Fundamentals of Biology 1A – Part II

16.11.2016

Chapter 14: Structure and Analyses of Genomes

Def. Genome: The unique totality of all the DNA (the genetic material) within a cell, organism or organelle (and very rarely it is the RNA, but only in viruses), which can be passed on to the next generation. One should not forget that the genetic material, the genome, which can be passed on to the next generation is thought of "haploid" contexts.

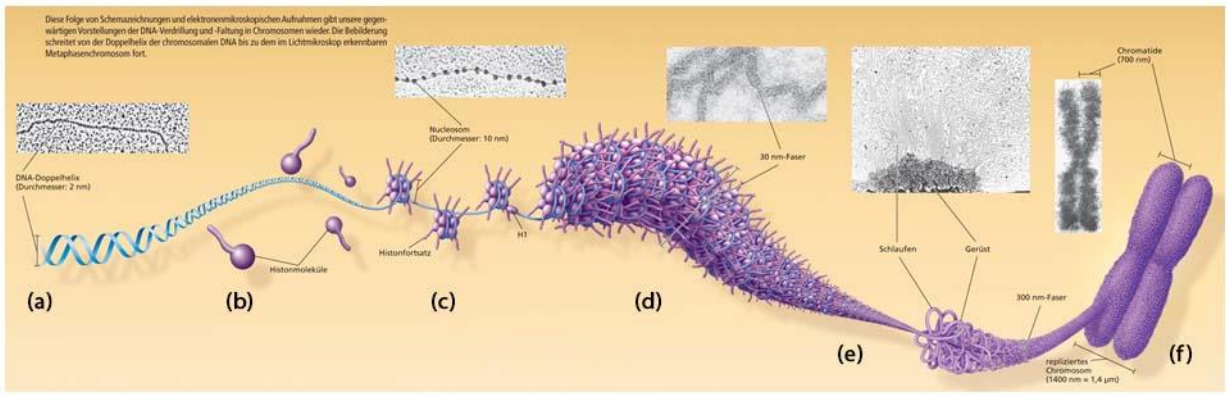


Abbildung 16.21: Aufbau des Chromatins eukaryotischer Chromosomen. (a) Die DNA-Doppelhelix (b) Histone (c) Nucleosomen (d) die 30 nm-Faser, die durch Wechselwirkungen zwischen Histonfortsätzen eines Nucleosoms mit der DNA eines anderen Nucleosoms entsteht. Durch diese Wechselwirkungen wird die gestreckte Faser aufgerollt. (e) Schlaufendomäne (300 nm-Faser). Die 30 nm-Faser legt sich in Schlaufen, die an einem Proteingerüst verankert werden. (f) Das Metaphasechromosom. Während der Zellteilung falten sich die Schlaufendomänen weiter und das Chromosom wird bis zum charakteristischen Metaphasenchromosom komprimiert. (adaptiert von Abbildung 16.21, Biologie Campbell, 8. Auflage).

Corollary.: Most cells have the exact same genome. Exceptions are the gametes. These cells only have half the chromosome set of diploid cells. In addition, since meiosis shuffles the genetic information in order to maximize genetic variation, the gametes all have different genomes.

Genomes of prokaryotes: They do not possess a nuclear envelope with DNA contained in it (no cell nucleus). The genome is made of a long string of a circular DNA molecule. The DNA is not coiled around histones, but there are specific proteins that are able to coil and even supercoil the DNA molecule. There are also smaller parts of circular DNA which are called plasmids.
One does not speak of chromosomes, since it is not made of chromatin. One sometimes refers to it as bacterial chromosomes.

Viral genomes: The are often only one DNA string or sometimes RNA packed in a protein shell called capsid. They are not considered life forms since they do not have metabolism.

Def. Sanger-Sequencing/Chain termination method (first generation): The classical chain-termination method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleosidetriphosphates (dNTPs), and modified di-deoxynucleosidetriphosphates (ddNTPs), the latter of which terminate DNA strand elongation. These chain-terminating nucleotides lack a 3'-OH group required for the formation of a phosphodiester bond between two nucleotides, causing DNA polymerase to cease extension of DNA when a modified ddNTP is incorporated.
DNA is denaturated, so it becomes two linear strands. The strand to be sequenced is copied with the

addition of chemically modified bases (dNTPs; since it is in some sort of container, the dNTPs and ddNTPs can be simply added to the overall solution). dNTP and ddNTP are equivalent, but only ddNTP will terminate the addition of further bases. The replication process is stopped each time a modified ddNTP is added. One repeats this process for all 4 bases and then simply puts them together like a puzzle.

Def. Pyrosequencing (second generation): DNA polymerase extends the complementary DNA string nucleotide by nucleotide starting at a primer. Nucleotides are built in by an enzyme system, luciferase, which makes luciferin with ATP to oxyluciferin. This reaction emits a light signal which can be detected using detectors.
What nucleotides are being built in? So-called deoxynucleosidetriphosphates such as dATP, dGTP, dTTP and dCTP (the nucleotides in the DNA are only deoxynucleosidemonophosphates). When DNA polymerase uses these other nucleotides, a reaction occurs that produces pyrophosphate (PP). PP is transformed to ATP which fuels the luciferin-oxyluciferin reaction. The intensity of the light signal is proportional to the ATP used.

Def. Nanopores (third generation): Pores of the order of 1 nm. In DNA sequencing, there is an electric current passing through the nanopore When a DNA molecule passes through it, there is a significant change in the magnitude of the electric current which can be measured. One can do this for individual bases in order to sequence DNA.

Def. Splicing: Occurs during DNA synthesis. After RNA is made, non-coding parts are cut off by spliceosomes, a small nuclear ribonucleoproteins complex. The RNA is now called mRNA (messenger RNA), which is being used for the protein synthesis for it is only made of the coding parts of the DNA.

Def. Exon: The defining parts of the DNA respectively RNA that are used for protein synthesis. These parts are left over after splicing and are joint together to create the mRNA which is used for the synthesis of proteins. Although exons are often referred to as the coding parts of the DNA, there are also non-coding exons (keep that in mind).

