Euphytica (2019) 215:195
https://doi.org/10.1007/s10681-019-2525-z

Effects of parental genetic distance on offspring growth performance in *Pinus massoniana*: significance of parental-selection in a clonal seed orchard

ZhouXian Ni · TianDao Bai · YaBin Chen · YiFa Huang · Li-An Xu

Received: 24 May 2019 / Accepted: 31 October 2019 / Published online: 11 November 2019

Abstract *Pinus massoniana* is the most widely cultivated tree in southern China and an important afforestation, timber species. Seed orchards are the main source of high-quality seeds of *P. massoniana,* and parental selection is key to seed orchard construction. In this study, the relationships among genetic-quality stability, growth-performance and parental genetic-distance (GD) were analyzed based on 18 shared half-sib families from 3 progeny tests (established in 1995, 1996 and 1997) of *P. massoniana* seed orchard. Based on the growth stability of the offspring among different years, the stable groups (SGs) (9 families with the lowest growth fluctuation) and unstable groups (USGs) (9 families with the greatest growth fluctuation) were selected. Using 27 simple sequence repeat markers, we identified the parental-combinations for the most individuals in SGs and USGs from 18 known females and 217 candidate males. The parental GDs in the SGs changed less than those in USGs. Furthermore, based on the growth-performance of offspring, the excellent group (EG) (4 families in the top level), middle group (MG) (4 families in the medium level) and worst group (WG) (4 families in the last level) were selected. Compared with the MG and WG, the EG contained more parental-combinations with GD > 0.6, indicating this group could produce more high-quality variation. Hence, parents with a small range of GDs should be selected to ensure the sustainable and stable production of seeds with excellent genetic quality in the construction of *P. massoniana* seed orchards, while hybrid parents with greater GDs should be selected in the breeding process of *P. massoniana* to obtain more offspring with excellent genetic quality.

Keywords *Pinus massoniana* · Parental genetic distance · Progeny performance · SSR markers · Breeding

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10681-019-2525-z) contains supplementary material, which is available to authorized users.

Z. Ni · T. Bai · L.-A. Xu (✉)
Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China
e-mail: laxu@njfu.edu.cn

T. Bai
Forestry College, Guangxi University, Nanning 530004, China

Y. Chen · Y. Huang
Wuyi National Forest Farm Fujian Province, Zhangping 364400, China

Introduction

Masson pine (*Pinus massoniana*) is an important afforestation and timber species, particularly in China,

where it is widely distributed in all regions south of the Yangtze River Basin. Additionally, this species occupies the most extensive afforestation area in southern China. The breeding of Masson pine has always been a key emphasis in the work of Chinese tree breeders. Seed orchards are the main source of excellent seeds of *P. massoniana* and are thus established with the purpose of sustained and efficient production of seeds with high genetic quality. Parental selection is an important process that should not be overlooked during the construction of seed orchards. This process can be approached in two ways, namely, a priori and a posteriori selection. A posteriori selection is always used to assess the genetic quality of parents based on progeny performance, but it is a long and tedious process for perennial woody plants. By contrast, a priori selection based on parent–offspring relationships may simplify and shorten the process of parental selection (Dias et al. 2004).

Based on molecular data, some previous studies have observed significant positive correlations between parental genetic distance (GD) and the growth performance of hybrid offspring (Tian et al. 2016; Zhang et al. 2013). Furthermore, these correlations are known to differ among plant species, but few studies on tree species have been reported. Although relationships between parental GD and offspring performance have been reported in trees such as *Liriodendron* (Yao et al. 2015) and *Eucalyptus grandis* (Abad et al. 2005), nonsignificant or weak correlations were observed between parental GD and offspring performance because of the limited number and type of molecular markers used. In some previous studies, random amplified polymorphic DNA (RAPD) (Krystkowiak et al. 2009; Kuczyńska et al. 2007), amplified fragment length polymorphism (AFLP) (Legesse et al. 2008) and sequence-related amplified polymorphism (SRAP) (Tian et al. 2016) markers were used to evaluate GDs and predict the performance of offspring. With the development of molecular markers and high-throughput sequencing technology, simple sequence repeat (SSR) (Yao et al. 2015; Xu et al. 2017; Li et al. 2018) and single nucleotide polymorphism (SNP) (Luo et al. 2016; Ndhlela et al. 2015) markers have gradually become mainstream tools used in related research. SSR markers in particular are widely used in such studies due to their advantages of codominance, high polymorphism and low cost (Xing et al. 2014; Wegary et al. 2013; Yao et al. 2015; Grattapaglia et al. 2014).

