigra* individuals collected from 12 populations in Western Europe, the same populations used in this study.

From 10 331 SNPs, several SNPs and genotypes were excluded due to missing data, SNPs segregation problem, SNPs deviation from Hardy-Weinberg equilibrium, and monomorphic SNPs. Therefore, 7 896 SNPs were retained in the genomic data of 706 *P. nigra*. The genomic data for 158 *P. nigra* were a subset from the genomic data of 706 *P.* nigra. There are four genotypes from this study whose SNPs were removed due to clearing reasons in Faivre-Rampant *et. al.,* (2016), thus we excluded them from further analysis.

2.4. The assessment of quantitative resistance in a laboratory

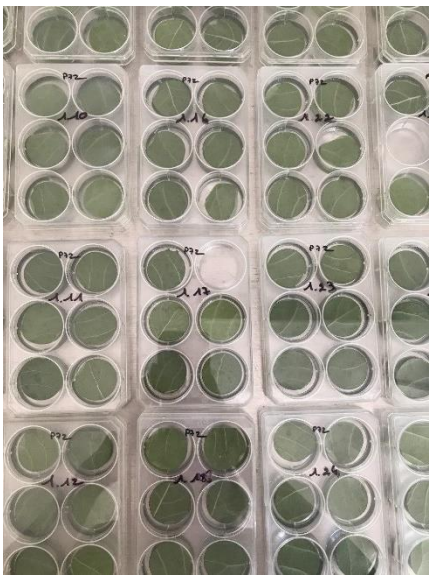


Figure 2 P. nigra leaves were cut into disks of 3-cm in diameter and were put in polycarbonate well cell culture plates

Quantitative resistance was divided into three traits: latent period (LP), uredinia number (UN), and uredinia size (US). The three traits were observed through a laboratory experiment. The inoculation and traits observation were carried out using leaf disk bioassay procedure (Dowkiw*, et al*., 2003). The experiment was carried out in a 5-block experimental design where each genotype was replicated five times and within each block, the genotypes were inoculated by three *Mlp* strains.

P. nigra leaves were taken from the fifth to the eight unrolled leaf below the apex because these leaves exhibit maximum sensitivity toward *Mlp*. Inoculations were then conducted on the sampled leaves that were cut into leaf disks with diameter of 3 cm (Figure 2). Inoculations were done in three separate time, each for one strain (i.e. 09AX27, 93JE3 and P72). The leaf disks were floated upside down on distilled water in polycarbonate well cell culture plates and were sprayed with suspension of *Mlp* urediniospore using a hand atomizer. The inoculum suspension concentration was 30 mg/liter with the following composition: 3 mg of urediniospores in 100 ml distilled water with 10 mg agar. The disks were kept for 14 days in a growth chamber under controlled conditions: temperature was set to 17℃ - 18℃ with 16-h photoperiod.

The number of urediniospore deposited in the inoculation of 09AX27, 93JE3 and P72 was calculated by placing 18 agar boxes among the genotypes. The calculation was done using microscope. The germination rate of the urediniospore was also calculated. Two agar boxes were placed among the genotypes during the inoculation. They were then kept in the growth chamber together with the other genotypes for 48 hours. The calculation of the germinated urediniospore was done in 24 and 48 hours after inoculation using microscope.

The components of quantitative resistance (Figure 3a) were measured within 14 days after the inoculation. Latent period (LP) was indicated as the time needed by the leaves until the first sporulating uredinium appears. It was measured on a half-day basis every day since D-7 until D-14. Any symptoms found before D-7 were evaluated in terms of contamination. Uredinia number (UN) was indicated as the total number of sporulating uredinia on the leaves and under the water. It was evaluated in D-13 after the inoculation through counting. Uredinia size (US) was indicated as the size of the sporulating uredinia on the leaves. It was done in D-14 after the inoculation through a scoring method with a scale from 1 (small) to 5 (large). The scoring system was specific to each strain, as each strain produces different variation in traits. The scoring system was determined by surveying the traits in all genotypes and sampling the genotypes that show small to large size, this was done for each strain separately (Figure 3b).

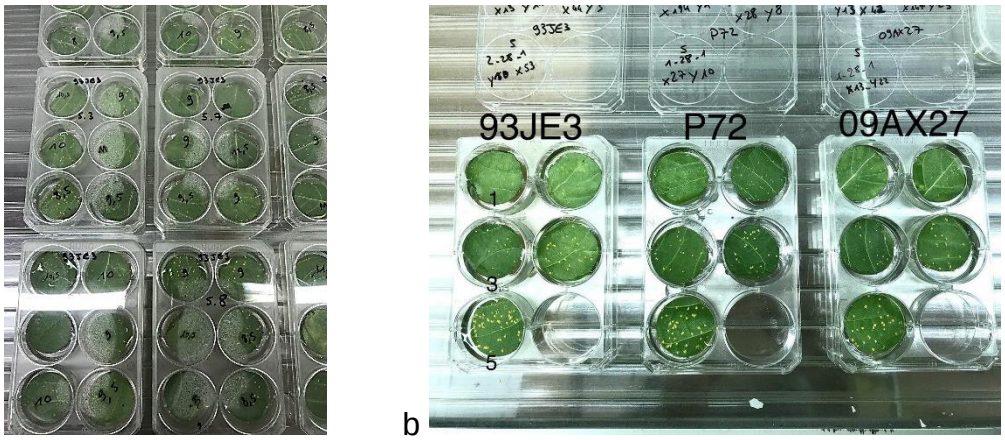


Figure 3 a) Observation of LP (numbers on top the disks), UN (counting the sporulating uredinia on the leaves and under the water), US (scoring uredinia size); b) Scoring system specific to each strain to score US

2.5. Data analysis

2.5.1. Distribution of trait variation, heritability and correlation

Data were analyzed using R programming language for statistical analysis version 4.0.1 (The R Development Core Team, 2020). Prior to population structure analysis, genome-wide association study (GWAS), and interaction modeling, data were analyzed for block corrections, heritability, trait variation, correlations (between traits and between strains), and summary statistics.

The measurement data were corrected for block effects using a linear mixed-effects model: $Y_{ijk} = \mu + B_i + G_j + \varepsilon_{ijk}$, where Y_{ijk} is the trait of interest of *j*-th genotype (*j* = 1, …, 154) in block *i*-th (*i* = 1 , … , 5) for strain *k* (*k* = 1, … , 3), μ is the grand mean, B_i is the block effects (fixed), G_j is genotype effects (random), and ε_{ijk} is residual effects. The model was run using remlf90() function from breedR package in R (Muñoz *et al.,* 2016). The significance of the random effects was evaluated using paired X^2 of the log-likelihood of mixed-effects model and null model. The significance threshold is α = 0.05. The residuals of the model was also evaluated for normality and homogenous variance. Broad-sense heritability for genotype additive effects was then calculated as $H^2 =$

$\frac{\sigma^2{}_G}{\sigma^2{}_G + \sigma^2{}_\varepsilon}$, where $\sigma^2{}_G$ is the additive effects from genotypes and $\sigma^2{}_\varepsilon$ is the environmental effects.

Trait variation (using adjusted data) was evaluated visually through histograms to see the pattern between traits and among strains. Traits were also evaluated in terms of their correlation using Pearson-Correlation method and its R^2 (from -1 to 1). Strains were evaluated in terms of their correlation using the same method. Summary statistics of each trait were also calculated in terms of mean, standard deviation, first quantile, and third quantile.

2.5.2. Analysis of P. nigra population structure through the estimation of admixture coefficients

Before proceeding the analysis to genome-wide association study (GWAS), the genomic data of 154 *P. nigra* collected from natural populations were evaluated for population structure. This step was done to correct for false-positives in GWAS due to genetic relatedness between populations. The evaluation of population structure was done in the sense of admixture coefficients.

Admixture coefficients provide the information about the presence of genes in an individual that come from a distantly-related population. This is based on the premise that the population today is the result from the mixtures of distant populations in the past through evolution processes, thus creating an admixture.

A genomic matrix Z of 154 *P. nigra* genotyped by 7 896 SNPs were used to perform the analysis of admixture coefficients in R. Each observation in matrix Z corresponds to the number of derived alleles at locus l for genotype j. As the genotypes are diploid organisms, the number of derived alleles at locus l is coded as 0, 1, or 2. These codes indicate heterozygous alleles (coded as 1) and homozygous alleles (coded as 0 or 2).

The admixture coefficients were estimated using snmf() program in LEA package which is based on sparse non-negative matrix factorization (described in Frichot *et. al.,* 2014) and least-squares optimization. The estimation of admixture coefficients generally supposes that the genome of j individuals is originated from K ancestral populations that are unknown *a priori*. Given K ancestral populations, the probability of individual j carrying n derived alleles at locus l can be described as (Frichot *et. al.,* 2014):

$$p_{il}(n) = \sum_{k=1}^{K} q_{jk} z_{kl}, \qquad n = 0, 1, 2; j = 1, \dots, 154$$

Where q_{jk} is the proportion of individual j's genome that originates from the ancestral population k, and z_{kl} is the frequency of homozygous ($n = 0$ or 2) or heterozygous alleles ($n = 1$) at locus l in population k.

The estimation of admixture coefficients was then solved in two steps. The algorithm begins with the sparse non-negative matrix factorization (sNMF) and continues with least-square optimization of Q matrix that is suitable to the Z matrix. The Q matrix contains the proportion of individuals' genome to K ancestral populations.

The K ancestral populations were determined by evaluating the prediction error of admixture estimates using a cross-validation technique that corresponds to cross-entropy. Ancestral population K that has the lowest cross-entropy was chosen. In this study, the genomic matrix was initially fitted to K = 1 – 15. The cross-entropy from this range was then evaluated. The K that corresponds to the lowest cross-entropy was then picked. The Q matrix that corresponds to this K was then used in GWAS to correct for false-positives.

2.5.3. Genome-wide association study (GWAS)

A genome-wide association study (GWAS) was intended to model the association between genes at segregating sites and phenotypes (herein are traits). In this study, GWAS was conducted in TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) version 5.2.64 (Bradbury, *et. al.,* 2007). TASSEL was running to test the association between the three traits (LP, UN, US) with a matrix of 154 genotypes x 7896 SNPs for each strain separately. This will result on three different models for each strain. The measurement data of the traits were adjusted for block effects and were averaged for five blocks, resulting in a mean measurement data for one genotype, this is true for all strains.

GWAS in TASSEL is using the approach of structured association. It means that the association model is integrating the variance arising from population structure which indicates the degree of membership in underlying populations. The population structure was derived from the Q matrix of admixture coefficients. By including the Q matrix, TASSEL is able to control the possible false positives that may arise from population structure and avoid any spurious conclusion regarding the association between SNPs and trait.

