Section nine

Chapter fifty-four

Basic Principles of Genetics

Amalia Martinez-Mir# and Angela M. Christiano#§

#Department of Dermatology and §Department of Genetics and Development, Columbia University, New York, New York, U.S.A.

Key features

- Mendelian diseases *vs* complex traits.
- Mendelian patterns of inheritance.
- Exceptions to the basic Mendelian patterns of inheritance.
- Chromosomal disorders.
- Disease gene identification.

Basic Principles of Genetics

"Everything is genetic (except trauma)"

Francis Collins, meeting of the American Academy of Dermatology, 1995

The completion of the human genome sequence, with an estimated size of 3.2 gigabases (Gb) and around 30000-35000 genes[1,2], presents a powerful tool in medicine, facilitating and speeding the identification of disease genes.

The rapid advance in the development of molecular biology techniques, together with the continuing flow of information and tools derived from the human genome project, puts the physician in a privileged position, having direct access to the genetic basis of disease. It is important, thus, for the physicians in general and dermatologists in particular, to understand the basic concepts of genetics, as well as to be familiar with the new techniques to correctly offer and interpret the results to their patients. The wealth of new information also emphasizes the necessity of combined efforts among the clinical and research arms of medicine.

Access to patients with genetic disorders and accurate clinical description is essential for a research study aimed at the identification of a disease-causing gene. The benefits that disease gene identification can contribute to medicine are numerous. Its chief importance in genetic counseling and prenatal diagnosis has already been shown. But it can also contribute to the identification of carriers, identification of at-risk individuals by population screenings, presympomatic diagnosis, prediction of possible complications and course of the disease, response to pharmacological therapy and finally, to the eventual development of a treatment.

In the last two decades, the molecular bases for numerous genetic disorders have been identified, such as epidermolysis bullosa (EB; chapter 34), ectodermal dysplasia (ED), Netherton syndrome or atrichia with papular lesions (APL). Although they do not account for the majority of patients in a dermatologist practice, they have helped to set the basis for the study of more common disorders, such as atopic dermatitis. These more common disorders can be major public health issues, but have a considerably more complex etiology. The gene-identification strategies developed for more simple diseases, and the identified genes themselves, have played a major role in the study of the underlying causes of common complex disorders. The clinician has had and will have a key role in the feasibility of such studies, completely dependent on a large and well-characterized collection of patient samples.

1. Basic Concepts in Genetics

In this section we provide some definitions to cover the basic concepts used in human genetics.

The nuclear DNA is packed into ***chromosomes*** and it is selectively folded and unfolded to allow the expression of hundreds of genes at different times and in different cell types. Somatic and germline cells contain around 2 meters of DNA[3], which gives an idea of the packaging level of the DNA. We all inherit one set of chromosomes from our father and one from our mother. The paternal and maternal gametes are ***haploid*** cells, with only one chromosome complement or ***n*** (n=23), whereas the resulting embryo is a ***diploid*** cell, with two chromosome complements or ***2n*** (2n=46). Each pair of chromosomes is composed of two ***homologous chromosomes***, one from each parent. A chromosome complement consists of 23

chromosomes, with 22 pairs of ***autosomes*** and one pair of ***sex chromosomes*** (XX in females and XY in males). Chromosomes X and Y share only two small regions known as the ***pseudoautosomal regions***. The ***karyotype*** describes the chromosome constitution of an individual.

The ***centromere*** of a chromosome is the site of a primary constriction that is essential for correct cell division. It is composed of a series of proteins called histones and of repetitive DNA, which can reach several thousand kilobases (Kb). The centromere divides the chromosome into two ***arms***, the short arm or ***p***, and the long arm or ***q***. According to the position of the centromere, the chromosomes can be classified as ***metacentric***, with the centromere near the middle, resulting in two arms more or less equivalent in length; ***submetacentric***, with the centromere between the center and the end, with two arms of different lengths; and ***acrocentric***, with the centromere near the end and the two arms with considerably different lengths. The ***telomere*** constitutes the end of each chromosome and it is formed by thousands of repetitions of the DNA sequence TTAGGG, as well as proteins. They play a key role in the maintenance of chromosome integrity, as well as ensuring a correct chromosome replication.

Meiosis is the process of cell division by which a diploid cell from the germline gives rise to haploid gametes. It results in a new combination of chromosomes in the daughter cells, as compared to that in the progenitor cells. ***Mitosis***, on the other hand, is the process of cell division of proliferating cells by which the genetic material of a cell is duplicated and equally distributed among the two resulting daughter cells. To achieve this duplication of the genetic material, each chromosome duplicates itself resulting in two ***sister chromatids***. Each of the sister chromatids will become the chromosome of the daughter cell.

The location of a particular sequence of DNA on a chromosome, a gene or any fragment of DNA, is designated as a ***locus*** (plural, loci). Since we have two copies of each chromosome, we also have two copies of each locus. These two copies can be identical in the two homologous chromosomes, or they can display some differences. These alternative forms of a particular DNA sequence or gene are called ***alleles***. An individual will have two alleles at any given autosomal locus, one from each parent. These two alleles will be identical in a ***homozygous*** individual or locus, and different in a ***heterozygous*** individual or locus (**Fig. 1**). Since males have only one copy of the X and Y chromosomes, they will be ***hemizygous*** for loci on these two chromosomes. Although an individual will have only two alleles for any given autosomal locus, being homozygous or heterozygous, when the population at large is considered, a particular locus may exhibit more than two alleles. By convention, when a locus exhibits at least two alleles with the frequency of one of them in the general population being greater than 1%, it is known as a ***polymorphism***. They can be located inside or outside a gene. It has been estimated that single nucleotide polymorphisms or SNPs, which are only one class of polymorphic sequences, occur on average every 1000-2000 nucleotides[4]. The total length of the human genome, 3.2 Gb gives an idea of the great possible genetic variability, emphasizing the singularity of each individual's genetic information.

Fig. 1 Example of a polymorphic locus.

The alleles present at a specific locus (or in the whole genome) constitute the ***genotype*** of an individual and the manifestation of a particular genotype and the environment in which it is expressed is the ***phenotype***. Different genotypes are not always recognizable as distinct

phenotypes. The genetic variation that results in the different alleles at a given locus arises by mutational events at the DNA level, which may or may not result in differences in the phenotype. If there is a change in the phenotype, this can be caused by variation (alleles) present in the general population (normal variation) or by pathogenetic variation leading to a disease. Although any change in the DNA is a ***mutation***, regardless of the resulting phenotype, if any, the terms ***polymorphism*** and ***mutation*** tend to be used in human genetics to refer to normal variation present in the general population versus variation resulting in a disease phenotype, respectively. It is important to note, though, that this use of the terms mutation and polymorphism is not strictly correct.

2. Genetic diseases

Although genetic diseases may manifest themselves in several members of a family, this is not the case for all of them. Likewise, it is not true that any disease manifested in more than one family member is genetic. Pedigrees in which a genetic disease is transmitted, i.e. ***segregates***, can be present in either one single affected member or up to several generations with many affected members in each generation.

These genetic diseases may be caused by genetic defects in one gene (***Mendelian*** or ***monogenic*** diseases) or more than one gene (***polygenic*** diseases), but they can also be the result of the interaction of environmental and genetic factors (***multifactorial*** diseases). In other cases, an aberrant structure or number of one or more chromosomes or a differential parental chromosomal contribution can be the cause of the disease.

In the following sections, we provide an overview of the different types of genetic diseases.

3. Mendelian Diseases: Patterns of Inheritance

A key initial step in the assessment of the risk for each individual from a family with a particular inherited disease is the establishment of the mode of inheritance of the condition. For this purpose it is essential to collect very accurate information on the clinical phenotype and to examine as many family members as possible to define their status, affected or unaffected. The family is usually ascertained through one family member, the ***proband***, from whom the ***pedigree*** is built with the information on the phenotype and the relationship among every individual.

Even when a single gene defect is enough to cause a disease phenotype (monogenic disorders), its expression can be highly variable (see sections 4.1 and 4.2 for incomplete penetrance and variable expression). Genes are expressed on a particular genetic background, which is different in every individual, and these genes and their products interact with other genes and proteins. Thus, one can expect that even monogenic disorders may show some extent of variability. If a gene can be subject to different genetic backgrounds in different individuals and if this genetic background can modify its expression, then slightly or even considerably different clinical entities can be due to mutations in the same gene, known as ***allelic heterogeneity***. This term also refers to those diseases caused by different mutations in the same gene. On the other hand, one could expect that defects in different components of the same metabolic pathway, for example, could result in the same phenotype, known as

genetic or ***locus heterogeneity***. Thus, a diagnosis based on pure clinical symptoms can lead to misclassification of a patient, which underscores the relevance of the identification of disease genes.

There are several examples of both allelic and locus heterogeneity in dermatology. Mutations in the *GJB3* gene coding for connexin 31, for example, can result in autosomal dominant erythrokeratodermia variabilis (EKV) and autosomal dominant or recessive deafness. At the other end, epidermolysis bullosa simplex (EBS) of the Koebner type can result from mutations on the genes for keratin 14 or keratin 5. **Figures 2** and **3** show examples of allelic and locus heterogeneity in skin disorders. Some diseases, such as EBS, display both allelic and locus heterogeneity.

Fig. 2 Examples of allelic heterogeneity in genodermatoses

Fig. 3 Examples of genetic or locus heterogeneity in genodermatoses

Once a phenotype is defined, including its range of variability, the disease phenotype can be followed along the pedigree. The transmission of the trait in the pedigree will define of the ***pattern of inheritance*** of the disease. The genes responsible for monogenic traits can be located either on autosomes, defining an ***autosomal*** phenotype, or on the X chromosome in the case of an ***X-linked*** phenotype. Irrespective of the location of the gene, when a mutation in ***one*** allele is enough for the phenotype to arise, the trait is ***dominant***. On the contrary, when ***both*** alleles of a gene must be mutated in order to alter the phenotype, the condition is called ***recessive***. Thus, a phenotype can be inherited as an ***autosomal dominant*** or ***recessive*** trait or

as an ***X-linked dominant*** or ***recessive*** trait. There are also Y-linked genes, normally affecting male fertility[25-27], although their relevance in more general diseases is discussed.

It is important to note that although this chapter is focused on genodermatoses, all the statements for Mendelian or complex traits apply to both normal and disease phenotypes.

3.1. Autosomal Dominant Inheritance

In a pedigree with an autosomal dominant disease, each affected person has an affected parent, both sexes are equally affected, and both sexes can transmit the trait; and there is transmission from generation to generation (**Fig. 4*A***).

A single mutant allele is enough to give rise to the phenotype. Thus, the offspring of an individual with an autosomal dominant disorder have a risk of 50% (or 1 in 2 children) of inheriting the mutant allele.

An autosomal dominant trait can be transmitted from male to male and it cannot be transmitted by an unaffected individual (see section 4.1 for incomplete penetrance). Examples of autosomal dominant skin diseases are EBS, Darier-White and Hailey-Hailey diseases, monhilethrix or autosomal dominant ichthyosis vulgaris (chapter 57).

A patient with an autosomal dominant disease is usually heterozygous for the disease-causing mutation. Although not very often, homozygotes for autosomal dominant diseases have been reported, which usually exhibit a more severe phenotype compared to heterozygotes. An example of this latter is the case of a family with mutations in the *KRT14* gene, which in heterozygosis gives rise to autosomal dominant EBS of the Weber-Cockayne type and in homozygosis results in the more severe EBS of the Koebner type[28].

There are, however, exceptions to the rules stated above, mainly explained by incomplete penetrance (section 4.1), variable expression (section 4.2), or *de novo* mutations (section 4.4).

Fig. 4 Examples of Mendelian patterns of inheritance.

3.2. Autosomal Recessive Inheritance

In a pedigree with an autosomal recessive disease, an affected individual has clinically unaffected parents; both sexes are equally affected; and there is no transmission from generation to generation (**Fig. 4*B***).

Both alleles of a given gene have to be mutated for the disease to develop. The recurrence risk for the offspring is 25% (or 1 in 4 children). APL (a rare form of total alopecia) and Naxos disease (woolly hair, palmoplantar keratoderma and heart disease) are examples of autosomal recessive genodermatoses.

There is an increased frequency of consanguinity (sharing a common ancestor) among families with recessive conditions. Consanguineous matings increase the probability that a mutated allele present in the common ancestor will become homozygous in the offspring of the related parents (homozygosity by descent, section 4.6). When more than one affected family member for an autosomal recessive disease is present in a single pedigree, they are usually part of the same generation. But the presence of consanguinity can also explain some pedigrees in which an autosomal recessive trait is observed in more than one generation (section 4.6).

A patient with an autosomal recessive disease has inherited two copies of the disease-causing mutation, while both parents are heterozygotes for the mutant gene, i.e. they are ***carriers*** of one mutant and one normal copy. Both parents can carry the ***same mutation***, thus the patient will be homozygous for the disease-causing mutation, or they can carry two different mutations affecting the same gene, and the patient will be ***compound heterozygous***.

