Collaborative genetic mapping of 12 forensic short tandem repeat (STR) loci on the human X chromosome

Collaborative genetic mapping of 12 forensic short tandem repeat (STR) loci on the human X chromosome

Michael Nothnagel*, Reinhard Szibor, Oliver Vollrath, Christa Augustin, Jeanett Edelmann, Maria Geppert, Iva Gomes, Leonor Gusmão, Marielle Heinrich, Yiping Hou, Uta-Dorothee Immel, Serena Inturri, Haibo Luo, Marielle Heinrich, Sabine Lutz-Bonengel, Carlo Robino, Lutz. Roewer, Burghard Rolf, Juliane Sanft, Kyoung-Jin Shin, Jeong Eun Sim, Peter Wiegand, Christian Winkler, Michael Krawczak, Sandra Hering

*: to whom correspondence should be addressed

M. Nothnagel, O. Vollrath, M. Krawczak
Institut für Medizinische Informatik und Statistik,Christian-Albrechts-Universität Kiel, Brunswiker Strasse 10, 24105 Kiel, Germany

R. Szibor, Institut für Rechtsmedizin, Otto-von-Guericke-Universität Magdeburg, Leipziger Straße 44, 39120 Magdeburg, Germany,

C. Augustin, Institut für Rechtsmedizin, Universitätsklinikum Hamburg-Eppendorf, Butenfeld 34, 22529 Hamburg

J. Edelmann, Institut für Rechtsmedizin, Universität Leipzig, Johannisallee 28, 04103 Leipzig, Germany

M. Geppert, L. Roewer, Abteilung für Forensische Genetik, Institut für Rechtsmedizin und Forensische Wissenschaften, Charité–Universitätsmedizin Berlin, Berlin, Germany

I. Gomes, L. Gusmão
IPATIMUP, Institute of Molecular Pathology and Immunology, University of Porto 4200-465 Porto, Portugal

Y. Hou , H. Luo, Institute of Forensic Medicine, SichuanUniversity
Ren Min Nan Lu 3-17; Chengdu 610041, Sichuan, China

S. Inturri, C. Robino, Dipartimento di Anatomia,Farmacologia e Medicina Legale, Università di Torino,
Corso Galileo Galilei 22, 10126 Torino

U. D. Immel, Institute for Legal Medicine, University Halle-Wittenberg,
Franzosenweg 1, 06112 Halle, Germany

S. Lutz-Bonengel, M. Heinrich, Institute of Legal Medicine, University of Freiburg, Albertstr. 9, 79104 Freiburg, Germany.

B. Rolf, Institute of Legal Medicine, Ludwigs-Maximilians-Universität, Frauenlobstr. 7 a, 80337 München, Germany.

J. Sanft, Institut für Rechtsmedizin, Friedrich-Schiller-Universität Jena, Fürstengraben 23, 07743 Jena, Germany

K.J. Shin, J.E. Sim, Department of Forensic Medicine, Yonsei University College of Medicine, 250 Seongsanno, Seodaemun-Gu, 120-752 Seoul, Korea

C. Winkler
Institut für Blutgruppenforschung LCG GmbH, Stolberger Str. 370 | 50933 Köln, Germany

P. Wiegand
Department of Legal Medicine, UniversityHospital, Prittwitzstr. 6, 89075 Ulm, Germany

S. Hering, Institut für Rechtsmedizin, Technische Universität Dresden, Fetscherstrasse 74, 01307 Dresden, Germany

Abstract

A large number of short tandem repeat (STR) markers spanning the entire human X chromosome have been described and established for use in forensic genetic testing. Due to their particular mode of inheritance, X-STRs often allow easy and informative haplotyping in kinship cases. Moreover, some X-STRs are known to be tightly linked so that, in combination, they constitute even more complex genetic markers than each STR taken individually. As a consequence, X-STRs have proven to be a particularly powerful tool for solving deficiency cases. However, valid quantification of the evidence provided by X-STR genotypes in the form of likelihood ratios requires that the recombination rates between markers are exactly known. In a collaborative family study, we used X-STR genotype data from 401 two- and three-generation families to derive valid estimates of the recombination rates between 12 forensic markers, namely DXS10148, DXS10135, DXS8378, DXS7132, DXS 10079, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS10134 and DXS7423. Our study is the first to simultaneously allow for mutation and recombination in the likelihood calculations, thereby obviating the bias-prone practice of excluding ambiguous transmission events from further consideration.

