Conserv Genet (2016) 17:1055–1066
DOI 10.1007/s10592-016-0843-4

RESEARCH ARTICLE

Conservation genetics of the endangered San Francisco Bay endemic salt marsh harvest mouse (*Reithrodontomys raviventris*)

M. J. Statham[1] · S. Aamoth[1] · L. Barthman-Thompson[2] · S. Estrella[2] · S. Fresquez[1,2] · L. D. Hernandez[1,2] · R. Tertes[3] · B. N. Sacks[1,4]

Received: 18 September 2015 / Accepted: 11 April 2016 / Published online: 19 April 2016

Abstract The salt marsh harvest mouse (SMHM, *Reithrodontomys raviventris*) is an endangered species endemic to the San Francisco Bay region of California, USA, where habitat loss and fragmentation over the past century have reduced the mouse's distribution to <25 % of its historical range. To aid in conservation prioritization, we first investigated the possibility of hybridization with the morphologically similar western harvest mouse (WHM, *R. megalotis*) in areas of sympatry and developed genetic tools to differentiate the two species. We then investigated the phylogeography and genetic structure of the SMHM, including support for currently recognized SMHM subspecies designations. Lastly, we evaluated the morphological criteria currently used for the identification of species in the field. Analyses using mtDNA cytochrome *b* sequences and 11 microsatellites from 142 mice indicated complete and substantial separation of the SMHM and WHM, with no evidence of hybridization. These genetic markers as well as the mtDNA control region also identified a deep genetic division within the SMHM concordant with the current subspecies designations, *R. r. raviventris* and *R. r. halicoetes*. We identified the lowest genetic diversity within the southern subspecies, which inhabits a much reduced and highly fragmented portion of the species range. Morphological field identification of harvest mouse species was more successful at identifying SMHM (92 %) than WHM (44 %), with a large portion of WHM being incorrectly identified as SMHM. Field identification of harvest mouse species in the range of the southern SMHM subspecies was just above 50 %, indicating that current methods for morphological differentiation of species in that area are insufficient. Our confirmation of genetically distinct SMHM subspecies highlights the importance of determining the status and genetic composition of relict populations in the remaining patches of marshland in the central San Francisco Bay where the two subspecies may occur, as well as developing better tools for the discrimination of species, particularly in the range of the southern subspecies

Keywords Field identification · Management · Morphology · Population subdivision · Species identification · Subspecies

Electronic supplementary material The online version of this article (doi:10.1007/s10592-016-0843-4) contains supplementary material, which is available to authorized users.

✉ M. J. Statham
Statham@ucdavis.edu

[1] Mammalian Ecology and Conservation Unit, Veterinary Genetics Laboratory, University of California, One Shields Avenue/Old Davis Road, Davis, CA 95616-8744, USA

[2] Suisun Marsh Unit, Bay Delta Region, California Department of Fish and Wildlife, 2109 Arch Airport Road, Stockton, CA 95206, USA

[3] Don Edwards San Francisco Bay National Wildlife Refuge, United State Fish and Wildlife Service, One Marshlands Road, Fremont, CA 94555, USA

[4] Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, One Shields Avenue/Old Davis Road, Davis, CA 95616-8744, USA

Introduction

Understanding the underlying genetic architecture of a species facilitates preservation of genetic diversity and maintenance of maximum evolutionary potential. The partitioning of genetic diversity within a species reflects its

history (Moritz 1994; Statham et al. 2014). Recent partitioning can result in subpopulations with similar but non-identical allele frequencies, whereas longer term population subdivision can allow for greater differentiation through genetic drift, accumulation of novel mutations, and local adaptation if exposed to different selective environments (Hewitt 2000). The magnitude of divergence or population subdivision also depends in part on the sizes of populations and the frequency of migration (Paetkau 1999). Habitat destruction and fragmentation associated with human alterations of landscapes increasingly threaten viability of endemic species, and necessitates managing populations to maintain genetic diversity (Allendorf and Luikart 2007). The coastal salt marshes of the San Francisco Bay Area [consisting of San Francisco Bay proper (hereafter, San Francisco), San Pablo Bay, and Suisun Bay] have been reduced and fragmented by urban and agricultural development, erosion, water management, and vegetation change (Shellhammer 1989; USFWS 2013). This fragmentation will likely be exacerbated by future sea level rise, which is predicted to range 50–200 cm during the twenty first century (US National Research Council 2010).

The salt marshes of the San Francisco Bay Area of California are home to five endangered species including the endemic salt marsh harvest mouse (SMHM, *Reithrodontomys raviventris*; USFWS 2013). The SMHM (Rodentia, Cricetidae) is restricted to salt and brackish marshes, and is highly dependent on dense cover for predator avoidance (Shellhammer 1989). These mice can swim well and tolerate high salt concentrations in their diets (Fisler 1963, 1965). Although unverified by genetic criteria, two subspecies of SMHM are recognized. The southern subspecies (*R. r. raviventris*) inhabits central and South San Francisco Bay, while the northern subspecies (*R. r. halicoetes*) occurs in San Pablo and Suisun Bays (Fig. 1; Shellhammer 1989). Fragmentation and loss of $\geq$75 % of the species habitat through reclamation of tidal areas led to the SMHM being state and federally listed as endangered (Goals project 2000; Shellhammer 1982, 1989; USFWS 2013). The situation is magnified in the range of the southern subspecies where >90 % of the historical habitat has been lost, with most of the remainder lacking in suitable high tide escape habitat (USFWS 2013). Despite being listed as endangered for over 40 years there are a number of key outstanding questions necessary for the management of the species.

Accurate species identification is one of the most basic requirements for population monitoring. Salt marsh harvest mouse are often difficult to identify in the field due to extensive sympatry with the morphologically similar western harvest mouse (WHM, *R. megalotis*). The WHM is abundant and widespread, with a range across much of the western U.S., extending into Canada and Mexico. Although the WHM is primarily a grassland species, it also inhabits the edges of salt and brackish marshes in the San Francisco Bay Area (Fisler 1963). Differentiating between the two harvest mouse species in the field requires several morphological measurements to assign individuals to a species (Shellhammer 1984). This method often identifies ambiguous individuals of intermediate morphology, classified as unknown harvest mouse species, which has contributed to speculation that the two species could be undergoing hybridization (USFWS 2013). Discrepancies regarding how personnel take measurements can result in assignment of the same mouse to different species. In addition, juvenile and sub-adult mice are especially difficult to identify to species. Thus additional tools are needed to ensure accurate species assignment of harvest mice.

