Supplementary Data to the Application Note

GSEA-SNP: Applying Gene Set Enrichment Analysis to SNP data from genome-wide association studies

Marit Holden[1], Shiwei Deng[2], Leszek Wojnowski[2] and Bettina Kulle[3,4]

[1]Norwegian Computing Center, Oslo, Norway

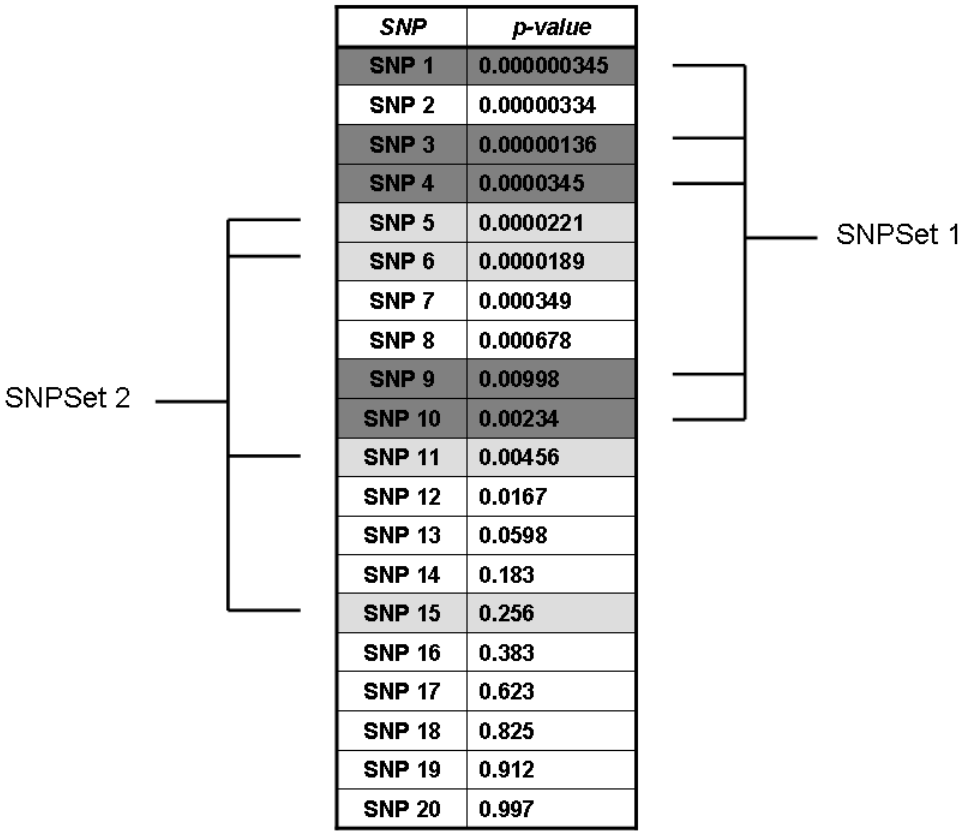
[2]Department of Pharmacology, University of Mainz, Mainz, Germany

[3]Epi-Gen, Faculty Division Akershus University Hospital, University of Oslo, Oslo, Norway

[4]Department of Biostatistics, University of Oslo, Oslo, Norway

1. Supplementary Figure

Graphical depiction of the GSEA-SNP method. SNPs from pathway 1 (Set 1) show a higher degree of enrichment among SNPs with low *P* values within the list *L* than SNPs from pathway 2 (Set 2). Pathway 1 is more likely to be involved in the phenotype under investigation than pathway 2.



SNP	*p-value*
SNP 1	0.000000345
SNP 2	0.00000334
SNP 3	0.00000136
SNP 4	0.0000345
SNP 5	0.0000221
SNP 6	0.0000189
SNP 7	0.000349
SNP 8	0.000678
SNP 9	0.00998
SNP 10	0.00234
SNP 11	0.00456
SNP 12	0.0167
SNP 13	0.0598
SNP 14	0.183
SNP 15	0.256
SNP 16	0.383
SNP 17	0.623
SNP 18	0.825
SNP 19	0.912
SNP 20	0.997

2. GSEA-SNP test results for a case-control data set

The GSEA-SNP method was applied to a SNP data set derived from a nested case-control study published elsewhere (Wojnowski, et al., 2005). This study includes 52 cases and 52 matched controls from a cohort of patients with non-Hodgkin lymphoma. Cases developed cardiotoxicity following chemotherapy with anthracyclines, whereas controls did not. Controls and cases were matched according to age, sex, and treatment. Each of the 104 DNA samples was hybridized to a GeneChip Human Mapping 10K Xba 131 Array (Affymetrix, Santa Clara, CA), which contains 11555 SNPs. The assignment of these SNPs to genes was taken from the Affymetrix annotation file (www.affymetrix.com). After excluding the monomorph SNPs, SNPs with unknown gene annotation, SNPs not in Hardy-Weinberg equilibrium (in controls) and SNPs with a call-rate below 80%, 9446 SNPs remained for the GSEA-SNP analysis.

