

## 2 Genetic Inheritance

---

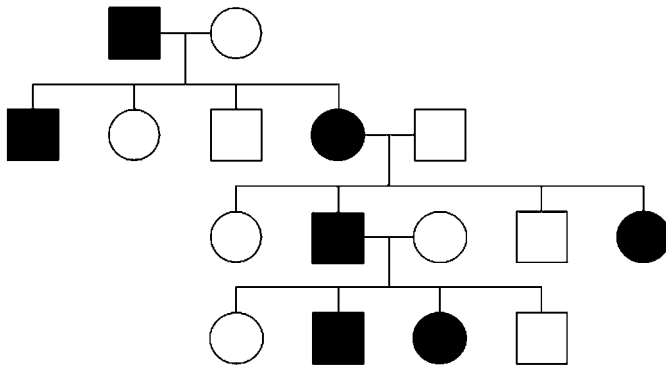
The patterns of genetic inheritance within families and the behavior of genes at the level of a population are of fundamental importance in the estimating the future risk of genetic diseases within families and in the general population. In this chapter, the common patterns of inheritance and the mechanisms of estimating risks to future offspring in affected families are described. The ability to perform simple bayesian calculations is of increasing importance in the molecular pathology laboratory as genetic testing becomes more widespread.

### PATTERNS OF MENDELIAN INHERITANCE IN FAMILIES

The patterns of inheritance and expression of genetic traits in families depends on whether a genetic locus is on an autosome, on a sex chromosome, or in the mitochondrial genome. Any genetic trait, including a disease, may be recessive or dominant depending on the locus and the mutation in question. The inheritance patterns of most diseases can be classified as autosomal or X-linked and as dominant or recessive. Diseases arising from mutations in mitochondrial DNA have a characteristic pattern of inheritance. Knowledge of the principal features of these inheritance patterns is essential for understanding the molecular basis of genetic diseases. Acquired mutations, such as those arising in malignant diseases, can also be described in terms of dominance and recessiveness at a molecular level.

#### *Autosomal Dominant Diseases*

In inherited autosomal dominant diseases, the presence of a single mutant allele on one chromosome is sufficient to cause disease. Although the allele on the homologous chromosome is normal, it does not prevent expression of the mutant phenotype. Autosomal dominant diseases affect both males and females in multiple genera-



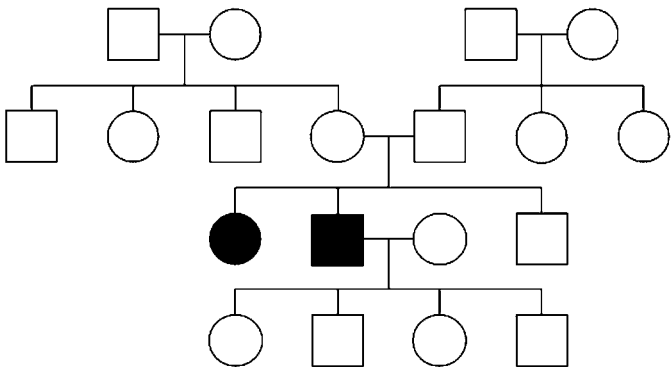
**Fig. 1.** Autosomal dominant pattern of inheritance. Individuals with the trait are shown in black. Fifty percent of the offspring of an affected individual have the trait. Male-to-male transmission is seen.

tions, and each offspring of a person with an autosomal dominant disease has a 50% chance of inheriting the mutation (Fig. 1). Autosomal dominant disorders can be transmitted from father to son, a feature that distinguishes autosomal dominant from X-linked patterns of inheritance.

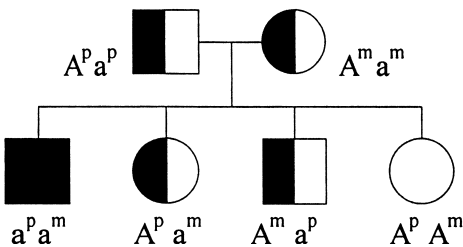
Most dominant disorders show a dosage effect so that individuals who are homozygous for a dominant mutation are more severely affected than are individuals who are heterozygous. For example, familial hypercholesterolemia, which is caused by mutations in the low-density lipoprotein receptor gene, shows a marked difference in the severity of hypercholesterolemia and clinical symptoms between heterozygotes and homozygotes. Diseases that have a dose-dependent severity are sometimes referred to as incompletely dominant or semidominant. A notable exception to the influence of mutation dosage is Huntington's disease, which is caused by mutations in the *IT15* gene (*see* Chap. 5). Homozygotes for this disorder are not more severely affected than are heterozygotes (1).

### ***Autosomal Recessive Diseases***

In autosomal recessive disorders, the alleles on both homologous chromosomes are mutated. There is therefore no normal allele to provide the missing function encoded by the mutated genes. Most inborn errors of metabolism are caused by autosomal recessive traits. Usually, the heterozygous carriers of autosomal recessive

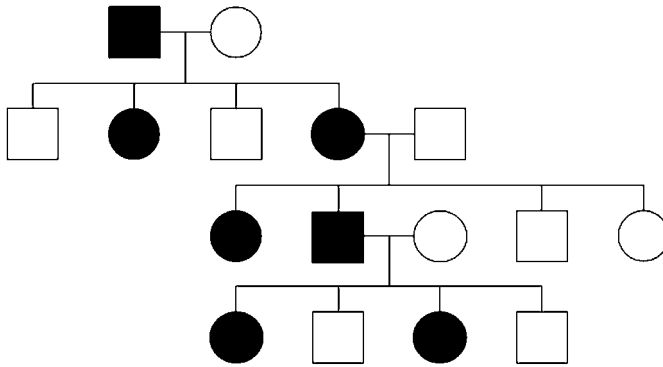


**Fig. 2.** Autosomal recessive pattern of inheritance. Only a single generation shows affected individuals. Both parents of these individuals are carriers. Other family members may be carriers.



**Fig. 3.** Transmission of alleles in an autosomal recessive trait. The  $a$  allele is recessive to the  $A$  allele. The paternal alleles are designated by  $p$  and the maternal alleles by  $m$ . Both parents are carriers, each having a single  $a$  allele. The children show the possible genotypes for offspring of two carrier parents: homozygous affected, two carriers, and a homozygous normal.

traits appear clinically normal, although closer laboratory testing may reveal a biochemical difference from individuals with two functional alleles. Autosomal recessive diseases affect males and females, and often there is no known family history of the disease before the first affected individual comes to medical attention (Fig. 2). Each offspring of the mating of two carriers of an autosomal recessive trait has a 25% risk of being affected, a 50% risk of being a carrier, and a 25% chance of inheriting a normal allele from each parent (Fig. 3). In many autosomal recessive disorders, multiple mutant alleles exist; frequently patients are compound heterozygotes for two distinct mutations.

