

CALIFORNIA LEAFY GREENS RESEARCH PROGRAM

April 1, 2013, to March 31, 2014

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. Tightly linked markers are now available for marker assisted selection of resistance to corky root and *Verticillium* race 1. 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, 2013 to March 31, 2014

PROJECT TITLE: GENETIC VARIATION IN LETTUCE

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

COOPERATING PERSONNEL: **María José Truco**
Keri Cavanaugh
Miguel Macias Gonzalez
Dean Lavelle
Oswaldo Ochoa
Marilena Christopoulou
Manju Govindarajulu
Lorena Parra
German Sandoya
Carlos Galeano
Pauline Sanders
Lutz Froenicke
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
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.
- 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 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 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.

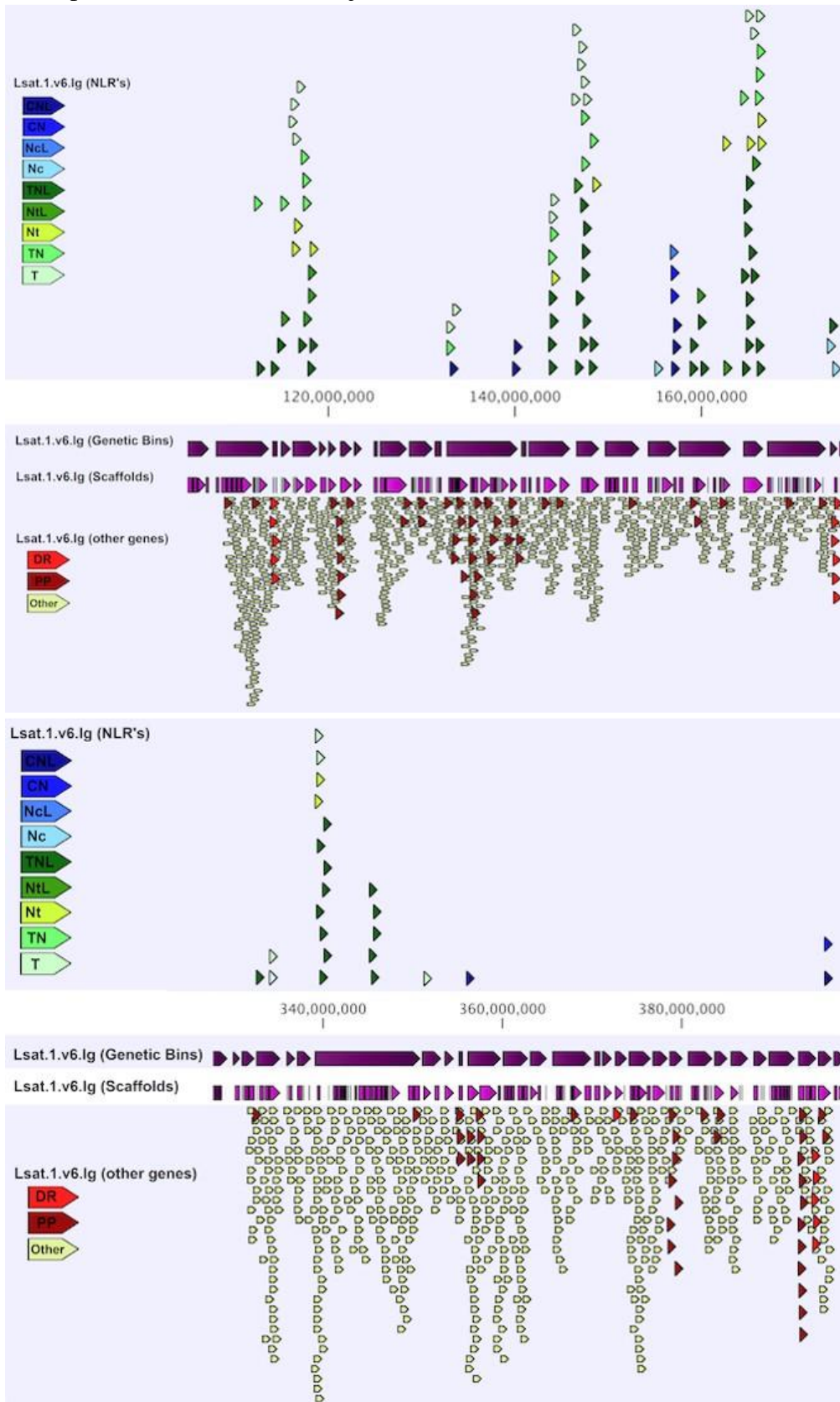
Resistance Candidate Gene Identification and Mapping

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, 25 co-localize to three major resistance clusters on chromosomes 1, 2, and 4 (MRC1, 2 and 4, respectively). The majority of these resistance phenotypes are linked to NBS-LRR-encoding (NLR) genes.

