## Molecular Biology

Department of Molecular Biology and Pharmaceutical Biotechnology (45-308)

> Mgr. Marie Brázdová, Ph.D. (FAF, BFU AVČR) RNDr. Jan Hošek Ph.D. (FAF) Mgr. Daniel Renčiuk, Ph.D., (BFU AVČR)

> > brazdovam@vfu.cz,

MB\_1\_2017

#### examination requirements:

- Exam has both written and oral part
- Test: min 60% to pass (E), ....90-95% (A)

#### Moodle- molecular biology 2016

 Molecular Biology of the Cell, Fifth Edition, Alberts, Johnson, Lewis, Raff, Roberts, Walter, Garland Science Publ. 2008



## Suggested literature

 Molecular Biology, Understanding the Genetic Revolution, Clark D.P., Elsevier, 2005





 Lewin's Cells, Second edition, Cassimeris L., Lingappa V.R., Plopper G (eds), Jones and Barlett Publishers, 2011

#### outline of the first lecture

- defining molecular biology
- historical context
- living systems: types, properties
- prokaryotic x eukaryotic cells
- cell-free systems
- model systems in molecular biology

## What is molecular biology?

- field studying relationships between physical structure, mutual interactions of biomolecules and the properties of living systems
- explores relationships between two levels of living systems: between physical/chemical (structure and function of biomacromolecules) and biological (properties and functions of cells/organisms

## Molecular biology x Biochemistry

- biomolecular interactions, mainly information biomacromolecules, the functions and properties of living systems.
- The goal of molecular biology is to explain the functions and features of living systems and the relationships in between the structure and interactions of their molecules (macromolecules).
- These relationships are explained from a complex point of view integrating physical, chemical and biological methods.
- Studies processes that take place in living systems at the molecular level and in which genetic information is processed

#### Biochemistry

• Biochemistry addresses chemical processes in living organisms. The field of study in biochemistry is the structure and function of the basic building blocks of living matter like sugars, fats, proteins, nucleic acids and other biomolecules.

#### Principles of Molecular Biology are making their way into all biology fields

- originally focused on heredity maintenance and expression of genetic information and its inheritance (molecular genetics)
- currently concerned with common characteristics of life at the molecular level, which are valid across the outer diversity of living creatures and the essence of processes taking place in cells and organisms
- practical applications: biotechnology, paleontology, archeology, evolution, medicine, anthropology, criminalistics, determining parenthood, etc.

## Importance of molecular biology

- research moves to a level that is common to all living systems
- Enables cellular manipulations- affect their characteristics in desired the direction (human and veterinary medicine, biotechnology, gene engineering, agriculture, etc.)
- Brings in new possibilities to study phylogenetics, evolution, etc.
- Molecular biology belongs along with information technology to the most progressive fields of present times

## Historical background

- Crossbreeding of plants and animals is done for thousand of years, yet genetics as a science arose only in the 80s of the 19<sup>th</sup> century
- Techniques for investigating the associations between genes and observable traits are available only a few decades
- The era of DNA cloning and executing targeted alterations to genetic information began in 70s of the 20<sup>th</sup> century (Stanford University, UCSF, Genentech) – biotechnology used to prepare drugs



Colony of cells each containing copies of the same recombinant plasmid

- Gathering information about molecules of life dramatically speeded up in the last several years
- milestones: DNA double-helix discovery, central dogma of molecular biology, restriction endonucleases, gene cloning, reverse transcription, DNA sequencing, h PCR, monoclonal antibodies, microarrays, RNA interference, stem cells, artificial nucleases, etc...





#### Important milestones in molecular biology

- 1944 purified DNA used to transform bacteria
- 1953 model of DNA structure (J. Watson, F. Crick, M. Wilkins)
- 1956 genetic information is written in DNA as a sequence of bases
- 1958 complementary DNA strands separate during replication
- 1958 isolation of DNA-polymeraseI and DNA synthesis in vitro
- 1958 postulating central dogma of molecular biology
- 1960 discovering mRNA and prooving its function
- 1961 mRNA used to decipher genetic code
- 1961 experimental evidence to central dogma of MB
- 1961 operon theory postulation regulation of gene expression
- 1966 complete deciphering of genetic code
- 1970 first restriction enzyme was isolated
- 1970 reverse transcriptase discovered in retroviruses
- 1972 first recombinant DNA molecules prepared *in vitro*
- 1973 beginnings of gene clonning GI basis

