Integrative Zoology 2015; 10: 376–388 doi: 10.1111/1749-4877.12136

ORIGINAL ARTICLE

The development and application of a multiplex short tandem repeat (STR) system for identifying subspecies, individuals and sex in tigers

Zheng-Ting ZOU,[1,*] Olga V. UPHYRKINA,[2] Pavel FOMENKO[3] and Shu-Jin LUO[1]

[1]College of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China, [2]Institute of Biology and Soil Sciences, Far Eastern Branch Russian Academy of Sciences, Vladivostok, Russia and [3]World Wide Fund for Nature – Russia Amur Branch, Vladivostok, Russia

Abstract

Poaching and trans-boundary trafficking of tigers and body parts are threatening the world's last remaining wild tigers. Development of an efficient molecular genetic assay for tracing the origins of confiscated specimens will assist in law enforcement and wildlife forensics for this iconic flagship species. We developed a multiplex genotyping system "tigrisPlex" to simultaneously assess 22 short tandem repeat (STR, or microsatellite) loci and a gender-identifying *SRY* gene, all amplified in 4 reactions using as little as 1 ng of template DNA. With DNA samples used for between-run calibration, the system generates STR genotypes that are directly compatible with voucher tiger subspecies genetic profiles, hence making it possible to identify subspecies via bi-parentally inherited markers. We applied "tigrisPlex" to 12 confiscated specimens from Russia and identified 6 individuals (3 females and 3 males), each represented by duplicated samples and all designated as Amur tigers (*Panthera tigris altaica*) with high confidence. This STR multiplex system can serve as an effective and versatile approach for genetic profiling of both wild and captive tigers as well as confiscated tiger products, fulfilling various conservation needs for identifying the origins of tiger samples.

Key words: Amur tiger, genetic, multiplex, sexing, short tandem repeat

INTRODUCTION

The tiger [*Panthera tigris* (Linnaeus, 1758)] , one of the world's most prominent charismatic megafauna species, has been the focus of conservation efforts for almost half a century. Historically, tigers inhabited much of Asia and likely numbered near 100 000 as recently as the 1930s (Seidensticker *et al.* 2010). However, today only approximately 3000 adult tigers survive in the wild (Dinerstein *et al.* 2007; Walston *et al.* 2010; Chundawat *et al.* 2011). By contrast, there is a much larger

Correspondence: Shu-Jin Luo, College of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.
Email: luo.shujin@pku.edu.cn
**Present address*: Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA.

captive population of over 13 000 individuals, most of which lack coordinated breeding management and are considered "generic" tigers (Luo *et al.* 2008; Nyhus *et al.* 2010).

Wild tigers are mainly threatened by growing human populations, habitat loss and fragmentation, and the hunting of both themselves and their prey. Poaching and illegal trade in tiger parts are critical factors pushing tigers to the brink of extinction (Dinerstein *et al.* 2007; Walston *et al.* 2010). Tiger body parts and derivatives, such as bones, tails, whiskers and even eyeballs, are claimed to have therapeutic effects in traditional oriental medicine, leading to a highly profitable black market in many countries (Mills & Jackson 1994). Open illegal trade in tiger skins also exists in China's Tibetan regions for use in household decorations and costumes (Verheij *et al.* 2010). Over the past decade, seizures of tiger parts being brought across the borders of India, Russia and other tiger range states into China have been increasing (Verheij *et al.* 2010). All these trends highlight the importance of enhancing law enforcement efforts to prevent poaching and trade, a process that will require the development of tools and assays for tracing sources of confiscated tiger parts.

Molecular genetic evidence suggests the existence of 6 living subspecies: Bengal tiger (*Panthera tigris tigris*), Amur tiger (*Panthera tigris altaica*), South China tiger (*Panthera tigris amoyensis*), Sumatran tiger (*Panthera tigris sumatrae*), Indochinese tiger (*Panthera tigris corbetti*) and Malayan tiger (*Panthera tigris jacksoni*) (Luo *et al.* 2004, 2010b; Chundawat *et al.* 2011). Mitochondrial DNA and microsatellite-based genetic marker systems have been applied widely to a variety of tiger conservation issues (Luo *et al.* 2010a), including reconstruction of demographic history (Mondol *et al.* 2009a), non-invasive assessment of population genetic status (Henry *et al.* 2009; Mondol *et al.* 2009b; Sharma *et al*. 2009, 2013; Borthakur *et al*. 2011), assignment of sub-specific origins of captive tigers (Luo *et al*. 2008), development of mtDNA SNP chips for forensic applications (Kitpipit *et al.* 2012) and inference of genetic ancestry of the extinct Caspian tiger using museum specimens (Driscoll *et al.* 2009). In particular, microsatellite markers in combination with non-invasive sampling and mark–recapture statistics have contributed significantly to individual identification and estimates of population size in India and Russia (Henry *et al.* 2009; Mondol *et al.* 2009b).

Due to their relatively high mutation rates and multi-allele nature, microsatellite or short tandem repeat (STR) markers are suitable for detecting genetic diversity at population and individual levels, and are, therefore, often used in molecular ecology and forensic studies (Butler *et al.* 2007; Wasser *et al.* 2008; Ogden 2011). Multiplex polymerase chain reaction (PCR) amplifications of STR loci are often applied to improve the efficiency of template DNA where the quantity of a sample is limited. Mondol *et al.* (2009b; 2012) report simultaneous amplification of STR loci for tiger fecal samples collected from India. Borthakur *et al.* (2011) conducted two sets of multiplex PCRs in tiger scat samples from eastern India, with 4 STR loci amplified in each set, but the loci chosen did not overlap with previous STR-based studies and, hence, were not comparable with genetic profiles from other populations. To date there is no well-established multiplex PCR system for tiger STR profiling across all subspecies.

We attempted to develop a multiplex system that can serve as an effective, robust and versatile approach to obtain genetic profiles of captive and wild tigers, as well as tiger parts of unknown origin. Here we present the results: a multiplex STR genotyping system for the tiger, which incorporates 1 gender-identification marker and 22 autosomal STR loci optimized for tiger genomic DNA, organized into four subset panels. We used three samples with known genotypes for calibration with genetic profiles reported by Luo *et al.* (2004), and then tested 12 tiger samples of undisclosed source for identification of individual, sex and subspecies to validate the system. Due to the degraded nature of non-invasively collected samples such as scats, a single multiplex assay including a large number of loci (i.e. >12) may not be realistic and could incur severe allele dropout. To increase the system's versatility for potential applications to various categories of genetic samples, the 23 markers were organized into four 5-plex or 6-plex panels. This way we may assess the level of precision with reduced numbers of markers, thereby allowing for adjustments in various situations depending on time, labor, sample quality or financial constraints. Finally, we also made serial dilutions of DNA to assess sensitivity and volumes required for reliable results.

