IASR
INTERNATIONAL ASSOCIATION OF SCIENTISTS & RESEARCHERS

Xournals

Academic Journal of Forensic Sciences

ISSN: 2581-4273 | Volume 07 | Issue 01 | April-2024

Impact of Linkage and Mutations on STR Typing and Interpretation of Results

Nadeem Mubarik[1]

Available online at: www.xournals.com

Received 1st April 2024 | Revised 9th April 2024 | Accepted 22nd April 2024

Abstract:

DNA typing is emerging as an important tool for boosting criminal justice delivery system of the whole world. The technique has made tremendous progress in the recent past. With the progress of time all possible efforts have been made to develop better multiplex systems to cater with new challenges. Efforts are made to include the impact of linkage and linkage disequilibrium in calculating likelihood ratios when dealing with close relations. Efforts are also underway to increase the strength of databases by making intensive and extensive studies covering different populations of the world. Extensive research is being carried out to get better information regarding the mechanism and marker specific mutation rates in STRs. Computer based models are designed to get better insights in calculating the impact of all the factors that have a potential to disturb the likelihood calculations. This chapter is designed to acquaint the reader about all these concepts.

Keywords: *STR typing, Linkage, Mutation, Recombination*

Authors:

1. *DNA Division, Forensic Science Laboratory, Jammu and Kashmir, Srinagar, INDIA.*

Introduction

The experiments performed by Gregor Johnn Mendel during 1860's on the garden pea plant laid a foundation of understanding the basic principles of genetics, commonly known as Medalian genetics. Based on simple pedigree analysis it was made clear that the characters are inherited from parents to their off springs. During his studies, Mendel postulated some basic principles like Dominance, Segregation and independent assortment of characters. In the modern era, the study of inheritance performed at the molecular level is based on some seemingly esoteric concepts, which of course need to be resolved at the outset so that the concept is understood as a whole. The human genome is composed of millions of nucleotides and only a small portion of it is utilized as a tool in differentiating individuals from each other, first done by Sir Aliec Jaffrey in 1984 **(Jeffreys *et al.*, 1985).** The technique however remained less appreciated and was mostly employed in demystifying the genetic disorders due to malfunction in some of the genes in the close proximity of the repetitive sequences. Soon after the value of genetic diversity among these repetitive sequences was recognized, more and more research started focusing on recognizing the potential hotspots for better variability and stability. The invention of Polymerase Chain Reaction in 1985 by Kary B. Mullis proved a milestone as the technique was based on a powerful concept of generating multiple copied from the desired segments of DNA. The invention was unequivocally a major breakthrough to defeat the constraints of low template DNA, met mostly in forensic samples. The quest of designing better and more reliable hotspots having better polymorphic characters and greater stability, more genetic markers were added in order to increase the discrimination power. Consequently, a huge number of markers have already found there place in various multiplex kits. Even the markers from same chromosome, having impact of linkage are being selected which of course are going to actually disturb the calculated value of evidence, if the linkage factor is ignored. In this chapter we shall try to evaluate the effect of two important concepts –Linkage and Mutations on STR typing and interpretation of results. However, before going in to the discussing precisely, there are certain basic concepts which need to be understood first especially for the beginners.

A diploid cell undergoes a series of sequential modifications before giving rise to a germ cell. The most important one of these changes takes place during prophase-I of meioses where the homologous chromosomes align to undergo exchange of segments during recombination. In human cells (whether male or female), the two homologs align adjacent to each other and take part in the formation of chiasmata during crossing over except sex chromosomes in human males, that will be discussed later. The autosomal (AS) markers have two alleles at one locus, of which one is from father (paternal) and the other one is from mother (maternal). After a successful recombination event, the chromosomes are separated towards opposite poles of the cell during anaphase-I. In AS chromosomes recombination serves as a driving factor towards polymorphism bringing in more and more combinations. Same is the case with sex chromosomes in human females (XX) which also undergo recombination leading to formation of recombinants. However, the Y-chromosomes, in human males do not have a counterpart to pair with and hence there is no recombination between XY. The only X chromosome in male child which has been contributed by the mother is a recombinant one while in case of a female child, one X chromosome contributed by the mother is a recombinant whereas the other one contributed by the father is a non-recombinant.

The genetic information stored in the DNA is distributed among 46 chromosomes having 44 autosomes and 02 sex-determining chromosomes, in humans. Only about 0.1 percent of this genetic information is found to be polymorphic, of which Short Tandem Repeats (STRs) are the most targeted hotspots in the contemporary forensic DNA typing. Short Tandem Repeats (STRs), are simple sequence repeats having a repetitive unit of 1-6 bp **(Tautz, 1993).** These repetitive sequences are scattered in the genome of both Prokaryotes as well as Eukaryotes including human. Most of these STR sequences are found in non-coding parts and only a small portion in coding part **(Ellegren, 2000).** These STRs are commonly found to be A-rich and their concentration is mostly on chromosome-19. On an average, in humans there is one STR sequence per 2000bp **(www.genome.gov)** There is a tendency for selection of these STR markers based on their polymorphic potential coupled with greater stability and reduced allele spread, as far as their role in forensics is concerned. A marker with greater discriminating power, low mutation rate and reduced length based alleles at a marker are the markers of choice. The reduced length based allele spread is desired because when there is less size difference between two heterozygous alleles, there are less chances of ambiguity in mixture analysis **(Walsh *et al.*, 1992).** A large number of autosomal STR markers have been selected for this purpose over the last three decades; however there are only a limited number of these markers which can actually pour in greater variability as far as the STR profile is concerned. The second generation multiplex (SGM) having TH01, VWA,

