American Immigration Lawyers Association
1331 G Street, NW, Suite 300
Washington, DC 20005

UNITED STATES DEPARTMENT OF JUSTICE
EXECUTIVE OFFICE FOR IMMIGRATION REVIEW
BOARD OF IMMIGRATION APPEALS
FALLS CHURCH, VIRGINIA

In the Matter of	
Nejat Ibrahim Ruzku, beneficiary of a visa petition filed by his brother, Abdalla Ibrahim Ruzku	A205-744-090 Visa Petition Proceedings

BRIEF OF AMICUS CURIAE THE AMERICAN IMMIGRATION LAWYERS ASSOCIATION

Russell Reid Abrutyn (DZ32031)
Marshal E. Hyman & Associates
3250 W. Big Beaver, Suite 529
Troy, MI 48084
(248) 643-0642
rabrutyn@marshalhyman.com

Attorney for Amicus Curiae

BRIEF OF AMICUS CURIAE

The United States Citizenship and Immigration Services' blanket refusal to accept sibling-to-sibling DNA tests is a result of the agency's misunderstanding of both the science behind those tests and the preponderance of the evidence standard that petitioners have to meet in visa petition proceedings. The agency's refusal to consider the results of these tests, called sibship DNA tests, whether as stand-alone evidence or in tandem with other evidence, will unnecessarily impede the ability of U.S. citizen petitioners to sponsor their brothers and sisters.

Amicus American Immigration Lawyers Association (AILA) is submitting this brief addressing the two questions asked by the Board in the supplemental briefing notice: (1) whether direct sibling-to-sibling DNA testing achieving a 99.5% or greater degree of certainty for the relationship, which is accepted as the industry standard in paternity/maternity cases, is sufficient, alone, to establish a biological sibling relationship by a preponderance of the evidence in visa petition cases? (2) if a greater degree of certainty should be required for sibling-to-sibling DNA test results, what percentage of probability, if any, is considered sufficient under the preponderance of the evidence standard for DNA evidence alone to establish a biological sibling relationship? Following the Board's request for supplemental briefing, the United States Citizenship and Immigration Services (USCIS) issued a

policy memo purporting to create a blanket rule of refusing to consider any sibling-to-sibling DNA test result regardless of the circumstances.[1]

The prevailing view in the scientific community is that a sibship DNA test that results in a probability of relationship percentage over 99.5% is sufficient to satisfy a petitioner's burden of proving the sibling relationship by a preponderance of the evidence. Because of the risk of false negatives, a test resulting in a lower percentage does not exclude the relationship; a lower test result must be considered along with the other evidence submitted. Amicus will first address the insurmountable problems that the USCIS policy creates for some U.S. citizen petitioners. Amicus will then address the USCIS's apparent belief that petitioners must prove the claimed relationship by an absolute certainty because this suggests that the agency misunderstands the preponderance of the evidence standard. Amicus will then explain the science behind a sibship DNA test and why it is possible to obtain a test result that satisfies the petitioner's burden of proof, or, where the test result falls below that threshold, a result that can be combined with other reliable and probative evidence to meet that burden.[2]

I. Statement of Interest

The American Immigration Lawyers Association (AILA) is a national association with more than 13,000 members throughout the United States and

[1] http://www.uscis.gov/sites/default/files/files/nativedocuments/2014-1017_DNA_Evidence_of_Sibling_Relationships_PM_Effective.pdf (last visited November 19, 2014).

[2] Amicus takes no position on the merits of the petitioner's appeal or the other issues raised in that appeal.

abroad, including lawyers and law school professors who practice and teach in the field of immigration and nationality law. AILA seeks to advance the administration of law pertaining to immigration, nationality and naturalization; to cultivate the jurisprudence of the immigration laws; and to facilitate the administration of justice and elevate the standard of integrity, honor and courtesy of those appearing in a representative capacity in immigration and naturalization matters. AILA's members practice regularly before the Department of Homeland Security, the Executive Office for Immigration Review, United States District Courts and Courts of Appeal, and the United States Supreme Court.

II. The USCIS's New Policy Creates an Insurmountable Burden for Some Sibling Petitioners

It is unreasonable and contrary to the prevailing view of the scientific community for the USCIS to reject all sibship DNA tests. This new blanket policy creates an insurmountable burden for some petitioners because sibship DNA testing may be the best available evidence to prove the claimed relationship if primary or secondary evidence is unavailable, the USCIS finds that evidence insufficient, or it is impossible or impractical to obtain a parental DNA test.

There are a number of situations where petitioners cannot obtain parental DNA tests. The most obvious one is if the parent is deceased. Second, because the U.S. citizen petitioner's parent is not a party to the visa petition proceedings, the petitioner cannot compel the parent to submit to a DNA test if the parent is unwilling to cooperate. DNA tests are intrusive, inconvenient, and expensive. Third, if the parent lives in a remote or dangerous part of the world, it may be impossible

or impractical for the parent to travel to the location of the DNA test.[3] Fourth, the parent may live in a part of the world where there are no acceptable DNA testing facilities.

If the USCIS does not believe that a petitioner's primary or secondary evidence is sufficient, a DNA test is often the only other way for siblings to establish their relationship. A DNA test conducted under the agency's oversight is more reliable and less subject to manipulation than other evidence and it can help resolve any doubts an adjudicator has about the reliability of the evidence.

III. The USCIS's New Position Rejecting all Sibling DNA Testing Springs from a Misunderstanding of the Preponderance of the Evidence Standard

The preponderance of the evidence standard applies to visa petition proceedings. *Matter of Martinez*, 21 I&N Dec. 1035, 1036 (BIA 1997).[4] This is the lowest standard of proof and is met when the trier of facts determines "that the existence of a fact is more probably than its nonexistence." *Siddiqui v. Holder*, 670 F.3d 736, 741-45 (7th Cir. 2012) *quoting Concrete Pipe & Prods. of Cal. Inc. v. Constr. Laborers Pension Trust for So. Cal.*, 508 U.S. 602, 622 (1993). If the evidence is "sufficiently reliable and sufficiently probative to demonstrate the truth

[3] In regions where DNA testing cannot be done under the supervision of a U.S. government official or authorized DNA testing facility, it will be impossible to establish a reliable chain of custody.

[4] A higher standard of proof only applies where the law specifies a different standard. *Matter of Patel*, 19 I&N Dec. 774, 782-83 (BIA 1988) (applying a higher standard where specifically required by statute). The "clear and convincing" standard only applies where a lawful permanent resident petitions for a new spouse within five years of obtaining his status through a prior marriage or if the marriage was entered into during proceedings. INA §§ 204(a)(2)(A)(ii), 245(e)(3). It is unlikely that a sibling petition will ever trigger this heightened standard of proof.

of the asserted proposition with the requisite degree of certainty," then the petitioner has met his or her burden. *Concrete Pipe*, 508 U.S. at 602.

The petitioner need only show that his or her claim is "probably true." *Matter of E-M-*, 20 I&N Dec. 77, 79-80 (Comm'r 1989). Even if the adjudicator has some doubt as to the truth of the claim, the petitioner has satisfied his or her burden of proof through the submission or relevant, probative, and credible evidence that "leads the [adjudicator] to believe that the claims is 'more likely than not' or 'probably' true." *Matter of Chawathe*, 25 I&N Dec. 369, 376 (AAO 2010)[5] *quoting INS v. Cardoza-Fonseca*, 480 U.S. 421, 431 (1987).

The preponderance of the evidence standard in immigration proceedings is the same standard that the federal courts use. Under this standard, the petitioner does not have to prove that something is conclusively true or true beyond a reasonable doubt. Rather, the petitioner only has to prove that something is more likely than not, or that there is a slightly greater than 50% chance of something being true. *Cardoza-Fonseca*, 480 U.S. at 431.

The USCIS's new policy rejecting all sibship DNA test results strays far from the preponderance of the evidence standard. Even if a DNA test result leaves "some doubt" as to the claimed relationship because there is a 1 or 2% chance of a false positive, this is an insufficient basis for denying a petition under the preponderance of the evidence standard. *Chawathe*, 25 I&N Dec. at 376. A DNA test does not have

[5] This decision is binding on the USCIS. 8 C.F.R. § 1003.1(i); *Matter of Al Wazzan*, 25 I&N Dec. 359, 359 n.1 (AAO 2010). It is also binding on the Board in visa petition proceedings. 8 C.F.R. § 1003.1(g).

to be 100% conclusive to meet the preponderance of the evidence test; it only has to show that it is more likely than not that the claimed relationship is true.[6]

IV. The USCIS's Policy Memo has no persuasive value

Despite considering this issue for at least three years, the USCIS did not announce its policy until after the Board issued this request for supplemental briefing.[7] The USCIS's policy was not made after a notice-and-comment regulatory process and is contrary to the recommendations of its own experts. The USCIS new policy does not bind the Board and given the circumstances under which it was made, it lacks any persuasive value.

Internal USCIS memoranda are not binding on the Board. *Matter of Arrabally*, 25 I&N Dec. 771, 776 n.4 (BIA 2013). These memoranda are entitled to respect only to the extent they have the power to persuade. *Id.*; *Skidmore v. Swift & Co.*, 323 U.S. 134, 140 (1944). As in *Arrabally*, the USCIS policy memo addressing sibship DNA testing is not binding on the Board and is unpersuasive because the memo is contrary to the recommendations of its own experts.

In 2011, government experts gave a presentation to the USCIS as it was considering its policy on DNA tests. *Interpretation of DNA Typing Results for*

---

[6] It is worth noting that the USCIS will accept a parent-child DNA test result of 99.5%.

[7] http://www.uscis.gov/sites/default/files/files/nativedocuments/2014-1017_DNA_Evidence_of_Sibling_Relationships_PM_Effective.pdf (last visited November 19, 2014).

Kinship Analysis, O'Connor, Kristen Lewis, Jan. 25, 2011 (O'Connor Report)[8]. As far as amicus can determine, the agency did not act on these recommendations until receiving the Board's request for supplemental briefing.

Contrary to the USCIS's position, the government experts concluded that it is possible to obtain sibship DNA test results that satisfy the preponderance of the evidence standard. The government experts noted that in the United Kingdom, probability results of 90% or greater are "very strong evidence" of a full or half-sibling relationship. O'Connor Report at p. 23. On page 24, the experts concluded that a probability result of 99.5% means that it is more likely than not that a pair of individuals are full siblings. A test result of less than 100% for a parent-child test is sufficient for immigration and other purposes. Properly conducted, a sibship DNA test can meet or exceed the 99.5% threshold that has been accepted as conclusive for parent-child testing. *Id.* at 26.

The test results become more reliable if more genetic markers, or alleles, are measured. The tests start from the position that there is an even chance that two individuals are or are not related. By then comparing their genetic markers, it is possible to determine how likely it is that the two individuals are or are not related. With paternity or maternity tests, a 100% match is possible but not required as the accepted threshold is at 99.5%, and even lower for some purposes. Sibship tests might not reach a certainty level of 100%, but it can reach the 99.5% threshold.

---

8 http://www.cstl.nist.gov/strbase/pub_pres/OConnor_USCIS_interpretation%20of%20DNA.pdf (last visited November 19, 2014) (Tab A).

Even at a 92% certainty level, the risk of a false positive is only about 0.02%

O'Connor Report at 26

V. Sibship DNA Tests are Sufficiently Reliable to Satisfy the Preponderance of the Evidence Standard and a 99.5% result should be accepted as conclusive.

The scientific community has agreed that properly conducted sibship DNA tests are sufficiently reliable for use in legal and other proceedings. The prevailing view in the scientific community is that a sibship DNA test result of 99.5% makes it more likely than not that two individuals are siblings. At that threshold, depending on the number of loci, or allele, tested, the risk of a false positive ranges from about 0.1% to about 2-3%. Properly conducted, sibship DNA testing is just as reliable as parent-child DNA testing and can produce results that are just as conclusive. DNA test results below 99.5% should not be taken to exclude the claimed relationship because there is a risk of false negatives. Rather, these test results should be considered along with the other evidence submitted.

To help the Board understand the science of sibship DNA testing, Amicus has attached an affidavit and curriculum vitae of Dr. Monte Wayne Miller at Tab B, an article by Dr. Kristen Lewis O'Connor at Tab C, and an excerpt of from the Annual Report Summary for Testing in 2008 by AABB[9] at Tab D.

Dr. Miller has been involved in DNA testing and analysis for more than 15 years. He has served as a lab director, consultant, and expert witness in military,

[9] The full report is available at https://www.aabb.org/sa/facilities/Documents/rtannrpt08.pdf (last visited November 19, 2014).

federal, and state courts on DNA issues. He concludes that a sibship DNA test result "can serve as sufficient evidence that it is more likely than not that two individuals are siblings." Tab B at p. 2. The risk of a false positive is within the limits that the courts find acceptable. *Id.* There is a greater risk of a false negative, so while a positive test can serve as evidence of a claimed relationship, a negative test should not serve as conclusive evidence that the relationship does not exist. Tab B at p. 3-4.

In terms of selecting an acceptable probability determination, Dr. Miller concludes that the 99.5% probability level for parent-child testing should apply to sibship testing. Tab B at p. 2. The mathematical methods used in sibship analysis are just as reliable as those used in parent-child testing. *Id.* While sibship testing may not provide the highest percentage certainty that a parent-child test theoretically can produce, it can provide probability determinations of 99.5% or greater. Tab B at p. 3-4. That threshold is more than sufficient to establish the claimed parent-child relationship in immigration and other legal proceedings.

These expert sources conclude that it is possible to obtain reliable sibship DNA tests if proper care is taken in conducting the test. The AABB, which the USCIS cites as an expert source in its policy memo, indicates that a 90% probability of a sibling relationship is "reasonable evidence of either a full or half sibling relationship." Tab D at p. 3. If the AABB is willing to accept a 90% probability as reasonable evidence, then there should be no doubt that 99.5% probability is sufficient by itself to meet the preponderance of the evidence standard. A

probability result below 99.5% should therefore not be dispositive of the claimed relationship but should instead be considered along with the other evidence presented.

At 99.5%, there is a one in 200 chance that two individuals who claim to be siblings and who receive a positive DNA test are not actually related as siblings. This means that it is more likely than not that they are related, especially since a false positive would result from chance and not through manipulation by the parties.

Dr. Kristen O'Connor, who presented to the USCIS on behalf of the National Institute of Standards and Technology, has likewise concluded that the likelihood of a false positive ranges from a high of 2-3% to a low of less than 1%, depending on the number of loci, or allele, that are tested. Tab C at p. 5. When 40 loci are tested, the chance of a false negative or positive is only about 0.1%. Tab C at p. 6. As additional loci are tested, there is an increased risk of a false negative, so adjudicators should not rely solely on a negative DNA test to conclude that the claimed relationship does not exist. Tab C at p. 7. On page 27 of Dr. O'Connor's report, she provides a chart showing the likelihood of false negatives or positive depending on the race of the individuals tested, the number of loci tested, and the claimed relationship.

CONCLUSION

Amicus Curiae the American Immigration Lawyers Association respectfully requests that the Board adopt the prevailing view within the scientific community that a sibling-to-sibling DNA test resulting in a 99.5% probability of the claimed relationship is sufficient to meet the petitioner's burden of proving the claimed relationship by a preponderance of the evidence. The Board should further conclude that a percentage less than 99.5 is not dispositive of the claimed relationship and should instead be considered along with the evidence and circumstance presented with each petition.

Respectfully submitted November ____, 2014

Russell Abrutyn
Marshal E. Hyman & Associates, PC
3250 West Big Beaver Road, Suite 529
Troy, Michigan 48084

Counsel for Amicus Curiae

CERTIFICATE OF SERVICE

On November ____, 2014 I, served a copy of the foregoing BRIEF OF AMICUS CURIAE THE AMERICAN IMMIGRATION LAWYERS ASSOCIATION on the following via first class mail:

Howard Robert Davis Esq.
Howard R. Davis
1351 Third Street, Suite 301
Santa Monica, CA 90401

DHS/CIS- California Serv. Ctr./CSC
24000 Avila Rd., Div. VII/CRU
Laguna Niguel, CA 92677

_____
Russell Abrutyn
Attorney for Amicus Curiae

INDEX

Tab A

Interpretation of DNA Typing Results for Kinship Analysis

Kristen Lewis O'Connor, Ph.D.

National Institute of Standards and Technology

USCIS Working Group on DNA Policy
Washington, DC
January 25, 2011

NIST
National Institute of Standards and Technology

Questions to Be Addressed

- How is DNA typing used to assess relatedness?
- How do we interpret kinship analysis results?
- What are some issues that need consideration?

What is kinship analysis?