Exceptions to these statements are seen in inbred populations, in which matings between unaffected carriers and affected patients give rise to affected family members in two consecutive generations (pseudodominant inheritance, section 4.7).

3.3. X-linked Recessive Inheritance

An X-linked recessive disease typically affects males, who will usually have unaffected parents; there is no male-to-male transmission; and the trait does not appear in successive generations (**Fig. 4*C***).

The recurrence risk for the offspring will depend on the parent transmitting the disease allele. An affected father will transmit the mutated allele, but not the trait, to all his daughters, but to none of his sons. All the daughters of an affected male will be ***obligate carriers***, since they will all inherit the mutated allele from the father. An obligate carrier female will transmit the mutated allele to 50% of her children. Thus, sons of a carrier mother will have a chance of being affected of 50%, while none of the daughters will be affected.

The two main characteristics of an X-linked recessive trait are the absence of male-to-male transmission and the female carriers being asymptomatic. One main exception is the

offspring of an affected male and a carrier female. Other exceptions where carrier females might be affected involve non-random X-inactivation (section 4.3).

Examples of recessive X-linked diseases in dermatology are X-linked ichthyosis (or steroid sulfatase deficiency; chapter 57) and anhidrotic ectodermal dysplasia (EDA).

3.4. X-linked Dominant Inheritance

In a pedigree with an X-linked dominant disease, affected individuals have at least one affected parent; it affects both males and females, although affected females are more common than affected males, but have a milder phenotype; there is no male-to-male transmission; and the trait appears in successive generations (**Fig. 4*D***).

The recurrence risk for the offspring will depend again on the transmitting parent. An affected father will transmit the trait to all his daughters, but to none of his sons. An affected female, on the other hand, will transmit the trait to 50% of her children, regardless the sex of the offspring.

Although both sexes can be affected, the trait is usually milder in females than in males due to non-random X-inactivation (section 4.3). Examples of an X-linked dominant disease for which females present with a less severe phenotype are focal dermal hypoplasia (DHOF or Goltz syndrome)[34] and incontinentia pigmenti [35]. Females affected with DHOF present with cutaneous, skeletal, dental, ocular, and soft-tissue defects, whereas for males, lethality in uterus is suspected.

3.5. Y-linked Inheritance

In a pedigree with a Y-linked disease (holandric inheritance), affected males always have affected fathers and only males are affected. All sons of affected males are affected (**Fig. 4*E***).

There have been only a few reports on Y-linked diseases, and most of them affect male fertility[25-27]. A few examples affecting other organs have been described[33], although their Y-linked inheritance cannot be proven

4. Exceptions to Basic Mendelian Inheritance Patterns

By definition, Mendelian diseases are the result of mutations in one single gene. But, as mentioned above, this gene can modify and be modified by other genes and by the environment. Thus, although the presence of the genotype at one single locus will still be responsible for the disease phenotype, the final phenotype can be subject to variation in its expression. The following sections (4.1 to 4.10) describe single-gene defects that by means of variability in the expression of the mutant allele or by the location of the mutant gene constitute apparent or real deviations to the basic Mendelian patterns of inheritance outlined above.

4.1. Incomplete or Reduced Penetrance and Late Onset Diseases

As defined above, a carrier of a dominant mutation or a homozygote (or compound heterozygous) for recessive mutations will manifest the disease phenotype. This is true for many genetic traits, but there are some exceptions to this rule. In terms of genetic counseling, these exceptions can present a dilemma.

Incomplete or ***reduced penetrance*** refers to those situations where not all the individuals who carry the disease genotype manifest the condition. In these cases, the value of the penetrance will indicate the probability of an individual with the disease genotype to develop the disease phenotype. It is important to note that penetrance refers to an "all or nothing" event, in which some individuals who have inherited the disease genotype do not express the disease phenotype at all. An example of the result of incomplete penetrance would be an autosomal dominant pedigree with an affected individual, unaffected parents, but an affected grandparent. As a consequence, the trait appears to "skip" a generation (**Fig. 5*A***).

A particular case of incomplete penetrance is the so-called ***late onset*** diseases, characterized by ***age-dependent penetrance***. Some diseases exhibit complete penetrance, but only starting from a certain age. In the case of congenital diseases, the trait becomes manifest from birth. But there are many other diseases where the symptoms appear later in life. Darier-White and Hailey-Hailey diseases are good examples, in which the main features of the disease develop around early adulthood and second to fourth decade, respectively (chapter 59)[36,37].

Age-dependent penetrance presents another difficulty for genetic counseling and, in general, for the establishment of the inheritance pattern of a trait, since some individuals may be too young for a definitive diagnosis.

Fig. 5 Examples of complications to the basic Mendelian patterns of inheritance.

4.2. Variable Expression

A disease genotype can fail to manifest itself completely (reduced penetrance), but it can also manifest in every individual with the disease genotype, but with variation in the severity of symptoms. This variation in the expression of a genotype can be interfamilial (between families) or even intrafamilial (within a family). The variable degree in which a particular genotype is expressed as a phenotype is called ***variable expression***. Darier-White disease, for example, displays complete penetrance in adults, but the severity of the cutaneous manifestations can be variable (chapter 59)[40].

The variable expression of some traits manifest as a tendency to increase in severity and decrease in the age of onset through successive generations. This phenomenon is referred to as ***genetic anticipation*** and it offers an additional challenge in genetic counseling. The molecular basis of genetic anticipation is now known. Certain genes show a variable number in the repetition of a three nucleotide sequence, known as trinucleotide. These repeats range in number in the general population, as part of the normal variation. Through a yet unknown mechanism (DNA polymerase "slippage" has been hypothesized), the number of these repeats can increase dramatically and enter the disease-associated range, where the number of repeats become unstable through successive generations. In general, the number of repeats within this range tends to increase through generations, although it can also decrease. The severity and age of onset of the resulting phenotype correlates with the length of the repeat. The paradigm

of genetic anticipation is myotonic dystrophy, with a "cascade of mild, adult, childhood, or congenital disease in subsequent generations"[41].

4.3. X-inactivation

Although females have a double dose of the X chromosome as compared to males, for most of the X-linked genes (except those present in the pseudoautosomal region) any particular cell in a female has only a single active X chromosome. The mechanism of ***X-inactivation*** or ***lyonization*** is responsible for this dosage compensation in females. During embryonic development, one of the two X chromosomes in each cell is randomly inactivated. From this stage on, all the derived cells will have the same inactivated X chromosome as their progenitor cell. As a result, a female has two alleles for every gene located in the X chromosome, but a given cell expresses only one of them. If one of these alleles happen to contain a disease-causing mutation, the resulting phenotype in the female will depend on the proportion of cells expressing that mutated allele in a particular tissue. This explains why female carriers for X-linked recessive mutations can manifest some symptoms of the disease, whereas an X-linked dominant phenotype in females is usually milder than in males. As an average, each allele of an X-linked gene is active in half of the cells in a female. But this proportion can be variable, also explaining why the expression of X-linked dominant traits is more variable in females than in males. In fact, it can range from a completely normal phenotype to a full spectrum of the disease phenotype.

There are exceptions, however, where a particular X-chromosome is preferentially inactivated, known as ***skewed*** or ***non-random X-inactivation***. Affected females with

incontinentia pigmenti type II (IP2) show highly skewed X-inactivation, suggesting a negative selection against cells with an active IP2 mutant allele[35]. Non-random X-inactivation in different hematopoietic cells has also been proved in obligate carrier females with dyskeratosis congenita[42].

Other situations prone to non-random X-inactivation are the X:autosome translocations. The normal X-chromosome is usually inactivated in the presence of a balanced X:autosome translocation to avoid monosomy for the genes on the translocated autosome. This is thought to be the explanation for affected females for X-linked recessive diseases. Such is the case for anhidrotic ectodermal dysplasia (EDA) in carriers of a t(X;9) translocation[43].

4.4. *De novo* Mutations

The different alleles for a locus present in the population, responsible for both normal variation or disease phenotypes, arise by mutational events in the DNA. This phenomenon can be observed in pedigrees in which a genetic disease appears with no previous family history for that condition. When this is the case, an affected child is born from unaffected parents who do not carry the mutation. ***De novo*** mutations occur during meiosis in one of the parents and are usually more common in severe diseases, with a marked decreased reproductive fitness. The identification of *de novo* mutations can have a profound impact in genetic counseling. A family with one single affected offspring is usually suspected to have a recessive condition, and thus, a risk of 25% for another affected child. The likelihood of the patient having affected offspring will be the same as in the general population. If, on the

contrary, the pedigree with one single affected individual is due to a *de novo* mutation, there is no recurrence risk for the generations of the parents and the siblings of the affected child, since a second *de novo* mutation at the same gene would be highly unlikely. However, the affected individual will transmit the mutation to 50% of its offspring, and therefore it will behave as a typical dominant trait. An exception to the recurrent risk for a *de novo* mutation is the case of mosaicism (section 4.5; chapter 62).

De novo mutations have been identified in different types of epidermolysis bullosa (EB). In dystrophic EB (DEB), for example, the same *de novo* mutation, Gly2043Arg, in the type VII collagen gene, *COL7A1*, has occurred independently throughout the world[44].

4.5. Mosaicism and Chimerism

A general form of mosaicism is that involving X-inactivation in females, with each cell expressing the genes from one single X-chromosome. But mosaicism can also refer to pathogenetic situations.

When a *de novo* mutation first appears in a pedigree, the recurrence risk for a second affected child is extremely low. But there are some instances in which a second affected child is born unexpectedly, although a *de novo* mutation was initially suspected. This is usually due to gonadal or germline ***mosaicism*** (chapter 62). When a mutation or chromosomal abnormality occurs in a cell during embryonic development, its descendent cells will carry the same mutation. On the contrary, the rest of cells, not affected by the mutation, will carry the wild-type (nonmutant) allele. The result of this mutational event will be an individual with two different cell populations. Although the mutation present in one of these populations can

affect one single gene, its behavior will differ from that of a typical Mendelian mutation. The location of the cell and the timing of the mutational event during development will determine the extent and distribution of the cells carrying the mutation. If a *de novo* dominant mutation occurred after the separation of the germline and somatic cell lineages and it took place in the latter, the gonads will not carry the mutated allele, and the risk for the offspring will be 0. On the contrary, if the mutation took place in a germline cell or precursor, a proportion of the resulting gametes will carry the mutated allele. In the first affected generation, the trait will mimic an autosomal recessive trait, since the affected children will be born from unaffected parents. But the trait will be inherited as a classical autosomal dominant trait in successive generations, as stated above for *de novo* mutations. The key information for genetic counseling will be to know whether the mutation is present in the germline of the mosaic individual. The extent of involvement of the gonads will determine the recurrence risk for the offspring.

Maternal germline mosaicism has been demonstrated in DEB, for example. A mutation was identified in an affected child born to unaffected parents [45]. Since the mutation was present only in the patient, while absent in the unaffected parents, the mutation was most likely a *de novo* event. However, the possibility of germline mosaicism was considered and prenatal diagnosis was performed for the next pregnancy. The results indicated that the fetus had, in fact, an identical mutation, indicating germline mosaicism in one of the parents. These findings have significant impact when counseling on the recurrence risk for affected offspring.

When mosaicism affects the skin, it can become apparent by the appearance of different cutaneous patterns, such as the lines of Blaschko (chapter 62)[46,47].

Genetic mosaicism affecting the skin has been classified in two types. ***Type I*** refers to the type of mosaicism delineated above, where the segmental affected skin is thought to carry the mutant allele, while the genotype of the skin outside the affected areas is thought to bear the wild-type allele. But there are examples where this segmental manifestation of the disease is expressed over a more diffused and milder involvement of the same phenotype. This is referred to as ***type II*** involvement. It has been hypothesized that an individual with a type II involvement would have a heterozygous dominant mutation in the germline, giving rise to the diffuse phenotype, and that a second mutation would inactivate or delete the wild-type allele, resulting in the segmental and more severe phenotype.

There are several examples of somatic mosaicism affecting the skin, such as epidermal nevus, Darier-White disease, neurofibromatosis, multiple cutaneous leiomyomatosis or linear porokeratosis superimposed on disseminated superficial actinic porokeratosis[48-52].

Finally, the different populations of cells present in a mosaic are thought to derive from a single zygote. In contrast, there are examples of individuals with cell populations with a complete different genetic composition. These are known as ***chimeras***, and the different cell populations are thought to derive from different zygotes.

4.6. Inbreeding

Inbreeding is one of the factors that can modify the frequency of alleles in a population, increasing the number of homozygotes. When a mating occurs between two individuals who share a common ancestor, this mating is referred to as ***consanguineous***.