Compared with crops, trees have a longer life history and breeding cycle, making the evaluation of offspring performance in the process of tree breeding extremely time consuming. Additionally, a shorter receptive period and high rainfall in pollination-time, cone maturation in late autumn next year make artificial controlled pollination of *P. massoniana* difficult. Moreover, manual controlled pollination of *P. massoniana* is cumbersome and numerous. All the factors mentioned above limit studies on the relationships between parental GD and offspring performance through artificially controlled hybridization. Hence, using half-sib families of *P. massoniana* is particularly important in related research. However, what effect does the parental combination have on an offspring's performance? Is the genetic quality of seeds in seed orchards consistent among years? What are the main factors that affect the stability of seed genetic quality among years? This study aimed to investigate the relationships of parental GD with the stability of seed genetic quality among years and with offspring performance in open-pollination (OP) seed orchards of *P. massoniana* to answer the questions presented above and to provide guidance for parental selection during the construction of seed orchards.

Materials and methods

Parents, progeny and field trials

The clonal seed orchard of *P. massoniana* was built in 1985 with rootstocks and is located in the Wuyi Forest Farm in Zhangping, Fujiang Province (24° 58.072′ N, 117° 27.509′ E). A total of 218 clones were grafted in 1987, with a planting density of 4 m × 4 m. The parental clones originated from some provinces in which the species is naturally distributed in China. OP seeds were harvested in 1993, 1994 and 1995, and the progeny tests were established in 1995, 1996 and 1997. Seedlings of half-sib progeny were planted in a randomized complete block design with 8 replicates and 5 seedlings in each plot. The growth traits [including height (H) and diameter at breast height (DBH)] were measured when the progeny were 14 years old. The wood volume (V) was estimated with the following formula (Zhang et al. 2013):

$V = 0.000062341803 \times \mathrm{DBH}^{1.8561497} \times \mathrm{H}^{0.9568492}$.

Selection of half-sib families

A total of 72 shared half-sib families were screened from the 3 progeny tests (2475 individuals). Standard deviations (SDs) of V were used to measure the growth stability of shared half-sib families across the 3 progeny tests. The 9 half-sib families with the lowest SD values were screened as the stable groups (SGs), and the 9 half-sib families with the highest SD values were screened as the unstable groups (USGs) (Table 1). There were 369 individuals of SGs and USGs used in paternity analysis and other further analysis. In addition, all progeny screened as described above were also divided into three groups based on a descending ranking of average V [the top-four half-sib families with the highest values as the excellent group (EG), the last-four half-sib families with the lowest values as the worst group (WG), and the medianfour half-sib families with intermediate values as the middle group (MG)] (Table 1). It should be noted that three less differentiated families were respectively removed between two groups of EG, MG and WG (EG–MG and MG–WG) in order to show significant differences among groups.

Sample collection and PCR amplification

All needles were collected in 2014 and stored at − 70 °C. Total genomic DNA was extracted from fresh leaves using a Plant Genomic DNA Kit (ZomanBio Inc., Beijing, China) and quantified with a NanoDrop 2000c spectrophotometer (ThermoFisher Scientific, Waltham, USA). A total of 27 polymorphic SSR markers were used in this study (Supplementary Table 1) (Bai et al. 2014). SSR-PCR was performed in a 10 μL reaction mixture containing 50 ng of DNA template, 1.0 μL of 10 × PCR buffer (10 mM Tris–HCl), 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 0.3 μM primer, and 0.5 U of Taq polymerase (Takara Bio Inc., Dalian). The PCR program was as follows: an initial denaturation at 94 °C for 5 min; 25 cycles of 30 s at 94 °C, 30 s at the annealing temperature, and 40 s at 72 °C; and a final extension at 72 °C for 5 min. PCR products were tested by using 8% polyacrylamide gels for 2.5 h at 150 V in a vertical gel electrophoresis apparatus (Liuyi Co., Ltd., Beijing). PCR products of 5 SSR markers were measured using a Qsep100 capillary gel electrophoresis system (BiOptic Inc., Taipei).

