Yi Li[1]
Yine Hong[1]
Xiujiang Li[1]
Jinmeng Yang[1]
Lanjiang Li[2]
Ying Huang[2]
Chuanchao Wang[3]
Hui Li[3]
Bingying Xu[2]

[1]Department of Criminal Science and Technology, Dali Police Station, Dali, P. R. China
[2]Kunming Medical University, Kunming, P. R. China
[3]State Key Laboratory of Genetic Engineering, and Ministry of Education Key Laboratory of Contemporary Anthropology, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai, P. R. China

Received March 11, 2015
Revised May 16, 2015
Accepted May 19, 2015

Research Article

Allele frequency of 19 autosomal STR loci in the Bai population from the southwestern region of mainland China

The aim of this study was to investigate a 19 STR loci database using the Bai population from China. This multiplex amplification kit included 13 CODIS STR markers and six plus STR markers (D19S433, Penta E, D2S1338, Penta D, D6S1043, and D12S391) that were successfully analyzed by using 1158 DNA samples from the Bai population from the southwestern part of mainland China. These results indicate that this multiplex amplification kit may provide significant polymorphic information for kinship testing and relationship investigations.

Keywords:
Bai / Goldeneye™ 20A kit / Population study / Short tandem repeats
DOI 10.1002/elps.201500129

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

STR is a repeated DNA sequence that has a 2–7 bp core repeat unit; the number of repeats is highly variable among individuals. Thus, it has a high practical value in forensic identification and paternity testing. The Goldeneye DNA 20A PCR amplification kit is a commercial PCR amplification kit that is widely used in the crime STR database in mainland China. It consists of 13 CODIS loci and six plus STR loci (D19S433, Penta E, D2S1338, Penta D, D6S1043, and D12S391) [1]. In this study, we reported the STR data of the Bai population using this commercial kit. There are nearly 1.56 million Bai people living in Yunnan province, and over 1.11 million live in the Dali Bai Autonomous Prefecture (Supporting Information Fig. 1). The Bai language is a particular dialect that belongs to the Sino-Tibetan language family [2, 3].

2 Materials and methods

2.1 Ethics statement

All participants provided written informed consent prior to inclusion. Blood sample collection was conducted in conformity with ethical and human research principles of the Institutional Ethics Committee, Kunming Medical University, China.

2.2 Sample collection and DNA extraction

We studied 1158 unrelated and healthy individuals of the Bai from Dali. Blood samples or buccal swabs were collected after informed consent.

2.3 PCR amplification and STR typing

Genomic DNA was extracted by the Chelex100 method [4]. PCR amplification was performed in a GeneAmp PCR SYSTEM 9700 (AB: Applied Biosystems, Foster City, CA, USA) following the manufacturer's instructions for the Goldeneye 20Akit (Peoplespot Company, Beijing, China). PCR products were separated by the GeneScan500 LIZ (AB) on a 3130 Genetic Analyzer (AB).

The laboratory participants in this study were accredited according to the ISO/IEC 17025:2005 General Requirements for the Competence of Testing and Calibration Laboratories (CNAS-CL01 Accreditation Criteria for the Competence of Testing and Calibration Laboratories).

Correspondence: Professor Bingying Xu, Kunming Medical University, Chunrong West Road No.1168 Yuhua Subdistrict, Chenggong New Town, Kunming 650500, P. R. China
E-mail: bingying_xu@126.com

Abbreviation: PE, probability of paternity exclusion

Colour Online: See the article online to view Figs. 1, 2 and Table 1 in colour.

