# Admixture mapping and selection scans identify genomic regions associated with stomatal patterning and disease resistance in hybrid poplars.

Karl C. Fetter[1,2] and Stephen R. Keller [*1]

[1]Department of Plant Biology, University of Vermont
[2]Department of Ecology and Evolutionary Biology, University of Connecticut

May 25, 2023

**Abstract**

Variation in fitness components can be linked in some cases to variation of key traits. Metric traits that lie at the intersection of development, defense, and ecological interactions may be expected to experience strong environmental selection, informing our understanding of evolutionary and ecological processes. Here, we use quantitative genetic and population genomic methods to investigate disease dynamics in hybrid and non-hybrid populations. We focus our investigation on morphological and ecophysiological traits which inform our understanding of physiology, growth, and defense against a pathogen. In particular, we investigate stomata, microscopic pores on the surface of a leaf which regulate gas exchange during photosynthesis and are sites of entry for various plant pathogens. Stomatal patterning traits were highly predictive of disease risk. Admixture mapping identified a polygenic basis of disease resistance. Candidate genes for stomatal and disease resistance map to the same genomic regions, and are experiencing positive selection. Genes with functions for guard cell homeostasis, the plant immune system, components of constitutive defenses, and growth related transcription factors were identified. Our results indicate positive selection is filtering genetic variance from one of the parental species maladpated to a novel pathogen, and changing suites of stomatal traits which contribute to disease variation in natural populations.

**Keywords:** Admixture mapping, stomata, trade-offs, *Populus*, *Melampsora*.
**Orcid:**
KCF - https://orcid.org/0000-0002-9234-9300
SRK - https://orcid.org/0000-0001-8887-9213

[*]Corresponding author: srkeller@uvm.edu

# 1 Introduction

The evolutionary and ecological consequences of hybridization across landscapes have received considerable attention in plants, particularly in sunflowers (Rieseberg *et al.,* 2007), monkeyflowers (Chase *et al.,* 2017), and poplars (Suarez-Gonzalez *et al.,* 2018). Hybridization studies in poplars have been influential in developing our understanding of how admixture between foundational tree species can alter biotic interactions between herbivores (Whitham *et al.,* 1996), fungi (Bailey *et al.,* 2005), and entire ecosystems (Bailey *et al.,* 2009). Underlying ecological change in hybrid populations are the novel phenotypes expressed as a result of gene flow between isolated genomes, heterosis, and transgressive segregation, among other phenomena.

Research from poplar hybrid zones indicate their genomes are porous to the movement of genetic variants (Suarez-Gonzalez *et al.,* 2018). Multiple mechanism of evolutionary change have been studied in poplar hybrid zones, in particular adaptive introgression, when a gene evolved in one species is introduced to another via reproduction, and subsequently experiences selection to maintain its function (Rieseberg & Carney, 1998). However, less attention has been paid to the maladaptation of hybrids and the phenotypic and genomic effects of hybridization and backcrossing in natural systems. Plant diseases and pests have been indicated as important players structuring ecological systems (Floate *et al.,* 2016). Hybrids, while often thought of as 'super genotypes' in plant breeding, are susceptible to decreased fitness as a results of trade-offs and mismatches of growth-defense syndromes that have evolved in different evolutionary and ecological contexts (Fetter *et al.,* 2021). Advanced generation hybrids offer a useful opportunity to employ methods to identify trade-offs and to use genomic methods to categorize the genes associated with ecologically relevant traits.

A major goal of ecological genomics is to link phenotypic and genomic variation to find genes underlying ecological processes. Advances in sequencing technology have enabled the discovery of genetic polymorphisms covering a large fraction of a genome. While many association genetic methods have been developed to help control for false positives arising from population structure within a species, a sample generated from hybrids is a considerably more difficult problem (Shriner *et al.,* 2011), due to large blocks of loci in high LD as a result of relatively few recombination events in hybrids. Association genetic methods based on allele frequency variation are likely to yeild many false positives in hybrid populations. Admixture mapping was invented to enable association genetics in hybrid or admixed populations (Smith & O'Brien, 2005). Rather than using the identity of a base pair at a locus (i.e. A, G, T, or C), a locus is represented as homozygous or heterozygous for ancestry from either parental species/population. Thus, the phenotypic associations are made to the local ancestry of a locus (Shriner *et al.,* 2011).

*Populus* is a genus of long-lived, wind-pollinated trees which exhibit remarkable ecological amplitude. Species in the genus occur beyond the tundra-boreal forest ecotone in northern Alaska and Canada (e.g. *P. balsamifera*, Breen, 2014). Poplars extend to the humid subtropical forests of the southeastern American coastal plain (e.g. *P. heterophylla*), and to the arid deserts of the the Middle East (e.g. *P. euphratica*). *Populus* species will readily form fertile hybrids, and bi-directional crossing is frequent (Suarez-Gonzalez *et al.,* 2018), although not all species can do this, (e.g. *P. balsamifera* x *deltoides* hybrids will only cross into *P. balsamifera*, Thompson *et al.,* 2010). Given the promiscuity of poplars and their ability to persist as clones on the landscape, hybrid backcrosses are expected to occur, creating the biological conditions to employ admixture mapping. In a previous study, we found evidence of hybrids expressing a strong genetic correlation between stomatal traits and resistance to the basidiomycete leaf rust, *Melampsora medusae* (Fetter *et al.,* 2021). *M. medusaue* is a foliar pathogen with two obligate hosts: a poplar as the telial host, and a larch (*Larix*) as the aecial host (Feau *et al.,* 2007). Here, we explore variation in disease and its relationship to stomatal and ecophysiological traits using quantitative genetic models. We then use admixture mapping to identify genomic region associated to traits. Finally, we test if candidate associations show evidence of positive selection that is potentially driven by pathogen-induced selection.

# 2 Materials & Methods

## 2.1 Plant collections, phenotypes, and sequencing

Plant material, genotypes, and phenotypes were previously described in Fetter *et al.* (2021). Briefly, plant material was collected in 2013 and 2014 from central and western North American populations of *P. balsamifera* (Fig. 1A). Cuttings were stored in a cold room until propagated in conetainers and planted in a common garden in Vermont, USA in June, 2014. After overwintering, ecophysiology, height growth, and bud phenology traits were measured (Table 1). Disease severity to *M. medusae* was measured in 2015 using the ordinal scale of La Mantia *et al.* (2013), and in 2016, using both the ordinal scales of La Mantia *et al.* (2013) and Dowkiw & Bastien (2004) (Table 1). Genomic DNA was extracted with DNeasy 96 Plant Mini Kits (Qiagen, Valencia, CA, USA) and libraries prepared for genotyping-by-sequencing (GBS). Sequence reads were obtained from an Illumina HiSeq 2500 to generate 100 bp single end reads. Reads were mapped to the *P. trichocarpa* reference assembly version 3.0 (Tuskan *et al.,* 2006) and SNPs obtained using a modified Tassel pipeline (Glaubitz *et al.,* 2014). SNPs with a minor allele frequency < 0.001 were removed, and only biallelic sites were retained. Sites with with a mean depth < 5, genotype quality > 90, and indels were removed. Missing data were imputed with Beagle v5.0 (Browning *et al.,* 2018), and sites with post-imputation genotype probability < 90 and sites with any missingness were removed. A total of 227,607 SNPs were called. Sequence reads are available at the SRA (SRX1605454-68).

## 2.2 Trait modeling

Best Linear Unbiased Predictors (BLUPs) were fit for each trait from data collected from clones. Each model included the garden row and column position as fixed effects and individual code as a random effect. To identify traits that predicted disease presence or absence, a logistic model was fit in R (R Core Team, 2021) using the glm function. Disease severity in 2016 (D2) was converted to a binary presence/absence trait by calling BLUPs greater than zero as disease presence. D2 was preferred to other disease traits, as its distribution more clearly lent itself to binary factorization (Fig. S1). Regression coefficients were standardized by dividing them by two standard deviations (sensu Gelman, 2008) before plotting.

After evaluating the logistic model, several traits suggested trade-offs with disease resistance. We fit random slopes and intercepts models from BLUPs with disease resistance (from D1) as the response, and relative growth rate, stomatal ratio, log-stomatal density as predictors in separate models. A model with stomatal ratio (response) and log-stomatal density (D, predictor) was also fit. The hybrid set was included as a grouping effect in the model. Models were fit with brms (Bürkner, 2017).

## 2.3 Admixture mapping

Admixture mapping requires phased reference haplotypes from unadmixed parental populations to estimate locus specific ancestry in a test population. To choose reference individuals for *P. balsamifera*, we estimated global ancestry using ADMIXTURE (Alexander *et al.,* 2009) from the 534 individuals we sequenced. We chose 25 individuals from the western populations of Duck Mountain, Manitoba (DCK) and Hudson Bay, Ontario (HBY) that exhibited minimal signs of admixture with other *Populus* species (i.e. ADMIXTURE q-matrix > 0.95 *P. balsamifera* at K = 2) (Fig. S2). For the *P. trichocarpa* reference set, 25 individuals with whole genome sequences publicly available were chosen from western Washington which were known to lack admixture with *P. balsamifera* (Evans *et al.,* 2014). Whole genome sequences were downloaded from https://phytozome.jgi.doe.gov/. The test population included 117 individuals from the Canadian prairie provinces and western and mid-western states (Table S1).

The locus positions of the test vcf files were used to filter the *P. trichocarpa* reference set, and after discarding flipped and multi-allelic sites, 74,878 homologous sites remained. Haplotypes of the reference populations were jointly estimated with fastPHASE (Scheet & Stephens, 2006), setting the following parameters: 20 random starts; 45 EM iterations per run; 5000 samples of the posterior haplotype distribution; the K-selection function was limited between 5 and 30, at 5 unit intervals; and loci with genotype probability <90 were flagged. Locus specific ancestry was determined for the test set via recombination ancestry switchpoint probability estimation with the program RASPberry (Wegmann *et al.,* 2011). The input data sets for RASPberry were: a test set of 485 individuals with 227,607 SNP loci and no site-wise missingness; two reference populations of 25 individuals each with 74,878 SNP loci; and the ADMIXTURE q-matrix at K = 2. The default recombination rate of 5 cM was used and population recombination rates set to 120 and 173 for *P. balsamifera* and *P. trichocarpa*. The mutation rate was set to 0.0079365, and miscopy rate to 0.01.

Mixed effects models were fit to identify association between 24 phenotypes (Table 1) and local ancestry genotypes corrected by the global ancestry of the individual using the bayesian mixed model (BMIX) of Shriner *et al.* (2011):

$$f(y_i) = \beta_0 + \beta_1 A_{ij} + \beta_2 \bar{A}_i + \varepsilon_i$$

Where, $y_i$ is a vector of BLUPs, $\beta_N$ are regression coefficients to estimate, $A_{ij}$ is a vector of local ancestries for the $j^{\text{th}}$ locus, $\bar{A}_i$ is a vector of global ancestries for the $i^{\text{th}}$ individual, and $\varepsilon_i$ is error variance for the $i^{\text{th}}$ individual. Global ancestry was again estimated for each individual with the RASPberry data by summing the frequency of the local ancestries of the homozygous and half the heterozyous genotypes corresponding to the *P. balsamifera* allele. Association probabilities were estimated from chi-sq test statistics of converted model p-values. The significance level for true associations was set to $\alpha$ / admixture burden (0.05 / 237.5), which represents the number of independent tests in the sample. Admixture burden was estimated from the first-order autoregressive (AR(1)) models for each locus summed across the genome for all individuals using the function ar from the stats package in R. Manhattan plots were used to visualize p-values of tests across the genome for each trait. Intersections between candidate gene lists were identified with UpSetR (Conway *et al.,* 2017).

