



## **A framework for describing genetic diseases**

Draft 7/8/02

Boston Cure Project, Inc.

The causes of multiple sclerosis, like the causes of many other diseases, are unknown. Both genetic and environmental factors likely interact over time to produce the disease. This belief is currently best supported by twin studies showing that an identical twin of a person diagnosed with MS is 30 times more likely than a non-identical twin to also develop MS. In other words, the concordance rate for developing MS is about 30x higher for identical twins who possess near identical genetic backgrounds than for non-identical twins who on average have half of their genes in common. Yet even for identical twins the concordance rate is not 100%. Therefore, while a person's genetic makeup substantially affects his/her relative risk, genetic susceptibility on its own is not sufficient to develop the disease.

Identification of the gene or genes that increase the susceptibility to developing MS will be extremely useful for defining relevant molecular and cellular pathways involved in the development of the disease (pathogenesis) that then serve as targets for potential therapeutics.

Our approach to determining the genetic factor(s) for MS is to first organize our knowledge about genetic diseases in general by grouping diseases into meaningful classes, and then determine what evidence there is for placing or not placing MS in these classes. Where there is some evidence for linking MS with a class, we can propose experiments that would rule it either in or out, based on strategies that have been previously successful. Where conclusive evidence exists for excluding MS from a class, we can cease our efforts in exploring that class. We will perform similar analyses for other possible causal factors for MS (pathogens, nutrition, toxins or trauma), combining the results into a cure map that will guide development of a research strategy to find the causes of MS. It is our intention that these efforts will be generically applicable to other diseases for which the causes are also not yet known.

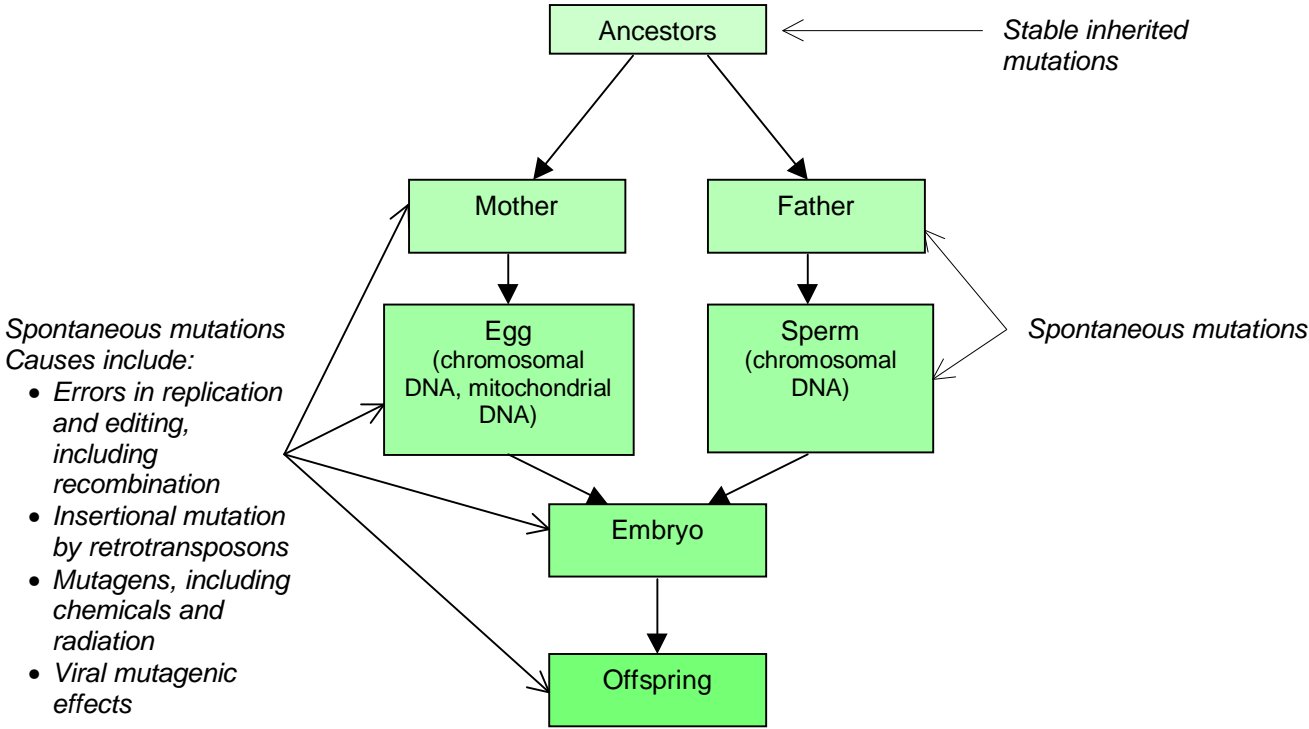
Our ultimate goal is an understanding of what factor(s) are responsible for MS, how people acquire these factors, and how they produce the disorder. With that understanding, we will have a much higher chance of finding a cure which prevents the disability caused by MS, permanently stops the progress of MS, and/or repairs the damage already caused in patients with MS.

The diagrams on the next two pages illustrate our framework for thinking about genetic diseases. They describe a genetic disease in terms of the discrete players and events that interact to cause a disorder. The arrows show causal relationships that exist between the entities (e.g., the specific genetic mutation will determine the defect in the protein produced by that gene).

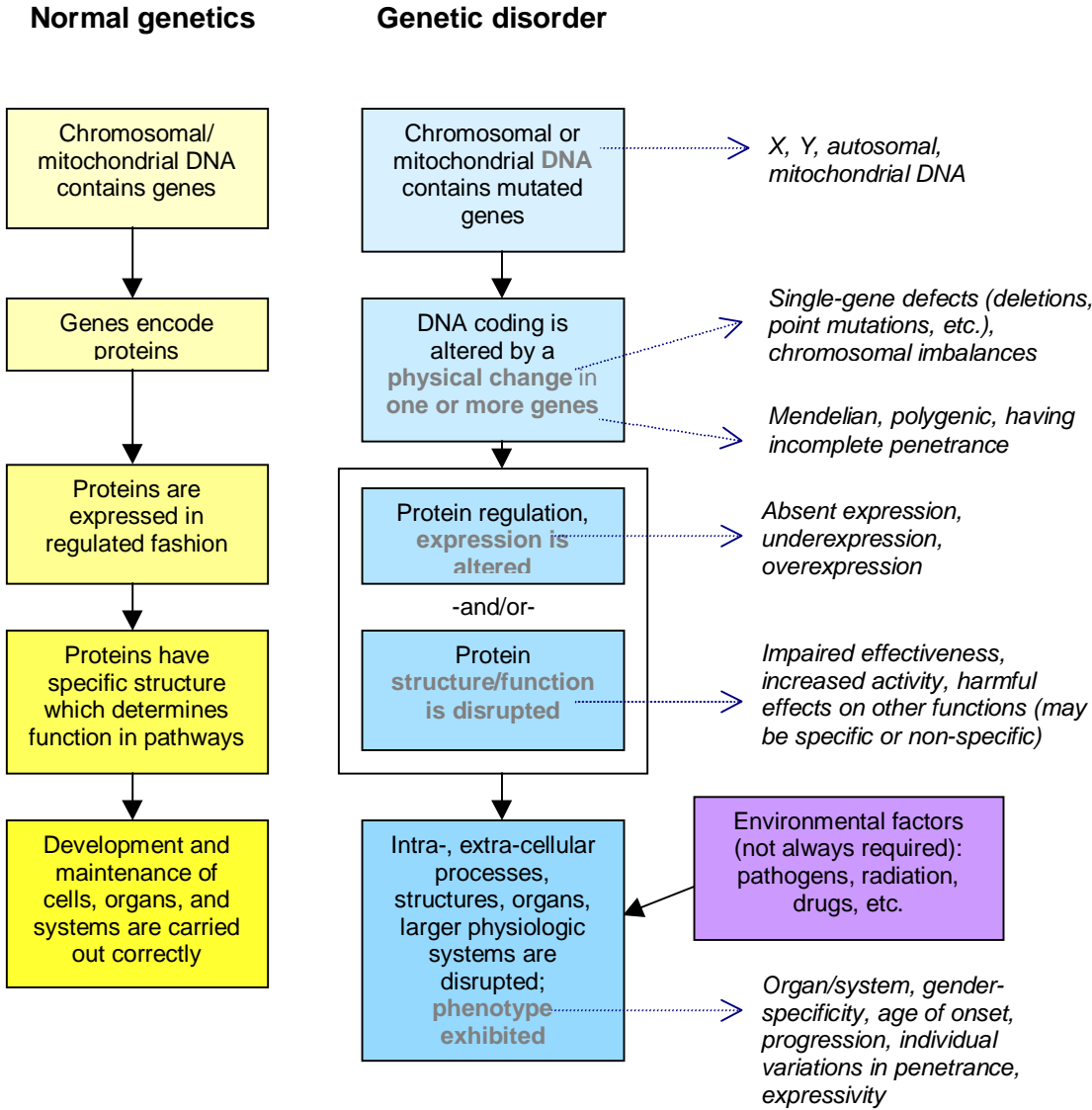
This component-oriented model fits well with the modular way diseases are usually investigated and the research techniques that are used – scientists investigate one or two of these aspects of a disease at a time, such as the location of the gene, inheritance patterns, the type of protein involved, or disease progression. Therefore this framework should be useful for analyzing and integrating individual research data points as they become available.