Evaluation of relatedness between individuals

<u>Applications</u>

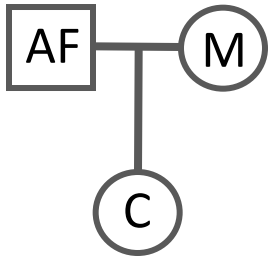
Parentage testing (civil or criminal)

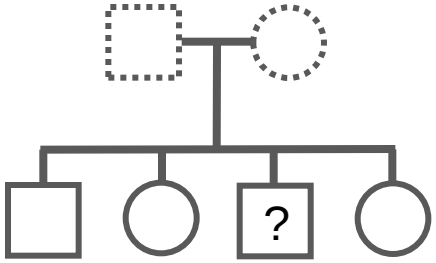
Disaster victim identification

Missing persons identification

Familial searching

Immigration

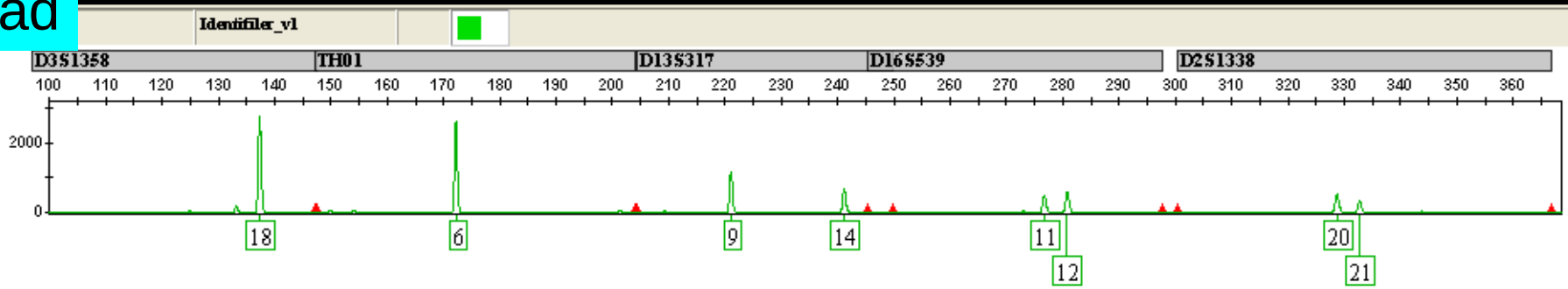




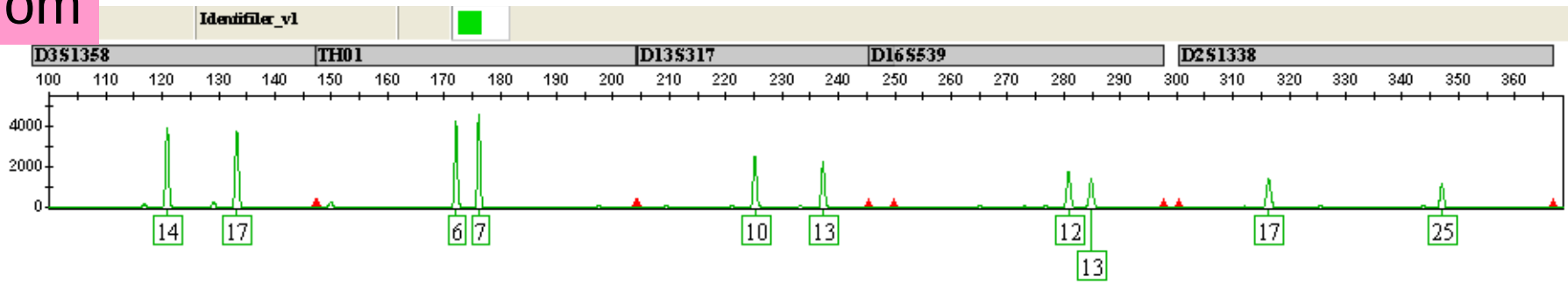
Fundamentals of Paternity Testing

Focusing on 5 markers…

Dad

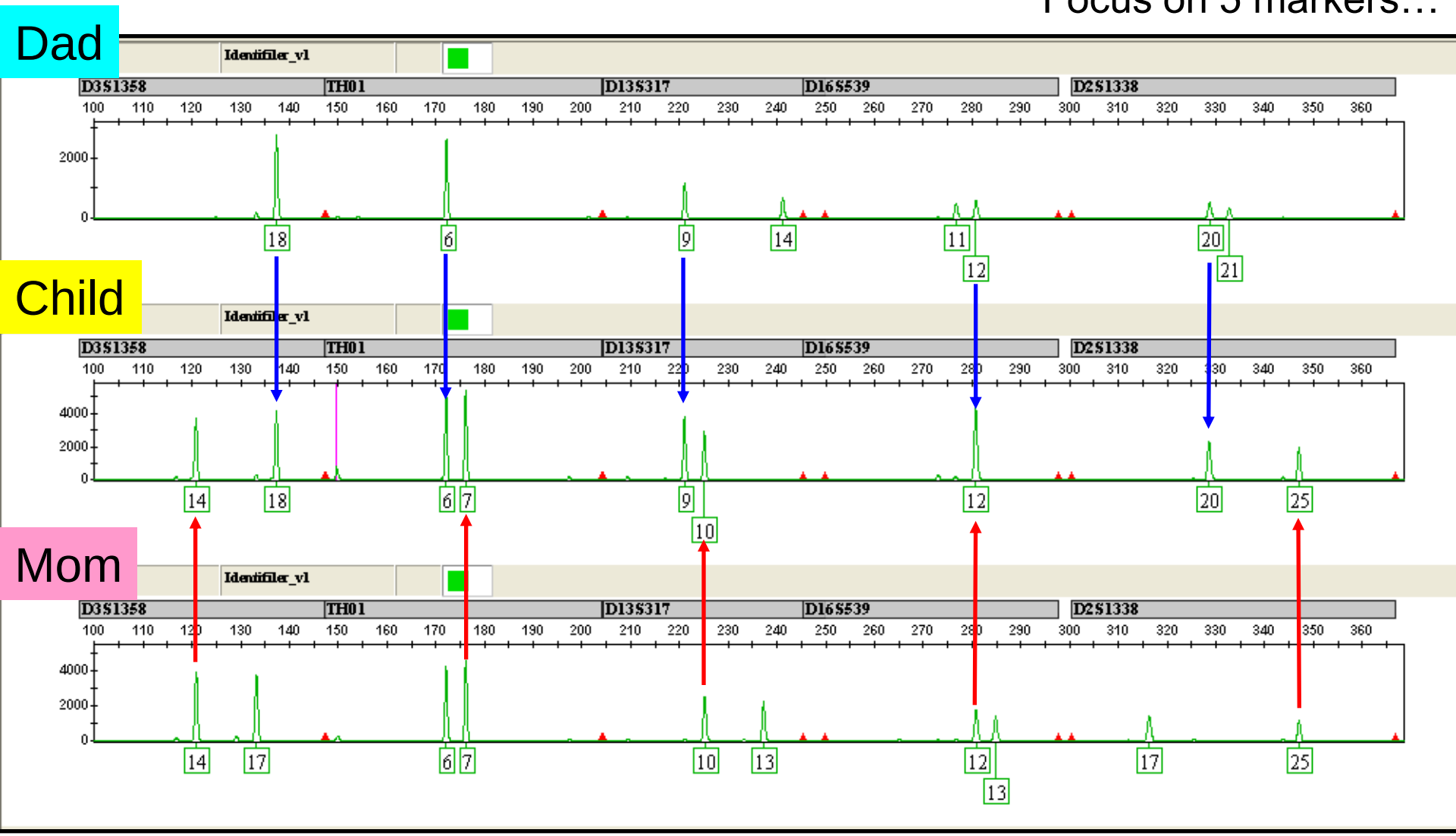


Mom



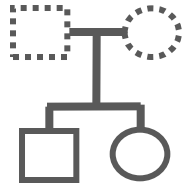
Fundamentals of Paternity Testing

Focus on 5 markers…

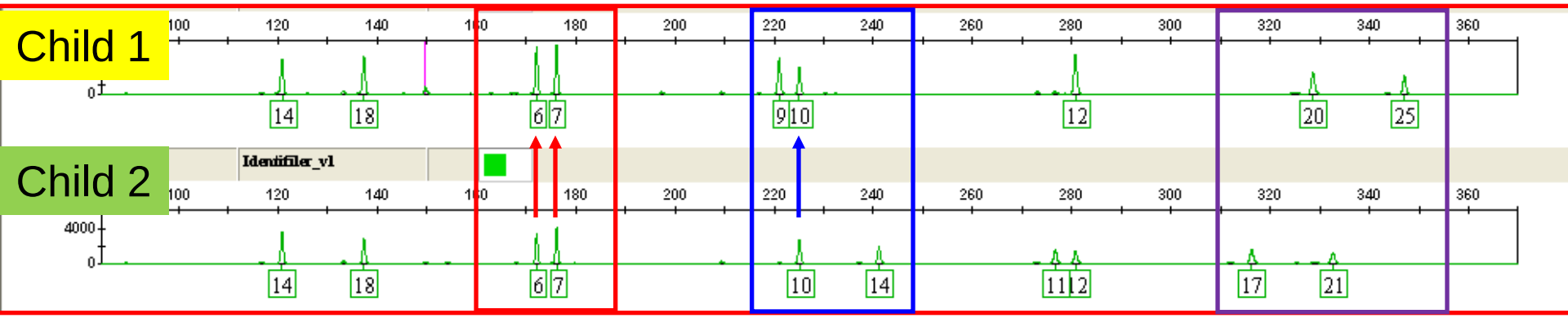


Parent-offspring will share **one** allele at every locus

Kinship Analysis: Full Siblings



Focusing on 5 markers…



Full siblings may share **two**, **one**, or **zero** alleles at a locus

For more distant familial relationships, allele sharing decreases

Uncertainty Increases

Why can kinship analysis be complex?

For more distant familial relationships,
allele sharing decreases → uncertainty increases

Probability of Sharing Alleles from a Common Ancestor

High

Level of Certainty

Low

Relationship	0 alleles	1 allele	2 alleles
Parent-child	0	1	0
Full siblings	1/4	1/2	1/4
Half siblings	1/2	1/2	0
Uncle-nephew	1/2	1/2	0
Grandparent-grandchild	1/2	1/2	0
First cousins	3/4	1/4	0

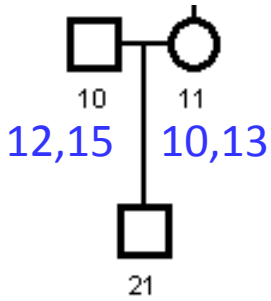
Half siblings, uncle-nephew, and grandparent-grandchild are genetically identical

AILA InfoNet Doc. No. 14112543. (Posted 11/25/14)

What information is required for kinship analysis?

1. Alleged relationship
2. Genotypes at specific markers
3. Method to assess the relationship

Paternity trio

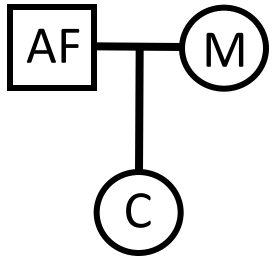


Marker D8S1179

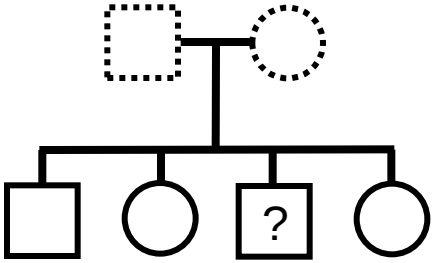
What information is required for kinship analysis?

1. Pedigree of claimed relationships

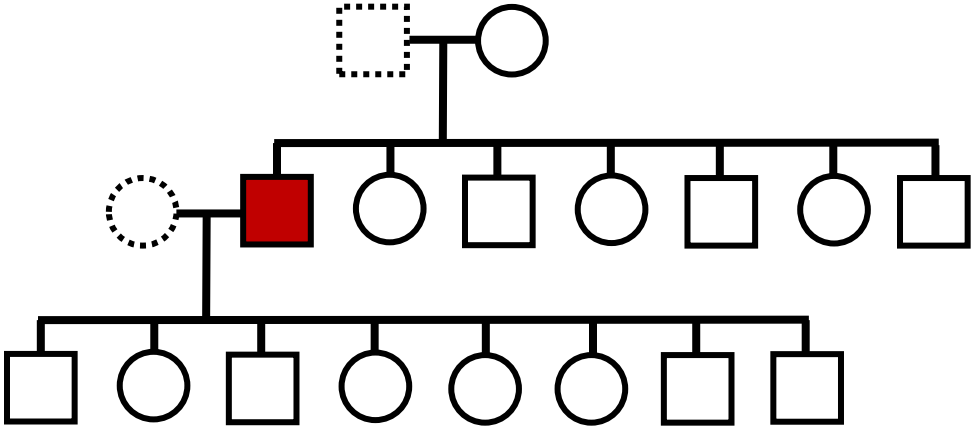
Paternity trio



Full siblings



Complex pedigree



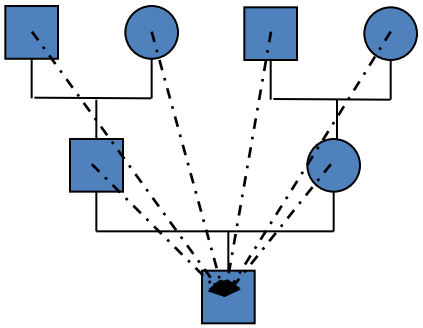
Define relationships in a pedigree (“family tree”)

Collect DNA samples from informative individuals

AILA InfoNet Doc. No. 14112543. (Posted 11/25/14)

What information is required for kinship analysis?

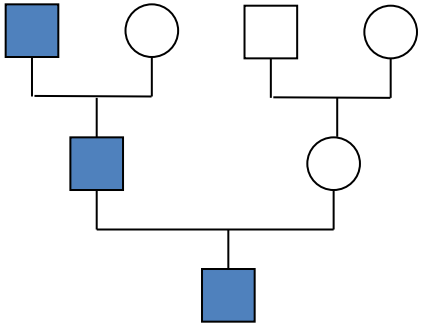
2. Genotypes for individuals making a claim



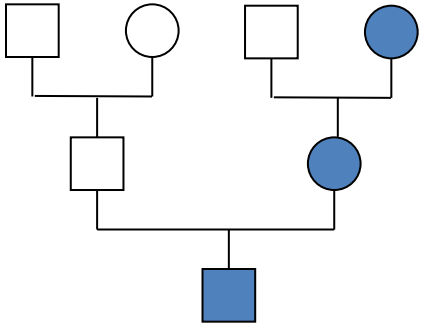
Autosomal
(**passed on in part**,
from all ancestors)

- Typically test 13-25 STR loci
- Work well for close relatives (parentage and full siblings)
- Need more family references for distant relatives

Lineage Markers



Y-Chromosome
(**passed on complete**,
but only by sons)



Mitochondrial
(**passed on complete**,
but only by daughters)

AILA InfoNet Doc. No. 14112543. (Posted 11/25/14)

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition,* Figure 9.1, ©Elsevier Science/Academic Press

What information is required for kinship analysis?

3. Method to assess the relationship

The question is **NOT** "Are they related?"

The question is "Is the **claimed** relationship supported by the genetic and non-genetic evidence?"

Remainder of this presentation will cover the method to assess relatedness:
Likelihood ratio, prior probability, posterior probability

Likelihood Ratio (LR)

Describes how strongly the genotypes support one relationship versus the other relationship

Expresses the likelihood of obtaining the DNA profiles under two mutually exclusive hypotheses

$$\mathbf{LR} = \frac{\text{Probability of genotypes if individuals are related as claimed}}{\text{Probability of genotypes if individuals are unrelated}}$$

The LR takes into account:

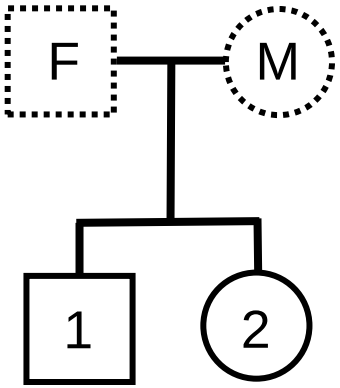
- the **probability of allele sharing** for individuals with a specific relationship
- the **allele frequency** of alleles
- a possible mutation event (if necessary)

Likelihood Ratio (LR)

The LR is also called the relationship index (RI) or kinship index (KI).

Each independent locus tested produces its own relationship index, which can be multiplied by those of other independent loci to calculate a combined relationship index (CRI).

$$\textbf{CRI} = \frac{\text{Probability of genotypes if 1,2 are full siblings}}{\text{Probability of genotypes if 1,2 are unrelated}}$$



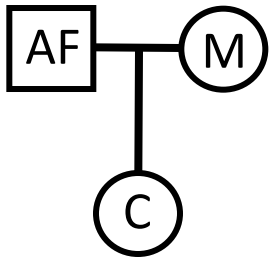
By the definition of a LR:
CRI > 1 supports the numerator (claimed relationship)
CRI < 1 supports the denominator (alternative relationship)

Larger CRI values provide more support for the claimed relationship

Likelihood Ratio (LR)

Hypothesis 1 = Paternity Trio, **Hypothesis 2 =** Unrelated

Paternity trio

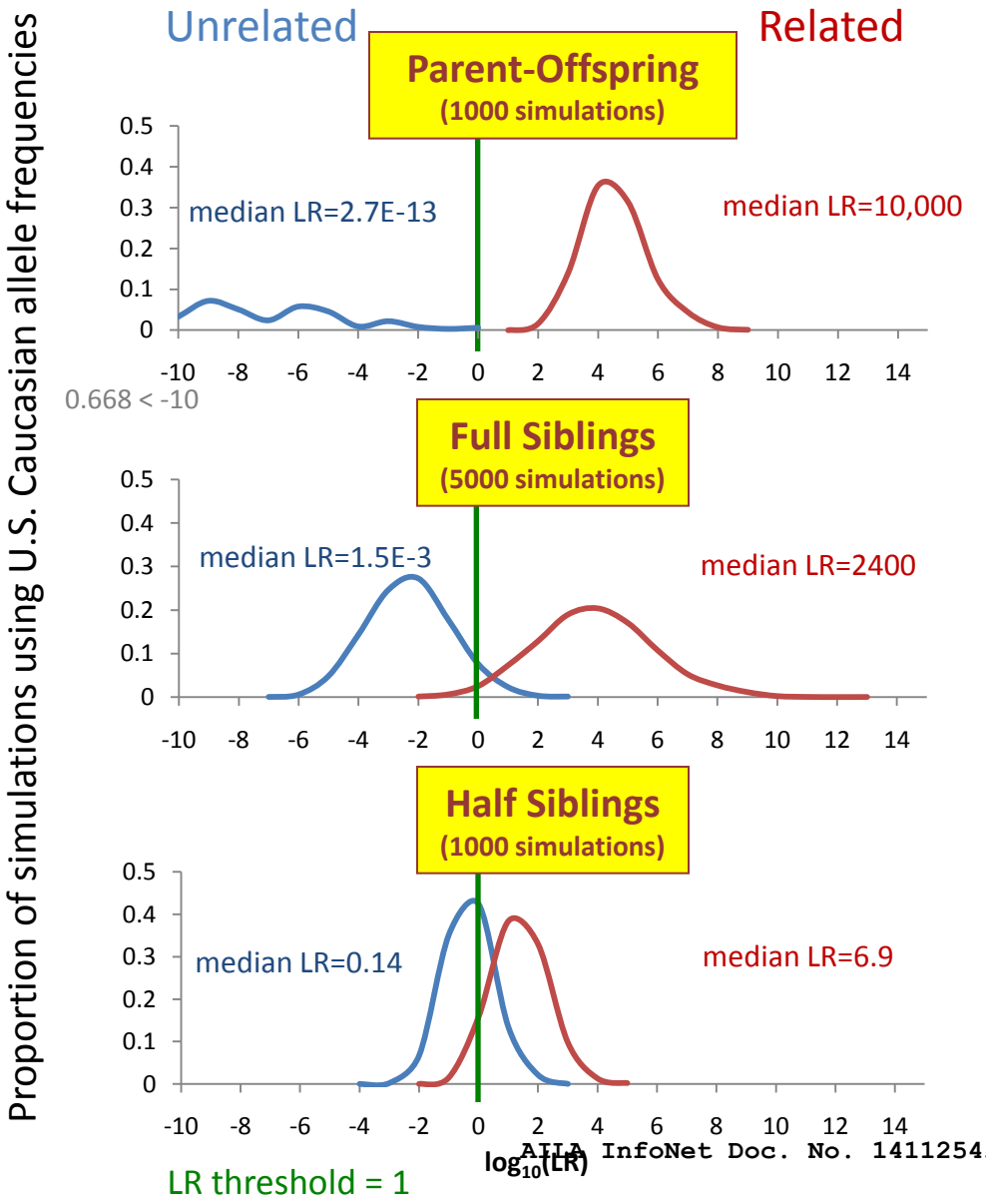


LR = 168,468,800

Locus	Probability (Hypothesis 1)	Probability (Hypothesis 2)	Likelihood Ratio
D8S1179	0.001545163	0.000574194	2.691012
D21S11	0.0003079	0.000171693	1.793322
D7S820	0.00078148	0.000138664	5.635774
CSF1PO	0.003673636	0.000798261	4.602047
D3S1358	0.002522579	0.001086988	2.320706
THO1	0.001420379	0.00032926	4.313852
D13S317	0.000454644	4.37E-05	10.39317
D16S539	9.47E-05	2.80E-05	3.38817
D2S1338	4.87E-05	1.15E-05	4.250356
D19S433	0.004076747	0.000661891	6.159245
VWA	0.000131184	5.26E-05	2.492709
TPOX	0.008606737	0.005087928	1.691599
D18S51	0.000328927	9.07E-05	3.625514
D5S818	0.002742154	0.000772507	3.549682
FGA	0.000532767	0.000198233	2.687581
Total	**2.27E-47**	**1.35E-55**	**168,468,800**

