Curr Genet
DOI 10.1007/s00294-015-0480-3

RESEARCH ARTICLE

Mitochondrial genome sequences reveal evolutionary relationships of the *Phytophthora* 1c clade species

Erica S. Lassiter · Carsten Russ · Chad Nusbaum · Qiandong Zeng · Amanda C. Saville · Rodrigo A. Olarte · Ignazio Carbone · Chia‑Hui Hu · Andaine Seguin‑Orlando · Jose A. Samaniego · Jeffrey L. Thorne · Jean B. Ristaino

Received: 20 February 2014 / Revised: 11 February 2015 / Accepted: 12 February 2015
© Springer-Verlag Berlin Heidelberg (outside the USA) 2015

Abstract *Phytophthora infestans* is one of the most destructive plant pathogens of potato and tomato globally. The pathogen is closely related to four other *Phytophthora* species in the 1c clade including *P. phaseoli*, *P. ipomoeae*, *P. mirabilis* and *P. andina* that are important pathogens of other wild and domesticated hosts. *P. andina* is an interspecific hybrid between *P. infestans* and an unknown *Phytophthora* species. We have sequenced mitochondrial genomes of the sister species of *P. infestans* and examined the evolutionary relationships within the clade. Phylogenetic analysis indicates that the *P. phaseoli* mitochondrial lineage is basal within the clade. *P. mirabilis* and *P. ipomoeae* are sister lineages and share a common ancestor with the Ic mitochondrial lineage of *P. andina*. These lineages in turn are sister to the *P. infestans* and *P. andina* Ia mitochondrial lineages. The *P. andina* Ic lineage diverged much earlier than the *P. andina* Ia mitochondrial lineage and *P. infestans*. The presence of two mitochondrial lineages in *P. andina* supports the hybrid nature of this species. The ancestral state of the *P. andina* Ic lineage in the tree and its occurrence only in the Andean regions of Ecuador, Colombia and Peru suggests that the origin of this species hybrid in nature may occur there.

Keywords *Phytophthora* 1 c clade · Mitochondria · Late blight · Phylogenetics

Communicated by L. Tomaska.

Electronic supplementary material The online version of this article (doi:10.1007/s00294-015-0480-3) contains supplementary material, which is available to authorized users.

E. S. Lassiter
Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695, USA

C. Russ · C. Nusbaum · Q. Zeng
BROAD Institute, MIT, Cambridge, MA 02142, USA

A. C. Saville · R. A. Olarte · I. Carbone · C.-H. Hu · J. B. Ristaino (✉)
Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA
e-mail: jean_ristaino@ncsu.edu
URL: http://www.plantpath.cals.ncsu.edu/faculty/jean-beagle-ristaino

A. Seguin-Orlando · J. A. Samaniego
Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

J. L. Thorne
Department of Biological Sciences and Department of Statistics, North Carolina State University, Raleigh NC 27695, USA

Introduction

Phytophthora infestans (Montagne) de Bary, the destructive pathogen responsible for the Irish potato famine of the mid-1800s (Bourke 1964), still causes widespread disease on both potato and tomato (Fry et al. 2013; Hu et al. 2012). *P. infestans* and other oomycetes are members of Stramenopila and are not true fungi (Chesnick et al. 1996; Gunderson et al. 1987; Föster et al. 1990). *P. infestans* is closely related to four other *Phytophthora* species that infect different hosts. The evolutionary relationships of these species has been investigated using multiple approaches including single and multi-gene genealogies and whole genome sequencing (Blair et al. 2012; Gomez et al. 2008; Goss et al. 2011; Raffaele et al. 2010). *P. andina* is a hybrid species, with *P. infestans* as one of the parents and there are two mitochondrial lineages (Ia and Ic) (Adler et al. 2004;

Published online: 10 March 2015

Gomez et al. 2008; Goss et al. 2011; Oliva et al. 2010). However, the whole mitochondrial genomes of all the species in the clade have not been sequenced previously and the evolutionary relationship of both mitochondrial lineages of *P. andina* with the other species in the clade remains unresolved.

The genus *Phytophthora* is divided into many clades and *P. infestans* and its four sister species are found in clade 1c (Kroon et al. 2004; Blair et al. 2008; Cooke et al. 2000). *P. infestans* shares clade 1c with *P. phaseoli*, *P. ipomoeae*, *P. mirabilis* and *P. andina*. *P. phaseoli,* first reported in the USA (Thaxter, 1889), is a species that infects lima and common bean (*Phaseolus limensis, P. lunatus, P. vulgaris*) and has been reported in Mexico and Central and South America, but its center of origin is unknown (Evans et al. 2007; Farr and Rossman 2014). *P. mirabilis* causes disease on *Mirabilis jalapa* (Galindo and Hohl 1985) and *P. ipomoeae* causes leaf blight on *Ipomoea longipedunculata* and these two species have only been reported in Mexico (Flier et al. 2002). On the other hand, *P. andina* has only been reported in Ecuador, Colombia and Peru, but not Mexico (Forbes et al. 2013). The Toluca Valley in central Mexico has been proposed as the center of origin of the entire 1c clade (Brasier and Hansen 1992; Flier et al. 2002, 2003; Goodwin et al. 1994; Goss et al. 2014; Grünwald and Flier 2005), but the Andean region has also been suggested as a center of origin for species in the clade including *P. andina* and *P. inf*estans (Adler et al. 2004; Gomez-Alpizar et al. 2007, 2008). *P. infestans*, *P. mirabilis* and *P. ipomoeae* share morphological characteristics, but have different mating systems. Sympatric evolution of three of the species (*P. infestans*, *P. mirabilis* and *P. ipomoeae*) found in the highlands of central Mexico has occurred (Grünwald and Flier 2005). Host jumps followed by host specialization through mutations of effector genes may be driving speciation in the clade (Dong et al. 2014; Flier et al. 2002; Raffaele et al. 2010).

In Ecuador, three clonal lineages of *P. infestans* (US-1, EC-1, EC-3) and one heterogeneous lineage (EC-2) consisting of two mitochondrial haplotypes (EC-2 Ic, EC-2 Ia) formerly named *Phytophthora infestans* sensu lato are found in association with different wild *Solanum* species (Adler et al. 2002, 2004; Forbes et al. 2013; Ordonez et al. 2000). Kroon et al. (2004) sequenced several nuclear and mitochondrial genes from the EC-2 Ic mitochondrial lineage and placed this lineage in the 1c clade and was the first to use the name "*P. andina*". Subsequently, Blair et al. (2008) referred to the species as "*P*. sp. '*andina";* however, the evolutionary relationships between the sister species in clade 1c remained unresolved (Kroon et al. 2004).

