California Leafy Greens Research Program
Final Report
April 1, 2020 to March 31, 2021

Project Title: HARNESSING THE LEAF MICROBIOME FOR IMPROVED FOOD SAFETY IN LETTUCE

Principle Investigator: Gail Taylor, University of California, Davis, CA **gtaylor@ucdavis.edu**
Cooperating Personnel: Annabelle Damerum, University of California, Davis, CA **adamerum@ucdavis.edu**; Professor Maeli Melotto, University of California, Davis, Professor Richard Michelmore, University of California, Davis, CA; Dr Ivan Simko, USDA, Salinas
Technical Staff: Suzanne Milner, University of California, Davis, CA **smilne@ucdavis.edu**

ABSTRACT

Leafy crops, including lettuce, have the potential to provide nutrient dense and versatile fresh produce, but eaten raw, with minimal processing, they present a risk to human health and recent events for romaine lettuce with several recalls have shown the devastation that this may have on the industry. New approaches are needed to address the delivery of safe and clean raw leaves for the processed salad leaf market. Leaves are host to a diverse and numerous community of microorganisms that sit closely adhered to leaves in a biofilm and as such they may be difficult to dislodge and remove, even in the most professional and high quality wash lines, presenting a safety risk to human health and to spoilage. The overarching aim of this new research project was to quantify the leaf microbial composition of commercial cultivars of lettuce and to elucidate the impact of cultivar, leaf characteristics and swap-over season on the lettuce leaf microbiome. In the first year, we screened a panel of commercial cultivars and the genome wide association study (GWAS) population developed by Ivan Simko in two field trials in the Salinas valley and in El Centro, Imperial valley during the fall and spring *swap over* seasons, respectively. For the first time, we assessed a suite of traits known to be involved in determining abundance and diversity of leaf bacterial taxa including leaf hydrophobicity, epidermal cell edges, stomatal numbers and size which are known to influence the attachment of bacteria on the leaf surface. Ongoing work is isolating the DNA from epiphytic (surface) and endophytic (internalized) microorganisms which will be sequenced using taxonomic biomarkers by the end of 2021. This will enable us to detect genetic loci in the lettuce genome influencing microorganism abundance and diversity by GWAS.

OBJECTIVES

1. Screen commercial cultivars of lettuce (the "commercial panel"), supplied on an agreed list by CLGRB to determine how the leaf microbiome varies with contrasting lettuce cultivars during swap over season in Salinas and Yuma/Imperial.
   Deliverable: A database that describes how the lettuce leaf microbiome varies, depending on cultivar with insight into the abundance and diversity of human pathogens with respect to crop and environment (between farm and across the season)
   Progress: field work complete, DNA sequencing in progress and database available end 2021.
2. Assess commercially grown lettuce cultivars (commercial panel) for leaf characteristics thought to impact the ability of microbes to attach and proliferate on lettuce leaves, quantified alongside the leaf microbiome in 1. This will include up to ten of the cultivars

being used in the laboratory of Maeli Melotto, grown at the Salinas and Yuma/Imperial Valley field site

Deliverable: *Identification of lettuce cultivars that are more able to withstand microbial attachment and proliferation, linked to the laboratory assay being test by Melotto, to determine viability of a lab assay.*

Progress: ***completed analysis of 20 commercial cultivars identifying differences in leaf surface traits.***

3. Quantify the lettuce leaf microbiome in a wide mapping population romaine type lettuce cultivars (the "diverse panel")

   Progress: ***completed traiting of leaf surface characteristics and QTL analysis will be completed early 2022.***

PROCEDURES

Objective 1

Along with the diverse panel of 493 lettuce accessions developed by the lab of Ivan Simko (Sthapit Kandel *et al.*, 2020) (USDA) used to conduct genome-wide association studies (GWAS; see objective 3), we grew 23 current cultivars utilized in the industry (commercial panel), provided to us by the Calgreens board (Table 1). This included Romaine, butter and red and green leaf types.

We completed 2 field trials at the USDA research site in Salinas (Fall; August-November, 2020) and the UC Desert Research Center in El Centro (Spring; February-April, 2021):

1) For the Salinas trial, seedlings were sown in a greenhouse at UC Davis on August 9th-10th and were watered by mist irrigation three times daily. On September 1st, seedlings were transported to Salinas and transplanted into the field and were harvested 9 weeks later on the 2-3rd of November.
2) For the El Centro trial, seedlings were sown from February 1st-3rd in a greenhouse at the UC Desert Research Center and were misted for 1 minute every hour. On February 16th, seedlings were transplanted into the field and were harvested 7 weeks later from the 5th-7th April.

For both trials, lettuce lines were planted in a fully randomized, three block trial design. Each 40-inch bed comprised two planting rows, with each row spaced 12 inches apart. The two outer beds acted as guard rows, along with four plants at the beginning and end of each row. For the commercial panel, six replicates were sewn per line. For the Salinas trial, two plants constituted one replicate (planted consecutively in a row), with six beds of experimental plants each 274 feet long and two beds of guard plants (Figure 1A). For the El Centro trial, three plants constituted one replicate (planted consecutively in a row), with nine beds of experimental plants 322 feet long with two guard rows. For both trials, we utilized the PaperPot Co. transplanting system (**https://paperpot.co/**). In this system, seedlings are sown into collapsible paper-chain pots into solid-bottom germination trays at 12 inch spacing. Seedlings in the biodegradable paper-chain pots can then be rapidly transplanted directly into the field using the mechanical transplanter.