Because complex traits, as resistance, are controlled by multiple quantitative trait loci (QTLs), the primary goal of GWAS was to find the causal markers (i.e. SNPs) for each QTL. GWAS was fitted using a generalized linear model (GLM). GLM performs the ordinary least squares solution that finds the suitable estimates to the association between SNPs and traits (Brabdury, *et. al.,* 2007). In this study, trait variation was dissected to SNPs as the main effects and population structure Q matrix as the covariate. The following is GLM structure used in this study.

$$Y = Z\beta_1 + X\beta_2 + \varepsilon$$

Where Y is the vector of observation of the trait of interest; Z is the known matrix for SNPs effects with dimension 154 genotypes by 7 896 SNPs, modeled as the main effects; X is the known matrix for population structure Q matrix with dimension 154 genotypes by 6 Q, modeled as a covariate; and ε is a vector of residuals. The parameters β_1 and β_2 are unknown vectors containing the estimated effects of SNPs and population structure.

The results are presented in Manhattan plot where the X-axis represents *P. nigra* chromosomes (19 chromosomes) and Y-axis represents the P-value in -log10 scale. Each P-value in the Y-axis is associated with the SNPs being modelled. The significance was evaluated in terms of P-values. For further analysis, we will evaluate the markers that have P-values lower than 0.001 and marker R^2 between 8 – 13%.

Further analysis involved the identification of the causal genes that are responsible for LP, UN, and US. The evaluation was carried out by searching the genetic codon

associated with the significant SNPs on an online database Phytozome version 12.1 (Goodstein *et. al.,* 2012) using the genome of *Populus trichocarpa* version 3.0 as the template species.

2.5.4. Estimation of genotype by strain (GxS) interaction effects

The analysis was continued to the interaction between 154 *P. nigra* genotypes and three *Mlp* strains for LP, UN, and US. This analysis was aimed to see the degree of interaction between *P. nigra* genotypes and *Mlp* strains to give an idea about the range of *Mlp* strains with which the *P. nigra* genotypes are actively interacting. The interaction was modeled using a linear mixed-effects model and was plotted as reaction norm with strain 09AX27, 93JE3, and P72 in the X-axis and degree of interaction in the Y-axis.

The effects of genotype by strain interaction (GxS) were modeled using linear mixed-effects model. The data was prepared by combining the un-adjusted measurement data of three strains, thus resulting to dataset that contains trait measurements for the three strains. The linear mixed-effects model is structured as follows: $Y_{ijk} = \mu + S_k + G_j + B_i + (GS)_{jk} + \varepsilon_{ijk}$, where Y_{ijkl} is the trait of interest for *j*-th genotype (*j* = 1, …, 154) in *i*-th block (*i* = 1, …, 5) for strain *k* (*k* = 1, …, 3), μ is the general mean, S_k is strain effects (fixed-effects), G_j is genetic effects (random-effects), B_i is block effects (random-effects), $(GS)_{jk}$ is GxS interaction effects (random-effects), and ε_{ijkl} is residual effects. The model was run using remlf90() function in breedR package in R (Muñoz et al., 2016). The significance of the variance components was evaluated using paired X^2 of the log-likelihood of mixed-effects model and null model. A threshold α = 0.05 was used to test the significance.

Interaction effects were then predicted through the model ($(GS)_{jk}$). Ecovalence (W) was then computed to see the relative contribution of each strain across genotypes ($n_{jk}\ {(GS)_{jk}}^2$) to GxS interaction ($SCE_{(GS)}$) using the following equation:

$$W_{rk} = \frac{\sum_{k=1}^{s} n_{jk}\ {(GS)_{jk}}^2}{SCE_{(GS)}}$$

The contribution of each genotype across strain (W_{rj}) was also computed to produce the rank of genotypes with high interaction with the strains. The threshold was arbitrarily chosen and it was W_{rj}= 0.01. The genotypes with W_{rj}> 0.01 were considered as highly interactive with the strains.

$$W_{rj} = \frac{\sum_{j=1}^{g} n_{jk}\ {(GS)_{jk}}^2}{SCE_{(GS)}}$$

Genetic correlation of traits among strains was evaluated by fitting a linear mixed-effects model for multivariate analysis:

$$Corr(S) = Y\beta_1 + G\beta_2 + I\beta_3 + \varepsilon;\ G = \frac{(M-2P)(M-2P)'}{2\sum p_i q_i}\ \text{(VanRaden's G matrix)}$$

Where Y is 3 x 154 matrix of the trait recorded in the three *Mlp* strains, K is 154 x 154 marker-based relationship matrix between genotypes, calculated using VanRaden's G

matrix in AGHmatrix R package (Amadeu *et. al.,* 2016), I is 154 x 154 incidence matrix with 1 in the diagonal and 0 elsewhere; and ε is residual effects.

3. Results

3.1. Data distribution

It was shown that block had significant effects on the traits. Therefore, the traits were adjusted for block effects. The traits used for further analysis are the adjusted traits, except in the case where it is indicated otherwise. It was also shown that genotype effects were significant for all traits.

Indicated by relatively small standard deviation, LP and US for all strains did not show high variation, whereas UN showed higher variation than the other two (Table 2). Genotypes showed longer LP and lower UN when they were infected by strain 09AX27 than in strain 93JE3 and P72. Between strain 93JE3 and P72, genotypes in 93JE3 had symptoms in earlier time (shorter LP) and higher UN than in strain P72. However, there was no visible differences between US in the three strains.

Table 2 Summary statistics of latent period (LP), uredinia number (UN), and uredinia size (US) in 154 P. nigra genotypes collected from natural populations across strain 09AX27, 93JE3, P72.

			LP (days)		UN		US	
	Inoc. Pres.[a]	Germ.[b]	Mean	Q1 – Q3[c]	Mean	Q1 – Q3	Mean	Q1 – Q3
Strain 09AX27	564 ± 94	42%	10.6 ± 0.6	10.1 - 10.9	7.0 ± 2.7	5.3 - 8.9	3.5 ± 0.5	3.2 - 3.8
Strain 93JE3	364 ± 66	96%	9.5 ± 0.4	9.3 - 9.7	11.4 ± 3.8	8.9 - 13.7	3.8 ± 0.4	3.5 - 4.1
Strain P72	472 ± 61	91%	9.3 ± 0.5	9.0 - 9.6	17.8 ± 6.1	14.2 - 21.1	3.2 ± 0.5	2.9 - 3.5

[a] Inoculum pressure: average number of urediniospore deposited in each agar box after inoculation process, ± standard deviation
[b] Germination rate: average percentage of spores that were successfully germinated 24 and 48 hours after inoculation process
[c] First (Q1) and third (Q3) quantiles of mean values

Table 2 also shows a contrasting inoculum pressure and germination rate among strains, which may influence the high variation in UN. The average number of urediniospore deposited in the agar boxes for 09AX27 was the highest, but only 42% of the urediniospore was germinated. The genotypes started showing symptoms on day 10 with moderately low uredinia size. However, because there was only 42% of the urediniospore germinated, it would be difficult to conclude the aggressivity of strain 09AX27. Genotypes were inoculated with 93JE3 with the lowest inoculum pressure. With 96% of it was germinated, the genotypes started showing symptoms on day 9 with the highest uredinia size. Therefore, 93JE3 could be considered as an aggressive strain. Strain P72 is less aggressive than 93JE3. Despite that P72 inoculum pressure is second highest and 91% of it was germinated, the genotypes showed the lowest uredinia size.

Visual representation of the LP, UN, and US in strain 09AX27, 93JE3, and P72 (Figure 5) confirms the summary statistics. Genotypes in strain 09AX27 started showing symptoms (first sporulating uredinia) in a later day (LP) than genotypes in strain 93JE3 and P72. There were also lower counts of sporulating uredinia (UN) in the genotypes in 09AX27 than in the other strains. For US, the genotypes were behaving similarly across strains.

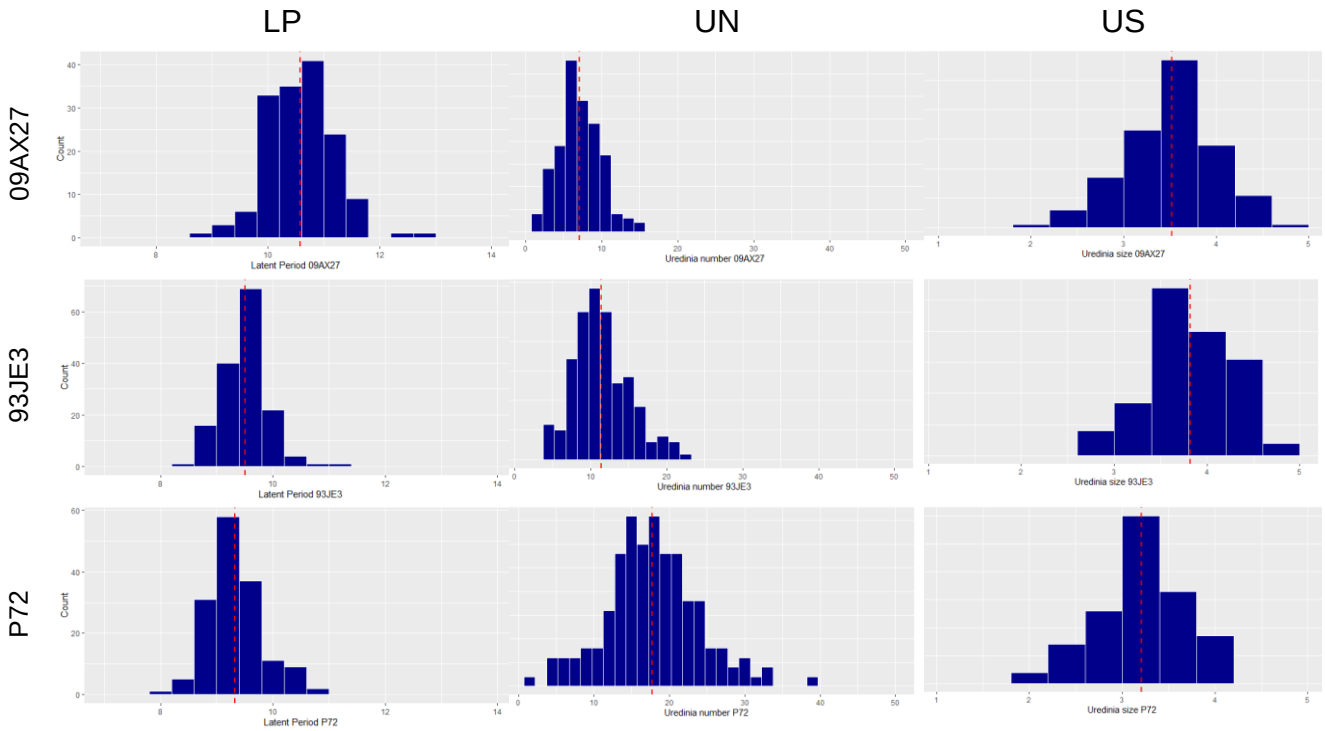


Figure 4 Distribution of latent period (LP) with limit on the X-axis 7 – 14 (days); uredinia number (UN) with limit on the X-axis 1 – 40 (counts); and uredinia size (US) with limit on X-axis 1 - 5 (scores) for strain 09AX27, 93JE3, and P72. Red line indicates the mean for the related traits.

Based on the mixed-effects model, genetic explained almost 50% of trait variance for three *Mlp* strains. Broad-sense heritability was then measured for each trait and each strain (Table 3). Heritability estimates were highest in US, followed by UN and LP. Although heritability for LP for strain P72 was the highest compared to the other strains. Overall, US had the highest heritability among all the traits, and this was true for strain 93JE3 and P72. US and UN had similar heritability in strain 09AX27.

