CALIFORNIA LETTUCE RESEARCH PROGRAM

April 1, 2007, to March 31, 2008

GENETIC VARIATION IN LETTUCE

Richard W. Michelmore
The Genome Center and
The Department of Plant Sciences
University of California, Davis
rwmichelmore@ucdavis.edu

SUMMARY:

Novel methods for detecting, analyzing, and manipulating genetic variation are being applied to lettuce. The aim of this project is to ensure that lettuce does not lag behind other crops in benefiting from the application of genomic and biotechnological techniques. We have continued four projects: (i) Introduction of genes into lettuce using *Agrobacterium tumefaciens* and analysis of their expression. (ii) Cloning and characterization of disease resistance genes. (iii) Genetic mapping using a variety of molecular markers with the goal of locating most of the disease resistance genes known in lettuce. (iv) Comparative genomics to identify candidate genes controlling horticultural traits. Over 200,000 ESTs have been generated representing over 35,000 unigenes and at least half of all genes in lettuce. These sequences have been mined for candidate genes for agriculturally important traits. Candidate genes have been mapped relative to loci for disease resistance, development, and horticulturally important traits. We are testing for function in resistance for candidate genes which cosegregate with disease resistance phenotypes. We have shown using post-transcriptional gene silencing that *Dm14, Dm18, Dm16,* and *Ra* are probably encoded by sequences similar to *Dm3* and that *Dm5/8* is encoded by a CC-NBS-LRR encoding gene. We have developed a lettuce affymetrix genotyping chip with over 6 million overlapping probes representing the > 35,000 unigenes from lettuce. We are in the process of generating an ultra-high density genetic map based on genotyping from hybridizations to the lettuce affymetrix chip (http://chiplett.ucdavis.edu). We have developed and are curating a database for the Compositae (http://compositdb.ucdavis.edu/) that includes genetic, molecular marker, cultivar, and sequence data for lettuce as well as information on lettuce researchers. We are developing a new lettuce database and genetic map viewer to facilitate analysis of phenotypic, genetic, and genomic data.

CALIFORNIA LETTUCE RESEARCH PROGRAM
April 1, 2007 to March 31, 2008

PROJECT TITLE: **GENETIC VARIATION IN LETTUCE**

PRINCIPAL INVESTIGATOR: **Richard W. Michelmore**
The Genome Center and
The Department of Plant Sciences
University of California, Davis
rwmichelmore@ucdavis.edu

COOPERATING PERSONNEL: **María José Truco,**
Leah McHale,
Dean Lavelle,
Marilena Christopoulou
Tadeusz Wroblewski,
Oswaldo Ochoa
Hans van Leuwen
Alex Kozik
Huaqin Xu
UC Davis Genome Center and
The Department of Plant Sciences
University of California, Davis
Kent Bradford
Allen van Deynze
Department of Plant Sciences, UC Davis
Ryan Hayes
UDSA-ARS, Salinas

OBJECTIVES:

To develop and apply new methods for detecting, analyzing and manipulating variation in lettuce. We continue to pursue these objectives in four projects:

1) Introduction of genes into lettuce using *Agrobacterium tumefaciens* and testing transgene activity.

2) Molecular cloning of disease resistance genes.

3) Development of a detailed genetic map using a variety of markers and identification of reliable readily-assayed markers linked to disease resistance genes.

4) Utilization of comparative genomics to identify candidate genes controlling horticultural traits.

In the first three projects we are emphasizing either novel forms of disease resistance or increasing the efficiency of selection for disease resistant genotypes. The fourth project includes

a wide range of horticultural traits. Some of these studies are currently funded from Federal grants, the UC Biostar Program and several seed companies. All projects were initiated with CLRP funds and application of the results to lettuce improvement is supported by CLRP funds. All projects impact both crisphead and leafy improvement.

PROCEDURES AND RESULTS:

Transgene Expression in Lettuce

Our transgenic studies remain a lower priority than the other objectives, in part because commercial application of transgenes in lettuce is not an immediate priority. No experiments specifically addressing this objective were conducted the past year. Introduction of genes into lettuce using *A. tumefaciens* is routine. However, introduced genes that can be expressed well in other plant species are expressed poorly over multiple generations in lettuce. Factors influencing stable transgene expression over multiple generations in lettuce are not well understood and require detailed investigation. We still do not have a protocol for reliable and predictable expression of transgenes in lettuce. We continue to generate dozens of transgenics as parts of other projects. These are providing data on transgene expression and stability. However, we do not currently have projects specifically focused on optimizing the stability of transgene expression over multiple generations.

