Identification of rare-disease genes in diverse undiagnosed cases using whole blood transcriptome sequencing and large control cohorts

Pathology retreat - October 6th 2018

Laure Fresard, Montgomery lab

Identification of rare-disease genes in diverse undiagnosed cases using whole blood transcriptome sequencing and large control cohorts

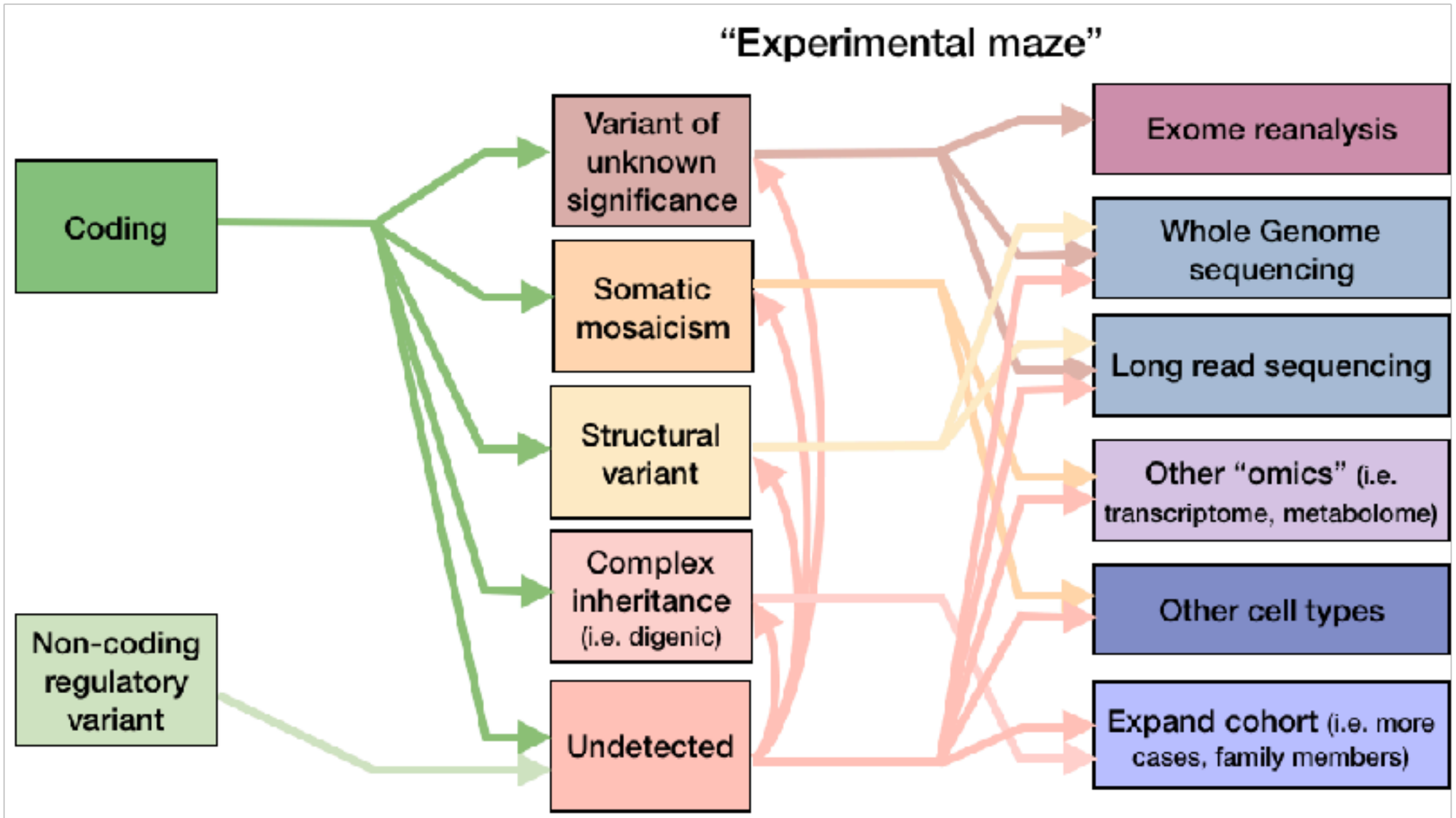
Pathology retreat - October 6th 2018
Laure Fresard, Montgomery lab

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

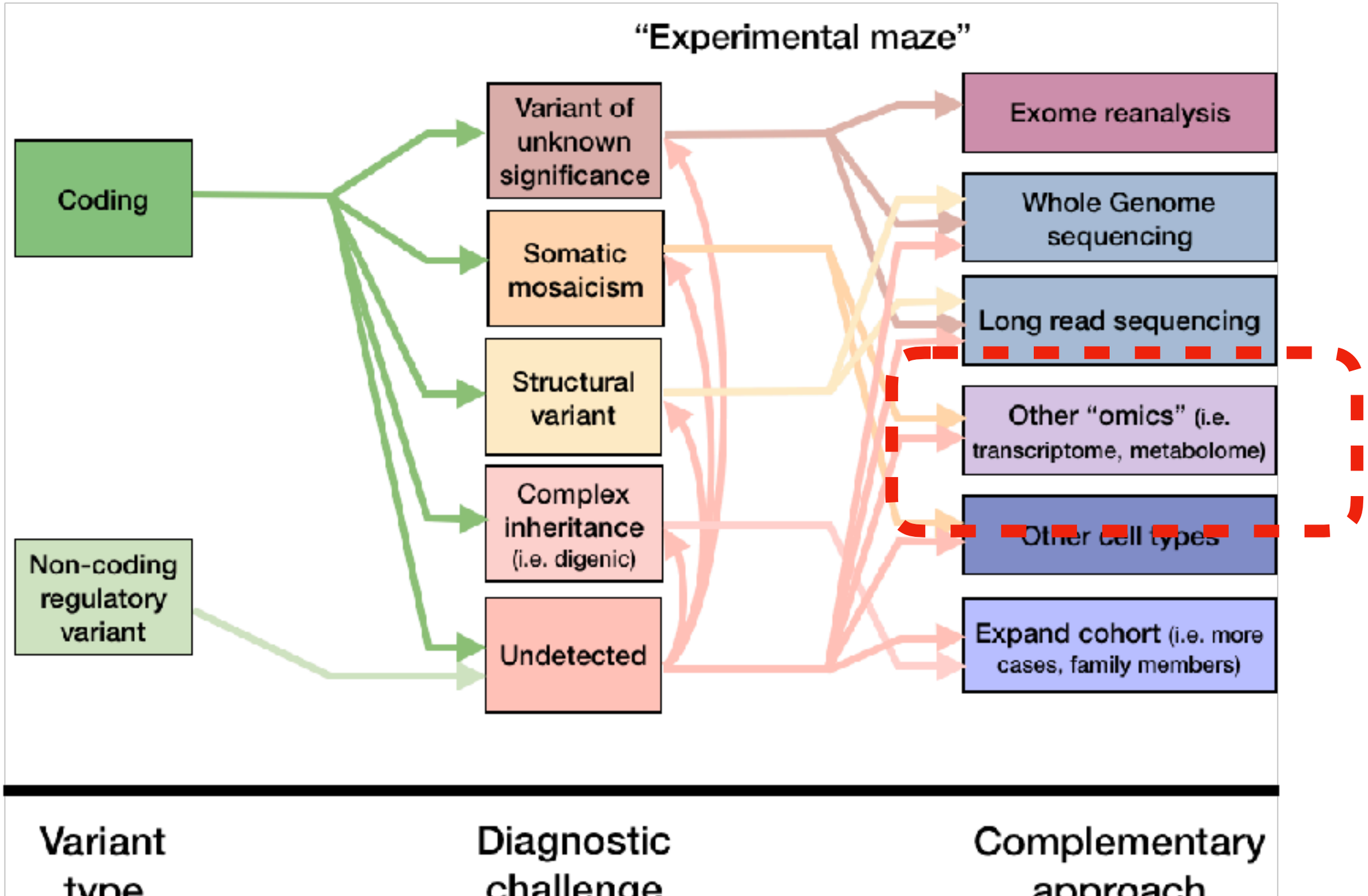
Identification of rare-disease genes in diverse undiagnosed cases using whole blood transcriptome sequencing and large control cohorts

