So just when will genomics revolutionise medicine?

Stephen Robertson
Department of Paediatrics and Child Health
University of Otago
NEW ZEALAND
The promise of genomics.....

- “all forms of disease, with the possible exception of some forms of trauma, have some genetic basis”

- “the sequencing of the human genome will usher in a revolution in medical practice – the age of personalised genomic medicine is upon us....”

Francis Collins, Director of The NIH
Genetic contributions to paediatrics

- Single gene disorders
- Chromosomal disorders
- Disorders with complex genetic architecture
- Drug response and safe prescribing
Established improvements to clinical practice

• Diagnosis of
  – Single gene disorders
  – High resolution definition of chromosomal disorders
  – Definition of susceptibility alleles
  – Tumour typing and stratification

Understanding the pathogenesis of many diseases
  – Obesity
  – Type 2 diabetes
  – Inflammatory bowel disease
  – …….and many non-paediatric examples – macular degeneration, cancers etc…….
The axes of variation and the technologies available

- Iterative (gene by gene) DNA sequencing
- Comparative genomic hybridisation (microarrays)
- Highly parallel DNA sequencing
  - Exomes
  - Genomes
Comparative genomic hybridisation

- Uses genomic DNA
- Two methods – Oligo arrays/SNP arrays
  - differential labeling of test and normal DNA
  - hybridisation of 1:1 mixture to spotted DNA array (matrix-CGH)
  - detect relative difference in DNA content of test versus normal by difference in red:green fluorescence ratio

Green = deletion in sample at that locus
Red = duplication in sample at that locus
Yellow = balanced at that locus

Current arrays have between 64 K – 5 million different individually assayed data points spread over the genome
Because it is digital it detects...

- Duplications
- Deletions
- Aneuploidies

…and can refine breakpoints

It doesn't detect:

- Balanced chromosome configurations
- Balanced Insertions
- Inversions
Chromosomal deletion on Chr 1 detected on aCGH: improved resolution

Diagnosis: 1p36 deletion (5 Mb in extent)
Copy number variation in the human genome

- Healthy humans carry significantly sized CNVs

- Listed and catalogued on a widely used “Database of Genomic Variants” – DGV

- 2-3% of the entire genome is subject to commonly found “benign” CNV

- Sorting these from CNVs that contribute causally to a phenotype is the principal challenge of applying aCGH clinically

![Graph showing a 2 Mb benign duplication on chromosome 16]
Where might we come unstuck
Anticipating our mistakes (or just learning from the past?)

CGH defines a causal factor in about 15% of children with mod-severe ID

Cooper et al 2001
Case 1 – Knowledge of gene content adds confidence to diagnosis

- Fetus in utero diagnosed with a pat (14;21) robertsonian translocation
- Severe Dandy Walker malformation on fetal ultrasound
- Post delivery – bilateral choanal stenosis, bilateral chorioretinal colobomata, heart defect, tracheomalacia
- CGH: \textit{de novo} 5.8 Mb deletion on chr8 including \textit{CHD7} – the major locus for CHARGE syndrome
- Lessons:
  - Fortuitous that a karyotype was performed rob(14;21)
  - Unequivocal consonance between deletion and clinical phenotype on basis of gene content
  - A firm diagnosis of CHARGE could have been made in utero
Case 2: a contributory, but not sole determinant of a phenotype?

- 18/12 female
- Developmental delay, dysmorphism, heart defect
- Father – learning delay, dysmorphism, otherwise healthy
- Clinical diagnosis 22q11 del
- Result: paternally inherited del of 1.5 Mb on 6q24
- No CNVs seen in this region before
- Dozens of genes deleted
- No grandparental samples available
- (A maternally inherited 32 Kb del was also detected on 3p26; not recorded in databases – no genes deleted)
Arrays in Autism Spectrum Disorders

- Substantial genetic determinants underlying ASDs
- Heritability (variance accorded to genetic factors) estimates dropping
- Sib recurrence risks being revised upwards in recent years
- The re-emergence of shared environments as important
- Surveys of genetic variation in ASDs using aCGH or newer exome sequencing technology (sequencing every gene in a subject’s genome) using a case control design have shown in individuals with ASDs:
  1. A 2-3 fold excess of CNV’s in individuals with ASDs
  2. A 2-4 fold excess of deleterious de novo mutations
  3. A rising risk with advanced paternal age (but epidemiology points to elevated maternal age as also being a risk factor)
  4. Families of genes affected; recurrent CNVs observed
ASD’s – the enthusiasm for genetic testing

- Caveat: ASD phenotypes in the presence of mod-severe ID are genetically more tractable than ASD alone
- Examine the motivation for testing – few CNVs will be able to be used in a diagnostic/pre-natal setting
- Risk variants exist (≠ they are worth measuring!)
ASDs and aCGH: Some cautionary tales

The conclusion: all three identified factors could plausibly contribute to her phenotype

Is that helpful?
Case 2: Deletion of a headline autism gene: CNTNAP2

- Proband: standard “narrow sense autism”
- aCGH – 1.2 Mb deletion of CNTNAP2
- Repeatedly reported deleted in autism
- Knock out in mouse
- Father carrier (IT specialist)
- Father in new relationship
  - what are my recurrence risks?
Will there be a hard answer for families with autism?

