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

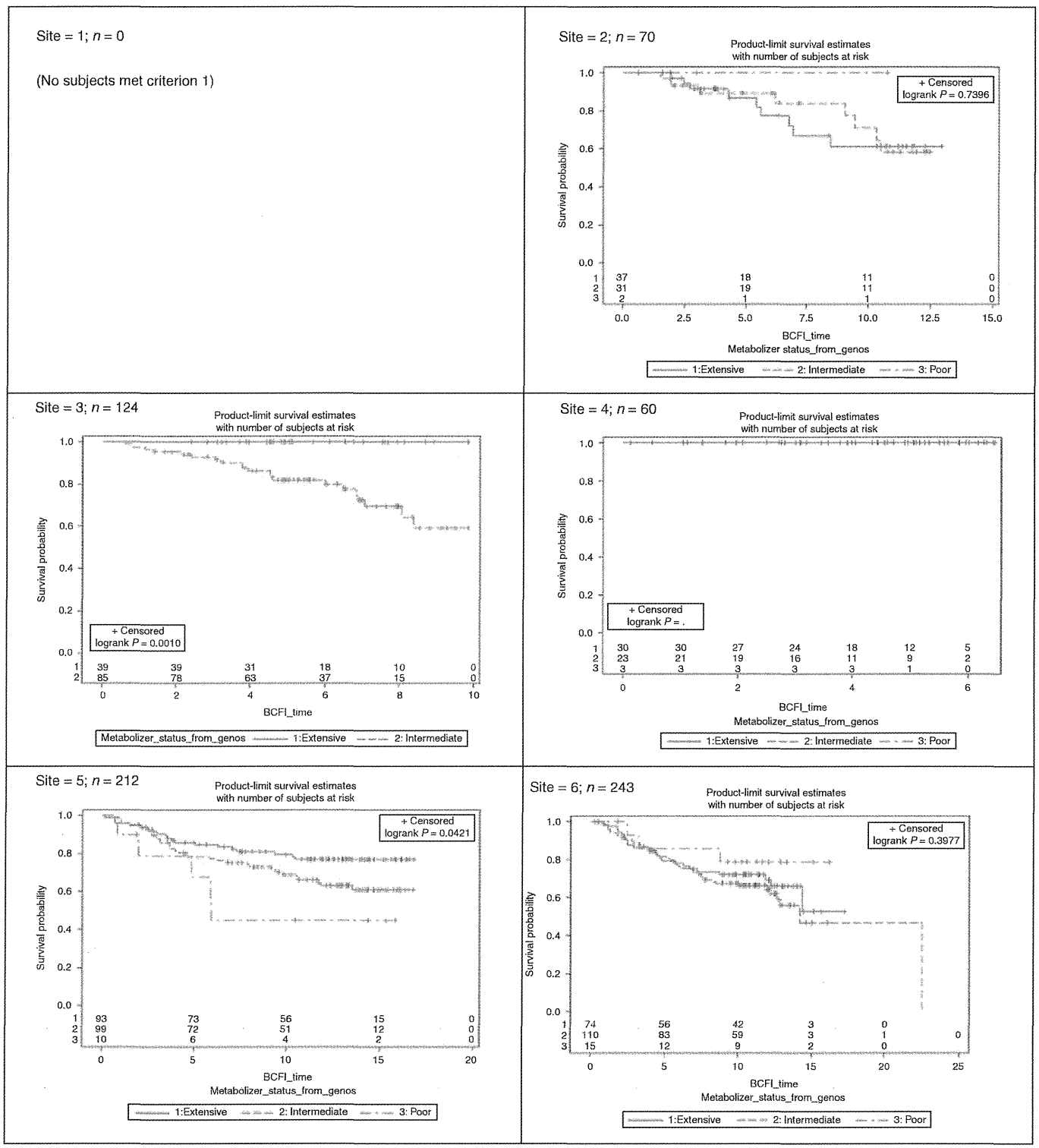

Figure 3 Site-specific effects of CYP2D6 metabolizer status on clinical outcomes for subjects meeting inclusion criterion 1 (outcome = breast cancer–free interval (BCFI)).

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

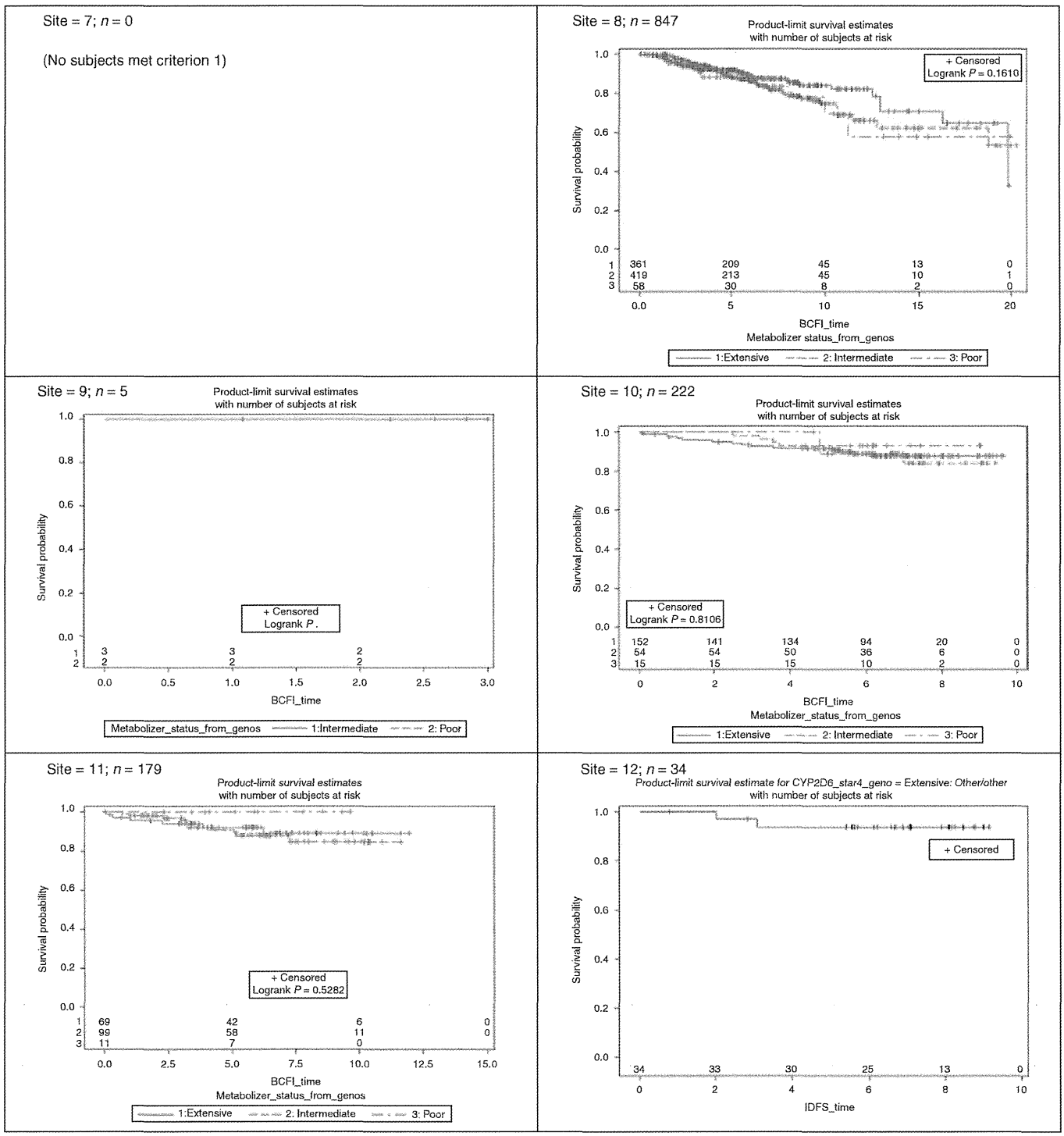

Figure 3 Continued

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).

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 C s are ca cu ated according to the associated genetic mode ; [b]individua s who deve oped grade 2 a opecia; [c]individua s who did not deve oped any ADRs after chemotherapy. CHR, chromosome; SNP, sing e nuc eotide po ymorphism; ADR, adverse drug reaction; RAF, risk a e e frequency; OR, odds ratio; C , confidence interva .

