Phytopathology® • 2022 • 112:1093-1102 • https://doi.org/10.1094/PHYTO-09-21-0382-R

e-Xtra*

Comparative Association Mapping Reveals Conservation of Major Gene Resistance to White Pine Blister Rust in Southwestern White Pine (*Pinus strobiformis*) and Limber Pine (*P. flexilis*)

Jun-Jun Liu,[1,†] Anna W. Schoettle,[2] Richard A. Sniezko,[3] Kristen M. Waring,[4] Holly Williams,[1] Arezoo Zamany,[1] Jeremy S. Johnson,[3,4] and Angelia Kegley[3]

[1] Canadian Forest Service, Natural Resources Canada, Victoria, British Columbia V8Z 1M5, Canada
[2] Rocky Mountain Research Station, Forest Service, U.S. Department of Agriculture, Fort Collins, CO 80526, U.S.A.
[3] Dorena Genetic Resource Center, Forest Service, U.S. Department of Agriculture, Cottage Grove, OR 97424, U.S.A.
[4] School of Forestry, Northern Arizona University, Flagstaff, AZ 86011-5018, U.S.A.
Accepted for publication 29 October 2021.

ABSTRACT

All native North American white pines are highly susceptible to white pine blister rust (WPBR) caused by *Cronartium ribicola*. Understanding genomic diversity and molecular mechanisms underlying genetic resistance to WPBR remains one of the great challenges in improvement of white pines. To compare major gene resistance (MGR) present in two species, southwestern white pine (*Pinus strobiformis*) *Cr3* and limber pine (*P. flexilis*) *Cr4*, we performed association analyses of *Cr3*-controlled resistant traits using single nucleotide polymorphism (SNP) assays designed with *Cr4*-linked polymorphic genes. We found that ~70% of *P. flexilis* SNPs were transferable to *P. strobiformis*. Furthermore, several *Cr4*-linked SNPs were significantly associated with the *Cr3*-controlled traits in *P. strobiformis* families. The most significantly associated SNP (M326511_1126R) almost colocalized with *Cr4* on the *Pinus* consensus linkage group 8, suggesting that *Cr3* and *Cr4* might be the same R locus, or have localizations very close to each other in the syntenic region of the *P. strobiformis* and *P. flexilis* genomes. M326511_1126R was identified as a nonsynonymous SNP, causing amino acid change ($Val_{376}Ile$) in a putative pectin acetylesterase, with coding sequences identical between the two species. Moreover, top *Cr3*-associated SNPs were further developed as TaqMan genotyping assays, suggesting their usefulness as marker-assisted selection (MAS) tools to distinguish genotypes between quantitative resistance and MGR. This work demonstrates the successful transferability of SNP markers between two closely related white pine species in the hybrid zone, and the possibility for deployment of MAS tools to facilitate long-term WPBR management in *P. strobiformis* breeding and conservation.

Keywords: association mapping, *Cronartium ribicola*, cross-species transferability of DNA markers, disease resistance, forest pathology, limber pine (*Pinus flexilis*), major gene resistance, marker-assisted selection, quantitative resistance, single nucleotide polymorphisms, southwestern white pine (*P. strobiformis*), white pine blister rust

Cronartium ribicola (J.C. Fisch. in. Rabh.) is the causal agent of white pine blister rust (WPBR). Since its accidental introduction to North America in the early 1900s, *C. ribicola* has spread across almost all geographical regions where eight of nine native white pines are distributed in North America, with the one exception being the Great Basin bristlecone pine (*Pinus longaeva*; Sniezko et al. 2011; Tomback and Achuff 2010). All native white pine species are highly susceptible to *C. ribicola*, including two closely related high elevation species, southwestern white pine (*P. strobiformis* Engelm.) and limber pine (*P. flexilis* James; Kearns et al. 2014; Tomback and Achuff 2010). In the central Rocky Mountains, 73% of the *P. flexilis* stands assessed have been invaded by WPBR and the disease continues to spread in the southern Rocky Mountains (Cleaver et al. 2015). By 2030, *P. flexilis* is expected to experience a 40% reduction in basal area in the United States (Krist and Romero 2014). Over 60% of *P. flexilis* trees were infected with >40% mortality in southern Alberta in 2009 (Smith et al. 2013). The Committee on the Status of Endangered Wildlife in Canada expects mortality of close to two-thirds of the mature *P. flexilis* over the next 100 years (COSEWIC 2014). In the province of Alberta, *P. flexilis* is listed as endangered and it is being considered for national listing as such under the Species at Risk Act (Government of Alberta 2014). WPBR infection has also been detected in *P. strobiformis* stands across the southwestern United States but has not yet been found in stands in Mexico (Geils et al. 2010; Looney et al. 2015).

Improvement of genetic resistance is considered as the most efficient approach for long-term WPBR management in *P. strobiformis* and other related species (Conklin et al. 2009; Sniezko et al. 2018). Both major gene resistance (MGR) and quantitative resistance (QR) to WPBR have been documented in several white pine species in screening and breeding programs (Sniezko et al. 2008, 2011, 2014). White pine MGRs to WPBR are controlled by a single dominant genetic locus, including *Cr1* in *P. lambertiana* (sugar pine; Kinloch 1970), *Cr2* in *P. monticola* (western white pine; Kinloch et al. 1999), *Cr3* in *P. strobiformis* (Kinloch and Dupper 2002), and *Cr4* in *P. flexilis* (Schoettle et al. 2014). *Cr1*, *Cr2*, and *Cr4* have been mapped on genetic linkage groups (LGs) using several types of DNA markers, including random amplification of polymorphic DNA (Devey et al. 1995; Liu et al. 2006), amplified fragment length polymorphism (Liu and Ekramoddoullah 2008), disease resistance

†Corresponding author: J.-J. Liu; jun-jun.liu@nrcan-rncan.gc.ca

Current address of Jeremy S. Johnson: Department of Environmental Studies, Prescott College, Prescott, AZ 86301.

Funding: This work was supported in part by the Canadian Forest Service Pest Risk Management Project, the U.S. Department of Agriculture's Forest Service Special Technology Development Program under grant number R2-2014-01, the U.S. National Science Foundation under grant number EF-1442597, and the U.S. Department of Agriculture's Forest Service Gene Conservation Program.

*The ***e*-Xtra** logo stands for "electronic extra" and indicates there are four supplementary figures and supplementary tables published online.

The author(s) declare no conflict of interest.

gene analog markers (Liu and Ekramoddoullah 2007), and single nucleotide polymorphisms (SNPs; Jermstad et al. 2011; Liu et al. 2014, 2016). Comparing *Pinus* consensus maps (Westbrook et al. 2015) by bridging genes of paired orthologs demonstrated that *Cr1*, *Cr2*, and *Cr4* were anchored on LG-2, LG-1, and LG-8, respectively (Liu et al. 2017, 2019; Stevens et al. 2016). This indicates that they have independently evolved with pleiotropic effects, presumably without historic white pine-blister rust coevolution, given that *C. ribicola* is an invasive pathogen that was introduced to North America in the early 1900s (Liu et al. 2016).

P. flexilis is a wide-ranging tree species, growing from Alberta and British Columbia of Canada, south to California, Arizona, and New Mexico, with extensive distributions in the U.S. Rocky Mountains (Colorado, Wyoming, Utah, Idaho, and Montana) and Great Basin. The distribution of *P. strobiformis* ranges from stands in mountainous areas of Arizona, New Mexico (with a scattering of trees in southern Colorado), and Texas, as far south as the Sierra Madre Occidental Mountains of northern Mexico. Most of the U.S. portion of the *P. strobiformis* distribution is a hybrid zone between *P. flexilis* and *P. strobiformis* with gene-flow between the two species, forming the *P. flexilis* complex (Menon et al. 2018, 2019). For *P. flexilis*, MGR has been documented in the southern Rocky Mountain region (Schoettle et al. 2014), and in Alberta, Canada (Sniezko et al. 2016), with presumably little or no gene-flow between these two areas because of geographic isolation since the last glaciation event. An association study revealed that MGR in both areas is controlled by the same *R* locus (*Cr4*) or genetic loci very close to *Cr4* (Liu et al. 2020). However, *Cr3* in *P. strobiformis* has not yet been genetically mapped, and its evolutionary relationship with other MGRs to *C. ribicola*, especially *Cr4*, is unknown.

Transcriptomes have been assembled and compared based on RNA-seq analysis of several white pine species for in silico mining of SNPs (Baker et al. 2018; Liu et al. 2014; Lorenz et al. 2012). Abundant SNPs, documented through RNA-sequencing (RNA-seq) and other next-generation sequencing (NGS) approaches in these species, offered a convenient and cost-effective opportunity for marker applications, facilitating studies of genetic diversity, map construction, development of molecular tools for marker-assisted selection (MAS), and other areas of research (Falk et al. 2018). Nonetheless, no SNPs were available in *P. strobiformis* until a 2018 development using double digest restriction-site associated DNA sequencing (ddRADseq; Menon et al. 2018). Neither a linkage map nor an annotated reference genome or transcriptome is available for *P. strobiformis*, limiting comprehensive understanding of the genomic basis underlying adaptive traits, including *Cr3*-mediated resistance and potential QR to *C. ribicola*. A subset of SNPs was extracted in *P. flexilis* genes conserved in the genus *Pinus* (BLASTn *E* values < 1e-100) and used in comparative genetic mapping in *P. flexilis* populations with different geographical origins (Liu et al. 2019), but the transferability and utilization of SNP markers across different white pine species have not been explored yet.

The primary goals of this study were to assess transferability of SNPs discovered in *P. flexilis* transcriptomes to *P. strobiformis*, and to explore evolution of MGRs in these two closely related white pine species. To meet these research goals, the SNP loci were selected based on their genetic linkage to *Cr4* as well as high conservation of the relevant gene sequences in *Pinus* species (Liu et al. 2020). An association study was performed by genotyping a *P. strobiformis* population with phenotypic segregation of MGR-related traits using a Sequenom MassARRAY iPLEX Genotyping System. TaqMan-based SNP arrays were further developed for MAS, allowing the identification of alleles useful in revealing genotypic differences underlying resistance-related phenotypes conferred by MGR and QR.

