Allelic Sharing Among Madurese as a Tool of Madurese Identification Using 11 Short Tandem Repeats and Amelogenin Gene: An Observational Analytical Study

Ahmad Yudianto1,2, Agung Sosiawan1,3, Abdul Hadi Furqoni1, Indah Nuraini Masjkur1, Qurrota A’yunil Huda1
1Human Genetic and Forensic Laboratory, Institute of Tropical Disease, 2Department of Forensic Medicine, Faculty of Medicine, 3Department of Dental Public Health, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

Abstract

Aim: To analyze the allelic sharing and a set of preferable short tandem repeats (STRs) loci among the Madurese full siblings.

Materials and Methods: This was an observational analytical study. The research subjects were determined from 20 Madurese families that originated from Madura Island, Indonesia consisting of father, mother, and two biological children. The blood samples were collected, and kinship relationship was examined by using the PCR-STR technique by 11 loci of genotyped STR alleles and a sex-typing amelogenin locus. Calculations of combined sibship indices (CSIs) for each pair were done to determine the likelihood ratios.

Results: Based on 220 observations (20 pairs of full sibling x 11 loci) of the STR allele genotype, 4 STR loci were found: D8S1179, F13, CSF1PO, and FES. A high-shared allelic frequency was found in two alleles that were shared among 11 STR loci by 57 times (25.9%), one allele shared among eleven STR loci by 130 times (59.1%), and zero alleles by 33 times (15.0%).

Conclusion: This study indicated the power of one allele in the sibship among a Madurese establishment through the 11 STR loci.

Keywords: Allele, Full Sibling, Madurese, STR Loci

Received: 07-05-2020, Revised: 10-08-2020, Accepted: 12-08-2020, Published: 28-01-2021.
There are 28 STR systems, namely D3S1358, TH01, D21S11, D18S51, D31S138, D35S810, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX, FGA, D19S43, D21S138, D21S1045, D1S1656, D10S1248, D2S441, D12S391, SE33, LPL, F13B, FESFPS, F13A01, and Penta C and 20,000 STR loci, that can be used and amplified by PCR. The STR loci are chosen based on the local dedicated project, because it can overlap with the other STR loci and the worldwide shared locus inclusivity. The variability of identity repeated in the unit of length in both allele and STR matching among people offers the frequency of likelihood disparity among the population in different territories. The characteristics of one population related to the other population suggest a localized population that is based on the STR derivative, even though it can be marginalized in the number of overlapping nations.[6,9]

The application of the amelogenin gene in forensic case works is using two types, such as amelogenin X-linked (AMELX) and amelogenin Y-linked (AMELY). The location of AMELX is on the distal part of the short arm of chromosome X p22.1–p22.3 region, and AMELY is near the centromer of the Y chromosome p11.2. AMELX and AMELY products can be distinguished by PCR that uses a primer flanking a 6 bp deletion in the third intron of AMELX. In AMELY, the deletion is not present.[10] The study aimed at establishing the allelic sharing and a set of preferable STR loci among the Madurese full sibling.

**Materials and Methods**

**Setting and design**

This observational analytical study was conducted from September to December 2019 at Human Genetic Laboratory, Institute of Tropical Disease, Universitas Airlangga, Indonesia. The sample size was 20 full-sibling individuals. Inclusion criteria consisted of full siblings of Madurese origin from Madura Island, Indonesia, which included alleged father, mother, and their biological two children who were nonidentical twins and agreed to participate in the study. In each of the cases, neither the alleged mother nor the alleged father was included as biological parents of either child, nor were mutational events detected. The parental alleles were ignored after case selection, as if the parents were unavailable.

