

Chapter 2

Why Telomeres Matter

Roger R. Reddel, MBBS, PhD (rreddel@cmri.org.au)
University of Sydney Children's Medical Research Institute

Introduction

Biologists have studied telomeres for many decades – from long before their connection with human disease was known. This chapter provides a background for the chapters that follow by summarizing information obtained from telomere research. It begins with an overview of a few key aspects of biology required to understand what telomeres are, how they function, and the processes that protect and maintain them, and it then describes how defects in these processes can cause disease.

Types of Cells

Organisms are composed of microscopic building blocks called cells and products (like the minerals and proteins that help form bone) that cells export into their environment. Each human contains many trillions of cells.

In the adult body, there are many types of cells with specialized roles, and many different shapes and sizes to suit their functions. One example is peripheral nerve cells, which are very long so they can carry electrical impulses from the spinal cord to a muscle. Another is red blood cells, which are essentially small bags containing the oxygen-binding protein hemoglobin that get pumped around the body inside blood vessels, carrying oxygen from the lungs to other tissues.

One fundamental distinction between cell types relates to whether they are designed to pass on genetic information to the next generation or not. Ova and sperm, the cells that fuse to form a zygote and start the process of embryo formation, and the specialized cells that give rise to ova and sperm, are referred to as germline cells. All other cells in the body are called somatic cells.

Why Cells Need to Make Copies of Themselves

Cells replicate themselves via a process which usually involves growing in size and then dividing into two cells. The process is known as cell division, which means that this is one situation where division is the same as multiplication!

Throughout our lifetime, our cells need to replicate themselves a large number of times. We start off as a single fertilized egg cell (zygote), which needs to undergo a very large number of cell divisions to produce a fully formed baby. Growth of a baby into an adult also occurs through the production of enormous numbers of new cells.

Further increasing the need for cell replication, most types of cells are replaced – in some cases many times over – during a normal life span. There are some exceptions, like certain types of nerve cells that are made early on in embryonic development and then can last for a lifetime, but cells of many other types are broken down and replaced by new ones.

Some of our cells are replaced regularly, which may be thought of as “programmed maintenance.” One example is red blood cells, which are replaced with new cells every four months on average; the old ones are broken down and their molecular components are recycled.

Another example is skin cells, which are programmed to die off and shed; they are continuously replaced by new cells. The lining of the bowel is also being continuously replaced. Other types of cells are either replaced when they are damaged or are produced to meet a particular demand. For example, certain types of white blood cells are produced in large numbers when they are needed to fight an infection, and then mostly die off when their job is completed.

The Molecules of Life

Most cells contain a complete copy of the entire instruction set required for them to function and interact with each other. This is encoded in a very long molecule – DNA (deoxyribonucleic acid). In contrast to the binary code used for computers (which consists of long strings of ones and zeros), DNA contains four component nucleotides represented by the letters A (adenine), C (cytosine), G (guanine), and T (thymine), which constitute a four-letter code. The complete set of DNA nucleotides is called the genome, and the DNA of the human genome is divided into 46 pieces called chromosomes. Twenty-three of these chromosomes come from one parent, and the other 23 from the other parent (see Chapter 5, Genetic Counseling for Families). Twenty-two chromosomes are paired (essentially slightly different versions of the same genetic

information) and referred to as autosomes; the remaining two chromosomes are the sex chromosomes, X and Y, which are similar in females (XX), but not in males (XY).

There are many other types of molecules within a living cell, but here just two other types will be mentioned: proteins and RNA (ribonucleic acid) molecules.

Proteins are particularly important because they form much of the intricate machinery that carries out a huge range of processes required for life, including copying DNA and other molecules when cells are getting ready to divide. The precise makeup of individual proteins is encoded within specific regions (genes) of the DNA. Genes are used as the templates for production of RNA molecules, which are transported to areas within the cell responsible for manufacturing proteins (Figure 1).