Data analysis

The genetic parameters, including the number of alleles (Na) per locus, Nei's gene diversity (h) and the polymorphism information content (PIC), were estimated with POPGENE 32 version 1.32 (Yeh 1999) and PowerMarker version 3.25 (Liu and Muse 2005). Nei's genetic distance of parents were estimated using PowerMarker version 3.25 (Liu and Muse 2005; Nei 1978). A total of 369 individuals of SGs and USGs were used for paternity analysis with 18 known females and 217 candidate males using the revised likelihood-based paternity inference program CERVUS 3.0.7 (Kalinowski et al. 2007) by using known females and 217 candidate males. Pearson's correlation coefficients between parental GD and the V of progeny were calculated using a custom script in R 3.4.0. The ggplot2 package in R was used to construct a scatter diagram and other plots (Wickham 2009).

Table 1 Half-sib families in groups based on different classification bases

Classification basis	Groups	Half-sib family IDs
Growth stability between years	Stable groups (SGs)	133, 277, 326, 335, 350, 404, 579, 589, 688
	Unstable groups (USGs)	1, 11, 53, 329, 393, 545, 686, 769, 773
Ranking of average values of wood volumes	Excellent groups (EG)	11, 326, 335, 350
	Middle groups (MG)	393, 404, 545, 769
	Worst groups (WG)	53, 277, 589, 686

Results

Genetic diversities of parents and progeny

In this study, 27 SSRs were used to determine the genetic diversity of parents and progeny. A total of 164 alleles were detected with SSR markers (2–18 alleles per locus), resulting in 5.86 alleles per locus on average. In addition, the PIC of polymorphic SSRs ranged from 0.269 to 0.869. Botstein et al. (1980) noted that codominant loci may be highly informative (PIC > 0.5), reasonably informative (0.5 > PIC > 0.25) and slightly informative (PIC < 0.25) (Botstein et al. 1980). Most of these polymorphic SSRs were highly or reasonably informative and could be useful in the study of genetic diversity or in parentage analysis. The primers PMa25a (10 alleles), PMa25b (12 alleles), PMa51 (14 alleles), PMa72 (18 alleles), PMa96 (12 alleles) and WX005 (14 alleles) amplified more alleles than the others because the capillary gel electrophoresis system was used for these markers (PMa25a and PMa25b are two loci of the primer P.Ma25) (Table 2).

High genetic diversity is an important genetic prerequisite for seed breeding that can produce more genetic variation. In this study, an analysis of genetic diversity among parents and progeny (1995, 1996 and 1997) in orchards was performed (Table 3). The Shannon's information index of genetic diversity among parents and for the progeny tests (1995, 1996 and 1997) was 1.1538, 1.1665, 1.1777 and 1.1683, respectively, and the corresponding values of Nei's gene diversity were 0.6027, 0.6011, 0.6079 and 0.6051, respectively. All these parameters of genetic diversity indicated that parents and progeny had high genetic diversity and were not significantly different.

Table 2 The polymorphism of SSR markers used in this study

Primer ID	Na	PIC	Primer ID	Na	PIC
PMa19	4	0.335	PMa96	12	0.777
PMa22	4	0.474	PMa116	4	0.59
PMa25a	10	0.765	PMa117	3	0.53
PMa25b	12	0.797	PMa216	4	0.575
PMa27	3	0.405	PMa221	4	0.425
PMa32	2	0.375	PMa229	6	0.590
PMa43	3	0.370	WX005	14	0.869
PMa45	3	0.476	WX013	7	0.626
PMa51	14	0.869	WX122	4	0.408
PMa65	3	0.373	WX127	3	0.265
PMa66	3	0.498	WX135	3	0.447
PMa72	18	0.852	WX212	3	0.372
PMa80	4	0.523	WX277	6	0.318
PMa86	3	0.286			
PMa95	5	0.554	Mean	5.86	0.527

Na number of alleles, *PIC* polymorphism information content; PMa25a and PMa25b are two loci of PMa25

Paternity analyses of samples

To investigate the parents of individuals in the SGs and USGs in the 3 progeny tests (1995, 1996 and 1997), parental analysis was performed with 217 candidate parents based on 27 pairs of SSR primers. Male parents of progeny in different groups were analyzed statistically under confidence levels of 95% and 80% (Table 4). The progeny of the SGs and USGs had different determination rates ranging from 0.1892 to 0.2429 (confidence level = 95%), while the determination rates were mainly in the range of 0.5405–0.8154 under a confidence level of 80%.

Distributions of parental genetic distances among groups

The purpose of seed orchards is to produce seeds with high genetic quality, and parents have important effects on the genetic quality of seeds. To investigate the relationships between parents and the stability of seed genetic quality among years, the relationships between parental GD and the stability of progeny performance among 1995, 1996 and 1997 was assessed by using parents determined in the parental analysis (confidence level = 80%).

