

UC Berkeley

UC Berkeley Previously Published Works

Title

Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees

Permalink

<https://escholarship.org/uc/item/1vn021p1>

Journal

Proceedings of the National Academy of Sciences of the United States of America, 118(17)

ISSN

0027-8424

Authors

Kling, Matthew M
Ackerly, David D

Publication Date

2021-04-27

DOI

10.1073/pnas.2017317118

Peer reviewed



Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees

Matthew M. Kling^{a,1} and David D. Ackerly^{a,b}

^aDepartment of Integrative Biology, University of California, Berkeley, CA 94720; and ^bDepartment of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720

Edited by Nils Chr. Stenseth, University of Oslo, Oslo, Norway, and approved March 17, 2021 (received for review August 14, 2020)

Wind disperses the pollen and seeds of many plants, but little is known about whether and how it shapes large-scale landscape genetic patterns. We address this question by a synthesis and reanalysis of genetic data from more than 1,900 populations of 97 tree and shrub species around the world, using a newly developed framework for modeling long-term landscape connectivity by wind currents. We show that wind shapes three independent aspects of landscape genetics in plants with wind pollination or seed dispersal: populations linked by stronger winds are more genetically similar, populations linked by directionally imbalanced winds exhibit asymmetric gene flow ratios, and downwind populations have higher genetic diversity. For each of these distinct hypotheses, partial correlations between the respective wind and genetic metrics (controlling for distance and climate) are positive for a significant majority of wind-dispersed or wind-pollinated genetic data sets and increase significantly across functional groups expected to be increasingly influenced by wind. Together, these results indicate that the geography of both wind strength and wind direction play important roles in shaping large-scale genetic patterns across the world's forests. These findings have implications for various aspects of basic plant ecology and evolution, as well as the response of biodiversity to future global change.

wind dispersal | wind pollination | landscape genetics | gene flow | genetic differentiation

Wind is a driving force in plant ecology and evolution. Outside the tropics, most forests, shrublands, and grasslands are dominated by plants whose seeds, pollen, or spores are transported by wind. In US forests, for example, wind-pollinated or wind-dispersed trees represent an estimated 95% of forest basal area and 75% of tree species (*SI Appendix, Appendix 1*), whereas across tropical forests ~3 to 30% of tree species are wind-dispersed and 1 to 10% are wind-pollinated (1–3). Wind-dependent plants include virtually all conifers, grasses, ferns, and mosses, as well as a smaller minority of flowering plants.

Strong geographic trends in wind speed and direction have shaped major patterns in plant biogeography, such as colonization dynamics of oceanic islands (4, 5) and latitudinal gradients in the prevalence of wind versus animal pollination (2, 3). While the role of wind in these taxonomic and functional diversity patterns is well established, comparatively little is known about its possible role in shaping genetic diversity patterns within species ranges. Isolated case studies have hinted that wind direction may influence gene flow in individual species (6–9), but others have concluded that it has no meaningful relationship with gene flow (e.g., ref. 10 and references therein), and most studies on the subject have focused on small spatial scales within populations rather than on large biogeographic scales. It has not been shown whether wind regimes systematically shape broad landscape genetic patterns in species that are dispersed or pollinated by wind.

One reason for this knowledge gap is that long-distance wind dispersal is difficult to directly observe and is challenging to model given the chaotic variability of weather patterns. (Note that for brevity we use “wind dispersal” throughout this paper as a generic term for wind transport of seeds, pollen, and spores.)

While wind dispersal modeling has a rich history in plant ecology (11), studies of long-term dispersal potential have focused mainly on temporal windspeed variability and vertical windspeed profiles and generally ignored geographic variation in wind speed and direction (12, 13). Dispersal studies that do incorporate wind geography have generally focused on individual weather events (e.g., ref. 14) or in some cases entire individual seasons (e.g., ref. 15), rather than the long climatic timescales that shape landscape genetics and biogeography. Only recently have methods emerged to combine large-scale, spatially explicit, high-resolution wind data with landscape connectivity algorithms to more rigorously model how long-term spatiotemporal wind variability influences biogeography (4, 16, 17).

These landscape wind connectivity (“windscape”) models use the time-integrated speed of wind diffusion between origin and destination locations as estimates of relative dispersal potential, opening a range of important questions about the role of wind in biogeography and spatial ecology. Wind connectivity has a natural correspondence to landscape genetic metrics because both atmospheric circulation and propagule dispersal can be considered processes of spatial diffusion and represented as pairwise relationships among populations; however, we are unaware of any prior study using wind-based landscape connectivity models to investigate gene flow. In this study, we use windscape models parameterized with a full 30 year climatology (1980 through 2009) of hourly wind data in a global analysis exploring how wind geography shapes gene flow in trees. The major patterns of atmospheric circulation are driven by the Earth’s shape and rotation and the locations of continents, and large-scale global

Significance

It is unknown whether wind currents shape large-scale gene flow in terrestrial organisms such as plants, though analogous river and ocean currents are known to strongly influence genetic patterns in aquatic organisms. We use newly developed “windscape” connectivity models in combination with a global multispecies forest genetics data set to demonstrate that wind shapes several distinct large-scale genetic patterns in many tree species, including population differentiation, migration direction, and genetic diversity. These findings advance our understanding of the spatial ecology and evolution of wind-dispersed and wind-pollinated plants. They also suggest that the geography of wind strength and direction could potentially influence patterns of forest vulnerability to human pressures like habitat fragmentation and climate change.

Author contributions: M.M.K. and D.D.A. designed research; M.M.K. performed research and analyzed data; and M.M.K. and D.D.A. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

¹To whom correspondence may be addressed. Email: mattkling@berkeley.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2017317118/-DCSupplemental>.

Published April 19, 2021.

prevailing wind patterns are thought to have changed relatively subtly over millennia (e.g., ref. 17); recent wind climatologies are thus assumed to be reasonable albeit imperfect proxies for the longer-term historic winds that shaped current forest genetic patterns.

Landscape genetic patterns have various facets, and we pose four hypotheses about how different spatial genetic patterns are shaped by separate facets of wind connectivity:

1. The “flow” hypothesis: Rates of directional gene flow are higher between source and destination populations connected by stronger directional wind flows.
2. The “isolation” hypothesis: Population pairs linked by weaker winds are more genetically differentiated.
3. The “asymmetry” hypothesis: Population pairs linked by more directionally asymmetrical winds have more imbalanced gene flow asymmetry ratios.
4. The “diversity” hypothesis: Downwind populations have higher genetic diversity.

As an example to help introduce these hypotheses (Fig. 1), we will consider the wind-dispersed, wind-pollinated tree species *Betula pendula*, silver birch. Nuclear microsatellite data were originally sampled from populations across this species’ range in western Eurasia by Tsuda et al. (18) in a study unrelated to wind. In this example, we reanalyze those data to derive measures of pairwise genetic relationships across populations and relate these to corresponding measures of time-integrated wind connectivity estimated by newly developed windscape models (17) for each of the hypotheses as follows. We use wind data only from the species’ pollination and dispersal months, since wind patterns vary seasonally.

The flow hypothesis (Fig. 1A) predicts that gene flow across a species range will be higher along routes with higher rates of wind flow. Gene flow is the directional movement of genetic material from one population to another resulting from seed and pollen dispersal over space and time. For a given pair of populations, rates of historic gene flow in each separate direction can be estimated from allele frequency data (SI Appendix, Fig. S1) (e.g., refs. 19, 20). Like gene flow, wind flow is a rate. It is quantified here as the inverse of the estimated time it takes an air parcel to diffuse from one location to another, averaged over spatial and temporal variation in wind speed and direction. In the birch example, regional winds are spatiotemporally variable but blow most frequently and strongly toward the east-northeast; for a focal population in the center of the species range, rates of outbound wind flow are thus highest to destinations toward the northeast, while rates of inbound wind flow are highest from origins toward the southwest. These wind flows are positively correlated with estimated rates of directional gene flow in this species after controlling for distance and environment (Fig. 1A), as predicted in the flow hypothesis.

Flow patterns like these are a composite of the speed and directionality of movement and can be decomposed into independent subpatterns associated with the isolation and asymmetry hypotheses, respectively. Genetic isolation patterns are a long-standing focus in landscape genetics and include common phenomena like isolation by distance (IBD) (21) and isolation by environment (IBE) (22). Our hypothesis of isolation by wind (Fig. 1B) posits that populations linked by higher wind speeds will be more genetically similar, after controlling for distance and environment. We can calculate a directionless measure of wind connectivity for a given pair of populations by averaging the two directional wind flows and compare this to genetic differentiation measures like F_{st} to test the isolation hypothesis. In the birch example, the central focal population is more wind-isolated from the northern portion of the species range than the southern portion. And across all population pairs of this species, we indeed see

the hypothesized positive relationship between wind connectivity and genetic similarity (Fig. 1B).

