Original Research article

TITLE

Within-family genomic selection in rubber tree (*Hevea brasiliensis*) increases genetic gain for rubber production

David CROS[a,b,c,*], Luther MBO-NKOULOU[d], Joseph Martin BELL[d] , Jean OUM[d], Aurélien MASSON[e], Mouman SOUMAHORO[f], Dinh Minh TRAN[g], Zeineb ACHOUR[b,h], Vincent LE GUEN[a,b], André CLEMENT-DEMANGE[a,b]

[a] CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement), UMR AGAP, F-34398 Montpellier, France

[b] University of Montpellier, CIRAD, INRA, Montpellier SupAgro, Montpellier, France

[c] CETIC (African Center of Excellence in Information and Communication Technologies), University of Yaoundé 1, Yaoundé, Cameroon

[d] Department of Plant Biology, Faculty of Science, University of Yaounde 1, Yaounde, Cameroon

[e] SOCFIN/SOGB (Société des caoutchoucs de Grand Bereby), Abidjan, Côte d'Ivoire

[f] SAPH (Société Africaine de Plantations d'Hévéas), Abidjan, Côte d'Ivoire

[g] RRIV (Rubber Research Institute of Viet Nam), Ho Chi Minh City, Viet Nam

[h] Montpellier SupAgro, Montpellier, France

* Corresponding author.

E-mail address: david.cros@cirad.fr (D. Cros).

ABSTRACT

Genomic selection (GS) could make more efficient the two-stage phenotypic breeding scheme used for rubber production in *Hevea brasiliensis*. It was evaluated using two trials in Côte d'Ivoire comprising 189 and 143 clones of the cross PB 260×RRIM 600, genotyped with 332 simple sequence repeat markers. The effect of statistical genomic prediction methods, training size, and marker data on GS accuracy was investigated when predicting unobserved clone production within and between sites. Simulations using these empirical data assessed the efficiency of replacing current first stage of phenotypic selection (evaluation of seedling phenotype) by genomic preselection, prior to clone trials. Genomic selection accuracy in between-site validations using all clones for training and all markers was 0.53. Marker density and training size strongly affected accuracy, but 300 markers were sufficient and using more than 175 training clones would have marginally improved accuracy. Using the 125 to 200 markers with the highest heterozygosity, between-site GS accuracy reached 0.56. Prediction methods did not affect GS accuracy. Simulations showed that genomic preselection on 3,000 seedlings of the considered cross would have increased selection response for rubber production by 10.3%. *Hevea* breeding programs can be optimized by the use of within-family GS. Further studies considering other crosses and traits, consecutive breeding cycles, more contrasted environments, and cost-benefit ratio are required.

Keywords: marker assisted selection, genomic predictions, selection response, clonal varieties

1. Introduction

The rubber tree (*Hevea brasiliensis*, hereafter *Hevea*) is almost the only source of commercial natural rubber (1,4 cis-polyisoprene), 70% of which is used by the tire industry. The production of natural rubber worldwide has increased steadily over time, and is now exceeding 12 Mt yearly (FAOSTAT, 2018). The total cultivated area, currently over 11 million hectares, is held by smallholders (80%) and industrial estates (20%). More than 90% of the production takes place in Asia, with Thailand and Indonesia as the largest producers. Côte d'Ivoire is the seventh world producer, and produced 420 thousand tons in 2017. Predictions indicate that demand for natural rubber will exceed 19 Mt in 2025 (Warren-Thomas et al., 2015), even though rubber plantations are already responsible for deforestation and pose threats to biodiversity, in particular in South-East Asia (Ahrends et al., 2015; Warren-Thomas et al., 2015). Yield therefore needs to be intensified in existing plantations to meet the expected demand while minimising environmental cost and increasing the income of poor producers. Genomic selection (GS) (Meuwissen et al., 2001), the state-of-the-art method of marker-assisted selection for quantitative traits, can play a key role in taking up this challenge.

Hevea is a diploid species ($2n = 36$) belonging to the family Euphorbiaceae and originating from the Amazonian forest. Vegetative multiplication by grafting permitted the development of clonal varieties from axillary buds grafted on seedling rootstocks (rubber clones). The initial 'primary' clones derived from ortet selection among populations of non-budded trees resulting from natural pollination. Controlled recombination by hand pollination was then applied to cross the best clonal parents for the generation of full-sib families. However, the naturally-low female fertility of *Hevea* makes it difficult to construct complex populations of connected families, with highly incomplete mating designs and strong imbalance in family sizes. This generally did not allow to accurately estimate parental genetic values and to take advantage of the large within-families variability. This prompted us to adopt a within-family clonal breeding program which focused specifically on certain large-sized F1 families (≥200 individuals) obtained with the few parent trees that combined good agronomic performances and female fertility sufficient to reach the targeted size. Since the 1990s and the development of molecular genetic markers, these large and highly performing F1 families also gave the opportunity to

acquire genetic information about the parents using genetic mapping and quantitative trait loci (QTL) detection (Clément-Demange et al., 2007). These strategies of QTL detection have been applied to various traits: resistance to *Pseudocercospora ulei* (Le Guen et al., 2011, 2007; Lespinasse et al., 2000) and to *Corynespora cassiicola* (Tran et al., 2016), vegetative growth and latex production (An et al., 2019; Rosa et al., 2018; Souza et al., 2013). However, for complex traits under the control of a large number of genes with small effects, such as yield, the efficiency of marker assisted selection approaches based on QTLs is limited, because it overestimates the effect of the strong QTLs while weak QTLs are not detected (Muranty et al., 2014).

Currently, *Hevea* breeding involves within-family two-stage phenotypic selection (PS) followed by large-scale agronomic evaluation (Figure 1, left). Although a large number, i.e. several thousand, of full-sibs can be evaluated in the first stage (seedling evaluation trial, SET, with non-replicated individuals), selection for rubber production at this stage is not very accurate (Bombonato et al., 2015; Gnagne, 1988). The second stage consists in small-scale clone trials (SSCT, with each genotype replicated in the form of several budded trees). The SSCTs make it possible to accurately assess clone yield, but the number of clones that can be evaluated in these trials is relatively low (< 200). This is followed by a long period of agronomic evaluation of growth rate, latex production, disease resistance, and other characteristics at the scale of tapped stands, in multi-local large-scale clone trials (LSCT).

GS is a very promising way to increase the rate of genetic progress in perennial crops (Grattapaglia, 2017; van Nocker and Gardiner, 2014) because it allows the genetic value of a large number of selection candidates to be estimated at an early stage. In *Hevea*, the current SETs prior to clone trials could thus be replaced by more accurate genomic preselection. If GS is sufficiently accurate, it could even replace SSCTs. However, the decision to shift from a conventional PS scheme to a GS alternative calls for detailed studies. Indeed, the relative rate of genetic gain of different breeding approaches depends on their respective selection accuracy, selection intensity, and generation interval, with a trade-off between these parameters due to practical and economic constraints in the breeding program and to biological constraints of the species. Standard statistical methods for GS predictions include random regression best linear unbiased predictor (RR-BLUP) (Meuwissen et al.,

2001), Bayesian least absolute shrinkage and selection operator regression (BLR) (de los Campos et al., 2009), and Bayesian reproducing kernel Hilbert Space (RKHS) (Gianola and van Kaam, 2008). BLR and RR-BLUP are linear approaches with different assumptions regarding the distribution of marker effects. Thus, RR-BLUP estimates marker effects following a normal distribution with common variance for all markers, while BLR uses a variance specific to each marker. RKHS is a semi-parametric and non-linear approach (i.e. using a non-linear genomic matrix) (Pérez-Rodríguez et al., 2012) that can capture both additive and non-additive effects (Zhang et al., 2016). When the purpose is to predict genetic values potentially including non-additive genetic effects (like clone values), it is appropriate to use models that take non-additive effects into account, either by modelling them explicitly (as it can be done in RR-BLUP and BLR) or implicitly (RKHS).

Despite the great economic importance of *Hevea*, no study has yet been published on the efficiency of GS compared with conventional PS in this species. Here, an alternative within-family breeding scheme for *Hevea* rubber production was suggested, in which the current phenotypic preselection of individual seedlings (SET) prior to clone trials (SSCT) would be replaced by genomic preselection in the nursery (Figure 1, right). As the GS model needs to be trained using phenotypic data, this alternative scheme would involve two SSCTs. The first, comprising a random sample of candidate clones (i.e. with no prior selection from SET results), would be used both to evaluate these candidates and to train the GS model; the second would be used to finalise the selection among the clones preselected by the GS model. The efficiency of the GS scheme compared with conventional PS will result from the accuracy and selection intensity (i.e. the number of clones genotyped to undergo genomic preselection) of GS. Genomic selection accuracy is usually estimated by within-site cross-validation. However, such estimates may be biased upwards (Beaulieu et al., 2014; Lorenz et al., 2011, p.94; Ly et al., 2013), and GS accuracy is consequently better estimated by validation using independent sites.

The aim of this study was to carry out the first evaluation of GS for *Hevea*, using genotypic and phenotypic data on one family at two sites. For this purpose, two within-family clonal selection strategies for rubber production were compared: a new breeding scheme combining genomic preselection and PS, and the current conventional PS scheme. More precisely, (1) within- and

between-site GS accuracy were estimated for rubber production in unobserved clones, (2) the effect of three parameters on GS accuracy was evaluated: statistical method of genomic prediction, size of training population, and molecular data (density and filtering), and (3) the increase in performance of the selected clones and in response to selection that could be expected from combining GS and PS compared with conventional PS was estimated using simulations based on the empirical data and on the between-site estimate of GS accuracy. Data on 330 clones from the F1 cross between two widely cultivated rubber clones (PB 260 × RRIM 600) were used, with phenotypic data collected from two independent clone trials in Côte d'Ivoire (189 clones at Site 1 and 143 clones at Site 2) and genomic data on 332 simple sequence repeat (SSR) markers.

2. Materials and methods

2.1 General overview

The study was divided into two parts. The aim of the first part was to obtain empirical estimates of GS accuracy for rubber production of unobserved clones of a F1 cross, with two independent field trials used for within- and between-site validations. The second part of the study aimed to estimate the additional annual response to selection that could be expected from combining GS and PS rather than using conventional PS. This was done by simulations based on the empirical data and on the GS accuracy estimated in the first part of the study.

For the first part (empirical estimation of GS accuracy), data on 330 clones from the F1 cross PB 260 × RRIM 600 were used. The clones were evaluated in two independent SSCTs in Côte d'Ivoire, with 189 clones at Site 1 and 143 at Site 2. The trials were implemented using conventional experimental designs, which allowed reliable estimations of clone values (hereafter referred to as phenotypes). The clones were also genotyped with 332 simple sequence repeat (SSR) markers. The GS model, trained using the molecular data and phenotypes of one part of the clones, predicted the phenotype of the other clones, for which molecular data only were used as inputs to the model. This made it possible to measure the accuracy of GS predictions of the performance of clones yet-to-be observed. The GS validation analyses were performed for predictions within and between sites to

158 assess the usefulness of within-site accuracies (i.e. obtained by cross validation) for decision-taking
159 regarding the practical implementation of GS. In addition, five standard statistical methods for
160 genomic prediction were compared in terms of GS accuracy. Different training size, marker density,
161 and SSR sampling method (sampling random SSRs or SSRs with the highest observed heterozygosity,
162 *Ho*) were also used to quantify the effect of these three parameters on GS accuracy.
163 In the second part of the study (comparison of GS and PS schemes using simulation), the
164 current conventional phenotypic breeding scheme (Figure 1, left) and an alternative scheme combining
165 genomic preselection and PS (Figure 1, right) were simulated, and the two approaches were compared
166 in terms of performance of the selected clones and of annual response to selection. The simulation was
167 calibrated with the empirical data and with the results obtained in the first part of the study (genetic
168 variance, PS accuracy, GS accuracy, etc.).
169
170 ## *2.2 Empirical estimation of GS accuracy*
171 ### *2.2.1 Plant material and phenotyping*
172 The parents of the F1 cross, PB 260 and RRIM 600, are two well-known and genetically unrelated
173 clones that were selected in Malaysia. RRIM 600 originated from a cross made in 1937 (TJIR 1 ×
174 PB 86) and is the most widely planted clone in the world due to its high latex yield generated soon
175 after tapping initiation and good adaptation to a variety of environments. Its potential for rubber
176 production is medium. PB 260, issued from the cross PB 5/51 × PB 49, was obtained in 1958. It is a
177 vigorous and high-yielding clone, largely used as female in crossings because it has one of the highest
178 female fertilities among the best rubber clones used as parents, thus allowing for much larger
179 progenies than other female parents (Baudouin et al., 1997). It was recommended for plantation in
180 Asia in the 1980s and 1990s and is still the second most widely planted clone in rubber-producing
181 countries.
182 The two study sites are located in the coastal area of south-western Côte d'Ivoire: Site 1
183 (latitude: 4°40'54" N, longitude: 7°06'05" W, on the SOGB [*Société des caoutchoucs de Grand*
184 *Bereby*] estate, elevation: 33 m a.s.l., gravelly clayey loam, with 189 clones), and Site 2 (latitude: 5°

19' 47.79" N, longitude: 4° 36' 39.74" W, on the SAPH [*Société Africaine de Plantations d'Hévéas*] estate, elevation: 89 m a.s.l. deep sandy soil, with 143 clones). The two sites lie approximately 300 km from each other. The sites have a similar tropical climate, with 1,600 mm mean annual rainfall and a mean annual temperature of 26°C. Two clones were used at both sites, giving a total of 330 clones. The two shared clones were not used for GS validation, but only to train the GS model. There was no preselection of the clones before the trials (i.e. no SET evaluation).

