Comprehensive genetic analyses reveal evolutionary distinction of a mouse (*Zapus hudsonius preblei*) proposed for delisting from the U.S. Endangered Species Act

Tim L. King[1,3], John F. Switzer[1], Cheryl L. Morrison[1],
Michael S. Eackles[1], Colleen C. Young[1], Barbara A. Lubinski[1],
and
Paul Cryan[2]

[1]US Geological Survey - Biological Resources Division
Leetown Science Center - Aquatic Ecology Branch
11649 Leetown Rd.
Kearneysville, WV 25430
USA

[2]USGS-BRD
Fort Collins Science Center
2150 Centre Avenue, Building C
Fort Collins, CO 80526
USA

Keywords: *Zapus hudsonius*, subspecies, phylogeography, control region, cytochrome *b*, microsatellite

Running title: evolutionary distinction of *Zapus hudsonius preblei*

[3]All correspondence to: Tim L. King, Phone: 304-724-4450; Facsimile: 304-724-4424;
E-mail: tlking@usgs.gov

ABSTRACT

Zapus hudsonius preblei, listed as threatened under the U.S. Endangered Species Act (ESA), is one of 12 recognized subspecies of meadow jumping mice found in North America. Recent morphometric and phylogenetic comparisons among *Z. h. preblei* and neighboring conspecifics questioned the taxonomic status of selected subspecies, resulting in a proposal to delist the *Z. h. preblei* from the ESA. We present additional analyses of the phylogeographic structure within *Z. hudsonius* that calls into question previously published data (and conclusions) and confirms the original taxonomic designations. A survey of 21 microsatellite DNA loci and 1380 base pairs from two mitochondrial (mt) DNA regions (control region and cytochrome *b*) revealed that each *Z. hudsonius* subspecies is genetically distinct. These data do not support the null hypothesis of a homogeneous gene pool among the five subspecies found within the southwestern portion of the species' range. The magnitude of the observed differentiation was considerable and supported by significant findings for nearly every statistical comparison made, regardless of the genome or the taxa under consideration. Structuring of nuclear multilocus genotypes and subspecies-specific mtDNA haplotypes corresponded directly with the disjunct distributions of the subspecies investigated. Given the level of correspondence between the observed genetic population structure and previously proposed taxonomic classification of subspecies (based on the geographic separation and surveys of morphological variation), we conclude that the nominal subspecies surveyed in this study do not warrant synonymy, as has been proposed for *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius*.

INTRODUCTION

Zapus hudsonius preblei is one of 12 recognized subspecies of meadow jumping mice found throughout North America (Hafner et al. 1981). The distribution of *Z. h. preblei* is confined to the riparian systems where moisture is most plentiful (Jones et al. 1983). At present, the subspecies' range stretches along the eastern slope of the Rocky Mountains from east-central Wyoming south to Colorado Springs, Colorado. The availability of suitable riparian habitat is declining throughout the range of *Z. h. preblei* due to degradation caused by agricultural, residential, and commercial development (U.S. Fish and Wildlife Service 1998). As a result of diminishing suitable riparian habitat and small population sizes, *Z. h. preblei* was listed as a threatened species under the U.S. Endangered Species Act (ESA) in 1998 (U.S. Fish and Wildlife Service 1998).

Effective programs for conserving threatened taxa require the identification of unambiguous units of management that reflect evolutionarily important lineages. The issue of defining appropriate units of management becomes acute when protection of a taxon under consideration relies on proof of distinction. Subspecies recognition within *Z. hudsonius* has been based primarily on geographic disjunction and morphological variation (Krutzsch 1954, Hafner et al. 1981) among relatively recently radiated populations. Given that morphological comparisons among accepted species of the genus *Zapus* have failed to provide reliably diagnostic characters (Jones 1981, but see Conner and Shenk 2003), it stands to reason that these malleable characters may not prove unfailing at distinguishing intraspecific taxa. Thus, the morphologically similar subspecies of *Z. hudsonius* present significant challenges to conservation biologists because of taxonomic uncertainty (Preble 1899, Krutzsch 1954, Ramey et al. 2005). Taxonomy based solely on morphological characters may not be consistent with phylogenetic relationships, as the rate of evolutionary change can vary among lineages and similar environmental influences may cause convergence (Grant 1987). This can complicate resource management efforts and ultimately jeopardize the ecological and evolutionary potential of a lineage (Moritz and Hillis 1996).

An integrative conservation approach that identifies and sustains ecological processes and evolutionary lineages is needed to protect and manage the biodiversity present in the southwestern portion of *Z. hudsonius*' range. Inherent in such an approach is the identification and characterization of associated migration, colonization, and extinction processes among

84 populations of these putative subspecies (Avise 2004). Molecular markers, with a clear heritable
85 genetic basis and the number of characters limited only by genome size (Moritz and Hillis 1996),
86 provide insight into these processes and can be used to reveal genetic discontinuities and
87 distinctiveness among or between taxa with subtle or undetectable morphological differentiation
88 (Clark et al. 2000). Patterns of gene exchange, the extent of genealogical relationships, and
89 accurate reflections of true evolutionary relationships (i.e., phylogeny) can be revealed through
90 use of the appropriate type and number of molecular genetic markers (Moritz and Hillis 1996,
91 Avise 2004).

92 Recent morphological and genetic comparisons among *Zapus hudsonius preblei* and
93 neighboring con-specifics have questioned the taxonomic status of several subspecies (Ramey et
94 al. 2005) (REA). The multidisciplinary study utilized a hypothesis testing approach to determine
95 uniqueness of subspecies, including analyses of cranial morphometrics and contemporary genetic
96 techniques. REA reported that multivariate morphometric analyses on selected cranial
97 measurements failed to support the original description of *Z. h. preblei* as a distinct subspecies
98 (but see Vignieri et al. 2006). Genetic components of the study compared haplotypes within the
99 mitochondrial (mt) DNA control region and multilocus genotypes at nuclear microsatellite DNA
100 loci. REA concluded that recent gene exchange and low levels of genetic structure among
101 subspecies supported synonymization of *Z. h. preblei*, *Z. h. campestris* (Bear Lodge meadow
102 jumping mouse), and *Z. h. intermedius* (meadow jumping mouse). REA's critical test of
103 uniqueness for *Z. h. preblei* and related taxa was that greater variance be demonstrated between
104 subspecies than within, a test criterion that, to our knowledge, has not been fully evaluated for
105 taxonomical, biological, or statistical relevance with molecular data. REA constitutes the lone
106 published molecular population genetic analysis of *Z. hudsonius*, with important implications for
107 the evolution, ecology, and conservation status of *Z. h. preblei*. The proposed synonymy of
108 these subspecies has prompted a subsequent reevaluation of the status of *Z. h. preblei* under the
109 ESA (US Fish and Wildlife Service 2005).

110 Studies of phylogeographic relationships among intraspecific taxa often exact a more
111 rigorous study design than that required for interspecific comparisons (Avise 2004). Further,
112 phylogeographic studies can be initially misled by dependence on tissues yielding insufficient
113 quality (Kirchman et al. 2001) and consistency of DNA (Steinberg 1999), inadequate portrayals
114 of genealogical relationships (Brower et al. 1996) through use of inappropriate methodology

(Posada and Crandall 2001), or insufficient resolution from too few molecular or morphological characters (Smouse and Chevillon 1998). REA utilized dried museum skins from selected collections, assessed the differentiation among selected subspecies of *Zapus hudsonius* with a hierarchical pairwise haplotypic distance approach, surveyed sequence variation for a 346 base pair (bp) fragment of the mtDNA control region, and represented nuclear DNA variation with five microsatellite DNA loci. In light of the unproven distinctiveness criteria applied to a relatively small fragment of mtDNA and minimal number of microsatellite loci, the methods applied and conclusions drawn by REA warrant independent verification.

We report additional results concerning the phylogeographic structure and evolutionary distinctiveness of *Zapus hudsonius* subspecies, placing particular emphasis on the relatedness of *Z. h. preblei* to each neighboring subspecies by comparing a larger collection of samples over a greater representation of both the mitochondrial and nuclear genomes than was previously investigated. Analyses consisted of evaluating sequence variation at approximately 374 bp of the mtDNA control region and 1000 bp of the mtDNA cytochrome *b* region, combined with the fragment analysis of a four-fold greater number of nuclear microsatellite DNA loci than surveyed by REA. We tested the null hypothesis that collections of *Z. h. preblei*, *Z. h. campestris*, *Z. h. intermedius*, *Z. h. pallidus*, and *Z. h. luteus* comprise a single homogeneous unit (i.e., fail to exhibit genetic discreteness) as reflected in the spatial distribution of mtDNA haplotypes and microsatellite DNA allele frequencies. Due to the taxonomic revision proposed by REA, importance was placed on comparisons among *Z. h. preblei, Z. h. campestris*, and *Z. h. intermedius*. In the present study we accepted as evidence of subspecific distinctiveness the conditions previously defined as significant phylogeographic separation of mtDNA alleles between subspecies (or populations), combined with congruent phylogeographic structure for nuclear loci (Avise and Ball 1990; Ball and Avise 1992; Moritz 1994a).

METHODS

Minimally invasive tissue samples (e.g., ear punches and blood) of *Zapus hudsonius campestris*, *Z. h. intermedius*, *Z. h. pallidus*, and *Z. princeps* were obtained from individuals trapped in summer 2005 (Table 1). Archived tissue samples were obtained from Z. *h. prebl*ei (ear punch), *Z. h. intermedius* (frozen liver) and *Z. h. luteus* (frozen liver) (Appendix A). In order to validate the haplotype data reported by REA as being shared among disjunct and/or highly differentiated

146 subspecies, tissue from 15 specimens was obtained from the University of Kansas Natural
147 History Museum, Lawrence, KS (KUNHM) (11 *Z. h. campestris*, two *Z. h. interemedius*, and
148 two *Z. h. pallidus*; Appendix B). These specimens represent seven of the 10 haplotypes reported
149 as being shared among subspecies by REA (haplotypes C/P1, C/P2, C/P3, C/P4, C9/INT-VII,
150 L/PAL/C1, and L/PAL/C2). All tissues were forwarded directly from the collector or museum to
151 the U.S.Geological Survey-Leetown Science Center, Kearneysville, WV for analysis. DNA was
152 obtained from ear punches (in 95% ethanol), frozen (-80° C) liver, or blood tissue (on FTA
153 cards; Whatman Inc., Clifton, NJ, USA) and extracted using the PUREGENE DNA extraction
154 kit (Gentra Systems, Inc., Minneapolis, MN) and resuspended in TE (10 mM Tris-HCl, pH 8.0, 1
155 mM EDTA). DNA was obtained from dried museum skin sections using the DNeasy^{R} Tissue
156 Extraction Kit (Qiagen, Valencia, CA, USA) with the manufacturer's protocol modified as
157 described in Iudica et al. (2001).
158
159 Microsatellite DNA Amplification and Fragment Analysis – Twenty-one microsatellite loci
160 developed by three different laboratories were screened in all individuals sampled (Appendix C).
161 The three sets of markers were: 1) Z.7, Z.20, Z.26, Z.48, and Z.52 (REA); 2) *Ztri*2, *Ztri*17,
162 *Ztri*19, and *Ztri*24 (isolated from *Z. trinotatus*; Vignieri 2003); and 3) *Zhu*C3, *Zhu*C6, *Zhu*C12,
163 *Zhu*C104, *Zhu*C119, *Zhu*C120, *Zhu*C129, *Zhu*C130, *Zhu*D107, *Zhu*D108, *Zhu*D109, and
164 *Zhu*D122 (King et al. 2006). Multiple requests by the U.S. Geological Survey-Leetown Science
165 Center laboratory for tissue or DNA samples to allow standardization with REA microsatellite
166 DNA allele scoring were not fulfilled. Microsatellite DNA amplification was conducted under
167 laminar flow and consisted of 100-200 ng of genomic DNA, 1X PCR buffer (10 mM Tris-HCl,
168 pH 8.3, 50 mM KCl), 2 mM $MgCl_2$, 0.25 mM dNTPs, 0.5 µM forward (fluorescently labeled)
169 and reverse primer, and 0.1 U *Taq* DNA polymerase (Promega, Madison, WI, USA) in 10 µl.
170 Amplifications were carried out on either a PTC-200 or PTC-225 thermal cycler (MJ Research,
171 Watertown, MA, USA) using the following: initial denaturing at 94°C for 2 min, 35 cycles of
172 94°C for 40 sec, 56°C for 40 sec, 72 °C for 1 min, and a final extension at 72 °C for 10 min.
173 Fragment analysis was performed on an Applied Biosystems (Foster City, CA, USA) ABI 3100
174 Genetic Analyzer, as described in King et al. (2001). Genescan™ 2.1 Analysis software and
175 Genotyper™ 3.6 Fragment Analysis software (Applied Biosystems) were used to score, bin, and
176 output allelic (and genotypic) data.

177

178 Mitochondrial DNA Amplification and Sequencing - Two regions of the mtDNA genome were
179 amplified and sequenced. A region of the non-coding control region (CR) was amplified by
180 double-stranded PCR using primers L15926 (5'-TCA AAG CTT ACA CCA GTC TTG TAA
181 ACC - 3') and H16498 (5'- CCT GAA CTA GGA ACC AGA TG -3') (Kocher et al. 1989;
182 Shields and Kocher 1991) for all *Z. hudsonius*, except the KUNHM specimens used to validate
183 the REA data. The CR sequence of interest could not be amplified for these specimens using the
184 primers L15926 and ZAP5Plr as described in REA. In order to amplify the CR fragment of
185 interest several internal primers were developed: DLIF1 (5'- TTT ACC ATT ATC CAT TCA
186 TGC TT -3'), DLIF2 (5'- CAG CAC CCA AAG CTG ATA TT -3'), DLIR1 (5'- TTA AGC
187 CTG ACT GAA TGT GG-3'). Ultimately a pairing of the primers DLIF1 and H16498 were able
188 to amplify a portion of the CR approximately 366 bp in length. *Z. princeps* mtDNA (*N*=7) was
189 amplified using primers L15398 (5'- ATC AGC ACC CAA AGC TGA TAT TC - 3') (REA) and
190 H16498. Polymerase chain reactions consisted of an initial denaturation at 94 ^{o}C for 3 min,
191 followed by 35 cycles of 1 min at 94 ^{o}C (denaturation), 1 min at 48 ^{o}C (annealing), and 2 min at
192 72 ^{o}C (extension), concluding with a final extension period of 5 min at 72 ^{o}C. The KUNHM
193 specimens were amplified using an annealing temperature of 55 ^{o}C . The mitochondrial
194 cytochrome *b* (cyt *b*) gene was amplified by double-stranded PCR using two primers designed
195 for this study: L14398A (5'- CCA ATG ACA TGA AAA ATC ATC G - 3') and H15634A (5'-
196 TGG TTT ACA AGA CCA GAG TAA - 3'). Polymerase chain reactions consisted of an initial
197 denaturation at 94 ^{o}C for 3 min, followed by 35 cycles of 1 min at 94 ^{o}C (denaturation), 1 min at
198 55 ^{o}C (annealing), and 2 min at 72 ^{o}C (extension), concluding with a final extension period of 10
199 min at 72 ^{o}C. Polymerase chain reactions consisted of 25 µl total volume, containing 2.5 µl of
200 $MgCl_2$-free buffer, 2.5 µl of $MgCl_2$ solution, 0.5 µl of dNTPs (2.5 mM each), 1.25 µl of each
201 primer (10 µM), one unit *Taq* polymerase, three µl of template (ca. 50-100 ng double-stranded
202 DNA), and 13.75 µl of sterile water. Negative controls, which did not include template DNA,
203 were set up alongside PCR reactions as checks for contamination of PCR reagents. Polymerase
204 chain reaction products were purified with exonuclease I and shrimp alkaline phosphatase. The
205 CR and cyt *b* PCR products were sequenced using ABI BigDye v3.1 terminator cycle sequencing
206 chemistry, with sequences read by an ABI 3100 Genetic Analyzer (Applied Biosystems). The
207 CR fragment was sequenced using primers L15926 or L15398 and H16498 for most specimens.

208 The KUNHM specimens were sequenced using the internal primers DLIF1, DLIF2, DLIR1 and
209 H16498. The cyt *b* gene was sequenced using primers L14398A and H15634A, and when
210 necessary, internal primers for the cyt *b* gene. Internal primers designed specifically for this
211 study were: CytbIF1 (5'- CCA TTC CAT ATA TTG GCT CA-3'), CytbIF2 (5'- TCC CAT TCC
212 ATC CTT ACT ACA - 3'), CytbIR1 (5' - CCA ATA TAT GGA ATG GCT GA - 3') and
213 CytbIR2 (5' – GGG GTA TTT AAT GGG TTT GC – 3'). Cycle sequencing reactions consisted
214 of 30 cycles of 20 sec at 96 ^{o}C (denaturation), 20 sec at 50 ^{o}C (annealing), and 4 min at 60 ^{o}C
215 (extension). Forward and reverse sequences for each individual were assembled using
216 Sequencher 4.5 (Gene Codes Corporation, Ann Arbor, MI), were aligned to a reference sequence
217 from GenBank (*Z. h. preblei*, AY598282 for CR and *Z. trinotatus*, AF119262 for cyt *b*), and
218 were double-checked by two researchers. Clustal X (Thompson et al. 1994) was used to obtain
219 multiple sequence alignments for CR and cyt *b*. Alignments of both datasets were performed
220 with default settings and were straightforward as only three sites had indels in the CR dataset and
221 no indels were encountered in cyt *b*, as would be expected for this protein-coding gene.
222 Sequences generated in this study are available in GenBank (accession numbers: CR DQ664546-
223 DQ664900; cyt *b* DQ664901- DQ665221).
224
225 *Data analysis*
226 Microsatellite DNA – Observed genotype frequencies were tested for consistency with Hardy-
227 Weinberg and linkage equilibrium expectations using randomization tests implemented by
228 GENEPOP 3.1 (Raymond and Rousset 1995). The Hardy-Weinberg test used the Markov chain
229 randomization test of Guo and Thompson (1992) to estimate exact two-tailed *p* values for each
230 locus in each sample. Tests for linkage equilibrium were conducted using the randomization
231 method of Raymond and Rousset (1995) for all pairs of loci. Bonferroni adjustments (Rice
232 1989) determined statistical significance for these and all other simultaneous tests. Average
233 observed (H_o) and expected (H_e) heterozygosities were calculated by BIOSYS-1 (Swofford and
234 Selander 1981). Amount of allelic diversity (expressed as allelic richness; El Mousadik and Petit
235 1996), estimates of population subdivision (F_{ST}; Weir and Cockerham 1984), and inbreeding
236 coefficients (F_{IS}) were determined using FSTAT (Goudet 1995). All pairwise F_{ST} estimates
237 between subspecies (N = 5) and other intraspecific groupings were tested for significance (i.e.,
238 difference from zero) by adjusted permutations using FSTAT. Pairwise R_{ST} values among

239 subspecies were calculated using GENEPOP 3.1 and are provided for comparison with F_{ST}
240 values. F_{ST} assumes allelic diversity results from migration and gene drift, while R_{ST} also
241 measures mutational differences between alleles. The statistical significance of genetic
242 differences between each pair of collections, clusters, and subspecies was tested using the genic
243 differentiation randomization test in GENEPOP. Results were combined over loci using Fisher's
244 method (Sokal and Rohlf 1994).

