Rare variants in antiviral response genes drive severe viral respiratory infections in children

Dylan Lawless[1], Samira Asgari[1,2], Luregn J. Schlapbach[3,4,5], Jacques Fellay[1,2,6]

A robust and self-limiting immune response is required for clearance of viral respiratory infections. In rare cases, life-threatening infections may occur in previously healthy children.

To uncover this susceptibility to severe disease, we searched for rare genetic variants in 120 children requiring intensive care support upon infection by a respiratory virus.

We used exome sequencing followed by protein network analysis to catalog rare coding variants and cluster them by known physical and functional associations. We identified potentially causal variants in genes involved in proinflammatory response and viral nucleic acid detection, including *DDX58* and *IFIH1,* encoding RIG-I and MDA5, respectively.

Both proteins share a common mechanism of RNA recognition and signal repression. In the absence of viral infection, each is maintained in an autoinhibited state, where the effector domain (CARD) and the ATP-binding helicase homology domain are masked by the C-terminal repressor domain (CTD). Binding of viral dsRNA at the CTD relieves repression and results in a proinflammatory cascade.

Three loss-of-function variants in *IFIH1* were previously reported for this cohort. A further four patients had helicase / ATP-binding domain variants in *IFIH1*, and one patient with a rare CTD variant. Rare variants were found in *DDX58* affecting the RNA binding motif; one patient harbored a variant predicted to disrupt the ATP-binding helicase. In total we identified 10 rare variants in 15 patients and several further candidate genes.

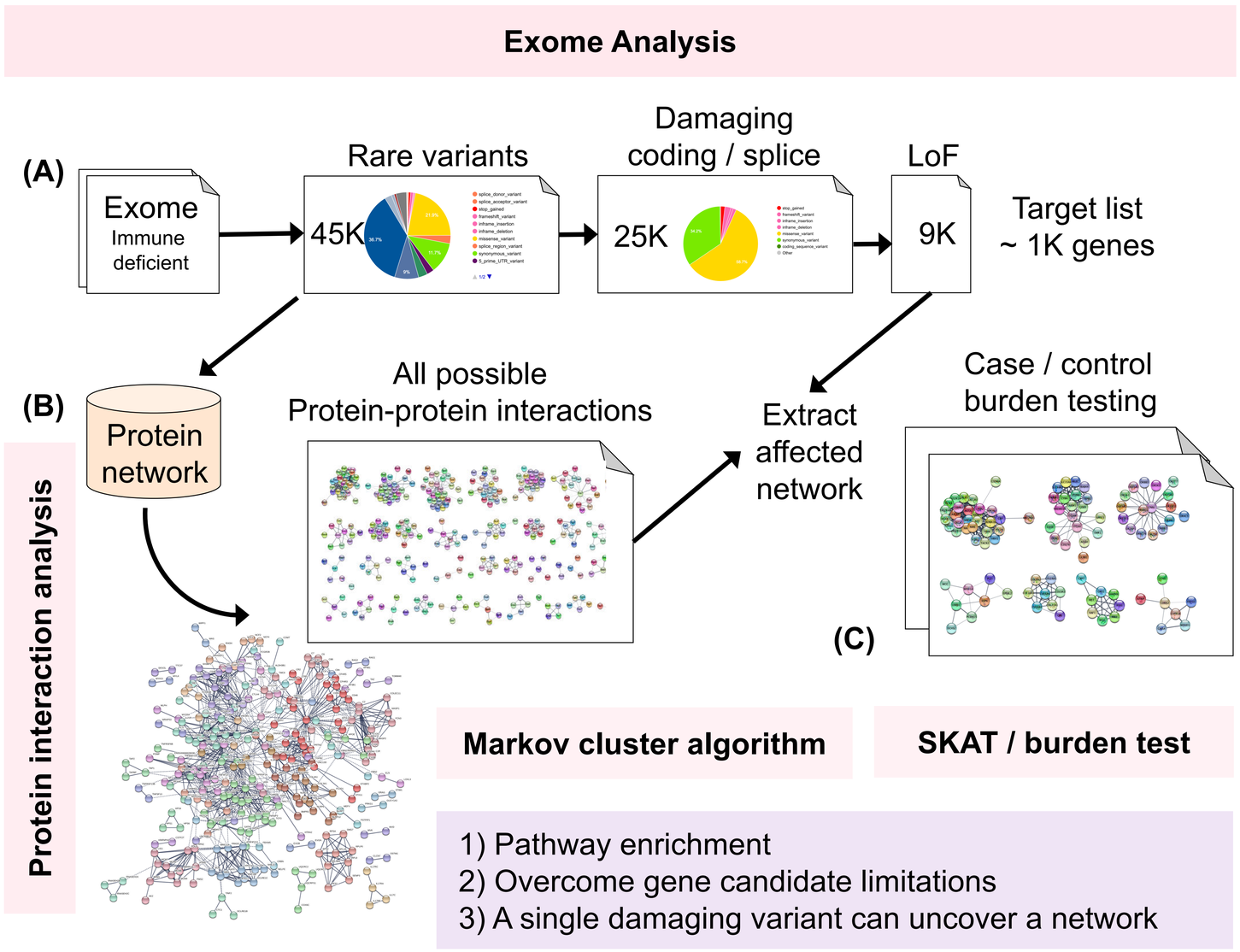
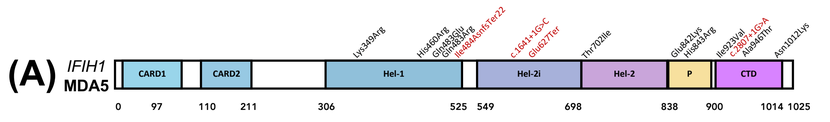