Laure Fresard, Craig Smail, Kevin S. Smith, Nicole M. Ferraro, Nicole A. Teran, Kristin D. Kernohan, Devon Bonner, Xin Li, Shruti Marwaha, Zachary Zappala, Brunilda Balliu, Joe R. Davis, Boxiang Liu, Cameron J. Prybol, Jennefer N. Kholer, Diane B. Zastrow, Dianna G. Fisk, Megan E. Grove, Jean M. Davidson, Taila Hartley, Ruchi Joshi, Benjamin J. Strober, Sowmithri Utiramerur, Care4Rare Canada Consortium, Undiagnosed Diseases Network, Lars Lind, Erik Ingelsson, Alexis Battle, Gill Bejerano, Jonathan A. Bernstein, Euan A. Ashley, Kym M. Boycott, Jason D. Merker, Matthew T. Wheeler, Stephen B. Montgomery

Exome sequencing has been estimated to lead to a diagnosis in 30 to 50% of rare Mendelian diseases



Exome sequencing has been estimated to lead to a diagnosis in 30 to 50% of rare Mendelian diseases



We used whole blood RNA-seq to identify rare disease genes

- We partnered with 3 rare disease programs with the goal to develop generalizable methods for a broad spectrum of diseases

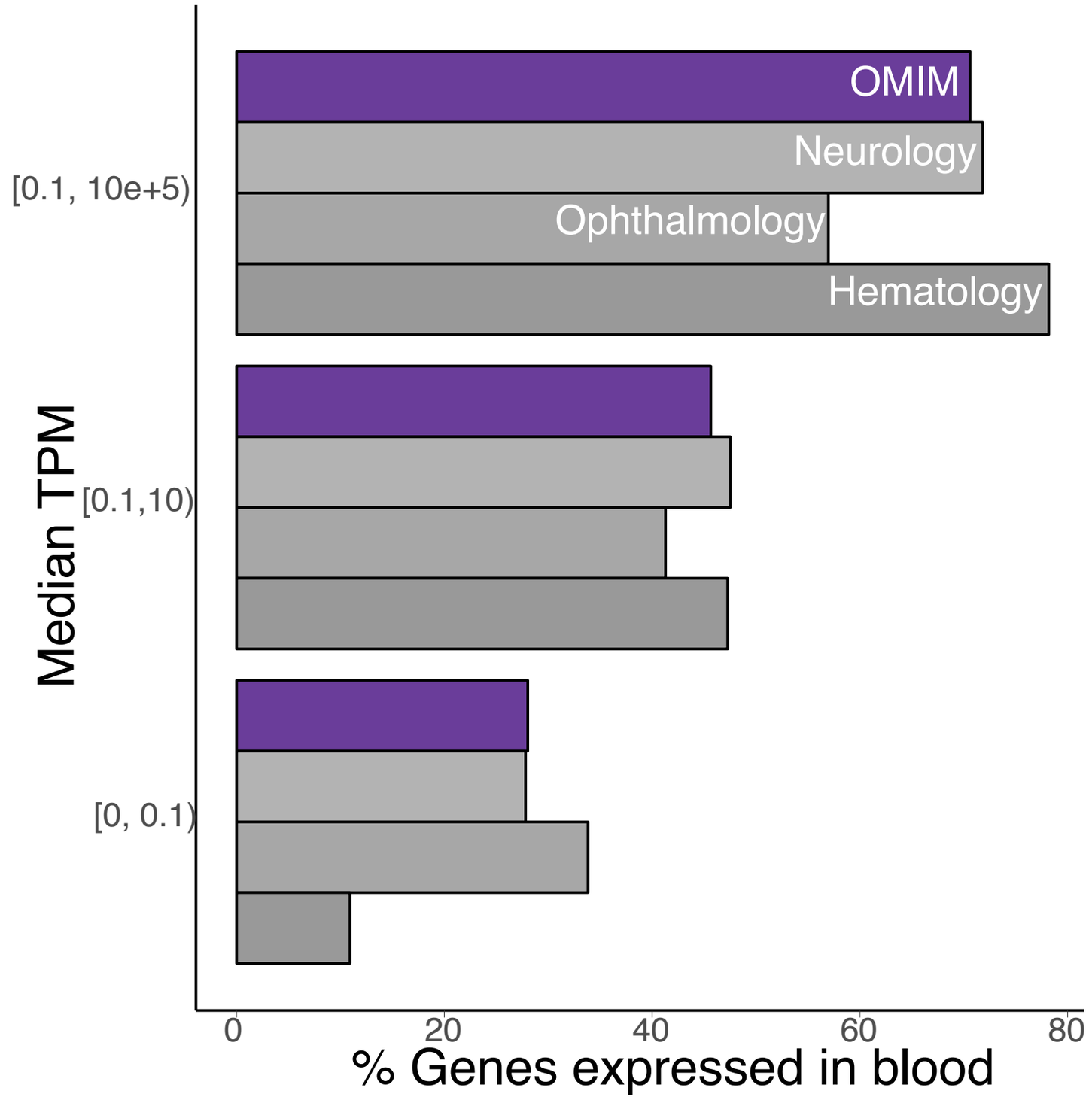
- We've sequenced 87 whole blood samples, 56 extracted from affected individuals and 31 unaffected family members

- The 56 cases represent a total of 47 independent diseases

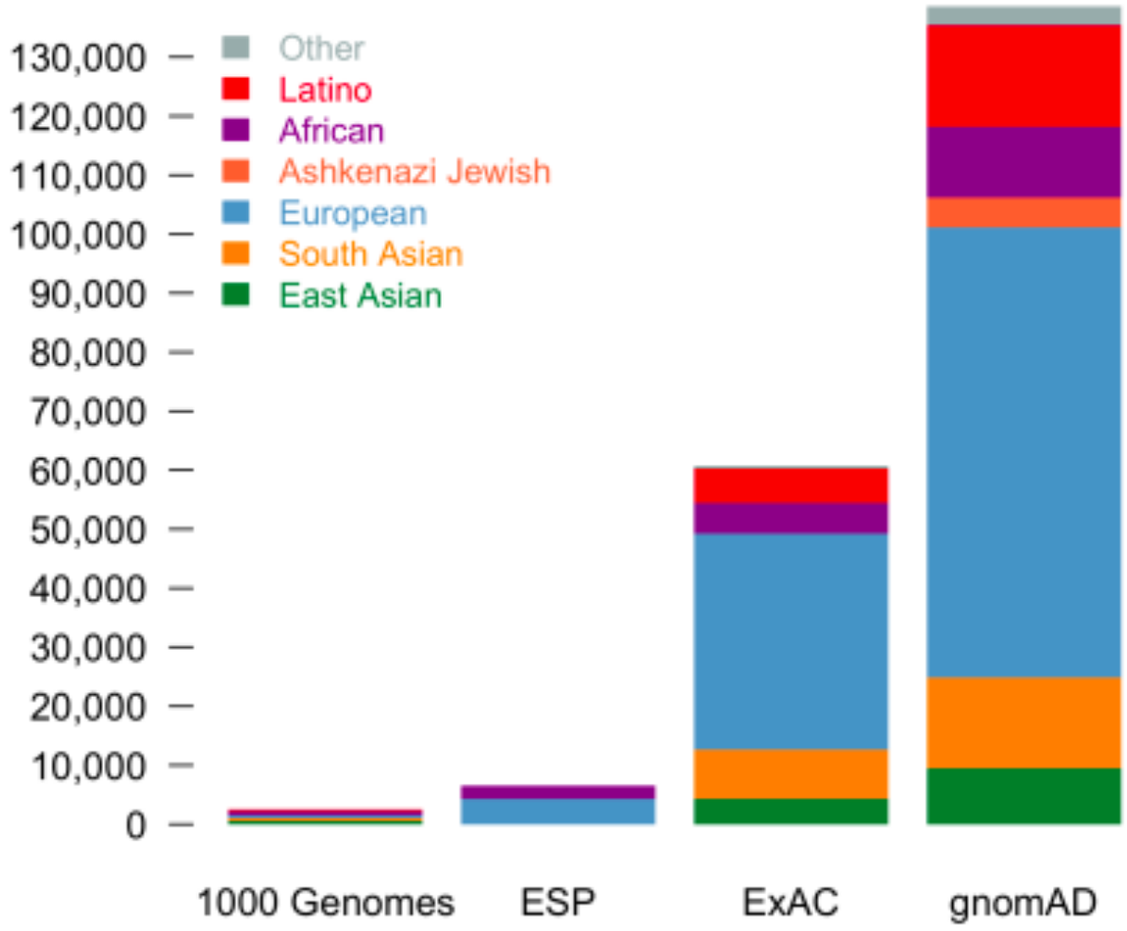
- We used large control cohorts to detect aberrant events in gene expression, splicing and allele specific expression