Table 3 Broad-sense heritabilities for latent period (LP), uredinia number (UN) , and uredinia size (US) in 154 P. nigra genotypes collected from natural populations and inoculated with three Mlp strains.

	Heritability H^2		
	LP	UN	US
Strain 09AX27	0.28	0.43	0.43
Strain 93JE3	0.47	0.43	0.57
Strain P72	0.54	0.48	0.60

Correlations between traits and between strains were also evaluated. The results are presented in Annex 2 Across traits, correlations between LP and UN, as well as between LP and US were negative. The results were true for all the *Mlp* strains. Correlations between LP and US were the highest, with R^2 of 0.54, 0.58, and 0.78 for strain between 09AX27 and 93JE3, 93JE3, and P72, respectively. Since all the correlations between LP and US were negative, long LP may cause low US, and vice versa. Correlations between US and UN were positive. This was also true for all the *Mlp* strains. The correlations between US and UNwere indicated by R^2 of 0.38, 0.40, and 0.52 for strain 09AX27, 93JE3, and P72, respectively. Since all the correlations between US and UN were

positive, high UN might cause high US, and vice versa. Across strains, correlations between all traits were the highest in strain P72, followed by strain 93JE3 and 09AX27.

Correlations between strains were evaluated by comparing the same trait over all the strains. For all the traits compared, strains do not seem to correlate to each other as high as traits are correlated. Visible correlation that explains 15 – 20% of trait variation can be observed when comparing LP in 09AX27 and P72 (16%, negative correlation), comparing UN in 93JE3 and in P72 (20%, negative correlation), and comparing US in 93JE3 and P72 (15%, positive correlation).

3.2. Admixture coefficients

The range for *K* ancestral population was between 1 – 15. This range was chosen because the *P. nigra* genotypes were collected from 12 natural populations. The hypothesis was that the genotypes should be grouped within their population of origin. Based on the evaluation of cross-entropy, the optimal number of ancestral clusters was *K* = 6. This is the *K* that corresponds to the lowest cross-entropy. Moreover, in the study of a larger population, *K* = 6 was considered suitable to the genomic data (Faivre-Rampant *et. al.,* 2016). Therefore, the *Q* matrix of admixture coefficients contains the coefficient of each Q_K (*K* = 1, … , 6) associated to each of 154 *P. nigra* genotypes. The Q matrix was then used as the covariate in GWAS (Figure 5).

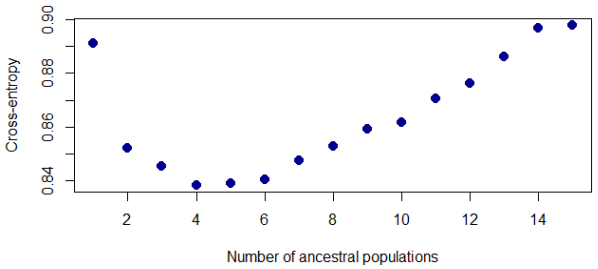


Figure 5 Cross-entropy to evaluate the number of ancestral populations

Figure 6 showed the admixture coefficients of each genotype that is grouped within their population of origin. Genotypes from the same population tended to have similar mixture and to have one major ancestral population specific to the population.

Genotypes from Adour (Southern France) probably share the same major ancestral population with those in Nohedes (Southern France), as well as with Loire (Central France), even though its share with Loire is not as big as it is with Nohedes. Genotypes from Dranse (Eastern France) have their major ancestral population that is also found in the genotypes from Rhin (Eastern France), although Rhin is more of a mixture of genes coming from Dranse and Kuhkopf. Genotypes from Kuhkopf (West Germany) showed a distinct ancestral population compared to the genotypes from France. This is the same case as the genotypes from Ticino (Northern Italy) which emerged as a distinct cluster. Ramieres (South Eastern France) had a major ancestral population that is quite different from the other populations in France. Val d'Allier and Loire show similar pattern of admixture.

However, genotypes from the Netherlands showed mixtures of all the ancestral populations without showing the domination of one ancestral population. A genotype from Garonne showed a special case as it completely showed a similarity with the genotypes from Adour (France).

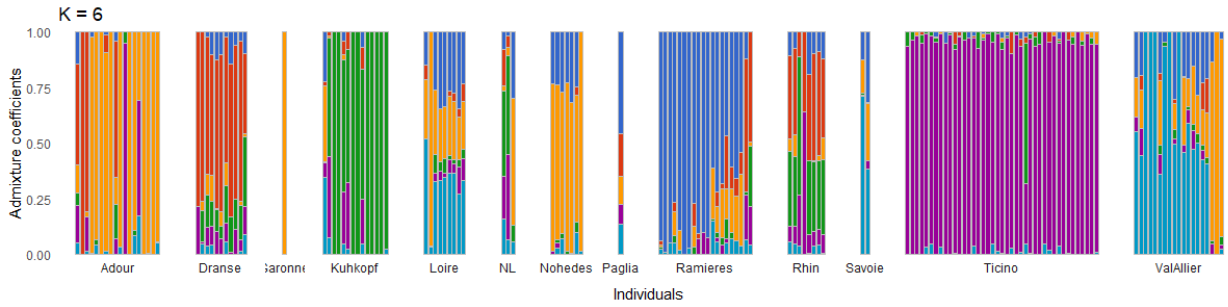


Figure 6 Population structure analysis using admixture coefficients with K = 6. The admixtures are resulted from 154 genotypes and 7 896 SNPs. Each color represents different ancestral populations . Each genotype is represented with a thin vertical line that is divided into color segments that are proportional to its degree of membership to an ancestral population.

3.3. Genome-wide association study (GWAS)

As the aim of GWAS was to find the association between SNPs and traits, the evaluation was based on SNPs' R^2 generated from the marginal models that were fitted to each of the SNPs. The R^2 marks the degree of trait variance that is explained by SNPs. These SNPs were then regarded as the candidate genes. The candidate genes will be used to identify the causal genes that are responsible to LP, UN, and US. To proceed to this stage, we considered the markers that have P-values lower than 0.001 (-log10(0.001) = -3) and R^2 between 8 – 13%. Figure 7 shows the Manhattan plots of the traits (horizontal order) for three *Mlp* strains (vertical order)

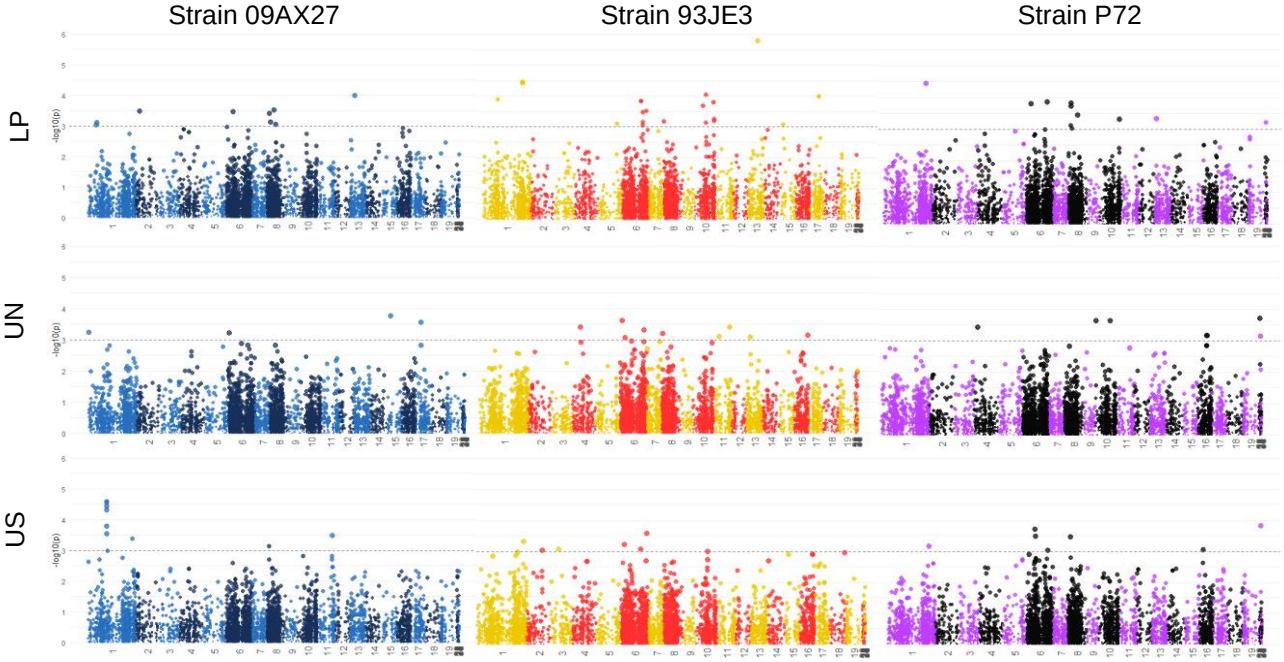


Figure 7 Manhattan plot generated from the genome-wide association study (GWAS) in TASSEL for latent period (LP), uredinia number (UN), and uredinia size (US) of 154 P. nigra genotypes and 7 800 SNPs. The GWAS was done for three Mlp strains. X-axis represents the chromosomes from 1 – 19, Y-axis represents the P-value in -log10 associated to each SNP, and points represent the SNPs in their location across the genome. The dashed grey line represents P-value threshold, which is $10^{-3}$and $-log(10^{-3}) = -3$. SNPs with P-value < than $-log(10^{-3})$ are considered as the candidate SNPs for further analysis.

Table 4 summarizes the number of significant SNPs that are associating with each trait in all the *Mlp* strains, their R^2 and the chromosome where they are residing. LP had the most number of SNPs that were associating with it. US and UN showed associations with relatively similar number of SNPs. In 09AX27, SNPs that were associating with US had the largest R^2 range, meaning that SNPs could explain 9 – 13% US variation. In 93JE3, SNPs that were associating with LP had the largest R^2 range, meaning that SNPs could explain 8 – 15% of LP variation. In P72, SNPs that were associating with LP had the largest R^2 range, meaning that SNPs could explain 8 – 13% LP variation. Between strains, R^2 between SNPs and traits was more variable in 09AX27 than in the other two.

Table 4 Count number of SNPs found significant in genome-wide association study (GWAS) between 154 P. nigra genotypes genotyped by 7 896 SNPs with latent period (LP), uredinia number (UN), and uredinia size (US) in strain 09AX27, 93JE3, and P72. P-value threshold = 0.001 (-log10(0.001) = -3).

