

# SCIENTIFIC REPORTS



OPEN

## Large sub-clonal variation in *Phytophthora infestans* from recent severe late blight epidemics in India

Tanmoy Dey<sup>1</sup>, Amanda Saville<sup>2</sup>, Kevin Myers<sup>3</sup>, Susanta Tewari<sup>4</sup>, David E. L. Cooke<sup>5</sup>, Sucheta Tripathy<sup>6,7</sup>, William E. Fry<sup>3</sup>, Jean B. Ristaino<sup>2</sup> & Sanjoy Guha Roy<sup>1</sup>

The population structure of the *Phytophthora infestans* populations that caused the recent 2013–14 late blight epidemic in eastern India (EI) and northeastern India (NEI) was examined. The data provide new baseline information for populations of *P. infestans* in India. A migrant European 13\_A2 genotype was responsible for the 2013–14 epidemic, replacing the existing populations. Mutations have generated substantial sub-clonal variation with 24 multi-locus genotypes (MLGs) found, of which 19 were unique variants not yet reported elsewhere globally. Samples from West Bengal were the most diverse and grouped alongside MLGs found in Europe, the UK and from neighbouring Bangladesh but were not linked directly to most samples from south India. The pathogen population was broadly more aggressive on potato than on tomato and resistant to the fungicide metalaxyl. Pathogen population diversity was higher in regions around the international borders with Bangladesh and Nepal. Overall, the multiple shared MLGs suggested genetic contributions from UK and Europe in addition to a sub-structure based on the geographical location within India. Our data indicate the need for improved phytosanitary procedures and continuous surveillance to prevent the further introduction of aggressive lineages of *P. infestans* into the country.

The late blight pathogen *Phytophthora infestans* (Mont.) de Bary has had a major impact on both science<sup>1,2</sup> and society<sup>3,4</sup> and continues to do so even 170 years after the first reported outbreaks. Its aggressiveness, co-evolution with its host and a high mutation rate make *P. infestans* the most important among oomycete pathogens<sup>1</sup>, and a persistent threat to potato production. Potatoes are the third most important staple crop behind wheat and rice, and the pathogen represents a significant threat to food security<sup>5,6</sup>. Compared to other stramenopiles, *P. infestans* has a large genome (240 Mb) with high levels of repeat rich sequences<sup>7</sup>. The origin of this 'economically important pathogen'<sup>4</sup> has been debated with work suggesting either an Andean origin in South America<sup>2,8–10</sup> or a central Mexican<sup>11–14</sup> origin. The pathogen caused widespread disease first in the US and then migrated to Europe and other parts of the world and has historically been disseminated as a series of clonal lineages. The lineage containing the HERB-1 mtDNA haplotype and the FAM-1 nuclear genotype initiated the Irish famine in 1840s<sup>9,10,15,16</sup> and was dominant and widespread in various parts of the world during the 19<sup>th</sup> century<sup>16</sup>. This lineage was later replaced by the US-1 genotype<sup>15–17</sup> which dominated until the 1980s<sup>18</sup> and was, in turn, displaced by other aggressive clonal lineages worldwide<sup>19,20</sup>.

The population structure of *P. infestans* has continued this rapid change over the past 40 years but variation in its reproductive biology and thus its genetic diversity is observed globally. The pathogen is heterothallic with two mating types, A1 and A2, and the presence of both types in central Mexico and in the Nordic countries of Europe and the Netherlands has led to sexual reproduction and high genetic diversity<sup>13,21,22</sup>. In Western Europe however, the *P. infestans* population is dominated by aggressive clonal genotypes including 13\_A2 (also known as Blue\_13) and 6\_A1<sup>23,24</sup>. The highly aggressive and dominant European genotype 13\_A2 was first recorded in 2004 in the Netherlands and Germany<sup>23</sup> and has subsequently been found in China<sup>25</sup> and India<sup>26</sup>. The spread of new aggressive lineages has led to an increase in late blight incidence in India and represents a current threat

<sup>1</sup>Department of Botany, West Bengal State University, Kolkata, 700126, India. <sup>2</sup>Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, North Carolina, Raleigh, NC, 27607, USA. <sup>3</sup>Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, NY14850, USA. <sup>4</sup>TeraBundle Analytics Pvt. Ltd, Haldia, 721602, India. <sup>5</sup>James Hutton Institute, Dundee, DD2 5DA, UK. <sup>6</sup>Indian Institute of Chemical Biology, Kolkata, 700032, India. <sup>7</sup>Academy of Scientific and Innovative Research, Ghaziabad, Delhi, India. Correspondence and requests for materials should be addressed to S.G.R. (email: [s\\_guharoy@yahoo.com](mailto:s_guharoy@yahoo.com))

to food security<sup>27</sup>. Many other late blight epidemics were also reported during 1990 to 2014 throughout the world. Among these outbreaks, the most recent and severe one in India occurred in the eastern and north-eastern regions, primarily in the state of West Bengal in 2014<sup>28</sup>.

India is the 2<sup>nd</sup> largest producer of potato with 80% of the total production in the northern and eastern parts of the country. Potato is a winter crop in most parts of India. Exceptions are in the northern and north-eastern hills, where potato is a summer crop, and the southern regions, where potato is an autumn crop. Late blight is a recurring problem in the northern hills every year. In the Indo-Gangetic plains, where the major growing regions of India are located, the disease is mild to sporadic each year. However, once every two to three years, the pathogen becomes epiphytic, causing up to 75% loss<sup>29</sup>. In the plains of eastern India, crop rotation is routine with potato grown between November and February in fields that are used for rice cultivation for the rest of the year. These host dynamics affect the pathogen population through repeated extinction and re-colonization events and this may affect gene flow as well as evolutionary trajectories. There are records of at least four migrations of *P. infestans* into India over the past 100 years. For example, the oldest samples of *P. infestans* from 1913 collected by J. F. Dastur in Bhagalpur (Bihar) were the Ia mitochondrial haplotype<sup>30</sup> and the US-1 clonal lineage (Ib haplotype) was present in India by the 1960s<sup>30</sup>. The presence of the A2 mating type in 1990s in northern hills provided additional evidence suggesting migration from an outside source<sup>31</sup> and more recently, the European 13\_A2 genotype was detected in southern India<sup>26</sup> and dominant on both potato and tomato between 2010 and 2012<sup>32</sup>. However, no clear population structures have been characterized in subtropical Indo-Gangetic regions and northeastern parts of India, which share many international borders with Nepal and Bangladesh. West Bengal, Assam, and Meghalaya are adjacent to Bangladesh where late blight is reported on potatoes cultivated under similar cropping regimes and climate. As the pathogen does not respect international boundaries, migration between regions is highly likely.

In eastern India the four major potato producing states are Assam, Bihar, Meghalaya, and West Bengal. These regions comprise approximately 45% of the total potato production in the country<sup>33</sup>. West Bengal is the second largest potato producing state after Uttar Pradesh. In 2014, a late blight epidemic in this region, primarily in West Bengal, led to dramatic social upheaval, including farmers committing suicide due to crop losses and policy changes such as setting minimum export values on potatoes<sup>28</sup>. During the epidemic year, the potato yield was approximately 8000 kg/ hectare less than what is expected for an average production year. A close examination of the pathogen populations during recent epidemics could provide insight into the source of *P. infestans* populations in India and spread within the country and help guide Indian growers to more effective disease management practices.

The objectives of this study were to (i) analyse the phenotypic and genotypic variation of the *P. infestans* population causing late blight in different regions of India (ii) develop a baseline of the current *P. infestans* genotypes in major potato growing regions for future studies and (iii) examine possible migration pathways for the pathogen in eastern and northeastern regions of India.