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)
					Comb ne	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)
					Comb ne	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)
					Comb ne	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 NCB 36 genome assemb y; [b]ORs and C s are ca cu ated according to the associated genetic mode ; [c]individua s who deve oped grade 2 a opecia; [d]individua s who did not deve oped any ADRs after chemotherapy. The same contro s were used in the GWAS and second stages ana ysis. SNP, sing e nuc eotide po ymorphism; CHR, chromosome; ADR, adverse drug reaction; RAF, risk a e e frequency; *P* min, minimum *P* va ue; OR, odds ratio; C , confidence interva .

4.44 in group 3 and 19.6 in group 4 ($P = 3.44 \times 10^{-9}$, 95% CI: 2.62 to 7.53; $P = 1.44 \times 10^{-21}$, 95% CI: 9.99 to 38.6, respectively), compared with patients in group 1.

Similarly, in the subgroup analysis, an individual belonging to group 4 with the highest risk score in each of the CEF, CAF, antimicrotubules, paclitaxel, and docetaxel analyses was estimated to have 86.2 times, 891 times, 858 times, 1,680 times, and 441 times higher risk for the drug-related alopecia than those in group 1, respectively (Additional file 8). Due to the clinical importance of antimicrotubule agents (paclitaxel and docetaxel), which cause chemotherapy-induced alopecia at nearly 80% frequency, we further investigated the wGRS scoring method using cases with grade 1 alopecia. Interestingly, the association levels and odds ratios of patients with grade 1 alopecia induced by the antimicrotubule agents were intermediate, compared with those of grade 2 alopecia (Table 3). Not only antimicrotubule agents, but other subgroups (all, CEA or CEF) also showed similar results, and the association level of grade 1 was intermediate compared with grade 2. These results further support a possible association of these variants in alopecia development (Additional file 9). As shown in Figure 1, the proportion of grade 2 alopecia increased according to the increase of the wGRS score; for example, in the case of docetaxel, only one (3.4%) of the 29 patients in group 1 revealed grade 2 alopecia, while 52 (83%) of 63 patients belonging to groups 3 and 4 developed grade 2 alopecia. These results indicate that our scoring system may be applied to predict severe chemotherapy-induced alopecia and might provide useful information for better understanding of the hair-loss mechanism, even though further verification using an additional independent set(s) of samples is warranted.

Finally, we simulated the sample number that is required to verify our scoring system. In BioBank Japan, a total of 279 patients received antimicrotubule agents (paclitaxel and/or docetaxel). Among them, 119 (43%) patients developed grade 2 alopecia, 55 (20%) developed grade 1 alopecia and 105 (37%) did not show any adverse events. Among 156 patients who received paclitaxel monotherapy, 57 (37%) developed grade 2 alopecia, 36 (23%) developed grade 1 alopecia and 63 (40%) did not develop any adverse reactions. When we assume that 100 patients who receive antimicrotubule agents (or paclitaxel monotherapy) are registered, the incidences of alopecia are estimated as shown in Table 4. If we categorize the patients by wGRS according to the data in Table 3, 100 additional patients should provide the sufficient statistical power to verify our results with P value of <0.01. Even if two individuals in each of groups 1 and 4 are not correctly predicted, the calculated P value is still 0.001 by Fisher's exact test.

Table 3 wGRS results of antimicrotubule agents, docetaxel, and paclitaxel-induced alopecia

Cat	Score	G2[a]	G1[b]	G0[c]	% G2	% G1	% G0	G2 vs. G0			G1 vs. G0		
								OR[d*]	95%CI	*P* value	OR[d*]	95%CI	*P* value
Antimicrotubule (6 SNPs)													
1	<5.56	2	7	34	0.05	0.16	0.79		Ref			Ref	
2	5.56 7.60	25	20	50	0.26	0.21	0.53	8.50	1.89 38.3	1.66E 03	1.94	0.74 1.42	2.52E 01
3	7.60 9.63	65	17	19	0.64	0.17	0.19	58.2	12.8 265	4.93E 14	4.35	1.53 12.4	6.42E 03
4	>9.63	26	6	1	0.79	0.18	0.03	442	38.0 5140	2.71E 14	29.1	3.02 282	8.39E 04
	Total	**118**	**50**	**104**									
Docetaxel (4 SNPs)													
1	<2.26	1	5	23	0.03	0.17	0.79		Ref			Ref	
2	2.26 4.70	9	6	14	0.31	0.21	0.48	14.8	1.69 130	4.39E 03	1.97	0.51 7.68	4.88E 01
3	4.70 7.15	33	5	4	0.79	0.12	0.10	190	19.9 1810	1.01E 11	5.75	1.12 29.4	4.08E 02
4	>7.15	19	2	0	0.90	0.10	0.00	611	23.5 15900	2.50E 11	21.4	0.89 511	4.83E 02
	Total	**62**	**18**	**41**									
Paclitaxel (7 SNPs)													
1	<3.24	2	4	28	0.06	0.12	0.82		Ref			Ref	
2	3.24 7.48	4	14	22	0.10	0.35	0.55	2.55	0.43 15.2	4.01E 01	4.46	1.28 1.92	2.60E 02
3	7.48 11.7	35	12	11	0.60	0.21	0.19	44.6	9.12 218	9.55E 10	7.64	2.02 28.9	2.30E 03
4	>11.7	16	5	0	0.76	0.24	0.00	376	17.0 8320	1.54E 10	69.7	3.26 1490	2.89E 04
	Total	**57**	**35**	**61**									

[a]Individuals who developed grade 2 alopecia; [b]individuals who developed grade 1 alopecia; [c]individuals who did not developed any ADRs after chemotherapy; [d]ORs and CIs are calculated using category (group) 1 as reference. *OR calculated after Haldane's correction: adding 0.5 to all the cells of a contingency table if any of the cell expectations would cause a division by zero error. Cat, category; OR, odds ratio; CI, confidence interval; Ref, reference.

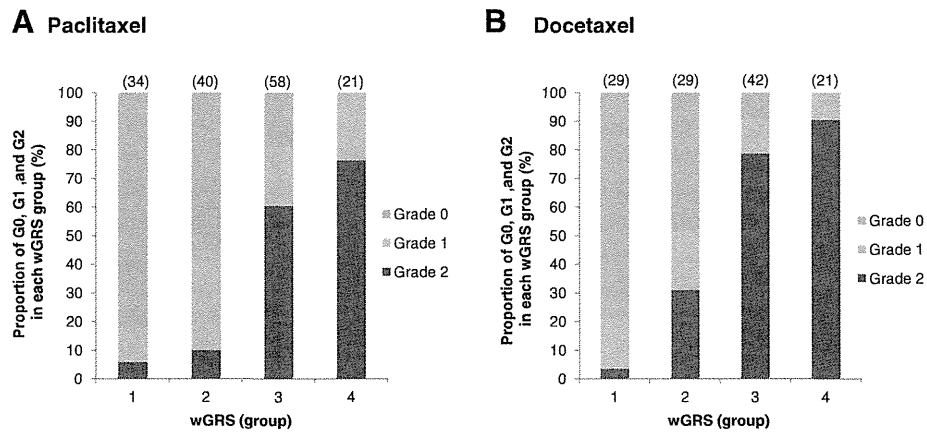

Figure 1 The proportions of patients by alopecia grade in each weighted genomic risk score. The proportions of patients who developed no adverse reaction (G0), grade 1 alopecia, or grade 2 alopecia in each of the weighted genomic risk score (wGRS) groups. The number in parentheses indicates the number of samples in each group. **(A)** paclitaxel monotherapy, **(B)** docetaxel monotherapy.

Discussion

Recent pharmacogenomics studies focus on prediction of drug response as well as the risk assessment of toxic events due to administration of drugs. Whole-genome association studies have been proven to be a powerful strategy to identify genetic factor(s) associated with various adverse reactions caused by certain drugs. In this study, we conducted the first GWAS for chemotherapy-induced alopecia in Japanese breast cancer patients, and identified one locus including two SNPs, rs3820706 on chromosome 2q23 and its nearby SNP rs16830728, which showed a strong association with genome-wide significance, and found several SNPs showing suggestive associations.

