Journal of Heredity, 2022, **113**, 706–711
https://doi.org/10.1093/jhered/esac051
Advance access publication 9 September 2022
Genome Resources

American Genetic Association

OXFORD

Genome Resources

The reference genome of the Vernal Pool Tadpole Shrimp, *Lepidurus packardi*

Shannon Rose Kieran Blair[1], Joshua Hull[2], Merly Escalona[3], Amanda Finger[1], Shannon E.K. Joslin[4], Ruta Sahasrabudhe[5], Mohan P.A. Marimuthu[5], Oanh Nguyen[5], Noravit Chumchim[5], Emily Reister Morris[6], Samantha Velazquez[6], Andrea Schreier[1]

[1]Genomic Variation Laboratory, Department of Animal Science, University of California, Davis, Davis, CA, United States,
[2]U.S. Fish and Wildlife Service, Sacramento Fish and Wildlife Office, Sacramento, CA, United States,
[3]Department of Biomolecular Engineering, University of California, Santa Cruz, Santa Cruz, CA, United States,
[4]U.S. National Park Service, Yosemite National Park, El Portal, CA, United States,
[5]DNA Technologies and Expression Analysis Cores, UC Davis Genome Center, University of California, Davis, Davis, CA, United States,
[6]Phase Genomics, Seattle, WA, United States

Address correspondence to S.R. Kieran Blair at the address above, or e-mail: sckieran@ucdavis.edu.

Corresponding Editor: Warren Booth

Abstract

In this paper, we report on the scaffold-level assembled genome for the federally endangered, California endemic crustacean *Lepidurus packardi* (the Vernal Pool Tadpole Shrimp). *L. packardi* is a key food source for other conserved California species including the California Tiger Salamander *Ambystoma californiense*. It faces significant habitat loss and fragmentation as vernal pools are threatened by urbanization, agricultural conversion, and climate change. This resource represents the first scaffold-level genome of any *Lepidurus* species. The assembled genome spans 108.6 Mbps, with 6 chromosome-length scaffolds comprising 71% of total genomic length and 444 total contigs. The BUSCO score for this genome is 97.3%, suggesting a high level of completeness. We produced a predicted gene set for this species trained on the *Daphnia magna* set of genes and predicted 17,650 genes. These tools can aid researchers in understanding the evolution and adaptive potential of alternative reproductive modes within this species.

Key words: branchiopod, California Conservation Genomics Project, CCGP, notostraca, triops

Introduction

The Vernal Pool Tadpole Shrimp (*Lepidurus packardi*) (Simon, 1886) (phylum: Crustacea, order: Branchiopoda, class: Notostraca, family: Triopsidae) is a freshwater microcrustacean. It is an ephemeral wetland specialist, occupying vernal pools, swales, and playas between Kern and Shasta Counties in California's Great Central Valley. It is California's only endemic notostracan (Rogers 2001). *L. packardi* is an important food source for the larval California Tiger Salamander (*Ambystoma californiense*) and has been found to comprise >90% of larval salamander diets when available (Messerman et al. 2021). It also feeds migratory waterfowl which occupy the pools during the wet season. It is an ecosystem engineer, creating bioturbation by burrowing and digging in vernal pool substrate (Croel and Kneitel 2011). It was protected under the Endangered Species Act in 1994 (United States Fish and Wildlife Service 1994).

The overwhelming majority of California's vernal pools have been lost over the past 3 centuries, and the habitat which remains is threatened by anthropogenic pressures including climate change, land conversion, and urbanization (AECOM 2009; Holland 2009). *L. packardi* was listed as federally endangered in 1994. Despite its protection, little is known about its biology, population genetics, or evolutionary history. A 2012 study using AFLPs found significant isolation by distance at small spatial scales (Aguilar 2012). A 2020 U.S. Fish and Wildlife Report (Kieran and Finger 2020) performed range-wide RAD-sequencing and found similarly high genetic differentiation between populations at small spatial scales, and low genetic diversity compared to other vernal pool crustaceans. Additionally, it is believed that populations of this species may possess alternative reproductive modes (fully bisexual, fully hermaphroditic, and mixed), but the extent, geographic variation and genetic basis of this variation is unknown. Understanding how populations differ across the genome is key to carrying out recovery actions such as introductions, translocations, and genetic rescue, and therefore the species has been included as part of the California Conservation Genomics Project (CCGP) (Shaffer et al. 2022).

