

Investigation of functional common and rare variant in cardiovascular disease

Currently genome wide association studies are targeted to common variants that are present in higher frequency in a population, which had lead into hypothesis of common variant common disease hypothesis ¹. However most of common variants identified so far are within the non-coding regions that are difficult to link with disease ². On the other end of spectrum, rare variants: less common in population , are less studied due to large sample requirement and higher cost of sequencing. However, recent studies have shown these genetic variants have higher effect on common diseases as well ³.

One of such method to study these variant using HiCap (variant of Hi-C) ⁴ methods, that identifies promoter-anchored interaction between variants that are thousand of bases apart from each other. Thus providing connectivity information from GWAS variants with potential genes.

In this current project, we will investigate role of both common and rare variant in one of common disease, cardiovascular diseases. Our group already carried out Hicap on large sample cohort thus having interaction data between putative promoters and enhancers.

The aim of this project would:

- Identify these variants as rare and common variants based on external databases such as (1000G project ⁵, SwedGen Variant frequency database ⁶).
- Annotate variants based on functional marks from ENCODE ⁷ database. Basically target common and rare variants with TFBSS (Transcription factor binding sites sequences) marks.
- From these common variant annotate variants with eQTL effect from GTEx ⁸ project on relevant cell types
- On the rare variants side, identify variants with functional effect on candidate genes

From this project,

I should be able to aptly use and manipulate all the large-scale databases from all the projects. Additionally, I should be programmatically manipulating all the analysis.

References:

1. Jonas, D.E., Wilt, T.J., Taylor, B.C., Wilkins, T.M., and Matchar, D.B. (2012). Chapter 11: challenges in and principles for conducting systematic reviews of genetic tests used as predictive indicators. *J Gen Intern Med* 27 *Suppl 1*, S83-93.
2. Saint Pierre, A., and Génin, E. (2014). How important are rare variants in common disease? *Brief Funct Genomics* 13, 353–361.
3. Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* 11, 415–425.
4. Sahlén, P., Abdullayev, I., Ramsköld, D., Matskova, L., Rilakovic, N., Lötstedt, B., Albert, T.J., Lundeberg, J., and Sandberg, R. (2015). Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biol* 16, 156.
5. 1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* 491, 56–65.
6. Ameer, A., Dahlberg, J., Olason, P., Vezzi, F., Karlsson, R., Lundin, P., Che, H., Thutkawkorapin, J., Kusalananda Kahari, A., Dahlberg, M., et al. (2016). SweGen: A whole-genome map of genetic variability in a cross-section of the Swedish population. *BioRxiv*.
7. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
8. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348, 648–660.