Journal of Forensic Research
ISSN: 2157-7145

Al-Azem et al., J Forensic Res 2020, 11:1

Research Article Open Access

Assessment of the Inbreeding Effect on Y-STR Profiles in the Lebanese Population

Mouayyad Al-Azem[1], Nassib Dagher[1], Marc Obeid[1], Sarah Abbas[1,3], Hiba Othman[2] and Issam Mansour[1*]

[1]*Department of Laboratory Science and Technology, Faculty of Health Sciences, American University of Science and Technology, Beirut, Lebanon*

[2]*Department of Mathematics, American University of Science and Technology, Lebanon*

[3]*School of Criminal Justice, University of Lausanne, Switzerland*

***Corresponding author:** Issam Mansour, Department of Laboratory Science and Technology, Faculty of Health Sciences, American University of Science and Technology, Beirut, Lebanon, Tel: +961218716230; E-mail: fhs@aust.edu

Received date: December 28, 2019; **Accepted date:** January 21, 2020; **Published date:** January 28, 2020

Abstract

Y-STR profiling is gaining interest in forensic investigations. However, differentiation resulting from genetic stratification by genetic relatedness could be very pronounced in specific populations and thus imposing a possible limitation to Y-STR applications. In Lebanon, published data showed that endogamy average rate amounts up to 88%. Thereby, it is essential to determine the effect of inbreeding on Y-STR haplotype match probability and its consequences upon the analysis of forensic and paternity studies in the Lebanese population. A total of 241 male samples from seven villages were tested with two different haplotype sizes using the Y-filer Kit (Applied Biosystems) which includes: DYS19, DYS390, DYS391, DYS392, DYS393, DYS389I/II, DYS385a/b, DYS437, DYS438, DYS439, DYS456, DYS456, DYS448 and Y-GATA-H4 systems, and the Y-23 Kit (Promega Corporation) which includes six extra systems: DYS576, DYS570, DYS549, DYS643, DYS533 and DYS481. Results showed that some haplotypes were common among unrelated males carrying different family names. The most common haplotype appeared as frequently as 19 times in 36 males from one village. The use of the Y-23 profile increased the haplotype diversity and discrimination power and decreased the match rate. Nevertheless, a significant number of profile matches among unrelated individuals belonging to the same village were still observed.

Keywords: Y-chromosome; Y-STR; Common haplotype; Endogamy; Lebanese villages

Introduction

During the past three decades there has been a dramatic increase in the use and reliance on forensic DNA analysis especially in the use of autosomal STRs. This has been especially true in sexual assault cases that routinely consist of evidentiary stains that are a mixture of body fluids from the victim and assailant. Current DNA differential analysis techniques permit the separation of the male and female components of these mixed stains [1]. Yet a complete separation is not always possible due to the size and condition of the evidence sample and the percentage of each component present in it [2-6]. Moreover, in some of these situations, autosomal STR fails to be informative due to several challenges such as failing to amplify the minor male component of DNA mixtures due to competition with alleles from the major female component, multiple genotype possibilities, homozygozity, allele sharing between contributors, allele drop-outs, allele stacking and stutter production which could lead to false inclusions using 15 and even 23 autosomal STR systems [7]. One approach to resolving these cases is to target male-specific polymorphisms such as Y-STRs found on the non-recombining portion of the Y chromosome (NRY) [8-12].

Hence, Y-STR analysis is used for forensic identity testing where a stain or other evidence sample must be compared to a suspect's profile. Moreover, it has even been suggested that the questioned Y haplotype might be able to predict a suspect's surname and provide an investigative lead when autosomal typing was unable to do so [13,14]. Y-STRs could also be used in kinship testing where questions of relatedness between individuals are asked especially in deficient paternity cases or motherless cases where autosomal STRs will not provide sufficient discrimination in order to assign paternity especially when using low number of STR systems (e.g. 16 systems) [15-19].

One limitation of Y-STR in forensic and paternity applications is that the majority of the Y-chromosome does not undergo recombination and is always in a haploid state with most of the polymorphisms lying in the non-recombining region of the Y-chromosome (NRY) [20,21]. NRY is inherited unchanged through paternal lineages unless a meiotic mutation occurs. So a match between an evidentiary sample and a suspect means that male members of the paternal lineage (father, brothers, sons, etc.) in addition to any male who shares a more distant ancestry with the suspect are also not excluded as potential sources because they all share the same Y chromosome [22].

