MOLECULAR ECOLOGY

Molecular Ecology (2017) **26**, 2498–2513 doi: 10.1111/mec.13999

Ecological genomics of tropical trees: how local population size and allelic diversity of resistance genes relate to immune responses, cosusceptibility to pathogens, and negative density dependence

J. H. MARDEN,*† S. A. MANGAN,‡§ M. P. PETERSON,*† E. WAFULA,* H. W. FESCEMYER,* J. P. DER,¶ C. W. DEPAMPHILIS*† and L. S. COMITA§††

*Department of Biology, Pennsylvania State University, University Park, PA 16802, USA, †Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA 16802, USA, ‡Department of Biology, Washington University in St. Louis, St. Louis, MO 63130, USA, §Smithsonian Tropical Research Institute, República de Panamá, 0843-03092 Panama, Panama, ¶Department of Biological Science, California State University, Fullerton, CA 92834, USA, ††School of Forestry and Environmental Studies, Yale University, New Haven, CT 06511, USA

Abstract

In tropical forests, rarer species show increased sensitivity to species-specific soil pathogens and more negative effects of conspecific density on seedling survival (NDD). These patterns suggest a connection between ecology and immunity, perhaps because small population size disproportionately reduces genetic diversity of hyperdiverse loci such as immunity genes. In an experiment examining seedling roots from six species in one tropical tree community, we found that smaller populations have reduced amino acid diversity in pathogen resistance (R) genes but not the transcriptome in general. Normalized R gene amino acid diversity varied with local abundance and prior measures of differences in sensitivity to conspecific soil and NDD. After exposure to live soil, species with lower R gene diversity had reduced defence gene induction, more cosusceptibility of maternal cohorts to colonization by potentially pathogenic fungi, reduced root growth arrest (an R gene-mediated response) and their root-associated fungi showed lower induction of self-defence (antioxidants). Local abundance was not related to the ability to induce immune responses when pathogen recognition was bypassed by application of salicylic acid, a phytohormone that activates defence responses downstream of R gene signalling. These initial results support the hypothesis that smaller local tree populations have reduced R gene diversity and recognition-dependent immune responses, along with greater cosusceptibility to species-specific pathogens that may facilitate disease transmission and NDD. Locally rare species may be less able to increase their equilibrium abundance without genetic boosts to defence via immigration of novel R gene alleles from a larger and more diverse regional population.

Keywords: demographic effects, eco-evolutionary feedback, ecological dynamics, forest research plot, fungi, genetic drift, immunity, long-term ecological study, plant–pathogen interaction, polymorphism, root microbiome

Received 8 November 2016; accepted 22 December 2016

Introduction

Determining the causes and consequences of differences in species abundances has long been a focus of theoretical and applied ecology (Preston 1948; Volkov *et al.* 2003). Early studies examining relative abundance

Correspondence: James H. Marden, Fax: 814-865-9131; E-mail: jhm10@psu.edu

within plant communities focused on species' abilities to tolerate local abiotic conditions and compete with co-occurring species for resources (Clausen *et al.* 1940), however, such differences should lead to competitive exclusion rather than stabilization of a diverse community. More recently, interactions with natural enemies have been shown to play a role in shaping plant species relative abundance (Klironomos 2002; Mangan *et al.* 2010; McCarthy-Neumann & Kobe 2010; van der Putten *et al.* 2013) and spatial dynamics (Gilbert *et al.* 1994; Burdon *et al.* 2006; Anderson *et al.* 2010; Brown & Tellier 2011; Laine *et al.* 2011; Moslonka-Lefebvre *et al.* 2011; Jousimo *et al.* 2014). For example, seedlings of locally less common species suffer more than common species from the density and proximity of conspecific neighbours in a diverse tropical forest (Comita *et al.* 2010). Shadehouse and field experiments revealed that this pattern was likely due to higher sensitivity to host-specific soil pathogens (Mangan *et al.* 2010). These findings, together with similar results from temperate plant communities (Klironomos 2002; Johnson *et al.* 2012; van der Putten *et al.* 2013), suggest that host-specific pathogens shape the relative abundance of plant species. Simulations provide theoretical support for the ability of plant–enemy interactions to determine relative abundance when plant species vary in susceptibility to their own enemies, resulting in stronger negative density dependence (NDD) and lower equilibrium population size for more susceptible species (Mangan *et al.* 2010; Chisholm & Muller-Landau 2011; Mack & Bever 2014). Importantly, stronger negative density dependence caused by specialized plant–pathogen interactions leads to the long-term persistence of rare species in the community (Mangan *et al.* 2010; Chisholm & Muller-Landau 2011). However, the mechanisms that cause less abundant plant species to be more susceptible to their own pathogens and therefore experience stronger NDD are unknown.

We hypothesized that the strength of pathogen-mediated NDD is greatest in locally less abundant plant species because founder effects (a form of genetic drift) reduce allelic diversity in a small local population compared with the regional metapopulation (Ellstrand & Elam 1993). Two features of locally reduced genetic diversity may lead to variation in pathogen susceptibility: reduction in the ability to activate defence responses due to general physiological deficiencies caused by inbreeding (Saccheri *et al.* 1998; Du *et al.* 2008; Seeholzer *et al.* 2010; Kariyat *et al.* 2012; Portman *et al.* 2015), and/or reduction in diversity of pathogen resistance (R) genes, which detect pathogens and determine allele-specific activation of defence (Hammond-Kosack & Jones 1997; Laine *et al.* 2011; Dangl *et al.* 2013). The former scenario is not supported by seedling survival data at our study site (Comita *et al.* 2010), which showed no relationship between species abundance in the community and seedling survival in the absence of conspecific neighbours (i.e. the intercept term in the survival model). In the latter scenario involving R genes, the probability of cosusceptibility to particular pathogen strains and transmission among conspecific neighbours increases when plants share the same genotype at loci critical for either complete immunity or quantitative resistance. This hypothesis involves genotype-specific host–pathogen interactions and follows from molecular genetic studies showing that diversification of R genes broadens the spectrum of pathogen recognition capability (Dodds & Rathjen 2010), conveying either complete or partial resistance (Poland *et al.* 2009) to a specific pathogen strain or functionally similar strains (Allen *et al.* 2004; Karasov *et al.* 2014). Hence, reduced genome-wide R gene allelic diversity of neighbouring plants should increase their cosusceptibility to particular strains within a highly diverse and rapidly co-evolving pathogen community and increase their transmission of pathogens and NDD.

R genes comprise a large and diverse set of gene families encoding proteins that perform various roles across multiple layers of plant immunity, including direct recognition and interaction with pathogen molecules (Jia *et al.* 2000) and indirect interaction via detection of host defence-signalling pathways inhibited by pathogen effector molecules (Simonich & Innes 1995). Multiple hormone signalling pathways function downstream of R genes to regulate defence responses, which in turn determine absolute or quantitative immunity (Knepper & Day 2010).

Among the many genes involved in plant defence, R genes should be especially susceptible to reductions in population size due to their unusually high levels of allelic diversity maintained by balancing selection (Bergelson *et al.* 2001; Rose *et al.* 2004; Meyers *et al.* 2005; Bakker *et al.* 2006; Seeholzer *et al.* 2010; Karasov *et al.* 2014). In a neutral model, small populations lose disproportionately more diversity at allele-rich loci than loci containing few alleles (Allendorf 1986). Muirhead (2001) modelled the effects of population structure and mutation rate on loci under balancing selection, concluding that 'the number of alleles maintained in a single deme is expected to decrease as the effective neighbourhood size of the deme decreases and individuals have a more restricted set of potential parents'. This has been confirmed empirically in studies of self-incompatibility genes, another class of plant loci with many alleles maintained by balancing selection. The smallest plant populations have the fewest S-alleles (Glemin *et al.* 2005; Brennan *et al.* 2006) and in the best-studied case a tight positive relationship between allele

number and population size (Young & Pickup 2010) despite strong fitness reductions in the smallest populations caused by frequent pollen incompatibility. Hence, balancing selection in small populations of plants does not overcome the diversity-reducing effect of bottlenecks and drift. In the case of R genes, erosion of diversity in small populations should make them more vulnerable to infection and transmission of their own co-evolving pathogens.

Previous studies of polymorphism at putatively neutral loci have shown that tropical trees tend to have reduced genetic variation in small local populations (Hamrick & Murawski 1991; Finkeldey & Hattemer 2007; Noreen & Webb 2013) and strongly increased seedling-to-adult survival for outcrossed progeny (Hufford & Hamrick 2003). Over a scale of a few tens of metres, tropical trees tend to be related, corresponding with seed dispersal kernels, with most pollen originating within ~100 to 1000 m of the mother tree, and detectable levels of long-distance pollen flow (Loveless 1992; Boshier *et al.* 1995; Hardy *et al.* 2006; Noreen & Webb 2013).

There is ample evidence indicating that pathogens play a strong role in driving density dependence in tree communities. At our study site, wood-decaying fungal incidence and diversity increases in a host density-dependent fashion (Gilbert *et al.* 2002), and damping-off disease in tree seedlings is negatively affected by dispersal distance and positively by seedling density [(Augspurger & Kelly 1984); see (Packer & Clay 2000; Hood *et al.* 2004; Bagchi *et al.* 2010; Reinhart *et al.* 2010) for similar results]. Collectively, these results indicate that seedlings are more vulnerable to disease when establishing around conspecifics, especially when they occur at high density.

