

A transcriptome-based genetic map of Chinese chestnut (*Castanea mollissima*) and identification of regions of segmental homology with peach (*Prunus persica*)

T. L. Kubisiak · C. D. Nelson · M. E. Staton ·
T. Zhebentyayeva · C. Smith · B. A. Olukolu ·
G.-C. Fang · F. V. Hebard · S. Anagnostakis ·
N. Wheeler · P. H. Sisco · A. G. Abbott · R. R. Sederoff

Received: 21 May 2012 / Revised: 2 October 2012 / Accepted: 16 October 2012 / Published online: 30 November 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract The Chinese chestnut (*Castanea mollissima*) carries resistance to *Cryphonectria parasitica*, the fungal pathogen inciting chestnut blight. The pathogen, introduced from Asia, devastated the American chestnut (*Castanea dentata*) throughout its native range early in the twentieth century. A highly informative genetic map of Chinese chestnut was constructed to extend genomic studies in the Fagaceae and to aid the introgression of Chinese chestnut blight resistance genes into American chestnut. Two mapping populations were established with three Chinese chestnut parents, ‘Mahogany’, ‘Nanking’, and ‘Vanuxem’, totaling 337 progeny. The transcriptome-based genetic map was created with 329 simple sequence repeat and 1,064 single nucleotide polymorphism markers all derived from expressed sequence tag sequences. Genetic maps for each parent were developed and combined to establish 12 consensus

linkage groups spanning 742 cM, providing the the most comprehensive genetic map for a Fagaceae species to date. Over 75 % of the mapped markers from the Chinese chestnut consensus genetic map were placed on the physical map using overgo hybridization, providing a fully integrated genetic and physical map resource for *Castanea* spp. About half (57 %) of the Chinese chestnut genetic map could be assigned to regions of segmental homology with 58 % of the peach (*Prunus persica*) genome assembly. A three quantitative trait loci (QTL) model for blight resistance was verified using the new genetic markers and an existing interspecies (*C. mollissima* × *C. dentata*) F₂ mapping population. Two of the blight resistance QTLs in chestnut shared synteny with two QTLs for powdery mildew resistance in peach, indicating the potential conservation of disease resistance genes at these loci.

Communicated by D. Chagné

Electronic supplementary material The online version of this article (doi:10.1007/s11295-012-0579-3) contains supplementary material, which is available to authorized users.

T. L. Kubisiak · C. D. Nelson (✉)
Southern Institute of Forest Genetics, Southern Research Station,
U.S. Forest Service, Saucier, MS 39574, USA
e-mail: dananelson@fs.fed.us

M. E. Staton · G.-C. Fang
Clemson University Genomics Institute, Clemson University,
Clemson, SC 29634, USA

T. Zhebentyayeva · B. A. Olukolu · A. G. Abbott
Department of Genetics and Biochemistry, Clemson University,
Clemson, SC 29634, USA

C. Smith
Bioinformatics Research Center, North Carolina State University,
Raleigh, NC 27695, USA

F. V. Hebard
Meadowview Research Farm, The American Chestnut Foundation,
Meadowview, VA 24361, USA

S. Anagnostakis
Connecticut Agricultural Experiment Station, New Haven, CT
06504, USA

N. Wheeler · R. R. Sederoff
Forest Biotechnology Group, Department of Forestry and
Environmental Resources, North Carolina State University,
Raleigh, NC 27695, USA

P. H. Sisco
The American Chestnut Foundation, Asheville, NC 28801, USA

Keywords Fagaceae · *Castanea* spp. · Chestnut · Genetic map · Comparative genomics · Disease resistance

Introduction

At the turn of the twentieth century, the American chestnut (*Castanea dentata*) was one of the predominant tree species in the deciduous forests of the eastern USA, estimated at 25 % of the standing timber in those forests (Little 1977; Russell 1987). The American chestnut had great value as a source of tannins for the leather industry and for many wood products, including pulp and paper, timber, and furniture (Buttrick 1915; Anagnostakis 1987). Its rot resistance made it desirable for construction, roofing, railroad ties, and fencing, while its nut production was a source of food for people, livestock, and a large and diverse spectrum of wildlife (Martin et al. 1951; Freinkel 2007). The supply of chestnuts was sufficiently abundant to be a source of trade in many areas (Cameron 2002). An accidental introduction of the fungal pathogen *Cryphonectria parasitica*, first noticed at the Bronx Zoo in 1904 (Merkel 1905), led to widespread devastation of the American chestnut during the first half of the twentieth century (Anagnostakis 1982). Billions of trees were lost and that loss extensively disturbed their ecosystem. Few mature American chestnut trees remain, usually at the margins of the original range.

Within the genus *Castanea*, Chinese chestnut (*Castanea mollissima*), Japanese chestnut (*Castanea crenata*), and Chinese chinkapin (*Castanea henryi*) have substantial levels of resistance to blight (Anagnostakis 1992). Hybrids of Chinese or Japanese chestnut with American chestnut are not as resistant as the Asian parent; however, Burnham (1981) and Burnham et al. (1986) proposed that backcross breeding might be used to introgress resistance into an American chestnut background. Genetic maps using dominant anonymous markers constructed for the parents of a *C. mollissima* × *C. dentata* interspecific hybrid cross were used to identify quantitative trait loci (QTL) for resistance to the pathogen (Kubisiak et al. 1997; Sisco et al. 2005). Kubisiak et al. (1997) proposed a three QTL model that explained about 70 % of the phenotypic variance for resistance to blight. Genetic maps constructed for ecologically diverse parents of the European chestnut, *Castanea sativa* (Casasoli et al. 2001), have been used to identify QTLs for various adaptive traits such as bud flush, growth, and carbon isotope discrimination (Casasoli et al. 2004).

Efforts to introduce resistance from Chinese chestnut into American chestnut by backcross breeding (Hebard 1994, 2006a, b; Diskin et al. 2006) have produced many promising backcross trees. Recent advances in genomics of chestnut raise the possibility of the identification and map-based cloning of disease resistance genes from Chinese chestnut (Wheeler and Sederoff 2009). Similarly, advances in transformation technology (Andrade and Merkle 2005; Polin et al. 2006)

provide the means for transferring genes conferring resistance from Chinese chestnut, or other sources, into American chestnut, with the potential for accelerating genetic improvement and species restoration (Merkle et al. 2007). Application of genomic technology, including the identification of disease resistance genes their marker-based selection in backcross breeding requires a high-resolution genetic linkage map. The construction of such maps for species and hybrids in *Castanea* should facilitate further QTL identification for use in marker-assisted selection for disease resistance and recurrent type, candidate gene selection, and map-based cloning.

The most suitable species for a genomic platform in *Castanea* is *C. mollissima*, given its importance as a source of host resistance to *C. parasitica* (Graves 1950; Clapper 1952; Kubisiak et al. 1997; Diskin et al. 2006). The first focus of our genomic approach was to develop a large set of expressed sequence tags (ESTs) from *C. mollissima* and *C. dentata* by high-throughput 454 sequencing (Barakat et al. 2009, 2012), resulting in a large database of ESTs for *Castanea* (<http://www.fagaceae.org>). Here, we use the assembled ESTs to develop and genetically map simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers. The resulting transcript-based genetic map was generated using two full-sib families of *C. mollissima*. We then used the new *C. mollissima* markers to verify the location of previously identified QTLs for blight resistance and to integrate the *Castanea* consensus genetic map with its physical map (Fang et al. 2012, companion manuscript). Finally, we identified conserved orthologs and surveyed the extent of marker synteny between *C. mollissima*, some related species within Fagaceae, and peach (*Prunus persica*) from the Rosaceae. We included peach for comparison since, of the completely sequenced and assembled tree genomes, it has the closest phylogenetic relationship to chestnut and is the smallest in size (~227 Mb, <http://www.rosaceae.org/peach/genome>, i.e., less than half the size of *Populus trichocarpa* (480 Mb, Tuskan et al. 2006), and only 68 % larger than *Arabidopsis* (135 Mb, AGI 2000)). These attributes facilitate genome comparisons even between families and potentially provide a valuable resource for candidate gene identification. In this regard, we present evidence of a large number of regions with significant segmental homology between the peach and chestnut genomes, accounting for slightly over half of their genetic and physical maps, and a list of candidate genes for chestnut blight resistance.

Materials and methods

Source of ESTs

A total of 25 cDNA libraries were prepared from various tissues of five species in the family Fagaceae—*C. mollissima*, *C. dentata*, *Fagus grandifolia* (American beech),

Quercus rubra (northern red oak), and *Quercus alba* (white oak). EST databases were created primarily by Roche 454 pyrosequencing and a limited amount of Sanger sequencing. A total of 172 Mb of cDNA sequence was obtained from *C. mollissima*. Detailed descriptions of each cDNA library including source species, tissue, sequence type, number of sequence reads, as well as individual EST assemblies are available on the Fagaceae Genomics website (<http://www.fagaceae.org>) and in part from Barakat et al. (2009).