MATERIALS AND METHODS

Biological samples

DNA from 3 tiger individuals (PTI-88, PTI-212 and PTI-270 [Table S1]) were used in pilot experiments and calibration for genotyping results between different STR runs, so that a comparison to genetic profiles of vouch-

Table 1 Marker panels and primers of the tiger short tandem repeat (STR) multiplex system

Locus	Cat chromosome	Repeat unit	Size range (bp)	Dye	Primer sequence (5'-3') †
Panel A					
FCA8	A1	(CA)n	132–148	ROX	F: ROX - ACTGTAAATTTCTGAGCTGGCC R: TGACAGACTGTTCTGGGTATGG
FCA32	A2	(CA)n	188–204	6-FAM	F: 6-FAM - GGCAATTCATGGTAGAGAAAAA R: CAAGAGTGCATTGGGCAGTA
FCA69	B4	(CA)n	97–115	6-FAM	F: 6-FAM - AATCACTCATGCACGAATGC R: AATTTAACGTTAGGCTTTTTGCC
FCA77	C2	(CA)n	144–160	HEX	F: HEX - GGCACCTATAACTACCAGTGTGA R: ATCTCTGGGGAAATAAATTTTGG
FCA105	A2	(CA)n	191–207	ROX	F: ROX - TTGACCCTCATACCTTCTTTGG R: TGGGAGAATAAATTTGCAAAGC
Panel B					
FCA5	E1	(CA)n	140–162	6-FAM	F: 6-FAM - TCCTGGCATCCTCCCCATTTCA R: AAGGCTGACACATCCATCTGGG
FCA43	C2	(CA)n	115–127	ROX	F: ROX - GAGCCACCCTAGCACATATACC R: AGACGGGATTGCATGAAAAG
FCA90	A1	(CA)n	115–125	6-FAM	F: 6-FAM - ATCAAAAGTCTTGAAGAGCATGG R: TGTTAGCTCATGTTCATGTGTCC
FCA91P	B4	(CA)n	128–146	HEX	F: HEX - ACTCCCAACTTTCACATTCTGACT R: TGCCCAAACATAATCTCTGCAT
FCA94	F2	(CA)n	194–206	ROX	F: ROX - TCAAGCCCCATTTTACCTTC R: CACCTGAGCCAAAGGCTATC
FCA290P	C1	(CA)n	208–226	6-FAM	F: 6-FAM - TCAGGCTACATCTGAAAGTGAGGA R: TGCCCATTTGAGAAAGGTCATC
Panel C					
FCA44	B4	(CA)n	110–126	HEX	F: HEX - AGGGCCTGAACCAAGAGAAT R: TATTTACAGAGTGCACAGAGGAGG
FCA126	B1	(CA)n	128–152	ROX	F: ROX - GCCCCTGATACCCTGAATG R: CTATCCTTGCTGGCTGAAGG
FCA161P	A3	(CA)n	169–187	HEX	F: HEX - CCGATACACACCTGCCAAGATT R: TCACAGACGTGCTCTAGCCAAA
FCA176	A1	(CA)n	200–222	ROX	F: ROX - GGAAACTTGGAAAGCAAAACC R: TCCACAGTTGGAGTTCTTAAGG
FCA220	F2	(CA)n	200–212	6-FAM	F: 6-FAM - CGATGGAAATTGTATCCATGG R: GAATGAAGGCAGTCACAAACTG
FCA441	D3	(ATAG)n	138–166	6-FAM	F: 6-FAM - ATCGGTAGGTAGGTAGATATAG R: GCTTGCTTCAAAATTTTCAC
Panel D					
FCA211P	B1	(CA)n	112–120	HEX	F: HEX - AGAACATAACGCCTCACCCAGT R: ATGGTGACTGCTTTCCTCCCTA
FCA293P	C1	(CA)n	196–208	ROX	F: ROX - ACAGATCGCCCAAAAGCACAC R: TCTCCACATCTTGTCAACAACG
FCA304P	A2	(CA)n	119–139	ROX	F: ROX - TCATTGGCTACCACAAAGTAGG R: TAGCTGCATGCCATTGGGTAAC
FCA310P	C2	(CA)n	123–133	6-FAM	F: 6-FAM - CTTTAATTGTATCCCAAGTGGTCA R: TCTTAATGCTGCAATGTAGGGCA
FCA391P	B3	(ATAG)n	200–224	6-FAM	F: 6-FAM - GGCCTTCTAACTTCCTTGCAGA R: CATTTAGTTAGCCCATTTTCATCA
SRY	Y	-	101	6-FAM	F: 6-FAM - AGCGAACTTTGCACGGAGAG R: GCGTTCATGGGCCGTTTGACG

†Primer sequences of the STR loci follow Menotti-Raymond *et al.* (1999), except for those labelled "P" at the locus names indicating redesigned primers.

er tiger specimens based upon the same loci (Luo *et al.* 2004) was possible. Twelve samples (RFET01-RFET12 [Table S1, Table 1]) of small pieces of tiger skin or muscle were collected in Russia in 2012 during biological examinations of animal parts and/or their derivatives. The specimens were of undisclosed identity (i.e. unknown number of individuals, sex or subspecies), and were provided for this study to validate the power of the STR multiplex genotyping system. DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen Inc., USA) following the manufacturer's protocols.

Development of the tiger short tandem repeat multiplex system

Short tandem repeat markers of the tiger multiplex genotyping system were selected from the 30 polymorphic markers that were first developed in the domestic cat (Menotti-Raymond *et al.* 1999) and then genotyped across all 6 extant tiger subspecies (Luo *et al.* 2004) and captive tigers with unknown genetic backgrounds (Luo *et al.* 2008). Selection criteria included coverage of different chromosomes, lack of linkage between loci according to the cat reference genome (i.e. located on different chromosomes or at least 12 centimorgans apart on the same chromosome), high levels of heterozygosity and assignment efficiency across multiple subspecies (Luo *et al.* 2008), and performance robustness in multiplex co-amplification. Sequences of the STR loci were verified by aligning them with the tiger reference genome (Cho *et al.* 2013) using BLASTN 2.2.26 (Altschul *et al.* 1997) with mismatched primers being redesigned accordingly with Primer3Plus (Untergasser *et al.* 2007) for an optimized performance in the tiger. Calibration of allele size using 3 reference samples (PTI-88, PTI-212 and PTI-270), which represent most of the allele range in tigers, enabled comparison of the genotypes obtained with modified primer sets to the previously published voucher tiger dataset (Luo *et al.* 2004). Primers for the male sex identification marker *SRY* were modified from the cat *SRY* primer (Menotti-Raymond *et al.* 2005) and female sex was determined by a negative result from *SRY* yet positive for other markers in the same multiplex panel.

