Learning Objectives

Understand:

• The distinction between **Quantitative Genetic** and **Molecular Genetic** research
• The structure of DNA
• How DNA functions as the hereditary material
  – How DNA is packaged
  – How DNA function is regulated, etc
• The different types of genetic diversity in human populations
• Common terms used
• The broad types of genotype-phenotype relationships
What are the origins of individual differences in human behaviour and complex traits?

Quantitative Genetic research

• How much do genetic (and environmental) factors contribute to individual differences in a measured trait or susceptibility to disease?

• Nature AND Nurture
  – and the correlation and interaction between the two
Why do we estimate heritability?

• Quantitative Genetic studies tell us:
  • The importance of genetic and environmental influence
  • Provides a best case scenario for genetic prediction
  • Allows us to estimate the success of gene identification efforts
Twins - a natural experiment

Genotype 100% vs 50%
Sex
Age
Parents
In utero environment?
Early life / home environment
Molecular Genetic Research

- The study of the structure and function of genes at the molecular level

- Builds on Quantitative Genetics research to tell us:
  - What are the specific genetic factors?
  - How big is their effect?
  - How do they influence human trait variation or liability to disease?
Why do we estimate heritability?

• **Quantitative Genetic studies tell us:**
  – The importance of genetic and environmental influence
  – Provides a best case scenario for genetic prediction
  – Allows an estimate the success of gene identification efforts

• They do **not** tell us:
  – Number of genes
  – Where they are located
  – Mechanism of action
  – Magnitude of effect
  – Types of genes
An atlas of genetic correlations across human diseases and traits
GWAS of Anorexia 2017

3,495 cases and 10,982 controls; Duncan et al 2017  Am J Psych
DNA Structure

- Chromosome
- Centromere
- Telomere
- Chromatin/Chromatid

Cell

- Nucleus
- Chromatin/Chromatid
- Telomere

Base Pairs

DNA (double helix)

Histones
DNA Structure
DNA to mRNA to Protein

Transcription

Translation
Transcription (DNA to RNA)

- The process of producing RNA from the DNA
- For each gene this is only from one strand of the DNA
- The pre-mRNA is “spliced” to form mature mRNA (removal of introns)
Several codons code for the same amino acid
“Third Base Wobble”
This means that some DNA variants in coding regions do not lead to an amino acid change
Amino acids can be grouped into chemically or structurally similar groups. Some DNA variation that leads to an amino acid change has little or no impact on protein structure/function.
From DNA to protein

http://www.yourgenome.org/video/from-dna-to-protein
https://www.youtube.com/watch?v=gG7uCskUOrA
Slightly more “out there”: https://www.youtube.com/watch?v=-ygpqVr7_xs&t=3s
Why are gene based studies important?
Genes to Behaviour

• **Fundamental insights into biology of traits and disease**

• **Diagnosis**
  - Gene-based (rather than symptom-based) diagnoses.
  - Disease heterogeneity.
    - Same aetiology, different phenotype
  - Genetic heterogeneity.
    - Different aetiology, same phenotype

• **Prevention (more likely, prediction)**
  - DNA as early warning system for behavioural and environmental (not genetic engineering)

• **Treatment**
  - Allow new directions for therapy
  - Tailored to individuals
    (e.g., pharmacogenetics)
  - Relevant environments
    (e.g., Tailoring Of Lifestyle choices)
Why are gene based studies important?

Genes to Behaviour

Genes

<table>
<thead>
<tr>
<th>Proteins</th>
</tr>
</thead>
</table>

Brain systems

Cells

Brain regions

Cognition/ emotion

Slide Jon Roiser UCL
Discovery of the Structure of DNA

Watson & Crick: 1953

Rosalind Franklin: X-ray images of DNA

Allowed scientists to ask how DNA functions as the hereditary material
DNA Structure

**DNA: Deoxyribonucleic Acid**  
**Structure:** double helix

DNA molecule is made up of *sugar* residues *phosphate* groups and *bases*

Attached to carbon atom 1’ of each sugar is a nitrogenous base:  
Adenine (A)  
Cytosine (C)  
Guanine (G)  
Thymine (T)

**Watson-Crick base-pairing rules:**  
A:T, C:G (hydrogen bonds)
The Human Genome Project

Size of the Human Genome
The Number of Genes
Variation between individuals
The Human Genome

- Published in 2003
- $3.1 \times 10^9$ bases (3 billion)
- 22,000 protein coding genes (~2%)
- $3$ billion USD, 15 years

ASCL1 (Achaete-scute complex-like 1) gene

https://www.youtube.com/watch?v=yqCJLRsBsFw
Announcement of a working draft June 2000
Navigating the Genome

• Genome Browsers
  – Gene Annotations
  – Phenotype Associations
  – Gene Expression
  – Etc...

