International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 5 Issue 4, April 2018

ISSN: 2394-4404

Allele Frequency Distribution Of The FGA STR Marker Among The Asante Population Of Ghana

Harriet Biaa Agyekum

Mphil Forensic Science, Department of Biochemistry and Biotechnology, Kwame Nkrumah University of Science and Technology, Ghana

Dr. S. Pandam Salifu

Lecturer, Department of Biochemistry and Biotechnology Kwame Nkrumah University of Science and Technology, Ghana

Abstract: ***Short tandem repeats (STRs) are highly polymorphic loci in the human genome and are used worldwide for forensic identification. STR loci or markers are motifs with 2-6 base pair repeats in the chromosome of an individual. A key problem with STR application is that; in most sub-Saharan African countries of which Ghana forms a part, there is limited knowledge and information concerning allele frequency distributions among different populations. In this study, the allele frequencies of 60 unrelated Asante individuals among the Ghanaian population were analyzed using polymerase chain reaction (PCR) and polyacrylamide gel electrophoresis (PAGE). Statistical parameters of forensic importance such as observed and expected heterozygosity (H_O and H_e), homozygosity and polymorphism information content (PIC) were calculated for the study marker, FGA. The parameters indicated the effectiveness of the markers in forensic study among the Asante population. The PIC, H_e and H_o of the FGA marker was recorded as 0.8580, 0.8704 and 0.3038 respectively. The obtained information on the FGA marker demonstrates that this locus is useful for forensic and identification purposes. This is the first time to establish the frequency distribution of FGA marker of which it is of much importance to the that Asante population and Ghana as a whole.***

I. INTRODUCTION

One of the most significant current discussions in the legal and moral philosophy that has revolutionized in the area of forensics over the past two decades is the discovery of Deoxyribonucleic acid (DNA) typing for human identification (Lynch, Cole, McNally, & Jordan, 2010).

DNA profiling has been used in generating Forensic DNA databases and are therefore now well established in many countries in the world but there is still less knowledge and information concerning these databases in Ghana.

Several methods and techniques including Restriction Fragment Length Polymorphisms (RFLPs), Single Nucleotide Polymorphisms (SNPs) and Short Tandem Repeats (STRs) have been exploited by geneticist and forensic investigators in human identification over the years. However, STRs have been proven to be the most acceptable among the rest and as such, widely used in forensic investigations (Kirby, 1992) since STR sequence data is expected to increase the effective number of alleles in forensic applications which improves the power of discrimination in many cases (Sun et al., 2012). The allelic frequency of the FGA STR markers was therefore assessed among the Asante population of Ghana.

II. MATERIALS AND METHOD

Ethical approval was sought from the KNUST medical school for this research. Human intravenous blood samples were collected from 60 Asante individuals between the ages of 18 and 65 years from the KNUST teaching hospital. Genomic DNA was extracted from each of the blood samples using the BLM method and protocol. The DNA was quantified and amplified using the FGA STR primer sequence**s;** 5'-GCCCCATAGGTTTTGAACTCA-3' and R: 5'-TGATTTGTCTGTAATTGCCAGC-3' for the forward and reverse primers respectively in the PCR reaction. The resulting PCR products were then run on 5% polyacrylamide gel electrophoresis and stained using 0.5ug.mL^{-1} of the ethidium bromide solution. The number of alleles, the allele frequency,

homozygosity ratio, heterozygosity ratio and the polymorphism information content (PIC) for the study marker were manually calculated using Microsoft Excel sheet.

III. RESULTS

Out of the sixty (60) Asante individuals selected for the study, thirty-five (35) were females and twenty-five (25) were males with their ages ranging between eighteen (18) and sixty-five (65) years. After the extraction of DNA from their blood samples and quantifying the extracted DNA, most of the concentration of the DNA extracts ranged between 2 $ng.\mu L^{-1}$ and 15 $ng.\mu L^{-1}$with a few measuring above 20 $ng.\mu L^{-1}$. Calculating the purity from the absorbance ratio, about half the DNA samples ha d their purity ranging between 1.8 and 2.0 with the rest being below 1.8.

Of the 60 samples amplified with FGA STR marker and run using 5% polyacrylamide gel electrophoresis, all samples showed bands, suggesting alleles. The electrophoresis of PCR samples for the population allele frequency study of the FGA markers revealed a range of 1 to 4 alleles per sample analysed (as shown on Plate 1). Over all, a total of 14 allele types were observed as 79 bands (alleles) for the marker. Most of the bands from the gel images were sharp indicating strong amplification with few being faint (weak amplification). The different alleles and STR allelic frequency for the study marker or loci are presented on the Table 1 below, showing the number of alleles for the FGA marker with their respective allele frequencies.

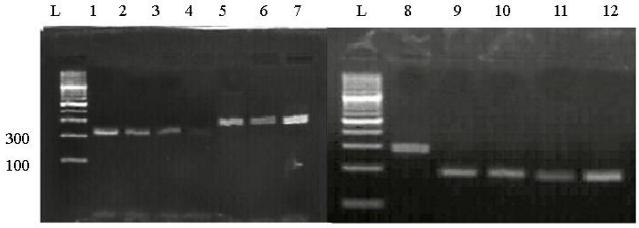

Plate 1: Polyacrylamide Gel Electrophoresis

Diagram showing bands for ladder (L) and PCR products for samples amplified with the FGA marker. These were run on 5% polyacrylamide gels at 100V for 120 mins along with 1kb full scale DNA ladder.

