Journal of Heredity, 2026, **117**, 450–458
https://doi.org/10.1093/jhered/esaf074
Advance access publication 3 October 2025
Genome Resources

American Genetic Association
OXFORD

Genome Resources

A high-quality reference genome for the ornate shrew (*Sorex ornatus*)

Isaac J. Linn[1,2,*], Merly Escalona[3], Oanh Nguyen[4], Mohan P.A. Marimuthu[4], Noravit Chumchim[4], Colin W. Fairbairn[5], William Seligmann[5], Eric Beraut[5], Christopher J. Conroy[1], Rauri C.K. Bowie[1,2], Michael W. Nachman[1,2]

[1]Museum of Vertebrate Zoology, University of California Berkeley, Berkeley, CA, United States, [2]Department of Integrative Biology, University of California Berkeley, Berkeley, CA, United States, [3]Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA, United States, [4]DNA Technologies and Expression Analysis Core Laboratory, Genome Center, University of California Davis, Davis, CA, United States, [5]Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, United States

*Corresponding author: Museum of Vertebrate Zoology, University of California Berkeley, 3101 Valley Life Sciences Building, Berkeley, CA 94720, CA, United States. Email: ijlinn@berkeley.edu

Abstract

The ornate shrew (*Sorex ornatus*) is a small predatory mammal with a broad distribution in northern, central, and southern California as well as in Baja California, Mexico. The ornate shrew is a highly productive consumer in wetland environments and is known to hybridize with the closely related vagrant shrew, *Sorex vagrans*. Here we present a high-quality de novo genome assembly for *S. ornatus* generated as a part of the California Conservation Genomics Project. The *S. ornatus* genome was generated using PacBio HiFi long reads and Omni-C chromatin interaction sequencing. The primary assembly is highly contiguous, with a contig N50 of 15.9 Mb, a scaffold N50 of 115.2 Mb, and a BUSCO completeness score of 95.10%. The ornate shrew genome will serve as a valuable resource for future North American *Sorex* conservation genomics as well as for research into shrew biology more generally.

Key words: California Conservation Genomics Project, CCGP, conservation, metabolic rate, salt marshes, Soricidae

Introduction

Long tailed shrews (*Sorex*) are small insectivorous mammals found primarily in temperate and boreal regions across Europe, Asia, and North America. The genus *Sorex* includes 88 recognized species found in a variety of habitats including woodlands, grasslands, wetlands, and alpine meadows (Mammal Diversity Database 2025). They are known for their high metabolic rates which can be up to 366% of the expected metabolic rate for a mammal of the same size (Morrison and Pearson 1946; Pearson 1948; Newman and Rudd 1978; Taylor 1998; Ochocińska and Taylor 2005).

The ornate shrew (*Sorex ornatus*) is broadly distributed across northern, central, and southern California west of the Sierra Nevada, and into Baja California, Mexico (Fig. 1) (Hall 1981; Owen and Hoffmann 1983; IUCN 2016). It primarily inhabits saline and brackish marshes, or riparian and palustrine habitats, though the species can be found in drier grassland or chaparral habitat (Grinnell 1933; Rudd 1955; von Bloeker, JC Jr. 1967). *S. ornatus* is an important predator in salt marsh environments (Newman 1970), consuming primarily invertebrates and requiring an ample supply of insect prey to maintain a constant high metabolic rate. Marsh and wetland environments provide food and dense ground cover, but these environments have been disturbed by human development and pollution for over the past 100 years (The Nature Conservancy 2011; Beagle et al. 2015; Benham and Bowie 2021). As populations continue to be threatened by human interactions, genomic resources for ornate shrews in California will be essential for understanding levels and patterns of genetic variation and thereby informing management decisions (Beninde et al. 2022; Fiedler et al. 2022).

While the ornate shrew has a broad distribution, the Buena Vista Lake ornate shrew, *S. o. relictus*, is endangered, and four other subspecies are listed as species of special concern by the California Department of Fish and Wildlife (Monterey shrew, *S. o. salarius*; southern California saltmarsh shrew, *S. o. salicornicus*; Suisun shrew, *S. o. sinuosus*; and Santa Catalina shrew, *S. o. willetti*; California Department of Fish and Wildlife 2025). The Suisun shrew (*S. o. sinuosus*) likely forms a hybrid zone with the closely related vagrant shrew (*Sorex vagrans*) though the extent of hybridization remains uncertain (Rudd 1955; Brown and Rudd 1981; Patton 2019). Natural hybridization can be both beneficial and detrimental to threatened species (e.g. Wilder et al. 2025), so characterizing this hybrid zone may be necessary to make an informed

Received on 19 May 2025; revised on 19 September 2025; accepted on 29 September 2025
© The Author(s) 2025. Published by Oxford University Press on behalf of The American Genetic Association.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

