Journal of Heredity, 2026, **117**, 526–536
https://doi.org/10.1093/jhered/esaf084
Advance access publication 22 October 2025
Genome Resources

American Genetic Association

OXFORD

Genome Resources

A genome assembly of Greene's tuctoria, *Tuctoria greenei*, an amphibious endemic and endangered California vernal pool grass

Daniel J. Toews[1,2,*,†], Merly Escalona[3,†], Lillie K. Pennington[1,4], Noravit Chumchim[5], Colin W. Fairbairn[6], Mohan P.A. Marimuthu[5], Courtney Miller[7], Oanh H. Nguyen[5], Dannise Ruiz-Ramos[8], William E. Seligmann[6], Molly Stephens[9], Robert D. Cooper[7], Erin Toffelmeier[7,10], H. Bradley Shaffer[7,10], Rachel S. Meyer[6], Jason P. Sexton[1]

[1]Department of Life and Environmental Sciences, University of California, Merced, CA, United States, [2]California Water Program, The Nature Conservancy, Sacramento, CA, United States, [3]Department of Biomolecular Engineering, University of California, Santa Cruz, Santa Cruz, CA, United States, [4]Genetics Department, University of Georgia, Athens, GA, United States, [5]DNA Technologies and Expression Analysis Core Laboratory, Genome Center, University of California, Davis, Davis, CA, United States, [6]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, United States, [7]Department of Ecology and Evolutionary Biology, University of California, Los Angeles (UCLA), Los Angeles, CA, United States, [8]Department of Natural Sciences, University of Maryland Eastern Shore, Princess Anne, MD 21853, United States, [9]Sierra Nevada Research Institute, University of California, Merced, Merced, CA, United States, [10]La Kretz Center for California Conservation Science, Institute of the Environment and Sustainability, University of California, Los Angeles (UCLA), Los Angeles, CA, United States

*Department of Life and Environmental Sciences, University of California, Merced, Science and Engineering 1 Building, Room 235, 5200 North Lake Road, Merced, CA 95343, USA Email: daniel.toews@tnc.org

†These authors contributed equally to this work.

Abstract

The Orcuttiinae subtribe of the grass tribe Cynodonteae (Poaceae) represents an ancient and unique group of amphibious grasses adapted to the winter-wet, summer-dry conditions of seasonally flooded vernal pools. The subtribe consists of nine species represented across three genera (*Neostapfia, Tuctoria, Orcuttia*), most of which are endemic to, and found exclusively in, vernal pools throughout the California Floristic Province (from the Modoc Plateau to Baja California, Mexico) and in the Magdalena Plain in the southern Baja California peninsula. All species are rare and most have state and federal Threatened and/or Endangered protected status in the U.S.—except *Tuctoria fragilis*, which inhabits Baja California Sur, Mexico, and does not have official protected status in Mexico. Here, we report a new chromosome–level reference genome assembly and annotation for Greene's tuctoria (*Tuctoria greenei*) developed in collaboration with the California Conservation Genomics Project. The assembly includes two haplotypes: haplotype one spans 2.59 Gb with contig N50 of 3.22 Mb, scaffold N50 of 216.09 Mb, largest contig N50 of 19.5 Mb, and BUSCO completeness of 96.8%. Haplotype two spans 258.89 Gb with contig N50 of 3.27 Mb and scaffold N50 of 213.15 Mb, with a BUSCO completeness of 97.4%. This genome assembly confirms earlier chromosome counts of $n=24$ for *T. greenei* and represents a powerful new tool that can be used to test hypotheses of gene flow, adaptation and comparative genomics between recently diverged species, and to assist in regional conservation priorities and restoration efforts.

Key words: amphibious, California conservation genomics project, CCGP, grass, Poaceae, vernal pool

Introduction

Vernal pools are ephemeral wetlands characterized by seasonal inundation and drying cycles that support a highly specialized and endemic biota, including numerous plant species (Keeler-Wolf et al. 1998; Keeley and Zedler 1998; Barbour et al. 2005). In California's Central Valley, these ecosystems have been severely impacted, with upwards of 95% of habitat lost due to urban development and destructive agricultural practices (Holland 2009; Witham et al. 2014). These losses have critically imperiled vernal pool specialist species, many of which exhibit narrow ecological niches and limited dispersal capacity (Keeley and Zedler 1998; Barbour et al. 2007).

Tuctoria greenei (Greene's tuctoria) is an annual, C_4 photosynthetic grass belonging to the subtribe Orcuttiinae within the tribe Cynodonteae (Poaceae: Chloridoideae). This subtribe includes three genera and nine species of semi–aquatic annual grasses that are endemic to vernal pool ecosystems in California and Baja California (Reeder 1982; Soreng et al.

Received on 16 May 2025; revised on 6 October 2025; accepted on 17 October 2025

2017). All species found in the United States are federally and state listed as Threatened and/or Endangered. *Tuctoria greenei* (Fig. 1) is one of three species in the genus *Tuctoria*; whereas *T. mucrunata* and *T. greenei* occur in California, *T. fragilis*—the type species of the genus—is restricted to Baja California, Mexico, making it the only member of the subtribe not found in the U.S. Like other members of Orcuttiinae, *T. greenei* is wind-pollinated and has amphibious life-history adaptations finely tuned to the extreme seasonal hydrological cycles of flooding and drought conditions of vernal pool ecosystems. Germination occurs under inundated and anaerobic conditions (Keeley 1988), with juvenile leaf formation while submerged; as pools dry, terrestrial leaves develop, allowing for rapid maturation and physiological drought tolerance (Keeley 1998). *Tuctoria greenei* also maintains a persistent seed bank, where seeds can remain dormant for several years and germinate when favorable conditions return (Griggs 1980). Nonetheless, habitat destruction and fragmentation have resulted in severe range contractions for *T. greenei*, leading to its listing as endangered under the U.S. Endangered Species Act in 1997 (USFWS 2024). Historically, *T. greenei* has been found across six of the sixteen California vernal pool ecoregions described by Keeler-Wolf et al. (1998). There are currently 31 known extant locations, many of which are in poor condition with decreasing population trends or have unknown status (CNDDB 2023). As with many organisms that inhabit vernal pool ecosystems, *T. greenei* is at risk of ongoing habitat loss, altered hydrology, and competition with invasive species.

Despite its critical conservation status, few genomic resources are available for *T. greenei*, which limits insights into its population genetic structure, adaptive potential, and evolutionary history. A population genetic study of Threatened and Endangered vernal pool grasses provided estimates of genetic variation, gene flow, and genetic structure for *T. greenei* (Gordon et al. 2012). It remains unclear, however, how much genetic variation is available in the seed bank, how variation may change between years, or if there is enough variation available to maintain evolutionary potential and respond to rapid climate change. The development of genetic resources is increasingly recognized as essential for informing conservation strategies, particularly for narrowly distributed taxa facing rapid environmental change (Fiedler et al. 2022; Shaffer et al. 2022). Reference genomes enable the identification of adaptive loci, reconstruction of historical demography, and delineation of conservation units, directly informing management actions. As part of the California Conservation Genomics Project (CCGP) (Shaffer et al. 2022), we present the first chromosome–scale reference genome for *T. greenei*. This resource will underpin future studies of ecological genomics, adaptive evolution, and conservation planning for vernal pool ecosystems.

