Ancient isolation and independent evolution of the three clonal lineages of the exotic sudden oak death pathogen *Phytophthora ramorum*

E. M. GOSS,* I. CARBONE† and N. J. GRÜNWALD*

Horticultural Crops Research Laboratory, USDA ARS, 3420 NW Orchard Avenue, Corvallis, OR 97330, USA, †Center for Integrated Fungal Research, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695-7244, USA

Abstract

The genus *Phytophthora* includes some of the most destructive plant pathogens affecting agricultural and native ecosystems and is responsible for a number of recent emerging and re-emerging infectious diseases of plants. Sudden oak death, caused by the exotic pathogen *P. ramorum*, has caused extensive mortality of oaks and tanoaks in Northern California, and has brought economic losses to US and European nurseries as well due to its infection of common ornamental plants. In its known range, *P. ramorum* occurs as three distinct clonal lineages. We inferred the evolutionary history of *P. ramorum* from nuclear sequence data using coalescent-based approaches. We found that the three lineages have been diverging for at least 11% of their history, an evolutionarily significant amount of time estimated to be on the order of 165 000 to 500 000 years. There was also strong evidence for historical recombination between the lineages, indicating that the ancestors of the *P. ramorum* lineages were members of a sexually reproducing population. Due to this recombination, the ages of the lineages varied within and between loci, but coalescent analyses suggested that the European lineage may be older than the North American lineages. The divergence of the three clonal lineages of *P. ramorum* supports a scenario in which the three lineages originated from different geographic locations that were sufficiently isolated from each other to allow independent evolution prior to introduction to North America and Europe. It is thus probable that the emergence of *P. ramorum* in North America and Europe was the result of three independent migration events.

Keywords: coalescent, emerging disease, oomycete, plant disease, recombination, RXLR-class effector

Received 18 July 2008; revision received 8 December 2008; accepted 18 December 2008

Introduction

Outbreaks of emerging and re-emerging infectious diseases are increasing in incidence (Morens *et al*. 2004; Jones *et al*. 2008). Much of the attention is on diseases of humans, but similar patterns can be observed among emerging diseases of plants, both on wild and cultivated species (Brown & Hovmoller 2002; Anderson *et al*. 2004; King *et al*. 2006). Some of the most destructive plant pathogens affecting agricultural and native ecosystems can be found in the genus *Phytophthora*, such as the potato late blight pathogen

P. infestans, the soybean pathogen *P. sojae*, and *P. cinnamomi*, which has killed large swaths of native forest in Australia and also infects many food crops. *P. infestans* is a classic example of an emerging and re-emerging pathogen (Fry & Goodwin 1997). In the 19th century, its initial migration from South or Central America to the USA and Europe resulted in explosive epidemics and the Irish Potato Famine (Fry *et al*. 1993; Goodwin *et al*. 1994; May & Ristaino 2004; Gomez-Alpizar *et al*. 2007). *P. infestans* has re-emerged due to recent migrations of the pathogen from Central Mexico, a possible centre of origin and the only location where the pathogen was known to reproduce sexually before the 1980s (Grünwald *et al*. 2001; Flier *et al*. 2003; Grünwald & Flier 2005), to export markets where historical clonal lineages were displaced by more virulent ones (Fry

Correspondence: Niklaus J. Grünwald. Fax: 541-738-4025; E-mail: Nik.Grunwald@ars.usda.gov

et al. 1992, 1993). *Phytophthora* are oomycetes (water moulds), diploid eukaryotes in the kingdom Stramenopila, and are more closely related to algae than to the fungi with which the oomycetes share life-history characteristics (Baldauf 2003).

A recently emerging *Phytophthora* is *P. ramorum*, the causal agent of sudden oak death, which is responsible for extensive mortality of coast live oak (*Quercus agrifolia*) and tanoak (*Lithocarpus densiflorus*) in northwest California (Rizzo *et al*. 2002, 2005; Grünwald *et al*. 2008b). *P. ramorum* also causes foliar lesions and twig dieback (Ramorum blight) on hosts in over 40 plant genera including many common ornamentals (Rizzo *et al*. 2005). Its infestation of nursery stock has provided a mechanism for long-distance dispersal and its quarantine has led to large economic losses by the nursery industry in North America and Europe. Its symptoms were first noticed in California in 1994–1995 (Garbelotto *et al*. 2001; Rizzo *et al*. 2002) and around the same time on nursery stock in Europe (Werres *et al*. 2001).

The available genetic data provide strong evidence that *P. ramorum* is an exotic pathogen that appears to have been introduced from an unknown origin at least three times (Rizzo *et al*. 2005; Ivors *et al*. 2006; Prospero *et al*. 2007; Mascheretti *et al*. 2008). There are three clonal lineages of *P. ramorum*, within which genetic variation has only been observed at rapidly evolving microsatellites (Ivors *et al*. 2006). Support for grouping isolates into three lineages comes from AFLP, microsatellite, and mitochondrial sequence data and is consistent across markers (Ivors *et al*. 2004, 2006; Prospero *et al*. 2007; Martin 2008). The EU1 lineage is responsible for all infestations in Europe but has also been found in nurseries on the West Coast of the USA. The NA1 genotype is the cause of the wildland epidemics in California and the southwest corner of Oregon and is also seen in nurseries (Prospero *et al*. 2007). The third genotype, NA2, has only been observed in a limited number of nurseries. *P. ramorum* is self-sterile, such that sexual reproduction must occur between individuals of two different mating types. The EU1 genotype is largely the A1 mating type and all tested NA1 and NA2 isolates have been A2 (Ivors *et al*. 2006; N. Grünwald, unpublished). The two mating types have been brought together by the nursery trade (Grünwald *et al*. 2008a) and have both been detected in a California creek (Frankel 2008). Yet oospores are not readily produced from crosses of the *P. ramorum* A1 and A2 mating types in the laboratory, leading to questions as to whether the sexual reproduction system is functional (Brasier & Kirk 2004; Werres & Kaminski 2005). The closest known relatives of *P. ramorum* are *P. lateralis*, *P. foliorum*, and *P. hibernalis*, which together make up *Phytophthora* clade 8c (Blair *et al*. 2008). Both *P. lateralis* and *P. foliorum* are also introduced pathogens from unknown origins (Erwin & Ribeiro 1996; Donahoo *et al*. 2006) and *P. hibernalis* is similarly

an exotic in US nurseries (Blomquist *et al*. 2005). Unlike *P. ramorum*, these species are self-fertile.

Here we examined the variation in five nuclear genes from a diverse sample of *P. ramorum* isolates in order to better understand the evolutionary history of the three clonal lineages and the relationships among them. We show that there was no variation in these genes among isolates within the lineages, that recombination has played a role in structuring the variation among lineages, and that the lineages experienced an extended period of isolation prior to introduction. We argue that this implies that the lineages were introduced from three geographically distinct and independently evolving populations that are descendents of an ancestral sexual population.

Materials and methods

Isolates

Phytophthora isolates were maintained on cleared 10% V8 agar medium (100 mL V8 juice; 2 g $CaCO_3$; 30 mg/L β-sitosterol (EMD Chemicals, Incorporated); 15 g agar; 900 mL deionized water) in a 20 °C incubator in the dark (Grünwald *et al*. 2008c). All isolates were maintained following the standard operating procedures associated with corresponding USDA APHIS permits and an exemption from the Director of the Oregon Department of Agriculture for work with *Phytophthora* under containment conditions. Twenty-one *P. ramorum* isolates were selected based on variation across seven microsatellite loci and represented the three clonal lineages (Fig. 1). The microsatellite loci used were PrMS39b, PrMS43a, PrMS43b, PrMS45 (Prospero *et al*. 2007) and 18, 64, 82 (Ivors *et al*. 2006), and were typed using primers and protocols as described elsewhere (Garnica *et al*. 2006; Ivors *et al*. 2006; Prospero *et al*. 2007; Grünwald *et al*. 2008c). A neighbour-joining tree was constructed from microsatellite allele frequencies to illustrate the variation among isolates (Fig. 1). Using the program MSA (Dieringer & Schlötterer 2003), genetic distance among isolates was calculated using Nei's chord distance, *D*a (Nei *et al*. 1983). The PHYLIP 3.67 package (Felsenstein 2004) was used to generate a neighbour-joining tree and produce a consensus tree from 1000 bootstrap replicates generated in MSA. Locus 82 was excluded from the tree because it amplified more than two alleles.

