

Zaira Seferbekova^{1,2}, Dylan Lawless², Jacques Fellay²

¹Lomonosov Moscow State University, sef.zaira@gmail.com

²École Polytechnique Fédérale de Lausanne

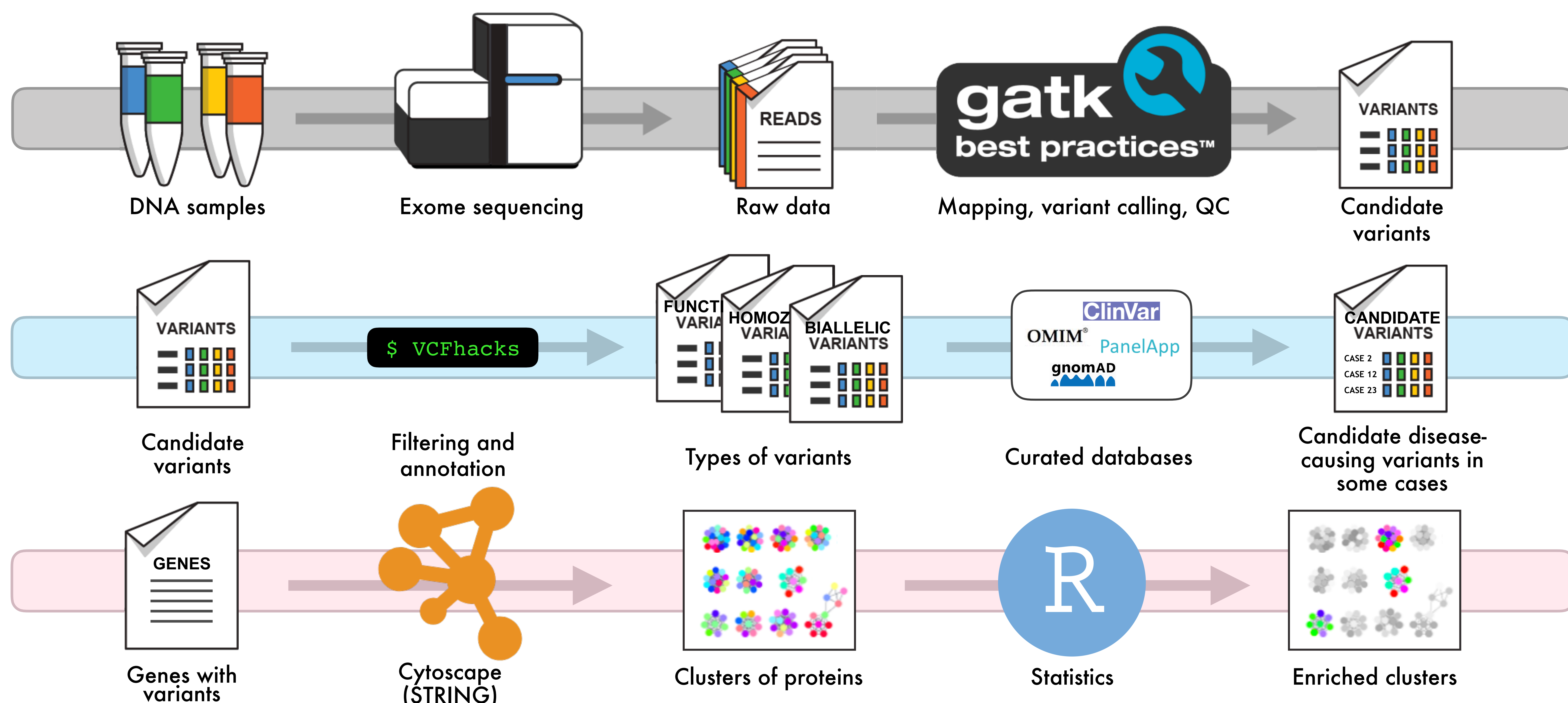
INTRODUCTION

Sepsis is a leading cause of childhood mortality worldwide. Rare genetic variants may cause a primary immunodeficiency and underlie fulminant sepsis. Such variants usually have low frequencies in the population due to purifying selection.

The aim of this study was to understand the genetic determinants that underlie disease using exome data from 270 patients.

CONCLUSIONS

- Candidate disease-causing variants that promote bacterial sepsis have been found for some cases.
- Pathway clustering and statistical testing have been developed and tested. This pipeline will also be used for the analysis of several other studies shortly.



METHODS

DNA samples from children admitted to paediatric intensive care due to sepsis were sequenced and raw exome data were analysed using GATK pipeline¹. GATK best practices provide step-by-step recommendations for performing variant discovery analysis in high-throughput sequencing data. This step produced a variants callset ready for downstream analysis.

Tailored analysis was used to identify cohort-specific rare variants. We used VCFhacks² for filtering and annotation of candidate variants. We filtered out common variants and technical artefacts, used population genetics frequencies, annotated with several curated databases of ontology and clinical pathogenesis, and predicted variants effect on protein function. This traditional approach identified top candidate variants.