Can whole blood be useful?

The majority of known rare disease genes are expressed in blood



Population scale sequencing analyses give access to gene sensitivity to mutations

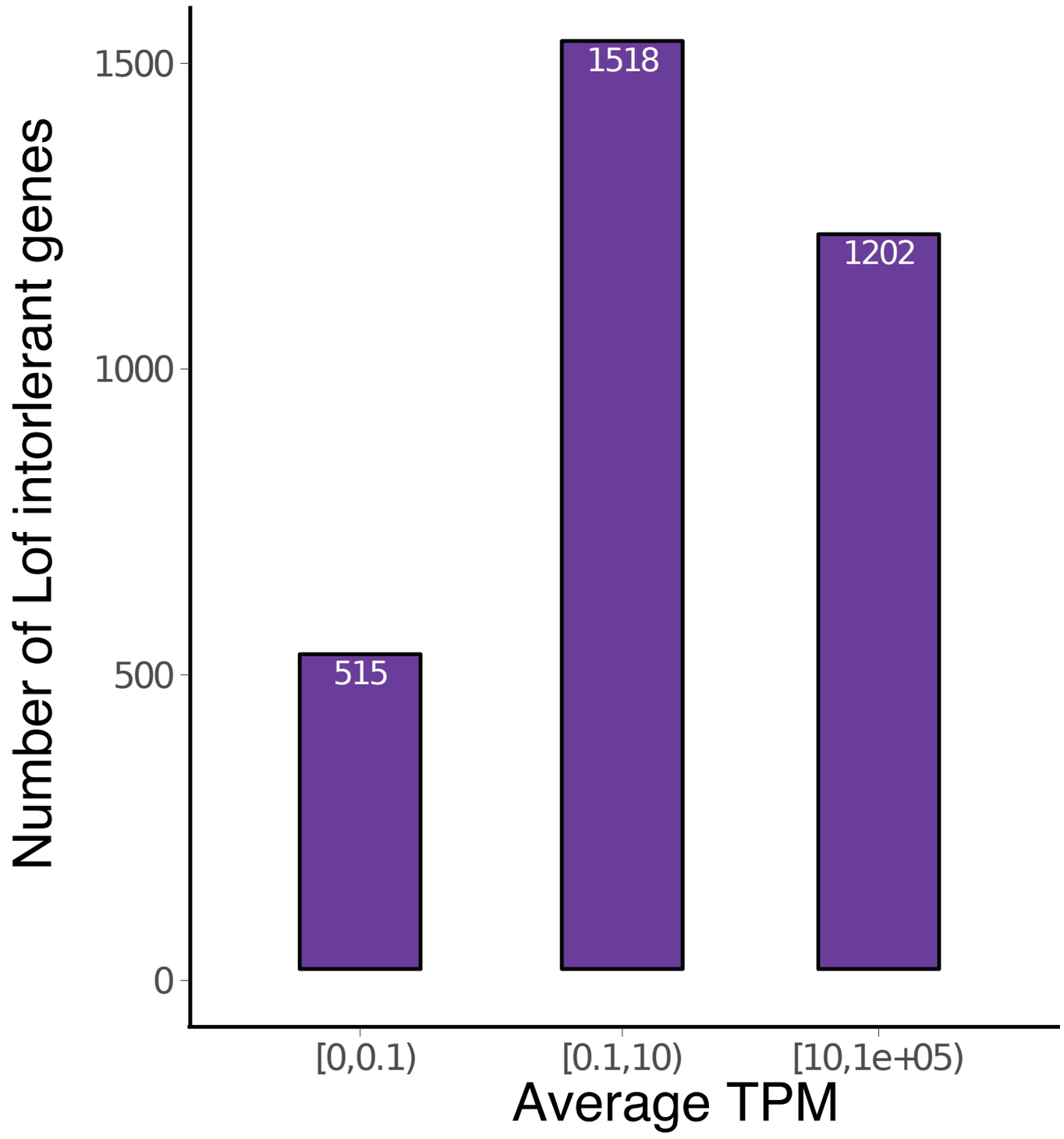


2017/02/27 release: exome sequence data from 123,136 individuals and whole genome sequencing from 15,496 individuals

Reference of allele frequency distributions across genes ==> Genes that are more sensitive to variants

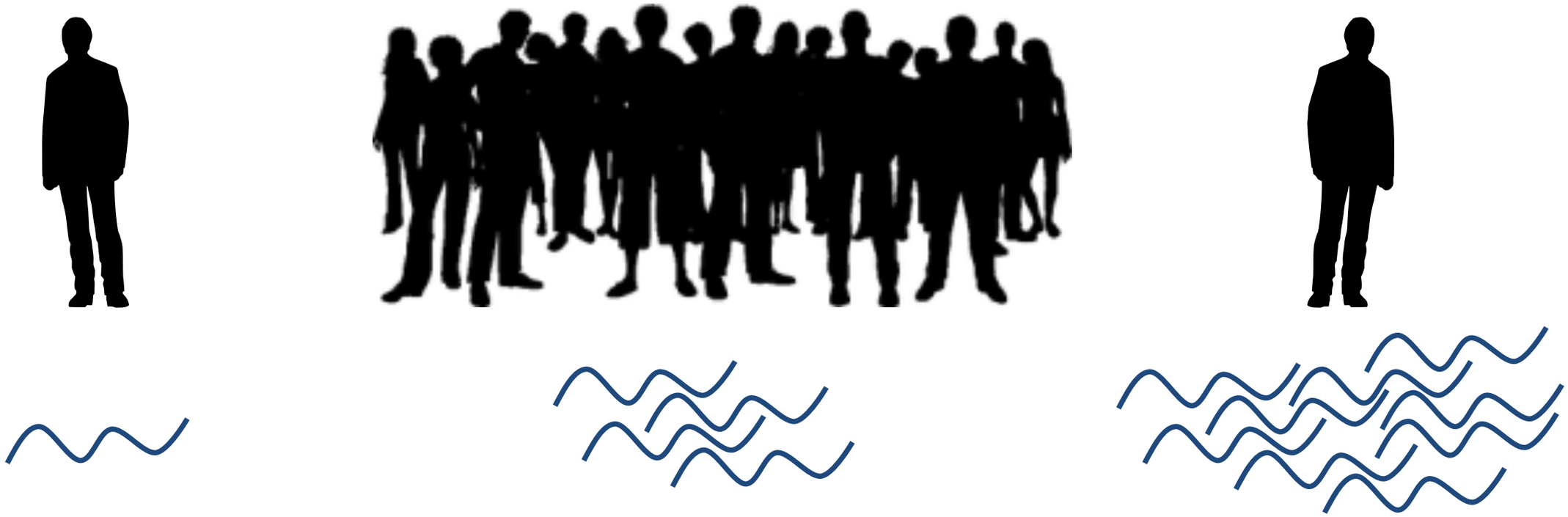
Konrad J. Karczewski, Eric V. Minikel, Kaitlin E. Samocha et al, Nature 2016

66% of Loss-of-Function intolerant genes are expressed in blood samples



Can RNA-seq be useful?

