CALIFORNIA LEAFY GREENS RESEARCH PROGRAM
April 1, 2022 to March 31, 2023

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 develop and 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) genetic mapping, cloning, and characterization of disease resistance genes; (iii) developing genomic resources; and (iv) comparative genomics to identify candidate genes controlling horticultural traits. We identified the promoter of the lettuce ubiquitin gene as providing high levels of expression; we are currently confirming the stability of its expression over generations using the chimeric *RUBY* transgene as a visible reporter. We continue to make extensive use of high-throughput sequencing, marker technologies, and genome editing for gene identification. Version 11 of the reference genome assembly of cv. Salinas, which has nine telomere-to-telomere scaffolds and improved contiguity and sequence accuracy, was uploaded to GenBank. A telomere-to-telomere, ~gapless v14 genome assembly based on HiFi and nanopore long-reads is nearing completion. We are currently generating T2T genome assemblies for ~150 additional lines to develop a pangenome for lettuce and a catalog of resistance gene candidates. Genome sequences are being mined for candidate genes for traits, such as disease resistance and development. We are continuing programs to study corky root resistance, Verticillium wilt resistance, and salinity tolerance in lettuce. We continue to curate several databases that include genetic, molecular marker, phenotypic, and sequence data for lettuce. Key sequences have been uploaded to GenBank to make them widely accessible.

CALIFORNIA LEAFY GREENS RESEARCH PROGRAM
April 1, 2022 to March 31, 2023

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: **Keri Cavanaugh, María José Truco, Dean Lavelle, Giulia Caliandro, Alex Kozik, Huaqin Xu, Sagayamary Sagayaradj, Kyle Fletcher, Maria Ferrer Ruiz, Beth Rowan,** UC Davis Genome Center and The Dept. of Plant Sciences, UC Davis.
Ivan Simko, Beiquan Mou, Steve Klosterman, UDSA-ARS, Salinas.
Michael Cahn, UC Cooperative Extension, Monterey Co.
Krishna Subbarao, Dept. of Plant Pathology, UC Davis.
Duke Pauli, School of Plant Sciences, U. Arizona.
German Sandoya, Horticultural Sciences Dept., U. Florida.
David Still, Dept. Plant Sciences, Cal Poly Pomona.

ABSTRACT:

We continue to develop and 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) genetic mapping, cloning, and characterization of disease resistance genes; (iii) developing genomic resources; and (iv) comparative genomics to identify candidate genes controlling horticultural traits. We identified the promoter of the lettuce ubiquitin gene as providing high levels of expression; we are currently confirming the stability of its expression over generations using the chimeric *RUBY* transgene as a visible reporter. We continue to make extensive use of high-throughput sequencing, marker technologies, and genome editing for gene identification. Version 11 of the reference genome assembly of cv. Salinas, which has nine telomere-to-telomere scaffolds and improved contiguity and sequence accuracy, was uploaded to GenBank. A telomere-to-telomere, ~gapless v14 genome assembly based on HiFi and nanopore long-reads is nearing completion. We are currently generating T2T genome assemblies for ~150 additional lines to develop a pangenome for lettuce and a catalog of resistance gene candidates. Genome sequences are being mined for candidate genes for traits, such as disease resistance and development. We are continuing programs to study corky root resistance, Verticillium wilt resistance, and salinity tolerance in lettuce. We continue to curate several databases that include genetic, molecular marker, phenotypic, and sequence data for lettuce. Key sequences have been uploaded to GenBank to make them widely accessible.

OBJECTIVES:

To develop and apply new methods for detecting, analyzing, and manipulating variation in lettuce. We continue to pursue these objectives in four overlapping sub-projects:

1) Analysis of transgenes in lettuce and genome editing.

2) Molecular identification of genes for disease resistance and other horticultural traits.

3) Development of genomic resources.

4) Utilization of comparative genomics to identify candidate genes controlling horticultural traits and development of molecular markers for them.

In the first three projects, we are mostly 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. Several of these studies have been funded from federal grants and support from seed companies. All projects were initiated with CLGRP funds and application of the results to lettuce improvement is supported by CLGRP 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. We continue to generate transgenics as components of other projects (see below); these provide additional data on transgene expression and stability.

Introduction of genes into lettuce using *A. tumefaciens* is routine. However, obtaining sustained high levels of transgene expression has been challenging, in part because the commonly used CaMV 35S promoter is frequently silenced in lettuce. Another commonly used promoter is from the ubiquitin gene. We cloned the ubiquitin promoter from lettuce as well as tested ubiquitin promoters from several species cloned by others. Last year, we initiated a study to determine which promoters can give high stable expression of transgenes, both within and across generations. A comparison of the 35S, *Arabidopsis thaliana* ubiquitin, and lettuce ubiquitin promoters revealed that the lettuce ubiquitin promoter drove high levels of expression in nearly all parts of the plant, as visualized by the RUBY reporter in the primary transgenics. We have now taken this study through to the next generation. This demonstrated that the lettuce ubiquitin promoter maintains high levels of expression in both self-progeny and in progeny derived from crosses with wild-type plants and was not silenced (Fig. 1), in contrast to the 35S promoter.

