

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

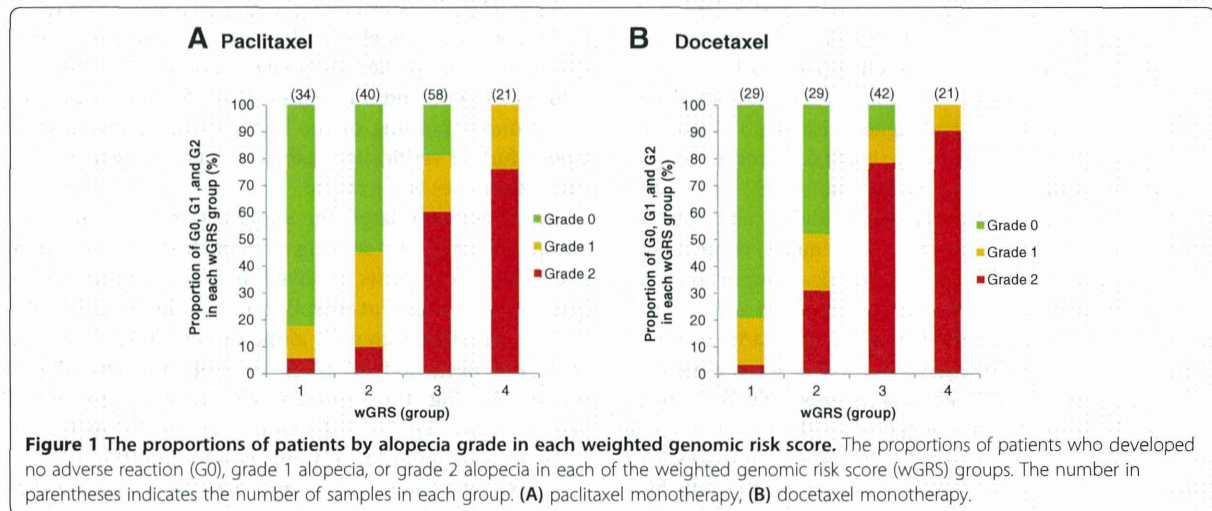
(83%) of 63 patients belonging to groups 3 and 4 developed grade 2 alopecia. These results indicate that our scoring system may be applied to predict severe chemotherapy-induced alopecia and might provide useful information for better understanding of the hair-loss mechanism, even though further verification 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										

^aIndividuals who developed grade 2 alopecia; ^bIndividuals who developed grade 1 alopecia; ^cIndividuals who did not developed any ADRs after chemotherapy; ^dORs 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.



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-FU; 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

¹Department of Medicine, The University of Chicago, 5801 South Ellis Avenue, Chicago, IL 60637, USA. ²Department of Surgery, The University of Chicago, A27 S Maryland Avenue, Chicago, IL 60637, USA. ³Laboratory for Statistical Analysis, Center for Genomic Medicine, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. ⁴Laboratory for Genotyping Development, Center of Genomic Medicine, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. ⁵Tokushima Breast Care Clinic, 4-7-7, Nakashimada-cho, Tokushima 770-0052, Japan. ⁶Laboratory 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

- 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.
- Sotiriou C, Pusztai L: Gene-expression signatures in breast cancer. *N Engl J Med* 2009, **360**:790-800.
- 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.
- Trueb RM: Chemotherapy-induced alopecia. *Semin Cutan Med Surg* 2009, **28**:11-14.
- Carrick S, Parker S, Thornton CE, Ghersi D, Simes J, Wilken N: Single agent versus combination chemotherapy for metastatic breast cancer. *Cochrane Database Syst Rev* 2009, **2**:CD003372.
- Batchelor D: Hair and cancer chemotherapy: consequences and nursing care—a literature study. *Eur J Cancer Care (Engl)* 2001, **10**:147-163.
- McGarvey EL, Baum LD, Pinkerton RC, Rogers LM: Psychological sequelae and alopecia among women with cancer. *Cancer Pract* 2001, **9**:283-289.
- 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.
- 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.
- 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.
- Freedman TG: Social and cultural dimensions of hair loss in women treated for breast cancer. *Cancer Nurs* 1994, **17**:334-341.
- Stenn KS, Paus R: Controls of hair follicle cycling. *Physiol Rev* 2001, **81**:449-494.
- Paus R, Haslam IS, Sharov AA, Botchkarev VA: Pathobiology of chemotherapy-induced hair loss. *Lancet Oncol* 2013, **14**:e50-e59.
- Chon SY, Champion RW, Geddes ER, Rashid RM: Chemotherapy-induced alopecia. *J Am Acad Dermatol* 2012, **67**:e37-e47.
- Grevelman EG, Breed WP: Prevention of chemotherapy-induced hair loss by scalp cooling. *Ann Oncol* 2005, **16**:352-358.
- 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.
- 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.
- 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.
- Wang J, Lu Z, Au JL: Protection against chemotherapy-induced alopecia. *Pharm Res* 2006, **23**:2505-2514.
- http://www.biobankjp.org.
- Nakamura Y: The BioBank Japan Project. *Clin Adv Hematol Oncol* 2007, **5**:696-697.
- 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.
- http://www.cran.r-project.org/.
- 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 WW, *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, Klopper 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.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



REVIEW

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

Kazuma Kiyotani^{1,2}, Taisei Mushiroda², Hitoshi Zembutsu³ 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 phase 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 17 β -estradiol.³

¹Division of Genome Medicine, Institute for Genome Research, The University of Tokushima, Tokushima, Japan; ²Research Group for Pharmacogenomics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan; ³Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan and ⁴Department 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

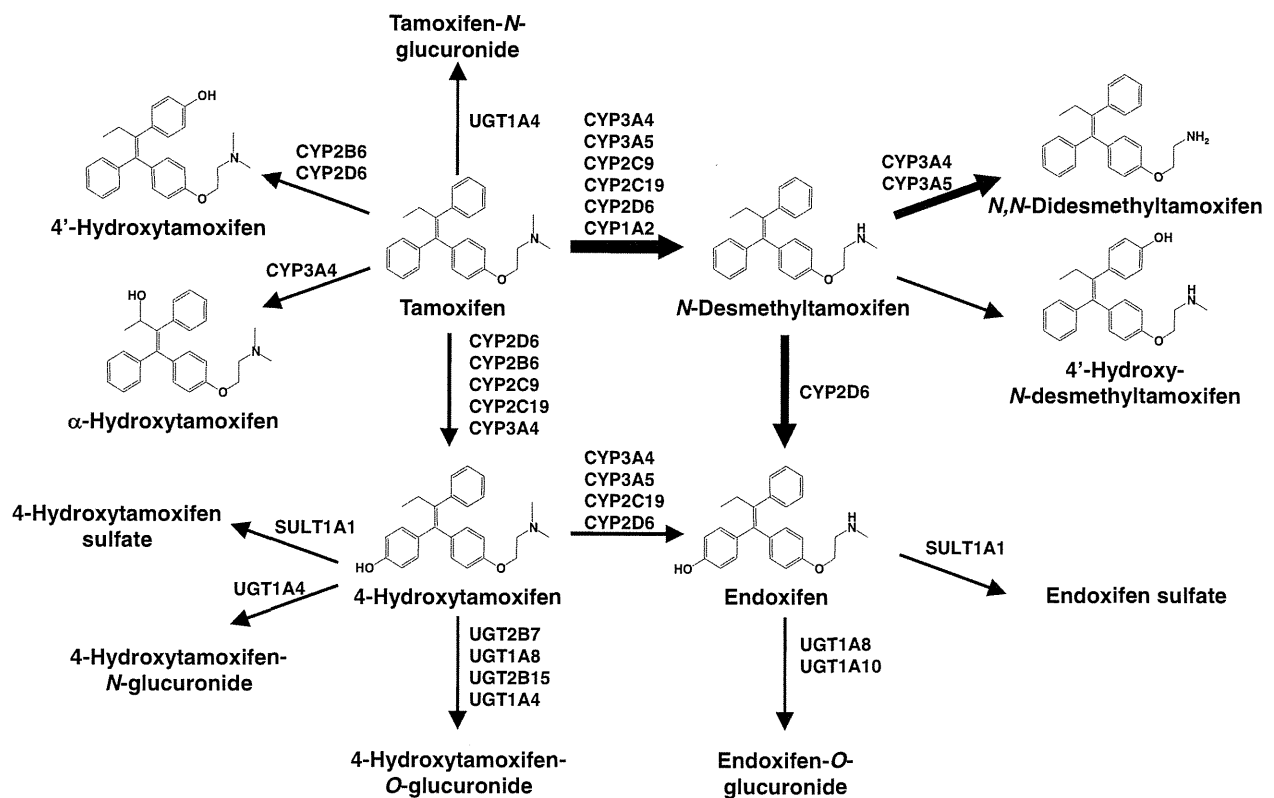


Figure 1 Metabolic pathways of tamoxifen in human. Major metabolic pathways are highlighted with bold arrows.

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-hydroxytamoxifen 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 metabolites, 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 intermediate 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 trials, the ATAC (Alimixen, 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 ($\chi^2 P=10^{-92}$ for *CYP2D6*4*) probably because they used the low-quality genomic DNA extracted from formalin-fixed paraffin-embedded tumor tissues.^{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 cancer-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

Studies	Number of		Tamoxifen therapy	% of		Outcome	Association results			
	patients	DNA source		monotherapy	Tamoxifen dose		Hazard ratio (95% CI)	P-value	CYP2D6*5 genotyping	CYP2D6 groups ^a
<i>Positive association</i>										
Goetz <i>et al.</i> 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 <i>et al.</i> 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 <i>et al.</i> ⁴⁵	206	FFPE tumors	Monotherapy	100%	– ^b	RFS	2.24 (1.16–4.33)	0.02	Yes	EM vs decreased
Newman <i>et al.</i> ⁴⁶	115	PBMC	+ Chemotherapy and/or radiation	63.5%	20 mg per day, median >4 years	RFS	1.9 (0.8–4.8)	0.19	Yes	wt/wt + wt/V vs V/V
Kiyotani <i>et al.</i> 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 <i>et al.</i> ⁵²	152	PBMC	Monotherapy	100%	–	DFS	4.7 (1.1–20.0)	0.04	No	100C/C + C/T vs T/T
Schroth <i>et al.</i> ⁴⁴	1,325	PBMC, 44.5% ⁶⁷ Tumor sections, 55.5% ⁶⁷	Monotherapy	100%	For 5 years	RFS	1.49 (1.12–2.00)	0.006	Yes	wt/wt vs hetEM/IM
							2.12 (1.28–3.50)	0.003		wt/wt vs PM
Bijl <i>et al.</i> ⁴⁷	85	PBMC	–	–	–	Breast cancer mortality	4.1 (1.1–15.9)	0.04	No	wt/wt vs *4/*4
Kiyotani <i>et al.</i> ⁴⁰	282	PBMC	Monotherapy	100%	20 mg per day for 5 years	RFS	4.44 (1.31–15.00)	0.017	Yes	wt/wt vs wt/V wt/wt vs V/V
							9.52 (2.79–32.45)	0.0032		
Ramon <i>et al.</i> ⁴⁸	91	PBMC	+ Chemotherapy	39.8%	–	DFS	–	0.016 ^c	Yes	Others vs PM
Park <i>et al.</i> ⁵³	110	PBMC	+ Chemotherapy	21.8%	20 mg per day, median 3.9 years	RFS	5.59 (0.93–33.5)	0.05	Yes	EM vs PM
Teh <i>et al.</i> ⁵⁴	95	PBMC	–	–	20 mg per day	Recurrence event	13.14 (1.54–109.94) ^d	0.004	Yes	EM vs IM
Sukasem <i>et al.</i> ⁵⁵	48	PBMC	+ Chemotherapy	6.3%	–	DFS	6.85 (1.48–31.69)	0.01	Yes	EM vs IM
Damodaran <i>et al.</i> 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 <i>et al.</i> ⁵⁰	453	FFPE tumors	Monotherapy	100%	20 mg per day for 5 years	Disease event	2.45 (1.05–5.73) ^d	0.04	No	EM vs PM
<i>Null association</i>										
Nowel <i>et al.</i> ⁵⁶	160	FFPE tumors	+ Chemotherapy and/or radiation	14.2%	–	DFS	0.67 (0.33–1.35)	0.19	No	wt/wt vs wt/*4 + *4/*4
Wegman <i>et al.</i> ⁵⁷	76	Fresh frozen tumors	+ Chemotherapy and/or radiation	–	40 mg per day for 2 years	RFS	< 1.0 ^e	–	No	wt/wt vs wt/*4 + *4/*4
Wegman <i>et al.</i> 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
Okishiro <i>et al.</i> ⁵⁹	173	PBMC	+ Chemotherapy and/or goserelin	42.2%	20 mg per day, median 52 months	RFS	0.60 (0.18–1.92)	0.39	No	100C/C + C/T vs T/T
Kiyotani <i>et al.</i> ⁷⁰	167	PBMC	+ Chemotherapy	0%	20 mg per day for 5 years	RFS	1.05 (0.48–2.27)	0.91	Yes	wt/wt vs wt/V wt/wt vs V/V
							0.64 (0.20–1.99)	0.44		
Abraham <i>et al.</i> 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 <i>et al.</i> ⁶¹	130	PBMC	+ Chemotherapy and/or aromatase inhibitors	18.2%	–	RFS	1.34 (0.42–4.28)	0.63	Yes	wt/wt + wt/V vs V/V
Rae <i>et al.</i> 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 <i>et al.</i> 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

Abbreviations: CI, confidence interval; DFS, disease-free survival; EM, extensive metabolizer; FFPE, formalin-fixed paraffin-embedded; IM, intermediate metabolizer; PBMC, peripheral blood mononuclear cell; PM, poor metabolizer; RFS, recurrence-free survival.

