High variation of mitochondrial DNA diversity as compared to nuclear microsatellites in mammalian populations

Takashi Saitoh

Field Science Center, Hokkaido University, Sapporo, Japan

Correspondence
Takashi Saitoh, Field Science Center, Hokkaido University, Kita-11, Nishi-10, Sapporo 060-0811, Japan.
Email: tsaitoh@fsc.hokudai.ac.jp

Funding information
JSPS KAKENHI, Grant/Award Number: 19H03005; The Ecological Society of Japan: The ESJ award 2015

Abstract
The effective gene number (the number of genes that can be inherited) of mitochondrial DNA (mtDNA) is one-fourth of that of nuclear DNA (ncDNA) in idealized populations. Therefore, mtDNA haplotype diversity ($h$) is predicted to be lower than ncDNA heterozygosity ($H_E$) because of the higher effect of genetic drift on mtDNA. This prediction has not yet been systematically tested. To this end, in this study, published data for 739 populations of 108 mammalian species (66 terrestrial and 42 marine species) revealed the following patterns: (a) $h$ was higher than $H_E$ in 54.9% of populations, (b) the variance of $h$ (0.097) was significantly higher than that of $H_E$ (0.018) and (c) the frequency distribution of $h$ differed between terrestrial and marine species. The terrestrial species exhibited a U-shaped distribution, whereas the marine species exhibited a right triangle shape. $H_E$ showed a unimodal distribution for both groups. (d) The mean of $H_E$ was similar between the terrestrial (0.668) and marine (0.672) species, whereas the mean of $h$ was significantly lower for the terrestrial species (0.578) than for the marine species (0.740). Two hypotheses were considered to explain the above-described patterns, one of which was based on the higher mutation rates of mtDNA, while the other was based on a nested subpopulation structure in which an ncDNA-based population includes several mtDNA-based subpopulations. Herein, the plausibility of these two hypotheses was discussed with a focus on the higher intraspecific variation of $h$.

KEYWORDS
effective gene number, haplotype diversity, heterozygosity, mutation rate, population structure

1 | INTRODUCTION

Genetic diversity is a source of adaptive traits in dynamically changing environments, thus contributing to higher individual fitness (Allendorf & Leary, 1986; Frankham, Ballow, & Briscoe, 2002; Reed & Frankham, 2003) and has been used as an indicator of the sustainability of populations (e.g., Frankham et al., 2002; Spielman, Brook, & Frankham, 2004). Further, genetic diversity may have important consequences at higher levels of ecological systems (e.g., community and ecosystem; Hughes, Inouye, Johnson, Underwood, & Vellend, 2008). Genetic diversity is, therefore, considered to be essential for the
sustenance of the earth system, and its degradation threatens the resilience of the earth system (Steffen et al., 2015). Leigh, Hendry, Vázquez Domínguez, and Friesen (2019) estimated a 5.4–6.5% decline in the population-level genetic diversity of wild organisms since the industrial revolution, and Miraldo et al. (2016) reported that the habitats affected by humans to a greater extent exhibit lower genetic diversity than that of wilder regions. The heterozygosity of nuclear DNA (ncDNA) is lower in threatened taxa than in related non-threatened taxa (Spielman et al., 2004) and is correlated with body size, which is a proxy of population size in mammals (Doyle, Hackin, Willoughby, Sundaram, & DeWoody, 2015; Wooten & Smith, 1985). Similar correlations describing mitochondrial DNA (mtDNA) diversity with a proxy of population size in mammals have also been reported (Mulligan, Kitchen, & Miyamoto, 2006; Piganeau & Eyre-Walker, 2009; Sato et al., 2017).

Some researchers have, however, questioned the relationship between population size and genetic diversity (e.g., Amos & Balmford, 2001; Wooten & Smith, 1985). Bazin, Glèmin, and Galtier (2006) showed that mtDNA diversity is relatively constant across species with different census populations sizes, whereas ncDNA diversity increases with the increase of population sizes. Pedreschi et al. (2018) did not find a consistent pattern of mtDNA diversity in relation to population size across species. Nabholz, Mauffrey, Bazin, and Glémin (2008); Nabholz, Glémin, and Galtier (2009) concluded that mtDNA diversity is essentially unpredictable. To address these inconsistent reports on the relationship between population size and genetic diversity, a deeper understanding of how genetic diversity is determined in wild populations is required.

Genetic diversity within a population is determined by selection, genetic drift, mutation and immigration (Ellegren & Galtier, 2016). For idealized finite populations, in which only genetic drift drives allele frequency changes, genetic diversity steadily decays in inverse proportion to the effective population size \( N_e \). Therefore, higher genetic diversity is predicted in a population with a larger \( N_e \).

The effects of \( N_e \) vary depending on the inheritance system of genes. Birky, Maruyama, and Fuerst (1983) proposed the concept of the effective gene number to clarify the effects of \( N_e \) based on the inheritance system. The effective gene number represents the number of genes that can be inherited, differing from \( N_e \). For example, the effective number of autosomal nuclear genes is a doubled \( N_e \) for diploids, whereas the effective gene number of maternally inherited organelle genes is equal to the number of breeding females (Birky et al., 1983). Therefore, the effective gene number differs between mtDNA and ncDNA even in the same population. In the populations with diploid ncDNA genes and completely maternally inherited mtDNA, the effective gene number for ncDNA is four times greater than that of the mtDNA, if the breeding sex ratio is even.

Birky et al. (1983) predicted that mtDNA diversity should be lower than ncDNA diversity in idealized populations because of the lower effective gene number of mtDNA and tested their theoretical prediction in a small number of empirical studies (Birky et al., 1983; Birky, Fuerst, & Maruyama, 1989). The diversity of mtDNA is substantially higher than that of ncDNA in a population of pocket gophers, Geomys pinetis (Avise, Giblin-Davidson, Laerm, Patton, & Lansman, 1979); and genetic diversity is nearly the same for the mtDNA and ncDNA in both human and Drosophila populations (Nei, 1983). Nei (1983) and Birky et al. (1989) attributed the divergence from the theoretical prediction to the higher mutation rate of mtDNA.

However, a standard pattern cannot be inferred from such a small number of examples, and alternative was not considered, although other factors (e.g., immigration) could also act as sources of genetic diversity.

The concept of the effective gene number has drawn the attention of various researchers. The key paper by Birky et al. (1983) has been cited by 520 articles in the “Web of Science” database as of August 2020, and various datasets of mtDNA and ncDNA diversity have accumulated for wild populations since the works of Nei (1983) and Birky et al. (1989). In light of the increasing body of available data on genetic diversity of populations, exploring the relationship between mtDNA and ncDNA diversity and testing the theoretical prediction proposed by Birky et al. (1983) are instrumental for advancing the field of population genetics.

In this study, the theoretical prediction proposed by Birky et al. (1983) was tested based on the meta-analysis of published data on mtDNA and ncDNA diversity. Mammalian populations were the main focus of the current meta-analysis, as substantial datasets on their genetic diversity are available. The empirical relationship between mtDNA and ncDNA diversity greatly differed from the theoretical prediction, which had assumed idealized populations. The effects of mutation rates and gene flow, as the sources of new alleles for ncDNA or mtDNA haplotypes, may explain the empirical patterns of genetic diversity. Comparing the plausibility of the two hypotheses based on mutation rates and gene flow, respectively, provides novel insights into the determinants of genetic diversity.