For the Salinas trial, at head harvest maturity, three leaves of varying ages (outer, middle, inner) were aseptically flash-frozen into liquid nitrogen and stored on dry ice. Samples were transported back to the Taylor laboratory at UC Davis on dry ice, where they were stored at -80ºC for microbiome analyses. The rest of the head was harvested to measure leaf surface traits including stomata density, length, and width, epidermal cell number and area (Objective

2). For the El Centro trial, at head harvest maturity, one whole head was aseptically frozen on dry ice and transported to the Taylor laboratory at UC Davis where they were stored at -20ºC until processing for microbiome analyses. Another whole head was harvested to measure stomata density, length, and width, epidermal cell number and area as well as contact angle; an indication of leaf hydrophobicity (Objective 2).

For isolation of microbial DNA from the lettuce leaves, the whole collected sample was ground to a fine powder in pre-sterilized (autoclaved) pestle and mortars in liquid nitrogen under a laminar flow hood to prevent sample contamination. Tissue was transferred into a 2mL Eppendorf tube and stored at -80˚C prior to DNA isolation. DNA was isolated using the DNeasy PowerSoil Pro 96 well extraction kit (Qiagen, USA). Samples were loaded into the PowerBead pro plate according to the protocol developed by Custer and Dibner (2020). Briefly, the PowerBead plate was covered with a perforated, sterile film. A flame-sterilized spatula was cleaned further in 10% v/v bleach and air dried. This spatula was used to transfer ~100mg tissue from the 2mL Eppendorf tube to a sterile 200µL PCR tube, filling approximately 60% of the PCR tube and this tube was inverted to ensure the tissue collected at the cap end. Surgical scissors were flame and bleach sterilized and used to cut the tip of the PCR tube, which was inserted into the well of the PowerBead plate, piercing through the perforated film. This process was repeated until 95 wells were filled, leaving one well for a negative control, and the PowerBead plate was centrifuged at 2,500 rpm to release tissue from the PCR tube into the well. All actions were performed under a laminar flow hood to prevent contamination. DNA was then isolated according to the manufacturers protocol.

Isolated DNA will be used to prepare sequencing libraries using the bacterial 16S primer pair 799F (5'-AACMGGATTAGATACCCKG-3') and 1115R (5'-AGGGTTGCGCTCGTTRC-3') to exclude chloroplast DNA and the fungal primers fITS7 (5'-GTGARTCATCGAATCTTTG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') and the QIIME2 bioinformatic pipeline will be used to assess the bacterial and fungal microbiomes as described in Damerum *et al.* (2021).

Table 1 Cultivars investigated for microbiome analyses

Variety	Type	S/C
Abilene	Romaine	18724
Milagro	Butter	12687
Steamboat	Lettuce	24876
True Heart	Romaine	18603
Solid King	Romaine	18812
Cuervo	Butter	11261
Omega 42	Lettuce	21550
Bergam's Green	Green Leaf	25010
Big Star	Green Leaf	18464
Pueblo	Lettuce	11327
Red Tide	Red Leaf	20776
Republic	Lettuce	18327
Red Mist	Red Leaf	25767
Solid Heart	Romaine	18455
Rio Bravo	Romaine	18926
Somerset	Lettuce	24317
Cielo Verde	Green Leaf	20248

Reliant	Lettuce	18718
Tropicana	Green Leaf	18575
Holbrook	Romaine	11275
Duqueshe	Romaine	18581
Solan	Butter	18331
Lucan	Butter	17520