Since 2012, technology for precise genome editing based on the CRISPR system has been developed for use in plants and animals. Gene knock-outs are currently much easier than sequence replacements or additions but the enabling technology is advancing rapidly. We continue to develop and apply genome editing technology for lettuce as part of our gene identification strategy. We also have on-going experiments aimed at base editing (changing single bases in the DNA sequence).

In addition to validating potential candidate genes for disease resistance, we are targeting a subset of the many genes that have been previously functionally characterized in other plant species, particularly Arabidopsis. We are using CRISPR-mediated gene knockouts to deduce whether the homologous genes in lettuce have a similar function. We are initially investigating genes potentially controlling traits such as nutrient content, leaf, flower, and root development.

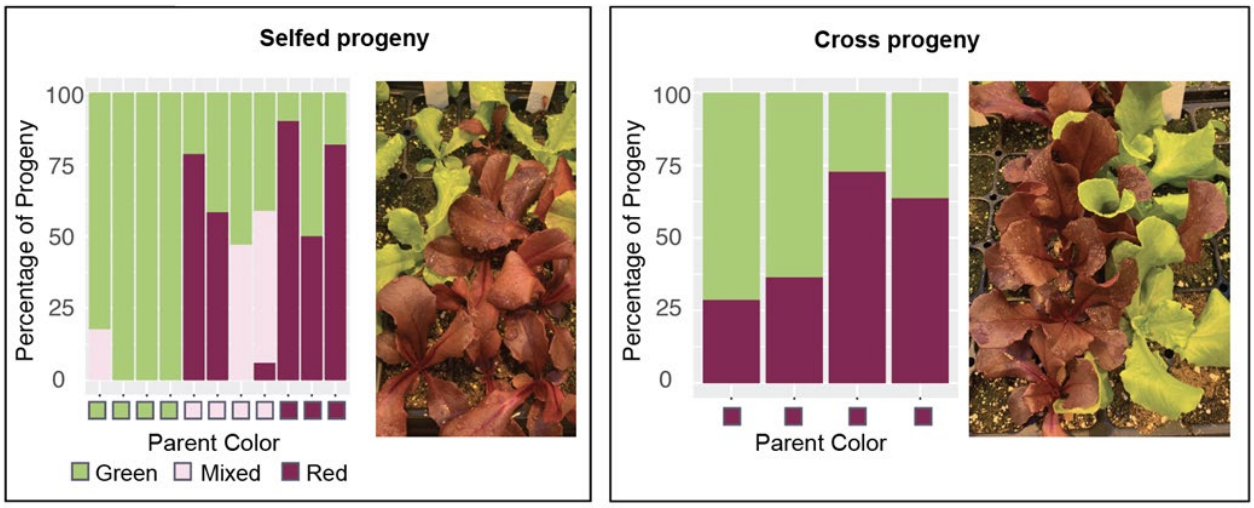


Figure 1. Expression of the RUBY marker driven by the lettuce ubiquitin promoter in selfed progeny of the primary transgenic lines (Cobham Green background) and the progeny of crosses between the primary transgenics (as females) and wild-type Cobham green (as males). Images show progeny from one of the lines in each case. The relative proportions of red and green progeny reflects the segregation of the transgene encoding RUBY.

In the longer term, once the technology has been adequately developed, we will use genome editing of lettuce to create stacks of resistance genes containing several resistance and other genes at single chromosomal positions so that they will be inherited as single Mendelian loci in breeding programs. 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, technology for inserting genes, preferably without tissue culture, needs to be developed and genes for resistance to each disease need to be identified at the molecular level.

Molecular identification of genes for disease resistance and other horticultural traits

We previously mapped multiple genes for resistance to downy mildew (DM), corky root, and *Fusarium* and *Verticillium* wilts onto the consensus genetic map and placed them on the genome sequence. This provides 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 (MRCs) on Chromosomes 1, 2, 3, 4, and 8. The majority of these resistance phenotypes are linked to NB-LRR-encoding (NLR) genes, as described in previous CLGRB reports, that provide markers for selecting for these resistances. Two populations were grown and sequenced to map new *Dm* genes in the past year; analysis is on-going.

Resistance to corky root

We previously showed that resistance to corky root was determined by the recessive *cor* locus on Chromosome 3. Fine mapping and genomic comparisons of resistant and susceptible lines identified four candidate genes for resistance to corky root at this locus, one or more of which could encode a susceptibility factor. Constructs for CRISPR-Cas9-mediated gene knock-outs with guide RNAs targeting each of these candidate genes in the *cor* region were transformed into cv. Salinas. We recovered T_2 seed from 32 transgenic lines that are ready to be evaluated for resistance to *Rhizorhapis suberifaciens* strain CA1.

