Paternity recovery in two maritime pine polycross mating designs and consequences for breeding

Marjorie Vidal, Christophe Plomion, Luc Harvengt, Annie A. Raffin, Christophe C. Boury, Laurent Bouffier

Paternity recovery in two maritime pine polycross mating designs and consequences for breeding

Marjorie VIDAL[1,2,3], Christophe PLOMION[1,2], Luc HARVENGT[3], Annie RAFFIN[1,2], Christophe BOURY[1,2], Laurent BOUFFIER[1,2§]

[1] **INRA**, BIOGECO, UMR 1202, F-33610 Cestas, France;

[2] **Univ. Bordeaux**, BIOGECO, UMR 1202, F-33400 Talence, France;

[3] **FCBA**, Biotechnology and Advanced Silviculture Dept., Genetics & Biotechnology team, F-33610 Cestas, France

[§]Corresponding author (laurent.bouffier@pierroton.inra.fr)

Abstract

Polycross mating systems are widely used in forest tree breeding for genetic testing. Backward selection based on polycross testing assumes equal male reproductive success and true half-sib progeny. The main objective of this study was firstly, to investigate the departure from these assumptions in a maritime pine polycross trial and, secondly, to evaluate the consequences for heritability and breeding values estimations.

A total of 984 offspring from 98 half-sib families was genotyped with single nucleotide polymorphism markers to recover the full pedigree. Paternity was assigned successfully for 89 % of the offspring at a 99% confidence level. We thus concluded there was a 11% pollen contamination rate, assuming contamination when no genotype from the polymix composition could be identified as a father. The paternal contribution to the offspring varied amongst the males, but the departure from half-sib assumption was moderate, since the average genetic correlation within the family was 0.26. Heritability and breeding values for girth at breast height and stem sweep were estimated using individual-tree mixed models with either partial or full pedigree information. The results highlighted a minor bias in heritability estimation due to unknown paternity, as well as a high correlation for estimated breeding values between the partial and full pedigree models, suggesting that the genetic merit of the parental generation for backward selection was adequately predicted using the partial pedigree model. Finally, pedigree recovery was also discussed in a perspective of forward selection.

Keywords: Pedigree reconstruction, SNP markers, paternal reproductive success, pollen contamination, breeding strategies, *Pinus pinaster* Ait.

Introduction

Tree breeding generally follows a classical recurrent selection scheme, which is characterized by repetitive cycles of breeding, testing and selection (White 1987; Zobel and Talbert 1984). In conifers, selection of best individuals for either further breeding or inclusion in a production population (seed orchard), is commonly performed based on additive genetic values since open-pollinated seed orchards are generally used for deployment. The breeding value of a genotype is defined as the sum of the average effects of its genes and is estimated here with the performances of its progeny. Thus, large numbers of polycross trials were established to assess breeding values and to rank female parents for backward selection (Burdon and Shelbourne 1971). Polycross mating is done by mixing pollen from several males and applying the resulting polymix to isolated female flowers. As the male parent of each individual offspring is unknown, estimates of genetic parameters and breeding values are based on the assumption that all progeny of a female are true half-sibs (Squillace 1974) and that an equal contribution of male parents takes place (Moriguchi et al. 2009). Departure from these assumptions could result in overestimation of both additive genetic variance and prediction of genetic gain, as relatedness would be underestimated.

Thanks to the availability of affordable and reliable DNA markers, coupled with powerful pedigree reconstruction methods, incomplete pedigree trials can be converted into complete pedigree trials. Most pedigree recovery studies have shown that paternal contributions in controlled polycross trials were unequal (Moriguchi et al. 2009); this has been confirmed for *Pseudotsuga menziesii* (Nakamura and Wheeler 1992), *Betula pendula* (Pasonen et al. 1999), *Picea abies* (Aronen et al. 2002), *Populus* spp. (Wheeler et al. 2006), *Pinus radiata* (Kumar et al. 2007) and *Picea rubens* Sarg, (Doerksen and Herbinger 2008). Equal paternal contribution was found in the progeny of a polycross (with nine pollen donors) in *Pinus taeda* (Wiselogel and Vanbuijtenen 1988). A differential reproductive success between males does not necessarily lead to a substantial increase of the genetic correlation amongst offspring within the same family (Doerksen and Herbinger 2010; Hansen and Nielsen 2010), or to an overestimation of additive genetic variance (Doerksen and Herbinger 2010; Hallingback and Jansson 2013). Recent studies have investigated the impact of unequal paternal contributions and/or selfing on genetic parameters and estimated breeding values (EBVs) by genotyping offspring originating from open-pollinated seed orchards (El-Kassaby et al. 2011; Gaspar et al. 2009; Hallingback and Jansson 2013; Hansen and McKinney 2010; Hansen and Nielsen 2010; Klapste et al. 2014; Korecky et al. 2013) or from polycross mating designs (Doerksen and Herbinger 2010; Kumar et al. 2007). These authors observed that individual heritability estimated without paternity information could be larger or smaller, in comparison with their respective estimates including paternity information.

In the French maritime pine (*Pinus pinaster* Ait.) breeding program, selection of the production population is based on polycross testing (backward selection) on the one hand, and on the other, selection of the following breeding generation is based on full-sib family testing (forward selection).

For this species, representing the most commonly planted forest tree in France, large genomic resources have been recently published including a unigene set (Canales et al. 2014) and a catalog of single nucleotide polymorphism (SNP) markers (Chancerel et al. 2013). These molecular markers constitute reliable tools to recover paternal identity in polycross trials. In this context, the main objective of this study was to investigate, for the first time in maritime pine, departure from assumptions of backward selection based on polycross trials. To this end, we firstly recovered paternity identities of 970 successfully genotyped offspring (out of 984 offspring sampled) obtained from two different polymixes (PMX) using SNP markers, we then assessed the pollen contamination rate (*i.e.* proportion of pollen donors not belonging to the PMX composition), studied the deviation from equal paternal contribution and tested whether the offspring of female parents were true half-sibs; finally we studied the bias in the estimation of heritability and breeding values due to unknown paternal identity.

Materials and Methods

Plant material and mating design

Six polycross trial series were established from 1994 to 2002 in South-western France to predict first-generation (G1) parental breeding values of the maritime pine breeding program. G1 trees are selected progeny from G0 trees, whereas G0 trees were mass-selected from the Landes provenance and constitute the base population of the breeding program (Illy 1966). Each series was composed of three polycross trials established on contrasted sites for water availability (humid, semi-humid and dry) with the same half-sib families. One trial of the third series was considered in this study (a humid site located 44°42'32"N / 0°46'8"W). It consisted of a 5 ha polycross trial established in 1996 with 174 half-sib (HS) families plus 5 checklots, planted in a randomized block design, totalizing 6,440 trees (35 complete blocks with one-tree plot per family and two trees per checklot in each block). Two different pollen mixes were used: 98 seed parents pollinated with Polymix A (PMX_A, 47 G1 pollen donors) and 76 seed parents pollinated with Polymix B (PMX_B, 43 G1 pollen donors). Among those G1 seed parents, 8 were pollinated with both PMX_A and PMX_B (families from identical seed parents but distinct PMX were considered as being different). There were no pollen donors in common between PMX_A and PMX_B.

Control-pollination was realized in April 1994 as described below. Isolated pollen cones were collected from the 47 (PMX_A) and 43 (PMX_B) pollen parents and processed following standard pollen extraction methods (pollen dried at 30°C with low relative humidity during 48-72 hours). Pollens were kept at 5°C until there were used. Both polymixes (PMX_A and PMX_B) consisted of equal volumes of their components. On each seed parent tree, seed cones were isolated with pollination bags approximatively 10-15 days before receptivity (scales still closed). The isolated cones were pollinated with one polymix three times during their maximum receptivity stages using a squeeze bulb syringe system to spray the pollen (holes in bags resulting from the pollen injection were plugged to prevent pollen contamination). Pollination bags were removed at the post-receptivity stage, generally 10-15 days after the last pollination and a durable identification tag was fixed. The cones resulting from the controlled pollinations were harvested 19 months later and then dried at 30°C during 3 months before extraction.

The offspring of both polycrosses (G2 trees) were phenotyped: tree girth at breast height was measured at the age of 12 years, and sweep (stem deviation from verticality at 1.5 m expressed in cm) was measured at the age of 8 years.

