

CALIFORNIA LEAFY GREENS RESEARCH PROGRAM

April 1, 2011, to March 31, 2012

GENETIC VARIATION IN LETTUCE

Richard W. Michelmore

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

SUMMARY:

We continue to apply novel methods for detecting, analyzing, and manipulating genetic variation in 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) Transgene expression in lettuce. (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. Our efforts have mainly been focused on projects (ii) through (iv). We continue to make extensive use of new high-throughput sequencing and marker technologies. Sequences have been and are being mined for candidate genes for agriculturally important traits. Candidate genes have been mapped relative to over 60 loci for disease resistance, development, and horticulturally important traits. Candidate genes which co-segregated with disease resistance phenotypes are being tested for function using RNA interference (RNAi) and several causal genes have been identified. We have utilized a custom lettuce Affymetrix genotyping chip representing >35,000 unigenes from lettuce and developed an ultra-high density genetic map of ~14,000 loci. The genome of lettuce has been sequenced and assembled and annotation is underway. We have developed and are curating several databases for lettuce as part of the Compositae Genome and Genes for Growers Projects that include genetic, molecular marker, cultivar, phenotypic and sequence data for lettuce.

CALIFORNIA LEAFY GREENS RESEARCH PROGRAM

April 1, 2011 to March 31, 2012

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**
Dean Lavelle
Marilena Christopoulou
Tadeusz Wroblewski
Oswaldo Ochoa
Alex Kozik
Huaqin Xu
Sebastian Reyes Chin Wo
UC Davis Genome Center and
The Department of Plant Sciences
University of California, Davis
Ryan Hayes
Ivan Simko
UDSA-ARS, Salinas
Richard Smith
Michael Cahn
UC Cooperative Extension, Monterey County
Krishna Subbarao
Thomas Gordon
The Department of Plant Pathology
University of California, Davis

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) Analysis of transgenes in lettuce.
- 2) Molecular cloning of genes for disease resistance and other horticultural traits.
- 3) Development of a detailed genetic map and identification of reliable, readily assayed markers linked to disease resistance genes.
- 4) Utilization of comparative genomics to identify candidate genes controlling horticultural traits and development of robust molecular markers for them.

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 and gifts from 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 studies on transgene expression remain a low priority, in part because there is not a major breeding objective that can only be addressed by transgenic lettuce and therefore commercial application of transgenes in lettuce is not an immediate priority. Introduction of genes into lettuce using *A. tumefaciens* is routine; in earlier experiments, however, transgenes that expressed well in some other plant species were expressed poorly after multiple generations in lettuce. Factors influencing the stability of transgene expression over multiple generations in lettuce are not well understood and require further investigation. We do not currently have projects specifically focused on transgene expression in lettuce. However, we continue to generate dozens of transgenics as parts of other projects (see below). These provide data on transgene expression and stability.

Resistance Candidate Gene Identification and Mapping

Loci for resistance to as many diseases as possible, particularly downy mildew (DM), LMV, corky root, *Fusarium* and *Verticillium* wilts, and big vein, are being placed onto the consensus genetic map and aligned with the ultra-dense map (see below). So far, a total of 54 phenotypic resistance loci have been mapped relative to candidate resistance genes. The majority of resistance phenotypes are linked to NBS-LRR-encoding genes.

Mapping of Resistance to Downy Mildew

We are continuing to map resistance phenotypes in order to develop molecular markers to assist the selection of resistance genes. In the past year, resistance to DM from a novel *L. saligna* donor was mapped to the major cluster of resistance genes on linkage group 1.

Identification of Markers Tightly Linked to Corky Root Resistance

We are continuing our efforts to identify markers tightly linked to the recessive resistance gene, *cor*, which confers resistance to corky root rot. There has been a paucity of markers closely linked to *cor* when analyzed in multiple crosses, possibly indicative of an elevated rate of recombination in this chromosomal region. Over the past year we have been fine mapping the chromosomal region containing *cor* by characterizing numerous recombinants in a Green Lakes x Diana population using flanking molecular markers. This data is consistent with *cor* being in a highly recombinogenic region and has provided multiple markers very tightly linked to *cor*. Several of these markers flanking the *cor* gene are located on the same ~500 kb genomic sequence scaffold assembled from cv. Salinas. We are now evaluating four candidate genes in this genomic scaffold. These markers will be used in our backcross programs to combine *cor* with downy mildew resistances from the novel sources.

