CALIFORNIA LEAFY GREENS RESEARCH PROGRAM

April 1, 2014, to March 31, 2015

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. This project aims to ensure that lettuce benefits from the application of genomic and biotechnological techniques and has four components: (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 been focused on components (ii) through (iv). We continue to make extensive use of high-throughput sequencing and marker technologies. The genome of lettuce has been sequenced and assembled and ~41,000 lettuce genes annotated. Genotyping by sequencing of the core mapping population has assigned over 95% of the assembled genome to genetic bins ordered along chromosomal linkage groups. We have sequenced additional lettuce genotypes to assess allelic variation for horticulturally important genes. Sequences have been and are being mined for candidate genes for traits such as disease resistance, development, and horticulturally important traits. Candidate genes which co-segregated with disease resistance phenotypes have been tested for function using RNA interference (RNAi) and several causal genes have been identified. 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, 2014 to March 31, 2015

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**
Keri Cavanaugh
Miguel Macias González
Dean Lavelle
Oswaldo Ochoa
Marilena Christopoulou
Manju Govindarajulu
Lorena Parra
Germán Sandoya
Pauline Sanders
Alex Kozik
Huaqin Xu
Sebastian Reyes Chin Wo
Lien Bertier
UC Davis Genome Center and
The Department of Plant Sciences
University of California, Davis
Ryan Hayes
Ivan Simko
Steve Klosterman
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 sub-projects:

1) Analysis of transgenes in lettuce and genome editing.

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 objective includes a wide range of horticultural traits. Some of these studies have been funded from Federal grants and gifts from 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 and Genome Editing

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 a near-term need. 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 warrant further investigation. We do not currently have projects specifically focused on transgene expression in lettuce. However, we continue to generate transgenics as components of other projects (see below); these provide data on transgene expression and stability.

Recently, a new technology for precise genome editing based on the CRISPR/Cas9 system of prokaryotic immunity has been developed and adapted for use in eukaryotic organisms, including plants. This technology can be used to create gene knockouts, gene deletions and replacements and for introduction of new genes or sequences. We are developing this technology for lettuce. Gene knock-outs are currently much easier than allelic replacements. In the long term, we aim to generate stacks of genes, particularly for disease resistance, at single chromosomal positions so that they will be inherited as single Mendelian loci in breeding programs.

Resistance Candidate Gene Identification and Mapping

We continue to map loci for resistance to downy mildew (DM), corky root, *Fusarium* and *Verticillium* wilts, onto the consensus genetic map and the ultra-dense map as well as align them with the genome sequence. We are continuing to develop molecular markers to assist the selection of resistance genes. Of the over 50 phenotypic resistance genes mapped in lettuce, most co-localize to one five major resistance clusters on chromosomes 1, 2, 3, 4, and 8 (MRCs 1, 2. 3, 4, and 8 respectively). The majority of these resistance phenotypes are linked to NB-LRR-encoding (NLR) genes.

Resistance to Downy Mildew

We are using a whole genome sequencing (WGS) approach to identify chromosome segments introgressed from wild species during breeding for DM resistance. Numerous genes for resistance to DM have been introduced into cultivated lettuce from wild species (*L. serriola, L. saligna* and *L. virosa*) by repeated backcrossing (often six or more backcrosses) by public and commercial breeders. This has resulted in near-isogenic lines (NILs) that only differ for small chromosome regions that are potentially associated with resistance. Low-pass WGS of NILs provides the opportunity to efficiently identify these regions in order to map the resistance

gene(s) and provide markers for marker-assisted selection (MAS). Over the past 30 years we have generated advanced breeding lines that constitute NILs of new *Dm* genes from the wild species, *L. serriola* and *L. saligna*, in the *L. sativa* cv. Salinas background. Thirteen of these advanced breeding lines were selected for WGS. In addition, resistance was reassessed using isolates representative of *B. lactucae* in California, including recent highly virulent isolates. Only three advanced lines were resistant to all the assayed isolates (Table 1), although most of the lines were resistant to all but the highly virulent isolates. We will test more isolates to characterize the specificity of interactions between these NILs and isolates of *B. lactucae*.

Table 1. Reactions of advanced breeding lines against old and recent isolates of *B. lactucae*. Green box = resistant reaction. Red box = susceptible reaction.

