Center for Produce Safety STEC Seasonality Project:
Romaine Lettuce Seasonal Risk in the California Central Coast Region
Prepared by Trevor V. Suslow, Professor of Cooperative Extension and Research, Emeritus, University of California, Davis | 2022

CPS STEC Issue Brief 2: Leafy Greens Production Patterns and Practices

Executive Statement: On April 6, 2021, the FDA communicated its position—to the leafy greens industry of California, the broader industry, and the public-at-large—that Shiga toxin-producing *E. coli* (STEC) in general and a reoccurring strain, in a reoccurring region, and linked by investigative sampling and detection to reoccurring issues with adjacent and nearby land use was a **known and reasonably foreseeable hazard** (FDA 2021a–c). At its most basic level, this determination by the FDA is not contested by the majority in industry, among food safety professionals, or the research community. At the same time, the FDA (FDA 2021a) released its updated STEC Leafy Greens Action Plan, which included a priority to advance root cause analysis/root cause investigation (RCA/RCI). The CPS STEC Seasonality Project was already in progress and was designed to address this aspect of the STEC Action Plan by the following science-informed and fact- and data-supported RCA steps:

STEC Seasonality Project Five-Step Process

1. Assemble
2. Aggregate
3. Align
4. Analyze
5. Act

The planned process and actions pipeline had two major prongs, which addressed RCA/RCI as the foundation for immediate to near-term industry implemented solutions and the best vehicle to develop a prioritized research roadmap for the Center for Produce Safety, to be shared with other research funding entities. Several disciplinary-defined clusters for data and fact aggregation were to provide an initial foundation for technical and practical cross-talk between these clusters. Broad and diverse stakeholder groups were contacted to analyze or react to the aligned facts, which included i) outbreak traceback and trace forward data; ii) genomic characterization of the regionally associated STEC isolates; and iii) specific data related to seasonality, production locations, practices at implicated ranches, and hypothesis generation around temporal root cause factors. It was quickly identified that overgeneralization and misinformation about production and practices was an ongoing barrier to both effective RCA and RCI, cross-STEC Project cluster group dialogue, and strategic identification of fundamental and applied research priorities.

Until outbreak, technical, and supply-chain information can be discussed in an integrated manner, across the five RCA/RCI steps mentioned above, and free from the understandable but solution-defeating thought silos, an architecture for regionally developed, long-term solutions will not be achieved.

Though not the focus of this Issue Brief, a recent study released by the UC Davis Department of Agricultural and Resource Economics details the drivers and factors involved in economic losses during the 2018 *E. coli* O157:H7 outbreak that was attributed to one sub-clade of the reoccurring strain and one production location on the southern Central Coast of CA (Kiesel et al. 2021). The conclusions of this study were that, due to the often-complex contractual agreements, growers were largely insulated from loss but shippers and especially processors realized the greatest economic impacts. It is not surprising, therefore, that processors have been most engaged in the CPS STEC project and independent efforts at RCA/RCI related to seasonality and discovery of linkages leading to solutions.

Purpose: This Issue Brief provides an overview of the general and regional production patterns and practices, and the seasonal influences and variability in cropping locations and management. The focus is primarily on romaine lettuce, consistent with the CPS STEC project focus. This Issue Brief is intended as a resource, and in combination with one focused on seasonality of regional cattle management (Issue Brief 3) and one on whole genome sequencing (Issue Brief 4), provides background for future discussions and strategic planning emerging from Issue Brief 1: Hypothesis Risk Matrix.

Background: As with any brief overview, the narrative is principally limited to general and descriptive and comparative industry profiles and trends for leafy greens in California and the CA Central Coast region. From its inception, CPS recognized that having this practical and regionally focused commercial profile was an essential element of shared knowledge to foster informed dialogue and fact alignment around outbreak root cause hypothesis generation. Equally, this baseline familiarity with production, practices, and seasonality along the CA Central Coast region, in relation to other ag-industries and climatic conditions, is important to the intelligent design of research proposals from the research community. The quality of review and selection of priority research for awards by CPS and other funding entities also benefits from this industry profile resource. Necessary granular details of commercial production and harvest operations are obviously not available for public disclosure in this Issue Brief. Other details, such as growing locations, are prone to highly dynamic changes and nuances as firms respond to the recurring concerns. It is important to bear in mind that past understandings of risk potential and risk exposure may no longer align with location and seasonality factors.

A few basics of crop traits and our understanding of the microbial communities associated with romaine lettuce and other leafy greens are noted below. These are very relevant to systematic approaches to RCA/RCI by the industry and to the review of research proposals, ideally designed and selected for funding to advance solutions to STEC contamination and outbreaks.

> <u>Varieties</u> – There are multiple varieties of commercial romaine (aka Cos) lettuce, full head and baby leaf, which are selected to best fit a planting and harvesting schedule, as production districts vary across seasonal cycles (See Figures 3 and 4). Pest and disease resistance and heat tolerance are among the key traits selected for to deliver generally consistent visual, shelf-keeping, and sensory quality year-round. Some varieties are optimized for specific value-added or processing yields

and customer or consumer traits. These factors are taken into consideration in Issue Brief 1: Hypothesis Risk Matrix.

<u>Impact of variety on *E. coli* and the associated microbial community</u> – A number of peer-reviewed studies have begun to characterize post-contamination behaviors of *E. coli* O157:H7 on romaine lettuce (or recent studies relevant to knowing the state-of-the-science) and the nature of the associated microbiome. A selection of some foundational studies (some funded by CPS) and more recent topic-aligned research outcomes are provided in Resources (See Burch et al. 2019, Kim et al. 2019, Leonard et al. 2021, Rastogi et al. 2012).

