Genetic Viability and Population History of the Giant Panda, Putting an End to the "Evolutionary Dead End"?

Baowei Zhang,* Ming Li,* Zejun Zhang,*† Benoît Goossens,‡ Lifeng Zhu,* Shanning Zhang,* Jinchu Hu,† Michael W. Bruford,‡ and Fuwen Wei*

*Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, the Chinese Academy of Sciences, Haidian, Beijing, People's Republic of China; †Institute of Rare Animals and Plants, China West Normal University, Nanchong, Sichuan, People's Republic of China; and ‡Biodiversity and Ecological Processes Group, School of Biosciences, Cardiff University, Cardiff, United Kingdom

The giant panda (*Ailuropoda melanoleuca*) is currently threatened by habitat loss, fragmentation, and human persecution. Its dietary specialization, habitat isolation, and reproductive constraints have led to a perception that this is a species at an "evolutionary dead end," destined for deterministic extinction in the modern world. Here we examine this perception by a comprehensive investigation of its genetic diversity, population structure, and demographic history across its geographic range. We present analysis of 655 base pairs of mitochondrial (mt) control region (CR) DNA and 10 microsatellite loci for samples from its 5 extant mountain populations (Qinling, Minshan, Qionglai, Liangshan, and Lesser Xiangling). Surprisingly, extant populations display average to high levels of CR and microsatellite diversity compared with other bear species. Genetic differentiation among populations was significant in most cases but was markedly higher between Qinling and the other mountain ranges, suggesting, minimally, that the Qinling population should comprise a separate management unit for conservation purposes. Recent demographic inference using microsatellite markers demonstrated a clear genetic signature for population decline starting several thousands years ago or even futher back in the past, and being accelerated and enhanced by the expansion of human populations. Importantly, these data suggest that the panda is not a species at an evolutionary "dead end," but in common with other large carnivores, has suffered demographically at the hands of human pressure. Conservation strategies should therefore focus on the restoration and protection of wild habitat and the maintenance of the currently substantial regional genetic diversity, through active management of disconnected populations.

Introduction

The giant panda (*Ailuropoda melanoleuca*) is often cited as one of the most endangered mammals in the world. It is a flagship species for global conservation and an endemic species to China, living in small, isolated populations with a severely fragmented distribution at the present time. As a result, this species has been a top priority for conservation action in China since the 1950s. However, the giant panda had a much wider range in the Pleistocene, with fossil records from Zhoukoudian, near Beijing to southern China, and into northern Myanmar, northern Vietnam, Laos, and Thailand (Wen and He 1981; Zhu and Long 1983; Hu et al. 1985; Wei et al. 1990; Hu 2001). Today, giant pandas are restricted to just 6 fragmented mountain ranges on the eastern edge of the Tibetan plateau, in Sichuan, Gansu, and Shaanxi Provinces (Hu 2001).

Fossil data indicate that this species originated in late Miocene, probably expanded during the middle Pleistocene, and started to contract in the late Pleistocene (Pei 1974; Wang 1974; Qiu and Qi 1989). The panda has been often described as a species whose body size, dietary specialization (living mainly on bamboo species, Hu and Wei 2004), reproductive constraints, and morphological adaptation have led it to an "evolutionary dead end" (Pei 1965, 1974; Wang 1974; Moran 1988; Wei et al. 1990; Huang 1993; Gittleman 1994). However, climate change (quaternary ice age oscillations), human activities (e.g., habitat destruction and fragmentation, poaching), and natural catastrophes (e.g., simultaneous large-scale bamboo flowering and die-off events) have also been cited as responsible for accelerating this "natural process," driving the species toword its supposedly inevitable extinction in the near future (Pei 1965, 1974; Wang 1974; Wei et al. 1990).

Researchers have previously attempted to examine the giant panda's evolutionary potential using molecular genetic approaches, by assessing its patterns of genetic diversity. Low genetic variability has been detected in giant panda populations using protein electrophoresis (Su et al. 1994), mtDNA control region sequence analysis (Zhang et al. 1997), and classical DNA fingerprinting (Fang et al. 1997), and as a result, its population decline has been partly ascribed to genetic factors, i.e., the relative absence of genetic variation (Zhang and Su 1997).

Effective conservation and management strategies for endangered species incorporate an understanding of the evolutionary and demographic history of populations and their present genetic structure, including the existence of demographic partitions throughout the geographical range and the distribution of genetic diversity (Avise 1989; Smith, Bruford, and Wayne 1993; O'Brien 1994; Beaumont and Bruford 1999; Groombridge et al. 2000). The application of a mixed approach, evaluating patterns of individual genotypes, allele frequencies, and demographic hypothesis testing using exploratory analysis is becoming a powerful approach to augment descriptive genetic analysis for inference of demographic dynamics and structure at the landscape scale, as is increasingly needed by conservation managers today (Manel et al. 2003; Bertorelle et al. 2004). To date, however, there have been few such efforts to infer detailed demographic processes from patterns of genetic variation and population structure in the giant panda.

Here we used mitochondrial DNA (mtDNA) control region (CR) sequences and analysis of 10 microsatellite loci to
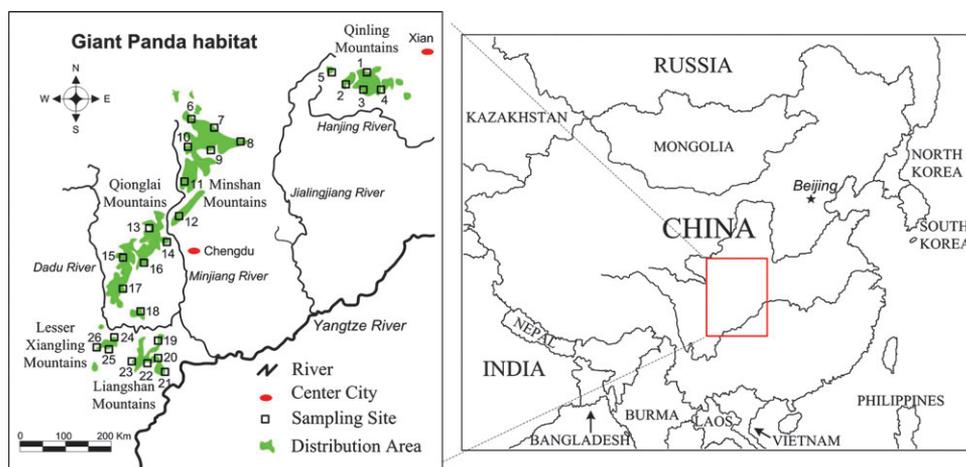
F<small>IG</small>. 1.—Map of the giant panda distribution and sampling locations of the present study. The sampling locations are indicated by black empty squares. Numerals indicate sample sites: (1) Taibai, (2) Yangxian, (3) Liuba, (4) Foping, (5) Fengxian, (6) Nanping, (7) Wenxian, (8) Qingchuan, (9) Pingwu, (10) Songpan, (11) Maoxian, (12) Dujiangyan, (13) Wenchuan, (14) Chongzhou, (15) Baoxing, (16) Lushan, (17) Tianquan, (18) Hongya, (19) Erbian, (20) Mabian, (21) Leibo, (22) Meigu, (23) Yuexi, (24) Shimian, (25) Mianning, (26) Jiulong.

