2008

Geographic diversity cline of R gene homologs in wild populations of Solanum pimpinellifolium (Solanaceae)

Ana Lucia Caicedo, *University of Massachusetts - Amherst*

American Journal of Botany 95(3): 393–398. 2008.

BRIEF COMMUNICATION

GEOGRAPHIC DIVERSITY CLINE OF *R* GENE HOMOLOGS IN WILD POPULATIONS OF *SOLANUM PIMPINELLIFOLIUM* (SOLANACEAE)[1]

ANA L. CAICEDO[2]

Biology Dept., 221 Morrill Science Center, University of Massachusetts, Amherst, Massachusetts 01003 USA

Plant resistance (*R*) genes tend to be highly variable within plant species and are thought to be under natural selection; however, little is known about the geographic distribution of *R* gene diversity within and among plant populations. To determine the possible roles of demography and selection on *R* gene evolution, patterns of diversity at the multigenic *Cf-2 R* gene family were studied in *Solanum pimpinellifolium* populations along the northern coast of Peru. Population diversity levels of *Cf-2* homologs follow a latitudinal cline, consistent with the species's history of gradual colonization of the Peruvian coast and population variation in outcrossing levels. Although previous evidence suggests that selection has shaped the DNA sequence content of the *Cf-2* genes, current results imply that the geographic distribution of *Cf-2* homolog diversity has been shaped primarily by demographic factors or by selective pressures with a clinal distribution.

Key words: *Cf*-2; diversity cline; *Hcr2-p*; *Lycopersicon pimpinellifolium*; Peru; *R* gene; Solanaceae.

In plants, resistance (*R*) genes are responsible for the recognition of invading pathogens, leading to a defense response (reviewed in Ellis et al., 2000; Michelmore, 2000; Hulbert et al., 2001; Meyers et al., 2005). In the past decade, studies of *R* gene sequence variation have shown that *R* genes tend to be highly polymorphic within species (e.g., Caicedo et al., 1999; Stahl et al., 1999; Bergelson et al., 2001; Mauricio et al., 2003; Rose et al., 2004). Studies have also shown that alleles of some *R* genes are long-lived, suggesting that *R* gene variation may often be maintained by balancing selection (Bergelson et al., 2001); other studies have shown that frequency-dependent or transient selection may also play a role in maintaining *R* gene diversity (Bakker et al., 2006). Further evidence that *R* gene evolution is adaptive is given by observations of high amino acid substitution rates in leucine-rich-repeat (LRR) domains of *R* gene protein products (e.g., Rose et al., 2004); these regions are believed to be directly involved in pathogen recognition (Jones and Jones, 1997).

Molecular population genetic studies of *R* genes to date have characterized species-wide levels of *R* gene variation. However, little is known about geographic patterns of diversity within and among populations, primarily because within-population sampling of *R* genes has not often been carried out. Attempts to understand *R* gene evolutionary dynamics can benefit from within- and between-population sampling for several reasons. Species-wide samples can occasionally obscure the type of selection responsible for patterns of variation; for example, balancing selection detected at the species level may be due to local adaptation (Nordborg and Innan, 2003; Wright and Gaut, 2005). Additionally, many evolutionary models of plant–pathogen interactions, such as frequency-dependent selection and the Red-queen's hypothesis, were developed within a population-level framework (Haldane, 1949; Hamilton, 1980). It is also likely that local population processes, such as gene flow, drift, and mating system, have an impact on *R* gene evolution, and detecting these processes requires characterizing diversity within and between populations.

To characterize geographic patterns of *R* gene variation among plant populations, I studied the distribution of *Cf-2* homologs in natural populations of *Solanum pimpinellifolium* L. [formerly *Lycopersicon pimpinellifolium* (L.) Mill.], a wild relative of cultivated tomato. The *Cf-2* locus was originally identified as comprising two paralogous genes, *Cf-2.1* and *Cf-2.2*, both conferring resistance to strains of *Cladosporium fulvum* (causal agent of leaf mold) carrying an *Avr2* avirulence gene (Dixon et al., 1996). A subsequent survey of *Cf-2* variation across the species detected 26 closely related homologs (maximum of 6.42% nucleotide divergence), establishing that *Cf-2.1* and *Cf-2.2* are part of a multigene family (hereafter referred to as the *Hcr2-p* family) (Caicedo and Schaal, 2004a). The occurrence of *R* genes in families is common and can have an impact on their evolutionary dynamics (Hulbert et al., 2001). The 26 *Hcr2-p* homologs display length variation due primarily to variation in the number of LRR-coding units within each gene and can be classified into nine different size classes according to length (Fig. 1); within size-classes, homologs differ from each other by one or a few single nucleotide polymorphisms (SNPs) (Caicedo and Schaal, 2004a) (Appendices S1, S2; see Supplemental Data with the online version of this article). *Solanum pimpinellifolium* individuals vary extensively in the number of *Hcr2-p* homologs they carry, with Southern blots results suggesting 1–5 genes per individual (Caicedo and Schaal, 2004a). Although the function of *Hcr2-p* genes other than *Cf-2.1* and *Cf-2.2* is currently undemonstrated, sequence similarities suggest that most are functional *R* genes. Four size classes, however, contain frame-shift mutations, three of which cause loss of a membrane spanning domain (Fig. 1) (Caicedo and Schaal, 2004a); given the substantial length of putative protein products of these size classes, it is possible that the truncated genes are still functional *R* genes or have some other functional role.

