# Population Genetic Analysis of *Phytophthora Infestans* in Sichuan, Southwestern China

**Sadia Hassan**
Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers
**Bingyan Xie**
Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers
**Ying Li**
Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers
**Hong Hao Li**
Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers
**Jian Ling** ( lingjian2005@126.com )
Chinese Academy of Agricultural Sciences https://orcid.org/0000-0001-9515-1204
**Muhammad Yaseen**
Ayub Agricultural Research Institute
**Muhammad Shahid Iqbal**
Ayub Agricultural Research Institute
**Qian Kun**
Chinese Academy of Agricultural Sciences Institute of Vegetables and Flowers
**Muhammad Ibrahim Tahir**
University of Arid Agriculture

**Research article**

**Keywords:** Phytophthora infestans, genetic diversity, Blue-13 lineage, mating type, self-fertile SSR genotype

**Posted Date:** September 3rd, 2020

**DOI:** https://doi.org/10.21203/rs.3.rs-52556/v1

## Abstract

**Background:** *Phytophthora infestans*, an oomycete causing late light disease, limits the potato production all over the world. It is the major food security threat across developing countries. Microsatellites (SSRs) have been considered as the most powerful markers to study population biology genetics, ecology, epidemiology, and evolution of Phytophthora. The genotypic population structure of *Phytophthora infestans* in Sichuan province of China was identified by using 12 highly informative polymorphic single sequence repeats (SSR) markers. The mating type analysis was carried out for each isolate as a phenotypic character. In the years 2010-2015, a collection of 209 single-lesion isolates was made by simple random sampling from potato fields.

**Results:** Genotypic results revealed 114 unique genotypes and two main linages. One lineage resemble to reference strain T30-4 and other is dominated by 13-A2 (Blue_13) variants. 95% of the variation was observed within population. Only 27 isolates with A1 mating types were found followed by A2 (83) and self-fertile (99) isolates. The ratio of A1 as compared to A2 mating type was less in recent years showed limited chances for sexual reproduction but the dominating percentage of self-fertile isolates may rise the opportunity for the production of oospores and subsequently increase the genotypic diversity in the pathogen population.

**Conclusion:** The findings of the study provide recent genetic structure and mating type status of the late blight pathogen, are helpful for monitoring and evaluation of *P. infestans* populations in Sichuan province of China.

## Background

*Phytophthora infestans*, an oomycete, most importantly limiting the potato production all over the world (Fry, 2008a) and threatening the food security especially in the developing countries (Forbes, 2012). The late blight epidemics in 1840s caused potato famine in Ireland, resulted a huge migration and death of one million people (Ristaino et al., 2001). Potato, being a rich source of starch, is one of the most important source of food globally and playing significant role in food safety (Haas et al., 2009). The economic damage due to *P. infestans* in developing countries have been assessed 10 billion ponds, whereas, 12 billion pounds in total worldwide per year (Haverkort et al., 2009)

Being heterothallic, the mating system of *P. infestans* is dependent on the mating types (A1 & A2) to reproduce sexually (Judelson, 1997). Sexual reproduction gives rise to the genetic variation and evolution of the pathogen more rapidly. The clonal lineage, US-1, dominated the population of *P.infestans* outside central Mexico till 1970s (Fry, 2008b). Later on in 1980s, the mating type A2 was discovered from many countries in Europe which changed the pace of late blight disease (Goodwin et al., 1994; Garry et al., 2005). In only a few years, the old dominant population was replaced by the new population and spread worldwide (Fry, 2008b) with increased fitness (Vleeshouwers et al., 2011). In Netherlands, a new lineage of *P. infestans* was reported in 2004 termed 13-A2 (or "Blue_13"), is more aggressive and resistant (Cooke *et al.*, 2012b). The new lineage has been reported outside Europe and posing a severe threat to tomato and potato production worldwide (Chowdappa et al., 2013). Stirling, a resistant potato cultivar, have been overcome by the isolates belong to this lineage since 1990s in England and became threatening potato crop production (Cooke et al., 2012a)

China is the largest producer of potato in the world and its potato farming industry has grown rapidly in recent years (Dongyu et al., 2005; Le et al., 2008). The earliest report of the serious damage caused by late blight in China to potato production was in 1940s (Guo et al., 2010). A2 phenotypes were detected comparatively shortly in China once they appeared in Europe (Zhang et al., 1996). Many studies on *P.infestans* had been carried out from different provinces of China to reveal its population structure in recent years. Researchers (Li et al., 2013) did a nationwide analysis, closely clustered *P. infestans* genotypes on the basis of their geographical origins and made a first ever report on the presence of 13_A2 ('Blue_13') clonal lineage in China. A particular clonal lineage was identified by seven SSR markers for a population genetic analysis of *P. infestans* in Yunli, Northern Shanxi of China (Tian et al., 2015). Later, a study conducted on *P. infestans* isolates collected during 2009–2011 showed the population was strictly asexual and distant from the European lineage 13_A2 ('Blue_13') while the dissemination of mating types (A1& A2) transformed over the years with the dominating number of self-fertile isolates in northwest of China (Tian et al., 2016).

Sichuan Province (Southwest) is the first among top ten potato producing regions of China with the production of 9.48 million ton and an average yield of 15.0 tons/hectare. This province has a varied climate with a subtropical monsoon in its eastern portions, the western part is very cold in winter and mild in summer with the mountains and high altitude and subtropical in its southern part. The Late blight epidemic is more frequent and serious in the south part of China because of substantial precipitation and long rainy season. In this way, decrease of potato production is still extremely serious once potato late blight epidemic occurs.

Previously, from the southwestern (Sichuan) China, forty-three *P. infestans* isolates showed high proportion of A2 and self-fertile isolates grouped in a same cluster with ten genotypes identified by only five SSR markers (Zhao, 2016). As the previous study was based on only five SSR markers with limited number of isolates collected in years 2010–2012 from Sichuan, the current study analyzed a set of 181 isolates collected in years 2010–2015 combined with 28 isolates from the previous studies. Moreover, for throughput genotypic analysis, a fixed set of twelve SSR markers (Li et al., 2012) was used.

The genetic diversity of the pathogen and its population structure are required to understand the pathogenic potential against host resistance (Ranathunge et al., 2009; McDonald & Linde, 2002). High biodiversity may give grounds for the development of unique genotypes, distinctive biological characters and resistance against fungicides (Rampersad, 2013). Keeping in view the importance of Sichuan province for potato production and pace of late blight disease, the objectives of the current study were to: (1) explore the variations in *P. infestans* populations in Sichuan with 12 SSR markers; (2) reveal the genetic diversity among the isolates of *P. infestans*; (3) investigate the proportion of self- fertile oospores in contrast to the mating types A1 and A2; and (4) assess the existence of sexual reproduction in recent years.

