npg

COMMENTARY

Association studies getting broader: A commentary on A polymorphism of the *POLG2* gene is genetically associated with the invasiveness of urinary bladder cancer in Japanese males

Masakazu Nakano and Kei Tashiro

In this issue of the *Journal of Human Genetics*, Ratanajaraya et al.[1] present a case–control association study to identify genes associated with the 'invasiveness' of urinary bladder cancer (UBC). The authors focused on assessing the genes related to DNA repair and metabolic pathways, and have successfully discovered a variant on the *POLG2* gene associated with Japanese male UBC patients. The variant could provide supportive information for the pathological classification of UBC as a genetic marker; for example, to predict the prognosis after the therapeutic intervention against UBC. From the point of view of studies in the field of human genetics, whether the study is based on a genome-wide or a candidate gene analysis, this study seems to be following the current trend of the recent association studies that have arisen from the big wave of genome-wide association studies (GWAS); the case samples have been subdivided into smaller categories in an attempt to identify variants determining clinical entities, malignancy, progression, prognosis, and so on, that underlie a single disease.

Technological breakthroughs in DNA microarray experiments, together with the enrichment of the public database derived from the International HapMap Project,[2] permitted us to perform GWAS to discover common variants associated with complex traits. Over the past few years, GWAS have

allowed thousands of common variants to be identified from a variety of common diseases. In fact, some of the common variants on the genes with a large effect size (odds ratio, >2.0), as in the case of age-related macular degeneration, have extensively been characterized for not only revealing the pathogenesis of the disease[3] but also applying to the diagnosis of the disease onset.[4] However, most of the studies, including our GWAS of primary open-angle glaucoma,[5] have failed to link the common variants to the specific gene(s). This is due to the fact that the variants, and even the variants in a linkage disequilibrium, were discovered from non-coding 'gene desert' loci with a small effect size (odds ratio, 1.2–1.5). The variants might be on the regulatory sequence for controlling the expression of distant genes or on the unidentified transcripts hiding in the region. Rare variants, which could not be detected by the current DNA microarrays, might also be contributing to the disease pathogenesis. One of the few ways to unravel the 'missing heritability'[6] of common diseases has been proposed to be obtaining the highest-resolution data by performing an in-depth sequencing of the particular loci using next-generation sequencers.

Recently, Harismendy et al.[7] have succeeded in demonstrating the biological role of a variant located in human chromosome 9p21, which is one of the most compelling gene desert previously shown by GWAS to be associated with coronary artery disease (CAD) and type-2 diabetes. The authors have identified 33 enhancers in 9p21, and it turned out to be the second densest gene

desert for predicted enhancers, suggesting the regulatory role of sequences residing within non-coding loci. They further investigated the interval precisely using various approaches, including identification of the complete set of variants in 9p21 by next-generation sequencers. They computationally predicted and experimentally confirmed that a variant in one of the enhancers disrupted the binding of *STAT1*, a transcription factor involved in inflammatory responses mediated by interferon-gamma. Finally, they determined a few adjacent, as well as distant (>45 kb), target gene regions relevant to CAD biology physically interacting with the enhancer by a new chromatin conformation capture technology, 3D-DSL,[7] which enables the detection of long-range enhancer interactions, in human vascular endothelial cells. Overall, their study has provided an excellent example of how to link the statistically associated variants to biological meanings and will definitely accelerate the studies of other non-coding GWAS loci associated with a number of common diseases to reveal the molecular mechanism of their pathogenesis.

However, it should be noted that the research flow of the association studies from variant discovery to its functional annotation has begun with using distinct 'case' and 'control' populations. In the initial studies, physicians carefully diagnosed the patients appropriate for the case group based on the diagnosis standard and strictly excluded the healthy volunteers who had any signs of disease-related abnormalities from the control group in order to maximize the statistical power to detect the variants specific

M Nakano and Professor K Tashiro are at the Department of Genomic Medical Sciences, Kyoto Prefectural University of Medicine, Kawaramachi-Hirokoji, Kamigyo-ku, Kyoto 602-8566, Japan.
E-mail: tashiro@koto.kpu-m.ac.jp

for the case group. Moreover, most of the studies focused only on the 'top hits' of the statistically significant variants, followed by the meta-analyses from multiple institutions to gain confidence of the variants identified; now we obtain a solution to take down the statistical association to molecular biology as described by Harismendy et al.[7] As the framework of the horizontal research stream of association studies has become clear and solid, it is quite natural for the researchers to continue on in the vertical direction in order to deepen our knowledge of disease etiology. The researchers have started paying attention to the 'runner-up hits' with modest association identified in the initial GWAS. As Ratanajaraya et al. displayed in their recent study,[1] researchers have also started subdividing the case group to carry out the association studies within the group to reveal the genetic determinants for slight, but significant phenotypic differences within a single disease. By continuing to build upon the width and depth of our knowledge for understanding the etiology of common diseases, 'personalized medicine' will become a true reality in the near future by utilizing the information regarding genetic–phenotypic relationships.

1 Ratanajaraya, C., Nishiyama, H., Takahashi, M., Kawaguchi, T., Saito, R., Mikami, Y. et al. A polymorphism of the POLG2 gene is genetically associated with the invasiveness of urinary bladder cancer in Japanese males. J. Hum. Genet 56, 572–576 (2011).
2 Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A. et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 449, 851–861 (2007).
3 Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C. et al. Complement factor H polymorphism in age-related macular degeneration. Science 308, 385–389 (2005).
4 Maller, J., George, S., Purcell, S., Fagerness, J., Altshuler, D., Daly, M. J. et al. Common variation in three genes, including a noncoding variant in CFH, strongly influences risk of age-related macular degeneration. Nat. Genet. 38, 1055–1059 (2006).
5 Nakano, M., Ikeda, Y., Taniguchi, T., Yagi, T., Fuwa, M., Omi, N. et al. Three susceptible loci associated with primary open-angle glaucoma identified by genome-wide association study in a Japanese population. Proc. Natl. Acad. Sci. USA 106, 12838–12842 (2009).
6 Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J. et al. Finding the missing heritability of complex diseases. Nature 461, 747–753 (2009).
7 Harismendy, O., Notani, D., Song, X., Rahim, N. G., Tanasa, B., Heintzman, N. et al. 9p21 DNA variants associated with coronary artery disease impair interferon-gamma signalling response. Nature 470, 264–268 (2011).