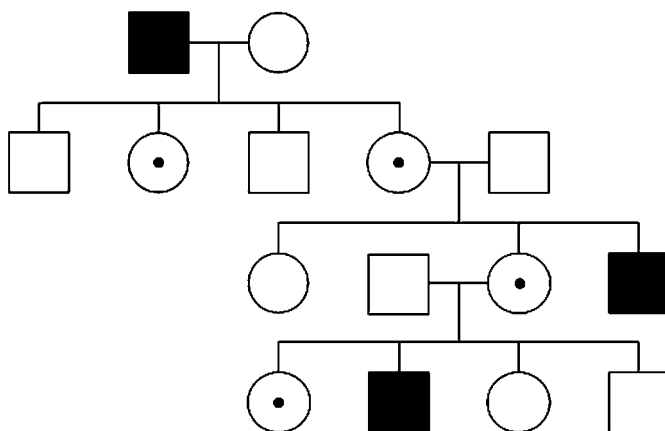


**Fig. 4.** X-linked dominant pattern of inheritance. Individuals with the trait are shown in black. There is no male-to-male transmission. All the daughters of affected males have the trait, whereas one-half of the offspring of affected females have the trait, regardless of sex.

### *X-Linked Dominant Diseases*

X-linked dominant disorders appear when a dominant mutation is present on the X chromosome. In males, the expression of the disease phenotype is inevitable. In females, although the homologous X chromosome is normal, it is insufficient to prevent the expression of the dominant phenotype (Fig. 4). When analyzing pedigrees, X-linked dominant inheritance can be suspected when there is no male-to-male transmission of a dominant phenotype. Affected females transmit the mutation to 50% of their offspring. Affected males transmit the mutation to all of their female offspring, but not to their male offspring.

X-linked dominance is not a frequent mode of inheritance. A form of hypophosphatemia with rickets (OMIM 307800) is inherited as an X-linked dominant disease. Another X-linked dominant disorder is Rett's syndrome, which is typically characterized by several months of normal infant development, followed by severe neurological decline and acquired microcephaly (2). Most cases of Rett's syndrome are caused by mutations in the *MECP2* gene (3). Nearly all reported cases have been in females, apparently because *de novo* mutations originate predominantly in the paternally derived X chromosome (4). *MECP2* mutations have been reported infrequently in males with symptoms ranging from severe neonatal encephalopathy to X-linked mental retardation (5).



**Fig. 5.** X-linked recessive pattern of inheritance. Heterozygous females are indicated by circles enclosing a dot. One-half of the sons of heterozygous females express the recessive trait, and one-half of the daughters of heterozygous females are carriers. All the daughters of affected males are carriers. There is no male-to-male transmission.

### *X-Linked Recessive Diseases*

X-linked recessive diseases are generally expressed only in males carrying a mutant X chromosome. Carrier females, who have one normal X chromosome in addition to the mutant chromosome, are generally unaffected, although in unusual situations they may express the phenotype. The absence of affected females and the lack of male-to-male transmission is characteristic of X-linked recessive traits (Fig. 5). Carrier females transmit the mutation to 50% of their offspring. Males who inherit the mutation invariably express the phenotype. Affected males transmit the mutation to all of their female offspring, who are therefore carriers, but not to their male offspring.

It is possible for females to express X-linked recessive traits through several mechanisms. In every somatic cell of a female, one of the two X chromosomes is physiologically inactivated. This X-chromosome inactivation is known as lyonization, after Dr. Mary Lyon, who proposed this concept in 1961. When female cells are viewed with special strains, the inactive X chromosome appears as a condensed piece of chromatin known as a Barr body. The selection of which of the two X chromosomes is inactivated is random, and the process begins in tissues during embryogenesis. Once a particular X chromosome is inactivated, all daughter cells maintain this

pattern of X-chromosomal inactivation. Inactivation of genes occurring during lyonization is the result of methylation of large stretches of chromatin that inactivates most, but not all, genes on the X chromosome. If lyonization is skewed such that the normal X chromosome is preferentially inactivated, then an X-linked recessive trait carried on the homologous chromosome may be manifest.

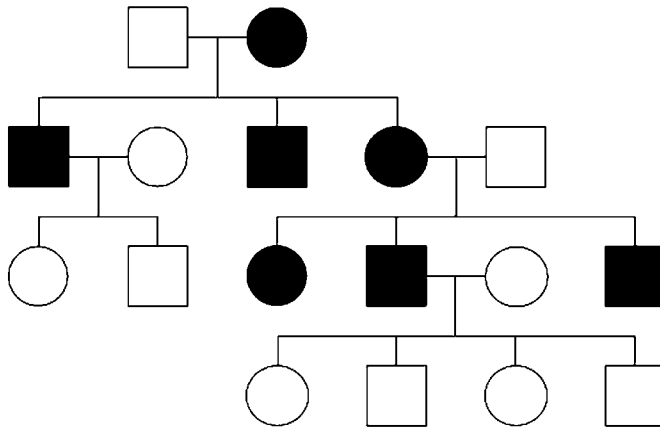
Females can also manifest an X-linked recessive disorder as a result of mating of an affected male with a carrier female. Such a mating will result in 50% of the female offspring being affected because of inheritance of an affected chromosome from each parent. The other 50% will be carriers because of inheritance of an affected chromosome from their father and a normal chromosome from their mother. Finally, females with X-chromosomal abnormalities such as Turner's syndrome (45, X), or microdeletions of X, or some unbalanced translocations, may express an X-linked recessive disease if the remaining X chromosome carries a recessive gene mutation.

### ***Mitochondrial Inheritance***

All mitochondria are derived from the oocyte at the time of fertilization and are therefore of maternal origin. Mitochondrial DNA undergoes mutation and accumulates mutations at a faster rate than does nuclear DNA (6). Mutations in mitochondrial DNA usually manifest as multiorgan disease with frequent neuromuscular abnormalities (*see* Chap. 5). A feature unique to mitochondrial disorders is variable expression of the disease depending on the ratio of normal to mutant mitochondria. This variability in mutation prevalence among mitochondria is termed heteroplasmy. Heteroplasmy depends on the ratio existing at the time of fertilization and on stochastic effects arising from random distribution of normal and mutant mitochondria during mitosis. These give rise to differences in the severity of the phenotype among siblings. Mitochondrial inheritance can be recognized by the exclusive transmission of the phenotype by females (Fig. 6). It is important to note that most proteins expressed in the mitochondria are encoded by nuclear genes. Genetic variation in these follows the usual patterns seen with nuclear genes (*see* Chap. 5).

