ORIGINAL PAPER

The genomic architecture of disease resistance in lettuce

**Leah K. McHale · Maria José Truco · Alexander Kozik · Tadeusz Wroblewski ·
Oswaldo E. Ochoa · Kirsten A. Lahre · Steven J. Knapp · Richard W. Michelmore**

Abstract Genbank and The Compositae Genome Project
database, containing over 42,000 lettuce unigenes from
Lactuca sativa cv. Salinas and *L. serriola* accession
UC96US23 were mined to identify 702 candidate genes
involved in pathogen recognition (*RGC*s), resistance signal
transduction, defense responses, and disease susceptibility.
In addition, to identify sequences representing additional
sub-families of nucleotide binding site (NBS)-leucine-rich
repeat encoding genes; the major classes of resistance
genes (*R*-genes), NBS-encoding sequences were amplified
by PCR using degenerate oligonucleotides designed to
NBS sub-families specific to the subclass Asteridae, which
includes the Compositae family. These products were
cloned and sequenced resulting in 18 novel NBS sequences
from cv. Salinas and 15 novel NBS sequences from
UC96US23. Using a variety of marker technologies, 294 of
the 735 candidate disease resistance genes were mapped in
our primary mapping population, which consisted of 119 F_7
recombinant inbred lines derived from an interspecific
cross between cv. Salinas and UC96US23. Using markers
shared across multiple genetic maps, 36 resistance pheno-
typic loci, including two new loci for resistance to downy
mildew and two quantitative trait loci for resistance to
anthracnose were positioned onto the reference map to pro-
vide a global view of the genomic architecture of disease
resistance in lettuce and to identify candidate genes for
resistance phenotypes. The majority but not all of the resis-
tance phenotypes were genetically associated with *RGC*s.

Introduction

The outcomes of interactions between plants and potential
pathogens are determined by numerous molecular events.
Variation in any of the genes encoding determinants on the
plant side could potentially influence the interaction and hence
be observed as variation in disease resistance or susceptibility.
Potential determinants of disease resistance can be broadly
classified as pathogen recognition genes (*RGC*s; resistance
gene candidates), signal pathway genes, defense response
associated genes, and susceptibility factors. The development
of comprehensive genomic resources as well as efficient map-
ping approaches now make it possible to dissect the genetic
determinants of disease resistance and to demonstrate the
types of genes responsible for various forms of resistance.

Disease resistance is often but not always determined by
single resistance genes (*R*-genes). Over 70 such genes have
now been characterized at the molecular level (Liu et al.
2007). Most of the genes conferring monogenic resistance
cloned to date encode proteins containing nucleotide bind-
ing site (NBS) and leucine rich repeat (LRR) domains with
an N-terminal region having a variable or coiled-coil (CC)
domain or having similarity to Toll and interleukin-1 recep-
tor (TIR) (Bent et al. 1994; Whitham et al. 1994). Other

L. K. McHale · M. J. Truco · A. Kozik · T. Wroblewski ·
O. E. Ochoa · K. A. Lahre · R. W. Michelmore (✉)
The Genome Center and Department of Plant Sciences,
University of California, Davis, CA 95616, USA
e-mail: rwmichelmore@ucdavis.edu

L. K. McHale
e-mail: lkmchale@ucdavis.edu

S. J. Knapp
Institute of Plant Breeding, Genetics and Genomics,
Center for Applied Genetic Technologies,
University of Georgia, Athens, GA 30602, USA

classes of *R*-genes are receptor-like proteins including intracellular serine/threonine protein kinases (PK) and genes encoding an extracellular-LRR and usually a trans-membrane- (TM-) domain but without NBS, CC, or TIR domains (Jones et al. 1994; Martin et al. 1993; Song et al. 1995). Genes of unknown function with sequence similarity to NBS-LRR encoding and other classes of *R*-genes are known as resistance gene candidates (*RGC*s) or resistance gene analogs (*RGA*s) and are generally thought to function in pathogen recognition (Tan et al. 2007). Monogenic resistance can also be conferred by genes which code for proteins outside of these R protein classes and do not function in pathogen recognition. Mutations in host susceptibility factors can result in resistance. Recessive resistance to viral pathogens is sometimes encoded by mutations in host proteins required by the virus for successful replication and spread (Albar et al. 2006; Nieto et al. 2006; Robaglia and Caranta 2006; Ruffel et al. 2006). In addition, recessive resistance to fungi can be encoded by *Mlo*-like genes, the protein products of which act as gateways for the pathogen entry (Buschges et al. 1997; Panstruga 2005).

The molecular bases of quantitative disease resistance that may be determined either monogenically or polygenically are unclear. In some cases, quantitative resistance may be determined by NBS-LRR encoding genes with incomplete penetrance; in other cases, there may be a different molecular basis. The co-evolution of macromolecular complexes may lead to epistatic interactions between genes at various loci (Bomblies et al. 2007). Genetic studies, particularly in *Arabidopsis*, have identified genes involved in signal transduction pathways downstream of cloned resistance genes (reviewed in Glazebrook 2001). Knock-out mutations in these resistance signaling pathway genes and defense response genes often result in broad spectrum disease resistance or susceptibility; however, there are few data available relating natural variation in these genes to phenotypic differences in disease resistance.

Pathogens secrete effector molecules which enter the plant cell and target host proteins in order to manipulate the host and create a favorable environment for pathogenesis (Grant et al. 2006). Plants probably have a limited number of points of vulnerability to pathogens and likely effector targets include proteins involved in plant disease resistance. Recognition of an effector molecule by a *RGC* results in triggering of plant defense (Jones and Dangl 2006; Mackey et al. 2002; Rooney et al. 2005; Shao et al. 2003). Plants have evolved the ability to recognize effectors either directly or indirectly via monitoring the activity of the effectors on their plant targets. However, only a few examples of direct or indirect recognition have been demonstrated at the molecular level (Belkhadir et al. 2004; Chisholm et al. 2006; Deslandes et al. 2003; Dodds et al. 2006; Jia et al. 2000).

Until recently, resistance phenotypes, particularly against different pathogens, have only rarely been mapped relative to each other and the generation of genotypes with resistance to multiple diseases has been slow (Grube et al. 2000; Pan et al. 2000; Zhang et al. 2002). The availability of cloned resistance genes and increasingly efficient molecular marker technologies provide opportunities for determining the genetic positions of phenotypic resistances and candidate genes (Naik et al. 2006). Such data will allow rapid characterization and exploitation of wild germ plasm as well as the pyramiding of resistance genes using marker-aided selection (Datta et al. 2002; Jain et al. 2004; Michelmore 1995, 2003).

Lettuce is an important crop worldwide and ranks as one of the top ten of most valuable crops in the US with an annual value of over $2 billion (Anonymous 2007). It is grown as extensive monocultures, often with several crops per year. Such intensive production makes the crop vulnerable to major epidemics and lettuce suffers from several economically important pests and diseases, of which lettuce downy mildew is the most serious (Davis et al. 1997). Breeding for disease resistance is a major objective of lettuce breeding programs and has involved the introgression of resistance from wild species, particularly *L. serriola* (Crute 1988). Previous studies have identified 25 loci in lettuce that contribute towards resistance to seven diseases, downy mildew (*Bremia lactucae*), root downy mildew (*Plasmopara lactucae-radicis*), corky root rot (*Sphingomonas suberifaciens*), lettuce mosaic virus (LMV), turnip mosaic virus (TuMV), lettuce dieback (Tomato bushy stunt virus; TBSV), and root aphid (*Pemphigus bursarius*) (Grube et al. 2005; Jeuken and Lindhout 2002; Kesseli et al. 1993; Landry et al. 1987; Maisonneuve et al. 1994; Montesclaros et al. 1997; Moreno-Vazquez et al. 2003; Paran et al. 1991). However, prior to this study, only 16 had been placed on the genetic map relative to anonymous molecular markers, mainly AFLPs (Kesseli et al. 1994; Truco et al. 2007).

In this study, we describe a detailed analysis of the genomic architecture of disease resistance in lettuce. Through mining of the lettuce EST database (http://cgpdb.ucdavis.edu) and cloning of new NBS-LRR encoding sequences using degenerate oligonucleotide primers, we identified numerous candidate genes for resistance and categorized them as *RGC*s, resistance signaling pathway genes, defense response associated genes, or susceptibility factors. A subset of these sequences was mapped relative to 26 previously and four newly phenotypically defined disease resistance loci. In addition, the genetic determinants of reactions to six different effector molecules from bacterial pathogens were mapped relative to *RGC*s and potential targets of pathogen effector proteins. The majority but not all of the resistance phenotypes were genetically associated with *RGC*s.

Materials and methods

Plant materials and DNA isolation

Four mapping populations were used for these studies (Table 1). The reference mapping population for *Lactuca* sp. was an $F_{7:8}$ population of 119 recombinant inbred lines (RILs) derived from *L. sativa* cv. Salinas and *L. serriola* acc. UC96US23 by single-seed descent from an F_2 population; 247 AFLP and SSR markers had been analyzed on this population to generate a detailed genetic map representing the nine chromosomal linkage groups (Truco et al. 2007). An F_2 population was generated from a cross between *L. sativa* cv. Salad Bowl and *L. serriola* acc. CGN14263. Ninety F_2 individuals were genotyped with 166 AFLP and 19 SSR markers and a genetic map was generated as described previously (Truco et al. 2007). F_3 families derived by self pollination from each of the 90 F_2 individuals were analyzed for resistance to *B. lactucae* and anthracnose. A second F_2 mapping population consisting of 105 individuals was derived from *L. sativa* cv. Ninja × *L. sativa* cv. Valmaine and scored for reactions to pathogen effectors. A third F_2 population was derived from from *L. serriola* LSE18 × *L. sativa* cv. Valmaine. F_3 families were generated from 101 F_2 individuals and scored for reactions to pathogen effectors. DNA was extracted from F_8 families and F_2 individuals using a modified CTAB method (Bernatsky and Tanksley 1986).