Two related *Lepidurus* species were recently sequenced at the contig level, these are *Lepidurus arcticus* and *Lepidurus apus lubbocki* (Savojardo et al. 2019). Together, these 3

Received August 16, 2022; Accepted September 7, 2022

species have non-overlapping ranges and an estimated divergence time of 65 million years (Mathers et al. 2013). Comparative genomics using these assemblies, combined with the high-quality, scaffold-level resolution of the genome presented here, have the potential to shed new light on the evolution and adaptation of these so-called "living fossils" (Fig. 1).

Methods

Biological materials

Live specimens were collected from the Jepson Prairie Preserve (38.274939, −121.823922) in Solano County on 19 February 2021 under Federal 10(A)1(a) collection permit TE-28101C-0. Specimens were collected via dipnet at transported live to the lab where they were immediately frozen in liquid nitrogen and held at −80 °C until extraction.

Nucleic acid library preparation

High molecular weight (HMW) genomic DNA (gDNA) extraction and nucleic acid library preparation were carried out by the University of California Davis DNA Technologies Core (Davis, CA). DNA was extracted from 50 mg of whole-body tissue using Nanobind tissue big DNA kit (Circulomics, Baltimore, MD; Cat. # SKU NB-900-701-01) following the manufacturer's guidelines. Extracted DNA was cleaned with equal volumes of phenol/chloroform using phase-lock gels (Quantabio, Beverley, MA; Cat. #2302830) and precipitated by adding 0.4× volume of 5 M ammonium acetate and 3× volume of ice-cold ethanol. The DNA pellet was washed twice with 70% ethanol and resuspended in an elution buffer (10 mM Tris, pH 8.0). The purity of the DNA was accessed using NanoDrop spectrophotometer (260/280 and 260/230 ratios) and the integrity of the HMW gDNA was verified on a Femto pulse system (Agilent Technologies, Santa Clara, CA).

Fig. 1. (A) Photo of adult hermaphrodite *Lepidurus packardi* from Solano County, CA. Photo courtesy Shannon Kieran Blair. (B) A vernal pool in Solano County, CA, hosting *L. packardi*.

DNA sequencing and genome assembly

The HiFi SMRTbell libraries were constructed using the SMRTbell Express Template Prep Kit v2.0 (PacBio, Cat. #100-938-900) according to the manufacturer's instructions. HMW gDNA was sheared to a target DNA size distribution between 12 and 20 kb. For library preparation input, the sheared gDNA was concentrated using 1.8× of AMPure PB beads (Pacific Biosciences—PacBio, Menlo Park, CA; Cat. #100-265-900) for the removal of single-strand overhangs at 37 °C for 15 min, followed by further enzymatic steps of DNA damage repair at 37 °C for 30 min, end repair and A-tailing at 20 °C for 10 min and 65 °C for 30 min, and ligation of overhang adapter v3 at 20 °C for 60 min. The SMRTbell libraries were purified and concentrated with 0.8× AMPure PB beads for size selection with 40% diluted AMPure PB beads to remove short SMRTbell templates, <3 kb. The 15 to 17 kb average HiFi SMRTbell libraries were sequenced on 8M SMRT cells (1 per library), Sequel II sequencing chemistry 2.0, and 30-h movies each at UC Davis DNA Technologies Core (Davis, CA) on a PacBio Sequel II sequencer.

Initial contig assembly

PacBio Hifi Reads were assembled into contigs using the PacBio "ipa" software program v. 1.3.1 with default parameters.