A recessive mutation can be passed on through generations in the heterozygous state. If a carrier for a recessive mutation has children with a second carrier of another recessive mutation in the same gene or even the same mutation, their children will have a 25% likelihood of inheriting both mutant alleles. Recessive disorders are not very frequent, since it requires these two carriers to mate. But in inbred populations, the probability that two carriers of the same mutation originating from a common ancestor mate is significantly higher. This is the explanation for the increased incidence of consanguinity among autosomal recessive diseases. The affected offspring from a consanguineous mating will be homozygous for the same mutation and also for the region surrounding that mutated gene, representing two copies of the same chromosomal region travelling from the common ancestor, down through different branches of the family and coming together again in the consanguineous mating. The affected offspring is called ***homozygous by descent***, reflecting the fact that both mutant alleles are copies from the same ancestral mutant allele. The more common a mutant allele is, the less incidence of consanguinity is observed in that disease, and vice versa. When a pedigree is ascertained with only one affected family member, the presence of consanguinity in the parents of the affected is suggestive of an autosomal recessive condition.

When building a pedigree, it is important to consider the possibility of consanguinity, even if denied by the parents, if these come from a small geographical area.

Since in an inbred population there is a higher probability for two persons to be related, it will also be possible that this can happen more than once in the same pedigree. In these situations, one can find affected individuals in different generations of a kindred, simulating an autosomal dominant trait (**Fig. 5*B***).

Most of the considerations applied above for consanguineous matings can also be applied to ***isolated populations***, such as the Finnish, the Icelandic or the French-Canadian populations. They usually derive from a small group of ancestors and are subject to geographical or linguistic isolation. If one of the ancestors was a carrier for a particular mutation, it will be overepresented in the present population. This effect, known as ***founder effect***, is the explanation for the fact that some diseases or mutations are present at very different frequency in isolated populations when compared to other groups.

In South Africa, for example, a large proportion of the present population descends from a small number of Dutch settlers and, in fact, the high proportion of cases of porphyria (chapter 50) variegata have been traced back to a few settlers in the 17th century[53].

4.7. Quasidominant or Pseudodominant Inheritance

There is another situation where the inheritance of a recessive trait mimics an autosomal dominant pattern (**Fig. 5*C***). This is the case for matings between an affected individual for a recessive trait and an unaffected carrier for the same or a different mutation in the same gene. In these situations, the offspring has a 50% probability of inheriting two mutant alleles. Although this is the same risk as for an autosomal dominant trait, this parent-to-child transmission does not occur in successive generations, unless the unlikely event of a second mating between a homozygote and a heterozygote. Instead, the risk of having an affected child became as low as for the general population. These situations can happen in both inbred and outbred populations, although the probability of such matings is increased in the first ones.

An example of pseudodominant inheritance has been reported in a family in which a mother and a son were affected with APL, suggesting dominant inheritance, whereas APL is commonly inherited as an autosomal recessive disease. The molecular analysis revealed, in fact, that both the affected mother and son were homozygous for mutations in the same gene, in accordance for the recessive nature of the disease [39].

4.8. Imprinting

Up to now, we have been discussing recurrence risk depending only on whether an individual has inherited the disease-associated genotype. But some genes are subject to an additional level of regulation by epigenetic phenomena: the sex of the transmitting parent will determine whether that gene will be expressed in the offspring, known as ***imprinting***. The phenomenon of imprinting influences the assessment of recurrence risk in genetic counseling. The risk for a parent of transmitting a mutated allele remains the same, 50%, for females and males, but the probability that the mutated allele will be expressed in the offspring will differ depending on the sex of the transmitting parent. These sex specific alterations involve DNA methylation.

The effect of imprinting is especially apparent in the case of ***uniparental disomy (UPD)***, where both homologous chromosomes are derived from the same parent, instead of one from each parent. In ***heterodisomy***, the two homologues from the same parent are inherited, as opposed to ***isodisomy,*** where one chromosome from one parent is duplicated. UPD can go undetected for some chromosomes. For others, it can result in a disease phenotype, such UPD for chromosome 15 in Prader -Willi and Angelman syndromes.

The appearance of UPD affecting recessive mutant alleles can bring the mutation to homozygosis in affected individuals. The Herlitz variant of junctional epidermolysis bullosa (H-JEB), also an example of locus heterogeneity, can be caused by mutations in one of the three genes coding for the different subunits of laminin, *LAMA3*, *LAMB3* and *LAMC2*. Both paternal and maternal UPD have been described in H-JEB patients. In two independent pedigrees, a complex combination of maternal uniparental hetero- and isodisomy resulted in homozygosity of a maternal *LAMB3* mutation in the affected child[54,55]. In a third case, paternal uniparental isodisomy resulted in a patient homozygous for a paternal *LAMC2* mutation[56].

4.9. Loss of Heterozygosity

There are some disorders in which the susceptibility for the development of certain tumors is inherited as a Mendelian trait. In these cases, affected individuals usually carry a germline mutation in a ***tumor suppressor gene***, involved in the regulation of cell growth. Inactivating (loss-of-function) mutations in both alleles of a tumor suppressor gene can contribute to tumor formation. These familial cases are, therefore, at a higher risk of developing a tumor, since a single somatic mutation inactivating the second, normal allele of the same gene, can trigger tumor formation. For this reason, inherited forms of cancer susceptibility can result in the development of multiple tumors in the same individual, as opposed to the single tumors found in the sporadic cancers. When studying these tumors, the usual finding is to identify two mutations in the two alleles of the gene. One of them will correspond to the germline mutation, whereas the second one will be a new somatic mutation, referred to as the "two-hit hypothesis"[57]. One common type of somatic mutation found in

tumors is known as ***loss of heterozygosity*** (***LOH***). When analyzing the DNA of tumor samples from an individual heterozygous for specific polymorphisms (usually microsatellites, section 6.1.1), the results from the tumor can show homozygosity for one single allele, instead. This phenomenon is known as LOH. Advanced tumor cells can show extensive LOH through the genome, however, there are certain regions for which LOH is common to several tumors. This usually happens in the vicinity of tumor suppressor genes.

4.10. Mitochondrial Inheritance

Mitochondrial inheritance is another example of a single gene defect that does not follow a classical Mendelian pattern of inheritance. Each cell has many mitochondria, each of them containing one or more copies of the mitochondrial genome, a 16-kilobase (Kb) circular DNA molecule. Only mothers pass mitochondria to the offspring, thus there is no paternal complement. During cell division, mitochondria will segregate randomly to each of the two daughter cells. Thus, if a mother carries a mutation in one gene located on the mitochondrial chromosome in only some of her mitochondria, the daughter cells can inherit a variable number of organelles carrying the mutation. A situation where all mitochondria carry the mutation is known as ***homoplasmy***, versus ***heteroplasmy***, referring to a mixture of mitochondria carrying the normal or the mutant version. The resulting phenotype will depend on the proportion of mutant and normal mitochondria in a specific tissue.

The result of mitochondrial inheritance is a pedigree where both females and males can be affected, but only females transmit the disease trait. The proportion of affected

children seems to be random, and there is a high degree of variability in the expression of the phenotype.

As an example, a transition from adenine (A) to guanine (G) at position 7445 of the mitochondrial DNA, affecting the gene encoding the tRNA for serine, has been identified as the cause for a form of palmoplantar keratoderma with or without sensorineural deafness. In these pedigrees the disease presents with incomplete penetrance and with varying severity[58-62].

4.11. Phenocopy

A ***phenocopy*** refers to those phenotypes caused by environmental factors that mimic a similar phenotype caused by genetic mutations. Porphyria cutanea tarda (PCT), for example, is an autosomal dominant disease caused by mutations in the *URO-D* gene, encoding the uroporphyrinogen decarboxilase. Affected patients are predisposed to light-sensitive dermatitis. It has been shown, however, that a similar phenotype can be caused by exposure to certain chemicals. Between 1955 and 1961, in fact, around 3000 patients developed PCT in Turkey. The cause was determined to be exposure to hexachlorobenzene, used to preserve the wheat seeds. PCT has also been reported in individuals with industrial exposure to certain hydrocarbons[63], as well as in workers exposed to some pesticides[64].

5. Chromosome Disorders

In the above sections, the disease-causing mutation affects a single gene. But this is not always the case. Around 65% of the spontaneous abortuses studied, for example, exhibit

chromosome anomalies[65], which can involve one or more genes contained in the affected region.

Chromosomal abnormalities can result in abnormalities in the number or in the structure of the chromosomes. The differences in DNA composition along a chromosome have allowed the development of various ***banding techniques***, which enable the identification of each particular chromosome in a karyotype, as well as the identification of chromosomal anomalies, both numerical or structural. In addition, the use of FISH (Fluorescent In Situ Hybridization), which can detect DNA sequences as short as 1 Kb^3, allows the identification of very small changes in the chromosome constitution, not visible by the use of conventional banding techniques.

Chromosomal numerical abnormalities can involve the whole genome, poliploidy or nullisomy, or just one pair of homologues, aneuploidy.

The only forms of ***poliploidy*** observed in humans are ***triploidy***, with three complete sets of chromosomes (3n), and ***tetraploidy***, with four chromosome complements (4n). However, most polyploid embryos are spontaneously aborted.

A particular case of cells with a different number of chromosome complement, yet viable and functional normal cells, is the case of the platelets and keratinocytes, which loose their nuclei and, as a consequence, are ***nullisomic***.

Aneuploidy, on the other hand, refers to the presence of an additional chromosome or its absence. The only observed aneuploidies are ***trisomies***, with three copies of a single chromosome and, very rarely, ***monosomies***. Trisomies for a whole chromosome are only observed for chromosomes 21 (Down syndrome), 18 (Edwards syndrome), and 13 (Patan syndrome). Regarding monosomies, only monosomy for chromosomes X (Turner syndrome)

and 21 have been observed, the latter being extremely rare. Partial trisomies or monosomies due to duplication, deletion or segregation of balanced translocations (section 5) can also be observed.

Structural abnormalities involve breakage of chromosomes, instead of a change in chromosome number. Most chromosomal breaks are normally reconstituted by the repair enzymes. When this is not the case, however, the breaks result in structural abnormalities. They will be ***balanced*** if there is no net gain or loss of genetic material, or ***unbalanced***, when they are accompanied by additional or missing genetic information. The stability of the resulting rearranged chromosome through cell division will depend on the presence of a centromere and two telomeres, the necessary elements for the correct segregation of chromosomes.

Balanced chromosomal abnormalities include some types of translocations and inversions. In a ***reciprocal translocation***, two breaks in two non-homologous chromosomes occur and the chromosomal segments distal to the break are exchanged. The total number of chromosomes remains the same. If the translocation breakpoint affects a gene or its regulatory sequence, or otherwise, places it under the control of novel regulatory elements, it can result in an observable phenotype. The main problem with reciprocal translocations arises in the offspring of the carriers. Depending on the combination of parental chromosomes inherited, they can result in partial trisomy and monosomy for the chromosomal regions involved in the translocation.

Inversions are balanced rearrangements in which two breaks occur in the same chromosome and the intervening segment is inverted before the chromosome is reconstituted. In a ***pericentric inversion***, the inverted segment includes the centromere, whereas in a

paracentric inversion, the two breaks happen in the same chromosome arm, without involvement of the centromere.

There are only a few examples of chromosomal abnormalities underlying skin or hair disorders as their major clinical feature. EDA, for example, has been diagnosed in several patients with rearranged X chromosomes, such as X;autosome translocations or genomic deletions involving the *EDA* gene [66]. Another example is Ambras syndrome, a unique form of congenital universal hypertrichosis (chapter 70). Although the molecular basis of the disease is still unknown, two unrelated cases have been reported with cytogenetic abnormalities involving one common chromosomal region, 8q22. One of the patients has a balanced pericentric inversion denominated (8) (p11.2, q22)[67] and a second patient has the paracentric inversion (8) (q12, q22)[68]. Although the specific gene defect underlying Ambras syndrome is still unknown, the fact of having two independent chromosomal abnormalities involving the same chromosomal region and both in patients with the same rare disease suggests that the chromosomal region 8q22 contains a candidate gene or regulatory region responsible for the Ambras syndrome phenotype[69].

Regarding ***unbalanced chromosomal abnormalities***, there are several mechanisms that can lead to a net gain or loss of chromosomal material. ***Deletions*** and ***duplications***, for example, can be the result of the segregation of a balanced translocation, among other possible causes.

6. Identification of Disease genes

As stated at the beginning of this chapter, disease gene identification is crucial for many areas of medicine, such as genetic counseling and prenatal diagnosis. There are two main approaches generally applied for the identification of the disease gene for a Mendelian disorder, in which a single mutated gene is expected to cause the disease phenotype. The approach will usually depend on the kind of information available on the particular trait.

Functional cloning requires some previous knowledge about the defective function, but no positional information is required. When the abnormal function leading to the disease phenotype is known, such as the biochemical defect or the defective protein itself, several approaches can be applied to clone the corresponding gene. However, even though functional cloning has provided many of the cloned disease genes, its application is very limited, since there are only a few disorders for which the biochemical defect is known. When the pre-existing information on the phenotype helps to point a particular gene, this gene becomes a ***candidate gene***, which can be analyzed directly for its involvement in the phenotype. The candidate gene approach can use information on the defective pathway, function, structure or tissue underlying the phenotype. The expression of a gene (in the same tissues as those affected by the disease phenotype), its function (its role in a given metabolic pathway or cellular structure), its homology or relation to mutated genes responsible for close phenotypes, for example, can be a good starting point to choose a candidate gene. An additional and greatly helpful source of information are animal models. There are many examples, although also some exceptions, where closely related phenotypes in mice and human, for example, are caused by mutations in orthologous genes (referring to homologous genes in different species).