Therefore, RNA acts as an intermediary molecule, taking instructions from the genome to the places where proteins are made. Biologists refer to the process of making RNA from the DNA code as transcription, and the process of making proteins according to the instructions in the RNA as translation. There are two functioning copies (alleles) of most genes, with many of the genes on the sex chromosomes being notable exceptions.

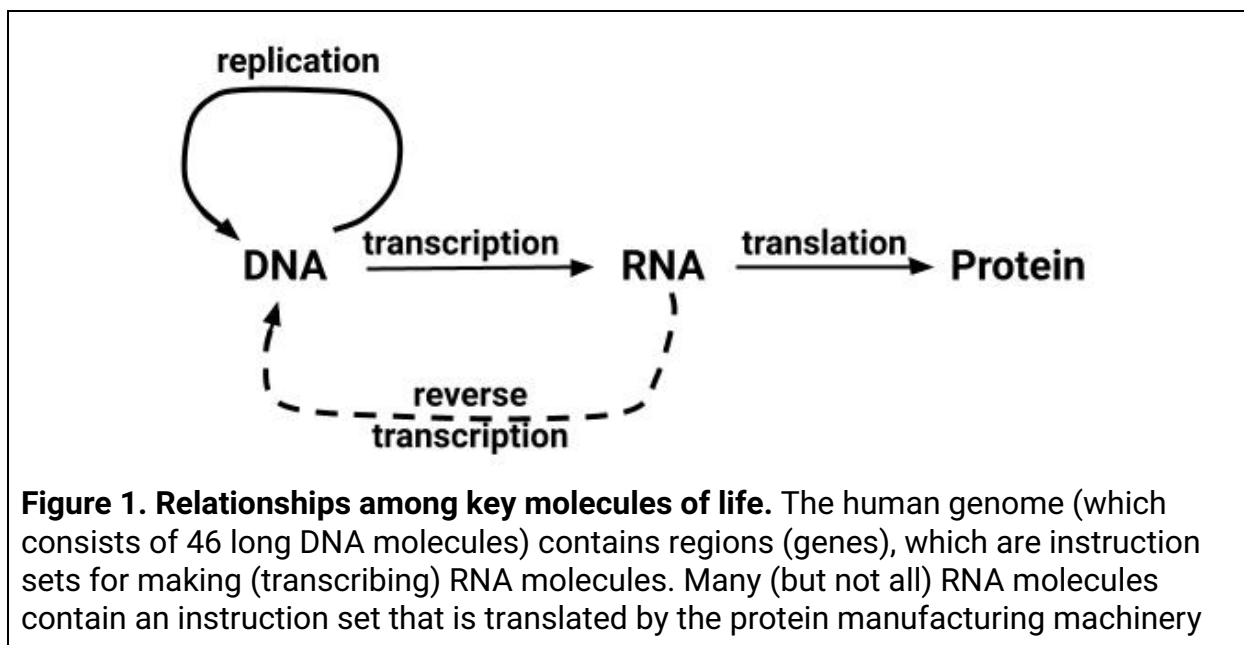


Figure 1. Relationships among key molecules of life. The human genome (which consists of 46 long DNA molecules) contains regions (genes), which are instruction sets for making (transcribing) RNA molecules. Many (but not all) RNA molecules contain an instruction set that is translated by the protein manufacturing machinery

into a specific protein. Specialized proteins and RNA molecules assemble into large molecular machines that make a replicate copy of a cell's DNA (using the existing DNA as a template) when it is getting ready to divide into two cells. Sometimes, an RNA molecule is used as the template for making relatively small pieces of DNA, a process known as reverse transcription.

The way DNA gets copied when cells are being replicated is particularly interesting. DNA has two side-by-side strands, twisted to form the well-known double helix structure discovered by Watson and Crick. The strands are not identical; instead, they are complementary to each other. The relationship between the two strands in the double helix is determined by a very simple rule: wherever one strand has an A the other strand must have a T (and vice versa), and wherever one strand has a C the other strand must have a G (and vice versa).

To copy DNA, specialized proteins pull the strands apart and synthesize a new, complementary strand on each of the existing strands. This means that each DNA copy contains one pre-existing strand plus one newly synthesized strand.

