Novel criteria of Egyptian population using Y-chromosome STRs by Yfiler Plus® System (A: Allele Frequency)

Samir A. Nassar [a*], Ali M. Abd Gawad[a], Mohamed F. Omar and Abdel-Aziz A. A. El-Sayed[a,b]
[a]Zoology Department, Faculty of Science, Zagazig University, Zagazig 44519, Egypt.
[b]Biology Department, Faculty of Science, Islamic University of Madinah, Madinah 42351, Saudi Arabia
Corresponding author: Samir A.Nassar
E-mail: sam59nas@gmail.com
Conflict of interest: None declared
Funding: No funding sources

Abstract

Background: The male-specific part of the human Y chromosome is globally used in forensic DNA analysis, identification of males for criminal justice purpose and population genetics, especially in situations where standard autosomal DNA profiling is not adequately informative. The aim of this study was to develop an allelic frequency database for the Egyptian population in order to obtain population data of 27 Y-STRs in order to appraise the resolution power of these loci in differentiating
Closely- related male individuals.

Materials and methods: In this study, we used the Yfiler® Plus kit, which includes 7 rapidly mutating loci (RM Y-STRs) and 20 standard Y-STR, to analyse 200 unrelated males coming from all the Egyptian governorates. For Y-STR analysis, 200 male buccal swap samples were used in this study, The recently introduced 6-dye Y filer Plus multiplex which includes 27 Y-STR loci (DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS460, DYS458, DYS19, YGATAH4, DYS448, DYS391, DYS456, DYS390, DYS438, DYS392, DYS518, DYS570, DYS437, DYS385 a/b, DYS449, DYS393, DYS439, DYS481, DYF387S1a/b and DYS533) has been used. These RM Y-STRs are useful for discriminating between closely related and unrelated males.

Results: Five thousand and four hundred alleles were detected at the 27 Y-STR loci in 200 samples. Alleles frequency ranged from 1 to 159 and the highest allele frequency registered by allele 14 was 159 at locus DYS437 recording a percentage of 79.5% inside this locus followed by allele 11 with 151 allele frequency recording a percentage of 75.5% at locus DYS392. Also, the highest predominant allele frequency regestered was 0.80 for the predominant allele 17 at locus DYS570, while the lowest frequency was 0.22 for the predominant alleles 18 and 25 at loci DYS385 B and DYS481 respectively.

Conclusion: The present study established the genetic information obtained using the Yfiler Plus® system for the Egyptian population and also created a database of 27 Y- STR markers in this population on the basis of allele dominance and frequencies.

Keywords: Y-chromosome STRs, Yfiler Plus® PCR Amplification Kit (Thermofisher Scientific), Locus, Allele frequency.

Tob Regul Sci.™ 2022;8(1): 3669-3686
DOI: doi.org/10.18001/TRS.8.1.278

1. Introduction:

The male-specific part of the human Y chromosome is globally used in forensic DNA analysis, especially in situations where standard autosomal DNA profiling is not adequately informative. Y-chromosome short tandem repeats (Y-STRs) are of great importance in forensic identification of male DNA from sexual assault cases; valuable in paternity and kinship tests via genealogy (**Roewer, 2009**). Since markers of Y-STRs are located on the non-recombining region (NRY) of Y chromosome; they are transmitted as haplotypes in the same way as single locus alleles, and hence their significance in tracing paternal lineages that help in investigations of missing persons and historical studies and to link families (**Jain** *et al.*, **2016**). **Aim**, The aim of this study was to develop an allelic frequency database for the Egyptian population so as to obtain population data of 27 Y-STRs in order to appraise the resolution power of these loci in differentiating Closely-related male individuals. Therefore, the goal was to establish an online accessible population database for Egyptians on the basis of Y-STR profiles.

2. Material and methods:

Subjects and sample collection

Buccal swap samples were randomly-collected from 200 unrelated adult males from the Egyptian population from 27 governorates. The consented samples were submitted to the central criminal lab "Ministry of Interior" at Abbassia, Cairo, Egypt.

Approval and preservation of samples

Samples were collected with an approval sheet from volunteers that were written in arabic language. The approval form includes volunteer name, family name and origin. Samples which collected from outside cairo were transferred in an icebox to the "central criminal lab" and then approval sheets were separated and samples were stored at -15°C for further use.

DNA Extraction and quantification:

DNA extraction was performed by an automatic BioRobot® EZ1 from Qiagen®, according to the manufactures protocol for isolation of genomic DNA **(Montpetit** *et al.*, **2005**) (Fig.1).

DNA Quantification was carried out by a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, Delaware, USA))(Fig.2) at 260 nm on a sample of DNA solution.

