A Genome-Wide Association Study Identifies Five Novel Genetic Markers for Trastuzumab-Induced Cardiotoxicity in Japanese Population

Mari Hara Nakano,a,b Chihiro Udagawa,a,c Arata Shimo,b Yasuyuki Kojima,b Reiko Yoshie,b Hisamitsu Zaha,d Norie Abe,d Tokiwa Motonari,d Mikiko Unesoko,d Kenji Tamura,e Tatsunori Shimo,i Masayuki Yoshida,f Teruhiko Yoshida,g Hiromi Sakamoto,g Ken Kato,h Taisei Mushiroda,i Koichiro Tsugawa,h Hitoshi Zembutsu*a,g

aProject for Development of Liquid Biopsy Diagnosis, Japanese Foundation for Cancer Research, Research Institute; 3-8-31 Ariake, Koto-ku, Tokyo 135-8550, Japan; bDivision of Breast and Endocrine Surgery, Department of Surgery, St.Marianna University School of Medicine; 2-16-1 Sugao, Miyamae-ku, Kawasaki 216-8511, Japan; cNew Business Development Life Science Group, Toyo Kohan Co., Ltd.; 2-18-1, Higashi-Gotanda, Shinagawa-ku, Tokyo 141-8260, Japan; dDepartment of Breast Surgery, Nakagami Hospital; 610, Noborikawa, Okinawa 904-2195, Japan; eDepartment of Breast and Medical Oncology, National Cancer Center Hospital; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan; fDepartment of Pathology and Clinical Laboratories, National Cancer Center Hospital; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan; gFundamental Innovative Oncology Core, National Cancer Center Research Institute; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan; hDepartment of Gastrointestinal Medical Oncology, National Cancer Center Hospital; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan; and iRIKEN, Center for Integrative Medical Sciences; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.
*Corresponding author.

Phone: +81-3-3570-0509 / FAX: +81-3-3570-0515 / E-mail: hitoshi.zembutsu@jfcr.or.jp
Summary

Trastuzumab has been administered to patients with HER2-positive cancer, however, the cardiotoxicity is identified as one of the life-threatening toxicities. Clinically useful biomarker for trastuzumab-induced cardiotoxicity has been expected to be developed. To identify a novel genetic marker(s) determining the risk of trastuzumab-induced cardiotoxicity, we performed a first genome-wide association study (GWAS) in Japanese population. We enrolled 481 patients who had been treated with trastuzumab and carried out a GWAS using 11 cases (with cardiotoxicity) and 257 controls (without cardiotoxicity). Top 100 single nucleotide polymorphisms (SNPs) which revealed the smallest $P$ values in GWAS ($P = 7.60 \times 10^{-7} - 2.01 \times 10^{-4} )$ were further examined using replication samples consisted of 14 cases and 199 controls. The combined analysis of the GWAS and replication study indicated possible association of five loci with trastuzumab-induced cardiotoxicity (rs9316695 on chromosome 13q14.3, rs28415722 on chromosome 15q26.3, rs7406710 on chromosome 17q25.3, rs11932853 on chromosome 4q25, and rs8032978 on chromosome 15q26.3, $P_{combined} = 6.00 \times 10^{-6}, 8.88 \times 10^{-5}, 1.07 \times 10^{-4}, 1.42 \times 10^{-4}, 1.60 \times 10^{-4}$, respectively). Furthermore, we developed a risk prediction model for trastuzumab-induced cardiotoxicity using the five marker SNPs. The incidence of trastuzumab-induced cardiotoxicity in patients with risk score $\geq 5$ was significantly higher (42.5%) compared to that in patients with score $\leq 4$ (1.8%) ($P = 7.82 \times 10^{15}$, odds ratio = 40.0). These findings suggest the potential to improve the ability of physicians to avoid the trastuzumab-induced cardiotoxicity for patients with HER2-positive cancer.
Key words

cardioxicity, human epidermal growth factor receptor 2 (HER2), breast cancer, trastuzumab, genome-wide association study (GWAS)
Introduction