Because RNA molecules are made according to the DNA code (but with different chemical building blocks), they can sometimes be "read backwards" by other specialized proteins to synthesize DNA. This process is called reverse transcription.

Telomeres

The arrangement of the human genome as a collection of 46 chromosomes, each with two ends, presents two big challenges that cells need to deal with. The first challenge is that the DNA copying machinery is incapable of copying all the way to the ends of a DNA molecule. The consequence is that chromosomes get slightly shorter every time cells are replicated.

The second is that cells need to be able to distinguish these 92 ends from accidental breaks elsewhere in the genome. Breaks in DNA have potentially serious consequences

for cells, and cells therefore have elaborate sets of machinery for rejoining fractured ends. It is very important, however, that ends of different chromosomes do not get "repaired" by being inappropriately joined together. This would result in one larger chromosome, which is prone to breakage at a random location, destabilizing the genome.

Cells solve these two problems in the following ways. First, there is a specific DNA code at every chromosome end, which consists of a string of six letters (TTAGGG) repeated many hundreds of times (Figure 2). Because this repetitive DNA does not code for a protein, it is partly dispensable: some of it can be lost due to normal shortening of chromosome ends without adverse consequences.



The diagram shows two DNA strands. The top strand is a sequence of six 'TTAGGG' units repeated, with ellipses at both ends. The bottom strand is a sequence of six 'AATCCC' units repeated, also with ellipses at both ends. Vertical lines connect the 'T' in the top strand to the 'A' in the bottom strand, and the 'A' in the top strand to the 'T' in the bottom strand, illustrating base pairing.

Figure 2. The DNA code at each chromosome end. One strand is a string consisting of hundreds of copies of the letters TTAGGG. The other strand contains the complementary letters, AATCCC.

Second, there are specialized proteins within the cell that recognize DNA containing this specific code and bind to it, and there are other proteins which bind to the DNA-binding proteins. Together, these proteins form a complex structure that coats the chromosome ends and protects (or shelters) them from being mistakenly recognized as an accidental DNA break. Collectively, these proteins are therefore called the shelterin complex, and within this complex there are six different proteins: TRF1, TRF2, TIN2, TPP1, POT1, and hRAP1.

The repetitive DNA at the end of a chromosome is called a telomere. Sometimes, however, biologists use the same word to refer to the DNA plus the proteins that bind to it. Although this is potentially confusing, the meaning of the word is usually clear from the context.

Although telomeres still function despite a considerable amount of shortening, the amount of shortening that can occur without consequence is not unlimited. If too much shortening occurs, there is not enough telomere DNA left for shelterin to bind to, which results in the chromosome ends losing their protection from being treated as a DNA break.

Cells have a built-in mechanism to deal with shortened telomeres. Once a telomere reaches a minimum length, the cell is no longer permitted to divide again. In effect, it reaches its use-by date for replication. Telomere shortening can thus contribute substantially to the aging process. When there is a significant decline in the number of cells that are able to divide, tissues and organs lose their capacity to undergo the renewal processes upon which healthy function depends. Telomere shortening can contribute to aging even in tissues that contain many non-dividing cells (like the brain) because their health depends on cells that do need to continue dividing. For example, the nutrition of the non-dividing cells of the brain depends on blood vessels lined by cells that must continue dividing to maintain normal function.

How Telomeres Can Last a Normal Lifespan

Telomeres are part of a finely tuned biological system. Under normal circumstances, telomeres continue to function, protecting the chromosomes throughout all cell divisions of a normal human lifespan. Two critically important factors in their continued competence are their starting length and the rate at which they shorten.

Germline cells require processes to provide an adequate starting length very early in the development of the embryo. These processes involve the action of a complex molecular machine (enzyme) called telomerase, which is able to add new DNA (containing many repeats of the TTAGGG sequence) to the ends of chromosomes, and thereby completely counteract the normal shortening process that accompanies cell division. The similarity between the words "telomere" and "telomerase" is potentially confusing, so it is important to keep in mind that telomeres are the DNA at chromosome ends (which

become shorter with cell division), whereas telomerase is an enzyme (molecular machine) which lengthens telomeres.