Table 1. Allele frequency value for the Bai samples at 19 autosomal STR loci in Goldeneye DNA 20A PCR amplification kit (1158)

Allele	D3S1358	FGA	D21S11	D5S818	D7S820	CSF1PO	D16S539	D19S433	Penta E	vWA	D8S1179	D18S51	D13S317	TPOX	TH01	D2S1338	Penta D	D6S1043	D12S391
5									0.0553										
6							0.0004							0.0004	0.0872		0.0030		
7				0.0177	0.0009	0.0017			0.0013			0.0009		0.0009	0.2517		0.0220		
8				0.0022	0.1287		0.0091		0.0125		0.0013		0.2759	0.5255	0.0639		0.0453		
9				0.0747	0.0743	0.0440	0.2586	0.0017	0.0073		0.0009		0.1503	0.1248	0.5190		0.3489	0.0030	
9.1					0.0004														
9.3															0.0540				
10				0.1770	0.1753	0.2362	0.1205	0.0035	0.0367		0.0954	0.0009	0.1364	0.0272	0.0237		0.1157	0.0276	
11				0.3515	0.3156	0.2621	0.2897	0.0039	0.1835		0.0794	0.0026	0.2245	0.2949	0.0004		0.1641	0.1377	
11.2								0.0013											
12	0.0004			0.2401	0.2660	0.3726	0.2232	0.0440	0.1118		0.1321	0.0281	0.1675	0.0255			0.1516	0.1395	
12.2								0.0186											
13				0.1222	0.0315	0.0760	0.0855	0.2517	0.0583	0.0022	0.2090	0.2055	0.0380	0.0004			0.1088	0.1239	
13.2								0.0510											
14	0.0406			0.0130	0.0073	0.0069	0.0125	0.2414	0.0984	0.2228	0.1986	0.2077	0.0069	0.0004			0.0320	0.1179	0.0009
14.2								0.1066											
15	0.3316			0.0017		0.0004		0.0816	0.0872	0.0289	0.1844	0.2060	0.0004				0.0056	0.0142	0.0155
15.2								0.1334		0.0004									
16	0.3541						0.0004	0.0138	0.0920	0.1740	0.0790	0.1403				0.0147	0.0026	0.0013	0.0069
16.2								0.0397											
17	0.2029							0.0017	0.0518	0.2560	0.0181	0.0704				0.0514	0.0004	0.0566	0.0993
17.2								0.0060											
17.3																			0.0004
18	0.0622	0.0492							0.0682	0.2012	0.0017	0.0315				0.0933		0.1701	0.2556
18.2	0.0004																	0.0043	
18.3																			0.0009
19	0.0078	0.0583							0.0497	0.0907		0.0402				0.1727		0.1503	0.2116
19.3																		0.0022	
20		0.0661							0.0440	0.0220		0.0220				0.1235		0.0436	0.1658
20.2		0.0009																	
20.3																		0.0009	
21		0.1032							0.0186	0.0017		0.0190				0.0458		0.0035	0.0959
21.2		0.0056																	
21.3																		0.0035	

Table 1. Continued

Allele	D3S1358	FGA	D21S11	D5S818	D7S820	CSF1PO	D16S539	D19S433	Penta E	vWA	D8S1179	D18S51	D13S317	TPOX	TH01	D2S1338	Penta D	D6S1043	D12S391
22		0.1615							0.0117			0.0164				0.0475			0.0855
22.2		0.0056																	
23		0.2180							0.0039			0.0052				0.2297			0.0453
23.2		0.0121																	
24		0.1667							0.0052			0.0022				0.1438			0.0112
24.2		0.0155																	
25		0.0877							0.0004			0.0013				0.0665			0.0048
25.2		0.0043																	
26		0.0332							0.0013							0.0082			0.0004
26.2		0.0022																	
27		0.0065	0.0009						0.0009							0.0030			
27.2		0.0004																	
28		0.0030	0.0358																
28.2			0.0138																
29			0.2530																
29.2			0.0048																
30			0.2608																
30.2			0.0229																
31			0.1015																
31.2			0.0747																
32			0.0225																
32.2			0.1494																
33			0.0026																
33.2			0.0514																
34			0.0009																
34.2			0.0043																
35.2			0.0009																
PD	0.867	0.968	0.947	0.910	0.912	0.881	0.913	0.953	0.984	0.934	0.955	0.958	0.933	0.796	0.834	0.965	0.936	0.971	0.952
PIC	0.668	0.855	0.803	0.732	0.742	0.683	0.742	0.817	0.901	0.776	0.824	0.824	0.774	0.560	0.608	0.846	0.776	0.862	0.813
PE	0.459	0.724	0.629	0.548	0.581	0.471	0.579	0.652	0.807	0.597	0.694	0.642	0.584	0.315	0.351	0.696	0.603	0.725	0.629
Ho	0.850	0.816	0.790	0.727	0.719	0.792	0.789	0.851	0.828	0.799	0.823	0.772	0.864	0.906	0.619	0.647	0.802	0.865	0.816
He	0.844	0.825	0.776	0.729	0.718	0.803	0.778	0.861	0.836	0.805	0.843	0.767	0.869	0.908	0.620	0.652	0.800	0.875	0.834
HWE	0.109	0.257	0.185	0.360	0.095	0.434	0.180	0.735	0.098	0.548	0.409	0.491	0.863	0.212	0.359	0.832	0.809	0.773	0.042