Site 1 and Site 2 trials were planted in July 2012 and July 2013 respectively, following almost complete block designs, with six blocks and individual trees randomised within each block, and with a planting density of 1,600 trees per hectare, with a spacing of 2.5 × 2.5 metres. Ramets were produced in the nursery by grafting on rootstocks generated from seeds issued from natural pollination of clone GT 1, and transplanted to the trials. The mean number of ramets per clone was 11 (range: 7 - 17) at Site 1 and 13 (5 - 20) at Site 2. This led to 2,016 ramets at Site 1 and 1,869 at Site 2.

Rubber production was recorded for each ramet according to the following protocol. The tapping system was in half-spiral on the trunk at 1 m above ground level, tapping every two days excepted on Sundays. In each trial, the six blocks were tapped by three tappers, with two blocks assigned to each tapper during the three consecutive months of the experiment. Each tapper tapped one block per day. No ethephon stimulation was applied to the trees in order to assess the natural latex flow of every tree. The latex was collected in plastic boxes of 180 ml attached to the trunks with rubber bands, and covers were screwed over the boxes between two tappings for preserving the latex production from rain. Every day in each trial, the boxes full of latex were collected and replaced by empty ones. The coagulated latex from the full boxes was extracted and manually pressed to eliminate the liquid serum. Tapping started 32 months after planting in Site 1 (end of dry season) and 38 months after planting in Site 2 (end of rainy season). For each ramet, the amount of rubber aggregated from all collected boxes during the 3-month tapping period (with a dry rubber content of around 65%) was computed. These raw production data were analysed for each site with a linear mixed model and the BLUP methodology, using the ASReml-R version 3.0 package (Butler et al., 2009). This gave the clone genetic values, adjusted for effects related to the experimental designs (blocks) and for variations in size among the trees at the time of tapping (i.e. variations in girth of the trunk measured

at 1 m above the ground just before initiation of tapping). These adjusted clone values are hereafter referred to as phenotypes. The broad sense heritability of clone mean level (H^2) was calculated at each site as per Eq. (1).

$$H^2 = \sigma_G^2 / (\sigma_G^2 + \frac{\sigma_E^2}{h_r}) \quad (1)$$

Where σ_G^2 is the genetic variance of the clones, σ_E^2 the residual variance and h_r the harmonic mean number of ramets per clone in the trial (Gonçalves et al., 2006), with σ_E^2 and σ_G^2 obtained from the linear mixed model.

2.2.2 Marker genotyping

Leaf samples were collected on the original mother-trees of the clones issued from the seeds of the cross. Genomic DNA extraction and SSR genotyping were carried out following the method described by Le Guen et al. (2009). Site 1 clones were genotyped with 332 SSRs (Tran et al., 2016), and Site 2 clones were genotyped with a subset of 296 SSRs (Achour, 2014). Table 1 lists the characteristics of the SSR molecular data obtained at each site. Sporadic missing SSR data were imputed with BEAGLE 3.3.2 (Browning and Browning, 2007), with parameters *niterations* set to 25 and *nsamples* to 20.

2.2.3 Statistical methods for genomic predictions

Three GS statistical methods were used to predict the genetic values of the validation clones: RR-BLUP, BLR and RKHS. In addition, BLR and RR-BLUP were carried out with two types of model, i.e. purely additive models (BLR_A and RR-BLUP-A) and additive plus dominance models (BLR_AD and RR-BLUP_AD). We did not consider the explicit modelling of epistatic effects for the sake of simplicity and assuming they would be negligible over additive and dominance effects.

For RR-BLUP_A and BLR_A, the model given by Eq. (2) was used.

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{Z_a m_a} + \boldsymbol{e} \quad (2)$$

Where $\boldsymbol{y}$ is the ($k \times 1$) vector of phenotypes of training clones, k the number of clones, μ the overall phenotypic mean, $\mathbf{1}$ a column vector of 1s, $\boldsymbol{m_a}$ the ($n \times 1$) vector of allele additive effects, with n the total number of alleles, $\boldsymbol{Z_a}$ the ($k \times n$) incidence matrix with elements $Z_{a_{ij}}$ = 0, 1 or 2 indicating the

number of alleles j for clone i, and $\boldsymbol{e}$ the vector of residual effects following $N(0, \sigma^2_e)$, with σ^2_e the residual variance. For RR-BLUP-AD and BLR_AD, the previous model was extended as indicated by Eq. (3).

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{Z_a m_a} + \boldsymbol{Z_d m_d} + \boldsymbol{e} \quad (3)$$

Where $\boldsymbol{m_d}$ is the ($p \times 1$) vector of dominance effects of all possible pairs of alleles at each SSR, with p the total number over all SSRs of possible combinations between two alleles of the same SSR, and $\boldsymbol{Z_d}$ the ($k \times p$) incidence matrix with elements $Z_{d_{ij}} = 1$ or 0, indicating whether clone i possesses allele combination (pair) j or not. The genomic estimated genetic value (GEGV) $\hat{g}_i$ of the validation clone i was obtained by Eq. (4) in RR-BLUP_A and BLR_A,

$$\hat{g}_i = \sum_{j=1}^{n} Z_{a_{ij}} \hat{m}_{a_j} \quad (4)$$

and by Eq. (5) in RR-BLUP_AD and BLR_AD,

$$\hat{g}_i = \sum_{j=1}^{n} Z_{a_{ij}} \hat{m}_{a_j} + \sum_{j=1}^{p} Z_{d_{ij}} \hat{m}_{d_j} \quad (5)$$

with $\hat{m}_{a_j}$ the estimated additive effect of allele j, and $\hat{m}_{d_j}$ the estimated dominance effect of the j^{th} pair of alleles. For RR-BLUP, the $\hat{\boldsymbol{m}}_{\boldsymbol{a}}$ and $\hat{\boldsymbol{m}}_{\boldsymbol{d}}$ vectors were the BLUP solutions; and for BLR they were the posterior mean values over the post burn-in iterations. For BLR, σ^2_e followed a scaled inverse chi-square prior distribution, and $\boldsymbol{m_a}$ and $\boldsymbol{m_d}$ followed conditional Gaussian prior distributions $N(0, \tau^2_{aj}\, \sigma^2_e)$ for allele j and $N(0, \tau^2_{dj}\, \sigma^2_e)$ for allele pair j, respectively. The τ^2_{aj} parameters were thus specific to each allele j, and the τ^2_{dj} to each allele pair j; and they followed exponential priors with rate $\lambda^2_a / 2$ and $\lambda^2_d / 2$, respectively, with the regularisation parameters λ^2_a and λ^2_d following gamma priors.

For RKHS, the model presented in Eq. (6) was used.

$$\boldsymbol{y} = \mathbf{1}\mu + \boldsymbol{g} + \boldsymbol{e} \quad (6)$$

Where $\boldsymbol{g} = \boldsymbol{K\alpha}$ is the vector of random genetic values of clones, $\boldsymbol{K}$ the ($c \times c$) kernel constructed from the SSR data of the c clones, c the total number of clones (i.e. training and validation clones) and $\boldsymbol{\alpha}$ the ($c \times 1$) vector of regression coefficients to be inferred, with prior distribution $N(0, \boldsymbol{K}\sigma^2_\alpha)$. $\boldsymbol{K}$ gave the covariance structure among clones and had elements given by Eq. (7).

$$K_{ij} = e^{-hd_{ij}^2} \quad (7)$$

With d_{ij}^2 the squared Euclidean distance between clones *i* and *j* computed from their SSR genotypes, and *h* a bandwidth parameter. A multi-kernel approach based on a set of values of *h* (0.1, 0.5, 2.5) was implemented, as explained in Pérez and de los Campos (2013). The vector of GEGV $\hat{\boldsymbol{g}}$ was obtained as per Eq. (8).

$$\hat{\boldsymbol{g}} = \boldsymbol{K}\hat{\boldsymbol{\alpha}} \quad (8)$$

The BGLR R package version 1.0.5 (Pérez and Campos, 2014) was used for BLR and RKHS with 30,000 iterations, with the first 9,000 as burn-in and a thinning interval of 10. For RR-BLUP, ASReml-R version 3.0 package (Butler et al., 2009) was used.

2.2.4 Validation approaches

The analyses were performed for predictions within and between sites, leading to four different validation approaches (Site 1 cross validation, Site 1 towards Site 2, Site 2 cross validation and Site 2 towards Site 1).

The clones from each site were randomly allocated into *k* sets used as validation replicates, with *k*=7 for Site 1 and *k*=5 for Site 2. In this way, the number of clones per set was similar in the validation experiments: 27 for Site 1 and 28 or 29 for Site 2, depending on the set. The allocation of clones to the validation sets was the same for all validation scenarios. Within-site validations were conducted using *k*-fold cross-validation approaches, successively using one of the *k* sets as the validation set and the remaining *k*–1 sets (or only some of them when varying the training size, see below) to train the GS model. For between-site validations, the *k* sets from one site (or some of the *k* sets when the training size was varied) were used to train the GS model, and the *k* sets of the other site were used for validation. The GS predictive ability was obtained for each set as the Pearson correlation between the GEGV ($\hat{g}$) and the phenotype (*y*) of the clones composing the set. Finally, GS accuracy was the predictive ability divided by the square root of the broad sense heritability H^2 (Lorenz et al., 2011, p. 94).

2.2.5 Effect of training size and SSR density

To quantify the effect of training size on GS accuracy, the number of sets composing the training population varied from one to k–1 (within-site validations) or k (between-site validations). For a given number of n_s sets used for training, the different possible combinations of n_s sets among the available sets (i.e. k–1 for within-site validations or k for between-site validations) were used successively. These combinations of sets became training replicates. The resulting total number of replicates (validation replicates × training replicates) per validation experiment and training size varied from 5 to 175 (see Supplementary Table S. 1 for details).

To investigate the effect of SSR density on GS accuracy, different numbers of SSRs were also used, considering six levels of number of SSRs, from 10 to all SRRs. For a given number of SSRs, eight replicates of random samples of SSRs were made.

2.2.6 Effect of SSR sampling method

To investigate whether sampling SSRs with high observed heterozygosity (*Ho*) would lead to higher GS accuracy than randomly selected SSRs, *Ho* was computed for each SSR as the mean percentage of heterozygous individuals and the validations described above were run considering 12 levels of number of SSRs, from 10 to all SRRs. For a given number of SSRs, eight replicates of random samples of SSRs were made. With SSR sampling selecting the highest *Ho*, four replicates were also made, as some SSRs had the same *Ho* (in which case, the SSRs were chosen randomly). Here, all the clones were used to train the GS model.