SSR identification, marker development, and detection

EST datasets consisting of sequences from *C. mollissima*, *C. dentata*, *Q. rubra*, and *Q. alba* were combined and assembled, and the consensus sequences searched for SSR motifs. Selected motifs had a minimum of either five di-nucleotide repeats, four tri-nucleotide repeats, three tetra- through hepta-nucleotide repeats, or two octa- or nona-nucleotide repeats. The presence of multiple reads with different numbers of repeats was taken as evidence for a polymorphic SSR. Primer pairs were designed for 455 SSRs using this approach. A second approach used only the *C. mollissima* CCall_Unigene_V2 EST assembly for SSR identification (assembly available at <http://www.fagaceae.org>). Consensus sequences were searched for repeat motifs and evidence for polymorphism was assessed. Using this approach, 492 additional, nonredundant SSRs were selected for primer design. The 947 SSRs were named by sorting the EST contig names from which they were developed and assigning the prefix “CmSI” (Cm=*Castanea mollissima* and SI=Southern Institute of Forest Genetics) followed by a four-digit number identifier (CmSI0001–CmSI0947). We note here and in [Supplemental File 2](#) (“markers-ESTs” tab) that markers CmSI0033 to CmSI0486 are from the first set and markers CmSI0001 to CmSI0032 and CmSI0487 to CmSI0947 are from the second set. Motifs reported utilize their alphabetic minimum (Jin et al. 1994; Echt and May-Marquardt 1997).

To reduce the costs associated with primer screening and to increase post-PCR multiplexing flexibility and capacity, an M13-specific sequence (5'-CACGACGTTGTAACGAC-3') was added to the 5' end of each forward primer as described by Schuelke (2000). To favor 3' adenylation of the forward amplified strand, all reverse primers were PIG-tailed with a 7-base sequence (5'-GTTTCTT-3') on the 5' end (Brownstein et al. 1996). For fluorescent detection, three-primer PCR was performed, which included a 5' dye-labeled M13-specific primer (same sequence as the M13 “tail” described above). PCR mixtures consisted of the following in 6 μ L total volume: 2.5 ng of template DNA, 0.16 μ M 5'-dye-labeled M13 primer DNA, 0.04 μ M of 5'-M13-tailed forward primer, 0.16 μ M of reverse PIG-tailed primer, 66 μ M of dNTPs, 0.6 μ L 10 \times *Taq* DNA polymerase reaction buffer (500 mM KCl, 100 mM Tris-HCl, 1.0 % Triton X-100, 15 mM MgCl₂), and 1.0 U of Hotstart *Taq* DNA polymerase. Reactions were loaded in 384-well microtiter

plates, covered with Mylar film, and PCRs run using MJ Research PTC-200 or PTC-225 thermal cyclers. The programmed thermal profiles were 4 min at 95 $^{\circ}$ C; 35 cycles of 20 s at 92 $^{\circ}$ C, 20 s at 55 $^{\circ}$ C, 20 s at 72 $^{\circ}$ C; 7 min at 72 $^{\circ}$ C; indefinite hold at 4 $^{\circ}$ C. Completed reactions were diluted with distilled water and 1 μ L was analyzed on an ABI PRISM 3130xl or ABI PRISM 3730xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. Allele sizes were determined using the LIZ600 internal size standard and the global southern algorithm implemented by ABI PRISM GeneMapper software version 3.7 (Applied Biosystems). Alleles were named according to Deemer and Nelson (2010) using the three parents (see below) as reference samples and alleles.

SNP identification, development, and detection

SNPs were identified using the *C. mollissima* CCall_Unigene_V2 assembly. PolyBayes v3.0 (Marth et al. 1999) was run on each contig to identify SNPs and compute SNP probability scores. Polymorphisms due to single base insertions and/or deletions were excluded, as were SNPs with probability scores <0.70, resulting in 25,904 SNPs. These SNPs were sent to Illumina (San Diego, CA, USA) for scoring with their in-house software and were further filtered using a cut-off of 0.70 for the Illumina quality score, yielding 21,390 SNPs for further consideration. A final set of 1,536 SNPs was selected for the GoldenGate BeadArray (Illumina) based on three factors: (1) first priority was given SNPs originating from unigene contigs found to be differentially expressed (Barakat et al. 2009); (2) second priority was given SNPs with the highest PolyBayes probability scores; and (3) only one SNP per contig was selected. A total of 205 SNPs met all three factors, while the remaining 1,421 SNPs met the second and third factors. Each SNP marker was named by first sorting the Ccall_Unigene_V2 contig names from which they were developed and then assigning the prefix “CmSNP” (CmSNP=*Castanea mollissima* SNP) followed by a five-digit identifier (CmSNP00001–CmSNP01536).

SNPs were interrogated using the GoldenGate BeadArray platform and automatically clustered, genotypes called, and confidence scores assigned using GenomeStudio software v2008.1 (Illumina). Although automated clustering using GenomeStudio generally produced one, two, or three distinct clusters corresponding to the expected genotypic classes based on parental genotypes, the data for all SNPs were inspected manually, and genotypic clusters were manually edited when necessary (Yan et al. 2010). Genotypes ambiguously located between clusters were scored as missing data.

C. mollissima mapping populations, plant material, and DNA extraction

Two *C. mollissima* full-sib families were used for genetic map construction. Both families [‘Mahogany’ $\text{\textcircled{f}}$ \times ‘Nanking’ $\text{\textcircled{m}}$ (M

$\times N$) and ‘Vanuxem’ ♀ \times ‘Nanking’ ♂ ($V \times N$)] were derived from controlled pollination between three *C. mollissima* cultivars being used as sources of resistance in The American Chestnut Foundation’s (TACF) backcross breeding program (Hebard 1995). DNA samples were extracted from young leaves using a CTAB-based protocol modified for use on a mixer mill (refer to Supplemental File 1). Marker segregation data were collected for a total of 179 progeny of the $M \times N$ family and 158 progeny of the $V \times N$ family.

Linkage mapping and QTL analysis

Linkage analyses were performed with JoinMap v3.0 (van Ooijen and Voorrips 2001). Data were coded separately for each parent in the $M \times N$ and $V \times N$ mapping families (script provided by C.D.N.). The four datasets were loaded into a single JoinMap session. Markers in each dataset exhibiting excessive segregation distortion ($P < 0.005$) were excluded from all further analyses. Maps were first constructed separately for each parent. For each dataset, linkage groups were established at log of the odds (LOD) 5.0. Syntenic groups were identified and combined using the JoinMap “Combine Groups for Map Integration” function. Linkage maps were calculated using default mapping parameters, i.e., all linkages with recombination estimates smaller than 0.4 and a LOD larger than 1.0 were used to determine marker orders. Map distances were calculated using the Kosambi mapping function. Only markers that could be placed during the first two rounds of JoinMap mapping, i.e., those that did not exhibit poor goodness-of-fit (χ^2 values > 5.0) or result in negative map distances, are reported in the final map. Prior to integrating maps, differences in recombination frequencies among shared markers were tested within JoinMap. Map graphics were generated with MapChart v2.1 (Voorrips 2002).

In order to correlate the *C. mollissima* EST-based genetic marker framework with a previous *C. mollissima* \times *C. dentata* F_2 -based map (Kubisiak et al. 1997; Sisco et al. 2005), the same F_2 mapping population ($n=89$) was genotyped using the 1,536 SNP GoldenGate array. The resulting SNP data were combined with the previously collected marker data consisting of 170 RAPDs, 12 RFLPs, 2 isozymes, and 16 genomic SSRs. A consensus genetic map was then constructed for the F_1 parent trees as described above. Linkage groups were named according to Kubisiak et al. (1997). The *C. mollissima* and *C. mollissima* \times *C. dentata* parental maps were aligned using the “Show Homologs” option available in MapChart v2.1. Using the new genetic maps, QTL mapping for blight resistance was recalculated using the F_2 genotypic and phenotypic data. QTL analyses were performed separately for each F_1 parent tree using both MapQTL v5 (van Ooijen 2004) and PLABQTL v1.2 (Utz and Melchinger 1994) and a set of 10 disease

metrics. All metrics were simple functions of the length and/or width of cankers induced by artificial inoculation of F_2 trees with two different isolates of *C. parasitica* (Ep155 and SG2-3), measured at 2 and 3 months post-inoculation (Kubisiak et al. 1997). Each of the 10 metrics was investigated using nonparametric analysis (Lehmann 1975), interval mapping (Lander and Botstein 1989; Haley and Knott 1992), and composite interval mapping (Jansen and Stam 1994; Utz and Melchinger 1994; Zeng 1994).

For composite interval mapping, the presence of putative QTLs was initially investigated using pre-selected marker cofactors (refer to the “cov SELECT” command in PLABQTL). In order to determine the 5 % genome-wide error rate for declaring significance of QTL, a permutation test was run that consisted of a minimum of 1,000 randomizations (Churchill and Doerge 1994). The most significant QTL interval, based on LOD peak height, was then identified and fixed as a cofactor. Genome scans were performed for each metric and the next most significant QTL was identified. This QTL was then fixed as an additional cofactor; permutation tests were again performed, followed by genome scans for additional significant QTL. This process was repeated until no additional QTL were declared significant.

Comparative mapping of the *Castanea* and *Prunus* genomes

BLASTn analyses were used to compare the genetic map for *C. mollissima* with the reference genome sequence of peach (*P. persica*). The peach genome (v1.0) scaffolds were downloaded from the Genome Database for Rosaceae website (<http://www.rosaceae.org>). Ungapped BLASTn v2.2.24+ analyses were performed using default criteria. Only alignments with e values $\leq 1.0e^{-40}$ and a greater than 80 % nucleic acid sequence identity were considered for comparative analysis. The *C. mollissima* EST contigs with only one significant alignment to the peach genome were considered for comparative purposes. The identification of putative segmental homologs between genomes was based on shared sequence identity and shared order, i.e., synteny and collinearity. Two-dimensional scatter plots, where the Y -axis represented marker position along the *C. mollissima* linkage group and the X -axis represented the best hit locations in the peach genome, were used to visually inspect marker order and collinearity.