In this study, we compared the distributions of parental GDs among groups in different years to assess relationships between parents and the stability of seed genetic quality among years (Fig. 1a). A significant difference in parental GD was found between the SGs and USGs ($p < 0.001$), and the average parental GD of the SG (mean = 0.4251) was much lower than that of the USGs (mean = 0.6019). Compared with the USGs (the proportion of parents (PPs) = 0.271 ~ 0.358), the SGs had more parent pairs (PPs = 0.702 ~ 0.750) for which the GD was less than 0.5, while the SGs had fewer parent pairs (t PPs = 0.191 ~ 0.225) than the

Table 3 The genetic diversities of parents and progeny in 1995, 1996 and 1997

Populations	Na	Ne	I	Obs_Het	Exp_Het	Nei
Parents	5.75	3.0014	1.1538	0.4253	0.6042	0.6027
95 progeny	6.10	3.0952	1.1665	0.4638	0.6032	0.6011
96 progeny	5.70	3.2173	1.1777	0.4586	0.6103	0.6079
97 progeny	5.62	3.1448	1.1683	0.4594	0.6044	0.6051

Na observed number of alleles, *Ne* effective number of alleles, *I* Shannon's information index, *Obs_Het* observed heterozygosity, *Exp_Het* expected heterozygosity, *Nei* Nei's (1973) gene diversity

Table 4 Paternity analyses of different groups of progeny in a *Pinus massoniana* seed orchard

Groups	Total numbers of individual	Individual numbers (95%)	Proportion in group (95%)	Individual numbers (80%)	Proportion in group (80%)
95SG	63	14	0.2222	47	0.7460
95USG	86	17	0.1977	59	0.6860
96SG	70	17	0.2429	54	0.7714
96USG	65	15	0.2308	53	0.8154
97SG	74	14	0.1892	40	0.5405
97USG	76	17	0.2237	53	0.6974

Groups are stable groups and unstable groups in 1995, 1996 and 1997; 95% and 80% are confidence levels in the paternity analyses

USGs (PPs = 0.396 $\sim$ 0.492) for which the GD was more than 0.6. All of these factors indicated that the distribution of parental GDs in the SGs were narrow, while that in the USGs were more dispersed.

Additionally, the relationships between parental GDs and progeny performance in terms of V were evaluated among the EG, MG and WG. The average GDs among groups were as follows: WG (0.4634) < MG (0.4993) < EG (0.5613). Additionally, the proportion of parent pairs for which the GD was less than 0.5 was highest in the WG (PPs = 0.639), followed by the MG (PPs = 0.549) and the EG (PPs = 0.486), while the proportion of parent pairs for which the GD was greater than 0.6 was larger in the EG (PPs = 0.419) than in the MG (PPs = 0.232) and WG (PPs = 0.301) (Fig. 1b). In total, the EG contained greater and more discrete parental GDs than the WG. Hence, use of the EG makes it possible to produce more genetic variation. A further comparison of the progeny performance in terms of V between the EG and WG was made (Fig. 2). The EG had superior performance in terms of V than the WG, and the maximum V was found in the EG. In summary, parents with distant GDs in the EG can produce more high-quality variation, and their offspring are better able to grow.

Correlations between parental genetic distance and progeny growth

Positive and weak correlations were detected between parental GD and progeny performance (H, DBH and V), and the correlation coefficients ranged from 0.146 to 0.225. Parental GD was significantly correlated with V and DBH ($p < 0.05$ or $p < 0.01$) among ranges of parental GDs, while no significant correlations with H were detected among most ranges of parental GDs (Table 5). Further regression analysis indicated that differentiation was also more extreme (pros and cons of coexistence) when the parental GDs were greater (especially when the GD > 0.6) (Fig. 3). The parental GDs of offspring with maximum values of V (0.4592 m^3) and DBH (0.276 m) were greater than 0.6. Hence, we should focus on choosing parents with greater GDs, which are likely to produce offspring with excellent performance.