Evaluation of resistance to lettuce downy mildew

Resistance to isolates of *B. lactucae*, the causal agent of lettuce downy mildew, was evaluated by inoculating 15–20 seven-day-old seedlings and assessing sporulation two weeks post-inoculation, as described previously (Hulbert and Michelmore 1985). To phenotype *Dm7*, *Dm5/8*, and *Dm13* segregating in the Salinas × UC96US23 population, 108 RILs were evaluated for resistance to isolates R60 (expressing *Avr7*), CG1 (expressing *Avr5/8*), and CS12 (expressing *Avr13* and *Avr5/8*).

The segregation of new genes for downy mildew resistance was evaluated in the Salad Bowl × CGN14263 F_2

population using four isolates of *B. lactucae,* C01O879 (pathotype CA-VIII), C98O648ED (pathotype CA-VII), C96O504 (pathotype CA-V), and C93D14 (pathotype CA-Novel) that were representative of pathotypes present at the time in California. CGN14263 was resistant and Salad Bowl was susceptible to all four isolates. Forty seedlings of each F_3 family were inoculated with each isolate and evaluated for resistance. The percentage of resistant seedlings was recorded for each F_2 progenitor for each isolate so the resistance phenotype could initially be analyzed as a quantitative trait.

Evaluation of resistance to anthracnose

Segregation of genes for resistance to anthracnose, caused by fungus *Microdochium panattoniana*, was evaluated in the F_2 population from Salad Bowl × CGN14263. Resistance was assessed by inoculating two Californian isolates of anthracnose, Ant 99-1 and Ant 83-5 onto approximately 50 plants per F_3 family as described previously (Ochoa et al. 1987). Salad Bowl was resistant to Ant 83-5, but susceptible to Ant 99-1. CGN14263 was resistant to both isolates. Plants at the 2–3 true leaf stage were sprayed with a suspension of conidia until droplets were visible on the leaf surface. Inoculated plants were incubated on a mist chamber at 18–21°C and evaluated after 10–12 days. Plants were scored individually using a disease index from 0–9 based on the appearance of lesions on the leaves. Scores of less than 4 were considered resistant. Asexual sporulation of *M. panatoniana* was associated with ratings of 5 and over. Due to segregation, the disease ratings of individual plants within F_3 families were heterogeneous. A disease severity index (DSI) was calculated by averaging scores for each family and was used for genetic analysis of the progenitor F_2 individual.

Evaluation of reactions to bacterial effectors

Reactions to bacterial effector proteins were assessed using *Agrobacterium tumefaciens*-mediated transient assays as described previously (Wroblewski et al. 2005). Effector proteins from *P. syringae*, *AvrRps4* (Genbank Acc. #

Table 1 Populations analyzed

Parental lines	Population	No. of lines	Traits phenotyped	Reference
L. sativa cv. Salinas × *L. serriola* acc. UC96US23	$F_{7:8}$ RILs	119	Downy mildew resistance Reaction to bacterial effectors	Truco et al. (2007)
L. sativa cv. Salad Bowl × *L. serriola* acc. CGN14263	$F_{2:3}$	90	Downy mildew resistance Anthracnose resistance	Truco et al. (2007)
L. sativa cv. Ninja × *L. sativa* cv. Valmaine	F_2	105	Reaction to bacterial effectors	This study
L. serriola acc. LSE18 × *L. sativa* cv. Valmaine	F_2	101	Reaction to bacterial effectors	This study

L43559), *AvrPto* (L20425), *AvrPpiC* (AE016853), *AvrRpm1* (X67808), *AvrB* (M21965), and *AvrRps2* (Z21715) were cloned into vector pBAV139 (Vinatzer et al. 2006) and expressed using *A. tumefaciens* strain C58C1 by leaf infiltration of RILs of Salinas × UC96US23 for *AvrRps4*, F$_3$ families of LSE18 × Valmaine for *AvrPto* and *AvrRps4*, or F$_2$ individuals of Ninja × Valmaine for *AvrRpm1*, *AvrB*, *AvrRps2*, and *AvrPpiC* (T. Wroblewski, in preparation). Elicitation of a hypersensitive response (HR) was evaluated 3 and 5 days post infiltration.

Identification of candidate resistance genes based on sequence similarity to genes with known function

The 76,043 and 52,034 ESTs from *L. sativa* cv. Salinas and *L. serriola* UC96US23, respectively, were assembled using CAP3 (Huang and Madan 1999) into 42,555 unigenes. These are downloadable as the CLX_S3 assembly from the Compositae Genome Project database (http://cgpdb. ucdavis.edu). These and other sequences from *L. sativa* and *L. serriola* retrieved in Genbank were searched for genes involved in resistance. The longest open reading frame translated from each unigene (http://cgpdb.ucdavis.edu) was used to conduct BLAST (blastp) searches (Altschul et al. 1990) against the protein sequences of *A. thaliana*. Lettuce sequences with the closest hits ($1 \times e^{-20}$ threshold) to sequences of *A. thaliana* annotated as involved in disease resistance or defense according to the *A. thaliana* Gene Ontology (GO) (Berardini et al. 2004) were identified as candidate lettuce resistance genes. Additional candidate lettuce resistance genes were identified as having significant BLAST (tblastn) hits ($1 \times e^{-20}$ threshold) to genes involved in resistance from other species. The lettuce sequences were also analyzed for TIR and NBS motifs using the HMMer program (Bateman et al. 2002; Eddy 1998); however, this identified no additional candidates.

Identification and phylogenetic analysis of new NBS sequences

Sets of NBS sequences from species in subclass Asteridae-containing few or no *Lactuca* sequences were identified

using neighbor-joining gene trees built from 1,600 NBS sequences from 24 plant families, described previously (McHale et al. 2006). Full sequences for each clade were realigned using ClustalX (Thompson et al. 1997) and degenerate oligonucleotide primers were designed for specific PCR amplification of each clade (Table 2). PCRs were performed on genomic DNA of Salinas and UC96US23 using HiFi *Taq* polymerase and an annealing temperature of 45°C according to the manufacturer's protocol (Invitrogen, Carlsbad, CA, USA). Products of the predicted size were purified from agarose using the QIAEX II Gel Extraction Kit (QIAGEN, Hilden, Germany) and cloned using the TOPO TA Cloning® Kit (Invitrogen). Sequencing reactions were carried out using Big-Dye™ 3.0 chemistry and T7 and M13 universal primers as per the manufacturer's protocol (Applied Biosystems, Foster City, CA). Sephadex G-50 columns (Millipore, MA, USA) were used to remove unincorporated dye-terminators. Electrophoresis was performed on an ABI 3700 capillary sequencer (Applied Biosystems) at the College of Agricultural and Environmental Sciences Genomics Facility, UC Davis (http://cgf.ucdavis.edu).

Chromatograms were analyzed using Sequencher v. 4.6 (Gene Codes Corporation, MI, USA) and BLAST (tblastn) searches to the *Arabidopsis* genome and known NBS-encoding sequences were used to determine similarity to NBS-encoding sequences. Novel NBS sequences were added to the multifamily NBS trees generated previously (McHale et al. 2006). At least 24 clones were sequenced from each amplified product and additional clones were sequenced until no new *Lactuca* NBS subfamilies were discovered. All sequences identified from only a single PCR amplification were re-amplified and re-sequenced using a specific, non-degenerate primer within the sequence and the degenerate primers at 5′ and 3′ ends to ensure all sequences were accurate and not artifacts of errors in the PCR. All *Lactuca* NBS sequences (novel, Genbank, and CGPDB derived) were translated into protein sequences and aligned using ClustalX v 1.83 with default options (Thompson et al. 1997). Nucleotide sequences were aligned based on protein alignments and a neighbor-joining tree was constructed using Phylip v 3.65 (Felsenstein 2004). Using ClustalX v 1.83, a protein alignment was constructed from a representative NBS sequence from each gene family for

Table 2 Degenerate oligonucleotides used for amplification of NBS-encoding sequences

Degenerate oligonucleotide 1		Degenerate oligonucleotide 2		Amplified sequences
F108	GGDGGDGTBGGDAARACHAC	R308	RTCRTTVACHARRTCRTGCAT	LsatNBS02–07, 09, 15, 19–22 LserNBS01–03, 05, 10, 16, 20, 24
F114	GGDATGGGDGGDGTBGGDAARAC	R114	YTGRAAHARYTCCCADGC	LsatNBS11
F114	GGDATGGGDGGDGTBGGDAARAC	R214	ARHGGHARHCCHCKRCAYTT	LsatNBS26
F001	GTBGTBCCDATHGTBGGDATG	R001	TTHCKHARRTCRTCCCAYTC	LsatNBS32, 40, 45, 46 LserNBS27,28, 33, 35, 39, 41, 42

Lactuca, *Helianthus* (Radwan et al. 2008), and *Cichorium* (Plocik et al. 2004), NBS sequences mined from the CGPDB for *Taraxacum officinale*, *L. perennis* and *L. virosa*, and NBS sequences functioning in resistance from other non-Compositae species. A neighbor-joining tree was built with Phylip v 3.65.

Candidate gene marker development and genotyping

Candidate genes were mapped over a period of five years using a variety of approaches as mapping technologies improved. All enabling marker information is available through the lettuce genetic map viewer on the CGP website (http://cgpdb.ucdavis.edu/database/genome_viewer/viewer). Oligonucleotide primers were designed to candidate genes with the aid of Primer3 (Rozen and Skaletsky 2000) to amplify products 250–500 bp in length, include predicted introns, and/or include in silico detected polymorphisms. Intron/exon splice sites and in silico polymorphisms were predicted with the aid of Python Contig Viewer (http://cgpdb.ucdavis.edu/SNP_Discovery). Genomic DNA (1 ng/μl) was used as template in 20 μl PCR volumes containing 1 mM $MgCl_2$, 0.05 μM forward primer, 0.05 μM reverse primer, 0.2 mM dNTPs, 1× PCR buffer, and 1 unit *Taq* DNA polymerase and cycling parameters of an initial denaturation for 5 min at 95°C, then 35 cycles of 30 s at 94°C, 30 s at 57°C and 1 min at 72°C, and a final extension for 5 min at 72°C.

