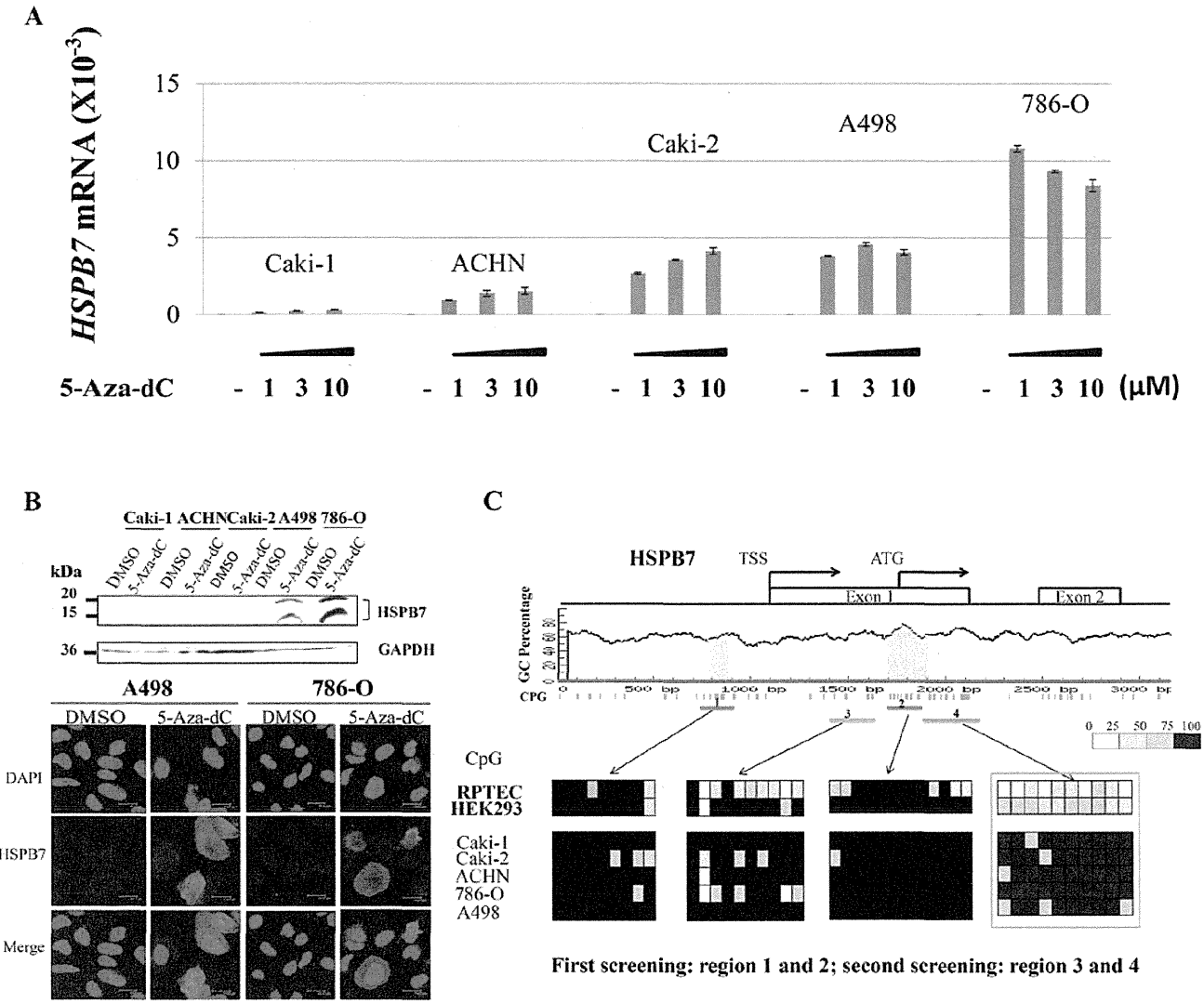

Figure 3. Epigenetic silencing of HSPB7 in RCC cell lines. (A) qPCR analysis and (B) western blot and ICC analysis of HSPB7 were performed in five RCC cell lines with treatment of the demethylating agent 5-Aza-dC. B2M was used for normalization of mRNA expression levels. GAPDH was used for normalization of protein expression levels. Values are expressed as the mean ± SD. (C) Hypermethylation of HSPB7 was confirmed by means of bisulfite sequencing. For each of the regions 1-4 in the cell lines, 10 or more colonies were randomly chosen and sequenced. Each square indicates a CpG site, and an average methylation level per CpG site is indicated by % methylation (shown in different color): white, 0-25% methylation; bright grey, 26-50% methylation; dark grey, 51-75% methylation; and black, 76-100% methylation. Region 4 showed higher level of methylation in the five RCC cell lines (Caki-1, Caki-2, ACHN, 786-O and A498) than in the two control cell lines (RPTEC and HEK293).

sion vector, pCAGGSnHC-HSPB7-HA. Introduction of HSPB7 into these two cancer cell lines caused significant decrease in the number of colonies, compared with corresponding mock-transfected controls (Fig. 4A). We also performed colony formation assay in 3 other RCC cell lines (Caki-2, A498 and 786-O) using the same vectors, and confirmed similar growth-suppressive effects (Fig. 4B), implying that HSPB7 may function as a tumor suppressor gene.

HSPB7 is regulated by p53. To further elucidate the biological significance, we first investigated its possible involvement in the p53-pathway because α B-crystallin, one of the small heat shock protein family members, was reported to be induced by p53 (22,23). We applied qPCR analysis to evaluate the expression of HSPB7 in NCI-H1299 (p53 null) cell lines with or without introduction of p53 using the adenovirus system. After the infection of Ad-p53, we observed induction of HSPB7 in a dose- and time-dependent manner (Fig. 5A and B), while no induction was observed in the control cells. After the 48-hour treatment with 40 MOI of Ad-p53, the expression level of HSPB7 became nearly 5 times higher than the control cells (Fig. 5A). Induction of HSPB7 was also confirmed under the treatment with relative lower dose of Ad-p53 (8 MOI) at different time points. Concordantly, DNA damage by adriamycin treatment induced HSPB7 expression in HCT116 cells with wild-type p53, but not in HCT116 cells without wild-type p53 (Fig. 5C and D), indicating that HSPB7 expression is regulated by wild-type p53. To further investigate whether HSPB7 is directly regulated by p53, we screened two possible p53-binding sites indicated by the p53-binding site search

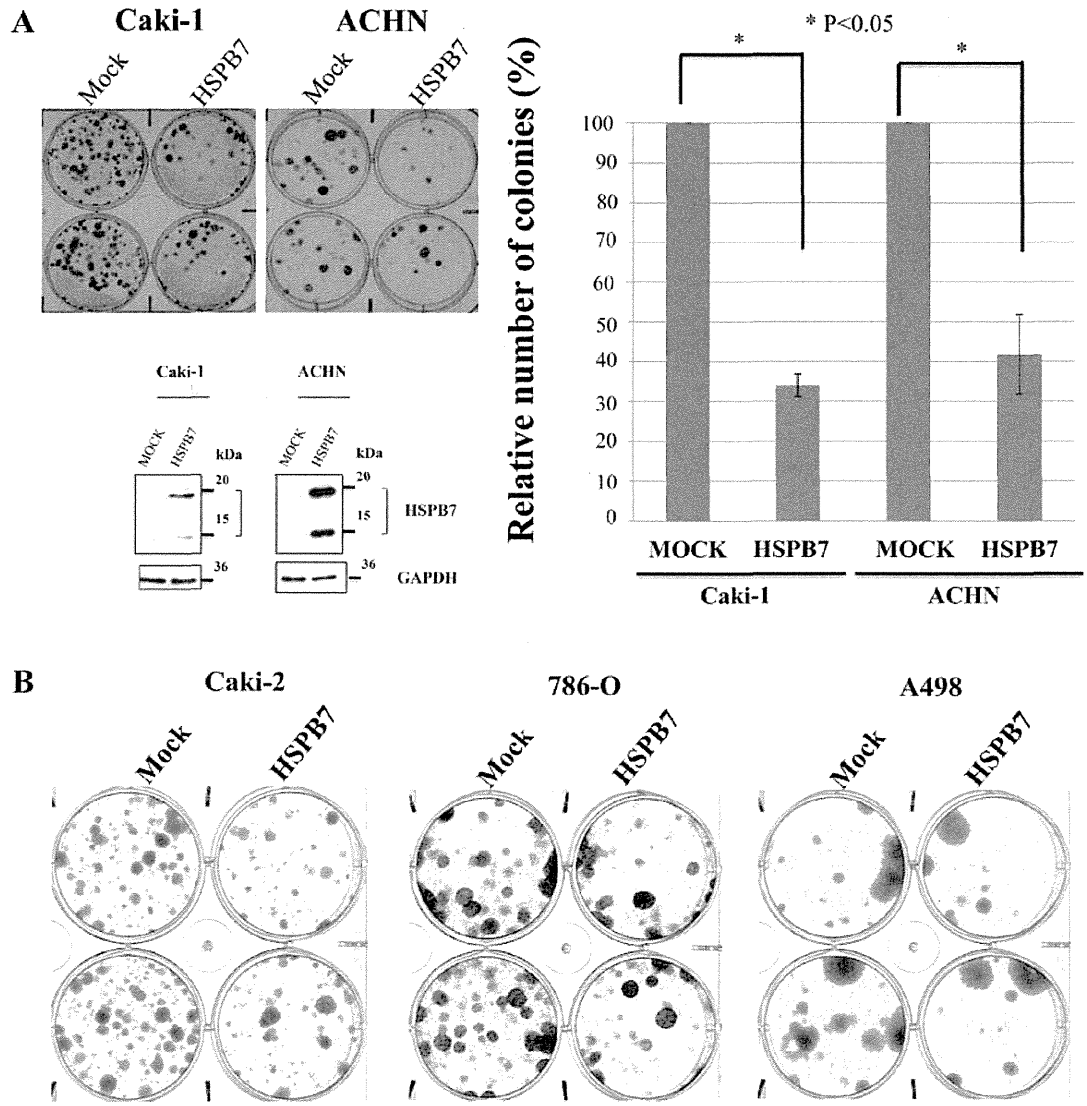

Figure 4. Ectopic HSPB7 expression suppresses RCC cell growth. (A) Colony formation assay showed that introduction of HSPB7 impaired colony-forming ability of Caki-1 and ACHN cells. Cells were transfected with plasmid expressing HSPB7 or mock plasmid, and colonies (>1 mm diameter) were counted after selection of 2-3 weeks with G418. At 48 h after transfection, total protein of cells was collected and applied for western blot to confirm the successful transfection. GAPDH was used for the normalization of protein expression levels. (B) Colony formation assay in Caki-2, 786-O and A498 RCC cell lines. Values are expressed as the mean ± SD.

software developed by us, but neither of these two candidate sites was confirmed to be a direct p53-binding site (data not shown). Although there might be another site(s) that p53 binds to, we are unable to conclude whether HSPB7 is directly or indirectly regulated by p53, it is certain that HSPB7 expression is inducible by wild-type p53.

Discussion

Scarce knownledge exists on the biological function of HSPB7, a member of the small heat shock protein family that is characterized by possessing a conserved α-crystallin domain. HSPB7 has been shown to interact with the cytoskeletal protein α-filamin (24) as well as other small heat shock proteins (25). HSPB7 belongs to a non-canonical HSPB protein that prevents the aggregation of polyQ proteins in an active autophagy machinery, but overexpression of HSPB7 alone did not affect the autophagy event (26). Several genome-wide association studies found that SNPs in the HSPB7 gene were strongly associated with idiopathic cardiomyopathies and heart failure (27-31). Recently, HSPB7 was suggested to regulate early developmental steps in cardiac morphogenesis (32). However, the involvement of HSPB7 in carcinogenesis has not been described.