### ***Dominance and Recessiveness of Mutations in Tumors***

Acquired mutations associated with tumor formation can also be classified as dominant or recessive at a cellular level. For example,



**Fig. 6.** Mitochondrial pattern of inheritance. The trait is transmitted exclusively by females, and to all of their offspring.

the *BCR-ABL* translocation seen in nearly all cases of chronic myeloid leukemia (CML) functions as a dominant mutation because of activation of the Abelson proto-oncogene. Mutation of a single allele resulting in overexpression of *ABL* leads to CML. Activation of many other oncogenes shows a similar genetic effect in that mutation of one allele is sufficient to manifest the phenotype. In familial cancer syndromes associated with germline oncogene activation, the pattern of inheritance is dominant.

In contrast to activation of oncogenes, tumor suppressor gene function is recessive at a cellular level. Both copies of the tumor suppressor gene must be inactivated for the associated malignant phenotype to develop. For example, retinoblastoma can develop when there is loss of function of both copies of the *RB* gene in the same cell of the developing retina. Having just one copy is sufficient to prevent this phenotype. Paradoxically, although loss of function of a tumor suppressor gene such as *RB* is a recessive trait with regard to tumor development at the molecular level, the inheritance pattern of familial cancers associated with loss of tumor suppressor function may be dominant.

Familial retinoblastoma is usually associated with germline deletions of *RB*. A “second-hit” mutation results in inactivation of the remaining *RB* allele on the homologous chromosome in a retinal cell, leading to tumor formation. Because of the high frequency of

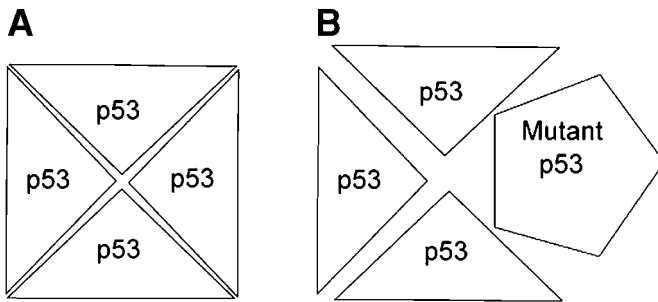
such “second hits,” the disease can affect multiple generations in a typical autosomal dominant pattern. Other familial cancer syndromes such as breast cancer associated with *BRCA* mutations show a similar pattern of dominance within families but recessiveness at a molecular level.

### *Mechanisms of Genetic Dominance*

A mutation that inactivates a gene on the X or Y chromosome, with the exception of genes located in the pseudoautosomal region of these chromosomes (*see* Chap. 1), might be expected to produce a phenotype in males because of the absence of an alternative allele. However, this explanation for dominant effects of gene mutations does not apply to autosomes, which are present in two copies in both males and females. Why are some mutations dominant whereas others are recessive? A review of the mechanisms of genetic dominance has resulted in the following classification of common reasons for this effect (7).

1. *Reduced gene dosage, also known as haploinsufficiency.* In this situation, loss of a single functional copy of a gene gives rise to disease. Examples of this include some forms of thalassemia, in which loss of a  $\alpha$ - or  $\beta$ -globin gene disturbs the normal ratio of  $\alpha$ -to- $\beta$  chains, and DiGeorge's syndrome, caused by loss of *Tbx1*, the dosage of which appears to be critical to normal development of neural crest-derived structures (8,9).
2. *Increased gene dosage.* An example is Charcot-Marie-Tooth disease, which is caused by a duplication of the peripheral myelin protein (*PMP-22*) gene. The reason why the presence of an extra copy of the gene gives rise to disease is not known. It is also of interest in this case that loss of a single copy of the *PMP-22* gene (i.e., haploinsufficiency) gives rise to another neurological disease, hereditary neuropathy with liability to pressure palsies (*HNPP*).
3. *Changes in the expression of mRNA.* An example of this is provided by the t(14;18) translocation, which brings the *BCL2* gene in proximity to the immunoglobulin heavy chain locus. The latter has strong transcription enhancers that are functional in B-cells and lead to increased production of bcl-2 protein, which is an inhibitor of apoptosis (*see* Chap. 7). This confers a selective growth advantage on cells that harbor this translocation.
4. *Constitutive activation of a protein.* This is seen in, for example, type 2 multiple endocrine neoplasia (MEN2), a group of hereditary cancer syndromes that involve constitutive activation of the *RET* proto-oncogene, a





**Fig. 7.** (A) Formation of a p53 tetramer. A single allele, encoding a mutant protein, can disrupt the formation of such a multimer, (B), thereby acting as a dominant mutation.

membrane receptor tyrosine kinase. Because of a mutation, the kinase is constitutively active, resulting in transmission of mitogenic signals.

5. *Dominant negative mutations.* When normal protein function requires assembly of several polypeptides, the presence of a population of mutant polypeptides may disrupt normal assembly of a disproportionate fraction of protein complexes. This deleterious effect of mutation of just one polypeptide on the function of such proteins is known as a dominant negative effect. As an example, certain p53 mutations are commonly found in cancer (*see* Chap. 7). p53 functions as a transcriptional activator requiring assembly of a tetramer in order to bind to DNA and stimulate transcription of genes that p53 activates in response to certain cell injuries such as ionizing radiation. Mutations that abolish the ability of p53 to form tetramers prevent this response (Fig. 7). Because of the requirement for formation of tetramers, the presence of a mutant p53 can disrupt the function of wild-type p53 polypeptides. Such mutations are therefore regarded as dominant negative mutations (10).
6. *Alterations in structural proteins.* This involves a similar effect as dominant negative mutations but is seen in structural proteins. For example, type I collagen, a major protein constituent of bone, is composed of a trimer consisting of two pro $\alpha$ 1(I) polypeptides and a pro $\alpha$ 2(I) procollagen polypeptide. A missense mutation in a pro $\alpha$ 1(I) gene that results in substitution of a glycine (a frequent amino acid in collagen) by another amino acid leads to assembly of structurally abnormal type I collagen proteins, and these comprise the majority (not just 50%) of all assembled proteins.
7. *Toxic effects of protein alterations.* This is exemplified by the unstable trinucleotide repeat disorders known as polyglutamine diseases. In these disorders, an increase in the number of glutamine residues in the mutant proteins is associated with cellular toxicity. For example, in

Huntington's disease, which is caused by an expansion of a polyglutamine-encoding region in *IT15*, cytoplasmic and nuclear deposits of proteolytic fragments containing the polyglutamine tract are believed to be involved in the pathogenesis of cell death (11).

8. *New functions.* In this form of mutation, the mutated protein either has a function not possessed by the wild-type protein, or a new protein is formed. This is commonly seen in translocations associated with malignancies. For example, the t(9;22) translocation associated with the Philadelphia chromosome encodes a chimeric protein, bcr-abl, that has a higher level of tyrosine kinase activity and different cellular distribution from the normal protein produced by the *ABL* proto-oncogene.

## NONMENDELIAN PATTERNS OF INHERITANCE

In addition to the classic patterns of inheritance, a few unusual types of inheritance pattern are recognized that do not conform to mendelian or multifactorial principles. These include diseases associated with unstable trinucleotide repeats, imprinting, and uniparental disomy.