## Results

## SSR markers Profile

Among 209 isolates, 51 alleles were yielded by a set of 12 SSR markers with an average of 4.3 each locus. The marker Pi70 was discovered as monomorphic amongst all isolates. Polymorphic markers detected 4.5 alleles per locus on an average. A mean PIC value was detected as 0.35. The highest PIC values, 0.60 and 0.58, were shown by D13 and SSR4 markers respectively. The averaged major allele frequency was 0.65. The numbers of multilocus genotypes (NG) observed were 71 with an average measured 5.9 for all microsatellite markers. The total numbers of effective alleles detected were ranging from 1.0 to 2.7. The private number of alleles (Np) refer to the unique number of alleles in a population were ranging from 1 to 8.7. The Gene diversity, being a possibility of alleles randomly selected in a population are different, was varying from 0.13 to 0.64 detected by PinfSSR11 and PinfSSR4 markers respectively averaging 0.42. The observed heterozygosity was 0.00 to 0.96, while expected heterozygosity value was 0.00 to 0.63. The highest level of observed and expected heterozygosity was shown by marker PinfSSR4 (0.96 and 0.63 respectively). The index of Shannon's information varied from 0.00 to 1.3. The coefficient of inbreeding (*F*is) was estimated by locus over populations, ranged from − 0.821 to 0.341. The negative values were observed for *Fis* as excess heterozygotes observed in the population. The list of loci and their detailed variability values are shown in Table 1 (end of document).

Table 1
Microsatellite polymorphism across loci

| Locus | D13 | G11 | Pi04 | Pi4B | Pi63 | Pi70 | PinfSSR2 | PinfSSR3 | PinfSSR4 | PinfSSR6 | PinfSSR8 | PinfSSR11 | Mean |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Dye color** | FAM | NED | VIC | PET | VIC | VIC | PET | NED | FAM | VIC | FAM | NED | |
| **Size range** | 100–185 | 130–180 | 160–175 | 200–295 | 265–280 | 185–205 | 165–180 | 255–275 | 280–305 | 230–250 | 250–275 | 325–360 | |
| **PIC value** | 0.6 | 0.55 | 0.43 | 0.4 | 0.37 | 0 | 0.041 | 0.422 | 0.58 | 0.315 | 0.43 | 0.13 | 0.36 |
| **MAF** | 0.57 | 0.514 | 0.5 | 0.51 | 0.55 | 1 | 0.98 | 0.54 | 0.5 | 0.74 | 0.49 | 0.93 | 0.65 |
| **NG** | 18 | 12 | 5 | 4 | 2 | 1 | 3 | 6 | 9 | 3 | 4 | 4 | 5.916666 |
| **GD** | 0.627 | 0.615 | 0.537 | 0.513 | 0.494 | | 0.042 | 0.527 | 0.641 | 0.386 | 0.536 | 0.135 | 0.421 |
| **Na** | 13 | 7 | 3 | 3 | 2 | 1 | 2 | 4 | 8 | 3 | 3 | 3 | 4.333 |
| **Np** | 8.750 | 4.500 | 3.000 | 2.750 | 2.000 | 1.000 | 2.000 | 3.000 | 5.250 | 2.750 | 2.750 | 2.500 | 3.354 |
| **Ne** | 2.730 | 2.434 | 2.147 | 2.053 | 1.972 | 1.000 | 1.049 | 2.118 | 2.738 | 2.165 | 1.181 | 1.654 | 1.937 |
| **I** | 1.333 | 1.020 | 0.822 | 0.756 | 0.686 | 0.000 | 0.110 | 0.813 | 1.172 | 0.822 | 0.293 | 0.598 | 0.702 |
| **Ho** | 0.619 | 0.878 | 0.972 | 0.883 | 0.896 | 0.000 | 0.031 | 0.840 | 0.963 | 0.931 | 0.124 | 0.540 | 0.640 |
| **He** | 0.611 | 0.581 | 0.534 | 0.513 | 0.493 | 0.000 | 0.046 | 0.526 | 0.634 | 0.537 | 0.144 | 0.394 | 0.418 |
| **Fis** | -0.014 | -0.511 | -0.821 | -0.722 | -0.817 | | 0.341 | -0.598 | -0.519 | -0.735 | 0.143 | -0.371 | |
| **Fst** | 0.041 | 0.025 | 0.001 | 0.004 | 0.004 | | 0.006 | 0.012 | 0.005 | 0.007 | 0.031 | 0.006 | |
| **P Fst** | 0.001 | **0.001** | 0.612 | 0.097 | 0.089 | | 0.801 | 0.013 | 0.196 | 0.008 | 0.012 | 0.398 | |

PIC = polymorphism information content,GD = genetic diversity, NG = number of genotypes, Na = observed number of alleles, Np = private alleles, Ne = effective alleles, I = Shannon's information, Ho = observed heterozygosity. He = expected heterpzygosity, Fis = fixation index, Fst = population differenciation

A non-significant result was observed at $P < 0.05$ for the Hardy- Weinberg Equilibrium test in the locus SSR11 in all four populations (S&SW, W&NW, E, N&NE) SSR6 in population E and N&NE, SSR2 in S&SW, N&NE and W&NW, SSR3 and SSR4 in W&NW and D13 only in S&SW populations.

## Genetic variation by region

Genotyping with 12 SSR markers, discovered 114 genotypes among 209 isolates. Out of total genotypes observed, 34 were revealed in the region E, 36 from N&NE, 27 in S&SE and 17 from W&NW region of Sichuan. 80 genotypes were found only once, and the other 31 genotypes showed matches ranging from maximum 12 to 2 in number. A closer look at genotypes revealed the variations among the *P. infestans* isolates were attributed by absence or presence of alleles at certain hypervariable locus (D13, G11 and SSR4).

The population variation among the four regions of Sichuan was discovered by AMOVA analysis. The variation among population ($P$ 0.001) was 5% while the variation within population was higher at 95%. The genetic variations among regions at $P < 0.001$ was found significant (Fst = 0.049) by randomization test. The estimated gene flow among populations was averaged Nm = 5.07. The genetic indices of diversity for the four populations of *P.infestans* are explained in Table 2. The total number of different alleles (Na) 3 to 3.75, private alleles (Np) 0.00 to 0.25 and effective alleles (Ne) 1.78 to 1.98, averaged across all loci among all four regions. The number of private and different alleles was higher for E and N&NE while the effective allele number was higher for W&NW region. The amount of observed heterozygosity (Ho) was more than expected heterozygosity (He) in all four regions that ranged from 0.62 to 0.65 and 0.38 to 0.44 respectively. The average percentage of polymorphic loci was 91.67% among all regions. A higher value of diversity was shown by the Shannon's Index (I) in all regions ranging 0.60 to 0.75.

Table 2
Polymorphism across regions

| Population | N | Na | Np | Ne | I | Ho | He | Gn |
|---|---|---|---|---|---|---|---|---|
| **E** | 58 | 3.5 | 0.250 | 1.989 | 0.735 | 0.629 | 0.431 | 34 |
| **N&NE** | 81 | 3.75 | 0.250 | 1.943 | 0.708 | 0.654 | 0.416 | 36 |
| **W&NW** | 27 | 3.16 | 0.000 | 2.027 | 0.757 | 0.62 | 0.443 | 17 |
| **S&SW** | 43 | 3 | 0.250 | 1.789 | 0.609 | 0.655 | 0.381 | 27 |

N = number of samples in each region, Na = number of different alleles, Np = private alleles, Ne = effective alleles, I = Shannon's Index, Ho = observed heterozygosity, He = expected heterozygosity, Gn = number of genotypes

# The Population Structure

The algorithm of model-based clustering in the software STRUCTURE v.2.2 did the analysis for genetic structure. The correction of STRUCTURE output was done (Fig. 2) by following as explained by Evano *et al.* (2005). The first high peak for $\Delta K$ and $K = 2$ revealed the two main clusters. Cluster 1 contained all 8 reference isolates with 49 isolates (23%) from Sichuan population and 28 genotypes in it. While Cluster 2 carries 160 isolates (76%) and 86 genotypes.