## 2.4 Selection Scans

Patterns of positive selection were identified with RAiSD (Alachiotis & Pavlidis, 2018) using the $\mu$ statistic, a composite statistic of the product of $\mu^{SFS}$ which measured shifts in the site frequency spectrum, $\mu^{LD}$ measuring linkage disequilibrium, and $\mu^{VAR}$ which measured genetic polymorphism. These statistics were calculated in overlapping 50 SNP sliding windows. RAiSD is ideal for identifying hard sweeps, but false positives can be generated by population bottlenecks, background selection, and population structure (Alachiotis & Pavlidis, 2018). As a result, we sub-sampled the *P. balsamifera* data set to a single deme in the western core which was identified by running ADMIXTURE (Alexander *et al.,* 2009) on all 534 *P. balsamifera* samples from K = 2 until resolution of known demes (*sensu* Keller *et al.,* 2010) within *P. balsamifera* was possible at K = 7 (Fig. S2). Individuals > 0.98 ancestry in the western core deme at K = 7 were selected. The hybrid set was selected from individuals with RASPberry global ancestry < 0.99. RAiSD was run separately for the *P. trichocarpa* reference individuals (N = 46), Western core *P. balsamifera* (N = 81), and hybrids (N = 39) (Table S4). The input sequence data set contained 31,545 SNPs common to both *P. trichocarpa* and *P. balsamiera* with no missingness. We considered windows in the top 1% of $\mu$ statistics as evidence of a hard selective sweep.

## 2.5 Gene annotation and candidate gene filtering

Gene annotations for the 227,607 SNPs were downloaded from the *P. trichocarpa* v3.0 genome and used to annotate significant loci. After the top 1% of $\mu$ statistic windows were identified, the overlap

of BMIX candidate genes and selection windows was determined using GenomicRanges (Lawrence *et al.,* 2013). Local ancestry sites from genes that passed the BMIX/selection filter were identified, and monomorphic sites removed. With the remaining polymorphic local ancestry sites, boxplots comparing local ancestry genotype and SR, D1, and G were made to evaluate distributions for signs of false positives. Sites that passed this final filter were mapped to *P. trichocarpa* gene annotations and the genes were manually investigated for gene function using www.popgenie.org (Sjödin *et al.,* 2009; Sundell *et al.,* 2015), atgenie.org (Sundell *et al.,* 2015), and TAIR (Berardini *et al.,* 2015). *Populus* gene orthologies to *Arabidopsis* were determined by the best BLAST hit on www.popgenie.org or via manually BLAST on The Arabidopsis Information Resource (www.arabidopsis.org).

# 3 Results

## 3.1 Global and local ancestry of hybrids

Sequence filtering and merging of the combined 117 test individuals and 50 reference individuals yielded 31,523 SNPs for global and local ancestry estimation. Global ancestry estimates calculated from the local ancestries indicated we sampled a range of hybrid ancestries from unadmixed (RASPberry K2 q-matrix = 0.9996) to admixed (RASPberry K2 q-matrix 0.5754) (Fig. 1C). Based on previous analyses , the filial generation of the test set was known to include 18 P1.$F_2$, 6 P1.P1$F_1$, 10 P1.P1$F_1$ and 83 unadmixed *P. balsamifera* (Fetter *et al.,* 2021).

The locus-specific ancestries revealed a patchwork of introgressing loci, where 97.2% of loci were homozygous for the *balsamifera*-ancestry allele (N = 3,417,375), 7.1% were heterozygous (N=261,555), and only 0.2% (N = 8,664) were homozygous for the *trichocarpa*-ancestry alleles (Fig. 1D). The predominance of heterozygous local ancestry sites was consistent with our expectations, given that the majority of samples are derived from advanced generation backcrosses into *balsamifera*.

## 3.2 Trait variation and models

The traits we measured fall into three general patterns of distributions: normal, left-skewed and zero inflated (Fig. 1B). Traits with normal distributions included the elemental and isotopic traits, budflush, relative growth rate and abaxial stomatal traits. Left-skewed traits include the three stomatal ratio traits (SR, LR, PR) and the adaixal stomatal traits which were left-skewed as a result of many unadmixed *balsamifera* lacking adaxial stomata. The disease phenotypes were all zero-inflated. After converting D2 into a binary disease presence/absence category, unadmixed *balsamifera* had 12 individuals with disease (9.8% of all *balsamifera*) and 110 without disease. Hybrids with *trichocarpa* ancestry had 19 individuals with disease (47.5%) and 21 individuals without disease sign (52.5%).

We fit a logistic model of the presence/absence of disease to the 19 traits and global ancestry (Fig. 2A, Table S2). Stomatal traits that were the sum of adaxial (upper) and abaxial (lower) traits were excluded from the logistic model. SR had the largest odds ratio (1.07e06), and PR had the smallest (4.66e-07), indicating variation of these traits contributed the most variance to the presence or absence of disease. Growth had a slight positive effect on disease presence (slope = 3.14, p-value = 0.038). Ecophysiology and bud flush traits explained little variance of disease presence, although SLA had a significant negative effect (slope = 7.1e-02).

Trade-offs (i.e. negative slopes) between disease resistance and growth, stomatal ratio, total stomatal density were recovered from random slope and intercept models of BLUPs with hybrid set as a grouping effect (Fig. 2B-D). In the growth-resistance model, the intercepts for each hybrid set were offset by a value of 0.65, but the slopes were similarly negative for BxB (-0.41) and BxT (-0.46) accessions. The similar slopes indicate the decline of resistance as growth increases has a similar effect in both genetic backgrounds, but a substantial offset in the intercepts can be explained by the different genotypic variance between each parental species. In contrast, the intercepts and slopes for the

SR-resistance model (Fig. 2C) were similar for both groups (Table S3), indicating the effect of adding more stomata to the upper leaf surface has a similar decay in resistance in both hybrid sets. Increased stomatal density has a nuanced effect on resistance, with only a slightly negative effect in unadmixed *balsamifera* (slope = -0.0023) but a significantly negative effect in *trichocarpa* hybrids (slope = -0.01 [-0.0157, -0.0067] 95% C.I.), suggesting the genotypic variance for this pair of traits is fundamentally different. The model for log total stomatal density and SR demonstrates how stomata are deferentially apportioned, with admixed genotypes shifting more stomata to the upper surface in response to increased stomatal density (Fig. 2E). Unadmixed *balsamifera* typically decrease the size of stomata to fit more on a leaf surface when density increases (linear model D∼S slope = -0.1074***).

## 3.3 Admixture mapping

Admixture mapping was performed with BMIX (Shriner *et al.,* 2011) using the 31,523 locus-specific ancestry estimates and 24 traits (Fig. 3). Out of 746,928 tests, 3.2% (23,670 tests) had p-values larger than the admixture burden corrected p-value threshold ($\alpha$ = 0.05 / 237.50) (Figs. 3A, S3, S4, S5). Significant loci were contained among 13,997 genes, of which 28% (3,877) were identified in only one test. Based on the evaluation of the Manhattan plots and the apparent co-localization of significant loci to regions within chromosomes (Fig. 3A), we clustered genes with UpSetR and revealed several groups of genes (Fig 3B). Notably, a group of genes was identified that were significantly associated in BMIX tests to D1, D2, disease presence/absence, PR, SR, LR, and stomatal density and pore length (N = 1142) . Another set of genes that only contained associations to disease traits independent of ecophysiology or stomatal traits was observed (N = 1562). These two sets of genes (disease plus stomata and disease-only) mapped to locations on 7 and 13 chromosomes, respectively, out of the 19 total chromosomes in the *Populus* genome (Fig. 3C). No significant loci were found for relative growth rate.

## 3.4 Selection scans

Positive selection was inferred using a sliding windows approach implemented with RAiSD (Alachiotis & Pavlidis, 2018). 31,545 SNPs were input into the selection scans of sliding windows of 50 SNPs in size yielded 13,395 windows in the hybrid set, 13,863 windows in the *P. trichocarpa* set, and 5,559 windows in the western core *P. balsamifera* set. The median value of the $\mu$ statistic for the *P. trichocarpa*, hybrid, and *P. balsamifera* data sets were 1.13, 2.42, and 3.97, respectively. $\mu$ statistics were plotted by window position (Fig. 4). Nine chromosomal regions contained candidate genes from BMIX analyses that overlapped with the top 1% of $\mu$ statistic windows and contained 271 unique genes (Table 2, Fig. 4).

## 3.5 Candidate genes and local ancestry class phenotypic distributions

To further investigate the 271 genes we identified with BMIX that contained sites within the top 1% of selection windows, we determined the local ancestry sites that mapped to those genes and filtered them for monomorphic sites. We had 535 local ancestry sites that mapped to the 271 genes, and 221 of those sites were polymorphic. Local ancestry genotypes for SR, D1 and G were evaluated to remove sites that had either no phenotypes for the heterozygote local ancestry genotype (53 sites), or had only one individual in the heterozygote local ancestry genotype (5 sites). After filtering, 163 sites remained which shared one of five patterns of phenotypic distributions (Fig. 5). Substituting a *P. trichocarpa* ancestry allele had a large average allelic effect for SR and D1, and to a lesser extent a negative effect for G for sites on chromosome 1 and 11 (Table 3).

We mapped the 163 sites to 99 genes and investigated each gene with popgenie.org (Sjödin *et al.,* 2009), atgenie.org (Sundell *et al.,* 2015), and TAIR (Berardini *et al.,* 2015). We recorded information on gene family, function, and expression profiles that were relevant to disease or stomatal patterning.

Among the 99 candidate genes, we identified genes involved in guard cell functioning, the immune system, detoxicants, lipid biosynthesis and trafficking, growth related proteins, cell wall production, abiotic/biotic stress response, epigenetic regulation, ubiquination, membrane transport, transcription factors, DNA replication, signal transduction, and floral development related genes (Table S5).

# 4 Discussion

Hybrid zones can act as natural evolutionary experiments to observe the effects of shuffling genomic regions into novel genetic backgrounds (Rieseberg *et al.,* 2007; Chase *et al.,* 2017; Christe *et al.,* 2016; Suarez-Gonzalez *et al.,* 2018). Hybrid populations have increased genetic variance compared to parental species, allowing recombination to uncover genetic variants linked to traits. These populations offer unique opportunities to study the genetic basis of disease. Our genetic models showed a trade-off between growth and disease resistance, influenced by stomatal traits, specifically the stomatal density ratio on upper and lower leaf surfaces. Using admixture mapping, we discovered candidate genes experiencing selection related to stomatal patterning, the immune system, and constitutive defenses. These findings provide insights into genomic mechanisms of trait evolution in response to selection from a foliar fungal pathogen.

