International Group 1-98 study, in which strong departures from HWE (to a magnitude of 10^{-92}) were observed, leading to a call for retraction of this article.[3,12,16]

For criterion 1, 49.9% of our patient DNA samples originated from blood, 21.6% from fresh-frozen tissues, and 28.5% from FFPE tissues. For criterion 2, 55.0% samples originated from blood, 20.5% were fresh-frozen tissues, and 24.5% from FFPE tissues. For criterion 3, 50.9% of DNA samples originated from blood, 31.9% from fresh-frozen tumor, 13.4% from FFPE tumor tissues, and 3.5% from FFPE normal tissue. Although we cannot exclude the presence of somatic events leading to misclassification of *CYP2D6* genotype, as evident from HWE deviation identified in data from

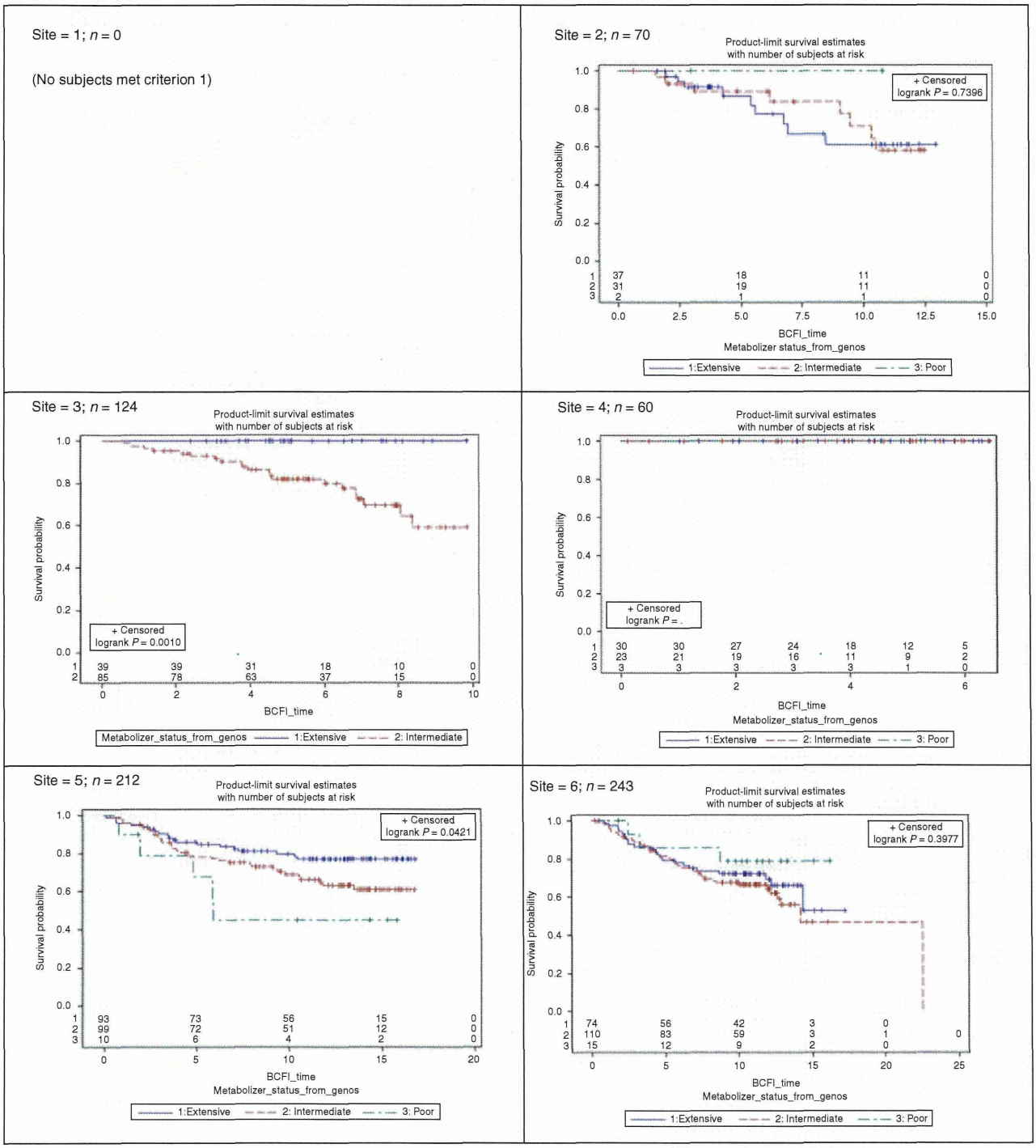

Figure 3 Site-specific effects of CYP2D6 metabolizer status on clinical outcomes for subjects meeting inclusion criterion 1 (outcome = breast cancer–free interval (BCFI)).

some sites, comprehensive testing for HWE did not reveal significant violations across most sites. Moreover, the extent of deviation from HWE in the **4* allele was not associated with sites that evinced less clinical benefit from tamoxifen in patients who were assessed to be PMs in terms of their *CYP2D6* status. This suggests that genotyping errors are unlikely to be a major issue in our analyses.

Our findings are subject to the shortcomings commonly encountered when performing retrospective "biomarker" studies. In our study, most sites were unable to collect or control for the factors known to alter endoxifen exposure, including dose and duration of tamoxifen administration and patients' adherence to the regimen. Although tamoxifen adherence is increasingly recognized as a critical factor for drug efficacy,[23] most studies evaluating tamoxifen biomarkers have not controlled for adherence. Other confounders include limited *CYP2D6* allele coverage and lack of information