- In general, variety and leaf age (developmental sequence in a head) appear to influence *E. coli* O157:H7 colonization in conjunction with other climatic and weather factors.
- When grown in the same field location, different varieties and different leafy greens types are reported to select for a statistically different associated microbiome and metagenomic profile.
- Nitrogen fertilization, in particular, has been reported to play a significant role in selecting for a reproducible phyllosphere (leaf surface) microbiome and functional metagenomics profile (what the microbial communities are doing), along with weather.
- Long-term, current studies are likely to identify some opportunities to direct the development of a protective microbiome to reduce the persistence of STEC on leafy greens and other crops.
- Romaine variety growth traits are influenced by seasonal weather and likely further influenced by microclimate variability. During on-farm site visits associated with this CPS STEC project, the openness of romaine heads at key stages of production, and especially as a ranch-block-lot approaches scheduled harvest, was noted as a potential risk/root cause factor to consider. The same variety may have a very different degree of openness during the same seasonal production schedule rotation if grown in the cooler northern region versus the much hotter southern valley region. Individual lots may have a lesser (left image below) or greater number (right image below) of open heads in a population. The specific relationship of open-head romaine to risk has not been established, but unpublished field research indicates a potentially elevated risk profile if overhead irrigation is being used. Markets and customers typically influence whether more open-heads are harvested by a crew. In general, they are not differentiated.

Production and Consumption Profile

Romaine shipments are highly seasonal, and outbreaks have acute, chronic, and oftentimes subtle impacts on industry production volume x location decisions and consumer purchase decisions (Astill 2019). The following figures provide a profile of annual lettuce production and consumption.

Figure 1. Decades spanning per capita consumption trends for loose leaf lettuces in the U.S. (includes and is dominated by romaine lettuce)

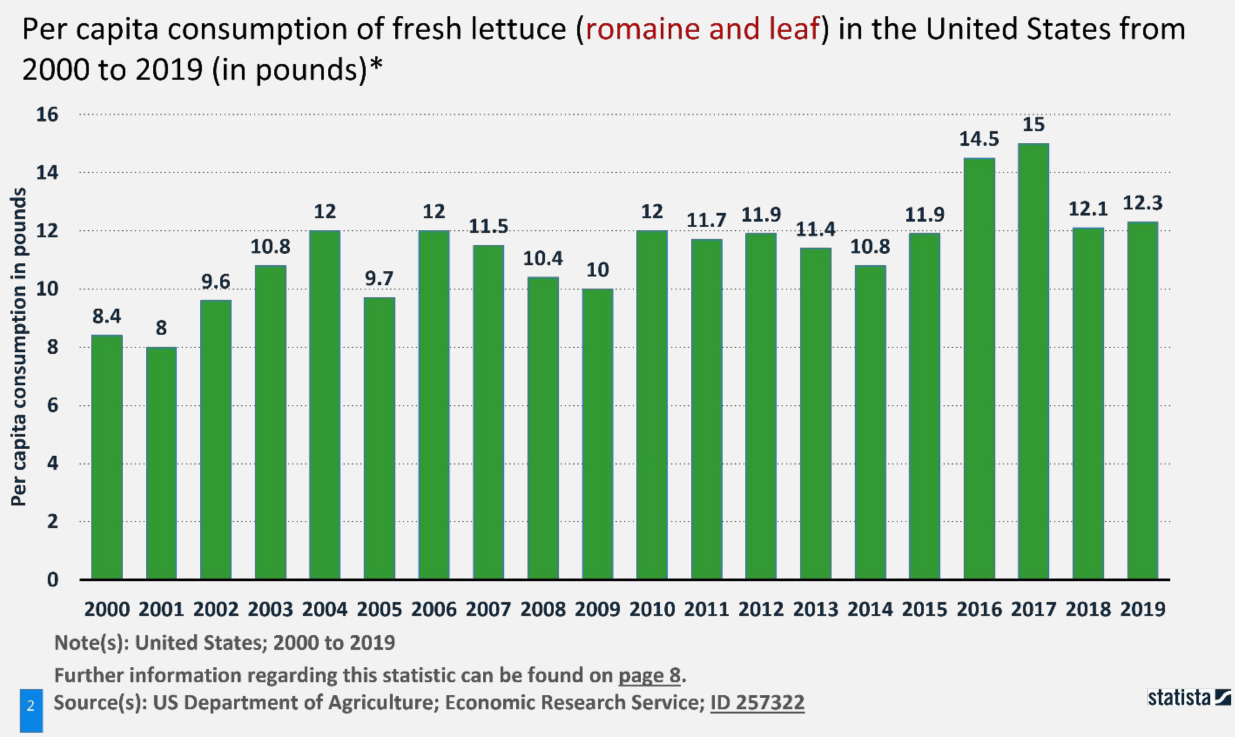
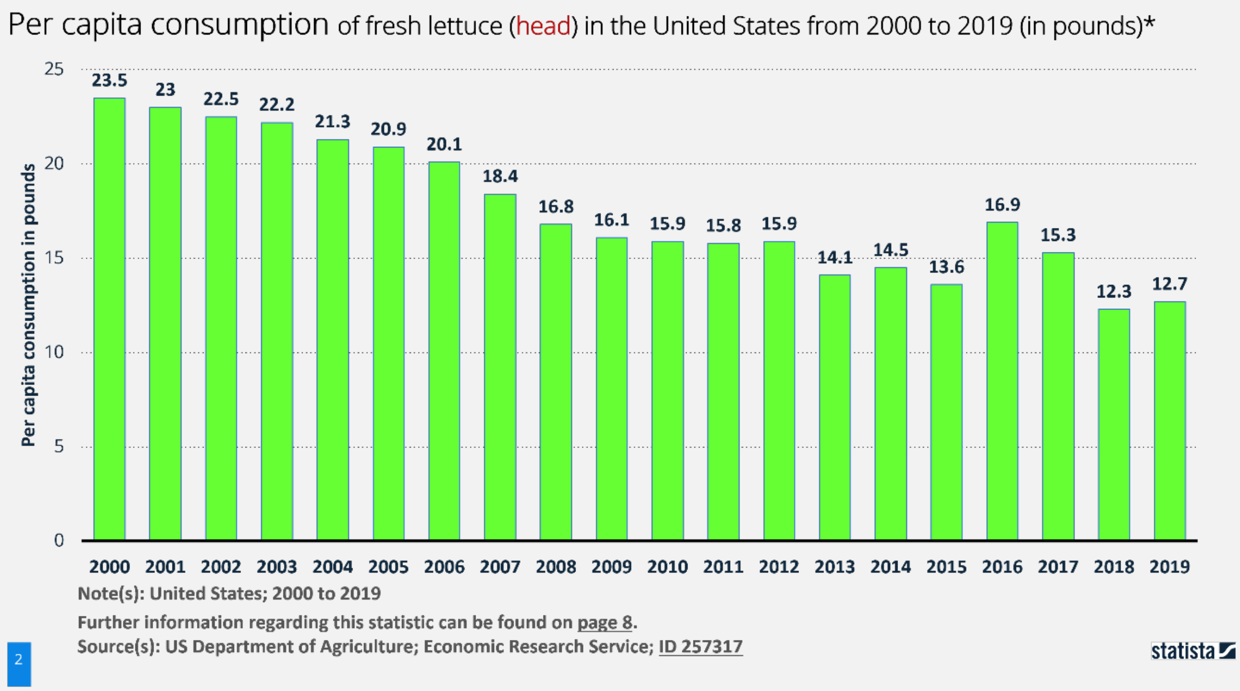


