

Research paper

Extended population genetic analysis of 12 X-STRs – Exemplified using a Norwegian population sample

Erik F. Bergseth^{a,1}, Andreas Tillmar^{b,c,1}, P. Jørgen T. Haddeland^a, Daniel Kling^{a,b,*}^a Department of Forensic Sciences, Oslo University Hospital, Pb. 4950 Nydalen, NO-0424 Oslo, Norway^b Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden^c Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Sweden

ARTICLE INFO

Keywords:

Argus X-12
X-chromosome
Linkage disequilibrium
Allelic association
Simulations

ABSTRACT

The use of X-chromosomal markers to resolve questions of relatedness has experienced a significant increase during the last years in forensic genetics. Perhaps primarily due to the emergence of commercial kits, but equally important due to an increased awareness of the utility of those markers. The X-chromosomal inheritance pattern entails that some cases, for instance paternal half-sisters, can potentially be resolved using a few X-chromosomal markers alone. For the statistical assessment in kinship cases it is of importance to have relevant population frequency data. In the present study 631 unrelated males from a Norwegian population sample are analyzed. The resulting haplotypes are compared to previously studied population samples and a deeper analysis of the linkage disequilibrium (LD) structure is conducted. We demonstrate that the power to detect LD will be low when few males, say below 300, are analyzed. We use entropy to describe the degree of LD between multiallelic loci and describe how this measure varies between different studied populations. Large population frequency databases have been recommended when using X-chromosomal markers, and we show that by combining reference databases from genetically similar populations, more precise haplotype frequency estimates can be obtained for rare haplotypes which improves the statistical assessment of the weight of evidence. In addition, we promote the use of simulations to assess the utility of STR markers in contrast to standard forensic parameters. Specifically we perform extensive simulations on cases where X-chromosomal markers are important and illustrate how the results can be used to infer the information gained from these markers.

1. Introduction

In forensic genetics, X-chromosomal markers have evolved to become a central battery of additional markers in more complex relationship cases [1–7]. The key lies in the particular inheritance pattern of the X chromosome. X-chromosomal markers display a hemizygous pattern where males, normally, possess one copy, inherited from the mother and females possess two copies, one from each parent. Gomes et al. [8] recently reviewed the use of X-chromosomal markers in forensics, while Pinto et al. [5,6] provide a general guideline to when X-chromosomal markers are useful in kinship testing. During meiosis, only the maternal X-chromosome is subject to crossovers, with the exception of a smaller section of homologous parts of the Y and X chromosomes in males. As a consequence the degree of association between alleles (also known as linkage disequilibrium, LD, or

gametic/allelic association) at different markers may be stronger in the population due to fewer recombination events.

Relevant and accurate population haplotype frequency estimates are of importance for the statistical assessment in kinship cases (e.g. calculations of case specific likelihoods and associated likelihood ratios). Tillmar et al. put forward that forensic X-STR testing generally requires larger population reference databases compared to when using standard autosomal STR [7] mainly a consequence from the requirement to model haplotypes. In particular, the large number of expected haplotypes, often exceeding 1000, entails that large databases are needed to capture the variation, but also appropriately address unobserved haplotypes (i.e. haplotypes not included in the population database). Although a statistical framework to handle this has been published (e.g. the lambda model [3,4]), frequency estimates based on observed haplotype data are usually more precise than estimates based on a statistical model. One

* Corresponding author at: Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden.

E-mail address: daniel.kling@rmv.se (D. Kling).

¹ Shares first authorship.

way to increase the size of a reference database is to merge observed haplotype data from genetically close populations. Prior to merging databases, in depth population comparisons shall however be performed to show the validity of the new super-population databases. In this study we exemplify how this can be done and study the effects of it, specifically for cases with rare haplotypes.

In this paper we analyze the genetic structure of a Norwegian population sample, based on X-chromosomal marker data. Specifically we aim to establish a better understanding of the LD structure for the set of markers studied in this paper and also in other similar studies. While genetic linkage is a concept related to the inheritance of haplotypes in a pedigree, LD is connected to the association of alleles in a population. Each population displays its own patterns of LD due to historical events, e.g. admixture, selection pressure etc. We explore different population genetic analyses and use entropy to measure the degree of LD between pairs of markers. Normalized entropy has previously been suggested as a measure of LD between multiallelic loci by Nothnagel et al. [9].

A standardized set of forensic parameters are commonly reported when population data on genetic markers are presented, see in particular Ferragut et al. for discussion on X-chromosomal markers [10]. These parameters typically relates to the frequencies of individual alleles or haplotypes in different combinations. For instance the exclusion power (PE) for a set of genetic markers can inform us how likely we are to exclude a false parent. Furthermore, typical paternity index (TPI) and power of discrimination (PD) yields other measures used to assess the expected power. Desmarais et al. described how to compute some forensic parameters for X-chromosomal markers [11]. However, since the application for X-chromosomal markers is mainly pedigrees more complex than standard paternity and maternity cases, Pinto et al. detailed how to calculate information measures for a number of kinship cases with high relevance for X-chromosomal data [12].

Simulations are a versatile tool to assess and understand the potential of a particular set of genetic markers for kinship investigations. In comparison to the afore-mentioned traditional forensic parameters, simulations will provide a better understanding of the expected information potential and the utility of our marker panel for any given case see for instance Ge et al. [13] summarizing data for autosomal markers and [Supplementary Tables 1 and 2](#) in Tillmar et al. for data relating to X-chromosomal markers [7].

The aim of this study is to 1) establish a haplotype frequency database (comprising 12 X-chromosomal STR markers) for the Norwegian population, 2) evaluate alternative approaches to analyze and detect allelic dependencies 3) use the combination of genetically proximal population to create larger databases and 4) provide evaluations of expected information power for a number of important cases for X-chromosomal markers. We use freely accessible software to perform all analyses and will supply all generated code as [Supplementary material](#) for others to reproduce or alter for their own purposes.

2. Material and methods

2.1. Data

Samples from 680 unrelated Norwegian males with self-declared Norwegian ancestry were randomly selected from paternity cases. As the ancestry is self-declared and no further investigations into the grandparental ancestries were conducted, a small bias towards other ancestries could be present, however most likely to a small degree. All samples had previously been analyzed for 23 overlapping autosomal markers in case work (unpublished data) which was used for an introductory blind search. The samples consisted of buccal cells on FTA-cards and were collected and analyzed at Oslo University Hospital, Norway according to the current Data Processing Agreements between Oslo University Hospital and the Norwegian Courts Administration, Tax Administration and Work and Welfare Administration respectively.

A sweep of published literature was conducted to map publications

on X-chromosomal STR data. In particular we focused on the markers in the Investigator Argus X-12 QS kit (Qiagen), see [Supplementary Table 1](#). Complete haplotype data from previous studies (spanning publication years 2012–2019) were included for subsequent comparison, listed in [Table 1](#). We refer to these data as our primary data set. Additional publications [14–25] listed in [Supplementary Table 1](#), with only cluster specific haplotypes were considered for a subset of the analyzes. We refer to these data as our secondary data set. The inclusion of published data is not exhaustive but provide a reasonable representation of what is published, see [Supplementary Table 2](#) in Gomes et al. [8] for a more comprehensive summary of studies.

2.2. Genetic analyses

The samples (680 unrelated Norwegian males) were amplified using the standard protocol of the Investigator Argus X-12 QS (Qiagen), according to manufacturer's protocol but using half-volume reactions. The kit divides 12 STR markers on the X-chromosome into four distinct clusters (also known as linkage group), each including three markers (see [Supplementary Table 1](#)). Subsequent capillary electrophoresis was performed on a 3500xl Genetic Analyzer (Thermo Scientific) with default settings on a POP4 36 cm capillary. Raw data was analyzed in GeneMapper ID-X 1.4 (Thermo Scientific) using RFU = 100 as threshold to determine genotypes.

2.3. Initial blind search

Using previously generated autosomal data (23 STRs) for the 680 unrelated Norwegian males, a blind search was performed with the intention to prune pairs of unknown relatives. Briefly the blind search creates all pairwise combinations, in total $679 \times 340 = 230,860$ comparisons. For each combination two likelihood ratios (LR) were computed, (i) one comparing the likelihood for a parent-child relation with the likelihood for unrelated, and (ii) one comparing the likelihood for a full sibling relation with the likelihood for unrelated. Using an arbitrary LR cut-off equal to 1000, 631 males were retained (59 individuals were excluded due to a potential parent/child or sibling relation). All LR calculations were performed in the software Familias [33,34] (version 3.2.9) using inhouse allele frequency data from a Norwegian population [35]. For the parent/child comparisons we considered an extended stepwise model with mutation parameters based on an inhouse dataset (unpublished data). A Familias project with all the parameters is freely available at https://familias.name/Familias_databases/Norwegian_DB.fam.

2.4. Biostatistical analyses

To analyze the resulting haplotype data, a number of different approaches were considered. First we compute summary statistics of the haplotype data in our Norwegian population sample using custom scripts in R. To this end we used formulas presented in Desmarais et al.

Table 1

Description of population samples with complete DNA-profiles used in this study. Number of male samples (Haplotypes) as well as reference is listed. All studies contain data analyzed with either the Argus X12 or the updated Argus X12 QS kit.

