Difference in genetic architecture is a key driver in plant-to-plant interaction: case study of height and rust damages in Populus nigra

Authors: Mathieu Tiret[1,3*], Leopoldo Sánchez[2], Martin Lascoux[1].

Affiliations:

- 1: Program in Plant Ecology and Evolution, Department of Ecology and Genetics, EBC and SciLife Lab, Uppsala University, Uppsala, Sweden
- 2: BioForA, INRAE, ONF, Orléans, France.
- 3: IGEPP, INRAE, Institut Agro, Université de Rennes, Domaine de la Motte, Le Rheu, France.

* corresponding author: mathieu.tiret@inrae.fr

Abstract: (179 words)

Main text: 5,100 words.

keywords: group selection, genomic selection, Populus nigra, quantitative genetics.

22   **Abstract**:

23

24   Adaptation to competition during the evolutionary history of plants leads to a reduction of the

25   overall group performance of plant elite material in the field. Traits related to competitive

26   ability can be affected by the presence or absence of kin in the neighborhood. Consequently,

27   local relatedness might reveal plant-to-plant interaction that can enhance the predictive

28   abilities of genomic models when accounted for. However, recurrent family selection

29   increases relatedness among all elite material and can blur the effect of relatedness on

30   neighbors' phenotypes. To overcome this difficulty, we analyzed data from the French

31   breeding program of *Populus nigra* L., where 1,452 genotypes were replicated six to eight

32   times, each time encountering a different neighborhood. We assessed local relatedness and

33   investigated genomic estimated breeding values on tree height and rust vulnerability with a

34   single-step GBLUP incorporating local relatedness as a covariate. Results suggests that

35   including local relatedness as an additional factor to GBLUP models is significantly more

36   important for rust resistance than for tree height, mainly due to their contrasting genetic

37   architecture – oligogenic for the former and polygenic for the latter.

38

Introduction

40 Breeding generally consists in redirecting evolutionary trade-offs towards new equilibria that

41 often decrease fitness in favor of other economically interesting traits, or towards paths that

42 were insufficiently explored by evolution in natural populations, such as group selection

43 (Weiner, 2019). Although the importance of group selection in nature is controversial

44 (Maynard Smith, 1964; Lewontin, 1970; Wilson & Wilson, 2007; Okasha, 2020), in a

45 breeding program, group selection can be used to improve overall performance (e.g., robust

46 productivity, tolerance to biotic or abiotic stress, or long-term response to selection), by

47 keeping only those candidates that contribute favourably to plant-to-plant interactions. The

48 performance of an individual can then be decomposed into direct and indirect breeding

49 values (also called direct and indirect genetic effects), where the former is the intrinsic value

50 of an individual without any competitors, and the latter the ability of the individual to increase

51 or decrease performance due to competition (Griffing, 1967; Bijma & Wade, 2008). Breeding

52 can benefit from accounting for both, especially in contexts in which the indirect genetic

53 effect is strong, such as when intraspecific competition is strong among candidates (Weiner

54 et al., 2017). Indeed, selfish strategies (in the sense of selfish genes), although increasing

55 individual fitness, are often costly for the populations in breeding schemes - and breeding

56 circumventing such strategies ended up with very successful results (e.g., Donald, 1981;

57 Donald and Hamblin, 1983; Weiner et al., 2010). Competitive behavior can, therefore, be a

58 burden that breeders might want to select against. Despite the early success of some

59 ideotypes optimizing group performance, current breeding programs do not necessarily

60 focus on integrating this dimension (e.g., Denison et al., 2003; Murphy et al., 2017;

61 Montazeaud et al., 2020). Consequently, the pool of elite genotypes used nowadays might

62 be mostly composed of selfish genotypes, since early mass selection, bypassing group

63 dynamic, favored selfish genotypes for their vigorous phenotypes (Murphy et al., 2017). It is

64   therefore increasingly important in order to reach sustained genetic gain in breeding

65   programs to counteract the effect of natural selection acting at the deployment phase, and to

66   correct past artificial selection trajectories focussed on competitive (i.e., selfish) individuals .

67

68   In the specific case of plant breeding, resource competition between candidates occurs with

69   the nearest plants (Casper et al., 1997; Milbau et al., 2007; File et al., 2011). The intensity of

70   competition among neighbors may depend on their relatedness, or local relatedness, either

71   because of kin selection, which predicts that relatives may cooperate (Hamilton, 1964), or

72   niche partitioning, which predicts, on the contrary, that similar individuals may compete for

73   similar resources (Silvertown, 2004). Therefore, plant-to-plant interaction can partly be due

74   to local relatedness in one way or the other (Cahill et al., 2011), so that classic models can

75   be improved by incorporating the information of neighboring genotypes. Although kin

76   recognition was reported to be not particularly relevant in crop species because of their

77   already high relatedness (e.g., Murphy et al., 2017), we might be able to detect it

78   nonetheless in breeding schemes that have been less intensive, such as forest tree

79   breeding. Group selection is in fact expected to give better genetic progress for perennial

80   species, since (i) inbreeding coefficient is much lower than in crop species (even though

81   inbreeding depression can be higher; Lesaffre & Billiard, 2021), and (ii) positive or negative

82   interactions between neighbors last a lifetime.

83

84   Part of the phenotypic variance is due to the micro-environmental heterogeneity, and

85   accounting for it has been shown to improve predictions (Cappa et al., 2022). Among the

86   micro scale heterogeneity, local relatedness by variable neighborhood can create

87   heterogeneity, hence a phenotypic variance that, when not accounted for, can bias genetic

88   estimates. To account for heterogeneity caused by local relatedness, we chose to conduct

89   our analyses on *Populus nigra* L.. Black poplar is a Eurasian riparian forest tree that

90   contributes, as a parent along with *Populus deltoides* Bartr. ex Marsh., to one of the most

91   widely used hybrid tree in forest breeding (*Populus x canadensis*), and is widely deployed as

92   clones, constituting a model in which kin recognition, if any, should be the strongest. The

93   dataset, stemming from the French breeding program with 1,452 genotypes replicated six to

94   eight times, created different neighborhoods for each replicated genotype (Pégard et al.,

95   2020), all resulting from controlled crosses of 34 parents from natural populations, some of

96   which were already used in the breeding programme for their performance. To detect any

97   signal of group selection, we focused on two quantitative traits relevant for breeding, tree

98   height and rust vulnerability. Growth of Black poplar is particularly susceptible to foliar rust,

99   caused by a fungus (*Melampsora larici-populina* Kleb.) that leads to a reduction in

100   photosynthesis efficiency, and might lead to the death of susceptible clones (Legionnet et

101   al., 1999). Rust development is strongly influenced by susceptibility of nearby trees, hence

102   putatively linked to local relatedness, whereas tree height is a poor indicator of competition

103   and is mostly related to the fertility of the trial. Height is known to have a polygenic

104   architecture (Du et al., 2016), and rust resistance to have a mono- or oligogenic architecture

105   (e.g., Jorge et al., 2005), implying that relatedness – as the amount of shared alleles – will

106   have different consequences on those two traits: genetic redundancy of height probably

107   making it less sensitive to local relatedness than rust resistance.

