Phytophthora infestans genomics: positional cloning of avirulence genes

S.C. Whisson, T. van der Lee[1], G.J. Bryan, R. Waugh, F. Govers[1] & P.R.J. Birch

Phytophthora infestans is the causal agent of late blight of potato (Fig.1) and tomato, and thus responsible for significant losses in these crops worldwide[1]. This is particularly the case for the cultivated potato, where it is regarded as a threat to global food security. *P. infestans* belongs to the Oomycetes, a class of organisms that includes many important plant pathogens. The Oomycetes exhibit a mycelial growth habit but are distinct from the Fungi[1]. *P. infestans* exhibits a gene-for-gene interaction with potato, in which avirulence gene (*Avr*) products from the pathogen are recognised by the host plant expressing the cognate resistance gene[2,3] (Fig. 1). This leads to the hypersensitive response (HR), a localised programmed cell death in the plant that inhibits further spread of the pathogen.

Control of late blight has been achieved mostly by the use of fungicide applications, and the deployment of resistance (either major resistance genes or field resistance). Wild species of potato, principally *Solanum demissum*, have been used as sources of resistance in breeding programs. Presently, at least 11 major genes for resistance (R1 to R11) have been identified and incorporated into differential lines for identifying races of *P. infestans*. Where R genes have been used to control late blight, they are quickly overcome by new races.

Figure 1 Gene-for-gene interaction between potato and *P. infestans*. Potato cultivars (L - R) Stirling, Bintje, 1512C, and Pentland Ace uninoculated (top row), and inoculated with different *P. infestans* races (middle and lower rows).

While the genetic basis of resistance to late blight in potato has been well characterised, the genetic basis of virulence/avirulence in *P. infestans* is less well defined. Only relatively recently has the genetic basis of avirulence been determined for more than a few avirulence genes matching R genes. Most recently, a genetic cross between two *P. infestans* races was extensively analysed by the group of Francine Govers at Wageningen University, the Netherlands[3,4]. Six dominant avirulence genes at single loci were shown to segregate in this cross. Three of the *Avr* genes, *Avr*3, *Avr*10, and *Avr*11 were tightly linked to each other. The other three *Avr* genes, *Avr*1, *Avr*2 and *Avr*4, all segregated independently.

At the molecular level, *P. infestans* is becoming a model Oomycete for the study of pathogenicity mechanisms and avirulence. Central to these studies are resources such as a molecular-genetic linkage map[4], an expressed sequence tag (EST) database[5], as well as techniques for the analysis of gene function such as transformation[6], gene silencing[7], *in planta* reporter systems[8] and heterologous expression of genes from other *Phytophthora* species[9]. DNA libraries, particularly large insert genomic libraries, are a critical resource for structural genomics and, by definition, for positional cloning.

The system most amenable to manipulating large cloned DNA inserts is the *Escherichia coli* bacterial artificial chromosome (BAC)[10]. These cloning vectors, based on a single copy F factor plasmid, are capable of stable maintenance of insert DNA up to 300 kb. Insert sizes above 100 kb are preferred for positional cloning applications, since fewer clones are required for chromosome walking or chromosome landing across linkage map intervals containing genes of interest. This is of particular consideration if the genome size is large, or if the gene of interest lies in a genomic region where the frequency of recombination is low. The genomes of many

1 University of Wageningen, the Netherlands.

fungal and Oomycete plant pathogens are relatively small compared to other organisms where positional cloning has been undertaken[11]. However, the genome size of *P. infestans* has been estimated to be 250 Mb[12], which is relatively large for an Oomycete plant pathogen, and larger than the genome of the model plant *Arabidopsis thaliana*.

When producing a BAC library that is primarily for the positional cloning of multiple target genes, it is more efficient to use a heterozygous individual from a selected F_1 mapping population as a source of high molecular weight DNA. An F_1 from a mapping population that possesses all required markers and phenotypic traits can be selected then if all the target genes segregate in the population. Recently, using a segregating population, the positions of six dominant *Avr* genes have been located on the molecular genetic linkage map of *P. infestans*. All have tightly linked AFLP markers[3]. Using one F_1 from this population, which contains all six *Avr* alleles from the genetic cross[3,4], we have produced a BAC library comprising ten-fold genome representation and an average insert size of 100 kb[13].

To convert from the genetic map to a physical map, BAC clones were pooled in an ordered manner to allow the library coordinates of *Avr*-linked BAC clones to be determined. The pools of BAC clones were screened for the AFLP markers determined to show close linkage to the six *Avr* genes. Positive clones were then fingerprinted individually and positioned relative to each other, the AFLP markers, and the *Avr* genes. For *Avr*1, which had only two linked markers 8 cM from the *Avr* locus, six BAC clones were identified. In this instance, both markers are one side of the gene, and it remains to be determined how far towards *Avr*1 the BAC contig extends (Fig. 2). This will determine how many further chromosome walking steps need to be taken to span this locus, and clone the gene.