Visual inspection of the comparative data was followed by statistical analysis with the software packages FISH (Calabrese et al. 2003) and LineUp (Hampson et al. 2003). For analysis using FISH, the minimum block size was set to 4 and the minimum score was set to 0, and all other parameters were set to default. For analysis using LineUp, a FastRun algorithm with a minimum run length set to 4 and distance set to 2 was used. The significance of segmental

homology was assessed by re-running the segment detection algorithm on 10,000 uniformly randomized gene maps. Finally, the markers in the QTL regions were manually examined to expand and refine the alignments to the peach genome assembly. Specifically, the QTL markers were queried against the peach genome with BLASTn using an e value cutoff $1.0e^{-10}$. If the marker's best match to peach was in the same peach region as other nearby QTL marker matches, then the new marker was included in the putative homologous segments. The genes from the putatively homologous regions from peach were collected and processed with Blast2GO in order to assign functions and GO terms (Götz et al. 2008).

Results

Identification and characterization of EST-derived SSR markers

SSR frequency in *C. mollissima* was assessed using the CCall_Unigene_V2 assembly available on the Fagaceae Genomics website (<http://www.fagaceae.org>). This assembly consisted of 838,472 454 pyrosequencing titanium reads and 9,480 Sanger reads which assembled into 48,335 EST unigene contigs. A total of 12,539 SSRs were identified in 8,737 unique ESTs. The most frequent SSR motif was trimeric (5,271 or 42 % of detected SSRs), followed by dimeric (4,793 or 38.2 % of detected SSRs), tetrameric (1,816 or 14.5 % of detected SSRs), and pentameric (659 or 5.3 % of detected SSRs). The most frequent di-nucleotide motif (alphabetic minimum) was AG (71 %), followed by AT (15 %), AC (14 %), and finally by CG (<0.1 %).

Of the 455 SSRs selected using a combined dataset of ESTs from several Fagaceae species, 241 (53 %) amplified a PCR product from *C. mollissima* DNA and 69 (15 %) of these were found to amplify a single polymorphic locus that was heterozygous in at least one of the three mapping parents. Of the 492 nonredundant SSRs identified using only EST sequences from *C. mollissima*, 345 (70 %) amplified a PCR product from *C. mollissima* DNA and 261 (53 %) were found to amplify a single polymorphic locus that was heterozygous in at least one of the three mapping parents. These 330 markers, consisting of 90 di-, 195 tri-, 26 tetra-, 11 penta-, and 8 hexa-nucleotide SSRs, were chosen for mapping. Of these SSRs, 78 (25.7 %) were heterozygous in all three mapping parents, 144 (43.6 %) were heterozygous in two parents, and the remaining 108 (32.7 %) markers were heterozygous in only one parent. Marker ID, GenBank ID, forward primer, reverse primer, motif, average allele size, gene diversity, heterozygosity, polymorphism information content (PIC), and whether null alleles were observed are reported for the 330 SSRs in [Supplemental](#)

[File 2](#) (*C. mollissima* SSRs tab). Markers identified using EST assemblies developed from across the Fagaceae genera should be useful for wider comparative analyses (within the inclusive range of markers from CmSI0032 to CmSI0486).

Identification and characterization of EST-derived SNP markers

SNP frequency in *C. mollissima* was assessed using the CCall_Unigene_V2 assembly. A total of 25,904 SNPs with a PolyBayes probability score >0.70 were identified in 8,890 unique EST contigs. The depth of the reads for any one SNP varied from 2 to 7,952, with a mean and median depth of 16.9 and 7 reads, respectively. The most frequent SNPs were C↔T transitions and A↔G transitions (8,195 and 8,193, respectively, or 31.6 % each of detected SNPs), followed by A↔T transversions (3,452 or 13.3 %), A↔C transversions (2,196 or 8.5 %), G↔T transversions (2,160 or 8.3 %), and C↔G transversions (1,462 or 5.6 %), with the remaining SNPs (247 or 1.0 %) being characterized by more than a single base change. A change in two adjacent bases is not a SNP strictly speaking, but these mutations are included in our analysis. As mentioned, after further culling based on Illumina's quality score (>0.70), a total of 21,390 SNPs were available for developing genotyping assays, and following prioritization (see "Materials and methods"), a final set of 1,536 SNPs was selected for the GoldeGate BeadArray. Within this final set of SNPs, 205 were represented among the 337 unigenes that Barakat et al. (2009) had reported as being differentially expressed between American chestnut canker tissue and Chinese chestnut canker tissue or between Chinese chestnut healthy stem and Chinese chestnut canker tissue.

Of the 1,536 SNPs chosen for array design, 213 (14 %) could not be scored due to poor cluster separation in the GoldenGate analysis (<http://dnatech.genomecenter.ucdavis.edu/illumina.html>). An additional 252 (16.4 %) were found to produce discrete clusters, but the markers were homozygous in all three mapping parents. Although the majority of these SNPs were monomorphic, being fixed for the same allele, 26 (10.3 %) were fixed for alternate alleles among parents. In total, 1,071 SNPs were found to be heterozygous in at least one of the three mapping parents (i.e., mappable), with 106 (9.9 %) heterozygous in all three parents, 414 (38.7 %) heterozygous in two parents, and the remaining 551 (51.4 %) heterozygous in one parent. Of these 1,071 SNPs, 362 (62 %, i.e., percent mappable or conversion rate) consisted of C↔T transitions, 355 (83 %) A↔G transitions, 112 (76 %) A↔T transversions, 94 (90 %) G↔T transversions, 83 (54 %) A↔C transversions, and 65 (54 %) C↔G transversions. We used a chi-square test to evaluate conversion rates for SNPs by mutation (transitions vs. transversion) and polymorphism (A/C vs.

A/G vs. A/T vs. C/G vs. C/T vs. G/T) types and found mutation type not significant ($P>0.05$) while polymorphism type was highly significant ($P<0.001$). Marker ID, GenBank ID, sequence context, gene diversity, heterozygosity, PIC, and observed alleles are reported for the 1,071 SNPs in [Supplemental File 2](#) (*C. mollissima* SNPs tab).

Construction and analysis of genetic linkage maps

Segregation data for 1,401 markers (330 SSRs and 1,071 SNPs) developed from 1,356 unique EST contigs were used for genetic mapping. In the $M \times N$ population, 539 markers were heterozygous in ‘Mahogany’ and 1,092 in ‘Nanking’, while in the $V \times N$ population, 590 markers were heterozygous in ‘Vanuxem’ and 1,088 in ‘Nanking’. Segregation data for all markers, coded separately for each parent, are provided in [Supplemental File 3](#). Allele calls of SSR and SNP markers for each parent are provided in [Supplemental File 4](#). Only 1.0–2.1 % of the alleles genotyped in the four datasets were missing. A majority of the missing data can be explained by null alleles (i.e., apparent mutation in primer sequence causing failure to PCR amplify) at SSR loci where the allelic configuration in the parents resulted in ambiguous progeny genotypes. Initially, maps were constructed for each parent separately. Across the four datasets, segregation data for 20 markers were eliminated from further analyses as the markers were significantly distorted ($P<0.005$) from their expected segregation ratios in parents of both crosses.

For each dataset, linkage groups were established using a two-point LOD threshold of 5.0 to obtain the 12 groups expected from karyotype ($2n=2x=24$) analyses (Jaynes 1962; Islam-Faridi et al. 2009). At this LOD threshold and considering all maps, only one marker (CmSI0518) remained unlinked. In addition, 224 markers were excluded from the final consensus map because they either had poor goodness-of-fit values or introduced negative map distances in the third round of marker ordering. Although these markers were not placed on the final map, their most likely map interval is provided in [Supplemental File 5](#) (along with positions for all mapped markers). The consensus map consists of 12 linkage groups with 1,156 markers (250 SSRs and 906 SNPs) mapping to 975 loci (i.e., discrete centimorgan positions) spanning 742.4 cM, with an average locus spacing of 0.7 cM ([Supplemental File 6](#)). The linkage groups ranged in size from 50.6 to 90.4 cM, with an average size of 61.8 cM.

Segregation data for 447 SNPs were suitable for genetic mapping in the interspecific (*C. mollissima* \times *C. dentata*) F_2 population. As noted previously (Kubisiak et al. 1997; Sisco et al. 2005), significant segregation distortion was observed in this cross, with 18.1 % of all SNPs being distorted at $P<0.005$. Distorted markers were excluded from all further analyses. Segregation data for 566 markers (366 SNPs, 170 RAPDs, 12

RFLPs, 2 isozymes, and 16 genomic SSRs) were then used for map construction. Initially, maps were constructed separately for each parent. Of the 566 markers used for genetic map construction, only two SNPs were unlinked to any other markers at LOD 5.0. An additional 44 markers were excluded from the final consensus genetic map because they either had poor goodness-of-fit values or introduced negative map distances in the third round of marker ordering. The interspecific F_2 -based consensus map consisted of 12 linkage groups with 520 mapped markers spanning 685.7 cM, somewhat less than the *C. mollissima* consensus map of 742 cM. The linkage groups ranged in size from 30.3 to 84.7 cM, with an average size of 57.1 cM. Alignments of the *C. mollissima* and interspecific F_2 maps are shown in [Supplemental File 6](#).

