

## Expansion of the Philippine Autosomal Short Tandem Repeat Population Database for DNA-based Paternity Testing

Sheena Marie B. Maiquilla, Jazelyn M. Salvador, Gayvelline C. Calacal,  
Frederick C. Delfin, Kristina A. Tabbada, Henry B. Perdigon, Minerva S. Sagum,  
Miriam Ruth M. Dalet and Maria Corazon A. De Ungria

DNA Analysis Laboratory, Natural Sciences Research Institute,  
University of the Philippines, Diliman, Quezon City, Philippines

Recent developments in forensic DNA technology have led to the expansion of the Philippine reference database of autosomal Short Tandem Repeat (aSTR) DNA markers currently being used by the DNA Analysis Laboratory, Natural Sciences Research Institute of the University of the Philippines for DNA-based paternity testing. A total of 502 unrelated individuals from different regional centers of the Philippines were analyzed at 15 aSTR DNA markers, which include the 13 DNA markers that comprise the Combined DNA Identification System (CODIS) of the US Federal Bureau of Investigations (FBI). Population statistical parameters were calculated and determined. Expansion of the population database resulted in the detection of additional alleles of the markers included in the database, increased power of discrimination and power of paternity exclusion compared to a smaller-sized Philippine database (n=100) that was reported previously (De Ungria et al. 2005, Salvador et al. 2007).

Key Words: human DNA typing, paternity testing, Philippine database, population genetics

### INTRODUCTION

DNA typing is the most powerful tool for paternity determination. DNA-based paternity testing uses highly polymorphic autosomal Short Tandem Repeat (aSTR) DNA markers which facilitates the discrimination of a father from a non-father. In 2007, the Supreme Court promulgated the Rule on DNA Evidence that prescribed the use of an appropriate reference population database in determining the weight of DNA evidence. If an alleged father is not excluded, the weight of matching DNA expressed as a Probability of Paternity ( $W$ ) and a Cumulative Paternity Index (CPI) are calculated using a reference population database (De Ungria et al. 2008). Paternity may be presumed if the Probability of Paternity is 99.9% or higher (Philippine Supreme Court 2007). Hence there is a need to establish a population

database with enough samples genotyped at a sufficient number of DNA markers to meet this demand.

Autosomal STR analysis uses small-sized DNA markers (100-500 bp) which are highly variable in any given population (Butler 2006). The US Federal Bureau of Investigations selected a core set of 13 aSTR markers called the Combined DNA Index System (CODIS) to establish their national DNA database system. Other countries also use these markers to allow comparisons of DNA information across borders. As a consequence, CODIS markers are included in commercially available DNA typing kits an example of which is the PowerPlex<sup>®</sup> 16 (Promega Corporation, Madison, WI).

In the Philippines, a population database composed of the DNA profiles of 100 unrelated individuals was previously published (De Ungria et al. 2005, Salvador et al. 2007) and maintained at the DNA Analysis Laboratory, Natural Sciences Research Institute of the

\*Corresponding author: mariadeungria@gmail.com

University of the Philippines (UP-NSRI-DAL), which will be arbitrarily referred to in this study as 'Database A'. This paper reports the expansion of Database A, to include a total of 502 samples from 12 regional centers in the Philippines and profiled at 13 CODIS and an additional two non-CODIS aSTR DNA markers. The utility of this expanded database (henceforth referred to as 'Database B') in paternity analysis is discussed.

## MATERIALS AND METHODS

### Population samples

Liquid blood samples from volunteers collected by the Philippine National Red Cross (PNRC) were transferred to FTA® blood cards and processed following the manufacturer's instructions (Whatman, BioSciences, USA). Human population sampling procedures followed the guidelines outlined by the Philippine National Bioethics Committee on Scientific Research.

### DNA typing

DNA samples were amplified at 15 aSTR DNA markers namely, D3S1358, HUMTH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, HUMCSF1PO, Penta D, HUMvWA, D8S1179, HUMTPOX and HUMFGA and the sex determining Amelogenin marker, HUMAMEL, using the PowerPlex® 16 multiplex system as recommended by the manufacturer (Promega Corporation, Madison WI). Amplified fragments were analyzed on the ABI Prism® 310 Genetic Analyzer with GeneScan® 3.7 and Genotyper® 3.7 softwares for automatic allele calling (Applied Biosystems, Foster City CA). Alleles and peaks were evaluated based on published guidelines (Bär et al. 1997, Butler 2005).

### Statistical analysis

Exact tests for Hardy-Weinberg Equilibrium (HWE) and Linkage Disequilibrium (LD) were done using the Arlequin software version 3.01 (Excoffier et al. 2005). To assess the level of population subdivision, exact test for population differentiation and Wright's  $F_{st}$  ( $\theta$ ) value were determined using the PowerMarker version 2.05 software (Liu & Muse 2005). Allele frequencies, Matching Probability (MP), Power of Discrimination (PD), Polymorphism Information Content (PIC), and Power of Paternity Exclusion (PE) were calculated using the PowerStats version 1.2 program (Tereba 1999, <http://www.promega.com/geneticidtools/powerstats>).

