TOPIC 1: Bioinformatics & extracting information from genome sequence & application to identifying genes associated with diseases

Lecture 1 – Introduction

Genomics = analysing genomes & finding genes

Functional genomics = understanding genomics & function of genes

- Finding out what genes are expressed & when
- The functions of genes regarding the proteins/functional RNAs they encode

Terminology:

- Genome = DNA sequence or hereditary material
- Transcriptome = RNAs that are transcribed from the genome sequence
- Proteome = proteins that are found in a cell, tissue or organism
- Metobolome = metabolites present in a cell or tissue
    - Metabolites (chemicals) affect the activity & state of cell
    - Proteins are often enzymes which act upon these metabolites
- Phenome = the state of the cell, tissue or organism determined by the genome, transcriptome, proteome & metabolome of cell
    - Eg. Genes mutated/deleted in genome or proteins absent from proteome would affect the overall phenome of the cell

Around half of genes in most organisms have unknown function, therefore to address this issue, conduct:

- Sequence analysis
    - See what other genes are similar to our gene of interest
    - See what domains does our gene/protein share with other genes
- Expression analysis
    - To study when & where gene is expressed (time & level of expression) → gives information about function

Gene to function = reverse genomics → know the gene, but don't know the function

- Therefore, mutate gene of interest in an organism → study the phenotype/result of mutation & compare with wild type

Function to gene = forward genetics → know the function, don't know what protein/gene involved (mutated)

- See what protein/gene is missing/mutated in this mutant organism

Disease to gene → know the disease, determine gene missing/mutated/affected

- Then find the cure. This is especially important for humans

Functional genomics most frequently conducted in model organisms. Why?

- There are too many species to study them all in detail → use representatives based on evolutionary tree
- Knowledge of one species provides tools and knowledge to extend knowledge on different parts of biology
- Difficult to perform studies on humans → public more accepting towards experiments on non-humans
- Some organisms have special characteristics → eg giant axons of squid, large oocytes of frogs (these characteristics might be useful to technological approaches)
- Most organisms share common molecular mechanisms (their genome has similarities)

Examples of model organisms

Organism	Characteristics
E. coli (bacteria)	- Genome completely sequenced - Grows rapidly in defined media → cell division occurs quickly, can use it as a replacement for organisms that do the same thing but require long time; changing types of nutrients in media can change the biochemical pathway of the E coli - - Easy to genetically manipulate, knockout of every gene is available
Arabidopsis (plant)	- Small size, rapid life cycle - Small genome, easy to genetically manipulate using Agrobacterium - T-DNA (transboson, transposable element) disruption available for most genes
Yeast (eukaryote)	- Single cell eukaryote - Grows rapidly in simple defined media - Small genome, easy to genetically manipulate - Can reproduce sexually or asexually - (small limitation → simple organism so it doesn't have as many genes as high eukaryotes; lass regulation but processes are basically similar)
C. elegans (worm)	- Simple worm, can be male or female (haemaphrodite) - Origin and fate of all 959 cells present are known - First multicellular organism to have its genome sequenced - Has 19,000 proteins → similar to humans - RNAi knockdown tool available for every gene in its genome
Drosophila (insect)	- First genetic organism - Many mutations identified and are available - Short life cycle (9 days) → easy & cheap to breed & keep
Mouse (mammal)	- Model organism for high organisms (especially mammals) - Easy to breed & house - Many small mutant strains available - Genome contains 21 chromosomes, 25,000 genes → similar to humans
Homo sapiens (humans)	- Not convenient to breed & house, difficult to experiment - Genome sequence completed - Great varieties of diseases → knowledge of affected genes provide insight to disease (cure) and how body works
Others	Zebra fish – model early vertebrate, model for developmental biology Squid – possess giant axon, model for neurobiology Pufferfish – simple gene structures, small genome

*there are many existing techniques available to study existing model organisms. If use new model organisms need to develop new tools & methods to study, sequence their genome, manipulate them and unable to use other literatures as reference

Limitations/pitfalls of using model organisms

- The focus on model organisms might actually narrow research to most easily addressed questions rather than the most important ones (ie take the route of low resistance, but this may cause us to ignore other important questions)
- In some cases, animal diseases do not reflect human diseases very accurately → (eg effective on mice but not humans; waste money)

In other cases, some organisms are studied for their own intrinsic interest rather than as a model for others

- Eg functional genomics conducted on viruses (HIV), parasites (malaria), domesticated plants & animals (agriculture purposes)
- Study them to develop therapies to stop them causing diseases or to maximise return on product