The entities in the diagrams constitute a classification system for describing the important genetic attributes of a disease. The classes and subclasses into which genetic diseases can be grouped are described in the pages following the diagrams. Common characteristics of the diseases in each subclass are listed, where possible; these characteristics may help indicate whether a disease whose cause is unknown could belong to the subclass.

# Origins of genetic mutations



# Factors involved in genetic disorders



## Classes of genetic diseases

As shown in the preceding diagrams, there are a number of different possible pathways and factors involved in producing a genetic disorder. Understanding which pathways and factors are involved in a particular disease is the goal of genetic research.

When studying a disease that has a genetic component, we want to learn:

- The origins of the genetic mutation
- The location of the genetic defect
- How many genes are involved
- What, if any, environmental factors are necessary for the genetic defect to manifest itself
- The exact physical change that has occurred in the gene
- How protein regulation or expression is altered, and/or how the protein structure is changed, interfering with its normal function or other cell functions
- Phenotypic characteristics of the disease

The remainder of this document discusses each of these elements in more detail. For each element, a set of classes is described that can be used to categorize diseases and possibly predict the nature of a disease of interest.

### Origins of genetic mutations

**Genetic mutations can either be inherited from generation to generation within a family, or they can occur spontaneously in an individual**

In investigating the genetic basis for a disease, we are often interested in knowing how a person acquires the mutation – is it generally passed down through parents and ancestors or is it more often due to a new mutation? This information is useful in predicting whether other family members may be affected, in conducting research to track down the gene(s) involved, and in understanding how mutations arise and exert their effect.

Most disorders with a known genetic cause have been studied extensively in terms of their inheritance patterns and rate of new mutations. For some disorders, most affected people have inherited the mutation from one or both parents; for other disorders, most cases are the result of new (spontaneous) mutations introduced through gametogenesis, during embryonic development, or in the course of a person's life. Each of these possible acquisition methods is discussed in the following paragraphs.

**Stably inherited mutations are passed down through multiple generations in families**

In an inherited disorder, an affected person has received a mutated gene from one or both parents, who in turn likely received the mutation from one of their parents, and so on. Inherited disorders may present inheritance patterns that provide clear clues that the

disease has a genetic component. Inheritance is investigated by mapping pedigrees showing the affected and unaffected people within a family; these pedigrees also provide some clues as to which genes and chromosomes may be involved and the strength of the mutation's effect (penetrance).

In order to be inherited, a genetic mutation must not destroy a person's ability to reproduce. If it did, there could be no offspring to inherit the mutation. Therefore inherited mutations must either be recessive (requiring two mutated genes to take effect) or not overly harmful to a person's physical or social reproductive fitness.

Tay-Sachs disease, MERRF (myoclonus epilepsy with ragged-red fibers) and Charcot-Marie-Tooth syndrome type X are all examples of disorders that are typically inherited.

### **Spontaneous mutations occur in individuals with no prior family history**

New mutations in a gene or chromosome can occur at a number of points during an individual's development – even prior to conception. The mutation can happen during germ line establishment and gametogenesis (the process of creating sperm and eggs) in either parent or as a result of mutations to somatic (non-gamete) cells occurring during embryonic development, childhood or adulthood.

Causes of spontaneous mutations include:

- **Errors in DNA replication and editing** introduced during mitotic or meiotic replication, meiotic recombination, or segregation during cell division. Related epigenetic processes such as the imprinting of genes during gametogenesis can also be subject to errors which create effects similar to alterations in the DNA itself. DNA replication and editing errors may be induced by mutagens, including chemicals and radiation, that alter the structure or sequence of DNA through a number of different mechanisms. For example, ionizing radiation (X-rays and gamma rays) can create free radicals of water which react chemically with DNA, causing chromosome breaks, genetic mutations, and other types of damage.
- **Viruses that have mutagenic effects on their hosts.** For example, hepatitis B virus (HBV) DNA is capable of being inserted into host DNA; this activity may contribute to the development of hepatocellular carcinoma, possibly by disrupting the regulation of human growth regulatory genes or by causing host sequence deletions, chromosomal translocations and other abnormalities. It has also been suggested that certain viruses (such as human papilloma virus and Epstein-Barr) may somehow induce genetic instability in the host genome and thus contribute to carcinogenic risk.
- **Retrotransposons.** The human genome itself also contains elements called retrotransposons that can be replicated and inserted into new locations in the genome. While this mechanism is not known to be a frequent cause of disease, it has been cited as operative in small numbers of cases in diseases such as hemophilia A, chronic granulomatous disease, and colon cancer.

Some disorders result primarily from new mutations. This is usually the case for those which affect reproductive ability in a dominant fashion. Even for those disorders which are predominantly inherited, some instances of spontaneous mutations are usually seen.

Presumably all inherited defects began as spontaneous mutations in the founding individual.

### **Errors in gamete production can result in new mutations**

Spontaneous mutations in gametes are usually the result of genetic errors (point mutations, deletions, etc.) introduced into DNA during gametogenesis. For some diseases, spontaneous mutation during gametogenesis is the most common acquisition pathway. For instance, in achondroplasia, 80% of all cases are due to new mutations in gametes (usually point mutations, and usually in sperm from fathers age 35 and older). The predominance of spontaneous mutation in achondroplasia stems from the fact that this disease inhibits reproductive success and that the gene involved has a mutation rate higher than the average rate for most human genes. Prader-Willi and Angelman syndromes are also examples of disorders that are primarily caused by new genetic mutations in gametes. As with achondroplasia, impaired reproductive fitness prevents these diseases from being stably inherited. Most individuals with Prader-Willi do not reproduce, and only one case of reproduction has been reported for Angelman.