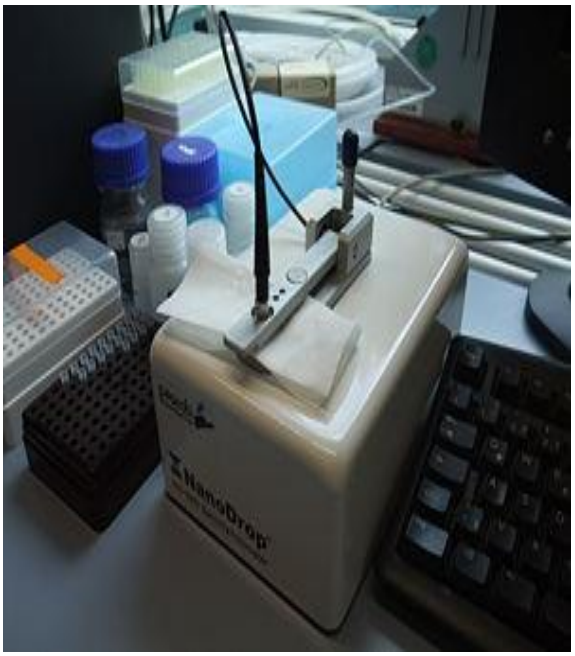


Fig. 1 : BioRobot EZ1 **Fig. 2: Nanodrop ND-1000 spectrophotometer**

Polymerase Chain Reaction (PCR):

Yfiler Plus® PCR Amplification Kit was used to amplify 27 Y-chromosome STR loci. These loci are: " DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS460, DYS458, DYS19, YGATAH4, DYS448, DYS391, DYS456, DYS390, DYS438, DYS392, DYS518, DYS570, DYS437, DYS385 a/b, DYS449, DYS393, DYS439, DYS481, DYF387S1a/b and DYS533".

Amplification was carried out by multiplex PCR in Veriti™ 96-Well Thermal Cycler from (Applied Biosystems) (Fig.3) using the commerical kit Yfiler Plus® system from (Thermofisher Scientific) . The amplified products were run on the ABI PRISM 3500 Genetic Analyzer (Applied Biosystems), and the obtained data was analyzed using the Gene Mapper ID Analysis Software (Applied Biosystems, USA).

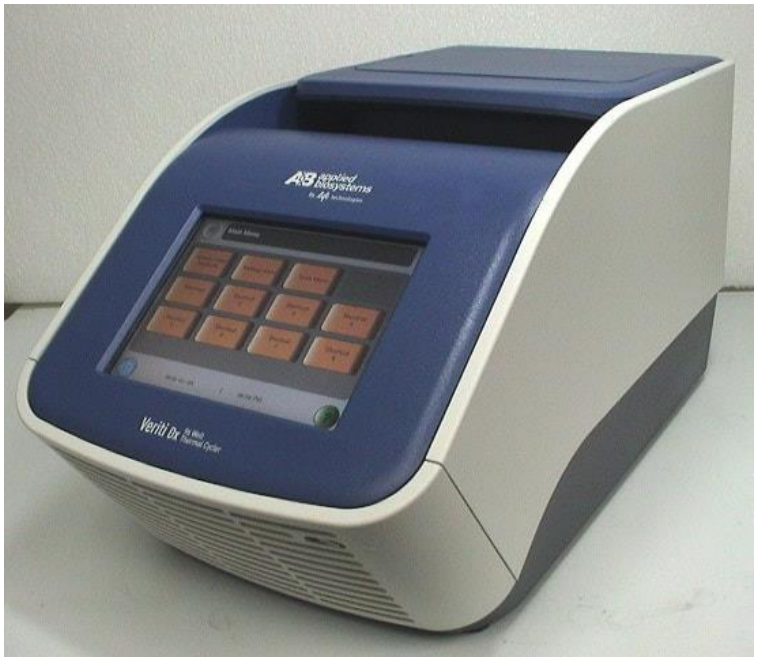


Fig. 3: Veriti™ 96-Well Thermal Cycler

Thermal Cycling Parameters for Yfiler Plus® PCR Kit:

All PCR reactions consisted of initial incubation cycle at 95°C for 1 minute, followed by 27 cycles of denaturation at 94°C for 4 seconds, and annealing and extension at 61.5°C for 1 minute per cycle. The PCR reactions were completed by post-extension for 22 minutes at 60°C as shown in (Table 1).

Table 1: Steps of thermal cycling

Initial incubation step	Optimum cycle number		Final extension	Final hold
	Denature	Anneal/Extend		
Hold	27 Cycle		Hold	Hold
95°C, 1 minute	94°C, 4 seconds	61.5°C, 1 minute	60°C, 22 minutes	4°C, up to 24 hours

Statistical analyses

Allele Frequency-based statistical analysis was calculated with the GenAlEx-6.5 Genetic Analysis software (**Peakall &Smouse, 2006**), while the haplotype diversity was calculated using the HapYdive software (**Parvathy *et al.,* 2012**). Analysis of molecular variance (AMOVA) was calculated by https://yhrd.org/amova/.yhrd.org.tools.

3. Results and discussion

Nowadays, the use of Y-chromosome polymorphisms forms an essential part of many forensic DNA investigations. **De Knijff (2022)** reported that, the use of Y-chromosome

polymorphisms is a very powerful forensic research tool, if used and interpreted with care. Particularly as an exclusion tool in complex male–female crime scene sample mixtures, it has proven to be invaluable.