Sampling

Two successive samplings were carried out:

- i/ In spring 2012, young green needles of 489 G2 individuals (49 families x 9-10 individuals/family) from HS families obtained with PMX_A were collected (Subset_A).
- ii/ In spring 2013, young green needles of 495 G2 individuals (49 families x 9-12 individuals/family) from HS families obtained with PMX_B were collected (Subset_B).

A total of 984 G2 genotypes (15 % of the progeny trial) was thus sampled, representing 98 HS families with 9-12 randomly sampled trees per family. Moreover, needles of the putative parental genotypes (seed and pollen donors) were collected (considering two ramets from two different clonal archives when available). Subset_A came from 49 theoretical females and 47 males (PMX_A). However, one pollen donor belonging to PMX_A could not be genotyped (since it was missing in the clonal archives). Four genotypes were used both as males and females, so 91 parental genotypes were collected to analyze Subset_A. Subset_B came from 49 theoretical females and 43 males (PMX_B). As 10 genotypes were used as males and females, 82 parental genotypes were collected to analyze Subset_B.

DNA extraction and fingerprinting

Needles were kept in a cool box during sampling, and then stored at -80°C. DNA extractions were carried out on fine powder of needle tissues using an Invisorb® DNA Plant HTS 96 Kit (Stratec Molecular, Berlin, Germany), in compliance with the manufacturer's instructions and quantified with a NanoDrop microvolume spectrophotometer (Thermo Fisher Scientific Inc., Waltham, CA, USA).

Genotyping A: From the 12k Infinium SNP-array (Illumina, San Diego, USA) developed by Chancerel et al. (2013), 2,600 polymorphic SNPs were available from a previous screening of 661 G0 and G1 trees (Plomion et al. 2014). A first filter was used to retain 150 SNPs with minor allele frequency (MAF) greater than 0.45 to maximize the discrimination power of the marker. Assay Design Suite 1.0 Software (https://www.mysequenom.com/Tools) from Sequenom was finally used to automatically design two sets of iPLEX single base extension primers comprising 32 and 33 SNPs. Subset_A was genotyped with these 65 SNPs using the Sequenom MassARRAY iPLEX Gold assays (Sequenom, San Diego, CA, USA) in the genotyping and sequencing facility of Bordeaux, France (http://www.pgtb.u-bordeaux2.fr/).

Genotyping B: Subset_B was genotyped with an optimized set of SNPs markers. The most informative and easily read SNPs (39 out of 65) from the above mentioned genotyping assay were retained and completed with an additional set of 41 SNPs, considering not only the MAF (>0.45) but also the extent linkage disequilibrium, a criteria that was not taken into account in Subset_A. Pairwise linkage disequilibrium among candidate SNPs was calculated with LDcorSV package (Mangin et al. 2012) implemented in R (R Core Team (2013)) that corrects for relatedness, and only unlinked SNPs

(r_v^2<0.3*)* were kept. Thus, two new multiplexes of 40 SNPs each were designed to genotype Subset_B (cf. supplemental information for SNPs details).

Parentage assignment

Likelihood inference methodology was carried out with Cervus 3.0 (Kalinowski et al. 2007; Marshall et al. 1998), both to check the female (seed parent) identity and to recover the paternal (pollen donor within each PMX) identity of G2 trees. Cervus was run assuming a 0.1% genotyping error rate (estimation based on repeated genotyping of 50 G0 and 136 G1 individuals). G2 trees with less than 42 loci successfully genotyped in both subsets were excluded from the parentage analysis.

The female parent was confirmed when the LOD score was positive and only one mismatch allele was allowed for each offspring and its supposed female parent.

For paternity recovery, we considered that 90% of the pollen donors had been sampled (we assumed a 10% contamination rate, namely the presence of pollen donors exterior to PMXs). The delta score (*i.e.* the difference in LOD scores of the two most likely candidate parents) was used as a criterion for assignment of paternity at 99% level of confidence. Critical values of delta scores were based on simulations of 100,000 offspring. One mismatch allele was allowed between a given offspring and its parent. The Cervus selfing option was used, as some potential pollen donors were also seed parents.

The effective number of males $N_e m$ (expected and observed values) was calculated for each subset using the following formula (Kimura and Crow 1963):

$$N_e m = \frac{N_m\, k_m - 1}{k_m - 1 + \frac{V_m}{k_m}} \quad (1)$$

Where N_m is the number of males, k_m the mean number of progeny produced by males, and V_m the variance of the number of progeny per male.

Relatedness within and between families

Relatedness within and between families was analyzed using the additive genetic relationship coefficients from the additive genetic relationship matrix (A). The A matrix (or numerator relationship matrix) stems from the pedigree and contains diagonal elements equal to $1+F_i$, where F_i is the inbreeding coefficient of individual i. The off diagonal elements (A_{ij}) specify the relationship coefficient between individuals i and j ($A_{ij}=2\Theta_{ij}$, where Θ_{ij} is the coefficient of coancestry). The coefficient of coancestry Θ_{ij} describes the relatedness between pairs of individuals that is equivalent to the probability that two gametes taken at random (one from each) carry alleles that are identical by descent (e.g. for half-sibs $2\Theta_{ij}=0.25$, for full-sibs $2\Theta_{ij}=0.5$, when parents are non-inbred and non-related) (Lynch and Walsh 1998).

The A matrix was calculated with the synbreed package in R (R Development Core Team 2013), taking into account full pedigree of G1 parents (*i.e.* considering their G0 parents) and partial or full

pedigrees (depending on the model considered) of G2 progeny. The average relatedness of a family was estimated with the additive relationship coefficient, which is the mean of the off diagonal elements of the genotyped progeny of this family.

Variance-component estimation and BLUP breeding value prediction

Mixed-models were solved by restricted maximum likelihood implemented in ASReml version 3.0 (Gilmour et al. 2009) to estimate variance components and BLUP breeding values for girth at 12 years of age and sweep at 8 years of age.

The analyses were based on individual-tree linear mixed-models of the form given in Equation 2:

$y = Xb + Za + e$ (2)

Where y is the vector of individual tree observations on each trait, b is a vector of fixed-effect estimates (block effect), a is a vector of random additive genetic effects, and e is a vector of random residual effects. X and Z are incidence matrices for fixed and random model terms.

The additive and error effects were assumed uncorrelated with zero means and normally distributed as $a \sim N(0, A\sigma_a)$ and $e \sim N(0, I\sigma_e)$, where I is an identity matrix and A was the additive relationship matrix.

Narrow-sense heritability (h^2) was calculated for each trait as:

$$h^2 = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_e} \quad (3)$$

Where σ^2_a is the additive genetic variance and σ^2_e is the residual variance.

The general mixed model was fitted to the quantitative data of the successfully genotyped G2 trees (Subset_A and Subset_B together), including different pedigree files according to the model considered.

Here, the term "model" was used to differentiate the varying degrees of pedigree information (with and without paternity identity), which were fit to the same individual-tree model described in equation 2. In the partial pedigree (PP) model, only the theoretical seed donors of the families were known. In the full pedigree (FP) model, the pedigree file contained both parents' identities (from paternity recovery).

In addition, for both models, the pedigree files in ASReml always contained the identity of the grand-parents (G0), when available.

Best linear unbiased predictions, named estimated breeding values (EBVs), were estimated for the parental clones as well as for their progeny. Female breeding values from both models were compared by Pearson product-moment correlation and Spearman rank-order correlation to examine whether the complete pedigree information caused significant changes in EBVs as well as ranking that would affect selection and predicted gains.

The accuracy of predicted breeding values was calculated as:

$$r = \sqrt{1 - \frac{PEV}{(1+F)\sigma^2_a}} \quad (4)$$

Where PEV is the prediction error variance, F is the inbreeding coefficient and σ^2_a is the additive genetic variance (El-Kassaby et al. 2011).

Results

SNP genotyping

489 offspring (Subset_A) and their respective 91 G1 parents were genotyped with 65 SNP markers, and 495 offspring (Subset_B) and their respective 82 G1 parents were genotyped with 80 SNP markers.

A SNP marker was considered unreliable (and therefore excluded from further analyses) when the proportion of individuals displaying assay failure was greater than 30% in the Sequenom output data. Overall, 56 and 63 SNPs were finally retained for Subset_A and Subset_B, respectively. Their positions on the maritime pine linkage map from Chancerel et al. (2013) are provided as supplemental information.

