Short communication

Targets of selection in a disease resistance network in wild tomatoesLAURA E. ROSE^{1,*}, LUKASZ GRZESKOWIAK^{1,†}, ANJA C. HÖRGER^{1,†}, MARTIN GROTH² AND WOLFGANG STEPHAN¹¹Section of Evolutionary Biology, LMU Munich, Grosshadernerstr. 2, 82152 Planegg, Germany²Section of Genetics, LMU Munich, Grosshadernerstr. 2, 82152 Planegg, Germany**SUMMARY**

Studies combining comparative genomics and information on biochemical pathways have revealed that protein evolution can be affected by the amount of pleiotropy associated with a particular gene. The amount of pleiotropy, in turn, can be a function of the position at which a gene operates in a pathway and the pathway structure. Genes that serve as convergence points and have several partners (so-called hubs) often show the greatest constraint and hence the slowest rate of protein evolution. In this article, we have studied five genes (*Pto*, *Fen*, *Rin4*, *Prf* and *Pfi*) in a defence signalling network in a wild tomato species, *Solanum peruvianum*. These proteins operate together and contribute to bacterial resistance in tomato. We predicted that *Prf* (and possibly *Pfi*), which serves as a convergence point for upstream signals, should show greater evolutionary constraint. However, we found instead that two of the genes which potentially interact with pathogen ligands, *Rin4* and *Fen*, have evolved under strong evolutionary constraint, whereas *Prf* and *Pfi*, which probably function further downstream in the network, show evidence of balancing selection. This counterintuitive observation may be probable in pathogen defence networks, because pathogens may target positions throughout resistance networks to manipulate or nullify host resistance, thereby leaving a molecular signature of host–parasite co-evolution throughout a single network.

Most proteins do not operate in isolation, but as components of complex pathways or metabolic networks. The ‘connectivity’ of a protein (i.e. the number of interactions of a protein with other components of a pathway) may determine the level of

constraint, and hence the rate of molecular evolution. Indeed, in yeast, the connectivity of genes in the network is correlated with their rate of evolution (Costanzo *et al.*, 2010). Similarly, the position in the pathway or network can affect the evolutionary constraint on the protein. For example, downstream proteins that serve as convergence points of a diverse group of signalling upstream molecules may be subject to greater evolutionary constraint than the upstream molecules (Alvarez-Ponce *et al.*, 2009). This can be viewed in terms of the extent of the pleiotropic effects of amino acid substitutions on proteins which serve as convergence points for different signalling molecules. Mathematically, it has been shown that highly pleiotropic genes ought to exhibit much reduced molecular variation (Waxman and Peck, 1998). Finally, the level of constraint may be affected by the degree of redundancy of genes in a pathway, which, in turn, may depend on whether the proteins are encoded by single-copy genes or by duplicate genes with overlapping functions (Costanzo *et al.*, 2010; Wagner, 2001).

Some of the first studies of pathways in plants indicated that upstream genes in biochemical pathways showed the greatest protein conservation when compared with downstream genes (Lu and Rausher, 2003; Rausher *et al.*, 1999, 2008). Recent studies in *Arabidopsis* have evaluated a number of defence genes, some of which are known to operate together in specific signalling pathways (Bakker *et al.*, 2006, 2008; Caldwell and Micheltore, 2009). Although these studies were not explicitly designed to test the effect of pathway position on evolutionary rates, the combined analysis of 27 resistance (*R*) genes and 27 downstream genes in *Arabidopsis* by Bakker *et al.* (2006, 2008) revealed that, although some *R* genes showed histories of transient balancing selection or partial selective sweeps, genes further downstream experienced almost exclusively purifying selection. These results are consistent with expectations that genes downstream in defence networks experience greater evolutionary constraint and upstream genes can be subject to adaptive change.

*Correspondence: Email: rose@bio.lmu.de

†These authors contributed equally to this work.

In this article, we describe sequence variation at genes operating at different points in a network controlling disease resistance in wild tomatoes. Such case studies complement the analyses of large protein databases, because, in case studies, the forces underlying evolutionary constraints can be analysed in much greater detail and can potentially capture a different evolutionary timescale (e.g. Wagner, 2000, 2001). Furthermore, the evolution of components of a signalling network may not conform to expectations based on biochemical pathways. As a first step in our analysis of this resistance network, we have previously characterized the sequence variation of the pathogen resistance *R* gene, *Pto*, and the functional consequences of this variation within and between populations of seven *Solanum* species (Rose *et al.*, 2005, 2007). We found evidence for elevated levels of amino acid polymorphism at this *R* gene, consistent with balancing selection at this locus. In this article, we describe the sequence variation and level of evolutionary constraint for an additional four genes involved in this signalling network: *Fen*, *Prf*, *Pfi* and *Rin4*.