PCR products were assessed for polymorphisms using a variety of methodologies. Large length polymorphisms and dominant markers were visualized and scored following electrophoresis in agarose gels (1.5% agarose, 1.5% Synergel; Diversified Biotech, Boston, MA) and staining with ethidium bromide. Some small length polymorphisms and SNPs were assayed using the REVEAL mutation detection system (Spectrumedix, College Park, PA), which employed temperature gradient capillary electrophoresis (TGCE). A variation of Spectrumedix's protocol was followed. Genotyping assays required two capillary runs per marker screened. The formation of a heteroduplex in the first TGCE run identified heterozygous individuals. For the second TGCE run, the PCR product from parental allele A was added to each of the original PCR samples to be genotyped. Individual samples with no heteroduplex formation were scored as parental allele A; samples forming a heteroduplex were scored as the alternate parental allele B.

Single-stranded conformational polymorphism (SSCP) analysis was used to score many markers. PCR products were diluted 1:1 with a loading buffer consisting of 40% formamide, 5 mM sodium hydroxide, 0.25% bromophenol blue, 0.25% xylene cyanol. PCR products were denatured at 95°C for 5 min and snap cooled on ice. DNA fragments were resolved by electrophoresis through 2.5× mutation detection enhancement gels (FMC BioProducts, Rockland, ME, USA). Gels were stained with silver nitrate by shaking for 7 min in 10% acetic acid, 20% ethanol and 0.25 M silver nitrate, followed by 5–10 min shaking in 1.2 M sodium hydroxide and 0.1% formaldehyde.

For each candidate gene in which polymorphism was not detected by any of the above methods, amplicons were sequenced from genomic DNA of Salinas and UC96US23. Following amplification, unincorporated primers were digested with Exonuclease I and Shrimp Alkaline Phosphatase (USB Corporation, Cleveland, OH, USA). Amplified products were sequenced and chromatograms analyzed as described above. Sequences with single nucleotide polymorphisms (SNPs) were submitted to Illumina, Inc. (San Diego, CA) for oligonucleotide design and marker detection using the Illumina Golden Gate® SNP assay (http://www.illumina.com). In brief, this technology simultaneously hybridizes hundreds (96, 384, 768, or 1,536) of allele-specific oligonucleotides to genomic DNA; these primers also contain a universal sequence which targets the primers to a particular address in the array. Allele-specific primer extension and ligation reactions are carried out before PCR with fluorescently labeled universal primers. Amplicons are hybridized to their address on the bead array and fluorescence is quantified resulting in an allele call for each bead. These markers were analyzed on the Illumina BeadStation by the DNA Technologies Core Facility at the UC Davis Genome Center (http://genomecenter.ucdavis.edu/dna_technologies/).

Where several markers for the same candidate gene were assayed, all markers were mapped to confirm co-segregation and then the most robust marker was chosen for inclusion in the genetic map. Redundant markers are documented at the Compositae Genome Project website map viewer (http://cgpdb.ucdavis.edu/database/genome_viewer/viewer/).

Linkage and bulked segregant analyses

A genetic linkage map based on the Salinas × UC96US23 $F_{7:8}$ RIL population and comprising 96 framework AFLP and 10 simple sequence repeat (SSRs) markers plus 291 resistance candidate genes, 481 non-resistance-related EST markers, and 4 phenotypic loci was constructed using JOINMAP v 4.0 with default settings and Kosambi mapping function (Stam and Van Ooijen 1995; Timms et al. 2006; Truco et al. 2007; J. Argyris and D.O. Lavelle unpublished). This map consisted of ten linkage groups spanning 1,296 cM; the two sections of linkage group 3 did not assemble together. The linear order of the map was compared to the lettuce integrated map (Truco et al. 2007) and there were no large rearrangements.

HR responses elicited by the bacterial effectors, AvrPto, AvrPpiC, AvrRpm1, AvrB, and AvrRps2, were mapped by bulked segregant analysis. Ninety-six PCR-based markers derived from resistance candidate genes were run on contrasting pairs of bulks of genomic DNA from 10 individuals that elicited an HR and from 10 individuals that elicited no HR in response to the bacterial effectors. Separate pairs of bulks were made for responses to AvrPto and AvrPpiC; a single pair of bulks was made for the responses to AvrRpm1, AvrB, and AvrRps2 as this co-segregated. Markers that distinguished the bulks were then analyzed on all phenotyped individuals in the populations and genetic linkage was determined with MapMaker v 3.0 (Lander et al. 1987).

QTL analysis of new resistances to anthracnose and downy mildew evaluated in Salad Bowl × CGN14263 were conducted using the composite interval mapping function in QTL cartographer v 2.5 (Wang et al. 2005) and based on the genetic map containing 166 AFLP and 19 SSR framework marker loci (Truco et al. 2007). Significance thresholds for identifying QTL were obtained by permutation analysis incorporating 1,000 permutations.

Sequences were submitted to Genbank: accession numbers EU889296 to EU889328 and FJ235160 to FJ235171.

Results

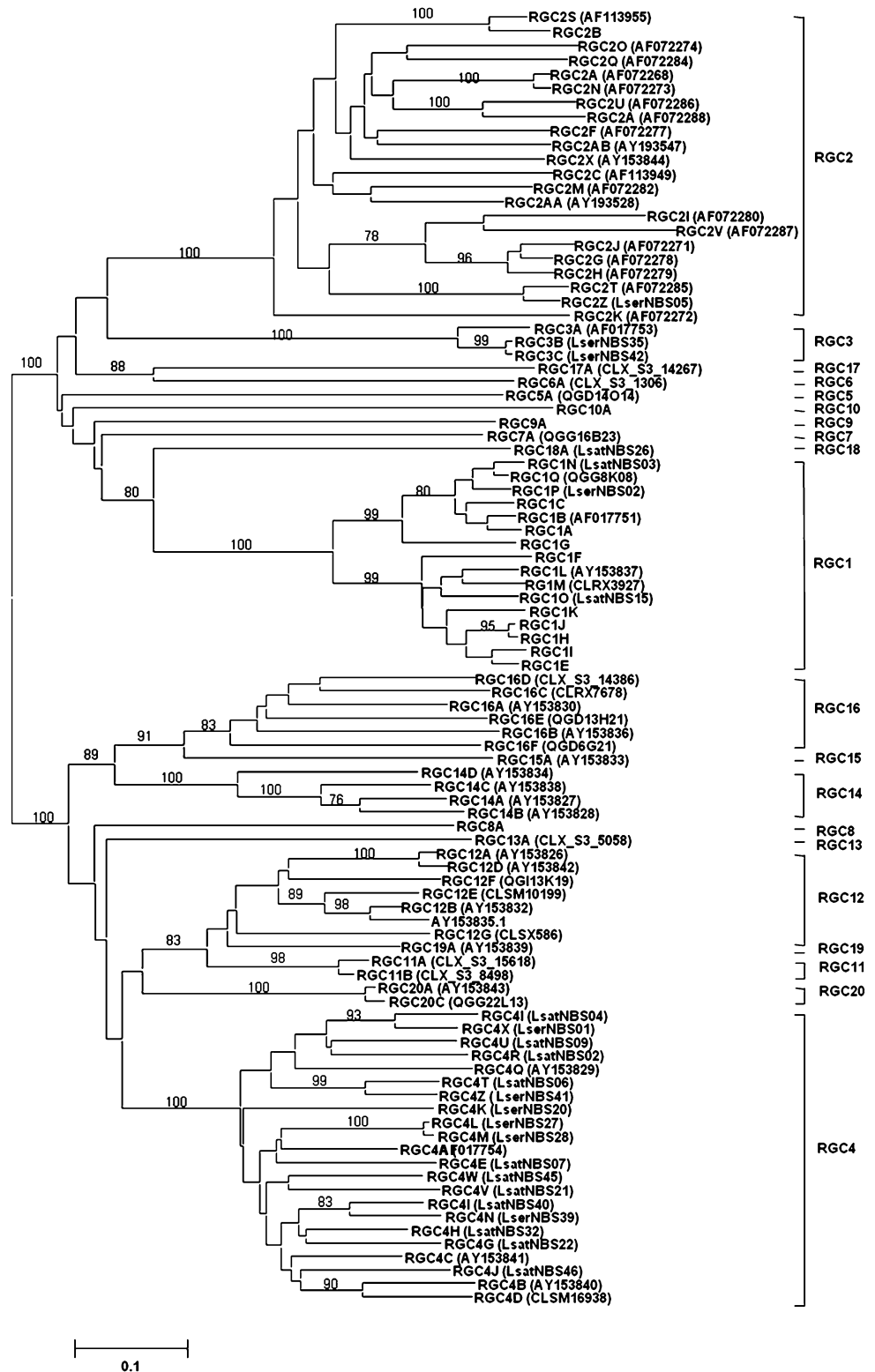### Identification of candidate resistance genes

Approximately 43,000 unigenes derived from ESTs of Salinas and UC96US23 generated by the Compositae Genome Project (http://cgpdb.ucdavis.edu) and *L. sativa* and *L. serriola* sequences retrieved in Genbank were searched for genes that may be involved in resistance. Translated sequences were used to conduct BLAST (blastp) searches (Altschul et al. 1990) against the inferred protein sequences of *A. thaliana*. Lettuce sequences with the closest hits ($1 \times e^{-20}$ threshold) to sequences of *A. thaliana* annotated as being involved in disease resistance or defense according to the *A. thaliana* GO annotations (Berardini et al. 2004) were identified as candidate lettuce resistance genes. Additional candidate lettuce resistance genes were identified as having significant BLAST (tblastn) hits ($1 \times e^{-20}$ threshold) to genes involved in resistance from other species identified by literature and keyword searches of PubMed and Genbank (http://www.ncbi.nlm.nih.gov). A total of 702 candidate resistance genes were identified from *Lactuca* Genbank sequences and the CGP database Salinas and UC96US23 EST assembly. These candidate resistance genes were classified according to sequence similarity to known genes from other species. The candidate genes included 207 *RGC*s, 266 resistance pathways, and 199

defense responses associated genes as well as 30 susceptibility factors (S1). *RGC*s were further categorized as similar to genes encoding CC-NBS-LRRs (80), TIR-NBS-LRRs (83), or receptor-like proteins (44).

