SUPPLEMENTARY MATERIAL 1

Table S1 Results from testing hypothesis about the adequacy of alternative genetic models fitted sequentially to breast-height diameter (DBH) growth, percentage of tree survival, percentage of alive trees that are reproductive, and capsule score, measured in the Tyenna trial at age 20 years from field planting. For a given trait, the most parsimonious genetic model that led to the largest reduction in significance of the lack-of-fit term (and also of the χ^2 statistic under the joint-scaling test procedure), to a value not significant at the 10% level was selected as the best model to explain the genetic basis of species divergence (Table 2 main paper). However, this criterion of model adequacy could not be achieved for percentage of tree survival, as there was still statistically significant lack of fit regardless of which digenic epistatic combinations were fitted.

Genetic models	DBH growth		Percentage of tree survival		Percentage of alive trees that are reproductive		Capsule score			
	F	χ^2	*F*	χ^2	*F*	χ^2	*F*	χ^2		
1	6.04 (*P*=0.007)	19.81 (*P*≤0.001)	16.70 (*P*≤0.001)	67.47 (*P*≤0.001)	6.24 (*P*=0.007)	22.62 *P*≤0.001)	14.47 (*P*≤0.001)	46.66 (*P*≤0.001)		
2.1	0.10 (*P*=0.906)	0.21 (*P*=0.900)	4.78 (*P*=0.026)	10.73 (*P*=0.005)	9.06 (*P*=0.003)	19.87 *P*≤0.001)	10.17 (*P*=0.002)	17.70 (*P*≤0.001)		
2.2	9.33 (*P*=0.003)	19.66 (*P*≤0.001)	24.24 (*P*≤0.001)	60.65 (*P*≤0.001)	2.54 (*P*=0.120)	5.45 (*P*=0.066)	19.38 (*P*≤0.001)	78.17 (*P*≤0.001)		
2.3	1.87 (*P*=0.190)	3.96 (*P*=0.138)	18.32 (*P*≤0.001)	44.28 (*P*≤0.001)	9.50 (*P*=0.004)	22.57 *P*≤0.001)	11.37 (*P*=0.002)	27.15 (*P*≤0.001)		
3.1	0.05 (*P*=0.829)	0.05 (*P*=0.823)	6.33 (*P*=0.019)	6.33 (*P*=0.012)	0.01 (*P*=0.929)	0.01 (*P*=0.929)	0.95 (*P*=0.348)	0.80 (*P*=0.371)		
3.2	0.15 (*P*=0.705)	0.15 (*P*=0.698)	4.14 (*P*=0.060)	4.15 (*P*=0.042)	17.28 (*P*≤0.001)	21.29 *P*≤0.001)	21.36 (*P*≤0.001)	22.85 (*P*≤0.001)		
3.3	3.84 (*P*=0.076)	3.87 (*P*=0.049)	35.12 (*P*≤0.001)	46.78 (*P*≤0.001)	1.06 (*P*=0.338)	1.06 (*P*=0.303)	1.93 (*P*=0.207)	1.72 (*P*=0.189)		

Genetic models involving different combinations of composite effects were compared for their adequacy to explain population differentiation on the basis of the significance of the lack-of-fit term. With data from six cross types, a genetic model with a maximum of five parameters being simultaneously fitted (i.e. the reference population mean μ plus four genetic terms) could be tested to allow at least one degree of freedom for the lack-of-fit term.

Genetic models: Model 1 is the base model, and included only additive and dominance effects; in addition to the composite genetic effects in the base model, Model 2 included two-locus epistatic effects fitted one at a time, i.e. additive x additive effects (α_2) in 2.1, additive x dominance effects ($\alpha_1\delta_1$) in 2.2, and dominance x dominance effects (δ_2) in 2.3; Model 3 is an extension of Model 2, by including two-locus epistatic effects fitted two at a time, i.e. α_2 and $\alpha_1\delta_1$ in 3.1, α_2 and δ_2 in 3.2, and $\alpha_1\delta_1$ and δ_2 in 3.3.

Calculated values and associated significance probabilities are presented for the Wald-type *F* statistic used to test the lack-of-fit term added to the mixed linear model after including composite genetic effects as covariates. Calculated values and associated significance probabilities for a chi-square (χ^2) test statistic, computed under the joint-scaling test procedure described by Lynch and Walsh (1998, pages 216-217), are also given for comparison. Both of these testing approaches resulted in similar conclusions in terms of model adequacy to explain the data.

For a given trait, the most parsimonious genetic model that led to the largest reduction in statistical significance of the lack-of-fit term (and also the χ^2 statistic under the joint-scaling test procedure), to a value not significant at the 10% level, is shaded and it was selected as the best model to explain the genetic basis of population divergence. However, this criterion of model adequacy could not be achieved for percentage of tree survival, as there was still statistically significant lack of fit under Model 3; this indicates that there was significant discrepancy between the observed cross-type means and the corresponding estimates obtained under the fitted

genetic model, which suggests that population differentiation for survival may comprise important epistatic effects involving more than pairs of loci.