Forensic Science International: Genetics 29 (2017) 269–275

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Research Paper

DNA Commission of the International Society for Forensic Genetics (ISFG): Guidelines on the use of X-STRs in kinship analysis

Andreas O. Tillmar[a,b,*], Daniel Kling[c], John M. Butler[d], Walther Parson[e,f], Mechthild Prinz[g], Peter M. Schneider[h], Thore Egeland[i,1], Leonor Gusmão[j,1]

[a] Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden
[b] Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Sweden
[c] Department of Forensic Services, Oslo University Hospital, Oslo, Norway
[d] National Institute of Standards and Technology, Special Programs Office, Gaithersburg, MD, USA
[e] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[f] Forensic Science Program, The Pennsylvania State University, PA, USA
[g] John Jay College of Criminal Justice, New York, USA
[h] Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Germany
[i] Norwegian University of Life Sciences, Oslo, Norway
[j] State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil

ARTICLE INFO

Article history:
Received 11 May 2017
Accepted 12 May 2017
Available online 13 May 2017

Keywords:
X chromosome
Linkage
Linkage disequilibrium
Haplotype
Recombination
Paternity
Kinship

ABSTRACT

Forensic genetic laboratories perform an increasing amount of genetic analyses of the X chromosome, in particular to solve complex cases of kinship analysis. For some biological relationships X-chromosomal markers can be more informative than autosomal markers, and there are a large number of markers, methods and databases that have been described for forensic use. Due to their particular mode of inheritance, and their physical location on a single chromosome, some specific considerations are required when estimating the weight of evidence for X-chromosomal marker DNA data. The DNA Commission of the International Society for Forensic Genetics (ISFG) hereby presents guidelines and recommendations for the use of X-chromosomal markers in kinship analysis with a special focus on the biostatistical evaluation. Linkage and linkage disequilibrium (association of alleles) are of special importance for such evaluations and these concepts and the implications for likelihood calculations are described in more detail. Furthermore it is important to use appropriate computer software that accounts for linkage and linkage disequilibrium among loci, as well as for mutations. Even though some software exist, there is still a need for further improvement of dedicated software.

1. Introduction

Short tandem repeat (STR) markers located on the X chromosome (X-STRs) may assist resolution of complex kinship cases as these markers can sometimes be more informative than autosomal STRs depending on the genetic relationship being explored [1–4]. In the analysis of X-chromosomal specific markers, the same type of genetic variation (e.g. STRs, SNPs and Indels) and genotyping methodologies as for autosomes is being used [2,5,6]. There are already a large number of markers, methodologies/protocols and databases described for forensic use [7–11], as well as commercial kits, for example a kit comprising the analysis of 12 X-STRs (Investigator Argus X-12 QS Kit, QIAGEN, Hilden, Germany). Recommendations established for autosomal and Y-chromosomal specific markers [12–14] can be applied to X-chromosomal ones without the need for special consideration, as far as validation requirements for laboratory methods, locus and allele nomenclature are concerned. The majority of the initial population genetic investigations were carried out among individuals of European ancestry, while it was later observed in studies from world-wide populations that some of the tandem repeat structures are more complex and variable than initially observed [15–18]. Thus, care must be taken to ensure that allele definitions correctly represent the existing level of genetic variation.

Nevertheless, due to their particular mode of genetic transmission, physical location on a single chromosome and the absence of

* Corresponding author at: Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Artillerigatan 12, SE-58758 Linköping, Sweden.
E-mail address: andreas.tillmar@rmv.se (A.O. Tillmar).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.fsigen.2017.05.005
1872-4973/

recombination in male meiosis, several specific considerations are required when using X-chromosomal markers in a forensic context, particularly for biostatistical evaluations. Therefore, along with recommendations on the use of X-chromosomal markers in kinship analysis, the aim of these guidelines is to review the main theoretical concepts associated to the specific transmission of the X chromosome, the possibility of linkage among loci, and the high degree of susceptibility to linkage disequilibrium (LD) between alleles at different loci.

2. Scope and limitations

The scope of this paper will be limited to the use of X-chromosomal markers in kinship analyses, but the general concepts apply also to other applications, with the exception of the analysis of X-chromosomal markers in mixtures (i.e. profiles with two or more contributors). The interpretation of DNA profiles from linked markers observed in forensic mixtures has been shown to contain an additional level of complexity [19], and will not be addressed in the present recommendations.

We follow the ISFG recommendations given by Gjertson et al. [20], unless otherwise stated. Therefore, a likelihood ratio (LR) approach is used in the biostatistical evaluations of kinships, based on X-chromosomal genetic profiles. Phylogenetic applications involving problems extending over a large number of generations are not considered.

This paper aims at putting together some general guidelines to be used as reference for those that are currently using X-STRs in kinship analyses. For new practitioners with limited experience on X-STRs, we recommend to first consulting some review articles [2,21] for a comprehensive introduction.