2.2.7 Analysis of results

To study the effect of the statistical method for genomic prediction on GS accuracy, analyses of variance (ANOVA) were performed separately for each validation approach on the accuracy obtained using all SSRs and all the clones for training, with statistical method and validation replicate as factors. The mean levels of factors in the ANOVAs were compared using Tukey's honest significant difference test. To assess the effect of the SSR sampling method, the Wald-type permutation test of the R package GFD (Friedrich et al., 2017) was used, given that the assumptions of normality and

variance homogeneity were not met, with SSR sampling method and validation replicate as factors. The tests were carried out separately for each SSR level of each validation approach. Prior to these analyses, GS accuracy underwent Fisher's Z transformation.

2.3 Comparison of combined GS/PS and PS breeding schemes

The application of the conventional PS scheme on cross PB 260 × RRIM 600 (see Figure 1, left) was simulated, as well as a GS scheme in which clones of the same cross evaluated in a first SSCT would be used to train a GS model to make a preselection among unobserved clones (seedlings) of the same cross prior to their final evaluation in a second SSCT (see Figure 1, right). The simulation was calibrated with the results of the linear mixed model initially implemented to obtain the phenotypes, and with the results of the between-site empirical GS validations.

The simulation procedure started with the joint simulation of the true genetic values (TGV) (g), the seedling phenotypes in SET (y'), the estimated genetic values in SSCT (EGV, i.e. phenotypes) (y), and the genomic estimated genetic values (GEGV) ($\hat{g}$) of n individuals as per Eq. (9).

$$n = max(3{,}000,\ 190+n_{GS}) \qquad (9)$$

With the 3,000 seedlings evaluated in SET, 190 clones evaluated in the first SSCT (used for both phenotypic selection and training of the GS model), and n_{GS} the number of additional selection candidates allowed by GS (i.e. candidates subjected to genomic preselection at the nursery stage, prior to the second SSCT), with n_{GS} varying from 100 to 5,000. These values were simulated using the *mvrnorm* function in the MASS R-package (Venables and Ripley, 2002). This required the variance-covariance matrix between g, y', y and $\hat{g}$ given in Eq. (10),

$$\begin{pmatrix} \sigma_g^2 & Cov(g,y') & Cov(g,y) & Cov(g,\hat{g}) \\ Cov(g,y') & \sigma_{y'}^2 & Cov(y',y) & Cov(y',\hat{g}) \\ Cov(g,y) & Cov(y',y) & \sigma_y^2 & Cov(y,\hat{g}) \\ Cov(g,\hat{g}) & Cov(y',\hat{g}) & Cov(y,\hat{g}) & \sigma_{\hat{g}}^2 \end{pmatrix} \qquad (10)$$

and the mean phenotypic value of the clones μ, which were obtained as follows. The correlation between y' and y ($r_{y',y}$) was 0.34, taken from Gnagne (1988), and gave the correlation between rubber production in SET and SSCT for an unselected population related to the cross used in this study. For each site, the clone phenotypes given by the initial mixed model analyses were used to compute the

associated variance, σ_y^2. This initial analysis also gave μ, and the accuracy of phenotypic selection ($r_{g,y}$, corresponding to the square root of H^2) in a SSCT with no preselection (genomic or based on SET evaluation). The variance of the TGV of the clones (σ_g^2) was obtained as per Eq. (11) (Clark et al., 2012, appendix 1).

$$\sigma_g^2 = \sigma_y^2 / {r_{g,y}}^2 \qquad (11)$$

The GEGVs obtained from the empirical between-site validations were used to compute the associated variance, $\sigma_{\hat{g}}^2$. The GS accuracy ($r_{g,\hat{g}}$) was taken from the between-site validations, and the correlation between y and $\hat{g}$ was obtained as per Eq. (12) (Lorenz et al., 2011, p. 94; Muranty et al., 2015, appendix).

$$r_{y,\hat{g}} = r_{g,\hat{g}}\, r_{g,y} \qquad (12)$$

Similarly, the variance of the seedling phenotypes ($\sigma_{y'}^2$) was calculated as per Eq. (13), the correlation between y' and g per Eq. (14), and the correlation between y' and $\hat{g}$ per Eq. (15).

$$\sigma_{y'}^2 = \sigma_y^2 / r_{y',y}^2 \qquad (13)$$

$$r_{g,y'} = \sigma_g / \sigma_{y'} \qquad (14)$$

$$r_{y',\hat{g}} = \sigma_{\hat{g}} / \sigma_{y'} \qquad (15)$$

The mean values over the two sites were computed for each of these parameters and were used to calibrate the simulation.

The n simulated individuals served as starting point for the simulation of the conventional PS scheme and the alternative scheme combining GS and PS. For PS, a random set of 3,000 individuals was sampled among the n simulated individuals. Among them, the 190 individuals with the highest performance in SET (i.e. highest y') were retained to make the first SSCT, and the n_{sel} = 10 clones with the highest EGV were finally selected among them. For combined GS/PS, a random set of 190 clones were sampled among the n simulated individuals to make the first SSCT. Then, the 185 clones with highest GEGV were selected among the n_{GS} simulated clones (i.e. among those that were not evaluated in the first SSCT), to make the second SSCT (only 185 instead of 190 in the first SSCT, since in practice some clones from the first SSCT would be repeated in the second). Finally, n_{sel} = 10 clones with the highest EBV were selected among the clones evaluated in the two SSCTs. The

performance of the clones selected in combined GS/PS and PS schemes was computed as the mean TGV of the n_{sel} selected clones. The annual selection response of the PS and combined GS/PS schemes was computed as the difference between the mean TGV of the n_{sel} selected clones and the mean TGV of the n initial clones, divided by the number of years required to complete the breeding cycle (25 years) (Figure 1). The selection intensity of PS and combined GS/PS was computed as the mean EGV of the n_{sel} selected clones and the mean EGV of the n initial clones, divided by the standard deviation of the EGV. The simulation process was repeated 5,000 times.

All analyses and simulations were conducted using the R software, version 3.4.1 (R Core Team, 2017).

3. Results

3.1 Phenotypic evaluations

Mean cumulated rubber production per tree was 78.7 g in Site 1 (range 0.50 – 318.0) and 244.6 g in Site 2 (range 0.25 - 840.1). The broad sense heritability of clone mean level (H^2) was 0.9 at each site.

3.2 Statistical methods for genomic predictions

The GS accuracies obtained for rubber production were not affected by the statistical method used for predictions. When training the GS models with all clones and using all the SSRs, the mean GS accuracy over validation replicates ranged from 0.33 to 0.60 (Figure 2). However, this variation was mostly due to the validation approach, with statistical methods having a negligible effect. The differences in accuracy between statistical methods were thus not significant, regardless of marker density and size of training set, with no interaction found between the GS prediction method and SSR number (Supplementary Fig. S 1), nor between the GS prediction method and training size (Supplementary Fig. S 2). For the rest of the study, only the BLR_A GS prediction method was used. Indeed, its mean accuracy across the four validation approaches when using all SSRs and all clones for training (0.498) was slightly higher than that of the other prediction methods (whose mean accuracy

400 ranged from 0.488 for BLR_AD to 0.495 for RR-BLUP_A). Furthermore, it was the method that came
401 out with the best average rank of the four validation approaches. When all the SSRs and all the clones
402 were used for training, BLR_A gave a mean GS accuracy of 0.594 in Site 1 cross validation, 0.509 in
403 Site 1 towards Site 2, 0.340 in Site 2 cross validation and 0.550 in Site 2 towards Site 1 validation.
404
405 *3.3 Training population size and molecular marker data*
406 GS accuracy for rubber production was strongly affected by the number of clones used to train the GS
407 prediction model (training size) and by the number of SSRs (Figure 3).
408 GS accuracy increased with the training size regardless of validation approaches and number
409 of SSRs used. For instance, when all the SSRs were used, increasing the training size from minimum
410 to maximum values (i.e. by an average of 447.7%, from 296% in Site 2 cross validation to 600% in
411 Site 1 towards Site 2), GS accuracy approximately doubled (mean of +93.6% across validation
412 approaches, from 72.8% in Site 1 cross validation to 111.7% in Site 1 towards Site 2). With all
413 validation approaches and numbers of SSRs, the increase in GS accuracy associated with increased
414 training size followed a diminishing returns pattern. Thus, when 296 SSRs were used, increasing the
415 training size from 28 to 56 clones increased GS accuracy by an average of 36.9% in the four validation
416 approaches, while doubling the training size again to reach 111 clones increased the GS accuracy by
417 slightly less (32.1%). Although usually GS accuracy did not reach a plateau, the shape of the curves
418 showed that further increases in training sizes would have led only to minor additional gains in GS
419 accuracy (except for Site 2 cross validation, due to the smaller overall population size). Similarly, with
420 all the validation approaches and training sizes, GS accuracy increased with the number of SSRs.
421 Thus, increasing the number of SSR from minimum to maximum values (i.e. by an average of 3.0%,
422 with 3.2% in Site 1 cross validation and 2.9% in other validations) when using the maximum training
423 sizes, the average GS accuracy over validation approaches increased by 201.6% (from 134% in Site 1
424 cross validation to 296.2% in Site 2 cross validation). Again, a diminishing returns trend was observed
425 for all validation approaches. For instance, with the largest training sizes, using 50 SSRs instead of 25
426 SSRs increased GS accuracy by 36.1% on average across all the validation approaches, while doubling
427 again SSR density increased GS accuracy by 16.5% only. In Site 1 cross validation, for which more

SSRs were available, using 332 SSRs resulted in the same accuracy as using 296. This indicated that no extra gain could be expected here from using more SSRs.

When SSR density was reduced, using the SSRs with the highest observed heterozygosity (*Ho*) generally resulted in significantly higher GS accuracies than using random SSRs (Figure 4). In particular, when the 125 to 200 SSRs with the highest *Ho* were used, GS accuracies were always significantly higher than the accuracies obtained with random SSRs, with an average increase of 13.9% (from 4.6% in Site 2 cross validation with 200 SSRs, to 21.1% in Site 2 towards Site 1 validation with 150 SSRs). Furthermore, in this range of number of SSRs, the *Ho* sampling approach led to almost always higher accuracies than using all SSRs, with an increase in GS reaching an average of 4.3% for the four validation approaches compared with using all the SSRs (the only exceptions being with 150 SSRs in Site 1 cross validation and with 200 SSRs in Site 2 cross validation, when GS accuracy with *Ho* SSR sampling was very slightly lower than when all SSRs were used). Mean GS accuracy of between-site validations thus reached 0.561, versus 0.530 using all SSRs. As expected, due to high variations in *Ho* among SSRs (Table 1), the SSR samples based on this parameter had a much higher mean *Ho* than the whole set of markers (Supplementary Fig. S 3). Thus, when using 125 to 200 SSRs, mean *Ho* was 0.78, as against 0.64 with all the SSRs.

3.4 Validation approach

The effect of the validation approach on GS accuracy was investigated by comparing accuracies among validation approaches using the same training size and number of SSRs. In this case, within-location analysis gave much higher accuracies for Site 1 than for Site 2 (Figure 3). For instance, using 296 SSRs, within-Site 1 GS accuracy was 0.54 with 108 clones for training, versus only 0.34 for within-Site 2 accuracy with 115 training clones. By contrast, between-locations accuracies were similar when making predictions from Site 1 towards Site 2 and from Site 2 towards Site 1; and between-location GS accuracies were intermediate between the two within-site accuracies.

Site 1 cross-validation accuracy overestimated Site 1 towards Site 2 accuracy for all training sizes and numbers of SSRs (Figure 3). Thus, when using all the clones for training and all the SSRs, Site 1 cross-validation accuracy was 0.60, while Site 1 towards Site 2 accuracy fell to 0.51 (-14.9%).

By contrast, Site 2 cross-validation accuracy largely underestimated Site 2 towards Site 1 accuracy, for all training sizes and numbers of SSRs. Thus, when using all the clones for training and all the SSRs, Site 2 cross-validation accuracy was 0.34, while Site 2 towards Site 1 accuracy reached 0.54 (+61.7%).