– http://genome.ucsc.edu/
– http://www.ensembl.org/index.html
http://genome.ucsc.edu/
We vary genetically

- No two people are genetically identical, except for MZ twins (with caveats)
- Humans are ~99.9% genetically identical
- We mostly know where the ~0.1% (3 million bases) of DNA variation resides
- Variation makes us unique

Classification of Genetic Variation

- ‘Mutation’ Change in DNA sequence from wild type. Often used as short-hand for pathogenic variant.

- ‘Polymorphism, or DNA variant’ A variation of the DNA sequence (at a specific locus) that is present in more than 1% of the population (minor allele frequency; MAF >1%).

- ‘Non-polymorphic’ The DNA base does not vary in the population.
Glossary of Terms

• **Genes:**
  – Units of inheritance that are passed down to the next generation

• **Locus:**
  – Specific position in the genome

• **Alleles:**
  – The genes found at the same locus on different homologous chromosomes are alleles.
  – *Allele frequency* refers to the number of alleles in the population

• **Genotype:**
  – Combination of the two alleles at a specific locus.
  – Alleles can be the same (*homozygous*) or different (*heterozygous*)
Many types of genetic variation

• **Small scale:**
  – Single base change (SNPs)

• **Medium scale:**
  – Copy number variation (CNVs), indels

• **Large scale:**
  – Chromosomal abnormalities
    • Change in the number of chromosomes (whole)
    • Change in the arrangement chromosomes (part)

• And everything in between...
SNPs: Alleles and Genotypes

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>CC</th>
<th>CT</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>500</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>Cases</td>
<td>400</td>
<td>500</td>
<td>200</td>
</tr>
</tbody>
</table>
## Genotypes to Alleles and MAFs

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>CC</th>
<th>CT</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>500</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>Cases</td>
<td>400</td>
<td>500</td>
<td>200</td>
</tr>
</tbody>
</table>

TT individuals have two T alleles = 200 T alleles in the controls. CT individuals have one T allele = 400 T alleles in the controls. There are 600 T alleles in the controls.

The total number of allele in the controls is 2000 (each person has two alleles. 600/2000 = 30%.

<table>
<thead>
<tr>
<th>Alleles</th>
<th>C</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>1400</td>
<td>600</td>
</tr>
<tr>
<td>Cases</td>
<td>1300</td>
<td>900</td>
</tr>
</tbody>
</table>

### Association Testing
- Chi-square 54.303  p=1.72x10^{-13}
- Odds ratio 1.615 (95% CI 1.421-1.836)
Biological insights from 108 schizophrenia-associated genetic loci

Schizophrenia Working Group of the Psychiatric Genomics Consortium

36,989 cases
113,075 controls
83 new loci
DRD2
several genes involved in glutamatergic neurotransmission
Schizophrenia GWAS 2009
2,601 cases, 3,345 controls
0 genome wide significant sites
Schizophrenia GWAS 2011

9,394 cases, 12,462 controls

5 genome wide significant sites
Schizophrenia GWAS 2012
25,785 cases, 28,441 controls
62 genome wide significant sites
Schizophrenia GWAS 2013
35,476 cases, 46,839 controls
97 genome wide significant sites
Biological insights from 108 schizophrenia-associated genetic loci

Schizophrenia Working Group of the Psychiatric Genomics Consortium*

36,989 cases
113,075 controls
83 new loci
DRD2
several genes involved in glutamatergic neurotransmission
Alterations in Chromosomal Number

• **Polyplody:**
  – Individuals have three copies of the autosomes, plus an extra sex chromosome (karyotype: 69, XXX, 69 XXY, 69 XYY)

• **Aneuploidy:**
  – One or more chromosomes are lacking or present in excess
    Examples: Trisomy 21 (Down’s Syndrome), Klinefelter’s syndrome (47, XXY), and Turner’s syndrome (45, X)
Chromosomal Abnormalities

Klinefelter’s Syndrome

Turner’s Syndrome
Copy Number Variation

Chromosomal micro-duplications and deletions

Deletion  Normal  Duplication

1 copy  2 copies  3 copies
CNV in schizophrenia

Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia

Rare chromosomal deletions and duplications increase risk of schizophrenia

Support for the involvement of large copy number variants in the pathogenesis of schizophrenia

The International Schizophrenia Consortium*
CNV in autism
## CNV in schizophrenia