MATERIALS AND METHODS

Plant materials and resistance assessment. Open-pollinated seeds were collected from the BRA08 parent tree in 2008 and 2013 as the seed families BRA08 (CY08) and BRA08 (CY13), respectively (Supplementary Table S1; Supplementary Fig. S1), from the same parental tree (BRA08) in Bradford Canyon in Lincoln National Forest (Latitude = 32.97851664; Longitude = –105.7209274; WGS = 84 datum), the first location where MGR was documented in *P. strobiformis* (Kinloch and Dupper 2002). Seed collection was allowed with a National Forest Permit issued by the U.S. Department of Agriculture's Forest Service. Although the seed families were from the same parent tree, they were open-pollinated, and could have a different mix of pollen parents contributing in different years. In this stand there is ~90% infection by WPBR, and many surviving non-cankered trees show MGR or QR genotypes (Johnson and Sniezko 2021; Sniezko et al. 2008). Seeds were sown at Dorena Genetic Resource Center in the spring of 2014 and 2015. Seedlings were inoculated for short-term MGR and longer-term operational trials in September of 2014 and 2015, respectively (Supplementary Table S1). The MGR trial was designed to detect MGR, while the operational trial was designed to detect QR. However, QR may express the stem-canker–free phenotype, which is phenotypically difficult to distinguish from MGR when virulent races are absent. The MGR trial was carried out by inoculating 6-month-old seedlings using basidiospores at high density >6,000 spores/cm^2, while the operational trial was carried out by inoculating 18-month-old seedlings using basidiospores at a lower density of ~3,000 spores/cm^2. Younger seedlings with primary needles are more susceptible, which resulted in many more infections with higher inoculum density. The MGR trials used younger seedlings and a higher spore density to minimize the percentage of seedlings with stem symptom-free traits associated with QR, but it is possible some were present in the MGR trial. The standard Dorena Genetic Resource Center protocols were used for inoculation with *C. ribicola* basidiospores in an inoculation room under controlled conditions at 18°C and 100% relative humidity (Schoettle et al. 2014; Sniezko et al. 2016). Given that no virulent races able to overcome *Cr3* or *Cr4* have yet been detected, the races used in inoculation trials were putatively avirulent races (*Avcr3*, *Avcr4*), collected in Oregon where *P. strobiformis* is not native and *P. flexilis* is rare.

After inoculation, WPBR disease development was assessed from October of the inoculation year to December of the following year, with at least three separate assessments for the MGR trials, and annually for 6 years to detect QR-related traits for the operational trials. Stem canker development was used as the main criteria to sort resistant from susceptible phenotypes (Schoettle et al. 2014; Sniezko et al. 2016). Using χ^2 tests, phenotypic segregation was estimated to determine if MGR-related traits were consistently expressed by comparing results between seed families from different trials as well from different inoculation years. Needle samples were collected from each seedling and stored at –20°C before DNA extraction.

Genomic DNA extraction and SNP genotyping. Nuclear DNA was extracted from 10 to 20 mg of dry needle tissue per individual seedling. Needle tissues were transferred into individual lysing matrix-A 2-ml tubes (MP Biomedicals, Irvine, CA) and an additional ceramic sphere was added. Tissues were homogenized using a FastPrep-24 instrument (MP Biomedicals) for 2 × 20 s pulses and 1 × 15 s pulse to fully grind the tissue into a fine powder. Tubes were frozen in liquid nitrogen before grinding and at intervals between tube-shaking. For DNA extraction, a DNeasy Plant Mini kit was used (QIAGEN, Hilden, Germany), following the manufacturer's instructions in general. Small modifications to the extraction protocol were made by increasing the incubation time of the samples in lysis buffer at 65°C and using a prewarmed elution buffer at 65°C to improve elution of gDNA from the column during the last step.

Integrity and purity of extracted genomic DNA were determined with a NanoDrop 2000/2000c Spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and agarose gel electrophoresis. A *P. flexilis* SNP assay with 72 SNPs developed in a previous *Cr4*-association study (Liu et al. 2016, 2020) was used for *P. strobiformis* genotyping in this study. This SNP set was localized in

P. flexilis genes conserved in the genus *Pinus* (BLASTn *E* values < 1e-100), providing potential for application in comparative genetic mapping in related *Pinus* species. The genetic position of these *Cr4*-linked genes on the *Pinus* consensus LG-8 had been determined in Liu et al. (2019, 2020). Based on flanking sequence of SNP loci, primers and probes (Supplementary Table S2) were designed and synthesized for SNP genotyping using a Sequenom MassARRAY iPLEX platform (San Diego, CA; Gabriel et al. 2009).

Association mapping of *P. strobiformis* disease resistance to WPBR. SNP loci with minor allele frequency (MAF) > 5% were extracted to estimate linkage disequilibrium (LD) by calculating the squared allele frequency correlations (r^2) for all pairwise combinations of SNP loci. Trait-SNP genotype association was examined using the general linear model (GLM) and mixed linear model (MLM) in the software TASSEL v.3.0 (Bradbury et al. 2007). Because population stratification or individual relatedness may cause spurious marker-trait associations, the $Q + K$ method was used in MLM, incorporating the Q matrix for population structure and kinship (K) matrix for average relationship between individuals (Eu-ahsunthornwattana et al. 2014; Yu et al. 2006). Significance of association tests was determined with a P value threshold by Bonferroni-correction at $\alpha = 0.05$.

In addition, allele frequency directional difference (AFDD) mapping (Fekih et al. 2013; Takagi et al. 2013) was used to search for *Cr3*-associated SNPs. The tool GenAlEx (Peakall and Smouse 2012) was used to evaluate allele frequency and Hardy-Weinberg equilibrium at each SNP locus across the total samples as well as across all seedlings of two groups: one with all resistant seedlings, and another with all susceptible seedlings. AFDD values were filtered with a threshold of 30 percentage points for the informative variants (Dougherty et al. 2018; Fekih et al. 2013), and those SNP loci were selected for a statistical test of phenotypic associations using X^2 tests to compare allele frequency differences between the two contrasted phenotypes. For mapping the *Cr3*-conferred resistant phenotype, the calculated AFDD values between resistant and susceptible phenotypes were plotted against the genetic positions of the polymorphic genes mapped previously to the *Pinus* consensus LG-8.

The multifactor dimensionality reduction statistical analysis for detecting and modeling gene–gene interaction. The multifactor dimensionality reduction (MDR) method was applied to identify SNP–SNP interactions (epistasis) using the software program MDR v.3.0.2 (Ritchie et al. 2001) with the following configurations: random seed = 10; attribute count range = 1 to 3; 10-fold cross-validation consistency (CVC); track top models = 10; and search type = exhaustive. The best model was selected on the basis of maximum values of CVC and testing balance accuracy.

Development of TaqMan-based SNP genotyping assays as MAS tools. The flanking sequences for SNPs, selected based on genotypic data from Sequenom SNP genotyping, were submitted to Custom TaqMan Assay Design Tools at Thermo Fisher Scientific (https://www.thermofisher.com/order/custom-genomic-products/tools/cadt/), where appropriate primers and TaqMan probes were designed (Supplementary Table S3). TaqMan minor groove binder probes were labeled with fluorescent dyes 6-carboxyfluorescein for allele-1 and VIC (Life Technologies, Carlsbad, CA) for allele-2, respectively, in the customized TaqMan SNP Genotyping Assays (Supplementary Table S3).

DNA samples were diluted in Milli-Q water (EMD Millipore/MilliporeSigma, Burlington, MA) to a final concentration of 20 ng/µl. Quantitative PCR (qPCR) was performed in a 10-µl final volume per PCR reaction, each containing 1 × TaqMan Genotyping Master Mix, 1 × Custom TaqMan SNP Genotyping Assay, and 20 ng of genomic DNA. A 7500 Fast Real-Time PCR system (Applied Biosystems, Waltham, MA) was used for qPCR runs with 96-well plates, and PCR conditions consisted of a hot-start activation at 94°C for 15 min, followed by 40 cycles at 94°C for 20 s and 60°C for 60 s. Negative controls (no template with water) were also included in each run of qPCR. TaqMan Genotyper Software (Applied Biosystems) was used to call genotypes at each SNP locus for each sample. To assess TaqMan assay-based selection for disease resistance, the ratio of the number of individual samples with agreement between genotypes and phenotype to the number of total individuals were calculated as genotype–phenotype matching rate.

Orthology analysis of the transferable SNP loci. An orthology analysis was performed to determine sequence similarity between *P. flexilis* and *P. strobiformis* genes. RNA-seq reads (SRX10204300 and SRX10204301) of the *P. strobiformis* megagametophyte tissues (Jin et al. 2021), released by GenBank on 4 June 2021, were downloaded and used for read mapping against *P. flexilis* gene sequences with targeted SNP loci using the tool CLC Genomics Workbench (v5.5; QIAGEN) with the following parameters: masking mode = no masking; mismatch cost = 2; insertion cost = 3; deletion cost = 3; length fraction = 0.9; similarity fraction = 0.95; global alignment = yes; and nonspecific match handling = ignore. The *P. strobiformis* consensus sequences from mapped RNA-seq reads (Supplementary Table S4) were extracted and used for BLASTn search (Altschul et al. 1990) to calculate sequence identities between *P. flexilis* and *P. strobiformis*. The software BLAST2GO (Götz et al. 2008) was used to annotate the associated SNPs and polymorphic genes based on sequence homologies to the available databases (NCBI-nr, PIR, KEGG, and GO).