The SNPs-in-pathways analysis was conducted using the C2 curated gene set for pathway annotation (http://www.broad.mit.edu/gsea/msigdb/collections.jsp#C2). This catalogue contained 1892 gene sets, which we converted into 1892 SNP sets. SNP sets with less than 15 SNPs or more than 500 SNPs were excluded, leaving 1038 sets for the analysis. Of those, 18 were significant at the nominal significance level of 0.01, including several directly or indirectly linked to anthracycline-induced cardiotoxicity [1] : CDMACPATHWAY is associated with cell proliferation through activation of MAP kinase pathway. Genes constituting the pathway (e.g. MAPK, TNF, PRKCA) have been implicated in anthracycline-induced cardiotoxicity (Kang, et al., 2000; Zhu, et al., 1999). OXSTRESS_BREASTCA_UP includes genes linked to electron transport and oxidoreductase activity, and thereby to oxidative stress underlying anthracycline-induced cardiotoxicity (Xiong, et al., 2006). NI2_MOUSE_DN contains a number of genes involved in cell adhesion which also contribute to the development of this phenotype (Abou El Hassan, et al., 2003; Deng, et al., 2007). Several genes from the LINDSTEDT_DEND_UP set are involved in apoptosis, cell death, cytokine and interleukin response which are also implicated in the cardiac response to anthracyclines (Kalyanaraman, et al., 2002). Due to the small sample size, these intriguing insights into the genetic causes of doxorubicin-induced cardiotoxicity remain to be verified in further, larger clinical cohorts.

References

Abou El Hassan, M.A., Verheul, H.M., Jorna, A.S., Schalkwijk, C., van Bezu, J., van der Vijgh, W.J. and Bast, A. (2003) The new cardioprotector Monohydroxyethylrutoside protects against doxorubicin-induced inflammatory effects in vitro, *Br J Cancer*, **89**, 357-362.

Deng, S., Kulle, B., Hosseini, M., Schluter, G., Hasenfuss, G., Wojnowski, L. and Schmidt, A. (2007) Dystrophin-deficiency increases the susceptibility to doxorubicin-induced cardiotoxicity, *Eur J Heart Fail*, **9**, 986-994.

[1] Detailed results are found on our webpage http://www.nr.no/pages/samba/area_emr_smbi_gseasnp.

Kalyanaraman, B., Joseph, J., Kalivendi, S., Wang, S., Konorev, E. and Kotamraju, S. (2002) Doxorubicin-induced apoptosis: implications in cardiotoxicity, *Mol Cell Biochem*, **234-235**, 119-124.

Kang, Y.J., Zhou, Z.X., Wang, G.W., Buridi, A. and Klein, J.B. (2000) Suppression by metallothionein of doxorubicin-induced cardiomyocyte apoptosis through inhibition of p38 mitogen-activated protein kinases, *J Biol Chem*, **275**, 13690-13698.

Wojnowski, L., Kulle, B., Schirmer, M., Schluter, G., Schmidt, A., Rosenberger, A., Vonhof, S., Bickeboller, H., Toliat, M.R., Suk, E.K., Tzvetkov, M., Kruger, A., Seifert, S., Kloess, M., Hahn, H., Loeffler, M., Nurnberg, P., Pfreundschuh, M., Trumper, L., Brockmoller, J. and Hasenfuss, G. (2005) NAD(P)H oxidase and multidrug resistance protein genetic polymorphisms are associated with doxorubicin-induced cardiotoxicity, *Circulation*, **112**, 3754-3762.

Xiong, Y., Liu, X., Lee, C.P., Chua, B.H. and Ho, Y.S. (2006) Attenuation of doxorubicin-induced contractile and mitochondrial dysfunction in mouse heart by cellular glutathione peroxidase, *Free radical biology & medicine*, **41**, 46-55.

Zhu, W., Zou, Y., Aikawa, R., Harada, K., Kudoh, S., Uozumi, H., Hayashi, D., Gu, Y., Yamazaki, T., Nagai, R., Yazaki, Y. and Komuro, I. (1999) MAPK superfamily plays an important role in daunomycin-induced apoptosis of cardiac myocytes, *Circulation*, **100**, 2100-2107.