The EC-1 clonal lineage of *P. infestans sensu lato* was later determined to be *P. infestans* based on sequences of the mitochondrial cytochrome oxidase 1 (*cox1*) gene and intron 1 of the *ras* gene (Gómez-Alpizar et al. 2008). However, the EC-2 Ic lineage formed a branch in the 1c clade distinct from *P. infestans, P. mirabilis, P. phaseoli and P. ipomoeae,* for both *cox 1* and *ras* intron 1, and these isolates were subsequently identified as the newly described species *P. andina* (Gómez-Alpizar et al. 2008; Oliva et al. 2010). A formal species description of *P. andina* was published and it was suggested that in addition to the EC-2 Ic lineage the EC-3 lineage (Ia) from *Solanum betaceum* was also *P. andina* (Oliva et al. 2010). Other authors have questioned (Cardenas et al. 2012) or defended (Forbes et al. 2012) the validity of *P. andina* as a species, while some of the same authors showed a hybrid origin of *P. andina* (Goss et al. 2011). Molecular evidence from several nuclear and mitochondrial genes from three different studies suggests that *P. andina* is a hybrid and that *P. infestans* is clearly one of the parents, but the other parental lineage is unknown (Blair et al. 2012; Gomez et al. 2008; Goss et al. 2011).

P. mirabilis is closely related to *P. infestans* and was first described in Mexico on *Mirabilis jalapa* also known as 4'o clock or the "Flower of Peru" (Galindo and Hohl 1985; Goodwin and Fry 1994). Studies have shown that *P. infestans* and *P. mirabilis* are now reproductively isolated, but the possibility of gene flow between the species was demonstrated since the two species can interbreed and produce viable offspring (Goodwin et al. 1999). *P. mirabilis* does not appear to be conspecific with *P. infestans* anywhere except Mexico. *Mirabilis jalapa* occurs widely in South America, but there have been limited surveys there for infection of the host by *P. mirabilis*. In contrast, the hybrid species *P. andina* is native to the Andean region of South America and has not been found in Mexico where *P. infestans* and *P. mirabilis* are sympatric. Thus, the hybridization between *P. infestans* and the unknown *Phytophthora* species that led to *P. andina* may have occurred in the Andes where both *P. andina* and *P. infest*ans are sympatric and infect common hosts (Forbes et al. 2013). The oldest mutations in the mitochondrial lineages leading to *P. andina* are of South American origin (Gomez et al. 2007; Martin et al. 2014). In addition, several Andean *Solanaceous* species including pear melon (*Solanum muricatum*), tree tomato (*S. betaceum*) and naranjilla—little orange (*Solanum quitoense)* are reported hosts to both *P. infestans* and *P. andina*, so gene flow to wild *Solanum species* is a possibility (Adler et al. 2004; Forbes et al. 2013). Fruits of all these species are imported into the USA posing phytosanitary hazards.

In this study we investigated the evolutionary history of *Phytophthora species* in the 1c clade by: (1) sequencing and annotating the complete mitochondrial genomes of the remaining four species of the *Phytophthora* 1c clade: *P. mirabilis*, *P. ipomoeae*, *P. phaseoli* and both the Ia and Ic mitochondrial lineages of *P. andina*; (2) aligning and

Table 1 Isolates and GenBank accession numbers of *Phytophthora species* used for mitochondrial genome sequencing

Species[a]	Isolate	Collection[b]	Origin, references	GenBank accession number (lineage)
P. infestans	West Virginia 4	ATCC16981	WVA, Lang and Forget (1993)	NC002387 (US-1)
P. phaseoli	Phy P18	WPC P11078	Delaware, Evans et al. (2007)	HM590418
P. andina (Ic)	EC 3425	CBS 122202	Ecuador, Adler et al. (2004)	HM590419 (EC-2)
P. andina (Ia)	EC 3394	NA	Ecuador, Adler et al. (2004)	KJ408269 (EC-3)
P. ipomoeae	Pic 99167	CBS 122203	Mexico, Flier et al. (2002)	HM590420
P. mirabilis	Pic 99114	CBS 122204	Mexico, Flier et al. (2002)	HM590421

[a] *Phytophthora* species in the 1c clade included in the study. Mitochondrial lineages are shown in parentheses

[b] *ATCC* American Type Culture Collection, Manassas, Virginia; *WPC* World *Phytophthora* Collection; *CBS* Fungal Diversity Centre, Netherlands; *NA* not applicable

comparing the genomes to the mitochondrial genomes of *P. infestans*; and (3) resolving the evolutionary relationships among these mitochondrial lineages using Bayesian statistical analyses. A preliminary abstract of a portion of this research has been published (Lassiter et al. 2010).

Materials and methods

Isolates

Five isolates of *Phytophthora* were used for this study including: *P. mirabilis*, *P. ipomoeae*, *P. phaseoli* and *P. andina* (mitochondrial lineages Ic and Ia) (Table 1). The mitochondrial genome sequence of the *P. infestans* Ib mt haplotype (NC 002387) was used as the reference genome for comparison (Lang and Forget 1993). All isolates were maintained on rye-V8 or lima bean agar at 18 °C. Two to three agar plugs containing mycelia were transferred into Petri dishes containing pea broth with 0.5 g/L sucrose and grown for 2–3 weeks. Mycelia from approximately 40 Petri dishes were harvested by filtration with Whatman #1 membrane filter paper and then frozen in liquid nitrogen. Mycelia were stored at −20 °C until further use. Approximately, 10 g of mycelia was used for DNA extraction.

DNA preparation

Previously collected mycelia were frozen in liquid nitrogen and then ground into a fine powder using a mortar and pestle. DNA extraction was performed using the Qiagen DNeasy Plant Maxi® kit (Qiagen Corporation, Maryland, USA). The mtDNA was purified in a cesium chloride (CsCl) density gradient (1.1 g CsCl/mL DNA-AE solution) with the addition of bisbenzimide (final concentration of 10 mg/mL) as described previously (Avila-Adame et al. 2006; Carter et al. 1990; Garber and Yoder 1983; Klimczak and Prell 1984). The DNA was purified by two consecutive centrifugations in a Beckman ultracentrifuge with a 70.1 Ti rotor at 200,000*g* for 42–48 h at 20 °C. The bisbenzimide and CsCl were removed with the use of isopropanol and dialysis in TE buffer (at 4 °C for 16 h including three changes of buffer), respectively. The mtDNA was concentrated by ethanol precipitation in the presence of 1/10 3.0 M sodium acetate (pH 5.2) and resuspended by dissolving in AE buffer provided in the Qiagen kit. The *P. andina* Ia mitochondrial DNA was prepped using a standard CTAB extraction (Ristaino et al. 2001).

Sequencing

The mitochondrial genomes of *P. ipomoeae, P. phaseoli* and *P. andina* (Ic haplotype) were sequenced at the BROAD Institute of MIT with a minimum of 100-fold sequence coverage with standard 454 technology fragment reads using methods described previously (Lennon et al. 2010). The mtDNA was sheared and size selected to generate fragments from 400 to 800 bp in length. Fragments were blunt ended, ligated on both ends with 454 adapter sequences and the resulting library was subjected to emulsion PCR. The DNA was then sequenced using the 454 Titanium sequencing platform. The *P. andina* Ia genome was sequenced at the Centre for GeoGenetics, University of Copenhagen. Initially, DNA was fragmented using a Bioruptor to a mean size of 200 bp, following which an Illumina-compatible library was constructed using the NEBNext library build kit 1 and Illumina PE adaptors, following the manufacturer's instructions. The library was subsequently index PCR amplified (Illumina index 1) for 20 cycles using Pfusion polymerase, prior to gel excision, quantification and sequencing over one lane of Illumina GAIIx (SR 76 bp) and one lane of Illumina HiSeq (SR100 bp) following the manufacturer's guidelines.