Resistance candidate gene identification and mapping

We have identified over 700 candidate resistance genes by mining the lettuce EST database (http://cgpdb.ucdavis.edu) and amplification of conserved domains using PCR with degenerate oligonucleotides primers for genes known to function in disease resistance. One hundred sixty seven genes involved in pathogen recognition, 272 disease resistance-related signal transduction, and 304 defense response have so far been identified. Of these candidate genes, 291 have been mapped. One hundred twenty of the pathogen recognition-type genes have high similarity to NBS-LRR encoding genes. Comparisons of these 120 sequences with over 1600 NBS sequences mined from Genbank reveals that the lettuce NBS sequences fall into distinct groups that can be separated into at least 20 families. Over half of the NBS-LRR encoding genes map to resistance phenotypes.

Mapped candidate genes are being placed onto the consensus map relative to loci for resistance to as many diseases as possible, particularly downy mildew, LMV, corky root, and anthracnose. So far, a total of 291 candidate resistance genes have been mapped relative to 36 phenotypic resistance loci (Fig. 1). We are continuing the functional analysis of clusters of co-segregating phenotypic resistances and candidate genes in detail (see below; Table 1).

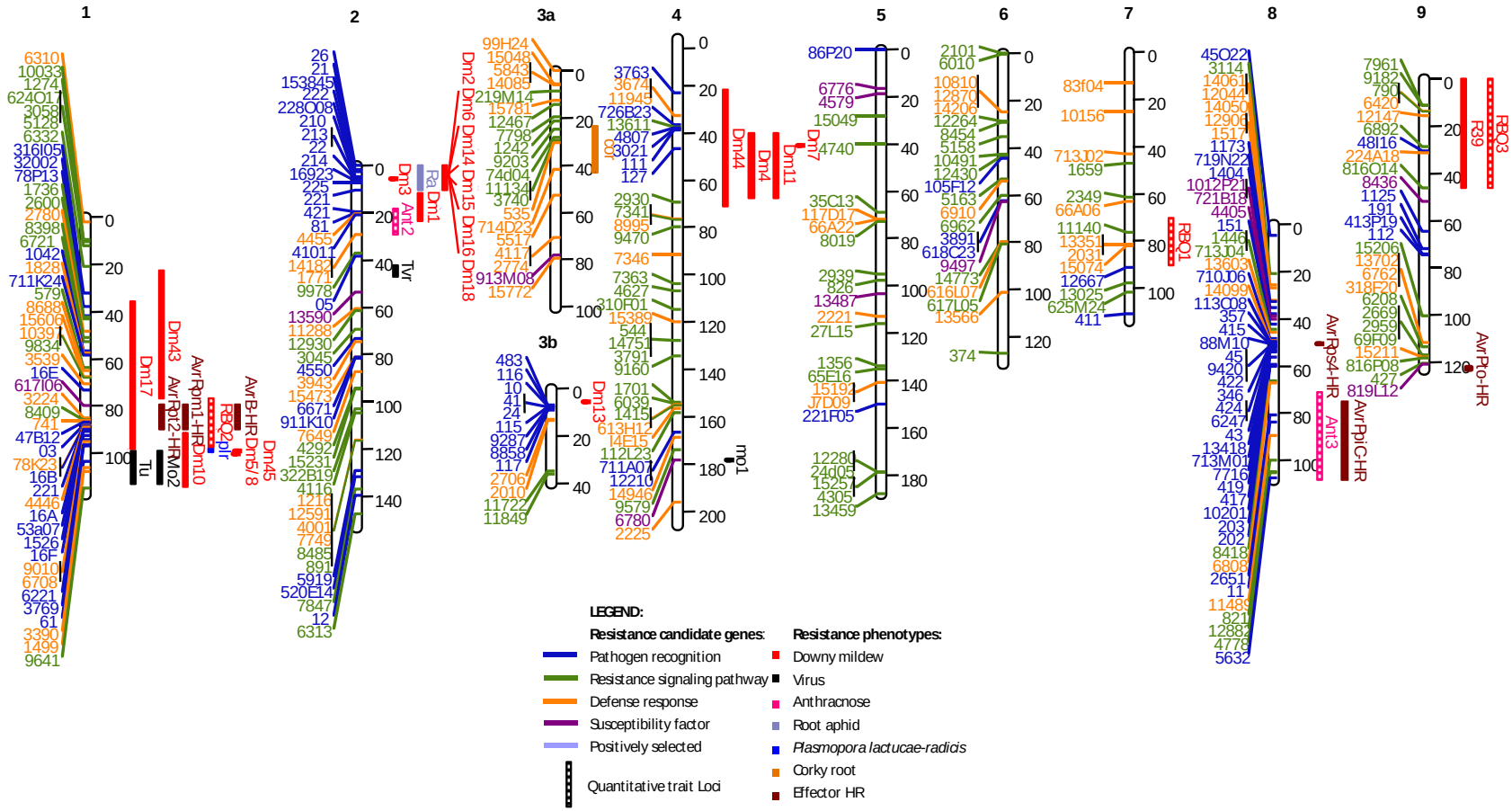
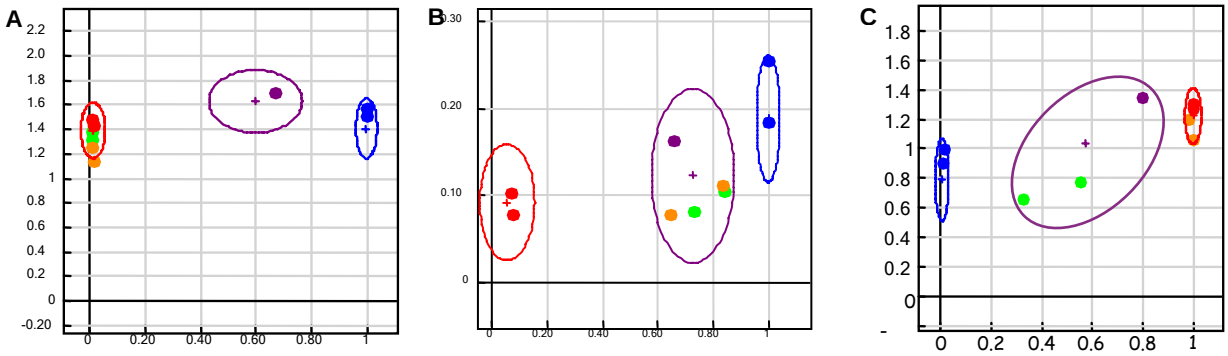