Through the genome-wide expression analysis in RCCs, we identified HSPB7 as a candidate tumor suppressor gene because of its common and significant downregulation in RCCs. Subsequent functional analysis revealed that HSPB7 was downregulated in cancer cells by hypermethylation. Bisulfite sequencing of genomic regions of HSPB7 confirmed hypermethylation in RCC cell lines. Although region 4 (Fig. 3C) contained no CpG Island, we observed significantly

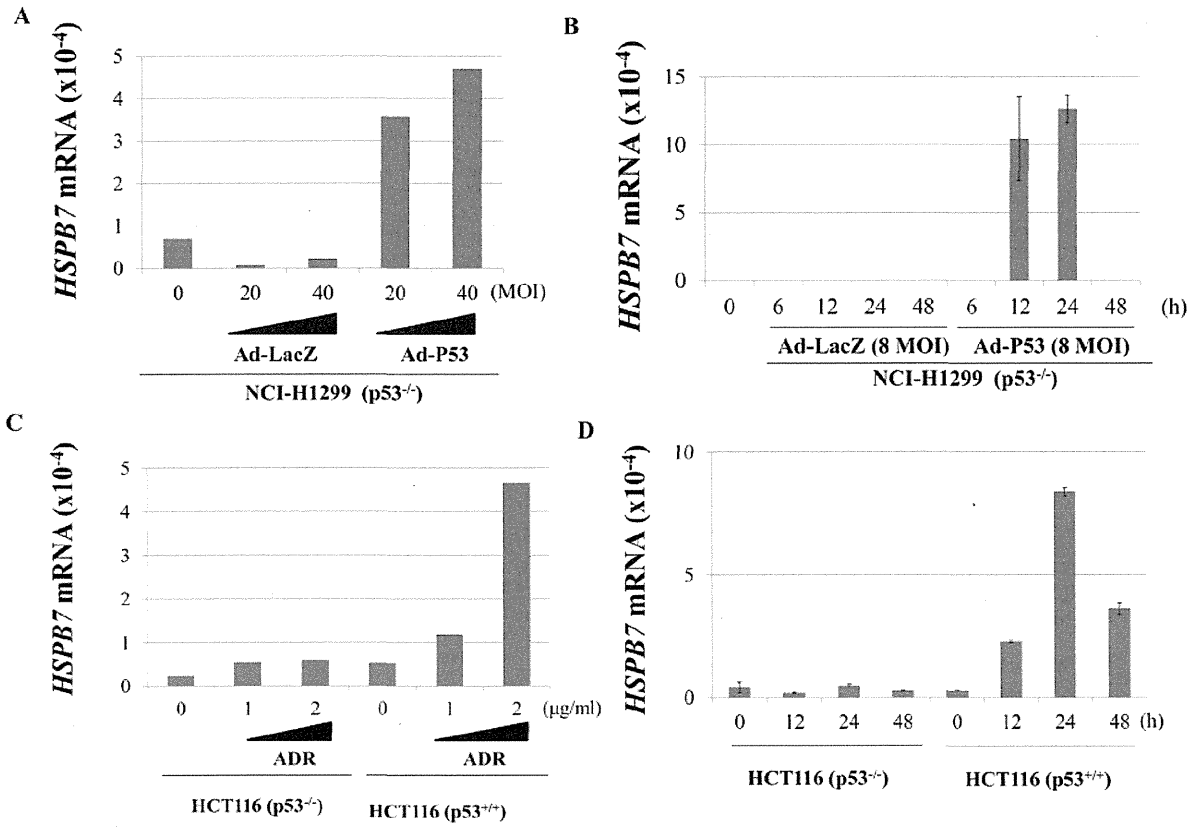

Figure 5. HSPB7 is regulated by p53. (A and B) HSPB7 expression in NCI-H1299 cells with or without p53 induction (A) dose- and (B) time-dependently. Cells were infected with replication-deficient recombinant adenovirus encoding p53 (Ad-p53) or LacZ (Ad-LacZ) at indicated doses, and the cells were collected 48 h later and qPCR analysis was performed (A). The cells were infected at 8 MOI and collected at different time points (B). (C and D) HSPB7 expression in HCT116 ($p53^{-/-}$) and HCT116 ($p53^{+/+}$) cells treated with adriamycin at indicated doses for 2 h and the cells were harvested at 48 h (C). The cells were treated with adriamycin at 2 μg/ml for 2 h and then harvested at different time points (D). B2M was used for normalization of expression levels. Values are expressed as the mean ± SD.

higher level of methylation in RCC cell lines than normal cell lines. Consistently, restoration of HSPB7 expression was observed by the treatment of cancer cells with 5-Aza-dC. In addition, since no somatic changes in coding regions of the HSPB7 gene were found in our sequence analysis of RCC cell lines or in the COSMIC database, HSPB7 in RCC is considered to be downregulated mostly by hypermethylation.

The second key finding in this study is that HSPB7 showed growth suppressive effect in cancer cells. Ectopic expression of HSPB7 significantly impaired colony-forming ability for 5 RCC cell lines, indicating that HSPB7 may function as a tumor suppressor gene. Similarly α B-crystallin, one of the small heat shock protein family members, was also indicated to function as a tumor suppressor in nasopharyngeal carcinoma cells (33). Furthermore, the region on chromosome 1p36.23-p34.3, where HSPB7 is located, showed frequent loss of heterozygosity in many types of solid tumors (34). However, further studies are needed to clarify the detailed tumor suppressor function of HSPB7 in RCC.

The third important finding in this study is that HSPB7 was likely to be involved in the p53 pathway. The expression of HSPB7 was significantly induced in p53-dependent manner that was clearly demonstrated by two experiments, i) that introduction of adeno-p53 in p53-negative cancer cells showed strong induction of HSPB7 and ii) that DNA-damage-dependent introduction of HSPB7 was observed in HCT116 cells with wild-type p53, but not in those lacking p53. Although we failed to identify the p53-binding site in or near the HSPB7 gene, these two pieces of evidence strongly imply a critical role of HSPB7 as the direct/indirect p53-signal transducer and its downregulation may be involved in the development of various types of cancer including RCC.

In conclusion, we carried out a genome-wide gene expression analysis and identified HSPB7 to be a candidate tumor suppressor gene in RCC. We confirmed downregulation of this gene caused by DNA hypermethylation, its growth suppressive effect in RCC cell lines and its p53-dependent expression, indicating the important roles of HSPB7 in renal carcinogenesis. Our finding could contribute to better understanding of the novel function of HSPB7 in cancer.

Acknowledgements

We thank Ryuji Hamamoto, Daechun Kang and Takashi Fujitomo for technological support of IHC. We also thank Jinichi Mori for making cDNA template after Adriamycin treatment. We thank Satomi Takahashi for providing general reagents, materials and sequencing.

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM: Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 127: 2893-2917, 2010.
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D: Global cancer statistics. CA Cancer J Clin 61: 69-90, 2011.
3. Naito S, Tomita Y, Rha SY, *et al*: Kidney Cancer Working Group report. Jpn J Clin Oncol 40 (Suppl 1): i51-i56, 2010.
4. Siegel R, Naishadham D and Jemal A: Cancer statistics, 2012. CA Cancer J Clin 62: 10-29, 2012.
5. Chow WH, Devesa SS, Warren JL and Fraumeni JF Jr: Rising incidence of renal cell cancer in the United States. JAMA 281: 1628-1631, 1999.
6. Hock LM, Lynch J and Balaji KC: Increasing incidence of all stages of kidney cancer in the last 2 decades in the United States: An analysis of Surveillance, Epidemiology and End Results program data. J Urol 167: 57-60, 2002.
7. Lindblad P: Epidemiology of renal cell carcinoma. Scand J Surg 93: 88-96, 2004.
8. Murai M and Oya M: Renal cell carcinoma: etiology, incidence and epidemiology. Curr Opin Urol 14: 229-233, 2004.
9. Motzer RJ, Agarwal N, Beard C, *et al*: NCCN clinical practice guidelines in oncology: kidney cancer. J Natl Compr Canc Netw 7: 618-630, 2009.
10. Ljungberg B, Cowan NC, Hanbury DC, *et al*: EAU guidelines on renal cell carcinoma: the 2010 update. Eur Urol 58: 398-406, 2010.
11. Escudier B, Eisen T, Porta C, *et al*: Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 23 (Suppl 7): vii65-vii71, 2012.
12. Lam JS, Leppert JT, Belldegrun AS and Figlin RA: Novel approaches in the therapy of metastatic renal cell carcinoma. World J Urol 23: 202-212, 2005.
13. Cairns P: Renal cell carcinoma. Cancer Biomark 9: 461-473, 2010.
14. Molina AM and Motzer RJ: Current algorithms and prognostic factors in the treatment of metastatic renal cell carcinoma. Clin Genitourin Cancer 6 (Suppl 1): S7-S13, 2008.
15. Ono K, Tanaka T, Tsunoda T, *et al*: Identification by cDNA microarray of genes involved in ovarian carcinogenesis. Cancer Res 60: 5007-5011, 2000.
16. Hirota E, Yan L, Tsunoda T, *et al*: Genome-wide gene expression profiles of clear cell renal cell carcinoma: Identification of molecular targets for treatment of renal cell carcinoma. Int J Oncol 29: 799-827, 2006.
17. Saito-Hisaminato A, Katagiri T, Kakiuchi S, Nakamura T, Tsunoda T and Nakamura Y: Genome-wide profiling of gene expression in 29 normal human tissues with a cDNA microarray. DNA Res 9: 35-45, 2002.
18. Oda K, Arakawa H, Tanaka T, *et al*: p53AIP1, a potential mediator of p53-dependent apoptosis, and its regulation by Ser-46-phosphorylated p53. Cell 102: 849-862, 2000.
19. Matsuda K, Yoshida K, Taya Y, Nakamura K, Nakamura Y and Arakawa H: p53AIP1 regulates the mitochondrial apoptotic pathway. Cancer Res 62: 2883-2889, 2002.
20. Tanikawa C, Ueda K, Nakagawa H, Yoshida N, Nakamura Y and Matsuda K: Regulation of Protein Citrullination through p53/PAD14 Network in DNA Damage Response. Cancer Res 69: 8761-8769, 2009.
21. Tanikawa C, Matsuda K, Fukuda S, Nakamura Y and Arakawa H: p53RDL1 regulates p53-dependent apoptosis. Nat Cell Biol 5: 216-223, 2003.
22. Watanabe G, Kato S, Nakata H, Ishida T, Ohuchi N and Ishioka C: alphaB-crystallin: a novel p53-target gene required for p53-dependent apoptosis. Cancer Sci 100: 2368-2375, 2009.
23. Evans JR, Bosman JD, Brown-Endres L, Yehiely F and Cryns VL: Induction of the small heat shock protein alphaB-crystallin by genotoxic stress is mediated by p53 and p73. Breast Cancer Res Treat 122: 159-168, 2010.
24. Krief S, Faivre JF, Robert P, *et al*: Identification and characterization of cvHsp. A novel human small stress protein selectively expressed in cardiovascular and insulin-sensitive tissues. J Biol Chem 274: 36592-36600, 1999.
25. Sun X, Fontaine JM, Rest JS, Shelden EA, Welsh MJ and Benndorf R: Interaction of human HSP22 (HSPB8) with other small heat shock proteins. J Biol Chem 279: 2394-2402, 2004.
26. Vos MJ, Zijlstra MP, Kanon B, *et al*: HSPB7 is the most potent polyQ aggregation suppressor within the HSPB family of molecular chaperones. Hum Mol Genet 19: 4677-4693, 2010.
27. Cappola TP, Li M, He J, *et al*: Common variants in HSPB7 and FRMD4B associated with advanced heart failure. Circ Cardiovasc Genet 3: 147-154, 2010.
28. Matkovich SJ, Van Booven DJ, Hindes A, *et al*: Cardiac signaling genes exhibit unexpected sequence diversity in sporadic cardiomyopathy, revealing HSPB7 polymorphisms associated with disease. J Clin Invest 120: 280-289, 2010.
29. Stark K, Esslinger UB, Reinhard W, *et al*: Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. PLoS Genet 6: e1001167, 2010.
30. Villard E, Perret C, Gary F, *et al*: A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. Eur Heart J 32: 1065-1076, 2011.
31. Li XP, Luo R, Hua W and Hua W: Polymorphisms of Hspb7 gene associate with idiopathic dilated cardiomyopathy susceptibility in a Chinese population. Heart 98: E47-E47, 2012.
32. Rosenfeld GE, Mercer EJ, Mason CE and Evans T: Small heat shock proteins Hspb7 and Hspb12 regulate early steps of cardiac morphogenesis. Dev Biol 381: 389-400, 2013.
33. Huang Z, Cheng Y, Chiu PM, *et al*: Tumor suppressor Alpha B-crystallin (CRYAB) associates with the cadherin/catenin adherens junction and impairs NPC progression-associated properties. Oncogene 31: 3709-3720, 2012.
34. Ragnarsson G, Eiriksdottir G, Johannsdottir JT, Jonasson JG, Egilsson V and Ingvarsson S: Loss of heterozygosity at chromosome 1p in different solid human tumours: association with survival. Br J Cancer 79: 1468-1474, 1999.

