Int J Legal Med (2012) 126:671–675
DOI 10.1007/s00414-012-0691-9

SHORT COMMUNICATION

Fifteen non-CODIS autosomal short tandem repeat loci multiplex data from nine population groups living in Taiwan

Hsiao-Lin Hwa · Yih-Yuan Chang · James Chun-I Lee · Chun-Yen Lin · Hsiang-Yi Yin · Li-Hui Tseng · Yi-Ning Su · Tsang-Ming Ko

Received: 11 November 2011 / Accepted: 17 January 2012 / Published online: 20 March 2012

Abstract The analysis of autosomal short tandem repeat (STR) loci is a powerful tool in forensic genetics. We developed a multiplex system in which 15 non-Combined DNA Index System autosomal STRs (D3S1744, D4S2366, D8S1110, D10S2325, D12S1090, D13S765, D14S608, Penta E, D17S1294, D18S536, D18S1270, D20S470, D21S1437, Penta D, and D22S683) could be amplified in one single polymerase chain reaction. DNA samples from 1,098 unrelated subjects of nine population groups living in Taiwan, including Taiwanese Han, indigenous Taiwanese of Taiwan Island, Tao, mainland Chinese, Filipinos, Thais, Vietnamese, Indonesians, and Caucasians, were collected and analyzed using this system. The distributions of the allelic frequencies and the forensic parameters of each population group were presented. The combined discrimination power and the combined power of exclusion were high in all population groups tested in this study. A multidimensional scaling plot of these nine population groups based on the Reynolds' genetic distances calculated from 15 autosomal STRs was constructed, and the genetic substructure in this area was presented. In conclusion, this 15 autosomal STR multiplex system provides highly informative STR data and appears useful in forensic casework and parentage testing in different populations.

Keywords Forensic genetics · Multiplex polymerase chain reaction system · Autosomal short tandem repeats · Discrimination power

Electronic supplementary material The online version of this article (doi:10.1007/s00414-012-0691-9) contains supplementary material, which is available to authorized users.

H.-L. Hwa · J. C.-I. Lee · H.-Y. Yin
Department and Graduate Institute of Forensic Medicine,
College of Medicine, National Taiwan University,
No. 7 Chung Shan S. Rd.,
Taipei, Taiwan

H.-L. Hwa
Department of Obstetrics and Gynecology,
National Taiwan University Hospital and College of Medicine,
National Taiwan University,
No. 7 Chung Shan S. Rd.,
Taipei, Taiwan

H.-L. Hwa · L.-H. Tseng · Y.-N. Su
Department of Medical Genetics,
National Taiwan University Hospital,
No. 7 Chung Shan S. Rd.,
Taipei, Taiwan

Y.-Y. Chang · T.-M. Ko (✉)
Genephile Bioscience Laboratory,
No. 10, 1 F, Lin Sen South Road, Zhong Zheng District,
Taipei 100, Taiwan
e-mail: tsangming@live.com

J. C.-I. Lee · C.-Y. Lin
Institute of Forensic Medicine, Ministry of Justice,
No. 123, Min'an St., Zhonghe Dist.,
New Taipei City, Taiwan

Introduction

Polymorphic autosomal short tandem repeat (STR) markers are presently the most effective and most widely used genetic markers for forensic human identity and parentage testing [1]. In forensic casework with a minimal amount of DNA sample, systems which type a large number of STR loci in one polymerase chain reaction (PCR) are required to increase the discrimination power and save the tested material [2]. Loci with short fragments aid the analysis of forensic samples especially because degraded DNA fragments are usually shorter than 300 bp [3, 4].

We have developed a 14 non-Combined DNA Index System (CODIS) autosomal STR multiplex system that can provide highly informative STR data and improve the ability to distinguish in addition to the commonly used

multiplex analysis systems including the 13 CODIS core STR loci [5]. In this study, we added another simple repeats non-CODIS autosomal STR loci (D10S2325) with a relatively short fragment (109–184 bp) to our multiplex system and allowed 15 STR markers to be analyzed in one PCR. Establishing the data of the allelic frequencies of different population groups is important for the further use of this newly developed STR multiplex system in forensic casework, possibly dealing with different population groups.

The aim of this study was to analyze the 15 non-CODIS autosomal STR markers of the different population groups living in Taiwan using the multiplex system [6–8]. The distribution of allelic frequencies and forensic parameters is presented. We also demonstrated the multidimensional scaling (MDS) plot based on the genetic distances between each population group.

