UMYU Scientifica, Vol. 3 NO. 1, March 2024, Pp 103 – 112
https://doi.org/10.56919/usci.2431.012
ISSN: 2955 – 1145 (print); 2955 – 1153 (online)
A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina

ORIGINAL RESEARCH ARTICLE

Forensic parameters of 10 Y-STR markers in the Nigerian Hausa, Igbo, and Yoruba populations

Samson Taiwo Fakorede[1,2]*, Khalid Olajide Adekoya[1], Onyekachi Ogbonnaya Iroanya[1], Mohaimin Kasu[2] and Maria Eugenia D'Amato[2].

[1]Department of Cell Biology and Genetics, Faculty of Science, University of Lagos, Lagos, Nigeria.
[2]Forensic DNA Laboratory, Department of Biotechnology, Faculty of Natural Sciences, University of the Western Cape, Bellville 7535, South Africa.

ABSTRACT

Y chromosome short tandem repeats, or Y-STRs, are interesting genetic markers suitable for identifying male offenders in forensic investigations because of the haploid nature of the human male determining chromosome. Furthermore, due to the non-recombining nature of the Y chromosome, Y-STRs are choice markers in paternity or kinship analysis, human evolutionary investigations, and population studies. The lack of reference Y-haplotype data for the populace currently hampers the use of these markers in forensic applications in Nigeria. In this study, the 10 Y-STR loci in the UniQTyper™ Y-10 system including DYS385ab, DYS447, DYS449, DYS481, DYS504, DYS518, DYS612, DYS626, DYS644 and DYS710 were studied in 461 males comprising 139 Hausa, 96 Igbo and 226 Yoruba ethnic populations. Sixty alleles were recorded for the pooled dataset across the 10 loci, with allele frequencies ranging from 0.0022 to 0.6052. There were 430 haplotypes detected, 403 unique (singletons), and 27 were shared. The discrimination capacity and haplotype diversity were 0.9330 and 0.9770, respectively. Allelic richness was highest in the locus DYS626 and lowest in DYS504. Private alleles were found in each group, with the Hausa population having the highest number. Duplications and microvariant alleles were also reported. These findings may be helpful for paternal lineage investigations, population genetics, and forensic applications among the Nigerian populations.

ARTICLE HISTORY

Received December 22, 2023.
Accepted March 01, 2024.
Published March 10, 2024.

KEYWORDS

Y chromosome STRs, discrimination capacity, haplotypes, private alleles, Nigeria

INTRODUCTION

Humans vary greatly from one another. These individual differences aid in distinguishing an individual or a member of a particular geographic place from another. Individuals' genetic makeup (genotype), physical characteristics (phenotype), and heritable characters vary. At the genetic level, the DNA is almost identical in all humans, with 99.9% of the DNA sequence having the same arrangement (Chakravarti, 2015; Duello *et al.*, 2021). Despite this overwhelming sequence similarity, there are regions on the human DNA molecule specific to each individual. These unique regions are at the core of DNA testing. They are commonly used in forensic genetics and genealogical research to analyze paternal lineage.

Short tandem repeats, also called microsatellites, are repetitive regions in the genome of tandemly repeated sequences of about 2-7 nucleotides in length (Fan and Chu, 2007). They are ideal genetic markers in forensic and population diversity research due to their relative abundance in the genome, high polymorphism, codominance, small length, and multi-allelism (Gharesouran *et al.*, 2021). The non-recombining region of the Y chromosome (NRY) is home to a unique type of tandem repeat known as Y chromosome short tandem repeats (Y-STRs), which are passed on directly and unaltered from father to son barring mutation (Jobling and Tyler-Smith, 2003; Batini *et al.*, 2015).

Y-STRs are valuable tools for biological evidence analysis because they help overcome the difficulties associated with autosomal STR profiling (Dooley, 2022). They assist in resolving the problem of preferential amplification of female DNA in mixtures including male DNA samples, which occurs when the victim's DNA is far more abundant than any traces of male DNA from sperm cells in male-on-female sexual assault cases (Roewer, 2019).

Correspondence: Samson Taiwo Fakorede. Department of Cell Biology and Genetics, Faculty of Science, University of Lagos, Lagos, Nigeria. ✉ tfakorede@unilag.edu.ng. Phone Number: +234 703 556 9436.
How to cite: Fakorede, S. T., Adekoya, K. O., Iroanya, O. O., Kasu, M., & D'Amato, M. E. (2024). Forensic parameters of 10 Y-STR markers in the Nigerian Hausa, Igbo, and Yoruba populations. *UMYU Scientifica*, 3(1), 103 – 112. https://doi.org/10.56919/usci.2431.012

The Y chromosome, which specifies maleness in humans by dictating the development of the testis early in embryogenesis, is found exclusively in males. As a result, the chromosome has certain distinct features that have drawn the interest of researchers interested in understanding human genetic and evolutionary history (Campbell and Tishkoff, 2010).

Because the Y chromosome does not partake in recombination, it is a gold standard in male individualization (Kayser, 2007; Syndercombe-Court, 2021). To do this, however, comprehensive population-specific Y-haplotype data are required. A growing number studies conducted in recent years have reported Y-haplotype data from different Nigerian populations using autosomal and Y chromosome short tandem repeat markers (Veeramah *et al.*, 2010; de Filippo *et al.*, 2011; Purps *et al.*, 2014; Willems *et al.*, 2014; Martinez *et al.*, 2017; Okolie *et al.*, 2018; Fakorede *et al.*, 2019; Akpan, 2022). However, some of these studies did not either take into account the heterogeneity of the Nigerian populations in their sampling, or the markers employed have low discriminating capacity in African populations. Additionally, not much research has been done on Nigerian populations using newer, highly polymorphic Y-STR markers. Hence, the present study was aimed at generating Y-STR haplotypes suitable for forensic and population analyses among Nigerian males using 10 Y-STR loci in order to gain more understanding of Y-STR landscape in the Nigerian diverse populations. This study utilized the UniQTyper™ Y-10 genotyping kit with improved discrimination capacities in African populations (D'Amato *et al.*, 2011; Lesaoana *et al.*, 2019; Kasu *et al.*, 2022).