2 | MATERIALS AND METHODS

2.1 | Datasets

Haplotype diversity \( (h) \) and expected heterozygosity \( (H_{E}) \) of the same populations were used to compare the
TABLE 1  The range of genetic diversity (ncDNA heterozygosity and mtDNA haplotype diversity) for each of 108 mammalian species. “n,” “Min” and “Max” indicate the number of populations investigated, the minimum value and the maximum value of the genetic diversity, respectively.

| ID | Species Common name | Latin name | Population n | ncDNA Heterozygosity Min | Max | mtDNA Haplotype diversity Min | Max |
|----|---------------------|------------|--------------|--------------------------|-----|-------------------------------|-----|
| 1  | Cheetah             | Acinonyx jubatus | 2             | 0.674                    | 0.698 | 0.551                      | 0.828 |
| 2  | Impala              | Aepyceros melampus | 8             | 0.680                    | 0.750 | 0.820                      | 1.000 |
| 3  | Red panda           | Ailuropoda fulgens | 4             | 0.634                    | 0.732 | 0.643                      | 0.942 |
| 4  | Giant panda         | Ailuropoda melanoleuca | 4             | 0.486                    | 0.610 | 0.736                      | 0.926 |
| 5  | Arctic fox          | Alopex lagopus   | 2             | 0.620                    | 0.836 | 0.650                      | 0.800 |
| 6  | European water vole | Arvicola amphibius | 4             | 0.750                    | 0.810 | 0.557                      | 0.786 |
| 7  | Indian hog deer     | Axis porcinus   | 2             | 0.543                    | 0.775 | 0.660                      | 0.831 |
| 8  | African golden wolf | Canis anthus    | 1             | 0.830                    | 0.830 | 0.907                      | 0.907 |
| 9  | Golden jackal       | Canis aureus    | 5             | 0.375                    | 0.611 | 0.000                      | 0.509 |
| 10 | Wolf                | Canis lupus     | 10            | 0.479                    | 0.782 | 0.000                      | 0.870 |
| 11 | European roe deer   | Capreolus capreolus | 30            | 0.530                    | 0.790 | 0.069                      | 0.970 |
| 12 | Eurasian beaver     | Castor fiber    | 2             | 0.514                    | 0.530 | 0.745                      | 0.761 |
| 13 | Red deer            | Cervus elaphus  | 44            | 0.405                    | 0.900 | 0.000                      | 0.914 |
| 14 | Sika deer           | Cervus nippon   | 7             | 0.075                    | 0.647 | 0.000                      | 0.791 |
| 15 | European hamster    | Cricetus cricetus | 8             | 0.111                    | 0.786 | 0.000                      | 0.816 |
| 16 | Southern tuco-tuco  | Ctenomys australis | 1             | 0.540                    | 0.540 | 0.700                      | 0.700 |
| 17 | Magellanic tuco-tuco| Ctenomys magellanicus | 2             | 0.570                    | 0.705 | 0.632                      | 0.715 |
| 18 | Talas tuco-tuco     | Ctenomys talaraum | 1             | 0.620                    | 0.620 | 0.820                      | 0.820 |
| 19 | Tiger quoll         | Dasyurus maculatus | 4             | 0.490                    | 0.670 | 0.257                      | 0.745 |
| 20 | Northern collared lemming | Dicrostonyx groenlandicus | 11         | 0.680                    | 0.860 | 0.000                      | 0.780 |
| 21 | Tufted deer         | Elaphodus cephalophus | 2             | 0.804                    | 0.835 | 0.980                      | 0.990 |
| 22 | Northern white-breasted hedgehog | Erinaceus roumanicus | 7             | 0.524                    | 0.763 | 0.192                      | 0.957 |
| 23 | European wildcat    | Felis silvestris | 3             | 0.560                    | 0.760 | 0.600                      | 0.890 |
| 24 | Roan antelope       | Hippotragus equinus | 5             | 0.360                    | 0.696 | 0.170                      | 0.923 |
| 25 | Indri               | Indri indri     | 1             | 0.871                    | 0.871 | 0.971                      | 0.971 |
| 26 | Southern brown bandicoot | Isoodon obesulus | 2             | 0.531                    | 0.605 | 0.000                      | 0.000 |
| 27 | Rufous hare-wallaby | Lagorchestes hirsutus | 1             | 0.610                    | 0.610 | 0.000                      | 0.000 |
| 28 | Brandt's vole       | Lasiopodomys brandti | 21           | 0.552                    | 0.755 | 0.000                      | 0.882 |
| 29 | North American brown lemming | Lemmus trimucronatus | 3             | 0.710                    | 0.790 | 0.000                      | 0.480 |
| 30 | Ocelot              | Leopardus pardalis | 4             | 0.362                    | 0.586 | 0.000                      | 0.679 |
| 31 | Snowshoe hare       | Lepus americanus | 30            | 0.510                    | 0.790 | 0.120                      | 0.990 |
| 32 | European hare       | Lepus europaeus  | 23            | 0.441                    | 0.680 | 0.000                      | 0.853 |
| 33 | Mountain hare       | Lepus timidus    | 3             | 0.441                    | 0.540 | 0.793                      | 0.910 |
| 34 | African bush elephant| Loxodonta africana | 7             | 0.510                    | 0.720 | 0.120                      | 0.880 |
| 35 | Eurasian otter      | Lutra lutra     | 5             | 0.460                    | 0.690 | 0.000                      | 0.224 |
| 36 | African wild dog    | Lycaon pictus   | 8             | 0.590                    | 0.760 | 0.000                      | 0.711 |
| 37 | Eurasian lynx       | Lynx lynx       | 10            | 0.477                    | 0.674 | 0.000                      | 0.820 |
| 38 | Bobcat              | Lynx rufus      | 18            | 0.610                    | 0.800 | 0.000                      | 0.933 |
| ID | Species Common name | Latin name | Population | n | ncDNA Heterozygosity Min | ncDNA Heterozygosity Max | mtDNA Haplotype diversity Min | mtDNA Haplotype diversity Max |
|----|---------------------|------------|------------|---|--------------------------|--------------------------|-----------------------------|-----------------------------|
| 39 | Common vole         | Microtus arvalis | 17 | 0.640 | 0.810 | 0.300 | 0.960 |
| 40 | Reeves’s muntjac    | Muntiacus reevesi | 7  | 0.798 | 0.848 | 0.788 | 0.985 |
| 41 | European mink       | Mustela lutreola | 3  | 0.379 | 0.539 | 0.000 | 0.939 |
| 42 | European polecat     | Mustela putorius | 3  | 0.556 | 0.595 | 0.741 | 0.952 |
| 43 | Western red-backed vole | Myodes Californicus | 4  | 0.850 | 0.880 | 0.000 | 0.590 |
| 44 | Gray-sided vole      | Myodes rufocanus | 36 | 0.577 | 0.938 | 0.000 | 1.000 |
| 45 | Southern plains woodrat | Neotoma Micropus | 3  | 0.837 | 0.840 | 0.915 | 0.933 |
| 46 | White-tailed deer    | Odocoileus virginianus | 9  | 0.780 | 0.837 | 0.602 | 0.964 |
| 47 | Bighorn sheep        | Ovis canadensis  | 20 | 0.390 | 0.720 | 0.000 | 0.910 |
| 48 | Chimpanzee           | Pan troglodytes  | 1  | 0.756 | 0.756 | 0.860 | 0.860 |
| 49 | Lion                 | Panthera leo     | 1  | 0.699 | 0.699 | 0.326 | 0.326 |
| 50 | Jaguar               | Panthera onca    | 3  | 0.622 | 0.724 | 0.846 | 0.933 |
| 51 | Collared peccary     | Pecari tajacu    | 3  | 0.601 | 0.685 | 0.072 | 0.652 |
| 52 | Amami rabbit         | Pentalagus furnessi | 2  | 0.306 | 0.400 | 0.000 | 0.139 |
| 53 | Yellow-footed rock-wallaby | Petrogale Xanthopus | 8  | 0.370 | 0.660 | 0.000 | 0.450 |
| 54 | Koala                | Phascolarctos cinereus | 8  | 0.346 | 0.847 | 0.000 | 0.590 |
| 55 | Long-nosed potoroo   | Potorus tridactylus | 12 | 0.472 | 0.800 | 0.000 | 0.950 |
| 56 | Raccoon              | Procyon lotor    | 6  | 0.630 | 0.726 | 0.000 | 0.717 |
| 57 | Bush rat             | Rattus fuscipes  | 15 | 0.030 | 0.830 | 0.000 | 0.889 |
| 58 | Swamp deer           | Rucervus duvaucelli | 3  | 0.542 | 0.593 | 0.154 | 0.720 |
| 59 | Chamois              | Rupicapra rapicapa | 5  | 0.539 | 0.624 | 0.193 | 0.893 |
| 60 | Sandhill dunnart     | Sminthopsis psammophila | 2  | 0.712 | 0.828 | 0.477 | 0.577 |
| 61 | Wild boar            | Sus scrofa       | 3  | 0.500 | 0.622 | 0.000 | 0.705 |
| 62 | American badger      | Taxidea taxus    | 5  | 0.678 | 0.794 | 0.000 | 0.810 |
| 63 | Brown bear           | Ursus arctos     | 15 | 0.265 | 0.792 | 0.000 | 0.820 |
| 64 | Polar bear           | Ursus maritimus  | 3  | 0.680 | 0.710 | 0.530 | 0.890 |
| 65 | Asian black bear     | Ursus thibetanus | 6  | 0.300 | 0.650 | 0.413 | 0.788 |
| 66 | Perote ground squirrel | Xeromysperophilus Perotensis | 3  | 0.603 | 0.735 | 0.325 | 0.758 |