The rate of telomere shortening is influenced by environmental and lifestyle factors. For example, toxic chemicals and cancer chemotherapy agents can cause tissue damage, and therefore induce a lot of cell division, thus increasing the overall rate of telomere shortening. Treatments that are normally used to get a patient ready for a bone marrow transplant can also cause this, and some viral infections can have this effect. In addition, lifestyle factors, such as physical inactivity, smoking, severe prolonged stress, and obesity are associated with shorter telomeres.

There are low levels of telomerase activity in the cells of many somatic tissues, which partially counteract normal telomere shortening. This slows down, but does not completely prevent the shortening in normal somatic cells. This is particularly important in organs that undergo a lot of cell division, including the bone marrow, which constantly produces huge numbers of new blood cells.

Telomerase is thus a key player in ensuring that our telomeres last for a normal human lifespan. It not only is important for providing a sufficient telomere length buffer at the start of life, but also for slowing down the rate of telomere loss through successive cell divisions. Inherited defects in telomerase can result in telomeres that are excessively shortened and therefore ineffective at preventing chromosome-related disease.

Telomerase

Telomerase is an enzyme that synthesizes the DNA sequence TTAGGG and adds it to the ends of telomeres. It does this by reverse transcribing an RNA molecule, which has a short sequence in the template region that is complementary to the TTAGGG DNA sequence. The RNA is sometimes referred to as hTER or hTERC, but usually hTR (human Telomerase RNA), and is encoded by the *TERC* gene (Telomerase RNA Component). When hTR molecules are being transcribed from the *TERC* gene, a string of additional

"letters" is added to its end. The letters are all "A", so this string is called the polyA tail. The level of hTR in a cell is finely tuned by enzymes that decrease (PARN) or increase (PAPD5) the length of the polyA tail. Increased length of the tail results in hTR being earmarked for destruction, and some drugs have recently been discovered which block the action of PAPD5 and therefore increase the level of hTR and of telomerase.

Telomerase also contains a protein subunit called TERT (Telomerase Reverse Transcriptase), which does the reverse transcribing. The name of the gene encoding this protein is *TERT*. A third component of the active telomerase complex is a protein called dyskerin, which binds to RNA molecules like hTR. The gene that encodes dyskerin is called *DKC1* because it was the first pathogenic variant confirmed to cause dyskeratosis congenita, one of the syndromes caused by short telomeres.

The active telomerase enzyme complex contains at least six molecules: two copies each of hTERT, hTR, and dyskerin. Assembly of this complex molecule requires the action of specialized proteins, including NOP10 and NHP2.

In order to lengthen telomeres, telomerase needs to be transported from the places where it is assembled to the ends of telomeres where it does its work. TCAB1 (encoded by the gene *WRAP53*) is a protein required for this transportation. TCAB1 (and also NOP10 and NHP2) may form part of the telomerase complex at various stages in its life cycle. Once telomerase arrives at the chromosome end, it needs to dock with the telomere. A protein critically important for this is TPP1 (encoded by the *ACD* gene). TPP1 has a small region on its outer surface (the "TEL patch") with which it latches onto the surface of the telomerase enzyme. TPP1 attaches to the telomere end by binding to another shelterin protein, POT1 (encoded by the *POT1* gene).

Other Molecules Needed for Normal Telomere Length

Considering data from other species, it would not be surprising to find that there are several hundred proteins that influence telomere length in humans to a greater or lesser extent.

Of those that are already known, the shelterin proteins are very important, because of their ability to influence telomerase activity and to protect the telomere. Telomerase synthesizes only the TTAGGG strand of telomere DNA. CTC1, STN1, and TEN1 are three proteins which form a molecular machine (the CST complex) which is thought to be involved in synthesizing the complementary CCCTAA DNA strand. The complex is also thought to be important in controlling the activity of telomerase.

