論文の内容の要旨

Identification of genetic variants associated with hormonal receptor positive breast cancer in the Japanese population

(日本人集団におけるホルモン受容体陽性乳がんに関わる遺伝子多型の同定)

エルガザール　シーハム　アブデル　ハルク

Background

Hormones including estrogen and progesterone have key roles both in the development and progression of the disease. Nearly 70% of breast cancersare known to express estrogen and progesterone receptors and their growth is dependent on the hormones. The exposure to high level and/or for long period of estrogen due to early menarche, late menopause, late age at first pregnancy, nulliparity, postmenopausal obesity, and high serum estrogen level in postmenopausal women,are considered to be risk factors for breast cancer. Furthermore, progestin, synthetic progesterone, was shown to markedly increase the risk of breast cancer in postmenopausal women when this hormonal therapy was provided for more than 10 years. In addition, genetic factorsplay an important role in the development of breast cancer. Although germ-line mutations in *BRCA1* and *BRCA2*, and rare variants in *CHEK2* and *ATM*have been shown to be implicated in high and moderate risks to breast cancer, respectively, several as yet unidentified common susceptibility variants are associated with low to moderate risk of the tumor.To identify genetic variants associated with hormonal receptor-positive (HRP) breast cancer, I performed a genome-wide association study (GWAS). Furthermore, we challenged to identify functional variant(s) that may be involved in the tumorigenesis of HRP-breast cancer through the locus identified by the GWAS.

Methodology

Most of the breast cancer cases and all the controls in this study were registered in BBJ and stored at the Institute of Medical Science, The University of Tokyo. All cases were diagnosed to have a HRP-breast cancer by the following examinations: examination of breast tissue (biopsy or cytology), estrogen receptor and progesterone receptor positivity, which was evaluated by immunohistochemistry.

In the initial GWAS, 1086 subjects with HRP-breast cancer had been selected as cases; 846 samples were obtained from BBJ and the remaining 240 samples were collected from collaborative hospitals.Controls were 1816 female subjects including 231 healthy volunteers from the Midosuji Rotary Club, Osaka, Japan,and 1585 subjectsthat were registered in BBJ as subjects with differenteightnon-cancerous diseases (hepatitis B, keloid, drug eruption, pulmonary tuberculosis, peripheral artery disease, arrhythmias, stroke, and myocardial infarction).In the replication study, 1547 cases were obtained from BBJ and additional 105 cases from the collaborative hospitals. A total of 2797 female subjects registered in BBJ for other diseases (rheumatoid arthritis, amyotrophic lateral sclerosis, and liver cirrhosis) were used as controls. 2266 cases with HRP-breast cancer and a total of 728 female controls were used for re-sequencing analysis.

All the samples were genotyped using Illumina Human Hap 610 Genotyping BeadChip. For sample quality

control, sample with call rate lower than 0.99 were excluded. Additionally, principal component analysis was performed to exclude individual who have admixture genetic component from the major Japanese Hondo cluster. For SNP quality control, SNPs which have call rate <0.99 in both cases and controls, *P*-value of Hardy-Weinberg equilibrium test of $<1.0x10^{-6}$ in control and minor allele frequency of SNP ≤0.01 were excluded from further analysis. A total of 453,627 SNPs on autosomal chromosomes and 10,525 SNPs on X chromosome passed the QC filters. P-values and OR with 95%CI were calculated by using PLINK program. Imputation analysis was carried out by referring to the genotype data of Japanese (JPT) individuals as deposited in the Phase II HapMap database using MACH v1.0 Genotypes of SNPs that are located in the genomic region within 500 kb upstream or downstream of the marker SNP (the SNP that showed the strongest association with HRP-breast cancer) were imputed. The imputed SNPs with an imputation quality score of $r^2>0.3$ were excluded from the subsequent analysis.

Six reporter plasmids containing one of the three regions; one in the 5' flanking region (region A: between -2,151 and -571)and two in the intron 1 (region B: between +10,063 and +11,152) and region C: between +13,956 and 15,864). These constructs harboredeither major or minor allele of the three SNP X, Y and Z, respectively. These SNPs showed perfect LD with the marker SNP rs6788895 and located in transcription factor binding sites. We further prepared four additional reporter plasmids, two containing 290 bp and two containing 60 bp in region B. All four constructs harbored either susceptible minor allele G or non-susceptible major allele T of SNPY. The Dual-Luciferase reporter assay was performed using BT-474 and T-47D breast cancer cells.