3. When X-chromosomal markers can be useful in kinship analysis

There are many standard cases where X-chromosomal analyses are superfluous. For instance, standard paternity cases (duos and trios) can usually be resolved based solely on autosomal markers. However, due to their specific mode of inheritance, there are several situations where the X-chromosomal markers can be more informative than autosomal, Y-chromosomal, or mitochondrial DNA (mtDNA) markers. In paternity cases, where the information from the autosomal markers is not sufficient to reach a conclusion, the X-chromosomal markers can represent a useful supplementary tool, e.g. in cases with few genetic inconsistencies [4]. When few inconsistencies are observed between the alleged father and the daughter, the most likely explanations are either that mutations have occurred or that the alleged father is closely related to the true father. For example, consider a father/daughter paternity case (Fig. 1). If the biological father is the father (with different mothers) of the alleged father, or if the biological father is the son of the alleged father, they will not share X-chromosomal alleles identical by descent (IBD). For these cases, an analysis using X-chromosomal markers can be much more informative than using autosomal markers.

There are also cases where the routinely used set of autosomal markers needs to be supplemented due to poor amplification results (for example, in paternity investigation from exhumed remains). For father–daughter duos, in case of inclusion, the analysis of X-chromosomal markers can give valuable additional information, since there is only one allele that can be transmitted from the father to the daughter, and more informative LRs can be expected than for the autosomal markers.

X-chromosomal markers can be used as a sole way of inferring relationships, particularly when high LR values are expected compared to the ones expected from autosomal markers. There are

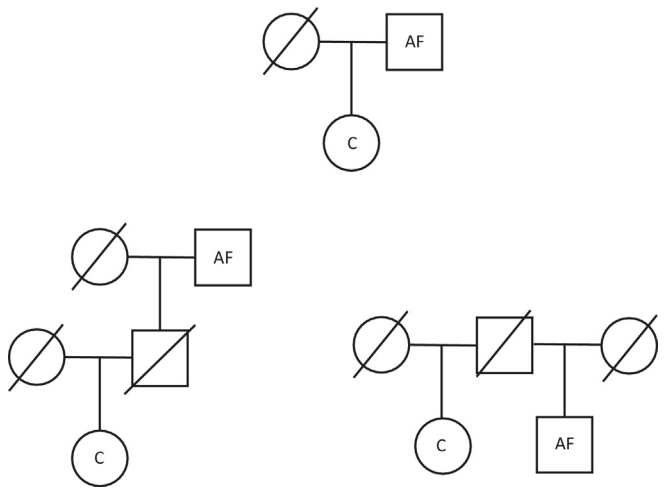

Fig. 1. When few inconsistencies exist between the alleged father (AF) and the daughter (C) in a standard paternity duo case (upper), the most likely explanations are either that mutations have occurred or that there is another close relationship between the alleged father and the child. If the true father is the father (with different mothers) (lower right), or son (lower left) of the alleged father, they will not share X-chromosomal alleles identical by descent (IBD) and the X-chromosomal markers can be much more informative than the autosomal markers.

different methods available that can be used a priori to evaluate the necessity and utility of X-chromosomal analyses. One way is to estimate exclusion probabilities [22]. An alternative is to perform simulations in order to obtain and evaluate distributions of LRs for particular case scenarios. As an example of the latter we performed simulations to obtain distributions of LRs and *exceedance probabilities*: the probability to exceed certain LR thresholds. If 10 of 1000 simulated LRs for a specific hypothesis exceed a threshold, say 10,000, the exceedance probability is 0.01 (1%). Such exceedance probabilities summarise simulation in a useful way as they are closer linked to conclusions and verbal statements than other summary statistics like median values or percentiles. Further examples of efficient calculation of exceedance probabilities are provided in Kruijver [23]. We performed simulations for two different X-chromosomal marker sets (Investigator Argus X-12 QS Kit (QIAGEN) and Decaplex [9]) for 28 different kinship scenarios (Supplementary Tables S1 and S2).

The use of X-chromosomal markers can be particularly important in some kinship cases, where the exclusion or inclusion powers are higher than those for autosomal markers with the same gene diversity [24]. For example, when comparing two alleged paternal half-sisters or an alleged paternal grandmother and a granddaughter (Figs. 2 and 3, respectively), other than for autosomal markers, the exclusion probability for X-chromosomal markers is not zero [2,4].

There are also some pedigrees that cannot be resolved with autosomal markers but can potentially be distinguished from each other by using information from the X chromosome. For instance, no matter the number of (unlinked) autosomal markers that we

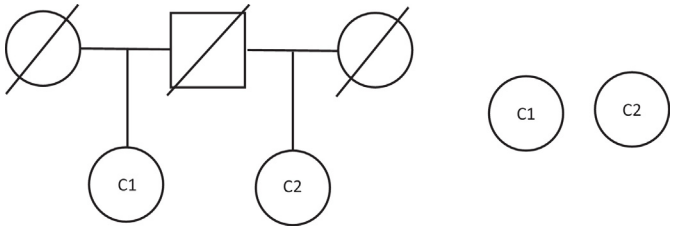

Fig. 2. A case scenario for which the exclusion probability (for the paternal half-sisters hypothesis (left)) is not null for X-chromosomal markers.