Short tandem repeat loci selected for the multiplex system were organized into 5-plex or 6-plex panels labeled with 3 different fluorescent dyes. Those with overlapping product size in the same panel were assigned different dye colors, while for non-overlapping loci the same dye color was used. In addition, amplicons within each panel were designed to leave at least 30 bp between loci of the same dye color, in order to reduce potential overlap if new alleles are discovered in the future. Forward primers were fluorescently labeled at the 5′ end with blue (6-FAM), green (HEX) or red (ROX) dyes. A 10× primer mix for each panel was prepared by diluting 100-μM primer stocks to reach a working concentration of 2 μM for each primer pair, except for FCA32 in Panel A whose concentration of the forward and reverse primer in the 10× mix was 20 μM each, to empirically adjust for an overall consistent amplification strength across loci.

The multiplex PCR amplifications were conducted using a Qiagen Multiplex PCR *Plus* Kit (Qiagen) in a 10-μL reaction system containing 5 μL of 2× Multiplex PCR Mastermix, 1 μL of 10× primer mix, and 6–10 ng of DNA. To evaluate the multiplex system's sensitivity to low DNA concentrations, 2 template DNA samples were diluted to yield a final mass of 5, 2 and 1 ng in the reaction system. Thermal cycling conditions were performed with the GeneAmp PCR System 9700 (Thermo Fisher Scientific Inc., USA) or Biometra TProfessional Thermocycler (Biometra, Germany) under the following conditions: 5 min denaturation at 95 °C, 30 cycles of 95 °C for 30 sec, 60 °C for 90 sec and 72 °C for 3 min, and a final extension at 68 °C for 30 min.

The multiplex PCR products were diluted appropriately and 3 μL of each product was mixed with 0.11 μL GeneScan 500LIZ Size Standard and 8.89 μL Hi-Di Formamide (Thermo Fisher Scientific Inc., USA). The mixture was then electrophoresed on an ABI 3730xl DNA Analyzer, which was fluorescently calibrated with the DS-33 GeneScan Installation Standards using 6FAM, VIC, NED and PET dyes (Thermo Fisher Scientific Inc., USA). Alleles were scored and analyzed using GENEMAPPER v4.0 (Thermo Fisher Scientific Inc., USA) and Allelogram v2.2 (Morin *et al.* 2009).

Population genetic analyses

Short tandem repeat genotypes from 113 voucher tiger specimens (Luo *et al.* 2004), which included Amur tigers (n = 34), captive South China tigers (n = 2), Indochinese tigers (n = 33), Malayan tigers (n = 22), Sumatran tigers (n = 16) and Bengal tigers (n = 6), were used to evaluate population statistics of the STR multiplex genotyping system. Average number of alleles per locus, polymorphism information content (PIC), and observed and expected heterozygosity were calculated using the Excel Microsatellite Toolkit v3.1.1 (Park 2001).

Table 2 Summary statistics of the tiger short tandem repeat (STR) multiplex panels

Locus	Number of Alleles Per Locus	Observed heteozygosity (H_O)	Expected heterozygosity (H_E)	Polymorphism information content (PIC)	F_{ST}	RHO_{ST}	Informativeness-for-assignment index (*In*)	Probability of identity (PID)	Probability of identity (PIDunbiased)	Probability of identity (PIDsib)
Panel A										
FCA8	9	0.422	0.685	0.654	0.383	0.351	0.713	0.1289	0.1219	0.4411
FCA32	9	0.524	0.762	0.723	0.321	0.402	0.784	0.0935	0.0887	0.3940
FCA69	9	0.610	0.775	0.742	0.175	0.244	0.694	0.0821	0.0772	0.3849
FCA77	7	0.450	0.693	0.642	0.172	0.132	0.437	0.1444	0.1390	0.4410
FCA105	7	0.555	0.823	0.795	0.235	0.352	0.645	0.0573	0.0533	0.3546
Overall	8.2	0.512	0.748	0.711	0.257	0.296	0.655	8.19×10^{-6}	6.18×10^{-6}	1.05×10^{-2}
Panel B										
FCA5	9	0.600	0.791	0.756	0.345	0.081	1.097	0.0764	0.0720	0.3753
FCA43	7	0.422	0.732	0.686	0.344	0.488	0.549	0.1159	0.1109	0.4147
FCA90	6	0.481	0.655	0.613	0.349	0.407	0.604	0.1605	0.1537	0.4643
FCA91P	10	0.631	0.793	0.761	0.279	0.506	0.890	0.0730	0.0684	0.3736
FCA94	5	0.598	0.723	0.677	0.161	0.237	0.669	0.1211	0.1158	0.4212
FCA290P	8	0.743	0.796	0.760	0.192	0.235	0.595	0.0752	0.0711	0.3727
Overall	7.5	0.579	0.748	0.709	0.278	0.326	0.734	9.44×10^{-7}	6.91×10^{-7}	4.24×10^{-3}
Panel C										
FCA44	8	0.613	0.779	0.743	0.280	0.572	0.683	0.0826	0.0780	0.3831
FCA126	7	0.473	0.691	0.636	0.473	0.304	0.780	0.1491	0.1439	0.4432
FCA161P	9	0.563	0.823	0.795	0.333	0.521	1.064	0.0564	0.0523	0.3546
FCA176	10	0.694	0.818	0.792	0.184	0.027	0.699	0.0569	0.0526	0.3573
FCA220	7	0.649	0.729	0.683	0.127	0.172	0.565	0.1176	0.1124	0.4166
FCA441	8	0.648	0.797	0.762	0.205	0.323	0.630	0.0738	0.0695	0.3720
Overall	8.2	0.606	0.773	0.735	0.267	0.320	0.737	3.43×10^{-7}	2.41×10^{-7}	3.33×10^{-3}
Panel D										
FCA211P	5	0.333	0.750	0.707	0.608	0.425	0.899	0.1044	0.0997	0.4026
FCA293P	4	0.541	0.592	0.526	0.198	0.323	0.412	0.2317	0.2259	0.5131
FCA304P	10	0.477	0.706	0.673	0.321	0.278	0.791	0.1179	0.1113	0.4283
FCA310P	6	0.491	0.756	0.711	0.322	0.648	0.607	0.1023	0.0978	0.3994
FCA391P	7	0.620	0.739	0.705	0.221	0.081	0.537	0.1002	0.0943	0.4078
Overall	6.4	0.492	0.709	0.664	0.334	0.351	0.649	2.92×10^{-5}	2.31×10^{-5}	1.44×10^{-2}
Overall	7.6	0.552	0.746	0.706	0.283	0.323	0.697	7.75×10^{-23}	2.38×10^{-23}	2.13×10^{-9}

[†]All values were calculated from the calibrated allele frequencies of the voucher tiger specimens used in Luo *et al.* (2004).