**Study method**

The blood from the peripheral blood system was extracted for DNA extraction. DNA extraction in this study used DNAzol Organic (Invitrogen, USA) method protocol.[11,12] After extracting the DNA concentration, purity was measured by using an ultraviolet spectrophotometer (Shimidzu UV-16, Kyoto, Japan).[11,12] Amplification of DNA was carried out by PCR-STR (PowerPlex® 21 Systems, Promega, USA) using 11 STRs combined DNA index (CODIS) systems with the Federal Bureau of Investigation standard and the amelogenin gene for determination of gender (GoTaq® Green Master Mix, Promega Corp., Madison, USA). The locusization of PCR-STR products was electrophoresed by using 6% poliacrylamide gel (Life Science Group, UK). The results of the visualization were then analyzed to determine the full-sibling pattern in the kinship analysis.[13]

**STR loci analysis**

The known siblings by the logic and calculations for the full-sibling index/sibship index (SI) were calculated based on the kinship analysis equations.[13] The allele probabilities were taken from allele frequencies among the Indonesian population. The descriptions of shared alleles were presented together with the likelihood ratios favoring sibship. Finally, the calculation of CSIs for each pair was done to determine the likelihood ratios. The CSIs were determined by multiplying the individual SI values of the chosen loci.[14]

**Results**

The allele-sharing degree in the present and previous published study is presented in Table 1. Among the 220 observations (20 pairs × 11 loci) in true siblings, two alleles per locus were shared 57 times (25.9%), one allele was shared 130 times (59.1%), and zero alleles were shared 33 times (15.0%).

The percentage of evidence strength found in this study is presented in Table 2. Overall, 100% of 20 true sibling pairs had likelihood ratios favoring sibship (i.e., their CSI values were greater than 1). The CSI ranges are from 1.4 to higher than 37 billion in the 20 known full-sibling pairs. All of the notable sibling pairs had CSIs greater than 10, and the majority of them (65%) had CSIs higher than 1000.

**Discussion**

The areas that contain repeated sequences of nucleotides such as STR sequences are of interest for forensic sciences, because they have a lot of information about the variation that makes them useful for human identification purposes.

The good indicators that are based on the allelic frequencies are homozygosity, heterozygosity, an

<table>
<thead>
<tr>
<th>Table 1: The probability of allele-sharing degree in sibling pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study</strong></td>
</tr>
<tr>
<td>Current study</td>
</tr>
<tr>
<td>Wenk, 1996</td>
</tr>
</tbody>
</table>
Table 2: The percentage of evidence strength found in this study

<table>
<thead>
<tr>
<th>CSIs</th>
<th>Evidence strength</th>
<th>Percentage in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–33</td>
<td>Weak</td>
<td>15</td>
</tr>
<tr>
<td>33–100</td>
<td>Fair</td>
<td>10</td>
</tr>
<tr>
<td>100–330</td>
<td>Good</td>
<td>0</td>
</tr>
<tr>
<td>330–1000</td>
<td>Strong</td>
<td>10</td>
</tr>
<tr>
<td>&gt;1000</td>
<td>Very strong</td>
<td>65</td>
</tr>
</tbody>
</table>

effective number of alleles, the content of polymorphism information, the discrimination power, and the exclusion power. A good indicator of genetic polymorphisms is based on the number of alleles. This can be caused by the presence of unique alleles in populations. Therefore, genetic screening must be done before conducting the research. It is theorized that a bigger population size will increase the number of observed alleles.[13]

Full siblings are intended to share zero or two identical alleles by descent at a certain locus with an equal probability of 0.5 and also share a single allele with a probability of 0.25.[16] In the current study, full siblings shared 15%, 60%, and 26% zero, single, and two identical alleles, respectively. The allele-sharing degree in the current study was comparable to that of other studies.[18] This variation was probably caused by the ethnicity that has contributed to different genetics between the populations. It can be linked to the historical demographic processes, which lead to genetic drift.