RESULTS

Phenotypic segregation of resistance trait. In the MGR trials, different sets of 6-month-old seedlings were inoculated independently in Septembers 2014 and 2015. After inoculation, resistance-related traits were assessed in the next summer based on cankered stem versus symptom-free stem. In the operational trials, two independent seed families (CY08 and CY13) of the same parent (collected in different years) were sown in 2014 and 18-month-old seedlings were inoculated in the fall of 2015, with six annual phenotypic assessments through 2020. A total of 158 seedlings were phenotypically evaluated and sampled for SNP genotyping, with 78 and 80 seedlings from the MGR and operational trials, respectively (Supplementary Table S1). Resistant seedlings with stem symptom-free traits accounted for 64.29 to 77.5% of the total in different trials, indicating that the parental tree was *Cr3/cr3* heterozygote. Although MGR trials detected phenotypic segregation ratios (stem symptom-free versus cankered stem) that fit with the expected Mendelian segregation ratio of 1:1 in different seed families in different years, in the total samples the phenotypic segregation ratios significantly deviated from 1:1, with higher proportions of resistant (stem symptom-free) than susceptible seedlings (stem-cankered) in both MGR and operational phenotyping trials. The operational trials detected a higher frequency of stem symptom-free seedlings than the MGR trials (76.25 versus 67.95%). The phenotyping assessment results suggest that ∼8.3% of total seedlings might have QR-conferred stemsymptom-free phenotypes in operational trials. Consistent phenotypic segregation indicates that stem symptom-free traits were stably expressed in two *P. strobiformis* seed families under different inoculation conditions.

Transferability of *P. flexilis* SNP assays in *P. strobiformis* genotyping. Using the Sequenom MassARRAY iPLEX arrays with 72 SNPs designed for *P. flexilis* (Supplementary Table S2), 50 SNPs (69.4% of the total) were successfully genotyped in *P. strobiformis*, with missing data <20%. This was slightly, but not significantly lower (X^2 test, $P = 0.348$) than successfully genotyped SNPs (76.4%) in *P. flexilis* populations (Liu et al. 2020). Using RNA-seq reads of *P. strobiformis* megagametophyte tissues available from GenBank, read mapping followed by BLASTn analysis of the *P. strobiformis* consensus sequences from read mapping revealed 60 of 72 *P. flexilis* SNP sequences with orthologous sequences expressed in *P. strobiformis* megagametophytes (Supplementary Tables S4 and S5). These orthologous pairs showed nucleotide identities from 88 to 100% between *P. flexilis* and *P. strobiformis*.

The transferable SNPs had nucleotide identities (average 98.97 ± 1.35%) significantly higher (*t* test, $P = 0.035$) than the failed and nontransferable SNPs (average 97.89 ± 3.25%). Sequence similarity between *P. flexilis* and *P. strobiformis* appeared to be an important factor affecting the SNP transferability from one species to another.

The r^2 statistic was used to estimate levels of LD between SNPs along LG. LD levels were low (average $r^2 = 0.048 \pm 0.104$) for all detected SNP pairs along the LG in the *P. strobiformis* open-pollinated seed families from parent BRA08 (Fig. 1A). The SNP pairs of the same genes or different genes with the same genetic positions on the LG showed complete or high LD ($r^2 = 1$). In filtering LD with significance ($P < 0.05$), SNP pairs had the average LD at the highest level (average $r^2 = 0.49$) within genetic distance of 1 cM. However, LD decayed significantly (one-way analysis of variance, $P = 1.23e-08$) to low levels (average r^2 ranging from 0.19 to 0.12) beyond 1 cM, without significant LD decay as genetic distances increased >1 cM (Fig. 1B). The fast LD decay indicated that the seed families were open-pollinated with a complex pollen source of high genetic diversity.

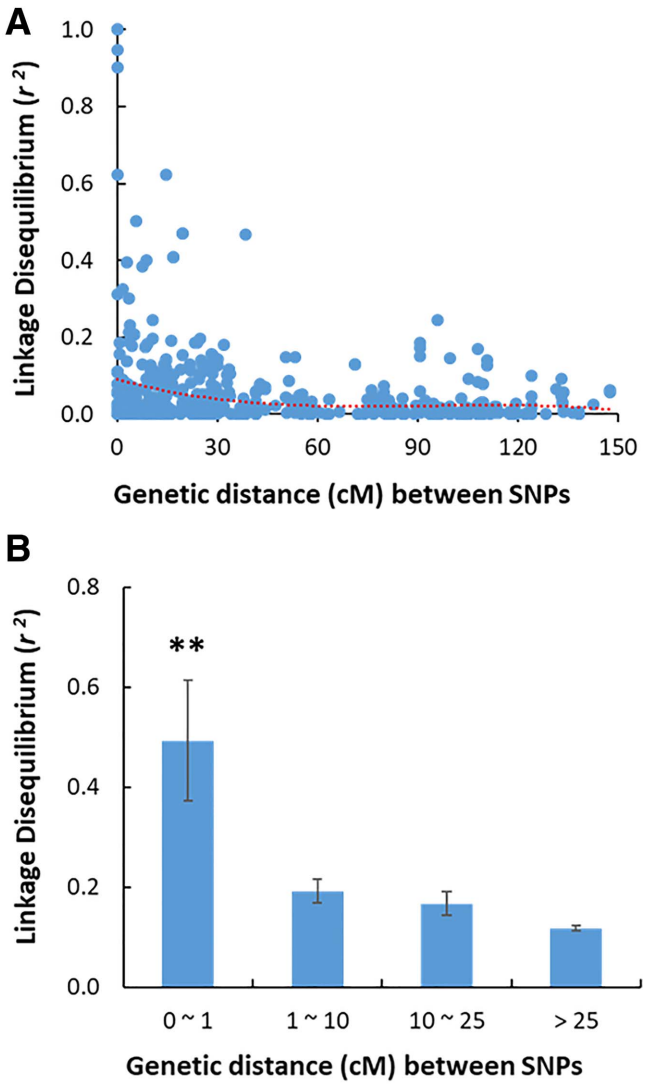

Fig. 1. Plotting the estimates (r^2) of linkage disequilibrium (LD) decay against genetic distance (in centimorgans, cM) between single nucleotide polymorphism (SNP) loci in southwestern white pine families. **A,** Linkage group (LG)-wide LD r^2 values plotted against the genetic distance between each intrachromosomal SNP pair on the *Pinus* consensus LG-8. The dashed red curve is a nonlinear regressions for expected decay of LD (r^2) in all pairs of SNP loci. **B,** Comparison of average LD decay estimates (r^2) based on genetic distances between two SNP loci. Estimate of LD decay (r^2) was grouped and compared among four intervals of genetic distances (1, 10, 25, and >25 cM). Two asterisks indicate significant difference by one-way analysis of variance at $P < 0.01$.

Association mapping. Both a GLM and an MLM were used to detect SNP markers in association with MGR-related traits. Of 50 genotyped SNPs, five SNP loci of four polymorphic genes (M326511_1126R, M416463_1129Y, M355530_134M, M355530_415R, and M438219_3042M) showed significant association with the resistant phenotype with marker effect (R^2) values ranging from 0.166 to 0.522 (Bonferroni-corrected $P < 0.05$). These five MGR-associated SNPs, from four unigene, were mapped on the *Pinus* consensus LG-8 map at positions from 117.85 to 151.41 cM with *P. flexilis Cr4* at 126.80 cM (Liu et al. 2019, 2020). Manhattan plot showed –log10 (*P* value) of the association statistical significance on the genetic map of LG-8 (Fig. 2A), revealing that the SNP of M326511_1126R had marker-trait association at the highest level of significance ($R^2 = 0.527$, $P = 1.45E-24$). This strongest association of M326511_1126R with resistant traits was further confirmed by MLM-based association analysis ($R^2 = 0.421$, $P = 4.78E-12$), followed by M416463_1129Y ($R^2 = 0.391$, $P = 3.97E-07$), M355530_134M ($R^2 = 0.269$, $P = 2.04E-05$), and M438219_3042M ($R^2 = 0.219$, $P = 1.36E-04$). In contrast, the SNP M355530_415R, detected by GLM, was not significant in marker-trait association as analyzed by MLM (Fig. 2B). M326511_1126R was mapped at 126.36 cM on the LG-8, very close to *Cr4* (126.80 cM) on the *P. flexilis* LG-8. M326511 showed 100% nucleotide sequence identity of the gene-coding region between *P. strobiformis* and *P. flexilis*.

AFDD mapping. Association of *Cr4*-linked SNPs with the *Cr3*-phenotype was further supported by AFDD mapping. Seven SNPs showed AFDD > 30 with significant difference between two seedling groups of resistant and susceptible phenotypes (χ^2 test, $P = 7.32e-06 \sim 1.49e-17$); including five SNPs detected by the GLM run, and two additional SNPs M463406_1290R and M286987_932Y (Fig. 2C). M326511_1126R had the highest AFDD value of 54.2, followed by M355530_134M/415R and M416463_1129Y with AFDD values at 44.2 and 43.6, respectively, closely positioned at either proximal or distal side of the *Cr4* locus (Fig. 3C). A similar pattern was observed when χ^2 tested *P* values were plotted against the genetic positions of the markers on the *Pinus* consensus LG-8, with M326511_1126R as the peak (Supplementary Fig. S2). As expected, those significant AFDD values of the associated SNPs showed negative correlation (Pearson correlation coefficient $R = –0.7374$, $P = 0.009$) with their genetic distances (in centimorgans) to the *Cr4* locus (Supplementary Fig. S3). These results implied that *Cr3* and *Cr4* might be an orthologous MGR locus, or located very close to each other on the *Pinus* consensus LG-8.

Impact of associated SNPs on gene functions. All seven SNPs identified above showing association with *Cr3/Cr4*-controlled MGR phenotypes were localized in the coding region of a unigene. Three of them (M416463_1129Y, M438219_3042M, and M355530_415R) were synonymous SNPs, not affecting putative functions of the proteins encoded by the corresponding genes. M416463 and M438219 encoded proteins with homologies to endoglucanase 16-like protein and nuclear matrix constituent protein 1-like protein, respectively. In contrast, four other SNPs were nonsynonymous SNPs with amino acid changes, which might impact putative functions of the proteins. They included M326511_1126R for $Val_{376}Ile$ in a pectin acetylesterase (PAE) encoded by M326511; M355530_134M for $Met_{12}Leu$ in a transport inhibitor response 1-like protein encoded by M355530; M463406_1290R for $Val_{407}Ile$ in a protein with domains of nucleotide-binding site and leucine-rich repeats (NLRs) encoded by M463406; and M286987_932Y for $Ile_{311}Thr$ in a pyrimidine reductase encoded by M286987.

