



# B4EST

## Adaptive BREEDING for productive, sustainable and resilient FORESTs under climate change

### *Deliverable D2.3*

## Novel selection methods combining individual and group performance

**Planned delivery date (as in DoA):** M49

**Actual submission date:** 30/09/2022 - M53

**Workpackage:** WP2

**Workpackage leader:** INIA-CSIC

**Deliverable leader:** INRAE

**Version:** 1.0

<b>Project co-funded by the European Commission within the Horizon 2020 Programme (2014 - 2020)</b>	
<b>Dissemination Level</b>	
<b>PU</b> Public	<b>PU</b>
<b>CI</b> Classified, as referred to Commission Decision 2001/844/EC	
<b>CO</b> Confidential, only for members of the consortium (including the Commission Services)	

Research and Innovation action: GA no. 773383

Start date of the project: May 1<sup>st</sup>, 2018



## TABLE OF CONTENTS

1	Summary.....	3
2	Introduction.....	3
3	Theoretical background.....	4
4	Case studies.....	4
5	Results .....	5
6	General Discussion .....	7
7	Conclusions.....	7
8	Dissemination.....	8
9	Partners involved in the work .....	8
10	References.....	9



## 1 Summary

Breeding is mainly about redirecting evolution towards new equilibria between traits that often decrease fitness in favor of economically interesting traits, or towards paths that were insufficiently explored in nature (Weiner, 2019), such as group selection. **Group selection**, when implemented artificially, consist in giving breeding advantages to a group of candidates because their performance as a group (productivity, tolerance to biotic or abiotic stress, or long-term response to selection) is higher than when selecting candidates by their individual performances (Wade et al, 2010). Increasing group yields through group selection rather than optimizing individual performance would have benefits on breeding. To that end, we explored two ways of incorporating group selection in a black poplar breeding scheme: (i) balancing selection with diversity for minimizing a long-term inbreeding depression; and (ii) measuring and accounting for how individuals react locally to their kin. We used for the two projects the dataset from the French breeding program of *Populus nigra* provided by INRAE Orléans (10,301 individuals, 1,452 different genotypes).

Regarding our first aim, with the use of computer simulations to show that the genetic gain will reach a higher selection plateau in the long run when reducing the accumulation of inbreeding than when the latter isn't controlled for. Breeding schemes focusing on individual breeding values will reach a lower selection plateau quickly. The best trade-off between genetic gain and inbreeding has been found to be obtained when mating an individual with high heterozygosity with an individual with high homozygosity.

Regarding our second aim, the benefit for genomic selection of accounting for plasticity to the presence of kin (local relatedness) was less obvious. According to genomic evaluations using a single step Genomic BLUP (ssGBLUP) approach, predictive ability was not significantly different when including local relatedness as a covariate to the ssGBLUP. Despite the fact that the presence of kin does not influence phenotypes enough to alter predictions, AIC of the models were slightly better (even when re-performing the analyses on a subset of SNP): this would suggest (i) a lack of power for prediction, or (ii) sensitivity to the presence of kin is either absent or not varying among individuals. This aspect warrants further studies.

## 2 Introduction

In breeding schemes, the phenotypes of planted individuals are the result of complex interactions between genetics, abiotic factors and biotic factors. In crop cultivation and forestry, practices aim to minimize biotic interactions and interspecific competition/parasitism: pesticides, fungicides, herbicides or weeding are commonplace. These practices have the secondary effect of drastically increasing intraspecific interactions, which become then the dominant biotic factor. In other words, due to agricultural and forestry practices, the main competitors of the selected and planted individuals are no longer the pests, but the *other* selected and planted individuals – at least in the vicinity. What are the



interactions between neighboring individuals, and how can a breeding scheme benefit from this knowledge?