Population	Haplotypes	Ref.
Norway	631	This study
Sweden	652	[26]
United Arab Emirates	501	[27]
Mexico	933	[28]
Germany	1034	[29]
Somalia	673	[30]
Hungary	219	[31]
Argentina	914	[32]

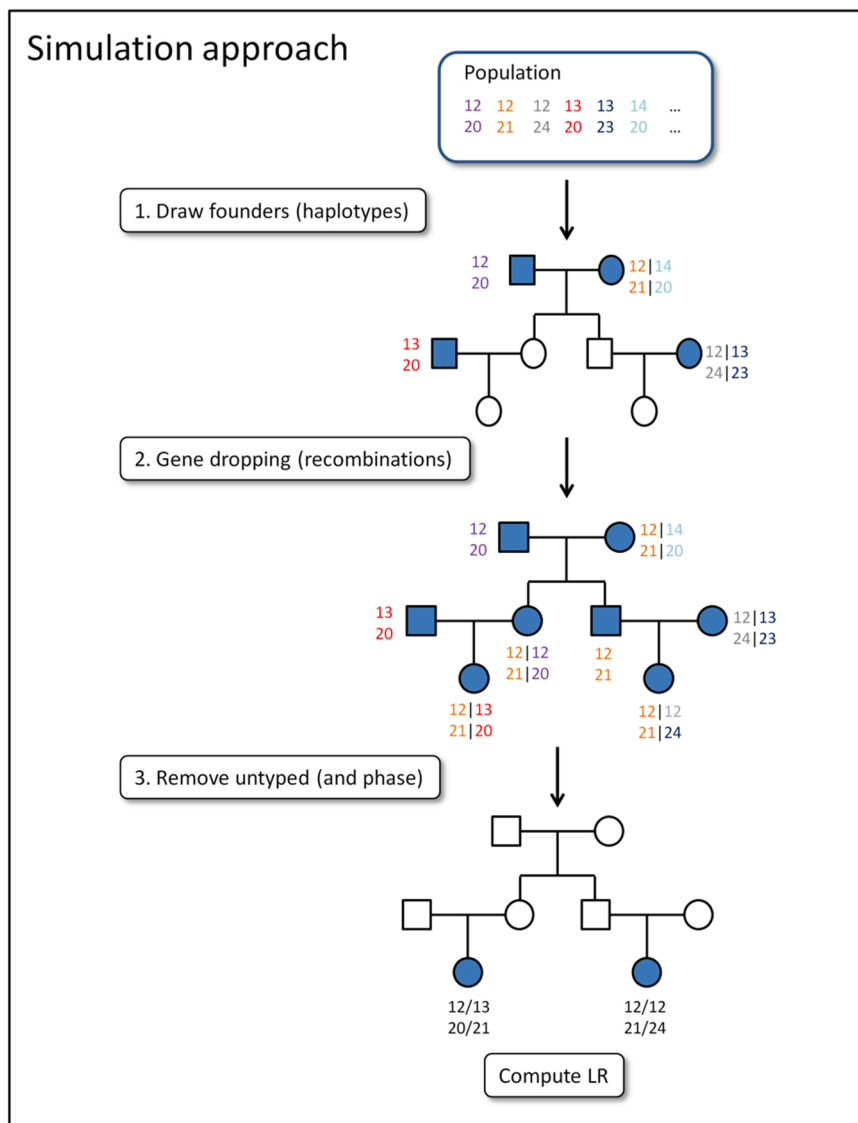


Fig. 1. Illustration of the simulation procedure employed in this study. The process start by randomly drawing haplotypes based on a frequency distribution. The haplotypes are transmitted throughout the pedigree in a process known as gene dropping. Individual alleles are transmitted throughout the pedigree using the standard rules of inheritance where recombinations are considered for each transmitted allele. In the final step, knowledge about phase, i.e. paternal and maternal chromosomes, are removed.

[11] where for instance the power of discrimination are computed for males (PDM) and females (PDF) separately. We further used formulas presented in Pinto et al. [12] to compute exclusion probabilities for paternal half sisters and paternal grandmother/granddaughter with data available for the mother, which providing measures with higher relevance to the application of X-chromosomal testing. Next we investigate the degree of linkage disequilibrium (LD) in our Norwegian population sample as well as in previous population studies (see Table 1 and Supplementary Table 2) using an approach previously described for Y-chromosomal markers [9] and laid out in detail below. Thirdly, we compute so called R_{st} values comparing the genetic distance between populations. In contrast to F_{st} values, the R_{st} also takes into account the repeat length and thus fully incorporating the STR marker data in the model. The R_{st} values are further used to create dendrograms as well as multidimensional scaling plots. Lastly we perform extensive simulations where we illustrate the potential of X-chromosomal markers, also contrasting their potential with autosomal markers in a selection of cases. All calculations and plotting are performed using inhouse scripts in the software R (version 4.1.1), available from the authors upon request.

2.4.1. Linkage disequilibrium

Association between alleles at two (or more) loci in a population is commonly referred to as linkage disequilibrium (LD). The name is un-

fortunately sometimes confused with genetic linkage, another concept related to inheritance in pedigrees. For biallelic markers the degree of LD can be defined as $D = p_{ab} - p_a p_b$ where D is the difference between the expected combination of alleles a and b , denoted p_a and p_b at two different markers and their observed rate of appearance in a population, denoted p_{ab} . Commonly the normalized measure r^2 is used. To extend this to multi-allelic markers and also to cover several loci, Nothnagel et al. [36] and Nothnagel and Rohde [37] detecting LD in haploblocks of SNP and later also Siegert et al. [9] measuring LD in Y-STRs, suggested the use of entropy. More specifically, op sic use Shannon's equivocation commonly employed in information theory,

$$H = - \sum_{s=1}^S p_s \log p_s$$

where H is the entropy for a genetic marker (or a combination of markers) and the allele (or haplotype) frequencies are given by p_s , where s represents each allele or haplotype.

To compare the information for two (or more) marker combinations Nothnagel et al. further used what they refer to as the normalized entropy difference (NED).

$$NED_{ij} = 2 \times \frac{(H_i + H_j) - H_{ij}}{H_i + H_j}$$

Where the NED for a pair of markers i and j is computed as the difference between the sum of their individual entropies and the two markers as a unit, i.e. with haplotypes and further divided by the sum of their individual entropies.

We contrast NED, which provides a quantitative measure of LD, with p-values, which only provides evidence for the presence of LD, commonly adopted to analyze if a certain combination of alleles deviates from linkage equilibrium. We use Fisher's exact test to best accommodate haplotypes with few or no observations. All analyses are performed in R using custom inhouse scripts and core functions.

2.4.2. Population differentiation and structure

There are several measures that can be used to describe a genetic population and its relation to other populations. We computed R_{st} using the software Arlequin (version 3.5.2.2) [38]. As previously mentioned, the R_{st} computes the distance between two haplotypes also taking the step-wise mutation mechanism of STR markers into account. Further, to account for the presence of microvariants, e.g. 13.1, all such variants were re-coded such that, for instance 13.1 became 113 and 14.1 became 114 to mimic the large genetic distance from alleles 13 and 13.1, but the short distance between alleles 13.1 and 14.1. We use core R functions *cmdscale* and *hclust* to visualize the resulting distance like matrix as a multidimensional scaling (MDS) and a dendrogram plot respectively. The *hclust* function is run with average linkage as clustering method.

We perform a population structure analysis based on our primary data set listed in Table 1 where complete haplotype data is available in the software STRUCTURE (version 2.3) [39]. Briefly the software uses a statistical model to assign the fraction of overarching populations for each included individual and iterates until the model is assumed to converge. It should be noted that STRUCTURE will not in its current implementation fully handle the haplotype structure of the current data set since it can only handle weakly linked markers and will only use allele frequency data rather than haplotypes. Given the above limitations, the results from STRUCTURE are only used for visualization purposes.

2.4.3. Description of simulations

To generate data we use a top-to-bottom simulation approach (see Fig. 1) where genotypes of founders (the outmost nodes of a pedigree) are sampled based on haplotype frequencies. We subsequently use gene-dropping [40] to iteratively generate the genotypes of non-founders assuming no mutations during transmissions. Throughout the simulations, information about the phase of the simulated haplotypes are contained and in combination with a genetic linkage map used to simulate crossover events. In fact, we use recombination values from a study by Nothnagel et al. [1]. In the final stage of the simulations, information about haplotype phase is removed as well as the genotypes of all untyped individuals. Finally, a likelihood ratio is computed weighting the evidence of the generated data under two competing hypotheses.

Specifically we use the software FamLinkX (version 2.9.2) allowing us to take the haplotype structure as well as recombination landscape fully into account [3]. FamLinkX uses a model (referred to as the lambda model) to assign frequencies to all theoretical haplotypes, see Kling et al. [3] for a description. We use $\lambda = 1$ yielding high weight to observed haplotypes and low weight to expected haplotype occurrence. To contrast the additive value of analyzing a set of 12 X-STR markers, see Supplementary Table 1, with a panel of 10 autosomal STR markers as well as a standard panel of 23 autosomal STR markers, we performed 1000 simulation for a number of relevant relationships (listed in the first column of Supplementary Table 5) and each marker set. We use an in-house Norwegian case-work frequency database [35] to generate data for autosomal markers and the 631 Norwegian males included in this study to generate a Norwegian X-chromosomal haplotype database. Mutations, subpopulation structure and other complicating factors were disregards (i.e. set to zero).

3. Results

We divide the results as follows; first we present some descriptive statistics commonly accompanying a forensic population paper. Secondly we lay out two different approaches to analyze linkage disequilibrium (LD). Thirdly we explore other topics related to population genetics and finally we present results from our simulations. As noted in previously, all analyzes and plotting are conducted using standard base functions and scripts in R (version 4.1.1), unless otherwise stated. In each analyses we clearly state whether we have used the primary data set (Table 1) or the primary and secondary data set (Supplementary Table 2).