This network has multiple signalling molecules, including *Pto*, *Fen* and *Rin4*, as potential inputs (Fig. 1). *Pto* is a receptor kinase that is activated on binding of pathogen-derived molecules, *AvrPto* or *AvrPtoB* (reviewed in Oh and Martin, 2010). *Fen* is a paralogue of *Pto* and confers sensitivity to the insecticide fenthion (Chang *et al.*, 2002; Martin *et al.*, 1994). *Fen* can also recognize and activate defence responses to versions of the *Pst* effector *AvrPtoB* lacking E3 ligase activity (Rosebrock *et al.*, 2007). However, wild-type forms of *AvrPtoB* ubiquitinate *Fen*, which leads to its degradation in plant cells. The *Rin4* gene was originally identified in *Arabidopsis* and plays a role in several *R* gene signalling pathways (Kim *et al.*, 2005). Identification and functional studies of *Rin4* in tomato indicate that it is degraded

in the presence of pathogen effectors, *AvrPto* and *AvrPtoB*, and this degradation depends on *Pto* and *Prf* (Luo *et al.*, 2009). As *Rin4* is believed to be a negative regulator of basal defence, its degradation is predicted to activate downstream defences.

Prf interprets and transduces these signals. *Prf* is located in the same cluster as the *Pto* gene family, although it is phylogenetically unrelated to *Pto* and its paralogues. The *Prf* protein contains regions showing homology to nucleotide-binding sites (NBS), coiled-coil (CC) domains and leucine-rich repeats (LRR). Silencing of *Prf* prevents signalling by *Fen* or *Pto*, indicating that *Prf* acts epistatically to *Fen* and *Pto*. It has been demonstrated that the two kinases, *Pto* and *Fen*, physically interact with the same N-terminal portion of *Prf* (Mucyn *et al.*, 2006; Ntoukakis *et al.*, 2009). The fifth gene in our study, *Pfi* (originally named *Prf*-interactor 30137), was cloned via yeast two-hybrid screen with a portion of the CC and NBS region of *Prf* as bait (Tai, 2004). This gene encodes a protein with homology to basic helix–loop–helix transcription factors. Functional testing of *Pfi* indicated that over-expression in tomato suppresses the hypersensitive response (HR), whereas viral-induced gene silencing of *Pfi* showed no phenotypic effect (Tai, 2004). As such, this gene appears to be a negative regulator of HR. Our further bioinformatic analyses of *Pfi* revealed two additional regions of interest: a nuclear localization signal (NLS) and a region with homology to hydrolases.

Based on this network structure, we might expect *Prf*, which acts epistatically to *Pto* and *Fen*, to show a stronger signature of purifying selection, because mutations in this gene affect at least two phenotypes (signalling via *Fen* and *Pto*). However, the genes *Pto*, *Fen* and *Rin4* may show a signature of relaxed constraint because they are less pleiotropic. Alternatively, these three genes may be subject to adaptive changes because they interact

Fig. 1 Schematic representation of the *Pto* signalling pathway in tomato.

directly with pathogen molecules. However, as the precise locations of both *Rin4* and *Pfi* in this network remain to be elucidated, definitive a priori predictions about the presence or absence of selective constraint operating at these genes are not possible.

For our study, we sampled multiple individuals of *S. peruvianum* from a single, large population in Tarapaca, Chile. *S. peruvianum* is endemic to the western coast of South America and is closely related to the cultivated tomato. This species is widespread and often occurs in large stands in central and southern Peru and northern Chile (reviewed in Chetelat *et al.*, 2009). Individuals of this species are diploid, obligate outcrossing, short-lived perennials, and census population sizes range from a few plants to several hundred (tgrc.ucdavis.edu). For this study, plants were grown under standard glasshouse conditions. DNA isolation, polymerase chain reaction (PCR) conditions and sequencing methods are described in Notes S1 (see Supporting Information). The entire coding regions of the five network genes were amplified and sequenced, yielding 20 fully resolved alleles for each gene (for a total of 14 167 base pairs). For comparisons between loci, the sequences of alleles of 14 other loci (Table 1) were obtained from Baudry *et al.* (2001) and Roselius *et al.* (2005). These reference genes were amplified from the same individuals of this *S. peruvianum* population.

