Final report to the West Virginia Department of Natural Resources

Project Title:
A genetic assessment of the population health and connectivity of a keystone species in high elevation Appalachian forest ecosystems: red spruce (*Picea rubens* Sarg.)

Principal Investigator:
Dr. Stephen R. Keller, Adjunct Faculty[1]

Key Personnel:
Regina Trott, Faculty Research Assistant

Supporting Institution:
Appalachian Laboratory
University of Maryland Center for Environmental Science
301 Braddock Rd.
Frostburg, MD 21532
301-689-7203
email: skeller@umces.edu

Award Number: WDU Project #13-R10

Date submitted: July 20, 2017

[1] Current Address:
Assistant Professor of Plant Biology
Department of Plant Biology
University of Vermont
111 Jeffords Hall, 63 Carrigan Drive
Burlington, VT 05465
Email: srkeller@uvm.edu
Phone: 802-656-5121

Background and Motivation

High elevation coniferous forests dominated by red spruce (*Picea rubens* Sarg.) are a biodiversity hotspot in the Central Appalachians, representative of boreal forest ecosystems typically restricted to more northerly regions of eastern Canada and New England. These forests are critical habitat for many species of plants and wildlife, and as such *P. rubens* represents a "keystone species" whose population viability and resilience to environmental change has far reaching conservation impacts (Byers et al. 2010).

The Central Appalachian red spruce ecosystem attains its greatest concentration in the high elevations (>2500 ft.) of the Allegheny Mountains in West Virginia, and western Maryland. Estimates of the original expanse of red spruce forest in this region suggest >400,000 hectares of this important ecosystem blanketed West Virginia. Today, <10% of this area is currently occupied by native red spruce (an estimated 30,000 hectares; WVWCAP), and much of this is second or third generation growth. This large historical reduction in distribution, along with its ecological sensitivity to a variety of sources of environmental change (e.g., acid rain, insect pathogens, deforestation, climate) has caused red spruce dominated forests to be regarded as endangered ecosystems at the state and global scales (Beane 2010; Byers & Norris 2011).

It is well known that demographic bottlenecks and habitat fragmentation can reduce genetic diversity within populations and alter gene flow across the landscape. When populations become smaller and more fragmented, diversity is lost through inbreeding and genetic drift, and differentiation among populations with limited connectivity increases as a result of restricted dispersal and gene flow (i.e., "population structure"). The role of genetic diversity in conservation is now well established, including the importance of heterozygosity to individual physiological performance and avoidance of inbreeding depression, and the necessity of genetic diversity for populations to be able to adapt to new selection pressures, such as changes in the type or abundance of insect herbivores, or shifts in the abiotic environment (Reed and Frankham 2003). The dominant view among conservation biologists and restoration ecologists is that the effective population size (N_e) is a key variable in to maintaining minimum viable populations (Schwartz et al. 2007). Best practices call for the minimum effective population size to be at least 500 in order to avoid the detrimental effects of inbreeding and genetic erosion, and to maintain viable populations that are adaptable and resilient in the face of environmental heterogeneity and disturbance (Shaffer 1981).

It is currently not clear what impacts historical decimation have had on the reservoir of genetic diversity present in Central Appalachian *P. rubens,* or if N_e in the region is large enough to meet conservation needs. Genetic diversity scales strongly and positively with reproductive fitness in *P. rubens* in other parts of its range, probably reflecting the degree of inbreeding depression arising among close relatives in low diversity populations (Rajora *et al.* 2000; Mosseler *et al.* 2003).

The ecological importance of *P. rubens* and its vulnerability to ecological disturbance has led to a large scale ecological restoration effort by the Central Appalachian Red Spruce Initiative (CASRI). The goal of this multi-agency group is to restore, via supplemental plantings and other silvicultural techniques, a functioning red spruce ecosystem throughout its former range in in the Central Appalachians (www.restoreredspruce.org). Genetic diversity plays a key role in restoration (Mckay *et al.* 2005), with restored populations benefitting from informed selection and planting of diverse, endemic genotypes. Thus, such restoration efforts could benefit from knowledge of the genetic diversity and N_e of restoration stock, and if certain existing stands suffer from usually low diversity or lack of genetic connectivity with other stands.

Given these conservation concerns, there is a need to assess the stability and health of existing natural stands of *P. rubens* across the Central Appalachian landscape. This project sought to fill some of these gaps by conducting a population genetic study to estimate levels of genetic diversity and degree of population structure among remnant stands of red spruce within WV and western MD, with the goal of characterizing the genetic diversity of *P. rubens* in the region, its effective population size and evidence for historical bottlenecks in N_e, and the extent to which gene flow maintains connectivity among remnant spruce stands. Specifically, we sought to address the following three sets of questions:

1. What are the standing levels of genetic diversity within stands? Do certain stands contain unusually low diversity, or exhibit evidence of inbreeding, and as such could be prioritized for genetic restoration?

2. What is the effective population size of Central Appalachian *P. rubens*? Is there evidence of temporal changes or bottlenecks in N_e? How does the N_e of Central Appalachian *P. rubens* compare to populations in other parts of its range?

3. How much genetic structure exists among existing stands in the Central Appalachians? What are the environmental variables (climate, geographic distance) that are most important to genetic isolation among remnant red spruce stands?

Methods

Field Sampling

Working in close collaboration with biologists at the WVDNR Natural Heritage Program, the Nature Conservancy, and the USDA Monongahela National Forest, we prioritized older growth or geographically isolated remnant stands of red spruce for field sampling of needle tissue for genetic analysis (Byers *et al.* 2010). Sampling locality information is summarized in Table 1.

Table 1. Sampling information for *P. rubens* populations collected during this study.

Site Code	State/ Prov	Location Name	NHP Plot Code[1]	N[2]	Latitude (dd.ddd)	Longitude (dd.ddd)	Elevation (m)	Minimum Stand Age (yrs)[3]
BAT	WV	Barlow Top	MONF.302	16	38.2292266	-80.2326555	1362	170
BLM	WV	Black Mountain	MONF.303	16	38.2951654	-80.2391810	1369	
BNT	WV	Boar's Nest Trail, Dolly Sods	MONF.313	16	38.9391179	-79.3913178	1315	
CBA	WV	Cranberry Glades Botanical Area	MONF.205	16	38.1980991	-80.2713728	1027	144
CMK	WV	Cheat Mountain Knob	MONF.272	16	38.6307291	-79.8953463	1212	
CNH	WV	Canaan Heights		16	39.0919726	-79.4498189	1155	
CNS	WV	Cranesville Swamp	CRSW.9	16	39.5338726	-79.4817813	787	
COW	WV	Cow Pasture	MONF.198	16	38.2022533	-80.2570746	1047	103
CPT	WV	Cow Pasture Trail	MONF.194	16	38.1955116	-80.2619217	1038	210
CSP	WV	Canaan State Park	CASP.30	16	39.0484004	-79.4739629	995	
FSR	WV	Forest Service Road 75		16	39.0111999	-79.3196768	1186	
GDK	WV	Gaudineer Knob		16	38.6141002	-79.8429144		
GLR1	WV	Glade Run-Shaver's Fork	MONF.288	16	38.6420113	-79.8414660	1193	215
GLR2	WV	Glade Run site 2		16	38.6571443	-79.8407687		
GRK	WV	Green Knob	MONF.318	16	38.8977829	-79.4411622	1406	
HAK	WV	Haystack Knob	MONF.320	16	38.9043187	-79.4406518	1336	140
HKB	WV	Huckleberry Trail	MONF.293	16	38.7347977	-79.5089717	1393	
KSF	WV	Kumbrabow State Forest	KUMB.22	17	38.6294772	-80.1327991	1058	301
LFW	WV	Laurel Fork Wilderness	MONF.165	16	38.6905218	-79.6991719	1060	90
OPR	WV	Old Piney Road	MONF.316	16	38.4729537	-79.7011336	1316	133
PHK	WV	Pharis Knob	MONF.266/267	16	38.7437972	-79.6307336	1287	196
PKB	WV	Panther Knob	PEND.14	16	38.5706214	-79.4900388	1295	
PRT	WV	Piney Ridge Trail	KUMB.23	16	38.6383999	-80.1315325	1105	287
RRN	WV	Red Run		16	38.6280000	-79.8995600		
RSK	WV	Red Spruce Knob	MONF.300	18	38.3346989	-80.1540319	1377	125

SEN	WV	Seneca Creek	MONF.321	16	38.7066954	-79.5473580	1201	135
SHV	WV	Shaver's Mountain	MONF.58/59	16	38.9896590	-79.6065495	1137	
SKB	WV	Stuart's Knob		16	38.9388145	-79.7135679	1206	
SKL	WV	Spruce Knob Lake		16	38.7180241	-79.6120161	1152	
SMR	WV	Saw Mill Run	PEND.17	16	38.6679848	-79.5704751	1165	90
SPK	WV	Spruce Knob	MONF.292	16	38.7166488	-79.5219644	1430	104
SPT	WV	South Prong Trail	MONF.295	16	38.9560255	-79.3565877	1220	
TKR	WV	Turkey (McGowan) Mountain	MONF.277	16	39.0192477	-79.6750557	1131	
TOA	WV	Top of Allegheny	MONF.315	16	38.4742984	-79.7166843	1299	116
WMT	WV	Whitmeadow	MONF.291	16	38.6670538	-79.8530919	1171	
WSW	WV	Whitmeadow Swamp	MONF.212	16	38.6720019	-79.8866321	1132	117
YCK	WV	Yellow Creek	MONF.103	16	38.9547280	-79.6637386	913	128
NGR	WV[4]	New Germany Seedlings		100				
FZL	MD	Finzel Swamp		16	39.7039308	-78.9384452	820	
GLD	MD	The Glades		16	39.5621173	-79.2781106	825	
WLF	MD	Wolf Swamp		16	39.6613345	-79.0917602	808	
BMB	PA	Bear Meadow's Bog		16	40.7303320	-77.7638050		
LAC	PA	Lackawanna		16	41.2008810	-75.6193360		
LKP	NY	Lake Placid		16	44.3312656	-73.9019769	494	
BPT	NH	Black Pond Trail		16	44.1057363	-71.5813700		
RAT	NH	Ripley-Arethusa Trail		16	44.1489594	-71.3884488		
SMN	VT	Smuggler's Notch		17	44.5734909	-72.7787015	595	
GT	ON	Gloucester Township		6	45.36199629	-75.53139755	90	

[1] Plot codes based on the WV DNR's Natural Heritage Program forest inventory plots

[2] Number of sampled individuals for genetic analysis

[3] Minimum stand age, based on the oldest dated tree core reported in the WV DNR Natural Heritage Program inventory plots (WVDNR 2017).