By means of the candidate gene approach, two separate groups identified mutations in the *TGM1* gene as the molecular basis for lamellar ichthyosis (LI; chapter 57), an autosomal recessive disease of cornification[70,71]. *TGM1* encodes the keratinocyte form of transglutaminase (TGK), involved in cross-link reactions during the formation of the cornified envelope. Huber *at al* analyzed the *TGM1* gene following studies that showed decreased or absent activity and expression of TGK in LI patients. Russell *et al*, on the other hand, studied *TGM1* and other genes on the basis of the role of the encoded proteins in the formation of the stratum corneum.

For ***positional cloning***, on the contrary, no previous knowledge on the defective function is required; it starts with the chromosomal assignment of the region containing the defective gene by means of linkage analysis. Positional cloning appears as a powerful tool in disease-gene identification, since for most of diseases the only information available is the observable phenotype segregating in pedigrees.

The identification of the gene responsible for the EDA phenotype is an example of positional cloning. The genetic locus was mapped to the chromosomal region Xq12-q13.1 by means of linkage analysis and translocations in female patients. After identifying putative new genes in the disease interval, Kere *et al* identified the *EDA* gene[72]. The location of the gene, the identification of mutations in the patients and its expression proved the pathogenetic role of the gene in the EDA phenotype.

The approach that have yielded more results, however, uses a combination of both functional and positional information. In the ***positional candidate approach***, once the chromosomal location of the disease locus is known, functional information is utilized to

directly analyze the candidate genes contained in that particular chromosomal region (section 6.1.3).

6.1. Positional Cloning

When no information is available on the possible defective function or product, the first step for the identification of the disease gene is the mapping of the chromosomal location that most likely contains the gene, the ***disease locus***. There are some instances where there is previous information on the location of the gene, such as the presence of chromosomal abnormalities or LOH in a patient, which can considerably accelerate the identification of disease genes. These chromosomal abnormalities are usually translocations, inversions, deletions or duplications.

The occurrence of X:autosome translocations in females affected with EDA and involving the same breakpoint on the X-chromosome helped to indicate the location of the causative gene. Eventually, a combination of linkage analysis, translocation breakpoints and small deletions in patients led to the identification of the *ED1* gene, responsible for the EDA phenotype[72].

In the same way, LOH can give a hint on where the disease gene might be. This has proven to be very useful in the identification of tumor suppressor genes. Familial cylindromatosis, for example, is characterized by multiple tumors of the skin appendages or cylindromas. The predisposition to develop the tumors is inherited as an autosomal dominant trait. Over 70% of the tumors show LOH in the vicinity of the gene that causes the disease and in all cases the allele lost in the tumors is the wild-type allele. Moreover, the tumors show

no other regions of LOH along the genome. In the example of cylindromastosis, linkage analysis combined with the presence of LOH led to the identification of the underlying mutated tumor-suppressor gene, *PTEN*[73,74].

6.1.1. Linkage Analysis

For most of the disorders, however, there is no previous information on the defective function or possible location of the disease gene. In such situations, a ***linkage analysis*** approach is performed to map the location of the disease gene. The aim of linkage analysis is the identification of a chromosomal region that is transmitted along with the disease phenotype (co-segregates). When this region is found, the disease is said to be ***linked*** to that genetic locus.

Gene mapping in general has proven to be very useful not only for the identification of the actual disease gene, but also for genetic counseling. Once the location of the gene is known, we can tract the disease gene in the families affected by following the inheritance of this chromosomal region.

The requirements to perform a linkage analysis include a thorough definition of the phenotype, the availability of pedigrees with enough linkage information (power), the availability of a dense map of highly polymorphic markers and, finally, the application of statistical analysis.

When defining the ***phenotype*** to be studied, it has to be considered that the correspondence between a gene and a disease is not always linear. There are numerous examples of allelic heterogeneity, with different mutations in the same gene resulting in

different phenotypes (**Figure 2**; section 3) and locus heterogeneity, in which the same or very close phenotypes are caused by mutations in different genes (**Figure 3**; section 3).

Since the purpose of linkage analysis is mapping a single gene, it is extremely important to work with as homogeneous of a sample as possible to avoid locus heterogeneity. It is important, thus, to define clear-cut phenotypic classes. But this is not always possible for all diseases. Some close phenotypes are represented by a continuum of symptoms, rather than by clear-cut phenotypic classes. For others, the age of the individual can lead to misdiagnosis because of age-dependent penetrance, for example, or because of the progressive nature of the disease, in which the patient can be at a too early or a too advanced stage for the disease to be distinguished from close phenotypes.

Finally, since in all these studies the disease gene is unknown and the phenotype itself will be used to track the disease gene along the pedigree, it is essential to know if special circumstances such as incomplete, age-dependent penetrance or phenocopies, for example, might be expected.

Once the phenotype has been clearly established, a large enough ***sample*** needs to be chosen to achieve statistically significant results. Linkage studies are usually performed either on a single large pedigree, or on several smaller ones. The first approach ensures genetic homogeneity of the sample, but it may not be available for every disease, whereas the second introduces the unwanted possibility of locus heterogeneity, but is usually more common for the majority of diseases.

With a large enough and well-characterized collection of samples, the linkage analysis itself can be performed. Linkage analysis, and genetic mapping methods in general, rely on ***genetic recombination*** or ***crossing-over***.

The rationale is that the closer two loci are together, the less frequently they will be separated by recombination during meiosis. Loci very close together will rarely be separated by recombination, while recombination will occur between loci separated by a certain distance on the same chromosome, the larger the distance, the higher the frequency of recombination. A particular combination of alleles for several close loci on the same chromosome is known as ***haplotype*** (**Fig. 6**). Finally, and due to the independent assortment of maternal and paternal homologous chromosomes during meiosis, loci located on different chromosomes will segregate independently.

Fig. 6 Examples of haplotypes for a chromosomal region containing four polymorphic loci, *A*, *B*, *C* and *D*.

The frequency of recombination, or θ, between two loci will reflect the genetic and physical distance between them. A recombination fraction of 1% or 0.01 is equivalents to 1 cM (cM: genetic distance measure unit).

When linkage analysis is performed, the disease gene is unknown. The inheritance of the disease phenotype is compared to the inheritance of polymorphic loci or ***markers*** along the genome. The goal is to find one or more polymorphisms that are inherited in the same way as would be expected for the disease gene. The disease locus will be then ***linked*** to such a marker, which will allow us to follow the disease gene through the pedigree until it is actually identified. It is important to note that the linked marker is not necessarily the disease gene itself, but a closely linked locus.

There are many types of DNA ***polymorphisms***, but not all of them are good markers for a linkage analysis. The two requirements are *i*) the informativeness and *ii*) the distribution of the polymorphisms.

The informativeness of a polymorphism refers to both the number and allelic frequency of the different alleles at a given locus. The more informative a polymorphism is, the greater the chance for an individual to be heterozygous or ***informative*** for linkage.

The two types of markers generally used in linkage analysis are ***microsatellites*** and ***SNPs*** (Single Nucleotide Polymorphisms) (**Fig. 7*A, B***). Microsatellites are extensively used in linkage analysis. They consist on a tandem repetition of a 1-13 nucleotides unit sequence[75]. Dinucleotides are the most commonly used microsatellites, with a repeat unit consisting of two nucleotides, followed by tetra- and trinucleotides, four and three-nucleotide repeats. Among the dinucleotides, the $(CA)_n$ repeats are by far the most numerous. For the SNPs the polymorphism involves the presence of a given nucleotide at a particular position, while the polymorphism for microsatellites refers to their length, due to the variation in the number of repeats of the basic unit sequence (**Fig. 7*C***).

Fig. 7 Examples of different types of polymorphisms.

The second requirement for a polymorphism to be a good marker in a linkage analysis depends on the density of the ***map***. The longer the distance between two loci, the larger the sample needed to detect linkage. In a typical linkage analysis, markers spaced at about 10 cM and distributed along the entire genome are analyzed, in what is known as a ***genome-wide scan***. This implies an average of 300 markers, analyzed using a very powerful technique

known as ***PCR*** (Polymerase Chain Reaction) that allows the analysis of a specific sequence of DNA by its exponential amplification.

Once the data from a genome-wide scan is gathered statistical analyses are performed to give us an idea of how significant our results are. When the genome as a whole is analyzed, the possibility of false positives has to be considered, since more than one region that mimics the inheritance of the disease gene in a particular pedigree can be found, although only by chance. The statistical value used is the ***LOD score*** or ***Z***, and is defined as the logarithm of the odds in favor of linkage. The LOD score will test the linkage between two loci in what is known as a ***two-point analysis***. A LOD score of 3 or greater is used to accept linkage between two loci, whereas a LOD score of –2 or less is used to exclude linkage. The values between 3 and –2 are inconclusive and are not sufficient to prove any hypothesis (**Fig. 8**). LOD score calculations are performed using statistical computer software, such as LINKAGE[76]. These calculations are done assuming a particular inheritance pattern, disease allele frequency, penetrance, possibility of phenocopies, etc. This kind of analysis, where the model under study is specified, is known as parametric analysis.

The use of more than one marker can be helpful, since a particular pedigree can be informative for different markers. The analysis of the data from different markers at a time is called ***multipoint analysis***, as opposed to two-point analysis as described above.

Fig. 8 Example of two-point LOD score values for markers on chromosome 1.

6.1.2. Refinement of the Linkage Interval

The identification of the linkage interval is not the end of the work, but just the beginning. It gives us an idea of the most likely location for the disease gene. But these intervals are normally too large to work with, 5-10 cM, or even larger, with hundreds of genes. For this reason, the step of narrowing down the disease locus is essential, unless an obvious candidate gene maps to the initially identified interval. This refinement of the disease locus can be achieved by different means, including the finding of LOH or a chromosomal abnormality in a subset of patients. But for the vast majority of genetic diseases, this is not the case.

As described above, in a genome-wide scan markers spaced at an average of 10 cM are used. But the genetic maps are much denser. Thus, once the disease interval has been identified, the best way to narrow it down is to increase the sample of individuals and the number of markers analyzed.

In this step of the study, the analysis of haplotypes is essential to find recombinations that can further define the linkage interval. The goal is to define the boundaries of the smallest interval, defined by the closest markers that show recombination with the disease phenotype. When several close markers are analyzed, the disease gene will be linked to a certain haplotype, instead of a single marker. This disease-associated haplotype can include a different number of markers in different members of the family, since recombination events can break this shared haplotype along the generations in the pedigree. Once an individual is identified where a recombination has occurred inside the disease-associated haplotype, known as a ***key recombination*** event, this will reduce the region where the disease gene could be located (**Fig. 9**). By identifying different key recombinants, the most likely location of the disease gene can be narrowed down.

Fig. 9 Example of haplotype analysis for the refinement of a linkage interval.

6.1.3. Positional Cloning and Positional Candidate Genes

When the linkage interval cannot be further reduced, the identification of the actual disease gene begins. Traditionally, this had to be done by an arduous and not very fruitful positional cloning strategy, where different techniques to isolate genes were applied. Although this is still necessary in some instances, the information and resources available from the public and private Human Genome Project efforts have made the positional candidate approach the most powerful strategy. In a positional candidate gene approach, the candidates will specifically map to the disease locus interval.

As a result of the public Human Genome Project effort, there are currently three internet servers providing free access to the up to date sequence data generated, National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov), University of California, Santa Cruz (UCSC; http://genome.cse.ucsc.edu/) and Ensembl (http://www.ensembl.org/). The user can browse integrated maps where the physical location of different elements, such as polymorphic markers, known genes, predicted genes, ESTs (Expressed Sequence Tags), or STSs (Single Tagged Sites), among others, are shown. Eventually, the sequence and the position of all human genes will allow the analysis of all the genes contained in a disease interval.

The identification of the genes responsible for Darier-White or Hailey-Hailey diseases are examples of positional candidate approaches (chapter 59). The genetic locus for the gene

underlying Darier-White disease was located to chromosome 12 by linkage analysis[77,78]. After refinement of the region, several genes were identified in the linkage interval. Among them, the *ATP2A2* gene, encoding the calcium pump SERCA2, was considered a good candidate due to the role of calcium in epithelial junctions (defective in the epidermis of Darier-White patients) and cell differentiation. In fact, *ATP2A2* was found to be mutated in Darier-White patients.

In a separate study, linkage of the similar Hailey-Hailey disease phenotype had been established to chromosome 3[79]. Fortunately, one gene in the linkage critical interval, *ATP2C1*, showed homology to calcium ATPases and, thus, it might have a function related to SERCA2. As expected, this gene turned out to be mutated in Hailey-Hailey patients[80,81].