FGA, D8S1179, D18S51 and D21S11 markers provided a match probability of about one in five million. Introduction of other highly polymorphic markers like SE33, D2S1338, D18S51, etc. have added a great deal to the level of polymorphism. (Table 1.1) gives the probability of identity, which is the pair wise comparison of genotypes at a locus. Markers having lower probability of identity are more variable and hence more valuable.

Linkage

A concept given by Morgan, states that when markers tend to remain together during their inheritance they are said to be linked. Markers that are physically near to each other are unlikely to be separated by the recombination event during crossing over, thus are said to be linked. When linkage plays a role, the results will differ from, as predicted by the law of independent assortment. The distance between any two loci is measured in centimorgans (cM) which is defined as the distance between markers in which there is one percent (1%) chance of one marker being separated from other as a result of recombination in one generation. 1 cM is equal to recombination of one percent. Markers with recombination frequency of less than 50% are believed to be present on same chromosome and are considered as linked, whereas a recombination frequency of 50% depicts that the markers are present on non-homologous chromosomes or, are far apart on the same chromosome and hence considered as unlinked. Therefore, recombination frequency gives an idea of the level of genetic linkage between the markers, and can be calculated as;

Recombination frequency (θ) = (Number of recombinant progeny)/(Total number of progeny) X 100

In STR typing, especially while dealing with close relationship studies the impact of linkage is obvious but to calculate exact likelihood, we must be able to assess the level of impact also. This subject will be dealt with, in detail in the proceeding discussions.

Linkage disequilibrium (LD)

Also known as allelic association, is another phenomenon that needs to be considered while calculating LR. LD is said to exist when alleles of different loci tend to remain together in a population study more often than expected. It reflects recombination in all generations since the evolution of the allele under study. In other words when the probability of the occurrence of two alleles together (P AB) is not equal to their individual probabilities (PA PB), believing the association as completely random, LD is said to exist. LD is influenced by the factors like demographic history, Admixtures, etc. **(Ardlie *et al.*, 2002; Laan and Pääbo, 1997; Zavattari *et al.*, 2000).**

Impact of linkage on STR typing

No doubt forensic evidence at the scene of crime is a potent and robust source of information linking crime with criminal, but its weight as an evidence is determined by the assessment of associated factors. In order to bring in more and more variability between any two DNA profiles, addition of more STR markers is made based on basic characters of marker designing. Although it is preferable to use markers from different chromosomes, however modern and emerging multiplex kits (Table 1.2) are using markers close to each other on same chromosomes which intern raises the question of allelic association. In forensic estimation of relatedness among closely related individuals, linkage becomes important in calculating probability of passing an allele from father to child. If the LR calculated involves H1, suspect as donor and H2 as an unrelated person to the suspect, as donor then there is no impact of linkage on LR calculations. However, if the alternative hypothesis accepts a close relative of the suspect as the donor, the linkage will undoubtedly have to be taken in to consideration while calculating LR **(Buckleton and Chris, 2006).**

$$LR=H1/H2$$

Same is the case with impact of linkage on paternity Trio/Duo cases. When the alternative hypothesis considers an unrelated person as the alleged father, linkage will have no impact on calculations. However, when the alternative hypothesis involves any close relative as the alleged father, then the linkage will come in to play and needs to be considered while interpretation of results. (Table 1.2) gives a list of AS STR markers which are considered in forensic multiplexes despite the fact that there are more than one markers selected from same chromosome **(Gill *et al.*, 2012).**

As we know the distance between any two markers in a genetic map has a direct impact on recombination frequency and thus on linkage. Therefore, the impact of linkage increases as the distance between the markers decreases.

An alternative approach as recommended by Budowle et al. **(Budowle *et al.*, 2011)** for kinship analysis while having impact of linkage, is to either incorporate the recombination rate for generating maximum likelihood estimates of haplotype frequencies for two

loci or opt for the most informative loci, out of the two in the calculations.

Table No. 1.1: Rank wise list of AS STR markers chosen for their variability (Butler *et al.*, 2009; strbase.nist.gov).