The highest and lowest PE, PD, and TPI are highlighted in red.
PD: power of discrimination, PIC: polymorphism information content, PE: power of exclusion, HWE: Hardy–Weinberg equilibrium (p), Ho: observed heterozygosity, He: expected heterozygosity.

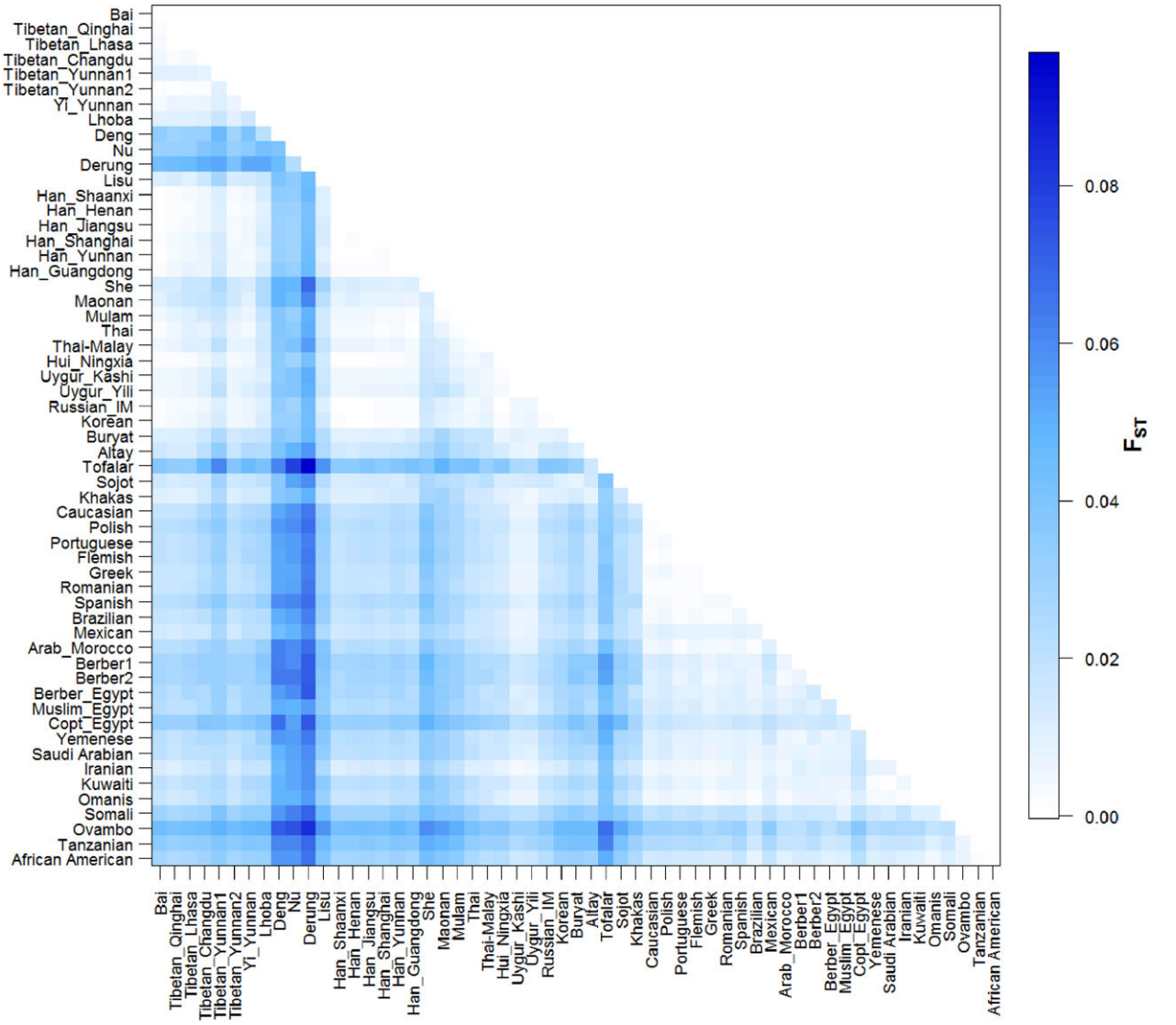

Figure 1. Pair Fst Matrix of Bai in Yunnan and other 56 worldwide populations.

2.4 Statistical analysis

Allele frequency and forensic parameters were calculated using Powerstats v1.2 software. The *p*-values of the exact test for Hardy–Weinberg equilibrium, observed heterozygosity and expected heterozygosity were calculated using CERVUS program [5]. The average number of pairwise differences, Slatkins linearized Fst, and coancestry coefficients were calculated using Arlequin v3.5 software [6] with the raw genotypic data of 13 STRs (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, D13S317, D16S539, D2S1338, D19S433, vWA, D18S51, D5S818, and FGA) from 57 populations all around the world (Bai and [7–35]). The detailed population genetic structure was performed using the above 13 STRs and model-based clustering method implemented in Structure 2.3.4 [36, 37] under assumptions of admixture, LOCPRIOR model, and correlated allele frequencies. Each run used 100 000 estimation iterations for $K = 2$–12 after a 20 000 burn-in length with three replicates. Posterior probabilities for each K were computed for each set of runs.

3 Results and discussion

3.1 Forensic parameter analysis

Allele frequencies and forensic parameters are presented in Table 1. The result shows that the highest value of the power of discrimination and probability of paternity exclusion were observed for Penta E, whereas the lowest value was observed for TPOX. The combined power of discrimination value and combined probability of paternity exclusion value were both >0.999999999. After Bonferroni's correction