Mouse models are also emerging as a powerful tool in the identification of disease genes. This approach relies on the existence of chromosomal regions that share the same origin in evolution, known as ***syntenic*** regions. Thus, once the genetic interval in human is identified, the syntenic interval in mouse can be analyzed to identify additional candidate genes. Also, if a similar phenotype in mouse has been mapped to this region and the mutated gene is known, this becomes an excellent candidate for the human disease counterpart, and viceversa. The identification of mutations in the hairless gene in patients with APL[82] is a good example for mutations in the human, mouse, rat and monkey orthologous genes resulting in a very similar phenotype.

If there are no obvious candidate genes in the region of interest or if none of the genes analyzed is the responsible for the disease, a positional cloning strategy will have to be performed for the identification of new genes in the interval.

6.1.4. Identification of Mutations

Once a candidate gene is identified, its involvement in the disease has to be proven. This involves the identification of DNA variants present only in patients and absent in the control population. There are several techniques that can be initially used to search for mutations, such as SSCP (Single Strand Conformation Polymorphism) or CSGE (Conformation Sensitive Gel Electrophoresis). However, eventually all of them lead to the determination of the precise DNA sequence of the mutant allele.

There are different types of ***mutations***. ***Point mutations*** consist of the substitution of a single nucleotide position. This change can lead to an amino acid substitution in the case of a ***missense*** mutation, to the generation of a termination codon, ***nonsense***, or to no change in the protein sequence, ***silent*** mutation. There can also be small duplications, insertions or deletions. If they affect a piece of the protein coding sequence that is not a multiple of three nucleotides (a codon), they will lead to a ***frameshift*** in the protein sequence and to the generation of a premature termination codon. These types of mutations can affect both coding and non-coding sequence, such as the promoter region, consensus splicing signals and other key intronic positions. ***Large rearrangements***, such as partial or complete deletions of the entire gene, can also occur.

It is important to note that some of these DNA mutations, especially missense and silent substitutions, can be non-pathogenetic variants, present in the general population as polymorphisms.

7. Complex Traits

Classical Mendelian diseases, where a genetic defect in one single gene is enough for the disease to develop, are only a small number of genetic diseases. There are many common conditions that, although being heritable, are not due to a single-gene defect. Instead, they are thought to be determined by the interaction of multiple genetic factors among them and with the environment, as well. This combination of genetic factors usually confers only susceptibility to a certain disease, but cannot be used as a diagnostic tool by itself. Such traits are referred to as ***complex*** or ***multifactorial traits***. They tend to cluster or aggregate in families, but they do not follow any Mendelian pattern of inheritance.

As opposed to Mendelian traits, recurrence risk assessment for multifactorial traits is empirical, i.e., it is based on epidemiological data, where the real frequencies for different degree relatives have been determined. Some examples of dermatological complex traits are atopic dermatitis, eczema, psoriasis and alopecia areata (chapters 9, 13 and 69). The absence of any known Mendelian inheritance pattern can be observed when a collection of pedigrees with the same complex trait is analyzed at once, as shown in **figure** 10.

Fig. 10 Pedigrees with alopecia areata as an example of multifactorial dermatological disease.

Why, then, if Mendelian disorders are rather rare compared to the more common complex traits, have they been, and still are, the focus of much research in Human Genetics? The complex nature of a multifactorial trait increases the difficulty of gene identification.

Although similar techniques are used if compared to single gene defects, they require different methods to interpret the results. The initial goal in the study of complex traits is the identification of different DNA variants in different genes, each of them with a contribution to the final phenotype and conferring a degree of susceptibility to the carrier, instead of a mutation in a single gene. As it has already been shown for some diseases, these variants can also be present in the general population, although in a significantly different frequency. Large collections of patients need to be studied in order to identify a statistically significant contribution of a given genetic factor. As it has already been shown, these susceptibility variants could be present in genes responsible for similar Mendelian diseases.

As an example, homozygosis for the mutation W185X, in the *PVLR1* gene, has been shown to cause autosomal recessive cleft lip/palate (CLP) with ectodermal dysplasia, very frequent on the Margarita island, north to Venezuela[83]. This same mutation has been shown to be overepresented among individuals with the multifactorial isolated CLP in northern Venezuela (5.8%), a, when compared to control individuals from the same population (0.4%). Although these results are highly significant, they constitute only a moderate genetic risk factor, since just a small number of CLP patients carry the mutation[84,85].

Another example for the role of Mendelian gene disorders in related complex traits is the case of the gene *SPINK5*, recessive mutations in which cause Netherton disease, a skin disorder associated with atopy[86,87]. The authors have shown a statistically significant association between the amino acid variant Glu420Lys in the *SPINK5* gene and atopy or atopic dermatitis[87]. Again, this does not mean that *SPINK5* is responsible for atopic dermatitis, but it could shed some light into the etiology of AD and atopy.

Studies like those described above underscore the value of the knowledge gathered from rare Mendelian disorders to derive important clues for more common related complex traits. In addition, the identification of genetic factors for multifactorial traits, although being just a susceptibility factor, can indicate unknown pathways susceptible for the development of targeted therapies and to a way to modify key environmental factors in these major public health issues.

Disease gene identification is of chief importance not only for genetic counseling and prenatal diagnostic, but it is also the first step in the elucidation of the mechanisms underlying a disease and the eventual development of a suitable treatment.

References

1. Baltimore D. Our genome unveiled. Nature 2001; 409**:**814-6.

2. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409**:**860-921.

3. Miller OJ, E. T. Human chromosomes. 4th ed. New York: Springer-Verlag New Yorc Inc, 2001.

4. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001; 409**:**928-33.

5. Richard G, Smith LE, Bailey RA, et al. Mutations in the human connexin gene GJB3 cause erythrokeratodermia variabilis. Nat Genet 1998; 20**:**366-9.

6. Xia JH, Liu CY, Tang BS, et al. Mutations in the gene encoding gap junction protein beta-3 associated with autosomal dominant hearing impairment. Nat Genet 1998; 20**:**370-3.

7. Liu XZ, Xia XJ, Xu LR, et al. Mutations in connexin31 underlie recessive as well as dominant non-syndromic hearing loss. Hum Mol Genet 2000; 9**:**63-7.

8. Lopez-Bigas N, Olive M, Rabionet R, et al. Connexin 31 (GJB3) is expressed in the peripheral and auditory nerves and causes neuropathy and hearing impairment. Hum Mol Genet 2001; 10**:**947-52.

9. Armstrong DK, McKenna KE, Purkis PE, et al. Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma. Hum Mol Genet 1999; 8**:**143-148.

10. Norgett EE, Hatsell SJ, Carvajal-Huerta L, et al. Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma. Hum Mol Genet 2000; 9**:**2761-2766.

11. Smith FJ, Eady RA, Leigh IM, et al. Plectin deficiency results in muscular dystrophy with epidermolysis bullosa. Nat Genet 1996; 13**:**450-7.

12. Pulkkinen L, Smith FJ, Shimizu H, et al. Homozygous deletion mutations in the plectin gene (PLEC1) in patients with epidermolysis bullosa simplex associated with late-onset muscular dystrophy. Hum Mol Genet 1996; 5**:**1539-46.

13. Koss-Harnes D, Hoyheim B, Anton-Lamprecht I, et al. A site-specific plectin mutation causes dominant epidermolysis bullosa simplex Ogna: two identical de novo mutations. J Invest Dermatol 2002; 118**:**87-93.

14. Lane EB, Rugg EL, Navsaria H, et al. A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering. Nature 1992; 356**:**244-6.

15. Coulombe PA, Hutton ME, Letai A, Hebert A, Paller AS, Fuchs E. Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: genetic and functional analyses. Cell 1991; 66**:**1301-11.

16. Dong W, Ryynanen M, Uitto J. Identification of a leucine-to-proline mutation in the keratin 5 gene in a family with the generalized Kobner type of epidermolysis bullosa simplex. Hum Mutat 1993; 2**:**94-102.

17. Bonifas JM, Rothman AL, Epstein EH, Jr. Epidermolysis bullosa simplex: evidence in two families for keratin gene abnormalities. Science 1991; 254**:**1202-5.

18. Chan YM, Yu QC, Fine JD, Fuchs E. The genetic basis of Weber-Cockayne epidermolysis bullosa simplex. Proc Natl Acad Sci U S A 1993; 90**:**7414-8.

19. Chen MA, Bonifas JM, Matsumura K, Blumenfeld A, Epstein EH, Jr. A novel three-nucleotide deletion in the helix 2B region of keratin 14 in epidermolysis bullosa simplex: delta E375. Hum Mol Genet 1993; 2**:**1971-2.

20. Fukai K, Oh J, Frenk E, Almodovar C, Spritz RA. Linkage disequilibrium mapping of the gene for Hermansky-Pudlak syndrome to chromosome 10q23.1-q23.3. Hum Mol Genet 1995; 4**:**1665-9.

21. Wildenberg SC, Oetting WS, Almodovar C, Krumwiede M, White JG, King RA. A gene causing Hermansky-Pudlak syndrome in a Puerto Rican population maps to chromosome 10q2. Am J Hum Genet 1995; 57**:**755-65.

22. Dell'Angelica EC, Shotelersuk V, Aguilar RC, Gahl WA, Bonifacino JS. Altered trafficking of lysosomal proteins in Hermansky-Pudlak syndrome due to mutations in the beta 3A subunit of the AP-3 adaptor. Mol Cell 1999; 3**:**11-21.

23. Anikster Y, Huizing M, White J, et al. Mutation of a new gene causes a unique form of Hermansky-Pudlak syndrome in a genetic isolate of central Puerto Rico. Nat Genet 2001; 28**:**376-80.

24. Suzuki T, Li W, Zhang Q, et al. Hermansky-Pudlak syndrome is caused by mutations in HPS4, the human homolog of the mouse light-ear gene. Nat Genet 2002; 30**:**321-4.

25. McElreavey K, Krausz C. Sex Chromosome Genetics '99. Male infertility and the Y chromosome. Am J Hum Genet 1999; 64**:**928-33.

26. Sun C, Skaletsky H, Birren B, et al. An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. Nat Genet 1999; 23**:**429-32.

27. Foresta C, Ferlin A, Moro E. Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. Hum Mol Genet 2000; 9**:**1161-9.

28. Hu ZL, Smith L, Martins S, Bonifas JM, Chen H, Epstein EH, Jr. Partial dominance of a keratin 14 mutation in epidermolysis bullosa simplex--increased severity of disease in a homozygote. J Invest Dermatol 1997; 109**:**360-4.

29. Martinez-Mir A, Liu J, Gordon D, et al. EB simplex superficialis resulting from a mutation in the type VII collagen gene. J Invest Dermatol 2002; 118**:**547-9.

30. Djabali K, Martinez-Mir A, Horev L, Christiano AM, Zlotogorski A. Evidence for extensive locus heterogeneity in naxos disease. J Invest Dermatol 2002; 118**:**557-60.

31. Zonana J, Elder ME, Schneider LC, et al. A novel X-linked disorder of immune deficiency and hypohidrotic ectodermal dysplasia is allelic to incontinentia pigmenti and due to mutations in IKK-gamma (NEMO). Am J Hum Genet 2000; 67**:**1555-62.

32. Kenwrick S. Survival of male patients with incontinentia pigmenti carrying a lethal mutation can be explained by somatic mosaicism or Klinefelter syndrome. Am J Hum Genet 2001; 69**:**1210-7.

33. Zhao G, Xia H, Xia Y. A pedigree of Y-linked retinitis pigmentosa. Investigation report of a family. Chin Med J (Engl) 1995; 108**:**631-3.

34. Goltz RW. Focal dermal hypoplasia syndrome. An update. Arch Dermatol 1992; 128**:**1108-11.

35. Smahi A, Courtois G, Vabres P, et al. Genomic rearrangement in NEMO impairs NF-kappaB activation and is a cause of incontinentia pigmenti. The International Incontinentia Pigmenti (IP) Consortium. Nature 2000; 405**:**466-72.

36. Burge SM, Wilkinson JD. Darier-White disease: a review of the clinical features in 163 patients. J Am Acad Dermatol 1992; 27**:**40-50.

37. Burge SM. Hailey-Hailey disease: the clinical features, response to treatment and prognosis. Br J Dermatol 1992; 126**:**275-82.

38. McGee TL, Devoto M, Ott J, Berson EL, Dryja TP. Evidence that the penetrance of mutations at the RP11 locus causing dominant retinitis pigmentosa is influenced by a gene linked to the homologous RP11 allele. Am J Hum Genet 1997; 61**:**1059-66.

39. Zlotogorski A, Martinez-Mir A, Green J, et al. Evidence for Pseudodominant Inheritance of Atrichia with Papular Lesions. J Invest Dermatol 2002; 118**:**881-886.

40. Munro CS. The phenotype of Darier's disease: penetrance and expressivity in adults and children. Br J Dermatol 1992; 127**:**126-130.

41. de Die-Smulders CE, Howeler CJ, Mirandolle JF, et al. Anticipation resulting in elimination of the myotonic dystrophy gene: a follow up study of one extended family. J Med Genet 1994; 31**:**595-601.

42. Ferraris AM, Forni GL, Mangerini R, Gaetani GF. Nonrandom X-chromosome inactivation in hemopoietic cells from carriers of dyskeratosis congenita. Am J Hum Genet 1997; 61**:**458-61.