Annotation

The 454 reads were reassembled using Newbler (Roche). The identification of genes, open reading frames and other

Fig. 1 Mitochondrial genome map of *P. andina* haploptype Ic (size 37,874 bp). Gene order, content and size are conserved among the 1c clade species of *Phytophthora*. Full mitogenome maps of other 1c clades species are in Supplemental Figure 1

features in the mitochondrial genomes was based on the previously published mitochondrial genome sequence of *P. infestans* (mtDNA haplotype Ib available in GenBank accession no. NC002387) (Lang and Forget 1993; Paquin et al. 1997). Portions of the four genomes were aligned with the *P. infestans* genome using Clustal W (Version 1.8.1) in BioEdit 7.0.5.3 (Hall 1999). Each gene within the *P. infestans* genome was locally aligned with the entire genome of each of the remaining species. The mitochondrial genomes of each species were analyzed and mapped using the software program Vector NTI version 11.52 (Fig. 1; Supplemental Figure 1). Annotated sequence data were submitted to GenBank and the accession numbers are reported in (Table 1).

The *P. andina* Ia genome was annotated using the *P. andina* Ic mitochondrial genome. AdapterRemoval (Lindgreen 2012) was used to trim adaptors, N bases and low-quality sequence from the raw Illumina reads using default parameters. Sequences were initially mapped using BWA against the *P. andina* Ic mitochondrial genome (HM590419). Subsequently, duplicate reads were removed using samtools (Li et al. 2009) and reads that mapped to more than one position were removed using shell scripts. The GATK pipeline (McKenna et al. 2010) was then used to realign the reads and to call SNPs and indels, following which a consensus sequence was generated using custom perl scripts. Bases were called consensus if a called SNP was present in more than 70 % of the reads mapping to one position. Post-trimming, 46,184 correctly indexed reads mapped uniquely to the reference, resulting in an assembly with 91.08× mean coverage (median 71×), and covering 99.735 % of the reference at minimum 1×.

Evolutionary analysis

Mitochondrial genome sequences were aligned using Align X in Vector NTI, version 11.52 edited, trimmed and collapsed into haplotypes using SNAP Map (Aylor et al. 2006) with the options of removing insertions and deletions and excluding infinite-sites violations. Base substitutions were categorized as transitions/transversions and phylogenetically informative/uninformative. Compatibility among pairs of informative sites and the largest nonrecombining partition were determined using the Clade and CladeEx programs (Bowden et al. 2008), respectively.

Table 2 Mitochondrial genome size and content of *Phytophthora* species in 1c clade

Species	Genome size (bp)	% Identity[a]	Protein coding		Non-coding		G + C content (%)
			bp	%	bp	%	
P. infestans (Ib)	37,957	–	34,209	90.1	3,748	9.9	22.3
P. phaseoli	37,914	99.0	34,284	90.4	3,630	9.6	22.4
P. andina (Ia)	37,883	99.9	34,272	90.4	3,611	9.5	22.2
P. andina (Ic)	37,874	99.3	34,272	90.4	3,602	9.5	22.1
P. ipomoeae	37,872	99.3	34,284	90.5	3,670	9.7	22.4
P. mirabilis	37,778	99.4	34,284	90.7	3,494	9.2	22.4

[a] % identity is calculated from alignment of protein coding regions and compared to the *P. infestans* Ib mitochondrial genome sequence

Table 3 Number of polymorphisms found among the *Phytophthora* 1c clade species

Species	*P. infestans*	*P. phaseoli*	*P. andina* (Ic)	*P.andina* (Ia)	*P. ipomoeae*	*P. mirabilis*
P. infestans	0	337	225	14	238	189
P. phaseoli	–	0	364	343	380	344
P. andina (Ic)	–	–	0	226	241	199
P. andina (Ia)	–	–	–	0	239	191
P. ipomoeae	–	–	–	–	0	213
P. mirabilis	–	–	–	–	–	0

Polymorphic sites include 672 single base pair substitutions within protein coding sites; these sites do not include indels greater than 1 bp within the protein coding region

Coalescent analysis using Genetree (Griffiths and Tavaré 1994) and implemented in SNAP Workbench (Price and Carbone 2005; Carbone et al. 2004) was used to infer the mutational history, time scale and evolution of the polymorphisms among the species. The largest nonrecombining partition was collapsed into nine haplotypes and 633 variable sites (Supplemental Figure 2). Only regions corresponding to protein coding genes, tRNAs or rRNAs were used for coalescent inference. These comprised nearly 90 % of the mitogenomes and gave good estimates of the species trees. The relative probabilities of all rooted genealogies were calculated by performing four independent runs of 10 million simulations of the coalescent assuming Watterson's (1975) estimate of theta, panmixis and constant population size. The mitochondrial genealogy with the highest root probability was examined to determine the relative ages of polymorphisms and clades in the evolutionary history of the species.

Analysis was also conducted using BEAST (Drummond and Rambaut 2007) to confirm the findings in the SNAP workbench. An xml file was created in BEAUti with simulations run for 10,000,000 generations of Markov chain Monte Carlo with a burn in of 1,000,000 under the HKY substitution model (Hasegawa et al. 1985). Trees were sampled every 1000 generations and summarized into a consensus tree using TreeAnnotator, with a burn in of 1000 trees and a posterior probability limit of 0. The tree was visualized in FigTree 1.4.6 (Rambaut 2007).

Results

We sequenced the complete mitochondrial genomes of *P. andina* (haplotypes Ia and Ic), *P. mirabilis*, *P. ipomoeae* and *P. phaseoli*, and genome maps and genome sizes are shown (Fig. 1; Supplemental Figure 1 A–C; Table 2). The mitochondrial genome sequence for the Ib haplotype of *P. infestans* was used for comparison (Lang and Forget 1993; Paquin et al. 1997). All the mitogenomes were similar in size and varied by only 279 bp. The mitochondrial genome of *P. infestans* was larger than the mitochondrial genomes of the other species and *P. mirabilis* had the smallest mitochondrial genome. The genomes of both the *P. andina* lineages were similar in size.