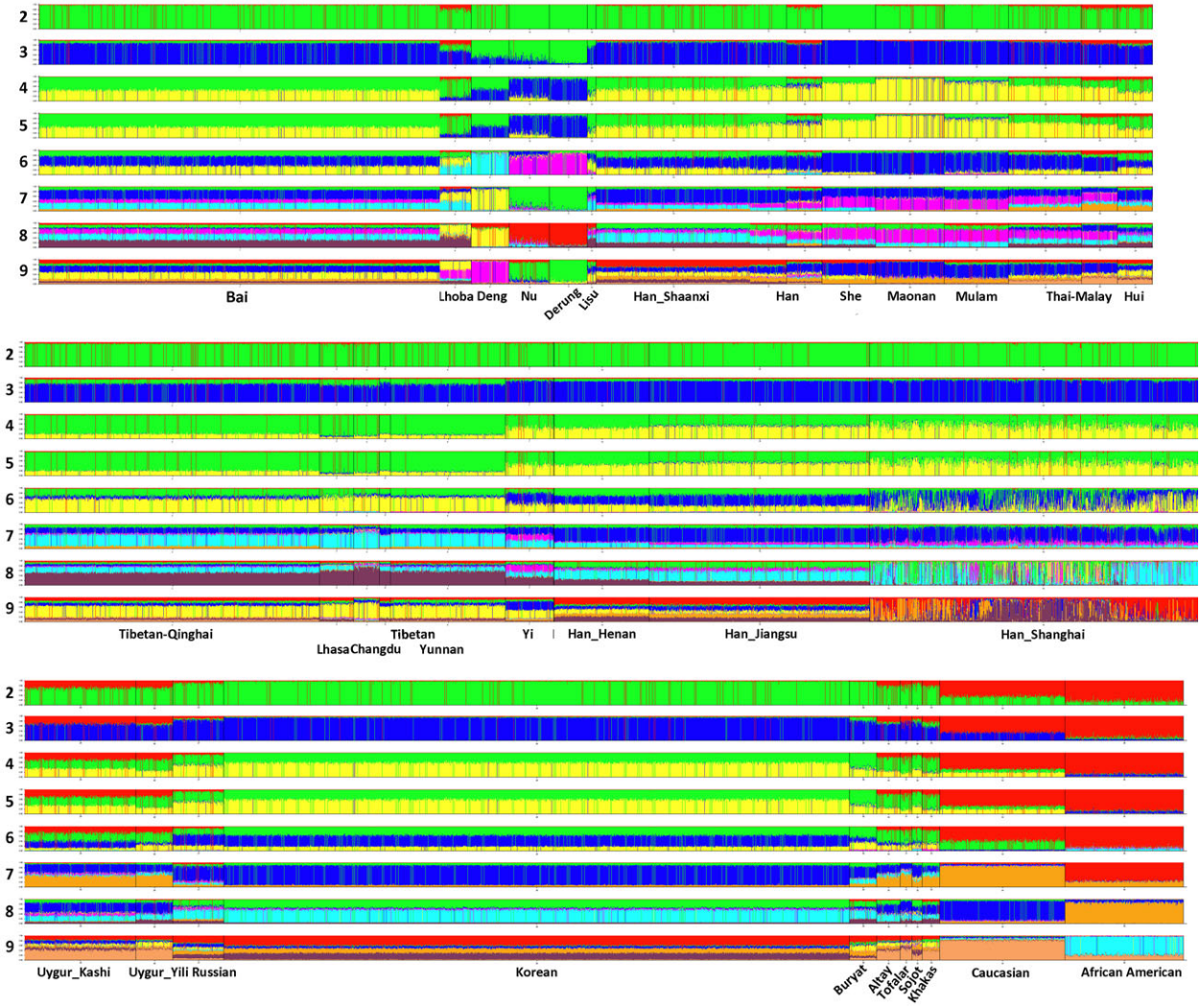

Figure 2. Estimated population genetic structure of Bai in Yunnan and other 30 worldwide populations using Structure 2.3.4 software.

($p = 0.00263$) [38], there are no deviations from Hardy–Weinberg equilibrium.

3.2 Population genetic distance study

The pair Fst matrix (Fig. 1 and Supporting Information Table 1) and detailed population genetic structure (Fig. 2) were calculated using genotype data from 57 populations from all around the world. The lowest Fst values have been observed in the Bai and Han Chinese in Henan, Shaanxi, Jiangsu, and Yunnan (all <0.001). The Fst in the Bai with Tibetan populations were slightly higher than the values with the Han Chinese but were still lower than those with Lolo-Burmese and Kam-Sui populations. The structure analysis also confirmed the genetic similarities of the Bai, Han Chinese, and Tibetan populations. Tibetan, southern Han Chinese (Yunnan and Guangdong), and southern indigenous populations (such as She, Mulam, and Maonan) contributed most to the gene pool of the Bai people. The results indicated that differences could be found in allelic frequency distribution between different populations of geographic locations and varied linguistic families.

4 Concluding remarks

In summary, we provided complete data for the 19 STR loci in the southwestern Bai population for the first time. Based on allelic frequency and statistical parameters for the Bai people, it can be concluded that these 19 autosomal STR loci indeed represent a robust and efficient approach in forensic human identification and parentage testing.

