The Search for Genetic Answers

Identifying candidate variants in families with unsolved rare genetic diseases

Rhett Marchant[1,2], Samantha Bryen[1,3], Leigh Waddell[1,3], Frances Evesson[1,4], Michaela Yuen[1,3], Himanshu Joshi[1,4], Sandra Cooper[1,3,4]

Overview

1. Rare Genetic Diseases

- **Rare genetic diseases** collectively affect **1 in 16** people worldwide.(1)
- Diagnosis of a patient's **causal genetic variant** is essential for disease management but only **50%** of patients are **diagnosed** following **initial** genetic evaluation.(2,3)

2. Pre-mRNA Splicing

- Normal protein production relies on correct mRNA processing including **pre-mRNA splicing**, a process primarily regulated by **intronic sequences**.
- Variants which disrupt these sequences are estimated to cause **15-60%** of genetic diseases, but are **underdiagnosed by current diagnostics**.(4,5)

3. Approach

- **Massively parallel sequencing** (MPS) has improved diagnosis by an additional 10-40%.(6,7)
- This study combined whole exome (WES), genome (WGS), and RNA sequencing (RNA-seq) data, to try **identify candidate causal variants in 5 undiagnosed families**.

Hypothesis

Causal genetic variants which **disrupt pre-mRNA splicing** and/or **occur in novel disease genes** are enriched in cohorts of undiagnosed families with rare genetic diseases.

Aim

Create a **bioinformatics workflow** using massively parallel sequencing data and ***in-silico* tools** to identify **candidate pathogenic variants**, especially those disrupting pre-mRNA splicing.

Results

The workflow utilised three MPS technologies, three computational tools, and three filtering parameters (Fig. 1).

For the three families with WGS and RNAseq, we focussed on mis-splicing variants via two separate approaches:

1. WGS + SpliceAI variant mis-splicing prediction (8)
2. RNAseq + FRASER mis-splicing event detection (9)

Both approaches returned candidate variants using low thresholds on tool score and variant population frequency.(10) The power of the tools and trio variant segregation was such that filtering by gene lists was not necessary. (Fig. 2) This facilitated novel gene discovery.

