COMPLEX TRAIT ALLELES ARE ENRICHED FOR CELL-SPECIFIC CHROMATIN MARKS

Cynthia Sandor[1-4], Gosia Trynka[1-4], Han Xu[5], Barbara E Stranger[1,4],

X Shirley Liu[5], and Soumya Raychaudhuri[1-4]

1. Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
2. Division of Rheumatology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
3. Partners Center for Personalized Genetic Medicine, Boston, MA, USA.
4. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA.
5. Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard School of Public Health, Boston, Boston, MA, USA

Complex trait alleles might act by altering gene regulatory regions, such as promoters and enhancers. In this case, chromatin marks highlighting active regulatory regions should overlap alleles in the specific cell-types most relevant to the trait. However, there are many potential marks, and those that are most useful in defining cell-types and fine-mapping common variants have not been defined. We hypothesized that those marks that have greatest utility must be “phenotypically cell-specific”, i.e. the overlap of alleles with marks is restricted by the phenotype to specific cell-types. We examined 15 chromatin marks across multiple cell-types and 510 independent variants associated with 31 phenotypes. Of 15 marks tested, H3K4me3 showed the highest phenotypic cell-specificity ($p<10^{6}$). We observed only those marks highlighting active gene regulation, such as H3K9ac, were phenotypically cell-specific. Then, we examined H3K4me3 marks to identify key cell-types for four specific phenotypes. We found that cell-specific H3K4me3 peaks most significantly overlapped 37 plasma low-density lipoprotein concentration alleles within the liver ($p<10^{-6}$), 40 rheumatoid arthritis alleles within CD4+ memory T-cells ($p=1x10^{-4}$), 32 body mass index alleles within pancreatic islet cells ($p=1x10^{-4}$) and 11 neuropsychiatric disease alleles within the mid frontal lobe ($p=0.02$). We illustrate how these results can be used to fine-map associated variants to causal variants. This study suggests that common variants implicate regions involved in cell-specific gene regulation, and that a limited number of high quality chromatin mark assays can be comprehensively applied to many cell-types to identify critical cell-types and specific regulatory elements in an unbiased fashion.