P

PARENTAGE TESTING

G S Williams, Northern Illinois University, DeKalb, IL, USA

Introduction

Parentage testing before 1986 involved only the use of serology to determine red blood cell (RBC) surface antigens, human leukocyte antigens (HLA), serum proteins, and RBC enzymes for identifying family members. Serology is literally defined as the study of serum (blood serum). However, forensic serology also included immunological study of biological tissues, body fluids, and biological stains.

Forensic serology has been changed to forensic biology as analysts now also perform deoxyribonucleic acid (DNA) testing on cells other than from blood. In 1986, nucleic acid testing (NAT) became legally acceptable to establish DNA markers, thus increasing the power to identify or exclude individuals as biologically related. The expanded definition is the natural outgrowth from the increase in knowledge of polymorphic genes from 1980 to 2003. The human genome has been sequenced and it is estimated that 0.1% or 3 million basepairs are polymorphic; one in every 1000 bases has more than one allele. These polymorphic sites are either true coding regions and transcribed into proteins or are in noncoding regions. Most DNA-based testing is on the highly polymorphic noncoding DNA regions due to the increased ability to exclude a wrongfully accused person. The most recent addition to DNA parentage testing is single nucleotide polymorphism (SNP) analysis which is automated.

The bulk of this article will focus on paternity testing of the typical trio of mother (Mo), child (Ch), and one alleged father (AF). In addition to the typical trio, parentage testing and identity testing within families include maternity confirmation, determination of identical versus fraternal twins, identification of recovered kidnapped children, matching grandparents to grandchildren, analysis of paternity without maternal specimens, settling immigration and inheritance disputes, and sibling confirmation.

History

Landsteiner discovered the ABO blood groups between 1900 and 1901. In 1926, the first use of the ABO system to exclude (rule out) an AF occurred in Vienna. The following year, M and N antigens were discovered. Ten years later, ABO and MN were first used in parentage testing in the USA. In 1939, RBCs were shown to have rhesus (Rh) system antigens also. Four Rh system antigens, C, c, D, and E, are used by most labs that do RBC antigen typing for parentage testing. Other RBC systems or antigens that have added more power to the exclusion rate include K in the Kell system in 1946, S in the MN system in 1947, Cellano (k) in the Kell system in 1949, the Duffy system in 1950, the Kidd system in 1951, and the addition of s to the MNS system, also in 1951.

The ability to rule out a wrongfully accused man (exclusion power) using RBC antigens was good but not good enough for many cases. In 1955, polymorphism of haptoglobin (Hp) was described and subsequently, other RBC enzymes and serum proteins were found to be polymorphic and useful for identity testing. The recognition that polymorphisms occur in serum proteins and RBC enzymes added more markers to help improve exclusion and inclusion percentages.

Prior to DNA testing, HLA testing was the best method to exclude and improve inclusion (possible biological father) probabilities of an AF. In 1972, HLA became available as a highly polymorphic system for use in paternity testing. By 1976, most laboratories used six RBC antigen systems of the more than 400 RBC surface antigens now known and HLA for parentage testing, and a few laboratories performed enzyme and protein testing.

DNA testing is now the norm throughout the world for determining genetic relatedness. In 1976, restriction fragment length polymorphism (RFLP) analysis was first described and it was used in 1980 to show

a highly polymorphic region of DNA. By 1986, RFLP was used for paternity testing. RFLP analysis requires large amounts of good-quality intact DNA and the procedure takes several days to perform. Newer testing methods for DNA are highly automated, do not require intact DNA, and are completed in a few hours.

When the polymerase chain reaction (PCR) became automated, new markers were amplified and used for identity testing. Variable number of tandem repeats (VNTR) genetic loci were identified as powerful markers for forensics and were soon used in parentage disputes. Both long tandem repeats (LTR) and short tandem repeats (STR) frequencies for most populations are established. These markers are now used more frequently than the traditional serological markers. By 1995, STR testing became practical and also replaced RFLP analysis for forensic identity testing and parentage studies. The most recent addition to parentage testing, SNP analysis, was approved for parentage testing by the American Association of Blood Banks (AABB) in 2003.

Current Practice

Most laboratories now use DNA analysis for parentage testing, some continue doing serological analysis of HLA, and fewer continue to offer the less expensive RBC antigen methods. Blood protein analysis has only been offered by specialized laboratories since 1976. Table 1 lists most markers currently used in parentage testing.

Specimen Collection

The specimens used for parentage testing vary depending upon the methods that will be used to analyze the sample. RBC antigen tests may be performed on clotted blood or anticoagulated blood in ethylenediaminetetraacetic acid (EDTA), or acid citrate dextrose (ACD). For HLA testing, ACD is

Table 1 Markers used in parentage testing[a]

Serology			*DNA*			
Red blood cell antigen systems	*Red blood cell enzymes and serum proteins*	*Human leukocyte antigens (HLA)*	*Restriction fragment length polymorphism*	*PCR SSP*	*PCR LTR*	*PCR STR*
ABO	PGM1	HLA-A	D1S339	HLA-DQA1	D1S80	FGA
Rh	ACP	HLA-B	D2S44	LDLR		HUMTHO1
MNSs	ESD		D4S139	GYPA		HUMTPOX
Kell	Hp		D4S163	HBGG		HUMCSF1PO
Kidd	GC		D5S110	D7S8		HUMF13AO1
Duffy	Gm		D6S132	GC		HUMVWA13/A
Lutheran	Am		D7S467			HUMFESFPS
Xg	Km		D10S28			HUMLIPOL
	AK		D12S11			D6S818
	ADA		D14S13			D9S302
	6-PGD		D17S26			D22S683
	TF		D17S79			D18S535
	BF					D7S1804
	M					D7S820
	C3					D3S2387
	GLO					D4S2366
	GPT					D5S1719
	UMPK					D3S1358
	PGP					D8S1179
						D13S317
						D16S538
						D18S51
						D21S11

[a]Tables 2, 3, and 6 have expanded listings of markers for HLA, and DNA analysis.

ACP, acid phosphatase; ADA, adenosine deaminase; AK, adenylate kinase; Am, immunoglobulin A polymorphic alleles; BF, Properdin factor B; C3, third component of complement; DNA, deoxyribonucleic acid; ESD, esterase D; GC, group-specific component; GLO, glyoxalase; Gm, immunoglobulin G polymorphic alleles; GPT, glutamate pyruvate transaminase; HBGG, hemoglobin G; Hp, haptoglobin; Km, light-chain polymorphic alleles; LDLR, low-density lipoprotein receptor; LTR, long tandem repeat; M, M subtyping of alpha$_1$-antitrypsin; PCR, polymerase chain reaction; 6-PGD, 6-phosphogluconate dehydrogenase; PGM, phosphoglucomutase; PGP, phosphoglycolate phosphatase; Rh, rhesus; SSP, sequence-specific primer or probe; STR, short tandem repeat; TF, transferrin; UMPK, uridine monophosphate kinase; GYPA, glycophorin A.

preferred for preservation of viability of white cells. ACD or Alsever's solution is also best for RBC enzyme studies to preserve the enzymes. Serum protein analysis is best done with serum from clotted blood.

For DNA tests, a wide variety of samples will do. If blood is drawn, 2–5 ml EDTA blood provides the best results because it preserves DNA by inhibiting DNase activity. FTA® paper also preserves DNA and samples can be stored for extended periods at room temperature. Alternative DNA samples include buccal cells or other tissue cells. When buccal samples are collected, different-colored swabs designated for each person help to avoid mix-up. For instance, a laboratory could have yellow, red, green, and blue swabs and always use yellow for a Ch, blue for AF 1, green for AF 2, and red for the mother.

After informed consent is obtained from each adult in the case, the collection of specimens must be performed by someone with no vested interest in the outcome and witnessed by another disinterested party. Each specimen must be properly labeled and processed.

Selection of Genetic Loci

Not all genetic loci are created equal. Even those with multiple alleles may have drawbacks. Establishing Hardy–Weinberg equilibrium for the locus increases the reliability of the data analysis. Several assumptions are made in using the Hardy–Weinberg approach: (1) random mating; (2) large population; and (3) migration unlikely. The frequency of a given allele at a polymorphic (at least two alleles) locus is established for each population.

Polymorphisms in the human genome occur in coding and noncoding regions. Polymorphisms of proteins, lipids, and sugars are differences in the structure that do not change their function significantly. The mutations in coding regions can lead to a functional gene product that will subsequently be passed on as an allele at that locus. Other mutations lead to changes incompatible with normal function resulting in a miscarriage or shortened life span. In parentage testing, normal alleles, resulting in expression of carbohydrates, lipids, and proteins on the RBC surface, inside the RBC, and in serum are useful markers. These traditional systems have been well evaluated, and exceptions to the normal expected outcomes of inheritance and testing have been well described. All of these parentage markers are genetically stable with rare mutations of less than one in a million individuals.