1958 - postulating central dogma of molecular biology

## The central dogma of molecular biology



• genetic information is passed by transfer of the DNA into RNA and protein (this postulate is called the central dogma of MB)

• In some viruses (e.g. HIV), RNA is used as a template for DNA synthesis by reverse transcription

 many genes encode a polypeptide, however RNA molecules in the cell play important roles



#### Important milestones in molecular biology

- 1975 Asilomar conference moratorium on work with recombinant DNA
- 1977 first recombinant molecules to carry mammal genes
- 1977 discovery of composed genes exons/introns pre-mRNA splicing
- 1977 introducing DNA sequenation
- 1981 discovery of RNA catalytic activity ribosomes
- 1982 human insulin produced commercially inside bacteria
- 1983 bacteriophage A whole DNA sequence acquired
  -starting projects to sequence genomes of model organisms

Genomic and postgenomic era

- analysing genome sequences
- genome × proteome
- bioinformatic approach



#### Present

- Since 2003 we have access to DNA nucleotide sequence that constitutes the human genome (approximately 20.000 genes, 23 chromosome pairs, 3x10<sup>9</sup> base pairs/haploid cell)
- •We thus have the information necessary for the establishment and functioning of human beings
- only 1,5% of the human genome consists of genes encoding the proteins, the rest consists of areas coding RNA, regulatory sequences, introns
- The functions of many human genes are not sufficiently explored, even less is known about non-coding sequences

#### Present - challenges

Understanding of the mechanisms coordinating gene expression and their relationship to human health (all diseases have hereditary component that gets co-decided by genes)

Curing hereditary diseases: resulting from faulty gene function carried over from parents to offspring – to understand the cause we must first understand the function of healthy genes

□Cancer treatment: also a consequence of gene dysfunctions, search for early diagnosis markers and appropriate targets for gene therapy, search for new chemotherapeutic agents

□Gene engineering and biotechnologies in agriculture: food source improvement (resistance to external adversities)

#### What is life?



**Usatisfactory** definition does not exist

Generally accepted characteristics of living organisms:

Dhighly organised

**Composed of one or more cells** 

□contain their own plan of arrangement, ie. the genetic program (genotype)

Dacquire and use energy

Dimplement and control many chemical reactions

- grow and change their appearance and abilities (ie. phenotype)
- keep a relatively constant internal environment
- □ reproduce
- **respond** to changes in the environment
- They may evolve over time

#### What is life?

- □ a key feature is the ability to reproduce. That is replicating the genetic information (genome), and the structure that is its bearer and protector (cell).
- □ growth and reproduction requires information and energy
- energy is one of the products of metabolism
- Metabolism is a set of processes, by which are nutrient molecules transmitted and converted so that the cells are provided with energy and new building materials

#### **Phylogenetic Tree of Life**

#### Living organisms evolve



- □ life on earth is changing by a process of evolution
- Organisms which are identical or very similar to parental organisms arise by reproduction
- but genetic information in progeny/offspring generations gradually accumulates changes
- accumulation of changes is related to the properties of molecules of nucleic acids and external environment
- □ a essential feature of life is that it presents a dynamic balance between its precise duplication and tolerance to changes



□Genetic information: biological information written in nucleic acid (DNA or RNA) oligonucleotide sequence

□gene: unit of genetic information determining protein/RNA structure (physically: nucleic acid segment)

**DNA**: serves long term genetic information storage

**DRNA:** is involved in mechanism by which genetic information is put into practical use

Protein

Gene

#### Terminology



- □genome: the sum of all the genetic material of an organism, cell or organelle, which copy is transmited to offspring (genes; non-coding sequences are included as well)
- □genotype: genetic constitution of an organism represented by a set of alleles of its genome (relative to individuals of that specie)
- phenotype: the sum of characteristics, genotype manifestation in the environment
- metabolism: a set of processes for energy and building the substances necessary for the biosynthesis of cellular components; term growth, reproduction and realization of genetic information

□ribosomes: key components in cell apparatus, create cellular proteins according to gene instructions

## Living organisms are made up by cells

 variability of the manifestations of life versus the unity of their foundations

 cellular structure based on similar ingredients occurs at all life forms

 Schleiden and Schwann in 30s of the 19th century: cells are the building blocks of life

Theodore Schwann (1810-1882)



 All living organisms are composed of one or more cells.
 The cell is the basic unit of structure

and organization in organisms.