3.1. Descriptive statistics

Table 2 summarizes the results from the analyzes of the 631 haplotypes in the Norwegian population sample in this study. The table divides the markers of the Investigator Argus X-12 QS kit into four clusters, also referred to as linkage groups (see Supplementary Table 1 for details on the clusters). It should be noted that the clusters are not freely recombining, which will affect likelihood ratio calculations, but not generally the metrics presented in the table. Henceforth we will use the concept theoretical haplotypes to denote all possible haplotypes that can be generated based on a set of markers and their alleles respectively, listed in Table 2 for the triad of markers in each cluster. This should not be confused with observed haplotypes which constitutes the observed haplotypes and singletons which refers to observed haplotypes only seen once in the data set. We note that there is high power of exclusion (> 95 %) for paternal half sisters and paternal grandmother/granddaughter (when the mother is typed) for all clusters and in particular the first cluster which has the greatest diversity on all metrics in the table.

Table 2

Population genetic data for the Norwegian population sample ($N = 631$), with relevant forensic efficiency parameters for four triads of X-chromosomal STR markers in the Argus X-12 QS kit. MEC = Mean exclusion chance, PE=Power of Exclusion. PDF/PDM/MEC Desmarais Trio and Duo are calculated using formulas in Desmarais et al. [11] whereas PE (Paternal half sisters) and PE (Paternal grandmother/granddaughter) are calculated using formulas in Pinto et al. [12].

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Theoretical haplotypes	5376	1560	1197	3456
Observed haplotypes	362	204	193	275
Most common haplotype	7	20	24	13
Singletons	218	91	91	142
Power of Discrimination in Females (PDF)	0.999969	0.999791	0.99972	0.99991
Power of Discrimination in Males (PDM)	0.996074	0.989685	0.988032	0.993256
MEC Desmarais Trio	0.996059	0.989582	0.987896	0.993212
MEC Desmarais Duo	0.99217	0.979544	0.976307	0.986584
PE (Paternal half sisters)	0.984411	0.959637	0.953365	0.973399
PE (Paternal grandmother/granddaughter) ^a	0.992139	0.979344	0.976042	0.986496

^a Data available for the mother of the granddaughter.

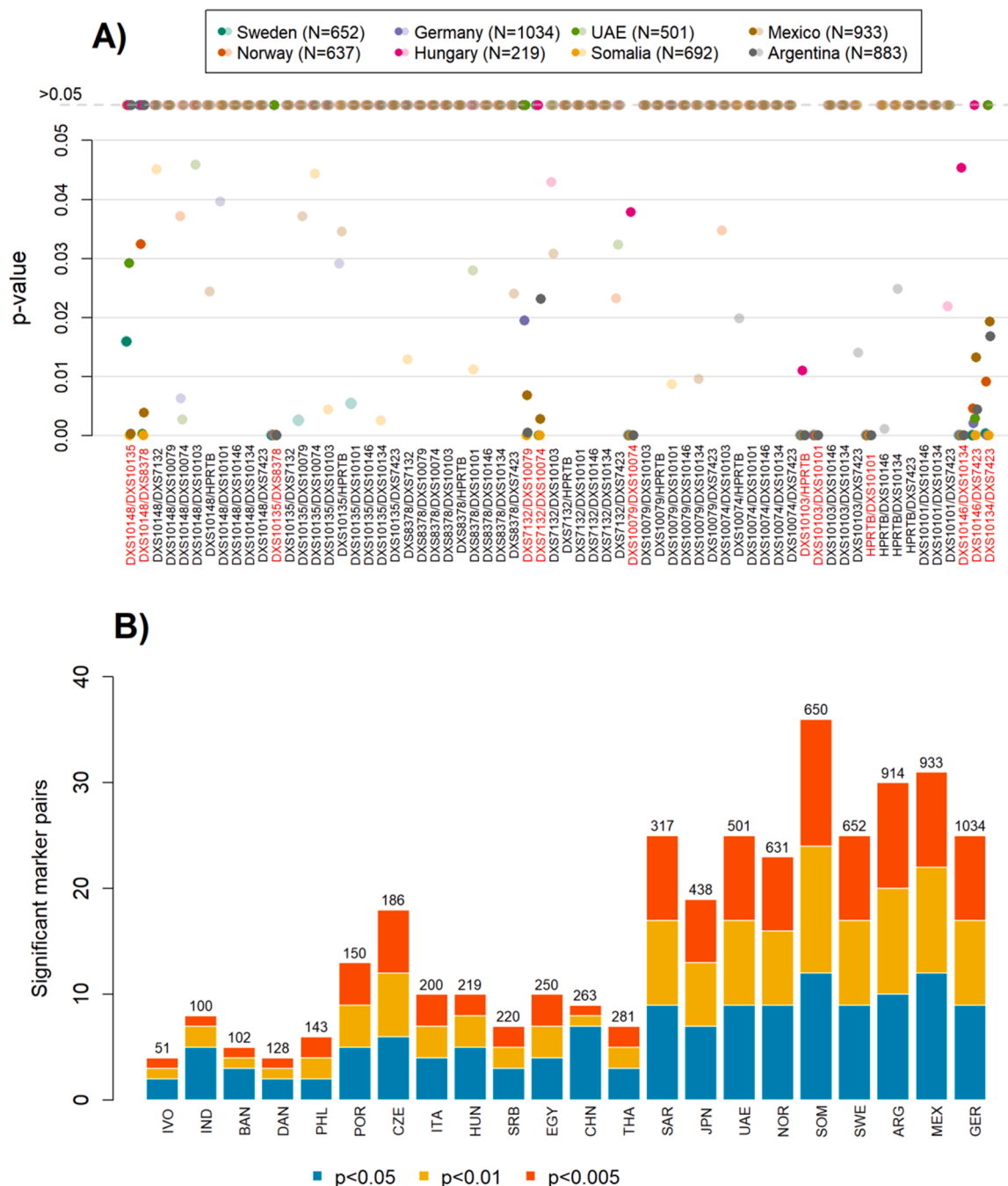


Fig. 2. Summary of p-values estimated for different populations. A) P-value versus pair of markers in consideration where a marker pair located in the same cluster is highlighted in red. For each marker pair, the p-value is calculated based on the observed versus expected haplotype frequency distributions. 8 different populations are considered and the faded dots represents marker combinations in different clusters. B) P-values within each cluster (total 12 comparisons) and number of significant loci using three different levels of significance (see legend). The number above each stacked bar represents the total number of observations. Each individual bar can have a maximal height of 12. SOM = Somalia, PHI = Philippines, CHN = China, JPN = Japan, BAN = Bangladesh, IND = India, THA = Thailand, MEX = Mexico, EGY = Egypt, SAR = Sardinia, HUN = Hungary, SER = Serbia, GER = Germany, ITA = Italy, SWE = Sweden, NOR = Norway (current study), UAE = United Arab Emirates, CZE = Czech Republic, POR = Portugal.

3.2. Linkage disequilibrium

The deviation from linkage equilibrium (LD) for alleles at different markers are commonly established through a statistical test. Due to the nature of the great number of haplotype combinations and the comparatively few number of observations in our study, we use an exact test (Fisher) to compute p-values for each combination of two markers. In particular we use the function *fisher.test* in R with observed haplotype frequencies as input. To illustrate for the complete set of markers in the

Investigator Argus X-12 QS kit, we analyze and visualize this using the primary set of populations listed in Table 1 where complete haplotype data is available. Fig. 2A illustrates the results where marker combinations located in the same cluster are highlighted in red (Supplementary Table 3 lists the p-values for each population and marker combination). As is expected there are significant p-values even for markers located in different cluster, either due to chance or too few samples.

Since LD between clusters is unlikely to persist due to large physical and genetic distance, we focus next on the individual linkage groups (see

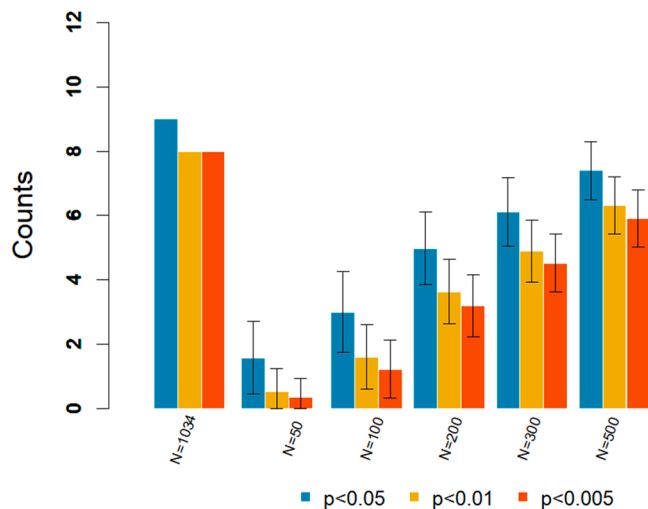


Fig. 3. Number of marker combinations with significant p-values (out of total 12). The leftmost juxtaposed bars are based on reference data using the total 1034 set of observed haplotypes. The remaining bars are generated through iteratively drawing N haplotypes from the total of 1034. The drawing is repeated 1000 times, the bars also displays the 95 % confidence region from the 1000 iterations.

Supplementary Table 1) and performed extensive p-value computations for the joint set of our primary data set and our secondary data set, listed in Supplementary Table 2. The results are illustrated in Fig. 2B where we have counted the number of significant comparisons (out of a total 12) for three different significance levels (0.05, 0.01 and 0.005). The figure provides an indication that few haplotypes (i.e. observations) entails undetected LD patterns (Spearman's correlation coefficient = 0.8255048, $p = 4e-06$ at 0.005 significance level). The result is not surprising since the statistical power is directly related to the sample size. To further explore this, we conducted a small study where the large German haplotype database ($N = 1034$) was used as a starting point [1]. We subsequently randomly draw N haplotypes from this database where $N = 50, 100, 200, 300, 500$ and computed the p-values for each draw. We repeated the sampling procedure 1000 times, and the averages as well as confidence intervals are summarized in Fig. 3 corroborating the observations in Fig. 2. All tests illustrate what is expected, small sample sizes will likely not have sufficient power to detect LD.

