Population Genetics, Genetic Epidemiology, Statistics, Bioinformatics

AB069. Proficiency of data interpretation: identification of signaling single nucleotides polymorphism for coronary artery disease

Asma Cheema[1], Samantha Rosenthal[2], M. Ilyas Kamboh[2]

[1]Department of Pathology, University Medical & Dental College, The University of Faisalabad, Faisalabad, Pakistan; [2]Department of Human Genetics, University of Pittsburgh, Pittsburgh, USA

Background: Coronary artery disease (CAD) is a complex disorder involving both genetic and non-genetic factors. Genome wide association studies (GWAS) have identified hundreds of single nucleotides polymorphism (SNPs) tagging over >40 CAD risk loci. We hypothesized that some non-coding variants might directly regulate the gene expression rather than tagging a nearby locus.

Methods: A total of 58 SNPs within 54 CAD loci was selected. Of these SNPs, 52 had accepted genome wide significant threshold ($P<5\times10^{-8}$) .We used SNAP webportal to identify SNPs in strong LD (r2 >0.80) with our SNPs of interest. We used RegulomeDB to identify potentially functional SNPs. RegulomeDB is a database that scores SNP's functionality based upon their existence in a DNAase hypersensitive site or transcription factor binding site. RegulomeDB grades the SNPs from 1 to 6. The SNP showing the strongest evidence of being regulatory is given the score of 1 and SNP demonstrating the least evidence of being functional is marked as 6.

Results: SNAP webportal gave a total of 1,200 SNPs. Among the 1,200 SNPs [58 were genome-wide significant ($P<5\times10^{-8}$) and the remaining were in linkage disequilibrium (r2 >0.80) with 58 SNPs]. Of these, 97 SNPs were predicted to have regulatory function with RegulomeDB score of <3, and only 8 of them were genome-wide significant (LIPA/rs2246833 = RegulomeDB score 1b; ZC3HC1/rs11556924, CYPA1/CNNM2/NT5C2/rs12413409, APOE-APOC1/rs2075650 and UBE2Z/rs46522 RegulomeDB score = 1f; ZNF259-APOA5-APOA1/rs964184, SMG6/rs2281727, and COL4A1-COL4A2/rs4773144 = RegulomeDB score 2b).

Conclusions: This study supports the hypothesis that some of the non-coding variants are true signals that regulate the gene expression at the transcriptional level. Our study indicates that RegulomeDB is a useful database to examine the large number of genetic variants and it may differentiate between true or tagged SNPs by defining the functional role of variants, particularly for variants in complex disease such as CAD.

Keywords: Coronary artery disease (CAD); genome wide association studies (GWAS); regulatory regions; RegulomeDB; SNPs

doi: 10.21037/atm.2017.s069

Cite this abstract as: Cheema A, Rosenthal S, Kamboh MI. Proficiency of data interpretation: identification of signaling single nucleotides polymorphism for coronary artery disease. Ann Transl Med 2017;5(Suppl 2):AB069. doi: 10.21037/atm.2017.s069

