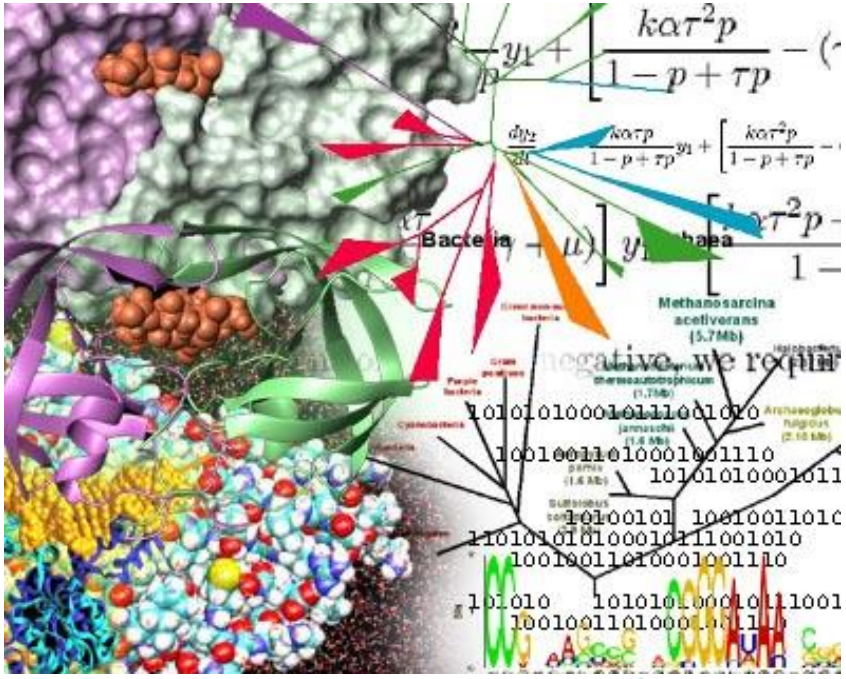
Chapter 6
Omics technology and Bioinfomatics



Outline:

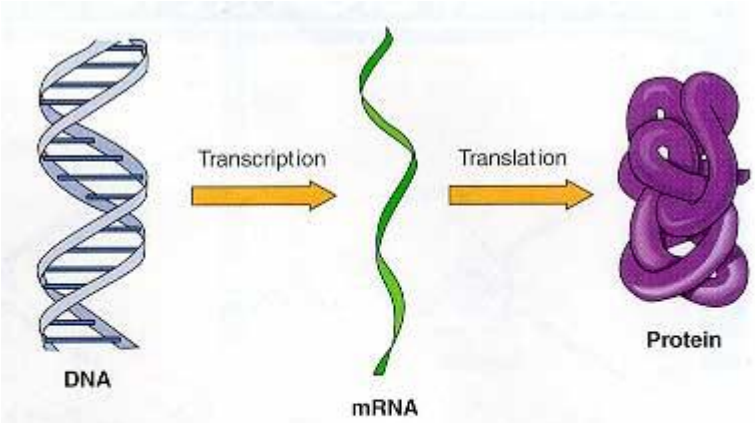
- 6.1 Introduction
- 6.2 Genomics
- 6.3 Transcriptomics
- 6.4 Proteomics
- 6.5 Metabolomics
- 6.6 Bioinformatics

Learning outcomes:

- Define genomics, transcriptomics, proteomics, metabolomics and bioinformatics.
- Describe the tools and applications of the omics technology.

