論文内容の要旨

Dissertation Abstract

Title: Identification of genetic variants associated with pancreatic cancer susceptibility and cyclophosphamide-induced adverse drug reactions in breast cancer

和文: ゲノムワイド関連解析による膵癌感受性遺伝子の同定および乳癌シクロホスファミド療法の副作用発現に関連する遺伝子多型の同定

氏名:羅　暁琪 (Low Siew Kee)

Part I: Genome-wide association study of pancreatic cancer in Japanese population

Background

Pancreatic cancer is the fifth leading cause of cancer death with an estimated death of 24,634 patients in Japan in year 2007. Its 5-year survival rate is as low as 6.7%. Since no specific symptom is observed at the early stage, most of the pancreatic cancer patients were diagnosed at their advanced stage with very low possibility of radical cure for the disease.

Previous reports indicated the involvement of both environmental and genetics factors in the etiology of this deleterious disease. Several case-control and cohort epidemiological studies have identified a number of possible risk factors such as smoking, diabetes, chronic pancreatitis, which are likely to predispose individual to the disease. Familial aggregation of the disease has implied the possible involvement of genetic factors in pancreatic cancer; approximately 10% of the patients were reported to have family history and individuals having first-degree relatives with pancreatic cancer revealed 2- to 4- fold higher risk of the disease. These data indicated that genetic factors are likely to play some roles in the development of pancreatic cancer.

Common genetic variations are known to associate with various common diseases and cancers. Single nucleotide polymorphism (SNP) is a type of common genetic variations which occurs in at least 1% of the human genome. The objective of this study is to identify SNPs associated with pancreatic cancer susceptibility in Japanese population through genome-wide association study (GWAS).

Methodology

This study is a collaborative study between The University of Tokyo and National Cancer Center, Japan. A total of 1006 invasive pancreatic ductal adenocarcinoma cases and 5311 controls were recruited. All the samples were genotyped using Illumina HumanHap550v3 or Illumina Human Hap 610 Genotyping BeadChip. For sample quality control, sample with call rate lower than 0.98 were excluded. Additionally, principal component analysis was performed to exclude individual who have admixture genetic component from the major Japanese Hondo cluster. After sample QC, association study was performed on 991 cases and 5209 controls. For SNP quality control, SNPs which have call rate <0.99 in both cases and controls, *P*-value of Hardy-Weinberg equilibrium test of $<1.0\times10^{-6}$ in control and minor allele frequency of SNP ≤0.01 were

excluded from further analysis. A total of 420,236 SNPs on autosomal chromosomes passed the QC filters. Case-control association study was performed by logistic regression analysis after adjustment of age (continuous), sex and smoking status (current/former, never). *P*-values and OR with 95%CI were calculated for allelic, dominant and recessive models by using PLINK program. We used the minimum *P*-values obtained from three models to evaluate the statistical significance of the association. Significance threshold was set at $P<5.0\times10^{-7}$. To infer untyped and missing genotypes around the candidate loci, genotype imputation was performed by utilizing a Hidden Markov model programmed in MACH version 1.0. By utilizing the genotype information from the HapMap database, maximum likelihood genotypes for the untyped SNPs were generated. For quality control, imputed SNPs with the estimated r^2 of >0.3 were retained.

Results and Discussion

The Q-Q plot for this GWAS based on allelic *P*-values by logistic regression revealed no significant population stratification with genomic inflation factor λ of 1.026 after sample and SNP QC. Three genomic regions 6p25.3, 12p11.21 and 7q36.2, shown to be significantly associated (P-value$<5.0\times10^{-7}$) with increased risk of pancreatic cancer in Japanese population were successfully identified (Table 1).

Table1: Association study of GWAS of pancreatic cancer in Japanese population

CHR*	SNP	Position*	Risk allele	RAF		Allelic				Dominant				Recessive				Pmin	Gene	Relative loc*
				Case	Control	P-value	OR	L95	U95	P-value	OR	L95	U95	P-value	OR	L95	U95			
6	rs9502893	1285189	G	0.411	0.351	3.30E-07	1.29	1.17	1.43	2.97E-05	1.36	1.18	1.57	2.18E-05	1.50	1.24	1.80	**3.30E-07**	*FOXQ1*	25196
12	rs708224	32327676	A	0.718	0.656	3.30E-07	1.32	1.19	1.47	8.54E-07	1.42	1.23	1.63	2.09E-03	1.46	1.15	1.86	**3.30E-07**	*BICD1*	0
7	rs6464375	153256776	A	0.116	0.103	1.15E-01	1.13	0.97	1.32	7.36E-01	1.03	0.87	1.22	4.41E-07	3.73	2.24	6.21	**4.41E-07**	*DPP6*	0

