Comprehensive analysis of molecular phylogeographic structure among the meadow jumping mice (*Zapus hudsonius*) reveals evolutionarily distinct subspecies

A Report submitted to the U.S. Fish and Wildlife Service

January 27, 2006

Tim L. King[1,3], John F. Switzer[1], Cheryl L. Morrison[1],
Michael S. Eackles[1], Colleen C. Young[1], Barbara 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]US Geological Survey - Biological Resources Division
Fort Collins Science Center
2150 Centre Avenue, Building C
Fort Collins, CO 80526
USA

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

Pre-publication Copy

ABSTRACT

Zapus hudsonius preblei is one of 12 recognized subspecies of meadow jumping mice found throughout North America. Recent morphometric, phylogeographic, and phylogenetic comparisons among *Z. h. preblei* and neighboring conspecifics have called into question the taxonomic status of selected subspecies. We present additional analyses of the phylogeographic structure within *Z. hudsonius* that contradict previous findings and clearly demonstrate that each subspecies considered is genetically differentiated from its neighboring subspecies. In a more comprehensive analysis of genetic data, we compared a larger collection of samples over a greater representation of both the mitochondrial and nuclear genomes than was previously investigated and placed particular emphasis on the relatedness of *Z. h. preblei* to each neighboring subspecies. A survey of 21 microsatellite DNA loci and 1380 base pairs from two mitochondrial DNA regions (control region and cytochrome *b*) revealed that each *Z. hudsonius* subspecies is genetically distinct to varying degrees, and thus our 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 was supported by highly significant findings for nearly every statistical comparison made, regardless of the genome or the taxa under consideration. We observed strong structuring of nuclear multilocus genotypes and subspecies-specific mtDNA haplotypes that corresponded directly with the disjunct distributions of the subspecies investigated. Given the level of correspondence between the observed genetic population structure and the 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*. The findings of this study differed in nearly every comparison with conclusions drawn from the only other molecular genetics data published on the taxonomic status of *Z. hudsonius* subspecies. The disparities between the previous and present study include differences in sampling design, tissues chosen for analysis, number of molecular characters sequenced and microsatellite fragments surveyed, methods used to portray genealogical relationships, test statistics applied, and criteria used to determine uniqueness.

Pre-publication Copy

INTRODUCTION

Preble's meadow jumping mouse (*Zapus hudsonius preblei*) is widely believed to be a remnant of past proglaciation of the Colorado Piedmont. The distribution of *Z. h. preblei* is confined to the riparian systems where moisture is most plentiful (Jones et al. 1983). At present, the range of the subspecies 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).

Zapus hudsonius preblei is one of 12 recognized subspecies of meadow jumping mice found throughout North America (Hafner et al. 1981). Subspecies recognition within *Z. hudsonius* has been based primarily on geographic disjunction and morphological variation (Krutzsch 1954, Hafner et al. 1981). Subsequently, morphological comparisons among species of the genus *Zapus* have failed to reliably identify diagnostic characteristics (Jones 1981, but see Conner and Shenk 2003). These morphologically similar subspecies present significant challenges to conservation biologists because of taxonomic uncertainty (Preble 1899, Krutzsch 1954, Ramey et al. 2005). A morphology-based taxonomy may not reveal true phylogenetic relationships as the rate of evolutionary change can vary among lineages and similar environmental influences may cause convergence (Grant 1987). Indeed, a taxonomy based solely on phenotypic characteristics 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 the range of *Z. hudsonius*. Inherent in such an approach is the identification and characterization of the associated migration, colonization, and extinction processes among populations of these putative subspecies (Avise 2004). Molecular markers, with a clear heritable genetic basis and the number of characters limited only by genome size (Moritz and Hillis 1996), provide insight into these processes and can be used to reveal genetic discontinuities and distinctiveness among or between taxa with slight morphological

Pre-publication Copy

differentiation (Clark et al. 2000). It has been repeatedly demonstrated that patterns of gene exchange, the extent of genealogical relationships, and accurate reflections of true evolutionary relationships (i.e., phylogeny) can be revealed through use of the appropriate type and number of molecular genetic markers (Moritz and Hillis 1996, Avise 2004).

Recent morphometric, phylogeographic, phylogenetic, and ecologic exchangeability comparisons among *Zapus hudsonius preblei* and neighboring con-specifics have called into question the taxonomic status of selected subspecies (Ramey et al. 2005). The multidisciplinary study utilized a hypothesis testing approach to determine uniqueness of subspecies, which included analyses of cranial morphometrics and contemporary genetic techniques. Results of multivariate morphometric analyses on general cranial measurements failed to support the original description of *Z. h. preblei* as a distinct subspecies. The genetic components of the study compared haplotypes generated from the mitochondrial (mt) DNA control region and multilocus genotypes at nuclear microsatellite loci utilizing DNA obtained from museum skins, frozen liver, or ethanol-preserved ear punch tissues. Ramey et al. (2005) concluded that recent gene exchange and low levels of genetic structure among subspecies supported synonymization of *Z. h. preblei*, *Z. h. campestris* (Bear Lodge meadow jumping mouse), and *Z. h. intermedius* (meadow jumping mouse). The study by Ramey et al. (2005) constitutes the lone published molecular population genetic analysis of *Z. hudsonius*, with important implications for the evolution, ecology, and conservation status of *Z. h. preblei*. The proposed synonymy of these subspecies has prompted a subsequent reevaluation of the status of *Z. h. preblei* under the ESA (US Fish and Wildlife Service 2005).

Studies of phylogeographic relationships among intraspecific taxa often exact a more rigorous study design than that required for interspecific comparisons (Avise 2004). Further, phylogeographic studies can be initially misled by dependence on tissues yielding insufficient quality (Kirchman et al. 2001) and consistency of DNA (Steinberg 1999), inadequate portrayals 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, Mitchell et al. 2000). Ramey et al. (2005) 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

Pre-publication Copy

represented nuclear DNA variation with the findings of five microsatellite DNA loci. Ramey et al.'s (2005) critical test of uniqueness for *Z. h. preblei* and related taxa was that greater molecular variance was demonstrated between subspecies than within, a test criterion that, to our knowledge, is not generally accepted. In light of the criteria applied to a relatively small fragment of mtDNA and minimal number of microsatellite loci, the methods applied and conclusions drawn by Ramey et al. (2005) warrant independent verification.

Here, we report on additional results concerning the phylogeographic structure and evolutionary distinctiveness of subspecies within *Zapus hudsonius*, 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 374 bp of the mtDNA control region and approximately 1000 bp of the mtDNA cytochrome *b* (cyt *b*) region, combined with the fragment analysis of a four-fold greater number of nuclear microsatellite DNA loci. We tested the null hypothesis that collections of *Z. hudsonius* distributed among *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 by the spatial distribution of mtDNA haplotypes and microsatellite DNA allele frequencies. Due to the taxonomic revision proposed by Ramey et al. (2005), importance was placed on comparisons among *Z. h. preblei, Z. h. campestris*, and *Z. h. intermedius*. In the present study we accept 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, Moritz 1994a).

METHODS

Tissue samples - 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 after emergence from hibernation in summer 2005 (Table 1). Archived tissue samples were obtained from *Z. h. prebl*ei (ear punch), *Z. h. intermedius* (frozen liver; one collection from Morrison County, Minnesota) and *Z. h. luteus* (frozen liver). All tissues were forwarded directly from the collector or museum to the U.S.G.S.-Leetown Science Center,

Pre-publication Copy

152 Kearneysville, WV for DNA extraction, fragment analysis, and DNA sequencing.
153
154 DNA extractions - DNA aliquots were obtained from ear punches (in 95% ethanol), frozen (-80°
155 C) liver, or blood tissue (on FTA cards; Whatman Inc., Clifton, NJ, USA) using the PUREGENE
156 DNA extraction kit (Gentra Systems, Inc., Minneapolis, MN) and resuspended in TE (10 mM
157 Tris-HCl, pH 8.0, 1 mM EDTA).
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 A).
161 The three sets of markers were: 1) Z.7, Z.20, Z.26, Z.48, and Z.52 (Ramey et al. 2005); 2) *Ztri*2,
162 *Ztri*17, *Ztri*19, and *Ztri*24 (isolated from *Z. trinotatus*; Vignieri 2003); and 3) *Zhu*C3, *Zhu*C6,
163 *Zhu*C12, *Zhu*C104, *Zhu*C119, *Zhu*C120, *Zhu*C129, *Zhu*C130, *Zhu*D107, *Zhu*D108, *Zhu*D109,
164 and *Zhu*D122 (King et al. in press). Attempts to obtain tissue or DNA samples for
165 standardization with the Ramey et al. (2005) microsatellite DNA allele scoring were
166 unsuccessful. Microsatellite DNA amplification reactions were conducted under laminar flow
167 conditions and consisted of 100-200 ng of genomic DNA, 1X PCR buffer (10 mM Tris-HCl, pH
168 8.3, 50 mM KCl), 2 mM $MgCl_2$, 0.25 mM dNTPs, 0.5 μM forward (fluorescently labeled) and
169 reverse primer, and 0.1 U *Taq* DNA polymerase (Promega, Madison, WI, USA) in a total
170 volume of 10 μl. Amplifications were carried out on either a PTC-200 or PTC-225 thermal
171 cycler (MJ Research, Watertown, MA, USA) using the following procedure: initial denaturing at
172 94°C for 2 min, 35 cycles of 94°C for 40 sec, 56°C for 40 sec, 72 °C for 1 min, and a final
173 extension at 72 °C for 10 min. Fragment analysis was performed on an Applied Biosystems
174 (Foster City, CA, USA) ABI 3100 Genetic Analyzer, under the conditions described in King et
175 al. (2001). Genescan™ 2.1 Analysis software and Genotyper™ 2.0 Fragment Analysis software
176 (Applied Biosystems) were used to score, bin, and output allelic (and genotypic) data.
177
178 Mitochondrial DNA Amplification and Sequencing - Two regions of the mitochondrial genome
179 were amplified under laminar flow conditions and sequenced for this study. A region of the non-
180 coding control region (CR) was amplified by double-stranded PCR using the primers L15926
181 (5'-TCA AAG CTT ACA CCA GTC TTG TAA ACC - 3') and H16498 (5'- CCT GAA CTA
182 GGA ACC AGA TG -3') (Kocher et al. 1989; Shields and Kocher 1991) for all *Z. hudsonius*.

Pre-publication Copy

183 Specimens of *Z. princeps* (*N*=7) were amplified using the primers L15398 (5'- ATC AGC ACC
184 CAA AGC TGA TAT TC - 3') (Ramey et al. 2005) and H16498. Polymerase chain reactions
185 consisted of an initial denaturation at 94 ^{o}C for 3 min, followed by 35 cycles of 1 min at 94 ^{o}C
186 (denaturation), 1 min at 48 ^{o}C (annealing), and 2 min at 72 ^{o}C (extension) followed by a final
187 extension period of 5 min at 72 ^{o}C. The mitochondrial cytochrome *b* (cyt *b*) gene of each
188 specimen was amplified by double-stranded PCR using two primers designed for this study
189 L14398A (5'- CCA ATG ACA TGA AAA ATC ATC G - 3') and H15634A (5'- TGG TTT
190 ACA AGA CCA GAG TAA - 3'). Polymerase chain reactions consisted of an initial
191 denaturation at 94 ^{o}C for 3 min, followed by 35 cycles of 1 min at 94 ^{o}C (denaturation), 1 min at
192 55 ^{o}C (annealing), and 2 min at 72 ^{o}C (extension) followed by a final extension period of 10 min
193 at 72 ^{o}C. Polymerase chain reactions consisted of 25 µl total volume, containing 2.5 µl of
194 $MgCl_2$-free buffer, 2.5 µl of $MgCl_2$ solution, 0.5 µl of dNTPs (2.5 mM each), 1.25 µl of each
195 primer (10 µM), one unit *Taq* polymerase, three µl of template (ca. 50-100 ng double-stranded
196 DNA), and 13.75 µl of sterile water. Negative controls, which did not include template DNA,
197 were set up alongside PCR reactions as checks for contamination of PCR reagents. Polymerase
198 chain reaction products were purified with exonuclease I and shrimp alkaline phosphatase. The
199 CR and cyt *b* PCR products were then sequenced using ABI BigDye v3.1 terminator cycle
200 sequencing chemistry, with sequences read by an ABI 3100 Genetic Analyzer (Applied
201 Biosystems). The control region fragment was sequenced using the primers L15926 or L15398
202 and H16498. The cyt *b* gene was sequenced using the primers L14398A and H15634A, and
203 when necessary, internal primers for the cyt *b* gene. Internal primers designed specifically for
204 the taxa in this study were: CytbIF1 (5'- CCA TTC CAT ATA TTG GCT CA-3'), CytbIF2 (5'-
205 TCC CAT TCC ATC CTT ACT ACA - 3'), CytbIR1 (5' - CCA ATA TAT GGA ATG GCT GA
206 - 3') and CytbIR2 (5' – GGG GTA TTT AAT GGG TTT GC – 3'). Cycle sequencing reaction
207 conditions consisted of 30 cycles of 20 sec. at 96 ^{o}C (denaturation), 20 sec at 50 ^{o}C (annealing),
208 and 4 min at 60 ^{o}C (extension). Forward and reverse sequences for each individual were
209 assembled using Sequencher 4.5 (Gene Codes Corporation, Ann Arbor, MI), were aligned to a
210 reference sequence from GenBank (*Z. h. preblei*, AY598282 for CR and *Z. trinotatus*, AF119262
211 for cyt *b*), and were double-checked by two researchers. Clustal X (Thompson et al. 1994) was
212 used to obtain multiple sequence alignments for CR and cyt *b*. Alignments of both datasets were
213 performed with default settings and were straightforward as only three sites had indels in the CR

Pre-publication Copy

214 dataset and no indels were encountered in cyt *b*, as would be expected for this protein-coding
215 gene.
216
217 *Data analysis*
218 Microsatellite DNA – Observed genotype frequencies were tested for consistency with Hardy-
219 Weinberg and linkage equilibrium expectations using randomization tests implemented by
220 GENEPOP 3.1 (Raymond and Rousset 1995). The Hardy-Weinberg test used the Markov chain
221 randomization test of Guo and Thompson (1992) to estimate exact two-tailed *p* values for each
222 locus in each sample. Global tests combined these results over loci and sampling locations using
223 Fisher's method (Sokal and Rohlf 1994). Tests for linkage equilibrium were conducted using the
224 randomization method of Raymond and Rousset (1995) for all pairs of loci. Bonferroni
225 adjustments (Rice 1989) were used to determine statistical significance for these and all other
226 simultaneous tests. Average observed (H_o) and expected (H_e) heterozygosities were calculated
227 using the program BIOSYS-1 (Swofford and Selander 1981). The amount of allelic diversity
228 (expressed as allelic richness; El Mousadik and Petit 1996), estimates of population subdivision
229 (F_{ST}), and inbreeding coefficients (F_{IS}) were determined using the program FSTAT (Goudet
230 1995). All pairwise F_{ST} estimates were tested for significance (i.e., difference from zero) by
231 adjusted permutations using FSTAT. The statistical significance of genetic differences between
232 each pair of collections, clusters, and subspecies was tested using the genic differentiation
233 randomization test in GENEPOP (Raymond and Rousset 1995). Results were combined over
234 loci using Fisher's method (Sokal and Rohlf 1994).
235 Several techniques were used to describe genetic relationships between subspecies and
236 collections. We used the model-based clustering method of the program *structure* (Pritchard et
237 al. 2000) to infer population structure among collections and to probabilistically assign all
238 individuals to the detected clusters (*k*). A sequential method of inferring *k* was used by first
239 identifying the "uppermost" hierarchical level of population structure followed by subsequent
240 analysis of each subset revealed to identify within-subset structure (Evanno et al. 2005). In the
241 initial phase, $k = 1$ to $k = 15$ clusters were considered for the 13 collections using a burn-in of
242 15,000 followed by 100,000 iterations. For each *k*, 10 independent analyses were performed.
243 The optimum number of clusters (subsets) in the initial phase was identified using the second
244 order rate of change of the likelihood function with respect to *k* (Δk) as described by Evanno et