Gene expression/splicing/ASE outliers can pinpoint to rare variants with large effects



Underexpression outlier

Overexpression outlier

Xin Li, Yungil Kim, Emily K. Tsang , Joe R. Davis, et al, Nature 2017

RNA-seq has shown diagnostic utility in specific tissue and diseases

SHARE

0

0

RESEARCH ARTICLE | GENETIC DIAGNOSIS

Improving genetic diagnosis in Mendelian disease with transcriptome sequencing

Beryl B. Cummings[1,2,3], Jamie L. Marshall[1,2], Taru Tukiainen[1,2], Monkol Lek[1,2,4,5], Sandra Don...
+ See all authors and affiliations

Science Translational Medicine 19 Apr 2017:
Vol. 9, Issue 386, eaal5209
DOI: 10.1126/scitranslmed.aal5209

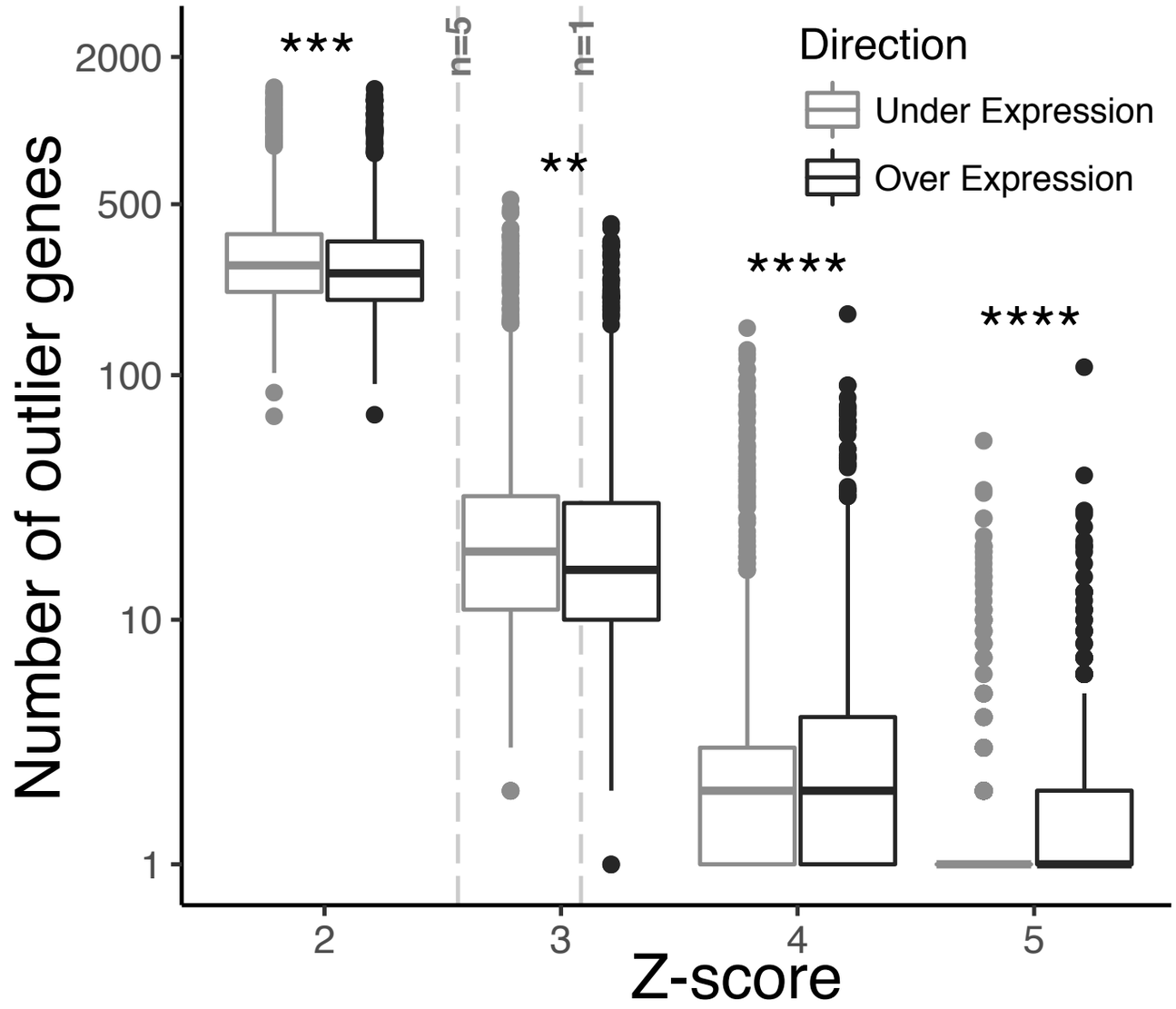
nature COMMUNICATIONS

Article | OPEN | Published: 12 June 2017

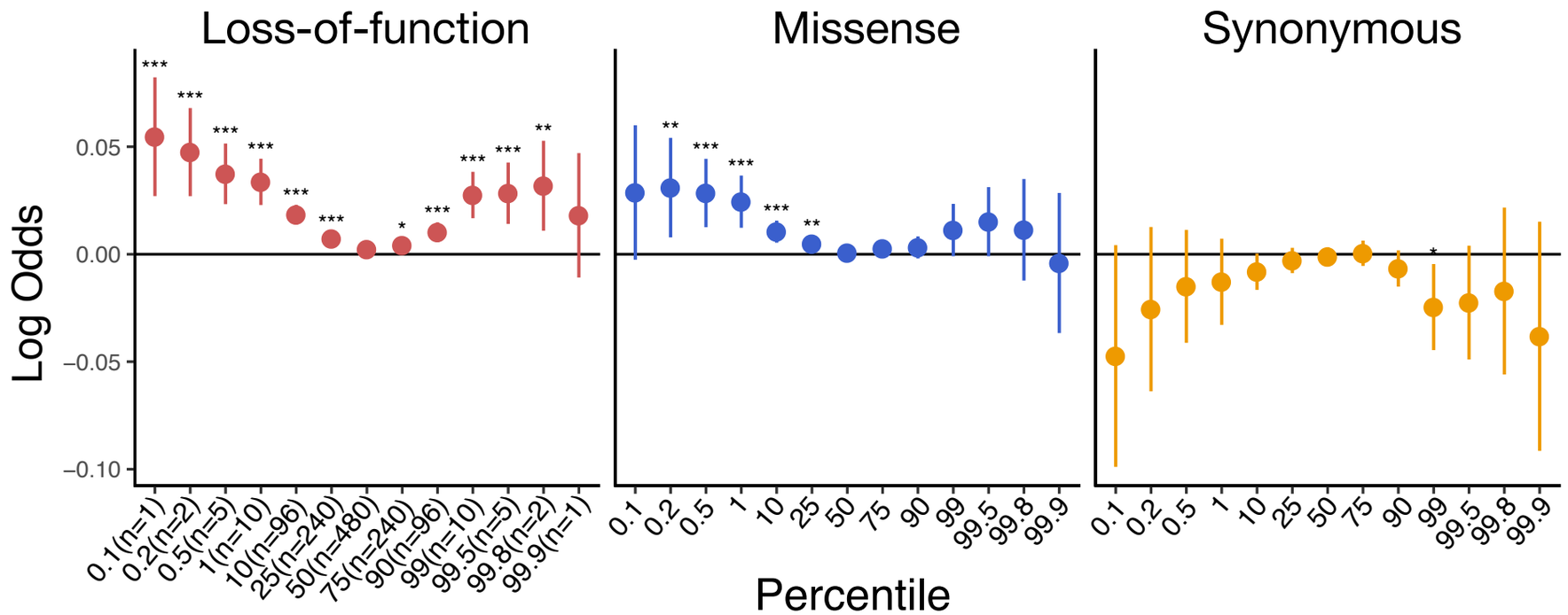
Genetic diagnosis of Mendelian disorders via RNA sequencing