Strain	Trait	Total significant SNPs[a]	R^2	Chromosome
09AX27	LP	9	7 – 11%	Ch. 1 (2)[b], Ch. 2, Ch. 6, Ch. 8 (4), Ch. 13
	UN	3	7 – 9%	Ch. 1, Ch. 6, Ch. 17
	US	8	9 – 13%	Ch. 1 (7), Ch. 8
93JE3	LP	18	8 – 15%	Ch. 1 (3), Ch. 6 (6), Ch. 8, Ch. 10 (6), Ch. 13
	UN	8	7 – 11%	Ch. 4, Ch. 6 (3), Ch. 8, Ch. 11, Ch. 13
	US	6	8 – 9%	Ch. 1, Ch. 2, Ch. 3, Ch. 6 (3)
P72	LP	10	8 – 13%	Ch. 1, Ch. 6 (2), Ch. 8 (5), Ch. 10, Ch. 13
	UN	5	8 – 11%	Ch. 4, Ch. 10, Ch. 16 (2), Ch. 22
	US	6	8 – 10%	Ch. 1, Ch. 6 (3), Ch. 8, Ch. 16

a Count number of SNPs with P-values lower than 0.001; b Count number of SNPs found clustering in the same chromosome

Causal SNPs were found in the same chromosomes across strains, especially for LP and US. There were seven causal SNPs found in chromosome 1 for LP and US in all the three strains. These SNPs were mostly found in clusters and explaining 9 – 13% of LP and US variation (not shown in the table). Moreover, seven causal SNPs were also found in chromosome 6 for LP and US in all the three strains. These SNPs were also found in clusters and explaining 8 – 11% of LP and US variation. There were several SNPs associated with UN that were unique to each strain. A significant SNP found in chromosome 17 was associated with UN in strain 09AX27, a SNP found in chromosome 11 was associated with UN in strain 93JE3, and a SNP found in chromosome 22 was associated with UN in strain P72. There was also a SNP in chromosome 3 that was uniquely associated with US in strain 93JE3. In P72, there were also several SNPs found in chromosome 16 that were associated with US and UN (Table 4).

Table 5 summarizes the list of the annotated genes for each causal SNP. There were eight genes related to serine-theonine protein kinase that were repeated in chromosome 1 for US in strain 09AX27 and LP in 93JE3, in chromosome 6 for LP in strain P72, and in chromosome 16 for UN in strain P72. There was also one gene related to leucine rich repeat (LRR) protein family in chromosome 11 for UN in 93JE3. Genes related to CGTHBA protein were repeated twice for LP in 93JE3 in chromosome 6. Genes related to YL1 nuclear family protein were repeated twice in chromosome 1 for LP and US, both in P72. As for the SNPs specific to each strain, UN in 09AX27 was related to a gene for F-box domain protein (R^2 = 0.08, chromosome 17) and UN in P72 was related to a gene that is in MYB transcription factor domain (R^2 = 0.11, chromosome 22).

Table 5 A list of the annotated genes that are related to the SNPs found significant in genome-wide association study (GWAS) for latent period (LP), uredinia number (UN) and uredinia size (US) in strain 09AX27, 93JE3, and P72.

Strain	Trait	SNP	Ch.	P-value	Marker R^2	Annotated gene	Function
09AX27	LP	SNP_IGA_1_6896006	1	9.34E-04	0.08	Potri.001G108400	Unknown function
	LP	SNP_IGA_1_6459551	1	7.65E-04	0.09	Potri.001G113100	Leucocyanidin oxygenase
	LP	SNP_IGA_6_7123706	6	3.48E-04	0.10	Potri.006G095300	O-Fucosyltransferase family protein
	UN	SNP_IGA_17_7183409	17	2.80E-04	0.08	Potri.017G058900	F-box domain (F-box)
	US	SNP_IGA_1_18941108	1	4.90E-05	0.12	Potri.001G200000	Serine-threonine protein kinase
	US	SNP_IGA_1_18939516	1	3.49E-05	0.13	Potri.001G200000	Serine-threonine protein kinase
	US	SNP_IGA_1_18936727	1	2.57E-05	0.13	Potri.001G200000	Serine-threonine protein kinase
	US	SNP_IGA_1_18937776	1	2.57E-05	0.13	Potri.001G200000	Serine-threonine protein kinase
93JE3	LP	SNP_IGA_1_40721150	1	3.65E-05	0.12	Potri.001G398500	Serine-threonine protein kinase
	LP	SNP_IGA_1_40722917	1	3.57E-05	0.12	Potri.001G398500	Serine-threonine protein kinase
	LP	SNP_IGA_6_18603837	6	1.51E-04	0.11	Potri.006G180800	CGTHBA protein -14 gene protein
	LP	SNP_IGA_6_18604833	6	1.51E-04	0.11	Potri.006G180800	CGTHBA protein -14 gene protein
	UN	SNP_IGA_11_16933109	11	3.87E-04	0.10	Potri.011G141100	Leucine rich repeat family protein (LRR)
	US	SNP_IGA_1_43685091	1	4.81E-04	0.09	Potri.001G443200	Cytoskeleton-associated protein 5
	US	SNP_IGA_3_6330558	3	8.44E-04	0.09	Potri.003G052400	Homeobox-leucine zipper protein glabra 2
	US	SNP_IGA_6_24951663	6	2.68E-04	0.10	Potri.006G250200	G-protein beta subunit
P72	LP	SNP_IGA_1_42176547	1	3.56E-05	0.13	Potri.001G412800	YL1 nuclear family protein;
	LP	SNP_IGA_6_4817356	6	1.56E-04	0.11	Potri.006G066900	Clathrin assembly protein
	LP	SNP_IGA_6_20564788	6	1.40E-04	0.11	Potri.006G199400	Serine-threonine protein kinase
	UN	SNP_IGA_16_9654896	16	6.84E-04	0.09	Potri.016G095700	Serine-threonine protein kinase
	UN	SNP_IGA_16_9549649	16	6.84E-04	0.09	Potri.016G095900	3-Deoxy-D-manno-octulosonic-acid transferase (kdotransferase)
	UN	SNP_IGA_19_504345	22	1.92E-04	0.11	Potri.T011400	MYB transcription factor
	US	SNP_IGA_1_42176547	1	6.85E-04	0.09	Potri.001G412800	YL1 nuclear family protein;
	US	SNP_IGA_6_8628573	6	1.94E-04	0.11	Potri.006G112200	Auxin efflux carrier family protein
	US	SNP_IGA_16_3627316	16	8.88E-04	0.09	Potri.016G056400	Unknown function

3.4. Estimation of genotype by strain (GxS) interaction effects

Annex 3 shows the dissection of variance components to trait variation. In order to depict genotype by strain (GxS) interaction effects, a linear mixed-effects model was fitted where strains were modeled in fixed effects, and genotypes, block, and GxS interaction were modeled in random effects. The significance of the variance components was evaluated based on α = 0.05. LP had the highest environmental variance (71.60%), UN had the highest GxS variance (20.66%), only a mere different than US, and US had the highest genotype variance (40.09%). However, even though genotype and GxS had significant effects on UN, block also had a significant effect on the trait. Therefore, UN will not be considered for further analysis on GxS. US was shown to have high variance both for genotype and GxS (Table 6) and had the highest genotypic heritability (Table 2).

Table 6 Relative effects of genotype variance (Var_g), genotype by strain variance (Var_{gxs}) and environmental variance (Var_e) to latent period (LP), uredinia number (UN), and uredinia size (US) variance.

Traits' variance components								
Var_g			Var_{gxs}			Var_e		
LP	UN	US	LP	UN	US	LP	UN	US
27.18%	33.49%	40.09%	19.70%	20.66%	20.38%	71.60%	64.75%	56.65%

Genotype by strain (GxS) interaction effects for LP and US were plotted in a reaction norm (Figure 8). Figure 8 clearly shows some genotypes that are highly interactive with the strains. They are shown as fluctuating lines across the strains. The weakly interactive strains are shown in the cluster of lines near Y = 0.0. These genotypes show similar GxS across strains, thus they are considered as being weakly interactive. In order to see the most interactive genotypes and strains, ecovalence of each was calculated.

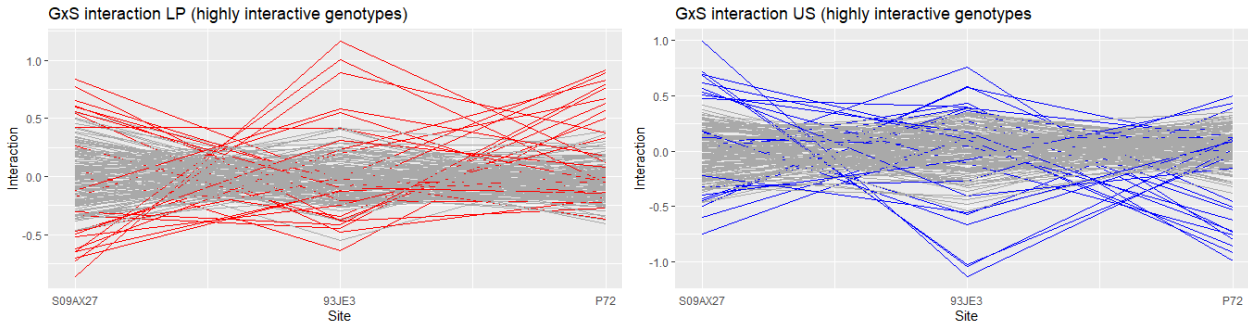


Figure 8 *Genotype by strain (GxS) interaction showing the highly interactive genotypes for latent period (LP) (left) and uredinia size (US) (right) for 154 P. nigra genotypes inoculated by three Mlp strains. Strain 09AX27, 93JE3, and P72 are plotted on the X-axis, and GxS effects are plotted on the Y-axis. Each genotype is represented by a line that spans across three strains. Red lines show the highly interactive genotypes for LP and blue lines show the highly interactive genotypes for US.*

There were 25 highly interactive genotypes with three *Mlp* strains for LP and 25 highly interactive genotypes for US (Figure 9, Table 7). Among these genotypes, 15 genotypes were highly interactive with *Mlp* strains for both LP and US simultaneously. The ecovalence of these genotypes were similar between LP and US. It is ranging between 1

– 6% with the highest is in genotype NOH-15, true for both traits. It means 15 genotypes contribute 1 – 6% to GxS interaction across strains for LP and US simultaneously.

Table 7 A list of highly interactive genotypes and their relative contribution to genotype by strain interaction (GxS) for both latent period (LP) and uredinia number (UN) simultaneously.

	1-A02	1-A08	1-J07	6-J32	KUH-54	N-11	N-22	N-35	NL-1138	NOH-10	NOH-15	NOH-25	SN-33	SPM-04	SSN-05
W_{rj} LP	2.23	1.53	2.68	2.13	2.47	2.16	3.40	3.20	4.40	4.29	6.35	3.92	2.72	1.06	1.16
W_{rj} US	3.59	2.14	4.06	3.91	1.04	1.76	1.21	2.56	4.50	4.75	6.73	3.32	2.41	1.51	1.07

W_{rj} Ecovalence showing relative contribution of each genotype to GxS

Beside ecovalence for each genotype, ecovalence for each strain was also evaluated. Table 8 shows that 09AX27 is the most interactive strain for LP with relative ecovalence 44%, meaning that 09AX27 contributes 44% to GxS interaction across genotypes. For US, 93JE3 shows the highest contribution to GxS interaction with ecovalence equals to 35%.

Table 8 Relative contribution of 09AX27, 93JE3, and P72 across genotypes to genotype by strain interaction (GxS) for latent period (LP) and uredinia number (UN).