Three parameters for evaluating population differentiation or assignment were calculated: measures of population subdivision F_{ST} and RHO_{ST} (the counterpart of F_{ST} with STR-specific mutation models, Slatkin 1995) were calculated in GENEPOP v4.2 (Raymond & Rousset 1995); the informativeness of genetic markers for inference of ancestry, or the index of power for correct subspecies assignment, *In*, was calculated as described by Rosenberg *et al*. (2003). The probability of identity (PID, Waits *et al.* 2001), which measures the probability that genotypes from two unrelated individuals will match by chance (error rate) and, hence, the STR markers' potential for correctly identifying individuals was calculated in GIMLET v1.3.3 (Valière 2002). South China tigers (n = 2) were excluded from these calculations due to small sample size.

Conservation application of the tiger short tandem repeat multiplex system

Genetic identification of individual tiger specimens with unknown background were conducted in CERVUS v3.0 (Kalinowski *et al.* 2007) using the "identity analysis" option. All pairwise comparison values were ranked by number of matched loci to check the reliability of the analysis. To determine the minimum number of STR multiplex panels required to resolve individual identification, we first conducted the analysis using only STR genotypes from the STR multiplex Panel C, which has the highest resolution of individual assignment (or the lowest PID) among the 4 (Table 2), and then data from Panel B, A and D were sequentially added to the dataset (Table 1) following an ascending order of PID.

For subspecies assignment, we applied a complete Bayesian clustering method implemented in STRUCTURE v2.3.3 (Pritchard *et al.* 2000) to calculate the probability (q) of a tiger with unknown identity belonging to each of the 5 putative subspecies, excluding South China tigers, or alternatively, the extent of admixture (i.e. generic tigers). We used the voucher tiger subspecies samples (n = 113 [Luo *et al.* 2004]) as a reference and calculated the probability of each individual being assigned to each population cluster (K = 7, with Indochinese tigers containing three populations [Luo *et al.* 2004]) without prior designation of the origin of the individuals. Seven (K = 7) clusters were chosen for interpreting the voucher tiger population structure, as it produced the highest likelihood among other choices of K (Luo *et al.* 2004). We assumed correlated allele frequencies under the admixed ancestry mode and set the number of burn-in and replication steps to be 100 000 and 1 000 000, respectively.

RESULTS

Selection of 22 short tandem repeat loci in the tiger multiplex genotyping panel

Initially, 30 domestic cat STR loci were validated and optimized for tigers, with genotype data accumulated from multiple past studies (Luo *et al.* 2004, 2008; Mondol *et al.* 2009a,b). Based upon this pilot work, 22 STR loci were selected for inclusion in a final tiger multiplex genotyping panel, among which primers for 8 loci were re-designed for an optimized performance in tiger DNA (labeled "P" in locus name, Table 1). The assay, named "tigrisPlex" with four 5-plex or 6-plex panels (Fig. 1, Panels A–D), contains 20 di-nucleotide and 2 tetra-nucleotide STR loci, and the *SRY* locus for gender identification with female sex determined by a negative result, together covering 11 of the 18 feline autosomes and the Y chromosome. The STR loci located on the same chromosome are greater than 12 centimorgans apart based on the domestic cat genetic linkage map (Menotti-Raymond *et al.* 1999) and are considered unlinked.

Short tandem repeat loci of the "tigrisPlex" exhibited high genetic variability based on population genetic data from the 113 voucher tiger specimens including all 6 extant subspecies (Luo *et al.* 2004). The overall probability of identity (PID) value for each panel was the product of PID per locus within this panel, while the overall values for other parameters were averaged across loci (Table 2). The number of alleles per locus ranged from 4 at FCA293 to 10 at FCA91P, FCA176 and FCA304P, with 7.6 alleles per locus on average. The locus-specific observed heterozygosity (H_O) ranged from 0.333 (FCA211P) to 0.743 (FCA290P) and the expected heterozygosity (H_E) ranged from 0.592 (FCA293P) to 0.823 (FCA105 and FCA161P). An overall average expected heterozygosity of 0.746 was obtained for the 22 loci from the voucher tiger samples. Consistent with H_E, PIC was greatest at FCA105 and FCA161P (0.795) and least at FCA293P (0.526). The values of F_{ST}, RHO_{ST} and information content for the assignment index (*In*) ranged from 0.28 to 0.61, 0.32 to 0.65, and 0.70 to 1.10, respectively, indicating sufficiently high genetic diversity of the panels of markers. The STR loci FCA5, FCA161P and FCA211P provided the highest resolution for distinguishing subspecies. Correlations (Spearman's rank correlation ρ, by R v3.0.1, R Core Team 2013) were significant between F_{ST} and RHO_{ST} (P = 0.013), and F_{ST} and *In* (P = 0.019), consistent with the notion that inter-population genetic divergence indexes reflect the statistical power of subspecies discrimination.

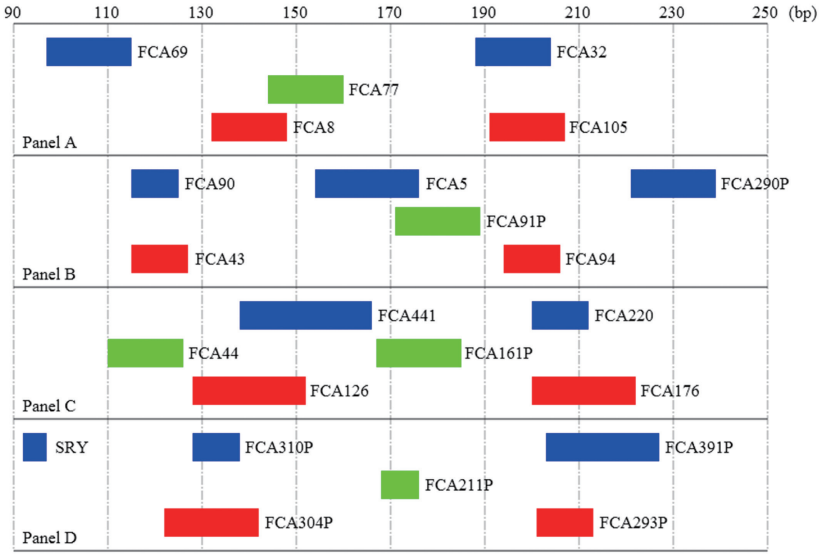

Figure 1 Fluorescent dye colors and size ranges for the 22 short tandem repeat (STR) loci and *SRY* based on allele frequencies of 113 voucher tiger specimens from Luo *et al.* (2004). Numbers on colored bars correspond to names of the STR loci.