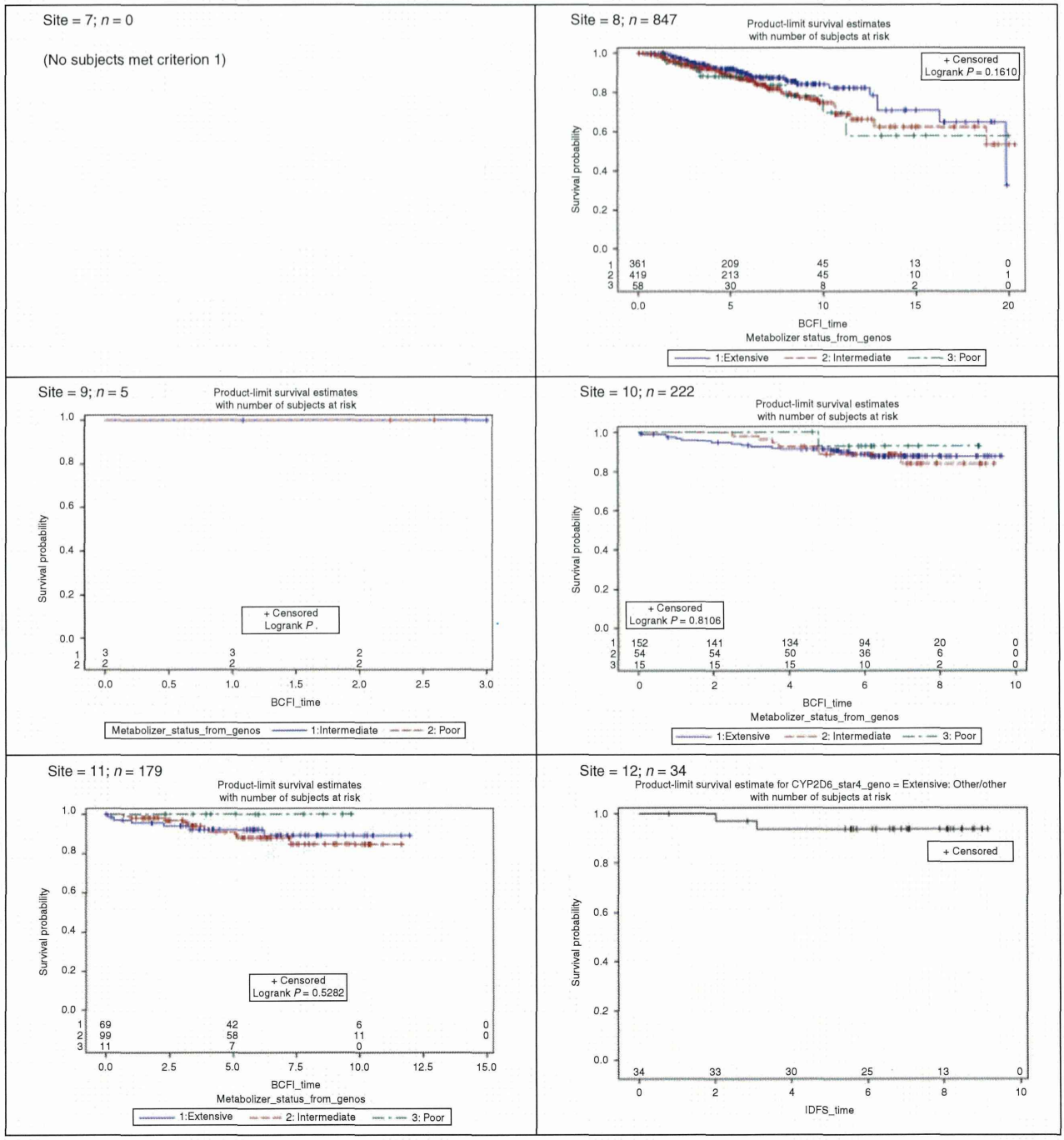

Figure 3 Continued

regarding the coadministration of CYP2D6 inhibitors, leading to potential misclassification of the CYP2D6 drug metabolism phenotype. Therefore, our meta-analysis results depend heavily on which subgroup of patients we include. If we accept that utmost precautions must be applied to avoid the distortion of results from influences derived from the aforementioned shortcomings, it follows that merely increasing the numbers of subjects without controlling the quality of input data, as done in our preliminary overview analysis,[5] may result in heterogeneity that masks the effect of a pharmacokinetic biomarker such as *CYP2D6*. From this, we conclude that until results from prospective adjuvant studies are available, women who meet criterion 1 as established in this and other independent cohorts (ABCSG 8) should be counseled regarding the potential impact of *CYP2D6* on the effectiveness of adjuvant tamoxifen, and potent CYP2D6 inhibitors should be avoided in these patients. Prospective adjuvant studies are needed to determine whether genotype-guided selection of hormonal therapy will improve the outcomes of women with early-stage ER-positive breast cancer, and results from ongoing prospective studies in the metastatic setting are eagerly awaited. A similarly motivated study on warfarin is currently being conducted in the Clarification of Optimal Anticoagulation through Genetics trial.[24]

By strict clinical and genotype criteria, reduced CYP2D6 metabolism is associated with a higher risk of recurrence (as measured by IDFS) in tamoxifen-treated women. However, the heterogeneity observed across sites contributing data to the ITPC points to the likely influence of critical confounding factors unlikely to be controllable in global retrospective studies. This study demonstrates the complexity of performing a retrospective biomarker study that focuses on the genetic factors that affect exposure to an active metabolite, endoxifen, for a drug, tamoxifen, administered for 5 years. Our observation that <50% of the patients in this study met the basic eligibility criteria—in terms of similar disease, treatment, and control for critical pharmacological factors such as dose and duration of tamoxifen—provides insight into possible reasons for the discrepancies in the literature on *CYP2D6* and tamoxifen. Although *CYP2D6* is a predictor of IDFS in a subset of patients treated with tamoxifen, the lack of an effect in the entire heterogeneous study population suggests that prospective studies are necessary to finally establish whether genotype-guided selection of hormonal therapy improves clinical outcomes of women with ER-positive breast cancer.

METHODS

Data collection and study cohorts. The ITPC invited any research group from across the world that had published or unpublished *CYP2D6* data to participate in this meta-analysis. The ITPC comprises 12 research projects for a total of 4,973 breast cancer patients treated with tamoxifen. This retrospective study does not include a control group not treated with tamoxifen. These data were curated at the PharmGKB (Pharmacogenomics Knowledge Base, http://www.pharmgkb.org). Consent for participation in the ITPC and DNA collection, *CYP2D6* genetic testing, and submission of data was obtained under local ethical review board permissions.

