Formal Genetics of Humans: Modes of Inheritance*

5

Arno G. Motulsky

Revised by S.E. Antonarakis and M.R. Speicher

Contents

> The law of combination of the differing traits, according to which the hybrids develop, finds its foundation and explanation in the proven statement that the hybrids produce germ and pollen cells … which originate from the combination of the traits by fertilization.
> G. Mendel, Versuche über Pflanzenhybriden, 1865

A.G. Motulsky (✉)
Departments of Medicine (Division of Medical Genetics) and Genome Sciences
University of Washington, School of Medicine, Box 355065, 1705 N.E. Pacific St. Seattle, WA 98195-5065, USA
e-mail: agmot@u.washington.edu

*This chapter is a revised and updated version of the third edition's Modes of Inheritance chapter originally authored by Friedrich Vogel and Arno Motulsky.

M.R. Speicher et al. (eds.), *Vogel and Motulsky's Human Genetics: Problems and Approaches,*
DOI 10.1007/978-3-540-37654-5_5, © Springer-Verlag Berlin Heidelberg 2010

5

5.1 Mendel's Modes of Inheritance and Their Application to Humans

Mendel's fundamental discoveries are usually summarized in three laws:

1. Crosses between organisms homozygous for two different alleles at one gene locus lead to genetically identical offspring (F_1 generation), heterozygous for this allele. It is unimportant which of the two homozygotes is male and which is female (law of uniformity and reciprocity). Such reciprocity applies only for genes not located on sex chromosomes.
2. When these F_1 heterozygotes are crossed with each other (intercross), various genotypes segregate: one-half are heterozygous again, and one-quarter are homozygous for each of the parental types. This segregation 1:2:1 is repeated after crossing of heterozygotes in the following generations, whereas the two types of homozygotes breed pure. As noted previously (Chap. 1), Mendel interpreted this result correctly, assuming formation of two types of germ cells with a 1:1 ratio in heterozygotes (law of segregation and law of purity of gametes).
3. When organisms differing in more than one gene pair are crossed, every single gene pair segregates independently, and the resulting segregation ratios follow the statistical law of independent segregation (law of free combination of genes).

This third law applies only when there is no linkage (Chap. 6). Human diploid cells have 46 chromosomes: the two sex chromosomes and 44 autosomes forming 22 pairs of two homologues each. The pairs of homologues are separated during meiosis, forming haploid germ cells or gametes. After impregnation, paternal and maternal germ cells unite to form the zygote, which is diploid again. Sex is determined genotypically; women normally have two X chromosomes, men have one X and one Y chromosome (Chap. 3).

For an understanding of the statistical character of segregation ratios in humans it is important to realize that the number of germ cells formed is very large, particularly among males. Only a very small sample comes to fertilization. Regarding single gene loci this sampling process can generally be regarded as random.

Two alleles may be termed A and A′. The set of combinations described in Fig. 5.1 are possible. As noted above, these theoretical segregation ratios are probabilities; segregation ratios found empirically should be tested by statistical methods to determine whether they are compatible with the theoretical ratios implied by the genetic hypothesis.

The mating type of identical homozygotes (AA × AA or A′A′ × A′A′) is uninteresting except where it permits conclusions regarding genetic heterogeneity of a recessive condition (Sect. 5.3.5). Mating between the two different homozygous types (AA × A′A′) is usually rare and is therefore of little practical importance. Matings between homozygotes and heterozygotes (AA′ × AA) and between two heterozygotes (A′A × A′A) are most important practically, as explained below.

Mendel found that a genotype does not always determine one distinct phenotype. Frequently heterozygotes resemble (more or less) one of the homozygotes. Mendel called the allele that determines the phenotype of the heterozygote dominant, the other recessive. With more penetrating analysis, some human geneticists have concluded that these terms may be misleading and should be abandoned. In fact, at the level of gene action, genes are not dominant or recessive. At the phenotypic level, however, the distinction is important and useful. Biochemical mechanisms of dominant hereditary diseases usually differ from those of recessive conditions. Hence the mode of inheritance gives a hint regarding the biochemical mechanism likely to be involved.

There are a number of instances in which each of two alleles in a heterozygous state has a distinct phenotypic expression. If both are inherited and phenotypically expressed, this mode of inheritance is sometimes called codominant.

5.1.1 *Codominant Mode of Inheritance*

The first examples of codominance in man were found in the genetics of blood groups; the MN blood types (111300; numbers refer to identifying numbers of diseases listed in [52]) may serve as an example (Table 5.1). When methods for genetic analysis at the protein level became available, many more examples were soon discovered. The example in Table 5.1 clearly points to a genetic model with two alleles, M and N, the phenotypes M and N being the two homozygotes and MN the heterozygote. This example is used below for a statistical comparison between expected and observed

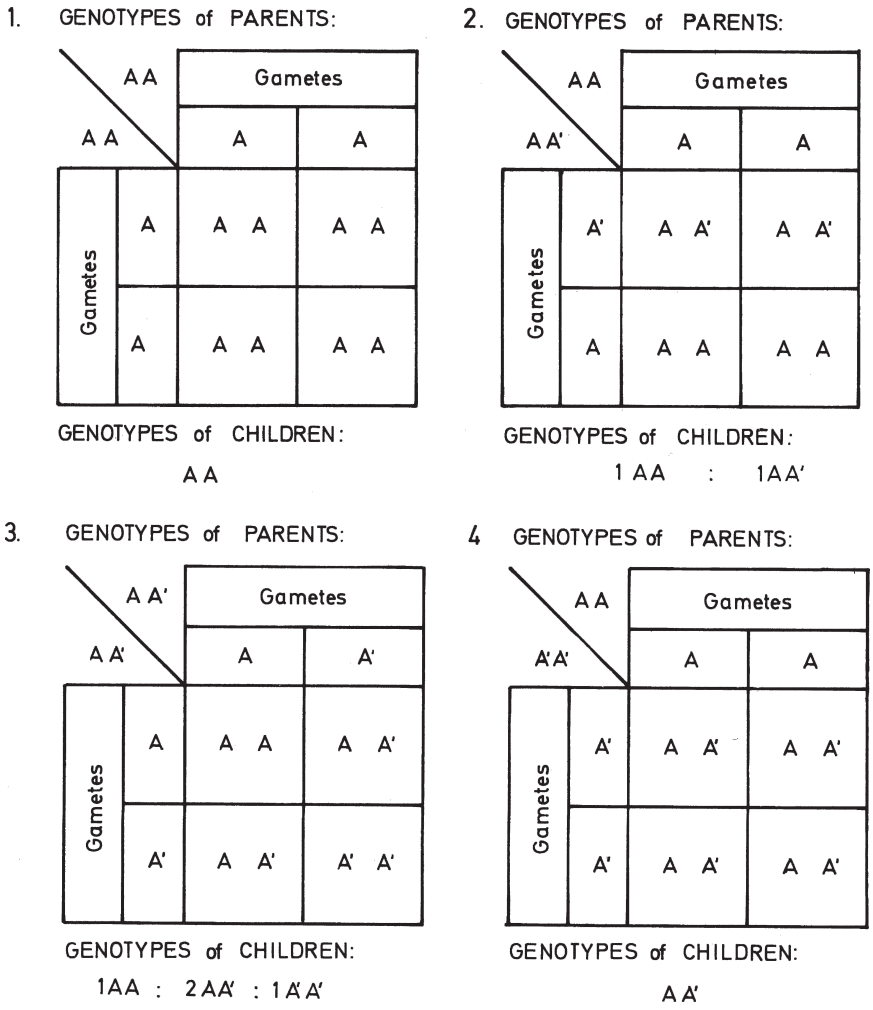

Fig. 5.1 Mating types with two alleles

Table 5.1 Family studies of the genetics of MN blood types from Wiener et al. [92]

Mating type	Number of families	Types of children			Total children
		M	N	MN	
M × M	153	326	0	(1)	327
M × N	179	(1)	0	376	377
N × N	57	0	106	0	106
MN × M	463	499	(1)	473	973
MN × N	351	(3)	382	411	796
MN × MN	377	199	196	405	800
	1,580	1,028	685	1,666	3,379

Parentheses, false paternity.

segregation ratios. The "aberrant" cases in parentheses, which at first glance seem to contradict the genetic hypothesis, were the result of false paternity – a frequent finding in most such investigations.

5.1.2 Autosomal Dominant Mode of Inheritance

The first description of a pedigree showing autosomal dominant inheritance of a human anomaly was Farabee's [22] paper in 1905 on "Inheritance of Digital

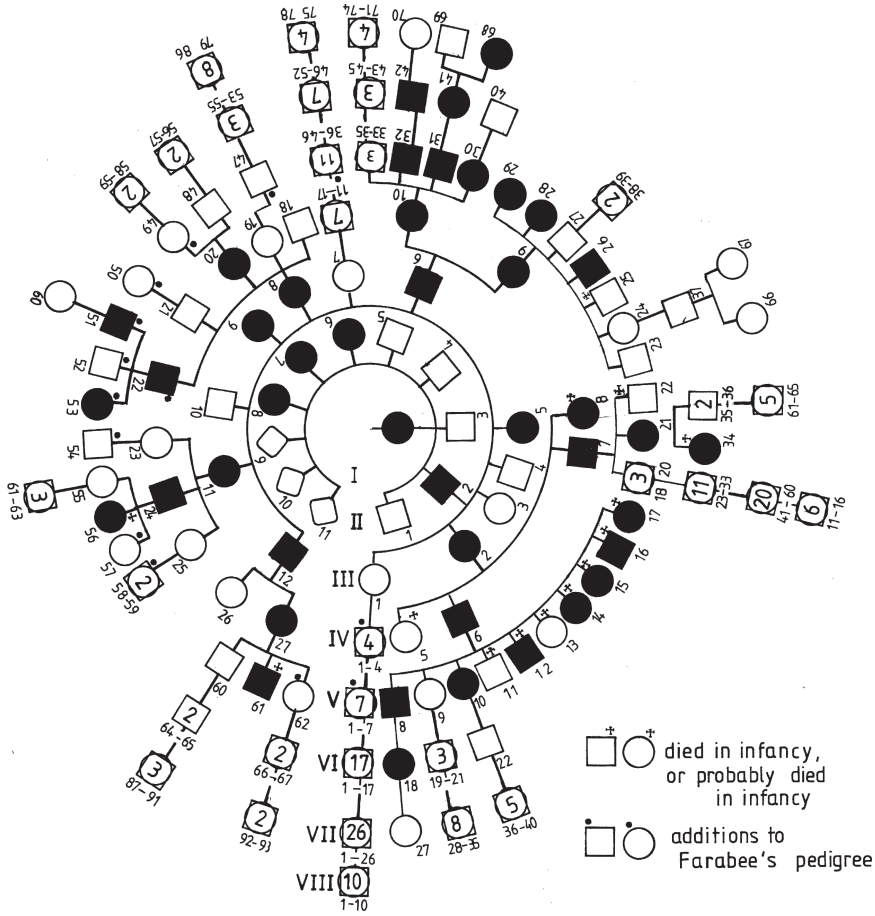

Fig. 5.2 The brachyphalangy pedigree of Farabee [22]. *Black symbols*, affected females (●) and males (■); *numbers,* point to their position in the pedigree

Malformations in Man" (Fig. 5.2). Textbooks usually refer to the condition as brachydactyly (short digits), but from the original paper it is clear that not only were the phalanges of hands and feet shortened, but the number of phalanges was also reduced (Fig. 5.3). In addition, stature was low (average of 159 cm in three males), apparently due to shortness of legs and inferentially also of arms. In every other aspect, Farabee wrote,

> The people appear perfectly normal …and seem to suffer very little inconvenience on account of their malformation. The ladies complain of but one disadvantage in short fingers, and that is in playing the piano; they cannot reach the full octave and hence are not good players.
>
> Figure 5.3 shows the pedigree. There are 36 affected in generations II–V, 13 of which are male and 23 female. Among the unaffected 18 are male and 15 female. The trait is transmitted from one of the parents to about half the children; transmission is independent of sex. Unfortunately, Farabee did not consider the children of the unaffected. Had he done so, he would have found them free from the anomaly. Many other pedigrees have shown absence of the trait among offspring of parents who do not carry the dominant gene. More recently the family has been reexamined [38]. The children of the unaffected family members and some affected family members were added, and X-ray examination confirmed that not only hands and feet were affected but the distal limb bones as well. The basic defect is thought to affect the epiphyseal cartilage.

The condition described by Farabee is now referred to as brachydactyly A-1 (BDA1; OMIM 112500). As pointed out by Farabee, characteristics include shortness of all middle phalanges of the hands and toes,

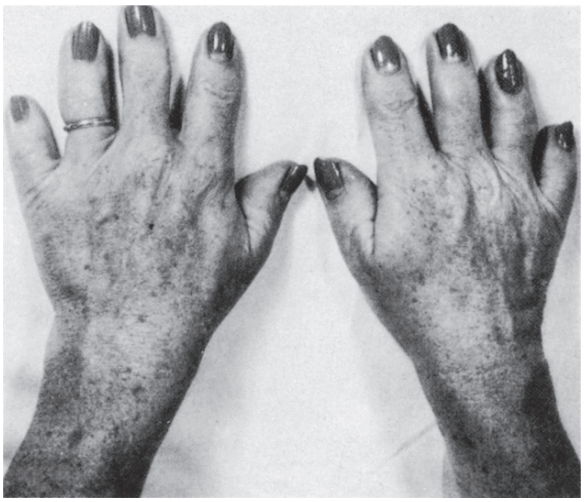

Fig. 5.3 Brachyphalangy in one member of a younger generation of Farabee's pedigree. From Haws and McKusick [38]

occasional terminal symphalangism, shortness of the proximal phalanges of the first digit, and short stature. In 2002 mutations in the Indian Hedgehog gene (*IHH*) were found in descendants of Farabee's family resolving an almost 100-year-old mystery [50, 51]. BDA1 is a heterogeneous condition as an additional locus in another BDA1 family was mapped to 5p13.3-p13.2 [1] (OMIM 607004), and in another BDA1-affected family both the *IHH* locus and the 5p13.3-p13.2 region were excluded [44] suggesting that other, yet unidentified mutations may cause the BDA1 phenotype.

Affected patients are heterozygous for an autosomal allele leading to a clearcut and regular abnormality in the heterozygote. Therefore the trait is, by definition, dominant. The family shows two other characteristics that have since been found to be widespread:

1. The anomalies were described as being almost identical in all family members, and in each person appearing in all four extremities. This is a frequent finding in malformations with a regular mode of inheritance. The reason for the symmetry is evident considering that the same genes act on all four extremities.
2. The anomaly affected the well-being of its bearers only very little. This lack of health impairment is typical for such extended pedigrees. Reproduction is normal. Otherwise the trait would not be transmitted and would soon disappear. This is why, especially in the more serious dominant conditions, extended pedigrees are the exception rather than the rule. Most diseases caused by mutations observed in the present generation have originated rather recently, often even in the germ cell of one of the parents.

5.1.2.1 Late Manifestation, Incomplete Penetrance, and Variable Expressivity

Sometimes a severe dominant condition manifests only during or after the age of reproduction. Here extended pedigrees are usually observed in spite of the severity of the condition. The classic example is Huntington disease (HD) (143100), a degenerative disease of the nerve cells in the basal ganglia (caudate nucleus and putamen) leading to involuntary extrapyramidal movements, personality changes, and a slow deterioration of mental abilities.

Wendt and Drohm [88] carried out a comprehensive study of all cases of HD in the former West Germany. The distribution of ages at onset is presented in Fig. 5.4. The great majority of their patients were married when they developed clinical symptoms. Even among thousands of patients the authors were not able to locate a single case that could be ascribed with confidence to a new mutation. For these reasons and based on results from other early studies, the existence of de novo mutations in HD had long been debated. HD is caused by an increased (CAG) trinucleotide repeat number within the huntingtin gene (*HD*) on 4p16. The unaffected range is $(CAG)_{6–35}$ repeats. Alleles with a length of $(CAG)_{40}$ and above are fully penetrant, i.e., they will cause HD within a normal lifespan. In contrast, alleles of $(CAG)_{36–39}$ confer an increasing risk of developing HD [3]. Analysis of apparently sporadic HD cases revealed that nonpathogenic alleles in the high normal range ($(CAG)_{27–35}$) have the potential to expand into the pathogenic range [57]. In fact $(CAG)_{27–35}$ alleles can be unstable during transmission and have a relatively high mutation rate for *HD* of $\geq 10\%$ in each generation [23]. Analysis of the gene is described in Chapter 9.

Another phenomenon occasionally encountered in dominant traits is incomplete penetrance [72]). Penetrance is a statistical concept and refers to the fraction of cases carrying a given gene that manifests a

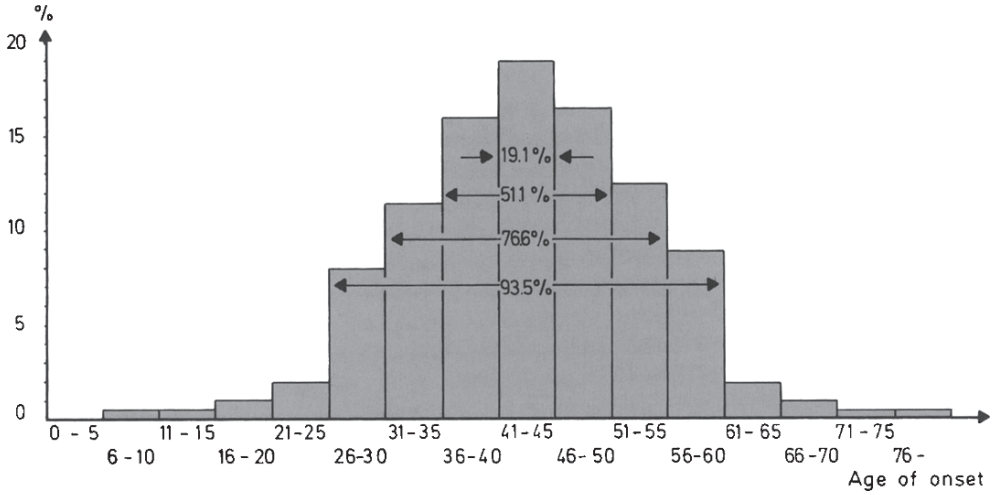

Fig. 5.4 Distribution of ages at onset in 802 cases of Huntington's disease. From Wendt and Drohm [88]

specified phenotype. The transmission seems occasionally to skip one generation, leaving out a person who judging from the pedigree must be heterozygous, or the fraction of those affected among sibs (after appropriate corrections, Sect. 5.3.4) turns out to be lower than the expected segregation ratio. An example is retinoblastoma (180200), a malignant eye tumor of children. Bilateral cases (and cases with more than one primary tumor) are always dominantly inherited, whereas most unilateral, single tumors are nonhereditary, probably being caused by somatic mutation (Chap. 10). Even in pedigrees otherwise showing regular dominant inheritance, however, apparent skipping of a generation is observed occasionally (Fig. 5.5). Calculation of the segregation ratio in a large sample showed that about 45% of sibs were affected instead of the 50% expected in regular dominant inheritance. The penetrance of all cases (unilateral and bilateral) is therefore about 90%. Penetrance in families with bilateral cases is higher than in those with unilateral cases.

In many cases, penetrance is a function of the methods used for examination; higher penetrance is observed with detection methods (clinical or laboratory) that are closer to gene action.

In many dominant conditions the gene may manifest in all heterozygotes, but the *degree of manifestation* may be different. An example is neurofibromatosis (162200). Some cases may show the full-blown picture with many tumors of the skin, café-au-lait spots, and

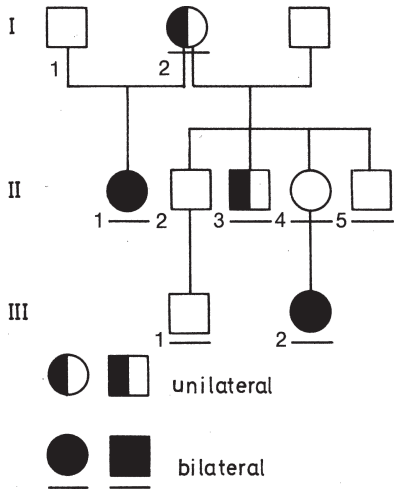

Fig. 5.5 Incomplete penetrance in retinoblastoma. The unaffected woman II,4 must be heterozygous, her mother I,2 and her daughter III,2 being affected; ❑ personally observed. (Personal observation, F. V.)

systemic involvement, whereas other cases – even in the same families – may show only a few café-au-lait spots. The term used to describe this phenomenon is "variable expressivity" [72]. While such terms as "incomplete penetrance" and "variable expressivity" are often needed to convey quick understanding about certain phenomena, they may become dangerous if we forget that they do not explain a biological mechanism but rather are labels for our ignorance.

It is indeed somewhat surprising that so many dominant conditions show such a large interindividual variability in age at onset and severity of manifestation. It would be more understandable if such variability were observed only between different families. Our knowledge of molecular biology (Chap. 10) suggests that the mutational events leading to these conditions are almost always slightly different between families. Indeed, there is usually an intrafamilial correlation between age at onset and severity of manifestation. For HD, for example, Wendt and Drohm [88] calculated a correlation coefficient of +0.57 for age at onset for affected family members. But there usually remains appreciable variability within families, in which the abnormal genes are identical by descent. It is again no more than a label for our ignorance when we invoke the "genetic background" or the action of all other genes for help. In HD, molecular analysis of the gene has provided at least a partial explanation: the number of repeats in the DNA triplet CAG is higher in patients with onset at a very young age. Alleles of $(CAG)_{70}$ repeats or more invariably cause a juvenile onset [3]. Unfortunately, there is no correlation between the number of repeats and age at onset in most patients who develop their clinical disease in the fourth to sixth decades of life.