Regarding the advantage of the *Ho* SSR sampling method over random sampling (Figure 4), consistent results were obtained between each within-site experiment and the between-site experiment in which the considered site was used for training and the other site for validation: in both cases, SSR sampling based on *Ho* gave higher accuracies than random sampling. In addition, the number of SSRs that gave the highest accuracy with *Ho* SSR sampling was the same in Site 2 cross validation and in Site 2 towards Site 1 validation (150), in both cases leading to higher GS accuracy than when all the SSRs were used. Similarly, although the number of SSRs that produced the highest accuracy with *Ho* SSR sampling in Site 1 cross validation and in Site 1 towards Site 2 validation differed (200 and 125, respectively), using the number of SSRs that gave the highest accuracy in Site 1 cross validation for Site 1 towards Site 2 validation would still have increased GS accuracy compared with using all SSRs.

3.5 Comparison of combined GS/PS and PS breeding schemes

The variance-covariance matrix between g, y, and $\hat{g}$ used to calibrate the simulation is given Figure 5. The mean phenotypic value (aggregated amount of rubber) was 186 g. The GS accuracy ($r_{g,\hat{g}}$) was 0.561, corresponding to the mean accuracy obtained in between-site validations with the 125 to 200 SSRs with the highest *Ho*. The accuracy of SET ($r_{g,y'}$) was 0.358.

The simulation showed that combining GS and PS outperformed conventional PS in terms of rubber production of the selected clones and annual selection response when genomic preselection was applied to a sufficient number of candidates, i.e. at least 1,000. In this case, additional rubber production was observed in the clones selected using GS (Figure 6). With 1,000 candidates, this additional production was very low (+0.4%) but increased when more candidates were used for preselection, and reached 5.9% when preselection was applied to 5,000 candidates. This led to an increase in annual response to selection when combining GS and PS compared with conventional PS,

483 which started from +1% with 1,000 clones subjected to genomic preselection and reached +15% with
484 5,000 candidates for genomic preselection (Figure 7). The results also indicated that using a larger
485 population of candidates for genomic preselection would have further increased the superiority of
486 combined GS/PS over conventional PS, albeit only slightly. Genotyping 3,000 candidates for genomic
487 preselection appeared as a good compromise between genotyping effort and efficiency of combined
488 GS/PS (+4.2% mean production for the selected clones, corresponding to a +10.3% increase in annual
489 selection response). In contrast, combining GS and PS performed worse than conventional PS when
490 100 and 500 candidates only were used for the genomic preselection.

491 As these values were means of 5,000 replicates of the simulation, they show the average extra
492 gain that would result from the application of genomic preselection in many replicates of the F1 cross
493 studied here. This is of major interest for breeders, but the actual gain that would be achieved in a
494 given breeding program is also crucial. To assess this (and in particular to assess the probability that a
495 given application of the GS scheme indeed performs better than the current PS scheme), Figure 7 also
496 shows the distribution of the relative performance of combining GS and PS compared with
497 conventional PS, in the form of a boxplot for each number of candidates for genomic preselection. For
498 instance, the figure shows that with 2,000 candidates for genomic preselection, although the mean
499 expected extra annual response to selection generated by GS reaches 7%, the first quartile is only
500 slightly above the value corresponding to a similar performance by PS. This indicates that, although
501 on average over a large number of replicates combining GS and PS will be better than conventional
502 PS, for a specific application there is an almost 25% risk that GS would actually not perform better, or
503 even worse, than PS (with the lowest value obtained being an annual selection response of GS
504 reaching only 78.3% that of PS). Therefore, the best way to decide on the size of the population of
505 selection candidates for genomic preselection is to consider both the mean expected annual selection
506 response of GS and the distribution of the possible values around this mean. Thus, using 3,000
507 candidates, 75% of the simulation replicates gave an annual selection response of combined GS/PS at
508 least 4.5% higher than when using conventional PS (with the maximum value reaching +45.6%), and
509 the risk of GS actually performing worse than PS was low, at 9.2%. When increasing the number of
510 candidates to 5,000, this risk dropped to 4.6%.

The increase in the relative performance of combined GS/PS compared with conventional PS when more candidates are used for genomic preselection resulted from the associated increase in selection intensity of combined GS/PS. Selection intensity in combined GS/PS was 15% lower than PS when 100 candidates were used for genomic preselection, but became roughly equivalent to PS with 1,000 candidates. It further increased to reach +14% when 5,000 candidates were used for the genomic preselection (see Supplementary Fig. S 4 for details). The fact that the selection intensity of the combined GS/PS scheme with 1,000 candidates to genomic preselection was similar to the selection intensity of the PS scheme with 3,000 individuals in SET resulted from the existence of two stages of selection. Indeed, genomic preselection can retain elite clones for SSCT that could have been discarded from the SET results, since the accuracy of genomic predictions is higher than SET accuracy. As a consequence, the 10 best clones selected at the end of the SSCT tend to perform better if the SSCT is preceded by genomic preselection rather than SET, leading to a higher selection differential and thus higher selection intensity in the combined GS/PS scheme than in the conventional PS scheme.

Finally, the better performance of the combined GS/PS scheme compared with conventional PS was the consequence of the greater selection accuracy of genomic preselection compared with phenotypic preselection with SET (GS accuracy being 56.7% higher) and of the greater selection intensity achieved when the number of candidates to genomic preselection was sufficiently high (≥1,000).

4. Discussion

The results presented here showed that applying the suggested breeding scheme combining GS and PS can increase rubber production in the cross PB 260 × RRIM 600. However, the advantage of this new breeding scheme over conventional PS resulted from GS accuracy, genetic variance and selection accuracy in SET and in SSCT, which vary among single crosses and traits. In particular, even in the case of GS implemented within full-sib families like here and despite the existence of deterministic equations, it remains difficult to predict GS accuracy for a particular trait in a given

family (Schopp et al., 2017). The study therefore needs to be extended to other families and traits, in particular using contrasted F1 crosses in terms of genetic and phenotypic variation.

4.1 Relevance of within-family GS for Hevea

The within-family GS scheme investigated here will not require restructuring breeding activities, already organised around full-sib families, and this is clearly a practical advantage for breeders. In addition, breeding schemes in which selection is applied within single crosses (i.e. full-sib families) are favourable situations for GS. In such biparental populations, there is a high linkage disequilibrium between marker alleles and gene alleles, which reduces the required marker density (as full-sibs share large chromosome segments), and there is no group structure (Crossa et al., 2017; Lin et al., 2014). Good results of within-family GS as implemented here have been reported in other plant species, with GS accuracies reaching moderate (i.e. between 0.5 and 0.7, as in the present study) to high values. For instance, GS accuracy estimated with a single-site cross validation was around 0.6 in a family of 500 Sitka spruce clones (Fuentes-Utrilla et al., 2017) and between 0.59 and 0.91 in a family of 180 *Citrus* clones (Gois et al., 2016).

A possible drawback of the within-family GS approach presented here is that it might not always be possible to obtain a training population of sufficient size. *Hevea* breeding programs use several families with limited resources, and the size of each family is therefore constrained. With the family used here, it appeared that using 175 clones to train the GS model was enough. However, this figure is close to the maximum amount of resources breeders can invest in a single family, and some families could require a larger training size, depending on their level of genetic variation. An alternative to the within-family approach studied here could be to implement GS in a population comprising several interconnected families, obtained using incomplete diallel or factorial mating designs. Although such a population would not be easy to obtain in *Hevea* due to the species' low female fertility, a comparison with within-family GS would be informative. This type of GS approach is implemented in a number of perennial species, including loblolly pine, spruce, eucalyptus (Grattapaglia, 2017), apple (eg Kumar et al., 2015; Muranty et al., 2015), and citrus (Minamikawa et al., 2017). This is interesting as it leads to a single (and therefore larger) training population compared with the various family-specific training

566 populations required for the within-family GS approach. However, this increase in training size,
567 although beneficial for GS accuracy, would be offset by a decrease in relatedness between the training
568 set and the application set, a situation known to have a negative impact on GS accuracy. Therefore, in
569 practice, a GS approach using a complex population involving several families could be more
570 complicated to manage, with GS accuracy varying among selection candidates depending on their
571 actual relationship with the training individuals. This could also actually lead to lower GS accuracies
572 than family-specific training populations (Crossa et al., 2017; Lenz et al., 2017; Schopp et al., 2017;
573 Toro et al., 2017; Würschum et al., 2017). In addition, from a practical point of view, the time needed
574 to achieve and release a commercial clone could be longer with a complex multiparental population
575 than with separate F1 families. This has to be taken into consideration as it represents a risk for *Hevea*
576 breeding, where cycles are long and the resources invested in breeding activities are very limited.
577
578 ### *4.2 Comparison of combined GS/PS and PS breeding schemes*
579 The most important point for breeders regarding GS is the annual selection response that could
580 result from its use, compared with the annual selection response of PS (Resende et al., 2017).
581 Although PS and GS selection accuracies play a crucial role in this comparison, other factors that
582 affect annual genetic gain must also be taken into consideration, i.e. relative generation interval and
583 selection intensity of PS and GS. A few studies have ventured beyond estimating empirical GS
584 accuracies and have used these estimates to evaluate the possible gain in annual selection response that
585 GS could elicit. In eucalyptus, GS annual selection response is expected to be 50% to 300% greater
586 than that of current PS, depending on the reduction in the duration of the breeding cycle and on GS
587 selection intensity (Resende et al., 2012, 2017). In black spruce, annual selection response should be
588 200% higher with the GS approach than with conventional selection, thanks to the shorter GS breeding
589 cycle (Lenz et al., 2017). In *Citrus*, annual selection response is expected to increase by 31% to 420%,
590 depending on how much the breeding cycle is shortened and on the trait concerned (Gois et al., 2016).
591 In *Hevea*, like in other perennial crops, the full potential of GS will be achieved over
592 consecutive breeding cycles. Given the data available for this first GS study in this species, it was only
593 possible to consider a single breeding cycle, whose duration could not be reduced due to the need for a

SSCT to train the GS model. This explains why the increase in annual selection response reported here may seem modest compared with that reported in studies on other perennial crops. However, beyond the first cycle, breeding cycles will become shorter: only one SSCT will be required, since the GS model will have been calibrated with data from the first cycle. In addition, the training population used in the second cycle will comprise the aggregated data of the two SSCTs of the first cycle, and in the following cycles the data of the new SSCTs will be added to the training population. This is known to enhance GS accuracy (Auinger et al., 2016; Cros et al., 2018; Denis and Bouvet, 2013). Further studies are needed to investigate the efficiency of GS over several cycles in *Hevea*.

Another possibility would be to consider a GS scheme with only one SSCT in which the genomic predictions would be used to select clones in the second nursery, before their final evaluation in LSCT, instead of using GS to make a preselection before SSCT. This would have the advantage of reducing the duration of the breeding cycle. However, a simulation similar to the one presented here showed that, within the range of the number of selection candidates that can reasonably be genotyped, this approach was not advantageous in terms of annual selection response because the steep decline in accuracy between the SSCT (0.95) and the genomic predictions (0.561 on average over the two between-sites validations, i.e. a 40.9% decrease) was not offset by the shorter generation interval and/or higher selection intensity made possible by GS (data not shown). Our study therefore focused on a GS scheme in which the use of GS methodology was limited to the replacement of the conventional seedling evaluation trials prior to clone trials, and it showed this was sufficient to enhance the efficiency of the breeding scheme. A similar result was obtained in an oil palm study (Cros et al., 2017), which evaluated the usefulness of genomic preselection prior to field evaluation, i.e. without reducing the breeding cycle duration, like in the present study. It thus showed that genomic preselection would increase bunch production by 6.5% to >10% when 2,000 to 10,000 candidates are used for genomic preselection.

Here, we used a single PS breeding scheme in order to benchmark the breeding scheme combining GS and PS. However, several PS schemes are possible. For instance, Gireesh et al. (2017) suggested the use of clonal nursery trials to optimize phenotypic breeding. It would therefore be

interesting to implement new simulation studies to consider a broader range of possible PS and GS schemes.