In the present study, genotyping for the 27 loci was obtained for all samples. A total of 206 alleles were detected at 27 Y-STR loci from 200 unrelated male Egyptian individuals. The number of alleles at each locus ranged from 4-15.The lowest number of alleles was 4 observed at loci DYS389I, DYS460, DYS391, DYS438 and DYS393. However, the highest number of alleles was 15 observed at locus DYS449 (Table 2). Concerning alleles frequency, it was ranged from 1 to 159 and the highest allele frequency registered by allele 14 was 159 at locus DYS437 recording a percentage of 79.5% inside this locus followed by allele 11 with 151 allele frequency recording apercentage of 75.5% at locus DYS392(Figure 5, Table, 3). Analysis of allele frequenies in the Egyptian population showed that every locus has a prodominant allelle. The current study illustrated the predominant alleles in each locus of the studied loci in the Egyptian poulation. The highest allele frequency regestered was 0.80 for the predominant allele 17 at locus DYS570, while the lowest frequency was 0.22 for the predominant alleles 18 and 25 at loci DYS385 B and DYS481 respectively (Table 4 and Figure 6). In the current experiment, the use of the Yfiler® Plus network of Y chromosomes, which contains seven rapidly mutating loci (RM Y-STRs) and 20 standard Y-STR, to analyse 200 unrelated males coming from all the Egyptian governorates, verified the high power of discrimination of this kit, thanks to higher variability of the RM Y-STRs. This is in agreement with **D'Atanasio** *et al.*(**2019**) who compared the use of both Yfiler® and Yfiler® Plus network of the E-M81 Y chromosomes and proved a high power of discrimination of the latter kit in discriminate between 477 male subjects belonging to 11 ethnic groups sampled from four northern African countries (Morocco, Algeria, Libya and Egypt). Also, **Ballantyne** et al. (**2010**) reported that, the most recent multiplex for the capillary electrophoretic analysis of Y-STRs is the Yfiler® Plus, its 7 rapidly-mutating STRs (RM-STRs) are of high mutation rate (higher than 1 × 10-2). The inclusion of the RM-STRs in the Yfiler® Plus multiplex increased the level of intra-population haplotype diversity and decreased the haplotype sharing between males from different populations to zero degree (**Iacovacci** et al., **2017**; **Khubrani** et al., **2018**) .

3.1. Number of alleles in the studied loci

Table 2. Total number of alleles for each locus in the Egyptian population

Locus	Number of alleles
DYS576	8
DYS389I	4
DYS635	9
DYS389II	8
DYS627	11
DYS460	4
DYS458	8

DYS19	6
YGATAH4	6
DYS448	7
DYS391	4
DYS456	6
DYS390	7
DYS438	4
DYS392	6
DYS518	11
DYS570	12
DYS437	5
DYS385 A	10
DYS385 B	11
DYS449	15
DYS393	4
DYS439	5
DYS481	11
DYS387S1 A	9
DYS387S1 B	9
DYS533	6
Total No. of allèles	206
Mean	7.630
S.D.	2.937