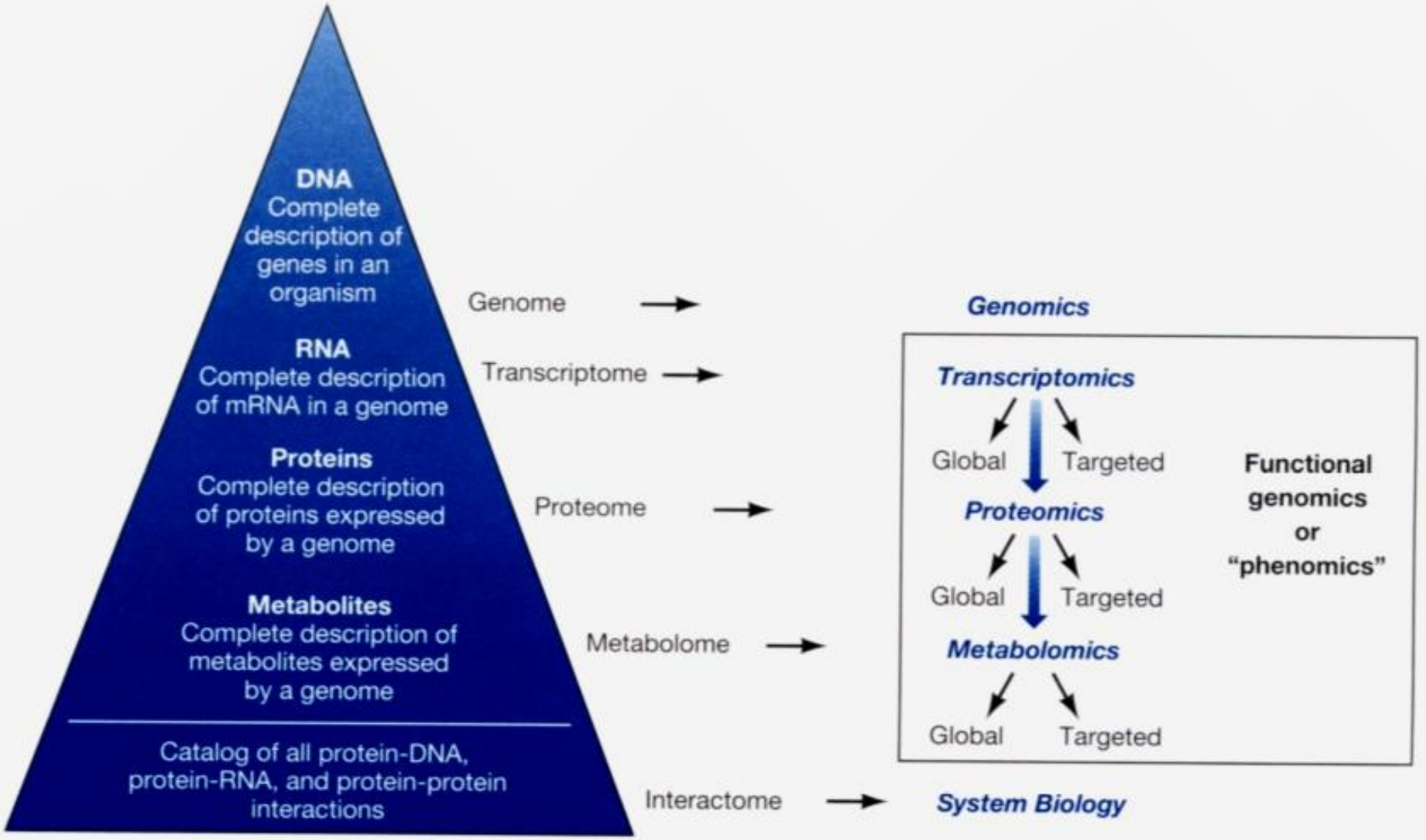
6.1 Introduction

- Omics technology:
- The universal detection of genes (genomics), mRNA(transcriptomics), proteins (proteomics) and metabolites (metabolomics)

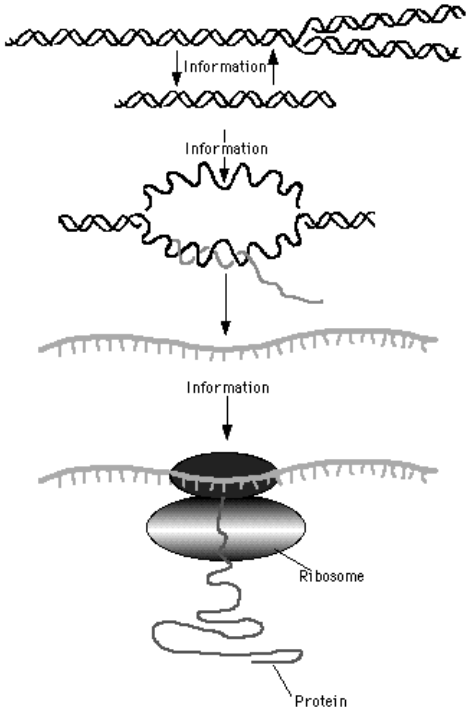


Central Dogma of Gene Expression.

Through the production of mRNA (transcription) and the synthesis of proteins (translation), the information contained in DNA is expressed.