MATERIALS AND METHODS

Sample collection and Ethical Statement.

The research employed a cross-sectional sampling method among unrelated males belonging to the three major ethnic groups in Nigeria, i.e., Hausa, Igbo, and Yoruba. Samples were collected in areas/towns with a high population density of each ethnic group. Participants' populations were spread across the six geographical regions (Figure 1) with varying numbers of sampled individuals per geographical location. Saliva samples were obtained from participants and preserved with a storage buffer at room temperature (Burrows *et al.*, 2019). All samples, comprising 139 Hausa, 96 Igbo, and 226 Yoruba populations, were collected with written informed consent and anonymized. The Health Research Ethics Committee of the College of Medicine of the University of Lagos (CMUL/HREC/11/17/307) and the Senate Research Committee of the University of the Western Cape (BM/16/3/18) approved the study. The study complied with international ethics standards in D'Amato *et al.* (2020).

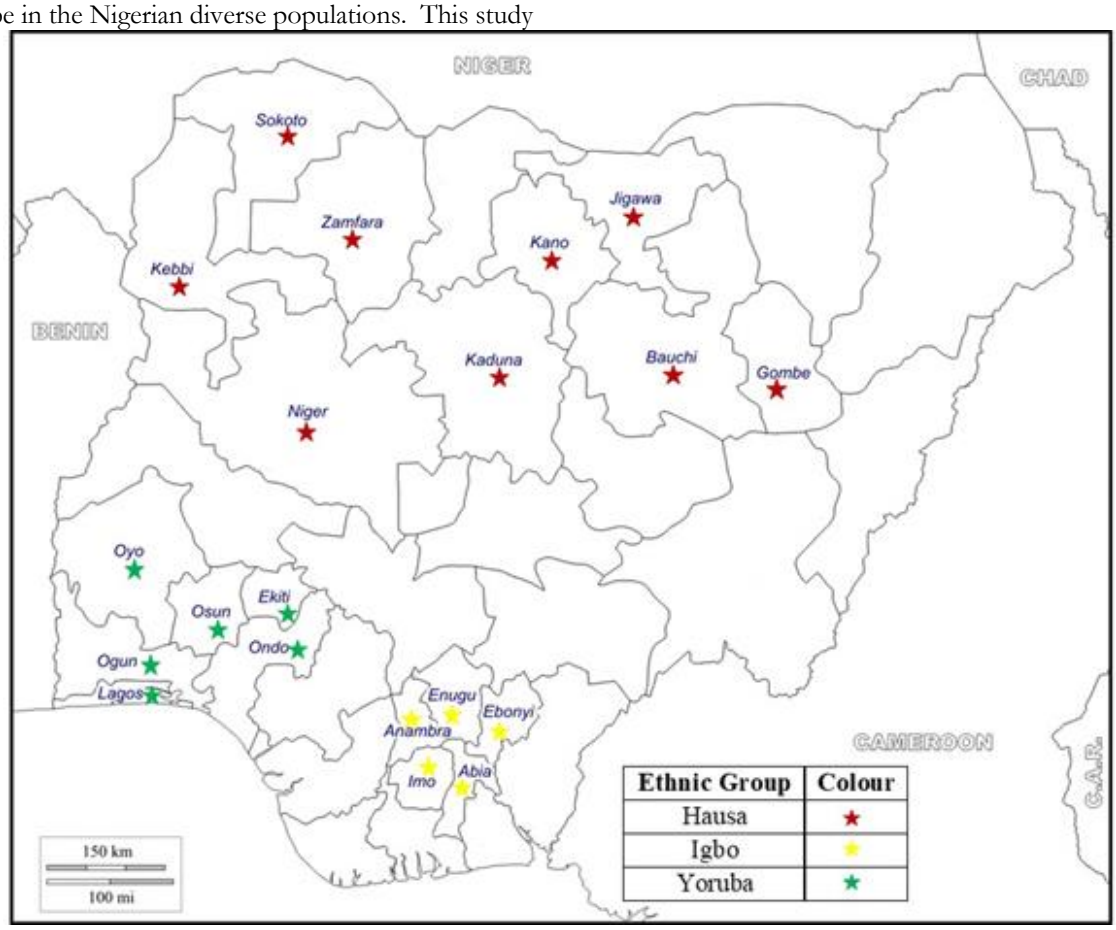

Figure 1: Map of Nigeria showing sampling geographical areas of the study populations.

DNA extraction and quantification

Genomic DNA was isolated from whole saliva samples using the salting out procedure (Medrano *et al.*, 1990). This technique utilizes high salt concentration to precipitate proteins from the DNA, and the DNA is recovered with ethanol precipitation. The DNA samples were quantified with the Nanodrop 2000 spectrophotometer (Thermo Fisher *Scientific*, MA, USA), averaging 130 ng/µL. Each DNA sample was diluted to 2 ng/µL for the multiplex PCR amplification.

Y-STR genotyping

Multiplex PCR amplification was performed using the UniQTyper™ Y-10 system (D'Amato *et al.*, 2011). The system contains 10 Y-STR loci (DYS385ab, DYS447, DYS449, DYS481, DYS504, DYS518, DYS612, DYS626, DYS644 and DYS710) that showed higher discrimination capacity and haplotype diversity than all known commercial systems in baseline South African data including the Xhosa, European, and Indian populations (D'Amato *et al.*, 2011). Genotyping and detection of Y-STR alleles were done at the Forensic DNA Laboratory of the Department of Biotechnology, University of the Western Cape Town, South Africa, following the Kasu *et al.* (2022) procedure. Capillary electrophoresis was performed using ABI 3500 Genetic Analyzer (Thermo Fisher *Scientific*, MA, USA) and analyzed with Genemapper® ID Software.