Laura S. Kremer, Daniel M. Bader [...] Holger Prokisch

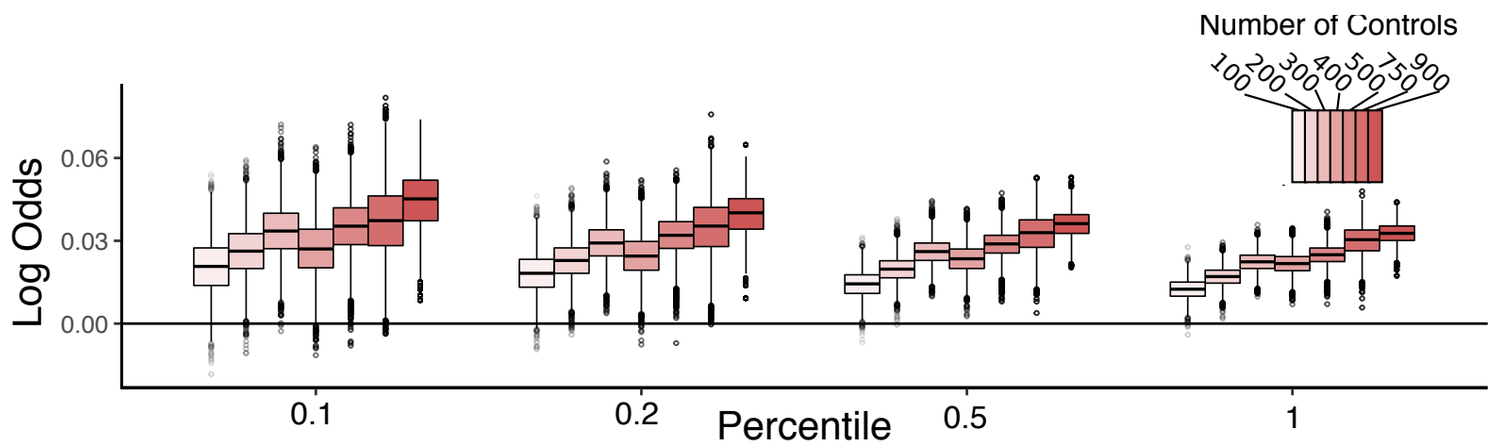
We observed an average of 350 outliers per sample (|Z-score| ≥2)



Gene expression outliers are enriched for loss-of-function intolerance

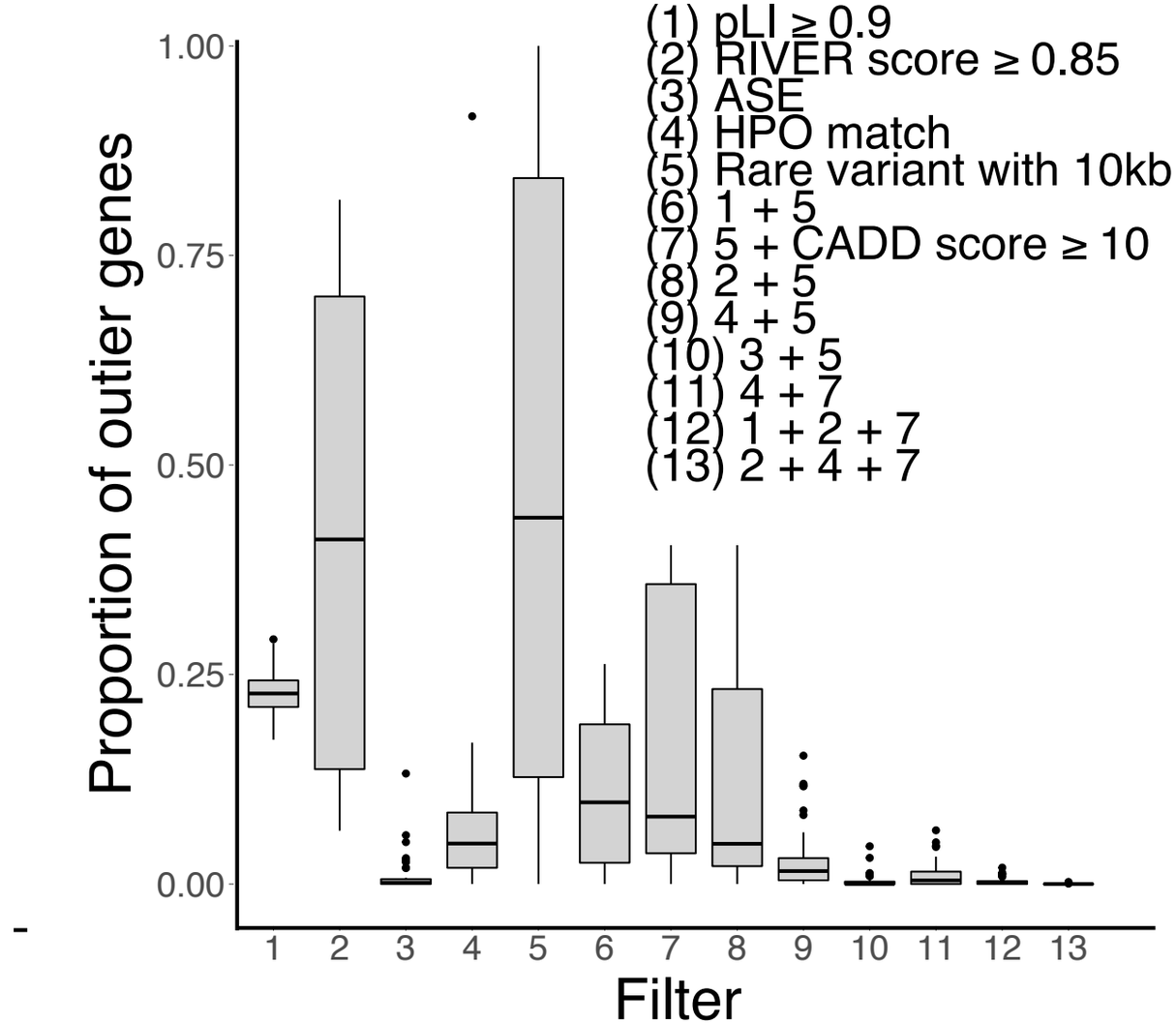


We observed an enrichment of case outliers in genes more sensitive to Loss-of-function mutations



