

Triangular fuzzy number for similarity measurement of Y chromosome DNA profile

Meira Parma Dewi¹, Aniati Murni Arymurthy¹, Suryana Setiawan¹, Nurtami Soedarsono²

¹Department of Computer Science, Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia

²Department of Oral Biology, Faculty of Dentistry, Universitas Indonesia, Depok, Indonesia

Article Info

Article history:

Received Nov 20, 2022

Revised Jul 13, 2023

Accepted Aug 2, 2023

Keywords:

DNA profile

Fuzzy similarity

Loci DNA

Short tandem repeat

Y-chromosome

ABSTRACT

This study measures the similarity of the short tandem repeat (STR) profile of human DNA. The similarity measurement had been done to the STR value of the allele loci in DNA profile between the query's DNA to the reference's DNA profile. The measurements were conducted on 27 DNA profile loci including the Y chromosome loci (YSTR). The YSTR loci were used as the main comparison of similarity measurements to determine the biological kinship relationship between the query DNA profile and the alleged male biological family. To measure the similarity of two STR values that have shifted due to several factors in the DNA source extraction process, a fuzzy similarity measure was used. The STR values of the DNA profile loci are described as triangular fuzzy numbers. Similarity value of the STR is the intersection of two isosecle that been compared. To conclude that the query has a biological relationship with the male reference, the similarity of the YSTR locus is equal or more than 0.75 and the similarity value of the other 24 DNA profile loci is greater or equal to 0.5. From the trial that have been done, 90% give the right results.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Meira Parma Dewi

Faculty of Computer Science, Universitas Indonesia

Depok, West Java, Indonesia

Email: meira.parma91@ui.ac.id

1. INTRODUCTION

The high number of recent terrorist bombings in Indonesia, natural disasters that result in many casualties, and the increasing number of killings with the modus operandi of mutilation have made it difficult or impossible for the victim to be recognized because of the damage to some or all their limbs [1]. The victims can be identified by examining their DNA [2]–[4]. Indonesia has carried out the process of identifying disaster victims and victims of crime by matching the victim's DNA profile with the alleged biological family. Identification of DNA profiles is conducted on biological evidence from the human body using polymerase chain reaction (PCR) technology and short tandem repeat (STR) sequences [4]–[6]. The DNA profile is a unique genetic fingerprint that distinguishes one individual from another because it is an inherited molecule.

The identification of DNA profiles, previously known as the combined DNA index system (CODIS), maps 16 STR loci, namely CSF1PO, D13S317, D16S539, D18S51, D19S433, D21S11, D2S1338, D3S1358, D5S818, D7S720, D8S1179, FGA, TH01, TPOX, vWA, and sex-determining amelogenin [2]–[7]. Previous studies showed that this 16 STR locus data has been successfully used to measure the similarity of the victim's DNA profile and their family reference. However, these loci sometimes cannot provide biological information from the victim due to a lack of required data or mixed interpretation in the identification process. This is the basis for the development of the PowerFlex 6C System Fusion which contains 27 STR loci, namely CSF1PO, FGA, TH01, vWA, D1S1656, D2S1338, D2S441, D3S1358, D5S818, D7S820,

D8S1179, D10S1248, D12S391, D13S317, D43S51539, D43S51, and D21S11 known as amelogenin and DYS391 as sex determinants, penta D, penta E, D22S1045, TPOX, SE33 two repeat mutations of Y-STRs, i.e., locus DYS570 and DYS576 [2], [3].

Previously, many studies have been conducted related to matching human DNA profiles. Some of them matched the DNA profile with the alleged victim's biological family used fuzzy inference system [6]–[10]; and DNA profile matching involving tribal information [8]; DNA matching with the victim's family using the Gaussian fuzzy number method [11]. From the previous method, the measurement of DNA similarity used 16 loci of DNA profile. While this study use Y-chromosome (YSTR) to measure the similarity of DNA profile involving family relationship.

YSTR is the male STR locus [12], [13]. Thus, the measurement of DNA profile similarity using YSTR should be conducted on male queries and male references. Y-STRs are Y chromosomes found in STRs used to identify male lineages. YSTR shows a high degree of variability which indicates the existence of a kinship relationship of male STR DNA in a population [14]. The advantage of this locus is the data consists of male and female reference DNA profiles with unclear traces or origins like in cases of sexual violence where the STR DNA profile data obtained is the DNA profile of a woman and a man.

One example of a crime that requires a YSTR identity is in a case of sexual violence. The evidence of DNA in the form of sperm, a DNA profile will be obtained which shows a difference in STR values on Y chromosome. From the YSTR value found [15], it will be known how many men were involved in this cases and the YSTR will make it easier to find out the identity of the perpetrators.