analyze genetic variation and explore the processes shaping genetic diversity in giant panda populations. Our results are used to evaluate whether the history and current genetic status of giant pandas are consistent with an "evolutionary dead-end" hypothesis, and to provide insights relevant for current and future conservation efforts in this high-profile species.

Materials and Methods
Sample Collection

Blood, tissue, skin, hair, and fecal samples of wild-born giant pandas were collected from zoos, the field, and museums, representing 26 geographic locations in the 6 mountain ranges (fig. 1). The samples from the field were confirmed as different individuals by individualization using microsatellites (Zhan et al. 2006). One hundred and sixty-nine individuals were analyzed in this study, comprising 3 muscle samples, 2 blood samples, 97 faecal samples, 3 hair samples, and 64 dried skins from museums (for details see supplementary table 1, Supplementary Material online). Most samples came from the major populations residing in 3 mountain ranges: Qinling (QIN, n = 41), Minshan (MIN, n = 50) and Qionglai (QIO, n = 58); but we were also able to include samples from the 2 smaller populations, Lesser Xiangling (XL, n = 11) and Liangshan (LIS, n = 9). A sample collected from Daxiangling Mountain was considered part of the Qionglai population for analysis (Location 18 in fig.1).

To test the robustness of our sampling, the software GENESAMP (Sjögren and Wyöni 1994) was implemented to calculate the sample size necessary to detect alleles at frequencies of 0.05 and 0.10 with 95% probability for CR and microsatellite markers, respectively.

Laboratory Methods

Genomic DNA was extracted from tissue, blood, or skin using SDS-phenol/chloroform (Sambrook, Fritsch, and Maniatis 1989). Each sample was homogenized overnight at 55°C in extraction buffer (EDTA 0.1 M, Tris-HCl 0.05 M pH = 8.0, SDS 1%, and Proteinase K 0.2 mg/ml). Subsequently, DNA was purified with a standard phenol/chloroform extraction protocol followed by an ethanol precipitation. Genomic DNA was isolated from plucked hair using Chelex® 100 (Walsh, Metzger, and Higuchi 1991), and the QIAamp® DNA Stool Mini Kit (QIAGEN GMBH, Hilden, Germany) was used for fecal samples according to the manufacturer's protocol. Extraction blanks were used as negative controls in downstream PCR amplifications.

The 5′ end of the mitochondrial control region was amplified by polymerase chain reaction (PCR) (35 cycles of 1 min at 94°C, 1 min at 55°C, and 1 min at 72°C) using the following primers: P-tp (5′- CTC CCT AAG ACT CAA GGA AG -3′, forward primer designed in present study), BEDL225 (5′- ATG TAC ATA CTG TGC TTG GC -3′, forward primer), and BEDH (5′- GGG TGA TCT ATA GTG TTA TGT CC -3′, reverse primer, Zhang et al. 2002). PCR fragments were purified using the E.Z.N.A. Cycle-pure Kit (Omega) and sequenced using the P-tp, BEDL225 and BEDH primers and the BigDye Terminator Sequencing Ready Reaction V3.0 kit (Applied Biosystems), following the manufacturer's instructions, using an ABI 3700 semi-automated DNA analyser.

Eighteen giant panda microsatellite loci (Lu et al. 2001) were initially assessed with 10 loci (Ame-μ05, Ame-μ10, Ame-μ11, Ame-μ13, Ame-μ15, Ame-μ16, Ame-μ24, Ame-μ25, Ame-μ26, Ame-μ27) being selected for this study on the basis of their efficiency using DNA from different sources, polymorphism, and yield. PCR amplification of 50 cycles was carried out for up to 4 loci simultaneously, with combinations selected based on fragment size range, Tm, and fluorescent dye (FAM, TET, or HEX), using the QIAGEN Mutliplex PCR kit according to the manufacturer's protocol at optimized annealing temperatures. The PCR products were separated on a polyacrylamide gel in ABI PRISM™ 377 DNA sequencer

Table 1
Genetic Variability Observed Within Populations of Giant Panda Using Mitochondrial Control Region and 11 Microsatellite Loci

Population	N(CR/MS)	H	$H(\pm$ SD$)$	S	$K(\pm$ SD$)$	$\Pi(\pm$ SD$)$	MNA	H_E	H_O	AR	F_{IS} (95% CI)
QIN	36/32	8	0.753 ± 0.058	9	2.55 ± 1.40	0.0039 ± 0.0024	3.50	0.486	0.525	2.58	−0.064 (−0.151 to −0.016)*
MIN	44/29	17	0.926 ± 0.019	11	3.28 ± 1.72	0.0051 ± 0.0029	4.80	0.559	0.561	3.03	0.019 (−0.079 to 0.062)*
QIO	63/40	18	0.885 ± 0.021	13	2.80 ± 1.50	0.0043 ± 0.0026	5.30	0.610	0.595	3.18	0.040 (−0.047 to 0.092)*
LIS	10/4	8	0.956 ± 0.059	8	2.36 ± 1.40	0.0036 ± 0.0024	2.30	0.366	0.425	2.30	−0.069 (0.013 to −0.148)*
XL	6/10	3	0.733 ± 0.155	3	1.40 ± 0.99	0.0022 ± 0.0018	4.00	0.635	0.685	3.29	−0.020 (−0.429 to 0.020)*
Total	159/115	39	0.943 ± 0.007	19	3.27 ± 1.69	0.0050 ± 0.0029	7.10	0.642	0.565	6.91	0.124 (0.081 to 0.160)*