The unbiased probability that two individual would match by chance (unbiased PID) using multiple panels of "tigrisPlex" varied from 2.31×10^{-5} (Panel D) to 2.41×10^{-7} (Panel C) (Table 2), which means using only the loci in the most powerful Panel C for individual identification; the chance that samples from two unrelated individuals matched at all 6 loci is less than 1/1 000 000. Even for related individuals from a family, the probability of mistakenly identifying two samples as the same individual using Panel C is on the scale of 1/1000 (PIDsib = 3.33×10^{-3}). If genotypes of Panel B, A and D are sequentially added to Panel C, the error rate would be subsequently reduced to 1.41×10^{-5}, 1.48×10^{-7} and 2.13×10^{-9}, respectively. These measurements indicate that the "tigrisPlex" offers high-level resolution for individual identification.

The tiger short tandem repeat multiplex "tigrisPlex" assay

The "tigrisPlex" assay typically requires 1–10 ng of extracted DNA and generates robust and easily interpretable product profiles (Fig. 2a). Although its sensitivity has not been tested extensively at sub-nanogram levels, the multiplex PCR amplifications gave reliable results with as low as 1-ng template DNA, as shown in the chromatograms from sample RFET04 from a female tiger (Fig. 2b). The PCR products from this multiplex amplification fell in the size range of 100 to 230 bp, with the *SRY* amplification detectable at 97 bp for male individuals (e.g. sample RFET05 in Fig. 2c). Although peak heights in 1 panel are not perfectly balanced (e.g. the signal for FCA304P is relative weak in Fig. 2), practically, the level of bias in amplification does not hinder allele reading. When the amount of template DNA was reduced, peak heights of different loci in 1 panel decreased almost proportionally, and there was no sign of further amplification bias (Fig. 2b).

A case study of confiscated tiger specimens

Genotypes from all 23 "tigrisPlex" loci for 12 tiger samples from Russia were clearly obtained (Table S1). Six pairs of samples were identified as complete matches, corresponding to six unique tigers (3 males and 3 females, Table 1). Values of PID for each of the perfectly matched sample pairs based on genotypes from 1 (panel C), 2 (panels C + B), 3 (panels C + B + A) and all 4 panels (C + B + A + D) of "tigrisPlex" are listed in Table 1, corresponding to high confidence in the individual identification results. Using all panels in "tigrisPlex" the probability of incorrectly identifying two siblings (PIDsib) as one individual is 1/100 000 and is 1/100 if using only panel C. Considering that there are currently only approximately 400 Amur tigers in the wild (Global Tiger Initiative Secretariat 2011), a PIDsib value of 1/1000 (panels C + B) would be sufficiently low to con-

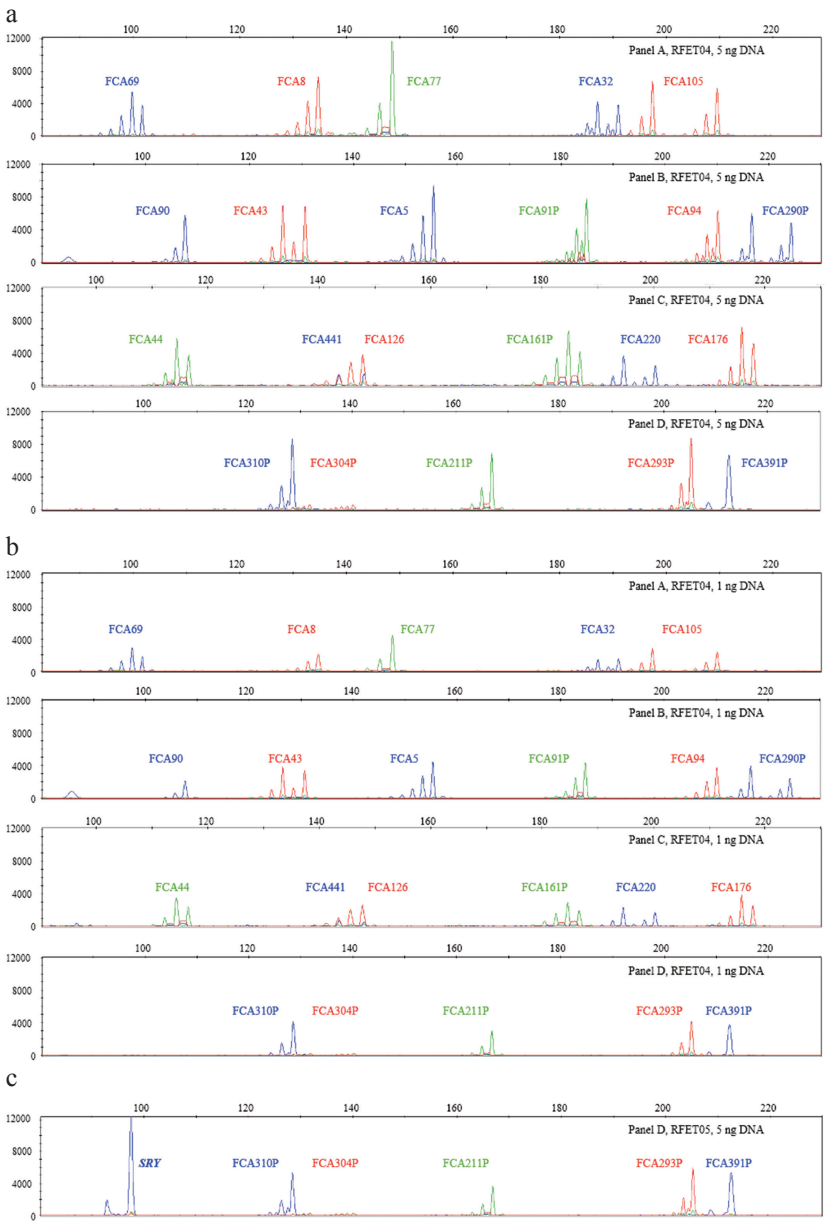

Figure 2 Chromatograms from 2 tigers amplified with the tiger short tandem repeat (STR) multiplex system ("tigrisPlex") and electrophoresed on an ABI 3730xl sequencer. The X-axis represents size of PCR products in base pairs; the Y-axis represents fluorescence units. (a) Multiplex amplifications from 5 ng of a female tiger DNA (RFET04). (b) Multiplex amplifications from the same female individual (RFET04) with template DNA diluted to 1 ng. (c) Multiplex amplifications from 5 ng of a male tiger DNA (RFET05) and only Panel D is shown, which included the sex-identifying marker *SRY*.