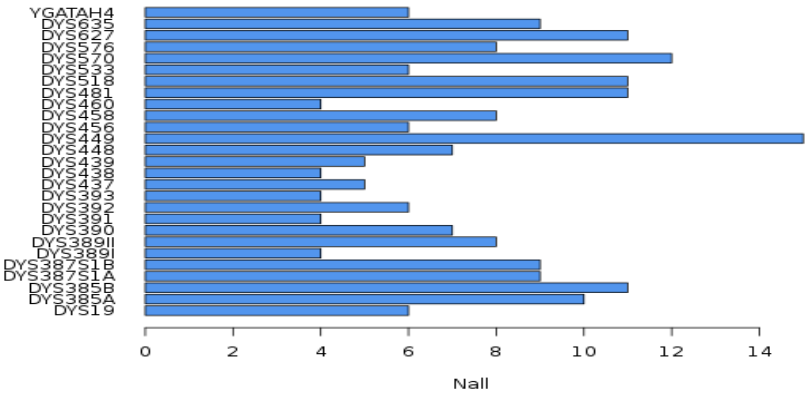


Figure. 4 Number of alleles at each locus

3.2. Allele frequencies of different loci for the Egyptian population

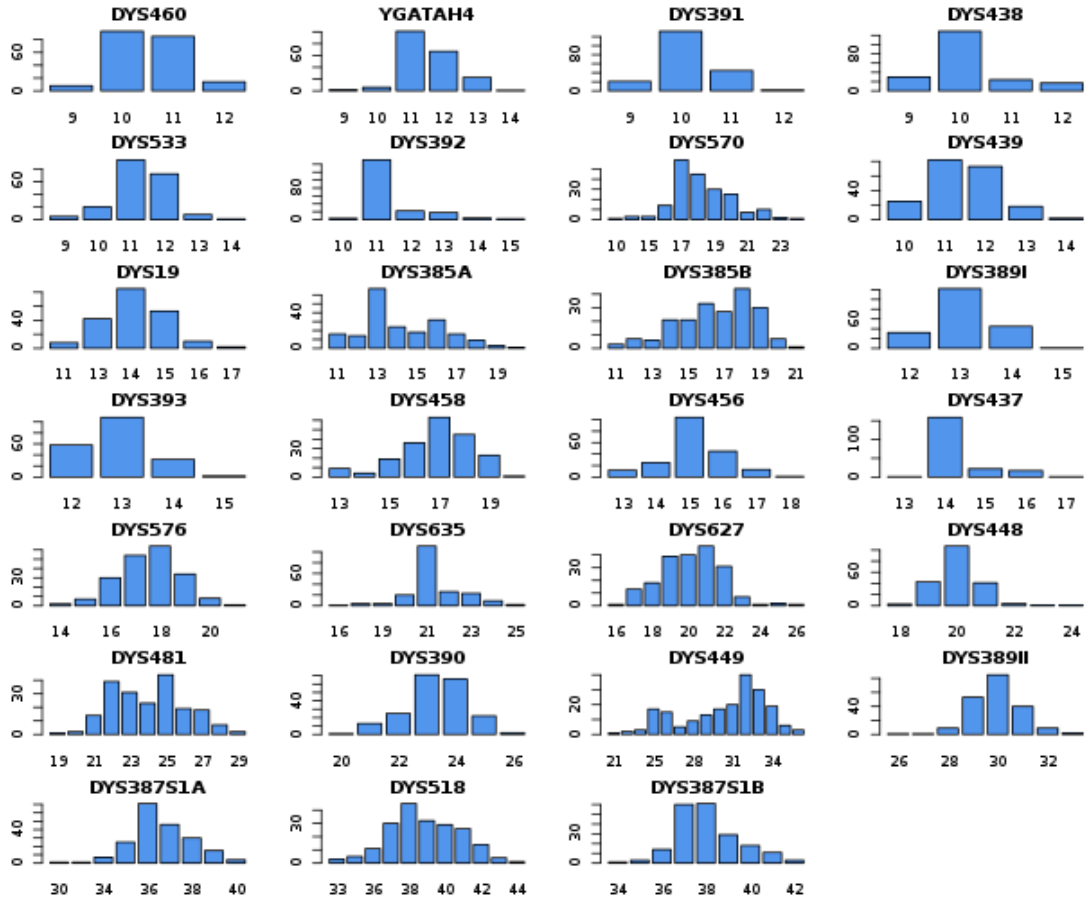
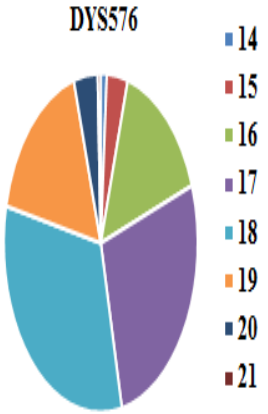


Figure. 5. Allele frequencies per locus in histograms

Table. 3. Allele frequency of different loci with pie charts for the Egyptian population

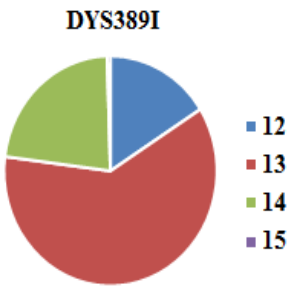
1. DYS576

Allele	Freq.	%	S.D.
14	2	1.0	0.007
15	7	3.5	0.013
16	30	15.0	0.025
17	54	27.0	0.031
18	64	32.0	0.033
19	34	17.0	0.027
20	8	4.0	0.014
21	1	0.5	0.005
Total	**200**	**100.0**	



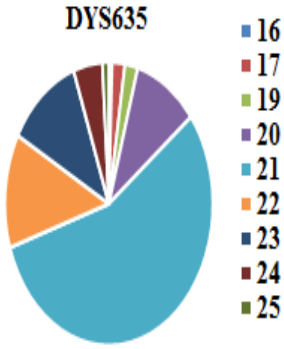
2. DYS389I

Allele	Freq.	%	S.D.
12	32	16.0	0.026
13	122	61.0	0.035
14	45	22.5	0.030
15	1	0.5	0.005
Total	**200**	**100.0**	



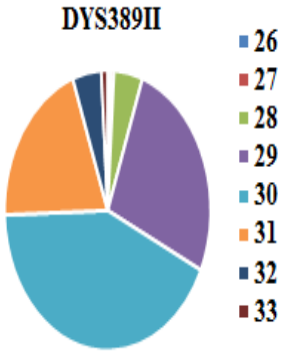
3. DYS635

Allele	Freq.	%	S.D.
16	1	0.5	0.005
17	4	2.0	0.010
19	4	2.0	0.010
20	20	10.0	0.021
21	111	55.5	0.035
22	26	13.0	0.024
23	23	11.5	0.023
24	9	4.5	0.015
25	2	1.0	0.007
Total	**200**	**100.0**	



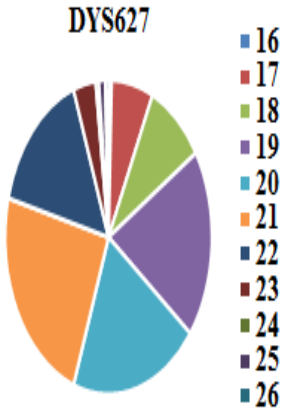
4. DYS389II

Allele	Freq.	%	S.D.
26	1	0.5	0.005
27	1	0.5	0.005
28	9	4.5	0.015
29	53	26.5	0.031
30	85	42.5	0.035
31	40	20.0	0.028
32	9	4.5	0.015
33	2	1.0	0.007
Total	**200**	**100.0**	