5.1.2.2 Effect of Homozygosity on Manifestation of Abnormal Dominant Genes

An abnormal gene is called dominant when the heterozygote clearly deviates from the normal. Indeed, almost all bearers of dominant conditions in the human population are heterozygotes. From time to time, however, two bearers of the same anomaly do marry and have children. One quarter of these are then homozygous. This has been observed in several instances, especially when the spouses were relatives. The first example was probably that described by Mohr and Wriedt in [54]. In a consanguineous marriage between two bearers of a moderate brachydactyly (112600) a child was born who not only lacked fingers and toes but also showed multiple malformations of the skeleton and died at the age of 1 year. A sister, however, had only the moderate anomaly, as did her parents [54].

Further examples of homozygosity of dominant anomalies are known. In one family, two parents with hereditary hemorrhagic teleangiectasia had a child showing multiple, severe internal and external telangiectasias who died at the age of 2.5 months [70]. Similarly, a very severe form of epidermolysis bullosa was observed in two of eight children of a couple, both of whom were afflicted with a mild type of this disease.

Another couple, both having a myopathy affecting the distal limb muscles, had 16 children, three of whom showed atypical and especially severe symptoms: the long flexors and the proximal hip muscles were also afflicted, and onset was earlier in life [87].

Epithelioma adenoides cysticum (132700) is a dominant skin disease characterized by multiple nodular tumors. One female patient, whose parents were both affected, had especially severe symptoms, and her eight children all showed this anomaly (Fig. 5.6) [28]. Further examples include achondroplasia (100800), Ehlers-Danlos syndrome (130000), and others. All these cases indicate that homozygotes of dominant anomalies are more severely affected than heterozygotes. It is therefore of interest that there appears to be no clinical difference between heterozygotes and homozygotes for HD, which is therefore a truly dominant disease as defined by Mendel. Clearly a different mechanism must apply to the pathogenesis of such a condition as compared with most other autosomal-dominant diseases, where dose effects are observed [89].

Given what we know about gene action, this is not surprising. In familial hypercholesterolemia (143890) for example, the mechanism of action of a dominant gene is known. A decreased number of receptors for a regulatory substance (low-density lipoprotein) showed the expected differences between heterozygotes and affected homozygotes: 50% decrease and complete absence or very much reduced activity of receptors,

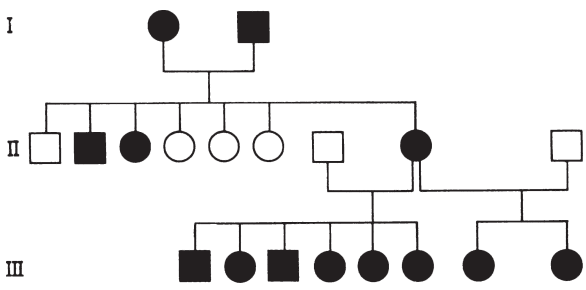

Fig. 5.6 Woman homozygous for epithelioma adenoides cysticum and her progeny in two marriages. From Gaul [28]. The pedigree was complemented in 1958 by Ollendorff-Curth [59]

respectively. Affected homozygotes show massive hypercholesteremia and usually die of myocardial infarction before the age of 30 years.

As noted above, Mendel called a gene dominant when the phenotype of the heterozygote resembled that of one homozygote. The examples of more severe manifestation of dominant genes in the homozygous than in the heterozygous state show that this strict definition is not maintained in human genetics. Here, all conditions are called dominant in which the heterozygote deviates consistently and perceptibly from the normal homozygote – irrespective of the phenotype of the anomalous homozygote. In Mendel's strict definition, most or even all dominant conditions in humans would be "intermediate." However, the more lenient connotation of "dominance" is now in general use.

5.1.3 Autosomal-Recessive Mode of Inheritance

The mode of inheritance is called recessive when the heterozygote does not differ phenotypically from the normal homozygote. In many cases special methods uncover slight detectable differences. Contrary to dominant inheritance, in which almost all crosses are between heterozygotes and homozygous normals (Sect. 5.1.2), the great majority of matings observed in recessive anomalies involve heterozygous and phenotypically normal individuals. Since the three genotypes AA, Aa, and aa occur in the ratio 1:2:1 among the offspring, the probability of a child's being affected is 25%. At the turn of the century when Garrod wrote his paper on alkaptonuria (Chap. 1) the "familial" character of recessive diseases was evident, as family size was large. Today, however, two-children families are generally predominant in industrialized societies. This means that the patient with a recessive disease is very often the only one affected in an otherwise healthy family. However, once an affected child has been born, the genetic risk for any further child of the same parents is 25%. This is important for genetic counseling.

Xeroderma pigmentosum is an autosomal recessive disease (278700). After exposure to ultraviolet light erythema develops, especially in the face, followed by atrophy and telangiectases (Fig. 5.7a). Finally, skin cancers develop that, if untreated, lead to death. Figure 5.7b shows a typical pedigree; here the parents are first cousins. The rate of consanguinity among parents of patients with rare recessive diseases is well above the population average. Usually these parents have inherited this gene from a common ancestor. In Garrod's days this was a powerful tool for recognizing rare recessive diseases; among ten families of alkaptonurics for which this information was available, the parents were first cousins in six cases. Today, however, the consanguinity rate has decreased in most industrialized societies. Hence, even if the rate of consanguinity in families with affected children is substantially increased above the population average, this does not

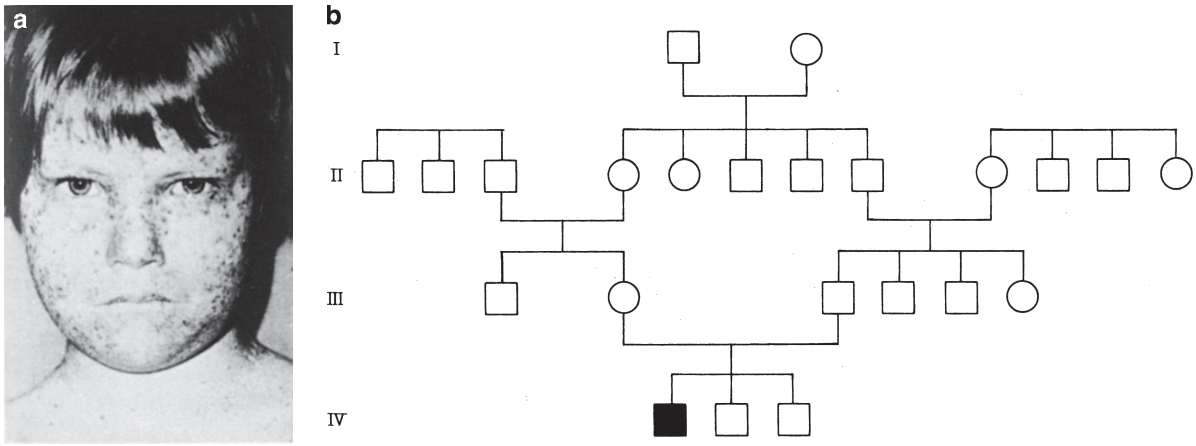

Fig. 5.7 (**a, b**) Xeroderma pigmentosum. (**a**) Girl with this condition (Courtesy of Dr. U. W. Schnyder) (**b**) Pedigree of single case with first-cousin marriage. From Dorn [19]

necessarily lead to the appearance of consanguineous mating when a limited number of families are studied particularly if the abnormal gene is not too rare. This phenomenon together with the small average family size makes it increasingly difficult to recognize an autosomal-recessive mode of inheritance with certainty. Fortunately, however, we no longer need to depend solely on formal genetics. When a rare disease, especially in a child, shows signs of being an inborn error of metabolism, and especially when an enzyme defect can be demonstrated, a recessive mode of inheritance can be inferred in the absence of evidence to the contrary. For purposes of genetic counseling, it must be assumed.

As a rule the vast majority of patients with autosomal-recessive diseases are children of two heterozygotes. Especially decisive for recessive inheritance are the rare matings of two homozygotes with the same anomaly. If both parents are homozygous for the same recessive gene, their mating should exclusively produce affected children. A number of such examples are reported in oculocutaneous albinism (OCA). Some marriages between albinos, however, have produced normally pigmented children [78]. Unless these children are all illegitimate, this proves that the parents must be homozygous for different albino mutations, i.e., more than one albino locus must exist in man. This is the kind of proof that formal genetics can provide to indicate genetic heterogeneity of diseases demonstrating an autosomal recessive mode of inheritance and the same (or a very similar) phenotype. Today, OCA is known as a group of inherited disorders of melanin biosynthesis characterized by a generalized reduction in pigmentation of hair, skin, and eyes. Several types of OCA can be distinguished, with OCA1A (OMIM: 203100) being the most severe type, while OCA1B (OMIM: 606952), OCA2 (OMIM: 203200), OCA3 (OMIM: 203290), and OCA4 (OMIM: 606574) represent milder forms. Each of these four types of OCA is inherited as an autosomal-recessive disorder and at least four genes are responsible for the different types of the disease (i.e., *TYR*, *OCA2*, *TYRP1* and *MATP*) [31].

Another condition for which genetic heterogeneity has been proven in this way is deaf-mutism (Fig. 5.8). Since environmental causes can also cause deafness, it is remarkable that in the pedigree shown here both spouses have an affected sibling, and both parents are consanguineous. Up to date at least 46 genes have been implicated in nonsyndromic hearing loss. The most frequent gene associated with autosomal-recessive nonsyndromic hearing loss is *GJB2*, which is responsible for more than half of cases. Other, relatively frequently implicated genes are *SLC26A4*, *MYO15A*, *OTOF*, *CDH23,* and *TMC1* [39]. Thus, it is likely that in the family shown in Fig. 5.8 the hearing loss was caused by mutations in different genes, e.g., by *GJB2* mutations in one family and *SLC26A4* mutations in the other family. In this scenario the two sons in generation IV would be heterozygous mutation carriers for these two genes; however, this does not result in hearing loss.

5.1.3.1 Pseudodominance in Autosomal Recessive Inheritance

Occasionally matings between an unaffected heterozygote and an affected homozygote are observed. One parent is affected, and the expected segregation ratio among children is 1:1. Since this segregation pattern mimics that found with dominant inheritance,

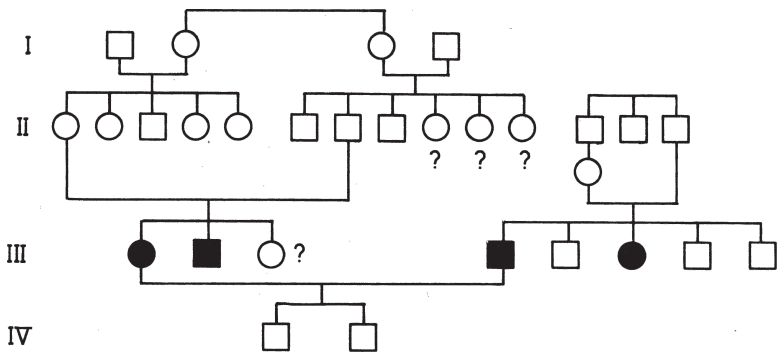

Fig. 5.8 Pedigree of deaf-mutism showing genetic heterogeneity. Both parents are affected with a hereditary type of deaf-mutism; they have affected sibs and come from consanguineous marriages; however, the two sons are not deaf. They are compound heterozygotes for different deaf mutism genes. From Mühlmann [56]

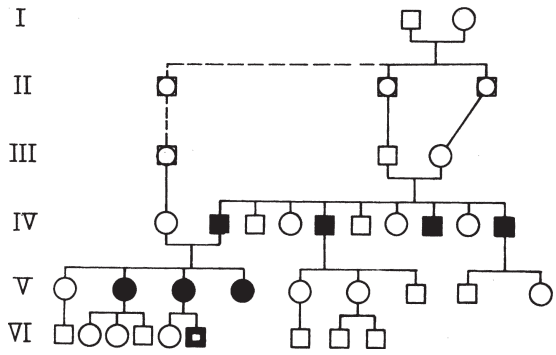

Fig. 5.9 Pedigree of pseudodominance of alkaptonuria, an autosomal-recessive condition. ▣, Suspected alcaptonuric; ⦿, sex unknown. (From Milch [53])

this situation is aptly named "pseudodominance." Fortunately for genetic analysis, such matings are very rare.

Garrod's alkaptonuria (203500) provides an example. In all families described since Garrod the autosomal-recessive mode of inheritance had been confirmed until 1956 when a family with a phenotypically similar but apparently dominant form was reported (Fig. 5.9) – a surprising finding. Some years later the authors had to disavow their conclusions: further family investigations had shown typical, recessive alkaptonuria. A number of marriages between relatives (homozygotes × heterozygotes) had led to pseudodominance. If an individual suffering from a recessive disease mates with a normal homozygote, all children are heterozygotes and hence phenotypically normal. As soon as we learn to treat recessive diseases successfully, marriages of affected but treated homozygotes will increase.

Expressivity is generally more uniform within the same family in recessive than in dominant disorders. Incomplete penetrance seems to be rare. Variability between families, however, may be appreciable.

5.1.3.2 Compound Heterozygotes

When a more penetrating biochemical analysis becomes possible, alleles of different origin frequently have slightly different properties. In an increasing number of instances when the gene is analyzed, and the mutations can be identified, such differences can be explained by the properties of the gene-determined proteins and the impairment of their specific functions.

The genes of hemoglobin α and β chains offer an extreme example. Homozygosity of a mutation within the Hbβ gene, for example, may lead to sickle cell anemia or thalassemia major, depending on the precise place of the base substitution. If there are different substitutions within the two alleles, the resulting phenotype might differ from any one of the two true homozygotes. The phenotype of the compound heterozygote who has the sickle cell mutation in one allele and the HbC mutation in the other is different from that of either homozygote (SS or CC). It depends on the population structure how often homozygous patients with a recessive disease are true homozygotes carrying precisely the same mutation twice, and how often they are compound heterozygotes who carry in their two chromosomes different mutations of homologous genes (Fig. 5.10).

We can be reasonably sure that an affected homozygote carries two copies of the same mutation if both copies have a common origin; for example, if his parents are first cousins and if the condition is very rare. Another source of identity by descent are cases from

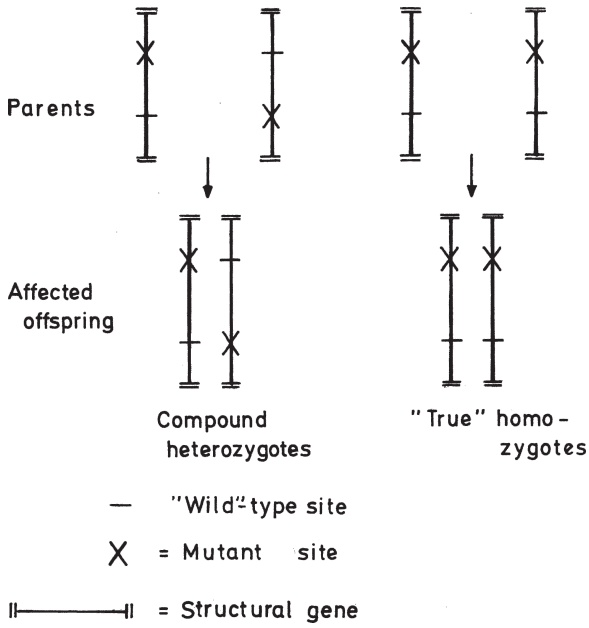

Fig. 5.10 Formation of a compound heterozygote. Each *line* represents the mutant locus on one chromosome in a parent. Among the many possibilities for mutation, two are shown. If parents are heterozygous for mutations that are at identical sites, the affected child is a "true" homozygote; otherwise, he or she is a compound heterozygote

an isolate in which a single mutation – which has been introduced by one individual – became frequent, such as the skin disease called Mal de Meleda (OMIM 248300) on the Croatian island of Mljet. Even in a larger and genetically heterogeneous population group, however, the majority of homozygotes may carry the same gene twice. This happens especially when the gene had a selective advantage some time in the past. The *CFTR* (cystic fibrosis) gene is one example: about 60–70% of all abnormal alleles in northwestern European populations are of the type delta 508, meaning that about 40–50% of patients are indeed homozygous for this mutation ($0.7 \times 0.7 = 0.49$). In other diseases the great majority of "homozygous" individuals are in fact compound heterozygotes. With the progress of DNA studies of human genes this question will be answered directly in an increasing number of instances.

5.1.4 X-Linked Modes of Inheritance

In humans, every mating is a Mendelian backcross with respect to the X and Y chromosomes:

		Paternal gametes	
		X	Y
Maternal gametes	X	$^1/_4$ XX	$^1/_4$ XY
	X	$^1/_4$ XX	$^1/_4$ XY
Total		$^1/_2$ XX♀ +	$^1/_2$ XY♂

This implies that on average female and male zygotes are formed at a 1:1 ratio. This, however, is not quite true. The sex ratio at birth (known as the secondary sex ratio in contrast to the primary sex ratio at conception) is slightly shifted in favor of boys (102–106 boys/100 girls). The primary sex ratio is not known exactly, but there are hints that it is also somewhat variable. The formal characteristics of X-linked modes of inheritance can easily be derived from the mode of sex determination. Many studies on the (primary and secondary) sex ratio have been published. Chromosome studies on abortions should reflect the primary sex ratio and point to a value not too far from 100 (boys and girls in a ratio of 1:1). However, the primary and secondary sex ratio also depend on the interval between sexual intercourse and ovulation, frequency of intercourse, general cultural conditions, and even war and peace. After artificial insemination, the fraction of male offspring appears to be appreciably increased.

5.1.4.1 X-Linked Recessive Mode of Inheritance

If we use A for the dominant, normal wild-type and a for the recessive alleles, the following matings are possible:

(a) AA♀ × A♂. All children have the phenotype A. Neither this nor the analogous mating aa × a is useful for genetic analysis.

(b) AA♀ × a♂. All sons have one of the mother's normal alleles. They are healthy. All daughters are heterozygous Aa. They are phenotypically healthy, but carriers of the abnormal allele. In the analogous, very rare mating aa♀ + A♂ all sons are affected (a), and all daughters are heterozygous (Aa).

(c) Aa♀ + A♂. This type is most important. All daughters are phenotypically normal; half are heterozygous carriers. Half of their sons are hemizygous a and affected. The analogous mating Aa♀ × a♂ is extremely rare. There is a 1:1 ratio of affected and heterozygotes among female children and an 1:1 ratio of affected and normals among males.

The principal formal characteristics of X-linked recessive inheritance can be summarized as follows: Males are predominantly – and in rare X-linked conditions almost exclusively – affected. All their phenotypically healthy but heterozygous daughters are carriers. If no new mutation has occurred, and the mother of the affected male is heterozygous, half of his sisters are heterozygous carriers. Among sons of heterozygous women, there is a 1:1 ratio between affected and unaffected.

Strictly speaking, transmission from affected grandfathers via healthy mothers to affected grandsons is helpful, but not altogether decisive for locating the gene on the X chromosome. An autosomal gene with manifestation limited to the male sex could show the same pattern. The fact that all sons of affected men are unaffected, however, is decisive unless the wife is a heterozygous carrier which may not be unusual for common X-linked traits. This criterion can create difficulties in interpretation when a disease is so severe that the patients do not reproduce.

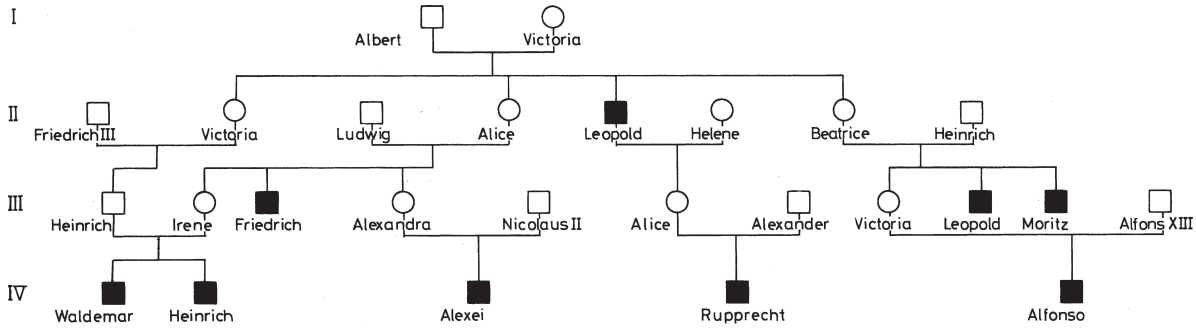

Fig. 5.11 Pedigree of X-linked recessive hemophilia A in the European royal houses. Queen Victoria (I,2) was heterozygous; she transmitted the mutant gene to one hemophilic son and to three daughters

The two most famous and, from a practical standpoint, very important examples are hemophilia A and B (306700, 306900). Due to its alarming manifestations, hemophilia has been known to doctors for a long time and has given rise to the formulation of Nasse's rule (Chap. 1). Figure 5.11 shows the famous pedigree of Queen Victoria's descendants in the European royal houses. One of the hemophilics was the Czarevich Alexei of Russia, and in this case genetic disease influenced politics. Rasputin's power over the imperial couple was based at least partially on his ability to comfort the Czarevich when he was frightened by bleedings. Much larger pedigrees have been described, probably the most extensive being that of hemophilia B in Tenna, Switzerland. As a rule, however, the pedigrees observed in practice are much smaller. Frequently there is only one sibship with affected brothers, or the patient is even the only one affected in an otherwise healthy family. Again, as in dominant conditions (Sect. 5.1.2), this is caused by the reduced reproductive capacity of the patients, which leads to the elimination of most severe hemophilia genes within one or a few generations after they have been produced by new mutation. As expected, almost all hemophilia patients are males. However, there are a few exceptions. Figure 5.12 shows a pedigree from former Czechoslovakia in which a hemophilic had married a heterozygote (who was his double first cousin because in their parents' generation two brothers had married two sisters). The homozygous sisters both had moderately severe hemophilia similar to their affected male relatives.