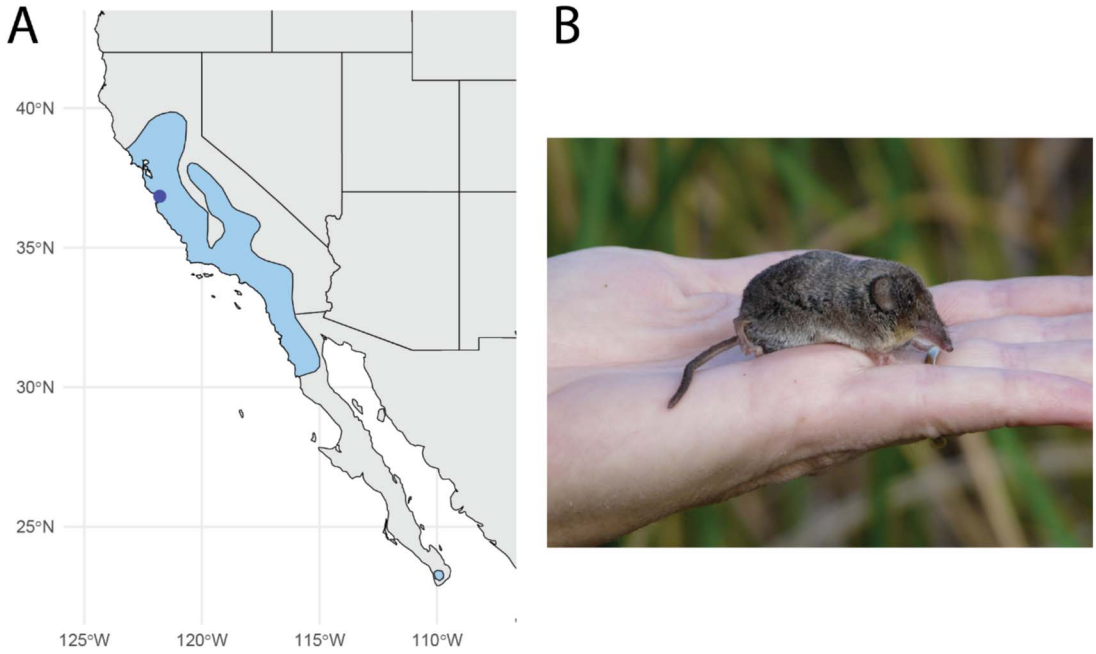

Fig. 1. (A) The ornate shrew (*Sorex ornatus)* is distributed across northern, central, and southern California west of the Sierra Nevada, and into Baja California, Mexico. Dot indicates the collection location for the *S. ornatus* individual from which the genome in this assembly was generated. Range data retrieved from IUCN. (B) an image of an endangered Buena Vista Lake shrew (*S. o. Relictus)* (photo by CSU Stanislaus, published by U.S. Fish and Wildlife Service, Pacific southwest region).

assessment of extinction risk or to take conservation action (Wayne and Shaffer 2016; Chan et al. 2019).

Here we present a whole-genome assembly for the ornate shrew, *Sorex ornatus salarius*, a subspecies of special concern, as part of the California Conservation Genomics Project (CCGP) (Shaffer et al. 2022). This reference genome will enable us to better understand the biogeography of North American shrews of conservation concern, examine the extent of gene flow between the Suisun shrew and the Vagrant shrew, and serve as a reference for studies on shrew biology more generally.

Methods

Biological materials

A female *S. ornatus salarius* was captured at McClusky Slough, along the coast in north Monterey County, California (36.84108465, −121.789764) on 10 March 2021 by Mark Allaback and Chad Steiner. The specimen was prepared by CJC as a museum study skin, skull, fluid preserved carcass, and tissues were frozen at −80°C for genomic and transcriptomic use as a part of the CCGP. The specimen was deposited as a vouchered specimen at the Museum of Vertebrate Zoology, Berkeley, California (MVZ: Mamm240319, CJC3344). For more details, visit https://arctos.database.museum/guid/MVZ:Mamm:240319.

We extracted high molecular weight (HMW) genomic DNA (gDNA) from 38 mg of liver tissue using the Nanobind Tissue Big DNA kit following the manufacturer's instructions (Pacific BioSciences—PacBio, Menlo Park, CA). Subsequently, the extracted HMW DNA underwent additional purification through the phenol-chloroform method (PacBio). DNA purity was estimated using absorbance ratios (260/280 = 1.84 and 260/230 = 2.40) on a NanoDrop ND-1000 spectrophotometer, and the yield (7.3 μg) was quantified using the Quantus Fluorometer with the QuantiFluor ONE dsDNA Dye assay (Promega, Madison, WI). Size distribution analysis of the HMW DNA was performed utilizing the Femto Pulse system (Agilent, Santa Clara, CA), revealing that 82% of the fragments were 75 kilobases (kb) or longer.

Nucleic acid library preparation

The HiFi SMRTbell libraries were constructed using the SMRTbell Express Template Prep Kit v2.0 (PacBio, Cat. #100–938-900) and SMRTbell prep kit 3.0 (PacBio, Cat. #102–182-700) according to the manufacturer's instructions. HMW gDNA was sheared to a target DNA size distribution between 15 kb and 18 kb using Diagenode's Megaruptor 3 system (Diagenode, Belgium; cat. B06010003). The sheared gDNA was concentrated by AMPure PB beads (PacBio, Cat. #100–265-900) for multiple enzymatic reactions such as removal of single-strand overhangs, DNA damage repair, end-repair/A-tailing, overhang adapter ligation, and nuclease treatment. The first two SMRTbell libraries were size selected using the BluePippin/PippinHT system (Sage Science, Beverly, MA; Cat #BLF7510/HPE7510) according to the v2.0 protocol to collect fragments >7–9 kb. The third SMRTbell library was size selected using 3.1X of 35% v/v diluted AMPure PB beads to progressively remove SMRTbell templates < 5 kb according to the SMRTbell prep kit 3.0 protocol. The HiFi SMRTbell libraries were sequenced at UC Davis DNA Technologies Core (Davis, CA) using nine SMRT cells (PacBio, Cat #101–389-001), Sequel IIe sequencing chemistry 2.0, and 30-hour movies each on a PacBio Sequel IIe sequencer.

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, specimen tissue (liver, ID: CJC3344) was thoroughly ground with a mortar and pestle while cooled with liquid nitrogen. Subsequently, chromatin was fixed in place in the

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

nucleus. The suspended chromatin solution was then passed through 100 μm and 40 μm cell strainers to remove large debris. Fixed chromatin was 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 an 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 the Vincent J. Coates Genomics Sequencing Lab (Berkeley, CA) on an Illumina NovaSeq 6000 platform (Illumina, CA) to generate ~100 million 2 × 150 bp read pairs per gigabase (Gb) of genome size.

Nuclear genome assembly

We assembled the genome of the *S. ornatus* specimen following the CCGP assembly pipeline, which uses PacBio HiFi reads and Omni-C data to produce high quality and highly contiguous genome assemblies. The pipeline is outlined in Supplementary Table 1 and lists the tools and non-default parameters used in the assembly process. First, we removed the remnant adapter sequences from the PacBio HiFi dataset using HiFiAdapterFilt (Sim et al. 2022) and generated the initial phased diploid assembly using HiFiasm (Cheng et al. 2021, Cheng et al. 2022, Cheng et al. 2024) on Hi-C mode, with the filtered PacBio HiFi reads and the Omni-C dataset. We aligned the Omni-C data to each assembly separately following the Arima Genomics Mapping Pipeline (https://github.com/ArimaGenomics/mapping_pipeline) and then scaffolded each assembly with SALSA (Ghurye et al. 2017; Ghurye et al. 2019).