Human epidermal growth factor receptor (EGFR) type2 (HER2) is a member of the EGFR family, which is a transmembrane tyrosine kinase receptor involved in cell differentiation and proliferation through the activation of intracellular signaling pathways in response to homo- and hetero-dimerization of HER2 with EGFR family members. It is reported that HER2 is overexpressed or amplified in approximately 25%-30% of breast cancer patients.1-3) Furthermore, patients with HER2-positive cancer have a higher risk of recurrence and mortality than those with HER2-negative cancer.1,4) Trastuzumab is a recombinant humanized monoclonal antibody and selectively targets the extracellular domain of the HER2 protein.1,3-5) Trastuzumab induces antibody-dependent cellular cytotoxicity (ADCC) and prevents HER2 receptor from its dimerization, resulting in inhibiting downstream signaling pathways: the phosphoinositide 3-kinase (PI3K) pathway, AKT pathway and MAPK pathway, which are associated with cell survival and proliferation.3,5,6) It is reported that patients treated with chemotherapy plus first-line trastuzumab showed a significantly better response than chemotherapy alone, and prolonged survival.7-9) However, it is reported that one of the most serious side effects of trastuzumab is cardiotoxicity, which causes cardiac dysfunction, either asymptomatic or symptomatic.6,10,11) Although trastuzumab is an effective drug for HER2-positive cancer, some patients cannot continue the trastuzumab therapy due to the cardiotoxicity. The mechanism of trastuzumab-induced cardiotoxicity is not yet fully understood, however, inhibition of HER2 signaling in cardiomyocytes has been suggested to be involved in this toxicity.12,13) Anthracycline-induced cardiotoxicity is thought to be cumulative, dose-dependent, and irreversible, whereas trastuzumab-induced cardiotoxicity is mostly reversible and
dose-independent, suggesting that genetic factors play important role in this toxicity.⁶,¹⁴,¹⁵

The association studies of candidate genes including Erb-B2 receptor tyrosine kinase 2 (ERBB2), which encodes HER2, have been reported, however, these results have not yet been sufficiently validated,¹⁶-¹⁸ and clinically available predictive marker for trastuzumab-induced cardiotoxicity has not yet been developed. Hence, it has been suggested that the other unknown genetic factors which could contribute to the risk of trastuzumab-induced cardiotoxicity should exist. In this study, to identify common genetic variants determining the risk of trastuzumab-induced cardiotoxicity, we carried out a genome-wide association study (GWAS) and replication study of 481 patients treated with trastuzumab, and identified five loci that were likely to be associated with risk of trastuzumab-induced cardiotoxicity.

**Materials and Methods**

**Patients**

To identify a genetic marker(s) for the risk of trastuzumab-induced cardiotoxicity, we selected 268 patients treated with trastuzumab from the samples registered in NCC biobank (http://www.ncc.go.jp/jp/biobank/) from February 2010 to December 2015 for GWAS stage. In the replication study, 213 patients (14 cases and 199 controls) treated with trastuzumab were collected from St. Marianna University School of Medicine Hospital and Nakagami Hospital from October 2017 to March 2018, and collected from the samples registered in NCC Biobank from January 2016 to October 2017. We defined trastuzumab-induced cardiotoxicity as left ventricular ejection fraction (LVEF) < 45% or LVEF < 50% with an absolute decrease of 10% from
baseline according to the criteria in Herceptin Adjuvant (HERA) trial.10,12,15,19) Transthoracic echocardiography was performed for assessment of LVEF at baseline and during follow-up (every 1.0 - 37.7 months, median; 3.1 months). Echocardiography was performed according to standard procedures by one or two sonographers and analyzed by one cardiologist using the iE33 (Philips, Netherlands), SONOS5500 (Philips, Netherlands) or VividE95 (GE, USA) system. All participants provided written informed consent to this study. The research protocol of the study was reviewed and approved by the Research Ethics Committee of the National Cancer Center (Tokyo, Japan), St. Marianna University School of Medicine (Kawasaki, Japan), Nakagami Hospital (Okinawa, Japan) and Japanese Foundation for Cancer Research (Tokyo, Japan), and the members of the committee included experts in the genetic research and ethics.

**Genotyping and quality control**

Genomic DNA was extracted from peripheral blood or formalin-fixed paraffin-embedded (FFPE) lymph nodes which did not contain cancer cells. For the GWAS analysis, 268 patients were genotyped using the Infinium OmniExpressExome-8 v1.4 DNA Analysis Kit (Illumina, San Diego, USA). We applied SNP quality control (call rate of ≥ 99%) in both cases and controls and a Hardy–Weinberg equilibrium ($P > 1.0 \times 10^{-6}$) in controls and SNPs with minor allele frequency of ≥ 0.01. A total of 543,807 SNPs in autosomal chromosomes passed the quality control filters. We utilized the identity-by-state method to evaluate cryptic relatedness for the samples included in this study. The population stratification was examined by principal component analysis (PCA) using the EIGENSTRAT software 6.0.1
The PCA was carried out by comparing the distribution of the sample populations with three reference populations from the 1,000 Genomes Project Phase3 database that included Europeans (represented by Caucasian from UTAH, CEU), Africans (represented by Yoruba from Ibadan, YRI) and East Asians (represented by Japanese from Tokyo, JPT, Han Chinese from Beijing, CHB, Southern Han Chinese, CHS, Chinese Dai in Xishuangbanna, CDX, and Kinh in Ho Chi Minh City, Vietnam, KHV). The PCA was performed on the basis of the genotype information from the samples included in this study. The quantile-quantile (Q-Q) plot was generated between observed $P$ values of Fisher's exact test against expected $P$ values and showed that genomic inflation factor was 1.172, representing no significant population stratification (Supplementary Fig.1).