[1] Manuscript received 2 May 2007; revision accepted 31 January 2008.

The author thanks P. Kover, K. Mather, R. Moore, K. Olsen, and A. Porter for useful comments on this manuscript. This research was supported in part by National Science Foundation Dissertation Improvement Grant DEB-0073082.

[2] E-mail: caicedo@bio.umass.edu

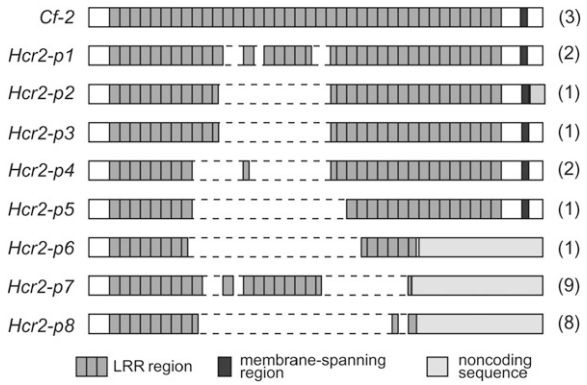

Fig. 1. Schematic alignment of the open reading frame (ORF) of the nine characterized *Hcr2-p* size classes in *Solanum pimpinellifolium*. Numbers in parenthesis refer to the number of different homologs found within each size class. In all *Hcr2-p* homologs, a single intron occurs downstream of the ORF and was not sequenced in its entirety.

Species-wide analyses of *Hcr2-p* sequence diversity suggest that selection has played a role in the evolution of the gene family. Patterns of amino acid substitution are consistent with purifying selection in the 5′ LRR-coding portion of the genes and positive selection on some amino acid residues in the 3′ region (Caicedo and Schaal, 2004a). Evolutionary relationships among homologs also suggest that balancing selection has shaped species-wide patterns of diversity (Caicedo and Schaal, 2004a). In this paper, I describe the geographic distribution of *Hcr2-p* diversity across a portion of *S. pimpinellifolium*'s native range, and use patterns of population variation to gain further understanding of the evolutionary processes affecting the evolution of this *R* gene family in the wild.

MATERIALS AND METHODS

A total of 16 *S. pimpinellifolium* populations were sampled along the northern coast of Peru as described in Caicedo and Schaal (2004b) (Fig. 2). From 120 collected plants, 185 *Hcr2-p* sequences were successfully isolated and sequenced using conditions described in Caicedo and Schaal (2004a). Due to PCR bias against efficient amplification of very large bands, the obtained products are a large, but likely not exhaustive, sample of *Hcr2-p* homolog diversity. These *Hcr2-p* homologs are known to be a part of a cohesive gene family because levels of divergence among them (0–6.4%) are lower than those compared the closest known paralog, *Hcr2–2A* (~11.5%) (Caicedo and Schaal, 2004a; Dixon et al., 1996). Sequences were classified as corresponding to one of 26 possible homologs and one of nine possible size classes. Across individuals, homolog combinations gave rise to 40 different genotypes (Appendix S3; see Supplemental Data with the online version of this article).

The variation in *Hcr2-p* gene number among *S. pimpinellifolium* individuals obscures relationships of orthology and paralogy between these genes, and assumptions cannot be made about the zygosity value represented by each sequence. Because traditional population genetic descriptors of variation are not suitable for gene families, levels of diversity were assessed with alternative measures. First, the frequency of homologous genes (of which there are 26 types) was determined in each population, without regard to the distribution within individuals. For each population, an index of diversity was calculated as the probability that two randomly chosen homologs are different from each other ($1 - \sum p_i^2$, where p_i is the frequency of the i^{th} homolog in the population); this measure is referred to as the homology diversity index. Second, an index of size-class diversity was calculated as above, but considering p_i as the frequency of the i^{th} homolog size class (i.e., the nine size classes determined by the length of the LRR coding region) in each population. Although these measures seem similar to the traditional population genetics descriptor, haplotype diversity (Nei, 1987: eq. 8.4), they should not be interpreted as such, as the categories are not alleles. Geographic correlations between both diversity indexes and latitude were examined using a nonparameteric Spearman rank correlation coefficient. Differences in levels of diversity among geographic regions were assessed with nonparametric Wilcoxon rank-sum tests.