This study was financially supported by National Natural Science Foundation of China, (81260467) and multiple Science and Technology Planning Projects of Yunnan Province, PR China, (2010CD226, 2011CA022, and 2011FZ113).

The authors have declared no conflict of interest.

5 References

[1] Huang, Y. M., Wang, J., Jiao, Z., Yang, L., Zhang, X., Tang, H., Liu, Y., *Int. J. Legal Med.* 2013, *127*, 587–590.

[2] Cang, M., *On the Migration Culture of the Ethnic Groups in Yunnan Kunming*. The Nationalities Publishing House of Yunnan. 1997.

[3] You, Z., *Yunnan Ethnic History Kunming*. Yunnan People's Publishing House. 1994.

[4] Walsh, P. S., Metzger, D. A., Higuchi, R., *Biotechniques* 1991, *10*, 506–513.

[5] Slate, J., Marshall, T., Pemberton, J., *Mol. Ecol.* 2000, *9*, 801–808.

[6] Excoffier, L., Lischer, H. E., *Mol. Ecol. Resour.* 2010, *10*, 564–567.

[7] Kang, L., Li, S., Gupta, S., Zhang, Y., Liu, K., Zhao, J., Jin, L., Li, H., *J. Hum. Genet.* 2010, *55*, 270–277.

[8] Yan, J., Shen, C., Li, Y., Yu, X., Xiong, X., Mu, H., Huang, Y., Deng, Y., Yu, J., *Forensic Sci. Int.* 2007, *169*, e3–e6.

[9] Wang, H.-D., Wu, D., Feng, Z.-Q., Jing, Z.-A., Li, T., Guo, Q.-N., Zhang, X.-P., Hou, Q.-F., Guo, L.-J., Kang, B., Zhang, H., Zhu, B.-F., Liao, S.-X., *Electrophoresis* 2014, *35*, 1509–1514.

[10] Zhai, D., Yang, J., Huang, Y., Chen, L., Wu, D., Wu, J., Xu, B., *Forensic Sci. Int. Genet.* 2014, *13*, e22–e24.

[11] Deng, Q., Xu, L., Gong, J., Zhou, L., Li, S., Deng, X., Luo, G., Xie, X., *J. Genet. Genomics* 2007, *34*, 1072–1079.

[12] Zhu, B. F., Yang, G., Shen, C. M., Qin, H. X., Fan, S. L., Deng, L. B., Chen, F., Chen, L. P., Wang, H. D., Wang, Z. Y., Lucas, R., Lian, Z. M., *Mol. Biol. Rep.* 2010, *37*, 3889–3895.

[13] Wu, Y. M., Zhang, X. N., Zhou, Y., Chen, Z. Y., Wang, X. B., *Forensic Sci. Int. Genet.* 2008, *2*, e15–e18.

[14] Chen, J. G., Pu, H. W., Chen, Y., Chen, H. J., Ma, R., Xie, S. T., Zhang, L. P., *Forensic Sci. Int. Genet.* 2012, *6*, e178–e179.

[15] Zhu, J., Li, J., Guo, Y., Liu, K., Zhu, B., Liu, Y., *J. Forensic Sci.* 2005, *50*, 1510–1511.

[16] Wang, Z., Huang, P., Chen, L., Xu, Y., He, K., Zhu, X., *J. Forensic Sci.* 2005, *50*, 1508–1509.

[17] Zhu, B., Wang, Z., Wu, Q., Sedike, M. E., Zhu, J., Huang, P., Xu, Y., Liu, Y., *J. Forensic Sci.* 2005, *50*, 1235–1236.

[18] Li, C., Li, L., Zhao, Z., Lin, Y., Que, T., Liu, Y., Xue, J., *Forensic Sci. Int. Genet.* 2009, *3*, e117–e118.

[19] Rerkamnuaychoke, B., Rinthachai, T., Shotivaranon, J., Jomsawat, U., Siriboonpiputtana, T., Chaiatchanarat, K., Pasomsub, E., Chantratita, W., *Forensic Sci. Int.* 2006, *158*, 234–237.

