

## A General Introduction to Genetics

### First stages of evolution

The earliest forms of life appeared about  $3.7 \times 10^9$  years ago. Development of life on earth (biogenesis) was first a chemical, abiotic evolution, taking place in the oceans in the presence of phosphates ( $XPO_4$ ), silicates ( $XSiO_4$ ), metal ions, an atmosphere of nitrogen ( $N_2$ ), ammonia ( $NH_3$ ), carbon dioxide ( $CO_2$ ), methane ( $CH_4$ ), sulfur hydrogen ( $H_2S$ ), hydrogen ( $H_2$ ) and energy sources of heat, radiation and electric discharges. Formed were mixtures of amino acids, proteinoid microspheres with first forms of a membrane, metabolism, growth by budding.

Biological evolution with nucleic acid chains capable of reproduction progressed in self-organization of matter with the basic forces of evolutionary drive of mutation, recombination, divergent development of structures and forms with specialization of functions, selection. Formed were protobiontes, containing a short DNA strand, which differentiated stepwise an improved metabolism, protein production, a multifunctional membrane. The first prokaryotes of blue algae and bacteria appeared. Following in the first evolutionary line were eukaryotes with differentiated organelles within the cell and a membrane enclosed nucleus, containing a chromosome set to control cell division by mitosis (start of phylogenesis). Their oldest known chalky fossils found in oceanic sediments are about  $1.5 \times 10^9$  years old. The spreading one cell organisms took up mainly carbon and hydrogen containing molecules in exchange for nitrogen and oxygen to radically change the composition of the atmosphere, starting  $2 \times 10^9$  years back, into the one we know today. During the Upper Precambrium of about  $9 \times 10^8$  years ago the evolutionary rate of the diverse aquatic one cell organisms accelerated, forming multi cellular eukaryotes with specialized cell functions. The branches of plants and animals separated about  $1 \times 10^8$  years back, introduced was sexual reproduction, further accelerating the evolutionary rate, and set up was a heterotrophic food chain with plants at the base, before the first forms of plant life appeared on land.

The first evolutionary line is continuously traceable in the development of genetic materials, their proteins (phylogenetic topology) and ensuing forms of life. The first nucleic acid chains grew by processes of base pair changes, addition, deletion, inversion, duplication, rearrangements, activation, deactivation, in later stages by production of catalyzing enzymes, by interaction of these factors, mostly enlarging the overall DNA material. The development of function specific genes and proteins is graphically demonstrated by a polygenetic tree (molecular phylogram), by branch order and lengths, indicating their degree and distance of relation, the evolutionary steps and the corresponding evolutionary rates. The molecular tree of lineage delineates copy true the evolutionary tree of comparative anatomy of all plant and animal species, stating a common ancestry of all living organisms in the bio-chemical building blocks, the genetic code, the bio-synthesis of proteins, the cata-lysis by enzymes, an energy metabolism with glycolysis.

The phylogenetic theory of descent serves as basis for description, denotation and categorization (taxonomy) of all organisms. As a result of evolutionary processes, there exist in discontinuous variability today about 500 000 plant and 2000 000 animal species. The degree of relationship between groups, traced in a hierarchical, monophyletic tree of lineage, is measured by singular, homologous, derived traits in descent of corresponding original traits (taxon, pl. taxa). The taxonomic categories are rooted in four kingdoms (regnum) of one cell organisms, prokaryotes, euka-ryotes and mushrooms.

