Help Me Understand Genetics

Variants and Health

Reprinted from MedlinePlus Genetics

U.S. National Library of Medicine

National Institutes of Health

Department of Health & Human Services

CONTENTS

Variants and Health

1 What is a gene variant and how do variants occur?

A gene variant is a permanent change in the DNA sequence that makes up a gene. This type of genetic change used to be known as a gene mutation, but because changes in DNA do not always cause disease, it is thought that gene variant is a more accurate term. Variants can affect one or more DNA building blocks (nucleotides) in a gene.

Gene variants can be inherited from a parent or occur during a person's lifetime:

- Inherited (or hereditary) variants are passed from parent to child and are present throughout a person's life in virtually every cell in the body. These variants are also called germline variants because they are present in the parent's egg or sperm cells, which are also called germ cells. When an egg and a sperm cell unite, the resulting fertilized egg cell contains DNA from both parents. Any variants that are present in that DNA will be present in the cells of the child that grows from the fertilized egg.
- Non-inherited variants occur at some time during a person's life and are present only in certain cells, not in every cell in the body. Because non-inherited variants typically occur in somatic cells (cells other than sperm and egg cells), they are often referred to as somatic variants. These variants cannot be passed to the next generation. Non-inherited variants can be caused by environmental factors such as ultraviolet radiation from the sun or can occur if an error is made as DNA copies itself during cell division.

Some genetic changes are described as new (de novo) variants; these variants are recognized in a child but not in either parent. In some cases, the variant occurs in a parent's egg or sperm cell but is not present in any of their other cells. In other cases, the variant occurs in the fertilized egg shortly after the egg and sperm cells unite. (It is often impossible to tell exactly when a de novo variant happened.) As the fertilized egg divides, each resulting cell in the growing embryo will have the variant. De novo variants are one explanation for genetic disorders in which an affected child has a variant in every cell in the body, but the parents do not, and there is no family history of the disorder.

Variants acquired during development can lead to a situation called mosaicism, in which a set of cells in the body has a different genetic makeup than others. In mosaicism, the genetic change is not present in a parent's egg or sperm cells, or in the fertilized egg, but happens later, anytime from embryonic development through adulthood. As cells grow and divide, cells that arise from the cell with the altered gene will have the variant, while other cells will not. When a proportion of somatic cells have a gene variant and others do not, it is called somatic mosaicism. Depending on the variant and how many cells are affected, somatic mosaicism may or may not cause health problems. When a proportion of egg or sperm cells have a variant and others do not, it is called germline mosaicism. In this situation, an unaffected parent can pass a genetic condition to their child.

Most variants do not lead to development of disease, and those that do are uncommon in the general population. Some variants occur often enough in the population to be considered common genetic variation. Several such variants are responsible for differences between people such as eye color, hair color, and blood type. Although many of these common variations in the DNA have no negative effects on a person's health, some may influence the risk of developing certain disorders.

For more information about variants:

The Centre for Genetics Education provides a fact sheet discussing variations in the genetic code (https://www.genetics.edu.au/SitePages/Variations-in-the-Genetic-Code.aspx).

An introductory explanation of what variants are and how they occur (https://www.yourgenome.org/facts/what-is-a-mutation) is provided by Your Genome from the Wellcome Genome Campus.

KidsHealth from Nemours offers an introduction to genes, genetics, and genetic changes (https://kidshealth.org/en/parents/about-genetics.html).

Additional information about genetic alterations is available from the University of Utah fact sheet "What is Mutation?" (https://learn.genetics.utah.edu/content/basics/mutation/)

Using animations and videos, DNA From the Beginning (Cold Spring Harbor National Laboratory) describes the early experiments that helped researchers understand genetic variants (http://www.dnaftb.org/27/).

2 How can gene variants affect health and development?

To function correctly, each cell depends on thousands of proteins to do their jobs in the right places at the right times. Sometimes, gene variants (also known as mutations) prevent one or more proteins from working properly. By changing a gene's instructions for making a protein, a variant can cause a protein to malfunction or to not be produced at all. When a variant alters a protein that plays a critical role in the body, it can disrupt normal development or cause a health condition. A condition caused by variants in one or more genes is called a genetic disorder.

In some cases, gene variants are so severe that they prevent an embryo from surviving until birth. These changes occur in genes that are essential for development, and often disrupt the development of an embryo in its earliest stages. Because these variants have very serious effects, they are incompatible with life.

It is important to note that genes themselves do not cause disease—genetic disorders are caused by variants that alter or eliminate a gene's function. For example, when people say that someone has "the cystic fibrosis gene," they are usually referring to a version of the *CFTR* gene that contains a variant that causes the disease. All people, including those without cystic fibrosis, have a version of the *CFTR* gene.

For more information about variants and genetic disorders:

The Genetic and Rare Diseases Information Center provides information to various resources about genetic disorders (https://rarediseases.info.nih.gov/).

The Centre for Genetics Education offers a fact sheet about genetic changes that lead to disorders (https://www.genetics.edu.au/SitePages/Types-of-Genetic-variation.aspx).

3 Do all gene variants affect health and development?

No; only a small percentage of variants cause genetic disorders—most have no impact on health or development. For example, some variants alter a gene's DNA sequence but do not change the function of the protein made from the gene.

Often, gene variants that could cause a genetic disorder are repaired by certain enzymes before the gene is expressed and an altered protein is produced. Each cell has a number of pathways through which enzymes recognize and repair errors in DNA. Because DNA can be changed or damaged in many ways, DNA repair is an important process by which the body protects itself from disease.

A very small percentage of all variants actually have a positive effect. These variants lead to new versions of proteins that help an individual better adapt to changes in his or her environment. For example, a beneficial variant could result in a protein that protects an individual and future generations from a new strain of bacteria.

Because a person's genetic code can have many variants with no effect on health, diagnosing genetic disorders can be difficult.

When determining if a gene variant is associated with a genetic disorder, the variant is evaluated using scientific research to date, such as information on how the variant affects the function or production of the protein that is made from the gene and previous variant classification data. The variant is then classified on a spectrum based on how likely the variant is to lead to the disorder.

Gene variants, as they relate to genetic disorders, are classified into one of five groups:

- Pathogenic: The variant is responsible for causing disease. There is ample scientific research to support an association between the disease and the gene variant. These variants are often referred to as mutations.
- Likely pathogenic: The variant is probably responsible for causing disease, but there is not enough scientific research to be certain.
- Variant of uncertain significance (VUS or VOUS): The variant cannot be confirmed to play a role in the development of disease. There may not be enough scientific research to confirm or refute a disease association or the research may be conflicting.
- Likely benign: The variant is probably not responsible for causing disease, but there is not enough scientific research to be certain.
- Benign: The variant is not responsible for causing disease. There is ample scientific research to disprove an association between the disease and the gene variant.

Evaluation needs to be done for each variant. Just because a gene is associated with a disease, does not mean that all variants in that gene are pathogenic. Additionally, evaluation of a variant needs to be done for all diseases with which it is thought to be associated. A variant that is pathogenic for one disease, is not necessarily pathogenic for a different disease. It is important to re-evaluate variants periodically; the

classification of a variant can change over time as more information about the effects of variants becomes known through additional scientific research.

Scientific journal article for further reading

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5): 405-24. doi: 10.1038/gim.2015.30. Epub 2015 Mar 5. PMID: 25741868; PMCID: PMC4544753.

For more information about the health effects of gene variants:

The University of Utah Genetic Science Learning Center provides information about genetic disorders (https://learn.genetics.utah.edu/content/disorders/) that explains why some variants cause disorders but others do not.

The National Human Genome Research Institute provides information about human genomic variation (https://www.genome.gov/dna-day/15-ways/human-genomic-variation).

Cold Spring Harbor National Laboratory's DNA From the Beginning explains the discovery of DNA repair mechanisms in cells (http://www.dnaftb.org/28/) and introduces the researchers who worked to understand these mechanisms.

FORCE (Facing Our Risk of Cancer Empowered) explains the significance of variants of unknown significance in cancer (https://www.facingourrisk.org/understanding-brca-and-hboc/information/hereditary-cancer/genetic-testing/basics/variants-of-uncertain-significance.php).

4 What kinds of gene variants are possible?

The DNA sequence of a gene can be altered in a number of ways. Gene variants (also known as mutations) can have varying effects --on health, depending on where they occur and whether they alter the function of essential proteins. Variant types include the following:

Substitution

This type of variant replaces one DNA building block (nucleotide) with another. Substitution variants can be further classified by the effect they have on the production of protein from the altered gene.

- Missense: A missense variant (Figure 1) is a typeof substitution in which the nucleotide change results in the replacement of one protein building block (amino acid) with another in the protein made from the gene. The amino acid change may alter the function of the protein.
- Nonsense: A nonsense variant (Figure 2) is another type of substitution. Instead of causing a change in one amino acid, however, the altered DNA sequence results in a stop signal that prematurely signals the cell to stop building a protein. This type of variant results in a shortened protein that may function improperly, be nonfunctional, or get broken down.

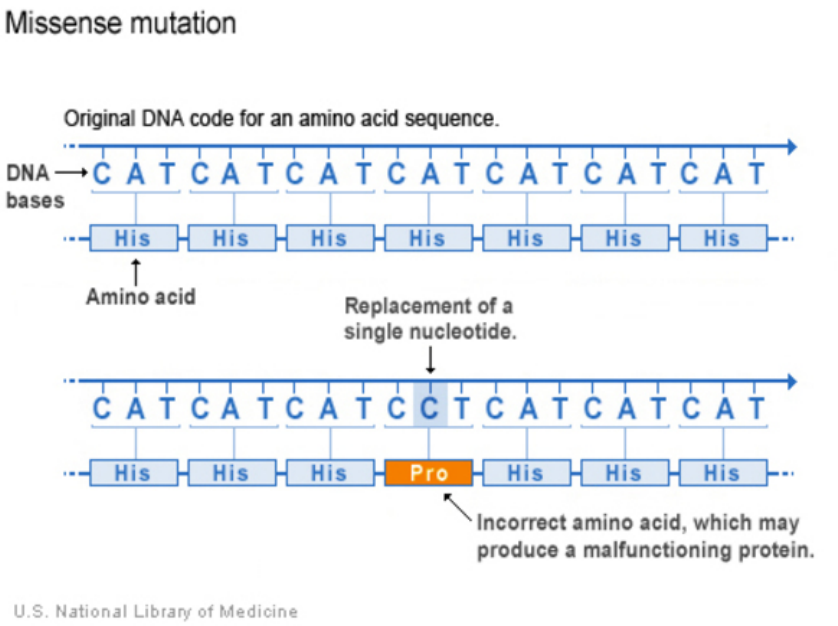

FIGURE 1: Replacement of a single nucleotide in the DNA sequence creates an incorrect amino acid, which may result in a malfunctioning protein.