Journal of Human Genetics (2013) 58, 353–361

www.nature.com/jhg

ORIGINAL ARTICLE

Impact of polymorphisms in drug pathway genes on disease-free survival in adults with acute myeloid leukemia

Sook Wah Yee[1], Joel A Mefford[2], Natasha Singh[3], Mary-Elizabeth Percival[3], Adrian Stecula[1], Kuo Yang[1], John S Witte[2], Atsushi Takahashi[4], Michiaki Kubo[4], Koichi Matsuda[5], Kathleen M Giacomini[1,6,7] and Charalambos Andreadis[3,6,7]

Acute myeloid leukemia (AML) is a clinically heterogeneous disease, with a 5-year disease-free survival (DFS) ranging from under 10% to over 70% for distinct groups of patients. At our institution, cytarabine, etoposide and busulfan are used in first or second remission patients treated with a two-step approach to autologous stem cell transplantation (ASCT). In this study, we tested the hypothesis that polymorphisms in the pharmacokinetic and pharmacodynamic pathway genes of these drugs are associated with DFS in AML patients. A total of 1659 variants in 42 genes were analyzed for their association with DFS using a Cox-proportional hazards model. One hundred and fifty-four genetically European patients were used for the primary analysis. An intronic single nucleotide polymorphism (SNP) in *ABCC3* (rs4148405) was associated with a significantly shorter DFS (hazard ratios (HR) = 3.2, $P = 5.6 \times 10^{-6}$) in our primary cohort. In addition, a SNP in the *GSTM1-GSTM5* locus, rs3754446, was significantly associated with a shorter DFS in all patients (HR = 1.8, $P = 0.001$ for 154 European ancestry; HR = 1.7, $P = 0.028$ for 125 non-European patients). Thus, for the first time, genetic variants in drug pathway genes are shown to be associated with DFS in AML patients treated with chemotherapy-based autologous ASCT.
Journal of Human Genetics (2013) **58,** 353–361; doi:10.1038/jhg.2013.38; published online 16 May 2013

Keywords: acute myeloid leukemia; autologous stem cell transplant; busulfan; cytarabine; etoposide; pharmacogenomics

INTRODUCTION

Adult acute myeloid leukemia (AML) is a hematologic malignancy with widely heterogeneous clinical outcomes. New treatments for AML are increasingly being tested in clinical trials of patients with specific tumor cell mutations.[1,2] Although there have been substantial improvements in the number of patients who achieve complete remission, the choice of induction and post-remission therapy for adult AML is still based on the 'one size fits all' principle. Most regimens incorporate antimetabolites (for example, cytarabine, fludarabine), topoisomerase II inhibitors (for example, etoposide, daunorubicin, idarubicin, mitoxantrone) and alkylating agents (for example, busulfan, cyclophosphamide) for the treatment of AML. Prognostic factors for treatment response include age, prior exposure to chemotherapy, cytogenetic markers and expression profiles, and appearance of specific genetic mutations in tumor tissue, such as mutation and translocation of particular genes (for example, *FLT3*, *NPM1*).[3,4] However, these prognostic factors do not adequately capture the wide diversity of clinical outcomes in this disease. The percent of adults with AML who can survive $\geqslant 3$ years and may be cured is ~ 5–70%.[5,6]

One possible explanation for the difference in response to AML treatment is germline genetic variation. Although the pharmacogenomics of AML drug response is an active area of research, there remain large challenges. These include the: (i) poor availability of uniform, well-collected and well-defined drug response phenotype information; (ii) lack of widely available germline DNA not contaminated with tumor cells (myeloblasts); (iii) limited availability of panels of genotype data in large patient cohorts; and (iv) inability to validate findings in replication studies. In this study, we overcome many of these challenges. In particular, to our knowledge, this is the first study that involved the analysis of large numbers of genetic polymorphisms in a cohort of AML patients treated with high-dose chemotherapy followed by autologous stem cell transplantation (ASCT). We present results from testing the association between

[1]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA, USA; [2]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA; [3]Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA, USA; [4]Center for Genomic Medicine, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan; [5]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan and [6]Institute for Human Genetics, University of California, San Francisco, CA, USA
[7]Co-senior authors.
Correspondence: Dr C Andreadis, Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA 94143, USA.
E-mail: candreadis@medicine.ucsf.edu
Received 18 March 2013; accepted 11 April 2013; published online 16 May 2013

germline variants in drug pathway and other genes with disease-free survival (DFS) in adult AML patients. We identified new associations between AML DFS and polymorphisms in several drug pathway genes for cytarabine, etoposide and busulfan, and also replicated single nucleotide polymorphisms (SNPs) previously reported to be associated with the AML response phenotypes.

MATERIALS AND METHODS

Clinical protocol, study criteria and patient cohorts

The population for the current study consists of AML patients who were enrolled in University of California, San Francisco (UCSF) study protocols 9203 or 9303 between 1988 and 2010. The study protocols and patient selection criteria have been previously described.[7–9] This treatment protocol was used in patients with low- and standard-risk *de novo* AML, including acute promyelocytic leukemia in first or second complete remission. It was also used in a small number of patients with high risk AML (that is, with secondary AML) if allogeneic stem cell transplantation was not an option for the patient (for example, unavailable donor).[7] In step 1, patients were treated with consolidation chemotherapy including cytarabine 2000 mg m^{-2} (intravenously) twice daily for 4 days concurrently, with etoposide 40 mg kg^{-1} by intravenous infusion over the 4 days. During the recovery period from chemotherapy, peripheral blood stem cells were collected under granulocyte colony-stimulating factor stimulation. In step 2, patients underwent ASCT, which involved the preparative regimen of busulfan (total dose 16 mg kg^{-1} orally or 12.8 mg kg^{-1} intravenously, over 16 doses in 4 days), followed by etoposide 60 mg kg^{-1} (intravenous bolus) and reinfusion of blood or marrow stem cells. Patients had to be in complete remission for at least 30 days before step 2 (Figure 1). Complete remission was defined as normal bone marrow morphology with <5% blasts, resolution of previously abnormal cytogenetics and no evidence of extramedullary leukemia. In addition, patients must meet criteria for neutrophil and platelet counts, liver and kidney function.[7–9] Detailed procedures of patient enrollment, diagnosis, data collection and follow-up have been previously described.[7–9] Briefly, patients were actively followed-up in the beginning within 6 months of diagnosis, with subsequent annual follow-up by clinic visits. UCSF electronic medical records, the UCSF Blood and Bone Marrow Transplant Clinic database and patients' medical charts were abstracted to determine patients' remission status. The UCSF Committee on Human Research approved the research protocol (institutional review board number 10-00649).

DNA isolation and genotyping

DNA was isolated from peripheral blood stem cells, which were collected during the recovery from step 1 consolidation chemotherapy. As noted in the above section, patients were in complete remission before consolidation chemotherapy, and hence the samples utilized in this step contained <5% leukemic cells. DNA was isolated at the UCSF DNA Banking and Extraction Services Lab. The lab followed standard DNA extraction protocol described in the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA). The DNA was then quantified using Picogreen (Life Technologies, Grand Island, NY, USA) and normalized to 50 ng μl^{-1}. For each sample, we genotyped 250 ng of DNA. The Illumina HumanOmniExpress v1.0 Beadchip was used, following the manufacturer's protocols, at the Center for Genomic Medicine, RIKEN, Yokohama, Japan. For quality control of the genotyping, we included one HapMap trio and three duplicates of the DNA samples from the AML patients. A total of 328 distinct DNA samples from patients were genotyped, along with three duplicates and three HapMap samples (trio).

Patients' ancestral origin

The genetic ancestral origin of patients was determined using principal component analysis implemented in Eigenstrat.[10] Genotype information on our 328 AML patients was analyzed in conjunction with SNP data from the HapMap project, which consists of Europeans (CEU and TSI), Asians (JPT and CHB), Africans (TSI and ASW) and Mexican (MXL). From these analyses we were able to distinguish 154 patients of European ancestry on the basis of their close clustering with the European HapMap samples.

Pathway, gene and SNP selection

A total of 42 genes were selected for analysis based on the following criteria: (i) genes in the pharmacokinetic and/or pharmacodynamics pathway of the drugs administered (cytarabine, etoposide and busulfan);[11–15] (ii) genes described in literature as having significant associations with drug cytotoxicity in lymphoblastoid cell lines (LCL);[16,17] (iii) genes involved in DNA mismatch repair;[18] and (iv) genes found previously to be associated with the AML response phenotype[19] (see Supplementary Table 1, Figure 1). After filtering the SNPs with low call rates (<90%) and SNPs with low minor allele frequencies (MAF; MAF<1%) in the 154 European ancestry patients, we selected the SNPs in the candidate genes and within 25 000 bp upstream and downstream flanking regions.

Statistical analysis of the associations

The primary analysis was to estimate the association between SNPs in the selected candidate genes and DFS in 154 patients of European ancestry. A Cox-proportional hazard model was used to estimate the hazard ratios (HR) and 95% confidence limits for the effect of genotype on DFS. An additive coding of genotypes was used in all analyses. The genetic effect estimates were adjusted for levels of a clinical risk score (see Table 1). The SNP associations with $P<0.01$ were also tested in the 125 non-European patients using a Cox-proportional hazard model. In light of the heterogeneous ethnicity of the samples, the genetic effect estimates were adjusted for the first 10 principal components calculated from the genome-wide association study data and the clinical risk score. The Cox-proportional hazards function from the R-project (version 2.15.1) was used. We used 3×10^{-5} (0.05/1659) as the significance level after Bonferroni correction for multiple testing.

Fine mapping of associations *via* imputation

In order to further clarify the association signals, we performed imputation on genes with $P<0.005$ in the primary analysis. For eight genes (*ABCC3*, *DCK*, *GSTM1*, *GSTT1*, *MSH3*, *RRM1*, *SLC22A12*, and *SLC28A3*), the genotypes at polymorphic sites known from the 1000 Genomes Project but not observed on the Illumina HumanOmniExpress v1.0 Beadchip were imputed using IMPUTE version 2 (version 2.3.0 for Mac OS X, http://mathgen.stats.ox.ac.uk/impute/impute_v2.html). The reference panel for the imputation was the 1000 Genomes Phase I-integrated variant set referenced to NCBI build b37 (March 2012 release, retrieved January 20 2013 from http://mathgen.stats.ox.ac.uk/impute/impute_v2.html#reference). Imputed variants with imputation quality scores<0.3 or MAF<0.01 were excluded. The remaining imputed SNPs were each used in Cox-proportional hazards models to predict DFS, as were the genotyped SNPs.

