Genetic Kinship Investigation from Blood Groups to DNA Markers

Gunther Geserick a  Ingo Wirth b

a Berlin,
bPolice Academy of the State of Brandenburg, Oranienburg, Germany

Summary
The forensic application of hereditary characteristics became possible after the discovery of human blood groups by Karl Landsteiner in 1901. The foundation for their use in kinship investigation was laid by Emil von Dungern and Ludwig Hirschfeld in 1910 by clarification of the inheritance of the ABO groups. Up to the middle of the 20th century further red cell membrane systems were discovered. From the 1920s Fritz Schiff and Georg Strassmann fought for the introduction of blood groups into forensic kinship investigation. A new era of hemogenetics was opened from 1955 as genetic polymorphisms were described in serum proteins. Starting in 1958 there followed the complex HLA system of white blood cells, which from 1963 was joined by polymorphisms in erythrocyte enzymes. Therefore, from the 1980s, it was possible to clarify the majority of kinship cases with a combination of conventional markers. From 1990 to 2000 the conventional markers were gradually replaced by the more effective DNA markers. Simultaneously typing shifted from the phenotype level to the genotype level. The genomic structure of conventional genetic markers could also now be explained. As a reflection of scientific progress the legal situation also changed, particularly in the form of the official guidelines for kinship investigation.
Introduction

Nowadays, a broad spectrum of efficient genetic markers is available to paternity experts which satisfy high quality demands in the following points: i) Formal genetics: Mendel’s inheritance of defined and if possible codominant alleles (genes) in discrete and well differentiated expression; if possible no (or only low) presence of mutations and silent alleles. ii) Population genetics: knowledge of genotype and phenotype frequencies in defined populations in which the polymorphism is in Hardy-Weinberg equilibrium and random mating occurs. iii) Internal and external reproducibility: the basis for this are standardized nomenclature and laboratory techniques, general availability of materials and reagents, as well as continuous internal and external quality control. Under the following suitable systems a selection according to cost-benefit criteria can be carried out. It was a long way before this state of knowledge and this comfortable situation in the practice for expert assessment had been reached. During the past 100 years thousands of scientists and experts in many countries have made uncountable contributions to the current structure of knowledge. Therefore, in this article, the development can only be briefly outlined, and the literature cited must be limited to milestones and selected reviews.

Discovery and Nomenclature of the ABO Blood Groups

The first clear indication of human blood groups came from Karl Landsteiner [1] (fig. 1) who in 1900 described the agglutination effect of ‘serum from healthy individuals (...) on human blood cells from other individuals’ [2]. He specified his observation 1 year later and divided the investigated blood samples into 3 groups which he named A, B, and C [3]. Landsteiner was born on June 14, 1868 in Baden near Vienna. He studied medicine in Vienna from 1885 to 1891. Following further clinical training he became an assistant at the Hygiene Institute in Freiburg/Breisgau. In 1906 he became the title of Professor of Pathological Anatomy in 1911. From 1908 to 1919 Landsteiner was director of a municipal institute in Vienna, and following that he was prosecutor in den Haag (Holland) until 1921. In 1923 he was appointed to the Rockefeller Institute for Medical Research in New York where he worked even after 1939 as emeritus until his death. Landsteiner died on June 26, 1943 in New York. The climax of his career was being awarded the Nobel Prize for Medicine in 1930. In addition to his studies on blood groups and further immunohematological topics his scientific work encompassed studies on poliomyelitis, paroxysmal hemoglobinuria, and syphilis [1].

The 4th blood group, now known as AB, was described by his coworkers von Decastello and Sturli in 1902 [4]. The current nomenclature of blood groups with the capital letters A, B, 0 (formerly C), and AB was formulated by von Dungern and Hirschfeld in 1910 [5, 6]. They designated the group most common in Europe as A, the rarer group as B. The blood group whose red cells were not agglutinatable was designated as 0. The group resulting from the simultaneous existence of A and B was therefore called AB. In addition, von Dungern and Hirschfeld’s suggestion to introduce symbols for the isoagglutinins was adopted: \( \alpha \) (alpha) for anti-A and \( \beta \) (beta) for anti-B. Von Dungern and Hirschfeld also discovered the A subgroups (A1, A2) [5–7]. Since the mid-1930s many rare markers of the ABO system have been discovered (reviews in [8–10]).

Inheritance and Biochemistry of the ABO System

Astonishingly the forensic introduction of blood groups in kinship analysis first occurred 2 decades after their use in stain testing [11, 12]. Obviously Landsteiner had not initially thought about the inheritance of the serological blood markers discovered by him. Compelling evidence for their hereditary nature was published by Epstein and Ottenberg in 1908 [13].

A convincing genetic theory was published by von Dungern and Hirschfeld in 1910 [5], who demonstrated in 72 families that dominant (A, B) and recessive (0) markers were inherited according to Mendel’s laws and could therefore be used in forensic investigations. They concluded that A and B had to be characteristics that segregated independently from each other. In accordance with the genetic thought of the times, the existence of 2 allele pairs was postulated. This laid the foundation for hemogenetic kinship analysis.

Emil Freiherr von Dungern (1867–1943) (fig. 2) worked from 1892 as assistant and from 1902 as professor at the Hygiene Institute in Freiburg/Breisgau. In 1906 he became director of the Biological-Chemical Department of the Institute for Cancer Research in Heidelberg, and from 1913 to 1918 he was director of the Cancer Research Institute in Hamburg-Eppendorf. He finally settled down in Ludwigs-hafen, Lake Constance [1].