Definition of alleles: 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.

Definition of genotype groups: wt/wt, 2 wt alleles; EM, wt/wt or wt/im; IM, im/im or im/pm; hetEM/IM, wt/im, wt/pm, im/im or im/pm; PM, 2 pm alleles; 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.

^aGenotype group was reassigned using reported data.

^bNot reported.

^cLog-link test P-value.

^dOdds ratio.

^eNot calculated hazard ratio according to *CYP2D6* genotypes.

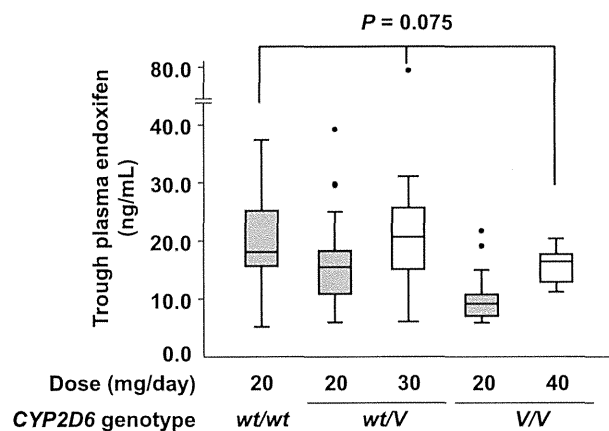


Figure 2 Steady state plasma concentration of endoxifen before and after dose escalation of tamoxifen in breast cancer patients. The horizontal line indicates the median concentration, the box covers the 25th–75th percentiles, and the maximum length of each whisker is 1.5 × the interquartile range, dots 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 contradictory 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, multidrug 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 $-1774\text{G}/\text{delG}$, 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 determinants 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 pharmacogenomics 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 pharmacogenomics 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.

- 1 Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. **365**, 1687–1717 (2005).
- 2 Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* **351**, 1451–1467 (1998).
- 3 Wakeling, A. E. & Slater, S. R. Estrogen-receptor binding and biologic activity of tamoxifen and its metabolites. *Cancer Treat. Rep.* **64**, 741–744 (1980).
- 4 Mürdter, T. E., Schroth, W., Bacchus-Gerybadze, L., Winter, S., Heinkele, G., Simon, W. *et al.* 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).
- 5 Boocock, D. J., Brown, K., Gibbs, A. H., Sanchez, E., Turteltaub, K. W. & White, I. N. Identification of human CYP forms involved in the activation of tamoxifen and irreversible binding to DNA. *Carcinogenesis* **23**, 1897–1901 (2002).
- 6 Crewe, H. K., Notley, L. M., Wunsch, R. M., Lennard, M. S. & Gillam, E. M. Metabolism of tamoxifen by recombinant human cytochrome P450 enzymes: formation of the 4-hydroxy, 4'-hydroxy and *N*-desmethyl metabolites and isomerization of *trans*-4-hydroxytamoxifen. *Drug Metab. Dispos.* **30**, 869–874 (2002).
- 7 Coller, J. K., Krebsfaenger, N., Klein, K., Wolbold, R., Nussler, A., Neuhaus, P. *et al.* Large interindividual variability in the *in vitro* formation of tamoxifen metabolites related to the development of genotoxicity. *Br. J. Clin. Pharmacol.* **57**, 105–111 (2004).
- 8 Borges, S., Desta, Z., Li, L., Skaar, T. C., Ward, B. A., Nguyen, A. *et al.* Quantitative effect of *CYP2D6* genotype and inhibitors on tamoxifen metabolism: implication for optimization of breast cancer treatment. *Clin. Pharmacol. Ther.* **80**, 61–74 (2006).
- 9 Jin, Y., Desta, Z., Stearns, V., Ward, B., Ho, H., Lee, K. H. *et al.* *CYP2D6* genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment. *J. Natl. Cancer Inst.* **97**, 30–39 (2005).
- 10 Johnson, M. D., Zuo, H., Lee, K. H., Trebley, J. P., Rae, J. M., Weatherman, R. V. *et al.* Pharmacological characterization of 4-hydroxy-*N*-desmethyl tamoxifen, a novel active metabolite of tamoxifen. *Breast Cancer Res. Treat.* **85**, 151–159 (2004).
- 11 Stearns, V., Johnson, M. D., Rae, J. M., Morocho, A., Novielli, A., Bhargava, P. *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).
- 12 Clarke, R., Liu, M. C., Bouker, K. B., Gu, Z., Lee, R. Y., Zhu, Y. *et al.* Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. *Oncogene* **22**, 7316–7339 (2003).
- 13 Crewe, H. K., Ellis, S. W., Lennard, M. S. & Tucker, G. T. Variable contribution of cytochromes P450 2D6, 2C9 and 3A4 to the 4-hydroxylation of tamoxifen by human liver microsomes. *Biochem. Pharmacol.* **53**, 171–178 (1997).
- 14 Dehal, S. S. & Kupfer, D. *CYP2D6* catalyzes tamoxifen 4-hydroxylation in human liver. *Cancer Res.* **57**, 3402–3406 (1997).
- 15 Coller, J. K., Krebsfaenger, N., Klein, K., Endrizzi, K., Wolbold, R., Lang, T. *et al.* The influence of *CYP2B6*, *CYP2C9* and *CYP2D6* genotypes on the formation of the potent antioestrogen *Z*-4-hydroxy-tamoxifen in human liver. *Br. J. Clin. Pharmacol.* **54**, 157–167 (2002).
- 16 Borgna, J. L. & Rochefort, H. Hydroxylated metabolites of tamoxifen are formed *in vivo* and bound to estrogen receptor in target tissues. *J. Biol. Chem.* **256**, 859–868 (1981).
- 17 Lien, E. A., Solheim, E., Lea, O. A., Lundgren, S., Kvinnsland, S. & Ueland, P. M. Distribution of 4-hydroxy-*N*-desmethyltamoxifen and other tamoxifen metabolites in human biological fluids during tamoxifen treatment. *Cancer Res.* **49**, 2175–2183 (1989).
- 18 Desta, Z., Ward, B. A., Soukhova, N. V. & Flockhart, D. A. Comprehensive evaluation of tamoxifen sequential biotransformation by the human cytochrome P450 system *in vitro*: prominent roles for *CYP3A* and *CYP2D6*. *J. Pharmacol. Exp. Ther.* **310**, 1062–1075 (2004).
- 19 Falany, J. L., Pilloff, D. E., Leyh, T. S. & Falany, C. N. Sulfation of raloxifene and 4-hydroxytamoxifen by human cytosolic sulfotransferases. *Drug Metab. Dispos.* **34**, 361–368 (2006).
- 20 Gerde, J., Hauglid, M., Breilid, H., Lundgren, S., Varhaug, J. E., Kisanga, E. R. *et al.* Effects of *CYP2D6* and *SULT1A1* genotypes including *SULT1A1* gene copy number on tamoxifen metabolism. *Ann. Oncol.* **19**, 56–61 (2008).
- 21 Nishiyama, T., Ogura, K., Nakano, H., Ohnuma, T., Kaku, T., Hiratsuka, A. *et al.* Reverse geometrical selectivity in glucuronidation and sulfation of *cis*- and *trans*-4-hydroxytamoxifens by human liver UDP-glucuronosyltransferases and sulfotransferases. *Biochem. Pharmacol.* **63**, 1817–1830 (2002).
- 22 Ogura, K., Ishikawa, Y., Kaku, T., Nishiyama, T., Ohnuma, T., Muro, K. *et al.* Quaternary ammonium-linked glucuronidation of *trans*-4-hydroxytamoxifen, an active metabolite of tamoxifen, by human liver microsomes and UDP-glucuronosyltransferase 1A4. *Biochem. Pharmacol.* **71**, 1358–1369 (2006).
- 23 Sun, D., Sharma, A. K., Dellinger, R. W., Blevins-Primeau, A. S., Balliet, R. M., Chen, G. *et al.* Glucuronidation of active tamoxifen metabolites by the human UDP-glucuronosyltransferases. *Drug Metab. Dispos.* **35**, 2006–2014 (2007).
- 24 Kaku, T., Ogura, K., Nishiyama, T., Ohnuma, T., Muro, K. & Hiratsuka, A. Quaternary ammonium-linked glucuronidation of tamoxifen by human liver microsomes and UDP-glucuronosyltransferase 1A4. *Biochem. Pharmacol.* **67**, 2093–2102 (2004).
- 25 Sun, D., Chen, G., Dellinger, R. W., Duncan, K., Fang, J. L. & Lazarus, P. Characterization of tamoxifen and 4-hydroxytamoxifen glucuronidation by human UGT1A4 variants. *Breast Cancer Res.* **8**, R50 (2006).
- 26 Ingelman-Sundberg, M. Genetic polymorphisms of cytochrome P450 2D6 (*CYP2D6*): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics* **5**, 6–13 (2005).
- 27 Bradford, L. D. *CYP2D6* allele frequency in European Caucasians, Asians, Africans and their descendants. *Pharmacogenomics* **3**, 229–243 (2002).
- 28 Sachse, C., Brockmoller, J., Bauer, S. & Roots, I. Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. *Am. J. Hum. Genet.* **60**, 284–295 (1997).
- 29 Griese, E. U., Zanger, U. M., Brudermann, U., Gaedigk, A., Mikus, G., Morike, K. *et al.* Assessment of the predictive power of genotypes for the *in-vivo* catalytic function of *CYP2D6* in a German population. *Pharmacogenetics* **8**, 15–26 (1998).
- 30 Broly, F., Gaedigk, A., Heim, M., Eichelbaum, M., Morike, K. & Meyer, U. A. Debrisoquine/sparteine hydroxylation genotype and phenotype: analysis of common mutations and alleles of *CYP2D6* in a European population. *DNA Cell Biol.* **10**, 545–558 (1991).
- 31 Sachse, C., Brockmoller, J., Hildebrand, M., Muller, K. & Roots, I. Correctness of prediction of the *CYP2D6* phenotype confirmed by genotyping 47 intermediate and poor metabolizers of debrisoquine. *Pharmacogenetics* **8**, 181–185 (1998).
- 32 Nakamura, K., Goto, F., Ray, W. A., McAllister, C. B., Jacqz, E., Wilkinson, G. R. *et al.* Interethnic differences in genetic polymorphism of debrisoquin and mephenytoin hydroxylation between Japanese and Caucasian populations. *Clin. Pharmacol. Ther.* **38**, 402–408 (1985).
- 33 Yokota, H., Tamura, S., Furuya, H., Kimura, S., Watanabe, M., Kanazawa, I. *et al.* Evidence for a new variant *CYP2D6* allele *CYP2D6J* in a Japanese population associated with lower *in vivo* rates of sparteine metabolism. *Pharmacogenetics* **3**, 256–263 (1993).
- 34 Tateishi, T., Chida, M., Ariyoshi, N., Mizorogi, Y., Kamataki, T. & Kobayashi, S. Analysis of the *CYP2D6* gene in relation to dextromethorphan *O*-demethylation capacity in a Japanese population. *Clin. Pharmacol. Ther.* **65**, 570–575 (1999).
- 35 Wang, S. L., Huang, J. D., Lai, M. D., Liu, B. H. & Lai, M. L. Molecular basis of genetic variation in debrisoquin hydroxylation in Chinese subjects: polymorphism in RFLP and DNA sequence of *CYP2D6*. *Clin. Pharmacol. Ther.* **53**, 410–418 (1993).
- 36 Yokoi, T., Kosaka, Y., Chida, M., Chiba, K., Nakamura, H., Ishizaki, T. *et al.* A new *CYP2D6* allele with a nine base insertion in exon 9 in a Japanese population associated with poor metabolizer phenotype. *Pharmacogenetics* **6**, 395–401 (1996).
- 37 Chida, M., Yokoi, T., Nemoto, N., Inaba, M., Kinoshita, M. & Kamataki, T. A new variant *CYP2D6* allele (*CYP2D6*21*) with a single base insertion in exon 5 in a Japanese population associated with a poor metabolizer phenotype. *Pharmacogenetics* **9**, 287–293 (1999).
- 38 Yamazaki, H., Kiyotani, K., Tsubuko, S., Matsunaga, M., Fujieda, M., Saito, T. *et al.* Two novel haplotypes of *CYP2D6* gene in a Japanese population. *Drug Metab. Pharmacokin.* **18**, 269–271 (2003).
- 39 Lim, H. S., Lee, J. H., Lee, S. K., Lee, S. E., Jang, I. J. & Ro, J. Clinical implications of *CYP2D6* genotypes predictive of tamoxifen pharmacokinetics in metastatic breast cancer. *J. Clin. Oncol.* **25**, 3837–3845 (2007).
- 40 Kiyotani, K., Mushirola, T., Imamura, C. K., Hosono, N., Tsunoda, T., Kubo, M. *et al.* Significant effect of polymorphisms in *CYP2D6* and *ABCC2* on clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients. *J. Clin. Oncol.* **28**, 1287–1293 (2010).
- 41 Lim, J. S., Chen, X. A., Singh, O., Yap, Y. S., Ng, R. C., Wong, N. S. *et al.* Impact of *CYP2D6*, *CYP3A5*, *CYP2C9* and *CYP2C19* polymorphisms on tamoxifen pharmacokinetics in Asian breast cancer patients. *Br. J. Clin. Pharmacol.* **71**, 737–750 (2011).
- 42 Goetz, M. P., Rae, J. M., Suman, V. J., Saigren, S. L., Ames, M. M., Visscher, D. W. *et al.* Pharmacogenetics of tamoxifen biotransformation is associated with clinical outcomes of efficacy and hot flashes. *J. Clin. Oncol.* **23**, 9312–9318 (2005).
- 43 Goetz, M. P., Knox, S. K., Suman, V. J., Rae, J. M., Saigren, S. L., Ames, M. M. *et al.* The impact of cytochrome P450 2D6 metabolism in women receiving adjuvant tamoxifen. *Breast Cancer Res. Treat.* **101**, 113–121 (2007).
- 44 Schroth, W., Goetz, M. P., Hamann, U., Fasching, P. A., Schmidt, M., Winter, S. *et al.* Association between *CYP2D6* polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA* **302**, 1429–1436 (2009).
- 45 Schroth, W., Antoniadou, L., Fritz, P., Schwab, M., Muerdter, T., Zanger, U. M. *et al.* Breast cancer treatment outcome with adjuvant tamoxifen relative to patient *CYP2D6* and *CYP2C19* genotypes. *J. Clin. Oncol.* **25**, 5187–5193 (2007).
- 46 Newman, W. G., Hadfield, K. D., Latif, A., Roberts, S. A., Shenton, A., McHague, C. *et al.* Impaired tamoxifen metabolism reduces survival in familial breast cancer patients. *Clin. Cancer Res.* **14**, 5913–5918 (2008).
- 47 Bijl, M. J., van Schaik, R. H., Lammers, L. A., Hofman, A., Vulto, A. G., van Gelder, T. *et al.* The *CYP2D6*4* polymorphism affects breast cancer survival in tamoxifen users. *Breast Cancer Res. Treat.* **118**, 125–130 (2009).
- 48 Ramón, Y., Altés, A., Paré, L., del Río, E., Alonso, C., Barnadas, A. *et al.* Impact of *CYP2D6* polymorphisms in tamoxifen adjuvant breast cancer treatment. *Breast Cancer Res. Treat.* **119**, 33–38 (2010).
- 49 Damodaran, S. E., Pradhan, S. C., Umamaheswaran, G., Kadambari, D., Reddy, K. S. & Adithan, C. Genetic polymorphisms of *CYP2D6* increase the risk for recurrence of breast cancer in patients receiving tamoxifen as an adjuvant therapy. *Cancer Chemother. Pharmacol.* **70**, 75–81 (2012).
- 50 Goetz, M. P., Suman, V. J., Hoskin, T. L., Gnant, M., Filipits, M., Saigren, S. L. *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).
- 51 Kiyotani, K., Mushirola, T., Sasa, M., Bando, Y., Sumitomo, I., Hosono, N. *et al.* Impact of *CYP2D6*10* on recurrence-free survival in breast cancer patients receiving adjuvant tamoxifen therapy. *Cancer Sci.* **99**, 995–999 (2008).