### 3 Theoretical background

Two evolutionary forces should be in opposition for intraspecific local biotic interactions: niche competition that is strongest when neighboring individuals are genetically close, and kin selection that should attenuate competition between related genotypes. In the specific case of plants, competition occurs with the nearest plants (Casper et al., 1997; Milbau et al., 2007; File et al., 2011). The intensity of competition among neighbors may depend on the relatedness among neighbors, hereafter referred to as local relatedness, either through kin selection, which predicts that relatives may cooperate (Hamilton, 1964), or, on the opposite, niche partitioning, which predicts that similar individuals may compete for similar resources (Silvertown, 2004). There are therefore reasons to expect local relatedness to impact phenotypes in one way or the other (Cahill et al., 2011), so understanding plasticity to the presence of related neighbors might provide valuable information for the implementation of group selection in plant breeding. Although kin recognition was reported to be not particularly relevant in crop species because of their already high relatedness (e.g., Murphy et al., 2017), we might be able nonetheless to detect plasticity to local relatedness in breeding schemes that have been less intensive, such as forest tree breeding.

In order to assess the extent to which local relatedness affected breeding in the context of genomic selection as defined by Meuwissen et al. (2001), we chose to conduct our analyses on the dataset from the French breeding program of *Populus nigra* L. (Pégard et al., 2020). Black poplar is a Eurasian riparian forest tree that contributes, as a parent, along with *Populus deltoides* Bartr. ex Marsh., to one of the most widely used hybrid trees in the wood industry (*Populus x canadensis*). Growth of Black poplar is particularly susceptible to foliar rust, a fungus that causes a reduction in photosynthesis efficiency, and might lead to the death of susceptible clones (Legionnet et al., 1999). We therefore focused our analyses on tree height and rust vulnerability of one-year-old trees. Our main objective was to assess whether estimating breeding values with a multitrait ssGBLUP was significantly outperformed when incorporating local relatedness as a covariate (representing the environment). Estimating breeding values with ssGBLUP is a solid baseline in forest trees as recently shown (e.g., Cappa et al., 2019; Ratcliffe et al., 2017), especially in its multitrait version which is known to give higher accuracy (Calus et al., 2011).

### 4 Case studies

The study was based on previously published data (Pégard et al., 2019, 2020). Seventeen males and 17 females *Populus nigra* were sampled in France, from natural populations and breeding programs. The sampled individuals were crossed, and clones of both parents and progeny were planted in four experimental trials at the same location (Guéméné-Penfao, France, 47°37'59"N, 1°49'59"W), each with