Third, the asymmetry hypothesis (Fig. 1C) posits that population pairs linked by directionally asymmetric winds will exhibit corresponding rates of asymmetric gene flow. Asymmetric gene flow between populations can have important evolutionary and ecological consequences (23–26) and has become an increasing area of focus in landscape genetics with the development of methods to estimate asymmetric gene flow from static population genetic data (e.g., refs. 19, 20). For a given population pair, we can calculate gene flow asymmetry as the ratio of gene flow in one versus the other direction, and wind flow asymmetry as the ratio of wind flow in one versus the other direction. These unitless ratios are independent of the absolute amounts of wind or gene flow and represent the relative directionality of flow, so high or low ratios can occur at any geographic distance. In the birch example, wind flow asymmetry patterns for the focal population emphasize the prevailing northeastward flow of wind in this region. These correlate positively with estimated gene flow asymmetry across the full set of populations (Fig. 1C), as expected under the asymmetry hypothesis.

Finally, the diversity hypothesis (Fig. 1D) predicts that downwind populations will tend to have higher genetic diversity. Standing genetic diversity influences a population’s evolutionary potential and its conservation importance and can vary widely across a species range. Our hypothesis is based on the idea that populations with higher rates of net immigration will accumulate genetic variation more rapidly than it is lost due to selection or drift, an effect that has been observed in empirical and modeling studies in river systems (27–29). We can assess the downwind diversity hypothesis by calculating the ratio of allelic richness measures for a given population pair, and comparing this to the same measure of wind flow asymmetry as in the asymmetry hypothesis above. The diversity prediction holds true in the silver birch example, in which downwind populations indeed have higher levels of allelic richness (Fig. 1D).

With four hypotheses all based on the same input data, it is important to ask whether these results are just alternative viewpoints on a single underlying pattern. The answer is that they are largely independent, with the exception of the flow hypothesis (SI Appendix, Fig. S2). Average wind connectivity (as used in the isolation hypothesis) and wind asymmetry ratios are mathematically uncorrelated by definition, and each contain half the information in the original wind flow data used to derive them. Whereas the wind asymmetry ratio emphasizes variation in wind directionality within individual population pairs, the mean wind connectivity emphasizes variation in wind speed among multiple population pairs; wind flow represents the combined effect. The diversity hypothesis is based on the same wind asymmetry metric as the asymmetry hypothesis, but estimated genetic diversity ratios and gene flow ratios are largely uncorrelated in many cases, including the silver birch example, making the two phenomena mostly independent.

Wind-genetic relationships in any individual data set like the silver birch example can be instructive, and the results for this species are robust to a variety of alternative model specifications (SI Appendix, Table S1), but they are also subject to numerous assumptions and uncertainties that could confound our ability to measure genetic effects of wind. While it must be true, at some level, that winds shape gene flow patterns in wind-dispersed and wind-pollinated taxa, it is far from clear that these effects will be detectable using available methods. Wind dispersal dynamics and millennial metapopulation histories are far more complex than wind connectivity models and gene flow models can hope to represent, and a variety of assumptions are thus necessary on both the wind and genetic sides of the modeling equation. For instance, current wind connectivity models are based on two-dimensional near-surface wind conditions, while wind dispersal

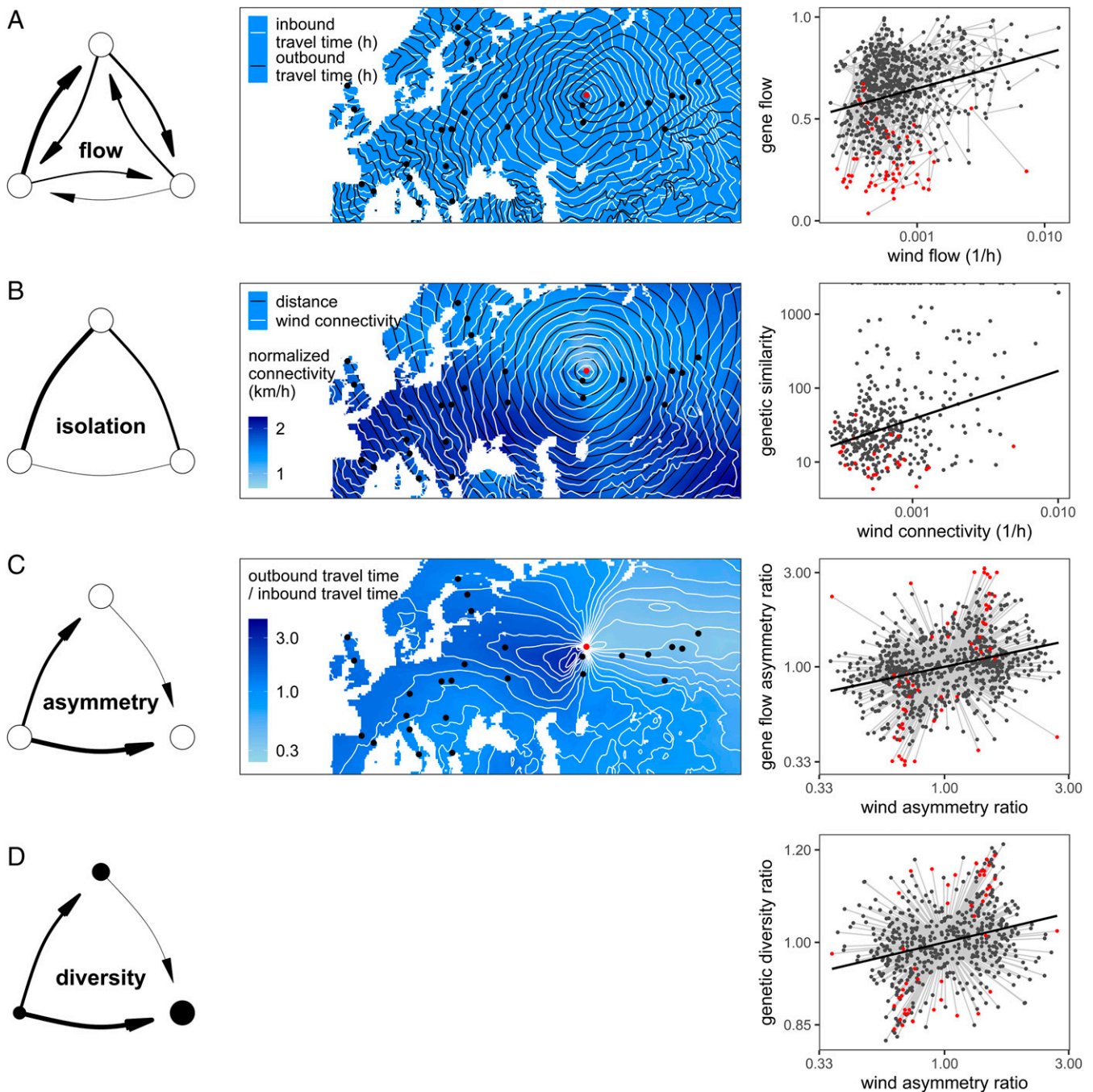


Fig. 1. Examples of the four facets of landscape genetics and wind patterns explored in this study: (A) flow, (B) isolation, (C) asymmetry, and (D) diversity. This example shows the wind-dispersed, wind-pollinated birch *Betula pendula*, one of the 120 data sets reanalyzed in this study; these genetic data were collected across the species range in western Eurasia by Tsuda et al. (18) at the populations shown in the maps for a study unrelated to wind. The schematic diagrams illustrate the four metrics. Wind connectivity landscapes are different for every reference location, and the maps here show the three wind connectivity metrics for one focal population (red point). The scatterplots show relationships between the wind and genetic metrics across the four facets, all of which are hypothesized to be positive for wind-dispersed genomes. In the scatterplots, the two directional values for each population pair are linked with line segments, and red points indicate relationships involving the red reference population in the maps; plots show first-order relationships, while the r and P values listed indicate the size and significance of partial correlations controlling for distance and climatic difference. Note that there are three wind metrics and four genetic metrics, since wind asymmetry is a predictor for both gene flow asymmetry and genetic diversity.

takes place in three dimensions (uplift and transport higher in the atmosphere can be very important); also, important long-distance dispersal events may occur under rare extreme conditions not fully captured by diffusion models that integrate over wind conditions across many decades, hours, and weather patterns (30). On the genetic side, inferring historic directional gene

flow from static snapshots of population genetic patterns can be attempted using a range of approaches (19, 20, 31), but all are subject to sampling uncertainty and make strong assumptions about evolutionary processes and metapopulation dynamics; even in contrived situations when these assumptions are met, there is substantial irreducible uncertainty in inferred gene flow

patterns. The idiosyncrasies of landscape genetic dynamics help explain why IBD and IBE, standard concepts of how dispersal and selection shape genetic differentiation, often explain only a small fraction of the observed variance in genetic patterns in trees. The concept of isolation by wind will advance our understanding of how wind patterns shape evolutionary ecology only if and when wind connectivity models can overcome these uncertainties enough to detect clear systematic signs of wind-genetic relationships after accounting for distance and environment.