Composite interval mapping identified three significant QTL for resistance to *C. parasitica* across the various cancer metrics (Table 1), one each on linkage groups B, F, and G (*Cbr1*, *Cbr2*, and *Cbr3*, respectively). The results of QTL mapping using alternative statistical approaches of nonparametric analysis and interval mapping were similar and consistent. The new map, with higher density and increased coverage, produced results similar to the earlier map (Kubisiak et al. 1997). For all three QTLs (*Cbr1*, *Cbr2*, and *Cbr3*), alleles conferring resistance were inherited from Chinese chestnut and LOD ± 1.0 support intervals were localized to regions less than 10 cM on the consensus genetic map (Table 1). Because of the increased density of EST-based SSR and SNP markers on the consensus map, many more ESTs can now be located within these QTL intervals. These additional markers can be used as a bridge to the physical map (Fang et al. 2012, [companion manuscript](#)) producing a means to obtain complete genome sequence data across these QTL.

Comparison of *C. mollissima* genetic map to the *Prunus* genome

Based on our BLASTn threshold criteria, 510 (46 %) of the mapped chestnut EST contigs had only one significant hit to the peach genome and hence were useful for comparative purposes ([Supplemental File 7](#)). Comparisons of the order of the EST contigs on each of the chestnut linkage groups to the order of putative orthologs in the peach genome can be visualized as two-dimensional scatterplots ([Supplemental File 7](#), tabs Graphic LG A–Graphic LG L). Regions of collinearity (i.e., potential regions of segmental homology) can easily be identified as diagonal lines. Similar segmental homologies were identified between chestnut and peach based on statistical analysis using FISH and LineUp. A total of 37 significant segmental homologous regions were identified between chestnut and peach using LineUp and 28 of these were verified by FISH (Table 2). Considering all 37 putative homologies (determined by LineUp), the average

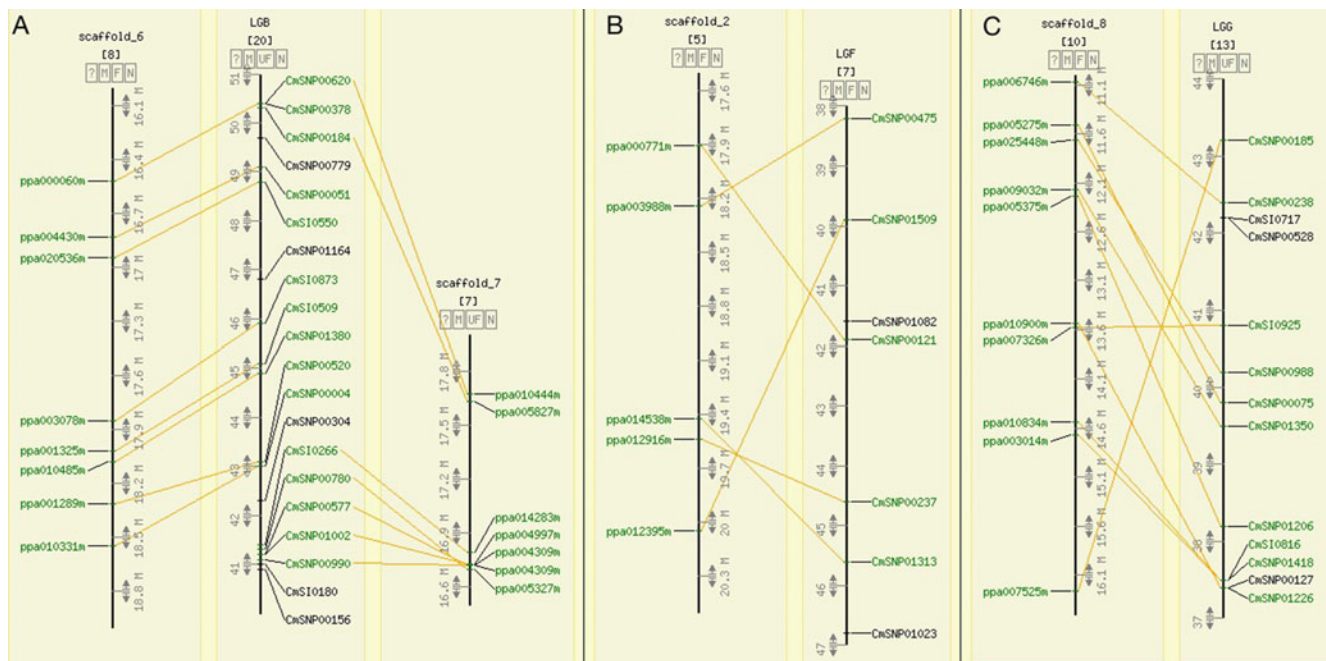


Fig. 1 Syntenic regions in chestnut and peach using the Chinese chestnut genetic map for three chestnut blight resistance QTL regions and the peach genome assembly. The chestnut regions are labeled **a** LGB, **b** LBF, and **c** LGG for linkage groups B, F, and G which

correspond to QTL regions *Cbr1*, *Cbr2*, and *Cbr3*, respectively (see text for details). The corresponding peach genome assemblies include **a** scaffold_6 and scaffold_7, **b** scaffold_2, and **c** scaffold_8 (see text for details), respectively

homologous segment contains nine loci covering 12.1 cM on the chestnut genetic map and 4.87 Mb in the peach genome. The combined segments span 422.5 cM (~57 %) of the chestnut genetic map and 131.8 Mb (~58 %) of the peach genome. The largest segment of significant collinearity is a region composed of 25 ESTs on chestnut linkage group D spanning 38.9 cM and 4.68 Mb on peach scaffold_5 and is significant at $P < 0.001$.

Marker loci associated with QTL for resistance to chestnut blight on linkage groups B, F, and G (*Cbr1*, *Cbr2*, and

Cbr3, respectively) were used to search the peach genome for orthologous sequences. Manual curation of the BLASTn results yielded more and longer putatively homologous segments between peach and chestnut in the QTL regions than software predictions alone. The QTL on linkage group B (*Cbr1*) matches two peach scaffolds (Fig. 1a). Fifteen of the 20 markers within this QTL match scaffold_6 or scaffold_7 in peach; the five markers without a match do not show sequence similarity to any peach scaffold. The pattern of matches suggests that a translocation may have occurred in

Table 1 Results of composite interval mapping for QTLs conferring resistance to *C. parasitica* in a Chinese chestnut × American chestnut F₂ population ($n=89$)

QTL ^a	Chestnut linkage group	Chestnut linkage interval (cM) ^b	Blight resistance metrics ^c	LOD scores ^d	Partial R ² (%) ^e	Homologous peach scaffold	Powdery mildew QTL ^f
<i>Cbr1</i>	B	40.9–50.4	4	3.45–5.10	12.65–21.73	7 (16.698–16.800 Mb) 6 (16.516–18.550 Mb) 7 (17.635–17.675 Mb)	G6 (16.0–22.94 Mb)
<i>Cbr2</i>	F	38.1–46.8	2	3.18–3.51	13.04–16.79	2 (17.905–20.049 Mb)	G2
<i>Cbr3</i>	G	35.7–39.5	3	3.24–5.06	13.45–23.65	8 (11.068–16.265 Mb)	G8 (11.48–17.21 Mb)

^a Chestnut blight resistance QTLs named in Kubisiak et al. (1997)

^b Most likely position of QTL in centimorgan (Kosambi) defined as a LOD ±1.0 support interval

^c Number of blight resistance metrics for which QTL was declared significant

^d Range of LODs observed for QTL across various metrics

^e Range of partial R², % variation explained by QTL across various metrics

^f Peach linkage groups for powdery mildew resistance QTLs as delineated by Foulongne et al. (2003)

Table 2 Significant homologous segments identified between the chestnut genetic map and the peach genome assembly

No.	Chestnut linkage group	Position on chestnut linkage group (cM)	Peach scaffold	Position on peach scaffold (Mb)	# of loci	Length of segment—chestnut (cM); peach (Mb)	LineUp % significance	Identified by FISH
1	A	32.7–39.7	1	32.35–44.47	7	7.0; 12.13	4.6	No
2	A	60.4–68.1	1	5.69–7.70	8	7.7; 2.01	0	Yes
3	A	70.1–88.0	1	29.06–43.19	7	17.9; 14.14	3.5	No
4	A	49.5–56.6	3	19.77–21.86	9	7.1; 2.08	0.2	Yes
5	A	47.6–60.7	7	17.97–21.83	9	13.1; 3.86	0.2	Yes
6	B	12.3–20.2	1	1.01–1.48	5	7.9; 0.47	0	Yes
7	B	23.1–31.8	1	2.18–11.72	6	8.7; 9.55	0	No
8	B ^a	38.6–50.4	6	16.52–21.54	11	11.8; 5.02	0	Yes
9	B	41.1–66.2	7	16.70–18.54	15	25.1; 1.84	0	Yes
10	C	25.0–29.1	1	30.43–31.44	6	4.1; 1.01	4.6	Yes
11	C	29.9–55.6	1	21.46–28.17	14	25.7; 6.71	0	Yes
12	C	0–24.7	6	25.54–28.21	12	24.7; 2.67	0	Yes
13	D	8.7–11.9	1	17.80–20.07	4	3.2; 2.27	1.5	Yes
14	D	16.9–55.8	5	13.10–17.78	25	38.9; 4.68	0	Yes
15	E	6.1–29.9	3	1.40–6.62	19	23.8; 5.22	0	Yes
16	E	29.4–36.4	4	2.29–4.50	11	7.0; 2.21	0	Yes
17	E	37.7–47.5	4	0.44–1.70	6	9.8; 1.26	0	Yes
18	F	13.9–26.6	4	8.97–17.04	13	12.7; 8.07	0	Yes
19	F	0–9.7	6	2.56–3.22	7	9.7; 0.67	0	Yes
20	G	24.0–27.6	1	0.19–5.54	5	3.6; 5.35	0	No
21	G ^a	35.7–39.5	8	12.17–15.95	10	3.8; 3.77	0	Yes
22	H	41.1–60.6	1	29.60–31.78	8	19.5; 2.19	0.1	Yes
23	H	7.0–14.0	6	0.25–13.48	10	7.0; 13.23	0	No
24	H	13.8–22.3	6	0.79–13.48	10	8.5; 12.69	0	No
25	H	28.8–40.4	6	0.96–4.08	7	11.6; 3.12	0.1	No
26	I	2.4–14.2	6	21.75–22.85	10	11.8; 1.10	0.1	Yes
27	I	17.4–21.5	6	23.49–25.42	8	4.1; 1.92	2.3	Yes
28	I	16.6–25.7	7	0.98–13.61	12	9.1; 12.63	0	Yes
29	I	33.6–48.8	7	7.71–9.48	7	15.2; 1.77	0	No
30	J	0–17.4	5	1.69–3.51	6	17.4; 1.82	0	Yes
31	J	27.4–40.1	5	4.00–11.00	12	12.7; 6.99	0	Yes
32	J	36.3–56.7	8	19.84–21.43	10	20.4; 1.60	0.5	Yes
33	K	20.9–25.0	2	8.28–19.22	6	4.1; 10.94	1.4	No
34	K	28.1–32.6	2	14.15–15.79	5	4.5; 1.64	0.6	Yes
35	K	48.6–59.6	2	9.52–10.28	6	11.0; 0.76	0	Yes
36	L	16.0–29.8	2	24.09–26.24	18	13.8; 2.14	0	Yes
37	L	31.7–35.6	8	17.65–19.55	5	4.0; 1.90	4.3	Yes

