

High-Resolution Genetic and Physical Mapping of Eastern Filbert Blight Resistance in Hazelnut

Vidyasagar Sathuvalli¹ and Shawn A. Mehlenbacher¹

Eastern filbert blight (EFB), caused by the pyrenomycete *Anisogramma anomala*, is a serious threat to the hazelnut (*Corylus avellana* L.) industry in the Pacific Northwest. A dominant allele at a single locus from the obsolete pollenizer ‘Gasaway’ confers a very high level of resistance, and has been extensively used in the hazelnut breeding program at Oregon State University. Several linked random amplified polymorphic DNA (RAPD) markers have been identified. Map-based cloning of the EFB resistance gene from ‘Gasaway’ was initiated by constructing a bacterial artificial chromosome (BAC) library for ‘Jefferson’ which is heterozygous for resistance (Sathuvalli and Mehlenbacher 2011). The BAC library was constructed using the cloning enzyme MboI and the vector pECBAC1 (BamHI site). The library consists of 39,936 clones arrayed in microtiter plates with an average insert size of 117 kb and estimated coverage of 12 genome equivalents.

A mapping population of 1,488 seedlings, which segregate for resistance, was developed and scored for flanking RAPD markers 152-800 and 268-580 to identify potential recombinants. Chromosome walking, initiated using primers designed from eight RAPD markers (Mehlenbacher et al. 2004, 2006) closely linked to resistance and extended with two further rounds of walking, identified a total of 93 BACs in the resistance region. The BAC library was screened using a PCR-based pooling and subpooling strategy (Sathuvalli and Mehlenbacher 2011). From the BACs, a total of 629 primers were designed, of which 63 pairs were polymorphic. Of these, we mapped 41 new markers to the resistance region. The new markers included 23 sequence-characterized amplified regions (SCARs), seven single-stranded conformational polymorphism (SSCP), seven high-resolution melting (HRM), and four simple-sequence repeat (SSR) markers. A high-resolution genetic map of the resistance region was created with 51 markers (41 from BACs, eight RAPD markers, and two SCAR markers derived from RAPDs) and the resistance phenotype in the mapping population using the program JoinMap v 4.0 (Van Ooijen 2006).

High-information content fingerprinting (HICF) was carried out on 93 BACs which were then assembled using the program FPC v 9.3 (Nelson and Soderlund 2009) into 22 contigs and 23 singletons. The combined maps (genetic and physical) identified a single contig that spans the EFB resistance locus. The 1,488 seedlings showed two recombination breakpoints within a single contig of three BACs (43F13, 66C22 and 85B7) and a physical distance of 135 kb. This region is the target for whole BAC sequencing and a search for disease resistance genes.

Whole BACs in the resistance region (< 1cM) were sequenced using an Illumina Ix genome analyzer, with multiplexing and barcoded adapters to reduce the cost, and paired-end reads to facilitate de novo sequence assembly. De novo assembly was carried out using the programs Velvet (Zerbino and Birney 2008) and SOPRA (Dayarian et al. 2010), and the resulting contigs were further aligned using CodonCode aligner software. The BAC sequencing and assembly generated contigs whose length ranged from 393 bp to 108,194 bp. Estimated coverage of the BACs ranged from 64 to 100 percent. The gene prediction program AUGUSTUS (Stanke et al. 2008) identified 233 genes from these sequences using *Arabidopsis* as the model. Of these, RNA-Seq data supported 32 genes at 100 percent support, of which five were in the contig that contains the resistance gene. The predicted gene sequences were compared with sequences in GenBank using a BLASTP search and identified two putative genes encoding a p-loop NTPase and F-box super family in the resistance region. Genes

¹ Department of Horticulture, 4017 ALS Building, Oregon State University, Corvallis, OR 97331.
Corresponding author: vidyasas@hort.oregonstate.edu.

in these two superfamilies have defense response properties. Future expression, complementation, and mapping studies are essential to confirm which gene confers resistance.

Literature Cited

- Dayarian, A.; Michael, T.P.; Sengupta, A.M. 2010.** SOPRA: scaffolding algorithm for paired reads via statistical optimization. *BMC Bioinformatics*. 11: 345.
- Mehlenbacher, S.A.; Brown, R.N.; Davis, J.W.; Chen, H.; Bassil, N.; Smith, D.C. 2004.** RAPD markers linked to eastern filbert blight resistance in *Corylus avellana*. *Theoretical and Applied Genetics*. 108: 651–656.
- Mehlenbacher, S.A.; Brown, R.N.; Nouhra, E.R.; Gökirmak, T.; Bassil, N.V.; Kubisiak, T.L. 2006.** A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. *Genome*. 49: 122–133.
- Nelson, W.; Soderlund, C. 2009.** Integrating sequence with FPC fingerprint maps. *Nucleic Acids Research*. 37(5): e36.
- Sathuvalli, V.R.; Mehlenbacher, S.A. 2011.** A bacterial artificial chromosome library for ‘Jefferson’ hazelnut: a resource for map-based cloning of eastern filbert blight resistance and pollen-stigma incompatibility genes. *Genome*. 54: 862–867.
- Stanke, M.; Diekhans, M.; Baertsch, R.; Haussler, D. 2008.** Using native and synthetically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*. 24: 637–644.
- Van Ooijen, J.W. 2006.** JoinMap 4.0, software for the calculation of genetic linkage maps. Wageningen, The Netherlands: Kyazama B.V.
- Zerbino, D.R.; Birney, E. 2008.** Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*. 18: 821–829.