- 52 Xu, Y., Sun, Y., Yao, L., Shi, L., Wu, Y., Ouyang, T. *et al*. Association between CYP2D6*10 genotype and survival of breast cancer patients receiving tamoxifen treatment. *Ann. Oncol.* **19**, 1423–1429 (2008).
- 53 Park, H. S., Choi, J. Y., Lee, M. J., Park, S., Yeo, C. W., Lee, S. S. *et al*. Association between genetic polymorphisms of CYP2D6 and outcomes in breast cancer patients with tamoxifen treatment. *J. Korean Med. Sci.* **26**, 1007–1013 (2011).
- 54 Teh, L. K., Mohamed, N. I., Salleh, M. Z., Rohaizak, M., Shahrin, N. S., Saladina, J. J. *et al*. The risk of recurrence in breast cancer patients treated with tamoxifen: polymorphisms of CYP2D6 and ABCB1. *AAPS J.* **14**, 52–59 (2012).
- 55 Sukasem, C., Sirachainan, E., Chamnanphon, M., Pechatanan, K., Sirisinha, T., Ativitavas, T. *et al*. Impact of CYP2D6 polymorphisms on tamoxifen responses of women with breast cancer: a microarray-based study in Thailand. *Asian Pac. J. Cancer Prev.* **13**, 4549–4553 (2012).
- 56 Nowell, S. A., Ahn, J., Rae, J. M., Scheys, J. O., Trovato, A., Sweeney, C. *et al*. Association of genetic variation in tamoxifen-metabolizing enzymes with overall survival and recurrence of disease in breast cancer patients. *Breast Cancer Res. Treat.* **91**, 249–258 (2005).
- 57 Wegman, P., Vainikka, L., Stal, O., Nordenskjold, B., Skoog, L., Rutqvist, L. E. *et al*. Genotype of metabolic enzymes and the benefit of tamoxifen in postmenopausal breast cancer patients. *Breast Cancer Res.* **7**, R284–R290 (2005).
- 58 Wegman, P., Elingarami, S., Carstensen, J., Stal, O., Nordenskjold, B. & Wingren, S. Genetic variants of CYP3A5, CYP2D6, SUL1A1, UGT2B15 and tamoxifen response in postmenopausal patients with breast cancer. *Breast Cancer Res.* **9**, R7 (2007).
- 59 Okishiro, M., Taguchi, T., Jin Kim, S., Shimazu, K., Tamaki, Y. & Noguchi, S. Genetic polymorphisms of CYP2D6*10 and CYP2C19*2, *3 are not associated with prognosis, endometrial thickness, or bone mineral density in Japanese breast cancer patients treated with adjuvant tamoxifen. *Cancer* **115**, 952–961 (2009).
- 60 Abraham, J. E., Maranian, M. J., Driver, K. E., Platte, R., Kalmyrzaev, B., Baynes, C. *et al*. CYP2D6 gene variants: association with breast cancer specific survival in a cohort of breast cancer patients from the United Kingdom treated with adjuvant tamoxifen. *Breast Cancer Res.* **12**, R64 (2010).
- 61 Park, I. H., Ro, J., Park, S., Lim, H. S., Lee, K. S., Kang, H. S. *et al*. Lack of any association between functionally significant CYP2D6 polymorphisms and clinical outcomes in early breast cancer patients receiving adjuvant tamoxifen treatment. *Breast Cancer Res. Treat.* **131**, 455–461 (2012).
- 62 Rae, J. M., Drury, S., Hayes, D. F., Stearns, V., Thibert, J. N., Haynes, B. P. *et al*. CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. *J. Natl Cancer Inst.* **104**, 452–460 (2012).
- 63 Regan, M. M., Leyland-Jones, B., Bouzyk, M., Pagani, O., Tang, W., Kammler, R. *et al*. 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).
- 64 Pharoah, P. D., Abraham, J. & Caldas, C. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial and Re: CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. *J. Natl Cancer Inst.* **104**, 1263–1264 (2012).
- 65 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 (2012).
- 66 Stanton Jr V. 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**, 1265–1266 (2012).
- 67 Brauch, H., Schroth, W., Goetz, M. P., Mürdtter, T. E., Winter, S., Ingle, J. N. *et al*. Tamoxifen use in postmenopausal breast cancer: CYP2D6 matters. *J. Clin. Oncol.* **31**, 176–180 (2013).
- 68 Schroth, W., Hamann, U., Fasching, P. A., Dauser, S., Winter, S., Eichelbaum, M. *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).
- 69 Hirano, A., Emi, M., Tsuneizumi, M., Utada, Y., Yoshimoto, M., Kasumi, F. *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).
- 70 Kiyotani, K., Mushihiro, T., Hosono, N., Tsunoda, T., Kubo, M., Aki, F. *et al*. Lessons for pharmacogenomics studies: association study between CYP2D6 genotype and tamoxifen response. *Pharmacogenet. Genomics* **20**, 565–568 (2010).
- 71 Kiyotani, K., Mushihiro, T., Nakamura, Y. & Zembutsu, H. Pharmacogenomics of tamoxifen: roles of drug metabolizing enzymes and transporters. *Drug Metab. Pharmacokinet.* **27**, 122–131 (2012).
- 72 Madlensky, L., Natarajan, L., Tchu, S., Pu, M., Mortimer, J., Flatt, S. W. *et al*. Tamoxifen metabolite concentrations, CYP2D6 genotype, and breast cancer outcomes. *Clin. Pharmacol. Ther.* **89**, 718–725 (2011).
- 73 Barginear, M. F., Jaremkó, M., Peter, I., Yu, C., Kasai, Y., Kemeny, M. *et al*. Increasing tamoxifen dose in breast cancer patients based on CYP2D6 genotypes and endoxifen levels: effect on active metabolite isomers and the antiestrogenic activity score. *Clin. Pharmacol. Ther.* **90**, 605–611 (2011).
- 74 Irvin, W. J. Jr., Walko, C. M., Weck, K. E., Ibrahim, J. G., Chiu, W. K., Dees, E. C. *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).
- 75 Kiyotani, K., Mushihiro, T., Imamura, C. K., Tanigawara, Y., Hosono, N., Kubo, M. *et al*. Dose-adjustment study of tamoxifen based on CYP2D6 genotypes in Japanese breast cancer patients. *Breast Cancer Res. Treat.* **131**, 137–145 (2012).
- 76 Tucker, A. N., Tkaczuk, K. A., Lewis, L. M., Tomic, D., Lim, C. K. & Flaws, J. A. Polymorphisms in cytochrome P4503A5 (CYP3A5) may be associated with race and tumor characteristics, but not metabolism and side effects of tamoxifen in breast cancer patients. *Cancer Lett.* **217**, 61–72 (2005).
- 77 Gjerde, J., Geisler, J., Lundgren, S., Ekse, D., Varhaug, J. E., Mellgren, G. *et al*. Associations between tamoxifen, estrogens, and FSH serum levels during steady state tamoxifen treatment of postmenopausal women with breast cancer. *BMC Cancer* **10**, 313 (2010).
- 78 Ruitter, R., Bijl, M. J., van Schaik, R. H., Berns, E. M., Hofman, A., Coebergh, J. W. *et al*. CYP2C19*2 polymorphism is associated with increased survival in breast cancer patients using tamoxifen. *Pharmacogenomics* **11**, 1367–1375 (2010).
- 79 Moyer, A. M., Suman, V. J., Weinsilboum, R. M., Avula, R., Black, J. L., Safgren, S. L. *et al*. SULT1A1, CYP2C19 and disease-free survival in early breast cancer patients receiving tamoxifen. *Pharmacogenomics* **12**, 1535–1543 (2011).
- 80 Hosono, N., Kato, M., Kiyotani, K., Mushihiro, T., Takata, S., Sato, H. *et al*. CYP2D6 genotyping for functional-gene dosage analysis by allele copy number detection. *Clin. Chem.* **55**, 1546–1554 (2009).
- 81 Hosono, N., Kubo, M., Tsuchiya, Y., Sato, H., Kitamoto, T., Saito, S. *et al*. Multiplex PCR-based real-time invader assay (mPCR-RETINA): a novel SNP-based method for detecting allelic asymmetries within copy number variation regions. *Hum. Mutat.* **29**, 182–189 (2008).
- 82 Kiyotani, K., Shimizu, M., Kumai, T., Kamataki, T., Kobayashi, S. & Yamazaki, H. Limited effects of frequent CYP2D6*36*10 tandem duplication allele on *in vivo* dextromethorphan metabolism in a Japanese population. *Eur. J. Clin. Pharmacol.* **66**, 1065–1068 (2010).
- 83 Goda, K., Bacso, Z. & Szabo, G. Multidrug resistance through the spectacle of P-glycoprotein. *Curr. Cancer Drug Targets* **9**, 281–297 (2009).
- 84 Teft, W. A., Mansell, S. E. & Kim, R. B. Endoxifen, the active metabolite of tamoxifen, is a substrate of the efflux transporter P-glycoprotein (multidrug resistance 1). *Drug Metab. Dispos.* **39**, 558–562 (2011).
- 85 Iusuf, D., Teunissen, S. F., Wagenaar, E., Rosing, H., Beijnen, J. H. & Schinkel, A. H. P-glycoprotein (ABCB1) transports the primary active tamoxifen metabolites endoxifen and 4-hydroxytamoxifen and restricts their brain penetration. *J. Pharmacol. Exp. Ther.* **337**, 710–717 (2011).
- 86 Kiyotani, K., Mushihiro, T., Nakamura, Y. & Zembutsu, H. ABC22 and clinical outcome of tamoxifen therapy: reply to T. Lang *et al*. *J. Clin. Oncol.* **28**, e449 (2010).
- 87 Choi, H. K., Yang, J. W., Roh, S. H., Han, C. Y. & Kang, K. W. Induction of multidrug resistance associated protein 2 in tamoxifen-resistant breast cancer cells. *Endocr. Relat. Cancer* **14**, 293–303 (2007).
- 88 Kiyotani, K., Mushihiro, T., Tsunoda, T., Morizono, T., Hosono, N., Kubo, M. *et al*. A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese. *Hum. Mol. Genet.* **21**, 1665–1672 (2012).
- 89 Musgrove, E. A. & Sutherland, R. L. Biological determinants of endocrine resistance in breast cancer. *Nat. Rev. Cancer* **9**, 631–643 (2009).
- 90 Giuliano, M., Schiff, R., Osborne, C. K. & Trivedi, M. V. Biological mechanisms and clinical implications of endocrine resistance in breast cancer. *Breast* **20** (Suppl 3), S42–S49 (2011).
- 91 Sørlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H. *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
- 92 Sørlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A. *et al*. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl Acad. Sci. USA* **100**, 8418–8423 (2003).
- 93 Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M. *et al*. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351**, 2817–2826 (2004).
- 94 Ma, X. J., Wang, Z., Ryan, P. D., Isakoff, S. J., Barmettler, A., Fuller, A. *et al*. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* **5**, 607–616 (2004).
- 95 Jansen, M. P., Sieuwerts, A. M., Look, M. P., Ritsier, K., Meijer-van Gelder, M. E., van Staveren, I. L. *et al*. HOXB13-to-IL17BR expression ratio is related with tumor aggressiveness and response to tamoxifen of recurrent breast cancer: a retrospective study. *J. Clin. Oncol.* **25**, 662–668 (2007).
- 96 Goetz, M. P., Suman, V. J., Couch, F. J., Ames, M. M., Rae, J. M., Erlander, M. G. *et al*. Cytochrome P450 2D6 and homeobox 13/interleukin-17B receptor: combining inherited and tumor gene markers for prediction of tamoxifen resistance. *Clin. Cancer Res.* **14**, 5864–5868 (2008).
- 97 Ellis, M. J., Ding, L., Shen, D., Luo, J., Suman, V. J., Wallis, J. W. *et al*. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature* **486**, 353–360 (2012).