Here, we use a transcriptome approach (Wang *et al.* 2009) to examine expression level and polymorphism in genes expressed in seedling roots of six species of tropical trees. We first present seedling root transcriptomes and validate the gene models and estimates of polymorphism. We use those data for an initial test of the hypothesis that smaller local populations have reduced protein-coding diversity in R genes. Finally, we use data from a replicated and controlled experiment to show how normalized amino acid coding diversity is associated with a number of immune-related phenotypes: NDD, induction of defence genes downstream of R genes following exposure to live soil or to a phytohormone that bypasses R genes, cosusceptibility of maternal cohorts to colonization by pathogenic fungi, self-defence responses by root-associated fungi, and seedling root growth arrest following live soil inoculation. We end by discussing how these initial findings (to be interpreted cautiously based on the small number of species sampled) may inform our understanding of the eco-evolutionary context of plant–pathogen interactions and community ecology.

Methods

Approach and rationale

We selected six focal species (Tables 1 and 2) representing a broad range of plant orders, families and relative abundances [ranked 10–97th by decreasing basal area of 180 tree species sampled in (Comita *et al.* 2010)] on the 50-ha Forest Dynamics Plot on Barro Colorado Island (BCI), Panama (Hubbell *et al.* 1999; Condit *et al.* 2000, 2012). All but one of these species was previously tested for effects of host-specific soil pathogens (Mangan *et al.* 2010) and all were included in a previous study of conspecific NDD in seedlings occurring naturally in the BCI forest (Comita *et al.* 2010). Using surface-sterilized seeds collected from the ground beneath canopies of five fruiting trees of each focal species (Figs 1 and S1 in Appendix S1), we first germinated cohorts of 15 seedlings per parent (75 seedlings per species) in sterile soil in trays and then transplanted them to individual pots after developing their first set of primary leaves. Pots contained equal amounts of the same homogenized and autoclaved field soil. Potted seedlings were assigned randomly among three soil treatments. Five seedlings of each cohort were untreated controls; five received a salicylic acid (SA) root drench 48 h prior to harvesting (0.3 g of SA diluted in 75 mL of distilled water); five were given 200 mL of a live soil inoculum 1 week prior to harvesting. The live soil inoculum contained equal portions of rhizosphere soil collected under a single adult of each our study species (Fig. S1 in Appendix S1, Table S1 in Appendix S2; one species, *Beilschmiedia pendula*, had too few fallen seeds not killed by insects to permit the live soil treatment). This inoculum was collected and fully homogenized (across source adults) immediately prior to inoculating seedlings.

Defence gene activation after live soil treatment depends on the ability of a plant to recognize pathogens or their effector proteins, whereas SA (a natural plant defence-signalling hormone) is taken up by plants and activates defences downstream of R genes (Alvarez 2000), including expression of pathogenesis-related (PR) genes (Vlot *et al.* 2009). Other defence-related phytohormones (e.g. jasmonic acid; ethylene) were not tested to keep the experiment tractable. SA is commonly triggered by microbes (Glazebrook 2005) and, although little studied in roots, can affect plant–microbe interactions in roots independently of any direct effect of the compound on microbes (Medina *et al.* 2003). Hence, SA treatment served experimentally to bypass pathogen recognition and directly activate antimicrobial

Table 1 Study species and assembly statistics for nonredundant gene models
[Correction added on 20 April, 2017, after first online publication: The values in Median length have been modified at the request of the authors]

Species (family)	*N* gene models	Median length, nt (aa)	Median read depth
Beilschmiedia pendula (Lauraceae)	25 747	1144 (248)	52
Virola surinamensis (Myristicaceae)	26 216	1249 (251)	33
Brosimum alicastrum (Moraceae)	19 689	1203 (252)	54
Eugenia nesiotica (Myrtaceae)	24 075	1514 (319)	49
Lacmellea panamensis (Apocynaceae)	24 415	1491 (298)	53
Dipteryx oleifera (Fabaceae)	30 251	1024 (209)	61

defences. The live soil and SA treatments therefore illuminated two physiologically distinct traits: (i) recognition-dependent inducible defence response to diverse soil microbes in live soil and (ii) recognition-independent SA-inducible defence.

Following treatments, we uprooted the 2-month-old seedlings, washed the roots free of soil and flash-froze them for RNA isolations, after which we pooled equal amounts of total RNA from each individual within combinations of parent tree ($N = 5$) and seedling treatment ($N = 3$) for sequencing (15 distinctly labelled pools per species, each sequenced in one lane [Illumina HiSeq 2500] using 2×150 nucleotide paired reads; Fig. 1).

We assembled transcriptomes from all trimmed and quality-filtered sequences across treatments using Trinity (Grabherr *et al*. 2011; Haas *et al*. 2013) (Tables 1 and 2). The transcriptome assemblies initially contained about 150 000 plant gene models per species, with redundancy caused by mRNA features such as alternative splicing and retained introns. To overcome redundancy and potential artefacts, we used a well-validated orthology analysis (Amborella Genome 2013; Honaas *et al*. 2013; Ming *et al*. 2013; Wickett *et al*. 2014) that employs a hidden Markov matrix procedure to detect deep homology and select the best representative gene model from each

Table 2 Polymorphism in nonredundant gene models from root transcriptomes of tropical trees. Data outside parentheses refer to the transcriptome exclusive of R genes; data in parentheses refer to R genes. (site-frequency spectra are shown in Fig. S6a in Appendix S2)

Species	*N* monomorphic	*N* polymorphic	Median *pN*	Median *pS*	Median *pN*/*pS*
Beilschmiedia (Bl)	16 556 (241)	8795 (155)	0.004 (0.041)	0.009 (0.042)	0.52 (0.84)
Virola (Vr)	18 930 (152)	7016 (118)	0.003 (0.036)	0.007 (0.058)	0.54 (0.73)
Brosimum (Br)	12 668 (234)	6624 (163)	0.004 (0.019)	0.009 (0.026)	0.49 (0.73)
Eugenia (Eu)	15 502 (435)	7780 (258)	0.004 (0.022)	0.009 (0.033)	0.50 (0.69)
Lacmellea (La)	17 454 (115)	6767 (79)	0.003 (0.024)	0.006 (0.048)	0.58 (0.68)
Dipteryx (Di)	18 199 (171)	11 785 (96)	0.007 (0.026)	0.018 (0.043)	0.46 (0.69)

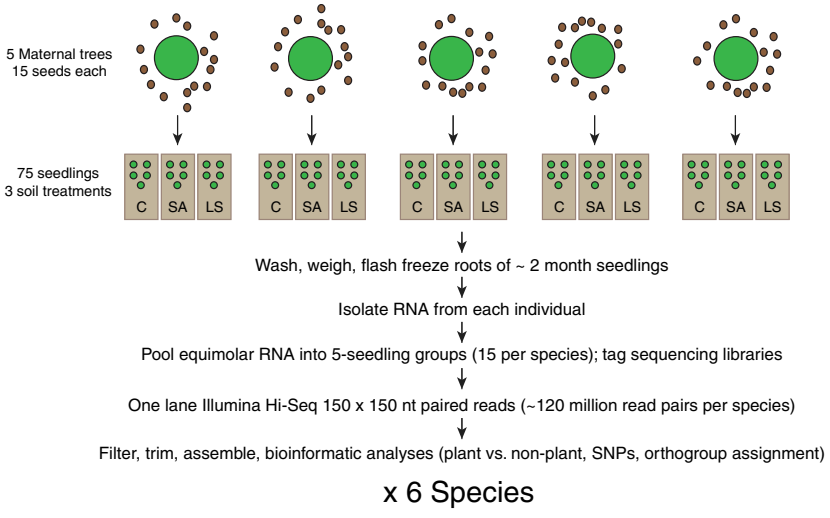

Fig. 1 Methods schematic showing the collection of maternal seed cohorts, random assignment of seeds to treatment groups within maternal cohorts, and downstream steps for transcriptome sequencing and analyses. Treatment abbreviations: sterile soil control: C; salicylic acid root soak: SA; live soil inoculation: LS. This protocol was repeated for six species, except *Beilschmiedia* for which we had only ten viable seeds per cohort and did not perform the live soil treatment.

orthogroup (narrowly defined gene families) from a global set of well-annotated orthogroups produced from 22 representative high-quality plant genomes. Only these highly filtered gene models (N = ~20–30K per species; Table 1) were retained for further analyses, including estimates of the completeness of transcriptome coverage (Table S2a–c in Appendix S2), well-supported single nucleotide (SNP) and insertion–deletion polymorphisms (Koboldt *et al.* 2009), and treatment-specific gene expression levels (Wang *et al.* 2009).

We estimated the normalized abundance of polymorphism in protein-coding regions from the number of nonsynonymous (amino acid changing) SNPs per nonsynonymous site (pN), synonymous SNPs per synonymous site (pS) and the ratio pN/pS (Table 2). pN/pS measures the abundance of protein-coding variation (pN) relative to neutral divergence (pS), which varies with allele age and across genomic locations. Compared with the rest of the genome, R genes have increased levels of nonsynonymous variation (pN) caused by diversifying selection (Michelmore & Meyers 1998; Bakker *et al.* 2006; Yang *et al.* 2006), particularly in protein regions that interact with pathogens. Different R gene loci have alleles of widely different age, some newly arisen and polymorphic for protein variants but with little synonymous divergence, whereas other loci contain older alleles that have accumulated synonymous polymorphisms and are more broadly diverged in amino acid sequence (Bakker *et al.* 2006; Yang *et al.* 2006). Protein amino acid divergence generally comprises a mix of neutral and functionally important differences (Shihab *et al.* 2013), so using only pN as a measure of diversity would bias towards the oldest and most divergent alleles, with additional detectability biases caused by variation in expression level and read depth in the transcriptome assembly. Hence, our analyses use pN/pS to normalize for the neutral divergence of alleles (Hughes & Nei 1988), assign weight to sequences according to their excess of nonsynonymous polymorphism and control for sequence depth-dependent SNP detectability.