	UC07105	UC07106	UC07107	UC07108	UC02202	UC02203	UC02204	UC02205	UC02206	UC12100	UC12101	UC12102	UC12103
DM gene	new	new	new	new	new	new	new	new	new	new	new	new	new
CAV(C04?1017)	-	-	-	-	-	-	-	-	-	-	-	-	-
CAVI(C05R1034)	-	-	-	-	-	-	-	-	-	-	-	-	-
CAVII(C98O648)	-	-	-	-	-	-	-	-	-	-	-	-	-
CAVIII(C01O879)	-	-	-	-	-	-	-	-	-	-	-	-	-
CAIX(C11O1327)	-	-	-	-	-	-	-	-	-	-	-	-	-
CAVIII-879	-	-	-	-	-	-	-	-	-	-	-	-	-
CAVII-1435	-	-	-	-	-	+	-	-	-	-	-	-	-
CAIIa-P24	-	-	-	-	-	+	-	-	-	-	-	-	-
CAIII-M47	-	-	-	-	-	+	-	-	-	-	-	-	-
CAV-1452	-	-	-	-	-	+	-	-	-	-	-	-	-
Novel-1557	-	-	-	-	-	+	-	-	-	-	-	-	-
Novel-1481	-	-	-	+	-	+	-	-	-	-	-	-	-
Novel-1324	-	-	-	-	-	+	-	-	-	-	-	-	-
Novel-1326	-	-	-	+	-	+	-	-	-	-	-	-	-
Novel-1485	-	-	-	-	-	-	-	-	-	-	-	-	-
Novel-1622*	+	+	-	+	+	+	+	+	+	+	-	+	-
Novel-1690*	+	+	-	+	+	+	+	+	+	+	-	+	-
*(highly virulent isolates)													

We have obtained low-coverage genomic sequence of these 13 advanced breeding lines by WGS. Mapping the sequencing reads from these advanced breeding lines against the *L. sativa* reference genome (cv. Salinas) allowed identification of introgressed regions as indicated by a high density of single nucleotide polymorphisms (SNPs; Figure 1). In some cases, NILs have only one highly polymorphic segment; in others there seem to be multiple polymorphic regions. We have sequenced 13 genotypes that were the wild donors of DM resistance for the development of these advanced breeding lines to assist in the identification of introgressed regions. In addition, we have sequenced 32 different lettuce cultivars, carrying known and new *Dm* genes (Table 2). These sequences will be used to characterize the *Dm* genes present in resistant cultivars and help distinguish between new and known *Dm* genes.

Figure 1. Breeding lines showing potential introgressions of regions encoding the Major Resistance Clusters (MRC) on chromosomes 1 and 2. A: Chromosome 1 for line UC02-105; B: Chromosome 2 for line UC12-101.