When all loci were considered, the combined power of exclusion was higher than 0.99 for both sets of markers. Moreover, the LD between the 63 SNPs of the optimized markers set used to genotype the Subset_B was very low (r^2: mean=2.5 10^{-3}, min=1.9 10^{-9}, max=0.24).

It should be noted that most (970 out of 984) offspring were genotyped with at least 42 SNPs. The 14 G2 trees below this threshold were simply removed from further analysis (Table 1). Similarly, all G1 parents (except one pollen donor from PMX_A which was not available in clonal archives) were successfully genotyped with at least 42 SNPs. A fraction of males (62 out of the 90 pollen donors) were genotyped twice (by collecting needles on two different ramets from two clonal archives) and compared: all the identities were confirmed.

Paternity recovery

The maternal identities were the first ones we checked. A total of 41 offspring (out of the 970 analyzed) did not conform to their nominal seed donor genotypes. Two types of pedigree errors were highlighted: i) when the error was related to only one G2 offspring in a given family, it was likely that needles of that offspring were collected on an erroneous genotype (this was the case for 11 offspring); ii) when the error was related to all offspring of a given family, mishandled maternal identities were highly probable (this was the case for 3 seed parents and their 10 respective offspring).

All these 41 offspring, identified as maternal mismatches, were excluded from paternity analyses, and thus the paternity recovery was performed on the 929 remaining progeny (460 in Subset_A and 469 in Subset_B, cf. Table 1). Paternity was successfully recovered for 828 offspring: 87% and 91.3% were assigned to a pollen donor member of the PMX (at a confidence of 0.99) in Subset_A and Subset_B, respectively.

Pollen contamination rate in the polymix mating

In this study, we considered the pollen contamination rate as the proportion of offspring which could not be assigned to a known pollen parent of either PMX with a statistical support >99%. Overall, 60

(13%) and 41 (8.7%) offspring in Subset_A and Subset_B, respectively, were clearly fathered by outside pollen (Table 1). The mean pollen contamination rate over the whole sampling was therefore 10.9%.

The contamination rate varied from 0 to 70%, depending on the family, (with 9 to 12 offspring genotyped per family). In Subset_A, 17 families were not contaminated by foreign pollen, 25 families had 1 or 2 offspring sired by foreign pollen parents, and 6 families presented 3 or more offspring sired by foreign pollen parents. In Subset_B, 18 families were not contaminated by foreign pollen, 28 families exhibited 1 or 2 offspring sired by foreign pollen parents and only 1 family had 3 offspring sired by foreign pollen parents.

Reproductive success of pollen parents

All the genotyped pollen donors were represented in the progeny pool (except the single pollen donor not genotyped belonging to PMX_A), although their contributions were variable.

Deviation from expected equal male reproductive success across all females was tested using the chi-square test (only offspring sired by a pollen donor from the PMX were considered for testing the null hypothesis of equal male contribution). The expected frequencies were obtained by dividing the number of offspring correctly assigned by the number of pollen parents in each PMX. In figure 1, the paternal contribution of each pollen donor is represented for both PMX, and the expected equal contribution is indicated by a continuous dark line. In Subset_A (Fig. 1a), paternal contribution varied from 1 to 26 (0.3 to 6.5 %), across all female parents, whereas the expected paternal equal contribution was 8.5 (2.1%). In Subset_B (Fig. 1b), paternal contribution varied from 1 to 24 (0.23 to 5.61 %), whereas the expected paternal equal contribution was 9.95 (2.33%). The null hypothesis that each pollen donor contributes equally to seed production was thus rejected in Subset_A (χ^2= *149; p= 8.6e-13)* and Subset_B (χ^2= *121; p = 1.2e-09*). However, it should be noted that only few pollen donors were out of the confidence interval (at 95% probability) indicated by dashed lines in Fig. 1. Four pollen donors pollinated more than expected in both subsets, while 9 and 8 in Subset_A and Subset_B, respectively, contributed less than expected. Thus, the majority of pollen donors fell within the expected range of variation.

Moreover, the disequilibrium in paternal contributions did not generate an important reduction in genetic variability. Indeed, the ratio between the observed and the expected (if equal contribution) effective number of males (N_em) was 0.81 for PMX_A and 0.84 for PMX_B. Therefore, the observed differential reproductive success did not lead to an important reduction of the effective number of males (N_em.)

Proportions of relatives and coancestry amongst offspring

A major output of this paternity recovery experiment was the reconstruction of the complete mating design. Among the 400 offspring of Subset_A, 352 different crosses could be identified (Fig. 2a): 313, 34, and 5 crosses were represented by one, two and up to six offspring, respectively. In Subset_B (Fig. 2b), 378 different crosses were identified among the 428 analyzed offspring: 331, 44 and 3 crosses were represented by one, two and three offspring, respectively. Thus, in both subsets, PMX families (size: 9-12 offspring) contained between 0 to 6 full-sibs: 32 families consisted solely of half-sibs, and only 8 families comprised more than 3 full-sibs. Moreover, the selfing rate was very low, in spite of 4 and 10 pollen donors in PMX_A and PMX_B, respectively, which also corresponded to seed parents. Only two offspring, belonging to Subset_A, were issued from a self-fertilization of the same genotype. Thus, for most of the HS families studied, the estimated coancestry between progeny within a family did not greatly differ from the expected additive relationship coefficient $2\Theta=0.25$. Indeed, mean coancestry among offspring within maternal families with full pedigree information was 0.260 (ranging from 0.250 to 0.365 according to the families). The family with the highest average coancestry ($2\Theta=0.365$) contained two progeny with an additive relationship coefficient equal to 1 because they were both issued from the selfing of the same female parent.

Moreover, because some G1 parents were related, some relationships among offspring between families were found (Fig. 3). For example, in the partial pedigree model, some individuals had one grand-parent in common ($2\Theta= 0.0625$) and some others had two grand-parents in common ($2\Theta =0.125$). Paternity recovery brought more complexity: relatedness among parents increased the coancestry among offspring within and between families.

Genetic parameters estimation

Genetic parameters were estimated with a partial pedigree (PP) model and a full pedigree (FP) model for the 929 G2 genotyped trees with confirmed seed donors. Subset_A and Subset_B were thus considered together in these analyses. Additive variance and heritability estimates for girth at the age of 12 years and sweep at the age of 8 years for each model are given in Table 2. Heritability estimates computed with the PP model were lower than heritability estimates computed with the FP model for both traits analyzed. The PP model tended to overestimate heritability, especially for sweep. However, the difference between heritability estimated with PP and FP models, were not significant for both traits, as shown by overlapping confidence intervals at 95% probability. The standard errors associated were slightly smaller when the FP model was used.

Estimated breeding values for parental population

Correlation between breeding values predicted using PP and FP models were high for both traits for G1 trees used as seed donors (Fig. 4). Pearson's correlation coefficients were 0.97 for sweep and 0.96

for girth. These high correlations indicated that female breeding values were adequately predicted using the PP model for most of the individuals. Moreover, the high value of Spearman rank correlations between EBVs predicted by FP or PP models (0.97 for sweep and 0.96 for girth) revealed that the ranking of seed donors was similar with both models, indicating that backward selection would be similar whatever the model (PP or FP). However, for a few seed donors, the estimation of breeding values varied substantially when the paternal information was added, and this could not be related to a high level of coancestry within the offspring (seed donors where the within family average relatedness, 2Θ, is greater than 0.27 are represented with closed circles in Fig. 4).

Mean EBVs' accuracy for G1 trees used as seed donors (female EBV) was similar for both models (Table 2). Thus, paternity recovery did not improve the accuracy of parental EBVs for seed donors. However, paternity recovery provided information for genetic evaluation of genotypes used as males, meaning that males could also be included in the backward selection pool. Nevertheless, mean accuracy for genotypes within the PMX (r=0.68 for sweep; r=0.61 for girth) was slightly lower than for seed donors (r=0.70 for sweep; r=0.63 for girth). Furthermore, EBV accuracy varied substantially between genotypes used as males: it ranged from 0.32 to 0.85 for sweep and from 0.27 to 0.80 for girth, depending on the number of progeny per male.