Marker	Rank	Probability of Identity	Number of alleles
SE33	1	0.0066	120
D2S1338	2	0.0220	22
D1S1656	3	0.0224	16
D18S51	4	0.0258	55
D12S391	5	0.0271	18
FGA	6	0.0308	98
D6S1043	7	0.0321	-
D21S11	8	0.0403	92
D8S1179	9	0.0558	19
D19S433	10	0.0559	30
vWA	11	0.0611	29
D7S820	12	0.0726	32
D16S539	13	0.0749	21
D13S317	14	0.0765	20
TH01	15	0.0766	21
D2S441	16	0.0841	11
D10S1248	17	0.0845	12
D3S1358	18	0.0915	27
D22S1045	19	0.0921	12
CSF1PO	20	0.1054	23
D5S818	21	0.1104	16
TPOX	22	0.1358	17

Table No. 1.2: List of AS STR markers chosen, despite an inter-marker distance Constraint, for forensic use (Gill *et al.*, 2012)

Location on chromosome	Locus pair	Distance in between
5q 23.2	D5S818	25 cM **(6)**
5q 33.1	CSF1PO	
2q 35	D2S1338	50cM **(6)**
2q25.3	TPOX	
21q21.1	D21S11	50cM **(7)**
21q33.1	Penta D	
2q35	D2S1338	**Located on different arms**
2p14	D2S441	
12p13.31	vWA	12cM **(8)**
12p12	D12S391	

Linkage in X STRs

Unlike autosomes, the X-chromosomes in males do not occur in pairs and thus cannot form homologs. Here we have only one X-chromosome and one Y-chromosome. In case of human females, there are two X-chromosomes forming a homologous pair which undergoes recombination during meioses. X-chromosomal STR markers have gained a special scope in forensic science especially in kinship cases where the alleged father cannot be traced and the child is a female **(Gusmão *et al.*, 2012; Szibor *et al.*, 2007)**. Although while dealing with standard paternity cases like duos/trios, it is highly desirable that AS STRs should be preferred. However, there are certain conditions when the use of X STRs becomes essential, either as additional or exclusive. For instance, in case of few genetic inconsistencies between the alleged father and the daughter, the inconsistencies are explained as the result of mutation or the alternative hypotheses is considered as the alleged father being genetically close to biological father **(Gomes *et al.*, 2012).** If the biological father is the father of the alleged father, with different mothers (Fig. 1.1) or if the biological father is the son of the alleged father, with different mothers (Fig. 1.2), they are not going to share X STR alleles, identical by descent (IDB).

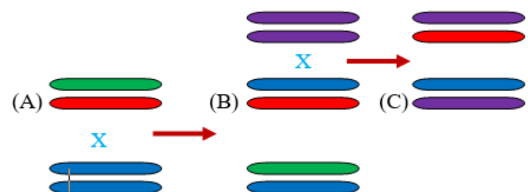

Figure No. 1.1: "A" the biological father is the father of the alleged father "B". "C" is the daughter in issue. "A" and "B" do not share the x-chromosome

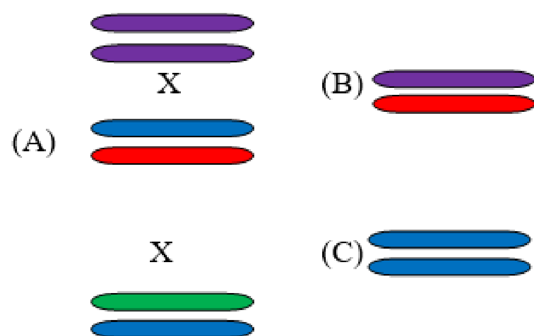

Figure No. 1.2: "A" being the biological father of "C" and "A",the alleged father. "A" and "B" do not share the X-markers

X STR typing is also important when there are no direct sample sources for matching the profile in a deficiency case. As discussed earlier, the sisters inherit partially matching haplotypes from the mother. In these cases when the putative father is not available, the mother of the putative father (grandmother of female child) is more informative and Mean Exclusion Chance (MEC) can be generated **(Krüger *et al.*, 1968).** which is a chance of excluding a man who is not the father by an inconsistency in at least three loci, considering mutations.

Here it must be remembered that these complications arise for AS STRs while dealing with close genetic relations. For example, in complete siblings the probability of having same allelic combination at any marker is ¼ (0.25). This means there are chances that the fifth child born from the same parents is expected to have the same combination of alleles as any one of the other four. With the increase in the number of markers used, of course the probability of repetition of same combination will go on decreasing as a result the AS STR profile gives a better conclusion and the authenticity of the profile generated increases (Fig.1.3). Nonetheless, due to low variation in between closely related individuals like siblings, it is preferable to opt for X STR typing when the above discussed scenarios are faced.

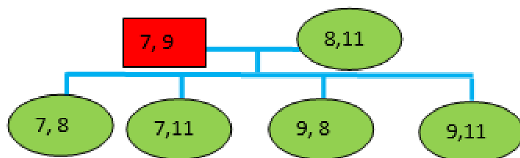

Figure No. 1.3: Using any marker in a full sibling case, a probability of 0.25 exists, that offspring beyond the four given in the figure, will repeat the genotype of any of the four given

Using X-STR markers is more authentic than AS STRs in such cases because of their inheritance pattern. The males contain one X-chromosome (XY) and hence the haplotype is transmitted from biological father to biological daughter without undergoing any recombination. However, in case of females there are two X-chromosomes (XX) and thus the recombination is possible. Therefore, while considering X STR markers for relationship tests two important factors need to be considered while interpreting the results. These are linkage (inheritance of closely related markers together with higher probability than physically separated markers) and linkage disequilibrium (At a population level, occurrence of alleles of different loci together, more or less frequently than expected by chance).

