

Genetic Variation using the 21 Str Codis Loci for Forensic Identification Examinations among Siblings of Madurese Living in Surabaya

Ahmad Yudianto^{1,2,4}, Arofi Kurniawan^{2,3}, Beta Novia Rizky^{2,3}, M. Kholil Ikhsan¹, Shella Morina¹, William Daniel Napitupulu^{1*}, Fery Setiawan², Indah Nuraini Masjkur⁴, Qurrotun A'yunil Huda⁴

1. Department of Forensic Medicine and Medicolegal, Faculty of Medicine, Universitas Airlangga, Surabaya, 60115 Indonesia.
2. Magister of Forensic Sciences, Postgraduate School, Universitas Airlangga, Surabaya, 60115 Indonesia.
3. Department of Forensic Odontology, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, 60115, Indonesia.
4. Human Genetic Study Group, Institut of Tropical Disease, Universitas Airlangga, Surabaya, 60115, Indonesia.

Abstract

Indonesia consists of various ethnic groups. However, well-established research on the genetic variation of each ethnic group in Indonesia has not yet been present. Madurese is one of the Indonesian ethnic groups inhabiting Madura island and living in remote areas throughout Indonesia. Surabaya- Madura's mobilization is tremendous as the closest city to Madura Island, connected by a bridge. The Madurese has a culture of didi and same-ethnic marriage patterns. Therefore, research was conducted on 21 STR panel loci to analyze Madurese living in Surabaya to increase genetic data libraries and genetic variation of Indonesian populations based on siblings.

PrepFiler kit to extract DNA from blood and GlobalFiler kit to amplify 21 loci were used. An ABI PRISM 3500 genetic analyzer was used to detect PCR products. Data were processed and analyzed with EasyDNA software and FORSTAT.

Allele frequency data for 21 autosomal loci in Madurese were obtained. The expected Heterozygosity (He) was found with the highest obtained at locus D19S433 (0.911), while the lowest was at locus D22S1045 (0.607). Power of Discrimination (PD) values ranged from the highest at 2 loci, namely D5S818, and D13S317 (0.968), while the lowest at locus D2S1338 (0.857). Polymorphism Information Content (PIC) value is highest at locus D22S1045 (0.765); all loci have PIC values > 0.5. The Sibship Index (SI) value of the Madurese in Surabaya is 0.88 – 1.20.

The 21 STR loci can be employed for population genetic variation studies and sibling forensic identification.

Experimental article (J Int Dent Med Res 2023; 16(4): 1616-1619)

Keywords: Genetic variation, siblings, Madurese, STR locus.

Received date: 20 October 2023

Accept date: 19 December 2023

Introduction

Using Deoxyribo Nucleic Acid (DNA) as a forensic molecular examination material is neither without constraints nor obstacles. Nevertheless, factors are causing forensic DNA experts to experience adversities, both in working with DNA samples and inferring the examination results. The lack of information from fathers and mothers or children to be used as a comparison in forensic DNA analysis is one of the problems

in forensic DNA analysis¹. The comparison between the alleles of the victim or perpetrator and their relatives is the basic principle of DNA examination, especially from parental lines per Mendelian law, such as in cases of "unborn child disputed," disputed paternity, or even in forensic DNA analysis in mass disasters or mass disasters^{2,3}. In these conditions, a comparator with a close family line is required as one of the ways to be taken in the forensic DNA analysis process, such as siblings, if a comparator from the parental or children line is not obtained. The identification process using siblings as a comparator will face the possibility of a mismatch or mismatch in the DNA locus profile used⁴. The problem with using sibling DNA is that not all forensic DNA examination loci in suspects are identical to their comparators as if using parent or

*Corresponding author:

William Daniel Napitupulu, MD
Department of Forensic Medicine and Medicolegal,
Faculty of Medicine, Universitas Airlangga,
Surabaya, 60115 Indonesia
E-mail: wildanapit@gmail.com

child comparators. This condition certainly creates its own difficulties in making conclusions about the examination when the identification process is carried out.

The Madurese can be found living in various parts of Indonesia. The nomadic spirit and economic pressure have caused them to migrate throughout the country. The migration is due to the desire to improve their socio-economic life and education. Surabaya is one of the places or cities that became a place for their permanent or semi-permanent migration due to its closest and more accessible transportation. This study aimed to determine the genetic variation and allele decline between siblings of the same parents among Madurese living in Surabaya.