clude that two perfectly-matched tiger samples are from one individual instead of a random match. For all other sample pairs that partially matched using panel C, incongruent alleles were detected from at least two of the 6 loci, at least one of which carried heterozygous genotypes in both samples, excluding the possibility that the mismatch was due to allele dropout.

We applied the tiger subspecies diagnostic system to assess genetic ancestry in these 12 tiger samples, as described previously (2008, 2010b). All 6 tigers were designated as Amur tigers with high confidence (Bayesian assignment probabilities are from 0.989 to 0.993 [Table 1, Fig. S1]).

DISCUSSION

Based on the mtDNA-based and microsatellite-based molecular genetic marker system that has been established and applied in the tiger (Luo *et al.* 2004, 2008, 2010b), we have now further developed a flexible and efficient genotyping system for identifying tiger subspecies, sex and individuals. The "tigrisPlex" system contains 22 STR loci and a gender-identifying *SRY* marker, and is organized into four multiplex panels that can be used singly, or in combinations as needed to increase precision and efficiency. Either all 4 panels or combinations of panels can be chosen to acquire the resolution needed in individual studies with various research purposes, and adjusted dependent on available time, labor and financial capacity. If sample quantity is constrained, we recommend beginning with panel C, which demonstrated the greatest capacity for individual identification, and adding other panels (B, A and D) subsequently.

In this study, we applied "tigrisPlex" to 12 specimens with undisclosed identity from the Russian Far East and correctly identified the number of individuals, their sex and subspecies. With reference samples available for calibration and comparison to voucher subspecies (Luo *et al.* 2004), this STR multiplex system is of significant practical value for wide-range applications in tiger conservation after the procedures previously described (Luo *et al*. 2008, 2010a,b), such as individual identification and verification of purebred or admixed genetic background from unknown captive tigers or confiscated tiger parts.

The Amur tiger (*P. t. altaica*) is currently found only in the Russian Far East and northeast China. According to investigations by TRAFFIC and other organizations, Russia is one of the key source countries involved in the illegal international tiger trade, (Verheij *et al.* 2010). The Amur tiger displays reduced genetic variability in comparison to other subspecies, possibly due to a post-glacial colonization of the region or a founder effect less than 10 000 years ago (Luo *et al.* 2004). Despite the low genetic variation, the "tigrisPlex" system proved sensitive and statistically powerful enough to distinguish Amur tiger individuals. Consequently, this approach has potential applications to the other extant tiger subspecies and even captive "generic" tigers, which number 13 000 to 20 000 worldwide (Luo *et al.* 2008; Nyhus *et al.* 2010).

One hurdle that has restricted cross-study comparisons of STR genetic profile is fluctuation in allele size,

Table 3 Application of the short tandem repeat (STR) multiplex system to tiger specimens with unknown identity from the Russian Far East

Sample	Matched sample	Panel C		Panel B + C		Panel A + B + C		Panel A + B + C + D		Sex	Bayesian assignment probability of subspecies (95% CI)
		PID	PIDsib	PID	PIDsib	PID	PIDsib	PID	PIDsib		
RFET01	RFET05	2.08×10^{-4}	2.41×10^{-2}	9.25×10^{-7}	1.40×10^{-3}	1.56×10^{-9}	7.84×10^{-5}	1.90×10^{-10}	2.58×10^{-5}	M	*P. t. altaica*, 0.989 (0.941,1.000)
RFET02	RFET12	5.65×10^{-4}	4.77×10^{-2}	2.97×10^{-7}	1.96×10^{-3}	5.13×10^{-10}	1.23×10^{-4}	3.28×10^{-11}	2.94×10^{-5}	M	*P. t. altaica*, 0.992 (0.956,1.000)
RFET03	RFET09	4.65×10^{-5}	1.49×10^{-2}	3.38×10^{-8}	5.87×10^{-4}	6.70×10^{-10}	9.14×10^{-5}	2.71×10^{-11}	2.28×10^{-5}	F	*P. t. altaica*, 0.993 (0.964,1.000)
RFET04	RFET11	2.53×10^{-4}	2.26×10^{-2}	8.40×10^{-7}	1.53×10^{-3}	4.04×10^{-9}	1.31×10^{-4}	4.90×10^{-10}	4.32×10^{-5}	F	*P. t. altaica*, 0.989 (0.939,1.000)
RFET06	RFET10	6.75×10^{-5}	2.55×10^{-2}	1.25×10^{-7}	1.35×10^{-3}	1.10×10^{-9}	1.47×10^{-4}	6.38×10^{-11}	4.37×10^{-5}	F	*P. t. altaica*, 0.990 (0.948,1.000)
RFET07	RFET08	9.93×10^{-4}	3.48×10^{-2}	4.23×10^{-6}	2.52×10^{-3}	8.05×10^{-8}	3.45×10^{-4}	3.26×10^{-9}	8.60×10^{-5}	M	*P. t. altaica*, 0.992 (0.959, 1.000)

PID, probability of Identity.

which can vary between different genotyping machines and even different runs on the same apparatus. In "tigrisPlex", three reference tiger DNA samples with known genotypes are provided for standardized calibration, therefore enabling the data generated with "tigrisPlex" to be readily compatible with the voucher tiger subspecies dataset (n = 113) from Luo *et al.* (2004). As information concerning geographic origin is critically needed for tiger management both in the wild and in captivity, as well as confiscated tiger parts from illegal trafficking, the standardized tiger STR multiplex system for subspecies identification represents a necessary toolkit for the tiger conservation community, due to the correlation between subspecies affiliation and their associated broad geographic range. However, in reality, shipping of the same reference tiger DNAs for use as calibration at different labs internationally may pose a challenge; therefore, an alternative option using synthetic DNA probes or other non-regulated DNAs for calibration is worth considering.

