Forests **2014**, *5*, 347-362; doi:10.3390/f5020347

OPEN ACCESS

forests

ISSN 1999-4907

www.mdpi.com/journal/forests

Article

Mapping Fusiform Rust Resistance Genes within a Complex Mating Design of Loblolly Pine

Tania Quesada [1], Marcio F.R. Resende Jr. [2], Patricio Muñoz [3], Jill L. Wegrzyn [4], David B. Neale [5], Matias Kirst [1,6], Gary F. Peter [1,6], Salvador A. Gezan [1], C. Dana Nelson [7] and John M. Davis [1,6,*]

1 School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611, USA; E-Mails: tquesada@ufl.edu (T.Q.); mkirst@ufl.edu (M.K.); gfpeter@ufl.edu (G.F.P.); sgezan@ufl.edu (S.A.G.)

2 Genetics and Genomics Graduate Program, University of Florida, Gainesville, FL 32611, USA; E-Mail: mresende@ufl.edu (M.R.)

3 Agronomy Department, University of Florida, Gainesville, FL 32611, USA; E-Mail: p.munoz@ufl.edu (P.M.)

4 Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, 60269 USA; E-Mail: jill.wegrzyn@uconn.edu (J.L.W.)

5 Department of Plant Sciences, University of California at Davis, Davis, CA 95616, USA; E-Mail: dbneale@ucdavis.edu (D.B.N.)

6 Genetics Institute, University of Florida, Gainesville, FL 32611, USA

7 USDA Forest Service, Southern Research Station, Southern Institute of Forest Genetics, Saucier, MS 39574, USA; E-Mail: dananelson@fs.fed.us (C.D.N.)

* Author to whom correspondence should be addressed; E-Mail: jmdavis@ufl.edu; Tel.: +1-352-846-0879; Fax: +1-352-392-1707.

Received: 8 October 2013; in revised form: 15 February 2014 / Accepted: 18 February 2014 / Published: 24 February 2014

Abstract: Fusiform rust resistance can involve gene-for-gene interactions where resistance (*Fr*) genes in the host interact with corresponding avirulence genes in the pathogen, *Cronartium quercuum* f.sp. *fusiforme* (*Cqf*). Here, we identify trees with *Fr* genes in a loblolly pine population derived from a complex mating design challenged with two *Cqf* inocula (one gall and 10 gall mixtures). We used single nucleotide polymorphism (SNP) genotypes at sufficient density to ensure linkage between segregating markers and *Fr* genes identifying SNPs that explained high proportions of variance in disease incidence using BayesCπ, that also were significant using Bayesian Association with Missing Data

(BAMD) software. Two SNPs mapped near *Fr1* and generated significant LOD scores in single marker regression analyses for *Fr1/fr1* parent 17 as well as four other parents. One SNP mapped near *Fr8* and was significant for parent 28. Two SNPs mapped to linkage groups not previously shown to contain *Fr* genes and were significant for three parents. Parent 2 showed evidence of *Fr* gene stacking. Our results suggest that it is feasible to identify trees segregating for *Fr* genes, and to map *Fr* genes, based on parental analysis of SNPs that cosegregate with disease incidence in designed resistance screening trials.

Keywords: *Pinus taeda*; major gene resistance; gene-for-gene interaction; single nucleotide polymorphism (SNP); BayesCπ; Bayesian Association with Missing Data (BAMD); logarithm of the odds (LOD)

1. Introduction

Fusiform rust is one of the most important pine diseases in the southeastern United States [1,2]. It is caused by the fungus *Cronartium quercuum* (Berk.) Miyabe ex Shirai f.sp *fusiforme* (*Cqf*), which alternates between oak and pine host species [1]. In pine, this fungus causes galls in stems and branches, reducing growth, reducing wood quality and making trees susceptible to breakage in windstorms, thereby generating significant economic losses [2].

Genetic resistance to fusiform rust can involve major genes in the pine host [3–6]. Resistant host genotypes carry one or more fusiform rust resistance (*Fr*) genes that interact with corresponding avirulence (*Avr*) genes in the pathogen. These are allele-specific interactions between host and pathogen—if the host is homozygous recessive (*fr/fr*) for an *Fr* gene, or if the pathogen carries an allele for virulence (*avr*) that can overcome the host *Fr* gene, the result in both cases is a diseased host. When the host carries a dominant allele (*Fr*) for resistance, and the pathogen is avirulent (*Avr*) to that gene, host resistance is expressed.

Given the difficulty of unraveling specific genetic interactions in populations with multiple *Fr* genes (at varying but unknown frequency) and multiple *Avr* genes (at varying but unknown frequency), mapping *Fr* genes has been best accomplished in families that segregate for single *Fr* genes after being challenged with genetically defined, single-spore-derived, inoculum [7,8]. A useful approach has been to separate galled and non-galled progeny of a suspected heterozygous *Fr/fr* tree, and screen the two pools with many hundreds of DNA markers using bulk segregant analysis. When the inoculum is avirulent to the *Fr* gene, markers linked to the *Fr* gene co-segregate with the resistance phenotype. This strategy was successfully applied to identify the first fusiform rust resistance locus in loblolly pine, *Fr1*, in progeny derived from a heterozygous (*Fr1/fr1*) genotype [6,8]. The DNA marker J7_470 cosegregated with resistance in seedlings inoculated with a *Cqf* single-spore isolate that was homozygous for the corresponding *Avr1* gene, and thus the marker was linked to *Fr1*. This finding was subsequently validated using the same parent [6] in the clonally-propagated CCLONES (Comparing Clonal Lines ON Experimental Sites) population [9] and in a slash × loblolly hybrid family [10].

Other *Fr* genes (numbered *Fr2* through *Fr9*) have since been discovered and mapped to single loci in the loblolly pine genome. These typically used the same strategy of bulk segregant analysis to identify linked markers in progeny of trees with *Fr* genes, having been challenged by single-spore cultures with distinct avirulence specificities [3,8]. When *Fr* genes are detected in two different parents at the same locus (*i.e.*, show linkage to the same DNA marker(s)) but show distinct interactions with inocula, they are assumed to be part of a cluster of resistance genes. Resistance gene clusters (e.g., complex resistance loci) contain two or more resistance genes as defined by their distinct avirulence specificities [8]. These gene clusters have been characterized for resistance genes in several different plant species in which the host has coevolved with the pathogen [11–14].