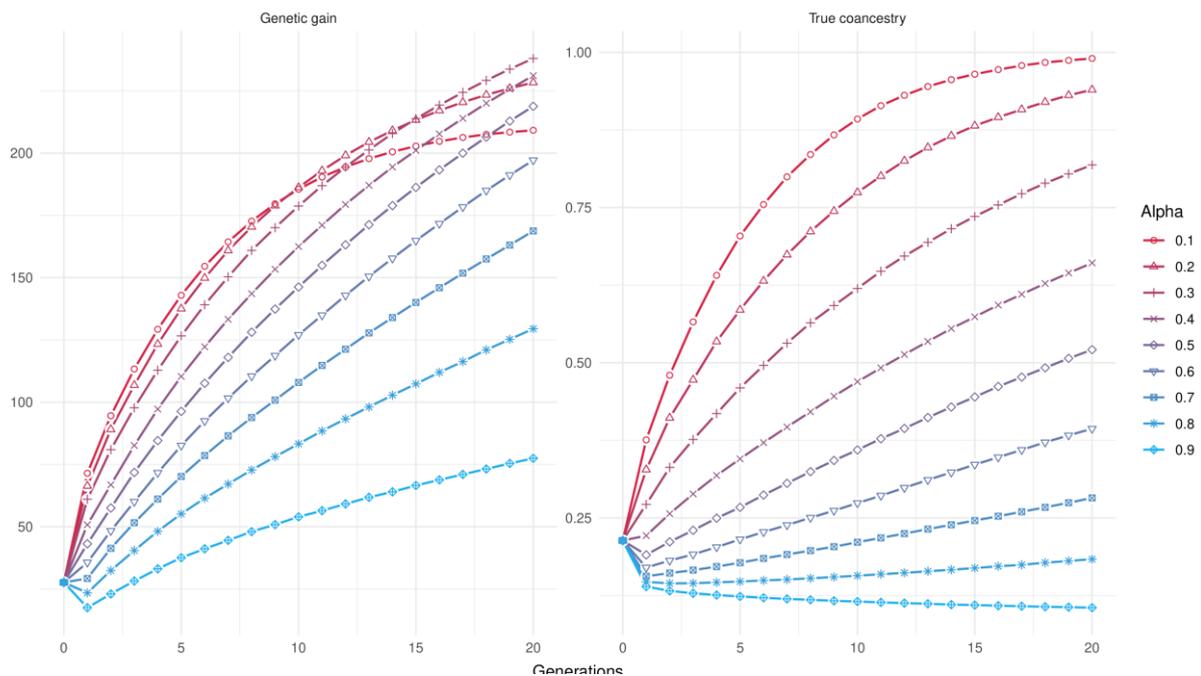


a random block design. In total, the experiment comprised 10,301 trees (including 7,169 genotyped trees), 42 full-sib families, 5 half-sib families, 1,452 genotypes (each replicated on average 7.09 times), and an Unknown Parent Group (UPG) of 105 genotypes (Table S1). Family size ranged from 1 to 119, with an average of 30.2 genotypes per family. Twenty genotypes and 17 families were shared across the trials.

## 5 Results

**First project:** balancing genetic gain and genetic diversity is an important part of the long-term breeding strategy, in the sense that the immediate gain might be low, but in the long term the genetic gain might be more advantageous than when not balancing (Fig. 1).

In these simulations, the genetic gain is penalized by the loss of diversity through mate optimization.



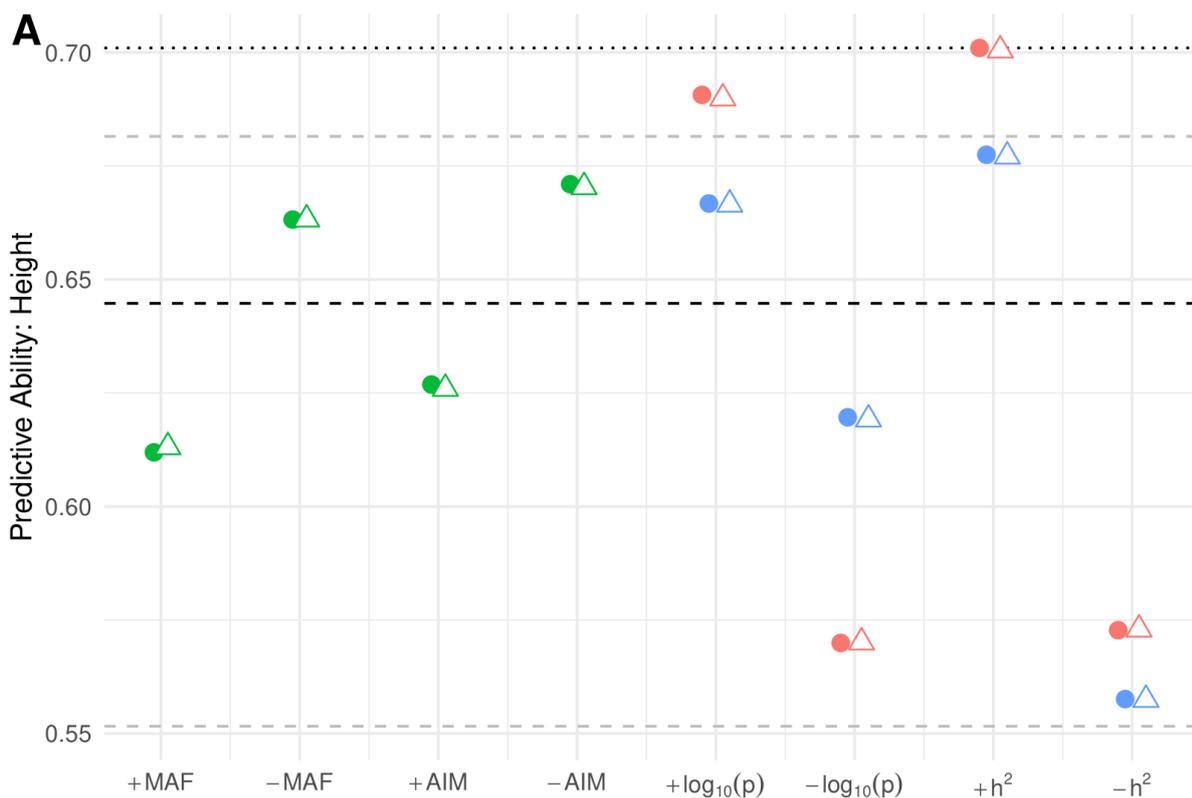
*Figure 1: Evolution over time of genetic gain (left panel) and coancestry (right panel), for different values of the trade-off  $\alpha$  (red for  $\alpha = 0.1$  favoring genetic gain till blue for  $\alpha = 0.9$  favoring genetic diversity).*