Genome-wide association study of chemotherapeutic agent-induced severe neutropenia/leucopenia for patients in Biobank Japan

Siew-Kee Low,^{1,4} Suyoun Chung,^{4,5} Atsushi Takahashi,¹ Hitoshi Zembutsu,⁴ Taisei Mushiroda,² Michiaki Kubo³ and Yusuke Nakamura^{4,5,6}

Laboratories for ¹Statistical Analysis; ²Pharmacogenetics; ³Genotyping Development, Center for Genomic Medicine, RIKEN, Yokohama; ⁴Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; ⁵Department of Medicine, The University of Chicago, Chicago, Illinois, USA

(Received March 21, 2013/Revised April 22, 2013/Accepted April 22, 2013/Accepted manuscript online May 4, 2013/Article first published online June 10, 2013)

Chemotherapeutic agents are notoriously known to have a narrow therapeutic range that often results in life-threatening toxicity. Hence, it is clinically important to identify the patients who are at high risk for severe toxicity to certain chemotherapy through a pharmacogenomics approach. In this study, we carried out multiple genome-wide association studies (GWAS) of 13 122 cancer patients who received different chemotherapy regimens, including cyclophosphamide- and platinum-based (cisplatin and carboplatin), anthracycline-based (doxorubicin and epirubicin), and antimetabolite-based (5-fluorouracil and gemcitabine) treatment, antimicrotubule agents (paclitaxel and docetaxel), and topoisomerase inhibitors (camptothecin and etoposide), as well as combination therapy with paclitaxel and carboplatin, to identify genetic variants that are associated with the risk of severe neutropenia/leucopenia in the Japanese population. In addition, we used a weighted genetic risk scoring system to evaluate the cumulative effects of the suggestive genetic variants identified from GWAS in order to predict the risk levels of individuals who carry multiple risk alleles. Although we failed to identify genetic variants that surpassed the genome-wide significance level ($P < 5.0 \times 10^{-8}$) through GWAS, probably due to insufficient statistical power and complex clinical features, we were able to shortlist some of the suggestive associated loci. The current study is at the relatively preliminary stage, but does highlight the complexity and problematic issues associated with retrospective pharmacogenomics studies. However, we hope that verification of these genetic variants through local and international collaborations could improve the clinical outcome for cancer patients. (*Cancer Sci* 2013; 104: 1074–1082)

It is now widely and well recognized that medication can cause distinct heterogeneity in terms of its efficacy and toxicity among individuals. These inter-individual differences could be explained in part by the common and/or rare genetic variants in the human genome. Pharmacogenomics aims to discover how genetic variations in the human genome can affect a drug's efficacy or toxicity, and thus brings great promise for personalized medicine in which genetic information can be used to predict the safety, toxicity, and/or efficacy of drugs.⁽¹⁾ Pharmacogenomics study for chemotherapeutic therapies is particularly important because these drugs are known to have a narrow therapeutic window; in general, a higher concentration causes toxicity and a lower concentration reduces the efficacy of the drug. Two of the well-described examples are the association of genetic variants in *TPMT* with 6-mercaptopurine-induced myelosuppression in treatment of pediatric acute lymphoblastic leukemia and that of *UGT1A1* variants with camptothecin-related neutropenia and diarrhea in treatment of

colorectal and lung cancers. The US Food and Drug Administration have recommended that variants on these two genes should be helpful for the prediction of severe adverse reactions prior to use of the drugs.^(2–7)

With advances in various technologies in the life sciences, it is now possible to accurately genotype more than a million common genetic variations by genome-wide high-density SNP array or to characterize all genetic variants in our genome by the next generation DNA sequencing methods. Although one of the greatest drawbacks of GWAS is the requirement of the large number of samples to achieve high statistical power,⁽⁸⁾ this issue could be overcome by the establishment of Biobank Japan in 2003 (<http://biobankjp.org/>).⁽⁹⁾ Biobank Japan collected approximately 330 000 disease cases (200 000 individuals) that had either one or multiples of 47 different diseases including cancers from a collaborative network of 66 hospitals throughout Japan, with the major aim to identify genetic variants associated with susceptibility to complex diseases or those related to drug toxicity. By using the samples from Biobank Japan, a significant number of insightful findings have been published in recent years for identification of common genetic variants associated with complex diseases including cancer.^(10–19) With a reasonable number of samples, it is also feasible to carry out pharmacogenomics studies on chemotherapy-induced toxicity.

Neutropenia and/or leucopenia are two of the most common drug adverse events after treatment with chemotherapeutic agents, which often cause life-threatening infections and the delay of treatment schedule that subsequently affect the treatment outcome. Although prophylactic granulocyte colony-stimulating factor has been given to the patients as a preventive measure,⁽²⁰⁾ the underlying mechanism and susceptible risk factors that cause neutropenia have not been fully elucidated. In this study, we carried out a total of 17 sets of GWAS using 13 122 cancer patients, who received various drug regimens, to identify genetic variants associated with the risk of chemotherapeutic agent-induced severe neutropenia/leucopenia in the Japanese population.

Subjects and Method

Study subjects. A total of 13 122 DNA samples from cancer patients, who received various chemotherapeutic agents, stored in Biobank Japan (University of Tokyo, Tokyo, Japan), were used in this study. Among them, 805 patients developed severe neutropenia and/or leucopenia (\geq grade 3), and 4804 patients

⁶To whom correspondence should be addressed.
E-mails: yusuke@ims.u-tokyo.ac.jp; ynakamura@bsd.uchicago.edu

were not reported to develop any adverse reactions after being given chemotherapeutic agents. The samples could be classified into subgroups according to the drugs used: an alkylating agent (cyclophosphamide); platinum-based (cisplatin and carboplatin), anthracycline-based (doxorubicin and epirubicin); antimetabolite-based (5-fluorouracil and gemcitabine), antimicrotubule-based (paclitaxel and docetaxel); and topoisomerase inhibitor-based (camptothecin and etoposide). The grade of toxicity was classified in accordance with the US National Cancer Institute's Common Toxicity Criteria version 2.0. The adverse event description is based on the medical records collected by the medical coordinator. The patients' demographic details are summarized in Table 1. Participants of this study provided written informed consent and this project was approved by the ethical committee from the Institute of Medical Sciences, University of Tokyo and the RIKEN Center for Genomic Medicine (Yokohama, Japan).