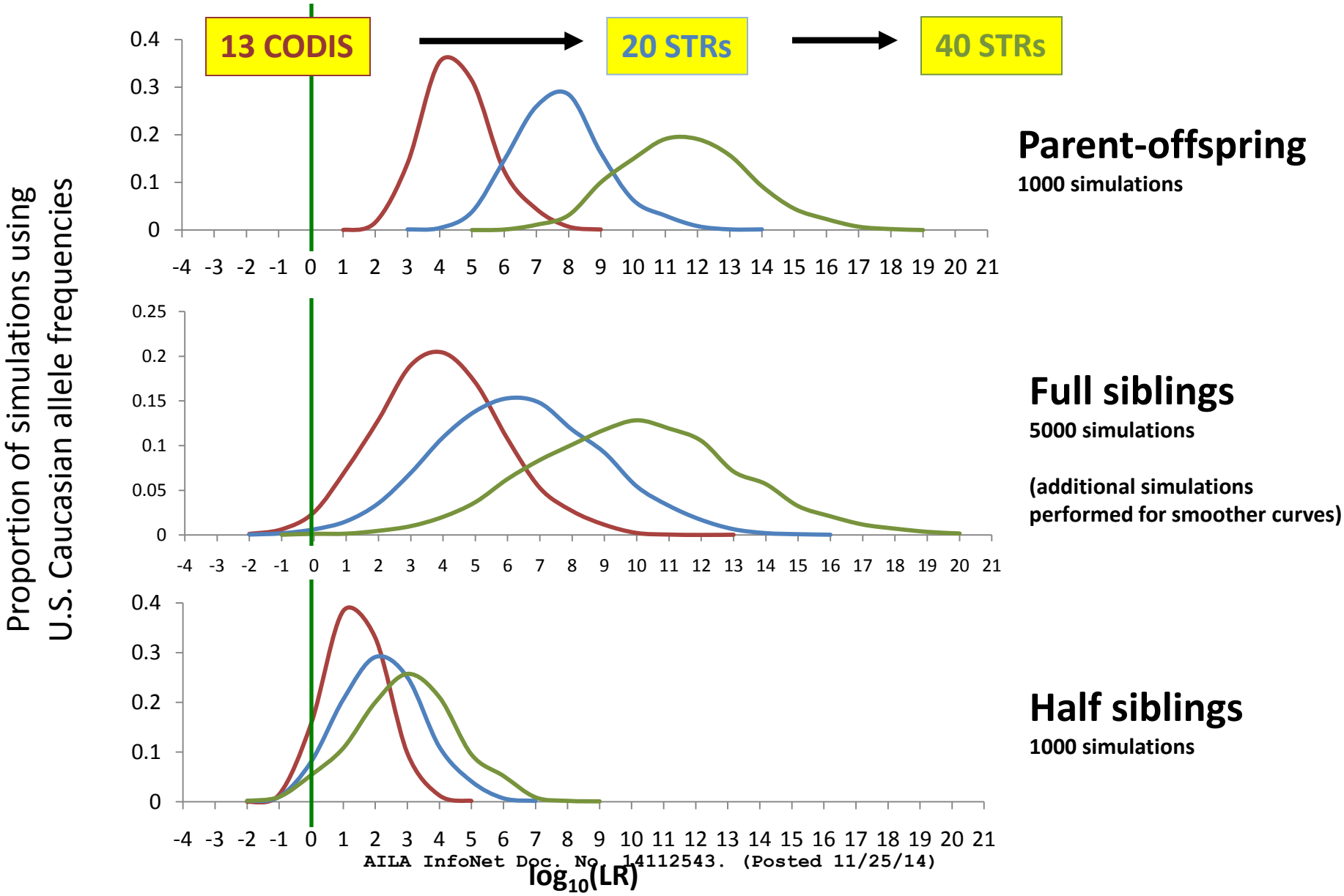
It is **168 million times** more likely that we observe these DNA profiles if the Alleged Father is the true father than if an unrelated man is the father of the child.

How do 13 loci perform for kinship analysis?



The degree of overlap corresponds with possible values for false positive or false negative results.

Do additional loci improve kinship determination?



Prior Probability

Describes the weight of non-genetic evidence **PRIOR** to DNA analysis

Case	**Prior Probability**	**Comment**
Paternity- U.S. courts	0.5	Both hypotheses are equally likely. Different priors could be claimed in court.
Missing Persons (ICMP)	1/N missing persons	Closed event (e.g., mass grave)
Immigration- U.S.	0.5	How do you assign weight to non-genetic evidence?

Relationship between Prior Probability and Prior Odds

Calculation of prior odds is necessary to combine the non-genetic information with the DNA information.

Prior odds are calculated using the prior probability as follows:

Prior Odds = Prior Probability/(1-Prior Probability)
= Pr/(1-Pr)

Example 1: Prior prob = 0.5

Prior Odds = 0.5/(1-0.5)
= 1

Example 2: Prior prob = 0.75

Prior Odds = 0.75/(1-0.75)
= 3

Posterior Odds

The posterior odds provide a numerical weight to the opinion of identification.

The mathematics for the combination of the kinship index and the prior odds is as follows:

Posterior Odds = Likelihood Ratio × Prior Odds

= CRI × P

Example with prior probability = 0.5 (prior odds = 1), and LR = 168,468,800

Posterior Odds = 168,468,800 *× 1*

= 168,468,800

Relationship between Posterior Odds and Posterior Probability

The probability of relationship (posterior probability) allows one to render an opinion about a relationship in understandable terms for the general public.

The probability of the relationship expressed as a percentage is calculated by the following equation:

Probability of Relationship = PO/(PO+1) × 100

or

Probability of Relationship = (CRI × Pr /[CRI × Pr + (1-Pr)]) × 100

where PO = Posterior Odds, Pr = Prior Probability, and CRI = Combined Relationship Index

Relationship between Posterior Odds and Posterior Probability

Example with prior probability = 0.5 (prior odds = 1), and LR = 168,468,800:

Probability of Relationship = (CRI × Pr /[CRI × Pr + (1-Pr)]) × 100

*= (*168,468,800 *× 0.5 /[*168,468,800 *× 0.5 + (1-0.5)]) × 100*

= 99.999999406418%

Posterior Probability

The probability of relationship (posterior probability) allows one to render an opinion about a relationship in understandable terms for the general public.

Case	Posterior Probability	Probability of Random Match
Paternity- U.S. courts	99.0-99.9%	0.1-1% (civil cases)
Missing Persons-ICMP	99.95%	0.05%
Immigration	99.5% (currently)	0.5%

Posterior Probability

The probability of relationship (posterior probability) allows one to render an opinion about a relationship in understandable terms for the general public.

Case	Posterior Probability	Conclusion
Paternity- U.K. (paternity or maternity)	99.99%	Positive: Very strong evidence of paternity/maternity
	0%	Negative: No support for relationship
Sibship- U.K. (full or half sibs)	90.00-99.99%	Positive: Very strong evidence of full/half siblingship
	10.00-89.99%	Inconclusive for relationship
	0-9.99%	Negative: No support for relationship

Alpha Biolabs

AILA InfoNet Doc. No. 14112543. (Posted 11/25/14)

http://www.alphabiolabs.com/assets/files/documents/DOT404VariousTypesofDNATestandtheTestingProcedureIssue01.pdf

Posterior Probability Varies with Different Priors

Table of posterior probabilities for different prior probabilities and likelihood ratios

Prior Probability	Paternity Index (LR)			
	1	**10**	**100**	**1,000**
0	0	0	0	0
0.001	0.001	0.00991	0.09099	0.5002501
0.010	0.010	0.09174	0.50251	0.9099181
0.100	0.100	0.52631	0.91743	0.9910803
0.500	0.500	0.90909	0.99009	0.9990010
0.900	0.900	0.98901	0.99889	0.9998889
0.990	0.990	0.99899	0.99989	0.9999899
0.999	0.999	0.99989	0.99999	0.9999990
1	1	1	1	1

AILA InfoNet Doc. No. 14112543. (Posted 11/25/14)

Evett and Weir, Interpreting DNA Evidence, 1998.

Range of Posterior Probabilities

Simulated pairs of individuals, either as true parent-child, full siblings, half siblings, or unrelated. **13 CODIS markers.**

Table shows the proportion of simulations within ranges of posterior probabilities (prior probability = 0.5)

Posterior Probability	True Parent-Child	Unrelated Parent-Child	True Full Siblings	Unrelated Full Siblings	True Half Siblings	Unrelated Half Siblings
0-10.0	0	0	0.0076	**0.9008**	0.017	**0.451**
10.0-20.0	0	**0.995**	0.0040	0.0356	0.030	0.161
20.0-30.0	0	0.002	0.0060	0.0170	0.034	0.099
30.0-40.0	0	0.002	0.0068	0.0096	0.035	0.074
40.0-50.0	0	0	0.0082	0.0096	0.057	0.060
50.0-60.0	0	0.001	0.0088	0.0056	0.055	0.039
60.0-70.0	0	0	0.0086	0.0060	0.077	0.035
70.0-80.0	0	0	0.0166	0.0060	0.090	0.027
80.0-90.0	0	0	0.0322	0.0050	0.137	0.028
90.0-95.0	0	0	0.0352	0.0020	0.145	0.017
95.0-99.0	0.019	0	0.1070	0.0018	**0.213**	0.009
99.0-99.5	0.024	0	0.0614	0.0006	0.046	0
99.5-99.9	0.121	0	0.1302	0.0004	0.049	0
99.9-100.0	**0.836**	0	**0.5674**	0	0.015	0

Caucasian genotypes simulated with NIST Caucasian allele frequency data. Mutations were not simulated.

Range of Posterior Probabilities

Simulated pairs of individuals, either as true parent-child, full siblings, half siblings, or unrelated. **20 markers (CODIS + 7 European markers).**

Table shows the proportion of simulations within ranges of posterior probabilities (prior probability = 0.5)

Posterior Probability	True Parent-Child	Unrelated Parent-Child	True Full Siblings	Unrelated Full Siblings	True Half Siblings	Unrelated Half Siblings
0-10.0	0	**1.000**	0.0022	**0.9724**	0.012	**0.683**
10.0-20.0	0	0	0.0018	0.0106	0.023	0.097
20.0-30.0	0	0	0.0008	0.0054	0.017	0.053
30.0-40.0	0	0	0.0014	0.0032	0.021	0.039
40.0-50.0	0	0	0.0022	0.0024	0.020	0.041
50.0-60.0	0	0	0.0004	0.0012	0.020	0.015
60.0-70.0	0	0	0.0020	0.0012	0.034	0.023
70.0-80.0	0	0	0.0026	0.0012	0.049	0.016
80.0-90.0	0	0	0.0092	0.0008	0.084	0.017
90.0-95.0	0	0	0.0094	0.001	0.101	0.008
95.0-99.0	0	0	0.0266	0.0004	**0.198**	0.007
99.0-99.5	0	0	0.0120	0	0.106	0.001
99.5-99.9	0	0	0.0578	0.0002	0.155	0
99.9-100.0	**1.000**	0	**0.8716**	0	0.160	0

Caucasian genotypes simulated with NIST Caucasian allele frequency data. Mutations were not simulated.

Issues to Consider

- Make sure the markers tested can meet/exceed your threshold for true relationships in question.

- What is the appropriate prior probability?
  - Prior probability of 0.5 may not adequately reflect prior information.
  - What if strong legal documents are presented?
  - What if you suspect fraud before DNA typing?

- What allele frequency databases will be used?
  - Need population-specific databases
  - Or calculate the range of relationship values using different databases and use the lowest value (most conservative)

- Mutations are possible and should be accounted for in the LR calculations

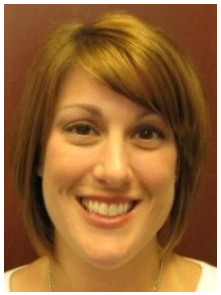
DNA Biometrics Project

Applied Genetics

Project Leader

Peter Vallone

Rapid PCR & Biometrics



Kristen Lewis O'Connor

Kinship Analysis

Recommended Reference

AABB (2010) Guidelines for mass fatality DNA identification operations. Available at http://www.aabb.org/programs/disasterresponse/Documents/aabbdnamassfatalityguidelines.pdf

Final version of this presentation available at:
http://www.cstl.nist.gov/strbase/NISTpub.htm

kristen.oconnor@nist.gov

Funding

NRC – Postdoctoral fellowship support for Kristen O'Connor
FBI – Application of DNA Typing as a Biometric Tool
NIJ – Forensic DNA Standards, Research, and Training

Tab B

Forensic DNA Experts
19069 Van Buren Blvd STE 114-498
Riverside, CA 92508
PH (951)275-2855
Fax (800)797-3013
www.ForensicDNAexperts.com
mmiller@ForensicDNAexperts.com

UNITED STATES DEPARTMENT OF JUSTICE

EXECUTIVE OFFICE FOR IMMIGRATION REVIEW

BOARD OF IMMIGRATION APPEALS

FALLS CHURCH, VIRGINIA

In the Matter of:

Nejat Ruzku, Beneficiary
of a Visa Petition filed by
Abdalla Ruzku, Petitioner

WAC-13-900-69825
A205-744-090

Visa Petition Proceedings

AFFIDAVIT OF

Dr. Monte Wayne Miller

I, Dr. Monte Miller of Forensic DNA Experts LLC, have been requested to opine on how the Board of Immigration Appeals should properly weigh kinship evidence based on the scientifically calculated statistical probability of such an event.

For my opinion I relied on: A) an October 17, 2014 U.S. Citizenship and Immigration Services Policy Memorandum on the subject of DNA evidence and sibling relationships; B) an October 1, 2014 letter from U.S. Department of Justice, Executive Office for Immigration Review, Board of Immigration Appeals regarding Nejat Ibrahim Ruzku; and C) an August 11, 2012 Universal Genetics DNA sibling test between Abdalla Ruzku Ibrahim and Nejat Ibrahim Ruzku. Other information was at my disposal, however no other information was used in the formation of my opinion.

In summary, The DNA test results between 2 siblings, by themselves, can serve as sufficient evidence that that it is more likely than not that two individuals are siblings. Sibling tests of this type are reliable and the risk of false positive is no greater than with parentage testing already accepted by the court. The inherent risk is in excluding true relatives.

Paternity and maternity are important issues in the courts on a regular basis and therefore a determination of what constitutes convincing and sufficient evidence became a necessity some time ago. Determining other types of kinship has been less vital and therefore, up until now, no standard has been set to guide the courts on other types of familial relationships. These relationships are routinely determined by DNA testing.

DNA evidence is strongly associated with its inherent ability to be given statistical weight. The statistical and mathematical methods used to evaluate all of the different kinships are reliable, well established, and uniform throughout the industry. What has not been established is the determination of the level of statistical probability that is convincing of a non-parentage relationship, in and of itself, and therefore sufficient without any further proof. There is always variability between individuals on what is convincing, however, this bar has already been set by the courts for parentage at 99.5% (the golden threshold) and we have no scientific basis for varying what the courts deem convincing based on specific kinships.

The mathematical methods used to establish parentage are no more or less reliable than any other kinship test. Parent-child relationships routinely furnish statistical probabilities higher than any other kinship and for this reason alone other relationships probabilities are subsequently thought of as less reliable. To be clear, this is not because scientists have less confidence in the numbers provided for other

types of kinship comparisons, but is instead a reflection that the probabilities are lower more often. Any probability that reaches this 99.5% threshold bears the same weight regardless of the relationship investigated. The difference is simply in how frequent a given kinship can reach this standard; parent-child kinship routinely lends itself to higher statistics. But, a high probability is convincing regardless of the kinship being tested and a low probability is less convincing based strictly on the probability given. A low probability does not mean that the claimed relationship does not exist, only that the relationship is less statistically likely. In summary of this issue, whether or not the kinship is convincing is solely determined by the statistical likelihood generated and is not derived from the relationship investigated. Therefore, a high probability from a sibling-only DNA test is equivalent in every aspect to the same probability from a parent-child DNA test.

The question of what probability should be accepted for sibling-only testing has already been answered, is already in place, and is already in use, namely the golden threshold of 99.5%. The lack of scientific consensus on a convincing statistical probability for other kinship tests is in no way due to the trustworthiness of the statistical numbers generated but instead is a reflection of the fact that, until now, the courts have narrowly applied this standard to the single kinship of parent-child when the scientific and mathematical models used for kinship probability extend equally to all relationship testing, including sibling tests.

The golden threshold used by the courts of 99.5% probability was put in place to guide the system with a useful number that could be relied upon to establish parent-child relationships and we have no scientific basis for altering this number for other well-accepted DNA based kinship testing. The testing is the same and the statistics generated carry the same trustworthiness; however, it is the courts who have not yet set the bar as to what is to be the legal standard of convincing and sufficient in regards to other kinships investigated.

It should also be noted that the inherent weakness in non-parental DNA kinship testing lies in excluding a true relative, not in including one that does not belong. In other words, there is an increased risk of false negatives, not false positives. Therefore, high probabilities can be trusted and low probabilities should be viewed as merely non-confirming; low probabilities should not be viewed as confirming non-

relationship status except in cases of direct parentage testing. Only parentage testing can unequivocally provide compelling proof of non-relationship status.

