CALIFORNIA LEAFY GREENS RESEARCH PROGRAM

April 1, 2015, to March 31, 2016

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 and genome editing in lettuce. (ii) Cloning and characterization of disease resistance genes. (iii) Genetic mapping with the goal of locating most of the disease resistance genes known in lettuce. (iv) Comparative genomics to identify candidate genes controlling horticultural traits. 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 genes annotated. Genotyping by sequencing of the core mapping population has assigned over 97% 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. We have demonstrated genome editing by knocking out a gene that inhibits germination at high temperatures. 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, 2015 to March 31, 2016

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 Macías González
Dean Lavelle
Marilena Christopoulou
Lorena Parra
Germán Sandoya
Pauline Sanders
Alex Kozik
Huaqin Xu
Sebastian Reyes Chin Wo
Lien Bertier
Adolfo Vargas
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 support 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 improvement of both crisphead and leafy types.

PROCEDURES AND RESULTS:

Transgene Expression and Genome Editing 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 deployment of transgenes in lettuce is not a near-term need. Introduction of genes into lettuce using *A. tumefaciens* is routine; 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 continue to generate transgenics as components of other projects (see below); these provide data on transgene expression and stability.

Since 2012, a new technology for precise genome editing based on the CRISPR/Cas9 system of bacterial immunity has been developed and adapted for use in plants and animals. This technology can be used to create gene knockouts, deletions and replacements and for introduction of new genes and sequences. Gene knock-outs are currently much easier than sequence replacements or additions but the enabling technology is advancing rapidly.

We are developing and applying genome editing technology for lettuce. To demonstrate the efficiency of generating monoallelic and biallelic gene knockouts in lettuce with CRISPR/Cas9, we targeted a gene involved in seed germination thermotolerance (*NCED4*); this gene had previously been cloned by Kent Bradford's group and provides a whole-plant selectable phenotype. A total of 65 transgenics were generated of two lettuce genotypes (cvs. Cobham Green and Salinas) transformed with three different *CAS9* constructs each containing one guide RNA targeting *NCED4*. The *NCED4* gene was sequenced in the T_1 primary transformants and in several hundred T_2 plants that showed editing of the target gene in the T_1. Of the 49 T_1 plants containing the CAS9 gene, 26 plants (53%) showed editing of at least one *NCED4* allele (several showed editing events in both alleles). Variable numbers of T_2 seed from approximately half of thirty-six families tested so far were able to germinate at a temperature that was non-permissive for the wildtype (Fig. 1).

In the longer term, we will develop genome editing of lettuce to create stacks of resistance genes containing several R genes at single chromosomal positions so that they will be inherited as single Mendelian loci in breeding programs (e.g. Fig. 2). Such stacks could contain multiple resistance genes effective against all known pathotypes of downy mildew and multiple viral, bacterial, fungal pathogens as well as insect pests and nematodes. This will enhance the durability of resistance by increasing the evolutionary hurdle that pathogens will have to overcome. Gene stacks could expanded as more resistance genes become available and genes replaced when overcome by changes in the pathogens. A herbicide resistant gene (e.g. *ALS*) could be used as selectable marker for the gene stack. Gene stacking will greatly simplify breeding for disease resistance so that breeders can focus on more complex traits such as water and nitrogen use efficiencies, and nutritional quality. However, in order for this to happen genes

for resistance to each disease need to be identified at the molecular level as well as intellectual property and regulatory issues clarified.

Figure 1. Editing of *NCED4* results in germination of lettuce cv. Salinas at high temperature. Germination at 35°C of T_2 seed from five independent lines plus an unedited control.

Figure 2. Example of a conceptual stack of resistance genes conferring resistance to multiple pathogens plus a herbicide resistance gene (*ALS*) as a selectable marker.

Resistance Gene Identification

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 of 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. Comprehensive descriptions of all resistance genes in lettuce have been published (Christopoulou *et al*., 2015. Dissection of two complex clusters of resistance genes in lettuce (*Lactuca sativa*). *Molec. Pl.-Micobe Interact*. **28**: 751-765; Christopoulou *et al.*, 2015. Genome-wide architecture of disease resistance genes in lettuce. *G3, Genes, Genomes, Genetics* **5**:2655-269.). These papers provide extensive enabling data for marker development.