245 Several techniques were used to describe genetic relationships between subspecies and
246 collections. We used the model-based clustering method of the program *STRUCTURE*
247 (Pritchard et al. 2000) to infer population structure among collections and probabilistically
248 assign all individuals to detected clusters (k). Due to complex migration patterns assumed to
249 exist among disjunct subspecies, a sequential method of inferring k was used by first identifying
250 the "uppermost" hierarchical level of population structure followed by subsequent analysis of
251 each cluster to identify within-cluster structure (Evanno et al. 2005). In the initial phase, $k = 1$
252 to $k = 15$ clusters were considered for the 14 collections using a burn-in of 15,000 followed by
253 100,000 iterations, and 100 independent runs for each k. The optimum number of clusters in the
254 initial phase was identified using Δk as described by Evanno et al. (2005). Subsequent analysis
255 of each cluster tested $k = 1$ to $k = C + 3$ (the number of collections (C) included in the subset plus
256 three), with a burn-in of 10,000 followed by 10,000 iterations, and 20 runs for each k. In the
257 within-cluster analyses, k was also determined using the Evanno et al. (2005) method .
258 Individual assignment success to the cluster or subcluster of origin was recorded as the highest
259 likelihood of assignment (q) and the percentage of individuals in a cluster with $q \geq 0.90$
260 (Pritchard et al. 2000).

261 Genetic distances between each pair of collections and subspecies were summarized with
262 D_a (Nei et al. 1983), calculated by DISPAN (Institute of Molecular Evolutionary Genetics, The
263 Pennsylvania State University, University Park, PA, USA). Simulation has shown that the
264 geometric-based D_a exhibits a stronger linear relationship over shorter divergence time, and
265 therefore, estimates better tree topology than other commonly-used genetic distances when
266 analyzing microsatellite DNA variation (Takezaki and Nei 1996). An unrooted phylogenetic
267 tree was fitted using the D_a distance matrix and neighbor-joining (NJ) algorithm. TreeView
268 (Page 1996) was used to visualize the tree. The strength of support for each node in the tree was
269 tested by bootstrapping over loci using NJBPOP (J.-M. Cornuet, INRA, Montpellier, France).

270 Analysis of molecular variance (AMOVA) was used to partition genetic variation among
271 clusters and subspecies (Excoffier et al. 1992). Arlequin 2.0 (Schneider et al. 2000) was used to
272 quantify and test statistical significance of observed differentiation between subspecies, between
273 clusters of collections, and within subspecies and clusters.
274
275 Mitochondrial DNA – Unique haplotypes for each data partition (CR, cyt *b*) were determined
276 using the program Collapse 1.2 (http://darwin.uvigo.es/software/collapse.html). Sequences of
277 each representative CR haplotype from this study were aligned with the REA data from
278 GenBank and the sequence data from the 15 KUNHM specimens to verify their sequences.
279 Clustal X (Thompson et al. 1994) was used to align the CR dataset with default settings as only
280 three sites had indels. Haplotypes were then compared using the program Collapse 1.2
281 (http://darwin.uvigo.es/software/collapse.html) to identify matching haplotypes. Sequences from
282 the 15 KUNHM specimens were not used in any subsequent analyses. To examine whether
283 differences in phylogenetic signal existed between the two mitochondrial data partitions,
284 incongruence length differences (ILD, Farris et al. 1994) were calculated in PAUP* by the
285 partition-homogeneity test. Settings for ILD tests were as in parsimony analysis, with
286 uninformative sites excluded, and 1,000 replicates per run. The combined data set (cyt b and
287 control region) was examined using maximum parsimony (MP) using PAUP* 4.0b10 (Swofford
288 2002). Parsimony analyses consisted of heuristic searches on unweighted, parsimony-
289 informative (PI) characters with starting trees obtained via stepwise addition, 100 random
290 additions of sequences per run, and tree bisection-reconnection (TBR) branch swapping on best
291 trees. Nodal support on parsimony trees was assessed using 1,000 bootstrap replicates
292 (Felsenstein 1985) with full heuristic searches and Bremer support (Bremer 1988, 1994) using
293 TreeRot v. 2b (Sorenson 1999). *Z. princeps* was used as the outgroup in all analyses.
294 Partitioned bayesian phylogenetic analysis of the combined data set was conducted using
295 the program MrBayes 3.0 (Huelsenbeck and Ronquist 2001). The data was partitioned into the
296 control region and first, second, and third codon positions of the cyt *b* gene. The appropriate
297 model of evolution for each partition was chosen with the Akaike information criterion
298 implemented by the program MrModeltest. A neighbor-joining tree for the combined dataset
299 was generated using the JC model and used for the analysis of each partition with MrModeltest.
300 Bayesian phylogenetic analyses were run for 1.0×10^7 generations with random starting trees,

default priors, four Markov chains and sampling every 1000 generations. Stationarity of the MCMC analyses was determined by plotting –lnL values and individual parameter estimates against generation times. Trees from the burn-in were discarded and the remaining trees used to calculate posterior probabilities for clades from their frequencies. Two separate analyses were performed and the resulting topologies and posterior probabilities compared.

We investigated intraspecific gene genealogies for the control region and cyt *b* data separately using the haplotype networking approach in the TCS computer program (Clement et al. 2000). This analysis implemented the statistical parsimony approach of Templeton et al. (1992) and Crandall et al. (1994).

Nucleotide diversity of the CR, cyt *b* and combined datasets and exact tests for subspecies and cluster (as defined by *STRUCTURE*) haplotype differentiation were performed pairwise with 1000 replications in Arlequin 2.0. Total haplotypic variation was partitioned into "among versus within" *Z. hudsonius* subspecies using AMOVA. Subspecies- and cluster-level differentiation was assessed using Arlequin 2.0. All AMOVA analyses were conducted in two stages to assess divergence from different evolutionary processes. The first analysis incorporated sequence divergence between haplotypes as well as their frequencies (Φ_{ST}) by calculating either Kimura 2-parameter (CR) or Tamura-Nei (Tamura and Nei 1993) (cyt *b*) estimates. The second analysis, which treated all haplotypes as equally differentiated (i.e., distance = 1.0), assessed the variance distribution based on haplotype frequencies alone (F_{ST}). Differences between haplotype frequencies are assumed to be due to genetic drift. The significance of pairwise Φ_{ST} and F_{ST} values were calculated by permuting haplotypes between populations 1,000 times.

RESULTS

Microsatellite DNA- Genotypes at 21 microsatellite DNA loci were determined for 348 *Z. hudsonius* from 14 locations representing five neighboring subspecies (see Table 1 for listing and abbreviations; Figure 1). A high level of genetic diversity was detected among the 14 collections; 280 alleles were observed across 21 loci ranging from 7 at *Zhu*C120 and *Zhu*C130 to 30 at *Z*.7 (Appendix C). The mean number of alleles per locus was 13.3. Allelic richness estimates for subspecies ranged from 4.6 (*Z. h. luteus*) to 9.4 (*Z. h. intermedius* and *Z. h. pallidus*) (Table 1). Average observed heterozygosity was on average 11.5% lower in *Z. h.*

332 *preblei* than in other subspecies. Observed mean heterozygosity ranged from 53.9% in *Z. h.*
333 *preblei* to 75.2% in *Z. h. pallidus*. The number of unique alleles observed ranged from 5 (*Z. h.*
334 *preblei*) to 33 (*Z. h. pallidus*). Estimates of individual pair-wise genetic distances, using the
335 proportion of shared alleles, indicated that levels of genetic diversity observed among the 21
336 microsatellite loci were sufficient to produce unique multilocus genotypes (i.e., genetic distances
337 > zero) for all animals surveyed (distances not presented).
338 When organized at the collection level, randomization tests showed that genotypes for
339 the majority of locus-by-collection comparisons were consistent with Hardy-Weinberg
340 expectations. A total of 294 locus-by-collection comparisons were made of which 10 (3.4%)
341 were statistically significant after Bonferroni adjustment for multiple tests at overall $\alpha = 0.05$ (P
342 < 0.0036). These comparisons consisted of six collections at six loci: CCWY at *Z*.26; BRCSD
343 and MCMN at *Zhu*C119; MCMN, BCSD, and KBCNE at *Zhu*C130; BCSD and KBCNE at
344 *Ztri*19; MCMN at *Zhu*C104; and KBCNE at *Z*.7 (all but the last occurrence due to heterozygote
345 deficiencies). This level of deviation is likely due to one or a combination of factors including
346 sub-structuring of the sample (i.e., Wahlund effect), inbreeding, or the presence of null alleles.
347 Interestingly, none of the deviations was observed in collections of *Z. h. preblei*, the subspecies
348 targeted for development of 17 of 21 markers surveyed. This observation, combined with the
349 heterozygote deficiencies observed, suggests that differentiation between *Z. h. preblei* and the
350 neighboring subspecies may have increased the likelihood that null alleles would be present in
351 the non-target subspecies. The number of deviations from Hardy-Weinberg expectations
352 increased ($N = 26$; 24.8%) when collections were pooled as putative subspecies. The majority
353 ($N = 15$) of the significant deviations (overall $\alpha = 0.05$; $P < 0.01$) were observed among the
354 collections pooled as *Z. h. preblei*. All deviations were the result of heterozygote deficiencies.
355 This result suggests that the allele frequencies of the populations pooled to form this subspecies
356 (*Z. h. preblei*-North and *Z. h. preblei*-South) have achieved different equilibria and that sub-
357 structuring (i.e., Wahlund effect) has been detected.
358 Minimal linkage disequilibrium was observed as five of 2700 (0.2%) comparisons of
359 each locus pair across all collections was found to be significant after correction for multiple
360 tests (overall $\alpha = 0.05$, $P < 0.0002$). Linkage disequilibrium was observed between *Z*.20 and
361 *Zhu*D122, *Z*.20 and *Ztri24*, *Zhu*D107 and *Zhu*D122, *Zhu*C12 and *Zhu*C120, and *Z*.48 and *Ztri*2
362 with the five occurrences distributed among four collections. These findings were likely the

363 result of sampling error, year-class mixing, population mixing, or a combination of the three
364 rather than physical linkage among loci.
365 Allele frequency heterogeneity was observed throughout the study area. Among 1866
366 single-locus pair-wise tests of allele frequency heterogeneity, 1382 (74.1%) indicated departures
367 from homogeneity after correction for multiple tests ($\alpha = 0.05$, $P < 0.0006$). When testing allele
368 frequency heterogeneity across 21-locus genotypes, highly significant heterogeneity was
369 observed in all 91 pairs of collections ($\alpha = 0.05$, $P < 0.001$).
370 Results from the *STRUCTURE* analysis identified that $k = 3$ was the appropriate number
371 of clusters to be recognized at the uppermost hierarchical level among the 14 collections of *Z.*
372 *hudsonius* genotyped at 21 microsatellite DNA loci. Ten of the twenty independent runs at $k = 3$
373 resulted in the assignment of the 14 collection to the following clusters: cluster [A] – SOWY,
374 LCCO1, LCCO2, DCCO1, DCCO2, ECCO1, ECCO2 (*Z. h. preblei*); cluster [B] - CCWY,
375 CCSD (*Z. h. campestris*) and BRCSD and MCMN (*Z. h. intermedius*); and cluster [C] - BCSD
376 and KBCNE (*Z. h. pallidus*) and SCNM (*Z. h. luteus*) (Figure 2). The other ten independent runs
377 at $k = 3$ resulted in the same clustering with the exception that 27 of 28 *Z. h. preblei* individuals
378 from collection SOWY were assigned to cluster [B], the *Z. h. intermedius*/*Z. h. campestris*
379 cluster. An additional 100 independent *STRUCTURE* runs were conducted at $k = 3$, with 57 of
380 the runs assigning all of the SOWY animals to the *Z. h. preblei* cluster [A]. All of the 57 runs
381 that assigned the SOWY mice to the *Z. h. preblei* cluster [A] did so with higher average q values
382 (0.93) than the runs that assigned the SOWY animals to the *Z. h. intermedius*/*Z .h. campestris*
383 cluster [B] (average $q = 0.85$). Given the higher number of runs assigning the SOWY collection
384 animals to the *Z. h. preblei* and the higher assignment values of these runs, this clustering was
385 considered the most appropriate. The three clusters detected as the uppermost hierarchical level
386 of population structure suggest the collections constituting *Z. h. preblei* form a distinct grouping
387 from the *Z. h. campestris* and *Z. h. intermedius* collections, which are also distinct from the *Z. h.*
388 *pallidus* and *Z. h. luteus* collections. This structure was confirmed by 100% correct assignment
389 of each mouse to the cluster-of-origin based on q values (Table 2). The average value of qMAX
390 for the clusters ranged from 0.96 ([B]) to 0.99 ([C]). Subsequent analysis of the three clusters
391 suggested a strong pattern of genetic differentiation throughout the study area and the presence
392 of $k = 6$ definable subclusters (Figure 2). This analysis identified phylogeographic discontinuities
393 present within *Z. h. preblei* that corresponded to the northern and southern collections

394 (subclusters 1 and 2, respectively), as well as among clusters allied with *Z. h. campestris*
395 (subcluster 3), *Z. h. intermedius* (subcluster 4), *Z. h. pallidus* (subcluster 5), and *Z. h. luteus*
396 (subcluster 6). A subsequent analysis of the clusters that assigned SOWY animals to cluster [B]
397 resulted in $k = 7$, where the clustering was the same, with the exception being that the SOWY
398 animals were assigned to a unique subcluster. Given these results, we determined that the
399 optimal number of subclusters for these data is six. The strength of the differentiation among
400 these six subclusters was also evident upon inspection of individual assignment results and the
401 average value of *q*MAX (Table 2). When each individual was assigned to subcluster based on
402 the largest value of *q*, average assignment success to subcluster of origin was 99.4% (346 of 348
403 individuals correctly assigned). Two *Z. h. preblei* individuals from *Z. h. preblei*-North
404 (subcluster 1) were incorrectly assigned to *Z. h. preblei*-South (subcluster 2). Average values of
405 *q*MAX for the six subclusters ranged from 0.94 (*Z. h. preblei - North*) to 0.99 (*Z. h. intermedius*,
406 subcluster 4, *Z. h. pallidus*, subcluster 5; *Z. h. luteus*, subcluster 6). When using $q \geq 90$ as an
407 assignment threshold, the percentage of correct assignment to subcluster ranged from 89.7% (*Z.*
408 *h. preblei* - North) to 100.0% (*Z. h. intermedius*, *Z. h. pallidus,* and *Z. h. luteus*). When
409 compared at the subspecies level ($q \geq 90$ criterion), assignment success ranged from 95.1% (*Z. h.*
410 *campestris*) to 100% (*Z. h. preblei*, *Z. h. intermedius*, *Z. h. pallidus,* and *Z. h. luteus*).
411 Pair-wise genetic distance values (D_a, Nei et al. 1983) were calculated between all
412 collections to investigate evolutionary relationships in allele frequencies. The greatest genetic
413 distances occurred between the *Z. h. luteus* collection and all other collections; the lowest
414 genetic distances were observed between collections from the same subspecies or cluster (as
415 defined by *STRUCTURE*). The underlying genetic structure of the D_a matrix is illustrated with
416 an unrooted neighbor-joining (NJ) tree (Figure 3). The patterns observed illustrate high levels of
417 differentiation among and within the five subspecies and reflects the patterns of genetic variation
418 identified by *STRUCTURE.* . The distinctiveness of the *Z. h. preblei* collections was confirmed
419 relative to other subspecies by 98% bootstrap support for separation of the subspecies from all
420 other collections. The SOWY collection from the northernmost portion of the subspecies' range
421 was intermediate between the other northern *Z. h. preblei* collections and the disjunct (by
422 approximately 150 km) *Z. h. campestris* collections. In addition, this figure depicts the
423 differentiation observed:1) among each of the five subspecies; 2) within each subspecies
424 (excluding *Z. h. luteus*); and 3) between two major groupings (*Z. h. prebl*ei, subclusters 1 and 2;

425 *Z. h. campestris*, subcluster 3; *Z. h. intermedius*, subcluster 4) and (*Z. h. pallidus*, subcluster 5; *Z.*
426 *h. luteus*, subcluser 6) (100% bootstrap support). The closest genetic relationship among
427 subspecies exists between *Z. h. campestris* and *Z. h. intermedius*.
428 A high level of genetic differentiation was also observed in F_{ST} comparisons at the
429 subcluster (data not presented) and subspecies (Table 3) scale. All 24 tests of significance in
430 pair-wise F_{ST} values (clusters and subspecies) were statistically greater than zero ($p < 0.001$)
431 indicating numerous significant genetic discontinuities throughout this portion of *Z. hudsonius*'
432 range. The lowest pair-wise F_{ST} value was observed between the two *Z. h. preblei* subclusters
433 (North and South, 0.10). F_{ST} estimates were highest between the *Z. h. luteus* subcluster and the
434 *Z. h. prebei, Z. h. campestris,* and *Z. h. intermedius* subclusters (all values ≥ 0.27). Moderately
435 high F_{ST} estimates were observed between the *Z. h. preblei* and *Z. h. campestris* subclusters
436 (average 0.12), and between *Z. h. pallidus* and the *Z. h. intermedius* and *Z. h. campestris*
437 subclusters (average 0.15). F_{ST} estimates at the subspecies level (all collections and subclusters
438 pooled) mirrored those observed among subclusters, with the highest estimates observed
439 between *Z. h. preblei* and *Z. h. luteus*, *Z. h. campestris* and *Z. h. luteus*, and *Z. h. intermedius* and
440 *Z. h. luteus* (Table 3). The lowest F_{ST} estimates were observed among *Z. h. preblei*, *Z. h.*
441 *campestris*, and *Z. h. intermedius* averaging 0.13. R_{ST} values were on average 2.1 times larger
442 than corresponding F_{ST} values and ranged from 0.16 (*Z. h. preblei* - Z. *h. campestris*) to 0.78 (*Z.*
443 *h. preblei – Z. h. luteus*) (Table 3). The ratio of R_{ST} to F_{ST} values ranged between 1.0 (*Z. h.*
444 *luteus* - Z. *h. pallidus*) and 2.9 (*Z. h. preblei – Z. h. pallidus*). The observed R_{ST} values suggest
445 that the differentiation observed among most of the subspecies is considerable as mutational
446 processes have acted to increase differentiation over that observed through random genetic drift.
447 Quantitative estimates of hierarchical gene diversity (AMOVA) among subspecies and
448 subclusters also identified statistically significant genetic structuring. A comparison between the
449 five subspecies (all collections pooled) determined that 18.4% ($P < 0.001$) of the genetic
450 variation occurred between subspecies and 81.6% ($P < 0.001$) was due to differentiation within
451 subspecies. Further partitioning resulted in 14.8% ($P < 0.001$) of the variation being distributed
452 between subspecies, 8.6% ($P < 0.001$) observed among collections within subspecies, and 76.6%
453 ($P < 0.001$) of the variance detected within collections. A comparison between the six
454 subclusters (all collections pooled) identified by *STRUCTURE* yielded results identical to those
455 observed for subspecies partitions - 18.4% ($P < 0.001$) of the variance due to differences

between subclusters and 81.6% ($P < 0.001$) was attributable to differentiation within subclusters. Partitioning variation among collections within subclusters resulted in slightly more variation being distributed among subclusters (15.4%, $P < 0.001$), less variation (6.7%; $P < 0.001$) among collections within subclusters, and more variation (77.9%; $P < 0.001$) attributed to variation in collections within subclusters than was observed for subspecies. When subclusters were grouped into subspecies, 11.9% ($P < 0.024$) of the variation was distributed between subspecies, 7.8% ($P < 0.001$) was between subclusters within subspecies, and 80.3% ($P < 0.001$) of the variance was observed within clusters.