Materials and methods

Observational laboratory research and the research design used were instantaneous. The research sample was the DNA buccal swab of a volunteer family of parents and children. The study has obtained ethical eligibility from the Faculty of Dentistry, Universitas Airlangga, number: 568/HRECC.FODM/V/2023. The research was conducted at the human genetic study group of the Institute of Tropical Disease, Universitas Airlangga. It was conducted during May – June 2023. The number of volunteers was 10 families [father, mother, first child, and second child].

| Loci | Chromosome location | Chromosome | Repeat Motive | Allele size (bp) |
|----------|--|------------|---------------|------------------|
| D1S1656 | 1q42 | 1 | TAGA | 159-207 |
| TPOX | 2p23.2 <i>thyroid peroxidase</i> intron ke 10 | 2 | AATG | 393-441 |
| D2S441 | 2p14 | 2 | TCTA/TCAA | 76-140 |
| D2S1338 | 2q35 | 2 | TGCC/TTCC | 281-350 |
| D3S1358 | 3p21.31 | 3 | TCTA/TCTG | 96-142 |
| FGA | 4q31.3 <i>alpha fibrinogen</i> intron ke 3 | 4 | CTTT/TTCC | 223-379 |
| D5S818 | 5q21-31 | 5 | AGAT | 138-183 |
| CSF1PO | 5q33.3-34 <i>c-fms protooncogene</i> , intron ke 6 | 5 | AGAT | 283-320 |
| SE33 | 6q14 <i>beta-actin related pseudogene</i> | 6 | AAAG | 307-439 |
| D7S820 | 7q11.21-22 | 7 | GATA | 262-299 |
| D8S1179 | 8q24.13 | 8 | TCTA/TCTG | 114-171 |
| D10S1248 | 10q26.3 | 10 | GGAA | 85-130 |
| TH01 | 11p15.5 <i>tyrosine hydroxylase</i> , intron ke 2 | 11 | TCAT | 179-219 |
| vWA | 12p13.31 <i>von Willebrand factor</i> , intron ke 40 | 12 | TCTA/TCTG | 156-209 |
| D12S391 | 12p13.2 | 12 | AGAT/AGAC | 216-269 |
| D13S317 | 13q22-31 | 13 | TATC | 198-244 |
| D16S539 | 16q24.1 | 16 | GATA | 227-268 |
| D18S51 | 18q21.33 | 18 | AGAA | 261-343 |
| D19S433 | 19q12 | 19 | AAGG/TAGG | 118-171 |
| D21S11 | 21q21.1 | 21 | TCTA/TCTG | 183-240 |
| D22S1045 | 22q12.3 | 22 | ATT | 88-122 |

Table 1. 21 STR Loci^{5,6}

Buccal swabs from volunteers were isolated via DNAzol kit and amplified using GeneAmp PCR System 9700. All 21 Short Tandem repeat (STR) loci were amplified simultaneously. 21 STR loci were used in STR genotyping (table 1), performed on ABI PRISM 3500 Genetic Analyzer and analyzed by EasyDNA and FORSTAT software programs.

Data Analysis

The final result is an electropherogram containing graphs from the Capillary Electrophoresis Genetic Analyzer 3500 ABI tool, which will show 21 individual loci, each of which has 2 allele fragments (heterozygous) or 1 allele fragment (homozygous). The graph will show the size of the DNA fragments in base pair units (bp) and the value of RFUs (Relative Fluorescence Units). Genetic variation using DNA siblings in Madurese living in Surabaya through analyzing alleles from 21 STR loci: allele frequency, the Homozygosity, the Expected Heterozygosity, the Power of Exclusion (PE), the Power of Discrimination (PD), the Paternity Index, the Polymorphism Information Content (PIC), the Allele Sharing and the Sibship Index (SI) using the EasyDNA and FORSTAT software.

Results

In this study, the total allele variance of 21 STR loci in Madurese is 220. The allele frequency of each locus is listed in Table 2, with the allele frequency range: 0.0021 – 0.6042. The expected heterozygosity (**He**), the Power of Discrimination (**PD**), and the Polymorphic Information Content (**PIC**) of 21 STR loci can be seen in Table 3.