[4] Source of NGR seedlings was from pooled seed collected in WV. Seedlings were sampled for analysis at New Germany State Park, MD, prior to planting.

Trees were selected for sampling across a range of representative diameter classes within each site, and based on feasibility of access to needle tissue. Consequently, we were not always able to sample tissue from some very large trees due to the height of the canopy. For each tree selected, we clipped a small amount of fresh, current year branch growth with healthy needles. We also measured DBH to the nearest 0.1 cm, and recorded individual tree GPS coordinates and elevation (m a.s.l.) using a Garmin GPSmap 60Csx. We also noted several categorical descriptors for each sample, including position on slope (riparian, flat, lower slope, upper slope, crest), aspect (N, NE, E, SE, S, SW, W, NW), abundance of *P. rubens* (dominant, abundant, common, uncommon, rare, solitary), life history stage of the sampled individual (juvenile, mature (cone bearing), over-mature (crown senescing)), and took a photograph of the sampled tree.

In total, we sampled forty sites across WV (N=37) and western MD (N=3), with an average of 16 individuals per site, for a total of 643 individuals (Figure 1). We also included additional sampling from 7 sites outside of the Central Appalachian region to provide a broader geographic and genetic context and to assess the distinctiveness of the Central Appalachian diversity. This consisted of *P. rubens* from NH (N=32), PA (N=32), NY (N=16), VT (N=16), and Ontario (ON; N=6). Lastly, to assess the diversity encompassed by seedlings distributed by CASRI's restoration seedling program, we sampled 100 individuals from a single seedling lot delivered to New Germany State Park, near Grantsville, MD (NGR). The project-wide sample size for genetic analysis was thus N=845 individuals representing 47 natural populations and 1 cohort of restoration seedlings. All sampled needle tissue was transported back to the University of Maryland's Appalachian Laboratory (Frostburg, MD) and stored frozen at -80 °C until further processing. Collection details for individual samples are given in Appendix A.

DNA Isolation and Microsatellite Genotyping

We extracted whole genomic DNA from needle tissue using the Qiagen DNeasy 96 Plant kit and quantified it fluorometrically (Invitrogen QUANT-IT). We used this purified DNA to PCR amplify 18 microsatellite loci developed in other *Picea* species (mostly *P. glauca*) that have been shown to reliably cross-amplify in *P. rubens* (Pfeiffer *et al.* 1997; Hodgetts *et al.* 2001; Rajora *et al.* 2001; Scotti *et al.* 2002; Rungis *et al.* 2004). These loci were tested in preliminary analyses and found to be polymorphic and produce reliable amplification products for scoring allele sizes. We employed a multiplexing strategy (Blacket *et al.* 2012) to amplify multiple loci simultaneously. Specifically, we used a combination of a universal primer fluorescently labeled with 1 of 4 dyes (6-FAM, NED, PET, and VIC) and a pair of locus-specific primers, with the forward locus-specific primer modified to have a 5' universal sequence tail complimentary to the labeled primer. This enabled a highly efficient and flexible system to develop multiplexed sets of loci labeled with different fluorophores (Table 2).

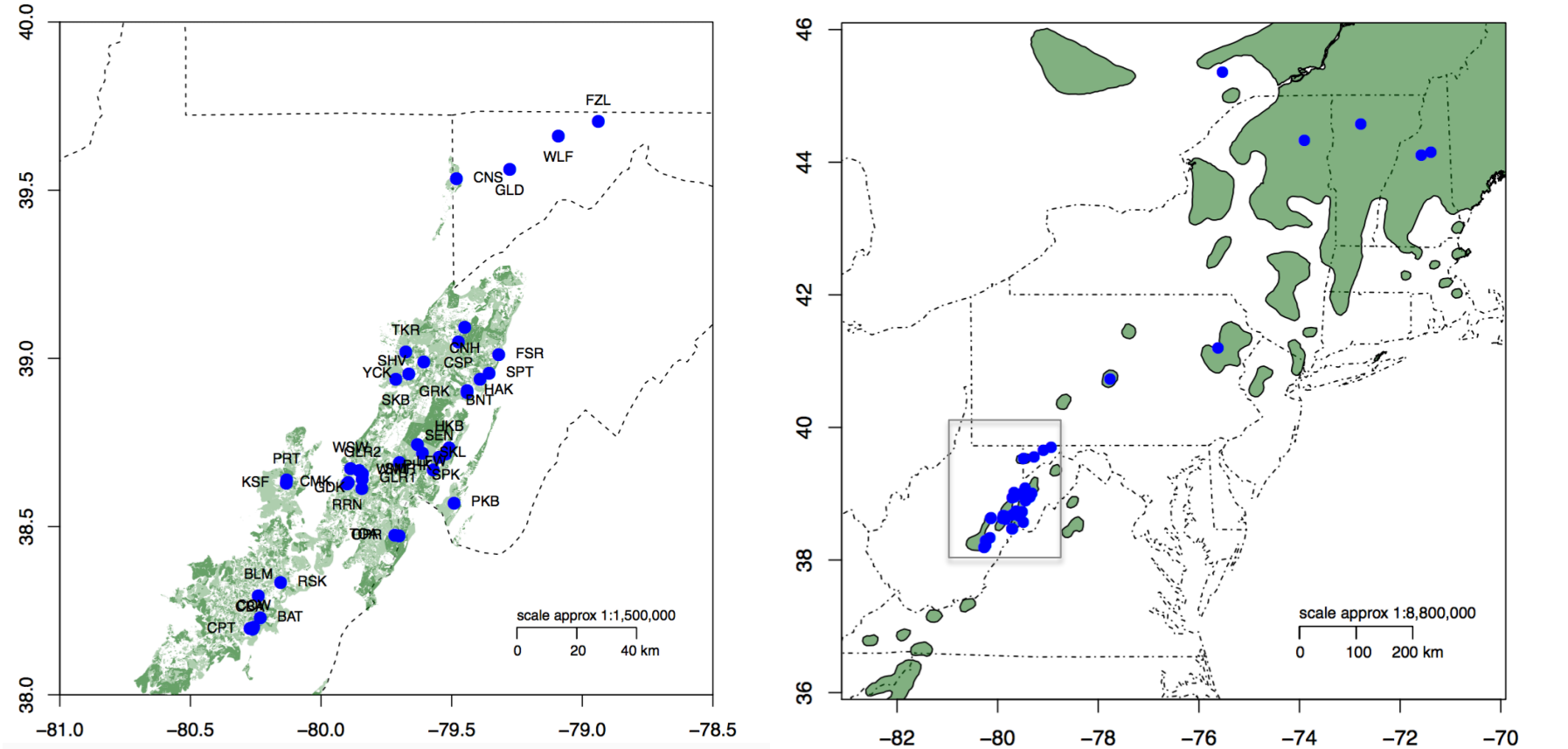

Figure 1. Sampling localities for red spruce. Blue points indicate sampling localities of trees used in this study. Left panel: Map of 40 sampling localities in the Central Appalachian region (WV and MD). Green shading indicates estimated density of red spruce cover (high, medium, low) based on WVDNR NHP mapping. Right panel: Regional map, showing additional sampling localities in the northeast. Inset shows the Central Appalachian region. Shading indicates the range-wide distribution of *P. rubens* (Little 1971).

Table 2. Primer information for 18 microsatellite loci amplified in *P. rubens.*

Locus Name	Motif	Size (bp)	Dye	Plex	T_A (°C)	Reference
GAT64	GAT	102-112	6-FAM	A	57	Hodgetts et al. (2001)
TG25	TG	94-106	PET	A	57	Hodgetts et al. (2001)
CT189B	CT	114-116	NED	A	57	Hodgetts et al. (2001)
CA24	AC	184-212	6-FAM	A	57	Hodgetts et al. (2001)
PGL12	AG	222-250	PET	A	57	Rajora et al. (2001)
PGL14	AG	136-180	VIC	A	50	Rajora et al. (2001)
CT144	CT	145-154	NED	A	50	Hodgetts et al. (2001)
SPAGG3	GA	130-146	6-FAM	B	57	Pfeiffer et al. (1997)
UAPgAG105	AG	175-179	NED	B	57	Hodgetts et al. (2001)
WS0082.E23	TA	248-266	VIC	B	57	Rungis et al. (2004)
PAAC23	GT	284-294	NED	B	57	Scotti et al (2000)
WS0022.B15	AG	210-225	NED	C	57	Rungis et al. (2004)
WS0092.A19	AC	241-245	6-FAM	C	57	Rungis et al. (2004)
EAC6B03	AC	129-135	VIC	C	57	Scotti et al. (2002)
SPL3AG1H4	GA	135-137	PET	C	57	Pfeiffer et al. (1997)
WS00111.K13	AT	234-238	PET	C	57	Rungis et al. (2004)
WS0082.023	TA	226-240	VIC	C	57	Rungis et al. (2004)

PCR amplifications used the Qiagen Multiplex PCR kit in 10ul reactions, consisting of 1 ul ddH_2O, 5 ul Multiplex PCR Master Mix, 1 ul Q-solution, and amplified on an Eppendorf MasterCycler pro PCR machine with the following conditions: 95 °C for 15 min., 30 cycles of 94 °C for 30 s, T_A (either 50 or 57 °C ; see Table 2) for 90 s, 72 °C for 90 s, followed by a final 72 °C extension for 30 min. and then a 4 °C hold. A sample of amplification products were screened using gel electrophoresis before combining with LIZ500 size standard and shipping to the West Virginia University Genomics Core Lab for fragment analysis on an ABI3130xl sequencer (Applied Biosystems). Raw data were sized using the software PeakScanner (Applied Biosystems) using automated scoring, followed by manual checking. Allele fragment sizes were binned using the program TANDEM (Matschiner & Salzburger 2009).