108

109   In order to assess the contribution of local relatedness in the context of genomic selection

110   (Meuwissen et al., 2001), we performed genomic evaluations with a multitrait single-step

111   Genomic Best Linear Unbiased Predictor (ssGBLUP; Legarra et al., 2009; Christensen et al.,

112   2010). Estimating breeding values with ssGBLUP is a solid baseline in forest trees as

113   recently shown (e.g., Cappa et al., 2019; Ratcliffe et al., 2017), especially its multitrait

114   version that is known to give higher accuracy (Calus et al., 2011). By including local

115   relatedness and spatial micro-environmental autocorrelation as random effects in ssGBLUP

116   models, we assessed whether plant-to-plant interactions – phenotypic covariance between

117   neighbors – could be due to local relatedness. We assessed local relatedness with different

118   subsets of SNPs, each maximizing or minimizing one of the following features: Minor Allele

119   Frequency, Ancestry Informativeness Coefficient, significance in a Genome Wide

120   Association Study, or heritability. We show that local relatedness significantly impacts the

121   phenotype of an individual and, when surrounded by related individuals, rust damage is

122   significantly more important, whereas no interaction was visible on tree height. Differences in

123   genetic architecture seem to play a role in plant-to-plant interactions.

124

Material and Methods

126   **Plant material.** The study is based on previously published data (Pégard et al., 2019, 2020).

127   Seventeen male and 17 female *Populus nigra* were sampled in France, from 21 natural

128   populations with no a priori selection and 13 from the French poplar breeding program for

129   their performance on a range of traits, including growth and rust resistance. These

130   individuals were used as parents in a factorial crossing plan and double-pair mating scheme.

131   Clones of both parents and progeny, and a hundred additional individuals from breeding

132   programs, were grown in four sequential experimental trials (parallel strips of two lines of

133   trees, with an inter-strip spacing of 1m and an inter-strip spacing of 2 metres) at the same

134   location (Guéméné-Penfao, France, 47°37'59"N, 1°49'59"W), each with a randomized

135   incomplete 26 block design of single tree-plots (0.9 m x 1.0 m). In total, the experiment

136   comprised 10,301 trees (including 7,169 trees with genotype information), 42 full-sib

137   families, 5 half-sib families, 1,452 genotypes (each replicated on average 7.09 times), and

138   an Unknown Parent Group (UPG) of 105 genotypes (Table S1). Family size ranged from 1 to

139   119, with an average of 30.2 genotypes per family. Twenty genotypes and 17 families were

140   shared across the four trials.

141

142   **Phenotype measurements.** We focused on two phenotypic traits on one-year-old trees:

143   height and rust vulnerability (both measured on field). Height was assessed with a

144   graduation rod (in cm), and rust vulnerability (natural infection) was assessed on a scale of 1

145   to 9 : 1 when no rust was observed on the tree and 9 when more than 75% of lamina was

146   covered by rust on more than 25% of the leaves (Legionnet et al., 1999). The homogeneity

147   of rust pressure across the trial was assured by the level of infection measured on control

148   individuals, being equally infected across blocks (clone ID: BDG). Some missing values were

149   present across the trials (dead trees or unexploitable data), but at least 95.29%, 96.05%,

150   and 95.13% of the trees were phenotyped for height, rust vulnerability, or both traits,

151   respectively. Given the low missing rate, we assumed that missing values only marginally

152   unbalanced the block design, and so we used default parameters of the software packages

153   for handling missing values when analyzing the data.

154

155   In each trial, every phenotype was first corrected for micro-environmental heterogeneity with

156   a multitrait mixed model as follows:

157

158    $Y_{ijkl} = \mu + BLOCK_i + SPLINE + e_{ijkl}$

159

160   where 'Y_{ijkl}' is the l-th phenotype of the k-th replicate of the j-th genotype in the i-th trial, μ is

161   the grand mean, '$BLOCK_i$' the random effect of the i-th block, '$SPLINE_i$' B-spline spatial

162   autocorrelation correction over the surface of rows and columns (Cappa & Cantet, 2007),

163   and 'e_{ijkl}' the residuals. Both the block effect and the residuals follow a Gaussian distribution

164   with diagonal covariance matrices, and the model was fitted separately by trial. We used the

165   function *remlf90* (R package *breedR*; Munoz and Sanchez, 2014) with the parameters

166   'model' set to 'splines' and 'method' set to 'em'. All further analyses were performed on the

167   residuals of this model, i.e., on spatially adjusted phenotypes.

168

169   **Genomic data.** Out of the 1,452 genotypes, 1,034 were genotyped (25 parents and 1,009

170   offspring), using a *Populus nigra* 12K custom Infinium Bead-Chip (Illumina, San Diego, CA)

171   (Faivre-Rampant et al., 2016). Details of the DNA extraction protocol, bioinformatic pipelines,

172   and SNP mapping on the 19 chromosomes are given in Faivre-Rampant et al. (2016) and

173   Pégard et al. (2019, 2020). It is worth noting that the SNP chip array included SNP in QTLs

174   and expression candidate genes associated with, among other traits, rust resistance, but not

175   with vertical growth. We retained SNPs with a Minor Allele Frequency (MAF) > 5%; in total,

176   7,129 SNPs were retained for further analyses (out of 7,513 initial SNPs).

177

178 **Genomic relationship matrix.** In order to build the relationship matrix necessary for

179 inferring breeding values, we combined pedigree and genomic information (Legarra et al.,

180 2009; Christensen et al., 2010). The pedigree relationship matrix A^Γ (following the notation of

181 Legarra et al., 2009) was built using the tabular rule method (Emik and Terrill, 1949),

182 modified to account for a single metafounder (assuming a single base population for the

183 Unkown Parent Group or UPG; Legarra et al., 2015), with a self-relationship of $\gamma = 8\sigma_p^2$, with

184 σ_p^2 the variance of allele frequencies across markers (Garcia-Baccino et al., 2017). The

185 combined relationship matrix H was summarized in Aguilar et al. (2019), and is:

186

187 $H^{-1} = (A^\Gamma)^{-1} + (0,0;0; G^{-1} - (A_{22}^\Gamma)^{-1})$ (1),

188

189 where A_{22}^Γ is the submatrix of A^Γ corresponding to the genotyped individuals, $G = (1-\alpha) (\omega_a +$

190 $\omega_b \times G_V) + \alpha A_{22}^\Gamma$, with G_V the genomic relationship matrix estimated from the SNPs and

191 scaled following the first scaling method of Van Raden (2008), α a scaling parameter (here

192 equal to 0.05), and ω_a and ω_b chosen to equate the average inbreeding and the average

193 relationships in G_V and A_{22}^Γ as in Christensen et al. (2012). As the sample size was relatively

194 small, the matrix H^{-1} was obtained by simply inverting H with the function *solve* (R package

195 *base*).

