Forensic Science International: Genetics 4 (2010) e145–e148

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Forensic Population Genetics - Short Communication

# Analysis of 14 highly informative SNP markers on X chromosome by TaqMan® SNP genotyping assay

Li Li [a,*], Chengtao Li [a], Suhua Zhang [a,b], Shumin Zhao [a], Yan Liu [a], Yuan Lin [a]

[a] Shanghai Key Laboratory of Forensic Medicine, Institute of Forensic Science, Ministry of Justice, PR China, Shanghai 200063, China
[b] Medical School of Soochow University, Suzhou 215123, PR China

ARTICLE INFO

*Article history:*
Received 27 August 2009
Received in revised form 26 January 2010
Accepted 17 April 2010

*Keywords:*
Forensic genetics
Genotyping
SNP
TaqMan® SNP genotyping assay
X chromosome
Population data

ABSTRACT

*Objective:* To genotype X-linked single nucleotide polymorphisms by TaqMan® SNP genotyping assay.
*Methods:* 14 SNP markers on X chromosome were selected and genotyped by TaqMan® SNP genotyping assay in Chinese Han population samples.
*Results:* According to the results of population studies, deviations from Hardy–Weinberg equilibrium could be found at rs1299087 SNP loci, no deviations from Hardy–Weinberg equilibrium could be found at the other 13 SNP loci. The 14 X-SNPs were high informative. The overall power of discrimination (CPD) in females and in males were 0.999998379 and 0.999899, respectively. The combined exclusion probability in paternity testing for trios and duos were 0.9983 and 0.9788, respectively.
*Conclusion:* SNP typing by TaqMan® SNP genotyping assay is suitable for low-throughput application. The forensic efficiency parameters showed that the 14 X-linked SNPs are highly informative and seem to be useful for forensic genetics except rs1299087.



## 1. Introduction

Single nucleotide polymorphisms (SNPs) are sites in the genome that have at least two different bases at the same location. These polymorphisms are highly abundant, on average, at every few hundred bases across the human genome [1]. In forensic and medical practice, because all daughters of the same biological father share at least one identical allele at every X-linked locus, SNPs on X chromosome may be more important than autosomal markers in deficiency cases when the alleged father is absent and there are two women who were separated as children have the same father. Additionally, identification some close blood relatives, for example, paternal grandmother and granddaughter, can be carried out only through the application of X-linked markers because one allele of the grandmothers was transmitted exactly to the granddaughters via the fathers in all the X-SNP markers. The availability of a theoretically unlimited number of SNPs should facilitate the selection of genetic markers [2]. Being biallelic markers, SNPs can effectively complement STR typing, since SNPs are robust in terms of data interpretation, stability of inheritance and population genetic analysis [3,4].

In the present SNP genotyping protocol, 14 SNPs on X chromosome were amplified and genotyped by TaqMan® SNP genotyping assay in Chinese Han population samples. Hardy–Weinberg equilibrium, linkage disequilibrium, allele frequencies and forensic efficiencies parameters were tested and estimated.

## 2. Materials and methods

### 2.1. Sample

238 blood samples were collected from unrelated Chinese Han individuals (113 females, 125 males). 41 family trios consisting of father–daughter–mother, 54 family duos consisting of 30 father–daughter pairs and 24 mother–child pairs were involved in this study after obtaining consent. Fathers and mothers of the duos cases were included in the 238 unrelated samples.

### 2.2. TaqMan assay design

14 SNP markers on X chromosome were selected from the SNP information (http://www.hapmap.org) supplied by NCBI (National Center for Biotechnology Information). The locus-to-locus space is about 3–20 Mb (Table 1). Locus-specific PCR primers and allele-specific TaqMan® probes were designed and supplied by Applied Biosystems (Foster City, CA, USA). TaqMan® probe specific to allele 1 had FAM as fluorescent reporter dye at the 5′ end, probe specific to allele 2 had VIC as fluorescent reporter dye at the 5′ end, and each of them had a non-fluorescent quencher (NFQ) with a minor groove binder (MGB) at the 3′ end. DNA was extracted by the

* Corresponding author. Tel.: +86 021 52352959; fax: +86 02153352959.
*E-mail address:* anneinchina@163.com (L. Li).

1872-4973/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigen.2010.04.004

**Table 1**
14 SNP loci for TaqMan assay.

| Assay ID | Mapped ID/dbSNP | Public location | SNP type | Locus-to-locus space (Mb) |
|---|---|---|---|---|
| C__30259984_10 | rs6639398 | chr. X 5151299 | A/G | |
| C__29152925_10 | rs5986751 | chr. X 25127167 | C/T | 19.98 |
| C__11778614_10 | rs6631828 | chr. X 33633370 | C/T | 8.51 |
| C__2462671_10 | rs5964206 | chr. X 42483666 | C/G | 8.85 |
| C__29642596_10 | rs9781645 | chr. X 53808325 | C/T | 11.32 |
| C__26782698_20 | rs2209420 | chr. X 68464886 | A/C | 14.66 |
| C__8897161_10 | rs1299087 | chr. X 82478126 | A/G | 14.01 |
| C__30037046_10 | rs5923750 | chr. X 86188996 | A/G | 3.71 |
| C__2456381_10 | rs2808742 | chr. X 95333052 | A/G | 9.14 |
| C__2553000_20 | rs6418251 | chr. X 106520819 | A/C | 11.19 |
| C__11224657_10 | rs2519557 | chr. X 116205417 | G/T | 9.68 |
| C__26801236_10 | rs149860 | chr. X 126221873 | C/G | 10.02 |
| C__15812121_20 | rs2085121 | chr. X 137006323 | A/G | 10.78 |
| C__2520158_20 | rs5954988 | chr. X 142169089 | C/T | 5.16 |

Chelex extraction method [5]. PCR was performed on 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA) in 25 μL reaction volume. For each PCR, 11.25 μL genome DNA (1–20 ng) was mixed with 12.5 μL 2× TaqMan universal PCR master mix, 1.25 μL 20× TaqMan SNP genotyping assay mix. All the reactions consisted of one cycle at 95 °C for 10 min, 40 cycles at 92 °C for 15 s, 60 °C for 1 min.