As complete genome information from thousands of individuals and species are becoming available, numerous tools for elucidating the origin, evolution, divergence and adaptation of free-ranging wildlife are also becoming available (Allendorf *et al.* 2010). Facilitated by the advances in next-generation sequencing technologies, single nucleotide polymorphism (SNP) markers are accelerating the field of conservation genomics due to their high-throughput capacity and genome-wide coverage (Allendorf *et al.* 2010). However, STR carries several apparent advantages over SNPs, including high variability and multi-allele characteristics, which will likely ensure their continued importance in forensic and conservation genetics (Butler *et al.* 2007). For instance, a subset of "tigrisPlex" markers was used to validate a tiger pedigree from captivity that included 38 closely related individuals (Xu *et al.* 2013). While 9 out of the 14 tested STR markers (>60%) detected two misplaced tigers from the family, only 6–9% of the genome-wide restriction-site-associated (RAD) DNA SNPs recognized the wrong paternity (Xu *et al.* 2013). Relative to SNP markers, much fewer STR loci are sufficient for resolving closely related populations, testing paternity or identifying individuals. In addition, STR genotyping at present is still more cost-effective compared to whole genome sequencing.

Commercial STR assays enabling co-amplification of up to 20 different loci have become widely used in forensic DNA typing in humans (Butler *et al.* 2007). Multiplex genotyping systems have also been developed for domestic animal species, such as the "meowplex" 12-plex STR assay for the domestic cat (Menotti-Raymond *et al.* 2005) and a 18-plex STR assay for the dog (Dayton *et al.* 2009). For many species, including tigers, non-invasively collected samples, such as hair and feces, represent a potential source of genetic materials that can be relatively easily collected in the field, but DNA is often degraded and in low quantities, requiring a highly sensitive DNA typing system. Although not yet tested specifically in fecal samples for potential allele dropout rate, the "tigrisPlex" system is robust and generates an easily interpretable product profile (Figs 1, 2) with as little as 1 ng of template DNA. However, with small amounts of DNA, balancing across all loci is often not perfect due to quality variation in STR amplicons. Further optimization in each individual case is likely needed by adjusting primer concentrations to obtain balanced signal strengths. Caution should also be taken for sex identification of non-invasive samples, as the absence of signal at the male-specific *SRY* locus might result from either female sex or null amplification. Multiple repeats are required to confirm the female sex. Alternatively, using a *ZFX*/*ZFY* or *AMELX*/*AMELY* sexing system, in which females are distinguished from males by two PCR products, may be worth considering (Pilgrim *et al*. 2005; McEwing *et al*. 2011).

In summary, the "tigrisPlex" tiger STR multiplex system is a flexible, cost-efficient and relatively high-throughput genome profiling approach that has been optimized and is ready for application to wildlife conservation and molecular ecological questions concerning the tiger. As a cautionary note, DNA degradation in non-invasively collected or historic specimens may impact the success rate of the multiplex PCR reactions, leading to biased amplification among loci, allele dropout or failure in amplification. Future improvement of amplification efficiency for degraded DNA may be realized by reducing the size of STR amplicons (Butler *et al.* 2003; Mondol *et al.* 2012). Primers can be designed readily from current "tigrisPlex" STR loci by moving the primer position towards the tandem repeat region. This approach will permit the use of STR loci already present in the system and generate genotyping data that is directly compatible to the "tigrisPlex" voucher tiger reference dataset. We envision that this assay will play an important role in facilitating informed conservation management decisions for one of the world's most iconic flagship species.

ACKNOWLEDGMENTS

This research was supported by the Peking-Tsinghua Center for Life Sciences at Peking University (PKU) and the National Natural Science Foundation of China (NSFC No. 31271320, No. 31471179). We thank Zhi Dong, Yan Zhuang, Yuechen Liu, Lin Miao and Galina Nemkova for technical assistance. Reference tiger samples for calibration were kindly provided by Dr Stephen O'Brien under a Material Transfer Agreement signed between PKU and the National Cancer Institute, USA, and in full compliance with the Convention on International Trade in Endangered Species. We also thank Dale Miquelle for helpful comments on an earlier version of the manuscript.

REFERENCES

Allendorf FW, Hohenlohe PA, Luikart G (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**, 697–709.

Altschul SF, Madden TL, Schäffer AA *et al.* (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–402.

Borthakur U, Barman RD, Das C *et al.* (2011). Noninvasive genetic monitoring of tiger (*Panthera tigris tigris*) population of Orang National Park in the Brahmaputra floodplain, Assam, India. *European Journal of Wildlife Research* **57**, 603–13.

Butler JM, Shen Y, Mccord BR (2003). The development of reduced size STR amplicons as tools for analysis of degraded DNA. *Journal of Forensic Sciences* **48**, 1054–64.

Butler JM, Coble MD, Vallone PM (2007). STRs vs. SNPs: Thoughts on the future of forensic DNA testing. *Forensic Science, Medicine, and Pathology* **3**, 200–5.

Cho YS, Hu L, Hou H *et al.* (2013). The tiger genome and comparative analysis with lion and snow leopard genomes. *Nature Communications* **4**, 2433.

Chundawat RS, Habib B, Karanth U *et al.* (2011). *Panthera tigris*. In: *IUCN Red List of Threatened Species. Version* 2013.2. [Cited 15 Mar 2014.] Available from URL: http://www.iucnredlist.org. IUCN, Gland, Switzerland and Cambridge, UK.

Dayton M, Koskinen MT, Tom BK *et al.* (2009). Developmental validation of short tandem repeat reagent kit for forensic DNA profiling of canine biological material. *Croatian Medical Journal* **50**, 268–85.

Dinerstein E, Loucks C, Wikramanayake E *et al.* (2007). The fate of wild tiger. *Bioscience* **57**, 508–14.

Driscoll CA, Yamaguchi N, Bar-Gal GK *et al.* (2009). Mitochondrial phylogeography illuminates the origin of the extinct Caspian tiger and its relationship to the Amur tiger. *PLoS ONE* **4**, e4125.

Global Tiger Initiative Secretariat (2011). *The Global Tiger Recovery Program*. World Bank, Washington, DC, 58 pp.

Henry P, Miquelle D, Sugimoto T *et al.* (2009). *In situ* population structure and *ex situ* representation of the endangered Amur tiger. *Molecular Ecology* **18**, 3173–84.

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099–106.

Kitpipit T, Tobe SS, Kitchener AC *et al.* (2012). The development and validation of a single SNaPshot multiplex for tiger species and subspecies identification-implications for forensic purposes. *Forensic Science International: Genetics* **6**, 250–7.