Special cases of defects introduced during gametogenesis include:

- **Triplet repeat expansion:** Some disorders are caused by repeated insertion of an existing group of three nucleotides into a gene during gametogenesis. In genes subject to repeat expansion, there is a “normal” range for number of trinucleotides which produces no disorder and a “symptomatic” range which does result in the disorder. There may also be an intermediate “premutation” range which does not result in the disorder but which may give rise to further repeat expansion during production of gametes, so that the carrier of a premutation gene is more likely to produce offspring with the full mutation. Triplet repeat expansion disorders are prone to anticipation, a phenomenon in which offspring of affected parents experience a more severe phenotype or earlier age of onset than the parent because the triplet repeat region expands further with each successive generation. Triplet repeat expansion disorders also sometimes exhibit anticipation characteristics that are specific to the gender of the parent transmitting the mutation. If the mutation tends to be more unstable during spermatogenesis than oogenesis, for example, then children of affected fathers will more likely be subject to anticipation than children of affected mothers. Examples of triplet repeat expansion disorders include Huntington’s disease (repeated CAG triplets with a normal range of 10-35 and abnormal range of 36-121) and fragile X syndrome (CGG repeats with normal range of 6-54, premutation range of ~55-230, and full mutation range of 231+).
- **Imprinting:** Imprinting is a process affecting a small number of genes whereby either the maternal or paternal copy of the gene is chemically inactivated during gametogenesis. Normally the other parent’s copy of the gene remains activated and is sufficient for proper functioning. However, if something happens to the parental gene that normally would be activated (e.g., it suffers a deletion), then no gene product is expressed. Likewise, it is thought that erroneous activation of a normally inactivated gene may lead to overexpression, also resulting in disease. If the imprinting process itself is not carried out properly, then genes are incorrectly activated or inactivated, potentially leading to a disorder. Faulty imprinting is one of the known causes of Prader-Willi syndrome.

## **Somatic mutation can occur in germline development, embryonic development, or later in life**

Somatic mutations are spontaneous genetic mutations that occur in cells other than gametes. Any post-zygotic mutation (one that occurs after development of the zygote) results in an individual who is mosaic; that is, they possess cell lines which are genetically distinct. Mosaicism may exist in germ cells or in somatic cells. Since the average rate of mutation for humans is estimated to be about  $1 \times 10^{-7}$  per nuclear gene per cell generation, and since humans contain about  $1 \times 10^{13}$  cells, in a sense we are all mosaics to some degree and our DNA keeps acquiring new mutations throughout our lifetime.

Somatic mutations are usually harmless because they typically either have no effect on the cell or cause the cell to die without affecting any larger systems. In fact, some somatic mutations are normal and helpful. For instance, B cells responsible for antibody production undergo a continual process of somatic mutation to fine-tune the body's immune response to pathogens; this is a highly beneficial process as long as the body's ability to eliminate or inactivate autoreactive B cells functions correctly (otherwise autoimmunity may develop).

Post-zygotic mutations that are harmful lead to disease under three circumstances:

- **When they occur in germ cells**, which does not affect the person in which the mutation originated but could affect any offspring he or she may have. In addition, a precursor germ cell involved in the earlier stages of gametogenesis can acquire a mutation that is passed on to all of its successor gametes, resulting in germline mosaicism. When germline mosaicism (two or more genetically different germ cell lines) exists, some but not all of the parent's gametes have the mutation, and therefore the mutation can be passed on to more than one child. For instance, of the 20-30% of Duchenne muscular dystrophy cases that are the result of new mutations, 10-20% arise from germline mosaics.
- **When they occur in early development**, leading to a large population of mutant cells ultimately contributing to the formation of the individual. Genetic mosaicism arising in early development might be suspected when the affected person has no family history of the disease and only partial effects are observed or only part(s) of the body are affected. Mutation in early development seems to be the predominant pathway for McCune-Albright syndrome. It also occurs occasionally in some other disorders; for instance, it is thought to represent up to 10% of cases of Alagille syndrome. Also, in the triplet repeat expansion disorder myotonic dystrophy, mitotic instability in the triplet repeat area of the gene can result in mosaicism because some cells undergo increased repeat expansions during embryonic development (and even in post-natal development).
- **When the mutation causes unrestricted cell division**, in which case a relatively small number of initial mutated cells can result in disease. For example, when genes that normally suppress tumor formation or trigger cell proliferation become mutated, cells are allowed to reproduce unchecked, resulting in tumors (cancer or neoplasia). Note: not all cancers are purely due to somatic mutations; some are the direct result of inherited cancerous mutations or inherited predispositions to cancer. In some cases, an inherited mutation that predisposes to cancer may require a subsequent additional somatic mutation (a "second hit") in order for a cancer to occur. Hereditary forms of cancer are

suspected when ratios of affected family members to unaffected family members are relatively high; hereditary forms of cancer can also sometimes have an earlier average age of onset than their non-hereditary forms.

## Location of the genetic defect

### **Genetic defects can be found on autosomal chromosomes, on X or Y chromosomes, or in mitochondrial DNA**

One of the most critical pieces of information in understanding any genetic disease is the location of the genetic defect. With this knowledge, we can create tests to determine who has the disorder or is at risk of having it, and can also begin to investigate the disease pathway.

Genes are found on each of the 22 autosomal chromosomes, on the X and Y chromosomes, and in mitochondrial DNA.

### **With a few exceptions, the autosomal chromosomes each contain functionally unrelated genes and are alike with respect to mutation**

For the most part, each autosomal chromosome contains an assortment of generally unrelated genes, and each is generally subject to the same types of mutations. However, there are a few differences between the autosomal chromosomes that are relevant to human genetic disorders. One distinction is the existence of gene families clustered within one region of a chromosome. For instance, the human leukocyte antigen (HLA) gene complex on chromosome 6 is associated with predisposition to several different diseases, especially inflammatory diseases. Another gene family, the growth hormone gene cluster, is localized on 17q22-24. Also, certain chromosomes are subject to particular kinds of chromosomal imbalances, which will be discussed in the section covering the physical nature of defects. Finally, a few chromosomes (such as 7, 11, 14, and 15) are known to contain imprinted genes. An important characteristic of disorders involving imprinted genes is that their expression depends on the gender of the parent from whom the gene is inherited.

In general, disorders due to defects in autosomal genes do not show a bias in disease frequency between genders. That is, recessive disorders occur with a frequency of 25% in each sex if both parents are carriers of the genetic defect, and dominant disorders occur with a frequency of 50% if a single parent is affected (assuming the genetic defect has 100% penetrance). Examples of disorders caused by mutated genes on autosomal chromosomes include cystic fibrosis (chromosome 7) and Canavan disease (chromosome 17).

### **X-linked diseases are identified by specific inheritance patterns**

Because males inherit one copy of the X chromosome (from their mother) and females inherit two copies (one from each parent), X-linked diseases exhibit certain inheritance patterns which provide clues as to the location of the defect.