We collected information on clinical factors previously shown to be associated with breast cancer therapy and prognosis that were available from the information received from the sites. These data included demographic characteristics, cancer history, cancer recurrence, use of other therapies, use of concomitant medications known to affect *CYP2D6* phenotype, ER status, and classic prognostic factors such as tumor size and number of affected lymph nodes. Information was also collected regarding the presence of *CYP2D6* genetic variants (**2*, **3*, **4*, **5*, **6*, **10*, **17*, and **41*, categorized by their DNA sources), for which coverage of these alleles varied by site. For 1,635 subjects, *CYP2D6* variants assessable from blood DNA using the AmpliChip CYP450 test (Roche) were collected. A complete list of the information collected is detailed in **S1–S3** online, including the project-specific *CYP2D6* genotype assays used and the DNA source. Independent confirmation of *CYP2D6* genotypes was not performed owing to lack of access to subjects' samples. The clinical outcome variable was either breast cancer–free interval or IDFS, as previously defined.[25] The complete data set of genotypes and clinical variables is available at http://www.pharmgkb.org.

Statistical analysis. Because the ITPC was not a prospectively defined multicenter study with a common protocol, there is potential for considerable study-to-study heterogeneity. Therefore, we did not analyze the combined data as a single series even though we had access to individual-level data from all studies. Rather, we applied a random-effects meta-analysis strategy. This provided estimates of the effect of *CYP2D6* in each study's data separately, allowing us to examine the consistency of the results across sites. The meta-analysis is a two-stage procedure. In the first stage, we fit proportional-hazards models to the data from each of the ITPC sites separately, predicting clinical outcome after surgery from *CYP2D6* genotype and other relevant covariates. These analyses produced a set of 12 parameter estimates of the HRs of *CYP2D6* genotypes on outcome, along with their corresponding SEs (one for each site). In the second stage, we used a random-effects meta-analysis procedure[26] to test for study heterogeneity (i.e., whether the 12 studies met the assumptions of the meta-analysis sufficiently so as to be combinable using that method). When the heterogeneity was not significant, we combined the log-HRs into a single, meta-analysis estimate of the effect of *CYP2D6* on tamoxifen-treated recurrence and/or survival outcomes. The DerSimonian and Laird method also provides a penalty in its test of overall association for moderate levels of study-to-study heterogeneity (i.e., for heterogeneity that is not so severe as to be statistically significant). This method is therefore conservative in its conclusions when heterogeneity is a potential issue.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/cpt

ACKNOWLEDGMENTS

The authors acknowledge useful conversations with Donald A. Berry (The University of Texas MD Anderson Cancer Center). The complete data set of genotypes and clinical variables, analysis codes, and full analyses is available to registered PharmGKB users at http://www.pharmgkb.org. We are grateful to all breast cancer patients for their participation. We thank the physicians and other hospital staff, scientists, research assistants, and study staff who contributed to the patient recruitment, data collection, and sample preparation.

This work was supported by grants from the National Institutes of Health (National Institute of General Medical Sciences, the National Cancer Institute, and the National Institute of Child Health and Human Development R24 GM61374, CA-25224, CA-37404, CA-15083, CA-35113, CA-35269, CA-63849, CA35103, CA-35195, CA-35272, CA-35101, CA-37417, CA-35415, CA-52352, CA-35448, CA-60274, 1R01CA133049-01, U19 GM61388, CA-116201, CA-58207, CA-51008, T32 GM007546,

and CA051008; Breast Cancer Research (Scotland); Tayside Tissue Bank; California Breast Cancer Research Program grant 14OB-0166; Cancer Research UK; Deutsches Krebsforschungszentrum, Heidelberg, Germany; Robert Bosch Foundation, Stuttgart, Germany; 7FP EU Marie Curie Initial Training Network "FightingDrugFailure" (GA 238132); GrantStichting Emmanuel van der Schueren (scientific partner of the Vlaamse Liga tegen Kanker); and the National Project for Personalized Genomic Medicine, Ministry for Health & Welfare, Republic of Korea (A111218-PG02). W.S. is supported by grant Deutsche Forschungsgemeinschaft SCHR 1323/2-1, Germany.

The institutions in the International Tamoxifen Pharmacogenomics Consortium are listed below.:

Writing committee:

Michael A. Province, Matthew P. Goetz, David A. Flockhart, Hiltrud Brauch, and Teri E. Klein

Data contributors:

Deutsches Krebsforschungszentrum, Heidelberg, Germany: Ute Hamann, Julia Boländer, and Hans-Ulrich Ulmer, Frauenklinik Städtisches Klinikum Karlsruhe, Germany

Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology Stuttgart and University of Tuebingen, Tuebingen, Germany: Hiltrud Brauch, Werner Schroth, Matthias Schwab, Stefan Winter, Michel Eichelbaum, Peter Fritz, and Wolfgang Simon

Division of Research, Federal Institute for Drugs and Medical Devices, and University of Bonn Medical Faculty, Bonn, Germany: Julia C. Stingl

Indiana University and University of Michigan, USA: David A. Flockhart and Anne T. Nguyen

Inje University College of Medicine and Department of Clinical Pharmacology, Busan Paik Hospital, Busan, Korea: Jae-Gook Shin and Ji-Yeob Choi

Mayo Clinic in Rochester, Minnesota, USA: Matthew P. Goetz, James N. Ingle, Vera J. Suman, and Richard M. Weinshilboum

Ninewells Hospital and Medical School, Dundee, UK: Colin A. Purdie and Lee B. Jordan

Örebro University, Sweden: Pia Wegman

RIKEN Center for Genomic Medicine and University of Tokyo, Japan: Hitoshi Zembutsu, Taisei Mushiroda, and Kazuma Kiyotani

Roswell Park Cancer Institute, USA: Christine B. Ambrosone

University Breast Center Franconia, University Hospital Erlangen, Erlangen, Germany: Peter A Fasching, Reiner Strick, and Matthias W Beckmann