### *Unstable Trinucleotide Repeat Diseases*

Several genetic disorders are characterized by instability of a repetitive sequence consisting of multiple copies of a trinucleotide. The instability is characterized by variation in the number of copies of the trinucleotide between generations. When the number of copies of the trinucleotide repeat expands beyond some critical threshold, disease results. Depending on the particular gene, the trinucleotide repeats may be situated in a coding region or in a noncoding region. If the trinucleotide codes for glutamine, such diseases are referred to as polyglutamine diseases. The unstable trinucleotide repeat diseases are considered in Chapters 4 and 5.

### *Imprinting*

Imprinting refers to the nonexpression of genes in a manner that is dependent on the parent of origin of the chromosome on which the gene is located. Imprinting violates a fundamental principle of classic mendelian genetics, namely, that genes are equally expressed from both members of a pair of homologous chromosomes. Genes on an imprinted chromosome, or region of a chromo-

some, are not expressed depending on whether they were contributed by the mother or by the father. Although the genes on the imprinted chromosome have a normal sequence and are therefore not mutated, they are not capable of being transcribed to form an mRNA and so are functionally absent. The mechanism that leads to imprinting involves methylation of regions of the chromosomes that are imprinted. This is somewhat similar to methylation of one of the two X chromosomes in the somatic cells of normal females, a process referred to as lyonization. However, whereas lyonization involves a random inactivation of an X chromosome, imprinting involves methylation of genes contributed by a specific parent. Depending on the locus, imprinting may involve either the maternal or paternal chromosome.

Several genetic diseases involve imprinted genes. The classic examples of diseases arising from imprinting are Prader-Willi and Angelman syndromes, which involve mutations on chromosome 15q11-q13. Prader-Willi syndrome (PWS; OMIM 176270), which affects approx 1 in 10,000 to 1 in 15,000 newborns, is characterized by hypotonia, short stature, polyphagia, obesity, small hands and feet, hypogonadism, and mild mental retardation. Most cases of PWS are sporadic. In 70% of cases, a cytogenetically visible deletion of 15q11-q13 is present in the paternal chromosome in a region that includes the *SNRPN* gene, which is a candidate gene for this disease. Whether deletion of this gene alone is responsible for the disorder is presently uncertain. The maternal chromosome 15q is imprinted in this region, and therefore subjects with a deletion of this region of the paternal chromosome have no functional copy of the genes that are deleted.

Angelman syndrome (AS; OMIM 105830) is characterized by mental retardation, ataxia, seizures, absence of speech, skin hypopigmentation, sleep disturbance, and spontaneous laughter. The disease has a prevalence of approx 1 in 20,000. Reminiscent of PWS, approx 70% of AS patients have a cytogenetically visible deletion of 15q11-q13 that involves loss of approx 4 Mb of DNA. However, in AS, the deletion is on the maternal chromosome and cannot be compensated for because of imprinting on the homologous region of the paternal chromosome. The specific gene that, when mutated, gives rise to AS is known to be *UBE3A*. This gene encodes a protein, E6-PA, that is involved in transport of other proteins to the pro-

teasome complex in the cytoplasm where protein degradation takes place. The extent of imprinting of *UBE3A* varies between tissues, but imprinting of the paternal allele in the brain is responsible for manifestations of clinical disease. Up to approximately 10% of patients with AS have a mutation in *UBE3A* that inactivates the gene (12). In families with more than one affected individual, mutations in *UBE3A* have been identified in up to 80% of patients. Most of the mutations are truncating. Approximately 7–9% of patients have an imprinting abnormality characterized by a paternal pattern of imprinting on both chromosomes. Some of these patients have defects in a putative imprinting center (IC) that may be responsible for coordinating the switch between maternal and paternal imprinting patterns.

### *Uniparental Disomy*

Both parents contribute equally to the genome of their children, with the exception of the Y chromosome, which is transmitted by fathers to sons, and the mitochondrial genome, which is transmitted exclusively by females. Rarely, individuals have both copies of a chromosome from one parent without the homologous chromosome from the other parent. This phenomenon is termed uniparental disomy (UPD). If UPD includes each member of the pair of chromosomes in a parent, it is termed uniparental heterodisomy. If an individual has two copies of one chromosome from a parent, it is termed uniparental isodisomy.

Several mechanisms might give rise to this phenomenon. At fertilization, an oocyte containing two copies of a chromosome (a result of nondisjunction during meiosis) might fuse with a normal sperm cell containing one copy of the chromosome. The zygote from such a fertilization would be trisomic for the chromosome in question. Re-establishment of a normal chromosome complement would require loss of one the three copies of the chromosome, an event sometimes referred to as “trisomy rescue.” Assuming that any of the three chromosomes is equally likely to be lost, in a third of cases the loss will lead to UPD, and because most nondisjunction occurs in female meiosis I, UPD arising by this mechanism will usually be maternal heterodisomy (13). Alternatively, an oocyte may be missing a chromosome, usually as a result of maternal nondisjunction during meiosis I. Fertiliza-

tion by a normal sperm would yield a zygote that is monosomic for a chromosome. Rescue of this zygote by duplication of the paternal chromosome would lead to paternal uniparental isodisomy. Although these appear to be the most common etiologies of UPD, they are not the only mechanisms. UPD may also arise from mitotic nondisjunction in a conceptus. This can give rise to pure uniparental isodisomy.

The most likely clinical outcome with UPD is that it is asymptomatic. Because uniparental isodisomy leads to homozygosity of genes on that chromosome, it may result in expression of a recessive trait if the involved chromosome carries a recessive mutation. The first patient to be identified with UPD was a child with cystic fibrosis. Her mother was a carrier for the common  $\Delta F508$  mutation in *CFTR* and the child was homozygous for this mutation. However, her father did not carry this mutation. Analysis revealed UPD for the maternal chromosome 7 (14).

UPD can also give rise to disease if the chromosome contains imprinted genes because the presence of two copies of an imprinted locus is functionally equivalent to the absence of both copies of the locus. Approximately 25% of patients with Prader-Willi syndrome have maternal UPD in the region of 15q11-q13, and 3–5% of patients with Angelman's syndrome have paternal UPD in this region. Recently, UPD for chromosome 6 has been shown to be associated with transient neonatal diabetes (15).

### ***Recommended Guidelines for Testing for UPD***

The American College of Medical Genetics has produced guidelines for laboratory diagnostic testing for UPD (13). These guidelines call for demonstration of UPD by use of polymorphic DNA markers such as microsatellites. At least two fully informative loci should demonstrate absence of inheritance of an allele from one parent for one chromosome, but normal biparental inheritance of markers on other chromosomes. Recommended indications for testing for UPD include prenatal identification of mosaicism for trisomy of chromosomes 6, 7, 11, 14, or 15 or robertsonian translocations involving chromosomes 14 or 15. These chromosomal abnormalities are commonly associated with UPD. Testing is also recommended for patients with features of disorders that are known to be associated with UPD, as shown in Table 1.

