

Genetic diversity and conservation of endangered animal species*

Ya-ping Zhang^{1,‡}, Xiao-xia Wang¹, Oliver A. Ryder², Hai-peng Li¹,
He-ming Zhang³, Yange Yong⁴, and Peng-yan Wang³

¹Laboratory of Molecular Evolution and Genome Diversity, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, P.R. China; ²Zoological Society of San Diego, San Diego, CA 92112, USA; ³Wolong Research and Conservation Center for the Giant Panda, Sichuan 623006, P.R. China; ⁴Foping National Nature Reserve, Shanxi 723400, P.R. China

Abstract: The loss of biodiversity resulting from extinctions is receiving increasing attention. Over several thousands of animal species have been evaluated and recognized as endangered species. Inbreeding depression has been demonstrated in many wild animal species. Here we sequenced 655–978 bp mitochondrial D-loop region of 32 individuals from four regional giant panda populations. Sixteen haplotypes were observed. AMOVE analysis demonstrated that genetic differentiation was not significant in the overall population, except the Qingling population. The current panda population may recover from a recent severe bottleneck that occurred about 43 000 years ago. Combining with our results on two endangered snub-nosed monkey species and one common hare species, different scenarios for low genetic variation have been discussed. Our results suggest that low genetic variation does not necessarily result from a recent bottleneck, and it is not necessarily an indication of the level of endangerment.

INTRODUCTION

The human population and human impact on natural habitat have significantly increased in the past few decades, therefore, it is well recognized that extinction now threatens a large number of species in the world [1]. May et al. estimated that impending extinction rates of animals and plants were at least four orders of magnitude greater than the background rate as judged by analyses of the fossil record [2]. Therefore, conservation of biodiversity becomes an urgent task for both governments and the public. With a good knowledge of the underlying processes of extinction, conservation efforts will be much more effective.

In the scientific effort for conservation of biodiversity, there has often been an emphasis on the importance of genetic factors in influencing extinction [3]. The cheetah is a good example. The systematic investigation of O'Brien's group on genetic diversity in the cheetah revealed low genetic variation in this species, which led to the conclusion that the lack of genetic variation is an important factor for the cheetah to be vulnerable to extinction (see [4]). However, controversy exists regarding whether genetic factors have been overemphasized in recommendations for conservation and management [3,5–7].

Lynch et al. showed that small populations might decline in fitness due to the accumulation of detrimental mutations [8]. Hedrick et al. suggested that low genetic variation in a species might be indicative of a recent population bottleneck, and such a bottleneck did potentially indicate vulnerabil-

*Lecture presented at the 3rd IUPAC International Conference on Biodiversity (ICOB-3), Antalya, Turkey, 3–8 November 2001. Other presentations are presented in this issue, pp. 511–584.

‡Corresponding author: E-mail: zhangyp@public.km.yn.cn

ity to extinction [9]. The more recent a bottleneck has been, the more we would expect the bottleneck to influence the future of the species. In small populations, genetic drift tends to reduce genetic variation, leading eventually to homozygosity and loss of evolutionary adaptability to environmental changes [10]. The knowledge of the population history and genetic structure of populations will be useful to develop sound conservation plans for endangered species.

The giant panda (*Ailuropoda melanoleuca*) is distributed only in China. One of the most widely recognized and cherished of all animal species, pandas have drawn much attention in conservation of endangered animals. About 1000 free-ranging pandas are threatened by extinction due to low members and highly fragmented distribution. Human activity is shrinking their habitat continually and cruelly. The wild pandas are subdivided into about 20 small populations by roads, rivers, and human settlements, which will confine the gene flow among populations and may cause inbreeding depression [11]. The information on the population genetic structure of the giant panda is limited.

As the most variable region in mitochondrial DNA, the control region (D-loop) was demonstrated to be an appropriate genetic marker for investigation of genetic variation in panda and bears [12–14]. Our preliminary study demonstrated low genetic variation in the giant panda [12].

In the present study, we sequenced 978 bp D-loop region, the entire D-loop except for the repeat region. 32 individuals from natural populations have been examined. Genetic variation of regional populations has been estimated. Combining the new data with our previous results on two endangered species and one common species, different scenarios for low genetic variation have been discussed.

MATERIALS AND METHODS

Thirty-two individuals from four large regional habitats were collected (Table 1). Total DNA was extracted from whole blood, hair, or dry skin. Because the conventional approach for collection of genetic materials, such as blood or fresh tissue, from the giant pandas may not be feasible in most cases, we made an effort to extract DNA from hair samples collected through a noninvasive approach [15] and from dry skin samples by the standard phenol/chloroform methods. Extraction without any material was used as a negative control in PCR amplification.