Macroecological approaches that test broad hypotheses across many species offer a partial solution to this uncertainty, by averaging over the idiosyncratic metapopulation histories and uncertainties that in any individual species may confound the signal of interest. In this study, we used this approach to test whether wind shapes large-scale genetic patterns in trees. We reanalyzed published landscape genetic data for more than 1,900 populations of 97 tree and shrub species from around the globe (Fig. 2A), integrating genetic metrics and dispersal-season wind connectivity models with functional trait data to test each of the four hypotheses described above.

The assembled data sets include a heterogeneous mix of nuclear and chloroplast DNA for species with varying reproductive ecology. For each species, the expected role of wind in shaping genetic patterns will depend on the combination of three traits: pollination syndrome, dispersal syndrome, and chloroplast DNA inheritance (Fig. 2B). We classify pollination and dispersal each as either wind or nonwind (mixed wind-animal pollination or dispersal syndromes were placed in the wind category). Chloroplast DNA is maternally inherited and dispersed through seeds

in most angiosperms but paternally inherited and dispersed through pollen in most conifers, which sets up important differences in the landscape genetics of these two groups (32). By combining these three traits, we can classify each genetic data set as falling into one of three “wind dispersal levels,” indicating whether wind is expected to drive spatial genetic patterns fully, partially, or not at all (Fig. 2B). All partially wind-dispersed data sets in this classification are diploid nuclear genomes in species that receive wind-dispersed genes from just one of their two parents, while fully wind-dispersed or nonwind-dispersed data sets include haploid plastid genomes as well as diploid nuclear genomes.

Focusing on different facets of this multispecies data set, three predictions can be made about the ways wind should influence a given genetic metric (Fig. 2C). The first prediction is that for fully or partially wind-influenced genomes, wind and genetic patterns will be positively correlated after controlling for distance and environment. The second approach compares the three wind dispersal levels, predicting that increasingly wind-dispersed genomes will exhibit increasingly strong genetic correlations with wind. The third prediction focuses on the subset of data sets where nuclear and plastid DNA were both collected for the same individuals and populations; in our case these all happen to be oak species, in which plastid DNA is exclusively animal-dispersed while nuclear DNA is transported by both wind and animals. The plastid genome in these oak populations can be used as an *in vivo* statistical control to isolate the wind-specific signal in the nuclear genome by removing the confounding effect of animal dispersal, holding everything else constant, with the prediction that the residual

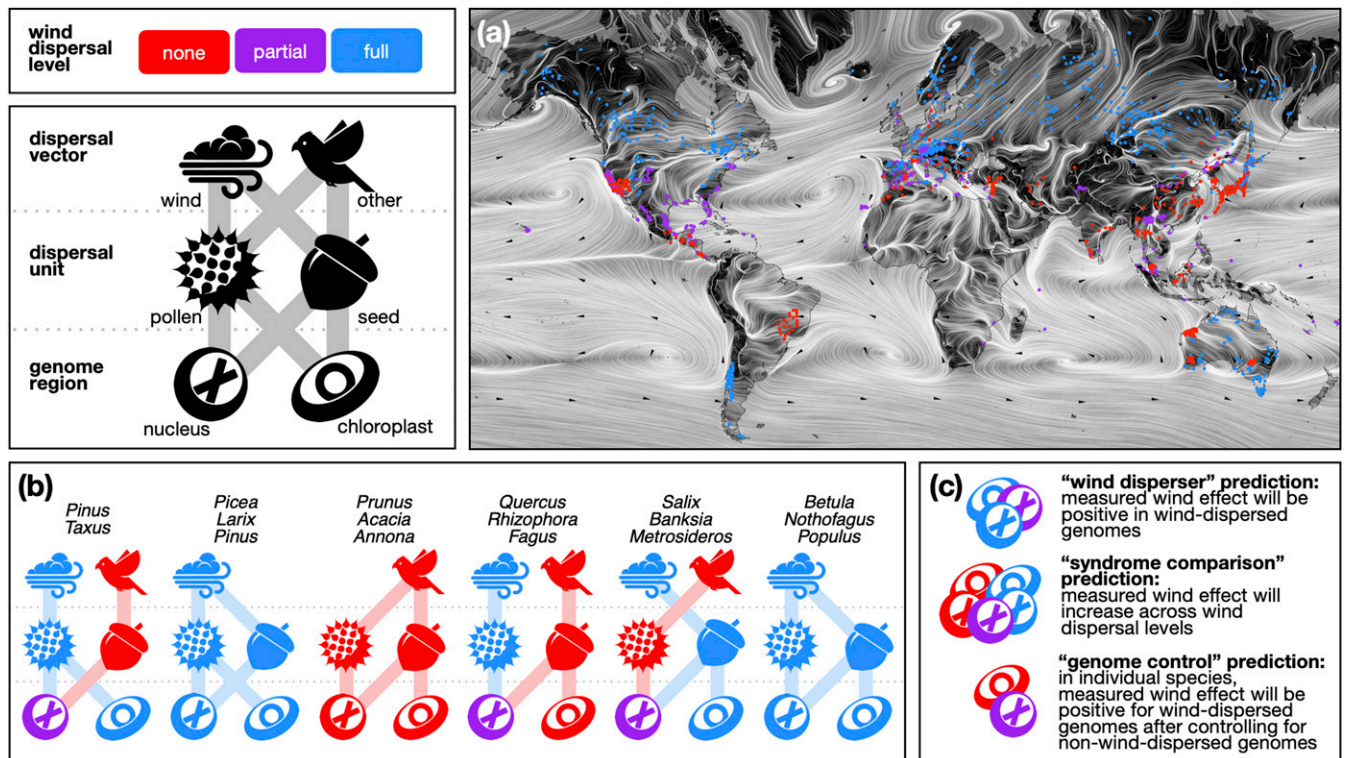


Fig. 2. Distribution and wind dispersal ecology of the genetic data sets analyzed in this study. (A) Population locations (points, colored by wind dispersal level) and global prevailing wind direction (black arrows and white paths; a Mercator projection is used to avoid distorting direction). (B) Schematics of relationships between seed plant genomes and wind dispersal. The influence of wind on a particular region of the genome (nuclear or chloroplast) is determined by species-specific differences in which dispersal unit (pollen or seed) carries the DNA and which dispersal vector (wind or nonwind) transports each dispersal unit. The six distinct syndromes exhibited by species in this study are shown, with example genera listed for each (eight syndromes are theoretically possible, but animal pollination is highly unusual in gymnosperms where chloroplast DNA tends to be paternally inherited). (C) Three predictions about relationships between wind dispersal level and landscape genetic patterns, each addressed using a different combination of data sets.

nuclear genetic signal will then be positively correlated with wind. We test each of these predictions for each of the four landscape genetic metrics described above (flow, isolation, asymmetry, diversity), for a total of 12 hypothesis tests (*SI Appendix, Table S1*).

In sum, our goals in this study are twofold. We begin with a basic quantification of wind connectivity patterns among populations of tree species to provide a descriptive assessment of the potential for variation in wind speed and direction to shape spatial genetic patterns. We then use these data to test the four major hypotheses about the way wind connectivity shapes gene flow, genetic differentiation, asymmetric gene flow, and genetic diversity across tree species ranges. These statistical tests required that we develop an extension of existing inference methods due to the structure of our data set, which comprises pairwise data for many species; while methods for analyzing pairwise data and for analyzing hierarchical multispecies data are both widely used, we are not aware of any published method for data that combines both these characteristics. We therefore introduce two alternative statistical tests for this purpose (*SI Appendix, Fig. S3*), each based on extending the traditional partial Mantel test to our multispecies case, and discuss differences between these approaches.

