

Supplemental File 1. SSR Classifier Tutorial

This program identifies unknown lineages of *Phytophthora infestans* based on 12 commonly used SSR markers (microsatellite data). The program is hosted on the USAblight.org website under "[Identify SSR Genotype](#)" and the database represents most of the currently named SSR lineages of *P. infestans*.

How it works

The SSR classifier compares the SSR data for your isolates of interest to over 2,500 described isolates from global populations in our *P. infestans* database. The closest genetic match is determined using Bruvo's distance (described [here](#)).

How to run SSRs

This program uses results obtained from a 12-plex microsatellite amplification for *Phytophthora infestans*. These 12 microsatellite loci are well described for *Phytophthora infestans* and vary across lineages, allowing identification of different genotypes of the pathogen. Genotyping using this SSR method first involves determining the microsatellite alleles for each of the 12 loci. The methods for running SSRs have been previously described [here](#). A simplified version of the protocol can be found on the USAblight.org website under "[SSR Genotyping Methods](#)"

How to format data

The raw microsatellite reads obtained from SSR genotyping must be processed and scored to create input data for the SSR classifier. Once the alleles are determined from the peak data, they can be formatted similar to the [example CSV](#) (see link or below) and uploaded to this tool for genotype identification. Three numbers are separated by forward slashes for each locus to represent all three alleles in triploid individuals. If only two alleles are present, the third number is a zero. There is one column for each locus, along with metadata columns for sample name, genotype, region, country, date, host, number of missing loci, and number of present loci. Each row beneath the header represents the data for one sample.

| Sample | Genotype | Region | Country | Date | Host | D13 | PinfSSR8 | PinfSSR4 | Pi04 | Pi70 | PinfSSR6 | Pi63 | PiG11 | Pi02 | PinfSSR11 | PinfSSR2 | Pi4B | Missing | Present |
|---------|----------|--------|---------|------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---------|---------|
| EXAMPLE | unknown | Europe | Poland | 2016 | unknown | 136/136/0 | 260/266/0 | 284/284/0 | 166/170/0 | 192/192/0 | 244/244/0 | 273/279/0 | 142/156/0 | 258/266/0 | 341/356/0 | 173/173/0 | 213/217/0 | 0 | 12 |

Fig. 1. Example formatting of an input CSV for the SSR classifier.

How to upload data

Upload a CSV file containing metadata about your sample as well as SSR data for the 12 loci of interest. Follow the format in the [example CSV](#). We recommend downloading this CSV, replacing the data with your own, and re-uploading it. The classifier will still work if your data is missing some loci, but at least 10 loci are recommended for the most accurate results. Include as many samples as you want.

How to interpret results

If your unknown sample has a close match, the program will return information about the match from our database of *Phytophthora infestans* genotypes (**Fig. 2**). The tool first presents a table of input samples with their closest match and metadata associated with the closest match. For example, in Figure 2, input sample EG_256 most closely matches an isolate of lineage EU_23_A1 found in Egypt. After this table follows a more in-depth description of each match and where to learn more about the matching genotype. Finally, a list of all submitted samples for which no match was found are listed. Note just because it shows a match to an Egyptian genotype, this does not mean that was the source. The region, country, sample date and host are additional metadata from the isolate we used in developing the classifier and can be used in downstream analysis.

| Results | | | | | | |
|--------------|-------------------|--------|---------|-----------|----------------|------------------|
| Input Sample | Matching Genotype | Region | Country | Date | Host | Bruvo's Distance |
| EG_256 | EU_23_A1 | Africa | Egypt | 2012/2013 | S_lycopersicum | 0.00 |
| EG_152 | EU_23_A1 | Africa | Egypt | 2013/2014 | S_lycopersicum | 0.07 |
| BU-11.22 | Unknown | Europe | UK | 2013 | S_tuberosum | 0.03 |
| BU-13.P04 | Unknown | Europe | UK | 2013 | S_tuberosum | 0.00 |
| BU-13.P10 | Unknown | Europe | UK | 2013 | S_tuberosum | 0.00 |

| Input Sample | Description |
|--------------|---|
| EG_256 | Based on the markers you provided, the genotype of this isolate is most likely EU_23_A1 or US_23. These two lineages are synonymous. In the eastern United States, US_23 has been found on both potato and tomato and has mating type A1. For more information, see https://doi.org/10.1094/PDIS-03-11-0156-RE . |
| EG_152 | Based on the markers you provided, the genotype of this isolate is most likely EU_23_A1 or US_23. These two lineages are synonymous. In the eastern United States, US_23 has been found on both potato and tomato and has mating type A1. For more information, see https://doi.org/10.1094/PDIS-03-11-0156-RE . |
| BU-11.22 | The markers you provided indicate your isolate is closely related to a sample in our database of Unknown genotype. This means your sample could be a clone that has been found but not described, or it could be from a sexually recombining population. Unknown. |
| BU-13.P04 | The markers you provided indicate your isolate is closely related to a sample in our database of Unknown genotype. This means your sample could be a clone that has been found but not described, or it could be from a sexually recombining population. Unknown. |
| BU-13.P10 | The markers you provided indicate your isolate is closely related to a sample in our database of Unknown genotype. This means your sample could be a clone that has been found but not described, or it could be from a sexually recombining population. Unknown. |

No close matches found for sample TU0899 .
 No close matches found for sample TU0761 .
 No close matches found for sample TU0803 .
 No close matches found for sample TU0750 .

Figure 2. Example output from SSR classifier.

Having issues?

If you are unable to upload your samples and run the SSR classifier, please email the CSV file you are using and the problem occurring to acoombe@ncsu.edu or acsavill@ncsu.edu.

References

Li Y, Cooke DE, Jacobsen E, van der Lee T. Efficient multiplex simple sequence repeat genotyping of the oomycete plant pathogen *Phytophthora infestans*. *Journal of microbiological methods*. 2013;92(3):316-22.

Saville A, Ristaino, JB. Genetic structure and subclonal variation of extant and recent US lineages of *Phytophthora infestans*. *Phytopathology*. 2019;109(9):1614-27.

Saville, A.C., La Spada, F., Faedda, R, Migheli, Q., Scanu, B, Ermacora, et al. Population structure of *Phytophthora infestans* in Italy, Algeria and Tunisia. Plant Pathol. 2021; 7: 2165-2178.

Other Resources

[USABlight](#): Genotype tracking website for North America

[EuroBlight](#): Genotype tracking website for Europe

[Tizón Latino](#): Genotype tracking website for Latin America

[Help notes for scoring SSR alleles in *Phytophthora infestans*](#): Allele calling guide developed by David Cooke.

For information on the SSR binning scheme used by this tool, see “[Population structure of *Phytophthora infestans* collected on potato and tomato in Italy](#)” by Saville et al. 2021, Supplemental Table 3.