Methods

Biological materials

To develop the high–quality reference genome for *T. greenei* we sampled leaf tissue from live plants grown from seed in lab conditions following methods adapted from Fisher (2013) and Griggs (1980). Obtaining fresh leaf material in the field from this species is challenging based on timing and unpredictable precipitation patterns that drive plant presence and phenology in this system. Access to land and the rarity of occurrences also contribute to challenges associated with sampling *T. greenei*. Thus, we obtained mature seeds by collecting whole inflorescences from senesced individuals at The Nature Conservancy's Vina Plains Preserve in summer (August) of 2018, a reliable and well-known occurrence for the species (CNDDB Element Occurrence No. 37). Briefly, plants were semi-randomly sampled by walking through the population and hand–picking intact inflorescences from five robust individuals and placing them in separate large coin envelopes. Seed envelopes were stored under cool, dry conditions in the Sexton Lab at University of California, Merced (Merced, CA). Whole inflorescences from individuals were placed into 6-inch sterile Petri dishes lined with Whatman No. 1 filter paper and submerged in ultra–purified molecular grade water. Petri dishes were immediately placed in a germination cabinet and cold stratified in the dark at 4°C for 7 weeks until seedlings emerged. Seedlings were transferred to trays containing native, saturated vernal pool soils and placed in growth chambers held at 23°C with 12-h days for 6 weeks. The youngest leaves were collected and flash-frozen in liquid nitrogen. Leaf tissue was split into sterile 1.5 mL cryo-tubes and submerged in a liquid nitrogen bath until frozen. Tubes containing flash–frozen leaf material were immediately packaged with dry ice and shipped to the University of California, Davis for high molecular weight (HMW) genomic DNA extraction and PacBio HiFi library preparation and sequencing and to the University of California, Santa Cruz for Dovetail Omni-C library preparation.

High-molecular-weight genomic DNA isolation

We extracted HMW genomic DNA (gDNA) from 140 mg of leaves using the cetyltrimethylammonium bromide (CTAB) method as described in Inglis et al. (2018) with the following modifications: 1) we used sodium metabisulfite (1% w/v) instead of 2-mercaptoethanol (1% v/v) in the sorbitol wash buffer and CTAB solution; 2) we repeated the tissue homogenate wash steps until the supernatant turned clear; 3) we performed the CTAB lysis step at 45°C and 4) we performed the chloroform extraction step twice using ice-cold chloroform. The DNA purity was estimated by absorbance ratios (260/280 = 1.82 and 260/230 = 2.40) measured using the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). The DNA yield (16 μg) was quantified using a Quantus Fluorometer (QuantiFluor ONE dsDNA Dye assay; Promega, Madison, WI), and the size distribution of the DNA was estimated using the Femto Pulse system (Genomic DNA 165 kb kit, Agilent, Santa Clara, CA), where 70% of the DNA fragments were found to be 30 kb or longer.

Nucleic acid library preparation

PacBio HiFi library preparation and sequencing

The HiFi SMRTbell library was constructed using the SMRTbell prep kit 3.0 (Pacific Biosciences (PacBio), Menlo Park, CA; Cat. #102-182-700) according to the manufacturer's instructions. HMW gDNA was sheared to a target DNA size distribution between 15 and 18 kb using Diagenode's Megaruptor 3 system (Diagenode, Belgium; cat. B06010003). The sheared gDNA was concentrated using 1X of SMRTbell cleanup beads provided in the SMRTbell prep kit 3.0 for the repair and a-tailing incubation at 37°C for 30 min and 65°C for 5 min, followed by ligation of overhang adapters at 20°C for 30 min, clean-up using 1X SMRTbell cleanup beads, and nuclease treatment at 37°C for 15 min. The SMRTbell library

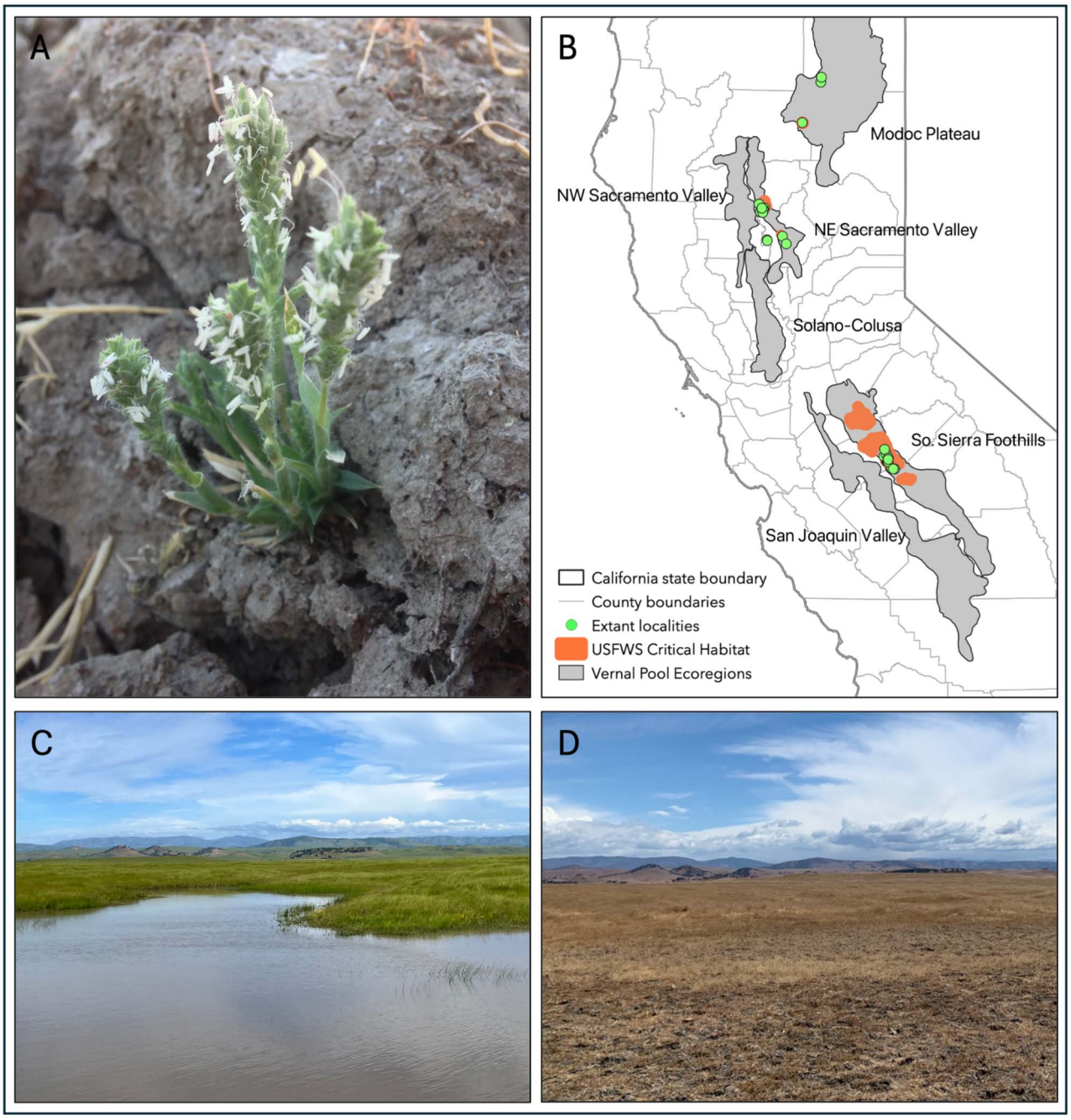

Fig. 1. Image of Green's tuctoria (*Tuctoria greenei*) used to generate the reference genome (A). Photo taken by Dr. Jason Sexton. Distribution map (B) for *T. greenei* based on California Native Diversity Database records (CNDDB 2023) with extant populations indicated by circles, polygons without border outlines representing U.S. Fish and Wildlife Service critical habitat (USFWS. 2005), and six vernal pool ecoregions shown with outlines, including the Modoc Plateau, Northwestern Sacramento Valley, Northeastern Sacramento Valley, San Joaquin Valley, Southern Sierra foothills, and Solano-Colusa regions (Keeler-Wolf et al. (1998). Photos of *T. greenei* habitat representing the fully aquatic phase (C) and dry terrestrial phase (D) of vernal pool seasonal conditions. Habitat photos by Dr. Daniel Toews.