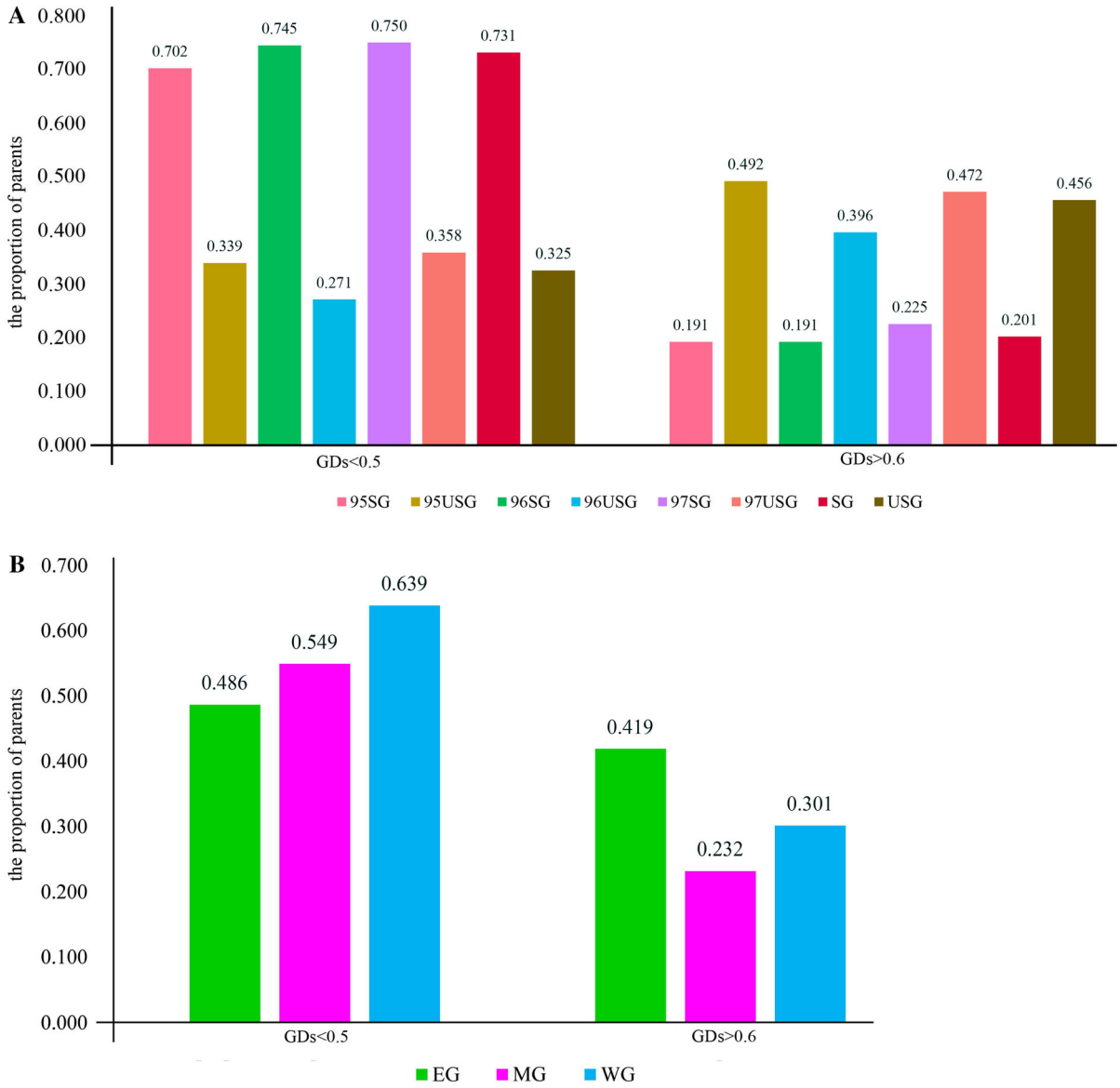

Fig. 1 Distribution of parental genetic distances between groups. **a** Distribution of parental genetic distances between the stable and unstable groups; **b** distribution of parental genetic distances among the best, middle and worst groups

Discussion

In the process of plant breeding, parental selection plays an important role in producing offspring with excellent performance (Xing et al. 2014; Hallingbäck and Jansson 2013). The relationships between parental GD and progeny performance have always been a key issue in cross-breeding (Legesse et al. 2008; Sang et al. 2015; Xing et al. 2014), and correlations have been detected in some plants (Tambasco-Talhari et al. 2005; Zhang et al. 2013; Hung et al. 2012; Wegary et al. 2013). Undoubtedly, significant and strong correlations between parental GD and progeny performance can provide guidance in predicting the growth performance of offspring (D'Andrea et al. 2013). Although a strong correlation between parental GD and progeny performance has been reported in some plants, such as *Salix babylonica* (Teklewold and Becker 2006), *Triticum aestivum* (Krystkowiak et al. 2009), and *Zea mays* (Makumbi et al. 2011), weak and