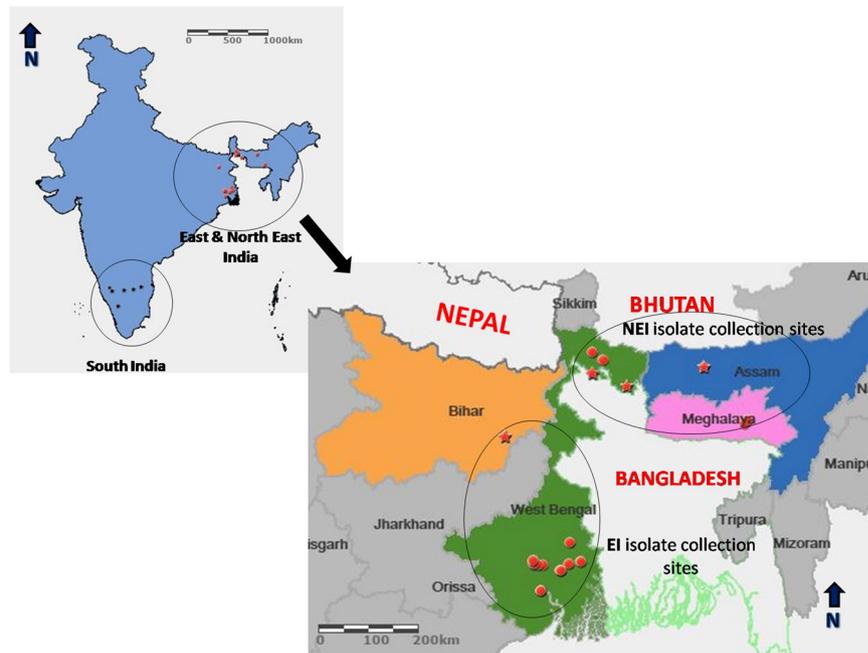
## Results

**Phenotypic and genotypic diversity.** A total of 59 isolates consisting of 19 from tomato and 40 from potato were collected between 2013–2014 growing season (Severe Late blight epidemic year) from the major potato and tomato growing regions (Fig. 1) (see Supplementary Table S1). All were of the A2 mating type and the Ia mtDNA haplotype. In addition, 57 of 59 isolates were resistant to metalaxyl, one was intermediately resistant, and one was sensitive (see Supplementary Table S1).

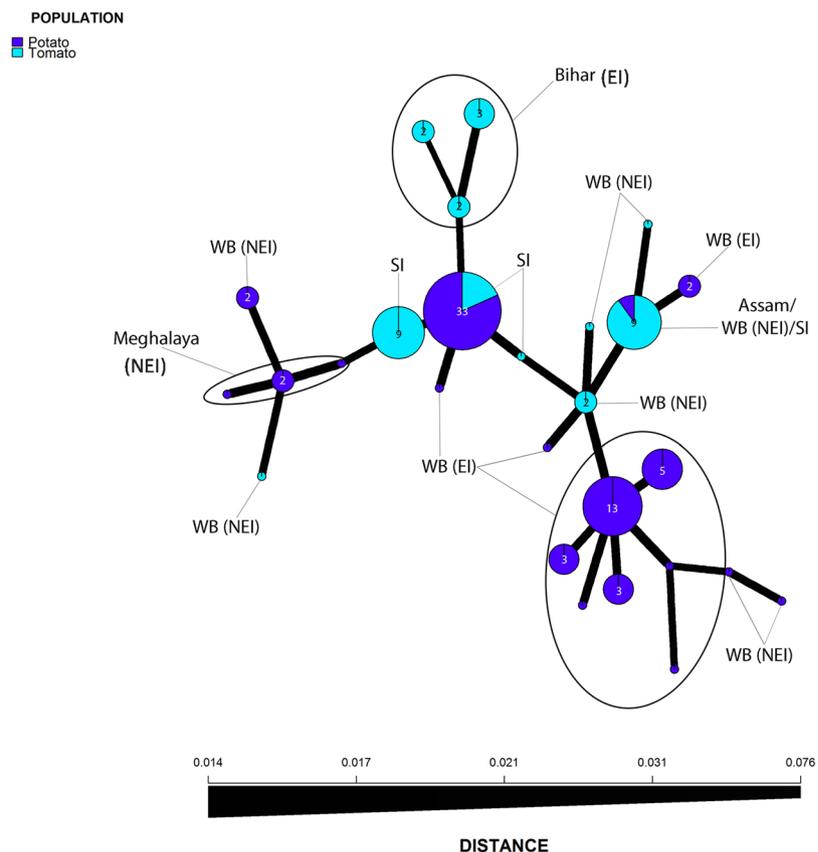
All the isolates were identified as the 13\_A2 genotype on the basis of the RG57 probe RFLP fingerprint, the Ia mtDNA haplotype and the 12-plex SSR data<sup>23,34</sup> (see Supplementary Table S2, see Supplementary Table S3). The SSRs also revealed mutations generating a third allele at the three most variable loci (D13, G11 and SSR4), allowing isolate discrimination and suggesting the clone was triploid.

**Population differentiation.** Examination of the population structure of *P. infestans* from Eastern India (EI) and North Eastern India (NEI) using SSRs identified a total of 24 multilocus genotypes (MLGs) amongst the 59 isolates. When another 45 isolates from south India (SI)<sup>32</sup> were incorporated, a total 27 MLGs were identified amongst 104 isolates from India. No single MLG dominated in the EI and NEI population; the three most frequently sampled MLGs were found 13, 9, and 5 times (Fig. 2). Little difference was observed between the diversity of MLGs measured in EI and NEI populations (Table 1). Multilocus genotypes in both regions showed the same level of evenness (E5) and clonality based on Ia and  $\bar{F}_d$ . The Ia and  $\bar{F}_d$  values were found to be significantly different from 0 based on 999 permutations, providing evidence in support of the presence of a clonal population (see Supplementary Table S4 and Supplementary Fig. S5). Due to its geography samples from the state of West Bengal (Fig. 1) were ascribed to Eastern and Northeastern subpopulations termed, WB-EI and WB-NEI, respectively. The WB-EI and WB-NEI samples were found to be the most diverse, with the greatest number of MLGs and a higher diversity index than other regions. Isolates collected from Assam were found to be the least diverse, with only one MLG observed (Table 1). A total of 38 alleles were identified across 12 loci with a mean of 3.17 (Table 2). Among these, the D13 locus had the most alleles, followed by PinfSSR4, PiG11, and Pi4B. Loci Pi70, PinfSSR11, and PinfSSR2 were the least diverse with 1 allele each. Of loci with more than two alleles, Pi4B was found to be the most evenly distributed (Table 2). In EI and NEI region isolates collected from potato were more diverse than isolates collected from tomato (Table 3). However, this finding may be due to the greater number of isolates collected from potato. When south Indian isolates<sup>32</sup> were incorporated into the analysis, almost equal diversity were identified on both hosts in India (see Supplementary Table S6).

The minimum spanning network (MSN) shows that no MLGs were shared between the eastern and northeastern regions, though MLGs from both regions were interspersed throughout the network (Fig. 2). Examination of the MSN based on host indicates that all except two of the ten MLGs from tomato were host-specific. One of the



**Figure 1.** Geographical locations for sample collections and populations corresponding to the potato and tomato growing regions. Eastern India (EI), northeastern India (NEI) and southern India (SI) populations have been depicted. EI & NEI were sampled in this study. SI was collected previously<sup>32</sup>, was used in data analysis. Only major sites have been shown; stars depict Tomato and dots Potato hosts. Map was constructed by using ArcGIS online<sup>78</sup> (<http://www.esri.com/software/arcgis/arcgisonline>).



**Figure 2.** Minimum spanning network (MSN) of multilocus genotypes of *Phytophthora infestans* from populations in India based on host. The size of the node is proportional to the number of individuals within the multilocus genotype.

Population	N <sup>a</sup>	MLG <sup>b</sup>	eMLG(SE) <sup>c</sup>	H <sup>d</sup>	Hexp <sup>e</sup>	Evenness	Ia <sup>f</sup>	$\bar{r}_d^g$
<b>Region</b>								
Eastern India	38	13	9.93(1.18)	2.18	0.426	0.647	0.997	0.238
Northeastern India	21	11	11(0)	2.05	0.432	0.634	0.947	0.236
All	59	24	12.90(1.61)	2.79	0.428	0.647	0.770	0.152
<b>State</b>								
Eastern West Bengal	31	10	5.45(1.092)	1.84	0.423	0.631	0.0967	0.030
Northeastern West Bengal	12	8	7.14(0.649)	1.98	0.445	0.889	0.9001	0.210
Assam	5	1	1(N/A)	0.00	0.417	N/A	N/A	N/A
Bihar	7	3	3(N/A)	1.08	0.409	0.969	0.7439	0.372
Meghalaya	4	3	3(N/A)	1.04	0.468	0.912	-0.333	-0.333
South India	45	4	2.54(0.672)	0.772	0.388	0.621	-0.0728	-0.0834
All	104	27	6.55(1.294)	2.55	0.42	0.515	0.9012	0.1944

**Table 1.** Diversity statistics for microsatellite data for 12 microsatellite loci in populations of *Phytophthora infestans* in India by region and by state. <sup>a</sup>*n*: number of individuals. <sup>b</sup>MLG: number of multilocus genotypes (MLG). <sup>c</sup>eMLG: expected number of MLGs at smallest size of at least ten; SE: Standard error. <sup>d</sup>H: Shannon-Weiner Index of MLG diversity. <sup>e</sup>Hexp: Nei's 1978 expected heterozygosity. <sup>f</sup>Ia: Index of Association. <sup>g</sup> $\bar{r}_d$ : standardized index of association.

