Int J Legal Med (2010) 124:205–215
DOI 10.1007/s00414-009-0413-0

ORIGINAL ARTICLE

Potentials and limits of pairwise kinship analysis using autosomal short tandem repeat loci

Michael Nothnagel · Jörg Schmidtke · Michael Krawczak

Received: 26 August 2009 / Accepted: 18 December 2009 / Published online: 10 February 2010

Abstract At least in principle, most instances of complex kinship testing can be reduced to pairwise kinship cases involving two critical family members that either link or separate presumed sub-branches of a family. In the European population, the 34 short tandem repeats (STRs) currently used in forensic genetics are sufficiently powerful to allow assessment of disputed first and second but not lower degrees of pairwise blood relatedness. We provide estimates of the means and variances of marker-specific log-likelihood ratios, using large-sample approximation and assuming different scenarios of pairwise kinship analysis. These estimates allow power calculations to be performed for any combination of the available STRs. Since some of the markers considered are physically linked, chromosome-wide likelihood calculations in kinship cases other than parent–child duos (and trios) have to take the reduced rates of meiotic inter-marker recombination into account. We show by simulation that this requirement may be ignored when discriminating distant hypotheses about kinship, but that linkage may play an important role in the biostatistical analysis of more intricate cases.

Electronic supplementary material The online version of this article (doi:10.1007/s00414-009-0413-0) contains supplementary material, which is available to authorized users.

M. Nothnagel · M. Krawczak (✉)
Institute of Medical Informatics and Statistics,
Christian-Albrechts University,
Campus Kiel, House 31, Arnold-Heller-Str. 3,
24105 Kiel, Germany
e-mail: krawczak@medinfo.uni-kiel.de

J. Schmidtke
Institute of Human Genetics, Hannover Medical School,
Hannover, Germany

Keywords Kinship · Linkage · Microsatellite · STR · Likelihood

Introduction

Most practically relevant cases of disputed kinship either concern the relationship between two individuals or can be confined to the clarification of a bilateral relationship. For instance, in classical trio cases of disputed paternity, the identity of the biological mother is rarely questioned so that the trio reduces to a duo comprising the child and the alleged father. The mother is only involved because knowledge of her genotype greatly increases the power to identify the true biological father of the child and because the maternal genotype may serve as an additional means of quality control. Similarly, the vast majority of complex kinship cases are concerned with a single disputed parent–child relationship that either connects or disconnects two groups of presumed blood relatives. Theoretically, most kinship cases could thus be confined to the analysis of two critically important individuals, one from each branch of the family, if the markers used for kinship testing were sufficiently powerful. The major advantage of trimming kinship cases down to pairwise analyses would be that, although a higher number of markers may be necessary to achieve sufficient power to solve the case, the number of individuals that need to be involved may be drastically reduced.

To our knowledge, autosomal genetic markers currently used for routine kinship analysis in humans comprise the 15 short tandem repeat (STR) loci of the PowerPlex® kit (Promega Inc., Madison WI, USA), the 12 STRs of the Humantype Chimera® kit (Biotype AG,

Dresden, Germany) plus eight additional STRs of comparable heterozygosity, bringing the total number to 34 (Table 1; note that D18S51 is included in both commercially available kits). In addition, a number of sex chromosomal STRs are available for routine use, including the Mentype® Argus X-8 set (Biotype AG, Dresden, Germany) and the Y-STRs constituting the 'core' and 'extended' haplotypes (see www.yhrd.org). Owing to their formal genetic particularities, however, sex-linked markers will not be considered here.

All markers listed in Table 1 have proven sufficiently informative for forensic case work and for solving trio cases of disputed paternity. At the same time, however, we are frequently asked to carry our pairwise kinship analyses, thereby repeatedly confronting us with the question of which types of disputed blood relatedness may be resolvable with a particular set of markers. Here, we evaluated this approach in four ways. First, we calculated for each marker the mean and variance of the log-likelihood ratio under different scenarios of disputed kinship. These parameters were then used to assess the overall probability of successfully solving the respective type of cases, using different thresholds for the log-likelihood ratio. Since many of the practically relevant markers are physically linked, the marker-specific likelihoods cannot simply be multiplied to obtain the overall likelihood in a given case. We therefore also explored for which purposes and under which circumstances linkage

Table 1 Autosomal STR markers under study

Marker	Chromosome	Physical map location (Mb)	Number of alleles[a]	Heterozygosity
F13B	1q31-q32	195.287	6	0.715
TPOX[b]	2p25	1.461	5	0.632
D2S1360[c]	2p24	17.267	14	0.837
D2S1338	2q35-q36	210.734	11	0.879
D3S1358[b]	3p21	45.627	8	0.792
D3S1744[c]	3q24	144.469	9	0.814
D4S2366[c]	4p16	6.418	7	0.787
FGA[b]	4q28	151.250	17	0.855
D5S2500[c]	5q12	55.655	10	0.807
D5S818[b]	5q23	118.299	9	0.702
CSF1PO[b]	5q33–q35	144.604	7	0.725
F13A1	6p23–p24	6.178	10	0.748
SE33[c] (ACTBP2)	6q15	89.043	35	0.948
D6S474[c]	6q22	110.454	6	0.780
D7S820[b]	7p12	78.398	9	0.809
D7S1517[c]	7q31	117.861	13	0.870
LPL	8q22	18.352	6	0.694
D8S1132[c]	8q23	102.651	12	0.862
D8S1179[b]	8q24	121.231	10	0.812
D10S2325[c]	10p14	12.706	11	0.870
TH (TH01)[b]	11p15	1.983	7	0.787
VWA[b]	12p13	5.945	9	0.778
D12S391[c]	12p13	12.215	15	0.882
PLA2A	12q24	119.249	7	0.718
D13S317[b]	13q21	63.429	7	0.776
CYP19 (CYAR04)	15q21	28.350	8	0.725
Penta E[b]	15q26	95.115	16	0.884
D16S539[b]	16q24	70.690	8	0.779
D18S51[b,c]	18q21	57.647	13	0.877
D19S253	19p13	15.296	9	0.803
D19S433	19q21	26.925	10	0.740
D21S11[b]	21q21	5.942	14	0.835
D21S2055[c]	21q22	26.661	21	0.877
Penta D[b]	21q22	43.853	12	0.829

Marker allele frequencies were obtained from (http://www.uni-duesseldorf.de/WWW/MedFak/Serology/) [1, 2]

[a] Number of alleles with a population frequency >0.1%, reported in the respective source

[b] PowerPlex® kit (Promega Inc, Madison WI, USA)

[c] Humantype Chimera® kit (Biotype AG, Dresden, Germany)

has to be taken into account in the respective likelihood calculations. Finally, for scenarios where the pairwise approach was found to have little power, we investigated the power gain to be expected from the inclusion of additional relatives into the analysis.