DNA
Complete description of genes in an organism
RNA
Complete description of mRNA in a genome
Proteins
Complete description of proteins expressed by a genome
Metabolites
Complete description of metabolites expressed by a genome
Catalog of all protein-DNA, protein-RNA, and protein-protein interactions
Genome
Transcriptome
Proteome
Metabolome
Interactome
Genomics
Transcriptomics
Global
Targeted
Proteomics
Global
Targeted
Metabolomics
Global
Targeted
Functional genomics or "phenomics"
System Biology

The –omics World



DNA ------> Genomics (Genome)

Transcription

mRNA ------> Transcriptomics (Transcriptome)

Translation

Protein ------> Proteomics (Proteome)

provide structure & drive metabolism

substrate ⟶ product ------> Metabolomics (Metabolome)

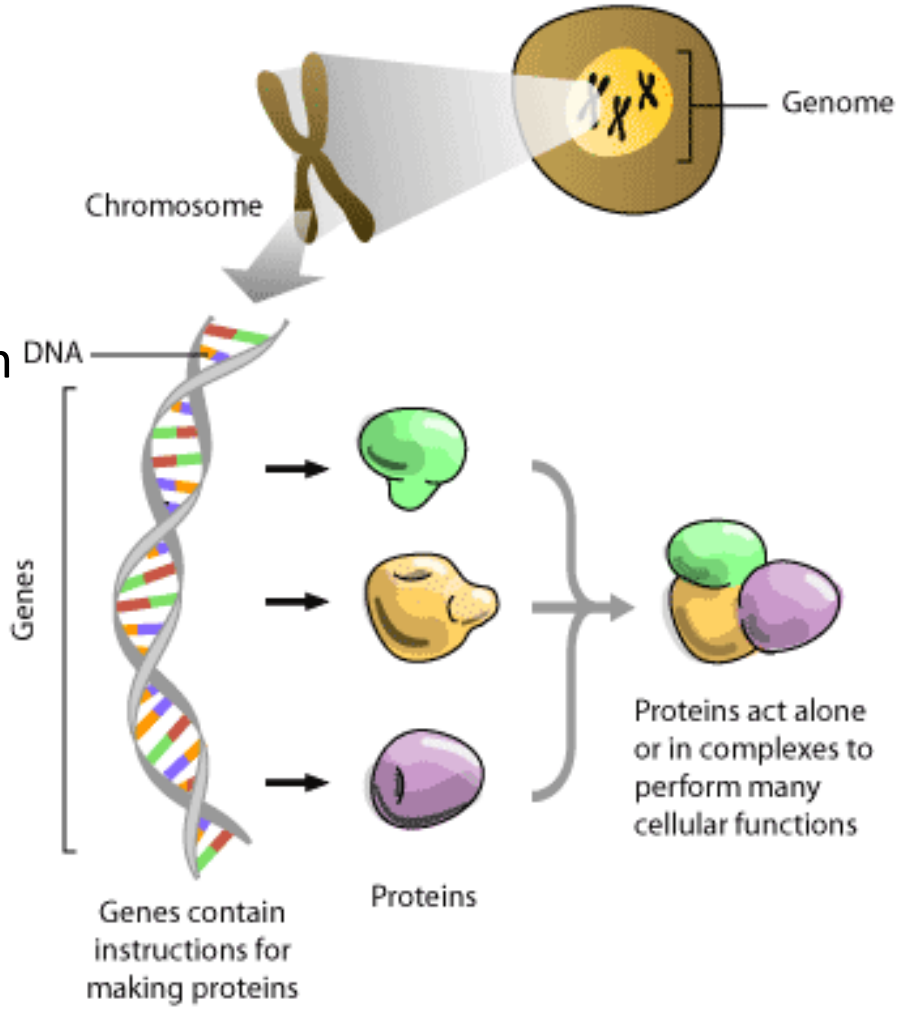
Genotype

Phenotype

morphology
physiology
behaviour
ecology

6.2 Genomics

- Genome
  - entire genetic complement of an organism
  - all the DNA contained in an organism or a cell, which includes both the chromosomes within the nucleus and the DNA in mitochondria.



http://www.dnai.org/c/index.html

6.2 Genomics

- Genomics:
  - the study of the genomes of organisms.
  - cloning, sequencing, and analyzing entire genomes
  - generation of information about living organisms by systematic approaches

6.2 Genomics

- Tools to study genomics
  - Shotgun sequencing or shotgun cloning
    - **The entire genome is cloned and sequenced**
    - **Produces thousands of fragments to be sequenced**
    - **Individual genes are sorted out later through bioinformatics**
      - **Computer programs are used to align the sequenced fragments based on overlapping sequence pieces**