We needed a pathogenic stock of *R. suberifaciens* CA1 in order to screen the transgenic knock-out lines for resistance. *R. suberifaciens* is difficult to maintain in culture and our stock of CA1 no longer elicited the characteristic reactions on susceptible and resistant lines of lettuce. Fortunately, we were able to obtain another sample of CA1 from Isolde Francis, California State University, Bakersfield that was derived from our original CA1 strain. We completely sequenced and annotated this and 12 other pathogenic and non-pathogenic strains of *Rhizorhapis* species and related genera, including two strains of *R. suberifaciens* that are being used by companies for screening for resistance to corky root (see last year's report). During the past year three new strains were received from Isolde Francis and sequenced. Companies interested in submitting additional strains for sequencing should contact Maria Ferrer (mferrerruiz@ucdavis.edu). We are also interested in receiving contemporary samples of corky root from infected fields.

This sample of CA1 (designated CA1F) elicits the appropriate reactions on susceptible and resistant lines and therefore will be used for screening the knock-out lines (Fig. 2). Over the past year we have been assessing the virulence of all sequenced isolates on a panel of resistant and susceptible lines as well as optimizing the protocol for reliable symptom development. We are also developing diagnostic assays for pathogenic and non-pathogenic strains of *R. suberifaciens, Rhizorhabdus* sp., *Sb. mellinum*, *Sb. xanthum*, and *Sphingopyxis* sp. based on their assemblies. These assays will also be useful for diagnosis in the field as well as epidemiological studies.

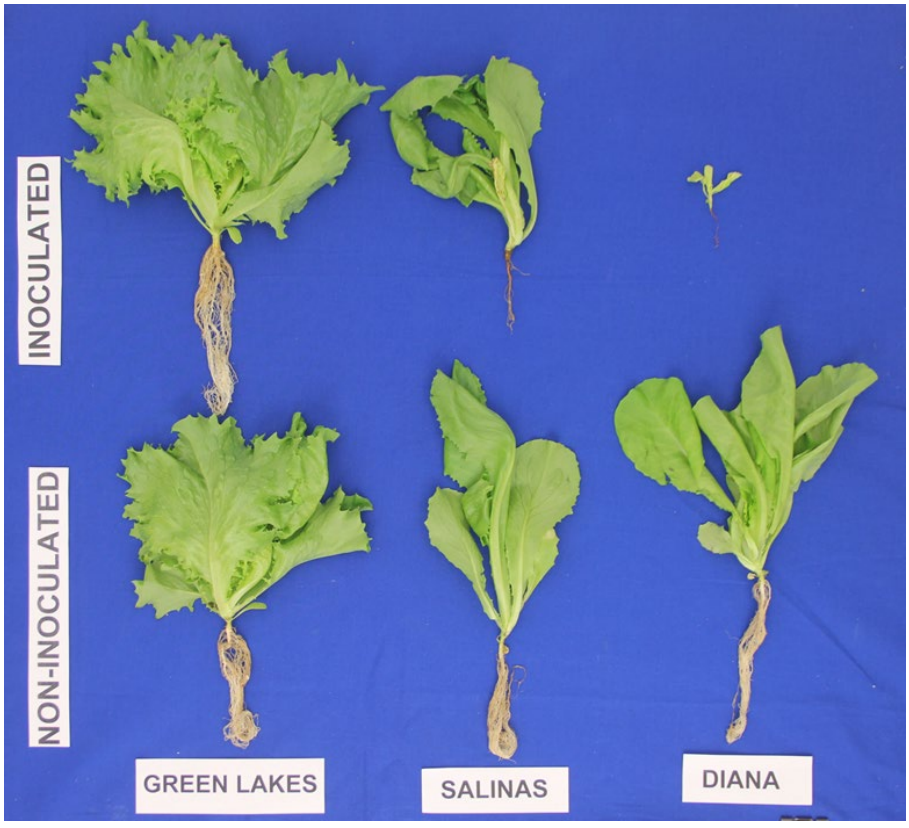


Figure 2. Reaction of resistant (Greenlakes) and susceptible (Salinas and Diana) cultivars to inoculation with *R. suberifaciens* CA1 five weeks post-inoculation of young seedlings.

Resistance to *Verticillium dahliae* race 1

We are dissecting the *Vr1* locus using *Agrobacterium*-mediated transient expression of the Ave1 effector. Previous studies identified the *Vr1* locus as the key determinant of resistance in lettuce against *Verticillium dahliae* race 1. The locus contains multiple *LsVe* genes that encode receptor-like proteins with sequence similarity to the *Ve* resistance genes found in tomato (Inderbitzin *et al.,* 2019. BMC Plant Biol 19, 305). However, it is unclear which of these *LsVe* genes contribute to resistance. We are primarily focused on characterizing the functionality of the *LsVe1* gene within the *Vr1* locus. The Ave1 effector is known to interact with the Ve1 protein in resistant tomato genotypes, activating a downstream hypersensitive response (HR) as part of the defense mechanism (de Jonge *et al.,* 2012. Proc. Natl. Acad. Sci. 109, 5110–5115). We infiltrated a panel consisting of resistant and susceptible lines with Agrobacterium strains carrying the Ave1 gene expressed from the lettuce ubiquitin promoter. The resistant lines recognized Ave1 and elicited necrotic response (Fig. 3). Screening for reaction to Ave1 is much simpler and faster than screening for resistance with fungal strains.