Materials and methods

Single-marker likelihood calculations for pairwise blood relationships

If two separate copies ('alleles') of a genetic marker are indistinguishable with respect to their marker-defining characteristics (e.g. the STR repeat number), they are usually referred to as being 'identical by state' (IBS). If, in addition, the two alleles originated from a single ancestral copy in the not too distant past, they are called 'identical by descent' (IBD). For a given marker, any pair of individuals therefore shares either none, one or two alleles IBS and IBD, respectively.

In a population at Hardy–Weinberg equilibrium (HWE), the likelihood of a certain blood relationship between two individuals, given their genotypes at a single genetic marker, is a simple function of the population allele frequencies of that marker and the IBD probabilities pertaining to the blood relationship in question, labelled r_0, r_1 and r_2 for sharing zero, one and two alleles IBD, respectively. Indeed, any pair of genotypes $a_{11}a_{12}$ and $a_{21}a_{22}$ can be encoded such that their IBS status falls into one of the four categories listed in Table 2, depending upon whether the genotypes share two (category I), one (categories IIa and IIb), or zero alleles IBS (category III). Categories IIa and IIb need to be distinguished because, in cases where exactly one allele is shared IBS, the likelihood of a presumed blood relationship depends upon whether genotype $a_{11}a_{12}$ is homozygous (category IIa) or heterozygous (category IIb). Table 2 lists coefficients b_0, b_1 and b_2 that allow the likelihood of a blood relationship, specified by r_0, r_1 and r_2, to be calculated as

$$L(r_0, r_1, r_2 | a_{11}a_{12}, a_{12}a_{22}) = \mathrm{f}(a_{11}a_{12}) \cdot [r_0 \cdot b_0 + r_1 \cdot b_1 + r_2 \cdot b_2] \quad (1)$$

for given genotypes $a_{11}a_{12}$ and $a_{21}a_{22}$. Here, $\mathrm{f}(a_{11}a_{12})$ denotes the population frequency of genotype $a_{11}a_{12}$, which can be calculated from known allele frequency estimates (http://www.uni-duesseldorf.de/WWW/MedFak/Serology/) [1, 2] because the population of interest is assumed to be at HWE anyway.

Multiple-marker power calculations by means of large-sample approximation

Rational decision making between two hypotheses H_0 and H_A about the blood relationship of two individuals is best based upon the likelihood ratio, $LR = L(H_A)/L(H_0)$, or its logarithm (logLR) [3]. In the following, we will take all logarithms to base 10. The probability π_A (π_0) of successfully solving a kinship case is then given by the probability under H_A (H_0), respectively, that the logLR exceeds a certain threshold T (in the right direction). Assuming independence, i.e. unlinked markers, the overall

Table 2 Likelihood of pairwise blood relationship in a population at Hardy–Weinberg equilibrium

Category	Identity by state (IBS)	Coefficient		
		b_0	b_1	b_2
I	$\{a_{11},a_{12}\}=\{a_{21},a_{22}\}$ e.g. AA–AA or AB–AB	$\mathrm{f}(a_{21}a_{22})$	$[\mathrm{f}(a_{21})+\mathrm{f}(a_{22})]/2$	1
IIa	$\{a_{11},a_{12}\}\cap\{a_{21},a_{22}\}=\{a_{11}\}$ $\{a_{11},a_{12}\}\Delta\{a_{21},a_{22}\}=\{a_{22}\}$ e.g. AA–AB	$\mathrm{f}(a_{21}a_{22})$	$\mathrm{f}(a_{22})$	0
IIb	$\{a_{11},a_{12}\}\cap\{a_{21},a_{22}\}=\{a_{11}\}$ $\{a_{11},a_{12}\}\Delta\{a_{21},a_{22}\}=\{a_{12},a_{22}\}$ e.g. AB–AC	$\mathrm{f}(a_{21}a_{22})$	$\mathrm{f}(a_{22})/2$	0
III	$\{a_{11},a_{12}\}\cap\{a_{21},a_{22}\}=\varnothing$ e.g. AB–CD	$\mathrm{f}(a_{21}a_{22})$	0	0

$M\Delta N$: set difference between M and N, defined as $(M \cap N^C) \cup (M^C \cap N)$. $M\Delta N$ contains all elements of M that do not belong to N and vice versa. $f(x)$: population frequency of genotype (or allele) x. r_0 (r_1, r_2): probability with which two individuals share 0 (1, 2) alleles identical by descent (IBD). In the case of full sibs, for example, $r_0=r_2=1/4$ and $r_1=1/2$. For two individuals with genotypes $a_{11}a_{12}$ and $a_{21}a_{22}$ at a single marker, the likelihood of a given blood relationship, specified by IBD probabilities r_0, r_1 and r_2 and given genotypes $a_{11}a_{12}$ and $a_{21}a_{22}$, equals $L(r_0, r_1, r_2 | a_{11}a_{12}, a_{21}a_{22}) = \mathrm{f}(a_{11}a_{12}) \cdot [r_0 \cdot b_0 + r_1 \cdot b_1 + r_2 \cdot b_2]$.

logLR equals the sum of the marker-specific logLRs so that the sought-for probability π_A equals

$$\pi_A = P(\log LR > T) = P\left(\sum_{i=1}^{n} \log LR_i > T\right). \quad (2)$$