As opposed to the low mutation rate of RBC antigens, HLA, or serum proteins by classical serological methods, mutations are a significant issue when using DNA-based methods. Noncoding DNA mutations do not have a selective disadvantage or advantage whereas mutations in the coding region often lead to a dysfunctional protein.

Co-dominant alleles inherited by classic Mendelian rules are best (segregation and independent assortment). Segregation refers to the mode of inheritance of alleles; two alleles at the same site (locus) are never found in the same gamete (ova or sperm). For instance in the MNS system, when a parent is genetically *MS/Ns*, only *Ns* or *Ms* will be in any one gamete. Independent assortment refers to the way genes for different traits are inherited. For instance, the inheritance of *O* is independent of the inheritance of *MS* because these genes are on different chromosomes. Other genes, like *D*, are far apart from *Fy* on chromosome 1. This is also referred to as linkage. *O* and *MS*, *D* and *Fy* are not linked to each other. Some genes are so closely linked that they are inherited together, like *MN* with *SsU*. Phenotypes that lead to probable genotypes are most useful. For instance, someone who phenotypes D+C+ c+e+ and is white is most likely *DCe/ce*. Silent alleles make the analytical process more complicated, as illustrated in a later section.

Genetic markers or loci should have databases established for all potentially tested races or ethnic groups. The distribution of the selected alleles should be effective in excluding an unrelated person (e.g., an AF who is not the biological father). In calculating probability of parentage or relatedness, the laboratory must take into account all loci that have linkage disequilibrium or are linked genetically. One common example of linkage disequilibrium occurs in the HLA system. HLA-A1 and HLA-B8 are found together more frequently than would be expected by their independent gene frequencies. In addition, HLA-DQA1 DNA testing results are not independent of HLA-A and HLA-B serological results because they are all closely linked on chromosome 6.

By contrast, although DNA testing for RFLP, PCR-VNTR, PCR-short tandem repeat (STR), PCR-SNP, or sequencing (SQ) is extremely powerful in including a possible father as perhaps the only person on the planet who could have contributed the sperm, it is also more likely to have confusing results from mutations in the noncoding regions that are most widely used for these tests. Some of the markers have a mutation rate of 0.03. Refer to Tables 2 and 3 for lists of mutation rates for DNA markers. DNA markers must also be stable under different host conditions and testing methods. This high rate of mutation could result in a male being excluded when he really is the father. An AF may have a very high inclusion rate of 99.9% and have one STR locus that appears to

Table 2 Mutation rates summarized for genetic markers analyzed by RFLP mapping[a]

System	*Maternal*[b] (%)	*Paternal*[b] (%)	*Null*[c] (%)	*Multi-banded*
D1S7	9/580 (1.55)	11/721 (1.52)	1/560 (0.17)	2/461 (<0.43)
D1S339	206/87600 (0.24)	388/104432 (0.37)	77/91846 (0.081)	143/69999 (0.20)
D2S44	335/203411 (0.17)	239/225733 (0.17)	465/233293 (0.20)	361/224260 (0.16)
D4S139	35/76809 (0.05)	951/100806 (0.94)	18/78545 (0.02)	778/83024 (0.94)
D4S163	4/21669 (0.02)	42/41635 (0.10)	32/46065 (0.07)	16/29731 (0.05)
D5S110	135/24567 (0.55)	412/23684 (1.74)	10/25879 (0.04)	502/30372 (1.65)
D5S43	0/525 (<0.191)	0/536 (<0.187)	UNK	UNK
D6S132	11/56265 (0.02)	67/84917 (0.08)	2/98399 (<0.01)	39/132922 (0.03)
D7S21	20/979 (2.04)	41/1317 (3.10)	UNK	1/1235 (0.081)
D7S22	15/2734 (0.55)	91/3187 (2.86)	UNK	UNK
D7S467	18/91022 (0.02)	156/140441 (0.11)	15/165771 (<0.01)	46/145815 (0.03)
D10S28	337/183546 (0.18)	180/188480 (0.10)	60/165265 (0.04)	144/167649 (0.09)
D12S11	5/15054 (0.03)	14/19043 (0.07)	3/19022 (0.02)	5/16199 (0.03)
D14S13	19/30596 (0.06)	108/33085 (0.33)	3/21391 (0.01)	119/26343 (0.45)
D16S309	0/176 (<0.06)	2/2129 (0.09)	UNK	UNK
D16S85	0/518 (<0.19)	2/542 (0.55)	0/676 (<0.148)	0/676 (<0.148)
D17S26	60/63059 (0.10)	157/65205 (0.24)	3/21165 (0.01)	32/55997 (0.06)
D17S79	7/15329 (0.05)	21/21222 (0.10)	12/10345 (0.12)	14/17582 (0.08)

[a]The mutation rates include data from the 2000 AABB Annual Report Summary.
[b]The data under these column headings refers to the number of inconsistencies/number of total meioses expressed as a percentage within the parentheses.
[c]Null alleles are assumed to exist in cases of paternal or maternal exclusion due to nonmatching homozygous banding patterns when there is otherwise overwhelming evidence in favor of paternity or maternity.
Reprinted from American Association of Blood Banks (2002) *Guidance for Standards for Parentage Testing Laboratories*, 5th edn., p. 140. Bethesda, MD: American Association of Blood Banks with permission.

exclude him because of a common mutation. Refer to the AABB *Guidance for Standards for Parentage Testing Laboratories* for examples.

Methodology

The methodology chosen for testing must comply with standards set by each country. In general, all samples should be assessed by two independent analysts using a different set of reagents from different sources. Quality assurance of testing personnel and procedures is essential. In order to trust data from different laboratories, the same procedures should be used by multiple laboratories for comparison by proficiency testing.

Serological RBC antigen testing, RBC enzymes, serum proteins, and HLA testing have been standardized for many decades and frequency tables have been established. NAT by STR or SNP databases on the other hand have discrepancies in nomenclature. For example, one group describes the repeat for HUMTH01 in the STR analysis as AATG and GenBank (a public database) uses the other strand of DNA and calls it TCAT. All investigators working on establishing these databases must agree upon nomenclature and frequencies to ensure accuracy.

RBC Typing

The six most common families of cell-surface proteins tested are ABO, MNS, Rh, Duffy, Kell, and Kidd. In the USA, the combined exclusion power for RBC antigens is 65.5%. The following is a list of the exclusion power of different systems of RBC antigens: (1) ABO: 20%, (2) MNS: 31%, (3) Rh: 25%, (4) Duffy: 7%, (5) Kell: 4%, (6) Kidd: 6%. In addition, Lutheran (3.5%) and Xg (varies) are tested by some laboratories. Table 4 lists RBC antigens, nomenclature, and where their genes are located. These are stable markers on intact RBCs usually available for most parentage testing. Table 5 lists the RBC phenotype frequencies for various ethnic groups in the USA.

Typing samples for RBC surface antigens involves the use of antisera from human or monoclonal sources. For tests using antihuman globulin, special controls of the individual's RBCs without reagent antibody must be included to ensure that they are not already coated with antibody. These Coombs-positive cells would test positive for every marker requiring antihuman globulin as part of the procedure.

ABO, Rh, and MNS systems are the most powerful RBC antigen systems to use for exclusions. The common phenotypes of the ABO system are O, A_1, A_2, B, A_1B, and A_2B. Several rare subgroups of

Table 3 Apparent mutations summarized for genetic systems analyzed by PCR[a]

