CALIFORNIA LEAFY GREENS RESEARCH PROGRAM

April 1, 2012, to March 31, 2013

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. This project 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 projects (ii) through (iv). We continue to make extensive use of new high-throughput sequencing and marker technologies. The genome of lettuce has been sequenced and assembled; annotation of the ~45,000 lettuce genes continues. Genotyping by sequencing of the core mapping population has assigned over 65% of the assembled genome to chromosomal positions. We are sequencing 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 are being 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, 2012 to March 31, 2013

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**
Oswaldo Ochoa
Marilena Christopoulou
Lorena Parra
Dean Lavelle
Miguel Macias Gonzalez
German Sandoya
Carlos Galeano
Alex Kozik
Huaqin Xu
Sebastian Reyes Chin Wo
Christopher Beitel
Teresa Jardini
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 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 a near-term 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 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), LMV, corky root, *Fusarium* and *Verticillium* wilts, and big vein, onto the consensus genetic map and the ultra-dense map as well as align with the genome sequence. The majority of resistance phenotypes are linked to NBS-LRR-encoding genes.

Mapping of Resistance to Downy Mildew

We are continuing to develop molecular markers to assist the selection of resistance genes. We are characterizing advanced breeding lines carrying a novel resistance from *L. saligna*. To fine map the resistance, progeny of a heterozygous line for the resistance was analyzed for molecular markers and phenotypic resistance. However, linkage to flanking markers was too loose for reliable use in marker assisted selection. Additional markers in the flanking regions are being assayed for polymorphism.

Identification of Markers for Corky Root Resistance

We are continuing to identify 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 have fine mapped the chromosomal region containing *cor* by characterizing numerous recombinants in a Green Lakes x Diana population using flanking molecular markers. These data were consistent with *cor* being in a highly recombinagenic region and has provided multiple markers very tightly linked to *cor*. We have localized the *cor* gene to two candidate genes on a single genomic scaffold sequence and are in the process of generating transgenics to identify the causal gene. Markers derived from these genes will be available for backcross programs to combine *cor* with downy mildew resistances from the novel sources.

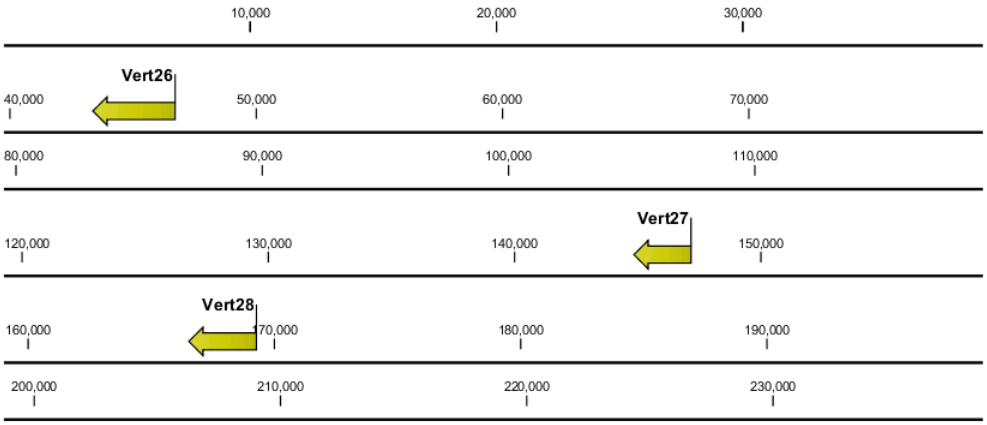
Genetic Analysis of Resistance to *Fusarium* Wilt

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 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 F_3 families from crosses between Lolla Rosa x Salinas and Salinas x Green Towers indicated one or two genes responsible for resistance. In 2012, we tested F_3 families from Lolla Rosa x Red Tide for resistance to *Fusarium* in a replicated field trial at UC D5uavis. We genotyped the F_2 population for molecular markers and constructed a genetic map. QTL analysis 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 results suggest that the QTLs from Lolla Rosa and Valmaine are in different regions of LG1.