Def. Intron: Parts of the RNA that are cut off and that are not being used for protein synthesis. Often referred to as non-coding parts of the DNA.

Def. Transposon: Also known as jumping genes. They have a defined length in the genome and they are able to change their location in the genome (transposition).

Def. Transposase: An enzyme that binds to the end of the transposon and allows it to change its location.

Non-coding sequences in the genome: Repetitive sequences. They often occur in centromeres and telomeres.

Mechanisms that lead to the evolution of genomes:

Duplication of chromosomes, new assembling of chromosomes (change of chromosome structure), duplication of genes and change through transposable elements.

Reassembling of chromosomes: Chromosomes can fuse at the telomere sequence region. Then there will be a telomere-like sequence region.
There are certain parts of human genes that are found in one chromosome, but in mice for example,

these parts can be allocated on several different chromosomes. So it means, that the human chromosome is somewhat more compact, since on 1 chromosome, there can be many different gene blocks, while mice have these gene blocks on different chromosomes.

Duplication: It can occur during unequal crossing-over in meiosis.

APPENDIX:

Properties of eukaryotes and prokaryotes:
Eukaryotes: Chromatin, Chromosomes, diploid, double-helix, double-stranded DNA and DNA polymerase, low gen-density, large genome, has histones, linear genome, has many genes, has many non-coding regions
Prokaryotes: diploid, double-stranded DNA and DNA polymerase, huge gen-density, small genome, has plasmids, few genes, few non-coding regions, circular genome

Def. Euchromatin: Loosely packed DNA in the chromosome, which is easy to access and often used.

Def. Heterochromatin: Very densely packed DNA in the chromosome, which is hard to access and often not used.

Def. Okazaki fragments: Discrete synthesized segments of the lagging strand during DNA replication (replicated by DNA polymerase III). Since the synthetization can only elongate from 5' to 3', the lagging strand has to be handled discretely because of the DNA molecule's antiparallel nature (3' to 5'). It always has to attach a new RNA fragment, so DNA polymerase III can synthesize the Okazaki fragments.

New-assembling chromosomes can be important for evolution since it leads to gene duplication which produces spare copies of the gene that can evolve freely and take on new functions.

DNA replication and involved enzymes:

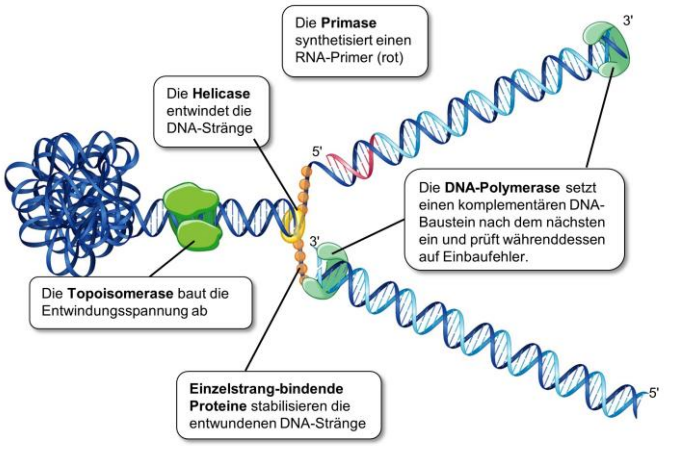


Abbildung 4: Proteine der DNA-Replikation. (adaptiert von Abbildung 12.8 How Life Works, Freeman and Company)

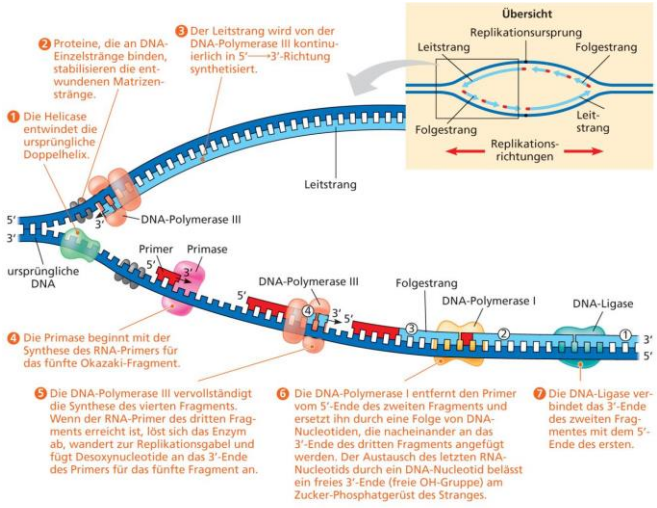


Abbildung 16.17: **Zusammenfassung der bakteriellen DNA-Replikation.** Obwohl hier in der höheren Auflösung wieder nur eine Replikationsgabel abgebildet ist, zeigt die Übersicht oben rechts, dass die Replikation für gewöhnlich gleichzeitig an zwei Gabeln erfolgt, die sich an den beiden Enden einer Replikationsblase befinden. Die Übersicht zeigt jeden Tochterstrang in seiner Gesamtheit, so dass zu erkennen ist, dass die eine Hälfte kontinuierlich gebildet wird (Leitstrang), während die andere Hälfte in Fragmenten synthetisiert den Folgestrang bildet.

17.11.2016

Chapter 15: Replication of Genomes

The process of DNA replication is called semiconservative, since the two new synthesized DNAs have one string as a template and the other one bound by hydrogen-bonds is the synthesized complementary one.

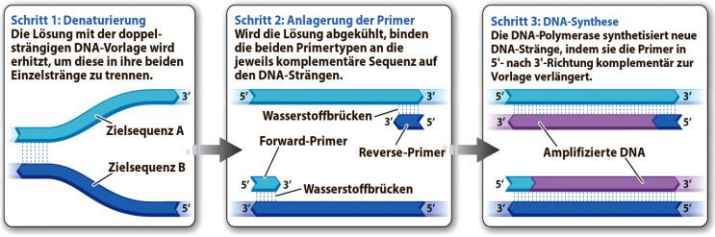
Def. recombinant DNA: Artificially created DNA from multiple sources with laboratory methods. Since DNA is universal, it can be joined together as one pleases (e.g. a part of human DNA can be joined to some fungal DNA). This recombination of DNA is called recombinant DNA. They do not have to be expressed necessarily. If expressed the protein is called recombinant protein.
This is not a trivial process, since the gene has to be restructured to include promoters, initiation signals etc. for the creation of the mRNA.