196

197 **Population structure.** In order to control for population structure and avoid spurious

198 associations, we computed a distance matrix $1 - H^*$, where H^* is the correlation matrix

199 obtained from H using the function *cov2cor* (R package *stats*), and the minus sign is to

200 convert similarities to distances. We then used Multidimensional Scaling (MDS) on the

201 distance matrix with the function *cmdscale_lanczos* (R package *refund*; Goldsmith et al.,

202 2022; Miller, 2022) with the parameter 'k' set to 5 (hence explaining > 90% of the variance),

203 and 'eig' set to 'TRUE'. Negative eigenvalues were set to null, and the variance explained by

204 an axis was computed as the ratio between its squared eigenvalue and the sum of the

205 squared eigenvalues. Additionally, to assess how well the family structure was captured, we

9   9

206 measured the variable V_F defined as the average within family variance on the first five axes

207 (the lower the V_F, the better the family structure is captured). To make MDSs comparable,

208 we normalized V_F by the sum of squares of the first five eigenvalues.

209

210 **Breeding value inferences.** The spatially adjusted phenotypes were then used to infer the

211 breeding values with a multitrait single-step Genomic Best Linear Unbiased Predictor

212 (ssGBLUP), as follows:

213

214 $Y_{ijkl} = \mu + mds1_j + mds2_j + TRIAL_i + ANIMAL_j + e_{ijkl}$,   (2)

215

216 where 'Y_{ijkl}' is the l-th phenotype of the k-th replicate of the j-th genotype in the i-th trial, μ is

217 the grand mean, '$mds1_j$' and '$mds2_j$' the fixed effects on the j-th genotype of the first and

218 second axes of the MDS, respectively, '$TRIAL_i$' the random effect of the i-th trial, '$ANIMAL_j$'

219 the polygenic effect of the j-th genotype, and 'e_{ijkl}' the residuals. Both the trial effect and the

220 residuals follow a Gaussian distribution with diagonal covariance matrices, and the polygenic

221 effect follows a Gaussian distribution with a covariance matrix equal to k.H x Σ, where k is a

222 scaling parameter equal to 1 - γ/2 (Legarra et al., 2015), H the combined relationship matrix,

223 x the Kronecker product, and Σ the 2x2 covariance matrix between height and rust

224 vulnerability. Estimates of coefficients and variance components of model (2) were obtained

225 with the function *remlf90* (R package *breedR*) through one run with the 'method' parameter

226 set to 'em' until convergence, then another run with the parameter 'method' set to 'ai' and the

227 parameter 'progsf90.option' set to 'maxrounds 1'. The output estimate of '$ANIMAL_j$' is the

228 Estimated Breeding Value (EBV) of the j-th genotype. The additive variance was obtained

229 from the diagonal elements of Σ, and used to estimate heritability (the denominator of which

230 also comprises the residual variance). The predictive ability (PA) of the model was defined

231 as Pearson's correlation coefficient between the breeding values re-estimated from the

232 effect sizes (see below) and the adjusted phenotypes corrected by the trial effect (i.e., Y_{ijkl} −

233 $TRIAL_i$).

10   10

234

235   **Genome-Wide Association Study (GWAS).** Following the equations of Aguilar et al.

236   (2019), the effect size of the SNPs was back-solved from the EBVs, along with their

237   standard error and their p-values. This process was performed trait by trait (i.e., without

238   accounting for the multitrait dimension) without loss of information, and taking all effects as

239   fixed, as the Pearson's correlation between the ssGBLUP's EBV and the breeding values re-

240   estimated by the effect sizes was greater than 0.999.

241

242   With the estimated effect sizes and the standard errors, we also computed an alternative to

243   the p-value accounting for multiple testing: the local false sign rate (lfsr) estimated with an

244   empirical Bayes approach for adaptive shrinkage using the function *ash* (R package *ashr*;

245   Stephens, 2017). In order to detect GWAS peaks (regions in the chromosomes with a high

246   density of significant SNPs), we used a custom peak detection script on the profile of lfsr

247   (Tiret & Milesi, 2021) with default parameters. The script filtered out significant hits that are

248   not statistically detected as a peak which is defined here as an abnormally high

249   concentration of significant SNPs in a small chromosomic area.

250

251   **Neighborhood.** In order to account for the effect of the neighborhood, we estimated the

252   breeding values with an additional fixed effect, the local relatedness, defined as the average

253   relationship (from the matrix H) between a focal individual and its neighbors (eight or less if

254   on boundaries; as the king's moves on a chessboard). Local relatedness was orthogonal to

255   any micro-environmental effect, because of the random block design. The re-estimated

256   breeding values were denoted nEBV.

257

258   **SNP subsets.** We investigated the effect of sampling SNPs on the ssGBLUP (i.e., weighted

259   ssGBLUP). We selected subsets based on population and quantitative genetics features

260   computed after a first run of ssGBLUP: MAF, Ancestry Informativeness Coefficient (AIM;

261   Rozenberg et al., 2003), and for each trait, significance in the GWAS, and additive variance.

262   For each feature, we kept one-tenth of the whole set (713 SNPs, a sample size close to

263   previous studies on weighted GBLUP, e.g., Li et al., 2018) with either the smallest or the

264   largest values, ending up with 12 subsets. AIM was estimated per family with the R script

265   provided by Cappa et al. (2022). Additive variance explained by a SNP was estimated as

266    $2p(1-p)a^2$, where p is the MAF, and a the effect size (assuming no dominance nor interaction

267   deviation of the genetic variance; p.129, Falconer, 1989). We re-estimated the combined

268   relationship matrix for each SNP subset by re-estimating in the equation (1) the genomic

269   relationship matrix G_v with the SNP subset. With the re-estimated matrices H, we re-

270   performed the MDSs, and re-estimated the PAs of the EBVs and nEBVs (where local

271   relatedness were measured with the re-estimated H).