A central goal of the US Fish and Wildlife Service's species recovery plan for the SMHM is to maintain genetic diversity to minimize the risk of inbreeding depression and enable future evolution and resilience to environmental change (USFWS 2013). This goal is intended to secure self-sustaining populations of SMHM across the geographic and ecological range, as well as distinct genetic units within the species. However, no phylogeographic work has been conducted to assess genetic subunits or genetic diversity within the SMHM (Shellhammer 1989; USFWS 2013). Although two subspecies of SMHM have been described, there is no morphological characteristic that discretely separates them. One of the main characteristics used to distinguish between the subspecies is the color of the ventral pelage. While many individuals of the southern subspecies have a striking cinnamon-colored ventral fur, others have white, which is also found in the majority of individuals of the northern subspecies (Shellhammer 1982). Thus the validity of currently recognized subspecies needs to be investigated with genetic data.

We undertook this study to provide support to conservation efforts for the endangered SMHM. The objectives were to: (1) create genetic tools for accurate species identification; (2) assess the possibility of hybridization between SMHM and WHM; (3) characterize population sub-division in SMHM; (4) characterize phylogeographic (prehistorical) patterns among populations, including assessment of the distinctiveness of the currently recognized northern and southern SMHM subspecies (*R. r. halicoetes* and *R. r. raviventris*); and (5) quantify and compare the genetic diversity in Suisun, San Pablo, and San Francisco bays, the three primary areas of habitat where the species occurs.

Materials and methods

Fieldwork and sampling

We trapped 142 harvest mice for genetic sampling in 2011–2013 in three locations within each of the three bays

Fig. 1 Map of sampling locations within the San Francisco Bay area. Sampling locations are indicated with a *red* pentagon. Coarse outline of the salt marshes is indicated with a *yellow line*. (Color figure online)

of the San Francisco Bay Area (Fig. 1). These three bays encompass the entire historical (and modern) SMHM range. We used up to >100 Sherman live traps (H.B. Sherman Traps, Tallahassee, FL), spaced at ~10 m intervals, with the layout depending on the wetland shape. We baited traps with mixed bird seed and ground walnut, added cotton or polyester batting for warmth, set the traps at dusk, and checked them at dawn. We measured the trapped mice for their total and tail lengths, tail diameter, and weight. Morphological assignment to species was based largely on these measurements, along with pelage coloration, and a number of characteristics of the tail as described in Shellhammer (1984). We plucked hair as a source of DNA. Prior to sampling from an individual we physically wiped down the forceps with a clean tissue, sterilized the forceps in a 2 % bleach solution, rinsed with water to remove the bleach, and dried the forceps with a second tissue. We stored the hair in 95–100 % ethanol until DNA extraction. Animal trapping, handing, and genetic sampling procedures were approved by UC Davis Institutional Animal Care and Use Committee (IACUC No. 17889) and authorized by the appropriate state and federal agencies.

Lab methods

Hair samples were digested overnight in solution consisting of 39 mM Dithiothreitiol (DTT), 100 mM NaCl, 3 mM $CaCl_2$, 8.13 mM tris EDTA pH 8 (TE), 2 % sodium dodecyl sulfate (SDS), and 0.2 mg proteinase K in a total volume of 400 µL. The digestion solution was purified using a modified phenol/chloroform method (Sambrook and Russell 2001), and then further cleaned using Amicon Ultra centrifugal filters (Millipore Ltd). Each extraction set of 11 samples included a negative control to monitor for

contamination. The resulting purified DNA was then used for polymerase chain reaction (PCR) amplification.

Mitochondrial DNA (mtDNA)

We amplified the 5′ portion of the cytochrome *b* gene and a small section of tRNA-Glutamate (hereafter, Cyt-b) in both target harvest mouse species using the primers MVZ-05 and MVZ-04R (Smith and Patton 1993; Brown 2003). This reaction was conducted in a total volume of 23 μL, consisting of 2 μL DNA extract, 1 × PCR buffer, 2.5 mm $MgCl^2$, 0.2 mm dNTPs, 0.1 μg/μL BSA, 0.5 μm of each primer and 1 U of AmpliTaq DNA Polymerase (Applied Biosystems, Inc), with the following cycle conditions: 94 °C for 5 min, followed by 40 cycles of 94 °C for 30 s, 56 °C for 30 s and 72 °C for 45 s; followed by a 10-min extension period at 72 °C. We amplified the control region in SMHM using the primers Nmic5′ (Méndez-Harclerode et al. 2005) and 3′-phenylalanine (Castro-Campillo et al. 1999), otherwise using the same PCR mixture and cycling conditions. PCR products were purified using ExoSap-IT (Affymetrix, Inc.) and sequenced in both directions using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, Inc.). We then electrophoresed products on an ABI 3730 capillary sequencer (Applied Biosystems, Inc). The resulting DNA sequences were aligned in the program Sequencher v5.1.

We used the resulting Cyt-b sequences to assign individuals to species by comparing to published sequences using BLAST in GenBank (Altschul et al. 1990). In addition, we investigated the phylogenetic relationship of the Cyt-b sequences we obtained with published ones from SMHM, WHM, and closely related harvest mouse species (Bell et al. 2001; Arellano et al. 2005). We created a phylogenetic tree using maximum likelihood in the program MEGA v5 (Tamura et al. 2011), using the most appropriate mutation model (GTR + I + G) as identified using Akaike information criterion in MEGA, with 1000 bootstrap replicates. For both mtDNA regions we created a median-joining network (Bandelt et al. 1999) in Network 4.2.0.1 (www.fluxus-engineering.com) to assess the phylogenetic relationships among haplotypes. We used Arlequin 3.5 (Excoffier and Lischer 2010) to estimate haplotype and nucleotide diversity, and to estimate the degree of subdivision among bay-specific populations (Φ_{ST}).