Additional NBS-encoding sequences were identified using PCR-based approaches. Twelve new NBS-encoding sequences were amplified from *L. sativa* cv. Diana using degenerate oligonucleotide primers to the P-loop and GLPL conserved motifs (Shen et al. 1998; M. Van Damme, unpublished). Analysis of a tree of 1,600 NBS-encoding genes from diverse species revealed that there were 19 clades that contained sequences derived from species in the subclass Asteridae but lacked sequences from *Lactuca* spp. (McHale et al. 2006). In order to have *Lactuca* sequences from each major NBS-encoding gene family, clade-specific degenerate oligonucleotides primers were designed based on NBS-encoding sequences in Asteridae-containing clades with few or no sequences from *L. sativa* or *L. serriola*. Eighteen new NBS-encoding sequences from Salinas and 15 new NBS-encoding sequences from UC96US23 were amplified using these primers (Table 2). NBS-encoding sequences were amplified from both TIR-NBS-LRR and CC-NBS-LRR classes. All of the new NBS-encoding sequences of the CC-NBS-LRR class encoded this sub-family's characteristic tryptophan in the kinase-2 motif (Meyers et al. 1999). These PCR-based approaches brought the total number of NBS-LRR encoding sequences from *L. sativa* and *L. serriola* to 90 CC-NBS-LRR encoding genes and 106 TIR-NBS-LRR encoding genes.

Ninety-one amino acid fragments from the NBS domain of unique *Lactuca* NBS-LRR proteins from *L. sativa* (70) and *L. serriola* (21) were used to align the corresponding nucleotide sequences and a neighbor-joining tree was produced. Using a 70% identity threshold, sequences fell into 20 distinct sub-families of NBS-LRR proteins (*RGC*s1–20) containing from 1 to 28 sequences (Fig. 1). *RGC* sub-families 1–3, 5–7, 9, 10, 17, and 18 contain sequences similar to CC-NBS-LRR proteins and sub-families 4, 8, 11–16, 19 and 20 contain sequences similar to TIR-NBS-LRR proteins. A neighbor-joining tree was then constructed using NBS amino acid sequences representing each *RGC* sub-family from *L. sativa*, *Helianthus* spp. (Radwan et al. 2008), and *Cichorium* spp. (Plocik et al. 2004), NBS sequences mined from the CGP database from *Taraxacum officionale*, *L. perennis* and *L. virosa*, and NBS sequences of known resistance genes from other non-Compositae species obtained from Genbank. *Lactuca* sequences were present in each clade separated by the 70% amino acid identity threshold (Fig. 2). Therefore, we concluded that representatives of most of the major sub-families of NBS-encoding sequences in *Lactuca* spp had been obtained.

Fig. 1 Neighbor-joining tree of a 240 nucleotide fragment of the NBS-encoding region of 91 sequences. Only bootstrap values greater than 75% are indicated. Sub-families are defined as distinct groups with 70% nucleotide identity and indicated with *brackets*. Sequence names have been abbreviated for display, see S1 for full names. The sequence alignment is available from http://cgpdb.ucdavis.edu


Linkage analysis and genetic mapping of candidate resistance genes

Markers were developed, genotyped, and mapped for 294 of the 735 resistance candidate genes. These comprised 96 *RGC*s, 101 resistance pathway genes, 83 defense response associated genes, and 14 susceptibility factors (Fig. 3). The remaining 441 resistance candidate genes could not be mapped: 111 were monomorphic for the amplicon sequenced, 278 were monomorphic in agarose, SSCP, and

Fig. 2 Neighbor-joining tree of a 80 amino acid fragment of the NBS region encoded by each gene family in *L. sativa*, *Helianthus* spp. (Radwan et al. 2008); *Cichorium* spp. (Plocik et al. 2004), NBS sequences mined from the CGPDB representing *Taraxacum officionale*, *L. perennis* and *L. virosa*, and NBS sequences functioning in resistance from other non-Compositae species. Sequences from *L. sativa* and *L. serriola* are labeled in *green*, other Compositae species are labeled in *red*, non-Compositae species are labeled in *black*. Sequence names have been abbreviated for display, see S1 for full names. The sequence alignment is available from http://cgpdb.ucdavis.edu
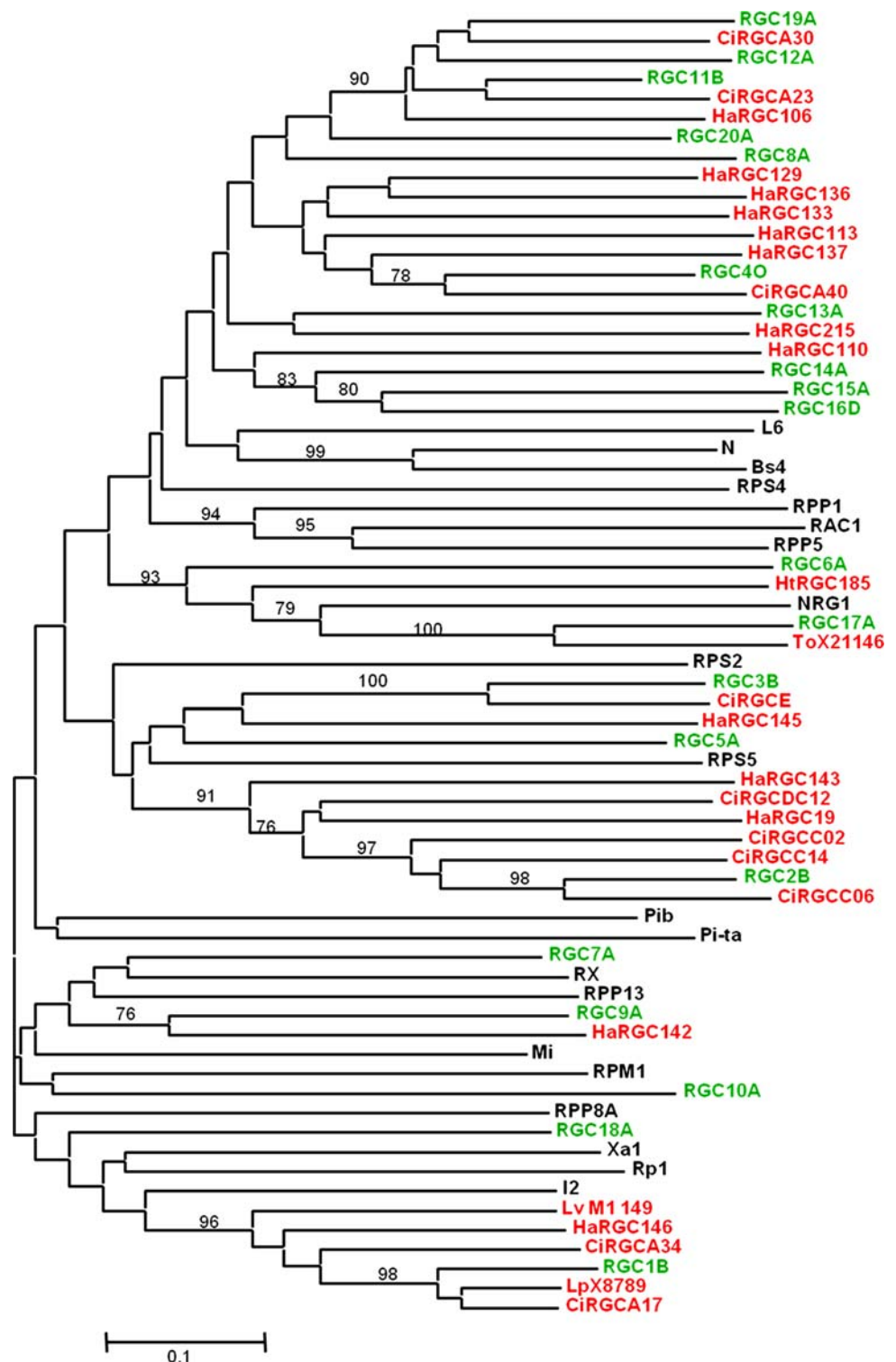
TGCE assays, but were not sequenced from both parental genotypes, and 62 were apparently polymorphic between the parents at the sequence level but attempts at marker development were unsuccessful. Markers were assayed using a variety of technologies: agarose electrophoresis (18), and TGCE (7), SSCP acrylamide electrophoresis (82), and the Illumina Golden Gate® SNP assay (187).