Six SNP loci targeted in two putative NLR genes (M463406 and M373574) were included in the SNP array (Supplementary Table S1). *P. strobiformis* and *P. flexilis* shared 99.83 and 92.08% nucleotide identities for M463406 and M373574, respectively. Only one of them (M463406_1290R) was associated with *Cr3*-controlled phenotype with a MAF = 0.42 and AFDD = 36.4 (χ^2 test, $P = 3.54E-07$). Of the

other five, two (M373574_318Y and M463406_171K) were not transferable; two (M373574_94K and M463406_337Y) were detected with *MAF* = 0; and one (M373574_220R) had a MAF = 0.08, probably with minor allele contribution as a result of open pollination.

Genetic interactions between *Cr3*-associated SNPs. MDR analysis was implemented to investigate if SNPs interact with each other (Supplementary Table S6). Consistent with association analysis and AFDD mapping, M326511_1126R was the best single-locus model to predict *Cr3*-confered WPBR resistance (CVC = 10/10; Testing balanced accuracy = 0.942; $P = 0.016$). A combination of M326511_1126R and M416463_1129Y fit in the best two-locus model (CVC = 9/10; Testing balanced accuracy = 0.942; $P = 0.016$) while the best three-locus model consisted of M326511_1126R, M355530_134M, and M416463_1129Y (CVC = 10/10; Testing balanced accuracy = 0.948; $P = 0.018$). Not surprising, little additive effect between the models was detected for MGR against *C. ribicola*, and SNP–SNP interactions were negligible (blue lines) with independent effects of the impact (Supplementary Fig. S4A). However, a combination of three SNP loci increased predicative value for selection of *Cr3/Cr4*-conferred MGR (Supplementary Fig. S4B).

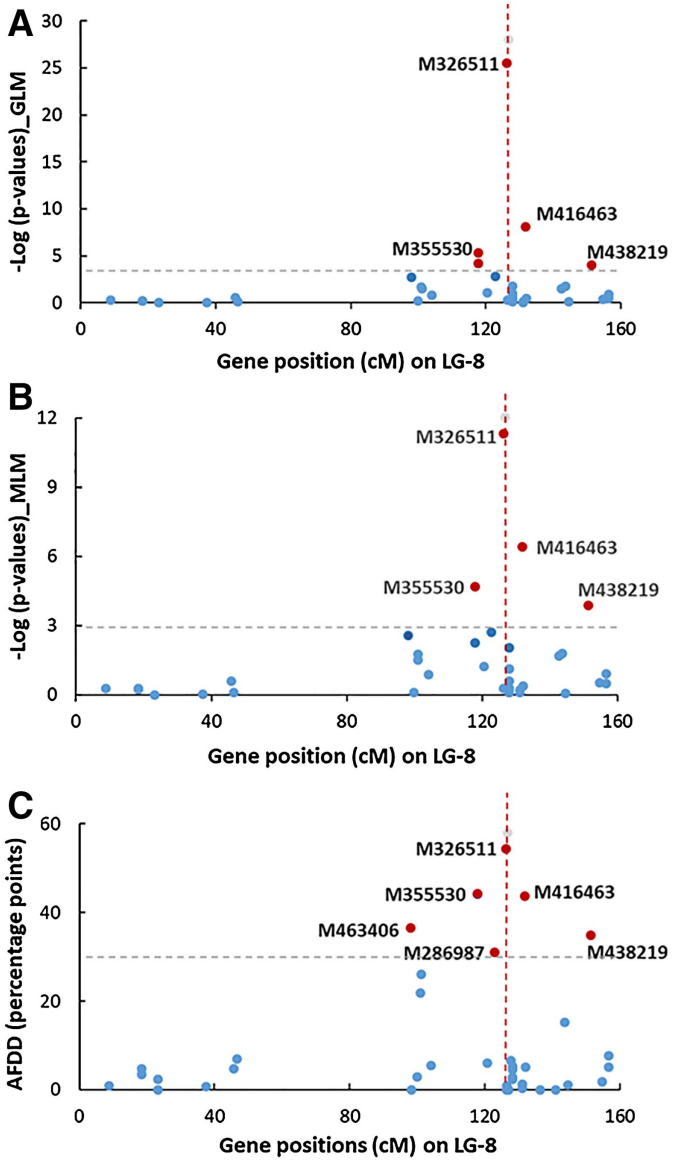

Fig. 2. Association analyses of single nucleotide polymorphism (SNP) genotypes with phenotypic traits conferred by southwestern white pine major gene (*Cr3*) resistance to *Cronartium ribicola*. The Manhattan plots of association levels (log10 *P* values) against genetic positions of SNP loci on the *Pinus* consensus linkage group 8 (LG-8). The *x*-axis represents SNP locations (in centimorgans) on LG-8 while the *y*-axis represents –log10 *P* values of association analyses with the phenotypic-resistant traits. The red-dashed vertical lines mark the genetic position of *Cr4* as mapped in limber pine in Liu et al. (2020). The gray-dashed horizontal lines mark the thresholds for significant association. Genes with significant associations are labeled. **A,** Association test using a general linear model (GLM). **B,** Association test using a GLM. **C,** Association test using genetic mapping based on allele frequency directional difference (AFDD).

Development of TaqMan arrays as MAS tools to distinguish MGR and QR genotypes. TaqMan assays were designed based on sequences around SNP loci M326511_1126R and M355530_415R (Supplementary Table S3). TaqMan-based SNP genotyping confirmed genotypic results as detected by the Sequenom MassARRAY iPLEX platform, and both genotyping technologies were highly reliable in *P. strobiformis*. Genotypic frequencies of A/A, A/G, and G/G were in Hardy-Weinberg equilibrium at both SNP loci (χ^2 test, $P = 0.7877$ at M326511_1126R, χ^2 test, $P = 0.4047$ at M355530_415R), suggesting the population size of this study is large enough with little sampling effect.

At the M326511_1126R locus, the BRA08 parent tree was an A/G genotype, and allele A was associated with resistant allele *Cr3*, while allele G was associated with susceptible allele *cr3*. Of its open-pollinated progeny, genotypes A/A and A/G were linked to the resistant phenotype while G/G was linked to the susceptible phenotype (Fig. 3A).

Because we hypothesized that the resistant phenotypes detected in the operational trial were contributed by genotypes of both MGR (*Cr3*) and QR, we performed a comparison of genotype-phenotype matching rates between the MGR- and operational trials (Fig. 3B). The genotype-phenotype matching rate was 93.6% in all seedlings from MGR trials, with 96.0 and 92.5% in susceptible and resistant phenotypes, respectively. As compared with MGR trials, operational trials showed the genotype-phenotype matching rate significantly lower in all seedlings (84.0 versus 93.6%; χ^2 test, $P = 0.011$) and in resistant seedlings (81.4 versus 92.5%; χ^2 test, $P = 0.022$), but no significant difference in susceptible seedlings between two types of inoculation trials (93.8 versus 96.0%; χ^2 test, $P = 0.953$). Because there is little question in phenotypic assessment of susceptibility during either MGR or operational trials, our results demonstrated that the TaqMan assay of M326511_1126R had an accuracy of 95.1% for selection of *Cr3* in both BRA08 (CY08) and BRA08 (CY13) families.

Of all resistant seedlings from both seed families of BRA08 (CY08) and BRA08 (CY13), the M326511_1126R genotype G/G accounted for 7.5% in the MGR trial, while it increased to 18.6% in the operational trial (Fig. 3A). Comparison of two seed families revealed that all susceptible seedlings of family BRA08 (CY13) had M326511_1126R genotype GG as assessed in both MGR and operational trials while 95.0 and 87.5% of susceptible seedlings of family BRA08 (CY08) had genotype G/G in MGR and operational trials, respectively (Fig. 3C and D). Percentages of the M326511_1126R genotype G/G in resistant seedlings showed no significant different between two seed families (7.0 versus 9.1%) in the MGR trials (Fig. 3C), but the percentage was much higher in BRA08 (CY08) than in BRA08 (CY13; 25.0 versus 12.9%; X^2 test, $P = 0.03$) in operational trials (Fig. 3D), suggesting a significant shift of genetic compositions between seed families collected from the same tree but in different years. Therefore, we hypothesize that for the increased proportion of resistant seedlings, mainly in family BRA08 (CY08), during the operational trials, it was likely QR rather than MGR that contributed the stem symptom-free traits of the resistant phenotypes. Thus, the resistant seedlings identified with the G/G genotype at the M326511_1126R locus may be a valuable resource to pinpoint QR genotypes for durable resistance to WPBR in future *P. strobiformis* breeding.

DISCUSSION

SNP transferability and consideration in design of genotyping arrays for cross-species application. *P. strobiformis* is considered an important component of mixed-conifer forests of the U.S. Southwest. However, in this nonmodel conifer species, available genomic

resources are very limited; neither a reference genome nor transcriptome have been reported. Although SNPs have been developed throughout the *P. strobiformis* genome by using ddRADseq (Menon et al. 2018), no genotyping assay has been designed yet with functional SNP annotation for comparative genetic mapping. This study tested and confirmed the transferability of a set of SNP markers between two white pine species of the *P. flexilis* complex. Overall, *P. flexilis* SNP markers showed high transferability and informativeness in *P. strobiformis* despite only open-pollinated progeny of one parental tree having been genotyped with limited sample sizes. We detected high transferability of cross-species genotyping, with 69.4% successful across total loci, making it possible to design SNP assays to cover a high proportion of highly polymorphic markers. The successful genotyping rate and marker informativeness in the cross-species genotyping were similar to those in the marker source species of *P. flexilis* (Liu et al. 2020), leading to identification of genotypes with significant phenotypic association in subsequent analyses of resistance genetic mapping.

An orthology analysis based on *P. flexilis* transcriptomes and *P. strobiformis* RNA-seq reads available from GenBank allowed us to determine which factors might affect SNP transferability between closely related white pine species and whether any of transferable SNPs have functional impacts on the proteins encoded by the polymorphic genes. We found that the transferable SNPs had flanking sequences with average nucleotide identities higher than the non-transferable SNPs (98.97 ± 1.35% versus 97.89 ± 3.25%), providing a useful index for selection of SNPs to design genotyping assays with potential applications across multiple species of white pines.