Pre-publication Copy

245 al. (2005). Subsequent analysis of each subset tested $k = 1$ to $k = c+3$ (the number of collections
246 (c) included in the subset plus three), with a burn-in of 10,000 followed by 10,000 iterations, and
247 20 analyses for each k. In the subsequent (within-subset) analyses, which included low numbers
248 of collections, the inferred k was determined by identifying the level of k in which q values
249 produced a consensus pattern across the 20 independent analyses. Individual assignment success
250 to the cluster of origin was recorded both as the highest likelihood of assignment (q) and the
251 percentage of individuals in a cluster with $q \geq 0.90$ (Pritchard et al. 2000).
252 Differences between each pair of collections and subspecies were summarized with D_a (Nei
253 et al. 1983), calculated by DISPAN (Institute of Molecular Evolutionary Genetics, The
254 Pennsylvania State University, University Park, PA, USA). Simulation has shown that the
255 geometric-based D_a exhibits a stronger linear relationship with shorter evolutionary times, and
256 therefore, estimates better tree topology than other commonly-used genetic distances when
257 analyzing microsatellite DNA variation (Takezaki and Nei 1996). An unrooted phylogenetic tree
258 was fitted to the data using the D_a distance matrix and neighbor-joining (NJ) algorithm.
259 TreeView (Page 1996) was used to visualize the tree. The strength of support for each node in
260 the tree was tested by bootstrapping over loci using NJBPOP (J.-M. Cornuet, INRA,
261 Montpellier, France). Multidimensional scaling (MDS), with a Kruskal linear regression loss
262 function, was used to summarize relationships among all collections (Systat 10.2; Systat
263 Software, Inc., Point Richmond, CA) based on the D_a. MDS was chosen to provide a
264 perspective of the underlying structure of the D_a matrix without imposition of a bifurcating
265 evolutionary history. Similarity between pairs of collections was represented by $1-D_a$.
266 Analysis of molecular variance (AMOVA) was used to partition genetic variation among
267 clusters and subspecies (Excoffier et al. 1992). Arlequin 2.0 (Schneider et al. 2000) was used to
268 quantify and test the statistical significance of the observed differentiation between subspecies,
269 between clusters of collections, and within subspecies and clusters.
270
271 Mitochondrial DNA - Each molecular data partition was tested by maximum parsimony (MP),
272 distance-based neighbor-joining, and maximum likelihood (ML) analyses using PAUP* 4.0b10
273 (Swofford 2002). For parsimony analyses, heuristic searches were run using unweighted,
274 parsimony-informative (PI) characters with the following settings: starting trees for branch
275 swapping obtained via stepwise addition, 100 random additions of sequences per run, and tree

Pre-publication Copy

bisection-reconnection (TBR) branch swapping on best trees. An estimate of support for each node on parsimony trees was assessed using bootstrap resampling (Felsenstein 1985) with 1,000 replicates and the full heuristic search algorithm. Additionally, Bremer support, or decay indices (Bremer 1988, 1994), were calculated for nodes occurring in the strict consensus tree of the combined dataset using TreeRot v. 2b (Sorenson 1999). *Z. princeps* was used as the outgroup in all analyses. To examine whether differences in phylogenetic signal existed between the two mitochondrial data partitions, incongruence length differences (ILD, Farris et al. 1994) were calculated in PAUP* by the partition-homogeneity test. Settings for ILD tests were as in parsimony analysis, with uninformative sites excluded, and 1,000 replicates per run.

The most appropriate model of DNA substitution for each species and each DNA region independently and for combined dataset was determined using Akaike Information Criterion (AIC) (Posada and Crandall 2001) in Modeltest 3.7 (Posada and Crandall 1998). ML analyses were run in PAUP* with the best-fit model from Modeltest and consisted of a heuristic search with the starting tree obtained by neighbor-joining and utilizing the tree-bisection-reconnection algorithm for branch swapping. ML bootstrap analyses were run with the "fast" stepwise addition algorithm and 500 replicates. NJ trees were generated using the most appropriate genetic distance method based on the results of Modeltest.

We investigated intraspecific gene genealogies for each gene region 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) and was run on representative haplotypes from each data partition which was determined using the program Collapse 1.2 (http://darwin.uvigo.es/software/collapse.html).

MtDNA nucleotide diversity was estimated for the CR, cyt *b* and combined datasets using Arlequin 2.0. 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 (Excoffier et al. 1992). Subspecies- and cluster-level differentiation was assessed using Arlequin 2.0 (Schneider et al. 2000). 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 (1993) (cyt *b*) estimates. The second

Pre-publication Copy

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 320 *Z. hudsonius* sampled from 13 locations representing five neighboring subspecies (see Table 1 for listing and abbreviations; Figure 1). A high level of genetic diversity was detected among the 13 collections; 279 alleles were observed across the 21 loci ranging from 7 alleles at *Ztri*19, *Zhu*C120, and *Zhu*C130 to 30 at *Z*.7 (Appendix A). The mean number of alleles per locus was 13.3. Allelic richness estimates for subspecies ranged from 4.3 (*Z. h.luteus*) to 9.4 (*Z. h. intermedius and Z. h. pallidus*) (Table 2). Average heterozygosity was lower in *Z. h. preblei* than in other subspecies. Observed mean heterozygosity ranged from 54.7% in *Z. h. preblei* to 75.2% in *Z. h. pallidus*. The number of unique alleles observed ranged from 4 (*Z. h. preblei*) to 38 (*Z. h. pallidus*). Estimates of individual pair-wise genetic distances, using the proportion of shared alleles, indicated that levels of genetic diversity observed among the 21 microsatellite loci were sufficient to produce unique multilocus genotypes (i.e., genetic distances > zero) for all animals surveyed (distances not presented).

When organized at the collection level, randomization tests showed that genotypes for the majority of locus-by-collection comparisons were consistent with Hardy-Weinberg expectations. A total of 273 locus-by-collection comparisons were made of which 11 (4.0%) were statistically significant after Bonferroni adjustment for multiple tests at overall $\alpha = 0.05$ ($p < 0.0038$). These comparisons consisted of six collections at six loci: CCWY at *Z*.26; BRCSD and MCMN at *Zhu*C119; MCMN, BCSD, KBCNE, and SCNM at *Zhu*C130; BCSD and KBCNE at *Ztri*19; DCC02 at *Ztri*24; and KBCNE at *Z*.7 (all but the last occurrence due to heterozygote deficiencies). This level of deviation is likely due to one or a combination of factors including sub-structuring of the sample (i.e., Wahlund effect), inbreeding, or the presence of null alleles. Interestingly, only one of the deviations was observed in among collections of *Z. h. preblei*, the subspecies targeted for development of 17 of 21 markers surveyed. This observation, combined with the heterozygote deficiencies observed, suggests that differentiation between *Z. h. preblei*

Pre-publication Copy

338 and the neighboring subspecies may have increased the likelihood that null alleles would be
339 present in the non-target subspecies. The number of deviations from Hardy-Weinberg
340 expectations increased ($N = 19$; 18.1%) when collections were pooled as putative subspecies.
341 The majority ($N = 13$) of the significant deviations (overall $\alpha = 0.05$; $p < 0.013$) were observed
342 among the collections pooled as *Z. h. preblei* ($N = 9$) and *Z. h. intermedius* ($N = 4$). All
343 deviations were the result of heterozygote deficiencies. This result suggests that the allele
344 frequencies of the populations pooled to form these two subspecies (*Z. h. preblei*-North and *Z. h.*
345 *preblei*-South; *Z. h. intermedius*-West and *Z. h. intermedius*-East) have achieved different
346 equilibria and that sub-structuring of the sample (i.e., Wahlund effect) has been detected.
347 Minimal linkage disequilibrium was observed as four of 210 (1.9%) comparisons of each
348 locus pair across all collections was found to be significant after correction for multiple tests
349 (overall $\alpha = 0.05$, $p < 0.0048$). The linkage disequilibrium was observed between *Z*.20 and
350 *Zhu*D122, *Z*.20 and *Z*.52, *Zhu*D107 and *Zhu*D122, and *Z*.7 and *Zhu*D109 with each occurring in a
351 single collection. This finding was likely a result of sampling error, year-class mixing,
352 population mixing, or a combination of the three rather than physical linkage among loci.
353 Allele frequency heterogeneity was observed throughout the study area. Among 1603
354 single-locus pair-wise tests of allele frequency heterogeneity, 1186 (74.0%) indicated departures
355 from homogeneity after correction for multiple tests ($\alpha = 0.05$, $p < 0.0006$). When testing allele
356 frequency heterogeneity across 21-locus genotypes, highly significant heterogeneity was
357 observed in all 78 pairs of collections ($\alpha = 0.05$, $p < 0.0006$).
358 Results from the program *structure* initially indicated that $k = 3$ was the appropriate
359 number of clusters to be recognized among the 13 collections of *Z. hudsonius* genotyped at 21
360 microsatellite DNA loci. The three clusters (subsets) resulted from the 320 individuals (and 12
361 collections) being combined (and designated) as follows: subset [A] - LCCO1, LCCO2, DCCO1,
362 DCCO2, ECCO1, ECCO2 (*Z. h. preblei*); subset [B] - CCWY, CCSD (*Z. h. campestris*) and
363 BRCSD and MCMN (*Z. h. intermedius*); and subset [C] - BCSD and KBCNE (*Z. h. pallidus*)
364 and SCNM (Z. h. luteus) (Figure 2). The three clusters (or subsets) detected the uppermost
365 hierarchical level of population structure, suggesting that the collections constituting *Z. h.*
366 *preblei* form a distinct grouping from the *Z. h. campestris* and *Z. h. intermedius* collections,
367 which are in turn distinct from the *Z. h. pallidus* and *Z. h. luteus* collections. This structure was
368 confirmed by 100% correct assignment of each mouse to the subset-of-origin based on q values

Pre-publication Copy

369 (Table 3). The average value of *q*MAX for the subsets ranged from 0.96 ([B]) to 0.99 ([A] and
370 [C]). Subsequent analysis of the three subsets suggested a strong pattern of genetic
371 differentiation throughout the study area and the presence of $k = 7$ definable clusters (Figure 2).
372 This analysis identified phylogeographic discontinuities present within *Z. h. preblei* that
373 corresponded to the northern and southern collections (clusters 1 and 2, respectively), and *Z. h.*
374 *intermedius* that associate with the western and eastern collections (clusters 4 and 5,
375 respectively), as well as clusters associated with *Z. h. campestris* (cluster 3), *Z. h. pallidus*
376 (cluster 6), and *Z. h. luteus* (cluster 7). The strength of the differentiation among these clusters
377 was also evident upon inspection of the individual assignment results and the average value of
378 *q*MAX (Table 3). That is, when each individual was assigned to cluster based on the largest
379 value of *q*, average assignment success to cluster of origin was 99.7% (319/320 individuals
380 correctly assigned). A lone *Z. h. preblei* individual from *Z. h. preblei*-South (cluster 2) was
381 incorrectly assigned to *Z. h. preblei*-North (cluster 1). Average values of *q*MAX for the seven
382 clusters ranged from 0.95 (*Z. h. intermedius*-West) to 0.99 (*Z. h. pallidus*, cluster 6; *Z. h. luteus*,
383 cluster 7). When using $q \geq 90$ as an assignment threshold, the percentage of correct assignment
384 to cluster ranged from 89.3% (*Z. h. intermedius*-West) to 100.0% (*Z. h. intermedius*-East, *Z. h.*
385 *pallidus,* and *Z. h. luteus*). The correct assignment success at six of the seven clusters exceeded
386 90.2%. When compared at the subspecies level ($q \geq 90$ criterion), assignment success ranged
387 from 90.2% (*Z. h. campestris*) to 100% (*Z. h. preblei*, *Z. h. pallidus,* and *Z. h. luteus*). The
388 highest rate of incorrect assignment occurred between *Z. h. campestris* and *Z. h. intermedius*.
389 Pair-wise genetic distance values (D_a, Nei et al. 1983) were calculated between all
390 collections to investigate evolutionary relationships in allele frequencies (Table 4). The greatest
391 genetic distances occurred between the collection representing *Z. h. luteus* and all other
392 collections; the lowest genetic distances were observed alongside the diagonal between
393 collections from the same subspecies or cluster (as defined by *structure*). The structure
394 underlying the D_a matrix is illustrated with an unrooted neighbor-joining (NJ) tree (Figure 3) and
395 by graphical depiction of the multidimensional scaling analysis (stress of final configuration =
396 0.072; $R^2 = 0.98$; Figure 4). The patterns observed in both graphical depictions are congruent
397 with each other and previous analyses, illustrating the high level of differentiation among and
398 within the five subspecies. The seven distinct clusters of collections identified by *structure* are
399 evident in both diagrams. The distinctiveness of the two clusters associated with *Z. h. preblei* is

Pre-publication Copy

confirmed relative to the other subspecies (100% bootstrap support) and each other (100% bootstrap support). Similarly, both figures depict the differentiation observed between the two *Z. h. intermedius* collections (lower bootstrap support) and the other subspecies. The collections of *Z. h. pallidus* and *Z. h. luteus* appear similarly distinct, while the closest genetic relationship exists between *Z. h. campestris* and *Z. h. intermedius*.

A high level of genetic differentiation was also observed in comparisons of F_{ST} values at the cluster (Table 5), and subspecies (Table 6) scale. All 21 tests of significance in pair-wise F_{ST} values among the scales (cluster and subspecies) were greater than zero ($p < 0.003$) indicating the presence of numerous significant genetic discontinuities throughout this portion of *Z. hudsonius*' range. Pair-wise F_{ST} values were lowest between clusters within subspecies (*Z. h. preblei*-North and *Z. h. preblei*-South, 0.11; *Z. h. intermedius*-West and *Z. h. intermedius*-East, 0.07) and between *Z. h. intermedius*-East and *Z. h. campestris* (0.10). F_{ST} estimates were highest between the *Z. h. preblei* clusters and the *Z. h. luteus, Z. h. intermedius* and *Z. h. pallidus* clusters (all values ≥ 0.20). Moderately high F_{ST} estimates were observed between the *Z. h. preblei* clusters and *Z. h. campestris*, and between *Z. h. pallidus* and the *Z. h. intermedius* and *Z. h. campestris* clusters. F_{ST} estimates at the subspecies level (all collections and clusters pooled) mirrored those observed among clusters with the highest estimates observed between *Z. h. preblei* and Z. h. luteus, *Z. h. intermedius* and *Z. h. pallidus*. The lowest F_{ST} estimate was observed between *Z. h. campestris* and *Z. h. intermedius*.

Quantitative estimates of hierarchical gene diversity (AMOVA) among subspecies and clusters also identified statistically significant genetic structuring, with the greatest amount of variation distributed within clusters and subspecies. Results of the comparison between the five subspecies (all collections pooled) determined that 17.4% ($p < 0.000001$) of the genetic variation occurred between subspecies and 82.6% ($p < 0.000001$) was due to differentiation within subspecies. When the data were partitioned into clusters as identified by the program *structure*, slightly more variation could be attributed to differentiation among clusters (18.1%, $p < 0.00001$) than among subspecies, and 81.9% ($p < 0.000001$) of the variance was found within clusters. When clusters are grouped into subspecies, 11.3% ($p < 0.024$) of the variation was distributed between subspecies, 7.9% ($p < 0.000001$) between clusters within subspecies, and 80.8% ($p < 0.000001$) of the variance observed within clusters.