Splitting time estimates of salt marsh harvest mouse populations

We estimated the effective population sizes and splitting times between the North and South Bay SMHM using Markov chain Monte Carlo (MCMC)-based simulations in the program IMa2 (Hey and Nielsen 2007; Hey 2010). Using jmodeltest v 0.1.1 (Posada 2008), we identified HKY as the most appropriate nucleotide substitution model on this single species dataset. We converted mutation scaled parameters to more biologically meaningful values using a range of mutation rates; we used 1 % per million years as the slowest rate, 2 % as the median, and 3 % as the highest given a 1.5–2 faster mutation rate in rodents than other mammals (Arbogast 1999; Nabholz et al. 2008). We tested both 'isolation only' and 'isolation with migration' models. Initial runs were conducted with default settings. We then used 40 geometrically heated chains sampling 10,000 steps (1 million generations) after a burn-in of 20,000 steps. We ran each analysis a second time using a different random seed to assess consistency. Two runs with high ESS values (>1000) and trend plots indicative of good mixing were combined in "Load Tree" mode to generate multi-dimensional parameter estimates. The 'isolation with migration' analyses failed to resolve the ancestral population size or population splitting time and results of this analysis were therefore not presented.

Microsatellite analyses

We screened all samples at 10 specifically designed harvest mice microsatellite loci (Reponen et al. 2014). In addition we investigated the utility of 10 microsatellite loci developed in *Reithrodontomys spectabilis*, a harvest mouse from Mexico (Vázquez-Domínguez and Espindola 2013). One locus, R34, was polymorphic in both SMHM and WHM and was subsequently used to screen all individuals. We tested for deviations from Hardy–Weinberg and linkage equilibrium using Genepop (http://genepop.curtin.edu.au/). We corrected for multiple test using the sequential Bonferroni method (Rice 1989). We calculated the Observed (H_o) and Expected Heterozygosities (H_e), and average number of alleles per locus (A) in Microsatellite tool kit (Park 2001). We calculated the inbreeding coefficient (F_{IS}) and allelic richness (Ar) in FSTAT v 2.9.3.2 (Goudet 1995). We calculated pairwise F_{ST} among sampling sites using Arlequin 3.5 (Excoffier and Lischer 2010).

We examined population substructure using a model based Bayesian clustering method implemented in the program STRUCTURE v 2.3.3 using the admixture model with correlated allele frequencies (Pritchard et al. 2000; Falush et al. 2003). This technique allowed us to evaluate population substructure without the need for a prior assignment of individuals to populations. Several loci amplified only in one of the harvest mouse species (as ascertained through mtDNA sequencing). Thus different loci were included depending on whether both species, just SMHM, or just WHM, were being analyzed. We first ran an analysis on all specimens, regardless of mtDNA species assignment, using only loci that amplified in both harvest

mice species. Subsequently, we ran independent analyses on each of the highest supported clusters evident in the first analysis. Iterations were run for K values of 1–5, with a burn-in of 100,000 followed by a run of 500,000 generations. Simulations were run 5 times at each value of K to assess consistency across runs. We determined the most meaningful K values by plotting the Ln P(D) values and determining where the greatest support was found (Pritchard and Wen 2002).

Results

Interspecific comparisons

We identified 8 SMHM haplotypes (Table 1) and 11 WHM haplotypes based on our 426 bp Cyt-b fragment. All haplotypes were deposited in GenBank (Accession Nos. KU528725-KU528750 and KU532155-KU532173). Only a single SMHM haplotype (B) had been previously identified (Arellano et al. 2005). The phylogenetic tree confirmed the species identifications of SMHM and WHM haplotypes, grouping them with published sequences from the same species with high bootstrap support (99 and 93 % respectively; Fig. 2). The SMHM haplotypes were most closely related to those from *R. montanus* (plains harvest mouse; Texas), whereas WHM haplotypes were most closely related to those from *R. zacatecae* (Zacatecas harvest mouse; Mexico). We observed 40 fixed nucleotide differences between the SMHM and WHM. The sequence divergence between SMHM and WHM (11.06 %) minus the average sequence divergence within species (0.21 %) resulted in a net sequence divergence between species of 10.85 %. Converting this value to years using mutation rates 1, 2, and 3 % per million years (see methods) resulted in divergence time estimates of 10.85, 5.43, and 3.62 million years between SMHM and WHM species. In contrast we observed only 1–5 nucleotide differences between any two SMHM haplotypes, equivalent to 0.23–1.17 % sequence divergence.

We excluded microsatellite locus Rrav15 from analyses due to ambiguous alleles. Within the SMHM nine loci amplified. Within the WHM six loci amplified. We observed fixed allelic differences between species at three loci: Rrav 6, 8, and 21 (Supplemental Table). We noted 5 occurrences of significant deviation from HWE, each of which occurred in just one of three bay populations (Supplemental Table). The only occurrences of significant deviation from LD occurred between loci Rrav6 and Rrav29 in the San Pablo Bay population of SMHM. None of the deviations from linkage and Hardy–Weinberg equilibria were consistent among populations, suggesting substructure rather than allelic dropout or physical linkage as the cause; therefore all loci (except Rrav15) were retained for subsequent analyses.

Bayesian cluster analysis of all harvest mice in the program STRUCTURE indicated highest support for $K = 2$, after which support values plateaued (Supplemental Figure). The discrete groups identified were consistent with the mtDNA species assignment indicating that they represented the SMHM and WHM species (Fig. 3). Individuals were assigned on average 99.6 % to a single species cluster.