Assemblies were manually curated by iteratively generating and visually analyzing their corresponding Omni-C contact maps. We generated two sets of contact maps, one to be visualized with HiGlass (Kerpedjiev et al. 2018), an interactive contact map viewer, and the second, with PretextView (https://github.com/wtsi-hpag/PretextView), which allows for interactive modifications of the contact maps.

To generate the contact maps for HiGlass, we aligned the Omni-C data with BWA-MEM (Li 2013), identified ligation junctions, and generated Omni-C pairs (Lee et al. 2022) using pairtools (Open2C et al. 2023). We generated a multi-resolution Omni-C matrix with cooler (Abdennur and Mirny 2020), balanced it with hicExplorer (Ramírez et al. 2018) and aggregated the data to be uploaded to the HiGlass viewer.

For the Pretext contact maps, we used the same Omni-C alignments as for the HiGlass contact maps and converted them into pretext format using PretextMap (https://github.com/wtsi-hpag/PretextMap), and visualized and modified the contact maps with PretextView (https://github.com/wtsi-hpag/PretextView).

By jointly visualizing and analyzing the contact maps from these two platforms we are able to properly modify the contact maps and later modify 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 PacBio HiFi reads and YAGCloser (https://github.com/merlyescalona/yagcloser). Finally, we checked for contamination using the BlobToolKit Framework (Challis et al. 2020).

Contact maps for the final version of the assemblies were generated as pretext contact maps and visualized with PretextSnapshot (https://github.com/wtsi-hpag/PretextSnapshot).

Genome quality assessment

We generated k-mer counts from the Omni-C 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. For 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 Mammalian ortholog database (mammalia_odb10) which contains 9226 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 the 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 = 54, known for the number of chromosomes for this species (Junge and Hoffmann 1981; Challis et al. 2023). Quality metrics for the notation were calculated on the primary haplotype assembly.

Collinearity analysis

We assessed synteny between *S. ornatus* and *Sorex fumeus* (GCA_029834395.1) using ntSyt (Coombe et al. 2024). We used Mash (Ondov et al. 2016) and estimated the mean sequence divergence between *S. ornatus* and *S. fumeus* to be 5.88%. Synteny blocks were estimated using ntSynt with the default parameters for taxa which are 1% to 10% diverged. Synteny blocks were then plotted using ntSynt-viz (Coombe et al. 2025) after removing scaffolds <5 Mb in length.

Genome annotation

We identified repetitive elements using the RepeatModeler2 LTR pipeline 2.0.7 (Flynn et al. 2020) and masked repetitive regions using RepeatMasker 4.1.7-p1 (Smit et al. 2013). Gene annotations were generated by using Liftoff 1.6.3 to map annotations from the *S. fumeus* assembly (GCA_029834395.1).

Mitochondrial genome assembly

We assembled the mitochondrial genome of *S. ornatus* from the PacBio HiFi reads using the reference-guided pipeline MitoHiFi (Uliano-Silva et al. 2023; Allio et al. 2020). The mitochondrial sequence of a *Sorex thibetanus* (NCBI:NC_064993.1) was used as the starting reference

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

Table 1. Metrics for the primary and alternate assemblies of the ornate shrew (*Sorex ornatus*) genome.

Metric	Primary[a]	Alternate
Number of contigs	983	676
Contig N50 (bp)	15 899 521	19 322 863
Contig NG50[b]	33 300 491	32 956 168
Longest contig	88 766 022	72 226 950
Number of Scaffolds	624	350
Scaffold N50	115 205 878	113 643 214
Scaffold NG50[b]	141 903 287	142 201 381
Largest scaffold	192 264 598	193 129 750
Phased block NG50[b]	33300.491	32 956 168
Indel QV (Frame shift)	41.13274692	40.01387523
Size of final assembly	2 814 783 814	2 788 702 286
Gaps per Gbp (#Gaps)	128(359)	117(326)
BUSCO		
Complete genes	95.1%	95.0%
Complete: single copy	93.2%	93.2%
Complete: duplicated	1.9%	1.8%
Fragmented	1.2%	1.2%
Missing	3.7%	3.7%
Base pair QV	64.1758	63.9719
k-mer completeness	88.067	88.0366

[a]The mitochondrial sequence was included in generating metrics for the primary assembly. For example, there were 623 scaffolds in the nuclear genome and 1 in the mitochondrial genome. [b]Read coverage and NGx statistics have been calculated based on the estimated genome size of 1.45 Gb.

sequence. After completion of the nuclear genome, we searched for matches of the resulting mitochondrial 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 mitochondrial assembly sequence.

Results

Sequencing data

The Omni-C library generated 492.77 million read pairs and the PacBio SMRTBell library generated 7.37 million HiFi reads. The PacBio HiFi reads yielded ~37X genome coverage and had an N50 read length of 15 447 bp; a minimum read length of 50 bp; a mean read length of 14 176 bp; and a maximum read length of 56 163 bp (see Supplementary Fig. 1 for read length distribution). Based on the Omni-C data, Genomescope 2.0 estimated a genome size of 2.7 Gb, a 0.836% heterozygosity rate and a 0.481% sequencing error rate. The k-mer spectrum shows a bimodal distribution with a major peak at ~36-fold and a minor peak at ~18-fold coverage, corresponding to homozygotes and heterozygotes, respectively (Fig. 2A).

Nuclear genome assembly

The final genome assembly (mSorOrn1) consists of two phased haplotypes; the assemblies have been tagged as primary and alternate based on our assessment of completeness and contiguity. The final haplotype assemblies are close in size, with a difference of ~26.08 Mb between haplotypes (Table 1).

The primary haplotype (mSorOrn1.0.p) consists of 623 scaffolds spanning 2.81 Gb with a contig N50 of 15.90 Mb, a scaffold N50 of 115.2 Mb, the largest contig size of 88.77 Mb, and the largest scaffold size of 192.26 Mb (Table 1). The alternate haplotype (mSorOrn1.0.a) consists of 350 scaffolds spanning 2.79 Gb with a contig N50 19.32 Mb, a scaffold N50 of 113.64 Mb, the largest contig size of 72.23 Mb and the largest scaffold size of 193.13 Mb.