Ludwig Hirschfeld (1884–1954) (fig. 3), also known as Ludwik Hirsfeld, worked from 1907 to 1911 in the
In conventional typing there is a formal genetic model of multiple alleles at the ABO gene locus: the genes *A and *B are codominant and dominant to *0, *A1 is dominant to *A2 and *0. The 6 serologically recognizable phenotypes represent 10 different genotypes. The gene locus is on chromosome 9 (review in [15] and gene database [16]).

The biochemical principles of the ABO system have been clarified since the 1950s (reviews in [8, 17]). The membrane properties are terminal carbohydrates: A = N-acetyl-D-galactosamine, B = D-galactose, and 0 = L-fucose. These sugars are however not a primary gene product. The genes responsible determine the presence and the activity of glycosyltransferases which transfer the terminal sugar residues N-acetyl-D-galactosamine or galactose to glycoproteins (for water-soluble blood group substances) or to glycolipids (for membrane-bound substances).

The ABO characteristics could be further categorized in the 1990s using DNA techniques (polymerase chain reaction, PCR), initially by Yamamoto et al. [18]. Soon after, this group and also other research groups produced results on sequencing of rare ABO variants so that many new alleles could be differentiated. Forensic use of the ABO system was finally simplified and accelerated thanks to more advanced PCR techniques. For example, in 2011, Lee et al. [19] announced a rapid direct ABO genotyping method that eliminates the need for DNA extraction, using a multiplex allele-specific primer set for the simultaneous detection of 3 single nucleotide polymorphism (SNP) sites. In 2011, another group, Watanabe et al. [20], developed a simple method for ABO genotyping using a DNA chip. This method is suitable for the simultaneous analysis of ABO genotyping and primate-specific species identification, especially to identify the ABO type of decomposed samples. This methods will however be mainly reserved for forensic stain analysis. In parentage investigation they might be used in special cases.

Further Blood Group Systems of Erythrocyte Membranes

After the ABO groups Landsteiner and coworkers discovered further serologically detectable blood group markers: MN and P [21], as well as the Rhesus system (RH) [22], which also became part of the kinship analysis (reviews in [8–10, 23]).

The MN system was extended from 1947 by the factors S and s. The biochemistry of the MNSs system could be clarified in the 1970s. The genetic differences are due to an amino acid polymorphism of the specific glycoproteins (glycophorin A and B) (reviews in [8, 17]).

The RH system, which was first recognized on the strongest antigen factor D, was revealed to be very complex by the discovery of further factors (C, Cw, C, E, e) and many variants. Through various nomenclatures and different genetic concepts the system was difficult to understand. According to
serologically assumed 2-locus model (D/CcEe) could be supported by sequencing of the cDNA, and for certain forensic cases RH genotyping has become possible using PCR-based methods (review in [28]). Since 1993 the MN group-specific sequences could be analyzed by PCR and be subdivided into new sub-alleles. A current overview of the molecular genetic relationships of the erythrocyte polymorphisms described can be found in the gene databanks [16, 25–27].

Initial Use of Blood Groups in Kinship Analysis

The Berlin laboratory physician Fritz Schiff [29] in particular was responsible for having initiated the use of blood groups in paternity cases [30, 31]. In 1924 he had already demonstrated the lifelong stability and regular inheritance of blood group characteristics in nearly 500 families [32, 33]. The inheritance of blood groups was confirmed by studies in other countries [34–37].

In September 1924 the forensic pathologist Georg Strassmann reported the first legal paternity cases in Germany, which were produced together with Schiff [38]. However, there were no exclusions of paternity. The surgeon Paul Moritsch described in 1926 the ‘first public case’ in which a putative father could be excluded by blood group determination [39]. At this time paternity exclusions by blood group determinations were also reported by the forensic pathologists Böhmer, Nippe and Ziemke [30]. Further reports on serological paternity tests followed by Schiff [40–47] and also from the forensic pathologists Fritz Strassmann [48] and Georg Strassmann [30], from lawyers, such as Hellwig [49] as well as medical officials from the State Investigation Office in Stuttgart [50–53]. Up to 1926 approximately 100 legal orders of evidence on paternity cases were received [50], and 3 years later Schiff cited the number in Germany and Austria as approximately 1,000 [45].
Fritz (Friedrich) Schiff (1889–1940) was briefly employed in Moabit hospital after participating in World War I, and from 1920 at the University of Greifswald. From there he went to Berlin in 1922 where he became head of the Bacteriology-Serology Department at the municipal hospital in Friedrichshain. In addition he taught at the University of Berlin. Because of his Jewish background, in 1935 he lost his position as department head and the teaching post. The next year he immigrated to the USA [29].

Georg Strassmann (1890–1972) received his training initially from his father in Berlin. Afterwards he began as assistant at the University of Vienna where in 1922 he obtained a postdoctoral qualification. The University of Breslau appointed him in 1928 as senior lecturer. At the beginning of the winter semester 1935/1936 he was suspended by the university administration due to his Jewish origins. Later he left Germany and continued his profession in the USA [11].

According to Schiff [32] the predicate ‘probability bordering on certainty’ for the exclusion constellation ‘child A, mother 0, putative father 0’ was valid. Jervell [35] held the opinion that for the same constellation the putative father was ‘excluded with absolutely certain’. The paternity exclusion chance (PEC) for non-fathers was given by Schiff as 16.7% or 17% [42, 45].