The distribution of sibling's allele sharing from all samples at all loci is mainly half sharing, or there is one allele in common from the two siblings compared with 58%. At the same time, the lowest percentage is sub-null sharing, or there is no common allele in a pair of alleles from the two siblings, compared to 8% (Figure 1). Meanwhile, the Sibship Index (SI) results are in Table 4. The SI value is 0.88 – 1.20. (Figure 1)

| Allele | D1S1656 | TPOX | D2S441 | D2S1338 | D3S1358 | FRA | D5S818 | CSF1PO | SE33 | D7S820 | D8S1179 | D10S1248 | THO1 | vWA | D12S391 | D13S317 | D16S539 | D18S51 | D19S433 | D21S11 | D22S1045 | |
|--------|---------|--------|--------|---------|---------|--------|--------|--------|--------|--------|---------|----------|--------|--------|---------|---------|---------|--------|---------|--------|----------|--------|
| 6 | | | | | | | | | | | | | 0.0729 | | | | | | | | | |
| 7 | | | | | | | 0.0083 | | | | | | 0.3750 | | | | | | | | | |
| 8 | 0.6942 | | | | | | 0.046 | 0.0729 | | 0.1894 | | | 0.147 | | | | | | | | | |
| 9 | 0.1708 | 0.0070 | | | | | 0.0771 | 0.0543 | | 0.0521 | | | 0.2833 | | | 0.3146 | 0.0063 | 0.0163 | 0.0063 | | 0.0063 | |
| 9.3 | | | 0.0640 | | | | | | | | | | 0.0604 | | | 0.2021 | 0.2250 | | | | | |
| 10 | 0.0188 | 0.0250 | 0.0722 | | | | 0.2688 | 0.0767 | | 0.2303 | 0.0854 | | 0.0875 | | | | | | | 0.0021 | | |
| 10.3 | | | | | | | | | | | | | 0.0021 | | | 0.0000 | 0.1542 | | | | 0.0021 | |
| 11 | 0.188 | 0.2125 | 0.2559 | | | | 0.2303 | 0.3078 | | 0.3803 | 0.0708 | | 0.0063 | | | | | | | | | |
| 11.3 | | | 0.1646 | | | | | | | | | | | | | 0.2458 | 0.3375 | | 0.0032 | 0.0063 | | 0.2333 |
| 12 | 0.0333 | 0.0375 | 0.118 | | 0.0021 | | 0.2675 | 0.240 | | 0.036 | 0.1438 | 0.0333 | | | | | | 0.0620 | 0.0366 | | 0.0271 | |
| 13 | 0.0792 | | 0.0542 | | | | 0.1657 | 0.0896 | | 0.0063 | 0.2292 | 0.104 | 0.0042 | | | 0.0879 | 0.0708 | 0.003 | 0.2271 | | | |
| 13.2 | | | | | | | | | | | | | | | | 0.0292 | 0.0338 | | 0.0250 | | | |
| 14 | 0.0250 | | 0.1836 | | 0.0156 | | 0.0021 | 0.0025 | 0.046 | | 0.1675 | 0.2333 | | 0.246 | | | | 0.2303 | 0.1838 | | 0.0354 | |
| 14.2 | | | | | | | | | | | | | | | | 0.0063 | 0.0091 | | 0.0667 | | | |
| 15 | 0.0771 | | 0.0354 | | 0.3449 | | | | 0.025 | | 0.0792 | 0.225 | 0.0166 | 0.0542 | 0.0063 | | | 0.2771 | 0.0638 | | 0.3729 | |
| 15.2 | | | | | 0.0082 | | | | 0.0342 | | | | | | | | | 0.0063 | 0.2104 | | | |
| 16 | 0.2146 | | 0.0083 | | 0.2688 | | 0.0186 | | 0.0354 | | 0.1083 | 0.0979 | | 0.0854 | | | | 0.0208 | 0.0438 | | 0.0866 | |
| 16.2 | | | | | | | | | 0.0063 | | | | | | | | | | 0.0333 | | | |
| 16.3 | 0.0146 | | | | | | | | | | | | | | | 0.0041 | | | | | | |