- Males with a genetic defect on their X chromosome always pass the mutation on to their daughters, but never to their sons.
- X-linked diseases which show recessive inheritance patterns in females are always dominant in males, because males do not have a second X chromosome to make up for the mutated gene. Therefore, they affect more men than women. It is unusual but still possible for women who are heterozygous carriers of a mutated recessive gene to manifest the disease. For instance, if the X-inactivation process in a female carrier results in the preferential inactivation of the normal copy of the gene, then the majority of the cells will express the mutated gene and the disease may be expressed to some degree. There may be additional reasons why female carriers sometimes manifest X-linked recessive diseases. Lesch-Nyhan syndrome and Duchenne muscular dystrophy are two examples of X-linked recessive diseases.
- X-linked dominant diseases in general affect twice as many women as men, because women have two X chromosomes and therefore have twice the chance of acquiring the disease. X-linked hypophosphatemia is an example of an X-linked dominant disease.

### **Y-linked diseases affect only males and are rare**

Because males have Y chromosomes and females do not, Y-linked diseases will be seen only in males. There are very few examples of Y-linked diseases; some affect male fertility, such as mutations in Azoospermia factor c.

Note that the Y chromosome also contains non-sex-related genes, called pseudoautosomal genes, that are homologous to genes on the X chromosome. These genes are contained on the tips of the chromosomes in the pseudoautosomal regions. Mutations in either of the homologs can result in disorders; but again, these are relatively rare because these regions are small.

### **Mitochondrial DNA disorders are usually transmitted by maternal inheritance and present variable phenotypes**

Mitochondria are organelles that serve as the site of ATP synthesis (the respiratory chain) as well as other activities. Like the cell's nucleus, mitochondria contain strands of DNA which contain genes which express proteins. Mitochondrial DNA is slightly different from nuclear DNA, being arranged in a ring and containing no introns.

Mitochondrial diseases exhibit certain patterns. For instance, inherited mitochondrial mutations seem to be only transmitted through the maternal line because mitochondria in sperm for some reason do not contribute significantly to the zygote. Also, inherited mitochondrial disorders can have highly variable phenotypes, even within families, in terms of types of symptoms, symptom severity, age of onset, and tissue distribution. This is partly due to the presence in maternal egg cells of both mutated and normal mitochondrial genotypes (heteroplasmy). Offspring who get higher ratios of the mutated DNA will have a more severe and/or widespread version of the disease. Other genetic and environmental factors may also play a role in phenotype variation. Finally, mitochondrial disorders may present elevated levels of lactate and pyruvate. When a disorder causes a deficiency in the respiratory chain, these metabolites, normally processed by respiratory enzymes, instead build up in the cells.

Examples of inherited mitochondrial DNA disorders include myoclonus epilepsy with ragged-red fibers (MERRF), Leber's hereditary optic neuropathy (LHON), and mitochondrial encephalopathy with acidosis and stroke-like episodes (MELAS). Mitochondrial DNA disorders that typically result from spontaneous mutations include Kearns–Sayre syndrome and Pearson's syndrome.

Note: Many of the proteins used in the mitochondria are encoded by nuclear genes. Thus some mitochondrial genetic diseases are caused by defects in nuclear DNA. Therefore the term “mitochondrial disease” doesn't necessarily mean a disorder caused by mutation of mitochondrial DNA. This distinction can be blurred at times. For example, a mutation in a nuclear gene required for maintaining mitochondrial DNA integrity can result in a dominant or recessive inherited predisposition to acquiring multiple mitochondrial deletions.

## Physical nature of the defect

**Genetic mutations range from those affecting individual genes to those affecting entire chromosomes**

In addition to knowing the site of the defect for a disorder (a region of a particular chromosome or mitochondria), it is necessary to know the exact nature of the defect: what is missing, added or altered. Having this information lets researchers develop tests for diagnosing the disorder. It is also sometimes possible to draw correlations between defect and phenotype which may be helpful in proposing possible defect types for disorders being investigated.

Some defects are localized to a single gene or a small number of genes; others affect larger chromosomal segments or even entire chromosomes. Both of these classes are discussed below.

**Defects involving individual genes vary in size, type, effect, and strength of phenotypic correlations**

Defects affecting individual genes can be located in exons, introns, splice sites, and poly(A) sites; they can also occur in non-coding regulatory (promoter) regions. Mutations can affect either the amino acid sequence of the protein or the level of expression of the protein. The different physical types of defects include:

- **Point mutations**, where one of the nucleotides in the gene is replaced with a different nucleotide. Point mutations that affect the coding region of a gene are classified as synonymous (silent), missense, nonsense, or splice site mutations depending on whether and how the nucleotide change alters the codon and resulting protein product.
  - Synonymous, or silent, mutations are so-called because the resulting altered codon specifies the same amino acid as the original codon. They generally have no effect on the resulting protein unless they happen to

- induce another type of sequence change (e.g., the coincidental creation of a new splice site sequence).
- Missense point mutations cause an incorrect amino acid to be substituted in the protein.
  - Nonsense point mutations replace an amino acid codon with a stop codon, causing premature termination of transcription and a truncated protein. The reverse situation can also occur (this is sometimes called a “sense” point mutation): a termination codon is changed into one that codes for an amino acid, resulting in extended transcription that continues until the next termination codon is reached.
  - Splice site abnormalities affect the splicing process and removal of introns, again changing the protein sequence.
  - **Deletions**, in which part of the genetic sequence is removed. These can vary greatly in size, ranging from removal of a single base pair to deletion of the entire gene. Sometimes a large deletion will remove more than one gene; the result is a contiguous gene syndrome, such as cri-du-chat syndrome. Deletions within the coding region of a single gene generally have a larger effect if they remove a number of base pairs that is not a multiple of three, because this shifts the codon reading frame and all of the codons from that point on are altered.
  - **Insertions**, where extra base pairs or extended genomic sequences are inserted into the genetic sequence. Like deletions, these can vary in size and generally cause more damage if they result in a frameshift. A special case of insertion mutation is triplet expansion, described in the section above on spontaneous mutation in gametes.
  - **Inversions**, where the DNA sequence in part of a chromosome is reversed. In one cause of hemophilia A, part of intron 22 in the F8C gene recombines with another sequence elsewhere on the X chromosome that is homologous but reads in the opposite direction relative to the intron 22 sequence. After recombination, the sequence is reversed and therefore incorrect.

A mutation causes disease when it critically affects the function of a protein, whatever that function may be. Some disorders are associated predominantly with one or two very specific defects. For instance, 98% of achondroplasia cases are caused by a particular point mutation (a G to A point mutation at nucleotide 1138 of the FGFR3 gene). One reason why a disorder might be associated with only a few specific mutations is that the gene involved contains only a limited number of “hot spots” (for example, CpG dinucleotides which are prone to mutation and thus underrepresented in translated regions of DNA).

For most genetic diseases, though, there is not just one known mutation in the gene; instead there are multiple defects at large in the affected population with potentially varying effects on the resulting phenotype (phenotypic heterogeneity). For example, in cystinosis, mutations in the CTNS gene include missense, nonsense, and splice site point mutations, deletions, and insertions. One deletion allele produces no CTNS mRNA, while the others produce some residual mRNA. For Mendelian recessive disorders, the presence of multiple possible mutations can lead to compound heterozygotes (a person having two mutated genes each with different mutations).

In addition to defects *within* a gene, defects can be produced through the duplication of an entire gene. This can happen during meiotic recombination, for example. The most

common mutation for Charcot-Marie-Tooth, type 1A is a duplicated gene for peripheral myelin protein. Having three genes instead of two leads to overexpression of the protein, producing the disorder.