Given the availability of a large number of single nucleotide polymorphic (SNP) markers for the loblolly pine genome [15], we hypothesized that associations could be detected based on SNP marker linkage to *Fr* genes. The CCLONES population was already genotyped for SNP markers on each chromosome, so that bulk segregant analysis need not be performed on pools of galled and non-galled progeny to identify *Fr*-linked markers. Instead, phenotypes and genotypes can be associated to identify SNPs that explain variance in (*i.e.*, co-segregate with) host resistance to the inoculum. We reasoned that the presence of parents 17 (known *Fr1/fr1*) and 32 (known *Fr8/fr8*) in the CCLONES population [16], coupled with the reported avirulence of both one gall and 10 gall inocula to *Fr1* [9], would create useful internal checks for validating the approach, and for the possible discovery of additional *Fr* genes within the population.

To identify SNPs associated with genetic resistance, we used a statistical method called BayesCπ that is applied in genomic selection experiments [17] to identify those SNPs with the largest effects on disease phenotypes [18]. We then used the association software BAMD (Bayesian Association with Missing Data), which performs multiple imputations to resolve for missing data, provides a simultaneous solution for all markers analyzed, and [19] generates a confidence interval around each SNP's effect. We detected linkage to *Fr* genes that were previously mapped to single loci using RAPD markers, which were subsequently localized on the genetic map of loblolly pine [20]. We also detected linkage to at least one *Fr* gene in seven other parents, and found evidence of *Fr* gene stacking in one parent. The approach described here should prove useful for identifying parents that harbor *Fr* genes, and in tracking their transmission in southern pine breeding populations.

2. Experimental Section

2.1. Plant Material and Disease Phenotyping

Phenotypic data from the 69 full-sib families of CCLONES were obtained from a previous study [16], where inocula were derived from aeciospores from a single gall (one gall inoculum), as well as with a mixture of aeciospores collected from 10 different galls (10 gall inoculum). The data consisted of results obtained six months after inoculation for both tests. Gall score was recorded at the ramet level: 1 for ramets with at least one gall or 0 for ramets with no galls, and then the gall score was calculated for each genotype as the proportion of galled ramets. Gall length (mm) was measured as a continuous trait on all galled ramets [16]. The initial measurements [16] had misclassified families that

were discovered after genotypic data became available. Therefore, a DNA marker-corrected pedigree structure of CCLONES was utilized for the current study [21].

2.2. Identification of Markers Using BayesCπ

Genotypic data for 4853 polymorphic SNPs were obtained using the Illumina Infinium platform (Illumina, San Diego, CA) [22]. Genotypic data were available for 803 of the 1360 clones inoculated with the 10 gall inoculum and for 467 of the 698 clones inoculated with the one gall inoculum.

The hierarchical Bayesian model BayesCπ [17] was used to identify SNPs with large effects, since it is computationally less intensive than the association testing platform BAMD. As a consequence, we used BayesCπ to initially detect a subset of potentially significant SNPs, which we then evaluated in BAMD. For each trait, deregressed breeding values [23] were regressed on all markers with BayesCπ simultaneously and an estimate of the effect of each marker was obtained. For gall score, breeding values were obtained using the ramet incidence (estimating clonal breeding values using individual ramet gall score data). For gall length data, breeding values were obtained directly from the gall length measurements only from the subset of trees with galls. In each case, markers explaining greater than 0.2% of the phenotypic variance (*i.e.*, the deregressed breeding values) were selected for association testing.

2.3. Association Testing Using BAMD

Association analyses were performed using BAMD (Bayesian Association with Missing Data), which finds simultaneous solutions for SNP effects, performs multiple imputation of missing SNP data and generates a confidence interval from the posterior distribution [19]. Significant associations were detected using the following linear mixed model:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon} \quad (1)$$

where $\boldsymbol{Y}$ is the vector of deregressed breeding values [23] for the trait, $\boldsymbol{X}$ is the matrix for population structure effects, $\boldsymbol{\beta}$ is the coefficient for population structure effects, $\boldsymbol{Z}$ is the matrix for SNP effects, $\boldsymbol{\gamma}$ is the coefficient for SNP additive effects with a single variance and $\boldsymbol{\varepsilon}$ is the residual ~ N (0, $\mathbf{I}\sigma^2_{\varepsilon}$). No specific population structure groups were identified using Structure software, version 2.3.4 [24,25], thus the $\boldsymbol{X}$ matrix represented a single population group. The matrix $\boldsymbol{I}$ is an identity matrix. Significant SNPs were obtained based on the 90%, 95%, 97.5%, 99% and 99.9% confidence intervals from the last 50,000 of a total of 100,000 iterations of the Gibbs sampler in BAMD.

2.4. Single Marker Regression Analyses and Mapping of Significant SNPs

Single marker regression was performed on the significant SNPs from the one gall and 10 gall score datasets. This was done to determine the likelihood of each marker being linked to a gene causing the phenotype. We used SAS software [26] to fit thefollowing linear model:

$$\boldsymbol{Y} = \mu + \boldsymbol{Xr} + \boldsymbol{Zb} + \boldsymbol{\varepsilon} \quad (2)$$

where $\boldsymbol{Y}$ is the vector of observations for the trait, μ is the population mean, $\boldsymbol{r}$ is the vector for fixed effect replication (1–5 reps), $\boldsymbol{b}$ is the vector for fixed effect SNP genotype (AA, AB, BB) and $\boldsymbol{\varepsilon}$ is

the vector of random residual effects, and $\boldsymbol{X}$ and $\boldsymbol{Z}$ are incidence matrices. A reduced model was then fit where the SNP genotype effect was omitted from the equation and a likelihood ratio test was performed.

$$\text{LR} = -2\ln(\text{full model}) + 2\ln(\text{reduced model}) \quad (3)$$

Logarithm of the odds (LOD) scores was obtained from likelihood ratios by multiplying LR by 4.61 for linear-log conversion. LOD scores of 3 or higher were noted, and LOD scores of 4 or higher were used in genetic interpretations.