Figure 2. Decades spanning per capita consumption trends for head lettuce (iceberg-type) in the U.S. Much of the loss volume over the years has been attributed to "cannibalization" of share of stomach by romaine lettuce, followed by baby spinach, and other tender greens.



Seasonality Profile

Figure 3. Generalized seasonal timeline cycles for California coastal region production. Localized microsites and speculative plantings, generally early season and late season, may be outlier sources of specific leafy green types (See Figures 7 and 8). The crop cycle timeline below is for iceberg lettuce; leaf lettuce and romaine harvest will continue through November. Depending on the specific sub-region, spring mix (tender greens) and baby spinach are planted from January through October and harvested from February through November. Each end of the crop cycle (earliest to latest) is typically in the southernmost sub-region of the Salinas Valley; warmer earlier and warmer later, with much higher temperatures late in the regional mid-season effectively dropping heading type lettuces from production schedules during this period. Two types of groundwork occur during the season: mid-season groundwork between crops, and year-end groundwork in preparation for the crops planted in the next year.

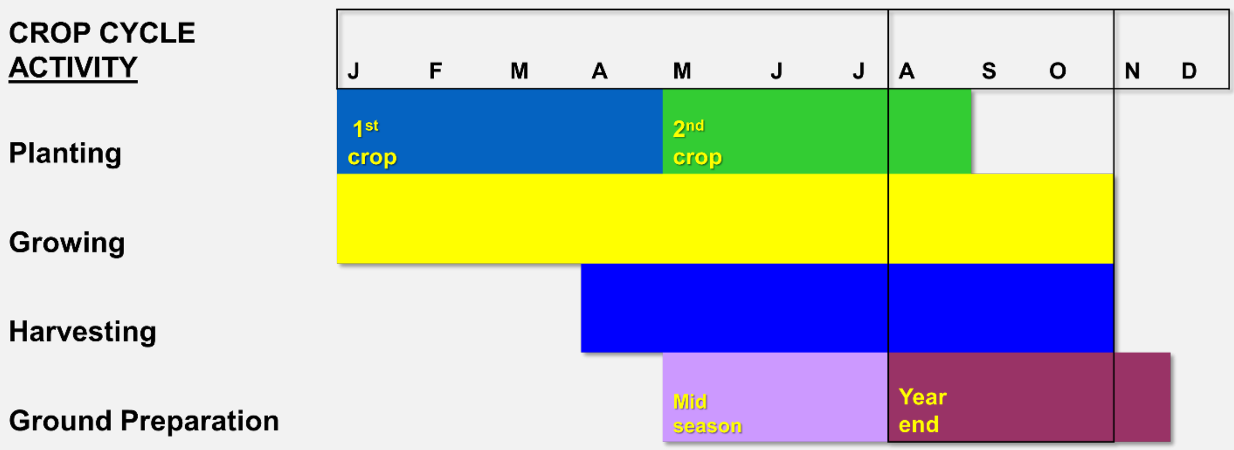


Figure 4. Generalized example of an individual ranch block crop rotation management program. The key variables to appreciate in relation to the romaine lettuce hypothesis generation and risk/research ranking matrix (CPS STEC Issue Brief 1) include:

- Multiple lots are in production simultaneously
- Different lots can be in production with different crops
- Ground preparation, pre-plant irrigation, planting, growing, side-dressing with organic amendments (such as thermally treated manure pellets in organic systems), and harvesting can occur simultaneously on the same ranch
- Variable factors determine whether a single lot or sub-lot will have two or three crops of a leafy green or even of romaine in a single seasonal cycle

Lot 1 Plant – Lettuce in Jan. **Harvest - April** Work Ground - May Plant – Lettuce in May **Harvest - Aug.** Work Ground. – Aug./Sept.	**Lot 2** **Strawberries** – planted previous Oct./Nov. **Harvest – April – Oct.**. Work Ground. – Nov.
Lot 3 Plant – **Broccoli** in Mar. **Harvest - June** Work Ground - July Plant – Lettuce in July **Harvest – Sept.** Work Ground. - Oct.	**Lot 4** Plant – Lettuce in April **Harvest - July** Work Ground - July Plant – Lettuce in Aug. **Harvest - Oct.** Work Ground. - Nov.

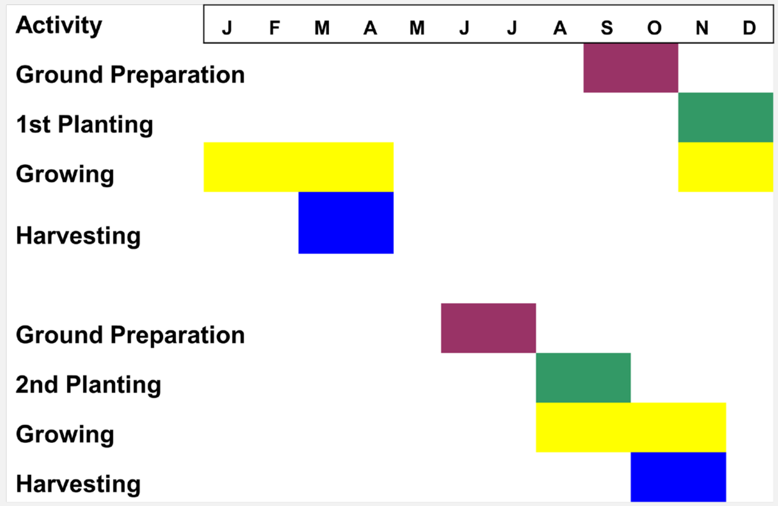


Figure 5. Generalized seasonal timeline cycles for the San Joaquin Valley (Huron) "transition gap-filling" production. The North to South cycle (upper timeline) and South to North cycle (lower timeline) reflect the short intervals of production volume sourcing between the primary Salinas region and the desert production seasonal cycles in the CA Imperial County and the Yuma, AZ regions.