A few conclusions can be made about the relationship between a particular defect involved in a disorder and the resulting phenotype. As noted above, deletions and insertions that cause frame shifts often are more damaging than mutations that leave the frame intact. In disorders that involve triplet expansions (such as myotonic dystrophy), more severe phenotypes and/or earlier age of onset tend to be associated with increasing length of repeat sequence. For some diseases associated with several types of mutations, there is a strong genotype-phenotype correlation, meaning that the course of the disease can be predicted from a specific mutation type (and vice versa). For others, no strong correlation exists, and patients with a single mutation type may demonstrate a wide spectrum of clinical variability.

### **Chromosomal defects include numerical imbalances and other abnormalities**

Defects that involve entire chromosomes or partial chromosomes often affect multiple genes. One category of chromosomal imbalance is aneuploidy (the addition or deletion of entire chromosomes from the normal chromosome complement, or euploid). Some of the known types of chromosomal defects involving aneuploidy and other abnormalities include:

- **Monosomy**, the absence of one copy of a chromosome. Monosomy of the X chromosome (without a Y chromosome) results in Turner syndrome and is the only monosomy that is known to be viable.
- **Trisomy**, the presence of an extra copy of a chromosome. Of the autosomal chromosomes, chromosome 21 is the only possible trisomy compatible with survival past infancy. This trisomy results in Down syndrome. Trisomy of the sex chromosomes produces relatively mild effects compared to that of the autosomal chromosomes. XXY results in Klinefelter syndrome; XXX and XYY produce phenotypes that are not yet well-defined.
- **Triploidy**, which arises when three sets of 23 chromosomes combine to form an embryo. Most triploid pregnancies result in miscarriage, and of the few triploid babies that are born, most survive only a few hours or days.
- **Partial trisomy and partial monosomy** resulting from chromosome breakage and rearrangement. These commonly produce developmental defects combined with mental deficiency. For example, partial trisomy of chromosome 11q is a disorder in which the distal portion of the long arm of the 11th chromosome appears three times rather than twice in an individual's cells. This condition often results in delayed growth and mental retardation.
- **Uniparental disomy**, a situation where both copies of a particular chromosome in an individual came from a single parent. This can cause problems when the chromosome in question has imprinted regions (regions that are inactivated when inherited from one gender but not the other). When one parent's chromosome bearing inactivated genes replaces the other parent's chromosome containing activated genes, the result will be loss of that gene product. Angelman and Prader-Willi syndromes are examples of disorders involving imprinted genes which can be caused by uniparental disomy (other causes include gene deletions and imprinting defects).

- **Ring chromosomes**, which are created when the two ends of a chromosome attach to one another to form a circular shape. Often this is caused by breakage at both ends of a chromosome, leaving two “sticky” ends which then join to form a ring. The resulting effect on the individual depends on the chromosome involved and the amount of material that is lost. For example, people who are diagnosed as having a chromosome 18 ring may exhibit a number of symptoms ranging from mental retardation to deafness to heart anomalies and so on, depending on the exact nature of their defect.
- **Translocations**, which result when chromosomes break and the fragments reattach to other chromosomes. Translocations in essence create new genetic sequences, which may have harmful effects. For instance, several types of cancer have been associated with translocations that either create a fusion gene (as in chronic myelogenous leukemia) or bring a gene under the influence of new transcriptional factors and thus increase the expression of its protein (as in Burkitt’s lymphoma).
- **Robertsonian translocation**, a special case of translocation whereby two chromosomes with very short short ends fuse at the centromere and the short ends are lost. This abnormality is only viable in those chromosomes whose short ends do not contain essential genetic material (13, 14, 15, 21, and 22). This condition can lead to monosomy or trisomy upon meiosis. Some cases of Angelman syndrome are due to Robertsonian translocation where the father transmits a 15:15 translocation.

Due to the extreme nature of their defects, disorders involving chromosomal abnormalities are usually easy to diagnose using genetic testing.

## Number and type of causal factors

### **Genetic disorders are categorized as Mendelian or incompletely penetrant depending on the number of causal factors**

Some genetic mutations are potent enough to have dramatic health effects on their own – a single change in a single nucleotide in a single gene can wreak havoc. Others are not as potent and require other factors to be present to have an effect. It is much harder to determine the causes of a genetic disorder that involves multiple genes and/or external factors than it is for disorders caused by a single mutated gene.

Diseases that have 100% penetrance and involve only a single gene are called Mendelian disorders. Disorders that require the presence of environmental and/or other genetic factors are said to have incomplete penetrance.

### **Mendelian disorders are recessive or dominant, depending on whether one or both gene copies must be defective**

Mendelian disorders are divided into two categories. Dominant disorders require only one copy of the gene to be mutated to produce the phenotype; recessive disorders require both copies to be defective. Whether a disorder is recessive or dominant

depends on the protein expressed by the gene and the nature of the mutation. For example, a mutation which decreases the production or activity of an enzyme creates a recessive disorder if a single unaffected gene can produce enough normal enzyme on its own. Cystic fibrosis and Canavan disease are both recessive disorders arising from protein loss of function. If the single gene cannot produce enough protein on its own, a situation called haploinsufficiency, then the disorder will be dominant. A mutation which changes the structure of a protein so that it interferes specifically or non-specifically with molecular pathways will also often be dominant. Huntington disease and Marfan syndrome are examples of dominant disorders thought to be caused by harmful structural defects.

Occasionally the task of identifying the genetic defect in Mendelian disorders is made more difficult by the fact that multiple genes involved in a common pathway can result in the same phenotype when mutated. This is called genetic heterogeneity; hemophilia is an example of such a disorder with several different genetic causes.

Mendelian disorders, when stably inherited, exhibit strong aggregation within families and identifiable inheritance patterns. By definition, penetrance is complete for all Mendelian disorders, whether stably inherited or not.

### **Incompletely penetrant disorders require more than one genetic and/or environmental factor to be present**

Disorders that have incomplete penetrance can only be exhibited if two or more causal factors are present. These disorders can be the result of either multiple genetic factors (polygenic disorders) or genetic and environmental factors combined.

### **Polygenic disorders are determined by multiple genes**

Whereas Mendelian disorders are caused by a mutation of a single gene, polygenic disorders require two or more genes. An individual gene on its own is insufficient to cause disease. Occasionally a distinction is made between oligogenic disorders (involving a “few” genes) and polygenic disorders (involving “many” genes whose effects are additive).

Unlike Mendelian disorders, which are due to relatively uncommon and harmful genetic mutations, it may be possible for some polygenic disorders to be the result of relatively common and normally innocuous alleles that are only harmful when combined with others in certain unlucky combinations.

Hirschsprung disease is thought to be an oligogenic disease for which susceptibility is determined by multiple genes, none of which is able to cause the disease by itself. There is also evidence that type II diabetes, or non-insulin dependent diabetes mellitus, may be polygenic (environmental factors may also be implicated).

A polygenic (or oligogenic) model has been proposed for the progression of chronic myelogenous leukemia (CML). In this model, a reciprocal translocation between chromosomes 9 and 22 creates a hybrid gene comprised of part of the ABL gene on 9 and part of the BCR gene on 22. This fused gene increases the rate of mitosis and protects the cell from apoptosis (cell death), resulting in leukemia. A mutation in

an additional gene such as a proto-oncogene or tumor suppressor causes the rate of mitosis to rise sharply and the daughter cells fail to differentiate, initiating the crisis phase of the disease.

A disorder is suspected of being polygenic when it exhibits weak aggregation within families but no single causal gene is apparent.