Figure 1. Lettuce genetic map displaying 291 mapped candidate resistance genes and 36 resistance phenotypes.

Mapping of resistance phenotypes

In order to increase the efficiency of bulked segregant analysis, bulks of resistant and susceptible progeny from a BC_2S_1 population segregating for resistance to downy mildew were genotyped with an Illumina GoldenGate SNP assay oligonucleotide pool designed to assay SNPs in candidate resistance genes (see above). A population derived from a BC_2S_1 cross between *L. sativa cv*. Salinas (recurrent parent) and *L. saligna* acc. UC96US2509 (donor parent) was evaluated for resistance to isolates prevalent in California (see Crisphead Report, 2008). Three markers linked to the resistance gene cluster on chromosome 1 were different between the resistant and susceptible bulks and were therefore identified as linked to the new resistance gene (Fig. 2C).

Figure 2. Bulked segregant analysis of a novel downy mildew resistance using the Illumina GoldenGate SNP assay. Examples of the three patterns observed. **Recurrent parent** allele is in **red**. **Donor parent** allele is in **blue**. **Parental heterozygote** allele is in **purple**. **Resistant bulks** are in **green** and **susceptible bulks** are in **orange**. A. Resistant and susceptible bulks have the heterozygote allele at genetic regions unlinked to resistance and the BC_2 line is not fixed for the recurrent parent. B. Resistant and susceptible bulks have the recurrent parent allele at genetic regions unlinked to resistance and the BC_2 line is fixed for the recurrent parent. C. Resistant bulks have the heterozygote allele and susceptible bulks have the recurrent allele at genetic regions linked to resistance.