Luo SJ, Johnson WE, Martenson J *et al.* (2008). Subspecies genetic assignments of worldwide captive tigers increase conservation value of captive populations. *Current Biology* **18**, 592–6.

Luo SJ, Johnson WE, O'Brien SJ (2010a). Applying molecular genetic tools to tiger conservation. *Integrative Zoology* **5**, 351–62.

Luo SJ, Johnson WE, Smith JLD *et al.* (2010b). What is a tiger? Genetics and Phylogeography. In: Tilson R, Nyhus PJ, eds. *Tigers of the World: The Science, Politics, and Conservation of Panthera tigris*, 2nd edn. Elsevier/Academic Press, New York.

Luo SJ, Kim JH, Johnson WE *et al.* (2004). Phylogeography and genetic ancestry of tigers (*Panthera tigris*). *Plos Biology* **2**, 2275–93/e442.

Menotti-Raymond M, David VA, Lyons LA *et al.* (1999). A genetic linkage map of microsatellites in the domestic cat (*Felis catus*). *Genomics* **57**, 9–23.

Menotti-Raymond MA, David VA, Wachter LL *et al.* (2005). An STR forensic typing system for genetic individualization of domestic cat (*Felis catus*) samples. *Journal of Forensic Sciences* **50**, 1061–70.

McEwing R, Ouitavon K, Rovie-Ryan JJ *et al.* (2012). Molecular sexing of tigers, *Panthera tigris*. *Conservation Genetics Resources* **4**, 299–301.

Mills JA, Jackson P (1994). *Killed for a Cure: A Review of the Worldwide Trade in Tiger Bone*. TRAFFIC International, Cambridge.

Mondol S, Karanth KU, Ramakrishnan U (2009a). Why the Indian Subcontinent holds the key to global tiger recovery. *Plos Genetics* **5**, e1000585.

Mondol S, Karanth UK, Kumar NS *et al.* (2009b). Evaluation of non-invasive genetic sampling methods for estimating tiger population size. *Biological Conservation* **142**, 2350–60.

Mondol S, Thatte P, Vadav P *et al.* (2012). A set of miniSTRs for population genetic analyses of tigers (*Panthera tigris*) with cross-species amplification for seven other Felidae. *Conservation Genetics Resources* **4**, 63–6.

Morin PA, Manaster C, Mesnick SL *et al.* (2009). Normalization and binning of historical and multi-source microsatellite data: overcoming the problems of allele size shift with allelogram. *Molecular Ecology Resources* **9**, 1451–5.

Nyhus PJ, Tilson R, Hutchins M (2010). Thirteen thousand and counting: how growing captive tiger populations threaten wild tigers. In: Tilson R, Nyhus PJ, eds. *Tigers of the World: The Science, Politics, and Conservation of Panthera tigris*, 2nd edn. Elsevier/Academic Press, New York.

Ogden ROB (2011). Unlocking the potential of genomic technologies for wildlife forensics. *Molecular Ecology Resources* **11**, 109–16.

Park SDE (2001). Trypanotolerance in west african cattle and the population genetic effects of selection (PhD thesis). University of Dublin, Dublin, Ireland.

Pilgrim KL, McKelvey KS, Riddle AE *et al.* (2005). Felid sex identification based on noninvasive genetic samples. *Molecular Ecology Notes* **5**, 60–1.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–59.

Raymond M, Rousset F (1995). Genepop (Version-1.2) – Population-genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248–9.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation Statistical Computing, Vienna, Austria.

Rosenberg NA, Li LM, Ward R *et al.* (2003). Informativeness of genetic markers for inference of ancestry. *The American Journal of Human Genetics* **73**, 1402–22.

Seidensticker J, Gratwicke B, Shrestha M (2010). How many wild tigers are there? An estimate for 2008. In: Tilson R , Nyhus P, eds. *Tigers of the World: The Science, Politics, and Conservation of Panthera tigris*, 2nd edn. Elsevier/Academic Press, New York.

Sharma R, Stuckas H, Bhaskar R *et al.* (2009). mtDNA indicates profound population structure in Indian tiger (*Panthera tigris tigris*). *Conservation Genetics* **10**, 909–14.

Sharma S, Dutta T, Maldonado JE *et al.* (2013). Forest corridors maintain historical gene flow in a tiger metapopulation in the highlands of central India. *Proceedings of the Royal Society B* **280**, 20131506.

Slatkin M (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457–62.

Untergasser A, Nijveen H, Rao X *et al.* (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research* **35**, W71–4.

Valière N (2002). Gimlet: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes* **2**, 377–9.

Verheij P, Foley KE, Engel K (2010). Reduced to Skin and Bones: An Analysis of Tiger Seizures from 11 Tiger Range Countries (2000–2010). TRAFFIC International, Cambridge.

Waits LP, Luikart G, Taberlet P (2001). Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* **10**, 249–56.

Walston J, Robinson JG, Bennett EL *et al.* (2010). Bringing the tiger back from the brink, the six percent solution. *Plos Biology* **8**, e1000485.

Wasser SK, Clark WJ, Drori O *et al.* (2008). Combating the illegal trade in African elephant ivory with DNA forensics. *Conservation Biology* **22**, 1065–71.

Xu X, Dong GX, Hu XS *et al.* (2013). The genetic basis of white tigers. *Current Biology* **23**, 1031–5.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.

Table S1 Samples used in the study and calibrated microsatellite genotypes using the tiger short tandem repeat (STR) multiplex system

Figure S1 Bayesian population clustering analysis of 119 tiger specimens (including 113 voucher tigers) based on "tigrisPlex" STR multiple assay assigned 6 individuals with unknown identity as Amur tigers (*Panthera tigris altaica*). Simulations were set at 100 000 burn-in period followed by 1 000 000 replicates in STRUCTURE (Pritchard *et al.* 2000). Here the population structure is shown when K = 7, which produced the highest probability among other choices of K. Three STRUCTURE runs produced almost identical individual affiliations. Each individual is represented by a vertical bar partitioned into 7 colored segments representing individual affiliations (q) to 5 designated tiger subspecies (*P. t. altaica*, *P. t. corbetti*, *P. t. jacksoni*, *P. t. sumatrae* and *P. t. tigris*).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing content) should be directed to the corresponding author for the article.

Cite this article as:

Zou Z-T, Uphyrkina OV, Fomenko P, Luo S-J (2015). The development and application of a multiplex short tandem repeat (STR) system for identifying subspecies, individuals and sex in tigers. *Integrative Zoology* **10**, 376–88.