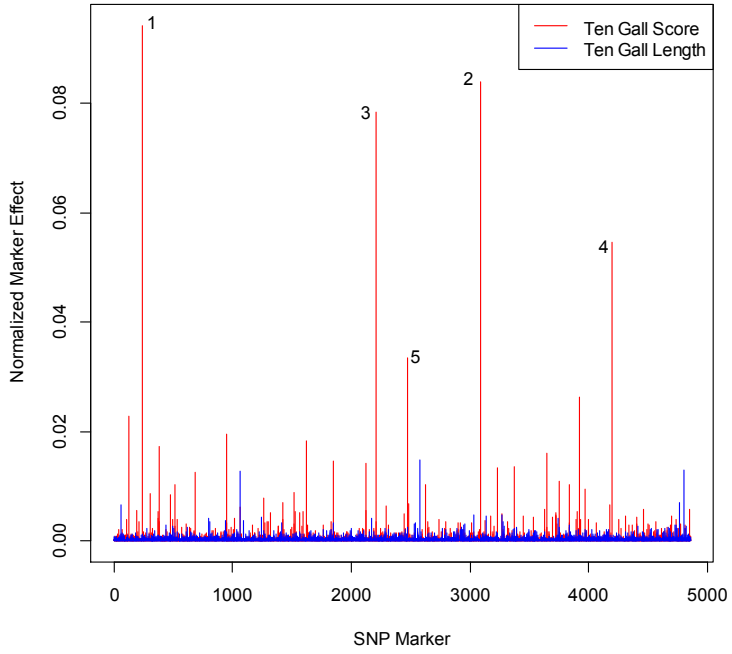
We conducted all possible single-marker regression analyses by parent for significant SNPs detected by both BayesCπ and BAMD. SNPs that were significantly associated with gall score were examined for their potential inclusion in genetic linkage maps relative to existing markers, which included RAPD markers previously recognized as linked to *Fr* loci [8,20,22] (Munoz P. and Peter G. unpublished data [27]) with our reference being the standard linkage group identifiers for loblolly pine [20]. Significant SNPs were recorded for all parents in the CCLONES population, and genetic inferences regarding *Fr* gene segregation were made based on the assumptions that: (1) parents segregating for an *Fr* gene would show a significant association with a marker linked to that *Fr* gene, and (2) the significant relationship would only be detected when the inoculum is avirulent to that *Fr* gene. These assumptions were based on a parent-by-parent basis given prior knowledge that parent 17 is *Fr1/fr1* [6] and parent 32 is *Fr8/fr8* (selection D in Amerson *et al.* [8]). We inferred that significant SNPs were associated with resistance when they cosegregated with an *Fr* gene (LOD > 4), and that SNPs would cosegregate with resistance only in cases where the inoculum was avirulent to the corresponding *Fr* gene. When a previously untested parent appeared to segregate for an *Fr* gene, we dropped out families (if any) shared with other parents that were either known to have an *Fr* gene, or inferred to have an *Fr* gene based on the parent-wise analysis. If the LOD score dropped <4, we inferred the previously untested parent did not have an *Fr* gene, but rather the association was due to the parent known to have an *Fr* gene..

3. Results

3.1. SNP Effect Size and Number Differs for Gall Score and Gall Length

The SNPs detected using BayesCπ that explained the greatest proportion of variance (over 0.2%) in gall score and gall length were normalized, superimposed, and the results for the 10 gall inoculum are presented in Figure 1. A comparatively coarser profile of SNP effects is observed for gall score as compared to gall length. Individual SNPs explained a greater proportion of the variance for gall score than for gall length, as the top five SNPs for gall score accounted for 29.9% of the variance, whereas the top five SNPs for gall length accounted for only 4.3%. In addition, a greater number of SNPs accounted for >0.2% of variance for gall score (N = 50) than for gall length (N = 16). The same trends were observed for the one gall inoculum, where SNP effects for gall score were larger (8.8% for the top five) and greater in number (N = 29) than for gall length (3.1% for the top five, N = 15; Figures S1 and S2). Taken together, more SNPs with major effects were detected for gall score than for gall length, which suggested that major effect SNPs for gall score warranted closer examination.

Figure 1. Comparative magnitude of the normalized effects (absolute values) for gall score (red markers) and gall length (blue markers) obtained using Bayes Cπ in the 10 gall datasets. SNP markers are represented by a single line and are in alphanumerical order. Numbers correspond to ranking of the top five SNPs with largest effects on gall score in the 10 gall dataset (See Figure S1 for identification of the top-ranked SNPs).



3.2. SNPs Differ in One Gall and 10 Gall Experiments

Considering all SNPs identified by BayesCπ with effects >0.2% on gall score (Figure S1), there were eight SNPs that were common for those detected in the one gall and 10 gall experiments, with their ranking being similar. Stated a different way, there was a total of 71 SNPs that interacted uniquely with either the one gall or 10 gall inocula, and eight SNPs that interacted similarly with both inocula.

SNPs selected using BayesCπ were analyzed in BAMD, and the results from both methods are presented in Figure S1. BAMD assigned a nonzero mean effect on gall score with a confidence interval of 90% or greater for a total of 16 SNPs (out of a total 29) for the one gall inoculum and 28 (out of a total 50) for the 10 gall inoculum. Because the SNPs declared by BAMD as significant represented a subset of the total number of SNPs detected by BayesCπ, we focused our attention on those SNPs that were detected by both methods. SNPs for both gall score and gall length identified by both BayesCπ and BAMD are tallied in Table 1.