**Genotype imputation**

Genome-wide imputation was performed for the 268 patients used in GWAS. Because 43 SNPs located in the 3 genomic regions (chromosome 13q14.3 and 2 independent loci on 15q26.3) showed greater significance than marker SNPs (tag-SNPs) at each locus in the genome-wide imputation analysis, these genomic regions (chromosome 13q14.3, 54593774-54618139, 15q26.3, 98578726-98646496, and 15q26.3, 101796748-101800094, respectively) were also imputed for 213 patients used in replication study. For imputation, we used the reference panel which was based on the 1,000 Genomes Project Phase 3 integrated release version 5 for individuals of East Asian descent comprising Japanese from Tokyo, Chinese from Beijing and Chinese from southern China. Missing genotypes were imputed with Minimac3 software. We extracted variants with imputation quality threshold of $RSQ > 0.3$. For imputation...
analysis, we performed SNP quality control by excluding SNPs that had low genotyping rate < 99%, and showed deviations from Hardy-Weinberg equilibrium ($P \leq 1.0 \times 10^{-6}$) in controls and SNPs with minor allele frequency of < 0.01.

Statistical analysis

In the GWAS and the replication study, a case-control study by Fisher's exact test was applied to three genetic models: an allelic frequency model, a dominant-inheritance model, and a recessive-inheritance model. Odds ratios (ORs) and confidence intervals (CIs) were calculated for the genetic model with smallest $P$ value in those three genetic models, using a non-risk allele or a non-risk genotype as a reference. Genome-wide significance levels after adjustment of Bonferroni correction for multiple testing of three genetic models were $P = 3.06 \times 10^{-8}$ [$0.05 / (543,807 \times 3)$] in the GWAS and $P = 1.67 \times 10^{-4}$ [$0.05 / (100 \times 3)$] in the replication study. For combination analysis, the genotype count of the replication study was added to that of the GWAS. The differences in the distribution of age, primary cancer site, pretreatment with anthracycline, HER2 status and hormone receptor status were evaluated by logistic regression analysis (Supplementary Table 1). For the predictive scoring system of the risk of trastuzumab-induced cardiotoxicity, we assigned a score of 2 to individuals homozygous for risk allele, 1 to individuals heterozygous for risk allele, and 0 to individuals homozygous for non-risk allele at rs9316695, rs11932853 and rs8032978, respectively. We further assigned a score of 2 to individuals homozygous for the risk allele and 0 to individuals with the other genotypes at rs28415722 and rs7406710 and summed up the scores for each SNP to obtain individuals’ scores. Based on this system, each patient was classified into any of the nine prediction score
groups (score 0, 1, 2, 3, 4, 5, 6, 7, or 8). All the statistical analyses were carried out using R statistical environment version 3.3.1 (http://www.r-project.org/), PLINK version 1.07,20 or the BellCurve for Excel (Social Survey Research Information Co., Ltd., Tokyo, Japan). Regional association plots were generated using Locus Zoom (http://locuszoom.sph.umich.edu/).

Results

Patient characteristics

To identify a genetic marker(s) determining the risk of trastuzumab-induced cardiotoxicity, we recruited 481 patients treated with trastuzumab, including 25 cases (with trastuzumab-induced cardiotoxicity) and 456 controls (without trastuzumab-induced cardiotoxicity). Table 1 showed the characteristics of these 481 patients receiving trastuzumab treatment. Their median age at the beginning of the trastuzumab treatment was 56 years old in both case and control groups. The distributions of gender (percentage of female patients) were 100 and 76.7% in the case and control groups, respectively, in the GWAS, and 100 and 91.5%, respectively, in the replication samples (Supplementary Table 2). The incidence of cardiotoxicity in gastric and breast cancers were 1.1% (1/90) and 6.2% (24/387), respectively (Table 1). Among the characteristics listed in Table 1, none of them showed significant association with trastuzumab-induced cardiotoxicity in logistic regression analysis (Supplementary Table 1).

