FULL RESEARCH ARTICLE

Development of a morphological key for the southern salt marsh harvest mouse using genetically verified individuals

MARK J. STATHAM[1]*, LAUREEN BARTHMAN-THOMPSON[2], SUSAN FRESQUEZ[1], AND BEN N. SACKS[1,3]

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

[2] *Bay Delta Region, Suisun Marsh Monitoring and Compliance Unit, California Department of Fish and Wildlife, 2109 Arch Airport Road, Stockton, CA 95206-3992, USA*

[3] *Department of Health and Reproduction, School of Veterinary Medicine, University of California, One Shields Avenue, Davis, CA 95616-8744, USA*

*Corresponding Author: statham@ucdavis.edu

The salt marsh harvest mouse (SMHM; *Reithrodontomys raviventris*) is a state and federally listed endangered species endemic to the coastal marshes of the San Francisco Estuary of California. Of two subspecies, the southern (*R. r. raviventris*) is most endangered and lacks reliable morphological field tools to distinguish from the sympatric western harvest mouse (WHM; *R. megalotis*). We trapped and collected genetic samples and morphological data from 204 harvest mice from 14 locations from across the range of the southern SMHM. Genetic species identification indicated these to be composed of 48 SMHM and 156 WHM, which we compared at ten morphological characters. Most continuous characters overlapped between species. Color characters were significantly differentiated and we identified a number of species-specific diagnostic pelage categories in both species. A random forest analysis indicated that ventral coloration of the abdomen and the ventral tail hair color were the most useful for differentiating between species. We used these two morphological characters to develop a decision tree which correctly classified 94% of harvest mice to species with 99% accuracy. These findings suggest that our decision tree can be used to reliably identify the species of most harvest mice in the range of the southern SMHM, with a small proportion (6% in our study) needing genetic confirmation. The decision tree should be tested on additional harvest mice that were not used in its development, particularly from novel locations across the range.

Key words: decision tree, endangered species, field identification, genetic identification, morphology, random forest, *Reithrodontomys megalotis, Reithrodontomys raviventris*, salt marsh harvest mouse, western harvest mouse

_____

Fundamental to monitoring, management, and conservation of endangered species is the ability to identify them in the field. Correct identification is essential for accurate characterization of the species' range, habitat, abundance, demography, population trends, and dietary requirements (Smith et al. 2018; Sustaita et al. 2018; Quinn et al. 2019; Statham et al. 2019). Erroneous identification leads to a false characterization of these parameters, which in turn invalidates conservation assessments and other scientific conclusions.

The salt marsh harvest mouse (SMHM, *Reithrodontomys raviventris*) is an endemic species restricted to the tidal and brackish marshes of the San Francisco Estuary (including San Francisco, San Pablo, and Suisun Bays; Fisler 1965; Shellhammer 1989). Fragmentation and loss of ≥75% of the species' habitat through reclamation of tidal areas led to it being listed as endangered under both the California and U.S. endangered species acts (Shellhammer 1982, 1989; USFWS 1970). The species is composed of two genetically and morphologically distinct subspecies: the southern *R. r. raviventris* of San Francisco Bay, and northern *R. r. halicoetes* of San Pablo and Suisun Bays (Fisler 1965; Statham et al. 2016).

Across much of its range the SMHM is sympatric with the morphologically similar non-endangered western harvest mouse (WHM; *R. megalotis*). In contrast to the SMHM, the WHM is abundant and ranges across a vast area comprising much of western U.S. and extending into Canada and Mexico (Jameson and Peeters 2004). While the WHM is primarily a grassland species, it also inhabits the edges of salt and brackish marshes in the San Francisco Estuary (Fisler 1963). Most biologists have considerable difficulty distinguishing between the two species (Shellhammer 1984), which has led to a number of efforts to provide morphological tools to facilitate differentiation.

Shellhammer (1984) described a method for the differentiation of the species based on the work of Fisler (1965). Shellhammer (1984) assigned values to tail traits: diameter 20 mm from the base of the tail, tail color pattern (i.e., bicolored, unicolored, or intermediate), ventral tail hair color (i.e., tan, intermediate, white to grayish-white), and tail tip (i.e., blunt, intermediate, or pointed). Different sets of tail characters were examined in the northern part of the species range (i.e., San Pablo Bay, Suisun Bay) versus the southern (i.e., San Francisco Bay). Based on the total score animals were assigned to SMHM, WHM, or intermediate/unknown. Although other characters were recorded (i.e., behavior, ventral coloration, tail to body length ratio, and presence of orange ear tufts) they were not included in the scoring system for species assignment, resulting in some confusion regarding application of these criteria in differentiation of the species.

More recently, Sustaita et al. (2018) examined morphological characters of harvest mice from Suisun Bay (in the range of the northern SMHM subspecies) that were genetically confirmed to species. Using these data, they identified tail length, body length, and tail diameter as the most useful characters for distinguishing species. They then built a multiple logistic regression model that correctly classified 90.1% of individuals. However, the validity of these characters for the identification of harvest mice beyond Suisun Bay was untested.