The DNA of telomeres can be configured in a number of ways other than the helical Watson-Crick structure. The presence of many consecutive Gs in the telomere sequence means that telomeres are able to form complex structures (known as G-quadruplexes), whereby the Gs bind to each other instead of forming G-C pairs. Telomeres can also form a “t-loop” structure.

A number of proteins, including RTEL1, are required to prevent the G-quadruplex and t-loop structures from causing problems when telomeres are being copied during cellular replication (Figure 3).

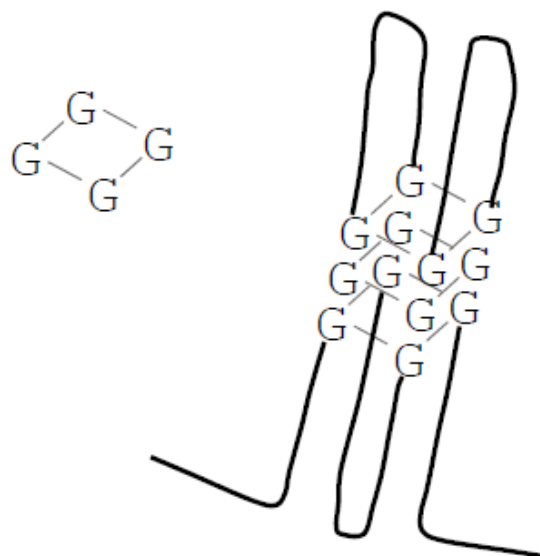
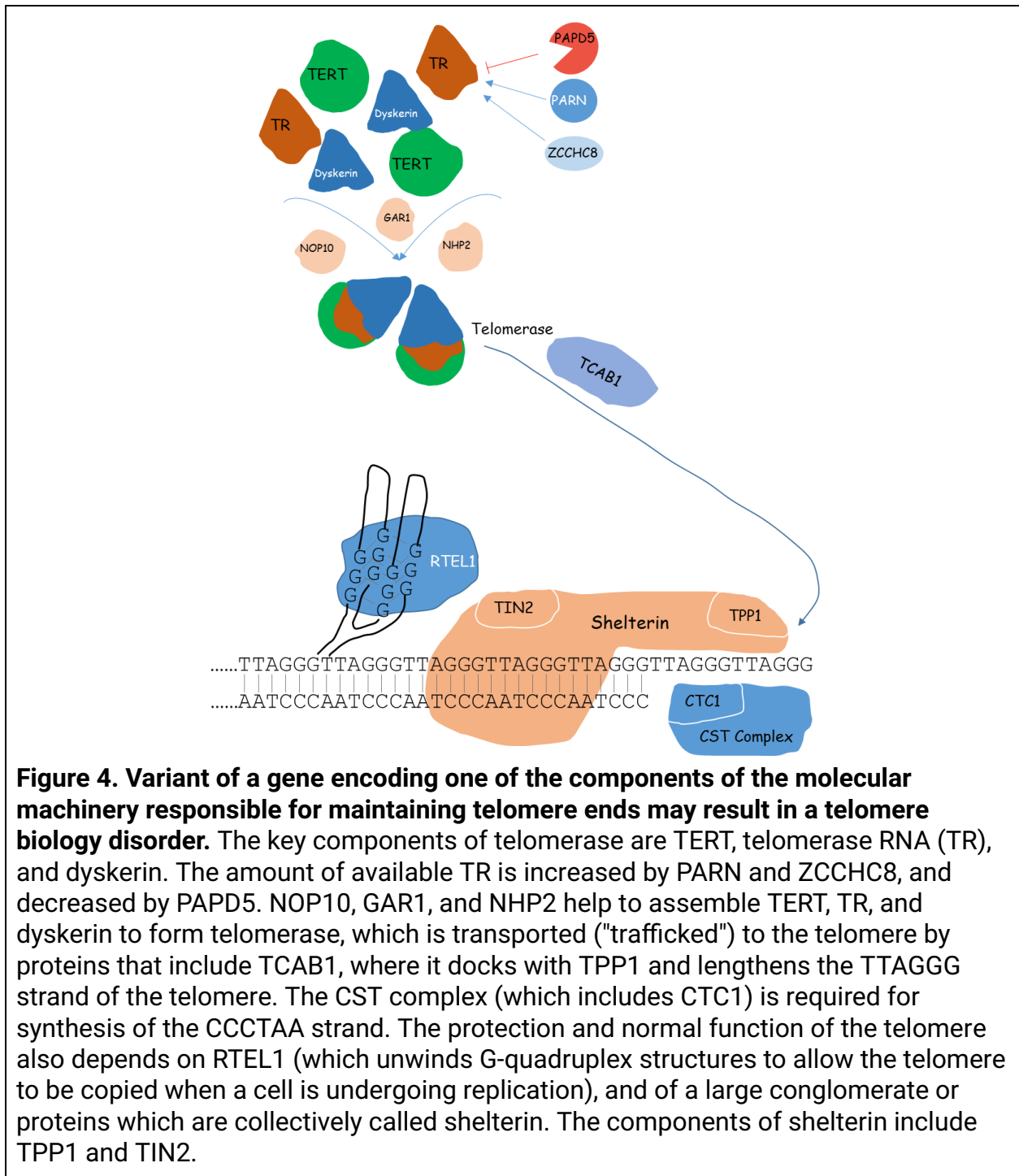