We used root RNA from a single individual of one of the focal species (*Dipteryx*) to perform PacBio single read sequencing to verify the Illumina+Trinity gene assembly in general and R genes in particular.

To validate our approach, we simulated the effect of a small number of founders on R gene allelic diversity using published data on *Arabidopsis thaliana* (see Supporting information, Fig. S3a in Appendix S2), which confirmed that pN/pS of the most diversified R gene loci is particularly reduced in the founding of a small population.

For hypotheses with a priori directional predictions, we used one-tailed statistical tests. A detailed description of methods is contained in the Supporting information.

Results

R gene allelic diversity in seedling root transcriptomes

Assembled transcriptomes from seedling roots yielded ~20 to 30 thousand nonredundant gene models in each of the six focal species (Table 1). Deep sequencing coverage (median read depth ~50; Table 2) enabled characterization of polymorphism in ~7 to 10 thousand genes per species, including substantial numbers (79–258) of R genes (Table 2). In all six species, R genes showed highly elevated polymorphism (pN, pS and pN/pS; Table 2, Fig. 2) compared with the remainder of the transcriptomes, consistent with long-term diversifying selection and accumulation of old alleles with amino acid diversity important for immune recognition and response to microbial pathogens (Bergelson *et al.* 2001; Allen *et al.* 2004; Rose *et al.* 2004; Meyers *et al.* 2005; Bakker *et al.* 2006; Seeholzer *et al.* 2010; Bakker *et al.* 2006; Karasov *et al.* 2014).

Verification of the assembly and gene models

R genes have a complex evolutionary history and are widely thought to be difficult to assemble correctly from short-read sequences. That conjecture was recently tested using the *Arabidopsis thaliana* genome to assess *de novo* assemblies of *A. thaliana* transcriptomes from Illumina sequencing. Those assemblies showed no examples of chimeric assembly of R genes in the postorthology filtered gene models (Honaas *et al.* 2016). We used the same Trinity assembly protocol and orthology filter in the present experiment, so potential chimera are unlikely to be retained as best representatives of narrowly defined orthogroups (see methods). Even so, we sought to verify gene models by performing single-molecule (PacBio) sequencing of one of the subject species (*Dipteryx*). Those data showed strong support for the assembly, expression level estimates, and polymorphism metrics. Notably, the subset of gene models from the Illumina assembly that were contained fully within a PacBio single-molecule read (i.e. fully verified; N = 8876, Table S4a,b in Appendix S2) had polymorphism metrics, including median pN/pS of R genes, nearly unchanged compared with the entire Illumina-based assembly (Table S4d in Appendix S2). This result is consistent with Honaas *et al.* (2016) and the observation that in *Arabidopsis thaliana*, the rapid divergence of R gene paralogs makes them more difficult to align than to tell apart (Mondragon-Palomino *et al.* 2002). We acknowledge the possibility that some misassembly of very recently diverged paralogs may have slipped through and remain present in our analysed data. Note, however, that all of our analyses of genetic data use

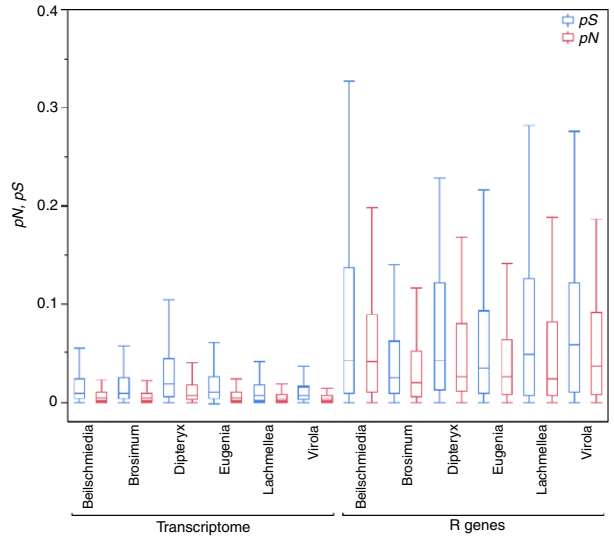

Fig. 2 Polymorphism in resistance (R) genes compared to the remainder of the transcriptome in roots of six species of tropical tree seedlings. Elevated *pN* and *pS* in R genes reflect long-term diversifying selection and the presence of old alleles, likely maintained by frequency-dependent selection imposed by interactions with rapidly evolving proteins of pathogens.

medians, which are highly robust to sporadic measurement errors, and there is no reason to suspect that assembly errors would be systematically biased in relation to the size of local populations or the phenotypes we measured.

Polymorphism metrics for R genes and the remainder of the transcriptome were concordant with whole genome-guided estimates from rice, *Picea,* and *Medicago* (see Fig. S4a and associated text in Appendix S2). There was no evidence of codon bias in any of the focal tree species (Table S5a in Appendix S2) and hence no evidence that *pS* varied in a species-specific manner due to historical differences in weak selection. Median *pN*/*pS* was replicable among cohorts (Fig. S5a in Appendix S2). Estimates of R gene polymorphism were little affected by lineage-specific differences in the evolutionary history of R gene families (Fig. S5c,d in Appendix S2), as evidenced by the nearly unchanged median *pN*/*pS* estimates of R genes when calculated from the single largest R gene orthogroup (canonical R genes containing NB-ARC and LRR domains; Fig. S5c in Appendix S2). As previously observed for protein-coding regions involved in direct interactions with pathogen effector molecules (Bakker *et al.* 2006), nonsynonymous polymorphism was significantly higher within leucine-rich repeat (LRR) domains of R genes ($P < 0.0001$; Fig. S5e in Appendix S2). Median *pN* within those LRR motifs was closely related to median *pN*/*pS* across all R gene regions (Fig. S5f in Appendix S2). Based on these analyses, we concluded that the assembly, gene expression measures and estimates of polymorphism were robust and that our metrics of R gene polymorphism were a valid proxy for functional differences related to plant–pathogen interactions.

Reduction in R gene protein-coding diversity in locally less abundant species

Median *pN*/*pS* in R genes of tropical trees was strongly related to local population size (Fig. 3A; one-tailed $P = 0.009$), but there was no such association between population size and *pN*/*pS* in the remainder of the transcriptome (one-tailed $P = 0.31$; Fig. 3B). Non-normalized amino acid diversity (*pN*) of R genes also increased with local population size (Fig. 3C; one-tailed $P = 0.01$) but that association was not present in the remainder of the transcriptome (Fig. 3D). Neutral diversity (*pS*) did not differ with local population size for either R genes or the remainder of the transcriptome ($P = 0.54$ and 0.42; here two-tailed in the absence of clear directional predictions).

Previous research on this plot (Hamrick & Murawski 1990, 1991) found a tendency for reduced polymorphism at allozyme loci in less abundant tree species. Our results replicated that pattern for nonsynonymous polymorphism in R genes but not in the remainder of transcriptome. This was likely due to the ~fivefold lower range in median *pN* among species in the transcriptome exclusive of R genes (Table 2). Low values and differences in among-species abundance of polymorphism applied also to synonymous variation, and hence, there was much less leverage for detecting species differences and population size effects on polymorphism in the transcriptome exclusive of R genes.

Subsampling to examine R gene protein-coding diversity in relation to other functional groups

We subsampled the transcriptome to test the possibility that lumping genes into two broad categories (R genes vs. the remaining transcriptome) may create a deceptive result or hide important associations with population size. In linear regressions testing the relationship between population size and median *pN*/*pS* for all well-represented protein families ($N = 133$ protein families with at least 10 gene models), R genes were in the top 2.5% for strength of association (Fig. S3b in Appendix S2). Only two protein families showed a stronger association; one was ribosomal protein L4/L1e, represented by 171 gene models but only 18 containing polymorphism. With such sparse data, this is most likely a false-positive association. The other family comprised proteins containing a thioredoxin domain, much better represented by 231 gene models (101 with