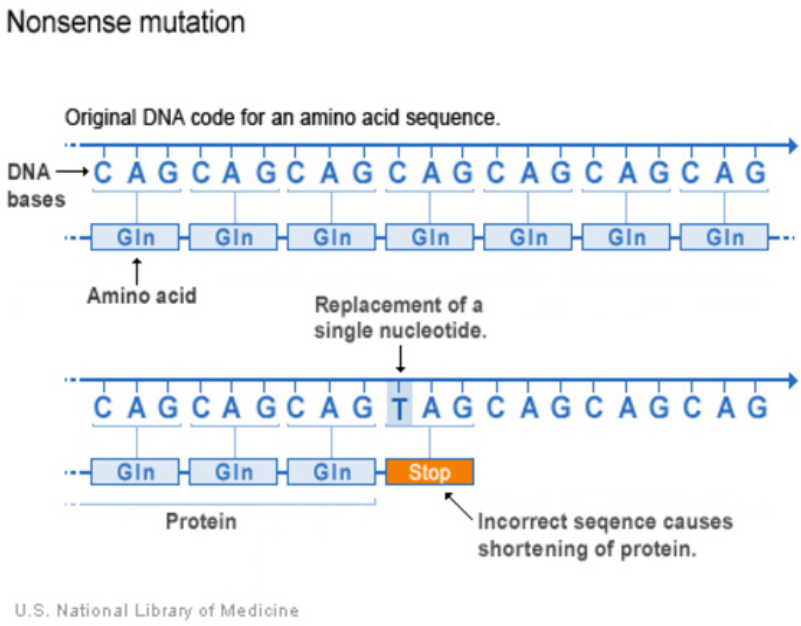

FIGURE 2: Replacement of a single nucleotide leads to a stop codon, which halts protein production at that point.

Insertion

An insertion (Figure 3) changes the DNA sequence by adding one or more nucleotides to the gene. As a result, the protein made from the gene may not function properly.

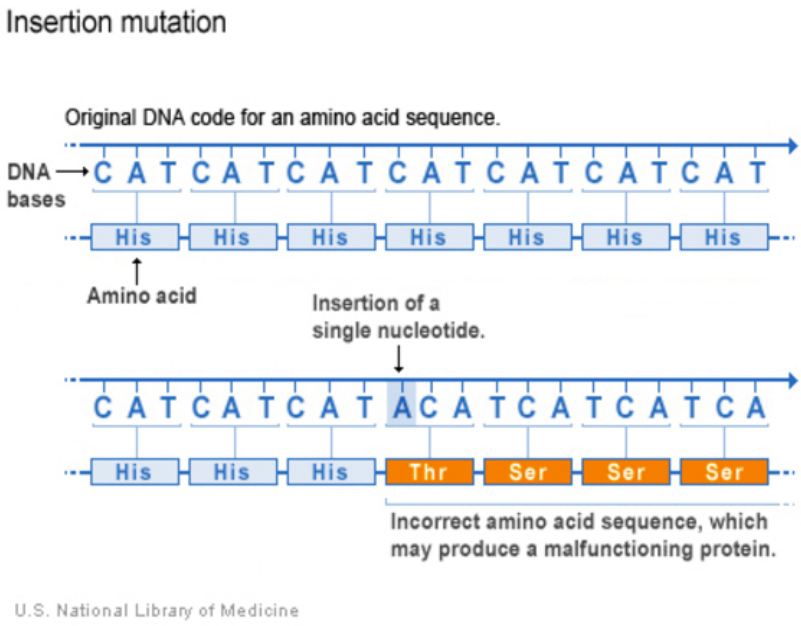

FIGURE 3: Insertion of a single nucleotide in the DNA sequence leads to an incorrect sequence of amino acids, which may produce a malfunctioning protein.

Deletion

A deletion (Figure 4) changes the DNA sequence by removing at least one nucleotide in a gene. Small deletions remove one or a few nucleotides within a gene, while larger deletions can remove an entire gene or several neighboring genes. The deleted DNA may alter the function of the affected protein or proteins.

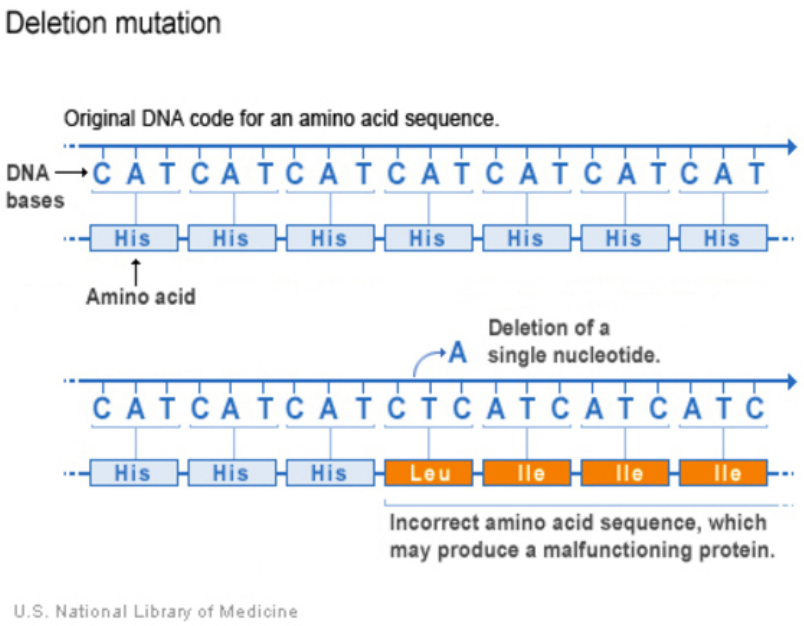

FIGURE 4: The deletion of a single nucleotide leads to an incorrect amino acid sequence, which may produce a malfunctioning protein.

Deletion-Insertion

This variant occurs when a deletion and insertion happen at the same time in the same location in the gene. In a deletion-insertion variant, at least one nucleotide is removed and at least one nucleotide is inserted. However, the change must be complex enough to differ from a simple substitution. The resulting protein may not function properly. A deletion-insertion (delins) variant may also be known as an insertion-deletion (indel) variant.

Duplication

A duplication (Figure 5)occurs when a stretch of one or more nucleotides in a gene is copied and repeated next to the original DNA sequence. This type of variant may alter the function of the protein made from the gene.

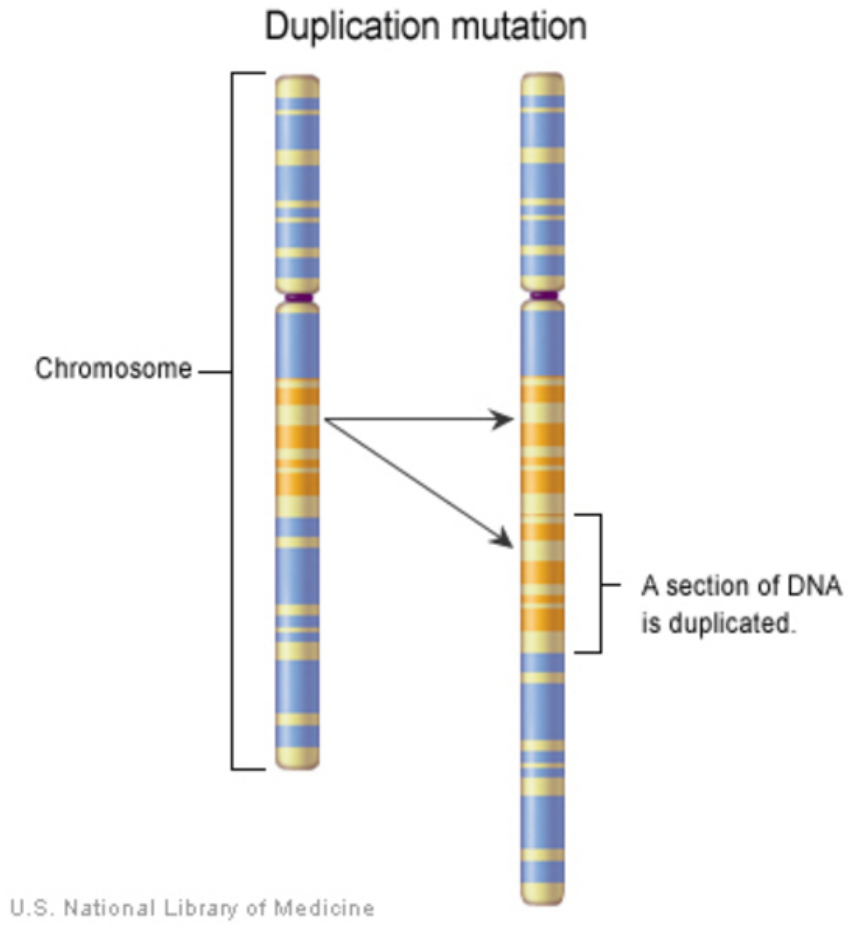

FIGURE 5: A section of DNA from a chromosome is duplicated.

Inversion

An inversion changes more than one nucleotide in a gene by replacing the original sequence with the same sequence in reverse order.

Frameshift

A reading frame consists of groups of three nucleotides that eachcode for one amino acid. Aframeshift variant (Figure 6) occurs when there is an addition or loss of nucleotides that shifts the grouping and changes the code for all downstream amino acids. The resulting protein is usually nonfunctional. Insertions, deletions, and duplications can all be frameshift variants.