was size selected using 3.1X of 35% v/v diluted AMPure PB beads (PacBio, Cat. #100-265-900) to progressively remove SMRTbell templates < 5 kb. The 15 to 18 kb average HiFi SMRTbell library was sequenced at UC Davis DNA Technologies Core (Davis, CA) using two 8 M SMRT cells (PacBio, Cat #101-389-001), Sequel II sequencing chemistry 2.0, and 30-h movies each on a PacBio Sequel IIe sequencer.

Omni-C library preparation and sequencing

The Omni-C library was prepared using the Dovetail™ Omni-C™ Kit (Dovetail Genomics, Scotts Valley, CA) according to the manufacturer's protocol with slight modifications. First, leaf tissue (Sample ID: TUGR_VPL22_DT) was thoroughly ground with a mortar and pestle while cooled with liquid nitrogen. Nuclear isolation was then performed using published methods (Workman et al. 2018). Subsequently, chromatin was fixed in place in the nucleus and digested under various conditions of DNase I until a suitable fragment length distribution of DNA molecules was obtained. Chromatin ends were repaired and ligated to a biotinylated bridge adapter, followed by proximity ligation of adapter containing ends. After proximity ligation, crosslinks were reversed, and the

DNA was purified from proteins. Purified DNA was treated to remove biotin that was not internal to ligated fragments. An NGS library was generated using a NEB Ultra II DNA Library Prep kit (New England Biolabs, Ipswich, MA) with an Illumina compatible y-adaptor. Biotin-containing fragments were then captured using streptavidin beads. The post capture product was split into two replicates prior to PCR enrichment to preserve library complexity, with each replicate receiving unique dual indices. The library was sequenced at Vincent J. Coates Genomics Sequencing Lab (Berkeley, CA) on an Illumina NovaSeq 6000 platform (Illumina, CA) to generate approximately 100 million 2 x 150 bp read pairs per gigabase (Gb) of genome size.

Transcriptome library preparation and sequencing

Total RNA was extracted from four tissue types—leaves, shoots, flowers, and roots—using the Qiagen RNeasy Mini Kit (Qiagen, Netherlands), following the manufacturer's protocol. RNA libraries were prepared with the KAPA mRNA HyperPrep Kit (Roche, Switzerland), also according to the manufacturer's instructions. Sequencing was performed on an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA) using 100 bp reads, generating approximately 50 million reads per library.

Nuclear genome assembly

We assembled the genome of *T. greenei* following the CCGP assembly pipeline, which uses PacBio HiFi reads and Omni-C data to produce a high-quality and highly contiguous assembly while minimizing manual curation, as outlined in Table 2 listing the tools and non-default parameters used in the process. We removed remnant adapter sequences from the PacBio HiFi dataset using HiFiAdapterFilt (Sim et al. 2022) and, joining the adapter-trimmed PacBio HiFi and the Omni-C datasets, we obtained the initial phased diploid assembly using HiFiasm (Cheng et al. 2021; Cheng et al. 2022).

We scaffolded each assembly using SALSA (Ghurye et al. 2017; Ghurye et al. 2019) with the Omni-C data and manually curated the assemblies by iteratively generating and analyzing their corresponding Omni-C contact maps. To generate the contact maps, we aligned the Omni-C data with BWA-MEM (Li 2013), identified ligation junctions, and generated Omni-C pairs using pairtools (Open2C et al. 2024). We generated a multi-resolution Omni-C matrix with cooler (Abdennur and Mirny 2020) and balanced it with hicExplorer (Ramírez et al. 2018). We used HiGlass (Kerpedjiev et al. 2018) and the PretextSuite (https://github.com/wtsi-hpag/PretextView; https://github.com/wtsi-hpag/PretextMap; https://github.com/wtsi-hpag/PretextSnapshot) to visualize the contact maps where we identified misassemblies and misjoins, and finally modified the assemblies using the Rapid Curation pipeline from the Wellcome Trust Sanger Institute, Genome Reference Informatics Team (https://gitlab.com/wtsi-grit/rapid-curation). Some of the remaining gaps (joins generated during scaffolding and/or curation) were closed using the adapter-filtered PacBio HiFi reads and YAGCloser (https://github.com/merlyescalona/yagcloser). Finally, we checked for contamination using the BlobToolKit Framework (Challis et al. 2020).

Genome size estimation and quality assessment

We generated k-mer counts from the PacBio HiFi reads using meryl (https://github.com/marbl/meryl). The k-mer counts were then used in GenomeScope2.0 (Ranallo-Benavidez et al. 2020) to estimate genome features including genome size, heterozygosity, and repeat content. To obtain general contiguity metrics, we ran QUAST (Gurevich et al. 2013). To evaluate genome quality and functional completeness, we used BUSCO (Manni et al. 2021) with the Embryophyta ortholog database (embryophyta_odb10), which contains 1,614 genes. Assessment of base level accuracy (QV) and k-mer completeness was performed using the previously generated meryl database and merqury (Rhie et al. 2020). We further estimated genome assembly accuracy via BUSCO gene set frameshift analysis using the pipeline described in Korlach et al. (2017). Measurements of the size of the phased blocks are based on the size of the contigs generated by HiFiasm on HiC mode. We follow the quality metric nomenclature established by Rhie et al. (2021), with the genome quality code x.y.P.Q.C, where, x = log10[contig NG50]; y = log10[scaffold NG50]; P = log10 [phased block NG50]; Q = Phred base accuracy QV (quality value); C = % genome represented by the first "n" scaffolds, following a karyotype of $2n = 24$, known for the number of chromosomes for this species (Reeder 1982). Quality metrics for the notation were calculated on the assembly for haplotype 1.

Genome annotation

We annotated the reference assembly for *T. greenei* using the NCBI Eukaryotic Genome Annotation Pipeline v0.3.2-alpha (hereafter, "EGAPx"), which is published in the NCBI RefSeq database (O'Leary et al. 2016) and accessible through the NCBI GitHub page ("ncbi/egapx"). Annotation features were identified by aligning transcripts and proteins from related taxa in the RefSeq database using BLAST (Camacho et al. 2009). Novel, species-specific RNA-Seq reads generated from five tissue types were also aligned to the assembly using the alignment software STAR (Dobin et al. 2013). Additional features are predicted using HMM–based gene models using the NCBI Gnomon software. We evaluated the quality and completeness of our annotation by comparing the longest protein for each annotated coding gene to Poales (odb10) using BUSCO v5.7.1 (Manni et al. 2021). We report the number of annotation features and the BUSCO results in Table 3.