Testing morphological identification of species

Species assignment of SMHM based on morphology correctly identified 92 % of individuals (83 of 90), with one

Table 1 mtDNA Cytochrome *b* haplotype occurrence in SMHM in all sampling locations across the three bays

Location	*n*	Haplotype							
		A	B	C	D	E	I	L	M
San Pablo Bay									
Napa/Sonoma Marshes	15	10	–	–	4	–	1	–	–
Fagan Slough	10	–	10	–	–	–	–	–	–
Napa Plant Restoration Area	5	3	–	–	2	–	–	–	–
Suisun Bay									
Hill Slough	10	6	–	–	2	2	–	–	–
Point Edith	10	10	–	–	–	–	–	–	–
Denverton	10	6	2	–	–	1	–	–	1
San Francisco Bay									
Warm Springs Mouse Pasture	4	–	–	2	–	–	–	2	–
Eden Landing	20	–	–	18	–	–	–	2	–
Mayhew's Landing	6	1	–	–	–	–	–	5	–
Total	90	36	12	20	8	3	1	9	1

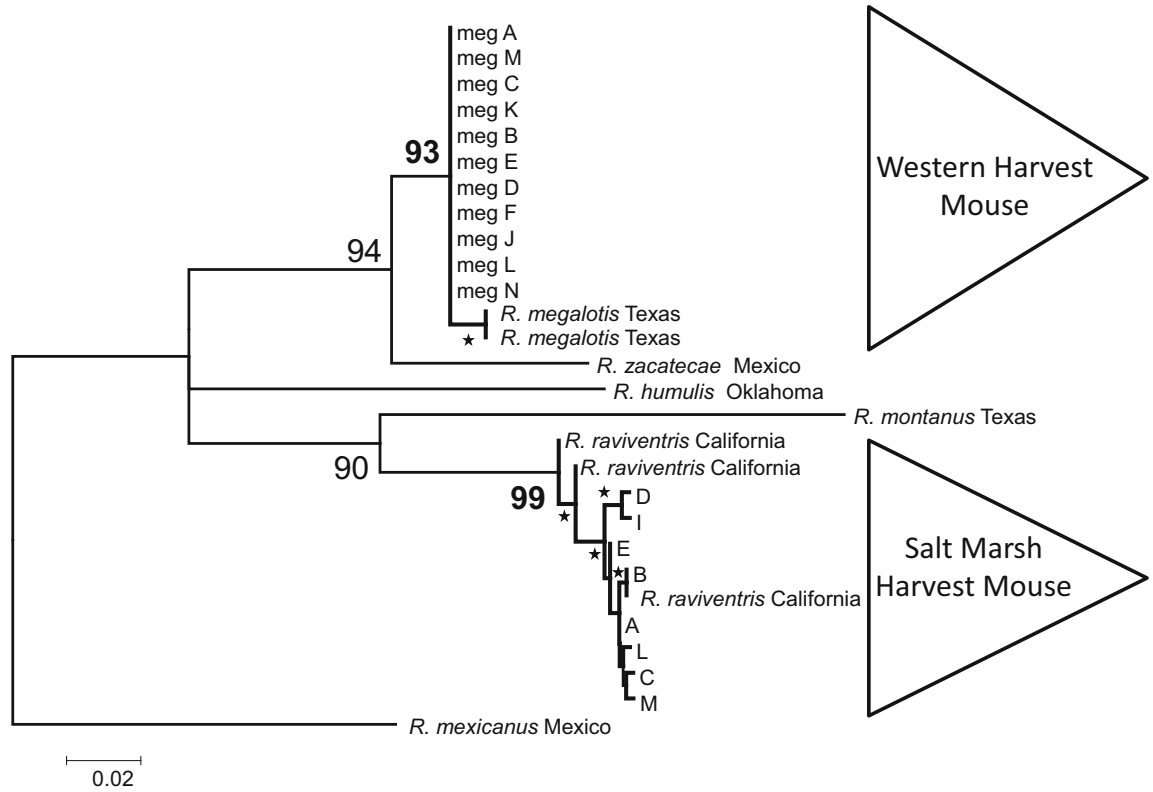

Fig. 2 Maximum likelihood consensus tree of salt marsh harvest mouse, western harvest mouse, and closely related members of the genus *Reithrodontomys*. Based on analyses of 403 bp of the cytochrome *b* gene, with 1000 BS replicates. The *asterisk* indicates nodal support of 68–94 %. Harvest mouse reference sequences are from Bell et al. (2001; AF176248-58), and Arellano et al. (2005; AY859470)

animal misidentified as WHM, and 6 animals unidentified to species. Species assignment of WHM based on morphology was much less successful, with 44 % (23 of 52) correctly identified and 27 % (14 of 52) misidentified as SMHM. In addition the proportion of harvest mice correctly assigned to species using morphology differed among the three bays. Correct species assignment was substantially lower in the San Francisco Bay (50.9 %) than either of the two northern bays (San Pablo, 100 %; Suisun, 84.6 %). The abundance of WHM relative to SMHM was similar in Suisun and San Francisco Bays (0.73:1, 0.9:1, respectively), suggesting greater difficulty in identifying harvest mouse species in the San Francisco Bay, which represented the southern portion of the SMHM range.

Salt marsh harvest mouse analyses

The SMHM Cyt-b haplotypes divided into two clades. Clade 1 was ubiquitous, while clade 2 was found only in San Pablo and Suisun bays (Fig. 4), indicating greater ancestral diversity in the northern portion of the species range. There was a greater degree of haplotypes sharing between the two northern Bays (San Pablo and Suisun), than between either of these Bays and the San Francisco Bay (Table 1). Individuals from San Francisco Bay had 3 closely related haplotypes, of which only 1, the ancestral A haplotype, was shared with the two northern bays. Haplotype A was only identified in a single San Francisco individual, while the remaining haplotypes (C and L) were found in multiple locations only in the San Francisco Bay (Fig. 2a). Analyses of the same individuals with the more highly resolving control region also identified only 3 haplotypes in San Francisco Bay, none of which were shared with populations in the northern bays (Fig. 4b). San Pablo and Suisun had substantially more control region haplotypes (11 and 13, respectively). Analyses of the combined mitochondrial gene regions (Cyt-b and control region) estimated that San Pablo and Suisun had nearly twice the haplotype diversity, and 4–5 times the nucleotide diversity found in San Francisco (Table 2).

Analyses of population subdivision using pairwise Φ_{ST} values indicated that all bays were significantly differentiated from one another (Table 3). Both mitochondrial gene regions indicated a substantially closer relationship between San Pablo and Suisun, than between either of these bays and San Francisco Bay. The control region haplotype network suggested little phylogenetic divergence between the two northern populations, including sharing of one haplotype.