The pattern of marker distribution throughout the genome is another factor that may affect cross-species transferability. Because amplification primers and/or hybridization probes designed for genotyping were more conservative across species and usually unique for the targeted loci, genic markers were probably more transferable with good reproducibility and increased resolution as compared with DNA markers in nongenic or random genomic regions (Poczai et al. 2013). Despite the small size of this genotyping assay, all 72 SNP loci were localized within putative protein-coding regions with mapping positions targeted on only one LG, thus the informativeness of the SNP assay was a key prerequisite for final identification of SNPs and polymorphic genes significantly associated with *Cr3*-controlled phenotypes. Genic markers are usually less polymorphic than random genomic ones, potentially limiting their utilization in genomic selection of quantitative traits (Abdollahi-Arpanahi et al. 2016). Nonetheless, genic SNPs are favorable for design of SNP assays and for targeted-NGS to search for functional markers and mutant-causing genes (Giral et al. 2018). Some genic SNPs within coding regions may cause amino acid changes of the encoded proteins while other genic SNPs within regulatory elements (such as promoter, enhancer regions, and long noncoding RNAs) affect expression of the causal genes; these may be better suited to identify significant association with informative annotations in a genome-wide association study (DeFaveri et al. 2013; Oetjen and Reusch 2007).

In addition to genetic mapping and association studies of disease resistance, confirmation of transferability of genomic resources among closely related species also facilitates the application of molecular markers in other studies, such as dissection of genetic architecture of other adaptive traits of economic and ecologic importance in a conifer hybrid zone formed between *P. strobiformis* and *P. flexilis* (Menon et al. 2021). Large SNP assays need to be developed based on in silico SNPs mined in *P. flexilis* and *P. strobiformis*

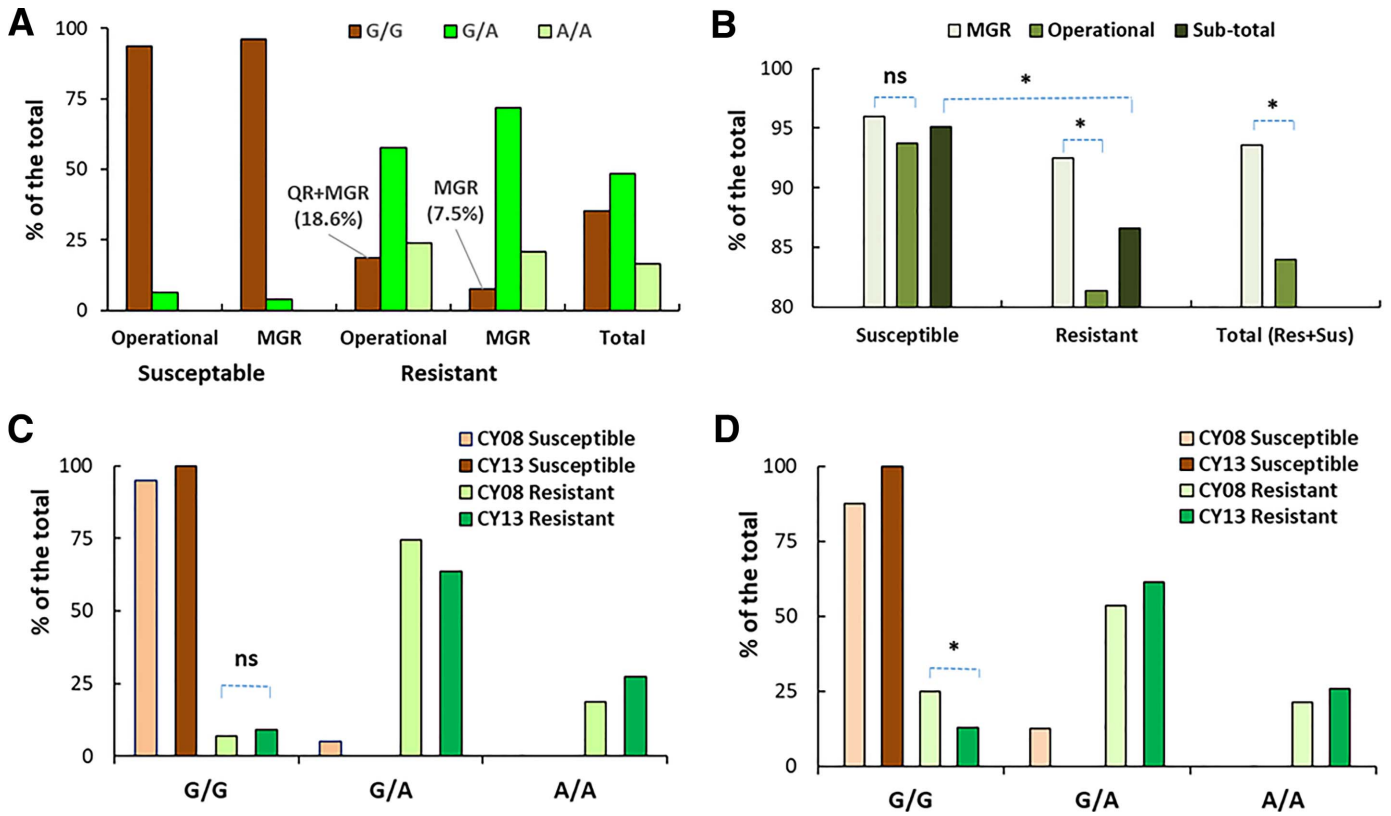

Fig. 3. Genotyping single nucleotide polymorphism loci M326511_1126R using a TaqMan assay. **A,** Genotypic frequencies in resistant and susceptible seedlings evaluated in major gene resistance (MGR) trials and operational trials. MGR trials detected resistant traits expressed exclusively by MGR phenotype while operational trials detected resistant traits expressed by genotypes of both MGR and quantitative resistance (QR) as inoculated avirulent races (*avcr3/4*). **B,** Genotype-phenotype matching rates for estimating accuracy for phenotypic prediction. **C,** Comparison of genotypic and phenotypic compositions between seed families BRA08 (CY08) and BRA08 (CY13) in the MGR trials; ns, not significant. **D,** Comparison of genotypic and phenotypic compositions between seed families BRA08 (CY08) and BRA08 (CY13) in the operational trials. An asterisk indicates significant difference between MGR trials and operational trials (χ^2 test, $P < 0.05$).

through various NGS approaches (Liu et al. 2016; Menon et al. 2018), especially by including those within candidate genes underlying phenotypic traits of interest with successful applications in population and conservation genetic studies in the *P. flexilis* complex (Menon et al. 2021). These invaluable genomic resources and tools allow for further examination of their cross-species transferability and utilization in studies of closely related native white pine species in North America.

Evolution of MGR to WPBR in *P. strobiformis* and other white pines. There is a dynamic "arms race" during plant–microbe interactions (Yang et al. 2013), resulting in adaptation of new R genotypes for plants to fight against new pathogenic races mainly through expansion and rearrangement of gene members within *R* loci (Leister 2004; Marone et al. 2013). Because *C. ribicola* was introduced to North America only ~100 years ago, *Cr1* to *Cr4* appeared to be ancient *R* genes that had occurred before native white pines were exposed to this exotic fungus; thus evolutionary dynamics would be actually more complex than a simple "arms race" model for the WPBR pathosystems in North America. On the other hand, emergence of virulent *C. ribicola* races *vcr1* and v*cr2*, which overcome *Cr1* and *Cr2*, respectively (Kinloch and Dupper 2002), probably resulted from selection pressure, considering that they were first detected at field sites where most susceptible trees were killed and MGR trees became dominant (Kinloch 1970; Kinloch et al. 1999).

Our comparative association study implied that *Cr3* and *Cr4* might be the same orthologous *R* locus, or different alleles very close to each other within the syntenic genomic region at least for the U.S. part of the range for *P. strobiformis*. MDR analysis showed no gene–gene interaction among *Cr3/Cr4*-associated SNPs, providing additional evidence that *Cr3/Cr4* locus is a single dominant *R* locus with resistant traits stably expressed in independent inoculation trials of seed families from one *P. strobiformis* seed tree. Despite MGR conservation between *P. strobiformis* and *P. flexilis*, we found that LD decayed faster along LG-8 in *P. strobiformis* than in *P. flexilis* despite sample types and sizes being similar in both studies (Liu et al. 2020), implying that they might have been subjected to different evolutionary forces (such as natural selection, mutation, and recombination rates). Because the U.S. part of the range is possibly the "ancient" hybrid zone for the *P. flexilis* complex (Menon et al. 2018), the question of whether *P. strobiformis* MGR is the same in the Mexican range, where there is no species hybridization, will have to be confirmed with utility of available SNP markers.

In contrast to *Cr3/Cr4* conservation, localizations of *Cr1*, *Cr2*, and *Cr3/Cr4* on different *Pinus* consensus LGs indicated an independent evolution of these three R loci. They may have functions other than genetic resistance to WPBR that evolved before *C. ribicola* was introduced through *sensu* exaptation (Gould and Vrba 1982). In addition, research suggests that *Cr4* in *P. flexilis* has a negligible fitness cost (Vogan and Schoettle 2016) and has a fitness benefit with respect to abiotic stress tolerances in *P. flexilis* (Vogan and Schoettle 2015), which may contribute to the much higher frequency of *Cr4* in predominantly naïve populations of *P. flexilis* compared with *Cr1* and *Cr2* in populations of *P. lambertiana* and *P. monticola*, respectively (Schoettle et al. 2014). *C. ribicola* virulent *vcr1* and *vcr2* races are also distinct, with different frequencies and spreading patterns in populations (Kinloch et al. 2004; Kinloch and Dupper 2002). Previous research has shown that *Cr4* in *P. flexilis* is not overcome by *vcr2* (Schoettle et al. 2014), contributing additional evidence to suggest that the genetic basis of *Cr3/Cr4* is different from that underlying MGRs in the other two white pine species.