^a Homologous segments with presumably overlapping QTLs for resistance to blight disease in chestnut and powdery mildew disease in peach (see text and Table 1)

this region. Five of the QTL markers have matches spanning from 16.7 to 16.8 Mb on scaffold_7 in peach. This region of peach has 18 genes that are likely to be retained in the same position in chestnut. Eight of the QTL markers have matches to scaffold_6 across more than 2 Mb (from 16.52 to 18.55 Mb). This region in peach has been annotated with 191 genes. Two markers match a region on scaffold_7 from

17.6 to 17.7 Mb encompassing six candidate genes. Other genes around these three regions in peach are possibly conserved in the QTL region as well, but the breakpoints of the rearranged segments cannot be more accurately determined.

The QTL identified on linkage group F (*Cbr2*) did not yield a corresponding segment in peach using the LineUp

software, but manual curation yielded a pattern of five markers with best matches to the peach genome in scaffold_2, all of which fall into a 2.14-Mb region (Fig. 1b). The other two markers of the seven markers defining this QTL have matches to other scaffolds, indicating that the homologous region found in peach may not span the entire QTL. While these matches are less convincing evidence for homologous segments than other regions of the genetic map, the five clustered matches to scaffold_2 indicate that the peach segment is worth further examination for candidate genes. This region has 309 genes in peach, and just under 20 % of it has been annotated as repetitive DNA. Linkage group G (containing the *Cbr3* QTL) has a strong homology to peach scaffold_8 with 10 of 13 markers anchored to this region (Fig. 1c). The markers span 5.2 Mb of scaffold_8, a large segment containing 776 genes and 21.1 % repetitive content. The three markers within the *Cbr3* QTL that did not have a match on scaffold_8 did not show strong sequence similarity to any other linkage group in peach.

Based on manual inspection of chestnut–peach homologous segments (Supplemental File 8), two of the three QTLs associated with blight resistance (linkage group B, *Cbr1*, and linkage group G, *Cbr3*) in chestnut were located in regions of segmental homology with peach that contain two major QTLs for resistance to powdery mildew disease (caused the fungal pathogen *Podosphaera pannosa* var. *persicae*) (Dirlewanger et al. 1996; Foulongne et al. 2003; Pascal et al., 2010). A third QTL for powdery mildew resistance in peach was reported on *Prunus* linkage group 2 (G2), which showed segmental homology with *Cbr2* (*Castanea* linkage group F). Delineating this segment in the peach genome was problematic due to inconsistency of the G2 QTL intervals across progenies and a discrepancy in the inferred parental origin of the resistance allele (Foulongne et al. 2003). In the following, we focus on the comparative results of the two major fungal resistance QTL intervals on *Prunus* G6 and G8 corresponding to *Castanea* QTL *Cbr1* and *Cbr3*, respectively.

The chestnut QTL *Cbr1* covering 9.5 cM has putative segmental homology to about 2 Mb on peach scaffold_6 (16.5 to 18.6 Mb). The powdery mildew resistance QTL is defined by peach scaffold_6 markers AG26, pchcms5, and PC73, located at 17.58, 19.16, and 22.94 Mb, respectively. Marker AA9-1.6, which defines the lower bound of this QTL interval in the peach genetic map (on G6), is not sequence-based and has not been located within the genomic sequence, but its location has been inferred from genetic mapping to be at about 16 Mb. Based on these marker locations, the powdery mildew resistance QTL on peach scaffold_6 spans 7 Mb, encompassing the homologous *Cbr1* region in chestnut. Similarly, the region containing resistance locus *Cbr3* is characterized by a segmental homologous region containing the powdery mildew resistance QTL on peach scaffold_8.

Chestnut linkage group G from 35.7 to 39.5 cM aligns to 5.2 Mb on peach scaffold_8 from 11.01 to 16.27 Mb. This region of scaffold_8 corresponds very well with the support interval defined by markers FG230 and FG37 (on G8), which are located at approximately 11.48 and 17.21 Mb, respectively. These results indicate that genes conferring resistance to unrelated pathogens may be the same or clustered and that comparative genomic approaches can help to identify candidate resistance genes.

Building on the assumption of chestnut–peach homology, markers spanning the three blight resistance QTLs in chestnut were used for mining homologous genomic regions in peach. *Cbr1* and *Cbr2* corresponded to regions of peach with 215 and 309 genes, respectively. *Cbr3*, the largest QTL region, also had the largest region of homology to peach, encompassing 776 genes. Blast2GO was able to assign GO terms to 1,140 of the 1,300 peach genes (Supplemental File 9). Of particular interest, 15, 21, and 59 genes from the three QTLs (*Cbr1*, *Cbr2*, and *Cbr3*, respectively) were annotated with “response to stress” (GO:0006952) encompassing a variety of different stress response functions. Another term, “response to biotic stimulus” (GO:0009607) was annotated to 5, 2, and 21 genes, respectively. A total of 24 genes can be identified that are located in one of the three QTL intervals and have both GO terms (Table 3). These genes can be further tested and utilized in candidate gene approaches for uncovering the molecular basis of blight resistance in chestnut.

Discussion

Marker identification and characterization

Small numbers of SSR markers had been developed previously for several chestnut species including *C. crenata* (Tanaka et al. 1999), *C. sativa* (Buck et al. 2003; Marinoni et al. 2003; Gobbin et al. 2007), and *C. mollissima* (Inoue et al. 2009). The frequencies of SSR types and motifs that we found in *C. mollissima* ESTs are in general agreement with those reported for other species in the Fagaceae (Barreneche et al. 2004; Ueno et al. 2009; Inoue et al. 2009; Cheng and Huang 2009; Durand et al. 2010). Based on our SSR search criteria, 18.1 % of the 48,335 *C. mollissima* CCall_Unigene_V2 ESTs contained at least one SSR. A search for SSRs in *Castanopsis sieboldii* ESTs (Ueno et al. 2009) found 13 % to contain at least one SSR. Similarly, 18.6 % of *Quercus* ESTs (Durand et al. 2010) contained at least one SSR. A search for SSRs in EST datasets available for four additional species in the Fagaceae (*C. dentata*, *Q. rubra*, *Q. alba*, and *F. grandifolia*: <http://www.fagaceae.org>) showed the frequency of SSR-containing ESTs to vary from 11 % in *F. grandifolia* to 16 % in *Q. rubra*. Similar to *Quercus* (Durand et al. 2010), the most frequent EST-SSR types in

C. mollissima were tri-nucleotides (42 %) followed by dinucleotides (38 %). Here we evaluated 947 EST-based SSR primer pairs and found 330 of them to be scoreable and polymorphic (35 %) in at least one of two *C. mollissima* full-sib families involving three parents. Of these 330 SSRs, 250 were confidently placed on the consensus genetic map and the remaining 79 were assigned to linkage groups. Of those 492 SSRs that were developed from *C. mollissima* EST contigs only, that is, (CmSI0001–CmSI0032 and CmSI0487–CmSI0947), a higher proportion was converted to mapped markers (260 of 492 or 53 %) compared to those developed from the multi-species EST contigs (69 of 455 or 15 %).

SSR identification typically requires fragment size identification, while SNP genotyping is amenable to a variety of higher throughput platforms. The available platforms have made it possible to carry out SNP genotyping for thousands of markers in months rather than years. The combination of high-throughput sequencing and genotyping can now reduce the time needed to produce maps to a fraction of what was required a few years ago. More than 30 different SNP detection methods have been developed and applied in different species, and several high-density platforms are now available (reviewed in Gupta et al. 2008). The Illumina GoldenGate BeadArray is a medium-density genotyping platform that can interrogate up to 1,536 SNPs per array. The GoldenGate technology is now being used for genetic analysis in several crop species including barley (Rostoks et al. 2006), soybean (Hyten et al. 2008), and maize (Yan et al. 2009; Yan et al. 2010) where the rates of successful scoring of SNP data were ≥ 90 % (Hyten et al. 2008; Rostoks et al. 2006; Yan et al. 2010). We found 1,064 of 1,536 tested SNPs (69 %) to be scoreable and mappable in at least one of two *C. mollissima* mapping families. Of these SNPs, 906 were confidently placed on the consensus genetic map and an additional 158 were assigned to linkage groups. Future SNP selection efforts may focus on A/C and A/G SNPs, as we found those to provide significantly higher conversion efficiencies. We also note that these two SNP types utilize only one bead type in Illumina's Infinium genotyping technology, making them even more efficient on this high-density platform. In the same way, avoiding C/G SNPs is advised given their low conversion efficiency and their requirement for two bead types on the Infinium platform.