Sequencing

Five single-copy loci were sequenced for the 21 *P. ramorum* isolates and the three other *Phytophthora* clade 8c species (Table 1). The genes sequenced were beta tubulin (*btub*), the indole-3-glycerol-phosphate synthase N-5'-phosphoribosyl anthranilate isomerase gene (*trp1*, involved in biosynthesis
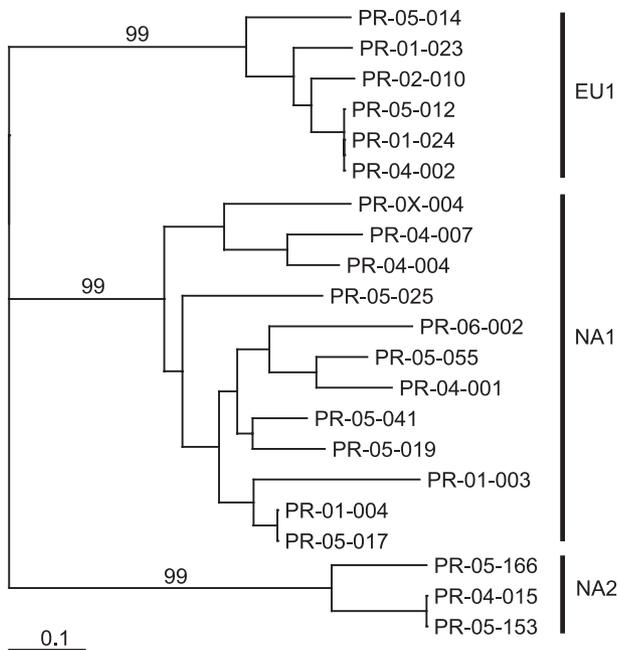
Fig. 1 Neighbour-joining tree of *Phytophthora ramorum* isolates used for sequencing based on variation at six microsatellite loci. Bootstrap values are shown for branches with greater than 80% support. A seventh locus (82, Ivors *et al.* 2006) that produces more than two alleles was not included in the tree, yet is variable among isolates and differentiated among lineages.

of tryptophan), two putative effectors [*Avr* gene homologues; *PrAvh120* and *PrAvh121* (Tyler *et al.* 2006; Jiang *et al.* 2008)], and a gene coding a hypothetical protein with a glycosyl transferase group 1 domain, named gwEuk.30.30.1 in the JGI annotation of the *P. ramorum* strain Pr102 genome sequence (Tyler *et al.* 2006). This gene contains a microsatellite (Pr9C3, Prospero *et al.* 2004) in an intron that shows variation among but not within *P. ramorum* lineages. The microsatellite itself was removed for analysis due to poor alignment. These genes were selected primarily because they showed intraspecific variation in preliminary sequencing. The putative effectors are homologous to known *Avr* genes in other *Phytophthora* species (Jiang *et al.* 2008) and were chosen because microarray data suggest that they are expressed in *P. ramorum* sporangia (C. Press and N. Grünwald, unpublished) and are expected to be rapidly evolving. The protein IDs for these genes in the Pr102 genome sequence are: *btub* 72114; *trp1* 71833; *PrAvh120* 76672; *PrAvh121* 81834; and gwEuk.30.30.1 46978. Each gene is located on a different scaffold in the genome assembly and was confirmed to be in single copy based on the genome sequence and corresponding trace archives (Table 2). The full gene was sequenced for the two effectors and a fragment of the gene was amplified for the three other loci. *Taq* DNA polymerase from GenScript Corporation

was used for polymerase chain reaction (PCR) amplification with 3 mM $MgCl_2$. PCR was performed with an initial denaturation for 3 min, followed by 35 cycles of denaturation for 1 min, annealing for 1 min, and extension at 72 C for 1 min, with one final 10-min extension. Primers and annealing temperatures are given in Table 2. PCR products were sequenced using ABI PRISM BigDye Terminator 3.1 chemistry (Applied Biosystems) and analysed on Applied Biosystems 3730 capillary sequencers by the Oregon State University Center for Genome Research and Biocomputing core laboratory. Sequences were edited using ContigExpress, a component of Vector NTI Advance 10.3.0 (Invitrogen Corporation), and BioEdit (Hall 1999). Sequences showing more than one heterozygous site were cloned in two isolates per lineage using TOPO TA Cloning (Invitrogen Corporation) or the pGEM-T Easy Vector System (Promega Corporation). ClustalW (Thompson *et al.* 1994) was used for multiple alignment of sequences. GenBank Accession numbers are provided in Table 2.

Polymorphism and selection

Within each lineage, all sequenced isolates were identical; therefore the data for each locus were collapsed to one individual per lineage. SITES (Hey & Wakeley 1997) and DnaSP (Rozas *et al.* 2003) were used to summarize the sequence data for each locus. Polymorphism was quantified using the population mutation parameter θ_W (Watterson 1975). We determined the minimum number of recombination events by the Rmin of Hudson & Kaplan (1985) and estimated the rate of intralocus recombination in *P. ramorum* with Hey & Wakeley's γ (1997). Tajima's *D* (Tajima 1989), Fu and Li's *D** and *F** (Fu & Li 1993), and Fu's F_S (Fu 1997) were used to test for selection or demographic effects on each locus. Significance was assessed using empirical distributions generated in DnaSP using the default settings given the data.

Phylogenetic inference

Standard phylogenetic analyses on each locus were conducted using PAUP* (Swofford 2002) and included all four 8c species. We used both maximum parsimony (MP) and maximum likelihood (ML) to infer gene trees using the branch- and-bound algorithm. For ML, we selected the best evolutionary model given the data using the AIC in ModelTest 3.7 (Posada & Crandall 1998). Support for branches was assessed using 500 bootstrap samples.

Coalescent analysis

The coalescent uses stochastic processes, based on a population genetic model, to approximate the ancestry of a sample of DNA sequences going backwards in time by

Table 1 Isolates sequenced for five nuclear genes

Isolate	Source name	Species	Origin	Year isolated*	Clonal lineage	Source†
PR-04-002	03-74-D12-A	*P. ramorum*	Oregon	2003	EU1	N. Osterbauer
PR-05-012	BBA9/95; CSL2266	*P. ramorum*	Germany	1995	EU1	H. DeGruyter
PR-01-023	BBA15/01-11a	*P. ramorum*	Germany	2001	EU1	S. Werres
PR-02-010	BBA2N0389	*P. ramorum*	France	2002	EU1	S. Werres
PR-01-024	BBA15/01-14	*P. ramorum*	Germany	2001	EU1	S. Werres
PR-05-014	CSL2267	*P. ramorum*	UK	n.a.	EU1	C. Lane
PR-06-002	04-189-B5	*P. ramorum*	Oregon	2004	NA1	E. Hansen
PR-04-007	04-207-Q	*P. ramorum*	Oregon	2004	NA1	E. Hansen
PR-04-004	Pr102	*P. ramorum*	California	n.a.	NA1	D. Rizzo
PR-0X-004	Pr106	*P. ramorum*	California	n.a.	NA1	D. Rizzo
PR-05-019	wsda4164	*P. ramorum*	Washington	2004	NA1	E. Hansen
PR-05-025	wsda4165	*P. ramorum*	Washington	2004	NA1	E. Hansen
PR-01-004	1020.1	*P. ramorum*	Oregon	2001	NA1	E. Hansen
PR-01-003	1033.1	*P. ramorum*	Oregon	2001	NA1	E. Hansen
PR-05-017	WA15.3-080403	*P. ramorum*	Oregon	2003	NA1	E. Hansen
PR-05-041	4361	*P. ramorum*	Oregon	2003	NA1	E. Hansen
PR-05-055	2092	*P. ramorum*	Oregon	2002	NA1	E. Hansen
PR-04-001	2027.1	*P. ramorum*	Oregon	2001	NA1	E. Hansen
PR-04-015	wsda3765	*P. ramorum*	Washington	2004	NA2	A. Wagner
PR-05-153	RHCC1	*P. ramorum*	California	2005	NA2	M. Garbelotto
PR-05-166	MR31	*P. ramorum*	Washington	2004	NA2	M. Garbelotto
PL-0X-001	366	*P. lateralis*	USA	n.a.	—	E. Hansen
PH-05-003	P3822	*P. hibernalis*	Australia	n.a.	—	M. Coffey
P-04-001	LT192-010505	*P. foliorum*	Tennessee	n.a.	—	K. Lamour

*n.a., not available.

†N. Osterbaur, Oregon Department of Agriculture, USA; H. DeGruyter, Dutch Plant Protection Service, the Netherlands; S. Werres, Federal Biological Research Centre, Germany; C. Lane, Central Science Laboratory, UK; E. Hansen, Oregon State University, USA; D. Rizzo, University of California, Davis, USA; A. Wagner, Washington State Department of Agriculture, USA; M. Garbelotto, University of California, Berkeley, USA; M. Coffey, University of California, Riverside, USA; K. Lamour, University of Tennessee, USA.

evaluating all possible mutational pathways back to a common ancestor. We used the coalescent to examine the ancestry of the three *P. ramorum* clonal lineages. The divergence of *P. ramorum* from the other three 8c species was too great to include them in the coalescent analysis.