Table 1  
Diseases Associated with Uniparental Disomy (UPD)<sup>a</sup>

<i>Disease</i>	<i>% with UPD</i>	<i>Chromosome</i>	<i>Parent</i>	<i>OMIM no.</i>
Transient neonatal diabetes mellitus	20	6	Mother	601410
Russell-Silver syndrome	10	7	Mother	180860
Beckwith-Wiedemann syndrome	20	11	Father	130650
upd(14)mat		14	Mother	
upd(14)pat		14	Father	
Prader-Willi syndrome	25	15	Maternal	176270
Angelman's syndrome	3–5	15	Paternal	105830

<sup>a</sup> Features of upd(14)mat and upd(14)pat include short stature, developmental delay, dysmorphic features, and skeletal and joint abnormalities. These disorders are defined by identification of UPD.  
Data from ref. 13.

MULTIFACTORIAL INHERITANCE

Most common diseases are not the result of single gene defects. Cancer, heart disease, and diabetes are generally the result of both hereditary and environmental factors. There are examples of single gene disorders that lead to forms of each of these diseases, for example, hereditary retinoblastoma, familial hypercholesterolemia, and maturity onset diabetes of youth (MODY), but these are exceptional forms of disease. Unfortunately, our current state of knowledge of the specific loci and alleles that contribute to most multifactorial disorders is very limited. We have a much better understanding of the molecular pathology of rare, single-gene disorders than we do of the common illnesses that affect large numbers of the population.

Multifactorial disorders tend to aggregate in families, but the risk of first-degree relatives (parents, children, siblings) of an affected patient developing the disease is less than the usual 50 or 25% found in mendelian genetic disorders. As the degree of relationship from an affected patient becomes more remote (e.g., uncles, cousins, second cousins), the prevalence of multifactorial disorders decreases. This has been interpreted as indicating a threshold effect, i.e., the overall genetic risk is influenced by the additive effects of many genes, some of which increase the disease

risk, and some of which may lessen the risk. When some critical threshold of genetic and environmental risks is crossed, the disease becomes manifest. There are several predicted consequences of this model (16):

1. As the degree of relatedness to an affected family member decreases, the frequency of disease will decline more rapidly than would occur if the disease were caused by a single major locus.
2. The risk of other family members' having a multifactorial disease depends on the degree of severity of the disease in the affected members of the family. If an affected family member has a severe form of the disease, the frequency of other affected individuals will be greater.
3. The risk to other family members also depends on the number of affected individuals. The larger the number of affected individuals, the greater the likelihood that a future sibling will also have the disease.
4. If a disease is normally found with greater frequency in a particular sex, relatives of patients of the less commonly affected sex will have a higher disease frequency than will relatives of patients of the more commonly affected sex. This is another manifestation of the second principle listed above. For a person of the less commonly affected sex to have the disease, the combination of genetic effects is probably more adverse than is required to manifest disease in a person of the sex in which the disease is naturally more prevalent.

Environmental factors are also of importance in multifactorial disease. For example, cigarette smoking, obesity, a high-fat diet, and physical inactivity are associated with increased risk of developing coronary artery disease. From a practical perspective, identification of these is important because environmental risks, unlike genetic risks, are often modifiable.

Traits that are either present or absent are known as qualitative traits. Traits that fall on a continuous scale such as blood pressure, plasma cholesterol, and height are known as quantitative traits, and loci that influence these are known as quantitative trait loci (QTLs). Because a multifactorial trait has both genetic and environmental influences, it is possible by genetic analysis of families to quantify the relative contribution of each to the variance of the trait or phenotype in the population (16). The proportion of variance of a phenotype that results from genetic factors is known as the heritability ( $h^2$ ). Heritability can vary from 0 if genes have no effect on the variance of the trait to 1 if genetic factors are exclusively responsible for variance. As examples,  $h^2$

Table 2  
Heritability ( $h^2$ ) of Lipids<sup>a</sup>

<i>Lipid</i>	<i>h<sup>2</sup> value</i>
Lipoprotein(a)	0.9
LDL and HDL cholesterol	0.5
Triglycerides	0.3

<sup>a</sup> For details, see ref. 23. LDL, low-density lipoprotein; HDL, high-density lipoprotein.

values for different types of plasma lipids are shown in Table 2. The fraction of variance that is not heritable is caused by environmental effects. This is sometimes called environmentability, defined as  $1 - h^2$ .

GENES IN THE POPULATION

Genetic diseases within families tend to follow the patterns of inheritance outlined above. The relationship between genes and their behavior in a population is governed by other factors, and the Hardy-Weinberg law provides a basis for understanding this behavior.

*Hardy-Weinberg Law*

At a given locus, if two possible alleles (A and a) can be found and the frequencies of these in a particular population are  $p$  and  $q$ , then it follows that  $p + q = 1$  because no other alleles exist in the population. The possible genotypes that any individual might have are A/A, A/a, a/A, and a/a. The expected frequencies of these genotypes are  $p^2$ ,  $pq$ ,  $qp$ , and  $q^2$ . Because no other genotypes exist,

$$p^2 + 2pq + q^2 = 1 \tag{1}$$

This simple relationship was independently described by Hardy and Weinberg in 1908. A population in which this formula holds true is said to be in Hardy-Weinberg equilibrium. In such a population, it can be shown that the ratio of the three genotypes will remain constant from one generation to the next. Although it is known as the Hardy-Weinberg law, it is not a law in the same sense that the term is used in physics, e.g., the second law of thermodynamics. Instead, it describes the most likely distribution of genotypes within a popu-



lation. However, for the Hardy-Weinberg law to apply, the population must meet certain criteria:

1. The population is sufficiently large that genetic drift does not occur. To understand genetic drift, consider a small population composed of only 10 mating pairs. If one member of this population of 20 people has a variant allele at some autosomal locus, the frequency of that allele is  $1/40$  in this population. If, by chance, that allele is not transmitted to any offspring, the allele will disappear from the population within a generation. On the other hand, if the person carrying the variant allele happens to transmit the allele to all of his or her offspring, the frequency of the allele may be greatly increased in the next generation. These dramatic changes in the frequency of an allele in this population are caused by the small size of the population. If the population consisted of 2 million people and 100,000 had the variant allele, the likelihood of such dramatic intergenerational changes in gene frequency in this large population would be much less.
2. There is no migration into or out of the population. Different populations may have different frequencies of alleles at a given locus and so migration and mating between populations will alter the frequencies of alleles within a population. Migration is not simply a geographical flow of people across national borders but includes admixture between ethnic groups that occupy a shared land region. For example, the Duffy blood system has three common alleles,  $FY^*A$ ,  $FY^*B$ , and  $FY^*O$ . The  $FY^*O$  allele does not encode an antigen because of a promoter mutation that abolishes gene transcription. Nearly all sub-Saharan African populations are  $Fy(a-b-)$ , i.e., homozygous for  $FY^*O$ , and this phenotype has been shown to confer resistance to infection with *Plasmodium vivax* (17). In the United States, the frequency of non- $Fy(a-b-)$  alleles among African Americans is approx 0.2, indicating that  $FY^+$  alleles have been introduced into this population through mating with Caucasians in the last several hundred years (18).
3. There is random mating of individuals without regard to their genotypes. For example, the expected likelihood of two carriers mating according to Hardy-Weinberg equilibrium is  $2pq \times 2pq$ . If the actual frequency of such matings were increased, then predictions of the frequency of genotypes in future generations based on the Hardy-Weinberg principle would be wrong. Preferential mating between individuals with similar genetic traits is known as assortative mating and commonly involves traits such as height and intelligence. Assortative mating of carriers of genetic diseases leads to a higher frequency of affected homozygotes among offspring. Inbreeding or consanguinity has the same tendency to increase the proportion of affected homozygotes in offspring of such matings. In general, the rarer a recessive disease in the population, the more frequently are parents of affected children related to each other.

4. No genotype has a significant reproductive advantage or disadvantage. This depends on the fitness associated with an allele. In genetic terminology, fitness ( $f$ ) is a measure of an individual's capacity to reproduce. A subject with  $f = 1$  has a similar number of offspring to other members of the general population. Often, individuals with genetic disorders have reduced fitness. For example, patients with type II osteogenesis imperfecta have very high childhood mortality, and so  $f = 0$ . This is an autosomal dominant disorder, and therefore most cases represent new mutations. Germline mosaicism in a parent accounts for a recurrence risk of 7% in future pregnancies (19). The coefficient of selection ( $s$ ) is defined as the proportional reduction in the gametic contribution of a particular genotype to the next generation (16).

$$s = 1 - f \quad (2)$$

In the case of autosomal recessive disorders, the heterozygotes may have a selective advantage that maintains a mutant allele at a high frequency. A good example of this is seen in sickle cell anemia. Homozygotes with the sickle cell mutation have a reduced fitness, but heterozygotes have a relative survival advantage in regions where malaria is found because of resistance to the parasite.

5. Mutation does not lead to a significant change in the frequency of alleles within the population.

A consideration of these criteria shows they are commonly not met in human populations. Despite this limitation, Hardy-Weinberg is important because it provides a means to solve clinically important calculations on gene frequencies in populations. It also allows for a comparison of experimental data of genotypes with what would be expected in a population that was in Hardy-Weinberg equilibrium, and this comparison is useful for understanding and examining deviations from an expected result.

### ***Examples of Calculations Involving Hardy-Weinberg Principles***

Hardy-Weinberg principles are frequently used in genetic calculations for clinical purposes. Examples of these calculations include the following.

*Example 1.* An autosomal recessive disease is observed to affect 1 in 10,000 people. What is the frequency of mutant alleles? What is the carrier frequency?

In autosomal recessive diseases, affected individuals carry two disease alleles. Using the designations  $p$  and  $q$  for the frequencies of the wild-type

and mutant alleles, respectively, the frequency of diseased patients,  $q^2$ , is 1/10,000 and therefore  $q$ , the frequency of the mutant allele, is 1/100.

Because  $p + q = 1$ ,  $p = (1 - q)$ , which is 99/100 or  $\cong 1$ . The carrier frequency is  $2pq$ , which is  $\cong 1/50$ .

Knowing the carrier frequency allows us to predict the frequency of affected individuals in a population, as in the following example.

*Example 2.* If the carrier frequency for an autosomal recessive disorder is 1/20, what is the frequency of affected individuals?

In a population with random mating, we expect that the frequency of mating of carriers will be  $2pq \times 2pq$  or, in this case,  $1/20 \times 1/20 = 1/400$ . One-fourth of the offspring of these matings will be affected, giving us a predicted 1/1600 affected individuals in the population.

Example 2 demonstrates the importance of random mating in a population for Hardy-Weinberg principles to apply. If, for example, carriers were much less likely (or much more likely) to mate with other carriers, then the predicted frequency of affected individuals would deviate from the above estimate.

Knowing the distribution of genotypes in a population allows us to determine whether the population is in Hardy-Weinberg equilibrium for the locus in question. For example, in a survey of the frequency of the  $\beta$ -globin sickle (S) allele and its wild-type counterpart (A) in 12,387 people in west Africa, the following numbers of genotypes were observed:

A/A 9365

A/S 2993

S/S 29

From these data, the frequency of the A and S alleles can be determined as follows:

In this population of 12,387 people, there are 24,774 alleles. Each A/A homozygote has two A alleles, and each A/S heterozygote has one A allele. The total number of A alleles is therefore 21,723. The total number of S alleles is 3051. The frequency of the A allele in this population is 21,723/24,774 or 0.877. The frequency of the S allele is 0.123.

Knowing the frequency of each allele in the population, we can determine the *expected* frequency of each genotype by using the Hardy-Weinberg equation. If  $p = 0.877$  and  $q = 0.123$ , then we *expect* the following frequencies of genotypes:

$$A/A (p^2) = 0.877^2 = 0.769$$

$$A/S (2pq) = 2 \times 0.877 \times 0.123 = 0.215$$

$$S/S (q^2) = 0.123^2 = 0.015$$

Multiplying each genotype frequency by the number of people in the population gives the *expected* numbers of people with each genotype:

$$A/A = 0.769 \times 12,387 = 9526$$

$$A/S = 0.215 \times 12,387 = 2663$$

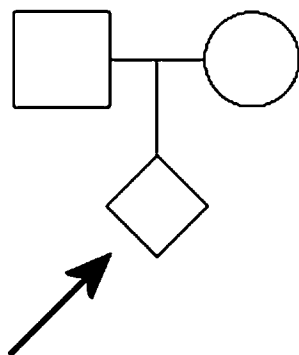
$$S/S = 0.015 \times 12,387 = 186$$

Comparison of the *observed* genotype data with the *expected* numbers of each genotype shows that there is a marked decrease in the number of observed S/S homozygotes and an increase in the number of observed heterozygotes relative to what one would expect for a population in Hardy-Weinberg equilibrium. The statistical significance of these differences can be determined by a chi square test, which, in this case, indicates that the observed data differ very significantly from the expected results, confirming that the observed population is not in Hardy-Weinberg equilibrium with respect to alleles at the  $\beta$ -globin locus. In this case, the high mortality associated with the S/S genotype and the relative survival advantage of A/S heterozygotes in a malarial region explain these findings.