Mitochondrial DNA - Sequence data from *Z. hudsonius* individuals from 13 collections representing five neighboring subspecies were analyzed for sequence variation at two mitochondrial genes, the 5'-end of the CR (374 bp, 332 individuals) and cyt *b* gene (1006 bp, 320 individuals). In the CR dataset 25 haplotypes were recovered with 28 (7.5%) sites variable, 27 (7.2%) sites parsimony-informative, and three indels parsimony-informative when coded as a fifth base (Appendix D). For the cyt *b* region 56 haplotypes were recovered with 116 (11.5%) sites variable, 84 (8.3%) sites parsimony-informative and no indels present (Appendix E).

Control region sequences were obtained for 15 of the KUNHM specimens utilized by REA. When these sequences were compared to the REA data from GenBank and representative CR haplotypes from this study, 13 of the sequences were different than those reported by REA (Table B.1, Appendix B). All seven of the *Z. h. campestris* that were reported as having *Z. h. preblei* haplotypes by REA were found to have common *Z. h. campestris* haplotypes. Given these results, the validity of all the REA data from the mitochondrial DNA genome is called into question, and therefore was not combined with the data from this study for any analyses.

Adjusted (net) mtDNA sequence divergence estimates (Kimura 2-parameter, CR and Tamura-Nei, cyt *b*) between *Z. hudsonius* subspecies were relatively low, averaging 3.37% (CR) and 4.43% cyt *b* (not shown). Net distances between *Z. hudsonius* subspecies ranged between 0.29% (CR) and 0.18% (cyt *b*) for the most closely related subspecies, *Z. h. campestris* and *Z. h. intermedius*, to 5.63% (*Z. h. campestris* - *Z. h. luteus,* CR) and 7.11% (*Z. h. campestris* - *Z. h. pallidus,* cyt *b*). *Z. h. preblei* was least divergent from *Z. h. intermedius* (0.57%, CR and 0.68%, cyt *b*), and most divergent from *Z. h. luteus* (5.08%, CR) or *Z. h. pallidus* (6.76%, cyt *b*). Divergence estimated from CR between *Z. h. preblei* and *Z. h. campestris* (1.03%) was nearly

487 twice the distance observed between *Z. h. preblei* and *Z. h. intermedius*. Estimates of divergence
488 between *Z. hudsonius* subspecies and *Z. princeps,* used as an outgroup in phylogenetic analyses,
489 were higher (10.33%, CR and 19.87%, cyt *b*).

490 Nucleotide diversity within subspecies was low across the study, and ranged from 0.0000
491 in *Z. h. luteu*s (both gene regions) to 0.0060 and 0.0030 in *Z. h. pallidus* (Table 4, CR and cyt *b*,
492 respectively)**.** Twenty-five (CR) and 56 (cyt *b*) haplotypes were observed study-wide for *Z.*
493 *hudsonius*, yet none were shared among the five subspecies (Appendices B and C). For *Z. h.*
494 *preblei*, 1 of 4 CR haplotypes and none of the 21 cyt *b* haplotypes were shared between the
495 northern and southern clusters.

496 Statistical parsimony analyses of sequence data from each mitochondrial gene region
497 produced largely concordant haplotype networks. In the CR analysis with gaps included as a
498 fifth character state the 0.95 limit for connections was eight steps and in the cyt *b* data set the
499 0.95 limit was 13 steps. Both analyses produced multiple networks, but both included a single
500 network including all *Z. h. preblei*, *Z. h. intermedius*, and *Z. h. campestris* haplotypes. In the CR
501 analysis, a second network was formed by *Z. h. pallidus* and *Z. h. luteus* haplotypes (Figure not
502 shown), where these haplotypes formed two separate networks corresponding to subspecies
503 based on cyt *b* (Figure 4). Within the (*Z. h. preblei* - *intermedius – campestris*) networks,
504 haplotypes made up of individuals from each subspecies clustered together (subspecies outlined
505 in boxes) and were separated from other subspecies' haplotypes by one (CR) to six (cyt *b*,
506 *preblei-intermedius,* Figure 4) mutational steps. The same three *Z. h. intermedius* individuals
507 had CR (ZhiCR_E) and cyt b (ZhiCB_A, Figure 4) haplotypes that had fewer mutational steps to
508 a *Z. h. campestris* haplotype than to other *Z. h. intermedius* haplotypes, yet these haplotypes
509 were not found in any *Z. h. campestris* individuals (N = 61).

510 Analysis of molecular variance of data from each mitochondrial gene region indicated the
511 presence of strong, significant genetic differentiation among the five *Z. hudsonius* subspecies
512 surveyed. Global Φ_{ST} (0.89 for CR and 0.96 for cyt *b*) and F_{ST} (0.36 for CR and 0.29 for cyt *b*;
513 not shown) values were high and significantly different from zero ($p < 0.00001$). As no
514 haplotypes were shared among the five *Z. hudsonius* subspecies, F_{ST} values (a measure of
515 haplotype frequencies only) generated by the AMOVA analysis have limited biological
516 resolution and are discussed only for comparison to previous findings (REA). A global Φ_{ST} of
517 0.96 indicates that nearly all of the haplotype variance (96%) was distributed between

518 subspecies, a result consistent with the absence of any shared haplotypes observed between
519 subspecies. In general, comparisons among all subspecies resulted in higher Φ_{ST} values for cyt *b*
520 than for CR but with one exception, the comparison between *Z. h. intermedius* and *Z. h.*
521 *campestris,* which also were the lowest Φ_{ST} estimates (CR = 0.59, cyt *b* = 0.53). A comparison
522 of haplotype variation among the three subspecies proposed for synonymy by REA, *Z. h. preblei*,
523 *Z. h. intermedius* and *Z. h. campestris*, resulted in 73% (CR) and 83% (cyt *b*) of the variance
524 being distributed among subspecies. Moreover, AMOVA results suggest that *Z. h. preblei* is
525 distinct from other neighboring subspecies and is evolutionarily more similar to *Z. h. intermedius*
526 than to *Z. h. campestris*, as Φ_{ST} estimates were lowest with *Z. h. intermedius* (Φ_{ST} = 0.67, CR
527 and Φ_{ST} = 0.82, cyt *b*) and slightly higher when compared with *Z. h. campestris* (Φ_{ST} = 0.79, CR
528 and Φ_{ST} = 0.89, cyt *b*). This finding differs from the pattern observed during analysis of the
529 nuclear DNA, which found *Z. h. preblei* to exhibit a lower genetic distance between *Z. h.*
530 *campestris* than *Z. h. intermedius*. All estimates of subspecies from different mitochondrial
531 clades (Figure 5A) were above 90%, with both *Z. h. preblei* and *Z. h. campestris* reaching 98%
532 in comparisons with *Z. h. luteus*. Haplotype variation distributed among the clusters indicated by
533 the nuclear microsatellite loci by the program *STRUCTURE* was similar to that observed among
534 the subspecies, as 90.0% of the CR variation and 95.4% of the cyt *b* variation ($P < 0.001$) was
535 attributable to differentiation among subclusters and the remainder due to variation within. These
536 findings mirrored the high level of genetic differentiation observed upon survey of these
537 subclusters with the nuclear microsatellite DNA markers. Exact tests for differentiation based on
538 haplotype frequencies for both gene regions and clusters (N = 15 comparisons) were statistically
539 significant ($P < 0.001$).
540
541 *Combined mitochondrial DNA*
542 Results of ILD tests were not significant (P = 0.23), plus phylogenetic analyses of the individual
543 data partitions were congruent with respect to the major clades recovered (data not shown), so
544 data was combined for individuals occurring in both data sets (303 individuals), then collapsed
545 into 69 unique haplotypes (68 *Zapus hudsonius*, 1 *Z. princeps* Table 4). The strict consensus
546 tree of 8032 most parsimonious trees (Figure 5A) revealed two well-supported clades within *Z.*
547 *hudsonius*. One clade, supported by 100 percent of bootstrap pseudoreplicates and a decay index
548 of 17, was comprised of *Z. h. pallidus* and *Z. h. luteus* haplotypes, and the two *Z. h. luteus*

haplotypes formed a well supported monophyletic group. The other major clade recovered was comprised of *Z. h. intermedius*, *Z. h. campestris*, and *Z. h. preblei* haplotypes (bootstrap = 100, decay index 36, Figure 5A). Within this clade none of the *Z. hudsonius* subspecies were recovered as monophyletic, yet the *Z .h. campestris/intermedius* haplotypes formed a clade paraphyletic to *Z. h. preblei* haplotypes (bootstrap values of 57, decay index of 4, Figure 5A). Generally, nodes within this major clade were poorly supported by bootstrap and decay indices.

Models select using the AIC by MrModeltest for the control region and first, second, and third codon positions of the cyt *b* gene were HKY+I+G, K80+I, GTR+I, and GTR resepectivley. In all Bayesian analyses stationarity was reached by 50,000 generations, therefore the first 50 trees were discarded as burn-in and the remaining trees used to calculate topology and posterior probabilities. Topologies of the consensus trees from the two Bayesian analyses were identical and posterior probabilities for each clade varied by no more than 0.01 (Figure 5B, one of two bayesian consensus trees). The trees resulting from the Bayesian analysis were similar in topology to the MP consensus (Figure 5A), with the major difference being the relationship between *Z. h. pallidus* and *Z. h. luteus*- they are not sister taxa in the Bayesian analysis (Figure 5B). In the phylogram of the Bayesian tree (Figure 5B), it is notable that branch lengths were short within the *Z. h. campestris/intermedius/preblei* clade, with the longest branch (and the only one with substantial bootstrap support) separating *Z. h. preblei* from *Z. h. campestris/intermedius* haplotypes.

DISCUSSION

The disjunct collections represented in the present study may not qualify as distinct species because their reproductive barrier is extrinsic, but their apparently longstanding separation has resulted in the accretion of considerable genetic differentiation that is reflected in geographically congruent patterns of divergence at multiple DNA markers. The magnitude of the observed differentiation was considerable and supported by highly significant findings for nearly every statistical comparison, regardless of the genome or the taxa under consideration. Given the strong structuring of nuclear multilocus genotypes, the presence of only subspecies-specific mtDNA haplotypes, and that the genetic discontinuities correspond with previously proposed taxonomic classifications based on geographic separation and morphological variation among these subspecies (Krutzsch 1954, Hafner et al. 1981), we find no genetic evidence warranting

580 taxonomic revision among the *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius* as previously
581 proposed (REA), nor suggesting the need to alter *Z. h. preblei*'s standing under the ESA.
582 Moreover, the level of differentiation observed between *Z. h. preblei* inhabiting southern
583 Wyoming and Larimer County, CO (*Z. h. preblei*-North) and mice sampled from Douglas and El
584 Paso counties, CO (*Z. h. preblei*-South) indicates that this subspecies is comprised of at least two
585 genetically distinct populations worthy of individual management consideration. Spatial genetic
586 structuring apparent in *Z. hudsonius* subspecies is reminiscent of that observed in *Z. trinotatus*
587 (Pacific jumping mouse). Through a landscape genetics approach, Vignieri (2005) explored the
588 interactions between dispersal and environmental characteristics and found that both restricted
589 dispersal and habitat-directed movement were likely forces shaping fine-scale genetic patterns
590 between populations. Sufficient data have also been presented to suggest that the differentiation
591 observed between the *Z. h. pallidus* and *Z. h. luteus* clade and other *Z. hudsonius* subspecies
592 warrants further study and serious consideration given to species-level recognition for the
593 former.
594 The findings of this study are contradictory in nearly every comparison with the
595 conclusions drawn from the only other molecular genetics data available on the taxonomic status
596 of *Z. hudsonius* subspecies (REA). It is imperative that the disparities between the previous and
597 present study be discussed due to the conservation implications of synonymizing *Z. h. preblei*, *Z.*
598 *h. campestris*, and *Z. h. intermedius* as proposed by REA, and the misleading precedent that may
599 be accepted for evaluating the genetic basis for listing under the ESA unless these discrepancies
600 are understood by scientists and properly portrayed to resource managers. The largest
601 differences between the two studies dealt with the presence or absence of haplotype sharing
602 among subspecies and the philosophical perspective with which the researchers interpreted the
603 findings (see Conservation Implications). It was particularly troubling that the CR sequences of
604 13 specimens reported by REA to possess shared haplotypes between subspecies were not
605 substantiated when sequenced for this study (see Appendix B). Equally troubling was that 80 of
606 the 222 total sequences deposited to GenBank by REA cannot be directly linked to the
607 specimens listed in Appendix 2 of REA because accession numbers are not provided and the
608 isolate names listed on GenBank for the 80 sequences do not match ID numbers given in REA.
609 The inability to directly link the sequences deposited in GenBank to the specimens listed in REA
610 and the associated locality data renders the data useless. Possible reasons for the different

611 sequences reported by REA are contamination, mislabeling of samples, or other procedural
612 incongruity. Given that all the specimens re-evaluated in the current study were found by REA
613 to have haplotypes that were shared among subspecies, the disagreement in DNA sequences
614 reported for these individuals calls into question all of the results of REA based on the
615 mitochondrial DNA genome and prevents meaningful analysis of the combined data. Moreover,
616 this study and that of REA differ in sampling regime, number of molecular characters sequenced
617 and microsatellite fragments surveyed, test statistics applied to AMOVAs, methods used to
618 portray genealogical relationships, and criteria used to determine uniqueness. We will discuss the
619 nature of these discrepancies and suggest how each approach impacts the conclusions drawn
620 regarding subspecies distinctness.

621 An appropriate sampling strategy is central to the successful delineation of population
622 genetic and phylogeographic structures (Baverstock and Moritz 1996). The "population-
623 oriented" strategy used in the present study differed significantly from the "one or few
624 individuals per site across a broad geographical area" approach applied by REA. This study
625 focused on collections ranging from 14 to 33 individuals from recognized geographic
626 populations, increasing the likelihood that the majority of haplotypes within a population would
627 be surveyed. Total genetic variation within *Z. hudsonius* can be hierarchically partitioned: 1)
628 between subspecies, 2) among populations within subspecies, and 3) within populations. The
629 sampling strategy and subsequent analyses performed by REA likely underestimated the level of
630 within-population variation, effectively inflating within-subspecies variance, while potentially
631 lowering the total variance attributed to between-subspecies differentiation. This sampling
632 strategy may have decreased the likelihood that distinctiveness (as defined by Ramey et al. 2005)
633 would be demonstrated. While both sampling strategies have their strengths and weaknesses and
634 a combination of the two may ultimately be preferred, we believe the strategy used in the present
635 study is the more pragmatic approach for testing statistical significance under the hypothesis
636 testing approach espoused by REA, in which uniqueness was based on partitioning of variation
637 within and between subspecies. The systemic error identified in the REA CR region sequences
638 prevented an analysis of the combined data.

639 Large numbers of microsatellite DNA loci provide better estimates of population genetic
640 parameters given the stochastic variation (including size homoplasy) expected among
641 independent loci (Adams et al. 2004, Takezaki and Nei 1996). Given that microsatellite loci are

642 densely interspersed in eukaryotic genomes (Katti et al. 2001), five polymorphic microsatellite
643 DNA loci may not provide sufficient representation of the nuclear DNA variation present
644 (Beaumont and Nichols 1996, Pritchard and Rosenberg 1999) or provide a sufficiently robust
645 multilocus genotype for population/phylogeographic comparisons (Smouse and Chevillon 1998).
646 In a test to determine the number of diagnosable clusters present in their microsatellite DNA
647 dataset using *STRUCTURE*, REA found the variability of likelihood estimates to be high, and
648 concluded that the eight clusters identified by the analysis were poorly defined. This result is
649 consistent with the findings of Pritchard et al. (2000) in which simulations found five loci were
650 insufficient to resolve clusters. The 21 loci surveyed in the present study allowed identification
651 of both deep and shallower levels of population structure, consequently resulting in six definable
652 clusters. The present study also generated larger estimates of subspecies differentiation ($F_{ST,}$
653 0.11 – 0.34; R_{ST}, 0.17 – 0.78) than found in the previous study (F_{ST}, 0.01-0.16), which is
654 consistent with the AMOVA that found a higher percentage of variation was distributed between
655 subspecies (17.4%) than reported by REA (7.5% - 8.9%). The increased coverage of the nuclear
656 genome afforded by the larger number of microsatellite loci yielded increased resolution of the
657 phylogeographic structure present in *Z. hudsonius*. R_{ST} differentiation observed among most of
658 the subspecies is considerable, not of recent occurrence, and not simply the result of genetic
659 drift.