Standard population summary statistics and tests of neutrality were conducted using DnaSP v. 4.0 (Rozas *et al.*, 2003). Coales-

cent simulations were used to examine whether the patterns of substitution at nonsynonymous and synonymous sites at these genes differed from those of the 14 reference genes. We used the arithmetic mean of π (the average pairwise differences) of the reference genes as the estimate of θ (the population mutation parameter). A thousand simulations were executed in DnaSP and, subsequently, we determined whether the value of π observed at the network genes fell within the 95% confidence interval of the simulations based on θ estimated from the 14 reference genes. For these simulations, we assumed no recombination.

In our analysis of 20 alleles from *S. peruvianum* of these five genes, polymorphism, quantified by average pairwise differences across all sites (π), ranged from 0.6% (*Prf*) to 1.6% (*Pfi*) (Table 1). For comparison, the mean across the set of 14 reference genes from these same individuals was 1.3%. *Pfi* and *Pto* showed the highest polymorphism at synonymous sites (2.2% and 2.0%, respectively), as well as at nonsynonymous sites (1.3% at both loci). The ratio of π_{non} to π_{syn} was 0.57 for *Pfi* and 0.62 for *Pto*, whereas this ratio was consistently much lower at the 14 reference loci (mean π_{non} to π_{syn} = 0.10, Table 1).

We used neutral coalescent simulations to test whether the values of π observed at nonsynonymous and synonymous sites at the set of resistance genes differed from those of the 14 reference genes (Hudson, 1990). These simulations indicated that both *Pfi* and *Pto* show excess variation at nonsynonymous sites ($P < 0.001$), whereas, at synonymous sites, the observed level of

Table 1 Summary statistics across loci.

Gene	π^* total (sites) [†]	π_{syn} (sites)	π_{non} (sites)	$\pi_{\text{non}}/\pi_{\text{syn}}$	Haplotype diversity	$D^\ddagger$	$R^\S$	ZnS ¶
<i>Pto</i>	0.01450 (928)	0.02038 (210.66)	0.01278 (710.34)	0.62	0.912	0.06383	0.0168	0.1698
<i>Fen</i>	0.00871 (966)	0.0156 (214.54)	0.00676 (745.46)	0.43	0.942	-0.55114	0.0406	0.1241
<i>Prf</i>	0.00667 (5541)	0.01386 (1206.59)	0.00448 (4211.41)	0.32	0.947	-0.27218	0.0117	0.1125
<i>Pfi</i>	0.01662 (5556)	0.02233 (494.17)	0.01277 (1707.83)	0.57	0.971	-0.59193	0.0108	0.1243
<i>Rin4</i>	0.00924 (1176)	0.01984 (221.34)	0.00320 (714.66)	0.16	0.947	-0.88121	0.0517	0.1113
<i>CT066</i>	0.00984 (1346)	0.03366 (332.10)	0.00204 (1011.90)	0.06	0.933	-0.30729	0.0306	0.1871
<i>CT093</i>	0.00568 (1388)	0.01763 (248.30)	0.00105 (780.70)	0.06	0.956	-0.14066	0.0591	0.1677
<i>CT099</i>	0.01827 (1198)	0.02138 (238.97)	0.00709 (649.03)	0.33	0.978	-0.81459	0.0408	0.1606
<i>CT114</i>	0.00820 (1165)	0.01605 (162.00)	0.00000 (510.00)	0.00	0.933	0.18940	0.0188	0.2535
<i>CT143</i>	0.01839 (1616)	0.01652 (113.00)	0.00000 (355.00)	0.00	0.978	0.25007	0.0141	0.2173
<i>CT148</i>	0.01433 (1394)	0.02010 (135.98)	0.00524 (407.02)	0.26	1.000	-1.05934	0.1020	0.1348
<i>CT166</i>	0.01429 (2602)	0.00699 (209.50)	0.00078 (732.50)	0.11	0.967	-0.03372	0.0198	0.2316
<i>CT179</i>	0.01069 (958)	0.03457 (153.00)	0.00000 (426.00)	0.00	0.911	-0.16307	0.0524	0.1980
<i>CT189</i>	0.01010 (1353)	0.00726 (82.55)	0.00000 (271.45)	0.00	1.000	-1.57754	0.0047	0.2295
<i>CT198</i>	0.02924 (760)	0.05648 (76.33)	0.00182 (256.67)	0.03	0.911	0.06862	0.0619	0.1958
<i>CT208</i>	0.00746 (1746)	0.00674 (168.10)	0.00000 (566.90)	0.00	0.778	-0.59745	0.0021	0.2866
<i>CT251</i>	0.01400 (1678)	0.03448 (317.78)	0.00721 (1005.22)	0.21	0.933	-0.24984	0.0125	0.2098
<i>CT268</i>	0.00941 (1884)	0.02587 (435.60)	0.00446 (1448.40)	0.17	1.000	-0.73472	0.0089	0.2027
<i>sucl</i>	0.01406 (1434)	0.02692 (267.43)	0.00401 (842.57)	0.15	1.000	-0.66941	0.1519	0.1358

*Average pairwise differences.