Locus	<i>n</i> <sup>a</sup>	1-D <sup>b</sup>	Evenness
Pi02	2	0.50	1
Pi4B	3	0.51	0.95
PiG11	6	0.63	0.78
Pi04	2	0.50	1
Pi63	2	0.50	1
Pi70	1	N/A	N/A
D13	9	0.72	0.67
PinfSSR11	1	N/A	N/A
PinfSSR2	1	N/A	N/A
PinfSSR4	7	0.66	0.80
PinfSSR6	2	0.50	1
PinfSSR8	2	0.50	1
Mean	3.17	0.42	0.91

**Table 2.** Population statistics for clone corrected microsatellite data for 12 microsatellite loci in populations of *Phytophthora infestans* in India. <sup>a</sup>*n*: number of individuals (not clone corrected). <sup>b</sup>1-D: Simpson index.

Population	N <sup>a</sup>	MLG <sup>b</sup>	eMLG(SE) <sup>c</sup>	H <sup>d</sup>	Hexp <sup>e</sup>	Evenness	Ia <sup>f</sup>	$\bar{r}_d^g$
<b>Original</b>								
Potato	40	17	10.6(1.41)	2.39	0.434	0.592	0.748	0.224
Tomato	19	8	8(0)	1.84	0.411	0.749	0.922	0.183
All	59	24	12(1.57)	2.79	0.428	0.647	0.770	0.152

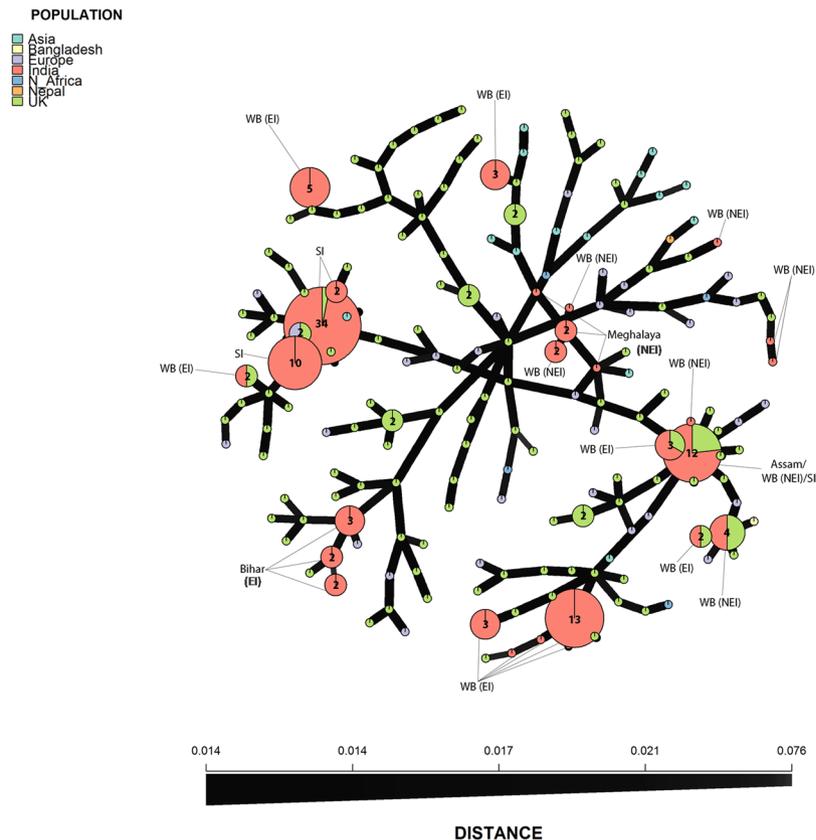
**Table 3.** Diversity statistics for all 12 loci in Eastern and North Eastern Indian populations of *Phytophthora infestans* based on host. <sup>a</sup>*n*: number of individuals (not clone corrected). <sup>b</sup>MLG: number of multilocus genotypes (MLG). <sup>c</sup>eMLG: expected number of MLGs at smallest size of at least ten; SE: Standard error; <sup>d</sup>H: Shannon-Weiner Index of MLG diversity. <sup>e</sup>Hexp: Nei's 1978 expected heterozygosity. <sup>f</sup>Ia: Index of Association. <sup>g</sup> $\bar{r}_d$ : standardized index of association.

exceptions was from the NEI region where there was a single isolate from potato in the same MLG as eight isolates from tomato probably because the potato isolate was from a field adjacent to the infected tomato crop. The other, was a MLG composed of only south Indian isolates<sup>32</sup>.

In addition, one of the four MLGs detected in south India matched the MLGs of *P. infestans* from disease outbreaks sampled in eastern and northeast India in the current study (Fig. 2). An examination of MLG relatedness by region indicated a greater range of MLG diversity in West Bengal (both WB-EI and WB-NEI) than in Bihar and Meghalaya. However, as the administrative region of West Bengal (WB) covers a large geographic distance (Fig. 1). MLGs from WB-NEI and WB-EI were interspersed throughout the network. The cluster of three related

Region	Assam	Eastern West Bengal	Northeastern West Bengal	Meghalaya	Bihar	South India
Assam		0.00673	0.00767	0.03680	0.02819	0.00903
Eastern West Bengal	24.607211		0.00759	0.01327	0.02300	0.02973
Northeastern West Bengal	21.55849	21.80198		0.01012	0.02262	0.01396
Meghalaya	4.36286	12.39732	16.30388		0.03537	0.01065
Bihar	5.74591	7.77150	7.20171	4.54520		0.01321
South India	18.28453	5.43934	11.77300	15.48405	12.44742	

**Table 4.** Pairwise calculations of  $F_{ST}$  (top diagonal) and  $N_m$  (bottom diagonal) for populations of *Phytophthora infestans* in multiple regions of India.

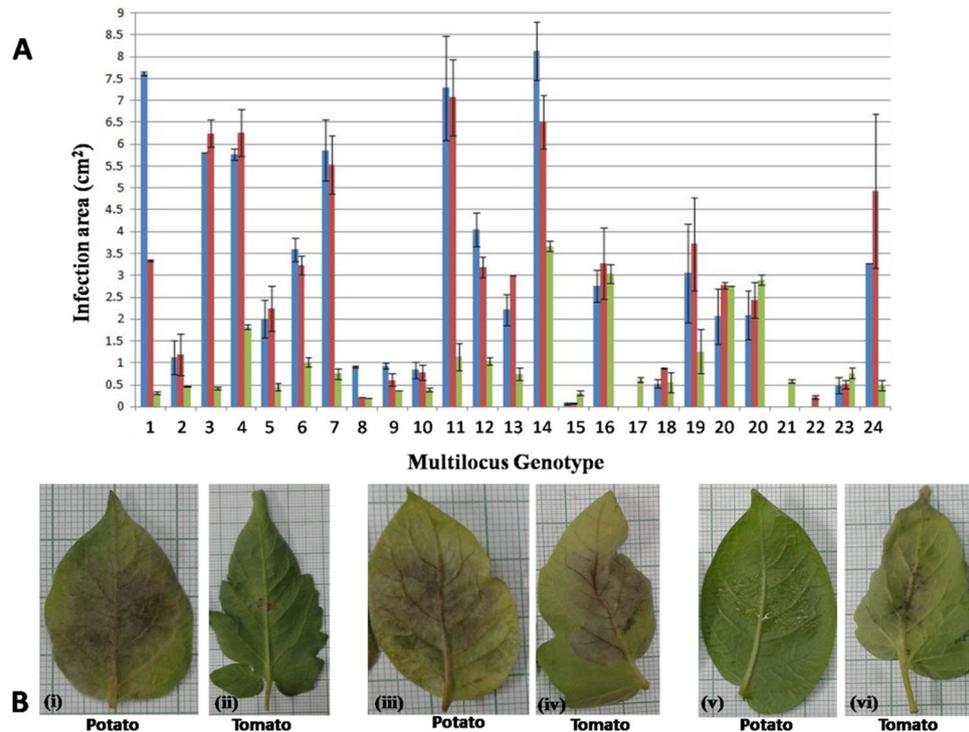


**Figure 3.** Minimum Spanning Network (MSN) of *Phytophthora infestans* 13\_A2 MLGs from India compared to representative MLGs found amongst a global sample of isolates.