3. Cells arise from pre-existing cells.



Natthias Jakob Schleiden (1804-1881)

#### Cells can differ substantially from each other...

- variability in size and shape...
- can live freely or be tied to a matrix or other cells



#### ...still they have much in common

- Are separated from the outer world by plasma membrane (with characteristic phospholipid structure)
- Plasma membrane systems control import into the cell and export from the cell
- Cellular structures are build up from food molecules processed by cells own systems using energy
- Contain genetic material carrying information needed to create/restock all cell components
- Dispose of system for <u>gene expression</u> through which genetic information is conveyed into practical use
- Particular proteins or RNAs (defined by genes) can form structures of higher order

# <u>Minimal requirements</u> for the existence of a living cell

- Plasma membrane (to separate cell from outer environment)
- Ability to build up cellular structures from simple molecules via utilizing external energy sources
- Presence of genetic information defining cell characteristics
- Presence of system to express genetic information



## Unicellular organisms



- the simplest organisms
- they live as separate units
- reproduce
- successfully adapt to extreme conditions (high / low temperature, aerobic / anaerobic, etc.).
- often live inside other organisms

## Multicellular organisms



- cells specialize in certain functions differentiate, that is changing gene expression in different cells of the same organism, leads to their phenotypic diversification
- some cells maintain the undifferentiated state (stem cells)
- significant structural and functional specialization of cells, leads to division of labor inside the body (often accompanied by a loss of ability growth and division)
- cells communicate with each other to ensure proper function of the whole organism



#### Plasma membrane - function

- ensures cell autonomy (cellular components exist in a limited space)
- ensures that the aqueous medium within the cell differs from the outer medium
- contains protein complexes that control import and export of molecules through the membrane
- transmits signals between the external and internal cell environment

#### Plasma membrane - structure

two-layer unit composed of phospholipids and proteins

#### Phospholipids

- Contain a water soluble head, which includes phosphate, exposed towards inner and outer membrane surface; and a hydrophobic part (a pair of fatty acid chains), which form membrane body



Molecules that carry both hydrophobic and hydrophilic area called amphipathic

#### Plasma membrane - structure

 aquatic environment inside and outside the cells results in aggregation of hydrophobic lipid strings (forming the inner environment of the membrane), hydrophilic (charged phosphate) heads form the outer surfaces of the membrane

 lipid bilayer membrane allows for merging membranes and creates suitable environment for binding proteins



#### Plasma membrane - permeability

- membrane is not freely permeable to ions, small charged (hydrophilic) molecules and large molecules
- water molecules and hydrophobic molecules can pass through the membrane
- different ion concentrations on both sides of the membrane may create osmotic pressure (the movement of water molecules through the membrane towards the environment with a higher concentration of ions)
- risk of cell damage





## Cells manage their content



- intracellular pH and ion content are subjected to tight regulation
- specialised membrane transporters help to achieve required balance
- metabolism substrates (energy sources) must also pass through plasma membrane, as well as basic components of cellular structures
- unwanted metabolic products or ions must be exported through the membrane
- both directions have a high transport selectivity
- the ability of cells to maintain stable internal environment is called homeostasis

#### Protein channels control ion passage through plasma membrane

- •create a channel in lipid bilayer
- •the outer surface of protein channel is in contact with membrane lipids
- inner surface is surrounded by aqueous medium
- ions and hydrophilic molecules pass through the channel without getting into contact with membrane lipids



#### Ion channels and protein carriers

- Two mechanisms to transfer ions across membrane: according to gradient direction or against the gradient
- Ion channels are used to transport ions (without energy requirements) from places with higher ion concentration to places with lower ion concentration (in gradient direction)
- Protein carriers transport ions against electrochemical gradient, which requires energy

#### Controlling ion channels

- Cannot be open all the time (ion concentrations would be equalized)
- Open up by undertaking a conformational shift
- conformational shift can be induced by various stimuli (signal molecule, electric voltage, temperature)