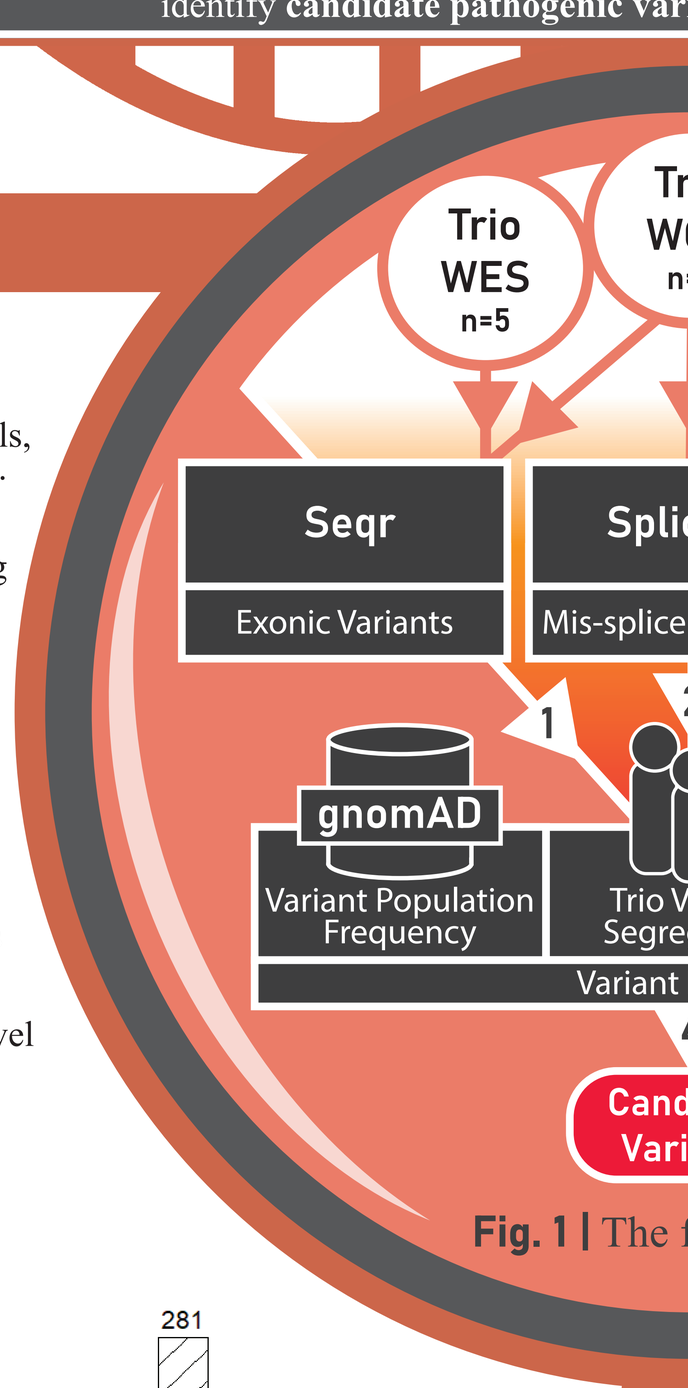

Fig. 1 | The final workflow

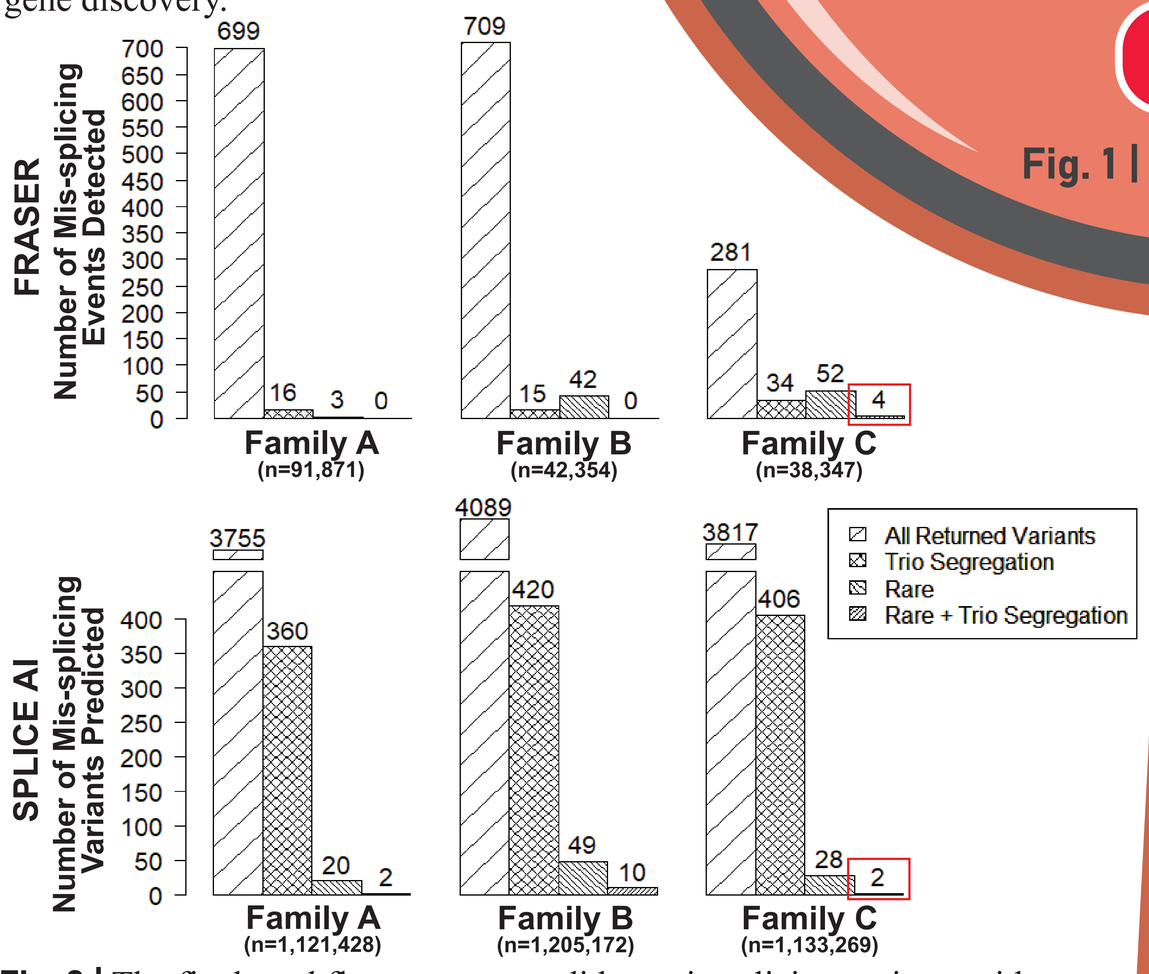

Fig. 2 | The final workflow returns candidate mis-splicing variants without gene list filtering. Red boxes highlight *CDK5RAP3* variant in Family C

CDK5RAP3

This mis-splicing analysis workflow revealed 1 candidate variant in a novel disease gene, *CDK5RAP3,* in Family C (Fig. 3).