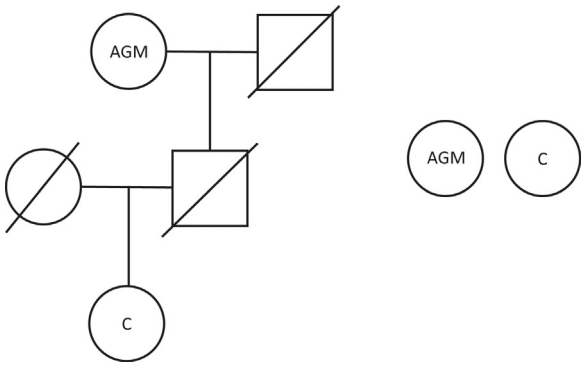

Fig. 3. A case scenario for which the exclusion probability (for an alleged paternal grandmother (AGM) and a granddaughter (C) hypothesis (left)) is not null for X-chromosomal markers.

use, the hypotheses paternal grandfather/granddaughter and paternal uncle/niece will be always equally likely. However, the likelihoods obtained for these two hypotheses can be different for X-STRs (Fig. 4). Other examples of pedigrees distinguishable by X-chromosomal markers but not by unlinked autosomal markers can be found in [3].

X-chromosomal markers can further be useful for inbred cases involving a female child as shown in Supplementary Fig. S1. The question is whether the child was fathered by the brother of the mother or by the father of the mother, in cases where data is available only from the grandmother and the child (and possibly also the mother). If the brother is the true father, the grandmother and child must share X-alleles, not so for the father.

Recommendation #1

In paternity cases (duos or trios, 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.

Recommendation #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 paternal half sibling duos involving two females, and paternal grandmother/granddaughter duos. Furthermore, X-chromosomal markers should be used in situations where two alternative hypotheses possess the same likelihood for autosomal 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.

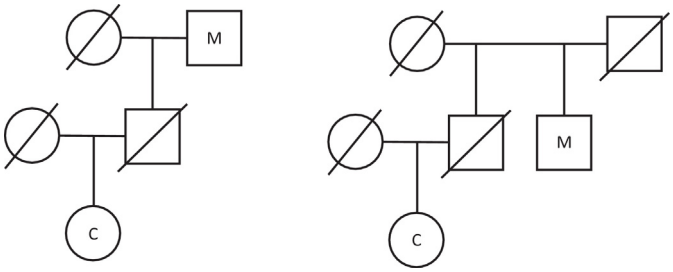

Fig. 4. Two pedigrees (paternal grandfather/granddaughter (left) and paternal uncle/niece (right)), which cannot be resolved with autosomal (unlinked) markers but could be distinguished when using X-chromosomal markers. DNA data are available for a male individual (M) and a child (C).

4. Linkage and linkage disequilibrium

Linkage and linkage disequilibrium (or allelic association) are two concepts of dependencies that become relevant when analysing multiple loci with close physical location, such as for X-STR multiplexes. Although linkage and LD have some properties in common it is important to keep the discussion about them separate, since there are differences regarding their definition, their impact and in how they are accounted for in the calculation of the LR. Linkage is a consequence of genetic recombination causing closely located markers to be inherited as a unit from parent to child with a higher probability than physically separated or independent markers. Linkage is accounted for via transmission probabilities where the recombination rate, between the linked loci, is a key component. LD, on the other hand, exists when alleles at different loci occur together, at a population level, more (or less) often than expected by chance. This non-random association can be caused by linkage or other population genetic effects like population substructure, non-random mating, migration etc. LD is accounted for by using haplotype frequencies instead of allele frequencies [25].

4.1. Linkage

As noted above linkage is a consequence of the recombination process occurring between two homologous chromosomes during meiosis (Fig. 5). If more than 50% of the gametes are expected to have the same segment as the parental chromosome (on which the two markers are located), the two markers are linked. Consequently, two markers are considered unlinked when an odd number of recombinations is expected to occur in 50% of the times during meioses.

There are several ways to study if loci are linked or not. Most often it is performed via segregation analysis in multi-generation families in order to estimate the recombination rate (also referred to as recombination frequency or recombination fraction), but linkage can also be studied based on larger population based studies, like the HapMap project [26], to create maps of genetic distances. Traditionally, such distances are given in centiMorgans (cM) and there are several mapping functions designed to convert genetic distances to recombination rates (e.g. Haldane's [27] and Kosambi's [28]). For markers on the same chromosome, a general rule of thumb is the assumption of no linkage, when the genetic distance is at least 50 cM, resulting in independent transmission to the offspring. Even though this rule is sometimes applicable, 50 cM is roughly equal to a recombination rate of 32%. This follows using Haldane's mapping function with $x = 50$:

$$r = \frac{1 - \exp(-2\frac{x}{100})}{2} = \frac{1 - \exp(-1)}{2} \approx 0.32$$

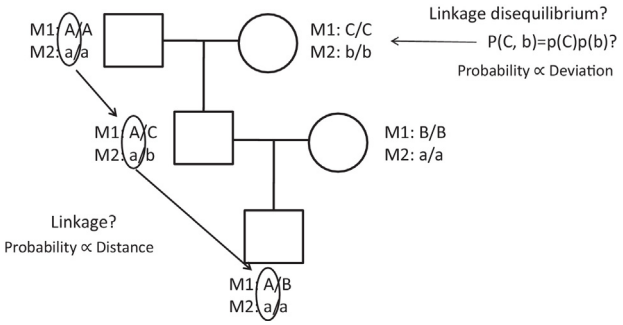

Fig. 5. An illustration of linkage and linkage disequilibrium.
The figure is reproduced with permission from T. Egeland, D. Kling and P. Mostad, In Relationship Inference with Familias and R, Academic Press, San Diego, 2016.

where *r* is the recombination rate and *x* is the genetic distance in cM. Supplementary Fig. S2 illustrates an example of the effect of the recombination rate on the LR, even for markers at 50 cM. When using Haldane's or Kosambi's mapping functions, recombination rates of approximately 50% are only obtained for genetic distances larger than 200 cM. In principle an independent transmission cannot, therefore, be assumed for markers separated by less than 200 cM, unless a recombination rate of 50% is established using genotype data from informative families.