Key words: X chromosome, haplotyping, recombination, STR, kinship testing

Introduction

Genotyping of X-chromosomal short tandem repeats (X-STRs) has become a useful tool in forensic genetics. In the recent past, a large number of X-STRs spanning the entire human X chromosome have been described [1-5]. Joint consideration of X-STRs can yield even more complex and informative genetic systems provided that the probability of inter-marker recombination is negligible and that haplotypes therefore segregate stably within families. In fact, sharing of rare X-STRs haplotypes has been recognized as being strongly indicative of kinship [6].

Haplotyping is easier with X-STRs than with autosomal markers. Since males are hemizygous for all X-chromosomal loci, their X-STR haplotypes are revealed directly by genotyping. In addition, for pedigrees in which critical kinship relationships are beyond doubt, it may be possible to determine female X-STR haplotypes with sufficient accuracy as well. Women always carry the entire paternal X chromosome so that their X-STR haplotypes can be assessed, for example, by genotyping their biological father. Likewise, genotyping of a mother and at least one son may also reveal both maternal haplotypes albeit with some residual uncertainty due to the possibility of female recombination.

Recently, eight X-STRs were evaluated for forensic use and were included into the Mentype® Argus X-8 PCR amplification kit [7]. For practical purposes, the eight markers have been group so far into pairs constituting four (presumably) independent linkage groups, namely DXS10135 - DXS8378, DXS7132 - DXS10074, HPRTB - DXS10101, and DXS10134 - DXS7423 [6]. Linkage within groups has been regarded to be sufficiently tight for the chance of intra-pair recombination to be negligible in practice. These postulates complied with published recombination data [8-10], including our own earlier family studies [6, 11], although only a small number of meioses were studied.

To allow female recombination to be taken properly into account in quantitative kinship analyses using X-STRs, the recombination fractions between the respective markers need to be known precisely [12]. Recently, two studies [13, 14] strongly suggested that our abovementioned rough and preliminary estimates of the recombination fractions between the Argus X-8 markers needed to be modified. This led us to publish the recombination pattern of 39 X-STRs as observed in German three-generation families, comprising a total of 135 meioses [15]. However, since the accuracy of the recombination fraction estimates ensuing from such analyses is

critically dependent upon the number of meioses studied, we choose to expand upon the aforementioned German study. Here, we present the results of an international multi-center study of X-STR recombination in a much larger number of female meioses, thereby allowing more precise estimation of the recombination fractions of interest. Moreover, by following a comprehensive likelihood-based approach, we were able for the first time to allow for meiotic mutation in the estimation of X-STR recombination fractions. With a view to increase the informativity of currently used marker panels, we included four additional STRS in our study, namely DXS10148 [16], DXS10079 [3], DXS10103 [2] and DXS10146 [17], to complement the Argus X-8 kit (Table 1).

Material and Methods

Families

Individuals investigated in the present study originated from one of six European or two Asian centers routinely involved in kinship testing. Probands belonged to the testing clientele of the centers or came from families of either students or friends of the authors. Specimens of the latter were de-identified before genotyping. Samples were collected from two types of families. Type I families were three-generation pedigrees comprising a man, one or more of his daughters and several of the daughters' sons. Type I families with more than one daughter were split into sub-families comprising a single mother, her father and her sons. After splitting, the total number of type I families available for analysis equaled 216. Type II families comprised a single mother and two or more sons (n=185). In total, our data included genotypes from 1284 individuals, all of whom had given their written informed consent prior to the study.