## The cell nucleus

The human genetic material, the genome, is stored on 2 sets (diploid) of 23 homo-logous chromosomes (22 autosome and 1 sex chromosome) and like all eukaryotes confined in a cell nucleus. The hereditary information is passed on as coded sequences and lengths, triplets or codons, of Desoxyribonucleic acids, DNA, selected from two purine bases guanine (G), adenine (A) and two pyrimidine bases cytosine (C), thymine (T). The codons, yielding  $4^3 = 64$  possibilities, encode for regulative signals and 20 essential proteins, the base group, accounting for about 200 000 proteins of the human body. The codons are arranged commaless, non-overlapping. They constitute a universal transcription code for all living organisms. About  $6 \times 10^7$  bases, one the primary structure, are joined by addition polymerization to form out a single macromolecular, chromosomal strand. Two complementary strands are arranged to a right handed, antiparallel double helix, the secondary structure. The interwinding strands are held together by hydrogen bonds between opposite C-G and A-T base pairs (bp). Each chromosome carries about 50 000 genes, one the functional unit, forming out a trait that is being passed on by Mendelian inheritance, consisting of single copy sequences with about 1 000 bp or polygenetically of repetitive, multicopy sequences with 2 – 10 gene copies with 20 – 500 bp each or of interspersed, with non-functional groups (intron) alternating, multicopy sequences. The double helix with a diameter of  $2 \times 10^{-9}$  m coils itself on nucleosomes, the tertiary structure, compressing the entire genome (chromatin) into a cell nucleus of  $6 \times 10^{-6}$  m in size.

## Proteins

Proteins, polymeric amino acids, containing a peptide bond in the repeat unit  $-(RCH(NH_2)COOH)-$  with a molecular mass of 10 – 100 000 and a chain length of 50 – 1000 units, constitute the basic building blocks of all organisms. They represent up to 50% of structural cell material and serve as regulative, storage, immune active proteins. They are synthesized mainly in two steps, a process called gene expression, by transcription within the nucleus and after transport by translation in the cytoplasm of the cell, both proceeding over the phases of initiation, elongation and termination. In transcription, realizing the encoded genetic information, a gene sequence of a locally unwound chromosome string, the template, is copied base by base onto a single stranded messenger RNA (mRNA), the matrix. In translation, assisted by ribosomal binding sites, the matrix directs bio-synthesis, the produced amino acid units being polymerized by addition, unidirectionally to a polypeptide chain, followed by folding, function specific modifications, transport.

## Growth processes

An organism's life cycle over the stages of zygote, embryo, youth, adult, death (biology of development) in regular succession of generations is fueled by species specific, somatic (non-germ) cell growth, quantitative increase of cell tissue and by differentiation, qualitative expression of specific cell functions and organs. The morphological changes (morphogenesis), development and arrangement of cell populations in precise positioning and organized manner, are regulated by temporal genes for cell specific initiators, transcriptional and translational control and by external factors like intercellular signals, often hormones, in equilibrium with anabolic and catabolic metabolism. Growth of a cell type is achieved by cell division and subsequent increase of cytoplasmic volume. The periodic cell cycle proceeds in the steps of cell division (mitosis (M)), gap (G1), synthesis (S), gap (G2). In nucleus (karyokinesis) and cell division (cytokinesis) the chromosome set is separated to distribute the two homologous halves to the daughter nuclei. In synthesis, the haploid sets in each nucleus are replicated over 10 000 replication units per chromosome simultaneously for a copy true, continuous passing down of the genetic information to the next cell generation.

## Sexual reproduction

Animal cell hybridization takes place in the sexual replication process in purpose of reproduction, the production of new living organisms to guarantee the continuity of the species. It proceeds in the sexual cycle in three successive stages: The male (spermatozoa) and female (ova) germ cells grow in germ cell production (gametogenesis) mainly out of meiosis, two cell divisions of meiosis I, proceeding in 9 phases and meiosis II, a mitotic division, proceeding in 5 phases, achieving random assortment of chromosomes in the germ cells and a reduction of the diploid chromosome set to one half. In cell fusion (karyogamy, conjunction of nuclei in copulation) the gametocytes with their haploid chromosome sets are brought together to form a fertilized egg cell (zygote), a randomly recombined diploid chromosome complement, preserving a constant number of chromosomes. In the third, diploid phase, the zygote develops into the embryo, the daughter generation (ontogenesis), by successive cell divisions and development to an adult with formation of sex organs to complete the sex cycle.