Table 1 Origins and haplotypes of the giant panda (*Ailuropoda melanoleuca*).
The studbook number is marked with #.

No.	Original location	Mountain system	Haplotype	Sample type
#202	Baoxing	Qionglai	1	Hair DNA
#247	Baoxing	Qionglai	3	Hair DNA
#390	Baoxing	Qionglai	3	Hair DNA
#308	Baoxing	Qionglai	5	Hair DNA
#320	Baoxing	Qionglai	7	Hair DNA
#329	Baoxing	Qionglai	8	Hair DNA
#371	Baoxing	Qionglai	9	Hair DNA
#121	Baoxing	Qionglai	11	Hair DNA
#41	Baoxing	Qionglai	1	Dried skin
#305	Baoxing	Qionglai	8	Blood
#414	Baoxing	Qionglai	3	Blood
#432	Baoxing	Qionglai	1	Blood
#382	Wolong	Qionglai	5	Blood
#343	Yuexi	Liangshan	1	Hair DNA
#386	Mabian	Liangshan	10	Hair DNA
#374	Leibuo	Liangshan	3	Blood
#230	Qingchuan	Minshan	2	Hair DNA

(continues on next page)

Table 1 (Continued)

No.	Original location	Mountain system	Haplotype	Sample type
#310	Qingchuan	Minshan	6	Hair DNA
#296	Nanping	Minshan	4	Hair DNA
#357	Nanping	Minshan	16	Blood
#298	Baishuijiang	Minshan	1	Blood
p11	Foping	Qinling	12	Dry skin
p12	Foping	Qinling	12	Dry skin
p14	Foping	Qinling	12	Dry skin
p17	Foping	Qinling	12	Dry skin
p20	Foping	Qinling	12	Dry skin
p21	Foping	Qinling	12	Dry skin
p13	Foping	Qinling	13	Dry skin
p16	Foping	Qinling	13	Dry skin
p18	Foping	Qinling	13	Dry skin
p15	Foping	Qinling	14	Dry skin
p8	Foping	Qinling	15	Dry skin

Six primers were designed based on a published sequence [16] and the new sequence generated in the present study (Table 2). The PCR was done under the condition of 95 °C for 2 min, then 38 cycles at 95 °C for 50 s, 53 °C for 1 min, and 72 °C for 1 min followed by a completion reaction at 73 °C for 5 min. The PCR productions were directly sequenced using the 377 automatic sequencer (Perkin-Elmer) with the Bigdye™ Terminator cycle sequencing kits (PE Biosystems).

Table 2 The DNA sequence of primers.

Primer name	Primer sequence
BED	5'-CTCCACTACCAGCACCCAAAG-3'
BEDH	5'-GGGTGATCTATAGTGTATGTCC-3'
R2L	5'-CTTCAAGAAGCTTACATATAC-3'
R2	5'-TCTAGGCATTTTCAGTGCCTTGC-3'
BEDL225	5'-ATGTACATACTGTGCTTGGC-3'
BEDh470	5'-GTCATTAGTCCATCGAGATG-3'

Sequences were aligned using the MegAlign module of DNASTar and adjusted by eye. Analysis of molecular variance (AMOVA) [17] was calculated using ARLEQUIN 1.1. A reduced median network was constructed by network 2.0b [18].

Both Tajima's [19] and Fu's [20] neutrality tests have been performed to examine the neutral mutation hypothesis under the assumption of population stationarity (i.e., constant size).

The time of the most recent common ancestor (MRCA) of the mtDNA haplotype was estimated following the method of Sailard et al. [21]. The scale $\rho = 1$ transition in 20 180 years within region 16090–16365 in human mtDNA [22] was used in converting the mutational scale into evolutionary time, to show a time scale of the MRCA of the haplotypes identified.