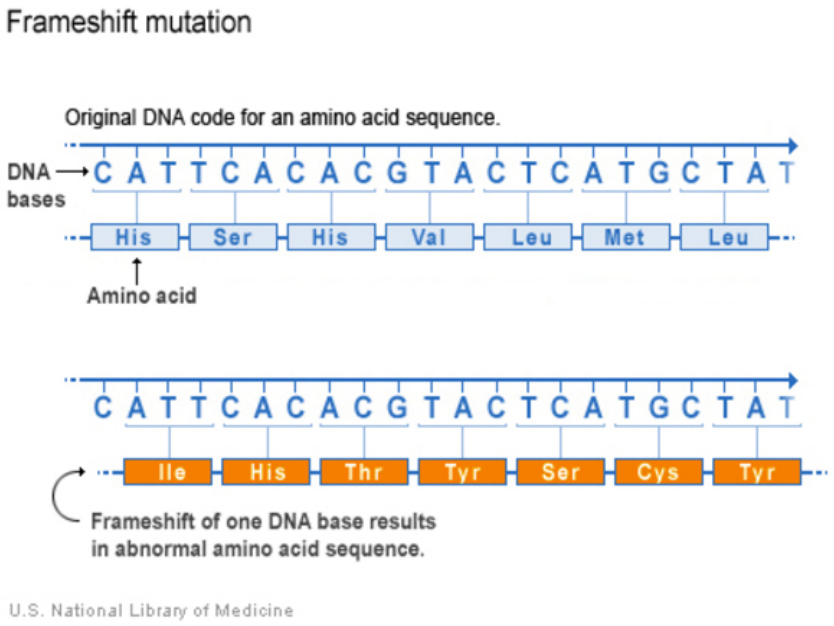

FIGURE 6: A framshift of one DNA base results in an abnormal amino acid sequence.

Repeat expansion

Some regions of DNA contain short sequences of nucleotides that are repeated a number of times in a row. For example, a trinucleotide repeat is made up of sequences of three nucleotides, and a tetranucleotide repeat is made up of sequences of four nucleotides. Arepeat expansion (Figure 7) is a variant that increases the number of times that the short DNA sequence is repeated. This type of variant can cause the resulting protein to function improperly.

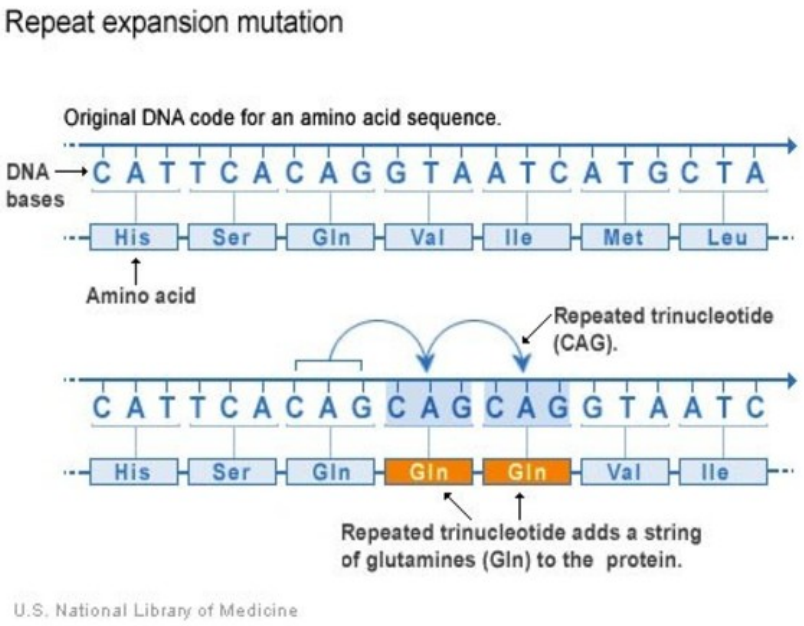

FIGURE 7: A trinucleotide (CAG) is repeated multiple times in a DNA sequence, adding a string of glutamines to the protein.

For more information about the types of gene variants:

The National Human Genome Research Institute offers a Talking Glossary of Genetic Terms (https://www.genome.gov/genetics-glossary). This resource includes definitions, diagrams, and detailed audio descriptions of several of the gene variants listed above.

A brief explanation of different variants types (https://www.yourgenome.org/facts/what-types-of-mutation-are-there) is available from yourgenome.org, a service of the Wellcome Trust.

The Khan Academy has a video describing the different types of gene variants (https://www.khanacademy.org/test-prep/mcat/biomolecules/genetic-mutations/v/the-different-types-of-mutations).

5 Can a change in the number of genes affect health and development?

People have two copies of most genes, one copy inherited from each parent. In some cases, however, the number of copies varies—meaning that a person can have one, three, or more copies of particular genes. Less commonly, both copies of a gene may be missing. These types of genetic difference are known as copy number variations (CNV).

Copy number variation results from insertions, deletions, and duplications of large segments of DNA that are at least one thousand nucleotides (also called 1 kilobase or 1kb) in length. These segments are often big enough to include whole genes. Variation in gene copy number can influence the activity of genes and the functioning of proteins made from them, which may affect body processes.

Copy number variation accounts for a significant amount of genetic difference between people. More than 10 percent of the human genome appears to contain differences in gene copy number. While much of this variation does not affect health or development, some differences influence a person's risk of disease, particularly some types of cancer, or response to certain drugs.

For more information about copy number variation:

The Howard Hughes Medical Institute discusses the results of recent research on copy number variation in the news release, Genetic Variation: We're More Different Than We Thought (https://www.hhmi.org/news/genetic-variation-were-more-different-we-thought).

More information about copy number variation (https://dnalc.cshl.edu/view/552-Copy-Number-Variants.html) is available in a video from Cold Spring Harbor Laboratory.

A definition of copy number variation (https://www.genome.gov/genetics-glossary/Copy-Number-Variation) is included in the Talking Genome Glossary from the National Human Genome Research Institute.

For people interested in more technical data, several institutions provide databases of structural differences in human DNA, including copy number variation:

- The Centre for Applied Genomics Database of Genomic Variants (http://dgv.tcag.ca/dgv/app/home)
- The Sanger Institute: Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER (https://decipher.sanger.ac.uk/))

6 Can changes in the number of chromosomes affect health and development?

Human cells normally contain 23 pairs of chromosomes, for a total of 46 chromosomes in each cell (Figure 8). A change in the number of chromosomes can cause problems with growth, development, and function of the body's systems. These changes can occur during the formation of reproductive cells (eggs and sperm), in early fetal development, or in any cell after birth. A gain or loss in the number of chromosomes from the normal 46 is called aneuploidy.

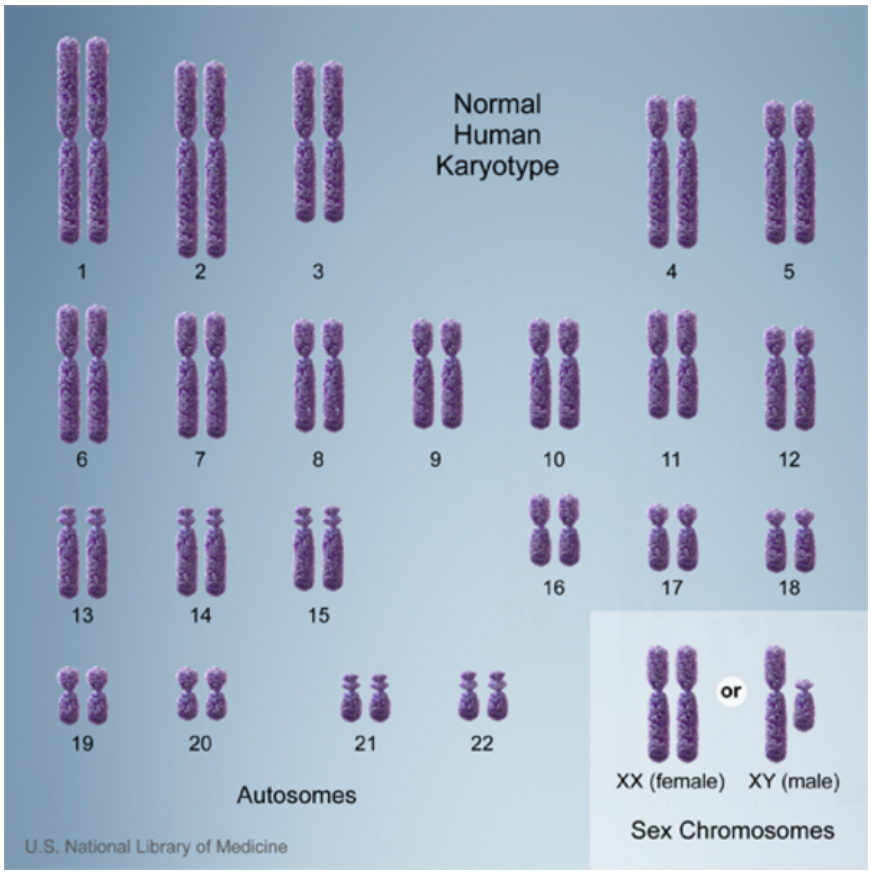

FIGURE 8: Normal human karyotype with 22 pairs of autosomes and two sex chromosomes, either two X chromosomes or an X and a Y chromosome.

A common form of aneuploidy is trisomy, or the presence of an extra chromosome in cells. "Tri-" is Greek for "three"; people with trisomy have three copies of a particular chromosome in cells instead of the normal two copies. Down syndrome (also known as trisomy 21) is an example of a condition caused by trisomy (Figure 9). People with Down syndrome typically have three copies of chromosome 21 in each cell, for a total of 47 chromosomes per cell.

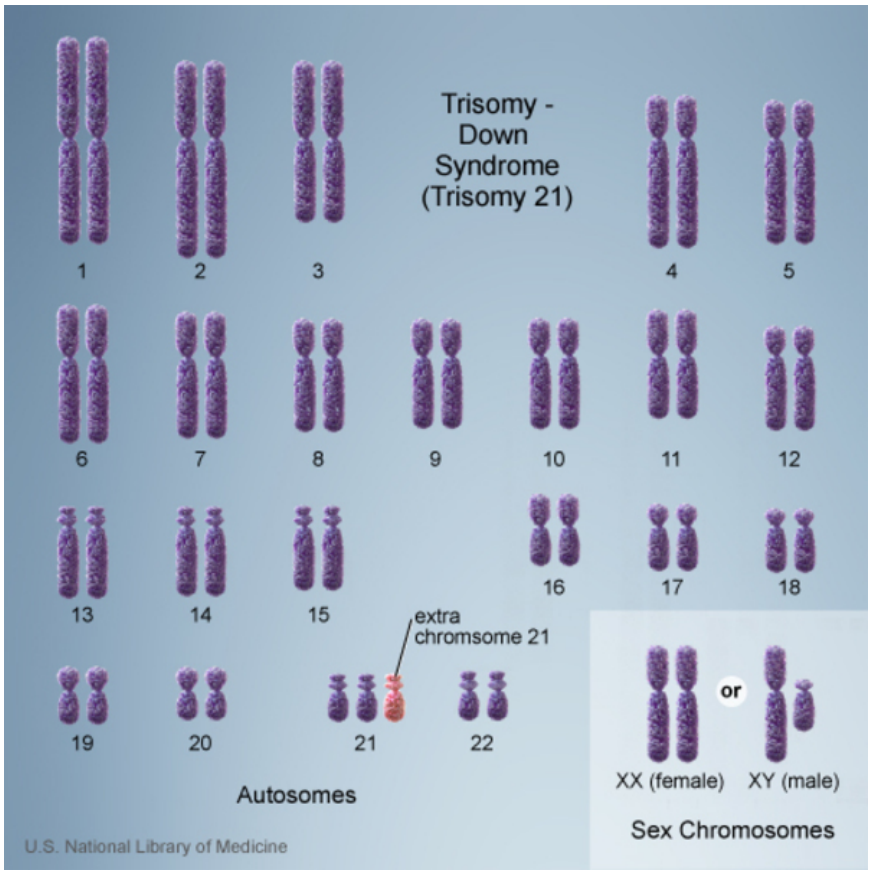

FIGURE 9: A karyotype showing an extra copy of chromsome 21, for a total of 47 chromosomes instead of the usual 46.