The enrichment becomes stronger as we increase the number of controls

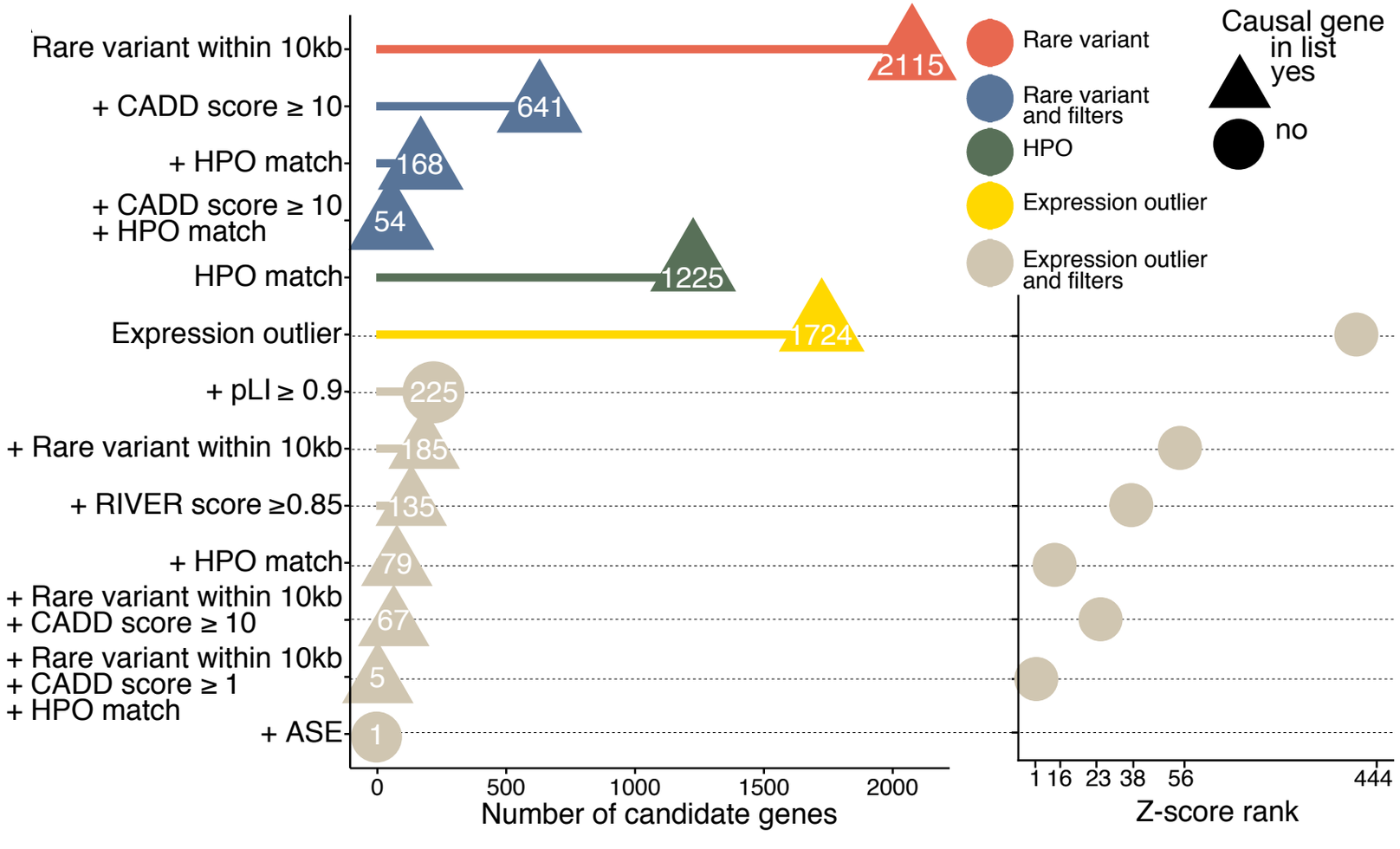
Filtering for relevant attributes helps narrow down to a handful of candidates



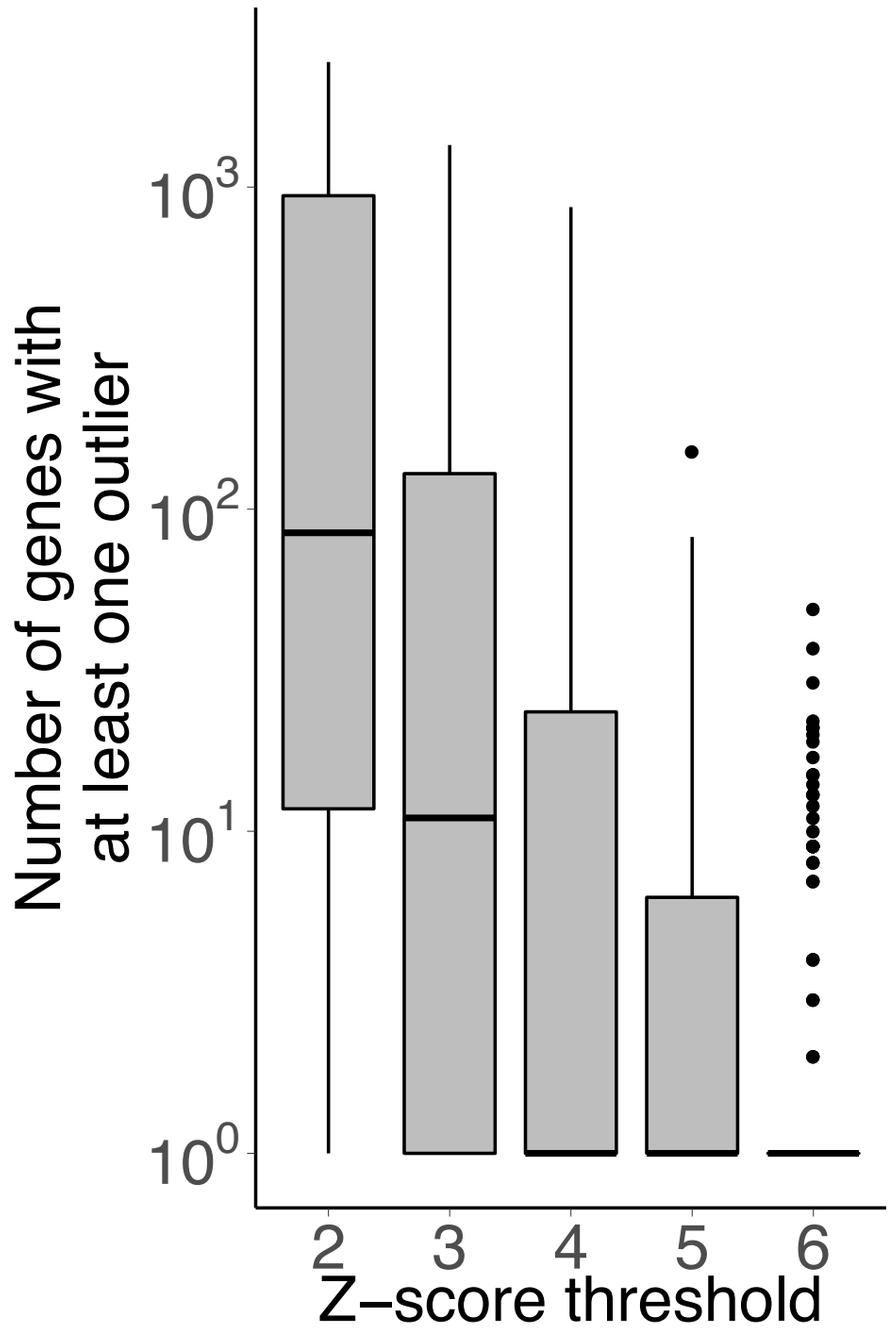
When restricting to **under-expression outlier genes** with **HPO matches** and a **deleterious rare variant** nearby, we reduce the candidate genes list to less than 10% of the initial set of outliers

Identification of disease gene through expression outlier detection: the *RARS2* case.