## 4.1 Stomatal morphology and disease resistance

Stomatal morphology is an important trait evolves in response to a plant's growth strategy (McKown *et al.,* 2019), body type (Muir, 2015), or disease environment (Melotto *et al.,* 2008). Simulation models indicate pathogens play an important role in determining the ratio of stomata on the upper to lower leaf surface (Muir, 2020). Using empirical disease and stomatal morphology data, we identified stomatal traits as being particularly important in explaining disease risk. Increasing stomatal density on the upper leaf surface was observed to correlate with increased growth in *P. trichocarpa* (McKown *et al.,* 2019), while here, we observe a large increase in the risk of disease (log-odds ratio = 1.07e+06) with no benefit of increased growth (random slopes and intercepts output: y-intercept for BxT = -0.44, 95% CI: {-2.3, 1.4}; slope = -0.05, 95% CI: {-0.08, -0.012}). When isolated from the numerous genetic effects of ancestry, stomatal ratio has a similar negative effect on resistance in both genetic groups (Fig. 2D), supporting simulation models which were conducted in the absence of genetic architecture (Muir, 2020).

Interestingly, the effect of increasing total stomatal density on resistance is not the same between hybrid sets (see slopes in Fig. 2D). *P. balsamifera* genotypes tend to pack stomata more tightly on the lower leaf surface with increased density, rather than shifting stomata to the upper leaf surface in *P. trichocarpa* hybrids (Fig. 2E). Contrasting stomatal-growth-defense trait syndromes are a feature of poplars species (Fetter *et al.,* 2021). Trade-offs observed in hybrid and non-hybrid populations indicate stomatal traits are constantly shifting in response to the underlying genotypic variance and biotic environment (Fetter *et al.,* 2021).

## 4.2 Growth-resistance trade-off

We found a trade-off between disease resistance and growth, stomatal ratio, and stomatal density that differed between unadmixed and admixed poplars. Trade-offs have been well-studied in the evolutionary literature as they indicate physiological limits to adaptation, and are responsive to ecological and environmental contexts (Cope *et al.,* 2021). The admixture mapping results indicate the genetic structure of the trade-offs are highly polygenic. Simulation studies indicate polygenic traits under phenotypic selection have higher evolutionary rates than traits with large-effect loci (Kardos & Luikart, 2021). The positive selection we identified tends to support the hypothesis that allele frequencies underlying ecologically important polygenic traits experience positive selection and are capable of rapid

evolution. These data support the growing consensus that many traits important for adaptation in natural environments are polygenic (Bomblies & Peichel, 2022).

## 4.3 Advanced generation hybrids and selection against *trichocarpa* ancestry

We collected a sample from across the southern and western range of *P. balsamifera* with the intent of collecting unadmixed *P. balsamifera* genotypes. However, the prevalence of hybrid zones in *Populus* and sampling dormant cuttings increases the likelihood of collecting hybrids. Genotyping these accessions revealed the tail of a hybrid ancestry distribution across a geographically large region. Backcrossing was observed in the sample, with hybrid ancestry starting at 0.57 and increasing (Fig 1). In a genetic landscape of hybrids, introgression can be expected to occur, and has been demonstrated to underlie important ecological adaptations in poplars. Introgression of an 880-kb genomic region on chromosome 15 from *P. balsamifera* into *P. trichocarpa* was demonstrated to confer increased ecological differentiation, perhaps allowing genotypes with the introgressed region to inhabit climatically challenging sites (Suarez-Gonzalez *et al.,* 2016). In a *P. trichocarpa*, *P. angustifolia*, and *P. balsamifera* trihybrid zone, introgression of soil ion detoxification and photoperiod regulation genes was observed (Chhatre *et al.,* 2018).

Given the documented importance of adaptive introgression in *Populus*, we expected to find genes from *P. tricochrpa* conferring an adaptive advantage in our sample. However, we found little evidence of a fitness component advantage in admixed genotypes with genomic blocks of *trichocarpa* ancestry. Mixed-effect models demonstrated *trichocarpa* hybrids had lower disease resistance overall (see intercept of Fig. 2B). Additionally, the disease cost of increasing stomatal density was higher in *trichocarpa* hybrids (see slope of Fig. 2D), further indicating global *trichocarpa* ancestry was selected against. At a finer scale, admixture mapping identified five chromosomal regions significantly associated with decreased disease resistance that are also associated with decreased growth and increased stomatal ratio (Fig. 5). In all five of these chromosomal regions, the heterozygote has decreased resistance and growth, indicating they are being selected against. These data generally suggest hybrid breakdown in novel disease communities occurs, and limits introgression and gene flow between these species. Hybrid breakdown has been previously reported in *Populus alba* x *treumla* and invoked to explain reproductive isolation between species (Christe *et al.,* 2016). These data suggest negative selection will act to protect a species' genome from introgression in response to increased mortality from disease.

## 4.4 Genomic basis of disease resistance

Admixture mapping revealed a complex, polygenic basis of disease resistance. Candidate genes within regions tagged by the SNP genotyping fell into broad categories related to stomatal function, the plant immune system, constitutive defenses, and growth regulators. While some of the Pfam descriptions of genes associated to disease resistance are obviously involved in defense (e.g LRR-N terminal domain), others are not, and indicate a plant's overall physiology contributes to resistance.

We identified several genes with known functions for stomatal guard cell regulation. Guard cells open and close the aperture pore of a stoma via reversible changes in the concentration of ions, subsequently altering cellular turgor pressure. Reactive oxygen species (ROS) and calcium ions can function as messenger molecules in stomatal signaling pathways, and can be pumped into guard cells to change the ionization of the cell (Lecourieux *et al.,* 2006). Among our candidates are an ion transmembrane transporter (Potri.016G115500) and an ROS-mediated signal transduction protein (Potri.011G112700).

Plant defenses against pathogens include both constitutive and induced defenses. A successful host induced immune response is initiated by recognizing the presence of pathogen associated molecular patterns (PAMPs). PAMP-triggered immunity (PTI) can be induced by the detection of PAMPs, which results in a signaling cascade to initiate a broad-spectrum defensive response by the host plant (Jones & Dangl, 2006). We detected candidate genes for the plant immune system, including signal transduction

proteins (Potri.011G116200), LRR proteins (Potri.011G116900), oxidative stress detoxifying proteins (Potri.011G113000), and a negative regulator of pathogenesis responsive genes (Potri.011G121200). Constitutive defenses can include morphological or chemical defenses that limit colonization or growth of a pathogen. The plant cuticle is composed of lipids which can limit colonization of pathogens on a leaf surface. Differences in cuticle lipid chemistry have been linked to variation of *Melampsora* infection in *P. trichocarpa* (Gonzales-Vigil *et al.,* 2017). We observed six candidate genes involved in lipid biosynthesis or transport (Potri.001G317400, Potri.016G115800, Potri.016G116400, Potri.016G113800, Potri.016G118000, and Potri.001G316600), and two candidate genes involved in cell wall homeostasis (Potri.013G056800, Potri.016G114300).

Finally, we identified several candidate genes involved in transcriptional regulation of growth (Potri.011G115400), cell boundary specification (Potri.011G121300), and an auxin transmembrane tranporter (Potri.016G113600) which may potentially lie at the intersection of growth-defense tradeoffs.

### 4.5 Conclusion

We used admixture mapping to identify genes under selection that are associated with disease severity to a fungal pathogens. These result provide evidence that admixture mapping can be used to find ecologically relevant genes, and supports the hypothesis that variation of loci within genes can have effects that cascade to the ecological relationships and the broader environment (Wymore *et al.,* 2011). In this study, hybrid genotypes serve as a reservoir of disease, an observation shared by other studies in poplars (Whitham, 1989). The shared signals of genomic association between disease and stomatal patterning likely indiates stomatal traits are under strong positive selection, in particular, the stomatal ratio. These results suggests natural selection can effectively purge maladaptive genetic variation in hybrid populations, and that stomata are components of an integrated network of physiological regulation of growth and defense.

## 5 Acknowledgements

This study was funded by a National Science Foundation Award #1461868 to S.R.K. Matt Fitzpatrick, Andrew Elmore, Raju Soolanayakanahally, Steve Guinn aided in collections of dormant cuttings. Dave Nelson provided isotope data. Terrence Delaney provided access to microscopy facilities. Sven Eberhardt helped create an early version of StomataCounter used to phenotype stomata density. Madie Hassett collected disease phenotypes in 2015. Many volunteers from the Plant Biology department planted cuttings in 2014 to establish the Spear Street common garden.

## 6 Data Accessibility

Raw sequence reads are available for download at NCBI SRA (SRP070954). Phenotypes from the common garden are available as supporting information in Fetter *et al.* (2021). Cuticle micrographs are deposited on Dryad (doi:10.5061/dryad.kh2gv5f).

## 7 Author contributions

KCF and SRK wrote and edited the manuscript and phenotyped accessions. KCF performed the statistical analyses. SRK established the common garden and obtained DNA sequenc resources. Both authors read and approved the final manuscript.

# 8 Competing Interests Statement

The authors have no conflict of interest to declare.

# Tables

Table 1: Trait definitions, abbreviations, and units.

| Definition | Abbvr. | Units |
|---|---|---|
| Disease | | |
| Disease severity scale 1 (2015) | D1 | ordinal |
| Disease resistance scale 1 (2015) | R1 | ordinal |
| Disease severity scale 1 (2016) | D2 | ordinal |
| Disease severity scale 2 (2016) | D3 | ordinal |
| Disease presence/absence (2016) | dis_pres | ordinal |
| Stomatal patterning | | |
| Stomatal ratio | SR | none |
| Pore length ratio | LR | none |
| Porosity ratio | PR | none |
| Abaxial (upper) stomatal density | SD_AB | ($mm^2$) |
| Adaxial (lower) stomata density | SD_AD | ($mm^2$) |
| Total stomatal density | D | ($mm^2$) |
| Abaxial (upper) pore length | PL_AB | ($\mu$m) |
| Adaxial (lower) pore length | PL_AD | ($\mu$m) |
| Abaxial (upper) porosity | PO_AB | (none) |
| Adaxial (lower) porosity | PO_AD | (none) |
| Total porosity | TO_PO | (none) |
| Ecophysiology | | |
| Relative growth rate | G | cm |
| Carbon:Nitrogen | CN | none |
| Leaf percent carbon | C | % |
| Leaf percent nitrogen | N | % |
| Carbon isotope discrimination | $\Delta^{13}$C | ‰ |
| Nitrogen isotope value | $\delta^{15}$N | ‰ |
| Specific leaf area | SLA | $mm^2\ mg^{-1}$ |
| Chlorophyll content index | CCI | none |
| Cumulative growing degree days to bud flush | cGDD-15 | days |
| Cumulative growing degree days to bud flush | cGDD-16 | days |

Table 2: Summary of the number of genes contained within each chromosomal region containing both candidate genes from admixture mapping, and $\mu$ values in the top 1%. The number of genes found in each chromosomal region are given in parentheses.

| BMIX Gene Set | *P. trichocarpa* | Hybrids | *P. balsamifera* |
|---|---|---|---|
| Stomata + Disease | Chr11 (46) | Chr11 (46) | |
| Disease | Chr9, Chr11 (2, 57) | Chr11 (18) | Chr1, Chr13, Chr16 (21, 12, 28) |

Table 3: Average allelic effects for substituting a *P. balsamifera* ancestry site (0) for *P. trichocarpa* (1). Phenotypes were scaled and standardized before calculating the allelic effect. The chromosomal coordinates for each region are provided in Fig. 5.

| Trait | Chr1 | Chr9 | Chr11 | Chr13 | Chr16 |
|---|---|---|---|---|---|
| SR | 1.87 | 1.61 | 2.00 | 1.17 | 1.17 |
| D1 | 2.21 | 1.26 | 2.21 | 1.02 | 1.02 |
| G | -1.88 | -1.61 | -2.00 | -1.16 | -1.16 |