Since the X chromosome is approximately 155 Mb long (approx. 180 cM) [26], we can only select a maximum of 3–4 X-chromosomal markers separated by more than 50 cM that would segregate reasonably independent. However, to increase the evidential weight, additional markers have been introduced within the so called linkage groups (or clusters) [2]. Historically, it has been suggested that markers within such linkage groups could be assumed to transmit without recombination through generations [29]. However, later studies of recombination rates based on family studies have shown that this assumption may be misleading [30,31], and the recommended approach is to account for recombination even for markers within a linkage group. This can be done in the biostatistical calculation in the same manner as for linked markers located outside a linkage group [32]. In other words, haplotypes are not necessarily transmitted unchanged within a pedigree but may change due to recombination events.

For a valid interpretation of the weight of evidence, the rate of recombination between adjacent markers in X-STR multiplexes needs to be determined. For example, Nothnagel and colleagues [30,31] estimated recombination rates for X-STR markers based on family studies. Phillips and co-workers used another approach and employed high density multi-point SNP data from population studies in order to estimate the genetic positioning of X-STRs [33].

For some pedigrees it is not always necessary to account for linkage when computing the LR. For example, under the assumption of linkage equilibrium (LE), in a maternity duo case scenario with X-chromosomal data from a putative mother and a male child, the recombination parameter will not be in the final formula (Box 1). Other situations where linkage will have no effect on the LR include scenarios with certain genotype constellations that cancel out the recombination rate parameter. There is, however, no definitive rule determining the effect of linkage on the results.

Software and tools are available for calculation of the LR for X-chromosomal markers taking linkage into account (see more details in Section 5). If linkage is not accounted for, the LR can either be over-estimated or under-estimated compared to the "true" LR as demonstrated in simulation studies [32,34].

Recommendation #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 via mapping functions based on genetic distances. A recombination rate below 0.5 indicates linkage.

Recommendation #4

Linkage should be accounted for when calculating LRs 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.

4.2. Linkage disequilibrium

Linkage disequilibrium is exhibited by a non-random association of alleles from different markers at a population level (Fig. 5). These markers may or may not be located on the same chromosome. If LD exists, a certain allelic combination (e.g. haplotype) will be more or less common in the population compared with the assumption of random association.

Since X recombination only occurs in women, alleles at markers located on the X chromosome are generally expected to display a higher degree of LD than autosomal markers. The lower mutation rates observed in the female germ line (which apply to 2/3 of the X chromosomes present in a population) also contribute to higher LD values for X-chromosomal markers [35,36]. The study of LD structure in one population should not be extrapolated to others [36] since, in addition to the physical distance between markers, other factors, typically related to the population history, will contribute to LD. Population sub-structure will also have a strong impact on the degree of LD [37]. It should also be noted that recombination events break down LD structures over generations [25,38].

Linkage equilibrium can be tested for by comparing differences between observed and expected haplotype frequencies. This can, for example, be achieved via an exact test, or a similar test, using a dedicated software like e.g. Arlequin [39] and PLINK [40] (http://pngu.mgh.harvard.edu/purcell/plink). There are also several libraries in R (https://www.r-project.org/) with wide ranging relevant

Box 1.

Consider a maternity case involving a putative mother and a child (male). The data consist of two linked markers (separated with a recombination rate, *r*) on the X chromosome. The putative mother has genotype *a*/*b* at marker 1 and *c*/*d* at marker 2, and the child has genotype *a* at marker 1 and *d* at marker 2. The formula for the LR can then be written as (assuming that *a*, *b*, *c* and *d* are different alleles):

$$LR = \frac{\Pr(DNA|mother\ of\ child)}{\Pr(DNA|unrelated)} = \frac{2 \cdot p_{ac} \cdot p_{bd} \cdot 0.5 \cdot r + 2 \cdot p_{ad} \cdot p_{bc} \cdot 0.5(1-r)}{(2 \cdot p_{ac} \cdot p_{bd} + 2 \cdot p_{ad} \cdot p_{bc})p_{ad}} = \frac{p_{ac} \cdot p_{bd} \cdot r + p_{ad} \cdot p_{bc} \cdot (1-r)}{(2 \cdot p_{ac} \cdot p_{bd} + 2 \cdot p_{ad} \cdot p_{bc})p_{ad}}$$

