COMMENTARY

Commentary to 'A remark on rare variants'

Takahiro Nakamura

The article entitled 'A remark on rare variants' by Oexle[1] is an interesting report in which the author describes an approach to the study of common diseases caused by multiple rare variants. The article contains three parts. The first part is a discussion of the concept of population attributable risk (PAR) proposed by Bodmer and Bonilla.[2] PAR is a kind of index that represents the contribution of a variant to the onset of the disease. For calculating the cumulative genetic effect of rare variants, an approximate expression for estimating PAR is presented. However, the explanation of PAR in the original article is difficult to understand because it does not go into enough detail. The author explained the process of obtaining the equations described in the original article. The second part is a discussion of the efficiency of genetic analysis for rare variants with strong effect size. The powers of both affected sib-pair analysis and transmission/disequilibrium test (TDT) were calculated based on the methods of Risch and Merikangas.[3] The author illustrated that affected sib-pair analysis is more sensitive to a decrease in frequency or effect size than TDT. The third part proposes a disease model based on Kimura's infinite sites model. The author derived a simple relationship between the variant's selection coefficient and its effect size. The number of contributing genetic variants can be estimated by this model. Finally, TDT was applied to the disease model. The author derived the required sample size for the test.

The first two parts are very informative and contribute to the understanding of ori-ginal articles. Moreover, I find the disease model proposed in the third part very intriguing.

Recently, many genetic variants that cause multifactorial disease have been detected by genome-wide association study (GWAS). To increase the power of GWAS, high-frequency single-nucleotide polymorphisms (SNPs) are used as markers. This means that common variants with low odds ratio (OR) can be detected by GWAS. However, it is also possible to discover rare variants with high OR by sequencing around the positive markers. For example, the association between *ABCG2* and serum uric acid levels was identified by GWAS.[4] The functional variant Q141K in *ABCG2* increases the serum uric acid level. We searched for the causative SNPs of gout and hyperuricemia in the *ABCG2* gene. The common variant Q141K was associated with gout (minor allele frequency (MAF)=0.32; OR=2.23). Moreover, the rare variant Q126X was detected as a disease-related SNP (MAF=0.03; OR=4.25). Functional analysis showed that Q141K reduced activity by half and that Q126X was associated with no activity.[5]

Discoveries such as these may increase in the future. However, there may be many rare causative SNPs that cannot be detected by GWAS. Genetic researchers are accordingly interested in the question of how many causative variants exist with a particular OR and MAF. The mathematical model proposed in the third part may give an approximate answer to this question. The proposed model contains complicated equations; in particular, there are many equations about selection parameter *s*. However, the parameter *s* was integrated out and two parameters remained: relative fertility and affection rate of the disease. These two parameters alone are required to estimate the number of the causative variants. The model therefore seems to be simple.

The simulation study for the power of TDT was performed using estimates of the number of causative SNPs in the third part. The estimated number can be applied not only for a power calculation of TDT but also for a power calculation of an association study using chi-square test. Therefore, the model can help in study design. Cumulative PAR, which represents the summary genetic effect of particular variants, can also be calculated by the estimate. Moreover, if we can obtain a more exact estimate of the number of causative SNPs with some genetic effect, the heritability of the disease may be estimated.

This model is based on the assumption that relative fertility and affection rate determine the ORs and the frequencies for the variants. Moreover, it is assumed that the variants associated with disease are under selection, even if the disease onset is late. These two assumptions seem not to be robust. Therefore, evidence that this model represents the real structure of genetic disease is required for genetic researchers to use the model confidently.

One method for confirming the validity of the model is to check the inheritance of the variants. The author claims that disease susceptibility genes may be associated with a selective disadvantage even if the average onset is late. If this assumption is true, the rare variants contributing to diseases are not inherited with a probability of 1/2. It appears that the fitness of the model for the real genetic architecture could be verified by investigating the probability of inheritance of the disease-related variant, genetic effect, frequency and relative fertility.

Dr T Nakamura is at the Laboratory for Mathematics, National Defense Medical College, 3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan.
E-mail: nakataka@ndmc.ac.jp

The real structure of genetic disease would not appear as simple as described in the article. However, for the purpose of theoretical study it is useful to start from the simplest case. I hope that this article serves as the basis from which to develop research on rare causative variants.

1 Oexle, K. A remark on rare variants. *J Hum Genet* **55,** 219–226 (2010).
2 Bodmer, W & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* **40,** 695–701 (2008).
3 Risch, N & Merikangas, K The future of genetic studies of complex human diseases. *Science* **273,** 1516–1517 (1996).
4 Dehghan, A, Köttgen, A, Yang, Q, Hwang, S J, Kao, W L, Rivadeneira, F *et al.* Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *Lancet* **372,** 1953–1961 (2008).
5 Matsuo, H, Takada, T, Ichida, K, Nakamura, T, Nakayama, A, Ikebuchi, Y *et al.* Common defects of ABCG2, a high-capacity urate exporter, cause gout: a function-based genetic analysis in a Japanese population. *Sci Transl Med* **1,** 5–11 (2009).