Genotyping and quality controls. DNAs obtained from the patients' blood were genotyped using Illumina OmniExpress BeadChip (San Diego, CA, USA) that contained 733 202 SNPs. Sample quality control was carried out by methods including identity-by-state to evaluate cryptic relatedness for each sample and population stratification by the use of principal component analysis to exclude genetically heterogeneous samples from further analysis.^(21,22) Then, our standard SNP quality control was carried out by excluding SNPs deviating from the Hardy-Weinberg equilibrium ($P \leq 1.0 \times 10^{-6}$), non-polymorphic SNPs, SNPs with a call rate of <0.99 , and those on the X chromosome.^(21,22) Q-Q plot and lambda values, which were calculated between observed P -values from Fisher's test allelic model against expected P -values, were used to further evaluate population substructure.

Statistical analysis. Genome-wide case-control association analyses were evaluated using Fisher's exact method considering allelic, dominant, and recessive genetic models. Manhattan plots of the study were generated using the minimum P -value among the three genetic models for each SNP.

Scoring system using wGRS. The scoring analysis was carried out using SNPs with P_{min} of $<1.0 \times 10^{-5}$ after exclusion of SNPs that are in strong linkage disequilibrium ($r^2 > 0.8$) in each GWAS. The wGRS were calculated according to De Jager

et al.⁽²³⁾ Briefly, we first calculated the weight of each SNP that is the natural log of the odds ratio for each allele/genotype, considering the associated genetic model. For an additive model, we assigned a score of 2 to an individual with two risk alleles, 1 to that with one risk allele, and 0 to that with no risk allele. For a dominant model, we assigned a score of 1 to an individual with one or two risk alleles, and 0 to that with no risk allele. For a recessive model, we assigned a score of 1 to an individual with two risk alleles, and 0 to that with no or one risk allele. Then the cumulative genetic risk scores were determined by multiplying the number of risk alleles/genotype of 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 genetics risk score into four different groups created from the mean and SD: group 1, $< \text{mean} - 1\text{SD}$; group 2, $\text{mean} - 1\text{SD}$ to mean; group 3, mean to $\text{mean} + 1\text{SD}$; and group 4, $> \text{mean} + 1\text{SD}$. Odds ratio, 95% confidence interval, P -value, sensitivity, and specificity were calculated using group 1 as a reference. To calculate the OR in which one of the cells in the contingency table is zero, we applied the Haldane correction, used to avoid error in the calculation by adding 0.5 to all of the cells of a contingency table.

Results

After subdividing the patients by administered drugs/major drug subgroups, as previously mentioned, a total of 17 GWAS analyses were carried out by comparing the allele/genotype frequency between the patients who had developed severe neutropenia/leucopenia (grade 3/4) to those who had not developed any adverse drug reactions. The Q-Q plots of each GWAS and the calculated lambda value of below 1.00 indicated no significant population stratification in each of these GWAS analyses (Fig. S1). From this study, although we could not identify any SNPs that surpassed the genome-wide significant threshold (P -value $< 5 \times 10^{-8}$) for showing association with the risk of neutropenia/leucopenia induced by the certain type of drug or regimen, several possible candidate loci were identified. The results of the GWAS are summarized in Table 2, Table S1, and Figure S2; the results of wGRS are summarized in Table S2.

Table 1. Demographic details of cancer patients treated with chemotherapeutic agents, whose DNA samples are stored in Biobank Japan (The University of Tokyo, Tokyo, Japan)

Category	Controls†	Grade 1/2	Grade 3/4	Category	Controls†	Grade 1/2	Grade 3/4
All	4804	1253	805	Drug subtype			
Age, years (mean)	62.9	58.7	59.6	Alkylating agent	346	266	176
Gender				Cyclophosphamide	335	255	168
Male	2604	424	318	Platinum-based	743	429	428
Female	2200	829	487	Cisplatin	471	191	176
Cancer subtype				Carboplatin	262	207	261
Lung cancer	587	259	266	Anthracycline	459	240	184
Breast cancer	876	388	204	Doxorubicin	66	85	83
Ovarian cancer	140	124	74	Epirubicin	370	132	83
Gastric cancer	827	100	56	Antimetabolite	2249	512	294
Esophageal cancer	208	65	53	5-Fluorouracil	952	331	177
Colorectal cancer	1573	161	50	Gemcitabine	226	111	80
Endometrial cancer	78	72	45	Antimicrotubule agent	825	468	371
Cervical cancer	129	57	35	Paclitaxel	364	321	218
Prostate cancer	91	13	21	Docetaxel	233	143	147
Pancreatic cancer	83	36	20	Topoisomerase inhibitor	187	123	106
Liver cancer	366	16	9	Camptothecin	155	106	59
Gallbladder cancer	56	9	1	Etoposide	39	19	54
				Paclitaxel + carboplatin	166	161	150

†Individuals who did not develop any adverse drug reactions after chemotherapy.

Table 2. Association analysis of single nucleotide polymorphisms (SNPs) with different chemotherapeutic drugs/drug subgroups known to induce severe neutropenia/leucopenia

CHR	SNP	BP	RA	NRA	RAF_Case	RAF_Ctrl	P_allelic	P_dom	P_rec	Pmin	OR	L95	U95	Gene	rel.loci
Cyclophosphamide															
16	rs2519974*	22889186	T	C	0.503	0.381	2.52E-04	4.35E-06	2.77E-01	4.35E-06	1.647	1.264	2.146	<i>HS3ST2</i>	0
1	rs10922438*	198469162	T	C	0.214	0.106	6.01E-06	1.71E-05	7.10E-02	6.01E-06	2.301	1.608	3.293	<i>ATP6V1G3</i>	23190
19	rs3745571*	6475613	T	C	0.778	0.670	4.05E-04	7.72E-06	1.00E+00	7.72E-06	1.730	1.276	2.345	<i>DENND1C</i>	0
All platinum-based drugs															
15	rs4886670*	75449674	A	C	0.320	0.227	9.86E-07	1.43E-05	8.14E-04	9.86E-07	1.605	1.330	1.937	<i>RPL36AP45</i>	29318
19	rs33428*	30937843	G	A	0.481	0.403	2.71E-04	2.78E-06	3.11E-01	2.78E-06	1.375	1.160	1.629	<i>ZNF536</i>	0
14	rs12589282*	22937656	G	T	0.535	0.437	6.30E-06	5.42E-03	4.11E-06	4.11E-06	1.480	1.250	1.752	<i>TRA@</i>	0
3	rs3845905*	66525963	G	A	0.915	0.850	4.12E-06	7.45E-05	3.65E-04	4.12E-06	1.894	1.433	2.503	<i>LRIG1</i>	0
5	rs1895302*	169542600	C	T	0.551	0.478	6.95E-04	7.41E-06	4.11E-01	7.41E-06	1.340	1.132	1.587	<i>FOXI1</i>	5871
1	rs16825455*	21837755	T	C	0.686	0.605	8.85E-05	8.62E-06	1.90E-01	8.62E-06	1.425	1.193	1.702	<i>ALPL</i>	0
Cisplatin															
7	rs10253216*	16861849	T	C	0.565	0.468	2.18E-03	1.68E-07	1.00E+00	1.68E-07	1.478	1.155	1.891	<i>AGR2</i>	-17111
4	rs11944965*	63424089	T	C	0.807	0.678	3.45E-06	1.68E-06	6.65E-02	1.68E-06	1.986	1.475	2.676	<i>LOC644534</i>	47600
7	rs7797977*	16862235	C	A	0.668	0.580	4.06E-03	5.23E-01	2.17E-06	2.17E-06	1.457	1.127	1.883	<i>AGR2</i>	-17497
18	rs2406342*	74488280	T	G	0.605	0.475	3.59E-05	2.48E-06	6.71E-02	2.48E-06	1.697	1.323	2.177	<i>ZNF236</i>	-47836
20	rs6077251*	7752366	T	C	0.271	0.153	2.50E-06	3.64E-06	3.06E-02	2.50E-06	2.065	1.537	2.773	<i>SFRS13AP2</i>	59982
8	rs11774576*	27740417	A	G	0.702	0.581	6.78E-05	2.82E-06	2.62E-01	2.82E-06	1.699	1.307	2.208	<i>SCARA5</i>	0
11	rs4627050*	18822037	G	A	0.781	0.649	4.43E-06	9.01E-06	2.17E-02	4.43E-06	1.932	1.452	2.572	<i>PTPN5</i>	-8648
1	rs12142335*	108302922	A	G	0.040	0.004	9.91E-06	8.45E-06	1.00E+00	8.45E-06	9.713	3.175	29.710	<i>VAV3</i>	0
Carboplatin															
15	rs11071200*	55950082	T	G	0.060	0.008	1.25E-06	8.51E-07	1.00E+00	8.51E-07	8.241	2.888	23.520	<i>PRTG</i>	0
5	rs3822735*	35799994	G	A	0.862	0.752	7.24E-06	1.68E-06	3.50E-01	1.68E-06	2.062	1.500	2.834	<i>SPEF2</i>	0
3	rs1623879*	58027197	G	A	0.441	0.321	7.69E-05	1.89E-02	3.75E-06	3.75E-06	1.669	1.297	2.148	<i>FLNB</i>	0
15	rs936229*	75132319	G	A	0.713	0.595	7.27E-05	4.41E-06	3.01E-01	4.41E-06	1.685	1.302	2.180	<i>ULK3</i>	0
13	rs7989332*	21050575	G	T	0.853	0.738	5.47E-06	1.16E-04	1.74E-03	5.47E-06	2.056	1.507	2.806	<i>CRYL1</i>	0
3	rs3845905*	66525963	G	A	0.921	0.828	5.99E-06	2.20E-05	1.50E-02	5.99E-06	2.433	1.645	3.598	<i>LRIG1</i>	0
8	rs1714746*	4105147	G	A	0.554	0.435	1.59E-04	6.63E-02	7.44E-06	7.44E-06	1.610	1.261	2.056	<i>CSMD1</i>	0
16	rs12446319*	81774798	A	G	0.253	0.143	8.97E-06	1.27E-04	3.16E-04	8.97E-06	2.026	1.480	2.774	<i>CMIP</i>	29431
1	rs1277203*	109392837	A	G	0.730	0.626	3.50E-04	3.51E-02	9.38E-06	9.38E-06	1.615	1.243	2.098	<i>AKNAD1</i>	0
All anthracycline-based drugs															
5	rs10040979*	158424391	G	A	0.701	0.618	4.68E-03	5.35E-01	4.60E-07	4.60E-07	1.452	1.120	1.883	<i>EBF1</i>	0
2	rs12615435*	200638509	T	G	0.883	0.773	3.95E-06	4.09E-06	9.17E-02	3.95E-06	2.214	1.555	3.154	<i>LOC348751</i>	0
5	rs7720283*	158459721	C	T	0.775	0.706	1.29E-02	3.37E-01	4.15E-06	4.15E-06	1.431	1.078	1.898	<i>EBF1</i>	0
1	rs1367448*	68633924	C	T	0.633	0.526	5.02E-04	5.32E-06	5.12E-01	5.32E-06	1.554	1.212	1.993	<i>LOC100289178</i>	0
6	rs2505059*	98495952	G	A	0.538	0.398	5.41E-06	2.29E-04	2.04E-04	5.41E-06	1.765	1.383	2.252	<i>MIR2113</i>	23455
12	rs4149639*	6442001	C	T	0.120	0.047	7.39E-06	1.52E-05	8.16E-02	7.39E-06	2.763	1.781	4.287	<i>TNFRSF1A</i>	0
19	rs1654260*	20329111	A	G	0.625	0.488	8.49E-06	3.03E-03	2.18E-05	8.49E-06	1.749	1.365	2.240	<i>LOC100421704</i>	-3576
Doxorubicin															
15	rs11857176*	78164706	A	G	0.657	0.515	1.74E-02	1.00E+00	8.08E-07	8.08E-07	1.800	1.127	2.874	<i>LOC100302666</i>	-6274
2	rs4380275*	773278	G	A	0.392	0.152	4.99E-06	1.54E-05	2.05E-02	4.99E-06	3.604	2.041	6.365	<i>LOC339822</i>	6559
11	rs2512987*	86414282	T	C	0.681	0.417	7.02E-06	6.04E-03	3.42E-05	7.02E-06	2.985	1.855	4.803	<i>ME3</i>	-30604
Epirubicin															
12	rs4149639*	6442001	C	T	0.163	0.042	2.89E-07	8.31E-07	3.32E-02	2.89E-07	4.443	2.571	7.677	<i>TNFRSF1A</i>	0
5	rs2964475*	5407814	C	A	0.615	0.415	4.13E-06	3.95E-04	1.10E-04	4.13E-06	2.248	1.592	3.174	<i>KIAA0947</i>	-14993
13	rs1923834*	28360487	G	A	0.916	0.770	9.40E-06	4.61E-06	3.33E-01	4.61E-06	3.236	1.823	5.744	<i>GSX1</i>	-6293
10	rs908366*	126144839	A	G	0.518	0.328	7.04E-06	1.08E-04	8.69E-04	7.04E-06	2.199	1.564	3.092	<i>LHPP</i>	-5502