At least two of the loci showed evidence of recombination in initial tests (Table 3); thus, we included recombination in our coalescent simulations. We used recom version 5.8 for coalescent analysis with recombination, assuming the infinite-sites model, neutral evolution, panmixia, and constant population size (Griffiths & Marjoram 1996). recom simulates the coalescent given values of the population mutation rate θ ($2N_e\mu$, as recom uses a haploid model) and the population recombination rate ρ ($2N_e r$) and estimates a joint-likelihood surface for these parameters given the data. It assumes that the ancestral state of each site is known. By comparing each segregating site in *P. ramorum* to its state in *P. lateralis*, *P. hibernalis*, and *P. foliorum*, we assigned the ancestral state. Two sites in *trp1* (428 and 431) had unclear ancestral states; one was interpreted as an allele-specific mutation in EU1 and for the other NA1 was

assigned to the derived state. One site with three states was removed in *PrAvh121*. Sequence data were converted to the recom input file format using SNAP Workbench (Price & Carbone 2005). For each locus, θ_W and ρ = 1.0 were used to generate initial maximum-likelihood estimates (MLEs) of ρ. Then, θ_W and the ρ estimate were used for four runs of 10 million to 50 million simulations to generate joint estimates of both parameters. Finally, these estimates were used to obtain estimates of recombination events, mutation times, and time to the most recent common ancestor (TMRCA, scaled by population size) also using four independent runs. Two to five times more simulations per run were used for *btub* and gwEuk.30.30.1 to improve convergence of estimates among runs. *PrAvh121* crashed the program unless the +b switch was used, which aborts low probability paths and returns zero. We compared parameter estimates obtained from recom to those from sequenceLD (Fearnhead & Donnelly 2002) and LDhat (McVean *et al.* 2002). Methods and results for the sequenceLD and LDhat analyses are described in the Supplementary Data.

Table 2 Primers, amplification conditions, and GenBank Accession numbers for resulting sequences

Gene	Amplification primers	Primer position in Pr102 genome	T_a (°C)	Amplicon size (bp)	GenBank Accession(s)
beta-tubulin*	TUBUF3 CGGTAACAACTGGGCTAAGG	88:92179-92198	55	988	EU850939–EU850974
	TUBUR1 CCTGGTACTGCTGGTACTCAG	88:93147-93167			
trp1	Trp1F1 CCGAACACGCAAAGTGCAAATAACAAGATGGG	52:331060-331091	60	972	EU850975–EU850996;
(*P.r.*, *P.l.*)	Trp1R1 CGTCACGGCTGAGCGAAAGTCCCCG	52:332008-332032			EU850999–EU851002
trp1	Trp1F2 GAACATCCTGGAGGAGATCG	52:331092-331111	58	1041	EU850997–EU850998
(*P.f.*, *P.h.*)	Trp1R2 CATCAAGTACTCGCCRACCA	52:332114-332133			
gwEuk.	Pr9C3c_F TTCAAGTGGGAGGAGAGGAA	30:312889-312908	56	748	EU851003–EU851038
30.30.1	Pr9C3c_R GCTCCAGGATGATCTCTCCA	30:313618-313637			
PrAvh120	PrRXLR120_-242F CTGTGCGAATCTTGCAACC	19:437966-437984	58	880	EU850875–EU850895;
(*P.r.*)	PrRXLR120_639R GCGAATCATTCCTACGCTTG	19:437104-437123			EU850899–EU850902
PrAvh120	PlRXLR120_-301F CGGCATGAACACCATATCAT	NA	58	929	EU850896
(*P.l.*)	PlRXLR120_+629R GACGCTAGGTGGGTCAAAAC				
PrAvh120	PfRXLR120_-301F GAGACAGCATCACGTGCCTA	NA	58	905	EU850898
(*P.f.*)	PfRXLR120_+605R GGTTCTTTGGCGTCAGTAGC				
PrAvh120	PhRXLR120_-125F CTATCTCCTCCCGCACTCTG	NA	58	804	EU850897
(*P.h.*)	PhRXLR120_+680R GCCACTGAACACAGTTTTGG				
PrAvh121	PrRXLR121_-244F CATGCTGACCCATTCAGTACC	64:182656-182676	58	866	EU850903–EU850923;
(*P.r.*)	PrRXLR121_623R TCAAAGGGCCTTCTGCATAC	64:183503-183522			EU850927–EU850938
PrAvh121	PhPrRXLR121_Pr-72F GCAARCTCATTCYTCGATCA	NA	58	808	EU850924
(*P.l.*)	PlRXLR121_+737R GCTACGCCACATCCAAAAAT				
PrAvh121	PfRXLR121_-291F GGCATTTAGGGAGACGTGAA	NA	58	817	EU850926
(*P.f.*)	PfRXLR_121_+527R AAGAAACCTTGTCGCTCGAA				
PrAvh121	PhRXLR121_-280F GCAACCTTTGCTCCTTTGAC	NA	55	937	EU850925
(*P.h.*)	PhRXLR121_+658R ACGAGGTTCTGCGTGAGTTT				

*primers modified from Kroon *et al.* (2004).

T_a, annealing temperature.

Table 3 Summary statistics for sequenced loci

	btub	trp1	gwEuk.30.30.1	PrAvh120	PrAvh121
P. ramorum only					
Length	908	807	630	387	399 (387)
Segregating sites	4	8	10	7	35
θw (per gene)	1.75	3.50	4.38	3.07	15.33
Rmin	0	0	1	0	1
γ	0	0	17.66	0	11.45
Tajima's D	0.768^{ns}	-0.399^{ns}	0.577^{ns}	-0.125^{ns}	-0.053^{ns}
Fu's F_S	-2.112^{**}	1.992^{ns}	1.101^{ns}	-1.283^{ns}	1.514^{ns}
Fu and Li's D^*	0.769^{ns}	-0.302^{ns}	0.699^{ns}	-0.138^{ns}	-0.151^{ns}
Fu and Li's F^*	0.814^{ns}	-0.347^{ns}	0.727^{ns}	-0.145^{ns}	-0.142^{ns}
With other 8c spp.					
Segregating sites	56	123	125	114	129
Mutations	63	133	140	125	145
Parsimony informative characters†	25	40	33	38	45
Consistency index for inform. char.†	0.80	0.88	0.85	0.85	0.68

[ns]$P > 0.1$.

**0.01 < P < 0.05.

†gaps removed.

```
                        btub   trp1       gwEuk30.30.1 PrAvh120 PrAvh121
                        678    444667     11112355     113      111111112222222222233333333333333333333
                        6339   45235352   6912492625   5679163  2366892233566755777899990112222222222333333445556
              Position  4108   13818125   6232071037   6400824  628461908155387680790278910723456789012384558 92
         Coding status^ SSSS   SSSSRIRS   RSSSSSSIII   RRRSRSR  SSRRSSS&RRSRRSSRSSRSRSRRSRRRRRDDDDDDDDDDDDDDDRRSRRRR
          Substitution* VNVN   NNVNVVNN   VNVVNNNNNN   NVNNVNN  NNNNNNN&NVNVNVNVNVNNVNNNNNNNVVV-----------VVVVNVN
Inferred ancestral state GTGT  GCGAATAG   AGGGGTAGTG   GGGCAGT
              Consensus GTCC   GCGAATAA   TATGGTAGTG   AGGCAGT  TTCAGCATTATGAATGAATGGGAATCGAGTCGGCAAGAAGGTATACG
                  EU1a  .CGT   ........   ........CA   ....TA.  CCTG.TGACTCTGTC................................
                  EU1b  ....   .TC.CAG.   AGGTAC.A..   ....T..  CCTG.TGGCT.TGTC..GGAAAGGCATC-----------TGTAGA.
                  NA1a  ..GT   A..G...G   ........CA   G.A....  ......GA.......CG.............................
                  NA1b  T...   A..G...G   ......G...   G.A....  .............................................G.A
                  NA2a  T...   ........   ......G...   ...T..C  .............................................G.A
                  NA2b  ..G.   ........   AGG.A.G...   GT.....  ....A..........C.............................
```

  ^ Coding status: R amino acid replacement (nonsynonymous); S synonymous; I intron; D indel.
  * Substitution: N transition; V transversion; & site has 3 states including both types of
             substitutions; - indel.

Fig. 2 Sites segregating in *Phytophthora ramorum*, shown by position in gene for *PrAvh120* and *PrAvh121* and by position in sequenced fragment for *btub*, *trp1*, and gwEuk.30.30.1. Identity to consensus nucleotide is indicated by a dot.