Resistance to Downy Mildew

We used the lettuce genome sequence to analyze MRC1 and MRC4 in detail. These two MRCs span over 65 and 64 Mb containing 104 and 24 NLR encoding genes, respectively, as well as 850 and 681 genes which are not related to NLRs (above and below the genetic bins and scaffolds in Fig. 1 respectively). We applied forward and reverse genetic approaches to dissect MRC1 and MRC4. Silencing of two of the five NLR-encoding gene families tested resulted in abrogation of 9 of 14 tested resistance phenotypes mapping to these two regions. At MRC1, members of the *RGCI* gene family were implicated in host and non-host resistance through silencing *Dm5/8* and *Dm45* mediated resistance to *Bremia lactucae* as well as the hypersensitive response to secreted proteins of the bacterium, *Pseudomonas syringae*, that is a non-pathogen of lettuce. At MRC4, silencing indicated that *RGCI2* family members conferred *Dm4*, *Dm7*, *Dm11*, and *Dm44* mediated resistance to *B. lactucae*. Changes in the DNA were identified in *dm7* loss-of-resistance mutant lines in *RGCI2G* within MRC4, indicating that it confers *Dm7*.

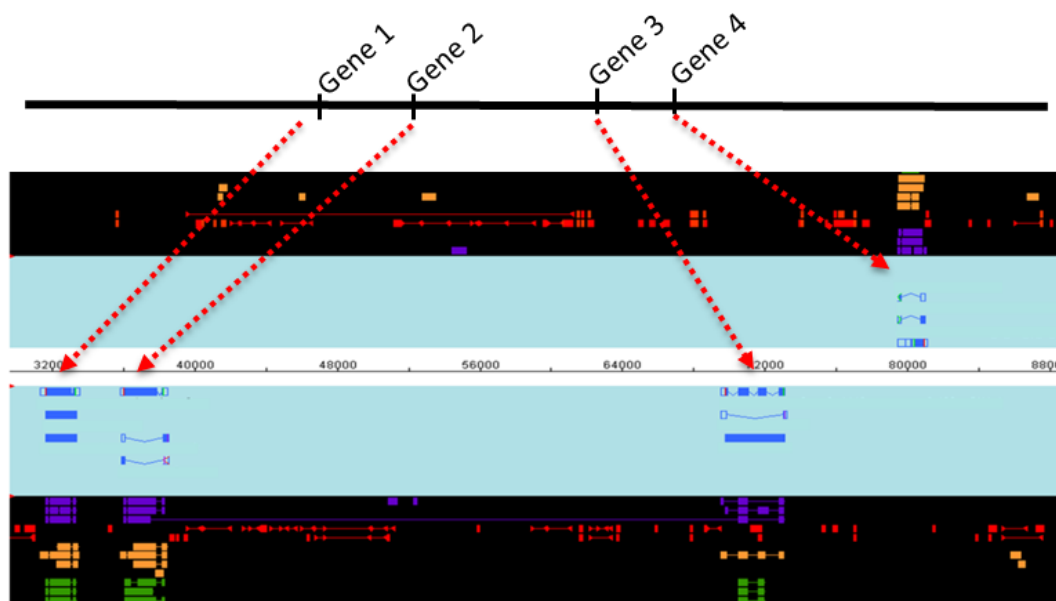
Figure 1. Graphical overview of the major resistance clusters on chromosomes 1 and 4.



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 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 x Diana population using flanking molecular markers. These data were consistent with *cor* being in a highly recombinogenic region and has provided multiple markers very tightly linked to *cor*. We have localized the *cor* gene to a genomic region containing four tightly linked genes on a single genomic scaffold sequence. These markers are now available for marker assisted selection by breeding companies so that *cor* can be efficiently combined with other genes for disease resistance.

Figure 2: The *cor* genomic region showing four candidate genes.



We generated stable transgenic plants carrying RNAi constructs with each of the four candidate genes and obtained multiple transgenic lines for each construct that showed good silencing of the GUS reporter gene. Testing of T₃ progeny of these silenced plants for resistance to corky root continues. We also sequenced the genomes of Green Lake (resistant to corky root) and Diana (susceptible to corky root) to determine the haplotypes of each in the *cor* region. This provided additional sequence information but no definitive indication as to which gene determines corky root resistance.