## Cell membrane is not a firm/hard structure

- But a weak, flexible and fragile one
- Mechanic support is often provided by cell wall, placed outside cell membrane, common in bacterial and plant cells
- Animal cells are supported by cytoskeleton components


### Two types of living cells



- Classification by internal division, compartmentalization (compartment = membrane enclosed space)
- prokaryotic cell: simpler, consists of only one compartment with the genetic material, apparatus, and products of gene expression
- eukaryotic cell: complex, comprising at least two compartments, one of which contains the genetic material

#### Prokaryotes

- Unicellular organisms
- cells are surrounded by a membrane
- membrane surrounded by a cell wall providing protection
- DNA located in the cytoplasm
- it includes all chemical and structural components necessary for life
- all the genetic information present on one chromosome



### Two types of prokaryotes: bacteria and archaea



 phylogenetic relationship of these groups was determined relatively recently modern methods (particularly DNA sequencing): archaea are a separate group of prokaryotes

- archaea have the appearance and structure similar to bacteria: Small single-celled organisms without internal membranes
- they live in extreme conditions (high temperatures, acidity, salt content)
- archaea often use unusual metabolic pathways, show chemical differences in the construction of cell walls, have an apparatus for gene expression, which is more like in eukaryotes than bacteria

## The genetic material of prokaryotes



- the simplest genome are present in bacteria that live freely but within other organisms (Mycoplasma)
- Host provides compounds which bacteria need, but cannot create
- their genome contains only about 500 genes that encode a basic structural cellular components
- genome of bacteria comprises at least 1500 genes that encode the structural elements in addition to enzymes and also more advanced system for gene expression regulation

## classification of bacteria



Hans Joachim Gram (1853-1938)

2 groups that diverged about 2 billion years ago:

- Gram-negative (e.g. E. coli) and Gram-positive (e.g. B. subtitlis)
- depending on how they react with Gram stain
- sensitivity given by dye interaction with the cell wall

# Gram staining



G+ bacteria have plasma membrane surrounded by a cell wall formed by proteoglycans and polysaccharides (blue / violet color) Gram-negative bacteria have the outside wall surrounded by lipopolysaccharide layer (red / pink color)

# Bacteria – a model for research of basic cellular processes

- advantages:
- unicellular microorganisms, uniformity of response to external stimuli
- low number of genes
- haploid state (only one copy of each gene)
- possibility of culturing under strictly controlled conditions (defined precisely by medium with the content of salts and carbon source)
- high growth rate, doubling time is only 20 minutes
- the possibility of storage in cold boxes (-70 °C) for 20 or more years

### Model bacterium - Escherichia coli

- rod-shaped bacteria (size 1 x 2.5 micron)
- the natural environment of the intestine (colonbecause "coli")
- belongs among Gram-negative bacteria





### Where can we find bacteria in nature?

- Almost everywhere (60 km high in the atmosphere, 11 km below sea level, in both fresh and salt water or sewage, soil, plant roots or in bodies of animals)
- the number of bacteria on earth is huge (amounts of bacterial carbon on Earth corresponds to the amount of carbon in plants)
- the bacteria is likely to constitute more than half of live mass of Earth

### Prokaryotes are very adaptable

- their lifestyle is varied and can adapt to extreme conditions (pH, presence of oxygen, temperature, etc.).
- Classification according to the ability to grow at different temperatures:
- mesophilic: grow best between 25 and 40 of °C (this includes the human pathogens)
- psychrophilic: grow best between 15 and 20 °C (but there are also those that live at 0 °C) a favorite environment cold water and land
- thermophilic: grow best between 50 and 60 °C (some tolerate even 110 °C)

### Bacteria are sensitive to natural substances

- Various bacteria in the same environment compete for resources
- excretion of toxic proteins, e.g. bacteriocins kill related bacteria, but no strains - producers (utilization of bacteriocin plasmids as vectors)





• Creation of antibiotics (clinical use)

### Some bacteria are harmful, others are useful

- a small portion of pathogenic bacteria: the originator of infectious diseases (cholera, tuberculosis, anthrax, syphilis, gonorrhea, whooping cough, diphtheria, etc.), what contributes to the eradication: sanitation, clean water, soap, flush toilets, as well as vaccinations and antibiotics
- Most bacteria have a positive meaning: it contributes to the balance of ecosystem (decomposition of dead bodies of plants and animals)
- destruction of waste products of human activity and pollution
- higher life forms could not survive without bacteria

