Perspectives Cheat Sheet

Hankertrix

November 9, 2023

Contents

1 Definitions

1.1 Genome

A genome refers to all the genetic information of an individual (or a species).

1.2 Genomics

Genomics is the study of the entire genome of an organism or genomes of several organisms.

1.3 Genetics

Genetics is the study of an individual gene or a specific group of genes.

1.4 "-omics" suffix

The "-omics" suffix refers to using a holistic approach, like using a whole genome or a whole cell, to study complex interactions within biological systems. This approach is also called holism. Dealing with the complexity of multiple factors in a meaningful way is one of the challenges in this approach.

All "-omics" studies ideally should have:

- Technology in handling large numbers of samples to acquire data. This is also known as high throughput sample management and data acquisition, which makes use of robotic and automated systems.

- Means of processing and analysing the large volume of data generated and acquired, which is made possible with computer programming, in the field known as bioinformatics.

1.5 "-etics" suffix

The "-etics" suffix reflects the more traditional reductionist approach, whereby the focus of the study is "reduced" to one or few components, like one or a few genes. It is as if we are mentally dissecting a gene from the rest of the pool of genes in an organism and finding out about its function and regulation.

Because of the interconnectivity of gene regulation, some studies cannot be done by examining one or two components through the "-etics" type of analysis. Instead, "-omics" level of study is needed to reveal complex patterns. However, without the data and conclusions based on "-etics" studies, no matter how much "-omics" data we may have, significant conclusions cannot be drawn. Thus, the two approaches complement each other.

1.6 The Human Genome Project (HGP)

The goals of the Human Genome Project is to:

- Determine the base pair sequence of the entire human genome

- Construct a map showing the sequences of genes on specific chromosomes

The Human Genome Project started in 1990, and it took 13 years to completely sequence the haploid set of the human genome.

From the sequence data of the human genome, we now know that:

- The human genome (the haploid set, which is 23 chromosomes) contains more than 3 billion base pairs.

- It is estimated that there are 20,000 to 25,000 protein-encoding genes, which take up only 1% of the genome.

- 95% of the average protein-coding gene in humans exists as intron sequences.

- Much of the human genome was therefore formerly described as "junk", and does not specify the order of amino acids in a polypeptide.

- It is now recognised that many "junk" regions have regulatory roles.

The human genome project has moved beyond the initial goal of simply determining the base sequences to the next phase that aims to bring clarity to the genes' functions and their complexity of interaction. This has spurred the concept of personalisation in healthcare. The data from the human genome project are also put through comparative genomic analysis.

1.7 Comparative genomic analysis

Comparative genomics compares the genome of one organism to the genomes of other organisms. Many model organisms have genetic mechanisms and cellular pathways in common with humans and comparative genomic data gives us clues on the functions of many human genes. Comparative genomics also offers a way to study evolutionary changes in the genome over time.

1.8 Functional genomics

Functional genomics aims to understand the role of the genome in cells or organisms through the study of expressed genes, like when mRNA transcripts are produced. However, this method may not correctly pick up the expression of genes that are regulated at the translational level.

1.9 DNA microarrays

DNA microarrays are a technique for identifying genes.

1. Microscope amounts of known DNA are fixed onto a small glass slide or silicon chip in known locations.

2. Mixtures of mRNAs collected from sample cells bind to DNA sequences on the chip through complementary base pairing and are detected through fluorescence.

3. The presence of mRNA will indicate the expression of a particular gene under the test condition.

1.10 Next-generation sequencing

Next-generation sequencing concepts are different from the traditional Sanger method of sequencing (which the human genome project was using) that allows several million to a billion bases to be sequenced within one run. Some of these methods can be used to determine the sequence of DNA and those of mRNA origin (i.e. genes which are expressed).

1.11 Primary protein structure

The primary protein structure is a sequence of a chain of amino acids.

1.12 Secondary protein structure

The secondary protein structure occurs when the sequence of amino acids is linked by hydrogen bonds.

1.13 Tertiary protein structure

The tertiary protein structure occurs when certain attractions are present between alpha helices, pleated sheets and other secondary structures.

1.14 Quaternary protein structure

A quaternary protein structure is a protein structure consisting of more than one amino acid chain.

1.15 Proteomics

Proteomics is the study of the structure, function and interaction of cellular proteins. Understanding protein function is essential to the development of better drugs. Often, **matrix-assisted desorption/ionisation (MADLI) - based mass spectrometry** is used to obtain the amino acid sequence (primary structure) of all proteins present.

Once the primary structure of a protein is known, it should be possible to predict its tertiary structure. Computer modelling of the tertiary structures of proteins is an important part of proteomics.

