審査の結果の要旨

氏名　劉暁渓

To better understand the underlying genetic architecture of autism spectrum disorder (ASD), especially in East Asian populations, a two-stage genome wide association study (GWAS) and genome wide copy number variation (CNV) study were performed using 2780 individuals. The following results were obtained.

(1) In the discovery stage of GWAS, after stringent sample quality control procedures, a total of 156 cases and 427 controls were used for case-control analysis and 156 cases and 295 parents were available for family based association test (FBAT) analysis. After SNP quality control, 588,073 SNPs remained for further analysis. The estimated inflation factor (λ) was 1.0069 and 1.0186 for FBAT and for case-control respectively, which indicated negligible population stratification. A number of nominal associations were observed within or near genes including *GPC6, NTN4, JARID2, AIM2, ROR2, CNTN4*. Of these genes, *JARID2* and *CNTN4* are known ASD candidate genes. The other genes might be novel ASD candidate genes. No single SNP was found to be associated with ASD at a genome-wide significance level ($P=5\times10^{-8}$) in both FBAT and case-control analysis.

(2) After the integration of case-control and FBAT, a combined *P* value was re-assigned to each SNP. Based on the ranking of the combined *P* value and other criteria, 86 SNPs were chosen for replication in the second stage in independent Japanese ASD trios (n=205) and Taiwanese ASD trios (n=410). No SNP was replicated either in the Japanese or the Taiwanese trio set. The results indicated that ASD has a very complicated genetic architecture. The lack of association was also likely to be caused by insufficient sample size and limited probe density of the genotyping platform.

(3) In the CNV analysis, three algorithms including PennCNV, QuantiSNP and Birdsuite were utilized for CNV calling to increase detection accuracy and to reduce false positive calls. A total of 27,720 CNVs were detected by all three algorithms in all subjects including ASD family samples and controls. Subsequently in a total of 5,653 CNVs identified in probands, 311 putative CNVs were suggested as *de novo* by the trio calling algorithm in PennCNV software (5.5%). Ten CNVs were not present in the Database of Genomic Variants (DGV hg18 ver.10) and were regarded as rare *de novo* CNVs. All 10 *de novo* CNVs were validated by SYBR Green-based quantitative PCR assay. Eight CNVs out of 10 from 8 ASD probands were confirmed (80%). Besides the well-known 15q11.2-13.1 duplication, other 7 regions were identified and experimentally confirmed, providing novel clues to the pathogenesis of ASD. Two novel genes *Gene A* and *Gene B* were indentified in the *de novo* loci and serve as novel ASD candidate genes for further functional study.

(4) Eight inherited CNVs (5 duplication and 3 deletions) in ASD probands were found to overlap with known ASD loci or ASD candidate genes. The maternal transmissions were observed three times more

frequently than paternal transmissions (6:2), which suggests that females might be protected from such genetic risks, in a way unknown yet. Of these CNVs, 16p11.2, *NLGN4X* and *MCPH1* have already been well-established to be associated with ASD. 22q11.21 and 17p12 have been reported in a variety of neuropsychiatric disorders including ASD. *PRKCB1* and *SLC25A12* were both ASD candidate genes but no CNVs affecting these two genes have previously been reported. Of all the above CNVs, *NLGN4X* is of particular interest. The deletion was transmitted from unaffected mother to the male proband, which is in an X-linked recessive inheritance mode. To confirm this deletion, PCR was carried out and *NLGN4X* deletion was confirmed in the proband, which indicated that a female could be a risk carrier of ASD. Collectively, 19 out of 158 patients were found to carry a potential ASD susceptible CNV (12%). The results confirmed the significant contribution of CNV in the etiology of ASD.

In conclusion, comprehensive analysis was conducted in this study to understand the genetic underpinning for ASD. The study provided a detailed genetic landscape of ASD in East Asian populations for the first time. The results from GWAS in this study, combined with previous GWAS findings, indicate that common variants captured by the current genotyping platform may play a more subtle role in the etiology of ASD than expected. In contrast, CNVs were found with a more critical role in the pathogenesis of ASD. The data generated in this study provide a solid basis for future study and contributes important new data to the ASD scientific research community. Considering the importance and novelty of the work, it is worthy of the award of Doctor of Philosophy to the candidate.