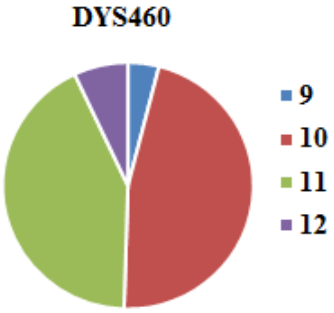
5. DYS627

Allele	Freq.	%	S.D.
16	1	0.5	0.005
17	13	6.5	0.017
18	18	9.0	0.020
19	39	19.5	0.028
20	40	20.0	0.028
21	47	23.5	0.030
22	31	15.5	0.026
23	7	3.5	0.013
24	1	0.5	0.005
25	2	1.0	0.007
26	1	0.5	0.005
Total	**200**	**100.0**	



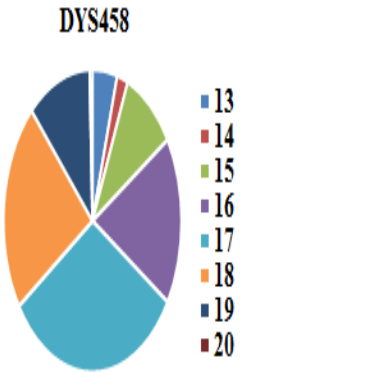
6. DYS460

Allele	Freq.	%	S.D.
9	8	4.0	0.014
10	93	46.5	0.035
11	85	42.5	0.035
12	14	7.0	0.018
Total	**200**	**100.0**	



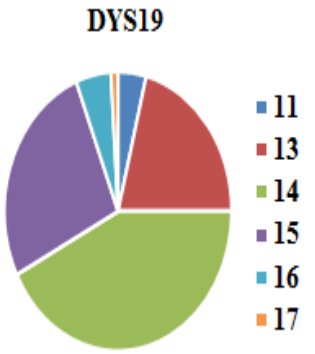
7. DYS458

Allele	Freq.	%	S.D.
13	9	4.5	0.015
14	4	2.0	0.010
15	19	9.5	0.021
16	36	18.0	0.027
17	63	31.5	0.033
18	45	22.5	0.030
19	23	11.5	0.023
20	1	0.5	0.005
Total	**200**	**100.0**	



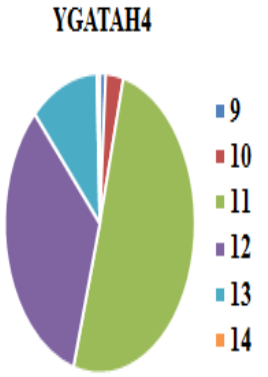
8. DYS19

Allele	Freq.	%	S.D.
11	8	4.0	0.014
13	42	21.0	0.029
14	85	42.5	0.035
15	53	26.5	0.031
16	10	5.0	0.015
17	2	1.0	0.007
Total	**200**	**100.0**	



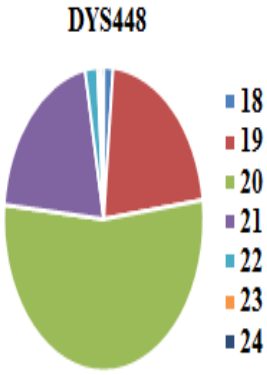
9. YGATAH4

Allele	Freq.	%	S.D.
9	2	1.0	0.007
10	6	3.0	0.012
11	101	50.5	0.035
12	67	33.5	0.033
13	23	11.5	0.023
14	1	0.5	0.005
Total	**200**	**100.0**	



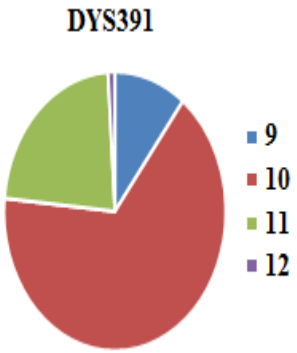
10. DYS448

Allele	Freq.	%	S.D.
18	3	1.5	0.009
19	43	21.5	0.029
20	107	53.5	0.035
21	41	20.5	0.029
22	4	2.0	0.010
23	1	0.5	0.005
24	1	0.5	0.005
Total	**200**	**100.0**	



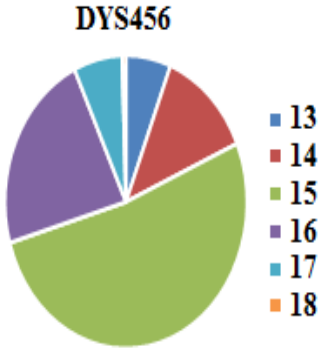
11. DYS391

Allele	Freq.	%	S.D.
9	21	10.5	0.022
10	132	66.0	0.034
11	45	22.5	0.030
12	2	1.0	0.007
Total	**200**	**100.0**	



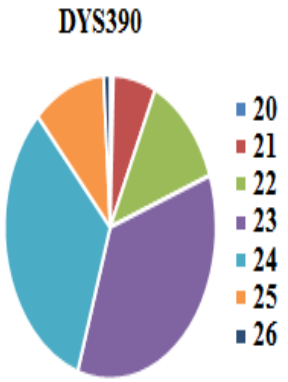
12. DYS456

Allele	Freq.	%	S.D.
13	12	6.0	0.017
14	25	12.5	0.023
15	104	52.0	0.035
16	45	22.5	0.030
17	13	6.5	0.017
18	1	0.5	0.005
Total	**200**	**100.0**	