Schiff’s publications began to attract increasing interest in English speaking countries from the mid-1920s [54], especially his article from 1929 in *The Lancet* [46]. In 1931 the Italian government approved the forensic application of blood groups, in which Leone Lattes [55] played a large role. Soon after, Ireland (1932) [56] and the USA (1935) followed with the legal application of blood groups. This was not least because of the efforts of Alexander S. Wiener and Philip Levine [57]. In England legal paternity testing by means of blood groups was first approved in 1939 [54].

Up until the 1950s paternity experts were dependent on the use of agglutinatable blood group markers (table 1). Using this combination of markers a cumulative PEC of approximately 70% could be achieved. Even if the weight of evidence was still relatively low, these systems offered an advantage of simple and cheap technology. Notwithstanding, an expert had to have a great deal of practical experience. Also flawless test sera were necessary which were initially manufactured by many experts themselves. Increasing commercially available and quality checked sera as well as erythrocyte test panels began to appear only after World War II.

**Argument over the Evidential Value of Blood Group Expert Opinions and First Quality Demands**

Apparent exceptions from the laws of inheritance were soon found in genetic studies. In connection with this in 1922 Dyke [34] pointed out that serum markers are not always expressed in children in the first months of life. In their family studies in 1924 Schiff and Adelsberger [33] explained some of the deviations by illegitimacy or faulty investigation techniques. This led to demands for an error-free methodology [41, 58, 59] and a theoretical and practical qualification of the experts [60, 61]. Other authors also warned about exercising caution with forensic use [62–69].

In a comprehensive literature review in 1926, which is still worth reading even now, Reinheimer [70] dealt with the chemical and physical stability of isoagglutinins as well as their ontogenetic development. Doubts on the lifelong stability of blood group markers as in a regular mode of inheritance were considered by him to be excluded. He argued for the forensic use in kinship analysis. This recommendation was underlined by Schiff in the same year [40, 71].

However, through the decision of the Prussian Chamber Court from October 11, 1927 the reliability of blood group determination as legal evidence was placed in question [72]. There was resistance from the experts against this court decision, especially from Schiff [44], Böhmer [60, 73], and other experts from Germany, Austria, and Switzerland [67]. The criticized court decision could not however prevent that courts in various regions of Germany still made use of blood group reports and ordered them to be carried out in central institutes [49, 60, 65, 74]. In 1927 the first conviction in Germany due to a blood grouping result occurred in a High Court in Württemberg. A child’s mother accused of perjury was given a prison sentence after exclusion of the putative father (mother A, child AB, putative father A) [74].

In 1930 the Prussian Minister of Justice published the resolution of a committee for blood group research: ‘(…) blood group determination by competent implementation is a reliable investigation procedure which (…) can be advantageously included for the exclusion of paternity’ [68]. A short time later a decision by the Chamber Court from 4 April 1930 corrected its earlier decision [75] ‘(…) and therefore complies with the dominating opinion with respect to the evidential value of blood group investigations. Subsequently the blood grouping procedure has been included to an increasing extent as evidential material for courts’ [66]. This trend was also reflected by the increasing number of investigations made by the University Institute for Legal Medicine in Berlin: from 288 commissions in 1935 the number increased to 861 in 1937.

**Serum Protein Polymorphisms**

In 1955 Smithies [76] opened a new era of hemogenetics with the development of electrophoresis in starch gels and the detection of a genetic polymorphism in haptoglobin (HP).

By using refined electrophoresis techniques (zone electrophoresis also in agarose and polyacrylamide gels and immunoelectrophoresis) and by selective development procedures...
Table 2. Separation of serum protein polymorphisms by isoelectric focusing (IEF)

<table>
<thead>
<tr>
<th>System</th>
<th>Main types (Zone electrophoresis)</th>
<th>Subtypes (IEF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP</td>
<td>1 2</td>
<td>1F 1S 2FF 2FS 2SS</td>
</tr>
<tr>
<td>GC</td>
<td>1 2</td>
<td>1F IS 2</td>
</tr>
<tr>
<td>TF</td>
<td>C</td>
<td>C1 C2 C3</td>
</tr>
<tr>
<td>PI</td>
<td>M</td>
<td>M1 M2 M3</td>
</tr>
<tr>
<td>ORM1</td>
<td>F S</td>
<td>F1 F2 S</td>
</tr>
<tr>
<td>BF</td>
<td>F S</td>
<td>FA FB S</td>
</tr>
</tbody>
</table>

for separated proteins (e.g. by immunoprecipitation), further protein polymorphisms could be recognized (reviews in [8, 23, 77]). Important for kinship testing was that because of the co-dominant inheritance the displayed phenotypes also reflected the genotypes. Many authors carried out population genetic frequency analyses and genetic studies, and some of the systems such as GC, PI, and TF showed a large number of variants the identification of which was only possible by direct comparison. Indications of deficient types were also soon discovered by so-called null alleles. A big advantage of electrophoretic typing was the possibility of documentation of the results by storing the gels or by photographs.

A new dimension of the separation of serum proteins was achieved by isoelectric focusing (IEF). In several systems (HP, GC, TF, PI, ORM1, BF), the main allele could be divided into 2 or 3 subtype alleles, as shown in table 2. The IEF separation in the systems A2HS, BF, F13A, F13B, ORM1, and PLG could be improved by combination with immunofixation or immunoprinting.

The PI polymorphism was also important in clinical medicine because null alleles, some variants, and especially the Z allele were linked to a deficit of alpha-1-antitrypsin (protease inhibitor, PI) which results in a tendency to develop chronic obstructive pulmonary diseases.