Statistical analysis

A total of 461 haplotypes were generated and considered for forensic statistics analysis. For frequency analysis, duplications at some loci were represented with fantasy numbers. Allele frequencies, number of alleles (Na), and the polymorphic information content (PIC) were computed using an online tool, STR Analysis for Forensics (STRAF 2.1.5) (Gouy and Zieger, 2017) based on direct counting methods. Haplotype diversity was calculated using the formula: $HD = n(1 - \sum pi^2)/(n - 1)$, where n = sample size, and pi = the frequency of the i-th haplotype (Nei and Tajima, 1981). Discrimination capacity (DC) was calculated as the ratio $\frac{h}{n}$ where h = the total number of distinct haplotypes, and n = the total number of individuals/samples. The allelic richness (Ar) of the 10 markers was calculated using Fstat version 2.9.4 (Goudet, 2003).

RESULTS

Types and frequencies of the 10 Y-STR markers alleles

Sixty different alleles, with sizes ranging between 9 to 46, were recorded across the 10 loci for the entire Nigerian dataset (Table 1). The locus DYS626 was the most diverse in the Nigerian population, with 18 alleles, followed by DYS710 with 17 alleles, while DYS504 had the lowest number of alleles (n=9). Microvariant alleles were observed at DYS710, DYS385ab, DYS644, and DYS626 distributed across the three ethnic groups in varying proportions. The locus DYS626 showed seven distinct duplicated profiles. These were found across the Hausa (n = 7), Igbo (n = 1) and Yoruba (n = 1) ethnic groups. Two microvariant alleles (16.2 and 14.2) were also observed at the locus DYS385b. Bimodal allele frequency was observed among locus DYS644 alleles comprising integers and .4 partial repeat alleles. Allele frequencies in the combined population ranged from 0.0022 to 0.6052.

One pattern of distribution was seen in DYS710 alleles among the three populations, with allele 33.2 being the most common in the Yoruba with an allele frequency of 0.389, Igbo (0.344) and Yoruba (0.252) populations, followed by allele 32.2 in the same order. For the DYS518 locus, two patterns of allelic distribution were observed. The Hausa and Igbo populations had allele 38 as the most common, followed by allele 39, whereas allele 39 was the most common among the Yorubas, followed by allele 40. The DYS504 allele 13 was the most prevalent among the three ethnic populations, including Hausa (0.331), Igbo (0.750), and Yoruba (0.712). Among the DYS481 alleles, 25 was the most common among the Igbo and Yoruba populations, and allele 22 was predominant among the Hausas. Allele 27 was largely observed at locus DYS447 among the Hausa and Igbo populations, while the most common allele for this locus in the Yoruba population was the allele 25. DYS449 had alleles 28, 29, and 31 as the most common among the Yoruba, Hausa, and Igbo populations, respectively.

A total of fifty-one (51) different alleles were observed among the 139 males of the Hausa population genotyped in this study. These allele sizes ranged from 12 (DYS504) to 46 (DYS518). DYS504 and DYS626, with 6 and 16 different alleles, represent loci with the lowest and highest number of alleles, while the average number of alleles observed in the Hausa population was 11±3.0. A total of forty-six (46) different alleles of sizes ranging from 9 in DYS504 to 46 in DYS518 were observed, distributed among the 96 male individuals that were successfully genotyped from the Igbo population. The locus with the lowest number of alleles (6) was DYS504 followed by DYS447 with 7 alleles. The loci DYS518, DYS644, and DYS612 have 10 alleles each, while four loci (DYS710, DYS626, DYS481, and DYS449) displayed the highest number of alleles with 11 each. For the Yoruba population (n = 226), forty-six (46) unique alleles were identified across the 10 Y-STR loci evaluated in this study. The allele sizes ranged from 10 (DYS504) to 45 (DYS518). The locus with the lowest number of alleles was DYS504 (7), while locus DYS710 had the highest number of alleles totaling 15.

Table 1: Allele frequencies of Y-STR loci in Nigerian Hausa, Igbo, and Yoruba males (n = 461)

Allele	DYS 710	DYS 518	DYS 644	DYS 612	DYS 626	DYS 504	DYS 481	DYS 447	DYS 449
9						0.002			
10						0.002			
11						0.002			
12						0.048			
13						0.605			
14						0.184			
15			0.033			0.106			
16			0.091			0.035			
17						0.015		0.004	
17.3			0.002						
18			0.004						
19			0.002				0.004		
20							0.002		
20.4			0.033						
21							0.011	0.002	
21.4			0.399						
22							0.091	0.007	
22.3			0.002						
22.4			0.336						
23							0.048	0.061	
23.4			0.069						
24							0.085	0.043	
24.3			0.002						
24.4			0.017						
25					0.017		0.282	0.291	0.011
25.4			0.004						
26				0.002	0.004		0.130	0.206	0.009
26.4			0.004						
27					0.015		0.130	0.332	0.050
28				0.002	0.056		0.161	0.052	0.148
28.2	0.007								
29				0.002	0.221		0.041	0.002	0.141
29.2	0.024								
30	0.007				0.275		0.011		0.169
30.2	0.041				0.002				
31	0.015			0.011	0.228		0.004		0.145
31.2	0.089								
32	0.048			0.024	0.106				0.128
32.2	0.191								
33	0.037			0.119	0.043				0.108
33.2	0.338								
34	0.033			0.178	0.011				0.067
34.2	0.111								
35	0.024	0.011		0.187					0.007
35.2	0.026								