We know that recombination is a direct measure of linkage between any two markers. A recombination frequency of less than 50% (50 cM) is considered as the indication of linkage between any two markers while a 50% recombination frequency is an indication that the two markers are either present on different chromosomes or are far enough on the same chromosome to assort independently. While dealing with X-chromosome with an approximate length of about 155 Mb (180 cM) **(International HapMap Consortium *et al.*, 2007).** only 3-4 STRs can be considered which are at least separated by a minimum distance of 50 cM, required to ensure their independent assortment. However, additional markers are introduced (**Szibor *et al.*, 2007**). which can greatly diminish the results, if linkage factor is not taken in to consideration. It is therefore recommended that while dealing with X STRs in kinship analysis the guidelines put forth by DNA commission for International Society for Forensic Genetics (ISFG) must be adhered to, before interpreting the results. The recommendations are summarized as under **(Tillmar *et al.*, 2017).**

Guidelines put forth by DNA commission for International Society for Forensic Genetics (ISFG)

1. In paternity cases (duo's or trio's, with a daughter), X STR analysis should be used to supplement DNA testing results when the information obtained from standard autosomal markers is inconclusive, such as may be observed in paternity cases with few genetic inconsistencies.
2. X-chromosomal markers should be used in specific kinship cases when the exclusion power does not equal null in contrast to the autosomal markers examined. Important examples include full or half sibling duos involving two females and paternal grandmother/granddaughter duos. Furthermore, X-chromosomal markers should be used in situations where two alternative hypotheses posses the same likelihood for AS markers but are expected to differ when X-chromosomal markers are examined. X-chromosomal analysis may also help to distinguish possible related fathers in incest cases.
3. Prior to using a X- chromosomal assay or commercial kit, markers should be evaluated to determine whether or not they are linked. Recombination rates should primarily be estimated from family studies or secondarily by mapping functions based on genetic distance. A recombination rate below 0.5 indicates linkage.
4. Linkage should be accounted for when calculating LR given that the X-chromosomal markers are linked and that linkage will have an impact on the final LR. This also includes accounting for recombination events within a cluster of X-chromosomal markers, known as linkage group.
5. Linkage equilibrium tests should be performed when generating population frequency data for the markers in a X-chromosomal marker multiplex.
6. X-chromosome markers that are located closely to each other and not in linkage equilibrium should be reported as haplotype frequencies rather than single locus frequencies for population databasing.
7. Haplotype frequencies should be used for likelihood calculations when LD exists.
8. Appropriate software should be used when calculating LR's based on X-chromosomal markers in kinship analysis to avoid manual calculation errors. The software should rely on likelihood calculations and should be able to accommodate linkage, linkage disequilibrium and mutations.
9. As far any other software calculating likelihood ratios to evaluate competing kinship scenarios, use of software for X-chromosome applications should follow the recommendations from the DNA commission of the ISFG on the validation of software programs.
10. Individual autosomal LR and X-chromosomal LR results should only be combined whenever equivalent (and clearly) defined hypothesis are used for both autosomal and X-chromosomal data, and when it is appropriate to assume that substructure and LD between autosomal and X-chromosomal alleles do not play a role.

Mutations and their impact on the interpretation of STR results

Although STRs are wide spread in both prokaryotes as well as in eukaryotes, they are mostly considered as junk DNA having no biological function. However, some recent advancements have revealed that in fact these STRs play an important role in many organisms **(Escher *et al.*, 2000; Gebhardt *et al.*, 1999; Meloni *et al.*, 1998).** As discussed earlier, an STR marker is selected considering a basic set of characters and a novel polymorphic STR marker is the one with lowest mutation rate. However, this criteria is not immune everywhere and in case of Y STR study, the role of mutations gains more focus and plays an important role in various cases. The normal DNA sequences

show a mutation rate of about 10-9 per generation as compared to the mutation rate of STRs, 10-6-10-2, which is several orders greater than the normal DNA sequence mutation rate **(Ellegren, 2000).** These mutations and more precisely, the mutation rates play an important role as far as the interpretation of results of a DNA profile is concerned. However, before going in to this discussion there are certain facts based on research, which should be taken in to account while framing the results.