Def. Restriction enzyme: Found in prokaryotes. These enzymes are used as protection against viral genetic material. These enzymes have a special mechanism of recognizing certain base sequences which can be removed from the plasmids and thus rendering it ineffective since it is not expressed into a protein anymore.

The polymerase chain reaction (PCR) is used to create recombinant DNA artificially, thus it is central for cloning.
One starts denaturing the DNA so it loses its double helix form in a solution. After it has cooled off and remained in its denatured state (just two long strings connected by the usual h-bonds at the bases), special primers bind to it by h-bonds. A forward primer binds to the 3' site and the reverse primer to the 3' site, too. A heat-stable DNA polymerase can now synthesise that parts of the DNA, thus amplifying the targeted gen. A heat-stable DNA polymerase does not denature at high temperatures.

Those produced parts will be copied with the aid of prokaryotes. These possess plasmids, circular DNA, that works independently of the main genome. One can cut and glue the plasmids, thus insert the targeted gen, which will be expressed later on and can be collected.
How does one realize that? It is realized by using special enzymes, called restriction enzymes. These enzymes hydrolyse the phosphodiester bonds, thus cutting the plasmids. The targeted gen is inserted with a cell-specific ligase. Since the cut plasmid has an open 3' and 5' site, one can bind the targeted DNA to it with the ligase (the OH-groups are bound to the phosphate residue to an ester bond).

Prokaryotes can be then used to express the gen, but this process is not so trivial, since gens often contain introns



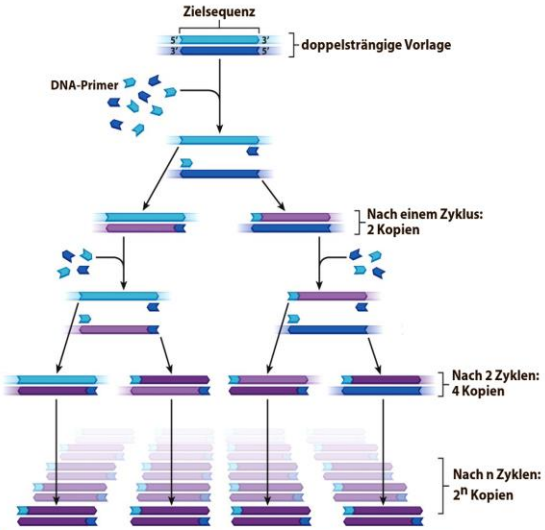


Abbildung 12.13mod: **Polymerasekettenreaktion (PCR).** (adaptiert von Abbildung 12.13, How LifeWorks, Freeman and Company)

(non-coding parts) that can be troublesome for the prokaryote's expression system.
Yeast is often used to express DNA with introns. It has the unusual ability for a eukaryotic lifeform to accept plasmids. Yeast can cut off the introns by themselves, since they are eukaryotes themselves and it is a common characteristic amongst all eukaryotes.
The gen can be expressed now.

Def. Gene-knockout: A gen in an organism is turned off. It is then observed how the organism reacts to it and changes its behaviour. It is an alternative to inserting genes in plasmids if one only wishes to observe the gene's function (and has no interest in producing more of the expressed protein).

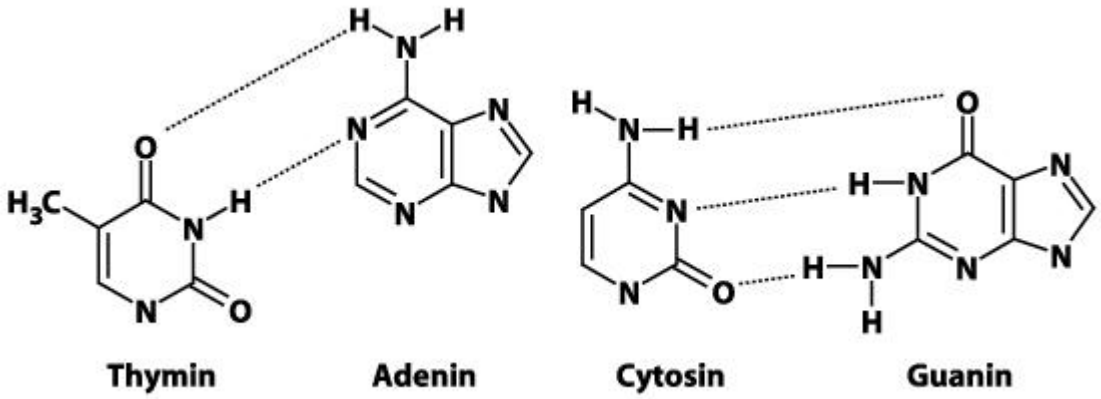
Def. Genetic marker: DNA variations that are often accompanied with a phenotype such as an illness.

23.11.2016

Regenerating cells: blood cells, skin cells, intestine cells, gametes etc. Every second 10**6 cells are replaced.

Steps of cell division: DNA replication, organelle replication, cell cycle (control mechanisms), reparations.

On DNA replication:
What is the specificity of the bonds between the bases? C binds to G, because both can make 3 h-bonds, while A and T can only make 2 h-bonds. Pyrimidin and a purin will always go together (because pyrimidin-pyrimidin or purin-purin, because one be too big, the other too small which would be problematic for DNA
polymerase).



Def. Brownian motion: Random movement of a molecule suspended in a liquid (or also gas) which is brought about the collision of other fast moving atoms or molecules.

How do mutations occur?
Spontaneously (most of the time), mutagen substances such as UV-light, x-ray waves, chemical substances. Single bonds can break by chemical modifications and double bonds can cease to exist by a too high intensity of UV-light exposure.