Visualizing the evolutionary history using genetree and beagle

For loci with no conflicting sites, we also ran coalescent analyses in genetree version 9.0 (Bahlo & Griffiths 2000), which assumes no recombination, to compare the results to those from recom and to obtain coalescent gene genealogies. The results are expected to be slightly different due to differences in the computational methods used by the programs (Griffiths & Marjoram 1996). Ancestral states were coded the same as for recom. An initial run using θ_W and 10 million simulations was used to generate an MLE of θ (using option –2 to calculate the exact likelihood when only 2 ancestors remain). This MLE was used to estimate mutation times and TMRCA, also using 10 million simulations (but not option –2).

The coalescent history of a sample that has experienced recombination cannot be visualized in a bifurcating tree, rather it is represented by an ancestral recombination graph (ARG). A true ARG would show all inferred recombination events involving ancestral sequences. A minimal ARG shows only the recombination events that are required given the data. The program beagle uses a branch-and-bound algorithm to compute an exact minimum number of recombination events required for the data and produces an associated minimal ARG (Lyngsø *et al.* 2005). The resulting ARG is a schematic of inferred relationships, showing the inferred order of recombination and coalescent events back to the root, but is not scaled to time like a true coalescent tree. We analysed the two loci with conflicting sites, gwEUK30.30.1 and *PrAvh121*, with beagle. Like recom, beagle assumes the infinite-sites model. The ARG can be rooted using a specified outgroup or beagle can infer the root. We rooted the ARG with *P. lateralis* (in effect assigning *P. lateralis* to be the ancestral sequence). Two sites with three states were excluded from *PrAvh121*. Sequence data

were reduced to binary variable sites using Map (Aylor *et al.* 2006) as implemented in SNAP Workbench (Price & Carbone 2005) for input into beagle.

Results

Polymorphism in P. ramorum *nuclear genes*

All of the observed variation was between the two alleles within each lineage or among the three lineages. There were no differences in the sequenced genes among individuals within the lineages such that each lineage produced one or two alleles at each locus with no variation among isolates. Across lineages, we found a total of four to six haplotypes at each sequenced locus (Fig. 2). Segregating sites per locus ranged from 4 to 35 and θ_W from 1.75 to 15 (Table 3). *PrAvh121* was much more variable than any of the other nuclear loci examined. Variable sites included noncoding, synonymous, and nonsynonymous (replacement) mutations (Fig. 2). Neutrality was not rejected at any locus as measured by Tajima's *D* or Fu and Li's statistics (Table 3). Fu's F_S was significant for *btub* only, but given the nonsignificant results for the other tests and loci, this result may be due to the sensitivity of F_S to recombination (Fu 1997). Both gwEuk.30.30.1 and *PrAvh121* contained a minimum of one recombination event (Table 3). The same two loci also had non-zero values of γ (Table 3).

The putative effector *PrAvh121* appeared to have a complex structure and history. The alleles in the EU1 lineage were highly diverged from the other two lineages for approximately the first third of the gene, while in the last third of the gene, one allele was more similar to the *Phytophthora lateralis* gene sequence than to the other *P. ramorum* sequences, including the presence/absence of a 12-bp indel (Figs 3 and 4). Yet, the 3'-end flanking regions of the EU1 alleles were identical over 150 bp and quickly became

```
                      10        20        30        40        50        60        70
              ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
NA1a          MRMNQLLLLVIATFIARCSTLTDAERSIQVDTLATDSITESAIEGHRYLKGSKTTTALDTVDEERIGATT
NA1b          ..............................................D......................
NA2a          ..............................................D......................
NA2b          ..............................................D......................
EU1a          ...........I........A...................TV......I..A..................
EU1b          ...........I........A...................TV......I..A..................
P. lateralis  ....F....V....C....A.......N..T...........T.........A..D.............
P. foliorum   ...R..F....S....L...IAS.DH.T..NAM....T..F.T..C....G..A.D..L.N...F---.
P. hibernalis ....*.....V...VVH.NPIAA......ANS.V.EA.A...TT.........A.D..AA.........

                      80        90        100       110       120       130
              ....|....|....|....|....|....|....|....|....|....|....|....|....|...
NA1a          PNFGMLKGLFKLPKFQSLSKLPLVKQLVAIQKKVGKKVGKKIGDIYLKWTRKRYESNPQNFM*
NA1b          ...................E..............................AK..........*
NA2a          ...................E..............................AK..........*
NA2b          .............................................................*
EU1a          ...................E.........................................*
EU1b          ...................E......VI...I.AK..F.T.----..V.S...RE.........*
P. lateralis  ....I..A...........E......VI...A.AK..F...----..V.Y...RE.........*
P. foliorum   .S....T.AL.......G.........VA.SIA.F..T----A..Q...LMK...Q.......*
P. hibernalis .K........R....KN.........VA..S..F.N.----V..V...AK...........*
```

Fig. 3 Peptide sequence of PrAvh121 for each haplotype of the three *Phytophthora ramorum* lineages and closest known relatives. Identity to the first sequence is indicated by a dot, dashes indicate gaps, and asterisks indicate stop codons.
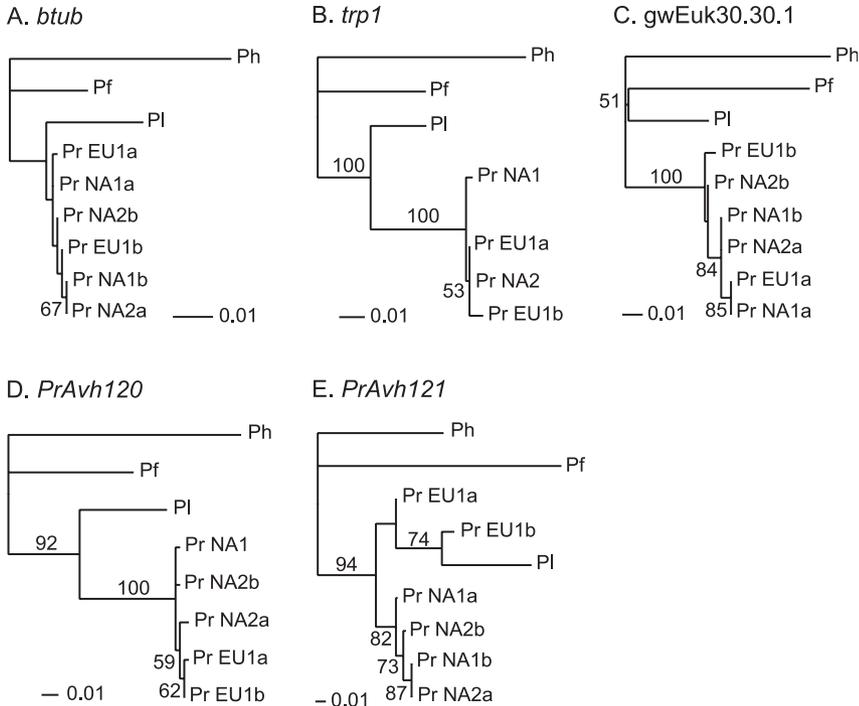
Fig. 4 Maximum-likelihood phylogenetic trees for each sequenced locus rooted with *Phytophthora hibernalis*. Bootstrap values are shown for branches that had greater than 50% support. Species names are abbreviated: Ph, *P. hibernalis*; Pf, *P. foliorum*; Pl, *P. lateralis*; Pr, *P. ramorum*. Shown are each haplotype observed for *P. ramorum* as well as a single representative sequence for the other three species.

highly diverged from the corresponding *P. lateralis* sequence (by 50 bp from the stop codon; data not shown). When only the third codon positions were examined, there were still 13 segregating sites, more variation than was observed across all sites at any of the other four loci.

