Supplemental materials and methods

Selection of genetic variants and genotyping

Selection of genetic variations of NQO2 and GSTM1-5 has been described elsewhere (1, 2). The 19 polymorphisms analyzed are listed in **Table 2**. For the NQO2 gene, single-nucleotide polymorphisms (SNPs) spanning a 22.3-kb region from 2 kb upstream to 0.5 kb downstream of NQO2 were surveyed in the NCBI-dbSNP and the International HapMap websites (2). We used the HapMap database of the Han Chinese population (HapMap Data Rel 21a/phaseII). We selected tagging SNPs (tSNPs) using the pairwise method under a restriction of minor allele frequencies (MAFs) >0.05 and r^2 ≥0.8, aiming to identify a set of tSNPs that efficiently captures all known common variants and is likely to tag most unknown variants. In all, nine tSNPs were identified. These tSNPs capture all 31 alleles with a mean r^2 of 0.979. Three tSNPs only tag themselves, while one tags itself and one additional SNP. Each of these four tSNPs was located in an intron and was excluded from the study. Therefore, we chose five representative tSNPs that effectively captured 26/31 SNPs (84%). In addition, six variants in NQO2 with potential functional effects were chosen for genotyping. After genotyping, of the 11 variations, three non-synonymous SNPs were excluded due to non-polymorphism in the Shanghai population, and one SNP in an intron failed for technical reasons. The remaining seven polymorphisms could capture the un-genotyped SNPs well with coverage rate of 77%. The seven polymorphisms were rs2070999 (-338G>A, the transcriptional start site is designated as +1), I-29/D (-60 to -32), rs2071002 (+237A>C), rs1143684 (+10324C>T, Phe47Leu), rs4149367 (+15799C>T), rs1885298 (+17514G>T) and rs9501910 (+18708G>C). Of note, one recent study (3) demonstrated that the previously reported I-29/D is indeed a tri-allelic polymorphism consisting of a 29 bp insertion (I-29), a 29 bp deletion (D), and a 16 bp insertion (I-16). We used a sensitive genotyping strategy that permits identification of I-29, I-16 and D alleles in all of the samples (4). We have excluded individuals harboring I-16 alleles among the 806 patients.

For the GSTM4-M2-M1-M5-M3 gene cluster, SNPs spanning the 89 Kb of this region, from 2 Kb upstream of GSTM4 transcriptional region to 2 Kb upstream of GSTM3 transcriptional region (GSTM3 gene oriented opposite to the other GSTMs), were surveyed (1). The SNPs in the fragment of the GSTM1 gene with adjacent recombination genetic regions (~25 Kb) were

excluded because this fragment was absent in the GSTM1-null allele. The program Tagger used a strategy of pairwise methods under a restriction of MAFs>0.05 and $r^2 \geq 0.8$. Ten tSNPs (rs1010167, rs560018, rs535537, rs655315, rs4970773, rs17014553, rs1571858, rs11807, rs1332018, rs4970737) that successfully captured 28 of all 35 (28/35=80%) common SNPs (MAF>0.05) were identified. In addition, we selected seven potentially functional SNPs. In all, 17 SNPs were originally chosen for genotyping. After genotyping, we removed two SNPs that either were monomorphic or had MAF<1% in our pilot population, and we discarded four SNPs due to repeated technical genotyping failure. Therefore, eleven SNPs that captured 27 of the 35 common SNPs (77%) were available for prognosis analysis.

References

1. Yu KD, Fan L, Di GH, Yuan WT, Zheng Y, Huang W, et al. Genetic variants in GSTM3 gene within GSTM4-GSTM2-GSTM1-GSTM5-GSTM3 cluster influence breast cancer susceptibility depending on GSTM1. Breast Cancer Res Treat. 2010;121:485-96.

2. Yu KD, Di GH, Yuan WT, Fan L, Wu J, Hu Z, et al. Functional polymorphisms, altered gene expression and genetic association link NRH:quinone oxidoreductase 2 to breast cancer with wild-type p53. Hum Mol Genet. 2009;18:2502-17.

3. Wang W, Le WD, Pan T, Stringer JL, Jaiswal AK. Association of NRH:quinone oxidoreductase 2 gene promoter polymorphism with higher gene expression and increased susceptibility to Parkinson's disease. J Gerontol A Biol Sci Med Sci. 2008;63:127-34.

4. Yu KD, Di GH, Fan L, Hu Z, Chen AX, Shao ZM. Caution regarding genotyping methodology for a tri-allelic polymorphism in the novel breast cancer susceptibility gene NQO2. Breast Cancer Res Treat. 2009;118:647-9.