272

273   **Statistical tests.** For each SNP subset, we computed the confidence interval of the

274   Predictive Ability (PA) with a 1000 iteration bootstrap (re-sampling individuals), without re-

275   estimating the covariance matrix Σ. Here, the bootstrap was only used to assess the

276   variance, so that we centered the bootstrapped PA around the true PA. To assess the effect

277   of sampling, these subsets were compared to 1000 random samples of 713 SNPs (the size

278   of the SNP subsets). Comparisons were performed with a Student's one-sample t-test

279   (denoted t_1 with a degree of freedom or df of 999 corresponding to the number of bootstrap

280   iterations), or a Welch two-sample t-test (denoted t_2 with a varying df), both using the

281   function *t.test* (R package *stats*). The sign of the t statistic is arbitrarily reported as positive

282   when greater than a baseline (e.g., the true PA), and negative when not. Models were

283   compared with a likelihood ratio test with a df of 1 using the function *lrtest* (R package *lmtest*;

284   Zeileis & Hothorn, 2002). All scripts were written in R v4.1.3 (R Core Team, 2022).

285

286

Results

288 Tree height and rust vulnerability both exhibited high heritabilities (0.469 and 0.666,

289 respectively), high predictive abilities (0.677 and 0.755 respectively), and a strong

290 independence, with an estimated genetic correlation of 0.022. The GWAS did not overfit

291 phenotypic variation: swapping phenotypes (permutation test) among family members of the

292 same block (i.e., only randomly mismatching Mendelian sampling) significantly reduced the

293 average significance of the SNPs ($t_1 < -2.59 \times 10^2$, $p < 0.001$). Only one genetic variant was

294 significantly associated with rust vulnerability, located on chromosome 1, and none for tree

295 height (Fig. 1A, 1B). The Q-Q plot on rust vulnerability revealed that some SNPs were

296 significantly associated with a level of significance well above the 95% confidence interval

297 (Fig. 1D); most of them being in the identified peak. None of the SNPs were significantly

298 associated with tree height after Bayesian shrinkage, and Q-Q plot revealed a classic

299 population structure overcorrection (Fig. 1C).

Common and most heritable alleles structured families

301 The population under study is structured in a hierarchical way: first into populations, then into

302 families. Population structure, captured by the first eigenvalues (Patterson et al., 2006),

303 substantially shaped genetic diversity, as the first two eigenvalues alone explained more

304 than 50% of the variance (40.0% and 24.5% respectively; Fig. S1). Because the population

305 underwent family selection during breeding in the past, family structure is strong and can

306 even be apprehended visually (Fig. 2). Several statistics significantly supports the strong

307 family structure: (i) randomly shuffling the family ID increased by 9.96 times the within

308 "family" variance ($t_1 = 3.70 \times 10^3$, $p < 0.001$); (ii) when restricting to common alleles, the first

309 two eigenvalues of the PCA still explained a large part of the variance (56.2%), showing that

310 family structure is bound to selection, as expected under family selection.

311

312　Structure is generally assumed to be reflected by polymorphism at neutral genes scattered

313　across the genome, to the point where samples of SNPs can reveal population structure

314　quite accurately if they are sufficiently large. In order to understand what drove the family

315　structure, similarly to using annotation to focus on some SNPs, we subset 10% of the SNPs

316　first randomly and then according to their MAF, AIM, significance in a GWAS, or additive

317　variance, thus redefining relatedness with a different set of linkage disequilibrium. In our

318　dataset, random samples of SNPs, performed quite poorly compared to the whole set of

319　SNPs ($t_1 = 2.12 \times 10^2$, $p < 0.001$), although the SNPs were scattered along the genome.

320　However, SNPs had a different structuring power at the family level according to the way

321　they were sampled (Fig. 3): SNPs with the largest AIM, performing better than random

322　samples ($t_1 = 3.86 \times 10^2$, $p < 0.001$), and even than the whole set; and SNPs with the largest

323　MAF and the largest additive variance, performing better than random samples ($t_1 > 1.09 \times$

324　10^2, $p < 0.001$). This result suggests that family structure information is carried by patches

325　over the genome instead of being scattered across the genome.

326　## Local relatedness improves ssGBLUP fit for rust vulnerability

327　Although the growth of focal individuals was not affected by their neighbors, damage due to

328　rust was more severe the more related the focal individuals were with their neighbors.

329　Indeed, the genomic models incorporating local relatedness as a covariate had significantly

330　higher likelihood than without for rust vulnerability ($\chi^2 = 9.83$, $p = 0.002$), but not for tree

331　height ($\chi^2 = 0.09$, $p = 0.236$). As for family structure which is driven by some SNPs, different

332　categories of SNPs had a different impact on focal phenotypes. When accounting for

333　relatedness via the different subsets of SNPs, the Akaike Information Criterion (AIC)

334　generally decreased (Fig. S2). However, for rust vulnerability, only the subset of SNPs with

335　the largest MAF significantly increased the likelihood of the model ($\chi^2 = 4.02$, $p = 0.04$), that

336　is the one that best captured the family structure. Likelihood was not significantly improved

337　by SNPs with the largest AIM ($\chi^2 = 0.15$, $p = 0.703$) or the largest additive variance ($\chi^2 =$

338　2.65, $p = 0.104$, respectively), and was significantly worse with other subsets ($\chi^2 > 6.66$, $p <$

339   0.010). For tree height, no subset of SNPs improved the likelihood (χ^2 < 3.48, p > 0.062). In

340   other words, when focal individuals had related individuals in their vicinity, the focal

341   phenotypes were more affected: in this case, the vicinity of related individuals caused

342   significantly more rust damage.

343

344   The results suggest that shared vulnerability is the cause of rust damage amplification in the

345   related neighborhood. In support of this claim, the covariance between random pairs of

346   neighboring individuals' rust vulnerability was enriched by local relatedness: from an average

347   of 0.28 for the whole set to 0.43 when only the most related neighbors (highest quartile)

348   were retained, suggesting that local relatedness was a proxy for phenotypic similarity for the

349   trait rather than the cause of plant-to-plant interactions (through kin selection or niche

350   partitioning).

351

352   In spite of genomic models being sensitive to local relatedness, improvement of predictive

353   abilities (PA) was limited (Fig. S3) - probably due to a lack of power. Indeed, incorporating

354   local relatedness did not significantly change the PA neither with the whole set of SNPs ($|t_2|$

355   < 0.594, p > 0.553), nor with any of the alternative subsets ($|t_2|$ < 1.03, p > 0.301), with a

356   slightly better PA when incorporating local relatedness for SNPs with the smallest MAF on

357   rust vulnerability (t_2 = 1.73, df = 2.00 x 10^3, p = 0.0841), and SNPs with the smallest tree

358   height significance on rust vulnerability (t_2 = 1.97, df = 2.00 x 10^3, p = 0.0491). See the

359   Supplementary materials for further details.

360

361

362

363

364

365

366

367

368

369

370

Discussion

Genomic selection in forest tree breeding has successfully predicted growth and wood quality for many species (Lenz et al., 2020), such as Eucalyptus (e.g. Resende et al., 2012), pines (e.g., Isik et al., 2016), or spruces (e.g., Beaulieu et al., 2014). The predictive ability (PA) of the underlying genomic models relies on the ability of genotypic data to capture the covariance between candidate loci, thereby reflecting how much genetic similarity explains phenotypic similarities (Powell et al., 2010). An additional source of phenotypic covariances are micro-environmental effects, and accounting for spatial autocorrelation can enhance genomic predictions (Cappa et al, 2022). In our case, on top of spatial autocorrelation, we accounted for local relatedness, another source of heterogeneity. Results show that part of the plant-to-plant interaction can be ascribed to the relatedness with neighbors. In the case of the Black poplars under study, for rust vulnerability the interaction was more likely due to relatedness being a proxy rather than revealing kin selection/niche partitioning. Additionally, relatedness increases rust damages, whether focal individuals and neighbors are resistant or not, suggesting that inoculum pressure partly increases with relatedness. The lack of diversity at the local scale can hence be detrimental, whatever the level of resistance of individuals.