43. Zankl A, Addor MC, Cousin P, Gaide AC, Gudinchet F, Schorderet DF. Fatal outcome in a female monozygotic twin with X-linked hypohydrotic ectodermal dysplasia (XLHED) due to a de novo t(X;9) translocation with probable disruption of the EDA gene. Eur J Pediatr 2001; 160**:**296-9.

44. Wessagowit V, Ashton GH, Mohammedi R, et al. Three cases of de novo dominant dystrophic epidermolysis bullosa associated with the mutation G2043R in COL7A1. Clin Exp Dermatol 2001; 26**:**97-9.

45. Cserhalmi-Friedman PB, Garzon MC, Guzman E, et al. Maternal germline mosaicism in dominant dystrophic epidermolysis bullosa. J Invest Dermatol 2001; 117**:**1327-8.

46. Happle R. Mosaicism in human skin. Understanding the patterns and mechanisms. Arch Dermatol 1993; 129**:**1460-1470.

47. Happle R. A rule concerning the segmental manifestation of autosomal dominant skin disorders. Review of clinical examples providing evidence for dichotomous types of severity. Arch Dermatol 1997; 133**:**1505-1509.

48. Paller AS. Expanding our concepts of mosaic disorders of skin. Arch Dermatol 2001; 137**:**1236-8.

49. Konig A, Happle R. Two cases of type 2 segmental manifestation in a family with cutaneous leiomyomatosis. Eur J Dermatol 2000; 10**:**590-592.

50. Agarwalla A, Thakur A, Jacob M, Joshi A, Garg VK, Agrawal S. Zosteriform and disseminated lesions in cutaneous leiomyoma. Acta Derm Venereol 2000; 80**:**446.

51. Konig A, Happle R. Type 2 segmental cutaneous leiomyomatosis. Acta Derm Venereol 2001; 81**:**383.

52. Freyschmidt-Paul P, Hoffmann R, Konig A, Happle R. Linear porokeratosis superimposed on disseminated superficial actinic porokeratosis: report of two cases exemplifying the concept of type 2 segmental manifestation of autosomal dominant skin disorders. J Am Acad Dermatol 1999; 41**:**644-7.

53. Dean G. Porphyria variegata. Acta Derm Venereol Suppl (Stockh) 1982; 100**:**81-5.

54. Pulkkinen L, Bullrich F, Czarnecki P, Weiss L, Uitto J. Maternal uniparental disomy of chromosome 1 with reduction to homozygosity of the LAMB3 locus in a patient with Herlitz junctional epidermolysis bullosa. Am J Hum Genet 1997; 61**:**611-9.

55. Takizawa Y, Pulkkinen L, Shimizu H, et al. Maternal uniparental meroisodisomy in the LAMB3 region of chromosome 1 results in lethal junctional epidermolysis bullosa. J Invest Dermatol 1998; 110**:**828-31.

56. Takizawa Y, Pulkkinen L, Chao SC, et al. Mutation report: complete paternal uniparental isodisomy of chromosome 1: a novel mechanism for Herlitz junctional epidermolysis bullosa. J Invest Dermatol 2000; 115**:**307-11.

57. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 1971; 68**:**820-3.

58. Hutchin TP, Lench NJ, Arbuzova S, Markham AF, Mueller RF. Maternally inherited hearing impairment in a family with the mitochondrial DNA A7445G mutation. Eur J Hum Genet 2001; 9**:**56-8.

59. Martin L, Toutain A, Guillen C, et al. Inherited palmoplantar keratoderma and sensorineural deafness associated with A7445G point mutation in the mitochondrial genome. Br J Dermatol 2000; 143**:**876-83.

60. Hatamochi A, Nakagawa S, Ueki H, Miyoshi K, Iuchi I. Diffuse palmoplantar keratoderma with deafness. Arch Dermatol 1982; 118**:**605-7.

61. Fischel-Ghodsian N, Prezant TR, Fournier P, Stewart IA, Maw M. Mitochondrial mutation associated with nonsyndromic deafness. Am J Otolaryngol 1995; 16**:**403-8.

62. Reid FM, Vernham GA, Jacobs HT. A novel mitochondrial point mutation in a maternal pedigree with sensorineural deafness. Hum Mutat 1994; 3**:**243-7.

63. Bleasel NR, Varigos GA. Porphyria cutanea tarda. Australas J Dermatol 2000; 41**:**197-206; quiz 207-8.

64. Spiewak R. Pesticides as a cause of occupational skin diseases in farmers. Ann Agric Environ Med 2001; 8:1-5.

65. Ladda RL. Principles of genetics. In: Alper JC, ed. Genetic disorders of the skin. Chicago: Year Book Medical Publishers, 1991.

66. Thomas NS, Chelly J, Zonana J, et al. Characterisation of molecular DNA rearrangements within the Xq12-q13.1 region, in three patients with X-linked hypohidrotic ectodermal dysplasia (EDA). Hum Mol Genet 1993; 2:1679-85.

67. Baumeister FA, Egger J, Schildhauer MT, Stengel-Rutkowski S. Ambras syndrome: delineation of a unique hypertrichosis universalis congenita and association with a balanced pericentric inversion (8) (p11.2; q22). Clin Genet 1993; 44:121-8.

68. Balducci R, Toscano V, Tedeschi B, et al. A new case of Ambras syndrome associated with a paracentric inversion (8) (q12; q22). Clin Genet 1998; 53:466-8.

69. Tadin M, Braverman E, Cianfarani S, et al. Complex cytogenetic rearrangement of chromosome 8q in a case of Ambras syndrome. Am J Med Genet 2001; 102:100-4.

70. Huber M, Rettler I, Bernasconi K, et al. Mutations of keratinocyte transglutaminase in lamellar ichthyosis. Science 1995; 267:525-8.

71. Russell LJ, DiGiovanna JJ, Rogers GR, et al. Mutations in the gene for transglutaminase 1 in autosomal recessive lamellar ichthyosis. Nat Genet 1995; 9:279-83.

72. Kere J, Srivastava AK, Montonen O, et al. X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein. Nat Genet 1996; 13:409-16.

73. Takahashi M, Rapley E, Biggs PJ, et al. Linkage and LOH studies in 19 cylindromatosis families show no evidence of genetic heterogeneity and refine the CYLD locus on chromosome 16q12-q13. Hum Genet 2000; 106**:**58-65.

74. Bignell GR, Warren W, Seal S, et al. Identification of the familial cylindromatosis tumour-suppressor gene. Nat Genet 2000; 25**:**160-5.

75. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409**:**860-921.

76. Lathrop GM, Lalouel JM, Julier C, Ott J. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci U S A 1984; 81**:**3443-6.

77. Craddock N, Dawson E, Burge S, et al. The gene for Darier's disease maps to chromosome 12q23-q24.1. Hum Mol Genet 1993; 2**:**1941-1943.

78. Bashir R, Munro CS, Mason S, Stephenson A, Rees JL, Strachan T. Localisation of a gene for Darier's disease. Hum Mol Genet 1993; 2**:**1937-1939.

79. Ikeda S, Welsh EA, Peluso AM, et al. Localization of the gene whose mutations underlie Hailey-Hailey disease to chromosome 3q. Hum Mol Genet 1994; 3**:**1147-1150.

80. Hu Z, Bonifas JM, Beech J, et al. Mutations in ATP2C1, encoding a calcium pump, cause Hailey-Hailey disease. Nat Genet 2000; 24**:**61-65.

81. Sudbrak R, Brown J, Dobson-Stone C, et al. Hailey-Hailey disease is caused by mutations in ATP2C1 encoding a novel Ca(2+) pump. Hum Mol Genet 2000; 9**:**1131-1140.

82. Ahmad W, ul Haque MF, Brancolini V, et al. Alopecia universalis associated with a mutation in the human hairless gene. Science 1998; 279**:**720-724.

83. Suzuki K, Hu D, Bustos T, et al. Mutations of PVRL1, encoding a cell-cell adhesion molecule/herpesvirus receptor, in cleft lip/palate-ectodermal dysplasia. Nat Genet 2000; 25**:**427-430.

84. Aldred MA. Cleft lip and palate: new genetic clues. Trends Mol Med 2001; 7**:**539-40.

85. Sozen MA, Suzuki K, Tolarova MM, Bustos T, Fernandez Iglesias JE, Spritz RA. Mutation of PVRL1 is associated with sporadic, non-syndromic cleft lip/palate in northern Venezuela. Nat Genet 2001; 29**:**141-2.

86. Norgett EE, Kelsell DP. SPINK5: both rare and common skin disease. Trends Mol Med 2002; 8**:**7.

87. Walley AJ, Chavanas S, Moffatt MF, et al. Gene polymorphism in Netherton and common atopic disease. Nat Genet 2001; 29**:**175-8.

Fig. 2 Examples of allelic heterogeneity in genodermatoses

Examples of allelic heterogeneity in genodermatoses		
Gene[a]	**Phenotype**[b]	**Reference**
GJB3	Erythrokeratodermia variabilis	5
	AD nonsyndromic sensorineural deafness	6
	AR nonsyndromic hearing loss	7
	AD deafness with peripheral neuropathy	8
Desmoplakin	Dermatosis palmoplantaris striata	9
	Dilated cardiomyopathy with woolly hair and keratoderma	10
Plectin	EBS and limb-girdle muscular dystrophy	11,12
	EBS of the Ogna type	13

[a] *GJB3*: gene encoding connexin 30.3
[b] AD: autosomal dominant; AR: autosomal recessive; EBS: epidermolysis bullosa simplex

Fig. 3 Examples of genetic or locus heterogeneity in genodermatoses

Examples of genetic or locus heterogeneity in genodermatoses		
Phenotype	**Gene**[a, b]	**Reference**
EBS Dowling-Meara variant	*KRT5*	[14]
	KRT14	[15]
EBS Koebner variant	KRT5	[16]
	KRT14	[17]
EBS Weber-Cockayne variant	*KRT5*	[18]
	KRT14	[19]
Hermansky-Pudlak syndrome	HPS1	[20,21]
	AP3B1	[22]
	HPS3	[23]
	HPS4	[24]

[a] *AP3B1*: gene for beta-3A subunit of AB3; *KRT5*: keratin 5 gene; *KRT14*: keratin 14 gene
[b] *HPS1*, *HPS3* and *HPS4*: genetic loci or genes for Hermansky-Pudlak syndrome.

Fig. 1 Example of a polymorphic locus. Locus *A* represents a polymorphic locus with three different alleles. The example shows a family in which the father is homozygous for allele 1 and the mother is heterozygous for alleles 2 and 3. The son of this family has inherited alleles 1 from the father and 3 from the mother, and is, thus, heterozygous (2,3).

Fig. 2 Examples of allelic heterogeneity in genodermatoses

Fig. 3 Examples of genetic or locus heterogeneity in genodermatoses

Fig. 4 Examples of Mendelian patterns of inheritance. ***A*** Autosomal dominant epidermolysis bullosa simplex superficialis (modified from ref. [29]); ***B*** Autosomal recessive woolly hair, palmoplantar keratoderma and heart disease (modified from ref. [30]); ***C*** X-linked recessive immune deficiency and hypohidrotic ectodermal dysplasia (modified from ref. [31]); ***D*** X-linked dominant incontinentia pigmenti (modified from ref. [32]); ***E*** Y-linked retinitis pigmentosa (modified from ref. [33]). The symbols commonly used for pedigree drawing are indicated at the bottom of the figure.

Fig. 5 Examples of complications to the basic Mendelian patterns of inheritance. ***A*** Reduced penetrance in a pedigree with autosomal dominant retinitis pigmentosa (modified from ref. [38]); ***B*** High degree of consanguinity in a family with hypotrichosis, with affected family members in several generations; ***C*** Pseudodominant inheritance in autosomal recessive atrichia with papular lesions (APL). The inheritance of APL in this pedigree mimics an autosomal dominant pattern of inheritance resulting from the mating

between a patient homozygous for the causative genetic mutation (II:2) and an unaffected carrier (modified from ref. [39]).

Fig. 6 Examples of haplotypes for a chromosomal region containing four polymorphic loci, *A*, *B*, *C* and *D*.

Fig. 7 Examples of different types of polymorphisms. ***SNP***s (***A***, ***B***) consist of single base substitutions. Panel ***A*** shows a point substitution (underlined) replacing a cytosine, C, for a thymidine, T, identified by DNA sequencing. A polymorphism that affects the recognition site for a restriction enzyme is referred to as an RFLP (Restriction Fragment Length Polymorphism). In this particular example, the sequence *ggatcc* is the target site for the restriction enzyme *BamH*I. In the presence of this site, the enzyme will cut the DNA into two fragments, detected as two bands in an agarose gel; ***C*** Example of a microsatellite. The repeat unit of the microsatellite shown here is the dinucleotide (CA). The number of tandem repeats of the (CA) unit can vary in each allele. The figure shows two individuals heterozygous for this dinucleotide. The difference in length between these three alleles appears as bands of DNA of different sizes in a polyacrylamide gel.