Devlin and Scherer Curr Op Genet Dev 2012
Rules of thumb for interpreting array results

- **Deletions** have a generally more severe effect than **duplications** (but duplications outnumber deletions as polymorphisms)
- **Size** matters…but it is not everything. Deletions > 100kb are often significant; duplications are more variable – treat anything > 500kb with respect
- Thinking about **gene content** is important but pre-existing comprehensive knowledge about underlying genes is normally limited
- Well established (but clinically variable) syndromes are emerging – every microdeletion syndrome has a duplication reciprocal
- **de novo status** is a strong but not absolute indicator of pathogenicity for a variant
- Remember **oligogenic inheritance underlie many (?)most** phenotypes under study – the discovery of a CNV (even if de novo) does not necessarily implicate it as the **sole unitary cause** of the phenotype
- Recourse to genome variant databases provides important evidence for pathogenicity – the lab do this prior to release of a result
- Because the use of parental samples is an important tool, counselling re, **incidental findings** and the implications of **non-paternity** is important
- **Beware of the X!**
A survey of human genetic variation

• Size of the genome: 3 billion bases
• Proportion subject to copy number variation – 2-10%
• Number of variable bases in the human - ?60 million
• Number of “variants” in a typical human – 4-5 million
• Number of new mutations in the human exome per generation – 0-5

When looking for a mutation the problem becomes sorting the wheat from the chaff
Whole-exome sequencing

• Un-biased approach to sequence “total exons” (1% of human genome).

• Analysis generates a huge amount of variation that needs to be interpreted with reference to:
  – Inheritance pattern
  – Parental exomes
  – Background variation (ethnicity may be important)
  – Phenotype

Whole-exome sequencing
So how do exomes perform?

- Non-syndromic MR (currently ~30% hit rate)
- “Syndromic” entities refractory to diagnosis
- Disorders with extreme phenotypic heterogeneity
  - Deafness
  - Retinitis pigmentosa
- The problems
  - Coverage, therefore variable confidence re. exclusion
  - Definition of the exome
  - Incidental findings
  - Private variants and variants of unknown significance
- Not the ultimate test
- Reiterative analysis continues to yield new hypotheses
Massively parallel sequencing
Is there a clinical sweet spot?

• Focused gene panels
  – Deafness, epilepsy, longQT syndrome, retinitis pigmentosa

• De novo dominant mutational mechanisms relatively easier to discriminate

• Finding mutations in known disease genes that retrospectively “fit” with the phenotype
Lennox Gastaut/infantile spasms

- N = 264
- Sequenced according to a trio design
- 329 de novo mutations
- Recurrent mutations in 2 previously unassociated genes – *GABRB3*/*ALG13*
- As a group genes were
  - Intolerant of mutation
  - Grouped into functions
- At an individual level, not much clinical utility
  - Not a genetic disorder?
  - ?extreme heterogeneity
So what about epigenetics?
Epigenetics today – why is it so limited in its utility clinically?

- Imprinting disorders
- Little else diagnostically useful in germline genetics currently
- Disorders of growth
  - SGA
  - Overgrowth
  - Tumour predisposition
- Dependence on tissue
- Epigenetic variation is a quantitative trait
Complex traits and personalised medicine

• The most preponderant diseases in Western society exhibit genetic susceptibility factors
• These factors are multiple, therefore susceptibility as a whole does not segregate like mendelian disorders
• The factors individually contribute very small effects
• Collectively they constitute a substantial contribution to whether an individual will contract a disease process or not
• Does that mean we can measure it in aggregate for an individual?
• If so, when will this vision be realisable?
So in 2020……

- Mendelian disease, most notably non-specific presentations
  - Use of targeted sequencing panels
- Epigenetics – limited usefulness
- A specialist role for exome analysis; little role for genome sequencing
- Genome driven personalised medicine still out of reach to a large degree

Anticipating the problems
- Heterogeneity (both genotypic and phenotypic)
- Oligogenic phenotypes
- Background variation
- Private disorders