	09AX27	93JE3	P72
W_{rk} LP	44	28	27
W_{rk} US	34	35	30

W_{rk} Ecovalence showing relative contribution of each strain to GxS

Genetic correlation was also evaluated for LP across strains and US across strains. Genetic correlation between LP across strains was not as high as the genetic correlation between US across strains. The correlation across strains for US was ranging between 81 to 91% (irrespective to the direction). Only US between 93JE3 and P72 were positively correlated. On the other hand, LP between 93JE3 and P72 were not really correlated, even though the direction was also positive. US between 09AX27 and 93JE3 were negatively correlated while LP between the same strains were positively correlated. Although the strength was higher in US than in LP (Figure 9 and 10).

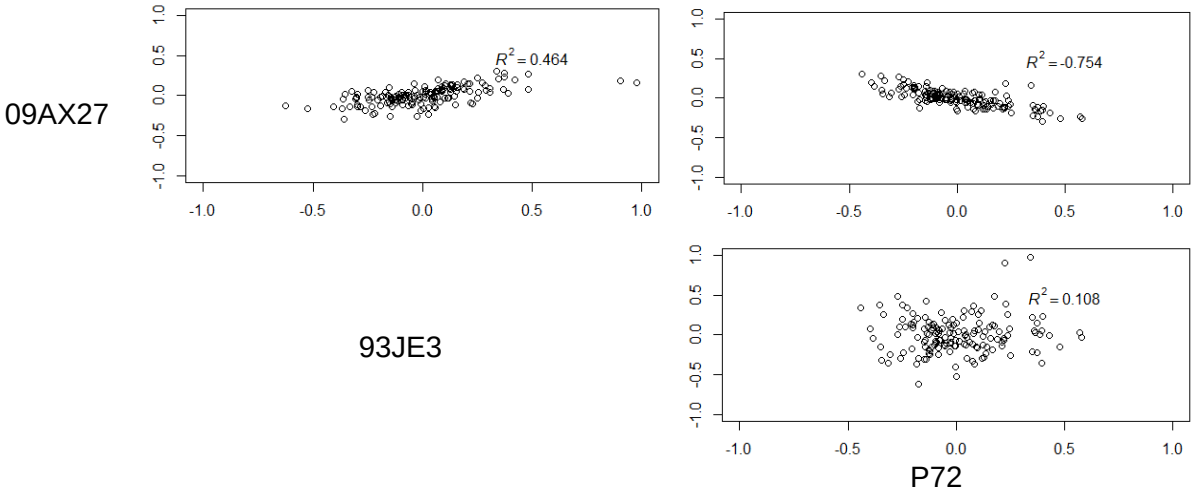


Figure 9 Genetic correlation between latent period (LP) in across strain 09AX27, 93JE3, and P72. R² shows the degree of relationship between LP in two strains.

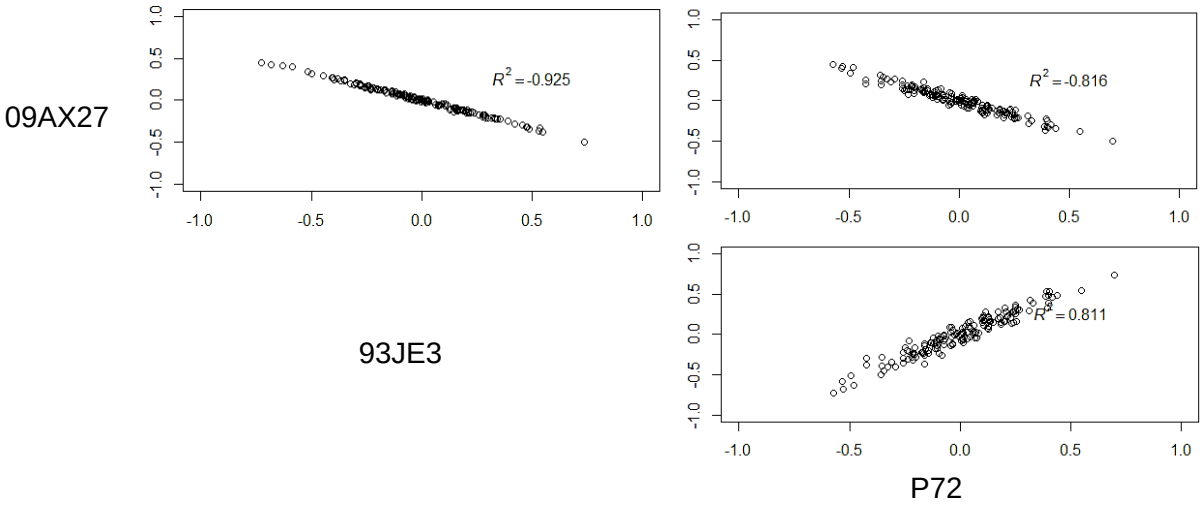


Figure 10 Genetic correlation between latent period (LP) in across strain 09AX27, 93JE3, and P72. R² shows the degree of relationship between LP in two strains.

4. Discussion

This study illustrated the genetic control of quantitative resistance in 154 *P. nigra* genotypes originating from 12 natural populations in Western Europe. The quantitative resistance was studied in terms of latent period (LP), uredinia number (UN) and uredinia size (US) by inoculating *P. nigra* genotypes with three *Mlp* strains, namely 09AX27, 93JE3, and P72. The experiment was carried out in an controlled laboratory experiment in a 5-block experimental design. The genetic control of the quantitative resistance was studied using genome-wide association study (GWAS) by dissecting trait variance into genetic effects (main effects) and population structure effects (covariate) in a generalized linear model (GLM). Using significance threshold of P-value = 0.001, this study found 25 causal SNPs in common chromosomes across traits and strains, which are chromosome 1 and 6. This study also found the annotated genes that may be responsible for the traits. Furthermore, a reaction norm that depicts strain-specificity of the quantitative resistance was also developed. The study confirmed the findings in Bastien *et. al.,* (2015) that the quantitative resistance in *P. nigra* is highly strain-specific as its qualitative resistance.

4.1. Responses of *P. nigra* genotypes to the different strains of *Melampsora laricii-populina* Kleb.

Uredinia number showed higher variation than latent period and uredinia size. On average, genotypes started showing symptoms on day 10 after inoculation for strain 09AX27 and on day 9 for strain 93JE3 and P72. The genotypes were behaving similarly in terms of uredinia size, although genotypes inoculated with P72 showed the smallest uredinia size. However, the ranking of strain aggressiveness cannot be concluded by looking through the differences in genotype responses. Moreover, the counts of uredinia number cannot provide reliable information as the parameter is influenced by inoculum pressure. Dowkiw *et. al.,* (2003) showed that double inoculum pressure had a significant

impact on uredinia number. Primary inoculum pressure and its germination rate should give more information to conclude the ranking.

Primary inoculum pressure is important as it determines the density of the first sporulating uredinia which leads to the settlement and the colonization of the uredinia on *P. nigra* leaves (Maupetit *et. al.,* 2018; Barrès *et. al.,* 2012; Pariaud et al., 2009; Pei et al., 2003). The germination rate is also important as it gives the idea of the relative amount of *Mlp* inoculum is able to germinate. As there was only 42% of inoculum pressure in 09AX27 was germinating, it is difficult to conclude the aggressivity of strain 09AX27. Strain 93JE3 seems to be more aggressive than P72 because despite its inoculum pressure being the smallest, the genotypes started showing symptoms on day 9 and on day 13, their uredinia size was the highest. For genotypes inoculated with P72, they also started showing symptoms on day 9 but their uredinia size was the smallest.

Across traits, latent period with uredinia number, as well as with uredinia size are negatively correlated. Uredinia number and uredinia size are positively correlated. The pattern is the same for strain 09AX27, 93JE3, and P72. However, even though latent period showed negative correlation with uredinia number, data distribution showed that long latent period could result to either small or large uredinia number, but long latent period could result to small uredinia size. Dowkiw *et. al.,* (2003) described this relationship as 'triangular'. This happens because of the limiting factor of latent period to the development of uredinia size.

Positive correlation between uredinia number and uredinia size had also been reported between *Populus deltoides x P. nigra* 'Robusta' and *Melampsora laricii-populina* Kleb. strain 98AG31 (Maupetit *et. al.,* 2018), on willow rust (Pei *et. al.,* 2002) and on wheat leaf rust (Robert *et. al.,* 2002). Across strains, it seemed that the traits were independent from the strains since there was no visible pattern of trait relationship across strains.

For heritability, uredinia size showed the highest heritability. Uredinia size is usually used to observe the quantitative resistance in *Populus* spp. study. It was shown to be segregating and inherited from *P. trichocarpa* parent in *P. x interamericana* cultivars (Bresson *et. al.,* 2011). Uredinia size is also the component of quantitative resistance that correlates with field resistance, making it a good indicator of *P. nigra* resistance in the field (Jorge *et. al.,* 2005).

4.2. Geographical distributions of *P. nigra* genetic diversity

Admixture coefficients were computed to observe the geographical distribution of *P. nigra* genetic diversity. Six ancestral populations were chosen because it had the lowest cross-entropy that minimizes the prediction error from the model. Another reason is that six ancestral populations were within the suitable range for the genomic data that were used in this study, and this has been evaluated previously by Faivre-Rampant *et. al.,* (2016) as the genomic data used in this study is originating from that study.

The admixture coefficients from this study showed the same pattern as in Faivre-Rampant *et. al.,* (2016). The genotypes from Kuhkopf (Western Germany) and Ticino (Northern Italy) emerged as distinct populations than those in France. There is no mixtures with the

other populations from France in these two populations. On the other hand, the genotypes that come from outside France, namely the Netherlands population, was a mixture of France, Germany, and Italy populations.

In France, Adour and Nohedes shared a major ancestral population. This is understandable as even though Adour is located in South-Western France and Nohedes in South-Eastern, they are connected by the Pyrenees. Additionally, these two populations were not studied in Faivre-Rampant *et. al.,* (2016) and they are distinct from the other populations described hereafter. The genotypes in Loire was a mixture of Val d'Allier, Ramieres and Adour. The reason might be because Loire and Val d'Allier are connected with Loire and Allier river in Central France. Genotypes from Dranse showed a distinct ancestral population than the other genotypes from France. This confirms the findings in Faivre-Rampant *et. al.,* (2015). Genotypes from Rhin showed a mixture from mostly from Dranse and Kuhkopf (Western Germany). This is probably due to its location that is near to Dranse, and is connected by Rhine river to Kuhkopf. Finally, Ramieres and Val Allier had their unique ancestral populations.

As a tree species that is grown along the river, admixture in *P. nigra* population is facilitated by the river that is stretching for kilometers, connecting one *P. nigra* population with the other. Principal component analysis (PCA) in Faivre-Rampant *et. al.,* (2015) for the same set of genomic data showed that there was a high level of admixture and low level of genetic differentiation between populations. This shows that there is important gene flow in *Populus* spp. population in general, and in *P. nigra* populations in France (Ballian and Tröber, 2017; Zheng *et. al.,* 2015; Smulders *et. al.,* 2008). Individuals belonging to the same river basin, as in this study are Allier, Loire, Rhin, and Rhone rivers, are clustering together and are sharing similar ancestral population. However, there is also a clear pattern of genetic differentiation between populations from different river basins, as shown by individuals from Val Allier and Ramieres.