Statistical Analysis

1. Genetic Diversity within Populations and Individuals

We estimated population genetic diversity statistics averaged across loci for each site using the '*diveRsity*' package (Keenan *et al.* 2013) in R v.3.3.2 (R Core Team 2014). Diversity measures included the raw number alleles per locus unadjusted for sample size (A), the proportion of all alleles observed within sites (%A), allelic richness adjusted for sample sized based on rarefaction (Ar), observed heterozygosity (H_o), expected heterozygosity (H_e), and the population inbreeding coefficient (F_{IS}). Significance of F_{IS} was tested with 1,000 bootstrap resamples.

We were also interested in assessing genetic diversity at the individual level, to test for an association between size class (DBH) and heterozygosity. Such an association might be expected if selection acting across life history stages favors more heterozygous, less inbred genotypes, yielding an enrichment in heterozyogisty among the largest DBH trees (a proxy for age) relative to more recently recruited small DBH trees. To estimate multi-locus heterozygosity for each individual across the 18 loci, we used the R package '*Rhh*' (Alho *et al.,* 2010) to estimate the multi-locus standardized heterozygosity measure (SH; Coltman *et al.,* 1999). In comparing SH to DBH, we recognize that DBH is going to be highly influenced by environmental differences, both across sites and within a site, due to variation in release rates of subcanopy trees. Therefore, we tested for heterozygosity~DBH associations by restricting the sample to just the lower (<9.60 DBH) and upper quartile of tree DBH (>34.40 DBH), and included site as a random effect in a mixed linear model using the '*lme4*' package in R (Bates *et al.* 2015).

2. Population Structure and Genetic Ancestry

We estimated genetic differentiation among all populations using Nei's standardized estimate of allele frequency divergence among populations (G_{ST}), and Hedrick's re-scaled estimate that determines G_{ST} relative to its maximum possible level (G'_{ST}). Estimation was done using the '*diveRsity*' package, with significance determined with 1,000 bootstrap resamples. To assess differences in genetic ancestry without our sample, and to assign individuals to their most likely gene pools, we used the Bayesian clustering program STRUCTURE version 2.2 (Pritchard *et al.,* 2000; Falush *et al.,* 2003), using the admixture model with correlated allele frequencies. We performed 10 independent runs for each *K* (1-10) with 1,000,000 MCMC iterations after a burn-in period of 200,000 iterations. Post-processing of STRUCTURE runs and ad-hoc estimation of the number of clusters (*K*) based on the delta-K method (Evanno *et al.* 2005) was implemented in the software STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Ancestry coefficients across runs were combined using CLUMPP (Jakobsson & Rosenberg, 2007).

3. Genetic Connectivity and Isolation by Environment

To estimate the factors contributing to historical connectivity and gene flow among Central Appalachian stands, we used a landscape genetics approach to relate population structure to environmental features across the landscape (Manel *et al.* 2003). Specifically, we used Generalized Dissimilarity Modeling (GDM) to relate pairwise genetic distances among sites to differences in their local climate environment (Ferrier *et al.* 2007; Fitzpatrick & Keller 2015). The GDM approach is a form of multivariate matrix regression that fits non-linear I-splines to model the relationship between biological distance and multiple environmental predictors. Climate data consisted of 19 bioclimatic variables that summarize seasonal aspects of temperature and precipitation (http://www.worldclim.org/bioclim), based on the WorldClim dataset downscaled to 30s resolution (Hijmans *et al.* 2005). We also included pairwise Euclidean geographic distance as an additional predictor to

accommodate effects of isolation by distance. We used Hedrick's G'_{ST} as the genetic distance metric, normalized between 0-1 by scaling based on the largest pairwise value observed. GDM models were fit using the R package '*gdm*' (Manion *et al.* 2016). After model fitting, we retained all significant environmental predictors and used them to transform continuous bioclimatic space in the Central Appalachian region based on the modeled association between genetic and climatic distances. The resulting transformation was then mapped to visualize the continuous surface of genetically-transformed environmental differences between sites. Such an analysis can be used to highlight which regions of the landscape are estimated to be well connected by gene flow through climatic space, and which are likely to be genetically isolated by environmental differences.

We also explored the extent of genetic connectivity among Central Appalachian *P. rubens* using Dyer's population graph method (Dyer & Nason 2004; Dyer 2014). Population graphs employ a graph-theoretic framework to model genetic covariance in allele frequencies among sites using network theory, and estimate conditional genetic distances that reflect the pattern of connectivity and gene flow among sites. Population graphs for *P. rubens* in the Central Appalachians were estimated using the '*popgraph*' R package, with an alpha threshold tolerance of 0.025, and overlaid onto GDM model predictions to map how population network connectivity corresponds to genetic isolation by climatic distance.

4. Current and Historical Effective Population Size (N_e)

To estimate the effective population size (N_e) in the Central Appalachians, we employed the LDNe method implemented in NeEstimator (Do *et al.* 2014). This method estimates current N_e by looking at the extent of linkage disequilibrium present within a population, relative to what would be expected under genetic drift in a population of size N_e. We initially attempted to estimate N_e separately for each site, but the average sample size of 16 individuals (Table 1) led most values to be inestimable. Thus, we instead considered 4 different data subsets for estimating N_e: (1) a single Central Appalachian "population" consisting of 643 sampled individuals from WV and MD (excluding the restoration seedlings from NGR); (2) the 100 NGR restoration seedlings, to assess what fraction of total diversity was represented by restoration stock; (3) the 102 northeastern samples from PA, NY, VT, NH, and ON, to relate Central Appalachian diversity to that present further north; and (4) the entire pooled sample of 845 individuals.

As a compliment to the LDNe method, we also estimated current and historical N_e using a coalescent approach. Unlike the LDNe method based on patterns of linkage disequilibrium, the coalescent approach models the genealogical relationships giving rise to the distribution of allele sizes sampled in the population. The "birth-death" process of alleles then provides an estimate of the rate at which the population was subject to genetic drift, which is proportional to N_e. Unlike the LDNe method however, the coalescent can be used to infer past temporal changes in N_e that have shaped the genealogy of alleles. We used the coalescent method

implemented in the R package '*VarEff*' to model current N_e as well as ancestral N_e using a size-change model (Nikolic & Chevalet 2014). The tested model allowed for two historical size changes leading up to the current N_e in the Central Appalachians. Thus, there were 3 effective population sizes estimated looking backwards in time: current N_e, N_e after the size change (N_e-2), and the ancestral N_e prior to the size change (N_e-3). As is standard in coalescent approaches, estimation was for the compound parameter θ (=4 $N_e\mu$, where μ is the microsatellite mutation rate per locus per generation). Similarly, timing of the size change was estimated as the parameter τ (=gμ, where g is time measured in generations). We first estimated appropriate ranges for priors on N_e and g from preliminary runs using the Theta() function. We then performed Bayesian estimation of parameters with Markov Chain Monte Carlo, with a burn-in of 10,000 iterations followed by 10,000 sampling iterations and a thinning interval of 10. As advised by Nikolic and Chevalet (2014), we took the median value from the posterior distribution of each parameter as our demographic estimate, and integrated across the 95% posterior probability to obtain confidence intervals. To convert compound parameters to an absolute demographic scale, we substituted a mutation rate of $5x10^{-4}$ for μ. This is value an over-simplification that does not accommodate uncertainty in the mutation rate nor variance among loci, but is consistent with microsatellite mutation rates reported in the literature (Marriage *et al.* 2009), and estimation of N_e in this way provides a point of comparison to the LDNe method which does not require use of a mutation rate. Therefore, agreement between the two different N_e estimators, which make use of different signatures in the genetic data, lends confidence to our inference of effective population size.

Results

Genetic Diversity

Microsatellite diversity within sites showed limited variability across sampling locations, with several notable exceptions (Table 3). Allelic richness (Ar) within Central Appalachian sites (WV and MD) varied from 2.8 to 3.86, with a mean of 3.12. Sites that stood out as having unusually high diversity of alleles included SKB (Stuart's Knob), CNH (Canaan Heights) and YCK (Yellow Creek), while KSF (Kumbrabow State Forest) and PKB (Panther Knob) possessed the lowest values for this region, consistent with the geographic isolation of these localities. This was additionally reflected by the proportion of total allelic richness present within sites (%A), in which SKB stood out clearly with >45% of the total diversity observed within this one site. The NGR restoration seedlings possessed the largest absolute number of alleles (111), likely reflecting the large samples size (N=100) and pooling of seeds across collecting localities. After rarefaction, allelic richness of NGR was still relatively high but lower than SKB, CNH, or YCK (Table 3). Outside of the Central Appalachians, LAC (Lackawanna) also possessed high allelic diversity (A, %A, and Ar) while populations further north such as LKP (Lake Placid, NY), SMN (Smuggler's Notch, VT) and GT (Gloucester Township, ON) exhibited some of the lowest values

Table 3. Population genetic diversity within sites. Boldface values of F_{IS} are significant based on 95% confidence intervals from 1000 bootstrap replicates.