660 The more than two-fold differential in the estimate of variation between subspecies
661 provided by Φ_{ST} (this study) and F_{ST} (REA) broaches important issues regarding the conclusions
662 drawn from the respective studies and underscores the fundamental differences between the two
663 test statistics (Φ_{ST} and F_{ST}) generated by AMOVA. Φ_{ST} incorporates sequence divergence
664 between haplotypes (providing the option of several distance metrics), as well as determining
665 haplotype frequencies. Evolutionary differences among the haplotypes (i.e., mutations) are
666 incorporated into calculations of the test statistics. F_{ST}, which utilizes conventional F-statistics,
667 treats all haplotypes, regardless of their evolutionary interrelationships, as equally differentiated
668 (i.e., distance = 1.0), and assesses variance distribution based on haplotype frequencies alone.
669 Differences between haplotype frequencies are assumed to be due to genetic drift. Given that no
670 haplotypes were shared among the five *Z. hudsonius* subspecies in the present study, Φ_{ST} would
671 appear to be the most appropriate statistic as F_{ST} values (a measure of haplotype frequencies
672 only) have limited biological resolution or significance in this particular situation. When Φ_{ST}

673 and F_{ST} values are similar in magnitude, any population differentiation is likely due to genetic
674 drift, but when Φ_{ST} is large (as is the case for all subspecies comparisons in this study),
675 regardless of the F_{ST}, evolutionary depth in separations among haplotypes is implied. Moreover,
676 if REA's critical test of uniqueness for *Z. h. preblei* mtDNA sequence data were applied (i.e.,
677 greater molecular variance be demonstrated between subspecies than within), the results
678 obtained between all subspecies compared in the present study exceed this uniqueness test
679 criterion (all Φ_{ST} values were > 0.50). By opting to utilize F_{ST} and failing to recognize the
680 evolutionary differences among observed haplotypes, we contend that REA have utilized an
681 inappropriate criterion for determining subspecies uniqueness.
682 For microsatellite DNA comparisons, REA's critical test of uniqueness for subspecies
683 and historic genetic exchangeability was two-fold: most variation was observed between
684 subspecies in pairwise AMOVA comparisons (i.e., $F_{ST} \geq 0.50$) and multiple private alleles are
685 found at higher frequency than shared alleles at a majority of loci. These criteria often are not
686 met among accepted species due to constraints on fixation indices generated from microsatellite
687 DNA loci created by high heterozygosity levels and homoplasy (Hedrick 1999; Balloux et al.
688 2000). Although several R_{ST} values observed among the five subspecies compared in this study
689 approached or exceeded the proposed criterion (Table 3), we contend the proposed criteria for
690 microsatellite markers are unfounded because of the longer coalescent times associated with the
691 nuclear genome (four times longer than mitochondrial DNA), and they likely have little
692 biological meaning because a fundamental evolutionary process (e.g., mutation) is ignored.
693 In phylogeographic studies, principles of genealogical concordance have become a
694 conceptual framework for identifying deeper units within a species (Ball and Avise 1992; Avise
695 2000). Data presented here for *Z. hudsonius* subspecies are genealogically concordant in 3 out
696 of 4 levels defined by Avise (2000): concordance between DNA characters (I); concordance
697 between genomes (II); and concordance of gene tree partitions with independent biogeographic
698 and systematic information (IV). First, there was concordance between the patterns of change in
699 DNA sequences from the two mtDNA regions surveyed, seen in the similar, subspecies-specific
700 patterns apparent in analyses of each gene region (Appendices D and E; Figure 4) and from the
701 fact that branch support for several nodes distinguishing subspecies increases, and becomes
702 significant, when datasets are combined (Figure 5A). The addition of more characters (e.g. cyt *b*
703 and CR) shows that there is more "depth" to the divergence between *Z. h. preblei*, *Z. h.*

campestris and *Z. h. intermedius* (i.e. greater number of inferred substitutions separating haplotypes and greater genetic distances) than examination of fewer bases of mtDNA revealed. Fine-scale patterns also become apparent, such as genetic structuring between northern and southern *Z. h. preblei* populations. Because intraspecific phylogeographic datasets are predisposed to have fewer informative characters and relationships between haplotypes are unlikely to be hierarchical in nature, traditional phylogenetic methods can lead to inadequately portrayed genealogical relationships (Brower et al. 1996). The haplotype network analysis adopted here depicted ancestral relationships and the nonrandom distribution of mutations among lineages, which resulted in the diagnostic haplotypic structure observed between subspecies and not observed in the more traditional parsimony analyses performed in this study or the distance-based approach utilized by REA. When forced into a phylogenetic comparison (e.g., parsimony analysis), *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius* exhibited shallow gene genealogies (at both CR and cyt *b*) – an intraspecific pattern consistent with ancestral polymorphism and incomplete lineage sorting; patterns that would be expected among recently diverged groups (Maddison 1997, Arbogast et al. 2002) and not uncommonly seen within rodent species (e.g., *Peromyscus boylii*, Tiemann-Boege et al. 2000; *Neotoma floridana*, Edwards and Bradley 2001; *Microtus agrestis*, Jaarola and Searle 2002). For the reason that evolution occurs at varying rates depending primarily on environmental- and demographic-based dynamics, we contend that sequence variation/divergence thresholds are inappropriate for intra- or interspecific taxonomic designations. Alternatively, as genetic variation is apportioned among and within taxa, decisions on taxonomic recognition should be rendered on a case-by-case basis. Lastly, concordant patterns observed in the two mtDNA genes and nuclear microsatellites (Figures 2 and 3) are in agreement with the disjunct geographic distributions and morphological differences that led to the subspecies descriptions (Krutzsch 1954, Hafner et al. 1981) and lends confidence that actual evolutionary relationships have been recovered.

CONSERVATION IMPLICATIONS

Criteria chosen for delineating isolated intraspecific groupings (e.g., subspecies, phylogroups, distinct population segments, evolutionary significant units, designatable units, or management units) become of critical importance when the entity's continued existence is at risk (Moritz

734 1994b, Paetkau 1999). Given that the issue at hand constitutes an intraspecific comparison, we
735 are concerned about the methodological pitfalls of applying approaches based on species-level
736 inference. Mayr (1942) defined a subspecies as "a geographically localized subdivision of the
737 species, which differs genetically and taxonomically from other subdivisions of the species."
738 While this definition infers that subspecies have unique ranges and are diagnosable (Patten and
739 Unitt 2002), they are not necessarily reproductively isolated or they would qualify as species.
740 Systematists have taken diverse positions on how "taxonomically" may be interpreted. Some
741 have asserted that this term infers long-term reproductive isolation and that the various
742 components should be monophyletic (see Zink 2004 and references therein). Recently diverged
743 intraspecific taxa can be characterized by multiple ancestral lineages at the time of divergence
744 and might not exhibit reciprocal monophyly (Hudson and Coyne 2002; Funk and Omland 2003;
745 Green 2005). We contend that since adaptation, diagnosability and biological speciation may
746 exist well before achieving reproductive isolation and monophyly, these latter criteria have no
747 basis under which to become standards for subspecies designation.
748 Similarly, genetic distinctness criteria that are inflexible represent a conservation
749 stratagem that could impact the potential for future evolutionary change within the intraspecific
750 unit of management and may lead to extirpation or extinction of discrete and evolutionarily
751 significant intraspecific diversity. We found Ramey et al.'s (2005) core genetic-based criteria
752 for *Z. hudsonius* subspecies distinctness (i.e., greater variation between subspecies than within)
753 to be 1) too stringent to account for relatively recently radiated taxa, 2) scientifically unfounded
754 as this criterion would not distinguish certain well-established species (Forbes et al. 1995; see
755 Vignieri et al. 2006 and references therein), 3) unachievable under some circumstances with the
756 test statistics applied (F_{ST}; Hedrick 1999), and 4) untenable given that important evolutionary
757 processes (e.g., mutation) are disregarded. Consequently, this criteria should not be considered
758 as an accepted standard when addressing ESA actions or for international conservation laws.
759 Alternatively, we recommend application of the requirements for describing subspecies under
760 the ESA proposed by Haig et al. (In Press), which state that all such designations should meet
761 the "discreteness" and "significance" criteria defined for listing a distinct population segment
762 (DPS) under the Act.
763 The level of discontinuity observed among the subspecies in this study should not be
764 considered minor or "shallow" (Avise 2004) simply because the level of DNA sequence

differentiation is relatively low, and not accompanied by the presence of reciprocal monophyly, demonstrable phenotypic divergence (but see Vignieri et al. 2006), or obvious adaptive significance. The differential magnitude observed between Φ_{ST} (distance and frequency) and F_{ST} (frequency) values for both mtDNA CR and cyt *b* suggests moderate microevolutionary depth among the five subspecies accompanied by phylogeographical structuring of haplotypes and multilocus genotypes within *Z. h. preblei*. The diagnostic nature of the haplotype variation between *Z. h. preblei* and other subspecies (there are no credible published data suggesting otherwise) and large degree of nuclear genetic differentiation at microsatellite loci suggests that these subspecies (and certain geographic populations within) have been reproductively isolated for such time that they appear to be on independent evolutionary trajectories, while having nearly achieved complete lineage sorting. We have identified strongly differentiated units of conservation (which could qualify as DPSs) that are entirely consistent with current taxonomy and support protection of the threatened population segment under the ESA. We conclude the subspecies surveyed in this study do not warrant synonymy, as proposed for *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius*; rather, they constitute distinct evolutionary lineages that merit separate management consideration, and those populations facing demographic challenges (e.g., *Z. h. preblei*-North, *Z. h. preblei*-South) should be afforded high conservation priority. The strong concurrence among patterns of mitochondrial and nuclear DNA variation observed in this study suggests that formal recognition of the relationships revealed will assist in preserving the potential for future evolutionary change within and among these subspecies.

A detailed comparison similar to that performed here for all subspecies of *Z. hudsonius* is warranted. Such a study would likely shed light on the significance of the considerable differentiation observed among the *Z. h. preblei, Z. h. campestris,* and *Z. h. intermedius* clade and will allow differentiation observed for *Z. h. pallidus* and *Z. h. luteus* to be observed in the best available context.

ACKNOWLEDGMENTS

The authors are deeply indebted to: Tanya Shenk, Colorado Division of Wildlife and Robert Schorr, Colorado National Heritage Program for providing *Z. hudsonius preblei* tissue samples; and Cheryl Parmenter, Terry Yates, and the Division of Genomics Resources, Museum of

Southwestern Biology, University of New Mexico, Albuquerque, NM, Paula Cushing, Denver Museum of Nature and Science, Denver, CO, Robert Timm, University of Kansas Natural History Museum, Lawrence, KS, and Alex Schubert, Region 6, US Fish and Wildlife Service (USFWS) for provision of *Zapus* tissues. Field assistance was kindly provided by Laura Ellison, J.P. Farrar, Keith Geluso, Britta Schielke, and Ernie Valdez. Sampling by USGS field crews would not have been possible without the help of: Doug Backlund, South Dakota Department of Game, Fish, and Parks; Cara Staab, Black Hills National Forest, USFS; Tom Koerner, Shilo Comeaux-Kingfisher, and Scott Gonnion, Lacreek National Wildlife Refuge; Tom Whitford and Rhonda O'Byrne, Custer National Forest; Eugene Williams, Sand Lake National Wildlife Refuge; Martin Grenier, Wyoming Game and Fish Department; and Jerry Pecha, Nebraska Game and Parks Division. The authors would like to thank Fred Allendorf, Peter Dratch, Steve Chambers, Laurent Excoffier, Guillaume Laval, Brian Bowen, Peter Smouse, and Joseph Quattro for helpful discussions regarding study design and statistical analysis. David Hafner, Eric Hallerman, Andrew Martin, Sara Oyler-McCance, Ed Pendleton, Sacha Vignieri, and Rocky Ward provided constructive comments on the first draft of this manuscript. Ronald Adkins, Joel Cracraft, Hopi Hoekstra, Maarit Jaarola, Kim Scribner, Robin Waples, and Mike Wooten provided substantive peer reviews of a later draft. The authors are enormously indebted to all of these scientists for their unselfish efforts to improve this presentation. The authors thank Robin Johnson for laboratory assistance and Anne Henderson for assistance with editing. This project was jointly funded by the USFWS, US Geological Survey (USGS)-Leetown Science Center's Aquatic Ecology Branch, and the USGS Science Support Program. Lastly, the authors wish to thank Dr. Loren Rieseberg and three anonymous reviewers for their constructive criticisms and meaningful contributions to this manuscript.

LITERATURE CITED

Adams RI, Brown KM, Hamilton MB (2004) The impact of microsatellite electromorph size homoplasy on multilocus population structure estimates in a tropical tree (*Corythophora alta*) and an anadromous fish (*Morone saxatilis*). *Molecular Ecology*, **13,** 2579-2588.

Anderson KW, Jone JK (1971) Mammals of northwestern South Dakota. University of Kanas Publications, Museum of Natural History, **19**, 361-393.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics*, **33**, 707-740.

Avise JC (2000) *Phylogeography.* Harvard University Press, Cambridge, Massachusetts.

Avise JC (2004) Molecular markers, natural history, and evolution. Second edition. Sinauer Associates, Sunderland, Massachusetts.

Avise JC, Ball RM (1990) Principles of genealogical concordance in species concepts and biological taxonomy. In: Oxford Surveys in Evolutionary Biology. (eds. Futuyma D, Antonovics J), pp. 45-67. Oxford University Press, Oxford, United Kingdom.

Ball RM, Avise JC (1992) Mitochondrial DNA phylogeographic differentiation among avian populations and the evolutionary significance of subspecies. *Auk*, **109**, 626-636.

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J (2000). Microsatellites can be misleading: an empirical and simulation study. *Evolution*, **54,** 1414-1422.

Baverstock PR, Moritz C (1996) Project design. In: Molecular Systematics, Second Edition (eds. Hillis DM, Mortiz C, Mable BK) pp.17-27. Sinauer Associates Inc., Sunderland, Massachusetts.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **263,** 1619-1626.

Bremer K (1988) The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution*, **42,** 795-803.

Bremer K (1994) Branch support and tree stability. *Cladistics*, **10,** 295-304.

Brower AVZ, DeSalle R, Vogler A (1996) Gene trees, species trees, and systematics. *Annual Review of Ecology and Systematics*, **27,** 423-450.

Clark CM, Wentworth TR, O'Malley DM (2000) Genetic discontinuity revealed by chloroplast microsatellites in eastern North American *Abies* (Pinaceae). *American Journal of Botany*, **87,** 774-782.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9,** 1657-1659.

Connor MM, Shenk TM (2003) Distinguishing *Zapus hudsonius preblei* from *Zapus princeps princeps* by using repeated cranial measurements. *Journal of Mammalogy*, **84,** 1456-1463.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution*, **15,** 290-295.

Crandall KA, Templeton AR, Sing CF (1994) Intraspecific cladogram estimation: problems and solutions. In: Models in Phylogeny Reconstruction. (eds. Scotland RW, Siebert DJ, Williams DM), pp. 273-297. Clarendon Press, Oxford, United Kingdom.

Edwards, CW, Bradley RD (2001) Molecular phylogenetics of the *Neotoma floridana* species group. *Journal of Mammalogy*, **82,** 791-798.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L) Skeels] endemic of Morocco. *Theoretical and Applied Genetics*, **92,** 832-839.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14,** 2611-2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131,** 479-491.

Farris JS, Kallersjo M, Kluge AJ, Bult C (1994) Testing significance of incongruence. *Cladistics*, **10**, 315-319.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39,** 783-791.

Forbes HS, Hogg JT, Buchanan FC, Crawford AM, Allendorf FW (1995) Microsatellite evolution in congenerics mammals: Domestic and Bighorn sheep. Molecular Biology and Evolution, **12**, 1106-1113.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: Frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics,* **34**, 397-423.

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F statistics. *Journal of Heredity*, **86,** 485-486.

Grant WS (1987) Genetic divergence between congeneric Atlantic and Pacific Ocean fishes. In: Population genetics and fishery management. (eds. Ryman N, Utter F) pp. 225-246. University of Washington Press, Seattle, Washington.

Green DM (2005) Designatable units for status assessment of endangered species. *Conservation Biology*, **19**, 1813-1820.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics*, **48,** 361-372.

Hafner DJ, Peterson KE, Yates TL (1981) Evolutionary relationships of jumping mice (Genus *Zapus*) of the southwestern United States. *Journal of Mammalogy,* **62,** 501-512.

Haig SM, Beever EA, Chambers SM et al. (In Press) Review: Taxonomic considerations in listing subspecies under the U.S. Endangered Species Act. *Conservation Biology*.

Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53,** 313-318.

Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. Evolution, **56**, 1557-1565.

Huelsenbeck JP, Ronquist F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, **17**, 754—755.

Iudica W, Whitten M, Williams NH (2001) Small Bones from Dried Mammal Museum Specimens as a Reliable Source of DNA *BioTechniques*, **30,** 732-736.

Jaarola M, Searle JB (2002) Phylogeography of field voles (*Microtus agresti*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology*, **11,** 2613-2621.

Jones GS (1981) The systematics and biology of the genus *Zapus* (Mammalia, Rodentia, Zapodidae). PhD Thesis, Indiana State University.

Jones JK, Armstrong DM, Hoffmann RS, Jones C (1983) *Mammals of the Northern Great Plains*. University of Nebraska Press, Lincoln, Nebraska.

Katti MV, Ranjekar PK, Gupta VS (2001) Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution*, **18,** 1161-1167.

King TL, Lubinski BA, Spidle AP (2001) Microsatellite DNA variation in Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) and cross-species amplification in the Acipenseridae. *Conservation Genetics*, **2,** 103-119.

King TL, Eackles MS, Young CC (2006) Microsatellite DNA markers for assessing phylogeographic and population structure in Preble's meadow jumping mice (*Zapus hudsonius preblei*) and cross-amplification among neighboring taxa. *Molecular Ecology Notes*. Published on-line 06 July 2006.

Kirchman JJ, Hackett SJ, Goodman SM, Bates JM (2001) Phylogeny and systematics of ground rollers (Brachypteraciidae) of Madagascar. *The Auk*, **118,** 849-863.

Kocher, TD, Thomas WK, Meyer RA, Edwards SV, Paabo S, Villablanca FX, Wilson AC (1989) Dynamics of Mitochondrial-DNA Evolution in Animals - Amplification and Sequencing with Conserved Primers. *Proceedings of the National Academy of Sciences of the United States of America*, **86,** 6196-6200.

Krutzsch PH (1954) North American jumping mice (genus *Zapus)*. University of Kansas Publication Museum of Natural History, **4,** 349-472.

Maddison WP (1997) Gene trees in species trees. *Systematic Biology* **46**, 523-536.

Mayr E (1942) *Systematics and the origin of species*. Columbia University Press, New York.

Moritz,C (1994a) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology*, **3,** 401-411.

Moritz C (1994b) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology & Evolution*, **9,** 373-375.

Moritz C, Hillis DM (1996) Molecular systematics: context and controversies. In: Molecular Systematics, Second Edition (eds. Hillis DM, Mortiz C, Mable BK) pp.1-13. Sinauer Associates Inc., Sunderland, Massachusetts.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, **19,** 153-170.

Paetkau D (1999) Using genetics to identify intraspecific conservation units: a critique of current methods. *Conservation Biology*, **13,** 1507-1509.

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, **12,** 357-358.

Patten MA, Unitt P (2002) Diagnosability versus mean differences of Sage Sparrow subspecies. Auk **119:**26–35.

Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. *Systematic Biology,* **50,** 580-601.