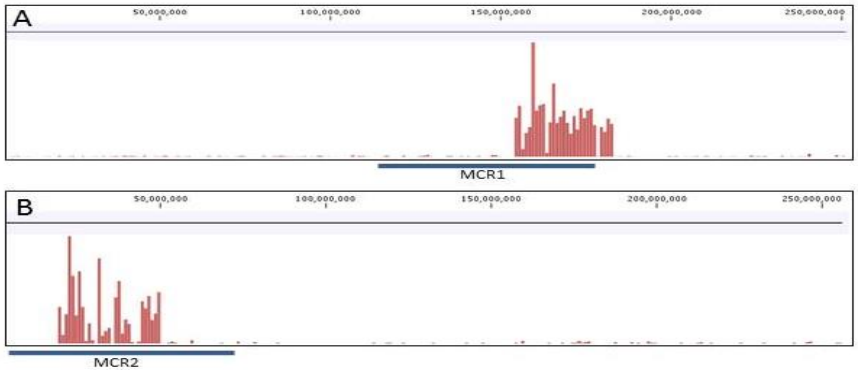


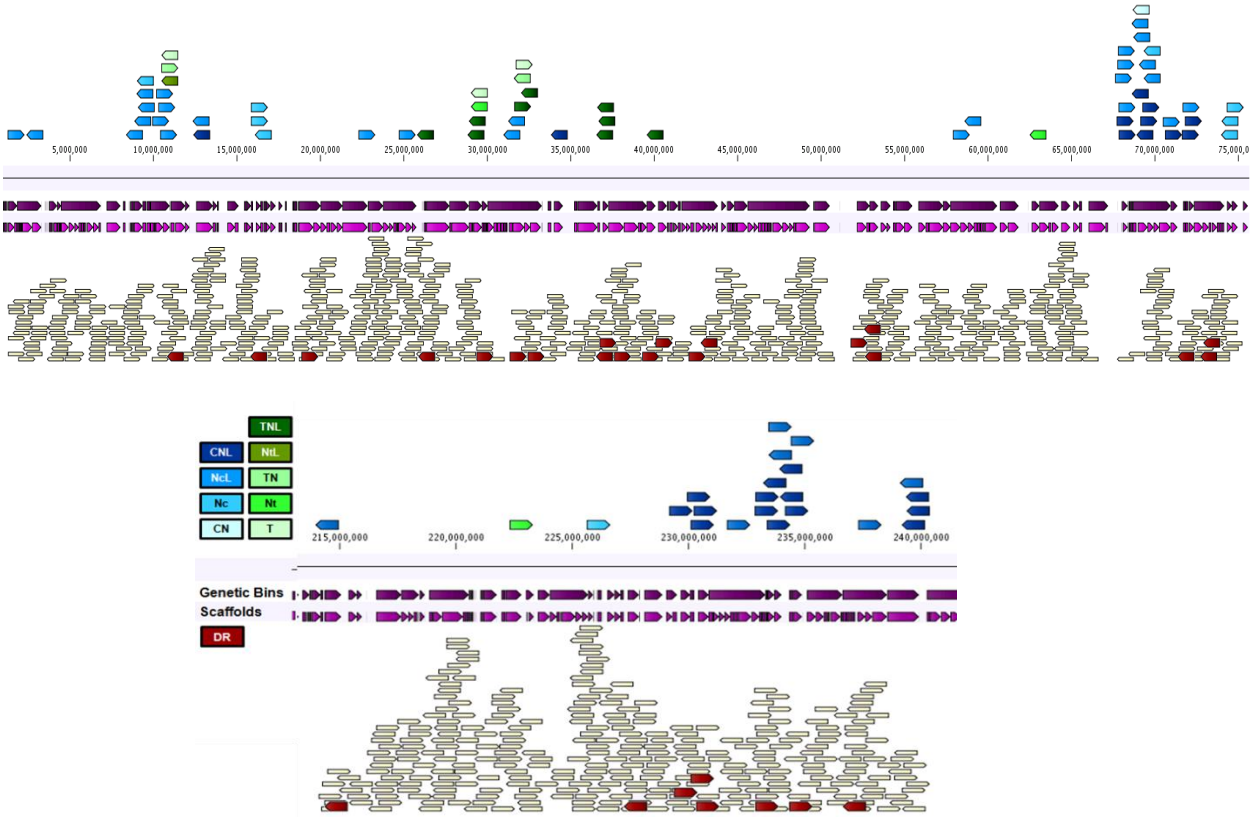
Table 2. Differential cultivars and lines carrying known or new *Dm* genes sequenced.

Cultivar/line	*Dm* gene	Cultivar/line	*Dm* gene
Cobham Green	*none*	CG Dm16	*Dm16*
Lednicky	*Dm1*	LSE/18	*Dm16*
UCDM2	*Dm2*	Saffier	*Dm1, Dm 5/8, Dm16*
Dandie	*Dm3*	Nun Dm17	*Dm17*
R4T57D	*Dm4*	El Dorado	*Dm18*
Valmaine	*Dm5/8*	Colorado	*Dm18*
Sabine	*Dm6*	R32	*Dm32, Dm18*
G. Blonde d'Hiver	*Dm7*	Ninja	*Dm3, Dm11, Dm36*
LSE 57/15	*Dm7*	Discovery	*Dm37*
UCDM10	*Dm10*	Argeles	*Dm38*
Capitan	*Dm11*	Amplus	*Dm40*
Hilde	*Dm12*	RYZ 2164	*?*
Pennlake	*Dm13*	RYZ910457	*?*
UCDM14	*Dm14*	Bedford	*?*
PIVT 1309	*Dm15*	Ballesta	*?*
Nun Dm15	*Dm15*	Bellisimo	*?*

.

We used the lettuce genome sequence to analyze the five MRCs in detail. MRC1 and MRC4 were described in the 2013-2014 CLGRB report and in Christopolou *et al.* (2015. *Molec. Pl.-Micobe Interact*. http://dx.doi.org/10.1094/MPMI-06-14-0175-R). MRC2 spans 73 Mb and contains 61 NLRs of six different gene families that co-segregate with nine disease resistance specificities. MRC3, which is 25 Mb, contains 22 *RGC21* genes and co-locates with *Dm13*. There are three smaller MRC on chromosome 8. A library of 33 transgenic RNAi tester stocks has been generated for functional analysis of NLR-encoding genes that co-segregated with disease resistance phenotypes in each of the MRCs. Members of four NLR-encoding families, *RGC1*, *RGC2*, *RGC21* and *RGC12* were shown to be required for 16 disease resistance phenotypes in lettuce.