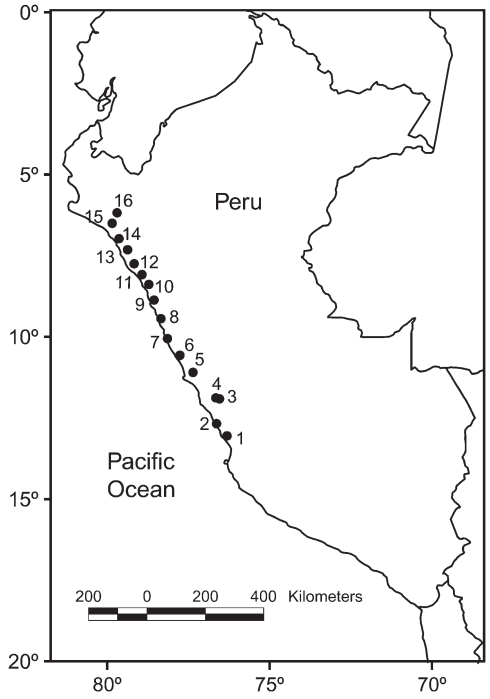

Fig. 2. *Solanum pimpinellifolium* sampling localities along the Peruvian coast.

Diversity indexes based on multihomolog *Hcr2-p* genotypes were also calculated for each population and examined in a geographic context. To determine how genotype variation is distributed among and between populations, individual genotypes were further assessed by scoring each *Hcr2-p* size class as presence–absence data; genotypes were thus made up of combinations of one of two possible states in nine separate categories. An analysis of molecular variance (AMOVA) (Excoffier et al., 1992) using pairwise differences between genotypes was carried out with the program Arlequin (Schneider et al., 2000), using 16 000 permutations. An AMOVA was not carried out on multihomolog genotypes, as homologs partition variation too finely to allow appreciable sharing of genotypes across populations.

RESULTS AND DISCUSSION

The geographic distribution of Hcr2-p variation—Populations of *S. pimpinellifolium* along the sampled range were found to contain many different combinations of *Hcr2-p* homologs and *Hcr2-p* size classes (Table 1). Levels of homolog diversity and size-class diversity vary greatly between populations, with homolog diversity values ranging from 0.31 to 0.87, and size-class diversity ranging from 0 to 0.76 (Table 2, Fig. 3). *Hcr2-p* diversity values are significantly different in southern and northern populations (Wilcoxon rank sum test: $P_{homolog} = 0.040$, $P_{size\text{-}class} = 0.018$), and tend to be higher in northern Peru (Table 2, Figs. 2, 3). Both homolog diversity and size class diversity were found to be positively and significantly correlated with population latitudinal position on a south-north axis ($\rho_{homolog} = 0.7607$, $P < 0.001$; $\rho_{size\text{-}class} = 0.7703$, $P < 0.001$), as determined by a nonparametric Spearman rank correlation coefficient.

Different *Hcr2-p* size classes code for protein products that differ in the number of LRRs. Within size classes, however, some homologs differ by synonymous substitutions and code

TABLE 1. Geographic distribution and frequencies of *Hcr2-p* homologs in sampled *Solanum pimpinellifolium* populations.

					Cf-2[d]			***Hcr2-p1***		***Hcr2-p2***	***Hcr2-p3***	***Hcr2-p4***		***Hcr2-p5***	***Hcr2-p6***	***Hcr2-p7***								***Hcr2-p8***					
Population[a]	Latitude (degrees South)	*N*[b]	Homolog totals	Size-class totals	*Cf-2.1*[c]	*Cf-2.2*	*Cf-2.3*	*Hcr2-p1.1*	*Hcr2-p1.2*	*Hcr2-p2*	*Hcr2-p3*	*Hcr2-p4.1*	*Hcr2-p4.2*	*Hcr2-p5*	*Hcr2-p6*	*Hcr2-p7.1*	*Hcr2-p7.2*	*Hcr2-p7.3*	*Hcr2-p7.4*	*Hcr2-p7.5*	*Hcr2-p7.6*	*Hcr2-p7.7*	*Hcr2-p7.8*	*Hcr2-p7.9*	*Hcr2-p8.1*	*Hcr2-p8.2*	*Hcr2-p8.3*	*Hcr2-p8.4*	*Hcr2-p8.5*	*Hcr2-p8.6*
1 Roma, Lima	13°03′18″	10	4	1															7	2		1	1						
2 Mala, Lima	12°40′57″	10	3	2	1	1					9																		
3 Cocachacra, Lima	11°54′53″	8	2	2													6												8
4 Sta. Eulalia, Lima	11°52′49″	8	2	1	8	8																							
5 Huacho, Lima[b]	11°06′02″	1	1	1																			1						
6 Pativilca, Lima	10°34′17″	8	7	4		2	2							2				3		1							1	3	
7 Huarmey, Ancash	10°03′57″	6	3	2		5	5			1																			
8 Casma, Ancash	09°27′11″	7	4	3		1	1			1					6														
9 Vinzos, Ancash	08°52′39″	5	2	2						5																	2		
10 Virú, La Libertad	08°23′46″	7	5	4					1	5					1												1		1
11 Laredo, La Libertad	08°05′09″	6	3	2					4																		5	1	
12 Chocope, La Libertad	07°45′39″	9	8	5		3	2				3			3		1										3	1		1
13 Jequetepeque, La Libertad	07°19′14″	8	4	3				2									4								3	2			
14 Ucupe, Lambayeque	06°59′11″	9	4	3						5							3									2		4	
15 Túcume, Lambayeque	06°30′36″	9	8	4	2	1	1					2	2				2									3			2
16 Tzu Chi, Lambayeque	06°11′14″	9	9	5	2	1	1			2	1										2			1		3			2
Homolog totals					13	22	12	2	5	19	13	2	2	5	7	1	15	3	7	3	2	1	2	1	3	13	10	3	5	14
Size-class totals							47		7	19	13		4	5	7									35						48