Protein coding regions accounted for over 90 % of the mitogenomes and sequence identity was greater than 99 % (Table 2). The *P. andina* Ia mitogenome had the greatest sequence similarity to the *P. infestans* mitogenome and only 14 polymorphisms were found between the mitogenomes of the two species (Table 3). All five mitogenomes were adenine and thymine rich and G-C content was less than 23 % in each of the species (Table 2). A total of 61 genes with known function were identified including 18 genes involved in electron transport, 2 ribosomal RNA genes, 16 ribosomal protein genes and 25 transfer RNA genes (Supplemental Table 1; Fig. 1). Six ORFs which code for proteins of unknown function were also identified. Genes were coded in

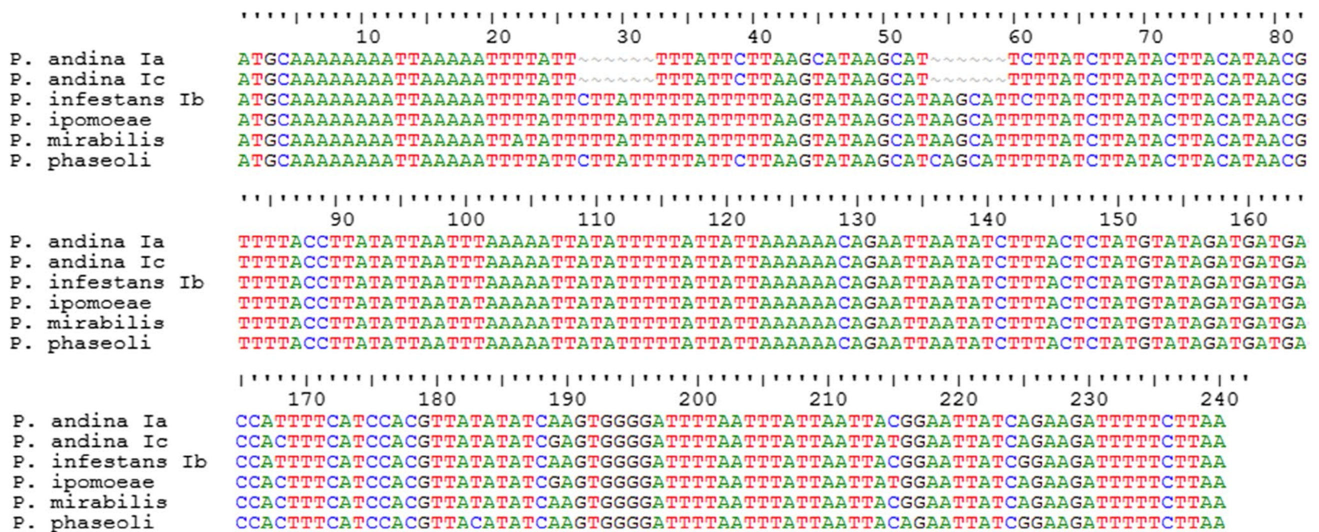


Fig. 2 Schematic of *P. andina* ORF 79 and the other 1c clade species showing a 12 nucleotide deletion present only in the two *P. andina* mitochondrial lineages and not the other species in the 1c clade

both strands of the DNA, and gene order was identical among all five 1c clade species (Fig. 1; Supplemental Figure 1A–C; Supplemental Table 1). None of the seven ORFs identified previously in the 2 kb insertion found in type II mitochondrial genomes of *P. infestans* were found among the four sister species (Avila-Adame et al. 2006). The genes coding for the ribosomal proteins *rps*12 and *rps*7 overlapped. Similarly, there was an observed overlap between the genes coding for the NADH dehydrogenase subunits 1 (*nad*1) and 11 (*nad*11). Two indels in ORF 79 were detected in both haplotypes of *P. andina* corresponding to a 12 nucleotide deletion in the region (Fig. 2). These indels were not found in any of the mitogenomes of the other 1c clade species and did not correspond to any of the known restriction enzyme sites used to differentiate the *P. andina* Ic mitochondrial lineage from *P. infestans* (Ordonez et al. 2000). Many restriction sites were identified among the mitogenomes and can be visualized by downloading the genomes into Vector NTI.

A total of 672 polymorphic sites were detected among the mitochondrial genomes of five species in the clade (Supplemental Figure 2). Among these polymorphic sites were 198 transversions and 474 transitions (Supplemental Figure 2). A total of 129 nonsynonymous substitutions were identified (Supplemental Table 3). None of these polymorphisms produced frameshifts. The highest numbers of SNPs were found in the *nad*11 ($n = 50$), *nad*5 ($n = 38$), *cox*1 ($n = 36$) and *atp*1 ($n = 30$) genes (Fig. 3; Supplemental Table 2). ORF217 ($n = 19$) and ORF79 ($n = 13$) had higher numbers of SNPs than the other hypothetical ORFs. There were 19 and 14 SNPs, respectively, in the genes coding for *rnl* and *rns*. Fourteen of the 25 genes coding for tRNAs had SNPs, but there were fewer than five SNPs in these genes.

We collapsed the sequences into nine haplotypes with 633 variable sites (Fig. 4). Coalescent analysis indicates that *P. phaseoli* is the basal lineage of the clade and diverges first (Fig. 4). The lineage leading to *P. andina* Ic diverged earlier than the other species. The *P. andina* Ic lineage shares a common ancestor with *P. mirabilis* and *P. ipomoeae*. *P. mirabilis* and *P. ipomoeae* are inferred to be sister lineages distinct from the *P. andina* Ic lineage (Fig. 4). The *P. infestans* type I and *P. andina* Ia mitochondrial lineages are of more recent origin, but share a common ancestor with the lineages leading to the remaining three species (*P. andina* Ic, *P. ipomoeae* and *P. mirabilis*). *P. infestans* is the most recently diverged lineage in coalescent time and includes the four mitochondrial lineages (Ia, Ib, IIa and IIb) and the *P. andina* Ia lineage (Fig. 4). The tree resolved using BEAST had a similar topology to the coalescent tree, but shows the *P. andina* Ia lineage diverging before the type I *P. infestans* lineages. However, both are inferred as sister lineages (Supplemental Figure 3).

Each branch on the haplotype tree indicated multiple fixed substitutions that accumulated in the mitochondrial genomes (Fig. 4). *P. phaseoli* had the greatest number of fixed substitutions (190) and this number was higher than any other species of the clade. *P. ipomoeae* had the second highest number of fixed substitutions (112) followed by the *P. andina* Ic lineage (94) and *P. mirabilis* (69). The lineage that led to the separation of the other five members of the clade from *P. phaseoli* included 28 substitutions (Fig. 4). The locations of the fixed substitution and polymorphisms in each mitochondrial genome are shown in Supplemental Table 2.