Genealogies

Maximum-likelihood genealogies for each locus showed that the two alleles of each clonal lineage did not always cluster together (Fig. 4). Furthermore, the relationships

Table 4 Results of coalescent analyses on individual loci

	btub	*trp1*	gwEuk.30.30.1	*PrAvh120*	*PrAvh121*
recom					
Generating θ	1.75	3.7	4.3	3.9	13.0
Generating ρ	3.5	0	1.5	0.5	1.0
Mean TMRCA (sd)	2.64 (0.88)	1.49 (0.47)	2.49 (0.71)	1.22 (0.45)	2.28 (0.37)
Mean recombination events (SD)	6.96 (2.20)	0.00 (0.00)	4.03 (1.35)	0.96 (1.00)	3.00 (0.002)
Estimated likelihood	1.21×10^{-4}	1.32×10^{-3}	2.91×10^{-5}	3.96×10^{-5}	3.15×10^{7}
(standard error)	(3.84×10^{-5})	(1.18×10^{-6})	(8.48×10^{-6})	(1.46×10^{-6})	(3.15×10^{7})
θ estimate	1.9	3.70	4.44	3.90	11.0
ρ estimate	2.88	0.00	1.90	0.875	0.90
genetree					
TMRCA (SD)	1.27 (0.52)	1.48 (0.48)	NA	1.03 (0.35)	NA
Likelihood of tree	4.72×10^{-5}	5.33×10^{-6}	NA	1.76×10^{-5}	NA
(standard error)	(2.07×10^{-8})	(3.46×10^{-9})		(1.09×10^{-8})	
θ estimate	2.05	3.675	NA	4.00	NA

NA, not applicable due to conflicting sites in gene.
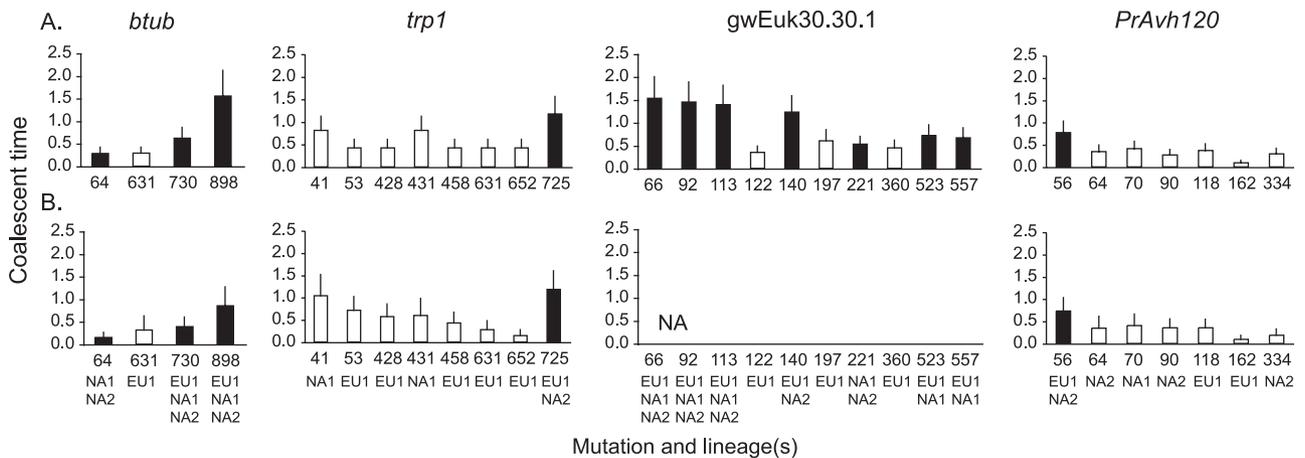
Fig. 5 Mutation ages estimated by (A) recom and (B) genetree in one representative run for each locus. Bars indicate mean estimates of mutation age, error bars show standard deviation, and scale is in coalescent time, which is a function of population size and generation time (2N × time × generation time). Filled bars are mutations shared between lineages and empty bars are lineage-specific mutations, with the lineage(s) in which the mutation occurred shown below. Ancestral state for each site was inferred by comparison to *Phytophthora lateralis*, *P. hibernalis*, and *P. foliorum*.

between the lineages changed among loci, although with weak bootstrap support. The results of maximum parsimony were similar except for *btub*, for which parsimony gave much higher bootstrap support values than the likelihood model. *P. ramorum* is sufficiently diverged from *P. lateralis*, *P. hibernalis*, and *P. foliorum* that multiple mutations per site and homoplasy were observed (Table 3).

Coalescent analyses

We obtained the TMRCA, in scaled coalescent units, of the *P. ramorum* clonal lineages at each locus as well as the

age of each mutation using the coalescent with recombination. In most cases, convergence was observed among independent runs for each locus. A summary of representative runs for each locus is given in Table 4 and illustrated in Fig. 5. Estimates of TMRCA and number of recombination events for *btub* were similar among runs, but estimates of θ and ρ ranged from 1.8 to 2.4 and 1.9 to 3.5, respectively. In addition, the ages of the four *btub* mutations varied among runs; therefore, one of three scenarios is given in Fig. 5. *trp1*, with an estimated population recombination rate of zero, gave nearly identical results among replicate runs. Yet, the estimates of mutation ages by recom for *trp1* (with
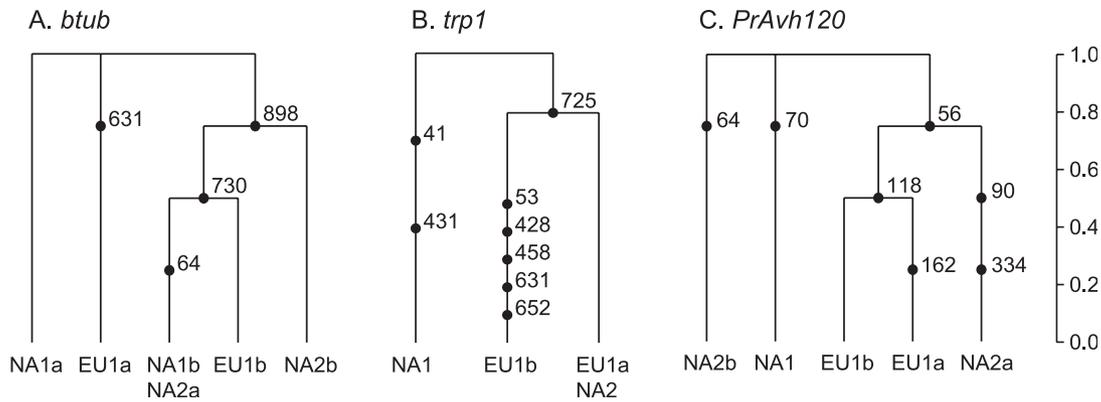
Fig. 6 Coalescent-based gene genealogies for the three loci that had no conflicting sites. Trees were generated in genetree, which assumes no recombination, and scaled to TMRCA of 1.0 for each locus. Mutations are labelled by their location in the sequence, as in Fig. 2.

no recombination) were problematic; different mutations occurring on a given lineage were assigned nearly identical mutation times (Fig. 5). The number of recombination events was estimated to have been greater than one for all loci except *trp1* and *PrAvh120* (Table 4).

We attempted to reconstruct the evolutionary history of *PrAvh121* using the coalescent with recombination. However, the structure of the variation appeared to cause problems for recom as the program consistently crashed. When we aborted low probability paths, the runs that were successful in producing output showed few unaborted paths, ranging from 35 to 88 out of 50 million simulations. *PrAvh121* apparently had a difficult history to reconstruct using a coalescent model, suggesting that evolution by point mutation and neutral evolution were not appropriate assumptions for this locus. While the locations of probable recombination events varied among runs, each run showed at least one event around the middle of the gene, a location one would expect based on the structure of the variation.

We obtained coalescent gene genealogies for the three loci with no conflicting sites (Fig. 6). *PrAvh120* and *trp1* showed similar TMRCA in genetree and recom (Table 4), since these loci appear to have experienced little recombination. In contrast, the TMRCA of *btub* differed by a factor of two between the analyses with and without recombination (Table 4).

The minimal ARG generated in beagle for gwEuk.30.30.1 showed that the NA2b haplotype may be the result of a recombination event between the EU1b haplotype and a shared NA1/NA2 haplotype (Fig. 7A). For *PrAvh121* rooted with *P. lateralis*, beagle estimated a minimum of five recombination events (Fig. 7B) in contrast to the single event (after position 177) inferred when only the *P. ramorum* sequences were used. The ARG shows the N-terminal region of the gene associated with the EU1 alleles evolving

with the C-terminal region shared by most of the *P. ramorum* alleles (dual gray and black paths in Fig. 7B). The EU1 N-terminal region is brought together with the *P. lateralis*-like C-terminal region by a recombination event after site 131 while the NA1 and NA2 alleles are each joined with their N-terminal regions by a different recombination event (at locations 249, 256, and 258). The fifth recombination node in the ARG (at site 135) is solely to address the mutation at this site, which is shared between one of the EU1 alleles and the NA1 and NA2 alleles.