Figure 3. Telomeres can form structures called G-quadruplexes. DNA that contains many runs of the letter G, can fold into complex structures, where the Gs bond together (instead of binding to the Cs on the strand of complementary DNA). Four Gs can bond to form a square-like structure called a G-quartet (left), and stacks of G-quartets (right) are called G- quadruplexes.

Inherited Causes of Excessively Short Telomeres

Inherited defects (pathogenic variants) in any of the genes that encode components of telomerase (*DKC1*, *TERC*, or *TERT*) or of specific genes that encode proteins that control the level of hTR (*PARN*), hTR maturation (*ZCCHC8*), or are involved in telomerase's assembly (*NOP10*, *NHP2*, or *NAF1*), transportation to the telomere (TCAB1 encoded by *WRAP53*), or docking with the telomere end (*TPP1*, encoded by *ACD*), can result in telomeres being too short (Figure 4). In addition, defects in genes that encode specific proteins involved in other aspects of telomere protection and synthesis (*TINF2* [the gene encoding TIN2, one of the shelterin proteins], *CTC1*, *STN1*, *POT1*, and *RTEL1*) may also cause excessive telomere shortening. Some individuals who have excessively short telomeres do not appear to have variants in any of these genes, so it is likely that there

are additional inherited causes of short telomeres resulting in telomere biology disorders that have not yet been found.



The consequences of deficient telomerase activity are two-fold (Figure 5).

First, the normal shortening that accompanies cell division is counteracted less effectively than when telomerase is normal, so telomeres shorten faster. Second, when there is less telomerase activity than normal, it may not be possible to restore telomere length in the germline cells and early embryo. The result is that an individual in the next generation inheriting the defective gene may also inherit telomeres that start off shorter than normal, a "double whammy" effect. This results in a tendency for some diseases with short telomeres to get worse from generation to generation, an unusual pattern of inheritance called genetic anticipation.

Defects in the other genes which cause short telomeres may follow a similar pattern, even though they may not directly affect telomerase activity. Gene defects causing failure to protect the telomere or to copy the telomere DNA properly during cell division may speed up the rate of telomere shortening to an extent that normal levels of telomerase are not able to adequately counteract, either in somatic or germline cells that pass DNA on to the next generation.