The cell line sample 9947A (Applied Biosystems, Foster City, CA) was used as control DNA for TaqMan assay. The results were confirmed by sequencing. The primers for sequencing were shown in Table 2.

### 2.3. Statistical analysis

Basic statistical computations including allele frequency, observed heterozygosity (H), average power of paternity exclusion in father–daughter duos lacking maternal genotype information ($PE_{duos}$) and in trios involving daughters ($PE_{trios}$), discrimination power in males ($PD_{male}$) and in females ($PD_{female}$), polymorphism information contents (PIC) were determined using the appropriate formulas [6–9]. Mutation frequency was calculated by the number of mutation event per locus per meiosis. Hardy–Weinberg equilibrium (HWE) tests and multi-locus linkage disequilibrium tests [10,11] were performed using the Powermarker v3.25 program [12]. The *p*-values were corrected for the multiple tests.

## 3. Results

### 3.1. Validation

Validating carried out on DNA extracts prepared from blood samples with known SNP genotypes from TaqMan® SNP genotyping assay at the 14 loci by sequencing revealed fully consistent and reproducible results. To enable the further comparison of population, control DNA 9947A with a diploid state was genotyped (Table 2). The loci were also genotyped in 54 family duos and 41 trios. The SNP genotypes were fully consistent with mendelian segregation. No mutation was found in the 14 SNP markers.

### 3.2. Allelic distributions and tests of Hardy–Weinberg and genotypic disequilibriums

A total of 238 related bloodstain samples were genotyped at the 14 SNP loci among Chinese Han population. The allele frequencies respectively obtained from females and males were shown in Table 3. The result of HWE test was shown in Table 4. The exact test of 13 SNPs showed no significant deviation from Hardy–Weinberg equilibrium ($p > 0.05$), but rs1299087 was not in Hardy–Weinberg equilibrium ($p < 0.001$). Allele frequencies between female and male were not significantly different in all SNP markers with Hardy–Weinberg equilibrium.

**Table 2**
SNP genotype of control DNA 9947A by TaqMan assay and sequencing.

| SNP | Genotype |
|---|---|
| rs6639398 | GG |
| rs5986751 | TT |
| rs6631828 | TT |
| rs5964206 | CC |
| rs9781645 | CC |
| rs2209420 | AC |
| rs1299087 | GG |
| rs5923750 | AA |
| rs2808742 | AG |
| rs6418251 | CC |
| rs2519557 | TT |
| rs149860 | CG |
| rs2085121 | GG |
| rs5954988 | CC |

### 3.3. Linkage disequilibrium and haplotype frequencies

Multi-locus linkage disequilibrium was analyzed by SNP genotyping 113 female DNA samples. The result is shown in Table 5. LD was observed for three markers, namely rs5964206, rs9781645 and rs2209420 ($p < 0.05$) although the locus-to-locus space >10 Mb. So, we propose to use haplotype frequencies instead of frequencies of single SNP in calculation of likelihood. The haplotype frequencies are shown in Table 6.

**Table 3**
Allele frequencies of 14 X-SNPs in unrelated females ($n = 113$) and males ($n = 125$).

| SNP | Females | | Males | |
|---|---|---|---|---|
| | Allele 1 | Allele 2 | Allele 1 | Allele 2 |
| rs6639398 | A:0.4912 | G:0.5088 | A:0.4160 | G:0.5840 |
| rs5986751 | C:0.4602 | T:0.5398 | C:0.4720 | T:0.5280 |
| rs6631828 | C:0.5487 | T:0.4513 | C:0.5440 | T:0.4560 |
| rs5964206 | C:0.4336 | G:0.5664 | C:0.3920 | G:0.6080 |
| rs9781645 | C:0.5265 | T:0.4735 | C:0.4720 | T:0.5280 |
| rs2209420 | A:0.5310 | C:0.4690 | A:0.5360 | C:0.4640 |
| rs1299087 | A:0.1637 | G:0.8363 | A:0.4720 | G:0.5280 |
| rs5923750 | A:0.6549 | G:0.3451 | A:0.6640 | G:0.3360 |
| rs2808742 | A:0.5354 | G:0.4646 | A:0.5280 | G:0.4720 |
| rs6418251 | A:0.6504 | C:0.3496 | A:0.6080 | C:0.3920 |
| rs2519557 | G:0.3451 | T:0.6549 | G:0.3680 | T:0.6320 |
| rs149860 | C:0.4425 | G:0.5575 | C:0.4320 | G:0.5680 |
| rs2085121 | A:0.5575 | G:0.4425 | A:0.5280 | G:0.4720 |
| rs5954988 | C:0.5354 | T:0.4646 | C:0.5200 | T:0.4800 |