Since statistical tests using p-values have drawbacks (e.g. lack of power), we suggest an alternative exploration of the degree of LD in contrast to stating if deviation from equilibrium is significant. We use entropy to measure the degree of information in a combination of haplotypes in contrast to markers considered separately. Specifically we use the normalized entropy (NED), described by Siegert et al. for Y STR markers [9], which assumes a value in the range 0–1. Similar to the analyzes conducted for p-values above, we first focus on the primary data set with complete haplotype data (see Table 1). Fig. 4A displays the NED computed for all pairs of markers in the Investigator Argus X-12 QS kit and Supplementary Fig. 1 illustrates the same data but with sorting based on the number of theoretical haplotypes in each marker combination. The figures strongly suggests a correlation between NED values and number of theoretical haplotypes for each marker combination (Spearman's correlation coefficient > 0.8 and $p < 1e-14$ for all populations). Fig. 4B illustrates the NED values for the secondary data set, listed in Supplementary Table 2, where sample sizes range from 51 to 1034, further suggesting inflated NED values for small population samples and a correlation between sample size and NED values (Spearman's correlation coefficient = -0.809 and $p = 8.846e-06$ at $NED > 0.1$). We explore these phenomena by generating data in LE and compute the NED value for each generated data point. We use four levels

on the number of observed haplotypes (49, 168, 576 and 768) and 20 levels of database size ranging from 50 to 2000. That is we draw n number of haplotypes based on LE data, generated from the expected haplotype frequencies, and where n ranges from 50 to 2000. Fig. 5A illustrates the results and indicate that high to moderate NED (> 0.1) is detected even for samples from a pool of haplotypes in LE when the database size is small and the number of observed haplotypes is large. The figure suggests that a database size of say at least 1000 might be needed to provide useful data for estimation of NED, for this particular set of markers. We note that 1000 is very crude estimate and is affected by the population in consideration.

In order to further explore the impact of the number of haplotypes on the NED score, we singled out data for the marker combinations DXS10103/HPRTB (with a total of 56 observed haplotypes in our Norwegian population sample) and DXS10148/DXS10135 (with a total of 768 observed haplotypes in our Norwegian population sample). We further compute the NED for all populations encompassing our secondary data set (Supplementary Table 2). The results, illustrated in Fig. 5B further corroborate our previous findings where greater number of possible haplotypes and lower sample size potentially inflates the NED score. Finally, to exhaust the investigations, we singled out data for the markers DXS10103/HPRTB (total of 56 observed haplotypes) located closely in the same cluster with expected LD and DXS8378/DXS7423 (total of 42 observed haplotypes) located far apart on the chromosome with expected LE for our primary data set (Table 1). The results, visualized in Fig. 5C, show that with larger sample sizes (e.g. > 500) and few haplotypes (e.g. < 100) NED provides a good measure of the degree of association between multi-allelic STR markers.

3.3. Measuring population similarities

3.3.1. Genetic distance

We compute the R_{st} (using Arlequin [38]) which contrasts with the commonly used F_{st} that does not account for the repeat like structure of STR markers. The raw output is a distance-like matrix (per cluster of three markers in the Argus X-12 QS kit), with each row and column corresponding to a population (Supplementary Table 4). We conduct analyzes separately for each cluster of the Argus X-12 QS kit and create an average based on the four clusters.

Fig. 6B illustrates a dendrogram based on the averaged R_{st} matrix. The dendrogram uses an average linkage approach to create a phylogenetic-like tree where the distance (arbitrary units) on the y-axis depicts the genetic distance between populations (branches). In short, the iterative approach clusters populations with the shortest distance and then creates an average distance between the formed cluster and the remaining, un-clustered, populations. Fig. 6A alternatively displays a multidimensional scaling (MDS) plot using the same R_{st} data.

We further combined geographically proximal population samples into larger ones. We used $R_{st} = 0.005$ as an upper threshold to combine populations, listed in Table 3. For the East Asian (EAS) populations, we deliberately did not include the Mexican (MEX) or Argentinian (ARG) populations, even though the R_{st} values might suggest such connection. We argue that the geographical location of MEX and ARG compared to EAS is too distant to justify a merge. Furthermore, the indicated South American populations did exceed the R_{st} threshold perhaps suggesting that the included populations are from different ancestries, perhaps European versus indigenous origins. The R_{st} data also suggest that NEU and SEU could further be combined into a European super-population, although we decided to treat them separately in this study.

We used the combined data sets and performed a new set of R_{st} analyses (see Supplementary Fig. 2) as well as summarizing significant p-values (see Supplementary Fig. 3) and NED values (see Supplementary Fig. 4). In particular, Supplementary Fig. 4 provides an interesting observation, the baseline, i.e. the background degree of association of alleles (NED), is deflated and approaches zero even for combinations of markers with larger number of theoretical haplotypes while still