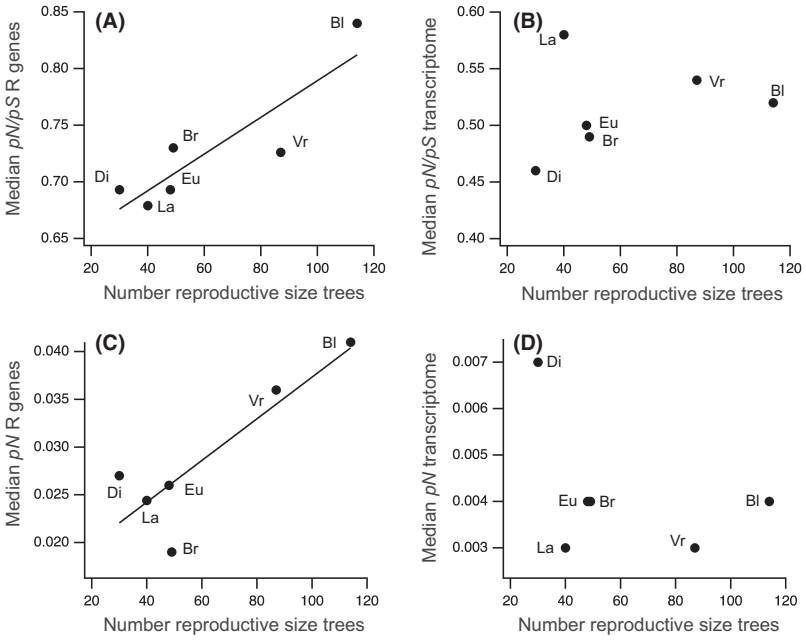

Fig. 3 Local population size [number of reproductive-size adult trees of each species in the 2010 census of the BCI 50-ha plot (Condit *et al.* 2012)] is positively associated with the relative (A) and absolute (C) abundance of nonsynonymous polymorphism (*pN*/*pS*) in R genes but not in the remainder of the transcriptome (B, D). Species abbreviations on plots refer to: Bl: *Beilschmiedia pendula* (Lauraceae); Vr: *Virola surinamensis* (Myristicaceae); Br: *Brosimum alicastrum* (Moraceae); Eu: *Eugenia nesiotica* (Myrtaceae); La: *Lacmellea panamensis* (Apocynaceae); Di: *Dipteryx oleifera* (Fabaceae).

polymorphism, 29 with pN/pS > 1). A flurry of recent studies demonstrate interaction of thioredoxin proteins with pathogen effector molecules, with outcomes that affect immune signalling, regulation of a pathogen-favourable redox environment, and susceptibility/resistance (Lorang *et al.* 2012; Yang *et al.* 2014; Carella *et al.* 2016; Faheem *et al.* 2016; Floryszak-Wieczorek *et al.* 2016; Fujiwara *et al.* 2016; Luna-Rivero *et al.* 2016; Mukaihara *et al.* 2016; Popa *et al.* 2016; Sang *et al.* 2016). Thioredoxin domain proteins are therefore likely to have undergone diversifying selection in a fashion similar to R genes and appear to contain an elevated level of amino acid polymorphism that is similarly underrepresented in small local populations. In summary, the relationship between normalized protein-coding diversity of R genes and local population size was extreme compared with individual protein families but was replicated in one protein family that similarly mediates immunity.

R gene protein-coding diversity and negative density dependence

A potential ecological outcome of reduced R gene diversity is that neighbouring plants possessing more similar R gene genotypes may be more likely to transmit and succumb to particular pathogen genotypes. If so, there should be an inverse relationship between the local population R gene protein-coding diversity and the negative impact of conspecific neighbours. To test this hypothesis, we compared median *pN*/*pS* of R genes with two independent measures of NDD for these species from the same study site (Comita *et al.* 2010; Mangan *et al.* 2010). Species with reduced normalized protein-coding diversity in R genes suffered more negative effects when exposed to soil collected beneath trees of the same species (i.e. plant–soil feedback) in a shadehouse experiment (Fig. 4A; one-tailed $P = 0.045$) and had stronger conspecific negative density-dependent mortality of naturally occurring seedlings in the BCI forest (Fig. 4C; one-tailed $P = 0.041$). Using *pN* instead of *pN*/*pS* produced a weaker relationship for plant–soil feedback (one-tailed $P = 0.12$) but a similar result for density-dependent seedling mortality (one-tailed $P = 0.04$). In the remainder of the transcriptome, there were weak or no associations between these measures of negative conspecific effects and median *pN*/*pS* (Fig. 4B,D; one-tailed $P > 0.5$ in both cases), *pN* (one-tailed $P = 0.12$, 0.48), or *pS* (one-tailed $P = 0.10$, 0.47, respectively; plots not shown) and hence little indication that differences in overall genetic diversity explain species differences in density dependent mortality and plant–soil feedback.

Increased pathogen cosusceptibility in locally less abundant species

Stronger negative impacts of conspecific soil and conspecific neighbours suggest that individuals of species with reduced R gene protein-coding diversity are indeed more cosusceptible to particular microbial pathogens. To test experimentally for differences in cosusceptibility, we used transcriptome data to quantify variability among maternal seedling cohorts in

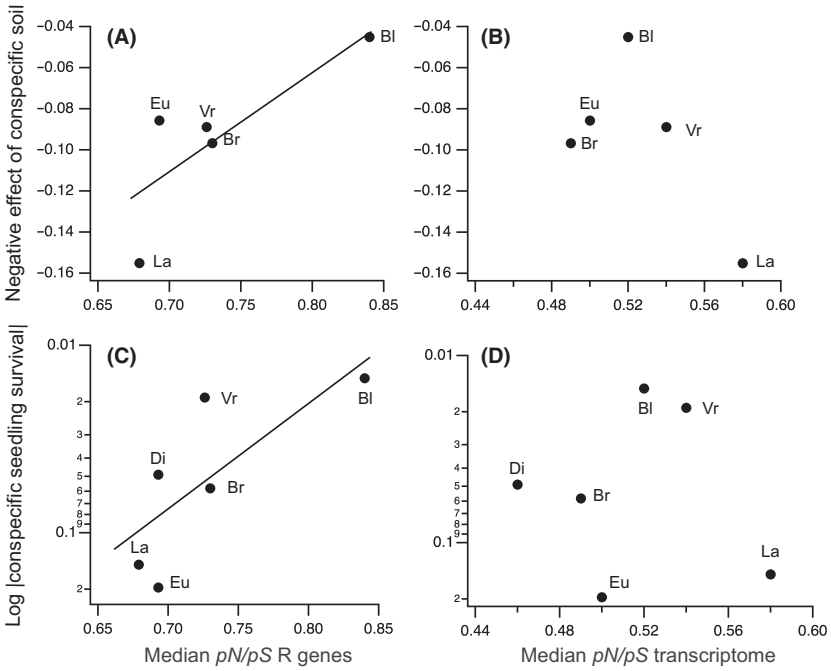

Fig. 4 Association of tree species genetic diversity with independent measures of the negative effect on growth caused by (A, B) inoculation of conspecific soil [vertical axis shows data from (Mangan *et al.* 2010)] in a shade-house experiment (that did not include *Dipteryx*) and (C, D) negative conspecific effects (shown here as increasingly positive numbers to accommodate log transformation) on survival of naturally occurring seedlings [vertical axis shows data from (Comita *et al.* 2010)]. All data shown here involve material from the same tropical forest site. Note that these indices of NDD are associated with relative abundance of nonsynonymous polymorphism in R genes (A, C), but not in the remainder of the transcriptome (B, D).

colonization of roots by fungi after exposure to a standardized mixture of live soil. This soil was collected in such a way that it contained the diverse community of microbes naturally associated with our study species (Fig. S1 in Appendix S1; Table S7a,b in Appendix S2).

There were 35 putatively pathogenic fungal taxa that passed the threshold for gene number and mean increase following live soil inoculation (Table S7a in Appendix S2); these showed fairly little overlap among tree species. Among-cohort variability in colonization by these potentially pathogenic fungi was significantly different among tree species ($P = 0.02$; Table S7c in Appendix S2), and decreased significantly with decreasing *pN*/*pS* of R genes ($R^2 = 0.65$, one-tailed $P = 0.05$; Fig. 5A). Such a relationship is expected when different maternal cohorts of seedlings share more of the same R gene alleles and are therefore sensitive to the same pathogens (presumably strains contained within these OTUs). Among-cohort variability in colonization by saprophytic fungi (Table S7b in Appendix S2) also differed significantly among tree species ($P = 0.002$; Table S7d in Appendix S2), but in a manner unrelated to *pN*/*pS* of R genes ($P = 0.95$, here a two-tailed test because there is no directional prediction for the way tree immune gene diversity should affect fungi feeding extracellularly on dead tissue).

We did not design the experiment to measure tree species differences in overall microbial diversity or level of colonization. Differences in plant RNA abundance and/or the surface/volume ratio of roots should affect the relative abundance of plant vs. nonplant RNA in the samples and therefore the detectability of fungi. For this reason, we have not attempted analyses of total fungal colonization or diversity.