Site Code	State	N	A	%A	Ar	Ho	He	F_{IS}
BAT	WV	15.67	83	31.52	3.23	0.51	0.54	0.0607
BLM	WV	15.44	74	29.29	3.04	0.51	0.52	0.0172
BNT	WV	15.5	79	31.12	3.27	0.52	0.53	0.0254
CBA	WV	13.78	76	29.44	2.94	0.51	0.52	0.0264
CMK	WV	15.44	78	30.29	3.08	0.53	0.53	-0.01
CNH	WV	14.89	100	38.96	3.4	0.49	0.56	0.1344
CNS	WV	15.78	88	35.15	3.27	0.51	0.54	0.0582
COW	WV	15.78	77	30.18	3.25	0.57	0.54	-0.0676
CPT	WV	15.67	78	29.9	3.03	0.48	0.5	0.0394
CSP	WV	15.5	79	30.76	3.11	0.52	0.52	0.0138
FSR	WV	15.5	74	27.92	2.98	0.52	0.51	-0.0258
GDK	WV	15.17	81	31.4	3.14	0.54	0.53	-0.0068
GLR1	WV	14.06	77	29.07	2.98	0.54	0.5	-0.0775
GLR2	WV	15.61	79	31.07	3.13	0.53	0.53	0.016
GRK	WV	15.5	78	30.72	3.05	0.5	0.53	0.0667
HAK	WV	15.61	76	30.19	3.04	0.5	0.5	0.0098
HKB	WV	15.83	70	27.59	3	0.5	0.52	0.0327
KSF	WV	14.5	72	28.47	2.8	0.47	0.51	0.0672
LFW	WV	14.5	79	30.51	3.04	0.49	0.5	0.0266
OPR	WV	15.33	77	30.02	3.07	0.54	0.54	-0.013
PHK	WV	15.11	78	30.36	3.2	0.54	0.54	-0.0017
PKB	WV	15.28	67	26.09	2.8	0.48	0.47	-0.0246
PRT	WV	14.78	75	29.96	2.95	0.51	0.54	0.0419
RRN	WV	15.72	77	30.38	3.11	0.54	0.52	-0.0459
RSK	WV	16.17	80	31.13	3.02	0.53	0.54	0.024
SEN	WV	15.17	83	32.59	3.13	0.54	0.53	-0.0133
SHV	WV	15.67	83	31.98	3.25	0.55	0.54	-0.016
SKB	WV	14.94	112	45.19	3.86	0.52	0.64	**0.1894**
SKL	WV	15.33	80	31.13	3.12	0.48	0.52	0.0646
SMR	WV	15.5	84	33.24	3.14	0.5	0.53	0.042
SPK	WV	15.5	83	31.27	3.2	0.5	0.53	0.057
SPT	WV	14.94	74	28.37	3.03	0.55	0.53	-0.0394
TKR	WV	15.33	78	30.82	3.19	0.56	0.53	-0.0448
TOA	WV	15.67	75	29.01	3.02	0.56	0.52	**-0.0789**
WMT	WV	15.33	79	31.14	3.12	0.53	0.53	-0.0028
WSW	WV	15.28	77	29.29	3.17	0.55	0.53	-0.0415
YCK	WV	15.22	90	34.73	3.4	0.57	0.56	-0.0084
FZL	MD	14.44	81	33.56	3.09	0.47	0.55	**0.1369**
GLD	MD	15.39	82	33.07	3.17	0.5	0.55	0.0843
WLF	MD	14.72	71	28.97	2.94	0.47	0.52	**0.0967**
NGR	WV	96.5	134	49.71	3.31	0.51	0.56	**0.086**
BMB	PA	14.61	86	35.19	3.23	0.52	0.57	**0.0918**
LAC	PA	14.67	118	48.97	3.88	0.56	0.66	**0.1466**
SMN	VT	14.78	75	30.61	2.79	0.49	0.48	-0.0122
LKP	NY	15.22	67	27.49	2.71	0.43	0.49	**0.1228**
BPT	NH	15.61	77	31.87	3.05	0.53	0.51	-0.0387
RAT	NH	15.67	78	31.45	3.04	0.53	0.52	-0.0244
GT	ON	5.06	63	27.09	2.87	0.48	0.53	0.1005

observed across the entire dataset, possibly reflecting a loss of diversity along the pathway of post-glacial range expansion.

Heterozygosity showed similar trends among sites, with SKB and CNH having high values (Table 3). However, levels of observed heterozygosity at these sites caused a significant departure from Hardy-Weinberg equilibrium (significantly positive F_{IS} value), indicating a deficit of heterozygotes. While selfing or biparental inbreeding is one possible explanation for positive F_{IS} values, an alternative explanation consistent with the high diversity at these sites is a recent admixture of multiple sources with divergent allele frequencies (e.g., a Wahlund effect), causing allelic diversity and heterozygosity to increase, but with observed heterozygosity lagging due to insufficient time for mating to homogenize allele frequencies among offspring. The explanation of recent admixture both raising diversity while also creating a deficit of observed heterozygotes is also consistent with the diversity of the NGR seedlings, which showed significant F_{IS}. In this case, we know that "admixture" occurred during the pooling of seeds from multiple different source localities. In contrast, FZL (Finzel Swamp) and WLF (Wolf Swamp) in western MD both showed significant F_{IS} accompanied by relatively low allelic richness, suggesting that heterozygote deficit within these sites at the northern edge of the Central Appalachians may be attributable to inbreeding. At the other end of the spectrum, site TOA (Top of Allegheny) showed a significantly *negative* F_{IS}, indicating an excess of observed heterozygotes relative to expectations under random mating. One possibility is selection for heterozygous genotypes at TOA, but given that this is the only population to show such a pattern, the significance of F_{IS} here may simply be attributable to multiple testing across a large number of sites.

As a more rigorous test of the hypothesis that selection may favor the eventual dominance of more heterozygous genotypes, we tested for an association between individual multi-locus heterozygosity and DBH class. Heterozygosity was weakly but not significantly correlated across loci within individuals (r = 0.042; 95% CI: -0.016-0.099), providing limited support for selection favoring individuals that are heterozygous across their genomes. Consistent with this, we observed no support for an association between multi-locus heterozygosity and DBH size class after controlling for site effects (β = 0.019; t = 0.729; P = 0.467) (Figure 2). Thus, unlike previous reports from *P. rubens* populations in Canada (Rajora *et al.* 2000; Mosseler *et al.* 2003), we find little evidence of increased heterozygosity with size class.

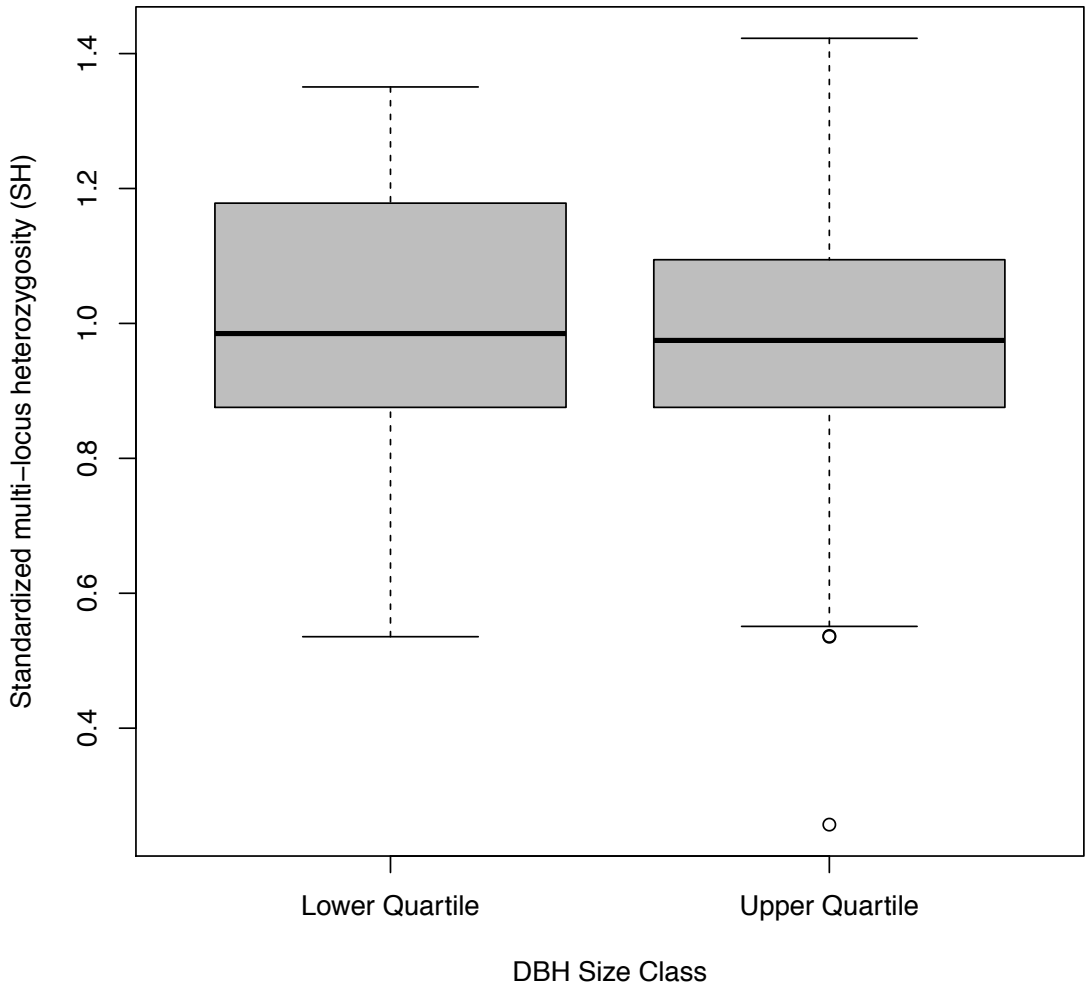

Figure 2. Association between standardized heterozygosity (SH) and the lower and upper quartile of tree DBH, used here as a rough proxy of age.

Population Structure

Genetic differentiation among sites was low overall (G_{ST} = 0.0109, 95% C.I. 0.0051-0.0179; G'_{ST} = 0.0245, 95% CI: 0.0122-0.0386), but in line with low differentiation reported for other wind-pollinated, predominantly outcrossing forest trees (Savolainen *et al.* 2007). These G_{ST} values indicate that roughly 1-2% of the total genetic diversity present in the Central Appalachians is partitioned among different sites, with the majority of variability existing within sites. Estimates of pairwise divergence among sites showed more nuanced variability, with most site-pairs showing little divergence, and a few showing relatively large genetic divergence, often involving either FLZ or WLF and other Central Appalachian populations (Appendix B).

STRUCTURE analysis of genetic ancestry returned K=3 clusters based on the delta-K method of model selection (Evanno *et al.* 2005). At K=3, individual ancestry assignments showed most sites in the Central Appalachians contained mixtures of 2 different genetic clusters, possibly reflecting admixture between two different refugial regions in the southeastern U.S. during the Pleistocene glacial maximum (Appendices C1 and C2). Interestingly, many individuals in SKB and CNH shared ancestry in a genetic cluster that is predominantly found in northeastern sites (red in K=3; Appendix C1). At K=4, this cluster separates into a fourth group at highest frequency in the LAC site, and it is with this group that some SKB individuals share genetic ancestry. Along with evidence of elevated allele diversity and departures from Hardy-Weinberg equilibrium in these sites, this provides indication that SKB and CNH were probably supplemented at some point in the past with seedling material originating from outside the Central Appalachian region, likely from northeastern *P. rubens*. Also of interest is the ancestry of NGR restoration seedlings, which has approximately equal ancestry across all three clusters, including the northeastern cluster (Appendix C2). This could indicate that the seed sources of restoration plantings may themselves have been supplemented at some time in the past with genetic ancestry from the northeast, but this remains speculative without additional testing.