The end of a DNA is open. For that, telomeres are present to protect that end.

Mechanisms of DNA repair:
DNA ligase: seals break in the sugar – phosphate backbone.
Mismatch repair: single mispaired base repaired by removing and replacing a DNA segment.
Base excision repair: incorrect base and its sugar are excised from the strand and then replaced.
Nucleotide excision repair: recognizes several mismatched pairs.
How is a new DNA strand recognized? It is differently methylated than the old strand, so it can be distinguished from the template strand.

For a population, a certain frequency of mutations is beneficial, since it allows the species to become more easily adaptable to changing environmental factors. Too little mutations per genome size will render a population unable to adapt (or too adapt too slowly), eventually leading to its extinction. Too much will cause chaos though.

In gametes, it is aimed to minimize the number of mutations, since it will lead to defective offspring (or to somatic variations, which is not deadly for the offspring). A mutation in a gamete will persist, which will eventually lead to every cell in the offspring possessing the mutation.

On Telomerase
Linear chromosomes become shorter. At the end of the template strand, an RNA primer will be on the copied strand, then removed, thus replication ends. This means that there is an unreplicated part at the end of the replicated strand. Effectively, after a few generations, chromosomes would be severely shortened. This disaster is prohibited by the telomerase.
Telomerase is an enzyme and an RNA component. (**FINISH THE PROCESS**)

In a nutshell, telomere protect the end of the DNA – unprotected DNA are reactive and will bind to other open ends leading to chromosomal mutation.
In must somatic cells, telomerase is inactive. Shortening of the telomere will lead to deletion of the last few DNA sequences which leads to a loss of information and open ends leading to chromosomal mutations again.
Increased affinity of mutation will lead to senescence. Cancer cells can reactive telomerase.

Def. CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats. A repeat is about 20-40 bp.
(DOWNLOAD GERALD SCHWANK'S PAPER OFF MOODLE AND WORK THROUGH IT)

(update your notes on CRISPR and Cas if there is a lot in the Campbell)

APPENDIX: Human DNA polymerase can replicate the whole human genome in about 8 hours theoretically. It still takes longer than this since not all origins of replication are active at the same time, they are unevenly spaced throughout the genome and the DNA is contained in several chromosomes. Effectively, there are approximately 2161 origins of replication. One DNA polymerase can replicate about 33 nucleotides per second and the genome is 3.08*10^9 base pairs large.

In order for a primer to bind uniquely it only needs to be 17 nucleotides long. Statistically, this sequence will only occur once in the human genome.

25.11.2016

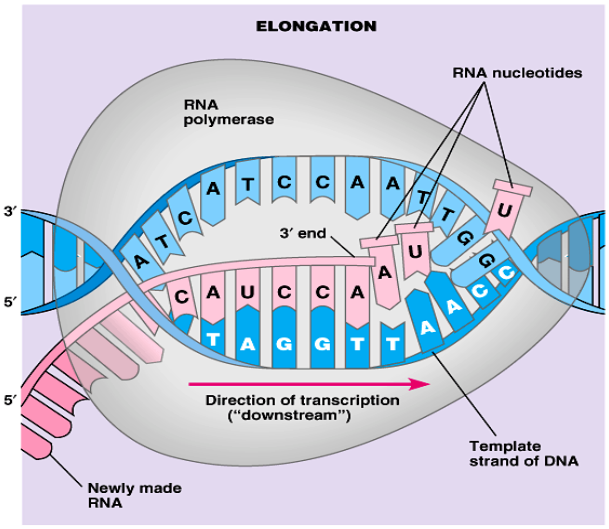
Chapter 16: Interpretation of Genomes

Proteins are the link between the genotype and the phenotype.
Genes are a specific sequence of hundreds to thousands of nucleotides. Genes are a defined DNA sequence in other words.

Def. Promoter: A sequence of about 100 nucleotides on a DNA strand, typically before a gene. RNA polymerase has to bind to it, so it can safely start transcription. There are distinct regions on the promoter such as the TATA-box (specific adenin-thymin sequence).

Kor: In eukaryotes, there are specific groups of proteins that help RNA polymerase bind to a promoter. They are called **transcription factors**.

Transcription: RNA polymerase accumulates at a promoter on the DNA. The DNA unwinds (about 10-20 nucleotides) and the RNA polymerase can start transcribing the DNA template strand into a mRNA. Depending which of the two strands is the template strand, the RNA polymerase will create a mRNA with different genes on it (the two strands are complementary to one another, this means that their base sequence is not the same and thus the genes are different as well). The mRNA is not the copy of the template strand, rather it is the "copy" of the non-template strand. Keep in mind that the mRNA has no thymin base T, instead it has uracil as its base.



The RNA polymerase propagates in the 3' to 5' direction. If the RNA polymerase goes from left to right the lower DNA strand is the template strand. If the RNA polymerase goes from right to left the upper DNA strand is the template strand. The DNA winds behind the RNA polymerase. The orientation of the promoter is the defining factor.
Termination is the final step when the mRNA is released and RNA polymerase leaves. In prokaryotes, there is a termination sequence that signals the termination of transcription. It will then be translated into a protein by translation (recall that prokaryotic mRNA is mostly exons anyway).
In eukaryotes, there is a polyadenyl signal (another sequence after the protein-coding part has been transcribed) which will be transcribed from the DNA and added to the mRNA, which cuts off the mRNA from the RNA polymerase. There are other helper proteins necessary so that the mRNA can leave the cell nucleus and continue with translation. This initial mRNA is called **pre-mRNA**.

The pre-mRNA is processed in the cell nucleus. A modified guanine nucleotide is added to the 5' end. This is also known as a 5' cap. The 3' end gets a sequence of 50-200 adenines added to it. That is why the poly-A-signal is important at the end of transcription, so that the poly-adenine tail can bind to it correctly.