Following optimization with the panel, we will proceed to screen knock-out lines and complementation lines. We have created knock-out lines of lettuce through CRISPR-Cas9 gene editing with two guide RNAs targeting the LsVe1 gene that were transformed into cv. La Brillante. Additionally, we have generated complementation lines by introducing the *LsVe1* gene, cloned from the resistant La Brillante cultivar into the susceptible Salinas and Cobham Green cultivars. We recovered T_3 seed from 5 transgenic knock-out lines and T_3 seed from 17 transgenic complementation lines that are ready to be evaluated. By comparing the responses of these lines to the Ave1 effector, we will be able to evaluate the contribution of *LsVe1* to the observed resistance response against *V. dahliae* race 1.

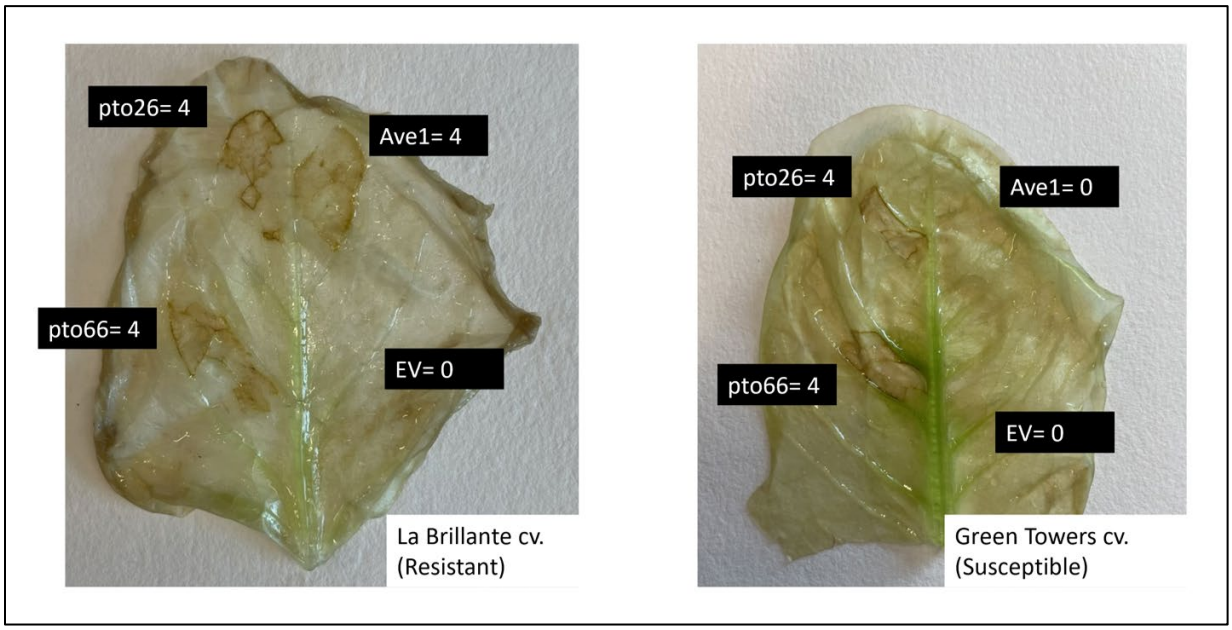


Figure 3. Agroinfiltration results of two lettuce genotypes from the panel with the Ave1 effector from *V. dahliae* race 1. The Ave1 effector, the positive controls pto26 (HopM1) and pto66 (AvrPto) effectors, and the negative control EV (empty vector) were agroinfiltrated in lettuce leaves and leaves were scored for their reaction five days post-infiltration. The leaves were cleared in 70% EtOH prior to imaging. Qualitive scores were converted into numeric scores for

data analysis: 0 = no reaction, 1 = mild chlorosis, 2 = chlorosis, 3 = mild necrosis or mixture of chlorosis and necrosis, and 4 = full necrosis. The resistant La Brillante cv. showed necrosis in the areas agroinfiltrated with the Ave1 effector and positive controls in contrast to the susceptible Green Towers cv. that only showed necrosis in the positive controls.

Remote Sensing

In collaboration with Dr. Duke Pauli (University of Arizona), we conducted a field trial at the Maricopa Agricultural Center, AZ to study water use efficiency in lettuce. The Maricopa Agricultural Center hosts the largest field analytical robot in the world. The high-throughput phenotyping field-scanning robot has a gantry that moves on rails while imaging the crops growing below with a diverse array of cameras and sensors, including scanners for hyperspectral and thermal imaging, 3D laser, and chlorophyll fluorescence (Fig. 3). This is the third trial conducted at the Maricopa Agricultural Center to understand the genetic basis controlling agronomic and stress-adaptive traits, to monitor growth and development, and to quantify water deficit stress and the impacts on photosynthetic efficiency as well as to develop analytical tools for remote sensing.

