**Supplementary information**

**De novo identification of complex multimorbid conditions by integration of gene regulation and protein interaction networks with genome-wide association studies**

Roan E. Zaied1, Tayaza Fadason1,2, Justin M. O’Sullivan1,2,3,4 \*

1 The Liggins Institute, The University of Auckland, Auckland, Private Bag 92019, New Zealand

2 The Maurice Wilkins Centre, The University of Auckland, Auckland, Private Bag 92019, New Zealand

3 Australian Parkinsons Mission, Garvan Institute of Medical Research, Sydney, New South Wales, 384 Victoria Street, Darlinghurst NSW 2010, Australia

4 MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton, SO16 6YD, United

Kingdom

\* To whom correspondence should be addressed: [justin.osullivan@auckland.ac.nz](mailto:justin.osullivan@auckland.ac.nz)

Chart

Description automatically generated

**Supplementary Figure 1. Protein coding eGenes are significantly more intolerant to loss of function compared to protein coding non-eGenes.** A)Number of eQTL-eGene interactions in GTEx v8 compared to spatial eQTL-eGene interactions seen in whole blood, classified by interaction type. B) Proportion of eQTL-eGene interactions across chromosomes. C) Proportion of non-eGenes and statistically significant eGenes expressed in whole blood (median TPM >0.1) classified by biotype. D) expression level and E) tolerance to loss of function of protein coding eGenes and non-eGenes expressed in whole blood. In both D and E, the mean is shown in red while the median is shown in black. NS, not significant, t test (\*\*\*\*P <1 × 10−4).

Graphical user interface

Description automatically generated with medium confidence

**Supplementary Figure 2. Tissue-specific spikes in the ratio of eQTL-eGene interactions to SNPs seen in chromosomes 7, 18 and 19.** Ratio of whole blood, adult brain, and left ventricle spatial eQTL-eGene interactions to SNPs across A) chromosome 7, B) chromosome 18 and C) chromosome 19.

Chart

Description automatically generated with medium confidence

**Supplementary Figure 3. Asthma eGenes are enriched for immune-related processes and lipid metabolism**. GO enrichment analysis of asthma eGenes having significant enrichment of A) biological processes B) molecular functions and C) reactome pathways.

A picture containing table

Description automatically generated

**Supplementary Figure 4. Blood GRN *de novo* identifies conditions that are multimorbid with ALL.** GWAS traits enriched in each of the four expanded PPIN neighbors of the ALL disease module. Circle size indicates the number of eQTLs, and circle color indicates statistical significance. ACPA; anti-citrullinated protein antibodies, OCB; oligoclonal band, EBNA-1; Antibodies against Epstein-Barr nuclear antigen 1, AQP4; Aquaporin 4, FEV1; forced expiratory volume in the first second, FVC; Forced vital capacity, COPD; chronic obstructive pulmonary disease.