The replication of the genome is essential for the continuity of life. The molecular mechanism is very similar in all groups of organisms. Although the basics of replication are already well understood, researchers are still focusing on questions relating to DNA replication. These questions not only deal with the understanding of a basic biological process, but also with related medical aspects.

One attribute of living things is the ability to reproduce. The information required to pass on traits to the next generation is mainly stored in cellular DNA. Daughter and parent cells need to be equipped with an identical copy of DNA during cell division. The molecular basis of this process is the replication of DNA. In 1958, Matthew Meselson and Franklin Stahl showed that newly replicated bacterial DNA consists of a new and an old DNA strand. In the 1970s and 1980s, further evidence was found relating to the mechanism of ‘semiconservative’ replication. Replication consists of three phases: initiation, elongation and termination. The basis of replication is the pairing of the four bases found in DNA: adenine pairs with thymine, cytosine with guanine. The process, which results in two DNA helices instead of one, is mediated by several dozen proteins. A simplified view of what happens during replication is sufficient to reveal the complexity of this process.

One becomes two

During initiation, so-called initiator proteins bind to the DNA double strand of the parent cell in the region of the origin of replication. The enzyme helicase separates the individual strands at this site and unwinds them together with the enzyme topoisomerase. This process makes the two complementary parts of the double helix accessible to the replication enzymes. The enzyme DNA primase combines a short fragment of RNA (primer) with the complementary constituents of the parent DNA. DNA polymerase, which is the major replication enzyme, needs a primer at which it can add the first nucleotide. The DNA and the replication enzymes form the so-called replication fork, which moves along the template DNA during the elongation of the DNA. The DNA polymerase then follows and pairs new nucleotides with their complementary counterparts on the template strand. Continuous synthesis occurs in the direction in which the replication fork is moving. The DNA polymerase also registers mismatches and repairs them, thus enabling extension to continue. The ends of the new strand are oriented in the opposite direction of the parent strand, which scientists refer to as antiparallel pairing.

The process of elongation is complicated. Each old strand becomes paired with a new strand copied from it. What makes elongation so complicated is the function of DNA polymerase which can only add nucleotides in a specific direction: the enzyme adds new nucleotides in the 5'→3' direction in a continuous manner since it uses the free 3'-OH group donated by a single RNA primer. Since the two DNA strands are arranged in an anti-parallel manner, the direction of synthesis on the parent strand is therefore predetermined. On the leading strand, the new DNA strand is continuously synthesised in the 5'→3' direction, i.e. the direction in which the replication fork is moving. The lagging strand, which is the DNA strand opposite of the leading strand, runs in the 3'→5' direction. DNA polymerase cannot synthesise in this direction. Therefore, the lagging strand is synthesised in short segments, known as Okazaki fragments. The RNA primers disappear after a short time and the enzyme DNA ligase links the Okazaki fragments. The termination of DNA synthesis is mediated by an enzyme that blocks the enzyme helicase, which results in the replication fork coming to an end.

Differences between bacteria and eukaryotes

Replication was initially investigated in bacterial cells. The mechanism of replication is a lot simpler in bacteria than in eukaryotic cells. For example, a bacterial cell has five polymerases that slightly differ from each other in structure and function. Human cells have more than a dozen different polymerases. The E. coli genome is considerably smaller than the eukaryote genome. In addition, bacteria have a circular genome (plasmid) while the eukaryote genome is linear and distributed amongst several chromosomes. In E. coli, DNA replication progresses bidirectionally away from the origin of replication. In these circular DNA molecules, one origin often suffices and the two resulting growing forks merge on the opposite side of the circle to complete replication. However, the long linear chromosomes of eukaryotes have thousands of origins which are created at different points in time; the two growing forks from a particular origin continue to advance until they meet the advancing growing forks from neighbouring origins. Replication then comes to a halt.
Replication and ageing

However, the process is not the same at the end of the chromosomes, where a replication fork can no longer progress due to the lack of bases. At the site of the lagging strand where the DNA primase places the last RNA primer, the DNA polymerase is no longer able to continue replication. The terminal DNA segment cannot be replicated. The DNA strand gets shorter and shorter as the number of cell division rounds increases. To protect themselves against the rapid shortening of the DNA, eukaryotic chromosomes possess sequence repeats (telomeres) at their extremities, which do not code for proteins. Since the telomeres do not contain any important information, the key parts of the DNA are protected. The telomeres get shorter each time a cell divides. The length of these so-called telomere caps defines the number of possible divisions and hence the lifespan of a cell. Some cell types (for example maturing sex cells or certain tumour cells) contain the enzyme telomerase, which prevents telomere shortening and thus protects the cell from cell ageing and programmed cell death.

Order in the cells

A cell must at all costs prevent errors from occurring during the replication of DNA. The strict order of the copying process is therefore essential. In eukaryotic cells, the DNA is kept as clearly arranged as possible: genome regions that are not undergoing replication are densely packed in chromosomes. Scientists also assume that the chromosomes in eukaryotic cells are spanned over a cytoskeleton consisting of protein tubes and wires. The replication enzymes are bound to this so-called nuclear matrix and motor proteins pull the genome past them. It appears that bacterial cells have similar mechanisms.