<table>
<thead>
<tr>
<th>Locus</th>
<th>P-value in previous studies</th>
<th>CNV frequency, % (n/N)</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.1 del</td>
<td>1.3 \times 10^{-9}</td>
<td>0.17 (33/19056)</td>
<td>8.35 (4.65–14.99)</td>
<td>4.1 \times 10^{-13}</td>
</tr>
<tr>
<td>1q21.1 dup</td>
<td>2.0 \times 10^{-4}</td>
<td>0.13 (21/16247)</td>
<td>3.45 (1.92–6.20)</td>
<td>9.9 \times 10^{-5}</td>
</tr>
<tr>
<td>NRXN del</td>
<td>7.9 \times 10^{-9}</td>
<td>0.18 (33/18762)</td>
<td>9.01 (4.44–18.29)</td>
<td>1.3 \times 10^{-11}</td>
</tr>
<tr>
<td>3q29 del</td>
<td>2.3 \times 10^{-8}</td>
<td>0.082 (14/17005)</td>
<td>57.65 (7.58–438.44)</td>
<td>1.5 \times 10^{-9}</td>
</tr>
<tr>
<td>WBS dup</td>
<td>5.5 \times 10^{-5}</td>
<td>0.066 (14/21269)</td>
<td>11.35 (2.58–49.93)</td>
<td>6.9 \times 10^{-5}</td>
</tr>
<tr>
<td>VPR2 dup</td>
<td>0.006</td>
<td>0.11 (15/14218)</td>
<td>1.54 (0.77–3.09)</td>
<td>0.27</td>
</tr>
<tr>
<td>15q11.2 del</td>
<td>2.2 \times 10^{-7}</td>
<td>0.59 (116/19547)</td>
<td>2.15 (1.71–2.68)</td>
<td>2.5 \times 10^{-10}</td>
</tr>
<tr>
<td>AS/PWS dup</td>
<td>0.014</td>
<td>0.083 (12/14464)</td>
<td>13.20 (3.72–46.77)</td>
<td>5.6 \times 10^{-6}</td>
</tr>
<tr>
<td>15q13.3 del</td>
<td>2.1 \times 10^{-11}</td>
<td>0.14 (26/18571)</td>
<td>7.52 (3.98–14.19)</td>
<td>4.0 \times 10^{-10}</td>
</tr>
<tr>
<td>16p11.2 dup</td>
<td>0.03</td>
<td>0.31 (37/12029)</td>
<td>2.30 (1.57–3.36)</td>
<td>5.7 \times 10^{-5}</td>
</tr>
<tr>
<td>16p11.2 distal del</td>
<td>0.0014</td>
<td>0.063 (13/20732)</td>
<td>3.39 (1.21–9.52)</td>
<td>0.017</td>
</tr>
<tr>
<td>16p11.2 dup</td>
<td>3.2 \times 10^{-14}</td>
<td>0.35 (58/16772)</td>
<td>11.52 (6.86–19.34)</td>
<td>2.9 \times 10^{-24}</td>
</tr>
<tr>
<td>17p12 del</td>
<td>0.0004</td>
<td>0.094 (12/12773)</td>
<td>3.62 (1.73–7.57)</td>
<td>0.0012</td>
</tr>
<tr>
<td>17q12 del</td>
<td>0.004</td>
<td>0.036 (5/14024)</td>
<td>6.64 (1.78–24.72)</td>
<td>0.0072</td>
</tr>
<tr>
<td>22q11.2 del</td>
<td>1.0 \times 10^{-30}</td>
<td>0.29 (56/19084)</td>
<td>NA (28.27–∞)</td>
<td>4.4 \times 10^{-40}</td>
</tr>
</tbody>
</table>

del, deletion; dup, duplications; NA, not applicable; WBS, Williams–Beuren syndrome; AS/PWS, Angelman/Prader–Willi syndrome.

a. For a more detailed version of this table that includes the CNV frequency, % (n/N) from previous studies see online Table DS6. P-values are based on Fisher exact test, 2-tailed.
16p11.2 CNV region

Effect of the chromosome 16p11.2 duplication on BMI and head circumference.

16p11.2 Duplications and Deletions produce mirror phenotypes on obesity and many measures of brain imaging

Similar phenotypes are seen in a zebrafish model using the KCTD13 gene

doi:10.1038/nature10406

Many phenotype associated loci effect non-coding genomic regions

• There are important non-coding regions that regulate gene expression (e.g., promoters, UTRs), and regions of the genome that are transcribed but not translated

• May alter epigenetic processes and the structure of chromatin?
Learning Objectives

Understand:

• The distinction between **Quantitative Genetic** and **Molecular Genetic** research
• The structure of DNA
• How DNA functions as the hereditary material
  – How DNA is packaged
  – How DNA function is regulated, etc
• The different types of genetic diversity in human populations
• Common terms used
• The broad types of genotype-phenotype relationships
Online Resources

- Help Me Understand Genetics – ebook (free) primer on genetics

- Your Genome: Resources on DNA, Genomes and Proteins
  - http://www.yourgenome.org/

- DNA to Protein

- DNA Replication
  - http://www.yourgenome.org/video/dna-replication

- DNA Sequencing
  - http://www.yourgenome.org/video/dna-sequencing

- Also: Genetic Science Learning Centre
  - http://learn.genetics.utah.edu/
Learning Questions

• Why might a genetic variant in a coding region have no impact on a protein?
• How would you determine whether a gene is expressed in the brain?
• Describe an experiment that would allow you to determine whether common genetic variants were associated with a trait of interest.