Functional studies

The potential functional effects of the SNPs with $P<0.01$ associated with DFS in European population were examined using the following steps:

(1) All tag-SNPs in linkage disequilibrium to the SNPs with $P<0.01$ in our primary analysis were identified using the Proxy Search in the Broad Institute SNAP (SNP Annotation and Proxy Search; version 2.2), http://www.broadinstitute.org/mpg/snap/ldsearch.php. The search options used in this step were: SNP data set = 1000 Genomes Pilot 1 in CEU population panel, r^2 threshold = 0.8 and distance limit = 500 kbp.

(2) Potential regulatory functions were identified by searching the following databases: Regulome database[20] and eQTL browser (http://eqtl.uchicago.edu, http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/).

(3) Literature searches were used to identify previous reports of associations with drug response phenotypes of any of the SNPs associated with AML response in this study.

RESULTS

This study investigates the potential associations between variants in 42 candidate genes and DFS in adult AML patients treated with a two-step approach to ASCT. High-dose cytarabine, etoposide and busulfan were used in this treatment approach. Table 1 describes

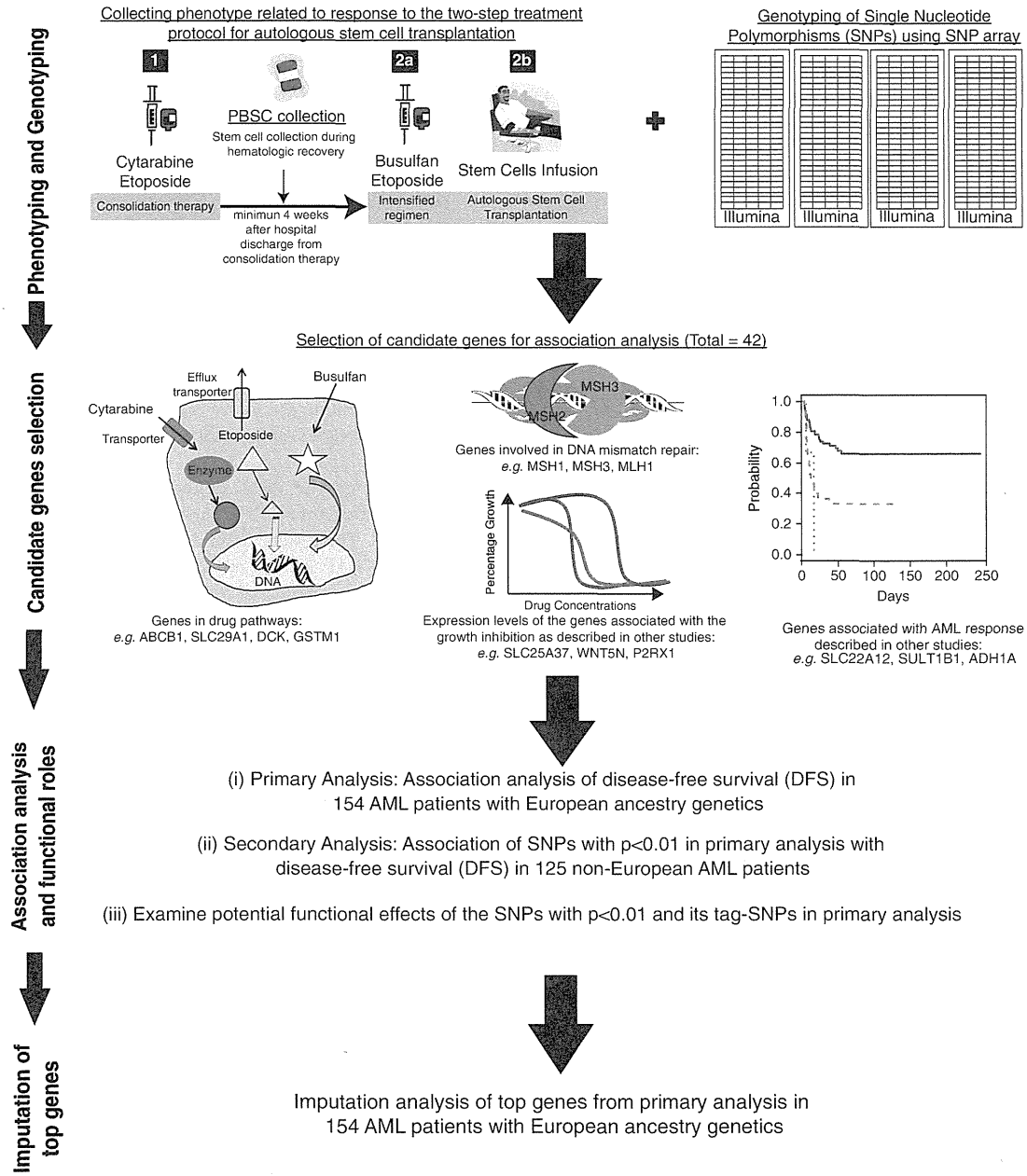

Figure 1 Schematic of workflow applied to determine the association of genetic variations in 42 candidate genes with disease-free survival (DFS) in acute myeloid leukemia (AML) patients treated with a two-step treatment protocol before autologous stem cell transplantation (ASCT). The workflow begins with phenotype data collection and genotyping of DNA samples using a genome-wide SNP array. Candidate genes were selected on the basis of their roles in the: drugs' pharmacokinetic/pharmacodynamics pathway; DNA mismatch repair mechanism; association with the drug cytotoxicity in lymphoblastoic cell lines previously identified in literatures; and pharmacogenomics studies of AML drug response. After association of each SNPs with DFS in the 154 AML patients of European ancestry, the SNPs with $P<0.01$ and their tag-SNPs were examined for their potential functional roles using databases to identify eQTL SNPs and predicted regulatory elements, such as binding sites of transcription factors and for their replications of previous studies. The SNPs with $P<0.01$ were also examined for their associations with DFS in 125 non-European ancestry. Finally, imputation was performed to identify other SNPs with stronger associations with DFS in European ancestry.

the demographic characteristics of this cohort, which consisted of 154 patients of European ancestry and 125 patients of non-European ancestry (African, Asian and Mexican). In the DFS analysis, 55 patients of European ancestry (35.7%) and 47 of non-European ancestry (37.6%) relapsed during the observation period, which extended from 1988 to 2010.

We applied quality control criteria to the SNP data, with a genotype call rate of 0.99 and minor allele frequency $\geqslant 0.01$. After filtering, a total of 42 genes covered by 1659 SNPs were included in this association analysis. Furthermore, none of the DNA samples showed chromosomal abnormality by GenomeStudio (Illumina). Results in Figure 2 and Table 2 showed that among the 40 SNPs

with $P<0.01$, the SNPs in the pharmacokinetic/pharmacodynamic pathway genes have stronger associations compared with SNPs not in the drug pathway. After Bonferroni correction for multiple testing (1659 tests), a SNP in the first intron of *ABCC3* (rs4148405) was significantly associated with DFS, with the minor allele (G) associated with shorter time to relapse (*P*-unadjusted $= 9.5 \times 10^{-6}$, Figure 3a). Although other SNPs did not reach significance after Bonferroni correction, overall there were 23 SNPs associated with DFS at $P<0.005$ (Table 2). These 23 SNPs are in or within 25 000 bp of eight genes: *SLC28A3*, *DCK*, *RRM1*, *GSTM1*, *ABCC3*, *MSH3*, *GSTT1*, or *SLC22A12*. The majority of the minor alleles were associated with poor outcome (shorter DFS). Kaplan–Meier estimate plots of DFS are shown in Figures 3a–d for four of the top SNPs with MAF $\geqslant 3\%$. Other SNPs in the genes of the cytarabine pathway including *NT5C2* and *RRM2B* were also associated with DFS but with weaker *P*-values ($P<0.01$, Table 2). Interestingly, SNPs in three out of eight selected genes (*SLC25A37*, *WNT5N* and *P2RX1*), for which expression levels have previously been correlated with either etoposide or cytarabine IC_{50} values in LCL, showed significant but weaker association ($P<0.01$) compared with genes in the drug pathways. Next, we examined the 40 SNPs in patients of non-European ancestry. Only one SNP, in *GSTM1-GSTM5* locus (rs3754446), was significantly associated with DFS in patients of non-European ancestry (HR = 1.7, $P=0.028$). Overall, in the entire cohort the minor allele of the SNP (rs3754446) was significantly associated with shorter DFS (HR = 1.7, $P=0.00027$).

Imputations of the eight candidate genes were performed to determine whether other SNPs in the regions have stronger association with DFS. The results showed that there are 234 imputed SNPs with $P<0.01$ (MAF $\geqslant 1\%$), and among these there are 93 SNPs with improved *P*-values compared with the genotyped SNPs (Figure 4). Several SNPs in deoxycytidine kinase (*DCK*) have significant *P*-values $< 1.0 \times 10^{-4}$ (Figure 4), with MAF 5–10%. Although none of these 93 SNPs are in exonic regions, on examination in the eQTL browser, Regulome database and GTEx eQTL browser, we determined that SNPs in GSTM1 (rs929166, rs11101989) and MSH3 (6151896) are associated with their respective gene expression levels in liver[21] or lymphoblastoid cells[22] (data not shown). In the *ABCC3* and *SLC28A3* regions, imputation analysis did not identify other more significant SNPs in addition to the most significant genotyped SNPs, rs4148405 and rs11140500, respectively (Figure 4).

Using *in silico* analysis, we determined whether the 40 SNPs and their tag-SNPs were in known or predicted functional regions of the genome. Several of the SNPs were in DNA regions predicted to have binding sites for transcription factors. Some of these regions appear in

Table 1 Demographic characteristics of patients undergoing autologous stem cells transplantation for AML from 1988 to 2010

Variables	*Patients of European ancestry (N = 154)*	*Patients of non-European ancestry (N = 125)*
Sex, N (%)		
Male	76 (49.3%)	67 (53.6%)
Female	78 (50.6%)	58 (46.4%)
Age (years)		
Median (s.d., range)		
Age at diagnosis	47.0[a] (13.0, 18–72)	40.5[b] (12.8, 17–68)
Age at autologous bone marrow transplantation	47.0 (13.0, 19–72)	41.0 (12.8, 19–69)
Year of transplantation		
1988–1995	36 (23.4%)	38 (30.4%)
1996–2000	41 (26.6%)	31 (24.8%)
2001–2010	77 (50.0%)	56 (44.8%)
De novo/secondary AML, N (%)		
De novo	147 (95.5%)	110 (88.0%)
Secondary	7 (4.5%)	15 (12.0%)
Risk, N (%)[c]		
Acute promyeloid leukemia	15 (9.7%)	12 (9.6%)
Low	18 (11.7%)	18 (14.4%)
Standard	114 (74.0%)	80 (64.0%)
High	7 (4.5%)	15 (12.0%)
DFS (months)		
Median (s.d.)	21.4 (43.9)	16.7 (44.8)

Abbreviations: AML, acute myeloid leukemia; DFS, disease-free survival.
[a]*N* = 152.
[b]*N* = 124.
[c]Defined by the Cancer and Leukemia Group B (CALGB) criteria.