4.3. Genome-wide association study: finding the candidate genes for rust resistance in *P. nigra*

The data used in this study had high SNP number for *P. nigra* which was even higher than those reported in previous study (Chu *at. al.,* 2014). Moreover, this study is the first study that uses genome-wide association study (GWAS) in *P. nigra* genotypes from natural population. The genomic data of 154 *P. nigra* genotypes from 12 natural population showed a clear admixture within their populations. Including the admixture information such as Q matrix generated from admixture coefficients estimation is expected to increase the statistical power of GWAS in detecting causal SNPs and eliminating false positives. In this study, we would like to find the causal SNPs that are associating with epidemiological traits to give an idea about the genetic control of quantitative resistance in *P. nigra*.

This study found 25 causal SNPs in common chromosomes, which are chromosome 1 and 6 for trait latent period and uredinia size. These SNPs were found in clusters and they explained 8 – 11% of latent period and uredinia size variation. This study detected the repeats of serine theonine protein kinase genes. The genes are usually found in

resistance (R) genomic region containing NBS-LRR (nucleotide-binding site leucine-rich repeat) family of R genes. They were found in qualitative (R_1) resistance genes in chromosome 19 in *Populus trichocarpa* (Bresson *et. al.,* 2011), and in barley resistance genes *Rpg1* and *Rpg5* towards barley stem rust (Brueggeman *et. al.,* 2002, 2008).

Serine theonine protein kinase genes are grouped in plant's receptor class and they are interacting with a wide array of processes, including disease resistance. They are mainly responsible for signal reception in plant's cell and for interaction with a diverse group of proteins to mediate various combinations of signal response that is specific to the pathogen or the factor that comes into contact with the plant's cell. This group of protein is responsible to recognize the pathogen that comes into contact with the plant and activate the specific defense mechanism to resist or tolerate pathogen's infection (Afzal *et. al.,* 2007). In this study, serine theonine protein kinase was found in chromosome 1 and 6 in all strains and was related to uredinia size in 09AX27, to uredinia number in P72, and to latent period in 93JE3. There was also leucine rich repeat (LRR) protein family that was related to uredinia number in 93JE3, which is usually regulated for plant defense mechanism.

This study also found genes that are unique to each strain. A causal SNP found in chromosome 17 for uredinia number in 09AX27 was related to F-box protein domain. The same protein was encoded by the candidate genes for resistance in barley previously. The F-box protein is known to be a "part of functional components of the multiprotein E3 ubiquitin ligase complex known as SCF" (Solanki *et. al.,* 2019). The SCF complex itself plays a role in mediating regulation of different proteins that are responsible for plant immunity. It mediates the dual defense responses of special protein NLR in barley resistance genes *Rpg5* and *Rga1* (Solanki *et. al.,* 2019). The same role was also observed in tobacco (Liu *et. al.,* 2002). The presence of this protein in relation to uredinia number in 09AX27 may have similar regulation, that is mediating the interaction between proteins that are responsible for *P. nigra* defense. However, it should be investigated further as the protein is functioning as a component of another complex component.

A causal SNP in scaffold 22 (unassembled genome) for uredinia number in P72 was found to be related with genes in the domain MYB transcription factor. This is a large family of transcription factor genes which are usually responsible in many cellular processes. It is easily inducible by environmental conditions that the plants are exposed to and is able to promote the regulation of various functional genes to modify the traits. It is found to have an important role in plants' defence mechanism, plant development and stress tolerance (Ambawat *et. al.,* 2013). In *Populus,* Wilkins *et. al.,* (2009) found that there are five 3R-MYB genes and 192 R2R3-MYB genes. The R2R3-MYB genes have been extensively studied and are related to phenylpropanoid metabolism and biotic and abiotic stress (Segarra *et. al.,* 2009). In *Arabidopsis thaliana,* the genes were found to be related to *Arabidopsis* response against aphid (Ambawat *et. al.,* 2013, Liu *et. al.,* 2011) and insect herbivore *Pieris rapae* (De Vos *et. al.,* 2006), as well as wound response. Moreover, the gene MYB transcription factor can also activate ABA signals to induce the

release of salicylic acid, thus enhancing plant defense mechanism towards biotic attack (Ambawat *et. al.,* 2013, Seo and Park*,* 2010).

4.4. Strain-specificity of *P. nigra* quantitative resistance to different *Mlp* strains

This study showed that genotype by strain interaction effects (GxS) were significant, and thus can be concluded that the variation in quantitative resistance can be influenced by the interaction between the genotypes and the strains. The interaction was the strongest for uredinia number variation, although experimental design also played a role in it, so the conclusion of strain specificity based on uredinia number might not be reliable. Uredinia size can be the suitable indicator for the strain-specificity of quantitative resistance since its variation was mostly genetic and interactive. The variation in latent period was more influenced by environmental variation rather than genetics or interaction.

The interaction effects were visualized in a reaction norm. The reaction norm showed that most genotypes had stable relationship across the *Mlp* strains. However, there were several genotypes that showed clear distinction in their interaction with each strain. It implied that these genotypes were behaving differently across different strains, thus showing their high specificity to different strains. There were 15 genotypes that showed high strain-specificity for latent period and uredinia size simultaneously. Based on their geographical distribution, these genotypes are coming from mostly Ramieres and Nohedes, as well as from Ticino, the Netherlands, Kuhkopf, Loire and Adour. Strain 09AX27 is the most interactive for latent period and strain 93JE3 is the most interactive for uredinia size.

The genetic correlation between traits across strains was also evaluated to see if traits are dependent on the strains at genetic level. The correlation between latent period across strains was lower than the correlation between uredinia size across strains. This confirms that at genetic level, uredinia size is not only dependent on genetic variation of the genotypes but also on the interaction between genotypes and strains. The result also implied that probably there were different genes that control the interaction between genotypes and strain for different traits.

Furthermore, it can be concluded that quantitative resistance that were split into latent period and uredinia size were genetically controlled and had strain-specificity, especially in genotypes from Ramieres and Nohedes. This confirms the strain-specificity of these genotypes in the smaller set of experimental data in Bastien *et. al.,* (2015).

Usually reaction norm is used to observe the response of individuals' traits under different environmental gradients (Marchal *et. al.,* 2019; and between life-history traits in Brommer *et. al.,* 2017). It is generally used in studies about the range of individuals' capacities in responding to various environmental conditions which usually have two extremes: plasticity, defined as the capacity of an individual to generate a range of alternative phenotypes over a heterogenous environment, and canalization, defined as individuals' insensitivity to the changes in the environment (Marchal, *et al.,* 2019; Arnold, *et al*., 2018; Lacaze et al. 2009). In this study, we used *Mlp* strains to define the different conditions of which the *P. nigra* genotypes were exposed to. This is a new approach of using reaction

norm, and can provide the idea of: 1) the optimal range of *Mlp* strains where *P. nigra* quantitative resistance is stable, i.e. is not strain-specific, and 2) the groups of genotypes whose quantitative resistance is strain-specific.

4.5. Conclusion and perspectives for efficient breeding programs

The study found 25 causal SNPs that were associated with quantitative resistance components. The causal SNPs were found in the chromosome 1 and 6 for latent period and uredinia size in all strains. The genes for these SNPs annotated serine theonine protein kinase and leucine rich repeat (LRR) protein family, the types of protein that is usually found in resistance (R) genomic region. There were also genes that were specific for uredinia number in strain 09AX27 and P72. The gene in 09AX27 annotated F-box protein and the gene in P72 annotated MYB transcription factor. These are the types of protein that mediate plant defense mechanism. Genotype by strain (GxS) effects were significant, indicating the strain-specificity of quantitative resistance components (especially latent period and uredinia size). Most genotypes were not highly interactive, but genotypes from Ramieres and Nohedes were highly interactive in terms of the expression of the latent period and uredinia size It indicates that quantitative resistance components in the genotypes from Ramieres and Nohedes were highly strain-specific.

The results from this study can be confirmed by increasing the sample size of *P. nigra* genotypes and by testing them with more *Mlp* strains. By increasing the number of individuals, it will hopefully increasing the statistical power of genome-wide association study (GWAS) and increasing the chances to find the causal SNPs. The genes found here can be used as the starting point to study their regulation pathways in plant defense mechanism, particularly in *P. nigra.* The construction of reaction norm where the different responses of genotypes in terms of their fitness (depicted by latent period, uredinia number, and uredinia size in the Y-axis) across different *Mlp* strains may create more analytical results. In this way, one can conclude the optimal range of *Mlp* strains where the *P. nigra* genotypes can be resistant to. The study can be further extended to genomic selection of the reaction norm parameters that construct the optimal range of *Mlp* strains by extracting the best linear unbiased predictions (BLUPs) of these parameters and fitting them in genome-wide association model. Finally, this study provides the insights about the genetic control of quantitative resistance in *P. nigra* genotypes from natural populations and the list of genotypes whose quantitative resistance is strain-specific.

References

Afzal, A. J., A. J. Wood, and D. A. Lightfoot. 'Plant Receptor-Like Serine Threonine Kinases: Roles in Signaling and Plant Defense'. *Molecular Plant-Microbe Interactions*® 21, no. 5 (2008): 507–17. https://doi.org/10.1094/MPMI-21-5-0507.

Albert, E. 'Etude de la variabilité génétique de la résistance partielle du peuplier noir (Populus nigra L.) à différentes souches de l'agent pathogène Melampsora larici-populina Kleb. responsable de la rouille foliaire et validation d'un QTL de résistance partielle.' (2013): Mémoire de fin d'Etudes. Ecole Supérieure d'Agriculture d'Angers. Angers, France.

Amadeu, R. R., C. Cellon, J. W. Olmstead, A. A. F. Garcia, M. F. R. Resende, and P. R. Muñoz. 'AGHmatrix: R Package to Construct Relationship Matrices for Autotetraploid and Diploid Species: A Blueberry Example.' *The Plant Genome* 9 (2016). doi:10.3835/plantgenome2016.01.0009

Ambawat, S., P. Sharma, N. R. Yadav, R. C. Yadav. 'MYB Transcription Factor Genes as Regulators for Plant Responses: An Overview'. *Physiol Mol Biol Plants* 19, no. 3 (2013): 307 – 321.

Aylott, M., J. E. Casella, I. Tubby, N. R. Street, P. Smith, and G. Taylor. 'Yield and Spatial Supply of Bioenergy Poplar and Willow Short-Rotation Coppice in the UK.' *The New phytologist* 178, no. 2 (2008): 358–370. doi: 10.1111/j.1469-8137.2008.02396.x.

Ballian, D., and U. Tröber. 'Genetic Characterization of European Black Poplar (*Populus nigra* L.) in Bosnia and Herzegovina'. *Original Scientific Paper* 141, no. 7–8 (2017): Abstracts. https://doi.org/10.31298/sl.141.7-8.2.

Barrès, B., C. Dutech, A. Andrieux, F. Halkett, and P. Frey. 'Exploring the Role of Asexual Multiplication in Poplar Rust Epidemics: Impact on Diversity and Genetic Structure'. *Molecular Ecology* 21, no. 20 (2012): 4996–5008. https://doi.org/10.1111/mec.12008.