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. In 2013, we retested F₃ families from Lolla Rosa x Red Tide for resistance to *Fusarium* in a replicated field trial at UC Davis. Current results indicate that the QTLs from Lolla Rosa and Valmaine are in different regions. However, precise integration of data across populations awaits more complete maps for the individual 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-segregated on LG9 with an EST (QGJ16G22), which has sequence similarity to the *Ve1* gene for resistance to *V. dahliae* in tomato (McHale *et al.*, 2008. *Theor. Appl. Genet.* **118**: 565-80). This lettuce EST is located on a genomic scaffold that has three genes similar to the *Ve* genes in tomato (Fig. 3). The three lettuce homologs are similar in size to the tomato genes (*Vrc1*: 3.4 , *Vrc2*: 3.2 and *Vrc3*: 3.2 Kb). *Vrc1* is highly expressed in cv. Salinas. We have sequenced the genome of cv. La Brillante and determined that the resistant haplotype in the *Vrc* region is similar to that of the reference genome of the susceptible cv. Salinas. *Vrc3* has a frameshift mutation in La Brillante indicating that it is not responsible for resistance to *Verticillium* race 1. *Vrc1* in La Brillante has a 6 bp insertion that results in an *EcoRI* site (Fig. 4); this correlated with resistance in other cultivars. 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.

Figure 3. The positions and relative orientations of three genes with sequence similarity to the *Ve* gene of tomato on a genomic scaffold of *L. sativa* cv. Salinas.

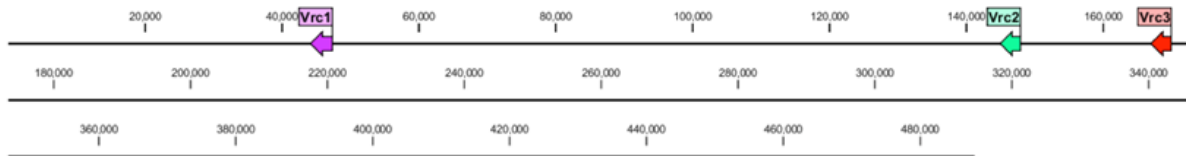


Figure 4. Alignment of PCR products for *Vcr1* from cvs. La Brillante and Salinas. The La Brillante allele contains an indel (boxed in red) that generates a unique *EcoRI* site (GAATTC).

La_Brillante	1	CAGATTGCTGCACCTGGATAGGTGTTAATTGCAGCATCGGGGGTCAGGTTATCGGCC TAGATTTAAACAACGA	73
Salinas	1	CAGATTGCTGCACCTGGATAGGTGTTAATTGCAGCATTGGGGGTCAGGTTATCGGCTAGATTTAAGCAACGA	73
La_Brillante	74	CGCTATATCTGGTGGTATTGATGGTTCCTACTTCTCTTTCCGTTTAGAGAATCTTCAGATGCTGAATCTGGCT	146
Salinas	74	AACGATATCTGGTGGTATTGATGATTCAGTTCTCTCTCCGTTTAGAGAGTCTTGAGACCGCTGAATCTGGCT	146
La_Brillante	147	GGAAATAACTTCAATTTCCACACAGATTCCTTCGATATTTGGCAGTCTGACTAGTTTGAGGAGTTTGAACCTTGT	219
Salinas	147	GGAAATAACTTCAATTTCCACGCCGATTCCTTCAGGATTTGGCAGTTTGACTAGTTTAAAGGAATTTGAACCTTGT	219
La_Brillante	220	CAAATTCGTTGTTTTCCAGGGCAGATTCAGGAGAATTGTCACGACTGACAAAAGCTTGAAGTCTTGTATTTATC	292
Salinas	220	CAAATTCGTTGTTTTCTGGCCAGATTCCTGGAGAATTGTCGCATCTGACAAAAGCTTCAAGTCTTGTATCTGTC	292
La_Brillante	293	TTTCGTTTTTCCCAATGGGAATTCGCTCACTGAAACTTGAGAAAACCAATCTAGCCATGCTTCTTAGGAACCTC	365
Salinas	293	TTCTCTTTTCTCCATGGGAATTCGCTCACTGAAACTTGAGAGCCCCAATCTAGCCATGCTTATCAAAAACCTC	359
La_Brillante	366	ACACAACCTTAGAGGTCTTTATCTGGATAGTGTGAACATATCAGCACAAAACCTCTGTTTGGTGCCAGGTTTTAT	438
Salinas	360	ACACAGCTTAAAGTCTTTCATCTGGATAGTGTGAACATATCGGCACAAAACCTCAGATTGGTGCCAGGCTTTAT	432
La_Brillante	439	CCTCATCTTTGCCACACTTAGAAGCTTTGAGCTTGTCAAATTTGCAACTTTTCAGGCCCTTTAG	501
Salinas	433	CCTCTCTTTGCTTGATTTGAGGTTTTGAGTTTGTCAAATTTGCAACTTTTCAGGCCCTTTAG	495