The last modification process of the pre-mRNA is splicing. Since the pre-mRNA is made of exons and introns, the introns need to be removed, since they do not code for a protein (they are not used in translation). There are small nuclear ribonucleoproteins (snRNPs) that react with the splicing locations of the introns catalyse the removal of them and at the same time help the exons to bind covalently.

The end product is called mature mRNA (or simply mRNA).

Def. Alternative splicing: Some of the exons of the pre-mRNA are also removed, which leads to multiple mature mRNAs that can be used to create proteins. Even if humans only have about 22000 coding genes, alternative splicing allows for a lot mature mRNAs to be made and therefore the variety of synthesized proteins is increased.

Def. Codon: A triplet of the nucleotides A, U, C, G that is translated into 1 amino acid ⇔ 3 nucleotides together define 1 amino acid (this function is clearly surjective since there are only 20 amino acids, but 4**3 combinations).

Kor: AUG = starting signal => Methionine. UAA, UAG, UGA = termination signal.

Translation: The mRNA is translated in a ribosome. The ribosome has 4 distinct binding sites – 1 for the mRNA and 3 for the tRNA. The tRNA carries an amino acid rest at one of its ends. At the other end, it carries 3 nucleotides, which are called anticodon. The anticodon can hybridise with the codon of the mRNA (keep in mind that the anticodon has the complementary nucleotides to the codon – otherwise it cannot hybridise). This hybridisation of the tRNA and mRNA happens at the A-location. Then, the tRNA is passed to the P-location, which is the place of the synthesis of polypeptides (proteins). Keep in mind that the tRNA carries an amino acid on its other end. It is then passed to the E-location where it exits the process. If another match occurs for the next codon at the A-location the process repeats itself. At the P-location, the released amino acids are always bonded together and kept there, always growing, until a tRNA hybridises to a stop-codon and the polypeptide chain (protein) is then released.

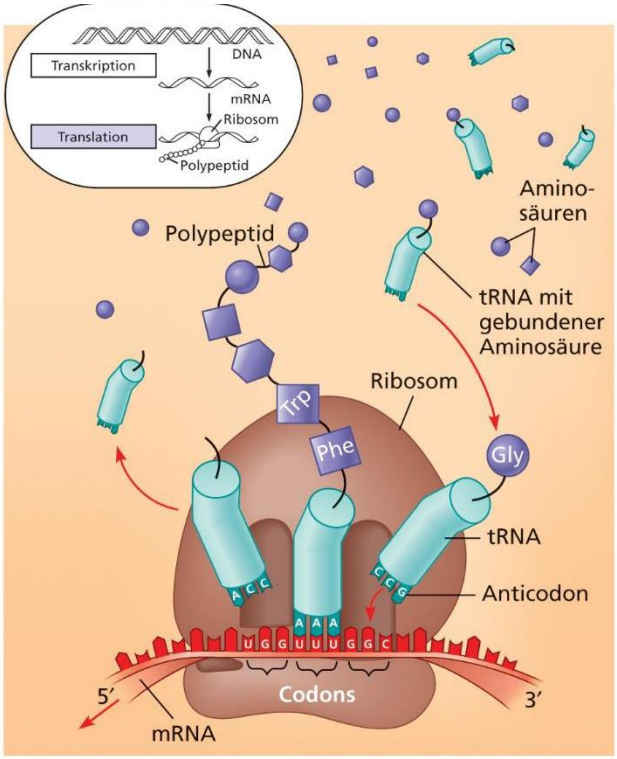


Abbildung 17.13: Translation: das grundlegende Konzept. Während ein mRNA-Molekül durch ein Ribosom gleitet, werden die Codons in Aminosäuren translatiert. Die „Übersetzer" sind tRNA-Moleküle, von denen jeder Typ an einem Ende mit einem spezifischen Anticodon und am anderen Ende mit einer bestimmten Aminosäure versehen ist.

There are 61 codons, but only 45 anticodons. This is due to the uracil in the anticodon which can hybridise with either guanine or adenine for example. The bond of the third part of the codon and anticodon is not that strong (the first and second bond is quite strong though). So only the third base of the codon and anticodon allows flexibility.

Def. Wobble (base): The flexibility of the third base in a codon and anticodon. That is why different codons can code for the same amino acid, they are different in the third base. Only the first and second base are specific.

The translation is divided into 3 subprocesses: initiation, elongation, termination.
Initiation: The first tRNA is called initiator-tRNA. It has the corresponding anticodon to the codon start signal AUG (UAC – the initiator-tRNA carries methionine). The mRNA is read from the 5' to 3' direction. Otherwise, a completely different protein will be created. This occurs at the small ribosomal sub unit. The big ribosomal subunit complex is then added on top of it (with all the P-, A-, E-locations). GTP is the energy form used for this complex.

Elongation: It is known that the polypeptides have a C-terminus and an N-terminus. Methionine has an N-terminus and the next polypeptide that enter the P-location with the tRNA will bind with its C-terminus to the N-terminus of the previous one and so on.

Termination: When the stop codon is reached, there will be no tRNA, since no tRNA can make h-bond to that codon. The stop codon induces the binding of a release factor enzyme, which will hydrolyse the polypeptide chain from the complex and the last tRNA. The complex disintegrates into mRNA and small and big ribosomal subunit.

The end product is the protein in its primary form. Chaperones help the protein reach its final conformation (the protein is folded).
The protein undergoes further modification by the addition of glycosyl-rests, phosphate groups, lipids and other chemical modifications. Most of the time, the protein is only a subunit of a protein-complex. (Ribosomes are protein-RNA complexes for example.)

The protein then has to reach its final destination, some of them are even secreted out of the cell. The mRNA has the code for the initiation of corresponding signals. Proteins can still undergo further modification.

Def. Gene (updated): An area in the DNA which can be expressed and according to its template a polypeptide or RNA molecule will be produced that fulfils a function.

Gene regulation in prokaryotes:
Prokaryotes make use of operons and repressors in order to only transcribe a protein when it is scarcely or abundantly available.

Def. Operon: A specific DNA sequence which consists of a promoter that has an operator part and structural genes for a protein.