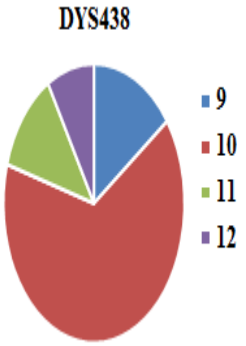
13. DYS390

Allele	Freq.	%	S.D.
20	1	0.5	0.005
21	13	6.5	0.017
22	25	12.5	0.023
23	71	35.5	0.034
24	66	33.0	0.033
25	22	11.0	0.022
26	2	1.0	0.007
Total	**200**	**100.0**	



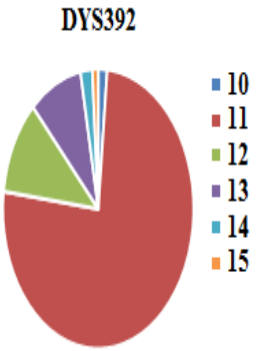
14. DYS438

Allele	Freq.	%	S.D.
9	30	15.0	0.025
10	129	64.5	0.034
11	24	12.0	0.023
12	17	8.5	0.020
Total	**200**	**100.0**	



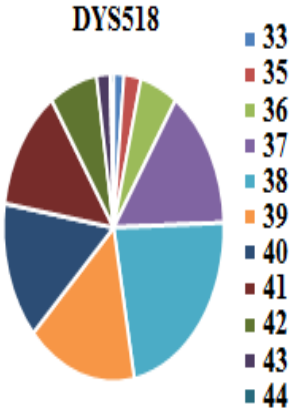
15. DYS392

Allele	Freq.	%	S.D.
10	3	1.5	0.009
11	151	75.5	0.030
12	22	11.0	0.022
13	18	9.0	0.020
14	4	2.0	0.010
15	2	1.0	0.007
Total	**200**	**100.0**	



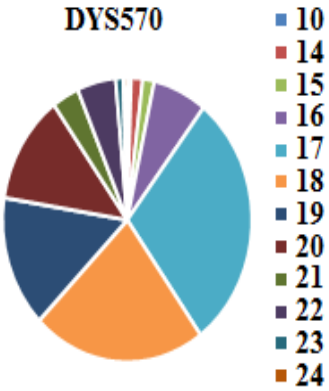
6. DYS518

Allele	Freq.	%	S.D.
33	3	1.5	0.009
35	5	2.5	0.011
36	11	5.5	0.016
37	30	15.0	0.025
38	45	22.5	0.030
39	32	16.0	0.026
40	29	14.5	0.025
41	26	13.0	0.024
42	14	7.0	0.018
43	4	2.0	0.010
44	1	0.5	0.005
Total	**200**	**100.0**	



17. DYS570

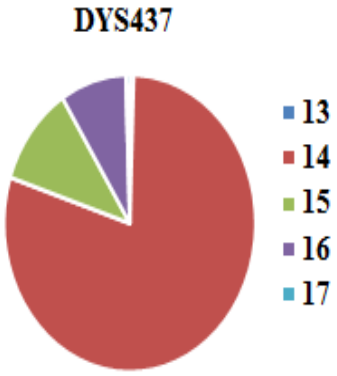
Allele	Freq.	%	S.D.
10	1	0.5	0.005
14	3	1.5	0.009
15	3	1.5	0.009
16	14	7.0	0.018
17	59	29.5	0.032
18	45	22.5	0.030
19	30	15.0	0.025
20	25	12.5	0.023
21	7	3.5	0.013
22	10	5.0	0.015



23	2	1.0	0.007
24	1	0.5	0.005
Total	**200**	**100.0**	

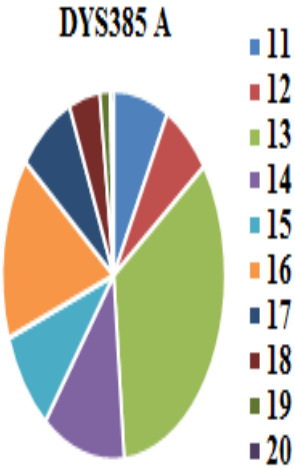
18. DYS437

Allele	Freq.	%	S.D.
13	1	0.5	0.005
14	159	79.5	0.029
15	22	11.0	0.022
16	17	8.5	0.020
17	1	0.5	0.005
Total	**200**	**100.0**	



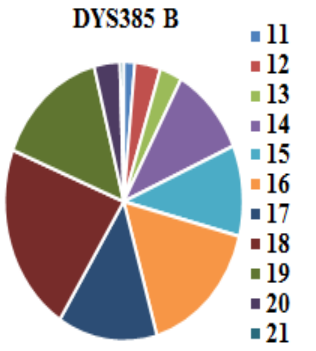
19. DYS385 A

Allele	Freq.	%	S.D.
11	16	8.0	0.019
12	14	7.0	0.018
13	67	33.5	0.033
14	24	12.0	0.023
15	18	9.0	0.020
16	32	16.0	0.026
17	16	8.0	0.019
18	9	4.5	0.015
19	3	1.5	0.009
20	1	0.5	0.005
Total	**200**	**100.0**	