Results

We found usable data from 72 publications, representing 97 tree species, 120 data sets, and 1,940 populations from around the world, with a total of 28,286 pairwise population comparisons within data sets (Fig. 2 and [Supplementary Data](#)). The data included 51 fully wind-dispersed genomes, 35 partially wind-dispersed genomes, and 34 nonwind-dispersed genomes. There were 100 (20) nuclear (chloroplast) data sets and 107 (13) simple sequence repeat (SSR) (single-nucleotide polymorphism [SNP]) data sets.

Measures of wind conductance among populations showed that wind flow rates are highly spatially variable and that this is a product of both strong directional asymmetry and high geographic variation in wind speed (Fig. 3). While wind travel time is correlated with distance, wind flow speeds, which express variation in wind travel times after controlling for distance, varied by a factor of more than 30 across the analysis. The median pair of populations had a wind asymmetry ratio greater than 2:1, while some had ratios greater than 10:1. Pairwise mean wind diffusion speed, representing the strength of wind connectivity after factoring out directionality, varied by a factor of more than 10 over the entire analysis and by a factor of more than 4 across population pairs within the median individual data set.

Each of the four hypotheses about how these wind patterns affect landscape genetics (flow, isolation, asymmetry, diversity) was tested against three predictions, for a total of 12 wind-genetic relationships (*SI Appendix, Table S2*). We used two variations of a hierarchical Mantel test (*SI Appendix, Fig. S3*), which we call the “data set null” test and “global null” test according to whether null distributions were evaluated separately for each data set and then combined or first combined into a single null distribution for a global summary statistic and then evaluated. The two inference methods yielded strongly correlated estimates of statistical significance across the 12 hypothesis tests ($r = 0.93$), with the global null method estimating more extreme P values (one-sided P values farther from 0.5) on average compared to the data set null method (*SI Appendix, Fig. S4*).

Of the 12 relationships we analyzed, one had an effect opposite the hypothesis, one was near zero, and the other 10 had effects in the hypothesized direction; eight of these 10 had at least marginally significant P values (Fig. 4). All four genetic hypotheses had results consistent with the “wind dispersers” prediction that the majority of wind-dispersed genomes have positive wind-genetic correlations. Results for all four hypotheses were also consistent with the “syndrome comparison” prediction that wind-

genetic correlations increase with wind dispersal level. For the “genome control” prediction focused on six oak species, we found positive correlations under the flow and asymmetry hypotheses, a result near zero for the isolation hypothesis, and a negative correlation for the diversity hypothesis; of these, only the flow result was marginally significant. Wind-dispersed SNP and SSR data sets exhibited similar results for all four hypotheses (*SI Appendix, Fig. S5*).

Patterns for flow, isolation, asymmetry, and diversity were largely independent. In the raw pairwise input data (*SI Appendix, Fig. S6A*), gene flow explained 34% of variation in genetic differentiation and 29% of variation in gene flow asymmetry ratios in the median data set, while all other combinations of genetic metrics had r -squared values less than 5%. In the results for each data set, partial correlation coefficients and Mantel P values did not correlate strongly among most genetic facets—excluding gene flow, r -squared values for all combinations of genetic facets were less than 4% (*SI Appendix, Fig. S6 B and C*).

In a set of three diagnostic analyses searching for evidence that nonwind spatial drivers of genetic patterns might be systematically confounded with wind patterns in ways that could bias our results, we found no indication of problematic patterns (see details in *SI Appendix, Appendix 3*).

Discussion

We found that wind patterns are both strongly directional and geographically variable, in contrast to the assumptions of many classical models that wind dispersal potential is isotropic and spatially uniform. Directional wind diffusion speeds between pairs of conspecific tree populations, reflecting wind accessibility after controlling for distance, varied more than 30-fold across the world. This variation results in part from highly asymmetric wind flows along many dispersal routes: Wind travel between populations took more than twice as long in one direction as the other for the median population pair, and more than ten times as long for some pairs. It is also a product of geographic variation in wind strength: There was a fourfold variation in pairwise mean wind speed (controlling for distance and directionality) among population pairs of the median species in our analysis, and more than tenfold variation across populations of all species. These estimates help to clarify the potential for anisotropic and spatially variable wind regimes to shape dispersal dynamics and biodiversity patterns. The strength and directionality of currents in river and ocean systems are known to have major evolutionary consequences (25, 29, 33); while the degree of asymmetry is lower for aerial dispersal due to the temporal variability associated with weather systems and fine-scale atmospheric turbulence, our models highlight the potential for important biological effects. Asymmetric and nonstationary gene flow rates that correspond to our estimated wind flow rates can have a strong influence on evolutionary outcomes.

Our results strongly suggest that these wind patterns do indeed influence forest genetics. We found evidence that wind influences gene flow, genetic isolation, asymmetric gene flow, and genetic diversity patterns in each of the ways we hypothesized. For each of these genetic facets, partial correlations between wind and genetics were positive in wind-dispersed genomes after controlling for distance and environment and were higher in wind-dispersed genomes than nonwind-dispersed genomes. Results of the genome control test (*Quercus* data) also trended in the hypothesized direction, with the exception of the diversity hypothesis, although these tests were mostly inconclusive due to small sample size and high variability in effect size among species. Because we tested the influence of wind on each genetic hypothesis in three distinct ways and found general agreement across these tests, we can have higher confidence that the observed effects of wind are real. The two new inference methods we used to estimate the statistical significance of these trends

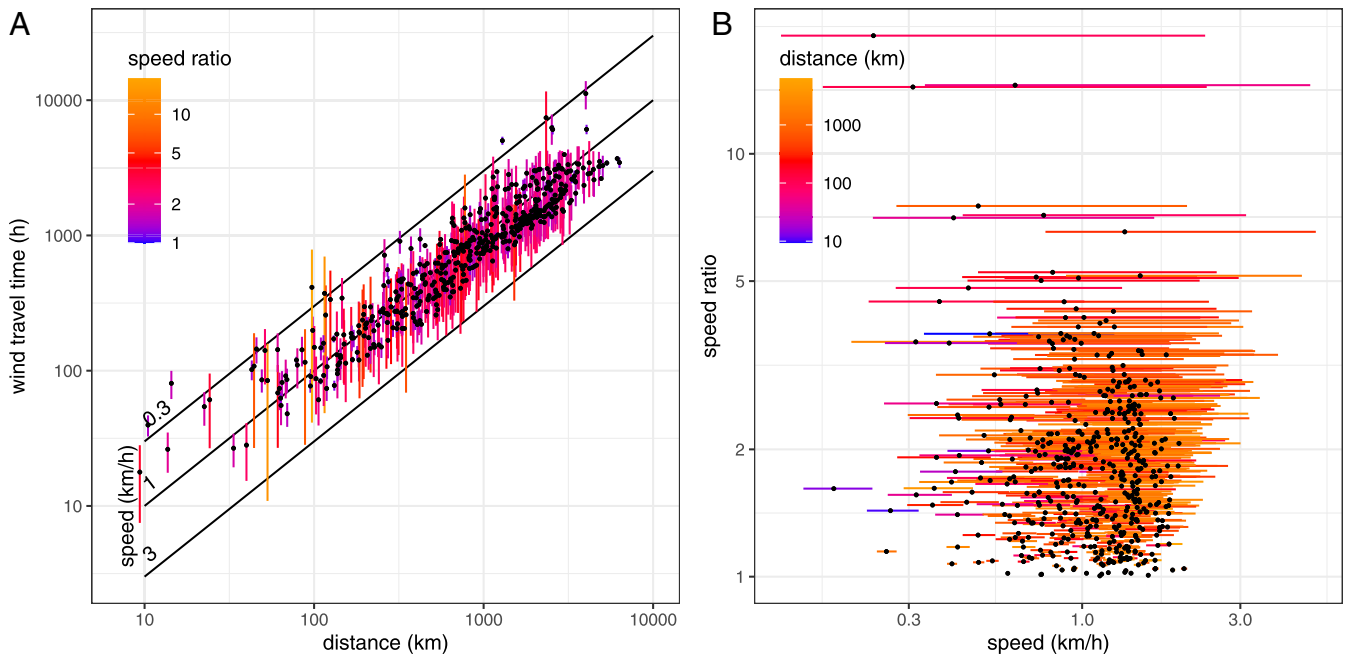


Fig. 3. Variation in, and relationships among, wind connectivity metrics. *A* and *B* show different views of the same data, representing geographic distance, wind travel time, wind speed, and wind speed ratios for 500 population pairs randomly selected across all species in the analysis. For each population pair, points indicate the average wind speed or wind travel time (used for the isolation hypothesis), while line segments link the wind flow metrics in the two directions (used for the other three hypotheses).

yielded highly correlated results, with the global null method generally estimating higher levels of significance. Taken together, these tests provide clear evidence for the effects of wind on large-scale landscape genetic patterns in trees.