Genetic Connectivity and Isolation by Environment

The climate zone inhabited by sampled *P. rubens* stands in this study show the Central Appalachians occupying a distinct climate niche from northeastern red spruce forests (Figure 3, top). Specifically, Central Appalachian sites were separated from other samples along the first Principal Component climate axis (54% of variance), with WV and MD sites characterized by higher temperatures, higher precipitation, and lower seasonality relative to northern sites. When the Central Appalachian region was analyzed separately, the three western MD sites (WLF, FLZ, and GLD) separated along a precipitation gradient represented by the first PC axis (51% of variance), indicating that these sites occupy a drier microclimate within the Allegheny Mountains (Figure 3, bottom).

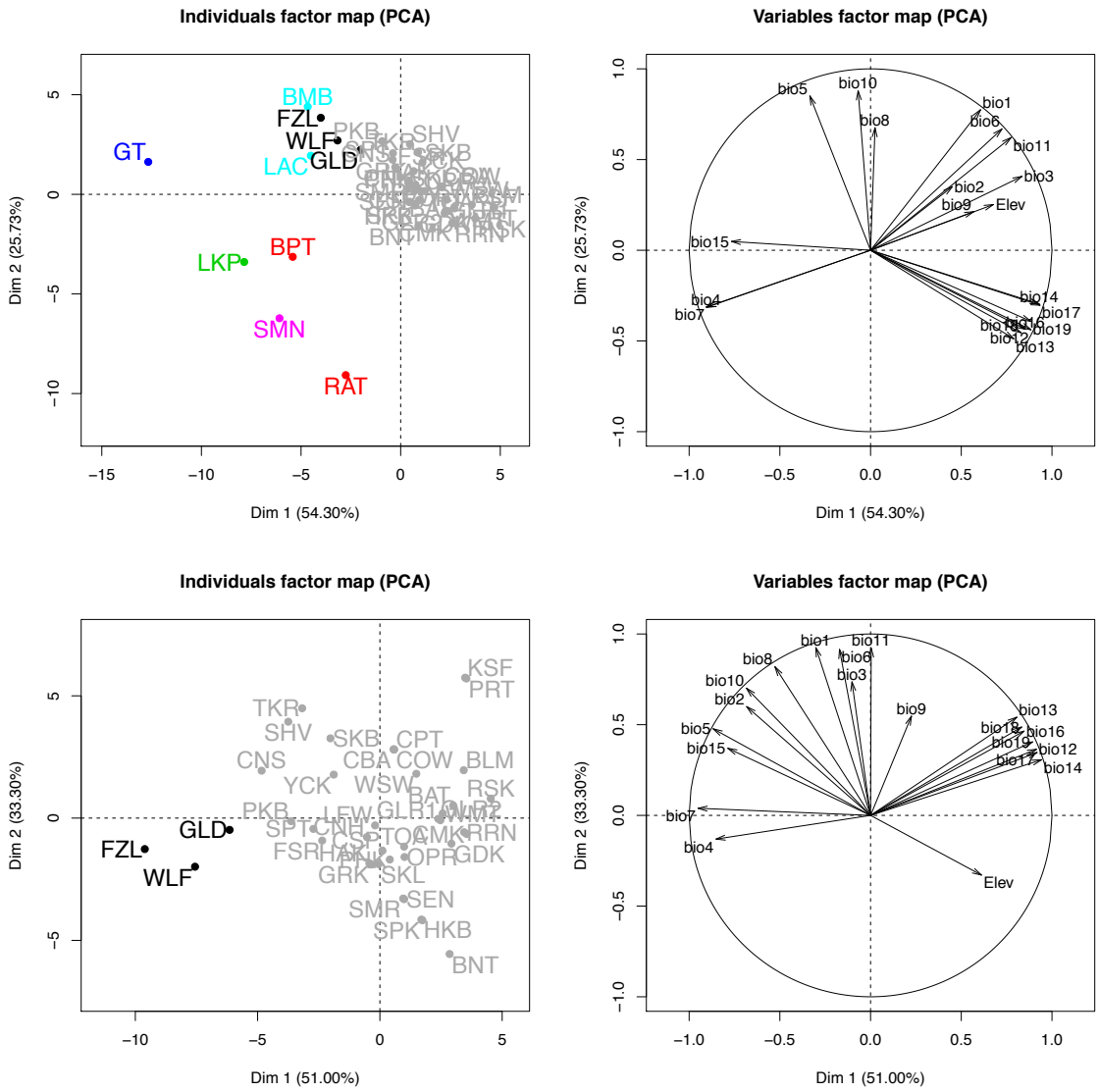

Figure 3. Principal Component Analysis (PCA) of climate space occupied by red spruce populations in this study, as defined by 19 bioclimatic variables. Top panels show range-wide PCA; bottom panels show PCA of Central Appalachian red spruce. Bioclim variable definitions: bio1 (mean annual temp.), bio2 (mean temp. diurnal range), bio3 (isothermality), bio4 (temp. seasonality), bio5 (max temp. of warmest month), bio6 (min temp. warmest month), bio7 (temp. annual range), bio8 (mean temp. wettest quarter), bio9 (mean temp. driest quarter), bio10 (mean temp. warmest quarter), bio11 (mean temp. coldest quarter), bio12 (annual precip.), bio13 (precip. wettest month), bio14 (precip. driest month), bio15 (precip. seasonality), bio16 (precip. wettest quarter), bio17 (precip. driest quarter), bio18 (precip. warmest quarter), bio19 (precip coldest quarter).

We then asked whether these bioclimatic differences in occupied climate niche space within the Central Appalachians have shaped patterns of gene flow and population connectivity across the region. Using Generalized Dissimilarity Modeling (GDM), we related pairwise population genetic divergence (G'_{ST}) to differences among sites for the 19 bioclimatic variables. The overall deviance explained by environmental differences was only 17% of the total deviance, suggesting most of the relatively weak genetic divergence among sites is left unaccounted for, probably due to other aspects of land use history as well as stochastic demographic and sampling processes. Interestingly, only four environmental predictors were significant in the model, showing non-zero deviance in explaining genetic connectivity among sites (Figure 4). Of these, the most important by far was precipitation during the warmest quarter of the year (bio18), with steep changes in genetic connectivity across the dry portion of the precipitation gradient from 32.0 – 36.0 cm, followed by little change in connectivity across the wetter portion of the gradient above 36.0 cm (Figure 4). The other important predictor of genetic connectivity was the mean temperature of the coldest quarter of the year (bio11), where connectivity changed the most along the cold portion of the gradient (below -3.5 °C). Deviance explained by the remaining variables showed relatively minor effects of geographic distance and precipitation seasonality (bio15).

The turnover functions produced by GDM allow transformation of the continuous bioclimatic surface into a genetically informed map of how the environment is expected to mediate genetic connectivity among sites (Fitzpatrick & Keller 2015). These transformations are based on the modeled turnover functions that associate specific portions along environmental gradients with high or low turnover in genetic diversity (Figure 4). In GDM transformed maps, pixels with similar colors are predicted to experience low genetic isolation by environment, and hence share high connectivity through the gene-environment space. GDM transformation in the Central Appalachians shows a high degree of connectivity among much of the core portion of *P. rubens* range in the region (Figure 5). The general trend in connectivity is along a SW to NE trending axis, with abrupt transitions towards the rain shadow of the Allegheny front, in the vicinity of PKB (Panther Knob) near the WV/VA border. Strong changes in connectivity are also apparent northward into Garret County, MD, where sites GLD (Glades), WLF, and FZL show low predicted connectivity with the rest of the range in WV. This drop in connectivity is likely driven by regions of reduced summer precipitation north and east of the Allegheny Front, although these specific stands of red spruce have found suitable sites mediated by local hydrology. Nevertheless, the GDM underscores the genetic and climatic isolation of the western MD spruce sites from the rest of the Central Appalachian region. This reduced connectivity is also accompanied by reduced within-site diversity and evidence for inbreeding (Table 3).

Overlaying the network topology of conditional genetic distances from the population graph analysis clearly shows the genetic isolation of western MD *P. rubens* sites, including CNS (Cranesville Swamp) on the WV/MD border (Figure 6).

