International Conference and Exhibition on
Obesity & Weight Management
December 3-5, 2012 DoubleTree by Hilton Philadelphia, USA

Replication and prioritization of biological pathways from genome-wide association analysis of extreme obesity

Sujoy Ghosh

North Carolina Central University, USA

Abstract

Pathway analysis addresses the missing heritability problem of SNP-centered GWAS data analysis by hypothesizing that a sufficient number of non-overlapping, non-concurrent disruptions in biological pathways, even with modest fractional contributions from single gene-variants, may be adequate to lead to a disorder. This approach is particularly relevant for disorders such as polygenic obesity where the contributions from individual SNP variants are typically quite low. However, a challenge faced by pathway analysis is that of replicating and prioritizing statistically associated pathways for downstream functional validation studies. To address this issue, we have developed a dual strategy of statistical replication and bioinformatic analysis, and applied this to identify and prioritize pathways based on GWA studies of extreme obesity. We employed 2 pathway analysis tools (iGSEA4GWAS and GSA-SNP) on an imputed 'Discovery' (985 cases/869 controls, BMI 43.1±8.7 and 20.3±1.84, respectively) and 'Replication' (540 cases/520 controls, BMI 49.4±8.8 and 20.7±1.8, respectively) cohort. Sixty-two (iGSEA4GWAS) and 22 (GSA-SNP) KEGG pathways from the Discovery cohort ($p<0.05$) were analyzed in the Replication cohort. A total of 19 pathways were replicated ($p<0.05$ level) between the 2 tools. The 19 replicated pathways were subsequently analyzed by the SNPNEXUS algorithm to calculate a 'SNP-burden' for each pathway, based on the estimated functional effects of the SNPs contained in them. A total of 16546 SNPs were analyzed. SNPs were scored for their location on chromosomal regions containing CPG islands, insertion-deletion regions, copy number polymorphisms, inversions, transcription factor binding sites, 3'- and 5'-UTRs, as well as amino-acid substitution effects on protein function. Finally, pathways were ranked by their predicted SNP-burden after normalization for the number of SNPs in each pathway. Two-way hierarchical clustering of pathway ranks and functional SNP categories identified the 'oxidative phosphorylation', 'purine metabolism' and 'adipocytokine signaling' pathways as closely clustered with high SNP-burden across several functional categories. In conclusion, a combination of independent validation and bioinformatic analysis allows prioritization of GWAS pathway analysis to aid functional validation studies.