Differences in genetic architecture driving local relatedness

It is generally assumed that tree height is a complex trait with a highly polygenic architecture, which might not be the case for rust vulnerability (Du et al. 2016; Jorge et al., 2005). The composition of our SNP array may also have had an influence on the way both architectures were '*captured*', as the chip was enriched during its development in QTLs and candidate expression genes related to rust resistance (Faivre-Rampant et al., 2016), but not for height, the most polygenic of the two traits. The low genetic correlation between both traits is consistent with some previous reports (e.g., Beaulieu et al., 2020; Liu et al., 2022), though

396   not all (e.g., Lenz et al., 2021; Cappa et al., 2022); we will assume in our case that

397   orthogonality is strong enough to show that the traits are non-redundant, with independent

398   architectures.

399

400   The differences in genetic architecture between the two traits studied, as well as the design

401   of the SNP array, may have determined the results in terms of association, with only one

402   major gene detected (as a peak) for rust, in contrast to the absence of associations for the

403   more polygenic trait. Another element that may have influenced the association for rust is its

404   phenotypic scoring, which is based on a qualitative scale of nine notations. Even if it should

405   be taken cautiously, it is not uncommon to encounter a simpler genetic architecture for

406   response to biotic constraints compared to complex quantitative traits (e.g., gene-for-gene

407   resistance; Flor, 1971; Jorge et al., 2005). Consequently, many haplotype combinations can

408   lead to tall individuals, but few to rust resistance, because a polygenic architecture is

409   redundant; in other words, when individuals are related, their probability of both being

410   resistant is higher than that of both being tall. Local relatedness thus reveals a shared

411   vulnerability within the neighborhood.

412

Selection of relatedness or co-selection of relatives?

413

414   Interactions between neighbors may depend on local relatedness, and whether this

415   interaction is positive or negative depends on the relative strength of kin selection (Hamilton,

416   1964) and niche partitioning (Silvertown, 2004). As mentioned above, the importance of

417   group selection in nature is still debated, but local relatedness has been shown to be

418   perceived through kin recognition in plant species (e.g., Dudley et al., 2007; Murphy et al.,

419   2009). Such recognition could then lead to favorable (kin selection) or unfavorable (niche

420   partitioning) interactions. One such detrimental interaction is competition. Competition can

421   be accounted for in a quantitative genetic model by assigning to an individual one direct

422   genetic effect (DGE) as the effect of its genes on its own phenotype, and one indirect

423 genetic effect (IGE) as the effect of its genes on phenotypes of surrounding individuals

424 (Griffing, 1967; potential applications in Muir, 2005). Competition occurs when DGE and IGE

425 are negatively correlated. A positive correlation from such a model would suggest a

426 cooperative synergistic interaction. Although adding complexity to the model, including

427 competition has been shown to reduce bias in covariance estimation (Costa e Silva et al.,

428 2013). In our case, in addition to attempts to include interactions, IGE resulted in a lower

429 predictive ability than with the ssGBLUP baseline (results not shown).

430

431 The simpler genetic architecture of rust vulnerability compared to that of tree height might

432 play a role in shaping the phenotypes. Family structure was driving similarity of rust

433 vulnerability, and combination of similar vulnerability resulted in increased rust damage. The

434 amplifying effect is even stronger in perennials as shared vulnerability can be perceived by

435 successive generations of pathogens. Consistent with the increased disease resilience

436 conferred by varietal association in a field (e.g., Smithson and Lenne, 1996; Burdon and

437 Thrall, 2009), the local aggregation of susceptible individuals amplified rust damage.

438 Resistance similarity and relatedness were strongly linked, but it is likely that the more

439 complex the genetic architecture, the easier it is to dissociate phenotypic similarity and

440 relatedness.

441

442 Once it has been shown that neighbourhood relatedness is indeed correlated with similarity

443 at focal phenotypes, the difficulty in disentangling kin selection from niche partitioning is to

444 determine whether the relatedness is a cause, as an expression of kin selection or niche

445 competition, or a consequence, as an expression of a confounding between relatedness and

446 phenotypic interaction. In the latter case, the focal phenotype would be modified by that of its

447 neighbors through plant-to-plant interactions, and if phenotypic interaction increases when

448 phenotypes are similar (i.e., directional selection, e.g., competition for resource uptake), it

449 can be confused with the effect of relatedness, as genetically related individuals tend to have

450 similar phenotypes.

451

452   Whether there was kin selection/competition in addition to the phenotypic interaction, or

453   whether we simply lacked statistical power, is difficult to assess. Selection on local

454   relatedness due to kin selection or niche partitioning - selection on relatedness - might be

455   confused with the fact that selection leads to phenotypic convergence, that in turn leads to a

456   genotypic convergence - a co-selection of related. The randomization in the block design

457   and repetition can partly limit this confusion, but it is evident that interpreting the data is

458   fraught with difficulties, as the tallest individuals showed also a higher relatedness with the

459   rest of the population. The fact that family structure was mainly carried by SNPs likely

460   associated with selection (following the hypothesis of Yang et al., 2010 and Biddanda et al.,

461   2020) suggest that relatives were co-selected, as is often the case in breeding programs

462   implementing family selection. An inevitable consequence is that, as seen in crop selection,

463   interesting genotypes for kin selection/cooperation have been lost during domestication

464   (Fréville et al., 2022), making it difficult for genomic selection to account for local relatedness

465   (as predictive abilities show).

466

467

468 Acknowledgements

476 Author Contributions

477 MT was responsible for writing the report, conducting the search, extracting and analysing

478 data, interpreting results, updating reference lists. LS was responsible for designing the

479 experiments and screening potentially eligible studies. ML and LS contributed to interpreting

480 results and provided feedback on the report.

481 Competing Interests

482 The authors declare that there is no conflict of interest.

483 Data Archiving

484 The data that support the findings of this study are openly available on DATA INRAE at

485 https://data.inrae.fr/privateurl.xhtml?token=b79ab1ca-ebb9-47c6-9272-1568c0c33d70.