Table 2. (continued)

CHR	SNP	BP	RA	NRA	RAF_Case	RAF_Ctrl	P_allelic	P_dom	P_rec	Pmin	OR	L95	U95	Gene	rel.loci
3	rs1553091*	187716886	G	A	0.452	0.358	2.65E-02	8.04E-01	7.46E-06	7.46E-06	1.480	1.053	2.080	<i>LOC100505844</i>	-22691
All antimetabolite drugs															
18	rs7228133*	4539085	C	A	0.733	0.686	2.26E-02	6.64E-01	1.70E-06	1.70E-06	1.255	1.034	1.522	<i>LOC284215</i>	243085
21	rs8127977*	26826514	A	G	0.804	0.722	1.30E-05	2.11E-06	2.67E-01	2.11E-06	1.587	1.282	1.966	<i>NCRNA00158</i>	-22501
12	rs894734*	54319727	G	A	0.849	0.776	3.84E-05	3.97E-06	6.63E-01	3.97E-06	1.619	1.279	2.050	<i>HOXC13</i>	-12849
13	rs9580312*	22754093	G	A	0.480	0.409	1.35E-03	8.09E-06	6.25E-01	8.09E-06	1.330	1.120	1.581	<i>LOC100506622</i>	-30331
21	rs2055011*	19481354	C	T	0.184	0.143	1.12E-02	2.12E-01	8.82E-06	8.82E-06	1.347	1.075	1.686	<i>CHODL</i>	-135796
12	rs12582168*	124894184	C	T	0.333	0.256	8.50E-05	9.31E-06	2.84E-01	9.31E-06	1.454	1.210	1.748	<i>NCOR2</i>	0
5-Fluorouracil															
7	rs10488226*	12713070	A	C	0.195	0.107	1.09E-05	3.54E-06	2.98E-01	3.54E-06	2.026	1.500	2.737	<i>LOC100505995</i>	-12175
2	rs6740660*	224943685	G	A	0.966	0.894	4.10E-06	8.83E-06	2.40E-01	4.10E-06	3.386	1.870	6.131	<i>SERPINE2</i>	-39649
4	rs1567482*	36026747	G	A	0.952	0.875	6.26E-06	1.44E-05	9.14E-02	6.26E-06	2.846	1.716	4.719	<i>LOC651644</i>	39948
2	rs6706693*	192465598	A	G	0.328	0.219	1.62E-05	2.12E-03	9.45E-06	9.45E-06	1.743	1.362	2.232	<i>OBFC2A</i>	-77200
Gemcitabine															
18	rs9961113*	75605399	C	T	0.625	0.403	1.43E-06	3.83E-04	3.73E-05	1.43E-06	2.473	1.706	3.584	<i>LOC100421527</i>	-260017
5	rs2547917*	58713680	A	G	0.350	0.212	9.06E-04	8.79E-02	3.33E-06	3.33E-06	1.997	1.345	2.965	<i>PDE4D</i>	0
15	rs12900463*	85415386	C	T	0.219	0.115	2.24E-03	1.02E-01	4.03E-06	4.03E-06	2.154	1.342	3.457	<i>ALPK3</i>	0
22	rs9609078*	31153276	T	C	0.089	0.009	4.32E-06	9.97E-06	2.59E-01	4.32E-06	10.890	3.528	33.610	<i>OSBP2</i>	0
5	rs6863418*	173625154	A	G	0.175	0.055	1.37E-05	6.98E-06	4.55E-01	6.98E-06	3.623	2.042	6.429	<i>HMP19</i>	88972
20	rs6037430*	344079	G	A	0.894	0.730	9.74E-06	1.75E-05	7.92E-02	9.74E-06	3.109	1.805	5.359	<i>NRSN2</i>	8567
All antimicrotubule drugs															
17	rs11651483*	12777402	C	T	0.729	0.665	1.69E-03	2.60E-01	3.37E-07	3.37E-07	1.357	1.120	1.643	<i>RICH2</i>	0
6	rs4235898*	77266188	A	G	0.830	0.740	1.05E-06	3.34E-06	6.50E-03	1.05E-06	1.718	1.377	2.142	<i>LOC100131680</i>	-103976
13	rs4771859*	93088651	G	A	0.764	0.709	5.49E-03	2.60E-01	1.47E-06	1.47E-06	1.328	1.088	1.623	<i>GPC5</i>	0
1	rs12145418*	216716320	T	G	0.334	0.274	3.04E-03	3.17E-01	2.35E-06	2.35E-06	1.331	1.104	1.604	<i>ESRRG</i>	0
6	rs9386485*	106329055	T	C	0.596	0.492	2.65E-06	2.85E-05	8.59E-04	2.65E-06	1.524	1.279	1.817	<i>PRDM1</i>	-205140
16	rs12935229*	77328895	A	G	0.249	0.182	1.83E-04	2.23E-02	4.40E-06	4.40E-06	1.495	1.214	1.840	<i>ADAMTS18</i>	0
7	rs6961860*	17085321	G	A	0.557	0.495	5.33E-03	8.87E-01	4.67E-06	4.67E-06	1.283	1.078	1.527	<i>LOC100131425</i>	156806
14	rs12882718*	86902054	T	C	0.737	0.643	5.91E-06	5.55E-06	2.03E-02	5.55E-06	1.555	1.283	1.884	<i>LOC100421119</i>	-42891
12	rs1043763*	122630909	T	C	0.668	0.574	1.36E-05	6.51E-06	3.10E-02	6.51E-06	1.496	1.248	1.795	<i>MLXIP</i>	1920
2	rs4591358*	196365890	C	T	0.302	0.215	6.60E-06	6.89E-04	1.89E-04	6.60E-06	1.578	1.297	1.920	<i>LOC391470</i>	81627
14	rs8022296*	97987857	G	A	0.663	0.608	1.06E-02	6.54E-01	7.29E-06	7.29E-06	1.269	1.058	1.521	<i>LOC100129345</i>	111127
4	rs6817170*	154374984	G	A	0.377	0.286	9.29E-06	4.61E-05	3.27E-03	9.29E-06	1.517	1.264	1.822	<i>KIAA0922</i>	-12514
Paclitaxel															
1	rs922106*	90025519	T	G	0.298	0.202	2.17E-04	1.95E-02	9.28E-07	9.28E-07	1.679	1.277	2.207	<i>LRRC8B</i>	0
6	rs9386485*	106329055	T	C	0.624	0.477	1.17E-06	1.43E-05	5.24E-04	1.17E-06	1.821	1.429	2.320	<i>PRDM1</i>	-205140
8	rs2444896*	99022009	T	G	0.727	0.603	1.58E-05	2.43E-06	7.41E-02	2.43E-06	1.754	1.355	2.269	<i>MATN2</i>	0
2	rs4666360*	20335709	C	T	0.216	0.114	4.62E-06	3.24E-06	2.27E-01	3.24E-06	2.136	1.546	2.950	<i>RPS16P2</i>	19625
9	rs3138083*	35648950	A	G	0.220	0.117	3.78E-06	1.24E-05	1.09E-02	3.78E-06	2.136	1.551	2.942	<i>SIT1</i>	345
17	rs3786094*	9875205	C	T	0.528	0.422	5.30E-04	1.80E-01	5.83E-06	5.83E-06	1.531	1.206	1.944	<i>GAS7</i>	0
15	rs4886670*	75449674	A	C	0.353	0.229	7.26E-06	5.38E-05	4.26E-03	7.26E-06	1.835	1.412	2.383	<i>RPL36AP45</i>	29318
5	rs792975*	172271007	T	C	0.654	0.519	7.66E-06	1.19E-04	6.08E-04	7.66E-06	1.746	1.366	2.232	<i>ERGIC1</i>	0
Docetaxel															
9	rs3747851*	124521260	T	C	0.337	0.176	5.61E-07	1.12E-05	5.63E-04	5.61E-07	2.377	1.693	3.339	<i>DAB2IP</i>	0
7	rs4727963*	122759980	C	T	0.772	0.618	7.99E-06	1.04E-06	1.69E-01	1.04E-06	2.094	1.505	2.914	<i>SLC13A1</i>	0
14	rs1756650*	87741025	G	A	0.211	0.162	9.95E-02	9.10E-01	1.74E-06	1.74E-06	1.386	0.954	2.014	<i>GALC</i>	658333

Table 2. (continued)

CHR	SNP	BP	RA	NRA	RAF_Case	RAF_Ctrl	P_allelic	P_dom	P_rec	Pmin	OR	L95	U95	Gene	rel.loci
13	rs488248*	106596719	T	C	0.918	0.795	3.29E-06	3.23E-05	1.17E-02	3.29E-06	2.896	1.802	4.655	LOC728192	-432192
6	rs12660691*	130008445	A	C	0.935	0.819	3.62E-06	4.99E-06	1.62E-01	3.62E-06	3.199	1.899	5.391	ARHGAP18	0
18	rs4553720*	62170726	T	C	0.377	0.281	7.77E-03	4.56E-01	6.77E-06	6.77E-06	1.547	1.131	2.116	LOC284294	79890
6	rs2157460*	130021128	T	C	0.932	0.820	7.07E-06	1.07E-05	1.62E-01	7.07E-06	3.013	1.806	5.025	ARHGAP18	0
2	rs837841*	130034012	T	G	0.714	0.662	1.49E-01	5.26E-01	8.34E-06	8.34E-06	1.279	0.930	1.759	LOC151121	-33653
All topoisomerase inhibitors															
5	rs10074959*	104208013	T	C	0.321	0.237	3.23E-02	8.08E-01	1.13E-06	1.13E-06	1.524	1.048	2.217	RAB9P1	-227162
7	rs1035147*	12094966	T	G	0.981	0.877	4.01E-06	4.75E-06	5.56E-01	4.01E-06	7.294	2.587	20.559	TMEM106B	-155882
1	rs303386*	99589379	A	G	0.585	0.444	1.10E-03	4.30E-06	3.88E-01	4.30E-06	1.766	1.256	2.483	LOC100129620	0
14	rs7494275*	56231800	C	A	0.543	0.406	1.86E-03	4.26E-01	8.50E-06	8.50E-06	1.732	1.232	2.433	RPL13AP3	-1163
3	rs480409*	7010081	G	A	0.495	0.348	6.08E-04	1.63E-01	9.28E-06	9.28E-06	1.842	1.307	2.596	GRM7	0
Camptothecin															
6	rs17318866*	3837198	G	A	0.966	0.790	1.47E-06	1.78E-06	1.91E-01	1.47E-06	7.559	2.689	21.263	FAM50B	-12434
2	rs17027130*	41273631	C	T	0.644	0.387	2.54E-06	8.03E-04	3.91E-05	2.54E-06	2.865	1.844	4.452	LOC729984	-110074
1	rs303386*	99589379	A	G	0.627	0.429	3.34E-04	3.61E-06	1.06E-01	3.61E-06	2.238	1.448	3.460	LOC100129620	0
Etoposide															
20	rs6039763*	10183517	A	G	0.370	0.090	1.27E-05	1.54E-06	4.61E-01	1.54E-06	5.966	2.502	14.230	LOC100131208	0
1	rs2506991*	48098406	G	A	0.593	0.269	1.39E-05	1.11E-02	2.28E-06	2.28E-06	3.948	2.101	7.418	LOC388630	127794
2	rs12987465*	49715021	A	G	0.593	0.359	1.87E-03	6.26E-01	6.61E-06	6.61E-06	2.597	1.424	4.737	F5HR	-333355
7	rs3095008*	20255705	T	C	1.000	0.846	1.74E-05	9.39E-06	1.00E+00	9.39E-06	inf	N/A	N/A	MACC1	0
Paclitaxel + carboplatin															
12	rs12310399*	95490248	A	G	0.708	0.567	2.60E-04	1.08E-01	2.46E-07	2.46E-07	1.852	1.329	2.580	FGD6	0
9	rs10785877*	137125501	T	C	0.833	0.660	7.38E-07	1.54E-05	1.98E-04	7.38E-07	2.580	1.766	3.769	RXRA	-92815
19	rs995834*	28866596	C	T	0.580	0.425	1.28E-04	1.20E-06	1.51E-01	1.20E-06	1.871	1.364	2.566	LOC100420587	307385
1	rs922107*	90022796	G	A	0.323	0.211	1.55E-03	9.07E-02	2.73E-06	2.73E-06	1.788	1.250	2.558	LRRC8B	0
7	rs1425132*	37562368	T	C	0.740	0.666	4.55E-02	7.35E-01	4.68E-06	4.68E-06	1.430	1.013	2.017	LOC442668	62349
1	rs6429703*	15339960	T	C	0.200	0.078	8.40E-06	2.59E-05	5.61E-02	8.40E-06	2.942	1.802	4.804	RP1-21018.1	0