Variation in ability to activate immune responses

As a consequence of having reduced diversity of R genes, tree species that are locally less abundant should show reduced ability to recognize particular pathogen strains and mount physiological defences. To test for such deficits arising from R gene diversity while also testing the alternative hypothesis that smaller local populations have reduced physiological capability to mount responses (due to potential factors such as inbreeding or reduced adaptation to the local environment), we applied a SA root soak treatment to examine expression of a large and diverse set of pathogenesis-related (PR) genes (e.g. chitinases, peroxidases; Table S8a,b in Appendix S2) that act downstream of R gene detection and attack microbes [(Shah 2003; Loon *et al.* 2006); a foot-soldier role as opposed to the sentinel role of R genes]. Compared with untreated controls, PR gene expression after the SA root soak treatment was significantly upregulated ($P < 0.0001$; Table S8c in Appendix S2), but in a manner unrelated to transcriptome-wide genetic diversity, local population size or NDD (Table S8d–g in Appendix S2). We therefore found no evidence that locally rare species were less able to activate defence genes in response to an exogenously applied defence-signalling hormone (SA) that acts downstream of R genes. This finding

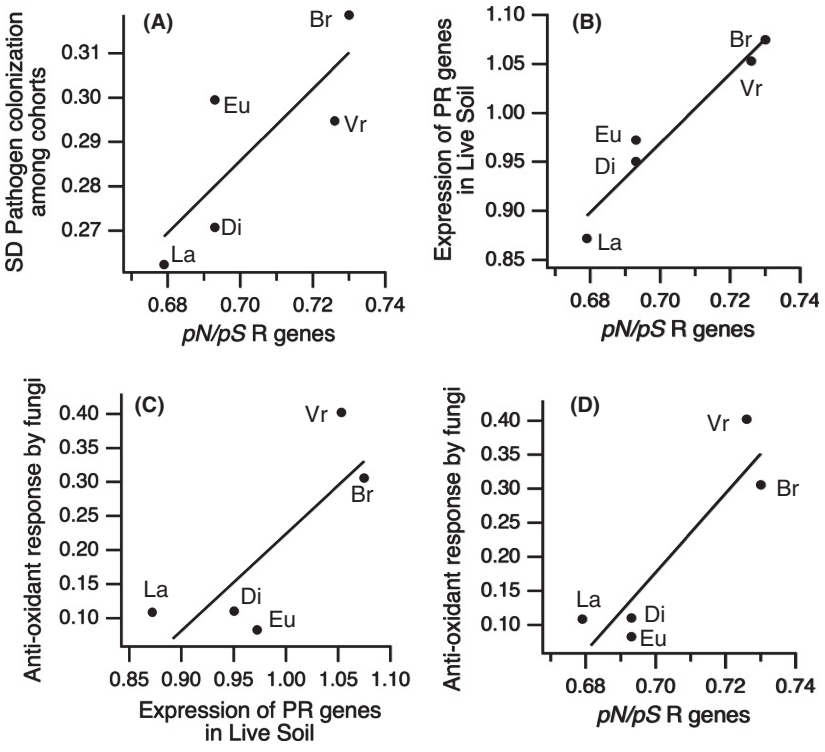

Fig. 5 (A) Association between relative abundance of nonsynoymous polymorphism (median *pN/pS*) of R genes and the cosusceptibility of maternal seedling cohorts to colonization by pathogenic fungi (Table S8c in Appendix S2). (B) Expression of pathogenesis-related (PR) genes ($N = 592$; Table S9a,i in Appendix S2) in roots after live soil inoculation, in relation to median *pN/pS* of R genes. (C) Fungal self-defence gene expression (antioxidant enzymes, $N = 799$; Table S10a–c in Appendix S2) in relation to the expression of host tree PR genes after live soil inoculation and (D) the median *pN/pS* of host tree R genes. As explained in the main text, *Beilschmiedia pendula* is missing from this analysis because it had too few fallen seeds not killed by insects to permit the live soil treatment.

strongly suggests that rare species were not deficient in their physiological ability to mount inducible defence responses, as can result from inbreeding (Du *et al.* 2008) or reduced adaptation to the local environment. Tree species did, however, differ in defence responses after exposure to live soil. When we controlled for response to SA (each gene's potential to be upregulated), the protein-coding diversity of R genes (*pN/pS*) explained variation in the expression of PR genes (Table S8i in Appendix S2) after live soil inoculation (one-tailed $P = 0.009$; species means shown in Fig. 5B). Increased R gene allelic diversity was therefore positively associated with the ability to detect diverse pathogenic fungi present in soil and activate inducible defence responses.

Defence gene activation affected root-associated microbes

To verify that species differences in PR gene expression affected plant–microbe interactions, we quantified countermeasures enacted by root-associated fungi. Fungi defend themselves (Molina & Kahmann 2007; Chung 2012) against peroxides and other oxidative plant defences by upregulating antioxidant genes (Chen *et al.* 2003) (Table S9a,b in Appendix S2). When challenged with live soil, root-associated fungi in tree species with higher PR gene expression and higher normalized protein-coding diversity of R genes had higher relative expression of antioxidant genes (median expression level of fungal antioxidant genes vs. all other detected fungal genes, Table S9c in Appendix S2; one-tailed $P = 0.047$, 0.02, respectively; Fig. 5C–D). This result indicates that plants better able to detect pathogens and mount defence responses provoked more self-defence in their root-associated fungi.

Variation in a whole-plant immune-related phenotype: root growth arrest

It was recently discovered that an R gene in *Arabidopsis* triggers, in an allele-specific fashion, an SA-independent root growth arrest in response to soil-borne pathogens (Kim *et al.* 2012; Kunz *et al.* 2016). This is presumably a temporary growth arrest while the immune system is activated and fighting an infection. To determine whether tropical tree seedlings also undergo an acute root growth arrest, we examined the final root mass of seedlings from the three different soil treatments (Table S10a in Appendix S2). Roots from control and SA-treated soil did not differ in size ($P = 0.53$), but roots from the live soil treatment where significantly smaller (Table S10a,b in Appendix S2; $P = 0.04$) and the effect varied significantly among species (a species*soil treatment interaction, $P = 0.035$). As the live soil was applied 1 week prior to harvest, growth arrest during

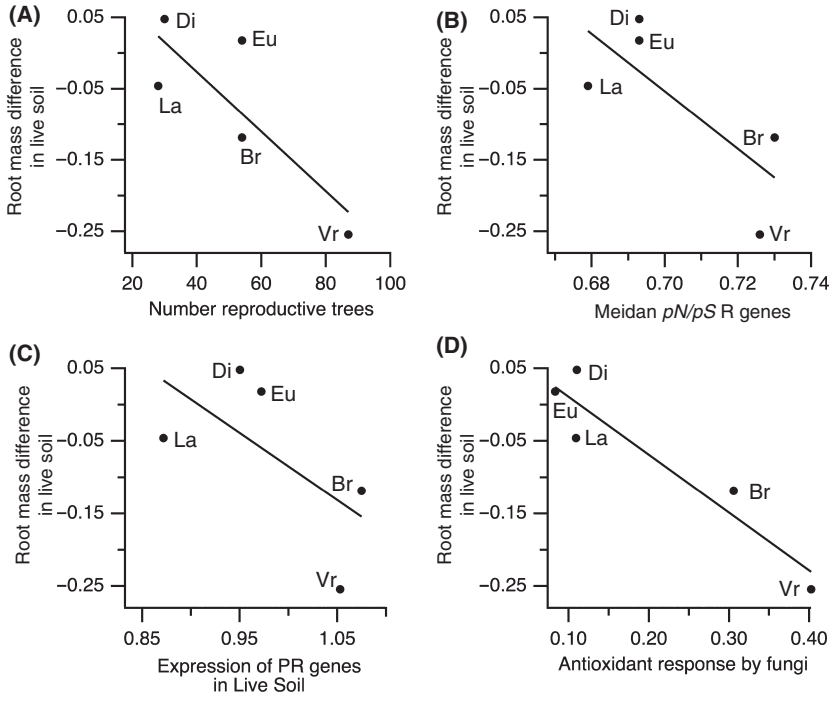

Fig. 6 Magnitude of root growth arrest (estimated from the difference in mean $\log_{10}$-transformed seedling root mass in live soil treatments compared with the mean of control and SA treatments (Table S11a,b in Appendix S2) compared with (A) local population size, (B) median-normalized amino acid coding diversity of R genes, (C) induction of pathogen response (PR) genes in response to live soil inoculation, and (D) induction of self-defence (antioxidant genes) by root-associated fungi.

this time was apparently sufficient to measurably reduce the final mass of the live soil-treated roots. Given the significant species*treatment interaction, we examined this phenotype from the perspective of population size, R gene diversity and molecular-level immune responses. As predicted based on the known R gene mediation of the phenotype, the magnitude of root growth arrest (Fig. 6) increased with population size (one-tailed $P = 0.02$), R gene diversity (one-tailed $P = 0.07$), PR gene induction by live soil (one-tailed $P = 0.04$) and induction of fungal self-defence (antioxidant gene expression) following live soil treatment (one-tailed $P = 0.008$). Hence, tree species that were more abundant, R gene diverse and immunologically responsive showed correspondingly more extreme root growth arrest following exposure to live soil, consistent with the hypothesis that those species have a greater ability to detect and respond to diverse microbes.

Potentially additive effects of R gene number

These focal tree species varied widely in their number and orthogroup distribution of R genes (Table 2), apparently as a result of evolutionarily independent expansions and contractions of different R gene families (Fig. S5d in Appendix S2). In addition to allelic diversity, R gene number (Michelmore & Meyers 1998) may be another axis for variation in immune function [along with epistatic effects of alleles at multiple loci (Saucet *et al.* 2015)]. Our present sample has limited power for testing species-level effects of multiple independent variables, but there are hints of support for the hypothesis that R gene protein-coding diversity and gene number have additive effects. Negative soil feedback, cosusceptibility to pathogens and PR gene expression after live soil inoculation had marginal associations with R gene number in models that included normalized R gene amino acid diversity (Table S11 in Appendix S2). Results from additional species will be required to rigorously determine whether R gene coding diversity combines with R gene number to affect pathogen resistance and NDD.

Discussion

Seedlings from local populations of a taxonomically diverse sample of tree species on the BCI 50-ha Forest Dynamics Plot in Panama had reduced levels of nonsynonymous polymorphism in R genes and differed in a number of immune-related traits. In controlled and replicated experiments, maternal cohorts of species with reduced protein-coding diversity of R genes (normalized for neutral divergence) were more cosusceptible to colonization by putatively pathogenic fungi. Higher cosusceptibility is likely to increase the density-dependent transmission of pathogens among neighbouring seedlings. In support of that hypothesis, we found that normalized protein-coding diversity of R genes also explained prior results showing that rarer species in this community have stronger negative effects of

conspecific soil (Mangan *et al.* 2010) and conspecific density-dependent seedling mortality (Comita *et al.* 2010).