Resistance to Downy Mildew

We have reviewed the available information on sources of resistance to downy mildew and rationalized it into 51 resistance genes and factors plus 10 new uncharacterized sources and 15 quantitative trait loci along with supporting documentation as well as genetic and molecular information. This involved multiple rounds of consultation with many of the original authors. This has been published and provides a foundation for future nomination of new *Dm* genes (Parra *et al*., 2016. Rationalization of genes for resistance to *Bremia lactucae* in lettuce. *Euphytica* http://link.springer.com/article/10.1007/s10681-016-1687-1).

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. Twelve of these advanced breeding lines show resistance to many isolates representative of *B. lactucae* in California, including recent highly virulent isolates; one of our advanced lines was resistant to all of the isolates assayed (Table 1).

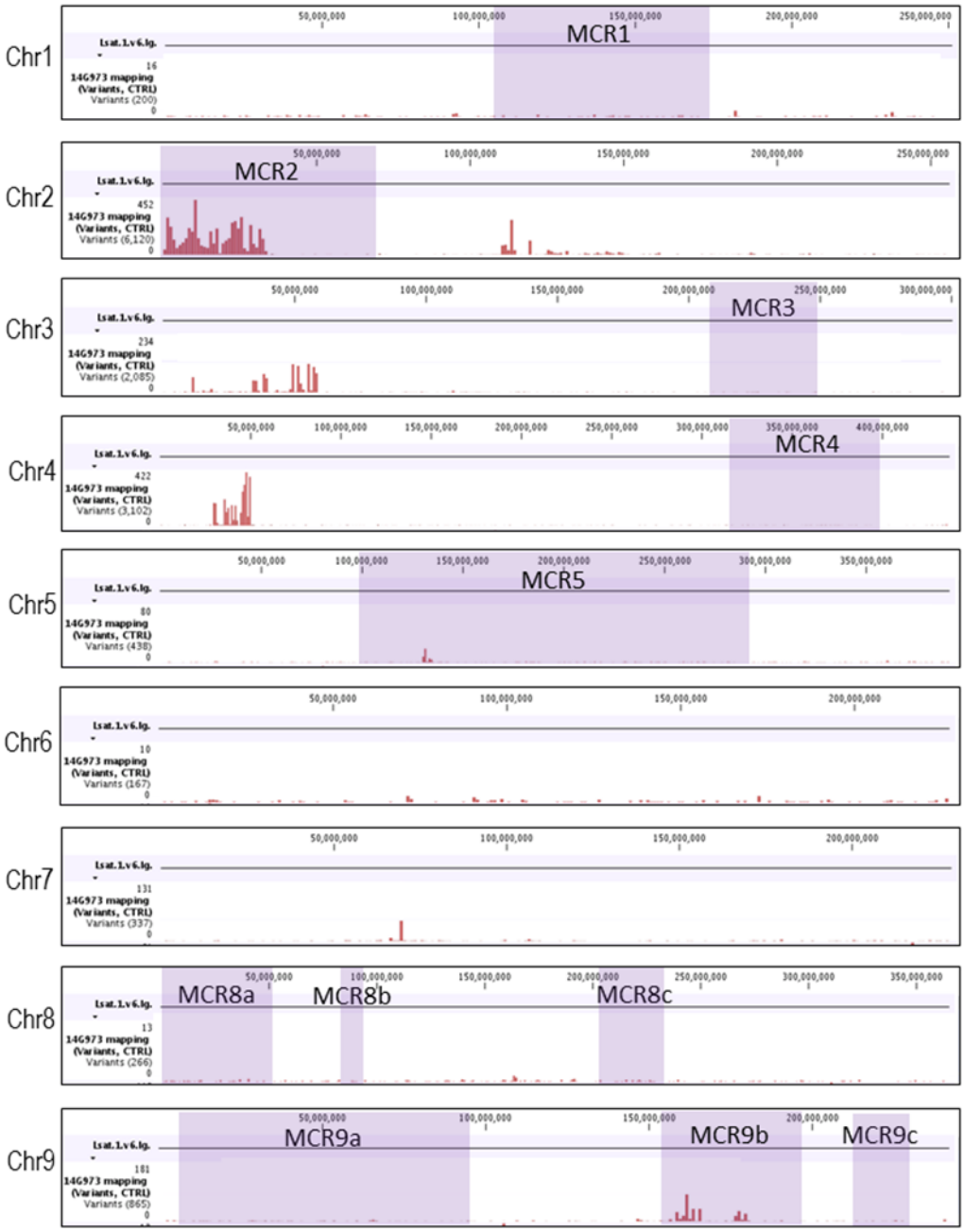
Table 1. Reactions of advanced breeding lines to old and recent isolates of *B. lactucae* from California. Green = resistant reaction. Red = susceptible reaction.