Recently, different subsets of morphological measurements have been recorded by agency (i.e., U.S. Fish and Wildlife Service [USFWS], California Department of Fish and Wildlife [CDFW]) personnel working on harvest mice in the San Francisco Estuary. Use of genetic species identification revealed that morphological assignment methods were inadequate in the range of the southern SMHM, where only ~50% were correctly identified (Statham et al. 2016). This was the case regardless of whether mice were being identified using the tail scoring method (Shellhammer 1984; USFWS and others), or the multiple logistic regression (Sustaita et al, 2018; CDFW). This finding essentially invalidated cur-

rent morphological methods for the identification of harvest mice species within the range of the southern SMHM, leaving the identification reliant solely on genetic analyses. The finding also threw into doubt historical records of species abundance and distribution, as well as scientific study reliant on correct identification, thus hampering recovery efforts.

The objectives of this study were to (1) assess whether morphological characters could be used to reliably distinguish between harvest mice species within the range of the southern SMHM, and (2) develop a morphological method for the identification of harvest mice species. To achieve these aims, we collected morphological and genetic data, and used both single variable and multivariable analyses to determine if morphological characters could be used to reliably differentiate species.

METHODS

Study Area

We conducted our study at 14 tidal and adjacent diked saline marsh locations across the south and central San Francisco Estuary of California (Fig. 1; 37.6º N, 122.1º W). These marshes encompass the putative geographic range of the southern SMHM. The marshes generally contained middle and high marsh where the vegetation cover was dominated by pickleweed (*Salicornia pacifica*).

Fieldwork and Genetic Sample Collection

Animal trapping, handling, and genetic sampling procedures were approved by UC Davis Institutional Animal Care and Use Committee and authorized by the CDFW and the USFWS. We collaborated closely with partners (USFWS, CDFW, H. T. Harvey & Associates) to obtain harvest mouse morphological measurements and genetic samples. Data and samples from many of the mice were obtained during ongoing trapping and monitoring efforts. Additional sites that were not part of ongoing monitoring activities were chosen to provide a diverse set of sampling locations spread across the range of the southern SMHM (Fig. 1). We used ~70-130 Sherman live traps (H.B. Sherman Traps, Tallahassee, FL) spaced at ~10-m intervals, although the exact number of traps and the layout depended 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. When harvest mice were captured and processed 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 (Statham et al. 2016). We stored the hair in 95-100% ethanol until DNA extraction.

Morphological Measurements

We collected a suite of morphological measurements for each harvest mouse, including those considered previously (Shellhammer 1984; Sustaita et al. 2018). The majority of the morphological data was collected by experienced CDFW and USFWS personnel, although >10 observers of varying skill levels contributed to data collection. Specifically, we recorded the Sex, Reproductive Condition, Mass, Total Length, Tail Length, Body Length (i.e., calculated by subtracting the tail length from the total length), Tail Diameter (20mm from base
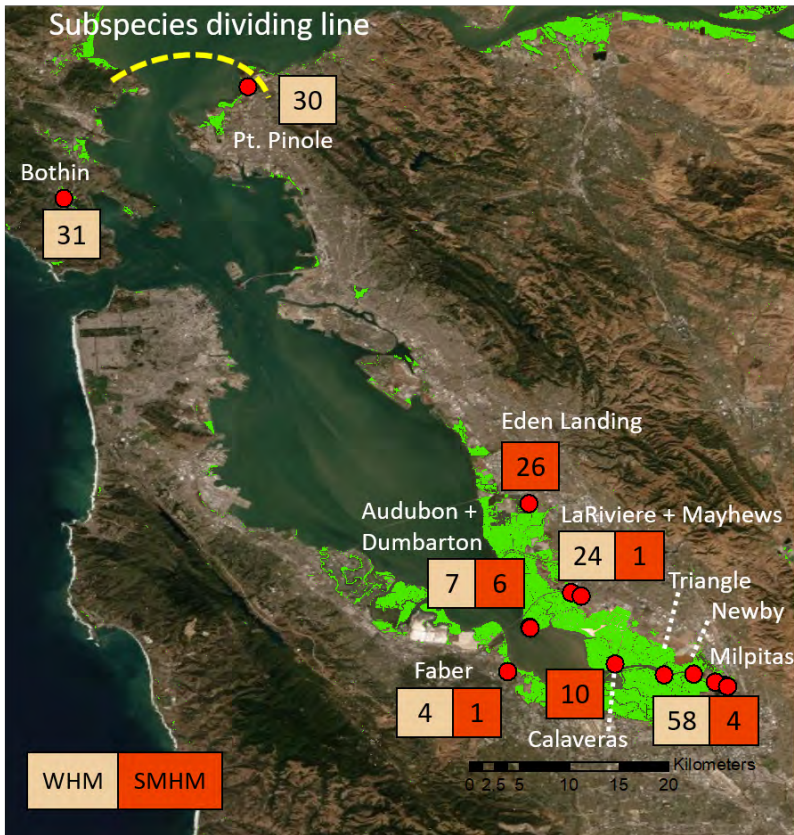
Figure 1. Map of genetically identified harvest mice. Numbers in the light-colored boxes are western harvest mouse (WHM; *Reithrodontomys megalotis*), numbers in the orange boxes are salt marsh harvest mouse (SMHM, *R. raviventris raviventris*). Red circles indicate trapping locations. Green highlighting indicates coastal wetlands.

of tail), Ventral Coloration (Fig. 2), Dorsal Hair Color(s), Tail Pattern (taking into account the hair color on both the dorsal and ventral surface of the tail [Fisler 1965] and categorized as bicolored, intermediate, unicolored), Ventral Tail Hair Color (i.e., all white, intermediate [few white hairs], tan [no white hairs]), Tail Tip Shape (i.e., pointed, intermediate, blunt), and Behavior (i.e., docile, intermediate, active). We did not conduct statistical analysis of Behavior or Dorsal Hair Color due to the relatively small number of records and the subjective nature of assigning animals to one category or another.