Recent studies showed that differentiation resulting from genetic stratification by genetic relatedness could be very pronounced in specific populations due to endogamy and introduces further limitations to Y-STR analysis [23]. Lebanon is a small country with an area of 10,452 km^2 that encompasses eighteen different religious communities officially recognized by the Lebanese government, with endogamous marriages averaging 88% [24,25]. These communities are either Christian (including Maronite, Greek Orthodox, Melkite Catholic, Armenian Orthodox and Syriac Catholic, Armenian Catholic, Syriac Orthodox, Roman Catholic, Chaldean, Assyrian, Copt and Protestant sub-communities) or Muslim (including Sunnite, Shiite, Druze and Alawite sub-communities). These subpopulations are distributed over distinct geographical areas (Figure 1), for several reasons such as historical distribution where for example the Maronites

settled first in the North of Lebanon for hundreds of years, in addition to some parts of Mounts of Lebanon. The Druze was in their majority in the Mounts of Lebanon. As for the Muslims they were mainly settled on the coast in addition to the South and the Bekaa regions. Political boundaries also played a role in this distribution where the Lebanese civil war was one of the main reasons to create boundaries and fix geographical enclosure, along with migration in both directions that increased the seclusions, especially in Beirut and Mount Lebanon. Finally, there is the cultural reason where the majority of the Lebanese marriages happen within the same religion (endogamous pattern), with a high percentage of marriages within the same family (consanguinity pattern). All these combined factors may add to the genetic isolation in the Lebanese population.

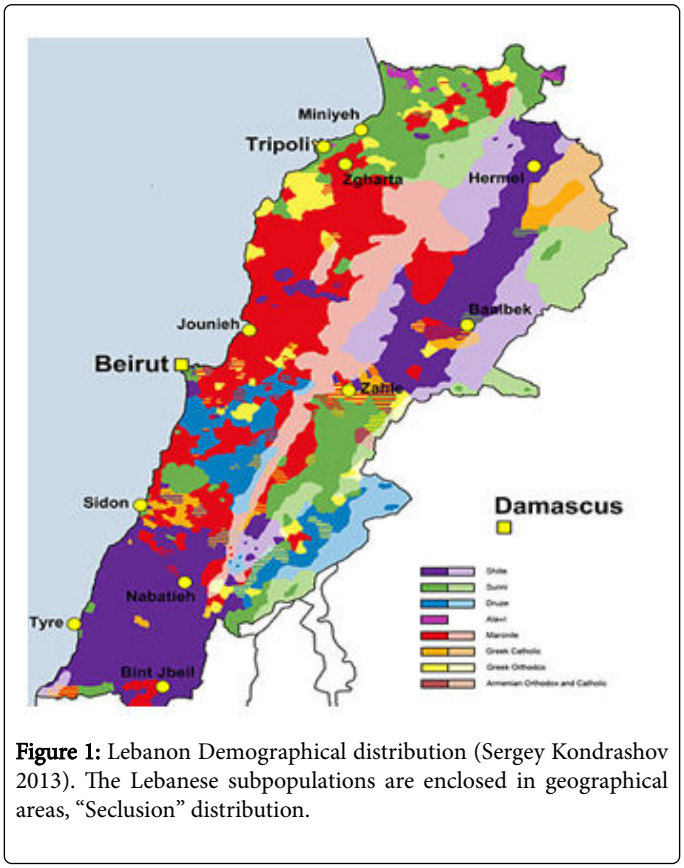

Figure 1: Lebanon Demographical distribution (Sergey Kondrashov 2013). The Lebanese subpopulations are enclosed in geographical areas, "Seclusion" distribution.

Moreover, this genetic isolation is more prominent in Lebanese villages and rural areas than in major cities. These areas are characterized by the highest rates of consanguineous and endogamous marriages which is associated with low socio-economic status, illiteracy, land-owning families and traditional ruling groups. In other words, villages are systematically formed around a few powerful families that force marriages with the same family and within the same village in order to maintain the family structure and property and strengthen the family ties. In addition, Lebanese villages are known to be divided based on religious sects. Lebanon's religious divisions are extremely complicated, and the country is made up by a multitude of religious groupings. The ecclesiastical and demographic patterns of the sects and denominations are complex. Divisions and rivalries between religious groups date back as far as the 15th century, and are still a factor till today. Hence, each religious community has the intention to increase the number of its own religious group, thus prohibits marriages from other religious groups which further restricts the genetic structure in villages. Since marriages in Lebanon follow the religious group of the males, this further increases the rate of marriages from the same family from the same village and religious group [26,27].

A previous study using 23 Y-STR systems was done to determine the allele and haplotype frequencies in the Lebanese population [28], whereby results showed a degree of endogamy at the level of religious sub-populations, and that non-related individuals carried common Y-STR profiles. Moreover, at the level of autosomal STR, studies done on highly consanguineous and endogamous villages showed that the use of 15 or 23 autosomal STR may fail to exclude in DNA mixtures [7]. Consequently, it becomes essential to determine if there is an effect of endogamous marriages on haplotype frequency and the match probability effect on Y-STR haplotype, and its consequences upon the analysis of forensic DNA evidences and paternity studies in the Lebanese population. The present study targets populations of villages knowing that villages represent more than half of the Lebanese population.