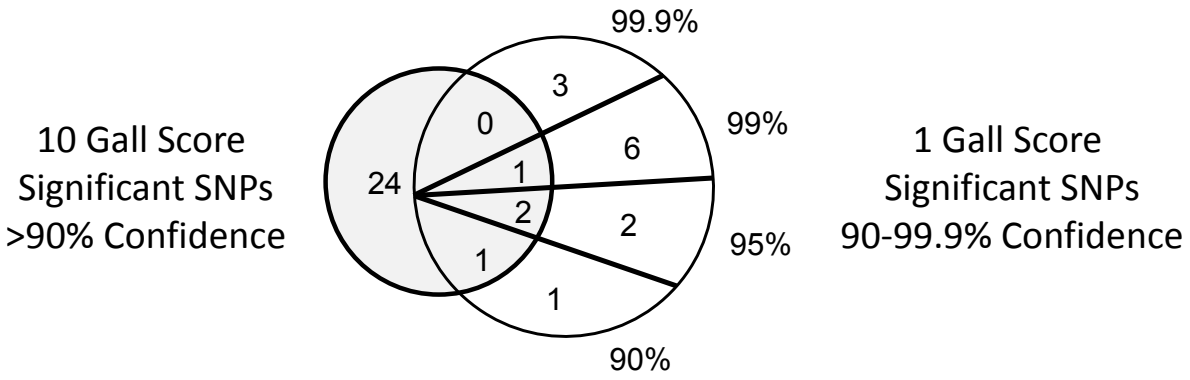
Since some of the significant SNPs presented in Table 1 were common for the two experiments, they are further compared in Figure 2. Of the 40 SNPs identified with high confidence (>90%) in one gall and 10 gall score experiments, four were shared. Collectively, these results imply there are some major SNP effects that explain host responses to both inocula. This would be expected when both inocula are predominantly avirulent to a specific *Fr* gene that is segregating in the population, e.g., SNP #4 (Figure 1). However, most SNPs are unique to the one gall or 10 gall inocula,

presumably reflecting differential interactions of *Fr* genes avirulent to only one inoculum. All significant SNPs were examined for potential linkage to one or more rust resistance genes.

Table 1. Significant SNPs after tandem BayesCπ and BAMD analysis for resistance to fusiform rust.

Inoculum	Phenotype	SNPs at Confidence Interval (%)			
		90–94.9	95–98.9	99–99.8	>99.9
1 gall	Gall score	2	4	7	3
10 gall		7	10	4	7
1 gall	Gall length	2	2	0	0
10 gall		1	4	3	2

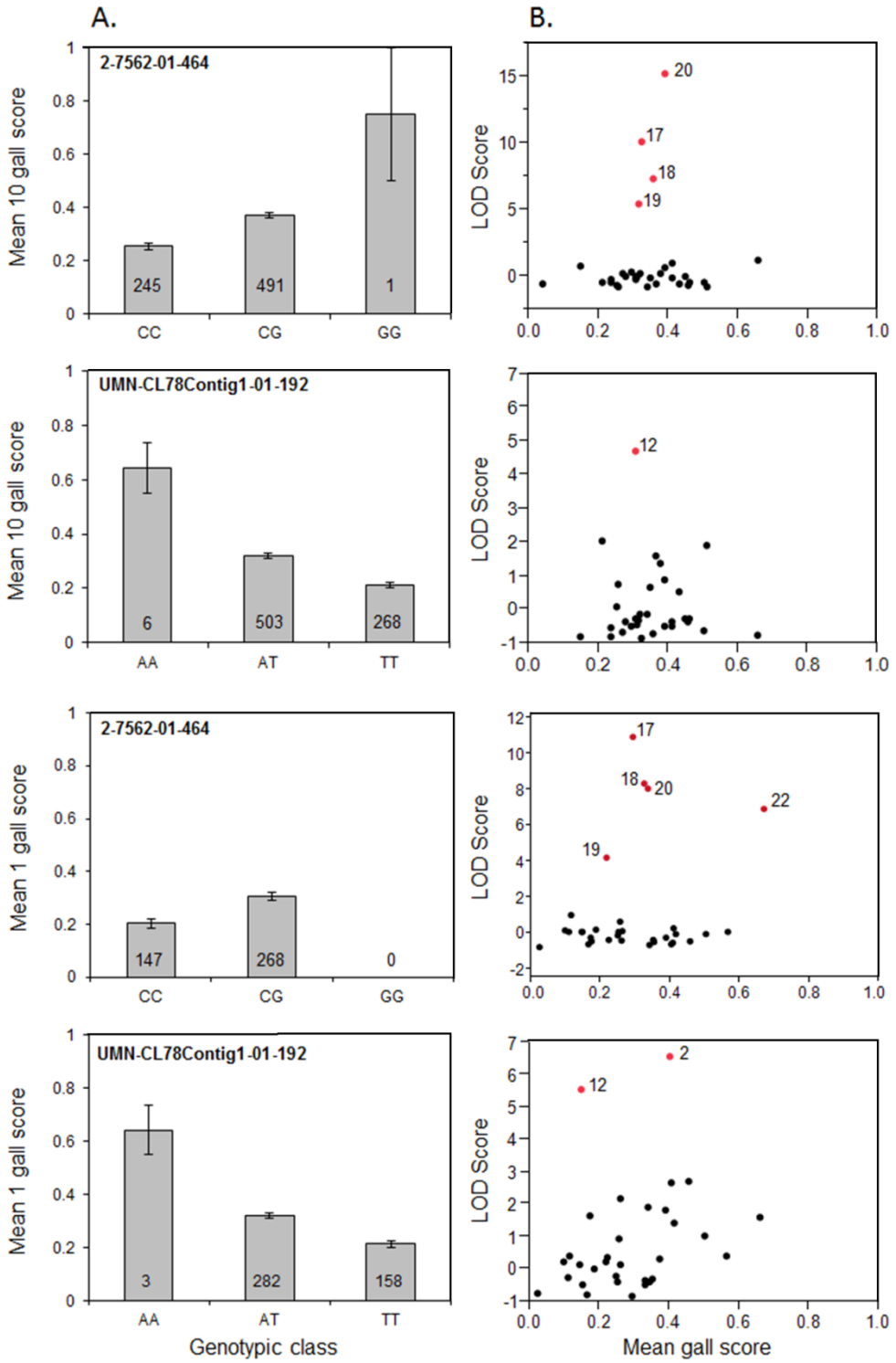
Figure 2. Venn diagram showing significant SNPs for rust resistance between gall score for 10 gall and one gall data that are summarized in Table 1. Of the total of 21 significant SNPs that were detected in the 10 gall inoculum test, four were also significant for the one gall inoculum at varying levels of significance, while 12 were only significant for the one gall inoculum.