*SNPs used for weighted genetic risk score analyses. BP, SNP genomic location; CHR, chromosome; inf, infinity; L95, lower 95% confidence interval; N/A, not applicable; NRA, non-risk allele; OR, odds ratio; P_allelic, *P*-value from allelic model; P_dom, *P*-value from dominant model; P_min, minimum *P*-value among the three models; P_rec, *P*-value from recessive model; RA, risk allele; RAF, risk allele frequency; rel.loci, distance of the SNP from the gene; U95, upper 95% confidence interval.

Among these datasets, GWAS carried out using samples who were given: (i) any kind of platinum-based chemotherapy (428 cases vs 743 controls); (ii) cisplatin-based chemotherapy (176 cases vs 471 controls); or (iii) carboplatin-based chemotherapy (261 cases vs 262 controls) identified SNPs showing the most significant association with chemotherapy-induced severe neutropenia/leucopenia are: rs4886670 ($P_{min} = 9.86 \times 10^{-7}$, OR = 1.61, 95% CI = 1.33–1.94) near *RPL36AP45* for (i); rs10253216 ($P_{min} = 1.68 \times 10^{-7}$, OR = 1.48, 95% CI = 1.16–1.89) near *AGR2* for (ii); and rs11071200 ($P_{min} = 8.51 \times 10^{-7}$, OR = 8.24, 95% CI = 2.89–23.5) on *PRTG* for (iii) (Table 2, Table S1, Fig. S2b). For the anthracycline-based regimen, we carried out GWAS with individuals given all anthracycline-based (184 cases vs 459 controls), doxorubicin-based (83 cases vs 66 controls), and epirubicin-based (83 cases vs 370 controls) chemotherapy, and identified three SNPs, rs10040979 ($P_{min} = 4.60 \times 10^{-7}$, OR = 1.45, 95% CI = 1.12–1.88) in *EBF1*, rs11857176 ($P_{min} = 8.08 \times 10^{-7}$, OR = 1.80, 95% CI = 1.13–2.87) near a hypothetical gene *LOC100302666*, and rs4149639 ($P_{min} = 2.89 \times 10^{-7}$, OR = 4.44, 95% CI = 2.57–7.68) in *TNFRSF1A*, to be most significantly associated with the risk of high-grade neutropenia/leucopenia, respectively (Table 2, Table S1, Fig. S2c). In the case of antimicrotubule agents, we carried out three different GWAS with individuals who were treated with antimicrotubule (371 cases vs 825 controls), paclitaxel-based (218 cases vs 364 controls), or docetaxel-based (147 cases vs 233 controls) regimens. We identified three SNPs, rs11651483 ($P_{min} = 3.37 \times 10^{-7}$, OR = 1.36, 95% CI = 1.12–1.64) in *RICH2*, rs922106 ($P_{min} = 9.28 \times 10^{-7}$, OR = 1.68, 95% CI = 1.28–2.21) in *LRRC8B* and rs3747851 ($P_{min} = 5.61 \times 10^{-7}$, OR = 2.38, 95% CI = 1.69–3.34) in *DAB2IP*, to be those most significantly associated with the increased risk of severe neutropenia/leucopenia, respectively (Table 2, Table S1, Fig. S2e). Our previous report by Kiyotani *et al.*⁽²⁴⁾ identified four SNPs to be associated with gemcitabine-induced hematological toxicities. Three of the four SNPs were included in the current study with suggestive association, rs12046844 ($P_{min} = 5.84 \times 10^{-4}$, OR = 2.53, 95% CI = 1.45–4.43), rs6430443 ($P_{min} = 8.61 \times 10^{-4}$, OR = 6.33, 95% CI = 1.90–22.2; $r^2 = 0.895$ with rs1901440) and rs11719165 ($P_{min} = 1.16 \times 10^{-2}$, OR = 2.36, 95% CI = 1.18–4.70) (Table S4). However, it is noted that some of the samples used in this study overlapped with those in the study reported by Kiyotani *et al.*, as both sourced samples from Biobank Japan.

Lastly, we also attempted to identify genetic variants associated with combined treatment of paclitaxel and carboplatin-induced severe neutropenia/leucopenia (150 cases vs 166 controls), as this combined treatment is commonly used as the standard therapy for both ovarian and lung cancers. We found the most significant association with the SNP rs12310399 ($P_{min} = 2.46 \times 10^{-7}$, OR = 1.85, 95% CI = 1.33–2.58) near the *FGD6* gene (Table 2, Table S1, Fig. S2a), which is

suggested to activate CDC42, a member of the Ras-like family of Rho and Rac proteins, and has a critical role in regulating the actin cytoskeleton. The second strongest association was observed at the locus encoding RXRA ($P_{min} = 7.38 \times 10^{-7}$, OR = 2.58, 95% CI = 1.77–3.77), an important transcriptional factor. We also calculated the cumulative genetic scores using SNPs on six loci and identified that individuals in group 4 could have 188 times (95% CI = 36.1–979) higher risk of developing severe neutropenia/leucopenia than those belonging to group 1 with the sensitivity of 95.9% and the specificity of 88.9% (Table S2). Because this drug combination is of clinical importance, we further investigated the association of these six selected loci using 161 individuals who developed grade 1/2 neutropenia/leucopenia, using cases registered in the Biobank Japan. Interestingly, the association results for the six loci were moderate for grade 1/2 neutropenia/leucopenia, with intermediate allele frequency and OR between individuals without any adverse reactions and those with neutropenia/leucopenia of \geq grade 3 (Table S3). In addition, as shown in Table 3 and Figure 1, the higher the calculated score becomes, the higher the proportion and grade of neutropenia/leucopenia. The intermediate scores for patients with grade 1/2 neutropenia/leucopenia could imply the possible usefulness of this scoring system for the prediction.

Furthermore, we used simulation to estimate how many samples are required to validate this scoring result. We started off by estimating the incidence of neutropenia/leucopenia by the combined treatment of paclitaxel and carboplatin. In Biobank Japan, a total of 477 individuals received this combined treatment; among them, 166 individuals (35%) did not develop any adverse drug reactions, 161 (35%) developed mild neutropenia/leucopenia (grade 1 or 2) and 150 (30%) developed severe neutropenia/leucopenia (grade 3 or higher). The frequency of developing severe neutropenia/leucopenia is in agreement with a multicenter study reported by Guastalla *et al.*⁽²⁵⁾ When we assume that 100 patients who receive this combination therapy are prospectively registered, the incidences of the adverse drug reactions are estimated as shown in Table 4. If we categorize the patients by wGRS according to the proportions indicated in Table 3 (and our hypothesis is right), the statistical power should be enough to validate by this small subset of patients. Even if two individuals in both group 1 and group 4 are incorrectly predicted, the calculated *P*-value is still 0.03 by Fisher's exact test.

Discussion

In this study, we carried out GWAS analyses for a total of 17 subsets of chemotherapies to identify genetic variants that might be associated with chemotherapeutic-induced neutropenia/leucopenia with grades 3 and 4, however, we could not identify any SNPs that surpassed the genome-wide significant threshold (P -value $< 5 \times 10^{-8}$). Through this study, we

Table 3. Weighted genetic risk score (wGRS) analysis of cancer patients who received combination treatment with paclitaxel and carboplatin

wGRS group	Score	G3G4	G1G2	G0	%_G3G4	%_G1G2	%_G0	G3/4 versus G0			G1/2 versus G0		
								OR	95%_CI	<i>P</i> -value	OR	95%_CI	<i>P</i> -value
1	<5.802	2	21	48	0.03	0.29	0.68	REF			REF		
2	5.802–7.665	36	58	77	0.21	0.34	0.45	11.2	2.58–48.70	8.69E-05	1.72	0.93–3.19	9.55E-02
3	7.665–9.528	64	62	33	0.40	0.39	0.21	46.5	10.60–204.00	2.36E-13	4.29	2.21–8.35	1.61E-05
4	>9.528	47	20	6	0.64	0.28	0.08	188.0	36.10–979.00	4.78E-20	7.62	2.68–21.70	6.08E-05
Total		149	161	164									

95%_CI, 95% confidence interval; G0, individuals who did not develop any adverse drug reaction; G1G2, grade 1 and grade 2 neutropenia (mild); G3G4, grade 3 and grade 4 neutropenia (severe); OR, odds ratio; REF, reference.

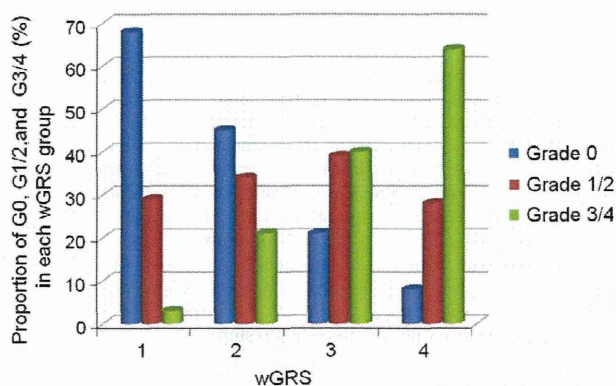


Fig. 1. Proportions of cancer patients who developed no adverse reaction (G0), mild neutropenia/leucopenia (G1/2), or severe neutropenia/leucopenia (G3/4) in each of the weighted genetic risk score (wGRS) score groups. All patients received combined treatment with paclitaxel and carboplatin and were registered with Biobank Japan. The total numbers of patients in scores 1, 2, 3, and 4 are 71, 171, 159, and 73, respectively.