Materials and Methods

Sampling and DNA extraction

Samples were collected from seven Lebanese villages from different geographic locations representing all Lebanese regions and the main religious groups (Table 1). These villages are also known to have high endogamous marriages [7]. Samples were collected from 518 individuals out of which 241 were males. 53 whole blood samples were collected from males in village A, 45 buccal swabs were sampled from males in village B, 41 buccal swabs were sampled from males in village C, 29 buccal swabs were sampled from males in village D, 36 buccal swabs were sampled from males in village E, 25 buccal swabs were sampled from males in village F and 12 blood samples were collected from males in village A.

Village	Population Size	Majorities Religion	Geographical Location	# of Collected Samples	# of Male Samples
A	15,318	Muslim Sunni	Bekaa	150	53
B	3774	Muslim Druze	Mount Lebanon	102	45
C	1406	Muslim Shia	Zahle	85	41
D	7075	Christian Maronite	Nabatieh	72	29
E	2000	Christian Orthodox	Denniyeh	59	36
F	2400	Christian Armenian Apostolic	Bekaa	28	25
G	6011	Muslim Shia	Bekaa	22	12

Table 1: Lebanese village's background and the number of collected samples.

DNA was extracted from whole blood leukocytes using the salting out method and from buccal swabs using a modified phenol-chloroform method. Samples were quantified using Nanodrop 2000 (Thermo Fisher Scientific Inc.) and diluted accordingly to approximately 1 ng/µl.

PCR amplification

Y-STR analysis was performed in two stages starting by a haplotype size of 17 systems and followed by a haplotype size of 23 systems. DNA amplification was performed using two commercial kits: the Applied Biosystems Y-Filer® multiplex PCR Amplification kit (Applied Biosystems, Foster City, CA) and the Promega PowerPlex® Y23 System (Promega, Madison, USA). The 23 Y-STR systems include the 11 core loci recommended by the SWGDAM: DYS19, DYS385a/b, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439 and additional markers DYS437, DYS448, DYS456, DYS458, DYS635 and Y-GATA-H4. Amplifications typically contained 0.5–1.0 ng of extracted DNA. The final reaction volume was 6.25 µl and PCR reactions were carried out using GeneAmp PCR System 9700 (Applied Biosystems) and the cycling conditions as described by the manufacturer [29,30].

DNA typing

Electrophoretic separation and detection were performed using the ABI PRISM® 3130 Genetic Analyzer4-capillary array system (ABI Prism 3130 Data Collection Software version 3.0) (Applied Biosystems, Foster City, CA). Size calling was performed using the GeneScan-500 Internal Lane Size Standard (LIZ-500) (Applied Biosystems) and CC5 Internal Lane Standard 500 (Promega, Madison, USA). Genotyping was performed by comparison with the provided allelic ladder and using Genemapper v4.0 (Applied Biosystems).

Statistical analysis

Y-STR data from the GeneScan® software was transferred to in-house software named Forensic Information Management System (FIMS) to estimate the haplotype matching for the 23 Y-STR systems. Haplotype diversity (HD) was calculated according to the formulas supplied by Nei [31]: HD = (n/n-1) (1-Σpi^2). Unique haplotypes (UH), random match probability (RMD) and discrimination capacity (DC) [32,33] were also calculated for the obtained data. Y-STR alleles are inherited in haplotypes, so their individual frequencies cannot be the product of the combined frequencies. Haplotype frequency was obtained using the counting method [34,35].

An extrapolation formula was used to determine the frequency of the most common haplotype that appeared in each of the villages. Extrapolation method was based on normal distribution law of proportions and it was done at 99% confidence level:

$$\hat{p} = \frac{k}{n}\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}} < \mathbf{p}\widehat{<p} + z_{\alpha/2}$$

whereby is the sample proportion (percentage), k is the number of haplotypes and n is the sample size. Thereby p is the true proportion (percentage) of the entire population.

Systems	9 Individuals	5 Individuals	4 Individuals	3 Individuals	3 Individuals	2 Individuals	2 Individuals	2 Individuals
DYS 456	16	14	17	16	17	15	15	16
DYS 389I	13	14	13	13	13	13	13	13
DYS 390	24	24	23	24	23	23	24	24
DYS 389II	30	31	29	30	29	30	30	30
DYS 458	15	17.2	16	15	17	18.2	15	15
DYS 19	14	14	15	14	15	15	14	14
DYS 385 a/b	18-19	13-18	13-16	18-19	13-16	13-19	16-17	18-19
DYS 393	13	12	12	13	12	12	15	14
DYS 391	10	11	9	10	9	11	10	10
DYS 439	14	11	12	13	12	12	11	14
DYS 635	22	22	22	22	22	21	23	22
DYS 392	11	11	11	11	11	11	11	11
Y GATA H4	13	11	12	13	12	11	12	13
DYS 437	14	14	14	14	14	14	14	14
DYS 438	10	10	9	10	9	10	10	10
DYS 448	19	20	21	19	21	20	20	19

Table 2: Different haplotypes observed in village A with extrapolation results for a haplotype occurring 9 times in this dataset (Extrapolation of frequency of the most recurrent haplotype (9 times): 3.6%<p<30.2%).