The *Avr*2 locus is flanked by two AFLP markers. One marker is 1 cM away and the other is 5 cM distant. BAC clones containing the AFLP markers were identified and aligned. By fingerprinting the BAC clones and by testing hybridisation to each other, it was shown that the two small BAC contigs do not overlap (Fig. 2). To close the gap between the BAC contigs, chromosome walking has been initiated from the ends of the BACs. Thus, four steps from the ends of the two BAC contigs have been taken. Again, the newly identified BACs do not overlap, and so further steps are required to close the gap. However, if the ends of the extended BAC contig closest to *Avr*2 can be mapped beyond *Avr*2, then no further chromosome walking is required to clone this locus.

Physical mapping surrounding the *Avr* gene cluster was considerably simpler than for *Avr*1 or *Avr*2 due to the greater density of AFLP markers linked to these genes. In total, 11 AFLP markers are located in a 5 cM region containing *Avr*3, 10, and 11. A BAC contig was constructed which contains 11 BAC clones, spanning approximately 300 kb of genomic DNA. This contig incorporates ten AFLP markers and should also contain the *Avr*11 locus (Fig. 2). *Avr*3 and *Avr*10 are distal to the last AFLP marker, and that end of the BAC contig must be mapped in order to ascertain the position of the *Avr* genes relative to the BAC contig.

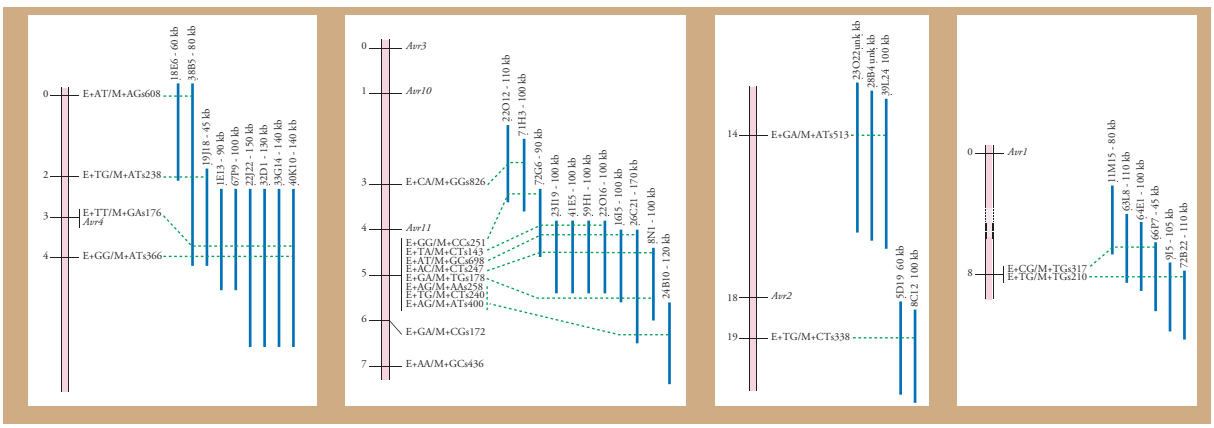

Figure 2 Section of *P. infestans* linkage groups (adapted from van der Lee et al. 2001). Indicated are the map positions of AFLP markers used in screening for identifying and verifying BAC clones spanning these genomic regions. Contigged BAC clones (not to scale) are shown as solid vertical lines to the right of the linkage group. Their respective insert sizes are indicated at the top of the solid vertical lines. The relative positions of AFLP markers in the BAC contigs are shown by dotted lines.

Similar to the *Avr* gene cluster, the genomic region surrounding *Avr*4 has several linked AFLP markers. A BAC contig across this 4 cM region has been constructed and contains 9 BAC clones spanning approximately 200 kb of genomic DNA. One BAC clone of 80 kb spans the entire 4 cM region containing the *Avr*4 locus (Fig. 2). The relatively small physical size of this region therefore makes *Avr*4 an attractive target for cloning.