Monosomy, or the loss of one chromosome in cells, is another kind of aneuploidy. " Mono-" is Greek for "one"; people with monosomy have one copy of a particular chromosome in cells instead of the normal two copies. Turner syndrome (also known as monosomy X) is a condition caused by monosomy (Figure 10). Women with Turner syndrome usually have only one copy of the X chromosome in every cell, for a total of 45 chromosomes per cell.

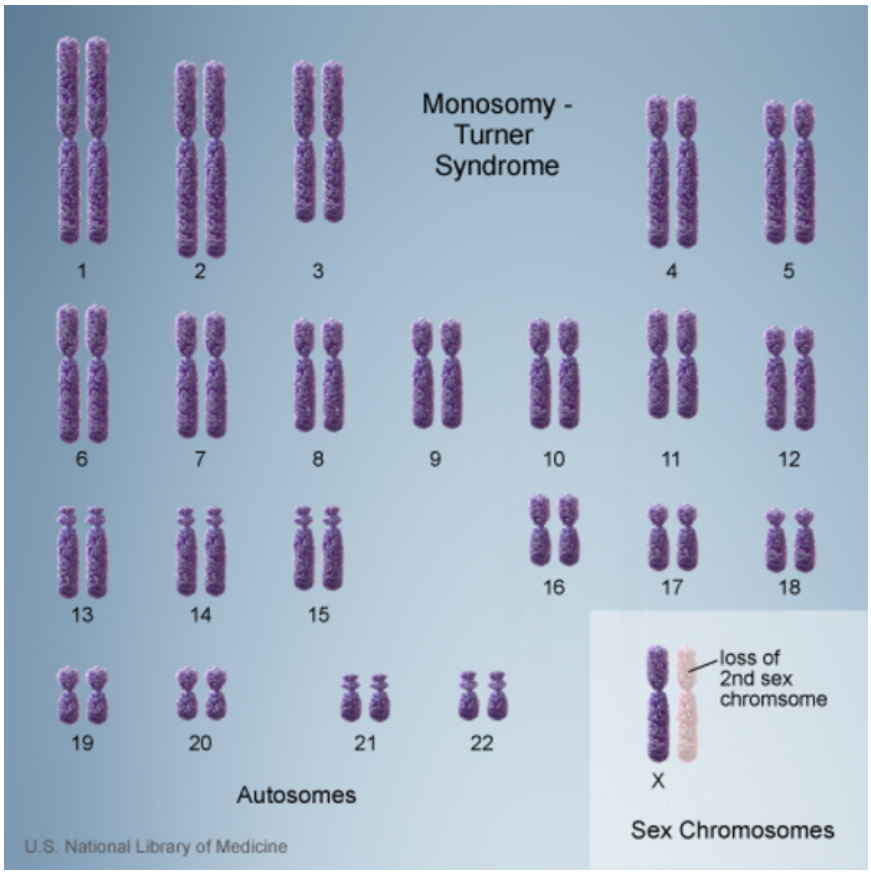

FIGURE 10: A karyotype showing a form of monosomy called Turner syndrome, in which the second sex chromosome is missing.

Rarely, some cells end up with complete extra sets of chromosomes. Cells with one additional set of chromosomes, for a total of 69 chromosomes, are called triploid (Figure 11). Cells with two additional sets of chromosomes, for a total of 92 chromosomes, are called tetraploid. A condition in which every cell in the body has an extra set of chromosomes is not compatible with life.

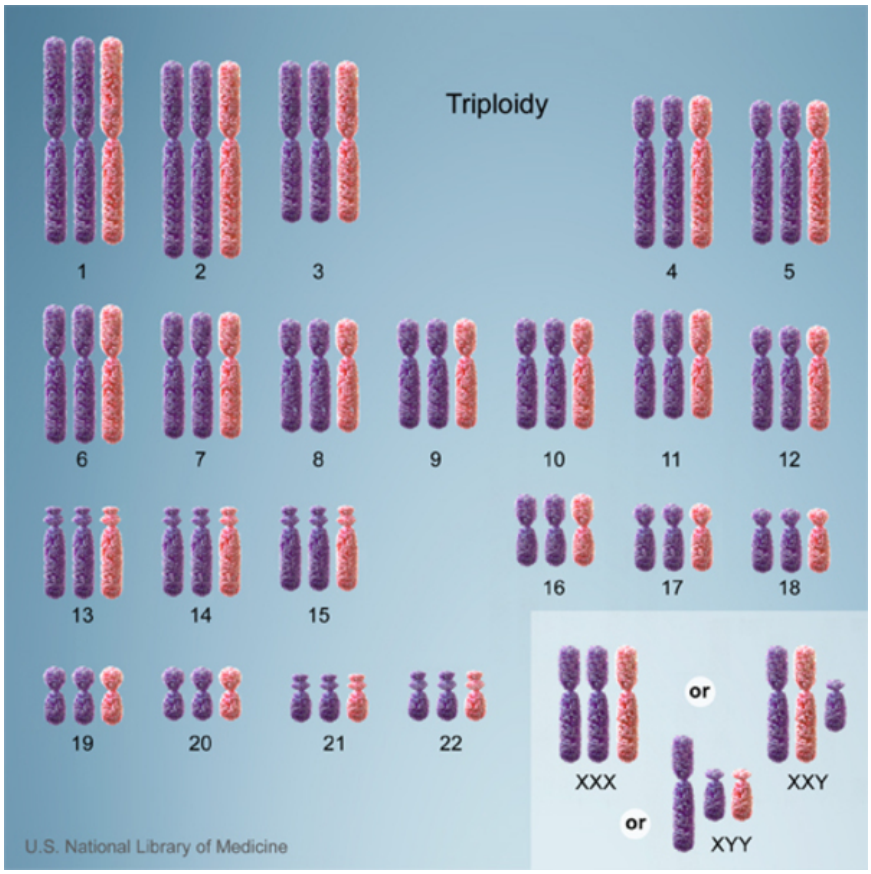

FIGURE 11: A karyotype with one extra copy of each chromosome, for a total of 69 chromosomes instead of the usual 46.

In some cases, a change in the number of chromosomes occurs only in certain cells. When an individual's cells differ in their chromosomal makeup, it is known as chromosomal mosaicism (Figure 12). Chromosomal mosaicism occurs from an error in cell division in cells other than eggs and sperm. Most commonly, some cells end up with one extra or missing chromosome (for a total of 45 or 47 chromosomes per cell), while other cells have the usual 46 chromosomes. Mosaic Turner syndrome is one example of chromosomal mosaicism. In females with this condition, some cells have 45 chromosomes because they are missing one copy of the X chromosome, while other cells have the usual number of chromosomes.

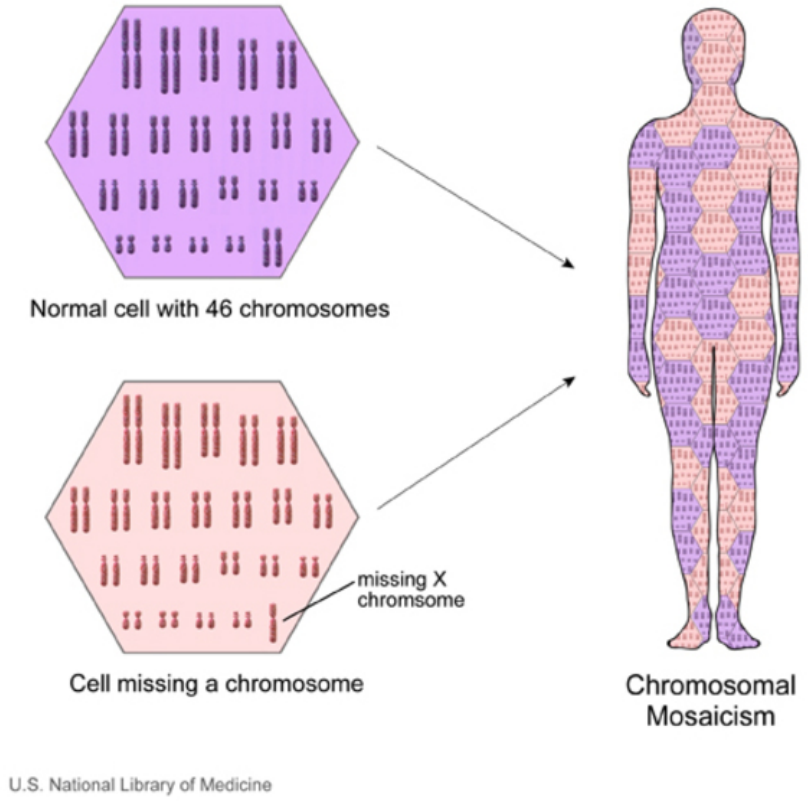

FIGURE 12: A person with mosaicism has some populations of cells with a chromsomal change and other cells without the change.

Many cancer cells also have changes in their number of chromosomes. These changes are not inherited; they occur in somatic cells (cells other than eggs or sperm) during the formation or progression of a cancerous tumor.

For more information about chromosomal disorders:

MedlinePlus provides a medical encyclopedia article about chromosomal mosaicism (https://medlineplus.gov/ency/article/001317.htm).

A discussion of how chromosomal abnormalities happen (https://www.genome.gov/about-genomics/fact-sheets/Chromosome-Abnormalities-Fact-Sheet) is provided by the National Human Genome Research Institute.