	UC07106	UC07107	UC07108	UC02202	UC02204	UC02205	UC02206	UC12100	UC12101	UC07105	UC12102	UC12103
C01O879	-	-	-	-	-	-	-	-	-	-	-	-
C82P24	-	-	-	-	-	-	-	-	-	-	-	-
C83M47	-	-	-	-	-	-	-	-	-	-	-	-
C11O1324	-	-	-	-	-	-	-	-	-	-	-	-
C11O1326	-	-	+	-	-	-	-	-	-	-	-	-
C13C1407	+	+	-	-	-	-	-	-	+	-	-	-
C14C1452	-	-	-	-	-	-	-	-	+	-	-	-
C14C1435	+	-	-	-	-	-	-	-	-	-	-	-
C14C1481	-	-	+	-	-	-	-	-	-	-	-	-
C14C1549	-	-	-	-	-	-	-	-	-	-	-	-
C14C1557	-	-	-	-	-	-	-	-	-	-	-	-
C15C1622	-	-	+	+	+	+	+	+	-	+	+	-
C15L1691	+	+	+	+	+	+	+	+	-	+	+	-
C15L1742	+	+	+	-	-	-	-	-	+	-	-	-
C15L1743	-	-	-	-	-	-	-	-	-	-	-	-
C15L1744	+	-	+	-	-	-	-	-	-	-	-	-
C15L1745	-	-	-	-	-	-	-	-	-	-	-	-
C15L1746	-	-	-	-	-	-	-	-	-	-	-	-
C15L1747	-	+	-	-	-	-	-	-	-	-	-	-
C15L1752	-	+	+	-	-	-	-	-	-	+	+	-
C15L1773	-	+	+	-	-	-	-	-	+	-	-	-
C15L1774	+	+	+	-	-	-	-	-	+	-	-	-
C15L1776	+	+	-	+	-	-	-	-	+	-	-	-
C15L1777	+	+	-	-	-	-	-	-	+	-	-	-
C15L1764	+	+	+	+	+	+	+	+	-	+	+	-
C15L1769	-	-	+	-	-	-	-	-	-	+	+	-

We have obtained low-coverage genomic sequence of these 12 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). In some cases, NILs have only one highly polymorphic segment; in others there seem to be multiple polymorphic regions. Currently,

we are generating F_2 populations from crosses between these NILs and cv. Salinas in order to determine which of these polymorphic regions are required for resistance.

Figure 3. Polymorphic regions resulting from introgression of resistance into an advanced breeding line from *L. serriola*. One region coincides with Major Resistance Cluster 2 (MRC2). Two additional regions on chromosomes 3 and 4 do not coincide with any canonical resistance gene sequences and may or may not be involved in resistance.



In order to identify the resistance genes present in the advanced breeding lines and distinguish these genes from resistance genes previously described in other cultivars, we have also sequenced 32 different lettuce cultivars, carrying known and new *Dm* genes. Additionally,

we will use long fragment sequencing (PacBio) for the identification of the repertoire of resistance genes present in each breeding line and the different cultivars carrying known *Dm* genes. These sequences will be used to characterize the *Dm* genes present in resistant cultivars and help distinguish between new and known *Dm* genes.

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). 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 indicated that the resistance QTLs from Lolla Rosa and Valmaine are in different regions. However, precise integration of data across populations awaited more complete genetic maps for each population. We have now genotyped three of the populations (Lolla Rosa X Salinas, Lolla Rosa x Red Tide and Salinas x Green Towers) using Genotyping by Sequencing (GBS) and have developed detailed genetic maps for each. New QTL analyses for resistance to *Fusarium* are in progress.

In 2014 and 2015 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. Selected lines are being trialed for uniformity in a non-infected field in Davis, CA (Fig. 4). Uniform lines will be planted in the *Fusarium* sick plot in Davis this summer to confirm resistance before releasing the best lines to seed companies.

Figure 4: Field trial in Davis, 2016 to test for uniformity of lines of several lettuce types previously selected for resistance to *Fusarium* in trials in infested fields.