Bastien C., A. Dowkiw, P. Faivre-Rampant, E. Albert, M. Villar, et al.. Are patterns of Populus nigra geographical differentiation for resistance to Melampsora larici-populina dependent of genetic variation of the pathogen populations?. *5th International Workshop on the Genetics of Tree-Parasite Interactions*, Orléans, France. (2015): 5th International Workshop on the Genetics of Tree-Parasite Interactions. Book of abstracts. ⟨hal-01269242⟩

Becheler, R., C. Xhaard, E. K. Klein, K. J. Hayden, P. Frey, S. De Mita, and F. Halkett. 'Genetic Signatures of a Range Expansion in Natura: When Clones Play Leapfrog'. *Ecology and Evolution 6*, no. 18 (2016): 6625–6632. https://doi.org/10.1002/ece3.2392.

Benetka, V., K. Černý, P. Pilařová, and K. Kozlíková. 'Effect of Melampsora Larici-Populina on Growthand Biomass Yield of Eight Clones of Populus Nigra.' *Journal of forest science (Praha)* 57, no. 2 (2011): 41–49. doi: 10.17221/51/2010-JFS.

Bradbury P.J., Z. Zhang, D. E. Kroon, T.M. Casstevens, Y. Ramdoss, E.S. Buckler. 'TASSEL: Software for association mapping of complex traits in diverse samples.' (2007): *Bioinformatics* 23: 2633-2635.

Bresson, A., V. Jorge, A. Dowkiw, V. Guerin, I. Bourgait, G. A. Tuskan, J. Schmutz, B. Chalhoub, C. Bastien, and P. Faivre Rampant. 'Qualitative and Quantitative Resistances to Leaf Rust Finely Mapped within Two Nucleotide-Binding Site Leucine-Rich Repeat (NBS-LRR)-Rich Genomic Regions of Chromosome 19 in Poplar'. *New Phytologist* 192, no. 1 (2011): 151–63. https://doi.org/10.1111/j.1469-8137.2011.03786.x.

Brommer, J. E., P. Kontiainen, and H. Pietiäinen. 'Selection on Plasticity of Seasonal Life-History Traits Using Random Regression Mixed Model Analysis: Selection on Reaction-Norm Properties'. *Ecology and Evolution* 2, no. 4 (2012): 695–704. https://doi.org/10.1002/ece3.60.

Brueggeman, R., A. Druka, J. Nirmala, T. Cavileer, T. Drader, N. Rostoks, A. Mirlohi, et al. 'The Stem Rust Resistance Gene Rpg5 Encodes a Protein with Nucleotide-Binding-Site, Leucine-Rich, and Protein Kinase Domains'. *Proceedings of the National Academy of Sciences* 105, no. 39 (2008): 14970–75. https://doi.org/10.1073/pnas.0807270105.

Chu Y, Q. Huang, B. Zhang, C. Ding, X. Su. 'Expression and molecular evolution of two DREB1 genes in black poplar (*Populus nigra*). *PLoS ONE* 9 (2014): e98334.

De Vos M et al. 'The Arabidopsis thaliana transcription factor AtMYB102 functions in defense against the insect herbivore Pieris rapae.' *Plant Signal Behav* 1 (2006):305–311

Dickmann D.I., and J. Kuzovkina. 'Poplars and Willow of the World, With Emphasis on Sylviculturally Important Species'. *Poplars and Willows: Trees for Society and the Environment.* (2014): 8 – 83.

do Vale, F. X. R., J. E. Parlevliet, and L. Zambolim. 'Concepts in Plant Disease Resistance'. *Fitopatologia Brasileira* 26, no. 3 (2001): 577–589. https://doi.org/10.1590/S0100-41582001000300001.

Dowkiw, A., and C. Bastien. 'Characterization of Two Major Genetic Factors Controlling Quantitative Resistance to *Melampsora Larici-Populina* Leaf Rust in Hybrid Poplars: Strain Specificity, Field Expression, Combined Effects, and Relationship with a Defeated Qualitative Resistance Gene'. *Phytopathology*® 94, no. 12 (December 2004): 1358–1367. https://doi.org/10.1094/PHYTO.2004.94.12.1358.

Dowkiw, A., C. Husson, P. Frey, J. Pinon, and C. Bastien. 'Partial Resistance to *Melampsora Larici-Populina* Leaf Rust in Hybrid Poplars: Genetic Variability in Inoculated Excised Leaf Disk Bioassay and Relationship with Complete Resistance'. *Phytopathology*® 93, no. 4 (April 2003): 421–427. https://doi.org/10.1094/PHYTO.2003.93.4.421.

Eberl, F., E. Perreca, H. Vogel, L. P. Wright, A. Hammerbacher, D. Veit, J. Gershenzon, and S. B. Unsicker. 'Rust Infection of Black Poplar Trees Reduces Photosynthesis but Does Not Affect Isoprene Biosynthesis or Emission'. *Frontiers in Plant Science* 9 (2018): 1733. https://doi.org/10.3389/fpls.2018.01733.

ElMalki R., V. Jorge, V. Segura, R. Bounon, A. Bérard, P. Faivre-Rampant, V. Guérin, P. Poursat, A. Dowkiw and C. Bastien. 'QTL mapping in *Populus nigra* reveals high strain specificity of partial resistance against the coadapted pathogen *Melampsora larici-populina'* (2013): *Plos one*.

Faivre-Rampant, P., G. Zaina, V. Jorge, S. Giacomello, V. Segura, S. Scalabrin, V. Guérin, et al. 'New Resources for Genetic Studies in *Populus Nigra* : Genome-Wide SNP Discovery and Development of a 12k Infinium Array'. *Molecular Ecology Resources* 16, no. 4 (2016): 1023–1036. https://doi.org/10.1111/1755-0998.12513.

Frichot, E., F. Mathieu, T. Trouillon, G. Bouchard, and O. François. 'Fast and Efficient Estimation of Individual Ancestry Coefficients'. *Genetics 196*, no. 4 (2014): 973–983. https://doi.org/10.1534/genetics.113.160572.

Gérard, P. R., C. Husson, J. Pinon, and P. Frey. 'Comparison of Genetic and Virulence Diversity of *Melampsora Larici-Populina* Populations on Wild and Cultivated Poplar and Influence of the Alternate Host'. *Phytopathology*® 96, no. 9 (2006): 1027–1036. https://doi.org/10.1094/PHYTO-96-1027.

Goodstein D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam, and D. S. Rokhsar. 'Phytozome: a comparative platform for green plant genomics'. *Nucleic Acids Research 40* (2012): 1178 – 1186.

Jorge, V., A. Dowkiw, P. Faivre-Rampant, and C. Bastien. 'Genetic Architecture of Qualitative and Quantitative Melampsora Larici-Populina Leaf Rust Resistance in Hybrid Poplar: Genetic Mapping and QTL Detection'. *New Phytologist* 167, no. 1 (2005): 113–27. https://doi.org/10.1111/j.1469-8137.2005.01424.x.

Lacaze, X, P M Hayes, and A Korol. 'Genetics of Phenotypic Plasticity: QTL Analysis in Barley, Hordeum Vulgare'. *Heredity* 102, no. 2 (February 2009): 163–73. https://doi.org/10.1038/hdy.2008.76.

Lefèvre, F., M. C. Goué-Mourier, P. Faivre-Rampant, and M. Villar. 'A Single Gene Cluster Controls Incompatibility and Partial Resistance to Various *Melampsora Larici-Populina* Races in Hybrid Poplars'. *Phytopathology*® 88, no. 2 (1998): 156–163. https://doi.org/10.1094/PHYTO.1998.88.2.156.

Liu R, B. Lü, X. Wang, C. Zhang, S. Zhang, J. Qian, L. Chen, H. Shi, H. Dong. 'Thirty-seven transcription factor genes differentially respond to a hairpin protein and affect resistance to the green peach aphid in Arabidopsis.' *J Biosci* 35 (2010):435–450. doi:10.1007/s12038-010-0049-8

Marchal A., C. D. Schlichting, R. Gobin, P. Balandier, F. Millier, *et al.*, 'Deciphering Hybrid Larch Reaction Norms Using Random Regression.' G3: Genes, Genomes, Genetics 9 (2019): 21–32. https://doi.org/10.1534/g3.118.200697

Maupetit, A., R. Larbat, M. Pernaci, A. Andrieux, C. Guinet, A-L. Boutigny, B. Fabre, P. Frey, and F. Halkett. 'Defense Compounds Rather Than Nutrient Availability Shape Aggressiveness Trait Variation Along a Leaf Maturity Gradient in a Biotrophic Plant Pathogen'. *Frontiers in Plant Science* 9 (2018): 1396. https://doi.org/10.3389/fpls.2018.01396.

Muñoz, F. and L. Sanchez. 'breedR: Statistical Methods for Forest Genetic Resources Analysts.' (2020) R package version 0.12-5. https://github.com/famuvie/breedR

Narasuka Y, K. Nakashima, Z.K. Shinwari, Y. Sakuma, T. Furihata, H. Abe, M. Narusaka, K. Shinozaki, K. Yamaguchi-Shinozaki K. 'Interaction between two cis-acting elements, ABRE and DRE, in ABA dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses.' *Plant J* 34 (2003):137–148

Pariaud, B., V. Ravigné, F. Halkett, H. Goyeau, J. Carlier, and C. Lannou. 'Aggressiveness and its role in the adaptation of plant pathogens.' *Plant Pathology* 58 (2009): 409–424. doi: 10.1111/j.1365-3059.2009.02039.x

Pei, M. H., C. Ruiz, J. Harris, and T. Hunter. 'Quantitative inoculations of poplars with Melampsora larici-populina.' *European Journal of Plant Pathology* 109 (2003): 269–276. doi: 10.1023/A:1022822503139.

Pei, M. H., C. Ruiz, T. Hunter, G. M. Arnold, and C. Bayon. 'Quantitative relationships between inoculum of *Melampsora larici-epitea* and corresponding disease on Salix'. *Plant Pathology* 51 (2002): 443 – 453. doi: 10.1046/j.1365-3059.2002.00739.x.

Persoons, A., K. J. Hayden, B. Fabre, P. Frey, S. De Mita, A. Tellier, and F. Halkett. 'The Escalatory Red Queen: Population Extinction and Replacement Following Arms Race Dynamics in Poplar Rust'. *Molecular Ecology* 26, no. 7 (2017): 1902–1918. https://doi.org/10.1111/mec.13980.

Polle, A., D. Janz, T. Teichmann, and V. Lipka. 'Poplar Genetic Engineering: Promoting Desirable Wood Characteristics and Pest Resistance.' *Applied microbiology and biotechnology* 97, no. 13 (2013): 5669–5679. doi: 10.1007/s00253-013-4940-8.

R Core Team. 'R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.' (2020) URL https://www.R-project.org/.

river systems: consequences for conservation and restoration.' *Forest Ecology and Management* 255 (2008): 1388–1399.

Robert, C., M. O. Bancal, and C. Lannou. 'Wheat leaf rust uredospore production and carbon and nitrogen export in relation to lesion size and density.' *Phytopathology* 92 (2002): 762–768. doi: 10.1094/PHYTO.2002.92.7.762.