Table 1 Continued

36	0.004	0.013		0.208					0.011
36.2	0.004								
37		0.082		0.111					0.007
38		0.197		0.085					
39		0.254		0.056					
39.2	0.002								
40		0.217		0.013					
41		0.115		0.002					
42		0.061							
43		0.033							
44		0.011							
45		0.002							
46		0.004							
31, 33*					0.007				
27, 29*					0.002				
28, 31*					0.002				
29, 32*					0.002				
29, 31*					0.002				
29, 33*					0.002				
29, 30*					0.002				
Na	17	12	14	14	18	9	13	10	13
Ne	5.6	5.7	3.5	6.8	5.2	2.4	6.3	4.1	7.9
He	0.820	0.824	0.712	0.853	0.806	0.585	0.841	0.755	0.873
PIC	0.803	0.802	0.667	0.835	0.780	0.548	0.824	0.715	0.860

Na = Number of alleles, Ne = Number of effective alleles, He: Expected heterozygosity, PIC: Polymorphic information content. *Duplicated profiles at the DYS626 locus

DYS385ab haplotypes

The multicopy locus DYS385, comprising DYS385a and DYS385b, was treated as a single marker (i.e., DYS385ab), and the allele frequencies were combined as haplotypes, as shown in Figure 2. 43 allelic combinations were observed at this locus out of which 16,17 for the entire Nigerian dataset. The haplotype 16,17 was found in 5.8% of the sampled Yoruba population, 18.8% in Hausa, and 18.6% among the Igbos. The most common haplotype among the Yoruba population for the locus DYS385ab was 16,17 (repeated 42 times), while the Hausa and Igbo populations had predominantly haplotypes 13,15 (repeated 20 times) and 17,18 (repeated 19 times), respectively as the most common among them. At the locus DYS385ab, 30 allelic combinations with 11 alleles were observed in the Hausa population. The Igbo population had 23 allelic combinations with 11 different alleles, while 30 allelic combinations with 12 separate alleles were observed in the Yoruba population. Nine (9) pseudo-homozygous haplotypes ranging from 11,11 to 19,19 were observed across the combined dataset, with haplotype 17,17 being the most prevalent with a frequency of 0.067.

Private alleles

Private alleles (PAs), which signify alleles that are unique or exclusive to a specific population, could measure genetic diversity between populations. This is because private alleles reduce the number of shared haplotypes within and between populations, thereby increasing the level of diversity in such population(s) and the overall discrimination potential of forensic testing kits.

Private alleles were observed in each of the studied populations. A list of the PAs for the ethnic groups and their respective frequencies is presented in Table 2. The Hausa population had the highest number of PAs at 12, while Igbo and Yoruba populations had 3 and 9, respectively, with 24 PAs recorded for the studied populations. However, the distribution of the alleles in a population may be a function of the sample size.

Haplotype diversity and discrimination capacity

The overall haplotype diversity (HD) for the Nigerian population was 0.977, while the HDs for the Hausa, Igbo, and Yoruba populations were 0.998, 0.999, and 1.000, respectively (Figure 3). However, the discrimination capacity (DC) ranged from 90.6% in Hausa samples to 95.8%, with an overall DC of 93.3%

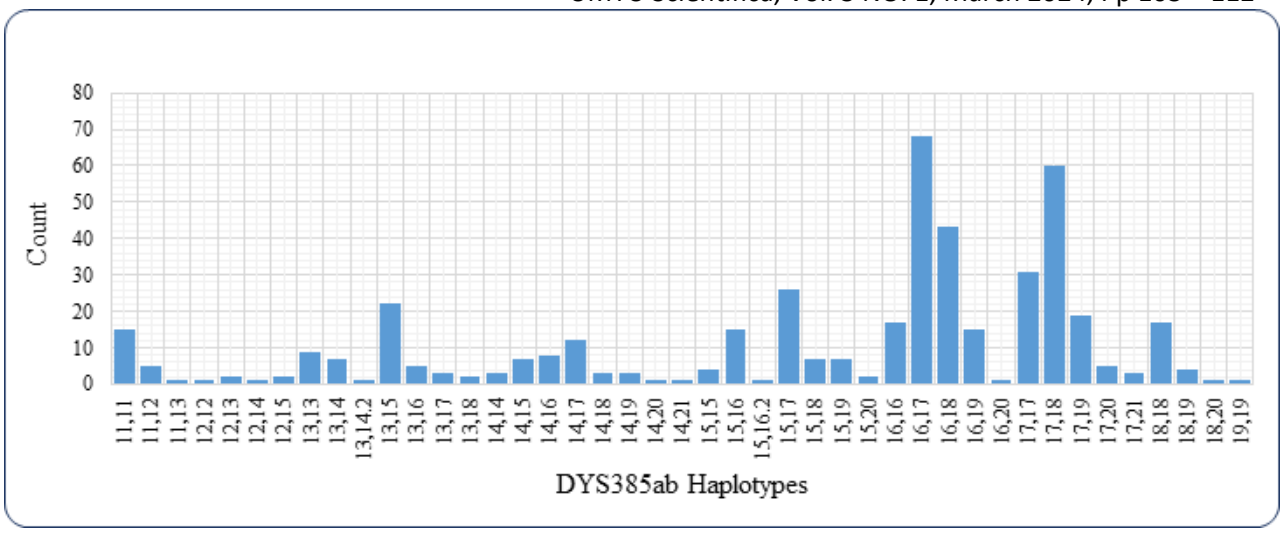

Figure 2: Haplotype count for the locus DYS385ab across the three studied populations