Markers for *RGC*s, resistance pathway, and defense response associated genes were distributed across all nine lettuce chromosomes. However, genes encoding NBS-LRR proteins were non-randomly distributed in the genome. Thirty-nine of the 78 mapped NBS-LRR encoding genes fell into seven clusters. Similar to clusters of NBS-LRR encoding sequences in *Arabidopsis* (Meyers et al. 1999;
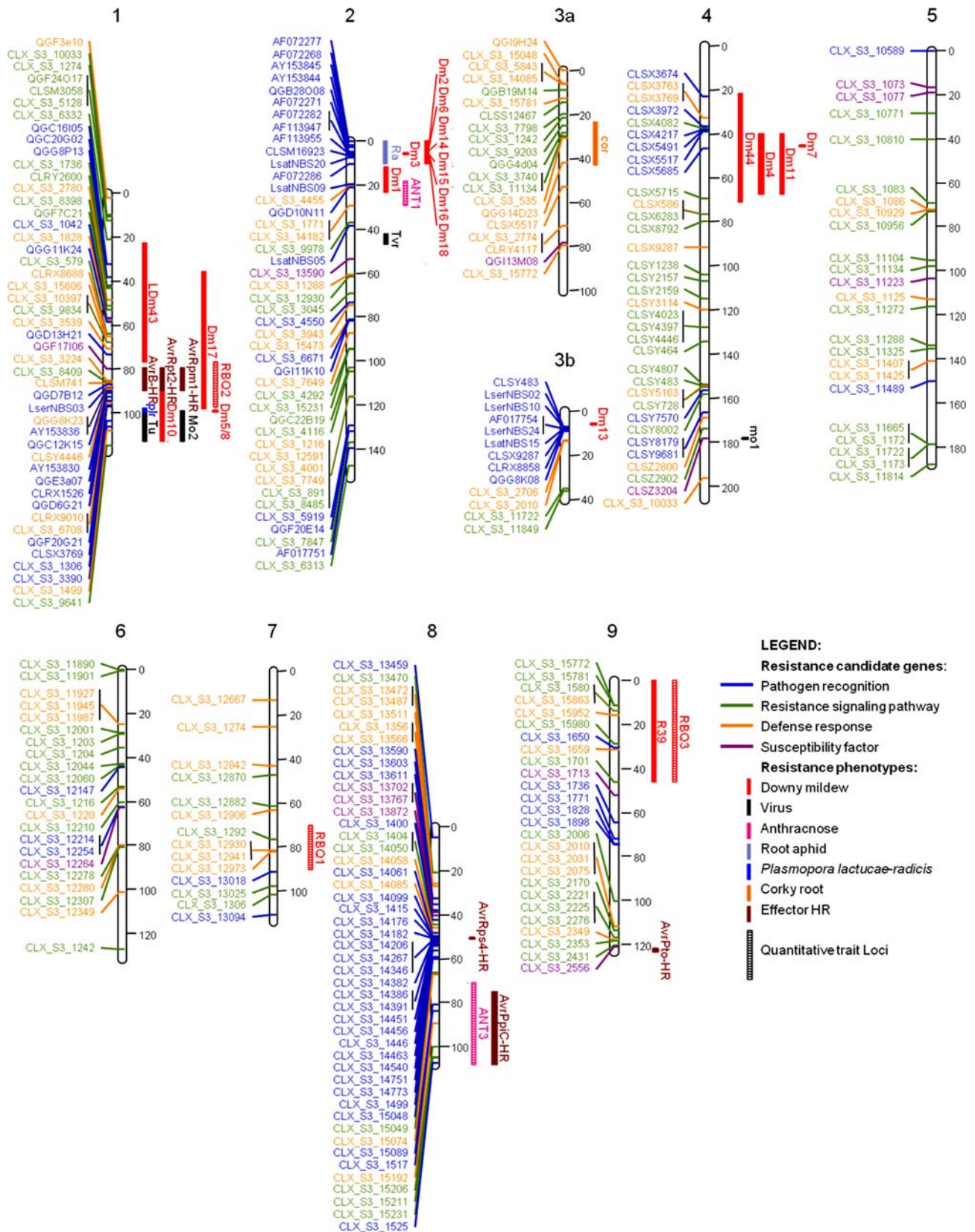
Fig. 3 Lettuce genetic map displaying 294 mapped candidate resistance genes and 35 resistance phenotypes. Sequence names have been abbreviated for display, see S1 for full names. Figure created with MapChart (Voorrips 2002)

Meyers et al. 2003), a cluster was defined as having at least three sequences of one type at a density greater than one marker per cM with no intervening EST-based markers of a different annotation. When mapped NBS-LRR encoding sequences represented the portion of the NBS domain used for phylogenetic analysis, it was noted which *RGC* subfamily was represented (Fig. 1); however, the subfamilies of many of the NBS-LRR genes is unknown because most of the EST sequences only encoded the 3′ LRR region due to the large size of NBS-LRR encoding genes. Chromosome 1 contains a cluster of four TIR-NBS-LRR encoding sequences that mapped to a 3 cM region; one is a *RGC16* member, the subfamily of the other three sequences is unknown. The previously characterized major resistance gene cluster that contains *Dm3* comprises of a large cluster of *RGC2* sequences on chromosome 2 (Meyers et al. 1998; Shen et al. 2002). All mapped members of the *RGC2* family mapped at this locus. In addition, one TIR-NBS-LRR mapped within a 7.5 cM region that also contained the *RGC2* cluster. Other CC-NBS-LRR type gene fragments also map to this region but do not encode an NBS domain; therefore, it is unknown if these sequences are members of the *RGC2* family. A large group of 16 TIR-NBS-LRR and 3 CC-NBS-LRR encoding sequences, which can be broken into three tightly linked clusters, mapped to an 11 cM region on chromosome 8. The first cluster in this large group includes eight TIR-NBS-LRR encoding sequences, seven of which are *RGC4* genes; the *RGC* subfamily of the eighth sequence is unknown because the EST did not encode an NBS domain. The second chromosome 8 cluster includes three sequences, a CC-NBS-LRR and TIR-NBS-LRR, both of unknown subfamilies, and an *RGC4* gene encoding TIR-NBS-LRR protein. The third cluster on chromosome 8 contains three *RGC20* sequences encoding TIR-NBS-LRR proteins. On chromosome 3, a mixture of 2 TIR-NBS-LRR and 7 CC-NBS-LRR encoding sequences map to a 2.5 cM region and represent both *RGC1* and *RGC4* sequences. While members of each *RGC* subfamily tended to map together, this is not always the case. Eleven of fourteen *RGC4* sequences mapped to a 6 cM region on chromosome 8, the other three sequences mapped individually to chromosomes 2, 3, and 7. Nucleotide identities parsed from BLAST (blastn) reports of all mapped lettuce sequences against each other was used to create an identity matrix. Sequence identity between mapped genes was plotted relative to genetic positions using GenomePixelizer (http://cgpdb.ucdavis.edu) and revealed no evidence from sequences flanking these *RGC4* family members of a segmental duplication between chromosome 8 and chromosomes 2, 3, or 7.

There were also a several small clusters of three non-*RGC* sequences. Some of these clusters may be artifacts of there being several non-overlapping unigenes representing a single gene or due to sequences from Salinas and UC96US23 having less than 95% sequence identity and therefore being identified as separate unigenes in the CAP3 assembly. In both cases, each sequence would have been mapped as a separate marker. However, four unigenes comprised of sequences from both Salinas and UC96US23 from an overlapping region encoding phenylalanine ammonia-lyase 1 (PAL1) mapped to a single locus on chromosome 2. From the BLAST (blastn)-derived identity matrix of mapped lettuce sequences it was determined that the four members of the *PAL1* cluster have the highest similarity to other sequences within the cluster and to three other *PAL1* sequences mapping to a single locus on chromosome 6. There was no evidence from the flanking mapped sequences for a segmental duplication between chromosomes 2 and 6.

Certain chromosomes had a paucity of candidate disease resistance genes. Although chromosomes 5 and 6 contained 9.5 and 7.9%, respectively of the 481 mapped EST sequences (http://cgpdb.ucdavis.edu/database/genome_viewer/viewer/), they had only 2 to 3% of the mapped *RGC*s. *RGC*s were significantly under-represented on these chromosomes ($P < 0.003$, $P < 0.03$; respectively; Fisher's exact test) based on a comparison of mapped *RGC*s versus mapped ESTs from other resistance categories and non-resistance related ESTs. This is in contrast to chromosomes 1, 2, and 8 that had a significant over-representation of *RGC*s ($P < 0.05$, $P < 0.02$, $P < 0.001$; respectively).

Mapping and distribution of phenotypic resistance loci

Dm5/8, *Dm7*, *Dm13*, and *AvrRps4-HR* segregated in the Salinas × UC96US23 RIL population and mapped to chromosomes 1, 4, 3, and 8, respectively, consistent with their positions determined using other populations (Landry et al. 1987; Kesseli et al. 1994; see below). QTL analysis of anthracnose isolate Ant 83-5 in the F_2 mapping population derived from *L. sativa* cv. Salad Bowl × *L. serriola* CGN14263 revealed QTLs on chromosomes 2 (*ANT2*) and 8 (*ANT3*) from CGN14263 accounting approximately for 40 and 30% of the variance, respectively. Only the QTL on chromosome 2 (*ANT2*) was significant on the analysis of isolate Ant 99-1. It accounted for approximately 55% of the variance. QTL analysis of downy mildew revealed resistance loci on chromosomes 1 and 4 from CGN14263. Both downy mildew resistance loci were dominant and sufficient to confer complete resistance to all *B. lactucae* isolates; they were therefore considered as distinct, independent *Dm* loci rather than QTL and designated *Dm43* and *Dm44* for the loci on chromosomes 1 and 4 respectively. HRs elicited by the bacterial effectors, AvrRpm1, AvrB, AvrRps2, and AvrPpiC, were analyzed by bulked segregant analysis followed by regression mapping of linked markers in the F_2

population derived from Ninja × Valmaine. The HRs elicited by AvrRpm1, AvrB, and AvrRps2 co-segregated with each other. This locus was mapped to a position coincident with the large cluster of resistances on chromosome 1 containing *Dm43* and *Dm5/8* among other resistances and was linked to *RGC16A*. *AvrPpiC-HR* was linked to *RGC5A* on chromosome 8 and coincident with *ANT3*. The HRs elicited by the bacterial effectors *AvrRps4* and AvrPto were analyzed in an F_2 population derived from LSE18 × Valmaine. *AvrRps4-HR* mapped to the same position as the HR mapped in the Salinas × UC96US23 population. *AvrPto-HR* was linked distally from QGG19L12 at the end of chromosome 9; no other resistance phenotypes mapped to this locus.

