Genomics and bioinnformatics

Jason Hodgson

6/11/2017

1 Human Genome project

- 1990-2003

- $3 billion

- Aim to sequence each base in the human genome at least once.

- Can now sequence entire human genome to 5-10x coverage much cheaper and all at once

2 Why do we want genoic data?

Population genetics: Allows us to uderstand relatedness, geographic structure, migrationm connectivity and demography.
Phylogenetics: Relationship between species.
Functional genetics: Relationships between genes and phenotypes.
Bahavioural genetics: Look at the relationship between genes and social structures. Conservation: Looking at genomic differentiation to help determine what species there are. See there are more species than we though. Smaller numbers of a species than thought as species are split up and so get given more conservation priority.
Single locus vs genomic data: Scale, important to get a full picture.
Single locus and statistical power

3 Whole genome sequencing

Shotgun genome sequencing: Take genome, chop into small bits, sequence fragments and then piece back together.
Most current techniqes produce short reads (15 to 100bp) some up to 10000

bp. Short reads generally have lower error rates. To counter errors sequence lots.

De Novo: sequencing a genome not done before. Re.sequencing: Sequencing a genome which we have previously sequenced. e.g human genome.

De novo is harder so require more data.

Interested in the polymorphisms, need to sample many times to increase likelihood to pick these up accurately. This becomes expensive.

4   Methods of sequencing

When choosing method need to consider many factors e.g read length, cost, speed etc.

RAD seq (Restriction site associated): Shrink genome into managable size, enables population genomics of non model organisms, SNP discovery method. Sufficient reads required to call heterozygotes, map poistions only....

Gene chip SNP typing: Very cheap

RNA-seq: Gather RNA from tissue and sequence it. Can compare tissues with each other.

DNA capture: Good for poor DNA sources.

5   How do you get DNA?

Depends on species, some easier than others.

Type of sample determines quality of DNA.

Blood, tissuem saliva, semen all have lots of DNA, DNA is intact, know the DNA is from the target species. Depending on species can be difficult to collect.

Sometimes need traping methods. This and other methods can potentially have an affect on the species.

Non invasive DNA sources: DNA gets left, no harm, sometimes easier to get, hair, feathers, faeces, urine, semen. Often low quality.

Ancient DNA: DNA does break down over time, rate depends on the environment its stored in.

The major difficulty in genomics is the amount of data: bioinformatics, programming, computer methods.