Table 2: List of private alleles for the Hausa, Igbo, and Yoruba populations

Population	Locus	Allele	Frequency
Hausa (n = 139)	DYS710	39	0.007
	DYS518	35	0.036
	DYS385b	12	0.043
	DYS644	17	0.007
	DYS644	26	0.014
	DYS612	28	0.007
	DYS612	41	0.007
	DYS504	17	0.050
	DYS481	19	0.014
	DYS481	20	0.007
	DYS447	17	0.014
	DYS449	35	0.022
Igbo (n = 96)	DYS385a	19	0.010
	DYS644	19	0.010
	DYS504	9	0.010
Yoruba (n = 226)	DYS710	36	0.018
	DYS518	45	0.004
	DYS612	26	0.004
	DYS612	29	0.004
	DYS504	10	0.004
	DYS504	11	0.004
	DYS447	21	0.004
	DYS447	29	0.004
	DYS449	26	0.018

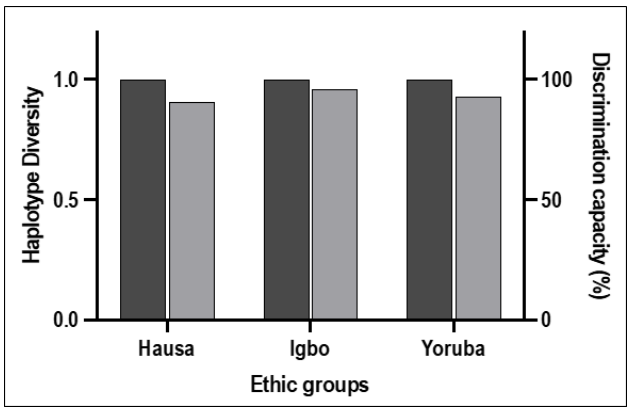

Figure 3: The haplotype diversity and discriminatory capacity for the Hausa, Igbo, and Yoruba ethnic groups

Gene diversities and allelic richness

Figure 4 depicts the graphical distinction of gene diversities (GD) of the studied ethnic groups. The GD values for the Hausa population ranged from 0.769 to 0.947, with an average of 0.850 representing the highest average GD value for the three populations. This establishes the heterogeneity of the Hausa population. The Igbo population had an average GD value of 0.757 with GDs between 0.419 and 0.909; for the Yorubas, the GDs ranged from 0.446 to 0.908 with an average value of 0.757. Mann-Whitney U test showed significant differences in the gene diversities of the Hausas against those of the Igbo ($z = -2.165$, $p = 0.030$) and the Yoruba ($z = -2.166$, $p = 0.030$) populations. However, no

significant difference was found between the southern populations of Igbo and Yoruba ($z = -0.044$, $p = 0.965$). The locus DYS385ab had the highest GD value (0.947) reported among the Hausas, while the Igbo and Yoruba populations had GD values of 0.909 and 0.908, respectively, for this locus.

All ten loci's GD values were higher than 0.5 except for the locus DYS504. The rapidly mutating loci DYS518, DYS612, DYS626, and DYS449 all have average GD values greater than 0.8 across the three populations. Allelic richness was highest in the locus DYS626 among the Hausa samples (Ar = 15.2), followed by DYS710 in both Yoruba and Igbo populations, with values of 14.0 and 11.0, respectively.