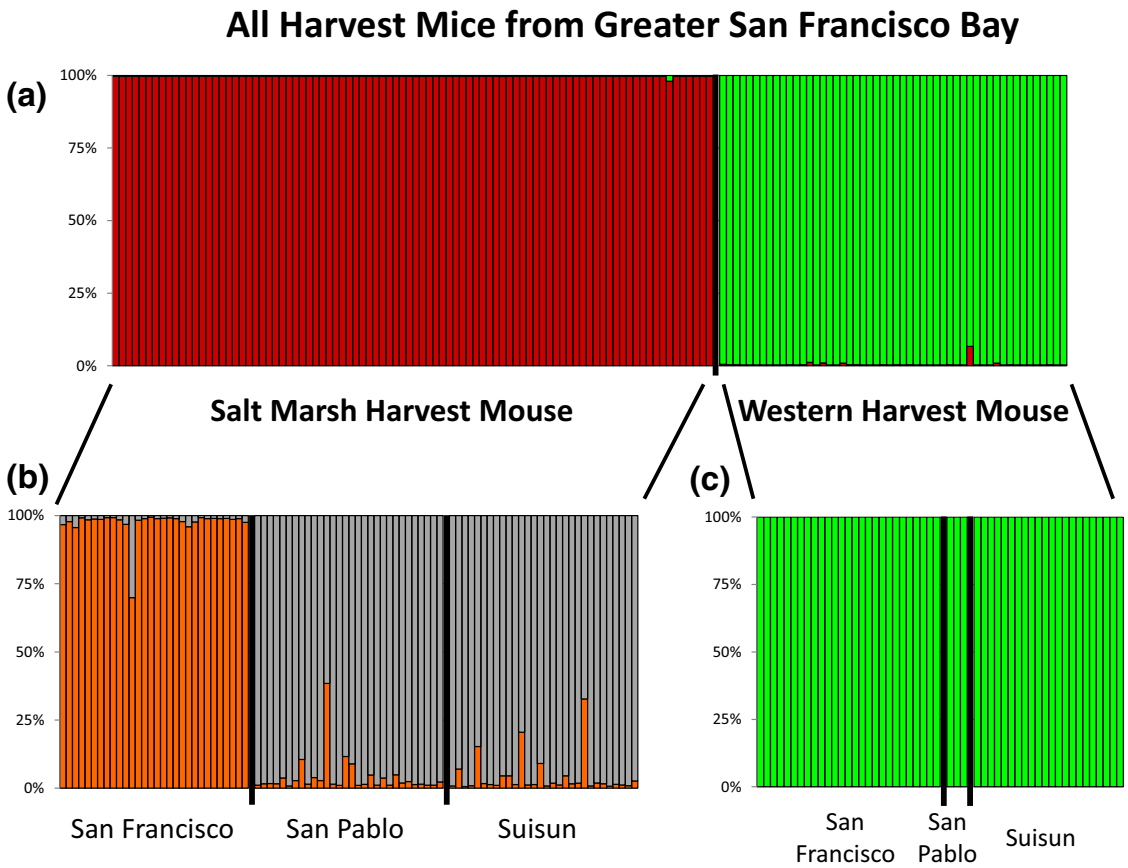

Fig. 3 Harvest mouse genetic subdivision as evident from microsatellite loci and cluster analysis using the program STRUCTURE. All analyses were performed without prior information regarding the population of origin or putative species. **a** Genetic cluster analysis of all harvest mice (based on 5 microsatellite loci) separated animals into two groups, salt marsh harvest mouse and western harvest mouse, consistent with the species assignment from mtDNA analyses. **b** Cluster analysis of the salt marsh harvest mouse (based on 8 microsatellite loci) separated animals from the south bay from those in the two northern bays, consistent with previously described subspecies. **c** There was no evidence of population subdivision within the western harvest mouse dataset (based on analyses of 6 microsatellite loci)

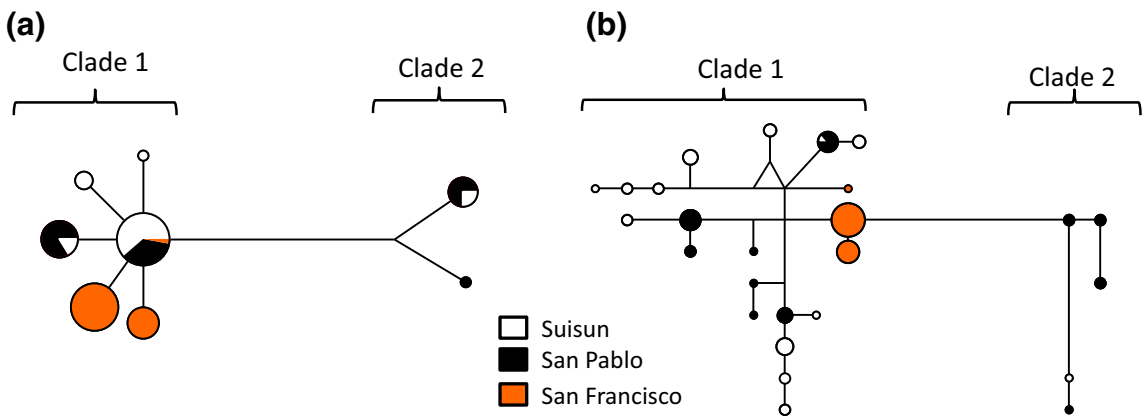

Fig. 4 Haplotype networks of salt marsh harvest mouse mtDNA. **a** Is based on 403 bp of cytochrome *b* from 90 animals. **b** Is based 785 bp of the control region from 87 animals. *Circle* sizes are proportional to the number of individuals represented

The San Francisco Bay population of SMHM exhibited the lowest Observed (H_o) and Expected Heterozygosities (H_e), and average number of alleles per locus (A; Table 4). This trend was particularly pronounced when the two northern bays were considered a single unit. Here the average number of alleles rarefied to control for sample

Table 2 Diversity statistics for combined mitochondrial cytochrome *b* and control region (1121 bp) within the salt marsh harvest mouse

Bay	*n*	Haplotypes	Haplotype diversity	SD	Nucleotide diversity	SD
San Pablo	29	11	0.89	0.03	0.0076	0.0040
Suisun	29	14	0.94	0.02	0.0066	0.0033
San Francisco	29	3	0.49	0.07	0.0014	0.0009

Table 3 Population subdivision within SMHM as evident with mtDNA and microsatellites

Population	Cytochrome *b*			Control region			Microsatellites		
	SP	SU	SF	SP	SU	SF	SP	SU	SF
San Pablo	–			–			–		
Suisun	0.09	–		0.11	–		0.019	–	
San Francisco	0.36	0.38	–	0.37	0.41	–	0.145	0.142	–