To validate the STRUCTURE output, Principal Component Analysis (PCA) was conducted. The three groups were resulted by the first two coordinates plotted, were contrary to the clusters shown by STRUCTURE analysis. PCA groups were defined by Population 1 (Pop. 1) consisting of 16 isolates including a reference isolate (T30-4), Population 2 (Pop. 2) gathers 188 isolates of Sichuan population with the reference 13-A2 variant (NL05246). The remaining 6 reference isolates and 5 isolates from Sichuan population were clustered in Population 3 (Pop. 3). The genotypes in the three populations were 9, 100 and 5 respectively.

A dendrogram was constructed for a detailed look and validate the STRUCTURE and PCA results. All 209 isolates with 8 references were included in analysis and tree constructed using NJ method. The six reference isolates (88133, 90128, F80029, IPO428-2, VK1.4 and NL08452) with 5 Sichuan isolates were clearly separated on the branches away from the isolates clustered with the reference T30-4 and NL05426 as shown by PCA results. The mating type of isolates, regional distribution, sampling years and structure K = 2 clusters were subsequently highlighted with different colors on the dendrogram (Fig. 4A B C D). The results showed no stricken geographic boundaries among the regional populations and isolates suggesting a complete gene flow between the isolates.

# Mating Type

Out of total 209 isolates of *P.infestans* examined, 27 were revealed A1, 83 were mating type A2 and 99 self-fertile isolates. The two mating types with self-fertile isolates prevailed in all four sub- regions (E, N&NE, W&NW and S&SW) of Sichuan but not in every year from 2010–2015. Out of 99 self- fertile isolates, 5 were found in the year 2010 from only West and Northwest region while other regions had mating types A1 and A2 in that year. During the years 2012–2015, a gradual rise in the mating type A2 and self-fertile isolates was observed as compare to A1 mating that was still present till the year 2015. A detailed proportion of mating types and self-fertile isolates in all four regions among the years 2010–2015 is presented in Fig. 5.

The mating types with self-fertile isolates were also compared and presented in the phylogenetic tree in Fig. 4A. The tree showed 16 isolates with mating types A1 and A2, mostly from the year 2013, were branched with the T30-4 reference strain derived from a cross between isolates 80029 × 88133 (Table 3).

Table 3
Origin and year of collection of *Phytophthora infestans* reference isolates used in the study.

| Isolate | Year | Origin | Mating type |
|---|---|---|---|
| 88133 | 1988 | WUR | A2 |
| 90128 | 1990 | WUR | A2 |
| F80029 | 1980 | WUR | A1 |
| IPO428-2 | 1992 | WUR | A2 |
| VK1.4 | 1958 | - | A1 |
| *T30-4 | 1992 | WUR | A1 |
| NL05246 | 2009 | WUR | A2 |
| NL08452 | 2009 | WUR | A1 |

* This is not a field isolate but derived from a cross between isolates 80029 × 88133

## Discussion

The Sichuan province of China had been under study for the late blight disease causing *P.infestnas* population patterns previously by Li et al., (2013), he used 10 SSR markers on *P. infestans* samples (69 isolates) collected during the years 2004–2009. Later (ZHAO Qing, 2016) used only 5 SSR markers on small set (43) of isolates collected during 2010–2012. In our study, an extended number of samples (209 isolates) from the years 2010–2015 were studied by 12 highly informative microsatellite markers (Li et al., 2012) for more detailed and recent insight in the structure of *P.infestans* populations in Sichuan province of China.

The 12 markers revealed low average of polymorphism (0.36) but the markers D13, G11 and PinfSSR4 showed significantly higher PIC values (0.60, 0.55 and 0.58 respectively). Moreover, 5 loci (D13, G11, Pi4B, PinfSSR3 and PinfSSR4) showed tri alleles in some isolates in the population. The results revealed that the genetic diversity by locus over populations was lower overall that suggested an asexual breeding of a clonal lineage in almost 90% of the isolates. Although the percentage of self-fertile isolates and genotypic diversity was found high but the significant negative values of coefficient of inbreeding and excess of heterozygotes indicated that the Sichuan population was clonal. Moreover, the AMOVA results showed a high level of variation within population than among populations. Also pair wise population differentiation analysis was significantly lower, varied from 0 to 0.093 at $P < 0.001$ that suggests no stricken boundaries among the geographic origins of the pathogen. The extent of gene flow was also found high (Nm = 5.07). It was suggested that if the value of Nm is found more than one, the amount of variation among populations will be reduced and the migration is farther significant than the genetic drift (McDermott & McDonald, 1993)

The results were more clearly validated by 3 populations of PCA and similar pattern of isolates on phylogenetic tree as compared to the clusters shown by STRUCTURE analysis. In our study, an interesting result found was the Pop.1 of PCA (Fig. 3), indicated a lineage resemble to the reference isolate T30-4 that is a derived isolate resulted by crossing 80029 and 88133 isolates (in the year 1992), with mating types A1 and A2 respectively belonging to 1980s. The 16 isolates of this lineage have 8 genotypes and mating types A1 and A2 only. Most of the isolates (14 of 16) belong to year 2013 and only one isolate from each of the year 2011 and 2014. These isolates belonged to all regions except the south and southwest (S&SW) region of Sichuan. This result suggested the presence of limited sexual reproduction during the year 2013. No self-fertile isolate was found in this lineage. To our knowledge, this result was not found in earlier study of Sichuan population by Li et al., (2013) and Qing *et al.*, (2016) from the field isolates during the years 2010–2012. The STRUCTURE analysis showed the same isolates in Cluster 1, and also clearly separated by the phylogenetic tree from the rest of the isolates and reference strains Fig. 4B.

The current study revealed the similar results for the presence of Blue_13 lineage and several genetic variants as described previously by Li et al., (2013). Almost 90% of the isolates were clustered with the reference strain of Blue_ 13 variant (NL05246) shown by Pop.2 in PCA and Cluster 2 of STRUCTURE analysis also highlighted on the phylogenetic tree. This lineage revealed 101 genotypes in 188 isolates. The lineage is more dominated by the self-fertile isolates as compare to A2 mating type. Although self-fertile isolates can undergo sexual production and can result into high level of genotypic diversity like in the central Mexico (Grünwald et al., 2003), but in this cluster the genotypes are clustered with 13-A2 or Blue_ 13 variant suggest that the isolates are surviving clonally and subclonal variation was high in this cluster.