### **Other disorders with incomplete penetrance require exposure to environmental factors**

Some disorders have a genetic basis but also involve one or more environmental factors. People who have the genetic defect(s) will exhibit the disorder only if exposed to the environmental factors.

One category of disorders with incomplete penetrance involves an abnormal response to drugs. Malignant hyperthermia is such a disorder – people with this disorder are at risk of dying under certain types of anesthesia. Also, the A1555G mutation in mitochondrial 12S rRNA confers susceptibility to deafness in people who are treated with aminoglycoside antibiotics. Sometimes the environmental factor in question is quite common. For instance, sunlight (UV radiation) is the triggering factor in xeroderma pigmentosum.

Disorders in this category show apparent aggregation within families but less than 100% penetrance; those who are affected have all been exposed to the same environmental factor.

## **Effect on protein expression**

### **Mutations can result in disorders by affecting the quantity of protein produced**

A key goal in genetic disease research is determining how a genetic defect results in physical harm to a person with the disease. Understanding this aspect of the disease is critical to developing effective treatments or cures.

Most genes encode proteins that serve a specific function. In many cases there are additional genes encoding proteins with similar or overlapping functions. This redundancy serves a vital role in reducing the likelihood that a single genetic defect will result in a deleterious phenotype. However, in other cases the gene encodes a unique protein indispensable to cell function, and in these cases producing the correct quantity of protein is critical.

One of the ways in which a mutation can cause physical harm is through affecting the expression of a gene, therefore altering the quantity of protein that is produced by that gene. Too much protein can cause problems; so can too little protein. The other way in which a mutation causes harm is through changing the structure and function of the protein, which is covered in the next section.

## **Insufficient protein expression results in loss of function**

Many disorders are caused by a total absence or insufficient production of a particular type of protein. The genetic defects involved either affect the quantity of protein produced or alter its structure so that it is unstable and degraded before it can perform its function. In X-linked dilated cardiomyopathy, the gene promoter for dystrophin in cardiac muscle is mutated, resulting in complete loss of production. In Duchenne muscular dystrophy, on the other hand, mutations produce a severely truncated dystrophin protein that is useless and gets degraded. Splicing mutations are often the cause of unstable proteins.

Mutations that affect the expression of proteins do not necessarily have to directly involve the genes for those proteins. Some protein deficits may be due to defects in *other* genes – for instance, the triplet expansions in myotonic dystrophy have been theorized to affect the chromosomal structure, suppressing other nearby genes. Mutations in non-coding RNA genes may also affect protein synthesis. For example, it has been suggested that the A1555G point mutation in the mitochondrial 12S rRNA gene that causes deafness in the presence of aminoglycosides does so by increasing the propensity of the rRNA molecules to bind to these antibiotics, which disrupts mitochondrial protein synthesis.

Most disorders caused by loss of protein expression are recessive because the protein produced is not a rate-limiting resource. One normal gene can produce enough protein on its own, so both genes need to be defective in order for the disease to be exhibited. However, there are a few examples of dominant diseases resulting from loss of production. Acute intermittent porphyria results when one copy of the gene for HBM synthase is knocked out, leading to half the normal level of the protein, which is insufficient.

## **Excessive protein expression leads to gain of function**

In some diseases, a gene is overexpressed and produces too much protein. Its function is therefore carried out to excess with damaging results. One example of such a disease is Charcot-Marie-Tooth, type 1A, caused by duplications of the PMP-22 gene. This disease exhibits a relative gene dosage effect (four copies of the gene results in a more severe phenotype than does three copies), implying that increased protein quantity is the basis for the disease. Diseases stemming from overproduction tend to exhibit dominant inheritance patterns.

## **Effect on protein structure and function**

**Mutations can also cause disorders by changing the structure of the protein, thereby affecting how well it functions**

The amino acid sequence of a protein determines its shape and structure, including how the protein is folded, which acids are on the inside and which are on the outside, whether the protein is fibrous, globular, or some other shape, etc. In turn, the structure of a protein is intrinsically related to its function. DNA mutations that cause alterations in the amino acid sequence can dramatically affect the protein's structure, thereby affecting how it functions and interacts with other molecules.

As with mutations that affect protein quantity, understanding the mechanism of mutations that affect protein structure is vitally important for developing therapies for genetic diseases.

Most mutations that change the structure of a protein make the protein less effective, resulting in a loss of function, but some mutations cause the protein to be too active and some even wind up impairing other functions.

**Most mutations that change protein structure reduce the protein's effectiveness**

There are a wide variety of mutations that result in loss of function due to alteration of functional sites in proteins. The table below lists some loss of function genetic disorders resulting from structure changes, categorized by protein type (enzyme, structural, etc.). Some mutations affect the protein's ability to fold into the proper conformation. For example, one common cause of collagen disorders such as osteogenesis imperfecta, Alport syndrome, and Ehlers-Danlos vascular type, is the replacement of the small amino acid glycine with a larger acid, restricting the protein's ability to fold and bend. Others disorders are caused by the protein's inability to target or bind to other molecules. In androgen insensitivity syndrome, the ability of an androgen receptor to bind to androgens or regulatory nucleotides is impaired. For many disorders the exact nature of the impairment is not yet known.

<b>Protein type</b>	<b>Disease</b>
<b>Enzyme</b>	<b>Canavan disease</b> Deficiency in enzyme aspartoacylase leads to build-up of N-acetylaspartic acid in the brain. Some mutations are null mutations which result in no production but some are missense that make less active forms of the protein.
	<b>Lesch-Nyhan syndrome</b> Multiple mutation types produce non-functional or very low function versions of the enzyme hypoxanthine/guanine phosphoribosyltransferase.
	<b>Galactosemia</b> Mutations reduce enzyme activity of GALT, for example by preventing formation of a stable GALT-UMP intermediate.
<b>Structural</b>	<b>Marfan syndrome</b> Fibrillin properties are altered in a way that reduces its deposition into extracellular matrix and also interferes with utilization of product from the normal allele.
	<b>Alport syndrome (X-linked)</b> Mutations typically result in a structural alteration of type III collagen that leads to intracellular storage and impaired secretion of collagen chains.

<b>Regulatory</b>	<b>Androgen insensitivity syndrome</b> Androgen receptor that exerts transcriptional control over target genes is mutated (usually by missense point mutations), impairing binding to either androgens or to regulatory nucleotides at the DNA-binding domain.
<b>Storage</b>	<b>Sickle cell anemia</b> Point mutation in beta subunit of hemoglobin leads to polymerization after oxygen release; repeated polymerization/depolymerization cycles eventually damage the hemoglobin and lead to destruction of the red cell itself.
<b>Transport</b>	<b>Cystinosis</b> Mutation in protein cystosin which is thought to transport the disulfide amino acid cystine out of the lysosome and into the cytoplasm. Some mutations are missense in or immediately before the transmembrane region; others affect non-functional areas and result in milder version; others are severely truncating resulting in loss of protein.
	<b>Wilson disease</b> Mutation in the gene that produces the main copper transporter moving copper from the hepatocyte into the bile.
	<b>Cystic fibrosis</b> Mutations in the gene for cystic fibrosis transmembrane conductance regulator that functions as a regulated chloride channel in epithelia. Mutations affect quantity and/or quality of the protein. One mutation may hinder the dissociation of the transport-regulator protein from one of its chaperones, preventing its maturation into the properly folded form.
<b>Receptors/ signalers</b>	<b>Alagille syndrome</b> Mutation in Jagged1 gene often results in severely truncated protein, which lacks the transmembrane region necessary to embed in the cell membrane and participate in Notch pathway signaling.
	<b>Hereditary hemochromatosis (HHC)</b> One mutation removes a highly conserved cysteine residue that normally forms an intramolecular disulfide bond, and thereby prevents the HFE protein from being expressed on the cell surface; the other may impair interaction of the HFE protein with the transferrin receptor on the cell surface.