Genetic maps and comparative analysis

The new parent-specific genetic maps and the consensus map developed here for *C. mollissima* represent a significant advance over previous maps for *Castanea* spp. (Kubisiak et al. 1997; Casasoli et al. 2001; Sisco et al. 2005) and a substantial advance in Fagaceae genomics. The previous *Castanea* genetic maps, largely composed of anonymous genetic markers (e.g.,

RAPDs, ISSRs, and AFLPs), were limited in their usefulness for comparative genomics and applications in molecular breeding. Advantages of the new maps include their higher densities (consensus map has 1,156 mapped markers, located at 975 map positions) and resolution (consensus map distances based on 158 and 179 progeny in two full-sib families) and increased sequence specificity (i.e., SSRs and SNPs) of the markers. These improvements allow for integration with physical maps (Fang et al. 2012, companion manuscript) and genome sequences as well as more informative comparative genomic analyses and molecular breeding applications.

The interspecific (*C. mollissima* \times *C. dentata*) F_2 map (F_1 parents, F_2 mapping population) was significantly improved over the original map (Kubisiak et al. 1997) with the addition of 447 SNPs. The revised map covers an additional 29 % in centimorgan distance (685.7 vs. 530.1) than the earlier estimate and contains an additional $2.8\times$ (520 vs. 184) number of markers. The average spacing is about 1 marker/1.5 cM vs. 1 marker/4.4 cM for the original map. However, the resolution of this higher density map remains the same as it is based on the same set of meioses (DNA samples) as the original population. The original map included part of linkage group E fused with linkage group B (Supplemental File 6), while the new map clearly separates these linkage groups and helps to resolve the absence of one of the 12 expected groups. It has been proposed that the absence of a linkage group in the earlier mapping study could be due to large structural genomic rearrangements (e.g., reciprocal translocations) between *C. mollissima* and *C. dentata* (Kubisiak et al. 1997; Islam-Faridi et al. 2008). The frequency of rearrangements between closely related species is high and has led to the hypothesis that rearrangements have a role in speciation (Rieseberg 2001). These data, clearly resolving linkage groups B and E, and new genetic mapping data for *C. dentata* based on a large full-sib family show no indication of a chromosomal translocation compared to *C. mollissima* (B.A.O. unpublished results), although cytogenetic verification is still needed. In addition, the new map based on the interspecific F_2 mapping population validates the three QTL model for blight resistance (*Cbr1*, *Cbr2*, and *Cbr3*) and better defines each QTL by providing many more sequence-specific genetic markers.

We found remarkable regions of synteny defined as segmental homology between the *C. mollissima* genetic map and the *P. persica* genome assembly. About 57 % of the chestnut genetic map in centimorgans could be assigned to a similar proportion of the peach genome in megabases. Moreover, careful manual curation in regions of interest yielded important extensions of this segmental homology. This high degree of homology will support comparative candidate genetic/genomic approaches (with peach and other Rosaceae species) in identifying the molecular networks involved in the chestnut–*C. parasitica* interaction. For

Table 3 Peach genes and their functional annotations that have homology to chestnut genes that are located within one of the three chestnut blight resistance QTL intervals (*Cbr1*, *Cbr2*, or *Cbr3*) and have GO terms “response to biotic stimulus” and “response to stress”

Peach gene	Chestnut QTL	Functional annotation in peach
ppa000612m	<i>Cbr1</i>	IQ calmodulin-binding motif
ppa010331m	<i>Cbr1</i>	Multifunctional chaperone (14-3-3 family)
ppa010444m	<i>Cbr1</i>	20S proteasome, regulatory subunit alpha type
ppa012082m	<i>Cbr1</i>	3R-hydroxymyristoyl ACP dehydrase
ppa018863m	<i>Cbr2</i>	Chitinase-related
ppa1027122	<i>Cbr2</i>	ATP synthase delta chain
ppa002568m	<i>Cbr3</i>	Long-chain acyl-CoA synthetases (AMP-forming)
ppa002730m	<i>Cbr3</i>	Alcohol dehydrogenase GroES-like domain
ppa002752m	<i>Cbr3</i>	Alcohol dehydrogenase GroES-like domain
ppa003014m	<i>Cbr3</i>	Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase
ppa005371m	<i>Cbr3</i>	Predicted RNA-binding protein
ppa005426m	<i>Cbr3</i>	LL-diaminopimelate aminotransferase
ppa006743m	<i>Cbr3</i>	Kynurenine aminotransferase, glutamine transaminase K
ppa009382m	<i>Cbr3</i>	AMP-dependent ligase/synthetase
ppa009922m	<i>Cbr3</i>	None
ppa010213m	<i>Cbr3</i>	AMP-dependent ligase/synthetase
ppa010395m	<i>Cbr3</i>	<i>cis</i> -Prenyltransferase, dehydrololichyl diphosphate synthase
ppa011261m	<i>Cbr3</i>	AMP-dependent ligase/synthetase
ppa013370m	<i>Cbr3</i>	Cysteine-rich secretory protein family
ppa013643m	<i>Cbr3</i>	None
ppa023235m	<i>Cbr3</i>	AP2 domain, transcription factor activity
ppa025016m	<i>Cbr3</i>	Histone H2A
ppa026281m	<i>Cbr3</i>	Histone deacetylase 6
ppa026313m	<i>Cbr3</i>	None

example, four of these segmentally homologous regions in peach span the three blight resistance QTLs in chestnut. Analysis of these regions in the annotated peach genome sequence shows that two regions contain genes for resistance to powdery mildew disease. Indeed, comparative genetic analyses of disease resistance in crop plants show that resistance genes are often clustered (Grube et al. 2000; Wisser et al. 2006a, b). Within the Solanaceae, clustering of genes conferring disease resistance to several unrelated pathogens often occurs and is conserved across tomato, potato, and pepper (Grube et al. 2000). The orthologous relationships supported by syntenic positions and sequence similarities between peach and chestnut suggest that these genomic regions may contain a set of conserved (prior to the divergence of the Fagaceae and Rosaceae) genetic elements whose products respond to fungal invasion.

The physical size of the *Prunus* genome is about 3.5× smaller than the size of the *Castanea* genome (Barow and Meister 2003; Kremer et al. 2007), yet regions of segmental homology were observed to account for roughly equivalent proportions of each genome. The *Prunus* reference genetic map covers ~520 cM (Dirlewanger et al. 2004; Howad et al. 2005), which is 40 % less than the size of the chestnut consensus map developed here. Independently computed sizes of significant segmental homologous regions between

these species (Table 2) are in agreement with comparative sizes of their genetic maps. Syntenic regions in *Castanea* may have undergone a general expansion relative to *Prunus*, possibly due to the acquisition and accumulation of repetitive DNA elements such as long terminal repeat retrotransposons. Accumulation of retrotransposon blocks between genes can play a significant role in genome evolution (Fedoroff 2000). In addition, they likely contribute to the larger sizes of syntenic genomic blocks observed in chestnut compared to peach.

To evaluate potential transferability of the EST-derived SSR markers from our dataset across the Fagaceae, we applied a “reverse bioinformatics” approach. Briefly, using the same BLASTn threshold criteria as used in the peach comparison, we completed homology searches for 301 EST-derived genetic markers previously mapped in other Fagaceae species (Casasoli et al. 2006; Durand et al. 2010). Of these, 25 SSRs (8.3 %) had strong sequence similarity to our mapped *C. mollissima* ESTs (Supplemental File 10). These markers along with the 16 genomic SSRs mapped in our *C. mollissima* × *C. dentata* F₂ population can now be tested between Fagaceae species such as *C. mollissima*, *C. dentata*, *C. sativa*, and *Q. robur* to develop a standard linkage group nomenclature. Previously, SSR markers mapped in *C. sativa* were placed on the same *C. mollissima* × *C. dentata* genetic

map allowing 11 of 12 homologous groups to be identified between *Castanea* species (Sisco et al. 2005). Further mapping of the *C. mollissima* genetic markers in other Fagaceae species will elucidate the genetic conservation across the family and extend the utility of genomic resources from the model species to less characterized species.

Identifying genes underlying the blight resistance QTLs in Chinese chestnut

Fine-scale genetic mapping

A rough estimate of the number of genes in the QTL intervals for blight resistance can be made based on approximate QTL size (twice the LOD 1 interval) and an estimated total number of genes (30,000). If this number of genes were equally distributed, a QTL covering 1.3 % of the genome (~10 cM) would contain about 400 genes. Increasing the resolution of QTL mapping by phenotyping and genotyping additional segregating progeny (e.g., at least twice as many) should reduce the potential genes in the interval. In future work, we anticipate resolving the three mapped QTLs (*Cbr1*, *Cbr2*, and *Cbr3*) to a higher degree to facilitate map-based cloning and marker-assisted selection as well as scanning for genes of lesser effect, including modifiers or genes that interact with the major QTLs. The markers and sequence resources reported here constitute a robust foundation for future fine mapping of QTLs for resistance and marker-assisted introgression activities in advanced generation *C. mollissima* × *C. dentata* hybrids. In addition, genome-wide genotyping is being carried out for the most and least resistant individuals from large BC₃ and BC₃-F₂ populations to identify the specific genomic segments of *C. mollissima* that have been maintained through four or more generations. It is anticipated that some of these segments will carry markers tightly linked to and within the QTLs.