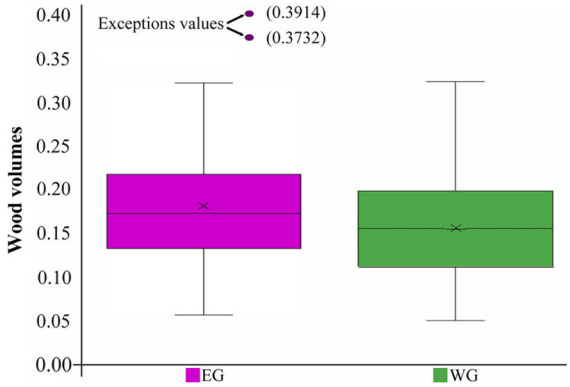

Fig. 2 Distribution of progeny wood volumes between the best groups and worst groups. *EG* excellent groups, *WG* worst groups

Table 5 Correlations between parental genetic distance and progeny performance

Ranges of genetic distance	Volume	Diameter	Height
0.20–1.00	0.146*	0.131*	0.165*
0.25–1.00	0.175**	0.162*	0.141*
0.30–1.00	0.169*	0.153*	0.135*
0.30–0.90	0.191**	0.180**	0.095
0.30–0.80	0.208**	0.211**	0.052
0.30–0.70	0.225**	0.222**	0.122
0.30–0.65	0.219**	0.228**	0.100

**Represents $p < 0.01$; *represents $p < 0.05$

nonsignificant correlations have also been found in many previous studies (Yao et al. 2015; Luo et al. 2016; Tian et al. 2016). In this study, significant but weak correlations were found between parental GD and the V and DBH of progeny, which was consistent with the results of previous studies (Luo et al. 2016; Ndhlela et al. 2015). However, weak correlations are not powerful in predicting the performance of progeny for two reasons: (1) the molecular markers used in studies are not associated with genes or loci controlling quantitative traits; and (2) genetic and environmental interactions can also result in weak correlations (Yao et al. 2015; Tian et al. 2016). Hence, new polymorphic markers (SNPs) (Sang et al. 2015; Luo et al. 2016; Ndhlela et al. 2015), quantitative trait locus (QTL) mapping (Marcotuli et al. 2017; Cai et al. 2017) and genome-wide association studies (GWAS) (Huang et al. 2012; Weng et al. 2011) may be powerful tools in predicting the correlations between parental GD and progeny performance. Although no significant and strong correlation were detected between parental GD and progeny performance (H, DBH and V), the distributions of parental GDs have influence on the stability of progeny performance among years. The parents with narrow distributions of parental GDs have more stable seed genetic quality among years than those parents with dispersed distributions of parental GDs in seed orchard. Moreover, more variation in progeny performance was also found when parental GD was greater than 0.6. And the parental GDs of the offspring with the maximum values of V and DBH were more than 0.6. Hence, we can assume that the individuals with excellent performance exhibited parental GDs greater than 0.6, which is similar to the result obtained in a study of *Liriodendron* (Yao et al. 2015).

The main purpose of seed orchards is to continuously and efficiently produce seeds with excellent genetic qualities. However, can the seeds maintain high genetic quality? How can fluctuations in the genetic quality of seeds across years be avoided? Answering these questions will require further studies. In this study, we investigated the relationships between parental GD and the stability of seed genetic quality in *P. massoniana* across years using 27 SSRs. The distributions of parental GDs of the USGs were

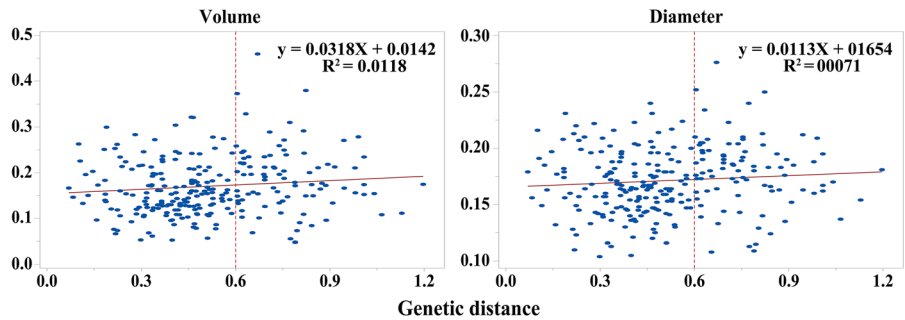

Fig. 3 Linear correlation between parental genetic distance and progeny performance

more discrete than those of the SGs, which indicated that more discrete distributions of parental GDs may result in fluctuations in seed genetic quality. Relationships between parental GD and the stability of seed genetic quality were not found in previous studies that focused on the correlation between parental GD and combining ability (specific combination ability (SCA) and general combining ability (GCA)) or heterosis (Tian et al. 2016; Luo et al. 2016; Yao et al. 2015).