Figure 2. Graphical overview of the major resistance clusters on chromosomes 2 (top) and 3 (bottom). T = TIR, C = non-TIR, N = nucleotide binding, L = leucine rich repeat domains present. DR = disease resistance related gene. Note the different scales. Genetic bins as defined by cross-overs and the sizes of genomic scaffolds across each region are shown.



Markers for Corky Root Resistance

We have identified markers tightly linked to the recessive resistance gene, *cor*, which confers resistance to corky root rot. There was a paucity of markers detected as closely linked to *cor* that was indicative of an elevated rate of recombination in this chromosomal region. We fine mapped the chromosomal region containing *cor* by characterizing numerous recombinants in a Green Lake (resistant to corky root) x Diana (susceptible to corky root) population that were selected using flanking codominant molecular markers. We localized the *cor* gene to a genomic region containing few scaffold sequences. These data provided multiple markers very tightly linked to *cor* that are now available for MAS by breeding companies so that *cor* can be efficiently combined with other genes for disease resistance. We have sequenced the genomes of Green Lake and Diana as well as eleven recombinant lines in the *cor* region to precisely map the *cor* gene. However, analysis of these recombinant lines has yet to resolve *cor* to a single candidate gene and additional neighboring scaffolds are currently being investigated. Once a candidate gene has been identified, we will validate it using genome editing to generate knock-outs.

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 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). Analysis in previous years of crosses between Lolla Rosa x Salinas and Salinas x Green Towers indicated one or two genes responsible for resistance. Also, QTL analysis of progeny from Lolla Rosa x Red Tide indicated the presence of one major QTL in LG1 and three minor ones, two on LG4 and one on LG8. We had previously identified a QTL for resistance to *Fusarium* also in LG1 from Valmaine. Current data indicate that the QTLs from Lolla Rosa and Valmaine are in different regions. However, precise integration of data across populations awaits more complete genetic maps for the individual populations that will be completed soon.

In 2014 we planted a subset of the best ~20 *Fusarium* resistant lines based on previous field data from each of Valmaine x Salinas, Salinas x Green Towers, Lolla Rosa x Salinas and Red Tide x Lolla Rosa populations in order to select resistant lines of different lettuce types. We selected 11 lines from Valmaine x Salinas, 13 lines from Salinas x Green Towers, 20 lines from Lolla Rosa x Salinas and 8 lines from Red Tide x Lolla Rosa populations.

Marker for Resistance to *Verticillium* Race 1

In collaboration with Dr. Ryan Hayes (USDA, Salinas), a major QTL for resistance against isolates of *V. dahliae* race 1 was previously mapped in a population from a cross between Salinas 88 (susceptible) x La Brillante (resistant) (CLGRP report 2008-2009; Hayes *et al*., 2011. *Theor. Appl. Genet.* **123**:509-17). This QTL co-segregates on LG9 with a scaffold that has three genes with sequence similarity to the *Ve1* gene for resistance to *V. dahliae* in tomato (McHale *et al*., 2008. *Theor. Appl. Genet*. **118**: 565-80). *Vrc1* in La Brillante has a 6 bp insertion that results in an *Eco*RI site that is correlated with resistance in other cultivars as described in last year's report. This provides an excellent molecular marker for resistance to *Verticillium* race 1. We are collaborating with the USDA group at Salinas to validate the resistance function of these genes using RNAi silencing. We will also confirm their function using knock outs generated by genome editing.

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

Starting in 2011, we have conducted a series of large field trials at Spence, USDA 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 are M. Cahn and R. Smith (UC Cooperative Extension, Salinas), R. Hayes and I. Simko (USDA ARS, Salinas) as well as personnel from this CLRRB project, specifically M. Macias González. See prior CLGRB reports for details of earlier experiments. In 2012, we analyzed a RIL population derived from Grand Rapids x Iceberg for nitrogen use efficiency (NUE) and other traits. A total of 94 RILs and the two parents were evaluated. We measured fresh and dry plant weight, percent solids, percent nitrogen, percent phosphorus, and percent potassium. NUE (dry plant weight / nitrogen content), phosphate use efficiency (dry plant weight / phosphorus content), and potassium (plant dry weight / potassium content) were calculated. QTLs for several traits were identified (See 2013-2014 report).