System	Maternal[b] (%)	Maternal null (%)	Paternal[b] (%)	Paternal null[c] (%)
D1S80	4/14052 (0.03)	UNK	75/199543 (0.04)	2/60372 (<0.01)
D1S2131	0/1212 (<0.08)	UNK	3/1240 (0.24)	UNK
D1S533	UNK	UNK	6/3830 (0.16)	UNK
D2S1338	UNK	UNK	9/45170 (0.02)	0/605 (<0.20)
D2S548	1/1212 (0.08)	UNK	0/1240 (<0.08)	UNK
D3S1358	5/18182 (0.03)	2/13293 (0.02)	77/52153 (0.15)	5/18094 (0.03)
D3S1744	16/10131 (0.16)	0/5697 (<0.02)	83/19555 (0.42)	0/8462 (<0.02)
D3S2386	0/1212 (<0.08)	UNK	1/1240 (0.08)	UNK
D5S818	29/127601 (0.02)	3/51327 (<0.01)	300/213999 (0.14)	18/67046 (0.03)
D7S820	19/120930 (0.02)	1/46032 (<0.01)	281/220861 (0.13)	5/73184 (<0.01)
D8S306	1/1212 (0.08)	UNK	3/1240 (0.24)	UNK
D8S1179	9/26510 (0.03)	3/14451 (0.02)	105/53114 (0.20)	4/22417 (0.02)
D9S302	19/8332 (0.22)	0/5669 (<0.02)	49/11179 (0.44)	0/8568 (<0.02)
D10S1214	28/2903 (0.97)	UNK	114/2938 (3.88)	UNK
D12S1090	9/4894 (0.18)	UNK	108/11957 (0.90)	0/5865 (<0.02)
D13S317	65/128422 (0.05)	112/68583 (0.16)	242/159361 (0.15)	131/134466 (0.10)
D13S764	0/1212 (<0.08)	UNK	0/1240 (<0.08)	UNK
D14S297	0/1212 (<0.08)	UNK	0/1240 (<0.08)	UNK
D16S539	19/97307 (0.02)	4/54649 (<0.01)	118/112872 (0.11)	14/58079 (0.02)
D17S5	0/228 (<0.44)	UNK	7/6568 (0.11)	UNK
D17S1185	0/1212 (<0.08)	UNK	0/1240 (<0.08)	UNK
D18S51	17/26804 (0.06)	2/12363 (0.02)	113/55362 (0.20)	3/20396 (0.02)
D18S535	1/2676 (0.04)	UNK	2/2624 (0.08)	0/5300 (<0.02)
D18S849	0/4281 (<0.03)	UNK	15/9594 (0.16)	0/5904 (<0.02)
D19S253	8/2997 (0.27)	1/1785 (0.06)	17/3247 (0.52)	7/2007 (0.35)
D21S11	31/28305 (0.11)	1/16244 (<0.01)	79/51202 (0.15)	2/18790 (0.01)
D21S1437	0/1212 (<0.08)	UNK	1/1240 (0.08)	UNK
D22S445	2/1212 (0.17)	UNK	1/1240 (0.08)	UNK
D22S683	2/2670 (0.08)	UNK	9/2625 (0.34)	0/5295 (<0.02)
ACTBP2	0/330 (<0.30)	UNK	330/51610 (0.64)	UNK
CYP19	6/343 (1.75)	UNK	205/177210 (0.12)	321/47259 (0.68)
CYAR04	2/3539 (0.06)	UNK	UNK	UNK
FGA	14/26123 (0.05)	0/15175 (<0.01)	690/236659 (0.29)	6/24689 (0.02)
HUMCSF1P0	21/109907 (0.02)	1/60275 (<0.01)	451/314702 (0.14)	5/68573 (<0.01)
HUMFESFPS	3/16264 (0.02)	1/7606 (0.01)	78/143297 (0.05)	0/11761 (<0.01)
HUMF13A01	0/8152 (<0.01)	0/283 (<0.40)	35/62820 (0.06)	0/2724 (<0.04)
HUMF13B	1/9857 (0.01)	0/4435 (0.02)	8/24314 (0.03)	0/7675 (<0.01)
HUMLIPOL	0/6200 (<0.02)	0/2311 (<0.05)	6/8918 (0.07)	0/2961 (<0.04)
HUMTHO1	10/100219 (0.01)	2/56371 (<0.01)	21/154685 (0.01)	2/74346 (<0.01)
HUMTPOX	2/79616 (<0.01)	0/50850 (<0.01)	21/112758 (0.02)	2/66052 (<0.01)
HUMvWA31	44/135789 (0.03)	1/66959 (<0.01)	1130/351664 (0.32)	22/120230 (0.02)
Penta E	1/6248 (0.02)	0/6248 (<0.02)	10/8315 (0.12)	0/8315 (<0.01)

[a]The mutation rates include data from the 2000 AABB Annual Report Summary.
[b]The data under these column headings refer to the number of inconsistencies/number of total meioses expressed as a percentage within the parentheses.
[c]Null alleles are assumed when cases of paternal or maternal exclusion occur due to nonmatching homozygous banding patterns in cases in which there is overwhelming evidence in favor of paternity or maternity.
Reprinted from American Association of Blood Banks (2002) *Guidance for Standards for Parentage Testing Laboratories*, 5th edn., p. 141. Bethesda, MD: American Association of Blood Banks with permission.

A and a few subgroups of B also exist. Rh typing for parentage testing is usually limited to testing for D, C, c, and E antigens. Antisera are not always monospecific and care must be taken when dual specificity is found, such as anti-Ce, for a person who is e-negative would not react well with this antiserum. It is so rare to be e-negative in most populations that anti-e is not used routinely. Exceptions include Mexicans, Native Americans, and Asians (Table 5). Testing for MNS system antigens is complicated by anti-N that cross-reacts with M. Careful controls must be run to ensure monospecificity. *MS*, *Ms*, *NS*, and *Ns* are the four haplotypes for MNSs and another antigen, U, may be useful to type in blacks who are S−s−. The combined exclusion power in the USA is 59%.

Table 4 Parentage markers and their chromosomal location

System/ marker	*ISBT symbol*	*Gene (ISGN)*	*Chromosome*
ABO	ABO	*ABO*	9
MNS	MNS	*GYPA, GYPB, GYPE*	4
Rh	RH	*RHD, RHCE*	1
Kell	KEL	*KEL*	7
Duffy	FY	*DARC*	1
Kidd	JK	*SLC14A1*	18
Lutheran	LU	*LU*	19
Xg	XG	*XG*	X
HLA			6

ISBT, International Society of Blood Transfusion; ISGN, International System for Human Gene Nomenclature.

Duffy, Kidd, and Kell system antigens are all detected by the antihuman globulin test. Common Duffy system genes include *Fy* (a silent allele found primarily in blacks), Fy^a, and Fy^b. Common Kidd antigens are Jk^a and Jk^b); however, anti-Jk^b is very rare. In the Kell system, the common antigens are K and k. The addition of Lutheran to the set increases the combined rate to 67% but antiserum is very rare.

Enzymes and Proteins

Although testing for polymorphisms of enzymes and proteins is rare in current practice, a discussion is included for historical completeness. The polymorphic RBC enzymes and serum proteins are listed in Table 1. Immunoglobulin polymorphisms can only be tested on someone over 6 months old because maternal immunoglobulin remains in the infant until then. In addition, rare variants exist for each enzyme and protein and these must be cataloged.

The procedure to determine different alleles involves electrophoresis of serum proteins and subsequent staining to reveal multiple bands of human proteins. Most proteins have alleles that migrate to sufficiently different distances in the gel. When electrophoresis is insufficient to resolve alleles, isoelectric focusing (IEF) can be used; PGM1i would designate the use of IEF rather than standard electrophoresis. The least used are Gm, Am, and Km due to lack of reagents and difficulty interpreting the results. NAT for PGM and GC now substitutes for protein electrophoresis.

HLA

Compared to the RBC antigen system, the HLA system is highly complex. Its power to exclude is much higher but the technical expertise to do the testing and analysis is also much greater. The microlymphocytotoxicity test takes much more time and is also very costly to perform. Antisera are derived primarily from multiparous women (women who have had many babies, preferably with the same father) and monoclonal preparations. Different sets of antiserum are used for different ethnic groups. Every part of the

Table 5 Frequencies for red blood cell phenotypes in the USA

Phenotype	*Blacks*[a]	*Whites*[b]	*Asians*[c]	*Native American*[d]	*Mexican*
ABO					
O	49	45	43		56
A	27	41	27		28
B	19	10	25		13
AB	4	4	5		4
Rh/DCE					
DCe	17	42	70	44	
DCE	<1	<1	1	6	
DcE	11	14	21	34	
Dce	44	4	3	2	
dCe	2	2	2	2	
dCE	<1	<1	<1	<1	
dce	<1	1	<1	6	
dcE	26	37	3	11	
MNS					
M+N+	44	50			
M+N−	26	28			
M−N+	30	22			
S+s+	28	44			
S+s−	3	11			
S−s+	69	45			
S−s− U−	<1	0			
Duffy					
Fy(a+b+)	1	49	8.9[e]		
Fy(a+b−)	9	17	90.8[e]		
Fy(a−b+)	22	34	0.3[e]		
Fy(a−b−)	<1	68	0		
Kell					
K−k+	96.5	91			
K+k+	3.5	8.8			
K+k−	<0.1	0.2			
Kidd					
Jk(a+b+)	34	49	50		
Jk(a+b−)	57	28	23		
Jk(a−b+)	9	23	27		
Jk(a−b−)	<1	<1	<1		

[a]Black refers to origins from any of the African Negro racial groups who indicate their race as ''black'' or African-American.
[b]White refers to people of original Caucasian (before massive immigration) European, Middle East, or North African origin who indicate their race as ''white''. This includes hispanic, unless otherwise indicated.
[c]Asian refers to those with origins in the Far East (Asian Indian, Chinese, Filipino, Korean, Japanese, Pakistani, Vietnamese, etc.) unless specifically indicated.
[d]Native American refers to those with origins from tribes in North or South America considered ''first people'' and who indicate their race as ''Native American'' or ''Alaskan Native'' or from a particular tribe.
[e]Chinese.