Pre-publication Copy

Mitochondrial DNA

Control region

A total of 305 *Z. hudsonius* individuals from 13 collections representing five neighboring subspecies were analyzed for sequence variation at the 5'-end of the mtDNA control region (CR). The number of polymorphic sites, haplotypes, and nucleotide diversity observed for the 374 bp of CR are provided in Table 7. For this region, 28 (7.5%) sites were variable, 27 (7.2%) sites were parsimony-informative, three sites included indels, and all were parsimony-informative when coded as a fifth base. Nucleotide diversity was low across the study, and ranged from 0.0000 in *Z. h. luteu*s to 0.0060 in *Z. h. pallidus (*Table 7*)***.** Twenty-seven CR haplotypes were observed study-wide for *Z. hudsonius*, yet none were shared among the five subspecies (Appendix B). Haplotype counts differed dramatically between collections for each subspecies and cluster (as defined for microsatellite DNA variation by the program *structure*) as 18 of 25 (72%) haplotypes were not shared among collections. For *Z. h. preblei*, 3 of 4 CR haplotypes observed were not shared between the northern and southern clusters.

A network of *Z. hudsonius* CR haplotypes was constructed using statistical parsimony methods (Figure 5). Gaps were included as a fifth character state and the 0.95 limit for connections was eight steps. Two unconnected haplotype networks resulted from the statistical parsimony analysis that corresponded to a *Z. h. preblei*, *Z. h. intermedius*, and *Z. h. campestris* network, and a *Z. h. pallidus*, and *Z. h. luteus* network. Within the *Z. h. preblei* - *Z. h. intermedius* - *Z. h. campestris* network, *Z. h. preblei* haplotypes formed a group that was one mutational step from a group comprised of eight *Z. h. intermedius* haplotypes, which was in turn separated from a group of five *Z. h. campestris* haplotypes and one *Z. h. intermedius* haplotype (ZhiCR_E) by one mutational step.

Adjusted (net) sequence divergences (Kimura 2-parameter) between *Z. hudsonius* subspecies for the mtDNA CR were relatively low (Table 8), averaging 3.37%. Net distances ranged between 0.29% (*Z. h. campestris* - *Z. h. intermedius*) and 5.63% (*Z. h. campestris* - *Z. h. luteus*). Distances between *Z. h. preblei* and other subspecies ranged between 0.57% (*Z. h. intermedius*) and 5.08% (*Z. h. luteus*). The genetic distance between *Z. h. preblei* and *Z. h. campestris* (1.03%) was nearly twice the distance observed between *Z. h. preblei* and *Z. h.*

Pre-publication Copy

462 *intermedius*. The average genetic distance between *Z. hudsonius* subspecies and *Z. princeps* was
463 10.33%.

464 Analysis of molecular variance of CR indicated the presence of strong genetic
465 differentiation among the five subspecies. Global Φ_{ST} (0.96) and F_{ST} (0.37) values, for distance
466 (Kimura 2-parameter) or frequency-based analysis (use of conventional F-statistics),
467 respectively, were high and significantly different from zero ($p < 0.00001$) (Table 9). As no
468 haplotypes were shared among the five *Z. hudsonius* subspecies, F_{ST} values (a measure of
469 haplotype frequencies only) generated by the AMOVA analysis have limited biological
470 resolution and are provided only for comparison to previous findings (Ramey et al. 2005). A
471 global Φ_{ST} of 0.96 indicates that nearly all of the CR haplotype variance (96%) was distributed
472 between subspecies, a result consistent with the absence of any shared haplotypes. Pair-wise
473 Φ_{ST} comparisons of *Z. h. preblei* with the other subspecies ranged from $\Phi_{ST} = 0.69$ (*Z. h.*
474 *intermedius*) to $\Phi_{ST} = 0.95$ (*Z. h. luteus*) and averaged $\Phi_{ST} = 0.85$. A comparison of CR
475 haplotype variation among the three subspecies proposed for synonymy by Ramey et al. (2005),
476 *Z. h. preblei*, *Z. h. intermedius*, and *Z. h. campestris*, resulted in 75% of the variance being
477 distributed among subspecies. Moreover, the AMOVA suggests that *Z. h. preblei* is distinct
478 from other neighboring subspecies and is evolutionarily more similar to *Z. h. intermedius* than to
479 *Z. h. campestris*. This finding differs from the pattern observed during analysis of the nuclear
480 DNA, which found *Z. h. preblei* to exhibit a lower genetic distance between *Z. h. campestris*.
481 The lowest level of differentiation among these three subspecies was observed between *Z. h.*
482 *intermedius* and *Z. h. campestris* ($\Phi_{ST} = 0.59$). Significant differentiation was recorded
483 between all subspecies comparisons with Φ_{ST} values ranging from 0.57 (*Z. h. pallidus – Z. h.*
484 *luteus*) to 0.97 (*Z. h. campestris – Z. h. luteus*).

485 CR haplotype variation distributed among the seven clusters was similar to that observed
486 among the subspecies, as 90.2% of the variance was distributed among clusters and 9.8% within
487 clusters. CR haplotype variation distributed among and within the seven clusters found that
488 82.8% ($p < 0.025$) of variance was distributed among defined clusters, 7.3% ($p < 0.00001$)
489 distributed among collections within clusters, and 9.9% ($p < 0.00001$) was attributed to
490 differences within collections. These findings mirrored the high level of genetic differentiation
491 observed upon survey of these clusters with the nuclear microsatellite DNA markers. Exact tests
492 for differentiation based on haplotype frequencies for all pairs of subspecies ($N = 10$

Pre-publication Copy

comparisons) and clusters ($N = 15$ comparisons) were statistically significant ($p < 0.00001$).

Cytochrome b

A total of 292 *Z. hudsonius* individuals from 13 collections representing five neighboring subspecies were analyzed for sequence variation at the mtDNA cyt *b* gene. For this region, 116 (11.5%) sites were variable, and 84 (8.3%) sites were parsimony-informative. The number of polymorphic sites, haplotypes, and nucleotide diversity observed for each subspecies for the 1006 bp of cyt *b* are provided in Table 7. Nucleotide diversity was low across the study and ranged from 0.000 in *Z. h. luteu*s to 0.003 in *Z. h. pallidus***.** Fifty-nine cyt *b* haplotypes were observed study-wide for *Z. hudsonius*, yet none were shared among the five subspecies (Appendix C). Haplotype counts differed dramatically between collections for each subspecies and cluster, as 45 of 52 (86%) haplotypes were not shared. For *Z. h. preblei*, none of the 17 haplotypes observed were shared between the northern and southern clusters.

Statistical parsimony analysis of the cyt *b* data set resulted in three haplotype networks: a *Z. h. preblei*, *Z. h. intermedius*, and *Z. h. campestris* network; a *Z. h. pallidus* network; and a *Z. h. luteus* network (Figure 6). These three networks were unconnected due to the number of mutational steps between them being greater than the $p = 0.95$ limit of 13 steps. Within the *Z. h. preblei*, *Z. h. intermedius*, and *Z. h. campestris* network, the *Z. h. preblei* haplotypes were recovered as a group, separated from a common *Z. h. intermedius* haplotype by at least six mutational steps. A group of 12 *Z. h. intermedius* haplotypes was separated from a group of 13 *Z. h. campestris* haplotypes and one *Z. h. intermedius* haplotype (ZhiCB_A) by two mutational steps. A single *Z. h. intermedius* haplotype (ZhiCB_A) was separated from the *Z. h. campestris* haplotype ZhcCB_B by 3 mutation steps, and was not found in any *Z. h. campestris*.

Adjusted (net) sequence divergences (Tamura-Nei) between *Z. hudsonius* subspecies for the mtDNA cyt *b* exhibited a similar pattern to that observed for CR, averaging 4.43% (not shown). Net distances ranged between 0.18% (*Z. h. campestris* - *Z. h. intermedius*) and 7.11% (*Z. h. campestris* - *Z. h. pallidus*). Distances between *Z. h. preblei* and other subspecies ranged between 0.69% (*Z. h. intermedius*) and 6.78% (*Z. h. pallidus*). The average genetic distance between *Z. hudsonius* subspecies and *Z. princeps* was 19.87%.

As was the case with the CR, AMOVA indicated the presence of strong genetic

Pre-publication Copy

differentiation among all five subspecies at cyt *b*. Global Φ_{ST} (0.96) and F_{ST} (0.33) values were high and significantly different from zero ($p < 0.00001$) (Table 9). Nearly all of the cyt *b* variance (96%) was distributed between subspecies, a result consistent with the absence of shared cyt *b* haplotypes and with the patterns observed in the CR. In general, comparisons among all subspecies resulted in higher Φ_{ST} values for cyt *b* than for CR but with one exception, the comparison between *Z. h. intermedius* and *Z. h. campestris*. Pair-wise Φ_{ST} comparisons of *Z. h. preblei* ranged from $\Phi_{ST} = 0.82$ (*Z. h. intermedius*) to $\Phi_{ST} = 0.99$ (*Z. h. luteus*) and averaged $\Phi_{ST} = 0.92$. A comparison of cyt *b* haplotype variation among the three subspecies proposed for synonymy by Ramey et al. (2005), resulted in 84% of the variance being distributed among subspecies. Cytochrome *b* AMOVA results confirm that *Z. h. preblei* is distinct from other neighboring subspecies and is evolutionarily most similar to *Z. h. intermedius*. As with the CR, the lowest level of differentiation among these three subspecies was observed between *Z. h. intermedius* and *Z. h. campestris* ($\Phi_{ST} = 0.53$).

Cytochrome *b* haplotype variation distributed among the seven clusters found that 96.1% ($p < 0.000001$) was attributable to differentiation among clusters and the remainder due to variation within. Cytochrome *b* haplotype variation distributed hierarchically among and within the seven clusters found that 92.7% ($p < 0.044$) of the variance was distributed among defined clusters, 2.5% ($p < 0.00001$) distributed among collections within clusters, and 4.8% ($p < 0.00001$) was attributed to differences within collections. These findings mirrored the high level of genetic differentiation observed upon survey of these clusters at the mtDNA CR and with the nuclear microsatellite DNA markers. Exact tests for differentiation based on haplotype frequencies for all pairs of subspecies ($N = 10$ comparisons) and clusters ($N = 15$ comparisons) were statistically significant ($p < 0.00001$).

Combined mitochondrial DNA

Results of ILD tests were insignificant ($p = 0.23$), plus phylogenetic analyses of the individual data partitions were congruent with respect to the major clades recovered (data not shown), so data was combined for individuals occurring in both data sets, then collapsed into 64 unique haplotypes for *Z. hudsonius* (Table 7). The strict consensus tree of 4,050 most parsimonious trees (Figure 7A; 205 steps, C.I. = 0.7171) revealed two well-supported clades within *Z. hudsonius*. One clade, supported by 100 percent of bootstrap pseudoreplicates and a decay index

Pre-publication Copy

of 17, was comprised of *Z. h. pallidus* and *Z. h. luteus* haplotypes. Within this clade, the two *Z. h. luteus* haplotypes formed a well supported monophyletic group. The other major clade recovered, supported by 100 percent of bootstrap pseudoreplicates and a decay index of 36, was comprised of *Z. h. intermedius*, *Z. h. campestris*, and *Z. h. preblei* haplotypes. Within this clade none of the *Z. hudsonius* subspecies were recovered as monophyletic, yet the *Z. h. campestris/Z. h. intermedius* haplotypes formed a clade separate from *Z. h. preblei* haplotypes, and this relationship was well supported (bootstrap values of 90/88, decay index of 4, Figure7A). Nodes within the *Z. h. campestris*/*Z. h. intermedius* clade were poorly supported by bootstrap and decay indices. Modeltest selected the Tamura-Nei + I model of evolution as the best fit for this combined mitochondrial data set using AIC. Model parameters were: base frequencies of A = 0.2752, C = 0.2633, G = 0.1217, T = 0.3399; proportion of invariable sites = 0.7272. ML analysis run with the best-fit model produced a tree similar in topology to the MP consensus (Figure 7A), with the major difference being the relationship between *Z. h. pallidus* and *Z. h. luteus*- they are not sister taxa in the ML analysis (Figure 7B). In the phylogram of the ML tree (Figure 7B), it is notable that branch lengths were short within the *Z. h. campestris*/*Z. h. intermedius*/*Z. h. preblei* clade, with the longest branch (and the only one with substantial boostrap support) separating *Z. h. preblei* from *Z. h. campestris*/*Z. h. intermedius* haplotypes.

DISCUSSION

The present survey of 21 microsatellite DNA loci and 1380 base pairs from two regions of mtDNA indicated that each *Z. hudsonius* subspecies is genetically differentiated from its neighboring subspecies to varying degrees, and therefore, does not support the null hypothesis of a homogeneous gene pool among the five subspecies from the southwestern portion of the species' range. The magnitude of the observed differentiation was considerable and was supported by highly significant findings for nearly every statistical comparison made, regardless of the genome or the taxa under consideration. Given the disjunct distribution of selected subspecies and the strong structuring of nuclear multilocus genotypes, combined with the presence of only subspecies-specific mtDNA haplotypes, the genetic discontinuities observed throughout the study area correspond with previously proposed taxonomic classifications based on the geographic separation and surveys of morphological variation among these subspecies (Krutzsch 1954, Hafner et al. 1981).

Pre-publication Copy

586 The microsatellite DNA dataset contained sufficient allelic diversity to detect a unique
587 multilocus genotype for each of 320 mice sampled, and to elucidate seven distinct clusters
588 among the 13 *Z. hudsonius* collections that corresponded to the five subspecies surveyed, while
589 delineating significant differentiation within two of these taxa (Figures 2-4). Microsatellite
590 DNA identified significant allele frequency heterogeneity among all collections, clusters (as
591 identified by the program *structure*), and subspecies, a high degree of population subdivision
592 (F_{ST}) within clusters (0.07-0.11) and among subspecies (0.11-0.27; all F_{ST} values statistically >
593 zero), large genetic (D_a) distances within (0.042-0.200) and among subspecies (0.218-0.732),
594 clear graphical separation of subspecies and clusters accompanied by high bootstrap support in
595 the NJ tree, and a high percentage of correct assignment to cluster (98%) and subspecies (100%)
596 of origin. *Z. h. preblei* and *Z. h. pallidus / Z. h. luteus* were observed to be the most distinct
597 subspecies in the comparison and *Z. h. campestris* and *Z. h. intermedius* appeared most closely
598 related. Significant population structuring was observed between the two *Z. h. preblei* clusters
599 (North and South), as well as between the two *Z. h. intermedius* collections/clusters (West and
600 East).

601 Phylogeographic analyses for both the mtDNA CR and cyt *b* regions were generally
602 congruent with each other and with the patterns observed in the nuclear DNA. Sequence
603 variation revealed unique sets of haplotypes for each subspecies (Figures 5 and 6), high
604 estimates of differentiation among subspecies (Φ_{ST}; 0.53-0.99; all values significantly > zero;
605 Table 9), and cyt *b* sequence divergences between most subspecies were within the range of
606 those observed in intraspecific comparisons of other rodents (e.g., *Peromyscus boyllii*, 0.3% -
607 1.8%, Tiemann-Boege et al. 2000; *Neotoma floridana*, 0.8% - 5.3%, Edwards and Bradley 2001;
608 *Microtus agrestis*, 0.5% - 5.2%, Jaarola and Searle 2002; see Bradley and Baker 2001 for
609 synthesis). Haplotype patterns also suggested that significant differentiation was observed
610 within selected subspecies, as little or no haplotype sharing was observed among clusters or
611 collections within subspecies (Appendices B and C). In general, the patterns of haplotype
612 variation and differentiation were greater and more informative overall for the larger cyt *b* region
613 (Appendix C; Figure 6).