The 2021–2022 trial consisted of 272 lines grown in two different water treatments, with three repetitions per treatment for a total of 1,632 plots (see 2021- 2022 CLGRB report). This year trial had partially the same layout. We repeated the set of ~ 200 diverse cultivars grown in 2021-2022 for the Genome Wide Association Study (GWAS). In addition, we tested 110 lines of a RIL population from a cross between cultivar Salinas and a *L. serriola* line segregating for root architecture. Dr. Duke Pauli's team developed a bed lifter for loosening the soil on the trial beds allowing the collection of a total of 1,008 roots for imaging and analysis for root architecture that is ongoing (Fig. 4).

Figure 4. Examples of roots collected for imaging.

Figure 5. Harvesting the trial at the University of Arizona Maricopa Agricultural Center in March 2022-2023. The Lemnatec Field Scanalyzer is at the end of the plot; the rails for the gantry are visible on either side.

We harvested one head of lettuce for each plot measuring fresh and dry weight and stem length (Fig. 5). These measurements will be used to ground truth the growth measurements made with remote sensors on the gantry. The remote sensor measurements of all plants in each plot will be used for the GWAS of WUE. We have finished whole genome sequencing of all of the lines for which sequence data was not already available. These sequences are being aligned to the reference genome of Salinas to detect sequence variation (SNPs). These will be used to identify genomic regions associated with phenotypic variation.

Genetics of Salt Tolerance in Wild *Lactuca* Species

We continue to investigate the genetic basis of salt tolerance as the foundation for breeding lettuce cultivars with increased salt tolerance. As described in last year's report, we initiated a genetic analysis of lettuce root system architecture (RSA) in response to salinity stress using an existing RIL population derived from Armenian 999 x PI 251246. Seedlings were grown on plates of ½ MS media with or without 100 mM NaCl for seven days. Photos were taken at multiple time points and images were analyzed using SmartRoot software to quantify root lengths (Fig. 6A). In initial screens of the parents, PI 251246 exhibited a significant decrease in primary root length, lateral root length, total root length, and lateral root number ($p < 1.0^{-3}$) compared to Armenian 999, which did not exhibit significant differences in the same RSA traits (Fig. 6B). Additionally,

Armenian 999 exhibited little lateral root production compared to PI 251246. There was a significant genotype x treatment effect ($p < 1.0^{-2}$) for all RSA traits (Fig. 6B). One hundred and twenty RILs were evaluated under similar conditions for nine days to identify QTLs for RSA changes due to salinity stress.