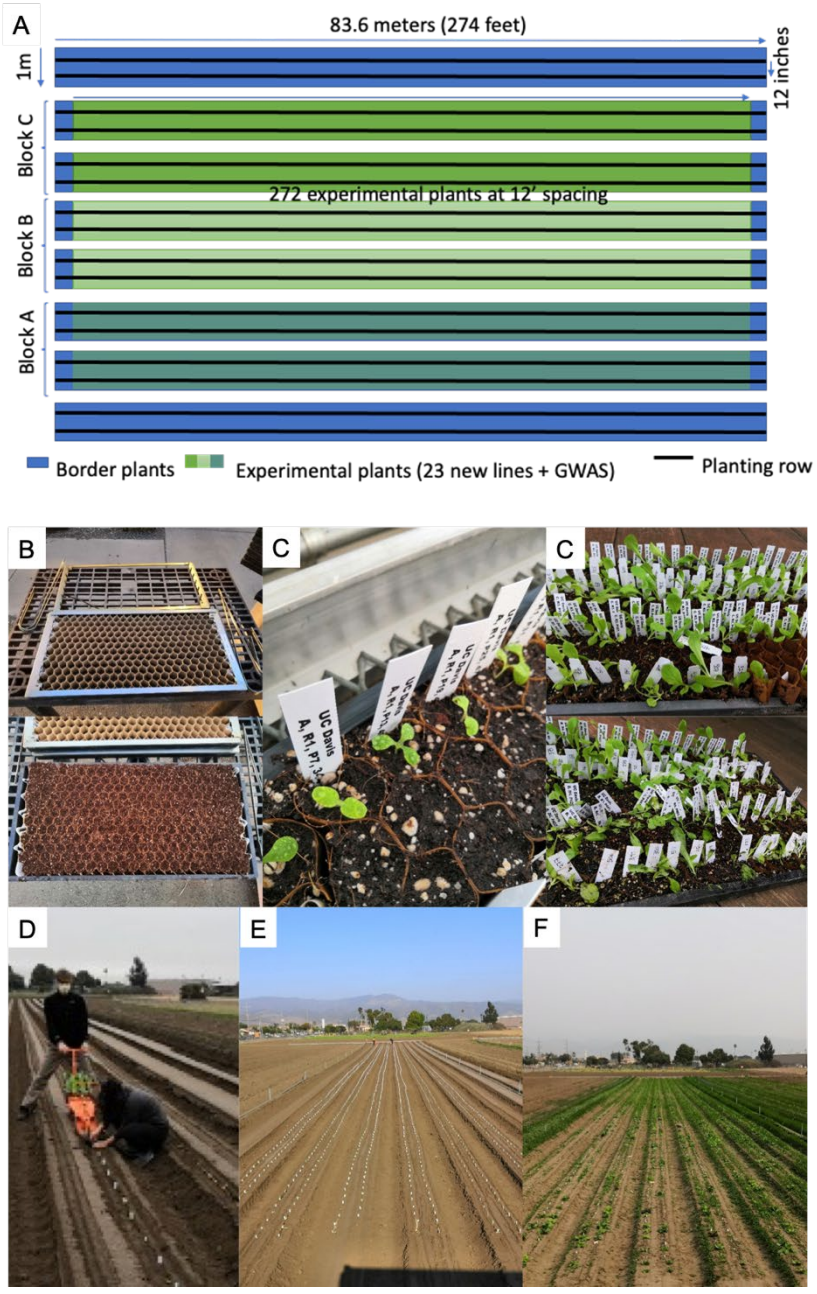


Figure 1 Example trial layout for the Salinas 2020 trial (A). Paper pot cells are spread and filled with soil (B) and a single seed sown per cell at 12 inch spacing (C). Seedlings ready to transplant (C) using the paper pot transplanter (D). Trial 1-week (E) and 4-weeks (F) post-transplant.

Objective 2

The commercial panel was grown in two field trials and samples were harvested as detailed in objective 1. For the Salinas 2020 trial, after collecting 3 leaves from a whole head at varying ages (outer, middle, inner), the rest of the head was utilized for phenotyping of leaf surface traits. An epidermal imprint was taken from one random leaf (as described in Damerum *et al.*, 2021) by applying a thin layer of clear nail polish to the abaxial leaf surface, enabling this to dry, removing with a piece of sellotape and adhering to a microscope slide. Slides were imaged using a light microscope attached to a lens camera and images were captured, which were analyzed for stomata density, length and width and epidermal cell area and perimeter using the ImageJ software (Figre 2; Schneider *et al.*, 2012).

For the El Centro 2021 trial, one whole head was harvested for phenotyping of leaf surface traits. As for the Salinas 2020 harvest, an epidermal imprint was taken from one random leaf. Additionally, contact angle, indicating the leaf surface hydrophobicity, was measured from both the abaxial and adaxial leaf surfaces as described in (Damerum *et al.*, 2021). Briefly, two leaf disks were cut from the same leaf which were affixed to a microscope slide with double-sided sellotape; one disk with the abaxial surface exposed and one disk with the adaxial surface exposed. This microscope slide was placed onto a pedestal and 10μL of H_2O was applied as a droplet on the leaf disk with a pipette. The side view of the slide was then imaged and the angle between the leaf surface and droplet was measured using the ImageJ software (see Figure 2 for example).

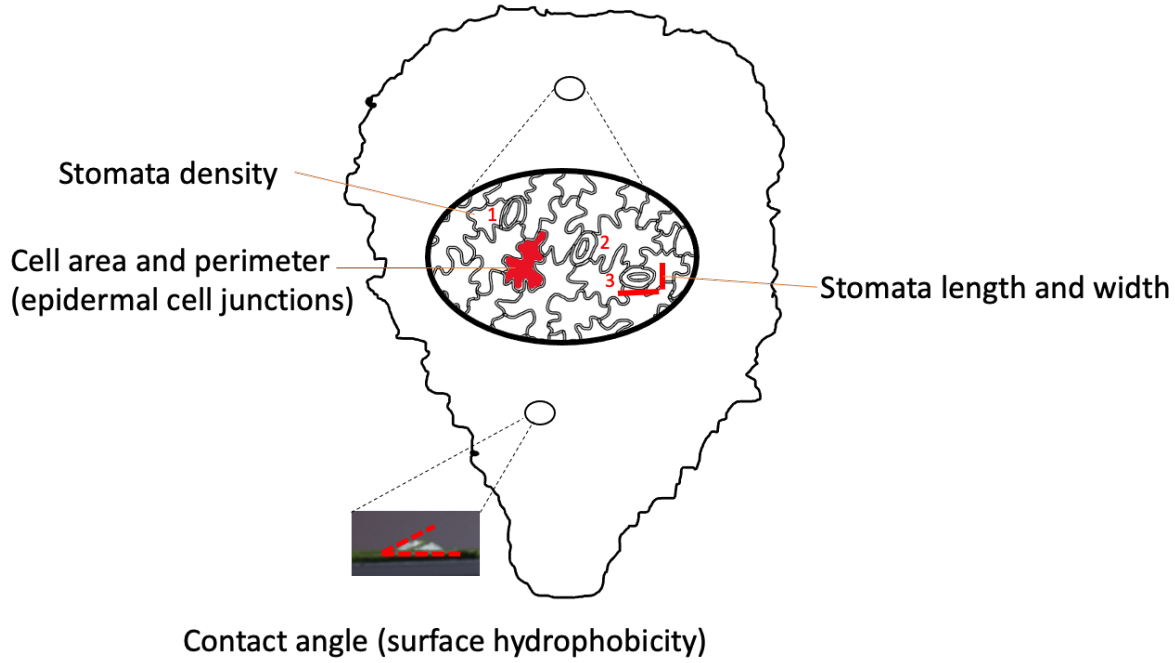


Figure 2 Leaf surface traits influencing microbe attachment and proliferation assessed.

Objective 3

Two field trials were conducted during the Salinas and Imperial valley late seasons and samples were collected for leaf surface characteristics and microbiome analyses, as detailed in Objective 1.

