arcOGEN, a large-scale GWAS for osteoarthritis

John Loughlin
Newcastle University
OA in the UK

- 5 million adults
- Majority aged > 60 years
- More females than males (3:1)
- Over 2 million GP appointments & 50,000 joint replacements annually
How do we know that genes influence the risk of someone getting OA?

- Epidemiological studies
Twin studies
1 in 89 deliveries

1/3 Monozygotic
2/3 Dizygotic
<table>
<thead>
<tr>
<th>Condition</th>
<th>MZ</th>
<th>DZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cystic Fibrosis</td>
<td>100</td>
<td>25</td>
</tr>
<tr>
<td>Die on a Tuesday</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>55</td>
<td>20</td>
</tr>
</tbody>
</table>
OA is multifactorial & polygenic

Environment

Gene 1

Gene 2

Gene 3

Gene 4

Etc, etc
Distribution of liability

- Low liability
- Average liability
- High liability

Threshold

- Unaffected
- Affected
Two broad approaches to identifying susceptibility loci

Genome

Target candidates

Genome-wide scan
arcOGEN

_arthritis research campaign_
_Osteoarthritis Genetics_

- Largest single grant ever awarded by the arc
- £2.2 million
What is the aim of arcOGEN?

- To identify the DNA changes present in, or near, our genes that increase the risk of us developing OA
How will this help people?

- Will lead to an understanding of the molecular basis of the disease
- Will suggest new therapies
- Will allow the development of DNA-based diagnosis and prognosis
The human genome has a large number of DNA sequence differences, *polymorphisms*

The most common are single nucleotide polymorphisms, *SNPs*

<table>
<thead>
<tr>
<th>T-allele</th>
<th>TCGAGAGGGC<em>TAGGCTAGGA</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>C-allele</td>
<td>TCGAGAGGGC<em>CAGGCTAGGA</em></td>
</tr>
</tbody>
</table>
arcOGEN will study SNPs across the whole human genome
In cases and in controls

Cases

Controls

SNP frequency 22%

SNP frequency 18%
We are therefore performing a genome-wide association scan.
The arcOGEN protocol

- 8000 hip or knee OA cases
  - Joint replacement surgery
  - Males and females
  - 3200 already collected (extant cases, stage 1)
  - 4800 being collected (new cases, stage 2)

- 8000 controls

- Illumina 610-Quad Array
610-Quad

- 620,000 polymorphic markers
  - Covers approximately 90% of genome

- 138 mitochondrial SNPs

- Copy number variation (CNVs)

- Common polymorphisms
  - mean minor allele frequencies of 0.23
Why the 610-Quad?

• February 2007
  – Two options: Illumina and Affymetrix
  – Chose Illumina HumanHap 300 array
    • 317,000 SNPs

• October 2007
  – Funding approved
  – We could now afford the 610-Quad!
There are 11 participating sites in arcOGEN.
9 sites are collecting new cases

- Oxford
- London
- Southampton
- Nottingham
- Cambridge
- Manchester
- Edinburgh
- Newcastle
- Northumberland
- Sheffield
- Oxford
- Cambridge
- London
- Worcester
- Newcastle
- Northumberland
- Sheffield
- Oxford
- Cambridge
- London
- Worcester
DNA extraction, genotyping & analysis

- Edinburgh
- Manchester
- Worcester
- Northumberland
- Newcastle
- Sheffield
- Nottingham
- Cambridge
- Oxford
- London
- Southampton
Current status

• Stage 1 genotyping complete
  – 515,000 SNPs in 3,177 cases and 4,894 controls
  – *in silico* replication on 102 independent SNPs
  – *de novo* genotyping underway on 36 SNPs in 16,700 cases & controls

• 3,851 of the 4,800 new cases so far collected
What next

• Complete the scan, replicate hits and then search for the functional variants

• Sounds so easy!
Three scenarios

Ideal

OK

Nightmare!

Associated SNP
New concern

The case of the missing heritability
Potential remedies

- **Rare variants of high penetrance**
  - Whole genome & candidate gene sequencing in large cohorts

- **Common variants of even lower penetrance than those already studied**
  - Much larger case-control cohorts

- **Copy number variation (CNVs)**
  - Both large and small ones
Concluding remarks

• A GWAS is just the beginning of a long, long journey......
Acknowledgements

arcOGEN consortium

Nigel Arden  Stuart Ralston
Fraser Birrell  Ashok Rai
Andrew Carr  Tim Spector
Kay Chapman  Ana Valdes
Panos Deloukas  Gillian Wallis
Michael Doherty  Ele Zeggini
Bill Ollier