SNP rs3820706 is located near a gene encoding calcium channel voltage-dependent subunit beta 4 (*CACNB4*), a member of a beta subunit family of the voltage-dependent calcium channel (VDCC) complex. Calcium (Ca^{2+}) functions as a second messenger in many cellular signal transduction pathways such as cell proliferation and apoptosis. When VDCC is activated it depolarizes membrane potentials, it allows Ca^{2+} to enter into cells [27]. We are not aware of any previous reports indicating that there is a relationship between the Ca^{2+} channel and alopecia. However, a potassium channel opener, minoxidil, was approved for the treatment of alopecia by the US FDA in 1988 [28] and has proven to be effective in a subset of alopecia patients. Although the mode of action of minoxidil is still not well known, the clinical outcome implies the involvement of ion channels for K^+ and probably Ca^{2+} in the pathogenesis of alopecia. Intriguingly, the second most significantly associated locus that we found in our study is a region containing the *PCDH15* gene on chromosome

Table 4 Estimation of required sample number for verification

	Cat	G2	G0	OR	95% CI	*P* value*
Antimicrotubule (paclitaxel and docetaxel) (N = 100)						
	1	1	12		Ref	
	2	9	18	6.00	0.67 53.7	1.24E 01
	3	24	7	41.1	4.53 374	2.48E 05
	4	9	1	108	5.92 1970	1.15E 04
	Total	43	38			
Paclitaxel (N = 100)						
	1	1	18		Ref	
	2	3	15	3.60	0.34 38.3	3.40E 01
	3	23	7	59.1	6.66 525	8.02E 07
	4	10	0	259	9.66 6950	5.49E 07
	Total	37	40			

**P* values are calculated by Fisher's exact test. Cat, category; OR, odds ratio; CI, confidence interval; Ref, reference.

10. *PCDH15* encodes a protocadherin-related protein, which is involved in calcium-dependent cell-cell adhesion. Additionally, among the 70 loci in the top 100 SNPs found in our GWAS study, five loci are implicated to be ion channels or proteins related to ion channels (data not shown). Ion channels have shown to have important roles not only in cell maintenance but also in stem/progenitor cells [29]. Because cytotoxic agents damage the proliferating progenitor cells in the hair matrix [13], we suspect that several ion channels might be involved in chemotherapy-induced alopecia and be promising targets for development of novel treatments.

However, since rs3820706 is strongly linked to rs16830728, which is located within a gene encoding a signal transducing adaptor molecule 2 (*STAM2*), we cannot exclude the possibility that *STAM2* is a candidate gene for chemotherapy-induced alopecia. STAM2 is a member of the STAM family, which is an adaptor protein involved in the downstream signaling of cytokine receptors that contain an SH3 domain and the immunoreceptor tyrosine-based activation motif (ITAM). STAM2 is involved in the signaling through GM-SCF and IL-2 stimulation, and has a crucial role in T cell development [30,31]. As most studies of STAM2 focused on immune cells, its functions in other cell types like hair follicle cells are not fully understood.

In addition, we performed subgroup analyses in which we identified multiple loci that might be associated with drug-induced alopecia. rs3885907, which was most significantly associated in CEF-treated patients, was located in an intron of *ALOX5AP*. *ALOX5AP*, arachidonate 5-lipoxygenase-activating protein, is related to the inflammatory responses and possibly to vascular diseases [32,33]. Detailed biological mechanisms in hair growth cycle are not well characterized, but one paper reported involvement of the *ALOX5AP* upregulation in scarring alopecia [34]. According to GWAS, for alopecia areata [35] that identified genes related in both innate and adaptive immunity, inflammatory or immune responses seem to be important in alopecia development. The mechanisms of hair loss in alopecia areata and in drug-induced alopecia may not be same, but our result suggests a possible relationship of the immune response with chemotherapy-induced alopecia.

A SNP in the *BCL9* gene was most significantly associated with hair loss in the CAF-treated group with very high OR of 36.3. The *BCL9* gene encodes B-cell lymphoma 9 which was reported to interact with β-catenin. The β-catenin signaling pathway is involved in hair follicle morphogenesis during embryogenesis and, interestingly, hair is completely lost when β-catenin is depleted even after hair follicles have been formed [36,37]. Similarly, *CDH7*, one of the cadherin family members, showed an association with severe hair loss in the CAF-treated group with high OR of 32.5. This cadherin has been reported to be expressed in hair follicles and regulate hair growth [38,39]. These results, in combination with our GWAS results, imply possible roles for BCL9 and CDH7 in chemotherapy-induced alopecia. If so, these two molecules as well as CACNB4 and other ion channel proteins could be promising targets for the development of new treatments. However, further validation is still needed.

Our approach of using retrospective BioBank samples is not ideal for addressing this type of clinical problem and certainly a prospective analysis with well-defined clinical information would reduce the possibility of false-positive and false-negative results. However, considering the rapid progress of drug development or new combination therapies in recent years, it may not be wise to spend lots of effort, time and budget to do a prospective study, because the investigated regimen may not be used years later when the research results come out. One of the ways to effectively use the data and samples from the retrospective study is shown by the application of our wGRS system. The wGRS system indicated cumulative effect of multiple genetic variants for alopecia prediction. For example, the patients in group 4 who received paclitaxel showed 376 times increased risk of alopecia, compared with those belonging to group 1. Similarly, the patients in group 4 who received docetaxel showed 611 times higher risk of alopecia than those belonging to group 1. We understand the disadvantages and pitfalls of the retrospective design for the pharmacogenomics study such as the higher risk of false results. However, considering the very high OR obtained by the wGRS system, the advantage of this approach is that we are able to verify the results by using a relatively small number of additional prospective samples. We simulated the sample size needed to verify our results, as shown in Table 4, and suggest that the statistical power should be sufficient to validate with this small number of samples. We recognize that the clinical utility for this wGRS may not be as high as in other studies looking at life-threatening adverse events. However, identification of genetic factors associated with drug-induced hair loss should be the first step to understand the molecular mechanism and to contribute to the development of new drugs to prevent or treat alopecia.

For many years, breast cancer patients have had to accept the psychologically stressful side effect of alopecia caused by cytotoxic chemotherapies. It is known that a subset of patients will refuse to have chemotherapy because they do not want to lose their hair and therefore may lose the opportunity to receive the benefit of the chemotherapy and a chance to be cured of their disease. The QOL of these patients is extremely important and we believe it is urgent that we work to develop new treatment or prevention strategies to manage chemotherapy-induced alopecia. Although further validation of our findings is

required, our study identified some significant molecular alterations in genes such as ion channel-related genes and genes in the β-catenin signaling pathway. We welcome other groups to examine and validate our results and hope these findings will contribute to the development of interventions that will improve the quality of life (QOL) of breast cancer patients.

Conclusions

In summary, we identified strongly associated genetic variants near gene *CACNB4* and several suggestively associated SNPs with chemotherapy-induced alopecia in breast cancer patients. These results provide new information of the pathogenesis of chemotherapy-induced alopecia.

Additional files

Additional file 1: Table S1. Patients' characteristics.

Additional file 2: Quantile quantile plot of the genome wide association study.

Additional file 3: Manhattan plot of the genome wide association study for chemotherapy induced alopecia in breast cancer.

Additional file 4: Haplotype analysis.

Additional file 5: Table S2. Haplotype analysis of two SNPs.

Additional file 6: Table S3. Summary of genome wide association study for chemotherapy induced alopecia with each drug subgroup (P <10 6).

Additional file 7: Table S4. Association of rs3820706 in subgroups.

Additional file 8: Table S5. Weighted genomic risk score of each genome wide association study for chemotherapy induced alopecia.

Additional file 9: Table S6. Weighted genomic risk score results of all, CAF and CEF induced alopecia.

Abbreviations

ADR: Adverse drug reaction; CAF: Cyclophosphamide + doxorubicin +/ 5 FU; CEF: Cyclophosphamide + epirubicin +/ 5 ; CI: Confidence interval; GWAS: Genome wide association study; LD: Linkage disequilibrium; OR: Odds ratio; QOL: Quality of life; QQ plot: Quantile quantile plot; SD: Standard deviation; SNP: Single nucleotide polymorphism; wGRS: Weighted genomic risk score.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YN planned and supervised the study and obtained funding. SC and SKL designed the experiments and performed the GWAS and combined analysis. SKL performed the wGRS and statistical analysis. HZ and MS collected additional samples and medical information. MK genotyped all BioBank Japan samples. AT performed sample quality control. YN, SC and SKL wrote the manuscript. All authors revised and approved the manuscript for publication.