Odds ratios, 95% confidence limits and P-values were obtained using logistic regression analysis according to allelic, dominant and recessive model after adjustment of age, sex and smoking.
RAF, risk allele frequency; OR, odds ratio; L95 U95, lower and upper confidence limits; Pmin, minimum P-value among three genetic models.
**Position and relative loci (Relativeloc) based on NCBI Human Genome Build 36.*

The most significantly-associated SNP, rs9502893, is located within a 75-kb linkage disequilibrium (LD) block on chromosome 6p25.3. This LD block includes *FOXQ1* (forkhead box (Fox) Q1) gene, which is located 25 kb upstream to this marker SNP. Imputation analysis revealed modest association at SNPs located near to or on the *FOXQ1* gene suggesting it to be one of the causative genes for pancreatic cancer. *FOXQ1* encodes for protein forkhead box (Fox) Q1. The Fox family of transcription factors consists of at least 43 members and mutations in Fox genes can cause significant effects on human common diseases and cancers. A recent study showed that FoxQ1 is overexpressed in pancreatic cancer, suggesting its role in pancreatic tumorigenesis.

The second significantly-associated SNP, rs708224, located in the second intron of the gene *BICD1* (Bicaudal-D homolog 1) on chromosome 12p11. The 80-kb LD block showing the association corresponds to the second intron of *BICD1* as revealed by the imputation analysis. Bicaudal-D homolog 1 protein plays a role in vacuolar trafficking. A recent study suggested that genetic variations within the *BICD1* gene could alter its transcriptional levels and in turn influence telomere length in humans. In addition, several recent studies have documented reduced telomere length in pancreatic ductal adenocarnoma specimens, suggesting telomeric dysfunction in pancreatic cancer cells.

The third locus is marked by rs6464375 in the first intron of *DPP6* gene. The SNP indicated suggestive associations only under recessive model. *DPP6* encodes protein dipeptidyl-peptidase 6, which binds to specific

voltage-gated potassium channels and alters their expression and biophysical properties. A recent study on core signaling pathways in human pancreatic cancers found three somatic mutations in *DPP6* among 24 pancreatic cancer samples examined by detailed sequence analyses. This report also suggested that *DPP6* might play a crucial role in regulation of invasion of pancreatic cancer cells.

Conclusion

This study represents the first GWAS to identify common genetic variants associated with pancreatic cancer in Japanese population. Genes that identified by this study have been implicated to play important roles in the pathogenesis, telomere dysfunction and invasion of pancreatic cancer cells. Our findings may contribute to a better understanding of pancreatic carcinogenesis.

Part II: Identification of genetic variants associated with cyclophosphamide-induced adverse drug reactions in breast cancer

Background

Cyclophosphamide (CPA) is one of the most widely used anticancer drugs in the treatment of hematological malignancies and a variety of solid tumors including breast cancer. The CPA-based combination treatment has known to be effective for breast cancer, but often causes adverse drug reactions (ADRs), such as leucopenia/neutropenia and gastrointestinal symptoms such as vomiting, anorexia and nausea. Most of the drug-metabolizing enzymes and transporters contain a wide range of genetic polymorphisms, which might cause a large interindividual variability in the plasma concentration of drugs. Furthermore, anticancer therapies are notoriously known to have narrow therapeutic range; higher concentration in patients' body causes toxicity and lower concentration reduces the efficacy of the drugs. Hence, the role of pharmacogenomics which is expected to provide a predictive way for severe drug toxicity is greatly essential. The objective of this study is to discover SNPs associated with CPA-induced ADRs in patients with breast cancer using a case-control association study, focusing on not only the drug-metabolizing enzymes, but also the transporters, which might also play an important role in pharmacokinetics of CPA or its active forms.