Here, n denotes the number of markers used, and T is a large positive number (e.g. +3). A similar formula applies to π_0, using negative values of T. Let $\mu_{i,*}$ and $\sigma^2_{i,*}$ be the mean and variance under H_*, respectively, of the $\log LR_i$ (where * stands for either 'A' or '0'). Then, according to the Central Limit Theorem of probability theory [4], the sum of the $\log LR_i$ follows a normal distribution with mean and variance

$$\mu_* = \sum_{i=1}^{n} \mu_{i,*} \text{ and } \sigma^2_* = \sum_{i=1}^{n} \sigma^2_{i,*}, \quad (3)$$

respectively, if n is sufficiently large. If $\Phi(x)$ denotes the distribution function of the standard normal distribution, i.e. $\Phi(x)=P(X\leq x)$ for a random variable with standard normal distribution, this means that π_A and π_0 can be approximated by

$$\pi_A = P_A\left(\sum_{i=1}^{n} \log LR_i > T\right) = 1 - \Phi\left(\frac{T-\mu_A}{\sigma_A}\right) \quad (4a)$$

for large positive values of T and by

$$\pi_0 = P_0\left(\sum_{i=1}^{n} \log LR_i < T\right) = \Phi\left(\frac{T-\mu_0}{\sigma_0}\right) \quad (4b)$$

for correspondingly small negative values of T.

It must be emphasised, however, that the Central Limit Theorem requires the mean and variance of all $\log LR_i$ to be finite, which may not always be the case. If the possibility of mutation and genotyping error is ignored in cases of a disputed parent–child relationship, for example, any pair of genotypes that share zero alleles IBS would yield a logLR of $-\infty$, but such genotypes will occur with probability strictly larger than zero if the two individuals of interest are unrelated. Therefore, the above approach cannot be used to calculate the power to detect false parenthood in dyads, and published formulas of the paternity exclusion probability have to be used instead [5–8].

Likelihood calculations for multiple linked markers

For more than one marker, the overall likelihood equals the product of the marker-specific likelihoods, provided that the markers are both in gametic equilibrium and physically unlinked (i.e. stochastically independent). For linked markers, however, likelihoods have to be calculated for whole chromosomes taking the pairwise genetic distances between markers into account (see below). Furthermore, a single set of IBD probabilities is not generally sufficient to determine the IBD distribution along an entire chromosome. The only exception to this are relationships for which $r_i=1$ for exactly one $i\in\{0,1,2\}$ and $r_j=0$ for all $j\neq i$. For example, likelihoods multiply even for linked loci if a decision has to be made between parenthood ($r_1=1$, $r_0=r_2=0$) and non-relatedness ($r_0=1$, $r_1=r_2=0$) in classical duo (or trio) cases of disputed paternity, which is probably why the problem of linkage has been widely ignored in kinship testing in the past.

In order to assess the effects of linkage between markers upon the likelihood calculations in pairwise kinship analysis, we performed a simulation study of five practically relevant case scenarios: parent–child vs unrelated (PC-U); full sibs vs half sibs (FS-HS); full sibs vs unrelated (FS-U); half sibs vs unrelated (HS-U; which is the same as aunt/uncle–niece/nephew vs unrelated and grandparent–grandchild vs unrelated); first cousins vs unrelated (FC-U). For each scenario, genotypes of the two individuals involved were simulated 1,000 times, adopting either of the two hypotheses about their blood relationship. We also investigated the effect of including additional relatives into the analysis of two scenarios, namely those distinguishing an aunt–niece pair and a pair of first cousins, respectively, from unrelated individuals.

Unfortunately, no comprehensive linkage map of the STRs in current forensic use is available. We therefore exploited the genome-wide average relationship between physical and genetic distance, implying that 1 Mb roughly corresponds to 1 cM, to extrapolate pairwise recombination fractions between adjacent markers from known NCBI coordinates (Table 1), using Kosambi's mapping function [9]. Whilst our genotype simulations were based upon the extrapolated linkage relationships of the 34 markers in Table 1, assuming gametic equilibrium, the likelihood calculations were carried out twice, once ignoring linkage ('rough') and once taking linkage properly into account ('exact'). Genotypes were simulated with SimPed [10]. All likelihood calculations were carried out using MLINK v5.10 or FASTLINK v4.1P [11–14], while MAKEPED v2.21 [12, 13] and PEDCHECK v1.1 [15] were used for data preparation and error checking. Graphs were prepared using R v2.8.2 [16].

In order to assess the convergence properties of the large-sample approximations in Eqs. 4a and 4b, we also simulated 1,000 genotype pairs for the ten and 20 most polymorphic STRs, respectively, each time assuming one of the case scenarios described above. We then compared the empirical distribution of the rough logLR values ensuing for each scenario to the normal distribution with the appropriate mean and variance.

Results

Approximate power calculations

Table 3 lists the mean and variance of the $\log_{10}$-likelihood ratio (logLR) for various scenarios of pairwise kinship testing, assuming that either the null hypothesis H_0 or the alternative hypothesis H_A about the blood relationship in question is correct. The results indicate that, for a disputed parent–offspring relationship (PC-U), the PowerPlex® set of markers is sufficiently powerful to provide a positive proof of parenthood. Thus, if $T=3$,

$$\pi_A = 1 - \Phi\left(\frac{-3 - 5.0640}{\sqrt{1.4057}}\right) = 1 - \Phi(-1.7409)$$
$$= 1 - 0.0409 = 0.9591,$$

i.e. the likelihood ratio would exceed 1,000:1 with more than 95% probability if the presumed parent–child relationship (H_A) is correct. If $T=2$, corresponding to a likelihood ratio of 100:1, the power would equal $\pi_A=1-\Phi(-2.5843)=0.9951$ or 99.5%.