Clues from gene expression studies

The genes located within the three chestnut blight resistance QTL intervals on the genetic map provide an extended list of candidates for blight resistance, given a QTL size of about 5 to 10 cM and an EST-based marker density of 1 per 0.7 cM. Many additional genes will be identified in the QTL interval from chestnut genome sequence when it becomes available. Gene expression can be correlated with the induction of disease and with differences in the response of resistance and susceptible species. A list of such differentially expressed candidate genes has been obtained from studies of EST abundance in control and infected Chinese and American chestnuts (Barakat et al. 2009). Only a small number of the candidate genes could be directly involved in the genetic basis of the QTLs and determine the response to

C. parasitica. Many others would be “downstream” effects that are part of the host response to the disease. Cloning and transfer of the indirect response genes would not confer resistance, but identification of such genes would provide useful biomarkers for evaluation of the disease response.

Clues from synteny with Prunus

Comparative genomics offers an additional path and new insights into candidate gene identification. Here we used newly developed, genetically mapped EST-based markers to bridge results from our relatively coarse-scale QTL intervals of blight resistance in *Castanea* to the very fine-scale mostly assembled and annotated genomic sequence in *Prunus*. This comparison allowed us to immediately look in these syntenic regions to search for disease resistance-like genes. Finding powdery mildew resistance QTLs colocalized to these regions gave us additional information about the potentially conserved nature of these genomic regions. Further inspection revealed known resistance and resistance-like genes that can now be considered advanced candidates for blight resistance in *Castanea*. Similarly, for the candidate genes identified from QTLs and expression studies (Table 3), several of these genes have been cloned from *C. mollissima* cDNA libraries and are now being transformed into *C. dentata* to evaluate their functional resistance to chestnut blight disease (W.A. Powell and S.A. Merkle, personal communication).

Physical mapping and sequencing

The genetic map size of 742.4 cM and an estimated genome size of 794 Mb (Kremer et al. 2007) give an overall ratio of genetic distance and physical size of 0.93 cM/Mb. The current average marker spacing of 0.7 cM (about 0.75 Mb) provides a feasible basis for map-based cloning. The consensus genetic map presented here has been aligned with the BAC-based physical map using hybridization of “overgo” probes representing genetic markers (Fang et al. 2012, companion manuscript). In brief, 691 linkage group-assigned markers (Supplemental File 11) were assigned to BAC contigs in the physical map and 350 BAC contigs were assigned to discrete genetic map positions. A graphical display of the genetic map using the CMap framework including its alignment with the physical map is available on the Fagaceae Genomics website (<http://www.fagaceae.org>). The availability of the integrated genetic and physical map puts a moderate number of BACs within these QTLs (*Cbr1*, *Cbr2*, and *Cbr3*), and the minimal number of overlapping BACs (minimum tiling path) across the QTLs can be determined. Selecting such physical contigs has been completed for each blight resistance QTL and sequencing these BAC clones is in progress (M.E.S unpublished data; J.E. Carlson, personal communication) in an attempt to identify all

the genetic elements within these important genomic regions. In addition, the integrated genetic and physical maps, along with BAC-end sequence information, are providing a framework for assisting whole genome sequence assembly (Fang et al. 2012, companion manuscript; J.E. Carlson, personal communication).

Towards American chestnut restoration

Understanding the underlying genetic mechanisms of chestnut blight resistance would greatly facilitate the efficient and effective transfer of blight resistance to *C. dentata*. The genomic resources and analyses presented here promise to advance this collaborative, multifaceted effort with an ultimate goal of restoring *C. dentata* and its ecosystem across its native range.

Acknowledgments We thank Charles Burdine and Dennis Deemer (SIFG) for technical assistance, including data collection using ABI 3130xl and 3730xl Genetic Analyzers and data scoring using ABI GeneMapper software. We thank Pamela Sletten (CAES) and Cornelia Pinchot, Kendra Gurnery, William White, and David Bevins (TACF) for technical assistance in crossing, growing progeny, and collecting leaf samples for DNA analysis. Primary support for this project was provided by National Science Foundation's Plant Genome Research Program (NSF-PGRP 0605135) awarded to R. R.S. (NCSU).

Data Archiving Statement We followed standard Tree Genetics and Genomes policy. In addition all GenBank accession numbers are listed in Supplemental File 2.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Anagnostakis SL (1982) Biological control of chestnut blight. *Science* 215:466–471
- Anagnostakis SL (1987) Chestnut blight: the classical problem of an introduced pathogen. *Mycologia* 79(1):23–37
- Anagnostakis SL (1992) Measuring resistance of chestnut trees to chestnut blight. *Can J For Res* 22:568–571
- Andrade GM, Merkle SA (2005) Enhancement of American chestnut somatic seedling production. *Plant Cell Reports* 24:326–334
- Barakat A, DiLoreto DS, Zhang Y, Smith C, Baier K, Powell WA, Wheeler N, Sederoff R, Carlson JE (2009) Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection. *BMC Plant Biol* 9:51
- Barakat A, Staton M, Cheng C-H, Park J, Yassin NBM, Ficklin S, Yeh C-C, Hebard F, Baier K, Powell W, Schuster S, Wheeler N, Abbott A, Carlson J, Sederoff R (2012) Chestnut resistance to the blight disease: insights from transcriptome analysis. *BMC Plant Biol* 12:38. doi:10.1186/1471-2229-12-38
- Barow M, Meister A (2003) Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. *Plant Cell Environ* 26:571–584
- Barreneche T, Casasoli M, Russell K, Akkac A, Meddour H, Plomion C, Villani F, Kremer A (2004) Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theor Appl Genet* 108:558–566
- Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq polymerase: primer modifications that facilitate genotyping. *Biotechniques* 20:1004–1010
- Buck EJ, Hadonou M, James CJ, Blakesley D, Russell K (2003) Isolation and characterisation of polymorphic microsatellites in European chestnut (*Castanea sativa* Mill.). *Mol Ecol Notes* 10:1046–1048
- Burnham CR (1981) Blight-resistant American chestnut: there's hope. *Plant Dis* 65:459–460
- Burnham CR, Rutter PA, French DW (1986) Breeding blight-resistant chestnuts. *Plant Breed Rev* 4:347–397
- Buttrick PL (1915) Commercial uses of chestnut. *Am For* 21:961
- Calabrese PP, Chakravarty S, Vision TJ (2003) Fast identification and statistical evaluation of segmental homologies in comparative maps. *Bioinformatics* 19(suppl 1):i74–i80
- Cameron DD (2002) The drovers of Appalachia. *J Amer Chestnut Found* 15(2):10–13
- Casasoli M, Mattioni C, Cherubini M, Villani F (2001) A genetic linkage map of European chestnut (*Castanea sativa* Mill.) based on RAPD, ISSR and isozyme markers. *Theor Appl Genet* 102:1190–1199
- Casasoli M, Pot D, Plomion C, MonteVerdi MC, Barreneche T, Lauteri M, Villani F (2004) Identification of QTLs affecting adaptive traits in *Castanea sativa* Mill. *Plant Cell Environ* 27:1088–1101
- Casasoli M, Derory J, Morera-Dutrey C, Brendel O, Porth I, Guehl J-M, Villani F, Kremer A (2006) Comparison of quantitative trait loci for adaptive traits between oak and chestnut based on an expressed sequence tag consensus map. *Genetics* 172:533–546
- Cheng L, Huang W (2009) An evaluation of SSR markers isolated from *Corylus avellana* for genetic analysis within four Asian species of the genus *Castanea*. *Acta Hort* 844:185–188
- Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. *Genetics* 138:963–971
- Clapper RB (1952) Relative blight resistance of some chestnut species and hybrids. *J For* 50:453–455
- Deemer DL, Nelson CD (2010) Standardized SSR allele naming and binning among projects. *Biotechniques* 49:835–836
- Dirlwanger E, Pascal T, Zuger C, Kervella J (1996) Analysis of molecular markers associated with powdery mildew resistance genes in peach (*Prunus persica* (L.) Batsch) × *Prunus davidiana* hybrids. *Theor Appl Genet* 93:909–919
- Dirlwanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arús P (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proc Natl Acad Sci (USA)* 101:9891–9896
- Diskin M, Steiner K, Hebard FV (2006) Recovery of American chestnut characteristics following hybridization and backcross breeding to restore blight-ravaged *Castanea dentata*. *For Ecol Manag* 223:439–447
- Durand J, Bodénès C, Chancerel E, Frigerio J-M, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn H-P, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herrán A, Ikarán Z, Cabané C, Ueno S, Alberto F, Dumoulin P-Y, Guichoux E, de Daruvar A, Kremer A, Plomion C (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. *BMC Genomics* 11:570
- Echt CS, May-Marquardt P (1997) Survey of microsatellite DNA in pines. *Genome* 40:9–17
- Fang G-C, Blackmon BP, Staton ME, Nelson CD, Kubisiak TL, Olukolu BA, Henry D, Zhebentyayeva T, Saski CA, Cheng C-H, Monsanto M, Ficklin S, Atkins M, Georgi LL, Barakat A, Wheeler N, Carlson JE, Sederoff RR, Abbott AG (2012) A physical map of the Chinese