Discussion

Development of a cost efficient genotyping technology for pedigree recovery in maritime pine

According to Glaubitz et al. (2003), "*at least 32 independently segregating SNPs (each with a minor allele frequency of 0.2 or greater) are required to achieve an exclusion probability greater than 0.99*" when one parent is known. In our study, paternity analysis was based on 42 SNPs or more (with MAF>0.45). Thus, our set of markers had very strong discrimination power (probability of exclusion close to 1). This led us to consider a 99% confidence level to determine the delta score used as a criterion for assignment of paternity. This confidence level was higher than the levels commonly used in most of other pedigree recovery studies (95% as high or "strict" confidence and 80% as good or "relaxed" confidence were often used), (Doerksen and Herbinger 2008; Doerksen and Herbinger 2010; El-Kassaby et al. 2011; Funda et al. 2014; Grattapaglia et al. 2014; Hansen and McKinney 2010; Hansen and Nielsen 2010; Klapste et al. 2014; Lai et al. 2010).

Moreover, because the pollen donors were not related in PMX_B and poorly related in PMX_A (coancestry coefficient of 0.036 in PMX_A and 0 in PMX_B), discrimination power was thus enhanced. To conclude, the optimized set of 63 SNPs developed in this study for maritime pine provides a reliable genotyping tool for pedigree recovery, at least for the Landes provenance. The genotyping cost using Sequenom technology was less than €7 per individual. This set of markers is now used routinely for the French maritime pine breeding program: for example, to check pedigrees or tree identities or to evaluate pollution rates in seed orchards.

High rate of pollen contamination

We consider here that the pollen contamination rate is the proportion of progeny with no father identified within the PMX composition. In this study, 13.0% and 8.7% of the progeny were fathered by outside pollen in Subset_A and Subset_B, respectively. However, this contamination rate was overestimated in Subset_A, since one pollen donor belonging to PMX_A was not genotyped (genotype not available). However, if we assign the average number of expected progeny (*i.e.* 8.5 offspring) to this pollen donor, the contamination rate of Subset_A decreases only to 11.2%.

The contamination rate varied greatly according to the family considered. It was higher than contamination rates usually reported in polycross trials in the literature: 4.1% in red spruce (Doerksen and Herbinger 2008), 5% in poplar (Wheeler et al. 2006), 13% in radiata pine (Kumar et al. 2007) and no contamination in Japanese cedar, except in one family with a contamination rate of 22% (Moriguchi et al. 2009).

Numerous hypotheses could explain this high rate of contamination. For the highly contaminated families (for example contamination rate of 70% for one family belonging to Subset_A), it is likely that there was a problem with flower isolation. Other possibilities include the collection of pollens on parents with erroneous identities in the clonal archives. Indeed, mishandled pollen or mislabeled pollen donors identities were possible and some foreign pollens could have been introduced accidentally in the PMX, contributing to increase the contamination. Some studies have already reported mislabeling errors in clonal archives, seed orchards and trials. In a detailed study using allozymes in conifer tree improvement programs in the United States, Adams et al. (1988) found that about 30% of the controlled crosses in Douglas-fir and loblolly pine had at least two progeny different from expectations. Mislabeling of individuals during various field operations was the likely explanation for discrepancies of this sort. In a red spruce polycross trial, three types of pedigree errors were detected (Doerksen and Herbinger 2008): volunteers ("*wild tree unconnected to the pedigree detected as individual progeny mismatching at more than two loci with their supposed mother*"), mishandling and foreign male pollination. Kumar and Richardson (2005) reported that about 5% of the documented relationships (e.g. full-sibs, parent–offspring, etc.) in the New Zealand radiata pine breeding population were identified as errors by molecular markers. In a black walnut progeny test, Zhao et al. (2013) showed that 20% of the genotyped seedlings were not assigned to their theoretical families, and that these misidentifications had a significant effect on the rank of families and potential selection of individuals. They recommended to genotype progeny soon after trial establishment or at least to genotype all selections to verify their identity and pedigree in order to avoid perpetuating an error into future generations and confounding breeding plans. Finally, it is important to note that in this study, the G2 individuals analyzed constituted the third generation of the French maritime pine breeding program. So errors during various fields operations as pollen harvesting, grafting in clonal archives or controlled crossing were accumulated during the three selection cycles leading to different types of mislabeling. This mislabeling hypothesis was strengthened by the 41 genotyped offspring which did not conform to their nominal seed donor genotypes.

Low departure from equal male reproductive success

Both subsets came from two different PMX, but the distribution of paternal contributions was similar in both cases (Fig.1). Even though the equal paternal contribution hypothesis was rejected in both subsets, a homogenous contribution close to the expected mean was observed, and only a few males contributed more or less than expected. This result was in accordance with the majority of pedigree recovery studies that had demonstrated that paternal contributions in controlled polycross trials were unequal (Moriguchi et al. 2009). Wheeler et al. (2006) studied male reproductive success in a 16-parent pollen polymix for Populus: 50% of the pollen donors contributed to 83% of the progeny, and the contribution of each father varied from 1 to 13%. Doerksen and Herbinger (2008) reported that in a

polycross red spruce progeny test, the 21 pollen donors constituting the polymix had unequal reproductive success: while the expected paternal equal contribution was 4.8%, each pollen donor contributed from 1.4 to 11% to the progeny. One father contributed more than twice as much as had been expected; 3 fathers contributed less than twice as much as had been expected. Thus, our results agree with these studies: in Subset_A and Subset_B respectively, 2 and 1 pollen donors contributed more than expected and 10 and 8 pollen donors contributed less than expected.

The causes of unequal pollen contribution have been studied in a large number of species. Aronen et al. (2002), Snow and Spira (1996), and Pasonen et al. (1999) reported that pollen-tube growth rate affects paternal contribution in *Picea abies, Hibiscus moscheutos*, and *Betula pendula,* respectively. Moreover, Parantainen and Pasonen (2004) found pollen-pollen interactions in artificial crossing experiments using pollen mixtures in *Pinus sylvestris*. Nakamura and Wheeler (1992) reported that genetic incompatibility between male and female gametophytes affects paternal contributions in *Pseudotsuga menziesii*. In addition, Moriguchi et al. (2009) found that the contribution of each pollen donor in a polycross was related to their germination rate, but Wheeler et al. (2006) reported that paternal success appeared to be unrelated to pollen vigor.

Moreover, it is known that the number of males constituting the pollen mix has an impact on the differential reproductive success (DRS) (Kumar et al. 2007; Moriguchi et al. 2009). Therefore, using a large number of pollen donors is a general recommendation to decrease variations in the level of paternal contribution. In radiata pine for example, Kumar et al. (2007) recommended using at least 15 males. Finally, the DRS could also be explained by genetic incompatibility (pre- and/or post-zygotic reproductive barriers). But in our study, the sampling did not contain enough progeny per family compared to the number of potential pollen parents to solve this question.

It is also important to note that since only healthy and vigorous seedlings are typically planted in progeny trials and generally not all seedlings survive by the selection age, the relative contributions of pollen parents could deviate from those observed at the sound-seeds stage (Kumar et al. 2007).

Pedigree reconstruction showed that a large number of crosses were actually realized with the polycross mating design, even if few individuals were available for each cross. The 828 progeny analyzed represented 730 different crosses, and 88% of them were realized once (*i.e.* a unique progeny/cross). Despite the DRS of pollen parents, the HS family assumption did not deviate from the null hypothesis and only a few families contained several full-sibs.

Finally, the selfing rate was very low: only two progeny were issued from self-mating, while there was a significant possibility of obtaining selfed offspring (4 and 10 pollen donors in PMX_A and PMX_B, respectively also appeared in the seed donor list). This was not surprising because self-fertilization causes inbreeding depression in conifers, associated with high seedling mortality at an early stage and reduced vigor in plants (Chancerel et al. 2013). In maritime pine, most lethal or sub-lethal alleles are

probably eliminated during seed formation and germination as well as during the first growing season, so that the proportion of selfed offspring is expected to be very low in mature populations (Gonzalez-Martinez et al. 2003). Durel et al. (1996), however, noticed that maritime pine survival rates after a first growing season in the nursery were similar independently of the level of inbreeding, and also observed that there were large differences of inbreeding depression among inbred families for the same level of inbreeding. In our study, the offspring genotyped were 16 years old at the time of pedigree recovery, so several selection steps could have occurred before our analysis (seed development, germination, plantation, *etc*.), possibly resulting in a lower selfing rate than that at seed stage. Noteworthy, only 10 offspring per family were genotyped in our study, so the number of offspring was lower than the number of potential pollen parents. Thus, our sampling was not really appropriate to study the selfing rate and it is likely that the HS family assumption would have deviated more significantly from the null hypothesis if more offspring per family had been genotyped.