# Mating type

Moreover, this study shows higher percentage of self-fertile isolates (47.3%) followed by A2 (39.7%) and A1 (12%) mating types in the population with a gradual increase in every successive years (Fig. 5). The study by Li et al., (2013) found 91% (21 of 23 isolates) A2 mating type and no self-fertile isolates during the years 2004–2009 while our study revealed 5 self-fertile isolates in the very next year (2010) alone from west and northwest (W&NW) region of Sichuan. In contrast, the study carried by ZHAO Qing et al., (2016) revealed mating types A2 (69%), A1 (4.65%) and SF isolates (25.5%) in the years 2010–2012. Comparing previous results with the recent once showed a significant increase in SF isolates over A2 mating types. Moreover, the overall percentage of A1 mating type was more than the previous once. Some other studies from different provinces of China also showed an increased frequency of A2 and SF isolates in successive years of study (Han et al., 2013) found high genetic diversity and increased self-fertility in Gansu in 2007. A notable shift was observed from A1 to A2 mating type and increased number of SF isolates during the years 2009–2011 in northwestern China (Tian et al., 2016). The population in Fujian province was found dominated by SF isolates (63.11%), less A1 (36.89%) and no A2 mating type was found in the years 2010–2012 (Zhu et al., 2015)When look back to some of older studies, it was found that in years 1995–1998, the rate of mating type A2 in Hebei, Shanxi and Inner Mongolia was higher than in Yunnan, Chongqing and Sichuan, only 1.4% self-fertile isolates were found all over (Zhang et al., 2001). The south part of China, Yunnan, was found with 3 SF isolates among the samples collected during the years 1998–2004 (Guo et al., 2010). The scenario changed over years, as mention above, the previous studies showed A1 mating type was overcame by A2 and now recent study revealing A2 is being dominated by self- fertile isolates. This situation is quite alarming because self-fertile isolates can produce sexual oospores independently and also more with A1 mating types that is still prevalent in China (Han et al., 2013; Tian et al., 2015). In this study, although the proportion of A1 was less but its presence in recent years can give opportunities to A2 and SF isolates for sexual mating and production of more oospores with increased genetic variability. It will make the control and management strategies of the late blight disease more complicated because oospores have the capability to survive harsh environments and more virulent to its hosts.

## Conclusions

The study provides phylogenetic tree and mating type status showing a dominating number of isolates resemble the variant Blue_13 A2 of *P. infestans* and self-fertile in nature posing a serious threat to potato crop worldwide. The insight to these results will be helpful to modify the management strategies for the late blight pathogen.

## Methods

# Sampling of P. infestans isolates

A set of 181 isolates were collected by visiting the fields in different regions of Sichuan province in the years 2010–2015 along with 28 isolates from the previous studies (ZHAO Qing, 2016) during the years 2010–2012. The number of isolates for each region and year of collection is shown in (Table 4).

Table 4
Geographic origins, the number of isolates in each region and year of collection of *Phytophthora infestans* population.

| Region | No.of isolates | Origin | Duration of collection | Year of collection | Total isolates in a year |
|---|---|---|---|---|---|
| **North &North East (N&NE)** | 81 | Pengzhou | 2010–2015 | 2010 | 22 |
| | | Beichuan | 2012 | 2011 | 13 |
| | | Xindu | 2010 | 2012 | 40 |
| | | Guangyuan | 2012 | 2013 | 50 |
| **East (E)** | 58 | Xuanhan | 2013/2014 | 2014 | 76 |
| | | Shangba | 2012–2014 | 2015 | 8 |
| | | Wanzhou | 2012 | | |
| | | Zhongxian | 2012–2013 | | |
| | | Shizhu | 2012–2013 | | |
| | | Wanyuan | 2014 | | |
| **West & North west (W&NW)** | 27 | Kangding | 2010 | | |
| | | Luding | 2011–2013,2015 | | |
| | | Daofu | 2010,2013 | | |
| **South and Southwest (S&SW)** | 43 | Yibin | 2013 | | |
| | | Ebian | 2014 | | |
| | | Zhaojue | 2010 | | |
| | | Puge | 2010 | | |
| | | Yanyuan | 2012 | | |
| | | Emei | 2014 | | |
| **Total** | **209** | | | | |

For the ease of analysis, the 20 different sampling locations were categorized into four regions according to their geographical locations, that is, North and North-East (N&NE), West and North-West (W&NW), South and South-West (S&SW) and East (E) of Sichuan province (Fig. 1).

The farmers of the sampling area reported late blight disease outbreaks in potato production fields. Infected leaves were collected on the basis of symptoms during the growing season, usually during the months of May to July. Records were made side by side for the sampling location, date of collection and number of samples collected in one region. Each sample collected represents distinct plant.

## Isolate purification and storage

Collecting infected potato leaves with single lesion from the field and put in the plastic bags has performed the isolate purification and storage process. In laboratory, the infected leaflets were placed on water agar media (1.5%) in Petri dishes and sealed. A small amount of infected tissue of 0.5 cm thickness, was picked from the lesion and kept underneath the fresh potato slice in Petri dish under sterilized condition. The process was completed for each sample separately in a separate petri dish and kept at 18ºC for at least 5–7 days in the incubator in dark. After the incubation time, the fungal mycelium emerged on the tuber slice, were transferred to the Rye Sucrose Agar (RSA) media amended with ampicillin 100 µg/ mL (Shattock et al., 1990). Single zoospore cultures of all isolates per infected leaves were maintained separately. Fresh mycelial agar plugs of each *P. infestans* sample were picked and stored in separate vials (2 ml) containing rye seeds broth (1.5 ml) for further analysis. The stored isolates were refreshed on RSA medium in 9 cm Petri dishes, kept and maintained at 18 ºC in the incubator for further use.

## Determination of mating types

The unknown mating types of *Phytophthora infestans* were investigated by combining the test isolates to A1 (VK98014) or A2 (NL88133) reference isolates (Table 3) on RSA medium separated by 2 cm. Plates were kept for incubation at 18ºC for 2–3 weeks in the dark. As the mycelial contact was observed after 5–7 days, the zone of contact was checked at 100x magnification under the microscope for the next seven days. The unknown test isolate when found to have oospores in the in the media with a known A1 reference isolate, the tested isolate was categorized as mating type A2 and vice versa. The isolates under test, when producing oospores both with known A1 and A2 mating types, were classified and confirmed for their self-fertile nature by isolating single zoospore to eliminate the possibility of mixing the cultures (Smart et al., 1998).

## DNA extraction

The mycelial agar plug of *P. infestans* individual isolates were taken from the margins of actively growing colony on RSA for seven days and transferred to rye broth. The mycelium was collected after 3–4 days of incubation in dark at 20 ºC and subjected to lyophilization and DNA extraction by CTAB method (Judelson & Tooley, 2000). Elution and suspension was performed using 200 *µL* ultra-pure water and kept at -20ºC for further analysis.