Figure 1. Deleterious rare variants uncovered damaged protein pathways in rare disease. A. GATK best practices were used for whole exome analysis[1] with joint genotyping for cases and controls; 240 in total. Custom filtering[2] extracted variants of high impact consequence (ostensibly loss of function (LoF)), present only in cohort cases. **B.** Genes harboring rare predicted LoF variants were grouped based on protein-protein interactions[3] using a Markov cluster algorithm[4]. **C.** Case-control testing was performed on each protein pathway cluster, as described in Figure 2.



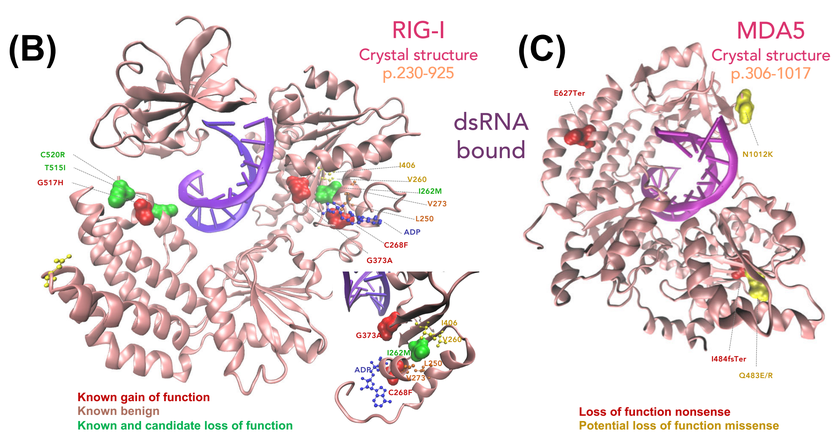


Figure 3. Viral recognition receptor rare variants in immune deficient patients. **A.** Gene structures of *DDX58* (RIG-I) and *IFIH1* (MDA5) listing variants found in immune deficient cases. **B.** RIG-I crystal structure. Both loss and gain of function variants are known in *DDX58*, RIG-I. Ile262Met is thought to directly affect the ATP binding site, nested between the sites of GoF and benign variants. Two further C-terminal variants are thought to interrupt the RNA binding face. **C.** MDA5 crystal structure. Nonsense variants were found in disease cases as well as missense variants affecting the Helicase ATP-binding and the RIG-like receptor C-terminal regulatory domain.