In 2014 we repeated the field trial of this RIL population for NUE to validate the prior results and to compare QTL stability over years and perform a QTL analysis on WUE on the same population. We conducted the following treatments: a high water high nitrogen treatment, a high water low nitrogen treatment, and a low water high nitrogen treatment. The same traits

were measured as in 2012. Unfortunately, due to residual nitrogen remaining in the soil after the dry winter, the low nitrogen treatment was not sufficient to stress the plants despite repeated leaching of the trial site prior to planting; therefore the low nitrogen treatment was dropped from the analysis. A total of 64 significant QTLs were identified using composite interval mapping. QTLs identified were associated with average plant weight, average dry plant weight, percent solids ([fresh plant weight/ dry plant weight] x 100), percent nitrogen content, percent phosphorus content, percent potassium content, NUE of dry plant weight (dry plant weight/ nitrogen content), NUE of fresh plant weight (fresh plant weight/ nitrogen content), relative efficiency index, and stress susceptibility index. Several of these QTLs were identified across both years; these included QTLs for plant fresh weight, dry plant weight, percent solids, and percent nitrogen.

Mapping of Candidate Genes Relative to Horticultural Traits

We continue to genetically analyze and develop markers for horticultural traits. Previous QTL analyses 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 in the process of a meta-analysis to integrate QTL information from multiple years, locations, and populations.

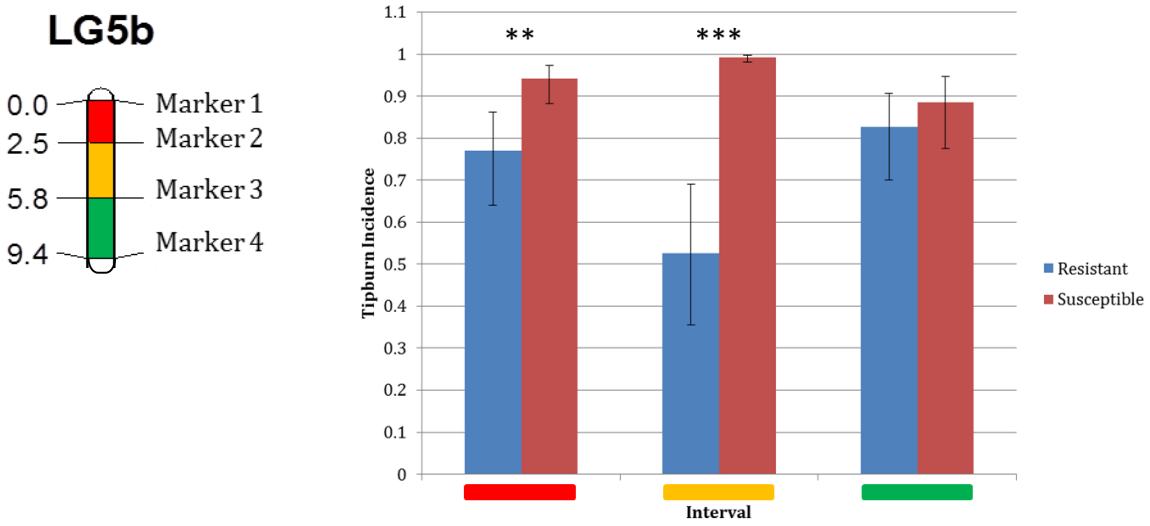
We are working to identify the candidate genes responsible for some of the QTLs identified in this study. We are currently targeting traits that are determined by QTLs with large phenotypic effects: shattering (LG6), lobed leaf (LG3) and spines on the stem (LG5). Using the same strategy described above 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 are conducting similar studies heading.

We welcome collaborations with researchers who wish to analyze existing populations, particularly the core 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.