Low bias in heritability and EBV due to unknown paternity

In most forest tree breeding programs, genetic parameters estimation relies on open-pollinated or polycross testing schemes. To evaluate the breeding value of selected parents, one can proceed by collecting seeds from desirable parents and establishing HS progeny trials, where a particular parent's breeding value is calculated as the deviation of its progeny mean from the population mean weighted by heritability. These estimations assume true HS families and equal contribution of male parents. Here, we studied the bias in genetic parameters estimation by comparison of heritability estimates calculated with and without paternity information and by calculating the correlation between female EBVs estimated by one model or the other.

Despite the unequal contribution of pollen parents, the average relatedness within families was 0.26 (slightly superior to the additive relationship coefficient of 0.25 expected for half-sibs) and the estimation of heritability was not significantly overestimated for girth and sweep. These trends were already reported in the literature. For example, Doerksen and Herbinger (2010), concluded that in a red spruce PMX trial, the heritability estimate increased slightly when adding pedigree information, mostly due to the capture of genetic variations between males (and not due to the correction of within-family genetic correlation coefficient). Other studies found either no significant bias in heritability estimation due to the inclusion of full-sibs information (Hansen and Nielsen 2010), or an overestimation of heritability estimates (El-Kassaby et al. 2011; Gauzere et al. 2013; Hansen and Nielsen 2010; Korecky et al. 2013) when paternity information was unknown. It should be stated that the sampling effect, the population size, the trait considered and the number of parents involved, all have important effects on heritability estimates.

Moreover, we showed that breeding values estimated with and without pollen parent identity were highly correlated. We also noticed that the presence of full-sibs within families did not consistently

lead to a biased estimation of breeding values. So ranking obtained from polycross designs (without paternity recovery), such as in this study, would be satisfactory to implement an efficient backward selection, provided that the number of pollen parents constituting the PMX is large. Doerksen and Herbinger (2010) reported that correlations between female EBVs estimated through partial and full pedigree models were close to one in a red spruce PMX design. Other studies with open-pollinated families (Hallingback and Jansson 2013; Hansen and Nielsen 2010; Klapste et al. 2014; Korecky et al. 2013) concluded that the ranking of the females based on their EBVs estimated from partial and full pedigree models, was very similar as well.

In many tree breeding programs, the use of polycross designs is restricted to the estimation of female EBVs only, because the lack of information on the male parentage strongly restrains any advanced-generation selection. Lambeth et al. (2001) proposed the use of molecular markers to identify the male parents of forward selections, allowing an enhanced application of polycross designs. They presented practical scenarios for using molecular markers to identify male parentage in polycross designs. Successful implementation of polymix breeding and testing system requires an accurate pedigree reconstruction, low differential in male reproductive success, low pollen contamination rate (since selection is performed only when a complete pedigree is known) and a low selfing rate. This study demonstrates that our polycross progeny trial complied with most of these requirements. Based on these results, implementing a polymix breeding strategy in the maritime pine breeding program will be investigated in the near future.

To conclude, we developed an affordable SNP marker technology providing accurate pedigree recovery for maritime pine. We showed that the pollen contamination rate was rather high (11%), and that the differential reproductive success of pollen parents used in polymixes was significant, due to a few pollen donors contributing more or less than expected. We also demonstrated that heritability and breeding values estimates in this polycross progeny trial are not biased when the pollen donor identity is unknown. Finally, despite recovering paternal identities, this did not result in a significant improvement for backward selection, the new pedigree information enabled forward selection of offspring in the maritime pine breeding program and opened new perspectives for future breeding strategies.

Acknowledgments

This study was funded by INRA (EFPA division "*projet innovant*"), the European-Union (ProCoGen project: n° 289841) and *Conseil Regional d'Aquitaine* (IMAF project co-funded by FCBA: n° 120009468-052). Marjorie Vidal received a CIFRE PhD fellowship (Public/Private Research Partnerships between FCBA and the French Ministry of Higher Education and Research). This study would not have been possible without the support of the maritime Pine Breeding Cooperative (GIS "*Pin Maritime du Futur*"). We gratefully acknowledge all its members. The authors also thank the INRA Experimental Unit (UE0570) for field measurements, Sophie Gerber (INRA) for helpful discussions on paternity recovery, Adline Delcamp (INRA) for the genotyping stage, Jean-Mathieu De Boisseson (FCBA) for needle sampling, Maël Ruby for his preliminary study on paternity recovery, and Fikret Isik (NCSU, USA) for his advice in shaping the content of the paper and his help in ASReml analysis.

The genotyping was performed at the Genomic Facility of Bordeaux (grants from the *Conseil Regional d'Aquitaine*: n°20030304002FA and 20040305003FA, the European Union: FEDER n°2003227 and ANR: n°ANR-10-EQPX-16 Xyloforest).

Data Archiving Statement

All SNPs used in this study were deposited in dbSNP (NCBI database), available on http://www.ncbi.nlm.nih.gov/.

The supplemental data "Supplemental_data_SNP_description.txt" contains the dbSNP accession ID for each SNP (#ss) and some useful information (SNP ID, set identity, contig name, linkage group LG, position on the LG, SNP type, sequence surrounding the SNP).

The genotypes of G2 trees and their potentials parents (G1 trees) are available in the supplemental data file "Subset_A_genotypes_56SNPs.csv" for Subset A and "Subset_B_genotypes_63SNPs.csv" 'for subset B. The first column contains the individual IDs, and the first line contains the SNP IDs. Blank cells are missing data.

References

Adams WT, Neale DB, Loopstra CA (1988) Verifying controlled crosses in conifer tree-improvement programs Silvae Genet 37:147-152

Aronen T, Nikkanen T, Harju A, Tiimonen H, Haggman H (2002) Pollen competition and seed-siring success in *Picea abies*. Theor Appl Genet 104:638-642. doi:10.1007/s00122-001-0789-9

Burdon RD, Shelbourne CJA (1971) Breeding populations for recurrent selection: conflicts and possible solutions. New Zealand Journal of Forestry Science 1: 174

Canales J, Bautista R, Label P, Gomez-Maldonado J, Lesur I, Fernandez-Pozo N, Rueda-Lopez M, Guerrero-Fernandez D, Castro-Rodriguez V, Benzekri H, Canas RA, Guevara MA, Rodrigues A, Seoane P, Teyssier C, Morel A, Ehrenmann F, Le Provost G, Lalanne C, Noirot C, Klopp C, Reymond I, Garcia-Gutierrez A, Trontin JF, Lelu-Walter MA, Miguel C, Cervera MT, Canton FR, Plomion C, Harvengt L, Avila C, Gonzalo Claros M, Canovas FM (2014) De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. Plant Biotechnol J 12:286-299. doi:10.1111/pbi.12136

Chancerel E, Lamy JB, Lesur I, Noirot C, Klopp C, Ehrenmann F, Boury C, Le Provost G, Label P, Lalanne C, Leger V, Salin F, Gion JM, Plomion C (2013) High-density linkage mapping in a pine tree reveals a genomic region associated with inbreeding depression and provides clues to the extent and distribution of meiotic recombination. BMC Biol 11:19. doi:10.1186/1741-7007-11-50

Doerksen TK, Herbinger CM (2008) Male reproductive success and pedigree error in red spruce open-pollinated and polycross mating systems. Can J For Res-Rev Can Rech For 38:1742-1749. doi:10.1139/x08-025

Doerksen TK, Herbinger CM (2010) Impact of reconstructed pedigrees on progeny-test breeding values in red spruce. Tree Genet Genomes 6:591-600. doi:10.1007/s11295-010-0274-1

Durel CE, Bertin P, Kremer A (1996) Relationship between inbreeding depression and inbreeding coefficient in maritime pine (*Pinus pinaster*). Theor Appl Genet 92:347-356

El-Kassaby YA, Cappa EP, Liewlaksaneeyanawin C, Klapste J, Lstiburek M (2011) Breeding without Breeding: Is a Complete Pedigree Necessary for Efficient Breeding? PLoS One 6:11. doi:10.1371/journal.pone.0025737