In our experience, the second most frequent type of pairwise kinship test in practise involves maternal half sibs that may (H_A) or may not (H_0) have the same biological father (scenario FS-HS). In this situation, the PowerPlex® set alone is not powerful enough to solve the case with sufficient certainty. For $T=2$, π_A is only $1-\Phi(0.5553)=0.2894$ or 28.9%. Even worse, if H_0 is correct and $T=-2$, corresponding to a likelihood ratio of 1:100 against a common biological father, then

$$\pi_0 = \Phi\left(\frac{-2 + 1.0483}{\sqrt{0.6879}}\right) = \Phi(-1.1475) = 0.1256.$$

However, if all 34 markers listed in Table 1 are being used, $\pi_A=1-\Phi(-0.6564)=0.7442$ and $\pi_0=\Phi(0.3705)=0.6445$ for the same T values. This means that a common biological father can be both demonstrated and excluded with reasonable power.

Only marginally worse power is obtained in cases where two individuals with different mothers wish to clarify whether they had the same biological father (H_A) or not (H_0). This scenario (HS-U) is also numerically equivalent to cases of a disputed grandparent–grandchild or aunt–niece relationship. Using all 34 markers, $\pi_A=0.6907$ and $\pi_0=0.6417$ for $T=2$ and $T=-2$, respectively. The PowerPlex® set alone would again be insufficient to solve such cases with reasonable power ($\pi_A=0.2143$ and $\pi_0=0.1392$ for $T=2$ and $T=-2$, respectively). Finally, as can be inferred from Table 3, full sibs can be distinguished reasonably well from unrelated individuals using the PowerPlex® set alone (scenario FS-U; $\pi_A=0.7168$ and $\pi_0=0.5947$ for $T=3$ and $T=-3$, respectively), whereas even the full marker set would not be sufficient to identify first cousins with reasonable certainty and power (scenario FC-U; $\pi_A=0.0943$ and $\pi_0=0.0279$ for $T=2$ and $T=-2$, respectively).

The mean and variance of the logLR_i for individual markers, assuming the five scenarios of kinship testing discussed above, can be found in Supplementary Table S1. These data allow summary statistics similar to those in Table 3 to be calculated for any combination of the 34 markers considered. As can be inferred from Supplementary Figs. 1 and 2, the quality of the normal approximation is far more than sufficient for the purpose of power calculations even if only the ten most polymorphic STRs were taken into consideration. With the top 20 markers, the quality of the approximation improves even further. We did not investigate any smaller marker numbers here because these would be irrelevant for practical case work anyway.

Table 3 Mean (μ) and variance (σ^2) of the $\log_{10}$-likelihood ratio (logLR) under different scenarios of disputed kinship

Scenario		Markers	μ_A	σ^2_A	μ_0	σ^2_0
PC-U	Parent–child (H_A) vs unrelated (H_0)	all	12.3204	3.1864	$-\infty$	na
		PowerPlex®	5.0640	1.4057	$-\infty$	na
FS-HS	Full sibs (H_A) vs half sibs (H_0)	all	3.2358	3.5450	−2.4648	1.5738
		PowerPlex®	1.3371	1.4249	−1.0483	0.6879
FS-U	Full sibs (H_A) vs unrelated (H_0)	all	10.2764	11.2616	−7.9773	5.2194
		PowerPlex®	4.2289	4.5915	−3.3609	2.2702
HS-U	half sibs (H_A) vs unrelated (H_0)[a]	all	2.8370	2.8244	−2.5012	1.9075
		PowerPlex®	1.1507	1.1508	−1.0277	0.8040
FC-U	First cousins (H_A) vs unrelated (H_0)	all	0.7963	0.8382	−0.6598	0.4913
		PowerPlex®	0.3171	0.3397	−0.2675	0.2021

na not applicable

[a] Same as aunt/uncle–niece/nephew vs unrelated and grandparent–grandchild vs unrelated

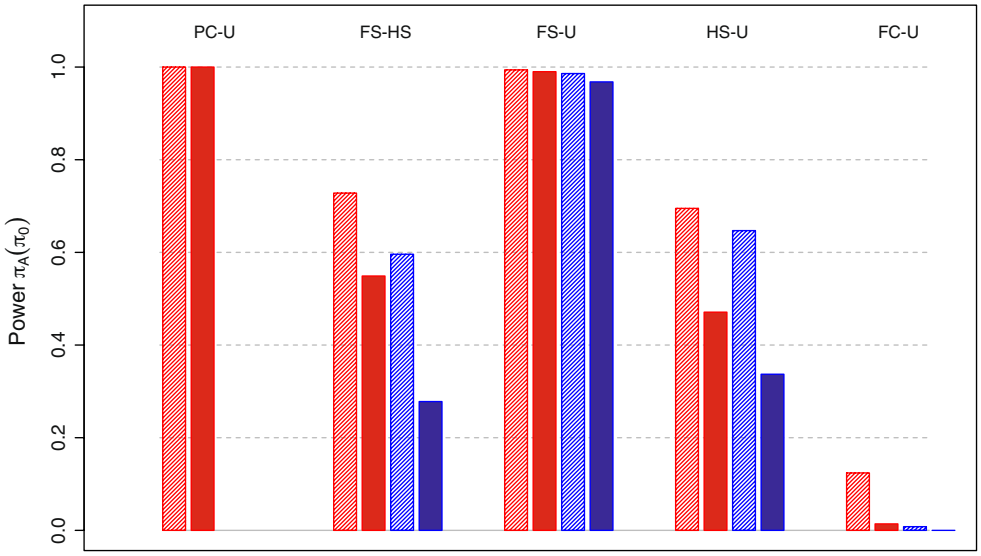

Fig. 1 Power of 34 autosomal STRs to confirm different hypotheses about pairwise kinship. *PC-U* parent–child (H_A) vs unrelated (H_0), *FS-HS* full sibs (H_A) vs half sibs (H_0), *FS-U* full sibs (H_A) vs unrelated (H_0), *HS-U* half sibs (HA) vs unrelated (H_0), *FC-U* first cousins (H_A) vs unrelated (H_A); *red bars*: genotype data simulated under H_A, *blue bars*: genotype data simulated under H_0; *hatched bars*: decision threshold T for the logLR set to either 2 or −2, *solid bars*: T=3 or T=−3, respectively. All likelihood calculations were performed taking inter-marker linkage appropriately into account (exact LR calculation). For scenario PC-U, the bar charts corresponding to H_0 are missing because Mendelian incompatibilities would occur with positive probability in such cases

Kinship inference using linked markers

Our simulation study employing exact (i.e. linkage-based) likelihood calculations generally confirmed the results of the approximate power calculations. For scenarios PC-U and FS-U, the combined power π_A of all 34 markers to prove close kinship at T=3 was found to exceed 95% (Fig. 1). The same applies to π_0, i.e. the power to exclude kinship under H_0, adopting T=−3 in scenario FS-U. It is worth noting that π_0 was practically 100% for scenario PC-U in the simulation study, i.e. a parent–child incompatibility arose for at least one marker in each of the 1,000 simulations.