Preble EA (1899) Revision of the jumping mice of the genus *Zapus*. *North American Fauna,* **15,** 1-41.

Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. *American Journal of Human Genetics,* **65:** 220-228.

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

Ramey RR, Liu HP, Epps CW, Carpenter LM, Weshausen JD (2005) Genetic relatedness of the Preble's meadow jumping mouse (*Zapus hudsonius preblei*) to nearby subspecies of *Z. hudsonius* as inferred from variation in cranial morphology, mitochondrial DNA and microsatellite DNA: implications for taxonomy and conservation. *Animal Conservation,* **8,** 329-346.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86,** 248-249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43,** 223-225.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN: A software for population genetics data analysis*. University of Geneva, Genetics and Biometry Laboratory, Geneva, Switzerland.

Shields GF, Kocher TD (1991) Phylogenetic-Relationships of North-American Ursids Based on Analysis of Mitochondrial-DNA. *Evolution*, **45,** 218-221.

Smouse PE, Chevillon C (1998) Analytical aspects of population-specific DNA fingerprinting for individuals. *Journal of Heredity*, **89,** 143-150.

Sokal RR, Rohlf ZJ (1994) *Biometry: The Principles and Practice of Statistics in Biological Research.* Third Edition. W.H. Freeman & Company, New York, NY.

Sorenson MD (1999) *TreeRot, version 2*. Boston University, Boston, MA.

Steinberg EK (1999) Characterization of polymorphic microsatellites from current and historic populations of North American pocket gophers (Geomyidae: *Thomomys*). *Molecular Ecology*, **8,** 1075-1092.

Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0.* Sinauer Associates, Sunderland, MA.

Swofford DL, Selander, RB (1981) BIOSYS-1: A Fortran program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity*, **72,** 282-283.

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, **144,** 389-399.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10,** 512-526.

Templeton AR, Crandall KA, CF Sing (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132,** 619-633.

Thompson JD, Higgins DG,Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research***, 22,** 4673-4680.

Tiemann-Boege I, Kilpatrick CW, Schmidly DJ, Bradley RD (2000) Molecular phylogenetics of the *Peromyscus boylii* species group (Rodentia: Muridae) based on mitochondrial cytochrome *b* sequences. *Molecular Phylogenetics and Evolution*, **16,** 366-378.

U.S. Fish and Wildlife Service (1998) Final rule to list the Preble's meadow jumping mouse as a threatened species. *The Federal Register*, **53,** 26517-26530.

U.S. Fish and Wildlife Service (2005) Endangered and threatened wildlife and plants; 12-Month Finding on a Petition to Delist the Preble's Meadow Jumping Mouse (Zapus hudsonius preblei) and Proposed Delisting of the Preble's Meadow Jumping Mouse. *The Federal Register*, **70,** 5404-5411.

Vignieri SN (2003) Isolation and characterization of eight highly variable microsatellite markers in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology Notes*, **3,** 638-640.

Vignieri SN (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology* **14**, 1925-1937.

Vignieri SN, Hallerman EM, Bergstrom BJ, Hafner DJ, Martin AP, Devers P, Grobler P, Hitt N (2006) Mistaken view of taxonomic validity undermines conservation of an evolutionarily distinct mouse: a response to Ramey et al. (2005). Animal Conservation. Published on-line 10 May 2006.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38,** 1358-1370.

Zink RM (2004) The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society B*, **271**, 561–564.

Table 1. Sample size, allelic richness, number of private alleles, observed and expected heterozygosity, and estimates of F_{IS} observed for 14 collections of *Zapus hudsonius* representing five neighboring subspecies surveyed at 21 microsatellite DNA loci.

Subspecies	Collection abbreviation	Collection locality	Cluster[1]	N	A^2	Private alleles	H_o	H_e	F_{IS}
Z. h. preblei				**170**	**6.7**	**[5]**	**0.539**	**0.624**	**0.136**
	SOWY	Laramie, Albany, Platte, and Converse Cos., WY		(28)	4.7	(1)	0.499	0.553	0.099
	LCCO1	N. Fk. Cache la Poudre River, Larimer Co., CO		(14)	3.4	(2)	0.507	0.502	-0.010
	LCCO2	Stove Prairie Creek, Larimer Co., CO		(16)	3.4		0.536	0.528	-0.015
			Z. h. preblei - North	58	5.4		0.511	0.614	0.169
	DCCO1	East Plum Creek, Douglas Co., CO		(34)	4.3		0.538	0.535	-0.005
	DCCO2	Indian Creek, Douglas Co., CO		(30)	4.0	(1)	0.538	0.540	0.004
	ECCO1	U.S. Air Force Academy, El Paso Co., CO		(22)	4.4	(1)	0.563	0.584	0.037
	ECCO2	Monument Creek, El Paso Co., CO		(26)	4.0		0.586	0.559	-0.050
			Z. h. preblei - South	112	5.7		0.553	0.583	0.052
Z. h. campestris			*Z. h. campestris*	**61**	**7.2**	**[2]**	**0.637**	**0.670**	**0.051**
	CCWY	Beaver Creek, Crook Co., WY		(30)	6.2		0.648	0.662	0.021
	CCSD	Iron /Willow Creeks, Custer/Pennington Cos., SD		(31)	6.0	(1)	0.625	0.654	0.045
Z. h. intermedius			*Z. h. intermedius*	**49**	**9.4**	**[22]**	**0.649**	**0.703**	**0.079**
	BRCSD	Columbia Road Reservoir, Brown Co., SD		(28)	6.1	(8)	0.619	0.637	0.029
	MCMN	Camp Ripley, Morrison Co., MN		(21)	8.3	(13)	0.687	0.735	0.067
Z. h. pallidus			*Z. h. pallidus*	**48**	**9.4**	**[33]**	**0.752**	**0.790**	**0.049**
	BCSD	Cedar Creek, Bennett Co., SD		(16)	6.9	(9)	0.738	0.757	0.026
	KBCNE	N. Channel Platte River, Kearney/Buffalo Cos., NE		(32)	7.5	(12)	0.759	0.789	0.038
Z. h. luteus	SCNM	Multiple sites, Sandoval Co., NM	*Z. h .luteus*	**20**	**4.6**	**[8]**	**0.576**	**0.623**	**0.076**

[1]Clusters (k = 6) were determined using the program *STRUCTURE*.
[2]Allelic richness
[3][ ] indicates the number of private alleles limited to 1 or all collections within a subspecies.

Table 2. Percentage of *Zapus hudsonius* individuals correctly assigned to one of $k = 3$ initial clusters and $k = 6$ subsequent clusters (subclusters) identified by the program *STRUCTURE* from a survey of 14 collections of mice representing five geographically proximal subspecies surveyed at 21 microsatellite loci. Individuals were assigned to cluster or subcluster based on the largest value of q. Average qMAX and percentage of *Zapus hudsonius* individuals correctly assigned to one of $k = 2$ subclusters identified by the program *STRUCTURE* based on the $q > 0.90$ criterion are also provided.

Subspecies	**Initial cluster (k = 3)**			**Mean qMAX**	**Cluster**	**[A]**		**[B]**		**[C]**		**Subspecies/ cluster designation**	**Mean qMAX**	**Percentage assigned at $q > 0.90$**
Collection	**[A]**	**[B]**	**[C]**		**subcluster (k = 6)**	**1**	**2**	**3**	**4**	**5**	**6**			
Z. h. preblei												***Z. h. preblei***	**0.97**	**100.0% (170)**
SOWY	**1.00**				**1**	**1.00**								
LCCO1	**1.00**				**1**	**0.93**	**0.07**							
LCCO2	**1.00**			**0.98**	**1**	**0.94**	**0.06**					**North**	**0.94**	**89.7% (52/58)**
DCCO1	**1.00**				**2**		**1.00**							
DCCO2	**1.00**				**2**		**1.00**							
ECCO1	**1.00**				**2**		**1.00**							
ECCO2	**1.00**				**2**		**1.00**					**South**	**0.98**	**98.2% (110/112)**
Z. h. campestris												***Z. h. campestris***	**0.98**	**95.1% (58/61)**
CCWY		**1.00**			**3**			**1.00**						
CCSD		**1.00**			**3**			**1.00**						
				0.96										
Z. h. intermedius												***Z. h. intermedius***	**0.99**	**100.0% (49)**
BRCSD		**1.00**			**4**				**1.00**					
MCMN		**1.00**			**4**				**1.00**					
Z. h. pallidus												***Z. h. pallidus***	**0.99**	**100.0% (48)**
BCSD			**1.00**		**5**					**1.00**				
KBCNE			**1.00**	**0.99**	**5**					**1.00**				
Z. h. luteus												***Z. h. luteus***	**0.99**	**100.0% (20)**
SCNM			**1.00**		**6**						**1.00**			

Table 3. F_{ST} values (below the diagonal) generated from a survey of 21 microsatellite loci in five geographically proximal subspecies of *Zapus hudsonius*. All F_{ST} estimates were statistically significant from zero ($\alpha = 0.05$, $P < 0.001$) after 1,000 permutations. R_{ST} values (above the diagonal) are provided for comparison with F_{ST}.

Subspecies	***Z. h. preblei***	***Z. h. campestris***	***Z. h. intermedius***	***Z. h. pallidus***	***Z. h. luteus***
Z. h. preblei		0.1561	0.4442	0.6259	0.7708
Z. h. campestris	0.1063		0.2075	0.4759	0.6474
Z. h. intermedius	0.1810	0.1069		0.3455	0.4910
Z. h. pallidus	0.2146	0.1812	0.1555		0.1816
Z. h. luteus	0.32500	0.3032	0.2714	0.1832	

Table 4. Sequence diversity at the mitochondrial DNA control (372 base pairs) and cytochrome *b* (1006 base pairs) regions observed in five neighboring subspecies of *Zapus hudsonius*.

	Control Region				**Cytochrome B**				**Combined**			
Taxon	**N**	**Polymorphic sites**	**Haplotypes**	**Nucleotide diversity**	**N**	**Polymorphic sites**	**Haplotypes**	**Nucleotide diversity**	**N**	**Polymorphic sites**	**Haplotypes**	**Nucleotide diversity**
Z. h. preblei	**160**	**3**	**4**	**0.0030**	**146**	**16**	**21**	**0.0013**	**133**	**19**	**25**	**0.0017**
Z. h. intermedius	**47**	**9**	**9**	**0.0041**	**47**	**22**	**13**	**0.0028**	**45**	**31**	**17**	**0.0032**
Z. h. campestris	**61**	**5**	**5**	**0.0025**	**60**	**13**	**13**	**0.0010**	**59**	**18**	**15**	**0.0013**
Z. h. pallidus	**47**	**9**	**5**	**0.0060**	**48**	**13**	**8**	**0.0030**	**47**	**22**	**9**	**0.0038**
Z. h. luteus	**19**	**1**	**2**	**0.0000**	**19**	**0**	**1**	**0.0000**	**18**	**1**	**2**	**0.0000**

FIGURE LEGENDS

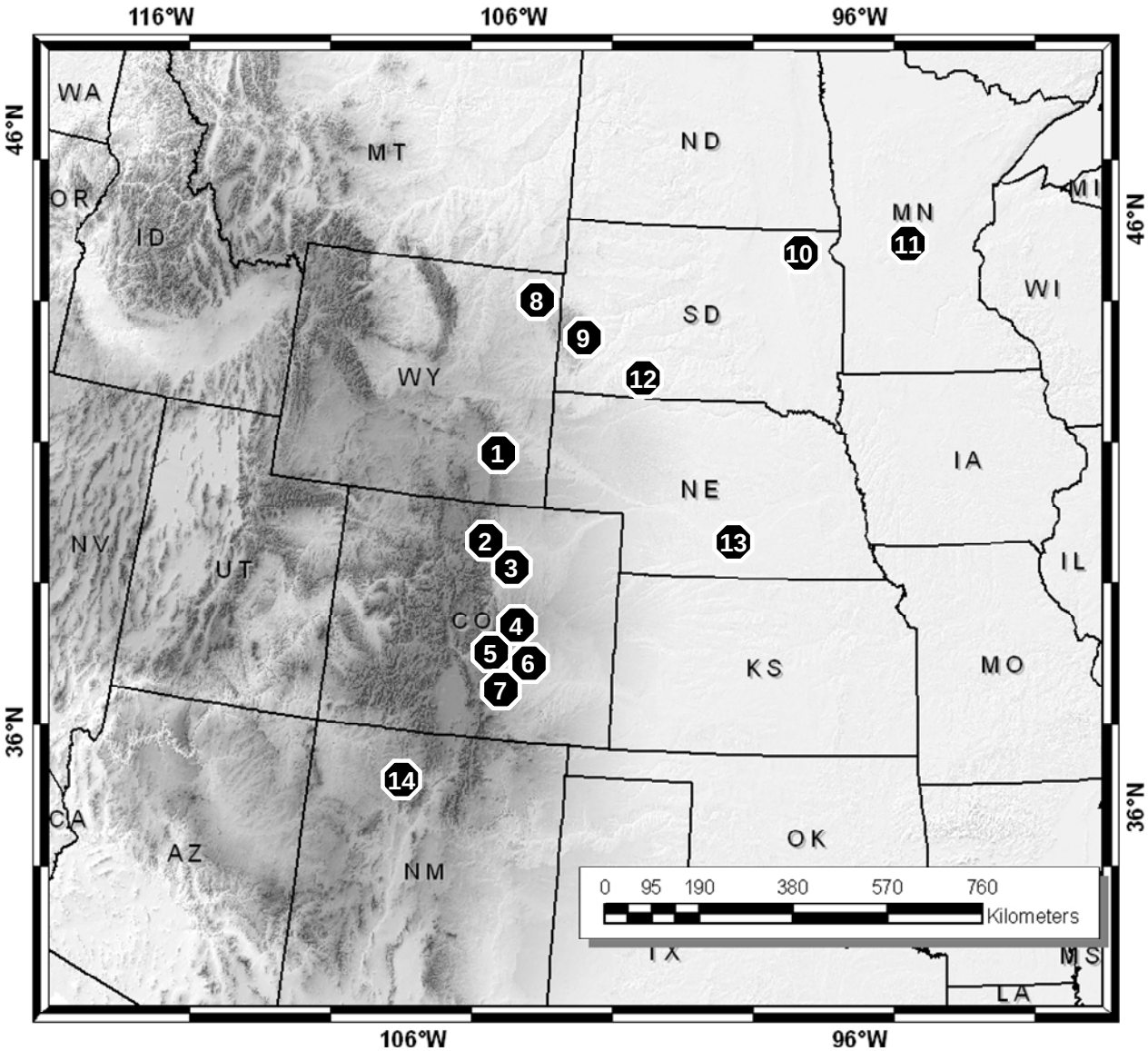
Figure 1. Generalized collection sites (N=14) representative of five nominal subspecies of *Zapus hudsonius* utilized in this study.

Figure 2. Summary plots of *q* estimates generated by the sequential cluster analysis of the program *STRUCTURE* performed on the multilocus (N=21) genotypes of 14 collections of *Zapus hudsonius*. The number of inferred clusters (*k*) in the initial (uppermost hierarchical level) analysis was three (clusters [A-C]). Each initial cluster was subsequently analyzed for within-cluster structure. The sequential analysis further subdivided each cluster into two subclusters for a total of six clusters (1-6). Each individual is represented by a single vertical line, broken into *k* colored segments, the length of which is proportional to the membership fraction in each of the *k* clusters. Individuals are grouped by populations and subspecies as indicated by brackets.

Figure 3. Unrooted neighbor-joining tree generated from pairwise genetic distance (*D*a; Nei et al. 1983) values between all collections of *Zapus hudsonius* generated from multilocus microsatellite genotypes. Branch lengths are proportional to *D*a units. Numbers along branches represent bootstrap support for nodes.

Figure 4. *Zapus hudsonius* haplotype networks resulting from the statistical parsimony analysis of the cyt *b* data set. The set of haplotypes belonging to each subspecies of *Z. hudsonius* are outlined by boxes. No haplotypes were shared among subspecies. Haplotypes identified as the root of each network by the program TCS (Clement et al. 2000) are represented by a square and all other haplotypes are represented as circles. The size of each haplotype symbol is proportional to the number of copies observed in the data set. Haplotypes for *Z. h. pallidus* and *Z. h. luteus* are not connected to the network containing *Z. h. preblei* due to the differentiation exceeding the 95% limit for connections for cyt *b* (13%).

Figure 5. Phylogenetic hypotheses for five *Z. hudsonius* subspecies based upon combined mitochondrial DNA data: A) strict consensus of 8032 most parsimonious trees (359 steps, CI = 0.8273) and B) phylogram resulting from partitioned Bayesian analysis . Numbers above branches in A are bootstrap proportions and italicized numbers below branches in A are decay indices, numbers above branches in B are posterior probabilities of nodes.