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

A lettuce field trial was conducted in summer 2011 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 (see previous CLGRB reports). 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. In 2012, we repeated the 2011 experiment with a reduced number of genotypes as well as analyzed a RIL population derived from Grand Rapids x Iceberg for nitrogen use efficiency (NUE) and other traits. The experiment consisted of a split-plot RCBD with three replications where nitrogen treatments were randomized within blocks and genotypes were randomized within nitrogen treatments. The nitrogen treatments consisted of a low nitrogen application (36 lbs / acre applied as preplant) and a high nitrogen treatment (236 lbs/acre split into four applications). 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.

In 2013 using the field data from experiments described above, we conducted a Principal Component Analysis (PCA) for fresh weight at low and high nitrogen treatments. The first principal component (PC1) represents the largest amount of variation, whereas the second principal component (PC2) represents the remaining variation (Fig. 5). RIL families above the regression line have higher fresh plant weight than expected at low nitrogen. PC1 describes the Relative Efficiency Index (REI; Rosales-Serna *et al.* 2000. *Agrociencia* **34**:153-165): RIL families with higher plant weight at high N tend to also have higher fresh plant weight at low N describing the efficiency of N usage to produce biomass. PC2 describes the Susceptibility Index (SI): How tolerant a RIL family is at maintaining its fresh plant weight under low N. REI measures the efficiency of a genotype to produce yield. SI measures the susceptibility of a genotype to stress. Well performing genotypes under stressful conditions are those with high REI and low SI values. QTL analysis was performed with QTL Cartographer using 1000 permutations to estimate the significance threshold. QTLs for several traits were identified (Fig. 6). This RIL experiment is being repeated this summer to validate the results and to compare QTL stability over years.

Figure 5. Linear regression of fresh plant weight at low and high N. Each data point represents a RIL family from the cross of Grand Rapids x Iceberg.