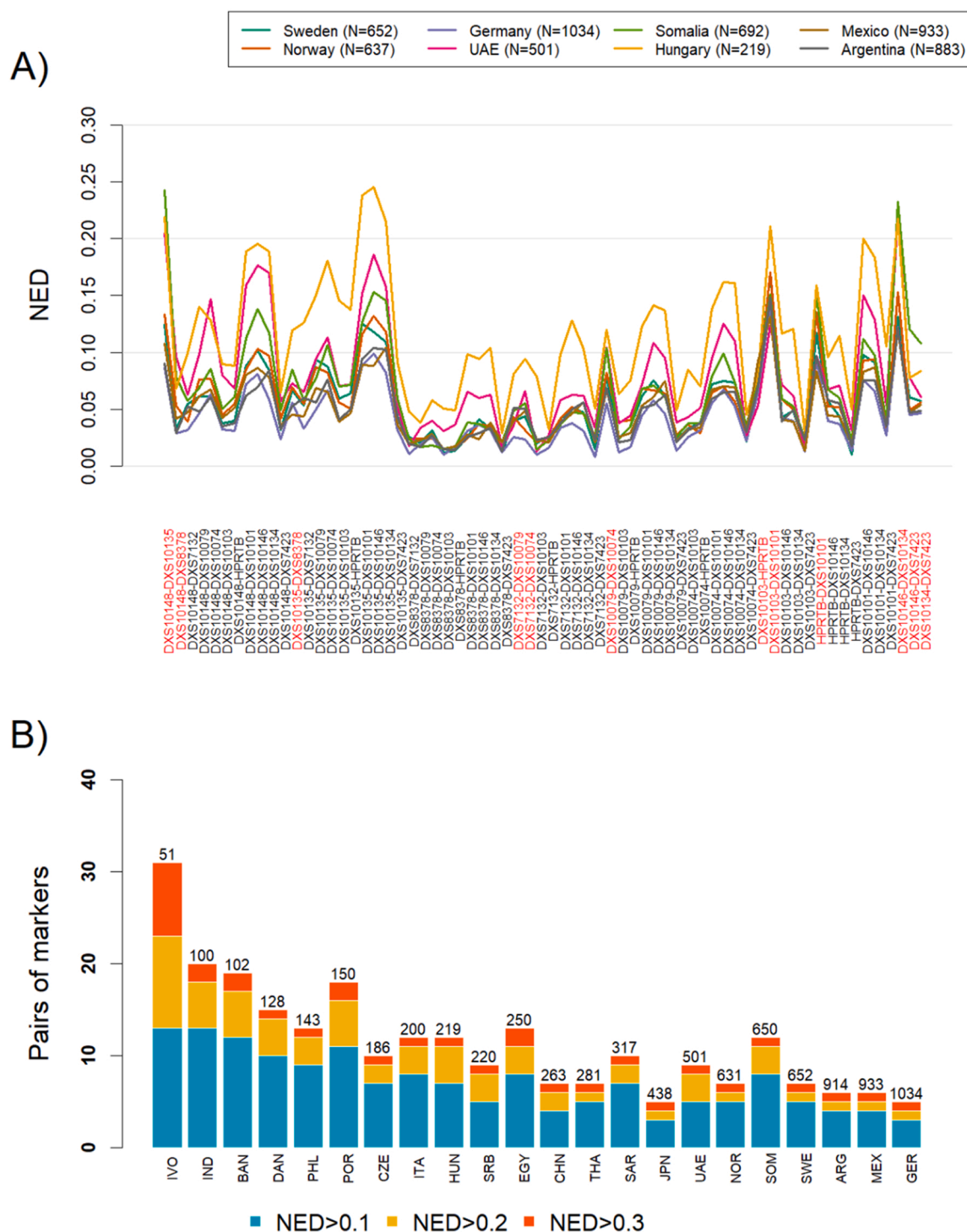


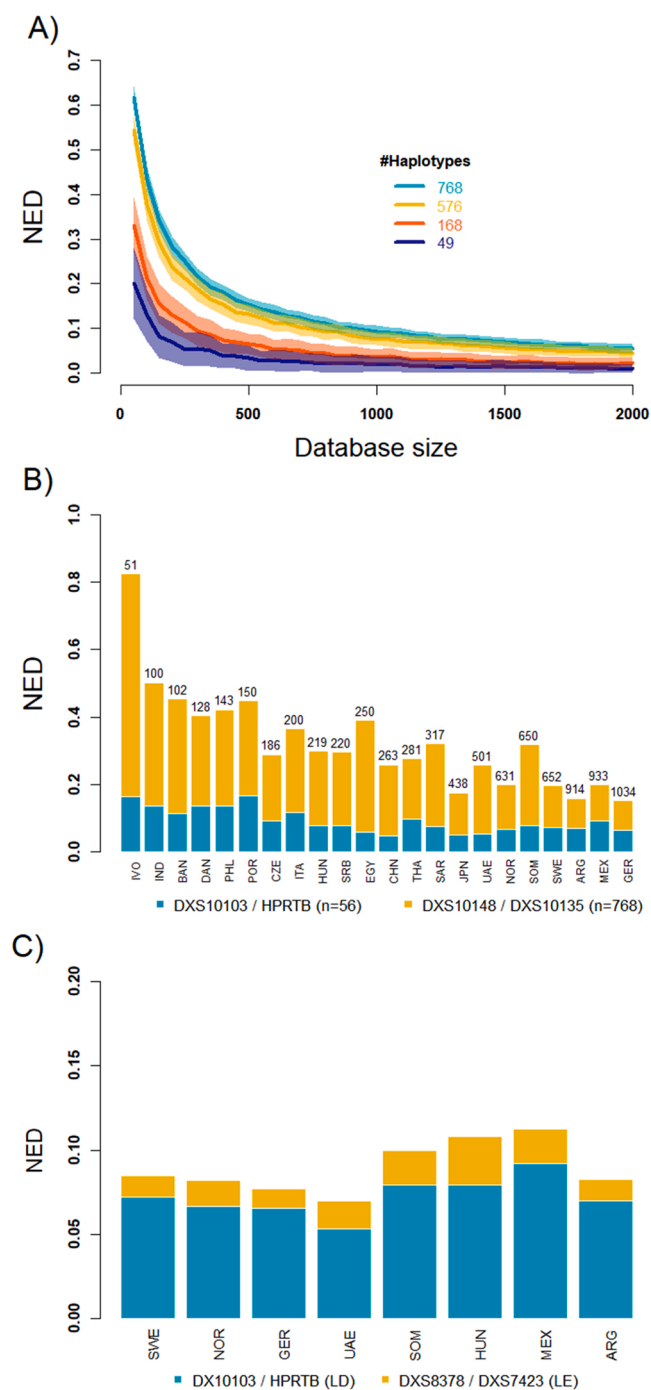
Fig. 4. Summary of analyses of linkage disequilibrium using normalized entropy difference (NED). A) A selection of eight population with complete haplotype data. The x-axis indicates the marker pair and markers in the same cluster are highlighted in red. B) Summary of a linkage disequilibrium analysis on haplotype data from the populations in this study. The results are further divided into three different classes, low degree ($NED > 0.1$), moderate degree ($NED > 0.2$) and high degree ($NED > 0.3$). The bars represents the counts (i.e. the number of marker combinations out of a total 36) of marker combinations in each class. If a marker combination is classified as high degree it cannot not be also classified as low or moderate degree. SOM = Somali, PHI = Philippines, CHN = China, JPN = Japan, BAN = Bangladesh, IND = India, THA = Thailand, MEX = Mexico, EGY = Egypt, SAR = Sardinia, HUN = Hungary, SER = Serbia, GER = Germany, ITA = Italy, SWE = Sweden, NOR = Norway (current study), UAE = United Arab Emirates, CZE = Czech Republic, POR = Portugal, IVO = Ivory Coast.

showing peaks at the combination of markers located in the same clusters, in particular for the joint Northern European population (NEU).

3.3.2. Population structure

The results are plotted in Fig. 7A, based on individual decomposition

into three assumed overarching populations and Fig. 7B, based on average population decomposition into 2–5 assumed overarching populations. The results from the STRUCTURE analysis, which as discussed above should be interpreted with cautions, are striking as they provide a surprising difference based on the mere twelve genetic markers



(caption on next column)

contained in the Argus X-12 QS multiplex, perhaps explained by the relatively large sample size and the high polymorphism of the markers. In particular, with $K = 3$, $K = 4$ and $K = 5$ in Fig. 7, a division into a European, a South American and a African subpopulation is observed, with the UAE being a mixture of European and African populations. Precaution is needed when evaluating these figures, in particular when evaluating the best fit in terms of number of subpopulations, K (i.e. the highest likelihood), since the model does not fully account for the haplotype structure in the data, which applies in particular to the likelihood.

Fig. 5. Summary of analyses of linkage disequilibrium using normalized entropy difference (NED). A) Degree of linkage disequilibrium measured using normalized entropy (NED) versus the size of the sampled database. Data is repeatedly sampled from a Norwegian population sample 1000 times for each size (0–2000) assuming linkage equilibrium (LE) between alleles at pairs of loci. The number of theoretical haplotypes, derived from observed alleles in the Norwegian population sample, are indicated in the legend for each combination of loci included in the simulations. Data is plotted for the mean as well as 95 % confidence bands. B) The degree of linkage disequilibrium, measured using normalized entropy difference (NED) is displayed on the y-axis for a number of population listed on the x-axis. Data is singled out for two combination of markers both pairs located in the same clusters (linkage groups), DXS10103/HPRTB with few theoretical haplotypes ($n = 56$ in the Norwegian population sample) whereas DXS10148/DXS10135 has a much greater number of possible haplotypes ($n = 768$ in the Norwegian population sample). The number above each bar represents the size of each population sample. SOM = Somali, PHI = Philippines, CHN = China, JPN = Japan, BAN = Bangladesh, IND = India, THA = Thailand, MEX = Mexico, EGY = Egypt, SAR = Sardinia, HUN = Hungary, SER = Serbia, GER = Germany, ITA = Italy, SWE = Sweden, NOR = Norway (current study), UAE = United Arab Emirates, CZE = Czech Republic, POR = Portugal, IVO = Ivory Coast. C) The degree of linkage disequilibrium, measured using normalized entropy difference (NED) is displayed on the y-axis for a number of population listed on the x-axis. Data is singled out for two combination of markers, DXS10103/HPRTB located closely in the same cluster and previously demonstrated to have significant association of alleles (LD), whereas DXS8378/DXS7423 are located far apart with no evidence of an association between alleles (LE). Both pairs of markers have a small number of theoretical haplotypes (56 and 49 respectively in our Norwegian population sample).

3.4. Simulations

The results from the simulations are listed in [Supplementary Table 5](#) (X-chromosomal data only) and illustrated in [Fig. 8](#) (comparison of autosomal and X-chromosomal information) where we have used a sliding threshold (log LR) on the x-axis and computed the fractions of simulations exceeding the threshold for each marker set respectively. We contrast the information content of 23 standard autosomal STRs with a battery of 10 additional autosomal STRs and the 12 X STRs described in this study ([Supplementary Table 1](#)). We see that 12 X-STRs are generally more informative compared to 23 aSTR for cases like paternal half sisters, maternal half brothers/siblings, paternal grandmother/granddaughter, maternal grandmother/grandson, maternal grandfather/grandchild, maternal aunt/nephew and paternal uncle/niece. We further note that the information content of the X-STRs always exceeds that of the additional 10 aSTRs, suggesting that for the listed relationships, X-chromosomal analysis should be the first choice.

To explore the evidential impact of combining databases (based on population similarity measures), see [Section 3.3.1](#), we used an approach whereby data were simulated 1000 times in the expanded Northern European database (combining Norway, Sweden, Germany and Czech Republic population data) for paternal half sisters and LRs computed in the Norwegian database described in this study. In addition, and for comparison, we also computed the LR in the Northern European database. As pointed out in Kling [\[43\]](#), due to phase uncertainty X-chromosomal haplotype data can provide unintuitive results, which in particular can occur for small size haplotype databases. In op sic, Kling particularly exemplifies a case of paternal half-sisters, where unobserved haplotypes can result in an exclusionary conclusion even though a complete haplotype is shared between the sisters. Kling further discusses that increasing the size of the database could provide a better estimate of the true haplotype frequencies. For comparison, we enumerated all theoretical haplotypes based on the Norwegian haplotype data, and matched these to the observations in the Norwegian haplotype data and the joint Northern European data sets. The results are listed in [Table 4](#). As expected, the number of haplotypes with at least one observation increased for all clusters when combining the

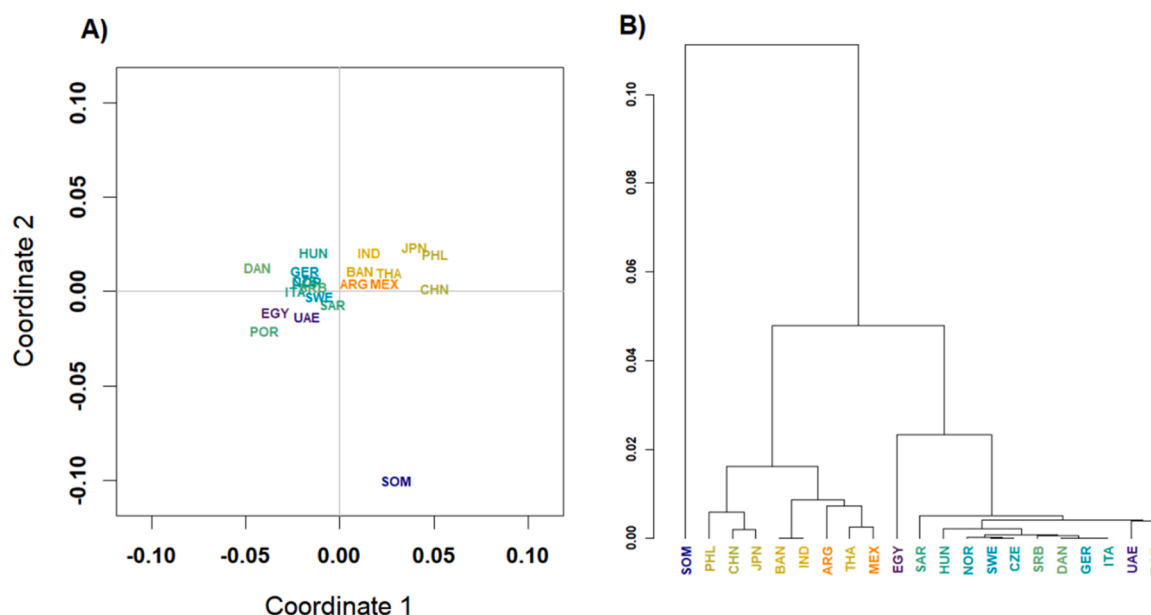


Fig. 6. Illustration of distance between population as measured by the R_{st} parameter. A) Multidimensional scaling (MDS). B) Dendrogram – created using hierarchical clustering (average linkage). SOM = Somali, PHI = Philippines, CHN = China, JPN = Japan, BAN = Bangladesh, IND = India, THA = Thailand, MEX = Mexico, EGY = Egypt, SAR = Sardinia, HUN = Hungary, SER = Serbia, GER = Germany, ITA = Italy, SWE = Sweden, NOR = Norway (current study), UAE = United Arab Emirates, CZE = Czech Republic, POR = Portugal.