### Evaluating paternity inclusions using the expanded database

To assess the utility of Database A and Database B, DNA profiles from 30 paternity trios were used. The CPI and  $W$  of complete trios and simulated motherless cases were calculated based on equations provided by Buckleton et al. (2005). A  $\theta$  value of 0.03 or 0.05 was used to correct for population sub-structuring when the alleged father is a Filipino or a foreigner, respectively (US NRC 1996). The significance of the differences in CPI using Database A or Database B was evaluated using T-tailed tests.

## RESULTS

### The newly expanded database includes samples from Luzon, Visayas and Mindanao

Database B consists of individuals from six Luzon regions ( $n=196$ ), two Visayas regions ( $n=123$ ), and four Mindanao regions ( $n=183$ ). The frequencies of the alleles are shown in Table 1. Hardy-Weinberg Equilibrium (HWE) analysis showed that there is no deviation from HWE ( $p>0.05$ ) for each aSTR DNA marker. Linkage disequilibrium analysis showed the absence of linkage ( $p=1.00000$ ) among all 15 aSTR DNA markers tested. Exact tests for population differentiation revealed the presence of significant differences between the allele frequencies of all regional populations within Database B ( $p<0.05$ ). The overall  $F_{st}$  value was calculated to be 0.002.

### Previously unobserved alleles were detected in the expanded Philippine database

Thirty seven new alleles at 13 aSTR DNA markers were detected in Database B. Additional alleles ranged from one new allele at D3S1358 to 11 new alleles at HUMFGA (Table 2). Table 3 shows the forensic and paternity statistical parameters of Database A and Database B.

### Application to paternity case analysis

In paternity trios, statistically significant difference was observed between the CPI values calculated using Database A and Database B ( $p=0.02$  and  $p=0.045$ , for a one-tailed T-test and two-tailed T-test, respectively). Using Database B, 10% of the simulated motherless cases have  $W$  less than 99.9% (Table 4). In general, the value of CPI decreased when the expanded database was used.

## DISCUSSION

The results presented in this paper indicate that the expanded Philippine reference database is suitable for DNA-based paternity testing with all the aSTR DNA

**Table 1.** Allele frequencies and forensic statistics of the expanded Philippine population database (n=502) at 15 STR markers.

Allele	Autosomal Short Tandem Repeat (aSTR) Markers														
	D3S1358	TH01	D21S11	D18S51	Penta E	D5S818	D13S317	D7S820	D16S539	CSF1PO	Penta D	vWA	D8S1179	TPOX	FGA
3															
3.2															
4															
5		0.0020			0.0388						0.0010				
6		0.1125												0.0010	
7		0.2938			0.0020	0.0149		0.0020	0.0010	0.0020	0.0060		0.0010		
8		0.0906			0.0030		0.2122	0.2261	0.0020	0.0090	0.0916		0.0010	0.4871	
8.3								0.0010							
9		0.4004		0.0010	0.0498	0.0508	0.1683	0.0369	0.2709	0.0309	0.3715		0.0020	0.1375	
9.3		0.0398													
10		0.0608		0.0010	0.0448	0.2789	0.1016	0.2022	0.1683	0.2839	0.1703		0.0807	0.0139	
11				0.0110	0.1544	0.2779	0.2968	0.3616	0.2799	0.2659	0.0807		0.1096	0.3317	
12				0.0558	0.0777	0.2430	0.1803	0.1434	0.1444	0.3466	0.1713	0.0010	0.0916	0.0279	
13	0.0020			0.0807	0.0339	0.1275	0.0299	0.0219	0.1125	0.0538	0.0916	0.0010	0.2560	0.0010	
14	0.0349			0.2112	0.0777	0.0050	0.0110	0.0050	0.0199	0.0080	0.0129	0.1633	0.1803		
15	0.2520			0.2112	0.1106	0.0020			0.0010		0.0030	0.0827	0.1444		
16	0.3108			0.1474	0.0608							0.1594	0.1066		0.0030
17	0.3287			0.1145	0.0867							0.2251	0.0259		0.0020
18	0.0707			0.0588	0.0847							0.2769	0.0010		0.0100
19	0.0010			0.0538	0.0767							0.0797			0.0837
20				0.0229	0.0588							0.0100			0.0757
20.2															0.0030
21				0.0110	0.0189										0.1285
21.2															0.0020
22				0.0110	0.0149										0.2102
22.2															0.0090
23				0.0040	0.0040							0.0010			0.2072
23.2															0.0020
24				0.0050	0.0010										0.1305
24.2															0.0010
25															0.0697
25.2															0.0010
26				0.0010											0.0418
26.2															0.0020
27				0.0110	0.0010										0.0110
28				0.0468											0.0010
28.1				0.0010											
29				0.2082											
30				0.2351											
30.2				0.0100											
31				0.1952											
31.2				0.0697											
32				0.0339											
32.2				0.1325											
33				0.0030											
33.2				0.0408											
34				0.0020											

table 1 continued next page . . .