1.16 Proteome

The entire collection of a species' proteins is its proteome. The proteome is larger than the genome. Mechanisms such as **alternative pre-mRNA splicing** increase the number of possible proteins.

1.17 Glycomics

Glycomics is the study of a whole collection of carbohydrates from a cell type, tissue, or organ.

1.18 Lipidomics

Lipidomics is the study of a whole collection of lipids from a cell type, tissue or organ.

1.19 Metabolomics

Metabolomics is the study of a whole collection of metabolites from a cell type, tissue or organ.

1.20 Metagenomics

Metagenomics are DNA sequences of the whole community of various microorganisms.

1.21   Bioinformatics

Bioinformatics is the application of computer technologies, software and statistical techniques to the study of biological information.

- Genomics, proteomics and other "-omics" produce an enormous amount of raw data.

- These fields depend on computer analysis to find significant patterns in the data.

Scientists hope to find relationships between genetic profiles and genetic disorders. New computational tools will be needed to achieve these goals.

1.22   Personalised medicine (PM)

Personalised medicine is a recently proposed medical model involving the customisation of healthcare. Medical decisions, practices and medical products are to be **tailored to the individual patient**, and not simply based on the disease or symptoms he is to be treated for.

In order to select the most appropriate and effective therapies for a patient:

- Diagnostic testing of the patient is carried out extensively to collect data on his physiology and molecular or cell biology.

- The genetic profile of the patient is put through analysis.

The term personalised medicine was first coined in the context of classical genetics, like by considering whether a patient carries a particular allele of a gene to decide if a particular drug should be used to treat him.

It has since broadened to encompass all sorts of measures making use of systems biology (the "-omics") approach, thanks to the progress made with the Human Genome Project (HGP).

1.22.1 Rationale for the systems biology ("-omics") approach

- Most of the genetic variation between individuals has no effect on health. Looking at specific genetic variations may not reveal much.

- However, an individual's health is a combined product of genetic variation, personal behaviours and influences from the environment. The collective outcome of these influences is better revealed at the systems biology level.

- The combination of genetic variations of an individual affects the responsiveness to some drugs. For example, most cancer drugs are only effective in 25% of patients, and some anti-depression drugs are effective in only 60% of patients.

- These differences in the responsiveness to drugs are not revealed through a single gene analysis of the patient, but possibly through a genomic level of analysis.

1.22.2 Reliance on "-omics" technology

Modern personalised medicine relies on technology that confirms a patient's fundamental biology holistically at the DNA, RNA, protein and metabolite levels, which then leads to confirming a disease and diagnostic approach. For example:

- Genome sequencing can reveal mutations in DNA that influence diseases.

- Transcriptomic analysis can show which RNAs are involved with specific diseases.

- Variations in specific proteins that are drug targets may reveal whether this drug is effective with this patient.

- Analysis of metabolomics profiles may reveal the type, extent and severity of the disease.

1.22.3 Challenges

- We still do not know enough to make holistic patient profiling and treatment customisation. Even though the sequence of the whole genome may be obtained, many diseases and genes have not yet been associated.

- The analysis of acquired diagnostic data, like genetic data obtained from next-generation sequencing, requires computer-intensive data processing prior to its analysis. Current versions are limited in their capability.

- We try to use other "-omics" analyses to make up for the limitations of data processing, but it is still not sufficient until we have increased our knowledge base.

- In the future, more sophisticated and adequate tools will be required to accelerate the adoption of personalised medicine.

1.22.4 Regulatory issues

The current approaches to intellectual property rights, health insurance policies and personal data protection (data of patients obtained for personalised medicine profiling are considered **"personal data"**), will have to be redefined and restructured to accommodate the changes personalised medicine will bring to healthcare. The FDA has already started to take initiatives to integrate personalised medicine into its regulatory policies. They developed a report in October 2013 titled, **"Paving the Way for Personalised Medicine: FDA's Role in a New Era of Medical Product Development"**.

1.23 Personalised health plans

Personalised health plans can help patients mitigate risks, prevent disease, and treat it with precision when it occurs.

1.24 Genome-wide association study (GWAS)

A genome-wide association study will look at one disease, and then sequence the genome of many patients with that particular disease to look for shared mutations in the genome (comparative genomics). Mutations that are determined to be related to a disease by a genome-wide association study can then be used to diagnose that disease in future patients, by looking at their genome sequence to find that same mutation. The first genome-wide association study was conducted in 2005 and studied patients with age-related macular degeneration.

1.25 Pharmacogenomics

Pharmacogenomics is an approach that uses an individual's genome to provide a more informed and tailored drug prescription. Since how well individuals respond to a certain treatment depends on their genetic variation, knowing their genetic content can change the type of treatment they receive. This will help prevent adverse events like side effects, allow for appropriate dosages, and create maximum efficacy with drug prescriptions.