Fig. 8 Example of two-point LOD score values for markers on chromosome 1. Those values greater than 3 (in red) indicate linkage between the phenotype and the marker. The negative values for markers D1S209 and D1S2798 at a recombination fraction of 0 indicate that genetic recombination has been observed between the phenotype and the marker.

Fig. 9 Example of haplotype analysis for the refinement of a linkage interval. The disease associated haplotype, 4 4 1 1 5 2 (shaded in blue) has been broken by genetic recombination in the family members (stars). The two key recombinant individuals, III:8 and IV:1, have allowed the narrowing of the disease interval.

Fig. 10 Pedigrees with alopecia areata as an example of multifactorial dermatological disease. Although some of the pedigrees individually can mimic Mendelian pattern of inheritance (***C*** and ***D***), others show a more complex pattern (***A***), in which the disease is inherited through two unrelated branches of the family. A single mode of inheritance cannot be established for the disease as a group.

Fig. 1

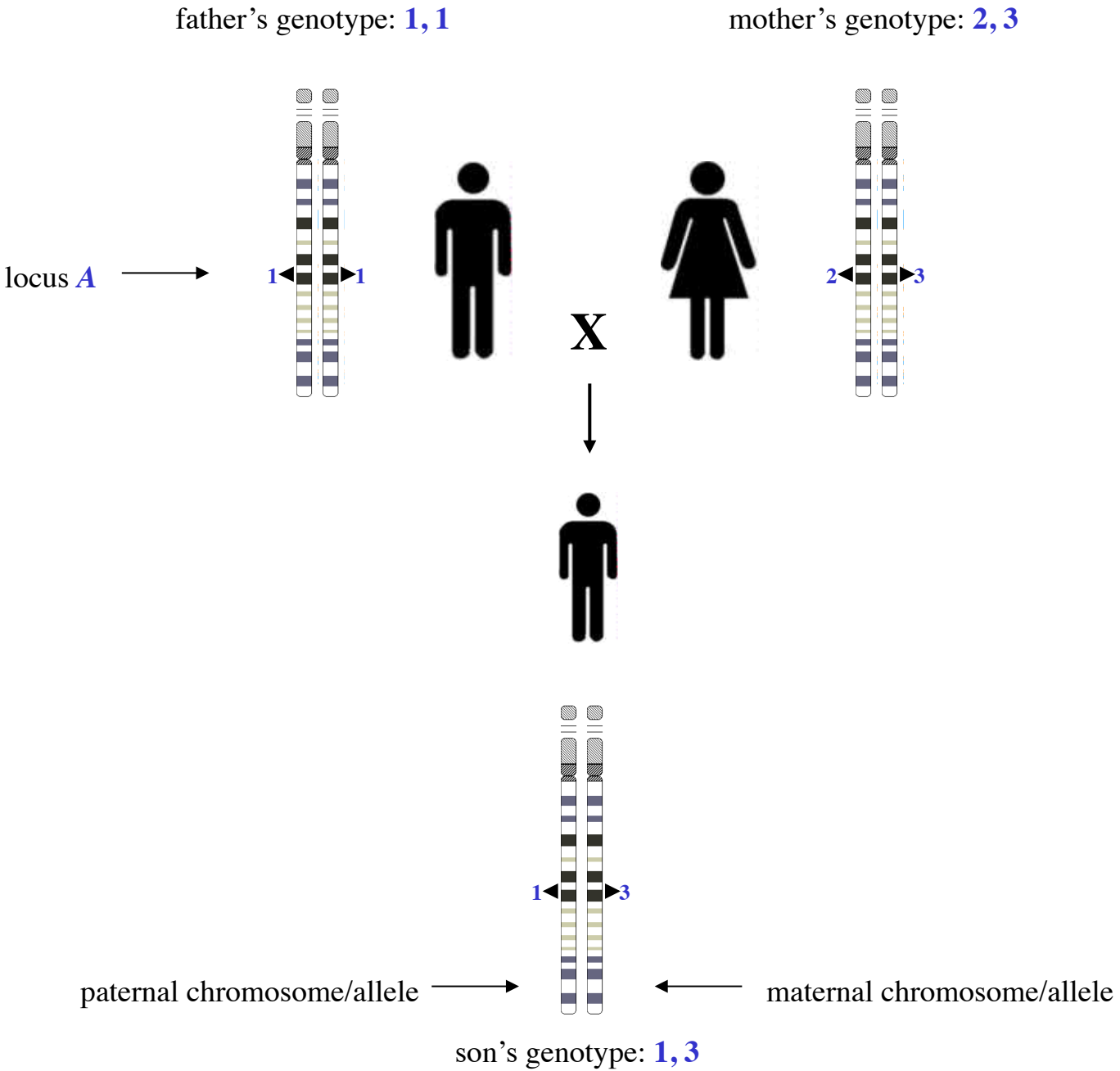

father's genotype: 1, 1
mother's genotype: 2, 3
locus A
1
1
X
2
3
1
3
paternal chromosome/allele
maternal chromosome/allele
son's genotype: 1, 3

Fig. 4

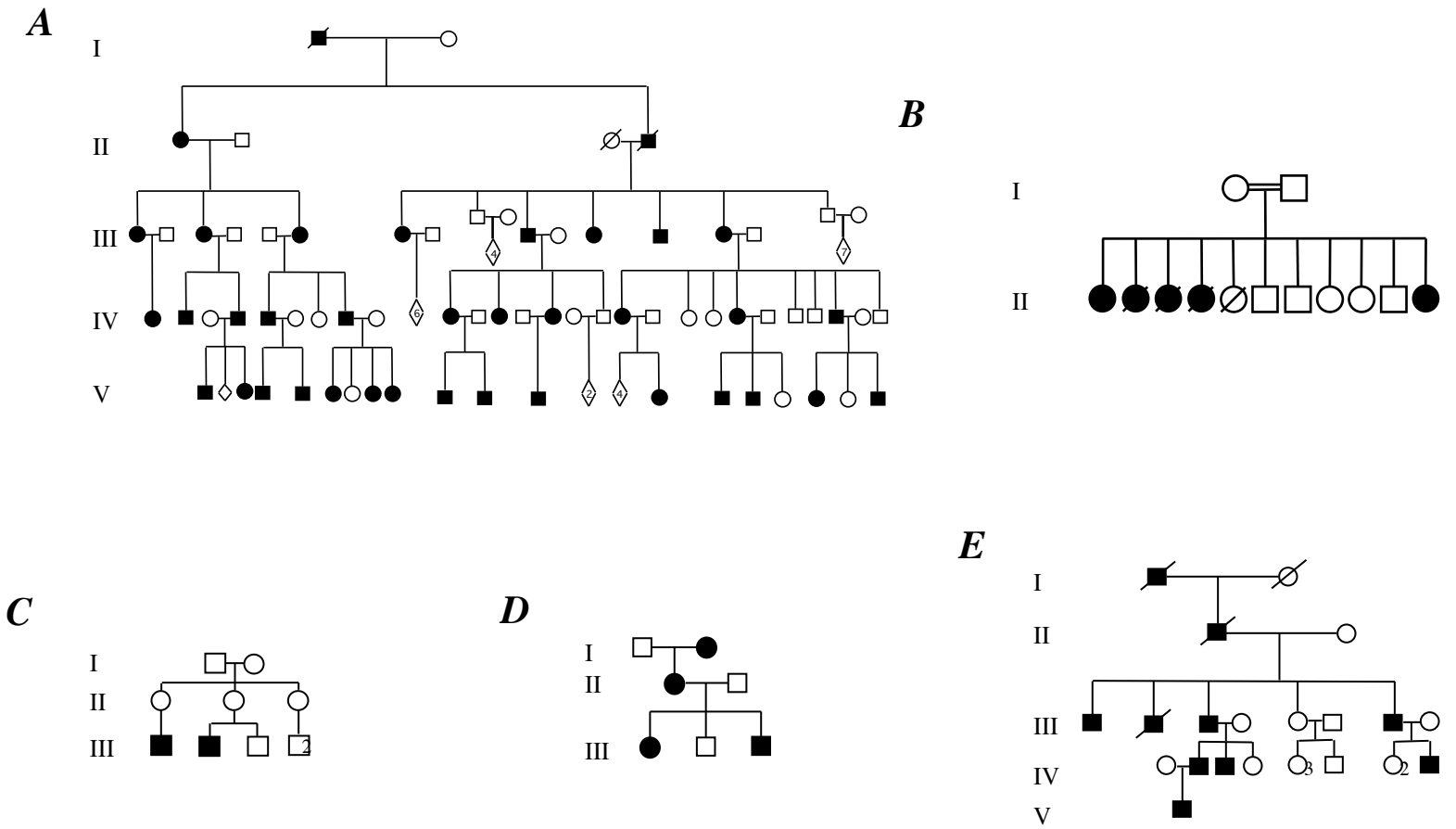
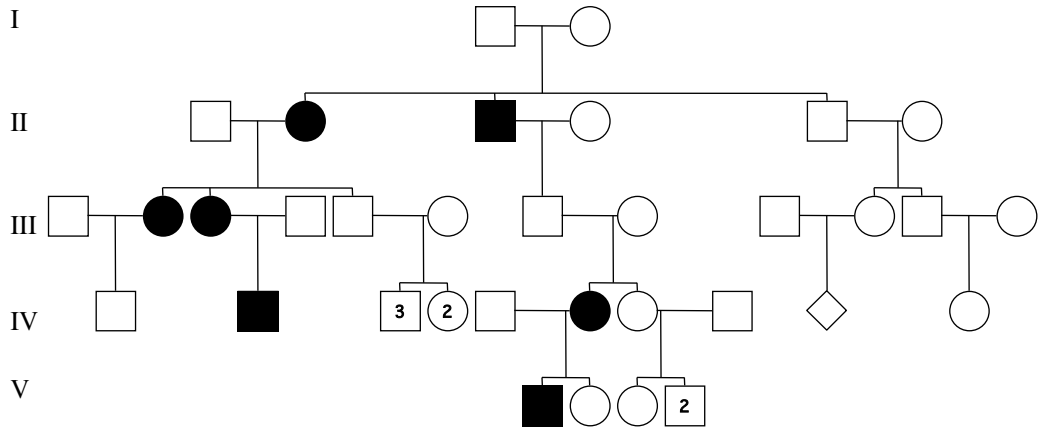
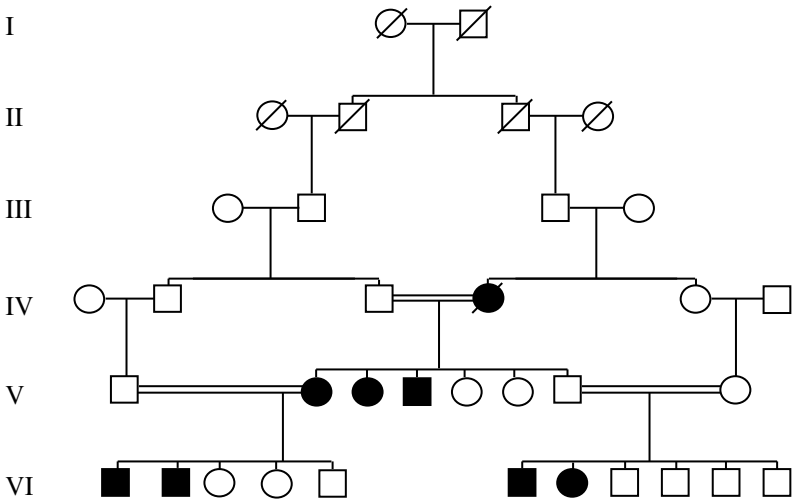

A
I
II
III
IV
V
B
I
II
C
I
II
III
D
I
II
III
E
I
II
III
IV
V
unaffected male/female
affected male/female
deceased individual
obligate female carrier
for an X-linked recessive trait
unspecified sex
number of male, female
or unspecified sex individuals
I, II ... generation
proband
consanguineous mating