Funda T, Liewlaksaneeyanawin C, El-Kassaby YA (2014) Determination of paternal and maternal parentage in lodgepole pine seed: full versus partial pedigree reconstruction. Can J For Res-Rev Can Rech For 44:1122-1127. doi:10.1139/cjfr-2014-0145

Gaspar MJ, de-Lucas A, Alia R, Paiva JAP, Hidalgo E, Louzada J, Almeida H, Gonzalez-Martinez SC (2009) Use of molecular markers for estimating breeding parameters: a case study in a *Pinus pinaster* Ait. progeny trial. Tree Genet Genomes 5:609-616. doi:10.1007/s11295-009-0213-1

Gauzere J, Oddou-Muratorio S, Pichot C, Lefevre F, Klein E (2013) Biases in quantitative genetic analyses using open-pollinated progeny tests from natural tree populations. Acta Bot Gall 160:225-236. doi:10.1080/12538078.2013.822827

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) ASReml User Guide Release 3.0. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK

Glaubitz JC, Rhodes OE, Dewoody JA (2003) Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Mol Ecol 12:1039-1047. doi:10.1046/j.1365-294X.2003.01790.x

Gonzalez-Martinez SC, Gerber S, Cervera MT, Martinez-Zapater JM, Alia R, Gil L (2003) Selfing and sibship structure in a two-cohort stand of maritime pine (*Pinus pinaster* Ait.) using nuclear SSR markers. Ann For Sci 60:115-121. doi:10.1051/forest:2003003

Grattapaglia D, Diener PSD, dos Santos GA (2014) Performance of microsatellites for parentage assignment following mass controlled pollination in a clonal seed orchard of loblolly pine (*Pinus taeda* L.). Tree Genet Genomes 10:1631-1643. doi:10.1007/s11295-014-0784-3

Hallingback HR, Jansson G (2013) Genetic information from progeny trials: a comparison between progenies generated by open pollination and by controlled crosses. Tree Genet Genomes 9:731-740. doi:10.1007/s11295-012-0588-2

Hansen OK, McKinney LV (2010) Establishment of a quasi-field trial in *Abies nordmanniana*-test of a new approach to forest tree breeding. Tree Genet Genomes 6:345-355. doi:10.1007/s11295-009-0253-6

Hansen OK, Nielsen UB (2010) Microsatellites used to establish full pedigree in a half-sib trial and correlation between number of male strobili and paternal success. Ann For Sci 67:10. doi:10.1051/forest/2010028

Illy G (1966) Recherches sur l'amélioration génétique du pin maritime. Ann Sci forest 23:765-948. doi:10.1051/forest/19660401

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16:1099-1106. doi:10.1111/j.1365-294X.2007.03089.x

Kimura M, Crow JF (1963) The measurement of effective population number. Evolution 17:279-288

Klapste J, Lstiburek M, El-Kassaby YA (2014) Estimates of genetic parameters and breeding values from western larch open-pollinated families using marker-based relationship. Tree Genet Genomes 10:241-249. doi:10.1007/s11295-013-0673-1

Korecky J, Klapste J, Lstiburek M, Kobliha J, Nelson CD, El-Kassaby YA (2013) Comparison of genetic parameters from marker-based relationship, sibship, and combined models in Scots pine multi-site open-pollinated tests. Tree Genet Genomes 9:1227-1235. doi:10.1007/s11295-013-0630-z

Kumar S, Gerber S, Richardson TE, Gea L (2007) Testing for unequal paternal contributions using nuclear and chloroplast SSR markers in polycross families of radiata pine. Tree Genet Genomes 3:207-214. doi:10.1007/s11295-006-0056-y

Kumar S, Richardson TE (2005) Inferring relatedness and heritability using molecular markers in radiata pine. Mol Breed 15:55-64. doi:10.1007/s11032-004-2059-4

Lai BS, Funda T, Liewlaksaneeyanawin C, Klapste J, Van Niejenhuis A, Cook C, Stoehr MU, Woods J, El-Kassaby YA (2010) Pollination dynamics in a Douglas-fir seed orchard as revealed by pedigree reconstruction. Ann For Sci 67:8. doi:10.1051/forest/2010044

Lambeth C, Lee BC, O'Malley D, Wheeler N (2001) Polymix breeding with parental analysis of progeny: an alternative to full-sib breeding and testing. Theor Appl Genet 103:930-943. doi:10.1007/s001220100627

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Inc., Sunderland, MA, USA, 980 p.

Mangin B, Siberchicot A, Nicolas S, Doligez A, This P, Cierco-Ayrolles C (2012) Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. Heredity 108:285-291. doi:10.1038/hdy.2011.73

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639-655. doi:10.1046/j.1365-294x.1998.00374.x

Moriguchi Y, Ishiduka D, Kaneko T, Itoo S, Taira H, Tsumura Y (2009) The Contribution of Pollen Germination Rates to Uneven Paternity Among Polycrosses of *Cryptomeria japonica*. Silvae Genet 58:139-144

Nakamura RR, Wheeler NC (1992) Pollen competition and paternal success in Douglas-fir Evolution 46:846-851. doi:10.2307/2409655

Parantainen A, Pasonen HL (2004) Pollen-pollen interactions in *Pinus sylvestris*. Scand J Forest Res 19:199-205. doi:10.1080/02827580410029336

Pasonen HL, Pulkkinen P, Kapyla M, Blom A (1999) Pollen-tube growth rate and seed-siring success among *Betula pendula* clones. New Phytol 143:243-251. doi:10.1046/j.1469-8137.1999.00451.x

Plomion C, Chancerel E, Endelman J, Lamy JB, Mandrou E, Lesur I, Ehrenmann F, Isik F, Bink M, van Heerwaarden J, Bouffier L (2014) Genome-wide distribution of genetic diversity and linkage disequilibrium in a mass-selected population of maritime pine. BMC Genomics 15:16. doi:10.1186/1471-2164-15-171

RCoreTeam (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Snow AA, Spira TP (1996) Pollen-tube competition and male fitness in *Hibiscus moscheutos*. Evolution 50:1866-1870. doi:10.2307/2410744

Squillace AE (1974) Average genetic correlations among offspring from open-pollinated forest trees. Silvae Genetica 23:149-156

Wheeler N, Payne P, Hipkins V, Saich R, Kenny S, Tuskan G (2006) Polymix breeding with paternity analysis in *Populus*: a test for differential reproductive success (DRS) among pollen donors. Tree Genet Genomes 2:53-60. doi:10.1007/s11295-005-0033-x

White T (1987) A conceptual framework for tree improvement programs. New For 1:325-342. doi:10.1007/BF00031742

Wiselogel AE, Vanbuijtenen JP (1988) Probability of equal mating in polymix pollinations of loblolly pine (*Pinus taeda* L.). Silvae Genet 37:184-187

Zhao P, Zhang SX, Woeste K (2013) Genotypic data changes family rank for growth and quality traits in a black walnut (*Juglans nigra* L.) progeny test. New For 44:357-368. doi:10.1007/s11056-012-9343-7

Zobel BJ, Talbert JT (1984) Applied Forest Tree Improvement (p. 528). New York: Wiley.

Paternity recovery in two maritime pine polycross mating designs and consequences for breeding

Marjorie VIDAL[1,2,3], Christophe PLOMION[1,2], Luc HARVENGT[3], Annie RAFFIN[1,2], Christophe BOURY[1,2], Laurent BOUFFIER[1,2§]

Tables

Table 1 Descriptive statistics about paternity recovery (with CERVUS) in two maritime pine polymix breeding populations. Subset_A and Subset_B correspond to 2 different PMXs.

Parameter	Subset_A	Subset_B
Total nb of sampled HS families	49	49
HS family size	9-10	9-12
Nb of offspring sampled	489	495
Nb of reliable SNPs genotyped	56	63
Nb of offspring successfully genotyped	478	492
Nb of offspring with seed donor confirmed	460	469
Nb of offspring sired by a PMX pollen donor	400	428
Nb of offspring sired by foreign pollen	60	41
Pollen contamination rate	0.13	0.087
Nb of contributing males	46	43
Mean male contribution (Nb of offspring)	8.70	9.95
Variance in male contribution	26.53	28.76

Table 2 Results from ASReml analyses for girth (age 12) and stem sweep (age 8). Comparison of additive variance components (and corresponding standard errors SE), coefficient of variation (CVa), heritability estimates (and their associated standard errors) and accuracy of the female predicted breeding values (r), according to the pedigree information model (PP: partial pedigree and FP: full pedigree).

