Improving genomic diagnoses through accurate, specific phenotype information

Lisa Ewans
Clinical Geneticist, RPAH, Sydney
PhD student in genomics, KCCG, Garvan; UNSW

https://vlab.org
Overview

- Phenotype in the context of genomics
- How phenotype can be utilised to improve genomic diagnosis
- Focus on phenotypic features & capture
Scale of variant interpretation problem in genomics

- Start with 6 billion base pairs
- Determine variation from reference sequence
- 5 million base pairs after bioinformatics
- Clinical Interpretation apply filters:
  - Inheritance models
  - Allele frequencies (population databases)
  - Known disease genes
  - Phenotype is a key final filter
- Aim: find the needle in the haystack

![Diagram showing the scale of variant interpretation problem in genomics]

- WGS Variant Filtering
  - ~ 3,000 variants
- Inheritance pattern
  - ~ 900 variants
- Population database
  - ~ 200 variants
- Known disease gene
  - 1-2 variants
- Diagnosis
Interpreting variants in genomics
Why is phenotype important?

- Having accurate phenotypic information is crucial in medical diagnostic investigation
  - Directs appropriate testing & aids interpretation
- Important for pre-test selection in genetic testing & interpretation of identified genetic variation
  - Particularly crucial in genomic tests due to high numbers of gene variants
- Phenotyping increases accuracy of genetic testing (Baynam et al. 2014)
- In genomic testing, phenotype is the starting point for assessing post-filtering variant relevance

https://monarchinitiative.org/
What does phenotype mean?

• Phenotype definition:
  – “Set of observable characteristics of an individual resulting from the interaction of its genotype with the environment” – Encyclopedia Britannica
  – In medicine: “Deviation from normal (or average) morphology, physiology, and behaviour” – Robinson et al. 2012

• Medical phenotype obtained by clinician: history, examination, etc.
  1. Symptom
  2. Sign
  3. Physical feature e.g. micrognathia, microcephaly
  4. Structural malformation e.g. coloboma
  5. Diagnosis e.g. spinal canal stenosis, renal tubular acidosis, anaemia
  6. Investigation result e.g. hyperlactinaemia

• Phenotypes may not be static over the course of a lifetime
• Deep phenotyping – precise & comprehensive analysis
How can phenotypes be utilised to make genomic diagnoses?
Accuracy & specificity of phenotype is important

- Variant interpretation is often guided by the individual’s phenotype
- Incorrect phenotype → relative importance of a gene variant may be over or under-valued
- If phenotype is absent, or insufficient information, variant interpretation is difficult
- An individual’s genomic testing results often contain variants in genes with overlapping phenotype – best phenotype match?
- More specific & rare the phenotype, the more likely a linked phenotype-genotype correlation is causative
Down a rabbit hole

- WES case – clinical diagnosis of vitamin B12 deficiency “please assess CUBN gene”
- 1 pathogenic-looking CUBN gene variant identified (AR disorder)
- Search for second variant fruitless
- WES reanalysis 12 months later identified a heterozygous pathogenic RPS26 gene
- Clinical file reviewed: Diamond-Blackfan anaemia diagnosis had been suggested
- Resulted in delay in diagnosis
Phenotype ontologies

• Categorise phenotypes & show relationships between them & disease
• Human phenotype ontology (HPO)
• Software/platforms to link listed phenotypes with genes – either known human disease genes or biological plausibility
  – Monarch initiative, PhenomeCentral
  – Exomiser (Sanger Institute), Phevor
  – Limitations e.g.:
    • Up-to-date information in databases such as OMIM, Orphanet
    • Patient gestalt may be missing if only phenotype information available is e.g. HPO terminology

https://monarchinitiative.org/
Using phenotype to refine big data

- 4 year old Aboriginal girl with undiagnosed disorder, likely genetic
- Undiagnosed diseases program node in Western Australia
- Unusual constellation of symptoms
- Trio WGS – Genome.One
- 2 interesting candidate genes
  - Schwartz-Jampel syndrome
  - Autosomal recessive – 1 pathogenic variant
- De novo splice region variant
- Russell-Silver like phenotype
- Linked to hypoglycaemia
Collaboration crucial for understanding genomic variation

- MDT virtual meeting with Dr. Gareth Baynam reviewed variants
- IGF2 gene clinical overlap considered significant
- Variant could not be classified as pathogenic due to splice region location (c.157+3A>C)
- Paternally inherited?
- PathWest – cell line available to perform RNA studies
  - Confirmed splicing defect
- Mother amazed at facial similarity
- IGF2 in growth hormone pathway & patient has responded to GH treatment

“When I saw the other children’s faces, it was like oh my god....what the h**l...I couldn’t believe it...”

- Patient’s Mother

Poulton et al. 2018
Begemann et al. 2018
Collaboration leads to better results

- Clinical – lab liaison results in optimal outcomes
- Local communication lines more straight forward, however, clinicians/labs at a distance just an email or phone call away
- Mutual beneficial education
- For example, clinicians understanding benefits of providing some but not too much phenotype information
  - Multisystem adult onset diseases – listing all features may confuse interpretation
Keeping an open mind – Genetic pleiotropy

• Referral to a genomic service of single affected patient aged 18 months
  – Developmental delay, dysmorphic features (incl. short palpebral fissures & short columella), microcephaly, gut malrotation
• Following WES, 3 variants of interest in ID-related genes were identified
  – How to differentiate?
• Genotype-phenotype assessment of 1 CREBBP gene variant
  – Initial call: Likely unrelated as associated disorder Rubinstein-Taybi syndrome had insufficient overlap
  – Clinician familiar with this gene suggested closer inspection as particular locus had been recently described with different phenotype
Keeping an open mind – Genetic pleiotropy (2)

Classic Rubinstein-Taybi syndrome
- Long columella, grimacing smile, broad & angulated great toes & thumbs
- CREBBP/EP300 genes

Variants within exons 30 & 31 of CREBBP confer a different phenotype to RTS
Keeping an open mind - Exceptions to the rule

- 6 year old girl with undiagnosed syndrome
- Global developmental delay, complex cardiac defect, distinctive facial features
- No family history of relevance
- UDP patient - trio whole genome sequencing at Genome.One
- Interesting de novo variant in TAF1 gene, but only reported in males with intellectual disability
- Striking overlap between patient features & publications
- Phenotype critical in linking genomic variant to patient diagnosis & patient included in follow up case series

O'Rawe et al. 2015
Important phenotypic features in genomics

- Phenotypic features should be accurate and available to those analysing genomic data
- The specificity of rare phenotypes aids variant interpretation for diagnosis
- Leveraging deep phenotyping enables broad approaches to testing (WES, WGS) and improves diagnostic yields
- Phenotype sub-characteristics
  - E.g. congenital vs. acquired microcephaly
- Overall “gestalt” of referred individual
Enabling capture of phenotypic information

- Electronic request forms, prompts for clinicians, free-text space
- Patient-reported phenotypes? (Phenotypr, https://phenotypr.com/; MyGene\(^2\), http://mygene2.org/)
- Phenotype mining from electronic health record vs genetic letters (Son et al. 2018)
Summary

• Specific phenotypic features are important in linking variants with patient diagnosis
• Combinations of AI & specialty knowledge, particularly with teams of experts maximises results
• Laboratory-clinical liaison/dialogue facilitates capturing of phenotypic information & diagnosis
References

- G Baynam et al. Phenotyping: Targeting genotype’s rich cousin for diagnosis. *JPCH.* 2014. 51(4)