Fig. 5

A



B



C

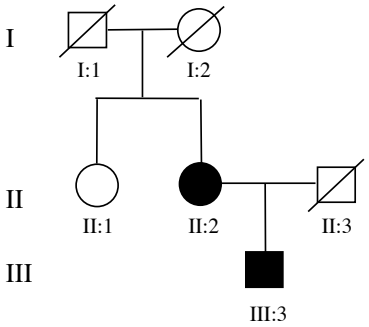


Fig. 6

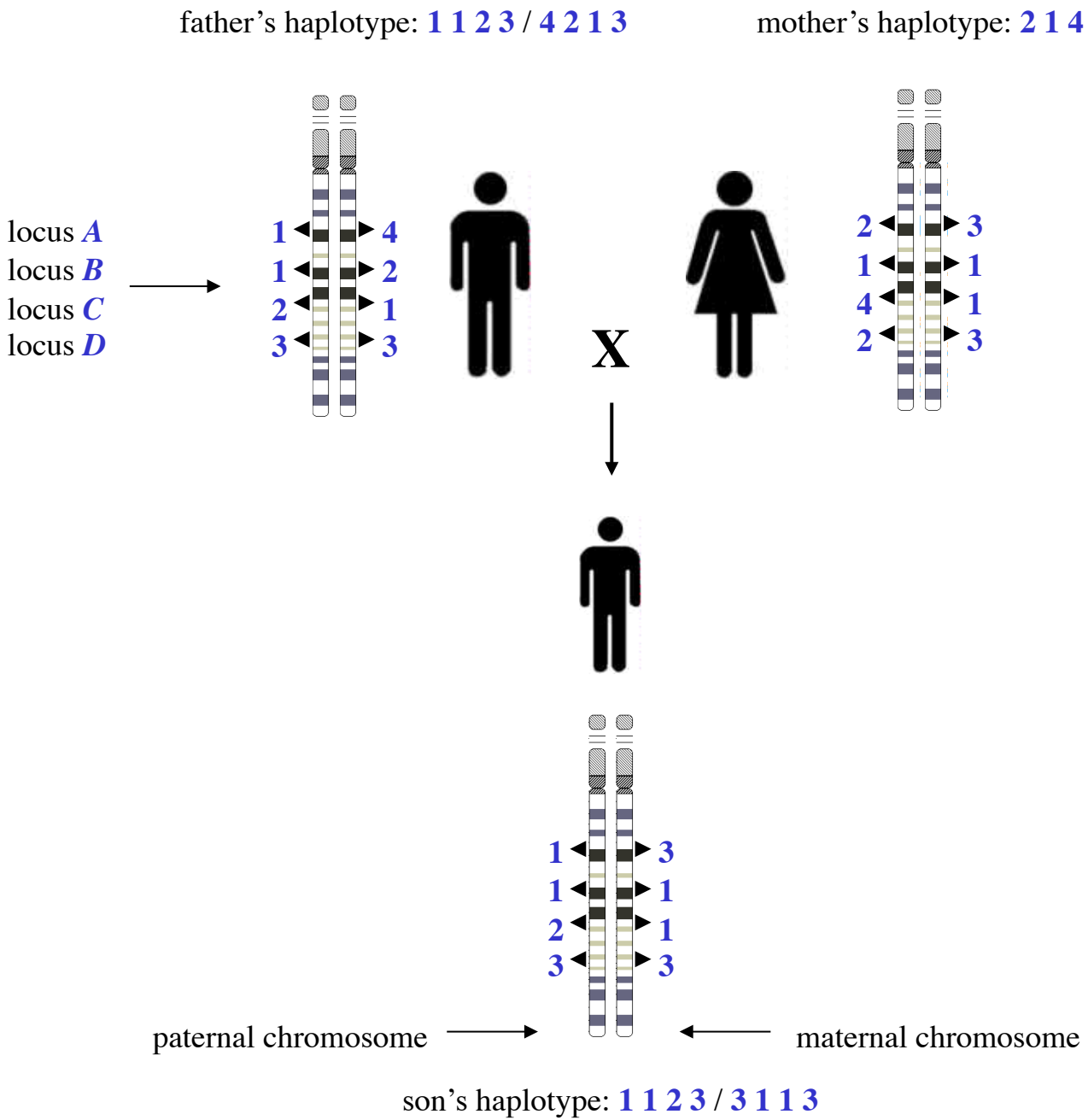
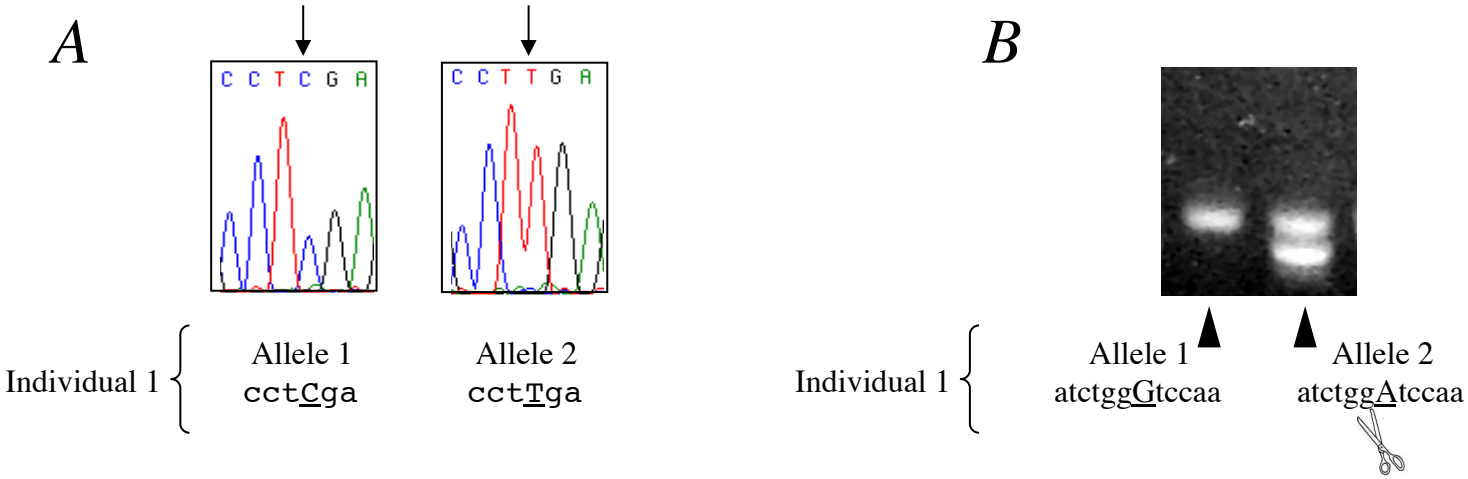

father's haplotype: 1 1 2 3 / 4 2 1 3
mother's haplotype: 2 1 4 2 / 3 1 1 3
locus A
locus B
locus C
locus D
1 1 2 3
4 2 1 3
X
2 1 4 2
3 1 1 3
1 1 2 3
3 1 1 3
paternal chromosome
maternal chromosome
son's haplotype: 1 1 2 3 / 3 1 1 3

Fig. 7

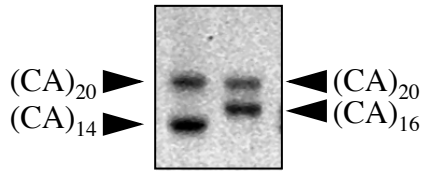


C

Individual 1

Allele 1 ttccccCACACACACACACACACACACACACACACACACACACACAgatgggc $(CA)_{20}$

Allele 2 ttccccCACACACACACACACACACACACACACACAgatgggc $(CA)_{14}$



Individual 2

Allele 1 ttccccCACACACACACACACACACACACACACACACACACACACAgatgggc $(CA)_{16}$

Allele 2 ttccccCACACACACACACACACACACACACACACACACACACACAgatgggc $(CA)_{20}$

Fig. 8

LOD score values at recombination fraction (θ)

	.00	.01	.05	.1
D1S209	-15.34	2.20	2.62	2.55
D1S2835	**4.46**	**4.30**	**3.99**	**3.50**
D1S198	**3.29**	**3.01**	2.68	1.85
D1S2798	-17.55	2.42	2.96	2.83

Fig. 9

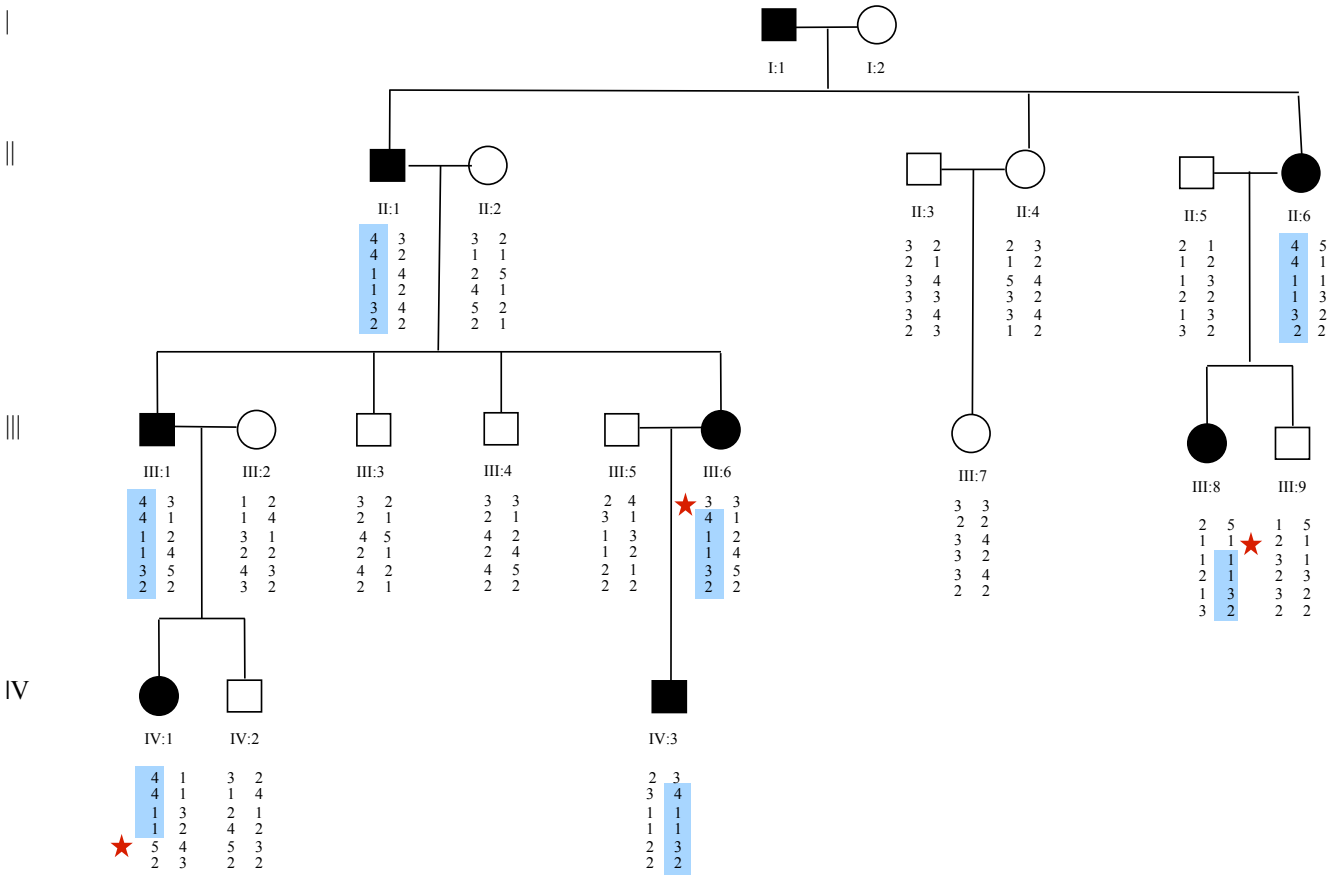
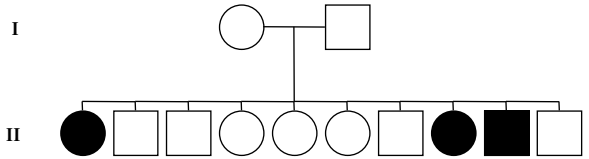
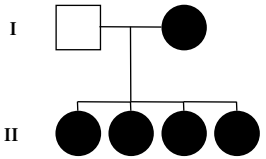

I
II
III
IV
I:1
I:2
II:1
II:2
II:3
II:4
II:5
II:6
III:1
III:2
III:3
III:4
III:5
III:6
III:7
III:8
III:9
IV:1
IV:2
IV:3

Fig. 10

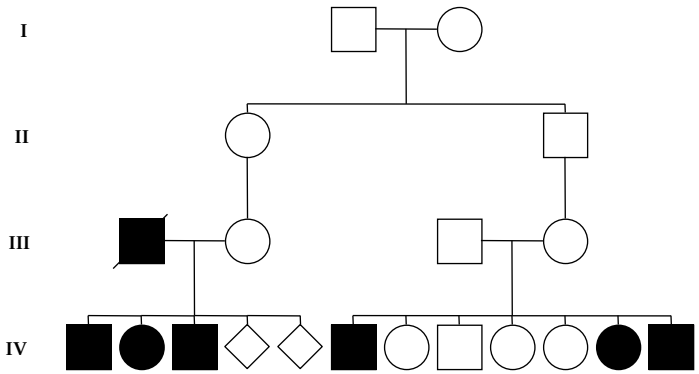
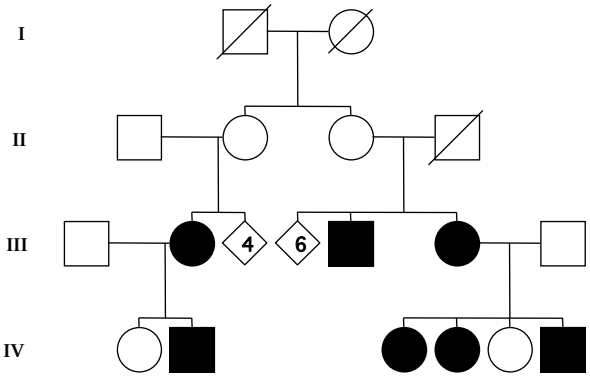
A



B



C



D