## SSR analysis

The genomic DNA taken from each of the *P. infestans* isolate was amplified with a set of known microsatellite markers (Knapova & Gisi, 2002; Lees et al., 2006; Li et al., 2010). The procedure for the amplification of twelve microsatellite markers was followed as explained in the previous studies (Li et al., 2012). QIAGEN Multiplex PCR kit was used and the total volume for reaction was 20 µl. The fluorescent labels are described in Table 1. The final concentration of primers D13, G11, Pi4B, Pi63, Pi70, PinfSSR2, PinfSSR3, PinfSSR4, PinfSSR6 and PinfSSR11 was 0.05 µmol/L an for PinfSSR8 and Pi04 was 0.3 µmol/L. Amplification were run in thermal cycler T100 (*BIO-RAD*) with 15 min denaturation at 95ºC, 30 cycles at 95ºC for 20 s, 58ºC for 90 s, 72ºC for 60 s and a final extension at 72ºC for 20 min. The PCR products were sent to the company (GENEWIZ, Suzhou Industrial Park, Suzhou, China) for genotyping. The allele scoring was evaluated by software GENEMAPPER v. 3.7 (Applied Biosystems). In order to understand and find the genetic relationship among *P.infestans* isolates, a total of eight already known isolates (Table 3) included in the study as a reference with a Blue_13 genotype variant (NL05246) from Europe previously used by Li et al., (2013)

## Data analysis

## Ploidy level:

Being a diploid organism, *P.infestans* showed two alleles per locus. However, in few cases, some of the loci showed more than two alleles, and these isolates are considered as a result of polyploidy or aneuploidy. This situation makes the data analysis complicated, as majority of the analytical software accept haploid or diploid based data. A modification was done for the loci showing three alleles to look like a diploid by assigning specific allele sizes. To prevent the simulated fault of heterozygous decrease, a diploid set of data was created by adding only the minimum and maximum allele sizes of the specific loci with three stable alleles (Chen et al., 2008). This self-styled set of diploid data was used to determine the genetic diversity and distance based clustering.

## Genetic diversity:

The variation in microsatellite loci was analyzed by using diploid set of data as described above, with all 209 *P. infestans* isolates and calculated the Polymorphic Information Content (PIC) value, Major Allele Frequency (MAF), the number of alleles observed (Na), number of genotypes (G), genotype diversity (GD) at each locus using Power Marker software version 3.25 (Liu & Muse, 2005). Effective number of alleles (Ne), Shannon's information index (I), observed heterozygosity (Ho), expected heterozygosity (He), coefficient of inbreeding (*F*is) and Hardy-Weinberg equilibrium (HWE) with $P \geq 0.05$, was estimated for all populations per locus in GENALEX 6.5 (PE, 2012). Analysis of molecular variance (AMOVA) was conducted to investigate the genetic variation among and within populations of P.infestans by GENALAX v.6.5.

## Genetic Structure analysis

The genetic structure of population was determined using the SSR data by STRUCTURE v. 2.3.4 by a Bayesian model based algorithm of clustering (Pritchard et al., 2000; Falush et al., 2003). The analyses were run separately with and without background information of the populations. The first run of analyses was completed by allocating the geographic regions considering the area of sample collection as supposed origin of populations and the second one was done by giving no such information so the STRUCTURE analysis can allocate the populations for each individual. The reference isolates were also combined with the population data. As suggested in the STRUCTURE manual, the admixture model was used with correlated frequencies of alleles. The number of populations (K) was determined using a burn-in period of 10,000 in each run and the collection of data completed around 100,000 Markov Chain Monte Carlo (MCMC) from K = 1 to K = 10 replications. For each cluster, the probability value was averaged across runs. This process assemblies all the isolates into populations and assess the their association in each population (Pritchard et al., 2000). Based on the rate of change in LnP(D) values between successive populations (K), the K value was determined by the log probability od data (LnP(D)). The peak value of K was determined by the simulation method (Evanno et al., 2005) using the software STRUCTURE HARVESTER v.0.6.92(Earl, 2012).

The Principal Coordinate Analysis (PCA) was completed based on matrix of genetic distance in statistical software *GenAlEX v.6.5*. The self-styled special data set was used to calculate the genetic based distance between isolates with references and the method of Neighbor Joining (NJ) was performed using Power Marker 3.25 software based on shared allelic distance (Liu & Muse, 2005). The phylogeny tree was formed by MEGA 6 (Tamura et al., 2013) expoloiting the distance based matrix created in Power Marker software.

### Abbreviations

*p. infestans: Phytophthora infestans,* SSR: Single sequence repeats, N&NE: North and North-East, W&NW: West and North-West, S&SW: South and South-West, E: East, RSA: Rye Sucrose Agar, PIC: Polymorphic Information Content, MAF: Major Allele Frequency, Na: number of alleles observed, G: number of genotypes, GD: Genotype diversity, Ne: Effective number of alleles, I: Shannon's information, index, Ho: Observed heterozygosity, He: Expected heterozygosity, *F*is: Coefficient of inbreeding, HWE: Hardy-Weinberg equilibrium, AMOVA: Analysis of molecular variance, K: number of populations, NJ:Neighbor Joining, NG: Multilocus genotypes, Np: Private number of alleles.

### Declarations

## -Ethics approval and consent to participate

Ethics approval does not apply to this study as it has not directly involvement of humans or animals. *P. infestans* isolates have been taken from the Institute of Vegetables and Flowers, Key Laboratory of the Ministry of Agriculture on Genetic Improvement for Vegetable Germplasm, Chinese Academy of Agricultural Sciences, Beijing 100081, China, The field sampling had been conducted in accordance with the institutional and national guidelines set for the research station/institutes involved in the current study. There was no need to get specific/additional permission to conduct the lab experiments or molecular analyses. The field studies did not involve endangered or protected species

## -Consent for publication

Not Applicable

## -Availability of data and material

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## -Competing interests

The authors declare that no competing interests exist.

## -Funding

The research was supported by grants from the National Key Research and Development Project (2018YFD0201200) and the Key Laboratory of Biology and Genetics Improvement of Horticultural Crops, Ministry of Agriculture. Moreover, the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## -Authors' contributions

BX, YL conceived and designed the research. BX, YL and JL managed the project, HHL performed leaf sampling from farmers field area of Sichuan province, Isolation purification and storage process. SH, MY and QK performed lab experiments and different molecular analysis. SH and MSI performed genetic data analysis. SH, MIT and MSI analyzed and interpreted data and prepared figures and tables. SH, MY, MIT drafted and processed the manuscript and all authors helped throughout this process and take active part in critical revisions and improvements in important intellectual contents. All authors read the manuscript critically and approved the final version of manuscript for publication. All authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## -Acknowledgements

We are thankful to National Key Research and Development Project administration and Key Laboratory of Biology and Genetics Improvement of Horticultural Crops, Ministry of Agriculture for providing financial assistance for publication of research findings

### References

1. Chen C-H, Sheu Z-M, Wang T-C. Host specificity and tomato-related race composition of Phytophthora infestans isolates in Taiwan during 2004 and 2005. Plant Disease. 2008;92:751–5.
2. Chowdappa P, Kumar NB, Madhura S. 2013. Emergence of 13_A2 blue lineage of Phytophthora infestans was responsible for severe outbreaks of late blight on tomato in south-west India. Journal of Phytopathology. 2013;161:49–58.
3. Cooke DE, Cano LM, Raffaele S. Genome analyses of an aggressive and invasive lineage of the Irish potato famine pathogen. PLoS pathogens. 2012a;8:e1002940.
4. Cooke DE, Cano LM, Raffaele S. Genome analyses of an aggressive and invasive lineage of the Irish potato famine pathogen. PLoS Pathog. 2012a;8:e1002940.
5. Dongyu Q, Kaiyun X, Liping J. Development of potato industry and food security in China. Scientia Agricultura Sinica. 2005.
6. Earl DA. Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method. Conservation genetics resources. 2012;4:359–61.
7. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14:2611–20.
8. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003;164:1567–87.
9. Forbes G. Using host resistance to manage potato late blight with particular reference to developing countries. Potato research. 2003;55:205–16.
10. Fry W. Phytophthora infestans: the plant (and R gene) destroyer. Mol Plant Pathol. 2008a;9:385–402.