Chloroplast genome assembly

The chloroplast sequence for *T. greenei* was generated with Oatk (Zhou et al. 2025) and we used GeSeq (Tillich et al. 2017) to generate a draft genome annotation. We used the chloroplast genome assembly of *Arabidopsis thaliana* (NCBI:NC_000932.1; Sato et al. 1999) as a guide for manual curation, in which we aligned the generated sequence against the guide, using lastz (Harris 2007), extracted contigs, and fixed orientation when needed using samtools (Danecek et al. 2021) and seqtk (https://github.com/lh3/seqtk). Visual validation of the alignment was done using LAJ (Wilson et al. 2001). The resulting assembly was annotated using the online version of GeSeq (Tillich et al. 2017) and visualized using the online version of OGDRAW (Greiner et al. 2019). After completion of the nuclear genome, we searched for matches of the resulting chloroplast assembly sequence in the nuclear

genome assembly using BLAST+ (Camacho et al. 2009) and filtered out contigs and scaffolds from the nuclear genome with a percentage of sequence identity > 99% and size smaller than the chloroplast assembly sequence.

Results

The Omni-C and PacBio HiFi sequencing libraries generated 182.66 million read pairs and 4.31 million reads, respectively. The latter yielded ~23-fold coverage, with N50 read length of 14,686 bp; minimum read length 118 bp; mean read length 13,765 bp; maximum read length of 53,438 bp (see Supplementary Fig. S1 for read length distribution). Based on the PacBio HiFi long reads, we estimated with Genomescope 2.0 a genome assembly size estimation of 2.56 Gb, 0.205% sequencing error rate, and 0.01% nucleotide heterozygosity rate. The k-mer spectrum shows a bimodal distribution with two major peaks at ~11 and ~22-fold coverage (Fig. 2A). Sequencing of mRNA libraries for SAMN41791172, SAMN41406448, SAMN41791171, SAMN41791173 yielded 54M, 48M, 51M, 48M paired-end reads, respectively.

The final assembly (lpTucGree1) consists of two phased haplotypes that vary slightly in size compared to the estimated value from GenomeScope2.0 (Fig. 2A), as has been observed in other taxa (see Pflug et al. 2020 for example). Haplotype 1 consists of 599 scaffolds spanning 2.59 Gb with contig N50 of 3.22 Mb, scaffold N50 of 216.09 Mb, largest contig of 19.5 Mb, and largest scaffold of 308.98 Mb. On the other hand, haplotype two consists of 451 scaffolds, spanning 2.58 Gb with contig N50 of 3.27 Mb, scaffold N50 of 213.15 Mb, largest contig 20.68 Mb and largest scaffold of 303.04 Mb.

During manual curation, we generated a total of 219 breaks and 534 joins, where 118 breaks were made on haplotype one and 101 were made on haplotype two, while we made 286 joins on haplotype one, and 248 joins on haplotype two. We were able to close a total of 111 gaps, 55 on haplotype one and 56 on haplotype two. Finally, we removed a total of 5 contigs due to contaminants from the Aves Phylum, 3 on haplotype one and 2 on haplotype two; and removed 15 contigs matched to the chloroplast, 8 from haplotype one and 7 from haplotype two.

The haplotype one has a BUSCO completeness score of 96.8% using the Embryophyta gene set, a per base quality (QV) of 63.3, a k-mer completeness of 83.82 and a frameshift indel QV of 44.91. The haplotype two has a BUSCO completeness score of 97.4% using the same gene set, a per base quality (QV) of 63.24, a k-mer completeness of 83.93 and a frameshift indel QV of 46.27. The Omni-C contact maps show that both assemblies are highly contiguous, suggesting that the genome of *T. greenei* is organized in 12 chromosomes based on the number of major bins along the diagonal of the contact maps (Fig. 2C and D; Supplementary Table S1). Assembly statistics are reported in Table 3, and graphical representation for the haplotype 1 assembly in Fig. 2B (see Supplementary Fig. S2 for haplotype 2 graphical representation). We have deposited both assemblies on NCBI (see Table 3 and Data Availability for details).

Our final genome annotation for *T. greenei* included 33 566 genes, with a Poales BUSCO completeness of 96.1%. A list of annotation statistics and BUSCO score breakdowns for Poales is reported in Table 3.

Chloroplast genome assembly

The final chloroplast genome assembly spans 133,732 bp, has a nucleotide composition of A = 31.02%, C = 19.09%, G = 19.17%, T = 30.73%, and 3 gaps (sequence of running Ns of 10 bp). The large single copy region is 80,712 bp long, the small single copy region is 12,770 bp long, and the inverted repeat (IR) is 20,110 bp long. This plastome contains a total of 299 genes including rRNAs and tRNAs (not counting the duplication of the IR, Supplementary Fig. S3). We have deposited the chloroplast assembly on NCBI (See Table 3 and Data Availability for details).

Discussion

This is the first reference genome published for a *Tuctoria* species, and one of the first for California vernal pool endemic plants in general (Pennington et al. 2025, *Neostapfia* genome assembly), and will be useful for further population genomics and phylogenetics research. This work establishes a genomic baseline for *Tuctoria* that can be used by management to set priorities and goals across endangered ecosystems such as those found in California's Great Central Valley. Moreover, this genome can be used to compare herbarium specimens from extirpated and historical records of extant populations, allowing for an understanding of genomic changes that have occurred across the species' ranges of these species over time.

The grass subtribe Orcuttiinae has evolved to specialize in the unique hydrology of vernal pool ecosystems (Boykin et al. 2010), yet due to modern land changes, all species that live within California within this group are now listed by the State of California or the Federal Government as Rare, Threatened or Endangered (Table 1). *Tuctoria greenei* appears to be diploid, with $2n = 24$ chromosomes (Fig. 2C, D), which is consistent with past treatments based on cell squash techniques (e.g. Reeder 1982). The estimated genome size of *T. greenei* is relatively large at 2.59 Gb, compared to 2.21 Gb for another Orcuttiinae grass, *Neostapfia colusana* ($2n = 40$) (Pennington et al. 2025).

Orcuttiinae grasses have been the subject of conservation genetics (e.g. Gordon et al. 2012) and phylogenetics research (e.g. Boykin et al. 2010), but to date, few genomic resources exist. Gordon et al. (2012) conducted a population genetics study of *T. greenei* (and *N. colusana*) using five microsatellite markers. They found high within-population diversity in this wind-pollinated species, evidence for regional genetic structure, and temporal genetic variation between sampling years. Griggs and Jain (1983), in a prior study using allozyme markers, found high intrapopulation genetic variation in a study of two populations. Boykin et al. (2010) conducted a phylogenetic study of the Orcuttiinae clade and concluded that *Tuctoria* is paraphyletic, raising the need for further studies of this group.