4.3 Within-site and between-site accuracies

The between-site GS accuracies obtained in this study and the resulting estimate of annual selection response are appropriate for the environment considered here. The higher GS accuracy obtained in Site 1 cross validation compared with Site 1 to Site 2 validation was expected from the literature, which indicated that within-site cross-validations can lead to upward biases in GS accuracy (Beaulieu et al., 2014; Lorenz et al., 2011, p.94; Ly et al., 2013). For instance, in another perennial crop, black spruce, Lenz et al. (2017) obtained GS accuracy from between-site validation lower than accuracies obtained within the same site. The relatively small difference between the accuracies of Site 1 cross-validation and of Site 1 to Site 2 validation, and the similar accuracies when making predictions from Site 1 towards Site 2 and from Site 2 towards Site 1, indicated that genotype × environment (G × E) interactions, that could have been generated by differences in locations and years, were weak – probably because the two environments were similar. However, significant G × E interactions can occur in *Hevea* (see for example Costa et al., 2000; Gonçalves et al., 2006; Tan, 1995), and in this case the between-site GS accuracy would certainly be lower. In this case, the solution would be to take the environment into account in the prediction model. For this purpose, rubber geneticists will benefit from the methodology developed in cereals and legumes, where G × E modelling in the context of GS has been extensively studied (Crossa et al., 2017). Surprisingly, GS accuracy obtained in Site 2 cross-validation was lower than the GS accuracy found in Site 2 towards Site 1 validation. What determined this result at this site remains unclear.

The effect of number of markers and SSR sampling method (random or based on high *Ho*) observed for a single-site cross validation was in good agreement with the results obtained when a GS model was calibrated at this site to predict the values of clones evaluated at the other site. This indicated that, in the environment considered here, a single-site cross validation experiment made it possible to identify the number of SSRs and the method for choosing the SSRs that would yield the

best GS accuracy that can be expected from using this experiment to train a GS model for predicting the rubber production of clones at another site.

4.4 Molecular data

In this study SSR markers were used, whereas in the vast majority of GS experiments in animals and plants, genotyping is carried out with single nucleotide polymorphism (SNP) markers. Simple sequence repeats were used here because this type of marker has already been shown to be efficient in GS validation studies in oil palm (Cros et al., 2015; Marchal et al., 2016), table grapes (Viana et al., 2016), and flax (You et al., 2016); and also because the biparental nature of the plant material used here suggested that the marker density achievable with SSRs could be sufficient. In the present study, 300 SSRs, which is a rather low density compared with what is usually found in GS studies, turned out to be sufficient to achieve the maximum GS accuracy that could be reached here. However, this result holds for the F1 cross and for the training population size considered here, and it is possible that, in a different situation (for example with a larger training population), the GS accuracy would benefit from the use of more markers. Also, with the dataset considered here, it was possible to further reduce marker density with a slight increase, or at least no loss, in GS accuracy by using a subset of the 125 to 200 SSRs with the highest *Ho*. With multi-allelic markers in a single cross between heterozygous parents, *Ho* actually indicates how informative the markers are. Thus, the SSRs with *Ho*=1, which was the case for 25 to 50 SSRs per validation (Supplementary Fig. S 3), were those for which the two parents RRIM 600 and PB 260 had no alleles in common. When the two parents were heterozygotes, this corresponded to a situation with a balanced representation of the four alleles in the cross (the frequency of each allele being around 25%). This suggests that the marker density required to reach maximum GS accuracy is likely to vary among F1 crosses, depending on parental relatedness and heterozygosity. Other parameters were used for SSR screening (polymorphism information content (Botstein et al., 1980, p. 320) and expected heterozygosity, *He*) but preliminary analyses indicated that filtering using *Ho* yielded better results (data not shown).

The practical implementation of GS will require a high throughput and a cost-effective genotyping method to make the screening of large populations of selection candidates feasible. Even a

reduced panel of SSRs might not be competitive in terms of cost compared with genotyping approaches involving SNPs. In addition, if the method is implemented over several generations, it will probably be necessary to increase marker density in order to limit decline in accuracy (Grattapaglia, 2017, p. 216). To our knowledge, there is currently no SNP array available in *Hevea*, but genotyping by sequencing (GBS) (Elshire et al., 2011), which has already been used in this species to construct a high density linkage map (Pootakham et al., 2015), could generate the molecular data required for GS in *Hevea*. Furthermore, approaches specific to biparental crosses that combine GBS and a relevant imputation methodology could be used to further increase the cost efficiency of large-scale genotyping (Gorjanc et al., 2017; Technow and Gerke, 2017).

4.5 Models and statistical methods for genomic predictions

Like in the present study, empirical GS accuracies have frequently been found to be unaffected by the statistical method of prediction (here BLR, BRR, and RKHS). Several examples with similar results are thus available in perennial crops. For various growth traits in eucalyptus, similar accuracies were obtained using BLR, RR-BLUP, and RKHS by Tan et al. (2017a), and using BLR and Bayesian random regression (BRR, similar to RR-BLUP) by Müller et al. (2017). In oil palm, BLR and BRR gave the same accuracies for yield components (Cros et al., 2015).

The explicit modelling of dominance effects with BLR_AD and RR-BLUP_AD had no effect on GS accuracy. Simulations in eucalyptus showed that including dominance in the GS model for prediction of clone performance improved accuracy when dominance effects were preponderant (ratio of dominance to additive variance of 1.0) and heritability was high (H^2=0.600) (Denis and Bouvet, 2013). Simulations in loblolly pine showed that including dominance in the GS prediction model improved accuracy when the ratio of dominance to phenotypic variance was over 20% (de Almeida Filho et al., 2016). With empirical data on eucalyptus, Tan et al. (2017b) reported that GS accuracy for traits with large dominance variance was increased by including dominance effects in the model. However, in apple, Kumar et al. (2015), empirically obtained similar GS accuracies with models with or without non-additive effects for fruit quality traits, despite a high proportion of non-additive variance in some traits. This apparent discrepancy could come from the fact that Kumar et al. (2015)

used a training population of around 230 individuals, much smaller than that used by the previously cited authors, who used training populations of at least 800 individuals. It can therefore be hypothesised that, in the present study, including dominance effects in the GS models did not affect accuracy because dominance variance was not large enough and/or because the training populations were too small (from 114 to 189 individuals). Similar reasons are likely to explain the fact that RKHS did not perform better than the other methods.

5. Conclusions

The within-family GS strategy investigated here will lead to the release of more productive *Hevea* clones than clones selected with the current PS scheme. This will increase the yield of rubber from existing plantations, and thus help to meet the demand for natural rubber while minimising environmental costs. With a F1 cross between two widely cultivated clones, PB 260 × RRIM 600, a mean empirical GS accuracy of 0.53 was obtained in predictions between two independent sites when using all the clones for training and all the SSRs. SSR density and training size markedly affected GS accuracy. Mean between-site GS accuracy reached 0.561 when using the 125 to 200 SSRs with the highest *Ho*. In contrast, the statistical method used to obtain the genomic predictions of clone values did not affect GS accuracy. Based on this empirical result, simulations showed that by applying a genomic preselection among 3,000 seedlings in the nursery prior to clone trial, instead of the current low-accuracy phenotypic preselection on 3,000 seedlings, the rubber yield of the clones selected in the F1 cross considered would have been 4.2% higher, corresponding to a 10.3% increase in annual selection response. This resulted from the greater selection accuracy of genomic preselection compared with phenotypic preselection.

The results presented here showed that combining GS and PS can increase rubber production in the cross PB 260 × RRIM 600. However, before generalising GS in rubber breeding, this study needs to be extended to other families because the results obtained, and in particular the GS accuracies and selection response, are affected by the genetic characteristics of the parents of the F1 cross used. Similarly, studies considering other traits, such as growth and architecture, are needed. It is also

necessary to compare GS and PS in terms of selection response per unit cost and to investigate the efficiency of GS over consecutive breeding cycles, which will make it possible to shorten the breeding cycle in the cycles following model training. Furthermore, using a broader range of environments for between-site validations will be of major interest.

Acknowledgments

We thank the *Institut Français du Caoutchouc* (IFC), and the companies Michelin, SIPH, and SOCFIN for funding the part of this research devoted to *Hevea* rubber in the framework of the IFC project "Hevea varietal creation in West-Africa". We also thank the companies SOGB (*Société Grand Bereby*) and SAPH (*Société Africaine de Plantations d'Hévéa*) for their logistical assistance in the field experiments in Côte d'Ivoire. We acknowledge Laurence Dedieu for reviewing the paper, and the CETIC (African Center of Excellence in Information and Communication Technologies) for support.

Data availability

The datasets generated and analysed during the current study are available from the corresponding author.

Conflict of interests

The authors declare no conflict of interest.

Author contributions

DC carried out data analysis and wrote the paper, with help of ACD and VLG. LM, JO, and JB carried out preliminary data analysis. AM and MS provided assistance and logistics for trial setting up and phenotyping. VLG and ACD designed field experiments, supervised collection of phenotypic data and generation of molecular data by DMT and ZA.

References

Achour, Z., 2014. Cartographie génétique de la famille F1 PB260 x RRIM600 (*Hevea brasiliensis*) et détection de QTLs associés à la croissance. Mémoire de Master 2 SEPMET, Montpellier-Supagro, UMR AGAP, CIRAD, Montpellier, 43 p.

Ahrends, A., Hollingsworth, P.M., Ziegler, A.D., Fox, J.M., Chen, H., Su, Y., Xu, J., 2015. Current trends of rubber plantation expansion may threaten biodiversity and livelihoods. Glob. Environ. Change 34, 48–58. https://doi.org/10.1016/j.gloenvcha.2015.06.002

An, Z., Zhao, Y., Zhang, X., Huang, X., Hu, Y., Cheng, H., Li, X., Huang, H., 2019. A high-density genetic map and QTL mapping on growth and latex yield-related traits in Hevea brasiliensis Müll.Arg. Ind. Crops Prod. 132, 440–448. https://doi.org/10.1016/j.indcrop.2019.03.002

Auinger, H.-J., Schönleben, M., Lehermeier, C., Schmidt, M., Korzun, V., Geiger, H.H., Piepho, H.-P., Gordillo, A., Wilde, P., Bauer, E., Schön, C.-C., 2016. Model training across multiple breeding cycles significantly improves genomic prediction accuracy in rye (*Secale cereale* L.). Theor. Appl. Genet. 129, 2043–2053. https://doi.org/10.1007/s00122-016-2756-5

Baudouin, L., Baril, C., Clément-Demange, A., Leroy, T., Paulin, D., 1997. Recurrent selection of tropical tree crops. Euphytica 96, 101–114. https://doi.org/10.1023/A:1002908918879

Beaulieu, J., Doerksen, T., MacKay, J., Rainville, A., Bousquet, J., 2014. Genomic selection accuracies within and between environments and small breeding groups in white spruce. BMC Genomics 15, 1048.

Bombonato, A.L., Gouvêa, L.R.L., Verardi, C.K., Silva, G.A.P., de Souza Gonçalves, P., 2015. Rubber tree ortet-ramet genetic correlation and early selection efficiency to reduce rubber tree breeding cycle. Ind. Crops Prod. 77, 855–860.

Botstein, D., White, R.L., Skolnick, M., Davis, R.W., 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 32, 314–331.

Browning, S.R., Browning, B.L., 2007. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. Am. J. Hum. Genet. 81, 1084–1097. https://doi.org/10.1086/521987

Butler, D.G., Cullis, B.R., Gilmour, A.R., Gogel, B.J., 2009. Mixed models for S language environments: ASReml-R reference manual (Version 3). Queensland Department of Primary Industries and Fisheries, 398 p.

Clark, S., Hickey, J., Daetwyler, H., van der Werf, J., 2012. The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. Genet. Sel. Evol. 44, 4.

Clément-Demange, A., Priyadarshan, P.M., Thuy Hoa, T.T., Venkatachalam, P., 2007. Hevea Rubber Breeding and Genetics. In: Plant Breeding Reviews. John Wiley & Sons, Inc., p. 177–283.

da Costa. R.B., de Resende, M.D.V., de Araujo, A.J., Gonçalves, P.S., Martins, A.L.M, 2000. Genotype-environment interaction and the number of test sites for the genetic improvement of rubber trees (*Hevea*) in São Paulo State, Brazil. Genet Mol Biol 23:179–187.