Traditionally, a deficiency in statistical power exists for identifying variants of unknown significance. To overcome this obstacle we used a second, novel method, to cluster genes into functional pathways using Cytoscape³ and STRING⁴ database. To perform statistical testing and pathway-enrichment analysis we used R-studio.

RESULTS

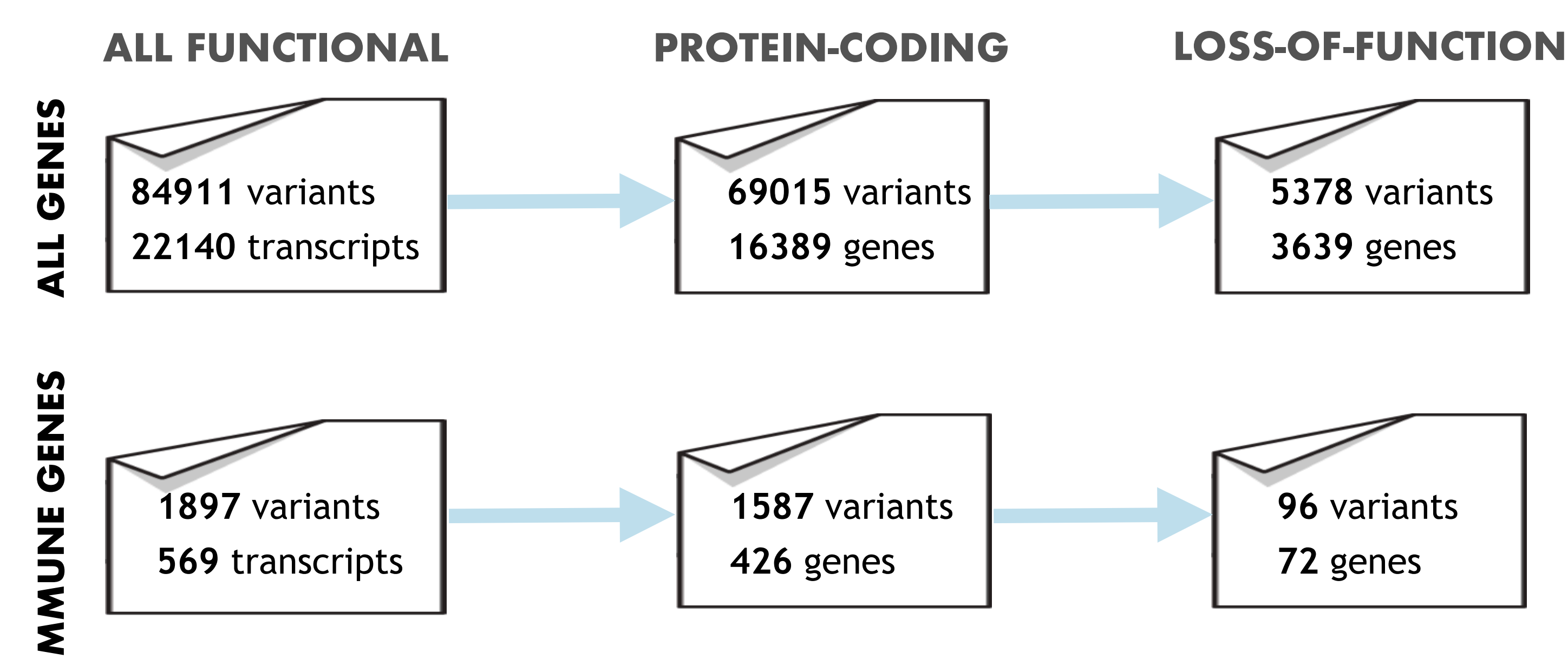


TABLE. Examples of identified pathogenic variants

GENE	REF/ALT	PROTEIN POS	POSSIBLE PHENOTYPE
TNFAIP3	T/A	544	autoinflammatory upregulation
	V/M	613	
NFKB1	E/K	600	common variable immunodeficiency (autosomal dominant)
	R/S	878	
TNFRSF13B	H/D	212	common variable immunodeficiency (compound heterozygosity)
	C/R	104	

single gene examples from wider cohort-level protein pathway variant enrichments

REFERENCES

1. **GATK**; A framework for variation discovery and genotyping using next-generation DNA sequencing data. DePristo M, et al. *Nature Genetics* 2011.
2. **VCFhacks**; David A. Parry. <https://github.com/gantzgraf/vcfhacks>.
3. **Cytoscape**: a software environment for integrated models of biomolecular interaction networks; Shannon P, et al. *Genome Research* 2003.
4. **STRING**: protein-protein interaction networks, with increased coverage and integration, Franceschini, A et al. *Nucleic Acids Research* 2013.