Importantly, our results show that wind has distinct, independent effects on genetic isolation, gene flow asymmetry, and genetic diversity. Wind-genetic correlations for wind-dispersed genomes were almost entirely uncorrelated among these three hypotheses. (Results for gene flow, the fourth hypothesis, were correlated with asymmetry and isolation, as expected given mathematical relationships among the metrics). This confirms that these results are not simply alternative measures of the same underlying pattern in the raw genetic data that were predetermined to yield similar results but instead represent three genuinely distinct tests of the influence of wind on different facets of landscape genetics. It also implies that while all three patterns are consistently influenced by differences among the life history traits we used to determine wind dispersal level, variation among species within a given wind dispersal group does not consistently shape the overall influence of wind. Whichever aspects of life history or biogeography do shape the variation among species appear to operate independently on these three landscape genetic metrics.

While the measured effects of wind on these genetic metrics are clear and significant, they are not especially strong or consistent. The average partial correlation between wind and genetic metrics is low, which is to be expected given that wind flow is highly correlated with geographic distance (Fig. 3*A*), and our results reflect only the portion of the wind effect that is not jointly explained by distance. There is also wide variability among species, with the role of wind becoming clear only when data are pooled across large numbers of taxa. While this analysis reveals that wind connectivity models are useful for understanding landscape genetic patterns at macroecological scales, our results imply that they still fall short for many individual species. On average across the four facets, partial correlations for roughly two-thirds (64%) of fully wind-dispersed genomes had Mantel *P* values in the hypothesized direction. Given the substantial uncertainties in estimates of both

wind connectivity and genetic patterns discussed above and given the low statistical power for many species with small sample sizes, it is perhaps unsurprising that a third of species did not follow predicted trends.

Landscape genetic patterns arise from complex interactions among a variety of spatially structured historical processes in addition to wind dispersal, and these other processes have the potential to confound inferences about wind. This possibility merits careful consideration. Confounding factors operating on individual species are probably common in our analysis and are likely a contributor to the wide variance in patterns among data sets. In order to bias our multispecies inferences rather than simply adding noise to our results, these confounding factors would need to operate similarly across many species, making them statistically nonindependent. This could be the case, for example, if multiple species expanded from shared glacial refugia in parallel, giving them similar genetic patterns that happened to align with wind in ways that mimic our hypothesized wind effects.

Possibilities like this cannot be entirely eliminated in an observational study, but several points argue against systematic bias. First, our study design hedges against this by testing each hypothesis in three different ways. Second, three of the genetic facets were almost entirely uncorrelated across species as noted above, meaning they are unlikely to be jointly biased by any single confounding phenomenon. And third, a set of diagnostic analyses searching for evidence of systematic, spatially structured confounding factors yielded very little cause for concern (*SI Appendix, Appendix 3*). For these reasons, we conclude that it is quite unlikely that the main patterns in our results are spurious. The alternative, biologically unsurprising possibility that wind itself has shaped these genetic patterns seems far more likely.

While the overall results are consistent with our core hypotheses for all four genetic facets, other aspects of the data are unexpected. Surprisingly, the majority of nonwind dispersers had significantly negative (rather than zero) wind-genetic correlations for the isolation and asymmetry analyses. This raises the possibility of biases in the measured correlations and underscores the value

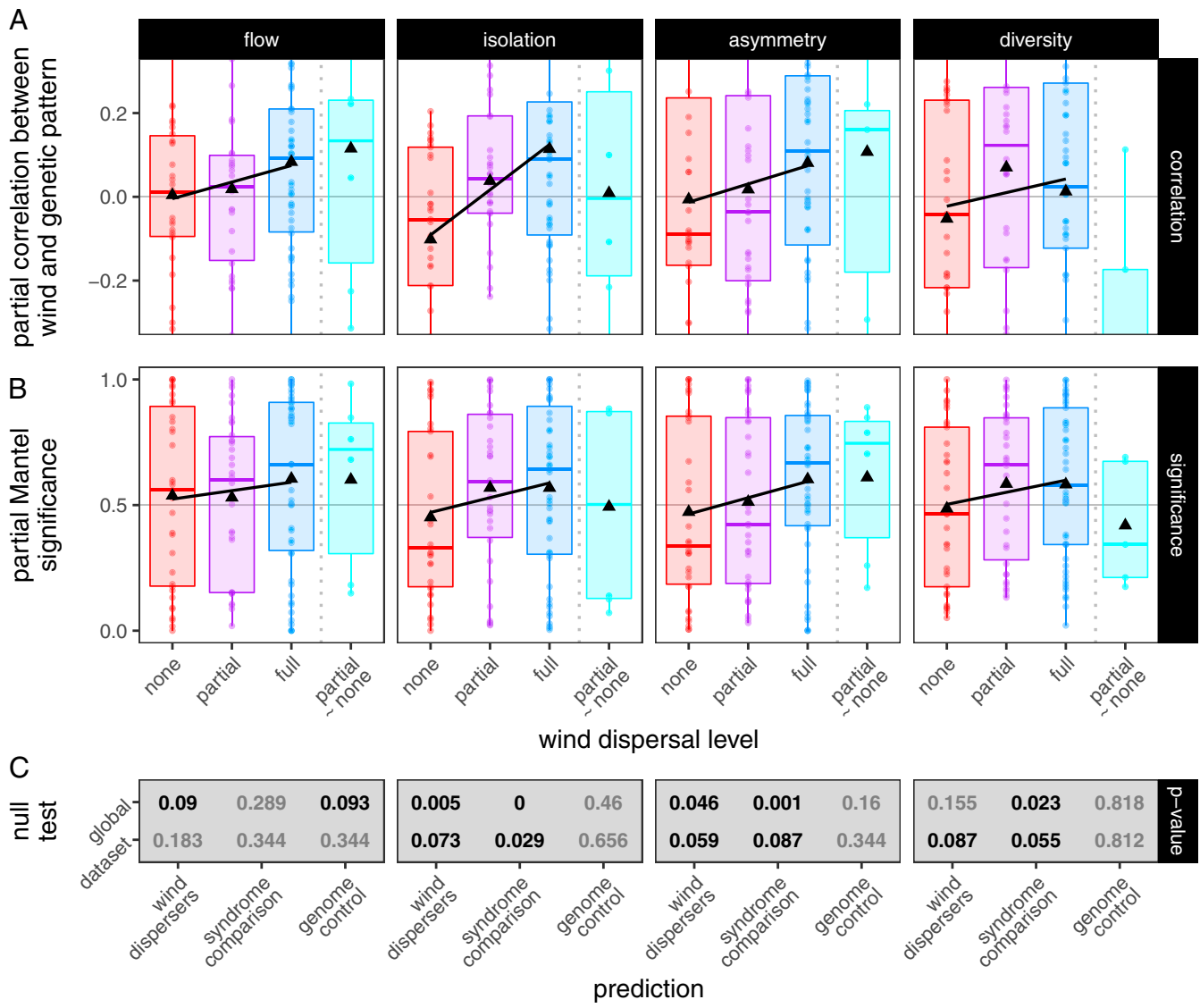


Fig. 4. Partial correlations between wind and each of the four landscape genetic patterns: flow, isolation, symmetry, and diversity. (A) Distributions of partial correlation coefficients for each data set; plots hide absolute values above 0.3 to emphasize central patterns over outliers. (B) Distributions of data set-level Mantel significance for each partial correlation; these are one-sided *P* values, with the direction set so that higher values correspond to more positive correlation coefficients. In A and B, red-purple-blue boxplots represent correlations for data sets with different wind dispersal levels as indicated on the x-axis, while cyan boxplots represent correlations for the six *Quercus* nuclear data sets in the “genome control” analysis; triangles are averages across data sets. (C) Statistical significance of the three distinct predictions for each of the four hypotheses, each tested using a “global null” test based on the correlations in A and a “data set null” test based on the *P* values in B; tests not significant at *P* < 0.1 are grayed out.

of comparing genomes with different wind dispersal levels in addition to analyzing wind dispersers on their own. Modeling the potential causes of these unexpected isolation and asymmetry results is beyond the scope of this study, but we can speculate about how they might contribute to negative correlations in nonwind dispersers and dampen positive correlations in wind dispersers. For the asymmetry hypothesis, negative wind-genetic correlations indicate asymmetric gene flow against the prevailing wind direction. Founder effects, which may be common in postglacial range expansions, can generate a bias in perceived migration rates, incorrectly estimating net migration from the newly founded population toward the older source population (19); if prevailing winds tended to blow opposite the direction of recent range expansions, the expected genetic signal of founder effects and wind effects would thus be similar, and the result could be confounded. But in the northern temperate latitudes representing the large majority of our

data sets, where postglacial range expansions and prevailing meridional winds both move in the poleward direction, founder events would instead be expected to cause negative bias in the estimated effect of wind asymmetry. For the isolation hypothesis, negative relationships between genetic similarity and wind connectivity in nonwind-dispersed species imply that genetic exchange resulting from animal movement is either higher in less windy portions of a species range or is higher along routes perpendicular rather than parallel to prevailing winds. This latter explanation could potentially result from east-west windspeeds in most places being stronger than north-south windspeeds, while seasonal migration pathways of animals, as well as range expansion pathways between the Last Glacial Maximum (LGM) and present day, tend to be oriented in the north-south direction.