Genetic Species Identification

We used DNA sequence data to discriminate between harvest mouse species following the methods of Statham et al. (2016). Briefly, we extracted DNA from the mouse hair samples and then PCR-amplified and sequenced a 426 bp fragment of the cytochrome *b* gene and a small section of tRNA-Glutamate using the primers MVZ-05 and MVZ-04R (Smith and Patton 1993; Brown 2003). SMHM and WHM form well defined reciprocally monophyletic clades (Statham et al. 2016) at this gene region. We used the associated >10% sequence divergence between SMHM and WHM to discriminate between the species with certainty.
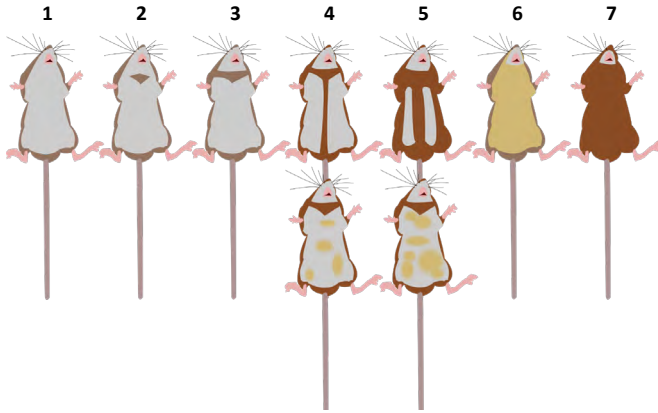
Figure 2. Ventral Coloration categories in harvest mice from Central and South San Francisco Estuary. The categories are as follows: (1) white or greyish white; (2) cinnamon pectoral spot; (3) pectoral band (i.e., spot extended across chest between forelimbs); (4) pectoral band with center stripe extending down belly or ¾ of the ventral surface white with mottled light cinnamon; (5) wider ventral stripe than 4 or ½ of the ventral surface white with mottled light cinnamon; (6) light cinnamon wash, paler than 7, does not have a band, stripe, or spot; (7) all cinnamon (with rare trace of pale color) with no abrupt color change from lateral (sides) to ventral (belly). Spots and stripes can vary in color intensity. Example photographs of each category are provided in the accompanying document. The depiction and description of Ventral Coloration categories are based on descriptions by Fisler (1965) and Shellhammer (1984), with new clarifications and additional description.

Statistical Analyses

For each of the continuous morphological characters (e.g. Body Length, Mass), we used t-tests to compare means. For categorical characters (e.g. Ventral Coloration, Tail Tip Shape) we used χ^2 tests of independence. Statistical analyses and graphing were conducted using R v.3.6 (https://www.r-project.org/). We conducted a multiple variable analysis to examine the importance of each of the morphological characters for the differentiation between species using Random Forest (Breiman 2001) implemented in the R package randomForest (https://cran.r-project.org/web/packages/randomForest/index.html). Random forest is a machine learning approach for creating classification trees. Decision trees are generated for each bootstrapped sample (i.e., a random sampling of individuals with replacement) where a random subset of predictor variables is assessed at each node. The best performing predictor variable is chosen for use at each node. Each decision tree attempts to classify data points at each of the nodes and assesses the information gain. This process continues with additional nodes until there is no further information gain. The information gain at each node in each tree can then be used to assess which predictor variables are important (most informative) for the correct classification. Individuals are then run through each of the trees, and the aggregation decision (the majority) is used to classify them (termed 'bagging').

Our analysis was limited to adult individuals with no missing data, resulting in a dataset of 129 individuals. We randomly assigned 70% of individuals to a training dataset and 30% of individuals to a test dataset. We created a random forest model with the training data set, generating 500 trees (sampling individuals from the dataset with replacement), and randomly choosing 6 (of 10) predictor variables (i.e., morphological characters) at each split. Based on the model (containing all 500 decision trees) we assessed the classification error overall and for each species. We then assessed how well the model predicted species in the test dataset.

We calculated the importance of each of the 10 characters in the model using two different parameters: 'mean decrease in accuracy' and 'mean decrease in Gini'. The resolving power of a character was calculated as the 'mean decrease in accuracy' when a character is left out. The additive value of individual characters in the model is assessed using 'mean decrease in Gini' when that character is left out. Gini is effectively a measure of how homogenous or pure a group is at the end of the tree. The results of the single-variable and multivariable analyses were used to design a decision tree for the identification of harvest mice species in the field. We also took into consideration how subjective assignment to categories were for morphological characters.

For comparison to our southern-specific method, we also assigned these mice using the tail character scoring method of Shellhammer (1984), which was developed for both subspecies, and the multiple logistic regression method, which was developed on a population of the northern SMHM subspecies (Sustaita et al. 2018). Only adult mice with the full set of necessary characters were considered for each analysis.