Species differences in cosusceptibility covaried with a other immune-related traits. One of these was the induction of pathogenesis-related (PR) genes, which respond to signalling initiated by R genes and affect the outcome of plant–pathogen interactions (Loon *et al.* 2006). In roots exposed to live soil, PR gene induction was positively related to the coding diversity of R genes, an association that disappeared when we bypassed R gene function experimentally by applying salicylic acid (SA), a hormone that acts downstream of R genes to activate PR gene expression (Shah 2003). Upregulated PR genes included peroxides, so to verify effects on plant–microbe interactions we examined the induction of antioxidant self-defence genes in host-associated fungi after roots were exposed to live soil. Fungal antioxidant genes were upregulated in proportion to both soil-induced PR gene expression and R gene coding diversity. An R gene-mediated morphological phenotype, root growth arrest (Kim *et al.* 2012; Kunz *et al.* 2016), was also tightly associated with R gene coding diversity, and induction of both PR genes after live soil inoculation and antioxidant gene expression in root-associated fungi. Together, these results connect the coding diversity of R genes with ability of tree seedlings to mount immune responses and defend against potentially pathogenic soil microbes and provide a coherent mechanistic hypothesis to explain the relationship between abundance in the community and pathogen-mediated negative effects from conspecifics.

If reduced R gene protein-coding diversity in smaller local populations is a general characteristic of tropical tree communities, the relationship may arise as follows. Long-term diversifying selection caused by co-evolution with species-specific pathogens causes R gene loci to be unusually diverse (Rose *et al.* 2004; Bakker *et al.* 2006), but founding and drift in small local populations causes an especially strong stochastic loss of diversity in the most allele-rich genes (Allendorf 1986). This outcome has previously been observed for self-incompatibility genes, despite strong reductions in fitness (Young & Pickup 2010), and is further supported by our simulation of a random downsample from a large collection of R gene haplotypes (Fig. S3a in Appendix S2).

Both the absolute (*pN*) and normalized (*pN*/*pS*) protein-coding diversity of R genes increased with local population size. In evolutionary biology, changes in *pN*/*pS* are often interpreted as variation in the strength of purifying selection. However, protein-coding polymorphism in R genes affects immune function and is subject to diversifying selection. Also, unlike small populations on, for example oceanic islands, there is no evidence or reason to expect that these tree species have long existed regionally at relative abundances similar to those presently observed on the BCI Forest Dynamics Plot. The isthmus of Panama was not completely formed until 3 mya (O'Dea *et al.* 2016) and more recent glaciation affected climate and most likely the composition of tree communities (Bush & Colinvaux 1990). Hence, for reasons involving both the ecological setting and the stronger effect that small population size has on allelic representation at diversified vs. ordinary loci, long-term evolutionary change need not play any role in these results, other than setting the stage by creating highly diversified sets of alleles at R gene loci.

Although our results are supported by multiple lines of evidence, they are based on a sample size of six species (further limited in live soil treatments to five species). Furthermore, associations relating local population size and R gene diversity to immune-related phenotypes were not overwhelmingly strong (typically one-tailed P values ~ 0.05). Note, however, that the two-tailed probability of observing, by chance, a consistent directional association between local population size, R gene coding diversity and six independently measured immune-related phenotypes is $0.5^7 = 0.008$. Hence, the aggregated results are strong within this sample, but data from additional species will be required to confirm a general relationship between local population size of trees and the ability to respond to and resist diverse pathogens. A previous study of genetic diversity in seedlings from eight species on the BCI Forest Dynamics Plot, including two of the same species sampled here, found similar results: allelic diversity increased with local abundance, and 'maternal trees of species in high densities receive pollen from several individuals, while low-density species get a larger proportion of their pollen from a few individuals' (Hamrick & Murawski 1990). This concordance with the present results hints at generality, as does the community-wide relationship between abundance and NDD (Comita *et al.* 2010). Our sample of 5–6 tree species is equal to the number of mammal species ($N = 5$) initially shown to suffer increased infectious disease severity after a population bottleneck (O'Brien & Evermann 1988) and used commonly to illustrate ecological effects of reduced immunological diversity.

Connecting plant immunity, population size, genetic diversity and NDD

Here, we place our results in the context of landscape and community ecology to explicitly point out ecological implications and discuss how conclusions from this initial study can be tested in future work. We found that small local populations with limited connectivity to

a regional metapopulation had reduced R gene allelic diversity, decreased defence gene induction downstream of R gene surveillance and increased cosusceptibility of related neighbouring plants to species-specific fungal pathogens. Greater cosusceptibility of less abundant species to their own co-evolving pathogens was associated with stronger negative effects of conspecific soil and NDD, which, according to prior studies, causes lower equilibrium abundances in the community (Mangan *et al.* 2010; Chisholm & Muller-Landau 2011; Yenni *et al.* 2012; Mack & Bever 2014). Importantly, strong stabilizing effects of NDD due to interactions between plants and their specialized pathogens can prevent the local extinction of rare species. The likelihood of a species increasing its local equilibrium abundance should be favoured by immigration of additional R gene alleles through long-distance dispersal of pollen or seeds from the larger regional pool (Dlugosch & Parker 2008), a possibility that may be restricted in the case of regionally rare or uncommon species. For most species, local NDD should be related to local abundance rather than being a species-level trait, making local communities unique and potentially dynamic in terms of species differences in pathogen transmission and ability to tolerate high conspecific density. In other words, a locally rare species may be more common in other areas of its range, with different levels of R gene polymorphism and NDD in different-sized local populations. Herein lies a critical test of the hypothesis; the species examined here should differ in predictable ways when sampled in other locations where they differ in abundance in the local community.

Comparable results in other systems

Reduction in genetic diversity important for pathogen recognition can in principle occur in any wild plant species with spatial genetic structure and small local populations. Comparable genome-scale ecological studies are not available for comparison, but two recent landscape-scale studies of variation in plant–pathogen interactions support this conclusion. In the temperate herb *Plantago*, local populations are resistant to some strains of a fungus (collected across a metapopulation), but susceptible to others, and the most isolated plant populations have the highest average susceptibility (Jousimo *et al.* 2014). Note that this is different from uniformly reduced disease resistance as might be expected from inbreeding (Hoffman *et al.* 2014) while matching exactly what we would predict from the present study, as the least-connected populations (fewest immigrants) should undergo the strongest erosion of R gene allelic diversity at the time of founding and have a more restricted set of pollen donors in subsequent generations. Similarly, in two subtropical tree species, survival of seedlings exposed to soil from different conspecific populations increased with spatial and genetic distance, but those effects disappeared after soil treatment with a fungicide (Liu *et al.* 2012). We suggest that reduced R gene allelic diversity of isolated *Plantago* populations and reduced R gene allelic similarity between more distant tree populations may underlie the observed susceptibility differences in those studies.

A wealth of knowledge from agricultural systems provide further support: when crop plants with low R gene diversity are grown at high density in monospecific stands, intensive chemical inputs, including antifungals, are needed to counteract rapid pathogen transmission and prevent crop loss (Dangl *et al.* 2013), but mixed cropping of different genotypes of a single crop with different sensitivities to a pathogen can result in complete control and maintenance of health and yield (Mundt 2002) [a likely example of herd immunity (Fine *et al.* 2011)]. Crops can be rescued by breeding or transgenic manipulations that incorporate particular R gene alleles conferring resistance against virulent pathogen strains (Saintenac *et al.* 2013). Compared with agricultural systems, the molecular biology of plant–pathogen interactions in natural plant communities has received little attention [but see Karasov *et al.* (2014)].

Final thoughts

This study reveals possible connections between genetics and ecological dynamics of tropical trees. However, all of the time, effort and resources required to generate these data ultimately boiled down to single metrics per species, as required to make inferences about communities. Results so far provide impetus to increase the sample size of species (presently in progress) and to compare the same target species in additional communities where they differ in local abundance. In time, these studies may contribute to a growing body of work [e.g. Laine *et al.* (2011); Barrett & Heil (2012)] from wild plant populations indicating that genetic interactions between plant species and their pathogens play an important role in shaping natural communities. Further examination of these processes may bridge the gap between molecular evolution studies that indicate ongoing selection for multiple resistance alleles at many loci, and how this diversity functions in nature.

Acknowledgements

This study was funded by NSF grants DEB-1120476, DEB-1120121 and DEB-1457571 to the senior authors (JHM, SAM, CWD and LSC), using facilities and institutional resources of the Smithsonian Tropical Research Institute and our

respective universities. This study was executed using research and export permits and a materials transfer agreement from the Republic of Panama. We thank I. Torres, J. Eck, L. Rey Alvendaz for assistance with seed collection, seedling rearing and root collection, E. Shearer, P. Ralph, Y. Wang, T. Le for RNA extraction and N. Tran and L. Honaas for bioinformatics. Additional support was provided by NSF IOS-0950416, IOS-1354667, PGRR-0922742 and the Huck Institutes of the Life Sciences (JHM and CWD), NSF DEB-1257989 (SAM) and NSF DEB-1242622 and DEB-1464389 (LSC). Illumina sequencing was performed on an instrument funded by NSF MRI-1229046, with abundant technical advice and assistance from C. Praul and the Huck Genomics Core Facility. Single-molecule sequencing was facilitated by a gift of sequencing cells from Pacific Biosciences. This study was made possible by data from the BCI forest dynamics research project, made possible by grants from NSF, Center for Tropical Forest Science, STRI, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Small World Institute Fund, numerous private individuals, and the work of over 100 people from 10 countries over the past three decades.