Genetic Analysis of Resistance to *Fusarium* Wilt

We have continued our genetic analysis of resistance to *Fusarium oxysporum* race 1 in collaboration with Dr. Thomas Gordon (Dept. Plant Pathology, UC Davis). We previously

developed four populations involving resistant (R), partially resistant (PR) and susceptible (S) cultivars: Salinas (PR) x Green Towers (R), Lolla Rosa (R) x Salinas (PR), Lolla Rosa (R) x Green Towers (R) and Lolla Rosa (R) x Red Tide (S) (R = resistant, PR = partially resistant, S = susceptible). The summer field trial in 2010 revealed segregation in F₃ families from crosses between Lolla Rosa x Salinas and Salinas x Green Towers indicating one or two genes responsible for resistance. We re-tested these F₃ lines in our 2011 summer trial to further characterize this resistance and constructed genetic maps in these populations for QTL analysis of the resistance. We also tested F₂ progeny from Lolla Rosa (R) x Red Tide (S) in the 2011 summer field trial. Phenotypic segregation data indicated that resistance from Lolla Rosa is determined by two recessive genes. We genotyped the F₂ population for molecular markers and we will test F₃ families of this population for resistance to *Fusarium* in our 2012 summer trial.

Water Use Efficiency (WUE) and Nitrogen Use Efficiency (NUE)

A lettuce field trial was conducted in summer 2011 in Spence, Salinas to investigate NUE and WUE as part of the USDA SCRI-funded project (with matching support from the CLGRB) entitled Next-Generation Lettuce Breeding: Genes to Growers. The PIs particularly involved in this component were M. Cahn and R. Smith (UC Cooperative Extension, Salinas), R. Hayes and I. Simko (USDA ARS, Salinas) as well as personnel from this project. The experimental design was a split-split-plot with four replications consisting of two different treatments of nitrogen fertilization (0 and 230 lb N/acre applied over 30 lb N/acre pre-plant) and two different water regimes (50% and 130% of crop ET by drip irrigation) after overhead irrigation applied during seed germination (Figure 1).

Figure 1. 2011 summer field trial at Spence, Salinas to investigate WUE and NUE.



A total of 50 lettuce varieties comprising iceberg (11), romaine (26), Batavia (2), butterhead (2), latin (1) and leaf (4) lettuce types and 4 primitive or wild types were assayed.

Fresh and dry weights, total plant nitrogen, potassium and phosphorus, and plant morphology with regard to firmness, color, rate of development, leaf characteristics and head closure were measured. Preliminary data analysis showed significant diversity of NUE and WUE among the cultivars analyzed. The follow-up field trial this summer will have two components: one to confirm some of the results by re-analyzing ten cultivars exhibiting the most diverse responses to water and nitrogen treatments and the other to analyze a segregating RIL population to investigate the genetic basis of NUE.

Functional Analysis of Candidate Resistance Genes

Post-transcriptional gene silencing (PTGS), induced by RNA interference (RNAi), involves expression of double-stranded RNA fragments of target genes leading to the degradation of the mRNA transcribed from the endogenous (target) genes and consequently to the loss of the target gene's activity. RNAi targeted to the LRR-encoding region of *RGC2B* (which encodes the *Dm3* resistance specificity) had previously shown that *Dm14*, *Dm16*, *Dm18* and *Ra* are also members of the *RGC2* family (Wroblewski *et al.*, 2007. *Plant J.* **51**:803-18). We have been generating transgenic plants silenced for other resistance gene candidates (*RGCs*) that co-segregate with resistance phenotypes in order to test the resistance function of each *RGC*. These RNAi lines are being used as tester stocks in crosses to resistant lines to identify genes involved in resistance to several diseases.

Ultra-dense Genetic Map

We have continued to refine the ultra-dense map for lettuce. In collaboration with Dr. Allen van Deynze (UC Davis), with support from the UC BioStar program, Enza Zaden, Rijk Zwaan and Vilmorin, we previously developed a custom 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 has facilitated rapid SNP discovery, genotyping, and mapping. Genomic DNA of 213 RILs from the core Salinas x *L. serriola* acc. UC96US23 mapping population were hybridized in duplicate and an ultra-dense genetic map was developed based on ~14,000 transcribed sequences assigned to nine chromosomal linkage groups spanning a total of 1,561 cM. The mapped EST sequences are located by their genetic map position and displayed on chromosomal pseudomolecules using GBrowse (<http://gviewer.ucdavis.edu/cgi-bin/gbrowse/lettucePublic/>). The GBrowse display allows searching for EST nucleotide sequences and SNPs.