Twelve markers at the bottom of chromosome 1 in the consensus map (Fig. 1) were then analyzed by Single-Strand Conformation Polymorphism (SSCP) for polymorphism between Salinas and UC23US2509. Four of these were polymorphic between these parents, three of which had been previously identified as SNPs in the BSA. These three SSCP markers segregated on the Salinas x UC23US2509 BC_2S_1 population. The fourth marker was fixed for the Salinas allele when assayed on the BC_2S_1 population, indicating that only a small genomic region had been introgressed from the donor parent. The new downy mildew resistance, designated *Dm45*, was mapped to a cluster of resistance phenotypes and resistance candidate genes encoding NBS-LRR proteins. These sequences are suitable markers for selection of *Dm45*.

Functional analysis of resistance candidate genes

Post-transcriptional gene silencing (PTGS, RNAi) assays involve expression of double stranded RNA fragments of target genes leading to the degradation of the mRNA transcribed from the endogenous genes and consequently to the loss of the target genes' activity. RNAi to silence *RGC2B* (encoding *Dm3*) was used to show that *Dm14*, *Dm16*, *Dm18* and *Ra* are members of the *RGC2* family (Wroblewski *et al*., 2007 Plant J. 51:803-18). We have been generating RNAi transgenics for other resistance gene candidates (*RGC*s) that cosegregate with resistance phenotypes in order to test the resistance function of each *RGC* (Fig. 1, Table 1). These RNAi lines are being used as tester stocks in crosses to resistant lines to identify genes involved in resistance to several diseases.

Table 1. RNAi constructs for 15 RGCs that cosegregate with a resistance phenotype have been transformed into Cobham Green. T_1 lines have been crossed to multiple resistant cultivars. Silencing is monitored by analyzing activity of the GUS reporter and progeny are tested for loss of phenotype. The protein domain encoded by the gene segment used to generate the RNAi construct is indicated.

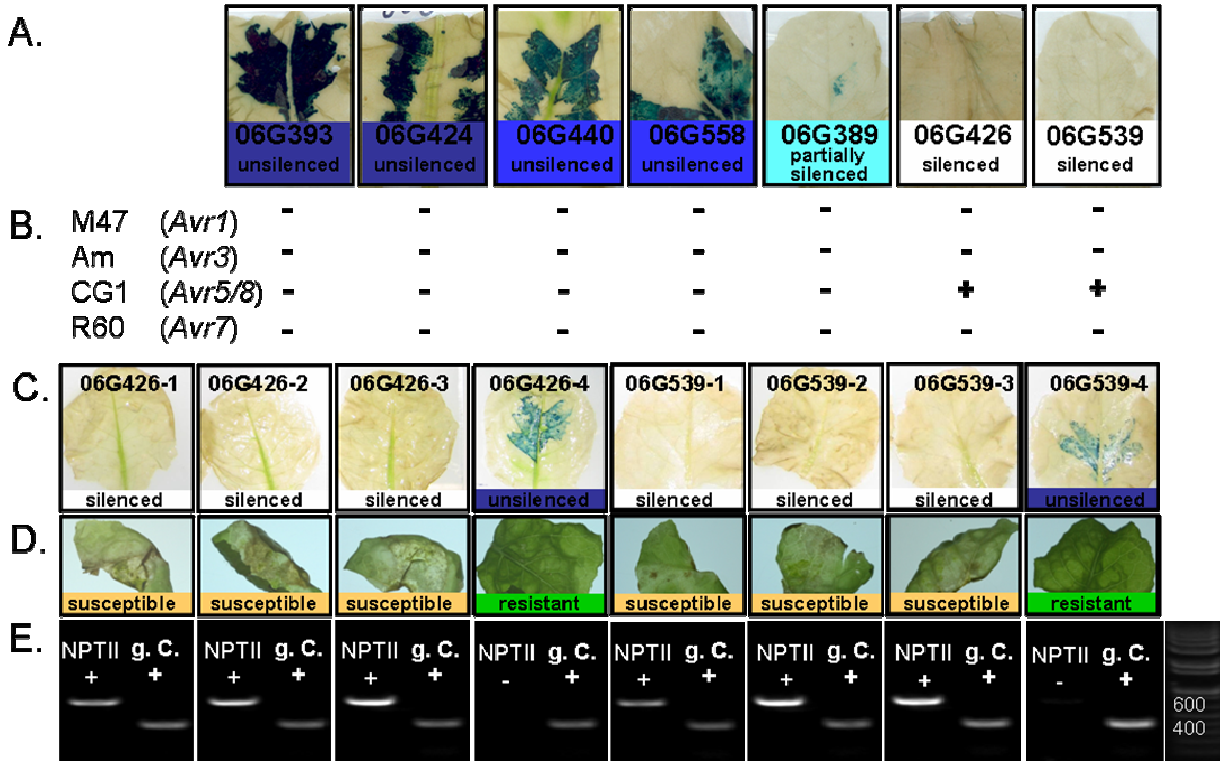
Resistance Locus	RGC Type	RGC family	Protein domain	Phenotypes silenced
Dm3 cluster	CC-NBS-LRR	*RGC2*	LRR	*Dm3, Dm14, Dm16, Dm18, Ra*
Dm3 cluster	CC-NBS-LRR	*RGC2*	NBS	*Dm3, Dm6, Dm18*
Dm5/8 cluster	CC-NBS-LRR		LRR	*Dm5/8*
Dm5/8 cluster	TIR-NBS-LRR	*RGC16*	NBS	
Dm5/8 cluster	CC-NBS-LRR		NBS	
Dm5/8 cluster	CC-NBS-LRR		NBS	
Dm5/8 cluster	TIR-NBS-LRR		TIR	
Dm5/8 cluster	TIR-NBS-LRR		LRR	
AvrRps4-HR	CC-NBS-LRR	*RGC4*	NBS	
AvrRps4-HR	CC-NBS-LRR		LRR	
AvrRps4-HR	CC-NBS-LRR		LRR	
AvrRps4-HR	CC-NBS-LRR		LRR	
Dm17 cluster	TM-LRR		LRR	
Dm17 cluster	TIR-NBS-LRR	*RGC16*	NBS	
AvrPpiC-HR, Ant3	TIR-NBS-LRR		NBS	