This project will provide genomic tools that can be used for future conservation genomics aims and goals for the rest of this special-status clade. These tools and databases will enable state and federal agencies to understand historical and contemporary gene flow patterns, patterns of genetic variation, and regions of highest genetic vulnerability and resilience. Among other aims, these data will be highly useful for the future development of conservation and restoration of these species and ecosystems, as well as to inform potential patterns of genetic diversity, current and past, within the Great Central Valley. Most of the plants that make California a

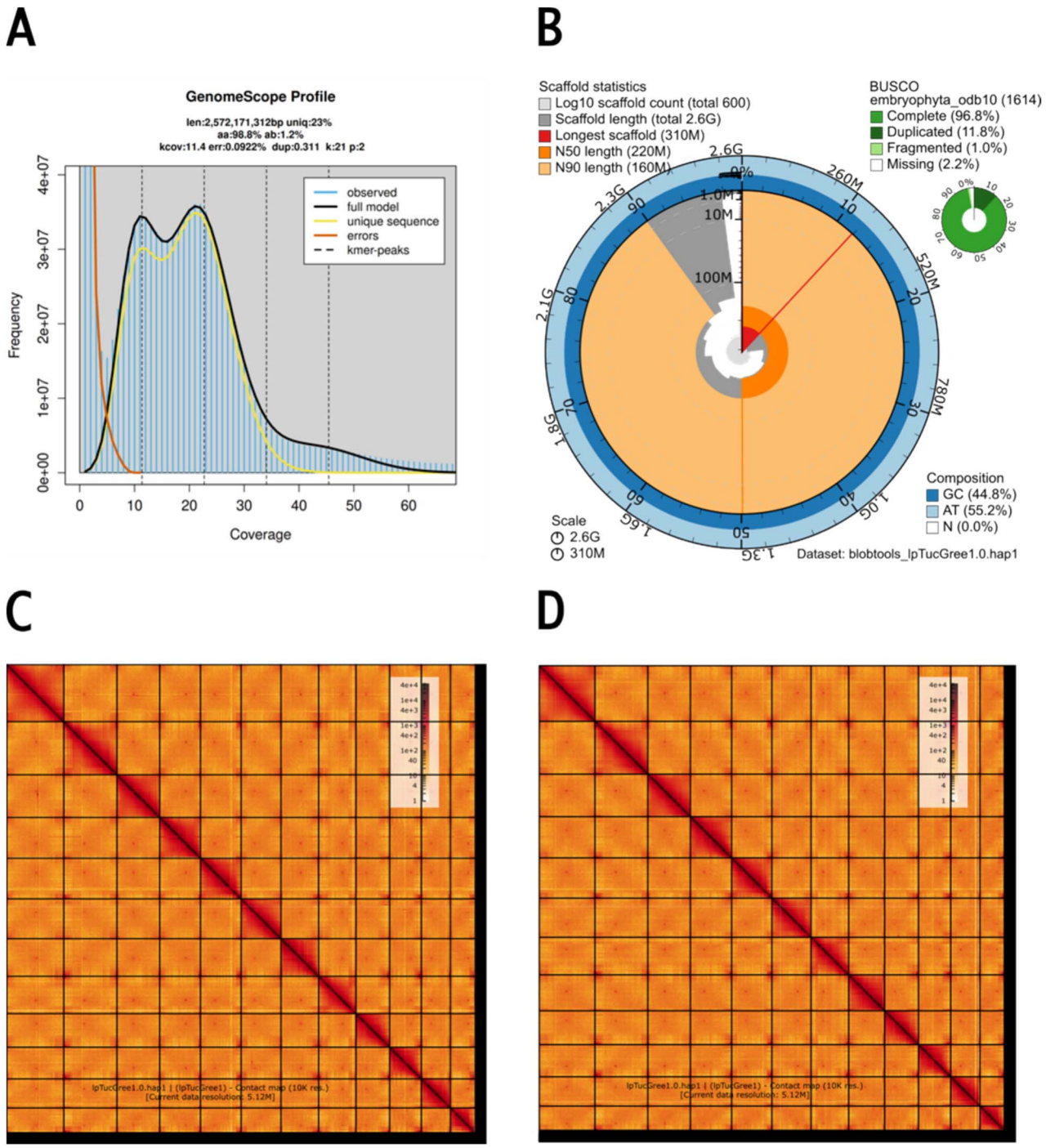

Fig. 2. Visual overview of the genome assembly for *Tuctoria greenei*. (A) K-mer spectrum with vertical bars representing the observed k-mer frequencies, the bold line shows the GenomeScope model fit, while the lines under and outside the bold model fit line indicate the contributions of unique sequences and sequencing errors, respectively, and the vertical dashed lines mark estimated k-mer coverage peaks. (B) BlobToolKit Snailplot showing quality metrics for *T. greenei* haplotype one assembly presented in Table 3. The circular plot presents scaffold statistics, with scaffold lengths shown in grey and scaffold counts (log10 scale) in white. Outer rings show GC content (44.8%) and AT content (55.2%), with no ambiguous bases (N). BUSCO scores for the Embryophyta set of orthologues indicate high assembly completeness. Omni-C contact heatmaps for the primary (C) and alternate (D) assemblies reflect chromatin interaction frequency, with red representing higher contact intensity. Strong diagonal patterns and distinct chromosomal interaction blocks indicate well-assembled, chromosome-scale scaffolds. The maps support the structural integrity of both assemblies and highlight the continuity and organization of genomic regions in the lpTucGree1 genome.

major biodiversity hotspot are annual species like members of the tribe Orcuttiinae. Thus, this newest genome reference will be beneficial by adding new, useful, relevant genomic resources for plant conservation in California.

Furthermore, understanding and managing rare Orcuttiinae amphibious grasses benefits other wetland species, including other plants, animals, fungi, and microbial communities in these ecosystems (Montiel-Molina et al. 2023; Ruiz-Ramos

Table 1. California state and U.S. federal government listing status for the eight Orcuttiinae species within California (USFWS, 2024).

Species	Common name	State; federal listing status
Neostapfia colusana	Colusa grass	State endangered; federally threatened
Orcuttia californica	California orcutt grass	State endangered; federally endangered
Orcuttia inaequalis	San Joaquin orcutt grass	State endangered; federally threatened
Orcuttia pilosa	Hairy orcutt grass	State endangered; federally endangered
Orcuttia tenuis	Slender orcutt grass	State endangered; federally threatened
Orcuttia viscida	Sacramento orcutt grass	State endangered; federally endangered
Tuctoria greenei	Green's tuctoria	State rare; federally endangered
Tuctoria mucronata	Prickly spiralgrass	State endangered; federally endangered

Table 2. Assembly pipeline and software used. Software citations are listed in the text.

Assembly	Software and any non-default options	Version
Filtering PacBio HiFi adapters	HiFiAdapterFilt	Commit 64d1c7b
K-mer counting	Meryl (k = 21)	1
Estimation of genome size and heterozygosity	GenomeScope (k = 21, l−7)	2
De novo *assembly (contiging)*	HiFiasm (Hi-C Mode, −primary, output p_ctg.hap1, p_ctg.hap2)	0.16.1−r375
Scaffolding		
Omni-C data alignment	Arima genomics mapping pipeline	Commit 2e74ea4
Omni-C Scaffolding	SALSA (-DNASE, -i 20, -p yes)	2
Gap closing	YAGCloser (-mins 2 -f 20 -mcc 2 -prt 0.25 -eft 0.2 -pld 0.2)	Commit 0e34c3b
Omni-C Contact map generation		
Short-read alignment	BWA-MEM (-5SP)	0.7.17-r1198-dirty
SAM/BAM processing	samtools	1.16-8-g629fb5e (using htslib 1.16-3-g5036186)
SAM/BAM filtering	pairtools	0.3.0
Pairs indexing	pairix	0.3.7
Matrix generation	cooler	0.8.10
Matrix balancing	hicExplorer (hicCorrectmatrix correct –filterThreshold -2 4)	3.6
Contact map visualization	HiGlass	2.1.11
	PretextMap	0.1.4
	PretextView	0.1.5
	PretextSnapshot	0.0.3
Manual curation tools	Rapid curation pipeline (Wellcome Trust Sanger Institute, Genome Reference Informatics Team)	Commit 4ddca450
Genome quality assessment		
Basic assembly metrics	QUAST (–est-ref-size)	5.0.2
Assembly completeness	BUSCO (-m geno, -l embryophyta)	5.0.0
	Merqury	29 January 2020
Chloroplast genome assembly		
de novo assembler of organelle genomes	Oatk (-k 1001, -c 300)	1.0 (Commit 6591c42)
Sequence editing	seqtk	1.3-r115-dirty
Sequence alignment	lastz (–nogapped,–notransition, –step = 20)	1.04.15
Alignment visualization	LAJ (http://globin.cse.psu.edu/dist/laj/)	14 December 2005
Genome annotation	GeSeq (https://chlorobox.mpimp-golm.mpg.de/geseq.html)	13 July 1905
Contamination screening		
Local alignment tool	BLAST+ (-db nt, -outfmt '6 qseqid staxids bitscore std', -max_target_seqs 1, -max_hsps 1, -evalue 1e-25)	2.10
General contamination screening	BlobToolKit (PacBIo HiFi Coverage, NCBI Taxa ID = 160,569, BUSCODB = embryophyta)	2.3.3