20. DYS385 B

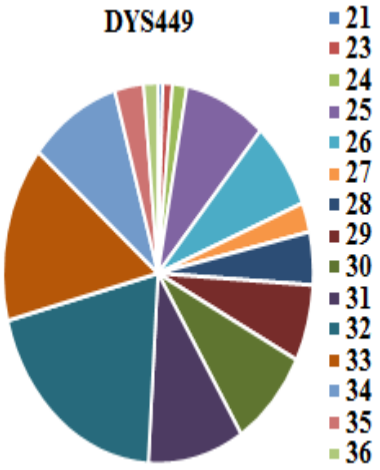
Allele	Freq.	%	S.D.
11	3	1.5	0.009
12	7	3.5	0.013
13	6	3.0	0.012
14	21	10.5	0.022
15	21	10.5	0.022
16	33	16.5	0.026
17	27	13.5	0.024
18	44	22.0	0.029
19	30	15.0	0.025



20	7	3.5	0.013
21	1	0.5	0.005
Total	**200**	**100.0**	

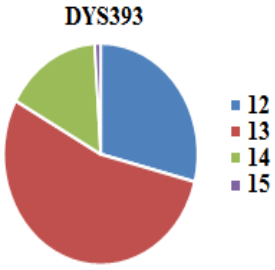
21. DYS449

Allele	Freq.	%	S.D.
21	1	0.5	0.005
23	2	1.0	0.007
24	3	1.5	0.009
25	17	8.5	0.020
26	15	7.5	0.019
27	5	2.5	0.011
28	9	4.5	0.015
29	13	6.5	0.017
30	17	8.5	0.020
31	20	10.0	0.021
32	40	20.0	0.028
33	30	15.0	0.025
34	19	9.5	0.021
35	6	3.0	0.012
36	3	1.5	0.009
Total	**200**	**100.0**	



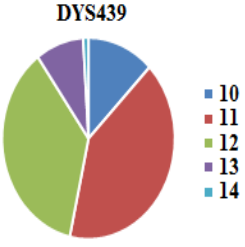
22. DYS393

Allele	Freq.	%	S.D.
12	58	29.0	0.032
13	108	54.0	0.035
14	32	16.0	0.026
15	2	1.0	0.007
Total	**200**	**100.0**	



23. DYS439

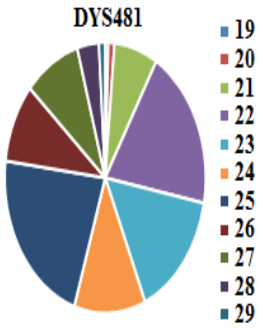
Allele	Freq.	%	S.D.
10	25	12.5	0.023
11	82	41.0	0.035
12	73	36.5	0.034
13	18	9.0	0.020
14	2	1.0	0.007



Total	**200**	**100.0**	

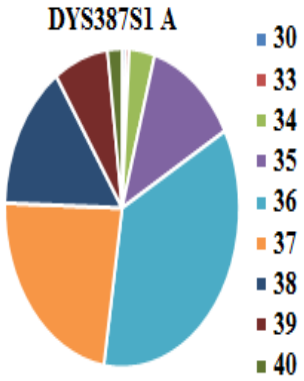
24. DYS481

Allele	Freq.	%	S.D.
19	1	0.5	0.005
20	2	1.0	0.007
21	14	7.0	0.018
22	39	19.5	0.028
23	31	15.5	0.026
24	23	11.5	0.023
25	44	22.0	0.029
26	19	9.5	0.021
27	18	9.0	0.020
28	7	3.5	0.013
29	2	1.0	0.007
Total	**200**	**100.0**	



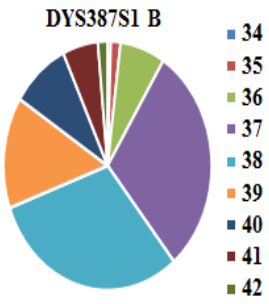
25. DYS387S1 A

Allele	Freq.	%	S.D.
30	1	0.5	0.005
33	1	0.5	0.005
34	7	3.5	0.013
35	25	12.5	0.023
36	71	35.5	0.034
37	46	23.0	0.030
38	30	15.0	0.025
39	15	7.5	0.019
40	4	2.0	0.010
Total	**200**	**100.0**	



26. DYS387S1 B

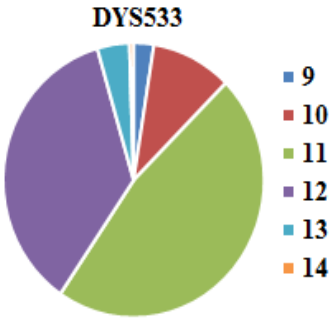
Allele	Freq.	%	S.D.
34	1	0.5	0.005
35	3	1.5	0.009
36	14	7.0	0.018
37	60	30.0	0.032
38	61	30.5	0.033
39	29	14.5	0.025



40	18	9.0	0.020
41	11	5.5	0.016
42	3	1.5	0.009
Total	**200**	**100.0**	

27. DYS533

Allele	Freq.	%	S.D.
9	5	2.5	0.011
10	20	10.0	0.021
11	94	47.0	0.035
12	72	36.0	0.034
13	8	4.0	0.014
14	1	0.5	0.005
Total	**200**	**100.0**	