804 Cros, D., Bocs, S., Riou, V., Ortega-Abboud, E., Tisné, S., Argout, X., Pomiès, V., Nodichao,
805 L., Lubis, Z., Cochard, B., Durand-Gasselin, T., 2017. Genomic preselection with
806 genotyping-by-sequencing increases performance of commercial oil palm hybrid
807 crosses. BMC Genomics 18, 839. https://doi.org/10.1186/s12864-017-4179-3
808 Cros, D., Denis, M., Sánchez, L., Cochard, B., Flori, A., Durand-Gasselin, T., Nouy, B.,
809 Omoré, A., Pomiès, V., Riou, V., Suryana, E., Bouvet, J.-M., 2015. Genomic selection
810 prediction accuracy in a perennial crop: case study of oil palm (*Elaeis guineensis*
811 Jacq.). Theor. Appl. Genet. 128, 397–410. https://doi.org/10.1007/s00122-014-2439-z
812 Cros, D., Tchounke, B., Nkague-Nkamba, L., 2018. Training genomic selection models across
813 several breeding cycles increases genetic gain in oil palm in silico study. Mol. Breed.
814 38, 89. https://doi.org/10.1007/s11032-018-0850-x
815 Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los
816 Campos, G., Burgueño, J., González-Camacho, J.M., Pérez-Elizalde, S., Beyene, Y.,
817 Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J.,
818 Varshney, R.K., 2017. Genomic Selection in Plant Breeding: Methods, Models, and
819 Perspectives. Trends Plant Sci. 22, 961–975.
820 https://doi.org/10.1016/j.tplants.2017.08.011
821 de Almeida Filho, J.E., Guimaraes, J.F.R., e Silva, F.F., de Resende, M.D.V., Munoz, P.,
822 Kirst, M., Resende Jr, M.F.R., 2016. The contribution of dominance to phenotype
823 prediction in a pine breeding and simulated population. Heredity 117, 33–41.
824 de los Campos, G., Naya, H., Gianola, D., Crossa, J., Legarra, A., Manfredi, E., Weigel, K.,
825 Cotes, J.M., 2009. Predicting Quantitative Traits With Regression Models for Dense
826 Molecular Markers and Pedigree. Genetics 182, 375–385.
827 https://doi.org/10.1534/genetics.109.101501
828 Denis, M., Bouvet, J.-M., 2013. Efficiency of genomic selection with models including
829 dominance effect in the context of *Eucalyptus* breeding. Tree Genet Genomes 9:37–
830 51. doi: 10.1007/s11295-012-0528-1
831 Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell,
832 S.E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high
833 diversity species. PLoS One 6, e19379.
834 FAOSTAT, 2018. Food and agriculture data [WWW Document]. URL
835 http://www.fao.org/faostat/en/#data/QC (accessed 2.26.18).
836 Friedrich, S., Konietschke, F., Pauly, M., 2017. GFD: An R Package for the Analysis of
837 General Factorial Designs. J. Stat. Softw. Code Snippets 79, 1–18.
838 https://doi.org/10.18637/jss.v079.c01
839 Fuentes-Utrilla, P., Goswami, C., Cottrell, J.E., Pong-Wong, R., Law, A., A'Hara, S.W., Lee,
840 S.J., Woolliams, J.A., 2017. QTL analysis and genomic selection using RADseq
841 derived markers in Sitka spruce: the potential utility of within family data. Tree Genet.
842 Genomes 13, 33. https://doi.org/10.1007/s11295-017-1118-z
843 Gianola, D., van Kaam, J.B.C.H.M., 2008. Reproducing Kernel Hilbert Spaces Regression
844 Methods for Genomic Assisted Prediction of Quantitative Traits. Genetics 178, 2289–
845 2303. https://doi.org/10.1534/genetics.107.084285
846 Gireesh, T., Meenakumari, T., Mydin, K.K., 2017. Fast track evaluation and selection of
847 Hevea brasiliensis clones from a clonal nursery. Ind. Crops Prod. 103, 195–201.
848 https://doi.org/10.1016/j.indcrop.2017.04.001
849 Gnagne, M.Y., 1988. Méthodologie de sélection de l'hévéa : évaluation des seedlings.
850 Caoutch Plast 681:113–121.
851 Gois, I., Borém, A., Cristofani-Yaly, M., de Resende, M.D., Azevedo, C.F., Bastianel, M.,
852 Novelli, V.M., Machado, M.A., 2016. Genome wide selection in *Citrus* breeding.
853 Genetics and molecular research 15(4). doi: 10.4238/gmr15048863

Gonçalves, P. de S., Silva, M. de A., Gouvêa, L.R.L., Scaloppi Junior, E.J., 2006. Genetic variability for girth growth and rubber yield in *Hevea brasiliensis*. Sci. Agric. 63, 246–254.

Gorjanc, G., Dumasy, J.-F., Gonen, S., Gaynor, R.C., Antolin, R., Hickey, J.M., 2017. Potential of Low-Coverage Genotyping-by-Sequencing and Imputation for Cost-Effective Genomic Selection in Biparental Segregating Populations. Crop Sci. 57, 1404–1420. https://doi.org/10.2135/cropsci2016.08.0675

Grattapaglia, D., 2017. Status and Perspectives of Genomic Selection in Forest Tree Breeding, in: Varshney, R.K., Roorkiwal, M., Sorrells, M.E. (Eds.), Genomic Selection for Crop Improvement: New Molecular Breeding Strategies for Crop Improvement. Springer International Publishing, Cham, pp. 199–249. https://doi.org/10.1007/978-3-319-63170-7_9

Kumar, S., Molloy, C., Muñoz, P., Daetwyler, H., Chagné, D., Volz, R., 2015. Genome-Enabled Estimates of Additive and Non-additive Genetic Variances and Prediction of Apple Phenotypes Across Environments. G3 GenesGenomesGenetics. https://doi.org/10.1534/g3.115.021105

Le Guen, V., Doaré, F., Weber, C., Seguin, M., 2009. Genetic structure of Amazonian populations of *Hevea brasiliensis* is shaped by hydrographical network and isolation by distance. Tree Genet Genomes 5:673–683. doi: 10.1007/s11295-009-0218-9

Le Guen, V., Garcia, D., Doaré, F., Mattos, C.R.R., Condina, V., Couturier, C., Chambon, A., Weber, C., Espéout, S., Seguin, M., 2011. A rubber tree's durable resistance to Microcyclus ulei is conferred by a qualitative gene and a major quantitative resistance factor. Tree Genet. Genomes 7, 877–889. https://doi.org/10.1007/s11295-011-0381-7

Le Guen, V., Garcia, D., Mattos, C.R.R., Doaré, F., Lespinasse, D., Seguin, M., 2007. Bypassing of a polygenic Microcyclus ulei resistance in rubber tree, analyzed by QTL detection. New Phytol. 173, 335–345. https://doi.org/10.1111/j.1469-8137.2006.01911.x

Lenz, P.R., Beaulieu, J., Mansfield, S.D., Clément, S., Desponts, M., Bousquet, J., 2017. Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (*Picea mariana*). BMC Genomics 18:335.

Lespinasse, D., Grivet, L., Troispoux, V., Rodier-Goud, M., Pinard, F., Seguin, M., 2000. Identification of QTLs involved in the resistance to South American leaf blight (Microcyclus ulei) in the rubber tree. Theor. Appl. Genet. 100, 975–984. https://doi.org/10.1007/s001220051379

Lin, Z., Hayes, B.J., Daetwyler, H.D., 2014. Genomic selection in crops, trees and forages: a review. Crop Pasture Sci. 65, 1177–1191.

Lorenz, A.J., Chao, S., Asoro, F.G., Heffner, E.L., Hayashi, T., Iwata, H., Smith, K.P., Sorrells, M.E., Jannink, J.-L., 2011. Genomic Selection in Plant Breeding: Knowledge and Prospects, in: Donald L. Sparks (Ed.), Advances in Agronomy. Academic Press, pp. 77–123.

Ly, D., Hamblin, M., Rabbi, I., Melaku, G., Bakare, M., Gauch, H.G., Okechukwu, R., Dixon, A.G.O., Kulakow, P., Jannink, J.-L., 2013. Relatedness and Genotype × Environment Interaction Affect Prediction Accuracies in Genomic Selection: A Study in Cassava. Crop Sci. 53, 1312–1325. https://doi.org/10.2135/cropsci2012.11.0653

Marchal, A., Legarra, A., Tisné, S., Carasco-Lacombe, C., Manez, A., Suryana, E., Omoré, A., Durand-Gasselin, T., Sánchez, L., Bouvet, J.-M., Cros, D., 2016. Multivariate genomic model improves analysis of oil palm (*Elaeis guineensis* Jacq.) progeny tests. Mol. Breed. 36, 1–13. https://doi.org/10.1007/s11032-015-0423-1

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157, 1819–1829.

Minamikawa, M.F., Nonaka, K., Kaminuma, E., Kajiya-Kanegae, H., Onogi, A., Goto, S., Yoshioka, T., Imai, A., Hamada, H., Hayashi, T., Matsumoto, S., Katayose, Y., Toyoda, A., Fujiyama, A., Nakamura, Y., Shimizu, T., Iwata, H., 2017. Genome-wide association study and genomic prediction in citrus: Potential of genomics-assisted breeding for fruit quality traits. Sci. Rep. 7, 4721. https://doi.org/10.1038/s41598-017-05100-x

Müller, B.S.F., Neves, L.G., de Almeida Filho, J.E., Resende, M.F.R., Muñoz, P.R., dos Santos, P.E.T., Filho, E.P., Kirst, M., Grattapaglia, D., 2017. Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of *Eucalyptus*. BMC Genomics 18, 524. https://doi.org/10.1186/s12864-017-3920-2

Muranty, H., Jorge, V., Bastien, C., Lepoittevin, C., Bouffier, L., Sanchez, L., 2014. Potential for marker-assisted selection for forest tree breeding: lessons from 20 years of MAS in crops. Tree Genet. Genomes 1–20. https://doi.org/10.1007/s11295-014-0790-5

Muranty, H., Troggio, M., Sadok, I.B., Rifaï, M.A., Auwerkerken, A., Banchi, E., Velasco, R., Stevanato, P., van de Weg, W.E., Di Guardo, M., Kumar, S., Laurens, F., Bink, M.C.A.M., 2015. Accuracy and responses of genomic selection on key traits in apple breeding. Hortic. Res. 2, 15060. https://doi.org/10.1038/hortres.2015.60

Pérez, P., Campos, G., 2014. Genome-wide regression and prediction with the BGLR statistical package. Genetics 198. https://doi.org/10.1534/genetics.114.164442

Pérez, P., de los Campos, G., 2013. BGLR: A Statistical Package for Whole Genome Regression and Prediction. http://R-Forge.R-project.org/projects/bglr/

Pérez-Rodríguez, P., Gianola, D., González-Camacho, J.M., Crossa, J., Manès, Y., Dreisigacker, S., 2012. Comparison Between Linear and Non-parametric Regression Models for Genome-Enabled Prediction in Wheat. G3 GenesGenomesGenetics 2, 1595–1605. https://doi.org/10.1534/g3.112.003665

Pootakham, W., Ruang-Areerate, P., Jomchai, N., Sonthirod, C., Sangsrakru, D., Yoocha, T., Theerawattanasuk, K., Nirapathpongporn, K., Romruensukharom, P., Tragoonrung, S., Tangphatsornruang, S., 2015. Construction of a high-density integrated genetic linkage map of rubber tree (*Hevea brasiliensis*) using genotyping-by-sequencing (GBS). Front. Plant Sci. 6, 367. https://doi.org/10.3389/fpls.2015.00367

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Resende, M.D.V., Resende, M.F.R., Sansaloni, C.P., Petroli, C.D., Missiaggia, A.A., Aguiar, A.M., Abad, J.M., Takahashi, E.K., Rosado, A.M., Faria, D.A., Pappas, G.J., Kilian, A., Grattapaglia, D., 2012. Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. New Phytol. 194, 116–128. https://doi.org/10.1111/j.1469-8137.2011.04038.x

Resende, R.T., Resende, M.D.V., Silva, F.F., Azevedo, C.F., Takahashi, E.K., Silva-Junior, O.B., Grattapaglia, D., 2017. Assessing the expected response to genomic selection of individuals and families in *Eucalyptus* breeding with an additive-dominant model. Heredity 119, 245–255.