In the identification of human DNA profiles, the STR value indicated by the DNA locus is an integer. However, due to several factors such as contaminated DNA sources, due to weathering and other causes, the STR value can shift to a decimal value. If the measurement of DNA profile similarity is done crisply, the decimal value indicated by the locus must be converted to an integer. To change the STR value, re-sampling and re-extraction of DNA sources must be conducted, which is costly and time-consuming. To accommodate the shift in the STR value at the DNA locus, the fuzzy method was used to measure the similarity of the STR value at the DNA profile locus.

2. MATERIAL AND METHOD

2.1. STR DNA profile

The DNA profile is an individual DNA structure that describes their biological identity. DNA profiles of somebody consist of 16 loci, each of which maps STR from the specifications of each locus. A person's DNA profile is identified by examining their biological evidence, also known as DNA evidence, which can be obtained from several parts of the body, such as blood, saliva, bones, muscles, sperm, teeth, hair, or body fluids such as urine and sweat [2], [16], [17].

Human DNA profiles can be identified through the examination of STR markers. STR is a repeating pattern of nucleotides that usually consists of 2–6 bases with the same pattern without any other sequence or intervention of a different sequence. Currently, the PowerPlex Fusion 6C System has been developed which contains 27 STR loci, namely CSF1PO, FGA, TH01, vWA, D1S1656, D2S1338, D2S441, D3S1358, D5S818, D7S820, D8S1179, D10S1248, D12S391, D13S317, D1816S5392119, D13S5392119 and known as amelogenin and DYS391 as sex determinants, penta D, penta E, D22S1045, TPOX, SE33 two repeating mutations of Y-STRs, i.e., locus DYS570 and DYS576. Each locus consists of a pair of alleles, each of which is inherited from the biological father and mother.

2.2. YSTR

YSTR is a Y chromosome found in the human DNA profile in male lineages. YSTR is the STR locus that is only owned by men. Thus, the measurement of DNA profile similarity using YSTR should be conducted on male queries and male references. YSTR analysis is commonly used for forensic purposes, paternity and genealogical DNA testing. YSTR is acquired from the paternal lineage. Therefore, a boy will have the same YSTR value as his biological father and his male sibling.

The following diagram in Figure 1 illustrates the linking of the Y chromosome (YSTR) in males, where squares represent males and circles represent females. YSTR is the male STR locus. Thus, the measurement of DNA profile similarity using YSTR should be conducted on male queries and male references. The Figure 2 shows that the measurement of DNA profile similarity using YSTR was conducted on male queries with reference to the DNA profile of the male victim's biological family. In this case, the reference used was the victim's grandfather's STR DNA profile from the father's side. This following diagram in Figure 3 describes the measurement of YSTR DNA with brother of victim's father as reference.

The references selected to compare DNA profile data if the similarity is measured based on YSTR are all males who have a biological relationship with the victim from the father’s lineage. Candidates for reference include the father, brother, father’s brother, grandfather from the father’s side, male cousin, and nephew.

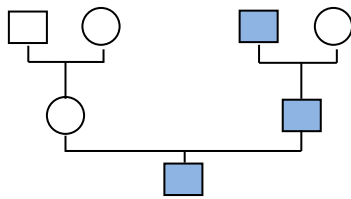


Figure 1. Y-chromosome

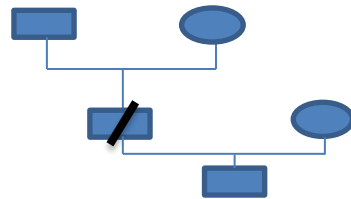


Figure 2. DNA similarity measurement with grandfather as reference

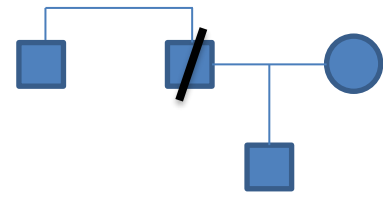


Figure 3. DNA similarity measurement with father’s brother as reference

2.3. Fuzzy inference system

The measurement of the similarity of STR DNA profiles at a DNA locus is done by measuring the similarity of each allele at 27 loci of DNA profiles. In measuring the similarity of human DNA profiles, reference is made to the victim’s biological father and mother. The rules for matching the victim’s DNA profile with the father’s and mother’s DNA profiles are:

- If [(victim_allele_1 ≈ allele_1_father) or (victim_allele_1 ≈ allele_2_father)] and [(victim_allele_2 ≈ allele_1_mother) or (victim_allele_2 ≈ allele_2_mother)] then match.
- If [(victim_allele_1 ≈ allele_1_mother) or (victim_allele_1 ≈ allele_2_mother)] and [(victim_allele_2 ≈ allele_1_father) or (victim_allele_2 ≈ allele_2_father)] then match.

However, when the father’s DNA profile data is not found, the similarity measurement can be made to the father’s biological sibling. In this case, the DNA profile data from the male reference is used.