MLGs each from Bihar and Meghalaya were distinct from other MLGs found in West Bengal. The five isolates from Assam, mostly from tomato, formed a single cluster along with three isolates from the neighbouring northern part of West Bengal from the northeastern population.

The lowest values of  $F_{ST}$  calculated were between the WB-EI and Assam populations ( $F_{ST} = 0.00673$ ) and between the WB-NEI and Assam populations ( $F_{ST} = 0.00767$ ). Conversely, the lowest level of gene flow was observed between Meghalaya and Assam ( $F_{ST} = 0.03680$ ). These values are reflected in the number of migrants calculated for each population pair, with Assam and WB-EI/WB-NEI exhibiting the most number of migrants per generation ( $N_m = 24.607211$  and  $21.55849$ , respectively) and Assam and Meghalaya exhibiting the fewest ( $N_m = 4.36286$ ) (Table 4).

Nineteen of the twenty four 13\_A2 variants identified in the sample from eastern and north eastern India were unique to this region (see Supplementary Table S3) when compared to a global sample of 172 13\_A2 variants (see Supplementary Table S7). The structure analysis did not reveal any evidence for more than one population. The MSN of Indian and global populations indicated that MLGs of Indian *P. infestans* populations were dispersed across the network rather than forming a distinct single cluster (Fig. 3). Samples from West Bengal were the most diverse and grouped with MLGs found in Europe, the UK and from neighbouring Bangladesh but were not linked directly to most samples from south India. A large MLG (MLG20) comprising isolates from Assam, West Bengal



**Figure 4.** (A) The mean infection area of the 24 MLGs of 13\_A2 lineage of *Phytophthora infestans* from eastern and northeastern India on two cultivars of potato ('Kufri pukhraj' & 'Kufri jyoti') and one cultivar of tomato ('Arka vikas'); shown as the blue, red and green bars, respectively. MLG20 had isolates from both potato and tomato and accordingly the MLG20 has been depicted twice here separately. (B) The corresponding pictures of the infection on potato and tomato in detached leaf assay; (i) and (ii) representative of MLGs more aggressive on potato, (iii) and (iv) equally aggressive on both hosts, (v) and (vi) no infection on potato.

and south India was shared with isolates previously found in the UK. Other samples tended to form discrete clusters such as those collected in Bihar, South India and Meghalaya. The two former regions had MLGs most similar to samples from Europe whereas the latter had closer genetic similarities with MLGs from Europe and Asia.

**Population virulence.** Differences in the aggressiveness phenotype, based on lesion area, incubation period, and latent period, on potato and tomato hosts were noted amongst the tested isolates of each sub-clonal variant of the 13\_A2 clonal lineage (Fig. 4). However, there was overall greater aggressiveness on both cultivars of potato than on tomato ( $P$  value < 0.01). The only exceptions were MLG16 and MLG20, which had almost equal aggressiveness on both hosts. No significant differences in aggressiveness on either host were observed based on the host of origin, but two MLGs (MLG17 and MLG21) from tomato, failed to infect potato, and one MLG from potato (MLG 22) did not infect tomato or the other 'Kufri Pukhraj' variety of potato. Differences in incubation period were also observed. MLGs 4, 5, and 6 had a two-day incubation period, while the rest had a three-day incubation period. All MLGs had a five-day latent period. However, the latent period for MLG 8 (tomato only), MLG 15 (potato only), and MLG 22 (Kufri Jyoti only) could not be determined, as no sporangia were seen after seven days of incubation (see Supplementary Table S8 and Fig. 4).

Analysis of *Avr3a* sequences identified the presence of the *Avr3a* EM type gene in EI and NEI populations. In addition, one synonymous (at base pair 363) and two nonsynonymous (at base pairs 43 and 370) mutations were also identified in the *Avr3a* EM type gene from fourteen isolates independently of their host of origin (see Supplementary Table S9). The polymorphism identified at position 370 leads to a change in the amino acid from arginine to glycine at position 124, located in the loop 3 region. Mutational analysis based on the model structure of *Avr3a* wild type (WT) predicted that the mutation with the highest likelihood to affect function at position 124 was glycine (see Supplementary Fig. S10).

## Discussion

The eastern region of India is the country's largest potato growing region and was the site of a severe late blight epidemic in 2014<sup>28</sup>. DNA fingerprinting of 59 isolates of *P. infestans* identified 13\_A2 (Blue\_13), a European genotype, as the only clonal lineage causing the outbreaks. Sub-clonal variation in SSR fingerprints revealed some sub-structure in the local population that was linked to region and host. Most isolates from eastern and north-eastern India differed from global populations and those previously observed in southern India. Rapid changes in population structure, in which dominant lineages are replaced by newer, fitter, more aggressive lineages has been observed previously in *P. infestans*<sup>16,35</sup> and the 13\_A2 lineage is a modern example of this pattern. Since its first isolation in Germany and the Netherlands in 2004, the lineage has successfully replaced other

genotypes and become dominant in parts of western Europe<sup>23,36</sup>. This lineage has subsequently spread to other parts of the world, including China and India<sup>25,26</sup>. Recent work from southern India reported that the importation of 13\_A2 most likely on seed potatoes from the NL or the UK was responsible for severe outbreaks of late blight on tomato and potato<sup>26,32</sup>. Four sub-clonal MLGs were identified amongst 45 isolates in the south. In contrast, the populations from north-eastern and eastern India examined in our study were more diverse with 24 MLGs identified from both hosts, 19 of which were unique to the region. This means that multiple mechanisms were probably at play; not only new introductions of variants from the UK or /and Europe may have occurred, but also local mutations and adaptations took place as discussed in detail later on.

Recent work has confirmed that the 13\_A2 lineage is triploid<sup>37</sup> which is consistent with the observed high levels of heterozygosity and SSR polymorphism<sup>23,25,36</sup>. In this study, some loci such as D13, PiG11 and PinfSSR4 were highly diverse and others such as PinfSSR11, Pi70 and PinfSSR2 were monomorphic and homozygous. The eastern and north-eastern Indian populations examined here showed high levels of diversity with novel mutations, in particular at the loci D13, PinfSSR4 and PiG11, that generated 19 MLGs not reported previously (Fig. 3). Such mutations are helpful in tracking populations of *P. infestans* on local and global scales.

Evidence of phenotypic changes in fungicide sensitivity have also been documented in earlier studies which showed that in the plains of northern India, the A1 mating type and metalaxyl sensitive isolates, presumably US-1, were prevalent, but later the population shifted to a metalaxyl resistant populations even though the A1 mating type was still observed<sup>38–40</sup>. Though no evidence of sexual reproduction has been recorded in the eastern and northeastern parts of India, the A2 mating type was observed in Meghalaya<sup>38</sup>. All our isolates were found to be of the A2 mating type and all except two tomato isolates were metalaxyl resistant. Mitochondrial haplotyping indicated that all isolates were the Ia mt DNA haplotype. Previous reports indicated that the A1 isolates were either the Ia or Ib haplotype<sup>38,41,42</sup>. This study thus confirms the presence of the A2 mating type in this region and the absence of the A1 mating type suggests the new A2 population is aggressive and has displaced the former population. Such a displacement by 13\_A2 is consistent with studies in Europe<sup>23</sup>. This aggressive and metalaxyl resistant pathogen population will make late blight more challenging to manage and it is important that local management decisions for growers should reflect this in these regions.