If LE holds (i.e. $p_{xy} = p_x \cdot p_y$) the LR becomes:

$$LR = \frac{2 \cdot p_{ac} \cdot p_{bd} \cdot 0.5 \cdot r + 2 \cdot p_{ad} \cdot p_{bc} \cdot 0.5(1-r)}{(2 \cdot p_{ac} \cdot p_{bd} + 2 \cdot p_{ad} \cdot p_{bc})p_{ad}} = \frac{p_a \cdot p_c \cdot p_b \cdot p_d \cdot r + p_a \cdot p_c \cdot p_b \cdot p_d \cdot (1-r)}{(2 \cdot p_a \cdot p_c \cdot p_b \cdot p_d + 2 \cdot p_a \cdot p_c \cdot p_b \cdot p_d)p_a \cdot p_d} = \frac{1}{4 \cdot p_a \cdot p_d}$$

For a maternity case, the recombination rate has no impact on the final LR if LE can be assumed. However if LD exists linkage needs to be accounted for even for a maternity case scenario as derived above.

functionality. Due to the relatively low power of these tests (i.e. the probability to detect LD is low, when LD exists) the population size should preferably be larger than the recommended size for standard autosomal STR allele frequency databasing. The recommended size of the database depends, as noted below, on several factors like the number of markers in the haplotype and the expected LD structure [24]. Large datasets are also preferred when estimating haplotype frequencies as further discussed below.

Since the gametic phase of a female X-STR profile is usually unknown (unless she is homozygous for all markers, for all markers but one, or if informative family members are genotyped), computations of likelihoods will involve iterations over all possible allelic combinations (i.e., haplotypes) given the genotypes of each individual (assuming Hardy-Weinberg equilibrium). It is therefore important to note that this may require estimates of all possible haplotype constellations given the detected alleles in the population at the different markers. As commented above, this requires large datasets in order to obtain accurate frequency estimates. We are aware of no published studies to find the optimal database size, but provide some crude guidelines: consider a three marker cluster with each marker having 10 alleles. In total 1000 possible haplotypes can be constructed from such cluster. In contrast, databasing for single marker frequencies normally encompasses at maximum of approximately 50 alleles [41]. It is up to each laboratory to consider the database size in light of these numbers together with data from empirical studies of the impact of various database sizes [24,42], and other experiences. The recommended approach for databasing is to analyse male individuals to generate data for haplotype frequency estimates. A Dirichlet based model has been proposed in order to also be able to account for unobserved haplotypes [24,32,34]. Supplementary Table S3 contains a reference list of more than 150 publications comprising population frequencies for various X-chromosomal markers.

When computing likelihoods (and LRs) LD is accounted for by using observed haplotype frequencies rather than those expected from the product of the allele frequencies (i.e. assuming linkage equilibrium). In the presence of LD, any likelihood ratio calculated based on allele rather than haplotype frequencies could be biased [32,34].

Furthermore, when LE can be assumed, the use of haplotype frequencies should give the same result as multiplying single locus allele frequencies. However, even though LE has been demonstrated, sampling effects (e.g. few samples) may falsely reject LD due to low statistical power. Therefore, whenever doubts exist concerning the presence of LD between closely linked markers, haplotype frequencies should be adopted.

If no adequate haplotype frequencies but only allele frequencies are available for a particular population background, the results of a single marker within each of the X-chromosomal linkage groups analysed in a given case could be used for biostatistical calculations. This approach should however be used with care, since the outcome might be biased. The result could depend on the pedigree to be evaluated, the impact of linkage and the allele frequencies. It is up to the laboratory the ensure an unbiased outcome if using this approach.

Recommendation #5

Linkage equilibrium tests should be performed when generating population frequency data for the markers in a X-chromosomal marker multiplex.

Recommendation #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 allele frequencies for population databasing.

Recommendation #7

Haplotype frequencies should be used for likelihood calculations when LD exists.

5. Tools available for LR calculation

Due to the necessity to account for linkage, manual calculations based on multiple X-chromosomal loci involve greater complexity and are therefore more prone to errors than non-linked markers. Therefore the use of appropriate computer software is highly recommended. If manual calculations (and algebraic formulas) are necessary, they should always be checked against software results.

Following the recommendations of the Paternity Testing Commission of the International Society for Forensic Genetics [20], biostatistical evaluations should be based on a likelihood ratio principle. For reasons mentioned above, it is important to use software that accounts for linkage and linkage disequilibrium among loci, as well as for mutations. When evaluating data from individuals from a stratified population, the software should ideally also be able to account for population substructure.

For obvious reasons, there are fewer available software packages for X-chromosome applications compared to the autosomal counterpart. Problems similar to those now faced by the forensic community have been encountered previously in other applied areas of genetics, including human genetics. MINX (MERLIN IN X) is the specific X-version of MERLIN, a program widely used for linkage analysis in human applications [43,44]. The early versions of the freely available FamLinkX [32] relied to a large extent on the likelihood calculations of MERLIN. While MERLIN is still an integrated part of FamLinkX, recent extensions including modelling of mutations (which is typically not accommodated in non-forensic applications) combined with linkage and linkage disequilibrium, have made the important parts of FamLinkX independent of MERLIN [24,32,45].