6.2 Genomics

- Research field

1. Comparative genomic

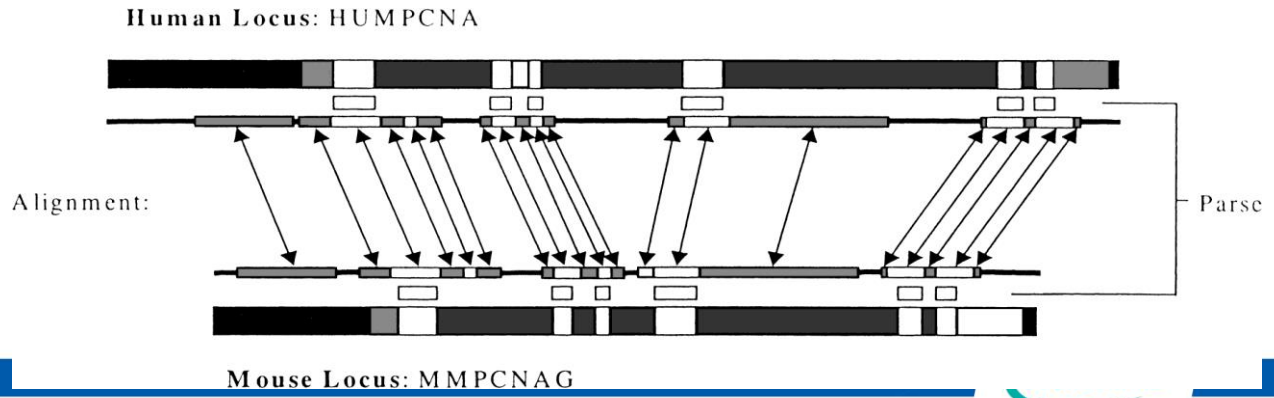
- The study of differences and similarities in genome structure and organisation of different organisms.
- The objective is to understand the process of evolution and to convert the DNA sequence into proteins of known function.

6.2 Genomics

- Research field

1. Comparative genomic

- Based on orthologous (similar genes found in different organims with the same function-signal from the same ancestor) and paralogous (similar genes found in the same organisms and encode a related proteins but different function) sequences



6.2 Genomics

- Research field

2. Gene prediction and counting

- identifying genes, their regulatory sequences and their functions.
- also involve identifying non-protein coding genes including those from ribosomal RNA, tRNA and small nucleus RNAs, mobile genetic elements and repetitive sequence families

6.2 Genomics

- Research field

3. Genome evolution

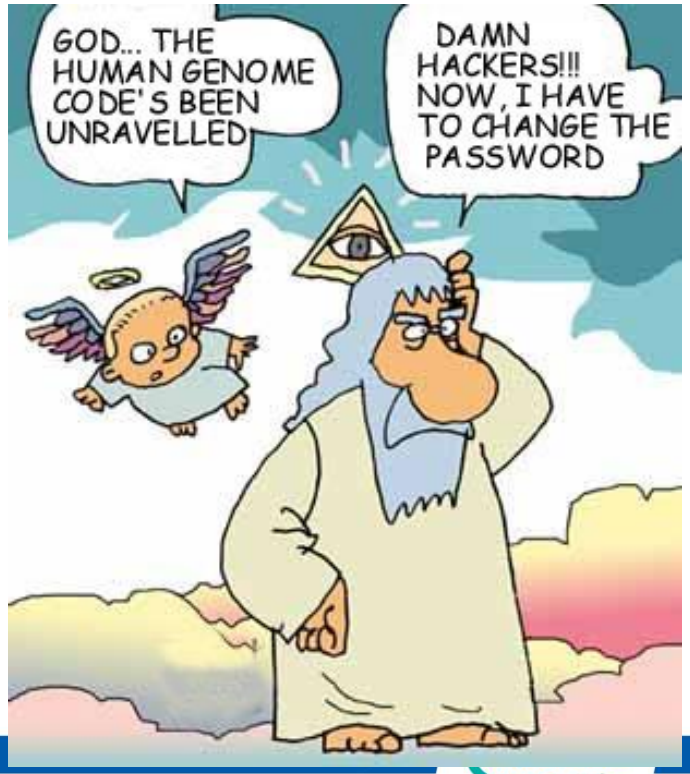
- Involve the study of genome evolution and the driving force in their evolution such as mutation, recombination, transportation, gene transfer as well as gene deletion and duplication.
- also involve the study of changes in prokaryot and eukaryot genome dynamically in terms of their size, shape and complexity

6.2 Genomics

- The Human Genome Project

Human Genome 10th Anniversary

http://www.genome.gov/25019885



6.2 Genomics

- The Human Genome Project
- started in the early 1990's, completed in April 2003
- involved the efforts of hundreds of scientists to generate high-quality reference sequence for the 3 billion base pairs of nucleotide sequence that make up the human genome.
- revealed that there are probably about 20,500 human genes.