Of particular interest were the proteins of the complement system. Polymorphisms were discovered for most of the 11 classical complement proteins (C1Q through C9) as well as components of the alternative pathway (factor B and factor D) and control proteins (C4BP). Taking population frequencies and detection techniques into consideration the systems C3, BF, and C6 became important for the routine practice in kinship analysis (reviews in [8, 78]). For the clarification of very complex cases symposia and workshops on complement genetics were carried out from 1972, the results of which led to publications involving statements on the international nomenclature for C systems, starting with C3 in 1973. For BF it is important that this C protein together with C2 and C4 (C4A and C4B) and the enzyme GLO1 share an adjacent gene locus on chromosome 6 with human leukocyte antigen (HLA) markers (the so-called major histocompatibility complex, MHC).

The immunoglobulin systems GM, KM, AM, and EM required totally different detection techniques (reviews in [8, 79]). Specific antibodies were required which were used in an agglutination inhibition test. In 1956 the factor GM(a) was first discovered which was represented as an IgG molecule. In the subsequent 2 decades many further GM factors were discovered as well as the independent KM (present on all Ig molecules), AM (present only on IgA molecules), and EM (present on IgE molecules) systems. Clarification of their identity, genetics and molecular structure resulted in a very complex system. Understanding of the system was complicated by changes in the nomenclature which changed from an alphabetic to a numerical classification to take the location on the immunoglobulin molecule into account. For daily practice, the most important GM markers were changed from GM(a), GM(x), GM(f), and GM(z) to GM(1), GM(2), GM(4), and GM(17), respectively. By considering the location on gamma-1 heavy chains, GM(a) for example was correctly called G1M(1). The use of different nomenclatures in expert reports and in publications made it very difficult to compare results. Due to the high degree of polymorphism and the relatively good stability of immunoglobulin G, GM markers became successfully established in kinship analysis and especially in stain investigations.

Of the lipoprotein markers AG (LDL fraction), Apo-E (VLDL fraction), and EL (HDL fraction), only the AG system became of some importance in the practice for kinship analysis (review in [8]). The first factor, now known as AG(a1), was discovered with specific human antibodies from transfused patients using immunoprecipitation. Further factors were described: AG(x), AG(z), AG(t), AG(y), AG(c), AG(m), AG(g), AG(d), AG(h), and AG(i). Most markers could be identified as allelic pairs, as for example AG(x) and AG(y). These markers also became useful in paternity testing.

The serum proteins substantially improved the evidential value of kinship analysis, and the total cumulative PEC for the systems listed in table 3 was approximately 94% when the most effective typing method was used for the calculation (PEC value in bold type in table 2). In particular, subtyping of the serum proteins HP, GC, TF, PI, and BF by IEF increased the forensic benefit of serum groups (reviews in [8, 23, 77]). The IEF techniques could also be standardized more easily than the immunogenetic procedures (for the systems GM, KM, AG). However, the technical requirements and practical details were more complicated and therefore became more cost-intensive. The typing of an extensive spectrum of serum proteins was now reserved for large specialized laboratories.

The expert had to, however, bear in mind that some of the serum proteins tested (e.g. HP, GC, GM) could not be detected at the time of birth or were only present in low concentrations which could lead to typing errors. This could be avoided if the children were tested when they were at least 6 months old. A further source of error was the presence of...
The serological typing of HLA antigens was carried out with specific antibodies and various techniques of which the lymphocyte toxicity test according to Terasaki became the most important. In this test viable lymphocytes are incubated with specific antisera and complement. In the presence of the specific antigen on the surface of the lymphocytes the cell then dies which can be recognized by staining. For HLA-D typing, the 2-color fluorescence or mixed lymphocyte culture technique is used.

The extremely complex HLA system was researched worldwide in a large number of laboratories. Knowledge on the HLA system has rapidly developed since 1964 due to a unique international cooperation and the ‘International Histocompatibility Workshops’ (current last HLA workshop 2008) held every 2–4 years. The results of the workshops which were published since 1965 in short intervals also contributed to this development [87, 88]. Since 1968 a new nomenclature was developed by the WHO [88, 89], in which the loci or regions were defined by letters A, B, C etc., and the specificity within a locus by numbers 1, 2, 3 etc. The current form of nomenclature of genes or alleles and specificities can be found in Marsh et al. [90] (also under [87]).

At the International HLA Workshop in 1970, 27 specificities of HLA-A, B, and C had already been defined. In the early 1980s a total of 14 HLA-A, 28 HLA-B, and 8 HLA-C markers were recognized as well as many supertypical specificities of A so-called silent genes or null alleles (or hyposynthetic variants) which were observed for HP, GC, C3, BF, C6, PLG, TF (very rare), and PI (very rare). The HP-0 gene was more commonly observed in African populations.

Since the mid-1980s serum proteins were increasingly being analyzed at the genome level. Using suitable primers, already known genotypes could be confirmed by PCR, and rare variants, for example in GC, PI, HP and A2HS, could be analyzed (reviews in [78, 80]).