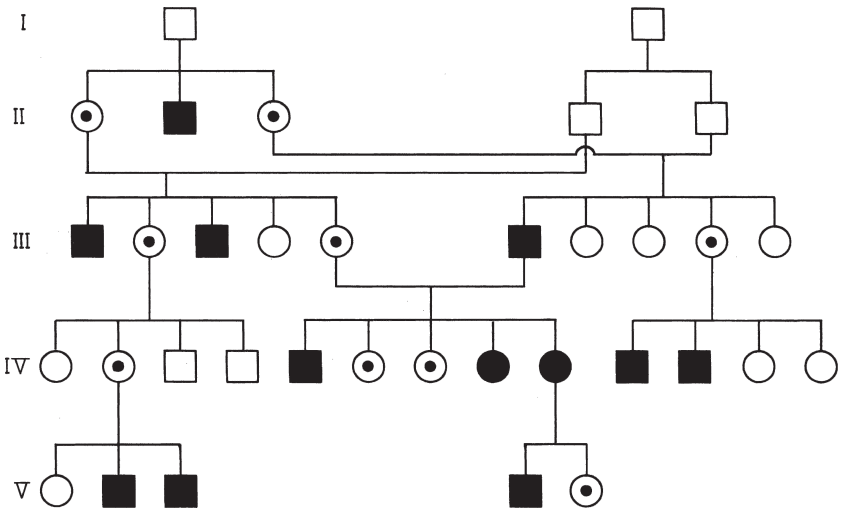

Fig. 5.12 Pedigree of two female homozygotes for X-linked hemophilia. The parents are double first cousins. ⊙, Obligatory heterozygotes. From Pola and Svojitka [64]

Some X-linked conditions have reached considerable frequencies. The most widespread are red-green color vision defects and variants of the enzyme glucose-6-phosphate dehydrogenase, but various types of X-linked mental retardation are also common.

5.1.4.2 X-Linked Dominant Mode of Inheritance

An X-linked dominant condition manifests itself in hemizygous men and heterozygous women. However, all sons of affected males are free of the trait unless their mothers are also affected, and the sons' children are also unaffected. On the other hand, all daughters of affected males are affected. Among children of affected women the segregation ratio is 1:1 regardless of the child's sex, just as in autosomal-dominant inheritance. If affected individuals have a normal rate of reproduction, about twice as many affected females as males are found in the population.

Since only children of affected males provide information in discriminating X-linked dominant from autosomal-dominant inheritance, it is difficult or even impossible to distinguish between these modes of inheritance when the available data are scarce.

The first clearcut example was described by Siemens in [67] in a skin disease that he named "keratosis follicularis spinulosa decalvans (KFSD) cum ophiasi" (308800). The disease manifests follicular hyperkeratosis leading to partial or total loss of eyelashes, eyebrows, and head hair. Severe manifestations were, however, confined to the male members of this family. KFSD is an extremely rare condition as in the last 50 years only 43 additional KFSD cases were identified. A disease-causing gene has not yet been identified [14].

Since then it has been confirmed for all traits with an X-linked dominant mode of inheritance that males are on average more severely affected than females. This finding is no surprise since heterozygous women have a normal allele for compensation, but a satisfactory explanation became possible only when random inactivation of one of the X chromosomes in females was discovered.

Another example of X-linked dominant inheritance is vitamin D-resistant rickets with hypophosphatemia (307800) [93]. In the pedigree shown in Fig. 5.13, all 11 daughters of the affected men suffered from rickets or had hypophosphatemia; all 10 of their sons, however, were healthy. The affected women have both affected and healthy sons and daughters. The probability for the mode of inheritance to be autosomal-dominant and for the affected males to have only affected daughters and only healthy sons is less than 1:10,000. Moreover, in this family male members also tended to be more severely affected than females. Meanwhile, it is established that X-linked hypophosphatemia is caused by mutations in the phosphate-regulating endopeptidase gene (PHEX) [41].

5.1.4.3 X-Linked Dominant Inheritance with Lethality of the Male Hemizygotes [90]

Females with X-chromosomal diseases tend to have milder symptoms than males, as noted above. In some cases the male zygotes may be so severely affected that

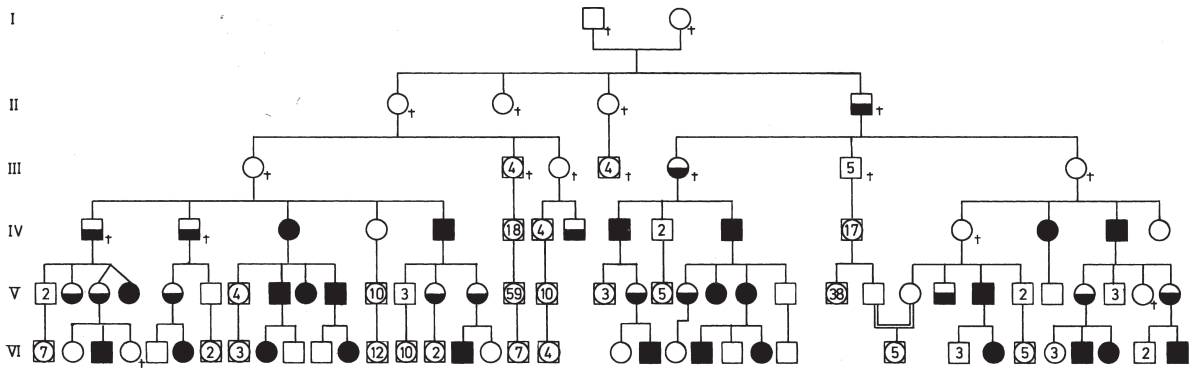

Fig. 5.13 Pedigree of X-linked dominant vitamin D resistant rickets and hypophosphatemia. ■, Hypophosphatemia and rickets; ⬓, hypophosphatemia without rickets. From Winters et al. [93]

they die before birth, and only the females survive. This would result in pedigrees containing only affected females, and among their children affected daughters, normal daughters, and normal sons would be found in the ratio of 1:1:1. Among the male hemizygotes who did not die in very early pregnancy, spontaneous abortions (or male stillbirths) would be expected. W. Lenz in [47] was the first to show that this mode of inheritance exists in humans in the condition known as incontinentia pigmenti (Bloch-Sulzberger; 308300).

Around the time of birth the girls affected with this disease develop inflammatory erythematous and vesicular skin disorders. Later, marblecakelike pigmentations appear (Fig. 5.14a). The syndrome additionally comprises tooth anomalies. Figure 5.14b shows a typical pedigree. The alternative hypothesis would be that of an autosomal-dominant mode of inheritance with manifestation limited to the female sex. The two hypotheses would have the following consequences:

a) With autosomal-dominant sex-limited inheritance, and after proper correction (Sect. 5.3.4), there would be a 1:1 ratio of affected to unaffected among sisters of propositae. All brothers would be healthy. If the population sex ratio is assumed to be 1:1, a sex ratio of 2♂:1♀ would be expected among healthy sibs. With X-linked inheritance, on the other hand, the expected number of healthy brothers is much lower, because one-half of the male zygotes are expected to die before birth (possibly leading to an increased rate of spontaneous miscarriages). Among healthy sibs a 1♂:1♀ ratio would be expected.

b) With autosomal-dominant inheritance the abnormal gene may come from the father or from the mother. Therefore more remotely related affected relatives are to be expected among paternal as well as among maternal relatives. With X-linked inheritance, on the other hand, the gene must come from the mother. Considering the rarity of the condition, additional cases would not occur in the father's family.

c) With autosomal-dominant inheritance the loss of mutant genes per generation would be relatively small compared to the total number of these mutations in the population, since the male carriers, being free of symptoms, would reproduce normally. Therefore, assuming genetic equilibrium, the number of new mutations would be small compared to the overall number of cases in the population. With X-linked inheritance, on the other hand, the loss of zygotes is high due to death

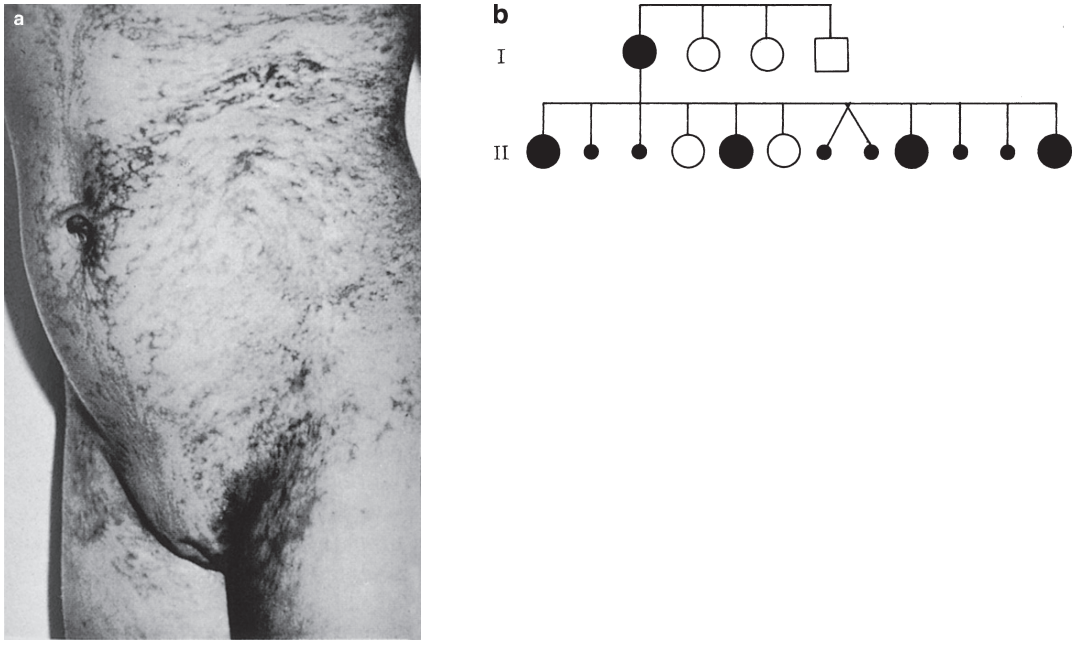

Fig. 5.14 (**a**) Incontinentia pigmenti (Bloch-Sulzberger; courtesy of Dr. W. Fuhrmann). Note the marble cake appearance of skin. (**b**) Pedigree of incontinentia pigmenti. •, Spontaneous abortion; ●, incontinentia pigmenti. From Lenz [47]

of the hemizygote. Hence many of the cases in the population are caused by recent mutation, and extensive pedigrees are rare [7].

The available statistical evidence has consistently supported the hypothesis of an X-linked dominant mode of inheritance with lethality of the male hemizygote. According to Carney et al. [13], 593 female and 16 male cases have been reported. Among the female patients 55% had a positive family history. How can the sporadic males be explained? Of course, the phenomenon of *Durchbrenners* (Hadorn [32] used the term "escapers" – the occasional survival of individuals affected with a lethal genotype) is well known, but Lenz [48] suggested a more specific explanation, assuming, on the basis of a suggestion by Gartler and Francke [27], that a mutation occurs in only one half-strand of the DNA double helix of either the sperm or the oocyte.

Several X-linked syndromes which occur predominantly among females have now been identified. The rareness of affected males in these syndromes is usually attributed to male lethality, which may often occur in the form of early pregnancy loss. About half of the X-linked conditions with predominant expression in females are associated with impairment of cognitive function. Examples include, in addition to the aforementioned incontinentia pigmenti: Aicardi syndrome (OMIM: 304050), focal dermal hypoplasia (Goltz syndrome; OMIM: 305600), Microphthalmia with linear skin defects syndrome (MLS; MIDAS; OMIM: 309801), oral-facial-digital syndrome I (OMIM: 311200), and Rett syndrome (OMIM: 312750). An example for an X-linked syndrome occurring predominantly among females without mental retardation is CHILD syndrome (Congenital hemidysplasia with ichthyosiform erythroderma and limb defects; OMIM: 308050) [74].

5.1.4.4 Genes on the Y Chromosome

Until the 1950s most geneticists were convinced that the human Y chromosome contained genes that occasionally mutate, giving rise to a Y-linked (or holandric) mode of inheritance with male-to-male transmission and males solely being affected. Stern in [72] reviewed the evidence with the result that the time-honored textbook example of Y-linked inheritance of the porcupine man (severe ichthyosis) could no longer be maintained as valid. The only characteristics for which Y-linked inheritance can still be discussed are hairy pinnae, i.e., hair on the outer rim of the ear. A number of extensive pedigrees have been published that show male-to-male transmission. However, the late onset, usually in the third decade of life, and the extremely variable expressivity and high prevalence in some populations (up to 30), makes distinction from a multifactorial mode of inheritance with sex limitation very difficult. Y-linkage can therefore not be fully accepted for this trait.

The Y chromosome contains genes for male differentiation as well as for spermatogenesis.

In experimental animals, segregation ratios deviating from those expected from Mendelian expectations were occasionally reported, one example being the T locus of the mouse [9].

Other cases for which abnormal segregation has been asserted are less well-documented. Since families with many children have become the exception in most industrial societies, the prospect for tracking down and verifying abnormal segregation of pathological genes is becoming more difficult.

5.1.5 "Lethal" Factors [32]

5.1.5.1 Animal Models

Mutations showing a simple mode of inheritance often lead to more or less severe impairment of their bearer's health. There is even evidence (Sect. 5.1.4) that some X-linked conditions prevent the male hemizygote from surviving to birth. It can be assumed that mutations exist which interfere with embryonic development of their carriers so severely as to cause prenatal death.

The first reported case of a lethal mutation in mammalian genetics was the so-called yellow mouse. L. Cuénot [17] reported an apparent deviation from Mendel's law in 1905. A mutant mouse with yellow fur color did not breed true. When yellow animals were crossed with each other, normal gray mice always segregated out. All yellow mice were heterozygous. They all had the same genetic constitution A^Y/A^+; A^Y is a dominant allele of the agouti series, the wild allele of which is termed A^+. When A^Y/A^+ heterozygotes were mated with A^+/A^+ homozygotes, the expected 1:1

ratio between yellow and gray mice was observed. In 1910 it was found that A^Y/A^Y homozygotes are formed but die in utero. Abnormal embryos were later discovered in the expected frequency of 25%.

In this case the allele that is lethal in the homozygous state can be recognized in the heterozygotes by the yellow fur color.

Cases of this sort are exceptional. Generally heterozygotes of lethals are not readily recognizable; therefore lethals occurring spontaneously are difficult to ascertain even in experimental animals and much more so in man.

Usually a lethal mutation kills the embryo in a characteristic phase of its development ("effective lethal phase" [32]). This can easily be explained by the assumption that the action of the mutant gene would be required for further development in this phase.

5.1.5.2 Lethals in Humans

In humans many different types of lethals must occur since many metabolic pathways and their enzymes are essential for survival. It is likely that many still undetected enzyme defects do indeed occur but are not compatible with zygote survival. Moreover, many types of defects of inducer substances needed during embryonic development, and enzymes involved in nucleic acid and protein synthesis, may occur and add to the high incidence of zygote death, which has so far been unexplainable genetically. This problem is discussed from a different standpoint in the context of population genetics (Chap. 16).

According to current estimates, about 15–20% of all recognized human pregnancies end in spontaneous miscarriage. Studies on other mammals suggest that an appreciable number of additional zygote losses go unnoticed, as death occurs during migration through the fallopian tubes. How much of this zygote wastage is due to genetic factors is unknown. A high proportion is caused by numerical or structural chromosome aberrations (Chap. 3). However, there are certainly other maternal causes for abortion as well. While it seemed hopeless to try to relate any proportion of antenatal (or even postnatal) zygote loss to autosomal-dominant or recessive lethals, it appeared more reasonable to speculate about X-linked lethals, as these could influence the sex ratio.

5.1.6 Modifying Genes

So far we have considered phenotypic traits depending on one gene only. However, the phenotypic expression of one gene is usually influenced by other genes. Experiments with animals, especially mammals, show the importance of this "genetic background." One way to overcome analytic difficulties caused by such variation is the use of inbred strains where all animals are genetically alike.

The genetic background is a fairly diffuse concept, but in a number of cases it has been possible to show that penetrance or expressivity of a certain gene can be influenced by another, which is called a "modifier gene" when expressivity is influenced. When penetrance is suppressed altogether, the term "epistasis" (and "hypostasis" of the suppressed gene) is used. In experimental animals cases have been analyzed in which the interaction of two mutations at different loci leads to a completely new phenotype. The classic example is the cross of chickens with "rose" combs and "pea" combs, which leads to the "walnut" comb in homozygotes for both of these mutations. To the best of our knowledge, a similar situation has not been described in man. Modifier genes and epistasis, however, have been demonstrated.

5.1.6.1 Modifying Genes in the AB0 Blood Group System

The best analyzed examples of modifying genes are offered by the AB0 blood group systems. Occurrence of the ABH antigens in saliva (and other secretions) depends on the secretor gene Se. Homozygotes se/se are nonsecretors; heterozygotes Se/se and homozygotes Se/Se are secretors. Hence, se is a recessive suppressor gene. Other rare suppressor genes even prevent the expression of ABH antigens on the surface of erythrocytes.

Bhende et al. [11] discovered a phenotype in 1952 which they called "Bombay" (211100). The erythrocytes were not agglutinated either by anti-A, anti-B or anti-H. The serum contained all three of these agglutinins. Later another family was discovered showing that the bearers of this unusual phenotype did have normal AB0 alleles, but that their manifestation was suppressed (Fig. 5.15; a woman, II, 6, has a Bombay phenotype but

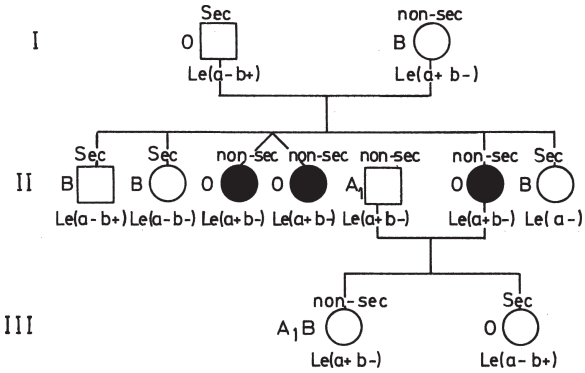

Fig. 5.15 The Bombay blood type. Manifestation of the B antigen is suppressed by a recessive gene x. Note that an O mother (II,6) has an A_1B child. From Bhende et al. [11]

transmitted the B allele to one of her daughters). It was further shown that A can also be suppressed, and the available family data suggested an autosomal-recessive mode of inheritance. In the family shown in Fig. 5.15, the parents of the proposita are first cousins.

The locus is not linked to the AB0 locus. The gene pair was named H, h, the Bombay phenotype representing the homozygote, h/h. The gene has been cloned (see [52]). Depending on the nature of the suppressed allele, the phenotype is designated O_hA_1, O_hA_2, or O_hB. The phenotype has a frequency of about 1 in 13,000 among Maharati-speaking Indians in and around Bombay. A variant with reduced activity is common in the population isolate on Reunion Island [29]. It is caused by the defect of an enzyme that converts a precursor substance into the H antigen, which in turn is a precursor of the A and B antigens [37, 60, 65]. A second gene pair Yy, the rare homozygous conditions of which partially suppresses the A antigen, has been postulated, and subsequently a number of additional families with this condition have been reported.

5.1.6.2 Modifying Genes in Cystic Fibrosis

Cystic fibrosis (CF) is characterized by progressive bronchiectasis, exocrine pancreatic dysfunction, and recurrent sinopulmonary infections. It is a common autosomal recessive disorder with significant morbidity and mortality. The gene, which causes CF, *CFTR*, was already identified 1989; however, the significant phenotypic variation observed in CF suggests that in addition to different mutations in the disease-causing gene and environmental factors, genetic modifiers may contribute to this variability. The identification of such modifiers would have a great potential to improve care for individuals with CF. However, such a modifying effect could to date only be established for a small number of genes. The majority of studies examined the phenotype of lung function and using this parameter, certain alleles of two genes, i.e., transforming growth factor $\beta 1$ (*TGFβ1*) and mannose binding lectin 1 (*MBL2*), were shown to have an effect on lung function [16]. The efforts of identifying modifier genes show some general problems: First, measurable parameters such as lung function are needed to establish a modifying effect. Second, in a disease affecting multiple organs, such as CF, a modifying effect may have an impact only on one organ but not on others. Third, effects of modifying genes are usually moderate and therefore difficult to identify. Newer tools, such as genome-wide association studies, may further contribute to the elucidation of such modifying genes.