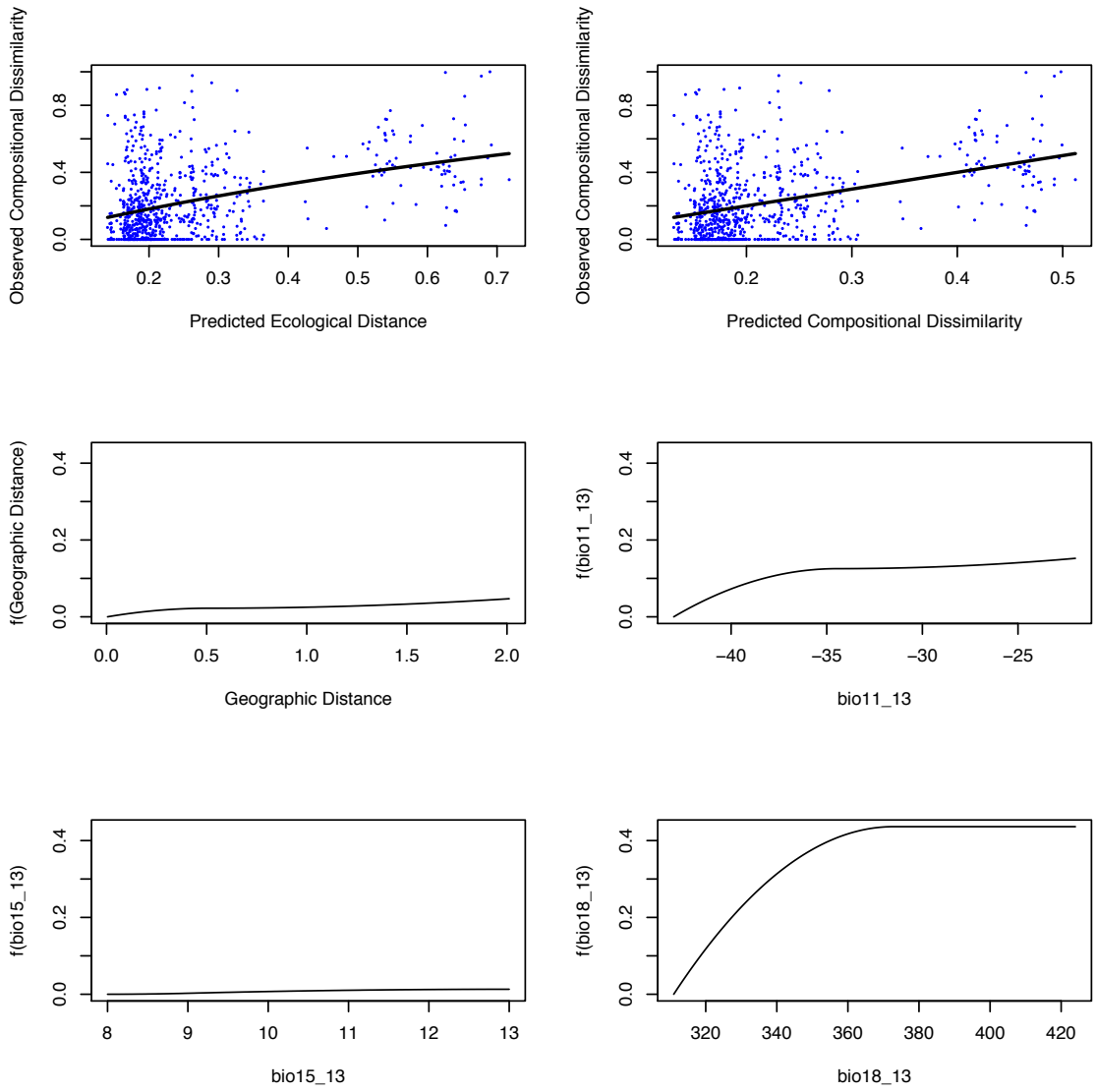

Figure 4. Generalized Dissimilarity Modeling (GDM) results showing influence of bioclimatic variables on population genetic structure. Top panels show model performance at predicting observed genetic distance among populations. Bottom two rows of figures show genetic turnover along environmental gradients of geographic distance, mean temperature of coldest quarter (bio11), precipitation seasonality (bio15), and precipitation of warmest quarter (bio18).

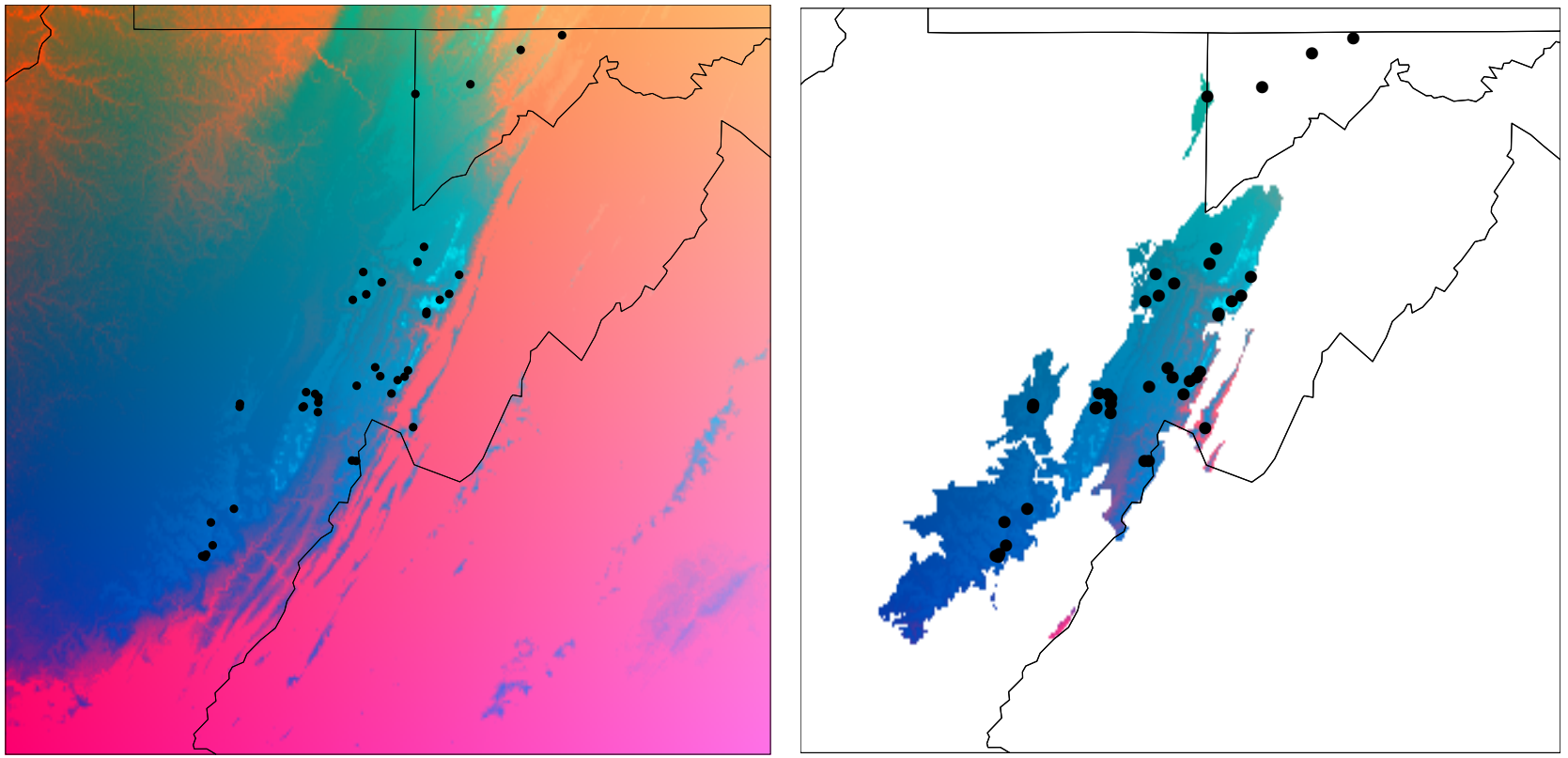

Figure 5. Predicted spatial variation in population genetic differentiation based on Generalized Dissimilarity Modeling (GDM). Areas with similar colors represent areas of predicted genetic similarity based on GDM transformation of bioclimatic variables and their association with genetic distance. The left panel shows the predicted gradient across the region. The right panel shows the same gradient cropped using the WV DNR Red Spruce (*Picea rubens*) Cover in West Virginia 2013 map layer in order to better show the predicted connectivity within the core distributional range of red spruce in WV.

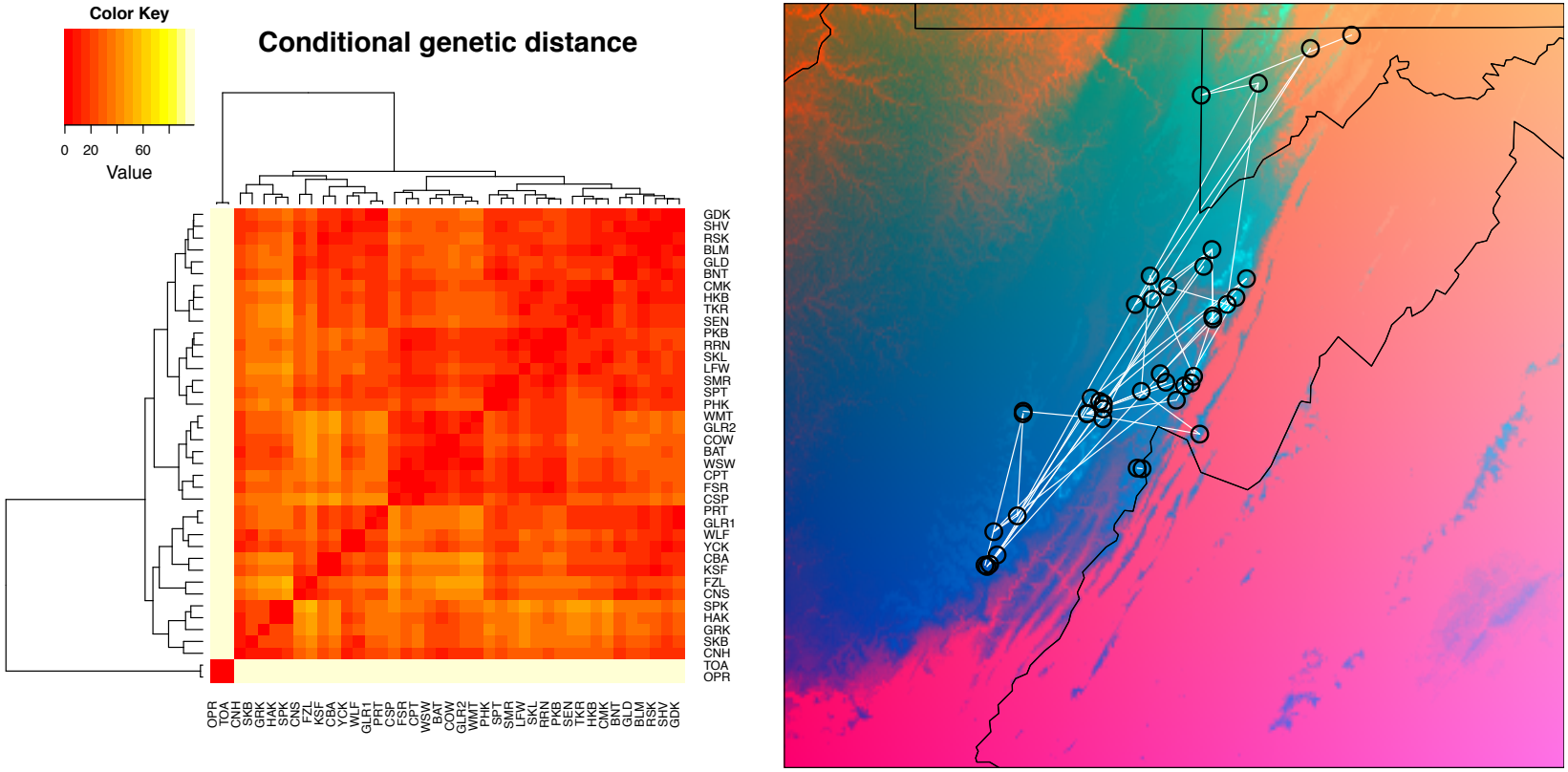

Figure 6. Population graph genetic network structure among Central Appalachian sites. Left panel shows heat map and hierarchical clustering of sampling sites according to their conditional genetic distances. Warmer colors indicate higher connectivity. Right panel spatially maps the population graph network topology of sites (nodes; black circles) linked by conditional genetic distances (white edges). Network topology is overlaid onto the genetically transformed climatic surface from GDM.