DNA extraction and genotyping

DNA was extracted from buccal cells using the Chelex method or QIAamp DNA extraction Kit (Qiagen, Hilden/Germany). The Mentype® Argus X-8 PCR amplification kit (Biotype AG, Dresden, Germany) was used to amplify the following STR loci: DXS10135, DXS8378, DXS7132, DXS10074, HPRTB, DXS10101, DXS10134 and DXS7423. Cycling conditions and allele calling were according to the producer's instructions. Four additional STRs, namely DXS10148, DXS10079, DXS10103 and DXS10146, were analyzed using PCR primers designed on the basis of GenBank information and making use of the Primer3 software (Table 2). Amplification was

carried out in a 25 µl quadroplex PCR setup using the Qiagen Multiplex-PCR Kit. The mixture contained 10.1 µl H_2O, 12.5 µl Qiagen-mastermix (containing Taq polymerase) and the following quantities of primers solution (50 µM each): 0.16µl DXS10148 F/R, DXS10079 F/R, DXS10103 F/R and 0.25µl DXS10146 F/R. The final PCR setup contained 24 µl of this mixture plus 1µl DNA (1-5 ng). The amplification protocol was as follows: 95°C for 15 min; 95°C for 30 sec, 60°C for 40 sec, 72°C for 40 sec ; 28 cycles ; 60°C for 20 min, 4°C for ever.

Since genotyping was carried out in different institutions, the laboratory equipment used varied somehow. Nevertheless, in all cases, automatic PCR cyclers were employed for DNA amplification. Fragment analysis was carried out on one of the genetic analyzers (ABI310, ABI3100 or ABI3130) produced by Applied Biosystems (Foster City, CA) [3]. Fragment sizes were determined using an internal lane standard (GS 400 HD) together with a set of standard DNA samples.

Of the 401 families in our study, 48 were genotyped for the Argus X-8 set only (1 of type I, 47 of type II). Another 17 families (all of type I) were also typed for markers DXS10079 and DXS10103. Six families (all of type I) were genotyped for all markers except DXS10148 whereas the remaining 330 families were typed for all 12 markers, with only sporadically missing genotypes. An overview of the genotyping scope is provided in Table 3.

Data quality control

One type I family had to be removed from the data because the maternal genotype information was incomplete for all loci. Genotypes at single loci were missing in several families (either sporadically or systematically). In order to allow these families to nevertheless contribute linkage information at other loci, all family members were assigned identical homo- or hemizygous genotypes, respectively, to render the locus in question uninformative for linkage.

Eight genotype incompatibilities (6 in type I families, 2 in type II families) could be attributed to single-step mutations (Table 3). Since the grandparental genotype in type I families was used only to infer maternal phase (see below), grandparental alleles were set equal to the maternal ones in these cases. One apparently erroneous offspring genotype in a type II family was corrected on the basis of the other genotypes in the family (arguing that the original genotype entry would have required either a double recombination or multiple mutation events, which was deemed extremely unlikely). In another type II family with five sons, one son was

excluded from the data because an apparently erroneous genotype at one STR could not be resolved with certainty. In all other cases of unresolved genotype incompatibilities, the respective marker was made uninformative for linkage by assigning all family members the same homo- or hemizygous genotype.

Likelihood formulation

Estimation of the recombination fraction between adjacent X-STRs was based upon comprehensive likelihood calculations taking the possibility of single-step mutation into account. Mutations of more than one repeat, or towards fractions of a repeat, were deemed too unlikely to warrant explicit consideration. In those rare instances where a multi-step mutation could not be ruled out, genotypes were made uninformative.

Families analyzed in the present study were of one of two types, either grandfather-mother-sons (type I) or mother-sons (type II). For both types, the likelihood of a family genotype can be explicitly formulated as follows: If there are n X-STRs of known physical order, let $(\theta_1,\ldots,\theta_{n-1})$ denote the pair-wise recombination fractions between adjacent markers, and let μ be the (uniform and symmetric) one-step mutation rate. For a given mother-son pair, let $V\in\{1,2\}^n$ denote the so-called 'inheritance vector' of the son, where V(i) indicates the grandparental origin of the allele at the i-th STR (i.e., V(i)=1 for grandpaternal, V(i)=2 for grandmaternal).

The conditional likelihood of a particular genotype $g_s(i)_{i=1..n}$ of a son, given the phased maternal genotype $g_m(i,j)_{i=1..n,j=1..2}$ and an inheritance vector V, equals

$$(1)\ L(g_s \mid g_m,\theta_1,\ldots,\theta_n,\mu,V)=\prod_{i=1}^{m-1}[\theta+(1-2\theta)\cdot 1_{V(i)=V(i+1)}]\times\prod_{i=1}^{m}[\mu+(1-3\mu)\cdot 1_{g_m(i,V(i))=g_s(i)}].$$