- Dinucleotide repeats are found to be showing greater mutation rates as compared to tetranucleotide repeats **(Chakraborty *et al.*, 1997).**
- Certain sequences like GC content have strong role in determining rate of STR mutations **(Gymrek *et al.*, 2017).**
- Mutations having negative effect on the fitness are going to be eliminated from the population and recombination has an effect on doing away with these mutations through recombination repair. Therefore, it becomes evident that these recombination events may have an impact on the amount of mutations that are carried on downward and thus on the interpretation of results especially in paternal biogeographic ancestry. For example, Y-chromosomal markers do not undergo recombination and hence mutations in these markers will get transmitted in to the offsprings 100 %. This feature of Y-chromosomes makes it fit for patrilineal studies and tracing of ancestry **(Kayser, 2017).**

STR mutation rate in case of AS, and Y-chromosomes falls in between 1-5x10-3 **((Dupuy *et al.*, 2004; Decker *et al.*, 2008).** and similar results are produced for X-chromosome STRs, however this needs extensive studies covering different populations to generate a concrete database. Study of mutations and their impact thereof, in STRs is probably the field of future research. (Table 1.3) gives the rate of mutation at the commonly used AS STR markers with TPOX and TH01 showing lowest rate while as the most polymorphic markers like FGA and SE33 showing the highest rate. (Table 1.4) gives the mutation rates at 15 X-STR markers.

At present, there are three possible mechanisms, which have been proposed for STR mutations. Although believed to have least effect, unequal crossover during meioses is considered one of the factors **(Huang *et al.*, 2002).** Similarly generation of A-rich STR's through 3′ extension of retrotranscripts by a mechanism known as retrotransposition **(Nadir *et al.*, 1996).** However, the most widely accepted mechanism for STR mutations is "Strand Slippage Model", proposed by Kornberg et al in 1964 **(Kornberg *et al.*, 1964).**

The DNA polymerase during replication uses the information of the template segment to generate a complementary sequence. Since in STRs, each repeat unit is the same, e.g. (GATA) x the DNA polymerase cannot differentiate between the succeeding units and thus there are ample chances that some repeat units may be skipped. Sometimes the DNA polymerase and the daughter strand complex temporarily dissociate and due to the same repeat unit ahead, the pairing resumes and annealing occurs normally, as a result the left out segment forms a loop. Now, if this loop formation occurs on the newly synthesized strand during replication, the synthesized strand (daughter strand) will be, the number of repeat units which were skipped, longer than the template strand and vice versa. (Fig.1.4 A, B). It is however believed that there is difference in the apparent rate of strand slippage and effective mutations **(Schlötterer and Tautz, 1992).** which is believed to be done by the strand specific mismatch repair system, reducing the effective mutation rate by about 100-1000 fold **(Strand *et al.*, 1993).**

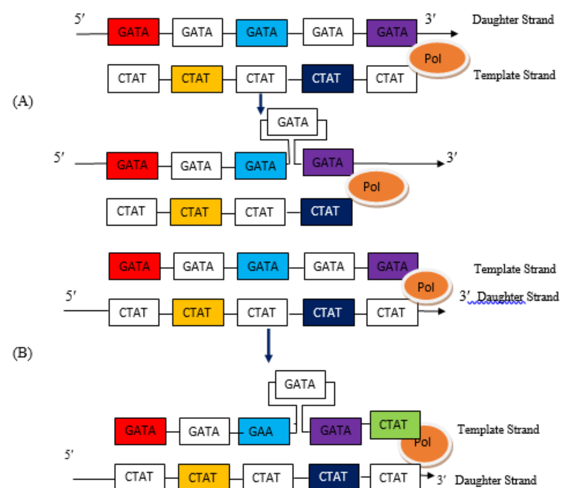

Figure No. 1.4: Mechanism of Strand Slippage (A) the formation of loop occurs in the daughter strand and hence it will be one repeat unit more as compared to the template. (B) formation of loop occurs in the template strand and hence the new strand will be one repeat unit short as compared to template strand.

Although AS STRs are the markers of choice due to their high polymorphism, there are certain conditions under which the study of Y STRs becomes essential. For instance in case of multisource samples (as in gang

rape), the concentration of source sample does not have uniform contribution from all the contributors. It is likely that the minor contributor will go scot free, using AS STRs. Similarly during rape, the slide prepared from the vaginal swab will no doubt contain the male contribution but, it is minor as compared to female background contribution. As a result there is serious competition for primer binding during PCR run with least chances of it (minor contribution) being primed. In order to do away with this, Y STR profiling becomes essential, first employed in 1992) **(Roewer and Epplen, 1992)** and considered three times more suitable in identifying male contributors as compared to AS STRs **(Purps *et al.,* 2015).** As discussed in the introduction, the human females do not possess the Y-chromosomes, thus the 50% specificity stands achieved. The human males normally possess one X-chromosome and one Y-chromosome and the Y-chromosome has no homolog to pair with and thus the markers are transferred from father to the male child without undergoing recombination. When there is no recombination, there is no variation and the only source of variation in Y STRs is therefore the effective mutation rate. At present up to 27 markers are utilized in commercial kits for forensic use (Yfiler Plus by Thermo Fisher Scientific) **(Gopinath *et al.,* 2016).** One of the striking features of Y STRs, which makes them essentially important to take up different assignments in forensic case work, is that different Y STR markers have different mutation rates. By using standard STRs with low-medium mutation rates (one or few mutations per 1000 per locus) **(Goedbloed *et al.,* 2009),** paternity cases can be solved where the putative father is not available to provide a direct sample for comparison with the male child. In this case any male relative of the unavailable putative father can be used as a source of sample providing Y haplotype similar to putative father. It is worth mentioning here that the rapidly mutating Y STR (RM Y STR) cannot be utilized for paternity and kinship tests because the higher mutation rate in these RM Y STRs will sometimes confuse with inconsistencies at specific loci and intern trouble the probability calculations. However, these RM Y STRs have an important role to play where the target is to work out the individual identification when there are large number of male relatives traced from long patrilineal tree. It becomes very confusing when there is no difference in Y STR haplotype in a huge number of male individuals sharing same Y STR profiles. Therefore in these cases RM Y STRs, having a mutation rate in the order of 10-2 **(Ballantyne *et al.,* 2010)** (Table 1.5) are very informative.