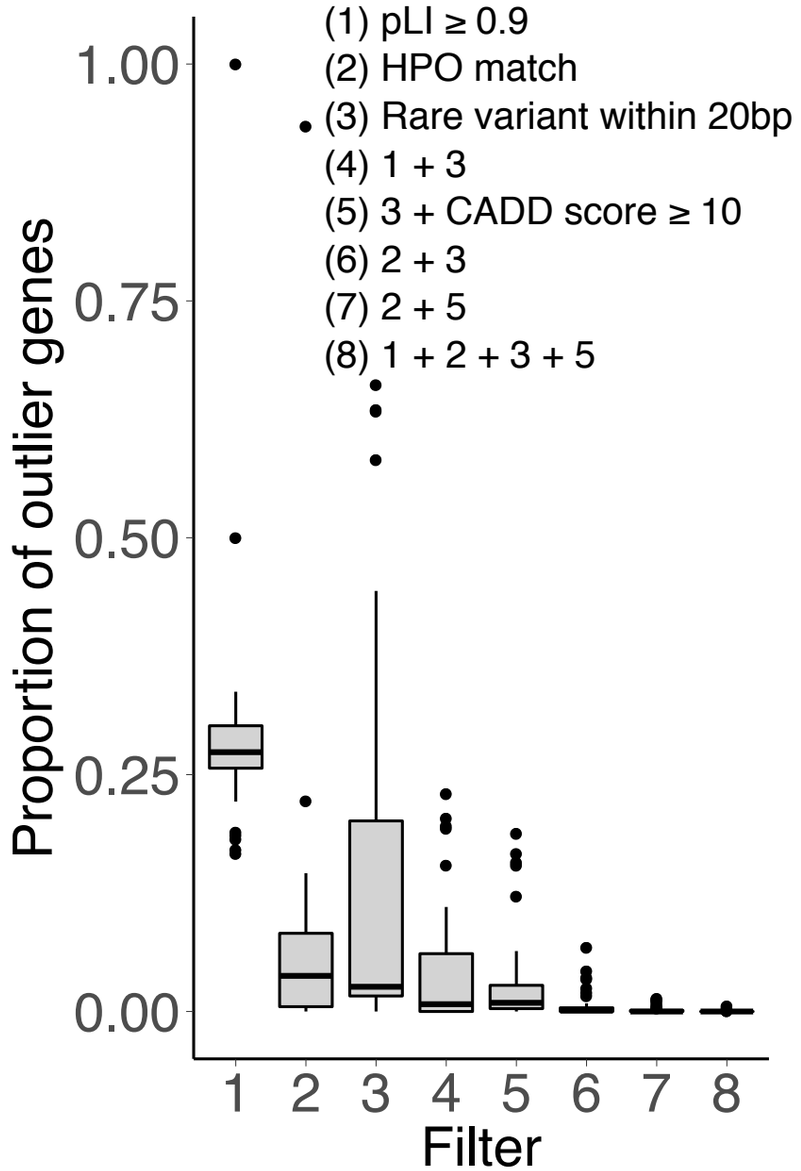
Two sisters exhibited **profound global developmental delay**, neonatal-onset **seizures**, acquired **microcephaly**, **hypotonia**, **G-/J-tube dependence**, and **progressive scoliosis** and had undergone a diagnostic odyssey including comprehensive metabolic evaluation, storage disorder enzymology, and genetic testing



We observed an average of 100 splicing outliers per sample (|Z-score| ≥2)



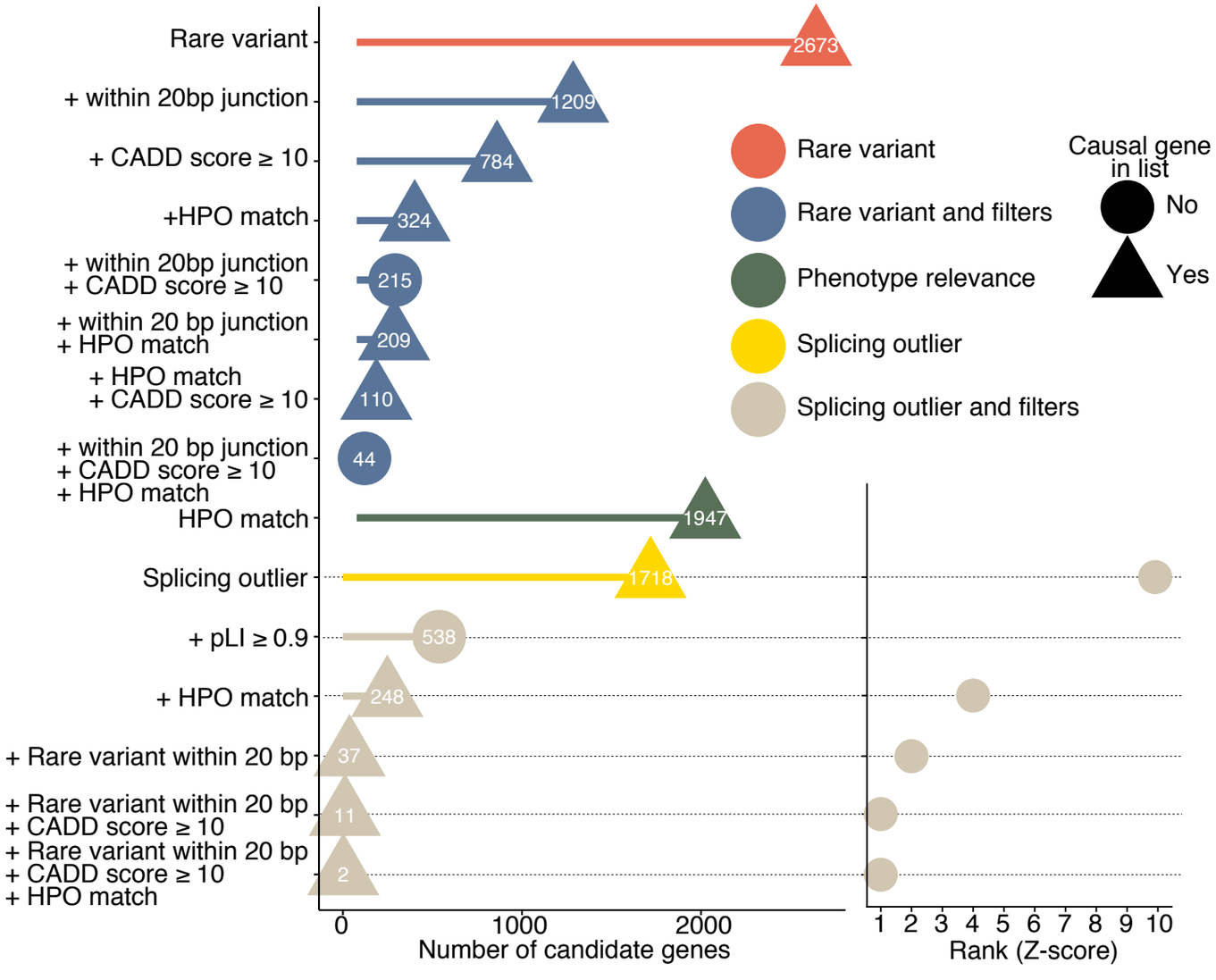
Filtering data for relevant criterion narrows down the number of potential candidates

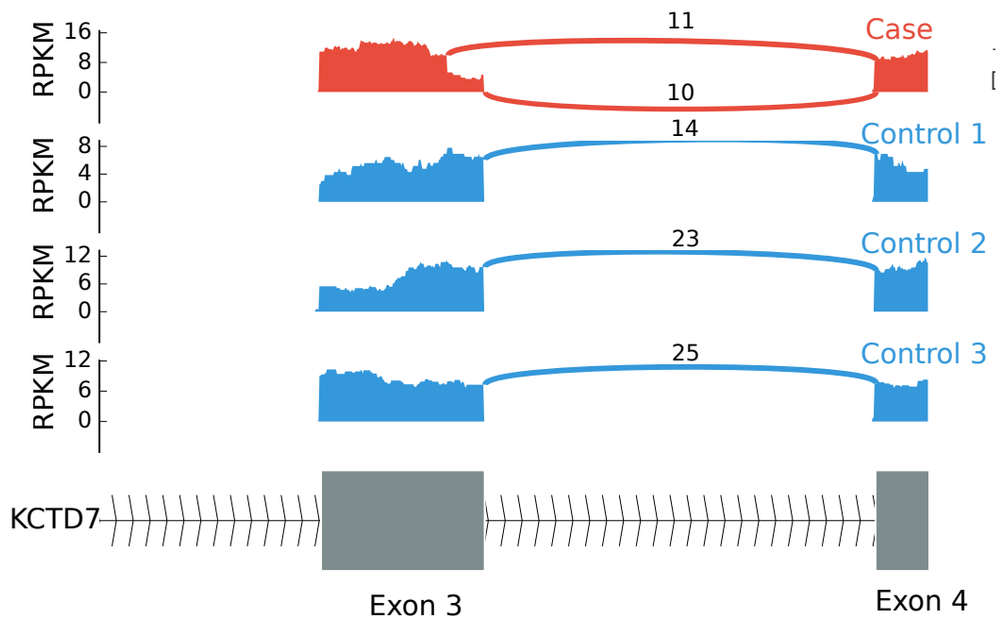


Limiting to genes r**elevant to the phenotype** and with a **deleterious rare variant** within **20 bp** of the splicing junction, we were able to narrow down to **only 0.05%** of potential candidate genes.