[a]Populations are ordered south to north; population numbers correspond to Caicedo and Schaal (2004a) and Fig. 2.
[b]Number of individuals sequenced per population
[c]Headings in italics refer to homologous genes.
[d]Headings in bold italics refer to homolog size classes.

for identical proteins (online Appendix S2; see Supplemental Data with the online version of this article). To determine if there is variation among populations for Hcr2-p proteins, I calculated *Hcr2-p* homolog diversity taking protein-coding variants in account (Table 2). Northern and southern populations differ significantly in protein variant diversity (Wilcoxon rank sum test: $P = 0.043$), and, as observed for *Hcr2-p* homolog and size-classes, protein variant diversity is strongly correlated with population latitudinal position ($\rho = 0.8061$, $P < 0.001$). Taken together, these results all suggest that the diversity of *Hcr2-p* genes increases in northern Peruvian populations.

Hcr2-p genotype diversity also varies greatly among *S. pimpinellifolium* populations and tends toward higher values in the north ($\rho_{genotype} = 0.7771$, $P < 0.001$; Table 2); this indicates that individuals in northern Peruvian populations tend to be more heterogeneous for *Hcr2-p* gene content. The geographic distribution of *Hcr2-p* diversity was further examined with an AMOVA based on class-size genotypes within individuals. Genotype diversity is almost equally distributed between (46.76%) and within (53.24%, $F_{ST} = 0.46$, $P < 0.0001$) populations in the complete population sample. Given that several populations, display very little variation, especially in the southern portion of the sample (Table 1), the analysis was repeated for the northern and southern geographic area separately. In the seven southern populations, most of the genetic variation occurred between populations (71.17%, $F_{ST} = 0.22$, $P < 0.0001$), while in the eight northern populations, most of the variation occured within populations (77.79%, $F_{ST} = 0.71$, $P < 0.0001$). No geographic break exists between these sets of populations, but these results, coupled with the observed correlations between latitude and diversity, suggest that there is an *Hcr2-p* diversity cline among sampled Peruvian *S. pimpinellifolium* populations.

The impact of demography on Hcr2-p variation—*Solanum pimpinellifolium* is a self-compatible plant, but is known to vary in breeding system; previous allozyme (Rick et al., 1977) and noncoding DNA sequence based studies (Caicedo and Schaal, 2004b) have shown that populations in northern Peru have lower inbreeding coefficients and higher levels of genetic diversity. Northern Peruvian populations also tend to have larger flowers and a greater degree of stigma exsertion (Rick et al., 1977). This cline in levels of diversity and outcrossing is believed to be due to *S. pimpinellifolium*'s sequential colonization of the Peruvian coast from a center of origin in northern Peru coupled with increased selfing favored in new populations (Rick, 1984; Caicedo and Schaal, 2004b). *S. pimpinellifolium* grows in river valleys found along the dry coastal lowlands of Peru, and the intervening desert regions likely limit gene flow between populations. It is thus possible that the observed geographic patterns of *Hcr2-p* diversity are a consequence of the demographic history and mating system variation in *S. pimpinellifolium*.

Hcr2-p population diversity values were compared to haplotype diversity measures based on the second intron of the vacuolar invertase gene, a neutral marker formerly shown to be affected by the demographic history of *S. pimpinellifolium* (Caicedo and Schaal, 2004b). The vacuolar invertase gene corresponds to *T1V1* (Klann et al., 1992; Chetelat et al., 1993; Elliott et al., 1993; R. Chetelat, University of California-Davis, personal communication), which has been mapped to chromosome 3. The two mapped *Cf-2* genes lie on chromosome 6 and

TABLE 2. Diversity statistics for *Hcr2-p* homologs in the populations of *Solanum pimpinellifolium* sampled.