# Figures

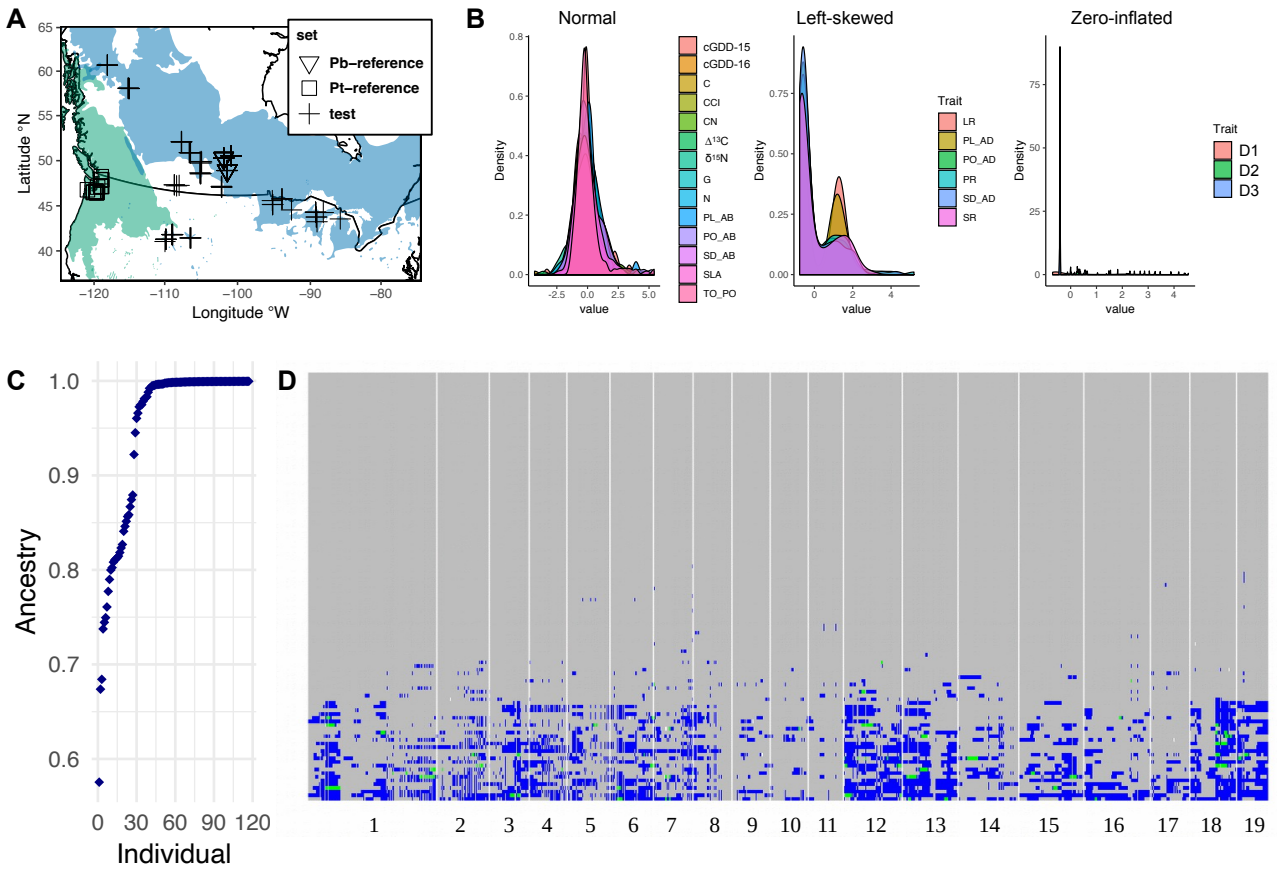


Figure 1: Map of collection localities in western North America with the range of *P. balsamifera* and *P. trichocarpa* colored in blue and green, respectively. Ranges from Little, 1971 (A). BLUPs from traits fall into three distribution categories: approximately normal, left-skewed, and zero-inflated (B). Global ancestry estimates from locus-specific ancestries estimated by RASPberry for the admixture mapping test set (N = 117) (C). Local ancestry for each individual (in rows) and for every site (in columns) colored as gray if the locus is homozygous for the *P. balsamifera* allele, blue for heterozygous loci, and green for homozygous for the *P. trichocrapa* allele (D).

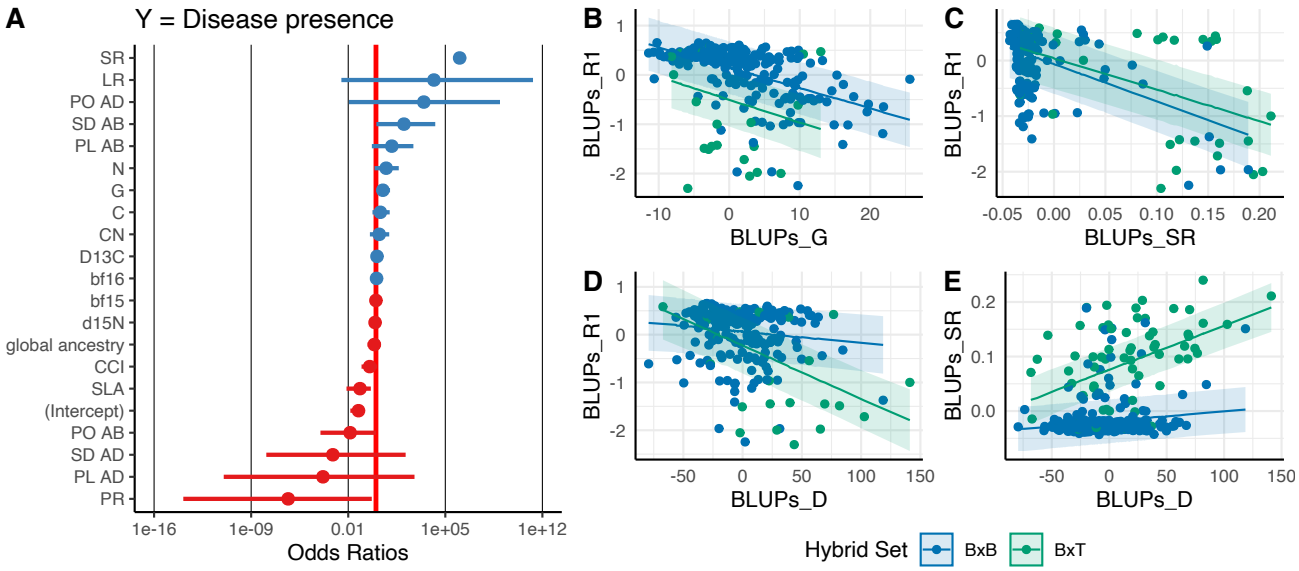


Figure 2: Forest plot of standardized regression coefficients from a logistic model with disease presence as the response. Blue and red points are positive and negative odds ratios, respectively. The red line is an odds ratio of 1. See Table S2 for logistic model output (A). Trade-offs underlying disease resistance from random slope and intercept models using best linear unbiased predictors (BLUPs) as input (Table S3). Shaded areas indicate $\pm$ standard error of the slope (B-E).

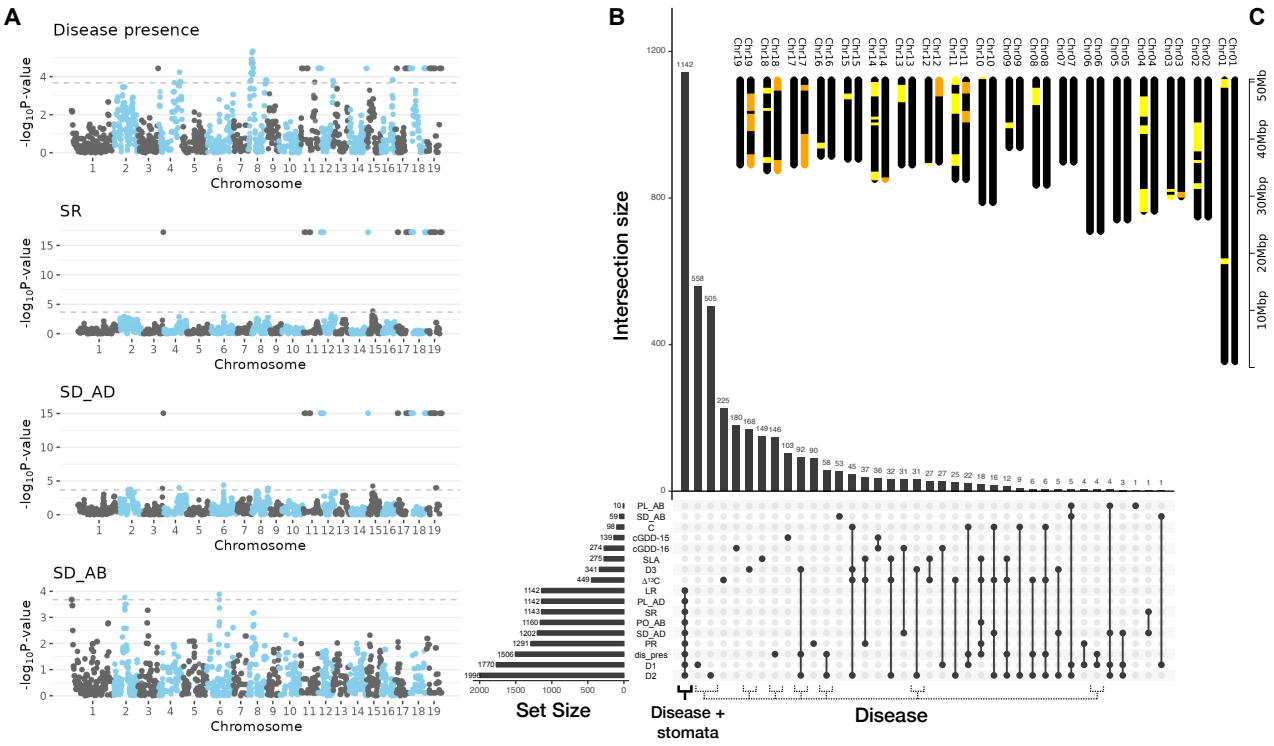


Figure 3: Results of admixture mapping genomic association tests. Manhattan plots for key traits. Cut-off value was $\alpha$=0.05/237.5 (grey dotted line) (A). UpSetR plots identified a shared, and potentially pleiotropic genomic basis for disease and stomatal trait variation ("Disease + Stomata"), and genes independent of stomata or ecophysiology traits ("Disease"). Overlap of gene lists from signficant SNPs of each trait were found. The set size refers to the number of genes containing significant SNPs (B). Chromosomal locations of windows containing candidate genes for the disease and stomatal set (orange), and disease-only set (yellow). A chromosome was plotted for each category of candidate genes and do not represent a diploid map (C).


Figure 4: Results of RAiSD selection analysis ($\mu$ statistic) organized by chromosome. Each line color indicates a different set of individuals: hybrids (magenta), unadmixed *P. balsamifera* (blue), and unadmixed *P. trichocarpa* (green). The overlap of BMIX candidate genes and the top 1% of $\mu$ statistic outliers are indicated by the solid or dashed boxes for the disease plus stomata gene set, and disease-only gene set, respectively. The density distribution of $\mu$-statistics is presented by set, and the vertical colored ticks on x-axis indicate the top 1% of values for each set.