The primary haplotype has a BUSCO completeness score for the Mammalian gene set of 95.1%, a base pair quality value (QV) of 64.18, a k-mer completeness of 88.07%, and a frameshift indel QV of 41.13. The alternate haplotype has a BUSCO completeness score for the same gene set of 95%, a base pair QV of 63.97, a k-mer completeness of 88.04%, and a frameshift indel QV of 40.01.

During manual curation we made a total of 294 joins (152 on the primary haplotype and 142 on the alternate haplotype) and 78 breaks (42 on the primary haplotype and 36 on the alternate haplotype) based on the signal from the Omni-C contact maps. We filtered out nine contigs corresponding to undefined sequences and viruses (Astashyn et al. 2024; details of the contamination screening is in Supplementary Table 3). In addition, we removed 44 contigs matching mitochondrial contaminants (22 per haplotype).

The Omni-C contact maps reveal a highly contiguous assembly, suggesting that the *S. ornatus* genome is organized in 27 chromosomes based on the number of major bins along the diagonal axis of the plot (Fig. 2C and D). Assembly statistics are reported in Table 1 and represented graphically in Fig. 2C. We also estimated a high level of synteny between this assembly and the PacBio assembly for *S. fumeus* (Supplementary Fig. 1), although we note that the large number of scaffolds for the *S. fumeus* assembly prevents a detailed assessment of synteny for some chromosomes. We have deposited the genome assembly in the NCBI GenBank (see Supplementary Table 2 and Data Availability for details).

Mitochondrial genome assembly

We assembled a mitochondrial genome with MitoHiFi. The final mitochondrial genome is circular and spans 17 106 bp. The base composition of the final assembly is A = 33.39%, C = 24.3%, G = 12.94%, T = 29.37%, and consists of 22 unique transfer RNAs, 13 protein coding genes and 2 rRNAs. Gene composition and order follow the animal mitochondrial genome structure (Boore 1999). We have deposited the mitochondrial genome assembly in NCBI GenBank (See Supplementary Table 2 and Data Availability for details).

Genome annotation

Through the RepeatModeler2 and RepeatMasker pipeline (Smit et al. 2015), we identified 43.07% of the genome as interspersed repeats, including 26.38% LINE elements, 1.03% LTR elements, 0.08% DNA transposons, 0.06% SINES and 15.53% as unclassified repeats. The high percentage of unclassified repeats and the low percentages of LTR elements, DNA transposons, and SINES suggest that many of these elements may be present but are not properly classified. Liftoff transferred the majority of annotations from *S. fumeus*, annotating 83.7% of genes and 96.1% of structural features.

Discussion

Here we present a high-quality reference genome for the ornate shrew, *S. ornatus*. The total length of the assembly