Classic breeding maximizes individual breeding values, but when deviating from such individual-based selection and accounting for population-wise measures – in other words, implementing a group selection based on group relatedness, genetic gain can reach a higher selection plateau while preserving genetic diversity – hence the ability to respond to selection. The ability to respond to selection was even more pronounced when the procedure favored mating between individuals sharing complementary loci (i.e., on with a homozygote stage and the other with a heterozygote stage), thus forming the perfect balance between gain and diversity.



**Second project:** In a first exploratory step, we assessed the advantage for breeding of optimizing for interactions between planted individuals. As a simplification, we tried to detect the effect of local relatedness as a proxy of the strength of kin selection/niche competition. For that, we compared the values of predictive abilities (PA) of different GBLUP models, some based on particular sets of SNPs that maximize a quantitative genetic feature (log-significance in the GWAS, heritability, MAF, Ancestor Informativeness Coefficient). These subsets were compared to a random subset of SNPs of equal size. We used these subsets as a way to understand for what SNPs and what function local relatedness might be important for.

The different sets of SNPs led to different predictive abilities (PA), depending on the combination of traits and SNP subsets (Fig. 2). The highest PA was reached when the entire set of SNPs were used for both tree height and rust vulnerability. When selecting the most significant SNPs, PA reached a significantly higher level than that of the random subsets, or when selecting the least significant SNPs, and nearly reached the level of the entire set. For tree height, however, although PA with the most significant SNPs was significantly higher than that of the random subsets, or when selecting the least significant SNPs, it was significantly lower than that of the entire set.



*Figure 2: Predictive ability on tree height for different subset of SNPs: the highest or the lowest MAF ( $\pm$  MAF), the highest or the lowest Ancestor Informativeness Coefficient ( $\pm$  AIM), the highest or the lowest significance ( $\pm$   $\log_{10}(p)$ ): red for tree height, blue for rust vulnerability), and the highest and the lowest heritability ( $\pm$   $h^2$ : same color than for significance). Circles are classic GBLUP, triangles include the local relatedness.*



Although PA was sensitive to the subset of SNPs, it was not quite so to the incorporation of local relatedness in the model, since doing so did not significantly modify the PA with the entire set of SNPs, nor any of the subset. In addition, when estimating breeding values with indirect genetic effect, the model reached a PA of 68.21% that was significantly lower than the baseline. The lack of sensitivity to local relatedness in regards to PA can be explained by a lack of power, since, despite the PA not being significantly different, the Akaike Information Criteria (AIC) was always higher when including local relatedness as a covariate, but not significantly so.

## 6 General Discussion

If local relatedness is quantitatively associated with a phenotype, it can either be the expression of relatedness acting directly upon the considered phenotype (niche competition or kin selection), or it can be a spurious association. For instance, if there is co-selection of related individuals on tree height, relatedness can reflect past selection and therefore be associated with large values of height, the more so if heritability is high. Teasing apart these two effects is fraught with difficulties, and requires to change the way relatedness is computed compared to what is done in genomic selection.