Table 6 Complete listing of WHO-Recognized HLA-A, -B serologic specificities[a]

1980		*2000*		
HLA-A	*HLA-B*	*HLA-A*		
HLA-A1	HLA-B5	*HLA-B*		
HLA-A2	HLA-B7	A1	B5	B51(5)
HLA-A3	HLA-B8	A2	B7	B5102
HLA-A9	HLA-B12	A203	B703	B5103
HLA-A10	HLA-B13	A210	B8	B52(5)
HLA-A11	HLA-B14	A3	B12	B53
HLA-Aw19	HLA-B15	A9	B13	B54(22)
HLA-Aw23(9)	HLA-Bw16	A10	B14	B55(22)
HLA-Aw24(9)	HLA-B17	A11	B15	B56(22)
HLA-A25(10)	HLA-B18	A19	B16	B57(17)
HLA-A26(10)	HLA-Bw21	A23(9)	B17	B58(17)
HLA-A28	HLA-Bw22	A24(9)	B18	B59
HLA-A29	HLA-B27	A2403	B21	B60(40)
HLA-Aw30	HLA-Bw35	A25(10)	B22	B61(40)
HLA-Aw31	HLA-B37	A26(10)	B27	B62(15)
HLA-Aw32	HLA-Bw38(w16)	A28	B2708	B63(15)
HLA-Aw33	HLA-Bw39(w16)	A29(19)	B35	B64(14)
HLA-Aw34	HLA-B40	A30(19)	B37	B65(14)
HLA-Aw36	HLA-Bw41	A31(19)	B38(16)	B67
HLA-Aw43	HLA-Bw42	A32(19)	B39(16)	B70
	HLA-Bw44(12)	A33(19)	B3901	B71(70)
	HLA-Bw45(12)	A34(10)	B3902	B72(70)
	HLA-Bw46	A36	B40	B73
	HLA-Bw47	A43	B4005	B75(15)
	HLA-Bw48	A66(10)	B41	B76(15)
	HLA-Bw49(w21)	A68(28)	B42	B77(15)
	HLA-Bw50(w21)	A69(28)	B44(12)	B78
	HLA-Bw51(5)	A74(19)	B45(12)	B81
	HLA-Bw52(5)	A80	B46	Bw4
	HLA-Bw53		B47	Bw6
	HLA-Bw54(w22)		B48	
	HLA-Bw55(w22)		B49(21)	
	HLA-Bw56(w22)		B50(21)	
	HLA-Bw57(17)			
	HLA-Bw58(17)			
	HLA-Bw59			
	HLA-Bw60(40)			
	HLA-Bw61(40)			
	HLA-Bw62(15)			
	HLA-B63(15)			
	HLA-Bw4			
	HLA-Bw6			

[a]1980 ''WHO-Recognized Specificities'' are those *without* ''w'' designations.

Reprinted from American Association of Blood Banks (2002) *Guidance for Standards for Parentage Testing Laboratories*, 5th edn., p. 109. Bethesda, MD: American Association of Blood Banks with permission.

procedure must be done in duplicate by two different analysts. The complexity of interpretation, which includes cross-reactivity, antigen splits, and linkage disequilibrium, requires highly skilled analysts to evaluate results. World Health Organization-specified HLA-A and HLA-B alleles are listed in Table 6.

DNA Testing

Most laboratories use DNA-based procedures to establish parentage or relatedness. RFLP was the first method used and it is still used to exclude or include siblings, fathers, and mothers. Only four or five RFLPs are required to have a high power of inclusion. Discrimination between individuals can also be achieved by detecting VNTR (STR and long tandem repeat (LTR)), SNP, and SQ. When using STR loci, 9–10 loci must be tested to have the same power of inclusion as LTR/RFLP analysis. When only two alleles are at one locus, as with SNP testing, 50 loci must be tested to have strong inclusion likelihoods. SQ is not yet widely used for parentage studies due to cumbersome procedures. New rapid sequencers are becoming available and SQ may be more widely used in the future. Mitochondrial DNA testing is extremely challenging and not routinely done for parentage studies. Mitochondrial DNA is used to identify decomposed or badly damaged remains of military personnel and other samples that are badly damaged, as in the World Trade Center attack on September 11, 2001. It requires special facilities and equipment so only special laboratories perform these tests. It is the only way to match grandmothers to grandchildren when the parents are missing.

VNTR may be LTR or STR. LTR are sequences of DNA that contain 9–80 bases as a core sequence that is repeated consecutively a few to hundreds of times. STRs are sequences of 2–5 bases repeated 4–40 times. LTRs are detected by either Southern blot for RFLPs or by PCR. STRs are detected by PCR procedures. STRs of four repeats are the most widely used in identity testing.

RFLP RFLP refers to the variable sizes of DNA fragments resulting from the cutting with a restriction endonuclease (also called restriction enzyme: RE). The Southern blot procedure is used for RFLP and is illustrated in Figure 1. Most regions of use in parentage testing are listed in Table 1. Sample cases are shown in Figure 2.

Restriction endonucleases are isolated from bacteria and named to indicate the source. Hae III and Pst I are two commonly used REs used for parentage testing in the USA. Hinf I is used in Europe. Hae III is the third RE to be isolated from *Haemophilus aegyptius*. Pst I is the first RE to be isolated from *Providencia stuartii*. REs recognize double-stranded DNA in 4, 6, 8, or 10 basepair sequences and cut at particular sites within that sequence. Hae III recognizes 5′-GG^CC/3′-CC^GG and cuts between the C and G. Pst I recognizes 5′-CTGCA^G/3′-G^ACGTC and cuts between G and A. Hae III cuts

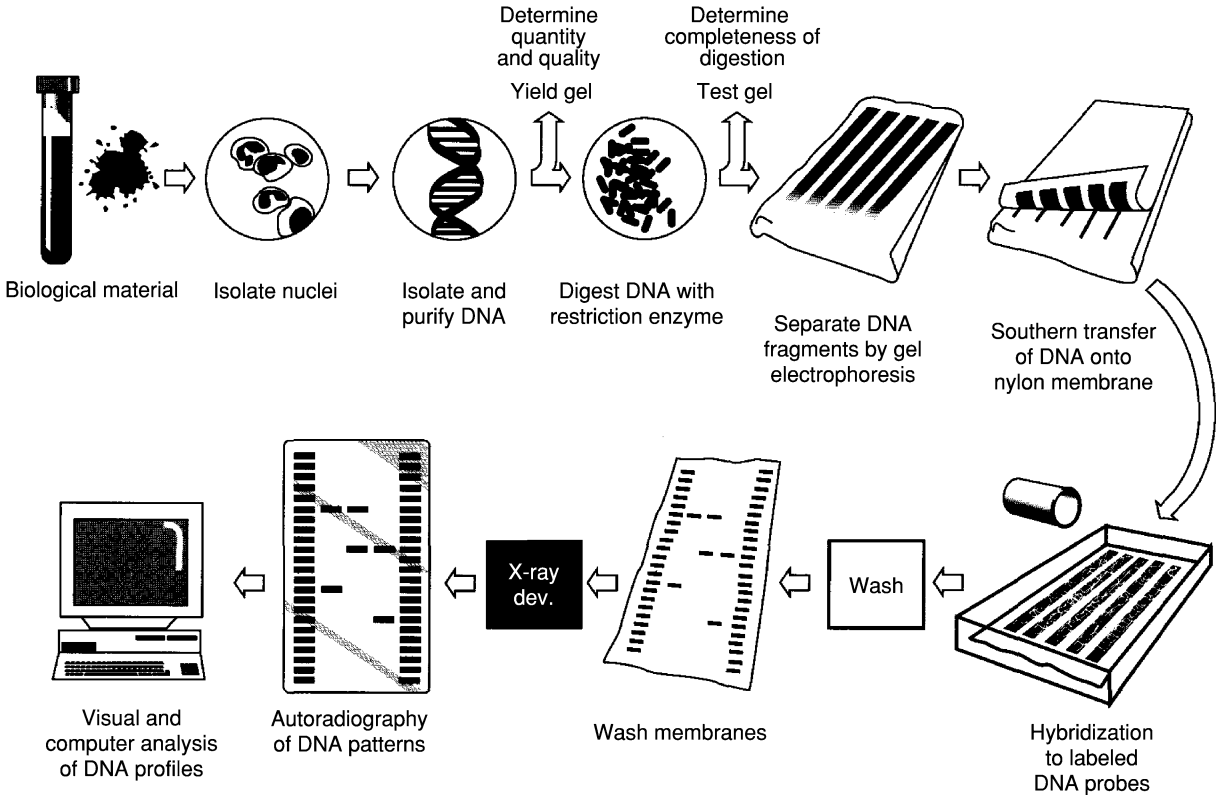

Figure 1 Diagram illustrating restriction fragment length polymorphism (RFLP) analysis used for DNA profiling testing. Reprinted with permission from Baird ML (1993) Quality control in DNA profiling tests. In: Farkas DH (ed.) *Molecular Biology and Pathology: A Guidebook for Quality Control,* p. 203. San Diego, CA: Elsevier.

so frequently that sometimes small fragments run right off the end of the gel. Hinf I is from *Haemophilus influenzae* and recognizes G^ANTC/CTNA^G and cleaves between A and G. REs can exhibit star activity (cut at inappropriate sites) if the conditions of the digestion vary from the optimal, as indicated by the manufacturer.