Materials and methods

This retrospective study was approved by the institutional review board and is in accordance with the Helsinki Declaration. A total of 1,098 DNA samples from apparently healthy and unrelated subjects were analyzed. There were 209 Taiwanese Han (TWH), 212 indigenous Taiwanese of Taiwan proper (TWI), 97 Tao (TAO), 193 mainland Chinese

Table 1 Ranking observed heterozygosity of (A) the 15 non-CODIS STRs of our system, and (B) the 15 STRs of the AmpF*l*STR Identifiler PCR Amplification Kit (Applied Biosystems) among nine population groups in this study

Locus	Observed heterozygosity (%)									
	TWH	TWI	TAO	CHI	FIL	THA	VIE	IND	CAU	Combined
A (no.)	(209)	(212)	(97)	(193)	(112)	(51)	(123)	(57)	(44)	(1,098)
Penta E	87.560	86.321	86.598	89.637	88.393	86.275	91.057	82.456	84.091	87.614
D12S1090	86.603	84.434	83.505	85.492	83.036	86.275	88.618	82.456	75.000	84.882
D20S470	88.995	79.717	72.165	81.347	85.714	88.235	84.553	78.947	86.364	82.878
D14S608	86.124	82.547	81.443	81.347	82.143	78.431	81.301	80.702	81.818	82.423
D10S2325	75.120	86.321	78.351	81.865	83.036	94.118	81.301	75.439	90.909	81.785
D18S1270	79.426	76.887	71.134	80.311	83.929	82.353	84.553	84.211	84.091	79.964
D3S1744	80.383	75.943	80.412	79.275	79.464	80.392	78.862	85.965	79.545	79.326
D22S683	83.254	77.830	62.887	78.756	76.786	84.314	76.423	85.965	88.636	78.597
Penta D	76.555	74.528	70.103	75.130	72.321	82.353	88.618	71.930	81.818	76.503
D4S2366	74.163	68.396	81.443	76.166	80.357	70.588	80.488	84.211	84.091	76.138
D8S1110	77.033	69.340	80.412	78.238	70.536	80.392	82.927	84.211	63.636	76.047
D17S1294	71.770	75.472	57.732	73.057	70.536	70.588	71.545	71.930	65.909	71.038
D13S765	69.856	75.000	55.670	73.057	75.893	66.667	67.480	66.667	79.545	70.583
D21S1437	73.684	68.868	49.485	74.611	62.500	66.667	60.976	66.667	77.273	67.668
D18S536	66.986	66.038	50.515	67.876	61.607	54.902	71.545	73.684	79.545	65.756
B (no.)	(199)	(212)	(97)	(191)	(111)	(51)	(123)	(54)	(44)	(1,082)
D2S1338	86.935	77.830	87.629	83.770	84.685	94.118	85.366	85.185	81.818	84.288
D18S51	83.920	70.283	89.691	87.958	87.387	84.314	83.740	85.185	86.364	82.994
D21S11	81.910	81.132	69.072	82.723	81.081	86.275	83.740	83.333	86.364	81.331
D19S433	79.899	80.189	78.351	78.534	83.784	78.431	80.488	88.889	90.909	80.869
vWA	83.417	76.887	79.381	78.534	79.279	90.196	80.488	75.926	86.364	80.222
FGA	78.894	81.132	60.825	84.293	78.378	82.353	78.862	83.333	93.182	79.575
D8S1179	83.920	81.132	47.423	84.293	72.973	92.157	85.366	81.481	70.455	78.928
D16S539	76.382	77.830	77.320	73.822	76.577	82.353	76.423	83.333	84.091	77.264
D7S820	79.397	77.358	68.041	78.534	72.072	78.431	77.236	77.778	81.818	76.802
D5S818	76.884	73.585	55.670	78.534	76.577	78.431	81.301	74.074	75.000	74.954
D13S317	78.392	67.925	62.887	77.487	80.180	64.706	71.545	77.778	81.818	73.660
CSF1PO	68.342	74.057	69.072	79.581	68.468	80.392	76.423	68.519	65.909	72.921
D3S1358	72.362	69.340	79.381	72.775	69.369	70.588	65.854	66.667	77.273	71.257
TH01	63.317	70.755	52.577	65.445	74.775	74.510	70.732	75.926	75.000	67.837
TPOX	56.281	46.226	57.732	57.592	59.459	64.706	52.033	66.667	70.455	56.007