Table 4. Simulation of weighted genetic risk score (wGRS) analysis for a prospective study of 100 patients who received combination treatment with paclitaxel and carboplatin

Estimated verification samples (n = 100; 35 expected to have grade 1/2 neutropenia)					
wGRS group	G3G4	G0	OR	95%_CI	P-value
1	0	10			
2	7	17	9.0	0.47–174.00	7.82E-02
3	13	7	37.8	1.93–740.00	1.06E-03
4	10	1	147.0	5.35–4040.00	3.40E-05
Total	30	35			

95%_CI, 95% confidence interval; G0, individuals without any adverse drug reaction; G3G4, grade 3 and 4 neutropenia (severe); OR, odds ratio.

encountered several important issues, which are now common problems in pharmacogenomics studies using retrospective clinical data, including confounding factors and heterogeneous treatments for individual patients (often given different combinations of drugs, different dosage of drugs, and different time-periods of treatment), that increase the complexity of studies and generate various noises in the analyses, and diminished the statistical power in the case-control association studies. We understand that our current approach was not an ideal study design, but it is not easy to perfectly standardize therapy in the daily clinical practice of cancer treatment. There are several factors contributing to the variability in treatments: (i) there is some preference by doctors or by hospitals to select a particular regimen among the various recommended standard treatments; (ii) the modifications (adjustments) of the dosage or schedule according to the patient's conditions (performance status, results of laboratory tests, etc.); and (iii) although we have been collecting the clinical information, it is not perfect to collect complete clinical information in some hospitals, particularly those that do not use electronic medical records. One can say that this kind of study should be performed as a prospective design, however, due to the very rapid advances in the development of novel molecular-targeted drugs and new regimens in the oncology area, the protocols have been and

will be modified or improved. Hence, spending many years and a huge budget on a prospective study may result in a clinically useless outcome, because the results are unable to be applied due to the replacement of the study protocol with a new protocol, when the results of association studies are available. Nevertheless, retrospective pharmacogenomic studies could be improved by implementing electronic medical record systems that could include detailed descriptions of patients' conditions and their responses to various drugs.

Although we understand the pitfalls in study designs like our present study, we need to seek possible ways to identify candidate genetic variants that might contribute to improvement in the clinical management of cancer patients, including chemotherapy-induced severe neutropenia/leucopenia. Nevertheless, some of the candidate genes that we identified are of interest, considering their known functions as well as their relations with drug actions. For example, the proto-oncogene *AGR2*, whose genetic variants were suggested to associate with cisplatin-induced neutropenia/leucopenia, encodes an anterior gradient 2 homolog (*Xenopus laevis*) that is known to play a critical role in cell migration, cell differentiation, and cell growth.⁽²⁶⁾ Cells stably expressing *AGR2* confer resistance to cisplatin *in vivo*, compared with control cells (empty vector) in a xenograft animal model.⁽²⁷⁾ The second example is *TNFRSF1A*, suggested to be associated with anthracycline-based and epirubicin-induced neutropenia/leucopenia. This gene encodes TNFRSF1A, which is a major receptor for TNF- α . The soluble TNFRSF1A level was found to be elevated after 1 month of anthracycline-based chemotherapy.⁽²⁸⁾ Additionally, both TNF- α and TNFRSF1A are known to play a critical role in doxorubicin-induced cardiotoxicity, in which doxorubicin stimulates an increase in circulating TNF and upregulates TNFRSF1A.^(29,30) Furthermore, genetic variants on *PDE4D*, which encodes for phosphodiesterase 4D, cAMP-specific, showed suggestive association with gemcitabine-induced severe neutropenia/leucopenia. Ablation of *PDE4D* has been reported to impair the neutrophil function with altered chemotaxis ability and adhesion capability as well as to reduce neutrophil recruitment to the site of inflammation.⁽³¹⁾ Besides, genetic variants on *RXRA* identified to be associated with combined treatment of paclitaxel and carboplatin-induced severe neutropenia/leucopenia, encodes retinoid X receptor alpha. Disruption of this gene in mouse models moderately alters lymphocyte proliferation and survival, and affects the T helper type1/type 2 balances.⁽³²⁾ All of these genes might provide some important insights into the mechanism of various chemotherapy-induced severe neutropenia/leucopenia, however, further validations are definitely essential.

As already described, the GWAS approach could provide a list of genetic variants that might be associated with complex phenotypes (drug responsiveness or drug-induced adverse reactions) in pharmacogenomics studies. One of the clinically important aims for identification of the associated genetic variants is to establish a prediction model to identify individuals who are at risk of adverse reactions with certain drugs or protocols. In this study, we have applied the wGRS system, by which we could distinguish high-risk patients from low-risk individuals by counting the number of risk alleles of the suggestively-associated SNPs in combination with estimating the effect size of each SNP. One of the best examples from this study was indicated by a scoring system using six candidate SNP loci that were identified through the GWAS of severe neutropenia/leucopenia caused by combination treatment of paclitaxel and carboplatin; among 53 individuals in the high-risk group (group 4) by this scoring method, 47 (89%) revealed high-grade neutropenia/leucopenia. In contrast, among 50 individuals in the low-risk group (group 1), only 2 (4%) revealed high-grade neutropenia/leucopenia, and

the odds ratio to have the severe adverse reaction in individuals belonging to group 4 was calculated to be 188 times higher than those categorized to group 1 (Table 3). Interestingly, individuals who developed grade 1/2 (mild neutropenia/leucopenia) were found to show intermediate risk scores between patients with severe neutropenia/leucopenia and those without any adverse reactions. Hence, we suggest that wGRS is an applicable method to evaluate the clinical utility of possible variants with specific phenotypes. However, the data are preliminary and require verification by an independent test sample(s) before any definitive conclusions can be drawn. But, considering that the OR of the high-risk group is very high, the number of samples required for the verification (if our hypothesis is right) is not so large. In fact, we have tried to simulate a prospective study design using a model of 100 patients according to the assumption that 35% individuals will not develop any adverse drug reactions, 35% individuals will develop mild neutropenia/leucopenia (grade 1/2), and 30% will develop severe neutropenia/leucopenia (grade 3/4). As shown in Table 4, the study of 100 patients should have very strong statistical power to verify. If this is verified, as we expect, it should improve the quality of lives of cancer patients and also contribute to reducing medical care costs by avoiding unnecessary adverse events. However, to achieve success in pharmacogenomics and personalized medicine, both local and international collaborative efforts are essential.

References

- 1 Wheeler HE, Maitland ML, Dolan ME, Cox NJ, Ratain MJ. Cancer pharmacogenomics: strategies and challenges. *Nat Rev Genet* 2013; **14**: 23–34.
- 2 Relling MV, Hancock ML, Rivera GK *et al*. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J Natl Cancer Inst* 1999; **91**: 2001–8.
- 3 Weinsilboum RM, Sladek SL. Mercaptopurine pharmacogenetics: monogenic inheritance of erythrocyte thiopurine methyltransferase activity. *Am J Hum Genet* 1980; **32**: 651–62.
- 4 Relling MV, Klein TE. CPIC: clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clin Pharmacol Ther* 2011; **89**: 464–7.
- 5 Innocenti F, Undevia SD, Iyer L *et al*. Genetic variants in the UDP-glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. *J Clin Oncol* 2004; **22**: 1382–8.
- 6 Iyer L, Das S, Janisch L *et al*. UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J* 2002; **2**: 43–7.
- 7 Hoskins JM, Goldberg RM, Qu P, Ibrahim JG, McLeod HL. UGT1A1*28 genotype and irinotecan-induced neutropenia: dose matters. *J Natl Cancer Inst* 2007; **99**: 1290–5.
- 8 McCarthy MI, Abecasis GR, Cardon LR *et al*. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**: 356–69.
- 9 Triendl R. Japan launches controversial Biobank project. *Nat Med* 2003; **9**: 982.
- 10 Takata R, Akamatsu S, Kubo M *et al*. Genome-wide association study identifies five new susceptibility loci for prostate cancer in the Japanese population. *Nat Genet* 2010; **42**: 751–4.
- 11 Akamatsu S, Takata R, Haiman CA *et al*. Common variants at 11q12, 10q26 and 3p11.2 are associated with prostate cancer susceptibility in Japanese. *Nat Genet* 2012; **44**: 426–9, S1.
- 12 Cui R, Kamatani Y, Takahashi A *et al*. Functional variants in ADH1B and ALDH2 coupled with alcohol and smoking synergistically enhance esophageal cancer risk. *Gastroenterology* 2009; **137**: 1768–75.
- 13 Cui R, Okada Y, Jang SG *et al*. Common variant in 6q26-q27 is associated with distal colon cancer in an Asian population. *Gut* 2011; **60**: 799–805.
- 14 Kumar V, Kato N, Urabe Y *et al*. Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet* 2011; **43**: 455–8.
- 15 Low SK, Kuchiba A, Zembutsu H *et al*. Genome-wide association study of pancreatic cancer in Japanese population. *PLoS ONE* 2010; **5**: e11824.
- 16 Elgazzar S, Zembutsu H, Takahashi A *et al*. A genome-wide association study identifies a genetic variant in the SIAH2 locus associated with hor-

Acknowledgments

We would like to express our heartfelt gratitude to all the patients who participated in this study. We convey our sincere appreciation to Dr Teruhiko Yoshida and Dr Hiromi Sakamoto from the National Cancer Center Research Institute for their kind support. Our thanks also goes to the members of the laboratory for statistical analysis and the laboratory for the genotyping development from the Center for Genomic Medicine for their kind support and fruitful discussions. We would like to extend our gratitude to the staff of Biobank Japan for their outstanding assistance. This work was carried out as part of the Biobank Japan Project, supported by the Ministry of Education, Culture, Sports, Sciences and Technology, Japan. In addition, this project was supported by the JSPS postdoctoral fellowship.

Disclosure Statement

The authors have no conflict of interest.

Abbreviations

GWAS	genome-wide association study
OR	odds ratio
SD	standard deviation
SNP	single nucleotide polymorphism
TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A
TNF- α	tumor necrosis factor- α
wGRS	weighted genetic risk score

- 17 Miki D, Ochi H, Hayes CN *et al*. Variation in the DEPDC5 locus is associated with progression to hepatocellular carcinoma in chronic hepatitis C virus carriers. *Nat Genet* 2011; **43**: 797–800.
- 18 Miki D, Kubo M, Takahashi A *et al*. Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. *Nat Genet* 2010; **42**: 893–6.
- 19 Cha PC, Zembutsu H, Takahashi A, Kubo M, Kamatani N, Nakamura Y. A genome-wide association study identifies SNP in DCC is associated with gallbladder cancer in the Japanese population. *J Hum Genet* 2012; **57**: 235–7.
- 20 Silvestris N, Del Re M, Azzariti A *et al*. Optimized granulocyte colony-stimulating factor prophylaxis in adult cancer patients: from biological principles to clinical guidelines. *Expert Opin Ther Targets* 2012; **16**(Suppl 2): S111–7.
- 21 Yamaguchi-Kabata Y, Nakazono K, Takahashi A *et al*. Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am J Hum Genet* 2008; **83**: 445–56.
- 22 Low SK, Takahashi A, Cha PC *et al*. Genome-wide association study for intracranial aneurysm in the Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. *Hum Mol Genet* 2012; **21**: 2102–10.
- 23 De Jager PL, Chibnik LB, Cui J *et al*. Integration of genetic risk factors into a clinical algorithm for multiple sclerosis susceptibility: a weighted genetic risk score. *Lancet Neurol* 2009; **8**: 1111–9.
- 24 Kiyotani K, Uno S, Mushirola T *et al*. A genome-wide association study identifies four genetic markers for hematological toxicities in cancer patients receiving gemcitabine therapy. *Pharmacogenet Genomics* 2012; **22**: 229–35.
- 25 Guastalla JP, Pujade-Lauraine E, Weber B *et al*. Efficacy and safety of the paclitaxel and carboplatin combination in patients with previously treated advanced ovarian carcinoma. A multicenter GINECO (Group d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens) phase II study. *Ann Oncol* 1998; **9**: 37–43.
- 26 Wang Z, Hao Y, Lowe AW. The adenocarcinoma-associated antigen, AGR2, promotes tumor growth, cell migration, and cellular transformation. *Cancer Res* 2008; **68**: 492–7.
- 27 Hengel SM, Murray E, Langdon S *et al*. Data-independent proteomic screen identifies novel tamoxifen agonist that mediates drug resistance. *J Proteome Res* 2011; **10**: 4567–78.
- 28 Perik PJ, De Vries EG, Boomsma F *et al*. The relation between soluble apoptotic proteins and subclinical cardiotoxicity in adjuvant-treated breast cancer patients. *Anticancer Res* 2006; **26**: 3803–11.