As with mutations resulting in a lack of production, mutations that impair effectiveness are recessive as long as the output of one normal gene is not rate-limiting and as long as the abnormal product does not interfere with the normal protein (known as a dominant negative effect).

### **Structure-altering mutations occasionally result in excess activity**

Some mutations have the opposite effect of those listed above and actually increase the activity of the protein. The resulting disorders typically display dominant inheritance patterns.

One mechanism by which this can occur is through the production of mutant proteins with increased activity compared to the wild type. In FGFR-related craniosynostosis syndromes such as Apert syndrome, mutated fibroblast growth factor receptors either have a higher affinity for their ligands than the wild type, or are able to activate on their own without the presence of the ligand. The same mechanism is also seen in multiple endocrine neoplasia type 2 – constitutive gain of function results from the mutant protein's ability to dimerize and activate independent of ligands.

Another mechanism that results in increased activity involves decreased ability to degrade. In factor V Leiden thrombophilia, factor V Leiden is inactivated at an approximately 10-fold slower rate than normal factor V and persists longer in the circulation, resulting in increased thrombin generation. The cause is a point mutation in one of the three cleavage sites in the molecule.

### **Harmful structural changes impair other cellular functions**

The third class of mutations involving structural protein changes includes those that inhibit or impair the functioning of other proteins.

Diseases involving triplet repeat expansions often fall into this category. Huntington disease, DRPLA, and spinocerebellar ataxia type 1 all produce expanded polyglutamine proteins that cause damage and are aggregated into clumps, which often incorporate other proteins such as chaperones, ubiquitin, and proteasomal subunits. The role of protein structure in the pathogenesis of these diseases is still under investigation. One hypothesis is that mutated proteins attract regulatory proteins such as CREB binding protein (CBP) which get pulled away from DNA and become useless.

Triplet repeat expansion is not the only cause of harmful structural changes. In sickle cell anemia, mutant hemoglobin molecules polymerize and depolymerize, affecting not only oxygen storage (hemoglobin's function) but also restricting the movement of red blood cells and eventually leading to red blood cell destruction via iron release and attraction of antibodies. And in transthyretin amyloidosis, mutations in the transthyretin proteins reduce its stability in a tetramer form and promote the polymerization of the monomer into amyloid structures.

## **Phenotypic characteristics**

**Diseases can be described by the systems/organs affected, age of onset, gender-specific effects, and variations in symptom expression**

Our final classification method for genetic diseases concerns their phenotypes, or the outward manifestations they produce. Some of the important phenotypic characteristics typically analyzed for a disease are the physiological systems and organs affected, the age of onset, and gender-specificity (whether or not the disease affects men and women equally). It is also important to know whether the appearance and severity of symptoms are relatively constant from person to person or vary widely across the affected population, and if they vary, what the causes of the variation may be.

Ideally, for a given disorder, we would like to understand the complete mechanism through which the genotype translates into the phenotype, and be able to explain why a genetic defect causes specific biochemical, cellular and tissue defects, why the symptoms appear at a certain age, and why certain people are affected differently from

others. We would also like to be able to apply our knowledge of these mechanisms to the development of ways to predict a disorder's genotypic features from its phenotypic features and vice versa. Although we are not yet at this level of understanding for most disorders, a few generalizations correlating phenotype with genotype have been developed; these are noted in the following paragraphs.

### **Disorders affecting a specific organ or system typically stem from a wide range of genetic causes**

The development and maintenance of any particular organ or physiologic system within the body requires the presence of a large number of different types of proteins (enzymes, receptors, regulatory proteins, etc.) operating and cooperating in many different pathways. Therefore, it is reasonable to suppose that a wide range of different genetic causes can be responsible for disorders affecting a given organ or system. For instance, heart defects are manifest in disorders involving a diverse set of proteins – examples include Marfan syndrome, caused by a structural protein defect, Friedreich ataxia, which involves a defective transport protein, and Alagille syndrome, which results from a mutation in a signaling protein gene. Conversely, genetic disorders often affect multiple systems. For example, cystinosis impairs a person's growth, renal function, vision, and pulmonary system, among other things. Therefore, it is usually difficult to draw any general conclusions about a disorder's genotype and protein involved based solely on analysis of the systems or organs affected.

### **Age of onset and progression of disease can sometimes be correlated with genotype or protein function**

The term "age of onset" refers to the age at which the initial symptoms or manifestations of a disease first appear in an individual. This is not necessarily the age at which the genetic defect starts affecting the individual; in fact, a person may harbor the effects of a genetic disease long before these effects become noticeable. For instance, in hereditary hemochromatosis, patients may exhibit abnormal serum iron levels years before actual symptoms are noticed, and in fact the iron levels of pre-symptomatic people thought to be at risk for this disease are often tested to see if symptoms are likely to develop later.

Knowing when a disease is most likely to manifest itself helps us understand what mechanisms may be involved. Some disorders manifest themselves from birth or soon thereafter. Others do not become apparent until childhood, adolescence, early adulthood or late adulthood. Most disorders have a typical age of onset, but there are some that manifest themselves at almost any age.

For most diseases, we do not know enough about the genetic defect, the protein involved, or the effect of the mutation to be able to adequately explain its age of onset. For instance, it is not known why the genetic disorder Leber's hereditary optic neuropathy (LHON) typically develops in early adulthood or why mitochondrial encephalopathy with acidosis and stroke-like episodes (MELAS) frequently appears in childhood.

However, through analyzing diseases where the disease mechanism *is* somewhat understood, a few generalizations can be made about genetic defect and age of onset. Genetic disorders resulting from enzyme deficiencies often become apparent soon after

birth. The reason may be that embryos have access to their mothers' enzymes via the placenta up until the time of birth, at which point the deficiency starts to become manifest. Canavan disease, Lesch-Nyhan syndrome, Angelman syndrome and galactosemia are all enzymatic deficiency diseases that strike in infancy. However, there are exceptions to this rule. For instance, while most cases of Tay-Sachs develop in infancy, some do not manifest themselves until adulthood; these are due to mutations that diminish but do not absolutely nullify the activity of the enzyme hexosaminidase A.

Triplet repeat expansion diseases also generally demonstrate a correlation between genotype and age of onset. For these diseases, which include Huntington disease and myotonic dystrophy, increased length of repeat expansion correlates with earlier age of onset. In other words, larger mutations on average will more quickly result in a noticeable effect, resulting in the phenomenon of anticipation as mutation size increases from one generation to the next.

Another phenotypic characteristic related to age of onset is progression of the disease. In some diseases, symptoms can stabilize at a certain level (especially if treatment is available). In others, they get progressively worse over time, or periodically relapse and remit. Most likely, these phenotypic characteristics relate to the molecular pathway or the organ system involved. For example, a genetic defect that leads to neuronal death may cause a progressive illness whereas a genetic defect leading to hypercoagulability may lead to episodic strokes. In addition, if the phenotypic presentation of a genetic defect requires an environmental factor and that factor is presented episodically, then the disease itself may be episodic. Amyotrophic lateral sclerosis (ALS) and Duchenne muscular dystrophy are steadily progressive diseases; systemic lupus erythematosus (SLE), rheumatoid arthritis, familial Mediterranean fever, and familial hemiplegic migraine are all relapsing/remitting diseases.