Population[a]	Type of diversity				
	Homolog	Size-class	Genotype	Protein variant	Putative functional homolog[b]
1 Roma, Lima	0.55	0.00	0.48	0.31	na
2 Mala, Lima	0.31	0.18	0.18	0.31	0.31
3 Cocachacra, Lima	0.49	0.49	0.38	0.49	na
4 Sta. Eulalia, Lima	0.50	0.00	0.00	0.50	0.50
5 Huacho, Lima[b]	—	—	—	—	—
6 Pativilca, Lima	0.84	0.72	0.72	0.84	0.67
7 Huarmey, Ancash	0.58	0.28	0.28	0.58	0.58
8 Casma, Ancash	0.52	0.41	0.45	0.52	0.67
9 Vinzos, Ancash	0.41	0.41	0.48	0.41	0.00
10 Virú, La Libertad	0.64	0.62	0.61	0.64	0.28
11 Laredo, La Libertad	0.58	0.48	0.50	0.58	0.00
12 Chocope, La Libertad	0.85	0.76	0.84	0.83	0.74
13 Jequetepeque, La Libertad	0.73	0.63	0.72	0.73	0.00
14 Ucupe, Lambayeque	0.72	0.64	0.72	0.72	0.00
15 Túcume, Lambayeque	0.86	0.72	0.81	0.86	0.78
16 Tzu Chi, Lambayeque	0.87	0.76	0.84	0.87	0.78

[a]Diversity indices were not calculated for population 5 due to low sample size.
[b]Putative pseudogenes were excluded from this analysis; na = only putative pseudogenes were found.

are physically unlinked from *T1V1*. As observed for the *Hcr2-p* gene family, distribution of *T1V1* genetic diversity varies according to geographical area: in the seven southernmost sampled populations, most genetic variation occurs between populations (82.88%; $F_{ST} = 0.83$, $P < 0.001$), while for the eight northernmost populations most of the variation occurs within populations (74.74%; $F_{ST} = 0.25$, $P < 0.001$) (Caicedo and Schaal, 2004b). Likewise, *T1V1* haplotype diversity (*h*) increases in northern populations, and is positively correlated with latitudinal population position (Caicedo and Schaal, 2004b). Spearman's rank correlation coefficients between *T1V1 h* values and *Hcr2-p* homolog and size-class diversity values were found to be positive and significant ($\rho_{(h;\text{homolog div.})} = 0.7455$ $P < 0.005$; $\rho_{(h;\text{size class div.})} = 0.6063$ $P < 0.02$). *Hcr2-p* diversity in *S. pimpinellifolium* thus follows the same geographical trend as that of an unlinked neutral marker, as well as previously studied allozymes (Rick et al., 1977).

The geographic diversity cline observed in *Hcr2-p* diversity is interesting for two reasons. First, it suggests that demographic history and local mating system may play a substantial role in of the amount and apportionment of variation for genes putatively involved in resistance to disease; this observation is unexpected, as natural selection is believed to play a crucial role in *R* gene evolution, and diversity patterns at *R* gene loci might not necessarily reflect population demographic processes. Second, it indicates that mating system may affect genetic diversity not only at single loci, but also at the level of gene families. While this may not be surprising, it should be noted that the mechanisms by which mating system affects diversity in each of these cases might be different. Increased inbreeding decreases effective population size through increased homozygosity, leading to a greater probability of loss of alleles through drift (Templeton, 2006); inbreeding should affect all loci (whether single-copy or in gene families) in the genome equally. Increased levels of outcrossing lead to increased heterozygosity, which in multigene families can promote unequal recombination, and the generation of diversity through the creation of new family members or multigenic haplotypes. This effect of outcrossing should be particular to multigene families. Current *Hcr2-p* diversity in sampled *S. pimpinellifolium* populations may thus have been shaped by both past sequential bottlenecks as the species spread south, as well as the local mating system.