NOTE.—N = number of individuals; CR = mitochondrial control region; MS = microsatellite; h = Number of haplotypes; H = Haplotype diversity; S = number of polymorphic sites; κ = the average number of pairwise nucleotide differences; π = nucleotide diversity ± variance. N = sample size; MNA = mean number of alleles per microsatellite locus; H_O and H_E = observed and expected heterozygosity. AR = allelic richness. Values of the inbreeding coefficient (F_{IS}) are reported with their 95% confidence intervals (CI). The asterisks (*) indicate significant ($P < 0.001$) departures from Hardy-Weinberg equilibrium, as estimated using GENEPOP.

with Tamra GS350 marker. All gels were analyzed using Genescan™ 2.0 and Genotyper® 2.0. Amplification was repeated minimally 3 times. We scored samples as heterozygous at 1 locus if both alleles appeared at least twice among replicates and as homozygous if all the replicates showed identical homozygous profile. If neither of those cases applied, we treated the alleles as missing data. Only individuals for which we obtained allele data information for 6 or more loci were included in the analysis, and therefore, 115 individuals were analyzed for microsatellites.

The CR sequences of 12 giant pandas (comprising MIN 3, QIO 7, LIS 2, supplementary table 2, Supplementary Material online) were used from Zhang et al. (2002) and analyzed together with 147 new sequences (QIN 36, MIN 41, QIO 56, LIS 8, XL 6). Sequences for each individual and consensus haplotypes were aligned using SeqManII in DNAStar (Burland 2000) and rechecked by eye. MtDNA CR sequences were aligned using ClustalW (Thompson, Higgins, and Gibson 1994) and visually checked. Initial sequence comparisons and measures of variability were performed using MEGA3.0 (Kumar, Tamura, and Nei 2004), and unique haplotypes were verified using DAMBE (Xia and Xie 2001). Nucleotide diversity (π, Nei 1987), mean number of pairwise differences (κ, Tajima 1983), haplotype diversity (H, Nei 1987), and the number of polymorphic sites (S) were estimated using ARLEQUIN

2.0 (Schneider, Roesslo, and Excoffier 2000). Microsatellite variation was assessed by summary statistics, including mean number of alleles per locus (MNA), expected (H_E) and observed (H_O) heterozygosities, and inbreeding coefficient (F_{IS}), which were computed for each population using GENETIX (Belkhir 2004). Allelic richness, an estimate of allelic diversity that compensates for unequal sample size, was calculate using FSTAT (Goudet 2002) and averaged across loci. Deviations from Hardy-Weinberg equilibrium for each population were assessed using an exact test implemented in GENEPOP (Raymond and Rousset 1995).

For CR sequences data, molecular variance analysis (AMOVA, Excoffier, Smouse, and Quattro 1992) was performed to test for differentiation between geographical units. The significance of φ-statistics generated by AMOVA was tested by random permutations of sequences among populations. In addition, a minimum spanning haplotype network was constructed to depict the relationships among haplotypes and their distribution using TCS (Crandall and Templeton 1993, Clement, Posada, and Crandall 2000), considering gaps as a fifth state. For microsatellites, population differentiation was assessed by Fisher's exact test, θ (Weir and Cockerham 1984), and by analysis of molecular variance using GENETIX. The association between the estimates of $F_{ST}/1-F_{ST}$ (Rousset and Raymond 1997)

Table 2
Summarized Results of Hierarchical AMOVA and Population Differentiation Analysis

(a)	Groupings	ϕ_{CT}	P
Group 1	[QIN] [MIN QIO LIS XL]	0.200	0.189
Group 2	[QIN] [MIN] [QIO LIS XL]	0.014	0.303
Group 3	[QIN] [MIN] [QIO] [LIS XL]	−0.150	0.602
Group 4	[QIN] [MIN QIO] [LIS] [XL]	0.240	0.000
Group 5	[QIN] [MIN QIO] [LIS XL]	0.189	0.000
Group 6	[QIN] [MIN QIO XL] [LIS]	0.174	0.098
Group 7	[QIN] [MIN QIO LIS] [XL]	0.274	0.000

(b)	MIN	QIO	LIS	XL
QIN	0.223 ($P = 0.00$)	0.186 ($P = 0.00$)	0.251 ($P = 0.00$)	0.166 ($P = 0.00$)
MIN		0.079 ($P = 0.00$)	0.286 ($P = 0.00$)	0.146 ($P = 0.00$)
QIO			0.149 ($P = 0.10$)	0.055 ($P = 0.10$)
LIS				0.123 ($P = 0.30$)

NOTE.—(a) AMOVA for groupings for populations estimate using φ-statistics based on CR sequence. (b) Pairwise F_{ST} based on microsatellite data.

Table 3
Posterior Distributions of the N_0, N_1 and T (Year) under the Exponential Model of Change

	N_0 50% $\pm$ SD (5% to 95%)	N_1 50% $\pm$ SD (5% to 95%)	T (years) 50% $\pm$ SD (GT = 5 / GT = 6)
QIN	170 $\pm$ 14 (12 to 854)	13,719 $\pm$ 206 (3,732 to 55,590)	3,682 $\pm$ 334 / 4,206 $\pm$ 415
MIN	1161 $\pm$ 42 (232 to 4,064)	18,273 $\pm$ 604 (4,560 to 89,126)	17,558 $\pm$ 708 / 20,800 $\pm$ 1,671
QIO	550 $\pm$ 30 (42 to 2,089)	13,660 $\pm$ 226 (4,056 to 46,990)	5,791 $\pm$ 516 / 6,893 $\pm$ 406

Note.—The posterior distributions of N_0, N_1 and T were described by the 5%, 50%, and 95% quantile, with values of 1 standard deviation ($\pm$ SD) computed across 5 replicates. T was estimated using generation time as 5 and 6 years, respectively.

and land-based Manhattan distance were assessed using the Mantel test, implemented in the IBD software (Bohonak 2002); the statistical significance of the values was obtained by 10,000 randomization steps.