Rosa, J.R.B.F., Mantello, C.C., Garcia, D., de Souza, L.M., da Silva, Carla Cristina, Gazaffi, R., da Silva, Cícero Casimiro, Toledo-Silva, G., Cubry, P., Garcia, A.A.F., de Souza, A.P., Le Guen, V., 2018. QTL detection for growth and latex production in a full-sib rubber tree population cultivated under suboptimal climate conditions. BMC Plant Biol. 18, 223. https://doi.org/10.1186/s12870-018-1450-y

Schopp, P., Müller, D., Wientjes, Y.C.J., Melchinger, A.E., 2017. Genomic Prediction Within and Across Biparental Families: Means and Variances of Prediction Accuracy and Usefulness of Deterministic Equations. G3 GenesGenomesGenetics 7, 3571. https://doi.org/10.1534/g3.117.300076

Souza, L.M., Gazaffi, R., Mantello, C.C., Silva, C.C., Garcia, D., Le Guen, V., Cardoso, S.E.A., Garcia, A.A.F., Souza, A.P., 2013. QTL mapping of growth-related traits in a full-sib family of rubber tree (Hevea brasiliensis) evaluated in a sub-tropical climate. PloS One 8, e61238–e61238. https://doi.org/10.1371/journal.pone.0061238

Tan, B., Grattapaglia, D., Martins, G.S., Ferreira, K.Z., Sundberg, B., Ingvarsson, P.K., 2017a. Evaluating the accuracy of genomic prediction of growth and wood traits in two *Eucalyptus* species and their F1 hybrids. BMC Plant Biol. 17, 110. https://doi.org/10.1186/s12870-017-1059-6

Tan, B., Grattapaglia, D., Wu, H.X., Ingvarsson, P.K., 2017b. Genomic relationships reveal significant dominance effects for growth in hybrid *Eucalyptus*. bioRxiv. https://doi.org/10.1101/178160

Tan, H., 1995. Genotype x environment interaction studies in rubber (*Hevea*) clones. J. Nat. Rubber Res. 10, 63–76.

Technow, F., Gerke, J., 2017. Parent-progeny imputation from pooled samples for cost-efficient genotyping in plant breeding. PLOS ONE 12, e0190271. https://doi.org/10.1371/journal.pone.0190271

Toro, M.A., Saura, M., Fernandez, J., Villanueva, B., 2017. Accuracy of genomic within-family selection in aquaculture breeding programmes. J. Anim. Breed. Genet. 134, 256–263. https://doi.org/10.1111/jbg.12272

Tran, D.M., Clément-Demange, A., Déon, M., Garcia, D., Le Guen, V., Clément-Vidal, A., Soumahoro, M., Masson, A., Label, P., Le, M.T., others, 2016. Genetic determinism of sensitivity to *Corynespora cassiicola* exudates in rubber tree (*Hevea brasiliensis*). PloS One 11, e0162807.

van Nocker, S., Gardiner, S.E., 2014. Breeding better cultivars, faster: applications of new technologies for the rapid deployment of superior horticultural tree crops. Hortic. Res. 1, 14022. https://doi.org/10.1038/hortres.2014.22

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth. Springer, New York, 495 p.

Viana, A.P., Resende, M.D.V. de, Riaz, S., Walker, M.A., 2016. Genome selection in fruit breeding: application to table grapes. Sci. Agric. 73, 142–149.

Warren-Thomas, E., Dolman, P.M., Edwards, D.P., 2015. Increasing Demand for Natural Rubber Necessitates a Robust Sustainability Initiative to Mitigate Impacts on Tropical Biodiversity. Conserv. Lett. 8, 230–241. https://doi.org/10.1111/conl.12170

Würschum, T., Maurer, H.P., Weissmann, S., Hahn, V., Leiser, W.L., 2017. Accuracy of within- and among-family genomic prediction in triticale. Plant Breed. 136, 230–236. https://doi.org/10.1111/pbr.12465

You, F.M., Booker, H.M., Duguid, S.D., Jia, G., Cloutier, S., 2016. Accuracy of genomic selection in biparental populations of flax (*Linum usitatissimum L.*). Crop J. 4, 290–303. https://doi.org/10.1016/j.cj.2016.03.001

Zhang, X., Sallam, A., Gao, L., Kantarski, T., Poland, J., DeHaan, L.R., Wyse, D.L., Anderson, J.A., 2016. Establishment and Optimization of Genomic Selection to Accelerate the Domestication and Improvement of Intermediate Wheatgrass. Plant Genome 9. https://doi.org/10.3835/plantgenome2015.07.0059

Tables

Table 1 Characteristics of the simple sequence repeat (SSR) molecular data obtained at each site. *Ho*: observed heterozygosity

	Site 1	Site 2
Number of SSRs	332	296
Missing data (%)	2.7%	2.2%
Range of missing data (%) per SSR	0.0% – 58.1%	0.0% – 51.0%
SSRs with ≤5% missing data (%)	87.3%	92.2%
Range of missing data (%) per clone	0.0% – 21.0%	0.0% – 32.8%
Clone with ≤5% missing data (%)	88.0%	95.1%
Mean number of alleles per SSR (range)	2.56 (2 – 4)	2.56 (2 – 4)
Total number of alleles	850	759
Mean allele frequency (range)	0.39 (0.14 – 0.86)	0.39 (0.15 – 0.84)
Mean *Ho* per SSR (range)	0.64 (0.34 – 1)	0.64 (0.33 – 1)

Figures

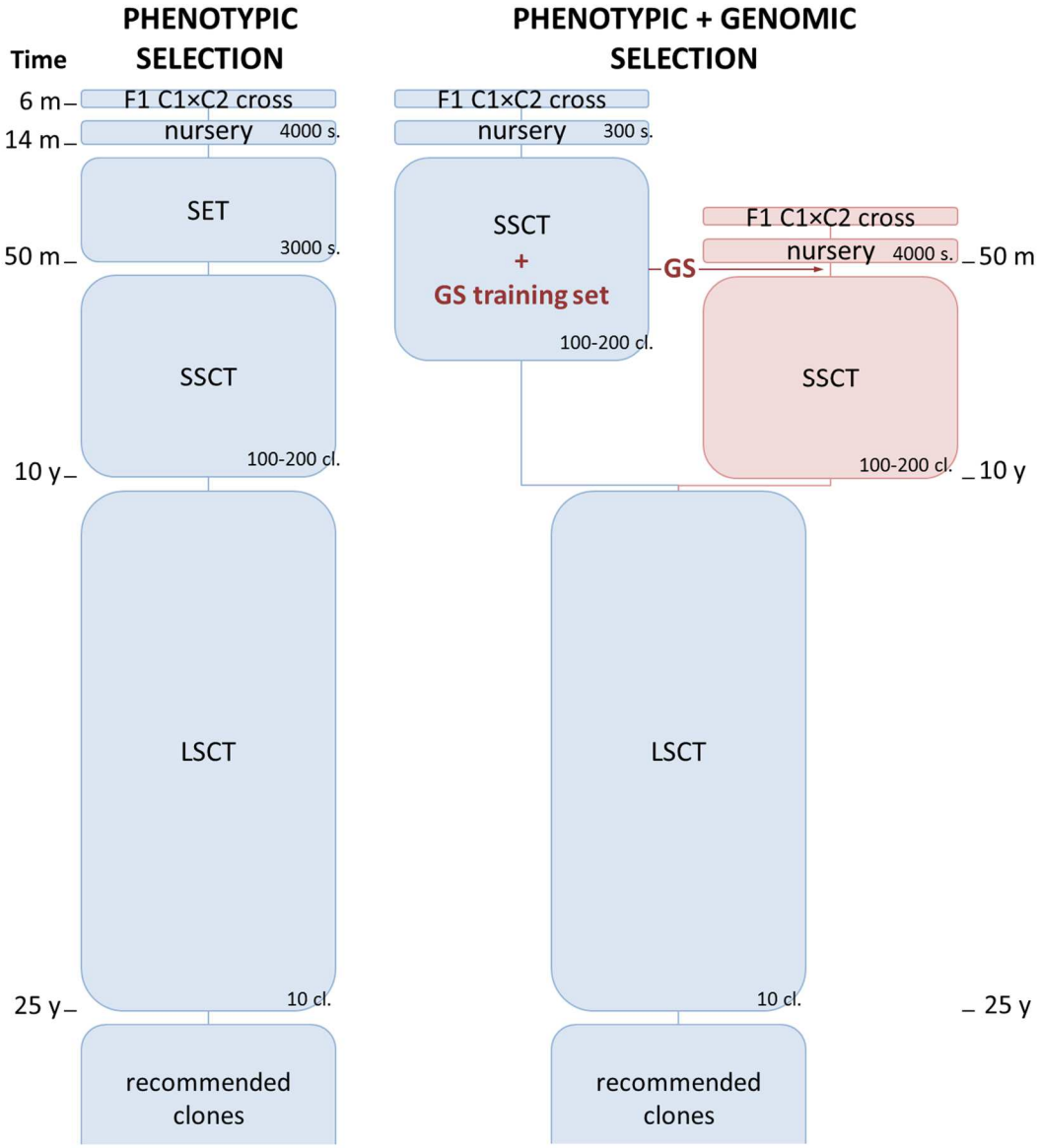
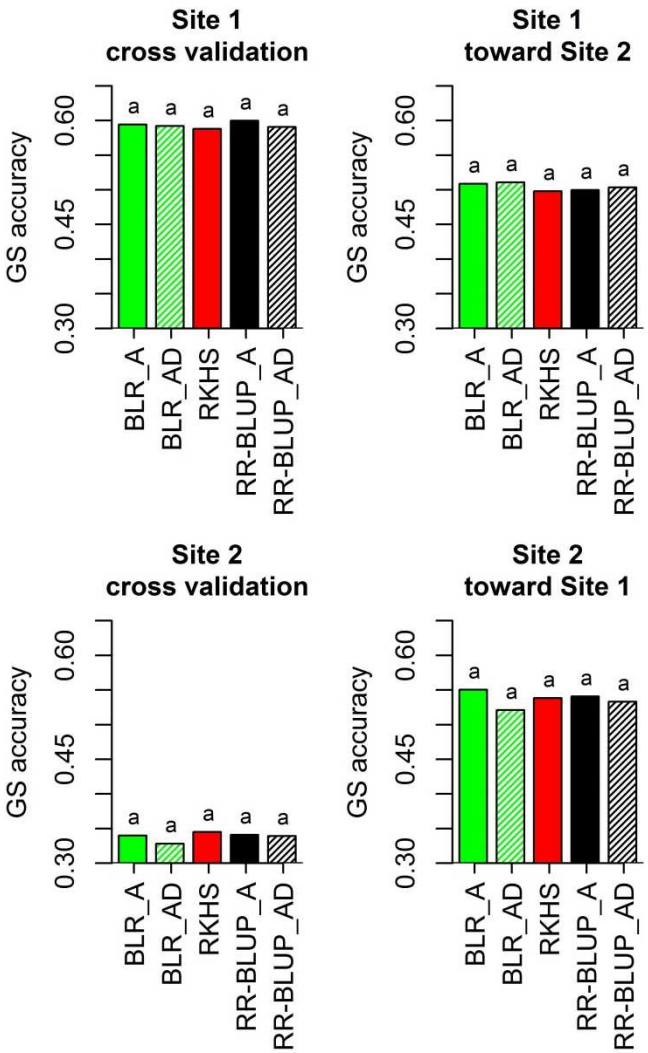


Figure 1. Conventional phenotypic selection (PS) and combined PS + genomic selection (GS) for a single F1 family (cross between C1 and C2 individuals). SET: seedling evaluation trial, SSCT: small-scale clone trial, LSCT: large-scale clone trial. The height of the boxes is proportional to duration. Blue boxes: usual steps of PS. Red: GS steps. Time is expressed in months (m) or years (y). Number of seedlings (s.) and clones (cl.) are given as an indication only.

1019



1020

1021 Figure 2. GS accuracy for rubber production according to statistical method of GS prediction, and
1022 validation approach. Values are means over seven replicates for Site 1 cross validation and Site 2 to
1023 Site 1 independent validation, and five replicates for Site 2 cross validation and Site 1 to Site 2
1024 independent validation. Values with the same letter within a given validation approach are not
1025 significantly different at P=0.05. All the clones were used to train the GS model. All the SSRs were
1026 used.