(CHI), 112 Filipinos (FIL), 51 Thais (THA), 123 Vietnamese (VIE), 57 Indonesians (IND), and 44 Caucasians (CAU, people with European and Near Eastern or South Asian ancestry, including ten from the USA, six from the UK, four from India, three from Australia, three from France, three from Germany, three from Brazil, two from New Zealand, two from Peru, two from South Africa, one from Denmark, one from Netherlands, one from Jordan, one from Syria, one from Canada, and one from Nicaragua). Thirty-seven parent–child pairs of parentage testing cases with a combined paternity index (CPI) below 1,000 and another 32 parent–child pairs with single-step mutations found in AmpF*l*STR Identifiler (Applied Biosystems, Foster city, CA, USA) loci were recruited for validation of the newly developed system. For these parent–child pairs, genotyping using the new 15 autosomal STR loci multiplex system was carried out. The CPI, using both our new system and the AmpF*l*STR Identifiler, was calculated and compared with the results from using the AmpF*l*STR Identifiler only.

One multiplex PCR for each sample was performed with 15 primer pairs (loci D3S1744, D4S2366, D8S1110, D10S2325, D12S1090, D13S765, D14S608, Penta E, D17S1294, D18S536, D18S1270, D20S470, D21S1437, Penta D, and D22S683) and a primer pair for amelogenin (AMEL). The 15 autosomal STRs and AMEL were typed following the methodology described previously, with minor modification [5]. Table S1 lists the primer sequences, the amount of each primer set, and dye labels used.

Genotyping was analyzed using either Genotyper or GeneMapper ID software (Applied Biosystems) by comparison with the allelic ladder and reference DNA control samples 9947A (female; Applied Biosystems), GM9948 (male; Coriell Institute for Medical Research, Camden, NJ), and GM3657 (male; Coriell Institute for Medical Research) as recommended [9]. The precision evaluation was performed as described previously [5, 10, 11]. DNA sequencing was performed for novel STRs, novel alleles, off-ladder alleles, and loci with limited sequence data using methods described previously [5].

Statistical analysis

The power of discrimination, the power of exclusion, and the mean exclusion chance were calculated [12–14]. The observed heterozygosity, the expected heterozygosity, the deviation from Hardy–Weinberg equilibrium (HWE) based on the exact test, the linkage disequilibrium analysis of these loci by the exact test, the Fst among different populations explored by analysis of molecular variance, and the Reynolds' genetic distances between population groups were calculated using Arlequin v3.1 software (http://cmpg.unibe.ch/software/arlequin3) [12–19]. A difference with a p value<0.05 was taken as statistically significant. Based on Reynolds' genetic distances, the MDS plot was constructed using the Statistical Package for the Social Sciences (SPSS) v16.0 (SPSS Inc., Chicago, IL, USA).

Results and discussion

The development of a new non-CODIS autosomal STRs multiplex can yield additional genotyping information in forensic casework. In this study, the 15 non-CODIS autosomal STR markers of the 1,098 studied DNA samples were successfully amplified in each multiplex PCR. Among the 15 STR loci, 11 revealed amplicons shorter than 300 bp.