Stable RNAi lines of the NBS-LRR encoding genes, *RGC221* were generated to test for their function in *Dm5/8* (in Diana). Cv. Diana, carrying *Dm1*, *Dm3*, *Dm5/8* and *Dm7*, and was transformed with RNAi constructs comprising inverted repeats from GUS and the LRR-encoding region of *RGC221*. T_1 plants were assayed for loss of GUS expression and loss of resistance to downy mildew. Two GUS silenced lines were identified (Fig. 2A).

Both T_1 transgenic Diana lines silenced for *RGC221* showed a loss of the resistance conferred by *Dm5/8*; these T_1s retained downy mildew resistance conferred by *Dm1*, *Dm3*, and *Dm7* and

therefore the loss of resistance was specific to *Dm5/8* (Figs. 2A and 2B). Loss of *Dm5/8* resistance was confirmed in the T_2 generation and co-segregated with lack of GUS expression (Figs. 2C and 2D). Amplification of the *nptII* gene in T_2 plants showed that a lack of GUS expression co-segregated with the transgene for all four constructs (Fig 2E). These RNAi results therefore indicate that *Dm5/8* is conferred by *RGC221* or a related sequence.

Figure 2. **A**. Seven Diana T_1 lines with the *RGC221* silencing construct were assayed for silencing of GUS activity. The lack of blue staining indicates successful silencing. **B**. The reactions of leaf discs from each T_1 line inoculated with isolates of *B. lactucae* to test for *Dm1*, *Dm3*, *Dm5/8*, and *Dm7*. + = susceptible reaction. - = resistant reaction. **C**. Four progeny from each of two T_1 lines, 06G426 and 06G539, silenced for *RGC221*, were tested for silencing of GUS expression. **D**. Detached leaf assays of 06G426 and 06G539 T_2 plants with *B. lactucae* isolate CG1 (carrying *Avr5/8*). **E**. In 06G426 and 06G539 T_2 plants, PCR amplification of *nptII* (600 bp) indicating presence or absence of transgene and a genomic control (g. C.) amplification (400 bp).