We estimated how long the lineages may have been isolated from each other by examining their relative time to coalescence, as indicated by the ages of mutations shared between lineages. Shared mutations were in most cases older than lineage-specific mutations (Fig. 5). *PrAvh120* and *trp1* each had a single shared mutation, in both cases shared by EU1 and NA2 (Figs 5 and 6). Different realizations of recom runs with *btub* showed either *btub*[64] (NA1/NA2) or *btub*[730] (all three lineages) to be the first shared mutation. Assuming no recombination, genetree estimated the first shared mutation to be *btub*[64] (Fig. 6). The first shared mutation in gwEuk.30.30.1 was consistently gwEuk.30.30.1[221] (NA1/NA2, Fig. 5). In absolute terms, the youngest shared mutation was *btub*[64] followed by gwEuk.30.30.1[221], both mutations shared by NA1 and NA2, suggesting that the first coalescence between lineages back in time may be of NA1 and NA2. When scaled to the TMRCA of the locus, the time to the first *btub* mutation varied between 11% and 18% (depending on the run) of the TMRCA. The mutation at gwEuk.30.30.1[221] was estimated to have occurred at 21% of the TMRCA of gwEuk.30.30.1. These time estimates assume panmixia, which would underestimate coalescent times if the lineages were in fact geographically structured and change the relative timing of mutation and coalescence events if migration rates varied.
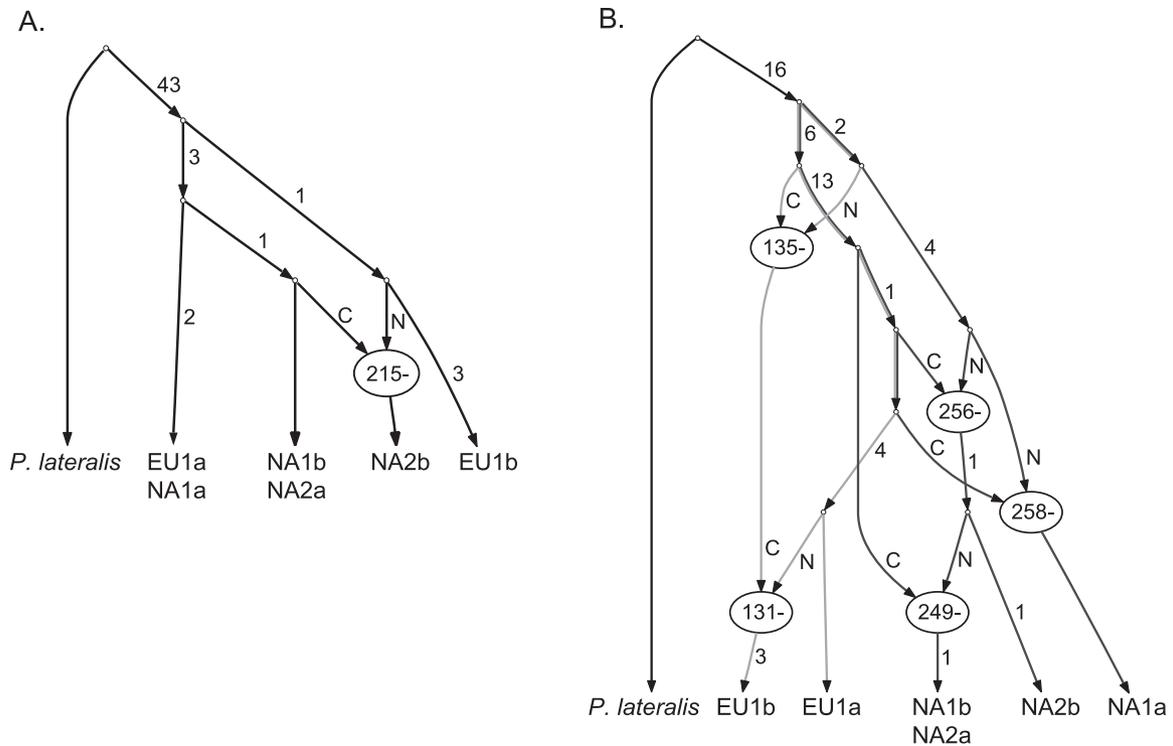
Fig. 7 Minimal ancestral recombination graph (ARG) for (A) gwEuk.30.30.1 and (B) *PrAvh121*. Recombination events are indicated by ovals containing the physical location of the recombination breakpoint in the sequence (breakpoint follows given site). The paths leading to a recombination node are labeled with 'N' or 'C' to indicate the origin of the N-terminal (5' to the breakpoint) and C-terminal (3') portions of the recombinant gene. Numbers to the right of paths are the number of mutations in that segment. The small open circles indicate coalescence events. The direction of the paths is from the past to the present. For *PrAvh121* the paths associated with EU1 alleles are light gray to distinguish them from the paths associated with the NA1 and NA2 alleles. The ARG is rooted with *Phytophthora lateralis*, therefore the total number of mutations is across the three *P. ramorum* lineages and two species.

Discussion

Our analysis of DNA sequence variation at five nuclear loci suggests that the *Phytophthora ramorum* clonal lineages have been isolated from each other for an evolutionarily significant amount of time, yet are descendents of an ancestral sexually recombining population. We did not find DNA sequence variation among isolates within each lineage, suggesting that the observed variation dates prior to introduction. Evidence for independent evolution of the three lineages can be seen by simply examining the distribution of mutations. We found that out of a total of 29 segregating sites across four loci, 17 mutations were specific to a single lineage when coded based on sequences from *P. ramorum*'s three closest known relatives. We used the coalescent to put a timescale on these mutations and estimated that the three lineages have been separated for approximately the last 11 to 21% of their history. Given the lack of a fossil record for oomycetes, we used a conservative approximation to date divergence. Based on synonymous substitutions across four genes and an assumed synonymous substitution rate between 2×10^{-9} and 7×10^{-9} mutations

per site per year based on the literature for plants, animals, and fungi (Li *et al.* 1987; Wolfe *et al.* 1989; Gaut *et al.* 1996; Kasuga *et al.* 2002), we estimate the history of the three lineages to be between 1.5 million and 5.4 million years old. This would make the most recent mutation shared between lineages a minimum of 165 000 years old. Given the divergence of the lineages, it is unlikely that the three lineages are introductions of three different genotypes from a single interbreeding population; rather the lineages are more likely to have originated from three geographically isolated populations. Alternatively, the lineages may have recently migrated to a single population in which they remained reproductively isolated (i.e. asexual, as is currently the case in US nurseries) and from which the introductions to North America and Europe originated. That the NA1 and EU1 genotypes were limited to North America and Europe, respectively, in the first years following introduction (the EU1 lineage has since migrated to the US) perhaps lends support to the hypothesis of different source populations. Studies of population genetic variation and dispersal in California and southwest Oregon indicate that *P. ramorum* may have limited intrinsic

ability to disperse long distances (Davidson *et al*. 2005; Prospero *et al*. 2007; Hansen *et al*. 2008; Mascheretti *et al*. 2008). Thus, it seems plausible that *P. ramorum* populations could be geographically structured across its native range.

The EU1 lineage is intriguing because it contained most of the allele-specific mutations and these were inferred from the coalescent analyses to be relatively recent. The EU1 lineage could have more of these mutations because it has been isolated from NA1 and NA2 longer than they have been isolated from each other, which is consistent with the coalescent analyses that dated shared mutations between NA1 and NA2 as the youngest of the shared mutations. This contrasts with analysis of variable regions in the mitochondrial DNA (mtDNA), which concluded that NA2 has the basal mtDNA haplotype (Martin 2008). The mtDNA haplotypes likely diverged separately from the nuclear genomes, which is possible since the mtDNA is maternally inherited and generally does not recombine. Alternative explanations for the EU1-specific mutations are an elevated mutation rate, a considerably larger effective population size, or more recent recombination in EU1 that maintained more variation in this lineage compared to the others.

We observed strong evidence of recombination in at least two genes, suggesting that there was sexual reproduction in *P. ramorum*'s past. This species should be capable of outcrossing, yet matings between A1 and A2 isolates proceed unusually slowly and it is unclear if oospores are viable (Brasier & Kirk 2004). Their lack of success in mating may be a consequence of their apparent long period of isolation. *P. cinnamomi* similarly appears to reproduce asexually across its introduced range even when both mating types coexist (Linde *et al*. 1999; Dobrowolski *et al*. 2003). *P. infestans* reproduces sexually in Mexico and Europe (Goodwin *et al*. 1992; Drenth *et al*. 1994; Grünwald & Flier 2005), but there are regions in the USA and Asia where both mating types have been present but with no detectable mating (Mosa *et al*. 1993; Miller *et al*. 1997). Therefore, *P. ramorum* is not exceptional in its current behaviour.