5.1.6.3 Sex-Limiting Modifying Genes

In other, less directly accessible traits the action of modifying genes has been analyzed with statistical methods.

Haldane [33] tried in 1941 to identify such genes in HD, using the family data assembled by Bell in [8]. Harris in [36] examined the problem in a condition called diaphyseal aclasis (133700), which is characterized by multiple exostoses near the cartilaginous epiphyses.

The mode of inheritance is dominant; however, the condition is about twice as common in males as in females. It may be transmitted in some families through unaffected females but not through unaffected males. Statistical analysis of the comprehensive pedigree data collected by Stocks and Barrington [75] suggests in part of the families independent segregation of a factor leading to incomplete penetrance only in females: a sex-limiting modifying gene.

5.1.6.4 Modification by the Other Allele

Phenotypic expression of a gene may be modified not only by genes at other loci but also by the "normal" allele. One example comes from the genetics of the Rh

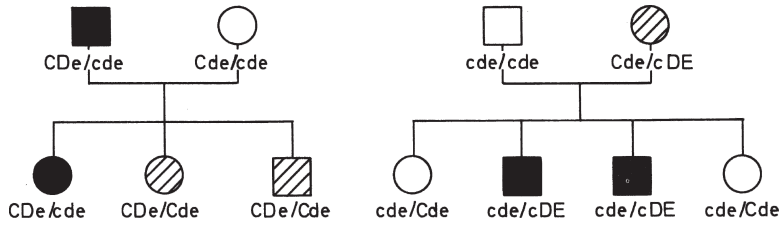

Fig. 5.16 Modification by the homologous allele in the Rh system. ●, D^+ blood with normal reaction; ▨, weak reaction (D^u variant); ○, D^- blood. The haplotype Cde reduces expression of the D factor to D^u. From Ceppellini et al. [15]

factor (Sect. 6.2.4). Occasional blood specimens, when tested with an anti-Rh D serum, give neither a strong positive nor a negative reaction but an attenuated positive reaction. These are called D^u. In most cases a special allele is responsible for this effect, but there are exceptions. In several families the D^u reaction was observed only in family members having Cde as the homologous allele (Fig. 5.16).

5.1.6.5 Modification by Variation in Related Genes

Sickle cell anemia caused by homozygosity for HbS (see Chap. 11) becomes clinically less severe in the presence of several genetic conditions that increase the amount of fetal hemoglobin in the affected red cells. Similarly, the presence of the common alpha thalassemia gene (see Chap. 11) makes for a milder disease manifestation.

5.1.6.6 Modification by a DNA Polymorphism Within the Same Gene

Analysis at the molecular level is revealing new and unsuspected phenomena, including those regarding modification of gene action. Prions are especially interesting proteins. Mutations within the prion gene (176 640) may cause hereditary diseases such as Creutzfeldt-Jakob disease (CJD), Gerstmann-Straussler disease (GSD), or familial fatal insomnia (FFI). The same mutation (Asp → Asn-178) may lead either to CJD or to FFI, depending on a normal polymorphism within the same gene but at a different site: the allele Val 129 segregated in CJD and the allele Met 129 segregated in FFI [30].

Study of various modifying genes and their mechanism is promising to be an important feature for our understanding of the variability of genetic diseases.

The causes of clinical variability in monogenic diseases are:

- Genetic heterogeneity
  - Intra-allelic: different mutations at same locus
  - Inter-allelic: different mutations at other loci
- Modifying genes
  - Additional polymorphisms altering protein conformation
  - Other, as yet unknown mechanisms
- Exposure to various environmental factors required for clinical end result
- Random additional somatic mutations of allele at same locus (e.g., tumors)
- Imprinting (parental origin of mutation)

5.1.7 Anticipation

A time-honored concept popular among physicians in the nineteenth and early twentieth centuries was anticipation. They observed that some hereditary diseases begin earlier in life and follow a more severe course as they progress through generations: the grandfather appeared to be mildly affected; the father was definitely ill, and in the son the disease manifests itself with full force. Anticipation was closely associated with another concept called "degeneration": in some families general, mental, and physical qualities were thought to deteriorate through the generations. These ideas became popular not only among physicians but also among the general public, and were expressed in literary works such as Thomas Mann's novel *Die Buddenbrocks.* In two diseases that tend to manifest during adult life, anticipation seemed to be obvious: HD and myotonic dystrophy (160900) [25]. In the latter, myotonia is associated with relatively mild muscular dystrophy, cataracts, and sometimes mental retardation, or dementia. This disease shows an unusual degree of variability in age at onset, and earlier onset

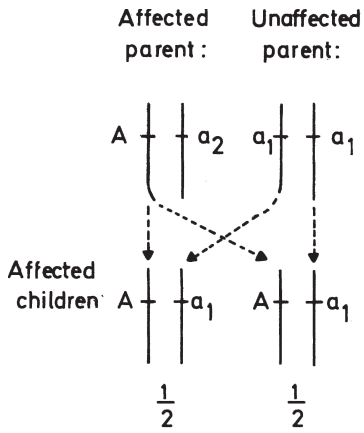

Fig. 5.17 Allelic modification. If manifestation of a dominant, abnormal gene A is modified by the normal allele, and if the allele a_1 causes severe and a_2 milder manifestation of A, there is a correlation in the degree of manifestation between affected sibs but not between affected parent and child. An affected child cannot receive the modifying a_2 allele

as well as a more severe course in some patients of the most recent generation.

When Mendel's laws were rediscovered, anticipation did not fit the new, and otherwise so successful, theory. Therefore scientists interested in genetic problems tried to explain these phenomena away with sophisticated arguments (which were also used in the first two editions of this book). Weinberg [86] pointed out, for example, that anticipation can easily be simulated if families were ascertained directly by patients of the youngest generation who were affected early in life. Their parents and grandparents, on the other hand, who were ascertained through these young probands, could be recognized only if the onset of the disease was so late that they had a chance to have children.

Penrose [63], one of the best human geneticists of his time, explained in great detail that anticipation could be mimicked if ascertainment through the youngest generation combined with dissimilarity of age at onset between parents and children, but similarity between sibs. This would be expected if, in a dominant condition, the normal allele influenced the degree of manifestation of the mutant allele (allelic modification; Fig. 5.17). There can be little doubt that explanations given by Weinberg and Penrose are correct in some instances. However, in HD and myotonic dystrophy molecular analysis revealed specific types of mutations whose effects increase with passage through succeeding generations.

In HD, patients with early onset are more likely to have inherited their mutant genes from the father, whereas late onset is more common when the gene comes from the mother. In myotonic dystrophy, on the other hand, cases of very early onset are less rare; the babies have signs of the disease even at birth. This occurs almost exclusively when the mothers are affected.

Such differences have also been observed in some other monogenic diseases (Table 5.2). On page 169 the huntingtin gene (initially designated "IT15" for "important transcript 15") and the mutations leading to HD were described: amplification of a $(CAG)_n$ repeat beyond 40 copies causes the disease. Moreover, these amplification products are unstable; the predominant tendency appears to be toward an increase in copy numbers by further rounds of amplification; a reduction in copy numbers may occur but is apparently rarer. Higher copy number, on the other side, correlates with earlier onset: a convincing explanation for anticipation.

In myotonic dystrophy an analogous explanation has been found [12, 26, 35]. Here an unstable, amplified sequence was found in the 3′ untranslated region of a gene whose product was predicted to be a member of a protein kinase gene family [12]. It is a $(CTG)_n$ repeat. In normal individuals between 4 and 37 CTG

Table 5.2 Dominant diseases in which parental origin influences the disease (modified from Reik [66])

Disorder	Chromosome	Observations
Huntington disease	4	Early onset frequently associated with paternal transmission
Spinocerebellar ataxia	6	Early onset with paternal transmission
Myotonic dystrophy	19	Congenital form almost exclusively with maternal transmission
Neurofibromatosis I	17	Increased severity with maternal transmission
Neurofibromatosis II	22	Earlier onset with maternal transmission
Wilms tumor	11	Loss of maternal alleles in sporadic tumor
Osteo-sarcoma	13	Loss of maternal alleles in sporadic tumors

repeats are found, 38 to 49 CTG repeats are a premutation. Affected patients may have between 50 and some 2,000 repeats or even more. The repeat number tends to increase over the generations [35]; it is correlated with age at onset and severity of the disease, explaining anticipation.

Thus, the sex of the transmitting parent is an important factor that determines the trinucleotide repeat allele size in the offspring. In the case of myotonic dystrophy it has been speculated that because expansion of the CTG repeat is more rapid with male transmission, negative selection during spermatogenesis may be required to explain the almost exclusive maternal inheritance of severe congenital onset myotonic dystrophy.

5.1.8 Total Number of Conditions with Simple Modes of Inheritance Known So Far in Humans

For many years McKusick has undertaken the task of collecting and documenting known conditions with simple modes of inheritance in man. This extremely valuable resource is now known as OMIM (Online Mendelian Inheritance in Man; www.ncbi.nlm.nih.gov/omim). This web-based full-text, referenced compendium of human genes and genetic phenotypes has the advantage that it can be updated daily, and the entries contain links to other genetics resources. OMIM contains information on all known Mendelian disorders. Table 5.3 provides the number of OMIM entries for autosomal, X-linked, Y-linked, and mitochondrial genes as of 25 May 2009 with information regarding known sequences and phenotypes. Enumeration of dominant and recessive entries was discontinued by OMIM 10 years ago. Note a total number of 19,462 entries, which should be compared with the estimation of about 25,000 human genes based on molecular data. While genetic polymorphisms are included, most conditions listed in this register are rare. Many are rare hereditary diseases. At first glance the list is impressive. However, more detailed scrutiny of the conditions shows that our knowledge of these rare diseases is not nearly as good as it should and could be. There are several reasons:

(a) Most hereditary diseases have become known by occasional observation of affected patients and their families. With rare diseases it is difficult to assess whether they do or do not have a genetic basis. Here, next-generation sequencing or third-generation sequencing may pave the way to finding possible genetic bases in rare diseases.
(b) Some recessive diseases have become known because they happened to be frequent in special populations, primarily in isolates. Isolate studies permit examination of the manifestation of recessive diseases caused by a single mutation. One problem with this approach is that chance determines which genes are studied.
(c) Most human and medical geneticists are working in relatively few industrialized countries. However, genes for rare diseases show a very unequal distribution in different populations. This is particularly true for recessives but has also been shown for dominants with normal or only slightly lowered biological fitness, i.e., when the incidence is not determined by the mutation rate. Hence the developing countries can be expected to abound with hereditary anomalies and diseases that are unclassified to date. Any medical geneticist who has ever walked through, say, an Indian village

Table 5.3 Number of OMIM Entries, 25 May 2009

	Autosomal	X-Linked	Y-Linked	Mitochondrial	Total
* Gene with known sequence	12,111	581	48	37	12,777
+ Gene with known sequence phenotype	347	25	0	0	372
# Phenotype description, molecular basis known	2,293	207	2	26	2,528
% Mendelian phenotype or locus, molecular basis unknown	1,598	141	5	0	1,744
Other, mainly phenotypes with suspected Mendelian basis	1,900	139	2	0	2,041
Total	18,249	1,093	57	63	19,462

knows that this suggestion is not merely a theoretical speculation.

(d) Genetic defects with simple modes of inheritance have a good chance of being detected when they show a clearcut phenotype that is readily recognizable. This is why the inherited conditions of the skin and eye are relatively well known. Other defects, however, may cause anomalies or diseases in some families that are precipitated by environmental factors. Most of such hidden defects are unknown at present.

(e) The real significance of hereditary disease and its total impact can be established only by studies in large populations, using epidemiological methods. Such studies offer the opportunity to detect heterogeneity in etiology and to aid in distinguishing genetic and nongenetic causes. They afford the only basis on which genetic parameters such as mutation rates, biological fitness, and the relative incidence of mild and severe mutations of the same gene can be established. They also help in predicting the long-term and public health effects of medical therapy and of genetic counseling for future generations.

5.1.8.1 Difference in the Relative Frequencies of Dominant and Recessive Conditions in Humans and Animals?

At first glance, there appears to be a difference between humans and experimental animals in the relative frequencies of dominant and recessive conditions. Of the better known mutants of *Drosophila melanogaster* 200 are recessive and only 13 (6.1%) dominant. In the chicken, 40 recessive and 28 dominant mutations have been reported. In the mouse only 17 of 74 mutants are dominant (23%) and the rest recessive. In the rabbit 32 recessive and 6 dominant mutations have been found. (Instances of multiple allelism are counted as one gene locus.) In humans, on the other hand, more dominant than recessive conditions are known. This discrepancy, however, is likely to be caused by diagnostic bias. Our species observes itself most carefully; therefore, defects are detectable that would probably escape observation when present in experimental animals. It would be difficult, for example, to detect brachydactyly in the mouse. This condition, however, leads to a much more severe defect when homozygous. Hence such a defect, dominant in man, would be counted as recessive in the mouse. Another reason might be that the population of industrialized countries is not in equilibrium for recessive genes. The frequency of consanguineous matings has dropped sharply, and therefore the chance of a recessive gene meeting another mutation in the same gene and becoming homozygous is reduced. A new equilibrium will be reached only in the very distant future when recessive genes could become sufficiently frequent again. In our opinion, there is no significant reason to assume that humans are unique in regard to the ratio of dominant and recessive mutations.

5.1.9 Uniparental Disomy and Genomic Imprinting

In 1980 Eric Engel of the University of Geneva published a paper in which he discussed the possibility of having a chromosomal pair derived from only one parent [21]. He termed this possibility "uniparental disomy" (UPD). The original article included calculations on the potential frequency of UPD; he predicted that as many as 3 individuals out of 10,000 might have UPD for one of the chromosomes involved in common aneuploidies such as 15, 16, 21, 22, and sex chromosomes. Eight years later, the team of Art Beaudet published in the *American Journal of Human Genetics* a case of UPD for chromosome 7 in a female with short stature, cystic fibrosis, and growth hormone deficiency [4]. The authors published a list of possibilities for the mechanism of UPD7 and favored a monosomy 7 conception followed by mitotic nondisjunction or replication of the solitary chromosome 7. The UPD7 in this case was maternal in origin, i.e., there were two chromosomes 7 from the mother and no chromosome 7 contribution from the father. Nonpaternity was obviously convincingly excluded. Isodisomy refers to the case in which the two homologues are identical in sequence (one parental chromosome duplicated); heterodisomy refers to the case in which the two homologues differ (both parental chromosomes inherited). Isodisomy and heterodisomy could be complete, i.e., for the entire chromosome, or partial (segmental) due to recombination events in the parental chromosomes.

The detection of UPD could be done with DNA analysis of the proband and the parents. Single

nucleotide polymorphisms, or short sequence repeat polymorphisms, could be used to mark the parental chromosomes, to follow the inheritance, and determine the UPD. In addition the genotyping of the DNA variants could determine the iso- or heterodisomy, either complete of segmental.

UPD has been observed for almost all human chromosomes [20]. The mechanisms resulting in UPD are multiple and include:

(a) "Trisomy rescue" refers to the loss of a chromosome from an initial trisomy (Fig. 5.18). Such reduction from a trisomy to a disomy results from two errors, one meiotic leading to a trisomy state after fertilization by a normal gamete, the other mitotic, removing the supernumerary chromosome by nondisjunction or anaphase lag. Trisomy rescue as a cause of UPD contributes primarily to cases of maternal UPD since most segregation errors occur in oogenesis.

(b) "Gamete complementation" is a mechanism by which a nullisomic gamete meets a disomy gamete. This mechanism implies two errors, one in each sex (Fig. 5.19).

(c) "Rescue of a monosomy" refers to the duplication of a singly inherited chromosome (Fig. 5.20). Such "correction" from a monosomy to a disomy results also from two errors, one meiotic leading to monosomy, the other mitotic duplicating the solitary chromosome.

(d) Somatic recombination (i.e., somatic crossing-over, the symmetrical "trading" of a paternal and maternal homologous chromatid segment) may also be the source of segregants producing cells with segmental UPD (Fig. 5.21).

(e) Chromosomal translocations particularly of acrocentric chromosomes have been found in numerous cases of UPD (Fig. 5.22). Heterologous Robertsonian translocations (of different acrocentrics), or homologous Robertsonian translocations (of the same acrocentric), as well as other translocations provide increased risk for UPD.

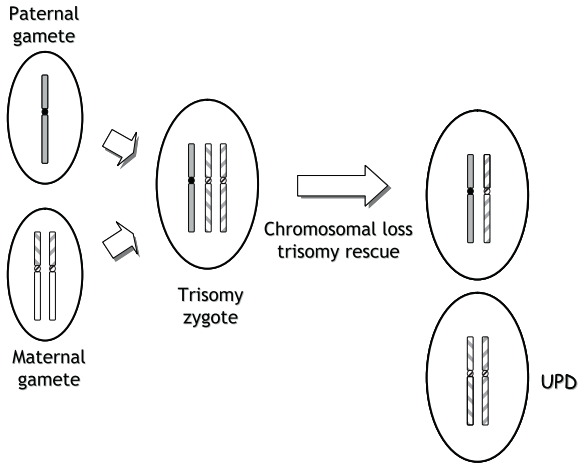

Fig. 5.18 Schematic representation of the mechanism of UPD due to trisomy rescue

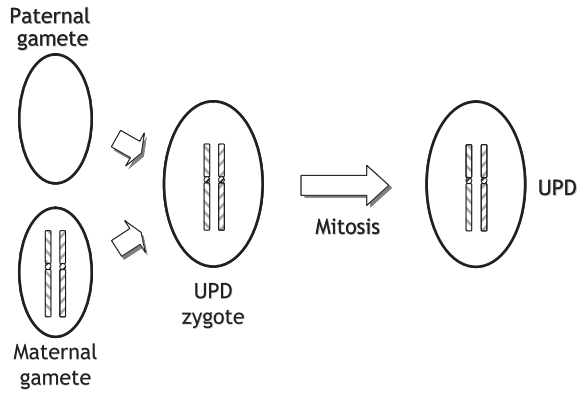

Fig. 5.19 Schematic representation of the mechanism of UPD due to gamete complementation

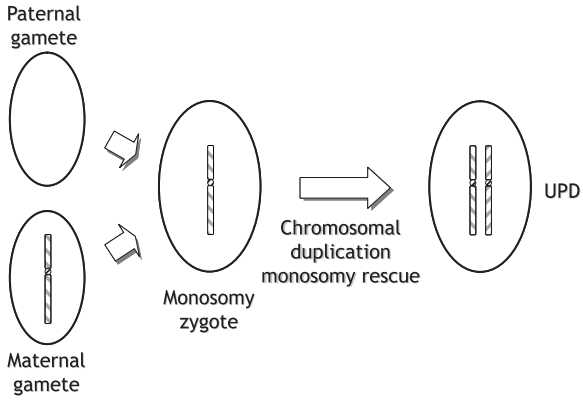

Fig. 5.20 Schematic representation of the mechanism of UPD due to monosomy rescue

5.1.9.1 Phenotypic Consequences of UPD

There are two main reasons for the phenotypic consequences of UPD:

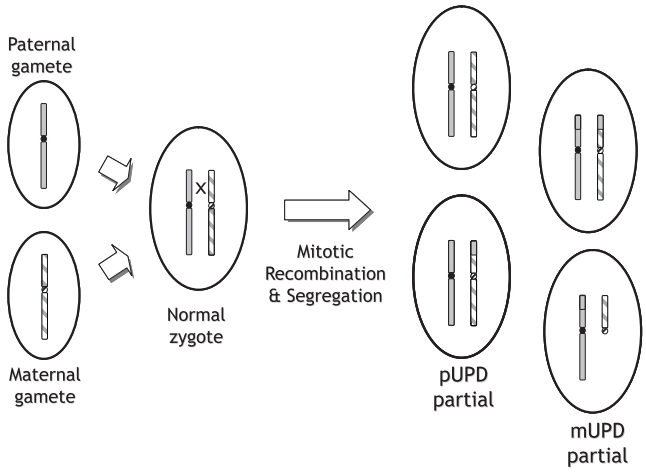

Fig. 5.21 Schematic representation of the mechanism of partial UPD due to somatic–mitotic recombination

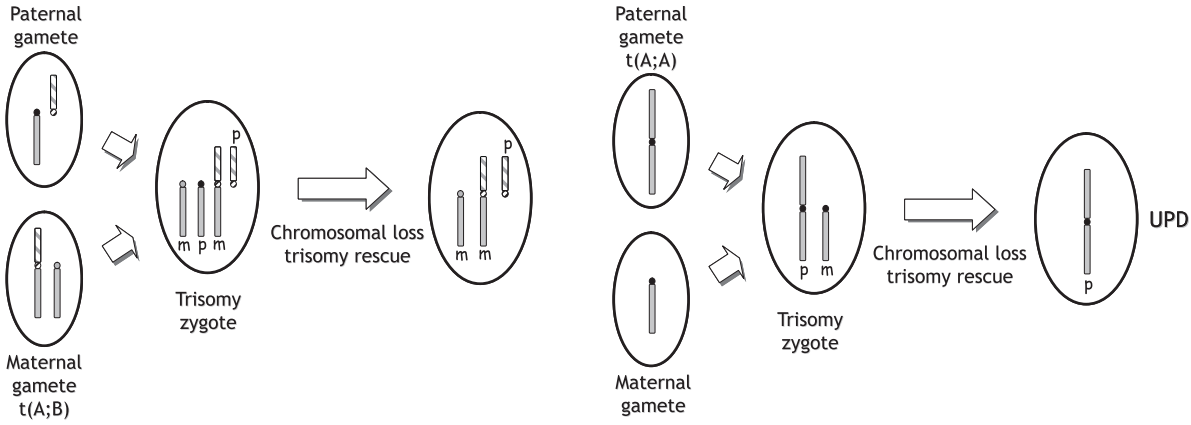

Fig. 5.22 Schematic representation of the mechanism of partial UPD due to translocations of acrocentric chromosomes. The *left panel* depicts a case of a translocation involving two different acrocentrics; the *right panel* shows a case of a translocation involving homologous acrocentrics

1. Duplication of autosomal recessive alleles. In isodisomy, two copies of a mutant allele would result in the disease phenotype. In the originally described case of maternal UPD7, cystic fibrosis was due to two maternally derived copies of the Gly542Ter mutation of the CFTR genes [4] (the mother in that case was a heterozygous carrier of this mutation).