The Centre for Genetics Education offers a fact sheet about changes in chromosome number or size (https://www.genetics.edu.au/SitePages/Chromosome-changes.aspx).

The University of Leicester's Virtual Genetics Education Center provides an explanation of numerical chromosome aberrations (https://www2.le.ac.uk/projects/vgec/healthprof/topics/patterns-of-inheritance/chromosomal-abnormalities#numerical-aberrations).

Your Genome from the Wellcome Genome Campus discusses chromosome disorders (https://www.yourgenome.org/facts/what-is-a-chromosome-disorder), including how changes in the number of chromosomes cause genetic disorders.

The National Organization for Rare Disorders offers an overview of triploidy (https://rare diseases.org/rare-diseases/triploidy/).

7 Can changes in the structure of chromosomes affect health and development?

Changes that affect the structure of chromosomes can cause problems with growth, development, and function of the body's systems. These changes can affect many genes along the chromosome and disrupt the proteins made from those genes.

Structural changes can occur during the formation of egg or sperm cells, in early fetal development, or in any cell after birth. Pieces of DNA can be rearranged within one chromosome or transferred between two or more chromosomes. The effects of structural changes depend on their size and location, whether gene function is interrupted, and whether any genetic material is gained or lost. Some changes cause health problems, while others may have no effect on a person's health.

Changes in chromosome structure include the following:

Translocations

A translocation occurs when a piece of one chromosome breaks off and attaches to another chromosome. This type of rearrangement is described as balanced (Figure 13) if no genetic material is gained or lost in the cell. If there is a gain or loss of genetic material, the translocation is described as unbalanced (Figure 14).

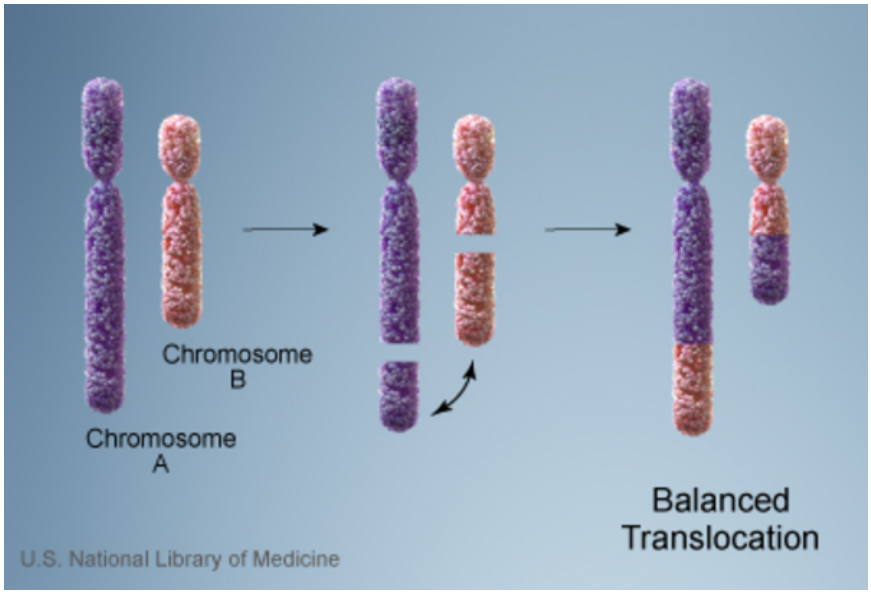

FIGURE 13: Two chromosomes break and swap pieces; no genetic material is gained or lost.

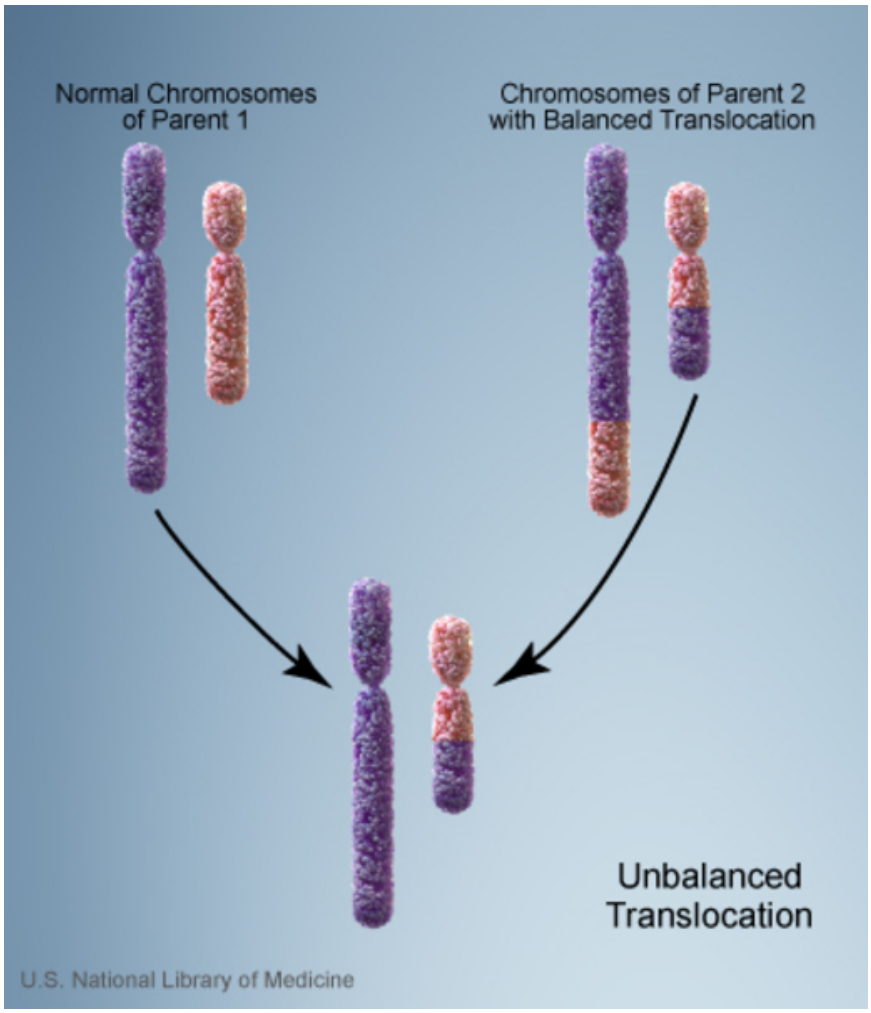

FIGURE 14: A balanced translocation can become unbalanced in the next generation, with extra or missing genetic material.

Deletions

Deletions (Figure 15) occur when a chromosome breaks and some genetic material is lost. Deletions can be large or small, and can occur anywhere along a chromosome.

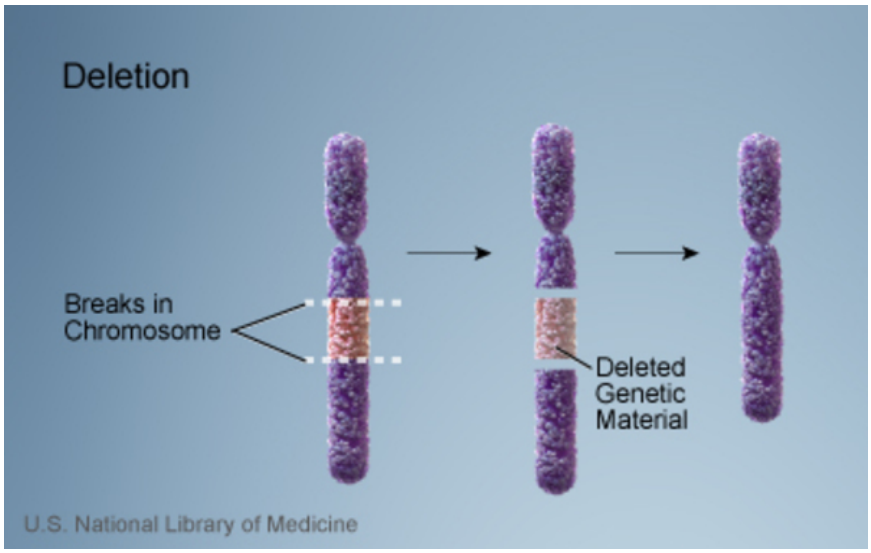

FIGURE 15: Genetic material is deleted from a chromosome.

Duplications

Duplications (Figure 16) occur when part of a chromosome is abnormally copied (duplicated). This type of chromosomal change results in extra copies of genetic material from the duplicated segment.

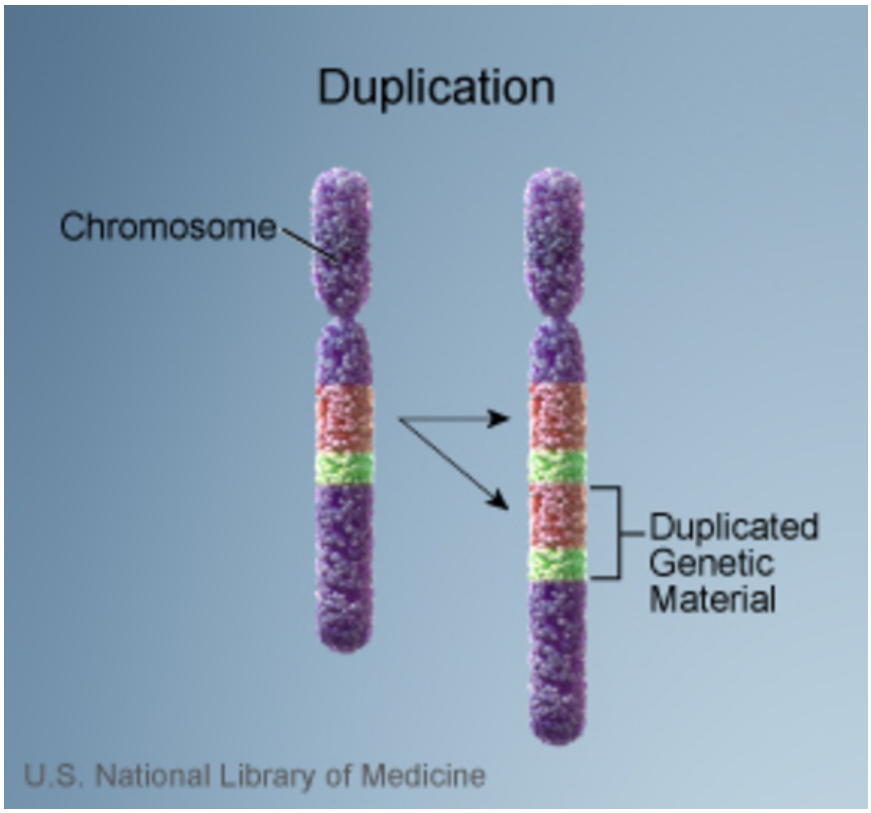

FIGURE 16: DNA from a chromosome is abnormally copied.