# Eukaryotic cell contains several compartments

- increasing the complexity of the cell the division of labor at the cellular level, interior space is divided into two main sections sealed by membrane: cytoplasm and nucleus
- macromolecules are transported into and from the nucleus via the nuclear pores (protein channels)

MB 1 2017

• pores are fully permeable to smaller molecules, i.e. the aqueous environment of the nucleus does not differ from the cytoplasm



# Genetic material is located in the nucleus

- unlike prokaryotes eukaryotes genome must ensure specification of new structural elements, define the location of proteins into the correct sections, use a more complex mechanism of gene regulation
- Genetic complexity of eukaryotes: the simplest single-celled organisms eukaryotes carry about 5,000 genes, human about 20000 genes encoding proteins



### Eukaryotic cell contains additonal organelles



- surrounded by membranes with the same structural features as the plasma membrane
- cytoplasmic = organelles + cytosol
- cytosol is an aqueous environment of the cytoplasm, where proteins are synthesized, which remain either in the cytosol or are transported to any organelle or outside of the cell
- organelle membranes are permeable, so the inner organelles environment differs from the cytosol (the exception is nucleus with its pores)

# Eukaryotic cell contains additonal organelles

- endoplasmic reticulum protein folding, assembling oligomers , (ribosomes-translation)
- Golgi apparatus secretion of proteins and other material outside the cell
- lysosomes containing digestive enzymes decay of molecules
- mitochondria size, shape, containing circular DNA, resemble bacteria - a specialist in energy recovery respiration
- chloroplasts specialized for photosynthesis, contain chlorophyll and other molecules necessary for capturing light, containing circular DNA like mitochondria

# The diversity of eukaryotes

- Unlike two genetically distinctive branches of prokaryotes, all eukaryotic organisms are genetically related
- originate from a common ancestor
- unicellular eukaryotes: yeast, protozoa
- multicellular eukaryotes: plants, fungi and animals

Two main cell types of eukaryotes

- somatic cells and gametes
- in most multicellular organisms, cells are specialized into tissues and organs
- •Gametes are part of the reproductive system and participate in the formation of the next generation
- somatic cells form the body, creating good conditions for the functioning of the gametes, do not participate in reproduction

# Model organisms

• For practical reasons they are preferred in research

- It assumes that the gained knowledge will apply in the other organisms, at least for those related to model organisms
- it is not always true for human medicine or agriculture that we should examine directly target organisms

### Yeast - model unicellular eukaryotes

- unicellular eukaryote similar benefits as bacteria
- genome was sequenced first (in eukaryotes) in 1996
- belong among fungi similar to animals and plants
- alternation of diploid and haploid phase, you can work with haploid cultures, which facilitates genetic analysis
- haploid genome contains only about 6,000 genes

## Yeast



- only few genes (about 5%) contain introns
- They grow in a chemically defined medium, forming colonies on agar plates
- the generation time of approximately 90 minutes
- the ability to easily preserve in frozen state
- specific multiplication by budding
- model is suitable for studying the function of genes and cell cycle

# Nematodes

- Caenorhabditis elegans
- non-pathogenic soil nematodes
- genome has a 7x higher DNA content than yeast
- higher content of introns and non-coding sequences
- 1 mm body composed of 959 cells
- evolution of each one from the original zygote is described
- model for studying development, apoptosis, aging
- RNA interference was first described in this model









S. Brenner (nar. 1927)<sup>8</sup>

# Flies





T.H. Morgan (1866-1645)

- Drosophila melanogaster:
- life cycle of two weeks
- •14 000 genes
- research on cell differentiation, development of an organism, cell signaling and behavior

# Zebrafish - a model for the study of evolution of vertebrates



- Zebrafish/Danio rerio: small freshwater fish (2,5 cm)
- fertilized egg develops outside the womb can be monitored by microscopy
- Development from egg to adult organism takes 3 months, thanks to the transparency we can follow the development of internal organs
- Easy microinjection of foreign DNA into the egg
- Molecular genetics of development
- 25 chromosomes, 75% homology with the human genome

