ALS Canada Genetics Glossary

Genetic-related acronyms

- **ASO** – antisense oligonucleotide
- **atDNA** – autosomal DNA
- **bp** – base pair
- **CNV** – copy number variation
- **DNA** – deoxyribonucleic acid
- **GTR** – genetic testing registry
- **GWAS** – genome wide association study
- **INDEL** – insertion or deletion
- **fALS** – familial ALS
- **FTD** – frontotemporal dementia
- **FUS** – fused in sarcoma
- **mRNA** – messenger RNA
- **nDNA** – nuclear DNA
- **NGS** – next-generation sequencing
- **PCR** – polymerase chain reaction
- **RNA** – ribonucleic acid
- **sALS** – sporadic ALS
- **SOD1** – superoxide dismutase 1
- **TDP-43** – TAR DNA-binding protein 43
- **SNP** – single nucleotide polymorphism
- **WES** – whole exome sequencing
- **WGS** – whole genome sequencing

Glossary of Genetic Terms

A

Adenine (A) - One of the four nucleotide bases in DNA and RNA. In DNA, it pairs specifically with the base thymine (T), and in RNA, it pairs with uracil (U).

Allele - The alternative versions of a gene. An individual will inherit **two alleles for each gene**, one from each parent. For example, different alleles explain why individuals have different blood types: they might have the A, B, or O blood type allele.

Amino acid - The 20 building blocks from which proteins are assembled. Just like letters of the alphabet can be combined to form a nearly endless variety of words, amino acids can be linked together to form a nearly endless variety of proteins.

Antisense Oligonucleotide (ASO) - Small pieces of DNA or RNA that can bind to specific molecules of RNA, blocking their ability to make proteins or perform other functions. Antisense oligonucleotides are being studied for the treatment of certain forms of ALS. Tofersen (Qalsody), a therapy approved in the USA and Europe to treat SOD1-ALS, is an ASO.

Autosomal DNA (atDNA) - DNA inherited from the autosomal chromosomes, rather than from the sex chromosomes.

Autosome - Non-sex chromosomes. Humans have 22 pairs of autosomes.

***ATXN2* -** The *ATXN2* gene provides instructions for making a protein called ataxin-2. This protein is involved in RNA regulation inside cells. Some variants in this gene are associated with an increased risk for ALS.

B

Base pair (bp) - Two nucleotide bases paired together in a DNA or RNA sequence. In DNA, adenine (A) pairs with thymine (T), and cytosine (C) pairs with guanine (G). In RNA, adenine (A) pairs with uracil (U), and cytosine (C) pairs with guanine (G).

C

Candidate Gene - A gene that is believed to be associated with a particular trait or disease, often based on its known biological function or its location in the genome.

Carrier - An individual carrying one copy of a genetic variant that is inherited in an autosomal recessive or sex-linked manner and does not manifest the disease. In some cases, carriers can have mild symptoms. Carriers can pass the variant to their children, even if they never manifest the disease.

Codon - A sequence of three nucleotide bases in DNA or RNA that encode for different amino acids. In humans, there are 64 different codons, each corresponding to an amino acid or a stop signal in protein synthesis.

Copy Number Variation (CNV) - A phenomenon where the number of copies of a specific gene or DNA sequence varies between individuals. CNVs are common features of the human genome and contribute to genetic diversity between individuals and groups.

Chromosome - Within the nucleus of every cell, our DNA is packaged up tightly into coiled-structures called chromosomes. Humans typically have 23 pairs of chromosomes, for a total of 46, inheriting half from each parent. These structures unravel when the DNA needs to be read by the cell.

CRISPR-Cas9 - CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats. CRISPR/Cas9 is a gene-editing technology used in a laboratory to modify DNA sequences in living organisms, allowing researchers to remove, add, or alter parts of the sequence.

Cytoplasm - The jelly-like material that fills the inside of a cell. The cytoplasm houses all the different parts of cells (called organelles). It is typically considered one of the two major compartments of cells, with the other being the nucleus.

Cytosine - One of the four nucleotide bases in DNA and RNA. It pairs specifically with the base guanine (G).