486

Rerefences

Aguilar, I., Legarra, A., Cardoso, F., Masuda, Y., Lourenco, D., & Misztal, I. (2019). Frequentist p-values for large-scale-single step genome-wide association, with an application to birth weight in American Angus cattle. *Genetics Selection Evolution*, *51*(1), 1-8.

Astle, W., & Balding, D. J. (2009). Population structure and cryptic relatedness in genetic association studies. *Statistical Science*, *24*(4), 451-471.

Beaulieu, J., Doerksen, T., Clément, S., MacKay, J., & Bousquet, J. (2014). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity*, *113*(4), 343-352.

Beaulieu, J., Nadeau, S., Ding, C., Celedon, J. M., Azaiez, A., Ritland, C., et al. (2020). Genomic selection for resistance to spruce budworm in white spruce and relationships with growth and wood quality traits. *Evolutionary applications*, *13*(10), 2704-2722.

Biddanda, A., Rice, D. P., & Novembre, J. (2020). A variant-centric perspective on geographic patterns of human allele frequency variation. *Elife*, *9*, e60107.

Bijma, P., & Wade, M. J. (2008). The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *Journal of evolutionary biology*, *21*(5), 1175-1188.

512   Burdon, J. J., & Thrall, P. H. (2009). Coevolution of plants and their pathogens in natural

513   habitats. *Science*, *324*(5928), 755-756.

514

515   Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: from

516   polygenic to omnigenic. *Cell*, *169*(7), 1177-1186.

517

518   Cahill Jr, J. F., & McNickle, G. G. (2011). The behavioral ecology of nutrient foraging by

519   plants. *Annual Review of ecology, evolution, and systematics*, *42*, 289-311.

520

521   Calus, M. P., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using

522   different methods. *Genetics Selection Evolution*, *43*(1), 1-14.

523

524   Cappa, E. P., & Cantet, R. J. (2007). Bayesian estimation of a surface to account for a

525   spatial trend using penalized splines in an individual-tree mixed model. *Canadian Journal of*

526   *Forest Research*, *37*(12), 2677-2688.

527

528   Cappa, E. P., de Lima, B. M., da Silva-Junior, O. B., Garcia, C. C., Mansfield, S. D., &

529   Grattapaglia, D. (2019). Improving genomic prediction of growth and wood traits in

530   Eucalyptus using phenotypes from non-genotyped trees by single-step GBLUP. *Plant*

531   *Science*, *284*, 9-15.

532

533   Cappa, E. P., Ratcliffe, B., Chen, C., Thomas, B. R., Liu, Y., Klutsch, J., et al. (2022).

534   Improving lodgepole pine genomic evaluation using spatial correlation structure and SNP

535   selection with single-step GBLUP. *Heredity*, *128*(4), 209-224.

536

537   Casper, B. B., & Jackson, R. B. (1997). Plant competition underground. *Annual review of*

538   *ecology and systematics*, 545-570.

539

540 Christensen, O. F., & Lund, M. S. 2010. Genomic prediction when some animals are

541 not genotyped. Genetics Selection Evolution 42, 2.

542

543 Christensen, O. F., Madsen, P., Nielsen, B., Ostersen, T., & Su, G. (2012). Single-step

544 methods for genomic evaluation in pigs. *animal*, *6*(10), 1565-1571.

545

546 Costa e Silva, J., Kerr, R.J. (2013) Accounting for competition in genetic analysis, with

547 particular emphasis on forest genetic trials. *Tree Genetics & Genomes, 9,* 1-17.

548

549 Denison, R. F., Kiers, E. T., & West, S. A. (2003). Darwinian agriculture: when can humans

550 find solutions beyond the reach of natural selection? *The quarterly review of biology*, *78*(2),

551 145-168.

552

553 Diouf, I., Derivot, L., Koussevitzky, S., Carretero, Y., Bitton, F., Moreau, L., & Causse, M.

554 (2020). Genetic basis of phenotypic plasticity and genotype× environment interactions in a

555 multi-parental tomato population. Journal of Experimental Botany, 71(18), 5365-5376

556

557 Donald, C.M. (1981). Competitive plants, communal plants, and yields in wheat crops.

558 L. Evans, W.J. Peacock (Eds.), Wheat Science – Today and Tomorrow, Cambridge

559 University Press.

560

561 Donald, C.M., Hamblin, J. (1983). The convergent evolution of annual seed crops in

562 agriculture. In Advances in agronomy (ed. NC Brady), pp. 97–143. New York, NY:

563 Academic Press.

564

565 Dong, L., Xiao, S., Chen, J., Wan, L., & Wang, Z. (2016). Genomic selection using extreme

566 phenotypes and pre-selection of SNPs in large yellow croaker (Larimichthys crocea). *Marine*

567 *Biotechnology*, *18*(5), 575-583.

568

569   Du, Q., Gong, C., Wang, Q., Zhou, D., Yang, H., Pan, W., et al. (2016). Genetic architecture

570   of growth traits in Populus revealed by integrated quantitative trait locus (QTL) analysis and

571   association studies. *New Phytologist*, *209*(3), 1067-1082.

572

573   Fisher, R.A. (1918). The correlation between relatives under the supposition of Mendelian

574   inheritance. *Trans. Roy. Soc. Edinburgh 52*, 399-433.

575

576   Gualdrón Duarte, J. L., Gori, A. S., Hubin, X., Lourenco, D., Charlier, C., Misztal, I., & Druet,

577   T. (2020). Performances of Adaptive MultiBLUP, Bayesian regressions, and weighted-

578   GBLUP approaches for genomic predictions in Belgian Blue beef cattle. *BMC genomics*,

579   *21*(1), 1-18.

580

581   Dudley, S. A., & File, A. L. (2007). Kin recognition in an annual plant. *Biology Letters*, *3*(4),

582   435-438.

583

584   Emik, L. O., & Terrill, C. E. (1949). Systematic procedures for calculating inbreeding

585   coefficients. *Journal of Heredity*, *40*(2), 51-55.

586

587   Faivre-Rampant, P., Zaina, G., Jorge, V., Giacomello, S., Segura, V., Scalabrin, S., et al.

588   (2016). New resources for genetic studies in Populus nigra: Genome-wide SNP discovery

589   and development of a 12k Infinium array. *Molecular ecology resources*, *16*(4), 1023-1036.

590

591   Falconer, D.S. (1989). Introduction to quantitative genetics (3rd edition, p. 129). Pearson

592   Education India.

593

594   Feau, N., Joly, D. L., & Hamelin, R. C. (2007). Poplar leaf rusts: model pathogens for a

595   model tree. *Botany*, *85*(12), 1127-1135.