GWAS and replication analysis

In this study, trastuzumab-induced cardiotoxicity was defined as LVEF <45%
or LVEF <50% with an absolute decrease of 10% from baseline according to the criteria in Herceptin Adjuvant (HERA) trial. 10, 12, 15, 19) We carried out a GWAS of 268 patients who received trastuzumab therapy using the Infinium OmniExpressExome-8 v1.4 DNA Analysis Kit (Illumina). After the standard quality control, an association analysis was carried out for 543,807 SNPs by Fisher’s exact test on the basis of three genetic models: allelic, dominant, and recessive. We generated a quantile-quantile plot and observed the genomic inflation factor of 1.172, indicating no population stratification (Supplementary Fig.1). However, we could not observe the SNP which reached the genome-wide significance level (Supplementary Fig.2). The top 100 SNPs which revealed the smallest $P$ value in GWAS showed possible associations ($P = 7.60 \times 10^{-7} - 2.01 \times 10^{-4}$, Supplementary Table 3). To further validate the results of the GWAS analysis, we carried out a replication study of the 100 SNPs using 213 independent patients, including 14 cases and 199 controls. Although no SNP reached the significance level for three genetic models after the Bonferroni correction ($P = 1.67 \times 10^{-4} [0.05/ (100 \times 3)]$) in the replication study, we identified 16 SNPs with suggestive associations ($P = 3.57 \times 10^{-2} - 9.37 \times 10^{-2}$, Supplementary Table 4). A combined analysis of the GWAS and the replication study for the 16 SNPs indicated that nine SNPs in five loci showed stronger association (smaller $P$ value) than those in GWAS results (rs9316695, rs9527156, rs12583122 and rs11617903 on chromosome 13q14.3, $P_{\text{combined}} = 6.00 \times 10^{-6}, 8.64 \times 10^{-6}, 1.92 \times 10^{-5}$ and $2.22 \times 10^{-5}$, respectively, rs1383149 and rs12372962 on chromosome 15q26.3, $P_{\text{combined}} = 1.01 \times 10^{-4}$ and $1.15 \times 10^{-4}$, respectively, rs7406710 on chromosome 17q25.3, $P_{\text{combined}} = 1.07 \times 10^{-4}$, rs11932853 on chromosome 4q25, $P_{\text{combined}} = 1.42 \times 10^{-4}$, and rs8032978 on chromosome 15q26.3, $P_{\text{combined}} = 1.60 \times 10^{-4}$, Supplementary Table 5, Fig. 1), however, none of them reached
the genome-wide significance level in this study ($P = 3.06 \times 10^{-8}$).

Furthermore, we carried out subgroup analysis according to cancer type and hormone receptor status (Supplementary Table 6), and found that rs8032978 on chromosome 15q26.3 showed stronger association in patients with hormone receptor positive breast cancer compared to that in original combined analysis ($P = 6.55 \times 10^{-5}$, OR=8.53, 95% CI=2.91-23.39, Supplementary Table 6). This finding suggests that rs8032978 might also be a promising marker of trastuzumab-induced cardiotoxicity in patients with hormone receptor positive breast cancer. Association analysis was also carried out by logistic regression analysis with adjusting medication for cardiovascular disease and gender which could be associated with the risk of cardiotoxicity (Supplementary Table 2), however, none of the above SNPs did not reach genome-wide significance level (Supplementary Table 7).

Imputation analysis

To identify additional candidate loci associated with trastuzumab-induced cardiotoxicity, we examined the associations by using genome-wide imputed genotypes of GWAS samples, however, we could not observe novel candidate locus associated with the trastuzumab-induced cardiotoxicity. Of the five candidate loci identified in GWAS (chromosome 4q25, 13q14.3, 2 independent loci on 15q26.3, and 17q25.3), imputation analysis identified 43 novel SNPs in three loci (chromosome 13q14.3 and 2 independent loci on 15q26.3) of which significance were equal to or greater than the marker SNPs (tag-SNPs) identified in GWAS (Fig. 1A, B, E, Supplementary Table 8). We further performed a replication study using 213 independent patients for the 43 SNPs and found that no SNP showed stronger association than each marker SNP in the
above three loci (Supplementary Table 8). However, a combined result of the GWAS and replication stage for the 43 SNPs revealed that the association of an imputed SNP, rs28415722 on chromosome 15q26.3, was the strongest in the region ($P_{\text{combined}} = 8.88 \times 10^{-5}$, Supplementary Table 5, 8). Hence, we regarded rs28415722 as a marker SNP in the region (Table 2, Fig. 1B).