Def. Operator: Part of the promoter sequence. It has a site, to which an active repressor can bind and thus RNA polymerase is prevented from further advancing on the DNA strand.

Prokaryotes possess such operons on their DNA.
E.g. tryptophan-operon or short trp-operon: When there is a lot of tryptophan, it will bind to a repressor which will become active. In this case, tryptophan is said to be a metabolite. This can bind to the operator and preventing the RNA polymerase to transcribe further. Thus, the metabolite concentration remains balanced.

There are two kinds of operons. Repressed operons and inducible operons. Trp-operons are an example for repressed operons. Inducible operons are normally not transcribed, because a repressor is already preventing its transcription. There needs to be an inductor molecule and once inactivated, transcription can take place as usual. Such enzymes are called induced enzymes sometimes.

26.11.2016

Gene regulation in eukaryotes:
Since there are more than 200 different main cell types, gene regulation has to happen according to the cell type, which required eukaryotic life forms to come up with new possibilities of gene regulation.

The first method is to influence the chromatin structure. There are very dense areas, heterochromatin, which is not accessible for RNA-polymerase. These areas can be chemically modified by neutralizing the positive charge on the histones, which will loosen up the density at that point. It is now possible to access these genes.

The second method is DNA-methylation. DNA-methyltransferase catalyses the methylation of the bases, mostly cytosine rests that decreases the gene expression. Removing the methyl groups increases the expression though.

Another method of gene regulation is to regulate the initiation and rate of transcription directly. There are general and specific factors. General factors are virtually present in all cell types, such as unwinding the DNA and helping RNA polymerase reach the promoter. Specific factors are only occur during the transcription of the respective gene. One example are **cis regulatory elements (CREs)**. These are a specific DNA sequence on the same gene which is being transcribed. These control elements are also called enhancers. They can either be located up or downstream of the gene to be regulated. They only occur in eukaryotic genomes and they can be several kbp away from the target gene (the DNA can bend so that a CRE is just above/below the target gene).

Gene regulation at the mRNA level:
Removing the poly-A-tail or the 5’ cap will lead to the reduction of the mRNA. The 5’ cap end can be modified, so it cannot enter the ribosome anymore. On the other hand, non-translated nucleotide sequences in the 3’ area can lead to increased life span of the mRNA, making several translations possible.

Def. RNA interference (RNAi): The DNA codes micro RNAs (miRNAs), which can hybridise with mRNA in order to create a double stranded RNA. This regulates gene expression. Such a complex is either reduced or the gene expression is impeded.

Another possibility is to add small interfering RNAs (siRNAs) to control gene expression.

30.11.2016

Analogies between CRISPR/Cas9 and PCR: In both, the specificity of the primer defines uniquely where to bind and the length of the RNA sequence. For a primer to be statistically unique, it must have at least 17 base pairs.

Differences transcription vs. replication: Replication the whole genome is doubled and DNA is the outcome. Both strands are replicated at the same time. Transcription only takes certain parts to produce

RNA.
Different enzymes are used (DNA polymerase vs. RNA polymerase).
DNA polymerase needs an RNA primer. RNA polymerase does not need an RNA primer, since it can freely bind to the targeted parts.

Properties of CREs:
50-200 bp long; possess several, partially overlapping, binding sites for transcriptionsfactors, which can be activators or suppressors (-> information integration); DNA looping makes CREs independent of position of promoters (advantages for evolution).
=> vast flexibility in gene regulation; information integration; DNA bridges lead to new combinations of CRE.

Evolutionary advantages of introns and alternative splicing: …

Sequence-independent protection from viruses and transposons – The Rise of double stranded RNA in cytoplasm:
The RNA can make a double strand, since it is complementary after all. Cells do notice that and they have a mechanism to break down the double strand RNA. They pack the double-stranded RNA in protein complexes and process them. They can cut them sequence specific (see siRNA and miRNA).

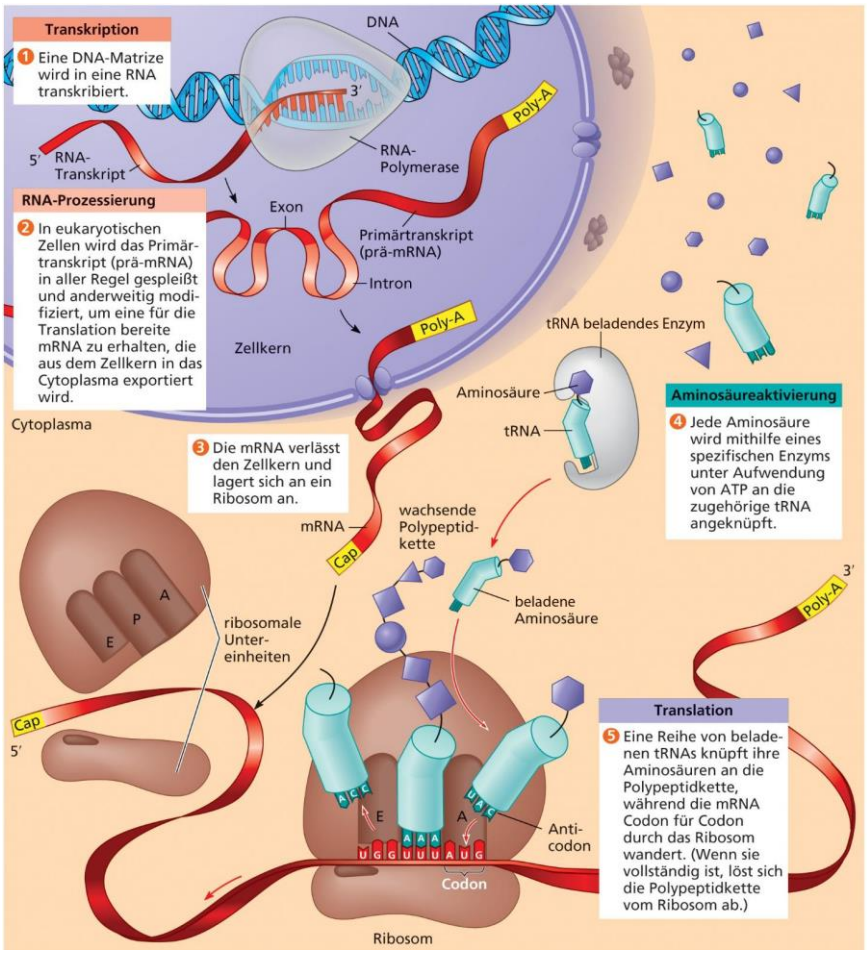
APPENDIX:

Def. upstream: In DNA, it is towards the 5’ end of the coding gene. Due to the DNA’s antiparallel nature, upstream is towards the 3’ end of the template strand.