[†]Total number of sites analysed excluding gaps.

[‡]Tajima's D (Tajima, 1989); all sites analysed.

[§]Hudson (1987), all sites analysed.

[¶]Kelly (1997), all sites analysed.

Locus	π_{syn}	Prob ($\pi_{\text{exp}} > \pi_{\text{obs}}$) [†]	π_{non}	Prob ($\pi_{\text{exp}} > \pi_{\text{obs}}$) [†]
Reference genes‡	0.023		0.0024	
<i>Fen</i>	0.016	0.668	0.0068	0.015
<i>Pfi</i>	0.022	0.413	0.013	<0.001
<i>Prf</i>	0.014	0.787	0.0045	0.085
<i>Pto</i>	0.020	0.390	0.013	<0.001
<i>Rin4</i>	0.020	0.520	0.0032	0.191

[†]Probability of observing a value of π greater than that observed at network genes in 1000 coalescent simulations, conditioned on the π values of the reference genes.

[‡]Arithmetic mean of π from 14 genes (*CT066*, *CT093*, *CT099*, *CT114*, *CT143*, *CT148*, *CT166*, *CT179*, *CT189*, *CT198*, *CT208*, *CT251*, *CT268* and *sucr*).

variation is within the 95% confidence interval based on θ across these 14 other genes (Table 2). Previous studies of *Pto* alleles from this tomato population have revealed that a larger proportion of ancestral variation is maintained at *Pto* when compared with these reference loci (Rose *et al.*, 2007). As the presence of elevated levels of amino acid polymorphisms is equally consistent with a hypothesis of relaxed constraint, we evaluated the functional consequences of these segregating amino acid variants. We found that the frequency spectrum of amino acid polymorphisms known to negatively affect *Pto* function is skewed towards low frequency compared with polymorphisms that do not affect function (Rose *et al.*, 2005, 2007). This reveals that deleterious mutations are not at high frequency in this population, and that the evolution of *Pto* appears to be influenced by a mixture of both purifying and balancing selection.

The evaluation of both polymorphism and divergence at *Pfi* using the McDonald–Kreitman test lent support to the hypothesis of balancing selection operating at *Pfi* (McDonald and Kreitman, 1991). Here, we detected significantly more variation at nonsynonymous sites (i.e. polymorphisms segregating in this population) than expected under neutrality (Table S1, see Supporting Information). A closer inspection of the distribution of variation across this large gene reveals that the region encoding the putative hydrolase motif and NLS harbours substantial amounts of nonsynonymous variation ($\pi_{\text{non}} = 0.0216$; Fig. 2a). Outside of this region, nonsynonymous variation is much lower: 0.00423. What is striking is that the region of elevated polymorphism is centred on the hydrolase region, but does not affect neighbouring coding regions. This could potentially be the target of natural selection driven by co-evolution with pathogens, as discussed below.

Of the five network genes, *Fen* and *Prf* showed the lowest levels of polymorphism, but intermediate values of $\pi_{\text{non}}/\pi_{\text{syn}}$. *Fen*, like *Pto*, is a small gene and encodes a functional protein kinase. Although *Fen* is known to interact with some pathogen ligands, no resistance function similar to that of *Pto* has been assigned to this gene. However, as this gene is expressed and retains its kinase activity across multiple wild tomato species, it is probably

Table 2 Results of coalescent simulations.

Fig. 2 Sliding window analysis of average pairwise differences among *Pfi* alleles (a) and *Prf* alleles (b) in *Solanum peruvianum*. Silent polymorphisms include both synonymous polymorphisms and polymorphisms in nonprotein coding regions, such as introns. The structure of the genes is shown below the graph. Boxes indicate exons, solid lines indicate introns. The putative functional regions are indicated below the appropriate exons. bHLH, basic helix–loop–helix; CC, coiled-coil; LRR, leucine-rich repeat; NB-ARC, nucleotide binding adaptor shared by APAF-1, certain R proteins, and CED-4; NLS, nuclear localization signal.

functionally important for tomatoes (Chang *et al.*, 2002; Riely and Martin, 2001; Rosebrock *et al.*, 2007). Selective constraint at this locus may explain the absence of a strong signature of either directional or balancing selection at *Fen*.