Y STRs also play an important role in studying paternal bio-geographic history. In absence of recombination in Y STR markers, when a mutation has occurred it cannot be removed from the gene pool, producing differences between the individuals of different geographic regions through factors like genetic drift. The impact of mutations and the selection of markers, from a variety of Y STRs available, becomes an essential step as far their use in forensic science is concerned. However, there are some specific guidelines provided by ISFG which need to be taken in to consideration while reporting a Y STR test. Given that different Y STR markers are transferred together in to, offsprings (single haplotype), it is not desired to multiply the individual allelic frequencies rather, the haplotype frequency is assessed against a relevant population in fairly large database. Similarly, mutation rates must also be taken in to consideration especially while dealing with kinship cases **(Gill *et al.,* 2001) (Butler *et al.,* 2009).** The profile of male staff members working in the laboratory should be recorded against Y STRs employing kits meant for general use. If increased PCR cycles are used the contamination factor must also be given due weightage **(Gill *et al.,* 2001).**

Table No. 1.3: Gives an idea about the mutation rates at commonly used AS STR markers in the course of paternity testing (strbase.nist.gov).

STR Marker	Total Number Of Mutations	Mutation Rate (%)
TPOX	100/857481	**0.01**
TH01	100/779554	**0.01**
D7S820	1089/1085305	**0.10**
D5S818	1259/1107339	**0.11**
D16S539	1041/962239	**0.11**
D119S433	187/173490	**0.11**
D3S1358	1152/964288	**0.12**
D2S1338	262/225140	**0.12**
D8S1179	1239/899837	**0.14**
D13S317	1558/1103282	**0.14**
Penta D	57/41202	**0.14**
CSF1PO	1487/947425	**0.16**
Penta E	163/100030	**0.16**
vWA	2480/1437945	**0.17**
D21S11	1816/962096	**0.19**
D18S51	1746/790342	**0.22**
FGA	3125/1101006	**0.28**
SE33	330/51940	**0.64**

Table No. 1.4: Gives an idea about the mutation rates at 15 X-STR markers (Diegoli *et al.*, 2014)

Marker	Mutations	Meioses	Rate of Mutation (x10 $^{-3}$)
GATA165B12	0	2333	0-1.6
DXS7130	0	1375	0-2.7
DXS10147	0	1429	0-2.6
DXS6795	0	1375	0-2.7
DXS101	1	3909	0.01-1.42
GATA172D05	1	2251	0.01-2.5
GATA31E08	1	2502	0.01-2.2
DXS7423	2	4890	0.05-1.5
DXS7424	2	3180	0.08-2.3
DXS6803	3	2390	0.26-3.7
DXS9902	4	1833	0.59-5.6
DXS6789	5	4853	0.33-2.4
DXS8378	8	5257	0.66-3.0
HPRTB	8	5905	0.59-2.7
DXS7132	15	6340	1.3-3.9

Table No. 1.5: Mutation rate at different Y STR markers used in different kits with corresponding mutation rates and number of RM YSTRs (Gymrek et al., 2017)

Y STR marker	AmpFlSTR® Yfiler®	PowerPlex® Y23ª	Yfiler® Plusᵇ	RM Y-STR setᶜ	Mutation rate (%)
DYS19	✓	✓	✓		0.4
DYS385a/b	✓	✓	✓		0.17/0.37
DYS389I	✓	✓	✓		0.51
DYS389II	✓	✓	✓		0.34
DYS390	✓	✓	✓		0.11
DYS391	✓	✓	✓		0.28
DYS392	✓	✓	✓		0.06
DYS393	✓	✓	✓		0.17
DYS437	✓	✓	✓		0.11
DYS438	✓	✓	✓		0.06
DYS439	✓	✓	✓		0.35
DYS448	✓	✓	✓		0
DYS456	✓	✓	✓		0.46
DYS458	✓	✓	✓		0.8
DYS635	✓	✓	✓		0.35
Y-GATA-H4	✓	✓	✓		0.28
DYS481		✓	✓		0.46
DYS533		✓	✓		0.46
DYS549		✓			0.42
DYS570		RM YSR	RM YSR	RM YSR	1.19
DYS576		RM YSR	RM YSR	RM YSR	1.39
DYS643					0.11
DYS449			RM YSR	RM YSR	1.18
DYS460					0.58
DYS518			RM YSR	RM YSR	1.8
DYS627			RM YSR	RM YSR	1.19
DYF387S1a/b			RM YSR	RM YSR	1.55
DYS526a/b				RM YSR	0.23/1.21
DYS547				RM YSR	2.32
DYS612				RM YSR	1.41
DYS626				RM YSR	1.18
DYF399S1				RM YSR	7.75
DYF403S1a/b				RM YSR	3.06/1.14
DYF404S1				RM YSR	1.21
Total RM YSR		**2**	**6**	**13**	