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

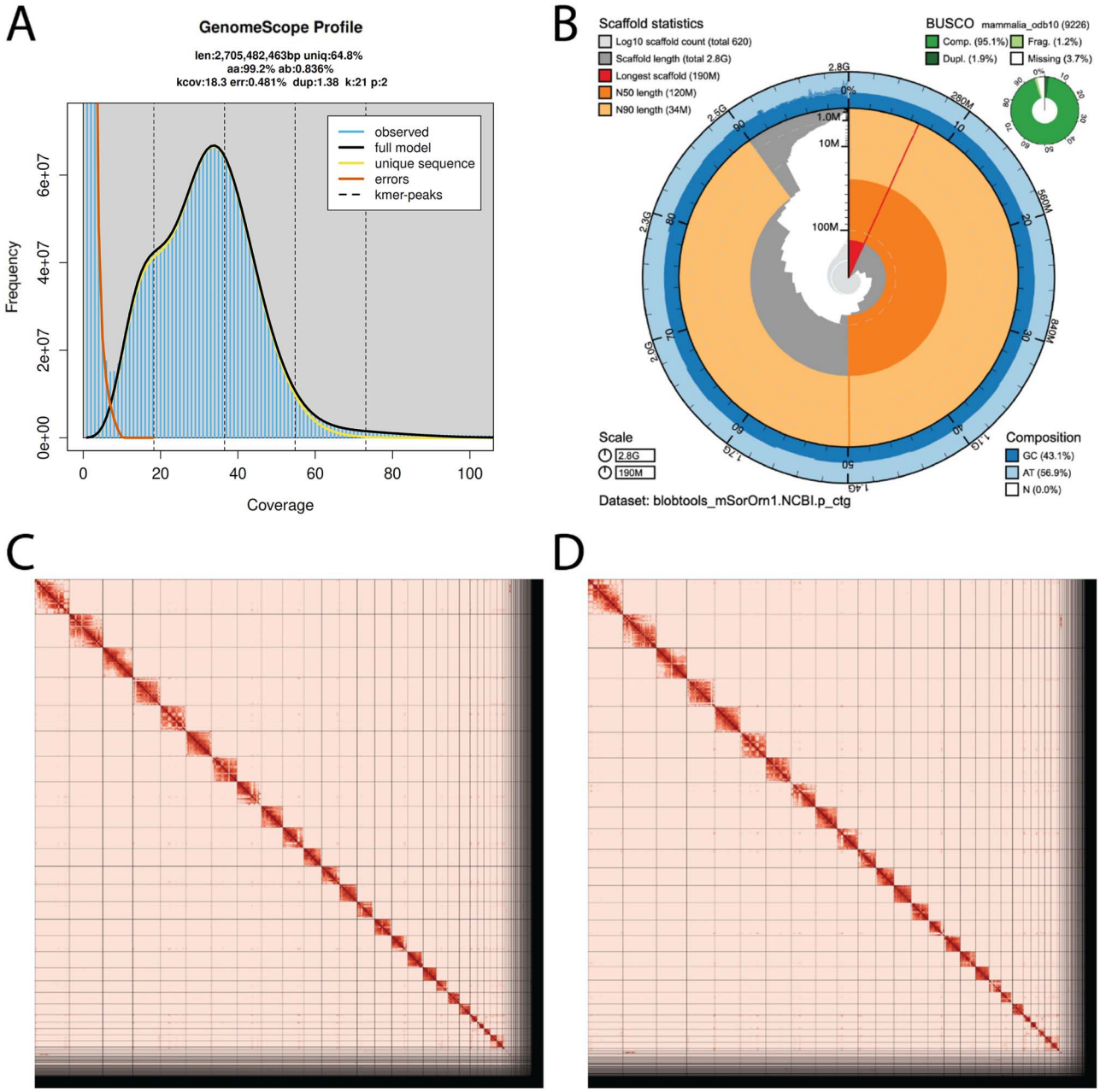

Fig. 2. Graphical representation of genome assembly and quality. (A) GenomeScope2.0 K-mer spectrum from Omni-C reads. (B) BlobToolKit Snailplot showing N50 metrics for *S. ornatus* assembly (mSorOrn1.0.p) and BUSCO scores for the mammalian set of orthologues. The outer circle displays the base pair composition across individual scaffolds, while the inner circle displays scaffold lengths. Scaffolds are shown clockwise from largest to smallest, with shading to indicate relevant lengths and counts. This indicates that the majority of the genome is in long scaffolds exceeding 34 mb and is highly contiguous. (C, D) Omni-C contact maps of primary and alternate assemblies respectively. Shaded values on the dot plot indicate physical contact between genomic sequences, so that shaded squares designate whole chromosomes. The 27 squares along the diagonal suggest a haploid genomic organization of 27.

is 2.81 Gb, with an N50 of 115.2 Mb, 623 scaffolds, and a BUSCO score of 95.1%. The assembly is highly contiguous, with 85.6% of the assembly present on 21 scaffolds > 50 Mb in length. Omni-C contact maps resolve an organization of 27 chromosomes, so many of these are likely chromosome-length scaffolds. This aligns with the reported karyotype of *S. ornatus* which has a haploid number of 27 chromosomes, including one small dot-like chromosome (Brown and Rudd 1981). The assembly consists of 43.07% repetitive elements, which is typical for a mammalian genome, although somewhat high for Eulipotyphlans (Osmanski et al. 2023). We do note that apart from LINES, the proportion of individual repeat elements is much lower than in other assemblies, however, there is a high proportion (15.53%) of unidentified repeats, so further manual curation may recover comparable repeat proportions.

This assembly joins five other species of *Sorex* with genomes on GenBank, resulting in a total of eight *Sorex* primary assemblies (Table 2). The whole-genome reported here is one of the highest quality *Sorex* assemblies, and the most contiguous assembly from the subgenus *Otisorex*, with the highest N50 and lowest number of scaffolds (Findley 1955; Repenning 1967; Esteva et al. 2010). One *Sorex araneus* assembly is more

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

Table 2. Whole-genome assemblies of different species of *Sorex*.

English name	Scientific name	Length (Gb)	# Scaffolds	Scaff. N50 (Mb)	Scaff. L50	Citation and Genbank ID
European shrew	*Sorex araneus*	2.41	65	374.4	3	Rhie et al. 2021 GCA_027595985.1
European shrew	*S. araneus*	2.42	12 845	22.8	36	Lindblad-Toh et al. 2011 GCA_000181275.2
Smoky shrew	*Sorex fumeus*	2.87	497	42.4	23	Cossette et al. 2024 GCA_029834395.2
Smoky shrew	*S. fumeus*	2.28	3379	2.6	209	Cossette et al. 2023 GCA_026122425.1
Pygmy shrew	*Sorex hoyi*	2.33	1289	6.3	92	Unpublished GCA_051176015.1
Maritime shrew	*Sorex maritimensis*	2.05	24 729	.105	5142	Cossette et al. 2024 GCA_030324115.1
American water shrew	*Sorex palustris*	2.3	498 345	.00964	11 458	Unpublished GCA_028565675.1
Ornate shrew	*Sorex ornatus*	2.81	623	115.2	10	This Paper GCA_041430695.1

contiguous, with an N50 of 374.4 Mb and 64 scaffolds (Rhie et al. 2021). This assembly is longer than the majority of other *Sorex* genomes, though slightly shorter than a 2.87 Gb *S. fumeus* assembly (Cossette et al. 2024).

Sorex shrews are known for having elevated metabolic rates and seasonal changes in body size and braincase depth (Dehnel 1949; Crowcroft and Ingles 1959; Pucek and Pucek 1963), a phenomenon which has been observed in other species with high metabolism and year-round activity (LaPoint et al. 2016; Nováková et al. 2022). Seasonal changes in body size, known as Dehnel's phenomenon, was first observed in the European common shrew (*S. araneus)*, and has also been documented in the ornate shrew (Newman 1977; Hays and Lidicker 2000). This genome assembly may allow for useful comparisons between species that do or do not exhibit high metabolic rates or Dehnel's phenomenon and may be used to identify conserved and distinct pathways that enable this lifestyle.

The ornate shrew is distributed across north, central, and southern California, and extends into Baja California, Mexico (Hall 1981; Owen and Hoffmann 1983; IUCN 2016). Multiple subspecies of the ornate shrew which occupy salt marsh environments are listed as California Department of Fish and Wildlife species of special concern (*S. o. salarius, S. o. salicornicus, S. o. sinuosus;* California Department of Fish and Wildlife 2025). Salt marshes in California have been disturbed by human development, and sea-level rise will continue to contribute to salt marsh habitat loss (The Nature Conservancy 2011; Thorne et al. 2018). These subspecies depend on dense vegetation and medium high marshes which provide habitat for invertebrate prey as well as shelter from flooding (Johnston and Rudd 1957).

The closely related vagrant shrew (*S. vagrans*) is also widely distributed, extending north from central California to southern British Columbia, occupying primarily damp meadows, forests and freshwater wetlands (Hennings and Hoffmann 1977; Junge and Hoffmann 1981; Gillihan and Foresman 2004). Molecular phylogenetic hypotheses suggest that the ornate and vagrant shrews are sister species (Esteva et al. 2010; Hope et al. 2014). The salt marsh wandering shrew *S. v. halicoetes* is a subspecies of special concern that commonly occupies some of the San Francisco Bay salt marshes, alongside *S. v. vagrans* (Johnston and Rudd 1957; CDFW 2025). As a part of the CCGP, this assembly will be integrated with whole-genome resequencing data from both *S. ornatus* and *S. vagrans* to aid in management decisions regarding endangered California *Sorex* subspecies, or those of special concern (Beninde et al. 2022; Fiedler et al. 2022; Shaffer et al. 2022).