The volumes of leaf greens or types of leafy greens from this gap-region have fluctuated significantly over the past decade but are generally greater than the volumes two years ago. Many factors determine land use for romaine and other major leafy greens in this area: drought conditions, water allocations, water quality, and conversion of arable land to tree nut production are among the key factors. Crop losses (largely due to pest issues) that are experienced by schedules planned for a shift to later plantings in the desert region have caused some shifts back to this region during the South to North transition. However, this remains a highly dynamic situation.

Figure 6a. Comparative cropping cycles per acre across the three major lettuce and leafy greens production districts in the U.S.

	Salinas/Santa Maria	Huron	Imperial/Yuma
No. of crops per land acre Per year	2 - 3	2	1
Same land in consecutive years	Yes	No	No
Direct Seed Crop Growth Periods	Early -120 days Mid - 60 days Late - 80 days	140 - 120 days 60 - 75 days	60 days 120 days 100 days

Figure 6b. Generalized and comparative climatic profiles across the major lettuce and leafy greens production district in the U.S. For temperature profiles during growing to harvest, the upper line is the North to South cycle and "gap" season (Huron) and the lower line is the South to North cycle.

	Salinas/Santa Maria	Huron	Imperial/Yuma
Temps during growing → harvest	Cool to Warm Warm to Cool	Cold to Warm Hot to Warm	Hot to Cold Cold to Warm
Rainfall amount and timing	12 – 15 inches Nov. – Mar.	8 – 12 inches Nov. – Mar.	2 – 3 inches July – Sept.
Possible Flooding Events	Jan. – Mar. adjacent to creeks/rivers	Excess rainfall can collect in individual fields	Excess rainfall can collect in individual fields

Figures 7-9. Production volume trends for California and the Salinas Central Coast Region. These graphs were created for this Issue Brief using tabular data available from the California Leafy Greens Research Board. (Note: In Figures 8 and 9, T&T refers to Topped & Tailed, an in-field processing term for removing all outer leaves and trimming both ends of the compact central head for subsequent packaged salad processing.)

Figure 7.

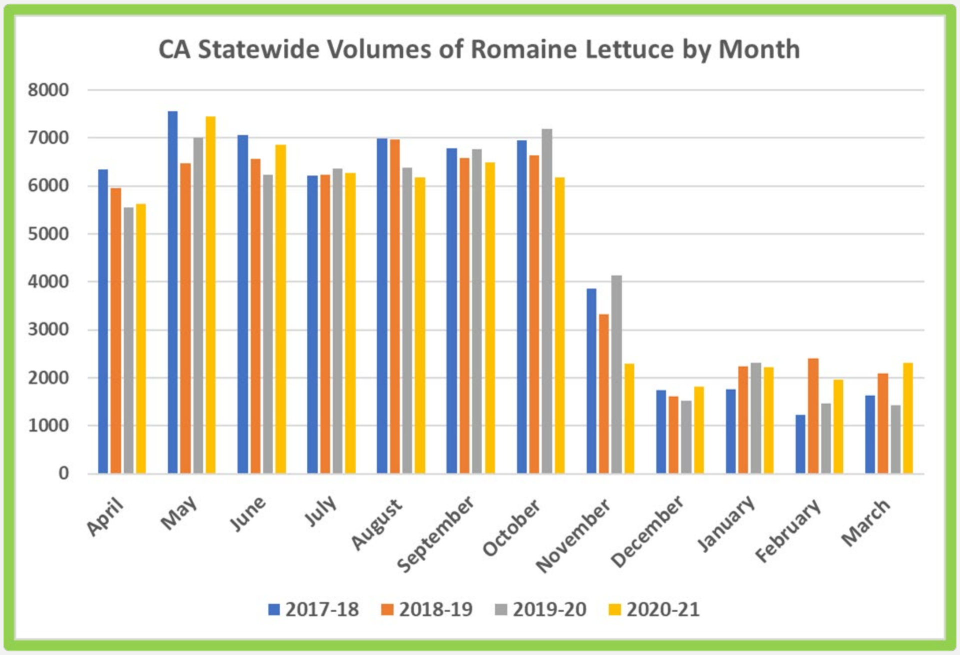


Figure 8.