Since asexual reproduction can cause a build-up of slightly deleterious mutations (Felsenstein 1974), it is interesting to note that we observed several lineage-specific amino acid replacements that are estimated to be evolutionarily recent mutations. Most of the replacement changes were in *PrAvh120* and clustered in the 5′ end of the gene (the first ~200 bp). This is notable because *Phytophthora* RXLR-class effectors are thought to be modular in structure: N-terminal domains are involved in targeting (secretion of the protein out of the pathogen and transportation into the host cell) while the C-terminal domain is responsible for the effector activity within the plant host (Whisson *et al*. 2007; Dou *et al*. 2008a, b). Given the function of the N-terminal sequence, one would hypothesize that this region would be under purifying selection and would exhibit little variation in amino acid composition, if the gene is functional, whereas the C-terminal region may be highly polymorphic (Shan *et al*. 2004; Rehmany *et al*. 2005; Allen *et al*. 2008). Thus, the amino acid polymorphism localized to the N-terminal is unexpected. We cannot determine whether recombination has occurred in the time since isolation of the lineages from our data, but if it is the case that these lineages have accumulated deleterious mutations, the shuffling of variation provided by sexual reproduction could lead to a significant improvement in fitness. Yet, *Phytophthora* species also exhibit mitotic recombination (Goodwin 1997; Dobrowolski *et al*. 2003) and gene conversion (Chamnanpunt *et al*. 2001), which may be effective in purging deleterious mutations.

The second effector that we sequenced, *PrAvh121*, provides a stronger case for selection shaping the patterns of variation observed in the three lineages. Both EU1 alleles share six amino acids that differ from the other lineages in the targeting region, but in the last 40 amino acids of the protein, one EU1 allele is identical to alleles found in the other lineages while the other shares eight amino acid changes and an indel with the orthologous *P. lateralis* protein (Fig. 3). Given the genetic distance between *P. lateralis* and *P. ramorum*, a recombination event between EU1 and *P. lateralis* seems unlikely. The most parsimonious explanation for the similarity in the C-terminus is that selection, specifically balancing selection, has maintained this shared domain since speciation. The neutrality tests we conducted may not have detected balancing selection since the pattern was limited to the EU1 lineage. The targeting region shared by the EU1 alleles contains more sites in the apparent ancestral state than NA1 and NA2, suggesting that this is the older *P. ramorum* allele. Clearly, recombination linked the one targeting region with the other functional region, and the question that remains is whether both alleles were maintained in one or more of the source populations or if this recombination occurred prior to the isolation of the lineages.

We used the coalescent to establish that the three clonal lineages of *P. ramorum* diverged long ago, prior to any possible human influence. However, there can be little doubt that the three introductions of *P. ramorum* to Europe and the USA were the result of human activity. Interestingly, zoonoses (diseases from non-human animal origin) make up a majority of the recent emerging infectious diseases of human populations (Jones *et al*. 2008) and these diseases, like *P. ramorum*, have comparatively broad host ranges (Woolhouse & Gowtage-Sequeria 2005). For some of these diseases, including West Nile virus and Lyme disease, humans are a dead-end host but it is their effect on humans that is the cause for concern. Similarly, *P. ramorum* would be a minor pathogen if it did not kill high-value oaks, which are dead-end hosts (Davidson *et al*. 2005).

Emerging diseases have been compared to weediness in plants, in that they are opportunistic and can take advantage of new ecological niches presented by human activity (Morens *et al.* 2004; Woolhouse & Gowtage-Sequeria 2005). *Phytophthora* species have repeatedly shown themselves to exhibit these characteristics and *P. ramorum* has already proven itself three times.

Acknowledgements

References

Allen RL, Meitz JC, Baumber RE *et al.* (2008) Natural variation reveals key amino acids in a downy mildew effector that alters recognition specificity by an *Arabidopsis* resistance gene. *Molecular Plant Pathology*, **9**, 511–523.

Anderson PK, Cunningham AA, Patel NG *et al.* (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution*, **19**, 535–544.

Aylor DL, Price EW, Carbone I (2006) SNAP: combine and map modules for multilocus population genetic analysis. *Bioinformatics*, **22**, 1399–1401.

Bahlo M, Griffiths RC (2000) Inference from gene trees in a subdivided population. *Theoretical Population Biology*, **57**, 79–95.

Baldauf SL (2003) The deep roots of eukaryotes. *Science*, **300**, 1703–1706.

Blair JE, Coffey MD, Park SY, Geiser DM, Kang S (2008) A multilocus phylogeny for Phytophthora utilizing markers derived from complete genome sequences. *Fungal Genetics and Biology*, **45**, 266–277.

Blomquist CL, Irving T, Osterbauer N, Reeser P (2005) *Phytophthora hibernalis*: A new pathogen on *Rhododendron* and evidence of cross amplification with two PCR detection assays for *Phytophthora ramorum*. *Plant Health Progress*, doi: 10.1094/PHP-2005-0728-1001-HN.

Brasier CM, Kirk SA (2004) Production of gametangia by *Phytophthora ramorum in vitro*. *Mycological Research*, **108**, 823–827.

Brown JKM, Hovmoller MS (2002) Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. *Science*, **297**, 537–541.

Chamnanpunt J, Shan WX, Tyler BM (2001) High frequency mitotic gene conversion in genetic hybrids of the oomycete *Phytophthora sojae*. *Proceedings of the National Academy of Sciences, USA*, **98**, 14530–14535.

Davidson JM, Wickland AC, Patterson HA, Falk KR, Rizzo DM (2005) Transmission of *Phytophthora ramorum* in mixed-evergreen forest in California. *Phytopathology*, **95**, 587–596.

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, **3**, 167–169.

Dobrowolski MP, Tommerup IC, Shearer BL, O'Brien PA (2003) Three clonal lineages of *Phytophthora cinnamomi* in Australia revealed by microsatellites. *Phytopathology*, **93**, 695–704.

Donahoo R, Blomquist CL, Thomas SL *et al.* (2006) *Phytophthora foliorum* sp. nov., a new species causing leaf blight of azalea. *Mycological Research*, **110**, 1309–1322.

Dou D, Kale SD, Wang X *et al.* (2008a) Conserved C-terminal motifs required for avirulence and suppression of cell death by *Phytophthora sojae* effector *Avr1b*. *Plant Cell*, **20**, 1118–1133.

Dou D, Kale SD, Wang X *et al.* (2008b) RXLR-mediated entry of *Phytophthora sojae* effector *Avr1b* into soybean cells does not require pathogen encoded machinery. *Plant Cell*, **20**, 1930–1947.

Drenth A, Tas ICQ, Govers F (1994) DNA fingerprinting uncovers a new sexually reproducing population of *Phytophthora infestans* in the Netherlands. *European Journal of Plant Pathology*, **100**, 97–107.

Erwin DC, Ribeiro OK (1996) *Phytophthora Diseases Worldwide*. American Phytopathological Society Press, St Paul, Minnesota.

Fearnhead P, Donnelly P (2002) Approximate likelihood methods for estimating local recombination rates. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, **64**, 657–680.

Felsenstein J (1974) The evolutionary advantage of recombination. *Genetics*, **78**, 737–756.

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) Version 3.6. Distributed by the author, Department of Genome Sciences, University of Washington, Seattle, Washington.

Flier WG, Grünwald NJ, Kroon LPNM *et al.* (2003) The population structure of *Phytophthora infestans* from the Toluca Valley of Central Mexico suggests genetic differentiation between populations from cultivated potato and wild *Solanum* spp. *Phytopathology*, **93**, 382–390.

Frankel SJ (2008) Sudden oak death and *Phytophthora ramorum* in the USA: a management challenge. *Australasian Plant Pathology*, **37**, 19–25.

Fry WE, Goodwin SB (1997) Re-emergence of potato and tomato late blight in the United States. *Plant Disease*, **81**, 1349–1357.

Fry WE, Goodwin SB, Matuszak JM *et al.* (1992) Population genetics and intercontinental migrations of *Phytophthora infestans*. *Annual Review of Phytopathology*, **30**, 107–130.

Fry WE, Goodwin SB, Dyer AT *et al.* (1993) Historical and recent migrations of *Phytophthora infestans*: chronology, pathways, and implications. *Plant Disease*, **77**, 653–661.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **1997**, 915–925.

Fu Y-X, Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Garbelotto M, Svihra P, Rizzo DM (2001) Sudden oak death syndrome fells three oak species. *California Agriculture*, **55**, 9–19.

Garnica DP, Pinzon AM, Quesada-Ocampo LM *et al.* (2006) Survey and analysis of microsatellites from transcript sequences in Phytophthora species: frequency, distribution, and potential as markers for the genus. *BMC Genomics*, **7**, 245.

Gaut BS, Morton BR, McCaig BC, Clegg MT (1996) Substitution rate comparisons between grasses and palms: Synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. *Proceedings of the National Academy of Sciences, USA*, **93**, 10274–10279.

Gomez-Alpizar L, Carbone I, Ristaino JB (2007) An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. *Proceedings of the National Academy of Sciences, USA*, **104**, 3306–3311.