In contrast to the preceding analyses (based on mountain-defined populations), a Bayesian clustering method was used to detect structure in the whole dataset and assign individuals to inferred clusters using STRUCTURE (Pritchard, Stephens, and Donnelly 2000). Eight independent runs of $K = 1$ to 10 were performed at 1,000,000 Markov Chain Monte Carlo (MCMC) repetitions with a 100,000 burn-in period using no prior information and assuming correlated allele frequencies and admixture. The posterior probability K [$P(K/X)$] was then calculated, and the log likelihood was used to choose the most likely value for K. In addition, the statistic ΔK, the second-order rate of change in the log probability of the data between successive values of K, was also estimated (Evanno, Regnaut, and Goudet, 2005), since this method is able to detect the appropriate number of clusters for simulated data sets under a number of gene exchange models. Because one cannot evaluate ΔK for $K = 1$, we explored the probability of the $K = 2$ to $K = 9$ on the basis of above results. After the analysis was carried out, individual admixture proportions were sorted and displayed by individual using DISTRUCT (Rosenberg 2004).

Recent population bottlenecks can produce distinctive genetic signatures in the distributions of allele size, expected heterozygosity and in the genealogy of microsatellite loci (Cornuet and Luikart 1996; Beaumont 1999; Garza and Williamson 2001; Goossens et al. 2006). Here, two approaches were used to detect population contractions. First, the now commonly used heterozygosity test implemented in BOTTLENECK (Piry, Luikart, and Cornuet 1999) was used to detect for the signature of a recent demographic bottleneck assuming infinite allele (IAM), stepwise mutation (SMM), and two-phase mutation models (TPM) with various (70% to 95%) single-step mutations (Di Rienzo et al. 1994). In addition, evidence for demographic change was inferred using MSVAR1.3 (Storz and Beaumont 2002), which implements a coalescent simulation-based Bayesian likelihood analysis, assumes a strict SMM, and estimates the posterior probability distribution of population parameters using Markov Chain Monte Carlo simulation, based on the observed distribution of microsatellite alleles and their repeat number. The most important parameters are N_0, N_1, and T, where (1) N_0 is the current effective number of chromosomes; (2) N_1 is the number of chromosomes at some point back in time T, and (3) T is the time since the population change. The reported values of giant panda generation time are 4.5 years (Pan et al. 2001) and 5.5 to 6.5 years (Hu et al. 1985), so here we set it as both 5 and 6 years for simulation. In every simulation, we ran each chain with 100,000 thinned updates and a thinning interval of 10,000 steps, leading to a total number of 1×10^9 updates. The first 10% of updates were discarded to avoid bias in parameter estimation at starting conditions, and the remaining data were used to obtain the lower (5%), the median (50%), and the upper (95%) quantiles of the posterior distributions. Five independent simulations were run on subsamples of QIN, MIN, and QIO using both an exponential and a linear model. In all MSVAR simulations only 9 loci were used, since the locus Ame-μ13 did not show evidence for a regular SMM. The other two subsample groups (XL, LIS) were ignored for BOTTLENECK and MSVAR due to small sample size.

Results
Genetic Variation

Six hundred fifty-five base pairs of mtDNA CR sequence were analyzed from 159 individuals after alignment. A total of 24 variable nucleotide sites, comprising 2 transversions, 17 transitions, and 5 insertions/deletions, defined 39 haplotypes (supplementary table 3, Supplementary Material online). While a small number of haplotypes showed a wide geographic distribution (e.g., GH32, GH33; fig.2), most were more restricted. Twenty-six of 39 haplotypes were found in only a single population. In the total sample, high haplotype ($H = 0.943 \pm 0.007$) but relatively low nucleotide diversities ($\pi = 0.0050 \pm 0.0029$) were found (table 1). A total of 71 alleles were detected using the 10 microsatellite loci from 115 individuals (QIN 32, MIN 29, QIO 40, XL 10, LS 4), with a mean allele number per locus of 7.10 (range 2.30 in LIS to 5.30 in QIO). Among the 10 loci, the number of alleles observed per locus varied from 5 (Ame-μ15) to 12 (Ame-μ10), and allelic richness ranged from 2.30 (LIS) to 3.29 (XL), with the overall allelic richness across loci being 6.91 (table 1). Values of observed (H_O) and expected heterozygosity (under Hardy-Weinberg equilibrium, H_E) ranged from 0.425 (LIS) to 0.685 (XL) and 0.366 (LIS) to 0.635 (XL, table 1), respectively. The mean overall values of H_O and H_E were 0.565 and 0.642, respectively (table 1). The mean overall F_{IS} (0.124) deviated significantly from zero (95% CI 0.081 to 0.160), although this was not the case in any of 5 populations taken separately (table 1).

Results of power analysis using the program GENE-SAMP indicated that sample sizes were not large enough
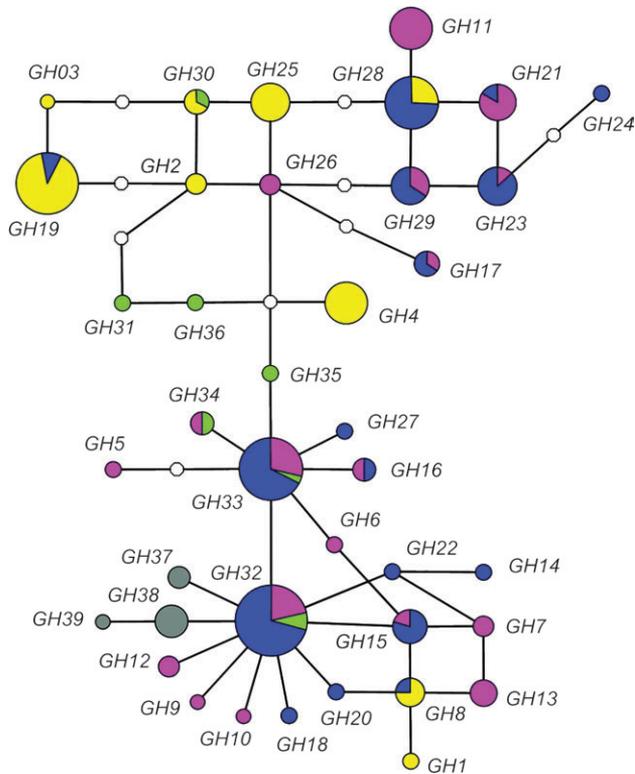
Fig. 2.—Haplotypes minimum spanning network based on statistical parsimony using TCS. In the network, the geographical origin of haplotypes is indicated by different colors (yellow-Qinling, purple-Minshan, blue-Qionglai, green-Liangshan, gray-Lesser Xiangling). Circles show the haplotype number and are proportional to the haplotype frequencies; empty circles indicate undetected intermediate haplotype states.

to reliably detect "low frequency" (0.1 or less) alleles for Liangshan and the Lesser Xiangling populations for both CR and microsatellites. In the samples for Qinling, Minshan, and Qionglai, our sampling was sufficient (with 95% probability) to detect such alleles for both CR and microsatellite except for CR haplotypes of frequency 0.05 or less in the Qinling and Minshan populations (data not shown).