The positions of the QTL for resistance to anthracnose (*ANT2* and *ANT3*), the two new downy mildew loci (*Dm43* and *Dm44*), the loci for the HR to bacterial effectors (*AvrRpm1-HR*, *AvrB-HR*, *AvrRps2-HR*, *AvrPpiC-HR*, and *AvrPto-HR*) and 28 resistance phenotypes mapped in other populations were transferred to the Salinas × UC96US23 reference map with the aid of flanking markers shared between the maps (Fig. 3). Some genes could only be located to fairly large intervals due to the lack of close flanking markers common to the source and reference populations. The 36 resistance phenotypes are on seven chromosomes; no phenotypic resistance loci mapped to chromosomes 5 or 6. Chromosomes 1 and 2 have the largest number of known resistances with eight resistance phenotypes mapping across chromosome 1 and ten phenotypes mapping to the region surrounding the *RGC2* cluster on chromosome 2 (Fig. 3).

Associations between phenotypic and candidate gene loci

More *RGC*s (57%) co-localized with at least one resistance phenotype than did other categories of candidate genes: resistance pathway genes (22%), defense response-associated genes (34%), susceptibility factors (15%), or sequences thought not to be involved in resistance (24%). For resistance phenotypes mapped in populations other than Salinas × UC96US23, co-localization was defined as being within the genetic interval defined by flanking framework markers present on multiple maps. Not including the recessive viral resistance gene *mo1* that is known to encode the elongation initiation factor 4e (Nicaise et al. 2003), all resistance phenotypes except *RBQ1*, *cor, Tvr1,* and *AvrPto-HR* co-localized with at least one *RGC*. A QTL for downy mildew resistance, *RBQ1* (Jeuken and Lindhout 2002) and the recessive bacterial resistance gene, *cor* (Moreno-Vazquez et al. 2003), were coincident with defense response-associated and candidate resistance pathway genes. Neither *Tvr1* (Grube et al. 2005) nor *AvrPto-HR* (T. Wroblewski, unpublished) were at genetic positions coincident with any

type of candidate resistance genes mapped in this study. Homologs of the tomato resistance gene *Pto* were not among the candidate genes identified in this study.

The large cluster of NBS-LRR encoding genes on chromosome 2 co-localized with a similarly large cluster of resistance phenotypes (Meyers et al. 1998). Eleven resistance phenotypes mapped to the region on chromosome 1 that contained 18 genes encoding nine TIR-NBS-LRR and three CC-NBS-LRR proteins, as well as six receptor kinases. The cluster of seven CC-NBS-LRR and two TIR-NBS-LRR encoding sequences on chromosome 3 only co-localized with a single resistance phenotype, *Dm13*. The cluster of nine TIR-NBS-LRR and one CC-NBS-LRR encoding genes on chromosome 8 co-localized with *AvrRps4-HR*, the genetic determinant of the induction of HR in response to the bacterial effector *AvrRps4*, but did not co-localize with any known *Dm* specificities. For each of these resistance loci, as well as, six resistance phentoypes mapping to chromosome 1, significantly more *RGC* than non-*RGC* EST markers co-localized with them (Fisher's exact test, Table 3). Twenty-four NBS-LRR encoding genes, mapping to 15 distinct loci did not co-localize with any resistance phenotype mapped in this study. Such NBS-LRR encoding genes without phenotypes co-segregating could have alternate functions, but more likely their resistance phenotype have not yet been identified and mapped.

Discussion

Two hundred and ninety-one resistance-related sequences were mapped relative to 36 phenotypic resistance loci in lettuce. Thirty-four of 36 resistance phenotypes co-localized with at least one candidate resistance gene; this establishes the basis for further genetic and functional analyses as well as provides breeders with markers and enabling information for marker assisted selection (MAS).

Several genotyping technologies were used depending on the type of polymorphism and the marker technologies available at the time of analysis. The majority of markers used the Illumina GoldenGate® SNP assay (http://www.illumina.com) that became available towards the end of the study. Where feasible, SNPs were assayed using this technology due to its speed, low cost, and data reliability; however, it is not amenable to assaying indels, sequences with many SNPs (more than 1 SNP per 25 bp), and multigene families. Alternative marker technologies were employed to assay these types of polymorphisms. Large indels were assayed on agarose. Sequences that were highly polymorphic or contained small indels were assayed using TGCE or SSCP analysis. SNPs in large multigene families such as some *RGC*s were predominantly assayed using

Table 3 Fractions of the total *RGC*s, non-*RGC*s, all resistance candidate genes, and non-resistance related ESTs which co-localized with each resistance phenotype

Resistance gene	Chrm.	Fraction *RGC*s co-localized	Fraction non-*RGC*s ESTs co-localized	Fraction all resistance candidates genes co-localized	Fraction non-resistance related ESTs co-localized
Dm43	1	0.061	0.055	0.062	0.052
Dm17	1	0.15[†]	0.064	0.12[‡]	0.048
AvrB-HR *AvrRpt2-HR* *AvrRpm1-HR*	1	0.061*	0.019	0.055[‡]	0.0062
RBQ2	1	0.10[‡]	0.015	0.062[‡]	0.0042
plr	1	0.061[‡]	0.0044	0.027[†]	0.0021
Dm5/8	1	0.010	0	0.0034	0
Tu *Mo2*	1	0.020	0.0044	0.010	0.0042
Dm3 cluster	2	0.13[‡]	0.0015	0.045[‡]	0.0021
Dm1	2	0.031*	0.0059	0.014	0.0062
ANT2	2	0.031	0.012	0.021	0.010
Tvr1	2	0	0.0015	0	0.0020
cor	3	0	0.018	0.021	0.012
Dm13	3	0.092[‡]	0	0.031[‡]	0
Dm44	4	0.061	0.055	0.034	0.069
Dm4 *Dm11*	4	0.010	0.013	0.0034	0.019
Dm7	4	0.010	0	0.0034	0
RBQ1	7	0	0.019	0.014	0.0083
AvrRps4-HR	8	0.010[‡]	0	0.034[‡]	0
ANT3	8	0.031	0.022	0.024	0.023
AvrPpiC-HR	8	0.031	0.021	0.024	0.021
R39 *RBQ3*	9	0.010	0.044	0.031	0.046
AvrPto-HR	9	0	0.0015	0	0.0020

Fisher's exact test was used to determine if the fraction of total *RGC*s or all resistance candidate genes were significantly greater than expected

* *P* value < 0.05; [†] *P* value < 0.01; [‡] *P* value < 0.005

SSCP analysis to avoid problems due to interactions between paralogous sequences.

NBS-encoding sequences are the major type of *R*-genes in plants. Thirty-three new NBS-encoding sequences were amplified and sequenced from lettuce in an effort to identify NBS-LRR encoding sub-families for which we previously had few or no representative lettuce sequences. The currently available NBS-encoding sequences from *Lactuca* can be grouped into 20 distinct *RGC* sub-families; five of these were comprised only of sequences identified from the EST database and four were comprised only of sequences identified using degenerate oligonucleotide primers (Fig. 1). NBS-LRR encoding genes are often present in large clusters of closely related sequences (Bhattacharyya et al. 2005; Meyers et al. 2003; Shen et al. 1998); however, TIR-NBS-LRR and CC-NBS-LRR classes can be co-located together as reported for other species, e.g., *Arabidopsis* and Cassava (Lopez et al. 2003; Meyers et al. 1999).

In lettuce, an *RGC4* TIR-NBS-LRR encoding sequence, for example, clusters with two *RGC1* sequences encoding CC-NBS-LRR proteins on chromosome 3 (Figs. 1, 2). This is in contrast to *Glycine max*, *Medicago truncatula*, and *Prunus* sp. where no co-segregation of TIR-NBS-LRRs with CC-NBS-LRRs was observed (Kanazin et al. 1996; Lalli et al. 2005; Zhu et al. 2002). Other types of genes, such as receptor-like proteins, have also been shown to be clustered in the genomes of many species (Martin et al. 1993; Ronald et al. 1992; Thomas et al. 1998). This study may not have identified the clustering of non-NBS-LRR encoding resistance-like genes due to the relatively small number mapped in comparison to the NBS-LRR encoding sequences.

As expected, *RGC*s provide candidates for most dominant resistances, 25 of the 27, yet to be cloned. *RGC*s also co-localized with resistance QTL and a recessive resistance locus. Several studies have reported the co-localization of resistance QTL with NBS-LRR encoding genes (Calenge

et al. 2005; Lein et al. 2007; Paal et al. 2004; Ramalingam et al. 2003). Of the five resistance QTL positioned on the lettuce map (*RBQ1*, *RBQ2*, *RBQ3*; Jeuken and Lindhout 2002; *ANT2*, *ANT3*), all but *RBQ1* co-localized with *RGC*s. Although *RGC*s generally encode dominantly inherited resistances, *RRS1-R* in *Arabidopsis* is an example of a *RGC* conferring recessive resistance (Deslandes et al. 2002). The recessive gene for resistance to *Plasmopara lactucae-radicis* (*plr*) maps to a 13 cM interval with 1 CC-NBS-LRR and 5 TIR-NBS-LRR sequences and may be encoded by a recessive NBS-LRR gene in a manner similar to RRS1-R in *Arabidopsis* (Deslandes et al. 2002).