Also, in agreement with the approximate results, the exact power of pairwise kinship testing turned out to be substantially lower for scenarios FS-HS and HS-U. For T=3, π_A was 54.9% for scenario FS-HS and 47.1% for HS-U; with T=2, π_A equalled 72.8% for FS-HS and 69.5% for HS-U. The power π_0 to exclude kinship if H_0 is correct was even lower: 27.8% for FS-HS and 33.7% for HS-U if T=−3 and 59.6% for FS-HS and 64.7% for HS-U if T=−2. Finally, the power to distinguish between first cousins and unrelated individuals (scenario FC-U) was <1.5% under both H_A and H_0, assuming either T=3 or T=−3, respectively, as a threshold for the exact logLR.

The logLR distributions obtained by exact and rough likelihood calculations were found to be very similar (Fig. 2). For theoretical reasons (see above), their overlap was expected to be perfect under scenario PC-U (data not shown), but not with more distant hypotheses about kinship (i.e. H_A). Nevertheless, the concordance between exact and rough logLR values was high (Fig. 3), and only some minor bias of the rough logLRs became evident in the assessment of the most disparate hypotheses about kinship, namely for scenario FS-U (Fig. 3). Here, the absolute difference between rough and exact logLR values exceeded 0.5 in 23.7% of the simulations under H_A and in 70.4% of the simulations under H_0. For practical purposes, however, the proportion of cases in which use of the rough, instead of the exact, logLR value would result in a wrong decision about kinship is more important. We therefore determined the percentage of simulations in which the logLR values fell on opposite sides of commonly agreed thresholds T for decision making, namely ±2 or ±3, which would be equivalent to the production of false-positive or false-negative evidence, respectively. While the percentage of such errors was most often below 2%, it still exceeded 5% for some scenarios (Table 4). Thus, rough instead of exact likelihood calculations may produce wrong evidence in a non-negligible proportion of cases.

Given the modest discriminatory power of pairwise kinship testing under scenarios HS-U, which is numerically equivalent to aunt–niece pairs in the absence of linkage, and FC-U, we also investigated the effect of including additional relatives into the analysis. In the first instance of a disputed aunt–niece relationship, we assumed that either the genotypes of the mother of the niece, of a second aunt or of both were also known. This additional information dramatically increased the

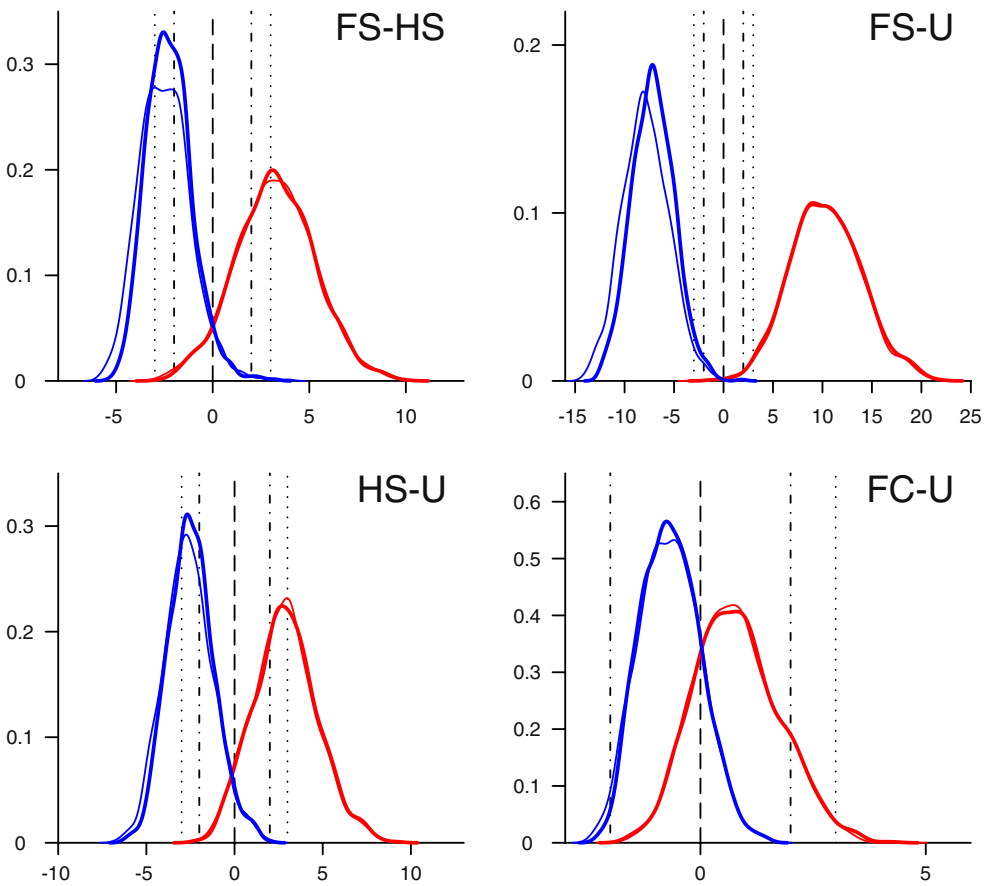

Fig. 2 Simulation-based densities of exact and rough logLR values under different kinship scenarios. *x-axis*: logLR, *solid curves*: exact logLR values, *thin curves*: rough logLR values calculated ignoring linkage, *vertical lines*: logLR thresholds $T=\pm2$ (*dash-dotted*), $T=\pm3$ (*dotted*) and $T=0$ (*dashed*); for the colour coding and the definition of kinship scenarios, see legend to Fig. 1

power to decide the case (Fig. 4a). The effects of including a second aunt or a mother were similar, with π_A increasing from 45.3% to 74.5% and 79.5%, respectively, if $T=3$. Simultaneous inclusion of both relatives raised the power to 96.1%. For scenario FC-U, we included either one or both sibling parents of the propositi into the analysis (Fig. 4b). Power π_A increased from 1.5% to 48.8% when only one parent was included and to 98.3% with both parents genotyped ($T=3$). The power to exclude the presumed relationships at $T=-3$ showed similar increases in both scenarios (Fig. 4).