The second project not only confirms this difficulty, but also pushes breeding programs towards a shift of paradigm: (mass) selecting based on individual breeding values – even when accounting for individual plasticity to neighbourhood – can be unsustainable (Weiner, 2019). The main reason is that response to a selection similar to natural selection (i.e., individual based) is low and can lead to evolutionary dead-ends. When such dead-ends are reached, individuals can no longer respond to abiotic changes, or biotic factors (e.g., intraspecific competition, or parasitism). This evolutionary aspect is especially important when selecting for resistant plants, as competition is an arm race – and neglecting it will only provide improved plants that are resistant for only a few generations. Deviating from individual based selection (i.e., group selection) is a way that is counterintuitive, difficult to implement and difficult to predict what population-based measure would bring the higher genetic gain, and yet very powerful. Implementing a group selection that maintains the ability to respond to selection seems to be the most valuable feature for a sustainable breeding, rather than accounting for interactions.

## 7 Conclusions

Implementing group selection would be fraught with difficulties if we try to breed for plants that are “naturally” inclined to cooperate with their neighbours (2nd project); such plants can hardly be found, as natural or artificial selection has likely selected competitive individuals. The most beneficial way for agriculture and forestry breeding to implement group selection would be to take advantage of information and predictions that are not available to “natural populations”, for instance, limiting genetic gain to reach a higher long-term genetic gain (1st project).



## 8 Dissemination

The results of this deliverable have been presented in the B4EST conferences (2019, 2021, 2022), EvolTree online conference (2020), and will be presented at the next PopGroup meeting (next session of 2023).

One paper reporting activities initiated in GENTREE project by the contributors of this deliverable served as seed to the work funded by B4EST project and helped design of simulation works specifically developed in this deliverable :

Tiret, M., Pégard, M., & Sánchez, L. (2021). How to achieve a higher selection plateau in forest tree breeding? Fostering heterozygote x homozygote relationships in optimal contribution selection in the case study of *Populus nigra*. *Evolutionary Applications*, 14, 2635– 2646. <https://doi.org/10.1111/eva.13300>

One paper will be submitted by the end of 2022:

Tiret, M., Sánchez, L., Lascoux, M. (work in progress). Local relatedness and genomic selection in *Populus nigra*.

## 9 Partners involved in the work

Martin Lascoux and Mathieu Tiret (UU), Véronique Jorge and Leopoldo Sanchez (INRAE)



## 10 .References

- 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., Nadeau, S., Ding, C., Celedon, J. M., Azaiez, A., Ritland, C., ... & Bousquet, J. (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.
- Cahill Jr, J. F., & McNickle, G. G. (2011). The behavioral ecology of nutrient foraging by plants. *Annual Review of ecology, evolution, and systematics*, 42, 289-311.
- Calus, M. P., & Veerkamp, R. F. (2011). Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution*, 43(1), 1-14.
- Cappa, E. P., de Lima, B. M., da Silva-Junior, O. B., Garcia, C. C., Mansfield, S. D., & Grattapaglia, D. (2019). Improving genomic prediction of growth and wood traits in Eucalyptus using phenotypes from non-genotyped trees by single-step GBLUP. *Plant Science*, 284, 9-15.
- Cappa, E. P., Ratcliffe, B., Chen, C., Thomas, B. R., Liu, Y., Klutsch, J., et al. (2022). Improving lodgepole pine genomic evaluation using spatial correlation structure and SNP selection with single-step GBLUP. *Heredity*, 128(4), 209-224.
- Casper, B. B., & Jackson, R. B. (1997). Plant competition underground. *Annual review of ecology and systematics*, 545-570.
- Christensen, O. F., & Lund, M. S. 2010. Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution* 42, 2.
- Christensen, O. F., Madsen, P., Nielsen, B., Ostersen, T., & Su, G. (2012). Single-step methods for genomic evaluation in pigs. *animal*, 6(10), 1565-1571.



Costa e Silva, J., Kerr, R.J. (2013) Accounting for competition in genetic analysis, with particular emphasis on forest genetic trials. *Tree Genetics & Genomes*, 9, 1-17.