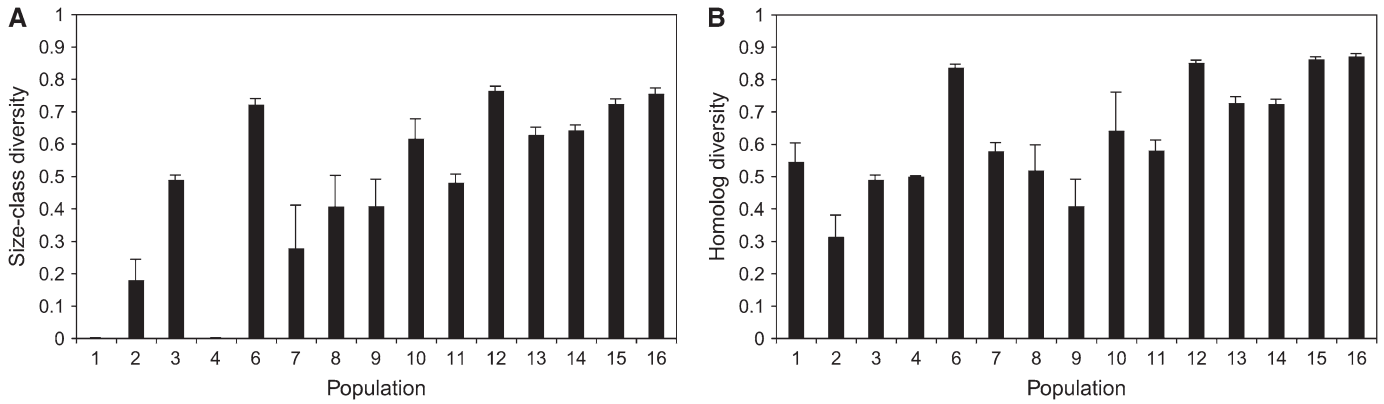

Fig. 3. Measures of (A) *Hcr2-p* homolog and (B) size class diversity in sampled Peruvian populations of *Solanum pimpinellifolium*. Populations are ordered south to north and population numbers correspond to those in Table 1 and Fig. 2. Population 5 was excluded due to low sample size. Error bars correspond to jackknife standard deviations.

The role of selection in Hcr2-p homolog evolution—Species-wide patterns of nucleotide diversity within the *Hcr2-p* family have suggested that there is selection against amino acid substitutions in the 5′ LRR-coding region of the genes, while positive selection has possibly occurred on some amino acid residues at the 3′ portion of the genes (Caicedo and Schaal, 2004a). Homolog nucleotide diversity is also significantly higher at the 5′ end of the genes, suggesting either a high mutation rate or maintenance of species-wide diversity by selection. The DNA sequence of the *Hcr2-p* genes thus seems to be determined by various types of natural selection. The geographic distribution of various levels of *Hcr2-p* diversity, on the other hand, is correlated with latitude and with geographic patterns of diversity observed for neutral loci. In light of the molecular and geographical characterization of *Hcr2-p* diversity, it would be of interest to determine whether demography or selection has played a larger role in shaping *Hcr2-p* family evolution.

Although the evidence for a latitudinal cline in *Hcr2-p* diversity is strong, local factors may nevertheless have an effect, causing some populations to deviate from the geographic trend. Among the populations sampled, for example, population 6 (Pativilca) has very high levels of *Hcr2-p* diversity for its latitude (Table 2, Fig. 3). Even more striking is that the *T1V1* inbreeding coefficient (F_{IS}) of this population is 1 (Caicedo and Schaal, 2004b). Thus, in population 6 there is no correlation between diversity at a neutral locus and *Hcr2-p* diversity, suggesting that local factors such as selection may be playing a role.

The best candidate for a selective agent on *Hcr2-p* homologs in the wild is *C. fulvum*, the fungus involved in the resistance response mediated by the *Cf-2.1* and *Cf-2.2* genes. Interestingly, in the sampled populations, one of the most common and widely distributed size classes is the one containing *Cf-2.1* and *Cf-2.2*, which was observed in eight of the populations across the sampled range (Table 1). However, two other widely distributed size classes are *Hcr2-p7* (nine populations) and *Hcr2-p8* (10 populations, Table 1), which code for truncated proteins and may be pseudogenes (Fig. 1). Surprisingly, two populations (Roma (1) and Cocachacra (3)) contained only *Hcr2-p7* and *-p8* genes. Additionally, if all genes coding for proteins lacking a transmembrane domain (*Hcr2-p6, -p7, -p8*) are ignored, an additional four populations, all from the northern half of the sampled range, display no homolog diversity (Tables 1, 2). Thus, the cline in diversity is only evident if putatively functional and nonfunctional genes are considered ($\rho_{\text{putative functional}} = 0.1367$ $P = 0.328$). It should be noted that it is possible that *Hcr2-p6, -p7, -p8* may still perform some function, as their predicted protein products range from 282 to 487 amino acids in length. However, if these genes are indeed nonfunctional, their wide geographic range and maintenance at high frequency levels in some populations, could suggest that demographic factors have played an important role in the distribution of *Hcr2-p* geographic variation.

Regardless of the exact role of each *Hcr2-p* gene, *C. fulvum* seems to have played a role in the evolution of wild tomatoes. Members of the tomato clade are the only known hosts of *C. fulvum* (Bond, 1938). Sources of *C. fulvum* resistance have been identified in many wild tomato species (Kerr and Bailey, 1964), and various *Cf*-type genes with demonstrated resistance function have been cloned from at least three species (Jones et al., 1994; Dixon et al., 1996, 1998; Thomas et al., 1997). Additionally, (Van der Hoorn et al., 2001) have found widespread recognition of a *C. fulvum* avirulence protein, Avr9, in accessions from *S. pimpinellifolium*'s geographic range. Thus, it is reasonable to believe that *C. fulvum* acts as a source of selective pressure on *S. pimpinellifolium*.