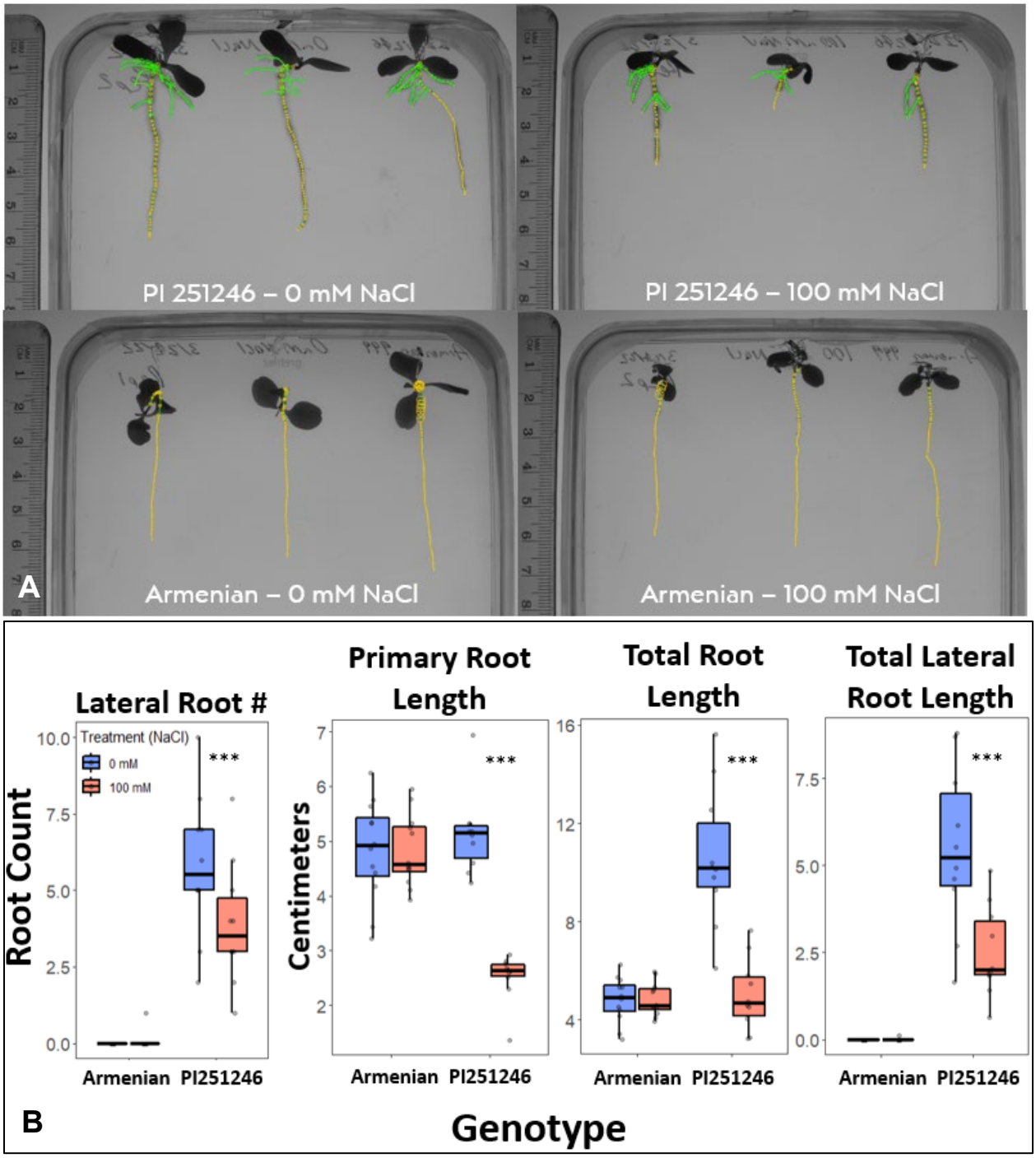
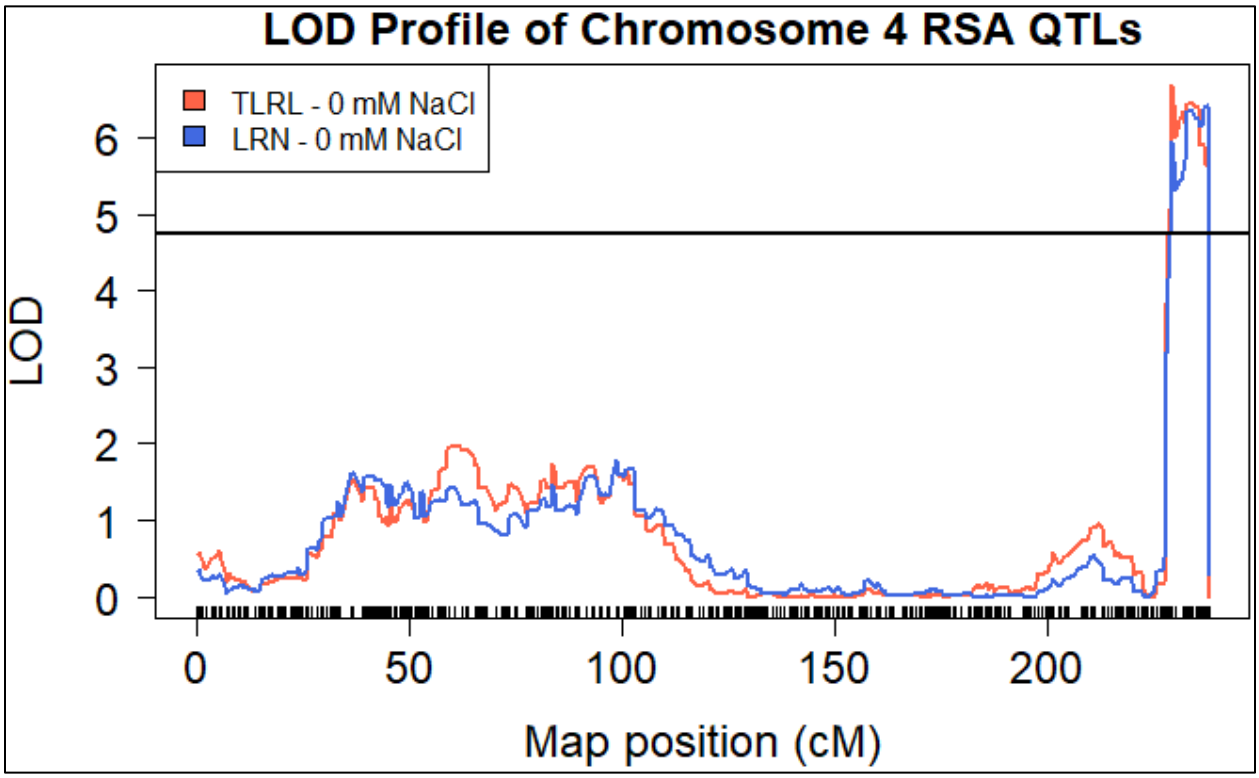


Figure 6. Root system architecture of Armenian 999 and PI 251246 grown under saline conditions. A) Roots of PI 251246 (top) and Armenian (bottom) seedlings grown on ½ MS media

with either 0 or 100 mM NaCl on day 7. Roots were traced with SmartRoot. B) RSA traits measured in Armenian 999 and PI 251246.