### Table 3. Selected serum protein polymorphisms

<table>
<thead>
<tr>
<th>System</th>
<th>Discoverer</th>
<th>Chromosomal location</th>
<th>PEC, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP (haptoglobin)</td>
<td>Smithies 1955</td>
<td>16q22.2</td>
<td>18</td>
</tr>
<tr>
<td>HP subtypes</td>
<td>Smithies, Connel, Dixon 1962</td>
<td>16q22.2</td>
<td>32</td>
</tr>
<tr>
<td>GM (chains IgG)</td>
<td>Grubb 1956</td>
<td>14q32.32-q32.33</td>
<td>11</td>
</tr>
<tr>
<td>KM (IoV, k chains IgG)</td>
<td>Roptart, Lenoir, Rivat 1961</td>
<td>2p11.2</td>
<td>6</td>
</tr>
<tr>
<td>TF (transferrin)</td>
<td>Smithies 1957</td>
<td>3q21</td>
<td>1</td>
</tr>
<tr>
<td>TF subtypes</td>
<td>Kühnl, Spielmann 1978</td>
<td>14q2.1</td>
<td>20</td>
</tr>
<tr>
<td>GC (vitamin D binding protein)</td>
<td>Hirschfeld 1959</td>
<td>4q12-q13</td>
<td>16</td>
</tr>
<tr>
<td>GC subtypes</td>
<td>Constans, Viau 1977</td>
<td>14q2.1</td>
<td>30</td>
</tr>
<tr>
<td>AG (LDL lipoprotein)</td>
<td>Allison, Blumberg 1961</td>
<td>2p23-p24</td>
<td>14</td>
</tr>
<tr>
<td>ORM1 (osoromucoid)</td>
<td>Schmid, Binette 1961</td>
<td>9q31-pter</td>
<td>21</td>
</tr>
<tr>
<td>PI (alpha-1-antitrypsin)</td>
<td>Fagerhol, Braend 1965</td>
<td>14q2.1</td>
<td>5</td>
</tr>
<tr>
<td>PI subtypes</td>
<td>Constans, Viau 1975</td>
<td>14q2.1</td>
<td>20</td>
</tr>
<tr>
<td>C3 (complement)</td>
<td>Alper, Propp 1968</td>
<td>19p13.3-p13.2</td>
<td>13</td>
</tr>
<tr>
<td>C3c</td>
<td>Azen, Smithies 1968</td>
<td>19p13.3-p13.2</td>
<td>13</td>
</tr>
<tr>
<td>BF (factor B properdin)</td>
<td>Alper, Boenisch, Watson 1972</td>
<td>6p21.3</td>
<td>14</td>
</tr>
<tr>
<td>BF subtypes</td>
<td>Geserick et al. 1982</td>
<td>6p21.3</td>
<td>23</td>
</tr>
<tr>
<td>C6 (complement)</td>
<td>Hobart, Lachmann, Alper 1975</td>
<td>5p12-p14</td>
<td>18</td>
</tr>
<tr>
<td>A2HS (glycoprotein)</td>
<td>Anderson, Anderson 1977</td>
<td>3q27.3</td>
<td>18</td>
</tr>
<tr>
<td>PLG (plasminogen)</td>
<td>Hobart 1979</td>
<td>6q26</td>
<td>20</td>
</tr>
<tr>
<td>F13B (coagulation factor 13)</td>
<td>Board 1980</td>
<td>1q31-q32.1</td>
<td>22</td>
</tr>
</tbody>
</table>

*Selected according to the usefulness in kinship analysis; the PEC value for IEF subtyping is in italic type. Middle European frequencies were used for calculation of the PEC.

PEC = Paternity exclusion chance.
and B alleles, which indicate several determinants on a defined HLA marker. The number of possible phenotypes at this time was calculated to be more than $23 \times 10^6$. A few years later 23 HLA-A, 46 HLA-B, and 9 HLA-C markers could be differentiated [85]. Currently 28 HLA-A, 62 HLA-B, and 10 HLA-C specificities are serologically defined [90]. Due to the high degree of polymorphism the HLA-A, B, and C markers were rapidly integrated into kinship analysis [83].

The alloantigens belonging to the later deciphered HLA-D region were also described as Class II antigens and are only detectable on particular cell types such as B lymphocytes and monocytes. At the 1975 workshop the HLA specificity Dw was described and by 1980 markers of the subregions DR and DQ had been defined. Serologically, 2 subregions with 17 HLA-DR and 7 HLA-DQ alleles could initially be differentiated [85]. The serological subtyping of the HLA-DR markers were also used in kinship analysis in addition to the HLA-A, B, and C alleles [91]. The combined use of HLA-A, B, C, and DR markers meant that a number of possible phenotypes greater than $20 \times 10^9$ could be differentiated [85]. The previously mentioned HLA-Dw factors could be detected using mixed lymphocyte culture with T-lymphocytes. Nowadays 26 HLA-D (Dw), 24 HLA-DR, 9 HLA-DQ, and 6 HLA-DP specificities can be serologically defined [90].

The frequency distributions of the HLA-A, B, C, and D markers show large differences between the 3 main human ethnic groups [25, 92].

The HLA system is located on the short arm of chromosome 6 in the form of closely linked loci within the MHC together with the complement proteins C2, C4A, C4B, and Bf (table 4). The HLA loci control gene products which differ in their biochemical structure and function. In this way HLA-A, B, and C loci encode for glycosilated polypeptides (A chain, molecular weight (MW) 44 kDa) on which the $\beta_2$-microglobulin binds. The HLA-D loci encode 2 different glycosilated polypeptides (A and B chains, MW 33 kDa and 29 kDa, respectively). The HLA gene products are topographically arranged in an intracellular, transmembranous, and intercellular part. Due to the multitude of alleles, especially at the gene loci A, B and DR, there is a gigantic number of possible phenotypes. The system is complicated by many cross-reactions which are particularly common among the A, B, and DR alleles. In addition there is a strong linkage disequilibrium within the HLA system [8, 84, 86] which must be taken into consideration for statistical calculations.