[20] Alenizi, M., Goodwin, W., Ismael, S., Hadi, S., *Leg. Med. (Tokyo)* 2008, *10*, 321–325.

[21] Coudray, C., Calderon, R., Guitard, E., Ambrosio, B., Gonzalez-Martin, A., Dugoujon, J. M., *Forensic Sci. Int.* 2007, *168*, e21–e24.

[22] Sanchez-Diz, P., Menounos, P. G., Carracedo, A., Skitsa, I., *Forensic Sci. Int. Genet.* 2008, *2*, e71–e72.

[23] Pereira, L., Alshamali, F., Andreassen, R., Ballard, R., Chantratita, W., Cho, N. S., Coudray, C., Dugoujon, J. M., Espinoza, M., Gonzalez-Andrade, F., Hadi, S., Immel, U. D., Marian, C., Gonzalez-Martin, A., Mertens, G., Parson, W., Perone, C., Prieto, L., Takeshita, H., Rangel Villalobos, H., Zeng, Z., Zhivotovsky, L., Camacho, R., Fonseca, N. A., *Int. J. Legal Med.* 2011, *125*, 629–636.

[24] Ossmani, H. E., Talbi, J., Bouchrif, B., Chafik, A., *Leg. Med. (Tokyo)* 2009, *11*, 155–158.

[25] Muro, T., Fujihara, J., Imamura, S., Nakamura, H., Yasuda, T., Takeshita, H., *Leg. Med. (Tokyo)* 2008, *10*, 157–159.

[26] Marian, C., Anghel, A., Bel, S. M., Ferencz, B. K., Ursoniu, S., Dressler, M., Popescu, O., Budowle, B., *Forensic Sci. Int.* 2007, *170*, 73–75.

[27] Kutanan, W., Kitpipit, T., Phetpeng, S., Thanakiatkrai, P., *J. Hum. Genet.* 2014, *59*, 675–681.

[28] Hernandez-Gutierrez, S., Hernandez-Franco, P., Martinez-Tripp, S., Ramos-Kuri, M., Rangel-Villalobos, H., *Forensic Sci. Int.* 2005, *151*, 97–100.

[29] Forward, B. W., Eastman, M. W., Nyambo, T. B., Ballard, R. E., *J. Forensic Sci.* 2008, *53*, 245.

[30] Coudray, C., Guitard, E., Keyser-Tracqui, C., Melhaoui, M., Cherkaoui, M., Larrouy, G., Dugoujon, J. M., *Forensic Sci. Int.* 2007, *167*, 81–86.

[31] Coudray, C., Guitard, E., El-Chennawi, F., Larrouy, G., Dugoujon, J. M., *Forensic Sci. Int.* 2007, *169*, 260–265.

[32] Butler, J. M., Schoske, R., Vallone, P. M., Redman, J. W., Kline, M. C., *J. Forensic Sci.* 2003, *48*, 908–911.

[33] Abu Halima, M. S., Bernal, L. P., Sharif, F. A., *Leg. Med. (Tokyo)* 2009, *11*, 203–204.

[34] Mertens, G., Mommers, N., Cardoen, E., DeBruyn, I., Jehaes, E., Rand, S., VanBrussel, K., Jacobs, W., *Int. Congr. Ser.* 2006, *1288*, 328–330.

[35] Immel, U. D., Erhuma, M., Mustafa, T., Kleiber, M., Klintschar, M., *Int. Congr. Ser.* 2006, *1288*, 421–423.

[36] Falush, D., Stephens, M., Pritchard, J. K., *Genetics* 2003, *164*, 1567–1587.

[37] Hubisz, M. J., Falush, D., Stephens, M., Pritchard, J. K., *Mol. Ecol. Resour.* 2009, *9*, 1322–1332.

[38] Dunn, O. J., *Ann. Math. Stat.* 1959, *30*, 192–197.