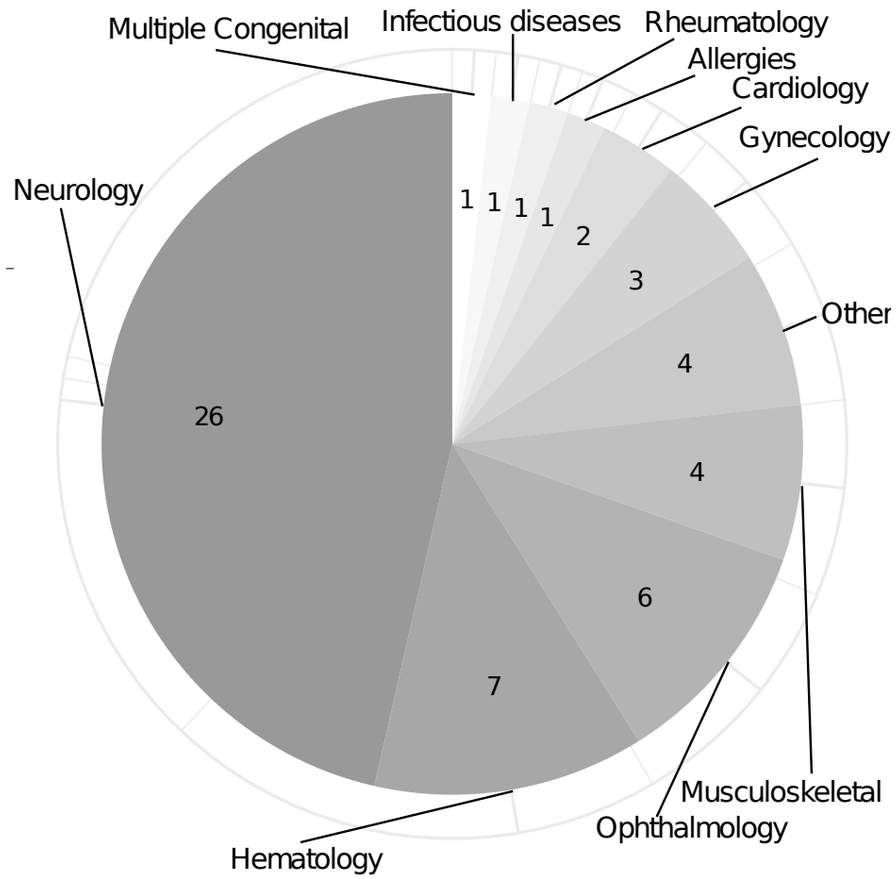
Identification of disease gene through splicing outlier detection: the *KCTD7* case

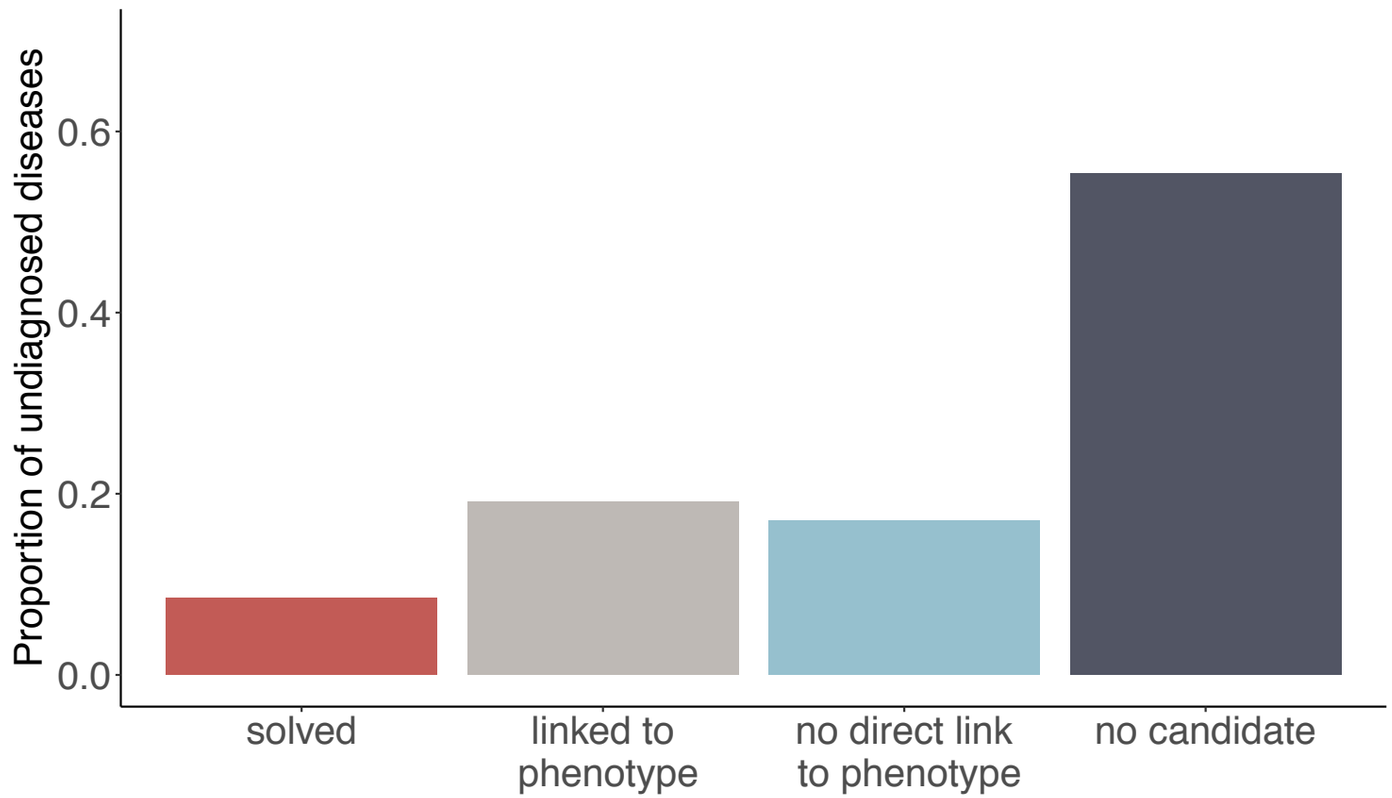
A 12 year old Hispanic female presented with **developmental regression** after typical development until age 18 months, manifesting with **loss of milestones** including head control, and speech. **Tremors** developed at 21 months; and **seizures** at 22 months. She also suffered from **occasional myoclonus**.





Across our cohort, we observed that RNA-seq yields a 8.5% diagnostic rate.





Exome sequencing has been estimated to lead to a diagnosis in 30 to 50% of rare Mendelian diseases