614 The phylogenetic (i.e., parsimony, maximum likelihood) analyses identified the existence
615 of two major evolutionary lineages within *Zapus hudsonius*: a *Z. h. preblei*, *Z. h. campestris*, and
616 *Z. h. intermedius* clade and a *Z. h. pallidus* and *Z. h. luteus* clade. Although phylogenetic

Pre-publication Copy

analyses did not reveal a strong geographic pattern of reciprocal monophyly, the diagnostic structure observed in the haplotype networks for both DNA regions combined with the large degree of differentiation (F_{ST}, AMOVA, D_a) and high assignment success observed in the microsatellite DNA analyses, indicate that an intermediate state of differentiation has been achieved within the two major lineages.

Comparisons with a Previous Study

The findings and conclusions of this study are contradictory in nearly every comparison with the conclusions drawn from the only other published molecular genetics data available on the taxonomic status of *Z. hudsonius* subspecies (Ramey et al. 2005). It is imperative that the disparities between the previous and present study be revealed due to the conservation implications of synonymizing *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius* as proposed by Ramey et al. (2005). We believe that fundamental differences exist between the two studies and that the lack of correspondence is not simply a matter of interpretation (or misinterpretation) between "lumpers and splitters". The two studies differ significantly in sampling regime, sources of tissues chosen for analysis, number of molecular characters sequenced and microsatellite fragments surveyed, test statistic applied to AMOVAs, method used to portray genealogical relationships, and the criteria used to determine uniqueness. We will discuss the nature of these discrepancies and suggest how each approach impacts the conclusions drawn regarding subspecies distinctness.

An appropriate sampling strategy is central to the successful delineation of population genetic and phylogeographic structures (Baverstock and Moritz 1996). The "population-oriented" strategy used in the present study differed significantly from the "one or few individuals per site across a broad geographical area" approach applied by Ramey et al. (2005). The sampling strategy used by Ramey et al. (2005), likely under-estimated the level of within-population variation, effectively inflating within-subspecies variance, while simultaneously lowering the amount of total variance attributed to between-*Z. hudsonius* subspecies differentiation and decreasing the likelihood that uniqueness would be demonstrated. The present study focused on collections of reasonable size (ranging from 14 to 33 individuals) from recognized geographic populations. This sampling strategy increased the likelihood that all haplotypes present within a population will be surveyed. While both sampling strategies have

Pre-publication Copy

648 their strengths and weaknesses, we believe the strategy used in the present study to be the more
649 pragmatic approach for testing statistical significance under the hypothesis testing approach
650 espoused by Ramey et al. (2005) in which uniqueness was based on the partitioning of variation
651 within and between subspecies.

652 Museum specimens are increasingly recognized as a source of DNA for phylogeographic
653 and phylogenetic investigations (Hoelzel et al. 2002). Genetic information garnered from such
654 tissues provides a means to test unique hypotheses designed to aid in reconstructing past
655 evolutionary relationships (Sorenson et al. 1999). However, DNA in archived specimens
656 exposed to traditional museum preservation and storage decays or becomes chemically modified
657 with time, such that PCR amplicon length and amplification efficiency are inversely proportional
658 to storage time (Glenn et al. 1999, Moum et al. 2002), which constrains the overall utility of such
659 biological material. In addition, the process is far from infallible, as studies are regularly marred
660 by erroneous findings in the form of base shifts, reference bias, phantom mutations, base mis-
661 scoring, artificial recombination (Yoa et al. 2004), and unrealized contamination at the sample
662 and laboratory level (Pääbo et al. 2004, Willerslev and Cooper 2005). Ramey et al. (2005) used
663 decades-old, dried museum skins as a principal source of genetic material for selected collections
664 used in the phylogeographic comparison among *Z. hudsonius* subspecies. Given the availability
665 of live specimens and the relative ease of obtaining minimally invasive tissue samples, this
666 methodological decision may have introduced unnecessary ambiguity to the findings. For
667 example, Ramey et al. (2005) reported the presence of *Z. h. preblei* haplotypes in DNA extracted
668 from five dried museum skins of Z. *h. campestris* collected from Custer County, SD. The
669 authors suggested this finding indicated recent gene flow and alluded to the presence of these
670 haplotypes as a critical element in the decision to recommend synonymy of these subspecies. In
671 the present study, 31 *Z. h. campestris* sampled recently from the same site in Custer County, SD
672 used by Ramey et al. (2005), along with 30 additional specimens from neighboring Crook
673 County, WY were subjected to mtDNA CR and cyt *b* sequence analysis. All 61 individuals were
674 determined to posses *Z. h. campestris*-specific mtDNA haplotypes. Moreover, the same
675 conclusion was reached with the microsatellite loci, as no *Z. h. campestris* individual from either
676 of these collections was assigned to *Z. h. preblei*. Given the prominent role the haplotypes
677 obtained for the five museum skins from Custer County, SD and two additional specimens from
678 Carter County, MT have played in the conclusions drawn by Ramey et al. (2005), it is

Pre-publication Copy

679 unsatisfactory that an *a posteriori* analysis was not considered as part of a routine quality
680 assurance/quality control effort. Since no attempts were made to reproduce the previous CR
681 results, to confirm the findings with another region of mtDNA, or to apply an additional finer-
682 resolution technique such as microsatellite DNA analysis, combined with our failure to detect *Z.*
683 *h. preblei* haplotypes among 61 *Z. h. campestris* from the same and an adjacent location, the
684 conclusions drawn by Ramey et al. (2005) should be considered questionable.

685 Weak phylogeographic signal may result in poor resolution or inaccurate gene trees as an
686 artifact of phylogenetic reconstruction (Avise 2004). When a small gene fragment is considered,
687 the collected data may provide too few synapomorphies to robustly recover the underlying gene
688 tree (Funk and Omland 2003). Too few characters of any single region of DNA may not reflect
689 actual systematic relationships (a "species tree") among taxa, but instead may reflect the history
690 of that gene lineage (a "gene tree"; Brower et al. 1996; see Edwards and Beerli 2000, Nichols
691 2001, Hudson and Turelli 2003 for reviews). Ramey et al. (2005), when considering the
692 phylogenetic status of *Z. hudsonius* subspecies, surveyed sequence variation at a 346 bp
693 fragment of the mtDNA CR. Ramey et al.'s (2005) critical test of uniqueness for *Z. h. preblei*
694 mtDNA sequence data was that greater molecular variance is demonstrated between subspecies
695 than within or that collections show nearly complete reciprocal monophyly. These criteria
696 appear disproportionately stringent given the small number of informative sites examined in the
697 346 bp fragment of the mtDNA CR. With such considerations in mind, and in order to establish
698 whether genealogical partitions among *Z. h. preblei* populations and collections of other putative
699 subspecies are, in fact, bona fide, patterns of divergence from additional mtDNA characters (or
700 genes) would have been appropriate. The present study included the survey of an additional
701 1006 bp of the cyt *b* region which served to augment phylogeographic resolution and verified the
702 major mitochondrial CR and microsatellite DNA findings. The cyt *b* results revealed: 1) more
703 resolved relationships among the *Z. h. pallidus* and *Z. h. luteus* haplotypes, those subspecies
704 being monophyletic in the parsimony tree (Figure 7) and supported by high bootstrap values; 2)
705 diagnostic haplotype patterns between the five *Z. hudsonius* subspecies; and 3) discontinuity
706 between the *Z. h. preblei*-North and *Z. h. preblei*-South clusters, as no cyt *b* haplotypes were
707 shared between the two groupings.

708 Large numbers of microsatellite DNA loci provide better estimates of population genetic
709 parameters given the stochastic variation (including size homoplasy) expected among

Pre-publication Copy

independent loci (Adams et al. 2004, Takezaki and Nei 1996). Ramey et al. (2005) presented the findings of five microsatellite DNA loci surveyed among *Z. hudsonius* subspecies. Given that microsatellite loci are densely interspersed in eukaryotic genomes (Katti et al. 2001), five polymorphic microsatellite DNA loci may not provide sufficient representation of the nuclear DNA variation present (Beaumont and Nichols 1996, Pritchard and Rosenberg 1999) or provide a sufficiently robust multilocus genotype for population/phylogeographic comparisons (Smouse and Chevillon 1998). Comparisons of microsatellite results between the present study (21 loci) and Ramey et al. (2005) indicate that some similarities and considerable differences exist. Among comparisons of basic Mendelian statistics, the present study observed fewer deviations from Hardy-Weinberg expectations, lower estimates of allelic richness and F_{IS}, and greater estimates of private alleles than reported by Ramey et al. (2005). Both studies demonstrated lower allelic richness, fewer private alleles, and lower average heterozygosity in collections of *Z. h. preblei*; findings indicative of the demographic status of the two distinct populations of this subspecies. In a test to determine the number of diagnosable clusters present in their microsatellite DNA dataset using the program *structure*, Ramey et al. (2005) found the variability of likelihood estimates to be high, and concluded that the eight clusters identified by the analysis were poorly defined. This finding seemed to be confirmed by the poor assignment to cluster reported. In the present study, seven distinct clusters were identified (Figure 2) among the five subspecies with *Z. h. preblei* (North and South) and *Z. h. intermedius* (West and East) collections each divided into two individual clusters. The significance of clustering was confirmed by the assignment test associated with the *structure* analysis (Table 3). When individuals were assigned to cluster based on the largest value of q, 99.7% of the individuals were correctly assigned to cluster and 100% were assigned to subspecies (Table 3). When the test is constrained by the $q > 0.90$ criterion, 96.1% of the *Z. hudsonius* individuals were assigned to cluster and 96.8% were assigned to subspecies. The present study generated larger estimates of subspecies differentiation (F_{ST}; 0.11 – 0.34) than found in the previous study (0.01-0.16), which is consistent with results of the AMOVA that found a higher percentage of the variation was distributed between subspecies (17.4%) and clusters (18.1%), than that reported by Ramey et al. (2005) (7.5% - 8.9%). It would appear that the increased coverage of the nuclear genome afforded by the larger number of microsatellite loci yielded increased resolution of the phylogeographic structure present in *Z. hudsonius*.

Pre-publication Copy

741 For microsatellite DNA comparisons, Ramey et al.'s (2005) critical test of uniqueness for
742 subspecies and historic genetic exchangeability was two-fold: that most variation was observed
743 between subspecies in pairwise AMOVA comparisons and that multiple private alleles are found
744 at higher frequency than shared alleles at a majority of loci. These criteria often are not met
745 even among accepted species due to the constraints on fixation indices generated from
746 microsatellite DNA loci created by high heterozygosity levels and homoplasy (Hedrick 1999;
747 Balloux et al. 2000). We contend that these criteria are unfounded and likely have little
748 biological meaning. Ramey et al. (2005) assessed the differentiation among selected
749 subspecies of *Z. hudsonius* with a traditional pairwise haplotypic distance approach designed to
750 elucidate the hierarchical nature of the inter-specific evolutionary relationships among genes
751 (Posada and Crandall 2001). Because intraspecific phylogeographic datasets are predisposed to
752 have fewer informative characters, traditional phylogenetic methods can lead to inadequately
753 portrayed genealogical relationships (Brower et al. 1996). Our study endeavored to elucidate
754 intraspecific gene genealogies that are not hierarchical in nature and assumed to be the result of
755 relatively recent evolutionary processes. Therefore, we opted to emphasize intraspecific gene
756 geneaologies for each subspecies at two mtDNA regions using a haplotype network analysis that
757 implemented the statistical parsimony approach of Templeton et al. (1992) and Crandall et al.
758 (1994). The network analysis better depicted ancestral relationships and the nonrandom
759 distribution of mutations among lineages, which resulted in the diagnostic haplotypic structure
760 observed between subspecies and not observed in the more traditional parsimony analyses
761 performed in our study or the distance-based approach utilized by Ramey et al. (2005).
762 Analysis of molecular variance applied to the mtDNA CR and cyt *b* sequences indicated
763 the presence of strong genetic differentiation among the five subspecies. It is instructive to
764 underscore the differences between the two test statistics (Φ_{ST} and F_{ST}) generated by the
765 AMOVA routine in Arlequin 2.0 (Schneider et al. 2000). Φ_{ST} incorporates sequence divergence
766 between haplotypes (providing the option of several distance metrics), as well as determining
767 haplotype frequencies. Evolutionary differences among the haplotypes (i.e., mutations) are
768 incorporated into calculations of the test statistics. F_{ST}, which utilizes conventional F-statistics,
769 treats all haplotypes, regardless of their evolutionary interrelationships, as equally differentiated
770 (i.e., distance = 1.0), and assesses the variance distribution based on haplotype frequencies alone.
771 Differences between haplotype frequencies are assumed to be due to genetic drift. In the present

Pre-publication Copy

study, Φ_{ST} observed for CR ranged from 0.57 (between *Z. h. pallidus* and *Z. h. luteus*) to 0.97 (*Z. h. campestris* and *Z. h. luteus*). The comparison among *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius* resulted in a Φ_{ST} value of 0.75 (or 75% of variation is distributed between subspecies) and a F_{ST} value of 0.33 (33% of the variation distributed between subspecies). The Φ_{ST} value for this three-subspecies comparison is more than two-fold greater than the value observed by Ramey et al. (2005) (F_{ST} = 0.31) for this comparison, while the F_{ST} value obtained in our study is similar to that observed in the previous study.

The more than two-fold differential in the estimate of variation between subspecies provided by Φ_{ST} (this study) and F_{ST} (Ramey et al. 2005) broaches important issues regarding the conclusions drawn from the respective studies. First, any phylogeographic or phylogenetic comparison should take into consideration the genealogical/evolutionary relationships among the observed haplotypes provided by Φ_{ST}. Second, no haplotypes were shared among the five *Z. hudsonius* subspecies in the present study, and therefore Φ_{ST} would appear to be the most appropriate statistic as F_{ST} values (a measure of haplotype frequencies only) generated by the AMOVA analysis have limited biological resolution or significance in a situation where no haplotypes are shared. Third, when Φ_{ST} and F_{ST} values are similar in magnitude, any population differentiation is likely due to shuffling of closely related haplotypes (i.e., gene drift). If Φ_{ST} is considerably greater than F_{ST} (as is the case for all subspecies comparisons in this study), this implies some level of evolutionary depth to the separations among haplotypes (Quattro et al. 2002). Lastly, Ramey et al.'s (2005) critical test of uniqueness for *Z. h. preblei* mtDNA sequence data was that greater molecular variance be demonstrated between subspecies than within. The results obtained between all subspecies compared in the present study exceed this test criterion, as all Φ_{ST} values were > 0.50. It would appear that by opting to utilize the F_{ST} statistic and failing to recognize the evolutionary differences among haplotypes observed in the previous study, Ramey et al. (2005) have created an unduly stringent (and possibly unachievable) criterion for determining subspecies uniqueness.

The null hypothesis stated by Ramey et al. (2005) was one of no significant intraspecific genetic differentiation (i.e., high exchangeability) among *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius*. We concur with these authors that "well-defined criteria and regulations" would be ideal when considering ESA issues. However, requiring unduly stringent or unachievable evidence to reject null hypotheses as implemented in Ramey et al. (2005) does not present a

Pre-publication Copy

tenable conservation strategy. We contend that due to the limited number of mtDNA CR nucleotides (346) and polymorphic microsatellite markers surveyed (five), reliance on questionable sources of tissue for DNA extraction, the misapplication of certain statistical analyses (e.g., F_{ST} instead of Φ_{ST}), and the excessively stringent requirements for uniqueness, Ramey et al. (2005) have in all likelihood failed to correctly reject the null hypothesis of no significant differentiation among the five subspecies and failed to recognize the presence of unique taxa among *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius*.