Zapus hudsonius **Collection Sites**

Z. h. preblei
1. **SOWY**
2. **LCC01**
3. **LCC02**
4. **DCC01**
5. **DCC02**
6. **ECC01**
7. **ECCO2**

Z. h. campestris
8. **CCWY**
9. **CCSD**

Z. h. intermedius
10. **BRCSD**
11. **MCMN**

Z. h. pallidus
12. **BCSD**
13. **KBCNE**

Z. h. luteus
14. **SCNM**

Figure 1

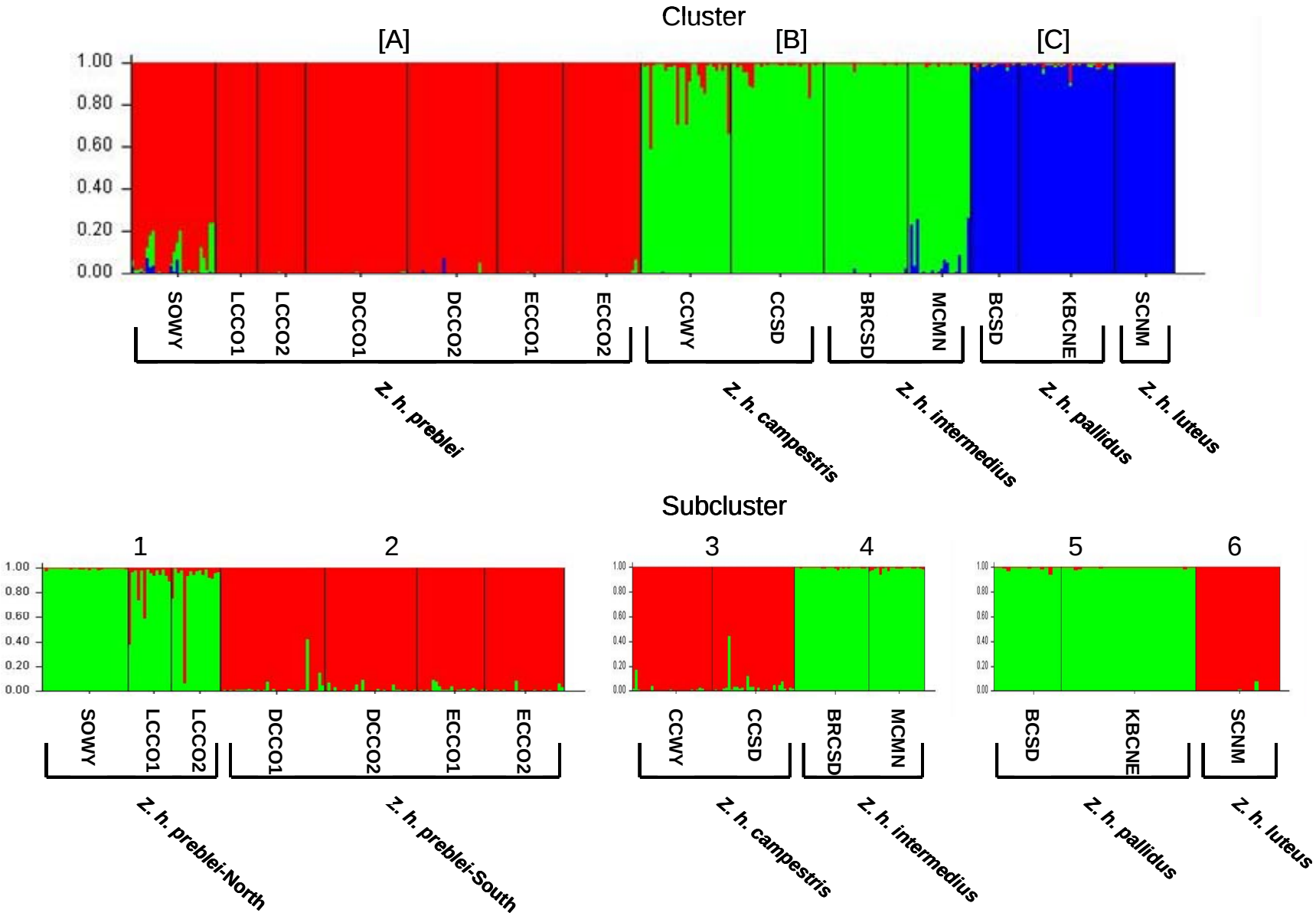

Cluster
[A]
[B]
[C]
1.00
0.80
0.60
0.40
0.20
0.00
SOWY
LCCO1
LCCO2
DCCO1
DCCO2
ECCO1
ECCO2
CCWY
CCSD
BRCSD
MCMN
BCSD
KBCNE
SCNM
Z. h. preblei
Z. h. campestris
Z. h. intermedius
Z. h. pallidus
Z. h. luteus
Subcluster
1
2
3
4
5
6
Z. h. preblei-North
Z. h. preblei-South
Z. h. campestris
Z. h. intermedius
Z. h. pallidus
Z. h. luteus

Figure 2

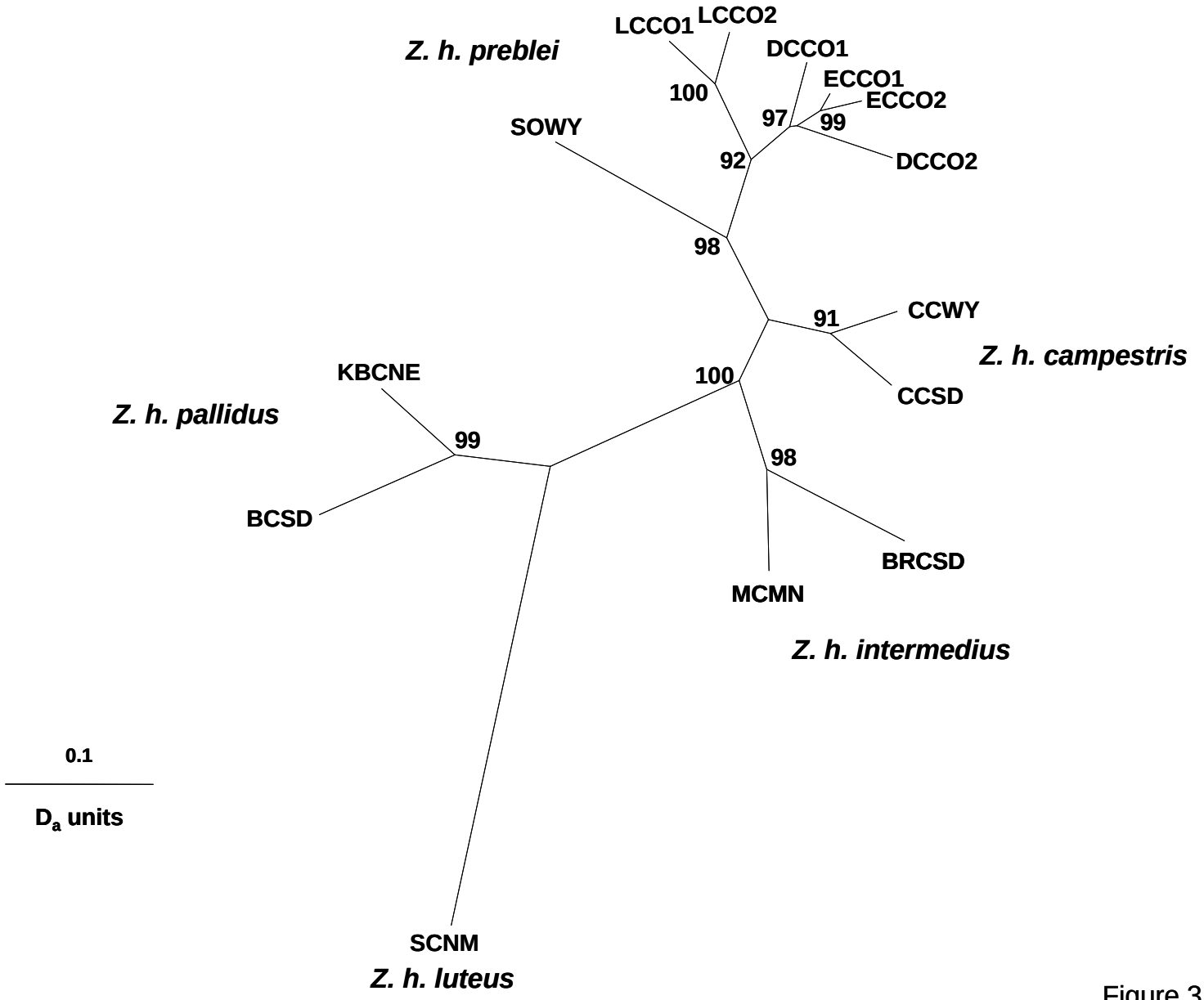

LCCO1
LCCO2
Z. h. preblei
DCCO1
ECCO1
ECCO2
100
97
99
SOWY
92
DCCO2
98
91
CCWY
Z. h. campestris
KBCNE
100
CCSD
Z. h. pallidus
99
98
BCSD
BRCSD
MCMN
Z. h. intermedius
0.1
Da units
SCNM
Z. h. luteus

Figure 3

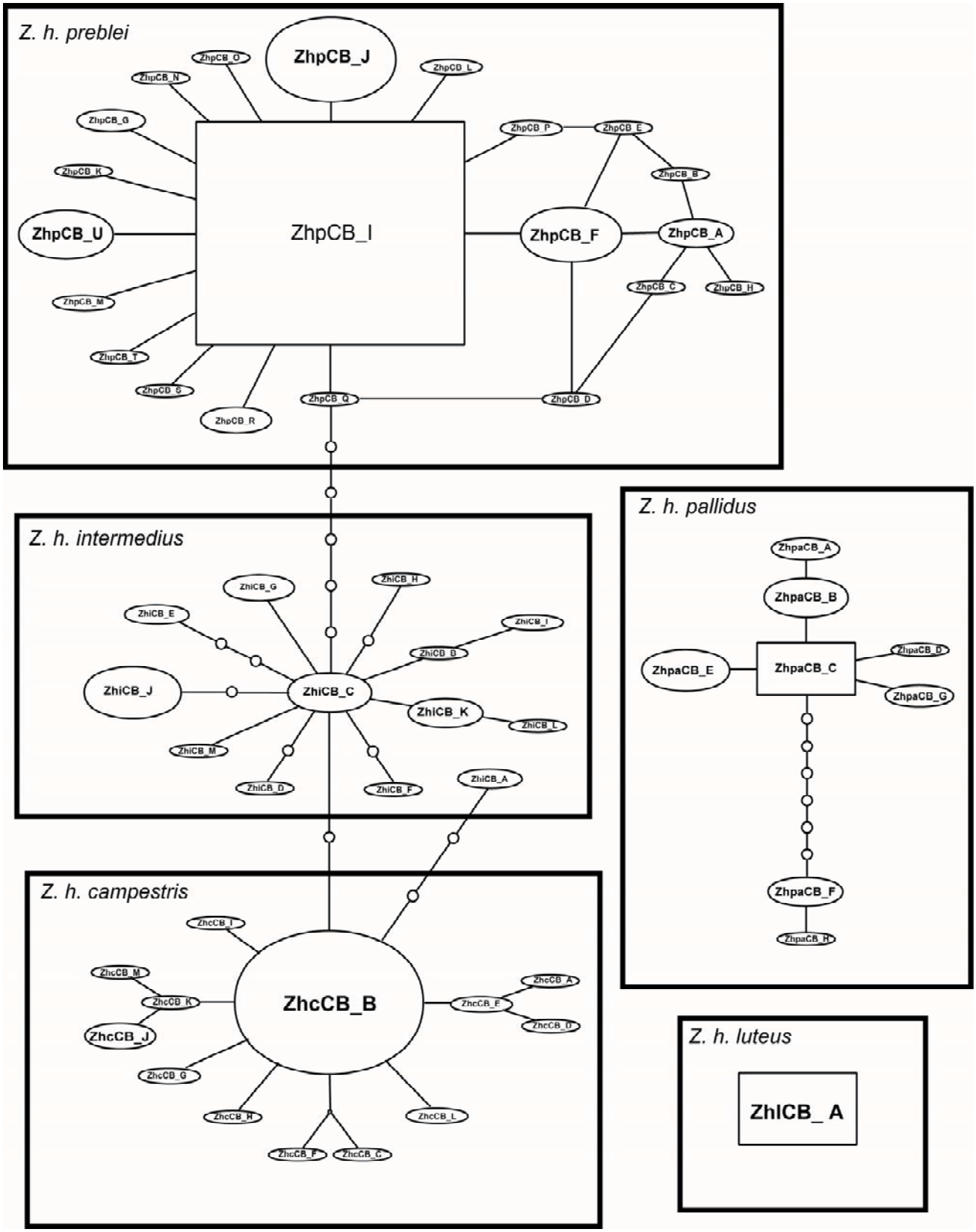

Z. h. preblei
ZhpCB_J
ZhpCB_O
ZhpCB_N
ZhpCB_L
ZhpCB_G
ZhpCB_P
ZhpCB_E
ZhpCB_K
ZhpCB_B
ZhpCB_U
ZhpCB_I
ZhpCB_F
ZhpCB_A
ZhpCB_C
ZhpCB_H
ZhpCB_M
ZhpCB_T
ZhpCB_S
ZhpCB_Q
ZhpCB_D
ZhpCB_R
Z. h. intermedius
ZhiCB_G
ZhiCB_H
ZhiCB_E
ZhiCB_I
ZhiCB_B
ZhiCB_J
ZhiCB_C
ZhiCB_K
ZhiCB_L
ZhiCB_M
ZhiCB_D
ZhiCB_F
ZhiCB_A
Z. h. pallidus
ZhpaCB_A
ZhpaCB_B
ZhpaCB_E
ZhpaCB_C
ZhpaCB_D
ZhpaCB_G
ZhpaCB_F
ZhpaCB_H
Z. h. campestris
ZhcCB_I
ZhcCB_M
ZhcCB_A
ZhcCB_K
ZhcCB_B
ZhcCB_E
ZhcCB_D
ZhcCB_J
ZhcCB_G
ZhcCB_H
ZhcCB_L
ZhcCB_F
ZhcCB_C
Z. h. luteus
ZhlCB_ A

Figure 4

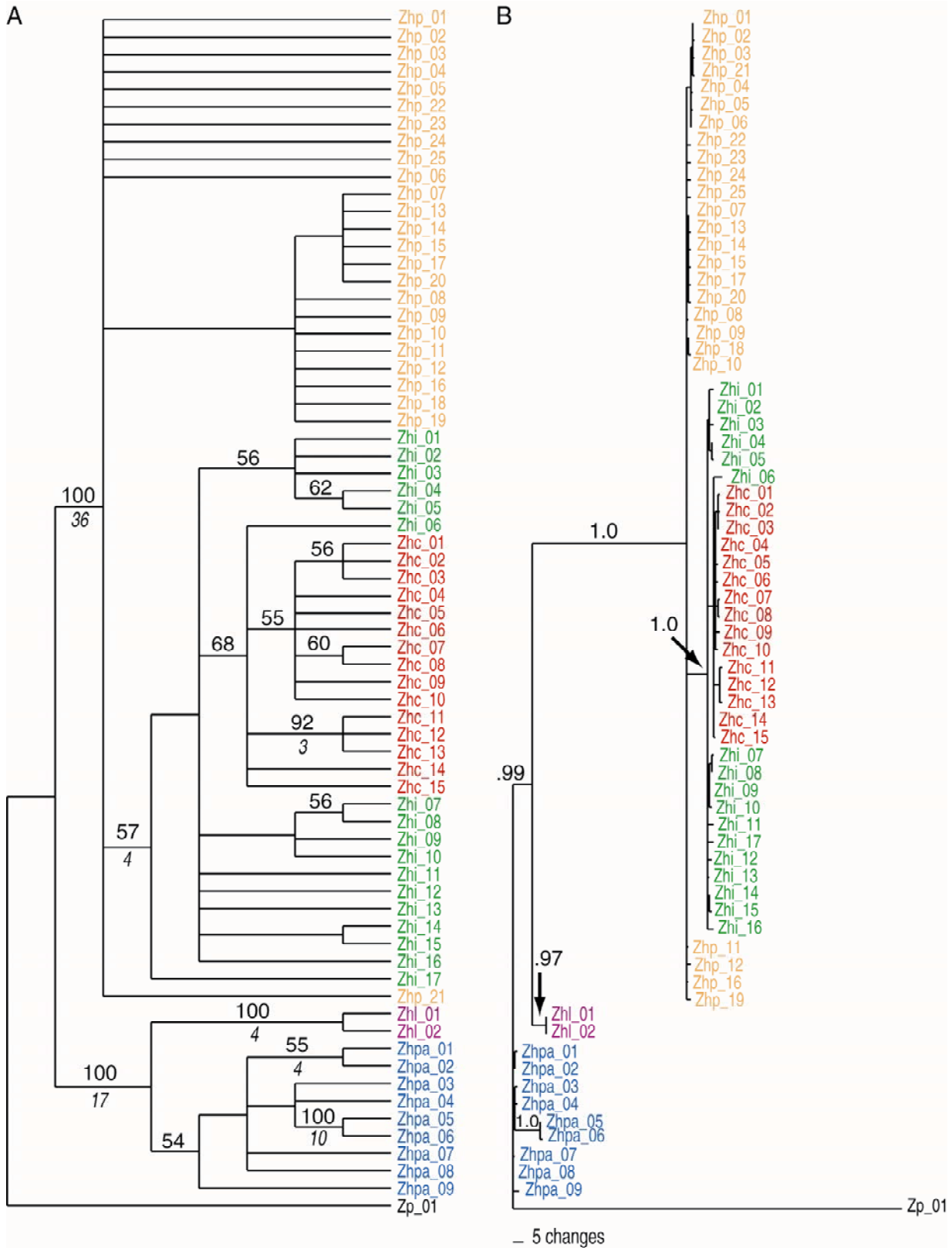

A
Zhp_01
Zhp_02
Zhp_03
Zhp_04
Zhp_05
Zhp_22
Zhp_23
Zhp_24
Zhp_25
Zhp_06
Zhp_07
Zhp_13
Zhp_14
Zhp_15
Zhp_17
Zhp_20
Zhp_08
Zhp_09
Zhp_10
Zhp_11
Zhp_12
Zhp_16
Zhp_18
Zhp_19
Zhi_01
Zhi_02
Zhi_03
Zhi_04
Zhi_05
Zhi_06
Zhc_01
Zhc_02
Zhc_03
Zhc_04
Zhc_05
Zhc_06
Zhc_07
Zhc_08
Zhc_09
Zhc_10
Zhc_11
Zhc_12
Zhc_13
Zhc_14
Zhc_15
Zhi_07
Zhi_08
Zhi_09
Zhi_10
Zhi_11
Zhi_12
Zhi_13
Zhi_14
Zhi_15
Zhi_16
Zhi_17
Zhp_21
Zhl_01
Zhl_02
Zhpa_01
Zhpa_02
Zhpa_03
Zhpa_04
Zhpa_05
Zhpa_06
Zhpa_07
Zhpa_08
Zhpa_09
Zp_01
100
36
56
62
56
55
68
60
92
3
56
57
4
100
4
55
4
100
17
100
10
54
B
Zhp_01
Zhp_02
Zhp_03
Zhp_21
Zhp_04
Zhp_05
Zhp_06
Zhp_22
Zhp_23
Zhp_24
Zhp_25
Zhp_07
Zhp_13
Zhp_14
Zhp_15
Zhp_17
Zhp_20
Zhp_08
Zhp_09
Zhp_18
Zhp_10
Zhi_01
Zhi_02
Zhi_03
Zhi_04
Zhi_05
Zhi_06
Zhc_01
Zhc_02
Zhc_03
Zhc_04
Zhc_05
Zhc_06
Zhc_07
Zhc_08
Zhc_09
Zhc_10
Zhc_11
Zhc_12
Zhc_13
Zhc_14
Zhc_15
Zhi_07
Zhi_08
Zhi_09
Zhi_10
Zhi_11
Zhi_17
Zhi_12
Zhi_13
Zhi_14
Zhi_15
Zhi_16
Zhp_11
Zhp_12
Zhp_16
Zhp_19
Zhl_01
Zhl_02
Zhpa_01
Zhpa_02
Zhpa_03
Zhpa_04
Zhpa_05
Zhpa_06
Zhpa_07
Zhpa_08
Zhpa_09
Zp_01
1.0
1.0
.99
.97
1.0
5 changes

Figure 5

Appendix A. Listing of *Zapus hudsonius* museum specimens used in the present study including the museum identifier, the U.S.G. S. individual identifier, the abbreviated collection name (Table 1), the county and state each specimen was collected, and the subspecies designation. Tissues of specimens from three museums were utilized: Denver Museum of Nature and Science, Denver, Colorado; University of Kansas, Natural History Museum, Lawrence, Kansas, and the University of New Mexico, Museum of Southwestern Biology, Albuquerque, New Mexico.