Trait	Model	Mean	Additive var	SE	CVa	Heritability (SE)	$r_{(BVs)}$
Girth	PP	53.37	15.90	6.10	0.07	0.25 (0.09)	0.63
	FP	53.37	13.54	3.98	0.07	0.21 (0.06)	0.63
Sweep	PP	19.02	65.57	17.37	0.43	0.47 (0.11)	0.75
	FP	19.02	43.29	10.41	0.35	0.31 (0.07)	0.71

Paternity recovery in two maritime pine polycross mating designs and consequences for breeding

Marjorie VIDAL[1,2,3], Christophe PLOMION[1,2], Luc HARVENGT[3], Annie RAFFIN[1,2], Christophe BOURY[1,2], Laurent BOUFFIER[1,2§]

Figures

Fig. 1 Paternal contribution of the 47 and 43 pollen donors of PMX_A (**a**) and PMX_B (**b**). The average expected contribution (continuous line), and 95% confidence intervals (dashed lines), are indicated. Each vertical bar represents a different pollen donor of the polymix.

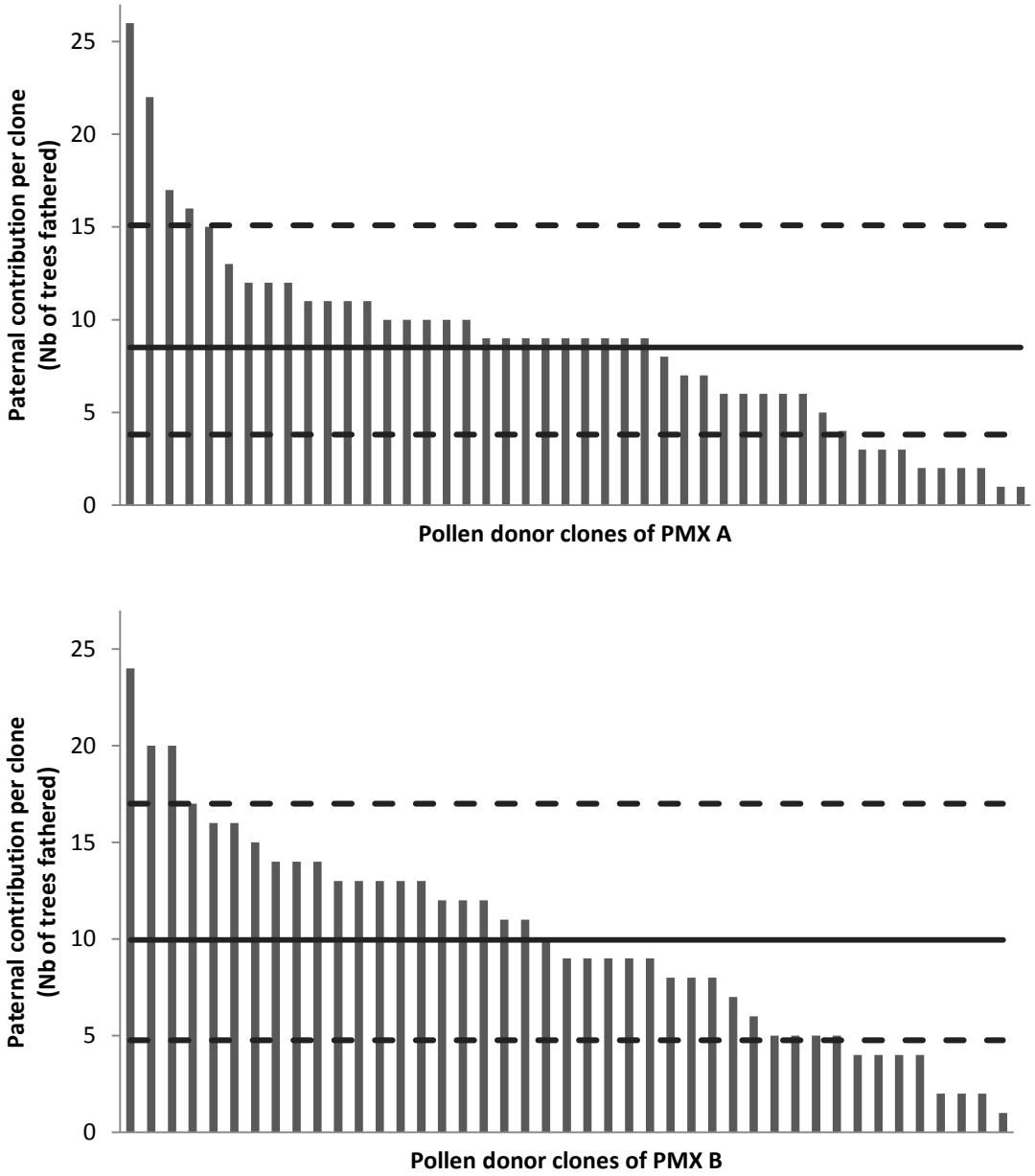


Fig. 2 Mating design revealed by pedigree recovery of the 400 and 428 crossings from Subset_A (**a**) and Subset_B (**b**), respectively.

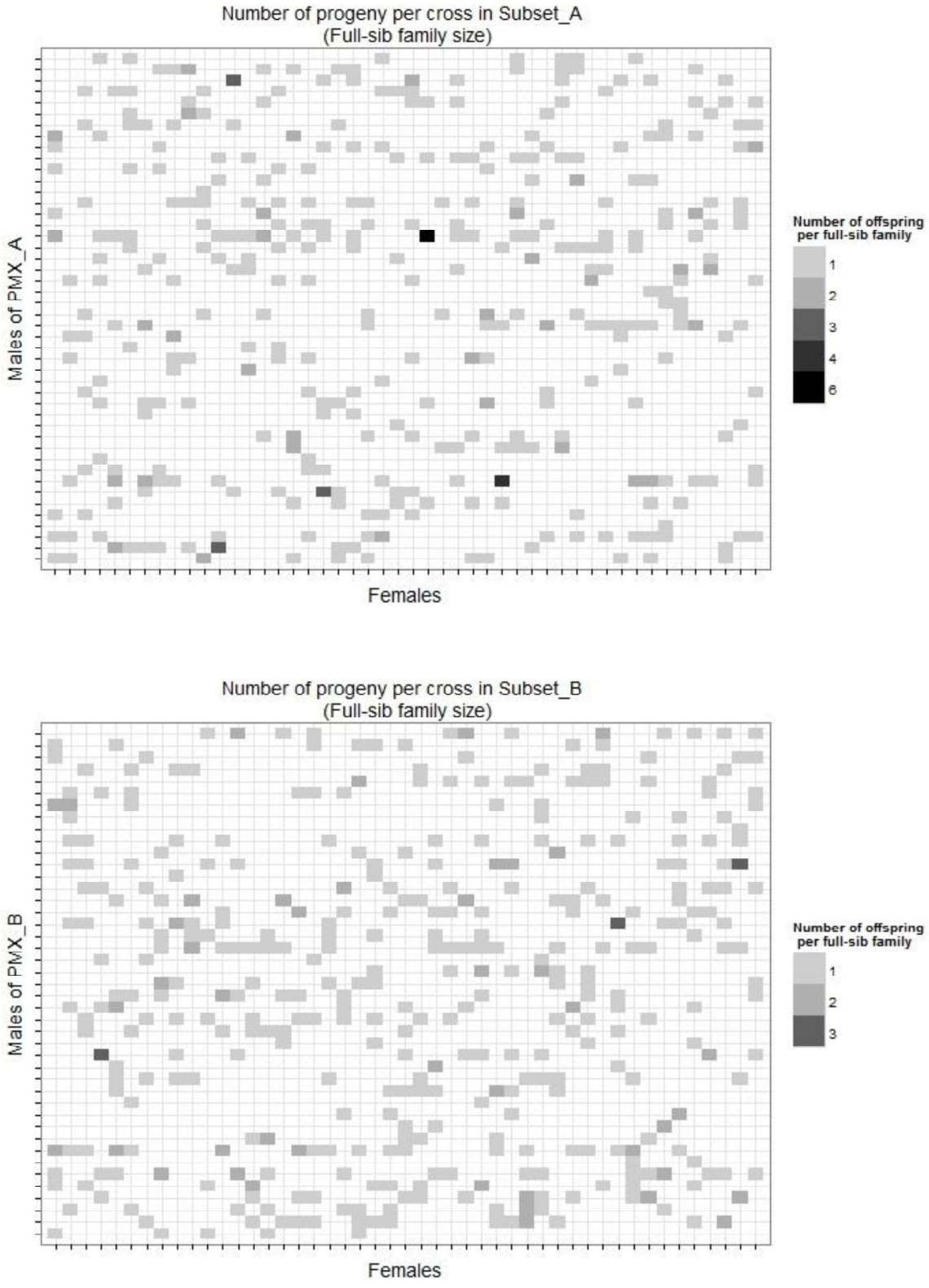


Fig. 3 Distribution of additive relationship coefficients (2Θ from A matrix of size 929) among genotyped G2 trees for partial and full pedigree models. The number of 2Θ=0 were removed from this figure.

