Making Sense of *Phytophthora infestans* diversity at national and international scales

DAVID E. L. COOKE¹, ALISON K. LEES¹, POUL LASSEN² & JENS GRØNBECH-HANSEN²

¹The James Hutton Institute, Invergowrie, Dundee, DD2 5DA, United Kingdom
²Aarhus University, Faculty of Science and Technology, Research Centre Foulum,
P.O. Box 50, DK-8830 Tjele, Denmark
E-mail: david.cooke@hutton.ac.uk

**SUMMARY**

In this paper we present the status of investigations into the genetic diversity of *Phytophthora infestans* populations in Europe and an update on the means of storing, collating and interpreting such information in a central database. Simple Sequence Repeat markers (SSR), also termed microsatellites, remain the method of choice for examining population diversity as they are rapid, robust and resolve population structure at a level appropriate for identifying and discriminating clonal lineages and exposing the signature of sexual recombination in highly diverse populations. SSR analysis has progressed from 3 panels of 3-5 multiplexed markers to a single multiplex assay in which 12 markers are run in a single PCR. Frequent discovery of more than two alleles at a single SSR locus in some *P. infestans* populations suggests changes in chromosome complements. These different ploidy levels within a population have proved challenging for most methods of population genetic analysis but recent publications provide a way forward and are discussed. Several comprehensive national studies on *P. infestans* diversity are underway within Europe and, combined with some international comparisons will clarify the recent evolutionary history of this damaging pathogen. SSRs are considered selectively neutral markers. There is interest in examining sequence diversity in effector genes that are under selection pressure, in order to understand the evolution of *P. infestans* virulence/pathogenicity. In combination with neutral SSR data, effector sequences will provide a different perspective on pathogen diversity at local and international scales. The Eucablight database that was developed in the 'Eucablight' EU funded Concerted Action project is also evolving and we briefly describe its status as it merges with the larger and more flexible 'cropproblem' database that holds data for the Wheat Rust Toolbox which forms part of the Global Cereal Rust Monitoring System.

**KEYWORDS**

Microsatellites, Simple Sequence repeats, ploidy, effectors, population biology, database

**BACKGROUND**

We have reported previously the rationale for examining *P. infestans* populations and the benefits of a centralised database to store information in a standard format (Cooke *et al.* 2007b; Cooke *et al.* 2009; Cooke and Lees 2004; Hansen *et al.* 2007). Global populations of *P. infestans* are
characterised by the migration and domination of clonal lineages in many regions (e.g. Fry et al. 2009) with evidence of sexual recombination in others such as Mexico (Goodwin et al., 1992; Grunwald and Flier 2005) and Scandinavia (Brurberg et al. 1999; Brurberg et al., 2011; Widmark et al. 2007). Well documented cases of host resistance breakdown and fungicide resistance are testament to the challenges of managing potato blight. The success of integrated control strategies depends on our understanding of the pathogen diversity at local, national and international scales and the mechanisms and rates of pathogen evolution. The theoretical advantages of planned host resistance deployment strategies have been demonstrated (Skelsey et al. 2010) but their success hinges on understanding the stability of such sources of resistance in the face of a diverse repertoire of rapidly evolving pathogen effectors (Haas et al. 2009).

UPDATES ON PROGRESS

SSR markers and their analysis
SSR markers have been developed by at least three research groups (Knapova et al. 2001; Lees et al. 2006; Li et al. 2010) and used in various combinations (Knapova and Gisi 2002; protocols in www.eucabligh.org). Multiplexing (the amplification of more than one locus in a single PCR mix) has been applied to increase the efficiency of the approach. Not all markers are equally informative (i.e. the polymorphism information content or PIC score varies) or easy to score due to levels of stutter (e.g. Brownstein et al. 1996). This has led to different numbers of markers being applied in different studies (e.g. Brurberg et al. 2011). A collaboration between Wageningen University and The James Hutton Institute has resulted in a multiplex assay which includes 12 of the most informative and easy-to-score markers (Li et al. 2012). Step-by-step protocols will be published in the near future on the Euroblight web site (www.eucabligh.net). This 12-plex assay has been used on several thousand isolates in the UK and the Netherlands and is proving a rapid and efficient means of genotyping and identifying lineages according to their multi-locus genotype (MLG).

In a diploid organism, such as P. infestans, a specific SSR locus may be homozygous, in which case PCR amplification generates a single SSR allelic peak, or heterozygous resulting in two peaks on the electropherogram (Fig. 1). The occurrence of more than two peaks has been observed in our studies and is reported in the literature (Lees et al. 2006; Brurberg et al. 2011). Differences in ploidy have been observed in P. infestans (e.g. Tooley and Therrien 1991; Whittaker et al. 1991) and are consistent with these observations (Fig 1). Most methods for population genetic analysis are designed for examining diploid or haploid populations but analysis of other levels of ploidy and, in particular, populations comprising mixtures of isolates of different ploidies have proved challenging. However, a recent method (Bruvo et al. 2004) has been implemented in R as the package POLYSAT (Clark and Jaseniuk 2011). POLYSAT is capable of appropriately interpreting SSR data in populations of mixed ploidy to generate Bruvo genetic distances, principal co-ordinate analysis and F-statistics and will improve the interpretation of such P. infestans populations.
Current studies on pathogen diversity