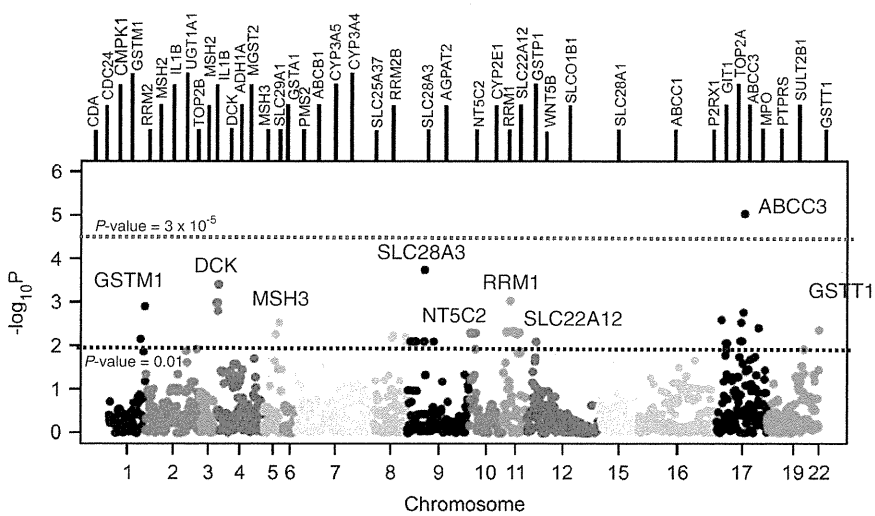

Figure 2 Plot showing the significance (−log10 of the *P*-value) of associations of the 1659 SNPs with DFS in 154 acute myeloid leukemia (AML) patients of European ancestries. Only SNPs with minor allele frequencies (MAF) of $\geqslant 1\%$ in the selected 42 candidate genes are shown. Each dot represents a SNP. SNPs above the black dotted line are SNPs with $P<0.01$, and the SNP above the red dotted line reached a *P*-value, which was significant after correction for multiple testing ($P<3 \times 10^{-5}$).

Table 2 Significant SNPs (P<0.01) associated with DFS in 154 AML patients of European ancestry

SNP	Chromosome	Gene	Feature	5′ flanking gene/ 3′ Flanking gene	Risk allele	MAF	P (unadjusted)	HR	(95% CI) of the HR	Genotype counts	Minor allele	Major allele	Reason for gene selection
rs4148405	17	ABCC3	Intron	CACNA1G/ANKRD40	G	0.12	9.45E-06	3.12	(1.88–5.15)	2/34/117	G	T	Drug pathway
rs11140500	9	SLC28A3	Intron	RMI1/NTRK2	T	0.01	0.00018	9.94	(2.99–33.09)	0/3/151	T	C	Drug pathway
rs10805074	4	DCK	Intron	MOBKL1A/ LOC100128311	A	0.03	0.00039	4.59	(1.98–10.67)	0/9/145	A	G	Drug pathway
rs7684954	4	DCK	Intron	LOC100128311/ LOC727995	A	0.03	0.00039	4.59	(1.98–10.67)	0/9/145	A	G	Drug pathway
rs4593998	11	RRM1	Intron	RRM1/LOC643244	A	0.14	0.00093	2.28	(1.40–3.71)	4/34/116	A	G	Drug pathway
rs6842838	4	DCK	3′UTR of MOB1B	MOB1B/DCK	G	0.05	0.0010	3.21	(1.60–6.45)	0/15/139	G	T	Drug pathway
rs1385985	4	DCK	3′UTR of MOB1B	GRSF1/DCK	C	0.05	0.0010	3.21	(1.60–6.45)	0/15/139	C	T	Drug pathway
rs3754446	1	GSTM1	Near-gene-5 [GSTM5]	GSTM1/GSTM5	G	0.38	0.0012	1.81	(1.26–2.59)	26/66/62	G	T	Drug pathway
rs7689093	4	DCK	Near-gene-5 [DCK]	MOBKL1A/DCK	G	0.01	0.0016	5.34	1.89–15.13)	0/4/150	G	A	Drug pathway
rs1989983	17	ABCC3	Near-gene-5 [ABCC3]	CACNA1G/ABCC3	A	0.11	0.0017	2.33	(1.37–3.96)	2/30/122	A	G	Drug pathway
rs2301835	17	ABCC3	Synonymous variant in coding of CACNA1G	CACNA1G/ABCC3	T	0.06	0.0029	2.60	(1.39–4.89)	1/16/137	T	C	Drug pathway
rs12515548	5	MSH3	Intron	LOC100128458/ RASGRF2	T	0.13	0.0029	2.07	(1.28–3.35)	6/27/121	T	C	DNA mismatch repair genes
rs2277624	17	ABCC3	Synonymous variant	CACNA1G/ANKRD40	A	0.25	0.0040	1.75	(1.20–2.55)	12/54/87	A	G	Drug pathway
rs11090305	22	GSTT1	Near-gene-5 [CABIN1]	GSTTP2/CABIN1	C	0.18	0.0044	1.98	(1.24–3.17)	3/50/101	C	T	Drug pathway
rs7130539	11	RRM1	Intron	STIM1/OR55B1P	C	0.06	0.0047	2.56	(1.33–4.91)	1/16/137	C	T	Drug pathway
rs11031136	11	RRM1	Intergenic	OR55B1P/LOC643244	G	0.06	0.0047	2.56	(1.33–4.91)	1/16/137	G	T	Drug pathway
rs528211	11	SLC22A12	Intergenic	SLC22A11/SLC22A12	G	0.29	0.0048	0.51	(0.31–0.81)	13/62/79	A	G	Other AML study
rs2360872	11	SLC22A12	Intergenic	SLC22A11/SLC22A12	C	0.29	0.0048	0.51	(0.31–0.81)	13/62/79	T	C	Other AML study
rs505802	11	SLC22A12	Near-gene-5 [SLC22A12]	SLC22A11/SLC22A12	A	0.29	0.0048	0.51	(0.31–0.81)	13/62/79	G	A	Other AML study
rs524023	11	SLC22A12	Near-gene-5 [SLC22A12]	SLC22A11/SLC22A12	G	0.29	0.0048	0.51	(0.31–0.81)	13/62/79	A	G	Other AML study
rs9734313	11	SLC22A12	5′UTR	SLC22A11/SLC22A12	T	0.29	0.0048	0.51	(0.31–0.81)	13/62/79	C	T	Other AML study
rs11231825	11	SLC22A12	Synonymous variant	SLC22A11/SLC22A12	C	0.29	0.0048	0.51	(0.31–0.81)	13/62/79	T	C	Other AML study
rs2268166	11	RRM1	Intron	STIM1/OR55B1P	G	0.06	0.0049	2.55	(0.20–0.75)	1/16/136	G	T	Drug pathway
rs11606370	11	SLC22A12	Intron of NRXN2	SLC22A12/RASGRP2	A	0.30	0.0050	0.51	(0.32–0.82)	14/64/76	C	A	Other AML study
rs11191547	10	NT5C2	Intergenic	CNNM2/NT5C2	T	0.31	0.0051	1.71	(1.17–2.49)	15/66/73	T	C	Drug pathway
rs11191549	10	NT5C2	Near-gene-3 [NT5C2]	CNNM2/NT5C2	T	0.31	0.0051	1.71	(1.17–2.49)	15/66/73	T	C	Drug pathway
rs11191553	10	NT5C2	Intron	CNNM2/ LOC100128863	T	0.31	0.0051	1.71	(1.17–2.49)	15/66/73	T	G	Drug pathway
rs10883836	10	NT5C2	Intron	LOC100128863/ LOC729081	C	0.31	0.0051	1.71	(1.17–2.49)	15/66/73	C	T	Drug pathway
rs7095304	10	NT5C2	Intergenic	NT5C2/LOC401648	A	0.31	0.0051	1.71	(1.17–2.49)	15/66/73	A	G	Drug pathway
rs6151816	5	MSH3	Intron	LOC100128458/ RASGRF2	T	0.12	0.0055	2.05	(1.24–3.41)	5/27/122	T	C	DNA mismatch repair genes
rs893006	11	SLC22A12	Intron	SLC22A11/NRXN2	T	0.28	0.0055	0.51	(0.32–0.82)	13/61/80	G	T	Other AML study
rs7818607	8	SLC25A37	Intergenic	SLC25A37/LOC646721	A	0.30	0.0057	1.75	(1.18–2.59)	16/59/79	A	C	Associated with drug cytotoxicy in LBL
rs2853229	8	RRM2B	Intron	NCALD/UBR5	A	0.49	0.0062	1.68	(1.16–2.44)	38/74/42	A	C	Drug pathway
rs8534	8	SLC25A37	Intergenic	SLC25A37/LOC646721	T	0.38	0.0067	1.68	(1.15–2.43)	24/69/61	T	C	Associated with drug cytotoxicy in LBL
rs8079740	17	ABCC3	Intron of CACNA1G	SPATA20ANCC3	G	0.31	0.0078	0.56	(0.37–0.86)	20/56/78	A	G	Drug pathway
rs757420	17	ABCC3	Intergenic	CACNA1G/ABCC3	T	0.30	0.0079	0.55	(0.36–0.86)	18/56/80	T	C	Drug pathway
rs2010851	12	WNT5B	Near-gene-3 [WNT5B]	WNT5B/ LOC100132548	A	0.30	0.0082	0.52	(0.32–0.84)	13/67/74	C	A	Associated with drug cytotoxicy in LBL
rs4995289	17	P2RX1	Intergenic	P2RX1/ATP2A3	T	0.28	0.0088	0.51	(0.31–0.85)	14/59/81	C	T	Associated with drug cytotoxicy in LBL
rs1516801	17	P2RX1	Intergenic	P2RX1/ATP2A3	G	0.28	0.0088	0.51	(0.31–0.85)	14/59/81	T	G	Associated with drug cytotoxicy in LBL
rs2607662	8	RRM2B	Intron of UBR5	NCALD/UBR5	T	0.46	0.0095	1.64	(1.13–2.37)	31/79/44	T	C	Drug pathway

Abbreviations: AML, acute myeloid leukemia; CI, confidence interval; DFS, disease-free survival; HR, hazards ratio; LBL, lymphoblastoid cell lines; MAF, minor allele frequencies; SNPs, single nucleotide polymorphisms.
The classifications near-gene-5 and near-gene-3 label SNPs that are outside transcribed regions, but within 2000 bp of a transcription region. Near-gene-5 includes upstream promoter region and untranslated 5′ mRNA.

the ENCODE Chip-Seq and DNase I peaks (Supplementary Table 2), suggesting that they could have regulatory functions. Interestingly, the GTEx (Genotype-Tissue Expression) eQTL browser (http://www.ncbi.nlm.nih.gov/gtex/GTEX2/gtex.cgi) and eQTL browser (http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/) showed that the SNP in *GSTM1*, rs3754446, is associated with *GSTM1* and GSTM5 expression levels in the liver,[21] brain and LCL[22–24] (Supplementary Table 2).

DISCUSSION

Previous pharmacogenomic studies of AML response to chemotherapy have been limited to a small number of candidate genes.[2,25–27] Although studies related to genes in drug pathways have been performed, many have not been replicated. In addition, some AML pharmacogenomics studies have been conducted using DNA from blast cells, which may have included somatic cell mutations in addition to germline polymorphisms.[28–30]

To determine whether germline genetic variations are associated with AML response, we designed our own pharmacogenomic study in AML patients treated with a chemotherapeutic regimen consisting of cytarabine, etoposide and busulfan, followed by ASCT. Our study was focused primarily on 154 AML patients of European ancestry. The analysis was centered on 42 genes related to the pharmacokinetic and pharmacodynamic pathways of the chemotherapy. A few other genes that had previously been associated with drug response in AML were also included (Supplementary Table 1, Figure 1 and Figure 2). The goals of this study were to identify new associations with DFS in AML patients and to determine whether SNPs previously reported to be associated with AML response could be replicated.