The HLA system is technically demanding. The correct determination of HLA antigens requires optimal serological test conditions, high-grade test sera, and extensive practical experience [8, 84]. Because of this the successful implementation of the HLA system in kinship testing in the 1970s and 1980s was extensively reserved for specialized laboratories. The HLA system was as a rule used as a supplement to the blood, serum, and enzyme groups. By combining these conventional systems with the HLA-A, B, C, and DR markers, a PEC of greater than 99.9% could be achieved. The HLA system proved to be particularly valuable for problem cases if no exclusion constellations could be found in one- or multiple-men cases or where there were indications of silent alleles. A further successful application area was in deficiency cases where close relatives had to be included instead of the putative father. The exclusion constellations in the HLA system are similar to blood groups with multiple alleles and inheritance by haplotypes: i) the child possesses a marker which neither the mother nor the putative father have; ii) the child does not have either of the markers (alleles) at one locus which the putative father has; iii) the child has a haploypotype which the putative father does not have, whereby the haplotypes of the persons involved were determined by family investigations [84, 93].

Since 1986 [94] previously known alleles of HLA-A, HLA-B, and HLA-C could be further subdivided by molecular genetic techniques (restriction fragment length polymorphisms (RFLP), nucleotide sequencing, and PCR), in particular recognition of the polymorphisms of HLA-DRA, DRB1, DQA1, DQB1, DPA1, and DPB1. Currently there are over 5,400 alleles of Class I markers in databases and over 1,500 alleles of Class II markers. New allele sequences are continually being found by sequence-based typing. In the results of the last 2 ‘International HLA and Immunogenetics Workshops’, 3,249 officially named Class I alleles and 1,198 Class II alleles were reported [90]. The HLA-A alleles (n = 616), HLA-B alleles (n = 913), and HLA-C alleles (n = 446) show a particularly high degree of polymorphism, as do the HLA-DRB alleles (DRB1 n = 368) in contrast to only 45 DQb1 and 22 DPB1 alleles [27, 87, 88, 90]. The assignment of serologically defined specificities to the sequence-based alleles is not always clear (reviews in [90, 95]).

### Table 4. HLA system

<table>
<thead>
<tr>
<th>Locus I</th>
<th>Occurrence</th>
<th>Gene locus</th>
<th>PEC, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I antigens</td>
<td>all nucleated cells</td>
<td>6p21.3</td>
<td>96</td>
</tr>
<tr>
<td>Class II antigens</td>
<td>B-lymphocytes, monocytes, activated T-lymphocytes, epithelial cells</td>
<td>6p21.3</td>
<td>98</td>
</tr>
</tbody>
</table>

PEC = Paternity exclusion chance.
patterns in a similar way to serum proteins. In addition, for phosphoglucomutase 2, new independent gene loci PGM2 and PGM3 (on chromosomes 4 and 6, respectively) could be discovered. Since 1968 rare null genes have continually been detected, initially for PGM1 and in subsequent years also for ACP, AK, ADA, GPT (frequency in European populations approximately 1–2%), ESD, and GLOI (frequency in European populations approximately 1.5%). The existence of null genes must be taken into consideration in kinship analysis because false exclusions can be produced by apparent contrasting homozygosity. In addition, for GLOI the linkage with the MHC on chromosome 6 must also be taken into account.

The phenotypes of PGM1, AK, and ADA are also detectable in other body cells apart from erythrocytes.

Erythrocyte Enzyme Polymorphisms

Since the discovery of the erythrocyte acid phosphatase (ACP) polymorphism in 1963 by Hopkinson, Spencer, and Harris [96] the field of human enzyme genetic polymorphisms developed rapidly. Kinship analysis received a substantial boost of further systems with confirmed codominant inheritance and useful population frequencies. This resulted in a further increase in PEC and biostatistical probabilities.

The technical foundation for this was the separation of stroma-free erythrocyte hemolysates by electrophoresis (initially in starch gels but later in polyacrylamide and agarose gels or cellulose acetate strips). Following electrophoresis the enzyme patterns were visualized with the zymogram technique, i.e. a functional development of enzyme bands with relatively enzyme-specific reactions on the separation medium. In later years alternative visualization methods were developed, such as electroblotting with specific antibodies. The electrophoresis techniques were also improved. Using thin-layer electrophoresis the amount of material used could be minimized, and with IEF the separation could be improved in some systems. This was especially true for PGM1 where the 2 most common gene products could be separated into 4 subtypes which led to a doubling of the PEC to 30%.

As for the serum groups, in the 2 decades following the discovery of the ACP system comprehensive studies were carried out on the frequency distribution in various populations and on the inheritance of enzyme systems. The combined (cumulative) PEC for the systems listed in table 5 was 75% (reviews in [8, 77, 97–99]).

The molecular and genetic basis of enzyme polymorphisms has been clarified since the 1980s, and in the following decade typing of enzyme polymorphisms was also possible at the DNA level using PCR, starting with ACP (reviews in [28, 100, 101]).

For most enzyme polymorphisms, from 1965 many rare variants were detected using electrophoresis: PGM1, AK, ADA, GPT, ESD, GLO I (only very rarely). Identification could only be made by a direct electrophoretic comparison of patterns in a similar way to serum proteins. In addition, for phosphoglucomutase 2, new independent gene loci PGM2 and PGM3 (on chromosomes 4 and 6, respectively) could be discovered. Since 1968 rare null genes have continually been detected, initially for PGM1 and in subsequent years also for ACP, AK, ADA, GPT (frequency in European populations approximately 1–2%), ESD, and GLOI (frequency in European populations approximately 1.5%). The existence of null genes must be taken into consideration in kinship analysis because false exclusions can be produced by apparent contrasting homozygosity. In addition, for GLOI the linkage with the MHC on chromosome 6 must also be taken into account.