This gene is not associated with human disease but is embryonic lethal in mouse knock-out models.

Additionally, several disease genes in the same biochemical pathway as *CDK5RAP3* cause a distinct phenotype similar to the patient's in this family.

Further in-vitro analysis revealed that the variant segregated with disease in the family and thus, this intronic variant is likely responsible for the rare disease.

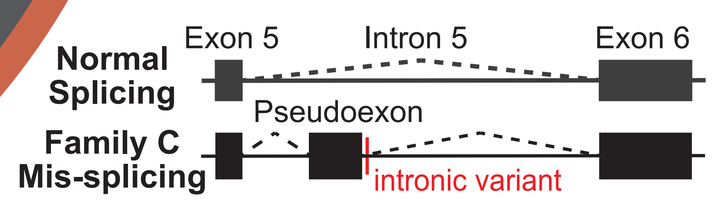

Fig. 3 | An intronic variant in Family C in *CDK5RAP3* disrupts normal splicing by activating a pseudoexon

Conclusions & Future Directions

By leveraging two powerful tools to **identify mis-splicing** at both the **DNA and RNA level**, this workflow has provided a genetic answer for **1 of 5** analysed families with unsolved rare genetic disease. This result supports our hypothesis.

Further validation with both solved and unsolved cases will be essential to understand the workflow's limitations. However, even in its current form, this approach is likely to **increase diagnostic rates of families with unsolved rare genetic diseases.**

References

1. Ferreira CR. The burden of rare diseases. Am J Med Genet Part A. 2019 Jun 1;179(6):885–92.
2. Hartley T, Lemire G, Kernohan KD, Howley HE, Adams DR, Boycott KM. New Diagnostic Approaches for Undiagnosed Rare Genetic Diseases. Annu Rev Genomics Hum Genet. 2020 Aug 31;21(1):351–72.
3. Shashi V, McConkie-Rosell A, Rosell B, Schoch K, Vellore K, McDonald M, et al. The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. Genet Med. 2014;16(2):176–82.
4. López-Bigas N, Audit B, Ouzounis C, Parra G, Guigó R. Are splicing mutations the most frequent cause of hereditary disease? FEBS Lett. 2005;579(9):1900–3.
5. Wai HA, Lord J, Lyon M, Gunning A, Kelly H, Cibin P, et al. Blood RNA analysis can increase clinical diagnostic rate and resolve variants of uncertain significance. Genet Med. 2020;22(6):1005–14.
6. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. npj Genomic Med. 2018;3(1):1–10.
7. Frésard L, Smail C, Ferraro NM, Teran NA, Li X, Smith KS, et al. Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. Nat Med. 2019;25(6):911–9.
8. Jaganathan K, Kyriazopoulou Panagiotopoulou S, Mcrae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019;176(3):535-548.e24.
9. Mertes C, Scheller IF, Yépez VA, Çelik MH, Liang Y, Kremer LS, et al. Detection of aberrant splicing events in RNA-seq data using FRASER. Nat Commun. 2021;12(1):1–13.
10. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434–43.

1 Kids Neuroscience Centre, Kids Research, The Children's Hospital at Westmead, Sydney, NSW 2145, Australia
2 Discipline of Applied Medical Science, Faculty of Medicine and Health, University of Sydney, Sydney, NSW 2006, Australia
3 Discipline of Child and Adolescent Health, Sydney Medical School, University of Sydney, Sydney, NSW 2006, Australia
4 The Children's Medical Research Institute, 214 Hawkesbury Road, Sydney, NSW 2145, Australia and Faculty of Medicine and Health, University of Sydney, Sydney, NSW 2006, Australia

THE UNIVERSITY OF SYDNEY
kr kids research
kids neuroscience centre
The Sydney children's Hospitals Network
care, advocacy, research, education