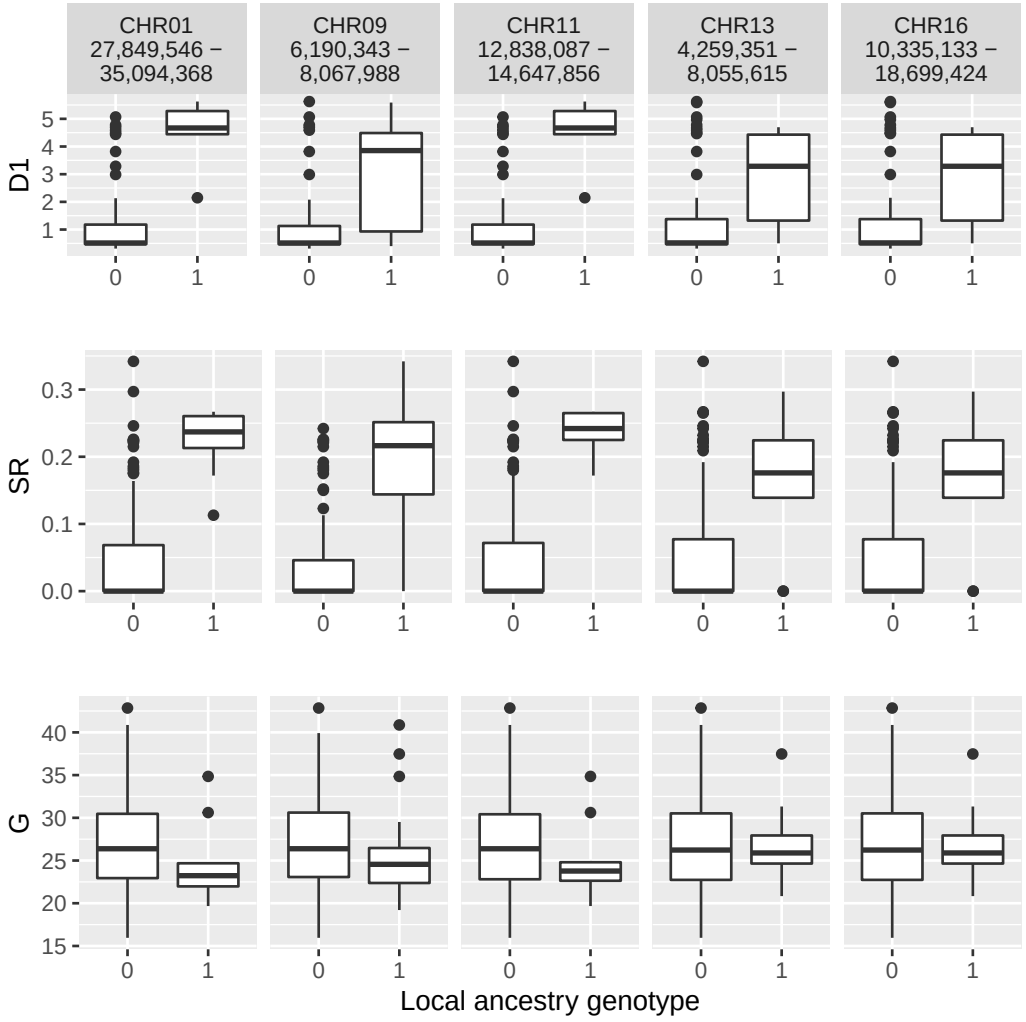


Figure 5: Allelic effects of polymorphic sites identified by admixture mapping and within the top 1% of selection scans. A shift in local ancestry genotype from homozygous for the *P. balsamifera* allele (0) to heterozygous (1) changes disease severity (D1), stomatal ratio (SR), and relative growth rate (G).

# References


1. Alachiotis, N. & Pavlidis, P. RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. *Communications biology* **1,** 79 (2018).
2. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome research* **19,** 1655–1664 (2009).
3. Bailey, J. K., Deckert, R., Schweitzer, J. A., Rehill, B. J., Lindroth, R. L., Gehring, C. & Whitham, T. G. Host plant genetics affect hidden ecological players: links among *Populus*, condensed tannins, and fungal endophyte infection. *Canadian Journal of Botany* **83,** 356–361 (2005).
4. Bailey, J. K., Schweitzer, J. A., Ubeda, F., Koricheva, J., LeRoy, C. J., Madritch, M. D., Rehill, B. J., Bangert, R. K., Fischer, D. G., Allan, G. J., *et al.* From genes to ecosystems: a synthesis of the effects of plant genetic factors across levels of organization. *Philosophical Transactions of the Royal Society B: Biological Sciences* **364,** 1607–1616 (2009).
5. Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E. & Huala, E. The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *genesis* **53,** 474–485 (2015).
6. Bomblies, K. & Peichel, C. L. Genetics of adaptation. *Proceedings of the National Academy of Sciences* **119,** e2122152119 (2022).
7. Breen, A. L. Balsam poplar (*Populus balsamifera* L.) communities on the Arctic Slope of Alaska. *Phytocoenologia* **44,** 1–24 (2014).
8. Browning, B. L., Zhou, Y. & Browning, S. R. A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics* **103,** 338–348 (2018).
9. Bürkner, P.-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80,** 1–28 (2017).
10. Chase, M. A., Stankowski, S. & Streisfeld, M. A. Genomewide variation provides insight into evolutionary relationships in a monkeyflower species complex (*Mimulus* sect. *Diplacus*). *American journal of botany* **104,** 1510–1521 (2017).
11. Chhatre, V. E., Evans, L. M., DiFazio, S. P. & Keller, S. R. Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of *Populus*. *Molecular ecology* (2018).
12. Christe, C., Stölting, K. N., Bresadola, L., Fussi, B., Heinze, B., Wegmann, D. & Lexer, C. Selection against recombinant hybrids maintains reproductive isolation in hybridizing *Populus* species despite F1 fertility and recurrent gene flow. *Molecular Ecology* **25,** 2482–2498 (2016).
13. Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33,** 2938–2940 (2017).
14. Cope, O. L., Keefover-Ring, K., Kruger, E. L. & Lindroth, R. L. Growth–defense trade-offs shape population genetic composition in an iconic forest tree species. *Proceedings of the National Academy of Sciences* **118,** e2103162118 (2021).
15. Dowkiw, A. & Bastien, C. Characterization of two major genetic factors controlling quantitative resistance to *Melampsora larici-populina* leaf rust in hybrid poplars: strain specificity, field expression, combined effects, and relationship with a defeated qualitative resistance gene. *Phytopathology* **94,** 1358–1367 (2004).
16. Evans, L. M., Slavov, G. T., Rodgers-Melnick, E., Martin, J., Ranjan, P., Muchero, W., Brunner, A. M., Schackwitz, W., Gunter, L., Chen, J.-G., *et al.* Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nature genetics* **46,** 1089 (2014).

17. Feau, N., Joly, D. L. & Hamelin, R. C. Poplar leaf rusts: model pathogens for a model tree. *Botany* **85,** 1127–1135 (2007).

18. Fetter, K. C., Nelson, D. M. & Keller, S. R. Growth-defense trade-offs masked in unadmixed populations are revealed by hybridization. *Evolution* **75,** 1450–1465 (2021).

19. Floate, K. D., Godbout, J., Lau, M. K., Isabel, N. & Whitham, T. G. Plant–herbivore interactions in a trispecific hybrid swarm of *Populus*: assessing support for hypotheses of hybrid bridges, evolutionary novelty and genetic similarity. *New Phytologist* **209,** 832–844 (2016).

20. Gelman, A. Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine* **27,** 2865–2873 (2008).

21. Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q. & Buckler, E. S. TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS one* **9,** e90346 (2014).

22. Gonzales-Vigil, E., Hefer, C. A., von Loessl, M. E., La Mantia, J. & Mansfield, S. D. Exploiting natural variation to uncover an alkene biosynthetic enzyme in poplar. *The Plant Cell* **29,** 2000–2015 (2017).

23. Jones, J. D. & Dangl, J. L. The plant immune system. *nature* **444,** 323 (2006).

24. Kardos, M. & Luikart, G. The genetic architecture of fitness drives population viability during rapid environmental change. *The American Naturalist* **197,** 511–525 (2021).

25. Keller, S. R., Olson, M. S., Silim, S., Schroeder, W. & Tiffin, P. Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera*. *Molecular Ecology* **19,** 1212–1226 (2010).

26. La Mantia, J., Klápště, J., El-Kassaby, Y. A., Azam, S., Guy, R. D., Douglas, C. J., Mansfield, S. D. & Hamelin, R. Association analysis identifies Melampsora× columbiana poplar leaf rust resistance SNPs. *PLoS One* **8,** e78423 (2013).

27. Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T. & Carey, V. J. Software for computing and annotating genomic ranges. *PLoS computational biology* **9,** e1003118 (2013).

28. Lecourieux, D., Ranjeva, R. & Pugin, A. Calcium in plant defence-signalling pathways. *New Phytologist* **171,** 249–269 (2006).

29. Little Elbert L., J. *Atlas of United States trees. Volume 1. Conifers and important hardwoods* **1146,** 320 (U.S. Department of Agriculture, Forest Service., Washington, D.C., 1971).

30. McKown, A. D., Klápště, J., Guy, R. D., Corea, O. R., Fritsche, S., Ehlting, J., El-Kassaby, Y. A. & Mansfield, S. D. A role for SPEECHLESS in the integration of leaf stomatal patterning with the growth vs disease trade-off in poplar. *The New phytologist* (2019).

31. Melotto, M., Underwood, W. & He, S. Y. Role of stomata in plant innate immunity and foliar bacterial diseases. *Annu. Rev. Phytopathol.* **46,** 101–122 (2008).

32. Muir, C. D. Making pore choices: repeated regime shifts in stomatal ratio. *Proceedings of the Royal Society B: Biological Sciences* **282,** 20151498 (2015).

33. Muir, C. D. A stomatal model of anatomical tradeoffs between gas exchange and pathogen colonization. *Frontiers in plant science* **11,** 518991 (2020).

34. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2021).

35. Rieseberg, L. H. & Carney, S. E. Plant hybridization. *The New Phytologist* **140,** 599–624 (1998).

36. Rieseberg, L. H., Kim, S.-C., Randell, R. A., Whitney, K. D., Gross, B. L., Lexer, C. & Clay, K. Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica* **129,** 149–165 (2007).

37. Scheet, P. & Stephens, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *The American Journal of Human Genetics* **78,** 629–644 (2006).

38. Shriner, D., Adeyemo, A. & Rotimi, C. N. Joint ancestry and association testing in admixed individuals. *PLoS computational biology* **7,** e1002325 (2011).

39. Sjödin, A., Street, N. R., Sandberg, G., Gustafsson, P. & Jansson, S. The *Populus* Genome Integrative Explorer (PopGenIE): a new resource for exploring the *Populus* genome. *New phytologist* **182,** 1013–1025 (2009).

40. Smith, M. W. & O'Brien, S. J. Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. *Nature Reviews Genetics* **6,** 623 (2005).

41. Suarez-Gonzalez, A., Hefer, C. A., Christe, C., Corea, O., Lexer, C., Cronk, Q. C. & Douglas, C. J. Genomic and functional approaches reveal a case of adaptive introgression from *Populus balsamifera* (balsam poplar) in *P. trichocarpa* (black cottonwood). *Molecular ecology* **25,** 2427–2442 (2016).