The conditional likelihood of the genotype of the son, given the phased maternal genotype alone, is obtained by summing the term in (1) over all 2^n possible inheritance vectors, i.e.

$$(2)\ L(g_s \mid g_m,\theta_1,\ldots,\theta_n,\mu)=\sum_{j=1}^{2^n}L(g_s \mid g_m,\theta_1,\ldots,\theta_n,\mu,V_j).$$

In type I families, knowledge of the grandpaternal genotype allows phasing of the maternal genotype so that the likelihood of the whole family is simply the product of the son-specific likelihoods given in (2). In type II families, in contrast, the maternal phase is usually unknown and the likelihood calculation has to take this uncertainty into account by summing the aforementioned likelihood products over all 2^n possible

maternal phases. Finally, the likelihood of the total data is obtained by multiplying all family-specific likelihoods.

Likelihood maximization

To our knowledge, there is currently no publicly available computer program that would allow estimation of recombination fractions between STRs using the fully comprehensive likelihood model formulated in the previous section. Therefore, maximum-likelihood estimates (MLE) of the recombination fractions between adjacent markers were obtained in the present study by means of statistic software R v 2.13.0 [18], using in-house scripts. More specifically, likelihood maximization with respect to recombination fractions was carried out with the *optim* function, which employs the so-called 'L-BFGS-B' method [19] to allow for box constraints on parameters. To avoid numerical underflow, recombination rates were limited to the interval between 10^{-8} and 0.5. The one-sided mutation rate μ was set equal to 0.001 for all markers, which represents an average of recent estimates for X-STRs [20, 21], and which also agrees well with the findings of the present study (see Supplementary Table 1). In fact, we observed 8 mutations in 10,290 informative meioses, which corresponds to a two-sided μ of 7.8×10^{-4} (Table 3). However, this figure is likely to be an underestimate of the true mutation rate because mutations were systematically overlooked in cases where the mutated allele was identical to the other maternal allele.

Recombination fractions were estimated in two phases, first considering only the eight markers of the Mentype® Argus X-8 kit, then comprising all 12 X-STRs. Three different sets of starting values were chosen for each maximization round.

1. Recombination fractions were interpolated from known physical inter-marker distances (Table 2) applying a rule-of-thumb whereby 1 Megabase (Mb) of DNA corresponds to a genetic distance of 1 centiMorgan (cM). Genetic distances are derived from recombination fractions, and *vice versa*, by so-called 'mapping functions' [22, 23]. Here, we transformed genetic distances into θ values using Kosambi's mapping function [24].
2. All recombination fractions between adjacent markers were set equal to 0.25.
3. Complete linkage was assumed within linkage groups (i.e., $\theta=10^{-8}$) and free recombination was assumed between linkage groups (i.e., $\theta=0.5$).

Results

The main goal of the present study was to estimate as precisely as possible the recombination fractions between 12 X-chromosomal STRs in current forensic use. Genotyping of three X-STRs per linkage group yielded a highly polymorphic system in all four groups. Thus, maternal genotypes that were completely uninformative for linkage between adjacent groups were only rare (Table 3). As a consequence, most recombination events between adjacent linkage groups could be identified directly in the 270 to 279 informative meioses comprising our type I families. In type II families, between 397 and 406 informative meioses were available for analysis (Table 3), but identification of recombination events was less certain because of the inherent uncertainty about maternal phase.

Interestingly, 45 apparent recombination events in our data set were also explicable by a single-step mutation. A particularly illustrative example of the resulting ambiguity is provided by Portuguese type I family 252 (Fig. 1). While the single male offspring in this case clearly has inherited markers DXS10148 to DXS10074 from his grandmother and markers DXS10101 to DXS7423 from his grandfather, the location of the intermittent recombination breakpoint is unclear. If allele 19 of DXS10103 was indeed of grandmaternal origin (scenario A), as suggested by its actual repeat number, then this would have required a recombination within linkage group III. Alternatively, the boy could have inherited the whole of grandpaternal linkage group III if allele 18 has mutated to allele 19 in his mother's meiosis (scenario B). In conclusion, since the likelihood of X-STR mutations may be of similar or even higher order than that of recombination within linkage groups, estimation of recombination fractions by mere counting and exclusion of ambiguous transmission events was deemed too unreliable. Instead, a comprehensive likelihood analysis of the available data was performed, including simultaneous allowance for both mutation and recombination in strictly formalized fashion.