Acknowledgements

We express our heartfelt gratitude to all the participants. We thank Miss Kumi Matsuda for her outstanding technical assistance. We also thank all other members and staff for their contribution to the sample collection. This study was supported by a leading project of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author details

[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. [3]Laboratory for Statistical Analysis, Center for Genomic Medicine, 1 7 22 Suehiro cho, Tsurumi ku, Yokohama, Kanagawa 230 0045, Japan. [4]Laboratory for Genotyping Development, Center of Genomic Medicine, 1 7 22 Suehiro cho, Tsurumi ku, Yokohama, Kanagawa 230 0045, Japan. [5]Tokushima Breast Care Clinic, 4 7 7, Nakashimada cho, Tokushima 770 0052, Japan. [6]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4 6 1 Shirokanedai, Minato ku, Tokyo 108 8639, Japan.

Received: 1 May 2013 Accepted: 18 July 2013
Published: 11 September 2013

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ: **Cancer statistics, 2008.** *CA Cancer J Clin* 2008, **58**:71 96.
2. Sotiriou C, Pusztai L: **Gene expression signatures in breast cancer.** *N Engl J Med* 2009, **360**:790 800.
3. Bosch A, Eroles P, Zaragoza R, Vina JR, Lluch A: **Triple negative breast cancer: molecular features, pathogenesis, treatment and current lines of research.** *Cancer Treat Rev* 2010, **36**:206 215.
4. Trueb RM: **Chemotherapy induced alopecia.** *Semin Cutan Med Surg* 2009, **28**:11 14.
5. Carrick S, Parker S, Thornton CE, Ghersi D, Simes J, Wilcken N: **Single agent versus combination chemotherapy for metastatic breast cancer.** *Cochrane Database Syst Rev* 2009, **2**:CD003372.
6. Batchelor D: **Hair and cancer chemotherapy: consequences and nursing care a literature study.** *Eur J Cancer Care (Engl)* 2001, **10**:147 163.
7. McGarvey EL, Baum LD, Pinkerton RC, Rogers LM: **Psychological sequelae and alopecia among women with cancer.** *Cancer Pract* 2001, **9**:283 289.
8. Tierney AJ, Taylor J, Closs SJ: **Knowledge, expectations and experiences of patients receiving chemotherapy for breast cancer.** *Scand J Caring Sci* 1992, **6**:75 80.
9. Munstedt K, Manthey N, Sachsse S, Vahrson H: **Changes in self concept and body image during alopecia induced cancer chemotherapy.** *Support Care Cancer* 1997, **5**:139 143.
10. de Boer Dennert M, de Wit R, Schmitz PI, Djontono J, v Beurden V, Stoter G, Verweij J: **Patient perceptions of the side effects of chemotherapy: the influence of 5HT3 antagonists.** *Br J Cancer* 1997, **76**:1055 1061.
11. Freedman TG: **Social and cultural dimensions of hair loss in women treated for breast cancer.** *Cancer Nurs* 1994, **17**:334 341.
12. Stenn KS, Paus R: **Controls of hair follicle cycling.** *Physiol Rev* 2001, **81**:449 494.
13. Paus R, Haslam IS, Sharov AA, Botchkarev VA: **Pathobiology of chemotherapy induced hair loss.** *Lancet Oncol* 2013, **14**:e50 e59.
14. Chon SY, Champion RW, Geddes ER, Rashid RM: **Chemotherapy induced alopecia.** *J Am Acad Dermatol* 2012, **67**:e37 e47.
15. Grevelman EG, Breed WP: **Prevention of chemotherapy induced hair loss by scalp cooling.** *Ann Oncol* 2005, **16**:352 358.
16. Duvic M, Lemak NA, Valero V, Hymes SR, Farmer KL, Hortobagyi GN, Trancik RJ, Bandstra BA, Compton LD: **A randomized trial of minoxidil in chemotherapy induced alopecia.** *J Am Acad Dermatol* 1996, **35**:74 78.
17. Granai CO, Frederickson H, Gajewski W, Goodman A, Goldstein A, Baden H: **The use of minoxidil to attempt to prevent alopecia during chemotherapy for gynecologic malignancies.** *Eur J Gynaecol Oncol* 1991, **12**:129 132.
18. Rodriguez R, Machiavelli M, Leone B, Romero A, Cuevas MA, Langhi M, Romero Acuna L, Romero Acuna J, Amato S, Barbieri M, *et al*: **Minoxidil (Mx) as a prophylaxis of doxorubicin induced alopecia.** *Ann Oncol* 1994, **5**:769 770.
19. Wang J, Lu Z, Au JL: **Protection against chemotherapy induced alopecia.** *Pharm Res* 2006, **23**:2505 2514.
20. http://www.biobankjp.org.
21. Nakamura Y: **The BioBank Japan Project.** *Clin Adv Hematol Oncol* 2007, **5**:696 697.
22. Ohnishi Y, Tanaka T, Ozaki K, Yamada R, Suzuki H, Nakamura Y: **A high throughput SNP typing system for genome wide association studies.** *J Hum Genet* 2001, **46**:471 477.
23. http://www.cran.r project.org/.
24. http://pngu.mgh.harvard.edu/purcell/plink/.

25. Purcell S, Neale B, Todd Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al*: **PLINK: a tool set for whole genome association and population based linkage analyses.** *Am J Hum Genet* 2007, **81**:559 575.
26. De Jager PL, Chibnik LB, Cui J, Reischl J, Lehr S, Simon KC, Aubin C, Bauer D, Heubach JF, Sandbrink R, *et al*: **Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score.** *Lancet Neurol* 2009, **8**:1111 1119.
27. Clapham DE: **Calcium signaling.** *Cell* 2007, **131**:1047 1058.
28. Price VH: **Treatment of hair loss.** *N Engl J Med* 1999, **341**:964 973.
29. Li GR, Deng XL: **Functional ion channels in stem cells.** *World J Stem Cells* 2011, **3**:19 24.
30. Endo K, Takeshita T, Kasai H, Sasaki Y, Tanaka N, Asao H, Kikuchi K, Yamada M, Chenb M, O'Shea JJ, *et al*: **STAM2, a new member of the STAM family, binding to the Janus kinases.** *FEBS Lett* 2000, **477**:55 61.
31. Yamada M, Ishii N, Asao H, Murata K, Kanazawa C, Sasaki H, Sugamura K: **Signal transducing adaptor molecules STAM1 and STAM2 are required for T cell development and survival.** *Mol Cell Biol* 2002, **22**:8648 8658.
32. Liu J, Sun K, Bai Y, Zhang W, Wang X, Wang Y, Wang H, Chen J, Song X, Xin Y, *et al*: **Association of three gene interaction among MTHFR, ALOX5AP and NOTCH3 with thrombotic stroke: a multicenter case control study.** *Hum Genet* 2009, **125**:649 656.
33. Manev H, Manev R: **5 Lipoxygenase (ALOX5) and FLAP (ALOX5AP) gene polymorphisms as factors in vascular pathology and Alzheimer's disease.** *Med Hypotheses* 2006, **66**:501 503.
34. Karnik P, Tekeste Z, McCormick TS, Gilliam AC, Price VH, Cooper KD, Mirmirani P: **Hair follicle stem cell specific PPARgamma deletion causes scarring alopecia.** *J Invest Dermatol* 2009, **129**:1243 1257.
35. Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, Kim H, Singh P, Lee A, Chen WV, *et al*: **Genome wide association study in alopecia areata implicates both innate and adaptive immunity.** *Nature* 2010, **466**:113 117.
36. Huelsken J, Vogel R, Erdmann B, Cotsarelis G, Birchmeier W: **Beta Catenin controls hair follicle morphogenesis and stem cell differentiation in the skin.** *Cell* 2001, **105**:533 545.
37. Zhang Y, Andl T, Yang SH, Teta M, Liu F, Seykora JT, Tobias JW, Piccolo S, Schmidt Ullrich R, Nagy A, *et al*: **Activation of beta catenin signaling programs embryonic epidermis to hair follicle fate.** *Development* 2008, **135**:2161 2172.
38. Young P, Boussadia O, Halfter H, Grose R, Berger P, Leone DP, Robenek H, Charnay P, Kemler R, Suter U: **E cadherin controls adherens junctions in the epidermis and the renewal of hair follicles.** *EMBO J* 2003, **22**:5723 5733.
39. Samuelov L, Sprecher E, Tsuruta D, Biro T, Kloepper JE, Paus R: **P cadherin regulates human hair growth and cycling via canonical Wnt signaling and transforming growth factor beta2.** *J Invest Dermatol* 2012, **132**:2332 2341.

doi:10.1186/bcr3475
Cite this article as: Chung *et al.*: A genome wide association study of chemotherapy induced alopecia in breast cancer patients. *Breast Cancer Research* 2013 15:R81.