- If [(victim_allele_1 ≈ allele_1_Ref_A) or (victim_allele_1 ≈ allele_2_Ref_A)] and [(victim_allele_2 ≈ allele_1_mother) or (victim_allele_2 ≈ allele_2_mother)] then match.
- If [(victim_allele_1 ≈ allele_1_mother) or (victim_allele_1 ≈ allele_2_mother)] and [(allele_2_victim ≈ allele_1_Ref_A) or (victim_allele_2 ≈ allele_2_Ref_A)] then match.

In the DNA profile matching using the Y chromosome (YSTR), similarity measurements can only be made against references from surrogate fathers.

- If [(victim_allele_1 ≠ allele_1_mother) or (victim_allele_1 ≠ allele_2_mother)] and [(victim_allele_2 ≈ allele_1_Ref_B) or (victim_allele_2 ≈ allele_2_Ref_B)] then match.

2.4. Membership functions of input and output variables

In fuzzy system there are membership functions for input and output. Input variables have three membership functions, namely small, medium, and large. Meanwhile the output variable consist of 2 membership functions, which are suitable and unsuitable.

$$f_{small}(x) = \begin{cases} 1 & x \leq 0.15 \\ \frac{0.3-x}{0.15} & 0.15 \leq x \leq 0.3 \\ 0 & x \geq 0.3 \end{cases}$$

$$f_{medium}(x) = \begin{cases} 0 & x \leq 0.2; x \geq 0.5 \\ \frac{x-0.2}{0.15} & 0.2 \leq x \leq 0.35 \\ \frac{0.5-x}{0.15} & 0.35 \leq x \leq 0.5 \end{cases}$$

$$f_{large}(x) = \begin{cases} 0 & x \leq 0.4 \\ \frac{x-0.4}{0.1} & 0.4 \leq x \leq 0.5 \\ 1 & x \geq 0.5 \end{cases}$$

2.5. Measurement of DNA profile STR similarity

In contrast to previous studies, measurement of allele similarity at each DNA profile locus will be carried out by considering references, one of which is a substitute for a biological father. Which this will affect the

similarity value generated by the process of measuring the overall DNA profile similarity [18]–[20]. Therefore a reference from biological mother is needed as a source of half of a pair of alleles at each DNA profile locus.

Measurement of the similarity of STR values at the DNA profile locus was conducted for each allele at each DNA locus [21]. The similarity measurement of the STR value was conducted using the fuzzy similarity measure [22], [23]. Fuzzy similarity is used to accommodate the shift in the STR value at the DNA locus [24], [25]. Shifts in STR values can occur due to several factors that cause the extraction results from DNA sources to be no longer pure. This causes the STR value, which should be an integer, to be a real number.

The STR value of the DNA profile is described as an isosceles triangle whose middle value is the STR value of the DNA profile. The height of the triangle set to 1 while the distance between the two legs of the triangle is set to 0.4. Figure 4 describes two isosceles which have a difference in STR value of 0.1. The measurement of similarity of two STR values that are suspected to have shifted the STR value, fuzzy similarity measure was used. The similarity value of the two STR values was obtained from the intersection of the two triangles use (1). Figure 5 shows the intersection of two triangle as the similarity value. The shift in the value of STR tolerated in this study was 0.2. The results of the similarity measurement using the fuzzy similarity measure gave a value in the range of 0 to 1.

$$t = \frac{1/2(a3-b1)}{a3-a2} \tag{1}$$

Where t is the intersection point of two allele; $a2$ is STR value of first allele; $a3$ is $a2 + 0.2$; $b1$ is STR value of second allele. So $0 \leq t \leq 1$.