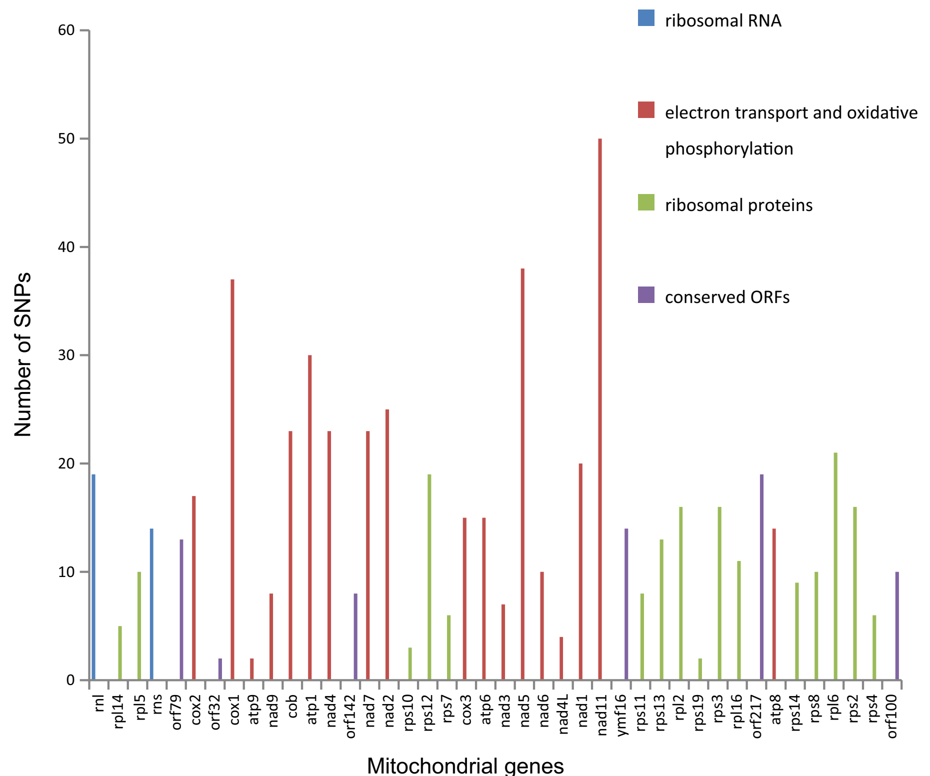


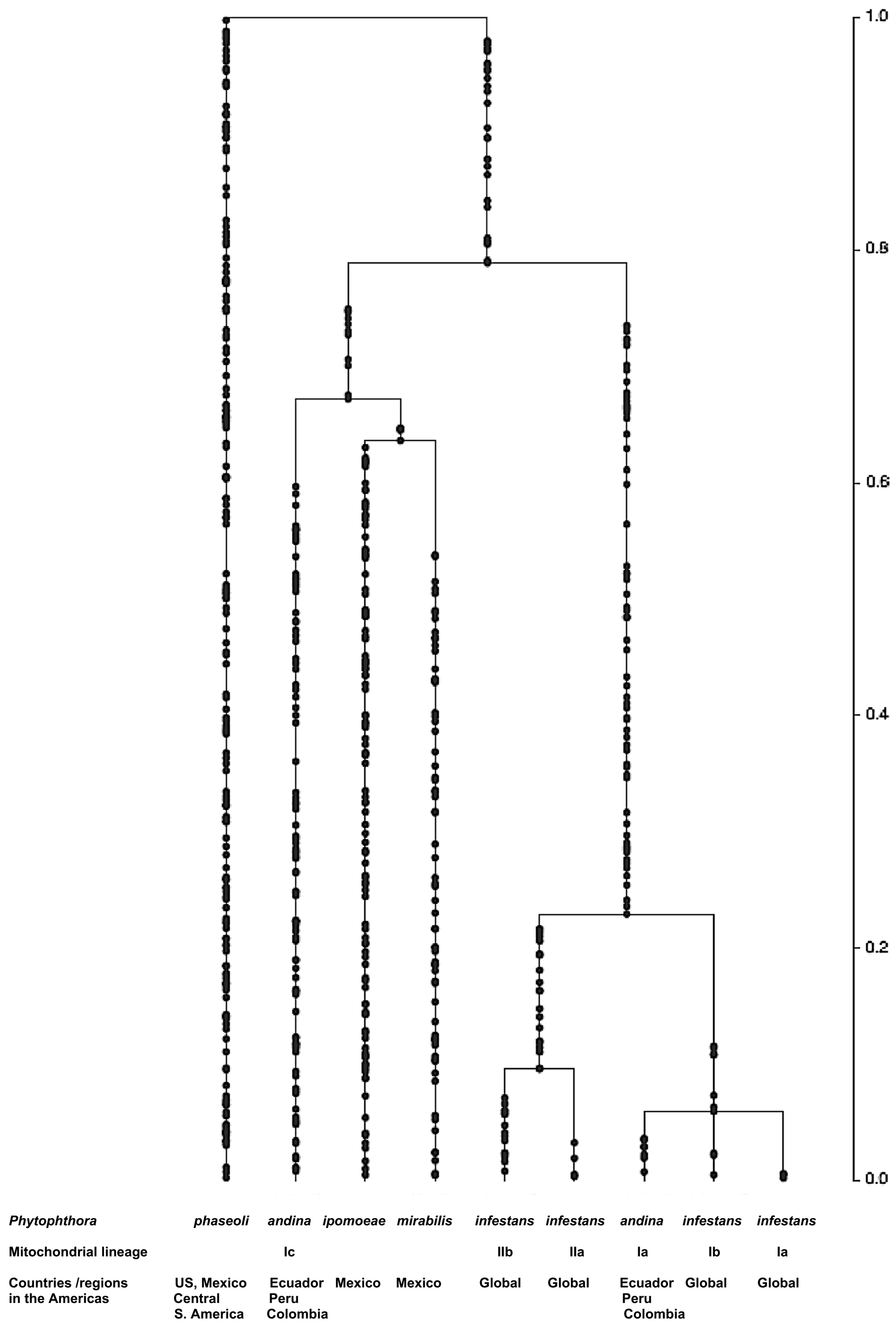
Fig. 3 Number of SNPs in genes involved in electron transport and oxidative phosphorylation, ribosomal proteins and ORFs in the *Phytophthora* 1c clade species compared to the *Phytophthora infestans* Ib mitochondrial genome

Discussion

The mitochondrial genomes of the four 1c clade species of *Phytophthora* were sequenced and annotated in this research. We mapped the mitochondrial genomes of *P. phaseoli*, *P. mirabilis*, *P. ipomoeae* and both haplotypes of *P. andina* (Ic and Ia), compared gene content, order and genome size, and mapped polymorphisms. The gene content was highly conserved among the species; however, there were clear differences and many polymorphisms in the protein-coding genes among the species. *P. phaseoli* was ancestral in the clade. The lineage leading to *P. andina* Ic diverged earlier then the lineage leading to *P. infestans and P. andina* (Ia) which were inferred as sister species. The two *P. andina* mitochondrial lineages contained over 200 SNPs that distinguished them from one another and the presence of two mitochondrial lineages documents the hybrid nature of this species.

The date of divergence of the *P. infestans* mitochondrial lineages has been estimated recently (Martin et al. 2014; Yoshida et al. 2013). Bayesian phylogenies using *P. infestans* mitogenomes from modern and historic lineages dated the estimated coalescence times of all the mitogenomes of *P. infestans* to 460–638 years. This includes the mitochondrial lineage reported to be present during famine-era epidemics (Herb-1) that is a sister lineage and a type Ia mitochondrial lineage of *P. infestans* (Martin et al. 2014; Yoshida et al. 2013). The divergence time of *P. infestans* from *P. mirabilis* was estimated to be 1318 years (Yoshida el al. 2013). Thus, the divergence time of the Ic lineage of *P. andina* and its sister species, *P. mirabilis* and *P. ipomoeae*, and the divergence time of the ancestral lineage in the clade, *P. phaseoli*, from all the other species must be much earlier. Although first reported in the USA, *P. phaseoli* is present in Mexico and has been reported in Central and South America. Studies on the center of the origin of *P. phaseoli* are needed and may be important to understand the origins of the entire clade, since this species is clearly ancestral. Research is in progress to date the divergence times of some of the species in the clade with a larger sample of recent and historic mitogenomes.