Journal of Human Genetics (2013) 58, 327 333

www.nature.com/jhg

REVIEW

Important and critical scientific aspects in pharmacogenomics analysis: lessons from controversial results of tamoxifen and *CYP2D6* studies

Kazuma Kiyotani[1,2], Taisei Mushiroda[2], Hitoshi Zembutsu[3] and Yusuke Nakamura[3,4]

Tamoxifen contributes to decreased recurrence and mortality of patients with hormone receptor-positive breast cancer. As this drug is metabolized by phase I and phase II enzymes, the interindividual variations of their enzymatic activity are thought to be associated with individual responses to tamoxifen. Among these enzymes, CYP2D6 is considered to be a rate-limiting enzyme in the generation of endoxifen, a principal active metabolite of tamoxifen, and the genetic polymorphisms of *CYP2D6* have been extensively investigated in association with the plasma endoxifen concentrations and clinical outcome of tamoxifen therapy. In addition to *CYP2D6*, other genetic factors including polymorphisms in various drug-metabolizing enzymes and drug transporters have been implicated to their relations to clinical outcome of tamoxifen therapy, but their effects would be small. Although the results of association studies are controversial, accumulation of the evidence has revealed us the important and critical issues in the tamoxifen pharmacogenomics study, namely the quality of genotyping, the coverage of genetic variations, the criteria for sample collection and the source of DNAs, which are considered to be common problematic issues in pharmacogenomics studies. This review points out common critical issues in pharmacogenomics studies through the lessons we have learned from tamoxifen pharmacogenomics, as well as summarizes the results of pharmacogenomics studies for tamoxifen treatment.
Journal of Human Genetics (2013) **58**, 327–333; doi:10.1038/jhg.2013.39; published online 9 May 2013

Keywords: ABCC2; C10orf11; endoxifen; estrogen receptor; genome-wide association study; P450 2D6; pharmacogenomics study design; single-nucleotide polymorphism

INTRODUCTION

Tamoxifen, a selective estrogen receptor (ER) modulator, has been widely used for the treatment and prevention of recurrence for patients with hormone receptor (ER or progesterone receptor) positive breast cancers. As >70% of breast cancers are hormone receptor positive, thousands of breast cancer patients worldwide initiate to take endocrine treatment including tamoxifen each year. In pre and postmenopausal patients with primary breast cancer, 5 years of adjuvant tamoxifen significantly reduced recurrence rate as well as cancer specific mortality for 15 years after their primary diagnosis.[1] However, approximately one third of patients treated with adjuvant tamoxifen experience a recurrent disease,[1,2] implicating possible individual differences in responsiveness to tamoxifen.

Tamoxifen is metabolized to more active metabolites or inactive forms by phase I and phrase II enzymes, including cytochrome P450s (CYPs), sulfotransferases (SULTs) and UDP glucuronosyltransferases (UGTs). The polymorphisms in these drug metabolizing enzymes are considered to affect individual differences in plasma concentrations of active tamoxifen metabolites and clinical outcome in breast cancer patients treated with tamoxifen. Among these enzymes, CYP2D6 has been most extensively investigated owing to its significant role in production of active metabolites, endoxifen and 4 hydroxytamoxifen.

This review summarizes current reports on the relationships of genetic polymorphisms and other biomarkers to individual differences in clinical outcome of breast cancer patients with tamoxifen treatment. In addition, we investigate reasons or causes of discordant results for the association between *CYP2D6* genetic variations and clinical outcome, and would like to highlight various problematic issues in pharmacogenomics studies.

TAMOXIFEN METABOLISM

Tamoxifen is extensively metabolized by phase I and phase II enzymes in the human liver (Figure 1).[3,4] Tamoxifen itself has low affinity to the ER as only 1.8% of the affinity of 17β estradiol.[3]

[1]Division of Genome Medicine, Institute for Genome Research, The University of Tokushima, Tokushima, Japan; [2]Research Group for Pharmacogenomics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; [3]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan and [4]Department of Medicine and Center for Personalized Therapeutics, The University of Chicago, Chicago, IL, USA
Correspondence: Dr Y Nakamura, Department of Medicine and Surgery, The University of Chicago, 900E 57th Street, KCBD 6130, Chicago 60637, IL, USA.
E mail: ynakamura@bsd.uchicago.edu
Received 19 March 2013; revised 11 April 2013; accepted 12 April 2013; published online 9 May 2013

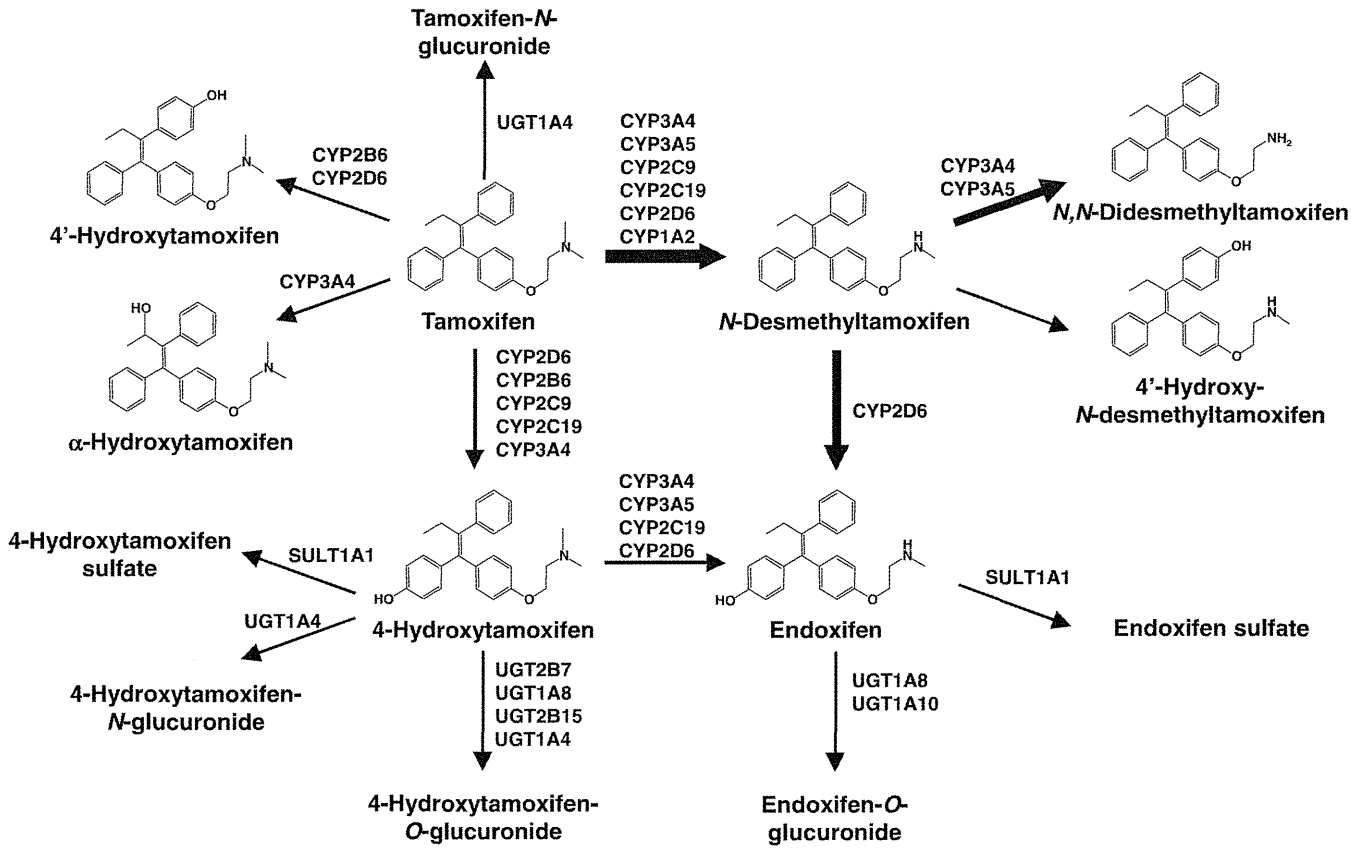

Figure 1 Metabolic pathways of tamoxifen in human. Major metabolic pathways are highlighted with bold arrows.