Population Structure

The minimum spanning network for CR haplotypes produced a complex pattern (fig. 2), with the network com-

prising a multiple configuration including two star-like networks, with the central sequences being haplotypes GH32 and GH33. MIN, QIO, and LIS haplotypes are distributed broadly throughout the network; however the distribution of XL and QIN haplotypes are more localized. Although 39 haplotypes were detected, the network shows relatively shallow genetic divergence and little evidence for overt geographic structure. Analysis of molecular variance showed the highest variance that could be explained among populations relative to the entire sample occurred for the groupings: [QIN] [MIN, QIO, LIS] [XL] (ϕ_{CT} = 0.274, P = 0.000, table 2). Among the 5 geographical foci (mountain ranges), there was significant isolation-by-distance using Rousset's distance (Rousset and Raymond 1997) for the microsatellite data set (r = 0.502, P < 0.05).

Bayesian STRUCTURE clustering analysis using multilocus microsatellite genotypes revealed a maximum posterior probability of the genetic data of K = 4 (log P [K/X] = −2371.1; Supplementary fig. 1, supplementary material online). Moreover, the ΔK value showed a clear maximum (ΔK = 662) at K = 4 (Supplementary fig. 1). The DISTRUCT diagram shows a strong genetic structure pattern, broadly following the mountain origins of each sample but with some misassignment of individuals (vertical bars) (fig. 3; supplementary table 4, Supplementary Material online). QIN is predominantly a separate cluster (in yellow), with MIN (purple), QIO (blue), and XL-LIS (green) being clearly differentiated. Individual reassignment rates (q) are 0.911 ("yellow" individuals) for QIN, 0.776 ("purple" individuals) for MIN, 0.682 ("blue" individuals) for QIO, and 0.658 and 0.975 for XL and LIS ("green," fig. 3; supplementary table 4), respectively. The higher misassignment levels for MIN, QIO, and XL, and especially for QIO and its adjacent populations imply significant recent gene flow. In comparison, limited gene flow can be inferred between QIN and the other populations.

Recent Population Demography

BOTTLENECK analysis did not provide convincing evidence for a recent population decline. No significant heterozygosity excess or deficit was found in QIN or MIN following all 3 mutation models. The QIO population showed a heterozygosity deficit under the SMM. However, no significant recent population bottleneck, denoted by

Table 4
A Summary of Genetic Diversity Parameters of Microsatellite Data of Several Ursine Species

| Species	Researcher	MNA	H_E
American black bear *Ursus americanus*	Paetkau and Strobeck 1994	8.75	0.82
Arctic grizzly bear *U. arctos*	Craighead et al. 1995	7.63	0.749
Brown bear *U. arctos*	Paetkau et al. 1998	2.13 to 7.63	0.298 to 0.788
Polar bear *U. maritimus*	Paetkau et al. 1999	6.5	0.68
Brown bear *U. arctos*	Waits et al. 2000	6.8	0.71
Asian black bear *U. thibetanus*	Saitoh et al. 2001	2.00 to 4.17	0.30 to 0.45
Kermode bear *U. a. kermodei*	Marshall and Ritland 2002	4.2 to 7.5	0.664 to 0.793
Spectacled bear *Tremarctos ornatus*	Ruiz-Garcia 2003	4.6	0.382
Spectacled bear *T. ornatus*	Ruiz-Garcia et al. 2005	5.67	0.56
Giant panda *A. melanoleuca*	This study	7.1	0.642
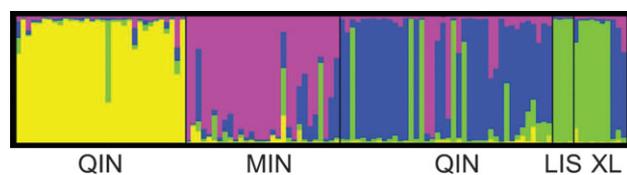
Fig. 3.—Bayesian cluster analysis of the microsatellite variation among 5 giant panda populations. The proportion of ancestry assigned to each of the 4 clusters was plotted by individuals. The 4 colors, yellow, purple, blue, and green, represent 4 genetic clusters.

heterozygosity excess, was detected in any locality under all 3 models.

However, Bayesian coalescent simulation provided compelling evidence for a more long-term population decline. In simulations using MSVAR, 5 independent replicates showed concordant results in QIN, MIN, and QIO. In these 3 populations, the posterior distribution of N_0 and N_1 did not overlap under an exponential or linear model (table 3; supplementary table 5, Supplementary Material online). Under the exponential model, the medians of the posterior distributions were approximately 170 (QIN), 1,161 (MIN), 550 (QIO) for N_0, and approximately 13,719 (QIN), 18,273 (MIN), 13,660 (QIO) for N_1 (table 3). Simulations under a linear model provided similar posterior distributions of N_0 and N_1 in the 3 populations (supplementary table 5). The MSVAR simulations also provided a consistent posterior distribution to date the point at which the population started to decrease. The median population decline time (T), for the two different generation times (5 and 6 years, respectively), was 3,682 and 4,206 (QIN), 17,558 and 20,800 (MIN), 5,791 and 6,893 (QIO) years, respectively (table 3). Under the linear model, however, even longer population decline times were inferred by MSVAR (supplementary table 5), although the linear model is considered more applicable to classical "open" populations and is unlikely to be appropriate for giant pandas (Storz and Beaumont 2002; Storz, Beaumont, and Alberts 2002).

Discussion
Genetic Diversity

Our data reveal surprisingly rich genetic variability in the 5 extant giant panda populations for both nuclear and mitochondrial DNA (table 1). Genetic variation at the 10 microsatellite loci analyzed can be compared with levels of diversity observed in surveys of other ursine species (table 4), although these results are not directly comparable because they used different microsatellite loci. The estimator expected to be the least biased of those compared here (average expected heterozygosity) was ranked mid to high for giant pandas in these studies (table 4).

Our results differ with previous studies on the level of genetic diversity of giant pandas. Compared with studies of Lu et al. (2001) and Zhang et al. (1997, 2002), many more sequence variants (CR haplotypes, microsatellite alleles) were found, and a higher gene diversity was observed in our data set. In most previous studies, low genetic diversity found was ascribed to a recent severe bottleneck (Su et al.