THE HUMAN GENOME - THE BLUEPRINT OF LIFE

The Human Genome project sequenced DNA, the molecules that make up chromosomes in cells. The information derived from this project presented scientists with a valuable opportunity to not only uncover the secrets of DNA but also the manner in which genes are associated with disease. Scientists now are able to compare the genomes of people who have a certain condition with those who do not, in order to determine whether genetic variation plays a role in that condition. This information will help them to predict and possibly prevent disease in the future.

1. Cell

Each of the trillions of cells in the human body contains 46 chromosomes packed tightly into the region called the nucleus.

2. Chromosomes

Half of the chromosomes in the nucleus come from your mother, and half from your father. Each chromosome is a long, tightly coiled molecule called DNA, or deoxyribonucleic acid.

3. DNA

If unwound, the DNA from all the chromosomes in a single cell placed end to end would stretch more than six feet.

4. Genome

DNA is made up of chemical building blocks abbreviated A, C, T, and G. The entire length of a DNA strand consists of these four blocks in different combinations. Together, all the DNA in all the chromosomes – more than 3 billion letters – makes up the human genome. When scientists say they have "sequenced" the human genome, they mean that they have figured out the order of all those A's, C's, T's, and G's in sequence.

6. Misspellings in the Sequence

The way the genes are "spelled" makes all the difference - one letter out of place in a gene can cause disease. Now that we know the normal sequence of the human genome, researchers can compare the DNA sequence from people who have a disease or condition to those who don't. If there are differences in the spelling of certain genes between the two groups, it's possible that the condition may be caused by or related to that misspelling in that gene.

5. Genes: 30,000 DNA Segments

Much of the DNA in the genome is organized into units called genes. There may be as many as 30,000 genes in the genome; they are the instruction manual for making all the proteins in the body. These proteins are the physical "stuff" that makes up our hair, skin, heart, and blood, among other things. They also control chemical reactions, regulate blood sugar and heart rate, and control how food or medicine is metabolized in the body.

7. Genes and Disease

Scientists have identified about 6000 diseases, such as Huntington disease and cystic fibrosis, that are directly caused by misspellings or physical problems in single genes.
But the genetic contribution to many common conditions – such as diabetes and heart disease – is part of a larger puzzle that could include diet, lifestyle, environment, and even other genes. For many of these common conditions, genetic misspellings probably make only a small contribution to disease relative to other factors, or work in concert with them to cause illness.

Produced by the Genetics & Public Policy Center at Johns Hopkins University with funding from The Pew Charitable Trusts.
Illustration by: BAKEDMEDIA, INC.