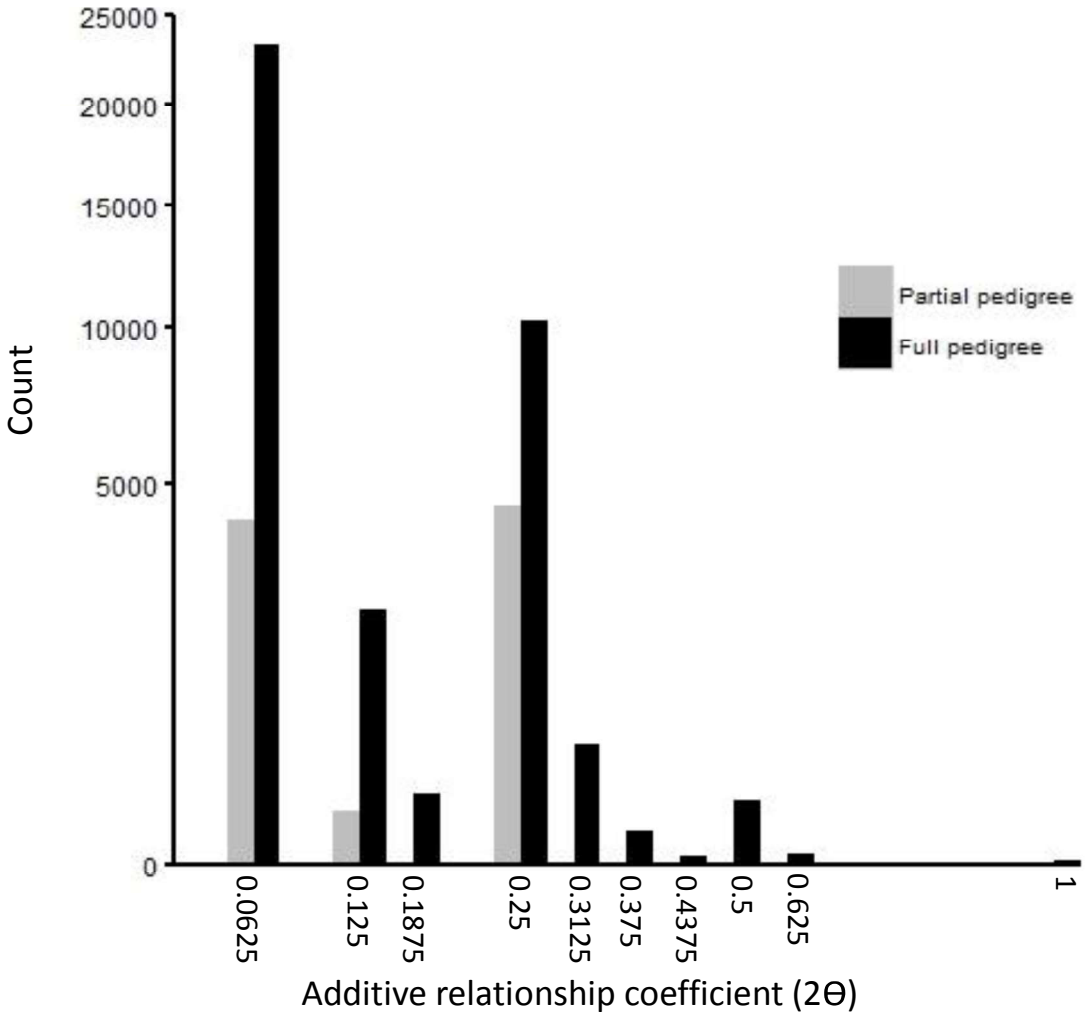
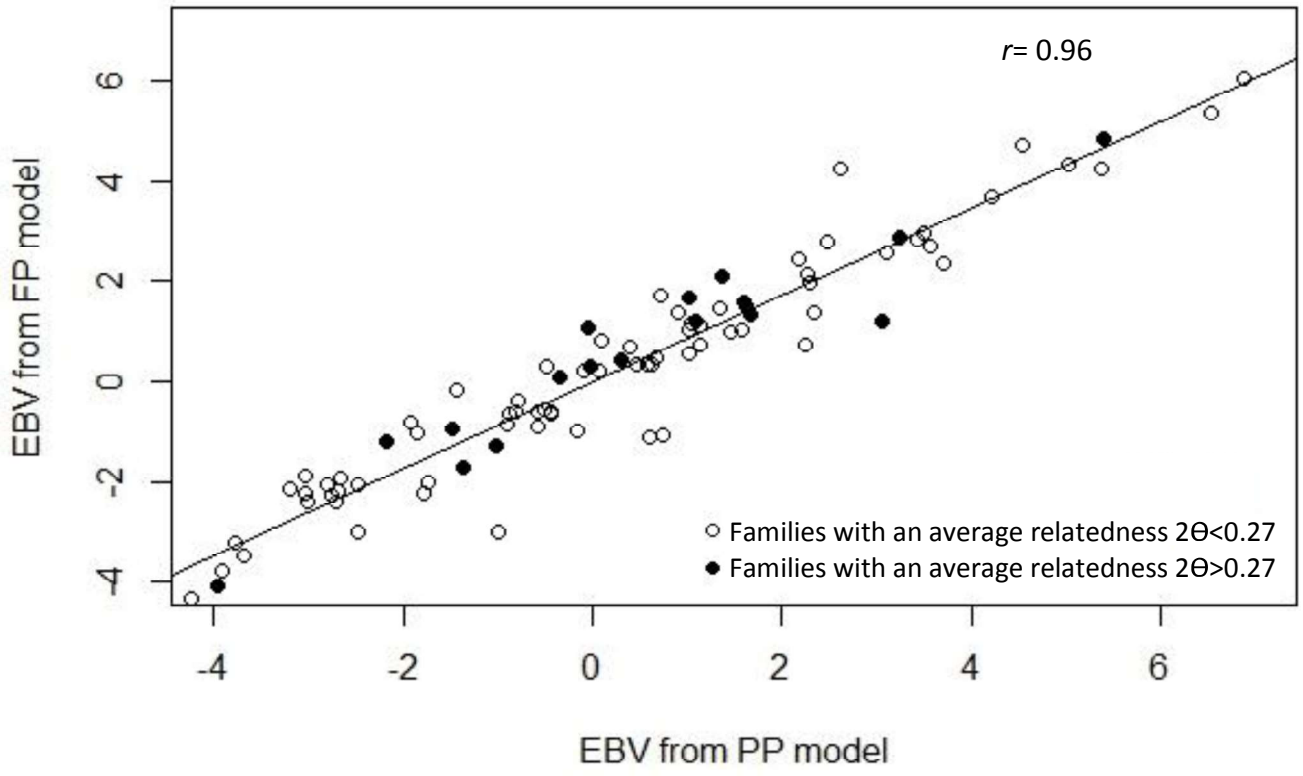


Fig. 4 Correlation between female EBVs from partial pedigree (PP) and full pedigree (FP) models for girth (a) and stem sweep (b). r is the Pearson product moment correlation coefficient. Closed circles represent the 18 families with the highest level of within family coancestry (*i.e.* the 18 seed donors with a family-average relatedness (2Θ) higher than 0.27).

(a) Correlation between EBVs of girth from PP and FP models.



(b) Correlation between EBVs of stem sweep from PP and FP models.