CONSERVATION IMPLICATIONS

The vast literature of molecular genetic comparisons amassed to date have demonstrated that many species in need of conservation are subdivided into genetically distinct geographic populations exhibiting considerable historical and phylogeographic separation (see Avise 2004 for synthesis). The genetic criteria chosen for delineating isolated intraspecific groupings (e.g., species, subspecies, phylogroups, distinct population segments, evolutionary significant units, or management units) becomes of critical importance when the entities continued existence is at risk (Moritz 1994a, Moritz 1994b; Paetkau 1999). In the initial comparison of genetic distinctness among subspecies of *Z. hudsonius*, Ramey et al. (2005) applied critical tests to the hypotheses of genetic and ecological exchangeability as proposed by Crandall et al. (2000). The authors suggested the establishment of a conceptually sound and consistent methodological approach for determining what constitutes species, subspecies, or DPS listings under the US ESA and international biodiversity laws (i.e., uniform criteria that apply equally to all taxa). We feel that this represents a laudable objective; however, as previous authors have observed (Waples 1991, Moritz 1994a, Pennock and Dimmick 1997, Waples 1998, Avise and Johns 1999, Avise and Walker 2000, Crandall et al. 2000, Goldstein et al. 2000, Hendry et al. 2000), operational thresholds for intraspecific designation are often necessarily arbitrary, inconsistent among genetic markers and genomes, not comparable across taxa, and it remains that all-encompassing criteria continue to be elusive.

We found the Ramey et al.'s (2005) core genetic-based criteria for subspecies distinctness (i.e., greater variation between subspecies than within) to be too stringent to account for relatively recently radiated taxa, scientifically unfounded as this criterion would not distinguish certain well-established species, and under certain circumstances unachievable with

Pre-publication Copy

the test statistics applied. Our concern is that if the criteria for genetic distinctness are inflexible, this can represent a conservation stratagem that impacts the potential for future evolutionary change within the intraspecific unit of management and that could lead to extirpation or extinction of discrete and evolutionarily significant intraspecific diversity. Consequently, these criteria should not be considered as an accepted standard when addressing US ESA actions or for international conservation laws.

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 marker types. Barriers to gene flow between populations of *Z. hudsonius preblei* and other subspecies have been of sufficient duration that these populations are now distinguishable by genetic characteristics and that the observed pattern is congruent with accepted subspecific designations (Krutzsch 1954; Hafner et al. 1981). This study found no evidence that warranted a taxonomic revision among the *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius* as previously proposed (Ramey et al. 2005), nor did this survey find data suggesting the need to alter *Z. h. preblei*'s standing under the ESA. Moreover, the level of differentiation observed between *Z. h. preblei* inhabiting Larimer County, CO (*Z. h. preblei*-North) and mice sampled from Douglas and El Paso counties, CO (*Z. h. preblei*-South) indicates that this subspecies is comprised of at least two genetically distinct populations worthy of individual management consideration. Sufficient data has also been presented to suggest that the differentiation observed between the *Z. h. pallidus* and *Z. h. luteus* clade and other *Z. hudsonius* subspecies warrants further study and serious consideration given to species-level recognition for the former.

We endorse the widespread opinion that decisions to protect a species and intraspecific divisions of vertebrate taxa under the ESA should be based on investigations of life history, ecology, population dynamics, and systematics (see Crandall et al. 2000 for review). However, molecular genetics is a tool with the demonstrated ability to identify evolutionarily divergent lineages that other methodologies failed to reveal (Moritz and Hillis 1996). 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 not uncommonly seen within rodent species (e.g., *Peromyscus boylii*, Tiemann-Boege et al.

Pre-publication Copy

865 2000; *Neotoma floridana*, Edwards and Bradley 2001; *Microtus agrestis*, Jaarola and Searle
866 2002). The level of discontinuity observed in this study among subspecies should not be
867 considered minor or "shallow" (Avise 2004) simply because the observed differentiation is not
868 accompanied by the presence of reciprocal monophyly, phenotypic divergence, or obvious
869 adaptive significance. The differential magnitude observed between Φ_{ST} (distance and
870 frequency) and F_{ST} (frequency) values for both mtDNA CR and cyt *b* suggests the presence of
871 moderate microevolutionary depth among the five subspecies accompanied by
872 phylogeographical structuring of haplotypes and multilocus genotypes within *Z. h. preblei* and *Z.*
873 *h. intermedius*. The diagnostic nature of the haplotype variation and large degree of nuclear
874 genetic differentiation at microsatellite loci suggests that these subspecies (and certain
875 geographic populations) have been reproductively isolated for such time that they appear to be
876 on independent evolutionary trajectories, while having nearly achieved complete lineage sorting.
877

878 A detailed comparison similar to that performed here for all subspecies of *Z. hudsonius* is
879 warranted. Such a study would likely shed light on the significance of the considerable
880 differentiation observed among the *Z. h. preblei, Z. h. campestris,* and *Z. h. intermedius* clade
881 and will allow the differentiation observed for *Z. h. pallidus* and *Z. h. luteus* to be observed in the
882 best available context. Moreover, any study of this type should focus on newly trapped
883 individuals where available rather than the use of archived museum skins. Newly acquired
884 tissue will provide for higher molecular weight DNA which will increase the diversity of
885 molecular techniques available for analysis, minimize the effects of curatorial (e.g.,
886 misidentification, mislabeling) and molecular error rates (Sefac et al. 2003, Yao et al. 2004), and
887 exclude genetic variation that may not be extant.

888 In summary, we conclude that the subspecies surveyed in this study do not warrant
889 synonymy, as has been proposed for *Z. h. preblei*, *Z. h. campestris*, and *Z. h. intermedius*; they
890 constitute distinct evolutionary lineages that merit separate management consideration, and those
891 populations facing demographic challenges (e.g., *Z. h. preblei*-North, *Z. h. preblei*-South) should
892 be afforded high conservation priority. The strong concurrence among patterns of mitochondrial
893 and nuclear DNA variation observed in this study suggests that formal recognition of the
894 relationships revealed will assist in preserving the potential for future evolutionary change
895 within and among these subspecies.

Pre-publication Copy

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 for provision of *Zapus* tissues. We would like to thank Alex Schubert, US Fish and Wildlife Service for providing *Z. princeps* tissues. 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 the second draft. The authors are enormously indebted to all of these scientists for their unselfish efforts to improve this presentation. The authors thank Robin Johnson and Abdel Bouberhan for laboratory assistance and Anne Henderson for assistance with editing. This project was jointly funded by the US Fish and Wildlife Service and the US Geological Survey-Leetown Science Center's Aquatic Ecology Branch.

Pre-publication Copy

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.

Avise JC (2004) Molecular markers, natural history, and evolution. 2nd edition. Sinauer Associates, sunderland, MA.

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.

Avise JC, Johns GC (1999) Proposal for a standardized temporal scheme of biological classification for extant species. *Proc. Natl. Acad. Sci. USA*, **96**, 7358–7363.

Avise JC, Walker D (2000) Abandon all species concepts? A response. *Conservation Genetics*, **1**, 77-80.

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, MA.

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

Bradley RD, Baker RJ (2001) A test of the genetic species concept: cytochrome-b sequences and mammals. *Journal of Mammalogy*, **82**, 960-973.

Pre-publication Copy

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 and 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.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescent time in phylogeographic studies. *Evolution*, **54**, 1839-1854.

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

Pre-publication Copy

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.

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.

Glenn TC, Stephan W, Braun MJ (1999) Effects of a population bottleneck on whooping crane mitochondrial DNA variation. *Conservation Biology*, **13**, 1097-1107.

Goldstein PZ, DeSalle R, Amato G, Vogler AP (2000) Conservation genetics at the species boundary. *Conservation Biology*, **14**, 120-131.

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.

Pre-publication Copy

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

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.

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

Hendry AP, Vamosi SM, Latham SJ, Heilbuth JC, Day T (2000) Questioning species realities. *Conservation Genetics*, **1**, 67-76.

Hoelzel AR, Fleischer RC, Campagna C, Le Boeuf B, Alvord G (2002) Impact of a population bottleneck on symmetry and genetic diversity in the northern elephant seal. *Journal of Evolutionary Biology*, **15,** 567-575.

Hudson RR, Turelli M (2003) Stochasticity overrules the 'three-times rule': genetic drift, genetic draft, and coalescent times for nuclear loci versus mitochondrial DNA. *Evolution*, **57,** 182-190.

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). Unpublished PhD Thesis: Indiana State University, Terre Haute, Indiana.

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

Pre-publication Copy

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 BL, 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*, In Press.

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.

Mitchell A, Mitter C, Regier JC (2000) More taxa or more characters revisited: combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). *Systematic Biology*, **49,** 202-224.

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. Sinmauer Associates Inc., Sunderland, MA.

Pre-publication Copy

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 and Evolution*, **9**, 373-375.

Moum T, Arnason U, Arnason E (2002) Mitochondrial DNA sequence evolution and phylogeny of the Atlantic Alcidae, including the extinct Great Auk (*Pinguinus impennis*). *Molecular Biology and Evolution*, **19,** 1434-1439.

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

Nichols R (2001) Gene trees and species trees are not the same. *Trends in Ecological Evolution,* **16**, 358-364.

Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M (2004) Genetic analyses from ancient DNA. *Annual Review of Genetics,* **38,** 1939-1943.

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.

Pennock DS, Dimmick WW (1997) Critique of the evolutionarily significant unit as a definition for "distinct population segments" under the U.S. *Endangered Species Act. Conservation Biology*, **11,** 611 619.

Pre-publication Copy

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics,* **14**, 817-818.

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.

Quattro JM, Greig TW, Coykendall DK, Bowen BW, Baldwin JD (2002) Genetic issues in aquatic species management: the shortnose sturgeon (*Acipenser brevirostrum*) in the southeastern United States. *Conservation Genetics*, **3,** 155-166.

Ramey RR, Liu HP, Epps CW, Carpenter LM, and 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.

Pre-publication Copy

Sefac KM, Payne RB, Sorenson MD (2003) Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. *The Auk*, **120,** 982-989**.**

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 FJ (1994) Biometry: The Principles and Practice of Statistics in Biological Research. Third Edition. W.H. Freeman & Co., New York NY.

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

Sorenson MD, Cooper A, Paxinos EE, Quinn TW, James HF, Olson SL, Fleischer RC (1999) Relationships of the extinct Moa-Nalos, flightless Hawaiian waterfowl, based on ancient DNA. *Proceedings of the Royal Society of London, Series B*, **266**, 2187-2193.

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, Massachusetts.

Swofford DL, Selander RB (1981) BIOSYS-1: A Fortran program for the comprehensive

Pre-publication Copy

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, Sing CF (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.

US 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.

US 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

Pre-publication Copy

in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology Notes,* **3**,638-640.

Waples R (1991) Genetic interactions between hatchery and wild salmonids: lessons from the pacific northwest. *Can. J. Fish. Aquat. Sci.*, **48**, (Suppl. 1): 124-133.

Waples R (1998) Evolutionarily significant units, distinct population segments, and the endangered species act: reply to Pennock and Dimmick. *Conservation Biology*, **12**, 718.

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

Willerslev E, Cooper A (2005) Ancient DNA. *Proceedings of the Biological Sciences B,* **272,** 3-16.

Yao Y-G, Bravi CM, Bandelt H-J (2004) A call for mtDNA data quality control in forensic science. *Forensic Science International***, 141,** 1-6.

Pre-publication Copy

Table 1. Collection information on specimens from four neighboring subspecies of *Zapus hudsonius* surveyed for mitochondrial and microsatellite DNA variation.

Subspecies	Collection abbreviation	Cluster[1] identification	Collection locality[2]	Collect year	N	Source[3]	Tissue Type[4]
Z. h. preblei	LCCO1	*Z. h. preblei*-North	N. F. Cache la Poudre River, Larimer Co., CO	1998	14	CDOW	Ear punch
	LCCO2		Stove Prairie Creek, Larimer Co., CO	1998	16	CDOW	Ear punch
	DCCO1	*Z. h. preblei*-South	East Plum Creek, Douglas Co., CO	1998	34	CDOW	Ear punch
	DCCO2		Indian Creek, Douglas Co., CO	1998	30	CDOW	Ear punch
	ECCO1		U.S. Air Force Academy, El Paso Co., CO	1998	22	CNHP	Ear punch
	ECCO2		Monument Creek, El Paso Co., CO	2005	26	CNHP	Ear punch
Z. h. intermedius	BRCSD	*Z. h. intermedius*-West	Columbia Road Reservoir, Brown Co., SD	2005	28	USGS	Ear/blood
	MCMN	*Z. h. intermedius*-East	Camp Ripley, Morrison Co., MN	1996-8	21	MSB	Ear/blood
Z. h. campestris	CCWY	*Z. h. campestris*	Beaver Creek, Crook Co., WY,	2005	30	USGS	Ear/blood
	CCSD		Iron and Willow Creeks, Custer and Pennington Cos., SD	2005	31	USGS	Ear/blood
Z. h. pallidus	BCSD	*Z. h. pallidus*	Cedar Creek, Bennett Co., SD	2005	15	USGS	Ear/blood
	KBCNE		North Channel Platte River, Kearney/Buffalo Cos., NE	2005	33	USGS	Ear/blood
Z. h. luteus[5]	SCNM	*Z. h. luteus*	Multiple sites, Sandoval Co., NM	1985	20	MSB	Liver
Z. princeps	NA	NA	Libby Creek, Albany Co., WY	2005	7	FWS	Ear/blood

[1]Clusters as identified by the program *structure* using microsatellite DNA allelic patterns.
[2]Detailed collection information available upon request from the corresponding author.
[3] Sources include Tanya Shenk, Colorado Department of Wildlife, Ft. Collins, CO (CDOW); Robert Schorr, CO National Heritage Program, Fort Collins, CO (CNHP); Paul Cryan, U.S. Geological Survey, Fort Collins Science Center, Fort Collins, CO (USGS); University of New Mexico Museum of Southwestern Biology, Albuquerque, NM (MSB); Alex Schubert, U.S. Fish and Wildlife Service, Cheyenne Field Office, Snowy Range, WY (FWS)..
[4]Samples labeled "Ear punch" were archived in 95% absolute ethanol. Samples labeled "Ear/blood" were collected during the summer of 2005 by Paul Cryan. Ear punches from these collections were preserved in 95% ethanol. Residual blood from the ear punch procedures was applied to FTA cards. Frozen liver tissue was obtained from the MSB.

Pre-publication Copy

Table 2. Sample size, allelic richness, number of private alleles, observed and expected heterozygosity, and estimates of F_{IS} observed for 12 collections of *Zapus hudsonius* representing four neighboring subspecies surveyed at 21 microsatellite DNA loci.