For data processing, all available phenotype data was tested for normality using the Shapiro-Wilk test and data transformation was applied if necessary, using R statistics (R Core Team, 2017). Genotype data for the diversity panel were collected using tunable Genotyping by Sequencing technology, as described in Kandel *et al.*, (2020), yielding 50,325 polymorphic sites. These sites were filtered to remove any with >10% missing data, >10% heterozygosity and a minor allele frequency of <0.05. The resultant 4,966 SNPs were utilized to determine the population structure using STRUCTURE (Pritchard *et al.*, 2000) using an admixture model, with a burn-in period of 100,000 and 100,000 Markov chain Monte Carlo repetitions. Optimal sub-population structure (K) was determined by the Evanno method, implemented in STRUCTURE HARVESTER (Earl & VonHoldt, 2012). A general linear model (GLM) was run using the software TASSEL (v5) (Bradbury *et al.*, 2007), correcting for population structure by incorporation of sub-population coefficients (Q) calculated in STRUCTURE. Significant marker-trait associations (SMTAs) were identified after Bonferonni P value correction and Manhattan plots were drawn in TASSEL.

RESULTS AND DISCUSSION

Objective 1

In the last year, leaf tissue of the commercial panel samples has been ground for both Salinas 2020 and El Centro 2021 trials. DNA has been isolated from samples collected in the Salinas 2020 trial and the next step is to isolate DNA from the El Centro 2021 samples and prepare libraries (for both Salinas and El Centro trials) for microbiome sequencing. Table 2 outlines a timeline for library preparation, sequencing and analyses, to be completed by the end of 2021.

Table 2 Gantt chart outlining the timeline for microbiome sequencing of the commercial panel for the Salinas 2020 and El Centro 2021 trials.

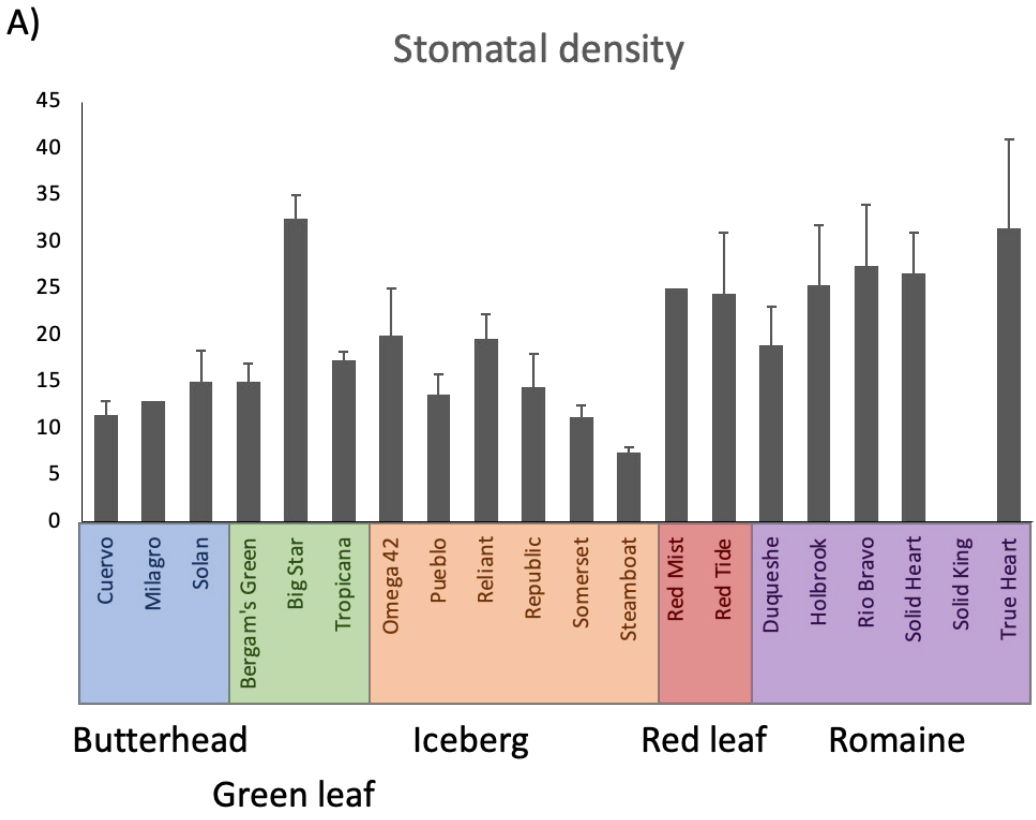
Task	July	August	September	October	November	December
Sample grinding						
DNA isolation						
Library preparation						
Sequencing (at facility)						
Data analyses						