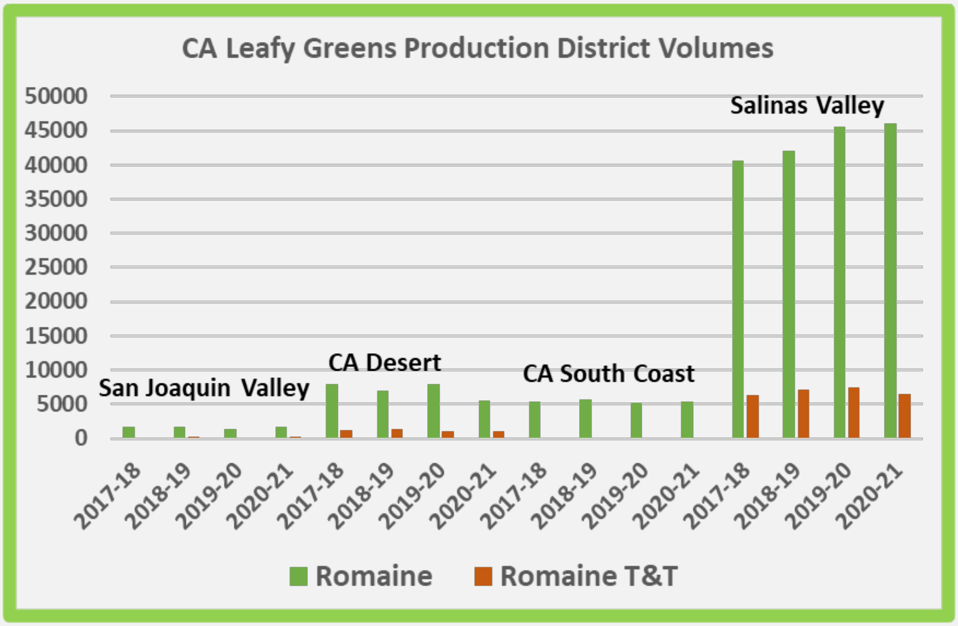
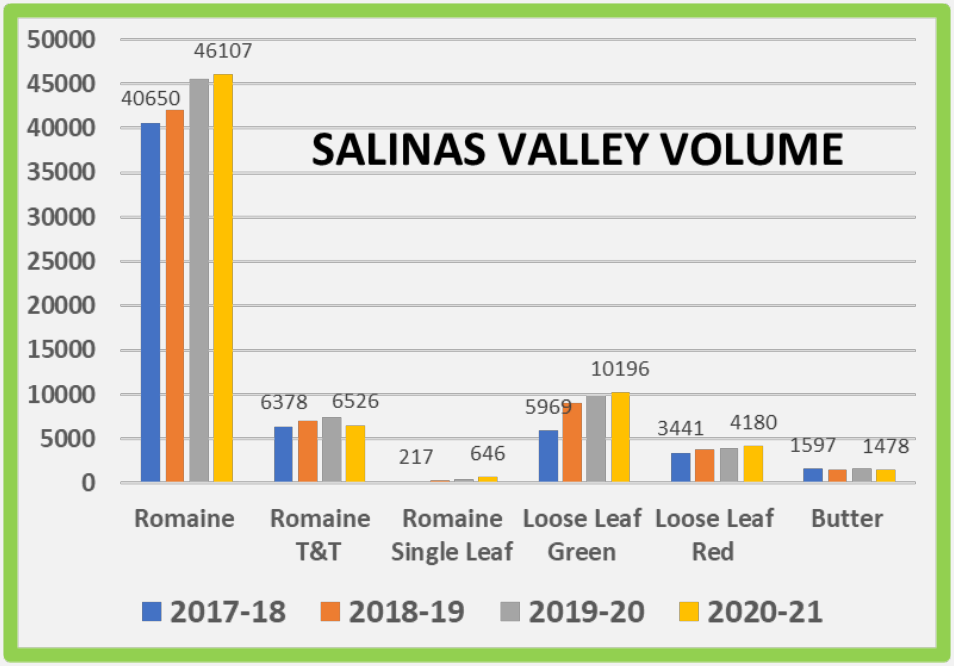


Figure 9.



Applying Knowledge from WGS Information to STEC Seasonality Root Cause Investigation

Over the timeline of the CPS STEC Seasonality Project the multifaceted conversations and some data and information exchanges identified several avenues and open questions to pursue. Pragmatically, the complexity and sensitivity surrounding much of these commercial, cross-agency, and technical disciplinary touchpoints were unlikely to curb the long-standing challenges to transparency. Regardless, in part due to industry leadership focus on resolving the reoccurring pattern of outbreaks, recalls, and individual lot pre-market destruction, several fact and knowledge discovery initiatives were taken. CPS participated in these efforts, providing technical input, facilitation, and coordination.

One area of CPS facilitation and partnership incubation continues to be in understanding and applying indicators and clues from regional and other whole genome sequencing (WGS) data, provided by the CDC, FDA, and available in public databases, to research prioritization (See Issue Brief 4: Genomics and SNPs). While beyond the scope of this Issue Brief, the tasks involved in the STEC Five-Step Process identified several areas essential to illuminating the relationship of genomic relatedness of the *E. coli* O157:H7 clade-clusters when addressing root cause assessment and investigation. The following figures provide some details of the current best representation of CDC and FDA for genome diversification, contrasting the 2018 to 2020 romaine outbreaks. However, currently, the granularity of data is inadequate to reveal necessary production, practices, and specific seasonality linkages to answer the STEC root cause open questions and refine an integrated research architecture and agenda.

The key areas of agreement that can be interpreted from the combined WGS information are as follows:

There are two primary subclusters of the REPEXH02 strain (Figure 10), which include isolates from product and those common-sense linked to "nearby" regional cattle. This group (Figure 11) is much less diverse than a more distantly related clade REPEXH01 (Figure 12).

The terminology surrounding "the same strain" … "highly related" (within agency perspectives) and "distinct subclusters" (within industry perspectives) is a considerable factor in dialogue between public health and industry, although both communities are seeking the same goals and moving to mutual understandings and common ground. The main issue is that genomic lumping of this regional diversity results in a very broad discussion of outbreak vehicles associated with the "same source" (cattle-region-producer), while the solutions-directed discussions are directed at "same seasonal source/factors" (region-production-practices). Fortunately, in ongoing interactions, the two camps are not that far apart. The next steps depend greatly on identifying the current state-of-the-effort to mine the extensive WGS database and deep comparative computational sequence discovery to understand the basis for persistence of the reoccurring *E. coli* O157:H7 subtype. Resources are provided that indicate how these questions have been approached in other systems. Most of these are highly technical journal publications, but several key take-aways are worth noting:

- Persistence of a specific subtype of *E. coli* O157:H7 in a closed herd of cattle has a characterized basis of genetic determinants
- These determinants may be borne on the accessory genome (acquired and transferred among cells)
- These accessory genomic elements have been shown to modify a variety of stress adaptations and determinants of environmental competitiveness
- Insights to rates of diversification (accumulation of SNP and wg/cg MLST differences; See Issue Brief 4) suggested comparative source-tracking research for growth conducive hosts and environments (more growth = greater rates of diversification), novel host-adaptation (greater environmental persistence and non-ruminant host prevalence), and niches for stable persistence but not growth conducive (limited growth = very slow SNP and MLST differentiation)
- Estimates for *E. coli* O157:H7 SNP accumulation (comparing isolates of human clinical and cattle origin) were 2.6 SNPs per genome per year (it's complicated but taking a simplified approach, >1 in 100,000,000 substitutions per year)

These values for DNA base-pair (ATGC) substitutions per year, resulting in SNPs, were similar to previous reports for *E. coli*. Several reports identified these accessory genomic regions as more represented in persistent isolates ($P \leq 0.05$) than in the sporadic and transient subtypes. Recent results suggest that STEC within a specific cattle population are highly clonal and may, therefore, reflect a level of relatedness appearing to be single-source outbreak-associated strains. The emerging science suggests caution about inferring genomic relatedness and, therefore, hyperlocal-region attribution of clinical strains. An identified need taken from this information is that long-term pathogen surveillance studies may be justifiable as one approach to better guide outbreak investigations. In addition, this branch of WGS research suggests that these accessory genome attributes may be potential targets for mitigation strategies directed at the foundational source: cattle.