Def. downstream: In DNA, it is towards the 3’ end of the coding gene. Due to the DNA’s anti-parallel nature, downstream is towards the 5’ end of the template strand.

Kor.: In RNA, upstream is towards the 5’ end and downstream is towards the 3’ end.

Summary of transcription and translation



1.12.2016

Chapter 17: Variability of Genomes

Def. Aneuploidy: The presence of an abnormal number of chromosomes in a cell (for example, trisomy 21).

What kind of mutation influences the function of a protein the least?
A nonsense mutation would mean that a codon is mistakenly turned into a stop codon which would render the current protein ineffective and useless.
Deletion or addition of a base pair would lead to the synthetization of another protein, since the codons would change.
The least effect would be a substitution of a base pair with another base pair. This can lead to different outcomes, such as a silent mutation (the mutation has no influence on the effect of the protein as a whole whatsoever) or to a missense mutation. The protein can still carry out its intended function, even if it does not possess the optimal form.

Def. Point mutation: Mutation in the DNA or RNA concerning only one base pair. There are 3 kinds: base pair insertion, deletion and substitution. Base pair substitution is the least detrimental for the function of

the protein (only 1 codon will change, all the other codons remain normal, while in insertion and deletion, every codon changes).

Bem.: Keep in mind that more than 98% of the human genome does not code for a protein which leads to a point mutation to occur rather seldom.

Differences in human genomes:
Point mutations: see above.
Differences in short tandem repeats: Short repetitive DNA sequences; they occur in different frequencies in different individuals.
Copy number variation: Longer gene regions have several copies in the DNA. Not all individuals have the same number of copies of these gene regions in their genome. Interestingly, this happens most commonly in the difference of human genomes. Normally, the gene only occurs twice – 1 for each chromosome. Some individuals possess the genes more than just once on their chromosome. There are several copies of the gene on the chromosome. The copies are normally right next to the original DNA region (they are not spread throughout the chromosome).

Keep in mind that two individuals are approximately 99.9% identical in their genome (this means only about 3 million base pairs are different).

Def. RFLP-analysis (restriction fragment length polymorphism analysis): The basic technique for detecting RFLP is to break down the DNA in smaller parts by restriction enzymes. These restriction enzymes can detect a specific short sequence of DNA and cut it accordingly. All the pieces are separated in a electrophoresis, where DNA pieces of similar length will be grouped together.

Def. Single-nucleotide polymorphism (SNP): Another word for the point mutation substitution case. SNP occurs every 100-300 base pairs. SNP in non-coding DNA regions can influence the genes in coding DNA regions. The SNP has to occur in at least 1% of the (human) population in order to be classified as a SNP.

Kor.: SNPs are useful for identifying diseases instead of analysing the entire genome. They do not say anything about the likelihood of the disease to occur. Certain diseases, such as cancer, need several SNPs and other environmental factors to occur.

Ex.: Such is the case with eye color. The gene OCA2 codes for the melatonin production for brown eyes. A SNP happens in an intron which somehow influences the exon containing the OCA2 gene.

Kor: SNP can also occur on a recognition site for restriction enyzmes, so that the restriction enzymes cannot bind to it anymore or a new recognition site can be made for restriction enzymes to bind.

7.12.2016

Genetic vatiations:
Chromosomal mutations: double strand break downs
Base substitutions: mutagene substances; errors in replications
Genomic mutations: aneuploidy
Deletions or duplications or insertions

Mutations are statistical entities.

Ex.: Assume eye color only depends on the presence of melanine. Blue eyes is the consequence of the absence of melanine. The most simple ansatz is that the transport protein is not working correctly and melanine is never transported to its location (eye). This could be due to a point mutation leading to an early stop codon. (This is a very simplified case.)

Def. Homozygote: A creature has the same allele twice (if it is a diploid organism).

Def. Heterozygote: A creature has two different alleles of the gene, one of them will be dominant, the other recessive. One of the alleles will be coding for the dominant phenotype.

Interaction between genotype and environment:

External chemical factors influence the expression of proteins. The affected protein is not expressed in sufficient quantities – the chemical factors reduce the expression.
(See the example of 1-alpha-trypsin and cigarette smoke).

APPENDIX: If a population has a especially low genetic variability then it is mostly because of the predecessor popluation that only had few individuals. This must have led to inbreeding, hence the low genetic variability.

Def. Proteasome: Protein complexes that occur in eukaryotes, archae and in some bacteria. In eukaryotes they are located in the cell nucleus and in the cytoplasm. They are responsible for degrading unneeded or misfolded proteins by a process called proteolysis – a chemical reaction that breaks down peptide bonds. These reactions are aided by enzymes called proteases. A protein to be degraded is tagged by the small protein ubiquitin.

A consequence of a non-reciprocal crossing-over is deletion and duplication. During crossing-over, it can be that similar but non-homolog sequences do crossing-over which then leads to an unequal exchange, leading to the deletion on one of the chromosomes (it lacks the sequence) and to duplication on the other chromsome. This phenomenon occurs in pairs.
After a chromosomal breakdown the DNA fragment can go over to a non-homolog chromosome. This is called translocation.
Most of genetic variation in a population is neutral.