Discussion

Our study has revealed that the 34 STR markers currently used in forensic practice are powerful enough to solve pairwise kinship cases if the degree of disputed blood relatedness is sufficiently high. Thus, parent–child dyads and full sibs can be distinguished firmly from pairs of unrelated individuals in more than 95% of cases, without involving additional relatives into the analysis. For distinguishing half sibs from either full sibs or unrelated pairs of individuals, the power would still be sufficient (~70%) if a less stringent criterion for decision making is employed (i.e. a likelihood ratio of 100:1 instead of 1,000:1). Wider kinship, however, cannot be reliably confirmed or excluded with the 34 markers considered.

In such instances, additional relatives may need to be involved, thereby transforming these cases into cases with higher degrees of disputed relatedness. For example, if two individuals wish to assess whether they are first cousins or unrelated, at least one of their sibling parents must be tested as well so that the case is essentially about an unclear uncle/aunt–niece/nephew relationship.

The genomic distribution of the currently used STRs implies that several markers will be located on one and the same chromosome. In principle, this means that likelihood calculations in complex kinship cases other than classical duos or trios almost always have to take the physical linkage of markers into account. It was recently reported that even groups of X-chromosomal STRs deemed to be virtually unlinked showed considerable inter-group linkage [17]. As we have observed, logLR values calculated either with ('exact') or without ('rough') the consideration of linkage are strongly correlated. However, at least in cases where half sibs are to be distinguished from full sibs or from unrelated pairs of individuals (and in equivalent cases), the discrepancy between exact and rough logLRs may be so large that it affects decision making in individual

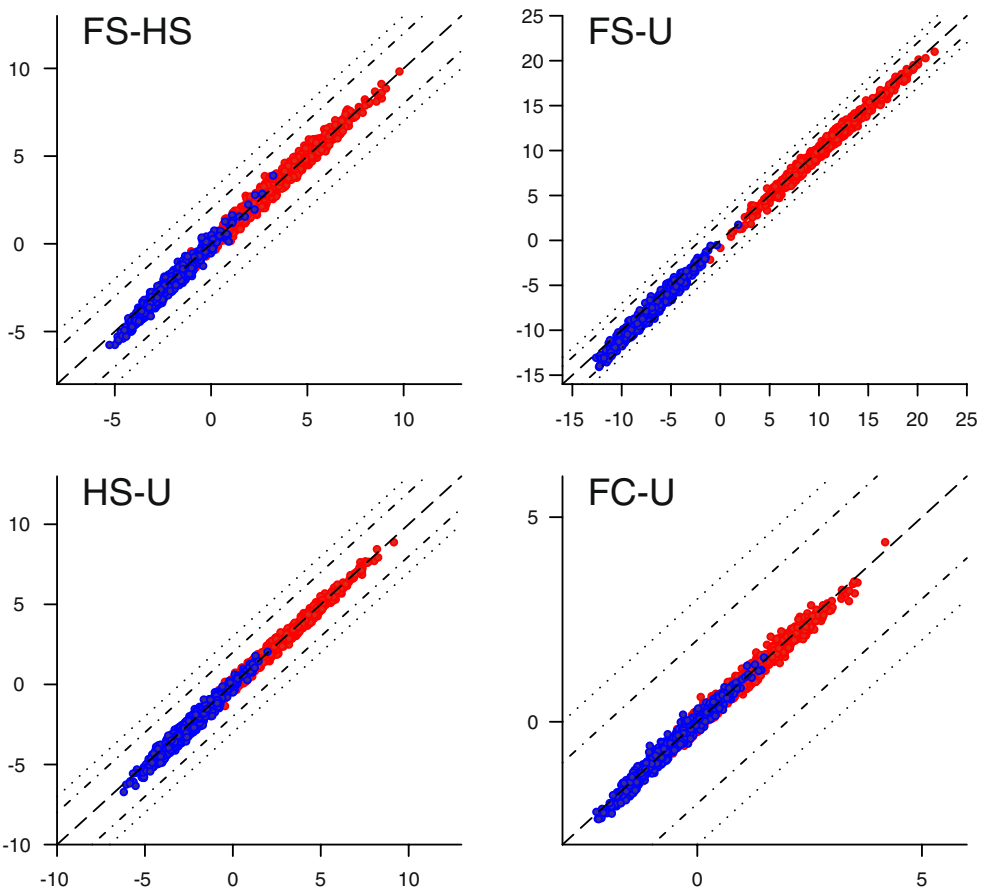

Fig. 3 Correlation between exact and rough logLR values. *x-axis*: exact logLR, *y-axis*: rough logLR, *black lines*: difference between exact and rough logLR values equal to ±2 (*dash-dotted*), ±3 (*dotted*) or zero (*dashed*), respectively; for the colour coding and the definition of kinship scenarios, see legend to Fig. 1

cases. Therefore, exact likelihood calculations appear mandatory in these cases. Anyhow, in our view, exact likelihood calculations are recommended for any type of kinship test because all propositi are entitled to the provision of as exact a biostatistical result as possible. Only in instances of disputed paternity can likelihoods continue to be multiplied in order to obtain overall likelihoods without invalidating the numerical outcome.