Alleles	No. of bp	No. of Alleles	Repeat Motif	Allele Frequency
43.2	282	3	[TTTC]4TTTT TT [CTTT]8 [CTGT]5 [CTTT]13[CTTC]4[CTTT]3CTCC[TTCC]4	0.037974684
41.2	274	4		0.050632911
39.3	267	9		0.113924051
38	260	12		0.151898734
36.1	253	19		0.240506329
34.2	246	9	[TTTC]4TTTT TT [CTTT]18[CTTC]3[CTTT]3 CTCC[TTCC]4	0.113924051
33.3	243	1	[TTTC]4TTTT TT [CTTT]17[CTTC]3 [CTTT]3 CTCC [TTCC]4.X	0.012658228
33	240	9	[TTTC]4TTTT TT [CTTT]16[CTTC]3 [CTTT]3 CTCC [TTCC]4.X	0.113924051
31.2	234	2	[TTTC]4TTTT TT [CTTT]15[CTTC]3 [CTTT]3CTCC [TTCC]4	0.025316456
30.1	229	1	[TTTC]3TTTT TTCT[CTTT]16 CCTT [CTTT]5CTCC[TTCC]2.X	0.012658228
28.3	223	2	[TTTC]3TTTT TT [CTTT]21CTCC [TTCC]2.X	0.025316456
27.2	218	4	[TTTC]3TTTT TT[CTTT]20CTCC[TTCC]2	0.050632911
26.1	213	3	[TTTC]3TTTT TTCT[CTTT]18CTCC [TTCC]2.X	0.037974684
25.1	209	1	[TTTC]3TTTT TTCT[CTTT]17CTCC [TTCC]2.X	0.012658228
Total number of alleles = 79				
homozygossity =			0.129626663	
Expected heterozygosity=			0.870373337	
Polymorphism Information Content (PIC)=			0.857972028	

Table 1: Allele Frequency Distribution of FGA Loci with repeat motif [CTTT]

Table showing the alleles present for the FGA marker with its respective number of alleles, number of base pairs, repeat motif and frequency distributions. 'X' representing an additional nucleotide C or Tin repeat motif. Both the frequency of each allele at the locus and the frequency of the marker within the study population are indicated (http://www.cybertory.org/resources/CODIS/index.htmlm and http://www.cstl.nist.gov/biotech/strbase/fbicore.htm)

Based on the genetic and forensic efficiency parameters calculated from the allelic frequency of the marker, the FGA marker had its most polymorphic alleles being allele 36.1 which occurred 19 times. The FGA marker had an allele frequency of 0.240506329

The number of heterozygote alleles for the FGA marker was 24 it's observed heterozygosity (Ho) which is the proportion of the number of heterozygotes alleles to the total number of alleles for a particular loci or marker being 0.3037974684.

The expected heterozygosity (He) was recorded as 0.870373337. The homozygosity which is also the sum of squares of all allele frequencies at a locus was also recorded as 0.129626663. Again, the polymorphism information content (PIC) which is the measure of the informativeness of a genetic marker was also calculated as 0.85797208 for the FGA marker as shown in the Appendix 1 and 2.

IV. DISCUSSION

This research was conducted with the objective of assessing the allele distribution of FGA marker within the Asante population of Ghana and to provide information for the

generation of a DNA database for the study population using the study marker.

The isolation of genomic DNA using the BLM protocol in this study yielded good amount of DNA attesting to Butler (2012) which states that for any successful molecular maker analysis, the isolation of high genomic DNA is very essential, hence the quantity or amount of DNA to be used in PCR is important for downstream application in forensic investigation (Butler, 2011).

Majority of the DNA extracts that had their purity ratios ranging between 1.8 to 2.0 with a few being very low or high than the purity range (1.8-2.0) suggests low amounts of DNA extracts in these samples and or possible protein contaminations resulting from the low effect of the activity of proteinase K on the extraction process of these samples. Upon the electrophoresis of the PCR products on the 5% polyacrylamide gel, all 60 PCR product samples used in this study showed various bands indicating the presence of the amplified sites of these individuals based on the FGA forensic marker used.

Although this study was in accordance with a study by White *et al.*, 1997 which stated that most STR markers analyzed as DNA fragments are made up of about 100-350 base pairs as seen in this study, there were few alleles which had their base pairs being below 100 bp and this suggests that within the Asante population, there are alleles of FGA having less than 100 bp which have not been reported in the STRBase as available on http://www.cstl.nist.gov/biotech/strbase/.

