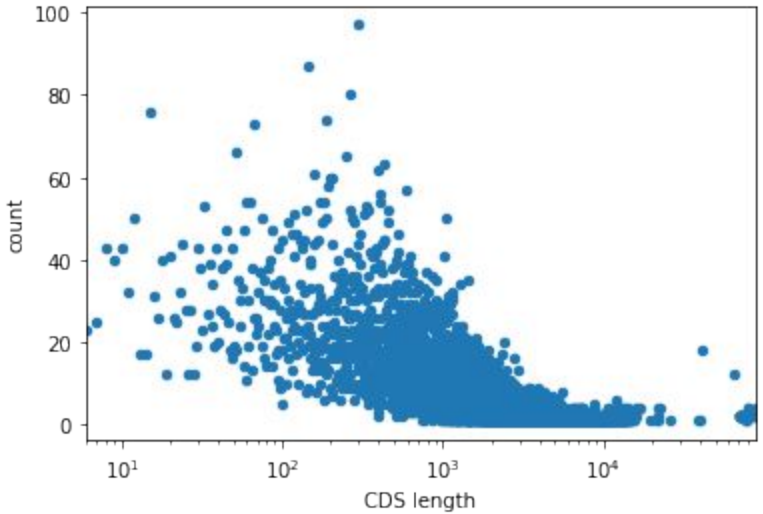
Genomic analysis for John M

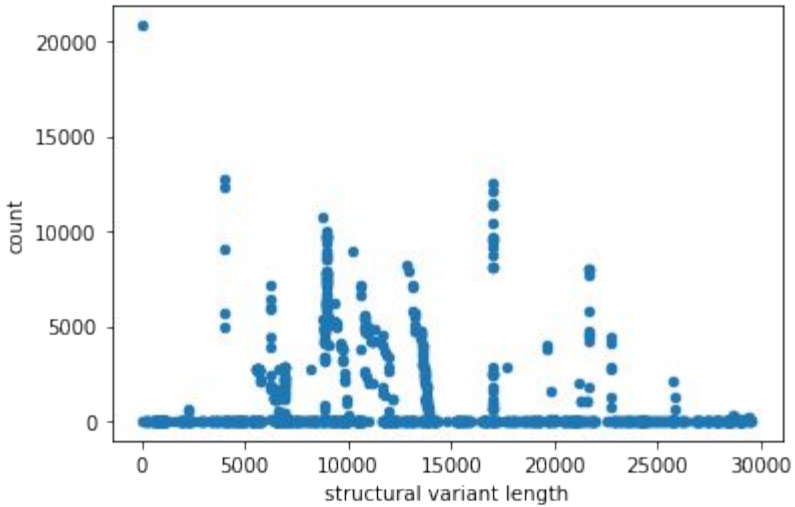
Start point: exome and structural variants (VCF)

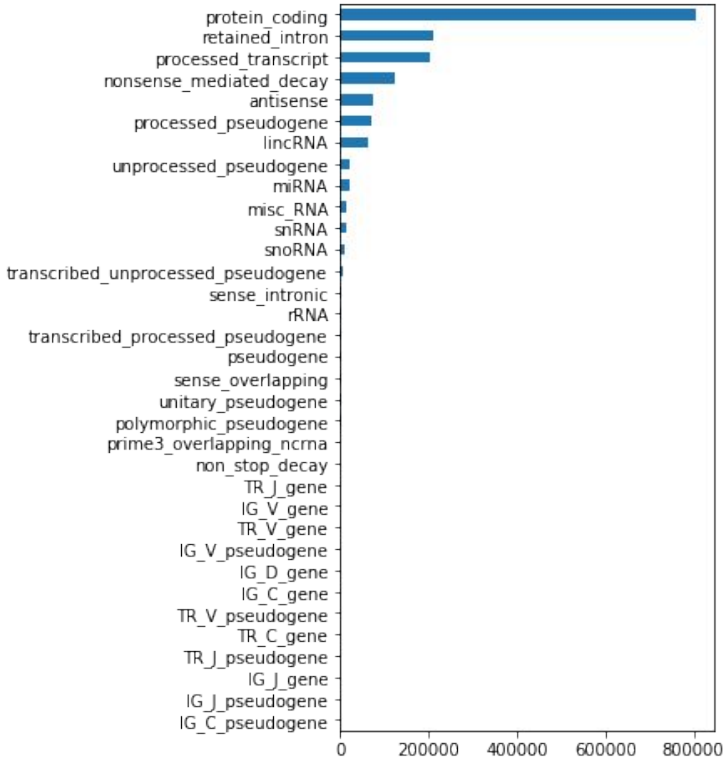
- Remove low quality variants and heterozygous loci