***C9ORF72* -** The *C9ORF72* gene provides instructions for making a protein called C9ORF72 (chromosome 9 open reading frame 72), found in many regions of the brain. Unique variants in *C9ORF72* have been identified as the most common cause of ALS and frontotemporal dementia (FTD). The variant is referred to as a repeat expansion variant, as the sequence of nucleotides GGGGCC is repeated multiple times. Variants in the *C9ORF72* gene are linked to approximately 25-40% of ALS cases with a family history, and 5-10% of ALS cases with no obvious family history.

D

Deletion - A type of genetic variant when one or more DNA or RNA bases are removed from the sequence. This can affect how the sequence is read and impact protein production.

DNA - A double-stranded molecule composed of nucleic acids, formed by base pairs. Our DNA sequence constitutes our genome, containing all the information about our traits and normal body functioning.

DNA Sequencing - Techniques used in a laboratory to determine the exact sequence of nucleotide bases in a DNA sequence.

Dominant pattern of inheritance - With dominant traits, only one copy of the gene containing the variant needs to be present for the trait or disease to be expressed in an individual. Most ALS-linked genes have a dominant pattern of inheritance, meaning if the parent has one copy of the variant, the child has a 50% chance of inheriting the gene variant. If the child inherits the gene variant, they are at chance of developing symptoms of the disease.

E

Environmental Risk Factors - Factors in an individual's environment that can increase their susceptibility or predisposition to disease. Examples include smoking or exposure to radiation.

Epigenetics - The study of how an individual's behavior and environment can cause changes in the way their DNA sequence is read. Epigenetic changes are reversible modifications, as they do not alter the DNA sequence itself but rather how the cell machinery accesses and reads each gene.

Exon - Regions of the genome that code for messenger RNAs (mRNAs). Most exons will code for proteins. Collectively, they make up a person's exome.

Exome - The complete set of exons in the genome.

F

Familial ALS (fALS) - The inherited form of ALS. Approximately 5-10% of all ALS cases show a family history indicating the disease is triggered by variants in an ALS-associated gene passed from parent to offspring.

Family History - Past documented cases of the disease in an individual's family, usually meaning multiple family members were affected by the disease, due to a hereditary variant in a gene related to the condition.

First-Degree Relative - An individual's parent, sibling, or child.

Frontotemporal dementia (FTD) - A group of disorders caused by progressive loss of neurons in the brain's frontal lobes (the areas behind your forehead) or its temporal lobes (the regions behind your ears). Symptoms include difficulties in critical thinking and problem solving, language/speech deficits and behavioral problems. FTD symptoms can exist on a spectrum with ALS, especially related to some cases of genetic ALS, where people can range from very mild, almost undetectable cognitive symptoms to a full FTD diagnosis.

***FUS* -** The *FUS* gene provides instructions for making a protein called FUS (Fused in Sarcoma, which is believed to contribute to normal cell function by playing a role in the biology between RNA and proteins. Variants in the *FUS* gene are linked to approximately 3-6% of ALS cases with a family history, and 1% of ALS cases with no obvious family history.

G

Gene - A segment of DNA that carries the genetic instructions for making a specific protein or RNA molecule. For example, the *SOD1* gene will hold the instructions to make the SOD1 protein, which has an important function in cells and can cause ALS if there is an error (variant) in that code. Genes are passed from parent to child and control the characteristics that the individual will have. Genes are bundled into chromosomes in the nucleus of cells.

Gene Therapy - The delivery of nucleic acids into a patient's cells as a treatment for a disease. Gene therapy is currently being explored to treat many diseases, including certain types of cancers as well as neurodegenerative diseases, such as ALS.

Genetic Counselling: Guidance and support relating to genetic disorders and genetic testing for an individual or a family. Pre- and post-counselling are always recommended for an individual considering genetic testing for ALS.

Genetic Predisposition - An inherited increase in the risk of developing a disease. Also known as genetic susceptibility.