RESULTS

Genetic Species Identification

We collected morphological data and genetic samples from 204 harvest mice from 14 sites within the putative range of the southern SMHM (Fig 1; Table 1). Using mitochondrial DNA sequence analysis, we identified 48 SMHM and 156 WHM. Despite catching 30 and 31 WHM respectively at Point Pinole Regional Shoreline (Contra Costa County) and

Table 1. Numbers of genetically identified salt marsh harvest mouse (SMHM; *Reithrodontomys raviventris raviventris*) and western harvest mouse (WHM; *R. megalotis*) from each location. Only mice that were included in statistical analyses are listed.

Site	WHM	SMHM
Audubon Marsh (Alameda County)	4	3
Bothin Marsh (Marin County)	29	0
Calaveras Marsh (Alameda County)	0	9
Eden Landing Ecological Reserve (Alameda County)	0	26
Dumbarton Marsh (Alameda County)	3	3
LaRiviere Marsh (Alameda County)	2	0
Mayhews Landing (Alameda County)	21	1
Milpitas - Coyote Creek Upland Ruderal Field (Santa Clara County)	10	0
Milpitas - Triangle Marsh (Santa Clara County)	0	2
Milpitas - Coyote Creek Reach 1A (Santa Clara County)	3	1
Milpitas - Lower Coyote Creek (Santa Clara County)	5	0
Newby Island (Santa Clara County)	36	0
Faber Marsh (San Mateo County)	4	1
Point Pinole Regional Shoreline (Contra Costa County)	31	0
Total	148	46

Bothin Marsh (Marin County), we did not catch any SMHM. All the genetically identified SMHM were restricted to the southern end of San Francisco Bay. We verified the continued existence of SMHM at Eden Landing (Alameda County), Dumbarton Marsh (Alameda County) Audubon Marsh (Alameda County), Mayhew's Landing (Alameda County), Triangle Marsh (Santa Clara County), and Milpitas (Santa Clara County) on the east shore of the South San Francisco Bay, and at Faber Marsh (San Mateo County) on the west shore of the South San Francisco Bay.

Single-variable Analyses

We obtained largely complete sets of measurements on 186 adult harvest mice (140 WHM; 46 SMHM). Preliminary analyses did not identify significant differentiation between the sexes within species. Therefore, we grouped by species for subsequent analyses. We identified overlapping ranges of measurement for both species for all six continuous variables (Fig. 3). Tail Length and Total Length did not differ significantly between the species
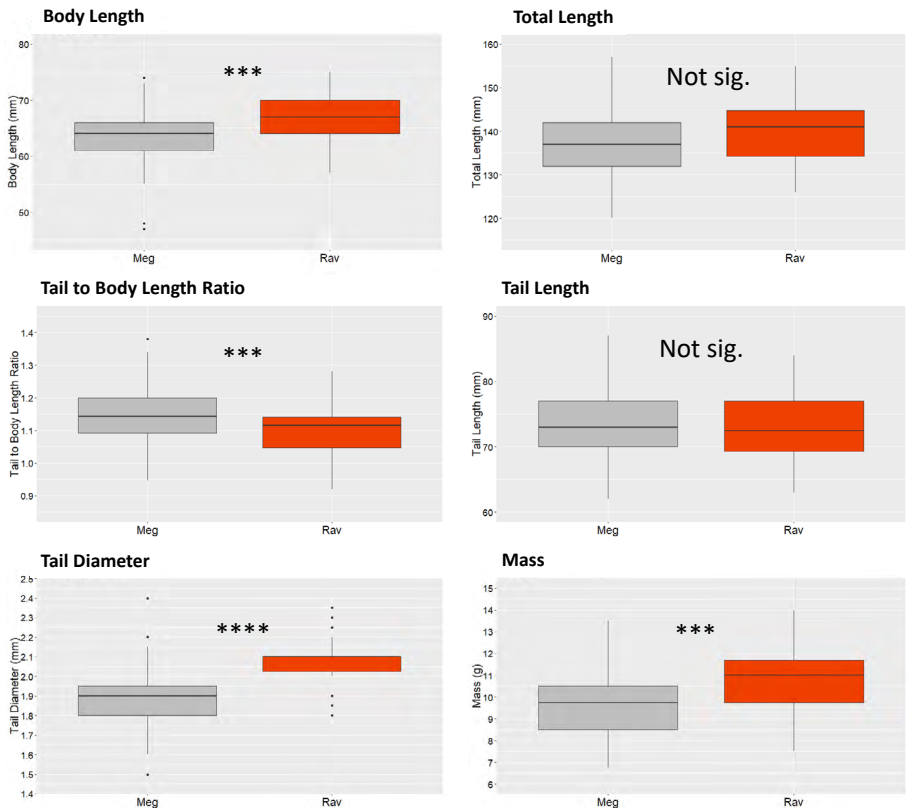

Figure 3. Box plots displaying the range of measurements for each of the continuous variables in harvest mice from the Central and South San Francisco Estuary. The significance of a t-test is indicated: WHM = western harvest mouse (*Reithrodontomys megalotis*), SMHM = salt marsh harvest mouse (*R. raviventris raviventris*), Not sig. = not significant, *** = significant at P < 0.001, **** = significant at P < 0.0001.

(Table 2). Body Length, Tail to Body Length Ratio, Tail Diameter, and Mass all differed significantly between species (Table 2). Tail Diameter was the most significantly differentiated continuous character between species, where SMHM averaged 2.1 (SD 0.12) mm and WHM averaged 1.9 (SD 0.14) mm.