The presence of functionally unique *Cr* genes, each with resistant and susceptible alleles that would have evolved over many thousands of years, raises the question of how these *R*-genes have been selected and how their polymorphic alleles have been maintained. The *Cr4* seed trees identified so far are not in the *P. flexilis–P. strobiformis* hybrid zone with distribution further north, even to Alberta, Canada. If *Cr3* and *Cr4* are confirmed to be the same *R* orthologous locus in the *P. flexilis–P. strobiformis* hybrid zone, as our research here suggests, it would be interesting to ascertain whether this conserved *R* locus is an old ancestral gene or had recently introgressed from one species to another. For example, higher drought and freeze tolerance of *P. strobiformis–P. flexilis* hybrid populations was likely generated through recent introgression of *P. flexilis* alleles into the *P. strobiformis* genomic background (Menon et al. 2021). A comprehensive comparison of the gene sequences and structures of the *Cr3/Cr4* region between *P. flexilis* and *P. strobiformis*, as well as their hybrid populations, would be an optimal way to address these questions and give insights into the long-term evolutionary dynamics of *R*-genes in the wild stands of the *P. flexilis* complex from wide geographical areas.

R genes of the NLR family are well characterized for their roles conferring MGR in various angiosperm plants, and they also have pleiotropic effects with other fitness benefits in the absence of pathogenic diseases (MacQueen et al. 2016), such as enhanced drought tolerance (Chini et al. 2004). Unlike other *Pinus* MGR loci where a number of NLR genes were genetically mapped in the syntenic regions, *R* candidates of the NLR family mapped at *Cr3/Cr4* locus were relatively limited, despite extensive NLR mapping efforts throughout the *P. flexilis* genome using whole exome-sequencing and targeted-sequencing of the NLR family (Liu et al. 2019, 2021). Our study revealed that one NLR gene (M463406) was highly conserved while another (M373574) was divergent between *P. strobiformis* and *P. flexilis*. BLAST analyses comparing the sequence data flanking genotyped SNP loci of other functional genes between the two species also confirmed the presence of dramatic genome divergence between species. Targeted NGS approaches, such as the *R* gene enrichment sequencing (or "RenSeq"), have been widely used to identify allelic variations of candidate *R* genes against various pathogens and pests in model and agricultural plant species (Jupe et al. 2013). The genomic variation documented here paves the way for future work to investigate the evolutionary mechanisms and adaptation of WPBR-resistance in the syntenic *Cr3/Cr4* region.

Previous studies have identified NLR genes as candidates in WPBR resistance (Liu et al. 2014; Stevens et al. 2016). However, we cannot exclude the possibility that *Cr3/Cr4* belongs to a non-NLR gene family, despite the prevalence of case studies with functionally confirmed *R* genes belonging to the NLR family. Our work identified a nonsynonymous SNP of a PAE homolog with the highest significance for association with *Cr3*-controlled resistance traits. PAEs are pectin-modifying enzymes catalyzing the deacetylation of pectin. Several studies have demonstrated modification of cell wall properties as having significant impacts on cell wall-mediated resistance. Over-expression of a fungal hemicellulose- and pectin-specific acetylesterase in transgenic plants led to increased resistance to microbial pathogens as the degree of polysaccharides acetylation decreased (Pogorelko et al. 2013). Arabidopsis PAE9 is required for constitutive up-regulation of defense-related compounds through its activity in oxidation-reduction reactions (Kloth et al. 2019). Fine cluster organization of NLRs and candidates belonging to other gene families at the *Cr3/Cr4* locus can be dissected by completely sequencing the targeted genomic regions independently in resistant and susceptible trees of both *P. flexilis* and *P. strobiformis*. Information obtained from this study using comparative genetic mapping as well as sequence comparison of both transferable and nontransferable SNPs will help functional confirmation of DNA variants and gene alleles that have indispensable roles in MGR, and contribute to a comprehensive understanding of short-term dynamics during molecular interactions of white pines with *C. ribicola*.

MAS tools to distinguish genotypes between MGR and QR. QR appears to be more durable than MGR at many field sites for long-term WPBR management (Sniezko et al. 2020). However, development of MAS tools for accurate prediction of MGR-related phenotypes and other complex traits conferring QR to WPBR is

quite challenging because of the huge genome size of white pines and availability of limited genetic resources. With this in mind, SNPs mined from comparative transcriptomics were selected to design SNP assays for genotyping white pine species. Validation and application of genic SNPs developed from transcriptomes probably have an increased chance to reveal DNA variations linked to causal genes that contribute to phenotypic variation of the morphological traits with economic or ecological interests, such as durable and stable resistance to environmental biotic or abiotic factors. In the context of white pine resistance to *C. ribicola*, distinguishing MGR genotypes from QR-related genotypes is important for molecular characterization of different resistance mechanisms, and for breeders to pinpoint germplasm resources as well as to develop either MAS or genomic selection tools.

Since the first report of WPBR infection in natural *P. strobiformis* stands in 1990 (Hawksworth 1990), *C. ribicola* continues to spread in New Mexico and Arizona landscapes (Looney et al. 2015), threatening the sustainability of *P. strobiformis* as a significant component in many high elevation ecosystems (Goodrich et al. 2018). Controlled seedling screening trials for *P. strobiformis* MGR and QR to WPBR has been carried out with significant progress in 2018 (Sniezko et al. 2018). Because of the relatively low frequencies of MGR and QR genotypes across *P. strobiformis*' geographical range, MAS tools may expedite seed family screening by genotyping parental trees and family progeny for planting resistant genotypes (Schoettle and Sniezko 2007). Previous studies in *P. flexilis* found that most *Cr4*-linked SNPs were family-dependent, suitable for resistance screening in the progeny of the tested families with specific genomic background and geographical origins (Liu et al. 2020). This study identified SNP locus M326511_1126R as the best marker associated with the *Cr3*-controlled stem symptom-free trait in the *P. strobiformis* two seed families of parent BRA08; we do not yet know if it is family-dependent. This marker was originally mapped closely to *Cr4* with a genetic distance of 0.44 cM in two *P. flexilis* families (LJ-112 and PHA-106) that originated in the southern Rocky Mountains, and one family (#6665) from Alberta (Liu et al. 2016, 2020). However, this SNP locus had no polymorphism in five other analyzed *P. flexilis* families. Second to M326511_1126R, M355530_134M was another top SNP locus in significant association with *Cr3*, mapped at 1.05 cM away from *Cr4*. However, no significant *Cr3*-association was detected in two SNPs (M296815_486Y and M362068_140R) that were close to *Cr4* within 1 cM, indicating that both SNP loci were monomorphic in the parent tree BRA08. These results implied that *Cr3/Cr4*-linked/associated alleles documented so far may have occurred through allelic mutation or/and recombination events during genetic differentiation among regions, populations, and seed families. Even ddRADseq-based genome-wide SNP genotyping detected little to no differentiation for most SNP loci between *P. strobiformis* and *P. flexilis* (Menon et al. 2018). Thus, a set of informative markers rather than one single marker appears to have a higher chance to achieve efficient MAS for accurate MGR prediction among parental trees without known resistance information, until markers are developed from *Cr3/Cr4* itself through positional sequencing and functional verification.

It is difficult to distinguish genotypes of QR from MGR in the WPBR pathosystems, especially for *P. flexilis* and *P. strobiformis*, where HR-like spots on infected needles, which are usually diagnostic for MGR in other species, are often not obvious in these two species. Identification of the MGR resistance genotype in *P. flexilis* and *P. strobiformis* therefore depends on the stem symptom-free trait. However, in the seedling trials, the progeny of some QR parents showed a relatively high frequency of canker-free seedlings, even when showing many susceptible needle spots, making the separation of MGR and some QR genotypes (the stem symptom-free phenotypes) problematic during phenotypic assessment. Phenotypic segregation ratios are commonly used to estimate presence of either MGR or QR in a seed family, but it is impossible to tell the difference between the two types of resistant genotypes if both resistant mechanisms are present in the same open-pollinated seed family. In addition, the frequency of canker-free seedlings from QR parents can vary from very low to intermediate levels (Johnson and Sniezko 2021). If QR approaches the 1:1 segregation ratio, it becomes phenotypically undistinguishable from MGR. This study assessed resistant traits of the same family using both MGR and operational trials. The susceptible ones would have neither MGR nor canker-free QR, but stem symptom-free traits of the resistant seedlings could be conferred by MGR, QR, or both MGR and QR genes. Some canker-free seedlings in MGR trial might also be developed by QR, although with the higher inoculum and younger age we would expect fewer. For an open-pollinated family, the seed parent is MGR, but pollen parents can be susceptible, QR, or MGR genotypes. Overall, 67.9% in MGR tests and 76.2% in QR tests were canker-free in this study, higher than the expected 50%, indicating either MGR or strong QR have some contribution through the pollen sources, especially obvious in family BRA08 (CY08). There are both MGR and QR parents, as well as susceptible ones, in the BRA08 stand, with MGR at a lower frequency. From our data, QR is more common than MGR in the Bradford Canyon region (Johnson and Sniezko 2021), where the parental tree was selected to harvest seeds for this study. The fact that we observed phenotypic segregation ratios significantly higher than 1:1 in the BRA08 family suggests that pollination of BRA08 occurs by MGR parents or QR parents or both MGR and QR parents, along with some susceptible genotypes. The MGR trials use younger seedlings and a much higher spore density in inoculation, factors that have little or no impact on expression of MGR, but that would be expected to yield relatively fewer canker-free seedlings from QR genotypes than in the operational trials. From the difference in percentage of canker-free seedlings in the two types of trials, it appears that ~10% of canker-free seedlings in the operational trials may be controlled by the QR genotypes in family BRA08 (CY08). In MGR trials, QR may also contribute a few seedlings, making marker-genotype–phenotyping matching rates lower in resistant phenotypes than in susceptible phenotypes in both MGR and operational trials.