Aggressiveness and host preferences are key factors that affect disease outbreaks. For example, US lineages US-8 and US-24 predominantly infect potato, while US-22, the primary lineage during a major outbreak in the US in 2009, infects both potato and tomato<sup>6,43</sup>. Such host adaptation either on potato or tomato has also been reported elsewhere<sup>19,44,45</sup>. The 13\_A2 lineages in southern India were equally aggressive on both hosts<sup>32</sup>. Our results show some variation in aggressiveness within the 13\_A2 clonal lineages. Detached leaf assays showed the pathogen to be more aggressive on potato than tomato irrespective of host of origin. However, we identified two MLGs from tomato that did not infect potato and one MLG from potato that did not infect tomato or the 'Kufri Pukhraj' cultivar of potato (Fig. 4). Such differences in host specificity may drive the observed subdivision of the clonal population by host (Fig. 2). Only one of the eight isolates of MLG20, for example, was from potato and the potato field where this isolate was collected was adjacent to tomato fields where other isolates of this MLG were collected. Similarly the distinct sub-clones in Bihar all originated from tomato. These findings provide some evidence of hitherto unseen sub-clonal variation based on host within the 13\_A2 clonal lineage. Other studies on Indian *P. infestans* isolates have shown there is no difference in aggressiveness and host specificity between tomato and potato isolates<sup>32,40,46</sup>. Further work needs to be done to explore this question.

The presence of the EM form of the AVR3a effector is consistent with its reported virulence against the R3a gene<sup>23</sup>. Additionally, the detection of two novel non-synonymous mutations at positions 15 (signal peptide) and 124 (loop 3) within the *avr3a* locus suggest these populations may be unique among current known 13\_A2 populations. Moreover, since it is possible that the mutation identified at position 124 may affect the function of this protein, further characterization of these isolates is needed.

Many factors regulate pathogen population dynamics including environmental forces, selection pressure, new aggressive strains, increasing global trade, and host dynamics<sup>20,47</sup>. Eastern and north eastern India are areas where movement of host material affects pathogen populations and changes in pathogen population dynamics can lead to severe disease outbreaks. The annual mean temperature in this sub-tropical region remains around 30 °C and during the summer temperatures can reach up to 42 °C. Potato cultivation occurs mainly in paddy fields that are waterlogged for most of the year. During the monsoon season, flooding is also a recurring problem in major potato growing regions of West Bengal. Since the pathogen population is reproducing asexually, oospores are not present for overwintering. In addition, sporangia cannot survive beyond 20 days in waterlogged conditions<sup>48</sup>. Due to these adverse environmental factors each year and the practice of rotating rice cultivation (where the field is also submerged) there is an almost complete elimination of both host and pathogen which means populations of *P. infestans* face a narrow genetic bottleneck. The pathogen survives only in infected stored seeds kept by small shareholders and use of infected seed impacts pathogen dispersal between cropping seasons. This asexual vertical transmission between seasons is partially facilitated by growers planting moderately resistant tubers<sup>47</sup> such as 'Kufri Jyoti' and 'Kufri Pukhraj' that are mainly cultivated in these regions. Since the tubers are not highly resistant to infection, less aggressive isolates with longer latent periods (e.g. those in our study with latent periods >7 days) are more likely to survive to spread, as they produce less severe symptoms. Thus, seed planted during the next growing season can survive<sup>47</sup> and act as a source of inoculum. In this region, new founder populations emerge each year, persist for some weeks, and then disappear. This pattern is made possible by the fact that potato is mainly a seasonal winter crop across most of this region (except for hill regions like Darjeeling and Shillong). There is also a difference in the way the host material (i.e. seed potato) is used in these regions. Seed in West Bengal comes almost every two to three years mostly from northern states (like Punjab). In contrast, farmers in Meghalaya tend to use potato seed which they have stored from past years<sup>49</sup>. These differential cultural practices, environmental factors, crop rotation and grower behaviour can compound to influence pathogen diversity and prevalence<sup>50</sup> and might explain the differences in observed genetic diversity of *P. infestans* in these regions.

Our data show that while the Indian late blight populations are clonal, there is more subclonal diversity within the 13\_A2 lineage than previously reported. Since north eastern and eastern India shares international borders with Bangladesh and Nepal, both of which regularly import potatoes from outside sources, and because of the practice among small marginal farmers of storing seed to use for the next season, the diversity observed may be the result of both local migrations of the pathogen across borders and common local agricultural practices. This diversity may be exacerbated by a local weather phenomena called the 'Western Disturbance'. This occurs in northern India, causing meteorological changes in eastern India and generating 5–6 day periods of wet, cool and windy weather suitable for *P. infestans* infection. The subsequent rapid dispersal and mixture of the rapidly expanding clonal populations due to this wind is likely to increase the diversity of *P. infestans* in this region.

Stepwise mutation of SSR loci generates variants that can be used to track local and migrant populations providing inferences on the spread of *P. infestans* inoculum via seed trade or airborne sporangial routes. The population structure of *P. infestans* in India indicates that the 13\_A2 lineage migrated recently to India likely due to movement of infected plant material from the UK and Europe<sup>26</sup>. Three probable pathways for interstate migration occur in India; one includes movement of *P. infestans* from the South Indian population upwards towards Bihar, the second from the bordering regions into West Bengal, possibly from neighboring countries like Bangladesh and Nepal. A third possibility includes direct import of infected tubers from Europe. The MSN for the 13\_A2 population across the whole of India revealed 27 MLGs with some evidence of a regional sub-structure (Fig. 2). One isolate from southern India shared an MLG with those from the northeastern India (MLG20 from Assam and NEI-West Bengal) and these were all isolated from tomato<sup>32</sup> which could support this as a migrant population. However, comparisons with the larger collection of 13\_A2 variants (Fig. 3) showed this was also shared with the population in Europe and was an MLG defined early in the history of 13\_A2 in Europe<sup>23,36</sup>. This MLG is thus equally likely to have been imported into northeastern India independently of the migrant pathogen population in the south. The other three MLGs from southern India were distinct from those in the current study also supporting an independent emergence of the population in the northeastern regions of India rather than a migration from the south. Only isolates from Bihar displayed a distinct local cluster and this was supported in comparisons with the wider sample of variants from global populations (Fig. 3). A comparison of the 13\_A2 variants identified in a cross section of global populations with those from north eastern and eastern India is informative. The presence of Indian MLGs across the MSN suggests either multiple independent imports or the importation of highly diverse populations of *P. infestans* 13\_A2 inoculum into the region. Some local clustering of northeast and eastern Indian populations from Bihar and Meghalaya is apparent but the variants from both West Bengal-EI and WB-NEI are scattered across the network. Shared allelic diversity with African and Asian isolates were noted but further more detailed analysis of greater sample numbers from these countries is required to determine those pathways of spread more precisely. It should also be noted that associations between MLG and location may be the result of chance mutations independently generating the same MLG. The resultant homoplasy can confound such analysis of dispersal pathways.

The data presented here can serve as a baseline for further studies of the diversity of Indian populations of *P. infestans* and for surveillance programs for improved management. The continued monitoring of these populations will provide forewarning of any new population shifts and potentially provide additional time to react in the face of another epidemic. India has yet to begin contributions to pathogen surveillance systems, including the Asiabligh program<sup>51</sup>, which mirrors the objectives of the Euroblight program (<http://euroblight.net/>) and USAblight (<http://USAblight.org>). However, the results of our study can now contribute to forming baseline microsatellite data for an Indiabligh network or a larger Asiabligh-based monitoring system in India. Our data also point to the need for improved phytosanitary measures to curtail importation of infected seed potatoes into India. The history of seed trade between India and Europe suggests that improved seed certification may be needed to suppress the additional introduction of new strains of *P. infestans* into the country.

## Materials and Methods

***Phytophthora infestans* collection, isolation and storage.** A single symptomatic sporulating late blight lesion from foliage of potatoes, and leaflets and fruits of tomato, were collected from farmers' fields in the major potato and tomato growing regions in eastern (EI) and north eastern India (NEI) during 2013–14 growing season. Each sample was carefully collected, placed in a plastic zipper bag, marked, and transported to the laboratory within 24 hours. The details for each *P. infestans* isolate collection are indicated in Supplementary Table S1.