So far, empirically derived inter-marker genetic distances have only been published for the minority of the forensic

Table 4 Percentage of simulations of kinship scenarios with false evidence produced by rough logLR value

Scenario		True hypothesis	T=±2		T=±3	
			% false positives	% false negatives	% false positives	% false negatives
FS-HS	Full sibs (H_A) vs half sibs (H_0)	H_A	1.6	2.0	1.8	2.6
		H_0	6.2	1.6	9.2	0.6
FS-U	Full sibs (H_A) vs unrelated (H_0)	H_A	0.0	0.3	0.0	0.4
		H_0	0.6	0.1	1.0	0.0
HS-U	Half sibs (H_A) vs unrelated (H_0)[a]	H_A	1.2	3.0	1.5	3.0
		H_0	5.5	1.9	6.7	1.0
FC-U	First cousins (H_A) vs unrelated (H_0)	H_A	1.1	1.7	0.0	0.2
		H_0	1.0	0.0	0.0	0.0

A positive sign of threshold T for decision making refers to the evaluation of H_A whereas a negative sign applies to H_0. A result was deemed false positive if the rough logLR exceeded T (in the right direction) while the exact logLR did not; a false negative result was obtained when the exact logLR exceeded T while the rough logLR did not.

[a] Same as aunt/uncle–niece/nephew vs unrelated and grandparent–grandchild vs unrelated

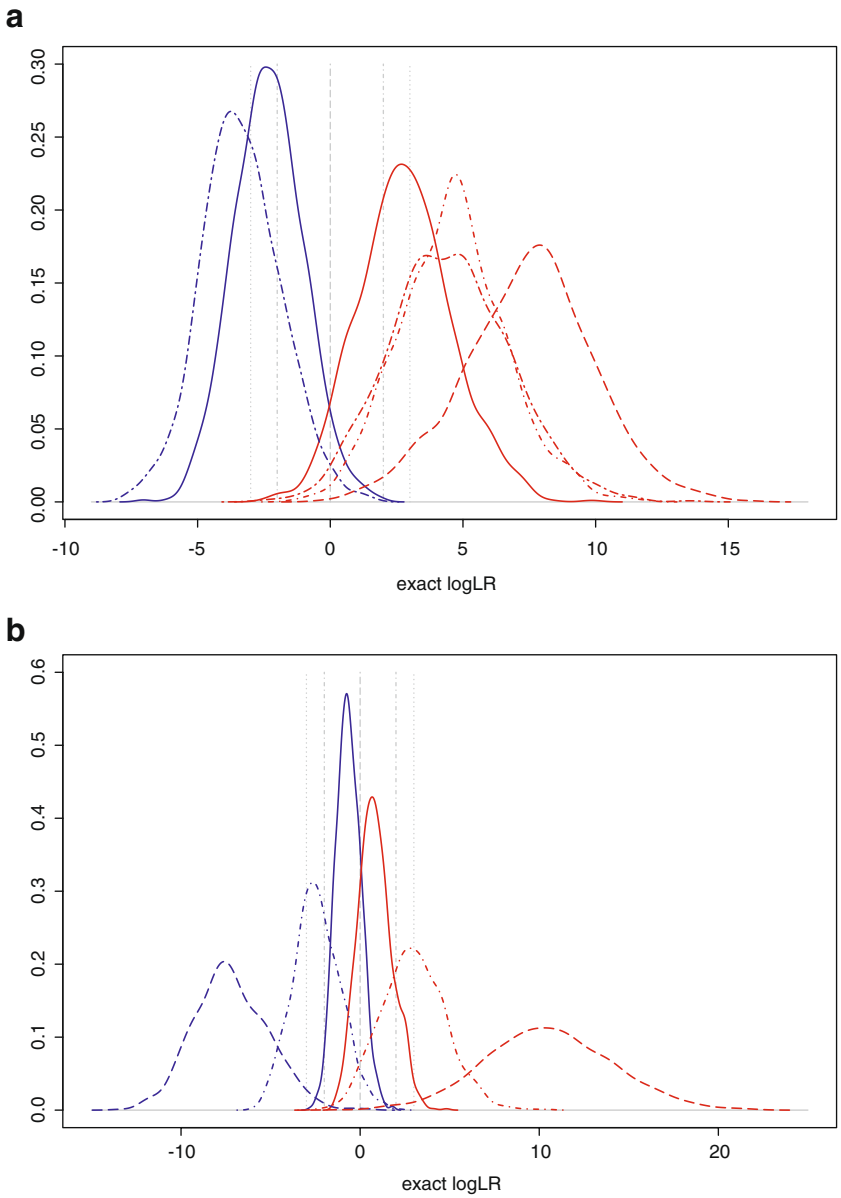

Fig. 4 Simulation-based exact logLR distribution when one or two additional relatives are included into two kinship scenarios. *Solid curves*: original scenario, i.e. aunt–niece *vs* unrelated (**a**) and first cousins vs unrelated (**b**), *short dash-dotted curves*: including one additional relative, namely a second aunt in **a** and one parent in **b**, *long-dash-dotted curve*: inclusion of mother instead of a second aunt (in **a** only), *dashed curves*: both additional relatives included, *vertical lines*: see legend to Fig. 2; for the colour coding, see legend to Fig. 1; for scenarios where the mother is included in **a**, the curves corresponding to H_0 are missing because Mendelian incompatibilities would occur with positive probability in such cases

STRs under study [18–22]. For the sake of consistency, we therefore choose not to use this fragmentary mapping information here but to approximate genetic distances by physical position instead. Given the considerable variation in recombination intensity known to exist along individual chromosomes [23, 24], this implies that even our 'exact' results may only represent first approximations themselves. A more precise assessment of the effects of linkage upon pairwise kinship testing and the introduction of linkage-based likelihood calculations into routine casework would both require the (highly warranted) construction of comprehensive genetic maps for forensic STRs.