Maximum likelihood estimation revealed that only linkage groups I and II are unlinked in the sense that the recombination fraction between them equals 0.5 (Table 4). This result was obtained irrespective of whether only the Mentype® Argus X-8 kit markers or all 12 STRs were taken into consideration in the analysis. We also found evidence for recombination within linkage group I, particularly between DXS10148 and DXS10135 (Table 4) for which the recombination fraction was estimated to be ~1%. Between linkage groups II and III, and between groups III and IV, a

considerably reduced recombination fraction was inferred in our study, with maximum likelihood estimates equal to 0.4274 for DXS10074 - DXS10103 and 0.3227 for DXS10101 - DXS10146 (Table 4). Intra-group linkage was also found to be less than perfect within linkage groups II to IV, with recombination fraction estimates ranging up to 0.0199 for DXS10146 - DXS10134.

Discussion

As has been demonstrated many times before, exact likelihood calculation in kinship testing with physically linked markers requires the consideration of both linkage and linkage disequilibrium (LD), not only for the X chromosome [1, 6, 12, 13], but in general [25]. The computational relevance of the two characteristics is a function of their actual tightness and strength, which implies that it may be admissible to numerically treat very loosely linked makers (or groups of markers) as if they were located on different chromosomes. For closely linked markers, however, it is usually impossible to make any *ex ante* predictions as to what extent negligence of linkage and/or LD will inflate or deflate the likelihood ratio in a given case. Therefore, it seems advisable to perform most precise likelihood calculations in all instances of kinship analysis, irrespective of the hypotheses under consideration.

The tightness of linkage between two genetic loci is measured by their recombination fraction, rather than their physical distance, because recombination intensity is known to vary in the genome, i.e. even equidistant loci on one and the same chromosome may recombine with different probability. Therefore, extrapolation of genetic distances from physical maps, or from existing genetic maps using physical distance as an extrapolation basis, can only provide a provisionary substitute of an empirical estimation of recombination fractions. Anyhow, publicly available genetic maps such as that of the Laboratory of Computational Genetics at Rutgers University [26] strongly suggested that X-STR linkage groups I to III were loosely linked whereas the genetic distance between groups III and IV was extrapolated as ~34 cM (Table 1). All intra-group genetic distances were reported to be well below 1 cM. In order to validate these figures, and to facilitate reliable use of X-STRs in practical kinship analysis, we choose to expand the existing German family data [15] into an international collaborative study.

As regards the observation of free recombination between linkage groups I and II, our analysis corroborates both the Rutgers map and findings by Tomas et al. [27], but

contradicts the results of both Pamjav et al. [28] and Tilmar et al. [13] who independently claimed a somewhat reduced recombination fraction between these two linkage groups (Table 4). Just the opposite picture emerged for linkage groups II and III, where the analyses by Tomas et al. [27] and Tilmar et al. [13] suggested free recombination, in line with the Rutgers map, whereas our study, like that of Pamjav et al. [28], was indicative of a recombination fraction of ~0.42. Finally, all studies were found to agree about the notably reduced probability of recombination between linkage groups III and IV, where our estimate of 0.3255 is close to that of Tomas et al. [27].

Our comprehensive linkage analysis has provided evidence for non-negligible internal recombination in all four linkage groups. Compared to other studies that suggested some intra-group recombination fractions to be notably larger than zero [13, 27], our estimates were predominantly higher (Table 4). This is not surprising because none of the other studies was simultaneously taking the possibility of mutation and recombination into account. Instead, ambiguous cases like the one depicted in Fig. 1 were either disregarded or treated as definitive mutations, owing to the small physical distance between STRs in the same linkage group. However, both approaches lead to a downward bias of the recombination fraction estimates and are therefore inferior to a thorough likelihood-based analysis of the complete genotype data.

As we have noted above, it may sometimes be impossible to distinguish between recombination and mutation as the true cause of an apparent recombination. In such cases, typing of flanking markers could be useful. International collaborations therefore seem warranted to establish additional X-chromosomal markers, even if they may not be used routinely in forensic case work. In any case, the most appropriate way to allow for ambiguities in the assignment of recombination breakpoints would be to use suitable pedigree analysis software. We have previously advocated use of the MLINK program originally developed for mapping and risk calculations in Mendelian disease genetics [12]. However, MLINK and other pedigree analysis packages have rather limited capabilities, not only in terms of handling mutations, but also as regards the number and variability of markers included. Eventually, this was the reason for using in-house scripts for likelihood calculations in the present study. There can be no doubt that the development of powerful software

tools specifically tailored to the needs of comprehensive STR analysis would an exercise highly welcome by the forensic genetics community.