Quality control

A proficiency testing quality control check was performed in conjunction with the YHRD [36].

Results and Discussion

Y-STR Haplotype determination in villages using 17 systems

17 Y-STR profiles were determined for all the males in each of the tested villages, and matching comparisons were performed simultaneously. In village A, out of 53 male samples, 31 different haplotypes were observed. The most common haplotype appeared nine times (17% of male samples). This haplotype belonged to unrelated individuals but carrying same surname. A second haplotype was observed four times (7% of total male samples), belonged to four non-related individuals with four distinct surnames. These results reflect the importance of common ancestry whereby although individuals may originate from distinct families, yet they might share a common ancestor yielding common Y-STR haplotype profiles. Results in Table 2 show haplotypes that are present more than once.

Extrapolation studies were performed to determine how frequent the most common haplotype (9 times) could occur in the entire village. Extrapolation results revealed that this haplotype would be carried at least by 3.6% and at best by 30.2%.

In Village B, out of 45 male samples, 17 different haplotypes were observed. The most common haplotype appeared 14 times (31% of male samples) and this haplotype originates from individuals from two different surnames. Extrapolation statistics showed that this haplotype would be carried by at least 13.3% and at best 48.9% of the total males in this village, meaning that almost half of the village's males could have this same Y-STR haplotype. Frequently inhabitants of this village with different family names had same haplotypes. Results in Table 3 show haplotypes that are present more than once.

Systems	14 Individuals	5 Individuals	4 Individuals	3 Individuals	3 Individuals	2 Individuals	2 Individuals	2 Individuals	2 Individuals
DYS 456	16	15	16	15	15	15	16	16	17
DYS 389I	13	13	13	13	13	13	12	13	13
DYS 390	25	22	24	22	24	25	23	25	23
DYS 389II	29	29	30	30	32	29	28	29	28
DYS 458	17	15	15	15	16	17	17	17	17
DYS 19	14	15	13	14	16	14	15	14	17
DYS 385 a/b	12-14	13-17	15-17	13-17	11-14	12-14	14-18	13-14	14-16
DYS 393	12	11	13	11	13	12	12	12	13
DYS 391	10	10	10	10	11	10	10	10	10
DYS 439	11	12	13	12	11	11	13	11	11
DYS 635	23	23	23	23	24	23	24	23	21
DYS 392	13	13	11	14	11	13	11	13	13
Y GATA H4	11	12	12	12	13	11	11	11	11
DYS 437	15	15	14	14	14	15	15	15	14
DYS 438	11	10	10	10	11	11	11	11	9
DYS 448	20	19	20	19	19	20	19	20	19

Table 3: Different haplotypes observed in village B with extrapolation results for a haplotype occurring 14 times in this dataset (Extrapolation of frequency of the most recurrent haplotype (14 times): 13.3%<p<48.9%).

In Village C, out of 41 male samples, 11 different haplotypes were observed. The most common haplotype appeared 15 times (36% of male samples) in this village and extrapolations (done at 99% confidence level) showed that this haplotype would be carried by at least 17.2% and at best 55.8% of the total males in this village. In addition, this haplotype originates from individuals from three different surnames. Results in Table 4 show haplotypes that are present more than once.

In Village D, out of 29 male samples, 14 different haplotypes were observed. The most common haplotype appeared 11 times (38% of male samples) and these haplotype originate from individuals from same the surname. Extrapolation statistics showed that this haplotype would be carried by at least 14.7% and at best 61.1% of the total males in this village. In any forensic case or human identification in this village, the current Y-STR analysis with 17 systems would not be useful to include any of the individuals in this village. Results in Table 5 show haplotypes that are present more than once.

Systems	15 Individuals	8 Individuals	5 Individuals	4 Individuals	3 Individuals
DYS 456	15	16	15	14	15
DYS 389I	14	13	14	12	13
DYS 390	23	22	24	23	23
DYS 389II	31	30	30	28	31
DYS 458	15	15	15	17	15
DYS 19	15	14	14	14	15
DYS 385 a/b	12-17	13-17	11-17	12-19	12-17
DYS 393	12	11	12	12	12
DYS 391	10	10	10	10	10
DYS 439	11	11	12	11	11
DYS 635	21	22	23	21	21
DYS 392	11	14	13	11	11
Y GATA H4	11	12	13	11	11
DYS 437	14	15	15	16	14
DYS 438	9	10	12	9	9
DYS 448	19	19	19	22	19

Table 4: Different haplotypes observed in village C with extrapolation results for a haplotype occurring 15 times in this dataset (Extrapolation of frequency of the most recurrent haplotype (15 times): 17.2%<p<55.8%).