3.3. Significant SNP Markers are Associated with Resistance

Significant SNPs were regressed on phenotypes (*i.e.*, gall score de-regressed breeding values) across the entire CCLONES population to determine the magnitude of the association, and by parent to identify those parents that may be segregating for resistance (Figure 3). In the population analysis (Figure 3A), the SNP genotype classes were typically represented by two major classes, with one of the homozygous classes in some cases represented by a small number of individuals, although that was not the case for all significant SNPs (Tables S2 and S3). Associations between gall score and the SNP marker are detected as significant mean differences between the major homozygous class and the heterozygous class (Figure 3A).

Figure 3. (**A**) Mean gall scores for 10 gall (top) and one gall (bottom) inoculum across genotypic classes for two significant SNPs for gall score. Error bars correspond to standard errors of the mean and the numbers within the graphs correspond to the number of clones for each genotypic class in the CCLONES population; (**B**) Single marker regression results showing LOD scores *vs.* mean gall score for 10 gall (or one gall) inoculum by parent for the two significant SNPs shown in A. Each point within the plot corresponds to a single parent. Red dots show parents with LOD scores above 3 with their corresponding identification number, while black dots show parents with LOD scores <3.



When single-marker regression was performed by parent, most parents showed non-significant relationships between the SNP class and gall score among their progeny (Figure 3B). However, significant relationships (LOD score>4) were detected for subsets of parents. Cases were found in which the significant relationship between a marker and gall score for a parent was observed for one gall inoculum only, or for 10 gall inoculum only, or for both inocula, presumably reflecting differences in virulence of the inocula to specific *Fr* genes.

We tabulated the results from single marker regression analyses into a summary table (Table S1) for all parents, including the families for which they were either the seed or pollen donor, the inoculum source (one gall or 10 gall) that appears avirulent to the corresponding *Fr* gene, the LOD scores for significant SNPs, the linkage group to which the significant SNP has been assigned, and our genetic interpretation of the results.

We detected a few cases in which a significant LOD score was inconsistent with what was already known about a parent. Most notably, parents 18 and 19 were known test cross parents for *Fr1* (*i.e.*, they are both *fr1/fr1*); however both parents showed significant LOD scores for an *Fr* gene linked to the *Fr1* locus. The most likely explanation for this inconsistency relates to the circular mating design within the CCLONES population. The parental analyses are not entirely independent, so parents adjacent to one another in the mating scheme tend to share families—e.g., family 400 (parent 17 × parent 18) and family 401 (parent 17 × parent 19). We eliminated shared families, repeated regression analyses for the relevant SNPs, and filtered the results in Table S1 to include only those parents for which we could confidently assign *Fr* genes to the correct parents.

The results from parents heterozygous for *Fr* genes, and for other parents that shared the same significant associations, are presented in Table 2. Two parents (parents 17 and 32) are known heterozygotes for *Fr* genes (*Fr1/fr1* and *Fr8/fr8*, respectively; selections 10–5 and D, in Amerson *et al.* [8]), while one parent (parent 23) is heterozygous for a potentially new *Fr* gene (tentatively defined as *Fr10*) mapping to linkage group 11, which distinguishes it from previously recognized *Fr* genes 1–9 [8]. Along with parent 17, four additional parents (parents 2, 12, 20 and 22) showed significant associations with *Fr1*-linked SNPs. Parent 32 (*Fr8/fr8*) showed no significant associations in the SNP marker analyses, while parent 28 showed significant associations with a SNP marker linked to a previously identified marker linked to *Fr8* [8]. Finally, SNPs associated with gall score in parent 23 (*Fr10*) were also associated in parents 2 and 4. Parent 2 showed evidence of *Fr* gene stacking, with *Fr* genes within the *Fr1* and *Fr10* complex loci. We eliminated parents 18 and 19 from Table 2 based on their known *fr1/fr1* genotype, which we confirmed by finding non-significant LOD scores for parents 18 and 19 when families shared with parents 17 and 20 were excluded from the regression analysis.

Table 2. Significant SNPs obtained using BayesCπ and BAMD that mapped to linkage groups (LG) containing *Fr* genes. Each parent was known from previous work to segregate for a major gene, or yielded a LOD score >4 when the genotypic classes for the SNP were regressed on gall score breeding values from one gall and 10 gall inoculation experiments.

Locus	LG	Parent	Families	Avr Source	Significant SNP (LOD 1Gall/ 10 Gall)	Genetic interpretation
Fr1	2	17	400, 401, 416, 436, 442	1 Gall/10 Gall	2-7562-01-464 (10.9/10.0)	Parent 17 known *Fr1/fr1*, 1/10 gall known avirulent to *Fr1*
		12	419, 428, 446, 466, 472	1 Gall/10 Gall	UMN-CL78 Contig1-01-192 (5.5/4.7)	
		20	408, 429, 430, 457, 461, 476, 477	1 Gall/10 Gall	2-7562-01-464 (8.9/15.1)	Parents 12, 20 may segregate for *Fr1* or non-*Fr1* at locus, 1/10 gall avirulent
		2	406, 412, 423, 433, 435, 454, 469	1 Gall	UMN-CL78 Contig1-01-192 (6.5/1.9)	Parents 2, 22 may segregate for non-*Fr1* at locus, 1 gall avirulent
		22	403, 408, 420, 451	1 Gall	2-7562-01-464 (6.6/1.2)	
Fr8	-	32	404, 461	-	-	Parent 32 known *Fr8/fr8*, neither 1/10 gall appear avirulent to *Fr8*
	10	28	402, 421, 433, 452, 453, 475	1 Gall	CL3925Contig1 -03-163 (6.4/-0.6)	Parent 28 may segregate for non-*Fr8* at locus, 1 gall avirulent
Fr-	11	23	420, 443, 448, 455, 467, 473	10 Gall	0-13348-01-203 (2.8/4.1)	Parent 23 known unnamed Fr resistance gene (*Fr-*) 10 gall appears avirulent to *Fr-*
		4	415, 437, 469	10 Gall	0-13348-01-203 (2.7/7.0)	Parent 4 may segregate for *Fr-* or non-*Fr-* at locus, 10 gall avirulent
		2	406, 412, 423, 433, 435, 454, 469	1 Gall/10 Gall	0-2040-01-149 (6.8/5.2)	Parent 2 may segregate for non-*Fr-* at locus, 1/10 gall avirulent

4. Discussion

Rapid progress has been made in genetic development of fusiform rust resistance with a few generations of selection [3,28,29]. Given this rapid progress, it seems reasonable to conclude that much of the improvement has been due to selection for major gene resistance [3,30–32]. Genomic markers now enable detection of parents that harbor specific *Fr* genes, so that their progeny deployed in plantation forests can be monitored for durability under field conditions.