1994; Zhang et al. 1997, 2002), and the lack of genetic diversity was further considered to be a contributory factor to the endangered situation of the giant panda (Zhang and Su 1997), reinforcing the notion of an "evolutionary dead-end species."

In contrast, our results imply that the giant panda does not possess low genetic diversity, that the tag of being genetically impoverished is unwarranted, and that diminished genetic diversity cannot obviously account for its endangered status. Earlier studies may have failed to uncover the genetic richness of this species due to low sample sizes and/or because of the more sensitive molecular markers applied in the present study. Although evidence for a population reduction is clearly apparent, this has not been severe enough either to be detected using heterozygosity tests designed to uncover very recent demographic contraction, using BOTTLENECK (in contrast, for example, to orangutans in Sabah, Malayisia; Goossens et al. 2006) or to have produced low values for standard estimates of genetic diversity (see above). Therefore, it is apparent that the impact of the population decline (see below) and habitat loss have not yet been severe enough to produce a significant signature in some elements of genetic diversity, and the present wild giant panda populations still comprise a surprisingly rich gene pool.

Population Structure and Genetic Differentiation

Features of the genetic structure in modern giant pandas are (1) the considerable genetic distinctiveness of QIN from other populations, and (2) substantial historic gene flow inferred among the 4 southern populations, MIN, QIO, LIS, and XL, despite significant genetic differentiation being detected in almost all pairwise comparisons using microsatellites (table 2). The genetic distinctiveness of QIN was confirmed in the mtDNA AMOVA and microsatellite STRUCTURE analysis (table 2; fig. 2) and has also been detected in previous studies (Lu et al. 2001; Wan et al. 2003). STRUCTURE, however, clearly identified 4 genetic clusters in giant panda populations using approaches of Pritchard, Stephens, and Donnelly (2000) and Evanno, Regnaut, and Goudet (2005). These clusters were largely associated with specific mountain ranges, although with evidence for gene flow among some of them (fig. 3; supplementary table 4). The above results imply that those populations may have lost contact relatively recently, perhaps as a result of habitat fragmentation, but that some incipient divergence was present even when they were in contact. Although large rivers may act as effective barriers to gene flow (Eriksson et al. 2004; Goossens et al. 2005), our data indicate that the rivers do not provide a complete barrier between adjacent giant panda populations.

Therefore, the apparent genetic divergence between QIN and the other populations and the genetic differentiation among current populations may not result from isolation by rivers alone. Continuous habitat connected the Qinling with the Minshan ranges until the middle twentieth century, and the QIN population is only expected to have become demographically entirely disconnected from MIN in the relatively recent past (Hu 2001; Hu et al. 1985).

Therefore, isolation by distance (IBD) may partly account for the genetic differentiation between QIN and the other mountain populations. In addition, increasing human movements began from the Warring State Period (475 B.C. to 221 B.C.), connecting the Sichuan basin to the northern Shaanxi area, and agriculture, which concomitantly developed in the valleys of the Hanjiang and Jialingjiang rivers, may be partly responsible for the further isolation of QIN.

Population Demography

Almost all published studies state that giant pandas have experienced a recent rapid population decline, which has been generally ascribed to the development of Prehistoric agriculture in the Neolithic (Pei 1974; Wang 1974; Wei et al. 1990; Hu 2001; Wan et al. 2003) or increasing human activity (Hu 2001; Lu et al. 2001). However, to date, no attempt has been made to estimate the extent of the decline, or to date the point at which the decline started.

In our results, the posterior distributions of N_0 and N_1 do not overlap in all 3 studied populations under both linear and exponential models, which provides strong evidence for a genetic signal of population decline (table 3; supplementary table 5). MSVAR simulations assuming an exponential model indicated large historical effective population sizes in the QIN, MIN, and QIO mountain ranges (table 3). Fossil and historical records indicate that these 3 mountain regions comprised most or all of the giant panda habitat before the population decline (Hu 2001); however obtaining precise estimates of the habitat area for the ancestral population is impossible. Pristine mountain habitat is estimated to likely have comprised approximately 130,000 km^2 (QIN), 110,000 km^2 (MIN), and 66,000 km^2 (QIO), based on predictions of habitat covering the whole areas of these mountain ranges, although these values may be underestimates, because pandas were also historically recorded in large areas close to mountains (Wen and He 1981). Taking the criterion of the core area of a panda individual (0.33 km^2, Hu et al. 1985) and a carrying capacity of approximately 1.2 km^2 per individual (Qin 1990), the ancestral habitat seems more than adequate to contain the effective population sizes inferred. Under the exponential model, the posterior distribution of N_0 and N_1 (50% quantile) indicates 80-fold (QIN), 16-fold (MIN), and 25-fold (QIO) effective population declines, respectively (table 3). Simulations using linear models also indicated similar population decreases (supplementary table 5). Comparing ancestral habitat (300,000 km^2 for 3 ancestral populations) with that of extant populations (21,600 km^2; Hu and Wei 2004), approximately 92% has disappeared, a factor which might have contributed most strongly to historical population declines. A similar (although stronger) population decrease was recently reported to relate to severe deforestation for orang-utans in Malaysia (Goossens et al. 2006).

MSVAR simulations also imply that population declines might have started several thousand years ago or even further back in the past. While it is difficult to pinpoint exactly why and how the population decline started (there is currently no evidence to eliminate anthopogenic or ecological/climatic explanations), we can postulate that these population declines would have been accelerated and enhanced by major expansions in ancient human civilizations (~1500 B.C.; Stavrianos 1998), the rapid development of agriculture, and concomitant landscape events in recent time. In the sixteenth century, the introduction of several high-yield crops of American origin (potato, sweet potato and maize) is known to have triggered an extensive human population growth in China (Ho 1959). Accompanied by the migration policy of the Qing Dynasty, the population of the Sichuan basin expanded rapidly, especially in the mountain areas, which may have resulted in anthropogenic habitat fragmentation and further population decline, and was likely to be responsible for extinction of populations at the eastern edge of the Sichuan basin (e.g., in Hunan and Hubei provinces; Wen and He 1981; Zhu and Long 1983).

Heterozygosity-based bottleneck tests (Cornuet and Luikart 1996), however, did not show signals of population declines. Although most empirical and theoretical work suggests that the SMM and the TPM are more appropriate mutation models for microsatellite loci than the IAM (Shriver et al. 1993; Di Rienzo et al. 1994; Schlotterer et al. 1998), these two models also did not show evidence for a recent population decrease. This might be ascribed to the fact that heterozygosity tests are most applicable to very recent events (usually a few dozen generations; Luikart et al. 1998), rather than the near thousand generations implied here.