Denver Museum of Nature and Science

Museum identifier	Individual identifier	Study collection	County	State	Designated subspecies
CHG9901	CHG9901	SOWY	Laramie	Wyoming	*preblei*
CTA9802	CTA9802	SOWY	Albany	Wyoming	*preblei*
CTN9901	CTN9901	SOWY	Albany	Wyoming	*preblei*
DOU9901	DOU9901	SOWY	Converse	Wyoming	*preblei*
DUC9901	DUC9901	SOWY	Albany	Wyoming	*preblei*
ELB9901	ELB9901	SOWY	Laramie	Wyoming	*preblei*
ELB9902	ELB9902	SOWY	Laramie	Wyoming	*preblei*
EPB9901	EPB9901	SOWY	Laramie	Wyoming	*preblei*
FRC9802	FRC9802	SOWY	Albany	Wyoming	*preblei*
LUM9901	LUM9901	SOWY	Platte	Wyoming	*preblei*
LUM9902	LUM9902	SOWY	Platte	Wyoming	*preblei*
LUM9903	LUM9903	SOWY	Platte	Wyoming	*preblei*
MCW9901	MCW9901	SOWY	Albany	Wyoming	*preblei*
MCW9902	MCW9902	SOWY	Albany	Wyoming	*preblei*
NLR9901	NLR9901	SOWY	Albany	Wyoming	*preblei*
NSB9901	NSB9901	SOWY	Laramie	Wyoming	*preblei*
RBC9901	RBC9901	SOWY	Platte	Wyoming	*preblei*
SBC9901	SBC9901	SOWY	Laramie	Wyoming	*preblei*
SSC9901	SSC9901	SOWY	Albany	Wyoming	*preblei*
SSC9902	SSC9902	SOWY	Albany	Wyoming	*preblei*
SSC9903	SSC9903	SOWY	Albany	Wyoming	*preblei*
SYB9901	SYB9901	SOWY	Platte	Wyoming	*preblei*
YCA9806	YCA9806	SOWY	Laramie	Wyoming	*preblei*
YCA9807	YCA9807	SOWY	Laramie	Wyoming	*preblei*
YCB9801	YCB9801	SOWY	Laramie	Wyoming	*preblei*
YCB9802	YCB9802	SOWY	Laramie	Wyoming	*preblei*
YCB9803	YCB9803	SOWY	Laramie	Wyoming	*preblei*
YCB9804	YCB9804	SOWY	Laramie	Wyoming	*preblei*
CER-9801	CER-9801	LCCO1	Larimer	Colorado	*preblei*
CER-9802	CER-9802	LCCO1	Larimer	Colorado	*preblei*
CER-9803	CER-9803	LCCO1	Larimer	Colorado	*preblei*
CER-9804	CER-9804	LCCO1	Larimer	Colorado	*preblei*
CER-9805	CER-9805	LCCO1	Larimer	Colorado	*preblei*
CER-9806	CER-9806	LCCO1	Larimer	Colorado	*preblei*

Denver Museum of Nature and Science

Museum identifier	Individual identifier	Study collection	County	State	Designated subspecies
HRK-9801	HRK-9801	LCCO1	Larimer	Colorado	*preblei*
HRK-9802	HRK-9802	LCCO1	Larimer	Colorado	*preblei*
HRK-9803	HRK-9803	LCCO1	Larimer	Colorado	*preblei*
HRK-9804	HRK-9804	LCCO1	Larimer	Colorado	*preblei*
CP9806, Random # 205	CP-205	LCCO1	Larimer	Colorado	*preblei*
MC-9801	MC-9801	LCCO1	Larimer	Colorado	*preblei*
MC-9803	MC-9803	LCCO1	Larimer	Colorado	*preblei*
NFP-9801	NFP-9801	LCCO1	Larimer	Colorado	*preblei*
NFP-9802	NFP-9802	LCCO1	Larimer	Colorado	*preblei*
BG-9801	BG-9801	LCCO2	Larimer	Colorado	*preblei*
BG-9802	BG-9802	LCCO2	Larimer	Colorado	*preblei*
PGC-9801	PGC-9801	LCCO2	Larimer	Colorado	*preblei*
SP9803, Random # 125	SP-125	LCCO2	Larimer	Colorado	*preblei*
SP9802, Random # 169	SP-169	LCCO2	Larimer	Colorado	*preblei*
SP9805, Random # 170	SP-170	LCCO2	Larimer	Colorado	*preblei*
SP9807, Random # 223	SP-223	LCCO2	Larimer	Colorado	*preblei*
SP9804, Random # 243	SP-243	LCCO2	Larimer	Colorado	*preblei*
SP9806, Random # 336	SP-336	LCCO2	Larimer	Colorado	*preblei*
SP9812, Random # 367	SP-367	LCCO2	Larimer	Colorado	*preblei*
SP9810, Random # 375	SP-375	LCCO2	Larimer	Colorado	*preblei*
SP9811, Random # 674	SP-674	LCCO2	Larimer	Colorado	*preblei*
SP9801, Random # 746	SP-746	LCCO2	Larimer	Colorado	*preblei*
SP9809, Random # 861	SP-861	LCCO2	Larimer	Colorado	*preblei*
YG-9801	YG-9801	LCCO2	Larimer	Colorado	*preblei*
YG-9803	YG-9803	LCCO2	Larimer	Colorado	*preblei*
MA98133, Random # 127	MAY-127	DCCO1	Douglas	Colorado	*preblei*
MA9809, Random # 165	MAY-165	DCCO1	Douglas	Colorado	*preblei*
MA98122, Random # 215	MAY-215	DCCO1	Douglas	Colorado	*preblei*
MA98201, Random # 229	MAY-229	DCCO1	Douglas	Colorado	*preblei*
MA98135, Random # 234	MAY-234	DCCO1	Douglas	Colorado	*preblei*
MA9801, Random # 254	MAY-254	DCCO1	Douglas	Colorado	*preblei*
MA98104, Random # 268	MAY-268	DCCO1	Douglas	Colorado	*preblei*
MA98108, Random # 281	MAY-281	DCCO1	Douglas	Colorado	*preblei*
MA98134, Random # 368	MAY-368	DCCO1	Douglas	Colorado	*preblei*
MA98132, Random # 374	MAY-374	DCCO1	Douglas	Colorado	*preblei*
MA9802, Random # 385	MAY-385	DCCO1	Douglas	Colorado	*preblei*

Denver Museum of Nature and Science

Museum identifier	Individual identifier	Study collection	County	State	Designated subspecies
MA98136, Random # 408	MAY-408	DCCO1	Douglas	Colorado	*preblei*
MA98120, Random # 416	MAY-416	DCCO1	Douglas	Colorado	*preblei*
MA98130, Random # 429	MAY-429	DCCO1	Douglas	Colorado	*preblei*
MA98138, Random # 452	MAY-452	DCCO1	Douglas	Colorado	*preblei*
MA98131, Random # 494	MAY-494	DCCO1	Douglas	Colorado	*preblei*
MA98124, Random # 497	MAY-497	DCCO1	Douglas	Colorado	*preblei*
MA98121, Random # 517	MAY-517	DCCO1	Douglas	Colorado	*preblei*
MA98102, Random # 532	MAY-532	DCCO1	Douglas	Colorado	*preblei*
MA98106, Random # 694	MAY-694	DCCO1	Douglas	Colorado	*preblei*
MA98204, Random # 706	MAY-706	DCCO1	Douglas	Colorado	*preblei*
MA98107, Random # 714	MAY-714	DCCO1	Douglas	Colorado	*preblei*
MA98123, Random # 748	MAY-748	DCCO1	Douglas	Colorado	*preblei*
MA9805, Random # 785	MAY-785	DCCO1	Douglas	Colorado	*preblei*
MA9810, Random # 798	MAY-798	DCCO1	Douglas	Colorado	*preblei*
MA98103, Random # 817	MAY-817	DCCO1	Douglas	Colorado	*preblei*
MA98203, Random # 822	MAY-822	DCCO1	Douglas	Colorado	*preblei*
MA9804, Random # 880	MAY-880	DCCO1	Douglas	Colorado	*preblei*
MA9806, Random # 933	MAY-933	DCCO1	Douglas	Colorado	*preblei*
MA98137, Random # 940	MAY-940	DCCO1	Douglas	Colorado	*preblei*
MA98202, Random # 946	MAY-946	DCCO1	Douglas	Colorado	*preblei*
MA9811, Random # 964	MAY-964	DCCO1	Douglas	Colorado	*preblei*
MA9813	MAY-9813	DCCO1	Douglas	Colorado	*preblei*
MA9814	MAY-9814	DCCO1	Douglas	Colorado	*preblei*
WH9801, Random # 715	WH-9801	DCCO2	Douglas	Colorado	*preblei*
WH9802, Random # 911	WH-9802	DCCO2	Douglas	Colorado	*preblei*
WH9803, Random # 629	WH-9803	DCCO2	Douglas	Colorado	*preblei*
WH9804, Random # 675	WH-9804	DCCO2	Douglas	Colorado	*preblei*
WH9805, Random # 849	WH-9805	DCCO2	Douglas	Colorado	*preblei*
WH9806, Random # 961	WH-9806	DCCO2	Douglas	Colorado	*preblei*
WH98100, Random # 573	WH-98100	DCCO2	Douglas	Colorado	*preblei*
WH98101, Random # 789	WH-98101	DCCO2	Douglas	Colorado	*preblei*
WH98102, Random # 672	WH-98102	DCCO2	Douglas	Colorado	*preblei*
WH98103, Random # 884	WH-98103	DCCO2	Douglas	Colorado	*preblei*
WH98104, Random # 719	WH-98104	DCCO2	Douglas	Colorado	*preblei*
WH98105, Random # 635	WH-98105	DCCO2	Douglas	Colorado	*preblei*
WH98106, Random # 603	WH-98106	DCCO2	Douglas	Colorado	*preblei*
WH98107, Random # 716	WH-98107	DCCO2	Douglas	Colorado	*preblei*

Denver Museum of Nature and Science

Museum identifier	Individual identifier	Study collection	County	State	Designated subspecies
WH98108, Random # 208	WH-98108	DCCO2	Douglas	Colorado	*preblei*
WH98109, Random # 958	WH-98109	DCCO2	Douglas	Colorado	*preblei*
WH98110, Random # 543	WH-98110	DCCO2	Douglas	Colorado	*preblei*
WH98120, Random # 768	WH-98120	DCCO2	Douglas	Colorado	*preblei*
WH98121, Random # 113	WH-98121	DCCO2	Douglas	Colorado	*preblei*
WH98300, Random # 140	WH-98300	DCCO2	Douglas	Colorado	*preblei*
WH98301, Random # 204	WH-98301	DCCO2	Douglas	Colorado	*preblei*
WH98302, Random # 814	WH-98302	DCCO2	Douglas	Colorado	*preblei*
WH98303, Random # 610	WH-98303	DCCO2	Douglas	Colorado	*preblei*
WH98304, Random # 190	WH-98304	DCCO2	Douglas	Colorado	*preblei*
WH98305, Random # 314	WH-98305	DCCO2	Douglas	Colorado	*preblei*
WH98306, Random # 612	WH-98306	DCCO2	Douglas	Colorado	*preblei*
WH98309, Random # 120	WH-98309	DCCO2	Douglas	Colorado	*preblei*
WH98311, Random # 860	WH-98311	DCCO2	Douglas	Colorado	*preblei*
WH98312, Random # 883	WH-98312	DCCO2	Douglas	Colorado	*preblei*
WH98313, Random # 924	WH-98313	DCCO2	Douglas	Colorado	*preblei*
MCA9801	Zhp-01	ECCO1	El Paso	Colorado	*preblei*
MCA9805	Zhp-03	ECCO1	El Paso	Colorado	*preblei*
MCA9806	Zhp-04	ECCO1	El Paso	Colorado	*preblei*
98KC01	Zhp-05	ECCO1	El Paso	Colorado	*preblei*
98PV01	Zhp-06	ECCO1	El Paso	Colorado	*preblei*
98SC01	Zhp-07	ECCO1	El Paso	Colorado	*preblei*
98SC02	Zhp-08	ECCO1	El Paso	Colorado	*preblei*
98SC03	Zhp-09	ECCO1	El Paso	Colorado	*preblei*
98SC04	Zhp-10	ECCO1	El Paso	Colorado	*preblei*
98SC05	Zhp-11	ECCO1	El Paso	Colorado	*preblei*
98SC07	Zhp-13	ECCO1	El Paso	Colorado	*preblei*
DMC01	Zhp-15	ECCO1	El Paso	Colorado	*preblei*
DMC02	Zhp-16	ECCO1	El Paso	Colorado	*preblei*
BS0198	Zhp-17	ECCO1	El Paso	Colorado	*preblei*
BS0298	Zhp-18	ECCO1	El Paso	Colorado	*preblei*
LR0198	Zhp-19	ECCO1	El Paso	Colorado	*preblei*
LR0298	Zhp-20	ECCO1	El Paso	Colorado	*preblei*
LR0398	Zhp-21	ECCO1	El Paso	Colorado	*preblei*
DC9802	Zhp-23	ECCO1	El Paso	Colorado	*preblei*
DC9803	Zhp-24	ECCO1	El Paso	Colorado	*preblei*
DC9804	Zhp-25	ECCO1	El Paso	Colorado	*preblei*
DC9805	Zhp-26	ECCO1	El Paso	Colorado	*preblei*
GCC0198	Zhp-27	ECCO1	El Paso	Colorado	*preblei*

University of Kansas Natural History Museum

Museum identifier	Individual identifier	Study collection	County	State	Designated subspecies
KU109972	Zhc-116	NA	Custer	South Dakota	*campestris*
KU109963	Zhc-124	NA	Lawrence	South Dakota	*campestris*
KU109978	Zhc-117	NA	Custer	South Dakota	*campestris*
KU109984	Zhc-118	NA	Custer	South Dakota	*campestris*
KU109985	Zhc-119	NA	Custer	South Dakota	*campestris*
KU110013	Zhc-115	NA	Custer	South Dakota	*campestris*
KU123592	Zhc-097	NA	Carter	Montana	*campestris*
KU123597	Zhc-095	NA	Carter	Montana	*campestris*
KU112661	Zhc-126	NA	Lawrence	South Dakota	*campestris*
KU112663	Zhc-127	NA	Lawrence	South Dakota	*campestris*
KU115700	Zhi-033	NA	Burleigh	North Dakota	*intermedius*
KU115730	Zhi-037	NA	Walworth	South Dakota	*intermedius*
KU112665	Zhc-128	NA	Lawrence	South Dakota	*campestris*
KU153706	Zhpa-050	NA	Leavenworth	Kansas	*pallidus*
KU110033	Zhpa-051	NA	Bennett	South Dakota	*pallidus*

Museum of Southwestern Biology, University of New Mexico

none	MSB-41518	MCMN	Morrison	Minnesota	*intermedius*
84916	MSB-41532	MCMN	Morrison	Minnesota	*intermedius*
84917	MSB-41533	MCMN	Morrison	Minnesota	*intermedius*
90860	MSB-80766	MCMN	Morrison	Minnesota	*intermedius*
90861	MSB-80767	MCMN	Morrison	Minnesota	*intermedius*
90862	MSB-80768	MCMN	Morrison	Minnesota	*intermedius*
90863	MSB-80769	MCMN	Morrison	Minnesota	*intermedius*
90864	MSB-80770	MCMN	Morrison	Minnesota	*intermedius*
90865	MSB-80771	MCMN	Morrison	Minnesota	*intermedius*
90866	MSB-80772	MCMN	Morrison	Minnesota	*intermedius*
90867	MSB-80773	MCMN	Morrison	Minnesota	*intermedius*
90868	MSB-80774	MCMN	Morrison	Minnesota	*intermedius*
90869	MSB-80778	MCMN	Morrison	Minnesota	*intermedius*
90870	MSB-80779	MCMN	Morrison	Minnesota	*intermedius*
90871	MSB-80780	MCMN	Morrison	Minnesota	*intermedius*
90872	MSB-80781	MCMN	Morrison	Minnesota	*intermedius*
90873	MSB-80782	MCMN	Morrison	Minnesota	*intermedius*
90874	MSB-80783	MCMN	Morrison	Minnesota	*intermedius*
90875	MSB-80784	MCMN	Morrison	Minnesota	*intermedius*
90876	MSB-80785	MCMN	Morrison	Minnesota	*intermedius*
90943	MSB-80786	MCMN	Morrison	Minnesota	*intermedius*

Museum of Southwestern Biology, University of New Mexico

Museum Identifier	Individual identifier	Study collection	County	State	Designated subspecies
56982	MSB-3826	SCNM	Sandoval	New Mexico	*luteus*
56993	MSB-3827	SCNM	Sandoval	New Mexico	*luteus*
56994	MSB-3828	SCNM	Sandoval	New Mexico	*luteus*
56996	MSB-3829	SCNM	Sandoval	New Mexico	*luteus*
56984	MSB-3830	SCNM	Sandoval	New Mexico	*luteus*
56995	MSB-3831	SCNM	Sandoval	New Mexico	*luteus*
56979	MSB-3832	SCNM	Sandoval	New Mexico	*luteus*
56983	MSB-3833	SCNM	Sandoval	New Mexico	*luteus*
56997	MSB-3834	SCNM	Sandoval	New Mexico	*luteus*
56980	MSB-3835	SCNM	Sandoval	New Mexico	*luteus*
None	MSB-3836	SCNM	Sandoval	New Mexico	*luteus*
56981	MSB-3837	SCNM	Sandoval	New Mexico	*luteus*
56985	MSB-3838	SCNM	Sandoval	New Mexico	*luteus*
56990	MSB-3839	SCNM	Sandoval	New Mexico	*luteus*
56991	MSB-3840	SCNM	Sandoval	New Mexico	*luteus*
56992	MSB-3841	SCNM	Sandoval	New Mexico	*luteus*
56986	MSB-3842	SCNM	Sandoval	New Mexico	*luteus*
56989	MSB-3843	SCNM	Sandoval	New Mexico	*luteus*
56987	MSB-3844	SCNM	Sandoval	New Mexico	*luteus*
56988	MSB-3845	SCNM	Sandoval	New Mexico	*luteus*

Appendix B. Validation of Ramey et al. (2005) control region sequence data.