## Hereditary traits:

The genotype of an organism, the complete set of genes, determines the hereditary traits and forms out in steps of species specific development the phenotype of the organism (phenogenesis), the visible and empirically verifiable manifestation of a morphological form. The phenotype is codetermined by a multitude of competing factors like the organism's environment, by humans also by anthropological conditions, especially social and personal environments, which change repeatedly over a life span. A genetically hereditary trait is based on an organism's identical replication and distribution of alleles to daughter cells, on a selected biosynthetic pathway (gene expressivity or penetrance), on timing of gene expression of the required gene at the required time of development from the complete genome present in each cell (totipotency). Growth and differentiation of functions over the stages of embryo, youth, adult (ontogenesis) form out the full complexity, capacity, coordination and flexibility of the phenotype's hereditary traits. In humans, least determined by its genotype are behavioral traits, because of the enormous variety of developmental pathways of the central nervous system.

## Laws of inheritance

Mendel's laws of inheritance (1865) describe the genetic recombinations of allele pairs in sexual reproduction over successive generations, visible as hereditary traits, where the parent generation P differs in one allele on their diploid chromosome set with a pure homozygote wildtype  $a^-a^-$  and a pure homozygote mutagenic type  $a^+a^+$ . The variability of the genome is passed on in new combinations, where the progeny's ratio of genotypes is statistically predictable. Mendel's laws therefore serve as the genetic basis for breeding technologies.

1<sup>st</sup> law of uniformity: Crossing of two pure bred homozygote strains P with the allele combinations  $a^-a^-$  and  $a^+a^+$  results in a first daughter generation  $F_1$ , which is uniform heterozygote in genotype  $2a^-a^+$ . The trait expressed allele is called dominant, the unexpressed recessive, codominant alleles will form out an intermediate quality or intensity.

2<sup>nd</sup> law of segregation: Crossing of the heterozygous  $F_1$  generation results in a second daughter generation  $F_2$  with randomly distributed allele pairs, in average a relation of genotypes of 1:2:1 or  $a^-a^- : 2a^-a^+ : a^+a^+$ . The phenotypes split correspondingly 3:1 with a trait dominant allele, 1:2:1 with trait codominant alleles.

3<sup>rd</sup> law of independent assortment: Crossing of polyhybrid  $F_n$  strains with the non-linked allele combinations ab and cd results in a daughter generation  $F_{n+1}$  with a free combination of allele pairs, where the gene loci separate and new genotypes and phenotypes may arise that are not present in the  $F_n$  generation.

Breeding technologies:

Breeding techniques have been employed since prehistoric times of about 10 000 years. Improving plant and animal traits of quality and form like nutritional content, yield, adaptability, resistivity, freshness, has given a major contribution to human civilization. Selection, crossing and cultivation, utilizing genetic variability and hereditary traits, reduce the genetic reserves, which are also depleted by destruction of biotopes of wildtypes. Breeding (mating) systems today describe all essential factors aside from mutation, which control population structure and evolutionary divergence.

Breeding of a phenotype is determined by the breeding value of the trait; on the morphological level by the kind of sex organs present, mostly dioecious, where a partner is required to contribute the second nucleus; on the genetic level by fertilization factors to inactivate cell specific restrictions for gametes to fuse; by contribution of the number of chromosomes and nuclei to karyogamy; by allele frequencies; by gene expressivity.

Main plant and animal breeding techniques comprise selection, cross, heterosis and bio-engineered breeding:

In selection breeding a phenotype is mass selected according to its desired trait from a mixture of a larger population for further cultivation. Directed (positive) selection improves the degree of efficiency of a trait by elimination of one extreme, shifting the average of the character within the population. Stabilizing (negative) selection eliminates deviant individuals from the population, narrowing the range of genetic variability. Disruptive selection of specific extremes leads to greater variability and to polymorphism. Through line breeding by selection over successive generations a group of identical pure bred individuals is obtained and the chosen trait then multiplied.

In mono- and polyhybrid crossbreeding of genetically different organisms a fusion of alleles, surpassing incompatibility barriers, achieves in the daughter generation the combined, desired traits in one heterozygous genotype. Genetic hybrids are mixo-ploid combinations (mosaics) from different genera, leading to new species (chimeras). Through convergence breeding by recurrent selection and intercrossing the new trait is stabilized in uniformity and consistency.

In heterosis breeding also a crossing of strains takes place, not to obtain a constant genotype, but for the heterosis effect, where the heterozygous mix in the genome is superior in a desired trait, which may be lacking in the P generation (hybrid vigor). The hybrid seed can only be gained from its parent populations, as the heterozygous state loses its specific mix relation by further intercrossing.