Proximo Hi-C sequencing and scaffolded assembly

Chromatin conformation capture data were generated using a Phase Genomics (Seattle, WA) Proximo Hi-C 4.0 Kit, which is a commercially available version of the Hi-C protocol (Lieberman-Aiden et al. 2009). Following the manufacturer's instructions for the kit, intact cells were crosslinked using a formaldehyde solution, digested using the DPNII, DDE1, HINF, and MSEI restriction enzymes, end repaired with biotinylated nucleotides, and proximity ligated to create chimeric molecules composed of fragments from different regions of the genome that were physically proximal in vivo, but not necessarily genomically proximal. Continuing with the manufacturer's protocol, molecules were pulled down with streptavidin beads and processed into an Illumina-compatible sequencing library. Sequencing was performed on an Illumina NovaSeq (San Diego, CA). Reads were aligned to the draft assembly also following the manufacturer's recommendations. Briefly, reads were aligned using BWA-MEM (Li and Durbin 2010) with the -5SP and -t 8 options specified, and all other options default. SAMBLASTER (Faust and Hall 2014) was used to flag PCR duplicates, which were later excluded from analysis. Alignments were then filtered with samtools (Li et al. 2009) using the -F 2304 filtering flag to remove non-primary and secondary alignments. Putative mis-joined contigs were broken using Juicebox (Rao et al. 2014; Durand et al. 2016) based on the Hi-C alignments.

Phase Genomics's Proximo Hi-C genome scaffolding platform was used to create chromosome-scale scaffolds from the corrected assembly as described in Bickhart et al. (2017). As in the LACHESIS method (Burton et al. 2013), this process computes a contact frequency matrix from the aligned Hi-C read pairs, normalized by the number of restriction enzyme cut sites on each contig, and constructs scaffolds in such a way as to optimize expected contact frequency and other statistical patterns in Hi-C data. Approximately 40,000 separate Proximo runs were performed to optimize the number of scaffolds and scaffold construction in order to make the scaffolds as concordant with the observed Hi-C data as possible. Finally, Juicebox was again used to correct scaffolding errors.

Assembly metrics and validation

The assembly completeness was estimated by running BUSCO (Waterhouse et al. 2018) version 5.2.2 in genome mode using the arthropoda_odb10 database. Assembly statistics were calculated using genometools (Gremme et al. 2013) version 1.5.9 and QUAST (Mikheenko et al. 2018) version 5.0.2. Further quality assessment was carried out following the frameshift pipeline described in Korlach et al. (2017).

Ab initio gene prediction

We performed ab initio gene prediction following the method of Savojardo et al. (2019). Briefly, we used RepeatModeler v1.0.11 (Smit and Hubley 2008-2015) to identify repeat content and create transposable element (TE) libraries. Repeats were masked using RepeatMasker v.4.1.2 (Smit et al. 2013-2015) and gene prediction was performed using Augustus v3.3.3 (Stanke et al. 2006) with default parameters, trained on the *Daphnia magna* set of annotated genes (NCBI accession GCA_001632505.1, accessed May 2022) (Table 1).

Results

Genome assembly

Proximo Hi-C Illumina Novaseq Sequencing generated a total of 90,119,568 PE150 read pairs. Juicebox contig breaking introduced a total of 9 breaks in 9 contigs, and the same alignment procedure was repeated from the beginning on the resulting corrected assembly. The Proximo Hi-C Scaffolding pipeline resulted in a set of 6 chromosome-scale scaffolds containing 79211423 Mbp of sequence (72.91% of the corrected assembly).