Fine Mapping of Tipburn Resistance in Emperor x El Dorado Population

A QTL for resistance to tip burn has been previously identified in linkage group 5 (Fig. 3) on a cross between cvs. Emperor and El Dorado (Jenni *et al*., 2013. *Theor. Appl. Genet*. 126:3065-3079). We have fine mapped this QTL by phenotyping lines selected as being recombinant using codominant molecular markers flanking the QTL region. Earlier data had indicated that the QTL was between markers 1 and 4. Ninety six lines with a recombination event between markers 1 and 4 (Fig. 3) were evaluated for tip burn resistance in Yuma, AZ in collaboration with Dr. Ryan Hayes. Differences for tip burn incidence were evaluated among groups of recombinants carrying different alleles at the three intervals defined by markers 1, 2, 3, and 4 (Fig. 3). Recombinants carrying the resistant allele at the interval between markers 2 and 3 showed a highly significant difference on tip burn incidence with those recombinants carrying the susceptible allele. A smaller significant difference was observed for the interval between markers 1 and 2. No significant difference was observed for the third interval. More molecular markers are being analyzed between markers 1 and 3 to refine the position of tipburn resistance and provide markers for MAS.

Figure 3. Fine mapping of the QTL for tipburn resistance in LG5b. Tip burn incidence in groups of recombinant lines carrying the resistant or susceptible allele at the intervals defined by markers 1, 2, 3, and 4. Colors indicate the intervals evaluated. Error bars represent the 95% confidence interval of the mean, **: p <0.01; ***: p<0.001.



Genome and Transcriptome Sequencing

The genome of lettuce was sequenced in collaboration with the BGI, Shenzen, China, funded by an international consortium of ten companies as well as the BGI. We now have 2.3 Gb of genomic sequence assigned to genetic bins ordered along the nine lettuce chromosomes representing 96.7% of the total length of assembled scaffolds. We have placed genetically validated scaffolds relative to numerous phenotypes. The genome has been annotated to provide *ca*. 41,000 high-confidence gene models. The genome sequence is publically available at https://lgr.genomecenter.ucdavis.edu/.

We have also sequenced and *de novo* assembled the genomes of *L. serriola* (acc. US96UC23), PI251246, Green Lake, Diana, La Brillante, Iceberg, and Valmaine. In addition, we have low-pass sequenced 16 additional genotypes that represent mapping parents and diversity across lettuce types (Table 3). These are providing the sequences of alleles and haplotypes for several of the genes of interest described above.

We have recently embarked on the second phase of the Lettuce Genome Project that will refine the draft genome sequence using several new approaches in order to provide greater contiguity and resolution of each genetic bin of scaffolds as well as extensive allelic sequence variation. This is a three-year endeavor funded by the International Lettuce Genome Consortium that comprises of 17 large and small breeding companies.

In collaboration with groups working on different aspects of lettuce biology, we are conducting RNAseq profiling experiments to provide an atlas of genes expressed in lettuce at different developmental stages, under different abiotic stresses, and during resistance and susceptibility to diverse diseases.

Table 3. Lettuce genotypes sequenced.

Paired-end and mate-pair reads, *de novo* assemblies:	*L. sativa* cv. Salinas (Mapping parent. Reference genotype.)
	L. serriola US96UC23 (Mapping parent, reference population.)
	L. sativa PI251246 (Mapping parent. Oil type)
	Valmaine (Mapping parent. Romaine)
	Diana (Mapping parent. Butterhead)
	Greenlakes (Mapping parent. Crisphead)
	Iceberg (Mapping parent. Batavia)
	La Brillante (Mapping parent. Batavia)
Low pass:	*L. serriola* Armenian Acc. (Resistance to *Vert*. race 2)
	PI171674 (Cos type. Resistance to *Vert*. race 2)
	Calicel (Crisphead. Calmar derivative. Parent for tipburn RILs.)
	Grand Rapids (Mapping parent)
	Ride Tide (Red leaf recurrent parent.)
	Western Redleaf (Mapping parent.)
	Tropicana (Greenleaf recurrent parent.)
	Green Towers (Romaine recurrent parent.)
	Bib (Butterhead recurrent parent.)
	Diplomat (Early desert crisphead. Selection from Empire. Mapping parent.)
	Coolguard (Winter desert crisphead.)
	L. serriola Lse102 (Donor of *Dm17*.)
	Mariska (Butterhead, *Dm18*.)
	Margarita (Butterhead.)
	Little Gem (Baby leaf.)

Databases

We continue to curate several publicly accessible 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. The lettuce v3 genome assembly is publically available at https://lgr.genomecenter.ucdavis.edu/. These databases are being modified to facilitate access to marker information for breeding purposes from disease-centric, breeder-oriented perspectives. The Bremia Database displays virulence phenotypes, mating type and fungicide sensitivity for Californian isolates of *B. lactucae* characterized from 2008 to the present (http://bremia.ucdavis.edu/bremia_database.php).