2. Parental imprinting effects. Genomic imprinting refers to parent-of-origin dependent gene expression. Some genes are monoallelically expressed either from the paternally or the maternally derived chromosome. Thus for a paternally-only expressed gene, maternal UPD would result in a null phenotype for this gene. On the other hand, for a maternally-only expressed gene, paternal UPD would result in a similar null phenotype. An example of the former is Prader-Willi syndrome caused by matUPD15; and of the latter is Angelmann syndrome caused by patUPD15 [57].

A considerable number of imprinted genes have been identified in human and mouse [55].

5.1.9.2 Human Disorders Involving UPD

Rare cases of UPD for almost all chromosomes have been identified; the phenotypes are variable. Among them there are some recognizable syndromes which include: (a) Prader-Willi syndrome (matUPD15); (b) Angelmann syndrome (patUPD15); (c) Beckwith-Wiedemann syndrome (patUPD11p15); (d) neonatal transient diabetes mellitus in patUPD6; (e) maternal and paternal UPD14 syndromes; (f) some cases of Russell-Silver syndrome (matUPD7).

5

5.1.10 Diseases Due to Mutations in the Mitochondrial Genome

As shown in Chapter 2, the mitochondrial genome, mtDNA, consists of a ring-shaped chromosome with 16 596 bp. It encodes a small (12 S) and a large (16 S) rRNA for mitochondrial RNA translation, 22 tRNAs, and 13 genes encoding subunits of the respiratory chain. All these polypeptides are subunits of the mitochondrial energy-generating pathway, oxidative phosphorylation (OXPHOS). OXPHOS encompasses five multiunit enzyme complexes, arrayed within the mitochondrial inner membrane; most of the peptides necessary for building these enzyme complexes are encoded in nuclear genes.

At fertilization the oocyte contains about 200,000 mtDNAs. Once fertilized, the nuclear DNA replicates and the oocyte cleaves, but the mtDNA does not replicate until after the blastocyst is formed. Since the blastocyst cells that are destined to become the embryo proper constitute only a small fraction of all blastocyst cells, and only a fraction of these cells enter the female germ line, few of the oocyte's mtDNA molecules are found in the primordial germ cells. However, it is questionable whether this mechanism is sufficient for creating an mtDNA "population" in human cells that is as homogeneous, as is normally found, especially if we consider the fact that a single mitochondrium contains 5–10 mtDNA molecules.

Most proteins necessary for development of the mitochondria themselves are produced by nuclear genes. Therefore some of the diseases due to malfunction of mitochondria are caused by defects of such genes; they follow classical Mendelian modes of inheritance [82, 84]. On the other hand, diseases due to defects of genes in the mitochondrial genome are transmitted as the mitochondria themselves, i.e., from the mother to all children, irrespective of sex. However, considering the great number of mitochondria that a oocyte contains, and the number of genomes per mitochondrium, it is not surprising that a child may inherit from its mother more than one type of mitochondrial genome; cells containing variable proportions of affected mitochondria are "heteroplasmic." During further development, one genome may become more abundant; different cell lineages may even become "homoplasmic" for different mitochondrial genomes. This may explain in part the enormous phenotypic variation between individuals with the same mitochondrial disease. A heteroplasmic mtDNA mutation may reduce the function of the gene-determined peptide. In most instances this is unimportant, but in a few cells the fraction of mitochondria containing the mutant increases to the extent that OXPHOS enzyme activity decreases until it falls below the cellular or tissue energetic threshold, i.e., the minimum activity necessary to sustain oxidative phosphorylation. Because OXPHOS is necessary for nearly all cells, any organ can be affected in mitochondrial diseases. Thus, respiratory chain deficiencies caused by mitochondrial disorders may generate almost any symptom, in any organ system, and at any stage of life. The heteroplasmy produces marked variability in the severity and symptom patterns of these conditions. The most severe inherited mitochondrial disorders become clinically apparent during infancy, whereas other disorders of mitochondrial function may have an adult onset.

Four categories of diseases due to mutations in the mitochondrial genome may be distinguished (Fig. 5.23) [84]. In the first we find missense mutations with relatively mild phenotypic effects. These are transmitted maternally and appear to be homoplasmic. The second category comprises deleterious point mutations. Of course they can be transmitted maternally only if they are heteroplasmic. The third category, deletion mutants, occur by new mutations during early development, and these are therefore heteroplasmic. In the fourth category of diseases, certain mutations may be present that diminish OXPHOS activity somewhat at onset but not sufficiently to cause functional damage. During life time, however, additional random mutations accumulate in somatic cells, reducing their OXPHOS capacity until the threshold is reached. Then a degenerative disease of advanced age such as Alzheimer or Parkinson disease might ensue.

5.1.10.1 Leber Optical Atrophy

An example of the first category is Leber's hereditary optical neuropathy, (LHON; 308900) [82, 83]. In this disease, rapid vision loss occurs during young adult age; cardiac dysrhythmia is common. Variation in severity of the disease is strong; males are more often and on average more severely affected than females; the proportion of transmitting females in the family is much larger than expected if the mutation were

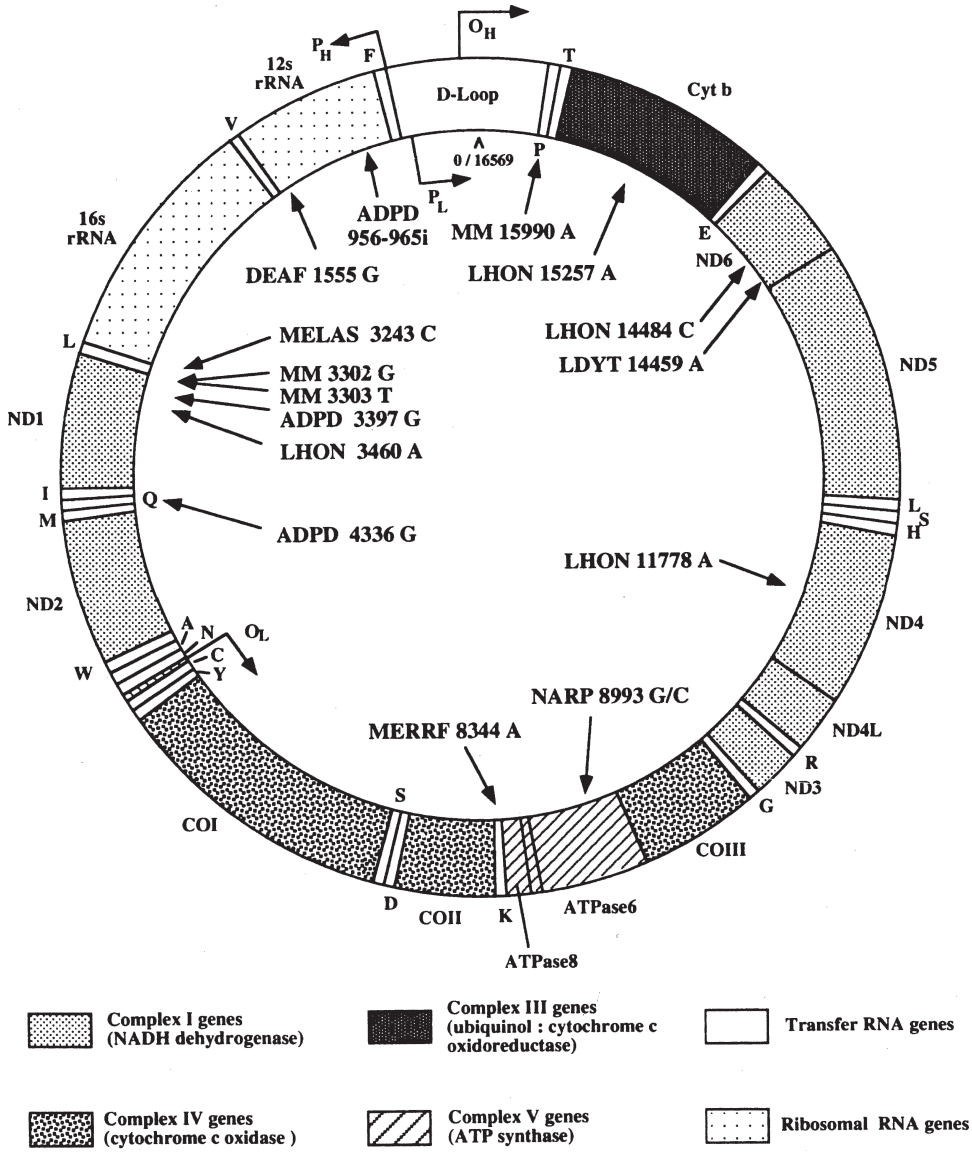

Fig. 5.23 Human tDNA map showing locations of genes and mutations Definitions of gene symbols and mutations, example: MTTK*MERRF8344A.MTTK is the altered mtDNA (MT) gene for tRNA (Lys) (TK); Myoclonic epilepsy and Ragged Red Fiber disease (MERRF) is the most characteristic clinical presentation, 8344 is the altered nucleotide, and A is the pathogenic base. From Wallace [84]

X-linked. Transmission, however, occurs exclusively through females [80]. Molecular analysis revealed a G → A transition (G3460A) leading to an Arg → His replacement in the gene for the NADH subunit 4. The Arg residue must be important for function since it has been conserved in evolution from flagellates and fungi to humans. The mutation is homoplasmic; hence the clinical variability as well as the sex difference must have other causes that are still unknown. Other mitochondrial mutations in closely related genes have occasionally been described [40].

Two diseases apparently belong to the second category – deleterious but heteroplasmic point mutations. In a large kindred, Leber disease was found to be associated with infantile bilateral striatal necrosis. In this family four phenotypes were found: normal, Leber disease, striatal necrosis, and the combination of the two diseases. All members were related through the

female line. Since careful analysis has shown no deletion, the disease appears to be due to a deleterious but heteroplasmic point mutation. Depending on the preponderance of the aberrant mtDNA, the clinical signs vary [82]. The second disease of this class is one combining myoclonic epilepsy and mitochondrial myopathy – both conditions with huge interindividual variation [18].

5.1.10.2 Deletions

The third category is that of sporadic and heteroplasmic deletions. These occur as somatic mutations; since all deletions in one individual are identical, they must have arisen by clonal expansion of a single molecular event. Therefore a selective advantage of mutant cells has been suggested [82]. Figure 5.24 shows such deletions. Clinical manifestations again depend on the distribution of mutant mitochondria. A family has been described [94] in which multiple deletions of mtDNA behaved as one autosomal-dominant trait. The affected individuals suffered from progressive external ophthalmoplegia, progressive proximal weakness, bilateral cateract, and precocious death.

5.1.10.3 Diseases of Advanced Age

The fourth category comprises diseases of advanced age that have not found satisfactory explanations so far. In both Alzheimer and Parkinson diseases, for example, pedigrees have been observed in which relatively early onset in middle age is combined with an autosomal-dominant mode of inheritance. In the majority of these cases, however, an accumulation of affected individuals within families is found but no combination of clearcut Mendelian mode of inheritance with onset at more advanced age. Here, mildly to moderately deleterious germ line mutations established in the distant past, and present in a certain proportion of the population in combination with somatic mutations occurring during lifetime of the individual, may lead to such degenerative diseases. For example, a homoplasmic mutation among whites at nucleotide base pair 4,336 leading to a tRNA mutant has been observed in 5% of Alzheimer and Parkinson disease mutations, but appears to be much rarer in the general population. It may contribute to the multifactorial origin of these diseases [76].

In general, mutations within the mitochondrial genome affect mainly organ systems that depend on intact oxidation – central nervous system and muscles. Probably the number of known diseases due to mutations in the mitochondrial genome will increase in future (Table 5.4).

5.1.10.4 Interaction Between Nuclear and Mitochondrial Genomes

Several subunits of the electron transport chain are not encoded within the mitochondrial DNA but by the nuclear DNA. As a consequence, mutations in the nuclear genome can cause secondary mitochondrial DNA information loss. Hence, there are some clinical syndromes in which defects in OXPHOS follow classic Mendelian patterns of dominant-recessive transmission and not the maternal pattern, which is usually associated with this group of disorders. An example is the mitochondrial neurogastrointestinal encephalopathy syndrome (MNGIE; OMIM: 603041) which can be caused by mutations in the gene encoding thymidine phosphorylase (ECGF1).

5.1.11 Unusual, "Near Mendelian" Modes of Inheritance

As a "bridge" between monogenic and polygenic phenotypes it is worth mentioning two concepts that provide an understanding of the increased complexity between genotype and phenotype [2].

5.1.11.1 Digenic Inheritance

In this case, the phenotype is due to one mutant allele in each of two different genes. The first example was that of one form of retinitis pigmentosa published in 1994 from the laboratory of T. Dryja [42]. Individuals with a mutation in the ROM1 gene (OMIM 180721 on chromosome 11q13) *AND* a mutation in the RDS gene (OMIM 179605 on chromosome 6p21) manifest the disease (Fig. 5.25). However, individuals with only heterozygosity of the ROM1 gene mutation, or with only heterozygosity of the RDS gene mutation, were not affected

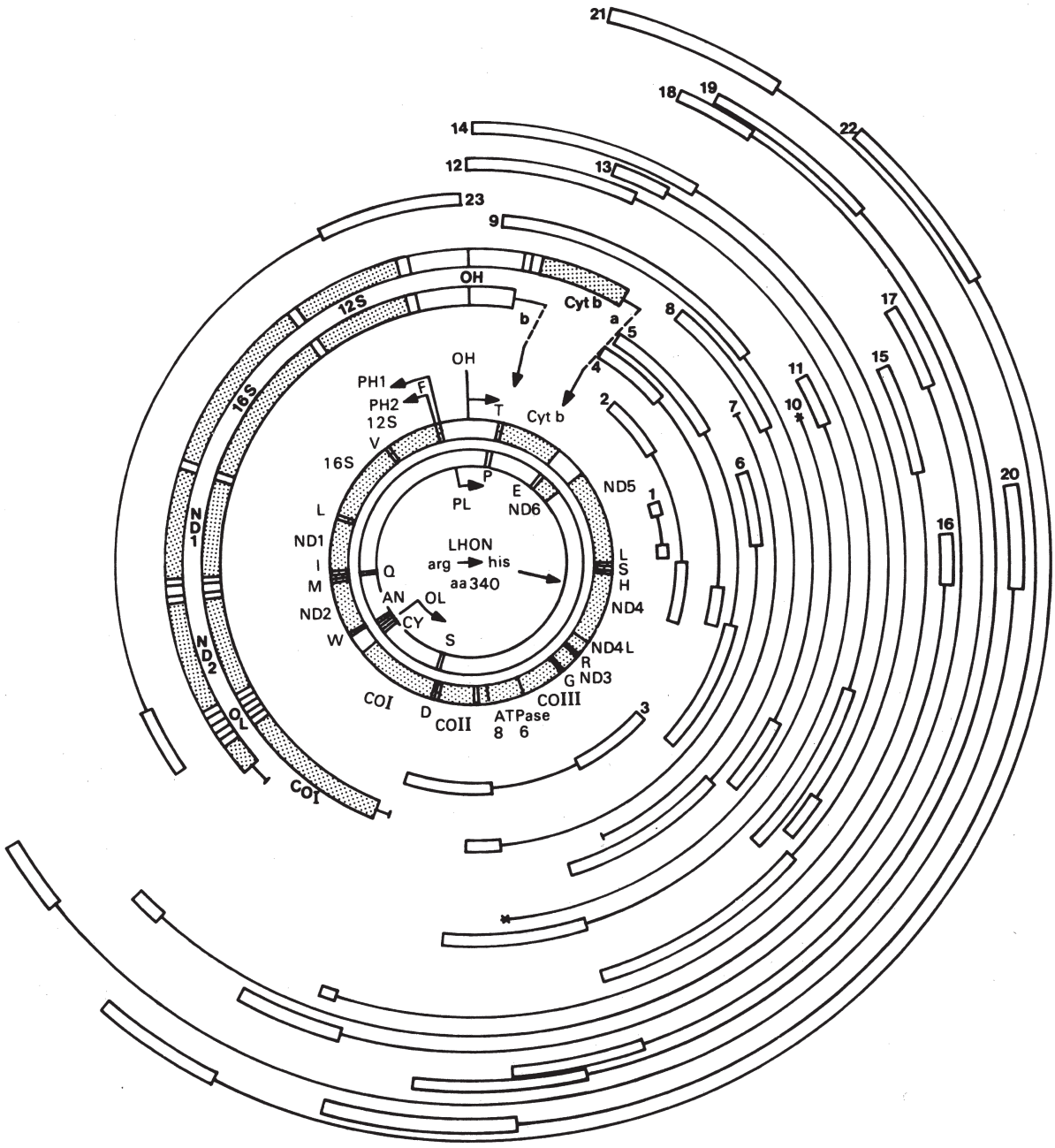

Fig. 5.24 Deletion map of human mtDNA. The *inner circles* show localization of genes and mutations. (See also Fig. 5.23). The arcs no. 1–23 show the mtDNA regions that were lost in various deletions. The *open bars* at the end of the arcs show regions of uncertainty. Deletion 1 was found in a patient with Myoclonic Epilepsy and Ragged Red Fibre Disease (MERRF) together with stroke-like symptoms. Deletions 2–23 were found in ocular myopathy patients with symptoms of varying severity. Deletion 10 was found in about one third of all ocular myopathy patients. The * at the ends of deletion 10 represents the associated 13 base pairs direct repeat. The two partial mtDNA maps labeled "a" and "b" to the left of the function map indicate the regions that were tandemly duplicated in patients with ocular myopathy associated with diabetes mellitus. The insertion sites around MTCYB (cytb) are indicated by *arrows*. From Wallace [83]

5

Table 5.4 The Mitochondrial Chromosome from McKusick [52]

Location (nt)	Symbol	Title	MIM	Disorder[a]
577–647	MTTF	tRNA phenylalanine	590070	
648–1,601	MTRNR1	12 S rRNA	561000	Deafness, aminoglycoside-induced, 580,000
1,602–1,670	MTTV	tRNA valine	590105	
1,671–3,229	MTRNR2	16 S rRNA	561010	Cloramphenicol resistance, 515,000
3,230–3,304	MTTL1	tRNA leucine 1 (UUA/G)	590050	MELAS syndrome, 540,000; MERRF syndrome, 545,000; Cardiomyopathy; Diabetes-deafness syndrome, 520,000
3,307–4,262	MTND1	NADH dehydrogenase 1	516000	Leber optic atrophy, 535,000
4,263–4,331	MTTI	tRNA isoleucine	590045	Cardiomyopathy
4,400–4,329[b]	MTTQ	tRNA glutamine	590030	Cardiomyopathy
4,402–4,469	MTTM	tRNA methionine	590065	
4,470–5,511	MTND2	NADH dehydrogenase 2	516001	Leber optic atrophy, 535,000
5,512–5,576	MTTW	tRNA tryptophan	590095	
5,655–5,587[b]	MTTA	tRNA alanine	590000	
5,729–5,657[b]	MTTN	tRNA asparagine	590010	Ophthalmoplegia, isolated
5,826–5,761[b]	MTTC	tRNA cysteine	590020	Ophthalmoplegia, isolated
5,891–5,826[b]	MTTY	tRNA tyrosine	590100	
5,904–7,444	MTCO1	cytochrome c oxidase I	516030	
7,516–7,445[b]	MTTS1	tRNA serine 1 (UCN)	590080	
7,518–7,585	MTTD	tRNA aspartic acid	590015	
7,586–8,262	MTCO2	cytochrome c oxidase II	516040	
8,295–8,364	MTTK	tRNA lysine	590060	MERRF syndrome, 545,000
8,366–8,572	MTATP8	ATP synthase 8	516070	
8,527–9,207	MTATP6	ATP synthase 6	516060	Leigh syndrome; NARP syndrome, 551,500
9,207–9,990	MTCO3	cytochrome c oxidase III	516050	Leber optic atrophy 535,000
9,991–10,058	MTTG	tRNA glycine	590035	
10,059–10,404	MTND3	NADH dehydrogenase 3	516002	Leber optic atrophy, 535,000
10,405–10,469	MTTR	tRNA arginine	590005	Leber optic atrophy, 535,000
10,470–10,766	MTND4L	NADH dehydrogenase 4 L	516004	
10,760–12,137	MTND4	NADH dehydrogenase 4	516003	Leber optic atrophy, 535,000
12,138–12,206	MTTH	tRNA histidine	590040	
12,207–12,265	MTTS2	tRNA serine 2 (AGU/C)	590085	
12,266–12,336	MTTL2	tRNA leucine 2 (CUN)	590055	
12,337–14,148	MTND5	NADH dehydrogenase 5	516005	
14,673–14,149[b]	MTND6	NADH dehydrogenase 6	516006	Leber optic atrophy, 535,000

(continued)

Table 5.4 (continued)

Location (nt)	Symbol	Title	MIM	Disorder[a]
14,742–14,674[b]	MTTE	tRNA glutamic acid	590025	
14,747–15,887	MTCYB	cytochrome b	516020	
15,888–15,953	MTTT	tRNA threonine	590090	
16,023–15,955[b]	MTTP	tRNA proline	590075	

[a] In addition to the disorders caused by point mutations in individual genes, deletions involving more than one mitochondrial gene have been identified in Pearson syndrome (557,000), early-onset chronic diarrhea with villus atrophy (520,100), and Kearns-Sayre syndrome (530,000), among others[b] Transcribed from light chain (L) in opposite direction from all the other genes which are transcribed from the heavy chain (H)

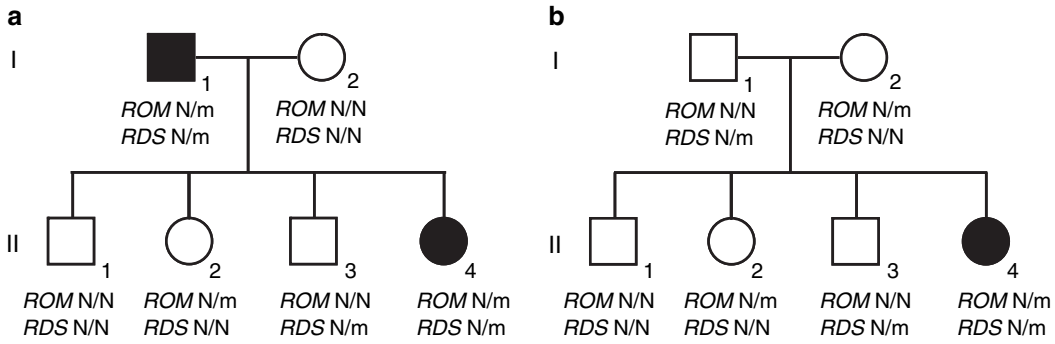

Fig. 5.25 Schematic representation of pedigrees with digenic inheritance; affected individuals are shown in *black*. (**a**) Pedigree with an apparently dominant inheritance. *N* normal allele, *m* mutant allele at the two different genes. (**b**) Pedigree with an apparently recessive mode of inheritance. From [43]

with retinitis pigmentosa. The mode of transmission of this condition resembles an autosomal dominant trait in some pedigrees (vertical transmission, males and females affected) but with 25% affected offspring from an affected parent. In some pedigrees, the transmission resembles a recessive trait when the two parents are heterozygotes for a pathogenic mutation in two different genes. In this case the affected offspring are also 25%.