Methodology

All the samples of this study were recruited from Biobank Japan, the University of Tokyo. In this study, patients who revealed ≥grade3 leucopenia or neutropenia, or those with ≥grade2 gastrointestinal toxicity induced by CPA combination therapy were defined as cases (ADRs), while controls (non-ADRs) were defined as patients who had shown no toxicity during CPA-based combination therapy. The first exploratory sample set consist of 76 cases (ADRs) and 140 controls (non-ADRs), and an independent second set samples (replication set) consist of 108 ADR cases and 79 non-ADR controls were recruited. A total of 141 SNPs (tagSNPs and functional SNPs) and two deletion polymorphisms in 13 candidate genes, which are involved in activation, detoxification and transportation of CPA (*CYP2B6*, *CYP2C9*, *CYP2C19*, *CYP3A4*, *CYP3A5*, *ALDH1A1*, *ALDH3A1*, *GSTA1*, *GSTM1*, *GSTP1*, *GSTT1*, *ABCC2*, and *ABCC4*), were genotyped. The selection criteria of the tagSNPs were based on the measures of linkage disequilibrium (LD) with r^2 value ≥0.8 and minor allele frequency (MAF) of >10% from the HapMap database. All the SNPs were genotyped using multiplex

polymerase chain reaction (PCR)-invader assay or direct sequencing. Case-control association study was evaluated using Cochran-Armitage trend test. Additionally, multiple testing, Bonferroni correction threshold (*P*-value<0.00035) were applied, to assess the significance level of the association. Subgroup analysis, such as types of ADR developed after receiving CPA combination therapy and types of chemotherapy regimen for breast cancer, were also evaluated.

Results and Discussion

Among the 143 common variations analyzed in this study, one SNP, rs9561778 in *ABCC4* is significantly associated with CPA combination therapy induced ADR after applying strict Bonferroni's correction (Table 2). For the subgroup based on types of ADRs, rs9561778 showed association with both the gastrointestinal toxicity and leucopenia/neutropenia, yielding similar trends of odds ratio (Table 2). Additionally, this SNP also revealed significant association with a higher odds ratio with patients treated with the CA(F) (Cyclophosphamide and Anthracyclin with or without 5-fluorouracil) drug regimen (Table 2) in which CA(F) regimen is one of the most major combination therapies for breast cancer.

Table 2: Association study of SNP rs9561778 with CPA combination therapy induced ADRs and its subgroup analysis.

Chr	SNP	Risk allele	Cat.	Risk allele frequency		Cochran-Armitage trend p-value	Odds ratio	95% CI	
				ADR	Non-ADR			Lower	Upper
ALL									
13	rs9561778	T	1st stage	0.24	0.13	**0.0086**	2.11	1.16	3.83
			2nd stage	0.26	0.18	**0.047**	1.72	0.95	3.13
			Total	0.25	0.15	**0.00031**	2.06	1.36	3.11
ADR subgroups									
13	rs9561778	T	GI*	0.27	0.15	**0.00019**	2.31	1.45	3.68
			LN**	0.23	0.15	**0.014**	1.83	1.10	3.05
Drug regimen subgroup									
13	rs9561778	T	CA(F)***	0.26	0.12	**0.00028**	3.13	1.68	5.83

GI*: ≥Grade 2 Gastrointestinal toxicity, LN**: ≥Grade3 leucopenia or neutropenia
CAF***: C; Cyclophosphamide, A; Anthracyclin (Epirubicin or Adriamycin), F; 5-fluorouracil

ABCC4 is a member of the superfamily of ATP-binding cassette (ABC) transporters. A recent study indicated that CPA and/or its active metabolites are the substrates to ABCC4 because the *in vitro* CPA cytotoxicity was significantly enhanced after addition of ABCC4 inhibitor. Since ABCC4 is known to express in kidney, the expression of ABCC4 in the kidney might play an important role in elimination of CPA and its metabolites from the body and genetic variations within this gene might affect the amount or nature of this transporter, resulting in impairment of excretion and subsequent overdose manifestation. In addition, the expression of ABCC4 in the sinusoidal membrane of hepatocytes might facilitate the secretion of active metabolites of CPA produced in the liver into the systemic circulation. Variants on this gene might cause excess efflux of CPA and its metabolites, which consequently increased systemic drug concentration in the body.

Conclusion

Through candidate gene approach, associations of *ABCC4* genotypes and CPA-induced ADRs were identified. Although the association as well as the mechanism to induce ADRs should be further validated by using larger number of samples or by molecular analysis, this study has contributed another piece of puzzle into the mist of prediction system which may help in identifying patients at risk of CPA-induced ADRs and lead to a better prognosis and quality of life for patients with cancer.