Unfortunately, little is known of the origin of *C. fulvum* and its behavior in natural populations. Most knowledge of the pathogen comes from observations of greenhouse tomatoes, and it has only rarely been reported in cultivated fields (Alexander, 1934; Esquinas-Alcazar, 1981; Joosten and de Wit, 1999). Because the tomato clade has a South American origin, *C. fulvum* is assumed to originate in South America (Joosten and de Wit, 1999). However, no studies of the pathogen have been carried out in its presumed native range. Interestingly, *S. pimpinellifolium* thrives in drier conditions than would be expected to favor *C. fulvum* infection and sporulation (Langford, 1937). Attempts to infect *S. lycopersicum* (cultivated tomato) control plants in the laboratory were only successful at humidity levels higher than those observed in collection localities (personal observation). Repeated attempted evaluation of *Hcr2-p* gene function in *S. pimpinellifolium* accessions with *C. fulvum* race 5, 9 (carrying the *Avr 2* avirulence gene) in the laboratory have been unsuccessful, possibly due to lack of isogenic fungal strains, leading to *S. pimpinellifolium* recognition of uncharacterized avirulence factors in the pathogen (personal observation). Thus, many questions still remain about *C. fulvum*'s action as a selective agent in the wild.

Based on the available data, three possibilities regarding the role of *C. fulvum* in the evolution and geographic distribution of *Hcr2-p* diversity come to mind. It is possible that molecular variation at the *Hcr2-p* genes has been shaped by past selection preceding speciation. In this case, the geographic distribution of *Hcr2-p* genes in *S. pimpinellifolium* would be due entirely to demographic processes and would reflect lack of current selection on these genes. Lack of current selection might also explain the high frequency of possible pseudogenes in the *Hcr2-p* family. A second possibility is that under periodic climatic conditions *C. fulvum* can sporadically attack *S. pimpinellifolium*, and thus maintain *Cf-2* homolog variation; if selection is irregular or not very strong, demographic factors may be dominant in shaping the geographic distribution of diversity, although the DNA sequence of the genes may be shaped by selection. A last possibility is that selection pressure on *Hcr2-p* function (whether it is resistance against *C. fulvum*, resistance to other pathogens, or an unknown function) has itself a clinal distribution; under this scenario, geographic diversity patterns could in fact be due to selection, and selection could be contributing to population variation in outcrossing. The future integration of field pathogen studies with *R* gene population studies should help shed further light on the various forces shaping the evolution of *R* genes in the wild.

LITERATURE CITED

ALEXANDER, L. J. 1934. Leaf mold resistance in tomato. Ohio Agricultural Experiment Station, Wooster, Ohio, USA.

BAKKER, E. G., C. TOOMAJIAN, M. KREITMAN, AND J. BERGELSON. 2006. A genome-wide survey of *R* gene polymorphisms in *Arabidopsis*. *Plant Cell* 18: 1803–1818.

BERGELSON, J., M. KREITMAN, E. A. STAHL, AND D. TIAN. 2001. Evolutionary dynamics of *R*-genes. *Science* 292: 2281–2285.

BOND, T. E. T. 1938. Infection experiments with *Cladosporium fulvum* Cooke and related species. *Annals of Applied Biology* 25: 277–307.

CAICEDO, A. L., AND B. A. SCHAAL. 2004a. Heterogeneous evolutionary processes affect *R* gene diversity in natural populations of *Solanum pimpinellifolium. Proceedings of the National Academy of Sciences, USA* 101: 17444–17449.

CAICEDO, A. L., AND B. A. SCHAAL. 2004b. Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology* 13: 1871–1882.

CAICEDO, A. L., B. A. SCHAAL, AND B. N. KUNKEL. 1999. Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, USA* 96: 302–306.

CHETELAT, R. T., E. KLANN, J. W. DEVERNA, S. YELLE, AND A. B. BENNETT. 1993. Inheritance and genetic mapping of fruit sucrose accumulation in *Lycopersicon chmielewskii. Plant Journal* 4: 643–650.

DIXON, M. S., K. HATZIXANTHIS, D. A. JONES, K. HARRISON, AND J. D. G. JONES. 1998. The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number. *Plant Cell* 10: 1915–1925.

DIXON, M. S., D. A. JONES, J. S. KEDDIE, C. M. THOMAS, K. HARRISON, AND J. D. G. JONES. 1996. The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* 84: 451–459.

ELLIOTT, K. J., W. O. BUTLER, C. D. DICKINSON, Y. KONNO, T. S. VEDVICK, L. FITZMAURICE, AND T. E. MIRKOV. 1993. Isolation and characterization of fruit vacuolar invertase genes from 2 tomato species and temporal differences in mRNA levels during fruit ripening. *Plant Molecular Biology* 21: 515–524.

ELLIS, J., P. DODDS, AND T. PRYOR. 2000. Structure, function and evolution of plant disease resistance genes. *Current Opinion in Plant Biology* 3: 278–284.