Inversions

An inversion (Figure 17) occurs when a chromosome breaks in two places; the resulting piece of DNA is reversed and re-inserted into the chromosome. Genetic material may or may not be lost as a result of the chromosome breaks. An inversion that includes the chromosome's constriction point (centromere) is called a pericentric inversion. An inversion that occurs in the long (q) arm or short (p) arm and does not involve the centromere is called a paracentric inversion.

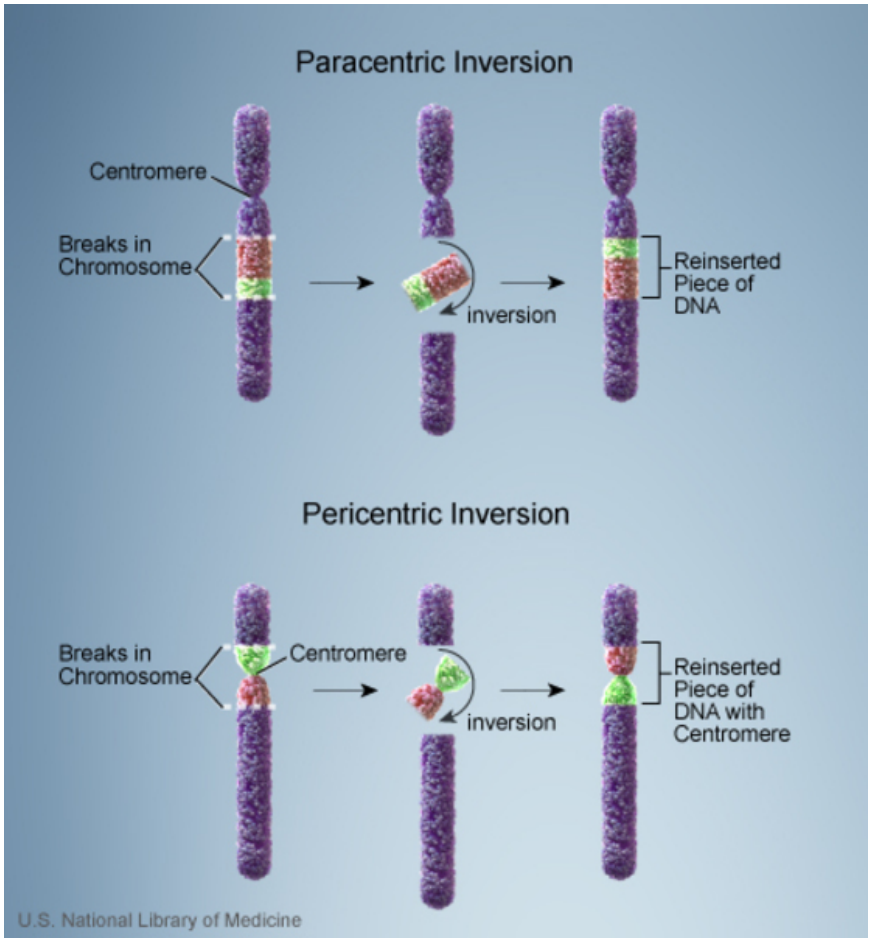

FIGURE 17: A paracentric inversion does not include the centromere; a pericentric inversion does include the centromere.

Isochromosomes

An isochromosome (Figure 18) is a chromosome with two identical arms. Instead of one q arm and one p arm, an isochromosome has two q arms or two p arms. As a result, these abnormal chromosomes have an extra copy of some genes and are lacking copies of genes on the missing arm.

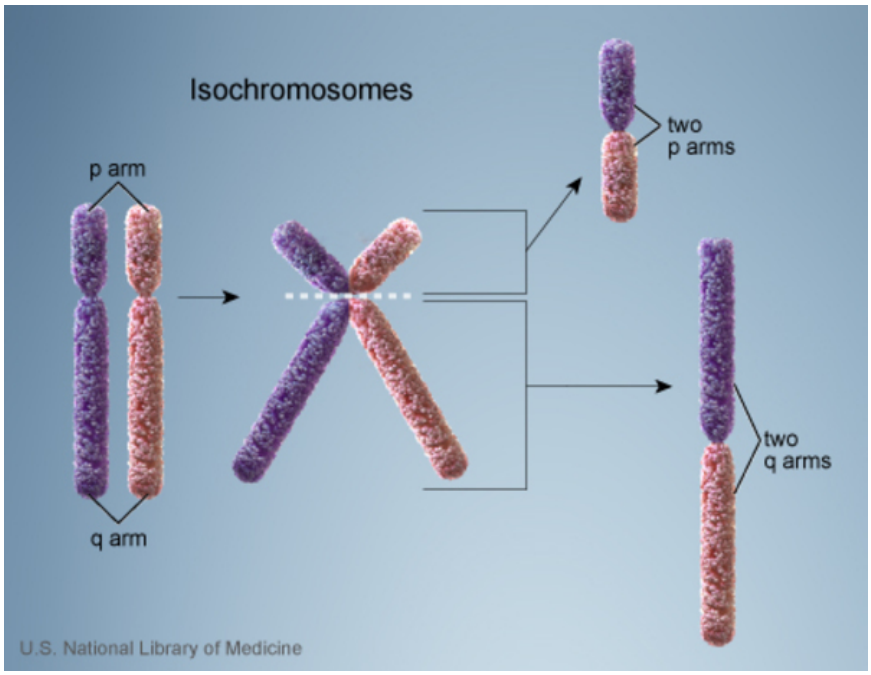

FIGURE 18: Isochromosomes have two p arms or two q arms instead of one p arm and one q arm.

Dicentric chromosomes

Unlike normal chromosomes, which have one centromere, a dicentric chromosome (Figure 19) contains two centromeres. Dicentric chromosomes result from the abnormal fusion of two chromosome pieces, each of which includes a centromere. These structures are unstable and often involve a loss of some genetic material.

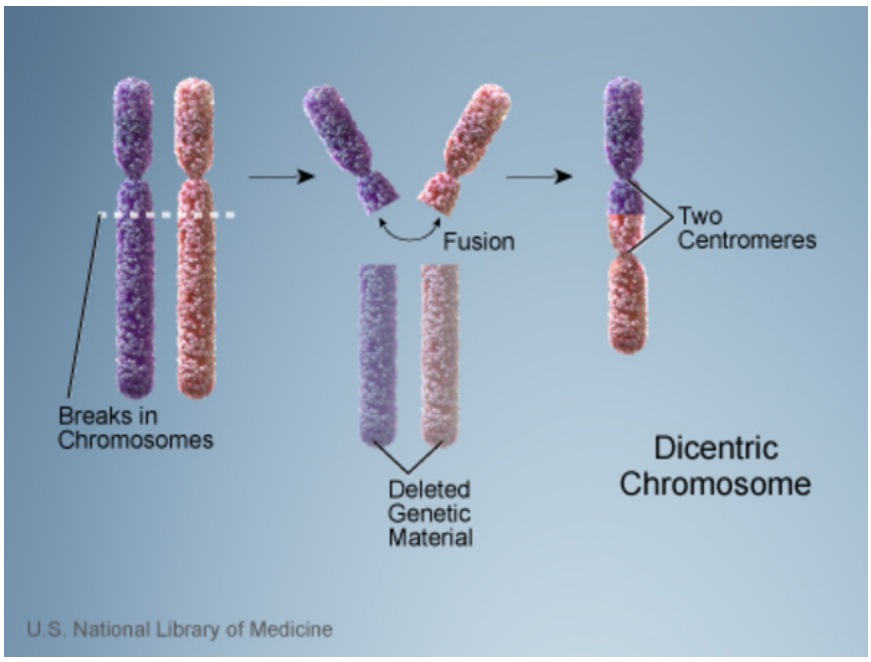

FIGURE 19: A dicentric chromosome results from the abnormal fusion of pieces of two chromosomes, creating a chromosome with two centromeres.

Ring chromosomes

Ring chromosomes (Figure 20) usually occur when a chromosome breaks in two places, typically at the ends of the p and q arms, and then the arms fuse together to form a circular structure. The ring may or may not include the centromere, depending on where on the chromosome the breaks occur. In many cases, genetic material near the ends of the chromosome is lost.

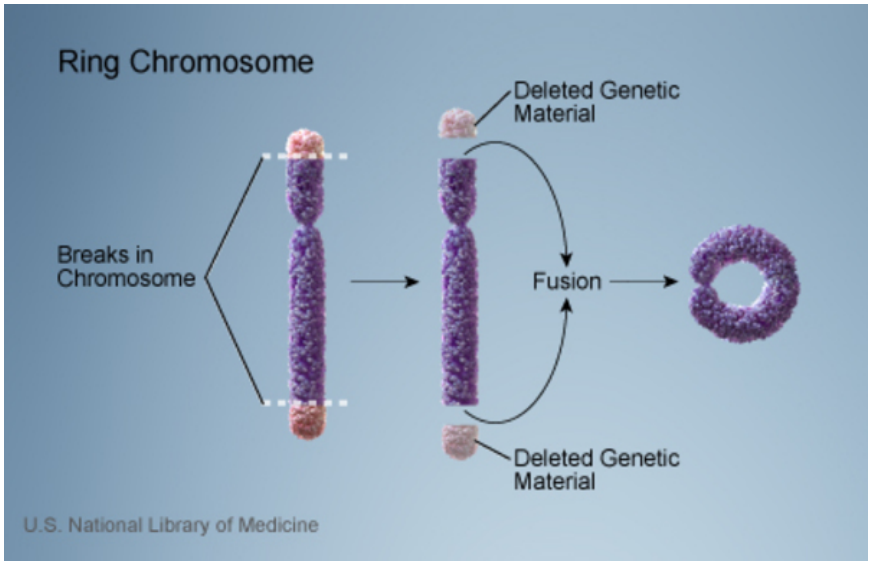

FIGURE 20: A ring chromosome occurs when the two ends of a chromosome break and then fuse together, creating a ring shape.

Many cancer cells also have changes in their chromosome structure. These changes are not inherited; they occur in somatic cells (cells other than eggs or sperm) during the formation or progression of a cancerous tumor.