Total (range) = 511 (1–44)

Marine species

| ID | Species Common name | Latin name | Population | n | ncDNA Heterozygosity Min | ncDNA Heterozygosity Max | mtDNA Haplotype diversity Min | mtDNA Haplotype diversity Max |
|----|---------------------|------------|------------|---|--------------------------|--------------------------|-----------------------------|-----------------------------|
| 1  | Galapagos fur seal  | Arctocephalus Galapagoensis | 3  | 0.660 | 0.690 | 0.760 | 0.810 |
| 2  | Common minke whale  | Balaenoptera Acutorostrata | 8  | 0.506 | 0.689 | 0.841 | 0.978 |
| 3  | Antarctic minke whale | Balaenoptera Bonaerensis | 2  | 0.854 | 0.858 | 0.982 | 0.985 |
| 4  | Sei whale           | Balaenoptera borealis | 3  | 0.630 | 0.660 | 0.480 | 0.610 |
| 5  | Bryde’s whale       | Balaenoptera brydei | 6  | 0.605 | 0.691 | 0.320 | 0.853 |
| 6  | Eden’s whale        | Balaenoptera edeni | 1  | 0.241 | 0.241 | 0.224 | 0.224 |
| 7  | Blue whale          | Balaenoptera musculus | 6  | 0.625 | 0.761 | 0.683 | 0.958 |
| 8  | Fin whale           | Balaenoptera physalus | 5  | 0.490 | 0.830 | 0.150 | 0.860 |
| 9  | Northern fur seal   | Callorhinus ursinus | 8  | 0.782 | 0.807 | 0.990 | 0.997 |
| 10 | Hector’s dolphin    | Cephalorhynchus hectori | 4  | 0.469 | 0.548 | 0.000 | 0.745 |
| 11 | Hooded seal         | Cystophora cristata | 4  | 0.730 | 0.740 | 0.991 | 1.000 |

(Continues)
genetic diversity between mtDNA and ncDNA. These indices are given by the same equation: \( h \) or \( HE \)

\[
H_E = \frac{N}{N-1} \left(1 - \sum_{i=1}^{k} x_i^2 \right),
\]

where \( x_i \) is the relative frequency of a haplotype (or allele), \( k \) is the number of haplotypes (or alleles) and \( N \) is the number of investigated individuals. To minimize the methodological variation, this meta-analysis focused on the control region (D-loop) of mtDNA and microsatellites of ncDNA, which are frequently used in genetic studies of wild mammals. The control region is the most polymorphic region of the mtDNA genome and is preferred for population-level

| ID | Species | Latin name | Population | Heterozygosity | Haplotype diversity |
|----|---------|------------|------------|----------------|---------------------|
|    |         |            |            | Min | Max | Min | Max |
| 12 | Beluga whale | Delphinapterus leucas | 12 | 0.592 | 0.748 | 0.294 | 0.786 |
| 13 | Common dolphin | Delphinus delphis | 14 | 0.631 | 0.820 | 0.342 | 0.977 |
| 14 | Sea otter | Enhydra lutris | 3 | 0.401 | 0.434 | 0.180 | 0.451 |
| 15 | Southern right whale | Eubalaena australis | 3 | 0.800 | 0.830 | 0.750 | 0.780 |
| 16 | Steller sea lion | Eumetopias jubatus | 3 | 0.614 | 0.659 | 0.872 | 0.922 |
| 17 | Risso’s dolphin | Grampus griseus | 5 | 0.691 | 0.711 | 0.868 | 0.960 |
| 18 | Gray seal | Halichoerus grypus | 7 | 0.656 | 0.790 | 0.835 | 0.978 |
| 19 | Northern bottlenose whale | Hyperoodon ampullatus | 3 | 0.624 | 0.655 | 0.430 | 0.550 |
| 20 | Atlantic white-sided dolphin | Lagenorhynchus acutus | 6 | 0.662 | 0.747 | 0.868 | 0.946 |
| 21 | White-beaked dolphin | Lagenorhynchus albirostris | 3 | 0.545 | 0.673 | 0.657 | 0.879 |
| 22 | Pacific white-sided dolphin | Lagenorhynchus obliquidens | 2 | 0.660 | 0.760 | 0.894 | 0.993 |
| 23 | Dusky dolphin | Lagenorhynchus obscurus | 3 | 0.701 | 0.735 | 0.950 | 0.970 |
| 24 | Gray’s beaked whale | Mesoplodon grayi | 3 | 0.659 | 0.675 | 0.949 | 0.987 |
| 25 | Australian sea lion | Neophoca cinerea | 6 | 0.560 | 0.648 | 0.000 | 0.744 |
| 26 | Yangtze finless porpoise | Neophocaena asiaeorientalis | 6 | 0.631 | 0.702 | 0.000 | 0.650 |
| 27 | Walrus | Odobenus rosmarus | 8 | 0.621 | 0.652 | 0.991 | 1.000 |
| 28 | Killer whale | Orcinus orca | 2 | 0.479 | 0.647 | 0.450 | 0.682 |
| 29 | South American sea lion | Otaria flavescens | 6 | 0.700 | 0.790 | 0.830 | 0.940 |
| 30 | Harbor seal | Phoca vitulina | 13 | 0.515 | 0.697 | 0.363 | 0.943 |
| 31 | New Zealand sea lion | Phocarctos hookeri | 1 | 0.642 | 0.642 | 0.441 | 0.441 |
| 32 | Sperm whale | Physeter macrocephalus | 4 | 0.711 | 0.791 | 0.000 | 0.677 |
| 33 | La Plata dolphin | Pontoporia blainvillei | 4 | 0.770 | 0.790 | 0.680 | 0.900 |
| 34 | Tucuxi | Sotalia flaviatilis | 1 | 0.681 | 0.681 | 0.863 | 0.863 |
| 35 | Guiana dolphin | Sotalia guianensis | 4 | 0.489 | 0.550 | 0.100 | 0.927 |
| 36 | Indo-Pacific humpback dolphin | Sousa chinensis | 1 | 0.268 | 0.268 | 0.285 | 0.285 |
| 37 | Striped dolphin | Stenella coeruleoalba | 5 | 0.691 | 0.749 | 0.981 | 1.000 |
| 38 | Atlantic spotted dolphin | Stenella frontalis | 3 | 0.593 | 0.794 | 0.729 | 0.964 |
| 39 | Spinner dolphin | Stenella longirostris | 13 | 0.707 | 0.835 | 0.200 | 0.975 |
| 40 | Indo-Pacific bottlenose dolphin | Tursiops aduncus | 2 | 0.510 | 0.520 | 0.600 | 0.647 |
| 41 | Burrunan dolphin | Tursiops australis | 11 | 0.541 | 0.699 | 0.000 | 0.803 |
| 42 | Common bottlenose dolphin | Tursiops truncatus | 24 | 0.210 | 0.830 | 0.000 | 1.000 |

Total (range) = 231 (1–24)

Abbreviations: mtDNA, mitochondrial DNA; ncDNA, nuclear DNA.
were excluded. However, (<10 individuals) or studies of introduced populations were selected. Data from studies with a small sample size (including unpublished datasets on gray-sided voles) were obtained from 739 populations of 108 species, ples size for the other value was ≥10. Consequently, data were obtained from 739 populations of 108 species, including unpublished datasets on gray-sided voles (Myodes rufocanus) and sika deer (Cervus nippon) from the author’s research group (Table 1). When an article provided the data on haplotype or allele frequency but no direct information on h or \( H_E \), the latter were calculated based on the frequency. Authors were contacted, when the article or supporting materials lacked the required information. A list of populations, including the data with references, is provided in the Supporting Information (Table S1).