Objective 2

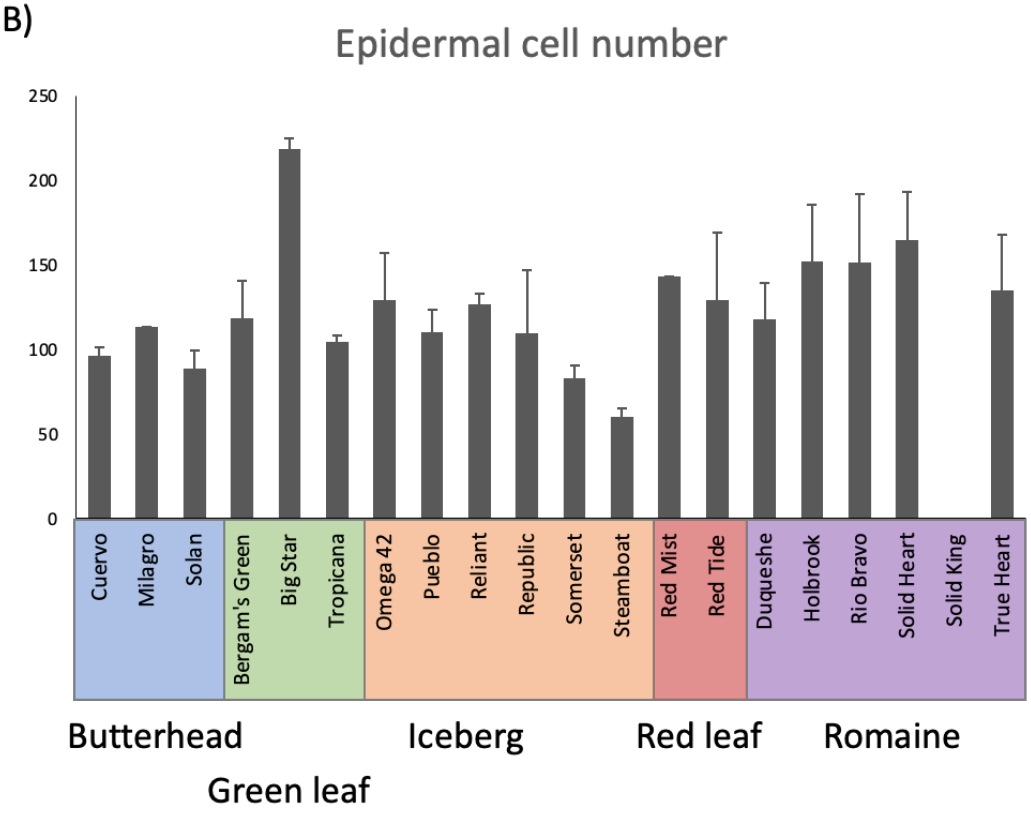
The investigation of leaf surface traits including stomatal density, epidermal cell number, area and perimeter for the Salinas 2020 trial identified differences amongst the cultivar panel (Figure 3). Stomatal density ranged from 7.5 in the iceberg type Steamboat to 32.5 in the green leaf cultivar Big Star per unit area. On average, stomatal density was higher in red leaf and romaine types compared to butterhead and iceberg. A similar pattern was observed for

epidermal cell number, which varied from 60 (Steamboat) to 218 (Big Star) and was higher on average for red leaf, romaine and green leaf types than butterhead and iceberg. Iceberg and butterhead types had higher average epidermal cell areas and perimeters (an indication of epidermal cell junctions), with romaine leaf types demonstrating smaller cell areas. **This indicates that romaine, green and red leaf types have leaves with a higher number of smaller sized cells on their surface compared to iceberg and butterhead types, which have a smaller number of cells with a bigger cell area.** This is important, given that bacteria such as the toxic *E. coli* O157:H7 has been shown to aggregate in epidermal cell junctions and evade washing (López-Gálvez *et al.*, 2010). **From this data, we might expect the washing of iceberg and butterhead types to be more effective at removing harmful bacteria from the leaf surface, due to the lower abundance of epidermal cell junctions per unit area.**

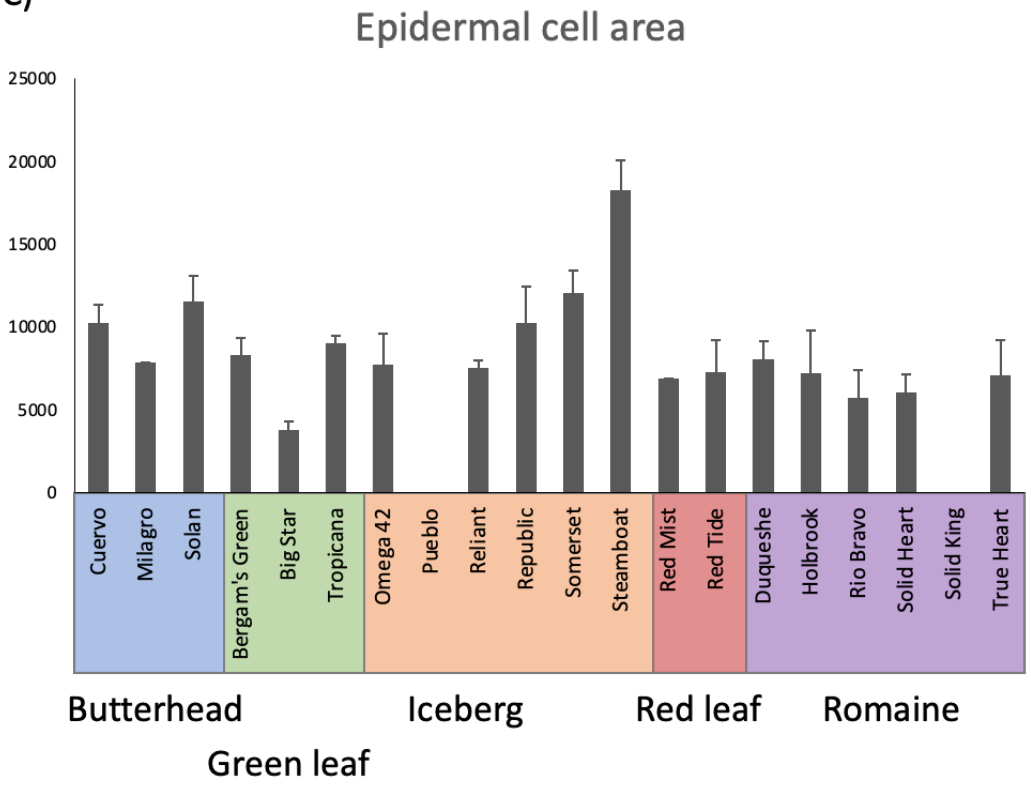
Ongoing work will continue to process phenotype data from the El Centro 2021 trial, including the imaging and analyses epidermal imprints and the image analyses of contact angle, to be completed by March 2022.