Figure 10. The "Strain" (REPEXH02) is characterized by cgMLST into two major groupings: one that is more associated with the southern Central Coast (blue subcluster) and one with secondary branching associated with the main region of the Central Coast (red subclusters) in California. (See Resources: Marshall et al. 2020)

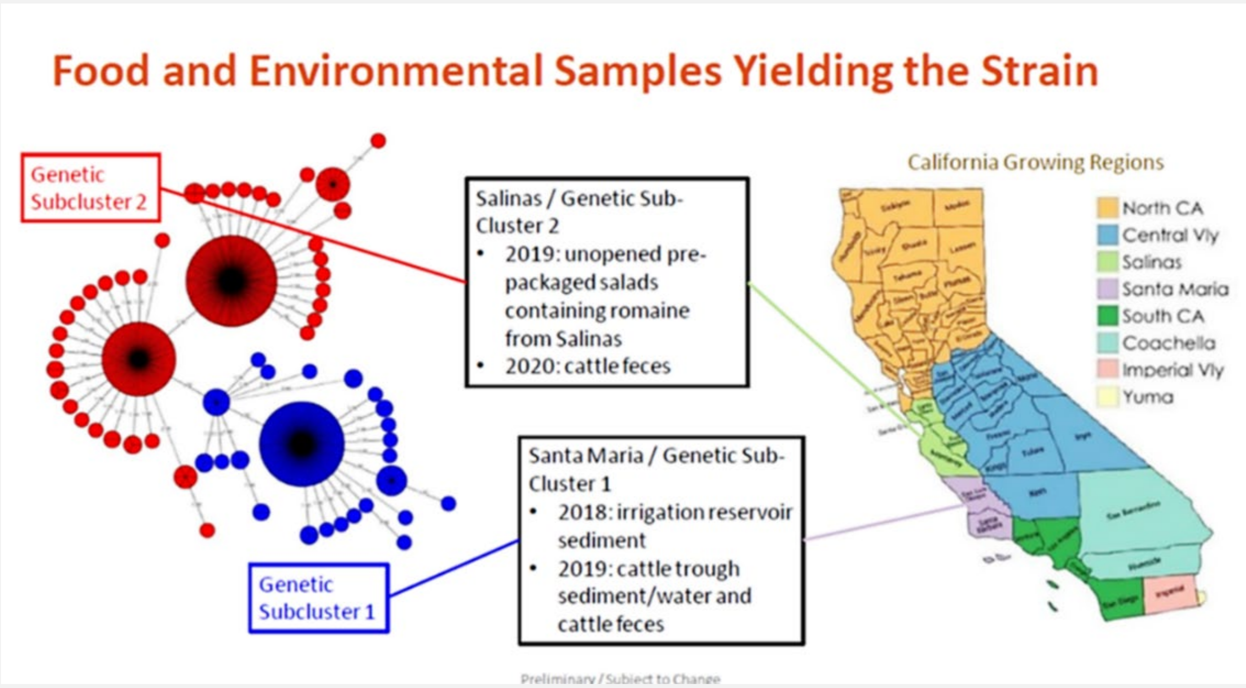


Figure 11. Detailed data mining and analysis of the REPEXH02 genomic profile beyond the phylogenetic relationships and clues to diversification from a central source may provide critical clues to future solutions. The shorter the branches on the clustering relationship tree, the more closely related (by a large database of alleles) the isolates are to the central reference isolate and those strains within each grouping.