In order to allow better judgment of the relative importance of meiotic X-STR mutation, either in population studies or in individual kinship case work, we have collected mutation rates from the literature and added this information to the present study as electronic supplementary material (Supplementary Table 1). It must be taken into account, however, that these figures may be underestimates of the true mutation rates because not all mutations that actually occurred may have been detected in the respective studies.

The Mentype® Argus X-8 kit has become well established in forensic case work, and its expanded successor, the Argus-X 12 kit, is now available. As regards recombination, our analysis presented here has helped to derive a basis for correct likelihood calculation in kinship testing using both kits. However, recombination fractions and mutation rates are only one half of the story. For the future, the most important challenge is to derive sufficiently accurate haplotype frequency estimates for the four linkage groups, in different world populations. The samples used here for studying recombination and mutation originated from different populations and ethnicities, and therefore cannot provide an appropriate population database for estimating haplotype frequencies. Therefore, we recommend that scientists who take the trouble to genotype their local populations share their data with the forensic genetics community at www.chrx-str.org.

References

[1] R. Szibor, S. Hering, E. Kuhlisch, I. Plate, S. Demberger, M. Krawczak, J. Edelmann, Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing, Int. J. Legal Med. 119 (2005) 363-369.

[2] H. Rodig, F. Kloep, L. Weissbach, C. Augustin, J. Edelmann, S. Hering, R. Szibor, F. Gotz, W. Brabetz, Evaluation of seven X-chromosomal short tandem repeat loci located within the Xq26 region, Forensic Sci. Int. Genet. 4 (2010) 194-199.

[3] 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.

[4] J. Edelmann, S. Hering, E. Kuhlisch, R. Szibor, Validation of the STR DXS7424 and the linkage situation on the X-chromosome, Forensic Sci. Int. 125 (2002) 217-222.

[5] C. Augustin, R. Cichy, S. Hering, J. Edelmann, E. Kuhlisch, R. Szibor, Forensic evaluation of three closely linked STR markers in a 13 kb region at Xp11.13, Int. Congr. Ser. 1239 (2006) 311-314.
[6] 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.
[7] D. Becker, H. Rodig, C. Augustin, J. Edelmann, F. Gotz, S. Hering, R. Szibor, W. Brabetz, Population genetic evaluation of eight X-chromosomal short tandem repeat loci using Mentype Argus X-8 PCR amplification kit, Forensic Sci. Int. Genet. 2 (2008) 69-74.
[8] R. Nagaraja, S. MacMillan, J. Kere, C. Jones, S. Griffin, M. Schmatz, J. Terrell, M. Shomaker, C. Jermak, C. Hott, M. Masisi, S. Mumm, A. Srivastava, G. Pilia, T. Featherstone, R. Mazzarella, S. Kesterson, B. McCauley, B. Railey, F. Burough, V. Nowotny, M. DUrso, D. States, B. Brownstein, D. Schlessinger, X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content, Genome Res. 7 (1997) 210-222.
[9] Kong, D.F. Gudbjartsson, J. Sainz, G.M. Jonsdottir, S.A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, A. Shlien, S.T. Palsson, M.L. Frigge, T.E. Thorgeirsson, J.R. Gulcher, K. Stefansson, A high-resolution recombination map of the human genome, Nat. Genet. 31 (2002) 241-247.
[10] S. Myers, L. Bottolo, C. Freeman, G. McVean, P. Donnelly, A fine-scale map of recombination rates and hotspots across the human genome, Science 310 (2005) 321-324.
[11] R. Szibor, X-chromosomal markers: Past, present and future, Forensic Sci. Int. Genet. 1 (2007) 93-99.
[12] M. Krawczak, Kinship testing with X-chromosomal markers: Mathematical and statistical issues, Forensic Sci. Int. Genet. 1 (2007) 111-114.
[13] A.O. Tillmar, P. Mostad, T. Egeland, B. Lindbolm, G. Holmlund, K. Montelius, Analysis of linkage and linkage disequilibrium for eight X-STR markers, Forensic Sci. Int. Genet. 3 (2008) 37-41.
[14] F.B. Machado, E. Medina-Acosta, Genetic map of human X-linked microsatellites used in forensic practice, Forensic Sci. Int. Genet. 3 (2009) 202-204.
[15] S. Hering, J. Edelmann, C. Augustin, E. Kuhlisch, R. Szibor, X chromosomal recombination-a family study analysing 39 STR markers in German three-generation pedigrees, Int. J. Legal Med. 124 (2010) 483-491.
[16] T. Hundertmark, S. Hering, J. Edelmann, C. Augustin, I. Plate, R. Szibor, The STR cluster DXS10148-DXS8378-DXS10135 provides a powerful tool for X-chromosomal haplotyping at Xp22, Int. J. Legal Med. 122 (2008) 489-492.
[17] J. Edelmann, S. Hering, C. Augustin, R. Szibor, Characterisation of the STR markers DXS10146, DXS10134 and DXS10147 located within a 79.1 kb region at Xq28, Forensic Sci. Int. Genet. 2 (2008) 41-46.
[18] R Development Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2011)
[19] R.H. Byrd, P.H. Lu, J. Nocedal, C.Y. Zhu, A Limited Memory Algorithm for Bound Constrained Optimization, SIAM J. Sci. Comput. 16 (1995) 1190-1208.
[20] Q.L. Liu, D.J. Lu, X.G. Li, H. Zhao, J.M. Zhang, Y.K. Lai, Y.F. Chen, Development of the nine X-STR loci typing system and genetic analysis in three nationality populations from China, Int. J. Legal Med. 125 (2011) 51-58.