Results and Discussion

In the first screening set, I selected 33 single-nucleotide polymorphisms (SNPs) with suggestive associations in GWAS (P-value of 1×10^{-4}) as well as 4 SNPs that were previously implicated their association with breast cancer.By the replication using an independent set of 1653 cases and 2797 controls, I found an association of the disease withtwo SNPs, rs3750817 and rs2981579(Pcombined= 8.47×10^{-8} and 1.77×10^{-06} with OR=1.22 and OR=1.20, respectively)in the *FGFR2* (fibroblast growth factor receptor 2) gene, and a SNP, rs6788895 (Pcombined= 9.43×10^{-8} with odds ratio (OR) = 1.22) in the *SIAH*2 (seven in absentia homolog 2) gene on chromosome 3q25.1.Although the SNP in *SIAH*2 was significantly associated with HRP-tumors, it was not associated with estrogen receptor negative or progesterone receptor negative breast cancer.

To identify functional variations of the *SIAH2* gene, I focused on three SNPs, SNPX, SNP Y, and SNP Zthat are located in transcription factors-binding sitesand showed strong LD with rs6788895,the marker SNP associated with HRP-breast cancer. I prepared 6 reporter plasmids containing one of the three regions, region A, B, and C, encompassing SNP X, SNP Y, and SNP Z, respectively, and carried out a reporter assay in breast cancer cells. These constructs carried either major or minor allele of each SNP.Although promoter activity of region A in the 5' flanking region and enhancer activity of region C in intron 1 were not significantly different between major and minor allele, the enhancer activity of region B in intron 1 showed significant difference between the major and minor allele; the minor non-risk allele of T showed higher enhancer activity than the major risk allele of G. In addition, the reporter activities are higher than the control (mock vector), suggesting that the variation may affect the transcriptional enhancer activity of *SIAH2*. Since another SNPY2 was located in region B in addition to SNP Y,and was complete

linkage disequilibrium with SNP Y, we analyzed the effect of Y2 on reporter activity. As a result, the reporter plasmids lacking SNP Y2 and substitution mutant of Y2 did not reveal significant change of the reporter activity, suggesting that Y2 may not be involved in the enhancer activity. These data may imply that the non-risk allele of SNP Y leads to higher expression of *SIAH2*, and that SIAH2 may play a suppressive role in HRP-breast carcinogenesis.

The role of SIAH2 in human carcinogenesis remains controversial. It is reported that Siah2 has tumor suppressor effect by elimination of oncogenic proteins through proteasomal and kinase signaling pathways in leukemia. On the contrary, increased SIAH2 expression and its oncogenic role were shown in human neuroendocrine (NE) prostate tumor samples. Additionally, the role of SIAH2 in breast carcinogenesis is also controversial.It was reported that SIAH2 was under-expressed in breast cancer and that the patients with low level of expression of SIAH2 had a high risk of recurrence and low chance of disease free survivalwith increased the aggressiveness of the breast cancer. Their report supported the notion thatSIAH2 functions as a tumor suppressorfor breast cancer.Our data is in good agreement with this notion, because the risk allele of SNP Y is expected to reduce the enhancer activity and may results in low or repressed SIAH2 expression.

It was reported that low levels of SIAH2 were associated with resistance of breast cancer cells to tamoxifen, an inhibitor of estrogen receptor. It was reported that low levels of SIAH2 were associated with resistance of breast cancer cells to tamoxifen, an inhibitor of estrogen receptor. Considering that SIAH2 expression is regulated by ER, insufficient under-regulation of SIAH2 by tamoxifen may render cancer cells to escape from apoptosis or growth suppression. Further investigation of the SNPs in SIAH2 may help the personalized tamoxifentreatment for patients with HRP-breast cancer.

The SNP Y was predicted to present within the binding site of transcription factor I, II and III using UCSC Genome Bioinformatics. Since transcription factor I and II are involved in estrogen receptor signaling, exposure to estrogen may affect the SIAH2 expression through the association with transcription factor I and/or II. Since SIAH2 is involved in protein degradation via ubiquitin-proteasome system, altered expression of SIAH2 may affect the protein stability of tumor suppressor genes or oncogenesassociated withHRP-breast carcinogenesis. Further studies are needed to disclose the role of SIAH2 in HRP-breast carcinogenesis.

Conclusion

In my study, I discovered a link between SIAH2 and HRP-breast cancer. Additionally, Ifound that SNP Y that is associated with HRP-breast cancer may be involved in the regulation of SIAH2 expression. Further analysis of SIAH2 function may be useful for the development of new therapeutic and/or preventive approaches to HRP-breast cancer.