Overall, the most significant SNP in our analysis was the intronic variant in *ABCC3*, rs4148405 (HR = 3.1, $P = 9.5 \times 10(-6)$). ABCC3 is a multidrug resistance-associated protein, which is known to transport the etoposide metabolite, etoposide glucuronide.[31] Several lines of evidence support a role for *ABCC3* in DFS in AML patients. First, in a previous study a promoter variant in *ABCC3*, rs4793665 was associated with a shorter survival time in Israeli AML patients.[29] Although this variant was not associated with AML response in the current study, the data support a role of the transporter in response to chemotherapy in AML. Second, following etoposide administration to *Abcc2* −/−, *Abcc3* −/− mice, higher etoposide glucuronide levels were observed in the liver,[32] consistent with a potentially important role of *ABCC3* in etoposide pharmacokinetics. Third, cell lines transfected with *ABCC3* show greater resistance to etoposide.[33] Finally, higher *ABCC3* expression levels in leukemia cells are associated with poor outcome in children with leukemia.[34,35]

Data from Regulome database suggest that the *ABCC3* variant, rs4148405, is in a functional location in the genome, as this region demonstrates direct evidence of binding through ChIP-seq studies. Transcription factors that have a role in hepatic gene regulation (for example, *CEBPB*, *USF1*, *FOXA1*) have DNA response elements within this gene region (see http://regulome.stanford.edu/snp/chr17/48713567).[20,36] On the basis of the results of our primary analysis with DFS, we speculate that the minor allele of rs4148405 is associated with higher expression levels of *ABCC3* in the liver and/or leukemia cells, and thus reduced levels of etoposide in the tumor cells. Although the SNP rs4148405 is found at a MAF of >10% in non-European populations, this SNP was not significantly associated with DFS in the AML patients of non-European ancestry. It is possible that different linkage disequilibrium patterns between rs4148405 and potential causative SNPs may have confounded the analysis.

In addition to a SNP in *ABCC3*, we also identified 23 SNPs in seven other genes (*SLC28A3*, *DCK*, *RRM1*, *GSTM1*, *GSTT1*, *MSH3* and *SLC22A12*) that were associated with DFS in the European AML patients (with $P<0.005$, Table 2). Interestingly, expression levels of these genes or other SNPs in these genes have been previously associated with response to chemotherapy in AML or other cancers.[19,29,37–39] Expression levels or SNPs in these genes have also been associated with IC_{50} values of various chemotherapy agents in cell lines.[40,41] Genetic polymorphisms in glutathione-S-transferases,

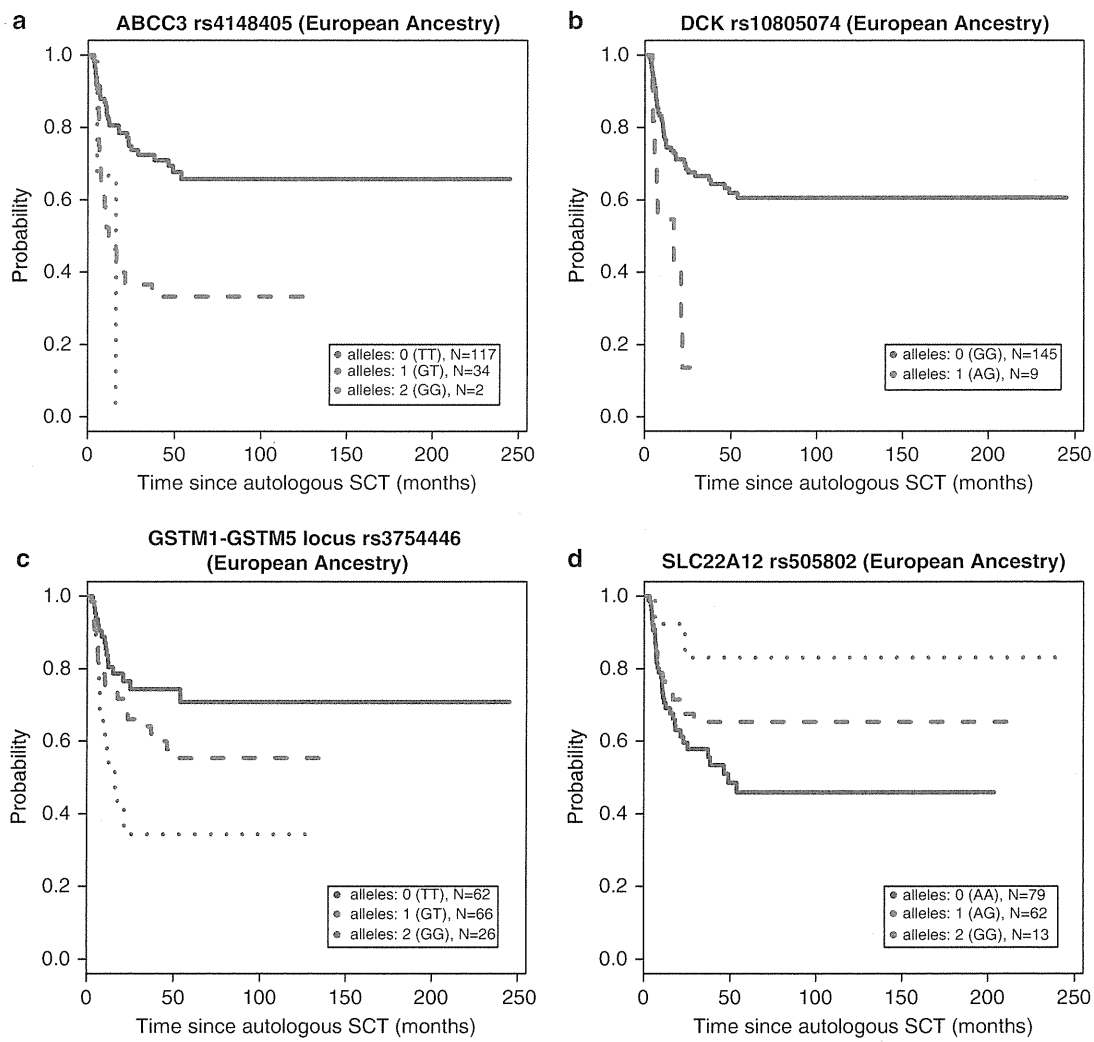

Figure 3 (**a**–**d**) Kaplan–Meier estimate of disease-free survival (DFS) stratified by the top SNPs (**a**) rs4148405 *ABCC3* (**b**) rs10805074 *DCK* (**c**) rs3754446 *GSTM1-GSTM5* locus and (**d**) rs505802 *SLC22A12* genotypes in patients of European ancestry.

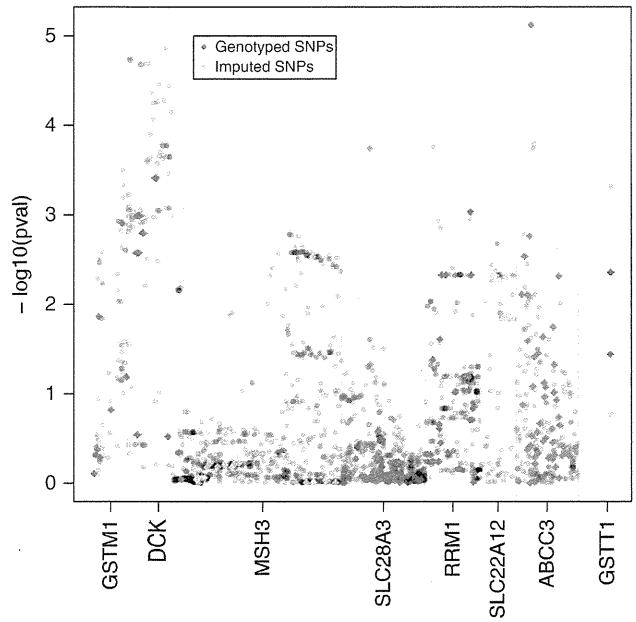

Figure 4 Plot showing the association of the genotyped and imputed SNPs with MAF $\geqslant 1\%$ in the selected top eight candidate genes (*ABCC3, DCK, SLC28A3, SLC22AA12, MSH3, RRM1, GSTM1, GSTT1*). Each colored diamond (not gray or black) represents a genotyped SNP and the gray/black dots represent imputed SNPs.

such as *GSTT1* and *GSTM1*, have been widely studied for their associations with drug toxicity, drug response and disease risk in leukemia patients.[29,39,42] Notably, *GSTM1* and *GSTT1* deletions have been implicated in various phenotypes associated with leukemia, including drug response,[43] busulfan pharmacokinetics[28] and disease risk.[44] Although we did not examine the effect of the *GSTM1* deletion in this study, we identified several SNPs (genotyped and imputed) in the *GSTM1-GSTM5* locus (rs3754446, rs929166 and rs11101989) associated with DFS in individuals of European ancestry that have not been previously reported (Table 2, Figure 2b). One of the SNPs in this locus, rs3754446, was also associated with AML response in the individuals of non-European ancestry (HR = 1.7, $P = 0.028$). Overall, in the entire cohort this minor allele SNP, rs3754446, was significantly associated with shorter DFS (HR = 1.7, $P = 0.00027$). Perusal of eQTL databases (Regulome database, GTEx browser and eQTL browser) suggests that the SNPs in the *GSTM1-GSTM5* locus are associated with expression levels of *GSTM1* and/or *GSTM5* in liver,[21] brain and LCL.[22–24] Thus, the SNPs in these genes, which are involved in drug metabolism, could affect AML response by affecting the pharmacokinetics of the drugs used in the treatment of AML. We examined 21 AML patients from our overall cohort, where we have their first-dose busulfan area under the curve. Interestingly, we observed a significant association between rs3754446 and reduced busulfan area under the curve in these 21 AML patients ($P = 0.03$, Supplementary Figure 1). The minor allele, G, in rs3754446, was associated with lower busulfan plasma levels (area under the curve), which was consistent with our observation that patients with the G allele had shorter DFS (Figure 3c). As higher busulfan plasma levels have been associated with busulfan liver toxicity,[14,45] future studies are needed to determine whether the SNPs in *GSTM1-GSTM5* are associated with liver toxicity. Collectively, these data suggest that *GSTM1* could have an important role in determining busulfan drug levels, drug response and/or drug toxicity.

Recently, using a drug-metabolizing enzyme/transporter gene SNP array, a synonymous variant in *SLC22A12* (rs11231825) was found to be associated with response in 94 AML patients treated with a combination drug regimen of gemtuzumab-ozogamicin with fludarabine-cytarabine-idarubicin.[19] Among the SNPs in the drug-metabolizing enzyme/transporter genes that were found to be significantly associated with AML response, we were able to replicate the SNP (rs11231825) in the uric acid transporter, *SLC22A12*. Notably, this synonymous variant, which is in linkage disequilibrium with a SNP upstream of *SLC22A12*, rs505802, has been found in various genome-wide association studies to be associated with uric acid levels.[46,47] In these genome-wide association studies,[46,47] the minor allele T has been associated with higher uric acid levels. Although speculative, our study and the previous study,[19] which demonstrated that patients with the T allele have a better response to chemotherapy, suggest that higher uric acid levels may be associated with longer DFS time. In our study, the T allele, which is associated with higher uric acid levels,[48,49] was associated with longer DFS time (see Figure 3d). Uric acid is a potent antioxidant, and it is possible that higher levels are beneficial for survival in AML patients.