We identified significant differences among categorical measurements between harvest mice species (Table 3). For Tail Tip Shape and Tail Pattern, all categories were shared between species, although the proportion of each species in each category differed (Fig. 4). For Ventral Tail Hair Color one of three categories was restricted to a single species. Specifically, all harvest mice with white Ventral Tail Hair Color were WHM, thus making this a diagnostic character for the species. For Ventral Coloration five of seven categories were restricted to a single species (Fig. 5). All mice with the lightest Ventral Coloration (categories 1, 2, and 3) were WHM, and all mice with the darkest Ventral Coloration (category 7) were SMHM. Category 5 was rare and only identified in SMHM. Together the four most common Ventral Coloration categories were diagnostic for species in 86% (143 of 166 mice) of harvest mice with Ventral Coloration data.

Table 2. Analysis of continuous morphological characters for salt marsh harvest mouse (SMHM; *Reithrodontomys raviventris raviventris*) and western harvest mouse (WHM; *R. megalotis*) in the South San Francisco Estuary. df and P are from a t-test between SMHM and WHM for each morphological character.

Character	SMHM mean	SMHM SD	WHM mean	WHM SD	df	P
Tail length (mm)	73.00	4.44	73.30	4.52	75	0.690
Body Length (mm)	66.46	4.18	63.75	4.61	81	< 0.001
Total length (mm)	139.46	7.16	137.07	7.48	80	0.057
Tail to Body Length Ratio	1.10	0.08	1.16	0.09	90	< 0.001
Tail Diameter (mm)	2.09	0.12	1.86	0.14	76	< 0.001
Mass (g)	11.24	1.90	10.02	1.80	62	< 0.001
Mass (g) without pregnant mice	10.77	1.56	9.70	1.44	51	< 0.001

Table 3. Comparison of categorical morphological characters between salt marsh harvest mouse (*Reithrodontomys raviventris raviventris*) and western harvest mouse (*R. megalotis*) in the South San Francisco Estuary.

Character	χ^2	df	P
Tail Tip Shape	32.36	2	< 0.001
Ventral Tail Hair Color	140.15	2	< 0.001
Tail Pattern	112.6	2	< 0.001
Ventral Coloration	139.05	6	< 0.001

Figure 4. The frequency of harvest mice in each of the categories of Tail Tip Shape, Ventral Tail Hair Color, and Tail Pattern in the Central and South San Francisco Estuary. The significance of a χ^2 test is indicated. Grey = western harvest mouse (*Reithrodontomys megalotis*), Orange = salt marsh harvest mouse (*R. raviventris raviventris*), **** = significant at P < 0.0001
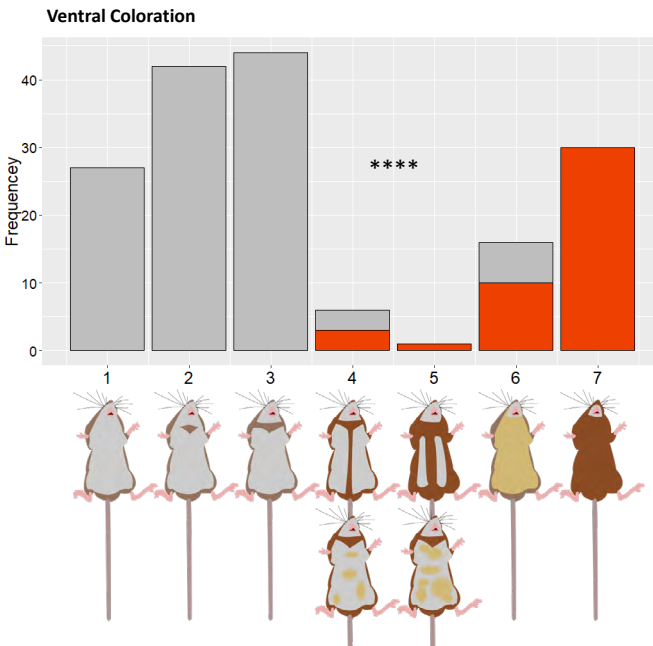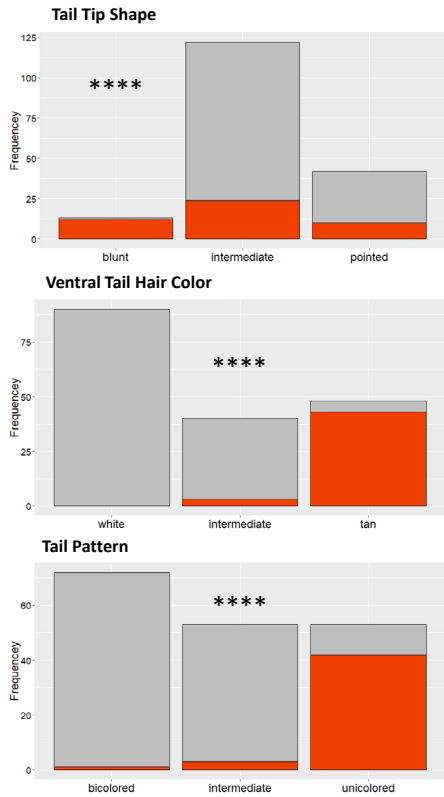
Figure 5. The frequency of Ventral Coloration categories in harvest mice from Central and South San Francisco Estuary. Orange = salt marsh harvest mouse (*R. raviventris raviventris*), Grey = western harvest mouse (*R. megalotis*). The significance of a χ^2 is indicated. **** = significant at P < 0.0001.