Again, in this study, the FGA marker showed higher diversity and polymorphism of the marker as in other studies incorporated into the STRBase. Comparing the forensic efficiency parameters in this study to that of a study population of the Mesan and Basra provinces in Southern Iraq which was also being studied for the first time by Hameed *et al.*, 2015, the homozygosity, heterozygosity and polymorphism information content (PIC) of these markers were almost the same. Within the Asante population, homozygosity, heterozygosity and PIC values of 0.1296, 0.8704 and 0.8580 were obtained for FGA loci whereas the study on these Southern Iraq populations had 0.141, 0.877 and 0.839 for FGA loci respectively. In the Southern Iraq populations, the total number of observed allele types for FGA was 8 (Hameed *et.al.*, 2015) whereas 14 different alleles types were observed for FGA in the Asante population.

Also, a study conducted in Botswana on five sub populations for the first time also reported 12 allele types for FGA. The homozygosity, heterozygosity and polymorphism information content (PIC) for these Botswana populations were 0.141, 0.5859 and 0.84 respectively for FGA (Mpoloka *et.al.*, 2008). This was also in accordance with the Asante population of Ghana and hence making this marker highly polymorphic as in other studies.

Again, the number of allele types observed in this study of the Asante population was in agreement with a study by Kashyap *et.al.*,(2004), which stated that most STR markers generally have about seven (7) to thirty (30) different alleles on the average and are highly polymorphic in various populations (Butler, 2006).

V. CONCLUSION

The findings of the study indicate that the FGA STR marker is important in forensic identification and can be used to estimate genotypic frequencies among the Asante population. The outcome of this study serves as a starting point for the generation of a database for STR markers among the Asante population of Ghana.

APPENDIX

A. FORMULAR FOR THE CALCULATION OF POLYMORPHISM INFORMATION CONTENT (PIC).

PIC =

$$1 - \sum_{i=1}^{n} p_i^2 - \left(\sum_{i=1}^{n} p_i^2\right)^2 + \sum_{i=1}^{n} p_i^4$$

B. CALCULATION OF FORENSIC EFFIECIENCY PARAMETERS FOR FGA MARKER.

Alleles	freq.	Freq.2	Freq.^{4}
25.1	0.01265823	0.000160231	2.56739E-08
26.1	0.0379747	0.001442077	2.07958E-06
27.2	0.0506329	0.002563692	6.57252E-06
28.3	0.0253165	0.000640923	4.10782E-07
30.1	0.0126582	0.000160231	2.56739E-08
31.2	0.0253165	0.000640923	4.10782E-07
33	0.1139241	0.012978689	0.000168446
33.3	0.0126582	0.000160231	2.56739E-08
34.2	0.1139241	0.012978689	0.000168446
36.1	0.2405063	0.057843294	0.003345847
38	0.1518987	0.023073225	0.000532374
39.3	0.1139241	0.012978689	0.000168446
41.2	0.0506329	0.002563692	6.57252E-06
43.2	0.0379747	0.001442077	2.07958E-06
Total	1	0.129626663	0.004401762
(SUM of Freq.2)2 =		0.016803072	
Homozygossity of FGA=		0.129626663	
Expected Heterozygosity=		0.870373337	
Polymorphism Information Content (PIC)of FGA =		0.857972028	

REFERENCES

[1] Butler, J. M. (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. Journal of Forensic Sciences, 51(2), 253–265. http://doi.org/10.1111/j.1556-4029.2006.00046.x

[2] Butler, J. M. (2011). Advanced Topics in Forensic DNA Typing: Methodology: Methodology. Academic Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=44tX9Ole490C&pgis=1

ISSN: 2394-4404

[3] Butler, J. M. (2012). Advanced Topics in Forensic DNA Typing. Advanced Topics in Forensic DNA Typing, 271–292. http://doi.org/10.1016/B978-0-12-374513-2.00009-9
[4] Hameed, I. H., Ommer, A. J., Murad, A. F., & Mohammed, G. J. (2015). Allele frequency data of 21 autosomal Short Tandem Repeat loci in Mesan and Basra provinces in south of Iraq. Egyptian Journal of Forensic Sciences. http://doi.org/10.1016/j.ejfs.2014.10.003
[5] Kashyap, V. K., Sitalaximi, T., Chattopadhyay, P., & Trivedi, R. (2004). DNA Profiling Technologies in Forensic Analysis, 4(1), 11–30.
[6] Kirby, L. T. (1992). DNA fingerprinting. WH Freeman.
[7] Lynch, M., Cole, S. A., McNally, R., & Jordan, K. (2010). Truth Machine: The Contentious History of DNA Fingerprinting. University of Chicago Press. Retrieved from https://books.google.com/books?id=EVmr3kyB7ewC&pgis=1
[8] Mpoloka, S. W., Kgotlele, T., & Wally, A. (2008). Determination of allele frequencies in nine short tandem repeat loci of five human sub-populations in Botswana. African Journal of Biotechnology, 7(8).
[9] Sun, J. X., Helgason, A., Masson, G., Ebenesersdóttir, S. S., Li, H., Mallick, S., … others. (2012). A direct characterization of human mutation based on microsatellites. Nature Genetics, 44(10), 1161–1165.
[10] White, H. W., Kusukawa, N., & Bioproducts, F. M. C. (1997). Agarose-Based System for Separation of Short Tandem Repeat Loci, 22(5), 976–980.