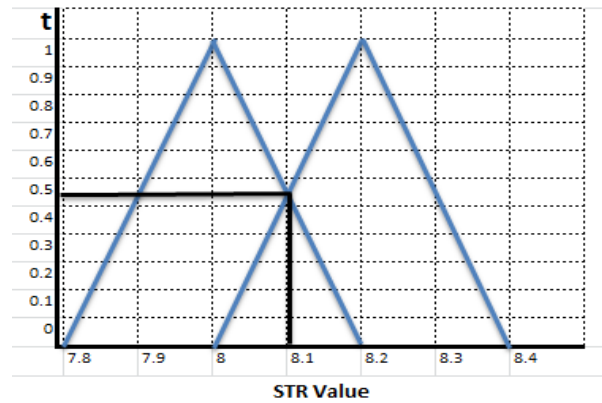


Figure 4. Triangular fuzzy number

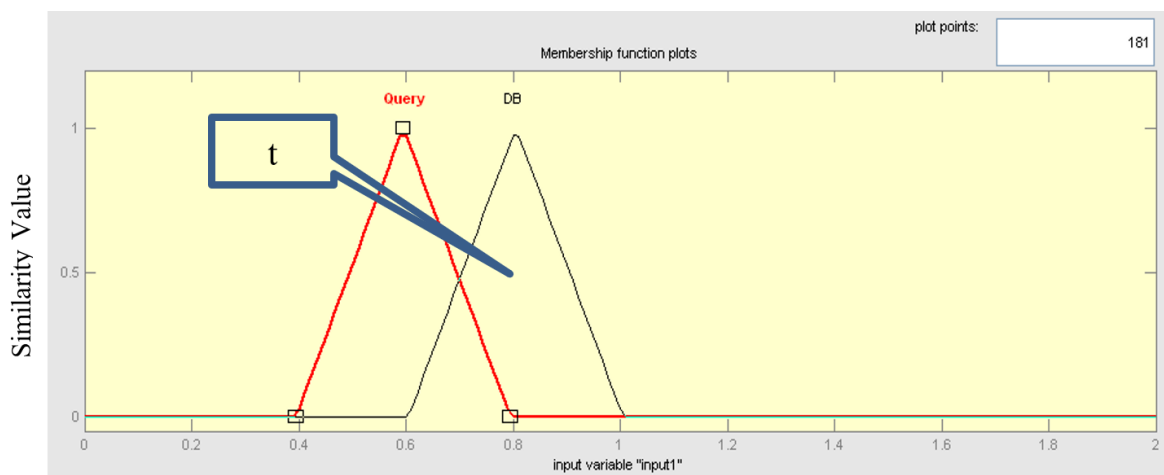


Figure 5. Similarity measurement of two DNA profile

The similarity measurement of the two DNA profiles was performed by measuring the similarity of the STR value of allele 1 to the STR value of alleles at the same locus from the first reference, and the second allele was compared or its similarity was measured with alleles at the same locus from the second reference. For the Y-STR locus, the similarity measurement of the STR value of a certain locus was compared against both alleles of the same locus from a male reference. Accordingly, for the YSTR locus, the locus similarity value is obtained from equation similarity in (2):

$$t = \frac{t_1+t_2}{2} \quad (2)$$

Where t_1 is similarity value of allele 1 and t_2 is similarity value of allele 2.

2.6. Data

The data used in this study are the DNA profiles of 27 loci of Javanese which consist of 100 samples. All data are obtained from the Faculty of Dentistry University of Indonesia. Table 1 is an example of DNA profile consist of victim data (query), biological mother data, and DNA profile data from biological relatives of the victim's father.

Table 1. Data of DNA profile

Loci	Victim		Mother		Grandfather		Uncle	
	Alel 1	Alel 2	Alel 1	Alel 2	Alel 1	Alel 2	Alel 1	Alel 2
AMEL	X	Y	X	X	X	Y	X	Y
D3S1358	15.2	16	15	16	15	20	15	16
D1S1656	16	17	13	16	13	15	13	16
D2S441	10	11.2	12	11	10	11	11	11
D10S1248	13	15	13	17	13	15	13	13
D13S317	10.2	10.2	10	11	8	10	10	10
Penta E	12	12	12	22	12	19	19	22
D16S539	9	13	9	10	9	11	9	13
D18S51	13.2	15	15	19	7	17	16	17
D2S1338	17.2	26	19	26	17	23	19	23
CSF1PO	11	11	10	11	11	12	10	11
Penta D	7.2	9	7	9	7	9	9	9
TH01	7.3	10	7	7	8	10	7	10
vWA	17	17	14	17	16	17	14	17
D21S11	30	31.2	30	32.2	29	32	30	32.2
D7S820	8	8.2	8	12	8	9	9	12
D5S818	12.2	12	10	12	12	13	10	12
TPOX	8	8	8	8	6	8	8	8
D8S1179	11.3	13	11	17	11	13	13	17
D12S391	19.2	19	19	23	19.2	20	20	20
D19S433	13	14	13	14	13	14	14	14
SE33	28.2	28.2	22	28	19	28	28.2	30.2
D22S1045	15	16	15	17	15	16.2	15	15
FGA	22	24.2	22	26	24	26	26	26
DYS391	9	9			9	9	9	9.2
DYS576	19	19			19	19	19	19
DYS570	22	22			22	22	22	22

3. RESULTS AND DISCUSSION

During the identification, the similarity of two DNA profiles must be measured by the similarity of DNA profile markers (STR). Measurement of the similarity of STR DNA profiles using Y chromosomes can be used for paternity tests, where DNA profile data for suspected biological fathers are not found. Thus, we need DNA profile data from the biological brother of the father. Measurement of DNA profile similarity was conducted on 27 DNA loci.

Similarity measurement of DNA profile using fuzzy similarity will give the similarity value of each allele in the loci in question. The similarity value of the DNA profile is the average of all allele similarity value. By using fuzzy similarity, the similarity value of two allele that compared are in the range of 0 and 1. Similarity measurement of DNA profile using fuzzy similarity will give the similarity value of each allele in the loci in question. The similarity value of the DNA profile is the average of all allele similarity value. Table 2 shows the comparison of the similarity measurements of DNA STR values using the crisp method and the fuzzy method.