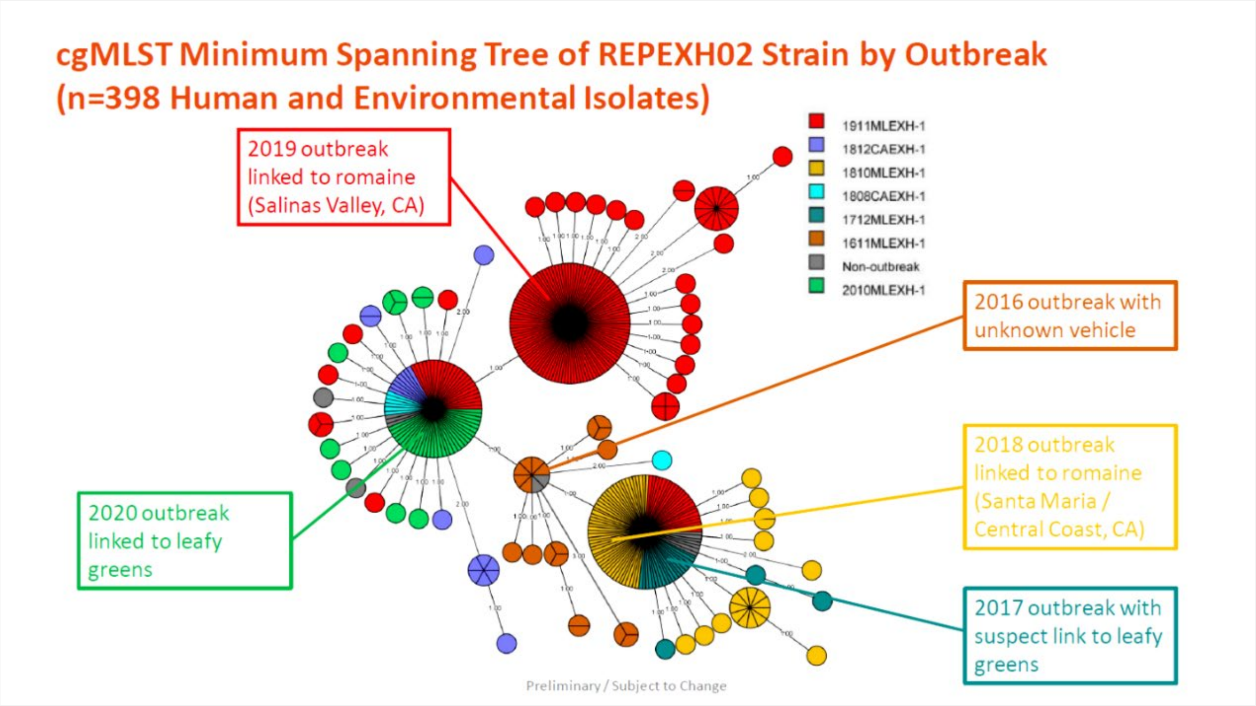
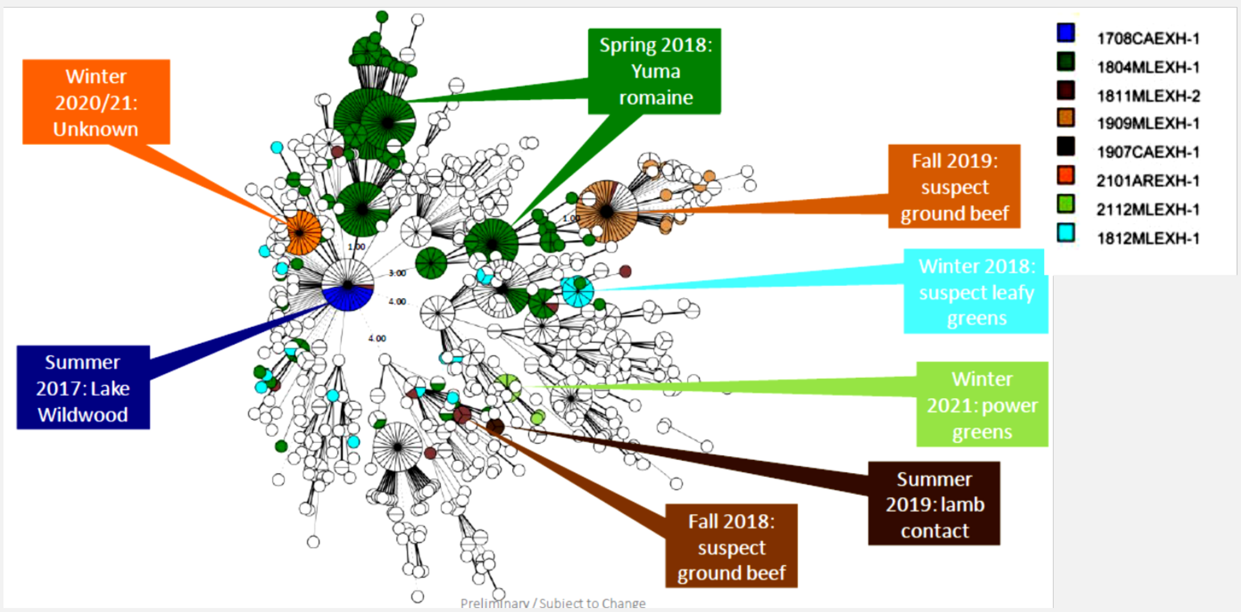


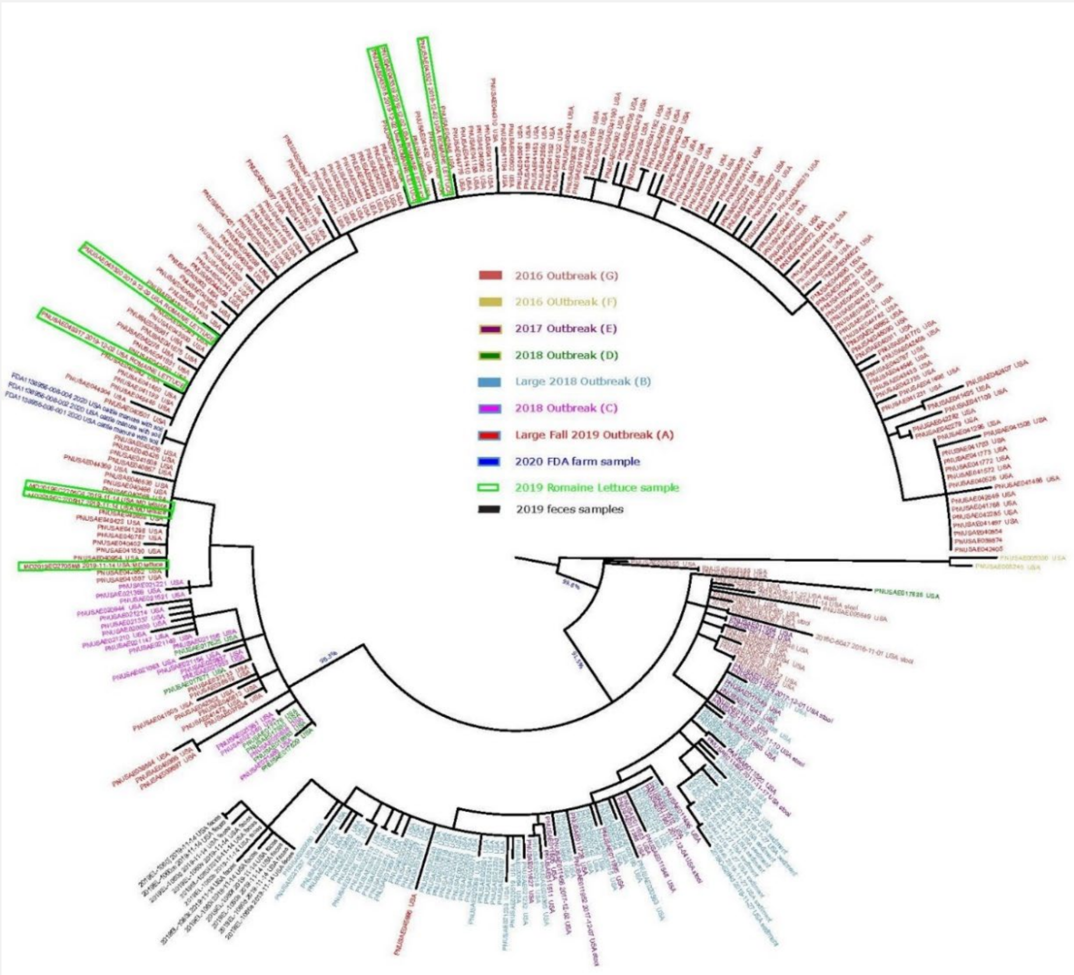
Figure 12. Compared to the REPEXH02 clusters, this REPEXH01 group is much more diversified. The smaller subcluster circles with longer spanning links suggest that host, environment, dispersal, and other unknown factors may promote more frequent growth opportunities for the *E. coli* O157:H7, leading to more mutations.