RESULTS

For the 11 skin specimen samples, we have obtained 655 bp sequence of the 5' control regions. For other samples, we have obtained 978 bp (655 bp plus 323 bp) of the control region (Figs. 1 and 2). The repetitive region of D-loop consists of tandem 10-base motifs. Our direct sequencing showed that each

1	ATACTATAAATCCAC	CTCTCATTTTATTCA	CTTCATACATGCTAT	TACACACTCTGTGCC	ATCATAGTATGTTTT	CATACATCCTCCCTT
2	T.....
3
4
5	T.....
6
7	T.....
8	T.....
9	T.....
10	T.....
11	T.....
12	T.....
13	T.....
14	T.....
15	T.....
16
1	CTTTC-ACACCCTAT	GTATATCGTACATTA	ATGGTGTACCCCCC	-T-CCCCCTATGTAT	ATCGTGCATTAATGG	CGTGCCCCATGCATA
2	C. C.....
3
4	C. C.....
5 C.....
6	C. C.....
7	C. C.....
8	C. C.....
9	C.....
10	C.....
11	C.....
12	G.....	C.....
13
14	C.....
15	C.....
16	C.....
1	TAAGCATGTACATAC	TGTGCTTGGCTTTAC	ATGAGGATACTCATT	ACAAGAACTTATTTTC	AAGCGATAGTCTATG	AGCATGTATTTCACT
2
3
4
5
6
7
8
9
10
11
12
13	T.....
14
15
16
1	TAGTCCAAGAGCTTG	ATCACCAAGCCTCGA	GAAACCAGCAATCCT	TGCGAGTACGTGTAC	CTCTTCTCGCTCCGG	GCCATAAATTTGTGG
2	G.....
3
4
5	G.....
6
7
8	G.....
9	G.....
10
11
12
13
14
15	G.....
16

(continued on next page)

Fig. 1 Alignment of 16 haplotype sequences of the giant panda in the 655-bp D-loop. Dots indicate agreement with haplotype 1 depicted in the first row.

Fig. 1 (Continued)

1	GGGTTTCTATACTGA	AACTATACCTGGCAT	CTGGTTCTTACCTCA	GGGCCATGTTAGCGT	CAACTCAATCCTACT	AACCCTTCAAATGGG
2T.....
3
4
5T.....
6
7
8
9
10
11
12T.....
13
14
15
16

1	ACATCTCGATGGACT	AATGACTAATCAGCC	CATGATCACACATAA	CTGTGGTGCATGCA	TTTGGTATTTTTTAA	TTTT-AGGGGGGAA
2
3
4
5
6
7
8
9
10
11
12T.....
13
14T.....
15
16

1	CTTGCTATGACTCAG	CTATGACCGTAAAGG	TCTCGTCGCAGTCAA	ATCAGCTGTAGCTGG	GCTTATTCATCTTTC	GAGGCTCCTCATGGA
2A.....
3
4
5A.....
6G.....
7A.....
8A.....
9A.....
10A.....
11A.....
12A.....
13A.....
14A.....
15G.....
16G.....

1	CACCCATAAGGTGCA	ATTCAGTCAA
2
3	T.....
4	T.....
5
6	T.....
7
8
9
10
11
12
13
14
15	T.....
16	T.....

```

11   CTTCAAGAAGCTTAC ATATACTTATGGATG TCCTGCCAAACCCCA AAAACAAGACTAAAT ATATGCGCAAACATG AAGTCACTTACACCT
others .....

11   AAACCGATATAATTA AGCTAACCCCCAGC CAATGTTGCAACAAC TACGGACATGGGACT CTAATTTTAATTTA TCTATAGATATTTTT
others .....

11   CTTTTACTGTGTCTG CCCAGCATTGATTTT TTAAGTATCATTATT GCACACCACCAATTT CCATTGAGCTATTTC ACATGAGTTCCAAAT
others ..... T.....

11   CAATTATGTTTCATGT AGCTTAACGAATAAA GCAAGGCACTGAAAA TGCCTAGA
others .....

```

Fig. 2 Alignment of haplotype 11 and other individuals in the 323bp D-loop. Dots indicate identity with haplotype 11. The second sequence was shared by other haplotypes.

panda possessed at least 29 copies of the motifs. However, we could not determine the sequence of entire repetitive region exactly from many individuals, probably due to heteroplasmy or artificial replication slippage by PCR. Therefore, we did not present the sequence data of repetitive region in this study.

Only one transition was observed in one individual from Qianglai in the 323 bp of the 3' D-loop region among 22 individuals from Liangshan, Qionglai, and Minshan. So we only used the 655 bp sequence of the 5' control region for the network analysis of the haplotypes.

Thirteen variable sites were found in the 655 bp sequence of the 5' control region, among which 4 are indels and 9 are transitions. These variable sites define 16 haplotypes. We did not find any transversion in the 978 bp D-loop region, while 10 transitions were observed. The cause for such transition bias for the D-loop, a noncoding region, is not clear.