The major metabolite *N* desmethyltamoxifen is formed by *N* demethylation, which is catalyzed mainly by CYP3A4 and CYP3A5, with small contribution by CYP2D6, CYP1A2, CYP2C9 and CYP2C19.[5–9] *N* desmethyltamoxifen shows weak affinity to the ER similar to tamoxifen.[3,4] However, 4 hydroxytamoxifen, which is formed by 4 hydroxylation of tamoxifen, has 100 fold higher affinity to the ER and 30 to 100 fold greater potency in suppressing estrogen dependent breast cancer cell proliferation than tamoxifen.[3,10–12] This conversion is catalyzed by CYP2D6, CYP2B6, CYP2C9, CYP2C19 and CYP3A4.[5,13–15] Endoxifen (4 hydroxy *N* desmethyltamoxifen) has a potency equivalent to 4 hydroxytamoxifen,[10,16,17] and its plasma concentration level exceed that of 4 hydroxytamoxifen by several folds, suggesting endoxifen to be a principal active metabolite.[9–11] Endoxifen formation from *N* desmethyltamoxifen is predominantly catalyzed by CYP2D6.[18] Several additional metabolites, such as *N*,*N* didesmethyltamoxifen, 4′ hydroxy *N* desmethyltamoxifen and α hydroxytamoxifen were reported, but no other highly active metabolite has been described so far.[4]

Tamoxifen and these metabolites are further metabolized by phase II enzymes, such as SULTs and UGTs. SULT1A1 is considered to be the primary SULT responsible for the sulfation of 4 hydro xytamoxifen and endoxifen.[19,20] UGT1A8, UGT1A10, UGT2B7, UGT2B15 and UGT1A4 are involved in the *O* glucuronidation of 4 hydroxytamoxifen and endoxifen.[21–23] Tamoxifen and 4 hydroxytamoxifen are glucuronidated by UGT1A4 to the corresponding N^+ glucuronides.[24,25] The genetic variations of these drug metabolizing enzymes are possible to affect tamoxifen metabolism.

GENETIC POLYMORPHISMS OF *CYP2D6*

CYP2D6 is one of the most important CYP isoforms owing to its central role in the metabolism of a number of clinically important drugs.[26] The *CYP2D6* gene is located on chromosome 22q13.1, containing two neighboring pseudogenes, *CYP2D7* and *CYP2D8*. This locus is extremely polymorphic with over 80 allelic variants, a subset of which should affect the gene product and result in wide interindividual and ethnic differences in CYP2D6 activity.[27] Commonly, four CYP2D6 phenotypes are defined on the basis of their *in vivo* metabolic capacities: poor metabolizer (PM), intermediate metabolizer (IM), extensive metabolizer (EM) and ultra rapid metabolizer (UM).[28,29] It has been reported that the PM phenotype, which is caused by carrying two null alleles, is present in 5 10% of Caucasians.[30] The *CYP2D6*3*, *CYP2D6*4*, *CYP2D6*5* and *CYP2D6*6* are major null alleles that are related to the PM phenotype and account for nearly 95% of the PMs in Caucasians.[31] Among them, *CYP2D6*4* shows the highest frequency as 17.5 23.0%.[27] *CYP2D6*5*, which is found at a frequency of ~5%, lacks an entire *CYP2D6* gene. In contrast, <1% of Asians show the PM phenotype,[32] and most Asians are categorized as IMs because of the high frequency of a *CYP2D6*10* allele.[33,34] The *CYP2D6*14*, *CYP2D6*18*, *CYP2D6*21*, *CYP2D6*36* and *CYP2D6*44* were null alleles found in Asian populations, although their frequencies are very low.[35–38] The frequencies of UMs, who carry a duplicated/multiplied

wild type *CYP2D6* gene(s), are 10 15% in Caucasian, whereas UMs are uncommon in Asians. As described here, because the *CYP2D6* gene locus is complex, genotyping of *CYP2D6* variants, especially *CYP2D6*5*, is technically not so easy. Although the accuracy of genotyping partly depends on the quality of DNAs and the platforms of genotyping, wrong genotyping results sometimes cause incorrect interpretation of the research outcome, and result in both false positive conclusions and false negative conclusions.

CYP2D6 GENOTYPE AND CLINICAL OUTCOME OF TAMOXIFEN THERAPY

In recent years, we have seen an explosion of interest in the clinical relevance of *CYP2D6* genotype on outcome of breast cancer patients who are treated with tamoxifen. Prospective cohort studies of adjuvant tamoxifen treatment have revealed a wide interindividual variation in the steady state plasma concentrations of active metabo lites, endoxifen and 4 hydroxytamoxifen during tamoxifen treatment in patients carrying *CYP2D6* genetic variants.[8,9,11] The patients homozygous for null alleles (categorized as PM) showed nearly one fourth of endoxifen concentration in plasma, compared with those carrying two normal alleles (categorized as EM).[8,9] The patients carrying two alleles that encode a low function enzyme, including *CYP2D6*10* and *CYP2D6*41* (categorized as inter mediate metabolizer), had nearly 50% of plasma endoxifen concentration compared with the controls.[4,39–41] These patients with low endoxifen concentration were suspected to have a poorer clinical outcome.

As shown in Table 1, a number of studies have reported the association between the *CYP2D6* genotype and clinical outcome of breast cancer patients receiving the tamoxifen therapy. One of the first studies reported by Goetz *et al.*[42,43] demonstrated that homozygous carriers of *CYP2D6*4* allele had a shorter relapse free survival (RFS) and disease free survival than the patients for heterozygous or homozygous for the wild type allele (hazard ratio (HR), 1.85; $P=0.18$ for RFS: HR, 1.86; $P=0.089$ for disease free survival). Following these reports, Schroth *et al.*[44] published retrospective analysis of 1,325 breast cancer patients with adjuvant tamoxifen monotherapy, and observed that PMs revealed a significantly higher risk of recurrence than EMs with HR of 2.12 for a time to recurrence ($P=0.003$). These associations were supported by several research groups.[45–50] In Asians, we reported the significant effects of *CYP2D6* genotype (especially *CYP2D6*10*) on RFS in Japanese patients receiving adjuvant tamoxifen monotherapy (HR, 9.52; $P=0.000036$).[40,51] The worse clinical outcome of tamoxifen therapy in the patients carrying *CYP2D6*10* was confirmed in Chinese, Korean, Thai and Malaysian populations.[52–55] However, several discordant results have been also reported.[56–61] More recently, two retrospective analyses of large prospective trails, the ATAC (Alimidex, Tamoxifen, Alone or in Combination) trial and the Breast International Group (BIG) 1 98 trial, were reported.[62,63] In the ATAC analysis, there was no significant association between any of CYP2D6 phenotypic groups and recurrence rates in 588 patients treated with tamoxifen (HR, 1.22; $P=0.44$; PM relative to EM).[62] Similarly, in the BIG 1 98 analysis, no significant difference was found among different CYP2D6 metabolizer groups and cancer free survival in 973 breast cancer patients (HR, 0.58; $P=0.35$; EM vs PM).[63] As discussed in previous reports, there may be several confounding factors or critical errors in the experimental designs to explain these discrepancies.