B)


Epidermal cell number
250
200
150
100
50
0
Cuervo
Milagro
Solan
Bergam's Green
Big Star
Tropicana
Omega 42
Pueblo
Reliant
Republic
Somerset
Steamboat
Red Mist
Red Tide
Duqueshe
Holbrook
Rio Bravo
Solid Heart
Solid King
True Heart
Butterhead
Green leaf
Iceberg
Red leaf
Romaine

C)


Epidermal cell area
25000
20000
15000
10000
5000
0
Cuervo
Milagro
Solan
Bergam's Green
Big Star
Tropicana
Omega 42
Pueblo
Reliant
Republic
Somerset
Steamboat
Red Mist
Red Tide
Duqueshe
Holbrook
Rio Bravo
Solid Heart
Solid King
True Heart
Butterhead
Green leaf
Iceberg
Red leaf
Romaine

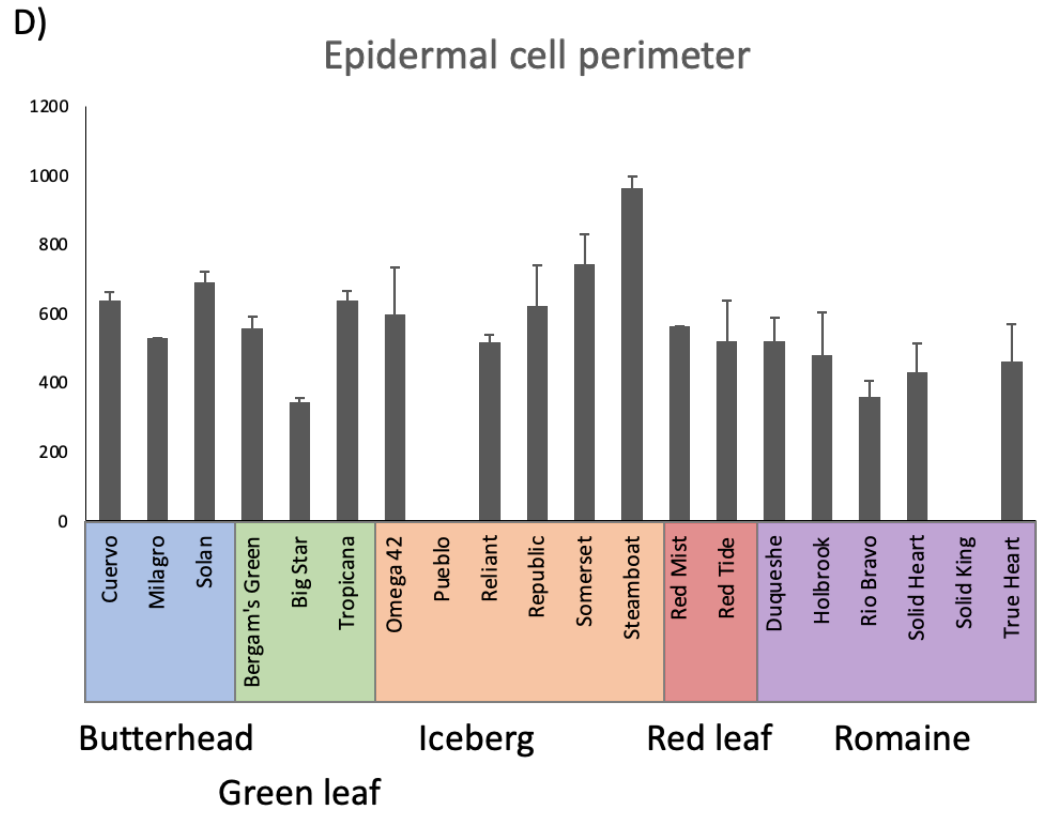


Figure 3 Leaf surface traits measured in the cultivar panel grown in Salinas, 2020.

Objective 3

In the last year, plant tissue of the diversity panel from both the Salinas 2020 and El Centro 2021 trials has been ground ready for sample preparation and DNA isolation. Ongoing work is isolating DNA from the El Centro 2021 samples, to be completed by December 2021, with sequencing libraries to be prepared and data to be analyzed by March 2022 (Table 3).

Table 3 Gantt chart outlining the timeline for microbiome sequencing of the diversity panel for the Salinas 2020 and El Centro 2021 trials.

Task	July	August	September	October	November	December	January	February	March
Sample grinding									
Sample preparation									
DNA isolation									
Library preparation									
Sequencing (at facility)									
Data analyses									

Leaf surface trait data has been collected for the Salinas 2020 trial, including stomatal density and epidermal cell number, area and perimeter. Due to high disease pressure in the Salinas a significant number of the diversity panel lines were dropped due to insufficient replication. Variation in stomatal density and epidermal cell number, area and perimeter was observed for the diversity panel (Figure 4). GWAS using a general linear model incorporating population structure revealed no significant marker-trait associations after FDR P value correction (Table 4, Figure 5), likely due to insufficient sample sizes. A total of 27 significant marker-trait associations were identified before P value correction (Table 4). Ongoing work will continue to process phenotype data for the El Centro 2021 trial, for which minimal disease pressure was observed.