Table 2. Comparison similarity value of two allele

Query	Mother	Similarity value	
		Crisp method	Fuzzy
15	15.2	0	0.5
9	9	1	1

From Table 2 we can compare the value of similarity measurement using the crisp method to give a value of 0 when STR value of query is 15 and reference STR value is 15.2. This is caused by $15 \neq 15.2$. Whereas in measurements using the fuzzy method, the similarity of the STR value of 15 with the STR value of 15.2 will give a similarity of 0.5. Which means 15 is similar or almost the same as 15.2.

The similarity measurement is carried out from the allele from the mother first, and the other alleles will be measured for similarity with the father's surrogate reference. The following Table 3 is a comparison table for measuring similarity values using the crisp method and the fuzzy method for all DNA profile loci. Table 3 illustrates a comparison of similarity values for the entire query DNA profile with the mother's DNA profile. The similarity value with the crisp method gives a value of 0 or 1 for each locus allele being compared. While the fuzzy method gives similarity values in the range 0 and 1. Similarity value of each allele will be accumulated and the average will show the similarity value of all locus alleles. So that the similarity value of the fuzzy method is greater than the similarity value using the crisp method. The same case can be seen in Table 4.

Table 3. Comparison of STR DNA profile similarity values using the crisp method and the fuzzy similarity measure method

Loci	Victim		Mother		Crisp similarity	Fuzzy similarity
	Alel 1	Alel 2	Alel 1	Alel 2		
AMEL	X	Y	X	X	1.00	1.00
D3S1358	15.2	16	15	16	0	1.00
D1S1656	16	17	13	16	1.00	1.00
D2S441	10	11.2	12	11	0	0.50
D10S1248	13	15	13	17	1	1.00
D13S317	10.2	10.2	10	11	0	0.50
Penta E	12	12	12	22	1	1.00
D16S539	9	13	9	10	1	1.00
D18S51	13.2	15.2	15	19	0	0.50
D2S1338	17.2	26	19	26	1	1.00
CSF1PO	11	11.2	10	11	1.00	1.00
Penta D	7.2	9	7	9	0	0.50
TH01	7.3	10	7	7	0	0.18
vWA	17	17	14	17	1	1.00
D21S11	30	31.2	30	32.2	1	1.00
D7S820	7	8.2	8	12	0	0.50
D5S818	12.2	12	10	12	0	0.50
TPOX	8	8	8	8	1	1.00
D8S1179	11.3	13	11	17	0	0.18
D12S391	19.2	19	19	23	0	0.50
D19S433	13	14	13	14	1	1.00
SE33	28.2	28.2	22	28	0	0.50
D22S1045	15	16	15	17	1	1.00
FGA	22	24.2	22	26	1	1.00
DYS391	9	9				
DYS576	19.2	19.2				
DYS570	22	22				
	Similarity value of DNA				0.54	0.79

Tables 4 show the value of the similarity measurement between query's STR DNA profile to the grandfather's DNA profile using crisp and fuzzy method. Similarity value of fuzzy method is greater than similarity value of crisp method in the same locus's alleles. This also applies to the comparison of crisp and fuzzy method for measurement the similarity between DNA profile of query and uncle.

Table 3 to Table 5 show that the similarity values of the two DNA profiles are compared using the crisp method and the fuzzy similarity measure method. The value of each allele at each locus compared with the crisp method is 0 or 1 while using the fuzzy similarity measure, the allele similarity value is in the range of 0–1. This will have an impact on the overall DNA profile similarity value where the value of the DNA profile similarity with the fuzzy similarity measure will be greater than that with the crisp method.

Table 4. Comparison of STR DNA profile similarity values using the crisp method and the fuzzy similarity measure method with the reference of grandfather from the father's side

Loci	Victim		Grandfather		Crisp similarity	Fuzzy similarity
	Allel 1	Allel 2	Allel 1	Allel 2		
AMEL	X	Y	X	Y	1.00	1.00
D3S1358	15.2	16	15	20	0.00	0.50
D1S1656	16	17	13	15	0.00	0.00
D2S441	10	11.2	10	11	0.00	1.00
D10S1248	13	15	13	15	1.00	1.00
D13S317	10.2	10.2	8	10	0.00	1.00
Penta E	12	12	12	19	1.00	1.00
D16S539	9	13	9	11	0.00	0.00
D18S51	13.2	15.2	7	17	0.00	0.00
D2S1338	17.2	26	17	23	0.00	0.50
CSF1PO	11	11.2	11	12	0.00	0.50
Penta D	7.2	9	7	9	1.00	1.00
TH01	7.3	10	8	10	1.00	1.00
vWA	17	17	16	17	1.00	1.00
D21S11	30	31.2	29	32	0.00	0.00
D7S820	7	8.2	8	9	0.00	0.00
D5S818	12.2	12	12	13	1.00	1.00
TPOX	8	8	6	8	1.00	1.00
D8S1179	11.3	13	11	13	0.00	1.00
D12S391	19.2	19	19.2	20	0.00	1.00
D19S433	13	14	13	14	1.00	1.00
SE33	28.2	28.2	19	28	0.00	0.50
D22S1045	15	16	15	16.2	0.00	0.50
FGA	22	24.2	24	26	0.00	0.50
DYS391	9	9	9	9	1.00	1.00
DYS576	19.2	19.2	19	19	0.50	0.50
DYS570	22	22	22	22	1.00	1.00
Similarity value of DNA					0.44	0.67