Subspecies	Collection	Cluster[1]	N	A^2	Private alleles	H_o	H_e	F_{IS}
Z. h. preblei			**142**	**5.3**	**[4]**	0.547	0.595	**0.081**
	LCCO1		(14)	3.4	(1)	0.507	0.502	-0.033
	LCCO2		(16)	3.3		0.536	0.528	-0.013
		Z. h. preblei - North	[30]	3.7	[1]	0.525	0.529	[0.008]
	DCCO1		(34)	3.8		0.538	0.535	-0.005
	DCCO2		(30)	3.7	(1)	0.538	0.540	0.007
	ECCO1		(22)	4.0	(1)	0.563	0.584	0.035
	ECCO2		(26)	3.7		0.586	0.559	-0.050
		Z. h. preblei - South	[112]	4.6	[2]	0.553	0.583	[0.052]
Z. h. campestris		*Z. h. campestris*	**[61]**	**7.0**	**[6]**	0.637	0.670	**0.049**
	CCWY		(30)	5.4	(2)	0.648	0.662	0.020
	CCSD		(31)	5.2	(2)	0.625	0.654	0.044
Z. h. intermedius			**49**	**9.4**	**21**	0.649	0.703	**0.078**
	BRCSD	*Z. h. intermedius* - West	[28]	5.7	[7]	0.619	0.637	[0.029]
	MCMN	*Z. h. intermedius* - East	[21]	7.0	[14]	0.687	0.735	[0.064]
Z. h. pallidus		*Z. h. pallidus*	**[48]**	**9.4**	**[38]**	0.752	0.790	**0.049**
	BCSD		(16)	6.5	(5)	0.738	0.757	0.012
	KBCNE		(32)	6.7	(13)	0.759	0.789	0.045
Z. h. luteus	**SCNM**	*Z. h .luteus*	**[20]**	**4.3**	**[8]**	0.576	0.623	**0.076**

[1]Clusters (k = 7) were determined using the program *structure*.
[2]Allelic richness

Pre-publication Copy

Table 3. Percentage of *Zapus hudsonius* individuals correctly assigned to one of $k = 3$ initial clusters (subsets) and $k = 7$ subsequent (within-subset) clusters identified by the program *structure* from a survey of 13 collections of mice representing five geographically proximal subspecies surveyed at 21 microsatellite loci. Individuals were assigned to subset or cluster based on the largest value of q. Average qMAX and percentage of *Zapus hudsonius* individuals correctly assigned to one of $k = 7$ clusters identified by the program *structure* based on the $q > 0.90$ criterion are also provided.

Subspecies	Initial subsets (k=3)			Mean *q*MAX	Subset	[A]		[B]			[C]		Cluster designation	Mean *q*MAX	Percentage assigned at *q* > 0.90
Collection	[A]	[B]	[C]		Cluster	1	2	3	4	5	6	7			
Z. h. preblei													***Z. h. preblei***	**0.97**	**100.0% (142)**
LCCO1	**1.00**				**1**	**1.00**									
LCCO2	**1.00**				**1**	**1.00**							**North**	**0.97**	**96.7% (29)**
DCCO1	**1.00**			**0.99**	**2**	**0.03**	**0.97**								
DCCO2	**1.00**				**2**		**1.00**								
ECCO1	**1.00**				**2**		**1.00**								
ECCO2	**1.00**				**2**		**1.00**						**South**	**0.98**	**96.4% (108)**
Z. h. campestris													***Z. h. campestris***	**0.96**	**90.2% (55)**
CCWY		**1.00**			**3**			**1.00**							
CCSD		**1.00**			**3**			**1.00**							
				0.96											
Z. h. intermedius													***Z. h. intermedius***	**0.98**	**93.8% (46)**
BRCSD		**1.00**			**4**				**1.00**				**West**	**0.95**	**89.3% (25)**
MCMN		**1.00**			**5**					**1.00**			**East**	**0.97**	**100.0% (21)**
Z. h. pallidus													***Z. h. pallidus***	**0.99**	**100.0% (48)**
BCSD			**1.00**		**6**						**1.00**				
KBCNE			**1.00**	**0.99**	**6**						**1.00**				
Z. h. luteus													***Z. h. luteus***	**0.99**	**100.0% (20)**
SCNM			**1.00**		**7**							**1.00**			

Pre-publication Copy

Table 4. Matrix of pairwise genetic distance (*D*a; (Nei et al. 1983) values by collection site for 21 microsatellite DNA loci surveyed in five neighboring subspecies of *Zapus hudsonius*.

Subspecies	***Z. h. preblei***						***Z. h. campestris***		***Z. h. intermedius***		***Z. h. pallidus***		***Z. h. luteus***
Cluster	**North**		**South**						**West**	**East**			
Collection	**LCCO1**	**LCCO2**	**DCCO1**	**DCCO2**	**ECCO1**	**ECCO2**	**CCWY**	**CCSD**	**BRCSD**	**MCMN**	**BCSD**	**KBCNE**	**SCNM**
LCCO1													
LCCO2	**0.0774**												
DCCO1	**0.1698**	**0.1706**											
DCCO2	**0.1976**	**0.2004**	**0.1225**										
ECCO1	**0.1675**	**0.1703**	**0.0767**	**0.1011**									
ECCO2	**0.1976**	**0.1830**	**0.0923**	**0.1147**	**0.0421**								
CCWY	**0.3471**	**0.3102**	**0.2770**	**0.3147**	**0.2769**	**0.2620**							
CCSD	**0.3633**	**0.3073**	**0.3021**	**0.3270**	**0.2871**	**0.2867**	**0.1019**						
BRCSD	**0.4718**	**0.4229**	**0.4304**	**0.4563**	**0.4336**	**0.4201**	**0.2663**	**0.2312**					
MCMN	**0.4475**	**0.4247**	**0.3895**	**0.4238**	**0.3781**	**0.3726**	**0.2429**	**0.2176**	**0.1738**				
BCSD	**0.5920**	**0.5745**	**0.5437**	**0.5396**	**0.5117**	**0.5220**	**0.5040**	**0.5075**	**0.4903**	**0.4334**			
KBCNE	**0.5629**	**0.5452**	**0.5091**	**0.5193**	**0.4940**	**0.4954**	**0.4469**	**0.4313**	**0.4311**	**0.3876**	**0.1669**		
SCNM	**0.7323**	**0.7255**	**0.6824**	**0.6623**	**0.6679**	**0.6961**	**0.6639**	**0.6430**	**0.6451**	**0.5743**	**0.4617**	**0.4695**	

Pre-publication Copy

Table 5. F_{ST} values (Wier and Cockerham 1984) generated from a survey of 21 microsatellite loci for k = 7 clusters identified by the program *structure* among *Zapus hudsonius* collections representing five geographically proximal subspecies. See Table 1 for explanation of clusters. All F_{ST} estimates were statistically significant from zero ($\alpha = 0.05$, $P < 0.003$) after 15,000 permutations.

Subspecies/Cluster	***Z. h. preblei***		***Z. h. campestris***	***Z. h. intermedius***		***Z. h. pallidus***	***Z. h. luteus***
	North	**South**		**West**	**East**		
***Z. h. preblei* - North**							
***Z. h. preblei* - South**	0.1099						
Z. h. campestris	0.1760	0.1329					
***Z. h. intermedius* - West**	0.2724	0.2512	0.1373				
***Z. h. intermedius* - East**	0.2291	0.2008	0.1036	0.0663			
Z. h. pallidus	0.2489	0.2373	0.1806	0.1861	0.1356		
Z. h. luteus	0.3722	0.3512	0.3015	0.3154	0.2517	0.1816	

Pre-publication Copy

Table 6. 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 ($\alpha = 0.05$, $P < 0.008$) after 6,000 permutations.

Subspecies	***Z. h. preblei***	***Z. h. campestris***	***Z. h. intermedius***	***Z. h. pallidus***	***Z. h. luteus***
Z. h. preblei					
Z. h. campestris	0.1277				
Z. h. intermedius	0.2038	0.1068			
Z. h. pallidus	0.2321	0.1806	0.1546		
Z. h. luteus	0.3412	0.3015	0.2701	0.1816	

Pre-publication Copy

Table 7. 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	**132**	**3**	**4**	**0.0030**	**119**	**12**	**17**	**0.0011**	**114**	**15**	**21**	**0.0016**
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**

Pre-publication Copy

Table 8. Pair-wise average net (adjusted) estimates of Kimura 2-parameter divergence for mitochondrial DNA control region (above the diagonal) between five subspecies of *Zapus hudsonius* and a collection of *Zapus princeps*. Estimates of net Tamura-Nei genetic divergence for combined mitochondrial DNA cytochrome *b* and control region are provided below (below the diagonal).

Subspecies	***Z. h. preblei***	***Z. h. campestris***	***Z. h. intermedius***	***Z. h. pallidus***	***Z. h. luteus***	***Z. princeps***
Z. h. preblei		**0.0103**	**0.0057**	**0.0501**	**0.0508**	**0.1072**
Z. h. campestris	**0.0094**		**0.0029**	**0.0558**	**0.0563**	**0.1077**
Z. h. intermedius	**0.0067**	**0.0024**		**0.0499**	**0.0497**	**0.1102**
Z. h. pallidus	**0.0622**	**0.0665**	**0.0635**		**0.0055**	**0.0942**
Z. h. luteus	**0.0613**	**0.0655**	**0.0623**	**0.0117**		**0.0971**
Z. princeps	**0.1762**	**0.1832**	**0.1812**	**0.1572**	**0.1630**	

Pre-publication Copy

Table 9. Comparison of Φ_{ST} and F_{ST} estimates observed in five neighboring subspecies of *Zapus hudsonius* sequenced at 372 base pairs of the mitochondrial DNA control region and at 1006 base pairs of the cytochrome b gene. Φ_{ST} and F_{ST}, while similar in that they constitute ratios resulting from the comparison of variation between subspecies, measure the effects of different evolutionary processes. Φ_{ST} measures the evolutionary relationship between haplotypes (i.e., mutation), as well as, haplotype frequency differences (i.e., genetic drift) between subspecies. F_{ST} quantifies only haplotype frequency differences between subspecies.

Subspecies comparison	**Control region**		**Cytochrome b**	
	Φ_{ST}	F_{ST}	Φ_{ST}	F_{ST}
Z. h. preblei - Z. h. intermedius - Z. h. campestris - Z. h. pallidus - Z. h. luteus	0.96	0.37	0.96	0.33
Z. h. preblei - Z. h. intermedius - Z. h. campestris	0.75	0.33	0.84	0.30
Z. h. preblei - Z. h. intermedius	0.69	0.22	0.82	0.21
Z. h. preblei - Z. h. campestris	0.81	0.39	0.90	0.36
Z. h. preblei - Z. h. pallidus	0.93	0.29	0.98	0.23
Z. h. preblei - Z. h. luteus	0.95	0.43	0.99	0.50
Z. h. intermedius - Z. h. campestris	0.59	0.38	0.53	0.31
Z. h. intermedius - Z. h. pallidus	0.91	0.25	0.96	0.15
Z. h. intermedius - Z. h. luteus	0.94	0.42	0.97	0.47
Z. h. campestris - Z. h. pallidus	0.93	0.45	0.98	0.32
Z. h. campestris - Z. h. luteus	0.97	0.64	0.99	0.64
Z. h. pallidus - Z. h. luteus	0.57	0.51	0.87	0.49

Pre-publication Copy

Figure Legends

Figure 1. Generalized collection sites (N=13) 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 13 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 (subset) was subsequently analyzed for within-cluster structure. The sequential analysis further subdivided [A] into two clusters, [B] was found to contain three clusters, and [C] consisted of two additional clusters for a total of seven clusters (1-7). 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.

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. Solid circles represent the clusters inferred by the program *structure*. The dashed circles represent the grouping of clusters consistent with the subspecies *Z. h. intermedius* and *Z. h. preblei*. *Z. h. campestris*, *Z. h. pallidus*, and *Z. h. luteus* (single collection) were recovered as single clusters by *structure*.

Figure 4. Graphical depiction of results from the multidimensional scaling analysis of pairwise 1-*D*a values among 13 collections of *Z. hudsonius* surveyed at 21 microsatellite DNA loci. Solid circles represent the clusters inferred by the program *structure*. The dashed circles represent the grouping of clusters consistent with the subspecies *Z. h. intermedius* and *Z. h. preblei*. *Z. h. campestris*, *Z. h. pallidus*, and *Z. h. luteus* (single collection) were recovered as single clusters by *structure*.

Figure 5. *Zapus hudsonius* haplotype networks resulting from the statistical parsimony analysis of the mitochondrial DNA control region 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.

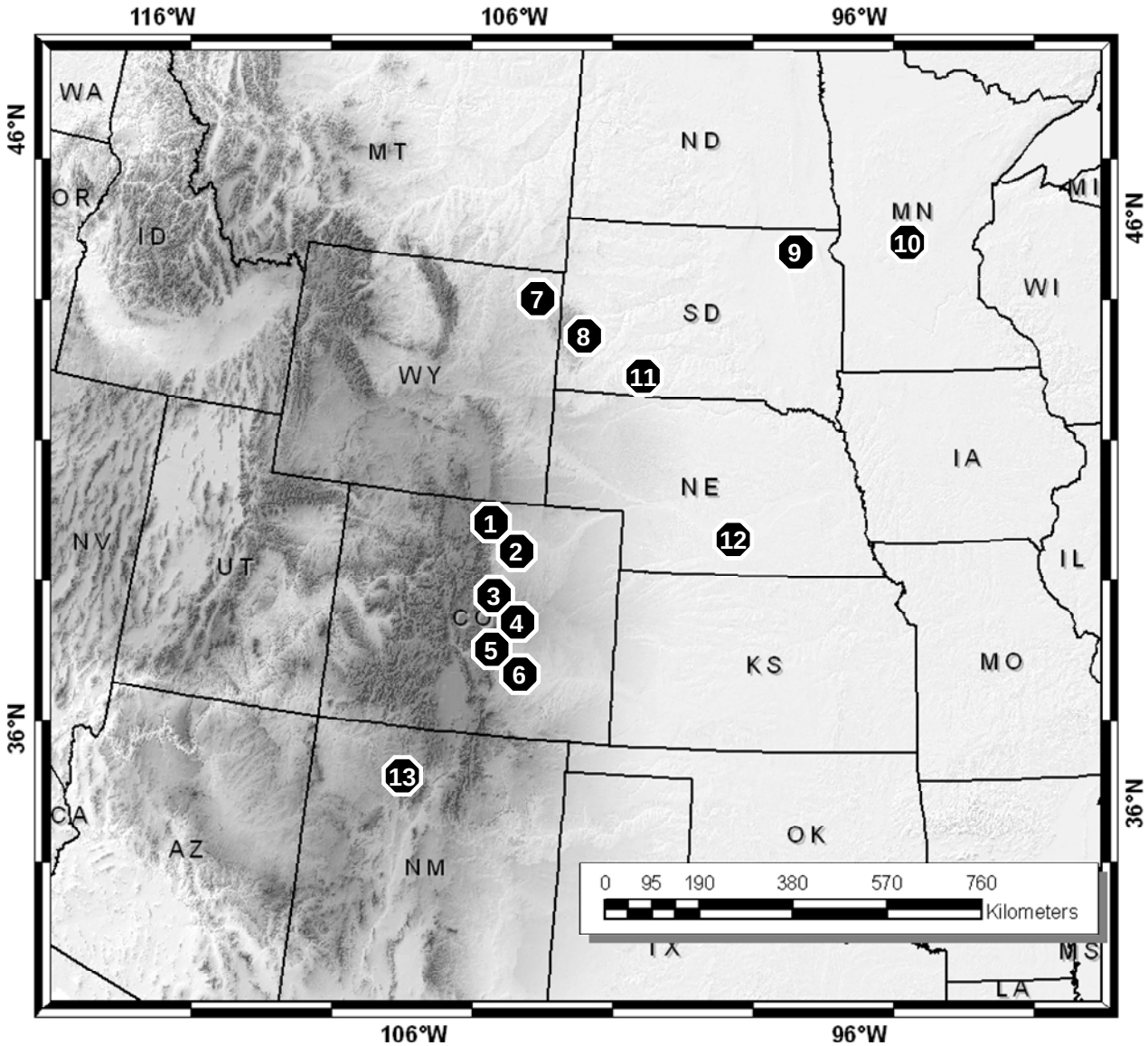
Figure 6. *Zapus hudsonius* haplotype networks resulting from the statistical parsimony analysis of the cyt *b* data set. The set of haplotypes belonging to each subspecies of *Z. hudsonius* are outlined by boxes. No haplotypes were shared among subspecies. Haplotypes identified as the root of each network 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.