As a step toward the goal of identifying *Fr* genes in breeding populations, we used two statistical approaches—Bayes Cπ and BAMD—to detect segregating *Fr* genes in a population generated by a complex mating design. BayesCπ is computationally rapid, and appears to efficiently identify SNPs associated with oligogenic traits [17,18]. BAMD is computationally more intensive as it performs formal multiple imputation for missing SNPs, and generates a Bayesian confidence interval to assess the quality of associations. By using these two methods in tandem, we identified SNPs that were highly significant for rust resistance. It was not our intention to formally compare and contrast BayesCπ and BAMD; rather we wanted to screen for SNPs that mapped near *Fr* loci. We assume that higher LOD scores for the association represents closer linkage of the SNP marker to the *Fr* gene, coupled with a small proportion of misclassified genotypes, for example due to a low level of virulence in the respective inoculum to a particular *Fr* gene that is being mapped.

4.1. SNPs Mapped to Fr1 (LG2)

The avirulence of the one gall and 10 gall inocula to most *Fr* genes was not known prior to this study; however, an exception was *Fr1*, where we previously reported [9] that both one and 10 gall inocula were avirulent to *Fr1* with rare exceptions. Therefore, we expected markers linked to *Fr1* in the progeny of parent 17 would show significant associations in both one and 10 gall experiments. No predictions about avirulence to other *Fr* genes were made; rather we made genetic inferences about avirulence based on significant associations that were detected.

We validated the overall computational approach by detecting significant SNPs (including SNP #4) that were linked to *Fr1* in parent 17 (*Fr1/fr1*). In addition to detecting *Fr1*, there is at least one other *Fr* gene at the *Fr1* locus. This *Fr* gene is in parents 2 and 22 and differs from *Fr1* in that only the one gall inoculum is avirulent to it (*i.e.*, in the 10 gall inoculum this *Fr* gene has been overcome). The available data do not allow us to identify this *Fr* gene in parents 2 and 22—it could be one previously identified that is linked to *Fr1* (*Fr3*, *Fr4*, *Fr6*, *Fr7*, *Fr9*) or a new one. More detailed analysis of parents 2 and 22 is warranted.

Parents 12 and 20 are similar to parent 17. The data are consistent with parents 12 and 20 harboring *Fr1*, but we cannot exclude a different *Fr* gene to which both inocula are avirulent.

4.2. SNPs Mapped to Fr8 (LG10)

Parents 28 and 32 appear to have different *Fr* genes near the *Fr8* locus on linkage group 10. Previous mapping experiments [8] determined that parent 32 is *Fr8/fr8*, and it appears neither inoculum is avirulent to *Fr8* because we detected no significant SNPs in this parent. By contrast, we detected significant SNPs linked to the locus in parent 28 (only one gall avirulent).

4.3. SNPs Mapped to LG11

Parents 2, 4, and 23 have at least one *Fr* gene that maps to linkage group 11. Previous inoculation trials and limited marker investigation (Amerson, unpublished) suggested that parent 23 is heterozygous for a previously unmapped *Fr* gene. For the resistance gene that we have recognized in the current study in parent 23, it appears only that the 10 gall inoculum is avirulent to this gene. This

pattern is shared among the other parents, but we cannot exclude that one or more may harbor a different *Fr* gene to which only the 10 gall inoculum is avirulent.

4.4. Genetic Architecture of Rust Resistance

Significant SNPs were detected in some parents at more than one known *Fr* locus. Parent 2 may harbor more than one *Fr* gene, based on segregation for genes at *Fr1* and the new locus on linkage group 11. As analyses proceed in more advanced pedigrees, it should be expected that parents with multiple *Fr* genes will be detected, especially given the importance of rust resistance as a selected trait and the opportunities for *Fr* gene "stacking" to have occurred during the first, second and/or third breeding cycle of southern pine tree improvement [28–30,33–35] Parent 2 is an interesting example in which the markers linked to *Fr1* and the linkage group 11 locus enabled the detection of these *Fr* genes in only the one gall and in both inocula, respectively. Thus, the inocula served as differentials for the *Fr* genes, and the markers provided the information on their locations. Based on our results, it seems feasible to identify these kinds of parents in complex populations, and then sort out the actual *Fr* identification using more laborious but precise methods [4,30].

We also analyzed gall length on susceptible genotypes, in which larger galls indicate greater severity of the disease, whereas smaller galls may reflect a certain degree of tolerance; a possible indicator of partial resistance [16,35,36]. No SNPs were shared between gall score and length datasets, which reinforces the observed lack of genetic correlation between the traits using quantitative genetic methods [16]. In contrast to gall score, the genetic architecture of gall length appears to be quantitative, with many genes of small effect [16,36].

4.5. Identifying Fr Genes: Paths Forward

It will be useful to map all significant SNPs, so that individuals can be genotyped by breeders for *Fr* genes using flanking markers. Identification of *Fr* genes themselves would obviate the need to use SNPs in linked genes as surrogates for marker assisted selection. In several plant systems, disease resistance genes have been molecularly identified and this creates useful tools for breeders to monitor their transmission to progeny [37–41]. Many of these resistance genes have been identified in genomic mapping studies, and as the loblolly pine draft genome continues to improve with additional information, integration of the genetic and physical maps around *Fr* loci should be feasible. The logic of genomic mapping is that an *Fr* gene must be present on any scaffold containing two known flanking markers. The loblolly pine genome is physically the largest sequenced to date at 21.7 Gb [42] with a genetic map distance of 899 cM [22]. In a hypothetical situation in which two SNPs flanking *Fr1* were 1 cM apart on a linkage map, the minimum length of a scaffold containing both SNPs would be ~20,000 Kb. Alternatively, candidate gene approaches based on homology to known R genes may help identify *Fr* genes. The continued refinement of the loblolly pine genome, coupled with increasingly dense genetic maps, should enable *Fr* genes to be identified in the near future.