The longstanding policy of accepting the 99.5% kinship probability as a proper standard to determine if a given relationship should be considered convincing is a reasonable threshold. On the surface that standard seems to meet the desired quality of difficult but attainable. Other kinship testing which can be said to meet or exceed this same 99.5% probability should likewise be designated as meeting the established criterion and be convincing as well. The mathematical models have already taken the specific familial relationship into consideration and the probability generated should be considered at face value as to whether or not it confirms the kinship investigated. In other words, look at the probability without regard to the relationship tested.

These matters never happen in a vacuum and other issues should be weighed if possible, but, if 2 people are not related the chances of an individual slipping past this rigorous test is extremely low. A 99.5% probability is equivalent to having 1 spin on a roulette table and having to hit 1 specific number with 200 spaces on the wheel. A high but reasonable bar for scientists, and therefore decision-makers, to rely on as both sufficient and convincing regardless of the relationship investigated. This appears to also satisfy the much lower standard of preponderance of the evidence as any probability greater than 50% would be more likely than not. As mentioned above, the weakness lies in excluding those who are true relatives and not in erroneously including non-kinfolk.

All scientific tests, where probability can be applied, should be screened on an individual basis to ensure that they are a good match for the criteria and situation outlined here. The specific instance of calculating a sibling kinship instead of a parent-child kinship is so similar in scope and in nature as to make application of this same standard reasonable. In fact, I could find no good argument to deny application of the same policy used on the parent-child test to the sibling test, namely that a 99.5% probability is sufficient and convincing, in and of itself, to establish as fact the relationship tested.

The probability of sibling kinship between Abdalla Ruzku Ibrahim and Nejat Ibrahim Ruzku exceeds the 99.5% probability mark and therefore exceeds the long standing

policy and standard and should subsequently be scientifically considered convincing and sufficient and therefore without need for further proof. The kinship tested here does not affect the scientific probability and therefore the veracity of the investigation. The probability given in this case is greater than 99.8%, and this means that we have 99.8% confidence that this information can be relied upon as accurate; far greater than the 50% necessary for preponderance of the evidence and much more likely than not.

Proper science does not disregard information. The kinship test cannot and should not be disregarded when a test of this nature falls below the golden threshold, but instead should be considered in the context of the situation. Simply because an item of evidence is not convincing and sufficient, in and of itself, automatically, to all people, does not render it weightless; the individual fact is merely insufficient to cause the issue to automatically pass inspection based on this criterion alone. For this reason the 99.5% probability should not be used as a pass-or-fail test, but instead should be used as scientists intend, namely to give proper weight to an investigative fact.

Any probability approaching 99.5% far exceeds a 50% chance that a given relationship exists. Many kinship tests will be unable to meet the golden threshold, however, this does not suggest that the individuals are not related as specified. What it means is that, due to the inheritance pattern of DNA, testing was unable to unequivocally confirm this relationship. Except in the case of parentage testing, statistical likelihoods run from the spectrum of "probable" to "looks random". "Probable" is not synonymous with confirmed, and "looks random" is not synonymous with false. A high probability means that the test should be given strong weight and a low probability means that the test may not be given much weight. The given probability of an individual test is merely a lens used to view the fact at hand and give it weight, nothing more and nothing less. Therefore no difference exists between a high probability parentage test and a high probability sibling test.

Most relationship testing laboratories routinely include a statement similar to "Person A is 258 times more likely to be related as a full-sibling to Person B than to be a random non-related person". Statements like these can assist the court in decision-making even when they fail to meet the golden threshold of 99.5%

probability. To ignore evidence falling below this standard simply because it alone cannot provide finality would be short-sighted and reflect a misunderstanding of the basis of scientific probability. How many cases in our courts today are definitively adjudicated by a single fact?

It is therefore my professional opinion that:

- No scientific basis exists that causes the probabilities of parent-child kinship tests to be more reliable than sibling-only kinship tests for the purpose of establishing relationships (though some low probabilities may be spurious).
- The long standing policy of accepting 99.5% probability for parent-child DNA relationship testing should be extended to cover sibling-only DNA Testing. Any sibling-only probability that is determined to be greater than 99.5% should be considered convincing and sufficient evidence that it is more likely than not that the sibling relationship exists without need for further proof.
- Relationship investigations falling below 99.5% probability should be considered in context and in the light of the other facts surrounding a case as well as weighed considering the relative strength of the given kinship probabilities. Care should be given to low probabilities when the DNA testing is not a parent-child test.

Dr. Monte Miller, Director

Nov 12, 2014
Date

Jurat

State of California

County of Riverside

Subscribed and sworn to (or affirmed) before me on this 12 day of November,

20 14 by Monte Miller,

proved to me on the basis of satisfactory evidence to be the person(s) who appeared before me.

Signature _____ (Notary seal)

J. MOORE
COMM. #2012690
NOTARY PUBLIC - CALIFORNIA
RIVERSIDE COUNTY
My Comm. Expires March 16, 2017

OPTIONAL INFORMATION

DESCRIPTION OF THE ATTACHED DOCUMENT

Affidavit
(Title or description of attached document)

(Title or description of attached document continued)

Number of Pages 6e Document Date 11/12/14

(Additional information)

Example of an oath or affirmation to be asked by the notary prior to signing: "Do you swear or affirm that the statements made in the attached document are true to the best of your knowledge?" (The affiant must reply affirmatively.)

INSTRUCTIONS FOR COMPLETING THIS FORM

The wording of all Jurats completed in California after January 1, 2008 must be in the form as set forth within this Jurat. There are no exceptions. If a Jurat to be completed does not follow this form, the notary must correct the verbiage by using a jurat stamp containing the correct wording or attaching a separate jurat form such as this one which does contain proper wording. In addition, the notary must require an oath or affirmation from the document signer regarding the truthfulness of the contents of the document. The document must be signed AFTER the oath or affirmation. If the document was previously signed, it must be re-signed in front of the notary public during the jurat process.

- State and County information must be the State and County where the document signer(s) personally appeared before the notary public.
- Date of notarization must be the date that the signer(s) personally appeared which must also be the same date the jurat process is completed.
- Print the name(s) of document signer(s) who personally appear at the time of notarization.
- Signature of the notary public must match the signature on file with the office of the county clerk.
- The notary seal impression must be clear and photographically reproducible. Impression must not cover text or lines. If seal impression smudges, re-seal if a sufficient area permits, otherwise complete a different jurat form.
  - Additional information is not required but could help to ensure this jurat is not misused or attached to a different document.
  - Indicate title or type of attached document, number of pages and date.
- Securely attach this document to the signed document

Tab C

Evaluating the effect of additional forensic loci on likelihood ratio values for complex kinship analysis

Kristen Lewis O'Connor, Erica Butts, Carolyn R. Hill, John M. Butler, and Peter M. Vallone

Biochemical Science Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8312

Disclaimer: This work was funded by the FBI Criminal Justice Information Services Division and by the National Institute of Justice through an interagency agreement 2008-DN-R-121 with the NIST Office of Law Enforcement Standards. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. Certain commercial equipment, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

Corresponding Author:
Kristen Lewis O'Connor kristen.oconnor@nist.gov Phone: (301) 975-5205 Fax: (301) 975-8505

A copy of the presentation given is available at:
http://www.cstl.nist.gov/strbase/pub_pres/OConnor-Promega2010-Additional-Loci-Kinship.pdf

Abstract

A study was conducted to evaluate the effect of adding autosomal STRs to the 13 U.S. core loci for complex kinship analysis. Additional genetic information may increase likelihood ratio values for true relationships in a pedigree, while reducing the chance of identifying false relationships. The clear discrimination of true versus false relationships is important for complex kinship cases, such as familial searches, paternity testing, missing persons work, and immigration testing. Allele frequency data was available for 46 forensic autosomal STR loci from U.S. Caucasian, Hispanic, and African American population samples. These loci were from commercial amplification kits (Identifiler®, PowerPlex® 16 System, and PowerPlex® ESI/ESX 17 Systems) and an in-house assay (NIST 26plex). With the large number of loci, various sets of loci were selected to simulate genotype pairs for related and unrelated individuals. The relationships evaluated were parent-offspring, full siblings, and half siblings. The expected likelihood ratio distributions were compared across sets of 13, 20, and 40 STR loci to evaluate the discrimination power gained by adding markers to the core U.S. and European forensic loci. Increasing the number of STR loci resulted in increased discrimination of true and unrelated pairs, although the results were more dramatic for parent-offspring and full siblings than for half siblings. With a likelihood ratio threshold of one, the use of 40 STR loci produced robust discrimination of parent-offspring (no false inclusions or exclusions) and full sibling pairs (less than 1% false inclusions and exclusions). However, half siblings demonstrated overlapping likelihood ratio distributions with 40 loci, resulting in false inclusion and exclusion rates of 5-7%. Suggestions are provided to further improve the discrimination power of autosomal STR loci for kinship determination.

Introduction

Since their selection in November 1997, genetic information from a core set of 13 short tandem repeat (STR) loci have been required for upload of DNA profiles to the national DNA database (Butler 2006). The 13 U.S. core loci used by the National DNA Index System are: CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11. These loci are commonly referred to as the 13 Combined DNA Index System (CODIS) loci. Forensic laboratories analyze the CODIS loci to identify the perpetrator of a

crime by directly matching an evidence profile and a suspect profile or by searching an evidence profile against an offender database. If a match is found between two DNA profiles, then the rarity of the genotype is determined by calculating the random match probability (RMP) (Butler 2009). To determine how much more likely it is that the evidence sample originated from the suspect than from a random individual, a likelihood ratio (LR) is used. In the forensic case, a LR is a comparison of the probabilities of the evidence under two alternative hypotheses. The numerator (prosecution's) hypothesis is that the DNA from the crime scene came from the suspect, while the denominator (defense's) hypothesis is that the DNA originated from an unrelated, random individual in the population. Thus, the LR is calculated as 1/RMP (Butler 2009).

In addition to individual identification for forensic purposes, the U.S. core loci are routinely used to evaluate the relatedness between individuals. Instead of identifying a perfect match between two DNA profiles, kinship analysis assesses any shared genetic information between putative relatives under alternative hypotheses or by estimating an unknown relationship (Weir et al. 2006). Applications for kinship analysis include civil paternity disputes, criminal paternity cases, immigration applications, missing persons cases, disaster victim identifications, and familial searches of offender databases.

In cases of alleged paternity, a putative father-child relationship can be tested using information obtained from DNA typing and statistical analysis. Typically, DNA samples from the mother, child, and alleged father(s) are tested with 15-20 forensic STR markers. If the alleged father is indeed the biological father, then at each tested marker the child will share one allele with the mother and one allele with the father (barring mutations). These shared alleles are identical by descent (IBD) since the allele originated from a common ancestor in the father-child and mother-child pairs. For pairs of relatives, the probabilities of sharing alleles IBD are known (Weir et al. 2006, Table 1). Using the genetic data, population allele frequencies, hypothesized relationships, and associated IBD probabilities, an LR calculation will evaluate how strongly the genetic data support the purported relationship (Evett and Weir 1998, Buckleton et al. 2005, Weir et al. 2006, Gjertson et al. 2007). In the case of a paternity test, a LR is commonly referred to as a paternity index (AABB 2009).

Figure 1 illustrates how the forensic core competency is being expanded from direct matching with 13 STR loci to applications of complex kinship analysis and familial searching for which indirect matches are made and uncertainty increases. As opposed to forensic identifications, no standard set of loci exists for kinship testing applications. However, for many kinship analyses, commercial forensic PCR kits are used to type 15 STR loci—13 CODIS loci plus D2S1338/D19S433 or Penta D/Penta E, depending on the manufacturer. In most paternity trio cases, 15 forensic loci are sufficient to provide positive proof of paternity (Poetsch et al. 2006). However, in cases where a relationship cannot be confirmed or refuted with statistical certainty with 13-15 STR loci (e.g., deficient paternity, pairs of relatives, or possible mutation events), additional STR loci or alternative DNA markers may be required (Wenk et al. 2003, Poetsch et al. 2006, Betz et al. 2007, AABB 2009). With familial searches and disaster victim identifications, a one-to-many search in a database of 13-15 STR profiles will produce a large number of fortuitous matches for close biological relatives, such as parent-offspring and full siblings (Brenner and Weir 2003, Bieber et al. 2006, Reid et al. 2008, Curran and Buckleton 2008). Furthermore, the number of false positive matches increases when the alleged genetic relationship becomes more distant due to the reduced amount of IBD allele sharing. However, tests of more distant relationships, such as for half siblings or uncle-nephew, may be of interest for inheritance disputes and immigration testing, for example. Additionally, the use of additional loci may decrease the number of false positive relationships and provide more support for true relationships than the 13-15 forensic STR loci currently used for familial searches in the U.S.

In April 2005, the European Network of Forensic Science Institutes selected five STR loci (D12S391, D1S1656, D2S441, D10S1248, and D22S1045) to add to their existing European Standard Set of seven STRs (TH01, vWA, FGA, D8S1179, D18S51, D21S11, and D3S1358) (Gill et al. 2006a, 2006b). Germany and several other European countries test the highly polymorphic locus SE33 (Hill et al. in press). To provide maximum overlap of core markers among the European community, STR kit manufacturers have produced multiplex PCR kits that include the 12 ESS loci, D16S539, D2S1338, D19S433, and SE33. In addition to the expanded set of European loci, 26 STR loci have been characterized by NIST for use in forensic typing, and an NIST in-house multiplex PCR assay has been developed with 25 of these loci (Hill et al. 2008, Hill et al. 2009). Combining the loci from the expanded ESS, the commercial U.S. and European kits, and the NIST-developed assay, 46 unique STR loci are available to the forensic and kinship testing communities. For the major U.S. populations, allele frequency data have been collected for these 46 loci (Butler et al. 2003, Hill et al. 2008, Hill et al in press, unpublished data). For challenging kinship

cases, the large set of available genetic data can be used to determine whether additional loci, beyond the 13-15 STR loci currently tested, can improve the discrimination of true relatives from unrelated individuals by reducing false inclusions and exclusions.

Ideally for use in forensic and kinship analyses, genetic markers on the same chromosome should be more than 50 centimorgans (cM) apart in genetic distance (approximately 50 Mb in physical distance). This distance ensures full recombination (recombination frequency = 0.50) and thus independent inheritance of alleles at multiple markers. The new ESS locus D12S391 occurs on the short arm of chromosome 12 and is only 6.3 megabases (Mb) from the established vWA STR locus that is widely used (Phillips et al. in press). Recent studies have detected significant linkage disequilibrium (O'Connor et al. in press) and linkage (recombination frequency estimate of 0.108, Budowle et al. in press) between the D12S391 and vWA loci. Consequently, the single-locus genotype probabilities for D12S391 and vWA should not be multiplied to determine the match probability of an autosomal STR profile when unrelated or related individuals are involved. More research is needed to determine the effect of linkage disequilibrium and linkage on match probability calculations. Until then it is convenient to simply exclude D12S391 from consideration as an additional marker to improve complex kinship analysis since vWA has been a core U.S. and European locus for years.

Materials and Methods

Allele frequencies for 46 STR loci were available from approximately 600 NIST U.S. population samples including Caucasians, Hispanics, and African Americans (Butler et al. 2003, Hill et al. 2008, Hill et al. in press, unpublished data). The NIST allele frequency data are available at http://www.cstl.nist.gov/biotech/strbase/NISTpop.htm. For the current study, the following marker systems were used:

- AmpFlSTR® Identifiler® (Applied Biosystems, Foster City, California): CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, and D19S433 (Butler et al. 2003)
- PowerPlex® 16 System (Promega Corporation, Madison, Wisconsin): CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, and Penta E (unpublished data)
- PowerPlex® ESI 17 System (Promega): D1S1656, D2S441, D2S1338, D3S1358, D8S1179, D10S1248, D12S391, D16S539, D18S51, D19S433, D21S11, D22S1045, FGA, TH01, vWA, SE33 (Hill et al. in press)
- NIST 26 miniSTR loci: D1GATA113, D1S1627, D1S1677, D2S441, D2S1776, D3S3053, D3S4529, D4S2364, D4S2408, D5S2500, D6S474, D6S1017, D8S1115, D9S1122, D9S2157, D10S1248, D10S1435, D11S4463, D12ATA63, D14S1434, D17S974, D17S1301, D18S853, D20S482, D20S1082, D22S1045 (Hill et al. 2008)

Allele frequencies of the three major U.S. population groups were kindly imported into DNA-VIEW™ v. 29.23 (Charles Brenner, Oakland, California) by Dr. Charles Brenner. In the Kinship Simulation module of DNA-VIEW™, pairs of genotypes were simulated using population-specific allele frequencies for sets of 13, 20, and 40 STR loci (Table 2). Genotype pairs were simulated as true parent-offspring, true full siblings, true half siblings, and unrelated individuals. For each simulation, a LR was calculated that compared the hypotheses of a specific relationship versus no relationship. The true relative genotype pairs were evaluated with a numerator hypothesis corresponding to their actual relationship (e.g., for a simulated parent-offspring pair, the LR evaluated the probabilities of the genotypes given a parent-offspring relationship versus no relationship). The unrelated genotype pairs were evaluated in separate simulation modules using the hypotheses of parent-offspring, full siblings, or half siblings versus no relationship. Although no mutations were created during genotype simulations, LR calculations accounted for possible mutation events as per AABB recommendations (AABB 2009). For each of the three relationship scenarios, 1000 independent genotype simulations and LR calculations were performed. However, due to the large range of LR values for full sibling comparisons and the resulting ragged LR distributions, 5000 simulations were necessary to increase the density of data points and to produce smoother curves, which allowed for a more direct visual comparison between test schemes (data not shown). For each set of simulations, log LR values were graphed to produce expected LR distributions in Excel®.

Results

To assess the discrimination power gained by increasing the number of STR loci for kinship analysis, pairs of genotypes with defined relationships (parent-offspring, full siblings, half siblings, and unrelated persons) were simulated using NIST U.S. Caucasian, Hispanic, and African American allele frequencies for 13, 20, and 40 STR loci (Table 2). For each simulation, LR values were calculated to compare the probabilities of the genotype pair under the hypotheses of a specific relationship versus no relationship. The distributions of LR values for each type of true relative pair and marker set were graphed separately for Caucasians, Hispanics, and African Americans (Figures 2-4, respectively). As the number of loci increased, the LR distributions shifted to the right for true relatives, indicating larger LR values that give more strength to the hypothesized relationship. The increase in LR values was greater for true parent-offspring and full sibling relationships than for half sibling relationships. All LR values were greater than one for true parent-offspring relatives for 13, 20, and 40 STR loci, while the LR distributions shifted below one for true full and half siblings, indicating a region of LR values that supported the hypothesis of no relationship. Additionally, the breadth of the true relative distributions increased with the number of loci typed, demonstrating a larger range of possible LR values when additional loci were typed.