For more information about structural changes to chromosomes:

The National Human Genome Research Institute provides a list of questions and answers about chromosome abnormalities (https://www.genome.gov/about-genomics/fact-sheets/Chromosome-Abnormalities-Fact-Sheet), including a glossary of related terms.

Chromosome Disorder Outreach offers a fact sheet on this topic titled Introduction to Chromosomes (https://chromodisorder.org/introduction-to-chromosomes/). This resource includes illustrated explanations of several chromosome abnormalities.

The Centre for Genetics Education provides a fact sheet about chromosome changes (https://www.genetics.edu.au/SitePages/Chromosome-changes.aspx).

The University of Leicester's Virtual Genetics Education Center provides an explanation of structural chromosome aberrations (https://www2.le.ac.uk/projects/vgec/healthprof/topics/patterns-of-inheritance/chromosomal-abnormalities#structural-aberrations).

Your Genome from the Wellcome Genome Campus discusses chromosome disorders (https://www.yourgenome.org/facts/what-is-a-chromosome-disorder), including types of structural abnormalities in chromosomes that are involved in genetic diseases.

8 Can changes in noncoding DNA affect health and development?

It is well established that changes in genes can alter a protein's function in the body, potentially causing health problems. Scientists have determined that changes in regions of DNA that do not contain genes (noncoding DNA) can also lead to disease.

Many regions of noncoding DNA play a role in the control of gene activity, meaning they help determine when and where certain genes are turned on or off. Other regions of noncoding DNA are important for protein assembly. By altering one of these regions, a variant (also known as a mutation) in noncoding DNA can turn on a gene and cause a protein to be produced in the wrong place or at the wrong time. Alternatively, a variant can reduce or eliminate the production of an important protein when it is needed. Not all changes in noncoding DNA have an impact on health, but those that alter the pattern of a critical protein can disrupt normal development or cause a health problem.

Variants in noncoding DNA have been linked to several types of cancer and developmental disorders such as isolated Pierre Robin sequence. This condition is caused by changes in regions of noncoding DNA that act as enhancer elements. Enhancers attach proteins that help turn on particular genes. The enhancers altered in isolated Pierre-Robin sequence control the activity of the *SOX9* gene.

In addition to enhancer elements, variants in noncoding DNA can disrupt other regulatory elements. These other elements include promoters, where proteins that turn on genes attach; insulators, where proteins that help shape the activity of genes in different ways attach; and silencers, where proteins that turn off genes attach.

Some regions of noncoding DNA provide instructions for making certain kinds of RNA molecules that play roles in regulating gene activity or assembling proteins. Variants that interrupt these functional RNA molecules, such as transfer RNAs, microRNAs, or long noncoding RNAs, have also been implicated in disease.

The same types of genetic changes that occur in genes or that alter the structure of chromosomes can affect health and development when they occur in noncoding DNA. These alterations include changes in single DNA building blocks (substitution variants), insertions, deletions, duplications, and translocations. Noncoding DNA variants can be inherited from a parent or acquired during a person's life.

Much is still unknown about how to identify regions of noncoding DNA that have a function in cells and the roles such regions play. As a result, linking genetic changes in noncoding DNA to their effects on certain genes and to health conditions is difficult. The roles of noncoding DNA and the effects that genetic changes in noncoding DNA have on the body are growing areas of research.

Scientific journal articles for further reading

Scacheri CA, Scacheri PC. Mutations in the noncoding genome. Curr Opin Pediatr. 2015 Dec;27(6):659-64. doi: 10.1097/MOP.0000000000000283. Review. PubMed: 26382709; Free full text from PubMed Central: PMC5084913.

Chatterjee S, Ahituv N. Gene Regulatory Elements, Major Drivers of Human Disease. Annu Rev Genomics Hum Genet. 2017 Aug 31;18:45-63. doi: 10.1146/annurev-genom-091416-035537. Epub 2017 Apr 7. Review. PubMed: 28399667.

Gordon CT, Attanasio C, Bhatia S, Benko S, Ansari M, Tan TY, Munnich A, Pennacchio LA, Abadie V, Temple IK, Goldenberg A, van Heyningen V, Amiel J, FitzPatrick D, Kleinjan DA, Visel A, Lyonnet S. Identification of novel craniofacial regulatory domains located far upstream of SOX9 and disrupted in Pierre Robin sequence. Hum Mutat. 2014 Aug;35(8):1011-20. doi: 10.1002/humu.22606. PubMed: 24934569; Free full text from PubMed Central: PMC4389788.

Read more about the role of noncoding DNA in health and disease:

University of California, San Francisco: The Mysterious 98%: Scientists Look to Shine Light on Our Dark Genome (https://www.ucsf.edu/news/2017/02/405686/mysterious-98-scientists-look-shine-light-our-dark-genome)

Duke University: Variation in “Junk” DNA Leads to Trouble (https://today.duke.edu/2016/08/variation-%E2%80%9Cjunk%E2%80%9D-dna-leads-trouble)

HEDD: Human Enhancer Disease Database (https://zdzlab.einsteinmed.edu/1/hedd.php)

9 Can changes in mitochondrial DNA affect health and development?

Mitochondria (Figure 21) are structures within cells that convert the energy from food into a form that cells can use. Although most DNA is packaged in chromosomes within the nucleus, mitochondria also have a small amount of their own DNA (known as mitochondrial DNA or mtDNA). In some cases, inherited changes in mitochondrial DNA can cause problems with growth, development, and function of the body's systems. These variants (also known as mutations) disrupt the mitochondria's ability to generate energy efficiently for cells.

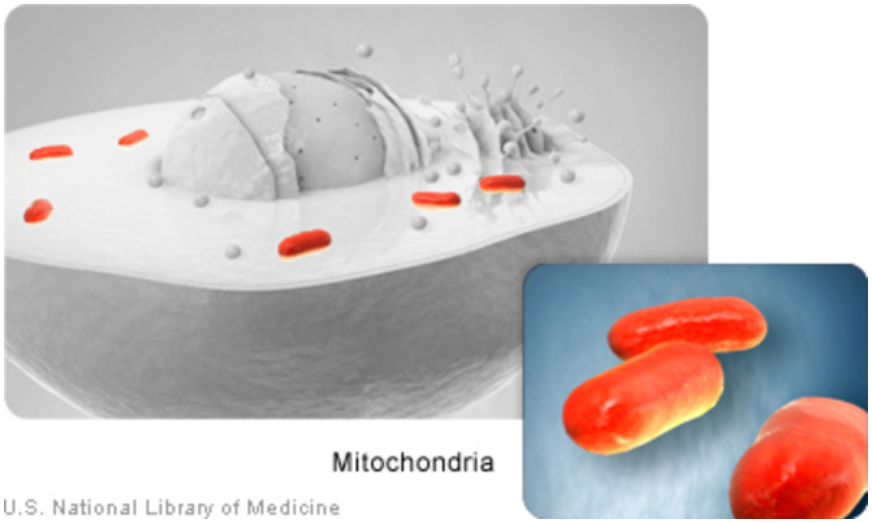

FIGURE 21: Mitochondria are structures within the cell cytoplasm.

Conditions caused by variants in mitochondrial DNA often involve multiple organ systems. The effects of these conditions are most pronounced in organs and tissues that require a lot of energy (such as the heart, brain, and muscles). Although the health consequences of inherited mitochondrial DNA alterations vary widely, frequently observed features include muscle weakness and wasting, problems with movement, diabetes, kidney failure, heart disease, loss of intellectual functions (dementia), hearing loss, and problems involving the eyes and vision.

Genetic changes that are not inherited (which are known as somatic variants) may also occur in mitochondrial DNA. Somatic variants occur in the DNA of certain cells (not sperm or egg cells) during a person's lifetime and are not passed to future generations. Because mitochondrial DNA has a limited ability to repair errors, these variants tend to build up over time. A buildup of somatic variants in mitochondrial DNA has been associated with some forms of cancer and an increased risk of certain age-related disorders such as heart disease, Alzheimer's disease, and Parkinson's disease. Additionally, research suggests that the progressive accumulation of these variants over a person's lifetime may play a role in the normal process of aging.

For more information about conditions caused by mitochondrial DNA changes:

An overview of mitochondrial disorders (https://www.ncbi.nlm.nih.gov/books/NBK1224/) is available from GeneReviews.

The Muscular Dystrophy Association offers an introduction to mitochondrial disorders as part of their fact sheet called Mitochondrial Myopathies (https://www.mda.org/disease/mitochondrial-myopathies).

The Neuromuscular Disease Center at Washington University in St. Louis provides an in-depth description of many mitochondrial disorders (https://neuromuscular.wustl.edu/mitosyn.html).

10 What are complex or multifactorial disorders?

Researchers are learning that nearly all conditions and diseases have a genetic component. Some disorders, such as sickle cell disease and cystic fibrosis, are caused by variants (also known as mutations) in single genes. The causes of many other disorders, however, are much more complex. Common health problems such as heart disease, type 2 diabetes, and obesity do not have a single genetic cause—they are influenced by multiple genes (polygenic) in combination with lifestyle and environmental factors, such as exercise, diet, or pollutant exposures. Conditions caused by many contributing factors are called complex or multifactorial disorders.

Although complex disorders often cluster in families, they do not have a clear-cut pattern of inheritance. It may be difficult to identify the role of genetics in these disorders, particularly because families often also share environments and may have similar lifestyles. This makes it difficult to determine a person's risk of inheriting or passing on these disorders. Complex disorders are also difficult to study and treat because the specific factors that cause most of these disorders have not yet been identified. Researchers continue to look for major contributing genes for many common, complex disorders.

For more information about complex disorders:

MedlinePlus (https://medlineplus.gov/) provides additional information about specific complex disorders such as diabetes and obesity and other reliable medical information

A fact sheet about the inheritance of multifactorial disorders (https://www.genetics.edu.au/SitePages/Environmental-and-Genetic-interactions.aspx) is available from the Centre for Genetics Education.

The Children's Hospital of Wisconsin provides basic information about multifactorial inheritance (https://childrenswi.org/medical-care/genetics-and-genomics-program/medical-genetics/multifactorial-inheritance) and examples of multifactorial disorders.

The National Human Genome Research Institute describes how researchers study complex disorders (https://www.genome.gov/10000865/complex-disorders-background).

The Centers for Disease Control and Prevention provides a list of diseases and conditions (https://www.cdc.gov/DiseasesConditions/) with additional information.