[21] M. Castañeda, V. Mijares, J.A. Riancho, M.T. Zarrabeitia, Haplotypic blocks of X-linked STRs for forensic cases: Study of recombination and mutation rates, J. Forensic Sci. doi:10.1111/j.1556-4029.2011.01916.x.
[22] 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.
[23] R. Nagaraja, S. MacMillan, C. Jones, M. Masisi, G. Pengue, G. Porta, S. Miao, A. Casamassimi, M. D'Urso, B. Brownstein, D. Schlessinger, Integrated YAC/STS physical and genetic map of 22.5 Mb of human Xq24-q26 at 56-kb inter-STS resolution, Genomics 52 (1998) 247-266.
[24] D.D. Kosambi, The estimation of the map distance from recombination values, Ann. Eugen. 12 (1944) 172-175.
[25] M. Nothnagel, J. Schmidtke, M. Krawczak, Potentials and limits of pair-wise kinship analysis using autosomal short tandem repeat loci. Int. J. Legal Med. 124 (2010) 205-215.
[26] T.C. Matise, F. Chen, W.W. Chen, F.M. De la Vega, M. Hansen, C.S. He, F.C.L. Hyland, G.C. Kennedy, X.Y. Kong, S.S. Murray, J.S. Ziegle, W.C.L. Stewart, S. Buyske, A second-generation combined linkage-physical map of the human genome, Genome Res. 17 (2007) 1783-1786.
[27] C. Tomas, V. Pereira, N. Morling, Analysis of 12 X-STRs in Greenlanders, Danes and Somalis using Argus X-12, Int. J. Legal Med. doi:10.1007/s00414-011-0609-y
[28] H. Pamjav, R. Kugler, A. Zalán, A. Völgyi, Z. Straky, P. Endrédy, Z. Kozma, X chromosomal recombination study in three-generation families in Hungary, Forensic Sci. Int. Genet. doi:10.1016/j.fsigen.2011.08.009

Table 1: Physical and genetic localisation of 12 X-STRs, according to NCBI built 36 and the Rutgers Map v.2, respectively