Conclusion

Over the years, more focus is laid on considering different potent factors which have strong correlation with STR typing and a viewed impact on the interpretation of results. Therefore, it is desired that the research in vogue regarding factors like linkage, linkage disequilibrium and mutations and their impact on the interpretation of results considering different scenarios faced in the forensic field, need a special focus. While dealing with linkage and linkage disequilibrium it is always desired to use computer generated programs for calculating LR to make them error free, especially in case of X STR analysis. In fact there are many software available but most of them assume markers as unlinked which intern makes the interpretation less reliable. However, software like Merlin include linkage and LD in the calculations but leave out the effect of mutations during LR calculations. FamLinkX is one such advancement which includes the impact of mutations also while dealing with X STR markers. Designing kits for X STR markers is one good step forward. Investigator Argus X-12 PCR amplification kit by Qiagen, uses twelve markers simultaneously and amelogenin for gender identification, can be helpful in various cases. While in case of Y STR markers it is always desired to have a larger database making extensive and intensive studies especially for rapidly mutating Y STRs covering different populations. The largest and widely used Y STR haplotype reference database is the YHRD.

References:

Ardlie, Kristin G et al. “Patterns of linkage disequilibrium in the human genome.” Nature reviews. Genetics vol. 3,4 (2002): 299-309. doi:10.1038/nrg777

Ballantyne, Kaye N., et al. “Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications.” American Journal of Human Genetics, vol. 87, no. 3, Sept. 2010, pp. 341–53. https://doi.org/10.1016/j.ajhg.2010.08.006.

References:

Buckleton, John, and Chris Triggs. "The effect of linkage on the calculation of DNA match probabilities for siblings and half siblings." Forensic science international vol. 160,2-3 (2006): 193-9. doi:10.1016/j.forsciint.2005.10.004

Budowle, Bruce et al. "Population genetic analyses of the NGM STR loci." International journal of legal medicine vol. 125,1 (2011): 101-9. doi:10.1007/s00414-010-0516-7

Butler, John M., et al. "The Single Most Polymorphic STR Locus: SE33 Performance in U.S. Populations." Forensic Science International. Genetics Supplement Series, vol. 2, no. 1, Dec. 2009, pp. 23–24. https://doi.org/10.1016/j.fsigss.2009.08.173.

Chakraborty, R et al. "Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci." Proceedings of the National Academy of Sciences of the United States of America vol. 94,3 (1997): 1041-6. doi:10.1073/pnas.94.3.1041

Decker, Amy E et al. "Analysis of mutations in father-son pairs with 17 Y-STR loci." Forensic science international. Genetics vol. 2,3 (2008): e31-5. doi:10.1016/j.fsigen.2007.08.016

Diegoli, Toni M et al. "Mutation rates of 15 X chromosomal short tandem repeat markers." International journal of legal medicine vol. 128,4 (2014): 579-87. doi:10.1007/s00414-014-1016-y

Dupuy, B Myhre et al. "Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci." Human mutation vol. 23,2 (2004): 117-124. doi:10.1002/humu.10294

Ellegren, H. "Heterogeneous mutation processes in human microsatellite DNA sequences." Nature genetics vol. 24,4 (2000): 400-2. doi:10.1038/74249

Escher, D et al. "Conservation of glutamine-rich transactivation function between yeast and humans." Molecular and cellular biology vol. 20,8 (2000): 2774-82. doi:10.1128/MCB.20.8.2774-2782.2000

Gebhardt, F et al. "Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1." The Journal of biological chemistry vol. 274,19 (1999): 13176-80. doi:10.1074/jbc.274.19.13176

Gill, P et al. "DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs." Forensic science international vol. 124,1 (2001): 5-10. doi:10.1016/s0379-0738(01)00498-4

Gill, Peter et al. "An evaluation of potential allelic association between the STRs vWA and D12S391: implications in criminal casework and applications to short pedigrees." Forensic science international. Genetics vol. 6,4 (2012): 477-86. doi:10.1016/j.fsigen.2011.11.001

Goedbloed, Miriam et al. "Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpFlSTR Yfiler PCR amplification kit." International journal of legal medicine vol. 123,6 (2009): 471-82. doi:10.1007/s00414-009-0342-y

References:

Gomes, Cláudia et al. "Comparative evaluation of alternative batteries of genetic markers to complement autosomal STRs in kinship investigations: autosomal indels vs. X-chromosome STRs." International journal of legal medicine vol. 126,6 (2012): 917-21. doi:10.1007/s00414-012-0768-5