5.1.11.2 Triallelic Inheritance

A more complicated case is that observed in several forms of Bardet-Biedl syndrome (BBS). The laboratories of N. Katsanis and J.R. Lupski described in 2001 BBS families in which *three* different mutations were necessary to cause the phenotypes of this syndrome [43]: for example, homozygous mutant alleles in the BBS2 gene (OMIM 606151 on chromosome 16q21) *AND* a heterozygous mutation in the BBS6 gene (OMIM 604896 on chromosome 20p12) needed to be present for the affected status (Fig. 5.26). Homozygous mutations only in the BBS2 gene and normal BBS6 gene were not sufficient for the phenotypic manifestations.

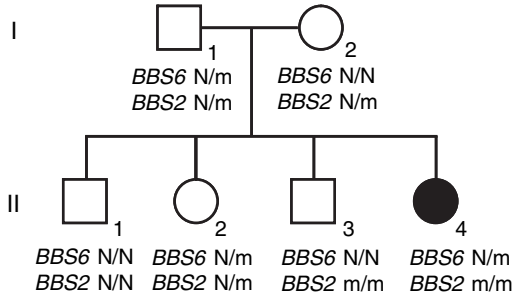

Fig. 5.26 Schematic representation of a pedigree with digenic inheritance; the affected individual II4 is shown in *black*. *N* normal allele, *m* mutant allele Only individuals with three mutant alleles are affected. Note that individual II2 who is heterozygous for mutant alleles in the *BBS6* and *BBS2* genes is not affected; furthermore, individual II3 who is homozygous for mutant alleles in the *BBS2* gene is also not affected. From [43]

Digenic and triallelic inheritance may not be rare in more complex phenotypes, and accumulation of mutations in a few or several genes (also referred to as oligogenic inheritance) may be necessary for complex polygenic disorders.

5.1.12 Multifactorial Inheritance

The majority of human phenotypic features, such as size, weight, intelligence, and others follow multifactorial inheritance. This mode of inheritance is described in Chaps. 4 and 8.

5.2 Hardy–Weinberg Law and Its Applications

5.2.1 Formal Basis

So far the application of Mendel's laws in man has been considered from the standpoint of the single family. What, however, are the consequences for the genetic composition of the population? The field of research that considers this problem is called population genetics. Some basic concepts are introduced here.

These concepts revolve around the so-called Hardy–Weinberg law, discovered by these two authors independently in 1908 [34, 85]. In 1904 Pearson [62] – in the process of reconciling the consequences of Mendel's laws for the population with biometric results – had already derived this law for the special case of equal gene frequencies of two alleles.

The law in its more general form may be formulated as follows: Let the gene frequencies of two alleles in a certain population be p for the allele A and q for the allele B; $(p+q=1)$. Let mating and reproduction be random with respect to this gene locus. The gene frequencies then remain the same, and the genotypes AA, AB, and BB in the F_1 generation occur in the relative frequencies p^2, $2pq$, and q^2, the terms of the binomial expression $(p+q)^2$. In autosomal genes, and in the absence of disturbing influences, this proportion is maintained through all subsequent generations.

5.2.1.1 Derivations from the Hardy–Weinberg Law

We assume that at the beginning the proportions of genotypes AA, AB, and BB in the population of both males and females are D, $2H$, and R, respectively. Symbolically, the distribution of genotypes in both sexes may be written as:

$$D \times \mathrm{AA} + 2H \times \mathrm{AB} + R \times \mathrm{BB} \qquad (5.1)$$

From this the distribution of mating types for random mating is obtained by formal squaring:

$$\begin{aligned}&(D \times \mathrm{AA} + 2H \times \mathrm{AB} + R \times \mathrm{BB})^2\\ &= D^2 \times \mathrm{AA} \times \mathrm{AA} + 4DH\\ &\quad \times \mathrm{AA} \times \mathrm{AB} + 2DR \times \mathrm{AA} \times \mathrm{BB} + 4H^2\\ &\quad \times \mathrm{AB} \times \mathrm{AB} + 4HR \times \mathrm{AB} \times \mathrm{BB} + R^2\\ &\quad \times \mathrm{BB} \times \mathrm{BB}\end{aligned}$$

The distribution of genotypes in the offspring of the different mating types is:

AA × AA	AA
AA × AB	$^1/_2$AA + $^1/_2$AB
AA × BB	AB
AB × AB	$^1/_2$AA + $^1/_2$AB + $^1/_4$BB
AB × BB	$^1/_2$AB + $^1/_2$BB
BB × BB	BB

Inserting these distributions for the mating types into (5.1) yields the distribution of genotypes in the F_1 generation:

$$\begin{aligned}&(D^2 + 2DH + H^2)\mathrm{AA} + (2DH + 2DR\\ &+2H^2 + 2HR)\mathrm{AB} + (H^2 + 2HR + R^2)\mathrm{BB}\\ &= p^2\mathrm{AA} + 2pq\mathrm{AB} + q^2\mathrm{BB}\end{aligned}$$

where $p=D+H$, $q=H+R$ are the frequencies of the alleles A and B, respectively, in the parental generation. Thus, the distribution of genotypes in the offspring generation is uniquely determined by the gene frequencies in the parental population:

$$D' = p^2, 2H' = 2pq, R' = q^2.$$

As:

$$p' = D' + H' = p^2 + pq = p,$$
$$q' = H' + R' = pq + q^2 = q,$$

the gene frequencies in the F_1 generation are equal to those in the parental generation. Thus, the genotype distribution in the next generation (F_2) is also the same as in the F_1 generation, and this holds true for all following generations.

This means that in autosomal inheritance these proportions are expected in the first generation and are maintained in the following generations. For X-linked genes the situation is slightly more complicated. At the same time, the concept of gene frequencies $p+q=1$ was created.

The Hardy–Weinberg law can also be rephrased, indicating that random mating is equivalent to drawing random samples of size 2 from a pool of genes containing the two alleles A and a with relative frequencies p and q. One of the advantages of this law is that frequencies of genetic traits in different populations can be expressed and compared in terms of gene frequencies.

Apart from making it possible to simplify population descriptions, the Hardy–Weinberg law can also help to elucidate modes of inheritance in cases where the straightforward approach through family studies would be too difficult. The classic examples are the AB0 blood types.

5.2.2 Hardy–Weinberg Expectations Establish the Genetic Basis of AB0 Blood Group Alleles

5.2.2.1 Multiple Allelisms

So far, only two different alleles for each locus have been considered. Frequently, however, more than two different states for one gene locus, i.e., more than two alleles, are possible. Examples of such "multiple allelism" in humans and experimental animals abound. Two of the classics are the white series in *Drosophila melanogaster* and the albino series in rabbits.

The formal characteristics can easily be derived:

(a) In any one individual a maximum of only two alleles can be present (unless there are more than two homologous chromosomes, as in trisomics).
(b) Between these alleles, crossing over can be disregarded as they are located at homologous loci. Here the simplest formal model is described, using the AB0 blood groups as an example.

5.2.2.2 Genetics of the AB0 Blood Groups

The AB0 blood groups were discovered by Landsteiner in [46]. Compared to other blood group systems their most important property is the presence of isoantibodies that have led to frequent transfusion accidents. These accidents helped in the discovery of blood groups. The first relevant genetic theory was developed by von Dungern and Hirszfeld in [79]. To explain the four phenotypes A, B, 0, and AB they assumed two independent pairs of alleles (A, 0; B, 0), with dominance of A and B. In 1925 Bernstein [10] tested this hypothesis using the Hardy–Weinberg expectations for the first time. He found their concept to be wrong and replaced it by the correct explanation – three alleles with six genotypes, leading to the four phenotypes due to the dominance of A and B over 0.

The most obvious method to discriminate between these two hypotheses is by family investigation. However, differences between them are to be expected only in matings in which at least one parent carries group AB (Table 5.5). The two-locus hypothesis allows for 0 children while the three-allele hypothesis does not. Although AB is the rarest group, the early literature contained some reports of supposedly 0 children with AB parents; these children were either misclassified or illegitimate. Bernstein, however, was not misled by these observations. His argument goes as follows. It may be assumed that the two-gene pair theory is correct; p may be the gene frequency of A, $1-p=p'$ of a; q the frequency of B, $1-q=q'$ of b. The frequencies to be expected in the population are presented in Table 5.6.

5

Table 5.5 Comparison of the two theories for inheritance of AB0 blood groups (adapted from Wiener [91])

	Children expected from the hypothesis of	
Parents	Two gene pairs	Multiple alleles
0 × 0	0	0
0 × A	0, A	0, A
0 × B	0, B	0, B
A × A	0, A	0, A
A × B	0, A, B, AB	0, A, B, AB
B × B	0, B	0, B
0 × AB	0, A, B, AB	A, B
A × AB	0, A, B, AB	A, B, AB
B × AB	0, A, B, AB	A, B, AB
AB × AB	0, A, B, AB	A, B, AB

Table 5.6 Expectations from multiple allele hypothesis for the AB0 system from Bernstein [10]

Phenotype	Genotype	Frequency	
0	aabb	$(1-p)^2(1-q)^2 = p'^2q'^2$	
B	aaBB	$(1-p)^2q^2$	$= p'^2(1-q'^2)$
	aaBb	$2(1-p)^2q(1-q)$	
A	AAbb	$p^2(1-q)^2$	$= (1-p'^2)q'^2$
	Aabb	$2p(1-p)(1-q)^2$	
AB	AABB	p^2q^2	
	AaBB	$2p(1-p)q^2$	$= (1-p'^2)(1-q'^2)$
	AABb	$2p^2q(1-q)$	
	AaBb	$2p(1-p)2q(1-q)$	

This leads to the following relationships ($\overline{A}$, $\overline{B}$: frequencies of phenotypes):

$$\overline{0} \times \overline{AB} = \overline{A} + \overline{B}$$

and

$$+A + +AB = 1 - p'^2; +B + +AB = 1 - q'^2$$

Thus, it follows:

$$(+A + +AB) \times (+B + +AB) = +AB$$

These identities can be tested. It turned out – and has turned out ever since – that $(\overline{A} + \overline{AB}) \times (\overline{AB} + \overline{AB}) > \overline{AB}$, and $\overline{0} \times \overline{AB} < \overline{A} + \overline{B}$. The differences are so large – and so consistent – that an explanation by chance deviations is inadequate. The first alternative possibility considered by Bernstein was heterogeneity within the examined population. This explanation, however, proved insufficient. On the other hand, it could be shown that the distributions in all populations for which data were available are in perfect agreement with expectations derived from the multiple-allele hypothesis.

To understand Bernstein's argument a fresh look at the Hardy–Weinberg law is necessary. Up to now it has been derived here for the special case of two alleles only. However, it can also be shown to apply for more than two alleles. Assuming n alleles p_1, p_2, …, p_n, the relative frequencies of genotypes are given by the terms of the expansion of $(p_1 + p_2 + \ldots p_n)^2$. It follows for the special case of A, B, and 0 with the frequencies p, q, and r that the distribution of genotypes is:

$$p^2(AA) + 2pq(AB) + 2pr(A0) + q^2(BB) + 2pr(B0) + r^2(B0).$$

Now, we follow Bernstein again (our translation): "for the classes" (phenotypes):

$$\overline{0} = 00 \quad \overline{B} = B0 + BB$$
$$\overline{A} = A0 + AA \quad \overline{AB} = AB$$

the following probabilities can be derived:

$$r^2 \quad 2qr + q^2 \quad 2pr + p^2 \quad 2pq$$

It follows:

$$\overline{0} + \overline{A} = (r+p)^2$$
$$\overline{0} + \overline{B} = (r+q)^2$$

and therefore:

$$q = 1 - \sqrt{\overline{0} + \overline{A}}$$
$$q = 1 - \sqrt{\overline{0} + \overline{B}}$$
$$q = 1 - \sqrt{\overline{0}}$$

and the relation:

$$1 = p + q + r = 1 - \sqrt{\overline{0} + \overline{B}} + 1 - \sqrt{\overline{0} + \overline{A}} + \sqrt{\overline{0}}$$

This can be tested using the AB0 phenotype distributions in various populations of the world. The criterion is that the gene frequencies calculated with this formula must add to 1. In addition, expected genotype

frequencies can be calculated from these gene frequencies and can be compared with observed frequencies. Apart from the correctness of the genetic hypothesis, however, this result requires still another condition. There must be random mating with regard to this characteristic.

In the data analyzed by Bernstein the agreement already was excellent, and this has proven to hold true for the huge amount of data collected ever since. One example may help in understanding the principle of calculation. The following phenotype frequencies were reported from the city of Berlin ($n = 21{,}104$): 43.23% A ($n = 9{,}123$), 14.15% B ($n = 2{,}987$), 36.60% 0 ($n = 7{,}725$), and 6.01% AB ($n = 1{,}269$).

Using Bernstein's formula, the gene frequencies are:

$$p = 1 - \sqrt{(0.3660 + 0.1415)} = 0.2876$$
$$q = 1 - \sqrt{(0.3600 + 0.4323)} = 0.1065$$
$$r = \sqrt{0.3660} = \frac{0.6050}{0.9991}$$

Thus:

$$p + q + r = 0.9991$$

At first glance, this result agrees well with the expectation, i.e., 1. As a statistical test for examining whether the deviation is significant, the χ^2 method can be applied [73]:

$$\chi_1^2 = 2n\left(1 + \frac{r}{pq}\right) D^2$$

$$D = 1 - (p + q + r)$$

In our example, the result is:

$$\chi_1^2 = 0.88$$

This confirms that the values found are in good agreement with the genetic hypothesis and with the assumptions of random mating for the AB0 system.

In a later paper Bernstein showed how the difference D may be utilized to correct the calculated gene frequencies. The uncorrected gene frequencies may be named p', q', and r', and the following formulas may be used:

$$p = p'(1 + D/2)$$
$$q = q'(1 + D/2)$$
$$r = (r' + D/2)(1 + D/2)$$

and for the example:

$$p = 0.2876(1 + 0.00045) = 0.2877$$
$$q = 0.1065(1 + 0.00045) = 0.1065$$
$$r = (0.6050 - 0.00045)(1 + 0.00045) = 0.6057$$

In the process of testing the two genetic hypotheses for the AB0 system Bernstein developed a method for calculating gene frequencies.

5.2.2.3 Meaning of a Hardy–Weinberg Equilibrium

Populations showing agreement of the observed genotype proportions with the expectations of the Hardy–Weinberg Law are said to be "in Hardy–Weinberg equilibrium." This equilibrium must be distinguished from that between alleles, which is discussed in the contexts of selection and of mutation. The Hardy–Weinberg equilibrium is an equilibrium of the distribution of genes in the population ("gene pool") among the various genotypes. Under random mating this equilibrium is reestablished after one generation, possibly with changed gene frequencies if it is disturbed by opposing forces.

It follows from our discussion, however, that the Hardy–Weinberg law can be expected to be valid only when the following prerequisites are not violated:

(a) The matings must be random with respect to the genotype in question. This can safely be assumed for such traits as blood groups or enzyme polymorphisms. It cannot be assumed for visible characteristics such as stature, and still less for behavioral characteristics such as intelligence. This should be kept in mind when measures used in quantitative genetics, (for example, correlations between relatives), are interpreted in genetic terms.

(b) A deviation from random mating is caused by consanguineous matings. If the consanguinity rate in a population is high, an increase in the number of homozygotes must be expected (Chap. 17). It is even possible to estimate the frequency of consanguinity in a population by

means of the deviations from the Hardy–Weinberg proportions.

(c) Recent migrations might disturb the Hardy–Weinberg proportions.

(d) Occasionally selection is mentioned as a factor leading to deviations. This may be true but need not necessarily apply. As a rule, selection tends to cause changes in gene frequencies; selection before reproductive age, for example, in the prenatal period, or during childhood and youth, does not influence the Hardy–Weinberg proportions in the next generation at all. If genotypes are tested among adults in a situation in which a certain genotype had been selected against in children, this genotype is found to decrease in frequency. Even assuming appreciable selection in a suitable age group, ascertainment of statistically significant deviations from Hardy–Weinberg proportions requires large sample sizes – larger than are usually available. Sometimes the absence of significant selection is inferred from the observation that Hardy–Weinberg proportions are preserved in a population. This conclusion, however, unless carefully qualified may easily be wrong. Considering all the theoretical possibilities for disturbance, it is indeed amazing how frequently the Hardy–Weinberg proportions are found to be preserved in the human population.

(e) Formally, a deviation from the Hardy–Weinberg law may be observed if the population is a mixture of subpopulations that do not completely interbreed (random mating only within subpopulations), and consequently the gene frequencies in these subpopulations differ. This was first described by Wahlund in [81], who gave a formula for calculating the coefficient F of the apparent inbreeding from the variance of the gene frequencies between the subpopulations.

(f) Another cause of deviation may be the existence of a hitherto undetected ("silent") allele, a heterozygous carrier of which cannot be distinguished from a homozygous carrier of the usual allele. C.A.B. Smith [69], however, has pointed out that a silent allele causes a significant deviation from the Hardy–Weinberg law only when it occurs at a sufficiently high frequency for the homozygote to be detected.

5.2.3 Gene Frequencies

5.2.3.1 One Gene Pair: Only Two Phenotypes Known

In rare autosomal-recessive diseases only one gene pair is present, and only two phenotypes are usually known when the heterozygotes cannot be identified, or, as is usually the case, when direct data on population frequencies of heterozygotes are not available. This also applies for blood group systems for which only one type of antiserum is available. Here the frequency of homozygotes aa being q^2, the gene frequency is simply. There is no way to test the assumption of random mating.

Table 5.7 [49] is slightly oversimplified; some of the frequencies given vary in different populations. However, the data point out how much more frequent the heterozygotes are, especially for rare conditions. This is important for genetic counseling, and for the much-discussed problem of the number of lethal or detrimental genes for which the average human being might be heterozygous.

5.3 Statistical Methods in Formal Genetics: Analysis of Segregation Ratios

5.3.1 Segregation Ratios as Probabilities

During meiosis – and in the absence of disturbances – germ cells are formed in exactly the relative frequencies expected from Mendel's laws. A diploid spermatocyte heterozygous for alleles A and a produces two haploid sperms with A, and two with a. If all the sperms of a given male come to fertilization, and none of the zygotes die before birth, the segregation ratio among his offspring would be exactly 1:1. There would be no place for any statistics.