In the analyzed publications, populations were defined according to the original purpose of studies, and thus the definitions varied between the ecological and the evolutionary paradigm. According to the ecological paradigm, a population is defined as “a group of organisms of the same species occupying a particular space at a particular time.” In contrast, according to the evolutionary paradigm, a population is defined as “a group of interbreeding individuals that exist together in time and space-time” (Waples & Gaggiotti, 2006). Under the ecological paradigm, individuals were usually grouped according to the independence of habitats, while most studies based on the evolutionary paradigm examined the genetic independence of groups based on genetic distance (e.g., \( F_{ST} \)). In all studies, some independence was confirmed for each group, and dispersal among groups was expected to be limited.

2.2 Effective gene number

The effective gene number (the number of genes that can be inherited) differs between mtDNA and ncDNA. The effective gene number of mtDNA is one-fourth of that of ncDNA in idealized populations, and, thus, mtDNA diversity is predicted to be lower than ncDNA diversity (Birky et al., 1983). The theory of Birky et al. (1983) can be confirmed in a finite population for which the following holds true: (a) diploid, dioecious, sexually reproducing; (b) complete maternal inheritance of mtDNA; (c) nonoverlapping generations; (d) random mating within a population with equal numbers of males and females; (e) no mutations; (f) a constant population size through generations (no bottlenecks); (g) no subpopulation structure, and, thus, no immigration and (h) no selection. In idealized populations, heterozygosity for ncDNA (\( H_E \) defined by Nei, 1987) steadily decays in inverse proportion to the effective population size (\( N_e \)): \[ \Delta H_E = \frac{1}{2 \times N_e^2} \], where \( \Delta H_E \) is the decay rate between generations. In this case, the effective gene number for ncDNA is \( 2 \times N_e \). The decay rate of haplotype diversity for mtDNA (\( \Delta h \); Nei & Tajima, 1981) is given by a similar equation with the different effective gene number (0.5 \( \times N_e \)): \[ \Delta h = \frac{1}{0.5 \times N_e} \]. Therefore, in idealized populations, \( h \) decays four times faster than \( H_E \).

2.3 Terrestrial versus marine mammals

Genetic discontinuities are often associated with landscape features (Manel, Schwartz, Luikart, & Taberlet, 2003), and population connectivity may differ among species inhabiting different landscapes. Mobile species that are distributed across continuous habitats may exhibit lower genetic differentiation and persistent gene flow in comparison to species in heterogeneous habitats (Amaral et al., 2012). Most marine mammal populations may be connected by the sea, whereas some populations of terrestrial mammals are isolated from others because of habitat fragmentation. Genetic diversity may reflect population connectivity with habitat features. Therefore, \( h \) and \( H_E \) were compared between the terrestrial and marine mammals.

2.4 Statistical tests

Differences between \( h \) and \( H_E \) within a population were tested using the asymptotic Wilcoxon signed rank test, which is a nonparametric test for paired samples. Differences in the mean of \( h \) or \( H_E \) between various populations were tested using the Brunner–Munzel test, which is a nonparametric test that adjusts for unequal variances. Differences in the variances of \( h \) or \( H_E \) were tested using the Kolmogorov–Smirnov test. The correlation between \( h \) and \( H_E \) was tested using the Pearson’s product–moment correlation. Species compositions were tested using the Fisher’s exact test. The effect
of sample sizes (the number of analyzed individuals in a population) on $h$ or $H_E$ was analyzed by the ordinary linear regression method. The effect of conservation status on $h$ or $H_E$ was also tested by the ordinary linear regression method, where the conservation status was treated as an ordinal scale variable. These statistical analyses were done using R version 3.6.3 (R Core Team, 2020).

3 | RESULTS

3.1 | Empirical patterns of mtDNA and ncDNA diversity

Based on the datasets of 739 populations of 108 mammalian species, mtDNA haplotype diversity ($h$) was plotted for microsatellite heterozygosity ($H_E$) for each population (Figure 1). $H_E$ ranged from 0.030 to 0.938, whereas $h$ ranged from 0 to 1. Many of $H_E$ values (63.7%) were between 0.6 and 0.8, whereas the percentage of $h$ values within this range was 22.9.

In 54.9% of populations, $h$ was higher than $H_E$, although a significant difference was not observed between the means of $h$ (0.629) and $H_E$ (0.669) (the asymptotic Wilcoxon signed rank test, $V = 127,299$, $p = .168$). These results challenged the theoretical prediction based on the difference in the effective gene number between mtDNA and ncDNA. The variance of $h$ $(0.097)$ was significantly higher than that of $H_E$ $(0.018; F = 5.557, p < 2.2 \times 10^{-16})$. A positive correlation between $h$ and $H_E$ was statistically supported ($r_p = .439, t = 13.257, p < 2.2 \times 10^{-16}$).

Large variations of $h$ and $H_E$ were also observed within a species. The range of $h$ and $H_E$ increased with population numbers for a species (Figure 2). The range of $h$ was particularly wide. In most species for which information from 10 or more populations was available, the range of $h$ exceeded 0.8, whereas that of $H_E$ appeared to reach a plateau at medium values. In 21 species with data from 10 or more populations, the variance of $h$ was higher than that of $H_E$. The higher variances of $h$ in all but one species (Rattus fuscipes, Figure 3) were statistically supported ($F$-test).

3.2 | Effects of sample size

Although the samples size (the number of analyzed individuals in a population) was generally set at >10, it greatly varied between 9 and 1,210 for $h$ and between 9.51 and 1,372 for $H_E$. Since $h$ and $H_E$ were predicted to increase with bigger sample sizes, the effects of sample sizes were analyzed using the ordinal linear regression (Supporting Information Figure S1). Sample size in the logarithmic scale did not have any significant effect on $H_E$ ($t = 0.161, p = .872$, adjusted $R^2 = -.001$), whereas it had a significant positive effect on $h$ ($t = 2.470, p = .014$, adjusted $R^2 = .007$). However, this model explained only 0.7% of the variation of $h$. Additionally, the effects of the number of microsatellite loci on $H_E$ and the length of the mtDNA region on $h$ were investigated. The number of loci in the logarithmic scale had no significant effect ($t = -1.668, p = .096$, adjusted $R^2 = .002$). The analyzed length of the mtDNA region in the logarithmic scale had a positive effect, but its explanatory power was limited ($t = 3.336, p = .001$, adjusted $R^2 = .014$).