The Eucablight database currently holds data on over 25,000 isolates, with SSR data for 5,776 isolates from fifteen European countries. This is a comprehensive resource but a combination of some notable absences in the geographic coverage of Europe and the problem of variable ploidy, discussed above, has hindered an analysis of the evolutionary history on a pan-European scale. Such issues should soon be resolved as many comprehensive datasets now fill these gaps. Examples include: the Netherlands (van den Bosch et al. 2012), Nordic regions (Brurberg et al. 2011; Grönberg et al. 2010), Northern Ireland and the Republic of Ireland (Kildea et al., 2010) France (Montarry et al. 2010; R. Corbiere pers. comm.), Poland (Chmielarz et al. 2010), Hungary (J. Bakonyi and Z. Nagy pers. comm.) as well as data generated at The James Hutton Institute on isolates provide by BayerCropScience and Syngenta as a part of their monitoring programmes. The inclusion of data from studies on isolates from North, Central and South America (F. Martin, M. Coffey, N. Grunwald and D. Cooke pers. comm.) will add a new dimension, placing the European population in a global perspective.

Effector sequencing

SSR markers as discussed above are highly polymorphic, making them a powerful tool for a relatively rapid and affordable view of population diversity and isolate discrimination even within clonal lineages; for example in P. ramorum clones, (Goss et al. 2009). For examining population
diversity at a slightly coarser phylogenetic resolution i.e. looking back further into the evolutionary history of the species, DNA sequence data has advantages (see Cooke and Lees 2004). Examples include sequencing selectively neutral genes to examine the genealogical history of \textit{P. infestans} (Gómez-Alpizar \textit{et al.} 2007) and as support for \textit{P. andina} being a hybrid of \textit{P. infestans} and another, as yet undescribed, taxa (Goss \textit{et al.} 2011). Rapid advances in the discovery of effector genes, the products of which affect host cell function to allow pathogen infection, are providing opportunities to explore the evolutionary drivers of pathogenicity in \textit{P. infestans}. The genes responsible for virulence against known resistances in the \textit{R} gene differential series have now been identified (reviewed in Vleeshouwers \textit{et al.} 2011) within a group of over 500 effectors with a distinct RXLR motif in the \textit{P. infestans} genome (Haas \textit{et al.} 2009). Sequencing a panel of these effectors amongst isolates of different MLG, defined by SSR markers, is underway at The James Hutton Institute. To date, the RXLR allelic diversity and MLG correspond closely. Avr2 provides an interesting example in which the R2 gene is ‘defeated’ by two mechanisms in different MLGs. In isolates of many MLGs the Avr2 gene is present and recognized by R2 resulting in resistance. The non-synonymous (replacement) single nucleotide polymorphisms (SNPs) reported did not affect this recognition. In isolates of other MLGs the Avr2 gene were deleted from the genome and R2 thus ‘defeated’. An alternative mechanism of overcoming R2 was also noted in isolates that were able to silence expression of the Avr2 gene (Gilroy \textit{et al.} 2010). Such studies of the sequence diversity provides a means of directly examining the evolution of pathogenicity and, in the longer term, supports strategies for selecting and deploying host resistance that is more likely to be durable (Vleeshouwers \textit{et al.} 2011).

\textit{The evolving Eucablight database}

A major achievement of the EU concerted action Eucablight project was databases of \textit{P. infestans} pathogen diversity and potato host resistance coupled with analysis tools and a web-based interface to summarise and view the data (www.eucablight.org). These databases and interfaces were planned by the project partners and implemented in 2003 at the University of Arhus (Lassen and Hansen 2005; Hansen \textit{et al.} 2007). Updates to the datasets of both host and pathogen databases have continued since the formal Eucablight project completion date in 2006. The web interface performs calculations and displays updates ‘on the fly’ so that summaries of all current data may be viewed. Both datasets are a valuable resource but it is recognised that improvements to the data upload facilities and the web interface for displaying the results are required for the databases to reach their full potential. Database and internet technology has evolved since these were designed and, as a result of additional funding, the team at the University of Arhus has designed and built a new ‘crop problem’ database. This database is based on.NET technology and has a broader scope to track changes in cereal pathogens within the ‘Eurowheat’ project (www.eurowheat.org; Jørgensen \textit{et al.} 2010) and the rust initiative (rustspore.cimmyt.org, www.wheatrust.org; Hodson \textit{et al.} 2011). Powerful mapping tools are now exploited to present wheat rust population data geographically. The Eucablight database will be migrated to this new format and alternative means of uploading data from the projects mentioned above will facilitate its expansion. The addition of data from other continents is also planned.

\textbf{ACKNOWLEDGEMENTS}

The ‘Eucablight – A Late Blight Network for Europe’ project (QLK5-CT-2002-00971) was supported by the European Commission under the Fifth Framework Programme. We gratefully acknowledge all project partners and those who have contributed data to the databases.
REFERENCES


Li Y., F. Govers, O. Mendes, A. Testa, E. Jacobsen, S.W. Huang and T.A.J van der Lee, 2010. A new set of highly informative SSR markers for Phytophthora infestans population analysis assembled into an efficient multiplex. Molecular Ecology Resources 10, 1098-105


Whittaker S.L., R.C. Shattock and D.S. Shaw, 1991. Variation in DNA content of nuclei of Phytophthora infestans as measured by a microfluorimetric method using the fluorochrome DAPI. Mycological Research, 95, 602-610