Mitochondrial haplotypes have been designated in modern populations of *P. infestans* using PCR and RFLP analysis of mitochondrial DNA (Carter et al. 1990; Griffith and Shaw 1998). Three extant mtDNA haplotypes of *P. infestans* were sequenced by Avila et al. (2006) and mutations leading to the evolution of both type I and type II mitochondrial lineages were clearly identified. The mitochondrial genomes of the other *Phytophthora* species in the 1c clade more closely resemble the type I mitogenomes of *P. infestans* than the type II mitogenomes, since all lacked the large 2 kb insertion sequence (Avila-Adame et al. 2006; Carter et al. 1990; Gavino and Fry 2002). Our data also suggest that the ancestral *P. infestans* parent of the *P. andina* hybrid was a type I mitochondrial lineage.


1.0
0.8
0.6
0.4
0.2
0.0
Phytophthora
phaseoli
andina
ipomoeae
mirabilis
infestans
infestans
andina
infestans
infestans
Mitochondrial lineage
Ic
IIb
IIa
Ia
Ib
Ia
Countries /regions in the Americas
US, Mexico Central S. America
Ecuador Peru Colombia
Mexico
Mexico
Global
Global
Ecuador Peru Colombia
Global
Global

◀ **Fig. 4** The rooted coalescent genealogy showing the distributions of mutations of *Phytophthora* 1c clade species using whole mitogenomes generated in GENETREE. The analyses were run a total of 10,000,000 simulations four separate times. *P. phaseoli* is basal in the clade. *Phytophthora andina* Ic lineage diverged next followed by the other two species (*P. mirabilis and P. ipomoeae*). The *Phytophthora* species, mitochondrial lineage and regions/countries where the species have been reported are indicated *below* the figure. The time scale is in coalescent units of effective population size and the direction of divergence is from *top* (past oldest) to the *bottom* (present youngest)

The center of origin of *P. infestans* and its close relatives is controversial and has been proposed as either central Mexico or the Andean region of South America (Abad and Abad 1997; Bourke 1964; Flier et al. 2003; Gavino and Fry 2002; Gómez-Alpizar et al. 2007; Goodwin et al. 1994; Goss et al. 2014; Grünwald and Flier 2005). Our goal in this research was not to address this question, but some of our data shed light on the controversy. *P. infestans*, *P. mirabilis*, *P. andina* and *P. ipomoeae* coalesced into the same group, indicating that these four species share a common ancestor. To date, *P. andina* has only been reported in the Andean region of Colombia, Ecuador and Peru in South America. This suggests that the earliest divergence of the *P. andina* Ic lineage from a *P. infestans* ancestor may have also occurred in the Andean region. *Phytophthora ipomoeae* and *P. mirabilis* are inferred as sister species and evolved more recently from a common ancestor of the Ic lineage of *P. andina*. It is unresolved whether these sister lineages occur in South America.

The host of *P. mirabilis, Mirabilis jalapa,* and the host plant for *P. ipomoeae*, *Ipomoeae longipedunculata,* both originated from South America and were transported to Mexico hundreds of years ago; none are native to Mexico (Hawkes 1990). One potential scenario includes the possibility that these species jumped hosts from potato to wild hosts in Mexico (Dong et al. 2014; Grünwald and Flier 2005). An alternative possibility is that the host jumps occurred in the Andean region and that all the 1c clade species evolved in South America on the hosts that are native there and were later moved northward on infected plants. The observation that *P. ipomoeae* and *P mirabilis* are extremely rare and difficult to find in Mexico suggests the idea of a limited introduction (Goss et al. 2014; Grünwald and Flier 2005).

The ORF 79 indel found in both *P. andina* mitochondrial lineages can readily be used to differentiate *P. andina* from the other 1c clade species. This site may be used with other known regions to differentiate *P. andina* from *P. infestans*. Creating a PCR primer specific to the *P. andina* ORF79 could enable a quick diagnostic tool for identification of *P. andina* in infected plant material. It is interesting that the indel in ORF 79 is present in both *P andina* mitochondrial lineages so this deletion event could have occurred multiple times in the evolutionary history of *P. andina*. Another possibility is that deletion event may have occurred in the *P. andina* mitogenomes by a hybridization event before the lineages diverged. In addition, the region of the mitochondrial genome close to ORF 79 is an area where other larger insertion events have been reported in *P. infestans* mitochondrial haplotypes (Avila-Adame et al. 2006). The type two lineages contain a 2 KB insertion near this region of the mitogenome.

There are many SNPs distributed through the mitochondrial genomes of *P. infestans* and its sister species, yet only a few of these have been used previously for PCR RFLP haplotyping (Griffith and Shaw 1998). In this report we identified some protein-coding genes that are more variable including *nad*11, *nad*5 and *atp* 1. These SNPs could be useful for SNP typing. The many SNPs reported here in the *P. andina* lineages may be useful for haplotyping populations of *P. andina* on various hosts and discerning the impact of host specialization on this species evolution. Unique SNPs that distinguish species within the 1c clade have been reported previously in the mitochondrial *cox* 1 gene and barcoding genes and nuclear genomes and include genes for host specialization and virulence (Dong et al. 2014; Robideau et al. 2011). A combination of SNP typing and species identification using ORF 79 could be useful in larger sampling studies to examine the role of host jumps and specialization in evolution of both *P. infestans* and its sister species.

A study by Farrer et al. (2011) on the pathogen *Batrachochytrium dendobatidis* demonstrated that multiple lineages of the fungus are being vectored between continents through the movement of the amphibian host. This is likely true with *P. infestans* as well as other species in the clade, since *Phytophthora* species move readily in infected plant material. Farrer et al. (2011) used whole genome sequences to determine global emergences of various lineages. A similar method, using the next generation sequencing of whole genomes of many isolates of 1c clade *Phytophthora* species, is underway to better characterize the evolutionary relationships and origin of the species in this clade. Martin et al. (2014) and Yoshida et al. (2013) examined the evolutionary relatedness of historic and present-day *P. infestans* using whole mitogenomes. Yoshida suggested that a single mitochondrial lineage named Herb-1 was present in the nineteenth century Europe and that this lineage was rare or even extinct in modern populations. Martin et al. (2014) subsequently documented that multiple mitochondrial haplotypes were present in nineteenth century *P. infestans* in Europe, suggesting either multiple introductions of different lineages or a single introduction of multiple lineages. Resolution of the phylogeography and center of origin debate for species in the lc clade of *Phytophthora* will have serious

implications for both plant breeding and phytosanitary regulations, since present-day movement of plant materials out of Central and South America to the USA could result in the introduction of novel species on exported Solanaceous fruits and tubers.