Denison, R. F., Kiers, E. T., & West, S. A. (2003). Darwinian agriculture: when can humans find solutions beyond the reach of natural selection?. *The quarterly review of biology*, 78(2), 145-168.

Donald, C.M. (1981). Competitive plants, communal plants, and yields in wheat crops. L. Evans, W.J. Peacock (Eds.), *Wheat Science – Today and Tomorrow*, Cambridge University Press.

Donald, C.M., Hamblin, J. (1983). The convergent evolution of annual seed crops in agriculture. In *Advances in agronomy* (ed. NC Brady), pp. 97–143. New York, NY: Academic Press.

Gualdrón Duarte, J. L., Gori, A. S., Hubin, X., Lourenco, D., Charlier, C., Misztal, I., & Druet, T. (2020). Performances of Adaptive MultiBLUP, Bayesian regressions, and weighted-GBLUP approaches for genomic predictions in Belgian Blue beef cattle. *BMC genomics*, 21(1), 1-18.

Dudley, S. A., & File, A. L. (2007). Kin recognition in an annual plant. *Biology Letters*, 3(4), 435-438.

Emik, L. O., & Terrill, C. E. (1949). Systematic procedures for calculating inbreeding coefficients. *Journal of Heredity*, 40(2), 51-55.

Faivre- Rampant, P., Zaina, G., Jorge, V., Giacomello, S., Segura, V., Scalabrin, S., et al. (2016). New resources for genetic studies in *Populus nigra*: Genome- wide SNP discovery and development of a 12k Infinium array. *Molecular ecology resources*, 16(4), 1023-1036.

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

File, A. L., Murphy, G. P., & Dudley, S. A. (2012). Fitness consequences of plants growing with siblings: reconciling kin selection, niche partitioning and competitive ability. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), 209-218.

Garcia-Baccino, C. A., Legarra, A., Christensen, O. F., Misztal, I., Pocrnic, I., Vitezica, Z. G., & Cantet, R. J. (2017). Metafounders are related to  $F_{st}$  fixation indices and reduce bias in single-step genomic evaluations. *Genetics Selection Evolution*, 49(1), 1-14.

Griffing, B. (1967). Selection in reference to biological groups I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences*, 20(1), 127-140.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of theoretical biology*, 7(1), 17-52.



Kijas, J., Elliot, N., Kube, P., Evans, B., Botwright, N., King, H., et al. (2017). Diversity and linkage disequilibrium in farmed Tasmanian Atlantic salmon. *Animal genetics*, 48(2), 237-241.

Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of dairy science*, 92(9), 4656-4663.

Legarra, A., Christensen, O. F., Vitezica, Z. G., Aguilar, I., & Misztal, I. (2015). Ancestral relationships using metafounders: finite ancestral populations and across population relationships. *Genetics*, 200(2), 455-468.

Legionnet, A., Muranty, H., & Lefèvre, F. (1999). Genetic variation of the riparian pioneer tree species *Populus nigra*. II. Variation in susceptibility to the foliar rust *Melampsora larici-populina*. *Heredity*, 82(3), 318-327.

Lenz, P. R., Nadeau, S., Mottet, M. J., Perron, M., Isabel, N., Beaulieu, J., & Bousquet, J. (2020). Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evolutionary applications*, 13(1), 76-94.

Li, B., Zhang, N., Wang, Y. G., George, A. W., Reverter, A., & Li, Y. (2018). Genomic prediction of breeding values using a subset of SNPs identified by three machine learning methods. *Frontiers in genetics*, 9, 237.

Liu, Y., Erbilgin, N., Ratcliffe, B., Klutsch, J. G., Wei, X., Ullah, A., et al. (2022). Pest defences under weak selection exert a limited influence on the evolution of height growth and drought avoidance in marginal pine populations. *Proceedings of the Royal Society B*, 289(1982), 20221034.

Luo, Z., Yu, Y., Xiang, J., & Li, F. (2021). Genomic selection using a subset of SNPs identified by genome-wide association analysis for disease resistance traits in aquaculture species. *Aquaculture*, 539, 736620.

Meuwissen, T. H., Hayes, B. J., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *genetics*, 157(4), 1819-1829.