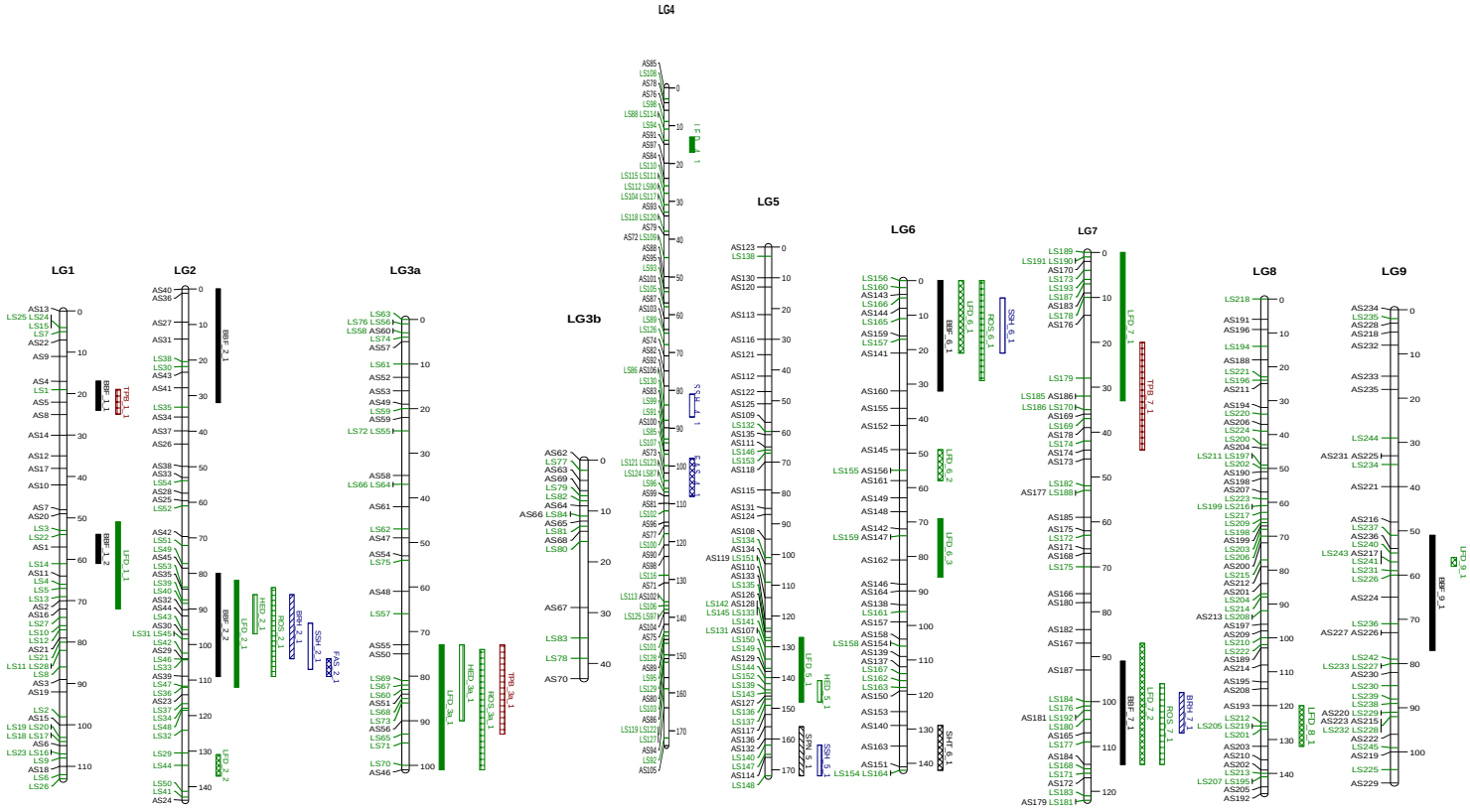
Mapping of candidate genes relative to horticultural traits

Candidate genes have been mapped relative to horticultural traits in our core $F_{7:8}$ RIL mapping population derived from a cross between *L. sativa* cv. Salinas and *L. serriola* acc. UC96US23. One hundred and fourteen candidate genes in lettuce with similarity to 93 genes of known function in Arabidopsis were mapped (Table 2). The majority (~90%) of these candidate genes were sequence verified and mapped with the Illumina GoldenGate® SNP assay. Additional markers for use with several technologies exist for many of candidate genes and will be made available at the project website (http://cgpdb.ucdavis.edu). QTL analysis performed on this population has identified chromosomal regions involved with horticultural traits such as

flowering time and leaf shape among others that are associated with 33 of these mapped candidate genes (Fig. 3).