Multivariable Analysis

On the training dataset the model of 500 decision trees achieved a 97.8% accuracy rate (90 of 92 mice correctly assigned). When we applied the model to the test dataset (i.e., mice excluded from model training) it achieved a similarly high accuracy rate with 97.3% (31 of 32 mice) correctly assigned. We then assessed which variables were most important for classification (Fig. 6). For mean decrease in accuracy, the greatest drop in correct classification occurred when Ventral Tail Hair Color and Ventral Coloration were excluded. Excluding Tail Pattern, Tail Tip Shape, and Tail Diameter also resulted in some loss of accuracy. Exclusion of the remaining five characters (Total Length, Mass, Body Length, Tail to Body Length Ratio, and Tail Length) did not have a substantial impact on the model accuracy, indicating that these characters were not important for classification of harvest mouse taxa in the range of southern SMHM. The mean decrease in Gini largely identified the same most important characters, however Tail Diameter was identified as more informative than Tail Tip Shape. The same following characters were largely identified as uninformative: Total Length, Mass, Body Length, Tail to Body Length Ratio, Tail Length, and Tail Tip Shape.
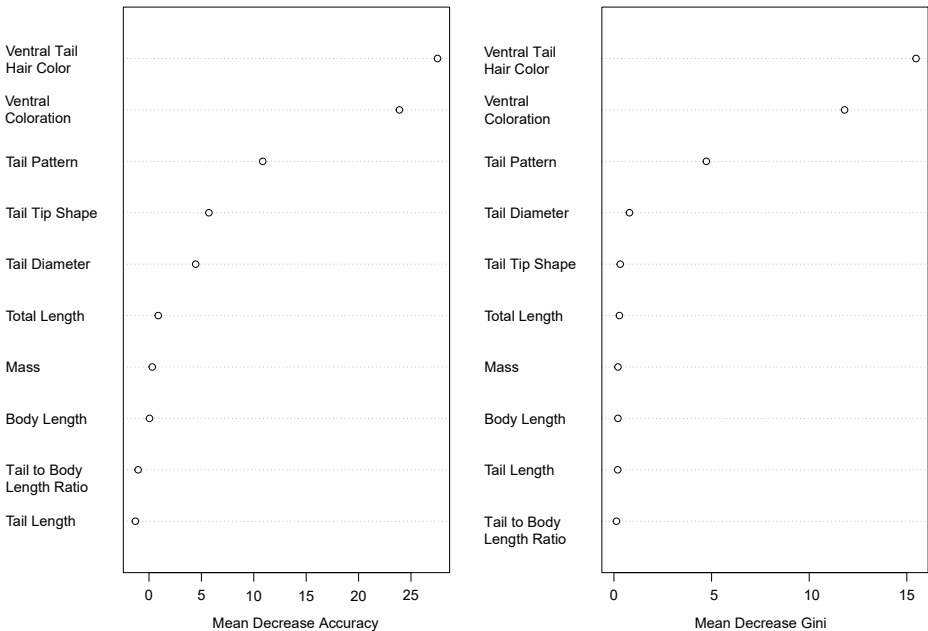
Morphological Characteristic Importance

Figure 6. Morphological character importance for the differentiation of southern salt marsh harvest mouse (*Reithrodontomys raviventris raviventris*), and western harvest mouse (*R. megalotis*) in the Central and South San Francisco Estuary. This analysis is based on results of a random forest analysis. Here 'Mean Decrease' indicates the reduction of resolving power if a character is not used for classification.

Decision Tree

We used the two most informative characters to design a decision tree for the identification of both species in the field (Fig. 7). Ventral Coloration was the primary character used because it was highly ranked in the multivariable analysis, it had well defined categories (Fig. 2), and on its own it was diagnostic for 86% of harvest mice sampled. We used Ventral Tail Hair Color as the secondary character because it was also highly ranked in the multivariate analyses, and white Ventral Tail Hair Ventral Color was diagnostic for WHM. The majority of SMHM had tan Ventral Tail Hair Color and overall it was 90% accurate for the identification of SMHM. Applying the decision tree to our dataset correctly identified 94% (169 of 179) of harvest mice to species, while the remaining 6% were unassigned.

For comparison, application of the tail character scoring method of Shellhammer (1984) correctly identified 79% of 146 individuals to species. The multiple linear regression method developed on a population of the northern subspecies (Sustaita et al. 2018) correctly identified 51% of 155 individuals to species.
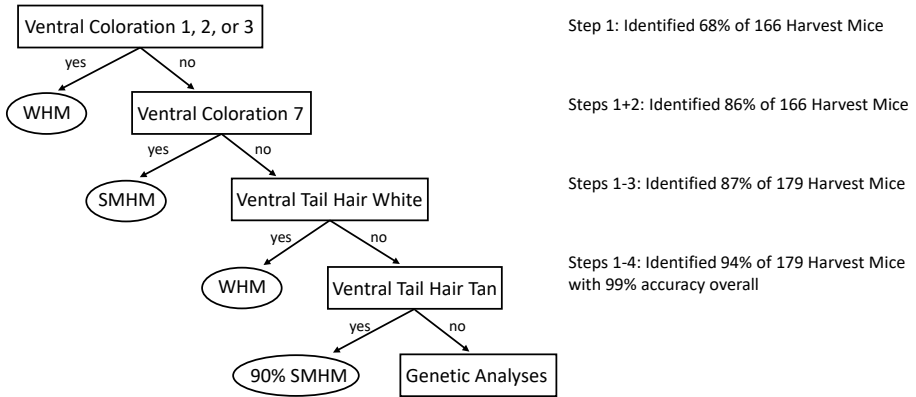