By the definition of a LR, any value greater than one supports the numerator hypothesis of a specific relationship. Conversely, any value less than one supports the denominator hypothesis of no relationship. By defining a specific LR threshold, the proportion of expected false inclusions (unrelated persons that appear related) and expected false exclusions (true relative pairs that appear unrelated) can be estimated from the simulation data. Additionally, the distributions of LR values can be graphed for unrelated persons and true relatives (Figure 5). Graphing the distributions allows for a convenient evaluation of whether increasing the number of STR loci, beyond the core 13 STRs used for forensic typing in the U.S., can produce robust discrimination of true relative pairs and unrelated persons. If the unrelated and related distributions overlap, then a range of LR values exists where false positive or false negatives can occur for a given set of loci, allele frequencies, and relationship scenario. For kinship analysis, one hopes to achieve complete separation of the distributions to reduce the chance of falsely identifying individuals as related or unrelated.

Discrimination among U.S. populations

To compare the power of various numbers of loci to discriminate true relatives from unrelated persons, distributions of LR values were graphed for related and unrelated genotype pairs under each relatedness scenario and allele frequency dataset. Figures 6-8 provide the simulation results using 13 STR loci for Caucasians, Hispanics, and African Americans, respectively. Similarly, LR distributions for 20 loci are shown in Figures 9-11, and results for 40 loci are shown in Figures 12-14. For each relationship scenario evaluated with 13, 20, or 40 loci, there was little difference between LR distributions with allele frequencies of the three population groups. In particular, the median LR values and breadth of the distributions were similar for specific relationship scenarios among the three populations. However, median LR values were slightly larger for African American and Hispanic genotypes than for Caucasian genotypes for a specific relationship scenario. For example, for true parent-offspring genotypes with 13 STRs, the median LR values were 10,000 for Caucasians; 15,000 for Hispanics; and 19,000 for African Americans (Figures 6-8, respectively). In the case of evaluating true full siblings with 20 STRs, the median LR values were 680,000 for Caucasians; 800,000 for Hispanics; and 1,750,000 for African Americans (Figures 9-11, respectively). For true half sibling genotypes with 40 STRs, the median LR values were 300 for Caucasians, 310 for Hispanics, and 610 for African Americans (Figures 12-14, respectively). Although differences were observed for median LR values between population groups, the proportion of false inclusions and exclusions remained similar between the groups (Table 3).

Since the LR distributions varied more with the number of loci and type of relationship than between population groups, data analysis efforts were focused on evaluating the differences between expected LRs when 13, 20, and 40 STR loci were used to assess specific relationships. The Caucasian dataset will be used as an example to illustrate the power of increasing the number of STR loci to discriminate true relatives from unrelated individuals. Results from analyses using Hispanic and African American genotypes are provided in the accompanying figures and Table 3.

Discrimination potential of 13 STR loci using Caucasian data

For a given set of loci, LR distributions exhibited increasing overlap between true relative and unrelated pairs when the hypothesized relationship became more distant. Using Caucasian 13-locus genotypes, LR values did not overlap between unrelated and related parent-offspring comparisons (Figure 6). However, the distributions were nearly overlapping, indicating a chance for false inclusions to occur. Note that false exclusions would only occur for a parent-offspring pair if a meiotic mutation was present and not accounted for in the likelihood ratio algorithm. No false positive and false negative parent-offspring relationships were indicated with 13 loci since LR values for unrelated pairs were less than one while values for true relatives were greater than one (Figure 6, Table 3). However, unrelated and related full and half sibling comparisons produced LR values that overlapped in the region defined by 0.01 to 1,000 (Figure 6). For full sibling comparisons, this area of uncertainty produced false positives and false negatives with a frequency of 0.027 and 0.033, respectively (Table 3). More extreme than the full sibling scenario, LR distributions overlapped for unrelated and related half sibling comparisons, producing false positives and false negatives with a frequency of 0.155 and 0.173, respectively (Figure 6, Table 3). The use of 13 locus profiles produced false inclusions and exclusions for pairs of full and half sibling relatives.

Discrimination potential of 20 STR loci using Caucasian data

The performance of 20 STR loci was investigated to determine if additional loci could improve the discrimination of true relatives from unrelated individuals. As with 13 loci, no overlap was observed between LR distributions of unrelated and related parent-offspring comparisons (Figure 9). Consequently, no false positive and false negative parent-offspring relationships were indicated with 20 loci (Table 3). As compared to results for full and half sibling hypotheses with 13-locus genotypes, the areas of overlap diminished with 20 loci, although the ranges of LRs defining the overlapping regions remained the same for full and half sibling comparisons (Figure 9). For full sibling comparisons, the reduced area of uncertainty produced false exclusions and false inclusions with a frequency of 0.006 and 0.008, respectively (Table 3). For half sibling comparisons, the reduced area of uncertainty produced false positive and false negative relatives with a frequency of 0.075 and 0.104, respectively (Table 3). The discrimination of parent-offspring, full siblings, and half siblings was improved with 20 STR loci due to increasing LR values of true relative pairs and decreasing LR values of unrelated individuals.

Discrimination potential of 40 STR loci using Caucasian data

An expanded set of 40 STR loci was used to evaluate the power to identify true relatives from unrelated individuals. As with the smaller numbers of loci, no overlap was observed between LR distributions of unrelated and related parent-offspring pairs (Figure 12). All unrelated comparisons produced LRs equal to zero when evaluated as parent-offspring with 40 loci. No false positive and false negative parent-offspring relationships were indicated with 40 loci (Table 3). When a full sibling relationship was evaluated with 40 loci, the area of overlap between unrelated and related distributions was reduced when compared with results for 13 or 20 loci. Moreover, the region of overlap decreased to LR values between 0.1 and 10 (Figure 12). For the full sibling hypothesis, the reduced area of uncertainty produced false positive and false negative relatives with a frequency of 0.001 and 0.002, respectively (Table 3). For half sibling comparisons, the reduced area of uncertainty was defined by the same LR range as with 13 and 20 loci (Figure 12) and produced false positives and false negatives with a frequency of 0.051 and 0.066, respectively (Table 3). The use of 40 STR loci provided robust discrimination of parent-offspring and full sibling relatives from unrelated individuals. Although 40 loci reduced the frequency of false inclusions and exclusions for half sibling comparisons to ~5-7%, discrimination was not markedly improved with a LR threshold of one when compared to 20 loci.

Discussion

Challenges of kinship analysis with limited numbers of STR loci

This study confirmed previous work that the 13 forensic STR loci are not sufficient to definitively discriminate between pairs of close relatives and unrelated individuals and may lead to false inclusions and exclusions (Wenk et al. 2003, Poetsch et al. 2006, Allen et al. 2007, Pu and Linacre 2008, González-Andrade et al. 2009). Although all true parent-offspring comparisons produced LRs that supported relatedness, a few false positives were observed

when unrelated pairs had LR values that favored a parent-offspring relationship. The chance for false inclusions (albeit small) illustrates the difficulty of determining paternity when deficient (motherless and reverse paternity) cases are attempted with a small set of STR loci (Wenk et al. 2003, Poetsch et al. 2006, González-Andrade et al. 2009). For full and half sibling cases, the chances of false inclusions and exclusions are even greater than for parent-offspring. True siblings and more distant relatives may not share an allele at every locus, potentially resulting in LR values that favor no relationship. Conversely, unrelated individuals can share alleles by chance to generate positive LR values that support relatedness.

In addition to the challenge of identifying close relatives through a pairwise (one-to-one) comparison with 13 loci, the task is more difficult when relatives are compared using a database (one-to-many) search (Figure 1). Databases are used to identify relatives using familial searches (Bieber et al. 2006, Reid et al. 2008, Curran and Buckleton 2008), missing persons cases (Gornik et al. 2002), and mass disaster identifications (Brenner and Weir 2003). When comparing one 13-15 locus profile against many profiles, the chance increases that two unrelated individuals will appear related due to fortuitous allele sharing. This will be especially problematic when testing for relationships in which allele sharing is not required at every locus (e.g., full siblings and more distant relationships).

Increased discrimination power with additional STR loci

To increase the discrimination power of identifying close relatives with 13 forensic loci, additional loci may be tested. This study shows that increasing the number of typed autosomal STR loci to 40 resulted in complete discrimination of true parent-offspring relatives from unrelated individuals. In fact, no unrelated pairs had a LR value greater than zero. Thus, when evaluating 40 markers for a parent-offspring relationship, there were no instances in which pairs of unrelated individuals shared an allele at every locus by chance. In full sibling comparisons, 40 STR loci produced robust discrimination of true relatives from unrelated individuals with false inclusion and exclusion rates of ~0.1%. Compared to typing with 13 or 20 STR loci, the use of 40 loci improved the discrimination of true half siblings and unrelated individuals. However, the false inclusion and exclusion rate of 5-7% reflected the lower power of this set of loci to identify half siblings due to the reduced amount of IBD allele sharing in half siblings. Additional STR loci would be required for further discrimination of half siblings and more distant relatives (Blouin 2003, Wilkening et al. 2006). In cases with one-to-many comparisons, the use of 40 STR loci could reduce fortuitous allele sharing between pairs of unrelated individuals, particularly for parent-offspring and full siblings where IBD allele sharing is higher than for more distant relationships (Weir et al. 2006).

When using additional loci for relatedness testing, one caveat to consider is that there are more opportunities for meiotic mutations to occur in true relatives. The recommendation currently employed—that at least two STR mutations are required to exclude a parent-offspring relationship (AABB 2009)—would likely need to be adjusted to allow for more potential mutations between true relatives. In this study, genotypes were not simulated with mutation events but possible mutations were accounted for in the LR algorithms. Further research is needed to evaluate the occurrence and impact of multiple mutations when genotyping large numbers of STR loci for kinship determination. Alternatively, SNP loci could be added to increase the discrimination power of forensic STR loci (Børsting and Morling 2010) due to the 5-fold lower mutation rate of SNPs compared to STR loci (Butler 2009).

Increased discrimination power with lineage markers

Unlike diploid autosomal markers, which are independently assorted during meiosis, loci on the Y chromosome and on the mitochondrial genome (mtDNA) are transmitted unchanged (barring mutation) from generation to generation. Y-chromosome haplotypes and mitochondrial sequences of hypervariable regions I and II (HVI/II) have proven highly informative for identifying related individuals, even distant relatives (Ginther et al. 1992, Junge et al. 2006, Butler et al. 2007). Y-haplotypes identify all relationships between males who are linked only by males. If a paternal relative exists in a database, a Y-STR match is always made (barring mutation). MtDNA sequences identify all relationships between any persons, male or female, who are linked only by females, including mother-son, mother-daughter, full siblings, maternal half siblings, maternal uncle-nephew, male or female cousins related entirely through mothers, and so on. If a mitochondrial relative exists in a database, an mtDNA sequence match is always made (barring mutation). The power to detect paternal or mitochondrial relatives depends on the population frequency of the shared Y-STR haplotype or mtDNA sequence. Population studies have indicated that 96% of Y-haplotypes defined by 17 STR loci are unique and all Y-haplotypes are individually rare (frequency of less than 1%) (Butler et al. 2007). Similarly, population studies have shown that approximately 70% of individuals have private

mtDNA haplotypes in samples from the three major U.S. populations (Budowle et al. 1999). In sampled Caucasians, one mtDNA HVI/II haplotype has a frequency of 7%; all other haplotypes are much more rare (Parsons and Coble 2001, Allard et al. 2002). All mtDNA haplotypes are individually rare in African and Hispanic samples (Allard et al. 2005, Allard et al. 2006). MtDNA sequences are particularly useful for identifying maternal half siblings or other relationships for which persons do not share a Y chromosome.

Increased discrimination power with additional relatives and non-genetic information

Before attempting a kinship analysis, it is necessary to determine the discrimination power of a particular set of loci. As was performed in this study, simulation experiments are valuable tools to evaluate the expected range of LR values for a defined set of loci, relationships, populations, and statistical algorithms (Blouin 2003). Depending on the simulation results, one could determine if additional loci, family members, or non-genetic data are needed to meet a predefined level of certainty. In some cases it is not always possible to type additional loci to increase the discrimination power due to limited resources or the type of relationship in question (e.g., distant relatives such as second cousins). The use of additional known or putative relatives can increase the power of correctly inferring relatedness (Nothnagel et al. 2010). Examining multiple relatives, such as several full siblings, can identify Mendelian discrepancies that may be missed in a pairwise analysis (Browning and Thompson 1999). Alternatively, non-genetic information may be available to improve confidence in a test result. Metadata, such as age, sex, ethnicity, and legal documentation, can be helpful in familial searching cases and immigration testing (Sheehan and Egeland 2007). When non-genetic information is used, Bayesian statistics are required to incorporate prior probabilities for the non-genetic with likelihood ratios for the genetic data (Shoemaker et al. 1999).

Defining thresholds for discrimination

Based on the definition of a LR, a LR threshold was selected in this study to define a positive or negative relationship as being greater than or less than one, respectively. However, in practice, different LR thresholds may be required to achieve particular specificity and sensitivity thresholds (Gaytmenn et al. 2002). If higher LRs are used, there will be a trade-off between decreasing false inclusions and increasing false exclusions. A balance must be sought and may vary across particular applications (e.g., paternity testing, immigration testing, familial searching, etc.). As opposed to a discrete LR threshold, a grey zone approach can be helpful for defining a likelihood ratio range that does not eliminate uncertainty about the relationship status (Coste and Pouchot 2003, Giroti et al. 2007).

References

AABB. Guidance for standards for relationship testing laboratories. 7th rev. ed. Bethesda: AABB, 2009.

Allard M, Miller K, Wilson M, Monson K, Budowle B. Characterization of the Caucasian haplogroups present in the SWGDAM forensic mtDNA dataset for 1771 human control region sequences. J Forensic Sci 2002;47: 1215-23.

Allard M, Polanskey D, Miller K, Wilson M, Monson K, Budowle B. Characterization of human control region sequences of the African American SWGDAM forensic mtDNA data set. Forensic Sci Int 2005;148:169-79.

Allard M, Polanskey D, Wilson M, Monson K, Budowle B. Evaluation of variation in control region sequences for Hispanic individuals in the SWGDAM mtDNA dataset. J Forensic Sci 2006;51: 566-73.

Allen RW, Fu J, Reid TM, Baird M. Considerations for the interpretation of STR results in cases of questioned half-sibship. Transfusion 2007 Mar;47(3):515-9.

Betz T, Immel U-D, Kleiber M, Klintscher M. "Paterniplex", a highly discriminative decaplex STR multiplex tailored for investigating special problems in paternity testing. Electrophoresis 2007;28:3868-74.

Bieber F, Brenner C, Lazer D. Finding criminals through DNA of their relatives. Science 2006;312:1315-6.

Blouin MS. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. Trends Ecol Evol 2003;18:503-11.

Børsting C, Morling N. Mutations and/or close relatives? Six case work examples where 49 autosomal SNPs were used as supplementary markers. Forensic Sci Int Genet doi:10.1016/j.fsigen.2010.02.007.

Brenner C, Weir B. Issues and strategies in the DNA identification of World Trade Center victims. Theor Pop Biol 2003;63:173-8.

Browning S, Thompson EA. Interference in the analysis of genetic marker data. Am J Hum Genet Suppl 1999;65:A244.

Buckleton J, Triggs C, Walsh S. Forensic DNA evidence interpretation. Boca Raton: CRC Press, 2005.

Budowle B, Wilson M, DiZinno J, Stauffer C, Fasano M, Holland M, et al. Mitochondrial DNA regions HVI and HVII population data. Forensic Sci Int 1999;103:23-35.

Budowle B, Ge J, Chakraborty R, Eisenberg AJ, Green R, Mulero J, et al. Population genetic analyses of the NGM STR loci. Int J Legal Med doi:10.1007/s00414-010-0516-7.

Butler, JM. Genetics and genomics of core STR loci used in human identity testing. J Forensic Sci 2006;51(2):253-65.

Butler, JM. Fundamentals of forensic DNA typing. San Diego: Elsevier Academic Press, 2009.

Butler JM, Schoske R, Vallone PM, Redman JW, Kline MC. Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. J Forensic Sci 2003;48:908-11.

Butler J, Hill C, Decker A, Kline M, Reid T, Vallone P. New autosomal and Y-chromosome STR loci: characterization and potential uses. In: Proc 18th Int Symp Hum Ident. Madison WI: Promega Corporation, 2007.

Coste J, Pouchot J. A grey zone for quantitative diagnostic and screening tests. Int J Epidemiol 2003 Apr;32(2):304-13.

Curran J, Buckleton J. Effectiveness of familial searches. Sci Justice 2008;48: 164-7.

Evett IW, Weir BS. Interpreting DNA evidence. Sunderland: Sinauer, 1998.

Gaytmenn R, Hildebrand DP, Sweet D, Pretty IA. Determination of the sensitivity and specificity of sibship calculations using AmpFISTR Profiler Plus, Int J Legal Med 2002;116:161-4.

Gill P, Fereday L, Morling N, Schneider PM. The evolution of DNA databases—recommendations for new European STR loci. Forensic Sci Int 2006a;156:242-4.

Gill P, Fereday L, Morling N, Schneider PM. New multiplexes for Europe—amendments and clarification of strategic development. Forensic Sci Int 2006b;163:155-7.

Ginther C, Issel-Tarver L, King M. Identifying individuals by sequencing mitochondrial DNA from teeth. Nat Genet 1992;2:135-8.

Giroti RI, Verma S, Singh K, Malik R, Talwar I. A grey zone approach for evaluation of 15 short tandem repeat loci in sibship analysis: a pilot study in Indian subjects. J Forensic Leg Med 2007 Jul;14(5):261-5.

Gjertson DW, Brenner CH, Baur MP, Carracedo A, Guidet F, Luque JA, et al. ISFG: recommendations on biostatistics in paternity testing. Forensic Sci Int Genet 2007 Dec;1(3-4):223-31.

González-Andrade F, Sánchez D, Penacino G, Martínez Jarreta B. Two fathers for the same child: a deficient paternity case of false inclusion with autosomic STRs. Forensic Sci Int Genet 2009 Mar;3(2):138-40.

Gornik I, Marcikic M, Kubat M, Primorac D, Lauc G. The identification of war victims by reverse paternity is associated with significant risks of false inclusion. Int J Legal Med 2002 Oct;116(5):255-7.