Segarra G, E. S; Van der, I. Trillas, C. M. J. Pieterse. 'MYB72, a node of convergence in induced systemic resistance triggered by a fungal and a bacterial beneficial Microbe.' *Plant Biol* 11 (2009):90–96

Seo PJ, and C.M. Park. 'MYB96-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in Arabidopsis.' *New Phytol* 186 (2010):471–483

Smulders M. J. M., J.E. Cottrell, F. Lefevre et al. 'Structure of the genetic diversity in black poplar (Populus nigra L) populations across European

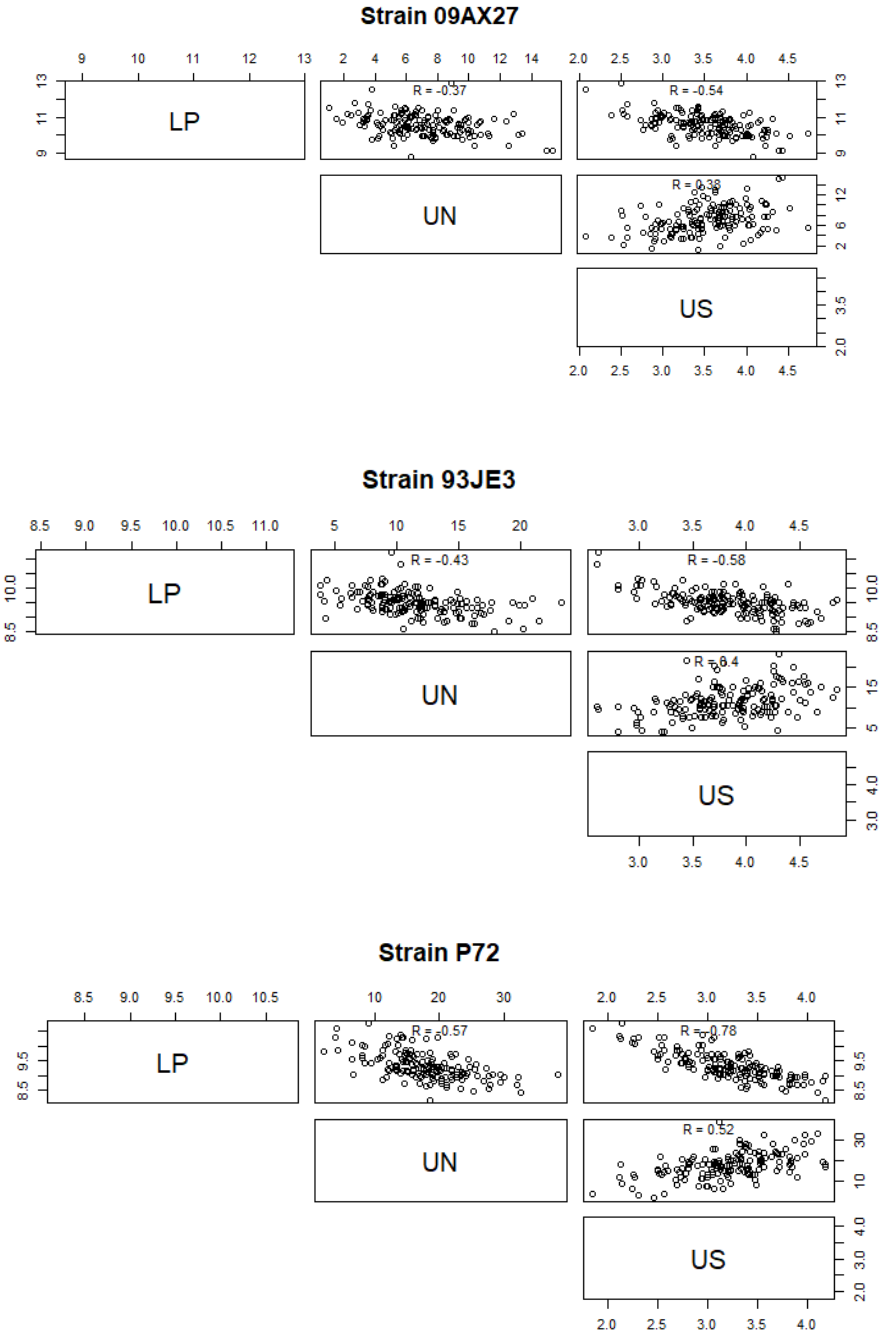
Solanki, S., J. Richards, G. Ameen, X. Wang, A. Khan, H. Ali, A. Stangel, et al. 'Characterization of Genes Required for Both Rpg1 and Rpg4-Mediated Wheat Stem Rust Resistance in Barley'. *BMC Genomics* 20, no. 1 (2019): 495. https://doi.org/10.1186/s12864-019-5858-z.

Zheng, S. -X., J. -G. Zhang, C. -Y. He, E. -J. Bao, A. -G. Duan, Y. -F. Zeng, and A. -M. Liu. 'Genetic Diversity of Section LLeuce in Populus along the Erqis River'. *Forest Research* 2 (2015): 222 - 229.

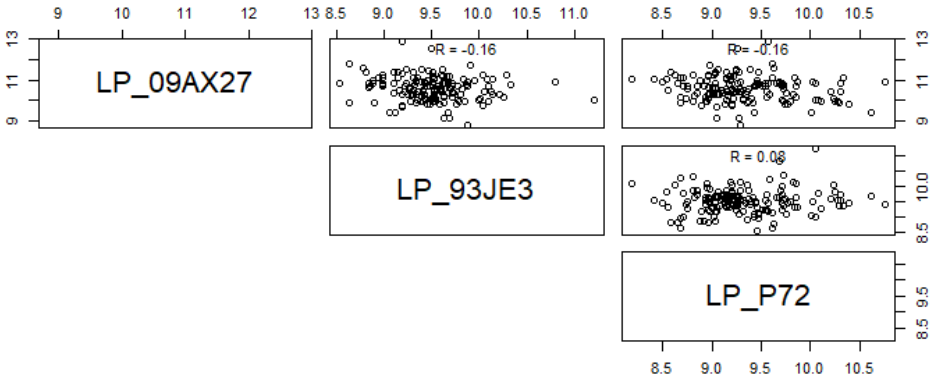
Annex 1 Linear model for trait variation in 09AX27, 93JE3, and P72 for 154 *P. nigra* genotypes

Variance components	Degree of freedom	09AX27			93JE3			P72		
		LP	UN	US	LP	UN	US	LP	UN	US
Genotype	153	0.412 ± 0.077**	16.281 ± 2.349**	0.451 ± 0.069**	0.385 ± 0.054**	28.217 ± 4.063**	0.531 ± 0.070	0.403 ± 0.054**	77.072 ± 10.582**	0.475 ± 0.062**
Block	4	0.049 ± 0.040	1.424 ± 1.107**	0.012 ± 0.012	0.029 ± 0.023	2.609 ± 2.016**	0.081 ± 0.059	0.004 ± 0.005**	1.764 ± 1.617**	0.024 ± 0.019
Residual		1.043 ± 0.065	21.342 ± 1.235	0.595 ± 0.037	0.441 ± 0.025	38.112 ± 2.157	0.395 ± 0.023	0.342 ± 0.020	82.443 ± 4.656	0.318 ± 0.018

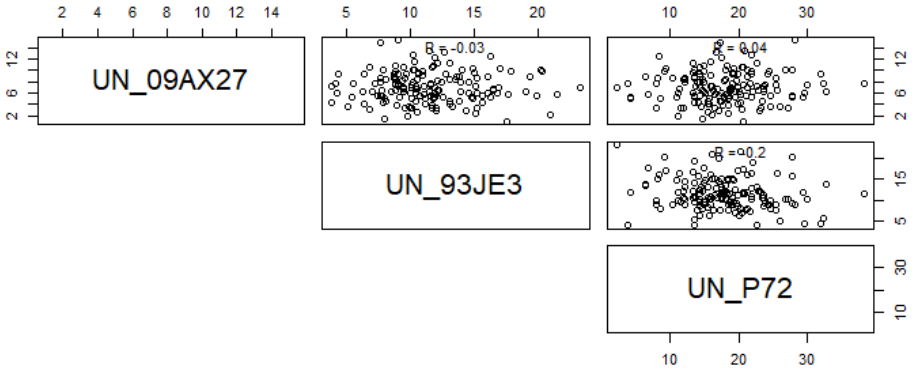
Annex 2 Correlation between traits and strains



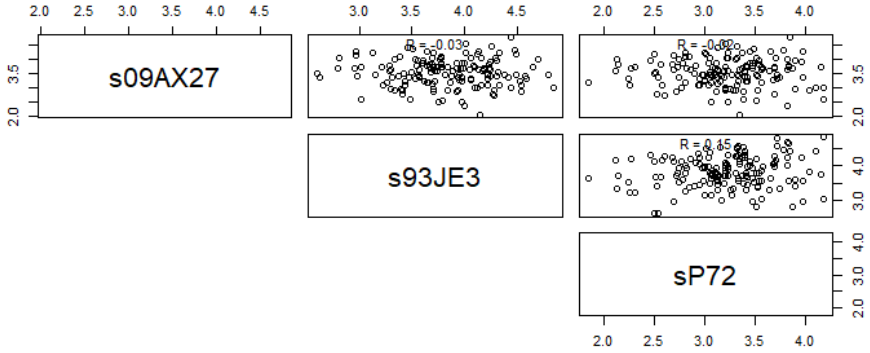
Latent period correlation between all strains



Uredinia number correlation between all strains



Uredinia size correlation between all strains



Annex 3 GxS model

LP		
Fixed effects		
Variance components	General mean	Standard error
09AX27	10.606	0.080
93JE3	9.508	0.080
P72	9.308	0.080
Random effects		
Variance components	Estimated variances	Standard error
Genotypes	0.230	0.038**
Block	0.015	0.011
GxS	0.167	0.025**
Residual	0.602	0.020
UN		
Fixed effects		
Variance components	General mean	Standard error
09AX27	7.021	0.762
93JE3	11.446	0.761
P72	17.797	0.760
Random effects		
Variance components	Estimated variances	Standard error
Genotypes	25.015	3.826**
Block	1.313	1.002**
GxS	15.433	2.043**
Residual	48.363	1.587
US		
Fixed effects		
Variance components	General mean	Standard error
09AX27	3.500	0.094
93JE3	3.823	0.093
P72	3.226	0.093
Random effects		
Variance components	Estimated variances	Standard error
Genotypes	0.315	0.046**
Block	0.026	0.018
GxS	0.160	0.021**
Residual	0.445	0.015

Annex 4 Reaction norm between 154 *P. nigra* genotypes and 3 *Mlp* strains

Genotype by strain interaction plot for latent period (LP) (left) and uredinia size (right) for 154 P. nigra inoculated with three Mlp strains. Strain 09AX27, 93JE3, and P72 are plotted on the X-axis and predicted GxS interaction generated from the linear mixed model is ploted on the Y-axis. The lines are representing P. nigra genotypes. Highly interactive genotypes are shown through their fluctuating lines across strains, and weakly interactive genotypes are clustering near Y = 0.0.