The fact that 75% and 65% of the sibling pairs had CSIs higher than 100 and 1000 (strong and very strong evidences), respectively, indicates that the current STR system is very promising to identify siblings of Madurese. This was strengthened by the United States of Citizenship and Immigration Services (USCIS) report, where it was found that only 9% of samples were excluded from the full-sibling relationships in tests using 20 loci. The report also shows that sibling kinship analysis is one of the personal identifications that needs to be considered, because until now there have been no standard locus for the STR examination for the purposes of full-sibling kinship analysis. The standard locus that can be used for the STR examination must have a 90% or more probability level.[17]

One thing that needs to be consider in the analysis of sibling kinship is the possibility of common allele in certain focus areas of STRs. The locus of STRs that have a tendency to have these common alleles should be aware before they are withdrawn as a conclusion in personal identification.[18] Furthermore, the family originating from marriage consanguinity increases the complexity of full sibling analysis by the increased of co-ancestry. Therefore caution is needed in the process of interpreting the results of the full-sibling STR examination.[19]

The results of this study can be either high or low: If the results are low, alternative choices can be got by two options. The first option is that the research must use another locus that is still related to the chosen locus. The second option is that the research can use additional DNA testing from the respective uncles, aunts, and cousins.[20] This is because shared alleles cannot be found in full-sibling kinship analysis.[6] More importantly, the use of 20 STRs examination locations in personal identification, based on full-sibling kinship analysis, becomes an unavoidable requirement to reach the 90% probability standard as determined by the USCIS.[19] It aims at increasing the power of discrimination for forensic DNA examination.[21]

This study is important as a basis for the personal identification process based on sibling kinship analysis using 11 STRs and the amelogenin gene, when the analysis of the parental inheritance of children is not possible. This study also has implications for oral health. This study can be used as a basis for whether a person has a tendency to have genetic disorders, especially in the oral cavity when she/he has the same or a close percentage of a locus pattern with STR determination. This study could be a predictor of similarity in siblings against certain genetic diseases, especially in the oral cavity such as the cleft palate. Microsatellites are markers that are commonly used in genetic analysis, as in the case of the cleft palate. This is related to the nature of these markers that have high mutation rates or, in other words, have a high polymorphism. The STR is one type of microsatellite marker that has a high polymorphism, so it is used for genetic forensic analysis. Indeed, there has been no genetic analysis in dentistry such as cleft palate using STR, but that does not mean it is not possible to be used in genetic analysis for cleft palate. So far, the cleft palate used five microsatellite markers such as DLX3, MSX1, RARA, BCL3, and EDN1.[22,23]

**Conclusion**

From the overview of the guidance for the sibling DNA test, it can be concluded that the one allele is the strongest one compared with the two and zero alleles. The mark of one allele is 59.1% and this indicates that full-sibling interpretation supports the fact that the claimed relationship exists. The difficulties experienced in the low number of DNA samples that have been obtained can be overcome by using another locus that is related to the chosen locus or additional DNA that is obtained from uncles, aunts, and cousins. In future, genetic analysis using STRs will be very useful in the dental field, especially for cleft palate study analysis.

**Acknowledgment**

We thank Rector Universitas Airlangga for financial support of this study.

**Financial support and sponsorship**

This research was funded by the Riset Mandat Grant (No 622/UN3.14/LT/2018).
Conflicts of interest
There are no conflicts of interest.

Author contributions
Ahmad Yudianto: study design, data collection and article writing; Agung Sosiawan: study design, data acquisition and analysis, and article writing; Abdul Hadi Furqoni: data collection, data interpretation; Indah Nuraini Masjukur: data collection, statistical analysis; and Qurrota Ayunil Huda: data collection. All authors approved the final version of the article for publication.

Ethical policy and institutional review board statement
This study was carried out under ethical committee clearance approval issued by Health Research Ethical Committee Faculty of Dental Medicine, Universitas Airlangga, Indonesia (524/HRECC.FODM/VII/2019).

Patient declaration of consent
Not applicable.

Data availability statement
Not applicable.

REFERENCES
21. Hares DR. Selection and implementation of expanded CODIS core loci in various Asian populations in or near Japan. Inter Con Series 2003;1239:117-20.