1027

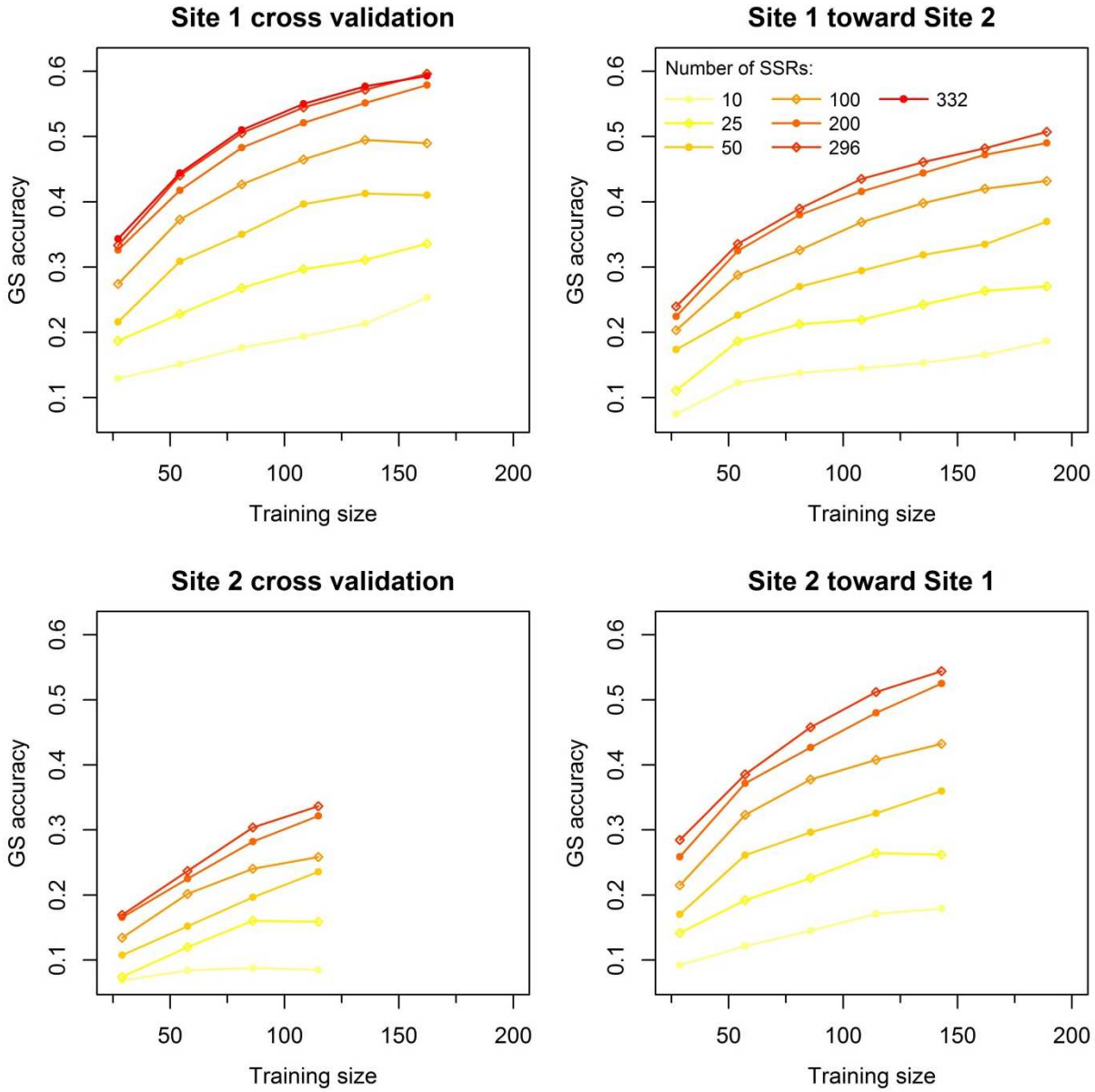


Figure 3. GS accuracy in predicting rubber yield according to number of clones used to train the GS prediction model (training size), number of SSRs, and validation approach. For a given number of SSRs, random SNPs were sampled. Values are means of seven to 1,400 replicates, depending on training size, number of SSRs, and validation approach.

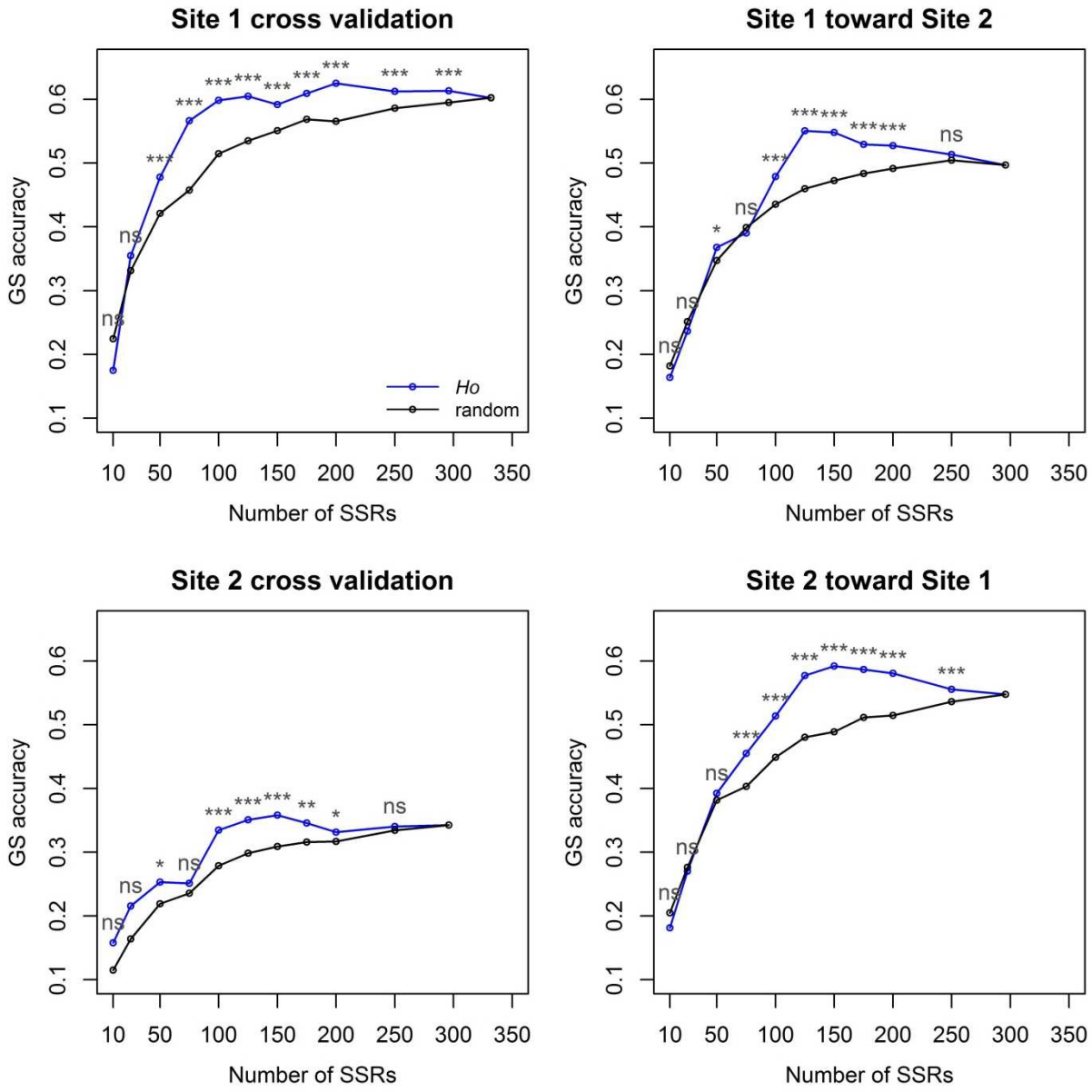


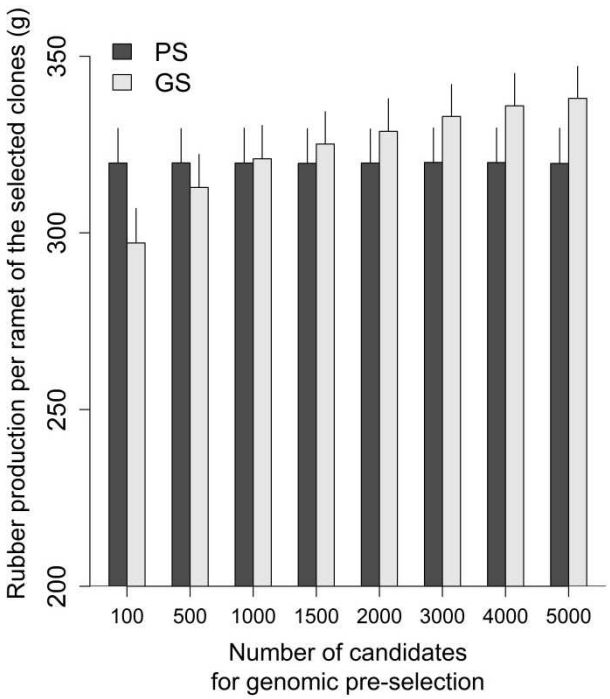
Figure 4. GS accuracy in predicting rubber yield according to SNP sampling method (highest observed heterozygosity [*Ho*] and random), number of SSRs, and validation approach. All available clones were used for training. Significance of Wald-type permutation test for method of SNP sampling: *** $P < 0.001$, * $0.01 \leq P < 0.05$, ns: not significant. Values are means of five to 56 replicates, depending on SNP sampling method, number of SSRs, and validation approach.

$$\begin{array}{c} \\ g \\ y' \\ y \\ \hat{g} \end{array} \begin{array}{c} \begin{array}{cccc} g & y' & y & \hat{g} \end{array} \\ \begin{pmatrix} 28{,}602{,}799 & 28{,}602{,}799 & 25{,}742{,}520 & 6{,}849{,}781 \\ 28{,}602{,}799 & 222{,}686{,}155 & 25{,}742{,}520 & 5{,}212{,}177 \\ 25{,}742{,}520 & 25{,}742{,}520 & 25{,}742{,}520 & 6{,}164{,}803 \\ 6{,}849{,}781 & 5{,}212{,}177 & 6{,}164{,}803 & 5{,}212{,}177 \end{pmatrix} \end{array}$$

Figure 5. Variance-covariance matrix between g, y, and $\hat{g}$ used to calibrate the simulation. g: true genetic values, y': seedling phenotypes in SET, y: estimated genetic values in SSCT, and $\hat{g}$: genomic estimated genetic values.

1050

1051



1052

1053 Figure 6. Rubber production per ramet of the clones selected using genomic selection (GS) and

1054 conventional phenotypic selection (PS) according to the number of candidate clones submitted to

1055 genomic preselection. Values are means over 5,000 replicates. Bars indicate standard deviations.

1056

1057

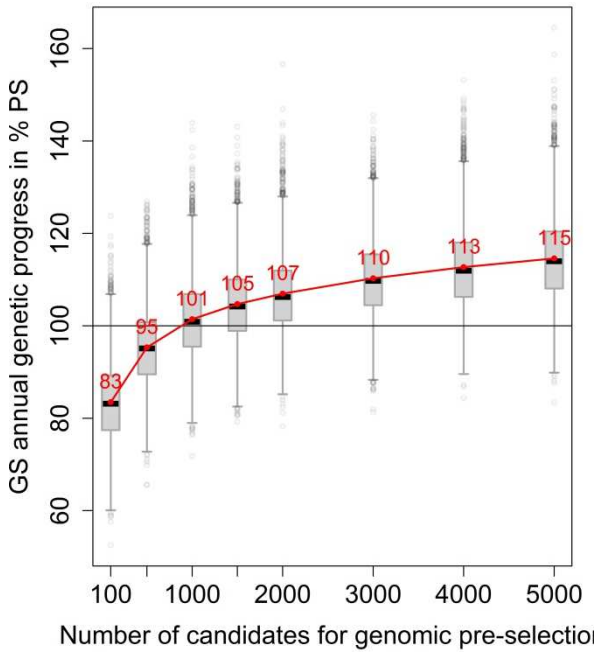


Figure 7. Annual response to selection in the GS scheme, expressed in % of the annual selection response in the conventional PS scheme, according to the number of candidates subjected to genomic preselection. Values in red are means of 5,000 replicates. The horizontal black line indicates annual selection response with GS equal to annual selection response with PS.