University Hospitals Leuven, Belgium: Patrick Neven and Anne-Sophie Dieudonné

University Institute of Medical and Chemical Laboratory Diagnostics, Paracelsus Private Medical University, Salzburg, Austria: Elisabeth Haschke-Becher

University of California, San Francisco, San Francisco, USA: Alan H.B. Wu, Wendy Lorizio, and Elad Ziv

University of Dundee, UK: Philip Quinlan, Lee B. Jordan, and Colin A. Purdie

Department of Gynecology and Obstetrics, University of Mainz, Mainz, Germany: Marcus Schmidt and Heinz Koelbl

University of Manchester, Manchester, UK: William G. Newman, Roberta Ferraldeschi, Anthony Howell, and Ayse Latif

Vesalius Research Center, Catholic University of Leuven, Leuven, Belgium: Diether Lambrechts

Yonsei University Health System, Korea: Byeong-Woo Park

Statistical Analysis:

Mayo Clinic in Rochester, MN, USA: Vera J. Suman

Stanford University, Stanford, CA, USA: Teri E. Klein and Ryan Whaley

Washington University in St. Louis, MO, USA: Michael A. Province

Data Curation:

Stanford University, Stanford, CA, USA: Teri E. Klein, Ryan Whaley, Joan M. Hebert, and Li Gong

Site Principal Investigators:

Russ B. Altman and Teri E. Klein (Stanford University), Christine A. Ambrosone (Roswell Park Cancer Institute), Hiltrud Brauch (IKP Stuttgart), Jae-Gook Shin (Inje University), Matthew P. Goetz (Mayo Clinic), William G. Newman (University of Manchester), Patrick Neven (Belgium), Alastair M. Thompson (University of Dundee), Sten Wingren (Örebro University), Hitoshi Zembutsu (University of Tokyo), and Elad Ziv (University of California, San Francisco)

Steering Committee:

Russ B. Altman, Hiltrud Brauch, Michel Eichelbaum, David A. Flockhart, Matthew P. Goetz, James N. Ingle, V. Craig Jordan, Teri E. Klein, Kent Osborne, Yusuke Nakamura, and Richard M. Weinshilboum

Participating sites:

The Christie NHS Foundation Trust, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY, USA

Department of Clinical Medicine, Örebro University, Örebro, Sweden

Department of Genetics, Stanford University, Stanford, CA, USA

Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

Department of Gynecology and Obstetrics, University Hospitals Leuven, Leuven, Belgium

Department of Oncology, Catholic University Leuven, Leuven, Belgium

Department of Medicine/GIM and Clinical Pharmacology and Experimental Therapeutics, and Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, USA

Department of Molecular Pharmacology & Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA

Department of Oncology, Georgetown University, Washington, DC, USA

Department of Oncology, Mayo Clinic, Rochester, MN, USA

Department of Oncology and Pharmacology, Mayo Clinic, Rochester, MN, USA

Department of Pharmacology; Mayo Clinic, Rochester, MN, USA

Department of Pharmacology and Pharmacogenomics Research Center, Inje University College of Medicine and Department of Clinical Pharmacology, Inje University, Busan Paik Hospital, Busan, Korea

Department of Surgery, Yonsei University Health System, Seoul, Korea

Deutsches Krebsforschungszentrum, Heidelberg, Germany

Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA

Division of Clinical Pharmacology, School of Medicine, Indiana University, IN, USA

Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology Stuttgart, and University of Tuebingen, Germany

Dundee Cancer Centre, Dundee, UK

Genetic Medicine, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

Laboratory for Pharmacogenetics, RIKEN Center for Genomic Medicine, Yokohama, Japan

Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum, Heidelberg, Germany

Ninewells Hospital and Medical School, Dundee, UK

CONFLICT OF INTEREST

A.M.T. and W.G.N. report Roche funding for genotyping H.B. and M. Schwab report that they have initiated scientific collaborations in 2009 with Roche Molecular Diagnostics and Siemens Healthcare Diagnostics Products, respectively P.A.F. and M.W.B. report Novartis research funding. M.-T.M.L. is a paid consultant of YongLin Health Foundation. R.B.A. is a founder, equity holder, and consultant for Personalis. The other authors declared no conflict of interest.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

✓ There has been extensive controversy with regard to the association between *CYP2D6* genetic variants and the clinical outcomes of tamoxifen use.

WHAT QUESTION DID THIS STUDY ADDRESS?

✓ The ITPC was established to address this controversy and to determine the association of *CYP2D6* status with IDFS in tamoxifen-treated early-stage, ER-positive breast cancer.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE

✓ We found that *CYP2D6* genotype was associated with a higher risk of recurrence in patients meeting the strict criterion. However, the observation of substantial heterogeneity in cohorts 2 and 3 suggests that study design factors that cannot be controlled retrospectively may obscure the predictive utility of *CYP2D6* genotype. This study demonstrates the complexity of performing a retrospective biomarker study.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS

✓ Although *CYP2D6* is a predictor of IDFS in a subset of patients treated with tamoxifen monotherapy, the lack of an effect in the entire heterogeneous study population suggests that prospective studies are necessary to fully establish the value of *CYP2D6* genotyping in tamoxifen therapy.