Genetic Analysis

We have been constructing detailed genetic maps using Genotyping by Sequencing (GBS) on populations segregating for a variety of traits (Table 2). GBS uses methylation-sensitive enzymes to avoid repetitive regions of the genome and reduce the amount of sequencing required. We implemented a bioinformatics pipeline using Tassel and custom scripts for fast and accurate SNP haplotyping and the development of maps for each population.

Population		Trait	Collaborators
PI251246 x Arm. *L. serriola*	RIL	Verticillium race 2 resistance	UCD
Lolla Rosa x Salinas	$F_{2:3}$	Fusarium resistance	UCD
Iceberg x PI491224	$F_{2:3}$	LDM field resistance	UCD/USDA
Grand Rapids x Salinas	$F_{2:3}$	LDM field resistance	UCD/USDA
Salinas x Calicel	RIL	Tipburn	UCD
PI171674 x PI204707	$F_{2:3}$	Verticillium race 2 resistance	USDA
Reine des Glaces x Delsay	RIL	Xanthomonas resistance	INRA, France
Salinas x Green Towers	$F_{2:3}$	Fusarium resistance	UCD
Red Tide x Lolla Rosa	$F_{2:3}$	Fusarium resistance	UCD
Vanda x Gisele	$F_{2:3}$	Fusarium resistance	Brazil
Grand Rapids x Iceberg	$F_{2:3}$	LDM field resistance/NUE,WUE	UCD/USDA
PI171674 x Salinas	$F_{2:3}$	Verticillium race 2 Resistance	USDA
Reine des Glaces x Eruption	$F_{2:3}$	Xanthomonas Resistance	USDA
RH08-0111 x PI491108-1	$F_{2:3}$		USDA
Saladin x Iceberg	RIL		U. of Warwick, UK
Salinas x L. serriola	RIL	Core population	UCD

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

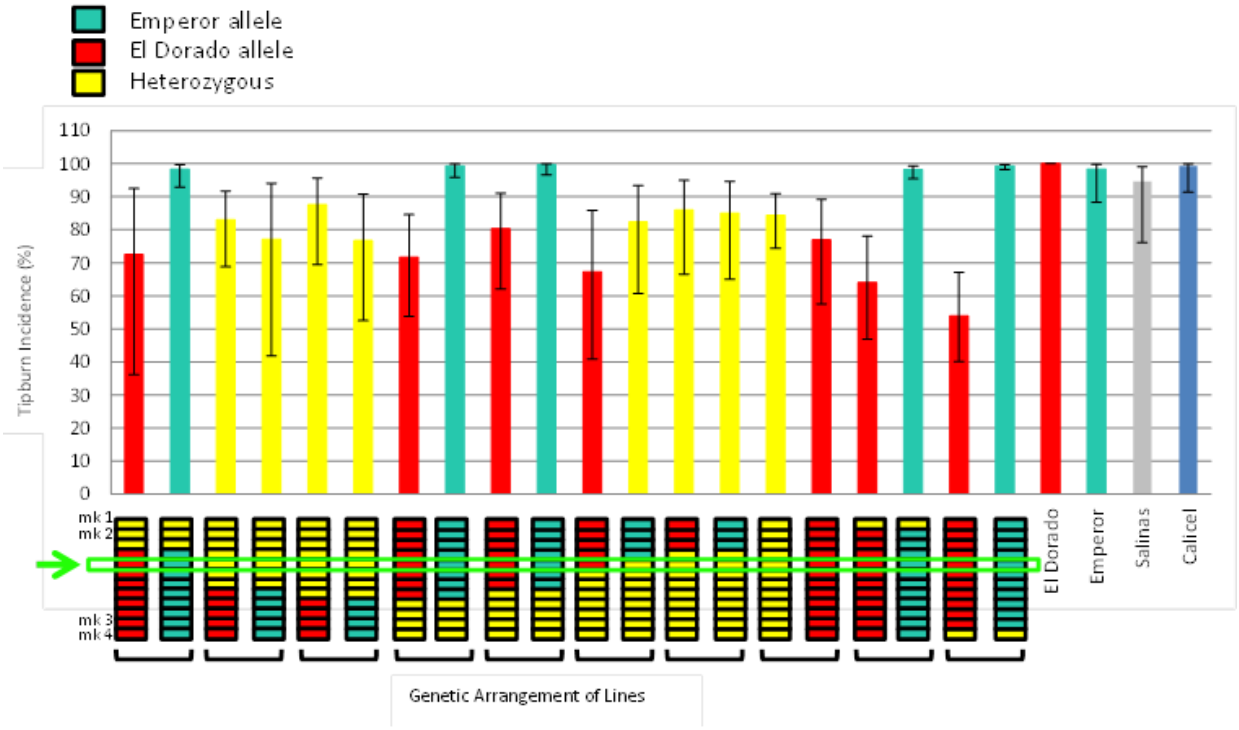
Starting in 2011, we have conducted a series of large field trials at Spence, USDA Salinas, to investigate NUE and WUE as part of USDA SCRI-funded projects (with matching support from the CLGRB) as described in prior CLGRB reports. QTLs for several traits were identified (See 2013-2014 report). A new trial focused on NUE will be conducted this summer at Spence, USDA Salinas to test if the QTL identified previously are reproducible across years.