The phenotypes of PGM1, AK, and ADA are also detectable in other body cells apart from erythrocytes.

Transition to the Use of DNA Systems

In the 1980s the structure of the human genome was increasingly being researched. It came to be recognized that repeat sequences occurred in the areas outside the genes of nuclear DNA, the so-called repetitive DNA. Under the tandem repetitive DNA regions the minisatellites (repeat length up to 100 bp) and microsatellites (repeats up to 6 bp) were of forensic interest.

The beginning of forensic DNA analysis is generally considered to be attributable to the procedure of DNA fingerprinting published by Jeffreys et al. in 1985 [102]. Using bacterial restriction enzymes the DNA double chains of the myoglobin gene were cut in a sequence-specific manner. The resulting RFLP revealed a multilocus polymorphism of great individual variability. Using such analyses up to 50 bands could be produced. A further form of multilocus analysis was carried out with oligonucleotide probes, the so-called simple repetitive sequences [103].

Using restriction enzymes, single locus polymorphisms could also be detected, systems which only marked 1 gene locus and could be formal genetically checked [104].
Whereas the RFLP method placed high demands on the amount and quality of extractable DNA, the amplification of extremely small amounts of DNA even in a degraded condition resulted in totally new possibilities. PCR [94] resulted in the enormous success of DNA systems due to an automated in vitro reaction. In the early phase of forensic PCR typing, DNA sequence polymorphisms were investigated using so-called oligotyping [105]. In the selected HLA-D region sequence polymorphisms of DQ and DR could be forensically applied (see section on HLA system).

A particular expansion of PCR systems in kinship analysis and criminalistics occurred due to the short tandem repeats (STRs) [106]. Due to the short fragment lengths these are particularly insensitive to degradation and can be typed from small amounts of material (picogram range). A large number of STR systems with numbers of alleles ranging from 5 to 30 were rapidly described. The first internationally established STR systems VWA and TH01 were followed in the 1990s by many further systems. Currently all routinely encountered cases in kinship analysis can be resolved using at least 15–25 validated STR systems, which can be determined using various commercially available kits, and a combined PEC of at least 99.99% can be reached. Further contributions to the clarification of kinship can be made by the gonosomal markers (X and Y STRs) [106, 107] and the maternally inherited mitochondrial DNA (mtDNA) [108].

New DNA polymorphisms, such as SNPs or insertion-deletion polymorphisms (indels) are also being applied. The DNA markers beginning in the 1990s have now in practical terms completely replaced conventional blood group markers, particularly in the form of STR systems, due to their higher efficiency (cost-benefit ratio).

**Biostatistics in Kinship Analysis**

Kinship analysis with blood group markers began in the 1920s with proof of non-paternity by exclusion of the putative father. With the inclusion of an ever increasing number of hemogenetic markers it became possible to succeed in finding positive evidence of paternity or of kinship relationships. The rapid succession of discovery of new hereditary markers in the blood following World War II allowed the increasing inclusion of biostatistics in kinship analysis.

When markers are concordant in the child and putative father, particularly when they are lacking in the mother and rarely occur in the population, this is a strong indication of paternity. Essen-Möller published a formula for this principle in 1938 [109] and introduced the concept of ‘probability of paternity’ (W). His formula was initially not generally accepted and first became a recognized principle for the biostatistical kinship analysis 2 decades later. In this formula it is presupposed that ‘true’ and ‘false’ fathers are available for investigation. A hereditary marker observed in both the child and the putative father must occur with a higher probability (X) in true paternity than when the child and the man are genetically independent (probability Y). The likelihood ratio $L = \frac{X}{Y}$ describes the relationship of these probabilities. The logarithm of the likelihood ratio is used as the EM value (named in honor of Essen-Möller). The chance of finding the ‘true’ father among the probands when ‘true’ and ‘false’ fathers present for examination in the same proportions (neutral a priori value of 0.5) is $W = \frac{X}{X+Y}$. From this the formula of Essen-Möller for probability of paternity $W = \frac{1}{1+Y/X}$ can be derived. The likelihood ratio for combined marker systems is formed from the product of the systems investigated. For the forensic application of biostatistical values, verbal predicates were beneficial for the explanation of W values. Verbal predicates had already been suggested by Essen-Möller, which were oriented on the sigma limits, but later a differentiated scale was established which was published in 1971 [110]. The Paternity Index $PI = \frac{X}{Y}$, which is recommended since 1952, is the inverse value of $Y/X$ and contains the same information as the W value.

$$P_{p} = \text{probability (types observed/ the hypothesis is that the tested man is the father)}$$

$$P_{f} = \text{probability (types observed/ the hypothesis is that a random man is the father)}$$

The PI was typically used in English speaking and Scandinavian countries. The relationship between W and the general paternity exclusion chance A ($AVACH = PEC$) can be represented by $W = \frac{1}{2} - A$ and the relationship to $PI$ as $PI = 1 - A$, i.e. the values $W$, $PI$, and $A$ are related to each other. The advantage of A is that it is independent of the case in question and can be considered as a measure for the efficiency of a hereditary system just as the W and PI values. Originally experts had to calculate the biostatistical values by hand but later tables and computer programs became available [110, 111], and HLA and DNA systems were also included in the calculation [112, 113] (further reviews in [114, 115]). Finally recommendations from the International Society for Forensic Genetics were formulated which corresponded to the recommendations of the International Standardization Organization ISO 17025 [114, 116].