1. Jordan, V.C. Tamoxifen: a most unlikely pioneering medicine. *Nat. Rev. Drug Discov.* **2**, 205–213 (2003).
2. Davies, C. *et al.* Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet* **378**, 771–784 (2011).
3. Johnson, M.D. *et al.* Pharmacological characterization of 4-hydroxy-N-desmethyl tamoxifen, a novel active metabolite of tamoxifen. *Breast Cancer Res. Treat.* **85**, 151–159 (2004).
4. Stearns, V. *et al.* Active tamoxifen metabolite plasma concentrations after coadministration of tamoxifen and the selective serotonin reuptake inhibitor paroxetine. *J. Natl. Cancer Inst.* **95**, 1758–1764 (2003).
5. Goetz, M.P., Berry, D.A. & Klein, T.E. 31136 Findings from the International Tamoxifen Pharmacogenomics Consortium. Proceedings of the San Antonio Breast Cancer Symposium, San Antonio, Texas, 9–13 December 2009. Abstract 33.
6. Mürdter, T.E. *et al.*; German Tamoxifen and AI Clinicians Group. Activity levels of tamoxifen metabolites at the estrogen receptor and the impact of genetic polymorphisms of phase I and II enzymes on their concentration levels in plasma. *Clin. Pharmacol. Ther.* **89**, 708–717 (2011).
7. Irvin, W.J. Jr *et al.* Genotype-guided tamoxifen dosing increases active metabolite exposure in women with reduced CYP2D6 metabolism: a multicenter study. *J. Clin. Oncol.* **29**, 3232–3239 (2011).
8. Rae, J.M. *et al.*; COBRA investigators. Cytochrome P450 2D6 activity predicts discontinuation of tamoxifen therapy in breast cancer patients. *Pharmacogenomics J.* **9**, 258–264 (2009).
9. Schroth, W. *et al.* Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA* **302**, 1429–1436 (2009).
10. Goetz, M.P. *et al.* Pharmacogenetics of tamoxifen biotransformation is associated with clinical outcomes of efficacy and hot flashes. *J. Clin. Oncol.* **23**, 9312–9318 (2005).
11. Rae, J.M. *et al.*; ATAC trialists. CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. *J. Natl. Cancer Inst.* **104**, 452–460 (2012).
12. Regan, M.M. *et al.*; Breast International Group (BIG) 1-98 Collaborative Group. CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the breast international group 1-98 trial. *J. Natl. Cancer Inst.* **104**, 441–451 (2012).
13. Brauch, H. *et al.* Tamoxifen use in postmenopausal breast cancer: CYP2D6 matters. *J. Clin. Oncol.* **31**, 176–180 (2013).
14. Nakamura, Y., Ratain, M.J., Cox, N.J., McLeod, H.L., Kroetz, D.L. & Flockhart, D.A. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial. *J. Natl. Cancer Inst.* **104**, 1264; author reply 1266–1268 (2012).
15. Stanton, V. Jr. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial. *J. Natl. Cancer Inst.* **104**, 1265–6; author reply 1266–8 (2012).
16. Goldberg, P. Experts claim errors in breast cancer study, demand retraction of practice-changing paper. *The Cancer Letter* **38**, 1–11 (2012).
17. Goetz, M.P. *et al.* CYP2D6 metabolism and patient outcome in the Austrian Breast and Colorectal Cancer Study Group trial (ABCSG) 8. *Clin. Cancer Res.* **19**, 500–507 (2013).
18. Schroth, W. *et al.* CYP2D6 polymorphisms as predictors of outcome in breast cancer patients treated with tamoxifen: expanded polymorphism coverage improves risk stratification. *Clin. Cancer Res.* **16**, 4468–4477 (2010).
19. Zanger, U.M., Turpeinen, M., Klein, K. & Schwab, M. Functional pharmacogenetics/genomics of human cytochromes P450 involved in drug biotransformation. *Anal. Bioanal. Chem.* **392**, 1093–1108 (2008).
20. Castells, A., Gusella, J.F., Ramesh, V. & Rustgi, A.K. A region of deletion on chromosome 22q13 is common to human breast and colorectal cancers. *Cancer Res.* **60**, 2836–2839 (2000).
21. Hirano, A. *et al.* Allelic losses of loci at 3p25.1, 8p22, 13q12, 17p13.3, and 22q13 correlate with postoperative recurrence in breast cancer. *Clin. Cancer Res.* **7**, 876–882 (2001).
22. Loo, L.W. *et al.* Differential patterns of allelic loss in estrogen receptor-positive infiltrating lobular and ductal breast cancer. *Genes. Chromosomes Cancer* **47**, 1049–1066 (2008).
23. Thompson, A.M. *et al.* Comprehensive CYP2D6 genotype and adherence affect outcome in breast cancer patients treated with tamoxifen monotherapy. *Breast Cancer Res. Treat.* **125**, 279–287 (2011).
24. French, B. *et al.*; COAG (Clarification of Optimal Anticoagulation through Genetics) Investigators. Statistical design of personalized medicine interventions: the Clarification of Optimal Anticoagulation through Genetics (COAG) trial. *Trials* **11**, 108 (2010).
25. Hudis, C.A. *et al.* Proposal for standardized definitions for efficacy end points in adjuvant breast cancer trials: the STEEP system. *J. Clin. Oncol.* **25**, 2127–2132 (2007).
26. DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177–188 (1986).

Chung et al. Breast Cancer Research 2013, 15:R81
http://breast-cancer-research.com/content/15/5/R81

Breast Cancer RESEARCH

RESEARCH ARTICLE Open Access

A genome-wide association study of chemotherapy-induced alopecia in breast cancer patients

Suyoun Chung[1,6†], Siew-Kee Low[3,6†], Hitoshi Zembutsu[6], Atsushi Takahashi[3], Michiaki Kubo[4], Mitsunori Sasa[5] and Yusuke Nakamura[1,2,6*]

Abstract

Introduction: Chemotherapy-induced alopecia is one of the most common adverse events caused by conventional cytotoxic chemotherapy, yet there has been very little progress in the prevention or treatment of this side effect. Although this is not a life-threatening event, alopecia is very psychologically difficult for many women to manage. In order to improve the quality of life for these women, it is important to elucidate the molecular mechanisms of chemotherapy-induced alopecia and develop ways to effectively prevent and/or treat it. To identify the genetic risk factors associated with chemotherapy-induced alopecia, we conducted a genome-wide association study (GWAS) using DNA samples from breast cancer patients who were treated with chemotherapy.

Methods: We performed a case-control association study of 303 individuals who developed grade 2 alopecia, and compared them with 880 breast cancer patients who did not show hair loss after being treated with conventional chemotherapy. In addition, we separately analyzed a subset of patients who received specific combination therapies by GWASs and applied the weighted genetic risk scoring (wGRS) system to investigate the cumulative effects of the associated SNPs.