11. Fry W. Phytophthora infestans: the plant (and R gene) destroyer. Mol Plant Pathol. 2008b;9:385–402.
12. Garry G, Forbes G, Salas A, Santa Cruz M, Perez W, Nelson R. Genetic diversity and host differentiation among isolates of Phytophthora infestans from cultivated potato and wild solanaceous hosts in Peru. Plant Pathol. 2005;54:740–8.
13. Goodwin SB, Cohen BA, Fry WE. Panglobal distribution of a single clonal lineage of the Irish potato famine fungus. Proceedings of the National Academy of Sciences. 1994;91:11591-5.
14. Grünwald NJ, Goodwin SB, Milgroom MG, Fry WE. Analysis of genotypic diversity data for populations of microorganisms. Phytopathology. 2003;93:738–46.
15. Guo L, Zhu X-Q, Hu C-H, Ristaino JB. Genetic structure of Phytophthora infestans populations in China indicates multiple migration events. Phytopathology. 2010;100:997–1006.
16. Haas BJ, Kamoun S, Zody MC. Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. Nature. 2009;461:393–8.
17. Han M, Liu G, Li JP. Phytophthora infestans field isolates from Gansu province, China are genetically highly diverse and show a high frequency of self fertility. J Eukaryot Microbiol. 2013;60:79–88.
18. Haverkort A, Struik P, Visser R, Jacobsen E. Applied biotechnology to combat late blight in potato caused by Phytophthora infestans. Potato research. 2009;52:249–64.
19. Judelson HS. The Genetics and Biology ofPhytophthora infestans: Modern Approaches to a Historical Challenge. Fungal Genet Biol. 1997;22:65–76.
20. Judelson HS, Tooley PW. Enhanced polymerase chain reaction methods for detecting and quantifying Phytophthora infestans in plants. Phytopathology. 2000;90:1112–9.
21. Knapova G, Gisi U. Phenotypic and genotypic structure of Phytophthora infestans populations on potato and tomato in France and Switzerland. Plant Pathol. 2002;51:641–53.
22. Le V, Ngo X, Brurberg M, Hermansen A. Characterisation of Phytophthora infestans populations from Vietnam. Australas Plant Pathol. 2008;37:592–9.
23. Lees A, Wattier R, Shaw D, Sullivan L, Williams N, Cooke D. Novel microsatellite markers for the analysis of Phytophthora infestans populations. Plant Pathol. 2006;55:311–9.
24. Li Y, Govers F, Mendes O. A new set of highly informative SSR markers for Phytophthora infestans population analysis assembled into an efficient multiplex. Mol Ecol Resour. 2010;10:1098–105.
25. Li Y, Lee T, Zhu J. Population structure of Phytophthora infestans in China–geographic clusters and presence of the EU genotype Blue_13. Plant Pathol. 2013;62:932–42.
26. Li Y, Van Der Lee T, Evenhuis A. 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. 2012;2:1529-40.
27. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 2005;21:2128–9.
28. Mcdermott JM, Mcdonald BA. Gene flow in plant pathosystems. Annu Rev Phytopathol. 1993;31:353–73.
29. Mcdonald BA, Linde C. Pathogen population genetics, evolutionary potential, and durable resistance. Annu Rev Phytopathol. 2002;40:349–79.
30. Pe PRS. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012;28:2537–9.
31. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945–59.
32. Rampersad SN. Genetic structure of Colletotrichum gloeosporioides sensu lato isolates infecting papaya inferred by multilocus ISSR markers. Phytopathology. 2013;103:182–9.
33. Ranathunge N, Ford R, Taylor P. Development and optimization of sequence-tagged microsatellite site markers to detect genetic diversity within Colletotrichum capsici, a causal agent of chilli pepper anthracnose disease. Molecular ecology resources. 2009;9:1175–9.
34. Ristaino JB, Groves CT, Parra GR. PCR amplification of the Irish potato famine pathogen from historic specimens. Nature. 2001;411:695.
35. Shattock RC, Shaw DS, Fyfe AM, Dunn JR, Loney KH, Shattock JA. Phenotypes of Phytophthora infestans collected in England and Wales from 1985 to 1988: mating type, response to metalaxyl and isoenzyme analysis. Plant Pathol. 1990;39:242–8.
36. Smart C, Willmann M, Mayton H. Self-fertility in two clonal lineages ofPhytophthora infestans. Fungal Genet Biol. 1998;25:134–42.
37. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular biology evolution. 2013;30:2725–9.
38. Tian Y, Sun J, Li H. Dominance of a single clonal lineage in the Phytophthora infestans population from northern Shaanxi, China revealed by genetic and phenotypic diversity analysis. Plant Pathol. 2015;64:200–6.
39. Tian Y, Yin J, Sun J. Population genetic analysis of Phytophthora infestans in northwestern China. Plant Pathol. 2016;65:17–25.
40. Vleeshouwers VGaA, Raffaele S, Vossen JH. Understanding and Exploiting Late Blight Resistance in the Age of Effectors. Annu Rev Phytopathol. 2011;49:507–31.
41. Zhang Z, Zhu J, Song B, Li Y, Tian S, Jiang H. Further investigations on A2 mating type of Phytophthora infestans in China. Journal of Hebei Agricultural University. 2001;24:32–7.
42. Zhang ZM, Li Y, Tian S, Zhu J, Wang J, Song B. The occurrence of potato late blight pathogen (Phytophthora infestans) A2 mating type in China. Journal of Agricultural University of Hebei. 1996;19:62–5.
43. Zhao Qing ZZ, Li Ying,YangYu-Hong. Li Hui-Xia, Xie Bing-Yan. An analysis on phenotypic and genotypic variation of Phytophthora infestans in Sichuan Province. Mycosystema. 2016;35:52–62.

44. Zhu W, Yang L-N, Wu E-J. Limited sexual reproduction and quick turnover in the population genetic structure of Phytophthora infestans in Fujian, China. Scientific reports. 2015;5:10094.

## Figures

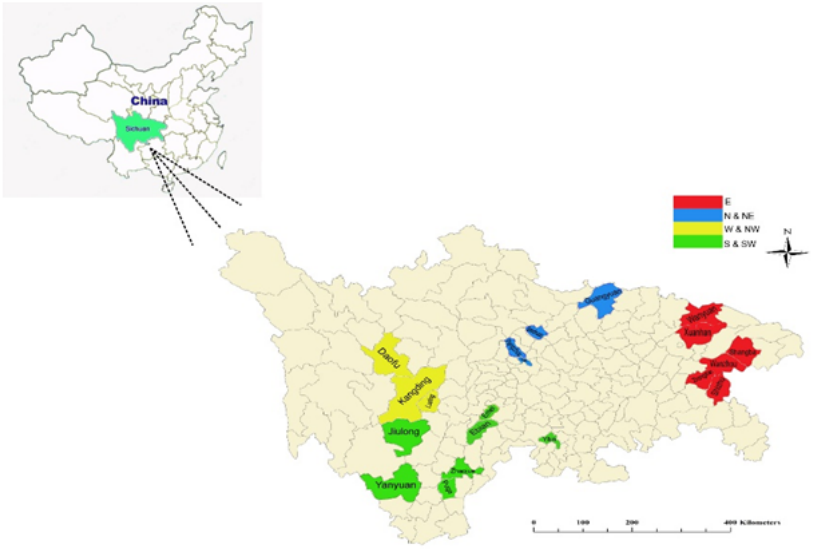


**Figure 1**

Map of Sichuan province showing the sampling regions of P.infestans isolates used to analyze in this study Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