| 17 | 0.146 | | | 0.0707 | 0.3021 | | | | 0.0792 | | 0.0186 | 0.0063 | | 0.2708 | 0.0542 | | | 0.0000 | 0.0021 | | 0.2167 | |
| 17.3 | 0.0163 | | | | | | | | | | | | | | | | | | | | | |
| 18 | 0.0208 | | | 0.0146 | 0.0583 | 0.0157 | | | 0.0438 | | 0.0020 | 0.0063 | | 0.2229 | 0.1638 | | | 0.0500 | | | 0.0083 | |
| 18.3 | 0.0123 | | | | | | | | | | | | | | | | | | | | | |
| 19 | | | | 0.2708 | | 0.0774 | | 0.0854 | | 0.0563 | | | 0.146 | 0.2563 | | | | 0.0521 | | | 0.004 | |
| 20 | | | | 0.0854 | | 0.0375 | | | 0.0604 | | | | 0.0042 | 0.0563 | | | | 0.0292 | 0.0497 | | | |
| 21 | 0.0042 | | | 0.0167 | | 0.055 | | | 0.0354 | | | | 0.0708 | | | | | | | | | |
| 22 | | | | 0.0792 | | 0.0750 | | | 0.0125 | | | | 0.082 | | | 0.0875 | | 0.004 | | | | |
| 22.2 | | | | | | 0.0042 | | | 0.0063 | | | | | | | | | | | | | |
| 23 | | | | 0.1896 | | 0.147 | | | | | | | | 0.0907 | | | | | | | | |
| 23.2 | 0.0102 | | | | | 0.0167 | | | 0.0729 | | | | | | | | | | | | | |
| 24 | | | | 0.1803 | | 0.0729 | | | 0.0063 | | | | | 0.0375 | | | | | | | | |
| 24.2 | | | | | | 0.004 | | | 0.0208 | | | | | | | | | | | | | |
| 25 | | | | 0.0333 | | 0.0875 | | | | | | | | 0.046 | | | | | | | | |
| 25.2 | | | | | | | | | 0.0521 | | | | | | | | | | | | | |
| 26 | | | | 0.0042 | | 0.0803 | | | | | | | | 0.0040 | | | | | | | | |
| 26.2 | | | | | | | | | 0.0375 | | | | | | | | | | | | | |
| 27 | | | | | | 0.0042 | | | | | | | | | | | | | | 0.0041 | | |
| 27.2 | | | | 0.0042 | | | | | 0.1521 | | | | | | | | | | | | 0.0792 | |
| 28 | | | | | | | | | 0.0354 | | | | | | | | | | | | | |
| 28.2 | | | | | | | | | 0.0063 | | | | | | | | | | | | 0.2833 | |
| 29 | | | | | | | | | 0.0729 | | | | | | | | | | | | 0.1675 | |
| 29.2 | | | | | | | | | 0.0063 | | | | | | | | | | | | 0.0338 | |
| 30 | | | | | | | | | | | | | | | | | | | | | 0.228 | |
| 30.2 | | | | | | | | | 0.0542 | | | | | | | | | | | | 0.0771 | |
| 31 | | | | | | | | | | | | | | | | | | | | | 0.0292 | |
| 31.2 | | | | | | | | | 0.004 | | | | | | | | | | | | 0.175 | |
| 32 | | | | | | | | | | | | | | | | | | | | | 0.0167 | |
| 32.2 | | | | | | | | | 0.0059 | | | | | | | | | | | | 0.0354 | |
| 33 | | | | | | | | | | | | | | | | | | | | | | |
| 33.2 | | | | | | | | | 0.0063 | | | | | | | | | | | | | |
| 34 | | | | | | | | | | | | | | | | | | | | | 0.004 | |
| 34.2 | | | | | | | | | 0.0063 | | | | | | | | | | | | | |
| 35 | | | | | | | | | | | | | | | | | | | | | | |
| 36 | | | | | | | | | | | | | | | | | | | | | | |
| 36.1 | | | | | | | | | | | | | | | | | | | | | 0.0042 | |
| 37 | | | | | | | | | | | | | | | | | | | | | | |
| 37.2 | | | | | | | | | | | | | | | | | | | | | 0.0042 | |