Several resistance phenotypes co-localize with non-NBS-LRR encoding candidate genes. Genes controlling resistance QTL have been identified as encoding receptor-like kinases and wound-inducible proteins in other species (Godiard et al. 2003; Llorente et al. 2005; Zhang et al. 2006). In lettuce, *RBQ1* co-localizes with the defense response candidates encoding defensin-fusion and allene oxide synthase proteins. Recessive resistances to viral pathogens often encode susceptibility factors required by the pathogen for successful propagation (Robaglia and Caranta 2006). There is also evidence that recessive resistance to bacterial pathogens may also be encoded by susceptibility factors (Iyer-Pascuzzi and McCouch 2007). Three recessive resistance genes were positioned on the lettuce map, *mo1*, *plr*, and *cor* (Kesseli et al. 1993; Moreno-Vazquez et al. 2003; Nicaise et al. 2003). The previously cloned *mo1* gene conferring resistance to LMV is encoded by eif4e, which in the context of LMV infection functions as a viral susceptibility factor (Nicaise et al. 2003). In addition to NBS-LRR encoding genes, *plr* also co-localized with the defense response related dirigent proteins (Burlat et al. 2001). The recessive bacterial resistance gene, *cor*, co-localized with the resistance pathway and defense response-associated genes, *RAP2.3*, *RCD1*, *CUL1*, and an epoxide hydrolase encoding gene (Büttner and Singh 1997; Gomi et al. 2003; Overmyer et al. 2000; Ren et al. 2005).

The candidate genes identified in this study that co-localize with resistance phenotypes provide numerous molecular markers in breeding for disease resistance. Historically, new *Dm* genes have been introduced sequentially and have a limited lifespan (Crute and Pink 1996). The durability of resistance should be improved by pyramiding *Dm* genes as well as introgressing a variety of *Dm* genes into different lettuce types to fragment selection pressures on the pathogen population (Michelmore 2003). Pyramiding *Dm* genes that are effective against all known isolates can only be achieved through MAS. By combining unlinked rather than linked resistance genes, population sizes in breeding programs can be reduced. However, where only tightly linked resistances are available for pyramiding, MAS can be utilized to select for rare recombinants that have resistance

genes in *cis*. By haplotyping germ plasm at resistance loci, novel sources of resistance can be distinguished and breeding programs can avoid repeatedly introgressing the same resistance gene from different sources (Kuang et al. 2008). Consequently, knowledge of the genetic positions of resistance phenotypes relative to each other and haplotyping of germplasm with *RGC* markers will aid in efficient breeding of lettuce cultivars with more durable resistances.

References

Albar L, Bangratz-Reyser M, Hebrard E, Ndjiondjop MN, Jones M, Ghesquiere A (2006) Mutations in the eIF(iso)4G translation initiation factor confer high resistance of rice to rice yellow mottle virus. Plant J 47:417–426

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Anonymous (2007) Crop values, 2006 summary. National Agricultural Statistics Service, USDA. http://usda.mannlib.cornell.edu/usda/current/CropValuSu/CropValuSu-02-15-2007.pdf

Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL (2002) The Pfam protein families' database. Nucleic Acids Res 30:276–280

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265:1856–1860

Belkhadir Y, Subramaniam R, Dangl JL (2004) Plant disease resistance protein signaling NBS-LRR proteins and their partners. Curr Opin Plant Biol 7:391–399

Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G, Moseyko N, Yoo D, Xu I, Zoeckler B, Montoya M, Miller N, Weems D, Rhee SY (2004) Functional annotation of the *Arabidopsis* genome using controlled vocabularies. Plant Physiol 135:745–755

Bernatsky R, Tanksley SD (1986) Genetics of actin-related sequences in tomato. Theor Appl Genet 72:314–321

Bhattacharyya MK, Narayanan NN, Gao H, Santra DK, Salimath SS, Kasuga T, Liu Y, Espinosa B, Ellison L, Marek L, Shoemaker R, Gijzen M, Buzzell RI (2005) Identification of a large cluster of coiled coil-nucleotide binding site-leucine rich repeat-type genes from the *Rps1* region containing *Phytophthora* resistance genes in soybean. Theor Appl Genet 111:75–86

Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, Weigel D (2007) Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol 5:e236

Burlat V, Kwon M, Davin LB, Lewis NG (2001) Dirigent proteins and dirigent sites in lignifying tissues. Phytochemistry 57:883–897

Buschges R, Hollricher K, Panstruga R, Simons G, Wolter M, Frijters A, van Daelen R, van der Lee T, Diergaarde P, Groenendijk J, Topsch S, Vos P, Salamini F, Schulze-Lefert P (1997) The barley *Mlo* gene: a novel control element of plant pathogen resistance. Cell 88:695–705

Büttner M, Singh KB (1997) *Arabidopsis thaliana* ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. Proc Natl Acad Sci USA 94:5961–5966

Calenge F, Van der Linden CG, Van de Weg E, Schouten HJ, Van Arkel G, Denancé C, Durel CE (2005) Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theor Appl Genet 110:660–668

Chin DB, Arroyo-Garcia R, Ochoa OE, Kesseli RV, Lavelle DO, Michelmore RW (2001) Recombination and spontaneous mutation at the major cluster of resistance genes in lettuce (*Lactuca sativa*). Genetics 157:831–849

Chisholm ST, Coaker G, Day B, Staskawicz BJ (2006) Host-microbe interactions: shaping the evolution of the plant immune response. Cell 124:803–814

Crute IR, Pink D (1996) Genetics and utilization of pathogen resistance in plants. Plant Cell 8:1747–1755

Datta K, Baisakh N, Thet KM, Tu J, Datta SK (2002) Pyramiding transgenes for multiple resistance in rice against bacterial blight, yellow stem borer and sheath blight. Theor Appl Genet 106:1–8

Davis RM, Subbarao KV, Raid RN, Kurtz EA (1997) Compendium of lettuce diseases. APS Press, St Paul

Deslandes L, Olivier J, Theulieres F, Hirsch J, Feng DX, Bittner-Eddy P, Beynon J, Marco Y (2002) Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive *RRS1-R* gene, a member of a novel family of resistance genes. Proc Natl Acad Sci USA 99:2404–2409

Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher C, Somssich I, Genin S, Marco Y (2003) Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. Proc Natl Acad Sci USA 100:8024–8029

Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CI, Ayliffe MA, Kobe B, Ellis JG (2006) Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. Proc Natl Acad Sci USA 103:8888–8893

Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14:755–763

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6 distributed by the author. Department of Genome Sciences. University of Washington, Seattle

Glazebrook J (2001) Genes controlling expression of defense responses in *Arabidopsis*-2001 status. Curr Opin Plant Biol 4:301–308

Godiard L, Sauviac L, Torii KU, Grenon O, Mangin B, Grimsley NH, Marco Y (2003) ERECTA, an LRR receptor-like kinase protein controlling development pleitropically affects resistance to bacterial wilt. Plant J 36:353–365

Gomi K, Yamamato H, Akimitsu K (2003) Epoxide hydrolase: a mRNA induced by the fungal pathogen Alternaria alternata on rough lemon (Citrus jambhiri Lush). Plant Mol Biol 53:189–199

Grant SR, Fisher EJ, Chang JH, Mole BM, Dangl JL (2006) Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. Annu Rev Microbiol 60:425–429

Grube RC, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the solanaceae. Genetics 155:873–887

Grube RC, Wintermantel WM, Hand P, Aburomia R, Pink DA, Ryder EJ (2005) Genetic analysis and mapping of resistance to lettuce dieback: a soilborne disease caused by tombusviruses. Theor Appl Genet 110:259–268

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

Hulbert SH, Michelmore RW (1985) Linkage analysis of genes for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). Theor Appl Genet 70:520–528

Iyer-Pascuzzi AS, McCouch SR (2007) Recessive resistance genes and the *Oryza sativa–Xanthomonas oryzae* pv. oryzae pathosystem. Mol Plant Microbe Interact 20:731–739

Jain A, Ariyadasa R, Kumar A, Srivastava MN, Mohan M, Nair S (2004) Tagging and mapping of a rice gall midge resistance gene, *Gm8*, and development of SCARs for use in marker-aided selection and gene pyramiding. Theor Appl Genet 109:1377–1384

Jeuken MJ, Lindhout P (2002) *Lactuca saligna*, a non-host for lettuce downy mildew (*Bremia lactucae*), harbors a new race-specific *Dm* gene and three QTLs for resistance. Theor Appl Genet 105:384–391

Jia Y, McAdams SA, Bryan GT, Hershey HP, Valent B (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J 19:4004–4014

Jones JD, Dangl JL (2006) The plant immune system. Nature 444:323–329

Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JD (1994) Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science 266:789–793

Kanazin V, Marek LF, Shoemaker RC (1996) Resistance gene analogs are conserved and clustered in soybean. Proc Natl Acad Sci USA 93:11746–11750

Kesseli RV, Witsenboer H, Stanghellini M, Vandermark G, Michelmore RW (1993) Recessive resistance to *Plasmopara lactucae-radicis* is mapped by bulked segregant analysis to a cluster of dominant disease resistance genes in lettuce. Mol Plant Microbe Interact 6:722–728

Kesseli RV, Paran I, Michelmore RW (1994) Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. Genetics 136:1435–1446

Kuang H, van Eck HJ, Sicard D, Michelmore RW, Nevo E (2008) Evolution and genetic population structure of prickly lettuce (*Lactuca serriola*) and its *RGC2* resistance gene cluster. Genetics 178:1547–1558

Lalli DA, Decroocq V, Blenda AV, Schurdi-Levraud V, Garay L, Le Gall O, Damsteegt V, Reighard GL, Abbott AG (2005) Identification and mapping of resistance gene analogs (RGAs) in *Prunus*: a resistance map for *Prunus*. Theor Appl Genet 111:1504–1513

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Landry BS, Kesseli RV, Farrara B, Michelmore RW (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331–337

Lein JC, Asbach K, Tian Y, Schulte D, Li C, Koch G, Jung C, Cai D (2007) Resistance gene analogues are clustered on chromosome 3 of sugar beet and cosegregate with QTL for *rhizomania* resistance. Genome 50:61–71

Liu J, Liu X, Dai L, Wang G (2007) Recent progress in elucidating the structure, function, and evolution of disease resistance genes in plants. J Genet Genomics 34:765–776