With the aid of SNP genotyping using a TaqMan assay for M326511_1126R, we demonstrated that seedlings hypothesized to be QR genotypes can be tagged and separated from MGR genotypes in this *P. strobiformis* family. The TaqMan assay of M326511_1126R thus provided a MAS tool with potential application for pyramiding different resistance mechanisms in future screening of *P. strobiformis* resistance to *C. ribicola*. Seedling families from hundreds of additional *P. strobiformis* families covering the entire range of the species, including its populations in Mexico, are currently undergoing WPBR screening (R. Sniezko, *personal communication*), and tools such as these may be useful in clarifying the resistance type and in making forward selections for seed orchards and clone banks after they have been shown to work in wide genetic backgrounds. As shown in MDR analysis, a combination of multiple *Cr3/Cr4*-linked SNP markers has the potential for more accurate MGR prediction and genotypic differentiation of QR from MGR. Integration of these modern genomics-based breeding tools, along with rapid phenotyping techniques (Haagsma et al. 2020), will accelerate long-term sustainable management of white pine species by selecting and utilizing a diversity of elite genotypes with a combination of multiple *R* genes underlying both MGR and QR and other well-adapted genes responsive to environmental stresses, such as drought and cold hardiness.

ACKNOWLEDGMENTS

We thank Benjamin Rancourt, Hao Chen, and Gary Chang at the Canadian Forest Service for bioinformatic analyses; colleagues at the Dorena Genetic Resource Center for sample collection and phenotype assessment; and Northern Arizona University graduate students and technicians for cone and seed collection and processing (in 2013).

LITERATURE CITED

Abdollahi-Arpanahi, R., Morota, G., Valente, B. D., Kranis, A., Rosa, G. J. M., and Gianola, D. 2016. Differential contribution of genomic regions to marked genetic variation and prediction of quantitative traits in broiler chickens. Genet. Sel. Evol. 48:10.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403-410.

Baker, E. A. G., Wegrzyn, J. L., Sezen, U. U., Falk, T., Maloney, P. E., Vogler, D. R., Delfino-Mix, A., Jensen, C., Milton, J., Wright, J., Knaus, B., Rai, H., Cronn, R., Gonzalez-Ibeas, D., Vasquez-Gross, H. A., Famula, R. A., Liu, J.-J., Kueppers, L. M., and Neale, D. B. 2018. Comparative transcriptomics among four white pine species. G3 8:1461-1474.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635.

Chini, A., Grant, J. J., Seki, M., Shinozaki, K., and Loake, G. J. 2004. Drought tolerance established by enhanced expression of the CC-NBS-LRR gene, *ADR1*, requires salicylic acid, EDS1 and ABI1. Plant J. 38: 810-822.

Cleaver, C. M., Jacobi, W. R., Burns, K. S., and Means, R. E. 2015. Limber pine in the central and southern Rocky Mountains: Stand conditions and interactions with blister rust, mistletoe, and bark beetles. For. Ecol. Manage. 358:139-153.

Conklin, D. A., Fairweather, M. L., Ryerson, D. E., Geils, B. W., and Vogler, D. R. 2009. White pines, blister rust, and management in the Southwest. USDA Forest Service, Southwestern Region, R3-FH-09-01. https://www.fs.usda.gov/treesearch/pubs/33753

COSEWIC. 2014. COSEWIC assessment and status report on the limber pine *Pinus flexilis* in Canada. Committee on the Status of Endangered Wildlife in Canada, Ottawa, Ontario, Canada.

DeFaveri, J., Viitaniemi, H., Leder, E., and Merilä, J. 2013. Characterizing genic and nongenic molecular markers: Comparison of microsatellites and SNPs. Mol. Ecol. Resour. 13:377-392.

Devey, M. E., Mix, A. D., Kinloch, B. B., Jr., and Neale, D. B. 1995. Random amplified polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. Proc. Natl. Acad. Sci. USA 92:2066-2070.

Dougherty, L., Singh, R., Brown, S., Dardick, C., and Xu, K. 2018. Exploring DNA variant segregation types in pooled genome sequencing enables effective mapping of weeping trait in *Malus*. J. Exp. Bot. 69:1499-1516.

Eu-ahsunthornwattana, J., Miller, E. N., Fakiola, M., Wellcome Trust Case Control Consortium 2, Jeronimo, S. M. B., Blackwell, J. M., and Cordell, H. J. 2014. Comparison of methods to account for relatedness in genome-wide association studies with family based data. PLoS Genet. 10: e1004445.

Falk, T., Herndon, N., Grau, E., Buehler, S., Richter, P., Zaman, S., Baker, E. M., Ramnath, R., Ficklin, S., Staton, M., Feltus, F. A., Jung, S., Main, D., and Wegrzyn, J. L. 2018. Growing and cultivating the forest genomics database, TreeGenes. Database (Oxford) 2018:bay084.

Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., Sharma, Sha., Sharma, Shi., Kanzaki, H., Matsumura, H., Saitoh, H., Mitsuoka, C., Utsushi, H., Uemura, A., Kanzaki, E., Kosugi, S., Yoshida, K., Cano, L., Kamoun, S., and Terauchi, R. 2013. MutMap+: Genetic mapping and mutant identification without crossing in rice. PLoS One 8:e68529.

Gabriel, S., Ziaugra, L., and Tabbaa, D. 2009 SNP genotyping using the Sequenom MassARRAY iPLEX platform. Curr. Protoc. Hum. Genet. 60: 2.12.1-2.12.18.

Geils, B. W., Hummer, K. E., and Hunt, R. S. 2010. White pines, ribes, and blister rust: A review and synthesis. For. Pathol. 40:147-185.

Giral, H., Landmesser, U., and Kratzer, A. 2018. Into the wild: GWAS exploration of non-coding RNAs. Front. Cardiovasc. Med. 5:181.

Goodrich, B. A., Waring, K. M., Auty, D., and Meador, A. J. S. 2018. Interactions of management and white pine blister rust on *Pinus strobiformis* regeneration abundance in southwestern USA. Forestry 91:492-505.

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talón, M., Dopazo, J., and Conesa, A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36:3420-3435.

Gould, S. J., and Vrba, E. S. 1982. Exaptation—A missing term in the science of form. Paleobiology 8:4-15.

Government of Alberta. 2014. Species assessed by Alberta's endangered species conservation committee: Short list. Edmonton, Alberta, Canada. https://open.alberta.ca/publications/species-assessed-by-the-conservation-committee-alberta-species-at-risk

Haagsma, M., Page, G. F. M., Johnson, J. S., Still, C. J., Waring, K. M., Sniezko, R. A., and Selker, J. 2020. Using hyperspectral imagery to detect an invasive fungal pathogen and symptom severity in *Pinus strobiformis* seedlings of different genotypes. Remote Sens. 12:4041.

Hawksworth, F. G. 1990. White pine blister rust in southern New Mexico. Plant Dis. 74:938.

Jermstad, K. D., Eckert, A. J., Wegrzyn, J. L., Delfino-Mix, A., Davis, D. A., Burton, D. C., and Neale, D. B. 2011. Comparative mapping in *Pinus*: Sugar pine (*Pinus lambertiana* Dougl.) and loblolly pine (*Pinus taeda* L.). Tree Genet. Genomes 7:457-468.

Jin, W. T., Gernandt, D. S., Wehenkel, C., Xia, X. M., Wei, X. X., and Wang, X. Q. 2021. Phylogenomic and ecological analyses reveal the spatiotemporal evolution of global pines. Proc. Natl. Acad. Sci. USA 118: e2022302118.

Johnson, J. S., and Sniezko, R. A. 2021. Quantitative disease resistance to white pine blister rust in southwestern white pine (*Pinus strobiformis*). Front. For. Glob. Change. https://doi.org/10.3389/ffgc.2021.765871

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G. J., Maclean, D., Cock, P. J., Leggett, R. M., Bryan, G. J., Cardle, L., Hein, I., and Jones, J. D. 2013. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. Plant J. 76:530-544.

Kearns, H. S. J., Jacobi, W. R., Reich, R. M., Flynn, R. L., Burns, K. S., and Geils, B. W. 2014. Risk of white pine blister rust to limber pine in Colorado and Wyoming, USA. For. Pathol. 44:21-38.

Kinloch, B. B., Jr. 1970. White pine blister rust: Simply inherited resistance in sugar pine. Science 167:193-195.

Kinloch, B. B., Jr., and Dupper, G. E. 2002. Genetic specificity in the white pine-blister rust pathosystem. Phytopathology 92:278-280.

Kinloch, B. B., Jr., Sniezko, R. A., Barnes, G. D., and Greathouse, T. E. 1999. A major gene for resistance to white pine blister rust in western white pine from the Western Cascade Range. Phytopathology 89:861-867.

Kinloch, B. B., Jr., Sniezko, R. A., and Dupper, G. E. 2004. Virulence gene distribution and dynamics of the white pine blister rust pathogen in western North America. Phytopathology 94:751-758.

Kloth, K. J., Abreu, I. N., Delhomme, N., Petřík, I., Villard, C., Ström, C., Amini, F., Novák, O., Moritz, T., and Albrectsen, B. R. 2019. PECTIN ACETYLESTERASE9 affects the transcriptome and metabolome and delays aphid feeding. Plant Physiol. 181:1704-1720.

Krist, F. J., and Romero, S. A. 2014. 2013–2027 National insect and disease forest risk assessment: Summary and data access. In: Forest health monitoring: National status, trends, and analysis 2014. K. M. Potter, and B. L. Conkling, eds. Forest Service, U.S. Department of Agriculture, Washington, DC.

Leister, D. 2004. Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. Trends Genet. 20:116-122.

Liu, J.-J., and Ekramoddoullah, A. 2007. The CC-NBS-LRR subfamily in western white pine in *Pinus monticola*: Targeted identification, gene expression, and genetic linkage with resistance against *Cronartium ribicola*. Phytopathology 97:728-736.

Liu, J.-J., and Ekramoddoullah, A. 2008. Development of leucine-rich repeat polymorphism, amplified fragment length polymorphism, and sequence characterized amplified region markers to the *Cronartium ribicola* resistance gene *Cr2* in western white pine (*Pinus monticola*). Tree Genet. Genomes 4:601-610.

Liu, J.-J., Ekramoddoullah, A., Hunt, R., and Zamani, A. 2006. Identification and characterization of random amplified polymorphic DNA markers linked to a major gene (*Cr2*) for resistance to *Cronartium ribicola* in *Pinus monticola*. Phytopathology 96:395-399.