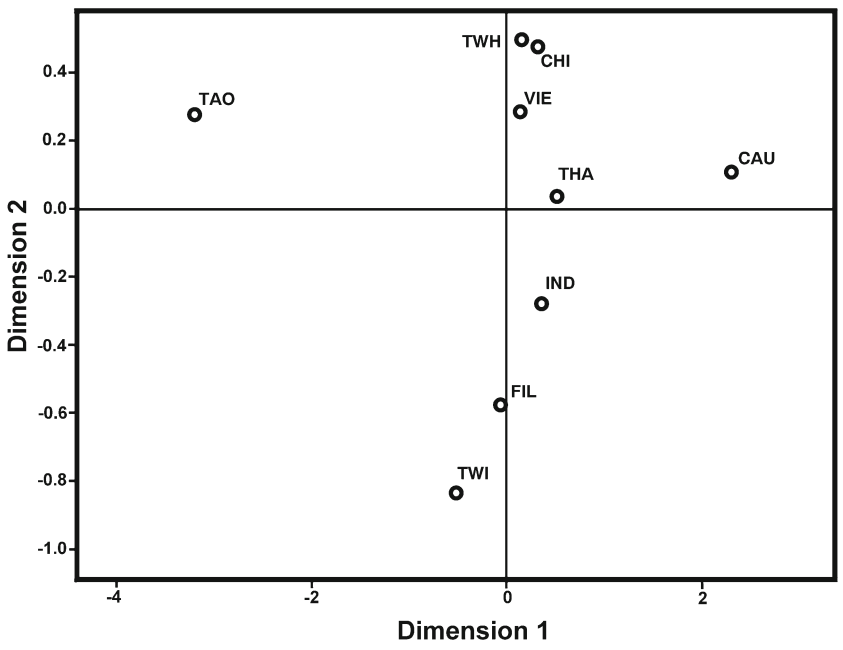

Fig. 1 The multidimensional scaling plot based on data from our 15 non-CODIS STRs of nine population groups

These 11 loci would be specifically useful to forensic casework dealing with degraded DNA samples. Examples of DNA profiles and chromatograms of the allelic ladder and an amplified DNA sample are illustrated in supplementary Fig. S1 and supplementary Fig. S2, respectively. All markers had high heterozygosity values and were located on different chromosomes or were separated by at least 50 cM from STRs of the commonly used AmpF*l*STR Identifiler PCR Amplification Kit (Applied Biosystems). Therefore, these markers may be unlinked to commonly genotyped STRs and can provide additional STR information for analysis of forensic casework and parentage testing. A detection limit of at least 125 pg of DNA was observed for this multiplex system. The sequence data and allelic nomenclature classified according to the guidelines of the International Society for Forensic Haemogenetics are presented in supplementary Table S2 [20]. The genotypes of common control DNAs (9947A, GM9948, and GM3657) are shown in supplementary Table S3 and can be used for calibration of these 15 non-CODIS markers.

Exact tests to determine the HWE were performed across all loci for each population, with a cutoff p value of 0.05. All 15 loci examined were regarded as adequate for this independence testing after Bonferroni correction [21, 22]. For pairwise linkage disequilibrium analysis of these loci, no statistically significant linkage disequilibrium was found (p values ranging from 0.05506 to 1.00000). These STR markers can be combined for biostatistical analysis.

The distributions of allelic frequencies and the forensic parameters for these 15 non-CODIS STRs of the 1,098 subjects from nine population groups are presented in Table S4. The loci with the highest or the lowest heterozygosity values differed among different population groups. The Penta E and D12S1090 loci showed the highest observed heterozygosity values in the combination of population groups (Table 1). The D18S536 and D21S1437 loci had the lowest observed heterozygosity values, especially among Asian groups (Table 1). The power of exclusion was high in all population groups analyzed in this study (Table S4).

A comparison of heterozygosity values between these 15 non-CODIS STRs and those of the AmpF*l*STR Identifiler PCR Amplification Kit (Applied Biosystems) from the same DNA samples was performed, with the results listed in Table 1. Almost all 30 markers had high enough observed heterozygosity values in different population groups to be considered useful in forensic casework. However, observed heterozygosity below 0.5 was noted in locus D21S1437 of our multiplex for TAO, in locus D8S1179 of CODIS for TAO, and in locus TPOX of CODIS for TWI (Table 1). These loci are less informative in specific population groups. All of our 15 non-CODIS loci had heterozygosities that were comparable to the CODIS loci. Therefore, our new STR multiplex can provide additional genotyping information for individual characterization.

In order to evaluate the forensic application of this newly developed STR multiplex in paternity testing, we genotyped 37 parent–child pairs with a CPI below 1,000, analyzed using the AmpF*l*STR Identifiler, and another 32 parent–child pairs with single-step mutations in loci of the AmpF*l*STR Identifiler. The CPI in these 37 parent–child pairs below 1,000 increased from 5,165 to 9,854,141.8 times (mean, 836,816.4 times). The CPI in 32 parent–child pairs with single-step mutations increased from 2,968.5 to 36,694,078.9 times (mean, 3,398,968.3 times). In addition to the autosomal STRs included in the AmpF*l*STR Identifiler, this set of autosomal STRs improved the ability to prove parentage and increased the CPI. They provide additional power to distinguish the possible single-step mutation in parent–child pairs.

