Parent of origin effects in obesity in the Sorbs

Xuanshi Liu

The existence of genetic factors in obesity is well supported by the tremendously successful genome wide association studies (GWAS). However, the genetic loci identified by GWAS can only marginally describe the heritability of body mass index (BMI). One possible reason is the missing family information in GWAS design.

Parent of origin effects (POE) which could explain part of heritability in obesity were studied. The candidate gene *FTO* (fat mass and obesity associated) was selected from the Sorbs, a self-contained population from eastern Germany. We hypothesized there were variants within robustly replicated *FTO* that conferred different risk of obesity depending on the parental origins of alleles. Genotypes from 525 Sorbs individuals were phased and parental origins were assigned to the alleles by using software AlphaImpute. Subsequently, standard association tests and parent-of-origin specific association tests were applied to 22 SNPs (single nuclear polymorphism) within *FTO* introns 1 to 3. Indications of parental origins were identified at several SNPs. To further support our findings, parental asymmetry tests in 705 German childhood obesity trios were performed and revealed further indications for POE effects.

In order to identify novel functional SNPs/regions in genome wide that contribute to obesity, GWAS considering POE (POE-GWAS) were performed in the Sorbs. We identified SNPs in linkage disequilibrium (LD), that is, correlated SNPs, with POE-GWAS top hits. Several bioinformatics tools and databases were then applied to explore the potentially functional roles on the correlated SNPs. For example, correlated SNP rs1204880 locates within the promoter of *PADI6* (peptidyl arginine deiminase, type VI) and *PADI6* was not described in obesity related studies earlier. Transcription factors SF1 (steroidogenic factor 1) and LRH1 (liver receptor homolog-1), which were showed to be involved in obesity related pathways in mouse studies, putatively bind at rs1204880. Our study provided a workflow on finding novel variants.

Additionally, incorporating epigenetic data with bioinformatics tools in the analyses could help to identify potential functional relevant variants playing a role in the etiology of obesity.

Finally, the Sorbs-specific association signal in *FTO* within intron 3 was studied for possible signatures of evolutionary selection. PAML (phylogenetic analysis by maximum likelihood) analyses revealed strong conservation in the coding region of *FTO*. Population genetic measures identified signatures of balancing selection at intron 3 of *FTO*. The data support the hypothesis that *FTO* associated with obesity may have been under evolutionary selective pressure. We observed several alterations in transcription factor binding sites, e.g. TCF3 (transcription factor 3) binding site was introduced by the rs17818902 minor allele. This suggested the potential function in intron 3 of *FTO*.

In conclusion, indications for POE in *FTO* were identified in Sorbs which were further supported by analyses from a German childhood obesity trios. A subsequent POE-GWAS proposed several potentially functional relevant regions/SNPs. Finally, a Sorbs specific association signal in intron 3 of *FTO* indicated balancing selection. Taken together, these data expand the current knowledge on genetic mechanisms contributing to obesity especially in the *FTO* gene and provide several novel candidate genetic regions for further exploration.