The final genome is 108,645,433 base pairs (108.6 Mbp). This is in line with what has been found in other notostracan taxa (*L. arcticus*, 73.2 Mbp, *L. apus lubbocki*, 90.3 Mbp) (Savojardo et al. 2019). Genomescope (Vurture et al. 2017) estimated the haploid length at 81,812,296 bp based on a 21mer spectrum produced by Jellyfish (Marçais and Kingsford 2011), also suggesting that the recovered genome size is appropriate. The N50 length is 12,481,572 bp. The scaffold L50 is 5. The genome was assembled into 6 chromosome-length scaffolds and 349 unscaffolded contigs. Because these scaffolds have not yet been assigned to chromosomes, the NCBI database reports 355 scaffolds (all scaffolds and unscaffolded contigs)

Table 1. Assembly pipeline and software used.

Assembly	Software	Version
Kmer counting	Jellyfish	2.2.6
Estimation of genome size and heterozygosity	GenomeScope	2
De novo assembly (contigging)	ipa	1.3.1
Long read, genome–genome alignment	Minimap2	2.22
Scaffolding		
HiC mapping	Phase Genomics Proximo HiC pipeline https://phasegenomics.github.io/2019/09/19/hic-alignment-and-qc.html	Commit 5f9d55ea3162f8d2 1988f486b5d012f0800abdc4
HiC scaffolding	Juicebox	2
HiC contact map generation		
Short-read alignment	BWA-MEM	0.7.17-r1188
SAM/BAM processing	SAMBLASTER	1.11
SAM/BAM filtering	samtools	0.3.0
Matrix generation and balancing	Phase Genomics Proximo HiC Pipeline	Commit 5f9d55ea3162f8d2 1988f486b5d012f0800abdc4
Benchmarking		
Basic assembly stats	QUAST	5.0.2
	GenomeTools	1.5.9
Assembly completeness	BUSCO	5.2.2
	Merqury	1.3
	Blobtoolkit	3.1.6
Repeat analysis	RepeatModeler	1.0.11
	RepeatMasker	4.1.2
Gene prediction	Augustus	3.3.3

Software citations are listed in the text.