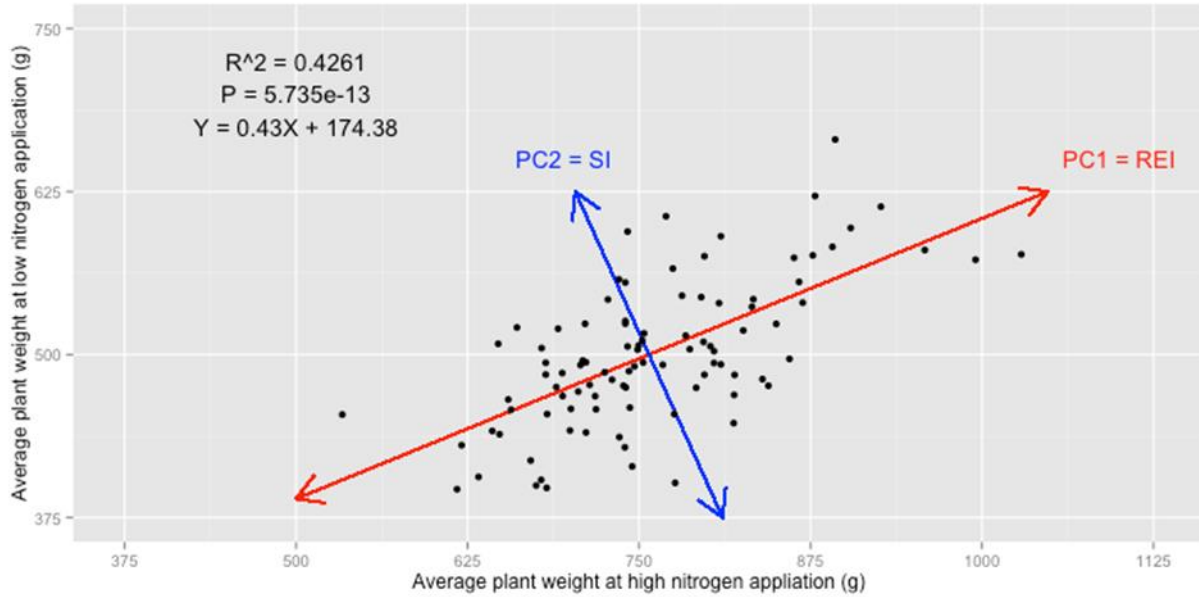
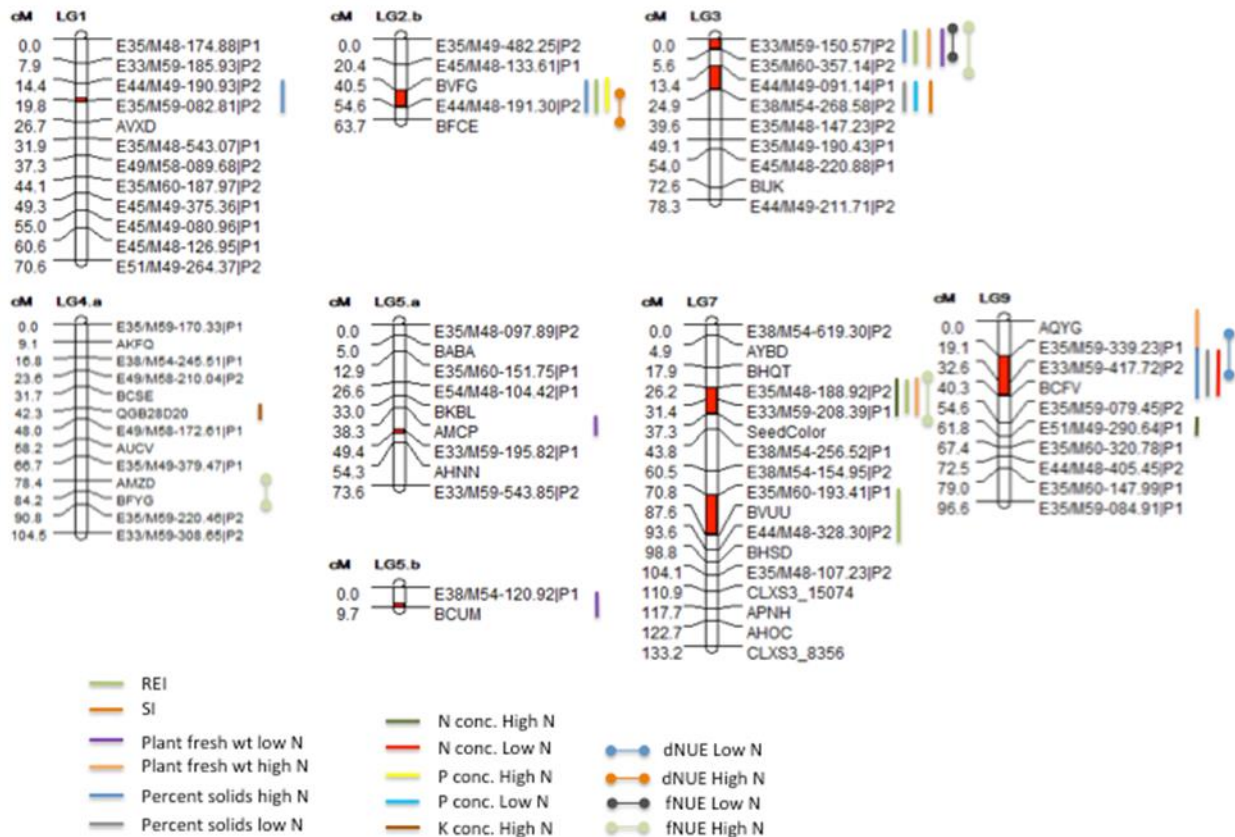


Figure 6. QTLs identified in the RIL population derived from Grand Rapids x Iceberg.



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. US96UC23. We are conducting similar studies for tipburn and 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. US96UC23 as well as additional populations, for traits that have not been studied yet.

Genome and Transcriptome Sequencing

The genome of lettuce was 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 reference genome of cv. Salinas 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. We sequenced the gene space of 96 RILs from the core mapping population derived from cv. Salinas x cv. acc. US96UC23 to validate genomic assemblies and develop a sequence-based ultra-dense genetic map. Very few chimeric genome scaffolds were identified and corrected, validating the quality of the assembly. We have now 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 these 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/>.

In addition, we have sequenced and are assembling the genomes of *L. serriola* (acc. US96UC23) and six additional lettuce genotypes (PI251246, Green Lake, Diana, La Brillante, Iceberg, and Valmaine). These are providing the sequences of alleles and haplotypes for several of the genes of interest described above.

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.

Databases

We continue to curate several publicly accessible databases for lettuce. The Compositdb contains several 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 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).