Each fragment will be assigned a size in kilobases by comparing the fragment to a known sizing ladder. The ladder is loaded into several wells of the gel about five lanes apart so that it is easier to identify the correct size. In Figure 2, one can see that some fragments migrate to nearly the same place on the gel and to be sure that the bands are identical in paternity testing, the AF and Ch's DNA will be combined and run in one lane. When the bands of the AF and Ch appear as one, they are the same. When an indentation occurs between them at the edge of the bands, they are different. Band sizes will vary depending on the RE used and variability in electrophoresis procedures and conditions. A band that is 2.0 kb at one lab using Hae III may be the same allele as one at 1.97 or 2.03 on a different day or a different lab. Therefore, the range of sizes (often 10% of average) that a laboratory decides is not separable is classified as one band and called a bin. In parentage testing, bins and running two samples of DNA together help cover for errors.

PCR Beginning in 1994, PCR testing for parentage markers has increased the speed and inclusion power of molecular testing. The markers used include D1S80, a VNTR that is an LTR, and numerous STRs. Like all DNA testing of noncoding regions, the mutation rate must be included in deciding which alleles to test and what conclusions may be made from the results. Sequence-specific primers under special conditions of high stringency (the primers bind only exactly matched complementary sequences) allow for the amplification of specific alleles without transferring and hybridizing probes. The amplicons (amplified DNA target sequence) are analyzed by electrophoresis and gels stained with fluorescent stains to visualize by eye and/or machine. Newer methods use fluorescent tags on the primers so that the fluorescing amplicons can be directly evaluated by an optical scanner or charge coupled device camera. These results may be analyzed by human eye or machine and calculated for paternity index (PI) by hand or computer. An alternative method for coding alleles involves the amplification of the gene region

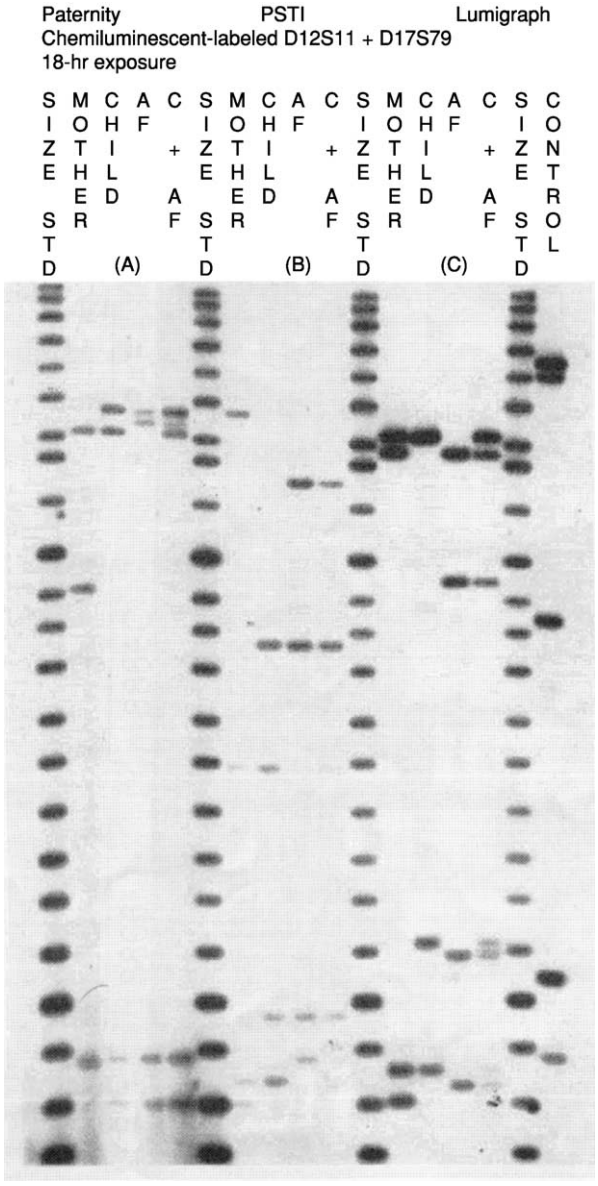

Figure 2 Photograph of a lumigraph from the analysis of three paternity trios A, B, and C. DNA was isolated from blood samples from the mother, child, and alleged father in each trio, digested with *PstI*, size-separated by electrophoresis, transferred to a nylon membrane, and hybridized with alkaline phosphatase-labeled, chemiluminescent probes which recognize the D12S11 (top) and D17S79 (bottom) loci. The size standard (STD) lanes contain fragments of known size used to measure the fragment lengths of the bands in the sample lanes. The control lane contains human DNA from cell line K562 digested with *PstI*. The alleged fathers in trios A and B are included while the alleged father in trio C is excluded as the biological father. C, child; AF, alleged father. Reprinted with permission from Baird ML (1993) Quality control in DNA profiling tests. In: Farkas DH (ed.) *Molecular Biology and Pathology: A Guidebook for Quality Control*, pp. 208–209. San Diego, CA: Elsevier.

using generic flanking primers and then reverse dot blots with allele-specific probes to determine phenotypes. These are usually read colorimetrically by hand by two independent observers. These markers include HLA-DQA1, low-density lipoprotein receptor (LDLR), glycophorin A (GYPA), hemoglobin G (HBGG), D7S8, and GC. GYPA will show the same data as for the classical serological testing of MNS system antigens, and thus should not be considered an independent marker. In addition, HLA-DQA1 is closely linked to HLA-A and HLA-B and must be figured separately when using both types of data.

Amplified fragment length polymorphisms from LTR: D1S80 or STRs have become the most widely used platforms for paternity testing up to 2003. Multilocus chips have been created that detect dozens of different alleles at once; therefore, the throughput has increased dramatically.

In May 2003, the AABB approved the used of SNP analysis called SNP-IT tag array (formerly called APEX). It is highly automated and fast in analyzing cases and has increased the number of cases that may be analyzed per day tremendously. SNP analysis has the potential to supplant STRs as the preferred multilocus testing procedure.

SNPs occur about every 100 bases in the human genome and many occur at a frequency of 0.5. There are only two alleles at each locus and the required number of tested loci must be large to equal the power of inclusion of five RFLPs or 10 LTRs or STRs. Suggested platforms use 30 as the starting number of loci and additional ones when the first 30 do not result in an adequate inclusion of paternity value. Automated testing for SNPs is now approved for use in parentage testing. Microchips with multiple SNP allele detection systems can determine parentage in a couple of hours. Figure 3 shows the results from a paternity test using APEX (now SNP-IT technology).

SQ is finding the exact spelling of the AGCT nucleotides in a piece of DNA. With the completion of the Human Genome Project, polymorphisms are being discovered rapidly. These can be used for all sorts of identity testing. New small rapid sequencers are now available, making this method of parentage testing practical.

Mitochondrial DNA Mitochondrial DNA is passed from the maternal ovum to children. When parents are unavailable and family matching to grandparents is possible, mitochondrial DNA from the grandmothers can be used to determine family linkage. A segment of noncoding DNA is amplified and sequenced. This has been used in South America where children were stolen from parents who were killed for political reasons, and to verify the remains of the family of Tsar Nicholas' family. The South American children were returned to their grandparents by matching not only chromosomal but mitochondrial DNA.