Genetic Testing – A type of medical test that identifies variants in an individual's chromosomes, genome (DNA sequence), or gene expression. The results of a genetic test can confirm or rule out a suspected genetic condition or help to determine a person's chance of developing or passing on a genetic disorder.

Genome - The complete set of genetic material (DNA sequence) present in a person. It comprises all 23 pairs of chromosomes.

Genotype – An individual's unique DNA sequence (genome).

Genome-wide association study (GWAS) – This method of study looks at the entire DNA sequence (the genome) of a group of people. These types of studies help researchers identify genes associated with a disease or trait.

Guanine - One of the four nucleotide bases in DNA and RNA (abbreviated as G). It pairs specifically with the base cytosine (C).

H

Hereditary - In medicine, describes the passing of genetic information from parent to child through the genes in sperm and egg cells. Also known as inherited.

Heterogeneity - The phenomenon by which the same disease can be caused, or contributed to, by a variety of different genes. The disease characteristics of ALS are also said to be heterogeneous as many people living with ALS will have varying ages of onset, rates of progression, and areas of onset (e.g. bulbar vs. limb), etc.

Heterozygous - An individual that inherited different versions (alleles) of a gene, one from each parent**. "**Hetero" means different.

Homozygous – An individual that inherited the same versions (alleles) of a gene, one from each parent**. "**Homo" means same.

I

Insertion or deletion (INDEL) – An insertion or deletion of a part of the DNA or RNA sequence. They can range from a single base to multiple bases. Indels ≥ 50 bases in length are classified as structural variants.

Intron – A segment of DNA that is not transcribed into the mRNA sequence containing the information to make proteins. They can be referred to as "non-coding regions."

K

Knockout – Gene knockouts (KO) are used in genetic engineering to remove or silence a specific gene. This is commonly used in research to assess the function of a gene by observing the effects of their absence or reduced activity.

M

Messenger RNA (mRNA) – To make proteins, the DNA sequence is read and transcribed into mRNA. This process is called transcription. The resulting mRNA carries the instructions to make a protein.

Mutagen: A chemical or physical agent that can induce variants in a DNA sequence. Examples include tobacco products or radiation.

Mutation - A mutation is now commonly referred to as a **"variant."** A variant is a change in the DNA sequence. This variation can be benign (does not impact an individual's health), pathogenic (increases an individual's susceptibility to the disease), or of uncertain significance (not enough information to assess if it is benign or pathogenic). Everyone's DNA sequence has many variants, and most are harmless.

N

Nucleus: A spherical structure inside of the cell that contains a person's DNA sequence.

Nuclear DNA (nDNA) – DNA found only inside the nucleus (the central part of the cell), and not **i**n the cytoplasm. Nuclear DNA is our genome.

Nucleic Acid – Polynucleotides (a chain of nucleotides). DNA and RNA are classes of nucleic acids.

Nucleotide – A base or "building block" that in a sequence, assembles DNA or RNA. DNA bases are adenine (A), cytosine (C), guanine (G), and thymine (T). RNA bases are A, C, G, and uracil (U).

Next-generation sequencing (NGS) – A technology that reads and sequences the genome. It can be used to detect genetic variants.

P

Pattern of inheritance – Describes how a trait or variant is inherited. The most common types are dominant, recessive, autosomal, or X-linked. The majority of ALS variants show an autosomal dominant inheritance, but documented cases have also shown autosomal recessive, or X-linked inheritances.

Polymerase chain reaction (PCR) – A technique used in a laboratory that allows for the amplification of a DNA fragment. PCR is essential for generating large quantities of DNA for various applications, including genetic analysis and research.

Penetrance - The proportion of individuals in a population who carry a specific genetic variant and develop clinical symptoms of the associated disease. For example, if a genetic variant is said to have 80% penetrance, then 80% of people with the variant will develop the disease and 20% will not.

Phenotype - The observable physical characteristics or traits of an individual, influenced by both genetic and environmental factors. ALS is often described as a disease with variable phenotypes, as symptoms can differ amongst people living with the disease.

Project MinE - A large-scale global research initiative aimed to determine the genetic basis of ALS. Find out more here and here.