3.3 | Effects of conservation status

The highest haplotype diversity ($h = 1$) was recorded in 14 populations of 6 species. Four of the six species were the marine species, namely, the hooded seal (Cystophora cristata), the walrus (Odobenus rosmarus), the Mediterranean striped dolphin (Stenella coeruleoalba) and the bottlenose dolphins (Tursiops truncates). The two terrestrial remaining species were the gray-sided vole (Myodes rufocanus) and the common impala (Aepyceros melampus).
The lowest diversity \((h = 0)\) was observed in 69 populations of 35 species. Many of them were small rodents inhabiting small islands (the Supporting Information Table S1), for example, lemmings \((Dicrostonyx\ groenlandicus, Lemmus\ trimucronatus)\), the Australian bush rat \((Rattus\ fuscipes)\) and vole species \((Myodes\ rufocanus, M.\ californicus)\). The following endangered species also exhibited the lowest \(h\) \((h = 0)\): the African wild dog \((Lycaon\ pictus; EN)\), the eco-mon hamster \((Cricetus\ cricetus; CR)\), the koala \((Phascolarctos\ cinereus; VU)\), the rufous hare-wallaby \((Lagorchestes\ hirsutus; VU)\), the European mink \((Mustela\ lutreola; CR)\), the Hector’s dolphin \((Cephalorhynchus\ hectori; EN)\), the sperm whale \((Physeter\ macrocephalus; VU)\), the Yangtze finless porpoise \((Neophocaena\ asiorientalis; EN)\), the Maui’s dolphin \((Cephalorhynchus\ hectori; EN)\), the Australian sea lion \((Neophoca\ cinerea; EN)\) and the bottlenose dolphin \((Tursiops\ truncatus; LC, T.\ australis; NT)\). Only six marine species were observed in the populations with \(h = 0\). Abbreviations following the Latin names indicate the conservation status of the species in the IUCN Red List website 2020. The IUCN Red List defines conservation status for extant species according to the following ranks from the highest to the lowest: critically endangered (CR), endangered (EN), vulnerable (VU), near threatened (NT) and least concerned (LC).

Since a population size was not recorded for most populations, the conservation status of species, as a proxy of population size, was considered as an explanatory variable of \(H_E\) and \(h\). Three species whose conservation status was “data deficient” were excluded from the analyses. Different conservation statuses (CR, EN, VU, NT and LC) accounted for 3, 15, 17, 11 and 59 species, respectively.

The lowest median of \(H_E\) \((0.481)\) was observed in 12 populations of 3 species with CR, while 59 species \((549\ populations)\) with LC had the highest median of \(H_E\) \((0.697, Figure\ 4)\). The magnitude relationship in \(H_E\) was consistent with the conservation status, except for NT species. Since conservation statuses can be treated as an ordinal scale variable, the linear regression analysis was conducted to test whether \(H_E\) increased with conservation status ranks. The effects of conservation status on \(H_E\)

**FIGURE 2**  The variations in genetic diversity in association with the number of observed populations. The variation is represented by the range between the minimum and maximum genetic diversity (mitochondrial DNA [mtDNA] haplotype diversity \([h]\) or microsatellite DNA heterozygosity \([H_E]\)) for each of the 108 mammalian species. Solid and open circles represent the range for \(h\) and \(H_E\), respectively.

**FIGURE 3**  Box plot of microsatellite DNA heterozygosity \((a): H_E)\) and mitochondrial DNA (mtDNA) haplotype diversity \((b): h)\) for 21 mammalian species with data on genetic diversity from 10 or more populations. A bold line and the upper and the lower edge of a box indicate the median, 75 percentile and 25 percentile, respectively. Circles denote outliers. The upper and the lower whiskers were drawn following the default of R version 3.6.3 (R Core Team, 2020)
was significant, but the explanatory power of the model was limited ($F = 8.839, p = 5.7 \times 10^{-7}$, adjusted $R^2 = .041$).

$H_E$ varied within a single conservation status (Table 1). The common hamster (*Cricetus cricetus*, $n = 8$), the indri (*Indri indri*, $n = 1$) and the European mink (*Mustela lutreola*, $n = 3$) are ranked as CR and had $H_E$ values of 0.111–0.786, 0.871 and 0.379–0.595, respectively. The variance of pooled $H_E$ for the three species was 0.062. The $H_E$ of 59 species with LC ranged between 0.030 and 0.938, with a variance of 0.016.

The relationship between $h$ and conservation status was similar to that observed for $H_E$. However, the explanatory power of the model was smaller than one-third of that for $H_E$ ($F = 3.249, p = .012$, adjusted $R^2 = .012$). The variation of $h$ within a conservation status was wide (Table 1). In the three species with CR, the pooled $h$ ranged from 0.000 to 0.971, with a variance of 0.139. The values of $h$ for 59 LC species ranged between 0 and 1, and the variance was 0.094.

### 3.4 Terrestrial versus marine mammals

$H_E$ exhibited a unimodal distribution pattern, in which the values for more than half of the populations were between 0.6 and 0.8, for both terrestrial and marine species (58.1 and 76.2% for the terrestrial and marine mammals, respectively; Figure 5a,c). Frequency distribution patterns did not significantly differ between the terrestrial and marine species (Kolmogorov–Smirnov test, $D = 0.099, p = .087$; this is an approximate $p$-value because of the presence of tie values). The mean of $H_E$ was similar at 0.668 and 0.672 for the terrestrial and marine species, respectively (Brunner–Munzel test, test statistic = 0.240, $p = .810$), while the variance of $H_E$ was significantly higher in the terrestrial species (0.021) than in the marine species (0.011; $F = 1.933, p = 2.5 \times 10^{-8}$).

The frequency distribution of $h$ showed different patterns. Terrestrial species values exhibited a U-shaped pattern (Figure 5b), whereas for the marine species, the frequency steadily increased from low to high and had a right triangular shape (Figure 5d, Kolmogorov–Smirnov test, $D = 0.224, p = 2.5 \times 10^{-7}$; this is an approximate $P$-value because of the presence of tie values). The mean of $h$ was significantly higher in the marine species (0.740) than in the terrestrial species (0.578, Brunner–Munzel test, test statistic = 7.329, $p = 1.1 \times 10^{-12}$), because of a larger number of the terrestrial species populations with $h = 0$. The variance of $h$ was significantly higher in the terrestrial species (0.104) than in the marine species (0.065; $F = 1.595, p = 6.2 \times 10^{-5}$).

Contrasting species compositions were observed between the highest and the lowest categories of haplotype diversity. Marine species occupied 16 of 27 species in populations with very high $h$ (>0.95), whereas they accounted for a small fraction (6 of 35 species) in populations with very low $h$ (<0.05). The compositions of the marine and terrestrial species significantly differed from each other (Fisher’s exact test, $p = .001$). The $H_E$ of

---

**Figure 4** Effects of conservation status on microsatellite DNA heterozygosity ((a): $H_E$) and mitochondrial DNA (mtDNA) haplotype diversity ((b): $h$). Conservation status (critically endangered [CR], endangered [EN], vulnerable [VU], near threatened [NT], least concerned [LC] or data-deficient [DD]) was given to each species based on the IUCN Red List website 2020. Species whose status was unknown (i.e., DD) were excluded. A bold line and the upper and the lower edge of a box indicate the median, the 75 percentile and the 25 percentile, respectively. Circles denote outliers. The upper and the lower whiskers were drawn following the default of R version 3.6.3 (R Core Team, 2020). Figures in parentheses indicate the number of populations (left) and the number of species (right).
populations with \( h > 0.95 \) was significantly lower for the marine species (0.738) than for the terrestrial species (0.785, Brunner–Munzel test, test statistic = 2.201, \( p = .031 \)), whereas the \( HE \) of populations with \( h < 0.05 \) did not show any significant difference between the terrestrial (0.542) and the marine species (0.600, Brunner–Munzel test, test statistic = 1.152, \( p = .264 \)).

4 | DISCUSSION

The genetic diversity of the 739 wild populations of 108 mammalian species showed the following features:

1. Haplotype diversity (\( h \)) was higher than heterozygosity (\( HE \)) in more than half of the populations (54.9%), and the mean of \( h \) (0.629) was not significantly lower than the mean of \( HE \) (0.669). These observations were not in agreement with the prediction based on the difference in the effective gene numbers between mtDNA and ncDNA.