One of the most important issues in the pharmacogenomics study is the quality of genotype data. This should be influenced by (i) the accuracy of genotyping methods, (ii) coverage of genotyped alleles and (iii) DNA source. In both of the ATAC and BIG 1 98 studies,[62,63] the authors mentioned the high reproducibility of genotyping methods because of the concordance of genotyping results in duplicate determinations. However, this does not fully guarantee the accuracy of their genotype results. Their genotype results were highly deviated from Hardy Weinberg equilibrium (χ^2 $P=10^{-92}$ for *CYP2D6*4*) probably because they used the low quality genomic DNA extracted from formalin fixed paraffin embedded tumor tis sues.[64–67] Therefore, they excluded *CYP2D6*5* from the analyses, and performed 60 cycle PCR to detect 1846 G>A (*CYP2D6*4*), which is likely to lead to the misgenotyping results. The importance of wide coverage of *CYP2D6* alleles was clearly demonstrated by Schroth *et al.*[68] In the report, the increase of genotyping coverage was shown to increase HR for RFS as well as enhance the statistical power. In our samples, we also detected a lower HR of 5.83 without *CYP2D6*5* genotyping data than that of 9.52 (*wt/wt* vs *V/V*, $N=282$; unpublished data). In addition, nearly 30% frequency of loss of heterozygosity at the chromosome 22q, where the *CYP2D6* gene is located, in breast cancer cells definitely causes misclassification of patients and leads to misinterpretation of the results if one uses DNAs isolated from tumor tissues (particularly caner cell rich samples).[69]

The second critical issue is selection of study participants. To evaluate the effects of *CYP2D6* genotype on tamoxifen efficacy, it is scientifically certain that the patients treated only with tamoxifen should be selected. As shown in Table 1, most of studies showing the 'null' association included the patients who were treated with a combination of tamoxifen and chemotherapy. We reported significant effects of *CYP2D6* genotypes on shorter RFS when we analyzed patients treated with the tamoxifen monotherapy (HR, 9.52; $P=0.0032$; $N=282$), but not when we analyzed those with the combination chemotherapy (HR, 0.64; $P=0.44$; $N=167$).[70,71] In a combined population (total 449 patents, including 37.2% of those with the combination therapy), HR dropped to 2.45 (95% confidence interval, 1.30 4.54) for *wt/wt* vs *V/V* (unpublished data).

These lines of evidence clearly tell us the importance of complete *CYP2D6* genotyping using germline DNAs isolated from very carefully selected samples with tamoxifen monotherapy. All of 'null' association studies lacked one or multiple elements of these essential factors, as shown in Table 1. Therefore, large prospective studies satisfying these conditions are needed to make a definite conclusion for the value of *CYP2D6* genotyping in tamoxifen therapy.

The patients carrying decreased or impaired function *CYP2D6* alleles consistently showed lower plasma endoxifen concentrations than those having the homozygous normal genotype.[4,8,9,11,39–41] Plasma endoxifen levels were suggested to associate with clinical outcome of tamoxifen treated patients.[72] Therefore, several research groups recently conducted *CYP2D6* genotype based dose adjustment studies.[73–75] Irvin *et al.*[74] demonstrated that endoxifen levels were significantly increased when the dose was increased from 20 40 mg in intermediate metabolizer and PM patients; however, endoxifen levels in PM patients were still significantly lower than the normal individuals. We also investigated the effects of the increase of tamoxifen dose from 20 to 30 mg or 40 mg in the patients heterozygous or homozygous for variant alleles, respectively, and demonstrated that endoxifen concentrations were significantly increased to a similar level of the CYP2D6 normal patients who took 20 mg of tamoxifen (Figure 2).[75] In these studies, the incidence of adverse events was not affected by the dose adjustment. Although further verification is required

Table 1 Summary of studies evaluating association of *CYP2D6* genotype with response to adjuvant tamoxifen therapy

							Association results			
Studies	*Number of patients*	*DNA source*	*Tamoxifen therapy*	*% of monotherapy*	*Tamoxifen dose*	*Outcome*	*Hazard ratio (95% CI)*	*P-value*	CYP2D6*5 *genotyping*	CYP2D6 *groups*[a]
Positive association										
Goetz *et al*[42]	190	FFPE tumors	Monotherapy	100%	20 mg per day for 5 years	DFS	2 44 (1 22–4 90)	0 012	No	*wt/wt + wt/*4 vs *4/*4*
Goetz *et al*[43]	180	FFPE tumors	Monotherapy	100%	20 mg per day for 5 years	RFS	3 20 (1 37–7 55)	0 007	No	*wt/wt vs* PM
Schroth *et al*[45]	206	FFPE tumors	Monotherapy	100%	–[b]	RFS	2 24 (1 16–4 33)	0 02	Yes	EM vs decreased
Newman *et al*[46]	115	PBMC	+ Chemotherapy and/or rad at on	63 5%	20 mg per day med an >4 years	RFS	1 9 (0 8–4 8)	0 19	Yes	*wt/wt + wt/V vs V/V*
K yotan *et al*[51]	58	PBMC	Monotherapy	100%	20 mg per day for 5 years	RFS	10 04 (1 17–86 27)	0 036	Yes	*wt/wt vs *10/*10*
Xu *et al*[52]	152	PBMC	Monotherapy	100%	–	DFS	4 7 (1 1–20 0)	0 04	No	*100C/C + C/T vs T/T*
Schroth *et al*[44]	1 325	PBMC 44 5%[67] Tumor sect ons 55 5%[67]	Monotherapy	100%	For 5 years	RFS	1 49 (1 12–2 00) 2 12 (1 28–3 50)	0 006 0 003	Yes	*wt/wt vs* hetEM/ M *wt/wt vs* PM
B j *et al*[47]	85	PBMC	–	–	–	Breast cancer morta ty	4 1 (1 1–15 9)	0 04	No	*wt/wt vs *4/*4*
K yotan *et al*[40]	282	PBMC	Monotherapy	100%	20 mg per day for 5 years	RFS	4 44 (1 31–15 00) 9 52 (2 79–32 45)	0 017 0 0032	Yes	*wt/wt vs wt/V wt/wt vs V/V*
Ramon *et al*[48]	91	PBMC	+ Chemotherapy	39 8%	–	DFS	–	0 016[c]	Yes	Others vs PM
Park *et al*[53]	110	PBMC	+ Chemotherapy	21 8%	20 mg per day med an 3 9 years	RFS	5 59 (0 93–33 5)	0 05	Yes	EM vs PM
Teh *et al*[54]	95	PBMC	–		20 mg per day	Recurrence event	13 14 (1 54–109 94)[d]	0 004	Yes	EM vs M
Sukasem *et al*[55]	48	PBMC	+ Chemotherapy	6 3%	–	DFS	6 85 (1 48–31 69)	0 01	Yes	EM vs M
Damodaran *et al*[49]	132	PBMC	+ Chemotherapy	6 8%	For 5 years	RFS	7 15 (1 77–28 89)	0 006	Yes	Score ⩽0 5 vs score ⩾1
Goetz *et al*[50]	453	FFPE tumors	Monotherapy	100%	20 mg per day for 5 years	D sease event	2 45 (1 05–5 73)[d]	0 04	No	EM vs PM
Null association										
Nowe *et al*[56]	160	FFPE tumors	+ Chemotherapy and/or rad at on	14 2%	–	DFS	0 67 (0 33–1 35)	0 19	No	*wt/wt vs wt/*4 + *4/*4*
Wegman *et al*[57]	76	Fresh frozen tumors	+ Chemotherapy and/or rad at on	–	40 mg per day for 2 years	RFS	< 1 0[e]	-	No	*wt/wt vs wt/*4 + *4/*4*
Wegman *et al*[58]	103	Fresh frozen tumors	–	–	40 mg per day for 2 years	RFS	0 87 (0 38–1 97)	0 74	No	*wt/wt vs wt/*4 + *4/*4*
	111	Fresh frozen tumors	–	–	40 mg per day for 5 years	RFS	0 33 (0 08–1 43)	0 14	No	*wt/wt vs wt/*4 + *4/*4*
Ok sh ro *et al*[59]	173	PBMC	+ Chemotherapy and/or gosere n	42 2%	20 mg per day med an 52 months	RFS	0 60 (0 18–1 92)	0 39	No	*100C/C + C/T vs T/T*
K yotan *et al*[70]	167	PBMC	+ Chemotherapy	0%	20 mg per day for 5 years	RFS	1 05 (0 48–2 27) 0 64 (0 20–1 99)	0 91 0 44	Yes	*wt/wt vs wt/V wt/wt vs V/V*
Abraham *et al*[60]	3 155	PBMC	+ Chemotherapy	48 4%	20 mg per day	RFS	1 57 (0 64–3 84)	0 32	Yes	*Others vs* PM
Park *et al*[61]	130	PBMC	+ Chemotherapy and/or aromatase nh b tors	18 2%	–	RFS	1 34 (0 42–4 28)	0 63	Yes	*wt/wt + wt/V vs V/V*
Rae *et al*[62]	588	FFPE tumors	+ Chemotherapy	95 7%	20 mg per day for 5 years	RFS	1 22 (0 76–1 96)	0 44	No	EM vs PM
Regan *et al*[63]	973	FFPE tumors	Monotherapy	100%	20 mg per day for 5 years	RFS	0 58 (0 28–1 21)	0 35	No	EM vs PM