Identification of candidate genes 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 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 EST is located on a genomic scaffold that has three genes similar to the two Verticillium resistance genes in tomato (*Ve1*: 3.1 Kb conferring resistance and *Ve2*: 3.4 Kb) (Fig. 1). The three lettuce homologs are similar in size to the tomato genes (Vert26: 3.4 , Vert27: 3.2 and Vert28: 3.2 Kb). Vert26 is highly expressed in cv. Salinas. We are currently in the process of identifying homologous copies of these genes in cultivars resistant to *V. dahlia* race 1, such as La Brillante, so we can determine the contribution of each gene to resistance. We are collaborating with the USDA group to study the function of these genes.

Figure 1. The positions and orientations of three genes with sequence similarity to the *Ve* gene of tomato on a genomic scaffold of *L. sativa* cv. Salinas scaffold.



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

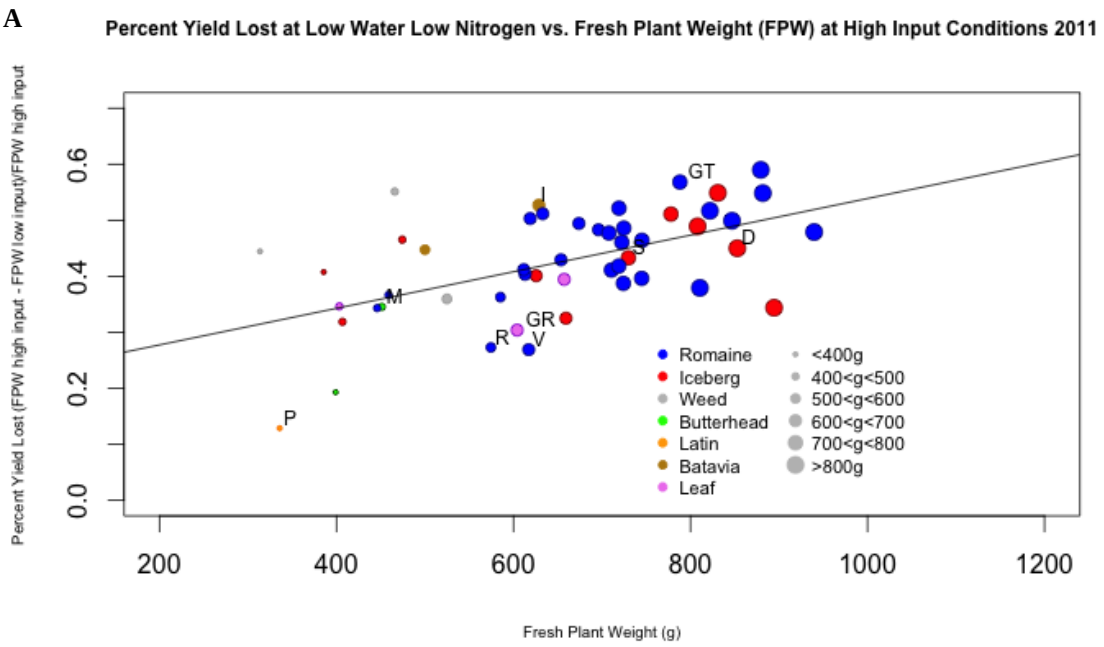
A lettuce field trial was conducted in summer 2011 in Spence, Salinas to investigate NUE and WUE as part of the USDA SCRI-funded project (with matching support from the CLGRB) entitled Next-Generation Lettuce Breeding: Genes to Growers. The PIs particularly involved in this component were M. Cahn and 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 after 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). A total of 50 lettuce varieties comprising iceberg (11), romaine (26), Batavia (2), butterhead (2), latin (1) and leaf (4) lettuce types and four primitive or wild types were trialled. 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. Data analysis showed significant diversity of NUE and WUE among the cultivars analyzed.

The follow-up trial in 2012 had two components: one to confirm some of the results from 2011 by re-growing and analyzing ten cultivars representing the most diverse responses to water and nitrogen treatments observed in 2011 and the other to analyze a segregating RIL population to investigate the genetic basis of NUE. Analysis of 2011 data showed a significant variation among the 50 varieties for sustaining plant weight at lower water and nitrogen inputs (stress conditions). A selection of 10 cultivars with diverse responses to water and nitrogen stress treatments in 2011 were re-trialed on 2012. The performance of these genotypes was similar in both years (Fig. 2.). Percent yield loss sustained was correlated with variety size. Bigger genotypes tended to lose more yield under stress conditions than smaller ones.