Fine Mapping of Tipburn Resistance in a Emperor x El Dorado Population

A QTL for resistance to tip burn had been previously identified in linkage group 5 in a cross between cvs. Emperor and El Dorado (Jenni *et al*., 2013. *Theor. Appl. Genet*. 126:3065-3079). We 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 (Fig. 5). Ninety six lines with a recombination event between markers 1 and 4 were evaluated for tip burn resistance in Yuma, AZ in collaboration with Dr. Ryan Hayes. 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. Recombinant lines were phenotyped in 2015-2016 for tip burn incidence in Yuma, AZ and genotyped for additional markers between markers 2 and 3 (Fig. 5) to refine the position of tipburn resistance and provide markers for MAS. Significant differences were observed among these lines depending on the parental allele they were carrying at the region highlighted in green in Figure 5. The lines carrying the El Dorado allele showed significantly

less tip burn incidence than heterozygous lines or those carrying the Emperor allele at this position.

Figure 5: Tipburn incidence of lines with recombination events between markers 2 and 3. Green box indicates the region most associated with tip burn incidence.



QTL Mapping of Tipburn Resistance in Salinas x Calicel RIL Population

We developed a RIL population of 382 families from a cross between the crisphead cultivars Salinas (tipburn resistant) and Calicel (tipburn susceptible). This population was evaluated in Salinas, CA in 2015 for tipburn resistance and genotyped using GBS. Five phenotypes were measured: plant fresh weight, tipburn incidence and severity, leaf serration and head firmness (Fig. 6).

QTLs were identified for all of these traits (Fig. 7). Four QTLs were associated with tipburn incidence and severity in linkage groups 2, 5, 7 and 8. The QTL in LG5 is in the same region as the QTL identified in the Emperor x El Dorado population (see above) that is being fine mapped.

The same RILs were trialed in Yuma, AZ in 2015-2016. These data are currently being analyzed.

Figure 6. Phenotypic distribution of five traits in the Salinas x Calicel RIL population.