It is clear there is room for improvement in future studies of landscape wind connectivity. Because trees tend to have high

levels of standing genetic variation within versus among populations (26, 34, 35), sampling uncertainty can be high, and so collecting data from larger numbers of populations, individuals, and/or loci could give more power to detect subtle patterns such as wind effects. Wind connectivity models could also be refined in a number of ways. For example, our modeled wind flow pathways were unconstrained by landscape features (other than large water bodies) and were based on all wind conditions during each species' dispersal or pollination season; more realistically, gene flow is perhaps likely to follow pathways through inhabited patches within the species range, and dispersal can be driven by specific weather conditions at specific times of day (36–38). These factors could be accounted for given sufficient information on a focal species. Another technical limitation is that the wind connectivity models used here rely on least-cost-path (LCP) distances; whereas resistance-based methods that account for additional dispersal pathways beyond the LCP are a better fit with gene flow theory (39), they are not compatible with the cyclic, asymmetric, directed connectivity graphs needed to represent long-term wind variation. There is potential for future windscape-based gene flow models to expand beyond LCP by incorporating random walks that simulate individual dispersal events.

It is worth noting that our use of three discrete wind dispersal levels is a simplification of functional trait variation across species. One issue is that propagules can be transported by both wind and nonwind vectors, making the relative contribution of wind to dispersal and pollination patterns a gradient rather than a discrete ranking; examples include pines with wind-dispersed seeds secondarily dispersed by rodents, willows pollinated by both insects and wind, and figs pollinated by wasps which are themselves subsequently transported by wind. Another issue is that the efficacy of long-distance wind dispersal varies enormously even among wind-dispersed taxa, ranging from heavy winged seeds and large pollen grains that generally disperse only short distances, to cottony seeds and small pollen grains capable of traveling vast distances. On the one hand, the use of three discrete levels makes sense in the context wind-genetic correlations, which measure the predictability of gene flow rather than absolute rates of gene flow—functional variation in wind dispersal ability should relate more to absolute rates, whereas predictability may simply relate to the degree of noise contributed by nonwind dispersal vectors like animals. On the other hand, genetic patterns reflect the balance between multiple evolutionary rates, and patterns in genomes with slower rates of wind dispersal should be relatively more influenced by processes such as selection and drift, making them less predictable from wind. It would be valuable for future work to integrate quantitative measures of wind dispersal ability.

Finally, our results imply that wind patterns are relevant to the conservation and management of forests vulnerable to multiple aspects of anthropogenic environmental change, with the caveat that our analysis reflects processes over timescales much longer than contemporary environmental changes. With ongoing climate change, the rate and direction of gene flow are important for transporting adaptive alleles to both the warm and cold edges of a species range (23, 40, 41), and the efficacy of these processes is likely to be shaped by global wind patterns (17). With widespread habitat destruction and fragmentation, isolated forest fragments will depend on long-distance wind dispersal and pollination for maintaining genetic diversity and reducing inbreeding depression (42, 43). And with increasing concern about engineered genes leaking from commercial forestry plantations into wild tree populations, wind strength and direction are likely to shape rates of genetic contamination (44, 45). To help managers assess and mitigate these risks, future studies will need to further develop our understanding of how wind geography shapes gene flow at different timescales.

Conclusions

Our results offer insight into the ways that wind patterns shape various aspects of landscape genetic patterns across populations of wind-dispersed and wind-pollinated plants. While the marginal effects of wind are subtle after controlling for distance and climate, they are clearly detectable using the combination of time-integrated wind connectivity models, a large global multi-species genetic data set, and extensions of existing statistical methods. This analysis presents evidence of the large-scale influences of wind on spatial genetic patterns, demonstrating that wind has distinct influences on genetic diversity, genetic differentiation, and asymmetric gene flow.

Materials and Methods

Genetic Data. Our study is based on reanalysis of previously published data sets available on the Dryad digital data repository. We compiled a list of 165 tree genera ([Supplementary Data](#)), including all native tree genera of the United States and Europe as well as a manageable set of major global tree genera abbreviated from the full set of over 4,000 tree genera worldwide. We entered each genus as a search term in Dryad and reviewed the abstracts of all results for each genus. Microsatellite (SSR) and SNP data sets from landscape genetics studies representing multiple individuals of a given species from multiple geographic locations were downloaded. We use the term “data set” to refer to the data of a single sequence type (SSR or SNP) representing a single genome (nuclear or chloroplast) of a single species from a single publication; some publications contained multiple data sets, which we disaggregated for most downstream analysis. Studies focusing on hybridization between species were excluded, except in cases where populations in the hybrid zone of overlap between species ranges could be identified and removed, retaining only the nonoverlapping portions of species ranges for analysis.

Each data set was individually restructured into a standardized format. All populations covered relatively small areas, and their locations were represented as point localities. Because wind connectivity cannot be computed between populations occurring in the same wind grid cell (see below), populations from a given genetic data set that fell within the same cell were merged prior to analysis by averaging the latitude and longitude of the constituent populations and pooling individual genotype data into a single population. After all filtering, cleaning, and reformatting, our final analysis was based on 120 data sets representing 1,940 populations of 97 tree and shrub species, sourced from 72 original publications (Fig. 2B; details in [Supplementary Data](#)). The median data set had 12 populations (interquartile range [IQR]: 6 to 20) and 19.8 individuals per population (IQR: 13.3 to 31.8); the median number of loci was 11 for SSR data sets and 267 for SNP data sets.

Genetic Metrics. We analyzed genetic patterns in each data set using a fully connected lattice model of populations, in which all population pairs are connected. For each data set, we calculated genetic differentiation and directional migration (gene flow) for every population pair and genetic diversity for every population. Diversity was measured as mean allelic richness for every population using the resampling method in the *divBasic* function in the R package *divRsity* (46); to convert these diversity values into the data structure of pairwise relationships between populations, we calculated genetic diversity ratios for each population pair in each direction as the ratio of allelic richness in the destination versus the origin population. Pairwise genetic similarity was measured as $1/F_{st}$.

Migration was estimated using the *divMigrate* method (19) with Jost's *D* metric of differentiation (47), as implemented in the R package *divRsity* (46). This approach uses allele frequency differences between population pairs to estimate rates of migration in each direction; note that these rates are relative to other population pairs in the same data set and cannot be compared across data sets. This method was chosen for its straightforwardness and computational tractability given the size of our analysis, compared to computationally intensive model-based alternatives (20, 31). We modified the *divMigrate* function in order to accommodate haplotype (chloroplast) data, which are consistent with the theoretical method but not supported in the original R package. Gene flow asymmetry was calculated for each population pair as the ratio of outbound to inbound gene flow; reciprocal values were calculated reversing outbound and inbound, for distance-based matrix analyses.

Traits. For each species, we compiled data on pollination and seed dispersal syndromes, the months of the year when pollination and seed dispersal occur,

and whether chloroplast DNA is transmitted via pollen or seed (Supplementary Data). Data were sourced from TRY (48) and BIEN (49) using bulk multispecies queries, and holes were filled by manual searching on Google Scholar and Google. Based on these traits, each data set was classified as nonwind dispersed, partially wind dispersed, or fully wind dispersed, according to the expected role of wind in shaping genetic patterns (Fig. 2).