We compared the performance of the genotypes under stress and non-stress conditions by calculating the Susceptibility Index (SI) and the Relative Efficiency Index (REI) as described by Rosales-Serna *et al*. (2000. *Agrociencia* **34**:153-165). SI measures the susceptibility of a genotype to stress. REI measures the efficiency of a genotype to produce yield. Well performing genotypes in stress conditions are those with low SI and a high REI values. Fig. 3 shows the values and relationship between the two indexes observed in 2011 and 2012. The data for the 10 cultivars correlated well between the two years.

Genetic analysis of the RIL population between Grand Rapids (GR) x Iceberg (I) and QTL analyses for WUE, NUE and other traits are in progress.

Figure 2. Percent yield loss at low water and low nitrogen conditions *versus* the potential yield at high water high nitrogen in the 2011 (A) and 2012 (B) field trials. Genotypes trialed in 2012: Diplomat (D), Iceberg (I), Margarita (M), Pavane (P), Rouge d'Hiver (R), Salinas (S), Valcos (V), Grand Rapids (GR), and Green Towers (GT). Bullet colors represent different lettuce types and bullet size the weight of the lettuce head.



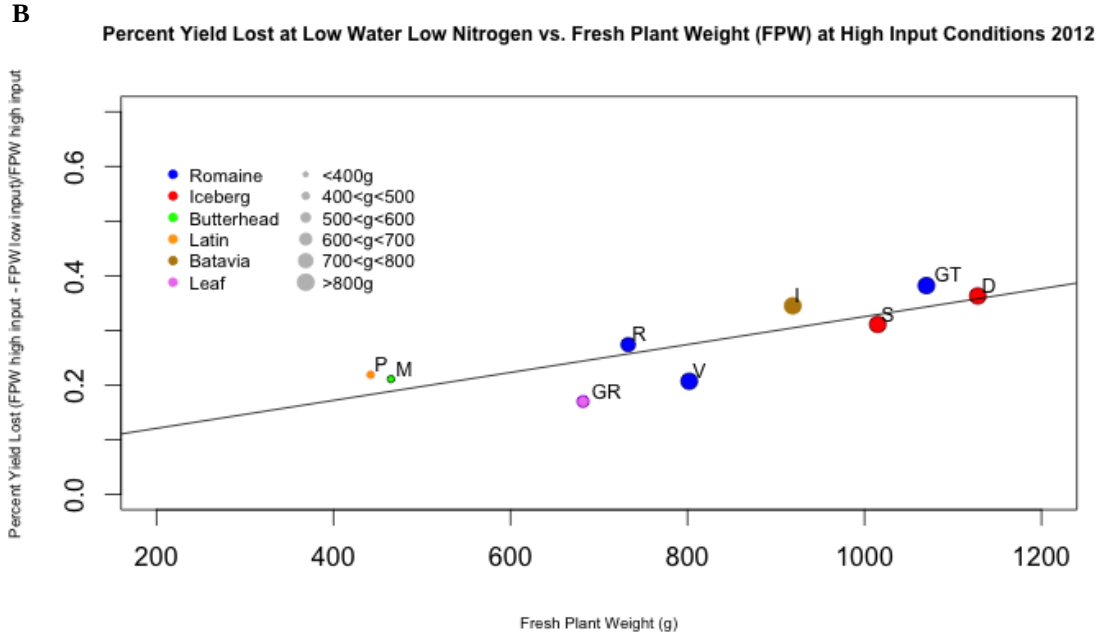
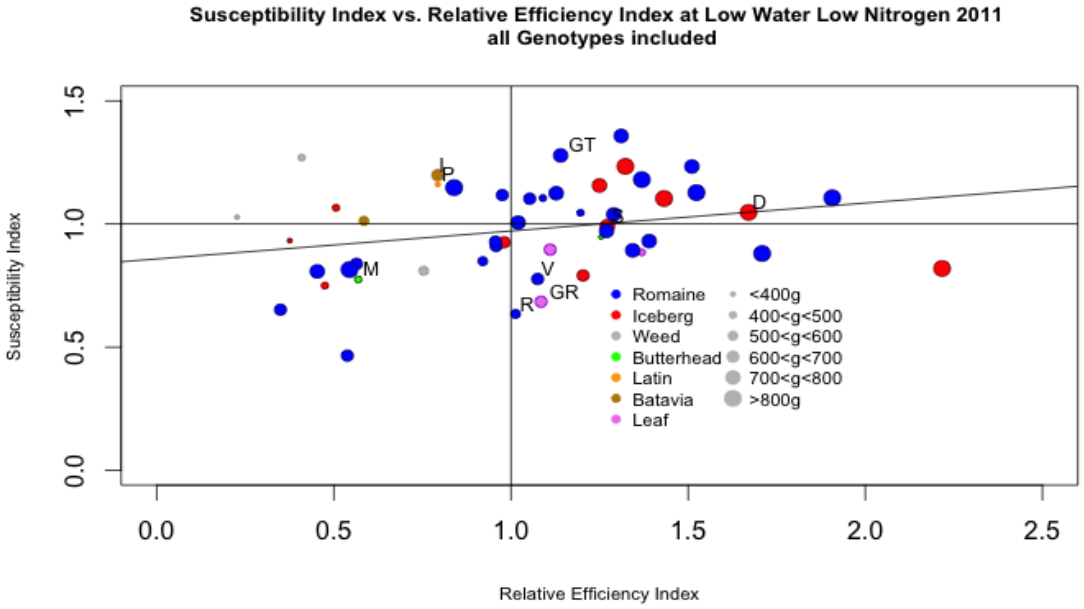
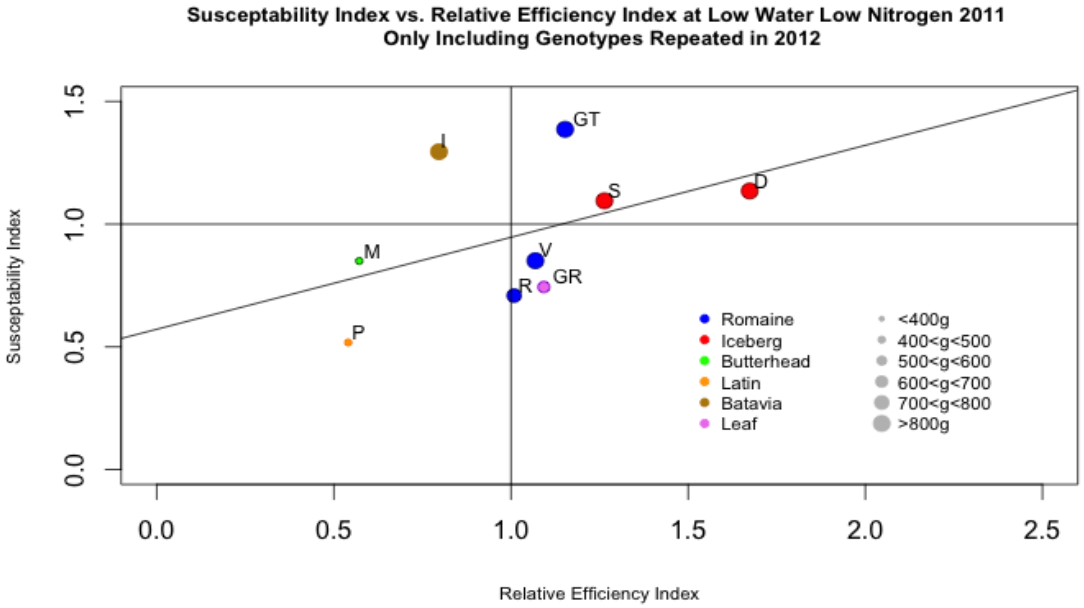


Figure 3. SI plotted *versus* REI for each genotype in the 2011 trial (A; 50 genotypes) and in the 2012 trial (B; 10 selected genotypes representative of the diversity observed in 2011).





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 financial 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. This lettuce chip contains sequences for detecting polymorphisms in approximately 35,000 unigenes in lettuce and

has facilitated rapid SNP discovery, genotyping, and mapping. 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.gc.ucdavis.edu/cgi-bin/gbrowse/lettucePublic/). The GBrowse display allows searching for EST nucleotide sequences and SNPs. Details of the ultra-dense map construction have been published in Truco *et al*. (2013. *G3, Genes, Genomes, Genetics.* **3**:617-631. doi:10.1534/g3.112.004929). Raw data used for map construction and maps of all the linkage groups can be viewed at http://chiplett.ucdavis.edu/map_2012. Over the past year we have sequenced the gene space of the same RILs to locate genomic assemblies and additional genes on the ultra-dense chromosomal map.