The population graph also reveals several additional insights into the genetic connectivity among sites that were not apparent from other analyses. The most striking feature being the almost complete lack of connectivity of a pair of populations along the eastern edge of *P. rubens* in WV, namely sites TOA (Top of Allegheny) and OPR (Old Piney Road), although both sites appear well connected to each other (Figure 6, left panel). Other sites also show somewhat limited connectivity within the network such as KSF and PRT in the Kumbrabow State Forest, and PKB to the east.

Effective Population Size and Demographic History

Current best estimates of the effective population size (N_e) of *P. rubens* in the Central Appalachians using the linkage-disequilibrium method implemented in NeEstimator are low: N_e = 322, with a 95% C.I. of 287-364 (Table 4). This is slightly lower than the minimum N_e recommended for long-term viable populations, and reflects the known low genetic diversity reported for *P. rubens* in other genetic studies (Hawley & DeHayes 1994; Perron *et al.* 2000). However, when put into a broader regional context, the Central Appalachian N_e is large relative to the N_e estimated for northeastern *P. rubens* (ca. 68 individuals, CI: 60-77), and represents about 76% of the total species-wide N_e estimated from the pooled sample of all individuals (Table 4). However, these conclusions regarding N_e in the northeast remain very tentative until much more thorough sampling can be conducted. The NGR restoration seedlings captured a surprisingly large fraction of the available N_e in the Central Appalachians, N_e = 157 (CI: 123-212), roughly half of the regional N_e (Table 4).

Table 4. Estimates of current effective population size (N_e).

Sample Group	Sample size (N individuals)	Effective population size (N_e)	Lower 95% C.I.	Upper 95% C.I.
Central Appalachians (WV & MD)	643	322	287	364
Restoration seedlings (NGR)	100	157	123	212
Northeast (PA, NY, NH, VT, ON)	102	68	60	77
Pooled	845	539	473	619

Estimates of current N_e based on the coalescent method in VarEff were highly congruent with those from NeEstimator. Assuming a mutation rate of $5x10^{-4}$, VarEff estimates a current N_e for the Central Appalachian region of 535.7 (95% CI: 88.5–1765.3)(Figure 7). This is slightly larger than the N_e from NeEstimator, but the confidence interval overlaps. Thus, given two alternate methods that make use of different signatures of genetic drift in the data, there emerges a relatively robust picture that current N_e is in the hundreds, and very close to the minimum recommended level advised by conservation biologists for long-term viability.

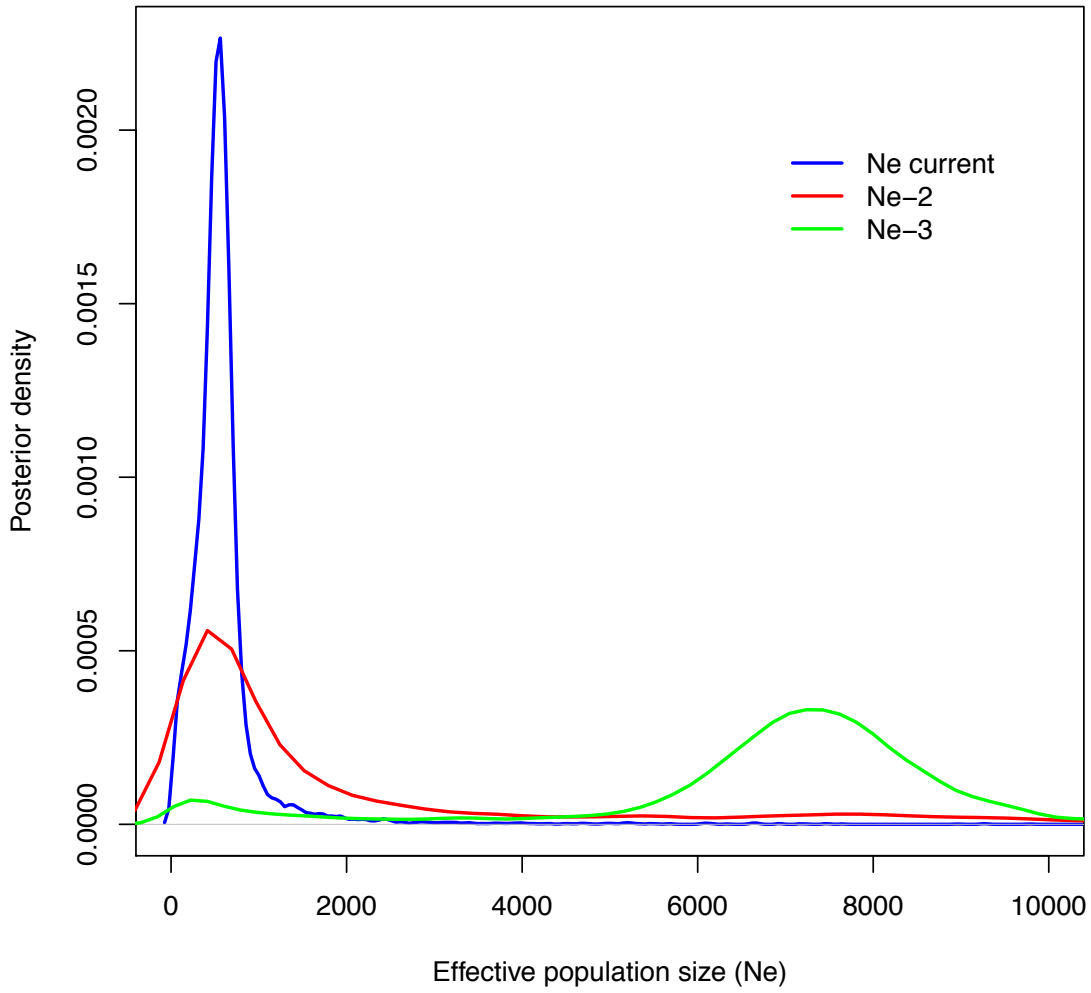

Figure 7. Coalescent model of effective population size (N_e) of red spruce at different time periods in the Central Appalachians. N_e is shown at time points corresponding to the ancestral N_e before the size change (N_e-3; green), the N_e following the size change (N_e-2; red), and the current N_e (blue).

Interestingly, N_e in the region has not always been at this low level. The size change model in VarEff estimates that the ancestral population to Central Appalachian *P. rubens* had a much larger N_e in the past (Figure 8). The size of this ancestral population to the Central Appalachian region is estimated to be 7,209 (95% CI: 268.3-10,248.5) -- an order of magnitude larger than current N_e estimates, but with a wide confidence interval reflecting the uncertainty inherent in estimating changes in the evolutionary past. The timing of the size change is estimated at 376 generations ago (95% CI: 13.3-1,229.0). Depending on the value assumed for generation time, this places the timing of the size reduction sometime in the mid to late Holocene, probably reflecting historical reductions in abundance as the post-glacial climate warmed. The dominant signature left in the genetic data reflect this more ancient event, and apparently overwhelm any effects on N_e from more recent events. This suggests that *P. rubens* was already quite bottlenecked in terms of its genetic diversity prior to the onset of land use change in the 19th and 20th centuries.

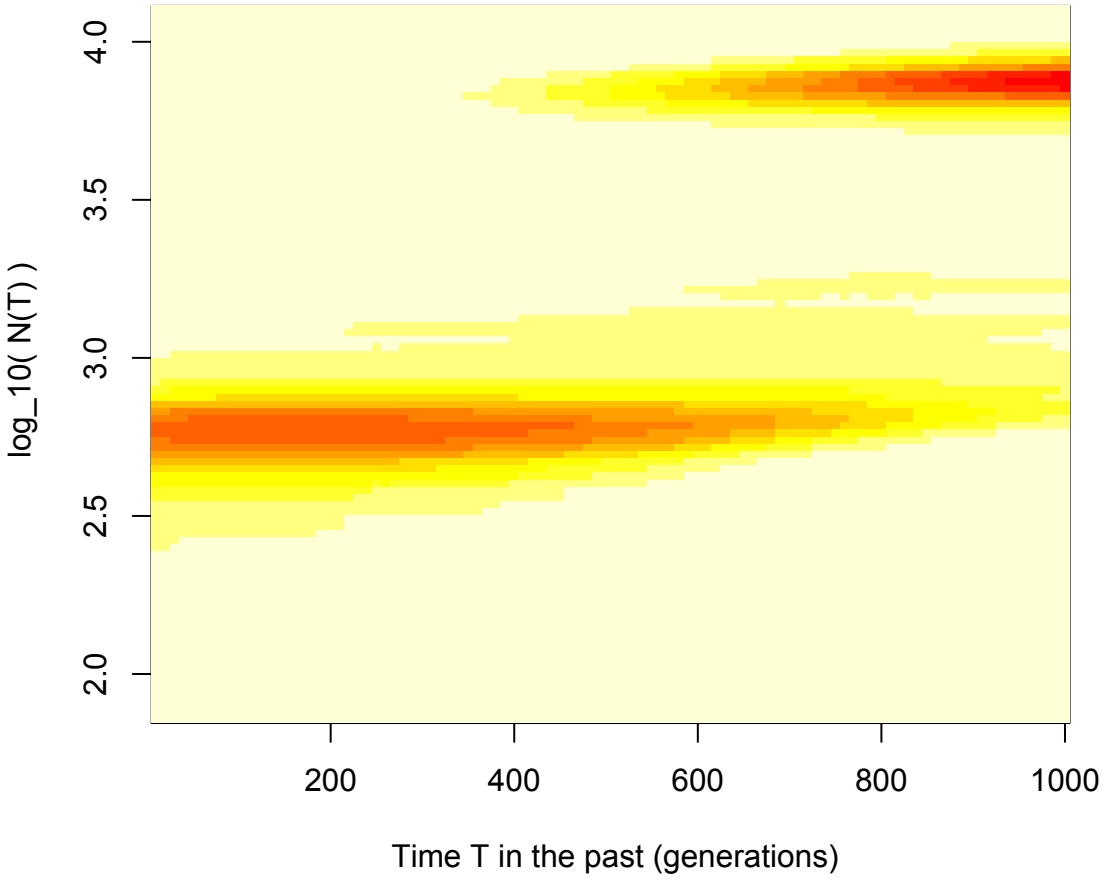

Figure 8. 2-D posterior density distributions of (log10) effective population size change through time., measured as generations in the past. Warmer colors indicate areas of greater posterior probability.