Wind. We used the *windscape* R package (50) in combination with three decades (1980 through 2009) of hourly global wind data ($n = 262,800$ timesteps in every location) from the Climate System Forecast Reanalysis (CFSR) (51) to estimate wind connectivity among sites. The CFSR is a gridded data set with a spatial resolution of ~ 35 km and is considered one of the best representations of actual atmospheric conditions over recent decades. In the *windscape* modeling framework, a connectivity graph is constructed in which each grid cell is connected to its eight neighbors in proportion to the frequency and speed with which wind blows in that direction, integrating over the full multidecadal time series of hourly wind conditions. Wind conductance over nonterrestrial cells was down-weighted by 90% to reflect the reduced likelihood of long-distance dispersal across large water bodies. The most efficient route between any two points was then identified using a LCP algorithm, with its “cost” quantified as the mean estimated number of hours wind would take to diffuse to that location based on the full spatiotemporal distribution of wind conditions across the landscape. We quantify wind flow as the inverse of this travel time, measured in units of h^{-1} .

We calculated wind flow in both directions for every pair of populations in each data set. For each population pair, we used these two flow values to calculate wind flow asymmetry (i.e., the ratio of outbound to inbound wind flow, which like gene flow asymmetry are reciprocals that contain the same information). We also calculated wind connectivity, the mean of wind flows in the two directions.

Wind patterns differ seasonally in many parts of the world, and for each data set we therefore calculated wind flow based on data only from months of the year when wind pollination and/or dispersal is thought to occur for that genome. For data sets where wind dispersal or pollination is relevant, wind data from only the relevant months was used. For nonwind-dispersed data sets, and for the small minority of data sets where pollination or dispersal phenology was unknown, wind data from all months of the year were used. This was determined for each data set based on the combination of genome type, plastid inheritance, and seed and pollen dispersal syndromes.

Climate. To control for IBD and IBE, we also calculated the pairwise geographic distance and the pairwise climatic difference between every pair of populations. Climatic difference was calculated based on four climate variables deemed likely to shape patterns of local adaptation in trees: maximum temperature of the warmest month, minimum temperature of the coldest month, annual actual evapotranspiration (AET, a measure of water available for plant growth), and annual climatic water deficit (CWD, a measure of dryness intensity). We used gridded 1 km global terrestrial climate data from CHELSA (52) representing mean climates from 1979 to 2013. We derived AET and CWD from monthly temperature and precipitation values and latitude following the methods of Wang et al. (53) and then transformed them for normality using $\log(x + 1)$. The four variables were then standardized using a principal component analysis across all terrestrial grid cells outside the (ant) arctic circles, yielding four orthogonal climate dimensions with equal variances. These principal component values were extracted for the point location of each population in the genetic data sets and used to calculate pairwise Euclidean distances representing climatic dissimilarity between populations.

Statistical Models. The steps described above generated nine pairwise matrices for each data set, including two asymmetric matrices with different values in the upper and lower triangles (gene flow and wind flow), three reciprocally symmetrical matrices with reciprocal values in the upper and lower triangles (gene flow ratio, genetic diversity ratio, and wind flow ratio), and four symmetric matrices with identical values in the upper and lower triangles (genetic similarity, geographic distance, climatic difference, and wind connectivity). All variables were log-transformed prior to statistical analysis, both to improve the normality of data distributions and so that correlations would reflect fine-scale variance among nearby/similar populations as well as large-scale variance among relatively distant populations, rather than being dominated by the latter.

These data were used to test the four hypotheses about how wind shapes different landscape genetic patterns. We tested the flow hypothesis as the partial correlation between wind flow and gene flow; the isolation hypothesis as the partial correlation of wind connectivity and genetic similarity;

the asymmetry hypothesis as the partial correlation between wind flow ratio and gene flow ratio; and the diversity hypothesis as the partial correlation between wind flow ratio and genetic diversity ratio. All of these partial correlations were hypothesized to be positive for wind-dispersed genomes. Geographic distance and climatic difference were included as controls in the partial correlations for the flow and isolation analyses, ensuring that only residual variation not associated with these predictors was tested. (Distance and climatic difference by definition have zero correlations with log-transformed wind and genetic asymmetry ratios, so it was not necessary to include them as controls in the asymmetry and diversity analyses.)

For each of these four hypotheses, we tested three predictions, for a total of 12 hypothesis tests. We refer to these as the “wind disperser,” “syndrome comparison,” and “genome control” predictions (Fig. 2C). The wind disperser prediction considers only the fully or partially wind-dispersed data sets, predicting that the majority of these will have positive partial correlations between the wind and genetic metrics. The syndrome comparison prediction considers all data sets, predicting that the partial wind-genetic correlation will increase across the three wind dispersal levels (none, partial, full). Finally, the genome comparison prediction considers only data sets from studies where partially wind-dispersed nuclear data and nonwind-dispersed chloroplast data were collected for the same individuals and populations, which comprised six *Quercus* species, predicting that the majority of nuclear data sets will have positive partial wind-genetic correlations after the chloroplast genetic signal is added as a control.

All tests were conducted using partial Mantel tests, the statistical approach that has been most widely used (54) to test patterns in pairwise matrix data for applications such as IBD and IBE. This test is based on the partial correlation statistic, which measures the correlation between two variables after controlling for additional covariates: conceptually, two focal variables (e.g., wind and genetic relationships) are each separately regressed against the other covariates (e.g., distance and environmental difference), and the correlation between the residuals from these two regressions is calculated. The Mantel test estimates significance using a null randomization, but because we are testing global hypotheses about macroecological patterns across the entire multispecies data set rather than hypotheses about any individual species in particular, an extension of the traditional single-species Mantel is needed. We used two alternative approaches that accommodate this hierarchical structure, which we call the “data set null” and “global null” approaches, and tested each of our 12 hypotheses using both approaches for a total of 24 hypothesis tests (SI Appendix, Table S2). Both approaches begin by using Mantel-style permutations of the rows and columns of the genetic matrix of every data set to derive a null distribution of 1,000 randomized partial correlation coefficients for every data set; they differ in how those randomized null values are summarized to derive a final global P value representing the significance of the overall hypothesis (SI Appendix, Fig. S3).

For the global null approach, we used these randomization data to calculate global summary statistics across data sets separately for each random iteration, to derive null distributions of these global test statistics that we then compared to the measured test statistic to test the overall hypothesis. Our test statistic for the wind disperser and genome control predictions was the median correlation coefficient across data sets. For the syndrome comparison prediction, our test statistic was the Spearman rank correlation between wind dispersal level and correlation coefficient.

For the data set null approach, we used the randomization data to calculate one-sided P values separately for each data set, comparing each data set’s measured correlation coefficient to its null distribution following the standard practice for the partial Mantel test. These one-sided P values contain information about both the direction and uncertainty of the relationship for each data set. To derive final global P values, we tested whether this collection of independent P values deviated significantly from null expectations. For the wind disperser and genome control predictions, the null expectation is that P values above and below 0.5 are equally likely; we tested this with simple one-sided binomial tests. For the syndrome comparison prediction, the null expectation is that P values do not differ across the three wind dispersal levels; we used one-sided Spearman’s rank correlation tests to determine whether P values decreased as wind dispersal level increased.

Lastly, we performed three analyses to evaluate whether confounding spatially structured phenomena are likely to have spuriously generated the observed relationships between wind and genetic patterns. These are described in SI Appendix, Appendix 3.

All analysis was done in R.

Data Availability. This study is based on previously published genetic data from 72 publications. Citations to these studies, and results from our analysis,

are included in [Supplementary Data](#)). The computer code for our analysis is deposited on Zenodo (DOI: [10.5281/zenodo.4660801](https://doi.org/10.5281/zenodo.4660801)).

ACKNOWLEDGMENTS. We are deeply indebted to the many authors of prior studies, listed in the [Supplementary Data](#), whose choice to openly publish

their landscape genetic data on Dryad made our reanalysis possible. We are grateful to research assistants Hudson Northrop and Yiwen Shi for their work locating and standardizing these data sets and to Ben Blonder, Perry de Valpine, and Ian Wang for helpful discussion about data analysis methods. M.M.K. was funded by an NSF Graduate Research Fellowship.