Several R libraries are useful. As an example, Supplementary Fig. S2 is plotted using paramlink [46] (https://cran.r-project.org/web/packages/paramlink/) and data simulated with the function "markerSim". As opposed to other software we are aware of, the mentioned function can do conditional simulations (i.e. some individuals, say children have been genotyped). Recently, paramlink was extended to allow for mutations. The Elston-Stewart algorithm [47] is also implemented for one and two markers (oneMarkerDistribution, twoMarkerDistribution) and this software was used to validate FamLinkX. However, a distinguishing feature of FamLinkX remains its ability to deal with mutation and LD (the approach to account for LD in MERLIN is not generally appropriate for the applications we have in mind, since MERLIN assumes that no recombinations occur within a LD cluster (i.e. linkage group)). The library IBDSIM, developed by Vigeland [48] can simulate as indicated by the title and also estimates kinship coefficients. The function "exclusionPower" computes the power (of a single marker) of excluding a claimed relationship, given the true relationship [46]. The mentioned R libraries work for general pedigrees and both in the autosomal and the X-chromosomal case.

With a more generalised use of X-STRs in the forensic field, especially to solve complex paternity cases and in the identification of missing persons, it is expected that other software will be available, incorporating the functionalities described above. As for any other software calculating likelihood ratios to evaluate competing kinship scenarios, the recommendations from the DNA Commission of the ISFG on the validation of software programs [49] should be followed.

Recommendation #8

Appropriate software should be used when calculating LRs 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.

Recommendation #9

As for 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 [49].

6. Combining LRs from X-chromosomal and autosomal information

If a population is not stratified, it is not expected to find LD between alleles from markers at different chromosomes and it is possible to multiply the LRs obtained for the autosomal and X-chromosomal markers respectively, without any kind of correction. If population stratification exists, it must be taken into account, requiring the use of correction parameters [50,51].

Moreover, in order to combine the LR from autosomal markers with the LR from X-chromosomal markers, the case specific hypotheses must be unambiguously formulated and be equivalent for the autosomal and for the X-chromosomal calculations. For example, one should not use a hypothesis like “not full siblings”, as the alternative hypothesis to “full siblings”, but instead “unrelated”, or “half siblings” (if applicable).

Recommendation #10

Individual autosomal LR and X-chromosomal LR results should only be combined whenever equivalent (and clearly defined) hypotheses 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.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the National Institute of Standards and Technology. Certain commercial entities are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that any of the entities identified are necessarily the best available for the purpose.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at (doi:10.1016/j.fsigen.2017.05.005).