Systems	11 Individuals	2 Individuals	2 Individuals	2 Individuals	2 Individuals	2 Individuals
DYS 456	16	12	14	15	15	16
DYS 389I	12	13	14	13	13	13
DYS 390	24	24	23	23	24	24
DYS 389II	28	30	32	30	29	30
DYS 458	16	15	17	17.2	19.2	16
DYS 19	13	14	15	14	15	14
DYS 385 a/b	11-14	11-14	13-17	13-19	11-18	18-18
DYS 393	12	12	12	12	12	14
DYS 391	11	10	10	10	11	10
DYS 439	13	13	11	12	12	11
DYS 635	24	23	21	21	21	20
DYS 392	13	13	11	11	11	13
Y GATA H4	13	12	11	11	11	12
DYS 437	15	15	15	14	14	14
DYS 438	12	12	9	10	10	11
DYS 448	20	19	19	20	20	20

Table 5: Different haplotypes observed in village D with extrapolation results for a haplotype occurring 11 times in this dataset (Extrapolation of frequency of the most recurrent haplotype (11 times): 14.7%<p<61.1%).

In Village E, out of 36 male samples, 11 different haplotypes were observed. The most common haplotype appeared 19 times (52% of male samples) and these haplotypes originate from individuals from two different surnames. Extrapolation statistics showed that this haplotype would be carried by at least 31.3% and at best 74.16% of the total males in this village. In any forensic case or human identification in this village, the current Y-STR analysis with 17 systems could not be useful to include any of the individuals in this village. Results in Table 6 show haplotypes that are present more than once.

In Village F, out of 25 male samples, 21 different haplotypes were observed. Themost common haplotype appeared 2 times (7% of male samples) and this haplotype originates from individuals from two different surnames. The results translate the low endogamy practices shown in this Christian Armenian Apostolic community where the rule is to prohibit consanguineous marriages especially between first cousins [37]. But most importantly, this community is newly settled in the village starting within the Armenian displacement one hundred years ago [38] Results in Table 7 show haplotypes that are present more than once.

Systems	19 Individuals	3 Individuals	2 Individuals	2 Individuals	2 Individuals	2 Individuals	2 Individuals
DYS 456	15	15	15	15	13	15	15
DYS 389I	14	13	12	13	12	14	14
DYS 390	23	23	24	23	24	23	23
DYS 389II	30	31	29	31	29	30	30
DYS 458	17.2	16.2	17	16.2	17	17.2	17.2
DYS 19	14	14	14	14	16	14	14
DYS 385 a/b	13-15	13-17	11-14	13-17	14-18	12-17	13-14
DYS 393	12	12	13	12	12	12	12
DYS 391	10	10	10	10	10	11	10
DYS 439	12	11	12	11	13	11	12
DYS 635	21	21	20	21	21	20	21
DYS 392	11	11	12	11	11	11	11
Y GATA H4	11	11	11	11	12	11	11
DYS 437	14	14	16	14	16	14	14
DYS 438	10	10	10	10	9	10	10
DYS 448	20	20	21	20	19	19	20

Table 6: Different haplotypes observed in village E with extrapolation results for a haplotype occurring 19 times in this dataset (Extrapolation of frequency of the most recurrent haplotype (19 times): 31.3%<p<74.16%).

In village G, out of 12 male samples, four different haplotypes were observed (Table 8). The most common haplotype appeared six times (50% of male samples). All individuals had the same "Family name"/ surname. It is worth noting that in the Lebanese society the surname is inherited from the paternal side only. These four different haplotypes differed in not more than one system with the difference most probably being due to a relatively recent mutation. Extrapolations (done at 99% confidence level) showed that the most common haplotype may be carried by at least 12.8% and at best 87.2% of the total males in this village. In any forensic case or human identification in this village, the current Y-STR analysis with 17 systems would not be useful to include any of the individuals in this village.

Y-STR haplotype diversity and match probability for 17 Y-STR systems in Lebanese villages

In general, profiles from villages have shown higher match probabilities when compared to the national population (Table 9). For example in village B, the haplotype frequency and match probability would increase by a three-fold when compared to the Lebanese

dataset. Based on the formula of 1-α1/N for a rare haplotype, the most common haplotype in a village E (19 times) when crossed with the Lebanese dataset, showed that one individual in every 170 would have the same profile; however, when crossed in the village dataset, the match probability is one individual match in every 2 sampled.

Systems	2 Individuals	2 Individuals	2 Individuals	2 Individuals
DYS 456	15	15	16	16
DYS 389I	12	12	13	13
DYS 390	22	24	24	24
DYS 389II	29	29	29	29
DYS 458	14	17	16	16
DYS 19	14	15	14	14
DYS 385 a/b	13-17	11-12	11-14	11-14
DYS 393	11	14	12	12
DYS 391	10	10	11	11
DYS 439	12	12	12	13
DYS 635	23	21	23	23
DYS 392	14	11	13	13
Y GATA H4	12	11	11	11
DYS 437	15	15	15	15
DYS 438	10	10	12	12
DYS 448	19	20	19	19

Table 7: Different haplotypes observed in village F.