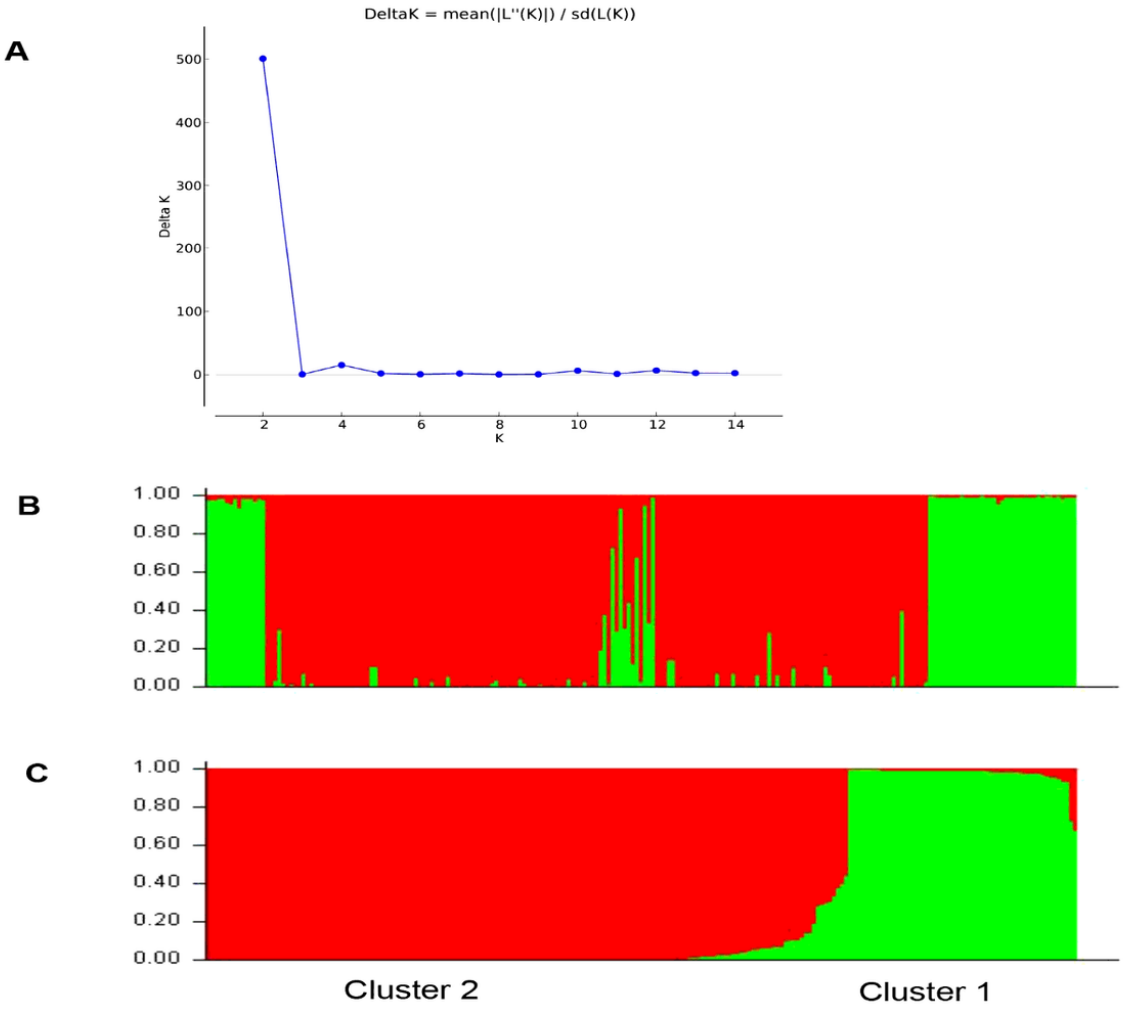


Figure 2

Structure analysis of the population STRUCTURE analysis of 12 microsatellite loci. (A) Results of ΔK calculation, the second order range of change of LnP(D) with respect to K (1-10); ΔK for K= 2-14 revealed a single distinct peak at K=2. (B) Plot example of the raw STRUCTURE output for one run (K=2) by Q value in multiple line, and (C) organized by Q value in single line.