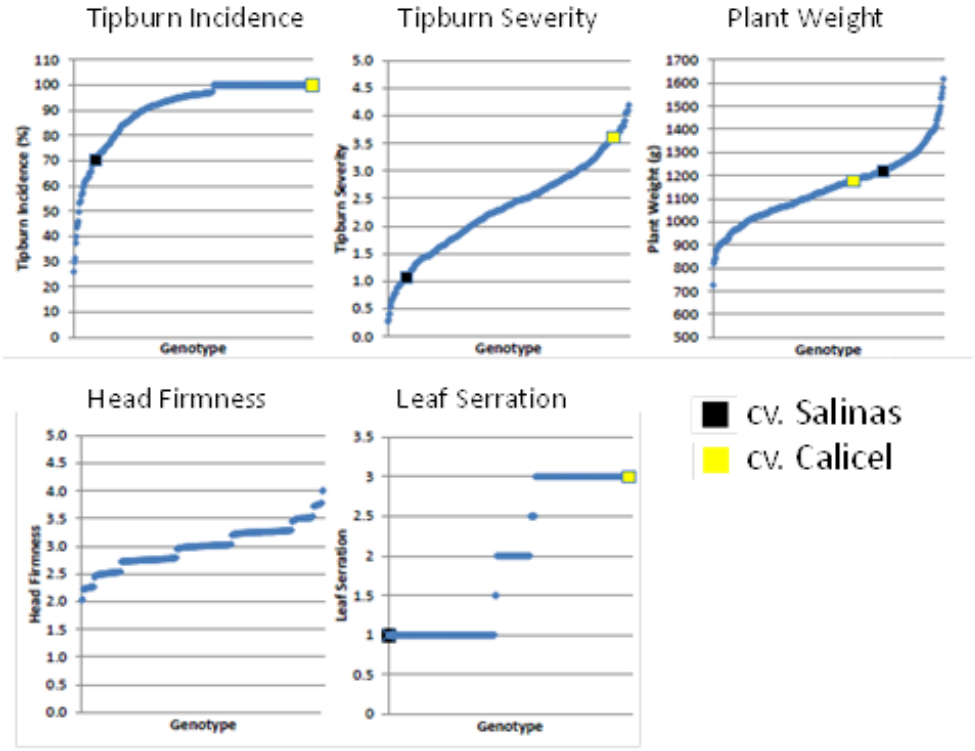
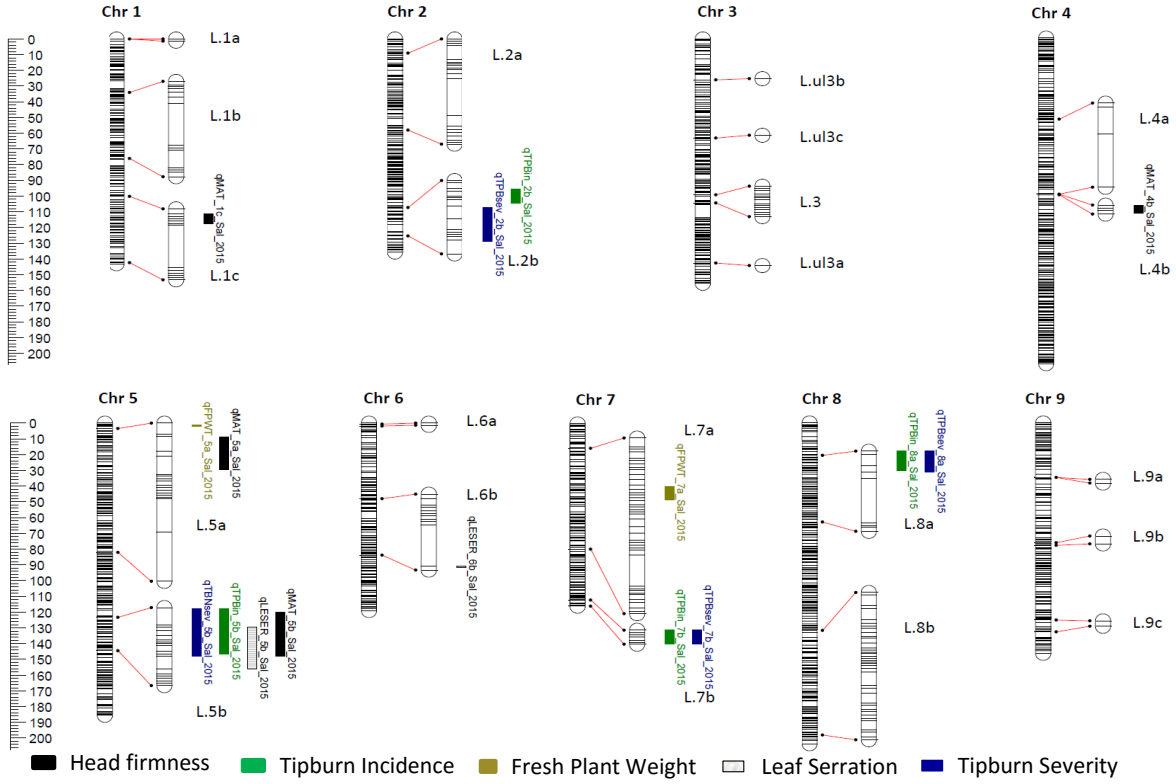


Figure 7. QTLs identified in the Salinas x Calicel population in 2015 in Salinas.



Genome and Transcriptome Sequencing

The genome of lettuce was previously 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 3.0 version of the genome sequence is publically available at https://lgr.genomecenter.ucdavis.edu/.

We have 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. These data are being made available through our website.

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 2001 to the present (http://bremia.ucdavis.edu/bremia_database.php).