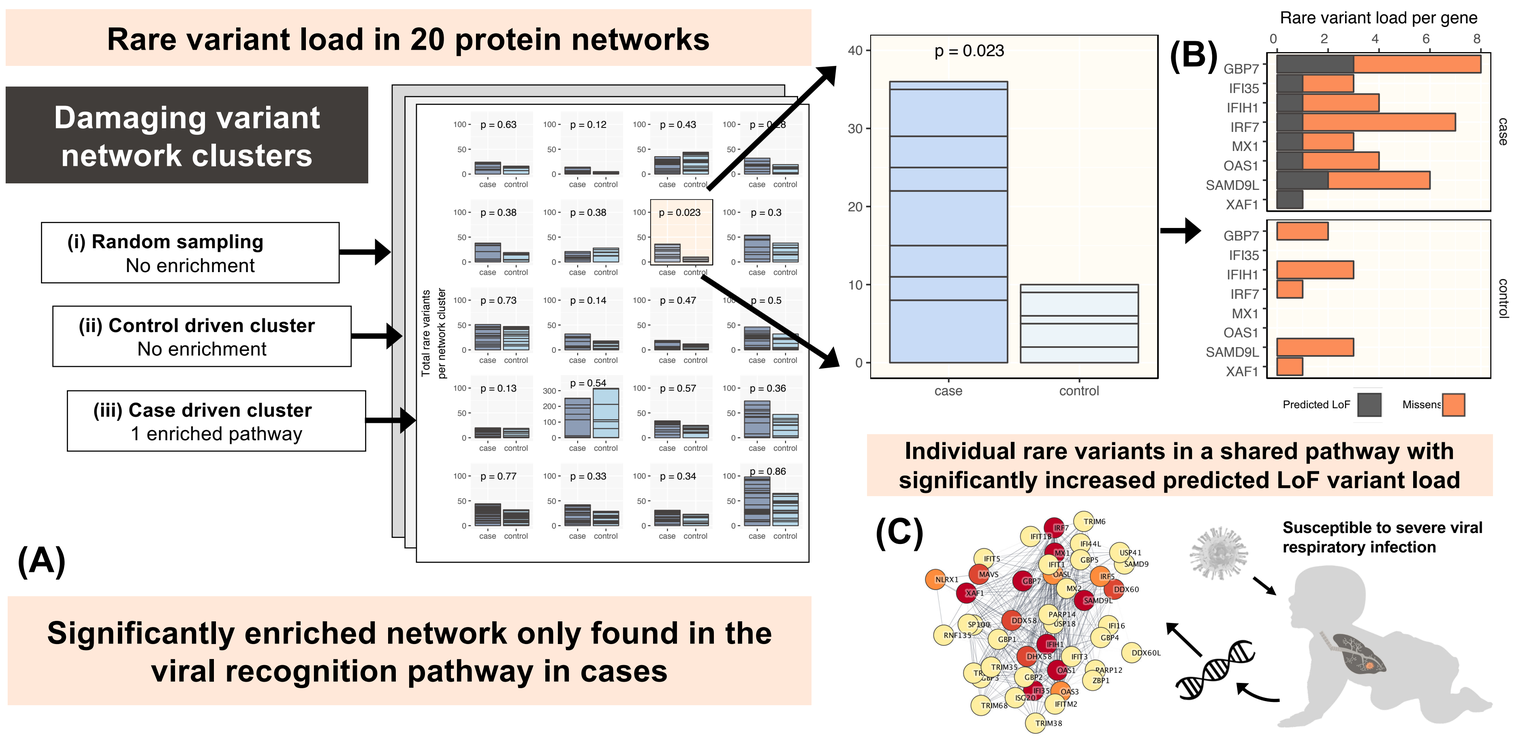


Figure 2. Damaging rare variants are enriched in the antiviral response pathway in immune deficient patients. **A.** Deleterious variant load per network was tested (i) using random sampling and comparing network clustering based on (ii) control versus cases, and (iii) case versus control. The antiviral response pathway was significantly enriched for potentially LoF variants in cases (p=0.023). **B.** The eight gene cluster harbored several other rare or novel missense variants. **C.** An extended antiviral immune pathway of 45 genes harboring rare variants, present only in cases, was uncovered after the significant deleterious load was flagged. Protein-protein interaction enrichment p-value <1.0e-16. Prioritised as: predicted LoF (red), missense top candidate (orange), and VUS (yellow).

Genomic analysis | **Functional organization** | **Statistical testing**

Translational application | **Clinical Interpretation** | **Custom annotation**

Summary

- We present a primary immunodeficiency resulting in extreme susceptibility to common respiratory RNA viruses, due to genetic variants in a common pathway that severely impairs viral recognition.
- A new method was developed for the unbiased detection of a protein network, driving disease, based on potential loss of function variants.

1. GATK; A framework for variation discovery and genotyping using next-generation DNA sequencing data. DePristo M, et al. Nat Genetics 2011
2. VCFhacks; David A. Parry, https://github.com/gantzgraf/vcfhacks
3. Franceschini, A et al, STRING: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res 2013.
4. Markov cluster algorithm; Stijn van Dongen, Graph Clustering by Flow Simulation. University of Utrecht, 2000.

1 Global Health Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, 1015, Switzerland; 2 Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland; 3 Paediatric Critical Care Research Group, Mater Research Institute, University of Queensland, Brisbane, QLD 4101, Australia; 4 Paediatric Intensive Care Unit, Lady Cilento Children's Hospital, Brisbane, QLD 4101, Australia; 5 Pediatric Intensive Care Unit, Department of Pediatrics, Inselspital, University Children's Hospital and University of Bern, 3010, Switzerland; 6 Precision Medicine Unit, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland. This study was funded by Swiss National Science Foundation Grant PP00P3_157529