This genome assembly will also be useful for investigating the extent of hybridization between these sister species in the San Pablo and Suisun Bays. Morphological and genetic study of this hybrid complex has revealed historical admixture between *S. o. californicus*, *S. o. sinuosus* and S. *v. vagrans*, likely followed by generations of backcrossing, resulting in variable combinations of characteristics from each (Rudd 1955; Maldonado et al. 2004; Patton 2019). Individuals in this complex all have karyotypes which correspond to the ornate shrew (2n = 54, FN = 76) rather than the vagrant shrew, which has a lower fundamental number (FN = 62) (Brown 1974; Brown and Rudd 1981). Mitochondrial genotyping has revealed that the majority of *Sorex* individuals involved in this hybrid complex have *S. vagrans* specific mitochondria, indicating that historical hybridization resulted in mitochondrial capture, but the extent of whole-genome admixture is unknown (Maldonado et al. 2001; Maldonado et al. 2004). The difference in chromosome arm number is suggested to be due to pericentric inversions (Patton 2019). If this is the case, then this system may provide a useful model to study the impact of structural variation on gene flow between recently diverged species (White 1978). Hybridization may also impact the potential management of the Suisun shrew, as it may provide pathways for genetic rescue or may lead to the loss of unique biodiversity Establishing the details of this contact zone will be necessary for future conservation (Wayne and Shaffer 2016; Chan et al. 2019). Overall, this genome will serve as a resource for future North American *Sorex* conservation genomics and may be useful for studies of high metabolism and other traits associated with shrew biology.

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

Acknowledgments

PacBio Sequel 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.
In addition, we thank Mark Allaback and Chad Steiner who collected the specimen used for de novo assembly.

Author contributions

Isaac Linn (Data curation, Investigation, Visualization, Writing—original draft, Writing—review & editing), Merly Escalona (Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Writing—original draft, Writing—review & editing), Oanh Nguyen (Investigation, Methodology, Writing—review & editing), Mohan P.A. Marimuthu (Investigation, Methodology), Noravit Chumchim (Investigation, Methodology), Colin W Fairbairn (Investigation, Methodology), William Seligmann (Investigation, Methodology), Eric Beraut (Investigation, Methodology), Chris John Conroy (Resources, Writing—review & editing), Rauri C.K. Bowie (Conceptualization, Funding acquisition, Resources, Writing—review & editing), and Michael W. Nachman (Conceptualization, Funding acquisition, Resources, Writing—review & editing)

Supplementary material

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

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].

Data availability

Data generated for this study are available under NCBI BioProject PRJNA777223. Raw sequencing data for sample MVZ:Mamm:240319 (NCBI BioSample SAMN40725982) are deposited in the NCBI Short Read Archive (SRA) under SRR30639629,SRR30639630 for PacBio HiFi sequencing data, and SRR30639627,SRR30639628 for the Omni-C Illumina sequencing data. GenBank accessions for both primary and alternate assemblies are GCA_041430695.1 and GCA_041430635.1; and for genome sequences JBCEVT000000000 and JBCEVU000000000. The GenBank genome assembly for the mitochondrial genome is CM084381.1. Assembly scripts and other data for the analyses presented can be found at the following GitHub repository: www.github.com/ccgproject/ccgp_assembly. Repetitive element annotation and Liftoff gene annotation were deposited on Dryad.

References

Abdennur N, Mirny LA. Cooler: scalable storage for hi-C data and other Genomically Labeled arrays. Edited by Wren J. *Bioinformatics*. 2020;**36**:311–316. https://doi.org/10.1093/bioinformatics/btz540

Allio R, Schomaker-Bastos A, Romiguier J, Prosdocimi F, Nabholz B, Delsuc F. MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment Phylogenomics. *Mol Ecol Resour*. 2020;**20**:892–905. https://doi.org/10.1111/1755-0998.13160

Astashyn A, Tvedte ES, Sweeney D, Sapojnikov V, Bouk N, Joukov V, Mozes E, Strope PK, Sylla PM, Wagner L, et al. Rapid and sensitive detection of genome contamination at scale with FCS-GX. *Genome Biol*. 2024;**25**:60. https://doi.org/10.1186/s13059-024-03198-7

Beagle J, Salomon M, Grossinger RM, Baumgarten SA, Askevold RA. Shifting shores: marsh expansion and retreat in San Pablo Bay. 2015; *San Francisco Estuary Institute*. Richmond, CA. No. 751: 75 pp. Available online: https://www.sfei.org/documents/shifting-shores-marsh-expansion-and-retreat-san-pablo-bay

Benham PM, Bowie RCK. The influence of spatially heterogeneous anthropogenic change on bill size evolution in a coastal songbird. *Evol Appl*. 2021;**14**:607–624. https://doi.org/10.1111/eva.13144

Beninde J, Toffelmier E, Shaffer HB. A brief history of population genetic research in California and an evaluation of its utility for conservation decision-making. *J Hered*. 2022;**113**:604–614. https://doi.org/10.1093/jhered/esac049

von Bloeker, JC Jr. 1967. Land mammals of the Southern California Islands. Pp. 245-266, in *Proceedings of the symposium on the biology of the California Islands*. Edited by Philbrick RN. Santa Barbara Botanic Garden, Santa Barbara, California, 245–266. Available online: https://sbbotanicgarden.org/symposium-on-the-biology-of-the-california-islands/

Boore JL. Animal mitochondrial genomes. *Nucleic Acids Res*. 1999;**27**:1767–1780. https://doi.org/10.1093/nar/27.8.1767

Brown RJ. A comparative study of the chromosomes of three species of shrews, *Sorex bendirii*, *Sorex trowbridgii*, and *Sorex vagrans*. *Wasmann J Biol*. 1974;**32**:303–326.