References

[1] R. Szibor, M. Krawczak, S. Hering, J. Edelmann, E. Kuhlisch, D. Krause, Use of X-linked markers for forensic purposes, Int. J. Legal Med. 117 (2003) 67–74.
[2] R. Szibor, X-chromosomal markers: past, present and future, Forensic Sci. Int. Genet. 1 (2007) 93–99.
[3] N. Pinto, L. Gusmão, A. Amorim, X-chromosome markers in kinship testing: a generalisation of the IBD approach identifying situations where their contribution is crucial, Forensic Sci. Int. Genet. 5 (2011) 27–32.
[4] C. Gomes, M. Magalhães, C. Alves, A. Amorim, N. Pinto, L. Gusmão, Comparative evaluation of alternative batteries of genetic markers to complement autosomal STRs in kinship investigations: autosomal indels vs. X-chromosome STRs, Int. J. Legal Med. 126 (2012) 917–921.
[5] C. Tomas, J.J. Sanchez, A. Barbaro, C. Brandt-Casadevall, A. Hernandez, M. Ben Dhiab, M. Ramon, N. Morling, X-chromosome SNP analyses in 11 human Mediterranean populations show a high overall genetic homogeneity except in North-west Africans (Moroccans), BMC Evol. Biol. 8 (2008) 75.
[6] R. Pereira, V. Pereira, I. Gomes, C. Tomas, N. Morling, A. Amorim, M.J. Prata, A. Carracedo, L. Gusmão, A method for the analysis of 32 X chromosome insertion deletion polymorphisms in a single PCR, Int. J. Legal Med. 126 (2012) 97–105.
[7] H. Asamura, H. Sakai, K. Kobayashi, M. Ota, H. Fukushima, MiniX-STR multiplex system population study in Japan and application to degraded DNA analysis, Int. J. Legal Med. 120 (2006) 174–181.
[8] S. Turrina, R. Atzei, G. Filippini, D. De Leo, Development and forensic validation of a new multiplex PCR assay with 12 X-chromosomal short tandem repeats, Forensic Sci. Int. Genet. 1 (2007) 201–214.
[9] L. Gusmão, P. Sánchez-Diz, C. Alves, I. Gomes, M.T. Zarrabeitia, M. Abovich, I. Atmetlla, C. Bobillo, L. Bravo, J. Builes, L. Cainé, R. Calvo, E. Carvalho, M. Carvalho, R. Cicarelli, L. Catelli, D. Corach, M. Espinoza, O. García, M. Malaghini, J. Martins, F. Pinheiro, M. João Porto, E. Raimondi, J.A. Riancho, A. Rodríguez, A. Rodríguez, B. Rodríguez Cardozo, V. Schneider, S. Silva, C. Tavares, U. Toscanini, C. Vullo, M. Whittle, I. Yurrebaso, A. Carracedo, A. Amorim, A GEP-ISFG collaborative study on the optimization of an X-STR decaplex: data on 15 Iberian and Latin American populations, Int. J. Legal. Med. 123 (2009) 227–234.
[10] Q.L. Liu, D.J. Lu, L. Quan, Y.F. Chen, M. Shen, H. Zhao, Development of multiplex PCR system with 15 X-STR loci and genetic analysis in three nationality populations from China, Electrophoresis 33 (2012) 1299–1305.
[11] E. Prieto-Fernández, M. Baeta, C. Núñez, M.T. Zarrabeitia, R.J. Herrera, J.J. Builes, M.M. de Pancorbo, Development of a new highly efficient 17 X-STR multiplex for forensic purposes, Electrophoresis 37 (2016) 1651–1658.
[12] W. Bär, B. Brinkmann, B. Budowle, A. Carracedo, P. Gill, P. Lincoln, W. Mayr, B. Olaisen, DNA recommendations. Further report of the DNA Commission of the ISFG regarding the use of short tandem repeat systems, Forensic Sci. Int. 87 (1997) 179–184.
[13] P. Gill, C. Brenner, B. Brinkmann, B. Budowle, A. Carracedo, M.A. Jobling, P. de Knijff, M. Kayser, M. Krawczak, W.R. Mayr, N. Morling, B. Olaisen, V. Pascali, M. Prinz, L. Roewer, P.M. Schneider, A. Sajantila, C. Tyler-Smith, DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs, Forensic Sci. Int. 124 (2001) 5–10.
[14] L. Gusmão, J.M. Butler, A. Carracedo, P. Gill, M. Kayser, W.R. Mayr, N. Morling, M. Prinz, L. Roewer, C. Tyler-Smith, P.M. Schneider, DNA Commission of the International Society of Forensic Genetics. DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, Forensic Sci. Int. 157 (2006) 187–197.
[15] C. Tomas, V. Pereira, N. Morling, Analysis of 12 X-STRs in Greenlanders, Danes and Somalis using Argus X-12, Int. J. Legal Med. 126 (2012) 121–128.
[16] S. Elakkary, S. Hoffmeister-Ullerich, C. Schulze, E. Seif, A. Sheta, S. Hering, J. Edelmann, C. Augustin, Genetic polymorphisms of twelve X-STRs of the Investigator Argus X-12 kit and additional six X-STR centromere region loci in an Egyptian population sample, Forensic Sci. Int. Genet. 11 (2014) 26–30.
[17] I. Gomes, A. Brehm, L. Gusmão, P.M. Schneider, New sequence variants detected at DXS10148, DXS10074 and DXS10134 loci, Forensic Sci. Int. Genet. 20 (2016) 112–116.
[18] D. Kling, B. Dell'Amico, P.J. Haddeland, A.O. Tillmar, Population genetic analysis of 12 X-STRs in a Somali population sample, Forensic Sci. Int. Genet. 11 (2014) e7–e8.
[19] G. Dørum, D. Kling, A. Tillmar, M.D. Vigeland, T. Egeland, Mixtures with relatives and linked markers, Int. J. Legal Med. 130 (2016) 621–634.
[20] D.W. Gjertson, C.H. Brenner, M.P. Baur, A. Carracedo, F. Guidet, J.A. Luque, R. Lessig, W.R. Mayr, V.L. Pascali, M. Prinz, P.M. Schneider, N. Morling, ISFG: recommendations on biostatistics in paternity testing, Forensic Sci. Int. Genet. 1 (2007) 223–231.
[21] T.M. Diegoli, Forensic typing of short tandem repeat markers on the X and Y chromosomes, Forensic Sci. Int. Genet. 18 (2015) 140–151.
[22] N. Pinto, L. Gusmão, L.T. Egeland, A. Amorim, Paternity exclusion power: comparative behaviour of autosomal and X-chromosomal markers in standard and deficient cases with inbreeding, Forensic Sci. Int. Genet. 7 (2013) 290–295.
[23] M. Kruijver, Efficient computations with the likelihood ratio distribution, Forensic Sci. Int. Genet. 14 (2015) 116–124.
[24] D. Kling, B. Dell'Amico, A.O. Tillmar, FamLinkX – implementation of a general model for likelihood computations for X-chromosomal marker data, Forensic Sci. Int. Genet. 17 (2015) 1–7.
[25] J. Ott, Analysis of Human Genetic Linkage, JHU Press, Baltimore, MD, 1985.
[26] The International HapMap Consortium, A second generation human haplotype map of over 3.1 million SNPs, Nature 449 (2007) 851–861.
[27] J.B.S. Haldane, The combination of linkage values, and the calculation of distances between the loci of linked factors, J. Genet. 8 (1919) 299–309.
[28] D.D. Kosambi, The estimation of map distances from recombination values, Ann. Eugen. 12 (1944) 172–175.
[29] S. Hering, C. Augustin, J. Edelmann, M. Heidel, J. Dressler, H. Rodig, E. Kuhlisch, R. Szibor, DXS10079, DXS10074 and DXS10075 are STRs located within a 280-kb region of Xq12 and provide stable haplotypes useful for complex kinship cases, Int. J. Legal Med. 120 (2006) 337–345.
[30] M. Nothnagel, R. Szibor, O. Vollrath, C. Augustin, J. Edelmann, M. Geppert, C. Alves, L. Gusmão, M. Vennemann, Y. Hou, U.D. Immel, S. Inturri, H. Luo, S. Lutz-Bonengel, C. Robino, L. Roewer, B. Rolf, J. Sanft, K.J. Shin, J.E. Sim, P. Wiegand, C. Winkler, M. Krawczak, S. Hering, Collaborative genetic mapping of 12 forensic