Abbrev at ons C confidence nterva DFS d sease-free surv va EM extens ve metabo zer FFPE forma n-fixed paraffin-embedded M ntermed ate metabo zer PBMC per phera b ood mononuc ear ce PM poor metabo zer RFS recurrence-free surv va

Defin t on of a e es *wt *1, *1-*1* or *2 im *9,*10 *10-*10, *17* or *41 pm *3 *4 *5 *6 *14 *21* or *36-*36 *V, im* or *pm*

Defin t on of genotype groups *wt/wt* 2 *wt* a e es EM *wt/wt* or *wt/im* M *im/im* or *im/pm* hetEM/ M *wt/im wt/pm im/im* or *im/pm* PM 2 *pm* a e es decreased *wt/pm im/im im/pm* or *pm/pm*; score ⩽0 5, *im/pm* or *pm/pm*; score ⩽1, *wt/wt, wt/im, wt/pm* or *im/im*

[a]Genotype group was reass gned us ng reported data

[b]Not reported

[c]Log- ank test *P*-va ue

[d]Odds rat o

[e]Not ca cu ated hazard rat o accord ng to *CYP2D6* genotypes

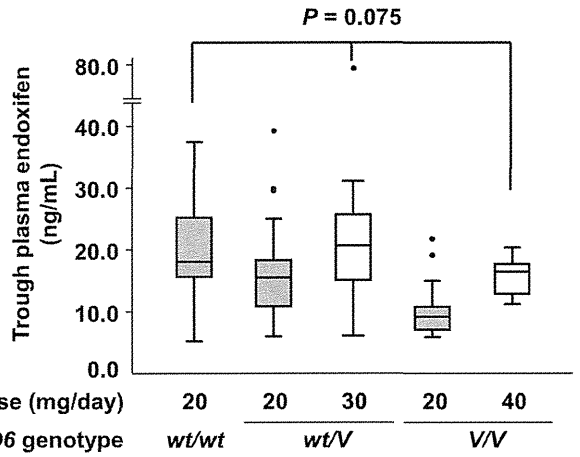

Figure 2 Steady state plasma concentration of endoxifen before and after dose escalation of tamoxifen in breast cancer patients. The horizontal line indicates the median concentration, the box covers the 25th 75th percentiles, and the maximum length of each whisker is 1.5× the interquartile range, dots outside the whiskers are outliers. Data from Kiyotani *et al.*[75]

especially for PM patients, these results suggest that increased tamoxifen dose is an effective way to maintain the effective endoxifen concentration for the patients carrying decreased function or null alleles of *CYP2D6*.

POLYMORPHISMS IN OTHER GENES AND CLINICAL OUTCOME OF TAMOXIFEN THERAPY

Other CYPs, including CYP2C9, CYP2C19, CYP3A4 and CYP3A5, UGTs and SULTs are also involved in the metabolism of tamoxifen. Among them, *CYP3A5*3* is well investigated in association with tamoxifen metabolism or clinical outcome of tamoxifen therapy; however, no significant association was observed.[4,9,42,45,76,77] For *CYP2C19*, a significant association with clinical outcome of tamoxifen treatment was found in carriers of *CYP2C19*17*,[45] but not in the carriers of *CYP2C19*2* or *CYP2C19*3*.[45,59] However, the results have also been contradictive and not conclusive.[78,79] Several investigations on genetic variations in the *SULT1A1* gene, including single nucleotide polymorphisms (SNPs) and copy number variations, found no clear association with tamoxifen efficacy[56,58,79] and tamoxifen metabolism.[9,57] Further analysis would be required by consideration of 'allele copy number' of *SULT1A1*, as demonstrated in the case of *CYP2D6*.[80–82]

There are several reports investigating the involvement of drug transporters in disposition of tamoxifen and its active metabolites, endoxifen and 4 hydroxytamoxifen. ABCB1 (P glycoprotein, multi drug resistance protein 1) is an ATP dependent, efflux transporter with broad substrate specificity widely appreciated for its role in mediating cellular resistance to many anticancer agents.[83] ABCB1 is reported to be involved in the transport of active tamoxifen metabolites.[84,85] Several *ABCB1* polymorphisms have been reported, including 2667 G>A/T and 3435C>T; however, no SNPs were significantly associated with clinical outcome of tamoxifen therapy.[40,54] ABCC2 (multidrug resistance associated protein 2) has an important role in the biliary excretion of glucuronides or sulfates of drugs, including tamoxifen and its metabolites.[17] We found an intronic SNP of *ABCC2* (rs3740065), which is in strong linkage disequilibrium ($r^2 = 0.89$) with 1774 G/delG, to be significantly associated with clinical outcome of patients with tamoxifen therapy through the screening using haplotype tagging SNPs.[40,86] An *in vitro* study reporting that ABCC2 was expressed at higher levels in tamoxifen resistant breast cancer cells suggests the possibility that active metabolites of tamoxifen are transported by ABCC2 from breast cancer cells.[87]

We also identified a novel locus, containing *C10orf11*, associated with RFS in the breast cancer patients treated with tamoxifen alone by the genome wide association study encompassing a total of 462 Japanese patients (HR, 4.53; $P = 6.28 \times 10^{-8}$).[88] At present, however, no report is available regarding the function of the C10orf11 protein. Large scale replication study and further functional analysis are required to verify these associations, and to clarify their biological significance and mechanisms that have effects on the clinical outcome of patients receiving tamoxifen therapy.

OTHER FACTORS AFFECTING CLINICAL OUTCOME OF TAMOXIFEN THERAPY

As well as the genetic polymorphisms modifying the tamoxifen pharmacokinetics, characteristic of cancers, including gene expression profiles or genomic alterations, are also one of important determi nants of individual response to tamoxifen. Many molecules have been identified to be involved in the tamoxifen resistance.[89,90] Several microarray analyses revealed the gene signatures to predict the outcome of adjuvant tamoxifen therapy, such as breast cancer intrinsic subtype,[91,92] 21 gene signature (used as OncotypeDX)[93] and *HOXB13/IL17BR* expression ratio.[94,95] Goetz *et al.*[96] reported that combination of *CYP2D6* genotype and *HOXB13/IL17BR* was significantly associated with disease free survival (log rank $P = 0.004$) and overall survival (log rank $P = 0.009$). More recently, Ellis *et al.*[97] clarified the elevated frequency of somatic mutations and genome structure changes in aromatase inhibitor resistant tumors by whole genome sequencing. Therefore, prediction of individual response to tamoxifen using cancer characteristics seems to be effective, and may affect the association results of genetic markers.

CONCLUSION

Although a large number of investigations on tamoxifen pharmaco genomics have been performed, the association results between *CYP2D6* genotype and clinical outcome are still controversial. However, accumulation of the evidence clarifies some of the causes of these controversial results, particularly some scientific issues in the false negative results, and implies the importance of the quality of genotyping as well as sample selections in the tamoxifen pharmaco genomics study. The important issues learned from the tamoxifen and breast cancer studies are commonly applicable in pharmacogenomics studies. As we are aiming to establish the personalized medicine system in which we select a right patient and provide an appropriate dose of a right drug, the pharmacogenomics study also requires the accurate genotyping using a sufficient number of appropriate patients in order to obtain truly positive results and avoid false positive and false negative results. Finally, genotype guided dose adjustment based on the *CYP2D6* genotypes will be a good example for the personalized medicine. To reduce the medical care cost without losing the quality of medical care, it is very important to use the drugs, which are available at lower cost, on the basis of individual genetic information. As several novel associated SNPs/loci have been identified, integration of genotypes of *CYP2D6* and other genes as well as tumor characteristics should be the future approach to predict clinical efficacy of tamoxifen and provide better quality of lives to breast cancer patients.