Brown RJ, Rudd RL. Chromosomal comparisons within the *Sorex ornatus-S. Vagrans* complex. *Wasmann J Biol*. 1981;**39**:30–35

California Department of Fish and Wildlife. Special Animals List. *California Natural Diversity Database*. Periodic Publication, Sacramento, CA; October 2025: 134 pp. Available online: https://wildlife.ca.gov/Data/CNDDB/Plants-and-Animals

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinform*. 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 Genes|Genomes|Genetics*. 2020;**10**:1361–1374. https://doi.org/10.1534/g3.119.400908

Challis R, Kumar S, Sotero-Ciao C, Brown M, Blaxter M. Genomes on a tree (GoaT): a versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic tree of life. *Wellcome Open Res*Jan 17. 2023;**8**:24. https://doi.org/10.12688/wellcomeopenres.18658.1

Chan WY, Hoffmann AA, van Oppen MJH. Hybridization as a conservation management tool. *Conserv Lett*. 2019;**12**:e12652. https://doi.org/10.1111/conl.12652

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

Cheng H, Asri M, Lucas J, Koren S, Li H. Scalable telomere-to-telomere assembly for diploid and polyploid genomes with double graph. *Nat Methods*. 2024;**21**:967–970. https://doi.org/10.1038/s41592-024-02269-8

Coombe L, Kazemi P, Wong K, Birol I, Warren RL. Multi-genome synteny detection using minimizer graph mappings. *bioRxiv*. 2024. https://doi.org/10.1101/2024.02.07.579356

Coombe L, Warren RL, Birol I. ntSynt-viz: visualizing synteny patterns across multiple genomes. *bioRxiv*. 2025. https://doi.org/10.1101/2025.01.15.633221

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

Cossette M-L, Stewart DT, Haghani A, Zoller JA, Shafer ABA, Horvath S. Epigenetics and island-mainland divergence in an insectivorous small mammal. *Mol Ecol*. 2023;**32**:152–166. https://doi.org/10.1111/mec.16735

Cossette ML, Stewart DT, Shafer ABA. Comparative genomics of the World's smallest mammals reveals links to echolocation, metabolism, and body size plasticity. Edited by Mello B. *Genome Biol Evolution*. 2024;**16**, **16**:evae225. https://doi.org/10.1093/gbe/evae225

Crowcroft P, Ingles JM. Seasonal changes in the brain-case of the common shrew (*Sorex araneus L.*). *Nature*. 1959;**183**:907–908. https://doi.org/10.1038/183907b0

Dehnel A. Studies on the genus *sorex* L. *Annales Universitatis Mariae Curie-Sklodowska, Sectio C*. 1949;**4**:17–97

Esteva M, Cervantes FA, Brant SV, Cook JA. Molecular phylogeny of long-tailed shrews (genus *sorex*) from México and Guatemala. *Zootaxa*. 2010;**2615**:47–65. https://doi.org/10.11646/zootaxa.2615.1.3

Fiedler PL, Erickson B, Esgro M, Gold M, Hull JM, Norris JM, 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

Findley JS. Speciation of the wandering shrew. In: Raymond Hall E, Byron Leonard A, Wilson RW, editors. *University of Kansas Publications, Museum of Natural History*. University of Kansas, Lawrence, Kansas; 1955;**9**:1–68.

Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci USA*. 2020;**117**:9451–9457. https://doi.org/10.1073/pnas.1921046117

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. Edited by Ioshikhes I. *PLoS Comput Biol*. 2019;**15**:e1007273. https://doi.org/10.1371/journal.pcbi.1007273

Gillihan SW, Foresman KR. *Sorex vagrans Mammalian Species*. 2004;**744**:1–5. https://doi.org/10.1644/744

Grinnell J. Review of the recent mammal Fauna of California. University of California Publications in Zoology. 1933;**40**:71–234

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

Hall ER. *The mammals of North America v 1*. 2nd ed. New York: Wiley; 1981.

Hays WST, Lidicker WZJ. Winter aggregations, Dehnel effect, and habitat relations in the Suisun shrew *Sorex ornatus sinuosus*. *Acta Theriol*. 2000;**45**:433–442. https://doi.org/10.4098/AT.arch.00-44

Hennings D, Hoffmann RS. A review of the taxonomy of the *Sorex vagrans* species complex from western North America. In: Wiley E, editor. *Occasional papers of the Museum of Natural History*. University of Kansas, Lawrence, Kansas; 1977;**68**:1–35.

Hope AG, Panter N, Cook JA, Talbot SL, Nagorsen DW. Multilocus Phylogeography and systematic revision of north American water shrews (genus: *sorex*). *J Mammal*. 2014;**95**:722–738. https://doi.org/10.1644/13-MAMM-A-196

IUCN, Álvarez-Castañeda ST, Matson J, Woodman N, de Grammont PC, Hammerson GA. *Sorex ornatus* The IUCN Red List of Threatened Species. 2016;**2016**. https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T41408A22313226.en

Johnston RF, Rudd RL. Breeding of the salt marsh shrew. *J Mammal*. 1957;**38**:157–163. https://doi.org/10.2307/1376305

Junge JA, Hoffmann RA. An annotated key to the long-tailed shrews (genus *sorex*) of the United States and Canada, with notes on middle America *sorex*. Occasional Papers of the Museum of Natural History, University of Kansas. 1981;**94**:1–48

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

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**:1–16. https://doi.org/10.1093/gigascience/gix085

LaPoint S, Keicher L, Wikelski M, Zub K, Dechmann DKN. Growth overshoot and seasonal size changes in the skulls of two weasel species. *R Soc Open Sci*. 2016;**4**:160947–160947. https://doi.org/10.1098/rsos.160947

Lee S, Bakker CR, Vitzthum C, Alver BH, Park PJ. Pairs and Pairix: a file format and a tool for efficient storage and retrieval for Hi-C read pairs. *Bioinform*. 2022; **38**:1729–1731. https://doi.org/10.1093/bioinformatics/btab870

Li, H. 2013. *Aligning sequence reads*, Clone Sequences and Assembly Contigs With BWA-MEM. *arXiv*. http://arxiv.org/abs/1303.3997

Lindblad-Toh K, Broad Institute Sequencing Platform and Whole Genome Assembly Team, Baylor College of Medicine Human Genome Sequencing Center Sequencing Team, Genome Institute at Washington University, Garber M, Zuk O, Lin MF, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478** : 476–482. https://doi.org/10.1038/nature10530

Maldonado JE, Vilà C, Wayne RK. Tripartite genetic subdivisions in the ornate shrew (*Sorex ornatus*). *Mol Ecol*. 2001;**10**:127–147. https://doi.org/10.1046/j.1365-294X.2001.01178.x

Maldonado JE, Hertel F, Vilà C. Discordant patterns of morphological variation in genetically divergent populations of ornate shrews (*Sorex ornatus*). *J Mammal*. 2004;**85**:886–896. https://doi.org/10.1644/1545-1542(2004)085<0886:DPOMVI>2.0.CO;2(2004)085<0886:DPOMVI>2.0.CO;2

Mammal Diversity Database. Mammal diversity database (version 2.0) [data set]. *Zenodo*. 2025. https://doi.org/10.5281/zenodo.15007505

Manni M, Berkeley MR, Seppey M, Simao 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

Morrison PR, Pearson OP. The metabolism of a very small mammal. *Science*. 1946;**104**:287–298. https://doi.org/10.1126/science.104.2700.287

Newman JR. Energy flow of a secondary consumer (*Sorex sinuosus*) in a salt Marsh Community. In: *Doctoral dissertation*. University of California, Davis, California; 1970. Available online: https://babel.hathitrust.org/cgi/pt?id=uc1.x30154&seq=12.