et al. 2023). Finally, this work complements existing genomic resources produced by the CCGP for vernal pool plant and animal species—including related Orcuttiinae grasses (Pennington et al. 2025; Toews et al., unpublished manuscript), fairy shrimp (Kieran Blair et al. 2023a; Kieran Blair et al. 2023b), tadpole shrimp (Kieran Blair et al. 2022), and others (www.ccgproject.org/species)—further enhancing our ability to understand and conserve the unique, ephemeral ecosystems that support these highly specialized and often imperiled taxa.

Table 3. Sequencing and assembly statistics, and accession numbers.

Bio projects & vouchers	**CCGP NCBI BioProject**		**PRJNA720569**			
	Genera NCBI BioProject		**PRJNA765880**			
	Species NCBI BioProject		**PRJNA808381**			
	NCBI Genome BioSample		SAMN38285869			
	Specimen identification (Genome)		TUGR_VPL22_DT			
	NCBI RNA BioSamples		SAMN41791172, SAMN41406448, SAMN41791171, SAMN41791173			
	NCBI Genome accessions		**Haplotype 1**		**Haplotype 2**	
	Assembly accession		JAYESD000000000		JAYESE000000000	
	Genome sequences		GCA_036927565.1		GCA_036927555.1	
Genome sequence	PacBio HiFi reads	Run	1 PACBIO_SMRT (Sequel IIe) run: 4.3 M spots, 59.3G bases, 35.5Gb			
		Accession	SRX25183507			
	Omni-C Illumina reads	Run	2 ILLUMINA (Illumina NovaSeq 6000) runs: 182.6 M spots, 55.2G bases, 18.2Gb			
		Accession			SRX25183508-9	
Genome assembly quality metrics	Assembly identifier (quality code*)		lpTucGree1(6.8.P6.Q63.C97)			
	HiFi Read coverage§				48.10X	
			Haplotype 1		**Haplotype 2**	
	Number of contigs		1,846		1,688	
	Contig N50 (bp)		3,223,507		3,270,295	
	Contig NG50§		5,770,543		5,618,589	
	Longest Contigs		19,508,695		20,687,851	
	Number of scaffolds		599		451	
	Scaffold N50		216,090,640		213,150,644	
	Scaffold NG50§		231,641,634		231,416,900	
	Largest scaffold		308,983,293		308,983,293	
	Size of final assembly		2,595,920,285		2,589,077,426	
	Phased block NG50§		5,676,136		5,635,868	
	Gaps per Gbp (# Gaps)		480 (1,247)		478 (1,237)	
	Indel QV (Frame shift)		44.91		46.27	
	Base pair QV		63.3079		63.2486	
			Full assembly = 63.2782			
	k-mer completeness		83.8236		83.9336	
			Full assembly = 97.8002			
	BUSCO completeness**		**C**	**S**	**D**	**F**	**M**
	(embryophyta) $n = 1\,614$	H1‡	96.80%	85.00%	11.80%	1.00%	2.20%
		H2‡	97.40%	85.60%	11.80%	0.80%	1.80%
	Organelles	1 partial chloroplast assembly				JAYESD010000599.1	
Genome annotation quality metrics				Count of features			
	Genes			33,566			
	Transcipts			4,717			
	mRNA			32,892			
	lncRNA			2,708			
	CDSs			32,892			
	BUSCO completeness**		**C**	**S**	**D**	**F**	**M**
	Poales (odb_10; $n = 4\,896$)		96.1%	73.8%	22.3%	0.4%	3.5%

[a] Assembly quality code x.y.P.Q.C derived notation, from (Rhie et al. 2021). x = log10[contig NG50]; y = log10[scaffold NG50]; P = log10 [phased block NG50]; Q = Phred base accuracy QV (Quality value); C = % genome represented by the first "n" scaffolds, following a known karyotype for this species of $2n = 24$ (Reeder 1982). Quality code for all the assembly denoted by primary assembly (lpTucGree1.0.hap1). [b] BUSCO Scores. Complete BUSCOs (C). Complete and single-copy BUSCOs (S). Complete and duplicated BUSCOs (D). Fragmented BUSCOs (F). Missing BUSCOs (M). [c] Read coverage and NGx statistics have been calculated based on the estimated genome size of 2.53 Gb [d] (H1) Haplotype 1 and (H2) haplotype 2 assembly values.

Acknowledgments

PacBio Sequel II/IIe library prep and sequencing was carried out at the DNA Technologies and Expression Analysis Core at the UC Davis Genome Center, supported by NIH Shared Instrumentation Grant 1S10OD010786-01. Deep sequencing of Omni-C libraries used the Novaseq S4 sequencing platforms at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 OD018174 Instrumentation Grant. We thank the staff at the UC Davis DNA Technologies and Expression Analysis Core and the UC Santa Cruz Paleogenomics Laboratory for their diligence and dedication to generating high-quality sequence data. We also acknowledge Robert Schlising with CSU Chico and Andrea Craig with The Nature Conservancy for guidance and access to Vina Plain Preserve, and Carol Witham and Dr Robert Holland for botanical expertise and land access to sampling locations on private lands. We thank the numerous landowners and current stewards of these unique vernal pool habitats for access to private land, including the Flying M, Morrison, and Roen ranches in East Merced County, CA. RNA extraction, library preparation, and sequencing were carried out at the UCLA Technology Center for Genomics & Bioinformatics.

Funding

This work was supported by the California Conservation Genomics Project, with funding provided to the University of California by the

State of California, State Budget Act of 2019 (UC Award ID RSI-19-690224), as well as the University of California President's Research Catalyst Award (CA-16-376437); US Fish and Wildlife Service and Bureau of Reclamation's Central Valley Project Conservation Program (CVPCP); Central Valley Project Improvement Act Habitat Restoration Program (HRP) (CESU—R17AC00044).