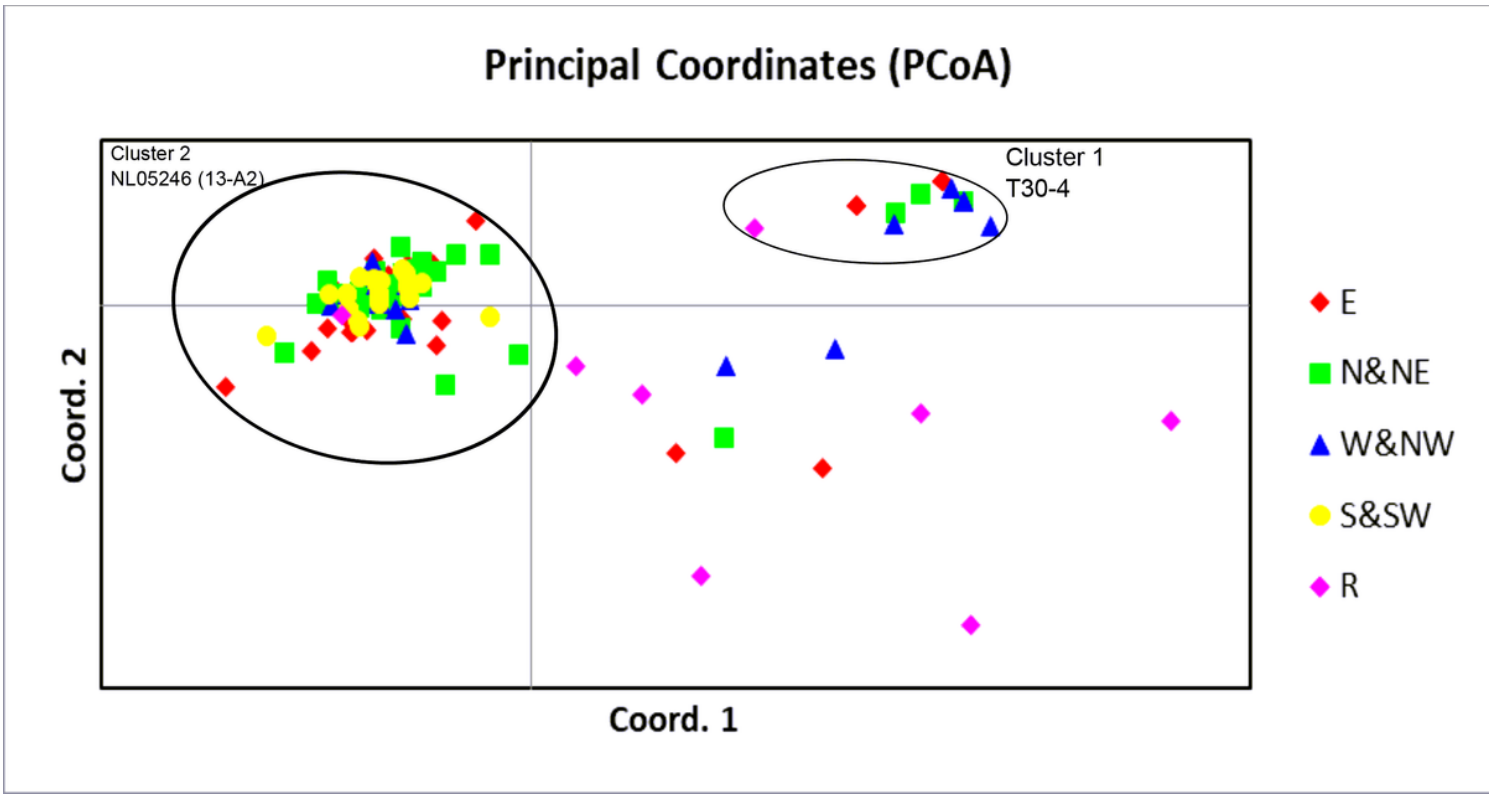


Figure 3

Principal coordinates analysis (PCoA) of Phytophthora infestans population in Sichaun China Each point represents a genotype. Percentages of variation explained by the first 3 axes (1, 2, and 3) are 35.19%, 17.37% and 7.27%, respectively. Sampling regions represented by shape and color: E, East; N&NE, North &Northeast; W&NW, West &Northwest; S&SW, South &Southwest; R, Reference strains

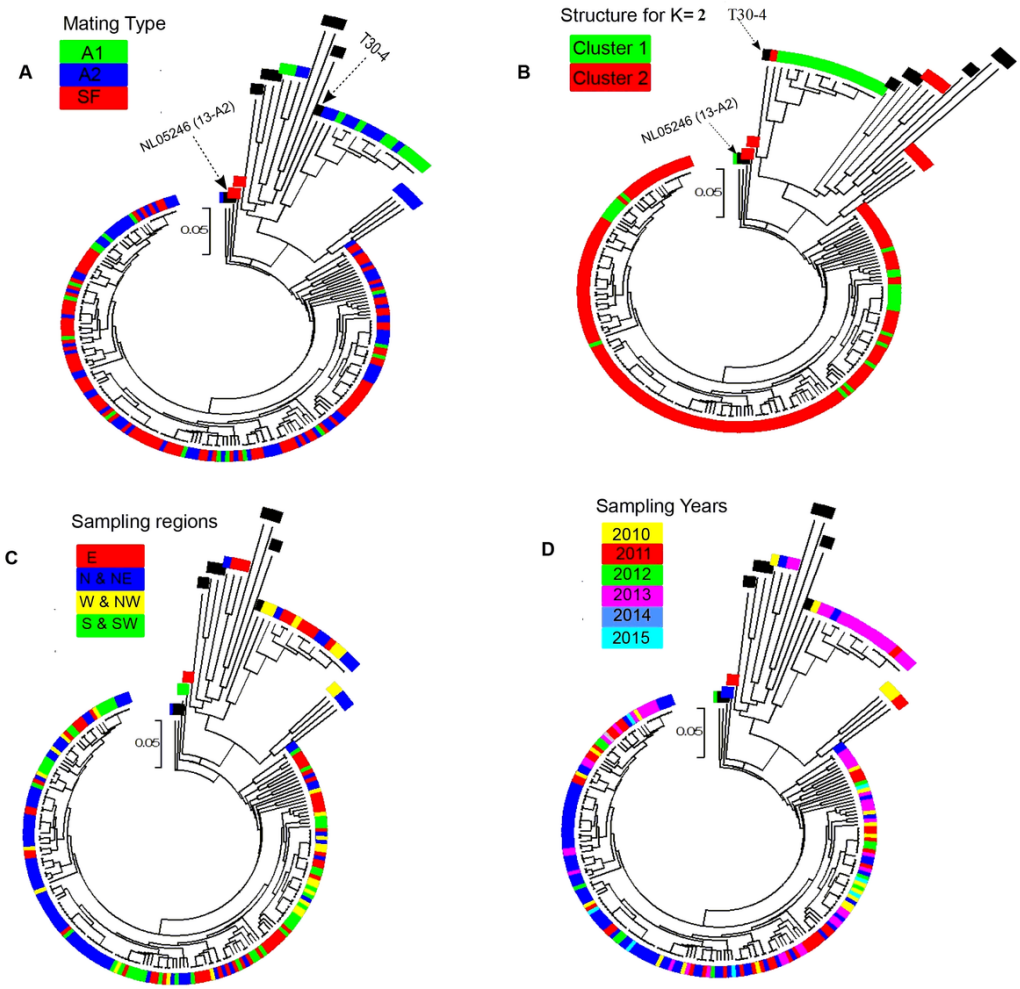


**Figure 4**

Dendrogram of all 209 isolates with 8 reference strains Represents the information highlighted in different colors for: (A) Mating type, (B) Clusters in structure K=2, (C) Sampling region (D) Sampling years of isolates. Reference strains are highlighted in black in each (A, B, C &D)

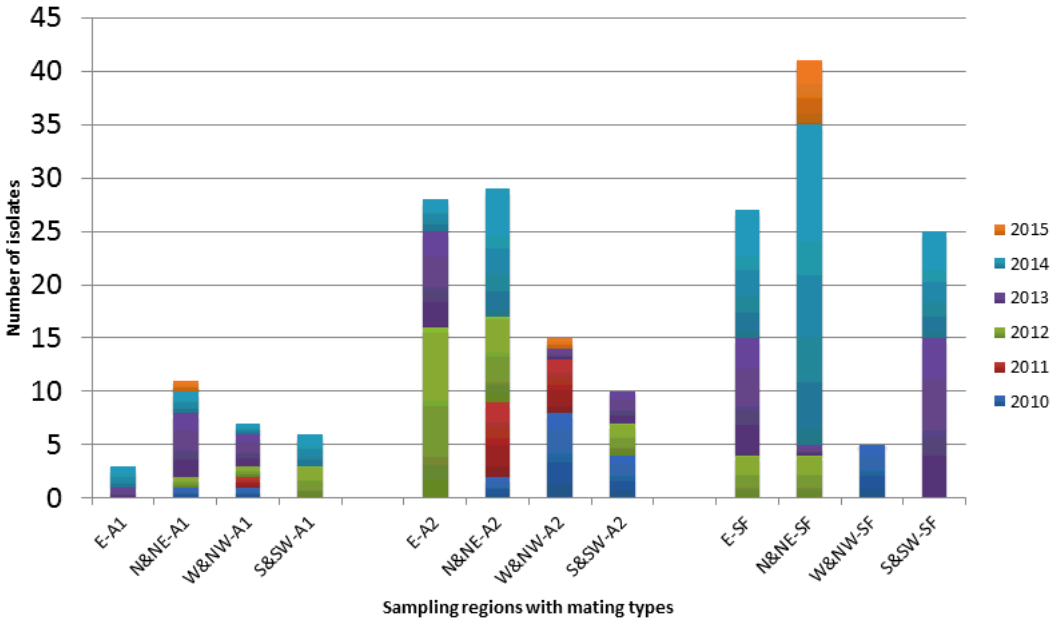


**Figure 5**

Mating type proportion in each sampling region during the years 2010-2015