Figure 7. Phylogenetic hypotheses for five *Z. hudsonius* subspecies based upon combined mitochondrial DNA data partitions and A) unweighted parsimony analysis (strict consensus of

Pre-publication Copy

4050 most parsimonious trees, 205 steps, CI = 0.7171) and B) maximum likelihood analysis using best-fit model of nucleotide substitution (-ln = 3682.52441). Numbers above branches in A are bootstrap proportions from maximum parsimony and maximum likelihood analyses, respectively, and italicized numbers below branches are decay indices. Stars indicate nodes that are of particular interest to this study and received high support.

Pre-publication Copy



Zapus hudsonius **Collection Sites**

Z. h. preblei
1. **LCC01**
2. **LCC02**
3. **DCC01**
4. **DCC02**
5. **ECC01**
6. **ECC02**

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

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

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

Z. h. luteus
13. **SCNM**

Figure 1

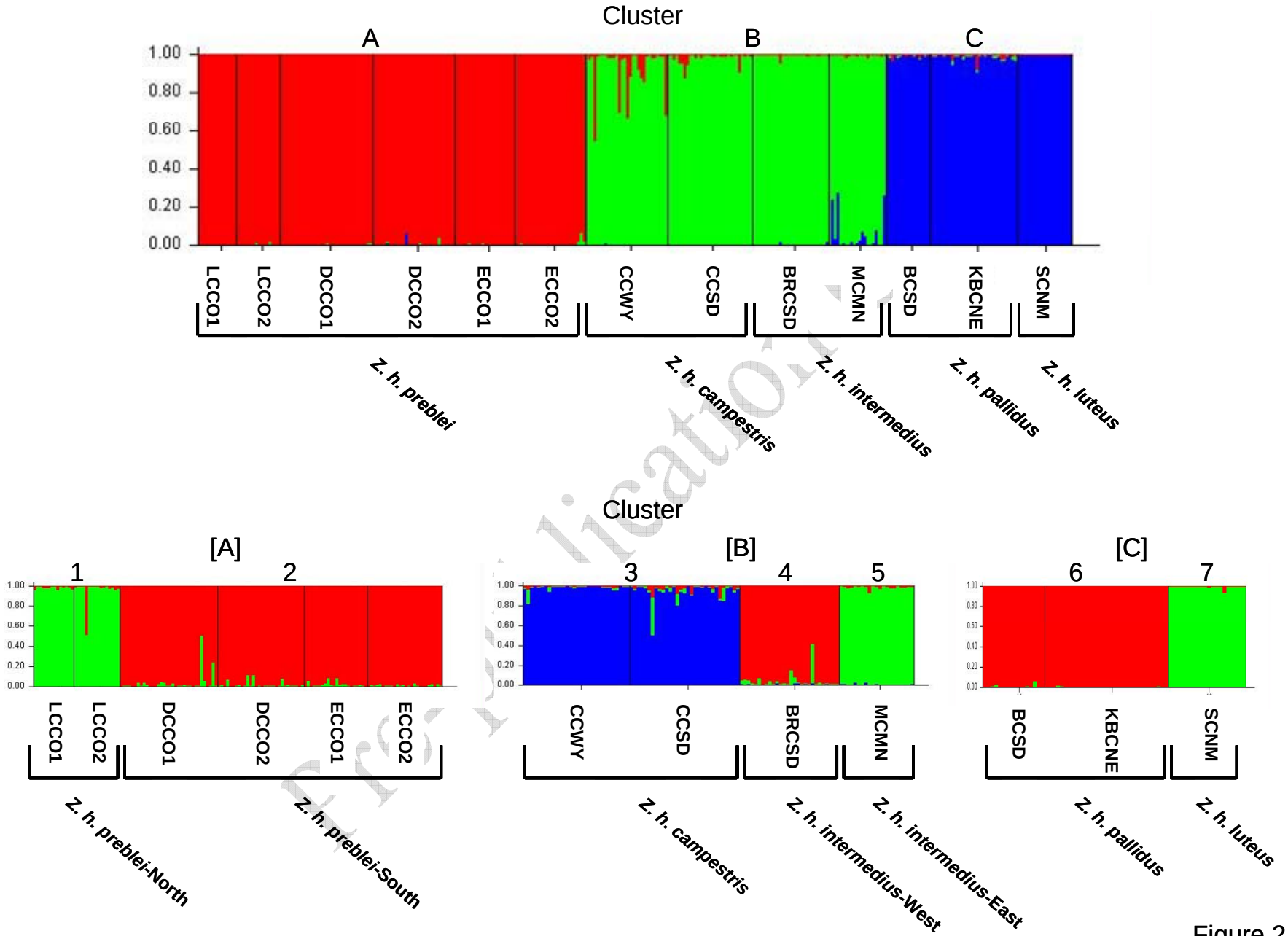
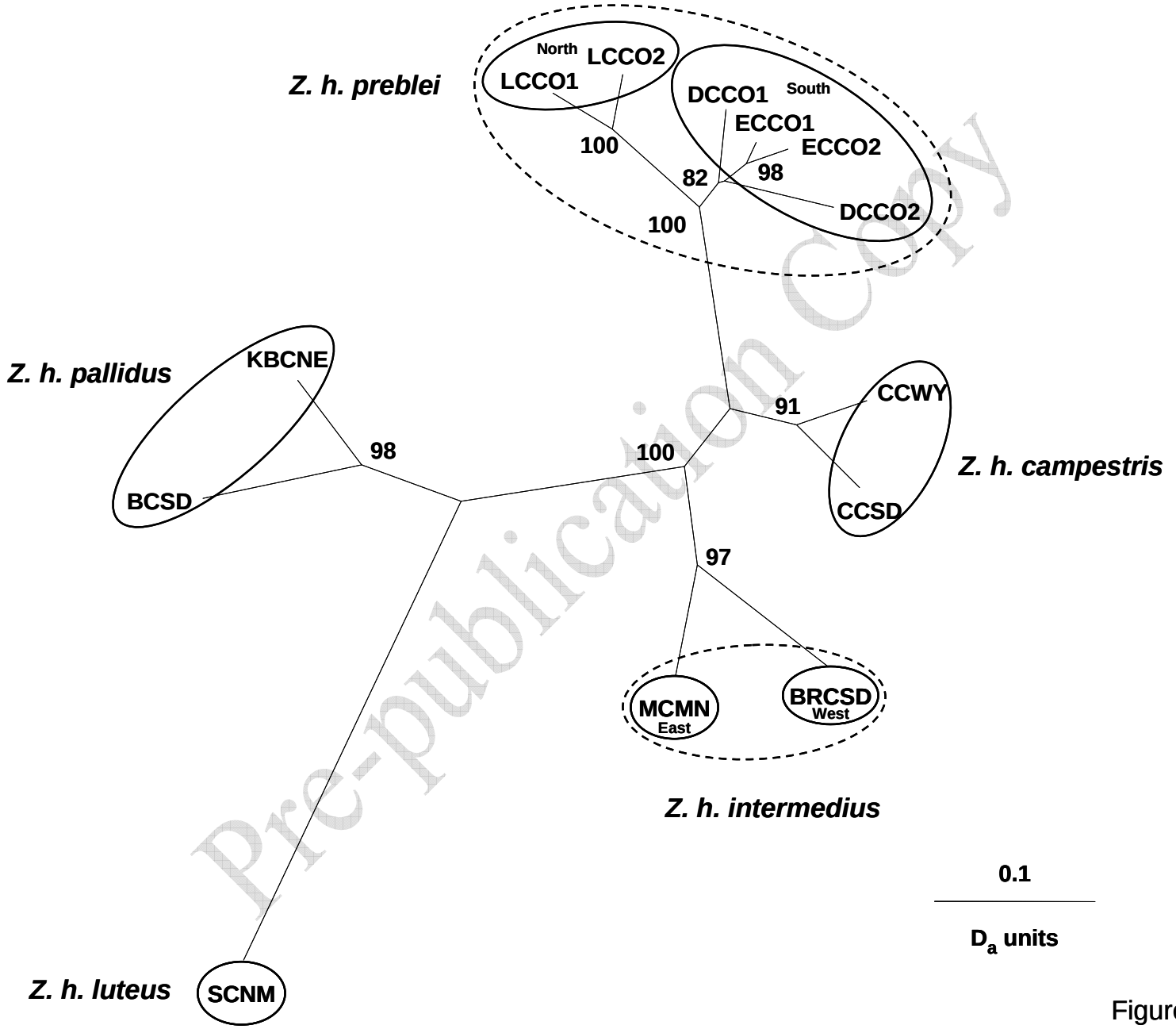
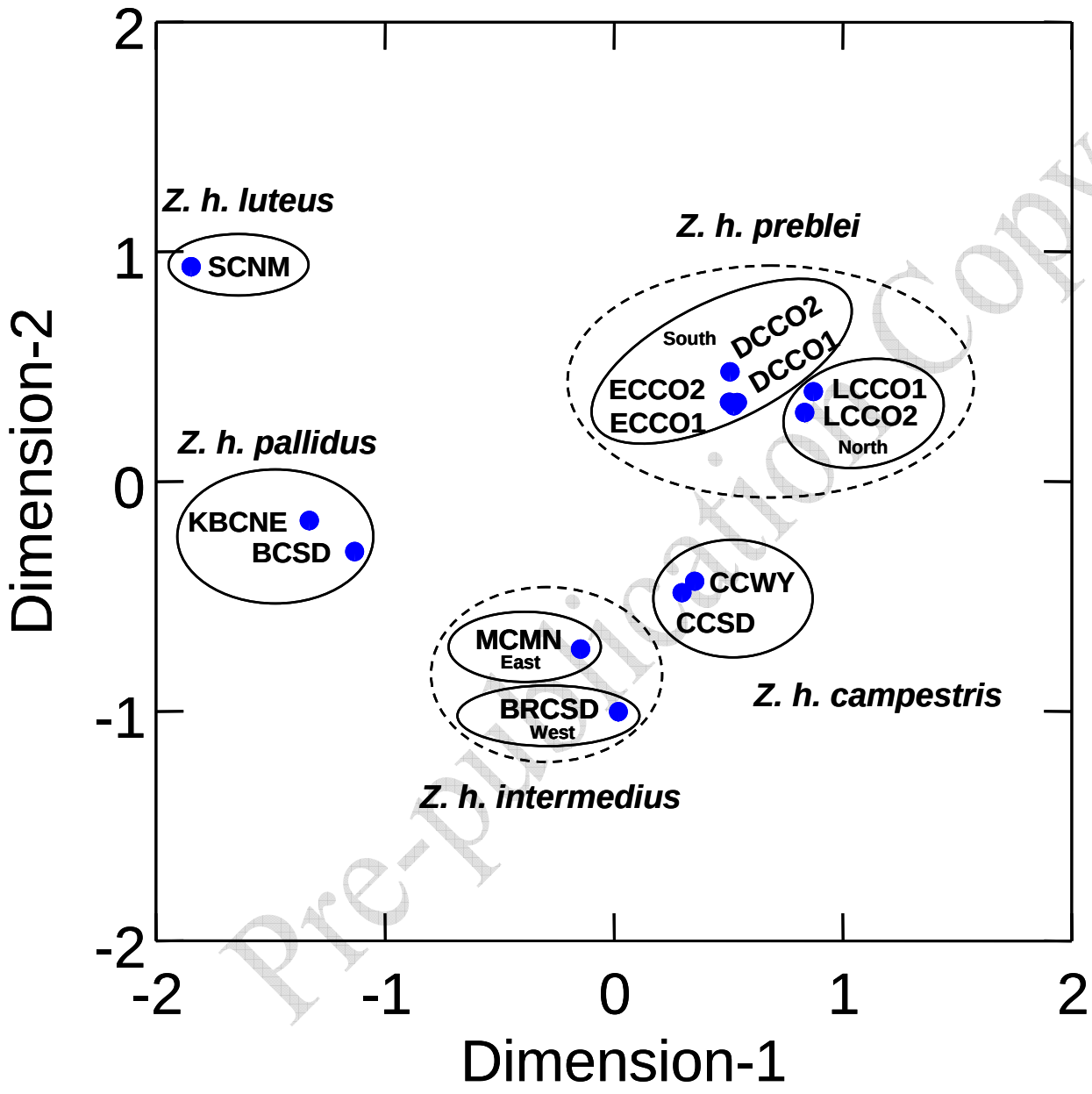

Cluster
A
B
C
1.00
0.80
0.60
0.40
0.20
0.00
LCCO1
LCCO2
DCCO1
DCCO2
ECCO1
ECCO2
CCWY
CCSD
BRCSD
MCMN
BCSD
KBCNE
SCNM
Z. h. preblei
Z. h. campestris
Z. h. intermedius
Z. h. pallidus
Z. h. luteus
Cluster
[A]
[B]
[C]
1
2
3
4
5
6
7
Z. h. preblei-North
Z. h. preblei-South
Z. h. campestris
Z. h. intermedius-West
Z. h. intermedius-East
Z. h. pallidus
Z. h. luteus

Figure 2



Pre-publication Copy

Figure 3



Pre-publication Copy

Figure 4

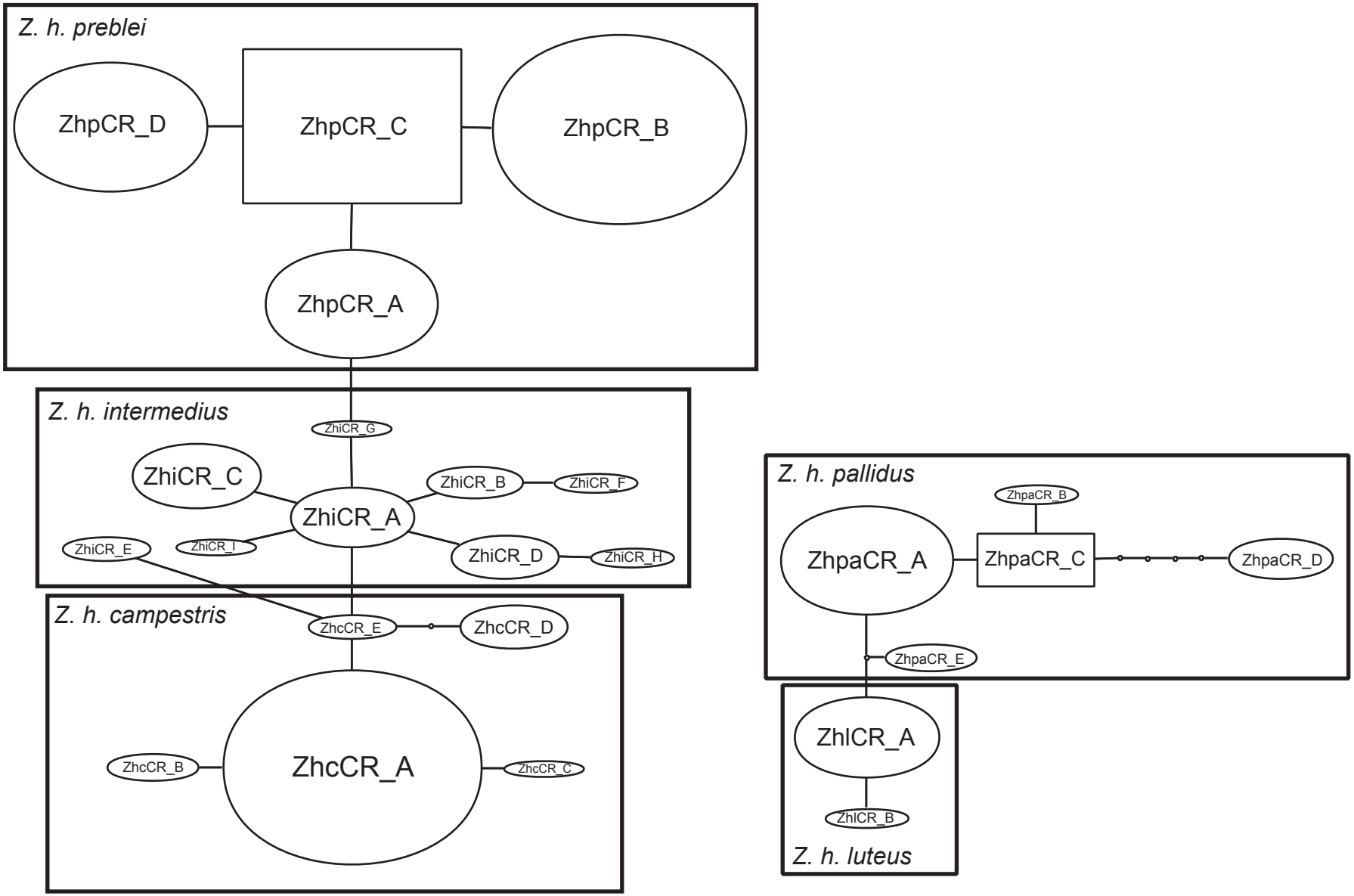

Z. h. preblei
ZhpCR_D
ZhpCR_C
ZhpCR_B
ZhpCR_A
Z. h. intermedius
ZhiCR_G
ZhiCR_C
ZhiCR_B
ZhiCR_F
ZhiCR_A
ZhiCR_E
ZhiCR_I
ZhiCR_D
ZhiCR_H
Z. h. campestris
ZhcCR_E
ZhcCR_D
ZhcCR_B
ZhcCR_A
ZhcCR_C
Z. h. pallidus
ZhpaCR_B
ZhpaCR_A
ZhpaCR_C
ZhpaCR_D
ZhpaCR_E
ZhlCR_A
ZhlCR_B
Z. h. luteus