**Predictive scoring system for trastuzumab-induced cardiotoxicity using the five marker SNPs**

Because the five SNPs, which showed smallest $P$ value in the combined study at each candidate locus (rs9316695, rs28415722, rs7406710, rs11932853 and rs8032978, Table 2), were independent predictors of trastuzumab-induced cardiotoxicity by logistic regression analysis ($P = 2.82 \times 10^{-4} - 4.15 \times 10^{-3}$, Supplementary Table 1), we investigated combined effects of the five loci on the risk of cardiotoxicity in patients treated with trastuzumab by using a scoring system. For the predictive scoring system, each patient was scored considering the total number of risk alleles and genotypes. We gave a score of 2 to individuals homozygous for risk allele and 0 to individuals with the other genotypes at rs28415722 and rs7406710 because these SNPs showed smallest $P$ value in recessive-inheritance model. We further gave a score of 2 to individuals homozygous for risk allele, 1 to individuals heterozygous for risk allele and 0 to individuals homozygous for non-risk allele at rs9316695, rs11932853 and rs8032978. The patients’ scores were obtained by summing up the scores of the five SNPs. The patients used in this study were classified into nine prediction score groups (score 0-8, Table 3). The proportion of patients with trastuzumab-induced cardiotoxicity was likely to be increased in groups with higher prediction scores; the incidences of the
cardiotoxicity were 1.8% (8/441) in the score 0-4 group, 36.4% (8/22) in the score 5 group, 22.2% (2/9) in the score 6 group, 75.0% (3/4) in the score 7 group and 80.0% (4/5) in the score 8 group (Table 3). We further classify the patients into two groups; score 0-4 group and 5-8 group, considering the best threshold for the diagnostic test (sensitivity, 68.0% and specificity, 95.0%). The incidences of trastuzumab-induced cardiotoxicity in score 5-8 group were 45.8% (11/24), 37.5% (6/16) and 42.5% (17/40), whereas those in score 0-4 group were 0% (0/244), 4.1% (8/197) and 1.8% (8/441) in the GWAS, replication and combined stage, respectively, suggesting that the predictive scoring system using the above five SNPs could predict the risk of cardiotoxicity prior to initiation of trastuzumab therapy ($P_{combined} = 7.82 \times 10^{-15}$, OR=40.0, 95% CI=15.6-102.3, Table 3, Fig. 2). Although we investigated the relationship between the prediction score and cardiotoxicity grade (LVEF) or time-to-event (cardiotoxicity), we could not observe significant relationship between them (Supplementary Fig.3). Moreover, we performed risk stratification analysis based on medication for cardiovascular disease because it showed significant association with cardiotoxicity ($P = 1.80 \times 10^{-4}$, Supplementary Table 2). However, this predictive scoring system appeared to be medication-nonspecific (independent of medication) because the associations of cardiotoxicity with prediction scores were comparable between medicated and unmedicated patients (medicated; $P = 6.82 \times 10^{-6}$, unmedicated; $P = 1.01 \times 10^{-6}$, Supplementary Table 9).

**Discussion**

The safe and effective treatment with trastuzumab based on individual germline and/or somatic genetic information is important to provide personalized medicine for
patients with HER2-positive cancer. Although a GWAS of cardiotoxicity in Adjuvant Trastuzumab Trial (NCCTG N9831) in Caucasian population has already been reported, this study is the first GWAS which attempts to identify common genetic variants associated with trastuzumab-induced cardiotoxicity in Japanese population. The genome-wide significant SNP was not identified in this study, however, we identified five independent loci showing possible association with trastuzumab-induced cardiotoxicity (Table 2, Fig. 1). Interestingly, of the above loci, proprotein convertase subtilisin/kexin type 6 (PCSK6) and small nuclear ribonucleoprotein polypeptide A (SNRPA1), which have been suggested to promote the proliferation of breast cancer and colon cancer, respectively, were included in the genomic region of chromosome 15q26.3 (Fig. 1E). 22,23) We observed significantly higher incidence of trastuzumab-induced cardiotoxicity in patients with higher score (≥ 5) compared to those with lower score (≤ 4) in predictive scoring system using the five marker SNPs (Fig. 2, Table 2, 3). This result suggests the potential to improve the ability of physicians to avoid the risk of trastuzumab-induced cardiotoxicity based on the result of this predictive scoring system for the treatment of patients with HER2-positive cancer.

rs9316695, the marker SNP on chromosome 13q14.3 (P_{combined} = 6.00 \times 10^{-6}, OR= 4.46, Table 2), which showed the strongest association in a combined analysis of GWAS and replication study, was located in intergenic region, and long intergenic non-protein coding RNA 558 (LINC00558) was included in this region (Fig. 1A). NOTCH3, which encodes the notch receptor 3, has been reported to be one of the target genes of LINC00558,24) and NOTCH3 signaling pathways are significantly activated in ErbB2-negative cells.25,26) The Notch signaling system is present in the cardiomyocytes and suggested to play a role in myocardial remodeling in heart
failure. Hence, genetic variation in LINC00558 might contribute to interindividual variability in trastuzumab-induced cardiotoxicity through effects on Notch signaling in the stressed (inactivation of HER2 signaling) cardiomyocyte, although further validation studies will be needed to prove the true association.