Figure 1 illustrates the MDS plot constructed on the basis of Reynolds' genetic distances using the data of our 15 STR loci of these nine populations and is similar to the pattern based on Y-STR [23]. The grouping together of the Taiwanese Han, mainland Chinese, and Vietnamese showed the ethnogeographic relationship between these populations. The long distances between the Caucasians and other groups represented the difference between Caucasian and Asian populations.

In conclusion, these 15 non-CODIS autosomal STRs were highly polymorphic in different population groups. The 15 autosomal STR multiplex system can provide further informative genotype data in addition to the commonly used 13 CODIS loci in forensic casework and parentage testing.

Acknowledgments We would like to thank Ms. Hong-Ciao Wan, Ms. Pi-Mei Hsu, and Ms. Shwu-Fang Li for technical support in DNA extraction. This study was supported by a grant (no. 101-1301-05-0505) from the Ministry of Justice, Taiwan, ROC.

References

1. Butler JM (2006) Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci 51:253–265
2. Grubwieser P, Zimmermann B, Niederstätter H, Pavlic M, Steinlechner M, Parson W (2007) Evaluation of an extended set of 15 candidate STR loci for paternity and kinship analysis in an Austrian population sample. Int J Legal Med 121:85–89
3. Brinkmann B (1992) The use of STRs in stain analysis. In: Proceedings from the third international symposium on human identification. Promega Corporation, Madison, USA, pp. 357–373
4. Coble MD, Butler JM (2005) Characterization of new MiniSTR loci to aid analysis of degraded DNA. J Forensic Sci 50:43–53
5. Hwa HL, Chang YY, Lee JC, Yin HY, Tseng LH, Su YN, Ko TM (2011) Fourteen non-CODIS autosomal short tandem repeat loci multiplex data from Taiwanese. Int J Legal Med 125:219–226
6. Department of Statistics, Ministry of the Interior, Taiwan (2011) http://www.moi.gov.tw/stat/. Accessed 12 Sep 2011

7. Council of Indigenous Peoples, Executive Yuan, Taiwan (2011) http://www.apc.gov.tw/. Accessed 12 Sep 2011
8. National Immigration Agency Taiwan (2011) http://www.immigration.gov.tw/. Accessed 12 Sep 2011
9. Szibor R, Edelmann J, Hering S, Plate I, Wittig H, Roewer L, Wiegand P, Cali F, Romano V, Michael M (2003) Cell line DNA typing in forensic genetics—the necessity of reliable standards. Forensic Sci Int 138:37–43
10. Frank WE, Llewellyn BE, Fish PA, Riech AK, Marcacci TL, Gandor DW, Parker D, Carter RR, Thibault SM (2001) Validation of the AmpFlSTR™ Profiler Plus PCR amplification kit for use in forensic casework. J Forensic Sci 46:642–646
11. Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Brown AL, Budowle B (2001) Validation of STR typing by capillary electrophoresis. J Forensic Sci 46:661–676
12. Fisher RA (1951) Standard calculation for evaluation a bloodgroup system. Heredity 5:95–102
13. Garber R, Morris J (1983) General equation for the average power of exclusion for genetic system of dominant alleles in one-parent case of dispute parentage testing. In: Walker R (ed) International workshop on the inclusion probabilities in parentage testing. American Association of Blood Banks, Arlington, pp 277–280
14. Krüger J, Fuhrmann W, Lichte KH, Steffens C (1968) Zur Verwendung des Polymorphismus der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung. Dtsch Z Ges Gerichtl Med 64:127–146
15. Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590
16. Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241–253
17. Guo SW, Tompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48:361–372
18. Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.0: an integrated software package for population genetics data analysis. Evol Bioinformatics Online 1:47–50
19. Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics 105:767–779
20. Bär W, Brinkmann B, Budowle B et al (1997) DNA recommendations. Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. Forensic Sci Int 87:179–184
21. Bland JM, Altman DG (1995) Multiple significance tests: the Bonferroni method. Br Med J 310:170
22. Hill CR, Kline MC, Coble MD, Butler JM (2008) Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. J Forensic Sci 53:73–80
23. Hwa HL, Tseng LH, Ko TM, Chang YY, Yin HY, Su YN, Lee JC (2010) Seventeen Y-chromosomal short tandem repeat haplotypes in seven groups of population living in Taiwan. Int J Legal Med 124:295–300