cgMLST Minimum Spanning Trees of REPEXH01 strain by Outbreak (confirmed or suspected vehicles; n = 1,052)



As part of FDA's GenomeTrakr Network and WGS database, the genome sequences associated with the isolates included in Figure 13 (below) can be accessed and are available to the public from the NCBI Pathogen Detection portal (https://www.ncbi.nlm.nih.gov/pathogens/. Additional analyses may be accomplished by downloading the STEC genomes and reconstruction of SNP-based phylogenetic assemblies using appropriate software and imaging tools. Both steps require a significant level of skill. Tutorials for using the NCBI Pathogen Detection portal are available at https://www.ncbi.nlm.nih.gov/pathogens.

Figure 13. The SNP analysis was done using CFSAN SNP Pipeline 2.0 with the complete genome of CFSAN101717 serving as the reference. In this example, the phylogenetic relationships are centrally anchored to the complete WGS of the isolate recovered from a clinically implicated and unopened prepackaged salad, exclusively containing romaine lettuce. Detailed characterization of SNP relatedness and diversification of the REPEXH02 clusters provides an important and complex mosaic of the *E. coli* O157:H7 REP group. The distribution and a more detailed analysis of the reoccurring strain cluster (the broadest and most simplistic way to refer to this collection of highly related isolates), around the circle, in relation to product source, geospatial details, season, adjacent land use seasonal profiles, practices and other factors, remain to be resolved. These data and the graphic visualization are useful to raise questions, in some cases, about whether an overlapping outbreak timeline is one, two, or more events.



Resources

Astill G. 2019. **Special Article: Seasonality in Romaine Outbreaks and Regional Shipments**. Vegetables and Pulses Outlook, VGS-362-SA, May 6, 2019, U.S. Department of Agriculture, Economic Research Service.

Barth SA, Menge C, Eichhorn I, Semmler T, Wieler LH, Pickard D, Belka A, Berens C, Geue L. 2016. **The accessory genome of Shiga toxin-producing Escherichia coli defines a persistent colonization type in cattle**. Appl Environ Microbiol. Aug 15;82(17):5455-64. doi: 10.1128/AEM.00909-16.

Burch AY, Do PT, Sbodio A, Suslow TV, Lindow SE. 2016. **High-level culturability of epiphytic bacteria and frequency of biosurfactant producers on leaves**. Appl Environ Microbiol. 82(19):5997-6009. doi: 10.1128/AEM.01751-16.

FDA. 2021a. CFSAN Constituent Updates. **FDA Takes Two Important Steps to Advance the Safety of Leafy Greens**. [Known and reasonably foreseeable hazard]

FDA. 2021b. **Executive Summary: Factors Potentially Contributing to the Contamination of Leafy Greens Implicated in the Fall 2020 Outbreak of E. coli O157:H7.**

FDA. 2021c. **Investigation Report: Factors Potentially Contributing to the Contamination of Leafy Greens Implicated in the Fall 2020 Outbreak of *E. coli* O157:H7**. https://www.fda.gov/media/147349/download

Kiesel K, Goodhue RE, Sexton RJ, Spalding A. 2021. E. coli in the Romaine Lettuce Industry: Economic Impacts from the November 2018 Outbreak. UC Davis Agricultural and Resource Economics.

Kim D, Hong S, Kim YT, Ryu S, Kim HB, Lee JH. 2018. **Metagenomic Approach to Identifying Foodborne Pathogens on Chinese Cabbage**. J. Microbiol. Biotechnol. 28:227-235. https://doi.org/10.4014/jmb.1710.10021

Leonard SR, Simko I, Mammel MK, et al. 2021. **Seasonality, shelf life and storage atmosphere are main drivers of the microbiome and *E. coli* O157:H7 colonization of post-harvest lettuce cultivated in a major production area in California**. *Environmental Microbiome 16, 25*. https://doi.org/10.1186/s40793-021-00393-y

Marshall KE, Hexemer A, Seelman SL, Fatica MK, Blessington T, Hajmeer M, et al. 2020. **Lessons learned from a decade of investigations of Shiga toxin-producing Escherichia coli outbreaks linked to leafy greens, United States and Canada.** Emerging Infectious Diseases 26(10), 2319–2328. https://doi.org/10.3201/eid2610.191418

Rastogi G, Sbodio A, Tech JJ, Suslow TV, Coaker GL, Leveau JH. 2012. **Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce.** *ISME J.* 6(10):1812-1822. doi:10.1038/ismej.2012.32

Smith OM, Olimpi EM, Navarro-Gonzalez N, Cornell KA, Frishkoff LO, Northfield TD, et al. 2022. **A trait-based framework for predicting foodborne pathogen risk from wild birds**. Ecological Applications. Accepted Author Manuscript, 17 December 2021. 2523. https://doi.org/10.1002/eap.2523

Wang LYR, Jokinen CC, Laing CR, Johnson RP, Ziebell K, Gannon VPJ. 2020. **Assessing the genomic relatedness and evolutionary rates of persistent verotoxigenic *Escherichia coli* serotypes within a closed beef herd in Canada.** Microb Genom. 6(6):e000376. doi: 10.1099/mgen.0.000376