short tandem repeat (STR) loci on the human X chromosome, Forensic Sci. Int. Genet. 6 (2012) 778–784.
[31] T.M. Diegoli, H. Rohde, S. Borowski, M. Krawczak, M.D. Coble, M. Nothnagel, Genetic mapping of 15 human X chromosomal forensic short tandem repeat (STR) loci by means of multi-core parallelization, Forensic Sci. Int. Genet. 25 (2016) 39–44.
[32] D. Kling, A. Tillmar, T. Egeland, P. Mostad, A general model for likelihood computations of genetic marker data accounting for linkage, linkage disequilibrium, and mutations, Int. J. Legal Med. 129 (2015) 943–954.
[33] C. Phillips, D. Ballard, P. Gill, D.S. Court, A. Carracedo, M.V. Lareu, The recombination landscape around forensic STRs: accurate measurement of genetic distances between syntenic STR pairs using HapMap high density SNP data, Forensic Sci. Int. Genet. 6 (3) (2012) 354–365.
[34] A.O. Tillmar, T. Egeland, B. Lindblom, G. Holmlund, P. Mostad, Using X-chromosomal markers in relationship testing: calculation of likelihood ratios taking both linkage and linkage disequilibrium into account, Forensic Sci. Int. Genet. 5 (2011) 506–511.
[35] S.F. Schaffner, The X chromosome in population genetics, Nat. Rev. Genet. 5 (2004) 43–51.
[36] M.A. Tariq, M.F. Sabir, S.A. Riazuddin, S. Riazuddin, Haplotype analysis of two X-chromosome STR clusters in the Pakistani population, Int. J. Legal Med. 123 (2009) 85–87.
[37] M. Laan, V. Wiebe, E. Khusnutdinova, M. Remm, S. Pääbo, X-chromosome as a marker for population history: linkage disequilibrium and haplotype study in Eurasian populations, Eur. J. Hum. Genet. 13 (2005) 452–462.
[38] P. Gill, C. Phillips, C. McGovern, J.A. Bright, J. Buckleton, An evaluation of potential allelic association between the STRs vWA and D12S391: implications in criminal casework and applications to short pedigrees, Forensic Sci. Int. Genet. 6 (2012) 477–486.
[39] L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour. 10 (2010) 564–567.
[40] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. de Bakker, M.J. Daly, P.C. Sham, PLINK: a tool set for whole-genome association and population-based linkage analyses, Am. J. Hum. Genet. 81 (2007) 559–575.
[41] C.R. Hill, D.L. Duewer, M.C. Kline, M.D. Coble, J.M. Butler, U.S. population data for 29 autosomal STR loci, Forensic Sci. Int. Genet. 7 (2013) e82–e83.
[42] A. Shyla, S.R. Borovko, A.O. Tillmar, N.N. Kuzub, S.A. Kotova, I.S. Tsybovsky, K. Rębała, Belarusian experience of the use of FamLinkX for solving complex kinship cases involving X-STR markers, Forensic Sci. Int. Genet. Suppl. Ser. 5 (2015) e539–e541.
[43] G.R. Abecasis, S.S. Cherny, W.O. Cookson, L.R. Cardon, Merlin – rapid analysis of dense genetic maps using sparse gene flow trees, Nat. Genet. 30 (2002) 97–101.
[44] M. Krawczak, Kinship testing with X-chromosomal markers: mathematical and statistical issues, Forensic Sci. Int. Genet. 1 (2007) 111–114.
[45] T. Egeland, D. Kling, P. Mostad, Relationship Inference with Familias and R. Statistical Methods in Forensic Genetics, 1st ed., Elsevier, 2016.
[46] T. Egeland, N. Pinto, M.D. Vigeland, A general approach to power calculation for relationship testing, Forensic Sci. Int. Genet. 9 (2014) 186–190.
[47] R.C. Elston, J. Stewart, A general model for the genetic analysis of pedigree data, Hum. Hered. 21 (1971) 523–542.
[48] M.D. Vigeland, IBDsim: an R package for simulation of chromosomal segments shared by family members, poster presentation at the The American Society of Human Genetics, (2011) .
[49] M.D. Coble, J. Buckleton, J.M. Butler, T. Egeland, R. Fimmers, P. Gill, L. Gusmão, B. Guttmank, M. Krawczak, N. Morling, W. Parson, N. Pinto, P.M. Schneider, S.T. Sherrys, S. Willuweit, M. Prinz, DNA Commission of the International Society for Forensic Genetics: recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications, Forensic Sci. Int. Genet. 25 (2016) 191–197.
[50] D. Balding, R. Nichols, A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity, Genetica 96 (1995) 3–12.
[51] J.C. de Zoete, Combining forensic evidence, (2016) PhD-thesis at University of Amsterdam.