Conclusions

In this study, significant and weak correlations were detected between parental GD and the growth performance of offspring (V and DBH), which are not sufficient to predict the performance of progeny (Ndhlela et al. 2015). However, we also found that a narrow distribution of parental GDs contributed to the sustainable production of seeds with excellent genetic quality. Additionally, we found some progeny with excellent performance when parental GDs were greater than 0.6. Hence, we should focus on the selection of parent pairs with a narrow range of GDs to ensure the sustainable and stable production of *P. massoniana* seeds with excellent genetic quality. However, in the breeding process, parental combinations with a high GD should be selected as hybrid parents to obtain offspring with excellent genetic quality.

Acknowledgements This work was supported by the Science and Technology Major Project of Guangxi (Project No. AA17204087-1), National Natural Science Foundation of China (No. 31400575) and the Program Development of Jiangsu Higher Education Institutions (PAPD).The authors were also grateful to the anonymous reviewers and editors for their constructive comments.

References

Abad J, Rocha R, Cruz CD, De Araújo EF (2005) Obtainment of *Eucalyptus* spp. hybrids aided by molecular markers—SSR analysis. Sci For 67:53–63

Bai TD, Xu LA, Xu M, Wang ZR (2014) Characterization of masson pine (*Pinus massoniana* Lamb.) microsatellite DNA by 454 genome shotgun sequencing. Tree Genet Genom 10:429–437. https://doi.org/10.1007/s11295-013-0684-y

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Cai L, Voorrips RE, van de Weg E, Peace C, Iezzoni A (2017) Genetic structure of a QTL hotspot on chromosome 2 in sweet cherry indicates positive selection for favorable haplotypes. Mol Breed 37:85. https://doi.org/10.1007/s11032-017-0689-6

D'Andrea K, Otegui ME, Cirilo AG, Eyhérabide GH (2013) Parent–progeny relationships between maize inbreds and hybrids: analysis of grain yield and its determinants for contrasting soil nitrogen conditions. Crop Sci 53:2147–2161. https://doi.org/10.2135/cropsci2013.02.0111

dos Santos Dias LA, de Toledo Picoli EA, Rocha RB, Alfenas AC (2004) A priori choice of hybrid parents in plants. Gen Mol Res 3:356–368

Grattapaglia D, do Amaral Diener PS, 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 Genom 10:1631–1643. https://doi.org/10.1007/s11295-014-0784-3

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

Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, Li W, Guo Y, Deng L, Zhu C, Fan D, Lu Y, Weng Q, Liu K, Zhou T, Jing Y, Si L, Dong G, Huang T, Lu T, Feng Q, Qian Q, Li J, Han B (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. Nat Genet 44:32–39. https://doi.org/10.1038/ng.1018

Hung H, Browne C, Guill K, Coles N, Eller M, Garcia A, Lepak N, Melia-Hancock S, Oropeza-Rosas M, Salvo S (2012) The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. Heredity 108:490–499. https://doi.org/10.1038/hdy.2011.103

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. https://doi.org/10.1111/j.1365-294x.2007.03089.x

Krystkowiak K, Adamski T, Surma M, Kaczmarek Z (2009) Relationship between phenotypic and genetic diversity of parental genotypes and the specific combining ability and heterosis effects in wheat (*Triticum aestivum* L.). Euphytica 165:419–434. https://doi.org/10.1007/s10681-008-9761-y

Kuczyńska A, Surma M, Kaczmarek Z, Adamski T (2007) Relationship between phenotypic and genetic diversity of parental genotypes and the frequency of transgression effects in barley (*Hordeum vulgare* L.). Plant Breed 126:361–368. https://doi.org/10.1111/j.1439-0523.2007.01367.x

Legesse BW, Myburg AA, Pixley KV, Twumasi-Afriyie S, Botha AM (2008) Relationship between hybrid performance and AFLP based genetic distance in highland maize

inbred lines. Euphytica 162:313–323. https://doi.org/10.1007/s10681-007-9503-6

Li Z, Zhong Y, Yu F, Xu M (2018) Novel SSR marker development and genetic diversity analysis of *Cinnamomum camphora* based on transcriptome sequencing. Plant Genet Resour Charact Util 16:568–571. https://doi.org/10.1017/S147926211800014X