Newman JR. Energy value of the shrew, *Sorex ornatus*. *Acta Theriol*. 1977;**22**:274–275. https://doi.org/10.4098/AT.arch.77-24

Newman JR, Rudd RL. Minimum and maximum metabolic rates of *sorex sinuosus*. *Acta Theriol*. 1978;**23**:371–380. Available online: https://www.europeana.eu/en/item/0940431/_nnbZ9W5. https://doi.org/10.4098/AT.arch.78-28

Nováková L, Lázaro J, Muturi M, Dullin C, Dechmann DKN. Winter conditions, not resource availability alone, may drive reversible seasonal skull size changes in moles. *R Soc Open Sci*. 2022;**9**:220652–220652. https://doi.org/10.1098/rsos.220652

Ochocińska D, Taylor JRE. Living at the physiological limits: field and maximum metabolic rates of the common shrew (*Sorex araneus*). *Physiol Biochem Zool*. 2005;**78**:808–818. https://doi.org/10.1086/431190

Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. Mash: fast genome and metagenome distance

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

estimation using MinHash. *Genome Biol*. 2016;**17**:132. https://doi.org/10.1186/s13059-016-0997-x

Open2C, Abdennur N, Fudenberg G, Flyamer IM, Galitsyna AA, Goloborodko A, Imakaev M, Venev SV. Pairtools: from sequencing data to chromosome contacts. *bioRxiv*. 2023. https://doi.org/10.1101/2023.02.13.528389

Osmanski AB, Paulat NS, Korstian J, Grimshaw JR, Halsey M, Sullivan KA, Moreno-Santillán DD, Crookshanks C, Roberts J, Garcia C, et al. Insights into mammalian TE diversity through the curation of 248 genome assemblies. *Science*. 2023;**380**:eabn1430. https://doi.org/10.1126/science.abn1430

Owen JG, Hoffmann RS. *Sorex ornatus Mammalian Species*. 1983;**212**: 1–5. https://doi.org/10.2307/3504070

Patton, J. L. 2019. Diversification within the ornate (*Sorex ornatus*) and vagrant shrew (*Sorex vagrans*) complex in the San Francisco Bay Area of California. *From field to laboratory: a memorial volume in honor of Robert J. Baker*. Edited by Bradley RD, Genoways HH, Schmidly DJ, Bradley LC. Special Publication, Museum of Texas Tech University, **71**: 85–129. Available online: https://www.depts.ttu.edu/nsrl/publications/downloads/SP71.pdf

Pearson OP. Metabolism of small mammals, with remarks on the lower limit of mammalian size. *Science*. 1948;**108**:44. https://doi.org/10.1126/science.108.2793.44

Pucek Z, Pucek Z. Seasonal changes in the braincase of some representatives of the genus *sorex* from the palearctic. *J Mammal*. 1963;**44**: 523–536. https://doi.org/10.2307/1377135

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. *Nat Commun*. 2020;**11**:1432. https://doi.org/10.1038/s41467-020-14998-3

Repenning CA. Subfamilies and genera of the Soricidae. In: *United States Department of the Interior: Geological Survey*. U.S. Government Printing Office, Washington, D.C.; 1967: No. 565. Available online: https://permanent.fdlp.gov/gpo80580/report.pdf

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

Rudd RL. Population variation and hybridization in some Californian shrews. *Syst Zool*. 1955;**4**:21–34. https://doi.org/10.2307/2411934

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

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

Taylor JRE. Evolution of energetic strategies in shrews. In: Wójick JM, Wolsan M, eds. *Evolution of shrews*. Polish Academy of Sciences: Mammal Research Institute; 1998:309–346.

The Nature Conservancy, Enright CPE. *Suisun marsh conservation assessment volume I: characterization report*. Wetland and Water Resources, Inc; 2011: Available online: https://nerrssciencecollaborative.org/resource/adaptation-planning-approach-vulnerable-infrastructure.

Thorne K, MacDonald G, Guntenspergen G, Ambrose R, Buffington K, Dugger B, Freeman C, Janousek C, Brown L, Rosencranz J, et al. U.S. Pacific coastal wetland resilience and vulnerability to sea-level rise. *Sci Adv*. 2018;**4**:eaao3270. https://doi.org/10.1126/sciadv.aao3270

Uliano-Silva M, João GRNF, Krasheninnikova K, Darwin Tree of Life Consortium, Blaxter M, Mieszkowska N, Hall N, et al. MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high Fidelity reads. *BMC Bioinform*. 2023;**24**:288. https://doi.org/10.1186/s12859-023-05385-y

Wayne RK, Shaffer HB. Hybridization and endangered species protection in the molecular era. *Mol Ecol*. 2016;**25**:2680–2689. https://doi.org/10.1111/mec.13642

White MJD. *Modes of speciation. A series of books in biology*. San Francisco: W. H. Freeman; 1978.

Wilder AP, Shier DM, King SND, Dudchenko O, Funk ER, Misuraca A, Houck ML, Miller WB, Curry CJ, Fronczek J, et al. Fitness benefits of genetic rescue despite chromosomal differences in an endangered pocket mouse. *Science*. 2025;**389**:885–839. https://doi.org/10.1126/science.adn4666

Downloaded from https://academic.oup.com/jhered/article/117/3/450/8272640 by University of California Library - Berkeley Library user on 08 June 2026

© The Author(s) 2025. Published by Oxford University Press on behalf of The American Genetic Association. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.
Journal of Heredity, 2026, 117, 450–458
https://doi.org/10.1093/jhered/esaf074
Genome Resources