Results: We identified an SNP significantly associated with drug-induced grade 2 alopecia (rs3820706 in *CACNB4* (calcium channel voltage-dependent subunit beta 4) on 2q23, $P = 8.13 \times 10^{-9}$, OR = 3.71) and detected several SNPs that showed some suggestive associations by subgroup analyses. We also classified patients into four groups on the basis of wGRS analysis and found that patients who classified in the highest risk group showed 443 times higher risk of antimicrotubule agents-induced alopecia than the lowest risk group.

Conclusions: Our study suggests several associated genes and should shed some light on the molecular mechanism of alopecia in chemotherapy-treated breast cancer patients and hopefully will contribute to development of interventions that will improve the quality of life (QOL) of cancer patients.

Introduction

Breast cancer is the most common malignancy among women worldwide [1]. Although treatment of breast cancer has been significantly improved by the development of molecular-targeted drugs in the past few decades, a subset of patients do not receive benefit from these modalities [2,3]. Such patients and the majority of relapsed patients are treated with conventional cytotoxic chemotherapy that can often cause various adverse events including hair loss.

Hair loss (alopecia) is one of the most common side effects caused by chemotherapy in cancer patients, particularly in women with breast cancer. Although molecular-targeted drugs such as trastuzumab do not cause alopecia, these drugs are given together with other chemotherapeutic agents. Most of the cytotoxic agents cause alopecia, but the severity in individual patients and the incidence by the types of drugs are significantly different: more than 80% of patients treated with antimicrotubule agents, more than 60% of those with

* Correspondence: ynakamura@bsd.uchicago.edu
†Equal contributors
[1]Department of Medicine, The University of Chicago, 5801 South Ellis Avenue, Chicago, IL 60637, USA
[2]Department of Surgery, The University of Chicago, A27 S Maryland Avenue, Chicago, IL 60637, USA
Full list of author information is available at the end of the article

alkylating agents, 60 to 100% of those with topoisomerase inhibitors, and 10 to 50% of those with antimetabolite-based drugs experience severe alopecia [4]. It is also well known that the incidence and the severity are increased when patients are treated with a combination of multiple drugs rather than a single agent [4,5]. Usually, hair loss begins one to two weeks after the start of chemotherapy and a patient's hair can be completely lost in a one- to two-month period. Hair starts to regrow after chemotherapy is completed or discontinued [6,7]. This drug-induced hair loss is not a life-threatening side effect, however, it can strongly influence cosmetic appearance and psychological stresses, and often affects the quality of life (QOL) of the patients [7]. Several studies have demonstrated that the majority of women patients are distressed due to treatment-related alopecia and that 8% of the women avoid chemotherapy because they are unwilling to deal with hair loss [7-10]. Moreover, one study reported that the hair loss was harder to manage than the loss of a breast in some patients [11].

It is known that there are three cycles during hair growth: anagen is the growth phase; catagen is the involuting or regressing phase; and telogen is the resting or quiescent phase [12,13]. It is thought that chemotherapeutic agents target highly proliferative hair matrix cells in the anagen phase, called the anagen effluvium [4,14], but the molecular mechanism is still largely unknown. Scalp cooling with cold air or liquid is the most widely used method since the 1970s to prevent or minimize drug-induced alopecia. However, it is not always effective and it is not easy to standardize the system of scalp cooling [4,15]. Since medications such as minoxidil or AS101, which are widely used for aging-related hair loss, failed to show any protective effect in the case of chemotherapy-induced alopecia [16-19], there is currently no good option to prevent or treat drug-induced alopecia.

In this study, we conducted a genome-wide association study (GWAS) using mono- or combination-chemotherapy-treated breast cancer cases to identify common genetic factors that are associated with drug-induced alopecia. We have identified some loci that are likely to be associated with increased risk of chemotherapy-induced alopecia. These results can provide new insight into the molecular mechanisms of hair loss induced by anticancer drugs and may contribute to development of drugs that can prevent or treat this emotionally devastating side effect.

Methods

Participants

All samples used in this study were obtained from the BioBank Japan located at the Institute of Medical Science at the University of Tokyo. The BioBank Japan project [20], which began in 2003, is a collaborative network of 66 hospitals in Japan [21]. The project achieved a collection of genomic DNA, serum, and clinical information from a total of 330,000 cases (200,000 patients) that had at least 1 of 47 defined diseases. Adverse drug reaction (ADR) information was collected from the patients' medical records by medical coordinators. From the BioBank Japan, we selected 1,367 individuals who had been diagnosed with breast cancer and had received conventional chemotherapy. Of them, 303 patients had experienced grade 2 alopecia (ADR), 184 revealed grade 1 alopecia, and the remaining 880 patients were reported to have had no alopecia (non-ADR). Grade 2 alopecia is defined as complete hair loss, which is the most severe grade in this adverse reaction (National Cancer Institute Common Toxicity Criteria (NCI-CTC) version 3.0). In addition, samples from 23 breast cancer patients with grade 2 alopecia were collected at the Tokushima Breast Care Clinic to further verify the findings of the initial GWAS study; all of the 23 patients were treated with a combination therapy of docetaxel and cyclophosphamide. The detailed clinical information is summarized in Additional file 1. All participants provided written informed consent. This project was approved by the Institutional Review Board of the Institute of Medical Science, the University of Tokyo, and RIKEN Center for Genomic Medicine.

Genotyping and quality control

For GWAS, all DNA samples were genotyped using Illumina Human OmniExpress BeadChip kits (Illumina, San Diego, CA, USA). Sample quality control was performed by identity-by-state clustering across all samples to evaluate cryptic relatedness for each sample and by use of principal component analysis to exclude genetically heterogeneous samples from further analysis. We applied SNP quality control by excluding SNPs with a call rate of <0.99, a P value of the Hardy-Weinberg equilibrium test of $\leq 1.0 \times 10^{-6}$, and non-polymorphic SNPs in the dataset. Quantile-quantile (Q-Q) plots and lambda values, which were used for further evaluation of population substructure, were calculated between observed P value from Fisher's exact test allelic model against expected P value. For genotyping of additional samples, we used the multiplex PCR-based Invader assay (Third Wave Technologies, Madison, WI, USA) as described previously [22].