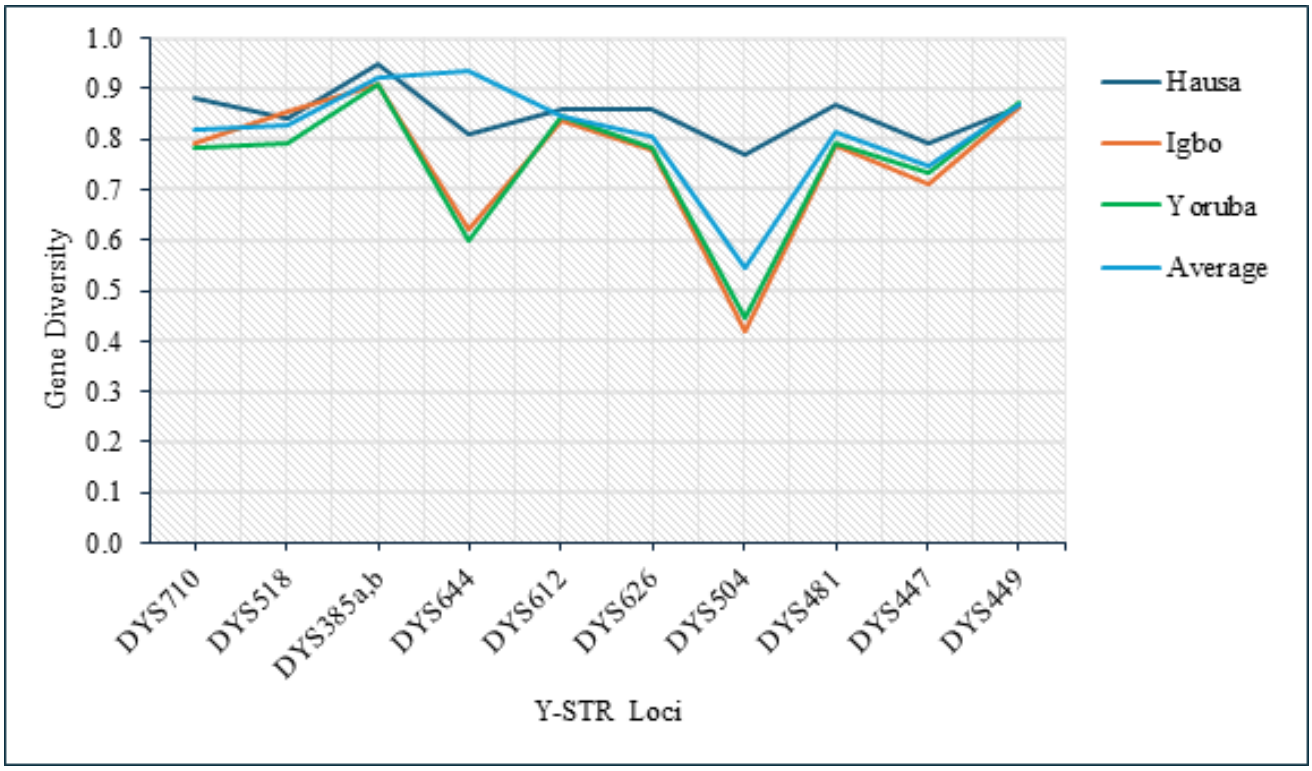

Figure 4: The gene diversity values at each locus for each population

Table 3: Allelic richness of the 10 Y-STR markers per locus and population

	Population			
Locus	Hausa	Igbo	Yoruba	Overall
DYS710	13.9	11.0	14.0	15.0
DYS518	10.8	10.0	8.3	10.7
DYS385a	7.0	9.0	6.7	8.3
DYS385b	10.8	9.0	9.2	11.5
DYS644	10.6	10.0	7.3	10.3
DYS612	11.8	10.0	10.9	11.3
DYS626	15.2	11.0	9.0	12.8
DYS504	6.0	6.0	6.3	7.1
DYS481	12.8	11.0	8.0	11.4
DYS447	7.9	7.0	8.0	8.1
DYS449	11.8	11.0	10.6	12.2

DISCUSSION

The present study evaluated the genetic diversity of the heterogeneous, multi-ethnic, and multicultural Nigerian populations. Robust haplotype data and genetic diversity evaluation from different populations are important for accurately estimating Y-chromosomal haplotype random match probability in investigating questioned samples. Such data are also expedient for biogeographic studies and familial searches. Likewise, because Y chromosome haplotypes exhibit varying degrees of frequency in different populations, there is a need to determine the population-specific Y-haplotypes distribution to create a Y-STR reference database. However, the various ethnic groups within a population must be well-represented in the reference database (Lewontin and Hartl, 1991).

Hammer and Redd (2006) posited that it is possible to underestimate the frequency of a Y-STR haplotype of an individual whose ethnic group is not represented in a database. Thus, the proportion of variation within and among the Nigerian ethnic groups viz-a-viz, the Hausa, Igbo, and Yoruba ethnicities needs to be evaluated for population differentiation and to avoid false positive or negative results in forensic DNA analysis. Additionally, it must be noted that the human Y chromosome is strictly inherited uniparentally from the father. These features make the Y chromosome enormously prone to genetic drift and a good genetic marker for studying human

evolution (Jobling and Tyler-Smith, 2003; Chiaroni *et al.*, 2009). The study generated Nigeria's largest known Y-STR dataset using highly polymorphic markers.

The study employed the UniQ-Typer™ prototype kit to determine the forensic value and reveal the population stratification within the Nigerian Hausa, Igbo, and Yoruba ethnic populations. A high level of allele diversity was observed among the studied populations as evident from the allelic richness (Ar) and haplotype and gene diversity values. These results agree with those obtained for Southern African populations (Lesaoana *et al.*, 2019; Kasu *et al.*, 2022) who employed the same genotyping kit. The overall DC was lower compared to what was obtained among 142 individuals belonging to the Hausa, Igbo, and Yoruba ethnic groups by Martinez *et al.* (2017) with Yfiler Plus (0.9998), a panel with 27 Y-STRs of which seven are rapidly mutating. Furthermore, an estimated DC of 0.9630 was recorded in 81 Yoruba individuals (Ibadan, Nigeria) using the PowerPlex® Y23 System (Purps *et al.*, 2014).

The Hausas harbor higher levels of allele richness or diversity than the other groups, which may be attributed to admixture, higher lineage diversity, marriage practices, or historical demographic processes. Allelic richness estimation is one of the various criteria for assessing within- and among-population genetic diversity and may also be used to estimate population evolutionary histories (Castric and Bernatchez, 2003; Ollivier and Foulley, 2005). However, the lowest haplotype diversity and DC were recorded among the Hausas. These are potentially attributable to factors such as effective population size and marriage practice (Greenberg, 1947).

The Hausas, being predominantly Muslims, practice polygyny, and consanguineous marriage is especially allowed (Swanson and Lagace, 1997). In addition, the Hausa households are patrilineal and patrilocal (Usman, 1997). Reduced haplotype diversities for Y-STR markers have previously been reported in samples from rural villages in Austria (Niederstätter *et al.*, 2016), central Sahel – northern Cameroon, and western Chad (Della Rocca *et al.*, 2020). The incidences of reduced HD and DC of Y-STR markers, especially among African and tribal populations, have been attributed to patrilocality, ethnic fragmentation, and genetic isolation (Khubrani *et al.*, 2018; D'Atanasio *et al.*, 2019; Della Rocca *et al.*, 2020; Kasu *et al.*, 2022).

Duplications at the DYS626 locus are rare. The reported duplications at the DYS626 locus are unique and have not been encountered in previous studies with the same genotyping system in Lesotho (Lesaoana *et al.*, 2019) and South African populations (Kasu *et al.*, 2022). Ballantyne *et al.* (2014) reported different duplications for this locus among Sub-Saharan (30, 32 and 31, 33), North African (31, 32), admixed native American (30, 32), and European (29, 31) populations. Turrina *et al.* (2015) reported simultaneous deletions and duplications (28, 32) in the DYS626 locus within the *AZFc* region of the Y chromosome in an Italian subject. Allelic variations such as duplications (and triplications), null and microvariant alleles are a major source of intra- and intra-specific variation in populations (Budowle *et al.*, 2008). Hence, this unique variation among the Nigerian population is promising for forensic applications.

CONCLUSIONS

This study generated 10 Y-STR allele frequency data for 461 unrelated males of the Hausa, Igbo, and Yoruba ethnic populations of Nigeria. The study identified private alleles and novel duplications among the sampled individuals which are forensically significant for the Nigerian populations. Results obtained from the study showed that the 10 Y-STR markers are highly polymorphic, informative, and suitable for forensic investigation, human identification and genetic diversity studies in the Nigerian populations. It is recommended that the forensic data generated from the study are utilized in crime investigation, especially as the country grapples with insecurity and plethora cases of gender-based violence. This requires collaboration between researchers, law enforcement agencies, and ethical bodies to ensure informed consent and data privacy.