Marker	Linkage group	Cytogenetic localisation	Physical localisation [Mb]	Genetic localisation [cM]
DXS10148		Xp22.31	9.198	19.84[a]
DXS10135	I	Xp 22.31	9.199	20.03[a]
DXS8378		Xp 22.31	9.330	20.20[b]
DXS7132		Xcen	64.572	90.75[b]
DXS10079	II	Xq12	66.632	90.82[a]
DXS10074		Xq12	66.894	90.83[a]
DXS10103		Xq26.2	133.246	149.37[a]
HPRTB	III	Xq26.2	133.443	149.66[b]
DXS10101		Xq26.3	133.482	149.75[a]
DXS10146		Xq28	149.335	183.72[a]
DXS10134	IV	Xq28	149.401	183.96[a]
DXS7423		Xq28	149.460	184.19[a]

a: calculated from the respective physical marker position using the Rutgers Map Interpolator (http://compgen.rutgers.edu/old/map-interpolator/); b: derived from actual recombination fraction estimates provided by the Rutgers Map v.2 using Kosambi's mapping function [24]

Table 2: Primer sequences and positions of four X-STRs not included in the Mentype® Argus X-8 kit

Marker	Primer labelling and sequence	amplicon length [bp]	3' primer position [bp]
DXS10148-F	Hex-AAAAAAGGGGGAAGGAAGGA	215 - 262	9,198,969
DXS10148-R	GGCTATTTCTCCTGCATAAG		9,199,205
DXS10079-F	Fam-GAGAATGGCTTGAACCTGG	313 - 357	66,632,537
DXS10079-R	GTTTGCCTGTGTTGTAACATCCT		66,632,882
DXS10103-F	Hex -TCATAATCACATATCACATGAGC	160 - 200	133.246.578
DXS10103-R	AAACAGAACCAGGGGAATGAA		133.246.757
DXS10146-F	Fam-CTGCCTTGCCCTTCCTACC'	178 - 268	149,334,927
DXS10146-R	GAAAAAGAAAGAAAGACAGAGA'		149,335,115

Table 3: X-STR genotype data used for linkage analysis

Marker[a]	No. linkage-informative meioses[b]				No. mutations[c]
	Type I families		Type II families		
DXS10148	208		260		2 (742)
DXS10135	255	274	378	403	1 (902)
DXS8378	177		278		1 (900)
DXS7132	213		323		1 (901)
DXS10079	235	277	237	406	0 (796)
DXS10074	242		336		1 (903)
DXS10103	234		244		0 (797)
HPRTB	214	279	312	398	0 (903)
DXS10101	245		335		0 (900)
DXS10146	228		248		0 (745)
DXS10134	226	270	353	397	2 (900)
DXS7423	173		270		0 (901)

a: STRs not included in the Mentype® Argus X-8 kit are highlighted. b: Given for each family type is the number of sons with a mother that was heterozygous for the marker in question (left column) or for at least one marker from the respective linkage group (right column). c: The figure in brackets is the number of mutation-informative meioses.

Table 4: Estimates of the recombination fraction between adjacent X-STRs

Marker interval	All 12 loci			Mentype® Argus X-8	
	Present study	Tomas et al. 2011 [27]	Pamjav et al. 2011 [28]	Present study	Tillmar et al. 2008 [13]
DXS10148 - DXS10135	0.0106	0.0001	-.-	-.-	-.-
DXS10135 - DXS8378	0.0000	0.0001	-.-	0.0015	0.00
DXS8378 - DXS7132	0.5000	0.5000	0.387 [a]	0.5000	0.45
DXS7132 - *DXS10079*	0.0064	0.0132	-.-	0.0106	0.01
DXS10079 - DXS10074	0.0080	0.0001	-.-		
DXS10074 - *DXS10103*	0.4274	0.5000	0.400 [a]	0.4144	0.50
DXS10103 - HPRTB	0.0095	0.0001	-.-		
HPRTB - DXS10101	0.0000	0.0001	-.-	0.0000	0.00
DXS10101 - *DXS10146*	0.3227	0.3142	0.367 [a]	0.3178	0.25
DXS10146 - DXS10134	0.0199	0.0001	-.-		
DXS10134 - DXS7423	0.0000	0.0078	-.-	0.0017	0.02

a: The paper by Pamjav et al. [27] reported only recombination fractions between, but not within, linkage groups.

Fig. 1: Example of an ambiguous X-STR transmission event provided by family 252. The genotype of individual 252–904 is explicable either by a maternal recombination between markers DXS10103 and HPRT within linkage group III (A) or by a recombination between DXS10074 (linkage group II) and DXS10103 (linkage group III) plus a single-step mutation from 18 repeats to 19 repeats at DXS10103.