rs8032978 on chromosome 15q26.3, which is one of the five marker SNPs in this study \( (P_{\text{combined}} = 1.60 \times 10^{-4}, \ OR= 5.83, \ \text{Table 2}) \), was located 44 kb downstream from the PCSK6 gene (Fig. 1E). PCSK6 activates corin which is a type-II transmembrane serine protease found mainly in the heart and converts it to a two-chain active enzyme through cleavage at Arg-801. Corin is known to play an essential role in the regulation of water and salt balance by converting natriuretic peptides to their active form, which promotes vasodilation, diuresis and natriuresis, and prevent cardiac remodeling, in patients with heart failure. These lines of evidence suggest that impaired activity and/or gene expression of PCSK6 might contribute to the inability to recover the cardiac function in damaged (inactivation of HER2 signaling) heart due to insufficient activation of natriuretic peptides.

To effectively clarify the mechanism of trastuzumab-induced cardiotoxicity, we performed expression quantitative trait locus (cis-eQTL) analysis for the five candidate loci. We observed cis-regulatory effect of rs7406710, which is located 1.5kb downstream from fascin actin-bundling protein 2, retinal (FSCN2) on chromosome 17q25.3 (Fig. 1C), on gene expression level of FSCN2 \( (P = 3.20 \times 10^{-10}, \ \text{Supplementary Fig.4}) \). FSCN2 is actin-binding proteins that cross-link filamentous actin into tightly packed parallel bundles. FSCN2 can bind with multiple actins and has been reported to bind with beta-, gamma-, and alpha-actin which is highly expressed in cardiomyocyte. Although the relationship between HER2 signaling and actin...
crosslinking protein, FSCN2, in myocardium is not yet clarified, lower expression of FSCN2 in cardiomyocyte might impair cardiac contraction under the condition of HER2 inhibition (Supplementary Fig.4). Further studies will be required to clarify the functional associations of FSCN2 with cardiotoxicity in patients treated with trastuzumab.

The candidate gene association studies previously reported that rs1058808 (P1170A) and rs1136201 (I655V) in ERBB2 (HER2) could be candidates that confer susceptibility to trastuzumab-induced cardiotoxicity. In our GWAS, the two SNPs showed no association with trastuzumab-induced cardiotoxicity (rs1058808; $P = 1.21 \times 10^{-1}$, rs1136201; $P = 6.86 \times 10^{-1}$). Moreover, a first GWAS of trastuzumab-induced cardiotoxicity in Caucasian population suggested that SNPs at six loci, LIM domain binding 2 (LDB2), BMP/retinoic acid inducible neural specific 1 (BRINP1), chr6 intergenic, RAB22A, member RAS oncogene family (RAB22A), transient receptor potential cation channel subfamily C member 6 (TRPC6) and long intergenic non-protein coding RNA 1060 (LINC01060), were associated with a decline in LVEF ($P = 8.93 \times 10^{-8}$ - $7.73 \times 10^{-6}$) in patients who received chemotherapy plus trastuzumab.

In our GWAS, the SNPs in these candidate genes showed no or weak association with trastuzumab-induced cardiotoxicity ($P = 3.33 \times 10^{-3}$ - $9.99 \times 10^{-1}$). Although the small sample size in our study might be one of the causes of the discrepant results, interethnic difference in allele frequency of the above SNPs could also affect the discrepancy. Moreover, mechanism of action involved in the trastuzumab-induced cardiotoxicity might differ between ethnic groups. The $P$ values of the SNPs in ATP binding cassette subfamily B member 1 (ABCB1), solute carrier family 28 member 3 (SLC28A3), carbonyl reductase 3 (CBR3), neutrophil cytosolic factor (NCF4), Rac family small
GTPase 2 (RAC2), retinoic acid receptor gamma (RARG) and UDP glucuronosyltransferase family 1 member A6 (UGT1A6), which were reported to be associated with anthracycline-induced cardiotoxicity, were from 1.68 x 10^{-2} to 9.99 x 10^{-1}, indicating no or weak association in our GWAS. The sample size used in our study might not be enough to detect true associations of the above SNPs because the effect sizes of these candidate genes were not large.16,17,21,39-42)

In conclusion, our GWAS using 481 patients treated with trastuzumab identified five novel candidate loci, chromosome 4q25, 13q14.3, 17q25.3, and 2 independent loci on 15q26.3, which were associated with trastuzumab-induced cardiotoxicity. Furthermore, the combined analysis of the five SNP loci revealed that the predictive score based on risk genotypes was significantly associated with trastuzumab-induced cardiotoxicity. These findings provide new insights into personalized trastuzumab therapy for patients with HER2 positive cancer. However, larger-scale replication study and further functional analysis are required to verify our results and to clarify their biological mechanisms which have effects on trastuzumab-induced cardiotoxicity.