42. Suarez-Gonzalez, A., Lexer, C. & Cronk, Q. C. Adaptive introgression: a plant perspective. *Biology letters* **14,** 20170688 (2018).

43. Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y.-C., Sjödin, A., Van de Peer, Y., Jansson, S., Hvidsten, T. R. & Street, N. R. The Plant Genome Integrative Explorer Resource: PlantGen IE. org. *New Phytologist* **208,** 1149–1156 (2015).

44. Thompson, S. L., Lamothe, M., Meirmans, P. G., Perinet, P. & Isabel, N. Repeated unidirectional introgression towards Populus balsamifera in contact zones of exotic and native poplars. *Molecular Ecology* **19,** 132–145 (2010).

45. Tuskan, G. A. *et al.* The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *science* **313,** 1596–1604 (2006).

46. Wegmann, D., Kessner, D. E., Veeramah, K. R., Mathias, R. A., Nicolae, D. L., Yanek, L. R., Sun, Y. V., Torgerson, D. G., Rafaels, N., Mosley, T., *et al.* Recombination rates in admixed individuals identified by ancestry-based inference. *Nature genetics* **43,** 847 (2011).

47. Whitham, T. G. Plant hybrid zones as sinks for pests. *Science,* 1490–1493 (1989).

48. Whitham, T. G., Floate, K. D., Martinsen, G. D., Driebe, E. M. & Keim, P. Ecological and evolutionary implications of hybridization: *Populus*–herbivore interactions. *Biology of Populus and its implications for management and conservation,* 247–275 (1996).

49. Wymore, A. S., Keeley, A. T., Yturralde, K. M., Schroer, M. L., Propper, C. R. & Whitham, T. G. Genes to ecosystems: exploring the frontiers of ecology with one of the smallest biological units. *New Phytologist* **191,** 19–36 (2011).

# Supporting Information

Table S1: Population locality and sample size summaries for individuals used in admixture mapping. Unadmixed *Populus balsamifera* (ref-*Pb.*, N = 25) and *P. trichocarpa* (ref-*Pt.*, N = 25) were selected to generate locus-specific ancestries for admixture mapping in the test set (N = 117).

| Pop. | Set | N | Lat. | Lon. |
|---|---|---|---|---|
| BESC | ref-*Pt.* | 15 | 47.81 | -122.01 |
| GW | ref-*Pt.* | 9 | 47.30 | -122.58 |
| Nisqually | ref-*Pt.* | 1 | 47.03 | -122.67 |
| DCK | ref-*Pb.* | 20 | 51.60 | -101.73 |
| CLK | test | 6 | 54.22 | -110.08 |
| HBY | ref-*Pb.*/test | 23 | 52.89 | -102.39 |
| CYH | test | 4 | 49.64 | -109.98 |
| FNO | test | 8 | 58.50 | -122.37 |
| JKH | test | 6 | 43.83 | -110.47 |
| MMT | test | 6 | 49.88 | -102.59 |
| MSG | test | 4 | 44.31 | -106.90 |
| OFR | test | 14 | 53.14 | -101.10 |
| OUT | test | 4 | 51.14 | -106.20 |
| SKN | test | 4 | 52.33 | -106.27 |
| SSR | test | 12 | 44.46 | -109.61 |
| TUR | test | 7 | 53.20 | -108.32 |
| USDA12 | test | 6 | 46.58 | -88.030 |
| USDA13 | test | 2 | 46.08 | -88.030 |
| USDA14 | test | 1 | 45.58 | -88.030 |
| USDA15 | test | 1 | 46.42 | -86.870 |
| USDA18 | test | 1 | 45.42 | -84.500 |
| USDA3 | test | 1 | 47.17 | -91.670 |
| USDA7 | test | 3 | 48.42 | -92.980 |
| USDA8 | test | 1 | 48.33 | -94.520 |
| USDA9 | test | 4 | 47.92 | -94.520 |
| WLK | test | 4 | 60.05 | -128.44 |
| Total | | 167 | | |

Table S2: Output of logistic model of disease presence and traits.

| term | estimate | std.error | p.value | p.stars |
|---|---|---|---|---|
| (Intercept) | 5.48e-02 | 0.57 | 0.000 | *** |
| SR | 1.07e+06 | 7.81 | 0.076 | |
| LR | 1.51e+04 | 7.70 | 0.212 | |
| PR | 4.66e-07 | 7.66 | 0.057 | |
| G | 3.14e+00 | 0.55 | 0.038 | * |
| SD_AB | 1.04e+02 | 2.40 | 0.054 | |
| SD_AD | 7.74e-04 | 5.42 | 0.187 | |
| PL_AB | 1.38e+01 | 1.69 | 0.121 | |
| PL_AD | 1.5e-04 | 7.65 | 0.250 | |
| PO_AD | 2.81e+03 | 5.96 | 0.183 | |
| PO_AB | 1.36e-02 | 2.28 | 0.061 | |
| D13C | 1.18e+00 | 0.45 | 0.711 | |
| d15N | 8.6e-01 | 0.38 | 0.698 | |
| CN | 1.71e+00 | 0.78 | 0.491 | |
| C | 2.06e+00 | 0.71 | 0.309 | |
| SLA | 7.05e-02 | 1.02 | 0.009 | ** |
| CCI | 3.51e-01 | 0.64 | 0.104 | |
| N | 5.42e+00 | 0.99 | 0.090 | |
| cGDD-15 | 9.92e-01 | 0.46 | 0.987 | |
| cGDD-16 | 1.1e+00 | 0.43 | 0.824 | |
| global_ancestry | 7.61e-01 | 0.55 | 0.622 | |

Table S3: Model output from random intercept and slope models. Models were fit in brms with BLUPs as input data. See Fig. 2 for plots of each model. Abbreviations: R1 = disease resistance (2015); G = relative growth rate; SR = stomatal ratio; D = log total stomatal density. Ancestry was either BxB or BxT.

R1∼ (1 + G | ancestry)

| Intercept | | | | |
|---|---|---|---|---|
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | 0.2067035 | 0.8917482 | -1.683928 | 2.118894 |
| BxT | -0.4433149 | 0.8964821 | -2.345828 | 1.458727 |
| Slope | | | | |
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | -0.04135374 | 0.005367762 | -0.05190709 | -0.0307486 |
| BxT | -0.04623021 | 0.017100377 | -0.07899603 | -0.0118752 |

R1∼ (1 + SR | ancestry)

| Intercept | | | | |
|---|---|---|---|---|
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | -0.2606446 | 0.8399156 | -2.184307 | 1.321452 |
| BxT | -0.1555295 | 0.8647422 | -2.212310 | 1.475180 |
| Slope | | | | |
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | -6.695734 | 1.109421 | -8.832040 | -4.505172 |
| BxT | -5.643501 | 1.103754 | -7.966181 | -3.590200 |

R1∼ (1 + D | ancestry)

| Intercept | | | | |
|---|---|---|---|---|
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | 0.03483479 | 0.6257267 | -1.490817 | 1.312665 |
| BxT | -0.25450876 | 0.6343904 | -1.808712 | 1.010025 |
| Slope | | | | |
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | -0.002386611 | 0.001398987 | -0.005101953 | 0.0003786823 |
| BxT | -0.011194649 | 0.002285893 | -0.015724202 | -0.0066023706 |

SR∼ (1 + D | ancestry)

| Intercept | | | | |
|---|---|---|---|---|
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | 0.4296254 | 0.9106484 | -0.7843780 | 1.948828 |
| BxT | 0.5240732 | 0.9123055 | -0.6918405 | 2.049142 |
| Slope | | | | |
| | Estimate | Est.Error | Q2.5 | Q97.5 |
| BxB | 0.0001817294 | 8.283463e-05 | 1.031871e-06 | 0.0003163581 |
| BxT | 0.0008086642 | 1.521083e-04 | 5.980606e-04 | 0.0010819346 |

Table S4: Summary of populations used in the RAiSD selection scan analysis.

| Pop. | Set | N | Lat. | Lon. |
|---|---|---|---|---|
| CLK | *P. balsamifera* | 12 | 54.22 | -110.08 |
| DCK | *P. balsamifera* | 12 | 51.60 | -101.73 |
| HBY | *P. balsamifera* | 6 | 52.90 | -102.39 |
| MMT | *P. balsamifera* | 13 | 49.88 | -102.59 |
| OFR | *P. balsamifera* | 12 | 53.14 | -101.10 |
| OUT | *P. balsamifera* | 5 | 51.15 | -106.26 |
| SKN | *P. balsamifera* | 10 | 52.35 | -106.64 |
| TUR | *P. balsamifera* | 10 | 53.20 | -108.36 |
| CYH | Hybrid | 4 | 49.64 | -109.98 |
| FNO | Hybrid | 8 | 58.51 | -122.38 |
| HBY | Hybrid | 1 | 52.93 | -102.39 |
| JKH | Hybrid | 6 | 43.84 | -110.47 |
| MSG | Hybrid | 4 | 44.32 | -106.91 |
| SKN | Hybrid | 1 | 52.22 | -106.27 |
| SSR | Hybrid | 7 | 44.46 | -109.60 |
| USDA12 | Hybrid | 3 | 46.58 | -88.03 |
| USDA18 | Hybrid | 1 | 45.42 | -84.50 |
| WLK | Hybrid | 4 | 60.05 | -128.44 |
| 13 | *P. trichocarpa* | 2 | 54.15 | -128.60 |
| 31 | *P. trichocarpa* | 8 | 49.71 | -125.06 |
| Nisqually | *P. trichocarpa* | 9 | 47.10 | -122.64 |
| Nooksack | *P. trichocarpa* | 8 | 48.80 | -122.17 |
| Olympic Penninsula | *P. trichocarpa* | 3 | 47.59 | -122.90 |
| Puyallup | *P. trichocarpa* | 5 | 47.09 | -122.20 |
| Skagit | *P. trichocarpa* | 5 | 48.51 | -122.07 |
| Skykomish | *P. trichocarpa* | 6 | 47.82 | -121.86 |
| Total | | 165 | | |

Table S5: Descriptions of candidate genes. *Arabidopsis* orthologs were identified by the best BLAST hit from www.popgenie.org or through BLAST results from The Arabidopsis Information Resource (www.arabidopsis.org). See text for details regarding candidate gene selection.