Collected samples were processed using a routine *P. infestans* isolation protocol<sup>52</sup>. Small freshly sporulating leaf pieces (5 mm<sup>2</sup>) were then placed on top or into a selective medium [pea agar amended with rifamycin (20 mg/L), vancomycin (50 mg/L), ampicillin (100 mg/L), Polymixin B (50 mg/L), pentachloronitrobenzene (50 mg/L), and carbendazim (100 mg/L)]<sup>53</sup>. The plates were incubated in the dark at 18 °C for 5–10 days. *P. infestans* colonies were selected from these plates and transferred onto a fresh pea agar plate without any antibiotics<sup>43</sup> for pure culture.

**Cryopreservation:** Agar plugs containing mycelium were suspended in 1 ml of cryoprotectant (10% glycerol) in screw-cap polypropylene vials, placed vials in 'Mr. Frosty' (Nalgene) tubes and then an uncontrolled cooling protocol<sup>54</sup> was followed.

**Extraction of total DNA.** Mycelial plugs from individual isolates were transferred to pea broth<sup>53</sup>. Mycelium from the 5 day old culture was harvested and excess medium was removed by filtration and washed in autoclaved distilled water. Approximately 0.1 g of mycelium was placed in a sterile 1.5 ml eppendorf tube and total DNA was extracted<sup>55</sup>.

**Determination of mating type.** Mating type was determined using molecular markers. The markers for A1 and A2 mating type are - A1- INF-1, INF-2<sup>56</sup>, A2 - PHYB-1,PHYB-2<sup>57</sup> primers were used. The validity of the marker method for testing mating type was tested using a set of control cultures of known mating type.

**Mitochondrial DNA haplotype and RG-57 fingerprinting.** Mitochondrial DNA (mt DNA) haplotypes were identified using the PCR-RFLP methods<sup>58</sup>. DNA fingerprinting was carried out using the RG57 multilocus nuclear DNA probe<sup>59</sup>. Total DNA from isolates along with a standard US-1 and 13\_A2 (2006\_3928A) isolate were used as standard.

**Multiplexed microsatellite marker analysis.** *P. infestans* simple sequence repeat (SSR) loci were genotyped using a modified version of the protocol for 12-plex single sequence repeat genotyping as described previously<sup>34</sup>. The *P. infestans* isolates were run alongside a standard 13\_A2 sample (2006\_3928A). The Qiagen Type-It Microsatellite PCR kit (Qiagen Corporation, Valencia CA) was used for PCR reactions, and sample volumes were modified to run a 12.5  $\mu$ L reaction by using 6.25  $\mu$ L  $2 \times$  Type-It Master Mix, 1.25  $\mu$ L of a  $10 \times$  multiplex primer master mix, 4  $\mu$ L PCR grade water, and 1–2  $\mu$ L of template DNA (5–10 ng). Thermal cycling conditions as described earlier<sup>60</sup>. Fragments were analyzed on an Applied Biosystems 3730xl DNA analyzer. The peak size was determined against a GeneScan 500 LIZ standard and alleles were scored manually using Peak Scanner 2 (Applied Biosystems, Foster City, CA), and fragment lengths were rounded to the nearest whole number for analysis.

**Metalaxyl sensitivity.** Sensitivity to metalaxyl was determined using an agar technique<sup>52</sup>. Pea agar plates amended with 5 and 100 ppm of metalaxyl (Glazer 35 WS<sup>®</sup>, Rallis India Ltd.) were prepared. Agar blocks (8 mm diameter) with actively growing mycelia were taken from the colony margin of each isolate and transferred to the centre of three replicate plates of metalaxyl-amended and non amended (served as control) pea agar plates. After 7 days of incubation at 18 °C in dark the metalaxyl sensitivity were checked as described earlier<sup>52</sup>. The metalaxyl sensitivity experiment was replicated twice for each isolate.

**Virulence and infection area determination.** A total of 25 isolates, one representative isolate from each MLG and two representative isolates from MLG 20 (one isolate from potato and one isolate from tomato), were selected for virulence and aggressiveness studies. Isolates were tested following the detached leaf assay on two cultivars of potato, Kufri Pukhraj and Kufri Jyoti, both classified as moderately resistant to late blight<sup>61,62</sup> and on the tomato cultivar Arka Vikas, which has no known resistance genes<sup>32</sup>. Tomato and potato plants were grown in a glasshouse in plastic pots. The plants were maintained at a mean daily temperature of 22–23 °C under a 16-h day photoperiod. Uniform leaflets of 6-week-old potato and tomato plants were harvested and placed, abaxial side up, on moist sterile filter paper in 90 mm sterile Petri dishes.

A sporangial suspension of a *P. infestans* isolate of each multi locus genotype (MLG) was prepared by scraping the surface of 15-day-old pea agar cultures in sterile distilled water. The concentration of the resulting suspension was determined with a haemocytometer and adjusted to  $1 \times 10^4$  sporangia  $\text{mL}^{-1}$ . The suspension was maintained at 4 °C for 2 h before inoculation. For inoculation, three leaflets were used per sample. Each leaflet was inoculated by placing one 20  $\mu$ L drop of a sporangial suspension on the abaxial surface near the midrib and incubated for 7 days at 18 °C with a light and dark cycle of 16 and 8 hours respectively as described earlier<sup>32,43,63,64</sup>. Plates were placed in a completely randomised design. The full experiment was repeated twice independently.

The virulence and aggressiveness of 25 isolates representing each MLG were evaluated daily after inoculation and incubation period (IP) and latent period (LP) were determined<sup>65</sup>. After 7 days of incubation, the infected leaflets were placed on 1  $\text{mm}^2$  graph paper and photographed. The lesion area was measured by using Image J software<sup>66</sup>. The measured lesion areas of both hosts (potato and tomato) were checked statistically to compare the aggressiveness of each MLGs and determine the significance of the effects of host of origin, target host and MLGs. For this, a non-parametric method of Wilcoxon rank sum test for comparing among groups were used<sup>67</sup>.

**Avr3a gene sequencing and annotation.** The Avr3a gene was amplified using the primers with a M13 tail to allow for amplification and sequencing of 453 bp corresponding to the entire gene<sup>68</sup>. Single band PCR products were further purified and sequenced for both strands by a commercial service (Xcelris labs Ltd, India). The sequences were then aligned by using tools available in Eumicrobe DB<sup>69</sup> and again compared manually with the supplied electropherogram data. A blast search was performed for each strand to compare the Avr3a sequences generated with those available in GenBank. Polymorphic positions were further analysed by using PHYRE2 investigator<sup>70</sup> to check which base has the highest likelihood of affecting function at respective polymorphic positions.

**Data analysis.** MLGs identified using SSRs were evaluated at regional, state, and global levels. At the regional level, data were partitioned into northeastern (NEI) and eastern (EI) regions for comparison. At the state level, analysis included haplotypic SSR data from southern India (SI)<sup>32</sup>. Due to the geographic distance between collection sites in West Bengal, the state was divided into WB-EI and WB-NEI. All SSR data collected from India were compared to a global set of single representative samples of 13\_A2 variants. To ensure consistent comparisons peak calls were calibrated to the allele naming scheme<sup>34</sup>. Allele sizes were also compared to other data sets from India<sup>32</sup>.

Analysis of SSR genotypes was conducted using the program Structure v.2.3.4<sup>71</sup>. The data were run using a 20,000 repeat burn-in and 1,000,000 MCMC repeats under both an admixture model and a model assuming no admixture. Structure was allowed to incorporate population data into the analysis (LOCPRIOR) as this setting can be used to discern more subtle population differences. Independent runs of the model used  $K$  values from 1 to 10 with 20 replicate runs at each value of  $K$ . The optimal  $K$  was estimated using the Evanno method in the web tool Structure Harvester<sup>72</sup>. In addition, the optimal  $K$  was inferred through direct observation of groupings of

the samples by their estimated likelihoods ( $\ln P(D)$ ). All runs for the optimal K values, as well as surrounding non-optimal K values, were averaged using CLUMPP v. 1.1.2<sup>73</sup> and visualized with the program Distruct v. 1.1<sup>74</sup>.