3.4. Frequency of predominant alleles the Egyptian population

Table 4. Predominant alleles in the Egyptian population

Locus	Predominant	Frequency
DYS576	18	0.32
DYS389I	13	0.61
DYS635	21	0.56
DYS389II	30	0.43
DYS627	21	0.24
DYS460	10	0.47
DYS458	17	0.32
DYS19	14	0.43
YGATAH4	11	0.51
DYS448	20	0.54
DYS391	10	0.66
DYS456	15	0.52
DYS390	23	0.36
DYS438	10	0.65
DYS392	11	0.76
DYS518	38	0.23
DYS570	17	0.30
DYS437	14	0.80
DYS385 A	13	0.34

DYS385 B	18	0.22
DYS449	32	0.20
DYS393	13	0.54
DYS439	11	0.41
DYS481	25	0.22
DYS387S1 A	36	0.36
DYS387S1 B	38	0.31
DYS533	11	0.47

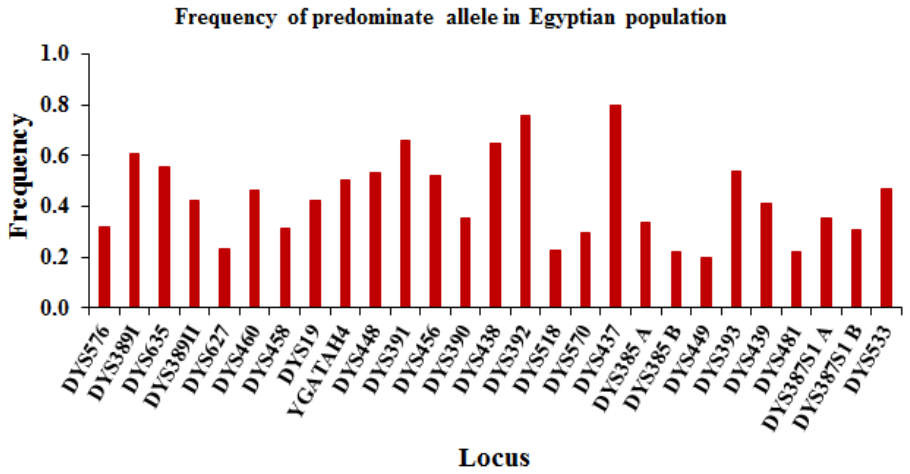


Figure 6. Frequency of predominant alleles in the Egyptian population

4. Conclusion:

A recent RM-YSTR multiplex assay tested in this study functioned well and efficiently generated genotyping data for all 200 Egyptian donors. Based on the results of this study, the genetic information obtained by using the Yfiler Plus® System for the Egyptian population was established and created a nucleus for database of 27 Y- STR markers in this population on the basis of allele frequencies and predominant alleles at each locus. Therefore, demonstrating their usefulness in forensic identification and parentage cases.

Acknowledgement

We thank all the donors for providing DNA samples that made this study possible.

References:

[1] **D'Atanasio E, Iacovacci G, Pistillo R, Bonito M, Dugoujon J-Michel, Moral P, El-Chennawi F, Melhaoui M, Baali A, Cherkaoui M, Sellitto D, Trombetta B, Berti A, Cruciani F**, Rapidly mutating Y-STRs in rapidly expanding populations: Discrimination power of the Yfiler Plus multiplex in northern Africa, Forensic Science International: Genetics (**2019**), https://doi.org/10.1016/j.fsigen.**2019**.11.002.

[2] **De Knijff, P.** On the Forensic Use of Y-Chromosome Polymorphisms. Genes **2022**, 13, 898. https://doi.org/10.3390/genes13050898
[3] **Ballantyne K.N. , Goedbloed M., Fang R., Schaap O.,** *et al,* Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications, Am. J. Hum. Genet. 87 (**2010**) 341–353. doi:10.1016/j.ajhg.2010.08.006.
[4] **Jain T, Shrivastava P, Bansal DD, Dash HR, Trivedi VB (2016)** PowerPlex Y23 system: a fast, sensitive and reliable Y-STR multiplex system for forensic and population genetic purpose. J Mol Biomark Diagn 7(281):2. https://doi.org/10. 4172/2155-9929.1000287
[5] **Khubrani Y.M., Wetton J.H., Jobling M.A.**, Extensive geographical and social structure in the paternal lineages of Saudi Arabia revealed by analysis of 27 Y-STRs, Forensic Sci. Int. Genet. 33 (**2018**) 98–105. doi:10.1016/j.fsigen.2017.11.015
[6] **Iacovacci G., D'Atanasio E., Marini O., Coppa A., Sellitto D., Trombetta B., Berti A., Cruciani F.**, Forensic data and microvariant sequence characterization of 27 Y-STR locianalyzed in four Eastern African countries, Forensic Sci. Int. Genet. 27 (**2017**) 123–131. doi:10.1016/j.fsigen.2016.12.015.
[7] **Montpetit S.A., Fitch I.T. and O'Donnell P.T. (2005):** A Simple automated Instrument for DNA Extraction in Forensic Casework. Journal of Forensic Sciences 50 (3):555:563.
[8] **Parvathy SN, Geetha A, Jagannath C (2012)** Haplotype analysis of the polymorphic 17 YSTR markers in Kerala nontribal populations. Mol Biol Rep 39(6):7049–7059.
[9] **Peakall ROD, Smouse PE (2006)** GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6(1):288–295.
[10] **Roewer L (2009)** Y chromosome STR typing in crime casework. Forensic Sci Med Pathol 5(2):77–84.