1. H. F. Howe, J. Smallwood, Ecology of seed dispersal. *Annual Review of Ecology and Systematics* **13**, 201–228 (1982).
2. J. Ollerton, R. Winfree, S. Tarrant, How many flowering plants are pollinated by animals? *Oikos* **120**, 321–326 (2011).
3. P. J. Regal, Pollination by wind and animals: Ecology of geographic patterns. *Annu. Rev. Ecol. Syst.* **13**, 497–524 (1982).
4. J. Muñoz, Á. M. Felicísimo, F. Cabezas, A. R. Burgaz, I. Martínez, Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* **304**, 1144–1147 (2004).
5. R. G. Gillespie *et al.*, Long-distance dispersal: A framework for hypothesis testing. *Trends Ecol. Evol.* **27**, 47–56 (2012).
6. S. Ahmed, S. G. Compton, R. K. Butlin, P. M. Gilmartin, Wind-borne insects mediate directional pollen transfer between desert fig trees 160 kilometers apart. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 20342–20347 (2009).
7. F. Austerlitz, C. Dutech, P. E. Smouse, F. Davis, V. L. Sork, Estimating anisotropic pollen dispersal: A case study in *Quercus lobata*. *Heredity* **99**, 193–204 (2007).
8. C. Born, P. C. Le Roux, C. Spohr, M. A. McGeoch, B. J. Van Vuuren, Plant dispersal in the sub-Antarctic inferred from anisotropic genetic structure. *Mol. Ecol.* **21**, 184–194 (2012).
9. Z. F. Wang *et al.*, Pollen and seed flow under different predominant winds in wind-pollinated and wind-dispersed species *Engelhardia roxburghiana*. *Tree Genet. Genomes* **12**, 19 (2016).
10. M. V. Ashley, Plant parentage, pollination, and dispersal: How DNA microsatellites have altered the landscape. *Crit. Rev. Plant Sci.* **29**, 148–161 (2010).
11. R. Nathan *et al.*, Mechanistic models of seed dispersal by wind. *Theor. Ecol.* **4**, 113–132 (2011b).
12. R. Nathan *et al.*, Spread of North American wind-dispersed trees in future environments. *Ecol. Lett.* **14**, 211–219 (2011a).
13. S. Thompson, G. Katul, Plant propagation fronts and wind dispersal: An analytical model to upscale from seconds to decades using superstatistics. *Am. Nat.* **171**, 468–479 (2008).
14. S. Schueler, K. H. Schlünzen, Modeling of oak pollen dispersal on the landscape level with a mesoscale atmospheric model. *Environ. Model. Assess.* **11**, 179 (2006).
15. M. I. Gassmann, C. F. Pérez, Trajectories associated to regional and extra-regional pollen transport in the southeast of Buenos Aires province, Mar del Plata (Argentina). *Int. J. Biometeorol.* **50**, 280–291 (2006).
16. J. Fernández-López, K. Schliep, rWind: download, edit and include wind data in ecological and evolutionary analysis. *Ecography* **42**, 804–810 (2018).
17. M. Kling, D. Ackerly, Global wind patterns and the vulnerability of wind-dispersed species to climate change. *Nat. Clim. Chang.* **10**, 868–875 (2020).
18. Y. Tsuda, V. Semerikov, F. Sebastiani, G. G. Vendramin, M. Lascoux, Multispecies genetic structure and hybridization in the *Betula* genus across Eurasia. *Mol. Ecol.* **26**, 589–605 (2017).
19. L. Sundqvist, K. Keenan, M. Zackrisson, P. Prodöhl, D. Kleinhans, Directional genetic differentiation and relative migration. *Ecol. Evol.* **6**, 3461–3475 (2016).
20. P. Beerli, J. Felsenstein, Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* **152**, 763–773 (1999).
21. S. Wright, Isolation by distance. *Genetics* **28**, 114–138 (1943).
22. I. J. Wang, G. S. Bradburd, Isolation by environment. *Mol. Ecol.* **23**, 5649–5662 (2014).
23. R. Aguilée, G. Raoul, F. Rousset, O. Ronce, Pollen dispersal slows geographical range shift and accelerates ecological niche shift under climate change. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E5741–E5748 (2016).
24. T. J. Kawecki, R. D. Holt, Evolutionary consequences of asymmetric dispersal rates. *Am. Nat.* **160**, 333–347 (2002).
25. J. M. Pringle, A. M. Blakeslee, J. E. Byers, J. Roman, Asymmetric dispersal allows an upstream region to control population structure throughout a species' range. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15288–15293 (2011).
26. O. Savolainen, T. Pyhäjärvi, T. Knürr, Gene flow and local adaptation in trees. *Annu. Rev. Ecol. Syst.* **38**, 595–619 (2007).
27. R. J. Gornall, P. M. Hollingsworth, C. D. Preston, Evidence for spatial structure and directional gene flow in a population of an aquatic plant, *Potamogeton coloratus*. *Heredity* **80**, 414 (1998).
28. E. Lundqvist, E. Andersson, Genetic diversity in populations of plants with different breeding and dispersal strategies in a free-flowing boreal river system. *Heredity* **135**, 75–83 (2001).
29. M. B. Morrissey, D. T. de Kerckhove, The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. *Am. Nat.* **174**, 875–889 (2009).
30. J. S. Clark, M. Lewis, L. Horvath, Invasion by extremes: Population spread with variation in dispersal and reproduction. *Am. Nat.* **157**, 537–554 (2001).
31. G. A. Wilson, B. Rannala, Bayesian inference of recent migration rates using multi-locus genotypes. *Genetics* **163**, 1177–1191 (2003).
32. R. J. Petit *et al.*, Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol. Ecol.* **14**, 689–701 (2005).
33. R. K. Cowen, S. Sponaugle, Larval dispersal and marine population connectivity. *Annu. Rev. Mar. Sci.* **1**, 443–466 (2009).
34. R. J. Petit, A. Hampe, Some evolutionary consequences of being a tree. *Annu. Rev. Ecol. Syst.* **37**, 187–214 (2006).
35. J. L. Hamrick, M. J. W. Godt, S. L. Sherman-Broyles, "Factors influencing levels of genetic diversity in woody plant species" in *Population Genetics of Forest Trees*, W. T. Adams, S. H. Strauss, D. L. Copes, A. R. Griffin, Eds. (Springer, Dordrecht, 1992), pp. 95–124.
36. D. F. Greene, The role of abscission in long-distance seed dispersal by the wind. *Ecology* **86**, 3105–3110 (2005).
37. K. D. Maurer, G. Bohrer, D. Medvigy, S. J. Wright, The timing of abscission affects dispersal distance in a wind-dispersed tropical tree. *Funct. Ecol.* **27**, 208–218 (2013).
38. S. J. Wright *et al.*, Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19084–19089 (2008).
39. B. H. McRae, Isolation by resistance. *Evolution* **60**, 1551–1561 (2006).
40. M. Bontrager, A. L. Angert, Gene flow improves fitness at a range edge under climate change. *Ecol. Lett.* **3**, 55–68 (2018).
41. J. P. Sexton, S. Y. Strauss, K. J. Rice, Gene flow increases fitness at the warm edge of a species' range. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 11704–11709 (2011).
42. A. S. Jump, J. Peñuelas, Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8096–8100 (2006).
43. A. J. Lowe, S. Cavers, D. Boshier, M. F. Breed, P. M. Hollingsworth, The resilience of forest fragmentation genetics—No longer a paradox—We were just looking in the wrong place. *Heredity* **115**, 97–99 (2015).
44. E. Small, T. Antle, A preliminary study of pollen dispersal in *Cannabis sativa* in relation to wind direction. *J. Ind. Hemp* **8**, 37–50 (2003).
45. S. P. DiFazio, G. T. Slavov, J. Burczyk, S. Leonardi, S. H. Strauss, "Gene flow from tree plantations and implications for transgenic risk assessment" in *Forest Biotechnology for the 21st Century*, C. Walter, M. Carson, Eds. (Research Signpost, Kerala, India, 2004), chap. 23, pp. 405–422.
46. K. Keenan, P. McGinnity, T. F. Cross, W. W. Crozier, P. A. Prodöhl, diveRsim: An R package for the estimation of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **4**, 782–788 (2013).
47. L. Jost, G(ST) and its relatives do not measure differentiation. *Mol. Ecol.* **17**, 4015–4026 (2008).
48. J. Kattge *et al.*, Nutrient Network, TRY plant trait database—Enhanced coverage and open access. *Glob. Change Biol.* **26**, 119–188 (2020).
49. B. S. Maitner *et al.*, The bien r package: A tool to access the botanical information and ecology network (BIEN) database. *Methods Ecol. Evol.* **9**, 373–379 (2018).
50. M. Kling, Windscape, Version 1.0.0, R package, Zenodo repository, <https://zenodo.org/record/3857730>. Accessed 6 April 2021.
51. S. Saha *et al.*, The NCEP climate forecast system version 2. *J. Clim.* **27**, 2185–2208 (2014).
52. D. N. Karger *et al.*, Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**, 170122 (2017).
53. T. Wang, A. Hamann, D. L. Spittlehouse, T. Q. Murdock, ClimateWNA—high-resolution spatial climate data for western North America. *J. Appl. Meteorol. Climatol.* **51**, 16–29 (2012).
54. J. P. Sexton, S. B. Hangartner, A. A. Hoffmann, Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution* **68**, 1–15 (2014).