2. Majority of \( HE \) (63.7%) ranged from 0.6 to 0.8, whereas the percentage of \( h \) within this range was 22.9% (Figure 1).

3. The variation of genetic diversity was considerable, even within a species. In particular, the variation of \( h \) was notably high (Figures 2 and 3).

4. The effects of conservation status, a proxy of population size, on \( HE \) and \( h \) were significant. \( HE \) and \( h \) decreased with conservation status rank (Figure 4). However, the explanatory power of conservation status was low, which was particularly true for \( h \).

5. The frequency distribution of \( h \) was distinct between the terrestrial and marine species because the former included various populations with extremely low diversity (Figure 5b,d). Such a contrasting pattern was not observed for the frequency distribution of \( HE \) (Figure 5a,c).

These observations cannot be interpreted via the effective gene number theory (Birky et al., 1983). In particular, the first and fifth features undoubtedly indicate that the genetic diversity of the studied empirical populations is determined not only by the effective gene numbers but also by other factors. Nei (1983) and Birky et al. (1989) suggested the deviation of genetic diversity from the theoretical prediction in empirical populations. However, their examination was very limited, as little data (see Section 1) were available to test the prediction at that time. Based on the large number of datasets from mammalian populations, the current work revealed that the mtDNA haplotype diversity was higher than expected from the theory of the effective gene number (Figures 1 and 5).

Mammalian species satisfy the idealized population conditions of diploidy and dioecism. Additionally, they are sexually reproducing, and their mtDNA is almost entirely maternally inherited (Avise, 2004). The criteria
of random mating are also satisfied in many populations. The effects of selection may be minimal for \( H_E \), as microsatellites are known to be under less selective constraints (Yashima & Innan, 2017). In contrast, the effects of selection on the mtDNA diversity should not be ignored. However, the mtDNA of mammals does not reject the nearly neutral model (Nabholz et al., 2008). Therefore, the effects of selection on \( h \) based on the control region of mammalian mtDNA may be minor. Through excluding those factors, the effects of mutation rates and subpopulation structures were the focus in this study, in addition to the genetic drift, Nei (1983) and Birky et al. (1989) suggested that a higher mutation rate could explain the unpredictably high genetic diversity of organelle genes observed in empirical populations. It is well established that mtDNA has a higher mutation rate than that of ncDNA (Allio, Donega, Galtier, & Nabholz, 2017). The biased mutation rate hypothesis may therefore explain the high \( h \) observed in the current study. In addition, mutation rates vary among species and populations (Ellegren, 2004; Galtier, Nabholz, Glemin, & Hurst, 2009), and Nabholz et al. (2008) asserted, based on indirect evidence from the species-level analyses, that mutation rates could be a determinant of mtDNA diversity under the nonequilibrium conditions of mutation and drift. Therefore, the biased mutation rate hypothesis could explain, at least in part, the variation of \( h \) and \( H_E \) observed in wild populations (Figures 2 and 3). However, the biased mutation rate hypothesis may not be strong at explaining the high intraspecific variations of \( h \) because mutation rate must show a considerable intraspecific variation corresponding to the intraspecific variation of \( h \) (Figure 3). Although knowledge on the intraspecific variation of mutation rates remains limited, there is little evidence suggesting that the variation in mutation rates could be a strong driver of the variation in genetic diversity (Ellegren & Galtier, 2016).

In addition to mutation, immigration should also be considered as a source of genetic diversity. Immigrants from another subpopulation may provide alien genes and enhance the genetic diversity of a focal subpopulation. To produce immigrants, a population must consist of genetically heterogeneous subpopulations. Male-biased dispersal (female philopatry) and polygyny are prevalent in mammals (Greenwood, 1980; Ishibashi & Saitoh, 2008; Lawson Handley & Perrin, 2007; Le Galliard, Remy, Ims, & Lambin, 2012; Mabry, Shelley, Davis, Blumstein, & van Vuren, 2013). Sex-biased dispersal may be related to social mating systems (Mabry et al., 2013) and result in different gene frequencies between sexes within and among populations or subpopulations (Prout, 1981). Recently, various studies reported intersexual differences in fine-scale spatial genetic structure in mammals (e.g., Banks & Peakall, 2012; Cooper et al., 2010; Ishibashi, Zenitani, & Saitoh, 2013; Peakall, Ruibal, & Lindenmayer, 2003; Temple, Hoffman, & Amos, 2006). Therefore, it is highly likely that, in mammals, several mtDNA-based subpopulations, which are shaped by female philopatry, are nested within an ncDNA-based population, in which the mtDNA-based subpopulations are linked by male dispersal.

The small effective gene number of mtDNA could be modulated by male immigrants between the mtDNA-based subpopulations. Male immigrants from other mtDNA-based subpopulations may carry alien mtDNA haplotypes to a focal mtDNA-based subpopulation, although they are not able to transfer the haplotypes to the next generation. In contrast, most microsatellite alleles brought in by immigrants may be familiar to the focal subpopulation because the accumulation of male immigrants has homogenized microsatellite composition among subpopulations in an ncDNA-based population. Therefore, male immigrants provide asymmetric genetic information in subpopulations. The diversity of mtDNA could be enhanced by male immigrants, whereas their effects on microsatellite diversity may be limited.

The effects of male immigrants on genetic diversity may be determined by dispersal rate. Different dispersal rates could cause varying \( h \) and \( H_E \). In addition, the female-biased breeding sex ratio contributes to reducing the differences in the effective gene number between mtDNA and ncDNA. Chesser and Baker (1996) considered the effects of dispersal rates and breeding sex ratios on genetic diversity. They described a model called “typical mammalian population structure,” wherein a high polygyny rate (0.4) and predominant male dispersal (\( d_m = 0.75 \) and \( d_f = 0.25 \), where \( d_m \) and \( d_f \) represent the dispersal rate of males and females, respectively) were assumed. They showed that effective population sizes became larger for maternally inherited genes than for genes inherited from both parent. Therefore, a higher mtDNA diversity is favored by the “typical mammalian population structure.” The results of the current study support their predictions, and the hypothesis based on nested population structure also has the potential to explain the empirical pattern of \( h \) and \( H_E \).

Both the biased mutation rate hypothesis and the nested population structure hypothesis discussed here may independently explain the empirical patterns of \( h \) and \( H_E \). However, the factors addressed by these hypotheses are not mutually exclusive, and thus the combined approaches, which include these factors, may more realistically address the empirical patterns. The biased mutation rate toward mtDNA could be a basic mechanism responsible for the high diversity of mtDNA. However, it may be challenging to explain the large
intraspecific variation of $h$ based on the biased mutation rate hypothesis alone (Figure 3). As a complementary mechanism, the nested population structure hypothesis has an advantage in explaining the high variation of $h$ because dispersal rates and breeding sex ratios may vary among wild populations.

4.1 Terrestrial versus marine mammals

Significant differences were revealed in the genetic diversity between the terrestrial and marine species in this study. The haplotype diversity ($h$) of the marine mammals was significantly higher than that of the terrestrial mammals (Figure 5b,d). The populations with $h < 0.05$ were frequently observed among the terrestrial species, while the marine species occupied a larger proportion of the populations with $h > 0.95$. There may be various obstacles preventing the dispersal of terrestrial mammals, and some terrestrial populations are likely to be isolated. As such isolated populations may become small and unstructured, their mtDNA genes may be severely affected by genetic drifts. The populations with $h = 0$, for example, are often observed on small islands (Table S1, see also Sato et al., 2017). In contrast, marine mammals may enjoy the open habitat without the interruption of movement. This may contribute to their larger effective population sizes and nested population structure, which in turn help to maintain high mtDNA diversity.