References

Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ *et al.* (2004) Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew. *Science*, **306**, 1957–1960.

Allendorf FW (1986) Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, **5**, 181–190.

Alvarez ME (2000) Salicylic acid in the machinery of hypersensitive cell death and disease resistance. *Plant Molecular Biology*, **44**, 429–442.

Amborella Genome P (2013) The *Amborella* genome and the evolution of flowering plants. *Science*, **342**, 1241089.

Anderson JP, Gleason CA, Foley RC *et al.* (2010) Plants versus pathogens: an evolutionary arms race. *Functional Plant Biology*, **37**, 499–512.

Augspurger CK, Kelly CK (1984) Pathogen mortality of tropical tree seedlings – experimental studies of the effects of dispersal distance, seedling density, and light conditions. *Oecologia*, **61**, 211–217.

Bagchi R, Swinfield T, Gallery RE *et al.* (2010) Testing the Janzen-Connell mechanism: pathogens cause overcompensating density dependence in a tropical tree. *Ecology Letters*, **13**, 1262–1269.

Bakker EG, Toomajian C, Kreitman M, Bergelson J (2006) A genome-wide survey of R gene polymorphisms in *Arabidopsis*. *The Plant Cell Online*, **18**, 1803–1818 1040–4651.

Barrett LG, Heil M (2012) Unifying concepts and mechanisms in the specificity of plant-enemy interactions. *Trends in Plant Science*, **17**, 282–292.

Bergelson J, Kreitman M, Stahl EA, Tian D (2001) Evolutionary dynamics of plant R-genes. *Science*, **292**, 2281–2285.

Boshier DH, Chase MR, Bawa KS (1995) Population genetics of *Cordia alliodora* (Boraginaceae), a neotropical tree. 3. Gene flow, neighborhood, and population substructure. *American Journal of Botany*, **82**, 484–490.

Brennan AC, Harris SA, Hiscock SJ (2006) The population genetics of sporophytic self-incompatibility in *Senecio squalidus* L. (Asteraceae): the number, frequency, and dominance interactions of S alleles across its British range. *Evolution*, **60**, 213–224.

Brown JK, Tellier A (2011) Plant-parasite coevolution: bridging the gap between genetics and ecology. *Annual Review of Phytopathology*, **49**, 345–367.

Burdon JJ, Thrall PH, Ericson AL (2006) The current and future dynamics of disease in plant communities. *Annual Review of Phytopathology*, **44**, 19–39.

Bush MB, Colinvaux PA (1990) A pollen record of a complete glacial cycle from lowland Panama. *Journal of Vegetation Science*, **1**, 105–118.

Carella P, Merl-Pham J, Wilson DC *et al.* (2016) Comparative proteomics analysis of phloem exudates collected during the induction of systemic acquired resistance. *Plant Physiology*, **171**, 1495–1510.

Chen D, Toone WM, Mata J *et al.* (2003) Global transcriptional responses of fission yeast to environmental stress. *Molecular Biology of the Cell*, **14**, 214–229.

Chisholm RA, Muller-Landau HC (2011) A theoretical model linking interspecific variation in density dependence to species abundances. *Theoretical Ecology*, **4**, 241–253.

Chung K-R (2012) Stress response and pathogenicity of the necrotrophic fungal pathogen *Alternaria alternata*. *Scientifica*, **2012**, 17.

Clausen J, Keck D, Hiesey W (1940) *Experimental Studies on the Nature of Species. I. Effect of Varied Environment on Western North American Plants*. Carnegie Institution of Washington Publication no. 520, Washington, District of Columbia.

Comita LS, Muller-Landau HC, Aguilar S, Hubbell SP (2010) Asymmetric density dependence shapes species abundances in a tropical tree community. *Science*, **329**, 330–332.

Condit R, Ashton PS, Baker P *et al.* (2000) Spatial patterns in the distribution of tropical tree species. *Science*, **288**, 1414–1418.

Condit RLS, Pérez R, Dolins SB, Foster RB, Hubbell SP (2012) Barro Colorado Forest Census Plot Data, 2012 Version. Center for Tropical Forest Science Databases. DOI https://doi.org/10.5479/data.bci.20130603.

Dangl JL, Horvath DM, Staskawicz BJ (2013) Pivoting the plant immune system from dissection to deployment. *Science*, **341**, 746–751.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.

Dodds PN, Rathjen JP (2010) Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics*, **11**, 539–548.

Du D, Winsor JA, Smith M, Denicco A, Stephenson AG (2008) Resistance and tolerance to herbivory changes with inbreeding and ontogeny in a wild gourd (Cucurbitaceae). *American Journal of Botany*, **95**, 84–92.

Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size – Implications for plant conservation. *Annual Review of Ecology and Systematics*, **24**, 217–242.

Faheem M, Li Y, Arshad M *et al.* (2016) A disulphide isomerase gene (PDI-V) from *Haynaldia villosa* contributes to powdery mildew resistance in common wheat. *Scientific Reports*, **6**, 24227.

Fine P, Eames K, Heymann DL (2011) "Herd immunity": a rough guide. *Clinical Infectious Diseases*, **52**, 911–916.

Finkeldey R, Hattemer HH (2007) *Tropical Forest Genetics*. Springer, Berlin.

Floryszak-Wieczorek J, Arasimowicz-Jelonek M, Izbianska K (2016) The combined nitrate reductase and nitrite-dependent route of NO synthesis in potato immunity to *Phytophthora infestans*. *Plant Physiology and Biochemistry*, **108**, 468–477.

Fujiwara S, Kawazoe T, Ohnishi K *et al.* (2016) RipAY, a plant pathogen effector protein, exhibits robust gamma-glutamyl cyclotransferase activity when stimulated by eukaryotic thioredoxins. *Journal of Biological Chemistry*, **291**, 6813–6830.

Gilbert GS, Hubbell SP, Foster RB (1994) Density and distance-to-adult effects of a canker disease of trees in a moist tropical forest. *Oecologia*, **98**, 100–108.

Gilbert GS, Ferrer A, Carranza J (2002) Polypore fungal diversity and host density in a moist tropical forest. *Biodiversity and Conservation*, **11**, 947–957.

Glazebrook J (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology*, **43**, 205–227.

Glemin S, Gaude T, Guillemin ML *et al.* (2005) Balancing selection in the wild: testing population genetics theory of self-incompatibility in the rare species *Brassica insularis*. *Genetics*, **171**, 279–289.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.

Haas BJ, Papanicolaou A, Yassour M *et al.* (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494–1512.

Hammond-Kosack KE, Jones JD (1997) Plant disease resistance genes. *Annual Review of Plant Physiology and Plant Molecular Biology*, **48**, 575–607.

Hamrick JL, Murawski DA (1990) The breeding structure of tropical tree populations. *Plant Species Biology*, **5**, 157–165.

Hamrick JL, Murawski DA (1991) Levels of allozyme diversity in populations of uncommon Neotropical tree species. *Journal of Tropical Ecology*, **7**, 395–399.

Hardy OJ, Maggia L, Bandou E *et al.* (2006) Fine-scale genetic structure and gene dispersal inferences in 10 neotropical tree species. *Molecular Ecology*, **15**, 559–571.

Hoffman JI, Simpson F, David P *et al.* (2014) High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences*, **111**, 3775–3780.

Honaas LA, Wafula EK, Yang Z *et al.* (2013) Functional genomics of a generalist parasitic plant: laser microdissection of host-parasite interface reveals host-specific patterns of parasite gene expression. *BMC Plant Biology*, **13**, 9.

Honaas LA, Wafula EK, Wickett NJ *et al.* (2016) Selecting superior *de novo* transcriptome assemblies: lessons learned by leveraging the best plant genome. *PLoS One*, **11**, e0146062.

Hood LA, Swaine MD, Mason PA (2004) The influence of spatial patterns of damping-off disease and arbuscular mycorrhizal colonization on tree seedling establishment in Ghanaian tropical forest soil. *Journal of Ecology*, **92**, 816–823.

Hubbell SP, Foster RB, O'Brien ST *et al.* (1999) Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*, **283**, 554–557.

Hufford KM, Hamrick JL (2003) Viability selection at three early life stages of the tropical tree, *Platypodium elegans* (Fabaceae, Papilionoideae). *Evolution*, **57**, 518–526.

Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, **335**, 167–170.

Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO Journal*, **19**, 4004–4014.

Johnson DJ, Beaulieu WT, Bever JD, Clay K (2012) Conspecific negative density dependence and forest diversity. *Science*, **336**, 904–907.

Jousimo J, Tack AJ, Ovaskainen O *et al.* (2014) Disease ecology. Ecological and evolutionary effects of fragmentation on infectious disease dynamics. *Science*, **344**, 1289–1293.

Karasov TL, Kniskern JM, Gao L *et al.* (2014) The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature*, **512**, 436–440.

Kariyat RR, Mena-Alí J, Forry B *et al.* (2012) Inbreeding, herbivory, and the transcriptome of *Solanum carolinense*. *Entomologia Experimentalis Et Applicata*, **144**, 134–144.