Table 3

List of super-population, their combined size and the included populations.

Super-population	Combined size	Populations included
Northern Europe (NEU)	2621	Norway, Sweden [26], Germany [29], Czech Republic [41]
Southern Europe (SEU)	887	Italy [42], Portugal [22], Sardinia [21], Serbia [23]
East Asia (EAS)	1126	Japan [19], Thailand [20], China [14], Philippines [17]

population samples. In fact, for clusters 1 and 4, the numbers are almost doubled.

We use decision rates to illustrate when the two reference databases yields identical or opposing conclusions in a case of simulated paternal half-sisters. That is, for a given decision threshold, we summarize the fraction of simulated cases where the results from each of the databases exceeds or falls below the threshold. So for instance, in one simulation the LR could be 10 when we use the Norwegian population database and 10,000 when we use the combined NEU database, which can further be used with a decision threshold of 100 stating that the LR in the former database is not sufficient to reach a conclusion. Fig. 9 illustrates the results from the simulations and suggests that 1) the LR using the NEU database is more likely to exceed the threshold compared to when using the Norwegian database (Fig. 9A), for instance at $LR = 1000$ roughly 90 % and 50 % of the cases will exceed the threshold in the two databases respectively, 2) the LR using the Norwegian database is below the inverse of the threshold to a much greater extent than when using the NEU database, for instance at $LR = 1/1000$, there is a roughly 15% chance that a case will yield an exclusion when using the former database.

4. Discussion

We provide an in-depth evaluation of the markers in the commercial Investigator Argus X-12 QS multiplex using 1) different approaches to explore linkage disequilibrium, 2) approaches to illustrate the population structure and differentiation 3) simulations as a versatile tool to assess the information content in different kinship scenarios.

We contrast commonly used statistical test for determining whether two genetic markers and their alleles are in linkage equilibrium (LE)

with a measure of the degree of linkage disequilibrium (LD) in multi-allelic loci, namely entropy [9,36,37]. The former yields a p-value that is sensitive to the size of the sample included in the test, see for instance Fig. 2B. In small population samples, the power to detect LD for the highly polymorphic STR loci in study is low [10]. Similarly, for entropy, a correlation is seen both to the numbers of theoretical haplotypes between a combination of markers, but also with the size of the population sample, see Figs. 3 and 4, suggesting that large databases are needed to fully capture LD structure between the highly polymorphic STR markers in our study. Our results further indicate that a sample size of at least 1000 is necessary to provide reliable measures of entropy as well as sufficient power to detect LD using the traditional statistical test. For databases with a sample size below 1000 and for some of the highly polymorphic marker combinations we stress that careful interpretation is needed.

We use extensive data from published population samples with genotype from the Investigator Argus X-12 QS and perform joint analyses with our Norwegian population sample with the ultimate aim to yield new databases consisting of not country specific, but databases spanning across borders forming so-called super-populations. This concept is not novel and has been studied, for instance, in the 1000 Genomes project [44] and elsewhere. Since most population studies rely on dense data from SNP microarrays or genome wide sequencing, it is hard to translate and compare the exact results from those studies. We study R_{st} metrics produced in the software Arlequin [38] which give a measure of the genetic proximity of population and further use this information to combine populations into such larger super-populations. We illustrate that combining haplotype data from genetically similar population provides a lower degree of spurious linkage disequilibrium while still being able to accurately detect true patterns of association of alleles (see

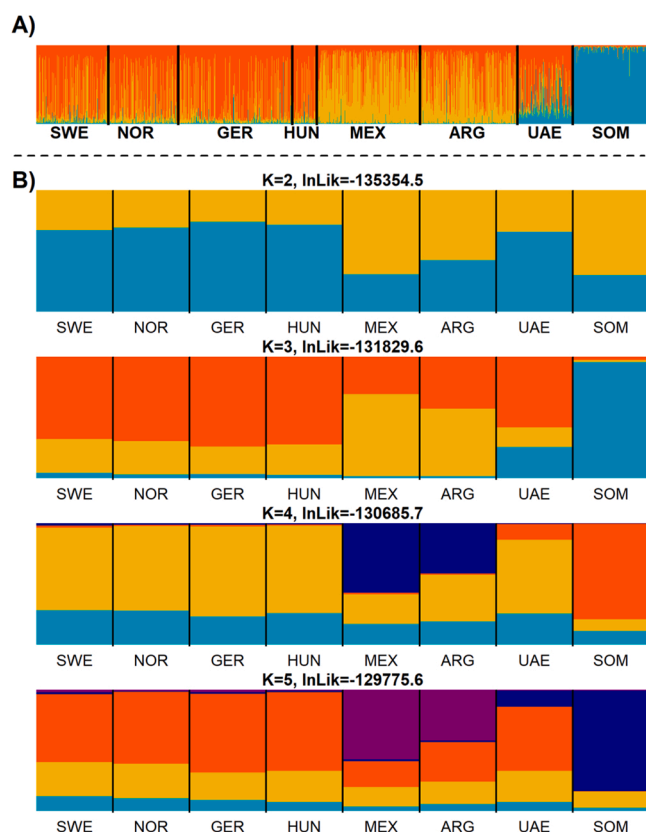


Fig. 7. STRUCTURE analysis using 12 X STR markers for different sets of population groups. A) Each bar represents an individual and its decomposition into different subpopulations. The analyzes has been conducted assuming three overarching populations ($K = 3$). B) Each bar represents instead a population its decomposition into different subpopulations. The analyzes has been conducted assuming 2–5 overarching populations and with the logarithm of the likelihood of the data given each K included. SWE = Sweden, NOR = Norway, GER = Germany, HUN = Hungary, MEX = Mexico, ARG = Argentina, UAE = United Arab Emirates, SOM = Somalia.

Supplementary Figs. 2–4). It should also be noted that additional population genetic analyses should be performed to fully validate the suggested approach to merge different datasets. An example of such analyzes is the testing of Hardy Weinberg disequilibrium and to study potential Wahlund effects [45]. This involves the comparison of expected and observed diplotype frequencies (i.e. pair of haplotypes in an individual). Since only male were typed in our study we have not been able to address this issue.

As previously demonstrated, X-chromosomal markers are undoubtedly a valuable addition of information in certain kinship scenarios [5,6,8]. In particular, some cases can only be distinguished using markers located on the X-chromosome, explored in Pinto et al. [5]. For instance, paternal half-sisters is a particularly interesting case where a complete

X-chromosomal haplotype must be shared, excluding mutations and errors in the data, which is also visualized in Table 2. In yet other cases it may provide null information due the loss of the X-chromosome in father-to-son meiosis, for instance paternal half-brothers. The use of a second battery of markers can expand the forensic kinship testing horizon to more challenging relationships, i.e. beyond standard paternity cases. However, an assessment should be made whether X-markers or autosomal markers are better suited if cost/benefit is a concern. In this paper, we illustrate the use of simulations as an excellent tool to explore the information content of genetic markers, particularly to better contrast expanded panels of different marker types. Whereas traditional forensic parameters, exemplified for our Norwegian population sample in Table 2, informs us of the general performance and exclusion power of the marker panel, simulations provide an additional depth of information. Indeed, simulations give the power under all hypotheses in considerations (inclusion and exclusion probabilities). Fig. 8 provides insight on some selected cases where using X-chromosomal markers compares or even outranks the information gained from running an expanded battery of autosomal markers. A caveat is that simulations are specific to the population used, in our study the Norwegian sample, to generate the data and some care should be taken to extrapolate our results to other populations.

To illustrate the power of a larger haplotype database, we performed simulations for a case with two paternal half sisters where data was generated in a large database ($N = 2631$) and LR's subsequently calculated in both the larger and a smaller ($N = 631$) database. As illustrated in Table 4, there is a substantial increase in the number of haplotypes with at least a single observation which in turn suggests that haplotypes not observed in the smaller database will be sampled from the larger database in the simulations. The results (visualized in Fig. 9) illustrate that 1) a larger database will result in a higher degree of cases with conclusion (Fig. 9A) and 2) a small database have an elevated chance of falsely excluding two paternal half sisters as related (Fig. 9B). The results are explained by the fact that unobserved haplotypes will have a low frequency (rare) in the smaller database and thus yield a low likelihood for relatedness where at least one individuals (the shared father) must have the haplotype. In contrast, the likelihood for the alternative hypotheses (unrelated) may be higher since the hypothesis does not require an obligate haplotype and in turn provides other possible haplotype combinations, potentially also observed in the smaller database. We speculate that this may be relevant also for other kinship problems, but to what extent remains to be tested.

Our study has a few limitations, first the studied populations samples (both our own Norwegian data set, but also data included from other studies) have varying degree of certainty in their ancestries. Our samples have self-declared Norwegian ancestry stated submission form, but with no further information on from where their ancestors are. Secondly, we performed a crude sweep of the literature searching for published data to include in our study. We acknowledge that this search was not exhaustive, Gomes et al. [8] list a more comprehensive summary in Supplementary Table 2.