596

597 File, A. L., Murphy, G. P., & Dudley, S. A. (2012). Fitness consequences of plants growing

598 with siblings: reconciling kin selection, niche partitioning and competitive ability. *Proceedings*

599 *of the Royal Society B: Biological Sciences*, *279*(1727), 209-218.

600

601 Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual review of*

602 *phytopathology*, *9*(1), 275-296.

603

604 Fréville, H., Montazeaud, G., Forst, E., David, J., Papa, R., & Tenaillon, M. I. (2022). Shift in

605 beneficial interactions during crop evolution. *Evolutionary Applications*, *15*(6), 905-918.

606

607 Garcia-Baccino, C. A., Legarra, A., Christensen, O. F., Misztal, I., Pocrnic, I., Vitezica, Z. G.,

608 & Cantet, R. J. (2017). Metafounders are related to F st fixation indices and reduce bias in

609 single-step genomic evaluations. *Genetics Selection Evolution*, *49*(1), 1-14.

610

611 Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J., et al. (2022). refund:

612 Regression with Functional Data. R package version 0.1-26.

613 https://CRAN.R-project.org/package=refund

614

615 Griffing, B. (1967). Selection in reference to biological groups I. Individual and group

616 selection applied to populations of unordered groups. *Australian Journal of Biological*

617 *Sciences*, *20*(1), 127-140.

618

619 Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of theoretical*

620 *biology*, *7*(1), 17-52.

621

622 Isik, F., Bartholomé, J., Farjat, A., Chancerel, E., Raffin, A., Sanchez, L., et al. (2016).

623 Genomic selection in maritime pine. *Plant Science*, *242*, 108-119.

624

625 Jorge, V., Dowkiw, A., Faivre-Rampant, P., & Bastien, C. (2005). Genetic architecture of

626 qualitative and quantitative *Melampsora larici-populina* leaf rust resistance in hybrid poplar:

627 genetic mapping and QTL detection. *New Phytologist*, *167*(1), 113-127.

628

629 Kijas, J., Elliot, N., Kube, P., Evans, B., Botwright, N., King, H., et al. (2017). Diversity and

630 linkage disequilibrium in farmed Tasmanian Atlantic salmon. *Animal genetics*, *48*(2), 237-

631 241.

632

633 Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and

634 genomic information. *Journal of dairy science*, *92*(9), 4656-4663.

635

636 Legarra, A., Christensen, O. F., Vitezica, Z. G., Aguilar, I., & Misztal, I. (2015). Ancestral

637 relationships using metafounders: finite ancestral populations and across population

638 relationships. *Genetics*, *200*(2), 455-468.

639

640 Legionnet, A., Muranty, H., & Lefèvre, F. (1999). Genetic variation of the riparian pioneer

641 tree species Populus nigra. II. Variation in susceptibility to the foliar rust Melampsora larici-

642 populina. *Heredity*, *82*(3), 318-327.

643

644 Lenz, P. R., Nadeau, S., Mottet, M. J., Perron, M., Isabel, N., Beaulieu, J., & Bousquet, J.

645 (2020). Multi-trait genomic selection for weevil resistance, growth, and wood quality in

646 Norway spruce. *Evolutionary applications*, *13*(1), 76-94.

647

648 Lesaffre, T., & Billiard, S. (2021). On Deleterious Mutations in Perennials: Inbreeding

649 Depression, Mutation Load, and Life-History Evolution. *The American Naturalist*, *197*(5),

650 E143-E155.

651

652 Lewontin, R. C. (1970). The units of selection. *Annual review of ecology and systematics*, 1-

653 18.

654

655 Li, B., Zhang, N., Wang, Y. G., George, A. W., Reverter, A., & Li, Y. (2018). Genomic

656 prediction of breeding values using a subset of SNPs identified by three machine learning

657 methods. *Frontiers in genetics*, *9*, 237.

658

659 Liu, Y., Erbilgin, N., Ratcliffe, B., Klutsch, J. G., Wei, X., Ullah, A., et al. (2022). Pest

660 defences under weak selection exert a limited influence on the evolution of height growth

661 and drought avoidance in marginal pine populations. *Proceedings of the Royal Society B*,

662 *289*(1982), 20221034.

663

664 Lu, S., Liu, Y., Yu, X., Li, Y., Yang, Y., Wei, M., et al. (2020). Prediction of genomic breeding

665 values based on pre-selected SNPs using ssGBLUP, WssGBLUP and BayesB for

666 Edwardsiellosis resistance in Japanese flounder. *Genetics Selection Evolution*, *52*(1), 1-10.

667

668 Luo, Z., Yu, Y., Xiang, J., & Li, F. (2021). Genomic selection using a subset of SNPs

669 identified by genome-wide association analysis for disease resistance traits in aquaculture

670 species. *Aquaculture*, *539*, 736620.

671

672 Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, *201*(4924), 1145-1147.

673

674 Meuwissen, T. H. E. (1997). Maximizing the response of selection with a predefined rate of

675 inbreeding. *Journal of animal science*, *75*(4), 934-940.

676

677 Meuwissen, T. H., Hayes, B. J., & Goddard, M. (2001). Prediction of total genetic value using

678 genome-wide dense marker maps. *genetics*, *157*(4), 1819-1829.

679

680   Milbau, A., Reheul, D., De Cauwer, B., & Nijs, I. (2007). Factors determining plant–

681   neighbour interactions on different spatial scales in young species-rich grassland

682   communities. *Ecological research*, *22*(2), 242-247.

683

684   Miller, D. L. (2022). porridge: Principal Co-Ordinate Ridge Regression. R package version

685   1.2.

686

687   Montazeaud, G., Rousset, F., Fort, F., Violle, C., Fréville, H., & Gandon, S. (2020). Farming

688   plant cooperation in crops. *Proceedings of the Royal Society B*, *287*(1919), 20191290.

689

690   Muir, W. M. (2005). Incorporation of competitive effects in forest tree or animal breeding

691   programs. *Genetics*, *170*(3), 1247-1259.

692

693   Muñoz, F., and Sanchez, L. (2020). breedR: Statistical Methods for Forest Genetic

694   Resources Analysts. R package version 0.12-5. https://github.com/famuvie/breedR

695

696   Murphy, G. P., & Dudley, S. A. (2009). Kin recognition: competition and cooperation in

697   Impatiens (Balsaminaceae). *American journal of botany*, *96*(11), 1990-1996.

698

699   Murphy, G. P., Swanton, C. J., Van Acker, R. C., & Dudley, S. A. (2017). Kin recognition,

700   multilevel selection and altruism in crop sustainability. *Journal of Ecology (Oxford)*, *105*(4),

701   930-934.

702

703   Okasha, S. (2020). Altruism, group selection and correlated interaction. *The British journal*

704   *for the philosophy of science*.