http://www.dnapolicy.org/resources/human_genome_final.jpg

6.2 Genomics

Pre-genomic era VS Post-genomic era

Gene VS Genome

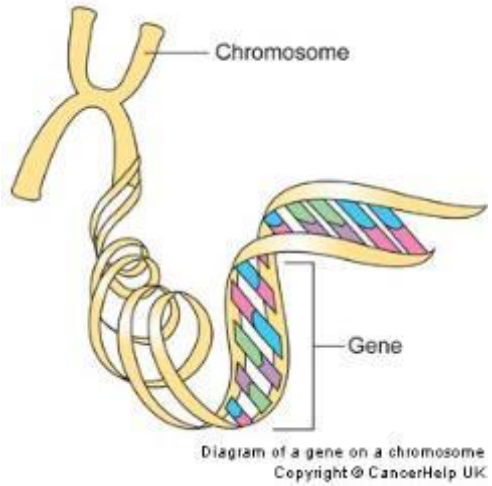


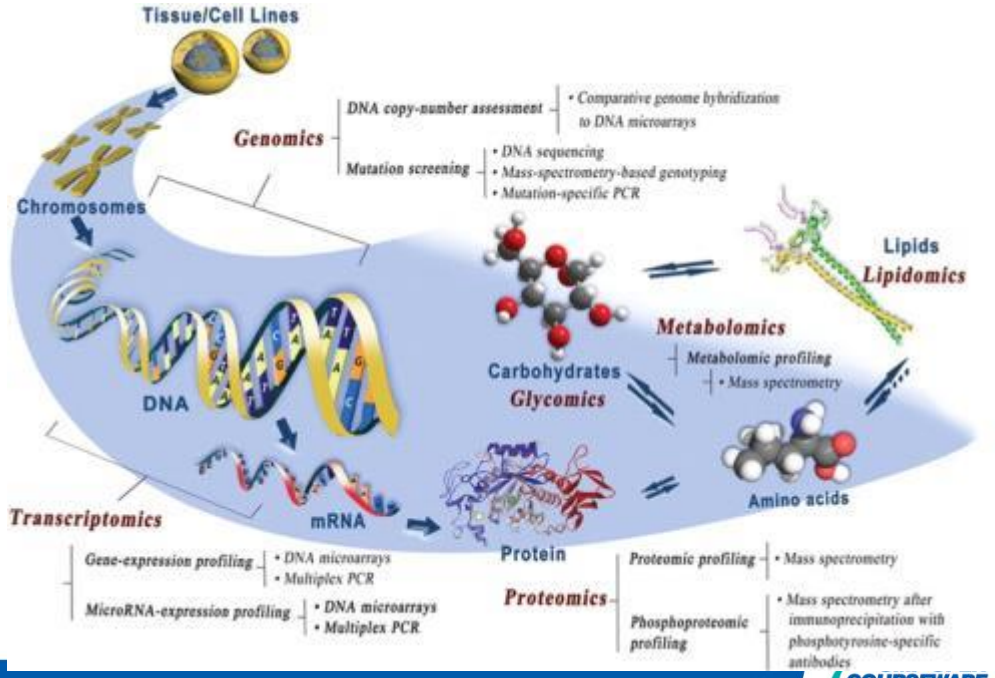
Diagram of a gene on a chromosome
Copyright © CancerHelp UK

6.3 Transcriptomics

- **Transcriptome**
- **percentage of the genetic code that is transcribed into RNA molecules (depends on development, environment, time of the day, tissue)**
- **collection of all gene transcripts present in a given cell/tissue at a given time ("snapshot")**

6.3 Transcriptomics

- **Transcriptomics**
- **Global analysis of gene expression**
- **genome-wide expression profiling**



6.3 Transcriptomics

- get an understanding of genes and pathways involved in biological processes ("guilt by association": genes with similar expression may be functionally related and under the same genetic control mechanism)
- help elucidating the function of unknown genes based on their spatial and temporal expression
- identifies marker genes for diagnosis of diseases

6.3 Transcriptomics

- Tools to study transcriptomics
- cDNA library
- PCR
- Reverse Transcription-PCR
- Microarray
- Real-time PCR
- Next generation sequencing

6.4 Proteomics

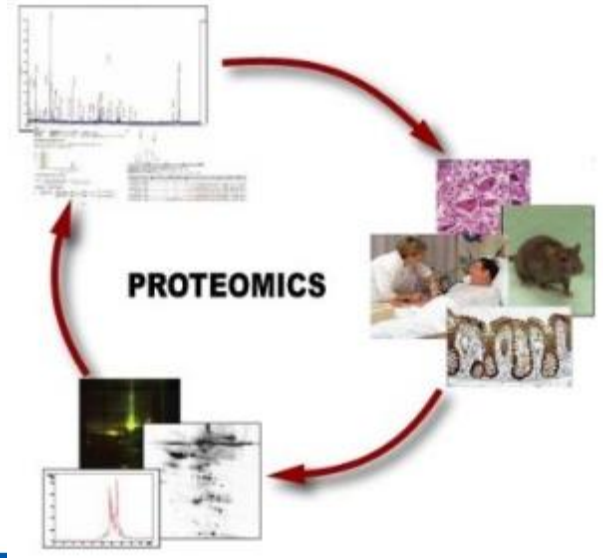
- Proteome
- refers to the entire protein complement of an organism

6.4 Proteomics

- Proteomics
- the global analysis of proteins
- large scale identification and characterization of proteins and all their properties
- unlike the genome, proteome varies with time

6.4 Proteomics

- Enables correlations to be drawn between the range of proteins produced by a cell or tissue and the initiation or progression of a disease state
- Permits the discovery of new protein markers for diagnostic purposes and of novel molecular targets for drug discovery.