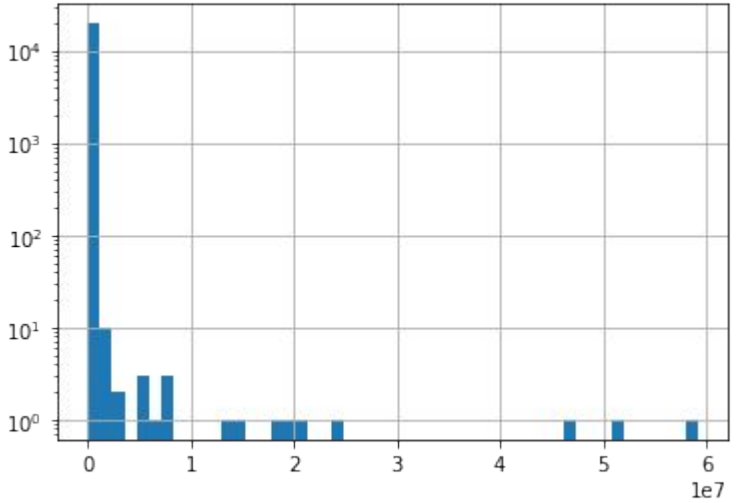
  ~ 2 M variants (1995031)
- Annotate using SnpEff
- Filter for genes highly expressed in GI tissues (714 genes)

  -> 35K variants (34650)
-


count
CDS length


count
structural variant length


protein_coding
retained_intron
processed_transcript
nonsense_mediated_decay
antisense
processed_pseudogene
lincRNA
unprocessed_pseudogene
miRNA
misc_RNA
snRNA
snoRNA
transcribed_unprocessed_pseudogene
sense_intronic
rRNA
transcribed_processed_pseudogene
pseudogene
sense_overlapping
unitary_pseudogene
polymorphic_pseudogene
prime3_overlapping_ncrna
non_stop_decay
TR_J_gene
IG_V_gene
TR_V_gene
IG_V_pseudogene
IG_D_gene
IG_C_gene
TR_V_pseudogene
TR_C_gene
TR_J_pseudogene
IG_J_gene
IG_J_pseudogene
IG_C_pseudogene
0
200000
400000
600000
800000


10^4
10^3
10^2
10^1
10^0
0
1
2
3
4
5
6
1e7

- Intersect with dbSNP: **24** variants
  - Non-benign: 3 variants -- Non-remarkable
- Intersect with OMIM database pathogenic genes: **58** variants:
  - HNF4A: Maturity onset diabetes of the young (MODY)
  - PAX4: MODY type IX, ketosis-prone diabetes
  - ALG14: Myasthenic syndrome, congenital
  - IL23R: IBD
  - MTTP deletion/nonsense-mediated decay: Abetalipoproteinemia / fat absorption disorder
  - SLC10A2 (3’UTR variant): Bile acid malabsorption
  - TREH: Trehalase deficiency
  -

- A4GNT (370 aa deletion, chr3):
    - Necessary for the synthesis of type III mucin which is specifically produced in the stomach, duodenum, and pancreatic duct (PubMed:10430883). May protect against inflammation-associated gastric adenocarcinomas.
- MUC13 (1240 aa deletion, chr3): IBD risk
- INSM1 (25810 nt deletion, chr29)
    - Promotes the generation and expansion of neuronal basal progenitor cells in the developing neocortex. Involved in the differentiation of endocrine cells of the developing anterior pituitary gland, of the pancreas and intestine, and of sympatho-adrenal cells in the peripheral nervous system. Promotes cell cycle signaling arrest and inhibition of cellular proliferation.

Future directions

Confirm mucin 3 and 13 deficiency in proteomics data