Table 5. Comparison of STR DNA profile similarity values using the crisp method and the fuzzy similarity measure with uncle from the father's side as reference

L O C I	Victim		Uncle		Crisp Similarity	Fuzzy Similarity
	Allel 1	Allel 2	Allel 1	Allel 2		
AMEL	X	Y	X	Y	1	1
D3S1358	15.2	16	15	16	1	1
D1S1656	16	17	13	16	0	0
D2S441	10	11.2	11	11	0	0
D10S1248	13	15	13	13	0	0
D13S317	10.2	10.2	10	10	0	0.5
Penta E	12	12	19	22	0	0
D16S539	9	13	9	13	1	1
D18S51	13.2	15.2	13	17	0	0.5
D2S1338	17.2	26	19	23	0	0
CSF1PO	11	11.2	10	11	0	0.5
Penta D	7.2	9	9	9	1	1
TH01	7.3	10	7	10	1	1
vWA	17	17	14	17	1	1
D21S11	30	31.2	30	32.2	0	0.18
D7S820	7	8.2	9	12	0	0
D5S818	12.2	12	10	12	1	1
TPOX	8	8	8	8	1	1
D8S1179	11.3	13	13	17	1	1
D12S391	19.2	19	20	20	0	0
D19S433	13	14	14	14	1	1
SE33	28.2	28.2	28.2	30.2	1	1
D22S1045	15	16	15	15	0	0
FGA	22	24.2	26	26	0	0
DYS391	9	9	9	9.2	1	0.5
DYS576	19.2	19.2	19	19	0	1
DYS570	22	22	22	22	1	1
Similarity value of DNA					0.48	0.56

Measurement of the similarity of the DNA profile between the query with the mother and the biological family of the father from the male side was conducted by first measuring the similarity of one allele from the same locus to the mother. The other alleles at the locus were measured for similarity to the

alleles from the surrogate biological father reference. As in the D21S11 locus, the first allele was compared with the allele at the same locus with the maternal DNA profile, namely the STR value of 30. The second allele with the STR value of 31.2 was compared to the allele at the D21S11 locus of the grandfather's DNA profile with a similarity value of 0. The Table 6 describes the comparison similarity value of YSTR between query's YSTR value to uncle's YSTR value and grand father's YSTR value.

By compare two DNA profile using fuzzy similarity, we obtained the similarity value using fuzzy is higher than using crisp method. Figure 6 illustrates the difference in DNA similarity values using the crisp and fuzzy methods. Figure 6 shows the similarity value with the fuzzy method is greater than the similarity value using the crisp method. Figure 7 describes the comparison of DNA YSTR similarity value between uncle and grandfather using crisp and fuzzy methods. To conclude the similarity of the query DNA profile with the biological family of the alleged father, the overall DNA profile similarity value must be greater than or equal to 0.5, and the similarity value for the 3 allele pairs of the Y chromosome (YSTR) is greater or equal to 0.75.

Table 6. Comparison similarity value of YSTR

Loci	Query		Uncle		Grandfather	
	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
DYS391	9	9	9	9.2	9	9
DYS576	19.2	19.2	19	19	19	19
DYS570	22	22	22	22	22	22
YSTR similarity value	Fuzzy		0.75		0.83	
	Crisp		0.5		0.67	