Author contributions

Daniel J Toews (Methodology, Resources, Visualization, Writing—original draft), Merly Escalona (Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review & editing), Lillie K Pennington (Resources, Writing—original draft), Noravit Chumchim (Data curation, Formal analysis, Methodology, Software, Writing—review & editing), Colin W Fairbairn (Data curation, Formal analysis, Methodology), Mohan PA Marimuthu (Data curation, Resources, Software), Courtney Miller (Funding acquisition, Project administration, Resources, Supervision), Oanh Nguyen (Data curation, Formal analysis, Writing—review & editing), Dannise Ruiz-Ramos (Formal analysis, Methodology, Resources, Writing—review & editing), William Seligmann (Data curation, Formal analysis, Methodology, Resources, Software, Writing—review & editing), Molly Stephens (Investigation, Writing—review & editing), Robert D Cooper (Data curation, Formal analysis, Methodology, Resources, Software, Validation, Writing—review & editing), Erin Toffelmier (Conceptualization, Investigation, Methodology, Project administration, Resources), H Bradley Shaffer (Conceptualization, Funding acquisition, Project administration, Supervision), Rachel Meyer (Data curation, Formal analysis, Supervision, Writing—original draft, Writing—review & editing), and Jason P Sexton (Investigation, Project administration, Supervision, Writing—original draft, Writing—review & editing)

Supplementary material

Supplementary material is available at *Journal of Heredity* online.

Data availability

Data generated for this study are available under NCBI BioProject PRJNA808381. Raw sequencing data for sample TUGR_VPL22_DT (NCBI BioSample SAMN38285869) are deposited in the NCBI Short Read Archive (SRA) under SRR29680270 for PacBio HiFi sequencing data, and SRR29680268, SRR29680269 for the Omni-C Illumina sequencing data. GenBank accessions for both primary and alternate assemblies are JAYESD000000000 and JAYESE000000000; and for genome sequences GCA_036927565.1 and GCA_036927555.1. Assembly scripts and other data for the analyses presented can be found at the following GitHub repository: www.github.com/ccgproject/ccgp_assembly. Data generated for the annotation in this study are available under NCBI BioProject PRJNA1013298. Raw RNA sequencing data (NCBI BioSample(s) SAMN41791172, SAMN41406448, SAMN41791171, SAMN41791173) are deposited in the NCBI SRA (SRR29366288, SRR29366286, SRR29366289, SRR29366287), respectively).

References

Abdennur N, Mirny LA. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics*. 2020;**36**:311–316. https://doi.org/10.1093/bioinformatics/btz540

Barbour MG, Solomeshch AI, Holland RF, Witham CW, Macdonald RL, Cilliers SS, Molina JA, Buck JJ, Hillman JM. Vernal pool vegetation of California: communities of long-inundated deep habitats. *Phytocoenologia*. 2005;**35**:177–200. https://doi.org/10.1127/0340-269X/2005/0035-0177

Barbour MG, Solomeshch AI, Buck JJ, Holland RF, Witham CW, MacDonald RL, Starr SL, Lazar KA. Classification, ecological characterization, and presence of listed plant taxa of vernal pool associations in California. In: *Final report, United States fish and wildlife service agreement/study*. 2007: No. 814205G238. University of California, Davis, California, USA.

Boykin LM, Kubatko LS, Lowrey TK. Comparison of methods for rooting phylogenetic trees: a case study using Orcuttieae (Poaceae: Chloridoideae). *Mol Phylogenet Evol*. 2010;**54**:687–700. https://doi.org/10.1016/j.ympev.2009.11.016

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;**10**:421. https://doi.org/10.1186/1471-2105-10-421

Challis R, Richards E, Rajan J, Cochrane G, Blaxter M. BlobToolKit—interactive quality assessment of genome assemblies. *G3 GenesGenomesGenetics*. 2020;**10**:1361–1374. https://doi.org/10.1534/g3.119.400908

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods*. 2021;**18**:170–175. https://doi.org/10.1038/s41592-020-01056-5

Cheng H, Jarvis ED, Fedrigo O, Koepfli K-P, Urban L, Gemmell NJ, Li H. Haplotype-resolved assembly of diploid genomes without parental data. *Nat Biotechnol*. 2022;**40**:1332–1335. https://doi.org/10.1038/s41587-022-01261-x

CNDDB. California natural diversity database, California department of fish and game. 2023; http://www.dfg.ca.gov/biogeodata/cnddb. Accessed July 2023

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. Twelve years of SAMtools and BCFtools. *GigaScience*. 2021;**10**:giab008. https://doi.org/10.1093/gigascience/giab008

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;**29**:15–21. https://doi.org/10.1093/bioinformatics/bts635

Fiedler PL, Erickson B, Esgro M, Gold M, Hull JM, Norris J, Shapiro B, Westphal M, Toffelmier E, Shaffer HB. Seizing the moment: the opportunity and relevance of the California conservation genomics project to state and federal conservation policy. *J Hered*. 2022;**113**:589–596. https://doi.org/10.1093/jhered/esac046

Fisher EG. *Road to recovery: introduction of two rare vernal pool grasses, Greene's tuctoria (Tuctoria greenei) and Colusa grass (Neostapfia colusana)*. 2013: [Master's thesis, California State University, Chico]. Chico, CA, USA. Available at ScholarWorks. https://scholarworks.calstate.edu/downloads/t148fh720.

Ghurye J, Pop M, Koren S, Bickhart D, Chin C-S. Scaffolding of long read assemblies using long range contact information. *BMC Genomics*. 2017;**18**:527. https://doi.org/10.1186/s12864-017-3879-z

Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, Koren S. Integrating hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput Biol*. 2019;**15**:e1007273. https://doi.org/10.1371/journal.pcbi.1007273

Gordon SP, Sloop CM, Davis HG, Cushman JH. Population genetic diversity and structure of two rare vernal pool grasses in Central California. *Conserv Genet*. 2012;**13**:117–130. https://doi.org/10.1007/s10592-011-0269-y

Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res*. 2019;**47**:W59–W64. https://doi.org/10.1093/nar/gkz238

Griggs FT. *Population studies in the genus Orcuttia (Poaceae)*. Davis: University of California; 1980.

Griggs FT, Jain SK. Conservation of vernal pool plants in California, II. Population biology of a rare and unique grass genus Orcuttia. *Biol Conserv*. 1983;**27**:171–193. https://doi.org/10.1016/0006-3207(83)90087-3

Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;**29**:1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Harris RS. *Improved pairwise alignment of genomic DNA*[Doctoral dissertation, The Pennsylvania State University]. 2007. ProQuest Dissertations Publishing. https://www.proquest.com/docview/304835295/abstract/E3B8E50565384537PQ/1

Holland RF. California's Great Valley vernal pool habitat status and loss: Rephotorevised., 2005. In: Unpublished technical report prepared for placer land trust. *Placer Land Trust. Auburn*, CA, USA; 2009:1–23.

Inglis PW, Pappas M D CR, Resende LV, Grattapaglia D. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. *PLoS One*. 2018;**13**:e0206085. https://doi.org/10.1371/journal.pone.0206085

Keeler-Wolf T, Elam D, Lewis K, Flint S. *California vernal pool assessment: a preliminary report*. The Resources Agency, Department of Fish and Game, Sacramento, CA: State of California; 1998.