### **Diseases can affect men and women differently due to sex chromosome linkage, hormonal effects, and other factors**

While many diseases affect men and women equally, there are some that do present a gender bias. Known reasons why a disease might affect one gender more often or more severely than the other include:

- **X- or Y-linkage:** X-linked diseases show a gender bias that depends on whether the genetic defect is recessive or dominant. X-linked dominant mutations will affect more women than men; X-linked recessive mutations will affect more men than women. Mutations on the Y chromosome will affect only men.
- **Hormonal effects:** The effects of some hormones are gender-specific, so mutations that affect those hormones or the pathways in which they operate may likewise have gender-specific effects. For example, a mutation of the gene for luteinizing hormone receptor (LHR) that renders it constitutively active results in precocious puberty for boys, but has no effect on females. LHR mutations that render the receptor inactive affect the sexual development of both males and females, but the clinical presentation is different for each gender.

However, for many diseases that exhibit gender biases, the reasons for differences in penetrance and expressivity between genders are not as clear-cut. Leber's hereditary optic neuropathy (LHON) exhibits several gender biases whose causes have not yet been explained: it affects more men than women and appears earlier in men than in

women, but the symptoms may be more severe in women than in men. Diseases that are suspected to have an autoimmune basis such as systemic lupus erythematosus and rheumatoid arthritis appear to affect more women than men (up to 15 times more women than men in the case of lupus). One theory is that gender-based differences in the immune system due to women's role in childbearing lead to higher autoimmune susceptibility for women, but this has not been proven.

### **Individual variations in expressivity can be caused by many factors**

Often the number, type and severity of the symptoms manifested by a specific disease vary from person to person. Factors that may account for variation between individuals include:

- **Allelic variances:** In some disorders, the specific allelic variant(s) acquired by an individual strongly influence the clinical presentation and progression of the disorder. For example, two disease-causing alleles (C282Y and H63D) have been identified in the HFE gene that produces hereditary hemochromatosis protein. People who inherit two copies of the C282Y allele have more severe symptoms than do those who inherit one of each allele. Similarly, of the three alleles responsible for LHON, G11778A generally causes the most severe visual failure, T14484C is associated with the best long-term prognosis, and G3460A results in an intermediate phenotype. And with triplet repeat expansion diseases, the severity of the phenotype sometimes correlates positively with the size of the expansion, as is the case in spinocerebellar ataxia type 1 and myotonic dystrophy.
- **Mitochondrial DNA heteroplasmy:** As mentioned previously, a person inheriting a mitochondrial DNA mutation can actually inherit a mixture of mutated and normal DNA. The ratio of mutated to normal DNA in an individual can affect the severity of his/her disease. In addition, unequal distribution of the mutated DNA to different tissues in the body can determine which symptoms are exhibited. Furthermore, the percentage of mutated DNA in a given tissue can change over time within an individual.
- **Somatic mosaicism:** When only some of an individual's cells contain a mutation, the result can be a varying form of the disorder or expression in only certain parts of the body. For example, McCune-Albright syndrome is thought to be the result of a postzygotic mutational event occurring in a mosaic fashion early in embryonic life. Because the mutation is distributed throughout the body sporadically, the severity of the symptoms and type of tissue affected can vary from individual to individual.
- **Environmental factors:** Some diseases require the presence of a particular environmental factor in order to become manifest. Therefore the exposure to that factor received by an individual will affect the expression of the disease. For instance, in the case of xeroderma pigmentosum, the severity of the disease increases with the amount of UV radiation received. Other diseases have been observed to have environmental "triggers" that may influence the severity or course of a disease. In acute intermittent porphyria, for example, attacks may be precipitated by drugs, alcohol, hormonal changes, starvation or infection. In addition, there are disorders whose causes are not yet determined but which are thought to have an environmental basis as well as a genetic basis. One of these disorders is type II diabetes, whose risk factors include obesity and lack of physical activity in addition to family history of the disease. In each of these

cases, phenotypic variations among individuals may be the result of variations in environmental factors.

- **Other genetic factors (epistasis):** The phenotype of some genetic diseases can be influenced strongly by the genotype at one or more other loci. The effect produced by these epistatic genes can be either beneficial or detrimental. Sickle cell anemia is an example of a disorder whose phenotype is influenced by other genes. The severity of the disease appears to be moderated by two genetically inherited conditions, alpha-thalassemia (the loss of one or more genes encoding the alpha globin chain) and hereditary persistence of fetal hemoglobin (in which expression of fetal hemoglobin is maintained at its original level even after infancy, when levels normally decline).

## Further Reading

For more information on the topics discussed in this paper, we recommend the following resources:

Books:

Strachan T and Read AP (1999) *Human Molecular Genetics*, 2<sup>nd</sup> edition. BIOS Scientific Publishers, Oxford.

Articles:

Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. *Journal of Medical Genetics* 38, 729-739.

Cartegni L, Chew SL, and Krainer AR. Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nature Reviews Genetics* 3, 285-298 (2002).

Crow J. The origins, patterns and implications of human spontaneous mutation. *Nature Reviews Genetics* 1, 40-47 (2000).

Cummings CJ and Zoghbi HY. Fourteen and counting: unraveling trinucleotide repeat diseases. *Human Molecular Genetics* 9, 909-916 (2000).

DiMauro S, Schon EA. Mitochondrial DNA mutations in human disease. *American Journal of Medical Genetics* 106,18-26 (2001).

Elliott B, Jasin M. Double-strand breaks and translocations in cancer. *Cellular and Molecular Life Sciences* 59, 373-385 (2002).

Emanuel BS and Shaikh TH. Segmental duplications: An “expanding” role in genomic instability and disease. *Nature Reviews Genetics* 2, 791-800 (2001).

Hassold T and Hunt P. To err (meiotically) is human: The genesis of human aneuploidy. *Nature Reviews Genetics* 2, 280-291 (2001).

Hesterlee S. Mitochondrial myopathy: An energy crisis in the cells. *Quest* 6 (1999).

Jimenez-Sanchez G, Childs B and Valle D. Human disease genes. *Nature* 409, 853-855 (2001).

Reik W and Walter J. Genomic imprinting: Parental influence on the genome. *Nature Reviews Genetics* 2, 21-32 (2001).

Sreekumar KR, Aravind L, Koonin EV. Computational analysis of human disease-associated genes and their protein products. *Current Opinion in Genetics and Development* 11, 247-57 (2001).

Wu S-M, Leschek EW, Rennert OM, and Chan W-Y. Luteinizing hormone receptor mutations in disorders of sexual development and cancer. *Frontiers in Bioscience* 5, d343-352 (2000).

Zlotogora J. Germ line mosaicism. *Human Genetics* 102, 381-6 (1998).

On-line references:

Online Mendelian Inheritance in Man™ hosted by the National Center for Biotechnology Information. <http://www.ncbi.nlm.nih.gov/omim/>

Neuromuscular Disease Center of the Washington University School of Medicine, St. Louis, MO. <http://www.neuro.wustl.edu/neuromuscular/index.html>

GeneTests·GeneClinics Web site developed by Children's Health System and University of Washington, Seattle. [www.geneclinics.org](http://www.geneclinics.org)