| Potri ID | *At*-ortholog | Pfam description |
|---|---|---|
| Guard cell function | | |
| 001G317800 | AT1G15690 | NA |
| 001G318800 | AT4G30990 | Down-regulated in metastasis |
| 011G112700 | AT3G14420 | oxidoreductase activity |
| 013G057500 | AT5G56540 | NA |
| 016G115500 | AT3G51860 | transmembrane transport |
| 016G117200 | AT3G51850 | Protein tyrosine kinase, protein phosphorylation EF hand |
| Immune system & detoxicants | | |
| 011G113000 | AT1G78380 | protein binding, Glutathione S-transferase, C-terminal domain |
| 011G116200 | AT4G27540 | PRA1 family protein |
| 011G116900 | AT5G53890 | Leucine rich repeat N-terminal domain |
| 011G117100 | AT3G23560 | MatE (multi antimicrobial extrusion protein) |
| 011G117200 | AT3G23550 | MatE (multi antimicrobial extrusion protein) |
| 011G117300 | AT3G23560 | MatE (multi antimicrobial extrusion protein) |
| 011G117400 | AT3G23550 | MatE (multi antimicrobial extrusion protein) |
| 011G118200 | AT1G28280 | VQ motif |
| 011G118900 | AT3G15353 | NA |
| 011G121200 | AT4G12560 | F-box associated, protein binding |
| 013G058500 | AT3G03960 | cellular protein metabolic process |
| 016G118100 | AT3G51830 | SacI homology domain |
| Lipid biosyntheis & transport | | |
| 001G316600 | AT3G05180 | NA |
| 001G317400 | AT4G22330 | ceramide metabolic process |
| 016G113800 | AT1G74720 | protein binding |
| 016G115800 | AT2G38180 | lipid metabolic process |
| 016G116400 | AT5G01410 | pyridoxal phosphate biosynthetic process |
| 016G118000 | AT3G51840 | acyl-CoA dehydrogenase activity |
| Growth related | | |
| 011G121300 | AT5G17260 | no apical meristem (NAM) protein |
| 011G115400 | AT5G53950 | regulation of transcription, DNA-dependent |
| 016G113600 | AT2G38120 | transmembrane amino acid transporter protein |
| 016G114600 | AT5G01270 | double-stranded RNA binding |
| 016G118400 | AT4G33270 | WD domain, G-beta |
| Cell wall related | | |
| 013G056800 | AT5G19780 | Tubulin/FtsZ family, GTPase domain |
| 016G114300 | AT2G20340 | carboxylic acid metabolic process |
| Abiotic/biotic stress responsive | | |
| 001G316900 | AT4G04980 | NA |
| 001G317000 | AT4G04980 | NA |
| 001G317300 | AT5G49210 | NA |
| 001G318900 | NA | NA |
| 011G115200 | AT3G30390 | NA |
| 013G056900 | AT5G18100 | superoxide metabolic process |
| 013G057700 | AT3G03890 | FMN binding |
| Epigenetics & DNA replication | | |
| 001G316200 | AT3G01320 | nucleus, Histone deacetylase (HDAC) interacting |
| 001G316300 | AT3G01320 | nucleus |
| 001G316500 | AT1g04840 | PPR repeat |
| 001G317500 | AT4G13780 | tRNA binding, aminoacyl-tRNA ligase activity |
| 001G317700 | AT2G31740 | methyltransferase activity |
| 009G085400 | AT1G44910 | protein binding, FF domain |
| 011G114100 | AT1G17160 | pfkB family carbohydrate kinase |
| 011G116000 | AT4G13870 | nucleobase, nucleoside, nucleotide and nucleic acid metabolic process |

Continued on next page

**Table S5 – continued from previous page**

| Potri ID | At-ortholog | Pfam description |
|---|---|---|
| 011G116100 | AT5G53920 | protein methyltransferase activity |
| 013G056700 | AT5G19790 | apetela 2 domain (AP2 domain) transcription factor |
| 013G057000 | AT5G18110 | translation initiation factor activity |
| 013G058800 | AT1G54390 | protein binding |
| 013G058900 | AT4G13650 | PPR repeat |
| 016G116900 | AT5G05610 | PHD-finger |
| 016G117300 | At5g01380 | Myb/SANT-like DNA-binding domain |
| Ubiquination | | |
| 001G316400 | AT1G04850 | protein binding, PUB domain |
| 011G112800 | AT3G14400 | ubiquitin thiolesterase activity |
| 016G115300 | AT5G01520 | zinc finger, C3HC4 type (RING finger) |
| Flower related | | |
| 001G316800 | AT1G04910 | GDP-fucose protein O-fucosyltransferase |
| 011G112500 | AT1G31660 | bystin |
| 011G115000 | AT5G57850 | catalytic activity |
| 016G116300 | AT5G01450 | NA |
| 016G117400 | AT5G01370 | NA |
| Membrane transporters | | |
| 001G316700 | NA | Rab GTPase activator activity |
| 001G317100 | AT4G13750 | NA |
| 001G317200 | AT4G13750 | NA |
| 001G318700 | AT1G71900 | NA |
| 016G115400 | AT5G01500 | Mitochondrial carrier protein |
| Secondary metabolism related | | |
| 001G317600 | AT1G04920 | sucrose metabolic process |
| 016G115600 | AT2G25300 | NA |
| Signal transduction | | |
| 011G114200 | NA | protein transport, Plug domain of Sec61p |
| 016G114800 | AT3G09010 | protein phosphorylation |
| 016G119300 | AT2G38280 | purine ribonucleoside monophosphate biosynthetic process |
| Unknown genes, functions, or enigmatic | | |
| 001G317900 | NA | NA |
| 001G318000 | NA | NA |
| 009G085500 | AT2G20240 | NA |
| 011G112600 | NA | NA |
| 011G115100 | AT5G53970 | transferase activity, transferring nitrogenous groups |
| 011G115300 | NA | Plant mobile domain |
| 011G115900 | NA | NA |
| 011G117000 | NA | NA |
| 011G118000 | AT5G53860 | NA |
| 011G118100 | NA | metal ion binding |
| 013G057100 | AT3G03860 | cell redox homeostasis |
| 013G057200 | AT5G18130 | NA |
| 013G058600 | AT1G64770 | NA |
| 016G113900 | NA | NA |
| 016G114000 | NA | NA |
| 016G114100 | NA | NA |
| 016G114700 | AT2G40060 | clathrin coat of trans-Golgi network vesicle |
| 016G115700 | NA | NA |
| 016G116200 | AT5G01460 | NA |
| 016G116800 | NA | NA |
| 016G117900 | NA | NA |

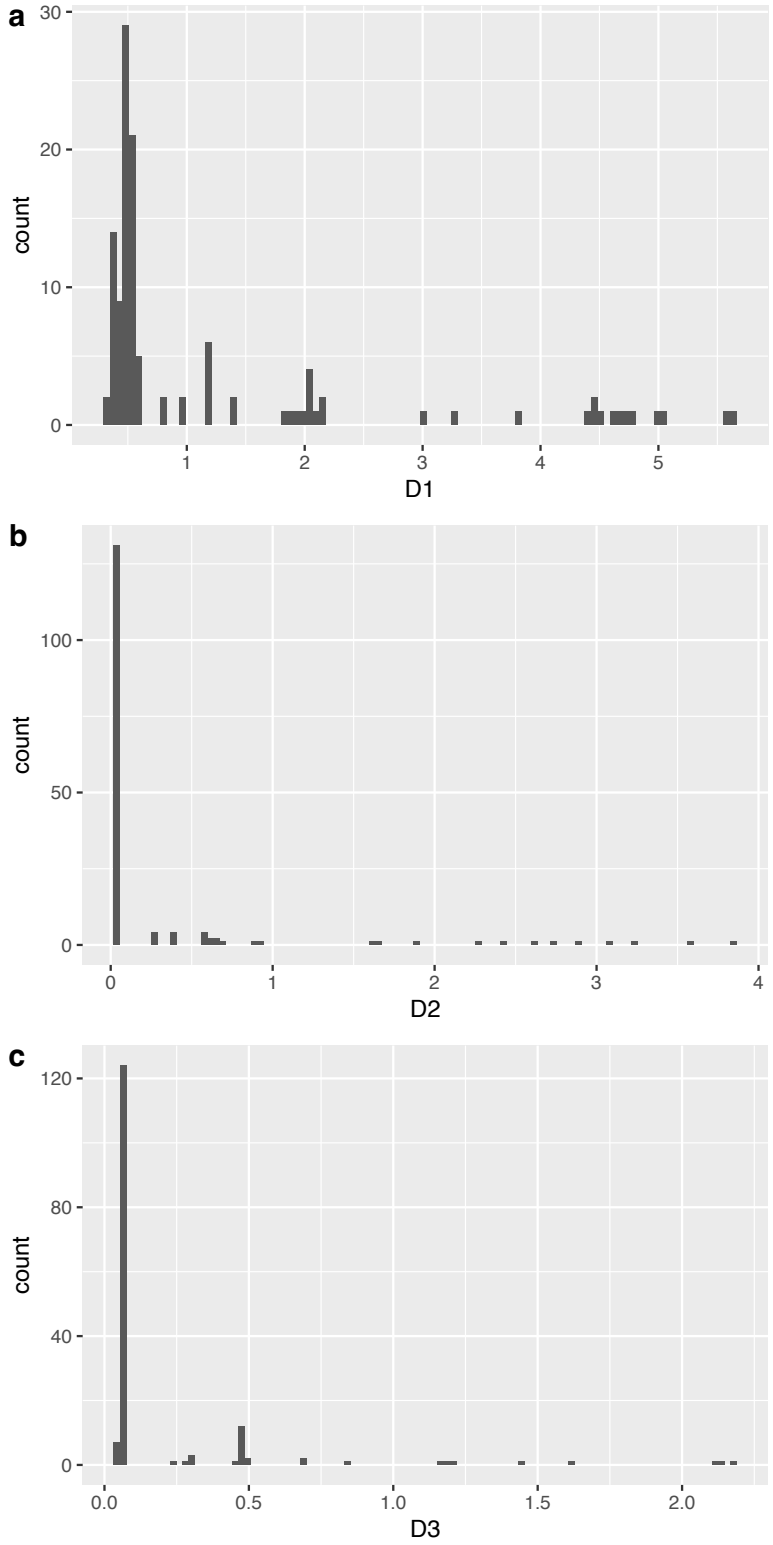


Figure S1: Disease severity BLUPs. The binary disease presence/absence response was converted from D2 (panel b).

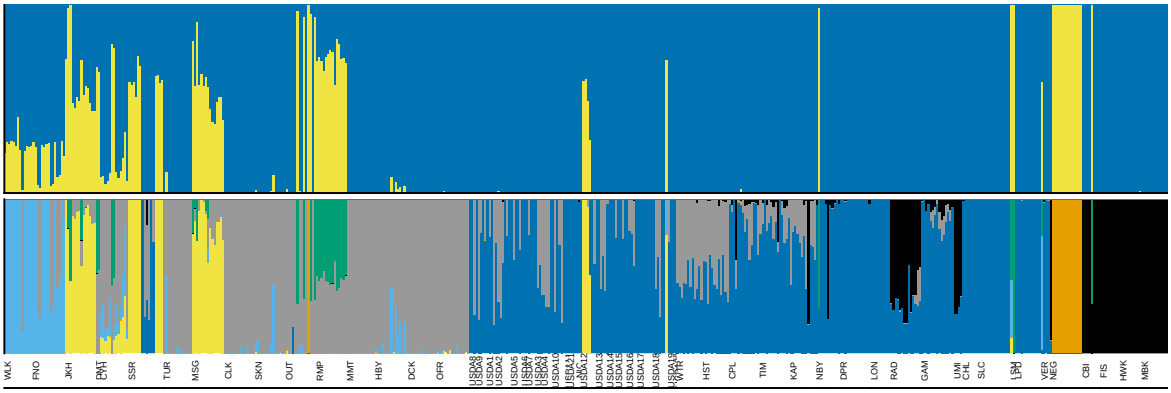


Figure S2: Global ancestry estimated from ADMIXTURE (Alexander *et al.,* 2009) at K = 2 (top) and K = 7 (bottom).