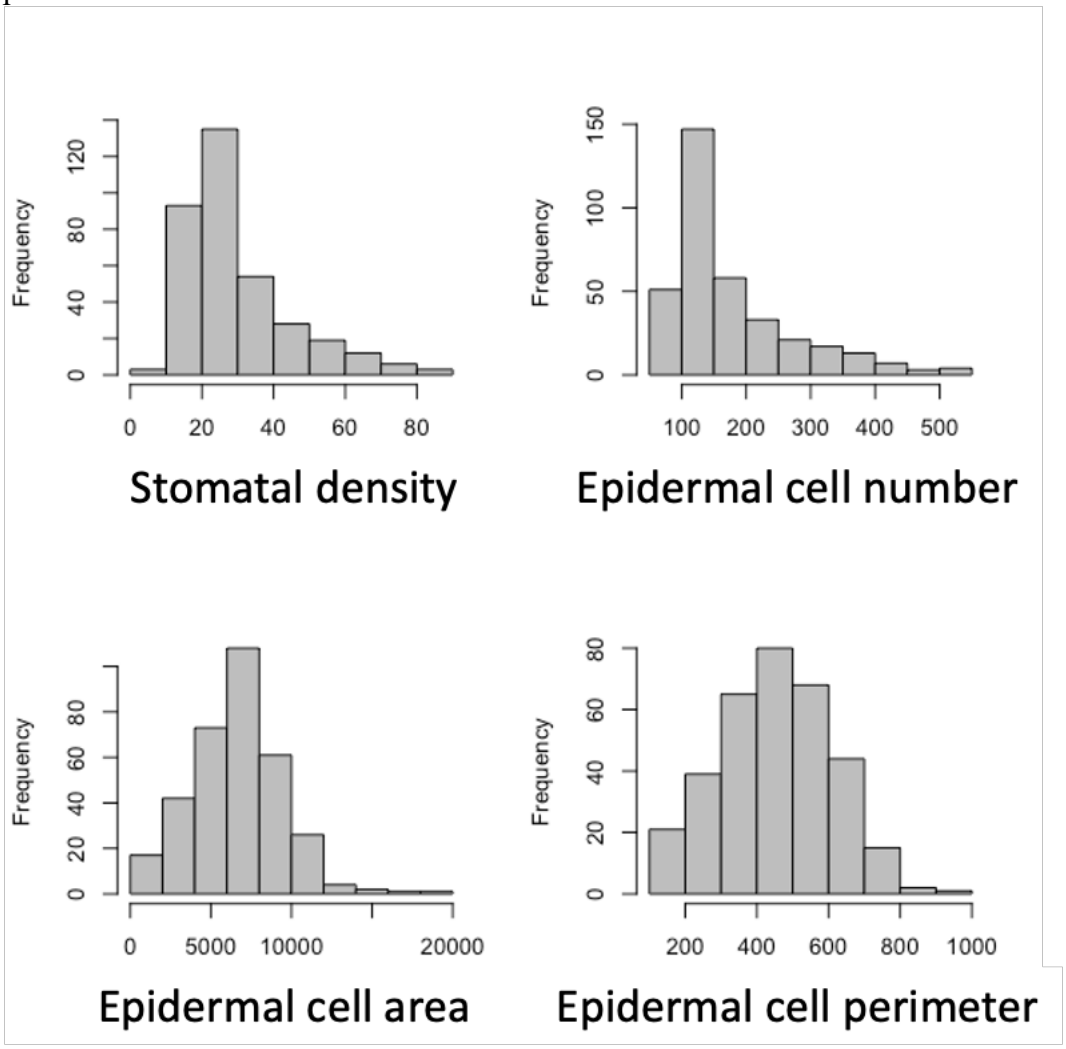


Figure 4 Frequency distributions of leaf surface traits measured in the Salinas 2020 trial for the diversity panel.

Table 4 Salinas 2020 marker trait associations

Trait	Chr	Pos	P value	P value (corrected)	Marker R^2
SD	LSAT_1_V8_LG_1	72340981	7.21E-04	0.848	0.04443
SD	LSAT_1_V8_LG_1	79539590	5.98E-04	0.82	0.03837
ECP	LSAT_1_V8_LG_1	112881896	7.34E-04	0.616	0.04829
ECA	LSAT_1_V8_LG_1	112881896	9.90E-04	0.72	0.04685
ECN	LSAT_1_V8_LG_2	56958584	2.52E-04	0.662	0.05022
ECN	LSAT_1_V8_LG_2	56958602	1.91E-04	0.61	0.05186
ECN	LSAT_1_V8_LG_2	56958617	2.88E-04	0.678	0.04957
ECN	LSAT_1_V8_LG_2	56958622	3.52E-05	0.357	0.06178
ECN	LSAT_1_V8_LG_2	56958696	1.82E-04	0.604	0.05282
ECN	LSAT_1_V8_LG_2	56958697	1.76E-05	0.269	0.06643
SD	LSAT_1_V8_LG_2	56958697	2.76E-04	0.676	0.04935
ECP	LSAT_1_V8_LG_2	56958697	6.02E-04	0.54	0.04765
ECN	LSAT_1_V8_LG_2	56958701	1.18E-04	0.512	0.05521
ECN	LSAT_1_V8_LG_2	56958724	1.18E-04	0.512	0.05521
ECA	LSAT_1_V8_LG_2	204857020	3.48E-04	0.379	0.05159
ECA	LSAT_1_V8_LG_2	204857025	3.73E-04	0.399	0.05115
ECA	LSAT_1_V8_LG_2	204857054	3.53E-04	0.384	0.0538
ECA	LSAT_1_V8_LG_4	218441223	7.68E-04	0.631	0.0486
ECA	LSAT_1_V8_LG_4	218441273	7.82E-04	0.636	0.04847
ECN	LSAT_1_V8_LG_4	307784120	8.90E-04	0.885	0.04597
ECA	LSAT_1_V8_LG_4	339113915	5.01E-04	0.48	0.05181
ECA	LSAT_1_V8_LG_5	282497058	6.53E-04	0.568	0.04968
ECA	LSAT_1_V8_LG_7	47111748	8.37E-04	0.659	0.0474
ECN	LSAT_1_V8_LG_7	132891719	2.14E-04	0.627	0.05144
ECA	LSAT_1_V8_LG_8	214893081	2.70E-04	0.309	0.05428
ECA	LSAT_1_V8_LG_8	220523518	5.11E-04	0.487	0.05178
SD	LSAT_1_V8_LG_9	74587599	4.26E-04	0.76	0.04632