DATA AVAILABILITY

The individual profiles are available at https://zenodo.org/records/8376039

FUNDING

This study was financially supported by the University of Lagos Central Research Committee grant (CRC No. 2018/18) awarded to OOI and KOA; the African German Network of Excellence in Science (AGNES), through the Programme AGNES Intra-Africa Mobility Grant for Junior Researchers (2021) as well as the University of Lagos Doctoral Assistance Grant awarded to STF; and the University of the Western Cape Senate Research funds awarded to MED.

REFERENCES

Akpan, U. U. (2022). Multiplex PCR and forensic potentials of nine short tandem repeats in the Hausa, Yoruba, Igbo, Ibibio and Tiv populations of Nigeria. PhD Thesis. University of Lagos, Nigeria.

Ballantyne, K. N., Ralf, A., Aboukhalid, R., Achakzai, N. M., Anjos, M. J., Ayub, Q., Balažic, J., Ballantyne, J., Ballard, J. D., Berger, B., Bobillo, C., Bouabdellah, M., Burri, H., Capal, T., Caratti, S., Cárdenas, J., Cartault, F., Carvalho, E. F., Carvalho, M., … Kayser, M. (2014). Toward male individualization with rapidly mutating y-chromosomal short tandem repeats. *Human Mutation, 35*(8), 1021-1032. **[Crossref]**

Batini, C., Hallast, P., Zadik, D., Delser, P. M., Benazzo, A., Ghirotto, S., Arroyo-Pardo, E., Cavalleri, G. L., De Knijff, P., Dupuy, B. M., & Eriksen, H. A. (2015). Large-scale recent expansion of European patrilineages shown by population resequencing. *Nature Communications, 6*(1), 1-8. **[Crossref]**

Budowle, B., Aranda, X. G., Lagace, R. E., Hennessy, L. K., Planz, J. V., Rodriguez, M., & Eisenberg, A. J. (2008). Null allele sequence structure at the DYS448 locus and implications for profile interpretation. *International Journal of Legal Medicine*, *122*, 421-427. **[Crossref]**

Burrows, A. M. Kasu, M., & D'Amato, M. E. (2019). Preservation of DNA integrity in biological material. *Forensic Science International Genetics: Supplemental Series, 7*(1), 416-418. **[Crossref]**

Campbell, M. C., & Tishkoff, S. A. (2010). The evolution of human genetic and phenotypic variation in Africa. *Current Biology, 20*(4), 166-173. **[Crossref]**

Castric, V., & Bernatchez, L. (2003). The rise and fall of isolation by distance in the anadromous brook charr (*Salvelinus fontinalis* Mitchill). *Genetics*, *163*(3), 983-996. **[Crossref]**

Chakravarti, A. (2015). Perspectives on human variation through the lens of diversity and race. *Cold Spring Harbor Perspectives in Biology*, *7*(9), a023358. **[Crossref]**

Chiaroni, J., Underhill, P. A., & Cavalli-Sforza, L. L. (2009). Y chromosome diversity, human expansion, drift, and cultural evolution. *Proceedings of the National Academy of Sciences*, *106*(48), 20174-20179. **[Crossref]**

D'Amato, M. E., Bajic, V. B., & Davison, S. (2011). Design and validation of a highly discriminatory 10-locus Y-chromosome STR multiplex system. *Forensic Science International: Genetics*, *5*(2), 122-125. **[Crossref]**

D'Amato, M. E., Bodner, M., Butler, J. M., Gusmão, L., Linacre, A., Parson, W., ... & Carracedo, A. (2020). Ethical publication of research on genetics and genomics of biological material: guidelines and recommendations. *Forensic Science International: Genetics*, *48*, 102299. **[Crossref]**

D'Atanasio, E., Iacovacci, G., Pistillo, R., Bonito, M., Dugoujon, J. M., Moral, P., ... & Cruciani, F. (2019). Rapidly mutating Y-STRs in rapidly expanding populations: Discrimination power of the Yfiler Plus multiplex in northern Africa. *Forensic Science International: Genetics*, *38*, 185-194. **[Crossref]**

de Filippo, C., Barbieri, C., Whitten, M., Mpoloka, S. W., Gunnarsdóttir, E. D., Bostoen, K., ... & Pakendorf, B. (2011). Y-chromosomal variation in sub-Saharan Africa: insights into the history of Niger-Congo groups. *Molecular Biology and Evolution*, *28*(3), 1255-1269. **[Crossref]**

Della Rocca, C., Cannone, F., D'Atanasio, E., Bonito, M., Anagnostou, P., Russo, G., ... & Cruciani, F. (2020). Ethnic fragmentation and degree of urbanization strongly affect the discrimination power of Y-STR haplotypes in central Sahel. *Forensic Science International: Genetics*, *49*, 102374. **[Crossref]**

Dooley, K. B., Madisha, M. T., Strümpher, S., & Ehlers, K. (2022). Forensic genetic value of 27 Y-STR loci (Y-Filer® Plus) in the South African population. *Science & Justice*, *62*(3), 358-364. **[Crossref]**

Duello, T. M., Rivedal, S., Wickland, C., & Weller, A. (2021). Race and genetics versus 'race'in genetics: a systematic review of the use of African ancestry in genetic studies. *Evolution, Medicine, and Public Health*, *9*(1), 232-245. **[Crossref]**

Fakorede, S. T., Adekoya, K. O., Akpan, U. U., & Ogunlusi, O. V. (2019). Allele frequencies and haplotype diversities of five Y-chromosome short tandem repeat loci in a random sample of Yoruba population in Lagos, Nigeria. *FUW Trends in Science & Technology Journal, 4*(2), 577-581