Table 2. Allele Frequency 21 STR Loci on the Madurese Populations in Surabaya.

| Loci | Expected heterozygosity (He) | Power of Discrimination (PD) | Power of Exclusion (PE) | Polymorphic Information Content (PIC) | Paternity Index (PI) |
|----------------|------------------------------|------------------------------|-------------------------|---------------------------------------|----------------------|
| D1S1656 | 0,840 | 0,962 | 0,673 | 0,756 | 0,144 |
| TPOX | 0,769 | 0,924 | 0,683 | 0,688 | 0,094 |
| D2S441 | 0,748 | 0,892 | 0,690 | 0,679 | 0,110 |
| D2S1338 | 0,832 | 0,857 | 0,666 | 0,673 | 0,113 |
| D3S1358 | 0,831 | 0,867 | 0,667 | 0,683 | 0,067 |
| FGA | 0,879 | 0,951 | 0,769 | 0,690 | 0,076 |
| D5S818 | 0,855 | 0,968 | 0,748 | 0,666 | 0,083 |
| CSF1PO | 0,792 | 0,867 | 0,832 | 0,704 | 0,111 |
| SE33 | 0,741 | 0,884 | 0,695 | 0,673 | 0,097 |
| D7S820 | 0,704 | 0,860 | 0,642 | 0,704 | 0,078 |
| D8S1179 | 0,705 | 0,954 | 0,688 | 0,686 | 0,129 |
| D10S1248 | 0,832 | 0,921 | 0,695 | 0,678 | 0,112 |
| THO1 | 0,715 | 0,921 | 0,642 | 0,653 | 0,093 |
| vWA | 0,662 | 0,953 | 0,673 | 0,673 | 0,031 |
| D12S391 | 0,842 | 0,919 | 0,683 | 0,707 | 0,124 |
| D13S317 | 0,769 | 0,968 | 0,690 | 0,656 | 0,106 |
| D16S539 | 0,748 | 0,957 | 0,666 | 0,688 | 0,097 |
| D18S51 | 0,866 | 0,866 | 0,704 | 0,695 | 0,129 |
| D19S433 | 0,911 | 0,876 | 0,673 | 0,642 | 0,144 |
| D21S11 | 0,716 | 0,897 | 0,695 | 0,665 | 0,254 |
| D22S1045 | 0,607 | 0,952 | 0,642 | 0,765 | 0,011 |
| Average | 0,743 | 0,910 | 0,691 | 0,687 | 0,105 |

Table 3. Expected Heterozygosity (He), Power of Discrimination (PD) and Polymorphic Information Content (PIC) on Madurese Populations in Surabaya.

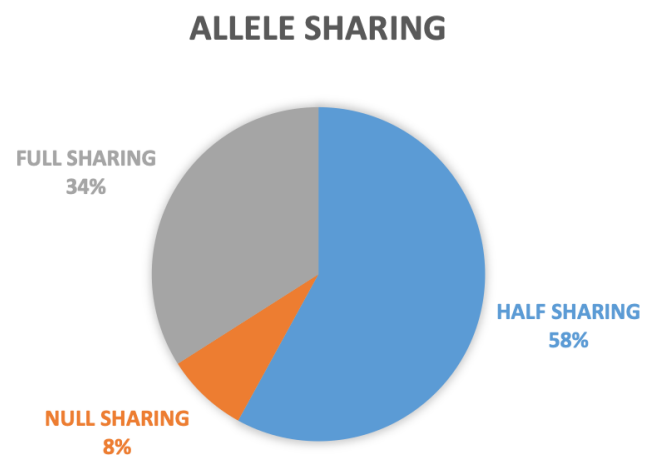


Figure 1. Percentage of 21 Loci STR Allele Sharing on Madurese Populations in Surabaya.

| Family | SI | Family | SI |
|--------|------|--------|------|
| 1 | 1,02 | 6 | 1,05 |
| 2 | 1,08 | 7 | 1,06 |
| 3 | 0,90 | 8 | 0,89 |
| 4 | 0,98 | 9 | 0,88 |
| 5 | 1,08 | 10 | 1,20 |

Table 4. Sibship Index (SI) on Madurese Populations in Surabaya.

Discussion

The TPOX locus has the lowest allele variant (5 allele variants), while the SE33 locus has the highest allele variant (24 allele variants). The allele frequency is employed to see the genetic diversity of each population. A locus is polymorphic if the number of alleles in the population at that locus is more significant than one with an allele frequency less than or equal to 0.95^{7,8}. The highest expected He value was obtained at locus D19S433 (0.911), while the lowest was obtained at locus D22S1045 (0.607). The highest PD value was obtained at 2 loci, namely D5S818, and D13S317 (0.968), while the lowest was at locus D2S1338 (0.857). The highest PIC value was at locus D22S1045 (0.765), with the average of all PIC values being > 0.5, indicating that all loci are highly polymorphic.

The absence of allele sharing in siblings implies that siblings do not always share their alleles with the other siblings, per the theoretical facts of Mendelian inheritance laws⁹. The SI value is 0.88 – 1.20. A more considerable SI value (>1.0) generally corresponds to a higher likelihood of a relationship. The higher the sibship index above 1.0, the higher the likelihood that the two individuals are related as siblings. Similarly, a smaller SI (<1.0) corresponds to a lower likelihood of a relationship¹⁰⁻¹².