Gopinath, Siddhita et al. "Developmental validation of the Yfiler(®) Plus PCR Amplification Kit: An enhanced Y-STR multiplex for casework and database applications." Forensic science international. Genetics vol. 24 (2016): 164-175. doi:10.1016/j.fsigen.2016.07.006

Gusmão, Leonor et al. "Capillary electrophoresis of an X-chromosome STR decaplex for kinship deficiency cases." Methods in molecular biology (Clifton, N.J.) vol. 830 (2012): 57-71. doi:10.1007/978-1-61779-461-2_5

Gymrek, Melissa et al. "Interpreting short tandem repeat variations in humans using mutational constraint." Nature genetics vol. 49,10 (2017): 1495-1501. doi:10.1038/ng.3952

Huang, Qing-Yang et al. "Mutation patterns at dinucleotide microsatellite loci in humans." American journal of human genetics vol. 70,3 (2002): 625-34. doi:10.1086/338997

International HapMap Consortium et al. "A second generation human haplotype map of over 3.1 million SNPs." Nature vol. 449,7164 (2007): 851-61. doi:10.1038/nature06258

International Human Genome Sequencing Consortium Publishes Sequence and Analysis of the Human Genome. Genome.gov, www.genome.gov/10002192/2001-release-first-analysis-of-human-genome.

Jeffreys, A J et al. "Hypervariable 'minisatellite' regions in human DNA." Nature vol. 314,6006 (1985): 67-73. doi:10.1038/314067a0

Kayser, Manfred. "Forensic use of Y-chromosome DNA: a general overview." Human genetics vol. 136,5 (2017): 621-635. doi:10.1007/s00439-017-1776-9

KORNBERG, A et al. "ENZYMATIC SYNTHESIS OF DEOXYRIBONUCLEIC ACID, XVI. OLIGONUCLEOTIDES AS TEMPLATES AND THE MECHANISM OF THEIR REPLICATION." Proceedings of the National Academy of Sciences of the United States of America vol. 51,2 (1964): 315-23. doi:10.1073/pnas.51.2.315

Krüger, J et al. "Zur Verwendung des Polymorphismus der sauren Erythrocytenphosphatase bei der Vaterschaftsbegutachtung" [On the utilization of erythrocyte acid phosphatase polymorphism in paternity evaluation]. Deutsche Zeitschrift fur die gesamte gerichtliche Medizin vol. 64,2 (1968): 127-46.

Laan, M, and S Pääbo. "Demographic history and linkage disequilibrium in human populations." Nature genetics vol. 17,4 (1997): 435-8. doi:10.1038/ng1297-435

Meloni, R et al. "A tetranucleotide polymorphic microsatellite, located in the first intron of the tyrosine hydroxylase gene, acts as a transcription regulatory element in vitro." Human molecular genetics vol. 7,3 (1998): 423-8. doi:10.1093/hmg/7.3.423

References:

Nadir, E et al. "Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications." Proceedings of the National Academy of Sciences of the United States of America vol. 93,13 (1996): 6470-5. doi:10.1073/pnas.93.13.6470

NIST STRBase. strbase.nist.gov.

Purps, Josephine et al. "Validation of a combined autosomal/Y-chromosomal STR approach for analyzing typical biological stains in sexual-assault cases." Forensic science international. Genetics vol. 19 (2015): 238-242. doi:10.1016/j.fsigen.2015.08.002

Roewer, L, and J T Epplen. "Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work." Forensic science international vol. 53,2 (1992): 163-71. doi:10.1016/0379-0738(92)90193-z

Schlötterer, C, and D Tautz. "Slippage synthesis of simple sequence DNA." Nucleic acids research vol. 20,2 (1992): 211-5. doi:10.1093/nar/20.2.211

Strand, M et al. "Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair." Nature vol. 365,6443 (1993): 274-6. doi:10.1038/365274a0

Szibor, Reinhard. "X-chromosomal markers: past, present and future." Forensic science international. Genetics vol. 1,2 (2007): 93-9. doi:10.1016/j.fsigen.2007.03.003

Tautz, D. "Notes on the definition and nomenclature of tandemly repetitive DNA sequences." EXS vol. 67 (1993): 21-8. doi:10.1007/978-3-0348-8583-6_2

Tillmar, Andreas O et al. "DNA Commission of the International Society for Forensic Genetics (ISFG): Guidelines on the use of X-STRs in kinship analysis." Forensic science international. Genetics vol. 29 (2017): 269-275. doi:10.1016/j.fsigen.2017.05.005

Walsh, P S et al. "Preferential PCR amplification of alleles: mechanisms and solutions." PCR methods and applications vol. 1,4 (1992): 241-50. doi:10.1101/gr.1.4.241

Zavattari, P et al. "Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: demography, chromosome recombination frequency and selection." Human molecular genetics vol. 9,20 (2000): 2947-57. doi:10.1093/hmg/9.20.2947