Previous studies have shown that the nucleoside transporter, *SLC28A3* (CNT3) has a role in cytarabine cytotoxicity and resistance.[40,50,51] In this study, we observed several low-allele frequency variants (MAF 1%) in *SLC28A3* are associated with DFS in AML patients on cytarabine and other chemotherapy. Therefore, we hypothesized that CNT3 may transport cytarabine. Supplementary Figure 2a shows that radiolabeled cytarabine was taken up into CNT3 stably expressing cells, and the uptake (over empty vector cells) was significantly enhanced in cell lines exposed to the equilibrative nucleoside transporter inhibitor *S*-(4-nitrobenzyl)-6-thioinosine, which reduced background uptake of cytarabine in the cells. Notably, cytarabine uptake decreased significantly in CNT3 stable cells treated with the *SLC28A3* inhibitor (phloridzin) or with both inhibitors (Supplementary Figure 2a). Fludarabine, a known substrate of CNT3 was used as a positive control in this study (Supplementary Figure 2b).[52] As low-allele frequency variant in CNT3 was found to be significantly associated with DFS (Table 2), we interpret the results with caution. Although our finding that *SLC28A3* transports cytarabine supports the association, functional studies of variants in this region and/or a larger sample size are required to determine whether these uncommon variants are associated with cytarabine response.

DCK has an important role in activating cytarabine to its active metabolite, cytarabine triphosphate. Two tag-SNPs, rs4308342 and rs3775289, in *DCK* that were associated with DFS in our AML patients of European ancestry have been previously associated with the IC_{50} of another nucleoside analog, gemcitabine, in LCL.[40] Furthermore, a previous study demonstrated that the level of cytarabine triphosphate in AML blast cells correlates with the ratio in expression levels of the cytarabine-activating enzyme, *DCK* and the -inactivating enzyme, 5′-nucleotidase, cytosolic II (*NT5C2*).[48] In our study, in addition to the two tag-SNPs in *DCK* associated with DFS in AML patients, five SNPs in the *NT5C2* region were associated with DFS (Table 2). Although several of these SNPs are eQTLs (see Supplementary Table 2), further studies are required to determine whether these SNPs have important roles in determining the levels of cytarabine triphosphate in AML blast cells. Other genes, which have important roles in the cytarabine pharmacodynamics pathway are ribonucleotide reductase, *RRM1* and *RRM2*, which are considered

targets of nucleoside drugs such as cytarabine. The role of this enzyme is to regulate intracellular pools of ribonucleotides, such as deoxycytidine triphosphate, which is important in building blocks for DNA replication. Studies have shown that AML blasts cells with high levels of deoxycytidine triphosphate are resistant to cytarabine, and that there is a significant correlation between *RRM1* and *RRM2* gene expression levels and deoxycytidine triphosphate levels after cytarabine treatment in AML blast cells.[15] Consistent with previous studies showing that SNPs in *RRM1* are associated with response or toxicity to gemcitabine-based chemotherapy in lung and breast cancer patients,[49,53] our findings suggest that the SNPs in *RRM1* are associated with AML response to chemotherapy that include cytarabine.

Although our current association analysis supports the important roles of drug pathway genes, mainly transporters and enzymes, in AML response, we also selected eight genes that have been associated with cytarabine or etoposide IC_{50} values in LCL.[16,17] A few SNPs in the eight genes were significantly associated with response in AML patients, suggesting that genes identified in *in vitro* assays in LCLs may also be important predictors of *in vivo* drug response in AML patients.[16,17]

In summary, in this genetic association study of DFS in AML patients, we identified polymorphisms that have not been previously associated with AML response, including SNPs in *ABCC3*, *DCK*, *GSTM1*, *MSH3*, *NT5C2*, *RRM1* and *SLC28A3*. A SNP in *ABCC3*, rs4148405, which remained significant after multiple testing, was associated with DFS, suggesting an important role of *ABCC3* in determining etoposide levels in the liver and other tissues, and hence AML response. Many of the significant SNPs or their tag-SNPs were eQTLs or located in functional regions in the genome. Finally, we determined for the first time that *SLC28A3* (CNT3) transported cytarabine into cells, suggesting an important role of this transporter in cytarabine cytotoxicity.

ACKNOWLEDGEMENTS

This research was supported in part by a grant from NIH NIGMS (GM61390). This study was supported by the NIH Pharmacogenomics Research Network (PGRN)-RIKEN Center for Genomic Medicine (CGM) Strategic Alliance. We are grateful to the UCSF AML Tissue Bank staff (Ms Joy Cruz, Ms Christine Cheng). SWY thanks Chris Gignoux and Pär Mattson for their help in creating plots in the R-project.

1 Daver, N. & Cortes, J. Molecular targeted therapy in acute myeloid leukemia. *Hematology* **17** (Suppl 1), S59–S62 (2012).
2 Patel, J. P., Gonen, M., Figueroa, M. E., Fernandez, H., Sun, Z., Racevskis, J. *et al.* Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N. Engl. J. Med.* **366,** 1079–1089 (2012).
3 Buccisano, F., Maurillo, L., Del Principe, M. I., Del Poeta, G., Sconocchia, G., Lo-Coco, F. *et al.* Prognostic and therapeutic implications of minimal residual disease detection in acute myeloid leukemia. *Blood* **119,** 332–341 (2012).
4 Martelli, M. P., Sportoletti, P., Tiacci, E., Martelli, M. F. & Falini, B. Mutational landscape of AML with normal cytogenetics: Biological and clinical implications. *Blood Rev.* **27,** 13–22 (2013).
5 Kroger, N., Brand, R., van Biezen, A., Cahn, J. Y., Slavin, S., Blaise, D. *et al.* Autologous stem cell transplantation for therapy-related acute myeloid leukemia and myelodysplastic syndrome. *Bone Marrow Transplant.* **37,** 183–189 (2006).
6 Novitzky, N., Thomas, V., du Toit, C. & McDonald, A. Is there a role for autologous stem cell transplantation for patients with acute myelogenous leukemia? A retrospective analysis. *Biol. Blood Marrow Transplant.* **17,** 875–884 (2011).
7 Linker, C. A., Damon, L. E., Ries, C. A., Navarro, W. A., Case, D. & Wolf, J. L. Autologous stem cell transplantation for advanced acute myeloid leukemia. *Bone Marrow Transplant.* **29,** 297–301 (2002).
8 Linker, C. A., Owzar, K., Powell, B., Hurd, D., Damon, L. E., Archer, L. E. *et al.* Auto-SCT for AML in second remission: CALGB study 9620. *Bone Marrow Transplant.* **44,** 353–359 (2009).
9 Linker, C. A., Ries, C. A., Damon, L. E., Sayre, P., Navarro, W., Rugo, H. S. *et al.* Autologous stem cell transplantation for acute myeloid leukemia in first remission. *Biol. Blood Marrow Transplant.* **6,** 50–57 (2000).
10 Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A. & Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38,** 904–909 (2006).
11 Yang, J., Bogni, A., Schuetz, E. G., Ratain, M., Dolan, M. E., McLeod, H. *et al.* Etoposide pathway. *Pharmacogenet. Genomics* **19,** 552–553 (2009).
12 Li, L., Schaid, D. J., Fridley, B. L., Kalari, K. R., Jenkins, G. D., Abo, R. P. *et al.* Gemcitabine metabolic pathway genetic polymorphisms and response in patients with non-small cell lung cancer. *Pharmacogenet. Genomics* **22,** 105–116 (2012).
13 Hassan, M. & Andersson, B. S. Role of pharmacogenetics in busulfan/cyclophosphamide conditioning therapy prior to hematopoietic stem cell transplantation. *Pharmacogenomics* **14,** 75–87 (2013).
14 Abbasi, N., Vadnais, B., Knutson, J. A., Blough, D. K., Kelly, E. J., O'Donnell, P. V. *et al.* Pharmacogenetics of intravenous and oral busulfan in hematopoietic cell transplant recipients. *J. Clin. Pharmacol.* **51,** 1429–1438 (2011).
15 Lamba, J. K. Genetic factors influencing cytarabine therapy. *Pharmacogenomics* **10,** 1657–1674 (2009).
16 Huang, R. S., Duan, S., Bleibel, W. K., Kistner, E. O., Zhang, W., Clark, T. A. *et al.* A genome-wide approach to identify genetic variants that contribute to etoposide-induced cytotoxicity. *Proc. Natl Acad. Sci. USA* **104,** 9758–9763 (2007).
17 Hartford, C. M., Duan, S., Delaney, S. M., Mi, S., Kistner, E. O., Lamba, J. K. *et al.* Population-specific genetic variants important in susceptibility to cytarabine arabinoside cytotoxicity. *Blood* **113,** 2145–2153 (2009).
18 Hewish, M., Martin, S. A., Elliott, R., Cunningham, D., Lord, C. J. & Ashworth, A. Cytosine-based nucleoside analogs are selectively lethal to DNA mismatch repair-deficient tumour cells by enhancing levels of intracellular oxidative stress. *Br. J. Cancer* **108,** 983–992 (2013).
19 Iacobucci, I., Lonetti, A., Candoni, A., Sazzini, M., Papayannidis, C., Formica, S. *et al.* Profiling of drug-metabolizing enzymes/transporters in CD33+ acute myeloid leukemia patients treated with Gemtuzumab-Ozogamicin and Fludarabine, Cytarabine and Idarubicin. *Pharmacogenomics J.* (e-pub ahead of print 15 May 2012; doi:10.1038/tpj.2012.13).
20 Boyle, A. P., Hong, E. L., Hariharan, M., Cheng, Y., Schaub, M. A., Kasowski, M. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–1797 (2012).
21 Schadt, E. E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P. Y. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6,** e107 (2008).
22 Stranger, B. E., Nica, A. C., Forrest, M. S., Dimas, A., Bird, C. P., Beazley, C. *et al.* Population genomics of human gene expression. *Nat. Genet.* **39,** 1217–1224 (2007).
23 Veyrieras, J. B., Kudaravalli, S., Kim, S. Y., Dermitzakis, E. T., Gilad, Y., Stephens, M. *et al.* High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet.* **4,** e1000214 (2008).
24 Gibbs, J. R., van der Brug, M. P., Hernandez, D. G., Traynor, B. J., Nalls, M. A., Lai, S. L. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6,** e1000952 (2010).
25 Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286,** 531–537 (1999).
26 Gaidzik, V. I., Paschka, P., Spath, D., Habdank, M., Kohne, C. H., Germing, U. *et al.* TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group. *J. Clin. Oncol.* **30,** 1350–1357 (2012).
27 Paschka, P., Marcucci, G., Ruppert, A. S., Whitman, S. P., Mrozek, K., Maharry, K. *et al.* Wilms' tumor 1 gene mutations independently predict poor outcome in adults with cytogenetically normal acute myeloid leukemia: a cancer and leukemia group B study. *J. Clin. Oncol.* **26,** 4595–4602 (2008).
28 Kim, S. D., Lee, J. H., Hur, E. H., Lee, J. H., Kim, D. Y., Lim, S. N. *et al.* Influence of GST gene polymorphisms on the clearance of intravenous busulfan in adult patients undergoing hematopoietic cell transplantation. *Biol. Blood Marrow Transplant.* **17,** 1222–1230 (2011).
29 Muller, P., Asher, N., Heled, M., Cohen, S. B., Risch, A. & Rund, D. Polymorphisms in transporter and phase II metabolism genes as potential modifiers of the predisposition to and treatment outcome of de novo acute myeloid leukemia in Israeli ethnic groups. *Leuk. Res.* **32,** 919–929 (2008).
30 Emadi, A. & Karp, J. E. The clinically relevant pharmacogenomic changes in acute myelogenous leukemia. *Pharmacogenomics* **13,** 1257–1269 (2012).
31 Zelcer, N., Saeki, T., Reid, G., Beijnen, J. H. & Borst, P. Characterization of drug transport by the human multidrug resistance protein 3 (ABCC3). *J. Biol. Chem.* **276,** 46400–46407 (2001).
32 Lagas, J. S., Fan, L., Wagenaar, E., Vlaming, M. L., van Tellingen, O., Beijnen, J. H. *et al.* P-glycoprotein (P-gp/Abcb1), Abcc2, and Abcc3 determine the pharmacokinetics of etoposide. *Clin. Cancer Res.* **16,** 130–140 (2010).
33 Belinsky, M. G., Dawson, P. A., Shchaveleva, I., Bain, L. J., Wang, R., Ling, V. *et al.* Analysis of the in vivo functions of Mrp3. *Mol. Pharmacol.* **68,** 160–168 (2005).
34 Steinbach, D., Wittig, S., Cario, G., Viehmann, S., Mueller, A., Gruhn, B. *et al.* The multidrug resistance-associated protein 3 (MRP3) is associated with a poor outcome in childhood ALL and may account for the worse prognosis in male patients and T-cell immunophenotype. *Blood* **102,** 4493–4498 (2003).
35 Steinbach, D., Lengemann, J., Voigt, A., Hermann, J., Zintl, F. & Sauerbrey, A. Response to chemotherapy and expression of the genes encoding the multidrug