Keeley JE. Anaerobiosis as a stimulus to germination in two vernal pool grasses. *Am J Bot*. 1988;**75**:1086–1089

Keeley JE. C4 photosynthetic modifications in the evolutionary transition from land to water in aquatic grasses. *Oecologia*. 1998;**116**:85–97. https://doi.org/10.1007/s004420050566

Keeley JE, Zedler PH. Characterization and global distribution of vernal pools. In ecology, conservation, and management of vernal pool ecosystems—proceedings from 1996 conference. *California Native Plant Society, Sacramento, CA*. 1998;**1**:1–14

Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, Azhir A, Kumar N, et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol*. 2018;**19**:125. https://doi.org/10.1186/s13059-018-1486-1

Kieran Blair SR, Hull J, Escalona M, Finger A, Joslin SE, Sahasrabudhe R, Marimuthu MP, Nguyen O, Chumchim N, Morris ER, et al. The reference genome of the vernal pool tadpole shrimp, Lepidurus packardi. *J Hered*. 2022;**113**:706–711. https://doi.org/10.1093/jhered/esac051

Kieran Blair SR, Schreier A, Escalona M, Finger AJ, Joslin SE, Sahasrabudhe R, Marimuthu M, Nguyen O, Chumchim N, Morris ER, et al. A chromosome-level reference genome for the versatile fairy shrimp, Branchinecta lindahli. *J Hered*. 2023a;**114**:74–80. https://doi.org/10.1093/jhered/esac057

Kieran Blair SR, Schreier A, Escalona M, Finger AJ, Joslin SE, Sahasrabudhe R, Marimuthu MP, Nguyen O, Chumchim N, Morris ER, et al. A draft reference genome of the vernal pool fairy shrimp, Branchinecta lynchi. *J Hered*. 2023b;**114**:81–87. https://doi.org/10.1093/jhered/esac056

Korlach J, Gedman G, Kingan SB, Chin C-S, Howard JT, Audet J-N, Cantin L, Jarvis ED. De novo PacBio long-read and phased avian genome assemblies correct and add to reference genes generated with intermediate and short reads. *GigaScience*. 2017;**6**:gix085. https://doi.org/10.1093/gigascience/gix085

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* preprint arXiv:1303.3997. 2013; Retrieved from http://arxiv.org/abs/1303.3997

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;**38**:4647–4654. https://doi.org/10.1093/molbev/msab199

Montiel-Molina JAM, Sexton JP, Frank AC, Beman JM. Archaeal and bacterial diversity and distribution patterns in Mediterranean-climate vernal pools of Mexico and the western USA. *Microb Ecol*. 2023;**85**:24–36. https://doi.org/10.1007/s00248-021-01941-2

O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 2016;**44**:D733–D745. https://doi.org/10.1093/nar/gkv1189

Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, Imakaev M, Venev SV. Pairtools: from sequencing data to chromosome contacts. *PLoS Comput Biol*. 2024;**20**,5:e1012164. https://doi.org/10.1371/journal.pcbi.1012164

Pennington LP, Escalona M, Toews DJ, Chumchim N, Cooper R, Fairbairn CW, Marimuthu MP, Miller C, Nguyen OH, Ruiz-Ramos D, et al. and J.P., Sexton. A reference genome for Colusa grass, *Neostapfia colusana*, a threatened and endangered California vernal pool plant. *J Hered*. 2025: https://doi.org/10.1093/jhered/esaf075

Pflug JM, Holmes VR, Burrus C, Johnston JS, Maddison DR. Measuring genome sizes using read-depth, k-mers, and flow cytometry: methodological comparisons in beetles (coleoptera). *G3 GenesGenomesGenetics*. 2020;**10**:3047–3060. https://doi.org/10.1534/g3.120.401028

Ramírez F, Bhardwaj V, Arrigoni L, Lam KC, Grüning BA, Villaveces J, Habermann B, Akhtar A, Manke T. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat Commun*. 2018;**9**:189. https://doi.org/10.1038/s41467-017-02525-w

Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nature. *Communications*. 2020;**11**:1432. https://doi.org/10.1038/s41467-020-14998-3

Reeder JR. Systematics of the tribe Orcuttieae (Gramineae) and the description of a new segregate genus, *Tuctoria*. *American Journal of Botany*. 1982;**69**:1082–1095. https://doi.org/10.1002/j.1537-2197.1982.tb13353.x

Rhie A, Walenz BP, Koren S, Phillippy AM. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol*. 2020;**21**:245. https://doi.org/10.1186/s13059-020-02134-9

Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021;**592**:737–746. https://doi.org/10.1038/s41586-021-03451-0

Ruiz-Ramos DV, Meyer RS, Toews D, Stephens M, Kolster MK, Sexton JP. Environmental DNA (eDNA) detects temporal and habitat effects on community composition and endangered species in ephemeral ecosystems: a case study in vernal pools. *Environ DNA*. 2023;**5**:85–101. https://doi.org/10.1002/edn3.360

Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. Complete structure of the chloroplast genome of Arabidopsis thaliana. *DNA Res*. 1999;**6**:283–290. https://doi.org/10.1093/dnares/6.5.283

Shaffer HB, Toffelmier E, Corbett-Detig RB, Escalona M, Erickson B, Fiedler P, Gold M, Harrigan RJ, Hodges S, Luckau TK, et al. Landscape genomics to enable conservation actions: the California conservation genomics project. *J Hered*. 2022;**113**:577–588. https://doi.org/10.1093/jhered/esac020

Sim SB, Corpuz RL, Simmonds TJ, Geib SM. HiFiAdapterFilt, a memory efficient read processing pipeline, prevents occurrence of adapter sequence in PacBio HiFi reads and their negative impacts on genome assembly. *BMC Genomics*. 2022;**23**:157. https://doi.org/10.1186/s12864-022-08375-1

Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Teisher JK, Clark LG, Barberá P, Gillespie LJ, Zuloaga FO. A worldwide phylogenetic classification of the Poaceae (Gramineae) II: an update and a comparison of two 2015 classifications. *J Syst Evol*. 2017;**55**:259–290. https://doi.org/10.1111/jse.12262

Tillich M, Lehwark P, Pellizzer T, Ulbricht -Jones ES, Fischer A, Bock R, Greiner S. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res*. 2017;**45**:W6–W11. https://doi.org/10.1093/nar/gkx391

USFWS. U.S. Fish and Wildlife Service. *Recovery plan for vernal pool ecosystems of California and southern Oregon*. Portland, Oregon;

2005: Retrieved August 15, 2025, from https://www.fws.gov/node/4746436.

USFWS. U.S. Fish and Wildlife Service. Technical corrections for seven species on the list of endangered and threatened wildlife and six species on the list of endangered and threatened plants. *Fed Regist*. 2024;**89**:73308–73311. Retrieved 2025 August 15, from https://www.fws.gov/species-publication-action/technical-corrections-seven-species-list-endangered-and-threatened-0

Wilson MD, Riemer C, Martindale DW, Schnupf P, Boright AP, Cheung TL, Hardy DM, Schwartz S, Scherer SW, Tsui LC, et al. Comparative analysis of the gene -dense ACHE/TFR2 region on human chromosome 7q22 with the orthologous region on mouse chromosome 5. *Nucleic Acids Res*. 2001;**29**:1352–1365. https://doi.org/10.1093/nar/29.6.1352

Witham C, Holland R, Vollmar J. *Changes in the distribution of Great Valley vernal pool habitats from 2005 to 2012*. Sacramento, California: Report prepared for US Fish and Wildlife Service; 2014.

Workman R, Fedak R, Kilburn D, Hao S, Liu K, Timp W. High molecular weight DNA. *Extraction from recalcitrant plant species for third generation sequencing*. Research Square. 2018. https://doi.org/10.1038/protex.2018.059

Zhou C, Brown M, Blaxter M, McCarthy SA, Durbin R. (2025). Oatk: a de novo assembly tool for complex plant organelle genomes. *Genome Biology*, **26**, 235. https://doi.org/10.1186/s13059-025-03676-6

Journal of Heredity, 2026, 117, 526–536
https://doi.org/10.1093/jhered/esaf084
Genome Resources