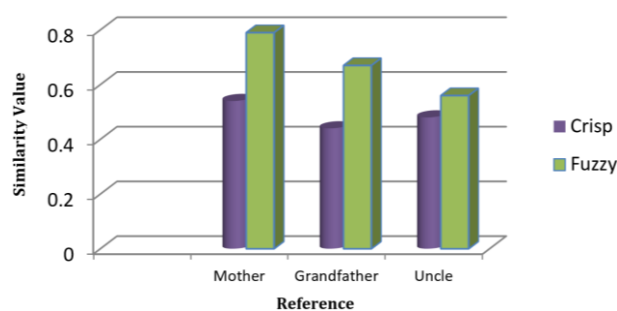


Figure 6. Comparison chart of DNA similarity value using fuzzy and crisp method

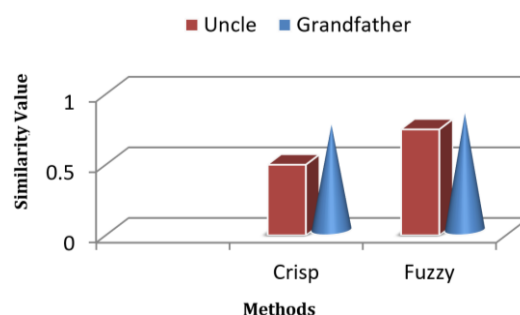


Figure 7. Comparison chart of YSTR similarity value using fuzzy and crisp method

4. CONCLUSION

Measurement of the similarity of human DNA profiles using the Y chromosome (YSTR) can be conducted to determine the kinship between the query DNA profile and the alleged biological family of the male. YSTR is passed down by males to their offspring without changing. However, the STR value of the DNA locus allele can experience a shift due to several factors. Measurement of DNA profile similarity using fuzzy similarity measure to STR values at DNA profile loci gives better results than using crisp or manual methods. With the fuzzy similarity measure, the shift of the STR value can be accommodated so that it can provide a similarity value in the range of values 0–1. To conclude whether the query has a kinship with a family reference, the DNA profile similarity value for the Y chromosome must be equal to or greater than 0.75.

ACKNOWLEDGEMENTS

Thank you to the Indonesia Endowment Funds for Education (LPDP), Ministry of Finance, Republic of Indonesia as my scholarship sponsor.




REFERENCES

- [1] M. P. Dewi and N. Soedarsono, "The Similarity Measurement of Human DNA Profile Using Fuzzy Similarity," *Eksakta : Berkala Ilmiah Bidang MIPA*, vol. 21, no. 1, pp. 46–53, Apr. 2020, doi: 10.24036/eksakta/vol21-iss1/216.
- [2] W. Goodwin, *Forensic DNA Typing Protocols*, vol. 1420. New York, NY: Springer New York, 2016, doi: 10.1007/978-1-4939-3597-0.