Fan, H., & Chu, J. Y. (2007). A brief review of short tandem repeat mutation. *Genomics, Proteomics and Bioinformatics, 5*(1), 7-14. **[Crossref]**

Gharesouran, J., Hosseinzadeh, H., Ghafouri-Fard, S., Taheri, M., & Rezazadeh, M. (2021). STRs: ancient architectures of the genome beyond the Sequence. *Journal of Molecular Neuroscience*, *71*(12), 2441-2455. **[Crossref]**

Goudet, J. 2003. Fstat (ver. 2.9.4), a program to estimate and test population genetics parameters. Available from **[Crossref]**

Gouy, A., & Zieger, M. (2017). STRAF—a convenient online tool for STR data evaluation in forensic genetics. *Forensic Science International: Genetics*, *30*, 148-151. **[Crossref]**

Greenberg, J. H. (1947). Islam and clan organization among the Hausa. *Southwestern Journal of Anthropology*, *3*(3), 193-211. **[Crossref]**

Hammer, M., & Redd, A. J. (2006). Forensic applications of Y chromosome STRs and SNPs. *Forensics in Law Enforcement*, *133*.

Jobling, M. A., & Tyler-Smith, C. (2003). The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, *4*(8), 598-612. **[Crossref]**

Kasu, M., Cloete, K. W., Pitere, R., Tsiana, K. J., & D'Amato, M. E. (2022). The genetic landscape of South African males: A Y-STR perspective. *Forensic Science International: Genetics*, *58*, 102677. **[Crossref]**

Kayser, M. (2007). Uni-parental markers in human identity testing including forensic DNA analysis. *Biotechniques*, *43*(6), S16-S21. **[Crossref]**

Khubrani, Y. M., Wetton, J. H., & Jobling, M. A. (2018). Extensive geographical and social structure in the paternal lineages of Saudi Arabia revealed by analysis of 27 Y-STRs. *Forensic Science International: Genetics*, *33*, 98-105. **[Crossref]**

Lesaoana, M., Kasu, M., & D'Amato, M. E. (2019). Forensic parameters and genetic structure based on Y-chromosome short tandem repeats in Lesotho populations. *Forensic Science International: Genetics Supplement Series*, *7*(1), 414-415. **[Crossref]**

Lewontin, R. C., & Hartl, D. L. (1991). Population genetics in forensic DNA typing. *Science*, *254*(5039), 1745-1750. **[Crossref]**

Martinez, B., Catelli, L., Romero, M., Okolie, V. O., Keshinro, S. O., Carvalho, E. F., ... & Gusmão, L. (2017). Forensic evaluation of 27 y-str haplotypes in a population sample from nigeria. *Forensic Science International: Genetics Supplement Series*, *6*, e289-e291. **[Crossref]**

Medrano, J. F., Aasen, E., & Sharrow, L. (1990). DNA extraction from nucleated red blood cells. *Biotechniques*, *8*(1), 43-43.

Nei, M., & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics*, *98*(3), 625-640.

Niederstätter, H., Berger, B., Kayser, M., & Parson, W. (2016). Differences in urbanization degree and consequences on the diversity of conventional vs. rapidly mutating Y-STRs in five municipalities from a small region of the Tyrolean Alps in Austria. *Forensic Science International: Genetics*, *24*, 180-193. **[Crossref]**

Okolie, V. O., Cisana, S., Schanfield, M. S., Adekoya, K. O., Oyedeji, O. A., & Podini, D. (2018). opulation data of 21 autosomal STR loci in the Hausa, Igbo and Yoruba people of Nigeria. *International Journal of Legal Medicine*, *132*, 735-737. **[Crossref]**

Ollivier, L., & Foulley, J. L. (2005). Aggregate diversity: new approach combining within-and between-breed genetic diversity. *Livestock Production Science*, *95*(3), 247-254. **[Crossref]**

Purps, J., Siegert, S., Willuweit, S., Nagy, M., Alves, C., Salazar, R., ... & Turrina, S. (2014). A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Science International: Genetics*, *12*, 12-23. **[Crossref]**

Roewer, L. (2019). Y-chromosome short tandem repeats in forensics—Sexing, profiling, and matching male DNA. *Wiley Interdisciplinary Reviews: Forensic Science*, *1*(4), e1336. **[Crossref]**

Swanson, E. C., & Lagace, R. O. (1997). *Hausa: Ethnographic Atlas.* Centre for Social Anthropology and Computing, University of Kent at Canterbury.

Syndercombe-Court, D. (2021). The Y chromosome and its use in forensic DNA analysis. *Emerging Topics in Life Sciences*, *5*(3), 427-441. **[Crossref]**

Turrina, S., Caratti, S., Ferrian, M., & De Leo, D. (2015). Deletion and duplication at DYS448 and DYS626 loci: unexpected patterns within the AZFc region of the Y-chromosome. *International Journal of Legal Medicine, 129*, 449-455. **[Crossref]**

Usman, H. (1997). Reproductive health and rights: The case of Northern Nigerian Hausa women. *Africa Development/Afrique et Développement*, *22*(1), 79-94

Veeramah, K. R., Connell, B. A., Pour, N. A., Powell, A., Plaster, C. A., Zeitlyn, D., ... & Thomas, M. G. (2010). Little genetic differentiation as assessed by uniparental markers in the presence of substantial language variation in peoples of the Cross River region of Nigeria. *BMC Evolutionary Biology*, *10*(1), 1-17. **[Crossref]**

Willems, T., Gymrek, M., Highnam, G., Mittelman, D., Erlich, Y., & 1000 Genomes Project Consortium. (2014). The landscape of human STR variation. *Genome research*, *24*(11), 1894-1904. **[Crossref]**