Values for mitochondrial regions indicate pairwise Φ_{ST}, while values for microsatellites indicate pairwise F_{ST} between the three bays

All pairwise comparisons were significant at $P \leq 0.05$

Table 4 Microsatellite diversity statistics within SMHM across eight polymorphic loci

Population	*n*	*A*	A_R^a	*Pr*	H_E	SD	H_O	SD	F_{IS}
San Pablo	30	4.9		3	0.63	0.06	0.56	0.03	0.11
Suisun	30	4.8		2	0.59	0.09	0.57	0.03	0.04
San Francisco	30	3.9	3.8	6	0.52	0.09	0.49	0.03	0.06
Northern 2 Bays	60	5.1	4.9	15	0.62	0.07	0.57	0.03	0.09

A average number of alleles per locus; A_R allelic richness, *Pr* private alleles, H_E expected heterozygosity, H_O observed heterozygosity, F_{IS} inbreeding coefficient

[a] Rarefaction estimate for a sample of 25 individuals

size indicated lower diversity in the San Francisco Bay relative to the northern two bays, in addition to a lower number of private alleles. Consistent with locus-specific tests of Hardy–Weinberg and linkage disequilibrium (above), all populations exhibited positive multilocus F_{IS} values, reflecting a deficit of heterozygotes apparently caused by population substructure. Similarly to mtDNA analyses, pairwise F_{ST} based on microsatellites revealed significant differentiation among all populations, with divergence of the San Francisco Bay population from the two northern populations (average $F_{ST} = 0.144$) greater than that between the two northern populations ($F_{ST} = 0.019$; Table 4).

Bayesian cluster analysis in the program STRUCTURE indicated that $K = 2$ had the highest support across five replicate runs. Without any prior geographic information this separated SMHM from the southern San Francisco Bay from those in the northern two bays (Fig. 3). Separate analyses conducted on SMHM from the southern San Francisco Bay, and from the combined samples of San Pablo and Suisun did not support additional subdivisions. Cluster analysis of the WHM dataset did not support subdivision of the population (Fig. 3).

Based on our coalescent analysis of Cyt-b data in IMa2, estimates of the splitting time between northern and southern populations ranged between from 3.7–194 kya, 5.6–291 kya, 11–582 kya, depending on the mutation rate assumed (3, 2, or 1 % per MY, respectively; Table 5). We estimated the ancestral effective size of the northern population at approximately 3 × that of southern population (Table 5).

Discussion

Our most significant findings were: (1) the clear and ancient difference between the sympatric and morphologically similar harvest mouse species; (2) the lack of hybridization between them; (3) the inadequacy of morphological tools to discriminate them in the field, especially in the range of the southern SMHM subspecies; and (4) the deep genetic subdivision between subspecies of the SMHM. Below, we discuss the evidence for and implications of each of these findings.

Distinguishing between harvest mouse species

Both mtDNA and nuclear DNA datasets clearly resolved SMHM and WHM species. We observed no haplotypes shared between species, and phylogenetic analyses of

Table 5 Mitochondrial DNA splitting time and demographic analyses of salt marsh harvest mouse populations in the North and South of the San Francisco Bay

Parameter	Mutation rate per million years		
	1 %	2 %	3 %
Splitting time	240,000	120,000	80,000
	(11,000–582,000)	(5,583–291,000)	(3,722–194,000)
Ne North Bay	609,000	304,000	203,000
	(123,000–1,195,000)	(61,000–597 000)	(41,000–398,000)
Ne South Bay	204,000	102,000	68,000
	(11,000–488,000)	(5,583–244,000)	(3722–163,000)
Ne Ancestral	1,136,000	568,000	378,000
	(0–4,076,000)	(0–2,038,000)	(0–1,359,000)

Analyses based on 403 bp of cytochrome *b* using the 'isolation only' model with the program IMa2. Parameter values were calculated using mutation rates of 1, 2, and 3 % per million years. The 95 % highest posterior density (HPD) intervals are given in parentheses after each of the parameter estimates

Cyt-b indicated that both species formed well resolved monophyletic lineages (Fig. 2). Moreover, the topology of the tree was consistent with previous analyses (Arellano et al. 2005) indicating that SMHM and WHM were not sister species. Rather the SMHM was the sister species to the plains harvest mouse (*R. montanus*) from the central U.S. and northern Mexico, and the WHM was the sister species to the Zacatecas harvest mouse (*R. zacatecae*) from Mexico. Coarse divergence time estimation indicated that SMHM and WHM most recently shared a common ancestor ~3–10 million years ago, during the Pliocene or Miocene, and, based on the phylogenetic tree, that ancestor would likely have been shared among multiple species within the genus *Reithrodontomys*. This estimate is similar to other divergence time estimates within the genus (Bell et al. 2001). Because mtDNA is a maternally inherited marker, we relied on autosomal nuclear microsatellites to assess the possibility of male mediated hybridization. Consistent with the mtDNA, microsatellites indicated unambiguous assignment of individuals to one or the other harvest mouse species. Despite their similar morphology and overlapping ranges our results make it clear that significant gene flow has not occurred historically between harvest mice species in the Bay Area.

The clear genetic differentiation between species allowed us to assess the efficacy of conventional species assignment based on morphology. Overall, morphological field identification of SMHM was more successful than identification of WHM. Specifically 92 % of SMHM, and 44 % of WHM were correctly identified. Our results indicated that a large portion (27 %) of WHM were incorrectly assigned as SMHM, whereas many WHM (29 %) were not assigned to one species or the other due to ambiguous characters. In areas of high WHM abundance such inaccurate species identification could falsely inflate the SMHM counts. This problem has potentially significant ramifications for ongoing surveying and monitoring efforts. We also noted a substantially lower species identification rate in the southern portion of the species range, despite a similar capture ratio of WHM to SMHM in Suisun Bay. In the North Bay (Suisun and San Pablo) tail length, and tail length to body length ratio (>79 mm and >115 %, respectively) are most heavily relied upon to identify SMHM (Brown 2003). In the South Bay (San Francisco) the species exhibits considerable morphological variation, including in coat color and tail length. Here the main characteristics used to differentiate species are tail diameter, pattern, coloration and shape (Shellhammer 1984). These characteristics are acknowledged to be difficult to use (Shellhammer 1984). With a success rate just 50.9 %, it is clear that the morphological species identification key used in the San Francisco Bay is insufficient to accurately discriminate between these species. Future efforts using genetically identified SMHM and WHM from the San Francisco Bay should be used to determine if there are distinguishing traits that can be used for morphological identification in the field.