- [3] J. M. Butler, *Forensic DNA Typing Biology, Technology, and Genetics of STR Markers*, 2nd Editio. United States of America: Elsevier Science Publishing, 2005.
- [4] E. Butler and R. Li, "Genetic Markers for Sex Identification in Forensic DNA Analysis," *Journal of Forensic Investigation*, vol. 2, no. 3, pp. 1–10, 2014.
- [5] S. J. Venables *et al.*, "Allele frequency data for 15 autosomal STR loci in eight Indonesian subpopulations," *Forensic Science International: Genetics*, vol. 20, pp. 45–52, Jan. 2016, doi: 10.1016/j.fsigen.2015.09.014.
- [6] M. R. Widyanto, R. N. Hartono, and N. Soedarsono, "A Novel Human STR Similarity Method using Cascade Statistical Fuzzy Rules with Tribal Information Inference," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, no. 6, Dec. 2016, doi: 10.11591/ijece.v6i6.11810.
- [7] M. S. Anggreainy, M. R. Widyanto, and B. Widjaja, "Weighting for DNA Profiling," in *Proceedings of the 1st International Conference on Algorithms, Computing and Systems*, New York, NY, USA: ACM, Aug. 2017, pp. 23–26, doi: 10.1145/3127942.3127960.
- [8] M. S. A. Anggreainy, M. R. Widyanto, B. H. Widjaja, N. Soedarsono, and K. Hirota, "Tribal Classification Using Probability Density Function (PDF) and Fuzzy Inference System (FIS)," in *2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS)*, IEEE, Dec. 2018, pp. 1181–1184, doi: 10.1109/SCIS-ISIS.2018.00186.
- [9] N. H. Faleh and K. H. Karidi, "Fuzzy Association Rule Classification System For Human DNA profile Identification," *IOP Conference Series: Materials Science and Engineering*, vol. 928, no. 3, Nov. 2020, doi: 10.1088/1757-899X/928/3/032076.
- [10] P. Shrivastava, T. Jain, and V. Ben Trivedi, "Structure and genetic relationship of five populations from central India based on 15 autosomal STR loci," *Annals of Human Biology*, vol. 44, no. 1, pp. 74–86, Jan. 2017, doi: 10.3109/03014460.2016.1151932.
- [11] M. S. Anggreainy, M. R. Widyanto, B. H. Widjaja, and N. Soedarsono, "Gaussian Fuzzy Number for STR-DNA Similarity Calculation Involving Familial and Tribal Relationships," *Advances in Bioinformatics*, vol. 2018, pp. 1–13, Jul. 2018, doi: 10.1155/2018/8602513.
- [12] A. Seman, Z. A. Bakar and A. M. Sapawi, "Centre-based clustering for Y-Short Tandem Repeats (Y-STR) as numerical and categorical data," *2010 International Conference on Information Retrieval & Knowledge Management (CAMP)*, Shah Alam, Malaysia, 2010, pp. 28–33, doi: 10.1109/INFRKM.2010.5466953.
- [13] J. Liu *et al.*, "The construction and application of a new 17-plex Y-STR system using universal fluorescent PCR," *International Journal of Legal Medicine*, vol. 134, no. 6, pp. 2015–2027, Nov. 2020, doi: 10.1007/s00414-020-02291-3.
- [14] D. Taylor, J. Curran, and J. Buckleton, "Likelihood ratio development for mixed Y-STR profiles," *Forensic Science International: Genetics*, vol. 35, pp. 82–96, Jul. 2018, doi: 10.1016/j.fsigen.2018.03.006.
- [15] S. M. B. Maiquilla *et al.*, "Y-STR DNA analysis of 154 female child sexual assault cases in the Philippines," *International Journal of Legal Medicine*, vol. 125, no. 6, pp. 817–824, Nov. 2011, doi: 10.1007/s00414-010-0535-4.
- [16] M. S. H. Abu Halima, L. P. Bernal, and F. A. Sharif, "Genetic variation of 15 autosomal short tandem repeat (STR) loci in the Palestinian population of Gaza Strip," *Legal Medicine*, vol. 11, no. 4, pp. 203–204, Jul. 2009, doi: 10.1016/j.legalmed.2009.02.072.
- [17] H. H. de Boer, G. J.R. Maat, D. A. Kadarmo, P. T. Widodo, A. D. Kloosterman, and A.J. Kal, "DNA identification of human remains in Disaster Victim Identification (DVI): An efficient sampling method for muscle, bone, bone marrow and teeth," *Forensic Science International*, vol. 289, pp. 253–259, 2018, DOI: 10.1016/j.forsciint.2018.05.044
- [18] M. Ahmad, L. T. Jung, and M. A.-A. Bhuiyan, "On fuzzy semantic similarity measure for DNA coding," *Computers in Biology and Medicine*, vol. 69, pp. 144–151, Feb. 2016, doi: 10.1016/j.combiomed.2015.12.017.
- [19] M. R. Widyanto, N. Soedarsono, N. Katayama, and M. Nakao, "Various Defuzzification Methods on DNA Similarity Matching Using Fuzzy Inference System," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 14, no. 3, pp. 247–255, Apr. 2010, doi: 10.20965/jaciii.2010.p0247.
- [20] V. K. Malhotra, H. Kaur, and M. Afshar Alam, "An Analysis of Fuzzy Clustering Methods," *International Journal of Computer Applications*, vol. 94, no. 19, pp. 9–12, May 2014, doi: 10.5120/16497-6578.
- [21] X. Yi, Y. Miao, J. Zhou, and Y. Wang, "Some novel inequalities for fuzzy variables on the variance and its rational upper bound," *Journal of Inequalities and Applications*, vol. 2016, no. 1, Dec. 2016, doi: 10.1186/s13660-016-0975-6.
- [22] D. Chandran, K. Crockett, D. Mclean, and Z. Bandar, "FAST: A fuzzy semantic sentence similarity measure," in *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, Jul. 2013, pp. 1–8, doi: 10.1109/FUZZ-IEEE.2013.6622344.
- [23] D. Chandran and K. Crockett, "Fuzzy ontologies in semantic similarity measures," in *2016 IEEE Congress on Evolutionary Computation (CEC)*, IEEE, Jul. 2016, pp. 4942–4949, doi: 10.1109/CEC.2016.7744424.
- [24] M. P. Dewi, A. M. Arymurthy, S. Setiawan and N. Soedarsono, "Human DNA Profile Identification Using DNA Database System," *2022 8th International Conference on Science and Technology (ICST)*, Yogyakarta, Indonesia, 2022, pp. 1-4, doi: 10.1109/ICST56971.2022.10136305.
- [25] D. Xu, J. M. Keller, M. Popescu, and R. Bondugula, *Applications of Fuzzy Logic in Bioinformatics*, vol. 9. in Series on Advances in Bioinformatics and Computational Biology, vol. 9. Published by Imperial College Press and Distributed by World Scient, 2008, doi: 10.1142/p583.

BIOGRAPHIES OF AUTHORS






Meira Parma Dewi    received S.Si from Padjadjaran University, and M.Kom from Universitas Indonesia in 2009. Currently pursuing doctor degree in University of Indonesia. And now still serving as lecturer in Department of Mathematic, Universitas Negeri Padang. She can be contacted at email: meira.parma91@ui.ac.id.






Aniati Murni Arymurthy    received Ir. from Universitas Indonesia, and MSc. from Department of Computer and Information Sciences, The Ohio State University. She received Doctor degree from Universitas Indonesia. Currently she is a Professor and Senior Lecturer at Faculty of Computer Science, Universitas Indonesia. She can be contacted at email: aniati@cs.ui.ac.id.



Suryana Setiawan    received Ir. from Institut Teknologi Bandung, M.Sc from Michigan State University and doctor degree from Universita Di Pisa. Currently he is senior lecturer at Faculty of Computer Science, University of Indonesia. He can be contacted at email: setiawan@cs.ui.ac.id.



Nurtami Soedarsono    received drg. from Faculty of Dentistry Indonesia University and doctor degree from Tokyo Medical and Dental University Japan in 2006. Currently she is a senior lecturer at Department of Oral Biology, Faculty of Dentistry, University of Indonesia. She can be contacted email: nurtami@ui.ac.id.