Statistical analysis

In the GWAS, Fisher's exact test was applied to three genetic models: an allele frequency model, a dominant inheritance model, and a recessive inheritance model. SNPs were rank-ordered according to the lowest P value among the three models. Odds ratio (OR) and confidence intervals (CIs) were calculated for the allelic model using a non-risk allele or a non-risk genotype as a reference. A Manhattan plot was generated by using the minimum

P value among three genetic models. For the combined analysis, the genotype count of the additional samples was added to that of the GWAS. All statistical analyses and plots were carried out using R statistical environment version 2.13.2 [23], and PLINK version 1.07 [24,25]. Haploview software was used for haplotype analysis, to draw the Manhattan plot and linkage disequilibrium (LD) map.

Scoring system using weighted genetic risk score (wGRS)

The scoring analysis was performed by utilizing SNPs with *P* min of $<1.0 \times 10^{-5}$ after exclusion of SNPs that show strong LD (r^2 >0.8) of each GWAS. wGRSs were calculated according to a method reported by De Jager *et al.* [26]. Briefly, we first determined the effect size of each SNP, calculated the cumulative genetic risk scores by multiplying the number of risk alleles for each SNP by its corresponding weight, and subsequently took the sum across the total number of SNPs that were taken into consideration of each GWAS set. We classified the genetic risk score into four different groups, which were created from the mean and standard deviation (SD) as follows: <mean −1 SD for group 1; mean −1 SD to average for group 2; average to mean +1 SD for group 3; >mean +1 SD for group 4. Odds ratio (OR), 95% confidence interval (CI), *P* value, sensitivity, and specificity were calculated using group 1 as reference.

Results

Genome-wide association for chemotherapy-induced alopecia in breast cancer

We performed a GWAS of 303 individuals who developed grade 2 alopecia, and compared them with 880 breast cancer patients who did not show any hair loss after being treated with conventional chemotherapy. The Q-Q plot and lambda (λ) value (λ <1.000) indicated no evidence of population stratification between the cases and controls we analyzed (Additional file 2). After the data was quality controlled, association analysis was carried out for 555,600 autosomal SNPs by Fisher's exact test on the basis of three genetic models: allelic-effect, dominant-inheritance, and recessive-inheritance models. Among the SNPs analyzed in the GWAS, we identified a locus that reached genome-wide significance (rs3820706 near *CACNB4*, minimum $P = 8.13 \times 10^{-9}$, ORrec = 3.71, 95% CI: 2.24 to 6.15) and five additional loci that revealed suggestive association with chemotherapy-induced alopecia with a *P* value of $<10^{-6}$ (Additional file 3 and Table 1). We further validated the top nine SNPs that revealed the smallest *P* value on the three loci in the GWAS result, using 23 additionally obtained alopecia cases. The combined analysis slightly improved the association with the rs3820706 locus (combined minimum $P = 1.85 \times 10^{-9}$, ORrec = 2.38, 95% CI: 1.44 to 3.93) and a nearby SNP rs16830728 (combined minimum $P = 2.60 \times 10^{-8}$, ORrec = 3.61, 95% CI: 2.17 to 5.98; Table 2). As these two SNPs are in strong LD with r^2 of >0.8, we performed haplotype analysis, but the association was not as strong as those of single SNPs (Additional file 4 and Additional file 5).

Association studies for drug subgroups and specific drugs

We also performed subgroup analyses for different types of chemotherapy, namely the CEF (cyclophosphamide + epirubicin +/− 5-FU)-treated and CAF (cyclophosphamide + doxorubicin +/− 5-FU)-treated groups. Detailed sample demographics are described in Additional file 1. In the GWAS of the CEF-treated group, genetic variants in the *ALOX5AP* gene on chromosome 13 were most significantly associated with chemotherapy-induced alopecia (rs3885907, minimum $P = 1.38 \times 10^{-6}$, OR = 2.66, 95% CI: 1.71 to 4.13). The GWAS analysis for the CAF-treated group identified SNP rs594206 located in an intronic region of *BCL9* on chromosome 1 to be most strongly associated (minimum $P = 5.91 \times 10^{-7}$, OR = 36.3, 95% CI: 4.58 to 287; Additional file 3 and Additional file 6). Although the *P* values for these variants did not exceed the genome-wide significance, it is notable that OR for the identified SNP for the CAF analysis is very large. In addition, we analyzed the association with antimicrotubule agents, paclitaxel monotherapy and docetaxel monotherapy because of their high incidence of alopecia, and found that rs1858231 (minimum $P = 1.95 \times 10^{-6}$, OR = 2.71, 95% CI: 1.79 to 4.12), rs11059635 (minimum $P = 2.05 \times 10^{-7}$, OR = 6.63, 95% CI: 2.95 to 14.9) and rs4262906 (minimum $P = 6.62 \times 10^{-7}$, OR = 4.36, 95% CI: 2.41 to 7.89) were most significantly associated, respectively (Additional file 6).

SNP rs3820706 on *CACNB4*, which showed the strongest association with chemotherapy-induced alopecia with the genome-wide significance in the analysis of all-combined samples, showed modest associations in all of the subgroup analyses (Additional file 7). Although the numbers of samples in these subgroup analyses were relatively limited, these data may provide fundamental information that will contribute to a better understanding of chemotherapy-induced alopecia.