Salt marsh harvest mouse

We identified major, long-standing, population differentiation within the salt marsh harvest mouse species. Multiple lines of evidence indicate that the species is divided into two distinct populations, one in the southern San Francisco Bay and one in the two northern Bays (San Pablo and Suisun). Most mtDNA haplotypes were found only in one population, with only a single, apparently ancestral, Cyt-*b* haplotype shared between northern and southern populations. Consequently we identified substantially higher Φ_{ST} values in pairwise comparisons between the San Francisco Bay population and either of the two northern bays relative to that between the two northern bays. The nuclear microsatellite dataset identified the same major subdivision within SMHM using both pairwise F_{ST} and

model based clustering methods. These patterns were consistent with previously recognized subspecies, *Reithrodontomys raviventris raviventris* from the San Francisco Bay and *R. r. halicoetes* from San Pablo and Suisun Bays.

The deep genetic subdivision between SMHM subspecies stands in contrast to patterns observed in other San Francisco Bay Area salt marsh endemics. Microsatellite analyses of morphologically distinct salt-marsh-associated subspecies of song sparrows (*Melospiza melodia*) indicated generally low levels of genetic differentiation and high gene flow across the San Francisco Bay Area, thus challenging current subspecies designations (Chan and Arcese 2002). Analyses of the endangered Ridgway's rail (*Rallus obsoletus*) within the San Francisco Bay found that population differentiation was strongly correlated with geographic distance, and that the single mtDNA haplotype present was shared with all other California subspecies (Chan et al. 2006; Maley and Brumfield 2013; Takekawa et al. 2014). The shallower subdivision in these bird species is likely affected by their relatively high dispersal ability. Thus the deeper subdivision we identify in the SMHM may be more representative of other, less vagile, marshland species.

The ancient split within the SMHM is more in line with pre-historical processes than with recent human induced habitat fragmentation. Even the extreme end of our 95 % HPD for splitting time, and assuming the upper end of reasonable mutation rates (3 % per MY), indicated that the split occurred at least 3700 years ago. More reasonably, our average estimates of the splitting time between northern and southern subspecies were on the order of tens to hundreds of thousands of years ago, during the last glacial period or earlier. This timeframe also agrees with our understanding of physical processes of the region during that period. During the Pleistocene, changes between glacial and interglacial periods led to fluctuations in sea level, which in turn impacted the extent and location of salt marshes of the San Francisco Bay. During the Wisconsin glaciation, when sea levels were considerably lower than the present, the area was a riparian valley through which the combined Sacramento and San Joaquin rivers flowed (Atwater 1979). Rising sea level after the last glacial maximum brought the sea into the golden gate around 10,000 years ago. Inundation slowed around 5000 years ago, at close to the present level, thus the current marshes surrounding the Bay stem from this time (Atwater 1979).

Endemism and diversity

Both SMHM subspecies exhibited private haplotypes and alleles, with the southern subspecies exhibiting substantially lower diversity in both genomic markers. Specifically, the multilocus nuclear microsatellite dataset indicated consistently lower values across all diversity estimates. The mitochondrial dataset supported this general trend, with the San Francisco Bay population having about 20 % of the nucleotide diversity found in either of the northern bays, thus reflecting a lower effective population historically. Such reduced genetic diversity has a bearing on the adaptive potential of the southern subspecies, which is particularly important in light of ongoing climate change. This concern is elevated by the reduced and fragmented nature of remaining salt marsh habitat in the southern San Francisco Bay (USFWS 2013).

Management implications and future work

Better tools for the differentiation of harvest mouse species are needed, especially in the range of the southern SMHM subspecies. Investigation of ultraviolet reflectance to differentiate between harvest mouse species has shown promise based on museum specimens (Ichiyanagi 2010). Further work involving live mice in the field (with genetic confirmation to species) will be necessary to determine the efficacy of this or other techniques. Until such tests are developed the use of genetic tools will be critical for the monitoring and surveying of the SMHM.

The magnitude of the genetic divergence we identified for the two SMHM subspecies was sufficient to warrant their consideration as evolutionary significant units (ESU, Moritz 1994, Paetkau 1999), which has important management implications. As suggested in guidelines set forth by U.S. Fish and Wildlife Service (2013), the range of genetic variation of the SMHM must be maintained to allow for evolution and response to environmental change. Thus, both units identified should be managed for persistence. This is important as the two subspecies have a degree of morphological differentiation and live in very different environments (water salinity and associated vegetation, as well as competitors and predators).

Although we have identified genetic subdivisions within the SMHM consistent with recognized subspecies the actual geographic distribution of each subspecies as well as strength of the subspecies boundary is unclear. The putative division between SMHM subspecies is considered to occur between San Rafael and Richmond Marsh (Fig. 1; Shellhammer 1989). Genetic sampling of individuals from these intervening marshes would allow assessment of whether the deep differentiation we observed corresponds to the putative subspecies dividing line, indicating a degree of reproductive isolation, or whether intervening populations facilitate gene flow between northern and southern populations more consistent with an isolation-by-distance pattern or porous boundary (i.e., hybrid zone).

Acknowledgments The primary funding for this research came from California Department of Fish and Wildlife Grant (P1282009). Additional funding came from the Veterinary Genetics Laboratory at UC Davis. Thank you for to Bill Burkhard (DWR), Peter Moyle (UC Davis) for provision of student funding. SMHM survey cooperators include Karen Taylor (DFW Napa/Sonoma Wildlife Area), Stacy Martinelli (DFW, Fagen Marsh), John Krause (DFW, Eden Landing Ecological Reserve), USFWS staff Joy Albertson (Don Edwards SFB NWR), Meg Marriott (USFWS, San Pablo Bay NWR), and Isa Woo (USGS, San Pablo Bay). Thank you to Natalie Goddard, Michelle Holtz, and Sini Reponen for their help in the field and laboratory. Thank you to two anonymous reviewers whose comments and suggestions improved the quality of this paper.