Locus statistics and population statistics were generated using the R library *poppr*<sup>75</sup>. Clone correction was performed as needed using the *clonecorrect* () function in *poppr*. Clone correction reduces the number of individuals in each population to one representative per MLG. *Poppr* was also utilized to generate minimum spanning networks (MSN) for MLGs at each spatial level. At the regional level, the index of association ( $I_a$ ) and the standardized index of association ( $\bar{F}_d$ ) were calculated and evaluated for significance using 999 permutations for both clone corrected and non clone corrected data. *Polysat*<sup>76</sup> was used at the state level to calculate pairwise fixation indices ( $F_{ST}$ ). Since the amount of selfing in the system is unknown, a simple frequency calculator in *polysat* was used to generate allele frequencies for the calculation of  $F_{ST}$ . It should be noted that this calculator assumes all alleles have the same chance to be present in more than one copy, and as such may result in an underestimated  $F_{ST}$ . For the purposes of the calculation, the populations were assumed to be autopolyploid.  $F_{ST}$  values were converted to the number of migrants per generation ( $Nm$ ) using the R library *StrataG*<sup>77</sup>. A ploidy of 3 was assumed for the calculation.

**Data Availability.** All data generated or analysed during this study are included in this published article (and its Supplementary Information files). Maps in Fig. 1 created using ArcGIS® online by Esri. ArcGIS® and ArcMap™ are the intellectual property of Esri and are used herein under trial version. Copyright® Esri. All rights reserved. For more information about Esri® software, please visit [www.esri.com](http://www.esri.com).

## References

- Kamoun, S. *et al.* The Top 10 oomycete pathogens in molecular plant pathology. *Mol Plant Pathol.* **16**, 413–434 (2015).
- Martin, M. *et al.* Genomic characterization of a South American *Phytophthora* hybrid mandates reassessment of the geographic origins of *Phytophthora infestans*. *Mol. Biol. Evol.* <https://doi.org/10.1093/molbev/msv241> (2016).
- Fry, W. E. *et al.* The 2009 late blight pandemic in the eastern United States: causes and results. *Plant Dis.* **97**, 296–306 (2013).
- Haverkort, A. J. *et al.* Societal costs of late blight in potato and prospects of durable resistance through cis gene modification. *Potato Res.* **51**, 47–57 (2008).
- Fisher, C. M. *et al.* Emerging fungal threats to animal, plant and ecosystem health. *Nature.* **484**, 186–194 (2012).
- Hu, C. H. *et al.* Recent genotypes of *Phytophthora infestans* in eastern USA reveal clonal populations and reappearance of mefenoxam sensitivity. *Plant Dis.* **96**, 1323–1330 (2012).
- Haas, B. J. *et al.* Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature.* **461**, 393–398 (2009).
- Gomez-Alpizar, L., Carbone, I. & Ristaino, B. J. An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. *Proc. Natl. Acad. Sci. USA* **104**, 3306–3311 (2007).
- Martin, M. D. *et al.* Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. *Nat. Commun.*, 10:1038/ncomms1372. (2013).
- Martin, M. D. *et al.* Persistence of the mitochondrial lineage responsible for the Irish potato famine in extant New World *Phytophthora infestans*. *Molecular Biol. Evol.* <https://doi.org/10.1093/molbev/msu08> (2014).
- Goss, E. M. *et al.* The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. *Proc. Natl. Acad. Sci. USA* **111**, 8791–8796 (2014).
- Grunwald, N. J. & Flier, W. G. The biology of *Phytophthora infestans* at its center of origin. *Annu. Rev. Phytopathol.* **43**, 171–90 (2005).
- Wang, J. *et al.* High levels of diversity and population structure in the potato late blight pathogen at the Mexico centre of origin. *Mol. Ecol.* **26**, 1091–1107 (2017).
- Fry, W. E. *et al.* Population genetics and population diversity of *Phytophthora infestans*. *Oomycete Genetics and Genomics*. John Wiley & Sons, Inc. pp. 139–164 (2008).
- Yoshida, K. *et al.* The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. *eLife.* **2**, e00731 (2013).
- Saville, A. C., Martin, M. D. & Ristaino, J. B. Historic late blight outbreaks caused by a widespread dominant lineage of *Phytophthora infestans* (Mont.) de Bary. *PLoS ONE* **11**(12), e0168381 (2016).
- Ristaino, J. B., Groves, C. T. & Parra, G. R. PCR amplification of the Irish potato famine pathogen from historic specimens. *Nature.* **411**, 695–697 (2001).
- Goodwin, S. B., Cohen, B. A. & Fry, W. E. Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. *Proc Nat Acad Sci, USA* **91**, 11591–5 (1994).
- Njoroge, A. W., Tusiime, G., Forbes, G. A. & Yuen, J. E. Displacement of US-1 clonal lineage by a new lineage of *Phytophthora infestans* on potato in Kenya and Uganda. *Plant Pathol.* **65**, 587–592 (2016).
- Fry, W. E. *et al.* Five reasons to consider *Phytophthora infestans* a re-emerging pathogen. *Phytopathol.* **105**, 966–81 (2015).
- Drenth, A., Goodwin, S. B., Fry, W. E. & Davidse, L. C. Genotypic diversity of *Phytophthora infestans* in The Netherlands revealed by DNA polymorphisms. *Phytopathol.* **83**, 1087–1092 (1993).
- Sjoholm, L. *et al.* Genotypic diversity and migration patterns of *Phytophthora infestans* in Nordic countries. *Fungal Biol.* **117**, 722–730 (2013).
- Cooke, D. E. L. *et al.* Genome analyses of an aggressive and invasive lineage of the Irish potato famine pathogen. *PLOS Pathog.* **8**, e1002940 (2012).
- Mariette, N. *et al.* Phenotypic and genotypic changes in French populations of *Phytophthora infestans*: are invasive clones the most aggressive? *Plant Pathol.* **65**, 577–586 (2016).
- Li, Y. *et al.* Population structure of *Phytophthora infestans* in China – geographic clusters and presence of the EU genotype Blue\_13. *Plant Pathol.* **62**, 932–942 (2013).
- Chowdappa, P. *et al.* Emergence of 13\_A2 blue lineage of *Phytophthora infestans* was responsible for severe outbreaks of late blight on tomato in south-west India. *J. Phytopathol.* **161**, 49–58 (2013).
- Derevnina, L. *et al.* Emerging oomycete threats to plants and animals. *Phil. Trans. R. Soc. B* **371**, 20150459 (2016).
- Fry, W. E. *Phytophthora infestans*: New tools (and old ones) lead to new understanding and precision management. *Ann. Rev. Phytopathol.* **54**, 529–547 (2016).
- Bhattacharyya, S. K., Shekhawat, G. S. & Singh, B. P. Potato Late Blight. Technical Bulletin No. 27. Shimla, India: CPRI (1990).
- Ristaino, J. B. & Hu, C. DNA Sequence analysis of the late blight-pathogen gives clues to the world-wide migration *Acta Horticulturae.* **834**, 27–40 (2009).
- Singh, B. P., Roy, S. & Bhattacharyya, S. K. Occurrence of the A2 mating type of *Phytophthora infestans* in India. *Potato Res.* **37**, 227–231 (1994).