Acknowledgments A graduate teaching assistantship from the Department of Genetics at NC State University funded Erica Lassiter's work. The senior author is now employed at Bayer Crop Science RTP, NC. Some of the research (J.L. Thorne) was partially supported by NIH Grant GM070806 and supplies were funded by USDA NIFA Grant #2006-55319-16550 and Agriculture and Food Research Initiative Competitive Grants Program (AFRI) Grant 2011-68004-30154. The Copenhagen-based sequencing was funded by Lundbeck Foundation Grant R52-5062 and appreciation is expressed to Tom Gilbert's laboratory for sequencing the Ia lineage of *P. andina*. Appreciation for providing cultures is expressed to: Gregory Forbes and Ricardo Oliva International Potato Center (Oliva now at the International Rice Research Institute, Philippines) for *P. andina;* Peter Bonants, Plant Research International, Wageningen for *P. ipomoeae*; Mike Coffey, University of California Riverside for *P. mirabilis*; and Tom Evans and Nancy Gregory, University of Delaware, for *P. phaseoli* used in this work. The technical support of Dr. Julia Hu, Dr. Monica Blanco Menenses and Caleb Pearce is appreciated. Dr. Geromy Moore, Dr. Ben Redelings, Dr. Reed Cartwright and Kristin Lamm helped troubleshoot analytical methods and simulations.

References

Abad ZG, Abad JA (1997) Another look at the origin of late blight of potatoes, tomatoes and pear melon in the Andes of South America. Plant Dis 81:682–688

Adler NE, Chacón G, Forbes G, Flier W (2002) *Phytophthora infestans* sensu lato in South America population substructuring through host-specificity. In: Lizárrega C (ed) Proceedings of the Global Initiative on Late Blight Conference. International Potato Center, Lima, pp 13–17

Adler NE, Erselius LJ, Chacon MG, Flier WG, Ordonez ME, Kroon LPNM, Forbes GA (2004) Genetic diversity of *Phytophthora infestans* sensu lato in Ecuador provides new insight into the origin of this important plant pathogen. Phytopathology 94:154–162

Avila-Adame C, Gómez-Alpizar L, Zisman V, Jones K, Buell CR, Ristaino JB (2006) Mitochondrial genome sequences and molecular evolution of the Irish potato famine pathogen, *Phytophthora infestans*. Curr Genet 49:39–46

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

Blair JE, Coffey MD, Park S-Y, Geiser DM, Kang S (2008) A multilocus phylogeny for *Phytophthora* utilizing markers derived from complete genome sequences. Fungal Gen Biol 45:266–277

Blair JE, Coffey MD, Martin FN (2012) Species tree estimation for the late blight pathogen, *Phytophthora infestans*, and close relatives. PLoS One 7:e37003

Bourke PM (1964) Emergence of potato blight. Nature 203:805–808

Bowden LC, Price EW, Carbone I (2008) SNAP Clade and Matrix, Version 2. Distributed over the Internet, http://www.cals.ncsu.edu/plantpath/faculty/carbone/home.html. Department of Plant Pathology, North Carolina State University

Brasier CM, Hansen EM (1992) Evolutionary biology of *Phytophthora*. Part II: phylogeny, speciation, and population structure. Ann Rev Phytopathol 30:173–200

Carbone I, Liu Y-C, Hillman BI, Milgroom MG (2004) Recombination and migration of *Cryphonectria* hypovirus I as inferred from gene genealogies and the coalescent analysis. Genetics 166:1611–1629

Cárdenas M, Tabmia J, Fry WE, Grünwald NJ, Bernal A et al (2012) Restrepo SDefining species boundaries in the genus Phytophthora: the case of *Phytophthora andina*. A response to '*Phytophthora andina* sp. nov., a newly identified heterothallic pathogen of solanaceous hosts in the Andean highlands' (Oliva, 2010). Plant Pathol 61:215–220

Carter DA, Archer SA, Buck KW (1990) Restriction fragment length polymorphism of mitochondrial DNA of *Phytophthora infestans*. Mycol Res 8:1123–1128

Chesnick JM, Tuxbury K, Coleman A, Lang BF (1996) Utility of the mitochondrial *nad4L* gene for algal and protistan phylogenetic analysis. J Phycol 32:452–456

Cooke DEL, Drenth A, Duncan JM, Wagels G, Brasier CM (2000) A molecular phylogeny of *Phytophthora* and related oomycetes. Fungal Gen Biol 30:17–32

Dong et al (2014) Effector specialization in a lineage of the Irish potato famine pathogen. Science 343:552–555

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214

Evans TA, Mulrooney R, Gregory NF, Kee E (2007) Lima bean downy mildew: impact, Etiology, and management strategies for Delaware and the Mid-Atlantic Region. US. Plant Dis. 91:128–135

Farr DF, Rossman AY (2014) Fungal Database Systematic Mycology and Microbiology Laboratory. ARS, USDA. http://www.nt.ars-grin.gov/fungaldatabases/

Farrer RA et al (2011) Multiple emergences of genetically diverse amphibian-infecting chytrids include a globalized hypervirulent recombinant lineage. Proc Nat Acad Sci 108:18732–18736

Flier WG, Grünwald NJ, Kroon LPNM, van den Bosch TBM, Garay-Serrano E, Lozoya-Saldana Bonants PJM, Turkensteen LJ (2002) *Phytophthora ipomoeae* sp. nov., a new homothallic species causing leaf blight on *Ipomoea longipedunculata* in the Toluca Valley of Central Mexico. Mycol Res 106:848–856

Flier WG, Grünwald NJ, Kroon LPNM, Sturbaum AK, van den Bosch TBM, Garay-Serrano E, Lozoya-Saldana H, Fry WE, Turkensteen LJ (2003) The population structure of *Phytophthora infestans* from the Toluca Valley of Central Mexico suggests genetic differentiation between populations of cultivated potato and wild *Solanum* spp. Phytopathology 93:382–390

Forbes GA, Ristaino JB, Oliva RF, Flier W (2012) A rebuttal to the letter to the editor concerning "Defining species boundaries in the genus *Phytophthora*: the case of *Phytophthora andina*". Plant Pathol 61:221–223

Forbes GA, Morales JG, Restrepo S, Perez S, Gamboa S, Ruiz R, Ceden L, Fermin G, Andreu A, Acuna I, Oliva R (2013) *Phytophthora infestans* and *Phytophthora andina* on Solanaceous hosts in South America. In: Lamour K (ed) Phytophthora: a global perspective. CABI International, Oxfordshire, pp 48–58

Föster H, Coffey MD, Elwood H, Sogin ML (1990) Sequence analysis of the small-subunit ribosomal rRNA's of three zoosporic fungi and implications for fungal evolution. Mycologia 82:306–312

Fry WE, McGrath MT, Seaman A, Zitter TA, McLeod A, Danies G, Small IM, Myers K, Everts K, Gevens AJ, Gugino B, Johnson SB, Judelson H, Ristaino JB, Roberts PB, Secor G, Seebold KW, Snover-Clift KL, Wyenandt A, Grunwald NJ, Smart CD (2013) The 2009 late blight Pandemic in Eastern USA-causes and results. Plant Dis 96:296–306