Liu, J.-J., Schoettle, A. W., Sniezko, R. A., Sturrock, R. N., Zamany, A., Williams, H., Ha, A., Chan, D., Danchok, D., Savin, D. P., and Kegley, A. 2016. Genetic mapping of *Pinus flexilis* major gene (*Cr4*) for resistance to white pine blister rust using transcriptome-based SNP genotyping. BMC Genomics 17:753.

Liu, J.-J., Schoettle, A. W., Sniezko, R. A., Williams, H., Zamany, A., and Rancourt, B. 2021. Fine dissection of limber pine resistance to *Cronartium ribicola* using targeted sequencing of the NLR family. BMC Genomics 22:567.

Liu, J.-J., Schoettle, A. W., Sniezko, R. A., Yao, F., Zamany, A., Williams, H., and Rancourt, B. 2019. Limber pine (*Pinus flexilis* James) genetic map constructed by exome-seq provides insight into evolution of disease resistance and a genomic resource for genomics-based breeding. Plant J. 98: 745-758.

Liu, J.-J., Sniezko, R. A., Sissons, R., Krakowski, J., Alger, G., Schoettle, A. W., Williams, H., Zamany, A., Zitomer, R., and Kegley, A. 2020. Association mapping and development of marker-assisted selection tools for the resistance to white pine blister rust in the Alberta limber pine populations. Front. Plant Sci. 11:557672.

Liu, J.-J., Sniezko, R. A., Sturrock, R. N., and Chen, H. 2014. Western white pine SNP discovery and high-throughput genotyping for breeding and conservation applications. BMC Plant Biol. 14:380.

Liu, J.-J., Sniezko, R. A., Zamany, A., Williams, H., Wang, N., Kegley, A., Savin, D. P., Chen, H., and Sturrock, R. N. 2017. Saturated genic SNP mapping identified functional candidates and selection tools for the *Pinus monticola Cr2* locus controlling resistance to white pine blister rust. Plant Biotechnol. J. 15:1149-1162.

Looney, C. E., Waring, K. W., and Fairweather, M. L. 2015. Monitoring the health of Pinus strobiformis: Early impacts of white pine blister rust invasion. Pages 167-174 in: Forest health monitoring: National Status, Trends, and Analysis 2013. K. M. Potter and B. L. Conkling, eds. General Technical Report SRS-207. Forest Service, Southern Research Station, U.S. Department of Agriculture, Asheville, NC.

Lorenz, W. W., Ayyampalayam, S., Bordeaux, J. M., Howe, G. T., Jermstad, K. D., Neale, D. B., Rogers, D. L., and Dean, J. F. D. 2012. Conifer DBMagic: A database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. Tree Genet. Genomes 8:1477-1485.

MacQueen, A., Sun, X., and Bergelson, J. 2016. Genetic architecture and pleiotropy shape costs of *Rps2*-mediated resistance in *Arabidopsis thaliana*. Nat. Plants 2:16110.

Marone, D., Russo, M. A., Laidò, G., De Leonardis, A. M., and Mastrangelo, A. M. 2013. Plant nucleotide binding site-leucine-rich repeat (NLR) genes: Active guardians in host defense responses. Int. J. Mol. Sci. 14: 7302-7326.

Menon, M., Bagley, J. C., Friedline, C. J., Whipple, A. V., Schoettle, A. W., Leal-Sàenz, A., Wehenkel, C., Molina-Freaner, F., Flores-Rentería, L., Gonzalez-Elizondo, M. S., Sniezko, R. A., Cushman, S. A., Waring, K. M., and Eckert, A. J. 2018. The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine (*P. flexilis*). Mol. Ecol. 27:1245-1260.

Menon, M., Bagley, J. C., Page, G. F. M., Whipple, A. V., Schoettle, A. W., Still, C. J., Wehenkel, C., Waring, K. M., Flores-Renteria, L., Cushman, S. A., and Eckert, A. J. 2021. Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. Commun. Biol. 4:160.

Menon, M., Landguth, E., Leal-Saenz, A., Bagley, J. C., Schoettle, A. W., Wehenkel, C., Flores-Renteria, L., Cushman, S. A., Waring, K. M., and Eckert, A. J. 2019. Tracing the footprints of a moving hybrid zone under a demographic history of speciation with gene flow. Evol. Appl. 13:195-209.

Oetjen, K., and Reusch, T. B. 2007. Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina*. Mol. Ecol. 16:5156-5157.

Peakall, R., and Smouse, P. E. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537-2539.

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J. P. T., and Hyvönen, J. 2013. Advances in plant gene-targeted and functional markers: A review. Plant Methods 9:6.

Pogorelko, G., Lionetti, V., Fursova, O., Sundaram, R. M., Qi, M., Whitham, S. A., Bogdanove, A. J., Bellincampi, D., and Zabotina, O. A. 2013. *Arabidopsis* and *Brachypodium distachyon* transgenic plants expressing *Aspergillus nidulans* acetylesterases have decreased degree of polysaccharide acetylation and increased resistance to pathogens. Plant Physiol. 162: 9-23.

Ritchie, M. D., Hahn, L. W., Roodi, N., Bailey, L. R., Dupont, W. D., Parl, F. F., and Moore, J. H. 2001. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am. J. Hum. Genet. 69:138-147.

Schoettle, A. W., and Sniezko, R. A. 2007. Proactive intervention to sustain high elevation pine ecosystems threatened by white pine blister rust. J. For. Res. 12:327-336.

Schoettle, A. W., Sniezko, R. A., Kegley, A., and Burns, K. S. 2014. White pine blister rust resistance in limber pine: Evidence for a major gene. Phytopathology 104:163-173.

Smith, C. M., Langor, D. W., Myrholm, C., Weber, J., Gillies, C., and Stuart-Smith, J. 2013. Changes in white pine blister rust infection and mortality in limber pine over time. Can. J. For. Res. 43:919-928.

Sniezko, R. A., Danchok, R., Savin, D. P., Liu, J.-J., and Kegley, A. 2016. Genetic resistance to white pine blister rust in limber pine (*Pinus flexilis*): Major gene resistance in a northern population. Can. J. For. Res. 46: 1173-1178.

Sniezko, R. A., Johnson, J. S., and Savin, D. P. 2020. Assessing the durability, stability, and usability of genetic resistance to a non-native fungal pathogen in two pine species. Plants People Planet 2:57-68.

Sniezko, R. A., Kegley, A. J., and Danchok, R. 2008. White pine blister rust resistance in North American, Asian and European species – Results from artificial inoculation trials in Oregon. Ann. For. Res. 51:53-66.

Sniezko, R. A., Kegley, A. J., Jacobs, J., Danchok, R., Luis, B., Long, S., and Savin, D. P. 2018. Variation in blister rust resistance and early height growth in three populations of southwestern white pine (*Pinus strobiformis*) and implications for management and conservation. Pages 140-148 in: Proceedings of the IUFRO Joint Conference: Genetics of Five-Needle Pines, Rusts of Forest Trees, and Strobusphere. Proceedings RMRS-P-76. A. W. Schoettle, R. A. Sniezko, and J. T. Kliejunas, eds. Forest Service, Rocky Mountain Research Station, U.S. Department of Agriculture, Fort Collins, CO.

Sniezko, R. A., Mahalovich, M. F., Schoettle, A. W., and Vogler, D. R. 2011. Past and current investigations of the genetic resistance to *Cronartium ribicola* in high-elevation five-needle pines. Pages 246-264 in: Proceedings of the High Five Symposium: The Future of High-Elevation, Five-Needle White Pines in Western North America. Proceedings RMRS-P-63. R. E. Keane, D. F. Tomback, M. P. Murray, and C. M. Smith, eds. Forest Service, Rocky Mountain Research Station, U.S. Department of Agriculture, Fort Collins, CO.

Sniezko, R. A., Smith, J., Liu, J.-J., and Hamelin, R. 2014. Genetic resistance to fusiform rust in southern pines and white pine blister rust in white pines—A contrasting tale of two rust pathosystems–Current status and future prospects. Forests 5:2050-2083.

Stevens, K. A., Wegrzyn, J. L., Zimin, A., Puiu, D., Crepeau, M., Cardeno, C., Paul, R., Gonzalez-Ibeas, D., Koriabine, M., Holtz-Morris, A. E., Martínez-García, P. J., Sezen, U. U., Marcais, G., Jermstad, K., McGuire, P. E., Loopstra, C. A., Davis, J. M., Eckert, A., de Jong, P., Yorke, J. A., Salzberg, S. L., Neale, D. B., and Langley, C. H. 2016. Sequence of the sugar pine megagenome. Genetics 204:1613-1626.

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., Innan, H., Cano, L. M., Kamoun, S., and Terauchi, R. 2013. QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J. 74:174-183.

Tomback, D. F., and Achuff, P. 2010. Blister rust and western forest biodiversity: Ecology, values, and outlook for white pines. For. Pathol. 40:186-225.

Vogan, P. J., and Schoettle, A. W. 2015. Selection for resistance to white pine blister rust affects the abiotic stress tolerances of limber pine. For. Ecol. Manage. 344:110-119.

Vogan, P. J., and Schoettle, A. W. 2016. Carbon costs of constitutive and expressed resistance to a non-native pathogen in limber pine. PLoS One 11:e0162913.

Westbrook, J. W., Chhatre, V. E., Wu, L. S., Chamala, S., Neves, L. G., Muñoz, P., Martínez-García, P. J., Neale, D. B., Kirst, M., Mockaitis, K., Nelson, C. D., Peter, G. F., Davis, J. M., and Echt, C. S. 2015. A consensus genetic map for *Pinus taeda* and *Pinus elliottii* and extent of linkage disequilibrium in two genotype-phenotype discovery populations of *Pinus taeda*. Genes Genomes Genet. 5:1685-1694.

Yang, S.-H., Li, J., Zhang, X.-H., Zhang, Q.-J., Huang, J., Chen, J.-Q., Hartl, D. L., and Tian, D.-C. 2013. Rapidly evolving *R* genes in diverse grass species confer resistance to rice blast disease. Proc. Natl. Acad. Sci. USA 110:18572-18577.

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., and Buckler, E. S. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38:203-208.