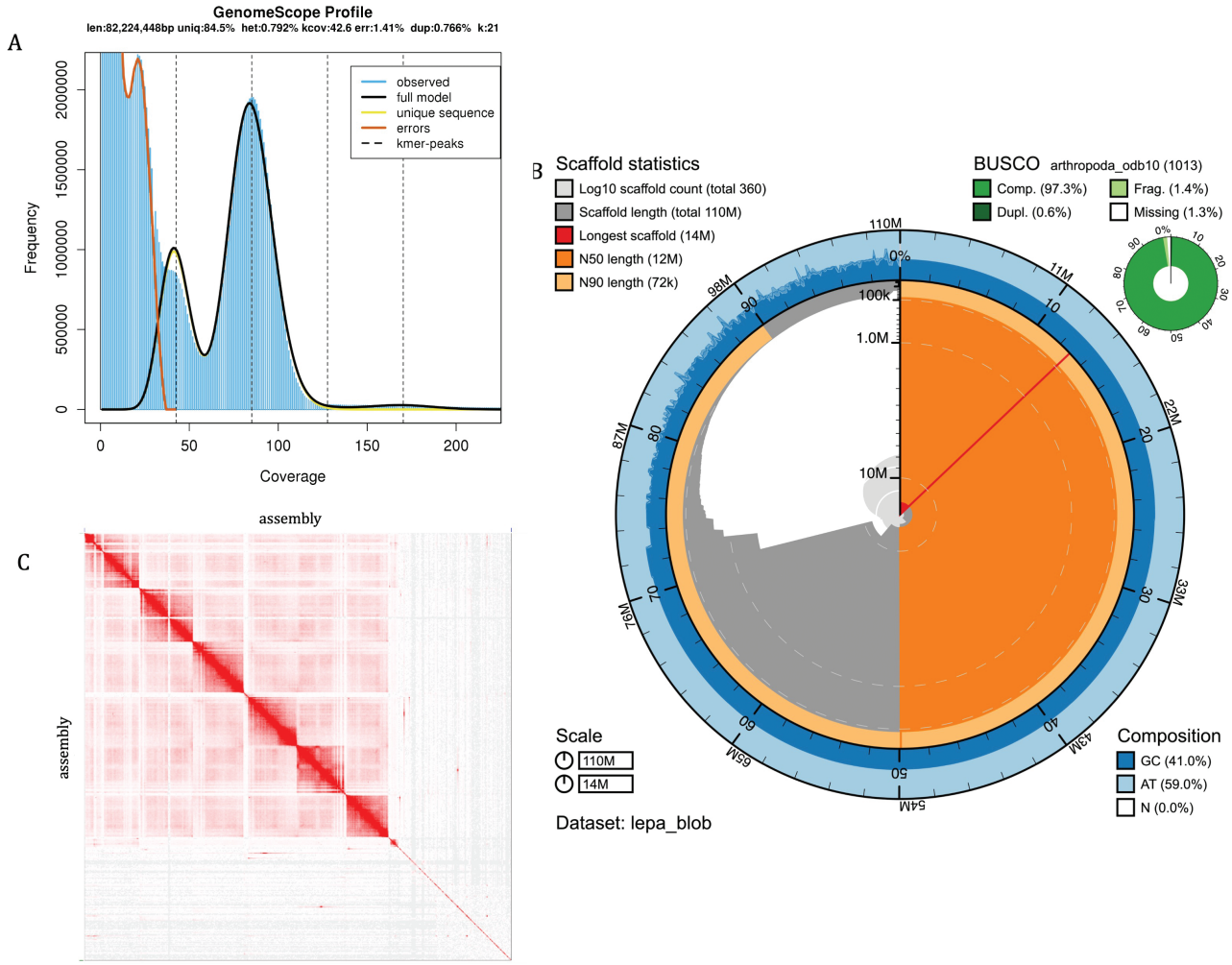

Fig. 2. (A) Genomescope k-mer spectrum of *Lepidurus packardi* sequence reads. (B) BlobToolKit snail plot showing N50 metrics for *L. packardi* assembly and BUSCO scores for the Arthopoda set of orthologs. (C) Contact map of *L. packardi* assembly. This map demonstrates the ~70% of sequence length contained in chromosome-level scaffolds, as well as the remaining sequence length clustered into contigs.

and 444 contigs (all scaffolds and unscaffolded contigs after spanning 89 gaps). 72.91% of all sequence length is contained in the 6 chromosome-length scaffolds. The longest scaffold is 14,048,704 bp and the scaffold N50 is 12,481,803 bp. After gap-splitting, the NCBI contig N50 is 1,298,445 bp. The BUSCO score for the completed assembly is 97.3% (96.7% complete and unduplicated, 0.6% duplicated, 1.4% fragmented, and 1.3% missing, n = 1013). GC content is 40.9% (Fig. 2).

Genomic repeat analysis

We identified 672 interspersed repeats. Based on RepeatMasker analysis, interspersed repeats made up 26.6% of the genome, nearly twice the other sequenced *Lepidurus* species. RepeatModeler was unable to classify 59.7% (401) into families. See Table 2 for the breakdown of gene families identified by RepeatModeler. Our Augustus gene model predicts 17,650 genes for *L. packardi*, slightly higher than predicted numbers for *L. arcticus* (10,718) and *L. apus lubbocki* (16,383) (Table 3).

Discussion

This resource has the potential to shed light on the understudied endangered *L. packardi*. Unexplored questions about sex determination and rumored variable reproductive mode can be answered using genomic tools and whole genome resequencing. Understanding the variation and genetic bases of reproduction in this species is vital before recovery actions such as genetic rescue or translocations can be carried out. Furthermore, a deep exploration of the genetic variation of this species across the landscape will help conservationists support the recovery of this species, which in turn will help support the species richness of crustaceans in vernal pools.