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 (See CLGRB 2011-2012 report). 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 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 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 and gene space from both *L. sativa* cv. Salinas and *L. serriola* acc. UC96US23 (Matvienko *et al*., 2013. *PLoS One.* **8**:e55913. doi:10.1371/journal.pone.0055913). 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.

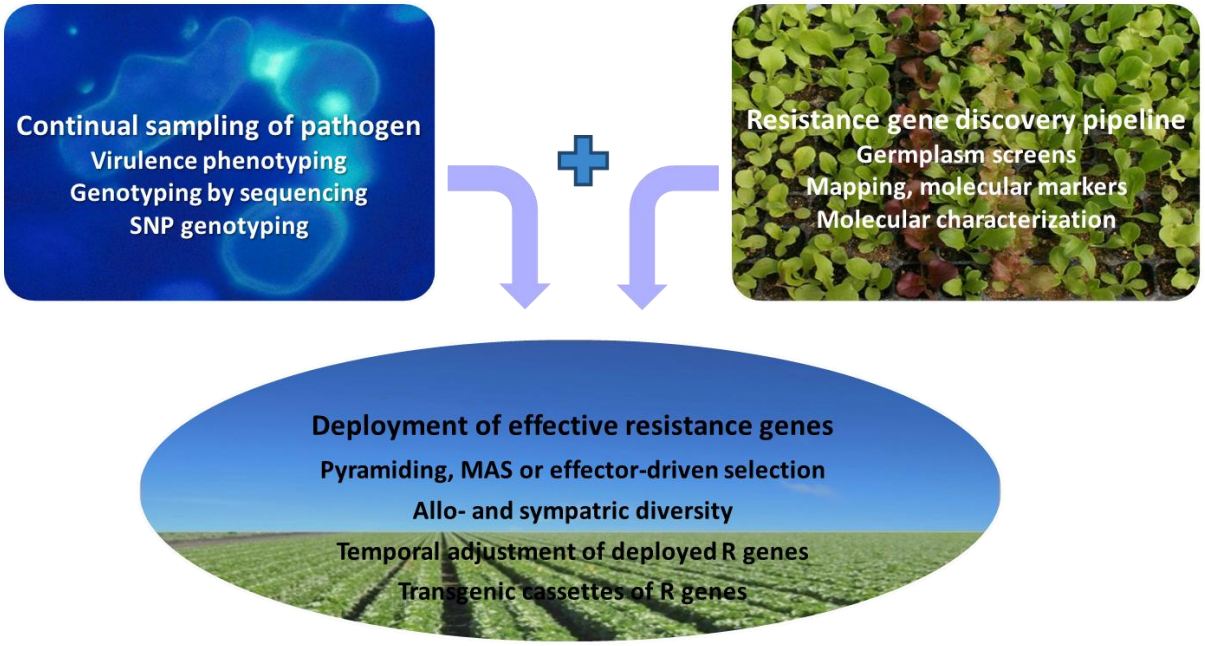
The genome of lettuce has been sequenced in collaboration with the BGI, Shenzen, China, funded by an international consortium of companies [Agrisemen (NL), Enza Zaden (NL), Gautier Semences (FR), Isi Sementi (IL), Monsanto Vegetable Seeds (USA), Rijk Zwaan (NL),

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*. 45,000 gene models; we are currently refining the gene annotations. The genome sequence is publically available (https://lgr.genomecenter.ucdavis.edu/). In addition, we are sequencing and assembling the genomes of *L. serriola* (acc. UC96US23) and six additional lettuce cultivars (PI251246, Greenlakes, Diana, La Brillante, Iceberg, and Valmaine).

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 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 (Fig. 4).

Figure 4. The influenza paradigm: using knowledge of pathogen population genetics to drive deployment of disease resistance genes.



Michelmore, Christopoulou, & Caldwell. Impacts of resistance gene genetics, function, and evolution on a durable future. *Ann. Rev. Phytopathol*. (2013).

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. 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.