Linkage analysis was achieved using composite interval mapping in Rstudio. A genetic map of the Armenian 999 x PI 251246 RIL population with 2,412 k-mer markers was used. Major QTLs (>15% PVE) were identified on chromosome 4 for total lateral root length (*qTLRL4*) and lateral root number (*qLRN4*) with 16.64% and 19.22% PVE respectively in 0 mM NaCl exclusively (Fig. 7A). Major QTLs were identified on chromosome 5 for primary root length (*qPRL5*) and total root length (*qTRL5*) with 21.42% and 21.52% PVE respectively in 100 mM NaCl exclusively (Fig. 7B). Major QTLs were identified on chromosome 9 for total root length (*qTRL9*), total lateral root length (*qTLRL9*), and lateral root number (*qLRN9*) with 24.4%, 25.5%, 23.59% PVE respectively in 100 mM NaCl exclusively (Fig. XC). A minor QTL was identified on chromosome 5 for lateral root number (*qLRN5*) with 12.4% PVE exclusively in 100 mM NaCl (Fig 7B).



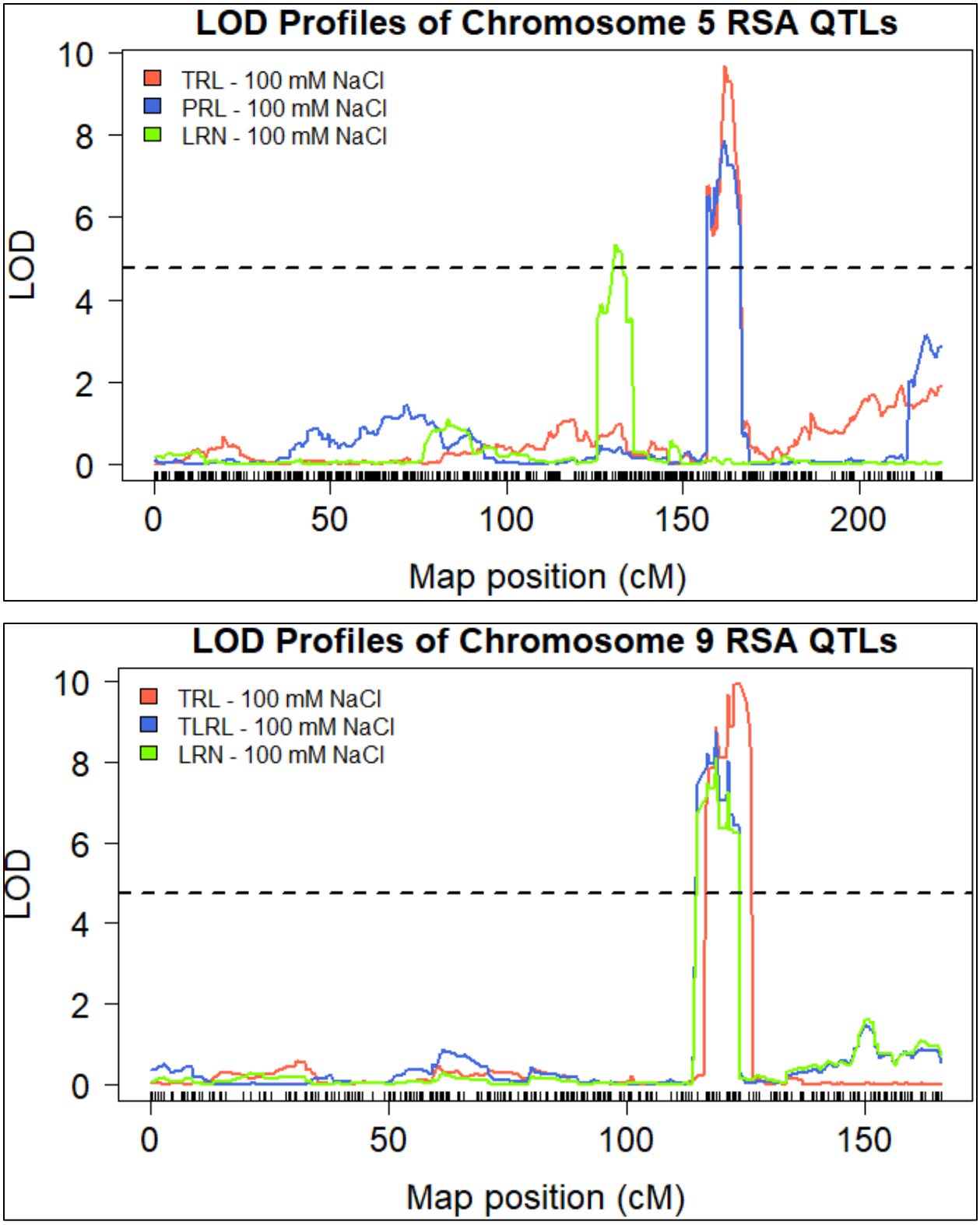


Fig. 7. LOD Profiles from composite interval mapping of root system architecture traits in Armenian 999 x PI 251246 RIL population. A) LOD profiles of *qTLRL4* and *qLRN4*. B) LOD profiles of *qTRL5*, *qPRL5*, and *qLRN5* C). LOD profiles of *qTRL9*

Development of genomic resources

We have continued to refine the reference genome of lettuce to provide the foundation for understanding the genetic architecture of traits in *Lactuca* spp. Last year we uploaded the v11 sequence of cv. Salinas to Genbank to make it widely available. Genbank then generated a Refseq annotation. After many years of effort and constantly improving sequencing technologies, we are finally approaching a complete, telomere-to-telomere (T2T), ~gapless assembly for the reference cultivar cv. Salinas (Fig. 8). The final v14 assembly is based on both PacBio HiFi and Oxford Nanopore long reads.