Figure 7. Decision tree for the identification of salt marsh harvest mouse (SMHM; *Reithrodontomys raviventris raviventris*) and western harvest mouse (WHM; *R. megalotis*) in the Central and South San Francisco Estuary. The tree uses just two morphological characters to assign 94% of mice to species with 99% accuracy. Harvest mice that are not assigned to species based on these characters should be identified using genetic analyses. This decision tree was validated only for the populations in this study. Application to other populations is not appropriate unless validated.

DISCUSSION

We identified a suite of morphological characters that can be used to reliably distinguish SMHM and WHM in the southern San Francisco Estuary. Multivariable analyses identified Ventral Tail Hair Color and Ventral Coloration as the most important characters, followed by Tail Pattern, Tail Tip Shape, and Tail Diameter. Single-variable analyses broadly agreed and identified the three color characters as the most differentiated between species, followed by Tail Diameter, then Tail Tip Shape.

Species-specific morphological categories were identified in Ventral Coloration and Ventral Tail Hair Color, thus providing diagnostic characters for the identification of harvest mouse species in the field. SMHM and WHM exhibited Ventral Coloration on opposite ends of the spectrum. All mice with the three lightest categories of Ventral Coloration (1, 2, and 3) were WHM, while all mice with the darkest Ventral Coloration (category 7) were SMHM, which together identified 86% of mice to species. Similarly, all mice with white Ventral Tail Hair Color were WHM, which was diagnostic for 67% of western harvest mice, while 94% of SMHM had tan Ventral Tail Hair Color. However, a small proportion of WHM were also identified with a tan tail, thus making tan tail 90% accurate for the identification of SMHM. Tail Pattern was also a very useful character; only a single SMHM out of 72 total harvest mice (1.4%) was identified with a bicolored tail. Among the continuous variables, Tail Diameter was the most differentiated between species. However, tail diameter is one of the most difficult dimensions to measure accurately in the field (Sustaita et al. 2018).

Ventral Coloration and Ventral Tail Hair Color identified the majority of harvest mice to species and therefore were used to design a decision tree for the dual identification of both harvest mouse species in the field. This tree identified 94% (169 of 179) of harvest mice with an overall accuracy of 99%. This is a substantial improvement over other methods for the morphological differentiation of the species in the Southern San Francisco Estuary. It is also simpler, being based on two characters that are easier to assess, as opposed to four or more.

Additional characters could be used for further classification to species of the remaining mice. However, each of the subsequent assignments would be probabilistic because of overlap in measurements or categories between species. This would result in a tentative assignment for the remaining 6% of harvest mice in the dataset. For this reason, we recommend that genetic analyses be used for species-identification of unresolved individuals.

The decision tree was validated only for adult individuals in the populations in this study. The northern SMHM subspecies and the sympatric WHM are known to display different morphological relationships than the southern populations studied here (Fisler 1965; Shellhammer 1984). Therefore, application to the northern SMHM subspecies is not appropriate. Moreover, we were unable to include SMHM specimens from the northern portion of the southern SMHM subspecies' range. Therefore, if putative SMHM are discovered in the northern portion of the southern subspecies range, genetic analyses should be used to classify them unless and until the decision tree is validated on those populations. Additionally, it would be useful to assess how well our approach works on non-adults.

Our findings were consistent with previous analysis of Ventral Coloration in San Francisco Estuary harvest mice (Fisler 1965), suggesting that the differences observed are longstanding characters of the populations. Stability of this character over time strengthens its use as the primary means for differentiating between harvest mouse species in the southern San Francisco Estuary. Fisler (1965) did not identify any WHM with red bellies (Category 7), but he did identify WHM with category 6 bellies on the Marin coast, which is consistent with our results. Shellhammer (1984) stated that all SMHM had red bellies (categories 5, 6, and 7), while WHM fell into categories ≤ 3, which is also close to our own findings. Shellhammer (1984) noted that color of the hair on the ventral portion of the tail and the tail diameter are easier to judge than the pattern of the tail and the tail tip. Thus, the characters we have selected for the decision tree were previously shown to be differentiated between species, and are among the easier ones to assign correctly.

Geographic Range of the Subspecies

Despite trapping across the putative range of the subspecies, SMHM were only identified in the marshes at the southern end of the San Francisco Estuary. The range of the southern SMHM subspecies is considered to extend north to the Central San Francisco Estuary near Point Pinole (Contra Costa County) and Point San Pedro (Marin County) (Fisler 1965; Shellhammer 1989; Fig. 1). Although we trapped >60 harvest mice in the northern end of the range (at Bothin Marsh and Point Pinole Regional Shoreline), all of them were genetically identified as WHM, suggesting that SMHM populations at these sites may be much reduced or extirpated. The remaining marshes in the Central San Francisco Estuary tend to be relatively small and fragmented, thus making resident SMHM populations more susceptible to extirpation. A survey in 2014 by the U. S. Geological Service at one such location on Marin County coast (Corte Madera, south of Pt. San Pedro, within the San Francisco Bay) did not result in any SMHM captures. Additional surveys should focus on remaining marshes in the Central San Francisco Estuary to determine if SMHM are still present.