Organisms in which such an analysis is indeed possible are yeast and the bread mould *Neurospora crassa,* which has become important in biochemical genetics. In the development of such an organism,

Table 5.7 Differing homozygote and heterozygote frequencies for different gene frequencies (with examples of recessive conditions; adapted from Lenz [49])

Homozygote frequency q^2	Gene frequency q	Heterozygote frequency $2pq$	Approximate homozygote frequencies in European populations
0.64	0.8	0.32	Lp(a-) lipoprotein variant
0.49	0.7	0.42	Acetyl transferase, "slow" variant (Sect. X56)
0.36	0.6	0.48	Blood group 0
0.25	0.5	0.50	Nonsecretor (se/se)
0.16	0.4	0.48	Rh negative (dd)
0.09	0.3	0.42	Lactose restriction (northwestern Germany)
0.04	0.2	0.32	Le(a–b–) negative
0.01	0.1	0.18	β-Thalessemia (Cyprus)
1:2500	1:50	1:25	Pseudocholinesterase (dibucaine-resistant variant), cystic fibrosis; α_1-antitrypsin deficiency
1:4,900	1:70	1:35	Adrenogenital syndrome (Canton Zurich)
1:10,000	1:100	1:50	Phenylketonuria (Switzerland; USA)
1:22,500	1:150	1:75	Albinism; adrenogenital syndrome with loss of NaCl
1:40,000	1:200	1:100	Cystinosis
1:90,000	1:300	1:150	Mucopolysaccharidosis type 1
1:1,000,000	1:1,000	1:500	Afibrinogenemia

there is a phase in which the diploid state has just been reduced to the haploid, and all four meiotic products lie in a regular sequence. They can be removed separately, grown, and examined ("tetrad analysis"). Expected segregation ratios are found with precision.

In higher plants and animals, including humans, only a minute sample of all germ cells comes to fertilization. In the human female about 6.8×10^6 oogonia are formed; the number of spermatogonial stem cells in the male is estimated at about 1.2×10^9; the actual number of sperm is a multiple of this figure. Hence any given germ cell has a very small probability of coming to fertilization. In addition, the sampling process is usually random with respect to a given gene pair A,a. This means that for the distribution of genotypes among germ cells coming to fertilization the rules of probability theory apply, and empirically found segregation ratios may show deviations from their statistical expectations.

Modern humans are fairly accustomed to thinking in statistical terms when solving daily problems. These experiences help us to understand simple applications of probability theory. Everyone, for example, readily recognizes that the following rationale is wrong.

A young mother had always wished to have four children. After the third, however, there was a long pause. The grandmother asked her daughter whether she had now decided differently. Answered the daughter: "Yes, in principle, I would still like four children. But I read in the newspaper that every fourth child born is Chinese. And a Chinese child …there I am reluctant."

In another example, the mistake is less obvious. The parents of two albino children visit a physician for genetic counseling. They wish to know the risk of a third child also being albino. The physician knows that albinism is an autosomal-recessive condition, with an expected segregation ratio of 1:3 among children of heterozygous parents. He also knows that sibships in which all sibs are affected are very rare. Hence, he informs the parents: "As you already have two affected children, the chance that the third child will also be affected is very small. The next child should be healthy." The actual risk, of course, remains 25% (Sect. 5.3.2).

A textbook on human genetics cannot teach probability theory and basic statistics. Therefore, it is assumed that the reader has some knowledge of the basic concepts of probability theory, that he knows the most important distributions (binomial, normal, and Poisson distribution), and has some idea of standard statistical methods. The following presents some applications to problems in human genetics. We are aware of the danger that this section may be used as a "cookbook," without understanding of the basic principles and recommend that the reader become familiar with these principles,

for example, in the opening chapters of Feller's *Probability Theory and Its Applications* [24].

5.3.2 Simple Probability Problems in Human Genetics

5.3.2.1 Independent Sampling and Prediction in Genetic Counseling

The physician who gave the wrong genetic counsel to the couple with two albino children did not take into account that the fertilization events leading to the three children are independent of each other, and that each child has the probability of $^1/_4$ to be affected, regardless of the genotypes of any other children. The probabilities for each child must be multiplied. He was right when he said that illness of all three children is rare in a recessive condition: The probability is $(^1/_4)^3 = {}^1/_{64}$ for all three children to be affected; the family to be counseled however, already had two such children and the probability of this occurring was only $(^1/_4)^2 = {}^1/_{16}$. It takes only one event with the probability $^1/_4$ to complete the three-child family with albinism, $^1/_{16} \times {}^1/_4 = {}^1/_{64}$. It is also intuitively obvious that there is no way for a given zygote to influence the sampling of gametes of the same parents many years later. Chance has no memory!

All possible combinations of affected and unaffected siblings in three-child families can be enumerated as follows (A = affected; U = unaffected):

UUU,AUU,UAU,AAU,\quad UUA,AUA,UAA, AAA

In recessive inheritance, the event U has the probability $^3/_4$. Thus, the first of the eight combinations *(UUU)* has the probability $(^3/_4)^3 = {}^{27}/_{64}$. This means that of all heterozygous couples having three children $^{27}/_{64}$, or fewer than 50% have only healthy children. On the other hand, all three children are affected in $(^1/_4)^3 = {}^1/_{64}$ of all such families. There remain the intermediate groups. Three-child families with one affected child and two healthy ones in that order obviously have the probability $^1/_4 \times {}^3/_4 \times {}^3/_4 = {}^9/_{64}$. However, we are not particularly interested in the sequence of healthy and affected children. Therefore the three cases of such families, *UUA, UAU,* and *AUU*, can be treated as equivalent, giving $3 \times {}^9/_{64} = {}^{27}/_{64}$. The group with two affected can be treated accordingly, giving $3 \times {}^1/_4 \times {}^1/_4 \times {}^3/_4 = {}^9/_{64}$. As a control, let us consider whether the various probabilities add up to 1:

$$\frac{27+27+9+1}{64}$$

This is a special case of the binomial distribution. There are two consequences for Mendelian segregation ratios one theoretical, the other extremely practical. First, it follows that among all families for which a certain segregation ratio must be expected, an appreciable percentage – 27 of 64 in a three-child family with recessive inheritance – cannot be observed because chance has favored them by not producing any affected homozygotes. Hence, the segregation ratio in the remainder is systematically distorted. Special methods have been devised to correct for this "ascertainment bias" (Sect. 5.3.4). Secondly, and this is a most practical conclusion, with limitation of the number of children to two or three, most parents both of whom are heterozygous for a recessive disease will not have more than one affected child. Since the probability of affected children occurring in another branch of the family is very low – and the rate of consanguinity in current populations of industrialized countries has likewise decreased – almost all affected children represent sporadic cases in an otherwise healthy family; there is no distinct sign of recessive inheritance. Any subsequent child, however, again runs the risk of $^1/_4$. The layman usually does not know that the condition is inherited. Therefore, genetic counseling must be actively offered to these families.

5.3.2.2 Differentiation Between Different Modes of Inheritance

In Sect. 5.1.4, an X-linked dominant pedigree is shown (Fig. 5.13) for vitamin D-resistant rickets and hypophosphatemia. What is the probability of such pedigree structure if the gene is in fact located on one autosome? Only the children of affected males are informative because among children of affected women a 1:1 segregation irrespective of sex must be expected. The seven affected fathers have 11 daughters, all of whom are affected. The probability of this outcome with autosomal inheritance is $(^1/_2)^{11}$. The same fathers have 10 sons who are all healthy, giving a probability of $(^1/_2)^{10}$. Hence, the combined probability of 11 affected daughters and 10 healthy sons is:

$$(^1/_2)^{21} = \frac{1}{2\ 097\ 152}$$

This probability is so tiny that the alternative hypothesis of an autosomal-dominant mode of inheritance is convincingly rejected. The only reasonable alternative is the X-linked dominant mode. This hypothesis is corroborated independently by the observation (Sect. 5.1.4) that on average male patients are more severely affected than female.

This is different for a rare skin disease (Brauer keratoma dissipatum). For this condition a Y-chromosomal mode of inheritance has been considered – and indeed all nine sons of affected fathers in a published pedigree show the trait, whereas five daughters in both generations are unaffected. This gives:

$$(^1/_2)^9 \times (^1/_2)^5 = (^1/_2)^{14} = \frac{1}{16\ 384}$$

Hence, the probability of this pedigree having occurred by chance as an autosomal-dominant trait is very low indeed. There is an important difference, however, from the example of vitamin D-resistant rickets. Other pedigrees showing autosomal-dominant inheritance are unknown for this type of rickets, and all observations confirm the location of this gene on the X chromosome. For Brauer keratoma dissipatum, on the other hand, some families have been observed exhibiting very similar phenotypes that show clearcut autosomal-dominant inheritance. It is therefore likely that the described pedigree has been selected from an unknown number of observations because of its peculiar transmission. The calculation is misleading as the "universe" from which this sample of observations was drawn (all pedigrees with the same phenotype) is much larger (and ill-defined), and the sample (the pedigree) is biased. The trait seems to be autosomal-dominant.

Another, more obvious example of an error in the definition of the sample space is the mother, above, who did not want a Chinese baby.

5.3.3 Testing for Segregation Ratios Without Ascertainment Bias: Codominant Inheritance

Apart from these limiting cases, calculation of exact probabilities for certain families or groups of families is usually impracticable. Therefore statistical methods are used that are either based on the parameters of the "normal" distribution, which is a good approximation of the binomial distribution (parametric tests), or derive directly from probabilistic reasoning (nonparametric tests). One method that is especially well suited for genetic comparisons is the χ^2 test. This enables us to compare frequencies of observations in two or more discrete classes with their expectations. The most usual form is:

$$\chi^2 = \Sigma \frac{(E-O)^2}{E}$$

(E = expected number; O = observed number). In Farabee's pedigree with dominant inheritance (Sect. 5.1.2), there are 36 affected and 33 unaffected children of affected parents. With dominant inheritance, E is $^1/_2$ of all children, i.e., 34.5:

$$\chi^2_1 = \frac{(36-34.5)^2}{34.5} = \frac{(33-34.5)^2}{34.5} = 0.13$$

The probability p for an equal or greater deviation from expectation can be taken from a χ^2 table for 1 degree of freedom. The number of degrees of freedom indicates in how many different ways the frequencies in the different classes can be changed without altering the total number of observations. In this case the content of class 2, unaffected, is unequivocally fixed by the content of class 1. Therefore, the number of degrees of freedom is 1. In general the number of degrees of freedom is equal to the number of classes less 1.

A second example is taken from the codominant mode of inheritance (Sect. 5.1.2). Table 5.1 summarizes Wiener's family data for the MN blood types. Are the resultant segregation ratios compatible with the genetic hypothesis? For this problem, matings MM × MM, MM × NN, and NN × NN give no information. Expectations in the matings MM × MN and NN × MN are 1:1, in the mating MN × MN 1:2:1. This leads to Table 5.8 for the χ^2 test: For 4 degrees of freedom we find in the χ^2 table: $p = 0.75$. This is very good agreement with expectation.

5.3.3.1 Dominance

The situation becomes slightly more complicated when one allele is dominant and the other recessive. This is the case, for example, in the AB0 blood group system. Here, the phenotype A consists of the genotypes AA and A0. The expected segregation ratios among their offsprings differ. Some of the heterozygous parents A0 can be

Table 5.8 Comparison between expected and observed segregation figures in the MN data of Wiener et al. (Table 4.1 [92])

Mating type	MM	MN	NN	χ^2	Degrees of freedom
MM × MN	$\frac{(499-486)^2}{486}$	$\frac{(473-486)^2}{486}$	–	0.6955	1
MN × MN	$\frac{(199-200)^2}{200}$	$\frac{(405-400)^2}{400}$	$\frac{(196-200)^2}{200}$	0.1475	2
MN × NN	–	$\frac{(411-396.5)^2}{396.5}$	$\frac{(382-396.5)^2}{396.5}$	1.0605	1

recognized, for example, in matings with 0 partners by the finding of 0 children. Others have only A children just by chance. Special statistical methods are necessary to calculate correct expectations and to compare empirical observations with these expectations [68].

5.3.4 Testing for Segregation Ratios: Rare Traits

5.3.4.1 Principal Biases

If the condition under examination is rare, families are usually not ascertained at random; one starts with a "proband" or "propositus," i.e., a person showing the condition. This leads to an *ascertainment bias,* which must be corrected. The bias can be of different kinds, depending on the way in which the patients have been ascertained.

(a) Family or truncate selection. All individuals suffering from a specific disease in a certain population at a certain time (or within certain time limits) are ascertained. The individual patients are ascertained independently of each other, i.e., the second case in a sibship would always have been found. Such truncate ascertainment is possible, for example, if the condition always leads to medical treatment, and all physicians report every case to a certain registry – as when an institute carries out an epidemiological study. As a rule, case collections approaching completeness are possible only in ad hoc studies of research workers specializing in a condition or group of conditions.
Here, the ascertainment bias is caused exclusively by the fact that only those sibships are ascertained that contain at least one patient. As noted above, however (Sect. 5.3.3), this leaves out all sibships in which no affected individual has occurred just by chance. Their expected number is:

$$\sum_s q^s n_s \qquad (5.2)$$

(s = number of siblings/sibship; p = segregation ratio; $q = 1 - p$; n_s = number of sibships of size s). In recessive disorders, $p = 0.25$. The smaller the average sibship size, the stronger is the deviation from the 3:1 ratio in the ascertained families.

(b) Incomplete multiple (proband) selection; single selection as limiting case. It is rare that all individuals in a population are ascertained; frequently a study starts, for example, with all patients in a hospital population who have a certain condition. Here an additional bias must be considered: the more affected members a sibship has, the higher is its chance to be represented in the sample. This bias causes a systematic excess of affected persons, which is added to the excess caused by truncate selection as explained above.

Koller [45] gave a simple example that demonstrates the nature of this excess. Let us assume that the probands are ascertained during examination of only a single year's group of conscripts. The population comprises a number of families with three children, at least one of whom has the disease, and one of whom is a member of the current year's group. Ascertainment of the family depends on the presence of an affected child in the 1 year group examined. Thus, all families with three affected siblings but only two-thirds of the families with two affected and one-third of those with only one affected are ascertained.
The methods of correction described below are reliable only if the probability for ascertainment

of consecutive siblings is independent of the ascertainment of the first one. In an examination of conscripts, as described above, this may be the case. Most studies, however, begin with a hospital population or some other group of medically treated persons. Here, according to general experience, subsequently affected children are much more frequently brought to a hospital when another child has been treated successfully. The opposite trend, however, is also possible. Becker [3], for example, collected all cases of X-linked recessive Duchenne's muscular dystrophy in a restricted area of southwestern Germany. He had good reason to think that ascertainment was complete for this area. Nevertheless, brothers developing muscular dystrophy as the second or later cases in their sibships were generally not ascertained as probands (i.e., through hospitals and physicians) but through the first proband in the family. In his interviews with the parents Becker found the reason. In the case of the first patient in the sibship the parents usually consulted a physician. Then, however, they discovered that in spite of examinations and therapeutic attempts, the course of the disease could not be influenced. Hence they refrained from presenting a second child to the hospital or the physician.

(c) Apart from these biases, which can be statistically corrected to a certain degree, there are other biases that cannot be corrected. Frequently, for example, a genetic hypothesis is discussed on the basis of families sampled from the literature. Experience shows that such sampling usually leads to reasonable results in autosomal-dominant and X-linked recessive disorders. Autosomal-recessive diseases, however, are more difficult to handle. Families with an impressive accumulation of affected sibs have a higher chance of being reported than those with only one or two affected members. This selection for "interesting" cases was more important early in the twentieth century because families generally had more children. Furthermore, recessive conditions discovered today are usually interesting from a clinical and biochemical point of view as well.

These biases can be avoided only by publishing all cases and by critical interpretation of data from the literature. A statistically sound correction is impossible, as such bias has no simple and reproducible direction.

To summarize, the method of segregation analysis depends on the way in which families are ascertained. It follows that the method of ascertainment should always be described carefully. Above all, the probands should always be fully indicated. It is also of interest whether the author during his case collection has become aware of any ascertainment biases.

These considerations show that complete (truncate) ascertainment of cases in a population, and within defined time limits, is the optimal method of data collection.

5.3.4.2 Methods for Correcting Bias

Two different types of correction are possible: test methods and estimation methods.

In a test method the observed values are compared with the expected values, which have been corrected for ascertainment bias. The first such test method was published by Bernstein in [45]; it corrected for truncate selection. The expected number of affected E_r is:

$$\mathrm{E_r} = \mathrm{sn_s}\,\frac{\mathrm{p}}{1-q^s} \tag{5.3}$$

in all sibships of size n (definition of symbols as in (5.2)). A similar test method can also be used for proband selection.

Test methods answer a specific question: do the observed proportions fit the expected values according to a certain genetic hypothesis?

In many if not in most actual cases, the question is more general: What is the unbiased segregation ratio in the observed sibships? This is an estimation problem. The earliest method was published in 1912 by Weinberg [86] and was called the sib method. Starting from every affected sib in the sibship, the number of affected and unaffected among the sibs is determined. This method is adequate for "truncate selection," i.e., when each affected person is, at the same time, a proband. The sib method is the limiting case of the "proband method" used when the families are ascertained by incomplete multiple proband selection. The number of affected and unaffected siblings is counted, starting from every proband. A limiting case is single selection. Here each sibship has only one proband, and the counting is done once among the sibs.

These estimates converge with increasing sample size to the parameter p, the true segregation ratio; they are *consistent.* It was realized early, however, that they are not fully *efficient,* except for the limiting case of

single selection, i.e., they do not make optimal use of all available information. Therefore improvements have been devised by a number of authors. Today such simple methods are no longer used. Moreover, the problems to be solved by segregation analysis are usually more complex. For example, the families to be analyzed may be a mixture of genetic types with various modes of inheritance; there may be admixture of "sporadic" cases, due either to new mutation or to environmental factors; penetrance may be incomplete, or the simple model of a monogenic mode of inheritance may be inadequate for explaining familial aggregation, and a multifactorial genetic model must be used (for the conceptual basis of such multifactorial models, see Chap. 8). Computer programs are now available for carrying out such analyses; they are available either from their authors' institutions or through an international network of program packages. Some of these also offer programs for comparing predictions from various genetic models.

5.3.5 Discrimination of Genetic Entities: Genetic Heterogeneity

It is a common experience in clinical genetics that similar or identical phenotypes are caused by a variety of genotypes. The splitting of a group of patients with a given disease into smaller but genetically more uniform subgroups has been a major topic of research in medical genetics over recent decades. Frequently such heterogeneity analysis is another aspect of the application of Mendel's paradigm and its consequences: carrying genetic analysis through different levels ever closer to gene action.

It appears at first glance that with modern biological methods discrimination of genetic entities on descriptive grounds, i.e., on the level of the clinical phenotype, would no longer hold interest. In our opinion, however, knowledge of the phenotypic variability of genetic disease in humans is needed for many reasons:

(a) Such knowledge provides heuristic hypotheses for systematic application of the more penetrating methods from biochemistry, molecular biology, immunology, micromorphology, and other fields.
(b) Treatment will often depend upon manipulation of gene disordered biochemistry and pathophysiology of a given disease.
(c) We require insight into the genetic burden of the human population.
(d) Better data are needed for many of our attempts to understand the problems of spontaneous and induced mutation.

5.3.5.1 Genetic Analysis of Muscular Dystrophy as an Example

One group of diseases in which analysis using the clinical phenotype together with the mode of inheritance proved to be successful are the muscular dystrophies. These conditions have in common a tendency to slow muscular degeneration, incapacitating affected patients who often ultimately die from respiratory failure. There are major differences in age at onset, location of the first signs of muscular weakness, progression of clinical symptoms, and mode of inheritance. These criteria were used by medical geneticists to arrive at the following classification of muscular dystrophies:

1. X-linked muscular dystrophies
    a. Severe type (Duchenne) (310200)
    b. Juvenile or benign type (Becker; 310100)
    c. Benign type with early contracture (Cestan-Lejonne and Emery-Dreifuss; 310300)
    d. Hemizygous lethal type (Henson-Muller-de Myer; 309950)
2. Autosomal-dominant dystrophy Facio-scapulo-humeral type (Erb-Landouzy-Déjérine; 158900)
3. Autosomal-recessive muscular dystrophies
    a. Infantile type
    b. Juvenile type
    c. Adult type
    d. Shoulder girdle type

This classification is based on many reports from various populations and, for the rarer variants, on reports of pedigrees. It does not include pedigrees in which affected members showed involvement only of restricted parts of the muscular system, such as distal and ocular types. Congenital myopathies were also excluded. The main criteria for discrimination are obvious from the descriptive terms used in the tabulation; for details, see Becker [6]. At present, various mutations of the X-linked dystrophin gene are known at the molecular level which lead to the Duchenne and Becker types. The gene for Emery-Dreyfus disease has been localized to distal Xq28.

5.3.5.2 Multivariate Statistics

The critical human mind is an excellent discriminator. However, statistical methods for identifying subgroups within a population on the basis of multiple characteristics are now available (multivariate statistics). Such methods can also be applied to the problem of making discrimination of genetic entities more objective.

5.3.6 Conditions Without Simple Modes of Inheritance

The methods discussed so far are used mainly for genetic analysis of conditions thought to follow a simple mode of inheritance. In many diseases, however, especially in some that are both serious and frequent, there are problems:

(a) Diagnosis of the condition may be difficult. There are borderline cases. Expressed more formally: the distribution of affected and unaffected in the population is not an outright alternative (examples: schizophrenia; hypertension; diabetes).
(b) It is known from various investigations, including twin studies, that the condition is not entirely genetic but that certain environmental factors influence manifestation (example: decline of diabetes in European countries during and after World War II).
(c) The condition is so frequent that clustering of affected patients in some families must be expected simply by chance (examples: some types of cancer).
(d) It can be concluded from our knowledge of pathogenic mechanisms that the condition is not a single disease but a complex of symptoms common to a number of different causes (example: epilepsy). In fact, it is becoming apparent that diagnoses such as hypertension and diabetes subsume groups of heterogeneous disease entities.