ESUs and MUs

Phylogeographic and population genetic studies can provide valuable information to help identify evolutionarily significant units (ESUs) and management units (MUs) for conservation (Moritz 1994). Accordingly, ESUs may be designated on the basis of reciprocal monophyly at mitochondrial DNA, while MUs are identified by significant differences in allele frequency distributions and significant divergence in mitochondrial or nuclear loci. Our data do not support the existence of major geographical partitions defining isolated groups which could be viewed as subspecies or ESUs. However, because of the significant genetic differentiations and reduced gene flow between Qinling and other populations, we suggest that QIN should certainly be considered as a separate MU for conservation purposes. The other populations are also significantly differentiated in most cases; however they do appear to have been exchanging genes recently in the past, so could be combined into same MU.

Conclusion

The results presented here are the first to (1) reveal relatively high genetic variation in giant pandas and suggest that genetic diversity has been underrecorded in previous studies; (2) depict the dynamic pattern of genetic structure present in current giant panda populations; (3) describe the demographic history of extent giant panda population; (4) demonstrate a clear genetic signature for the giant panda population decline starting several thousand years ago or even further back in the past, and being accelerated and

enhanced by the expansion of human populations. These results provide a new genetic profile for giant pandas, challenging the hypothesis that the giant panda is at an "evolutionary dead end." Indeed, a recent national giant panda population survey showed the first evidence of a population recovery (Yan 2005), and further detailed molecular census found that the national survey may have substantially underestimated the population, implying that the population may not have gone down to the tiny numbers previously thought (Zhan et al. 2006). These findings further support our conclusion and indicate that the species has a much better chance of long-term viability provided demographic stability and habitat protection remain in force. So the conservation imperative should focus on habitat protection and restoration and the protection of extant populations from threats of human activities.

Supplementary Material

Supplementary figures 1 and tables 1–5 are available at Molecular Biology and Evolution online (http://www.mbe. oxfordjournals.org/). GenBank accession numbers are included.

Acknowledgments

Literature Cited

Avise JC. 1989. Role of molecular genetics in recognition and conservation of endangered species. Trends Ecol Evol. 4:279–281.

Beaumont MA. 1999. Detecting population expansion and decline using microsatellites. Genetics. 153:2013–2029.

Beaumont MA, Bruford MW. 1999. Microsatellites in conservation genetics. In: Goldstein DB, Schlotterer C, editors. Microsatellites: evolution and applications. Oxford: Oxford University Press. p. 165–182.

Belkhir K. 2004. GENETX, Logiciel sous Windows TM pour la Génétique des Populations. Montpellier: Université de Montpellier II.

Bertorelle G, Bruford MW, Chemini C, Vernesi C, Hauffe HC. 2004. New, flexible Bayesian approaches to revolutionize conservation genetics. Conserv Biol. 18:1–2.

Bohonak AJ. 2002. IBD (Isolation by Distance): a program for analyses of isolation by distance. J Hered. 93:153–154.

Burland TG. 2000. DNASTAR's Laser gene sequence analysis software. Methods Mol Biol. 132:71–91.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol. 9:1657–1659.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics. 144:2001–2014.

Craighead L, Paetkau D, Reynolds HV, Vyse ER, Strobeck C. 1995. Microsatellite analysis of paternity and reproduction in Arctic grizzly bears. J Hered. 86:255–261.

Crandall KA, Templeton AR. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics. 134:959–969.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA. 91:3166–3170.

Eriksson J, Hohmann G, Boesch C, Vigilant L. 2004. Rivers influence the population genetic structure of bonobos (Pan paniscus). Mol Ecol. 13:3425–3435.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 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.

Fang S, Feng W, Zhang A, et al. (13 co-authors). 1997. The research on genetic diversity of the giant panda. Chengdu Zoo and Chengdu Research Base for Giant Panda Breeding, editors. Proceeding of the international symposium on the protection of the giant panda (Ailuropoda melanoleuca), Chengdu, China 1997. Chengdu: Sichuan Publishing House of Science and Technology. p. 141–147.

Garza JC, Williamson EG. 2001. Detection of reduction in population size using data from microsatellite loci. Mol Ecol. 10:305–318.

Gittleman JL. 1994. Are the pandas successful specialists or evolutionary failures. Bioscience. 44:456–464.

Goossens B, Chikhi L, Ancrenaz M, Lackman-Ancrenaz I, Andau P, Bruford MW. 2006. Genetic signature of anthropogenic population collapse in orang-utans. PLoS Biol. 4:285–291.

Goossens B, Chikhi L, Jalil MF, Ancrenaz M, Lackman-Ancrenaz I, Mohamed M, Andau P, Bruford MW. 2005. Patterns of genetic diversity and migration in increasingly fragmented and declining orang-utan (Pongo pygmaeus) populations from Sabah, Malaysia. Mol Ecol. 14:441–456.

Goudet J. 2002. FASTA, a program to estimate and test gene diversities and fixation indices. Lausanne: Institute of Ecology.

Groombridge JJ, Jones CG, Bruford MW, Nicols RA. 2000. Genetic analysis tells of conservation success in the Mauritius kestrel. Nature. 403:616.

Ho P. 1959. Studies on the population of China. Cambridge (UK): Harvard University Press. p. 1368–1953.

Hu J. 2001. Research on the giant panda. Shanghai: Shanghai Publishing House of Science and Technology.

Hu J, Schaller B, Pan W, Zhu J. 1985. The giant panda of Wolong. Chengdu: Sichuan Publishing House of Sichuan and Technology.

Hu J, Wei F. 2004. Comparative ecology of giant pandas in the five mountain ranges of their distribution in China. In: Lindburg D, Baragona K, editors. Giant pandas biology and conservation. Los Angeles: University of California Press. p. 137–148.

Huang W. 1993. The skull, mandible and dentition of giant panda (Ailuropoda): morphological characters and their evolutionary implications. Vertebrata Palasiatica. 31:1–9.

Kumar S, Tamura K, Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinformatics. 5:150–163.

Lu Z, Johnson WE, Menotti-Raymond M, et al. (12 co-authors). 2001. Patterns of genetic diversity in remaining giant panda populations. Conserv Biol. 15:1596–1607.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered. 89:238–247.