All in good time

Many bacteria divide once every thirty minutes, others replicate even faster. Eukaryotic cells only replicate their genome when new cells have to be created. This happens as a result of external signals, for example tissue loss or inflammation. The life cycle of eukaryotic cells underlies an accurately defined sequence of activities that can be divided into different phases: 1st resting (gap) phase (G1), DNA synthesis phase (S), 2nd gap phase (G2) and mitosis (M). The cells duplicate their genome in the S phase. During the gap between DNA synthesis (S) and mitosis (M), the cells check whether the DNA has been replicated and errors have been repaired in order to determine whether they are ready to proceed to mitosis and divide. This sequence of activities is controlled by enzymes of the group of cyclin-dependent kinases (CDKs). If errors occur during the control of the cell cycle, the cells might divide more quickly and more frequently as is the case with many cancer cells.

The replication of viruses

Viruses also reproduce, but they cannot do so on their own. That is why they cannot be called “alive” in the strictest sense of the word. They use the replication apparatus of the host cells, and have additionally developed a number of special characteristics. Scientists differentiate viruses according to the genome type – there are DNA and RNA viruses: viruses may have single-stranded or double-stranded linear RNA, single-stranded or double-stranded linear DNA, single-stranded or double-stranded circular DNA and other variations. Some viruses contain some of the enzymes required for their replication, for example the influenza virus, whose envelope not only contains an RNA genome but also an RNA polymerase. When the virus enters the host cell, the enzyme RNA polymerase starts to replicate the viral genome. The synthesis of the genome of DNA viruses usually begins at a replication origin that binds specific initiator proteins, which recruit replication enzymes of the host cell which then replicate the viral genome.

The HI virus is a retrovirus and thus a very exotic case. The virus got its name due to the fact that it reverses the normal process of transcribing DNA into RNA (transcription) during reproduction. The virus has a single-stranded RNA genome and an enzyme called reverse transcriptase. This enzyme copies the single-stranded RNA genome into a complementary DNA molecule, thereby enabling the integration of the viral genome into the host DNA. Once the viral genome is integrated into the host genome, it can be transcribed into RNA by the host enzymes at which point it can reproduce. Since viruses are able to use a broad range of replication mechanisms to reproduce, scientists working on the development of anti-viral drugs need to specifically investigate the individual viruses one by one. Many of the currently used drugs interfere with viral replication, for example the so-called nucleoside analogues which are used against the hepatitis B virus.

The replication of cancer cells – a target for therapy?

Replication is also of great interest in the field of medicine, in particular in the fight against cancer. Cancer cells are body cells that no longer behave normally - they replicate their genome and proliferate far more often than healthy cells. Researchers and physicians exploit this behaviour in their work on substances designed to interfere with cancer growth. Some substances inhibit replication, which prevents tumour growth. Modern chemotherapy uses alkylating agents (e.g., busulfan, ifosfamide). These substances bind to DNA via alkyl groups. Since these groups have two binding sites, the genome is joined together, thereby preventing it from replicating. Another example is platinum analogues that are among the most effective chemotherapeutic drugs. These substances have a platinum atom that binds to the DNA and joins it (e.g., cisplatin, carboplatin). Substances like anthracyclins and the antibiotic doxorubicin intercalate DNA and make it unaccessible to DNA polymerase, thereby preventing the enzyme from synthesising DNA.
Intercalate substances such as doxorubine are incorporated into the DNA bases (green) making them inaccessible to DNA polymerase, thereby preventing replication. The schematic shows a section of the DNA double helix into which two molecules have been intercalated.

**Domesticised replication**

The principle of replication was used for the first time in the 1980s by molecular biologists in the laboratory. When researchers are investigating DNA they usually have very small amounts of the molecule available. In order to be able to make effective statements, the quantity of DNA needs to be increased. This can be achieved in what is known as polymerase chain reaction (PCR) where DNA fragments of a certain sequence are amplified by the enzyme DNA polymerase. The DNA fragments of interest are used as a template, to which free DNA constituents (nucleotides) and specific primer sequences are added. The PCR process consists of several cycles of repeated heating and cooling of the reaction in order to melt and enzymatically replicate the DNA. As the reaction progresses, the DNA generated is used as a template for further amplification. This sets in motion a 'chain reaction' in which the original DNA template is exponentially amplified, thus achieving an amount that the researchers can work with. If the researchers do not know the sequence of the DNA template to be replicated, the primers used can help indicate the sequence as they consist of a specific sequence of nucleotides that enable the scientists to draw conclusions on the amplified material. PCR has been further developed in many ways and has now become a standard automated method in molecular biology laboratories around the world.

**The origin of replication**

What made autonomously replicating systems develop? Many scientists assume that RNAs were the first replicating macromolecules. RNAs, just like DNA, are in principle able to create negative blueprints of themselves through spontaneous base-pairing. In contrast to double-stranded DNA, RNA can also adapt different spatial shapes so that it can also function as an enzyme that catalyses its own replication. An enzymatic activity has already been shown for some RNAs, for example an RNA activity that cleaves other RNAs or mediates the binding between two amino acids. Whether the first cells actually had an RNA genome and DNA became their genome in a secondary step, is still a matter of controversy.

mn - 29.06.2009

© BIOPRO Baden-Württemberg GmbH

**Literature:**