Acknowledgments

We express our heartfelt gratitude to all the study participants. We thank Mr. Satoshi Kasahara, Ms. Aya Imai and Ms. Emi Shigino for their contribution to the collection of clinical information, and all staffs in Nakagami Hospital, St. Marianna University School of Medicine, and National Cancer Center Biobank for their contribution to the sample collection.
Conflict of Interest

The authors declare no conflict of interest.

Supplementary Materials

The online version of this article contains supplementary materials.
References


20) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.*, **81**, 559-575 (2007).


29) Khoury EE, Kinaneh S, Aronson D, Amir O, Ghanim D, Volinsky N, Azzam Z,


### Table 1. Patient demographics and clinical characteristics.

<table>
<thead>
<tr>
<th></th>
<th>GWAS (N=288)</th>
<th>Replication (N=213)</th>
<th>Total (N=483)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>59.0</td>
<td>57.0</td>
<td>55.5</td>
</tr>
<tr>
<td><strong>Primary cancer site</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>10 (90.9%)</td>
<td>185 (72.0%)</td>
<td>14 (100.0%)</td>
</tr>
<tr>
<td>Gastric</td>
<td>1 (9.1%)</td>
<td>68 (26.5%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>0 (0.0%)</td>
<td>4 (1.6%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>Pretreatment with anthracycline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>9 (81.8%)</td>
<td>148 (57.6%)</td>
<td>10 (71.4%)</td>
</tr>
<tr>
<td>No</td>
<td>2 (18.2%)</td>
<td>109 (42.4%)</td>
<td>4 (28.6%)</td>
</tr>
<tr>
<td><strong>HER-2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>10 (90.9%)</td>
<td>249 (96.9%)</td>
<td>14 (100.0%)</td>
</tr>
<tr>
<td>Negative</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (9.1%)</td>
<td>8 (3.1%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>ER status (Breast cancer)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>6 (60.0%)</td>
<td>119 (46.3%)</td>
<td>11 (78.6%)</td>
</tr>
<tr>
<td>Negative</td>
<td>3 (30.0%)</td>
<td>64 (25.6%)</td>
<td>3 (21.4%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (10.0%)</td>
<td>2 (1.1%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td><strong>PR status (Breast cancer)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>6 (60.0%)</td>
<td>104 (55.2%)</td>
<td>8 (57.1%)</td>
</tr>
<tr>
<td>Negative</td>
<td>4 (40.0%)</td>
<td>80 (43.2%)</td>
<td>6 (42.9%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>0 (0.0%)</td>
<td>1 (0.5%)</td>
<td>0 (0.0%)</td>
</tr>
</tbody>
</table>

ER, estrogen receptor; PR, progesterone receptor.
Table 2. Summary of association results of GWAS and replication study for five loci.