In conclusion, this study provides the fundamentals for exploring detailed population differences for a limited number of X-chromosomal

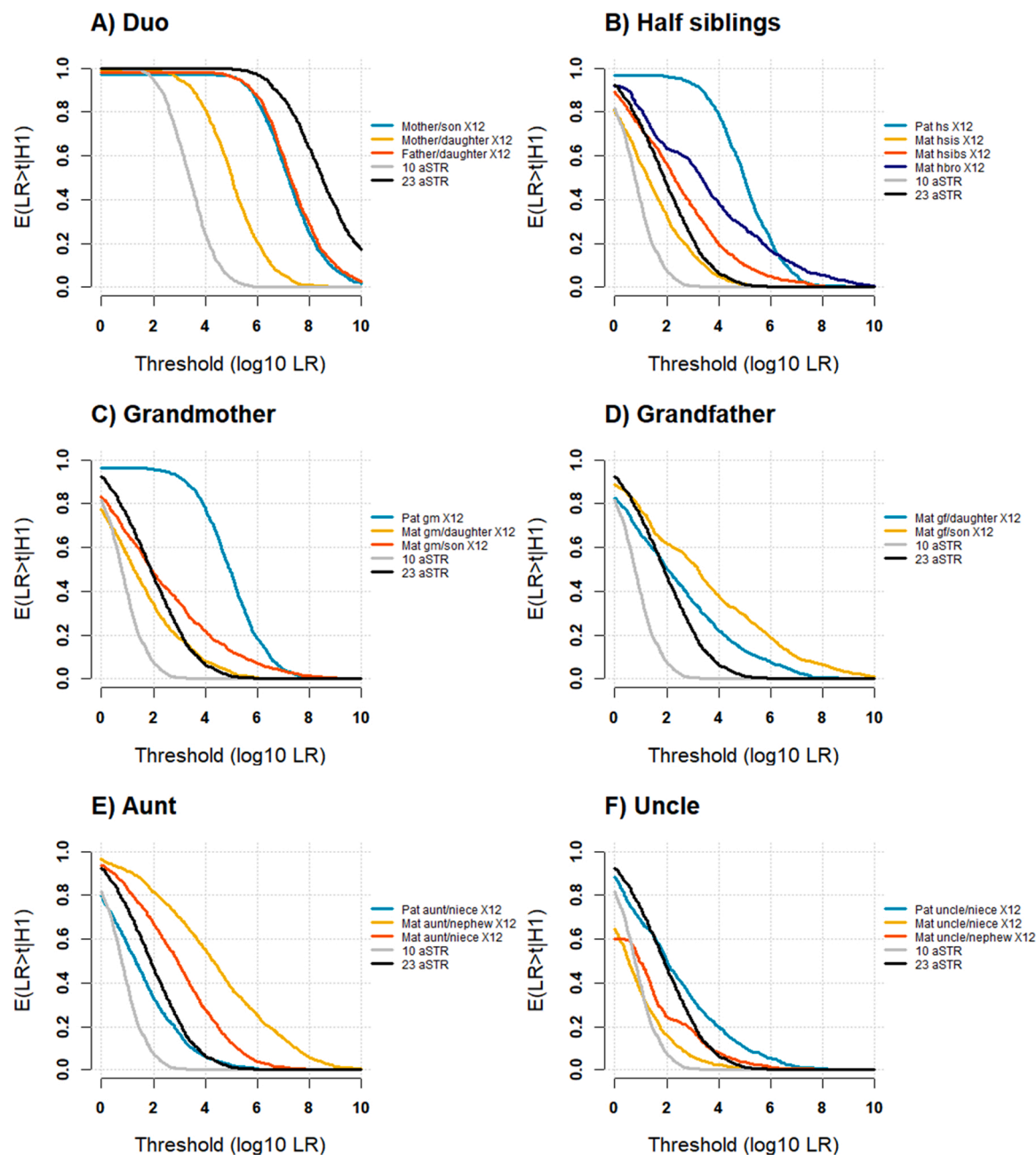


Fig. 8. Exceedance plots for a selection of pedigrees relevant for X-chromosomal marker data. The y-axes illustrate the exceedance probability, defined as the probability that a case will exceed a given LR threshold (x-axes). Data is simulated for the 12 markers in the Argus X-12 QS kit (denoted X12 in the legends) and using a standard battery of 23 aSTRs in addition to a set of 10 extra aSTRs. Abbreviations used: Pat = Paternal, Mat = Maternal, gm = grandmother, gf = grandfather, hs = half-sisters.

Table 4

Haplotypes in each cluster of the Argus X-12 QS kit based on the observed haplotypes in the Norwegian population (NOR) sample in this study. Unique haplotypes, matched to the theoretical haplotypes, are listed for the NOR population as well as the joint Northern European (NEU) population.

Cluster	Theoretical haplotypes	Observed haplotypes in NOR	Observed haplotypes in NEU	Increase in the number of haplotypes with at least one observation (%)
1	5376	362	707	95.3
2	1560	204	352	72.5
3	1197	193	314	62.7
4	3456	275	533	93.8

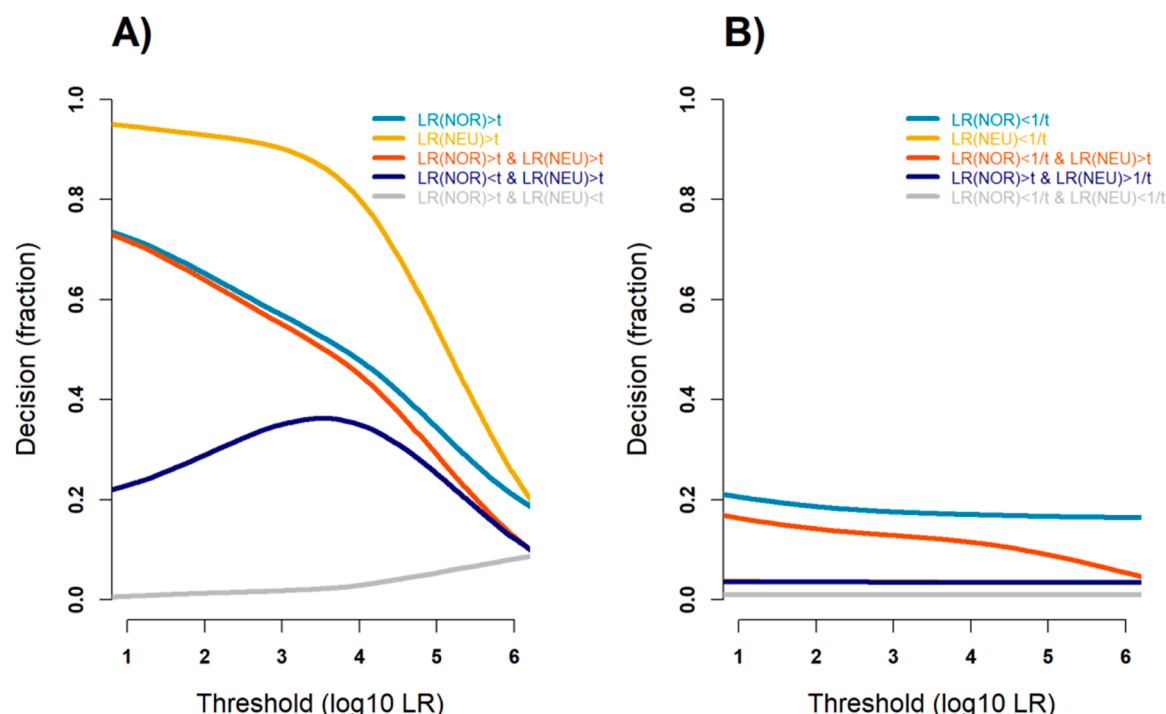


Fig. 9. Results from 100,000 simulations of paternal half-sisters. Data is simulated using a combined Northern European (NEU) database. The LR is computed assuming unrelated as the alternative hypothesis. A) Fraction of cases where the LR exceeds a threshold (t) using either the Norwegian population database (NOR) or the combined Northern European (NEU) population database in LR calculations. B) Fraction of cases where the LR is lower than the exclusion threshold, here defined as the inverse of the threshold to make an inclusion (1/t). For illustrative purposes the x-axes use log10 (t).

STR markers and lay out the way to combine haplotype data in turn providing a more comprehensive database for statistical calculations.

Acknowledgements

We are grateful for the comments from two anonymous reviewers, which greatly improved the final manuscript. The work was in part funded by a grant from Linköping University's Strategic Research Area in Forensic Science.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.fsigen.2022.102745](https://doi.org/10.1016/j.fsigen.2022.102745).