In contrast with *Fen*, *Prf* is a large gene, made up of both well-defined and poorly defined domains. These domains show different evolutionary histories, as captured in the sliding window analyses (Fig. 2b). In contrast with many other *R* genes, the LRR region of *Prf* does not show an excess of amino acid polymorphism. Instead, the N-terminal portion of the protein, known to bind *Pto* and *Fen*, has peaks of replacement site polymorphism. This elevated polymorphism provides hints that balancing selection at *Pto* may be carrying over to its interacting partner, *Prf*. These types of correlated selective history open the way to more complex forms of selection, such as epistatic selection between molecules. In comparison with the first half of *Prf*, the second half shows greater evolutionary constraint, consistent with its presumed role in downstream signalling.

The gene showing the greatest level of evolutionary constraint is *Rin4*. This gene has the lowest level of nonsynonymous polymorphism and the lowest $\pi_{\text{non}}/\pi_{\text{syn}}$ of the five genes (Table 1). Indeed, based on the distribution and levels of polymorphism, *Rin4* appears to be indistinguishable from the 14 reference loci. However, the frequency spectrum of mutations and the presence of a young, but divergent, allelic type, carrying several derived mutations, expose additional aspects of the *Rin4* history (Fig. S1, see Supporting Information). Alleles 7232.1, 7233.1 and 7240.1 show nine fixed differences relative to the other alleles. All nine nonsingleton polymorphisms distributed throughout the *Rin4* coding sequence are in significant linkage disequilibrium (LD). Two of these differences are nonsynonymous, whereas the others do not encode protein differences. These nine sites are derived relative to the alleles from the outgroup species, *S. habrochaites* and *S. pennellii*. Seven of these changes are derived relative to the more distant outgroup, *S. lycopersicoides* (Fig. S2, see Supporting Information). The absence of evidence of recombination between this sequence type and the others, the strong pattern of LD involving derived changes, two of which are nonsynonymous, and the low to moderate frequency of this sequence type are consistent with the presence of a partial or ongoing sweep at *Rin4*.

In our study of five genes involved in the *Pto* signalling network, we found two loci with elevated amino acid polymorphism, consistent with balancing selection, namely *Pto* and *Pfi*. A third gene, *Prf*, showed signatures of both balancing selection and purifying selection, whereas two other genes, namely *Fen* and *Rin4*, showed predominantly purifying selection. Previous studies have reported that *Pto* is subject to balancing selection within different wild tomato species and, given the substantial functional information available for *Pto*, a scenario of balancing selection is not surprising (Rose *et al.*, 2005, 2007). *Pto* binds and recognizes two different pathogen ligands and triggers a defence response in wild tomato. The maintenance of different host resistance proteins in natural populations is consistent with an ongoing co-adaptation between host and pathogen.

The second gene that showed elevated amino acid polymorphism relative to neutral expectations was *Pfi*. The protein product of *Pfi* physically interacts with *Prf*, has a putative NLS and is predicted to encode a transcription factor (Tai, 2004). As such, it may respond to an activated form of *Prf* by moving into the nucleus; however, the precise action of this protein remains to be elucidated. As a probable component of the signalling network, rather than a known pathogen target, it is surprising to uncover a signal of balancing selection at *Pfi*. The signature of balancing selection is located in a region that encodes a putative hydrolase, although enzymatic assays to confirm hydrolytic activity have yet to be conducted. Provided that this molecule is enzymatically active, it is possible that natural selection operates directly on the enzymatic function, and that protein variation is maintained in this region as a result of selection for different substrate specificities, perhaps involved in pathogen defence. Alternatively, this molecule could serve as a direct target by other tomato pathogens. Recent studies have revealed that all points in immune signalling networks can be vulnerable to pathogen manipulation (reviewed in Brodsky and Medzhitov, 2009). Pathogens may specifically secrete proteins (i.e. effector molecules) to target downstream points in the network to suppress host resistance (reviewed in Zhou and Chai, 2008). *Pfi* is described to be a negative regulator of defence; therefore, alteration of protein stability could result in suppression of the HR (Tai, 2004). In this case, balancing selection may not be specifically operating on enzymatic function, but rather on pathogen evasion. Alternative forms of *Pfi* found in these natural populations may vary in their 'resistance' to manipulation by pathogen molecules.