11 What does it mean to have a genetic predisposition to a disease?

A genetic predisposition (sometimes also called genetic susceptibility) is an increased likelihood of developing a particular disease based on a person's genetic makeup. A genetic predisposition results from specific genetic variations that are often inherited from a parent. These genetic changes contribute to the development of a disease but do not directly cause it. Some people with a predisposing genetic variation will never get the disease while others will, even within the same family.

Genetic variations can have large or small effects on the likelihood of developing a particular disease. For example, certain variants (also called mutations) in the *BRCA1* or *BRCA2* genes greatly increase a person's risk of developing breast cancer and ovarian cancer. Particular variations in other genes, such as *BARD1* and *BRIP1*, also increase breast cancer risk, but the contribution of these genetic changes to a person's overall risk appears to be much smaller.

Current research is focused on identifying genetic changes that have a small effect on disease risk but are common in the general population. Although each of these variations only slightly increases a person's risk, having changes in several different genes may combine to increase disease risk significantly. Changes in many genes, each with a small effect, may underlie susceptibility to many common diseases, including cancer, obesity, diabetes, heart disease, and mental illness. Researchers are working to calculate an individual’s estimated risk for developing a common disease based on the combination of variants in many genes across their genome. This measure, known as the polygenic risk score, is expected to help guide healthcare decisions in the future.

In people with a genetic predisposition, the risk of disease can depend on multiple factors in addition to an identified genetic change. These include other genetic factors (sometimes called modifiers) as well as lifestyle and environmental factors. Diseases that are caused by a combination of factors are described as multifactorial. Although a person's genetic makeup cannot be altered, some lifestyle and environmental modifications (such as having more frequent disease screenings and maintaining a healthy weight) may be able to reduce disease risk in people with a genetic predisposition.

For more information about genetic predisposition to disease:

The Genetic Science Learning Center at the University of Utah provides more information about calculating the risk of genetic diseases and predicting disease based on family history (https://learn.genetics.utah.edu/content/history/geneticrisk/).

More detailed information about the genetics of breast and ovarian cancer (https://www.cancer.gov/types/breast/hp/breast-ovarian-genetics-pdq), as well as other cancers (https://www.cancer.gov/publications/pdq/information-summaries/genetics), is available from the National Cancer Institute.

The National Human Genome Research Institute explains the calculation of polygenic risk scores (https://www.genome.gov/Health/Genomics-and-Medicine/Polygenic-risk-scores) and what information the scores can provide.

12 How are gene variants involved in evolution?

Evolution is the process by which populations of organisms change over generations. Genetic variations underlie these changes. Genetic variations can arise from gene variants (also called mutations) or from a normal process in which genetic material is rearranged as a cell is getting ready to divide (known as genetic recombination). Genetic variations that alter gene activity or protein function can introduce different traits in an organism. If a trait is advantageous and helps the individual survive and reproduce, the genetic variation is more likely to be passed to the next generation (a process known as natural selection). Over time, as generations of individuals with the trait continue to reproduce, the advantageous trait becomes increasingly common in a population, making the population different than an ancestral one. Sometimes the population becomes so different that it is considered a new species.

Not all variants influence evolution. Only hereditary variants, which occur in egg or sperm cells, can be passed to future generations and potentially contribute to evolution. Some variants occur during a person's lifetime in only some of the body's cells and are not hereditary, so natural selection cannot play a role. Also, many genetic changes have no impact on the function of a gene or protein and are not helpful or harmful. In addition, the environment in which a population of organisms lives is integral to the selection of traits. Some differences introduced by variants may help an organism survive in one setting but not in another—for example, resistance to a certain bacteria is only advantageous if that bacteria is found in a particular location and harms those who live there.

So why do some harmful traits, like genetic diseases, persist in populations instead of being removed by natural selection? There are several possible explanations, but in many cases, the answer is not clear. For some conditions, such as the neurological condition Huntington disease, signs and symptoms occur later in life, typically after a person has children, so the gene variant can be passed on despite being harmful. For other harmful traits, a phenomenon called reduced penetrance, in which some individuals with a disease-associated variant do not show signs and symptoms of the condition, can also allow harmful genetic variations to be passed to future generations. For some conditions, having one altered copy of a gene in each cell is advantageous, while having two altered copies causes disease. The best-studied example of this phenomenon is sickle cell disease: Having two altered copies of the *HBB* gene in each cell results in the disease, but having only one copy provides some resistance to malaria. This disease resistance helps explain why the variants that cause sickle cell disease are still found in many populations, especially in areas where malaria is prevalent.

To find out more about the role of gene variants in evolution:

National Institute of General Medical Sciences: The New Genetics Chapter 3: Life's Genetic Tree (https://www.nigms.nih.gov/education/Booklets/the-new-genetics/Pages/Home.aspx)

Learn.Genetics from the University of Utah: Evolution: DNA and the Unity of Life (https:/

/learn.genetics.utah.edu/content/evolution/)

Cold Spring Harbor Lab: Genetic Origins (http://www.geneticorigins.org/)

Understanding Evolution from the University of California Museum of Paleontology: Huntington’s Chorea: Evolution and Genetic Disease (https://evolution.berkeley.edu/evolibrary/article/0_0_0/medicine_05)

Understanding Evolution from the University of California Museum of Paleontology: Misconceptions About Natural Selection (https://evolution.berkeley.edu/evolibrary/article/evo_32)

13 What information can statistics provide about a genetic condition?

Statistical data can provide general information about how common a condition is, how many people have the condition, or how likely it is that a person will develop the condition. Statistics are not personalized, but they do offer estimates based on groups of people. By taking into account a person's family history, medical history, and other factors, a genetics professional can help interpret the statistics and explain what they mean for an individual.

Some statistical terms are commonly used when describing genetic conditions and other disorders. These terms include the following:

Incidence. The incidence of a gene variant (also called a gene mutation) or a genetic disorder is the number of people in a specified group who develop a variant or disorder during a particular time period. Incidence is often written in the form “1 in [a number]” or as a total number of a population.

Example: About 1 in 200,000 people in the United States are diagnosed with syndrome A each year. An estimated 15,000 people worldwide were diagnosed with syndrome B last year.

Prevalence. The prevalence of a gene variant or a genetic disorder is the total number of people in a specified group at a given time who are living with the variant or disorder. This term includes both newly diagnosed and pre-existing cases in people of any age. Prevalence is often written in the form “1 in [a number]” or as a total number of people who have a condition.

Example: Approximately 1 in 100,000 people in the United States have syndrome A at the present time. About 100,000 children worldwide currently have syndrome B.

Mortality. Mortality is the number of deaths from a particular disorder occurring in a specified group per year. Mortality is usually expressed as a total number of deaths.

Example: An estimated 12,000 people worldwide died from syndrome C in 2020.

Lifetime risk. Lifetime risk is the average risk of developing a particular disorder at some point during a lifetime. Lifetime risk is often written as a percentage or as “1 in [a number].” It is important to remember that the risk per year or per decade is much lower than the lifetime risk. In addition, other factors may increase or decrease a person's risk as compared with the average.

Example: Approximately 1 percent of people in the United States develop disorder D during their lifetimes. The lifetime risk of developing disorder D is 1 in 100.

For more information about understanding and interpreting statistics:

NIH News in Health offers an explanation of health statistics in their article "

Understanding Health Risks (https://newsinhealth.nih.gov/2016/10/understanding-health-risks)."

The New York Department of Health provides a basic explanation of statistical terms (ht tps://www.health.ny.gov/diseases/chronic/basicstat.htm), including incidence, prevalence, morbidity, and mortality.

More detailed information about health statistics is available from Woloshin, Schwartz, and Welch's Know Your Chances: Understanding Health Statistics (https://www.ncbi.nl m.nih.gov/books/NBK115435/), which is available through the NCBI Bookshelf.

The National Cancer Institute offers additional tools for understanding cancer statistics (https://www.cancer.gov/about-cancer/understanding/statistics).

14 How are genetic conditions and genes named?

Naming genetic conditions

Genetic conditions are not named in one standard way (unlike genes, which are given an official name and symbol by a formal committee). Doctors who treat families with a new, previously unknown disorder are often the first to propose a name for the condition. Later, healthcare professionals, researchers, people affected by the condition, and other interested individuals may come together to revise the name to improve its usefulness. Naming is important because it allows accurate and effective communication about particular conditions, which will ultimately improve care and help researchers find new approaches to treatment.

Condition names are often derived from one or a combination of sources:

- The basic genetic or biochemical defect that causes the condition (for example, alpha-1 antitrypsin deficiency);
- The gene in which the variant (or mutation) that causes the condition occurs (for example, TUBB4A-related leukodystrophy);
- One or more major signs or symptoms of the disorder (for example, hypermanganesemia with dystonia, polycythemia vera, and cryptogenic cirrhosis);
- The parts of the body affected by the condition (for example, brain-lung-thyroid syndrome);
- The name of a physician or researcher, often the first person to describe the disorder (for example, Marfan syndrome, which was named after Dr. Antoine Bernard-Jean Marfan);
- A geographic area (for example, familial Mediterranean fever, which occurs mainly in populations bordering the Mediterranean Sea); or
- The name of a patient or family with the condition (for example, amyotrophic lateral sclerosis is often called Lou Gehrig disease after the famous baseball player who was diagnosed with the condition).

Conditions named after a specific person are called eponyms. They can be in the possessive form (e.g., Alzheimer’s disease) or in the nonpossessive form (e.g., Down syndrome).

Naming genes

The HUGO Gene Nomenclature Committee (HGNC) designates an official name and symbol (an abbreviation of the name) for each known human gene. The HGNC is a nonprofit organization funded by the U.S. National Human Genome Research Institute and the UK's Wellcome Trust. The Committee has named more than 19,000 of the estimated 20,000 to 25,000 protein-coding genes in the human genome.

During the research process, genes often acquire several alternate names and symbols

from researchers investigating the same gene. To resolve this confusion, the HGNC assigns a unique name and symbol to each human gene, which allows effective organization of genes in large databanks, aiding the advancement of research. For specific information about how genes are named, refer to the HGNC's Guidelines for Human Gene Nomenclature.

Learn more about the naming of conditions and genes:

Common sources for finding the names of genetic conditions include PubMed (https://pubmed.ncbi.nlm.nih.gov/) , GeneReviews (https://www.ncbi.nlm.nih.gov/books/NBK1116/) , OMIM (https://omim.org/), GARD (https://rarediseases.info.nih.gov/) , MeSH (https://meshb.nlm.nih.gov/) , and Orphanet (https://www.orpha.net/).

The official names and symbols of all human genes are found on the HUGO Gene Nomenclature Committee (https://www.genenames.org/) (HGNC) website.