Scoring system for prediction of chemotherapy-induced alopecia

We then evaluated the cumulative effects of the candidate loci (SNPs showing $P < 10^{-5}$ in Table 1 and Additional file 6) using a weighted genetic risk scoring (wGRS) method [26]. We first selected eight SNPs from the GWAS of the combination of all samples and calculated wGRS. As shown in Additional file 8, only 17 of 190 patients belonging to group 1 showed severe hair loss (grade 2) while 54 of 82 patients in group 4 revealed it. Cumulative risk scores for the risk of drug-induced alopecia were calculated to be

Table 1 Summary of association results of the genome-wide association study

				ADR[b]			Non-ADR[c]			RAF		*P* value				
CHR	SNP	Gene	Allele 1/2 (risk)	11	12	22	11	12	22	ADR	Non-ADR	Allelic	Dominant	Recessive	OR[a]	95% CI
2	rs3820706	*CACNB4*	A/G (G)	18	169	116	167	421	291	0.66	0.57	8.26E-05	1.07E-01	8.13E-09	3.71	(2.24-6.15)
2	rs6725180	*CACNB4*	A/C (C)	17	152	134	135	429	316	0.69	0.60	7.90E-05	1.11E-02	3.84E-06	3.05	(1.81-5.14)
8	rs16908658	*FAM135B*	G/A (G)	30	93	180	23	286	571	0.25	0.19	1.07E-03	9.68E-02	9.93E-07	4.09	(2.34-7.17)
10	rs7476422	*PCDH15*	T/G (G)	4	47	252	34	245	601	0.91	0.82	1.20E-07	3.77E-07	3.58E-02	2.17	(1.60-2.93)
10	rs857373	*PCDH15*	G/A (A)	5	55	243	43	255	581	0.89	0.81	5.16E-07	3.15E-06	1.11E-02	2.00	(1.51-2.66)
10	rs857392	*PCDH15*	G/A (A)	5	55	243	42	252	584	0.89	0.81	9.08E-07	5.95E-06	1.60E-02	1.97	(1.48-2.62)
10	rs1319836	*PCDH15*	C/T (T)	5	55	243	42	254	583	0.89	0.81	9.10E-07	4.34E-06	1.60E-02	1.98	(1.49-2.63)
10	rs7919725	*PCDH15*	A/G (G)	5	56	242	42	256	580	0.89	0.81	9.94E-07	4.68E-06	1.60E-02	1.97	(1.48-2.60)
10	rs857369	*PCDH15*	T/C (C)	1	32	270	18	178	684	0.94	0.88	2.29E-06	7.25E-06	5.87E-02	2.33	(1.60-3.39)
10	rs9416306	*PCDH15*	G/T (T)	1	32	270	18	178	682	0.94	0.88	2.29E-06	7.13E-06	5.88E-02	2.34	(1.61-3.39)
10	rs1219862	*PCDH15*	C/T (T)	2	31	270	17	182	681	0.94	0.88	2.73E-06	5.08E-06	1.85E-01	2.28	(1.58-3.30)
13	rs7318267	*FARP1*	C/T (T)	11	149	143	108	387	385	0.72	0.66	6.69E-03	3.15E-01	4.09E-06	3.71	(1.97-7.01)
13	rs2282048	*FARP1*	T/C (C)	11	148	144	107	387	386	0.72	0.66	5.72E-03	2.84E-01	6.24E-06	3.68	(1.95-6.93)
17	rs1530357	*LOC100506974*	A/G (A)	57	170	76	114	417	349	0.47	0.37	1.11E-05	4.29E-06	1.39E-02	1.96	(1.45-2.63)
17	rs1530361	*LOC100506974*	A/G (A)	53	165	85	99	408	372	0.45	0.35	8.83E-06	1.12E-05	7.04E-03	1.54	(1.27-1.86)
19	rs11666971	*LASS4*	G/A (G)	46	119	138	56	379	445	0.35	0.28	1.64E-03	1.43E-01	8.13E-06	2.63	(1.74-3.96)

[a]ORs and CIs are calculated according to the associated genetic model; [b]individuals who developed grade 2 alopecia; [c]individuals who did not developed any ADRs after chemotherapy. CHR, chromosome; SNP, single nucleotide polymorphism; ADR, adverse drug reaction; RAF, risk allele frequency; OR, odds ratio; CI, confidence interval.

Table 2 Summary of combined results of the genome-wide association study and additional genotyped data

						ADR[c]				Non-ADR[d]				*P* value				
SNP	CHR	Chromosome position[a]	Gene	Allele 1/2 (risk)		11	12	22	RAF	11	12	22	RAF	Allelic	Dominant	Recessive	*P* min	OR[b] (95% CI)
rs3820706	2	152957411	*CACNB4*	A/G	GWAS	18	169	116	0.66	167	421	291	0.57	8.26E-05	1.07E-01	8.13E-09	8.13E-09	3.71 (2.24-6.15)
				(G)	2nd	1	12	10	0.70	167	421	291	0.57	9.80E-02	3.70E-01	1.00E-01	9.80E-02	1.72 (0.91-3.25)
					Combine	19	181	126	0.66	167	421	291	0.57	3.16E-05	7.65E-02	1.85E-09	1.85E-09	2.38 (1.44-3.93)
rs16830728	2	152981335	*STAM2*	G/T	GWAS	17	163	123	0.68	153	422	304	0.59	1.11E-04	6.16E-02	7.24E-08	7.24E-08	3.54 (2.11-5.96)
				(T)	2nd	1	11	11	0.72	153	422	304	0.59	9.40E-02	1.91E-01	1.55E-01	9.40E-02	1.79 (0.94-3.43)
					Combine	18	174	134	0.68	153	422	304	0.59	3.49E-05	4.30E-02	2.60E-08	2.60E-08	3.61 (2.17-5.98)
rs7476422	10	56204291	*PCDH15*	T/G	GWAS	4	47	252	0.91	34	245	601	0.82	1.20E-07	3.77E-07	3.58E-02	1.20E-07	2.17 (1.60-2.93)
				(G)	2nd	0	7	16	0.85	34	245	601	0.82	8.45E-01	1.00E+00	1.00E+00	8.45E-01	1.21 (0.53-2.72)
					Combine	4	54	268	0.91	34	245	601	0.82	2.63E-07	1.15E-06	2.41E-02	2.63E-07	2.06 (1.54-2.75)

[a]On the basis of NCBI 36 genome assembly; [b]ORs and CIs are calculated according to the associated genetic model; [c]individuals who developed grade 2 alopecia; [d]individuals who did not developed any ADRs after chemotherapy. The same controls were used in the GWAS and second stages analysis. SNP, single nucleotide polymorphism; CHR, chromosome; ADR, adverse drug reaction; RAF, risk allele frequency; *P* min, minimum *P* value; OR, odds ratio; CI, confidence interval.