To identify the gene from the non-coding DNA present on the BAC clone, we are first determining the DNA sequence of the entire 80 kb containing *Avr*4. From this, we will predict open reading frames (ORFs) for further analysis. Despite *P. infestans* being a eukaryote, introns are relatively rare in gene sequences analysed to date and so should not greatly interfere with gene prediction. Predicted ORFs can be tested for avirulence gene function in at least two approaches. Firstly, a virulent isolate can be genetically transformed with the candidate sequence. If the sequence is the *Avr* gene, then a shift to avirulence will be observed. Transformation of *P. infestans* has been used by numerous research groups for over ten years, and is being refined continually[6,14]. In the past, this has been a complicated procedure, but newer methods taking advantage of biolistic transformation are now available[15].

The second approach to proving *Avr* gene function relies on the assumption that the host plant recognises the *Avr* gene product and triggers the hypersensitive response to effect resistance. By transiently expressing the candidate *Avr* gene in a host expressing the R gene, tissue necrosis will be indicative of the HR, and therefore the interaction of host R gene with the pathogen *Avr* gene. Transient expression systems used to date have involved *Agrobacterium* delivery of the DNA expression cassette, or viral vectors such as potato virus X (PVX)[16].

Conclusions and future prospects Crop plant disease has a significant impact on the productivity and profitability of UK and overseas agriculture. Current approaches to disease control often rely on the application of expensive, highly toxic chemicals that pose serious health risks to humans and animals, and contribute to pollution of land and water supply. Understanding the molecular basis of plant resistance to pathogen challenge will allow us to identify key genes that integrate the plant defence-signalling networks, and will provide the groundwork to obtain broad-spectrum, durable resistance to crop pathogens using novel strategies that are both cost-effective and benign in the wider environment.

This project is providing the genetic and physical resources for isolation of avirulence genes (recognition of the products of which triggers resistance in the host) from *P. infestans*, the model Oomycete pathogen of potato. As yet, no avirulence genes have been isolated from Oomycetes, although considerable efforts have been directed towards obtaining the host *R* genes with which they interact. Isolation of the avirulence genes will allow the plant-pathogen gene-for-gene interaction to be studied in detail, and will provide genes that may be directly deployed to develop durable disease resistance in the plant. It is only

through a thorough understanding of the biology of *P. infestans*, the nature of its interaction with the host plant (potato), and the identification of genes involved in these processes, that we will be able to develop novel control strategies to protect potato production in Scotland.

References

[1] Erwin, D.C. & Ribeiro, O.K. (1996). *Phytophthora* diseases worldwide. APS Press, St Paul, Minnesota, USA.

[2] Al-Kherb, S.M., Fininsa, C., Shattock, R.C. & Shaw, D.S. (1995). *Plant Pathology* 44, 552-562.

[3] Van der Lee, T., Robold, A., Testa, A., van't Klooster, J.W. & Govers, F. (2001). *Genetics* 157, 949-956.

[4] Van der Lee, T., De Witte, I., Drenth, A., Alfonso, C. & Govers, F. (1997). *Fungal Genetics and Biology* 21, 278-291.

[5] Kamoun, S., Hraber, P., Sobral, B., Nuss, D. & Govers, F. (1999). *Fungal Genetics and Biology* 28, 94-106.

[6] Judelson, H.S. (1997). *Fungal Genetics and Biology* 22, 65-76.

[7] Van West, P., Kamoun, S., van't Klooster, J.W. & Govers, F. (1999). *Molecular Cell* 3, 339-348.

[8] Kamoun, S., van West, P. & Govers, F. (1998). *European Journal of Plant Pathology* 104, 521-525.

[9] Panabières, F., Birch, P.R.J., Unkles, S.E., Ponchet, M., Lacourt, I., Venard, P., Keller, H., Allasia, V., Ricci, P. & Duncan, J.M. (1998). *Microbiology* 144, 3343-3349.

[10] Shizuya, H., Birren, B., Kim, U-J., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992). *Proceedings of the National Academy of Science USA* 89, 8794-8797.

[11] Voglmayr, H. & Greilhuber, J. (1998). *Fungal Genetics and Biology* 25, 181-195.

[12] Tooley, P.W. & Therrien, C.D. (1987). *Experimental Mycology* 11, 19-26.

[13] Whisson, S.C., van der Lee, T., Bryan, G.J., Waugh, R., Govers, F. & Birch, P.R.J. (2001). *Molecular Genetics and Genomics* 266, 289-295.

[14] Randall, T.A. & Judelson, H.S. (1999). *Fungal Genetics and Biology* 28, 160-170.

[15] Cvitanich, C. & Judelson, H.S. (2001). *Fungal Genetics Newsletter* 48, (Suppl.) 129.

[16] Takken, F.L.W., Luderer, R., Gabriels, S.H.E.J., Westerink, N., Lu, R., de Witt, P.J.G.M. & Joosten, M.H.A.J. (2000). *Plant Journal* 24, 275-283.