Despite the obvious relevance of linkage for solving individual kinship cases, power considerations may safely be based upon the assumption of independent IBD relationships, i.e. of unlinked markers at gametic equilibrium. Based upon our simulations, we conclude that this simplification is sufficiently accurate in practise to judge whether a certain pairwise kinship case can be solved with a given set of markers or not. The mean and variance of the marker-specific logLR values provided in Supplementary Table S1 allow such power calculations to be performed for specific sets of markers with little extra effort. What is more, the approximate power estimates derived above were also found to correspond well to our own practical experience: of the 20 FS-HS cases analysed in our laboratory in the past, only three yielded a logLR larger than 2, using the PowerPlex® set alone, whereas one

yielded logLR<−2. In the remaining 17 cases, the results were inconclusive and required the inclusion of additional STRs. Assuming that 50% (i.e. ten) of the implied relationships were indeed correct and that 50% were incorrect, these observations agree perfectly well with the corresponding power estimates given above (28.9% and 12.6%).

One way to improve the power of wider pairwise kinship tests would be to increase the number and density of the available STRs. However, such an extension would only be sensible if, at the same time, sufficiently accurate estimates of the respective population allele frequencies and of the inter-marker genetic distances would become available. As a first step in this direction, a genetic map of 39 X-chromosomal STRs suitable for kinship analysis has recently been published [23]. However, since most applications of STRs in forensic practice, including trace donor matches and paternity tests, do not require additional markers over and above the 34 STRs taken into consideration here, it remains to be seen whether such additional efforts will also be made for the autosomal part of the genome. In any case, any inclusion of additional markers would inevitably increase the need to take linkage properly into account in the likelihood calculations because the discrepancy between exact and rough logLR values tends to increase with increasing marker density. Furthermore, increasingly dense marker maps will eventually create levels of inter-marker linkage disequilibrium that can no longer be ignored in likelihood calculations as well and that will require the construction of comprehensive haplotype instead of allele frequency databases [17]. Finally, inspection of Table 3 reveals that approximately three to four times as many markers as are currently available would be required to increase the mean logLR for the assessment of a first cousin relationship to that of half sibs. In most instances, it would undoubtedly be more economic to genotype one or two additional individuals, if available, rather than to triple or quadruple the genotyping load.

In summary, we may conclude that pairwise kinship analysis is a feasible option for the assessment of first- and second-degree relationships but that wider kinship has to be, and will probably continue to have to be, assessed by the involvement of additional individuals. Whilst power considerations can be based upon large-sample approximations assuming the stochastic independence of markers, likelihood calculations in individual cases should be performed taking the physical linkage of STRs properly into account.

Acknowledgements This study was supported by the German Ministry of Education and Research (BMBF) through National Genome Research Network (NGFNplus) grant 01GS0809.

References

1. Henke L, Aaspollu A, Biondo R, Budowle B, Drobnic K, van Eede PH, Felske-Zech H, Fernández de Simón L, Garafano L, Gehrig C, Luckenbach C, Malik N, Muche M, Parson W, Primorac D, Schneider PM, Thomson J, Vanek D (2003) Evaluation of the STR typing kit PowerPlex™ 16 with respect to technical performance and population genetics: a multicenter study. Int Congr Ser 1239:789–794
2. Schmid D, Anslinger K, Rolf B (2005) Allele frequencies of the ACTBP2 (=SE33), D18S51, D8S1132, D12S391, D2S1360, D3S1744, D5S2500, D7S1517, D10S2325 and D21S2055 loci in a German population sample. Forensic Sci Int 151:303–305
3. Cox DR, Hinkley DV (1974) Theoretical statistics. Chapman & Hall/CRC, Boca Raton
4. Feller W (1968) An introduction to probability theory and its applications. Wiley, New York
5. Essen-Möller E (1938) Die Beweiskraft der Ähnlichkeit im Vaterschaftsnachweis. Theoretische Grundlagen. Mitt Anthropol Ges 68:9–53
6. Ryman N, Hansson A, Hirschfeld J, Swan T (1981) Probability of paternity exclusion in different mother–child genotype combinations. Hereditas 94:99–104
7. Selvin S (1980) Probability of nonpaternity determined by multiple allele codominant systems. Am J Hum Genet 32:276–278
8. Selvin S, Grunbaum BW, Myhre BA (1983) The probability of exclusion or likelihood of guilt of an accused: paternity. J Forensic Sci Soc 23:19–25
9. Kosambi DD (1944) The estimation of the map distance from recombination values. Ann Eugen 12:172–175
10. Leal SM, Yan K, Muller-Myhsok B (2005) SimPed: a simulation program to generate haplotype and genotype data for pedigree structures. Hum Hered 60:119–122
11. Cottingham RW Jr, Idury RM, Schaffer AA (1993) Faster sequential genetic linkage computations. Am J Hum Genet 53:252–263
12. Lathrop GM, Lalouel JM (1984) Easy calculations of LOD scores and genetic risks on small computers. Am J Hum Genet 36:460–465
13. Lathrop GM, Lalouel JM, Julier C, Ott J (1984) Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 81:3443–3446
14. Schaffer AA, Gupta SK, Shriram K, Cottingham RW Jr (1994) Avoiding recomputation in linkage analysis. Hum Hered 44:225–237
15. O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63:259–266
16. R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
17. Tillmar AO, Mostad P, Egeland T, Lindblom B, Holmlund G, Montelius K (2008) Analysis of linkage and linkage disequilibrium for eight X-STR markers. Forensic Sci Int Genet 3:37–41
18. Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63:861–869
19. Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K (2002) A high-resolution recombination map of the human genome. Nat Genet 31:241–247
20. Kong X, Matise TC (2005) MAP-O-MAT: internet-based linkage mapping. Bioinformatics 21:557–559

21. Matise TC, Chen F, Chen W, De La Vega FM, Hansen M, He C, Hyland FC, Kennedy GC, Kong X, Murray SS, Ziegle JS, Stewart WC, Buyske S (2007) A second-generation combined linkage physical map of the human genome. Genome Res 17:1783–1786
22. Matise TC, Gitlin JA (1999) MAP-O-MAT: marker-based linkage mapping on the World Wide Web. Am J Hum Genet 65:A435
23. Machado FB, Medina-Acosta E (2009) Genetic map of human X-linked microsatellites used in forensic practice. Forensic Sci Int Genet 3:202–204
24. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P (2005) A fine-scale map of recombination rates and hotspots across the human genome. Science 310:321–324