Hill CR, Kline MC, Coble MD, Butler JM. Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. J Forensic Sci 2008;53(1):73-80.

Hill CR, Butler JM, Vallone PM. A 26plex autosomal STR assay to aid human identity testing. J Forensic Sci 2009;54:1008-15.

Hill CR, Duewer DL, Kline MC, Sprecher CJ, McLaren RS, Rabbach DR, et al. Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex® ESX 17 and ESI 17 Systems. Forensic Sci Int Genet doi:10.1016/j.fsigen.2010.03.014.

Junge A, Brinkmann B, Fimmers R, Madea B. Mutations or exclusion: an unusual case in paternity testing. Int J Legal Med 2006 Nov;120(6):360-3.

Nothnagel M, Schmidtke J, Krawczak M. Potentials and limits of pairwise kinship analysis using autosomal short tandem repeat loci, Int J Legal Med 2010;124(3):205-15.

O'Connor KL, Hill CR, Vallone PM, Butler JM. Linkage disequilibrium analysis of D12S391 and vWA in U.S. population and paternity samples. Forensic Sci Int Genet doi:10.1016/j.fsigen.2010.09.003.

Parsons T, Coble M. Increasing the forensic discrimination of mitochondrial DNA testing through analysis of the entire mitochondrial DNA genome. Croat Med J 2001;42:304-9.

Phillips C, Fernandez-Formoso L, Garcia-Magariños M, Porras L, Tvedebrink T, Amigo J, et al. Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel. Forensic Sci Int Genet doi:10.1016/j.fsigen.2010.02.003.

Poetsch M, Lüdcke C, Repenning A, Fischer L, Mályusz V, Simeoni E, et al. The problem of single parent/child paternity analysis--practical results involving 336 children and 348 unrelated men. Forensic Sci Int 2006 Jun 2;159(2-3):98-103.

Pu CE, Linacre A. Increasing the confidence in half-sibship determination based upon 15 STR loci. J Forensic Leg Med 2008 Aug;15(6):373-7.

Reid T, Baird M, Reid J, Lee S, Lee R. Use of sibling pairs to determine the familial searching efficiency of forensic databases. Forensic Sci Int Genet 2008;2:340-2.

Sheehan NA, Egeland T. Structured incorporation of prior information in relationship identification problems. Ann Hum Genet 2007 Jul;71(4):501-18.

Shoemaker JS, Painter IS, Weir BS. Bayesian statistics in genetics: a guide for the uninitiated. Trends Genet 1999 Sep;15(9):354-8.

Weir BS, Anderson AD, Hepler AB. Genetic relatedness analysis: Modern data and new challenges. Nat Rev Genet 2006 Oct;7(10):771-80.

Wenk RE, Chiafari FA, Gorlin J, Polesky HF. Better tools are needed for parentage and kinship studies. Transfusion 2003 Jul;43(7):979-81.

Wilkening S, Chen B, Hemminki K, Forsti A. STR markers for kinship analysis. Human Biology 2006;78(1):1-8.

List of Figures

Figure 1. Representation of the expansion of the U.S. forensic core competency.

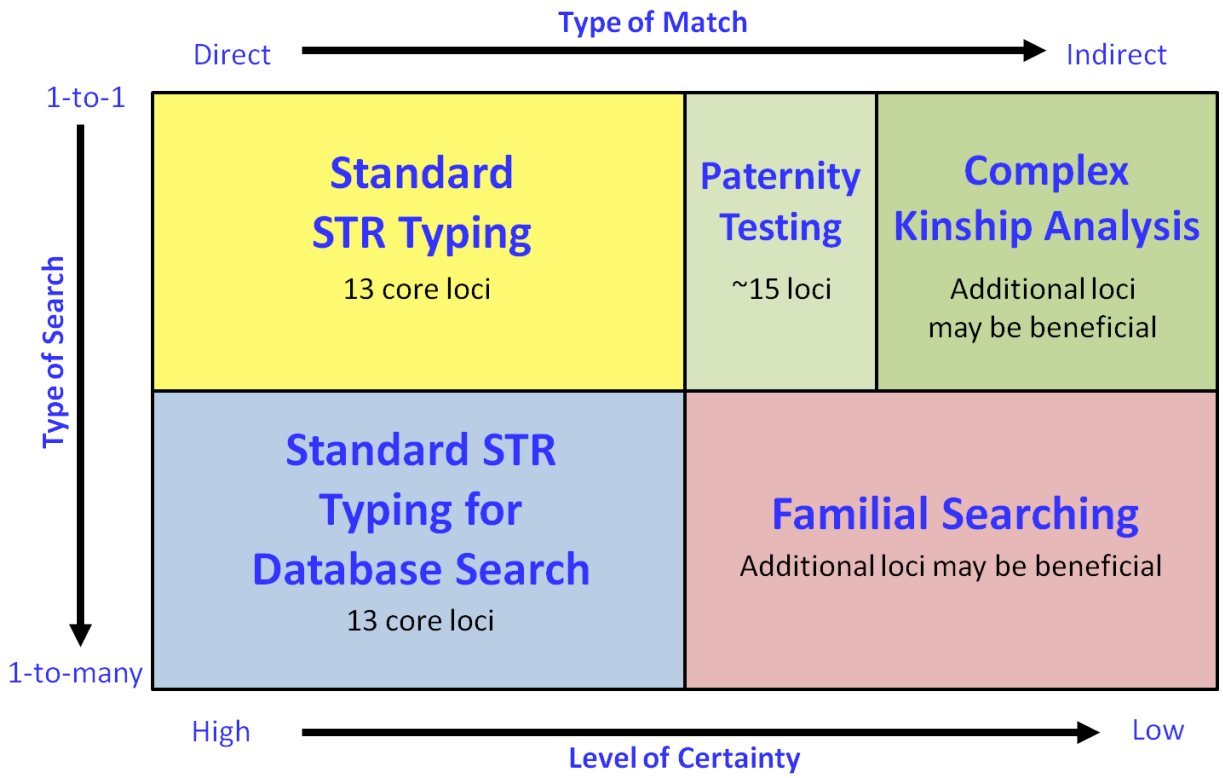


Figure 2. Distributions of likelihood ratio (LR) values for simulations of true parent-offspring, full sibling, and half sibling relationships using genotypes of 13 U.S. core (CODIS) STR loci, 20 STR loci, and 40 STR loci. Genotypes were simulated using NIST U.S. **Caucasian** allele frequency data. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

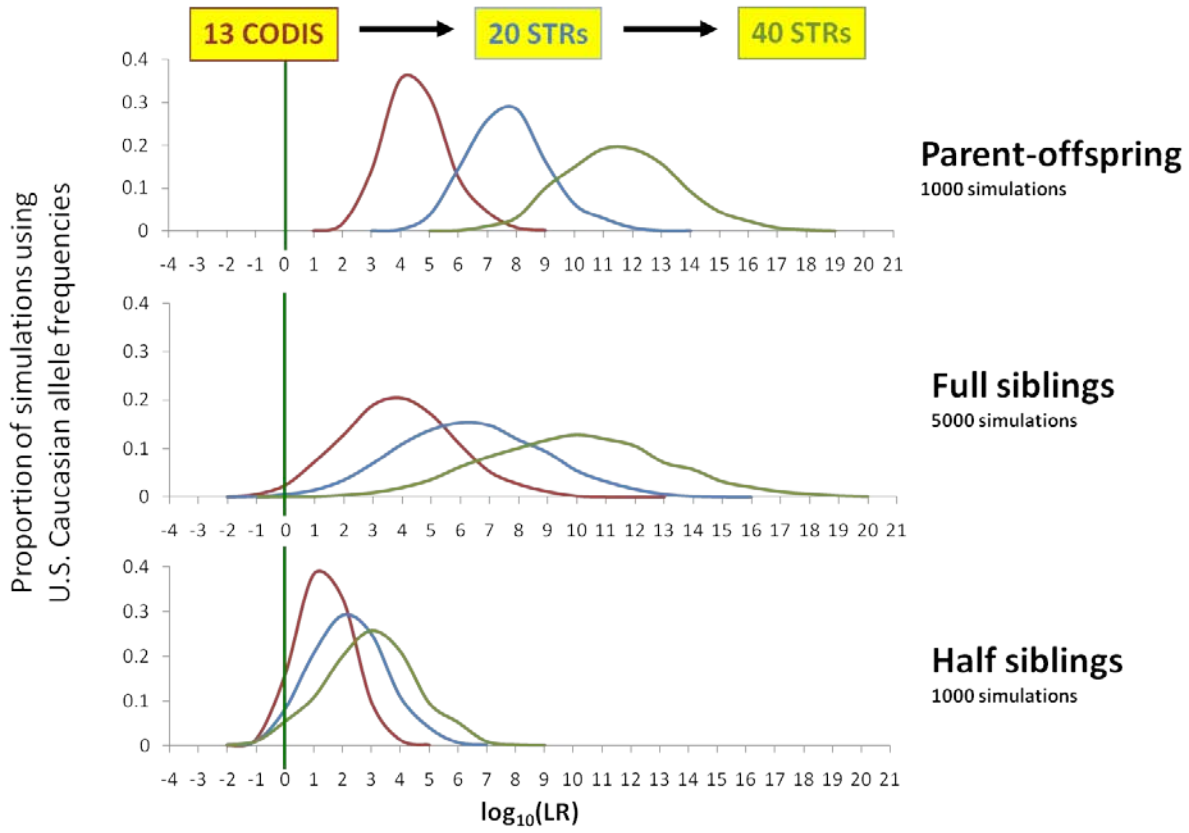


Figure 3. Distributions of likelihood ratio (LR) values for simulations of true parent-offspring, full sibling, and half sibling relationships using genotypes of 13 U.S. core (CODIS) STR loci, 20 STR loci, and 40 STR loci. Genotypes were simulated using NIST U.S. **Hispanic** allele frequency data. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

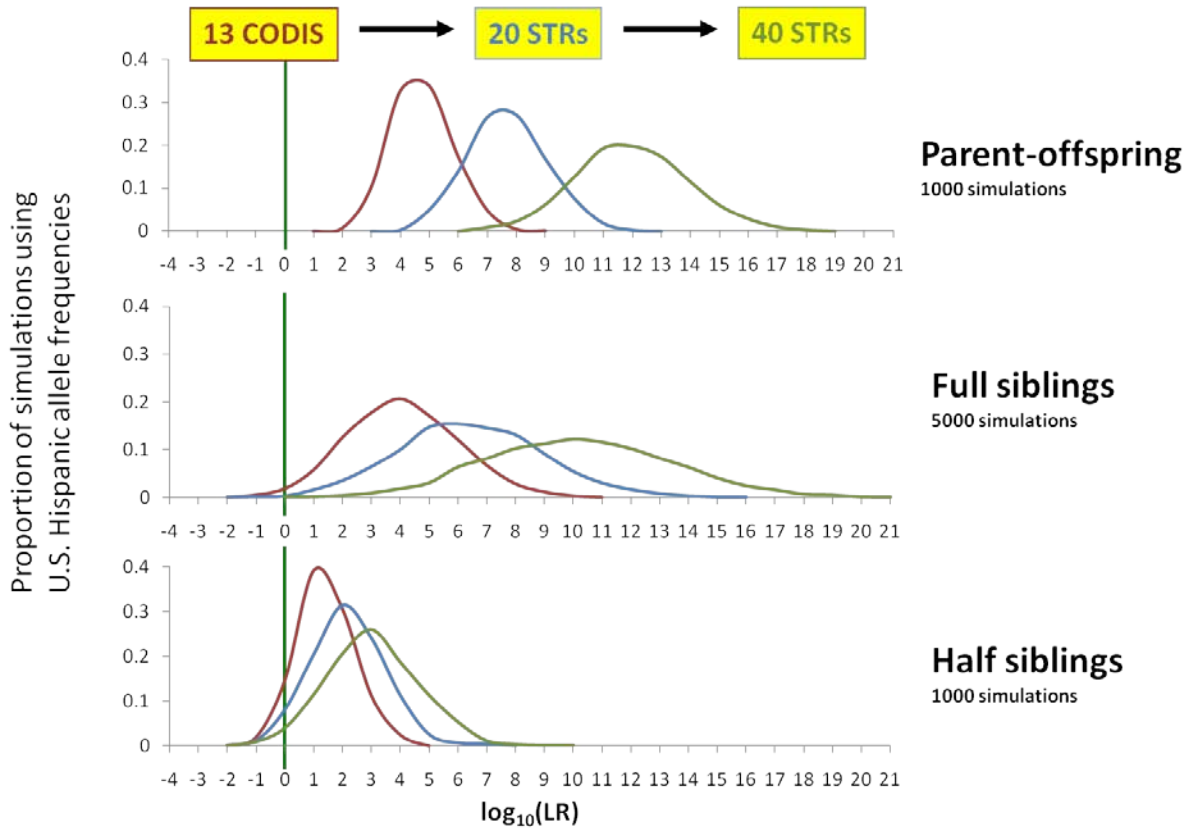


Figure 4. Distributions of likelihood ratio (LR) values for simulations of true parent-offspring, full sibling, and half sibling relationships using genotypes of 13 U.S. core (CODIS) STR loci, 20 STR loci, and 40 STR loci. Genotypes were simulated using NIST U.S. **African American** allele frequency data. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

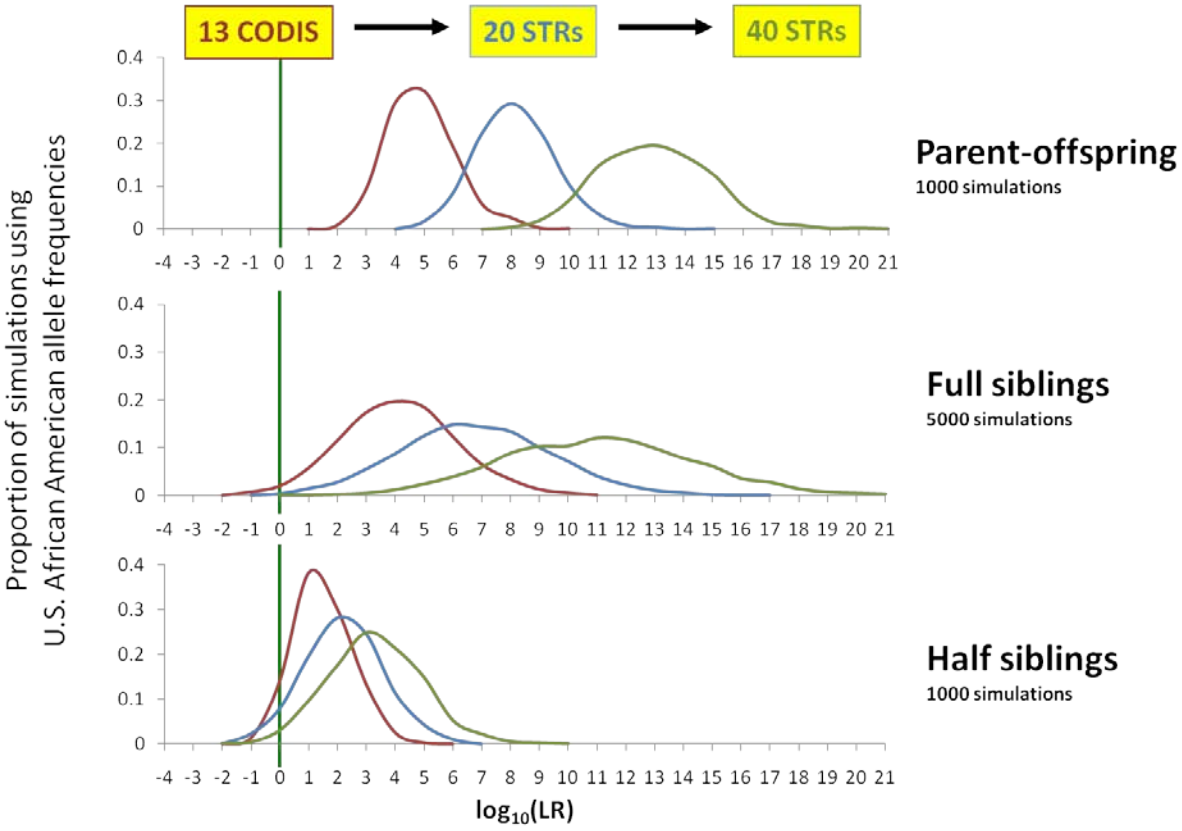


Figure 5. Schematic of overlapping likelihood ratio distributions after simulating related and unrelated genotypes for kinship analysis. A specific likelihood threshold will define regions of the distribution where false positives and false negatives occur. By definition, a likelihood ratio threshold greater than one supports the numerator (hypothesis of relationship), and a likelihood ratio threshold less than one supports the denominator (hypothesis of no relationship).