Prf, one of the central molecules of this network, showed two distinctive signals of natural selection. The region known to physically interact with *Pto* and *Fen* showed elevated amino acid polymorphism, providing the first hint that balancing selection at *Pto* may be carrying over to its interacting partner *Prf*. In comparison with the first half of *Prf*, the second half of this gene shows greater evolutionary constraint, consistent with its presumed role in downstream signalling.

The *Fen* and *Rin4* genes showed the greatest evolutionary constraint of the five genes. Although *Fen* is known to interact with some pathogen ligands, no resistance function similar to that of *Pto* has been assigned to this gene. It is possible that *Fen* does not operate in the same isolate-specific manner as *Pto*. If *Fen* is involved in basal defence and not in isolate-specific defence, it would be subject to different evolutionary forces. One such molecule that is known to contribute to basal defence and is involved in different resistance pathways (at least in *Arabidopsis thaliana*) is *Rin4*. We observed strong protein conservation at *Rin4* in our tomato population. However, the frequency spectrum of mutations and pattern of LD among *Rin4* alleles expose additional aspects of the history of *Rin4*, including the

presence of a young, but divergent, *Rin4* allelic type, carrying several derived mutations. One possible explanation of this pattern is that this divergent *Rin4* allele is passing through the population as an advantageous allele. However, detecting a selective sweep in progress is quite unlikely because the sojourn times of advantageous alleles are generally too fast. The fact that this *Rin4* allele with several derived changes is segregating with two other distinct alleles, all at moderate frequency, and none of these three allelic types shows any evidence of recombination, indicates that the frequency spectrum of these alleles has been perturbed in the recent history of this plant population.

Recent studies in *Arabidopsis* have evaluated a number of defence genes, some of which are known to operate together in specific signalling pathways (Bakker *et al.*, 2006, 2008; Caldwell and Michelmore, 2009). At a broad scale, these results are consistent with expectations that genes downstream in defence pathways experience greater evolutionary constraint, and upstream genes are subject to adaptive change. However, a subset of these same genes was recently evaluated more extensively by another team, and they came to slightly different conclusions (Caldwell and Michelmore, 2009). In a study of 10 downstream defence genes in *A. thaliana*, three genes (*NPR1*, *EDS1* and *PAD4*) showed interesting patterns of past adaptive evolution. This signature of balancing selection in these three genes may have been missed by Bakker *et al.* (2008) because, in the original study, only portions of the coding regions were analysed, rather than the entire genes. Interestingly, a fourth gene in the study by Caldwell and Michelmore (2009) overlapped with one in our study, namely *Rin4*. In their initial analyses, *Rin4* was identified as a potential outlier based on Hudson–Kreitman–Agaude (HKA) tests. The HKA test can be used to evaluate whether the level of intraspecific polymorphism at the locus of interest is elevated relative to the neutral expectation, and therefore serves as a powerful means to identify genes affected by balancing selection (Hudson *et al.*, 1987). Although the results from the HKA test for *Rin4* were inconclusive following correction for multiple testing, the authors reported that *Rin4* harbours substantial polymorphism within *Arabidopsis*, displaying more genetic variation than found at 93.5% of the genes in a set of 355 reference loci (Caldwell and Michelmore, 2009). To what degree this elevation in genetic diversity reflects past selective events has not been investigated.

Compared with these other studies, we did not find a strong correlation of selective constraint and network position. In a recent study of the genetic interactions in yeast, the authors reported that genes that have many interactions and serve as hubs are typically evolutionarily conserved (Costanzo *et al.*, 2010). However, not all genes with many interactions evolve slowly. For example, genes that encode proteins with disordered regions have a higher rate of evolution, despite having numer-

ous genetic interactions. Given the observation that the rate of evolution is not only a function of the number of genetic interactions, but also dependent on other features, such as protein structure, our current models of network evolution may be too simplistic and may not apply to defence pathways.

A second explanation for the lack of correlation between network position and selective constraint may reflect the biological reality that genes at several points in defence networks can be targets of adaptive evolution. Consequently, population genetic studies, such as ours, can uncover interesting candidates for future functional studies. For example, the consequences of *Rin4* protein polymorphism on *Pto*-specific resistance and possibly basal defence responses could be tested using the methods presented recently by Luo *et al.* (2009). Likewise, a better understanding of the functional consequences of protein polymorphism around the enzymatic core of the Pfi protein will probably reveal novel aspects of the defence repertoire of plants as, although this gene displays a signature of balancing polymorphism similar to other *R* genes in plants, it does not share the motifs of most other *R* genes.