<table>
<thead>
<tr>
<th>Chromosomal region</th>
<th>Position</th>
<th>SNP</th>
<th>Gene</th>
<th>Allele[a][b]</th>
<th>Risk allele</th>
<th>Stage</th>
<th>Genotyped/Imputed</th>
<th>Case Genotype</th>
<th>Control Genotype</th>
<th>P value</th>
<th>Allelic Odds ratio (95% CI)</th>
<th>Dominant Odds ratio (95% CI)</th>
<th>Recessive Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>13q14.3</td>
<td>54,615,773</td>
<td>rs9316695</td>
<td>No gene</td>
<td>C/A</td>
<td>A</td>
<td>GWAS</td>
<td>Replication combined</td>
<td>3 5 3 0.09</td>
<td>199 55 3 0.12</td>
<td>2.59 x 10^{-5}</td>
<td>8.22 x 10^{-5}</td>
<td>9.71 x 10^{-4}</td>
<td></td>
</tr>
<tr>
<td>15q26.3</td>
<td>98,609,746</td>
<td>rs28415722</td>
<td>No gene</td>
<td>G/A</td>
<td>A</td>
<td>GWAS</td>
<td>Imputed</td>
<td>1 2 8 0.82</td>
<td>95 119 43 0.40</td>
<td>1.97 x 10^{-4}</td>
<td>1.04 x 10^{-4}</td>
<td>1.10 x 10^{-4}</td>
<td></td>
</tr>
<tr>
<td>17q25.3</td>
<td>79,505,624</td>
<td>rs7406710</td>
<td>No gene</td>
<td>C/T</td>
<td>C</td>
<td>GWAS</td>
<td>Combined</td>
<td>11 0 0 1.00</td>
<td>111 119 27 0.66</td>
<td>2.38 x 10^{-4}</td>
<td>6.10 x 10^{-4}</td>
<td>1.35 x 10^{-4}</td>
<td></td>
</tr>
<tr>
<td>4q25</td>
<td>112,024,388</td>
<td>rs11932853</td>
<td>No gene</td>
<td>T/C</td>
<td>T</td>
<td>GWAS</td>
<td>Imputed</td>
<td>7 3 1 0.77</td>
<td>31 129 97 0.37</td>
<td>2.26 x 10^{-4}</td>
<td>6.04 x 10^{-4}</td>
<td>1.47 x 10^{-4}</td>
<td></td>
</tr>
<tr>
<td>15q26.3</td>
<td>101,799,790</td>
<td>rs8032978</td>
<td>No gene</td>
<td>A/G</td>
<td>G</td>
<td>GWAS</td>
<td>Replication combined</td>
<td>6 4 1 0.27</td>
<td>239 18 0 0.04</td>
<td>1.95 x 10^{-4}</td>
<td>9.84 x 10^{-4}</td>
<td>4.10 x 10^{-7}</td>
<td></td>
</tr>
</tbody>
</table>

[a] Based on GRCh37 genome assembly.

[b] Reference allele (GRCh37) was defined as allele 1.

[c] Odds ratios were shown for the model with minimum P values.

[d] Genotype of rs28415722 was validated using TaqMan SNP genotyping assay.

GWAS, genome-wide association study; SNP, single nucleotide polymorphism; RAF, risk allele frequency; CI, confidence interval; NA, not available.
Table 3. Prediction scores of trastuzumab-induced cardiotoxicity using rs9316695, rs28415722, rs7406710, rs11932853 and rs8032978.

<table>
<thead>
<tr>
<th>Score</th>
<th>N (%)</th>
<th>% of Case</th>
<th>P value</th>
<th>Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case</td>
<td>Control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0 (0.0)</td>
<td>53 (11.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0 (0.0)</td>
<td>93 (20.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3 (12.0)</td>
<td>109 (23.9)</td>
<td>1.8</td>
<td>1.00 (reference)</td>
</tr>
<tr>
<td>3</td>
<td>2 (8.0)</td>
<td>111 (24.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3 (12.0)</td>
<td>67 (14.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8 (32.0)</td>
<td>14 (3.1)</td>
<td>36.4</td>
<td>6.68 x 10^{-3}</td>
</tr>
<tr>
<td>6</td>
<td>2 (8.0)</td>
<td>7 (1.5)</td>
<td>22.2</td>
<td>1.47 x 10^{-2}</td>
</tr>
<tr>
<td>7</td>
<td>3 (12.0)</td>
<td>1 (0.2)</td>
<td>75.0</td>
<td>4.46 x 10^{-5}</td>
</tr>
<tr>
<td>8</td>
<td>4 (16.0)</td>
<td>1 (0.2)</td>
<td>80.0</td>
<td>1.50 x 10^{-6}</td>
</tr>
</tbody>
</table>

CI, confidence interval.
Fig. 1. Regional association plots of five loci associated with trastuzumab-induced cardiotoxicity.

Upper panel; $P$ values of genotyped SNPs (circle) and imputed SNPs (square) are plotted (as $-\log_{10}P$ value) against their physical location on chromosome 13q14.3(A), 15q26.3(B), 17q25.3(C), 4q25(D) and 15q26.3(E). The genetic recombination rates estimated from 1,000 Genomes samples (JPT + CHB) are shown with a blue line.
SNP's color indicates LD with rs9316695 (A), rs28415722 (B), rs7406710 (C), rs11932853 (D) and rs8032978 (E) according to a scale from $r^2 = 0$ to 1 based on pair-wise $r^2$ values from ASN data in 1,000 Genomes Project. Lower panel; gene annotations from the University of California Santa Cruz genome browser.
Fig. 2. Predictive risk score and incidence of trastuzumab-induced cardiotoxicity.

The incidence of trastuzumab-induced cardiotoxicity in patients with score ≥ 5 was significantly higher compared to those in patients with score ≤ 4 (Fisher’s exact test; $P = 7.82 \times 10^{-15}$). Stippled bar: GWAS stage, striped bar: replication stage, and black bar: combined stage.