主 論 文

MGST2 and *WNT2* are candidate genes for comitant strabismus susceptibility in Japanese patients
(斜視発症に関連する遺伝子候補 *MGST2* と *WNT2* を発見)

Introduction

Strabismus is a common condition with misalignment between two eyes that may lead to decrease of visual acuity, lack of binocularity, and diplopia. It is caused by heterogeneous environmental and genetic risk factors. Our previous research has identified new chromosomal susceptibility loci in 4q28.3 and 7q31.2 regions for comitant strabismus in Japanese families. We conducted a verification study by linkage analysis to narrow the chromosomal loci down to a single gene.

Materials and Methods

SNP Selection and Typing

From JSNP database for Japanese and U.S. National Center for Biotechnology Information (NCBI) databases, 24 rsSNPs and 233 rsSNPs were chosen from the 4q28.3 and 7q31.2 regions, respectively, and were typed in 108 affected subjects and 96 unaffected subjects of 58 families with primary and non-syndromic comitant strabismus including both esotropia and exotropia.

Genomic DNA that was isolated from peripheral blood leukocytes was amplified by multiplex polymerase chain reaction (PCR). The nucleic acids were detected by MassARRAY system, which is a high-throughput matrix-assisted laser desorption ionization and time-of-flight mass spectrometry (MALDI-TOF MS). We then proceeded to quality controls of SNPs and samples.

Statistical Methods

We firstly conducted Hardy-Weinberg equilibrium (HWE) and principal component analysis (PCA) by the genome-wide complex trait analysis (GCTA) program. Then three major analytical methods were used: Family-based association study: transmission disequilibrium test (TDT) and TDT allowing for errors (TDTae); and linkage analysis under dominant and recessive inheritance. Transmission disequilibrium test (TDT) is a test for association in the presence of linkage for a case-parent trio. *Plink* program version 1.9 was run to detect genotypes which violated the Mendelian rules in TDT analysis. In contrast, transmission disequilibrium test allowing for errors (TDTae) is an implementation of TDT which allows errors to be present in estimating their rates in the course of analysis by TDTae program. In the process of running the TDTae program, errors were estimated in the background of any one of a number of error models. We considered the two most reasonable and economical error models, which require few parameters: DSB (Douglas-Skol-Boehnke) allows for genotype errors and GHLO (Gordon-Heath-Liu-Ott) allows for allele errors. Under the two error models, we run TDTae program for dominant (d), recessive (r) and multiplicative (m) inheritance. Furthermore, we defined linkage

disequilibrium (LD) blocks using Haploview 4.2 on chromosome 4q28.3 and 7q31.2. Linkage analysis estimates recombination fractions between a putative disease locus and marker loci, and the results were output as LOD (logarithm of odds) scores. The Pseudomarker program estimates allele frequencies by maximum likelihood, separately under linkage and no linkage, which makes the results virtually independent of allele frequencies.

Results

Quality control of SNPs and samples

204 individuals and 237 SNPs remained after the quality control of samples and SNPs. And 229 SNPs were in HWE. PCA analysis showed that there was no population substructure among families.

TDT and TDTae and Linkage analysis

The SNPs with significant P values inTDT andTDTae were located solely at the gene, microsomal glutathione S-transferase 2 (*MGST2*), on chromosome 4q28.3 locus. In contrast, significant SNPs were dispersed in a few genes, containing wingless-type MMTV integration site family member 2 (*WNT2*), on chromosome 7q31.2 locus. The distribution of significant SNPs on the 7q31.2 locus showed that only the ST7 to *WNT2* region in the same big haplotype block contained significant SNPs for all three methods of linkage analysis.

The significant SNPs in the 4q28.3 locus were related to *MGST2* transcription in the search for expression quantitative trait locus (eQTL) in the Human Genetic Variation Database (HGVD) which displays the Japanese genetic variations and the association between the variations and transcription levels of genes.

Discussion

We firstly tried to use a method that did not depend on kinship, such as association study just adjusted by family. However, the false discovery rate (FDR) was too high to reduce the power in conducting multiple comparisons among SNPs, and therefore, we turned in the present study to focus on methods for linkage analysis.

When a few families or SNPs seem to contribute to the majority of errors, it is best to delete these families or SNPs firstly and then to carry out TDTae as the Pseudomarker program requires strict error-free data with no Mendelian errors. In contrast, TDT by *Plink* would handle errors by ignoring the offending genotypes.

In the present study, we clearly demonstrated that *MGST2* is a candidate for the chromosomal 4q28.3 locus. As for the 7q31.2 locus, in contrast, the results of different kinds of statistical analyses could not narrow the locus to a single gene. Under the circumstances, the distribution of significant SNPs in the locus showed that only the *ST7* to *WNT2* region contained significant SNPs for all three methods of linkage analysis. In the 7q31.2 locus, *ST7* is indeed in the same big haplotype block with *WNT2*.

Primary and non-syndromic comitant strabismus contains several different clinical entities, while patients with the same clinical entity or clinical diagnosis show varying degrees of manifestations. Under the circumstances, one way to define comitant

strabismus is as a disease with abnormal binocular vision, namely abnormalities in simultaneous perception, fusion and stereopsis. In the Japanese population, exotropia is more prevalent than esotropia. And there are indeed families which show mixed phenotypes of exotropia and esotropia, as observed in this study.

The large number of Mendelian errors seems to be the limitation of this study. The original data sets with Mendelian errors were used in TDT analyses. In contrast, the error-free data sets were used in TDTae and linkage analyses under dominant and recessive inheritance. The common use of the original data sets should have underlain the more consistent results whereas sharing of the same data sets would not necessarily mean that the data applied actually in analysis are the same. Or rather, the difference is merely the methods to handle the errors prior to software application or in the process by ignoring a single cell or deleting the whole series. In the present study, permutation tests were done to check the robustness of the results. Therefore, the results should be affected mainly by the methods, and would not be decided by the number of Mendelian errors.

Both *MGST2* and *WNT2* are known to be expressed in the brain and likely to be involved in the development of comitant strabismus. It should be noted that there is a limitation in applying the eQTL to the present study since the eQTL data have been obtained in analyses of blood cells.

Different analytical methods shed light on the data from different angles, therefore it is useful to apply more than one type of analysis. Strict Mendelian application as in this study, might not be appropriate in multifactorial disorders such as comitant strabismus, but would certainly provide a step to get guidance for detecting genetic risks of the disease. Since the proof of a responsible gene in a multifactorial disorder is difficult to be obtained in animal experiments, a different approach in patients, such as whole exome sequencing, would provide support for the present results of SNP typing. Further functional studies are necessary to clarify the mechanisms of the two genes on the susceptibility of comitant strabismus.

Conclusion

This study suggests that *MGST2* and *WNT2* are potential candidates for comitant strabismus in Japanese population.