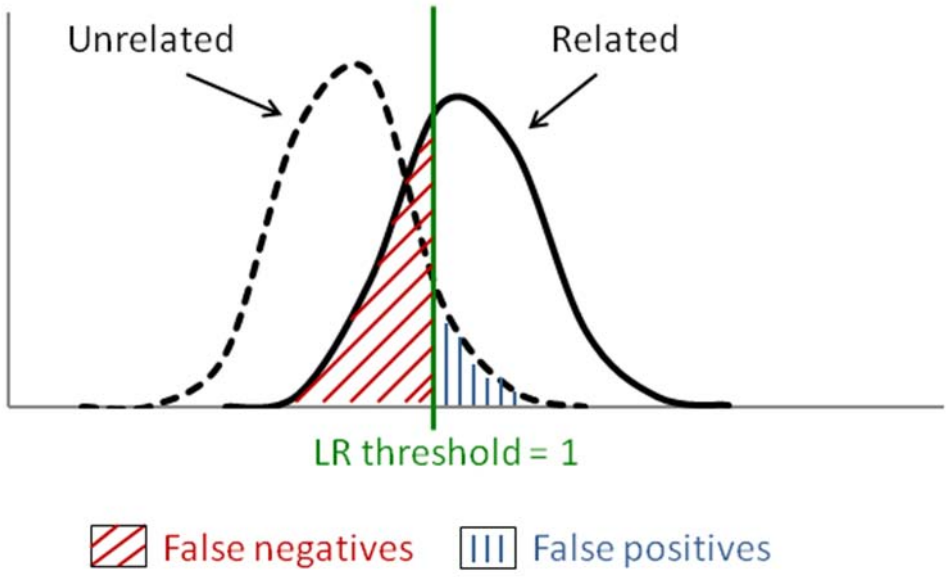


Figure 6. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **Caucasian** allele frequency data for **13** U.S. core STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

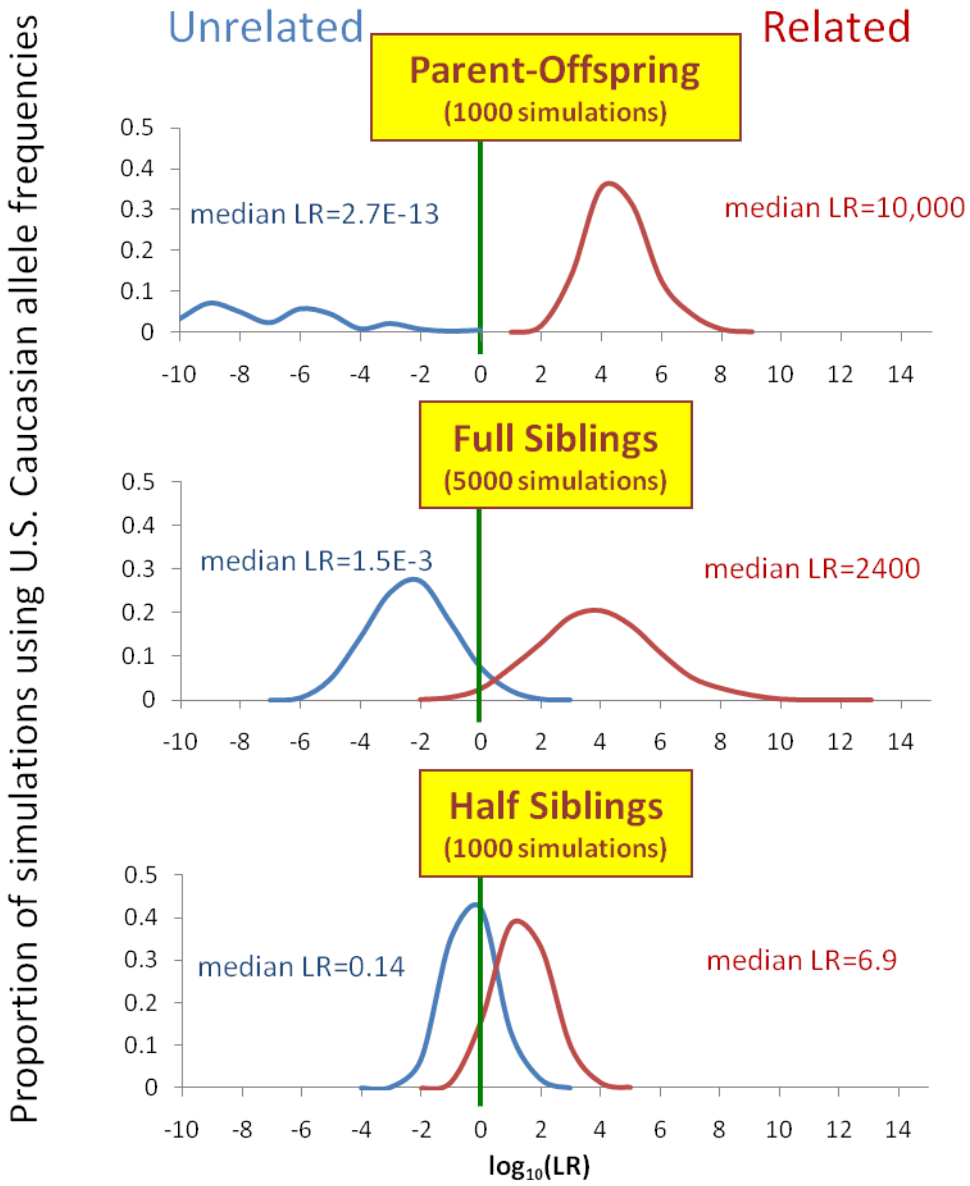


Figure 7. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **Hispanic** allele frequency data for **13** U.S. core STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

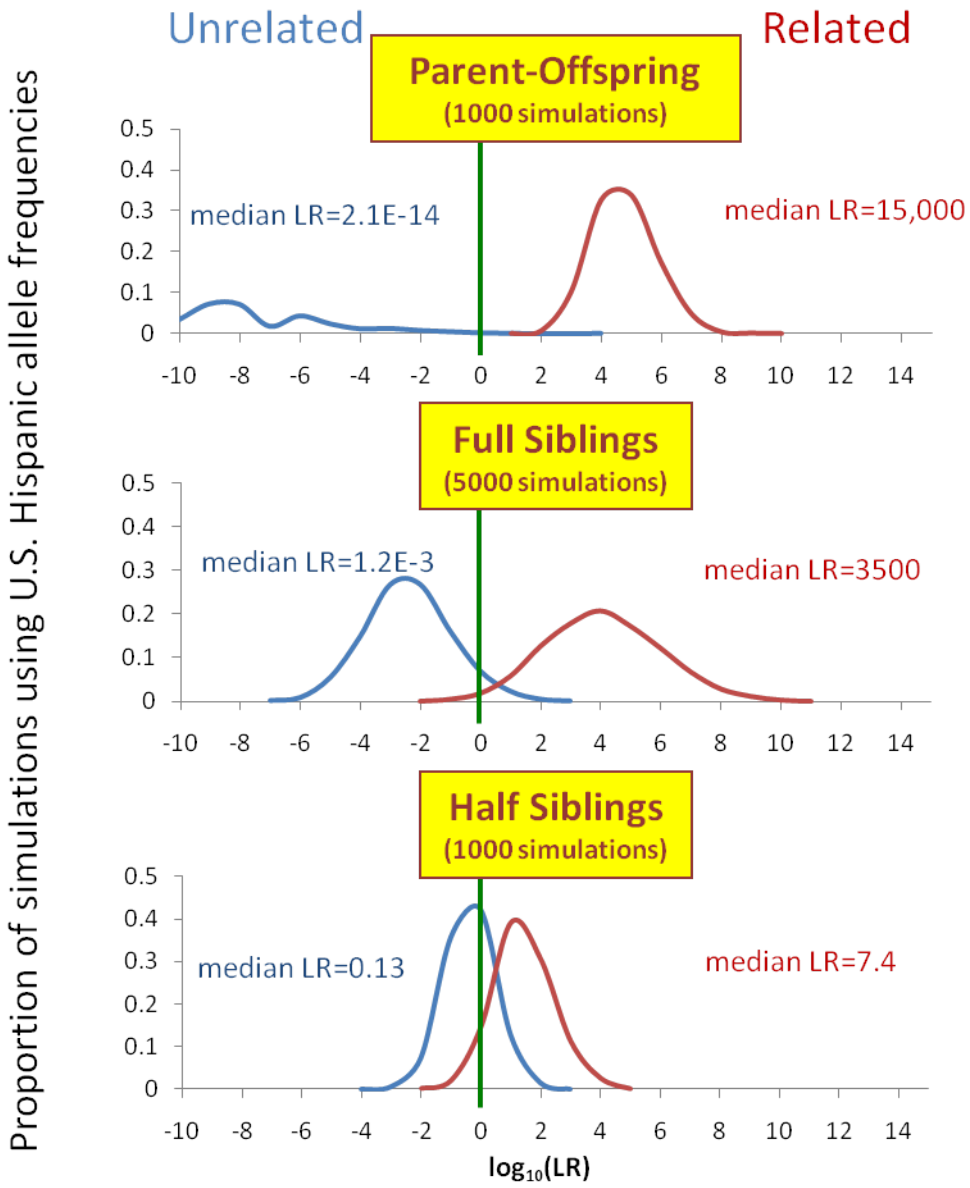


Figure 8. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **African American** allele frequency data for **13** U.S. core STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

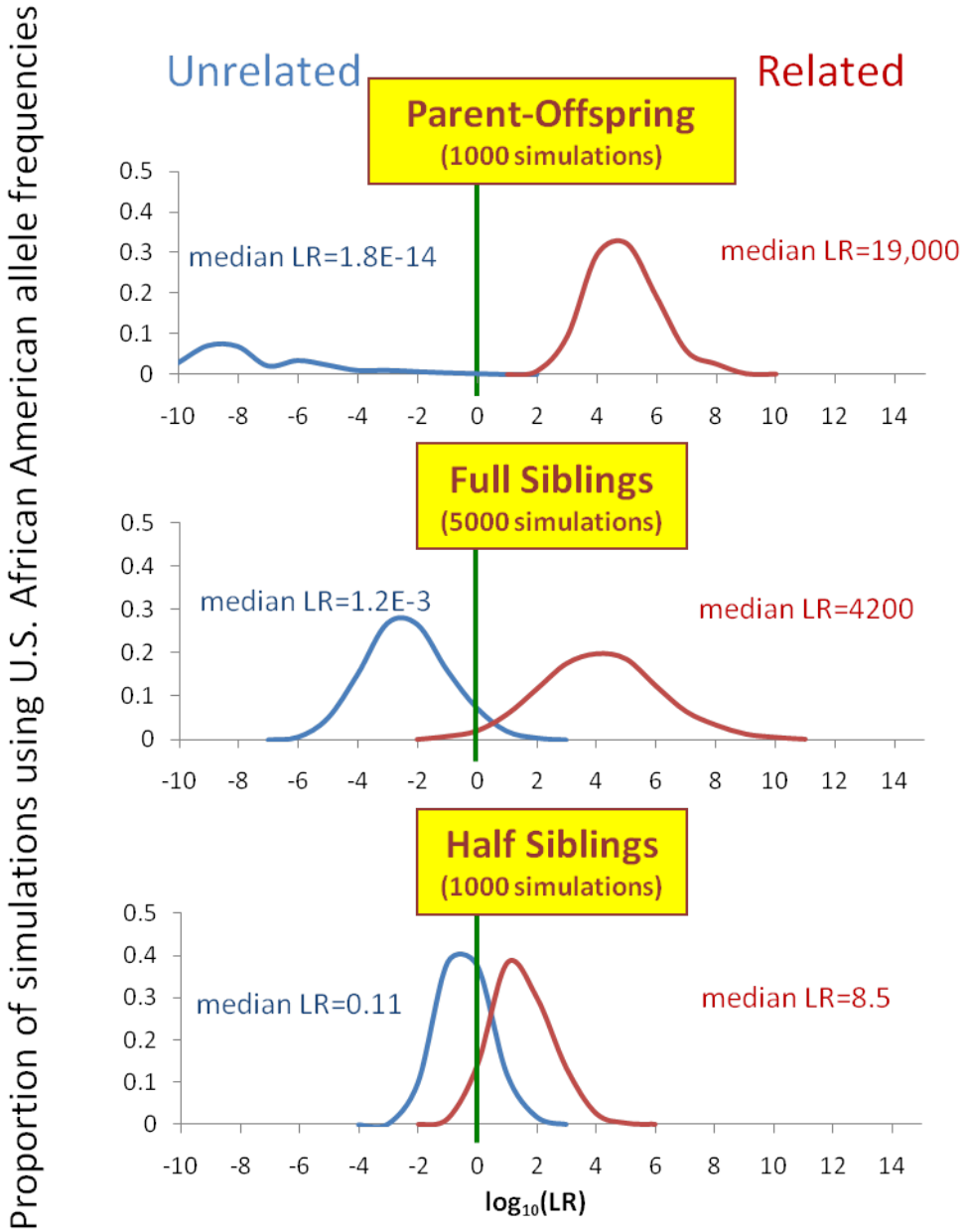


Figure 9. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **Caucasian** allele frequency data for **20** STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

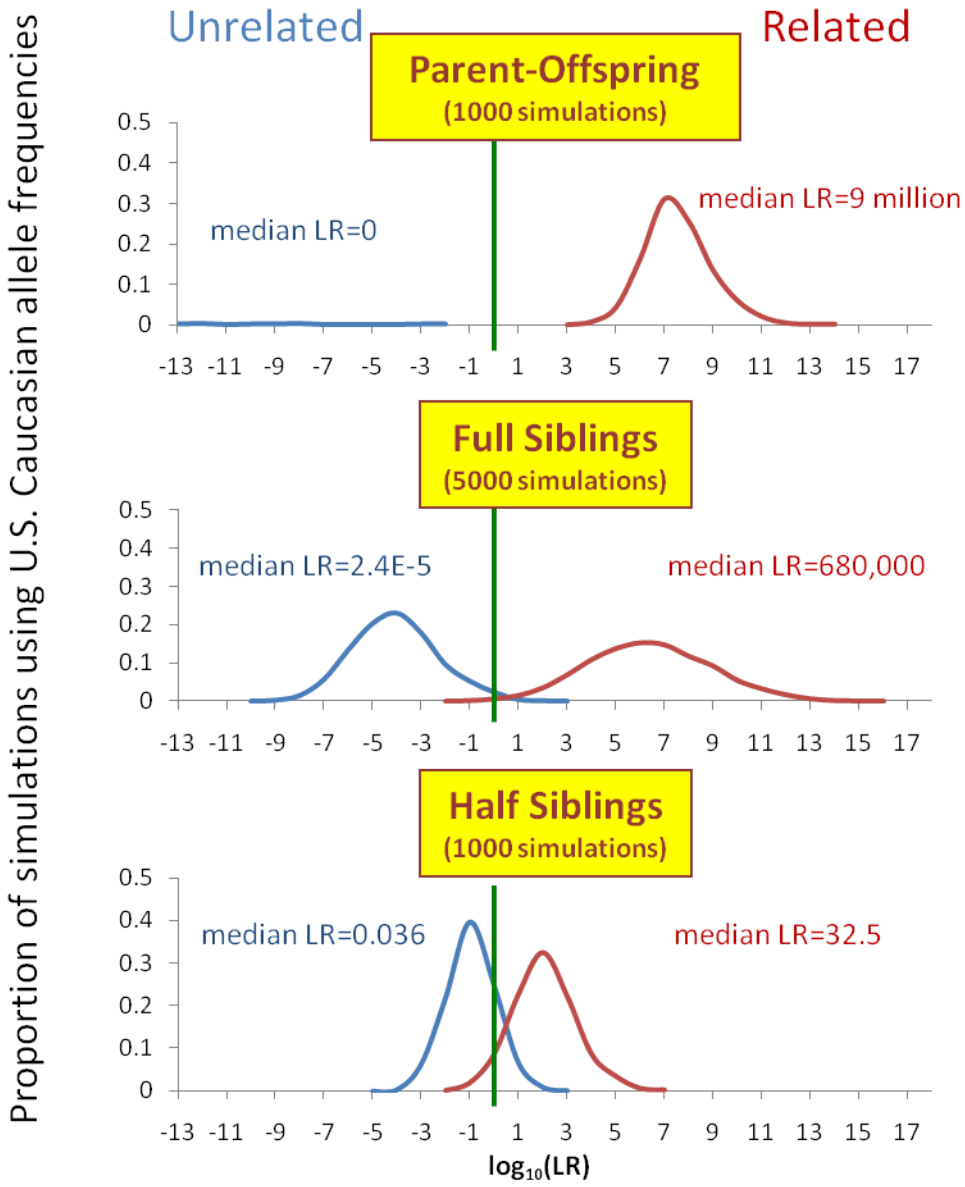


Figure 10. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **Hispanic** allele frequency data for **20** U.S. core STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

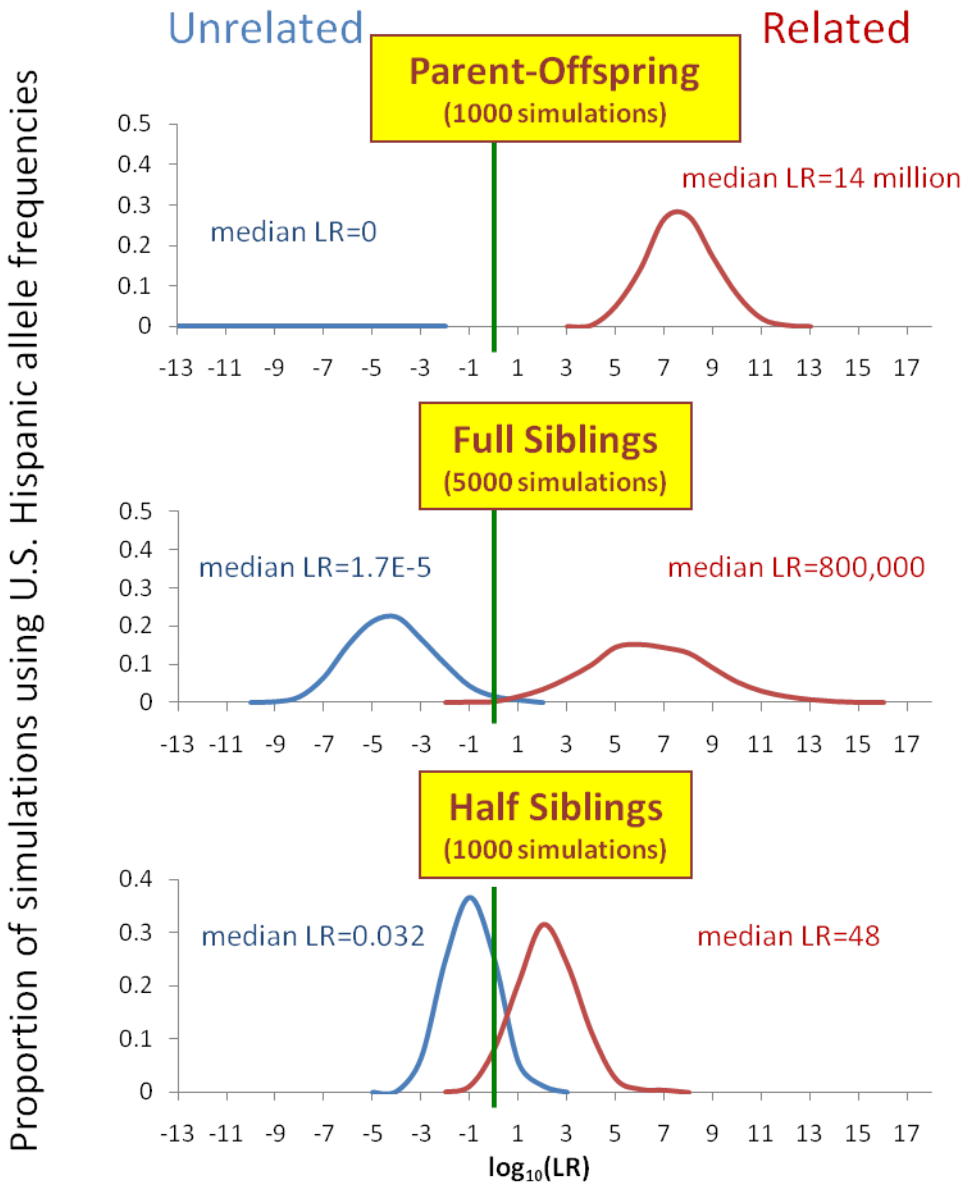


Figure 11. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **African American** allele frequency data for **20** U.S. core STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