Lopez CE, Zuluaga AP, Cooke R, Delseny M, Tohme J, Verdier V (2003) Isolation of resistance gene candidates (*RGCs*) and characterization of an *RGCs* cluster in cassava. Mol Genet Genomics 269:658–671

Llorente F, Alonso-Blanco C, Sánchez-Rodriguez C, Jorda L, Molina A (2005) ERECTA receptor-like kinase and heterotrimeric G protein from *Arabidopsis* are required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. Plant J 43:165–180

Mackey D, Holt BF 3rd, Wiig A, Dangl JL (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in *Arabidopsis*. Cell 108:743–754

Maisonneuve B, Bellec Y, Anderson P, Michelmore RW (1994) Rapid mapping of two genes for resistance to downy mildew from *Lact-*

uca serriola to existing cluster of resistance genes. Theor Appl Genet 89:96–104

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262:1432–1436

McHale L, Tan X, Koehl P, Michelmore RW (2006) Plant NBS-LRR proteins: adaptable guards. Genome Biol 7:212

Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelmore RW (1998) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10:1817–1832

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J 20:317–332

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell 15:809–834

Michelmore RW (1995) Molecular approaches to manipulation of disease resistance genes. Ann Rev Phytopathol 15:393–427

Michelmore RW (2003) The impact zone: genomics and breeding for durable disease resistance. Curr Opin Plant Biol 6:397–404

Montesclaros L, Nicol N, Ubalijora E, Leclerc-Potvin C, Ganivet L, Laliberte JF, Fortin MG (1997) Response to potyvirus infection and genetic mapping of resistance loci to potyvirus infection in *Lactucae*. Theor Appl Genet 94:941–946

Moreno-Vazquez S, Ochoa OE, Faber N, Chao S, Jacobs JM, Maisonneuve B, Kesseli RV, Michelmore RW (2003) SNP-based codominant markers for a recessive gene conferring resistance to corky root rot (*Rhizomonas suberifaciens*) in lettuce (*Lactuca sativa*). Genome 46:1059–1069

Naik S, Hampson C, Gasic K, Bakkeren G, Korban SS (2006) Development and linkage mapping of E-STS and RGA markers for functional gene homologues in apple. Genome 49:959–968

Nicaise V, German-Retana S, Sanjuan R, Dubrana MP, Mazier M, Maisonneuve B, Candresse T, Caranta C, LeGall O (2003) The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the potyvirus lettuce mosaic virus. Plant Physiol 132:1272–1282

Nieto C, Morales M, Orjeda G, Clepet C, Monfort A, Sturbois B, Puigdomenech P, Pitrat M, Caboche M, Dogimont C, Garcia-Mas J, Aranda MA, Bendahmane A (2006) An *eIF4E* allele confers resistance to an uncapped and non-polyadenylated RNA virus in melon. Plant J 48:452–462

Ochoa O, Delp B, Michelmore RW (1987) Resistance in *Lactuca* spp. to *Microdochium panattoniana* (lettuce anthracnose). Euphytica 36:609–614

Overmyer K, Tuominen H, Kettunen R, Betz C, Langebartels C, Sandermann H Jr, Kangasjärvi J (2000) Ozone-sensitive *Arabidopsis rcd1* mutant reveals opposite roles for ethylene and jasmonate signaling pathways in regulating superoxide-dependent cell death. Plant Cell 12:1849–1862

Paal J, Henselewski H, Muth J, Meksem K, Menendez CM, Salamini F, Ballvora A, Gebhardt C (2004) Molecular cloning of the potato *Gro1-4* gene conferring resistance to pathotype Ro1 of the root cyst nematode *Globodera rostochiensis*, based on a candidate gene approach. Plant J 38:285–297

Pan Q, Liu Y-S, Budai-Hadrian O, Sela M, Carmel-Goren L, Zamir D, Fluhr R (2000) Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and *Arabidopsis*. Genetics 155:309–322

Panstruga R (2005) Serpentine plant MLO proteins as entry portals for powdery mildew fungi. Biochem Soc Trans 33:389–392

Paran I, Kesseli R, Michelmore R (1991) Identification of restriction fragment length polymorphism and random amplified polymorphic DNA markers linked to downy mildew resistance genes in lettuce, using near-isogenic lines. Genome 34:1021–1027

Plocik A, Layden J, Kesseli R (2004) Comparative analysis of NBS domain sequences of NBS-LRR disease resistance genes from sunflower, lettuce, and chicory. Mol Phylogenet Evol 31:153–163

Radwan O, Gandhi S, Heesacker A, Whitaker B, Taylor C, Plocik A, Kesseli R, Kozik A, Michelmore RW, Knapp SJ (2008) Genetic diversity and genomic distribution of homologs encoding NBS-LRR disease resistance proteins in sunflower. Mol Genet Genomics. Epub ahead of print

Ramalingam J, Vera Cruz CM, Kukreja K, Chittoor JM, Wu JL, Lee SW, Baraoidan M, George ML, Cohen MB, Hulbert SH, Leach JE, Leung H (2003) Candidate defense genes from rice, barley, and maize and their association with qualitative and quantitative resistance in rice. Mol Plant Microbe Interact 16:14–24

Ren C, Pan J, Peng W, Genschik P, Hobbie L, Hellmann H, Estelle M, Gao B, Peng J, Sun C, Xie D (2005) Point mutations in *Arabidopsis Cullin1* reveal its essential role in jasmonate response. Plant J 42:514–524

Robaglia C, Caranta C (2006) Translation initiation factors: a weak link in plant RNA virus infection. Trends Plant Sci 11:40–45

Ronald PC, Albano B, Tabien R, Abenes L, Wu KS, McCouch S, Tanksley SD (1992) Genetic and physical analysis of the rice bacterial blight disease resistance locus, *Xa21*. Mol Gen Genet 236:113–120

Rooney HC, Van't Klooster JW, van der Hoorn RA, Joosten MH, Jones JD, de Wit PJ (2005) *Cladosporium Avr2* inhibits tomato *Rcr3* protease required for *Cf-2*-dependent disease resistance. Science 308:1783–1786

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, pp 365–386

Ruffel S, Gallois JL, Moury B, Robaglia C, Palloix A, Caranta C (2006) Simultaneous mutations in translation initiation factors *eIF4E* and *eIF(iso)4E* are required to prevent pepper veinal mottle virus infection of pepper. J Gen Virol 87:2089–2098

Shao F, Golstein C, Ade J, Stoutemyer M, Dixon JE, Innes RW (2003) Cleavage of *Arabidopsis* PBS1 by a bacterial type III effector. Science 301:1230–1233

Shen KA, Meyers BC, Islam-Faridi MN, Chin DB, Stelly DM, Michelmore RW (1998) Resistance gene candidates identified using PCR with degenerate primers map to resistance genes clusters in lettuce. Mol Plant Microbe Interact 11:815–823

Shen KA, Chin DB, Arroyo-Garcia R, Ochoa OE, Lavelle DO, Wroblewski T, Meyers BC, Michelmore RW (2002) *Dm3* is one member of a large constitutively expressed family of nucleotide binding site-leucine-rich repeat encoding genes. Mol Plant Microbe Interact 15:251–261

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804–1806

Stam P, Van Ooijen JW (1995) Joinmap version 2.0: software for the calculation of genetic linkage maps. Plant Res Int

Tan X, Meyers BC, Kozik A, West MA, Morgante M, St Clair DA, Bent AF, Michelmore RW (2007) Global expression analysis of nucleotide binding site-leucine rich repeat-encoding and related genes in *Arabidopsis*. BMC Plant Biol 7:56

Thomas CM, Dixon MS, Parniske M, Golstein C, Jones JD (1998) Genetic and molecular analysis of tomato Cf genes for resistance to *Cladosporium fulvum*. Philos Trans R Soc Lond B Biol Sci 353:1413–1424

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Timms L, Jimenez R, Chase M, Lavelle D, McHale L, Kozik A, Lai Z, Heesacker A, Knapp S, Rieseberg L, Michelmore R, Kesseli R (2006) Analyses of synteny between *Arabidopsis thaliana* and species in the Asteraceae reveal a complex network of small syntenic segments and major chromosomal rearrangements. Genetics 173:2227–2235

Truco MJ, Antonise R, Lavelle D, Ochoa O, Kozik A, Witsenboer H, Fort SB, Jeuken MJ, Kesseli RV, Lindhout P, Michelmore RW, Peleman J (2007) A high-density, integrated genetic linkage map of lettuce (*Lactuca* spp.). Theor Appl Genet 115:735–746

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Wang S, Basten CJ, Zeng ZB (2005) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh. http://www.statgen.ncsu.edu/qtl/crt/WQTL.htm

Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C, Baker B (1994) The product of the tobacco mosaic virus resistance gene *N*: similarity to toll and the interleukin-1 receptor. Cell 78:1101–1115

Wroblewski T, Tomczak A, Michelmore RW (2005) Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. Plant Biotechnol J 3:259–273

Wroblewski T, Piskurewicz U, Tomczak A, Ochoa O, Michelmore R (2007) Silencing of the major family of NBS-LRR-encoding genes in lettuce results in the loss of multiple resistance specificities. Plant J 51:803–818

Zhang LP, Khan A, Nino-Liu D, Foolad MR (2002) A molecular linkage map of tomato displaying chromosomal locations of resistance gene analogs based on a *Lycopersicon esculentum* × *Lycopersicon hirsutum* cross. Genome 45:133–146

Zhang Z, Ober JA, Kliebenstein DJ (2006) The gene controlling the quantitative trait locus EPITHIOSPECIFIER MODIFIER1 alters glucosinolate hydrolysis and insect resistance in *Arabidopsis*. Plant Cell 18:1524–1536

Zhu H, Cannon SB, Young ND, Cook DR (2002) Phylogeny and genomic organization of the TIR and non-TIR NBS-LRR resistance gene family in *Medicago truncatula*. Mol Plant Microbe Interact 15:529–539