The statistical calculations on the sample of Madurese living in Surabaya 21 STR loci showed significant results. Genetic variation in the sample closely correlates with several cultures inherent in the Madurese, commonly known as an endogamy marriage^{13,14}. This study also directly aligns with the migrating culture or venturing to other cities, such as migrating to Surabaya, seen from the SI value > 1.0 (60%), while < 1.0 (40%).

Conclusions

Based on the allelic frequency, the expected heterozygosity (He), the Power of Discrimination (PD), the Polymorphic Information Content (PIC), and the Sibship Index (SI), it can be inferred that this autosomal STR locus could be well applied in individual identifications and genetic variation studies through siblings examination, among Madurese population.

Declaration of Interest

The authors report no conflict of interest.

References

- Jacewicz R, Berent J, Prośniak A, Galecki P, Florkowski A, Szram S. Population genetics of the Identifiler system in Poland. *Int Congr Ser.* 2004;1261:229-232. doi:10.1016/S0531-5131(03)01543-7.
- Butler E, Li R. Genetic Markers for Sex Identification in Forensic DNA Analysis. *J Forensic Investigation.* 2014;2(3):3.
- Yudianto A, Kurniawan A, Furqoni AH, Rizky BN. Paternity Test Through Kinship Analysis and Cell Free Fetal DNA (Cff-DNA) as a Forensic Identification Technique. *Journal of International Dental and Medical Research.* 2022;15(4):1434-1441. Accessed November 13, 2023. <http://www.iidmr.com/journal/contents-of-jidmr-2022-vol-15-no-4/>
- Singh Negi D, Alam M, Bhavani SA, Nagaraju J. Multistep microsatellite mutation in the maternally transmitted locus D13S317: a case of maternal allele mismatch in the child. *Int J Legal Med.* 2006;120(5):286-292. doi:10.1007/s00414-006-0080-3
- Butler, John M. Advanced Topics in Forensic DNA Typing: Interpretation. In: Elsevier; 2015:19-20.
- Butler, John M. Advanced Topics in Forensic DNA Typing: Methodology. In: Elsevier; 2012:107.
- Reff F, Widyastuti SK, Wandia IN. Polymorphism of D10S1432 Microsatellite Locus on Long Tailed Macaque Population in Sangeh. *Jurnal Ilmu dan Kesehatan Hewan, Pebruari.* 2013;1(1):16-21.
- Iza N. Allele Frequency, Heterozygosity, and Allele Migration in Javanese and Madurese Population in Malang and Madura, East Java Indonesia. *JURNAL ILMIAH SAINS.* 2017;17(1):43. doi:10.35799/jis.17.1.2017.15289
- Reid TM, Baird ML, Reid JP, Lee SC, Lee RF. Use of sibling pairs to determine the familial searching efficiency of forensic databases. *Forensic Sci Int Genet.* 2008;2(4):340-342. doi:10.1016/j.fsigen.2008.04.008
- Maeda K, Murakami C, Irie W, et al. The case of 2 siblings that identified not only by DNA profiling. *Forensic Sci Int Genet Suppl Ser.* 2015;5:e555-e556. doi:10.1016/j.fsigs.2015.09.219
- Tamura T, Osawa M, Ochiai E, Suzuki T, Nakamura T. Evaluation of advanced multiplex short tandem repeat systems in pairwise kinship analysis. *Leg Med.* 2015;17(5):320-325. doi:10.1016/j.legalmed.2015.03.005
- Abbas S, Mourad L, Mansour I. Evaluation of Sibling-Ship Analysis in Secluded Lebanese Villages with Increased Mating Patterns. *Journal of Forensic Investigation.* 2018;6(1):1.
- Meianzi Yasak E, Indra Dewi S. Budaya Pernikahan Dini terhadap Kesetaraan Gender Masyarakat Madura. *Jurnal Ilmu Sosial dan Ilmu Politik.* 2015;4(3):426. doi:https://doi.org/10.33366/jisip.v4i3.123.
- Raharja MB. Fertility by Ethnicity in Indonesia: Analysis of 2010 Indonesian Population Census. *Jurnal Kependudukan Indonesia.* 2017;12(1):69. doi:10.14203/jki.v12i1.243.