Ramey et al. (2005) (REA) reported 10 haplotypes shared between and among *Zapus hudsonius* subspecies. REA used decades-old, dried museum skins as a principal source of genetic material for selected collections used in the phylogeographic comparison. Given the discrepancy between the results of REA and this study (King et al.; KEA) regarding haplotype sharing, we were concerned that this methodological decision may have introduced unnecessary ambiguity to the findings. For example, REA reported the presence of *Z. h. preblei* haplotypes in DNA extracted from five dried museum skins of Z. *h. campestris* collected from Custer County, SD. The authors suggested this finding indicated recent gene flow and alluded to the presence of these haplotypes as a critical element in the decision to recommend synonymy of these subspecies. KEA sampled 31 *Z. h. campestris* recently from the same site in Custer County, SD used by REA, along with 30 additional specimens from neighboring Crook County, WY. All individuals were subjected to mtDNA CR and cyt *b* sequence analysis. All 61 individuals were determined to posses *Z. h. campestris*-specific mtDNA haplotypes. Moreover, the same conclusion was reached with the microsatellite loci, as no *Z. h. campestris* individual from either of these collections was assigned to *Z. h. preblei*. Given the prominent role the haplotypes obtained for the five museum skins from Custer County, SD and two additional specimens from Carter County, MT have played in the conclusions drawn by REA and the absence of sharing observed in a large sample by this study, we felt compelled to validate the previous findings.

Here we present the results of attempts to validate the mitochondrial DNA control region (CR) sequences reported by REA for seven *Z. h. campestris* dried museum skin specimens obtained from the Kansas University Natural History Museum (KUNHM) and reported to possess four *Z. h. preblei* haplotypes. In addition, eight other specimens were obtained to assist in validation of four additional shared haplotypes. An alignment of the CR sequences for the 15 KUNHM specimens with the sequences generated for the five *Z. hudsonius* subspecies surveyed in this study (KEA) was performed. From this alignment, the identity of the 15 specimens was readily apparent. The follow description and associated table (Table B.1) are provided to assist the reader in assessing the validity of the mtDNA sequences reported by and conclusions made by REA.

Haplotype C/P1

Eleven specimens were listed by REA as having haplotype C/P1: nine *Z. h. preblei* (DMNH10405, DMNH10258, DMNH10270, DMNH10404, DMNH10406, DMNH10407, DMNH9568, PIONEER9568, PIONEER9B89) and two *Z. h. campestris* (KU109984, KU109985). When the REA GenBank data is compared with this study (KEA), all individuals were found to match the haplotype ZhpCR_A. Tissue samples from two of the specimens (KU109984 and KU109985) were obtained from the KUMNH and sequenced to verify the results of REA. The sequences obtained for the two specimens were not the same as those reported by REA (haplotype C/P1), both specimens had haplotype ZhcCR_D, a haplotype only observed in *Z. h. campestris* by KEA.

Haplotype C/P2

Fourteen specimens were reported by REA to have haplotype C/P2: 12 *Z. h. preblei* (DMNH9579, DMNH9313, DMNH1315, DMNH10380, DMNH9565, DMNH9563, DMNH9566, DMNH9573, DMNH9572, DMNH9571, DMNH9574, DMNH10607) and two *Z. h. campestris* (KU109978, KU123592). When the REA GenBank data is compared to that of KEA, all individuals were found to match the haplotype ZhpCR_B. Tissue samples from two of the specimens (KU109978 and KU123592) were obtained from the KUMNH and sequenced. The sequences obtained for the two specimens were not the same as those reported by REA (haplotype C/P2); both specimens had haplotype ZhcCR_D, a haplotype only observed in *Z. h. campestris* by KEA.

Haplotype C/P3

Twenty six specimens were listed as having haplotype C/P3 by REA: twenty four *Z. h. preblei*

and two *Z. h. campestris* (KU110013, KU123597). When the REA GenBank data is compared to that of KEA, all of these individuals were found to match the haplotype ZhpCR_C. Tissue samples from the two *Z. h. campestris* specimens (KU110013 and KU123597) were obtained from the KUMNH and sequenced. The sequences obtained for the two specimens were not the same as those reported by REA (haplotype C/P3); both specimens had haplotype ZhcCR_D, a haplotype only observed in *Z. h. campestris*.

Haplotype C/P4

Ten specimens are listed as having haplotype C/P4 by REA: nine *Z. h. preblei* and one *Z. h. campestris* (KU109972). When the REA GenBank data is compared to that of KEA, all of these individuals were found to match the haplotype ZhpCR_D. A tissue sample from the *Z. h. campestris* specimen (KU109972) was obtained from the KUMNH and sequenced to verify the results of REA. The sequence obtained for the specimen was not the same as that reported by REA (haplotype C/P4); the specimen had haplotype ZhcCR_A, a haplotype only observed in *Z. h. campestris*.

Table B.1 presents the results of the comparison among the seven *Z. hudsonius campestris* specimens. All seven individuals were observed to posses *Z. h. campestris* haplotypes consistent with the KUMNH identification. In addition to the CR comparison, microsatellite DNA was surveyed in these specimens. Multilocus genotypes of the seven individuals was subjected to a maximum likelihood assignment test to determine whether the specimen was most closely related to the baseline collections of *Z. h. campestris* or *Z. h. preblei* surveyed by KEA following the methods presented in the text. The results of the assignment testing indicated that all seven individuals were on average 2.0 times more likely to be *Z. h. campestris* than *Z. h. preblei*.

Haplotype L/PAL/C1

Haplotype L/PAL/C1 were reported to be shared among three subspecies of *Z. hudsonius* by REA. Eleven specimens were listed as having haplotype L/PAL/C1 by REA: seven *Z. h. luteus* (MSB58370, MSB56980, MSB56986, MSB56987, MSB56991, MSB56993, MSB62096, and NK856), two *Z. h. campestris* (KU112665, KU109963), and one *Z. h. pallidus* (KU110033). Of these 11 specimens only one, NK856, has consistent numbering between Appendix 2 of REA and data accessioned in GenBank that allows for direct comparison. When the REA GenBank data was compared to that of KEA, NK856 was found to match the haplotype ZhlCR_A. Ten other specimens were also found to match this haplotype: six *Z. h. luteus* (MSB20, MSB21, MSB23, MSB24, MSB25, MSB26, MSB27), two *Z. h. campestris* (KU25, KU28), one *Z. h. pallidus* (KU53). Given this evidence, we believe haplotype L/PAL/C1 reported by REA is the same as haplotype ZhlCR_A of KEA. Tissue samples from three of the specimens (KU112665, KU109963 and KU110033) were obtained from KUMNH and sequenced to verify the results of REA. The sequences obtained were not the same as those reported by REA (haplotype L/PAL/C1): specimen KU112665 (*Z. h. campestris*) had haplotype KU45, specimen KU109963 had haplotype ZhcCR_A and specimen KU110033 was found to have a unique haplotype not previously reported by REA or KEA.

Haplotype L/PAL/C2

Haplotype L/PAL/C2 was reported to be shared among three subspecies of *Z. hudsonius* by REA. Nine specimens are listed as having haplotype L/PAL/C2 by REA: six *Z. h. luteus* (MSB86344, MSB91627, MSB91675, NK1584, DMNH8635, and DMNH8633), two *Z. h. campestris* (KU41451, KU112661), and one *Z. h. pallidus* (KU153706). Of these nine specimens only three (NK1584, DMNH8635, DMNH8633) have consistent numbering between

Appendix 2 of REA and data accessioned in GenBank. When the REA GenBank data is compared to that of KEA, NK1584, DMNH8635 and DMNH8633 were not found to match any haplotypes, but did match a haplotype from REA. This haplotype is referred to as KU45; based on a specimen code in GenBank (REA accession) for an individual of *Z. h. pallidus* with this haplotype. Eight other specimens from GenBank were also found to match this haplotype: six *Z. h. luteus* (MSB8, MSB7, MSB5, NK1584, DMNH8635, and DMNH8633) and two *Z. h. campestris* (KU27, KU1). Given this evidence we believe haplotype L/PAL/C2 of REA is identical to haplotype KU45 of the REA GenBank accession. Tissue samples from two specimens (KU153706 and KU112661) were obtained from the KUMNH. Specimen KU153706 was found to have haplotype KU45 as reported by REA. However specimen KU112661 (*Z. h. campestris*) had haplotype ZhcCR_D, which is not the same haplotype as that reported by REA (haplotype KU45).

Haplotype C9/INT-VII

Haplotype C9/INT-VII was reported reported by REA to be shared by *Z. h. campestris* and *Z. h. intermedius*. Two specimens were listed by REA as having C9/INT-VII, KU112663 (*Z. h. campestris*) and KU115730 (*Z. h. intermedius*). Only KU115730 had consistent numbering between Appendix 2 of REA and data accessioned in GenBank. When the REA GenBank data was compared to that of KEA, KU115730 was not found to match any haplotype. One other REA sequence, KU26 (*Z. h. campestris*), had the C9/INT-VII haplotype, leading us to believe that KU26 of the GenBank alignment is KU115730. Tissue samples from the two specimens (KU112663 and KU115730) were obtained from the KUMNH. Both specimens, KU112663 and KU115730 were found by KEA to have haplotype ZhcCR_D, not C9/INT-VII as reported by REA.

Haplotype C5/INT-XIII

REA report haplotype C5/INT-XIII as being shared between *Z. h. campestris* and *Z. h. intermedius*. Specimens KU87040, KU83557, and KU87042 were collected from Harding County, South Dakota and cataloged by the museum as *Z. h. campestris*. KU115895, KU115896, and KU115897 were collected from the same general location in Harding County, South Dakota but were identified as *Z. h. intermedius*. The respective collectors of both sets of samples collaborated in a publication (Anderson and Jones 1971) in which they named all *Z. hudsonius* from this location as *Z. h. campestris*. This haplotype is not shared between the two subspecies.

Table B.1 provides the results of the validation for the 8 additional museum skins. In total, we investigated eight of 10 haplotypes reported by REA to be shared between or among selected subspecies of *Z. hudsonius*. Of the 15 specimens analyzed, 13 were found to have haplotypes distinctly different from those reported by REA. We believe these findings have identified the presence of a systemic error in the CR data reported by REA. Moreover, the inability to directly link the sequences deposited in GenBank to the specimens listed in Appendix 2 of REA and the associated locality data renders the data useless. Possible reasons for the different sequences reported by REA are contamination, mislabeling of samples, or other procedural incongruity. Given that all the specimens re-evaluated in the current study were found by REA to have haplotypes that were shared among subspecies, the disagreement in DNA sequences reported for these individuals calls into question all of the results of REA based on the mitochondrial DNA genome and prevents analysis of the combined data.

Table B.1. Results of 15 museum specimens (Kansas University Museum of Natural History; KUMNH) sequenced to verify selected results from Ramey et al. 2005 (REA). Information provided includes: museum catalog number, collection information, designated subspecies, haplotype reported by REA, corresponding haplotype in King et al. (this manuscript; KEA), the haplotype observed after re-analysis, and the results of an assignment to subspecies based on microsatellite DNA analysis for the seven specimens reported to possess *Z. h. preblei* haplotypes by REA. The number of loci used () and the ratio of likelihood of assignment to the designated subspecies divided by the score for the REA haplotype designation (i.e., *Z. h. preblei*)] are provided.

Catalog number	Collection county, state	Designated subspecies	REA haplotype	Coresponding KEA haplotype	Observed Haplotype[d]	Assignment to subspecies
KU123597[a]	Carter Co., MT	*campestris*	C/P1	ZhpCR_A	ZhcCR_D	*campestris* (8) [1.7]
KU123592[a]	Carter Co., MT	*campestris*	C/P1	ZhpCR_A	ZhcCR_D	*campestris* (13) [1.7]
KU110013[a]	Custer Co., SD	*campestris*	C/P2	ZhpCR_B	ZhcCR_D	*campestris* (12) [2.5}
KU109984[a]	Custer Co., SD	*campestris*	C/P2	ZhpCR_B	ZhcCR_D	*campestris* (11) [2.3
KU109985[a]	Custer Co., SD	*campestris*	C/P3	ZhpCR_C	ZhcCR_D	*campestris* (9) [1.1]
KU109978[a]	Custer Co., SD	*campestris*	C/P3	ZhpCR_C	ZhcCR_D	*campestris* (15) [2.2]
KU109972[a]	Custer Co., SD	*campestris*	C/P4	ZhpCR_D	ZhcCR_A	*campestris* (9) [2.8]
KU112665	Lawrence Co., SD	*campestris*	L/PAL/C1	ZhlCR_A	KU45	
KU109963	Lawrence Co., SD	*campestris*	L/PAL/C1	ZhlCR_A	ZhcCR_A	
KU110033	Bennett Co., SD	*pallidus*	L/PAL/C1	ZhlCR_A	New[e]	
KU112661	Lawrence Co., SD	*campestris*	L/PAL/C2	none[b]	ZhcCR_D	
KU153706	Leavenworth Co., KS	*pallidus*	L/PAL/C2	none[b]		
KU112663	Lawrence Co., SD	*campestris*	C9/INT-VII	none[c]	ZhcCR_D	
KU115730	Walworth Co., SD	*intermedius*	C9/INT-VII	none[c]	ZhcCR_D	
KU115700	Burleigh Co., ND	*intermedius*	C8/10/INT-VI	ZhcCR_D		

[a]*Z. h. campestris* specimens from KUMNH identified by REA as having *Z. h. preblei* haplotypes.
[b]Haplotype not observed by King et al. (this study); designated as KU45 by REA in GenBank.
[c]Haplotype not observed by King et al.; designated as KU26 by REA in GenBank.
dThirteen of 15 KUMNH specimens were observed (this study) to have different haplotypes than those reported by REA.
[e]This haplotype is unique as it was not observed by either study

Appendix C. Microsatellite DNA markers, the total number of alleles observed, and the range of amplified products in the survey of 348 *Zapus hudsonius*, and associated references.

Microsatellite locus	Alleles observed	Size range	Reference
Z.7	30	156-179	Ramey et al. 2005
Z.20	20	103-147	Ramey et al. 2005
Z.26	16	138-174	Ramey et al. 2005
Z.48	16	173-203	Ramey et al. 2005
Z.52	11	154-176	Ramey et al. 2005
*Zt*ri2	12	91-135	Vignieri et al. 2003
Ztri17	14	149-195	Vignieri et al. 2003
*Ztri*19	8	174-206	Vignieri et al. 2003
*Ztri*24	13	151-199	Vignieri et al. 2003
ZhuC3	16	204-264	King et al. 2006
*Zhu*C6	10	100-144	King et al. 2006
*Zhu*C12	8	96-124	King et al. 2006
*Zhu*C104	11	222-254	King et al. 2006
*Zhu*C119	17	207-263	King et al. 2006
*Zhu*C120	7	145-169	King et al. 2006
*Zhu*C129	10	200-236	King et al. 2006
*Zhu*C130	7	258-286	King et al. 2006
*Zhu*D107	13	213-261	King et al. 2006
*Zhu*D108	10	138-176	King et al. 2006
*Zhu*D109	13	133-177	King et al. 2006
*Zhu*D122	18	201-285	King et al. 2006

Appendix D. *Zapus hudsonius* mitochondrial DNA control region haplotype (372 base pairs) counts by subspecies and collection.

	Z. hudsonius preblei							*Z. h. intermedius*		*Z. h. campestris*		*Z. h. pallidus*		*Z. h. luteus*
Haplotype	SOWY	LCCO1	LCCO2	DCCO1	DCCO2	ECCO1	ECCO2	BRCSD	MCMN	CCWY	CCSD	BCSD	KBCNE	SCNM
N	28	12	16	32	26	19	26	26	21	29	31	16	31	19
*Zhp*CR_A	28	12	12											
*Zhp*CR_B				6	18	9	11							
*Zhp*CR_C			4	26	5									
*Zhp*CR_D					3	10	15							
*Zhi*CR_A								8	4					
*Zhi*CR_B								5						
*Zhi*CR_C								13						
*Zhi*CR_D									8					
*Zhi*CR_E									3					
*Zhi*CR_F									2					
*Zhi*CR_G									1					
*Zhi*CR_H									2					
*Zhi*CR_I									1					
*Zhc*CR_A										25	20			
*Zhc*CR_B										3				
*Zhc*CR_C										1				
*Zhc*CR_D											8			
*Zhc*CR_E											3			
*Zhpa*CR_A												14	10	
*Zhpa*CR_B												2		
*Zhpa*CR_C													10	
*Zhpa*CR_D													7	
*Zhpa*CR_E													4	
*Zhl*CR_A														17
*Zhl*CR_B														2

Appendix E. *Zapus hudsonius* mitochondrial DNA cytochrome *b* region haplotype (1006 base pairs) counts by subspecies and collection.

	Zapus hudsonius preblei							*Z. h. intermedius*		*Z. h. campestris*		*Z. h. pallidus*		*Z. h. luteus*
Haplotype	SOWY	LCCO1	LCCO2	DCCO1	DCCO2	ECCO1	ECCO2	BRCSD	MCMN	CCWY	CCSD	BCSD	KBCNE	SCNM
N	28	14	15	31	14	19	25	26	21	30	30	16	32	19
*Zhp*CB_A		1	5											
*Zhp*CB_B		1												
*Zhp*CB_C		1												
*Zhp*CB_D		1												
*Zhp*CB_E		1												
*Zhp*CB_F		9	5											
*Zhp*CB_G	9		4											
*Zhp*CB_H			1											
*Zhp*CB_I				7	11	14	24							
*Zhp*CB_J				23										
*Zhp*CB_K				1		1								
*Zhp*CB_L					2									
*Zhp*CB_M					1									
*Zhp*CB_N						1								
*Zhp*CB_O						1								
*Zhp*CB_P						2								
*Zhp*CB_Q							1							
*Zhp*CB_R	5													
*Zhp*CB_S	1													
*Zhp*CB_T	1													
*Zhp*CB_U	12													
*Zhi*CB_A									3					
*Zhi*CB_B									1					
*Zhi*CB_C								5	4					
*Zhi*CB_D									1					
*Zhi*CB_E									3					
*Zhi*CB_F									1					
*Zhi*CB_G									5					

*Zhi*CB_H		1					
*Zhi*CB_I		2					
*Zhi*CB_J	13						
*Zhi*CB_K	6						
*Zhi*CB_L	1						
*Zhi*CB_M	1						
*Zhc*CB_A			1				
*Zhc*CB_B			22	19			
*Zhc*CB_C			1				
*Zhc*CB_D			1				
*Zhc*CB_E			2				
*Zhc*CB_F			1				
*Zhc*CB_G			1				
*Zhc*CB_H				2			
*Zhc*CB_I			1				
*Zhc*CB_J				5			
*Zhc*CB_K				1			
*Zhc*CB_L				2			
*Zhc*CB_M				1			
*Zhpa*CB_A					4		
*Zhpa*CB_B					9		
*Zhpa*CB_C					2	11	
*Zhpa*CB_D					1		
*Zhpa*CB_E						10	
*Zhpa*CB_F						6	
*Zhpa*CB_G						4	
*Zhpa*CB_H						1	
*Zhl*CB_A							19