Mapping of Candidate Genes Relative to Horticultural Traits

We continue to genetically analyze and develop markers for horticultural traits. In past years, horticultural traits have been mapped relative to candidate genes in our core F_{7:8} RIL mapping population derived from a cross between *L. sativa* cv. Salinas and *L. serriola* acc. UC96US23. These studies have been expanded to other populations, often in collaboration with other groups (Table 1). QTL analysis identified chromosomal regions involved in horticultural traits such as heading, bolting, flowering time, (absence of) spines, leaf shape, branching, seed shattering, tipburn, high temperature seed germination, and several post-harvest disorders. We are now embarking on a meta-analysis to integrate QTL information from multiple years, locations, and populations.

We are working to identify the candidate genes responsible for some of the QTLs identified in this study. We are currently targeting traits that are determined by QTLs with large

Table 1. Populations being analyzed for disease resistance and horticultural traits.

Disease & Trait	Population	Collaborators
Downy mildew <i>Bremia lactucae</i>	Cobham Green x LS238 Cobham Green x LS241	B. Maisonneuve, INRA , France
	Iceberg x Saladin	P. Hand, HRI, UK
	Iceberg x Grand Rapids	I. Simko, USDA, Salinas
Corky root , <i>Sphingomonas suberifaciens</i>	Green Lakes x Diana	
Lettuce Mosaic Virus	Balady Aswan Green x Salinas PI226514 x Salinas 88 PI226514 x Clemente	R. Hayes, USDA, Salinas
<i>X. campestris pv. vitians</i>	Reine des Glaces x Delsay	B. Maisonneuve, INRA
	Little Gen x Salinas 88 Little Gem x Clemente	R. Hayes, USDA, Salinas
Big Vein	Thompson x Cisco	Y. Kawazu, NIV&TS, Japan
	Parade x Pavane	R. Hayes, USDA, Salinas
<i>Verticillium</i> wilt <i>Verticillium dahliae</i>	Pavane x Parade Salinas 88 x La Brillante Pacific x La Brillante	R. Hayes, USDA, Salinas
Lettuce drop <i>Sclerotinia minor</i> <i>Sclerotinia sclerotiorum</i>	PI261245 x Salinas Little Gen x Salinas 88 Little Gem x Clemente	R. Hayes, USDA, Salinas
<i>Fusarium</i> wilt <i>Fusarium oxysporum</i>	Valmaine x Salinas	T. Gordon, UC Davis
Basal rot <i>Botrytis cinerea</i>	Salad Bowl x CGN14263	
Lettuce dieback Lettuce necrotic stunt virus	Valmaine x Salinas	I. Simko, USDA, Salinas
Shelf life	Iceberg x Saladin	D. Pink, Harper Adams, UK
	Salinas x UC96US23	G. Taylor, U. Southampton
Rib discoloration & tipburn	Emperor x El Dorado	S. Jenni, AAFC, Canada
Salt Tolerance	Salinas x UC96US23	E. Schranz, U. Amsterdam
Nutritional content, Vit A, C & E, antioxidant & folate	Diplomat x Margarita Green Towers x Western Red Leaf	D. Still, CSU Pomona
Heat tolerance	Salinas x UC96US23	H. Jie, Nanyang Tech. U., Singapore

phenotypic effects: shattering (LG6), lobed leaf (LG3) and spines on the stem (LG5). Using the same strategy described above for *cor*, we have refined the position of each QTL and have reduced the region of interest using diversity panel association studies as well as genome assemblies of cv. Salinas and *L. serriola* acc. UC96US23. We have developed flanking markers in these regions and are refining the positions of these QTLs using recombinants derived from heterozygous RILs. We have initiated similar studies for tipburn and heading.

We welcome collaborations with researchers who wish to analyze existing populations, particularly the core F_{7:8} RIL mapping population derived from *L. sativa* cv. Salinas x *L. serriola* acc. UC96US23 as well as additional populations, for traits that have not been studied yet.