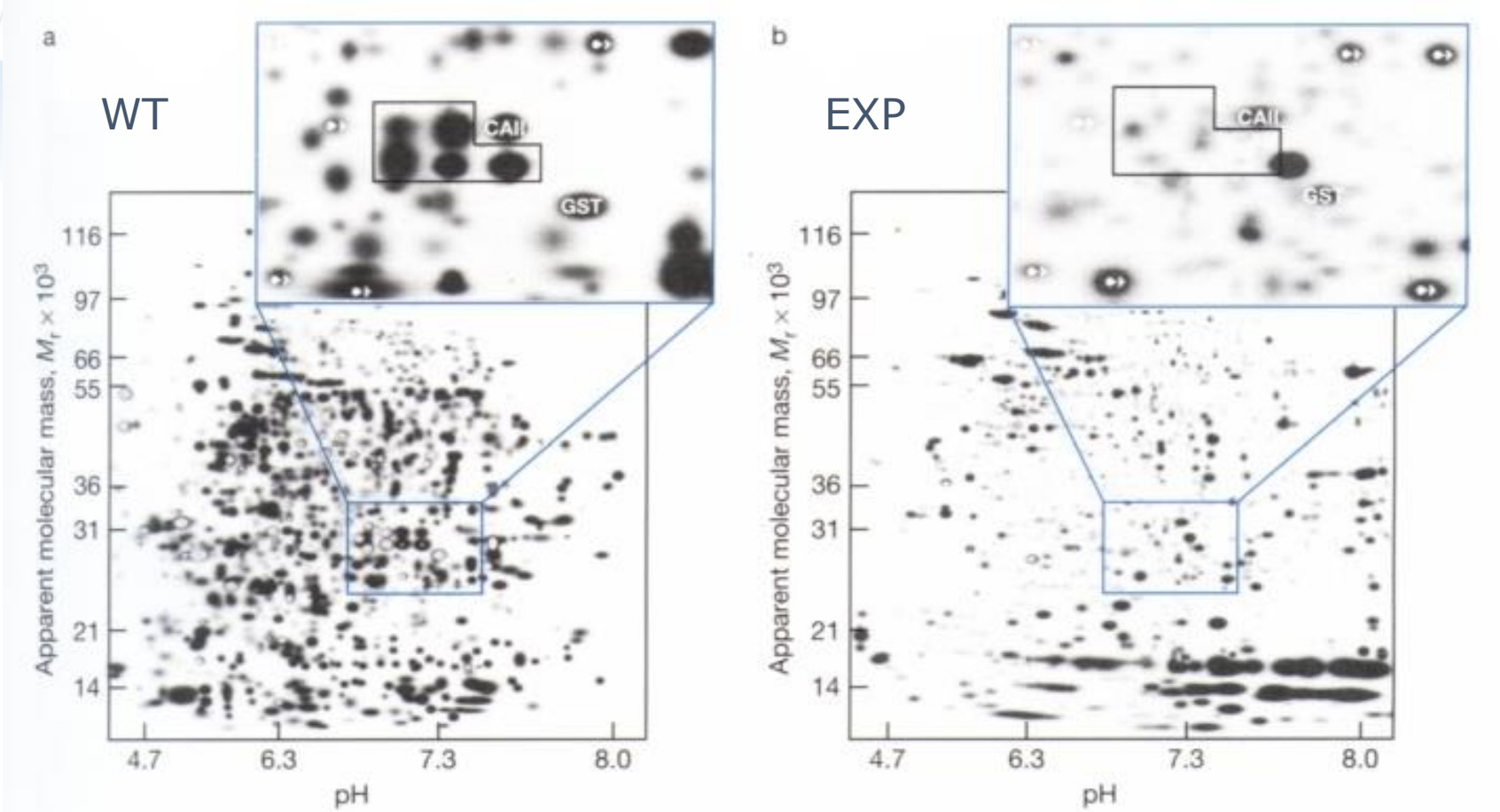


6.4 Proteomics

- Tools to study proteomics
- Separation of individual proteins by 2-D polyacrylamide gel electrophoresis (2D PAGE)
- Identification by mass spectrometry or N-terminal sequencing of individual proteins recovered from the gel
- Protein characterisation
- Storage, manipulation and comparison of the data using bioinformatics

6.3 Proteomics

Example 1, Wild type C57/Black/6J murine colonic crypt.



Global profiling proteomics

6.4 Proteomics

- Research field

1. Structural proteomics

- Prediction of 3-D structure
- Mapping out the 3-D structure and nature of protein complexes
- Main objective - to build a body of structural information that will help in predicting the probable structure and potential function for almost any protein from a knowledge of its coding sequence

6.4 Proteomics

- Research field

2. Expression protemics

- Refers to quantitative study of protein expression between samples.
- involves the study of pattern expression of a complete proteome or its part between samples.