The degree of polygyny influences the differences in the effective gene numbers between mtDNA and ncDNA (Chesser & Baker, 1996). When the breeding sex ratio is even, the effective gene number of mtDNA is one-fourth of that of ncDNA. However, the female-biased breeding sex ratio caused by polygyny reduces the difference in the effective gene number. Very low male ratios were reported for highly polygynous species: 0.030 and 0.024 for the northern Mirounga angustirostris and the southern elephant seal, M. leonina, respectively (Hoelzel, le Boeuf, Reiter, & Campagna, 1999). In the extreme case of the southern elephant seal, harem holders accounted for 89.6% of paternities (Fabiani, Galimberti, Sanvito, & Hoelzel, 2004). Males of various pinniped species compete to monopolize females (Cassini, 1999), and sexual size dimorphisms, which are a proxy of polygyny, are more notable in marine species than in terrestrial species (Weckerly, 1998). The current study’s observation that the $H_E$ values of the populations with $h > 0.95$ were significantly lower for the marine species than for the terrestrial species, may be attributed to the lower male ratio in the breeding populations of the highly polygynous marine mammals.

4.2 Effects of population size

It is generally accepted that genetic diversity is correlated with population size (Frankham, 1996), and the body of evidence supporting this relationship is increasing. Heterozygosity of ncDNA is negatively correlated with body size (an indicator of population size) in mammals (Doyle et al., 2015; Wooten & Smith, 1985). In the current study, the significant effects of conservation status on $H_E$ were also observed (Figure 4), although the explanatory power was limited (adjusted $R^2 = .041$). In addition to the difference between ncDNA and mtDNA, the effective ncDNA gene numbers differ between chromosomes. In a diploid species with X and Y sex chromosomes, the effective gene number ratio for autosomes, X, and Y chromosomes should be 4:3:1 when a population exhibits random mating and an even sex ratio. Therefore, a higher genetic diversity is expected in autosomes than in sex chromosomes. Ellegren and Galtier (2016) reviewed related research and reported that this prediction was generally supported in animal populations. Based on these observations, ncDNA diversity may be mostly determined by the effective gene number.

The effect of population size on mtDNA diversity remains controversial. In cetacean species, Vachon, Whitehead, and Frasier (2018) reported that the estimated population size has a significant effect on microsatellite diversity but not on mtDNA diversity. Bazin et al. (2006) suggested that mtDNA diversity does not reflect population size (see also Pedreschi et al., 2018), whereas Mulligan et al. (2006) asserted that mtDNA diversity might correlate with population size in mammals (see also Piganeau & Eyre-Walker, 2009; Sato et al., 2017). In the current work, the significant effects of conservation status on $h$ were observed (Figure 4), but conservation status only explained 1.3% of the variation of $h$ (adjusted $R^2 = .013$). The explanatory power of the model for $h$ was much smaller than the model for $H_E$ (adjusted $R^2 = .041$). Nabholz et al. (2008, 2009) concluded that mtDNA diversity was essentially unpredictable. One example of this unpredictability is unexpectedly high haplotype diversity of island populations. Sato et al. (2017) reported such populations of wood mice (Apodemus speciosus) on Hakatajima island, the haplotype diversity of which diverged from the general relationship between haplotype diversity and island size. A similar pattern was observed in sika deer (Cervus nippon) on Yakushima island (Table S1, Terada & Saitoh, 2018). The nested population structure hypothesis predicts that the structure of populations with higher haplotype diversity is more complex than that of those with standard haplotype diversity. Habitat heterogeneity, which helps populations to be structured, may be higher in those islands than in other similar-sized islands.
4.3 Suggestions for future studies

This study focused on mammalian populations and discussed the effects of the nested population structure on genetic diversity. In the bird-like populations with monogamous mating and predominantly female dispersal, the difference in effective gene numbers between mtDNA and ncDNA may be closer to that in the idealized populations, in comparison to mammalian populations (Chesser & Baker, 1996). A comparative study on the genetic diversity between mammals and birds is thus strongly encouraged.

The high variance (Figure 3) and discriminative frequency distribution of $h$ (Figure 5) may represent the complexity of mtDNA diversity. A greater number of factors may influence the mtDNA diversity in comparison to ncDNA diversity. Mitochondrial genomes evolve under different evolutionary rules as compared to nuclear genomes, because of their unique natural history (Ballard & Whitlock, 2004). The relatively small effective gene number is one of the specific features that define their evolutionary rules, and the mutation rate of mtDNA may vary in association with the effective gene number (Lynch, 2010). To fully understand mtDNA diversity, further studies are required on the evolution, ecology and biochemistry of the mitochondrion.

A significant variation was observed for both $h$ and $H_E$ in the current study (Figure 1). This variation could be partially explained by the methodological variation among studies. Mutation rates may depend on the length of analyzed mtDNA, and thus a higher $h$ is expected in studies analyzing longer mtDNA control regions. In fact, a positive effect of analyzed mtDNA length on $h$ was observed (Supporting Information Figure S1), although it explained only 1.4% of the variation of $h$. Microsatellites variation may depend on repeat motifs. $H_E$ increases with the number of repeats due to the higher mutation rates observed in microsatellites with longer repeats (Yashima & Innan, 2017). However, the intraspecific variation for both $h$ and $H_E$ observed in this study cannot be explained by methodological variation (Figures 2 and 3) because most research studies reporting intraspecific variation analyzed the same length of mtDNA and microsatellite loci across the studied populations.

Intraspecific variation prevents us from obtaining a representative value for the genetic diversity of a species, particularly for mtDNA. Therefore, species-level analyses based on the genetic markers examined in this study may not fully reveal the nature of genetic diversity. This study used conservation status as a proxy of population size. However, conservation status is an attribute of species and cannot reflect the intraspecific variation of population sizes. The low explanatory power of the conservation status model represents a limitation of studies based on species-level analyses (Figure 4). Information regarding ecological factors (population size, population structure, mating system and others) is needed for each population. Data accumulation on the population level is essential for a deeper understanding of the mechanisms determining genetic diversity in wild populations.

ACKNOWLEDGMENTS

The author thanks the Ecological Society of Japan for inviting this manuscript for submission to Ecological Research as a memorial article for the ESI award. Jun J. Sato read an early draft and provided helpful comments. The editor and anonymous reviewers provided suggestions to improve the manuscript. Gerald Heckel, Luis A. Pastene and Samuel Deakin helped the author with data compilation. Haruko Ando, Toshihito Takagi and Chisato Terada provided unpublished data. This work was partly supported by a Grant-in-Aid from the JSPS KAKENHI (No. 19H03005). The author also thanks Editage (www.editage.com) for the English language editing. The Ecological Society of Japan: The ESI award 2015.

ORCID

Takashi Saitoh © https://orcid.org/0000-0003-4085-5014

REFERENCES

Allendorf, F. W., & Leary, R. F. (1986). Heterozygosity and fitness in natural populations on animals. In M. E. Soulé (Ed.), Conservation biology: The science of scarcity and diversity (pp. 57–76). Sunderland, MA: Sinauer Associates.

Allio, R., Donega, S., Galtier, N., & Nahholz, B. (2017). Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. Molecular Biology and Evolution, 34, 2762–2772.

Amaral, A. R., Beheregaray, L. B., Bilgmann, K., Boutov, D., Freita, L., Robertson, K. M., ... Möller, L. M. (2012). Seascape genetics of a globally distributed, highly mobile marine mammal: The short-beaked common dolphin (genus Delphinus). PLoS One, 7, e31482-e31415.

Amos, W., & Balmford, A. (2001). When does conservation genetics matter? Heredity, 87, 257–265.

Avise, J. C. (2004). Molecular markers, natural history, and evolution (2nd ed., p. 684). Sunderland, MA: Sinauer Associates.

Avise, J. C., Giblin-Davidson, C., Laerm, J., Patton, J. C., & Lansman, R. A. (1979). Mitochondrial DNA clones and matriarchal phylogeny within and among geographic populations of the pocket gopher, Geomys pinetis. Proceedings of the National Academy of Sciences of the United States of America, 76, 6694–6698.

Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. Molecular Ecology, 13, 729–744.