The existence of genome-level resources for the non-California congeners *L. arcticus* and *L. apus lubbocki* will allow researchers to compare the divergence, adaptation, and sex determination systems of these species. This is an unusual richness of resources for branchiopod crustaceans, which are generally restricted to 1 or fewer reference genomes per genus. Branchiopod crustaceans are ancient lineages with deep interspecific and intergeneric divergence times, so the "closest available" genome is often 50 mya or more diverged from the target species. This will be a useful genomic resource for both targeted conservation and broader comparative crustacean research.

Funding

This work was funded by the U.S. Bureau of Reclamation [Grant #R20AP00037]. This work was supported by the

Table 2. Sequencing and assembly statistics, and accession numbers.

BioProjects and vouchers	CCGP NCBI Bio-project	PRJNA720569 http://www.ncbi.nlm.nih.gov/bioproject/PRJNA720569
	Lepidurus packardi NCBI Bio-project	PRJNA811174 https://www.ncbi.nlm.nih.gov/bioproject/PRJNA811174
	NCBI Bio-sample	SAMN26264343 https://www.ncbi.nlm.nih.gov/biosample/SAMN26264343
Genome sequence	PacBio HiFi long read runs	1 PACBIO_SMRT (Sequel II) run: 1.1M spots, 11.5G bases, 2.8 Gb downloads
	Proximo HiC Illumina sequencing	1 Illumina NovaSeq 6000 run: 65.3M spots, 19.6G bases, 6.1 Gb downloads
	PacBio HiFi NCBI SRA accession	SRX15225418 https://www.ncbi.nlm.nih.gov/sra/SRX15225418
	Proximo HiC Illumina NCBI SRA accession	SRX15225419 https://www.ncbi.nlm.nih.gov/sra/SRX15225419
	HiFi read coverage	109×
	Number of contigs	355
	Contig N50 (bp)	1,298,445
	Longest contig	14,048,704
	Number of scaffolds	444
	Scaffolds assigned to chromosomes	6
	Scaffold N50 (bp)	12,481,803
	Size of final assembly (bp)	108,645,424
	Gaps per Gbp	890
	NCBI Genome Assembly Accession	GCA_023053545.1 https://www.ncbi.nlm.nih.gov/assembly/GCA_023053545.1
Assembly quality	Base pair QV (merqury)	49.8
	Indel QV (frameshift analysis)	48.2
	K-mer completeness	71.1%
	BUSCO completeness (C:S:D:F:M)	97.30%:96.70%:0.6%:1.40%:01.30%

Table 3. Repeat families found in *L. packardi* assembly.

Gene family	% *L. packardi* genome
SINE	0.20%
RC/helicase	0.20%
LINE	2.29%
LTR	3.56%
DNA	3.92%
Unclassified	15.86%
Total	26.02%

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].

Acknowledgments

PacBio Sequel II library prep and sequencing were carried out at the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center, supported by NIH Shared Instrumentation Grant 1S10OD010786-01. Proximo Hi-C sequencing was carried out by Phase Genomics, Seattle, WA. We thank the staff at the UC Davis DNA Technologies and Expression Analysis Cores for their diligence and dedication to generating high quality sequence data. This work was performed in part at the University of California Natural Reserve System (Jepson Prairie Preserve) Reserve DOI: 10.21973/N3D082. We especially wish to acknowledge Virginia "Shorty" Boucher for her thoughtful and helpful assistance accessing the sampling site.

Data availability

Data generated for this study are available under NCBI BioProject PRJNA811174. Raw sequencing data for sample LEPA_1 (NCBI BioSample SAMN26264343) are deposited in the NCBI Short Read Archive (SRA) under SRR19158969.

References

AECOM. *Loss of Central Valley Vernal Pools Land Conversion, Mitigation Requirements, and Preserve Effectiveness Summary Report*. 2009. Prepared for the Placer Land Trust by AECOM. p. 16. http://www.placerlandtrust.org/uploads/documents/Vernal%20Pool%20Studies%20Report/VP%20Summary%20Report_Final.pdf.

Aguilar A. Range-wide and local drivers of genetic structure in an endangered California vernal pool endemic crustacean. *Conserv Genet*. 2012;13(6):1577–1588.