Seven haplotypes were observed in 12 individuals from Baoxing (belonging to the Qionglai mountain system) and 4 haplotypes existed in 11 pandas from Foping (located in the Qinling mountain system). Haplotypes 12, 1, and 3 are the major types and shared by 6, 5, and 4 individuals, respectively.

The median network generated by a table of binary data may include most parsimonious trees supported by the data. An approach using networks rather than trees can present concise and comprehensible information about consensus sequences, homoplasmy, haplogroups, and so on [23]. A possible relationship among 16 different haplotypes was estimated as a reduced median network shown in Fig. 3. The deletion and insertion sites were not included in the network analysis, and all the sites were treated with equal weights. Each node represents an observed haplotype, and the area of the circle is proportional to the number of individuals with that haplotype. The intersections represent the hypothetical unknown haplotypes. The numerous sequences 1 and 3 tend toward the center of the reticulation. The individuals possess them were found in Liangshan, Qionglai, and Minshan. Our network did not show much geographical pattern, except for Qingling population.

The time of the most recent common ancestor (MRCA) of the mtDNA haplotype was estimated [21]. The ρ is about 2.215. Using the rate $\rho = 1$ transition in 20180 years in human mtDNA [22], the age of the most recent common ancestor of the current panda populations is estimated to be 43000 years before present.

The geographic structure of sequence variation was analyzed from the nucleotide differences observed between Qionglai, Qinling, Minshan, and Liangshan. Genetic structure of sequence variation was studied using the AMOVA approach [17]. When the Qinling population was compared with other regional populations, the percentage of the total variance explained by the partition of the data into different populations was 23.44 % with a significant probability value ($P < 0.05$), which suggests limited gene flow between Qingling and other populations. However, AMOVA failed to detect significant differences among other three regional populations.

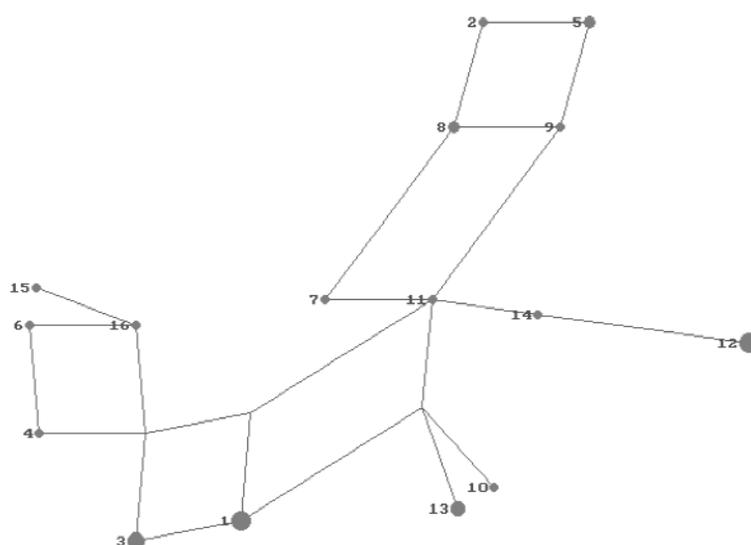


Fig. 3 The reduced median network for 16 haplotypes. Sequences are represented by circles, with area proportional to their frequency. Intersection represent hypothetical sequences, which have not been found in the sequencing exercise.

We use two methods to test the neutral mutation hypothesis under the assumption of constant population size. The Tajima's D value is 1.23, and $p > 0.1$, which suggests that the assumption of constant population size has not been rejected by Tajima's test [19]. However, the Fu's F_s value is -7.30 , and $p = 0.002$, which suggest that the assumption of constant population size has been rejected by Fu's test. Because the Fu's test is more powerful than Tajima's test, our results suggest that the giant panda population may have experienced population growth over the course of its history.

DISCUSSION

Population history of the giant panda

Su et al. [24] found low genetic variation in the giant panda compared with that in the black bear (*Selenarctos thibetanus*) by allozyme electrophoresis. In the present study, only 13 variable sites were found in the 655 bp sequence of the 5' control region, while 38 variable sites were observed in the 695 bp control region in the Japanese brown bear (*Ursus arctos*) [14]. Our sequence data also demonstrate low genetic variation in the giant panda.

The observed low genetic variation in the giant panda could be explained by a recent severe population bottleneck or a metapopulation structure.