Transcriptome and Genome Sequencing

We have used ‘next-generation’ DNA sequencing for several projects. We have sequenced and assembled the transcriptome from both *L. sativa* cv. Salinas and *L. serriola* acc. UC96US23. We now have over 52,000 assemblies representing the majority of the transcribed genes of lettuce as well as SNPs differentiating the parents of the core mapping population. The complete assembly had a total length of 53 Mb and provided 51,842 unigenes of 300 nt or longer with an average contig length of 1,020 nt and a median length of 736 nt. Approximately 67% of all the unigenes displayed significant similarity to known protein sequences in the Plant RefSeq database and over 16,000 unigenes have complete uninterrupted open reading frames (ORFs). The raw reads and assemblies are now available from NCBI GenBank (for *L. sativa* cv. Salinas: SRA SRX034873, SRX033326, and TSASD JI573761-JI625602; for *L. serriola* acc. UC96US23: SRX098217 and JO020427-JO087153).

We have also sequenced and assembled the gene space (genomic sequence depleted for highly repeated sequences) of cv. Salinas and *L. serriola* acc. UC96US23. For cv. Salinas, 876,110 contigs were assembled; the longest contig was 28 Kb and the average and median contig lengths were 1226 and 955 nt respectively. Total assembly length was 1.13 Gb (43% of the estimated length of the ~2.7 Gb lettuce genome). The reads and sequence assembly are available at NCBI GenBank (SRX099365, SRX099215, SRX099223, SRX098725; AFSA00000000). Alignments between the gene-space and transcriptome assemblies clearly identified the positions of the introns that for the most part exactly matched the positions of introns in the corresponding genes in Arabidopsis. Genomic reads for *L. serriola* acc. UC96US23 are available under accession # SRX098375.

The genome of lettuce has been sequenced in collaboration with the BGI, Shenzhen, China, funded by an international consortium of companies [Agrisemen (NL), Enza Zaden (NL), Gautier Semences (FR), Isi Sementi (IL), Monsanto Vegetable Seeds (USA), Rijk Zwaan (NL), Syngenta (USA), Taki & Co (JP), Tozers (UK), Vilmorin (FR)] as well as the BGI. The genome has been assembled into 15,471 scaffolds comprising 2.5 Gb of the 2.7 Gb genome with a contig N50 of 11.7 kb and a scaffold N50 of 461 kb. Of the scaffolds that contained multiple unigenes that had been mapped using the lettuce chip (over 60% of the assembled genome), 95% are genetically consistent with our ultra-dense map; most of the remaining scaffolds analyzed are simple chimeras. We have placed these validated scaffolds into chromosomal linkage groups relative to numerous phenotypes. The genome has been annotated to provide ca. 44,000 gene models; we are currently refining the gene annotations manually. These data are available for public release as of June 2012.

We have also sequenced the gene-space of 96 RILs from the cv. Salinas x *L. serriola* acc. UC96US23 mapping population. These data are being used to further validate the genome assemblies and order them in chromosomal linkage groups as well as place additional scaffolds into the chromosomal linkage groups.

In collaboration with groups working on various aspects of lettuce biology, we have initiated RNA-Seq profiling to generate an atlas of genes expressed in lettuce at different developmental stages, under different abiotic stresses, and during resistance and susceptibility to diverse diseases.

In addition, we are continuing to sequence isolates of *Bremia lactucae* and are using the sequence data to provide detailed fingerprints of the major pathotypes as well as to search for genes that determine the virulence of isolates of downy mildew to different cultivars of lettuce. This information will help us understand the evolution of downy mildew and design strategies for more durable deployment of disease resistance genes.

Databases

We have continued to curate publicly accessible, inter-operable 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). Chiplett provides access to data from the Affymetrix lettuce Genechip project as well as the ultra-dense genetic map (<http://chiplett.ucdavis.edu/>). The G2G site (<http://scri.ucdavis.edu/>) provides access to information generated as part of the Next-Generation Lettuce Breeding: Genes to Growers (G2G) and CLGRP-funded projects. Our GBrowse genome viewer (<http://gviewer.gc.ucdavis.edu/cgi-bin/gbrowse/lettucePublic/>) provides access to the ultra-dense map as genetic chromosomal pseudomolecules. Lettcv (<http://compositdb.ucdavis.edu/database/lettcv2/display/>) archives extensive genetic, passport and performance data on lettuce cultivars. The CGP database (http://compgenomics.ucdavis.edu/compositae_index.php) contains extensive sequence and related information as well as links to lettuce genetic maps (lettuce genetic map viewer) and marker information. Morphodb (http://compgenomics.ucdavis.edu/morphodb_index.php) is an archive of and provides access to phenotypic information on *Lactuca* species; this database utilizes standard ontologies to facilitate searches across databases. These databases are being revised to facilitate access to marker information for breeding purposes from disease-centric, breeder-oriented perspectives.