Bickhart DM, Rosen BD, Koren S, Sayre BL, Hastie AR, Chan S, Lee J, Lam ET, Liachko I, Sullivan ST, et al. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nat Genet*. 2017;49(4):643–650. doi:10.1038/ng.3802

Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol*. 2013;31(12):1119–1125. doi:10.1038/nbt.2727

Croel RC, Kneitel JM. Ecosystem-level effects of bioturbation by the tadpole shrimp *Lepidurus packardi* in temporary pond mesocosms. *Hydrobiologia*. 2011;665:161–181.

Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst*. 2016;3(1):99–101. doi:10.1016/j.cels.2015.07.012

Faust GG, Hall IM. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics*. 2014;30(17):2503–2505. doi:10.1093/bioinformatics/btu314

Gremme G, Steinbiss S, Kurtz S. GenomeTools: a comprehensive software library for efficient processing of structured genome annotations. *IEEE/ACM Trans Comput Biol Bioinf*. 2013;10(3):645–656.

Holland RF. California's Great Valley vernal pool habitat status and loss: rephotorevised 2005. 2009. p. 1–23. papers2://publication/uuid/F2C30E9C-9A68-451A-8972-742C1443ADB6. Accessed July 21, 2021.

Kieran SRC, Finger AJ. *Final report for Cesu R15AC00525: comparative population genetics across vernal pool branchiopod species reveals incongruous patterns of geographic structuring, genetic differentiation*. Prepared for the United States Fish and Wildlife Service. Sacramento, CA. 2020. p. 76.

Korlach J, Gedman G, Kingan SB, Chin CS, Howard JT, Audet JN, 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(10):1–16.

Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589–595. doi:10.1093/bioinformatics/btp698

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map (SAM) format and SAMtools 1000 Genome Project Data Processing Subgroup. *Bioinformatics*. 2009;25(16):2078–2079.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2013;326(5950):289–293. doi:10.1126/science.1181369

Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*. 2011;27(6):764–770.

Mathers TC, Hammond RL, Jenner RA, Hänfling B, Gómez A. Multiple global radiations in tadpole shrimps challenge the concept of "living fossils". *PeerJ*. 2013;2013(1):e62.

Messerman AF, Clause AG, Catania SVL, Shaffer HB, Searcy CA. Coexistence within an endangered predator–prey community in California vernal pools. *Freshw Biol*. 2021;66(7):1296–1310.

Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*. 2018;34(13):i142–i150.

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665–1680.

Rogers DC. Revision of the nearctic *Lepidurus* (Notostraca). *J Crustac Biol*. 2001;21(4):991–1006.

Savojardo C, Luchetti A, Martelli PL, Casadio R, Mantovani B. Draft genomes and genomic divergence of two *Lepidurus* tadpole shrimp species (Crustacea, Branchiopoda, Notostraca). *Mol Ecol Resour*. 2019;19(1):235–244.

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.

Smit A.F.A., Hubley R. *RepeatModeler Open-1.0*. 2008-2015. http://www.repeatmasker.org.

Smit AFA, Hubley R, Green P. *RepeatMasker Open-4.0*. 2013-2015. http://www.repeatmasker.org.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res*. 2006;34(Web Server issue):W435.

United States Fish and Wildlife Service. Endangered and threatened wildlife and plants; determination of endangered status for the conservancy fairy shrimp, longhorn fairy shrimp, and the vernal pool tadpole shrimp; and threatened status for the vernal pool fairy shrimp. *Fed Reg*. 1994;59(180):48136.

Vurture GW, Sedlazeck FJ, Nattestad M, Underwood CJ, Fang H, Gurtowski J, Schatz MC. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*. 2017;33(14):2202–2204.

Waterhouse RM, Seppey M, Simao FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva E, Zdobnov EM. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol*. 2018;35(3):543–548. doi:10.1093/molbev/msx319