Hence, due to endogamy and geographical genetic isolation, a haplotype may be common and abundant but would be unique in the total population. By testing whether the haplotype diversity in the villages matches the Lebanese population haplotype diversity and by setting a null hypothesis that both are equal, a chi-square test revealed a p value of 0.032. Since the p value is less than 0.05, this serves to reject the null hypothesis of equality and ascertains the difference between haplotype diversity observed and the haplotype diversity of the total population. A similar rational could be given for the discrimination capacity. Thus, the haplotype diversity (HD) and discrimination capacity (DC) were particularly low in comparison to the total population (Table 10).

Resolving common haplotypes using 23 Y-STR systems

Common haplotypes that were seen in the seven tested villages were further evaluated using 23 systems. In villages E and G, none of the common haplotypes was further differentiated. Even though the number of Y-STR systems was increased, the differentiation capacity was not enhanced. In the remaining five villages, the number of unique haplotypes increased when using the Y-23 Kit (Table 11).

Haplotype diversity and discrimination capacity were calculated for the Y-23 Kit in each village, and were compared to those obtained using the Y filer kit (Table 11). Results show that increasing the number of systems hasn't benefited the haplotype diversity or the discrimination capacity in villages E and G. In villages A, B, C, D, and F, the HD and DC improved with few individuals being differentiated.

Systems	6 Individuals	1 Individual	1 Individual	4 Individuals
DYS 456	15	15	15	15
DYS 389I	14	14	14	14
DYS 390	25	24	25	25
DYS 389II	32	32	32	32
DYS 458	17	17	17	17
DYS 19	14	14	14	14
DYS 385 a/b	16-17	16-17	17-17	16-17
DYS 393	13	13	13	13
DYS 391	10	10	10	10
DYS 439	12	12	12	12
DYS 635	21	21	21	22
DYS 392	11	11	11	11
Y GATA H4	11	11	11	11
DYS 437	14	14	14	14
DYS 438	10	10	10	10
DYS 448	19	19	19	19

Table 8: Different Y-STR haplotypes observed in village G with extrapolation results for a haplotype occurring 6 times in this dataset (Extrapolation of frequency of the most recurrent haplotype (6 times): 12.8%<p<87.2%).

Lebanese villages	Frequency in dataset	Rate of match in village	Rate of match in total population
Village A	9/53	1/6	1/170
Village B	14/45	1/3	1/170
Village C	15/41	1/3	1/170
Village D	2911	1/3	1/170
Village E	19/36	1/2	1/170
Village F	25/2	1/12	1/170
Village G	12/6	1/2	1/170

Table 9: Comparing a haplotype match probability in an isolated village dataset and the Lebanese dataset.

When looking closely at the new Y-STR loci found in the Y-23 kit, some systems showed to be more helpful than others in differentiating the haplotypes in some of the villages. The most discriminating were DYS576 and DYS570 which happen to be rapidly mutating Y-STRs (Table 12).

Population sample	Haplotype Diversity (HD)	Unique Haplotype (UHD)	Random Match Probability (RMP)	Discrimination Capacity (DC)
Village A	0.956	23 in 53 tested	0.043	58.40%
Village B	0.883	8 in 45 tested	0.117	37.70%
Village C	0.816	6 in 41 tested	0.184	26.80%
Village D	0.852	8 in 29 tested	0.148	48.20%
Village F	0.73	4 in 36 tested	0.269	30.50%
Village E	0.986	17 in 25 tested	0.014	84.00%
Village G	0.78	2 in 12 tested	0.22	33.30%
Lebanese population	0.9995	431 in 502 tested	0.0005	92.00%

Table 10: Statistical indices for the different tested villages and comparison to the Lebanese populations.

Villages	Systems	Statistical Indices			
		HD	UH	RMP	DC
Village A	17 Systems	0.956	23 out of 53	0.043	58.40%
	23 Systems	0.979	30 out of 53	0.021	69.80%
Village B	17 Systems	0.883	8 out of 45	0.117	37.70%
	23 Systems	0.965	13 out of 45	0.034	53.30%
Village C	17 Systems	0.816	6 out of 41	0.184	26.80%
	23 Systems	0.839	9 out of 41	0.161	34.10%
Village D	17 Systems	0.852	8 out of 29	0.148	48.20%
	23 Systems	0.899	10 out of 29	0.101	55.10%
Village E	17 Systems	0.73	4 out of 36	0.269	30.50%
	23 Systems	0.73	4 out of 36	0.269	30.50%
Village F	17 Systems	0.986	17 out of 25	0.014	84.00%
	23 Systems	0.993	21 out of 25	0.0007	92.00%
Village G	17 Systems	0.78	2 out of 12	0.217	33.30%
	23 Systems	0.78	2 out of 12	0.217	33.30%

Table 11: Comparison of HD, DC using 17 and 23 systems in Lebanese villages.

New Y-STRs	Number of Haplotypes Differentiated using the new Y-STR systems							
	Village A	Village B	Village C	Village D	Village E	Village F	Village G	Total
DYS576	3	1	0	1	0	0	0	5
DYS570	0	2	2	0	0	1	0	5
DYS533	1	1	1	0	0	0	0	3
DYS549	1	1	0	1	0	1	0	4
DYS481	1	0	1	0	0	0	0	2
DYS643	0	1	0	0	0	1	0	2

Table 12: Number of haplotypes differentiated using the new Y-STRs found in the Y-23 Kit.