ACKNOWLEDGEMENTS

We thank the Munich population genetics group for many valuable discussions, and excellent technical support from G. Büttner and S. Lange. The plant seeds were provided by the C.M. Rick Tomato Genetics Resource Center (TGRC; University of California, Davis, CA, USA). This work was supported by the Deutsche Forschungsgemeinschaft (DFG) grants: RO 2491/2-1 to LER and WS, and RO 2491/2-3 to LER.

REFERENCES

- Alvarez-Ponce, D., Aguade, M. and Rozas, J. (2009) Network-level molecular evolutionary analysis of the insulin/TOR signal transduction pathway across 12 *Drosophila* genomes. *Genome Res.* **19**, 234–242.
- Bakker, E.G., Toomajian, C., Kreitman, M. and Bergelson, J. (2006) A genome-wide survey of R gene polymorphisms in *Arabidopsis thaliana*. *Plant Cell*, **18**, 1803–1818.
- Bakker, E.G., Traw, M.B., Toomajian, C., Kreitman, M. and Bergelson, J. (2008) Low levels of polymorphism in genes that control the activation of defense response in *Arabidopsis thaliana*. *Genetics*, **178**, 2031–2043.
- Baudry, E., Kerdelhue, C., Innan, H. and Stephan, W. (2001) Species and recombination effects on DNA variability in the tomato genus. *Genetics*, **158**, 1725–1735.
- Brodsky, I.E. and Medzhitov, R. (2009) Targeting of immune signalling networks by bacterial pathogens. *Nat. Cell Biol.* **11**, 521–526.
- Caldwell, K.S. and Michelmore, R.W. (2009) *Arabidopsis thaliana* genes encoding defense signaling and recognition proteins exhibit contrasting evolutionary dynamics. *Genetics*, **181**, 671–684.
- Chang, J.H., Tai, Y.-S., Bernal, A.J., Lavelle, D.T., Staskawicz, B.J. and Michelmore, R.W. (2002) Functional analyses of the *Pto* resistance gene family in tomato and the identification of a minor resistance determinant in a susceptible haplotype. *Mol. Plant–Microbe Interact.* **15**, 281–291.