Conclusions

The results of this research provide a first, detailed assessment of the genetic status of remnant populations in *P. rubens,* a keystone species in high elevation conifer forests. While many questions remain, the results here should provide some guidance to current and future ecosystem restoration and management efforts in the Central Appalachian eco-region seeking to maximize the distribution of diversity and genetic connectivity in order to maintain a population that possess long-term viability. Below we provide some highlights and synthesis of the more striking results to come out of this study and their potential relevance to red spruce restoration:

- Genetic diversity of *P. rubens* is quite low, particularly for an outcrossing forest tree. This is reflected both in the low levels of allelic richness within sites (Table 3) and the very small values of its current effective population size (Table 4; Figure 7). Low levels of genetic diversity have also been found by other population genetic studies of red spruce in different portions of its range, and appear to be a unique feature of this taxon that stands out among other wind-pollinated outcrossing forest trees. Some researchers have posited that *P. rubens* is actually a recently derived (and not completely reproductively isolated) species from *P. mariana* (Perron *et al.* 2000; Jaramillo-Correa & Bousquet 2003; De Lafontaine *et al.* 2015).

- The low current N_e is near or slightly below what is generally recognized as a minimum acceptable effective population size for conservation of impacted taxa (Waples 2002; Reed & Frankham 2003; Frankham 2005; Latta 2008). These studies suggest that the ratio of N_e to census population size is often ~0.1, but such a ratio gives rise to unrealistically small census size estimates of red spruce in the Central Appalachians (on the order of a few thousand). Thus, it is more likely that the N_e/N ratio is therefore even smaller in *P. rubens*, possibly reflecting its demographic history of recent founder effect speciation during the Pleistocene with *P. mariana* as well as the historical bottleneck detected here in the mid-late Holocene that reduced N_e by an order of magnitude. Complications due to prolonged periods at low N_e include reduced ability to purge deleterious mutations and avoid inbreeding depression, as well as limited capacity for response to novel selection pressures, such as emerging pests and pathogens or a changing climate (Charlesworth 2009). Thus, restoration practices should work to maximize N_e of Central Appalachian *P. rubens* by drawing from a diverse seed stock within the region, and maximizing the allelic diversity of restoration seedlings. That said, the N_e present in the of Central Appalachians appears to be large relative to other parts of the species range, but confidence in this assessment will have to await further sampling of *P. rubens* in the northeast.

- Presence of genetic ancestry in the Central Appalachians characteristic of northeastern red spruce diversity suggests historic anthropogenic movement of genes from further north. The effect of this movement represents a potential opportunity to assess the consequences of long-distance gene flow as a mechanism of genetic rescue, or if such gene flow is maladaptive by diluting local adaptation. Detailed comparisons of trees with vs. without northern ancestry growing in the same micro-environment would be potentially revealing on this point.

- Genetic connectivity among Central Appalachian stands is high, but there are isolated pockets of populations separated from the core of the region by marginal climate conditions. The GDM analysis indicates summer precipitation is a key factor regulating this connectivity. Western MD sites are notable here, being climatically and geographically isolated from the rest of the Central Appalachian spruce, as well as showing reduced diversity and evidence for inbreeding. It is important to note that this study purposefully sampled trees deemed to be native and avoided the restoration plantings conducted by the Nature Conservancy in this region. Thus, the potential that these plantings might supplement the existing genetic diversity of these stands would seem to be high, and could be assessed as the plantings mature.

Acknowledgements

I am grateful for the assistance and advice of many people who contributed to this project. Most importantly, my research assistant Regina Trott who was highly committed to the collection effort and molecular analyses. I also appreciate many biologists with expertise on the red spruce ecosystem who assisted me in locating and accessing samples, including Kent Karriker, Jim Vanderhorst, Brian Streets, Elizabeth Byers, Dave Saville, Mike Powell, and Deborah Landau. Finally, I acknowledge the WVDNR and the University of Maryland Center for Environmental Science for funding support.

Appendices

Appendix A. Individual sampling information.

Appendix B. Excel file containing estimates of pairwise genetic divergence (G_{ST}, G'_{ST}) among populations.

Appendix C: STRUCTURE ancestry assignments of individuals to genetic clusters.

Appendix D. Microsatellite genotype data for all sampled individuals at 18 loci.

References

Bates D, Maechler M, Bolker B, Walker S (2015) lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8. *http://CRAN.R-project.org/package=lme4*.

Beane NR (2010) Using Environmental and Site-specific Variables to Model Current and Future Distribution of Red Spruce (Picea rubens Sarg.) Forest Habitat in West Virginia. PhD Thesis. West Virginia University.

Blacket MJ, Robin C, Good RT, Lee SF, Miller AD (2012) Universal primers for fluorescent labelling of PCR fragments-an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources*, **12**, 456–463.

Byers E, Norris S (2011) *Climate change vulnerability assessment of species of concern in West Virginia*. West Virginia Division of Natural Resources, Elkins, WV.

Byers BEA, Vanderhorst JP, Streets BP (2010) *Classification and Conservation Assessment of Upland Red Spruce Communities in West Virginia*. West Virginia Division of Natural Resources, Elkins, WV.

Charlesworth B (2009) Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature Reviews. Genetics*, **10**, 195–205.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Dyer RJ (2014) Population Graphs and Landscape Genetics. *Annual Review of*

Ecology, Evolution, and Systematics, **46**, annurev–ecolsys–112414–054150.

Dyer RJ, Nason JD (2004) Population Graphs: The graph theoretic shape of genetic structure. *Molecular Ecology*, **13**, 1713–1727.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–20.

Ferrier S, Manion G, Elith J, Richardson K (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, **13**, 252–264.

Fitzpatrick MC, Keller SR (2015) Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, **18**, 1–16.

Frankham R (2005) Genetics and extinction. *Biological Conservation*, **126**, 131–140.

Hawley GJ, DeHayes DH (1994) Genetic diversity and population structure of red spruce (Picea rubens). *Canadian Journal of Botany*, **72**, 1778–1786.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hodgetts RB, Aleksiuk MA, Brown A *et al.* (2001) Development of microsatellite markers for white spruce (Picea glauca) and related species. *Theoretical and Applied Genetics*, **102**, 1252–1258.

Jaramillo-Correa JP, Bousquet J (2003) New evidence from mitochondrial DNA of a progenitor-derivative species relationship between black spruce and red

spruce (Pinaceae). *American Journal of Botany*, **90**, 1801–1806.

Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA (2013) DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, **4**, 782–788.

De Lafontaine G, Prunier J, Gérardi S, Bousquet J (2015) Tracking the progression of speciation: Variable patterns of introgression across the genome provide insights on the species delimitation between progenitor-derivative spruces (Picea mariana × P. rubens). *Molecular Ecology*, **24**, 5229–5247.

Latta RG (2008) Conservation genetics as applied evolution: from genetic pattern to evolutionary process. *Evolutionary Applications*, **1**, 84–94.

Little EL (1971) *Atlas of United States trees: Vol. 1. Conifers and important hardwoods*. USDA Forest Service.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Manion G, Lisk M, Ferrier S, Nieto-Lugilde D, Fitzpatrick MC (2016) gdm: Functions for Generalized Dissimilarity Modeling.

Marriage TN, Hudman S, Mort ME *et al.* (2009) Direct estimation of the mutation rate at dinucleotide microsatellite loci in Arabidopsis thaliana (Brassicaceae). *Heredity*, **103**, 310–7.

Matschiner M, Salzburger W (2009) TANDEM: Integrating automated allele binning into genetics and genomics workflows. *Bioinformatics*, **25**, 1982–1983.

McKay JK, Christian CE, Harrison S, Rice KJ (2005) "'How Local Is Local ?'"— A

Review of Practical and Conceptual Issues in the Genetics of Restoration. *Restoration Ecology*, **13**, 432–440.

Mosseler A, Major J, Rajora O (2003) Old-growth red spruce forests as reservoirs of genetic diversity and reproductive fitness. *TAG Theoretical and Applied Genetics*, **106**, 931–937.

Nikolic N, Chevalet C (2014) Detecting past changes of effective population size. *Evolutionary Applications*, **7**, 663–681.

Perron M, Perry DJ, Andalo C, Bousquet J (2000) Evidence from sequence-tagged-site markers of a recent progenitor-derivative species pair in conifers. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 11331–11336.

Pfeiffer A, Olivieri AM, Morgante M (1997) Identification and characterization of microsatellites in Norway spruce (Picea abies K.). *Genome*, **40**, 411–419.

R Core Team (2014) R: A language and environment for statistical computing.

Rajora OP, Mosseler A, Major JE (2000) Indicators of population viability in red spruce, Picea rubens. II. Genetic diversity, population structure, and mating behavior. *Canadian Journal of Botany*, **78**, 941–956.

Rajora OP, Rahman MH, Dayanandan S, Mosseler A (2001) Isolation, characterization, inheritance and linkage of microsatellite DNA markers in white spruce (Picea glauca) and their usefulness in other spruce species. *Molecular and General Genetics*, **264**, 871–882.

Reed DH, Frankham R (2003) Correlation between Fitness and Genetic Diversity. *Conservation Biology*, **17**, 230–237.

Rungis D, Bérubé Y, Zhang J *et al.* (2004) Robust simple sequence repeat markers for spruce (Picea spp.) from expressed sequence tags. *Theoretical and Applied Genetics*, **109**, 1283–1294.

Savolainen O, Pyhajarvi T, Knurr T (2007) Gene flow and local adaptation in tees. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 595–619.

Scotti I, Paglia GP, Magni F, Morgante M (2002) Efficient development of dinucleotide microsatellite markers in Norway spruce (Picea abies Karst.) through dot-blot selection. *Theoretical and Applied Genetics*, **104**, 1035–1041.

Waples R (2002) Definition and estimation of effective population size in the conservation of endangered species. *Population viability analysis*.

WVDNR (2017) Plots2-WV database of community ecology plots. West Virginia Natural Heritage Program, WV Division of Natural Resources, Elkins, WV.