The fossil record demonstrated that the body size of the giant panda was very small in the early Pleistocene, and became bigger in the middle to late Pleistocene. The body size reached the biggest by the subspecies *baconi*, and then the size reduced a little to reach the present size [11]. In the middle to late Pleistocene, the giant panda was distributed widely in Southeastern China, and south to Burma and Northern Vietnam. In the historically written record, the giant panda was found in many provinces in China besides the current habitat, such as Henan, Hubei, Hunan, Guizhou, and Yunnan Provinces. It appears that the giant panda lost much of its habitat in the course of the development of human activity. The wide distribution of panda population in the history is not consistent with the metapopulation structure scenario.

The median network (Fig. 3) demonstrated that all the haplotypes observed in this study are closely related with each other. The coalescent time of the current panda populations is estimated to be 43 000 years before present. Even though this estimate may not be accurate because of the possible rate difference between human and panda D-loop, the magnitude of the estimates is likely to be correct. Our observation is consistent with the scenario of a severe bottleneck around 43 000 years ago. In such a scenario, a historical population expansion signal is expected, considering the wide distribution of panda population in the historical record. Bottlenecks typically result from environmental change. Therefore, such a bottleneck would occur not only in the giant panda, but also in some other mammals in this region.

Interestingly, our sequence data do suggest that the giant panda population has experienced population growth in the history. As discussed below, we have identified at least one primate species that suffered from a bottleneck around this time scale. All these data support the scenario of recent bottleneck.

Scenarios for low genetic variation in the snub-nosed monkeys and the Yunnan hare

The genus *Rhinopithecus* that consists of four species is only found in Asia. One member species *R. roxellana*, the Sichuan snub-nosed monkey, is only found in the central part of mainland China and overlaps partly with the giant panda. Similar to the situation of giant panda, fragmented and deteriorating habitat has severely threatened the existence of *R. roxellana* in the near future. *R. bieti* is only found in southwestern China and is in a worse situation. Current population sizes of *R. roxellana* and *R. bieti* are about 10 000–20 000 and 1000–1500, respectively [25].

No polymorphism has been detected in 44 allozyme loci from 32 individuals of *R. roxellana*. This exceedingly low polymorphism compared with that of other nonhuman primates is surprising, particularly considering that the current population size is many times larger than some other endangered species. Fossil records as well as the high prevalence of dental agenesis indicated that a late Pleistocene bottleneck occurred for the species. Our coalescent simulation suggested that the most recent severe bottleneck could have happened within the last 20 000 years with population size at bottleneck most likely around a few hundred individuals [26]. It appears that both *R. roxellana* and the giant panda have suffered from recent bottleneck, even though the current population size of *R. roxellana* is about 10–20 times bigger than the giant panda.

Genetic variation in several individuals of *R. bieti* was documented with allozyme electrophoresis and mitochondrial cytochrome b gene sequence [27,28]. Without conducting analytical tests, Su et al. suggested that surviving *R. bieti* is a remnant population having gone through a bottleneck [27]. Genetic variation in *R. bieti* is much higher than that in *R. roxellana* and the giant panda, even though its population size is only slightly higher than the giant panda. Because of the small sample size, we cannot do any rigorous test on different demography scenarios. We did not find support for recent severe bottleneck. The exponential decline model has been rejected by our simulation study, which suggests that an obvious population decline of *R. bieti* which happened recently is unlikely.

The Yunnan hare (*Lepus comus*) occurs throughout the Yunnan-Guizhou Plateau in China, and three subspecies have been recognized. It is a very common species in the area and a main animal for hunting. Surprisingly, no sequence variation in the mitochondrial cytochrome b gene was observed in 55 individuals from all three subspecies, which is the lowest in the mammal species reported. Slow mutation rate explanation has been rejected in the Yunnan hare. The possible explanation for the low genetic variation in this very common species is the recent origin of the species [29].

Conservation applications

Low genetic variation in a species may be an indication of a recent population bottleneck, and such a bottleneck could result in inbreeding depression. Species with low genetic variation may be more vul-

nerable to environment change, and consequently is vulnerable to extinction. However, our results demonstrate that low genetic variation does not necessarily result from recent bottlenecks, and it is not necessarily an indication of the level of endangerment. An increased knowledge of population demography could provide useful information for conservation management. For example, the demographic histories of both the giant panda and the Sichuan snub-nosed monkey demonstrated the ability of these species to expand if the habitat is restored, because they went through the severe Pleistocene bottleneck. Indeed, the recent field survey showed that the giant panda population in the well-protected areas recovered some, which is a good sign for the conservation effort.