Banks, S., & Peakall, R. (2012). Genetic spatial autocorrelation can readily detect sex-biased dispersal. Molecular Ecology, 21, 2092–2105.
Bazin, E., Glémín, S., & Galtier, N. (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science*, 312, 570–572.

Birky, C. W., Fuerst, P., & Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: Equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, 121, 613–627.

Birky, C. W., Maruyama, T., & Fuerst, P. (1983). An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics*, 103, 513–527.

Cassini, M. H. (1999). The evolution of reproductive systems in pinnipeds. *Behavioral Ecology*, 10, 612–616.

Chesser, R. K., & Baker, R. J. (1996). Effective sizes and dynamics of unparentally and parentally inherited genes. *Genetics*, 144, 1225–1235.

Cooper, J. D., Waser, P. M., Gopurenko, D., Hellgren, E. C., Gabor, T. M., & DeWoody, J. A. (2010). Measuring sex-biased dispersal in social mammals: Comparisons of nuclear and mitochondrial genes in collared peccaries. *Journal of Mammalogy*, 91, 1413–1424.

Doyle, J. M., Hackin, C. C., Willoughby, J. R., Sundaram, M., & DeWoody, J. A. (2015). Mammalian genetic diversity as a function of habitat, body size, trophic class, and conservation status. *Journal of Mammalogy*, 96, 564–572.

Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics*, 5, 435–445.

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, 17, 422–433.

Fabiani, A., Galimberti, F., Sanvito, S., & Hoelzel, A. R. (2004). Extreme polygyny among southern elephant seals on Sea Lion Island, Falkland Islands. *Behavioral Ecology*, 15, 961–969.

Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology*, 10, 1500–1508.

Frankham, R., Ballow, J. D., & Briscoe, D. A. (2002). *Introduction to conservation genetics* (p. 617). Cambridge: Cambridge University Press.

Galtier, N., Nabholz, B., Glémín, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology*, 18, 4541–4550.

Greenwood, P. J. (1980). Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, 28, 1140–1162.

Hoelzel, A. R., le Boeuf, B. J., Reiter, J., & Campagna, C. (1999). Alpha-male paternity in elephant seals. *Behavioral Ecology and Sociobiology*, 46, 298–306.

Hughes, A. R., Inouye, B. D., Johnson, M. T. J., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, 11, 609–623.

Ishibashi, Y., & Saitoh, T. (2008). Role of male-biased dispersal in inbreeding avoidance in the grey-sided vole (Myodes rufocanus). *Molecular Ecology*, 17, 4887–4896.

Ishibashi, Y., Zenitani, J., & Saitoh, T. (2013). Male-biased dispersal causes intersexual differences in the subpopulation structure of the gray-sided vole. *Journal of Heredity*, 104, 718–724.

Lawson Handley, L. J., & Perrin, N. (2007). Advances in our understanding of mammalian sex-biased dispersal. *Molecular Ecology*, 16, 1559–1578.

Le Galliard, J.-F., Remy, A., Ims, R. A., & Lambin, X. (2012). Patterns and processes of dispersal behaviour in arvicoline rodents. *Molecular Ecology*, 21, 505–523.

Leigh, D. M., Hendry, A. P., Vázquez Domínguez, E., & Friesen, V. L. (2019). Estimated six per cent loss of genetic variation in wild populations since the industrial revolution. *Evolutionary Applications*, 12, 1505–1512.

Lynch, M. (2010). Evolution of the mutation rate. *Trends in Genetics*, 26, 345–352.

Mabry, K. E., Shelley, E. L., Davis, K. E., Blumstein, D. T., & van Vuren, D. H. (2013). Social mating system and sex-biased dispersal in mammals and birds: A phylogenetic analysis. *PloS One*, 8, e57980–e57989.

Mandel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, 18, 189–197.

Miraldo, A., Li, S., Borregaard, M. K., Flórez-Rodríguez, A., Gopalakrishnan, S., Rizvanovic, M., ... Nogués-Bravo, D. (2016). An anthropocene map of genetic diversity. *Science*, 353, 1532–1535.

Mulligan, C. J., Kitchen, A., & Miyamoto, M. M. (2006). Comment on “population size does not influence mitochondrial genetic diversity in animals”. *Science*, 314, 1390a–1390a.

Nabholz, B., Glémín, S., & Galtier, N. (2009). The erratic mitochondrial clock: Variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC Evolutionary Biology*, 9, 54.

Nabholz, B., Mauffrey, J.-F., Bazin, E., Galtier, N., & Glémín, S. (2008). Determination of mitochondrial genetic diversity in mammals. *Genetics*, 178, 351–361.

Nei, M. (1983). Genetic polymorphism and the role of mutation in evolution. In P. K. Koehn & M. Nei (Eds.), *Evolution of genes and proteins* (pp. 165–190). Sunderland, MA: Sinauer Associates.

Nei, M. (1987). *Molecular evolutionary genetics* (p. 513). New York, NY: Columbia University Press.

Nei, M., & Tajima, F. (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics*, 97, 145–163.

Peakall, R., Ruibal, M., & Lindenmayer, D. B. (2003). Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution*, 57, 1182–1195.

Pedreschi, D., Garcia-Rodriguez, O., Yannic, G., Cantarella, E., Diaz, A., Golicher, D., ... Stewart, J. R. (2018). Challenging the European southern refugium hypothesis: Species-specific structures versus general patterns of genetic diversity and differentiation among small mammals. *Global Ecology and Biogeography*, 28, 262–274.

Piganeau, G., & Eyre-Walker, A. (2009). Evidence for variation in the effective population size of animal mitochondrial DNA. *PloS One*, 4, e4396–e4398.

Prout, T. (1981). A note on the Island model with sex dependent migration. *Theoretical and Applied Genetics*, 59, 327–332.

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, 17, 230–237.

Satō, J. J., Tasaka, Y., Tasaka, R., Gunji, K., Yamamoto, Y., Takada, Y., ... Yamaguchi, Y. (2017). Effects of isolation by continental islands in the Seto Inland Sea, Japan, on genetic diversity of the large Japanese field mouse, *Apodemus speciosus*.
(Rodentia: Muridae), inferred from the mitochondrial D-loop region. Zoological Science, 34, 112–121.

Spielman, D., Brook, B. W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences of the United States of America, 101, 15261–15264.

Steffen, W., Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., ... Sörlin, S. (2015). Sustainability. Planetary boundaries: Guiding human development on a changing planet. Science, 347, 1259855–1259855.

Stoneking, M., Hedgecock, D., Higuchi, R. G., Vigilant, L., & Erlich, H. A. (1991). Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. American Journal of Human Genetics, 48, 370–382.

Temple, H., Hoffman, J. I., & Amos, W. (2006). Dispersal, philopatry and intergroup relatedness: Fine-scale genetic structure in the white-breasted thrasher, Ramphocinclus brachyurus. Molecular Ecology, 15, 3449–3458.

Terada, C., & Saitoh, T. (2018). Phenotypic and genetic divergence among Island populations of sika deer (Cervus nippon) in southern Japan: A test of the local adaptation hypothesis. Population Ecology, 60, 211–221.

Vachon, F., Whitehead, H., & Frasier, T. R. (2018). What factors shape genetic diversity in cetaceans? Ecology and Evolution, 8, 1554–1572.

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology, 15, 1419–1439.

Weckerly, F. W. (1998). Sexual-size dimorphism: Influence of mass and mating systems in the most dimorphic mammals. Journal of Mammalogy, 79, 33–52.

Wooten, M. C., & Smith, M. H. (1985). Large mammals are genetically less variable. Evolution, 39, 210–212.

Yashima, A. S., & Innan, H. (2017). VARVER: A database of microsatellite variation in vertebrates. Molecular Ecology Resources, 17, 824–833.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Saitoh T. High variation of mitochondrial DNA diversity as compared to nuclear microsatellites in mammalian populations. Ecological Research. 2021;36:206–220. https://doi.org/10.1111/1440-1703.12190