5. Conclusions

We identified SNPs linked to known *Fr* genes in loblolly pine by using BayesCπ and BAMD in tandem. Specific parents that segregate for one or more *Fr* genes were identified using single-marker regression analysis. We also identified significant SNPs that map to presumed *Fr* loci in previously unreported regions of the pine genome. Four general principles seem the most important to consider in detecting *Fr* genes in breeding and natural populations: 1) Incorporation of host genotypes with known *Fr* genes as internal controls; 2) Integration of markers linked to *Fr* loci in the host population; 3) Use of differential inoculum, and 4) Reliable assessment of phenotypes to ensure correct classification of resistant and susceptible trees. Our approach supports the feasibility of using markers to guide breeding and selection for fusiform rust resistance.

Acknowledgments

This work was supported by USDA-CSREES-IFAFS (grant 2001-52100-11315) and by the National Science Foundation (IOS-PGRP-0501763). The Forest Biology Research Cooperative at the University of Florida provided the plant materials, and Dr. Robert Schmidt (retired, University of Florida) provided inoculum. We thank Carol Young (retired, US Forest Service, Resistance Screening Center), Kathy Smith (Southern Institute of Forest Genetics) and Chris Dervinis (University of Florida) for technical assistance. Finally, we thank Henry Amerson and Thomas Kubisiak for making their *Fr* gene mapping information available for cross-reference during the preparation of this manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Dwinell, L.D. *Biology of Fusiform Rust, Management of Fusiform Rust in Southern Pines*; Dinus, R.J., Schmidt, R.A., Eds.; University of Florida: Gainesville, FL, USA, 1977; p. 163.
2. Cubbage, F.; Pye, J.; Holmes, T.; Wagner, J. An economic analysis of fusiform rust protection research. *South. J. App. For.* **2000**, *24*, 77–85.
3. Amerson, H.V.; Kubisiak, T.L.; Garcia, S.A.; Kuhlman, G.C.; Nelson, C.D.; McKeand, S.E.; Mullin, T.J.; Li, B. Interacting Genes in the Pine-Fusiform Rust Forest Pathosystem. In proceedings of the 28th Biennial Southern Forest Tree Improvement Conference, Raleigh, NC, USA, 21–23 June, 2005; U.S. Department of Agriculture, Forest Service: Raleigh, NC, USA, 2005; p. 60.
4. Kubisiak, T.; Anderson, C.; Amerson, H.; Smith, J.; Davis, J.; Nelson, C. A genomic map enriched for markers linked to *Avr1 Cronartium. quercuum* f.sp. *fusiforme*. *Fungal Gen. Biol.* **2011**, *48*, 266–274.
5. Kubisiak, T.L.; Amerson, H.V.; Nelson, C.D. Genetic interaction of the fusiform rust fungus with resistance gene fr1 in loblolly pine. *Phytopathology* **2005**, *95*, 376–380.

6. Wilcox, P.L.; Amerson, H.V.; Kuhlman, E.G.; Liu, B.H.; O'Malley, D.M.; Sederoff, R.R. Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3859–3864.
7. Nelson, C.; Kubisiak, T.; Amerson, H. Unravelling and managing fusiform rust disease: A model approach for coevolved forest tree pathosystems. *For. Pathol.* **2010**, *40*, 64–72.
8. Amerson, H. North Carolina State University, Raleigh, NC, USA. *Forests* **2014**, unpublished.
9. Kayihan, G.; Nelson, C.; Huber, D.; Amerson, H.; White, T.; Davis, J. Clonal evaluation for fusiform rust disease resistance: Effects of pathogen virulence and disease escape. *Can. J. For. Res. Rev.* **2010**, *40*, 1042–1050.
10. Huber, D.; Amerson, H., Performance of the loblolly pine fusiform rust disease resistance gene (*Fr1*) in a slashXloblolly pine hybrid family. *Tree Genet. Genomes* **2011**, *7*, 535–540.
11. Thomas, C.M.; Vos, P.; Zabeau, M.; Jones, D.A.; Norcott, K.A.; Chadwick, B.P.; Jones, J.D. Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato *Cf-9* gene for resistance to *Cladosporium. fulvum*. *Plant J.* **1995**, *8*, 785–794.
12. Dodds, P.N.; Lawrence, G.J.; Catanzariti, A.M.; Teh, T.; Wang, C.I.; Ayliffe, M.A.; Kobe, B.; Ellis, J.G. Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8888–8893.
13. Bahri, B.; Kaltz, O.; Leconte, M.; de Vallavieille-Pope, C.; Enjalbert, J. Tracking costs of virulence in natural populations of the wheat pathogen, *Puccinia. striiformis* f.sp. *tritici*. *BMC Evol. Biol.* **2009**, *9*, 26.
14. McHale, L.K.; Truco, M.J.; Kozik, A.; Wroblewski, T.; Ochoa, O.E.; Lahre, K.A.; Knapp, S.J.; Michelmore, R.W. The genomic architecture of disease resistance in lettuce. *Theor. Appl. Genet.* **2009**, *118*, 565–580.
15. Eckert, A.; Pande, B.; Ersoz, E.; Wright, M.; Rashbrook, V.; Nicolet, C.; Neale, D. High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda* L.). *Tree Genet. Genomes* **2009**, *5*, 225–234.
16. Kayihan, G.; Huber, D.; Morse, A.; White, T.; Davis, J. Genetic dissection of fusiform rust and pitch canker disease traits in loblolly pine. *Theor. App. Genet.* **2005**, *110*, 948–958.
17. Habier, D.; Fernando, R.; Kizilkaya, K.; Garrick, D. Extension of the bayesian alphabet for genomic selection. *BMC Bioinform.* **2011**, *12*, 186.
18. Resende, M.; Munoz, P.; Resende, M.; Garrick, D.; Fernando, R.; Davis, J.; Jokela, E.; Martin, T.; Peter, G.; Kirst, M. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). *Genetics* **2012**, *190*, 1503–1510.
19. Li, Z.; Gopal, V.; Li, X.; Davis, J.M.; Casella, G. Simultaneous SNP identification in association studies with missing data. *Ann. App. Statist.* **2012**, *6*, 432–456.
20. Echt, C.S.; Saha, S.; Krutovsky, K.V.; Wimalanathan, K.; Erpelding, J.E.; Liang, C.; Nelson, C.D. An annotated genetic map of loblolly pine based on microsatellite and cDNA markers. *BMC Genet.* **2011**, *12*, 17.