In no such case can a genetic analysis that starts from the phenotype be expected to lead to simple modes of inheritance. However, for many such conditions, two questions of practical importance arise:

1. What is the risk of relatives of various degrees being affected? Is it higher than the population average?
2. What is the contribution of genetic factors to the disease? Under what conditions does the disease manifest itself?

Familial aggregation can be assessed by calculation of empirical risk figures. Twin studies and comparisons of incidence among relatives of probands with those in the general population are required to answer the questions. Here, we discuss risk figures.

5.3.6.1 Empirical Risk Figures

The expression "empirical risk" is used in contrast to "theoretic risks" as expected by Mendelian rules in conditions with simple modes of inheritance. The early methods were developed largely by the Munich school of psychiatric genetics in the 1920s with the goal of obtaining risk figures for psychiatric diseases.

The basic concept is to examine a sufficiently large sample of affected patients and their relatives. From this material, unbiased risk figures for defined classes of relatives are calculated. These figures are used to predict the risk for relatives in future cases. This approach makes the implicit assumption that risk figures are generally constant "in space and time", i.e., among various populations and under changing conditions within the same population. Considering the environmental changes influencing the occurrence of many diseases such as diabetes, this assumption is not necessarily true but is useful as a first approximation.

The approach can be extended to include the question of whether two conditions A and B have a common genetic component, leading to increased occurrence of patients with disease A among close relatives of patients with disease B.

5.3.6.2 Selecting and Examining Probands and Their Families

In conditions that have simple modes of inheritance, the selection of probands is usually straightforward. The modes of ascertainment are discussed in Sect. 5.3.4. For empirical risk studies the same rules apply. In fairly frequent conditions, complete ascertainment of

cases in a population is rarely if ever feasible and is also unnecessary in these investigations. In most situations, a defined sample of probands, such as all cases coming to a certain hospital for the first time during a predefined time period can be used. The mode of ascertainment is single selection, or very close to it. This approach simplifies correction of the ascertainment bias among sibs of probands. The empirical risk figures can be calculated by counting affected and unaffected among the sibs, excluding the proband. Risk figures among children ascertained through the parental generation are unbiased and need no correction.

Frequently, the diagnostic categories are not clearcut. In these cases, criteria for accepting a person as a proband must be defined unambiguously beforehand, and all possible biases of selection should be considered. Are more severe cases normally admitted to the hospital selected for study? Are patients selected from a particular social or ethnic group? Are there any other biases that might influence the comparability of the results? Genuinely unbiased samples are hardly if ever available, but the biases should be known. Most importantly, such biases should be independent of the problem to be analyzed. For example, it would be a mistake to consider only patients who have similarly affected relatives.

The goal of the examinations is to obtain maximal and precise information about the probands and their families as far as possible. Methods for achieving this goal, however, vary. Clinical experience and the study of publications on similar surveys are helpful.

Once the proband and his family are ascertained, the relatives should be noted as completely as possible, and information on their health status must be collected. Here, personal examination by the investigator and historical information provided by the patients and their relatives are indispensible. Such data should be backed by hospital records and various laboratory and radiological studies. Even results of clinical examinations should be regarded with scepticism since not all physicians are equally knowledgeable and careful, and official documents, such as death certificates, are often unreliable regarding diagnostic criteria.

In most cases, the determination of genetic risk figures answers the question of whether the risk is higher than in the average population. Sometimes adequate incidence and/or prevalence data from a complete population in which the study is carried out or a very similar one are available. More often than not, however, a control series must be examined with the same criteria as used for the "test" populations. If possible, examination on normal controls and their relatives should be performed in a "blind" way; i.e., the examiners should be unaware of whether the persons studied come from the patient or the control series. It is a good idea to use matched controls, i.e., to examine for every patient a control person matched in all criteria but not related to the condition to be investigated (such as age, sex, ethnic origin, etc.).

5.3.6.3 Statistical Evaluation, Age Correction

In conditions that manifest at birth, such as congenital malformations affecting the visible parts of the body, further calculations are straightforward: the empirical risk for children is given by the proportion of affected in the sample. In many cases, however, onset occurs during later life, and the period at risk may be extended. Here the question asked is: What is the risk of a person's becoming affected with the condition, provided he or she lives beyond the manifestation period? The appropriate methods of age correction have been discussed extensively in the earlier literature [45]; one much-used is Weinberg's "shortened method." First, the period of manifestation is defined on the basis of a sufficiently large sample (usually larger than the sample of the study itself). Then all relatives who dropped out of the study before the age of manifestation are discarded. The dropping out may be for any of a variety of reasons: death, loss of contact due to change of residence, or termination of the study. All persons dropping out during the age of manifestation are counted as one-half, and all who have survived the upper limit of manifestation age are counted full.

5.3.6.4 Example

Among children of schizophrenics, 50 were affected and 200 unaffected. Of these, 100 have reached the age of 45 and 100 are between the age of 15 and 45 (i.e., the age of manifestation for the great majority of schizophrenic cases).

Thus, the corrected number of unaffected is: $200 - {}^{1}/_{2} \times 100 = 150$; the empirical risk is:

$$\frac{50}{150 + 50} = 25\%$$

Chapter 23.7 deals in detail with practical problems, taking schizophrenia and affective disorders as examples.

5.3.6.5 Selection of Probands for Genome-Wide Association Studies

Genome-wide association studies have expanded our possibilities to identify new traits in conditions without simple modes of inheritance. The selection of probands or of big cohorts of individuals with a certain phenotypic feature follows different strategies to the aforementioned examples. This will be explained in detail in Chapters 8.1 and 8.

5.3.6.6 Theoretical Risk Figures Derived from Heritability Estimates?

There are suggestions that empirical risk figures should be replaced by theoretical risk figures computed from heritability estimates for the multifactorial model (Chap. 8), after data are found to agree with expectations from such a model. This could be done when the data compared with a simple diallelic model. Such heritability estimates can be achieved by comparing the incidence of the condition in the general population with that in certain categories of relatives, for example, sibs or, with caution, from twin data. In theory the method permits inclusion of environmental, for example, maternal, effects. Its disadvantage, however, is that it depends critically on the assumption that the genetic model fits the actual situation sufficiently well. Since the genetic model chosen may not apply to the data at hand, there is danger that the sophisticated statistical approach suggests a spuriously high degree of precision of the results.

5.4 Conclusions

The transmission of traits determined by single genes, including hereditary diseases, follows Mendel's laws. Autosomal-dominant, autosomal-recessive, and X-linked modes of inheritance can be identified on the basis of the location of mutant genes on autosomes or on the X chromosome, and noting the phenotypic distinction between homozygotes and heterozygotes. Mutations in mitochondrial DNA are transmitted from the mother to all children. Deviations from the classical Mendelian transmission scheme may occur as a consequence of "genomic imprinting," where the parental origin of the mutation determines the phenotype. "Anticipation," with earlier age of onset in succeeding generations, may owe its origin to unstable mutations. Genotype frequencies in populations follow the Hardy–Weinberg Law, which can be used to estimate gene frequencies. In rare traits, such as those in most hereditary diseases, pedigrees are often ascertained via affected individuals and their sibships; when such pedigrees are used to calculate Mendelian segregation ratios, the resulting "ascertainment bias" in favor of affected persons must be corrected by appropriate statistical methods. New sequencing approaches will now enable researches to find disease-causing genes even in relatively small families. Furthermore, genome-wide association studies have paved the way for identifying genomic loci associated with multifactorial inheritance.

References

1. Armour CM, McCready ME, Baig A, Hunter AG, Bulman DE (2002) A novel locus for brachydactyly type A1 on chromosome 5p13.3–p13.2. J Med Genet 39:186–188
2. Badano JL, Katsanis N (2002) Beyond Mendel: an evolving view of human genetic disease transmission. Nat Rev Genet 3:779–789
3. Bates GP (2005) History of genetic disease: the molecular genetics of Huntington disease – a history. Nat Rev Genet 6:766–773
4. Beaudet AL, Perciaccante RG, Cutting GR (1991) Homozygous nonsense mutation causing cystic fibrosis with uniparental disomy. Am J Hum Genet 48:1213
5. Becker PE (1953) Dystrophia musculorum progressiva. Thieme, Stuttgart
6. Becker PE (1972) Neues zur Genetik und Klassifikation der Muskeldystrophien. Hum Genet 17:1–22
7. Becker PE (ed) (1964–1976) Humangenetik. Ein kurzes Handbuch in fünf Bänden. Thieme, Stuttgart
8. Bell J (1934) Huntington's chorea. Treasury of human inheritance 4. Galton Laboratory
9. Bennett T (1975) The T-locus of the mouse. Cell 6:441–454
10. Bernstein F (1925) Zusammenfassende Betrachtungen über die erblichen Blutstrukturen des Menschen. Z Indukt Abstamm Vererbungsl 37:237
11. Bhende YM, Deshpande CK, Bhata HM, Sanger R, Race RR, Morgan WTJ, Watkins WM (1952) A "new" blood-group character related to the ABO system. Lancet 1:903–904

12. Brook JD, McCurrach ME, Harley HG et al (1992) Molecular basis of myotonic dystrophy. Cell 68:799–808
13. Carney G, Seedburgh D, Thompson B, Campbell DM, MacGillivray I, Timlin D (1979) Maternal height and twinning. Ann Hum Genet 43:55–59
14. Castori M, Covaciu C, Paradisi M, Zambruno G (2009) Clinical and genetic heterogeneity in keratosis follicularis spinulosa decalvans. Eur J Med Genet 52:53–58
15. Ceppellini R, Dunn LC, Turri M (1955) An interaction between alleles at the Rh locus in man which weakens the reactivity of the Rho factor (Du). Proc Natl Acad Sci USA 41:283–288
16. Collaco JM, Cutting GR (2008) Update on gene modifiers in cystic fibrosis. Curr Opin Pulm Med 14:559–566
17. Cuénot L (1905) Les races pures et leurs combinaisons chez les souris. Arch Zool Exp Genet 3:123–132
18. De Vries DD, de Wijs IJ, Wolff G et al (1993) X-linked myoclonus epilepsy explained as a maternally inherited mitochondrial disorder. Hum Genet 91:51–54
19. Dorn H (1959) Xeroderma pigmentosum. Acta Genet Med Gemellol (Roma) 8:395–408
20. Engel E (1993) Uniparental disomy revisited: the first twelve years. Am J Med Genet 46:670–674
21. Engel E (1980) A new genetic concept: uniparental disomy and its potential effect, isodisomy. Am J Med Genet 6:137–143
22. Farabee (1905) Inheritance of digital malformations in man. In: Papers of the peabody museum for american archeology and ethnology, vol 3. Harvard University Press, Cambridge/MA, 69
23. Falush D, Almqvist EW, Brinkmann RR, Iwasa Y, Hayden MR (2000) Measurement of mutational flow implies both a high new-mutation rate for Huntington disease and substantial underascertainment of late-onset cases. Am J Hum Genet 68:373–385
24. Feller W (1970/71) An introduction to probability theory and its applications, 2nd edn. Wiley, New York
25. Fleischer B (1918) Über myotonische Dystrophie mit Katarakt. Grafes Arch Ophthalmol 96:91–133
26. Fu Y-H, Pizzuti A, Fenwick RG et al (1992) An unstable triplet repeat in a gene related to myotonic dystrophy. Science 255:1256–1258
27. Gartler SM, Francke U (1975) Half chromatid mutations: Transmission in humans? Am J Hum Genet 27:218–223
28. Gaul LE (1953) Heredity of multiple benign cystic epithelioma. Arch Dermatol Syph 68:517
29. Gerard G, Vitrac D, Le Pendu J, Muller A, Oriol R (1982) H-deficient blood groups (Bombay) of Reunion island. Am J Hum Genet 34:937–947
30. Goldfarb LG, Petersen RB, Tabatan M et al (1992) Fatal familial insomnia and familial Creutzfeldt-Jakob disease: disease phenotype determined by a DNA polymorphism. Science 258:806–808
31. Grønskov K, Ek J, Brondum-Nielsen K (2007) Oculocutaneous albinism. Orphanet J Rare Dis 2:43
32. Hadorn E (1955) Developmental genetics and lethal factors. Wiley, London
33. Haldane JBS (1941) The relative importance of principal and modifying genes in determining some human diseases. J Genet 41:149–157
34. Hardy GH (1908) Mendelian proportions in a mixed population. Science 28:49–50
35. Harley HG, Rundle SA, Reardon W et al (1992) Unstable DNA sequence in myotonic dystrophy. Lancet 339:1125–1128
36. Harris H (1948) A sex-limiting modifying gene in diaphyseal aclasis (multiple exostoses). Ann Eugen 14:165–170
37. Harris H (1980) The principles of human biochemical genetics, 4 th edn. North-Holland, Amsterdam
38. Haws DV, McKusick VA (1963) Farabee's brachydactylous kindred revisited. Johns Hopkins Med J 113:20–30
39. Hilgert N, Smith RJ, Van Camp G (2009) Forty-six genes causing nonsyndromic hearing impairment: which ones should be analyzed in DNA diagnostics? Mutat Res 681:189–196
40. Huoponen K, Lamminen T, Juvonen V et al (1993) The spectrum of mitochondrial DNA mutations in families with Leber hereditary optical neuroretinopathy. Hum Genet 92:379–384
41. HYP Consortium (1995) A gene (PEX) with homologies to endopeptidases is mutated in patients with X-linked hypophosphatemic rickets. Nat Genet 11:130–136
42. Kajiwara K, Berson EL, Dryja TP (1994) Digenic retinitis pigmentosa due to mutations at the unlinked peripherin/RDS and ROM1 loci. Science 264:1604–1608
43. Katsanis N et al (2001) Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. Science 293:2256–2259
44. Kirkpatrick TJ, Au KS, Mastrobattista JM, McCready ME, Bulman DE, Northrup H (2003) Identification of a mutation in the Indian Hedgehog (IHH) gene causing brachydactyly type A1 and evidence for a third locus. J Med Genet 40:42–44
45. Koller S (1940) Methodik der menschlichen Erbforschung. II. Die Erbstatistik in der Familie. In: Just G, Bauer KH, Hanhart E, Lange J (eds) Methodik, Genetik der Gesamtperson. Springer, Berlin, pp 261–284 Handbuch der Erbbiologie des Menschen, vol 2
46. Landsteiner K (1900) Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe. Zentralbl Bakteriol 27:357–362
47. Lenz W (1961) Zur Genetik der Incontinentia pigmenti. Ann Paediatr (Basel) 196:141
48. Lenz W (1975) Half chromatid mutations may explain incontinentia pigmenti in males. Am J Hum Genet 27:690–691
49. Lenz W (1983) Medizinische Genetik, 6 th edn. Thieme, Stuttgart
50. McCready ME, Sweeney E, Fryer AE et al (2002) A novel mutation in the IHH gene causes brachydactyly type A1: a 95-year-old mystery resolved. Hum Genet 111:368–375
51. McCready ME, Grimsey A, Styer T, Nikkel SM, Bulman DE (2005) A century later Farabee has his mutation. Hum Genet 117:285–287
52. McKusick VA (1995) Mendelian inheritance in man, 11 th edn. Johns Hopkins University Press, Baltimore
53. Milch RA (1959) A preliminary note of 47 cases of alcaptonuria occurring in 7 interrelated Dominical families, with an additional comment on two previously reported pedigrees. Acta Genet (Basel) 9:123–126
54. Mohr OL, Wriedt C (1919) A new type of hereditary brachyphalangy in man. Carnegie Inst Publ 295:1–64
55. Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. Trends Genet 21:457–465

56. Mühlmann WE (1930) Ein ungewöhnlicher Stammbaum über Taubstummheit. Arch Rassenbiol 22:181–183
57. Myers RH et al (1993) *De novo* expansion of a (CAG)n repeat in sporadic Huntington's disease. Nature Genet 5:168–173
58. Nicholls RD, Saitoh S, Horsthemke B (1998) Imprinting in Prader-Willi and Angelman syndromes. Trends Genet 14:194–200
59. Ollendorff-Curth (1958) Arch Derm Syph 77:342
60. Ott J (1977) Counting methods (EM algorithm) in human pedigree analysis. Linkage and segregation analysis. Ann Hum Genet 40:443–454
61. Pauli RM (1983) Editorial comment: dominance and homozygosity in man. Am J Med Genet 16:455–458
62. Pearson K (1904) On the generalized theory of alternative inheritance with special references to Mendel's law. Philos Trans R Soc 203:53–86
63. Penrose LS (1947/49) The problem of anticipation in pedigrees of dystrophia myotonica. Ann Eugen 14:125–132
64. Pola V, Svojitka J (1957) Klassische Hämophilie bei Frauen. Folia Haematol (Leipz) 75:43–51
65. Race RR, Sanger R (1975) Blood groups in man, 6 th edn. Blackwell, Oxford
66. Reik W (1989) Genomic imprinting and genetic disorders in man. Trends Genet 5:331–336
67. Siemens HW (1925) Über einen, in der menschlichen Pathologie noch nicht beobachteten Vererbungsmodus: dominant geschlechtsgebundene Vererbung. Arch Rassenbiol 17:47–61
68. Smith CAB (1956/7) Counting methods in genetical statistics. Ann Hum Genet 21:254–276
69. Smith CAB (1970) A note on testing the Hardy-Weinberg law. Ann Hum Genet 33:377
70. Snyder LF, Doan CA (1944) Is the homozygous form of multiple teleangiectasia lethal? J Lab Clin Med 29:1211–1216
71. Spence JE et al (1988) Uniparental disomy as a mechanism for human genetic disease. Am J Hum Genet 42:217–226
72. Stern C (1957) The problem of complete Y-linkage in man. Am J Hum Genet 9:147–165
73. Stevens WL (1950) Statistical analysis of the AB0 blood groups. Hum Biol 22:191–217
74. Stevenson RE, Brasington CK, Skinner C, Simensen RJ, Spence JE, Kesler S, Reiss AL, Schwartz CE (2007) Craniofacioskeletal syndrome: an X-linked dominant disorder with early lethality in males. Am J Med Genet 143A:2321–2329
75. Stocks P, Barrington A (1925) Hereditary disorders of bone development. Treasury of human inheritence 3, part 1
76. Stoffel M, Froguel P, Takeda J et al (1992) Human glucokinase gene: isolation, characterization and identification of two missense mutations linked to early-onset non-insulin dependent (type 2) diabetes melitus. Proc Natl Acad Sci USA 89:7698–7702
77. Timoféef-Ressovsky NW (1931) Gerichtetes Variieren in der phänotypischen Manifestierung einiger Generationen von Drosophila funebris. Naturwissenschaften 19:493–497
78. Trevor-Roper PD (1952) Marriage of two complete albinos with normally pigmented offspring. Br J Ophthalmol 36:107
79. Von Dungern E, Hirszfeld L (1911) Über gruppenspezifische Strukturen des Blutes. III. Z Immunitatsforsch 8:526–562
80. Waardenburg PJ, Franceschetti A, Klein D (1961/1963) Genetics and ophthalmology. Blackwell, Oxford vols 1,2
81. Wahlund S (1928) Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. Hereditas 11:65–105
82. Wallace DC (1989) Report of the committee on human mitochondrial DNA. Cytogenet Cell Genet 51:612–621
83. Wallace DC (1989) Mitochondrial DNA mutations and neuromuscular disease. Trends Genet 5:9–13
84. Wallace DC (1994) Mitochondrial DNA sequence variation in human evolution and disease. Proc Natl Acad Sci USA 91:8739–8746
85. Weinberg W (1908) Über den Nachweis der Vererbung beim Menschen. Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg 64:368–382
86. Weinberg W (1912) Methoden und Fehlerquellen der Untersuchung auf Mendelsche Zahlen beim Menschen. Arch Rassenbiol 9:165–174
87. Welander L (1957) Homozygous appearance of distal myopathy. Acta Genet (Basel) 7:321–325
88. Wendt GG, Drohm D (1972) Die Huntingtonsche Chorea. Thieme, Stuttgart Fortschritte der Allgemeinen und Klinischen Humangenetik, vol 4
89. Wexler NS, Young AB, Tanzi RE et al (1987) Homozygotes for Huntington's disease. Nature 326:194–197
90. Wettke-Schöfer R, Kantner G (1983) X-linked dominant inherited diseases with lethality in hemizygous males. Hum Genet 64:1–23
91. Wiener AS (1943) Additional variants of the Rh type demonstrable with a special human anti-Rh-serum. J Immunol 47:461–465
92. Wiener AS, di Diego N, Sokol S (1953) Studies on the heredity of the human blood groups. I. The MN types. Acta Genet Med Gemellol (Rome) 2:391–398
93. Winters RW, Graham JB, Williams TF, McFalls VC, Burnett CH (1957) A genetic study of familial hypophosphatemia and viatmin D-resistant rickets. Trans Assoc Am Physicians 70:234–242
94. Zevani M, Serudei S, Gellera C et al (1989) An autosomal dominant disorder with multiple deletions of mitochondrial DNA starting at the D-loop region. Nature 339:309–311