References

- [1] 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 (6) (2012) 778–784.
- [2] R. Szibor, X-chromosomal markers: past, present and future, *Forensic Sci. Int.: Genet.* 1 (2) (2007) 93–99.
- [3] 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.
- [4] 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 (5) (2010) 506–511.
- [5] 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 (1) (2011) 27–32.
- [6] N. Pinto, P.V. Silva, A. Amorim, A general method to assess the utility of the X-chromosomal markers in kinship testing, *Forensic Sci. Int.: Genet.* 6 (2) (2012) 198–207.
- [7] A.O. Tillmar, D. Kling, J.M. Butler, W. Parson, M. Prinz, P.M. Schneider, T. Egeland, L. Gusmão, DNA Commission of the International Society for Forensic Genetics (ISFG): guidelines on the use of X-STRs in kinship analysis, *Forensic Sci. Int.: Genet.* 29 (2017) 269–275.
- [8] I. Gomes, N. Pinto, S. Antão-Sousa, V. Gomes, L. Gusmão, A. Amorim, Twenty years later: a comprehensive review of the X chromosome use in forensic genetics, *Front. Genet.* 11 (2020) 926.
- [9] S. Siegert, L. Roewer, M. Nothnagel, Shannon's equivocation for forensic Y-STR marker selection, *Forensic Sci. Int.: Genet.* 16 (2015) 216–225.
- [10] J.F. Ferragut, N. Pinto, A. Amorim, A. Picornell, Improving publication quality and the importance of post publication peer review: the illustrating example of X chromosome analysis and calculation of forensic parameters, *Forensic Sci. Int.: Genet.* 38 (2019) e5–e7.
- [11] D. Desmarais, Y. Zhong, R. Chakraborty, C. Perreault, L. Busque, Development of a highly polymorphic STR marker for identity testing purposes at the human androgen receptor gene (HUMARA), *J. Forensic Sci.* 43 (5) (1998) 1046–1049.
- [12] N. Pinto, L. Gusmão, 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 (2) (2013) 290–295.
- [13] J. Ge, B. Budowle, R. Chakraborty, Choosing relatives for DNA identification of missing persons, *J. Forensic Sci.* 56 (Suppl. 1) (2011) S23–S28.
- [14] F. Guo, Population genetic data for 12 X-STR loci in the Northern Han Chinese and StatsX package as tools for population statistics on X-STR, *Forensic Sci. Int.: Genet.* 26 (2017) e1–e8.
- [15] A. Zidkova, P. Capek, A. Horinek, P. Coufalova, Investigator® Argus X-12 study on the population of Czech Republic: comparison of linked and unlinked X-STRs for kinship analysis, *Electrophoresis* 35 (14) (2014) 1989–1992.
- [16] S. Elakkary, S. Hoffmeister-Ullrich, 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] J.M. Salvador, D. Apaga, F.C. Delfin, G.C. Calacal, S.E. Dennis, M. De Ungria, Filipino DNA variation at 12 X-chromosome short tandem repeat markers, *Forensic Sci. Int.: Genet.* 36 (2018) 8.
- [18] C. Bini, L.N. Riccardi, S. Ceccardi, F. Carano, S. Sarno, D. Luiselli, S. Pelotti, Expanding X-chromosomal forensic haplotype frequencies database: Italian population data of four linkage groups, *Forensic Sci. Int.: Genet.* 15 (2015) 127–130.
- [19] S. Uchigasaki, J. Tie, D. Takahashi, Genetic analysis of twelve X-chromosomal STRs in Japanese and Chinese populations, *Mol. Biol. Rep.* 40 (4) (2013) 3193–3196.
- [20] K. Vongpaisarnsin, A. Boonlert, K. Rasmeepaisarn, P. Dangkao, Genetic variation study of 12 X chromosome STR in central Thailand population, *Int. J. Leg. Med.* 130 (6) (2016) 1497–1499.
- [21] C. Robino, D. Lacerenza, S. Aneli, C. Di Gaetano, G. Matullo, R. Robledo, C. Calò, Allele and haplotype diversity of 12 X-STRs in Sardinia, *Forensic Sci. Int.: Genet.* 33 (2018) e1–e3.

- [22] L. Cainé, S. Costa, M.F. Pinheiro, Population data of 12 X-STR loci in a North of Portugal sample, *Int. J. Leg. Med.* 127 (2013) 1–2.
- [23] I. Veselinović, D. Vapa, M. Djan, N. Veličković, T. Veljović, G. Petrić, Genetic analysis of 12 X-STR loci in the Serbian population from Vojvodina Province, *Int. J. Leg. Med.* 132 (2) (2018) 405–408.
- [24] P. Shrivastava, T. Jain, U. Gupta, V.B. Trivedi, Genetic polymorphism study on 12 X STR loci of investigator Argus X STR kit in Bhil tribal population of Madhya Pradesh, India, *Leg. Med.* 17 (3) (2015) 214–217.
- [25] A. Sufian, M.I. Hosen, K. Fatema, T. Hossain, M.M. Hasan, A.K. Mazumder, S. Akhteruzzaman, Genetic diversity study on 12 X-STR loci of investigator® Argus X STR kit in Bangladeshi population, *Int. J. Leg. Med.* 131 (4) (2017) 963–965.
- [26] A.O. Tillmar, Population genetic analysis of 12 X-STRs in Swedish population, *Forensic Sci. Int.: Genet.* 6 (2) (2012) e80–e81.
- [27] M.A. Almarrí, R.A. Lootah, Allelic and haplotype diversity of 12 X-STRs in the United Arab Emirates, *Forensic Sci. Int.: Genet.* 33 (2018) e4–e6.
- [28] I. Cortés-Trujillo, F. Zuñiga-Chiquette, B. Ramos-González, M.L. Chávez-Briones, K. L. Islas-González, D.A. Betancourt-Guerra, R. Peralta-Coria, G. Martínez-Cortés, H. Rangel-Villalobos, Allele and haplotype frequencies of 12 X-STRs in Mexican population, *Forensic Sci. Int.: Genet.* 38 (2018) 11.
- [29] J. Edelmann, et al., X-chromosomal haplotype frequencies of four linkage groups using the Investigator Argus X-12 Kit, *Forensic Sci. Int.: Genet.* 6 (1) (2012) e24–e34.
- [30] 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.
- [31] G. Horvath, et al., A genetic study of 12 X-STR loci in the Hungarian population, *Forensic Sci. Int.: Genet.* 6 (1) (2012) e46–e47.
- [32] M.G. García, C.I. Catanesi, G.A. Penacino, L. Gusmão, N. Pinto, X-chromosome data for 12 STRs: towards an Argentinian database of forensic haplotype frequencies, *Forensic Sci. Int.: Genet.* 41 (2019) 8.
- [33] D. Kling, A.O. Tillmar, T. Egeland, Familias 3-Extensions and new functionality, *Forensic Sci. Int.: Genet.* 13 (2014) 121–127.
- [34] T. Egeland, P.F. Mostad, B. Mevåg, M. Stenersen, Beyond traditional paternity and identification cases. Selecting the most probable pedigree, *Forensic Sci. Int.* 110 (1) (2000) 47–59.
- [35] B.M. Dupuy, M. Stenersen, D. Kling, Frequency data for 35 autosomal STR markers in a Norwegian, an East African, an East Asian and Middle Asian population and simulation of adequate database size, *Forensic Sci. Int.: Genet. Suppl. Ser.* 4 (1) (2013) e378–e379.
- [36] M. Nothnagel, R. Fürst, K. Rohde, Entropy as a measure for linkage disequilibrium over multilocus haplotype blocks, *Hum. Hered.* 54 (4) (2002) 186–198.
- [37] M. Nothnagel, K. Rohde, The effect of single-nucleotide polymorphism marker selection on patterns of haplotype blocks and haplotype frequency estimates, *Am. J. Hum. Genet.* 77 (6) (2005) 988–998.
- [38] 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 (3) (2010) 564–567.
- [39] J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data, *Genetics* 155 (2) (2000) 945–959.
- [40] J.W. MacCluer, J.L. VandeBerg, B. Read, O.A. Ryder, Pedigree analysis by computer simulation, *Zoo. Biol.* 5 (2) (1986) 147–160.
- [41] A. Zídková, P. Capek, A. Horinek, P. Coufalová, Investigator® Argus X-12 study on the Population of Czech Republic: comparison of linked and unlinked X-STRs for kinship analysis, *Electrophoresis* 35 (2014) 1989–1992.
- [42] S. Inturri, S. Menegon, A. Amoroso, C. Torre, C. Robino, Linkage and linkage disequilibrium analysis of X-STRs in Italian families, *Forensic Sci. Int.: Genet.* 5 (2) (2011) 152–154.
- [43] D. Kling, Curiosities of X chromosomal markers and haplotypes, *Int. J. Leg. Med.* 132 (2) (2018) 361–371.
- [44] P.H. Sudmant, T. Rausch, E.J. Gardner, R.E. Handsaker, A. Abyzov, J. Huddleston, Y. Zhang, K. Ye, G. Jun, M.H. Fritz, M.K. Konkel, A. Malhotra, A.M. Stütz, X. Shi, F. P. Casale, J. Chen, F. Hormozdiari, G. Dayama, K. Chen, M. Malig, M. Chaisson, K. Walter, S. Meiers, S. Kashin, E. Garrison, A. Auton, H. Lam, X.J. Mu, C. Alkan, D. Antaki, T. Bae, E. Cerveira, P. Chines, Z. Chong, L. Clarke, E. Dal, L. Ding, S. Emery, X. Fan, M. Gujral, F. Kahveci, J.M. Kidd, Y. Kong, E.W. Lammeijer, S. McCarthy, P. Flicek, R.A. Gibbs, G. Marth, C.E. Mason, A. Menelaou, D. M. Muzny, B.J. Nelson, A. Noor, N.F. Parrish, M. Pendleton, A. Quitadamo, B. Raeder, E.E. Schadt, M. Romanovitch, A. Schlattl, R. Sebra, A.A. Shabalín, A. Untergasser, J.A. Walker, M. Wang, F. Yu, C. Zhang, J. Zhang, X. Zheng-Bradley, W. Zhou, T. Zichner, J. Sebat, M.A. Batzer, S.A. McCarroll, C. Genomes Project, R. E. Mills, M.B. Gerstein, A. Bashir, O. Stegle, S.E. Devine, C. Lee, E.E. Eichler, J. O. Korbel, An integrated map of structural variation in 2504 human genomes, *Nature* 526 (7571) (2015) 75–81.
- [45] J.S. Buckleton, C.M. Triggs, S.J. Walsh, *Forensic DNA Evidence Interpretation*, CRC press, 2004.