705

706   Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS*

707   *genetics*, *2*(12), e190.

708

709 Pégard, M., Rogier, O., Bérard, A., Faivre-Rampant, P., Paslier, M. C. L., Bastien, C., et al.

710 (2019). Sequence imputation from low density single nucleotide polymorphism panel in a

711 black poplar breeding population. *BMC genomics*, *20*(1), 1-16.

712

713 Pégard, M., Segura, V., Muñoz, F., Bastien, C., Jorge, V., & Sanchez, L. (2020). Favorable

714 conditions for genomic evaluation to outperform classical pedigree evaluation highlighted by

715 a proof-of-concept study in poplar. *Frontiers in plant science*, *11*, 581954.

716

717 Powell, J. E., Visscher, P. M., & Goddard, M. E. (2010). Reconciling the analysis of IBD and

718 IBS in complex trait studies. *Nature Reviews Genetics*, *11*(11), 800-805.

719

720 R Core Team (2022). R: A language and environment for statistical computing. R

721 Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/

722

723 Ratcliffe, B., El-Dien, O. G., Cappa, E. P., Porth, I., Klápště, J., Chen, C., & El-Kassaby, Y.

724 A. (2017). Single-step BLUP with varying genotyping effort in open-pollinated Picea glauca.

725 *G3: Genes, Genomes, Genetics*, *7*(3), 935-942.

726

727 Resende, M. D., Resende Jr, M. F., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A.,

728 Aguiar, A. M., et al. (2012). Genomic selection for growth and wood quality in Eucalyptus:

729 capturing the missing heritability and accelerating breeding for complex traits in forest trees.

730 *New Phytologist*, *194*(1), 116-128.

731

732 Rosenberg, N. A., Li, L. M., Ward, R., & Pritchard, J. K. (2003). Informativeness of genetic

733 markers for inference of ancestry. *The American Journal of Human Genetics*, *73*(6), 1402-

734 1422.

735

736 Sanchez, M. P., Govignon-Gion, A., Croiseau, P., Fritz, S., Hozé, C., Miranda, G., et al.

737 (2017). Within-breed and multi-breed GWAS on imputed whole-genome sequence variants

738 reveal candidate mutations affecting milk protein composition in dairy cattle. *Genetics*

739 *Selection Evolution*, *49*(1), 1-16.

740

741 Silvertown, J. (2004). Plant coexistence and the niche. *Trends in Ecology & evolution*,

742 *19*(11), 605-611.

743

744 Smithson, J. B., & Lenne, J. M. (1996). Varietal mixtures: a viable strategy for sustainable

745 productivity in subsistence agriculture. *Annals of applied biology*, *128*(1), 127-158.

746

747 Stephens, M. (2017). False discovery rates: a new deal. *Biostatistics*, *18*(2), 275-294.

748

749 Thistlethwaite, F. R., Gamal El-Dien, O., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., et al.

750 (2020). Linkage disequilibrium vs. pedigree: Genomic selection prediction accuracy in

751 conifer species. *PLoS One*, *15*(6), e0232201.

752

753 Tiret, M., & Milesi, P. Statistical Peak Detection for GWAS. Zenodo.

754 https://doi.org/10.5281/zenodo.5079556. Deposited 7 July 2021.

755

756 Tsai, H. Y., Matika, O., Edwards, S. M., Antolín–Sánchez, R., Hamilton, A., Guy, D. R., et al.

757 (2017). Genotype imputation to improve the cost-efficiency of genomic selection in farmed

758 Atlantic salmon. *G3: Genes, genomes, genetics*, *7*(4), 1377-1383.

759

760 VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of dairy*

761 *science*, *91*(11), 4414-4423.

762

763   Weiner, J., Andersen, S. B., Wille, W. K. M., Griepentrog, H. W., & Olsen, J. M. (2010).

764   Evolutionary Agroecology: the potential for cooperative, high density, weed-suppressing

765   cereals. *Evolutionary Applications*, *3*(5-6), 473-479.

766

767   Weiner, J., Du, Y. L., Zhang, C., Qin, X. L., & Li, F. M. (2017). Evolutionary agroecology:

768   individual fitness and population yield in wheat (Triticum aestivum).

769

770   Weiner, J. (2019). Looking in the wrong direction for higher-yielding crop genotypes. *Trends*

771   *in Plant Science*, *24*(10), 927-933.

772

773   Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology.

774   *The Quarterly review of biology*, *82*(4), 327-348.

775

776   Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al.

777   (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature*

778   *genetics*, *42*(7), 565-569.

779

780   Zeileis, A., & Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. R News

781   2(3), 7-10.

782

783   Zhang, Z., Liu, J., Ding, X., Bijma, P., de Koning, D. J., & Zhang, Q. (2010). Best linear

784   unbiased prediction of genomic breeding values using a trait-specific marker-derived

785   relationship matrix. *PloS one*, *5*(9), e12648.

786

787

Figures

Figure 1. GWAS on tree height and rust vulnerability in Populus nigra. **(A)**. Manhattan plot of the GWAS on tree height. The dots are the log significance of the markers. The solid horizontal line stands for the 5% significance level, and the dashed horizontal line stands for the 5% FDR level. The vertical dashed lines are position where significant hits were detected as being part of a peak. **(B)**. Idem as panel A, for rust vulnerability. **(C)**. Q-Q plot of the GWAS on tree height. The dots were the quantile of the markers; the solid line is the 1-slope segment; the gray cross is the position of the median; and the gray area is the 95% confidence interval. **(D)**. Idem as panel C, for rust vulnerability.

Figure 2. The first two axes of the Multidimensional Scaling (MDS) of Populus nigra. **(A)**. MDS with labels of the fathers (with black large shapes). Colored and smaller shapes are the offspring, the shape corresponding to its father. **(B)**. The same MDS as in panel A., but labeled for the mothers.

Figure 3. Within family variance (V_F) for each subset of SNPs. The black dashed horizontal line is the average V_F of randomly sampled SNPs, the gray dashed horizontal lines delimit the 95% confidence interval, and the black horizontal dotted line is V_F of the entire set of SNPs. Red dots are when selected for tree height, and blue for rust vulnerability. SNP subsets: largest or smallest MAF (+ or - MAF), largest or smallest AIM (+ or - AIM), most or least significant (+ or - $\log_{10}(p)$), and most or least heritable for tree height (+ or - h^2).

A

SAN-GIORGIO

SRZ

BDG

71077-308

72131-017

71034-2-406

72131-036

3824-3

92510-1

72156-003

H-487

72144-009

72159-004

72145-007

72146-11

73182-009

73193-056

B

662200216

VGN-CZB25

662200037

SSC

71072-501

71040

H480

72149-029

72131-001

71041-3-402

73193-089

73193-091

MDS1 (34.95%)

MDS2 (24.45%)