### ***Bayesian Calculations of Risk***

In genetics, the risk of having a disease can be assessed in various ways. For diseases that are characterized by a simple mendelian type of inheritance, the risk depends on the pattern of inheritance of the disease, as outlined previously. However, laboratory data are increasingly available that modify the risk of an individual's having a disease. Combining the risk based on the inheritance pattern with the modified risk based on laboratory data is achieved by performing a bayesian analysis.

*Example 3.* Consider the couple shown in Fig. 8. They want to know their risk of having a child affected with cystic fibrosis (CF), an autosomal recessive disease that is caused by mutations in the cystic fibrosis transmembrane conductance (*CFTR*) gene on chromosome 7. Both members of the couple are Caucasian, and this population has a CF carrier frequency of 1/29. Neither has a family history of cystic fibrosis. What is the probability of their fetus's having CF?



**Fig. 8.** The parents seek counseling regarding genetic risk of an autosomal recessive disorder to the fetus. See text for discussion.

The probability of either parent being a carrier is  $1/29$ . The probability that both are carriers is  $1/29 \times 1/29 = 1/841$ . For an autosomal recessive disease such as CF, the probability of any offspring of a mating of two carriers being affected is  $1/4$ . The combined probability of both parents being carriers and their child being affected is therefore  $1/841 \times 1/4 = 1/3364$ , which is the incidence of CF among Caucasian newborns.

Following recommendations from the American College of Obstetrics and Gynecology, the mother decides to undergo genetic testing for a panel of 25 mutations that includes 80% of all known *CFTR* mutations in Caucasians. Her test indicates that she does not have any of the 25 mutations. What is the revised probability that their child will have CF?

This is a typical problem that can be solved by a simple bayesian calculation. In this case we have two independent pieces of information to consider:

1. the *a priori* probability of being a carrier and
2. the probability of being a carrier without having a mutation identified in this testing procedure.

The term *a priori* means information that was available before any testing was performed. In this case it is the known frequency of carriers in the population, and the information that the patient has no family history of CF. We also know that the mechanism of inheritance is autosomal recessive and that CF carriers are asymptomatic.

The probability of being a carrier, yet not having a mutation identified by the panel of 25 mutations, is based on experimental observation. From studies of *CFTR* mutations in affected patients, we know what percentage of mutations will be detected by the panel of 25 mutations.

This percentage varies among ethnic populations because of variations in the frequencies of mutant alleles.

A bayesian calculation is performed as shown below. In step 1, we list the two mutually exclusive possibilities for the mother’s genetic status, i.e., carrier and noncarrier. In step 2, we list the *a priori* probability of her being a carrier (1/29). The probability of her not being a carrier is 1 – (1/29). In step 3, we list the conditional probabilities of her being a carrier and having a negative test result. Because we know that the test detects 80% of carriers, we know that it fails to identify 20% of carriers. We assume that the probability of a noncarrier having a negative result for *CFTR* mutations is 1. In step 4, we multiply the *a priori* and conditional probabilities for each genetic possibility (i.e., carrier or noncarrier). This value is known as the joint probability. Finally, in step 5, we determine the probability that the mother is a carrier. This probability is expressed as the ratio of the joint probability of being a carrier to the sum of the joint probabilities of being a carrier and of not being a carrier. This value is known as the posterior probability. In the illustration below, the posterior probability of the mother’s being a carrier is 1/141. The probability of her not being a carrier could be determined by either calculating the posterior probability of this, or by subtracting the probability of her being a carrier from 1, i.e., 1 – (1/141).

Step 1. List mutually exclusive possibilities.	Carrier	Not a carrier
Step 2. List a priori probability of each	1/29	28/29
Step 3. Conditional probability (i.e., of having a negative mutation test)	20/100	1
Step 4. Multiply line 2 and line 3		
Joint probability	(1/29)(20/100) = 20/2900	(28/29)(1) = 2800/2900
Step 5. Calculate posterior probability	(20/2900)/[(20/2900) + (2800/2900)] =20/2820 =1/141	

Because the mutation detection test did not show any mutations, the likelihood of the mother’s being a carrier is reduced from 1/29 to 1/141. The probability that both parents are carriers is therefore 1/29 × 1/141 = 1 in 4089, and the chance of their child’s having CF is approx 1 in 16,300, compared with 1 in 3364 before testing. If the father were tested and also found not to carry one of the panel of mutations, the risk of this couple having a child with CF would be 1/141 × 1/141 × 1/4 = 1/79,524.

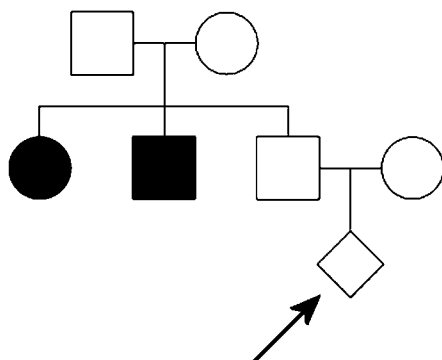
*Example 4.* Consider another couple who are Ashkenazi Jewish and who do not have a family history of CF. The frequency of CF carriers in this population is also approx 1/29, but because of a different distribution of mutations in this population, genetic testing can detect 97% of all mutations. What is the likelihood that an Ashkenazi Jewish mother who has no family history of CF and who is negative on *CFTR* mutation testing would have a child with CF?

	Carrier	Not a carrier
<i>A priori</i> probability	1/29	28/29
Probability of having a negative mutation test	3/100	1
Joint probability	(1/29)(3/100) = 3/2900	(28/29)(1) = 28/29
Final probability	(3/2900)/[(3/2900) + (2800/2900)] =3/2803 =1/934	

In this case, the probability that the mother is a carrier is 1/934. The probability of the couple having a child with CF is  $1/29 \times 1/934 \times 1/4 = 1$  in 108,344. Because mutation testing offers a higher detection rate in this population than in the northern European Caucasian population, this couple can be assured that their chance of having a child with CF is extremely small.

*Example 5.* What if the Caucasian mother described above is found to have a CF mutation, such as the common  $\Delta F508$ ? In this case, the national recommendations call for testing of the father. If his test result is negative, what is the probability of the couple having a child with CF? The probability that a Caucasian with a negative mutation test on the recommended panel is a CF carrier is 1/141 (see above). Knowing his wife is a carrier means that the risk of having a CF-affected child is  $1/141 \times 1 \times 1/4 = 1/564$ .

Now let’s consider scenarios that involve a positive family history. In the family shown in Fig. 9, the father of a fetus has two siblings with cystic fibrosis. What is this man’s risk of being a carrier? For autosomal recessive diseases, the risk to a child of carrier parents is 1/4 of being affected, 1/4 of being a carrier from the mother, 1/4 of being a carrier from the father, and 1/4 of being genetically normal at the locus in question (Fig. 3). Therefore the chance of being a carrier is 2/4 or 1/2. However, in this case we know that the man does not have CF. He can have one of only three possible genotypes: he might be a carrier by having inherited a mutation from his mother, or be a



**Fig. 9.** A father, who has two siblings with cystic fibrosis, seeks genetic counseling regarding the risk of the fetus's having this disease.

carrier by having inherited a mutation from his father, or he might be genetically normal. Of these three genotypes, two involve his being a carrier; therefore his risk of being a carrier is  $2/3$ . His chance of being genetically normal at the locus in question is  $1/3$ .