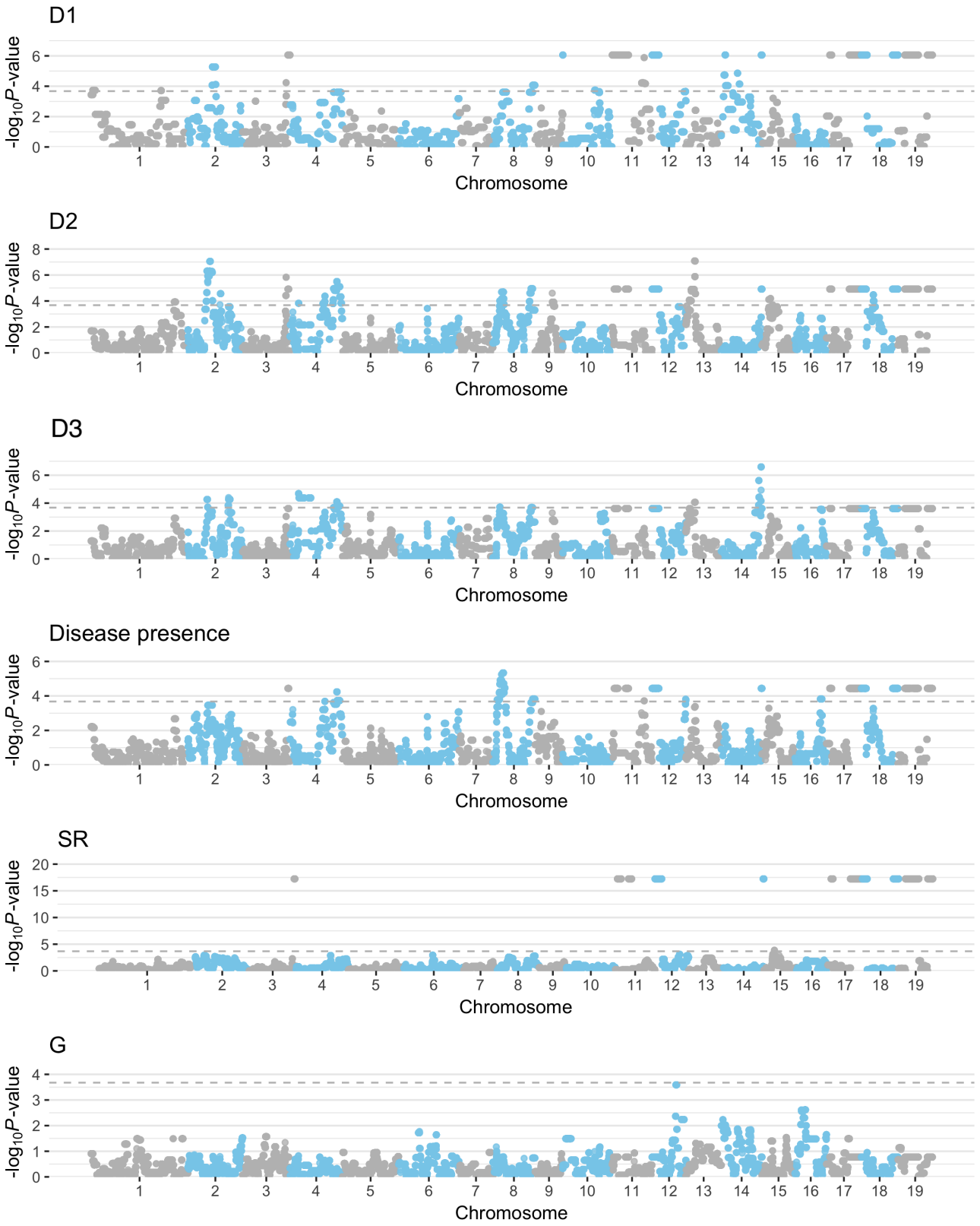


Figure S3: Individual manhattan plots of p-values from BMIX tests. See Table 1 for trait abbreviation definitions.

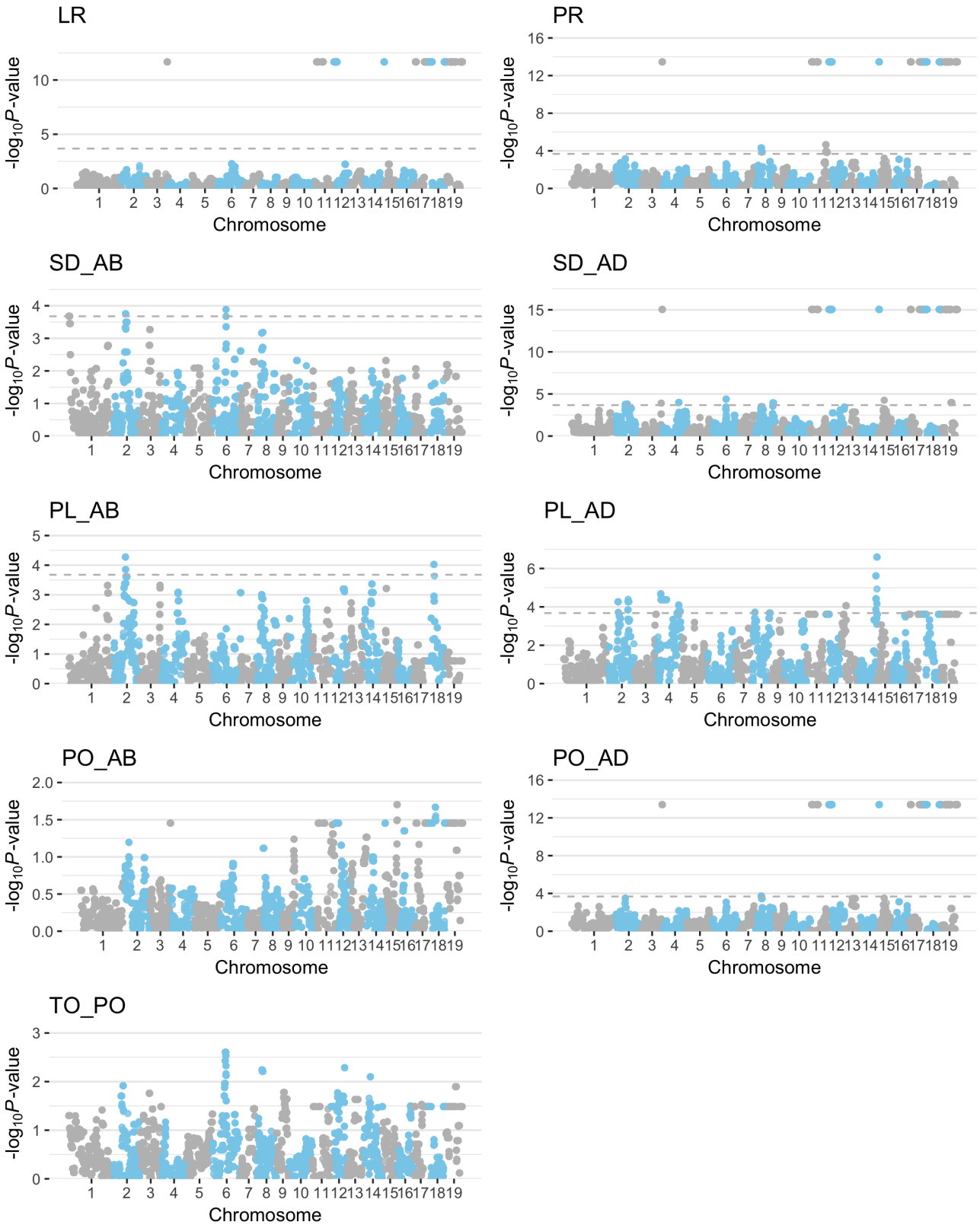


Figure S4: Individual manhattan plots of p-values from BMIX tests. See Table 1 for trait abbreviation definitions.

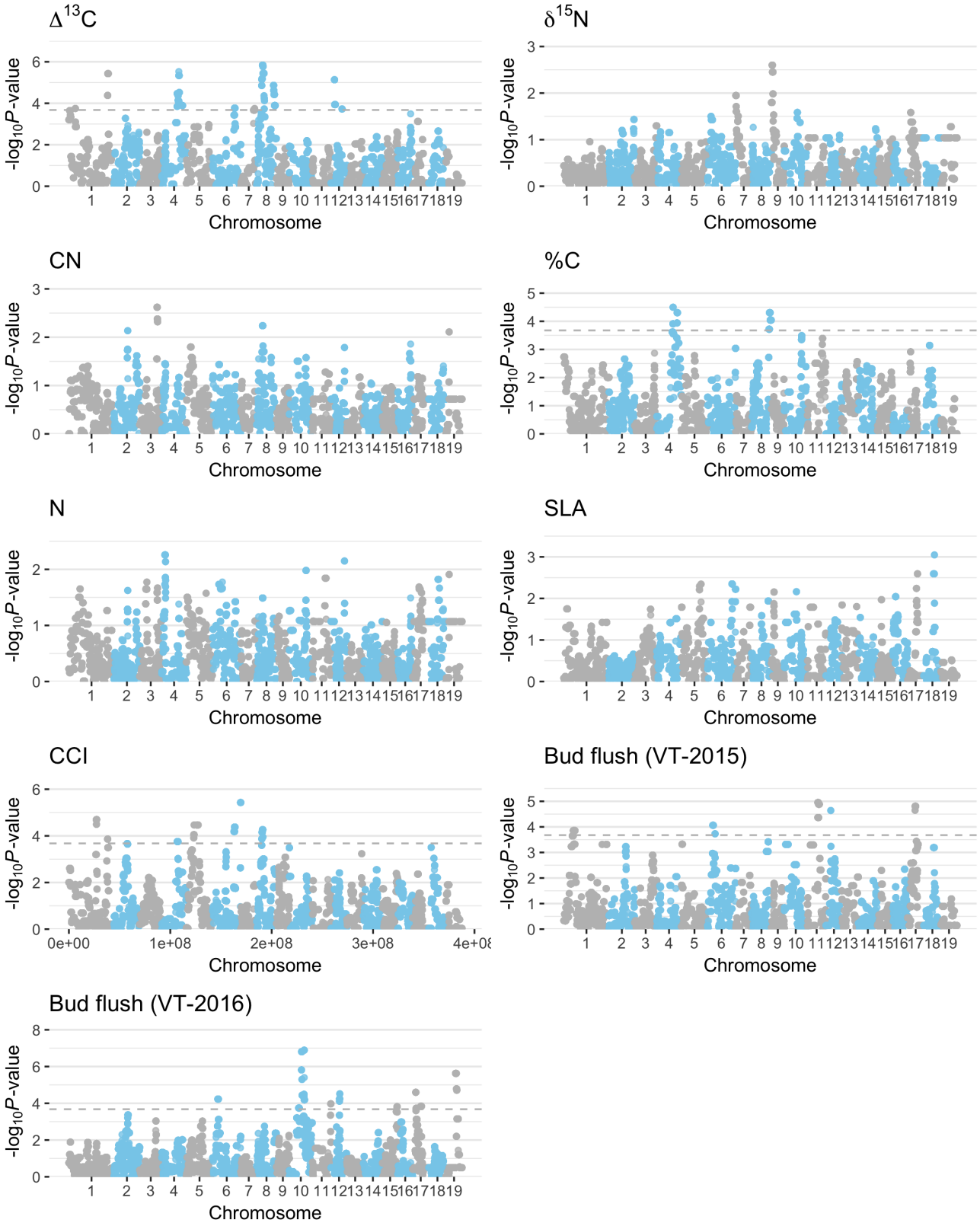


Figure S5: Individual manhattan plots of p-values from BMIX tests. See Table 1 for trait abbreviation definitions.