Newer breeding techniques employ in combination of mutagenic, recombinant and hybridization DNA technologies in vitro manipulation of cell cultures in an artificial nutrient, a semisolid or suspension medium under controlled environmental conditions. They allow for example large scale breeding of life stable colonies (colony breeding); somatic hybridization between alien gametes, bypassing fertilization barriers, where whole, isolated, by dissolution of their walls stripped cells (protoplasts) of different species are fused with ones containing a nucleus or with their nuclei removed to form a hybrid or a cybrid (cytoplasmic hybrid); embryo splitting, breaking up of embryos in the 2 - 4 cell phase, cultivation and reimplantation into two surrogate mothers; cloning, asexual reproduction of an identical, recombinant DNA molecule by mitosis out of a single somatic or germ cell.

## Gene technology

Genetic engineering as discipline of molecular genetics is a part of bio-technology. It comprises the theoretical and applied aspects of isolation, analysis, manipulation and recombination of structural and regulative genes and their introduction, expression and multiplication in other organisms apart from naturally occurring processes.

Molecular bio-technology furnishes a significant contribution to basic research in genetics. It developed methods for analysis of nucleotides and –sequences, their structures, functions, reactions and products with bio-synthetic pathways and interactions, as well as technologies for their a) isolation and identification, b) gene mapping, c) manipulation, d) synthesis, e) ligation, f) transfer, g) transformation and multiplication, f) test and production devices. Transformation following production of a passenger DNA sequence, a vector system and ligation is the last step of cloning procedures, the asexual, identical reproduction of a DNA sequence. It opened the way for gene libraries (colony banks) and commercial production.

Applications of gene technology, the 'soft' technology, concentrate on the fields of medicine, pharmacology, food production, human genetic diagnosis and therapy. They expand into reproduction technologies, forensic genetics, pest control. Patents are granted on their products and methods.

a) Methods for in vitro isolation of DNA segments are cleavage by pattern recognizing restriction enzymes together with separation of DNA fragments e.g. by blotting or polyacrylamide gel electrophoresis, separations by molecular weight with nucleic acid and protein identification.

b) Gene mapping of DNA segments on a chromosome proceeds in orientation of known genes by direct, fragmental DNA sequence determination or indirectly, e.g. by radioactive marking, cleavage and identification with a gene specific DNA probe.

c) Manipulation of a DNA segment to cause a specific change in a nucleotide or –sequence is based on the processes of DNA mutation by means of physical, chemical or bio-chemical incision, of DNA recombination by bio-chemical introduction, elimination or distortion, of DNA hybridization by cell and nucleic fusion of gene-tically close and distant (transgenic) materials.

d) Synthesis for construction of a specific DNA segment is achieved by bio-chemical de-novo synthesis of short oligotid sequences, followed by joining of the oligo- with polyotides, catalyzed by ligase enzymes, or by single strand synthesis through polymerization of complementary base pairs from a DNA matrix towards a complete DNA duplex, added by polymerase enzymes.

e) Ligation, joining of a passenger DNA segment into the open gap of a carrier DNA segment (replicon, vector), facilitating transformation and often with regulative sequences, stable gene expression, is achieved by covalent bonding, added by ligase enzymes, which represents via indirect integration the definite step towards recombinant and transgenic DNA.

f) Physical transfer of the vector system as an independent unit of replication into living host cells and cell nuclei is effected e.g. by concentration increase in form of a precipitate or charged complex or by in vitro laser poring, a micro-injection, physically opening the cell wall.

g) Transformation (a transposition) of a recombined vector system into the host genome aims at covalent bonding into the chromosome strands by cleavage and joining of both ends, assisted by restriction and ligase enzymes. The passenger DNA from in vitro cultivation is being multiplied in the nuclei of cell lines by repetitive transformations and cell divisions, also over the stages of ontogenesis.

h) Devices for testing and automated production demand accurate, sensitive, reliable, fast, miniaturized measurement and process control. Bio-chemical process parameters are taken up by bio-sensors, which contain two elements, an aggregate, recognizing the biological information with a molecular, cell like or microorganism bio-mass and a transducer for output of an electronic signal.