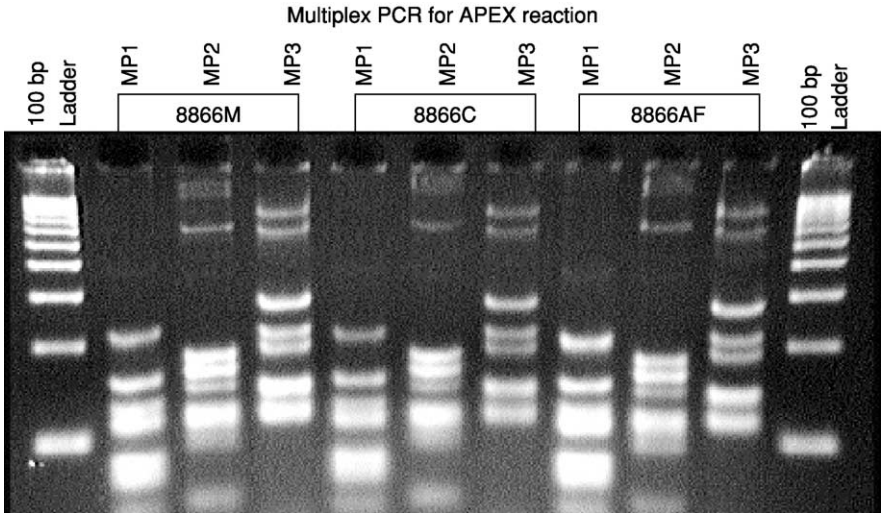

Figure 3 Image of ethidium bromide-stained 4% agarose gel with multiplex polymerase chain reaction (PCR) amplicons from a single case (#8866 TRIO). Three multiplex reactions were performed on genomic DNA from each individual in the case. MPI, multiplex 1 (ADH3, ARSB, LDLR, METH, PROS1, PRP, HSD3B, LPL, IGF2, BCL2); MP2, multiplex 2 (WI-1417, D3S2344, D2S1301, D7S1760, FUT1, DUF-1, TCRVB17); MP3, multiplex 3 (TCRVB12, DNASE1, CETP-1, CPT1, APOC3, CA2, COL2A1). 8866M, mother; 8866C, child; 8866AF, alleged father. Reprinted with permission from White LD, Shumaker JM, Tollett JJ, Staub RW (1998) *Human Identification by Genotyping Single Nucleotide Polymorphism (SNPs) Using an APEX Microarray.* Genetic Identity Conference Proceedings: Nineth International Symposium on Human Identification. Available at: http://www.promega.com.

Potential Errors

Many opportunities for errors occur in parentage testing. Several safeguards are built into specimen collection and analysis to catch most of these errors. The errors before and during testing include: (1) interchange of samples at collection accidentally or on purpose (fraud); (2) interchange of samples at the lab; (3) typing errors; (4) deficient test sera or technique; (5) deteriorated samples; (6) clerical errors; and (7) rare variants commonly missed by routine testing. During the interpretation phase, an analyst may disregard linkage disequilibrium, unbalanced gene frequencies, and silent alleles in the calculations, resulting in skewed calculations. Many of these problems are alleviated by automation; some are not. Prevention measures for errors at collection must be in place, including having a witness by a disinterested party, two independent specimen collections, photos, fingerprints, and signatures at collection. In the laboratory, errors may be prevented by using controls for antisera and other reagents, independent double testing using antisera from different sources, independent checking of results, and testing of duplicate samples. During analysis, errors may be prevented by consideration of the possibility of rare variants (Table 7) in the testing and analysis. For interpretation of the results, understanding the relationships between the people sampled including consanguinity, incest, and the ethnicities of all parties is important in selecting gene frequency databases. Correct evaluation of phenotypes and usage of more than one system to achieve an outcome are also essential. Evaluation of the probability of paternity of nonexcluding systems is also an important step. Further testing may be indicated and estimating the error risk is also required as a part of the final report.

Table 7 Rare variants of red blood cell (RBC) antigens

RBC system	*Examples of variants*
ABO	*cis*-AB, A_3, A_x, A_{el}, B_x
Duffy	Fy3, Fy4, Fy5, Fy6, Fy_x
Kell	Ko, Kx, KL
Kidd	Jk, Jk^3
MNS	M^g, M^k
Rh	C^w, hr^B, C^G

Analysis

For analysis of markers of expressed proteins on RBCs, white cells, or plasma, a single nonmaternal marker carried by the Ch that an AF does not carry is classified as an exclusion. Direct and indirect exclusions are terms used in classic protein-based testing. The markers used in these studies have well-established inheritance patterns. Using DNA technology, the mutation rate is too high to say that one nonmaternal band not shared by the AF is enough

to exclude; therefore, the term is mismatch and two mismatches are required to exclude an AF or declare nonmaternity. Several online databases exist, such as GenBank and ALFRED (allele frequency database for diverse populations and DNA polymorphisms) containing allelic frequencies for DNA markers for different populations.

The terms used for all paternity testing include PI, probability of paternity (PP) and probability of exclusion (PE), and random man not excluded (RMNE). Other terms sometimes used are cumulative paternity index (CPI) and cumulative probability of paternity (CPP). The PI refers to the ratio of the chance that the AF passed the gene to the chance that a random man could have passed it to the child. The cumulative PI is found by multiplying the PI for each locus. The PP includes the nongenetic evidence (prior chance) into the equation. It compares the AF chance to pass a gene to that of an untested man of the same race. PE is based exclusively on the maternal and child allelic frequencies. PE shows the chance that a falsely accused man would be excluded and is calculated as 1 − RMNE.

Direct Exclusion

There are two ways to have a direct exclusion: (1) when a Ch has a marker that neither the Mo nor AF carries; (2) the AF has two markers and the Ch has neither of them. An example of the first situation includes a Ch that is D-positive, the Mo is D-negative, and the AF is D-negative. The D must be contributed by one of the two parents of the Ch, therefore the AF is excluded. An example of the second situation includes an AF with Fy^a and Fy^b and a Ch who is Fy(a−b−) for the Duffy antigens on RBCs. The Ch would carry at least one of these markers if the AF was the biological father. Nonpaternity is established by one of these direct exclusions.

Indirect Exclusions

When a Ch has a single marker and the AF has a different single marker, homozygosity or a silent allele may be present at that locus. For an indirect exclusion the assumption is that both Ch and AF are homozygous. If there is a silent allele, the AF may indeed be the biological father and the test system cannot detect the marker. Suppose the Ch is Jk(a+b−) and assumed homozygous for Jk^a. The AF is Jk(a−b+) and assumed homozygous for Jk^b. There is a rare null allele *Jk* that is found in some parts of the Pacific Islands, Brazil, and in a few white families. The AF could be *Jk* Jk^b and the Ch *Jk* Jk^a. Thus the AF could be the true father. This allele is rare, thus, unless special circumstances are present in the racial profile or AF origin, an indirect exclusion would be accepted.

False Exclusions

When one of the tested parties has a chimera, dispermy, an unlinked suppressor gene, no precursor for another gene to act upon, or a significant mutation, an apparent direct exclusion may be false. In the ABO system, H substance is required for the expression of glycosyl transferase genes that modify H substance to become A or B or AB. Although very rare, some individuals do not have H substance. They are genetically *hh*. They can pass A or B genes to offspring that they do not express. Thus the mother may be O, the Ch B, and the father types as an O (genotype *hh*, *BB*) by both forward and reverse typing. The mother would provide an H for the Ch to be able to express the B gene product. This situation requires an additional testing with O cells in the reverse typing (not routinely used). The father and all others with Bombay phenotypes would have strong anti-H activity that would agglutinate H+ O cells, whereas a typical group O serum would not agglutinate group O cells.

For the Rh system it is important to type for all typical antigens to help rule out the possibility of a null phenotype. Someone could type D-negative and really be D-positive. The person may carry a suppressor gene located on an unlinked locus. The D gene may be passed without the suppressor gene so that the Ch successfully expressed the gene. Full typing of the DCcEe reveals the suppressor gene. The person would appear not just D-negative but C−c−E−e− as well.

False indirect exclusions occur because of weak, untyped, or silent alleles. Table 7 lists unusual variants of the typical alleles typed for parentage testing. Silent or weak or unusual variants occur in most RBC antigen systems. A single indirect exclusion is not adequate to establish nonparentage. Not all typing sera detect all variations of expression of weak alleles and most parentage testing does not test for rare variants like Ko, Fy4, or Fy5.

Ethnicity must be taken into consideration for antigen systems that are expressed in extremely different frequencies depending on the racial background. Blacks express a silent allele on RBCs, *Fy*, at a very high frequency of 82%. Middle Eastern populations and hispanics also express this silent allele due to natural selection in populations where malaria is endemic; Fy(a−b−) RBCs are not infected by *Plasmodium vivax*, the most common cause of malaria worldwide. The silent *Fy* is extremely rare in whites. Therefore, an indirect exclusion in a white trio with a

Ch who is Fy(a−b+) and an AF typing Fy(a+b−) would have a high level of confidence while the same phenotype in a black trio would not lead to the same conclusion. The AF is likely to be genetically Fy^aFy, the Ch Fy^bFy, and true paternity is highly possible.

The MNS system also has peculiarities based on linkage, linkage disequilibrium, and ethnic frequencies. M and N are antigens on glycophorin A, and S, s, and U are on glycophorin B. They are closely linked and only extremely rare recombinants or mutations have been described. Therefore, these linked loci are passed on as haplotypes of *MS*, *Ms*, *NS*, or *Ns* to our offspring. An illustration of linkage disequilibrium in the MNS system occurs because the frequency of *NS* (7%) is much less frequent than *Ns* (38%). The frequencies of the other two haplotypes are similar, *MS* (25%) and *Ms* (30%). Although rare phenotypes occur in other ethnic populations, the U− phenotype is present in 1% of blacks. Someone typing S−s− may also be U−.