Because of the severe habitat fragmentation, some isolated populations with small population numbers are vulnerable to extinction. Building the "Green Corridor" with the hope to promote gene flow among panda populations has been considered in conservation plans. Our data show that there is extensive gene flow among Qionglai, Minshan, and Liangshan populations. Therefore, building a corridor among these three regions may be feasible solely from a genetic viewpoint. However, the gene flow between Qingling and other regions appears to be limited. More careful and detail studies will be necessary to define if the Qingling population is a distinct evolutionary unit [30]. Such information is also necessary to make a sound breeding program for a captive population.

ACKNOWLEDGMENTS

This work is supported by State Key Basic Research and Development Plan (G2000046806), Program for Key International S&T Cooperation Project of China (2001CB711103), Chinese Academy of Sciences (KSCX 2-1-05), and NSFC.

REFERENCES

1. O. A. Ryder, A. McLaren, S. Brenner, Y-P. Zhang, K. Benirschke. *Science* **288**, 275 (2000).
2. R. M. May, J. H. Lawton, N. E. Stock. In *Extinction Rate*, J. H. Lawton and R. M. May (Eds.), Oxford University Press, Oxford (1994).
3. P. W. Hedrick. In *Molecular Genetic Approaches in Conservation*, T. B. Smith and R. K. Wayne (Eds.), pp. 459–477, Oxford University Press, New York (1996).
4. S. J. O'Brien. *Proc. Natl. Acad. Sci. USA*. **91**, 5748 (1994).
5. T. M. Caro and M. K. Laurenson. *Science* **263**, 485 (1994).
6. K. R. Crooks, M. A. Sanjayan, D. F. Doak. *Conserv. Biol.* **12**, 889 (1998).
7. W. Scheffrahn, J. Tomiuk, M. Bayes. *Folia Primatol.* **69** (suppl.), 117 (1998).
8. M. Lynch, J. Conery, R. Burger. *Am. Nat.* **146**, 489 (1995).
9. P. W. Hedrick, R. C. Lacy, F. W. Allendorf, M. E. Soule. *Conserv. Biol.* **10**, 1312 (1996).
10. R. Lande. *Science* **241**, 1455 (1988).
11. J-C. Hu. *Research and Progress in Biology of the Giant Panda*, Sichuan Publishing House of Science & Technology (1990).
12. Y-P. Zhang, O. A. Ryder et al. *Science in China (Series C)* **27**, 139 (1997).
13. L. P. Waits, S. L. Talbot, R. H. Ward, G. F. Shields. *Conservation Biology* **12**, 408 (1998).
14. T. Matsuhashi, R. Masuda, T. Mano, M. C. Yoshida. *Mol. Biol. Evol.* **16**, 676 (1999).
15. Y-P. Zhang and O. A. Ryder. *Zoo Biol.* **14**, 569 (1994).
16. Y-P. Zhang and O. A. Ryder. *Mol. Phylogenet. Evol.* **3**, 351 (1994).
17. L. Excoffier, P. E. Smouse, J. M. Quattro. *Genetics* **131**, 479 (1992).
18. A. Röhl. NETWORK: A program Package for Calculating Phylogenetic Network. Version 2.0b (1999).
19. F. Tajima. *Genetics* **123**, 585 (1989).
20. Y. X. Fu. *Genetics* **147**, 915 (1997).
21. J. Saillard, P. Forster, N. Lynnerup, H-J. Bandelt, S. Nørby. *Am. J. Hum. Genet.* **67**, 718 (2000).

22. P. Forster, R. Harding, A. Torroni, H-J. Bandelt. *Am. J. Hum. Genet.* **59**, 935 (1996).
23. H-J. Bandelt, P. Forster, B. C. Sykes, M. B. Richards. *Genetics* **141**, 743 (1995).
24. B. Su and L-M. Shi. *Chinese Sci. Bull.* **39**, 742 (1994).
25. S. L. Ma and Y. Wang. *Acta Theriologica Sinica* **8**, 250 (1988).
26. H. P. Li, S. Men, Z. M. Men, Y. X. Fu, Y-P. Zhang. *Genetics* (2002). Submitted for publication.
27. B. Su and L. M. Shi. *Conservation Biol.* **9**, 847 (1995).
28. Y. P. Zhang and O. A. Ryder. *Primates* **39**, 39 (1998).
29. C. H. Wu, Y. X. Wang, Y.-P. Zhang. *Biochem. Genetic.* **38**, 149 (2000).
30. O. A. Ryder. *Trends Ecol. Evol.* **1**, 9 (1986).