32. Chowdappa, P. *et al.* Severe outbreaks of late blight on potato and tomato in South India caused by recent changes in the *Phytophthora infestans* population. *Plant Pathol.* **64**, 191–199 (2015).
33. Anonymous 2016. Agricultural Statistics at a glance 2015. Government of India.
34. Li, Y., Cooke, D. E. L., Jacobsen, E. & van der Lee, T. Efficient multiplex simple sequence repeat genotyping of the oomycete plant pathogen *Phytophthora infestans*. *J Microbiol Methods.* **92**, 316–322 (2013).
35. Goodwin, S. B. *et al.* Genetic change within populations of *Phytophthora infestans* in the United States and Canada during 1994 to 1996: Role of migration and recombination. *Phytopathol.* **88**, 939–949 (1998).
36. Li, Y. *et al.* Population dynamics of *Phytophthora infestans* in the Netherlands reveals expansion and spread of dominant clonal lineages and virulence in sexual offspring. *G3: Genes|Genomes|Genetics* **2**(12), 1529–1540 (2012).
37. Li, Y. *et al.* Changing ploidy as a strategy: The Irish potato famine pathogen shifts ploidy in relation to its sexuality. *MPMI.* **30**(1), 45–52 (2017).
38. Chimote, V. P., Kumar, M., Sharma, P. K., Singh, P. H. & Singh, B. P. Characterization of changes in phenotype and genotype of *Phytophthora infestans* isolates from India. *J. Plant Pathol.* **92**, 669–77 (2010).
39. Arora, R. K. *et al.* Late blight disease of potato and its management. *Potato J.* **41**(1), 16–40 (2014).
40. Sharma, S. & Singh, B. P. Aggressiveness and host specificity of tomato and potato isolates of *Phytophthora infestans*. *Potato J.* **42**(2), 159–163 (2015).
41. Atheya, I. *et al.* Genetic diversity and differentiation of Indian isolates of *Phytophthora infestans* as revealed by RAPD analysis. *Indian J Exp Biol.* **43**, 817–823 (2005).
42. Gotoh, K. *et al.* Characterization of some Asian isolates of *Phytophthora infestans*. *Plant Pathol.* **54**, 733–739 (2005).
43. Danies, G., Small, I. M., Myers, K., Childers, R. & Fry, W. E. Phenotypic characterization of recent clonal lineages of *Phytophthora infestans* in the United States. *Plant Dis.* **97**, 873–881 (2013).
44. Vega-Sánchez, M. E. *et al.* Host adaptation to potato and tomato within the US–1 clonal lineage of *Phytophthora infestans* in Uganda and Kenya. *Plant Pathol.* **49**, 531–539 (2000).
45. Kröner, A., Mabon, R., Corbière, R., Montarry, J. & Andrivon, D. The coexistence of generalist and specialist clonal lineages in natural populations of the Irish Famine pathogen *Phytophthora infestans* explains local adaptation to potato and tomato. *Mol. Ecol.* **26**, 1891–1901 (2017).
46. Chen, C. H. *et al.* Phenotypic and genotypic changes in the *Phytophthora infestans* population in Taiwan – 1991 to 2006. *Phytopathol.* **157**, 248–255 (2009).
47. Pasco, C. *et al.* And the nasty ones lose in the end: foliar pathogenicity trades off with asexual transmission in the Irish famine pathogen *Phytophthora infestans*. *New Phytologist.* <https://doi.org/10.1111/nph.13581> (2015).
48. Porter, L. D. & Johnson, D. A. Survival of *Phytophthora infestans* in surface water. *Phytopathol.* **94**, 380–387 (2004).
49. CIP. <https://research.cip.cgiar.org/confluence/display/wpa/India> (2009).
50. Young, H. S., Parker, I. M., Gilbert, G. S., Guerra, A. S. & Nunn, C. L. Introduced species, disease ecology, and biodiversity-disease relationships. *Trends Ecol. Evol.* **32**, 41–54 (2017).
51. Cooke, L. R., Zhang, R., Wu, L., Chen, S. H. & Forbes, G. Recent developments: late blight in Asia AsiaBlight, EuroBlight, EuroBlight workshop 14–17 (2017).
52. Forbes, G. A. Manual for Laboratory Work on *Phytophthora infestans*. Quito, Ecuador: Centro Internacional de la Papa (1997).
53. Fry, W. E. Fry L: Biology of *Phytophthora infestans* and Management of Late Blight. <http://www.plantpath.cornell.edu/Fry/Protocols.html> (2014).
54. Tooly, P. W. Use of uncontrolled Freezing for liquid nitrogen storage of *Phytophthora sp.* *Plant Dis.* **72**, 680–682 (1988).
55. Cooke, D. E. L. & Duncan, J. M. Phylogenetic analysis of *Phytophthora* species based on the ITS1 and ITS2 sequences of ribosomal RNA gene repeat. *Mycological Research.* **101**, 667–677 (1997).
56. Jong, K. K. *et al.* A Genetic Marker Associated with the A1 Mating Type Locus. *Phytophthora infestans J. Microbiol. Biotechnol.* **15**(3), 502–509 (2005).
57. Lee, Y. S. & Kim, K. J. Genetic DNA marker for A2 mating type in *Phytophthora infestans*. *J. Microbiol.* **40**, 254–259 (2002).
58. Griffith, G. W. & Shaw, D. S. Polymorphisms in *Phytophthora infestans*: four mitochondrial haplotypes are detected after PCR amplification of DNA from pure culture or from host lesions. *Appl Environ Microbiol.* **64**, 4007–14 (1998).
59. Goodwin, S. B., Drenth, A. & Fry, W. E. Cloning and genetic analysis of two highly polymorphic, moderately repetitive nuclear DNAs from *Phytophthora infestans*. *Curr. Genet.* **22**, 107–115 (1992a).
60. Danies, G. *et al.* An ephemeral sexual population of *Phytophthora infestans* in the northeastern United States and Canada. *PLOS One. Online journal.*, <https://doi.org/10.1371/journal.pone.0116354> (2014).
61. Sharma, R. *et al.* Molecular characterisation of potato genotypes for late blight resistance. *Potato J* **40**(2), 164–172 (2013).
62. CPRI annual report 2012–13 (2013).
63. Klarfeld, S., Rubin, A. E. & Cohen, Y. Pathogenic fitness of oosporic progeny isolates of *Phytophthora infestans* on late-blight-resistant tomato lines. *Plant Dis.* **93**, 947–53 (2009).
64. Euroblight: A potato late blight network for Europe <http://euroblight.net/protocols/> (2014).
65. Chacón, M. G., Andrade-Piedra, J. L., Gessler, C. & Forbes, G. A. Aggressiveness of *Phytophthora infestans* and phenotypic analysis of resistance in wild *Petota* accessions in Ecuador. *Plant Pathol.* **56**, 549–561 (2007).
66. Abrámoff, M. D., Magalhães, P. J. & Ram, S. J. Image Processing with ImageJ. *Biophotonics International.* **11**(7), 36–42 (2004).
67. Hollander, M. & Wolfe, A. D. (1973) *Nonparametric Statistical Methods*. New York: John Wiley & Sons. Second edition (1999).
68. Cardenas, M. *et al.* Genetic diversity of *Phytophthora infestans* in the Northern Andean region. *BMC Genet.* **9**, <https://doi.org/10.1186/1471-2156-12-23> (2011).
69. Panda, A. *et al.* EumicrobeDBLite: a lightweight genomic resource and analytic platform for draft oomycete genomes. *Mol. Plant Pathol.* <https://doi.org/10.1111/mpp.12505> (2016).
70. Kelley, L. A. *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols.* **10**, 845–858 (2015).
71. Pritchard, J. K. & Stephens, M. Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics.* **155**, 945–959 (2000).
72. Earl, D. A. & vonHoldt, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* **4**, 359–361 (2012).
73. Jakobsson, M. & Rosenberg, N. A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* **23**, 1801–1806 (2007).
74. Rosenberg, N. A. DISTRICT: a program for the graphical display of population structure. *Mol. Ecol. Notes.* **4**, 137–138 (2004).
75. Kamvar, Z. N., Tabina, J. F. & Grünwald, N. J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and or sexual reproduction. *PeerJ.* e281 (2014)
76. Clark, L. & Jasieniuk, M. Polysat: an R package for polyploid microsatellite analysis. *Mol. Ecol. Res.* **11**(3), 562–566 (2011).
77. Archer, F. I., Adams, P. E. & Schneiders, B. B. STRATAG: An R package for manipulating summarizing and analysing population genetic data. *Mol. Biol. Res.* **17**, 5–11 (2017).
78. Esri. ArcGIS online. Environmental Systems Research Institute Inc. 380 New York Street, Redlands, CA 92373, USA. <http://www.esri.com/software/arcgis/arcgisonline> (2017).

## Acknowledgements

SGR, STr and TD acknowledge financial support from CSIR, New Delhi Grant Number 60(0108)/13/EMR-II. Funding for a portion of AS work from USDA AFRI Grant Number 2011-68004-30154. DC acknowledges financial support from the Scottish Government and AHDB Potatoes. We also sincerely acknowledge the anonymous reviewers for improving the manuscript.

## Author Contributions

T.D. co-wrote the main manuscript text and did the major experimental work. A.S. and J.B.R. conducted and interpreted the pop gen analyses and co-wrote the manuscript. K.M. did experimental work. STe did the statistical analyses. D.C. edited the manuscript and annotated the 13\_A2 variants. STr was the co-Principal Investigator of the project and annotated the effector sequences. W.E.F. edited the manuscript. S.G.R. conceived and designed the experiments, was the Principal Investigator for project, wrote the main manuscript text and edited it. All authors reviewed and approved the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-22192-1>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018