References

Allendorf FW, Luikart G (2007) Conservation and the genetics of populations. Blackwell publishing, Hoboken

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Arbogast BS (1999) Mitochondrial DNA phylogeography of the new world flying squirrels (Glaucomys): implications for pleistocene biogeography. J Mammal 80:142–155

Arellano E, Gonzalez-Cozatl FX, Rogers DS (2005) Molecular systematics of Middle American harvest mice *Reithrodontomys* (Muridae), estimated from mitochondrial cytochrome *b* gene sequences. Mol Phylogenet Evol 37:529–540

Atwater BF (1979) Ancient processes at the site of Southern San Francisco Bay: movement of the crust and changes in Sea Level. In: Conomos TJ (ed) San Francisco Bay: The urbanized estuary., San Francisco, Pacific Division, American Association for the Advancement of Science, pp 31–45

Bandelt HJ, Forester P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Bell DM, Hamilton MJ, Edwards CW, Wiggins LE, Martinez RM, Strauss RE, Bradley RD, Baker RJ (2001) Patterns of karyotypic megaevolution in *Reithrodontomys*: evidence from a cytochrome-b phylogenetic hypothesis. J Mammal 82:81–91

Brown SK (2003) Conservation genetics of salt marsh harvest mice (*Reinthrodontomys raviventris*) in Suisun Marsh, CA. Master Thesis, San Luis Obispo, Calif. California Polytechnic State University

Castro-Campillo A, Roberts HR, Schmidly DJ, Bradley RD (1999) Systematic status of *Peromyscus boylii* ambiguus based on morphologic and molecular data. J Mammal 80:1214–1231

Chan Y, Arcese P (2002) Subspecific differentiation and conservation of song Sparrows (*Melospiza melodia*) in the San Francisco Bay region inferred by microsatellite loci analysis. Auk 119:641–657

Chan YL, Hill CE, Maldonado JE, Fleischer RC (2006) Evolution and conservation of tidal-marsh vertebrates: molecular approaches. Stud Avian Biol 32:54–75

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–567

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

Fisler GF (1963) Effects of salt water on food and water consumption and weight of harvest mice. Ecology 44:604–608

Fisler GF (1965) Adaptations and speciation in harvest mice of the marshes of San Francisco Bay. Univ Calif Publ Zool 77:1–108. doi:10.2307/1932549

Goals Project (2000) Baylands ecosystem species and community profiles: Life histories and environmental requirements of key plants, fish and wildlife. Prepared by the San Francisco Bay Area Ecosystems Goals Project. In : P.R. Olofson (ed) San Francisco Bay Regional Water Quality Control Board, Oakland, California, USA

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Hewitt GM (2000) The genetic legacy of the quaternary ice ages. Nature 405:907–913

Hey J (2010) Isolation with migration models for more than two populations. Mol Biol Evol 27:905–920

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. Proc Natl Acad Sci USA 104:2785–2790

Ichiyanagi M (2010) The adaptive context of ultraviolet induced reflectance and fluorescence patterns within the order Rodentia. Master of Science Thesis, California State University, Sacramento

Maley JM, Brumfield RT (2013) Mitochondrial and next-generation sequence data used to infer phylogenetic relationships and species limits in the Clapper/King Rail complex. Condor 115:316–329

Méndez-Harclerode FM, Hanson JD, Fulhorst CF, Milazzo ML, Ruthven DC 3rd, Bradley RD (2005) Genetic diversity within the southern plains woodrat (*Neotoma micropus*) in southern Texas. J Mammal 86:180–190

Moritz C (1994) Defining 'evolutionary significant units' for conservation. Trends Ecol Evol 9:373–375

Nabholz B, Glémin S, Galtier N (2008) Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. Mol Biol Evol 25:120–130. doi:10.1093/molbev/msm248

National Research Council (2010) Sea level rise and the coastal environment. Advancing the science of climate change. The National Academies Press, Washington, DC, p 245

Paetkau D (1999) Using genetics to identify intraspecific conservation units: a critique of current methods. Conserv Biol 13:1507–1509

Park SDE (2001) The Excel microsatellite toolkit. Trypanotolerance in West African cattle and the population genetic effects of selection. PhD Thesis, University College Dublin

Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256

Pritchard JK, Wen W (2002) Documentation for structure software version 2. http://pritch.bsd.uchicago.edu

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Reponen SEM, Statham MJ, Barthman-Thompson L, Sacks BN (2014) Microsatellite primer development for the salt marsh harvest mouse (*Reithrodontomys raviventris*) and cross-amplification in the western harvest mouse (*R. megalotis*). Conserv Genet Resour 6:285–287. doi:10.1007/s12686-013-0120-5

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

Sambrook J, Russell D (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Shellhammer HS (1982) *Reithrodontomys raviventris*, vol 169., Mammalian speciesAmerican Society of Mammalogists, Washington, DC, pp 1–3

Shellhammer HS (1984) Identification of salt marsh harvest mice *Reithrodontomys raviventris*, in the field and with cranial characteristics. Calif Fish Game 70:113–120

Shellhammer HS (1989) Salt marsh harvest mice, urban development, and rising sea levels. Conserv Biol 3:59–65

Smith MF, Patton JL (1993) The Diversification of South American Murid Rodents: evidence from mitochondrial DNA sequence data for the Akodontine Tribe. Biol J Linn Soc Lond 50:149–177

Takekawa JY, Casazza ML, Overton CT, Bui TD, Vandergast A, Wood D, Estrella S (2014) Applied Studies on California Clapper Rail Population Dynamics in support of Comprehensive Multispecies Tidal Marsh Recovery. Unpublished Data Summary Report, U.S. Geological Survey, Western Ecological Research Center, Vallejo, Dixon, and San Diego, CA

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

U.S. Fish and Wildlife Service (2013) Recovery Plan for Tidal Marsh Ecosystems of Northern and Central California. U.S. Fish and Wildlife Service, Sacramento

Vázquez-Domínguez E, Espindola S (2013) Characterization of ten new microsatellite loci from the endangered endemic rodent, *Reithrodontomys spectabilis*. Conserv Genet Resour 5:251–253