Kim TH, Kunz HH, Bhattacharjee S *et al.* (2012) Natural variation in small molecule-induced TIR-NB-LRR signaling induces root growth arrest via EDS1- and PAD4-complexed R protein VICTR in *Arabidopsis*. *Plant Cell*, **24**, 5177–5192.

Klironomos JN (2002) Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature*, **217**, 67–70.

Knepper C, Day B (2010) From perception to activation: the molecular-genetic and biochemical landscape of disease resistance signaling in plants. *Arabidopsis Book*, **8**, e012.

Koboldt DC, Chen K, Wylie T *et al.* (2009) VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*, **25**, 2283–2285.

Kunz HH, Park J, Mevers E *et al.* (2016) Small molecule DFPM derivative-activated plant resistance protein signaling in roots is unaffected by EDS1 subcellular targeting signal and chemical genetic isolation of victr R-protein mutants. *PLoS One*, **11**, e0155937.

Laine AL, Burdon JJ, Dodds PN, Thrall PH (2011) Spatial variation in disease resistance: from molecules to metapopulations. *Journal of Ecology*, **99**, 96–112.

Liu X, Liang M, Etienne RS, *et al.* (2012) Experimental evidence for a phylogenetic Janzen-Connell effect in a subtropical forest. *Ecol Letters*, **15**, 111–118.

Loon L, Rep M, Pieterse CMJ (2006) Significance of inducible defense-related proteins in infected plants. *Annual Review of phytopathology*, **44**, 135–162.

Lorang J, Kidarsa T, Bradford CS *et al.* (2012) Tricking the guard: exploiting plant defense for disease susceptibility. *Science*, **338**, 659–662.

Loveless MD (1992) Isozyme variation in tropical trees: patterns of genetic organization. *New Forests*, **6**, 67–94.

Luna-Rivero MS, Hernandez-Zepeda C, Villanueva-Alonzo H *et al.* (2016) Expression of genes involved in the salicylic acid pathway in type h1 thioredoxin transiently silenced pepper plants during a begomovirus compatible interaction. *Molecular Genetics and Genomics*, **291**, 819–830.

Mack KML, Bever JD (2014) Coexistence and relative abundance in plant communities are determined by feedbacks when the scale of feedback and dispersal is local. *Journal of Ecology*, **102**, 1195–1201.

Mangan SA, Schnitzer SA, Herre EA *et al.* (2010) Negative plant-soil feedback predicts tree-species relative abundance in a tropical forest. *Nature*, **466**, 752–755.

McCarthy-Neumann S, Kobe RK (2010) Conspecific and heterospecific plant-soil feedbacks influence survivorship and growth of temperate tree seedlings. *Journal of Ecology*, **98**, 408–418.

Medina MJH, Gagnon H, Piche Y *et al.* (2003) Root colonization by arbuscular mycorrhizal fungi is affected by the salicylic acid content of the plant. *Plant Science*, **164**, 993–998.

Meyers BC, Kaushik S, Nandety RS (2005) Evolving disease resistance genes. *Current Opinion in Plant Biology*, **8**, 129–134.

Michelmore RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Research*, **8**, 1113–1130.

Ming R, VanBuren R, Liu Y *et al.* (2013) Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.). *Genome Biology*, **14**, R41.

Molina L, Kahmann R (2007) An *Ustilago maydis* gene involved in H_2O_2 detoxification Is required for virulence. *The Plant Cell Online*, **19**, 2293–2309.

Mondragon-Palomino M, Meyers BC, Michelmore RW, Gaut BS (2002) Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Research*, **12**, 1305–1315.

Moslonka-Lefebvre M, Finley A, Dorigatti I *et al.* (2011) Networks in plant epidemiology: from genes to landscapes, countries, and continents. *Phytopathology*, **101**, 392–403.

Muirhead CA (2001) Consequences of population structure on genes under balancing selection. *Evolution*, **55**, 1532–1541.

Mukaihara T, Hatanaka T, Nakano M, Oda K (2016) *Ralstonia solanacearum* type III effector RipAY is a glutathione-degrading enzyme that is activated by plant cytosolic thioredoxins and suppresses plant immunity. *MBio*, **7**, e00359–16.

Mundt CC (2002) Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytopathology*, **40**, 381–410.

Noreen AM, Webb EL (2013) High genetic diversity in a potentially vulnerable tropical tree species despite extreme habitat loss. *PLoS One*, **8**, e82632.

O'Brien SJ, Evermann JF (1988) Interactive influence of infectious disease and genetic diversity in natural populations. *Trends in Ecology & Evolution*, **3**, 254–259.

O'Dea A, Lessios HA, Coates AG *et al.* (2016) Formation of the isthmus of panama. *Science Advances* **2** no. **8**, e1600883.

Packer A, Clay K (2000) Soil pathogens and spatial patterns of seedling mortality in a temperate tree. *Nature*, **404**, 278–281.

Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ (2009) Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science*, **14**, 21–29.

Popa CM, Tabuchi M, Valls M (2016) Modification of bacterial effector proteins inside eukaryotic host cells. *Frontiers in Cellular and Infection Microbiology*, **6**, 73.

Portman SL, Kariyat RR, Johnston MA, Stephenson AG, Marden JH (2015) Cascading effects of host plant inbreeding on the larval growth, muscle molecular composition, and flight capacity of an adult herbivorous insect. *Functional Ecology*, **29**, 328–337.

Preston FW (1948) The commonness, and rarity, of species. *Ecology*, **29**, 254–283.

van der Putten WH, Bardgett RD, Bever JD *et al.* (2013) Plant-soil feedbacks: the past, the present and future challenges. *Journal of Ecology*, **101**, 265–276.

Reinhart KO, Royo AA, Kageyama SA, Clay K (2010) Canopy gaps decrease microbial densities and disease risk for a shade-intolerant tree species. *Acta Oecologica-International Journal of Ecology*, **36**, 530–536.

Rose LEB-E PD, Langley CH, Holub EB, Michelmore RW, Beynon JL (2004) The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics*, **166**, 1517–1527.

Saccheri I, Kuussaari M, Kankare M *et al.* (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491–494.

Saintenac C, Zhang W, Salcedo A *et al.* (2013) Identification of wheat gene *Sr35* that confers resistance to Ug99 stem rust race group. *Science*, **341**, 783–786.

Sang Y, Wang Y, Ni H *et al.* (2016) The *Ralstonia solanacearum* type-III effector RipAY targets plant redox regulators to suppress immune responses. *Molecular Plant Pathology*, doi:10.1111/mpp.12504.

Saucet SB, Ma Y, Sarris PF *et al.* (2015) Two linked pairs of *Arabidopsis* TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nature Communications*, **6**, 6338.

Seeholzer S, Tsuchimatsu T, Jordan T *et al.* (2010) Diversity at the Mla powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Molecular Plant-Microbe Interactions*, **23**, 497–509.

Shah J (2003) The salicylic acid loop in plant defense. *Current Opinion in Plant Biology*, **6**, 365–371.

Shihab HA, Gough J, Cooper DN *et al.* (2013) Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Human Mutation*, **34**, 57–65.

Simonich MT, Innes RW (1995) A disease resistance gene in *Arabidopsis* with specificity for the avrPph3 gene of *Pseudomonas syringae pv. phaseolicola*. *Molecular Plant-Microbe Interactions*, **8**, 637–640.

Vlot AC, Dempsey DA, Klessig DF (2009) Salicylic Acid, a multifaceted hormone to combat disease. *Annual Review of Phytopathology*, **47**, 177–206.

Volkov I, Banavar JR, Hubbell SP, Maritan A (2003) Neutral theory and relative species abundance in ecology. *Nature*, **424**, 1035–1037.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**, 57–63.

Wickett NJ, Mirarab S, Nguyen N *et al.* (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E4859–E4868.

Yang S, Feng Z, Zhang X *et al.* (2006) Genome-wide investigation on the genetic variations of rice disease resistance genes. *Plant Molecular Biology*, **62**, 181–193.

Yang P, Lupken T, Habekuss A *et al.* (2014) PROTEIN DISULFIDE ISOMERASE LIKE 5-1 is a susceptibility factor to plant viruses. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 2104–2109.

Yenni G, Adler PB, Ernest SK (2012) Strong self-limitation promotes the persistence of rare species. *Ecology*, **93**, 456–461.

Young AG, Pickup M (2010) Low S-allele numbers limit mate availability, reduce seed set and skew fitness in small

populations of a self-incompatible plant. *Journal of Applied Ecology*, **47**, 541–548.

JHM, LSC, SAM and CWD designed the project, obtained funding and wrote the manuscript. SAM directed seed collection, seedling growth and tissue collection. CWD directed transcriptome assembly and analyses. MPP helped conceive and wrote custom R scripts for expression and polymorphism analyses. EW wrote scripts and performed transcriptome assemblies, RNA-seq, sequence annotation, and orthogroup analyses. HWF performed RNA isolations, quality control and prepared RNA for sequencing. JPD designed orthology analyses and classification methods for R genes, and assisted with transcriptome processing and analysis. JHM analyzed the processed data, with input from MPP and LSC. All authors contributed to discussing the results and editing the manuscript.

Data accessibility

Sequence data are archived the NCBI Short Read Archive, SRA accession: SRP079907, Temporary Submission ID: SUB1719928, Release date: 2017-01-31 (or upon publication).

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Methods.

Appendix S2 Results.

Appendix S3 Transcriptomics analyses details.