Milbau, A., Reheul, D., De Cauwer, B., & Nijs, I. (2007). Factors determining plant–neighbour interactions on different spatial scales in young species-rich grassland communities. *Ecological research*, 22(2), 242-247.

Montazeaud, G., Rousset, F., Fort, F., Violle, C., Fréville, H., & Gandon, S. (2020). Farming plant cooperation in crops. *Proceedings of the Royal Society B*, 287(1919), 20191290.



Muir, W. M. (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics*, 170(3), 1247-1259.

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

Murphy, G. P., & Dudley, S. A. (2009). Kin recognition: competition and cooperation in *Impatiens* (Balsaminaceae). *American journal of botany*, 96(11), 1990-1996.

Murphy, G. P., Swanton, C. J., Van Acker, R. C., & Dudley, S. A. (2017). Kin recognition, multilevel selection and altruism in crop sustainability. *Journal of Ecology (Oxford)*, 105(4), 930-934.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS genetics*, 2(12), e190.

Pégard, M., Rogier, O., Bérard, A., Faivre-Rampant, P., Paslier, M. C. L., Bastien, C., et al. (2019). Sequence imputation from low density single nucleotide polymorphism panel in a black poplar breeding population. *BMC genomics*, 20(1), 1-16.

Pégard, M., Segura, V., Muñoz, F., Bastien, C., Jorge, V., & Sanchez, L. (2020). Favorable conditions for genomic evaluation to outperform classical pedigree evaluation highlighted by a proof-of-concept study in poplar. *Frontiers in plant science*, 11, 581954.

Powell, J. E., Visscher, P. M., & Goddard, M. E. (2010). Reconciling the analysis of IBD and IBS in complex trait studies. *Nature Reviews Genetics*, 11(11), 800-805.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL. <https://www.R-project.org/>

Ratcliffe, B., El-Dien, O. G., Cappa, E. P., Porth, I., Klápště, J., Chen, C., & El-Kassaby, Y. A. (2017). Single-step BLUP with varying genotyping effort in open-pollinated *Picea glauca*. *G3: Genes, Genomes, Genetics*, 7(3), 935-942.

Rosenberg, N. A., Li, L. M., Ward, R., & Pritchard, J. K. (2003). Informativeness of genetic markers for inference of ancestry. *The American Journal of Human Genetics*, 73(6), 1402-1422.

Silvertown, J. (2004). Plant coexistence and the niche. *Trends in Ecology & evolution*, 19(11), 605-611.

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



Tiret, M., & Milesi, P. Statistical Peak Detection for GWAS. Zenodo. <https://doi.org/10.5281/zenodo.5079556>. Deposited 7 July 2021.

Tsai, H. Y., Matika, O., Edwards, S. M., Antolín-Sánchez, R., Hamilton, A., Guy, D. R., et al. (2017). Genotype imputation to improve the cost-efficiency of genomic selection in farmed Atlantic salmon. *G3: Genes, genomes, genetics*, 7(4), 1377-1383.

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of dairy science*, 91(11), 4414-4423.

Weiner, J., Andersen, S. B., Wille, W. K. M., Griepentrog, H. W., & Olsen, J. M. (2010). Evolutionary Agroecology: the potential for cooperative, high density, weed- suppressing cereals. *Evolutionary Applications*, 3(5- 6), 473-479.

Weiner, J., Du, Y. L., Zhang, C., Qin, X. L., & Li, F. M. (2017). Evolutionary agroecology: individual fitness and population yield in wheat (*Triticum aestivum*).

Weiner, J. (2019). Looking in the wrong direction for higher-yielding crop genotypes. *Trends in Plant Science*, 24(10), 927-933.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature genetics*, 42(7), 565-569.

Zhang, Z., Liu, J., Ding, X., Bijma, P., de Koning, D. J., & Zhang, Q. (2010). Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. *PLoS one*, 5(9), e12648.

Wade M. J., P. Bijma, E. D. Ellen, and W. Muir, 2010 Group selection and social evolution in domesticated animals. 13.