ESQUINAS-ALCAZAR, J. T. 1981. Genetic resources of tomatoes and wild relatives. International Board for Plant Genetic Resources, Rome, Italy.

EXCOFFIER, L., P. E. SMOUSE, AND J. M. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haploytpes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

HALDANE, J. B. S. 1949. Disease and evolution. *La Ricerca Scientifica* 19 (Supplement): 68–79.

HAMILTON, W. D. 1980. Sex versus non-sex versus parasite. *Oikos* 35: 282–290.

HULBERT, S. H., C. A. WEBB, S. M. SMITH, AND Q. SUN. 2001. Resistance gene complexes: Evolution and utilization. *Annual Review of Phytopathology* 39: 285–312.

JONES, D. A., AND J. D. G. JONES. 1997. The role of leucine-rich repeat proteins in plant defences. *Advances in Botanical Research* 24: 89–167.

JONES, D. A., C. M. THOMAS, K. E. HAMMOND-KOSACK, P. J. BALINT-KURTI, AND J. D. G. JONES. 1994. Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266: 789–793.

JOOSTEN, M., AND P. DE WIT. 1999. The tomato–*Cladosporium fulvum* interaction: A versatile experimental system to study plant–pathogen interactions. *Annual Review of Phytopathology* 37: 335–367.

KERR, E. A., AND D. L. BAILEY. 1964. Resistance to *Cladosporium fulvum* Cke. obtained from wild species of tomato. *Canadian Journal of Botany* 42: 1541–1554.

KLANN, E., S. YELLE, AND A. B. BENNETT. 1992. Tomato acid invertase complementary DNA. *Plant Physiology* 99: 351–353.

LANGFORD, A. N. 1937. The parasitism of *Cladosporium fulvum* Cooke and the genetics of resistance to it. *The Canadian Journal of Research* 15: 108–128.

MAURICIO, R., E. A. STAHL, T. KORVES, D. C. TIAN, M. KREITMAN, AND J. BERGELSON. 2003. Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. *Genetics* 163: 735–746.

MEYERS, B. C., S. KAUSHIK, AND R. S. NANDETY. 2005. Evolving disease resistance genes. *Current Opinion in Plant Biology* 8: 129–134.

MICHELMORE, R. 2000. Genomic approaches to plant disease resistance. *Current Opinion in Plant Biology* 3: 125–131.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.

NORDBORG, M., AND H. INNAN. 2003. The genealogy of sequences containing multiple sites subject to strong selection in a subdivided population. *Genetics* 163: 1201–1213.

RICK, C. M. 1983. Evolution of mating systems: Evidence from allozyme variation. *In* V. L. Chapra, B. C. Joshi, R. P. Sharma, and H. C. Bansal [eds.], Genetics: New frontiers (Proceedings of VX International Congress of Genetics), vol. 4, 215–221. Bowker, Epping, Essex, UK.

RICK, C. M., J. F. FOBES, AND M. HOLLE. 1977. Genetic variation in *Lycopersicon pimpinellifolium*: evidence of evolutionary change in mating systems. *Plant Systematics and Evolution* 127: 139–170.

ROSE, L. E., P. D. BITTNER-EDDY, C. H. LANGLEY, E. B. HOLUB, R. W. MICHELMORE, AND J. L. BEYNON. 2004. The maintenance of extreme amino acid diversity at the disease resistance gene, *RPP13*, in *Arabidopsis thaliana*. *Genetics* 166: 1517–1527.

SCHNEIDER, S., D. ROESSLI, AND L. EXCOFFIER. 2000. Arlequin, version 2.000: a software for population genetic analysis, Genetics and Biometry Laboratory, University of Geneva, Geneva.

STAHL, E. A., G. DWYER, R. MAURICIO, M. KREITMAN, AND J. BERGELSON. 1999. Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* 400: 667–671.

TEMPLETON, A. R. 2006. Population genetics and microevolutionary theory. Wiley, Hoboken, New Jersey, USA.

THOMAS, C. M., D. A. JONES, M. PARNISKE, K. HARRISON, P. J. BALINT-KURTI, K. HATZIXANTHIS, AND J. D. G. JONES. 1997. Characterization of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in *Cf-4* and *Cf-9*. *Plant Cell* 9: 2209–2224.

VAN DER HOORN, R. A. L., M. KRUIJT, R. ROTH, B. F. BRANDWAGT, M. JOOSTEN, AND P. DE WIT. 2001. Intragenic recombination generated two distinct *Cf* genes that mediate AVR9 recognition in the natural population of *Lycopersicon pimpinellifolium*. *Proceedings of the National Academy of Sciences, USA* 98: 10493–10498.

WRIGHT, S. I., AND B. S. GAUT. 2004. Molecular population genetics and the search for adaptive evolution in plants. *Molecular Biology and Evolution* 22: 506–519.