- chestnut (*Castanea mollissima*) genome, and its integration with the genetic map. *Tree Genetics and Genomes*. doi:10.1007/s11295-012-0576-6
- Fedoroff N (2000) Transposons and genome evolution in plants. *PNAS (USA)* 97:7002–7007
- Foulongne M, Pascal T, Pfeiffer F, Kervella J (2003) QTLs for powdery mildew resistance in peach \times *Prunus davidiana* crosses: consistency across generations and environments. *Mol Breed* 12:33–50
- Freinkel S (2007) American chestnut: the life, death and rebirth of a perfect tree. University of California Press, Berkeley, p 284
- Gobbin D, Hohl L, Conza L, Jermini M, Gessler C, Conedera M (2007) Microsatellite-based characterization of the *Castanea sativa* cultivar heritage of southern Switzerland. *Genome* 50:1089–1103
- Götz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucl Acids Res* 36:3420–3435
- Graves AH (1950) Relative blight resistance in species and hybrids of *Castanea*. *Phytopath* 40:1125–1131
- Grube RC, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the Solanaceae. *Genetics* 155:873–887
- Gupta PK, Rustgi S, Mir RR (2008) Array-based high-throughput DNA markers for crop improvement. *Heredity* 101:5–18
- Haley CS, Knott SA (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Genetics* 132:1211–1222
- Hampson S, McLysaght A, Gaut B, Baldi P (2003) LineUp: statistical detection of chromosomal homology with application to plant comparative genomics. *Genome Res* 13:999–1010
- Hebard FV (1994) The American Chestnut Foundation breeding plan: beginning and intermediate steps. *J Amer Chestnut Found* 8 (1):21–28
- Hebard FV (1995) Meadowview notes 1993. *J Amer Chestnut Found* 8 (2):6–9
- Hebard FV (2006a) The backcross breeding program of the American Chestnut Foundation. *J Amer Chestnut Found* 19:55–77
- Hebard FV (2006b) The backcross breeding program of the American Chestnut Foundation. In: Steiner KC and Carlson JE (eds) Restoration of American chestnut to forest lands, Proceedings of a Conference and Workshop, May 4–6, 2004, The North Carolina Arboretum, Asheville, pp 61–77. Natural Resources Report NPS/NCR/CUE/NRR-2006/01. National Park Service, Washington
- Howard W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005) Mapping with a few plants: using selective mapping for microsatellite saturation of the *Prunus* reference map. *Genetics* 171:1305–1309
- Hyten DL, Song Q, Choi I-Y, Yoon M-S, Specht JE, Matukumalii LK, Nelson RL, Shoemaker RC, Young ND, Cregan PB (2008) High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. *Theor Appl Genet* 116:945–952
- Inoue E, Ning L, Hara H (2009) Development of simple sequence repeat markers in Chinese chestnut and their characterization in diverse chestnut cultivars. *J Amer Soc Hort Sci* 134:610–617
- Islam-Faridi N, Nelson CD, Banda H, Majid MA, Kubisiak TL, Hebard FV, Sisco PH, Paris RL, Phillips RL (2008) Cytogenetic analysis of a reciprocal translocation in F_1 hybrid between American and Chinese chestnuts. *Plant and Animal Genome Conference XVI Abstract W346*, San Diego CA
- Islam-Faridi N, Nelson CD, Sisco PH, Kubisiak TL, Hebard FV, Paris RL, Phillips RL (2009) Cytogenetic analysis of American chestnut (*Castanea dentata*) using fluorescent in situ hybridization. *Acta Hort* 844:207–210
- Jansen RC, Stam P (1994) High resolution mapping of quantitative traits into multiple loci via interval mapping. *Genetics* 136:1447–1455
- Jaynes RA (1962) Chestnut chromosomes. *Forest Sci* 8:372–377
- Jin L, Zhong Y, Chakraborty R (1994) The exact number of possible microsatellite motifs. *Amer J Human Genet* 55:582–583
- Kremer A, Casasoli M, Barreneche T, Bodénès C, Sisco P, Kubisiak T, Scalfi M, Leonardi S, Bakker E, Buiteveld J, Romero-Severson J, Arumuganathan K, Derory J, Scotti-Saintagne C, Roussel G, Bertocchi ME, Lexer C, Porth I, Hebard F, Clark C, Carlson J, Plomion C, Koelewijn H-P, Villani F (2007) Fagaceae trees. In: Kole C (ed) Genome mapping and molecular breeding in plants, vol 7, Forest trees. Springer, Leipzig, pp 161–180
- Kubisiak TL, Hebard FV, Nelson CD, Zhang J, Bernatzky R, Huang H, Anagnostakis SL, Doudrick RL (1997) Molecular mapping of resistance to blight in an interspecific cross in the genus *Castanea*. *Phytopath* 87:751–759
- Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199
- Lehmann EL (1975) Nonparametrics. McGraw-Hill, New York
- Little EL Jr. (1977) Atlas of United States trees, volume 4, Minor Eastern Hardwoods: U.S. Department of Agriculture Miscellaneous Publication 1342, 17 p, 230 maps
- Marinoni D, Akkac A, Bounous G, Edwards KJ, Botta R (2003) Development and characterization of microsatellite markers in *Castanea sativa* (Mill). *Mol Breed* 11:127–136
- Marth GT, Korf I, Yandell MD, Yeh RT, Gu Z, Zakeri H, Stitzel NO, Hillier L, Kwok P, Gish WR (1999) A general approach to single-nucleotide polymorphism discovery. *Nat Genet* 23:452–456
- Martin AC, Zim HS, Nelson AL (1951) American wildlife and plants, a guide to wildlife food habits. Dover Books, New York, p 500
- Merkel HW (1905) A deadly fungus on the American chestnut. New York Zoological Society, 10th Annual Report, pp 97–103
- Merkle SA, Andrade GM, Nairn CJ, Powell WA, Maynard CA (2007) Restoration of threatened species: a noble cause for transgenic trees. *Tree Genet Genomes* 3:111–118
- Pascal T, Pfeiffer F, Kervella J (2010) Powdery mildew resistance in the peach cultivar Pamirskij5 is genetically linked with the *Gr* gene for leaf color. *HortSci* 45:150–152
- Polin LD, Liang H, Rothrock RE, Nishii M, Diehl DL, Newhouse AE, Nairn CJ, Powell WA, Maynard CA (2006) *Agrobacterium*-mediated transformation of American chestnut (*Castanea dentata* (Marsh.) Borkh.) somatic embryos. *Plant Cell Organ Tissue Cult* 84(1):69–79
- Rieseberg LH (2001) Chromosomal rearrangements and speciation. *Trends Ecol Evol* 16:351–358
- Rostoks N, Ramsey L, MacKenzie K, Cardle L, Bhat PR, Roose ML, Svensson JT, Stein N, Varshney RK, Marshall DF, Graner A, Close TJ, Waugh R (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. *Proc Natl Acad Sci (USA)* 103:18656–18661
- Russell EWB (1987) Pre-blight distribution of *Castanea dentata* (Marsh.). *Borkh Bull Torrey Bot Club* 114:183–190
- Schuelke M (2000) An economic method for fluorescent labeling of PCR fragments. *Nature Biotech* 18:233–234
- Sisco PH, Kubisiak TL, Casasoli M, Barreneche T, Kremer A, Clark C, Sederoff RR, Hebard FV, Villani F (2005) An improved genetic map for *Castanea mollissima* / *Castanea dentata* and its relationship to the genetic map of *Castanea sativa*. *Acta Hort* 693:491–496
- Tanaka K, Tsumura Y, Nakamura T (1999) Development and polymorphism of microsatellites markers for *Fagus crenata* and the closely related species, *F. japonica*. *Theor Appl Genet* 99:11–15
- Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhaleerao RR, Bhaleerao RP, Blauze D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen G-L, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroev S, Déjardin A, dePamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehling J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L,

- Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leplé J-C, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson CD, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouzé P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai C-J, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). *Science* 313:1596–1604
- Ueno S, Aoki K, Tsumura Y (2009) Generation of expressed sequence tags and development of microsatellite markers for *Castanopsis sieboldii* var. *sieboldii* (Fagaceae). *Ann For Sci* 66:509
- Utz HF, Melchinger AE (1994) Comparison of different approaches to interval mapping of quantitative trait loci. In: van Ooijen JW, Jansen RC (eds) *Biometrics in plant breeding: applications of molecular markers*. CRPO-DLO, Wageningen, pp 195–204
- van Ooijen RW (2004) MapQTL 5.0, software for the mapping of quantitative trait loci in experimental populations. Plant Research International, Wageningen
- van Ooijen JW, Voorrips RE (2001) JoinMap® 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen
- Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77–78
- Wheeler N, Sederoff RR (2009) Role of genomics in the potential restoration of American chestnut. *Tree Genet Genomes* 5:181–187
- Wisser RJ (2006b) The genetic architecture of quantitative resistance in the cereals. PhD thesis, Cornell University, Ithaca, NY, p 195
- Wisser RJ, Balint-Kurti PJ, Nelson RJ (2006a) The genetic architecture of disease resistance in maize: a synthesis of published studies. *Phytopath* 96:120–129
- Yan J, Shah T, Warburton ML, Buckler ES, McMullen MD, Crouch J (2009) Genetic characterization and linkage disequilibrium estimation of a global maize collection using SNP markers. *PLoS One* 4(12):e8451
- Yan J, Yang X, Shah T, Sanchez-Villeda H, Li J, Warburton M, Zhou Y, Crouch JA, Xu Y (2010) High-throughput SNP genotyping with the GoldenGate assay in maize. *Mol Breed* 25:441–451
- Zeng Z-B (1994) Precision mapping of quantitative trait loci. *Genetics* 136:1457–1468