21. Munoz, P.R.; Resende, M.F.R.; Huber, D.A.; Quesada, T.; Resende, M.D.V.; Neale, D.V.; Wegrzyn, J.L.; Kirst, M.; Peter, G.F. Genomic relationship matrix for correcting pedigree errors in breeding populations: Impact on genetic parameters and genomic selection accuracy. *Crop Sci.* **2013**, doi:10.2135/cropsci2012.12.0673.
22. Eckert, A.; van Heerwaarden, J.; Wegrzyn, J.; Nelson, C.; Ross-Ibarra, J.; Gonzalez-Martinez, S.; Neale, D. Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., *pinaceae*). *Genetics* **2010**, *185*, 969–982.
23. Garrick, D.; Taylor, J.; Fernando, R. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genet. Sel. Evol.* **2009**, 41.
24. Falush, D.; Stephens, M.; Pritchard, J.K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **2003**, *164*, 1567–1587.
25. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959.
26. SAS Institute Inc. *SAS/STAT® 9.3. User's Guide*; SAS Institute, Inc: Cary, NC, USA, 2011.
27. Munoz, P.; Peter, G. University of Florida, Gainesville, FL, USA. Unpublished Data.
28. Li, B.; McKeand, S.; Weir, R. Tree improvement and sustainable forestry—Impact of two cycles of loblolly pine breeding in the USA. *For. Genet.* **1999**, *6*, 229–234.
29. McKeand, S.; Bridgwater, F. A strategy for the third breeding cycle of loblolly pine in the southeastern US. *Silvae Genet.* **1998**, *47*, 223–234.
30. Nelson, C.D.; Johnsen, K.H. Genomic and physiological approaches to advancing forest tree improvement. *Tree Physiol.* **2008**, *28*, 1135–1143.
31. Myburg, H.; Morse, A.; Amerson, H.; Kubisiak, T.; Huber, D.; Osborne, J.; Garcia, S.; Nelson, C.; Davis, J.; Covert, S.; *et al.* Differential gene expression in loblolly pine (*Pinus taeda* L.) challenged with the fusiform rust fungus, *Cronartium. quercuum* f.sp *fusiforme*. *Physiol. Mol. Plant Pathol.* **2006**, *68*, 79–91.
32. Cumbie, W.; Isik, F.; McKeand, S. Genetic improvement of sawtimber potential in loblolly pine. *For. Sci.* **2012**, *58*, 168–177.
33. McKeand, S.E.; Jokela, E.J.; Huber, D.A.; Byram, T.D.; Allen, H.L.; Li, B.; Mullin, T.J. Performance of improved genotypes of loblolly pine across different soils, climates, and silvicultural inputs. *For. Ecol. Manag.* **2006**, *227*, 178–184.
34. Schmidt, R.; Powers, H.; Snow, G. Application of genetic-disease resistance for the control of fusiform rust in intensively managed southern pine. *Phytopathology* **1981**, *71*, 993–997.
35. Schmidt, R.; Gramacho, K.; Miller, T.; Young, C. Components of partial resistance in the slash pine-fusiform rust pathosystem. *Phytopathology* **2000**, *90*, 1005–1010.
36. Poland, J.; Balint-Kurti, P.; Wisser, R.; Pratt, R.; Nelson, R. Shades of gray: The world of quantitative disease resistance. *Trends Plant. Sci.* **2009**, *14*, 21–29.
37. Ercolano, M.R.; Sanseverino, W.; Carli, P.; Ferriello, F.; Frusciante, L. Genetic and genomic approaches for r-gene mediated disease resistance in tomato: Retrospects and prospects. *Plant Cell Rep.* **2012**, *31*, 973–985.
38. Vleeshouwers, V.G.; Raffaele, S.; Vossen, J.H.; Champouret, N.; Oliva, R.; Segretin, M.E.; Rietman, H.; Cano, L.M.; Lokossou, A.; Kessel, G.; *et al.* Understanding and exploiting late blight resistance in the age of effectors. *Annu. Rev. Phytopathol.* **2011**, *49*, 507–531.

39. Xiao, W.; Yang, Q.; Wang, H.; Duan, J.; Guo, T.; Liu, Y.; Zhu, X.; Chen, Z. Identification and fine mapping of a major r gene to *Magnaporthe oryzae* in a broad-spectrum resistant germplasm in rice. *Mol. Breed.* **2012**, *30*, 1715–1726.
40. Vida, G.; Gal, M.; Uhrin, A.; Veisz, O.; Syed, N.; Flavell, A.; Wang, Z.; Bedo, Z. Molecular markers for the identification of resistance genes and marker-assisted selection in breeding wheat for leaf rust resistance. *Euphytica* **2009**, *170*, 67–76.
41. Panwar, P.; Jha, A.; Pandey, P.; Gupta, A.; Kumar, A. Functional markers based molecular characterization and cloning of resistance gene analogs encoding NBS-LRR disease resistance proteins in finger millet (*Eleusine coracana*). *Mol. Biol. Rep.* **2011**, *38*, 3427–3436.
42. Magbanua, Z.V.; Ozkan, S.; Bartlett, B.D.; Chouvarine, P.; Saski, C.A.; Liston, A.; Cronn, R.C.; Nelson, C.D.; Peterson, D.G. Adventures in the enormous: A 1.8 million clone BAC library for the 21.7 gb genome of loblolly pine. *PLoS One* **2011**, *6*, e16214.