Galindo J, Hohl HR (1985) *Phytophthora mirabilis*, a new species of *Phytophthora*. Sydowia 38:87–96

Garber RC, Yoder OC (1983) Isolation of DNA from filamentous fungi and separation into nuclear, mitochondrial, ribosomal, and plasmid components. Anal Biochem 135:416–422

Gavino PD, Fry WE (2002) Diversity in and evidence for selection of the mitochondrial genome of *Phytophthora infestans*. Mycologia 94:781–793

Gómez-Alpizar L, Carbone I, Ristaino JB (2007) An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene geneologies. Proc Natl Acad Sci 104:3306–3311

Gómez-Alpizar L, Hu CH, Oliva RF, Forbes GA, Ristaino JB (2008) Phylogenetic relationships of *Phytophthora andina*, a new species from the highlands of Ecuador that is closely related to the Irish potato famine pathogen *Phytophthora infestans*. Mycologia 100:590–602

Goodwin SB, Fry WE (1994) Genetic analysis of interspecific hybrids between *Phytophthora infestans* and *Phytophthora mirabilis*. Exp Mycol 18:20–32

Goodwin SB, Cohen BA, Fry WE (1994) Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. Proc Natl Acad Sci 91:11591–11595

Goodwin SB, Legard DE, Smart CD, Levy M, Fry WE (1999) Gene flow analysis of molecular markers confirms that *Phytophthora mirabilis* and *P. infestans* are separate species. Mycologia 91:796–810

Goss EM, Cardenas ME, Myers K, Forbes G, Fry WE, Restrepo S, Grünwald N (2011) The plant pathogen *Phytophthora andina* emerged via hybridization of an unknown *Phytophthora* species and the Irish potato famine pathogen *P. infestans*. PLos One 6:e24543

Goss EM, Tabima JF, Cooke DE, Restrepo S, Fry WE, Forbes GA, Fieland V, Cardenas M, Grünwald NJ (2014) The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. Proc Natl Acad Sci 24:8791–8796

Griffith GW, Shaw DS (1998) Polymorphisms in *Phytophthora infestans*: four mitochondrial haplotypes are detected after PCR amplification of DNA from pure cultures or from host lesions. Appl Envir Microbiol 64:4007–4014

Griffiths RC, Tavaré S (1994) Ancestral inference in population genetics. Stat Sci 9:307–319

Grünwald NJ, Flier WG (2005) The biology of *Phytophthora infestans* at its center of origin. Ann Rev Phytopathol 43:171–190

Gunderson JH, Elwood H, Ingold A, Kindle K, Soin ML (1987) Phylogenetic relationships between chlorophytes, chrysophytes, and oomycetes. Proc Natl Acad Sci 84:5823–5827

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser. 41:95–98

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Ecol 22:160–174

Hawkes JG (1990) The potato: evolution, biodiversity and genetics resources. Belhaven Press, London

Hu CH, Perez FG, Donahoo R, McLeod A, Myers K, Ivors K, Secor G, Roberts PD, Fry WE, Deahl KL, Ristaino JB (2012) Recent genotypes of *Phytophthora infestans* in eastern USA reveal clonal populations and reappearance of mefenoxam sensitivity. Plant Dis 96:1323–1330

Klimczak LJ, Prell HH (1984) Isolation and characterization of mitochondrial DNA of the oomycete fungus *Phytophthora infestans*. Curr Genet 8:323–326

Kroon LPNM, Bakker FT, van den Bosch BGM, Bonants PJM, Flier WG (2004) Phylogenetic analysis of *Phytophthora* species based on mitochondrial and nuclear DNA sequences. Fungal Genet Biol 41:766–782

Lang BF, Forget L (1993) The mitochondrial genome of *Phytophthora infestans*. In: O'Brien SJ (ed) Genetic maps: locus maps of complex genomes. Cold Spring Harbor Laboratory Press, Plan View, pp 3.133–3.135

Lassiter ES, Russ C, Nusbaum C, Zeng Q, Hu CH, Thorne JL, Ristaino JB (2010) Inferring evolutionary relationships of species in the *Phytophthora* 1c clade using nuclear and mitochondrial genes. Phytopathology 100:S68

Lennon NJ et al (2010) A scalable, fully automated process for construction of sequence-ready barcoded libraries for 454. Genome Biol 11:R15

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics 25:2078–2079

Lindgreen S (2012) AdapterRemoval: easy cleaning of next-generation sequencing reads. BMC Res Notes 5:337

Martin MD et al (2013) Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. Nat Commun 4:2172. doi:10.1038/ncomms3172

Martin MD, Ho SY, Wales N, Ristaino JB, Gilbert MTP (2014) Persistence of the mitochondrial lineage responsible for the Irish potato famine in extant New World *Phytophthora infestans*. Mol Biol Evol 31:1416–1420

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303

Oliva RF, Kroon LPNM, Chacón G, Flier WG, Ristaino JB, Forbes GA (2010) *Phytophthora andina* sp. nov., a newly identified heterothallic pathogen of solanaceous hosts in the Andean highlands. Plant Path 100:590–602

Ordonez ME, Hohl HR, Velasco JA, Ramon MP, Oyarzun PJ, Smart CD, Fry WE, Forbes GA, Erselius LJ (2000) A novel population of *Phytophthora*, similar to *Phytophthora infestans*, attacks wild Solanum species in Ecuador. Ecol Pop Biol 90:197–202

Paquin B, Laforest M-J, Forget L, Roewer I, Wang Z, Longcore J, Lang BF (1997) The fungal mitochondrial genome project:evolution of fungal mitochondrial genomes and their gene expression. Curr Genet 31:380–395

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

Raffaele S et al (2010) Genome evolution following host jumps in the Irish potato famine lineage. Science 330:1540–1543

Rambaut A (2007) FigTree. Mol Evol Phylogenet Epidemiol. http://www.tree.bio.ed.ac.uk/software/figtree/

Ristaino JB, Groves CT, Parra GR (2001) PCR amplification of the Irish potato famine pathogen from historic specimens. Nature 411:695–697

Robideau GP, de Cock AWAM, Coffey MD, Voglmayr H, Bonants PJM, Ristaino JB, Chitty D, Rintoul T, Désaulniers N, Eggertson Q, Bala K, Gachon CMM, Smith ML, Lévesque A (2011) DNA barcoding of oomycetes with cytochrome c oxidase subunit I (COI). Mol Ecol Res 11:1002–1011

Thaxter R (1889) A new American Phytophthora. Bot Gaz 14:273–274

Watterson GA (1975) On the number of segregating sites in genetic models without recombination. Theor Pop Biol 7:256–276

Yoshida K, Schuenemann VJ, Cano LM, Pais M, Mishra B, Sharma R, Lanz C, Martin FN, Kamoun S, Krause J, Thines M, Weigel D, Burbano HA (2013) The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. eLife 2:e00731