6.4 Proteomics

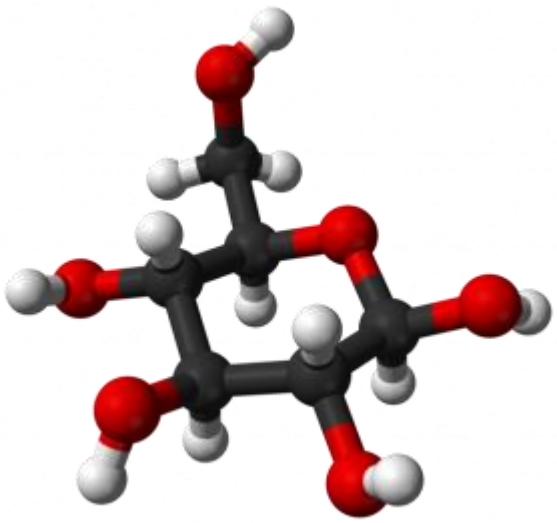
- Research field

2. Expression protemics

- Main objective is to identifying disease-specific proteins. For example, the over expression or under expression of proteins in cancerous and normal cells. Identification of disease markers.
- Application in drug action and effects in biological stimuli

6.5 Metabolomics

- Metabolites
  - Intermediates and products of metabolism
  - Examples include antibiotics, pigments, carbohydrates, fatty acids and amino acids

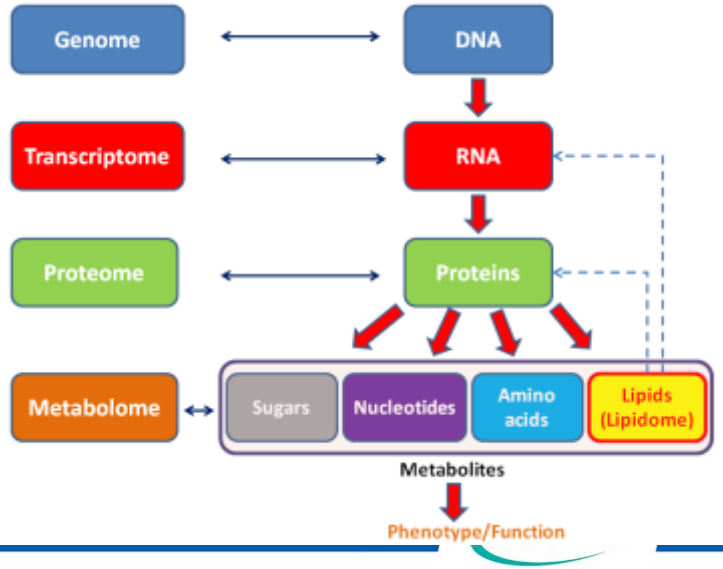


6.5 Metabolomics

- Metabolome
  - Refers to the complete set of small-molecule metabolites
  - results from the expression of the genome and proteome in response to the cellular environment
  - gives a direct picture of the cells activity and its environment.
  - represents a powerful portrait, reflecting health, aging and the effects of drugs and the environment

6.5 Metabolomics

- Metabolomics
- The study of metabolic profiling of an organisms
- Comprehensive and simultaneous systematic determination of metabolite levels in the metabolome and their changes over time as a consequence of stimuli



6.5 Metabolomics

- Tools to study metabolomic
- Gas chromatography (GC)
- High performance gas chromatography (HPLC)
- Mass spectrometry (MS)
- Nuclear magnetic resonance (NMR)
- spectroscopy

6.6 Bioinformatics

- Biology + Informatics (data processing)
- computational analysis and storage of biological data
- Aim: to discover new biological insights using computers and biology

6.6 Bioinformatics

- Datamining and screening of genes in a GENBANK
- Gene and protein modelling and simulation of how certain genes acts in certain biological pathway
- Gene prediction of functional protein using computer
- 3D structure of gene and protein prediction
- Comparative genome between species
- Evolutionary study

Discussion

- Distinguished genomics, transcriptomics, proteomics, metabolomics and bioinformatics.

Extra reading

- http://www.scq.ubc.ca/the-human-genome-project-the-impact-of-genome-sequencing-technology-on-human-health/
- http://www.genome.gov/10001772
- http://www.genome.gov/25019879
- http://www.genome.gov/10000464
- http://www.youtube.com/watch?v=9SzwiZMSBeQ

Universiti Malaysia PAHANG
Engineering • Technology • Creativity

THANK YOU