SD; stomatal density, ECA; epidermal cell area, ECN; epidermal cell number, ECP; epidermal cell perimeter

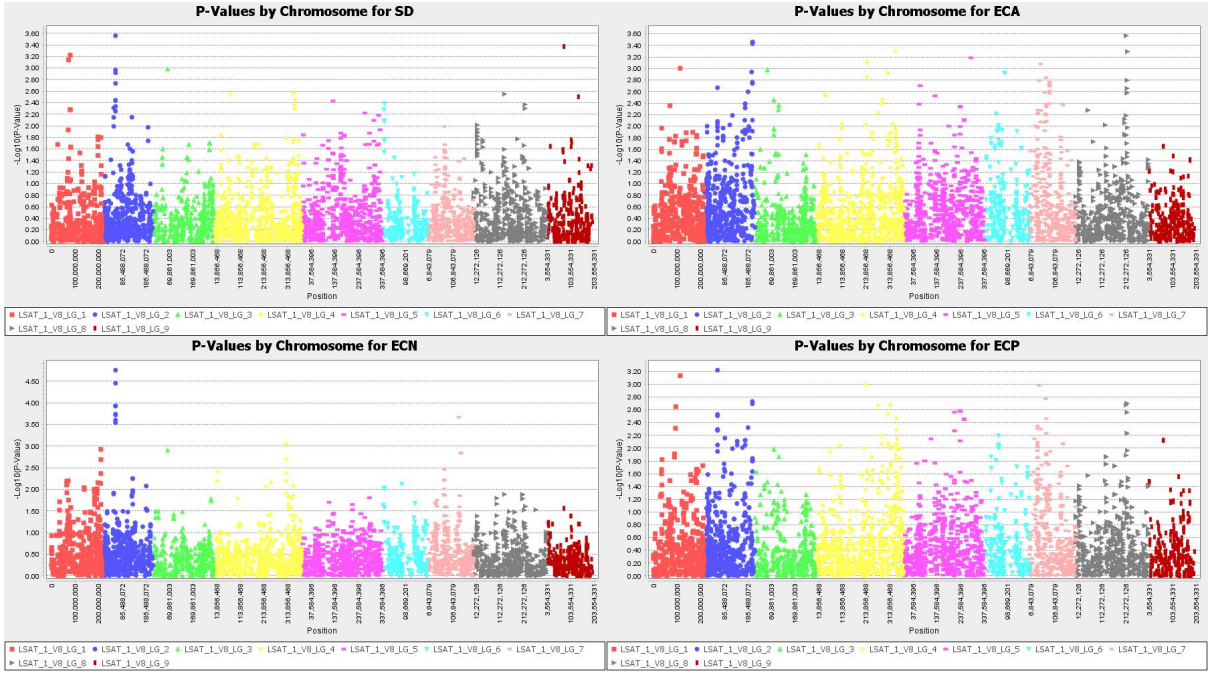


Figure 5 Manhattan plots for SD; stomatal density, ECA; epidermal cell area, ECN; epidermal cell number, ECP; epidermal cell perimeter collected from the diversity panel for the Salinas 2020 trial.

Conclusions

In the challenging first year of this project, during the 2020/21 season with a global pandemic in full swing, we were able to conduct two successful field trials in the Salinas and Imperial Valleys. For the first time, we collected data to assess leaf surface traits and the microbiome of the commercial (23 cultivars) and diversity (493 diverse accessions) panels, totaling over ~3000 lettuce plants. Ongoing work will continue to prepare sequencing libraries for microbiome sequencing of the commercial panel, before isolating DNA and preparing libraries. Microbiome sequencing of both the commercial and diversity panels is expected to be completed by March 2022. Once collected, data from the diversity panel will be utilized in GWAS to identify lettuce genomic regions influencing microbe abundance and diversity.

Literature

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**: 2633–2635.

Custer GF, Dibner RR. 2020. Modified Methods for Loading of High-Throughput DNA Extraction Plates Reduce Potential for Contamination. *JoVE (Journal of Visualized Experiments)* **2020**: e61405.

Damerum A, Arnold E, Bernard V, Smith H, Taylor G. 2021. Good and bad lettuce leaf microbes? Unravelling the genetic architecture of the microbiome to inform plant breeding for enhanced food safety and reduced food waste. *bioRxiv*: 2021.08.06.455490.

Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

López-Gálvez F, Gil MI, Truchado P, Selma M V., Allende A. 2010. Cross-contamination

of fresh-cut lettuce after a short-term exposure during pre-washing cannot be controlled after subsequent washing with chlorine dioxide or sodium hypochlorite. *Food Microbiology* **27**: 199–204.
Pritchard JK, Stephens M, Rosenberg N a, Donnelly P. **2000**. Association mapping in structured populations. *American journal of human genetics* **67**: 170–81.
R Core Team. **2017**. R: A Language and Environment for Statistical Computing. **7**: 275–286.
Schneider C a, Rasband WS, Eliceiri KW. **2012**. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.
Sthapit Kandel J, Peng H, Hayes RJ, Mou B, Simko I. **2020**. Genome-wide association mapping reveals loci for shelf life and developmental rate of lettuce. *Theoretical and Applied Genetics*.