MINTIE identifies novel structural and splice variants in **RNA-seq** data

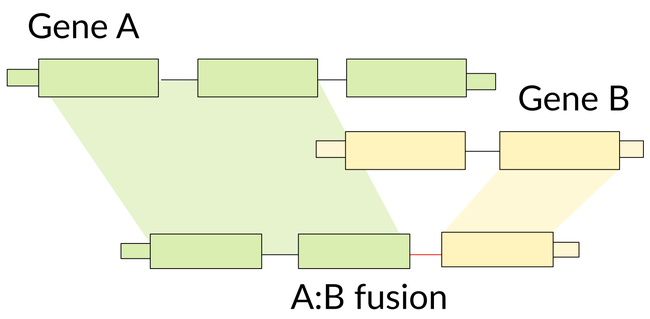
Marek Cmero, Breon Schmidt, Ian Majewski, Paul Ekert, Alicia Oshlack, Nadia Davidson

Motivation

- **Genomic rearrangements** can modify gene function, and have been shown to be **drivers** in both **cancer** and **rare disease**.
- Some transcriptomic variants (such as **irregular fusions** and **duplications**) are difficult to detect in **RNA-seq**.

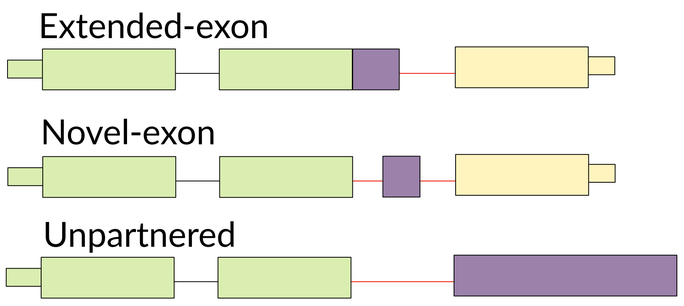
Transcriptomic variants

Let's consider the **canonical fusion** as a single gene product formed by two genes joined at an **exon-exon boundary**:



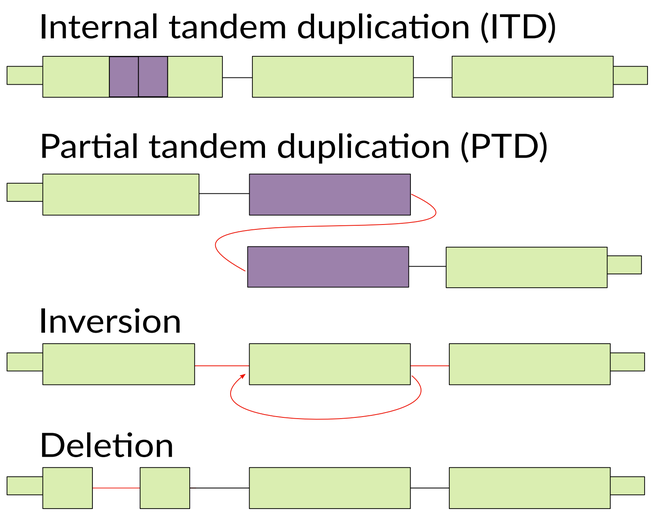
Fusion finders use very strict filters and do not consider non-standard fusions, such as:

Non-canonical fusions

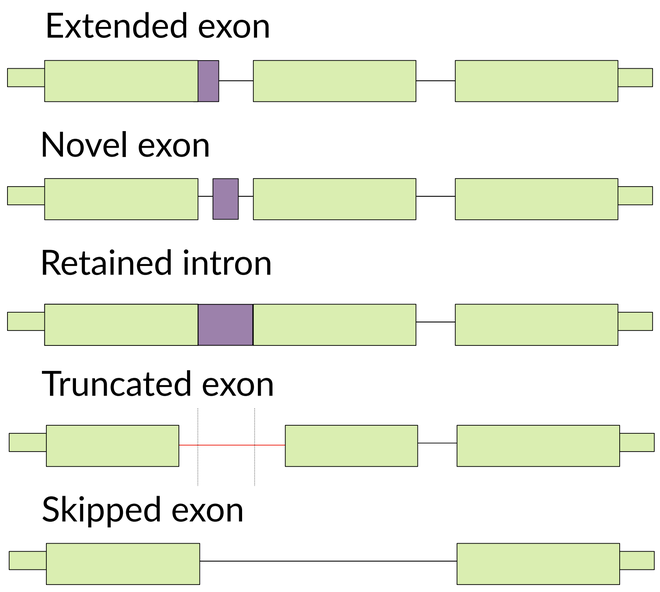


We also consider **transcriptomic variants** in single genes:

Transcribed structural variants

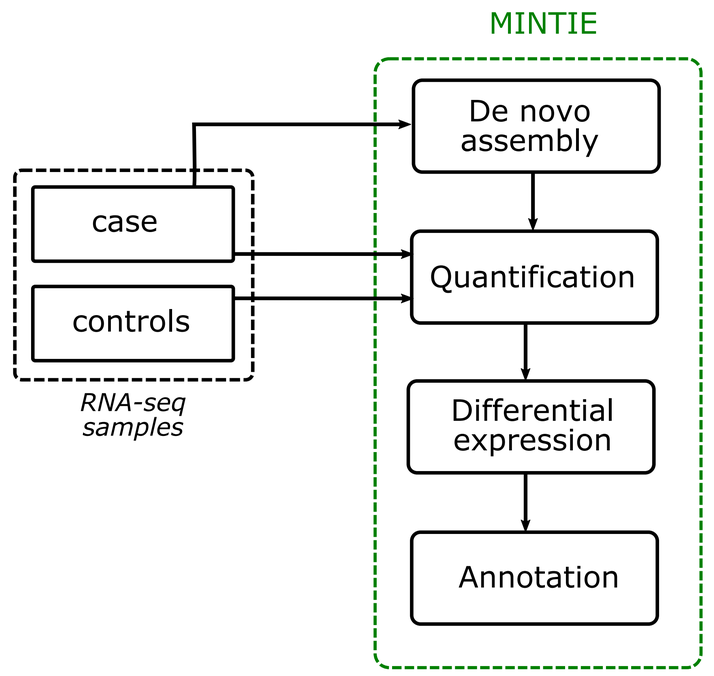


Novel splice variants



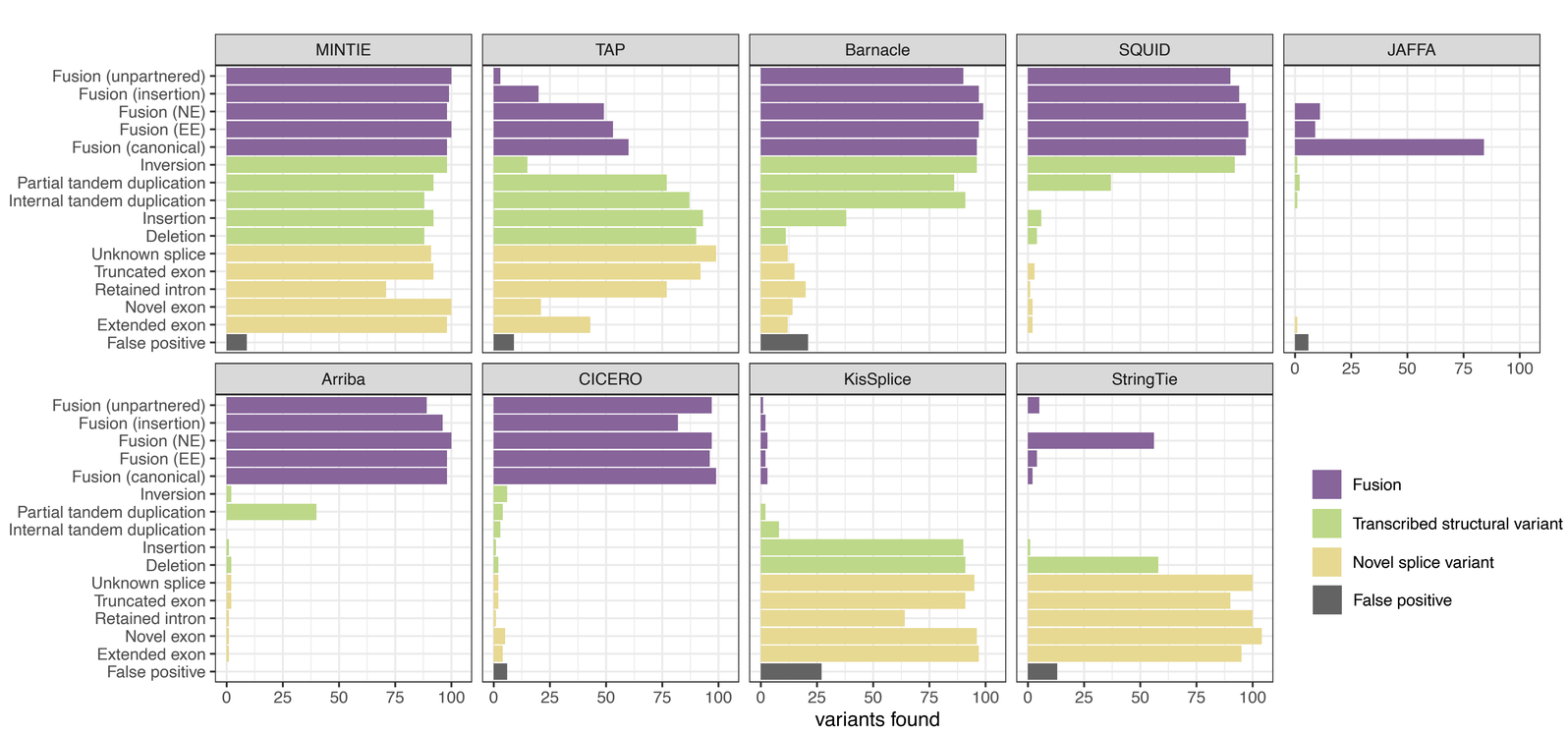
The MINTIE method

1. **De novo assemble** all transcripts in the case sample.
2. **Quantify** all assembled transcripts in case and a set of controls (do not need to be normals).
3. Perform **differential expression** on assembled transcripts (case vs. controls).
4. **Align** significant transcripts and **identify novel variants.**



Simulation results

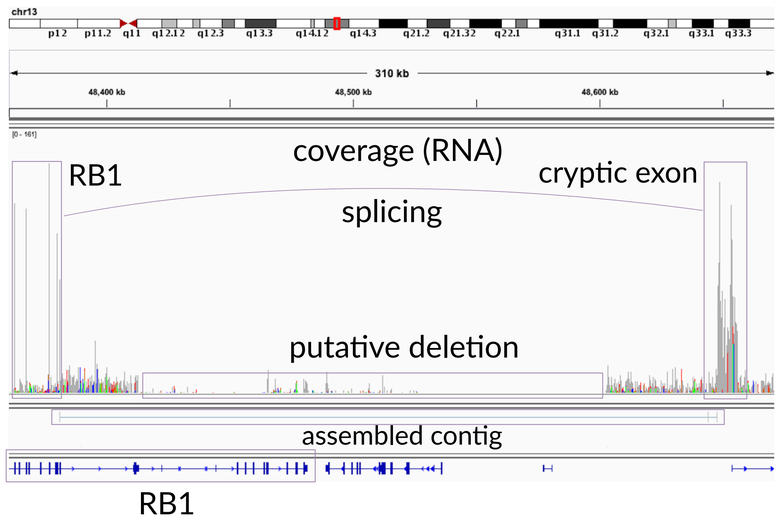
- We simulated **1500 variants** across **15 types**.
- **MINTIE** successfully found **>70%** of variants across all types.
- We ran these simulations on 8 other tools and found that **MINTIE** could find and annotate **more variants than any other method**.



Candidate variants in cancer and rare disease

B-cell Acute Lymphoblastic Leukaemia

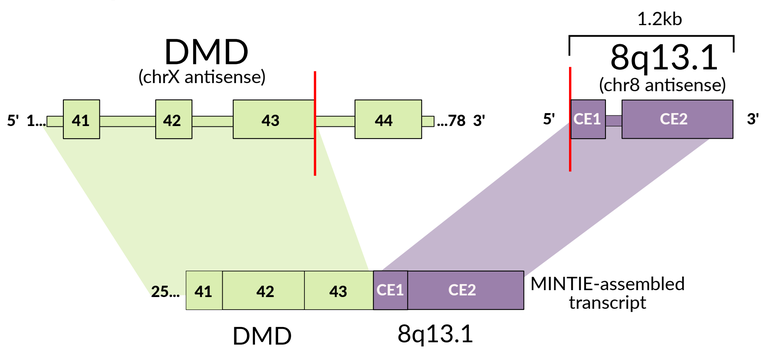
We ran **MINTIE** on 91 B-ALLs (Brown et al. 2020) from Melbourne's Royal Children's Hospital, and found several **clinically relevant** transcriptomic variants, including a **recurrent RB1 unpartnered fusion** found in three patients, novel **ETV6** splice variants found in two patients and an **IKZF1** and **PAX5 PTD**, each found in one patient.



Novel non-canonical fusion reveals likely **loss-of-function deletion** in **tumour suppressor RB1**.

Rare muscle disease

We ran **MINTIE** on the **RNA-seq** from a cohort of 46 **rare muscle disorder** patients from a prior study (Cummings et al. 2017), and found **three unpartnered fusions** involving the muscle disease associated **DMD** gene, two of which were **only identified in the DNA** in the original study, and one of which was **missed entirely**.



Previously undetected non-canonical inter-chromosomal fusion between muscle disease-associated gene **DMD** and intergenic region on chromosome 8.

Peter Mac
Peter MacCallum Cancer Centre
Victoria Australia

code: github.com/Oshlack/MINTIE
preprint: doi.org/10.1101/2020.06.03.131532