Protein - The molecules responsible for almost all cellular functions (often referred to as the workhorses of the cell). Proteins consist of long chains of smaller units called amino acids. DNA contains the code for protein formation in cells.

R

Recessive pattern of inheritance – With recessive traits, two copies of the gene containing the variant need to be inherited for the trait or disease to be expressed (one from each parent). This means that in a recessive pattern of inheritance, both parents need to have at least one copy of the variant, and both of those need to be passed on, for the child to be affected. If only **one copy** of the variant is inherited, the trait (or disease) will not manifest. In this case, the child will be considered a **carrier.**

Recombinant - In genetics, describes DNA, proteins, cells, or organisms that are made by combining genetic material from two different sources. Recombinant substances are made in the laboratory and have applications in research, medicine, and biotechnology.

Repeat expansion variants – Genetic variations characterized by the abnormal repetition of a specific sequence of DNA. These expansions can disrupt normal gene function, alter protein production, and potentially lead to disease. In ALS, the *C9*ORF*72* gene is affected by a repeat expansion (the nucleotide sequence GGGGCC is repeated multiple times).

Ribonucleic Acid (RNA) - A single-stranded nucleic acid molecule. RNA acts as a messenger molecule, carrying genetic information from the nucleus to the cytoplasm where protein assembly occurs.

S

Sex Chromosome – Non-autosome chromosomes. Humans will have one pair of sex chromosomes: XX (individuals assigned female at birth) or XY (individuals assigned male at birth).

***SOD1* –** The *SOD1* gene provides instructions for making a protein called SOD1 (superoxide dismutase 1), which plays an important role in reducing oxidative stress within cells. SOD1 destroys DNA-damaging superoxide, a highly reactive form of oxygen (referred to as a free radical). Variants in *SOD1* are linked to approximately 10-30% of ALS cases with a family history and 1-4% of ALS cases with no obvious family history.

Single Nucleotide Polymorphism (SNP) – A common change or substitution of one single DNA base (nucleotide) in the genome, that is present in at least 1% of the population. For example, the base guanine (G) may be changed to a cytosine (C). SNPs can be used as biological markers, helping researchers locate genes associated with a certain disease.

Susceptibility – Refers to an increased chance of developing a disease due to one or more genetic variants and/or a family history suggestive of an increased risk of the disease. Susceptibility factors increase the chances of disease occurrence but do not guarantee it.

T

***TARDP* -** The *TARDP* gene provides instructions for making a protein called TDP-43, which is an RNA-binding protein involved in ALS. Variants in the *TARDBP* gene are linked to approximately 4-5% of ALS cases with a family history, and 1% of ALS cases with no obvious family history. However, TDP-43 dysfunction is observed in about 98% of people living with ALS, even without a known *TARDBP* variant.

Thymine - One of the four nucleotide bases in DNA (abbreviated as T). It pairs specifically with the base adenine (A) in DNA. It is not an RNA base.

Tofersen – Sold under the brand name Qalsody, tofersen is an approved therapy in the US and Europe for the treatment of SOD1-ALS. It is an ASO (antisense oligonucleotide).

Transgenic - Describes a living organism that has been genetically engineered to contain DNA from another organism. Transgenic mice, for example, are often used in preclinical ALS research as a model to study the disease.

U

Uracil - One of the four nucleotide bases in RNA (abbreviated as U). It pairs specifically with the base adenine (A) in RNA. It is not a DNA base.

V

Variant - A change or variation in the DNA sequence (genetic code). Variants can be benign (having no impact on an individual's health), pathogenic (increases an individual's susceptibility to disease), or of uncertain significance (not enough information to assess if it is benign or pathogenic). The term "mutation" was traditionally used to refer to pathogenic variants. Everyone's DNA sequence has many variants, and most are harmless.

Variant of Uncertain Significance (VUS) - A variation or change in the DNA sequence that has an unknown effect on an individual's health.

W

Whole Exome Sequencing (WES) - The process of reading a person's complete sequence of exons (exome) at a single time.

Whole Genome Sequencing (WGS) - The process of reading a person's complete DNA sequence (genome) at a single time.