LINK TO ORIGINAL ARTICLE

Tests of association for rare variants: case control mutation screening

Sean V. Tavtigian, Mia Hashibe and Alun Thomas

Two recent reviews, one in *Nature Reviews Genetics* from Bansal *et al.* (Statistical analysis strategies for association studies involving rare variants. *Nature Rev. Genet.* **11**, 773–785 (2010))[1] and one elsewhere[2], examined the emerging area of rare variant association studies. These reviews nicely describe the progression from association studies for common SNPs towards those for rare variants. We would like to add to these discussions a strategy that has been used by several groups for rare variant case–control association studies. This strategy was developed independently of genome-wide association (GWA) studies and is largely confined to cancer genetics, and we refer to it here as case–control mutation screening (CCMS).

Ideas contributing to CCMS are as follows. First, linkage analysis shows that evidence from many, individually very rare sequence variants at the same locus can be combined[3]. Second, clinical testing of susceptibility genes such as breast cancer 1, early onset (*BRCA1*) and *BRCA2* has shown that testing can be based on sequencing rather than genotyping. Third, the integrated evaluation of unclassified variants in *BRCA1* and *BRCA2* has shown that *in silico* assessment of rare variants — currently, rare missense substitutions (rMSs) — can be used to grade variants on the basis of predicted severity without attempting to dichotomize them as deleterious versus neutral[4]. Finally, lessons from GWA studies tell us that well-powered CCMS studies will be large, usually multicentre and often multi-ethnic, and therefore must be analysed by statistical methods that allow for covariates.

The development of CCMS can be traced through the efforts of the genetics community to understand the contribution of heterozygous sequence variation in ataxia telangiectasia mutated (*ATM*) to risk of breast cancer (TABLE 1). Analysis of *ATM* CCMS data started with a case–control study that used a cohort allelic sums test limited to protein-truncating variants plus variants that clearly damage splice junctions (T+SJVs)[5], Analyses progressed to a two-pronged strategy of analysing the pool of *ATM* T+SJVs in one logistic regression and the pool of rMSs in a second logistic regression[6]. The subtlety in this latter approach lies in combining all of the rMSs into a single categorical variable that incorporates prior information, such as sequence conservation, and grades the severity of rMSs from probably harmful to probably benign[4,6]. This variable is easily assessed in a logistical regression test for trend, thus minimizing the multiple testing problem while accommodating epidemiologic covariates. We believe that this form of CCMS, augmented by steadily improving statistical methods[7,8], will be useful for identifying genes that harbour variants conferring intermediate risk, especially those in which most pathogenic variants are rare and either reduce or ablate function.

Table 1 | **Key analyses of association of rare ATM variants with breast cancer**

Study	Year	Key elements
FitzGerald *et al.*[5]	1997	CCMS study, CAST test of T+SJVs
Gatti *et al.*[9]	1999	Theoretical study predicting relative importance of *ATM* missense substitutions
Sommer *et al.*[10]	2003	CCMS study incorporating prediction of missense substitution severity
Renwick *et al.*[11]	2006	Well-powered CCMS study, combined CAST of T+SJVs and pathogenic rMSs
Tavtigian *et al.*[6]	2009	CCMS meta-analysis, combined CAST of T+SJVs and graded rMS trend test
Bernstein *et al.*[12]	2010	Case–case analysis, supports importance of 'predicted damaging' rMSs

ATM, ataxia telangiectasia mutated; CAST, cohort allelic sums test; CCMS, case–control mutation screening; rMSs, rare missense substitutions; T+SJVs, protein truncating plus splice junction variants.

Going forward, improving the accuracy and scope of methods for predicting sequence variant severity is a key goal. To this end, the Critical Assessment of Genomic Interpretation community exercise will illuminate the capabilities of current approaches and inform their further development. An important additional issue is that methods for predicting gene dysfunction must be sufficiently transparent to allow other researchers to readily replicate predictions and judge the effects of hidden multiple testing (which maybe introduced by the prediction of sequence variant severity) on CCMS data analysis.

Sean V. Tavtigian, Mia Hashibe and Alun Thomas are at the University of Utah School of Medicine, Salt Lake City, Utah 84112, USA.
Correspondence to S.V.T.
e-mail: sean.tavtigian@hci.utah.edu

Published online 1 February 2011

1. Bansal, V., Libiger, O., Torkamani, A. & Schork, N. J. Statistical analysis strategies for association studies involving rare variants. *Nature Rev. Genet.* **11**, 773–785 (2010).
2. Asimit, J. & Zeggini, E. Rare variant association analysis methods for complex traits. *Annu. Rev. Genet.* **44**, 293–308 (2010).
3. Terwilliger, J. D. & Ott, J. *Handbook of Human Genetic Linkage* (Johns Hopkins Univ. Press, Baltimore and London, 1994).
4. Tavtigian, S. V., Byrnes, G. B., Goldgar, D. E. & Thomas, A. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. *Hum. Mutat.* **29**, 1342–1354 (2008).
5. FitzGerald, M. G. *et al.* Heterozygous *ATM* mutations do not contribute to early onset of breast cancer. *Nature Genet.* **15**, 307–310 (1997).
6. Tavtigian, S. V. *et al.* Rare, evolutionarily unlikely missense substitutions in *ATM* confer increased risk of breast cancer. *Am. J. Hum. Genet.* **85**, 427–446 (2009).
7. Le Calvez-Kelm, F. *et al.* Rare, evolutionarily unlikely missense substitutions in CHEK2 contribute to breast cancer susceptibility: results from a breast cancer family registry (CFR) case-control mutation screening study. *Breast Cancer Res.* **13**, R6 (2011).
8. Price, A. L. *et al.* Pooled association tests for rare variants in exon-resequencing studies. *Am. J. Hum. Genet.* **86**, 832–838 (2010).
9. Gatti, R. A., Tward, A. & Concannon, P. Cancer risk in ATM heterozygotes: a model of phenotypic and mechanistic differences between missense and truncating mutations. *Mol. Genet. Metab.* **68**, 419–423 (1999).
10. Sommer, S. S. *et al. ATM* missense mutations are frequent in patients with breast cancer. *Cancer Genet. Cytogenet.* **145**, 115–120 (2003).
11. Renwick, A. *et al. ATM* mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nature Genet.* **38**, 873–875 (2006).
12. Bernstein, J. L. *et al.* Radiation exposure, the *ATM* gene, and contralateral breast cancer in the women's environmental cancer and radiation epidemiology study. *J. Natl Cancer Inst.* **102**, 475–483 (2010).

Acknowledgements
This work was supported by US National Institutes of Health grant R01 CA121245.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Critical Assessment of Genomic Interpretation:
http://genomeinterpretation.org

ALL LINKS ARE ACTIVE IN THE ONLINE PDF