Currently a kinship report should attain the minimum biostatistical value of $W = 99.9\%$ or an analogue statistical measure [117], but with modern methods (DNA systems) values of $A = 99.99\%$, $W = 99.99\%$, and $PI = 10,000$ are easily reached (according to recommendations by Baur et al. [118]).

**Guidelines for Kinship Analysis**

It is interesting to consider the official guidelines for expert opinions in Germany because they provide a good insight into the development of the science and techniques in the field of kinship testing (review in [119]).
Even in 1937 there were initial ministerial ‘Guidelines for carrying out blood group investigations and implementation of an official certification of the test sera used’ which included minimum requirements for the theoretical and practical training of investigators. In 1940 the Prussian Institute for Infectious Diseases formulated a ‘Working manual for carrying out forensic blood group investigations’.

After World War II amended guidelines were first published in 1960 in the ‘Bundesgesundheitsblatt’ (Federal Health Gazette) and subsequently updated at regular intervals, the last update having been in 2002 [117]. In the guidelines of the Federal Health Office from 1990 [120] a minimum standard consisting of the following systems was defined: i) ABO, MNSs, Rhesus (RH), Kell (K), Duffy (FY); ii) HP, GC, GM, INV; iii) ACP, PGMI, ADA, AK, PGD, ESD, GLO; iv) and for problem cases the serologically defined markers of the HLA-A, B, C, and DR locus. For isolated exclusions, extended investigations were recommended, in particular for exclusions due to opposite homozygosity.

In 1992 the guidelines were extended to DNA markers corresponding to the scientific progress [121], including single locus systems as well as multilocus systems. The DNA markers were initially recommended for use in extended investigations or for exceptional cases (e.g. material unsuitable for blood group analyses).

In the amendments to the guidelines from 1996 [122] the reservations regarding HLA and DNA systems were maintained. For expert opinions a minimum of 4 system categories and at least 10 loci with independent inheritance and a combined PEC of 99.99% were required. The markers in the 4 system categories were recommended as follows: i) Erythrocyte membrane antigens: ABO, MNSs, RH, Kidd (JK), Duffy (FY); ii) Serum protein systems: GC, PI, F13B, HP, A2HS, ORM, TF, PLG, BF, C3; iii) Erythrocyte enzyme systems: PGMI, ACP, GPT, GLO, ESD; iv) HLA system: serologically detectable antigens; v) DNA single locus polymorphisms.

In the current guidelines from 2002 [117] the following systems were recommended either alone or in combination: i) RFLP; ii) Microsatellite polymorphisms (STR); iii) HLA system; iv) Combination of erythrocyte membrane antigens, serum proteins, erythrocyte enzymes. A combined PEC of at least 99.99% was also required here, and at least 12 independent loci on a minimum of 10 different chromosomes should be investigated. For exclusions, at least 3 or more exclusion constellations on different chromosomes were necessary. For non-exclusions, a W value of at least 99.9% should be attained (or analogous statistical measures such as PI or A) in order to conform to the predicate ‘paternity practically proven’.

International recommendations were published in 2002 by the Paternity Testing Commission of the International Society for Forensic Genetics [114, 116].

In accordance with the German guidelines, which in 1992 first required an external control of laboratories, the Working Group of Experts for Kinship Analysis in the Federal Republic of Germany (Arbeitsgemeinschaft der Sachverständigen für Abstammungsgutachten in der Bundesrepublik Deutschland, AGSV), now called the German Society for Kinship Testing (Deutsche Gesellschaft für Abstammungsbegutachtung, DGAB), organized 2 proficiency testing trials per year. An examination of the necessary knowledge and practical qualifications of experts was introduced by the guidelines from 2002 and carried out by a commission under the leadership of the AGSV (now DGAB).

In the USA a similar program has been carried out by the American Association of Blood Banks (AABB) and the College of American Pathologists (CAP) since 1993 which also comprises an accreditation of successful participating laboratories. The ‘Standards for Parentage Testing Laboratories’ have been published by the AABB since 1985.

In Germany the Gene Diagnostics Act (GenDG) came into force on 1 February 2010 [123]. According to §5 of the GenDG, testing institutions must be in possession of an accreditation as from 1 February 2011. An interdisciplinary gene diagnostics commission (GeKo) will in the future formulate guidelines for the qualification of experts as well as for quality assurance measures (§23 GenDG).

International Society for Forensic Genetics

The International Society for Forensic Genetics is an independent society of scientists for the promotion of knowledge on human genetic markers (www.isfg.org) which was founded on June 24, 1968 at the University of Mainz (Germany) as the ‘Gesellschaft für forensische Blutgruppenkunde’ (Society for Forensic Blood Grouping). The expansion with increasing internationalization led in 1980 to transformation into the International Society for Forensic Hemogenetics which became the International Society for Forensic Genetics in 1991. The society organizes an international congress every 2 years as well as language-based working groups and has published numerous recommendations (e.g. [114, 116]). The society represents more than 1,000 members from over 50 countries. In 2007 the journal Forensic Science International: Genetics, which emerged from Forensic Science International, was founded as the official organ of the International Society for Forensic Genetics.

Disclosure Statement

The authors declared no conflict of interest.
References

Genetic Kinship Investigation from Blood Groups to DNA Markers

Transfus Med Hemother 2012;39:163–175