Recommendations

Future surveys of SMHM in the central and southern San Francisco Estuary should assess the performance of the decision tree. The decision tree should also be validated for application to harvest mice from novel locations across the southern SMHM range. The latter point is especially important in the area just south of the putative subspecies dividing line, where no SMHM were identified in the current study. SMHM were considered present historically (Fisler 1965), yet their continued presence has not been confirmed using genetically verified animals. Similarly, the subspecific status of SMHM in the area has never been assessed using genetic analyses. Therefore, we do not know whether SMHM historically or currently present in the northern end of the southern range belong to the northern subspecies, the southern subspecies, represent an intermediate between the two, or are part of a distinct lineage. Therefore, we also do not know which field identification method is the most appropriate for species identification in those locations. Our analyses only included a single SMHM from the western shore of the south San Francisco Bay necessitating further testing and potential refinement for use on harvest mice on the western shoreline.

Assigning the color characters of mice to categories can be partly subjective. To minimize observer biases, we recommend using a card, with white and tan colors, for improved assignment of Ventral Tail Hair Color. Similarly, a card for Ventral Coloration would aid in the consistency of assignment, especially for mice with fully-belly coloration (categories, 1, 6, and 7). Further, the decision tree could be validated using museum specimens that have been genetically confirmed to species. However, an assessment would need to be made to determine how well the color characters are maintained in preserved skins. Additionally, because both Ventral Coloration and Ventral Tail Hair Color were recorded in earlier surveys; the decision tree could be applied to those records to reassess historical population trends. Information from the museum specimens and historical surveys would provide an updated assessment of the SMHM presence, range, abundance, and population trends.

We have developed a simple and accurate method for the field identification of harvest mice in the southern San Francisco Estuary. The previous method required a greater number of morphological measurements and still resulted in a substantial proportion of mice unassigned to species. Following years of uncertainty, our improved and simplified field identification method will aid conservation efforts and enhance recovery of the endangered SMHM.

ACKNOWLEDGMENTS

LITERATURE CITED

Breiman, L. 2001. Random Forests. Machine Learning 45:5.

Brown, S. K. 2003. Conservation genetics of salt marsh harvest mice (*Reithrodontomys raviventris*) in Suisun Marsh, CA. Thesis, California Polytechnic State University, San Luis Obispo, CA, USA.

Fisler, G. F. 1963. Effects of salt water on food and water consumption and weight of harvest mice. Ecology 44:604–608.

Fisler, G. F. 1965. Adaptations and speciation in harvest mice of the marshes of San Francisco Bay. University of California Publications in Zoology 77:1–108.

Jameson, E. W., and H. J. Peeters. 2004. Mammals of California. University of California Press, Berkeley, CA, USA.

Quinn, C. B., P. B Alden, and B. N. Sacks. 2019. Noninvasive sampling reveals short-term genetic rescue in an insular red fox population. Journal of Heredity. 110:559–576.

Shellhammer, H. S. 1982. *Reithrodontomys raviventris*. Mammalian Species 169:1–3.

Shellhammer, H. S. 1984. Identification of salt marsh harvest mice, *Reithrodontomys raviventris*, in the field and with cranial characters. California Fish and Game 70:113–120

Shellhammer, H. S. 1989. Salt marsh harvest mice, urban development, and rising sea levels. Conservation Biology 3:59–65.

Smith, K. R., M. K. Riley, L. Barthman-Thompson, I. Woo, M. J. Statham, S. Estrella, and D. A. Kelt. 2018. Toward salt marsh harvest mouse recovery: a review. San Francisco Estuary and Watershed Science. 16:1–24.

Smith, M.F., and J. L. Patton. 1993. The diversification of South American murid rodents: evidence from mitochondrial DNA sequence data for the akodontine tribe. Biological Journal of the Linnean Society 50:149–177.

Statham, M. J., S. Aamoth, L. Barthman-Thompson, S. Estrella, S. Fresquez, L. D. Hernandez, R. Tertes, and B. N. Sacks. 2016. Conservation genetics of the endangered San Francisco Bay endemic salt marsh harvest mouse (*Reithrodontomys*

raviventris). Conservation Genetics 17:1055–1066.

Statham, M. J., D. Woollett (Smith), S. Fresquez, J. Pfeiffer, J. Richmond, A. Whitelaw, N. Richards, M. F. Westphal, and B. N. Sacks. 2019. Noninvasive identification of herpetofauna: pairing conservation dogs and genetic analysis. Journal of Wildlife Management. 84:66–74.

Sustaita, D., S. K. Brown, L., Barthman-Thompson, S Estrella, L. Patterson, P. Q. Finfrock, and F. X. Villablanca. 2018. Distinguishing between northern salt marsh and western harvest mice. Journal of Wildlife Management 82:723–733.

U.S. Fish and Wildlife Service (USFWS). 1970. Appendix D—United States List of Endangered Native Fish and Wildlife. Federal Register 35:16047–16048.