Conclusion

Knowing that endogamy is a common practice in the Lebanese population, it was necessary to assess its direct effect on Y-STR analysis and haplotype frequencies, rate of haplotype occurrence and match probability. Results showed that high co-ancestry rate has implications on the interpretation of forensic DNA evidence. Match probability increased in geographically isolated regions and in villages with high rates of endogamous marriages. The rate of common haplotypes between non-related male individuals was much higher in the isolated villages than in the entire Lebanese dataset. Hence, the match probability differed when comparing the national database with the village's database which raises the question which dataset to use in case of a match profile in a forensic case.

The current number of Y-STR markers of 23, is still not enough to discriminate between all individuals of the same geographical area whereby the discrimination capacity and haplotype diversity showed a significant decrease ($p=0.032$) comparatively to the total population with the discrimination capacity dropping by more than one fold in some villages. This could cause potential problems in the field of forensic casework or kinship testing in Lebanese villages and rural areas.

Results of the Y-23 Kit showed that the most helpful STRs were those that are rapidly mutating. This can pave the way for further studies using the 13 RM-YSTR to check whether the haplotype diversity would increase and consequently reduce the match probability in these highly endogamous villages. An in house 13 RM-YST kit has been previously tested and shown high rate of discrimination among non-related and related male individuals [39,40]. Moreover, in Forensic cases, the increased number of Y-STR would definitely help in increasing the exclusion rate and reduce the level of wrongful accusations in the court of law [41].

With the high rate of endogamy at the geographical level, almost half of the inhabitants of this specific area as shown in this study may have the same Y haplotype and eventually would match the Y profile from the evidence sample, which might be challenging specially when autosomal STRs are not being helpful and discriminatory [7].

In kinship cases, more than one individual might statistically appear as the biological parent of the child [7]. Although in these cases the use of lineage markers would be highly recommended, yet in the Lebanese situation and due to the low haplotype diversity in rural areas, this might further increase the problem.

In conclusion, this study assessed the impact of endogamy on Y-STR profiles in societies where the endogamous rate is high and showed poor discrimination power of Y-STR where it reached as low as 26.8% which is extremely problematic for forensic casework in such villages. For example, knowing that Lebanese villages form more than half of the Lebanese population, a sexual assault case in these populations would be limited by the power of Y-STR discrimination capacity where it would be nearly impossible to positively include the right perpetrator. Hence, further studies are needed to attempt and reach higher discrimination power for Y-STR analysis to reduce complications in match probabilities.

Acknowledgment

The authors wish to thank and acknowledge: all the Lebanese volunteers from the different studied villages for participating in this study, Applied Biosystems and Promega for granting and supporting this project.

References

1. Gill P, Jeffreys AJ, Werrett DJ (1985) Forensic application of DNA 'fingerprints'. Nature 318: 577–579.
2. Prinz M, Boll K, Baum H, Shaler B (1997). Multiplexing of Y chromosome specific STRs and performance for mixed samples. Forensic Sci Int 85: 209–218.
3. Honda K, Roewer L, De Kniff P (1999). Male DNA typing from 25-year-old vaginal swabs using Y Chromosomal STR polymorphisms in a retrial request case. J Forensic Sci 44: 868-872.
4. Betz A, Bassler G, Dietl G, Steil X, Weyermann G, et al. (2001) DYS STR analysis with epithelial cells in a rape case. Forensic Sci Int 118: 126–130.
5. Prinz M, Ishii A, Coleman A, Baum HJ, Shaler RC (2001) Validation and casework application of a Y chromosome specific STR multiplex. Forensic Sci Int 120: 177–188.
6. Parson W, Niederstatter H, Kochl S, Steinlechner M, Berger B (2001) When autosomal short tandem repeats fail: optimized primer and reaction design for Y-chromosome short tandem repeat analysis in forensic casework. Croat Med J 42: 285-287.
7. Abbas S, Mourad L, Mansour I (2017) Inbreeding effect on forensic investigations involving DNA mixtures: The Lebanese population case. Forensic Sci Int 277: 1-257.
8. Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, et al. (1996) Analysis of molecular variance (AMOVA) of Y--chromosome--specific microsatellites in two closely related human populations. Hum Mol Genet 6: 1029–1033.
9. Redd AJ, Clifford SL, Stoneking M (1997) Multiplex DNA typing of short-tandem-repeat loci on the Y chromosome. Biol Chem 378: 923-927.
10. Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ (2002) A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci Int 129: 10–24.
11. Martin P, Albarran C, Garcia O, Garcia P, Sancho M, et al. (2000) Application of Y-STR analysis to rape cases that cannot be solved by autosomal STR analysis. Prog in Forensic Genet 8: 526-528.
12. Corach D, Risso LF, Marino M, Penacino G, Sala A (2001) Routine Y-STR typing in forensic casework. Forensic Sci Int 118: 131–135.
13. King TE, Jobling MA (2009) What's in a name? Y chromosomes, surnames and the genetic genealogy revolution. Trends Genet 25: 351-360.
14. Dekairelle AF, Hoste B (2001) Application of a Y-STR-pentaplex PCR (DYS19, DYS389I and II, DYS390 and DYS393) to sexual assault cases. Forensic Sci Int 118: 122–125.
15. Lee HS, Lee JW, Han GR, Hwang JJ (2000) Motherless case in paternity testing. Forensic Sci Int 114: 57–65.
16. Gonzalez-Andrade F, Sanchez D, Penacino G, Jarreta BM (2009) Two fathers for the same child: A deficient paternity case of false inclusion with autosomic STRs. Forensic Sci Int Genet 3: 138–140.
17. Butler JM (2003) Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. Forensic Sci Rev 15: 91–111.
18. Andari AE, Mansour I (2013) Is he or is he not the father? Pros, cons effect of following standardized interpretations. Forensic Sci Int Genet 4: 186-187.