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129. https://doi.org/10.1093/bioinformatics/bti282

Luo X, Ma C, Yi B, Tu J, Shen J, Fu T (2016) Genetic distance revealed by genomic single nucleotide polymorphisms and their relationships with harvest index heterotic traits in rapeseed (*Brassica napus* L.). Euphytica 209:41–47. https://doi.org/10.1007/s10681-015-1629-3

Makumbi D, Betrán JF, Bänziger M, Ribaut J-M (2011) Combining ability, heterosis and genetic diversity in tropical maize (*Zea mays* L.) under stress and non-stress conditions. Euphytica 180:143–162. https://doi.org/10.1007/s10681-010-0334-5

Marcotuli I, Gadaleta A, Mangini G, Signorile AM, Zacheo SA, Blanco A, Simeone R, Colasuonno P (2017) Development of a high-density SNP-based linkage map and detection of QTL for β-glucans, protein content, grain yield per spike and heading time in durum wheat. Int J Mol Sci 18:1329–1344. https://doi.org/10.3390/ijms18061329

Ndhlela T, Herselman L, Semagn K, Magorokosho C, Mutimaamba C, Labuschagne MT (2015) Relationships between heterosis, genetic distances and specific combining ability among CIMMYT and Zimbabwe developed maize inbred lines under stress and optimal conditions. Euphytica 204:635–647. https://doi.org/10.1007/s10681-015-1353-z

Nei M (1978) The theory of genetic distance and evolution of human races. Jpn J Hum Genet 23:341–369. https://doi.org/10.1007/bf01908190

Sang S, Wang H, Mei DS, Liu J, Fu L, Wang J, Wang WX, Hu Q (2015) Correlation analysis between heterosis and genetic distance evaluated by genome-wide SNP Chip in *Brassica napus*. Sci Agric Sin 48:2469–2478. https://doi.org/10.3864/j.issn.0578-1752.2015.12.020

Tambasco-Talhari D, de Alencar MM, de Paz CCP, da Cruz GM, Rodrigues AD, Packer IU, Coutinho LL, de Almeida Regitano LC (2005) Molecular marker heterozygosities and genetic distances as correlates of production traits in F1 bovine crosses. Genet Mol Biol 28:218–224. https://doi.org/10.1590/s1415-47572005000200007

Teklewold A, Becker HC (2006) Comparison of phenotypic and molecular distances to predict heterosis and F1 performance in Ethiopian mustard (*Brassica carinata* A. Braun). Theor Appl Genet 112:752–759. https://doi.org/10.1007/s00122-005-0180-3

Tian HY, Channa SA, Hu SW (2016) Relationships between genetic distance, combining ability and heterosis in rapeseed (*Brassica napus* L.). Euphytica 213:1–11. https://doi.org/10.1007/s10681-016-1788-x

Wegary D, Vivek B, Labuschagne M (2013) Association of parental genetic distance with heterosis and specific combining ability in quality protein maize. Euphytica 191:205–216. https://doi.org/10.1007/s10681-012-0757-2

Weng J, Xie C, Hao Z, Wang J, Liu C, Li M, Zhang D, Bai L, Zhang S, Li X (2011) Genome-wide association study identifies candidate genes that affect plant height in Chinese elite maize (*Zea mays* L.) inbred lines. PLoS ONE 6:e29229. https://doi.org/10.1371/journal.pone.0029229

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Xing N, Fan C, Zhou Y (2014) Parental selection of hybrid breeding based on maternal and paternal inheritance of traits in rapeseed (*Brassica napus* L.). PLoS ONE 9:e103165. https://doi.org/10.1371/journal.pone.0103165

Xu M, Liu X, Wang J, Teng S, Shi J, Li Y, Huang M (2017) Transcriptome sequencing and development of novel genic SSR markers for *Dendrobium officinale*. Mol Breed 37(2):18–24. https://doi.org/10.1007/s11032-016-0613-5

Yao J, Li H, Ye J, Shi L (2015) Relationship between parental genetic distance and offspring's heterosis for early growth traits in *Liriodendron*: implication for parent pair selection in cross breeding. New For 47:163–177. https://doi.org/10.1007/s11056-015-9508-2

Yeh FC (1999) POPGENE: Microsoft Window-based freeware for population genetic analysis, version 1.31

Zhang Y, Yang Q, Zhou Z, Jin G (2013) Divergence among Masson pine parents revealed by geographical origins and SSR markers and their relationships with progeny performance. New For 44:341–355. https://doi.org/10.1007/s11056-012-9340-x