### Mouse



- model organism for humans
- lives 1-3 years, reaching sexual maturity after 4 weeks
- contains a 20 chromosome pairs
- less than 1% of mouse genes does not have a human homolog
- used to study gene function

# Human

- for ethical reasons we can not experiment with people
- it is possible to cultivate human or animal cells in culture
- immortal cell lines (e.g. HeLa cells) are formed by tumor cells
- HeLa cells are tumor cells derived from cervical cancer of Henrietta Lacks in 1951
- cell line is a suitable model for molecular biology studies

## Arabidopsis - model for plants

- molecular biology of plant has historically been somewhat lagging behind other organisms
- often they have many genes (rice: 40 000 to 50 000 genes)
- reason: they can not protect themselves by movement, thus accumulate genes for protection and adaptation
- the challenge now: genetic improvement of crops
- Arabidopsis thaliana: structural simplicity, small genome, five pairs of chromosomes and about 25,000 genes
- maturing plants to produce seeds takes 6 to 10 weeks
- can be cultivated in the haploid state advantage for genetic analysis





does not have the cell structure



- contain their own genes wrapped in a protein coat, but it can not express them
- lacks apparatus which ensures the cells energy
- cellular parasite, replication and expression of viral genes is carried out by infected cell
- viral genetic information is stored in the DNA or RNA
- he is able to cause disease
- applies in genetic engineering

### Bacterial virus = bacteriophage

- infects a bacterial cell that is caused to produce new bacteriophages, eventually bursts, releasing a new generation of bacteriophages
- each of the new virions can infect additional bacteria
- within a few hours phage epidemic can destroy bacteria cultures of the number exceeding several human population

### Human viruses



- infects human cells, originator of common diseases (measles, mumps, chickenpox, colds and flu) and serious diseases (polio, Ebola, AIDS)
- Infection of less dangerous virus may provide resistance against much more dangerous virus
- viral infection can not be cured; necessary prevention immunization
- can transfer genes from one organism to another host (importance for evolution and genetic engineering)
- Antibiotics have no sense in fighting viral infections, can only help with parallel bacterial infection

- are widespread in the biosphere
- various functions: they can cause serious illness or existence is almost unnoticeable
- They carry genetic information, but do not have tools for the life functions, do not exist outside the host cell
- Viruses are one of the most advanced
- viroids and plasmids: autonomous nucleic acid molecules that do not have protein shell
- viroids are RNA molecules that infect plants and force them to produce new viroids released into the environment

#### plasmids

- DNA molecules that are stably maintained within the host cell
- They may pass from one cell to another only if between them there is contact
- do not kill the host cell
- widely used in genetic engineering



- transposable elements (transposons)
- DNA molecules that do not replicate as separate units
- for its replication they require incorporation into other DNA molecules that have self replication ability
- have the ability to jump from one host DNA to another

### prions

- infectious protein molecules
- they contain no nucleic acid
- infect cells of the nervous system of animals and cause serious illness (eg. mad cow disease)
- represent incorrectly folded version of the normal protein of the nerve cells
- if they penetrate into the cell they cause incorrect folding of the corresponding normal protein which kills the cell

## The central dogma of molecular biology



• genetic information is passed by transfer of the DNA into RNA and protein (this postulate is called the central dogma of MB)

• In some viruses (e.g. HIV), RNA is used as a template for DNA synthesis by reverse transcription

 many genes encode a polypeptide, however RNA molecules in the cell play important roles

# p53 tumor suppressor and p53 binding to human genome



maintaining genomic stability

- transcription-dependent tumor suppressi
  - in response to various inputs, the p53 protein becomes stabilized
  - upon stabilization of p53, various transcriptional outputs determined by:
  - the strength of the p53 RE
  - posttranslational modification status of p53
  - specific p53 binding partners
  - the epigenetic landscape of target gene promoter
  - conformation of DNA, non-B DNA structures

#### How does p53 select its targets?

# Selected oncogenic properties of mutant p53 and their underlying mechanisms



#### Nature Reviews | Cancer

Ran Brosh & Varda Rotter Nature Reviews Cancer 9, 701-713 (2009)



#### Structure/sequence motifs in DNA

MAR/SAR elements in DNA (mutp53CD, C-terminus) (Muller  $Mgt_1al_{0.17}1996$ ; Deppert et al., 2000) 73

#### Mutation of TP53