Goodwin SB (1997) The population genetics of *Phytophthora*. *Phytopathology*, **97**, 462–473.

Goodwin SB, Spielman LJ, Matuszak JM, Bergeron SN, Fry WE (1992) Clonal diversity and genetic differentiation of *Phytophthora infestans* populations in northern and central Mexico. *Phytopathology*, **82**, 955–961.

Goodwin SB, Cohen BA, Fry WE (1994) Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. *Proceedings of the National Academy of Sciences, USA*, **91**, 11591–11595.

Griffiths RC, Marjoram P (1996) Ancestral inference from samples of DNA sequences with recombination. *Journal of Computational Biology*, **3**, 479–502.

Grünwald NJ, Flier WG (2005) The biology of *Phytophthora infestans* at its center of origin. *Annual Review of Phytopathology*, **43**, 171–190.

Grünwald NJ, Flier WG, Sturbaum AK *et al.* (2001) Population structure of *Phytophthora infestans* in the Toluca Valley region of Central Mexico. *Phytopathology*, **91**, 882–890.

Grünwald NJ, Goss EM, Larsen MM *et al.* (2008a) First report of the European lineage of *Phytophthora ramorum* on *Viburnum and Osmanthus* spp. in a California nursery. *Plant Disease*, **92**, 314–314.

Grünwald NJ, Goss EM, Press CM (2008b) *Phytophthora ramorum*: a pathogen with a remarkably wide host-range causing sudden oak death on oaks and ramorum blight on woody ornamentals. *Molecular Plant Pathology*, **9**, 729–740.

Grünwald NJ, Kitner M, McDonald V, Goss EM (2008c) Susceptibility in *Viburnum* to *Phytophthora ramorum*. *Plant Disease*, **92**, 210–214.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Hansen EM, Kanaskie A, Prospero S *et al.* (2008) Epidemiology of *Phytophthora ramorum* in Oregon tanoak forests. *Canadian Journal of Forest Research/Revue Canadienne de Recherche Forestiere* **38**.

Hey J, Wakeley J (1997) A coalescent estimator of the population recombination rate. *Genetics*, **145**, 833–846.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Ivors K, Garbelotto M, Vries ID *et al.* (2006) Microsatellite markers identify three lineages of *Phytophthora ramorum* in US nurseries, yet single lineages in US forest and European nursery populations. *Molecular Ecology*, **15**, 1493–1505.

Ivors KL, Hayden KJ, Bonants PJ, Rizzo DM, Garbelotto M (2004) AFLP and phylogenetic analyses of North American and European populations of *Phytophthora ramorum*. *Mycological Research*, **108**, 378–392.

Jiang RHY, Tripathy S, Govers F, Tyler BM (2008) RXLR effector reservoir in two *Phytophthora* species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proceedings of the National Academy of Sciences, USA*, **105**, 4874–4879.

Jones KE, Patel NG, Levy MA *et al.* (2008) Global trends in emerging infectious diseases. *Nature*, **451**, 990–993.

Kasuga T, White TJ, Taylor JW (2002) Estimation of nucleotide substitution rates in eurotiomycete fungi. *Molecular Biology and Evolution*, **19**, 2318–2324.

King DA, Peckham C, Waage JK, Brownlie J, Woolhouse MEJ (2006) Infectious diseases: preparing for the future. *Science*, **313**, 1392–1393.

Kroon LP, Bakker FT, van den Bosch GB, Bonants PJ, Flier WG (2004) Phylogenetic analysis of Phytophthora species based on mitochondrial and nuclear DNA sequences. *Fungal Genetics and Biology*, **41**, 766–782.

Li W-H, Tanimura M, Sharp PM (1987) An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *Journal of Molecular Evolution*, **25**, 330–342.

Linde C, Drenth A, Wingfield MJ (1999) Gene and genotypic diversity of *Phytophthora cinnamomi* in South Africa and Australia revealed by DNA polymorphisms. *European Journal of Plant Pathology*, **105**, 667–680.

Lyngsø RB, Song YS, Hein J (2005) Minimum recombination histories by branch and bound. *Lecture Notes in Computer Science*, **3692**, 239–250.

Martin FN (2008) Mitochondrial haplotype determination in *Phytophthora ramorum*. *Current Genetics*, **54**, 23–34.

Mascheretti S, Croucher PJP, Vettraino A, Prospero S, Garbelotto M (2008) Reconstruction of the Sudden Oak Death epidemic in California through microsatellite analysis of the pathogen *Phytophthora ramorum*. *Molecular Ecology*, **17**, 2755–2768.

May KJ, Ristaino JB (2004) Identity of the mtDNA haplotype(s) of *Phytophthora infestans* in historical specimens from the Irish potato famine. *Mycological Research*, **108**, 471–479.

McVean G, Awadalla P, Fearnhead P (2002) A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics*, **160**, 1231–1241.

Miller JS, Hamm PB, Johnson DA (1997) Characterization of the *Phytophthora infestans* population in the Columbia Basin of Oregon and Washington from 1992 to 1995. *Phytopathology*, **87**, 656–660.

Morens DM, Folkers GK, Fauci AS (2004) The challenge of emerging and re-emerging infectious diseases. *Nature*, **430**, 242–249.

Mosa A, Kobayashi K, Ogoshi A, Kato M, Sato N (1993) Isoenzyme polymorphism and segregation in isolates of *Phytophthora infestans* from Japan. *Plant Pathology*, **42**, 26–34.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19**, 153–170.

Posada D, Crandall KA (1998) ModelTest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Price EW, Carbone I (2005) SNAP: workbench management tool for evolutionary population genetic analysis. *Bioinformatics*, **21**, 402–404.

Prospero S, Black JA, Winton LM (2004) Isolation and characterization of microsatellite markers in *Phytophthora ramorum*, the causal agent of sudden oak death. *Molecular Ecology Notes*, **4**, 672–674.

Prospero S, Hansen EM, Grünwald NJ, Winton LM (2007) Population dynamics of the sudden oak death pathogen *Phytophthora ramorum* in Oregon from 2001 to 2004. *Molecular Ecology*, **16**, 2958–2973.

Rehmany AP, Gordon A, Rose LE *et al.* (2005) Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two *Arabidopsis* lines. *Plant Cell*, **17**, 1839–1850.

Rizzo DM, Garbelotto M, Davidson JM, Slaughter GW, Koike ST (2002) *Phytophthora ramorum* as the cause of extensive mortality of *Quercus* spp. and *Lithocarpus densiflorus* in California. *Plant Disease*, **86**, 205–214.

Rizzo DM, Garbelotto M, Hansen EM (2005) *Phytophthora ramorum*: integrative research and management of an emerging pathogen in California and Oregon forests. *Annual Review of Phytopathology*, **43**, 309–335.

Rozas J, BSanchez-De I, Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Shan W, Cao M, DL.Tyler BM (2004) The *Avr1b* locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene *Rps1b*. *Molecular Plant–Microbe Interactions*, **17**, 394–403.

Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**.

Thompson JD, Higgins DG, Gibson TJ (1994) ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Tyler BM, Tripathy S, Zhang X *et al.* (2006) Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science*, **313**, 1261–1266.

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Werres S, Kaminski K (2005) Characterisation of European and North American *Phytophthora ramorum* isolates due to their morphology and mating behaviour in vitro with heterothallic *Phytophthora* species. *Mycological Research*, **109**, 860–871.

Werres S, Marwitz R, Man In'T, Veld WA *et al.* (2001) *Phytophthora ramorum* sp. nov., a new pathogen on *Rhododendron* and *Viburnum*. *Mycological Research*, **105**, 1155–1165.

Whisson SC, Boevink PC, Moleleki L *et al.* (2007) A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature*, **450**, 115–118.

Wolfe KH, Sharp PM, Li W-H (1989) Rates of synonymous substitution in plant nuclear genes. *Journal of Molecular Evolution*, **29**, 208–211.

Woolhouse MEJ, Gowtage-Sequeria S (2005) Host range and emerging and reemerging pathogens. *Emerging Infectious Diseases*, **11**, 1842–1847.

Erica Goss is a post-doctoral researcher interested in the ecology and evolution of plant pathogens, currently focusing on the evolution and population genetics of the sudden oak death pathogen *Phytophthora ramorum*. Nik Grünwald's research program focuses on the ecology, population genetics and evolution of emerging and exotic oomycete plant pathogens including *P. ramorum*. Ignazio Carbone is an evolutionary biologist interested in the ecology and population biology of aflatoxigenic fungi. He is also developing software tools and workflows to streamline population genetic analyses.

Supporting Information

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

Supplementary Data Comparison of results from recom, LDhat, and sequenceLD

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.