Where does a mutation have to occur in order to be integrated into a gene pool?
A mutation must occur in during meiosis in the gametes in several individuals who then can reproduce amongst each other or if the mutation occurred in the family, they have to inbreed. If the mutation only occurs in the gametes of one individual alone, it will only be present on one homolog chromosome. The mutation does not necessarily need to be passed on to the second generation and so the mutation is lost.
Moreover, the mutation must not be disadvantegous for the individual. The person must still be able to reproduce and the mutation should not make the person less likely to survive under certain environmental circumstances or dead.
An advantageous mutation is more likely to be integrated in a population's gene pool.

10.12.2016

Chapter 18: Inheritance of Genomes

Def. Segregation: During meiosis 1 in anaphase 1, the alleles are seperated from the homologous chromosomes. A homozygotic being will only have one type of the allele, since it is a homozygote. The alleles will be at the same genlocus in all gametes. A heterozygote will have equally many alleles (in the most simple case, it is a 50:50 ratio).

Mendel's laws: Law of segregation of genes, law of independent assortment, law of dominance.

Not all phenomena can be deduced from the mendelian laws.

Alleles that are responsible for recessive phenotypes often code for a flawed protein or the protein is not expressed at all. In a heterozygote, it is not fatal at all, since it still possess a normally working allele, which can (most of the time) code sufficiently.

Sometimes, there are also codominant alleles. This is the case in sickel cell anaemia. Both normally functioning haemoglobin and defect haemoglobin is being produced in the case of a heterozygote. It makes them immune to malaria, while they can still live like people with normally functioning haemoglobin.

Def. incomplete dominance: An allele is neither completely dominant nor completely recessive. A heterozygote will have a different phenotype than the homozygotes.

Ex.: Let be C**w and C**r coding for white and red in plants. A heterozygote with the genotype C**rC**w will have neither red color nor white color but rose color, since it cannot produce enough red pigments with only one allele alone.

Kor.: In the above mentioned example, the ratio is 1:2:1 for red:rose:white in the parental generation. The filial generation will again have the phenotype segregation 1:2:1.

Def. Multiple alleles: A gene can have more than two alleles.

Ex.: Blood groups: The gene coding for specific carbonhydrate side chains on the cell surface can be of type A, B, AB, or 0.

Def. Epistasis: The phenotypic characterstic of one gene at its locus influences another gene.

Ex. (mice): Let B and b be the alleles coding for its fur color with B coding for black pigments and b for brown pigments. Let C and c be the alleles coding for the transport proteins whether or not the pigments reach the fur at all. This way, the C-c alleles will influence the phenotypic appearance of the color of the mouse, in which case it will be white in the case of BBcc or Bbcc.

Def. Pleiotropy: A single gene can influence more than one phenotype.

Def. Genomic imprinting: The phenomenon of the expression of genes (alleles) dependent of the origin of parent. If the allele is imprinted in the father (methylation, modification of histones, enhanced degradation of telomeres etc.), then only the allele coming from the mother will be expressed and vice versa. Therefore, one will be active and another one inactive (it can also be the other way around where imprinting an allele leads to its activation and a non-imprinted allele is inactive).

Kor.: This change (genomic imprinting) cannot be observed in the genotype for it does not change the DNA sequence – it can only be observed in the phenotype.

Def. Linkage: Genes that are located on the same chromosome. In many cases, they are physically close to one another and are thus often to be inherited together.

Kor.: Such genes are said to be coupled. During crossing-over, the genes are sometimes unlinked and distributed amongst both homologs. But it does not necessarily need to happen.

Def. Mitochondrial inheritance: The mitochondrial DNA (=: mtDNA) is located outside of the nucleus in the mitochondrion. The mitochondrion only possesses one round and small chromosome with 37 genes on it. The mtDNA is inherited only from the mother, since the egg cell has enough space to contain mtDNA etc – the sperm cell can only carry the father's genome, because it is so small. Thus, mtDNA is only passed on from the mother.

Def. Gonosome: Synonym for heterochromsome, heterosome, sex chromosome.

There are gender bound genes, which are only expressed in female when both X-chromosomes posses the recessive allele. In males, since they only have one X-chromosome and most of the time, there is no corresponding allele on the Y-chromosome, they phenotype will always be expressed and cannot be compensated by a second X-chromosome, since they do not posses as second one. One speaks of **hemizyogtes** and not homozygote.

The inheritance of the DNA of the organelles also diverges from the mendelian laws. Since a sperm cell is very small, it only carries the father's genome. But the egg cell is larger, thus enabling it to contain more information such as the DNA for mitochondria etc. Thus, the genes for the organelles are only passed on from the mother.

APPENDIX: What is the probability of my combination of chromosomes from my father and mother to have?
A human gets 23 from each parent. All the possible combinations from one parent is 1/2**23. Since humans are diploid, it follows: 0.5**23 * 0.5**23 = 1.42*10**-14.

Chapter 19: Genomes and Environment

Def. Complex characteristic: The phenotype is influenced by the genotype (often there are many genes involved) and by the environment.

Ex. Human growth: One cannot say that a 180 cm tall individual has gotten 90 cm from its genes and 90 cm from food. This is completely implausible to say. One has to compare a population in order to make conclusions regarding the genes (genetic variation in the population) and food. Answers will be of 80% of growth is caused by genes while the remaining 20% is caused by the food or so.

Homozygotic inbreds are used in order to compare influences of the environment on the genotype, thus changes in the phenotype can be observerd.

Def. Genotype-by-enviromnent-interaction: For two inbred populations, an enviromnental factor can interact differently with the genotypem causing different effects.

Def. Reaction norm: The interval of variation of a phenotype according to the interaction between genotype and environment.

Def. Heratibility: A measure for what percentage of differences in a characteristic is being caused by differences in the genotype of a population.

Kor.: Heratibility = 100% <=> all of the variation of a characteristic is caused by the genotype <=> the environment has no influence on the characteristica.
Heratibility = 0% <=> the variation of a characteristic is not influenced by different genotypes <=> all of the variations of a characteristica is caused by the environment.

SNPs do not influence the phenotype most of the time. They can only influence it if they occur in an exon.

Def. Concordance: % of the total number of affected pairs, in which both twins are affected.