Figure 5

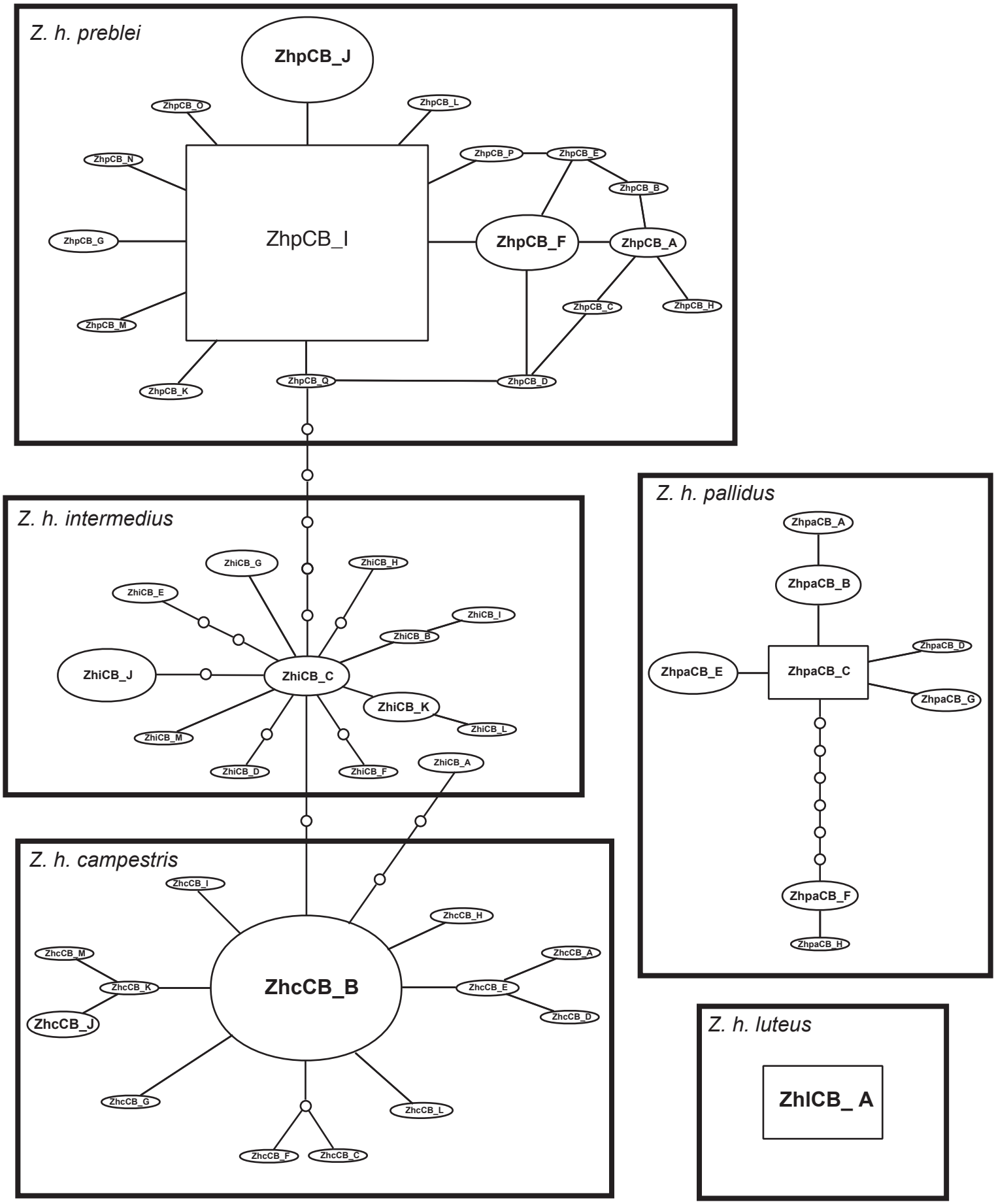

Z. h. preblei
ZhpCB_J
ZhpCB_O
ZhpCB_L
ZhpCB_N
ZhpCB_P
ZhpCB_E
ZhpCB_B
ZhpCB_G
ZhpCB_I
ZhpCB_F
ZhpCB_A
ZhpCB_C
ZhpCB_H
ZhpCB_M
ZhpCB_K
ZhpCB_Q
ZhpCB_D
Z. h. intermedius
ZhiCB_G
ZhiCB_H
ZhiCB_E
ZhiCB_I
ZhiCB_B
ZhiCB_J
ZhiCB_C
ZhiCB_K
ZhiCB_L
ZhiCB_M
ZhiCB_D
ZhiCB_F
ZhiCB_A
Z. h. pallidus
ZhpaCB_A
ZhpaCB_B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

Figure 6

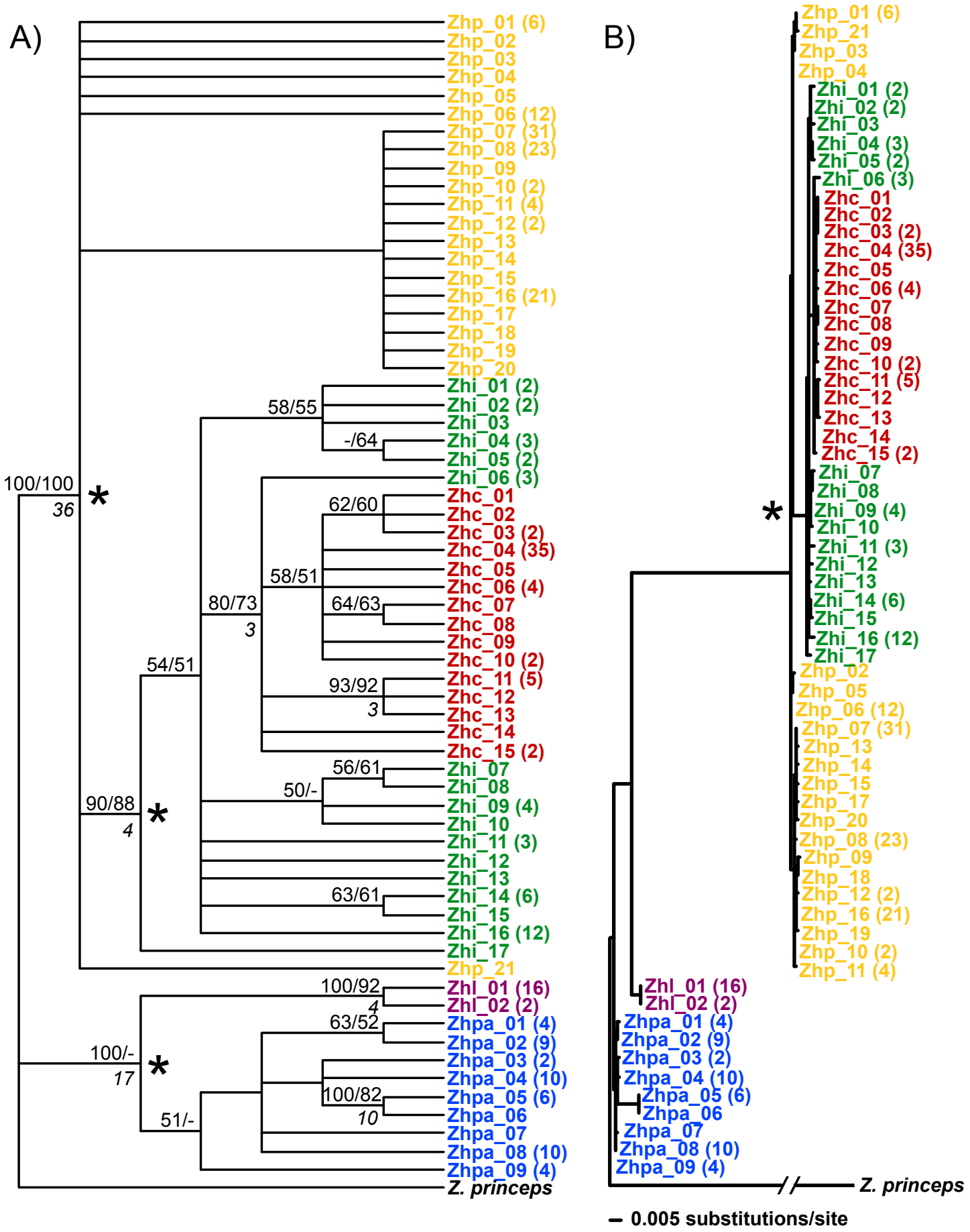

A)
Zhp_01 (6)
Zhp_02
Zhp_03
Zhp_04
Zhp_05
Zhp_06 (12)
Zhp_07 (31)
Zhp_08 (23)
Zhp_09
Zhp_10 (2)
Zhp_11 (4)
Zhp_12 (2)
Zhp_13
Zhp_14
Zhp_15
Zhp_16 (21)
Zhp_17
Zhp_18
Zhp_19
Zhp_20
Zhi_01 (2)
Zhi_02 (2)
Zhi_03
Zhi_04 (3)
Zhi_05 (2)
Zhi_06 (3)
Zhc_01
Zhc_02
Zhc_03 (2)
Zhc_04 (35)
Zhc_05
Zhc_06 (4)
Zhc_07
Zhc_08
Zhc_09
Zhc_10 (2)
Zhc_11 (5)
Zhc_12
Zhc_13
Zhc_14
Zhc_15 (2)
Zhi_07
Zhi_08
Zhi_09 (4)
Zhi_10
Zhi_11 (3)
Zhi_12
Zhi_13
Zhi_14 (6)
Zhi_15
Zhi_16 (12)
Zhi_17
Zhp_21
Zhl_01 (16)
Zhl_02 (2)
Zhpa_01 (4)
Zhpa_02 (9)
Zhpa_03 (2)
Zhpa_04 (10)
Zhpa_05 (6)
Zhpa_06
Zhpa_07
Zhpa_08 (10)
Zhpa_09 (4)
Z. princeps
100/100
36
*
58/55
-/64
62/60
58/51
80/73
3
64/63
54/51
93/92
3
56/61
50/-
90/88
4
*
100/92
4
63/52
100/-
17
*
100/82
10
51/-
B)
Zhp_01 (6)
Zhp_21
Zhp_03
Zhp_04
Zhi_01 (2)
Zhi_02 (2)
Zhi_03
Zhi_04 (3)
Zhi_05 (2)
Zhi_06 (3)
Zhc_01
Zhc_02
Zhc_03 (2)
Zhc_04 (35)
Zhc_05
Zhc_06 (4)
Zhc_07
Zhc_08
Zhc_09
Zhc_10 (2)
Zhc_11 (5)
Zhc_12
Zhc_13
Zhc_14
Zhc_15 (2)
Zhi_07
Zhi_08
Zhi_09 (4)
Zhi_10
Zhi_11 (3)
Zhi_12
Zhi_13
Zhi_14 (6)
Zhi_15
Zhi_16 (12)
Zhi_17
Zhp_02
Zhp_05
Zhp_06 (12)
Zhp_07 (31)
Zhp_13
Zhp_14
Zhp_15
Zhp_17
Zhp_20
Zhp_08 (23)
Zhp_09
Zhp_18
Zhp_12 (2)
Zhp_16 (21)
Zhp_19
Zhp_10 (2)
Zhp_11 (4)
*
Zhl_01 (16)
Zhl_02 (2)
Zhpa_01 (4)
Zhpa_02 (9)
Zhpa_03 (2)
Zhpa_04 (10)
Zhpa_05 (6)
Zhpa_06
Zhpa_07
Zhpa_08 (10)
Zhpa_09 (4)
Z. princeps
0.005 substitutions/site

Figure 7

Appendix A. Microsatellite DNA markers, the total number of alleles observed, and the range of amplified products in the survey of 300 *Zapus hudsonius*, and associated references.

Microsatellite locus	Alleles observed	Size range	Reference
Z.7	30	156-179	Ramey et al. 2005
Z.20	19	103-141	Ramey et al. 2005
Z.26	16	138-174	Ramey et al. 2005
Z.48	15	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	12	149-195	Vignieri et al. 2003
*Ztri*19	7	174-204	Vignieri et al. 2003
*Ztri*24	13	151-199	Vignieri et al. 2003
ZhuC3	16	204-264	King et al. In Press
*Zhu*C6	9	100-144	King et al. In Press
*Zhu*C12	8	96-124	King et al. In Press
*Zhu*C104	9	222-254	King et al. In Press
*Zhu*C119	17	207-263	King et al. In Press
*Zhu*C120	7	145-169	King et al. In Press
*Zhu*C129	10	200-236	King et al. In Press
*Zhu*C130	7	258-286	King et al. In Press
*Zhu*D107	13	213-261	King et al. In Press
*Zhu*D108	10	138-176	King et al. In Press
*Zhu*D109	13	133-177	King et al. In Press
*Zhu*D122	18	201-285	King et al. In Press

Pre-publication Copy

Appendix B. *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. pallidis*		*Z. h. luteus*
Haplotype	LCCO1	LCCO2	DCCO1	DCCO2	ECCO1	ECCO2	BRCSD	MCMN	CCWY	CCSD	BENCSD	KBCNE	SCNM
N	12	16	32	26	19	26	26	21	29	31	16	31	19
*Zhp*CR_A	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

Pre-publication Copy

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

	Z. hudsonius preblei						*Z. h. intermedius*		*Z. h. campestris*		*Z. h. pallidis*		*Z. h. luteus*
Haplotype	LCCO1	LCCO2	DCCO1	DCCO2	ECCO1	ECCO2	BRCSD	MCMN	CCWY	CCSD	BCSD	KBCNE	SCNM
N	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		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							
*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				

Pre-publication Copy

*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

Pre-publication C