A sample case using the MNS system follows. The Mo phenotypes MNSs and the undisputed biological father of the first three children is MSs. The Mo claims he is also the father of the fourth child. The children's phenotypes are as follows: Ch1 is MNS, Ch2 is MNS, Ch3 is MSs, Ch4 is MNs. The haplotypes passed by the mother must be *NS* and *MS* by the father for the first two. *Ms* is passed by the Mo for the third child and *MS* by the father. The fourth child received *Ms* from the mother so the biological father would have to pass *Ns* to Ch4. The biological father of the other three children is excluded as the father of the fourth child in this example.

Inclusion of Paternity

The Hardy–Weinberg formula and assumptions are used to calculate the chance of true paternity. There is a predictable relationship between observed allele frequencies and gene frequencies at any locus. This allows for the estimate of genotype frequency using the formula by multiplying the individual allele frequencies from one person's observed phenotype as follows:

$$p^2 + 2pq + q^2 = 1$$

where p = frequency of allele 1; q = frequency of allele 2; $2pq$ = frequency of heterozygotes (allele 1, allele 2); p^2 = homozygote for allele 1; and q^2 = homozygote for allele 2.

Linkage equilibrium is also assumed for these loci. Linkage equilibrium is established when haplotype frequencies match the expected frequency for each independent gene frequency multiplied. One key assumption is that the frequency of a phenotype with many alleles is the product of the individual allelic frequencies. The Hardy–Weinberg assumptions and linkage equilibrium do not always apply. Certain ethnic groups in a large population of multiple ethnic groups are more likely to mate within their group. However, in practice the formulas have been shown to be valid.

Paternity Index

Inclusion criteria for an AF (X) involves multiplying allele frequencies for each of the loci by each other and dividing by the chance these could be from a random man (Y) to arrive at a PI or system index. Table 8 shows the formulas to determine PI for various combinations of maternal phenotypes, child phenotypes, and AF phenotypes. The result gives an idea of the likelihood that the AF is the biological father.

The following is a sample case calculating the PI. Suppose, for a white trio, that the mother is type MS

Table 8 Formulas used to calculate paternity indices in biallelic SNP systems

Combination	*M*	*C*	*AF*	*PI*
1	X	X	X	$1/x$
2	X	X	XY	$0.5/x$
3	X	XY	XY	$0.5/y$
4	XY	XY	XY	$1/(x+y) = 1$
5	XY	X	X	$1/x$
6	XY	XY	X	$1/(x+y) = 1$
7	XY	X	XY	$0.5/x$
8	X	XY	Y	$1/y$
9	Not tested	X	X	$1/x$
10	Not tested	XY	X	$0.5/x$
11	Not tested	XY	XY	$0.25(x+y)/xy = 0.25/xy$
12	Not tested	X	XY	$0.5/x$

PI formulae for the eight possible combinations of SNP phenotypes in paternity trios and four possible combinations of SNP phenotypes where the mother is not tested. Homozygosity is assumed when individuals only display a signal for one nucleotide at any SNP locus. X and Y represent the nucleotide signals obtained while x and y are the gene frequencies for the respective SNP alleles. It should be noted that the calculations for combinations 10 and 11 assume that the mother and the alleged father are of the same race (as they are in all cases presented here). If they differ in race, the PI formula for Combination 10 becomes $PI = x_M/x_{AF}y_M + x_My_{AF})$ where x_M = freq(X) from mother's racial database, x_{AF} = freq(X) from AF's racial database, y_M = freq(Y) from mother's racial database and y_{AF} = freq(Y) from AF's racial database. Likewise, the PI formula for combination 11 becomes $PI = (x_M + y_M/[2(x_{AF}y_M + x_My_{AF})]$. Formulae derived according to Brenner in *Transfusion* (1993) 33:51–54. Reprinted with permission from White LD, Shumaker JM, Tollett JJ, Staub RW (1998) *Human Identification by Genotyping Single Nucleotide Polymorphism (SNPs) Using an APEX Microarray*. Genetic Identity Conference Proceedings: Nineth International Symposium on Human Identification. http://www.promega.com.

Table 9 Genotypes of individuals scored from 24 loci on identity chip

			Trio					*MNT*			
SNP locus	*Polym*	*%*	*8875M*	*8875C*	*8875AF*	*8875 PI calc.*	*8875 PI*	*8879C*	*8879AF*	*8879 PI calc.*	*8879 PI*
ADH3	A-G	56	AA	AA	AA	1/0.56	1.79	AA	AG	0.5/0.56	0.89
ARSB	A-G	67	GG	GG	FF	1/0.33	3.03	AG	AG	1/(4*0.67*0.33)	1.13
LDLR	T-C	45	TC	TC	TC	1	1.00	TC	TC	1/(4*0.45*0.55)	1.01
METH	T-C	58	TC	TC	TT	1	1.00	CC	CC	1/0.42	2.38
PROS1	T-C	58	TT	TT	TC	0.5/0.58	0.86	TC	TC	1/(4* 0.58* 0.42)	1.03
PRP	A-G	66	AG	AG	AA	1	1.00	AG	AG	1/(4* 0.66* 0.34)	1.11
HSD3B	A-C	77	CC	CC	AC	0.5/0.23	2.17	AC	AC	1/(4*0.77*0.23)	1.41
LPL	A-G	52	AG	AG	GG	1	1.00	GG	GG	1/0.48	2.08
IGF2	A-G	20	AA	AA	AG	0.5/0.2	2.50	AG	AG	1/(4*0.2*0.8)	1.56
BCL2	A-G	56	AA	AA	AA	1/0.56	1.79	AG	AA	0.5/0.56	0.89
W1-1417	C-T	48	TC	TC	TT	1	1.00	TT	TC	0.5/0.52	0.96
D3S2344	G-C	48	GC	GC	GG	1	1.00	CG	CG	1/(4*0.48*0.52)	1.00
D2S1301	G-A	55	AG	AG	GG	1	1.00	AG	AA	0.5/0.45	1.11
D7S1760	T-C	50	TT	TC	CC	1/0.5	2.00	TT	TC	0.5/0.5	1.00
DNASE1	A-G	56	AA	AA	AG	0.5/0.56	0.89	GG	GG	1/0.44	2.27
CETP-1	C-A	53	CC	CC	CC	1/0.53	1.89	CC	CC	1/0.53	1.89
FUT1	A-T	50	AA	AA	AA	1/0.5	2.00	AA	AA	1/0.5	2.00
DUF-1	A-G	41	AG	AA	AA	1/0.41	2.44	AA	AG	0.5/0.41	1.22
TRCVB17	C-T	48	TC	TC	TT	1	1.00	TT	TT	1/0.52	1.92
TCRVB12	C-T	53	TT	TT	TT	1/0.47	2.13	TT	TT	1/0.47	2.13
CPT1	G-A	49	AG	GG	AG	0.5/0.49	1.02	AG	AG	1/(4*0.49*0.51)	1.00
APOC3	T-C	58	TT	TT	TC	0.5/0.58	0.86	TT	TT	1/0.58	1.72
CA2	T-C	50	CC	CC	TC	0.5/0.5	1.00	CC	CC	1/0.5	2.00
COL2A1	C-T	48	TC	CC	CC	1/0.48	2.08	TC	TC	1/(4*0.48*0.52)	1.00
Cumulative PI							2901.30				1802.14

Reprinted with permission from Toll White LD, *et al.* (1998) *Human identification by genotyping single nucleotide polymorphism (SNPs) Using an APEX microarray*. Genetic Identity Conference Proceedings: Nineth International Symposium on Human Identification. Available at: http://www.promega.com.

for the RBC antigen system, the Ch is group MSs, and the father is group Ms. Since these are inherited as haplotypes, the mother contributed *MS* so the father contributed *Ms* and the chance for AF is 1 for contributing the *Ms* haplotype. A random man with this haplotype occurs at a frequency of 0.291. Therefore the PI (*X*/*Y*) is 1/0.291 = 3.43. In the same trio for D1S80, the mother is 24, 33, the Ch is 15, 24, and the AF is 15, 18. The AF has a 0.5 chance of passing either the 15 or the 18 to a Ch. The population frequency of the LTR D1S80 allele 15 is 0.0015. Therefore $X/Y = 0.5/0.0015 = 333.33$. In this same trio for HUMFESFPS the mother is 10, 11, the Ch is 10, 11, and the AF is 10, 12. The mother could contribute the 10 to the Ch and then the chance for the AF for contributing the 10 allele is 0, the chance for a random man to contribute the 10 allele is 0.225. If the mother contributed the 11 allele then the chance for the AF contributing the 10 is 0.5 and a random man's chance is 0.284. Therefore, $X = 0.5 \times 0.5 + 0.5 \times 0 = 0.25$. $Y = 0.5 \times 0.225 + 0.5 \times 0.284 = 0.2545$. $X/Y = 0.25/0.2545 = 0.982$. Since these are not linked loci, the PI is the product of these systems. $3.43 \times 333.33 \times 0.982 = 1122.7$. The AF is 1122.7 times more likely to have contributed these alleles to the Ch than a random man. Table 9 has examples using SNP analysis to determine PI in a case with a maternal sample available and without. Table 10 illustrates the calculations used when no maternal sample is available, the race of the father is either black or white, and LTR and STR loci are used.