Table 2. Ninety-three genes identified from Arabidopsis that have been mapped in $F_{7:8}$ RILs representing 114 different loci in lettuce. Arabidopsis gene identifiers are from TAIR (www.arabidopsis.org).

Mapped Candidate ESTs with Similarity to Arabidopsis Genes				
AFO	AT1G14340	ELF5	KAN	RPT3
AGL2	AT4G24760	EMF1	KNAT3	SEU
AGL20	AT4G31020	EMF2	KNAT7	SGT1B
AGL6	ATH1	ETR1	LBD25	SPY
AGL8	ATHB-15	FCA	LD	STM
AGL9	ATML1	FDH	MAF3	SVP
AGO1	BAM1	FIE	NPH3	TCP24-like
AGO1-like	BLH1	FKF1	PAP2	TCP4
AGO4-like	BP	FLC	PAT1	TIR1
AHK2	CLF	FLD	PHOT1	TOC1
AHK3	COL16	FT	PHYA	TOR
ANL2	COL4	FVE	PHYB	TRY
ANT	COL9	GA2	PHYC	TSL
AP1	COP1	GA3	PHYE	VIP4
AP2	CRC	GAI	PIE1	WUS
AP3	CRE1	GI	PLT1	ZLL
AS1	CRY1	GL2	PRR5	ZTL
ASK2	CRY2	HEN1	PRR7	



Abbreviation	*Trait*	*Abbreviation*	*Trait*
BBF	**Bolting, budding, flowering**	SSH	**Side-shoots**
LFD	**Leaf Development**	FAS	**Fasciation**
HED	**Heading**	TPB	**Tipburn**
ROS	**Rosette**	SPN	**Spines**
BRH	**Branching**	SHT	**Shattering**

Figure 3. Lettuce genetic map displaying 245 lettuce based markers (LS) and 236 anonymous genomic markers (AS) with QTL for 10 horticultural traits. 114 of the LS loci represent 93 Arabidopsis candidate genes (Table 2).

Massively parallel genetic analysis using an Affymetrix gene chip

In collaboration with Dr. A. van Deynze and Keygene, with support from the UC BioStar program, Enza Zaden, Rijk Zwaan and Vilmorin, we have developed a 6.6 million oligonucleotide Affymetrix array for high-throughput massively parallel genotyping of lettuce. The lettuce chip contains sequences for detecting polymorphisms in approximately 35,000 unigenes in lettuce and will allow for rapid SNP discovery, genotyping, mapping, and gene expression analysis. We have completed hybridizations in duplicate to genomic DNA of 115 RILs from our core Salinas x *L. serriola* UC96US23 mapping population. We have developed custom algorithms and scripts in order to process the data generated from the chip. We are now in the process of developing an ultra-dense genetic map based on transcribed sequences.

We have also hybridized in triplicate a diverse array of 50 *Lactuca* accessions to the Affymetrix chip. This hybridization data will aid in SNP discovery between parents of mapping populations and breeding lines. In addition, we are using the Affymetrix chip for bulked segregant analysis to identify genes linked to disease resistance genes (see CLRP Leafy Report, 2008). Bulks of DNA from plants resistant and susceptible to corky root have been hybridized to the lettuce Affymetrix chip. This should identify numerous unigenes tightly linked to corky root resistance.

The Compositae Database (Compositdb)

We have continued to curate electronic databases for lettuce. The Compositdb contains several searchable databases for lettuce accessible through http://compositdb.ucdavis.edu/ and is the main portal for distributing information generated by the Compositae Genome Project (CGP). Lettdb provides genetic and DNA sequence information as well as gel images for markers and descriptions and sequences of RFLP probes and oligonucleotide primer sequences for PCR-based markers. Lettcv archives extensive genetic, passport and performance data on lettuce cultivars. The CGP database contains extensive EST sequence information and related information as well as links to lettuce linkage maps (lettuce genetic map viewer) and easy to access marker information. A new database, morphodb, is being generated to archive and provide access to phenotypic information on *Lactuca* species.