The risk of the fetus's having CF is therefore  $2/3 \times 1/29$  (the mother's risk of being a carrier)  $\times 1/4 = 1/174$ . If the mutations in the affected siblings of the man were known, he could be tested for these mutations, and his genetic status for CF could be determined.

### *Empirical Risks*

A common task in genetic counseling is advising parents of a child with a genetic disorder of the recurrence risk for future children. When genetic diseases have a known inheritance pattern, it is usually possible to calculate the risk of recurrence, and if the results of genetic testing are available, the calculation can be refined by bayesian analysis. For some disorders that do not follow simple mendelian genetic patterns of inheritance, it may be possible to give an estimate of the risk of recurrence from empirical observation of other families with the disorder. For example, cleft lip with or without cleft palate (CL/P) is a common developmental anomaly that may be seen in both syndromic forms (i.e., with other developmental anomalies) and in isolated cases. Isolated CL/P has a poorly understood etiology, but familial clustering of cases indicates that heritable factors are contributory. Observation of recurrence rates in large numbers of families with isolated CL/P indicates that the



recurrence rate in families with an affected child or parent is approx 4%, whereas if both a parent and a child have CL/P, the recurrence risk is 17% (20). Such estimates depend on the accuracy of the diagnosis of nonsyndromic CL/P in affected family members.

### ***New Mutations, Germline Mosaicism, and Recurrence Risks***

In families in which a child with a genetic disease has a new mutation, the risk of recurrence in subsequent children is usually extremely low. Occasionally, however, more than one child in a family has a mutation that cannot be demonstrated in either parent. This is highly suggestive of germline mosaicism in a parent. In this situation, the germinal cells of a parent include a population of cells that harbor the mutation and that give rise to gametes with the mutation. This phenomenon might arise from a mutation that occurred during embryogenesis in the parent leading to the parent's being a mosaic with both somatic and germline mosaicism. Depending on the relative abundance of the mutation in tissues, particularly in white blood cells, which are the most common source of DNA for genetic analysis, the mutation may not be detectable by routine methods. Alternatively, a population of cells containing the mutation may be confined to the gonad (pure germline mosaicism) and may give rise to gametes with the mutation. When discussing recurrence risks for future pregnancies, the possibility of parental germline mosaicism should be considered, especially in the case of diseases such as osteogenesis imperfecta type II, tuberous sclerosis, Duchenne muscular dystrophy, and hemophilia A and B, in which germline mosaicism has been frequently documented (21,22). The presence of germline mosaicism for a disease-causing mutation greatly increases the likelihood of recurrence of the disease in future children. For diseases with reduced penetrance or variable expression, both parents should be thoroughly examined for subtle features of the disease in question. Identification of a mildly affected parent greatly increases the risk of recurrence.

## **REFERENCES**

1. Wexler NS, Young AB, Tanzi RE, et al. Homozygotes for Huntington's disease. *Nature* 1987;326:194-7.
2. Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* 1983;14:471-9.

3. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999;23:185–8.
4. Trappe R, Laccone F, Cobilanschi J, et al. MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. *Am J Hum Genet* 2001;68:1093–1.
5. Schanen C. Rethinking the fate of males with mutations in the gene that causes Rett syndrome. *Brain Dev* 2001;23 Suppl 1:S144–6.
6. Wallace DC. Mitochondrial DNA sequence variation in human evolution and disease. *Proc Natl Acad Sci USA* 1994;91:8739–46.
7. Wilkie AO. The molecular basis of genetic dominance. *J Med Genet* 1994;31:89–98.
8. Merscher S, Funke B, Epstein JA, et al. TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* 2001;104:619–29.
9. Lindsay EA, Vitelli F, Su H, et al. Tbx1 haploinsufficiency in the DiGeorge syndrome region causes aortic arch defects in mice. *Nature* 2001;410:97–101.
10. de Vries A, Flores ER, Miranda B, et al. Targeted point mutations of p53 lead to dominant-negative inhibition of wild-type p53 function. *Proc Natl Acad Sci USA* 2002;99:2948–53.
11. Davies S, Ramsden DB. Huntington's disease. *Mol Pathol* 2001;54:409–13.
12. Fang P, Lev-Lehman E, Tsai TF, et al. The spectrum of mutations in UBE3A causing Angelman syndrome. *Hum Mol Genet* 1999;8:129–35.
13. Shaffer LG, Agan N, Goldberg JD, Ledbetter DH, Longshore JW, Cassidy SB. American College of Medical Genetics statement of diagnostic testing for uniparental disomy. *Genet Med* 2001;3:206–11.
14. Spence JE, Perciaccante RG, Greig GM, et al. Uniparental disomy as a mechanism for human genetic disease. *Am J Hum Genet* 1988;42:217–26.
15. Hermann R, Laine AP, Johansson C, et al. Transient but not permanent neonatal diabetes mellitus is associated with paternal uniparental isodisomy of chromosome 6. *Pediatrics* 2000;105:49–52.
16. Emery AEH. *Methodology in Medical Genetics. An Introduction to Statistical Methods*. Churchill Livingstone, Edinburgh, 1986.
17. Miller LH, Mason SJ, Clyde DF, McGinniss MH. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood- group genotype, FyFy. *N Engl J Med* 1976;295:302–4.
18. Lautenberger JA, Stephens JC, O'Brien SJ, Smith MW. Significant admixture linkage disequilibrium across 30 cM around the FY locus in African Americans. *Am J Hum Genet* 2000;66:969–78.
19. Cole WG, Dalgleish R. Perinatal lethal osteogenesis imperfecta. *J Med Genet* 1995;32:284–9.
20. Curtis EJ, Fraser FC, Warburton D. Congenital cleft lip and palate. *Am J Dis Child* 1961;102:853–7.
21. Zlotogora J. Germ line mosaicism. *Hum Genet* 1998;102:381–6.
22. Leuer M, Oldenburg J, Lavergne JM, et al. Somatic mosaicism in hemophilia A: a fairly common event. *Am J Hum Genet* 2001;69:75–87.
23. Ozturk IC, Killeen AA. An overview of genetic factors influencing plasma lipid levels and coronary artery disease risk. *Arch Pathol Lab Med* 1999;123:1219–22.



<http://www.springer.com/978-1-58829-085-4>

Principles of Molecular Pathology

Killeen, A.

2004, IX, 333 p., Hardcover

ISBN: 978-1-58829-085-4

A product of Humana Press