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol. 18:189–197.

Marshall H, Ritland K. 2002. Genetic diversity and differentiation of Kermode bear populations. Mol Ecol. 11:685–697.

Moran NA. 1988. The evolution of host-plant alternation in aphids: evidence for specialization as a dead end. Am Nat. 132:681–706.

Moritz C. 1994. Application of mitochondrial DNA analysis in conservation: a critical review. Mol Ecol. 3:401–411.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

O'Brien SJ. 1994. A role for the molecular genetics in biological conservation. Proc Natl Acad Sci USA. 91:5748–5755.

Paetkau D, Amstrup SC, Born EW, et al. (11 co-authors). 1999. Genetic structure of the world's polar bear populations. Mol Ecol. 8:1571–1584.

Paetkau D, Strobeck C. 1994. Microsatellite analysis of genetic variation in black bear populations. Mol Ecol. 3:489–495.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Vyse E, Ward R, Strobeck C. 1998. Variation in genetic diversity across the range of north American brown bears. Conserv Biol. 12:418–429.

Pan W, Lu Z, Zhu X, Wang D, Wang H, Long Y, Fu L, Zhu X. 2001. A chance for lasting survival. Beijing: Beijing University Press.

Pei W. 1965. More on the problem of augmentation and diminution in size of Quaternary. Vertebrata Palasiatica. 9:37–46.

Pei W. 1974. Evolutionary history of giant panda. Acta Zool Sinica. 20:188–190.

Piry S, Luikart G, Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. J Hered. 90:502–503.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics. 155:945–959.

Qin Z. 1990. Bamboo food recourses of giant pandas and the regeneration of the bamboo groves in Sichuan. In: Hu J, editor. Research and progress in biology of the giant panda. Chengdu: Sichaun Publishing House of Science and Techology. p. 103–111.

Qiu Z, Qi G. 1989. *Ailuropoda* found from the late Miocene deposits in Lufeng, Yunnan. Vertebrata Palasiatica. 27:153–169.

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 86:248–249.

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. Mol Ecol Notes. 4:137–138.

Rousset F, Raymond M. 1997. Statistical analyses of population genetic data, new tools, old concepts. Trends Ecol Evol. 12: 313–317.

Ruiz-Garcia M. 2003. Molecular population genetic analysis of the spectacled bear (Tremarctos ornatus) in the northern Andean area. Hereditas. 138:81–93.

Ruiz-García M, Orozco-terWengel P, Castellanos A, Arias L. 2005. Microsatellite analysis of the spectacled bear (*Tremarctos ornatus*) across its range distribution. Genes Genet Syst. 80:57–69.

Saitoh T, Ishibashi Y, Kanamori H, Kitahara E. 2001. Genetic status of fragmented populations of the Asian black bear *Ursus thibetanus* in western Japan. Popul Ecol. 43:221–227.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Press.

Schlotterer C, Ritter R, Harr B, Brem G. 1998. High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. Mol Biol Evol. 15:1269–1274.

Schneider S, Roesslo D, Excoffier L. 2000. ARLEQUIN 2000. Geneva Switzerland: University of Geneva.

Shriver MD, Jin L, Chakraborty R, Boerwinkle E. 1993. VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics. 134:983–993.

Sjögren P, Wyöni PI. 1994. Conservation genetics and detection of rare alleles in finite populations. Conserv Biol. 8: 267–270.

Smith TB, Bruford MW, Wayne RK. 1993. The preservation of process: the missing element of conservation programs. Biodivers Lett. 1:164–167.

Stavrianos LS. 1998. A global history: from prehistory to 21st century, 7th edition. New Jersey: Prentice Hall Inc.

Storz JF, Beaumont MA. 2002. Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. Evolution. 56:154–166.

Storz JF, Beaumont MA, Alberts SC. 2002. Genetic evidence for long-term population decline in a savannah-dwelling primate: inferences from a hierarchical bayesian model. Mol Biol Evol. 19:1981–1990.

Su B, Shi L, He G, Zhang A, Song Y, Fei L. 1994. Genetic diversity in the giant panda: evidence from protein electrophoresis. Chinese Sci Bull. 39:1305–1309.

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics. 105:437–460.

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 Res. 22: 4673–4680.

Waits L, Taberlet P, Swenson JE, Sandegren F, Franzen R. 2000. Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (*Ursus arctos*). Mol Ecol. 9:421–31.

Walsh PS, Metzger DA, Higuchi R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques. 10:506–513.

Wan Q, Fang S, Wu H, Fujihara T. 2003. Genetic differentiation and subspecies development of the giant panda as revealed by DNA fingerprinting. Electrophoresis. 24:1353–1359.

Wang J. 1974. On the Taxonomic status of species, geological distribution and evolutionary history of Ailuropoda. Acta Zool Sinica. 20:191–201.

Wei F, Wu Y, Yuan C, Hu J. 1990. The change in body size of giant panda and its vicissitudes. J Sichuan Teach Coll. 11: 23–28.

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

Wen H, He Y. 1981. The distribution of giant panda in Henan, Hubei, Hunan and Sichuan in last five thousands years. J Southwest Teach Coll. 1:87–93.

Xia X, Xie Z. 2001. DAMBE: software package for data analysis in molecular biology and evolution. J Hered. 92:371–373.

Yan X. 2005. Status, challenge and prospect of wild giant pandas. Acta Theriol Sinica. 25:402–406.

Zhan X, Li M, Zhang Z, Goossens B, Chen Y, Wang H, Bruford MW, Wei F. 2006. Molecular censusing doubles giant panda population estimate in a key nature reserve. Curr Biol. 16:R451–452.

Zhang Y, Ryder O, Fan Z, Zhang H, et al. (17 co-authors). 1997. A study on sequence variation and genetic diversity of the giant panda. Sci China Series C. 27:139–144.

Zhang Y, Su B. 1997. Genetic diversity of giant panda. In: Hu Z, Zhang Y, editors. Genetic diversity of animals and plants in China. Hangzhou: Zhejiang Publishing House of Science and Technology. p. 7–18.

Zhang Y, Wang X, Ryder OA, Li H, Zhang H, Yong Y, Wang P. 2002. Genetic diversity and conservation of endangered animal species. Pure Appl Chem. 74:575–584.

Zhu J, Long Z. 1983. The vicissitudes of the giant panda. Acta Zool Sinica. 29:93–104.

Scott Edwards, Associate Editor