Utility of Breast Cancer Associated SNPs for Breast Cancer Risk Prediction

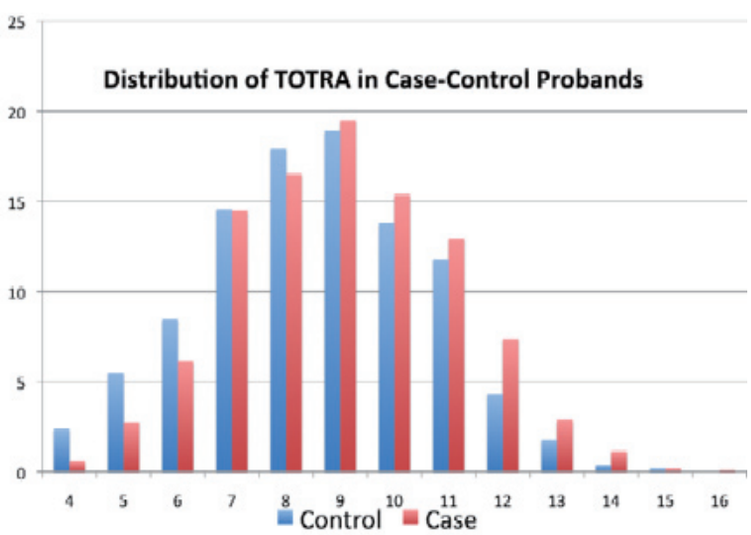
Wee Loon Ong[1], John Price[1], John Hopper[2], Georgia Chenevix-Trench[3], David Goldgar[4], Melissa Southey[5]

[1]Department of Biochemistry and Molecular Biology, Monash University, Australia, [2]Centre for Molecular, Environmental, Genetic and Analytical Epidemiology, The University of Melbourne, Australia, [3]Queensland Institute of Medical Research, Australia, [4]Department of Dermatology, University of Utah School of Medicine, USA, [5]Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Australia

Background: Recent genome-wide-association-studies (GWASs) investigating common genetic variants have successfully identified single nucleotide polymorphisms (SNPs) in a number of independent loci to be associated with breast cancer risk. Each SNP confers only a small increase in breast cancer risk (per-allele-OR<1.5), but the SNPs are purported to act multiplicatively, giving a higher risk in individuals carrying multiple susceptibility SNPs.

Methods: Samples for the study were obtained from the Australia Breast Cancer Family Study (ABCFS). Families with pathogenic BRCA1/2 mutations were excluded. DNA samples were extracted from Guthrie card blood spots, PCR-amplified using SNP-specific probes (Taqman), followed by end-point genotype analysis. 9 SNPs (verified by the Breast Cancer Association Consortium (BCAC)) have been selected for the study, namely rs2981582 (FGFR2), rs3803662 (TNRC9), rs3817198 (LSP1), rs889312 (MAP3K), rs13281615 (8q24), rs2107425 (H19), rs17468277 (CASP8), rs13387042 (2q35) and rs10941679 (5p12). For each individual, three SNP risk scores were computed: TOTRA (total number of risk alleles), TOTOR (total log odds ratio based on per allele OR reported by BCAC) and LOGOR (total log odds ratio based on per allele OR estimated using logistic regression on ABCFS samples). For relatives of case probands in the ABCFS database, whom SNPs genotype were not generated, simple SNP score imputation was done, whereby SNP score = K X (proband normalized SNP score) (K=kinship coefficient; K=1/2 if 1st degree relatives, 1/4 if 2nd degree relative, and 1/8 if 3rd degree relatives).

Results: Genotypes for all 9 SNPs were available for 2999 individuals: 1042 case probands, 508 control probands and 1449 relatives of case probands, of whom 111 had breast cancer diagnoses. The average age of breast cancer diagnosis for the case probands was 44, compared with 56 for the affected relatives. As expected, case probands have higher SNP scores compared to control probands ($p=2.5X10^{-6}$ for TOTRA; $p=6X10^{-6}$ for TOTOR; $p=9X10^{-12}$ for LOGOR), with cases carrying on average 0.5 more risk alleles than control probands (Figure 1). Similar risk score differences between case probands and control probands were observed using either all 9 SNPs data or 4 most significant SNPs data from stepwise analysis estimation. The highest



quartile of the LOGOR score was associated with increased breast cancer risk compared to the lowest quartile in the probands (OR=2.44; $p=8.5X10^{-9}$). Surprisingly, there was no difference in all three SNP scores between the relatives (with SNP data generated) as defined by their breast cancer status. In a Cox proportional hazards analysis of the risk of developing breast cancer in these relatives, none of the SNP scores significantly influenced the risk perhaps due to the small number of affected relatives. Comparisons of breast cancer incidence in relatives with SNP scores in the highest quartile and the lowest quartile also showed no significant difference. However, when relatives with predicted SNP scores were incorporated into the analysis (n=8416), the SNP scores appeared to influence the risk of developing breast cancer (HR=2.04; p=0.04).

Conclusion: Our results call into question the utility of SNP-based risk prediction, even in the familial setting, although the addition of many more such SNPs may improve their utility.