**Table 3.** Forensic and Paternity statistical parameters of two Philippine population databases using 15 aSTR DNA markers.

Sample size (n)	Philippine Database A n=100 (De Ungria et al. 2005, Salvador et al. 2007)	Philippine Database B n=502 (this study)
Combined Matching Probability	1 in 3.54 x 10 <sup>16</sup>	1 in 2.05 x 10 <sup>17</sup>
Combined Power of Discrimination	0.999999999999999717	0.999999999999999951
Combined Power of Paternity Exclusion	0.999998421	0.999999424

**Table 4.** Probability of paternity for 30 paternity trios (*W*%) and motherless cases (*W*<sub>-mother</sub>%) using Philippine Database B.

Mother-father-child trios		Father-child pairs (motherless cases)	
<i>W</i> (%)	% Families	<i>W</i> <sub>-mother</sub> (%)	% Families
99.8-99.9	0	99.8-99.9	10
99.9-99.99	0	99.9-99.99	27
>99.99	100	>99.99	63

In conclusion, a more informative Philippine aSTR reference database at 15 aSTR DNA markers that provides a better representation of the Philippine population has been established and may be used for parentage testing in the country.

## ACKNOWLEDGEMENTS

The authors acknowledge the support from the Natural Sciences Research Institute of the University of the Philippines (UP-NSRI) and the Department of Science and Technology - Philippine Council for Advanced Science and Technology Research and Development (DOST-PCASTRD). Samples were obtained from the Philippine National Red Cross (PNRC).

## REFERENCES

BÄR W, BRINKMANN B, BUDOWLE B, CARRACEDO A, GILL P, LINCOLN P, MAYR WR and OLAISEN B. 1997. DNA Recommendations: Further report of the DNA Commission of the ISFH regarding the use of short tandem repeat systems. *Int J Leg Med* 110: 175-76.

BUCKLETON J, TRIGGS C and WALSH S. 2005. *Forensic DNA Evidence Interpretation*. CRC Press USA. 534 p.

BUTLER JM. 2005. *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*, 2nd ed. Elsevier Academic Press USA. 660 p.

BUTLER JM. 2006. Genetics and genomics of core short tandem repeat loci used in human identity testing. *J Forensic Sci* 51(2):253-265.

DE UNGRIA MCA, CALACAL GC, DELFIN FC, TABBADA KA, SALVADOR JM, SAGUM MS, DALET MRM, FRANCO SAEL, MAIQUILLA SMB, PERDIGON HB, VILLAMOR LP, and HALOS SC. 2008. Overview of Forensic Services by the UP-NSRI DNA Laboratory: a 4-year Report. *Acta Medica Philippina* 42(2):34-38.

DE UNGRIA MCA, FRANI AM, MAGNO MMF, TABBADA KA, CALACAL GC, and HALOS SC. 2002. Evaluating DNA tests of motherless cases using a Philippine genetic database. *Transfusion* 42:954-957.

DE UNGRIA MCA, ROBY R, TABBADA KA, RAO-COTICONE S, TAN MM and HERNANDEZ KN. 2005. Allele frequencies of 19 STR loci in a Philippine population generated using AmpFISTR multiplex and ALF singleplex systems. *Forensic Sci Int* 152:281-284.

EXCOFFIER L, LAVAL G and SCHNEIDER S. 2005. Arlequin ver. 3.01: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1:47-50.

LIU K and MUSE SV. 2005. PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21(9):2128-2129.

PHILIPPINE SUPREME COURT. 2007. Rule on DNA Evidence. A.M. No. 06-11-5-SC. Manila, Philippines. 19p.

SALVADOR JM, CALACAL GC, VILLAMOR LP and DE UNGRIA MCA. 2007. Allele frequencies for two pentanucleotide STR loci Penta D and Penta E in a Philippine population. *Legal Medicine* 9:282-283.

SCHNEIDER PM. 2007. Scientific standards for studies in forensic genetics. *Forensic Sci Int* 165:238-243.

SHEPPARD C. 1999. How large should my sample be? Some quick guides to sample size and the power of tests. *Marine Pollution Bulletin* 38(6):439-447.

TABBADAKA, MAGNO MF, DELFIN FC, CALACAL GC, TAN MP, FERREON JC, HALOS SC and DE UNGRIA MCA. 2002. Allele frequencies of eight short tandem repeat loci in three Visayan regional populations of the Philippines. *Journal of Forensic Sciences* 47(6):1397-1398.

TEREBA A. 1999. Tools for analysis of population statistics. *Profiles in DNA* 2(3):14-16.

[US NRC] UNITED STATES NATIONAL RESEARCH COUNCIL. 1996. *The Evaluation of Forensic Evidence II*. Washington, DC: National Academy Press.

<http://www.cstl.nist.gov/div831/strbase/> on 16 March 2009.

<http://www.promega.com/geneticidtools/powerstats> on 20 May 2009.