Table 10a A sample, no mother case: phenotypes

System	*Child*	*Alleged father*
DIS80	17, 18	18, 21
YNZ	4, 5	4
FGA	3	2, 3
P450	2, 6	2, 6
THO1	6	6
VWA	3, 4	4, 5
FES	1, 3	3, 4
CSF	4, 6	4, 7

Probability of Paternity

A statistical theory created by Bayes uses an estimate of previous probability to come up with a percentage of paternity probability. This is based on social

Table 10b A sample, no mother case: allele frequencies[a]

System	*Allele*	*White*	*Black*	*System*	*Allele*	*White*	*Black*
DIS80	17	0.004	0.044	THO1	6	0.224	0.135
	18	0.248	0.075				
				VWA	3	0.115	0.257
YNZ	4	0.293	0.105		4	0.197	0.282
	5	0.043	0.102				
				FES	1	0.008	0.099
FGA	3	0.323	0.319		3	0.320	0.221
P450	2	0.131	0.477	CSF	4	0.037	0.133
	6	0.379	0.101		6	0.328	0.271

[a]Allele frequencies modified from data supplied by Laboratory Corporation of America.

Table 10c A sample, no mother case: results. There are three calculations: the first is for the case in which mother and alleged father are both White; in the second case they are both Black; in the third case the mother is Black and the alleged father White

System	*X*	*Y*	*PI*
Mother White, alleged father White, alternative father White			
DIS80	0.002000	0.001984	1.01
YNZ	0.043000	0.025198	1.71
FGA	0.161500	0.104329	1.55
P450	0.255000	0.099298	2.57
THO1	0.224000	0.050176	4.46
VWA	0.057500	0.045310	1.27
FES	0.004000	0.005120	0.78
CSF	0.164000	0.024272	6.76
Combined PI			205
Mother Black, alleged father Black, alternative father Black			
DIS80	0.022000	0.006600	3.33
YNZ	0.102000	0.021420	4.76
FGA	0.159500	0.101761	1.57
P450	0.289000	0.096354	3.00
THO1	0.135000	0.018225	7.41
VWA	0.128500	0.144948	0.89
FES	0.049500	0.043758	1.13
CSF	0.135500	0.072086	1.88
Combined PI			1042
Mother Black, alleged father White, alternative father White			
DIS80	0.022000	0.011212	1.96
YNZ	0.102000	0.034401	2.97
FGA	0.159500	0.103037	1.55
P450	0.289000	0.194014	1.49
THO1	0.135000	0.030240	4.46
VWA	0.128500	0.083059	1.55
FES	0.049500	0.033448	1.48
CSF	0.135500	0.053651	2.53
Combined PI			346

evidence such as fertility of the AF, possible access to the mother at the time of conception, whether other male family members might have had the opportunity, and other factors. Testing laboratories do not have access to these data and assign 0.5, assuming the AF being tested and any other man would have an equal chance of being the father. The PI is used for the calculation as well. The final $PP = PI/(PI + 1)$. For the previous example, $PP = 1122.7/(1122.7 + 1) = 1122.7/1123.7 = 99.9\%$. In the USA, the range of required PI and PP in different states is from "20 to 1 and 95%" to "1000 to 1 and 99.9%," respectively. In our test case, the AF would be held liable for child support in any state.

Probability of Exclusion

The probability of exclusion is determined to put a value on the chances of excluding a man based on the phenotypes of the Mo and Ch. Calculate the number of men who would not be excluded (RMNE) and subtract that number from 1. The exclusion formula excludes all males who do not have the paternal allele on either chromosome and this equals $(1 - (\text{frequency of paternal allele}))^2$. RMNE = 1 − exclusion value. In the sample case for Ms, the probability of exclusion is 0.7054. For D1S80, the exclusion value is $(1 - 0.0015)^2$, thus the RMNE is $1 - 0.997 = 0.003$. For HUMFESFPS, the exclusion value is $(1 - 0.284)^2$, the RMNE is $1 - 0.513 = 0.487$. The cumulative RMNE is $0.705 \times 0.003 \times 0.487 = 0.001$. The cumulative power of exclusion is $1 - \text{RMNE} = 0.999$ or 99.9%.

Calculating a PI is more challenging when there are homozygous alleles or silent alleles that are possible in the population in question. When only one of two potential co-dominant alleles is expressed or found, the possible options include homozygosity, a silent allele, or the system does not pick up some rare allele. The report of the case should indicate the possible errors due to these factors.

Other issues that must be addressed when setting up a parentage testing service and the analysis of samples include the complexity of testing procedure, amount and type of sample required, the complexity

of calculations, and the fragility of marker tested. RBC enzyme procedures, protein electrophoresis, HLA microlymphocytotoxicity testing, and RFLP analysis require many days to perform and highly skilled personnel such as a credentialed clinical laboratory scientist. While RBC antigens are easier to test for and less expensive in most cases, the analysis of the data can be quite challenging. Knowledge of rare alleles and variation in quality of antisera and antigens on the cell surface is important to a successful analysis. The background for these esoteric facts comes with a specialty in transfusion medicine also known as blood banking. For HLA testing, knowledge of cross-reactive groups and linkage disequilibrium is required to analyze a case correctly. HLA genes are all linked, but certain haplotypes (all tested genes close to each other on the same chromosome) are more frequent than would be expected if gene frequencies of each gene were multiplied. These situations must be carefully addressed for each case. In addition, HLA proteins show cross-reactivity with each other in serological systems. Cross-reactive groups have been established and the analysis must include each of these potential gene frequencies. Ethnicity must be factored in when known.

See Also

Anthropology: Role of DNA; **Blood Grouping**; **Children:** Legal Protection and Rights of Children; **Crime-scene Investigation and Examination:** Collection and Chain of Evidence; **DNA:** Basic Principles; Risk of Contamination; Mitochondrial; Postmortem Analysis for Heritable Channelopathies and Selected Cardiomyopathies; Hair Analysis; **Evidence, Rules of**; **Identification:** Prints, Finger and Palm; **Immunoassays, Forensic Applications**; **Mass Disasters:** Principles of Identification; **Serology:** Overview; Blood Identification

Further Reading

American Association of Blood Banks (2001) *Standards for Parentage Testing Laboratories,* 5th edn. Bethesda, MD: American Association of Blood Banks.

American Association of Blood Banks (2002) *Guidance for Standards for Parentage Testing Laboratories,* 5th edn. Bethesda, MD: American Association of Blood Banks.

Baird ML (1993) Quality control in DNA profiling tests. In: Farkas DH (ed.) *Molecular Biology and Pathology: A Guidebook for Quality Control*, pp. 201–215. San Diego, CA: Academic Press.

Brecher ME (ed.) (2002) *AABB Technical Manual,* 14th edn. Bethesda, MD: American Association of Blood Banks.

Gaensslen RE (1983) *Sourcebook in Forensic Serology, Immunology, and Biochemistry.* Washington, DC: US Government Printing Office.

Harrison CR (1999) Parentage testing. In: Harmening DM (ed.) *Modern Blood Banking and Transfusion Practices,* 4th edn., pp. 507–519. Philadelphia, PA: FA Davis.

Inman K, Rudin N (1997) *An Introduction to Forensic DNA Analysis*. Boca Raton, FL: CRC Press.

Issitt PD, Anstee DJ (1998) *Applied Blood Group Serology,* 4th edn. Durham, NC: Montgomery Scientific.

Mollison PL, Engelfriet CP, Contreras M (1997) *Blood Transfusion in Clinical Medicine,* 10th edn. Oxford, UK: Blackwell Scientific.

Parker PP, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: choosing and using molecular markers. *Ecology* 79: 361–382.

Race RR, Sanger R (1975) *Blood Groups in Man,* 6th edn. London: Blackwell Scientific.

Strachan T, Read AP (1996) *Human Molecular Genetics.* New York: BIOS Scientific/John Wiley.

Sussman LN (1976) *Paternity Testing by Blood Grouping,* 2nd edn. Springfield, IL: Charles C Thomas.

Walker RH (1983) *Inclusion Probabilities in Parentage Testing.* Arlington, VA: American Association of Blood Banks.