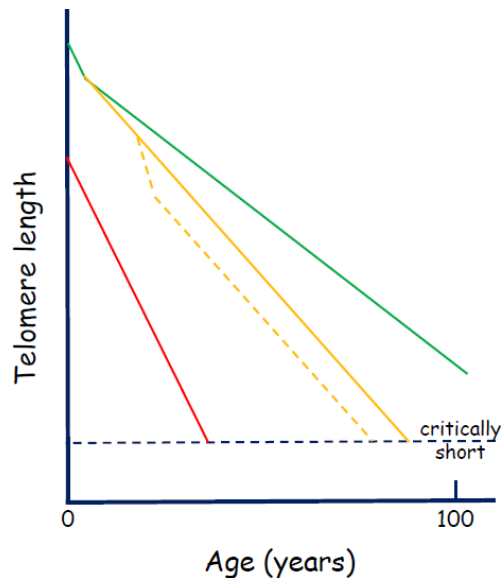


Figure 5. Telomeres need to last for a normal life span. The telomeres of many somatic cells shorten throughout life but, if they have a normal starting length and do not shorten too rapidly, we can live a long life and still have sufficient cells that do not have critically short telomeres (green line). Because of the amount of growth occurring, telomeres shorten more rapidly in early childhood than in adults. If the telomeres shorten faster than normal (which occurs, for example, when cells that undergo a lot of proliferation do not have enough telomerase), then some organs may lose the capacity to renew themselves later in life (amber line). Some environmental factors may cause cell death, which results in an increased need for proliferation and therefore, an increased rate of telomere shortening, and this may result in problems occurring earlier in life (dotted amber line). Individuals who start life with short telomeres and also have an increased rate of telomere shortening (red line) may have problems from short telomeres early in life.

Why Short Telomeres Cause Disease

Cells with excessively short telomeres reach their proliferative "use-by date" much earlier than normal, which means that various tissues and organs are not able to maintain themselves by normal numbers of cell divisions. Eventually, this may result in an insufficient number of cells (cytopenia) in various organs.

The severity of the condition tends to be related to how short the telomeres are. If the telomere length deficit is very severe, there may even be insufficient cell division for organs to develop normally in specific embryonic tissues (for example, in the cerebellum and other parts of the brain). When telomere biology disorders become manifest first in childhood, they often affect the bone marrow, which is normally one of the most highly proliferative organs. This causes a deficiency in numbers of red blood cells (anemia), white cells (neutropenia), and platelets (thrombocytopenia). The combination of these three deficiencies is called aplastic anemia or bone marrow failure. When telomere shortening is less severe, problems may not surface until later in life.

Excessively short telomeres are also associated with an increased risk of cancer. The reasons for this may include the following. First, when telomeres become excessively short, they lose their ability to protect chromosome ends from the DNA repair machinery. This results in end-to-end fusion of chromosomes and the potential for the joined chromosomes to break at a random location when the cell next divides. This sets up a cycle where continued random chromosome breakage and random rejoining creates unpredictable changes in the genome, which may increase the risk of cancer. Second, organs that are depleted of normal cells may send increasingly powerful signals to stimulate the remaining cells to divide, which may inadvertently favor the growth of rogue cells on their way to becoming cancerous. Third, normal cells in a normal organ can often restrain the growth of rogue cells, but this effect is progressively lost as surrounding cells die.

Excessive telomere shortening can affect almost any organ system, but it is still not clear why it causes bigger problems in some organs than others. It seems easy to understand why bone marrow, a highly proliferative tissue, may be affected. It is difficult to understand, however, why there are more often serious problems in lungs, which are thought to have only moderate rates of cell division, than in the skin, or in the lining of the gastrointestinal tract, both of which constantly undergo high levels of cell division. It is also difficult to understand, for example, why some individuals who have no major

problems with their bone marrow will develop lung disease. Even more baffling, there are individuals with very short telomeres who appear to be disease-free for a normal life span.

The answers may lie in part in the effects of environmental and lifestyle factors. In some families with short telomeres, pulmonary disease only occurs in individuals who have both inherited the defective gene and who smoke. In other families, an individual's short telomere-related problems may become manifest when treated with chemotherapy for cancer. The answers may also lie in the modifying effects of other genes we do not yet know about.

This uncertainty also provides some grounds for optimism. An individual who inherits a pathogenic variant associated with short telomeres may not necessarily develop any manifestations of the condition, and even if one or more of these manifestations occur, disease progress may be quite unpredictable. The more that is understood about interactions between the environment and telomere biology, the better we may be able to prevent or modulate the adverse effects of the genes.