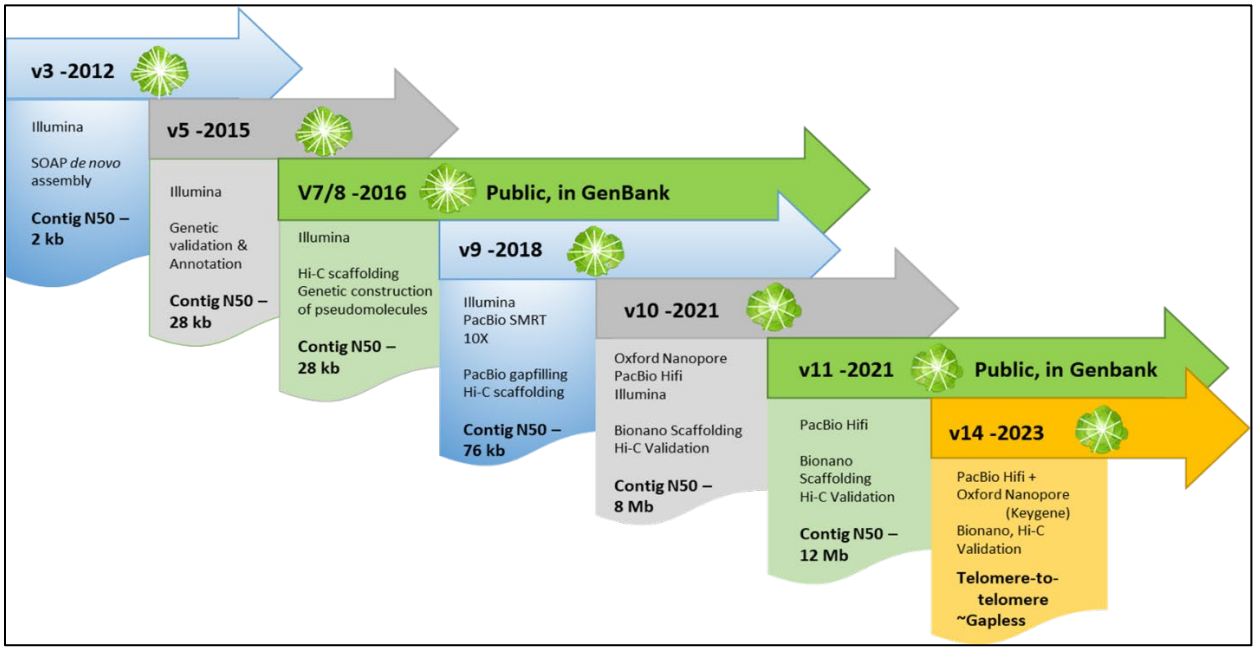


Figure 8. History of improvements in the genome sequence of *L. sativa* cv. Salinas. This was funded by the International Lettuce Genomics Consortium (ILGC 1, 2, and 3), with support from USDA FFAR Crops of the Future Program and 12, 17, and 14 companies, respectively.

It is now possible to rapidly and inexpensively generate T2T assemblies of any lettuce genotype. In April, the UC Davis Genome Center was one of the first places to get a PacBio DNA sequencer that efficiently generates large amounts of HiFi reads. We are currently using this to generate telomere-to-telomere assemblies of ~150 wild and cultivated *Lactuca* spp. that exhibit a broad diversity of disease resistance and developmental phenotypes. These data will provide an extensive catalog of resistance genes that will provide the foundation for pyramiding resistance genes using marker-assisted selection.

In order to generate a pan-genome that captures the genetic diversity within *Lactuca spp*., we have sequenced several diverse wild and cultivated lettuce lines using PacBio HiFi / Oxford Nanopore long-read technologies and Illumina short-read sequences. Currently high-quality reference assemblies of *L. sativa* cvs. Salinas, Ninja, *La Brillante*, PI251246, VIAE, and *L. serriola* accessions Armenian and US96UC23 are assembled (Table 4; Fig. 8). D*e novo* annotation is currently underway to reveal differences in gene content, particularly with regard to their resistance gene repertoires.

Databases

We continue to curate several publicly accessible databases for lettuce accessible through http://michelmorelab.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 of genetic chromosomal pseudomolecules. The majority of the sequencing data is uploaded to GenBank to facilitate broad access and avoid legacy issues. 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).