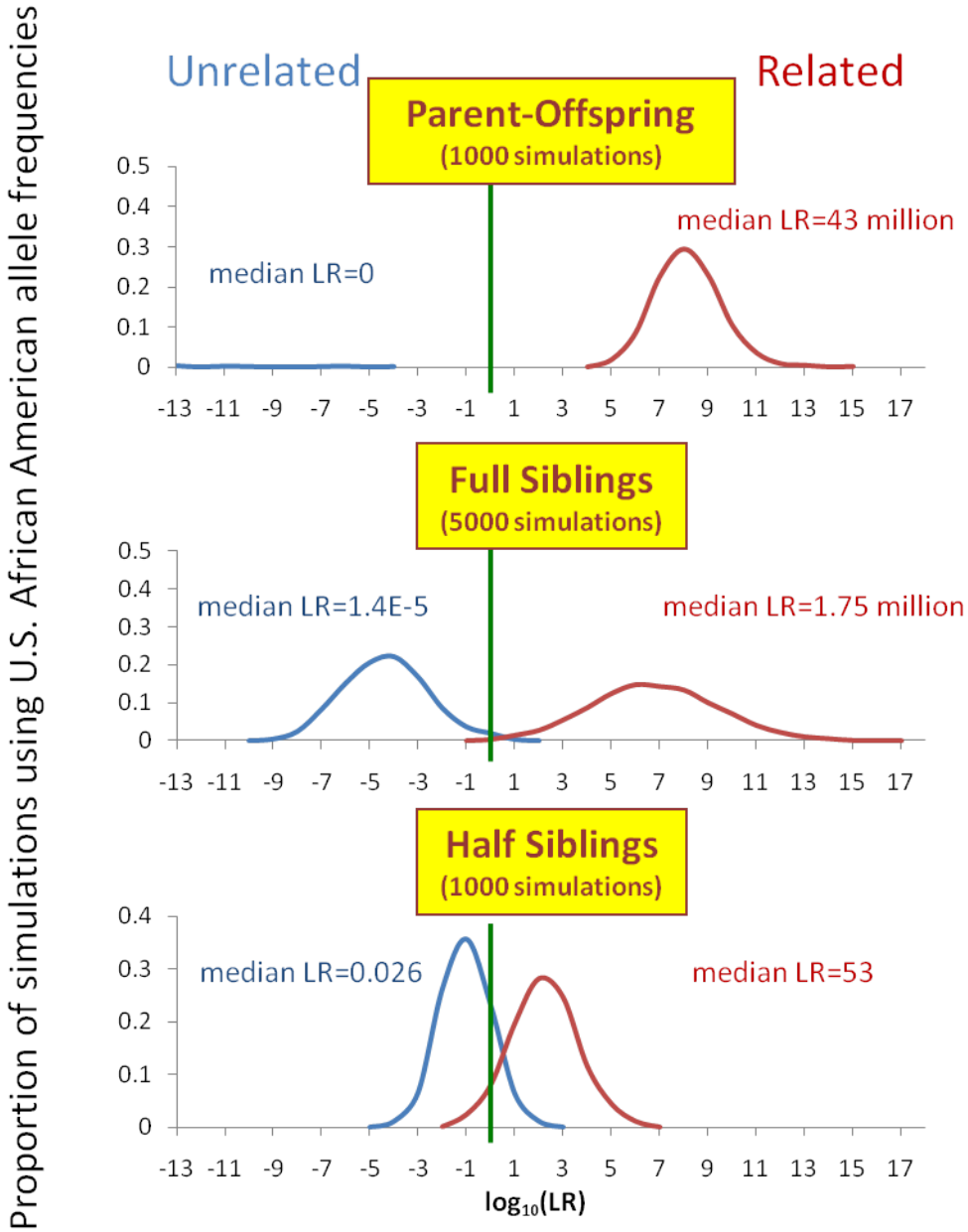


Figure 12. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **Caucasian** allele frequency data for **40** STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

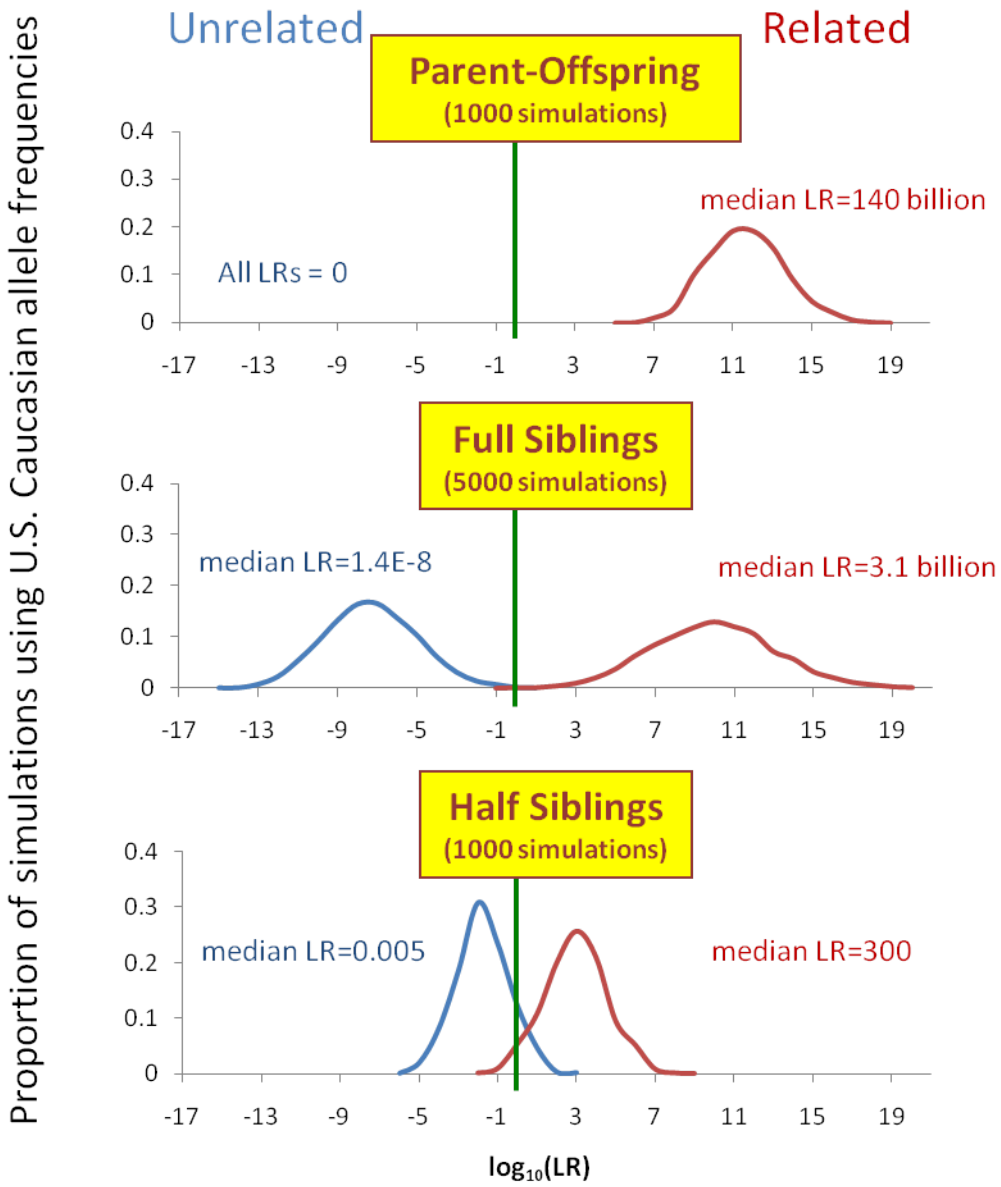


Figure 13. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **Hispanic** allele frequency data for **40** STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

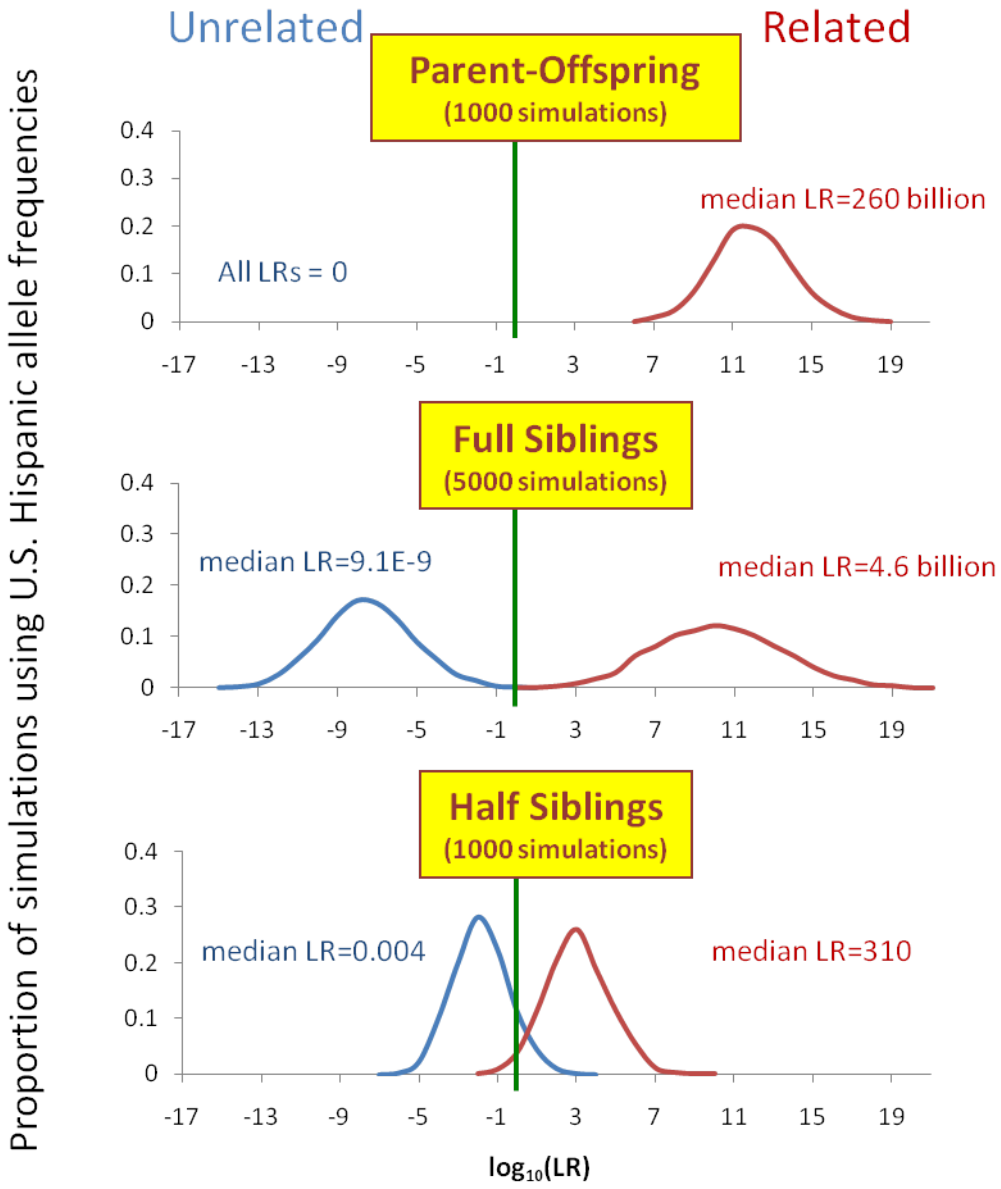


Figure 14. Distributions of likelihood ratio (LR) values for true relatives and unrelated persons. Pairs of genotypes were simulated from NIST U.S. **African American** allele frequency data for **40** STR loci. $Log_{10}(LR) = 0$ corresponds to a LR threshold of one.

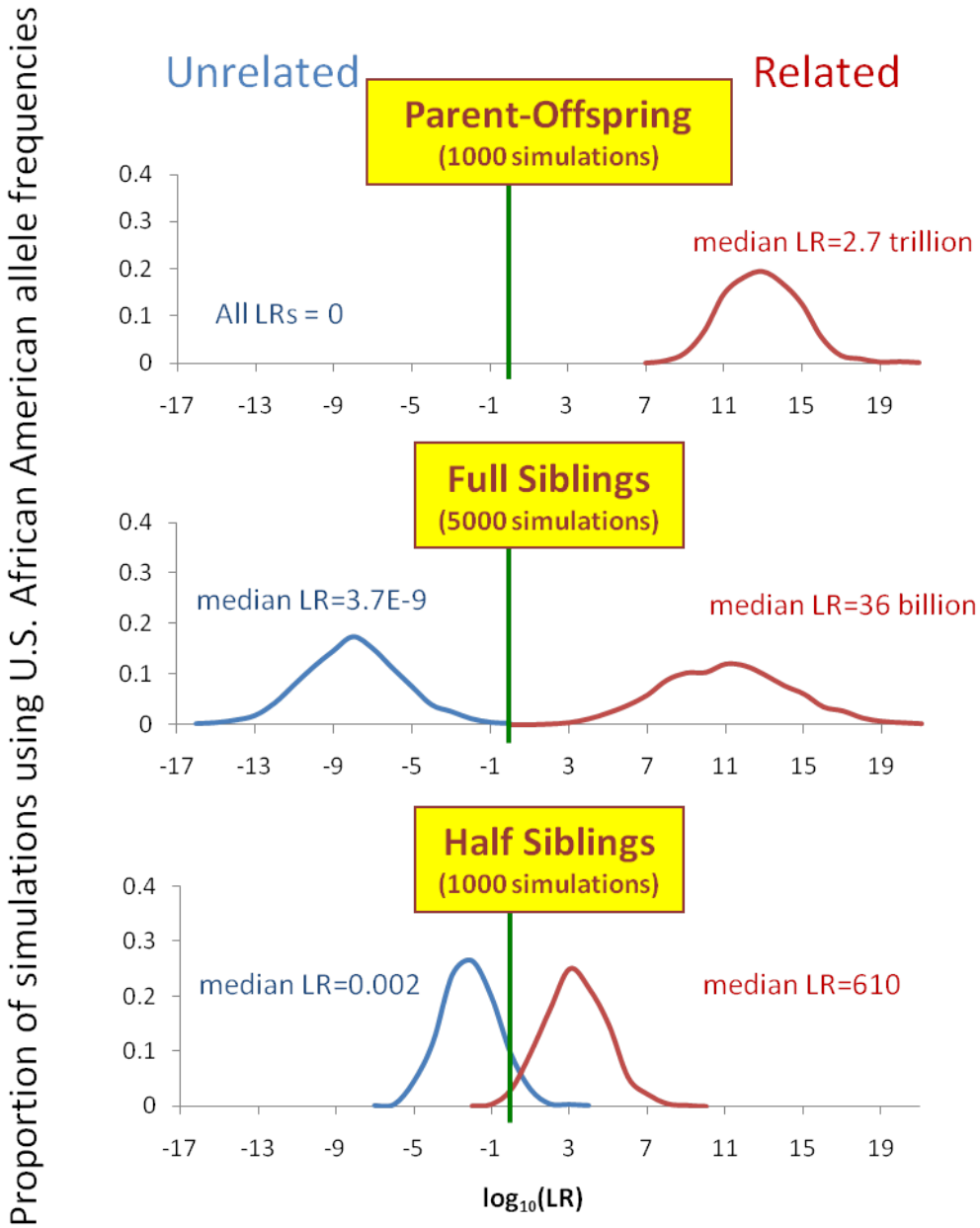


Table 1. Identity by descent (IBD) probabilities for pairs of non-inbred relatives (Weir et al. 2006). The values k_0, k_1, and k_2 correspond with the probabilities of sharing zero, one, and two alleles IBD, respectively.

Relationship	**k_0**	**k_1**	**k_2**
Identical twins	0	0	1
Parent-offspring	0	1	0
Full siblings	1/4	1/2	1/4
Half siblings	1/2	1/2	0
Avuncular	1/2	1/2	0
Grandparent-grandchild	1/2	1/2	0
First cousins	3/4	1/4	0
Double first cousins	9/16	3/8	1/16
Unrelated	1	0	0

Table 2. List of 13 U.S. core STR loci and expanded sets of 20 and 40 STR loci used in this study. The 12 European Standard Set loci are composed of the 11 loci in bold plus D12S391.

13 Loci	20 Loci	40 Loci
TPOX	TPOX	TPOX
CSF1PO	CSF1PO	CSF1PO
D5S818	D5S818	D5S818
D7S820	D7S820	D7S820
D13S317	D13S317	D13S317
FGA	**FGA**	**FGA**
vWA	**vWA**	**vWA**
D3S1358	**D3S1358**	**D3S1358**
D8S1179	**D8S1179**	**D8S1179**
D18S51	**D18S51**	**D18S51**
D21S11	**D21S11**	**D21S11**
TH01	**TH01**	**TH01**
D16S539	D16S539	D16S539
	D2S1338	D2S1338
	D19S433	D19S433
	D2S441	**D2S441**
	D10S1248	**D10S1248**
	D22S1045	**D22S1045**
	D1S1656	D1GATA113
	SE33	D1S1627
		D1S1677
		D2S1776
		D3S3053
		D3S4529
		D4S2364
		D4S2408
		D5S2500
		D6S474
		D6S1017
		D9S1122
		D9S2157
		D10S1435
		D11S4463
		D12ATA63
		D14S1434
		D17S974
		D17S1301
		D18S853
		D20S482
		D20S1082

Table 3. Frequency of false positive and false negative relatives defined with a likelihood ratio threshold equal to one. Genotypes of true relatives or unrelated persons were simulated from NIST U.S. Caucasian, Hispanic, and African American allele frequency data for 13 U.S. core STR loci, 20 STR loci, and 40 STR loci. The pairwise relationships evaluated were parent-offspring (PO), full siblings (FS), and half siblings (HS).

	13 Loci			**20 Loci**			**40 Loci**		
	PO	FS	HS	PO	FS	HS	PO	FS	HS
False positives									
Caucasian	0	0.027	0.155	0	0.006	0.075	0	0.001	0.051
Hispanic	0.002	0.026	0.143	0	0.007	0.067	0	0.001	0.058
African American	0.002	0.025	0.140	0	0.004	0.074	0	0.001	0.039
False negatives									
Caucasian	0	0.033	0.173	0	0.008	0.104	0	0.002	0.066
Hispanic	0	0.026	0.168	0	0.006	0.094	0	0.001	0.050
African American	0	0.029	0.156	0	0.005	0.105	0	0.001	0.036

Tab D