19. Mansour I, Andari AE (2017) Effect of motherless paternity cases on the interpretation of parentage investigations in a population with recurrent inbreeding practices. J Forensic Res 8: 1-4.
20. Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110: 118–124.
21. Prinz M (2003) Advantages and disadvantages of Y-short tandem repeat testing in forensic casework. Forensic Sci Rev 15: 189-195.
22. Diegoli TM (2015) Forensic typing of short tandem repeat markers on the X and Y chromosomes. Forensic Sci Int Genet 18: 140–151.
23. Zhivotovsky LA, Ahmed S, Wang W, Bittles AH (2001) The forensic DNA implications of genetic differentiation between endogamous communities. Forensic Sci Int 119: 269–272.
24. Klat M, Khudr A (1986) Religious endogamy and consanguinity in marriage patterns in Beirut, Lebanon. Soc Biol 33: 138–145.
25. Barbour B, Salameh P (2009) Consanguinity in Lebanon: Prevalence, distribution and determinants. J Biosoc Sci 41: 505–517.
26. Bittles AH (1994) The role and significance of consanguinity as a demographic variable. Popul Dev Rev 20: 561–584.
27. Bittles AH (1995) When cousins marry: A review of consanguinity in the Middle East. Perspect Hum Biol 1: 71–83.
28. Al-Azem M, El Andari A, Mansour I (2017) Estimation of allele and haplotype frequencies for 23 YSTR markers in the Lebanese population. Forensic Res Criminol Int J 5: 1-7.
29. AmpFlSTR® Yfiler™ PCR amplification Kit. User's manual. Applied Biosystems, USA.
30. Powerplex® Y23 system for use on the applied Biosystems® genetic analyzers Kit Technical Manual. Promega Corporation, Madison, USA.
31. Nei M (1987) Molecular Evolutionary Genetics. (1st edn), Columbia University Press, New York, USA.
32. Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, et al. (1997) Evaluation of Y-chromosomal STRs: A multicenter study. International Int J Legal Med 110: 125–133,141–149.
33. Schoske R, Vallone PM, Kline MC, Redman JW, Butler JM (2004) High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays. Forensic Sci Int 139: 107–121.
34. Budowle B, Adamowicz M, Aranda XG, Barna C, Chakraborty R, et al. (2005) Twelve short tandem repeat loci Y chromosome haplotypes: genetic analysis on populations residing in North America. Forensic Sci Int 150: 1–15.
35. Budowle B, Jianye G, Chakraborty R (2007) Basic principles for estimating the rarity of Y-STR haplotypes derived from forensic evidence. In: Proceedings of the Eighteenth International Symposium on Human Identification.
36. Willuweit S, Roewer L (2007) Y chromosome haplotype reference database (YHRD): Update. Forensic Sci Int Genet 1: 83–87.
37. Bedoyan H (1979) The social, political, and religious structure of the Armenian community in Lebanon. The Armenian Review 32: 119-130.
38. Harfouche JK (1965) Social structure of low-income families in Lebanon. Khayats; (1st edn), Jamaica, Massachusetts.
39. Alghafri R, Goodwin W, Hadi S (2013) Rapidly mutating Y-STRs multiplex genotyping panel to investigate UAE population. Forensic Sci Int: Genet Suppl Series 4: 200-201.
40. Zgonjanin D, Alghafri R, Almheiri R, Antov M, Toljic D, et al. (2017) Mutation rate at 13 rapidly mutating Y-STR loci in the population of Serbia. Forensic Sci Int: Genet Suppl Series 6: 377-379.
41. Hampikian G, Peri G, Lo SS, Chin MH, Liu KL (2017) Coincidental inclusion in a 17-locus Y-STR mixture, wrongful conviction and exoneration. Forensic Sci Int: Genet 31: 1–4.