- Chetelat, R.T., Pertuzé, R.A., Faúndez, L., Graham, E.B. and Jones, C.M. (2009) Distribution, ecology and reproductive biology of wild tomatoes and related nightshades from the Atacama Desert region of northern Chile. *Euphytica*, **167**, 77–93.
- Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L.Y., Toufighi, K., Mostafavi, S., Prinz, J., St. Onge, R.P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F.J., Alizadeh, S., Bahr, S., Brost, R.L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z.-Y., Liang, W., Marback, M., Paw, J., San Luis, B.-J., Shuteriqi, E., Hin Yan Tong, A., van Dyk, N., Wallace, I.M., Whitney, J.A., Weirauch, M.T., Zhong, G., Zhu, H., Houry, W.A., Brudno, M., Ragibizadeh, S., Papp, B., Pál, C., Roth, F.P., Giaever, G., Nislow, C., Troyanskaya, O.G., Bussey, H., Bader, G.D., Gingras, A.-C., Morris, Q.D., Kim, P.M., Kaiser, C.A., Myers, C.L., Andrews, B.J. and Boone, C. (2010) The genetic landscape of a cell. *Science*, **327**, 425–431.
- Hudson, R.R. (1987) Estimating the recombination parameter of a finite population model without selection. *Genet. Res.* **50**, 245–250.
- Hudson, R.R. (1990) Gene genealogies and the coalescent process. In: *Oxford Surveys in Evolutionary Biology* (Futuyma, D. and Antonovics, J., eds), pp. 1–44. New York: Oxford University Press.
- Hudson, R.R., Kreitman, M. and Aguade, M. (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.
- Kelly, J.K. (1997) A test of neutrality based on interlocus associations. *Genetics*, **146**, 1197–1206.
- Kim, M.G., da Cunha, L., McFall, A.J., Belkhadir, Y., DebRoy, S., Dangl, J.L. and Mackey, D. (2005) Two *Pseudomonas* type III effectors inhibit RIN4-regulated basal defense in *Arabidopsis*. *Cell*, **121**, 749–759.
- Lu, Y.Q. and Rausher, M.D. (2003) Evolutionary rate variation in anthocyanin pathway genes. *Mol. Biol. Evol.* **20**, 1844–1853.
- Luo, Y., Caldwell, K.S., Wroblewski, T., Wright, M.E. and Michelmore, R.W. (2009) Proteolysis of a negative regulator of innate immunity is dependent on resistance genes in tomato and *Nicotiana benthamiana* and induced by multiple bacterial effectors. *Plant Cell*, **21**, 2458–2472.
- Martin, G.B., Frary, A., Wu, T., Brommonschenkel, S., Chunwongse, J., Earle, E.D. and Tanksley, S.D. (1994) A member of the tomato *Pto* gene family confers sensitivity to fenthion resulting in rapid cell death. *Plant Cell*, **6**, 1543–1552.
- McDonald, J.H. and Kreitman, M. (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*, **351**, 652–654.
- Mucyn, T.S., Clemente, A., Andriotis, V.M.E., Balmuth, A.L., Oldroyd, G.E.D., Staskawicz, B.J. and Rathjen, J.P. (2006) The tomato NBARC-LRR protein Prf interacts with Pto kinase *in vivo* to regulate specific plant immunity. *Plant Cell*, **18**, 2792–2806.
- Ntoukakis, V., Mucyn, T.S., Gimenez-Ibanez, S., Chapman, H.C., Gutierrez, J.R., Balmuth, A.L., Jones, A.M.E. and Rathjen, J.P. (2009) Host inhibition of a bacterial virulence effector triggers immunity to infection. *Science*, **324**, 784–787.
- Oh, C.-S. and Martin, G.B. (2010) Effector-triggered immunity mediated by the Pto kinase. *Trends Plant Sci.* **16**, 132–140. Published online November 26, 2010.
- Rausher, M.D., Miller, R.E. and Tiffin, P. (1999) Patterns of evolutionary rate variation among genes of the anthocyanin biosynthetic pathway. *Mol. Biol. Evol.* **16**, 266–274.
- Rausher, M.D., Lu, Y.Q. and Meyer, K. (2008) Variation in constraint versus positive selection as an explanation for evolutionary rate variation among anthocyanin genes. *J. Mol. Evol.* **67**, 137–144.
- Riely, B.K. and Martin, G.B. (2001) Ancient origin of pathogen recognition specificity conferred by the tomato disease resistance gene *Pto*. *Proc. Natl. Acad. Sci. USA*, **98**, 2059–2064.
- Rose, L.E., Langley, C.H., Bernal, A.J. and Michelmore, R.W. (2005) Natural variation in the *Pto* pathogen resistance gene within species of wild tomato (*Lycopersicon*): I. Functional analysis of *Pto* alleles. *Genetics*, **171**, 345–357.
- Rose, L.E., Langley, C.H. and Michelmore, R.W. (2007) Natural variation in the *Pto* pathogen resistance gene within species of wild tomato (*Lycopersicon*): II. Population genetics of *Pto*. *Genetics*, **175**, 1307–1319.
- Rosebrock, T.R., Zeng, L., Brady, J.J., Abramovitch, R.B., Xiao, F. and Martin, G.B. (2007) A bacterial E3 ubiquitin ligase targets a host protein kinase to disrupt plant immunity. *Nature*, **448**, 370–374.
- Roselius, K., Stephan, W. and Städler, T. (2005) The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. *Genetics*, **171**, 753–763.
- Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X. and Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.
- Tai, Y.-S. (2004) Dissertation. The role of Prf and its partners in resistance to *Pseudomonas syringae* pv. *tomato*. Davis, CA: University of California.
- Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–596.
- Wagner, A. (2000) Mutational robustness in genetic networks of yeast. *Nat. Genet.* **24**, 355–361.
- Wagner, A. (2001) The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes. *Mol. Biol. Evol.* **18**, 1283–1292.
- Waxman, D. and Peck, J. (1998) Pleiotropy and the preservation of perfection. *Science*, **279**, 1210–1213.
- Zhou, J.-M. and Chai, J. (2008) Plant pathogenic bacterial type III effectors subdue host responses. *Curr. Opin. Microbiol.* **11**, 179–185.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Maximum parsimony tree based nucleotide sequences of *Rin4* alleles from *S. peruvianum*.

Fig. S2 Distribution of non-singleton polymorphisms at the *Rin4* gene among individuals of *S. peruvianum*.

Table S1 McDonald–Kreitman test on *Pfi*.

Notes S1 Experimental procedures.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.