resistance-associated proteins MRP2, MRP3, MRP4, MRP5, and SMRP in childhood acute myeloid leukemia. *Clin. Cancer Res.* **9,** 1083–1086 (2003).
36 Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).
37 Kim, K. I., Huh, I. S., Kim, I. W., Park, T., Ahn, K. S., Yoon, S. S. *et al.* Combined interaction of multi-locus genetic polymorphisms in cytarabine arabinoside metabolic pathway on clinical outcomes in adult acute myeloid leukaemia (AML) patients. *Eur. J. Cancer* **49,** 403–410 (2013).
38 Shi, J. Y., Shi, Z. Z., Zhang, S. J., Zhu, Y. M., Gu, B. W., Li, G. *et al.* Association between single nucleotide polymorphisms in deoxycytidine kinase and treatment response among acute myeloid leukaemia patients. *Pharmacogenetics* **14,** 759–768 (2004).
39 Xiao, Z., Yang, L., Xu, Z., Zhang, Y., Liu, L., Nie, L. *et al.* Glutathione S-transferases (GSTT1 and GSTM1) genes polymorphisms and the treatment response and prognosis in Chinese patients with de novo acute myeloid leukemia. *Leuk. Res.* **32,** 1288–1291 (2008).
40 Li, L., Fridley, B. L., Kalari, K., Jenkins, G., Batzler, A., Weinshilboum, R. M. *et al.* Gemcitabine and arabinosylcytosin pharmacogenomics: genome-wide association and drug response biomarkers. *PloS one* **4,** e7765 (2009).
41 Mitra, A. K., Crews, K. R., Pounds, S., Cao, X., Feldberg, T., Ghodke, Y. *et al.* Genetic variants in cytosolic 5'-nucleotidase II are associated with its expression and cytarabine sensitivity in HapMap cell lines and in patients with acute myeloid leukemia. *J. Pharmacol. Exp. Ther.* **339,** 9–23 (2011).
42 Mossallam, G. I., Abdel Hamid, T. M. & Samra, M. A. Glutathione S-transferase GSTM1 and GSTT1 polymorphisms in adult acute myeloid leukemia; its impact on toxicity and response to chemotherapy. *J. Egypt. Natl Canc. Inst.* **18,** 264–273 (2006).
43 Voso, M. T., Hohaus, S., Guidi, F., Fabiani, E., D'Alo, F., Groner, S. *et al.* Prognostic role of glutathione S-transferase polymorphisms in acute myeloid leukemia. *Leukemia* **22,** 1685–1691 (2008).
44 Das, P., Shaik, A. P. & Bammidi, V. K. Meta-analysis study of glutathione-S-transferases (GSTM1, GSTP1, and GSTT1) gene polymorphisms and risk of acute myeloid leukemia. *Leuk. Lymphoma* **50,** 1345–1351 (2009).
45 Zhang, H., Graiser, M., Hutcherson, D. A., Dada, M. O., McMillan, S., Ali, Z. *et al.* Pharmacokinetic-directed high-dose busulfan combined with cyclophosphamide and etoposide results in predictable drug levels and durable long-term survival in lymphoma patients undergoing autologous stem cell transplantation. *Biol. Blood. Marrow Transplant.* **18,** 1287–1294 (2012).
46 Takeuchi, F., Yamamoto, K., Isono, M., Katsuya, T., Akiyama, K., Ohnaka, K. *et al.* Genetic Impact on Uric Acid Concentration and Hyperuricemia in the Japanese Population. *J. Atheroscler. Thromb.* **20,** 351–367 (2012).
47 Kolz, M., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M. *et al.* Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet.* **5,** e1000504 (2009).
48 Yamauchi, T., Negoro, E., Kishi, S., Takagi, K., Yoshida, A., Urasaki, Y. *et al.* Intracellular cytarabine triphosphate production correlates to deoxycytidine kinase/cytosolic 5'-nucleotidase II expression ratio in primary acute myeloid leukemia cells. *Biochem. Pharmacol.* **77,** 1780–1786 (2009).
49 Kim, S. O., Jeong, J. Y., Kim, M. R., Cho, H. J., Ju, J. Y., Kwon, Y. S. *et al.* Efficacy of gemcitabine in patients with non-small cell lung cancer according to promoter polymorphisms of the ribonucleotide reductase M1 gene. *Clin. Cancer Res.* **14,** 3083–3088 (2008).
50 Errasti-Murugarren, E., Pastor-Anglada, M. & Casado, F. J. Role of CNT3 in the transepithelial flux of nucleosides and nucleoside-derived drugs. *J. Physiol.* **582,** 1249–1260 (2007).
51 Sarkar, M., Han, T., Damaraju, V., Carpenter, P., Cass, C. E. & Agarwal, R. P. Cytosine arabinoside affects multiple cellular factors and induces drug resistance in human lymphoid cells. *Biochem. Pharmacol.* **70,** 426–432 (2005).
52 Badagnani, I., Chan, W., Castro, R. A., Brett, C. M., Huang, C. C., Stryke, D. *et al.* Functional analysis of genetic variants in the human concentrative nucleoside transporter 3 (CNT3; SLC28A3). *Pharmacogenomics J.* **5,** 157–165 (2005).
53 Rha, S. Y., Jeung, H. C., Choi, Y. H., Yang, W. I., Yoo, J. H., Kim, B. S. *et al.* An association between RRM1 haplotype and gemcitabine-induced neutropenia in breast cancer patients. *Oncologist* **12,** 622–630 (2007).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

Journal of Human Genetics (2013) 58, 588–593

www.nature.com/jhg

ORIGINAL ARTICLE

A replication study for three nephrolithiasis loci at 5q35.3, 7p14.3 and 13q14.1 in the Japanese population

Takahiro Yasui[1], Atsushi Okada[1], Yuji Urabe[2], Masayuki Usami[1], Kentaro Mizuno[1], Yasue Kubota[1], Keiichi Tozawa[1], Shoichi Sasaki[1], Yoshihito Higashi[3], Yoshikazu Sato[4], Michiaki Kubo[5], Yusuke Nakamura[2], Koichi Matsuda[2] and Kenjiro Kohri[1]

A previous genome-wide association study (GWAS) reported three novel nephrolithiasis-susceptibility loci at 5q35.3, 7p14.3 and 13q14.1. Here, we investigated the association of these loci with nephrolithiasis by using an independent Japanese sample set. We performed case–control association analysis using 601 patients with nephrolithiasis and 201 control subjects. We selected seven single-nucleotide polymorphisms (SNPs): rs12654812 and rs11746443 from 5q35.3 (*RGS14-SLC34A1-PFN3-F12*); rs12669187 and rs1000597 from 7p14.3 (*INMT-FAM188B-AQP1*); and rs7981733, rs1170155, and rs4142110 from 13q14.1 (*DGKH* (diacylglycerol kinase)), which were previously reported to be significantly associated with nephrolithiasis. rs12654812, rs12669187 and rs7981733 were significantly associated with nephrolithiasis after Bonferroni's correction ($P = 3.12 \times 10^{-3}$, odds ratio (OR) = 1.43; $P = 6.40 \times 10^{-3}$, OR = 1.57; and $P = 5.00 \times 10^{-3}$, OR = 1.41, respectively). Meta-analysis of current and previous GWAS results indicated a significant association with nephrolithiasis ($P = 7.65 \times 10^{-15}$, 7.86×10^{-14} and 1.06×10^{-9}, respectively). We observed a cumulative effect with these three SNPs; individuals with three or more risk alleles had a 5.9-fold higher risk for nephrolithiasis development than those with only one risk allele. Our findings elucidated the significance of genetic variation at these three loci in nephrolithiasis in the Japanese population.
Journal of Human Genetics (2013) **58**, 588–593; doi:10.1038/jhg.2013.59; published online 30 May 2013

Keywords: genome-wide association study; nephrolithiasis; single-nucleotide polymorphisms; sodium-phosphate cotransporter; urinary calculi; urolithiasis

INTRODUCTION

Nephrolithiasis, also known as kidney stones, is a worldwide health problem that affects nearly all populations, causing severe acute back pain, and occasionally leading to more severe complications such as pyelonephritis or acute renal failure. The lifetime prevalence of nephrolithiasis in Japan is estimated to be 15.1% in men and 6.8% in women,[1] with a recurrence rate of nearly 60% in patients within 10 years after their initial treatment.[2]

Nephrolithiasis is a multifactorial disease resulting from a complex interaction between environmental and genetic factors. Environmental factors, such as lifestyle, obesity, dietary habits and dehydration, have been implicated in nephrolithiasis development,[3,4] while hormonal, genetic or anatomical factors may also influence its pathogenesis.[5] In addition, a family history of the disease has been reported to increase the disease risk in men by 2.57-fold,[6] and the concordance rate of the disease in monozygotic twins has been found to be higher than that in dizygotic twins (32.4% vs 17.3%),[7] suggesting a pivotal role for genetic factors as etiological factors for nephrolithiasis. Previous studies have indicated that genetic polymorphisms of the genes encoding the calcium (Ca)-sensing receptor, vitamin D receptor and osteopontin are highly correlated with kidney stone formation.[8] Thus, these genetic polymorphisms are important putative markers for nephrolithiasis risk among Caucasian populations.

In 2012, a genome-wide association study (GWAS) of nephrolithiasis in the Japanese population, including 5796 patients with nephrolithiasis and 17 344 healthy controls, was performed.[9] This study identified three novel loci for nephrolithiasis at 5q35.3 (*RGS14-SLC34A1-PFN3-F12*), 7p14.3 (*INMT-FAM188B-AQP1*) and 13q14.1 (*DGKH*).[9] However, to evaluate the role of these genetic factors, a validation study using an independent case–control sample set is essential. Here, we performed a replication study using 601 nephrolithiasis and 201 control subjects.

[1]Department of Nephro-urology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; [2]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan; [3]Department of Urology, Ijinkai Takeda General Hospital, Kyoto, Japan; [4]Department of Urology, Sanjukai Hospital, Sapporo, Japan and [5]Center for Genomic Medicine, The Institute of Physical and Chemical Research (RIKEN), Kanagawa, Japan
Correspondence: Dr T Yasui, Department of Nephro-urology, Nagoya City University Graduate School of Medical Sciences, 1, Kawasumi, Mizuho-cho, Mizuho-ku, Nagoya 467-8601, Japan.
E-mail: yasui@med.nagoya-cu.ac.jp
Received 13 January 2013; revised 23 April 2013; accepted 26 April 2013; published online 30 May 2013

MATERIALS AND METHODS

Ethics statement

All subjects provided written informed consent. The study protocol conformed to the Declaration of Helsinki, and the study was approved by the ethical committees at the Nagoya City University Graduate School of Medical Sciences, Ijinkai Takeda General Hospital, and Sanjukai Hospital.

Subjects

The clinical characteristics of case and control samples are shown in Table 1. In total, 601 unrelated Japanese patients with nephrolithiasis (mean age ± s.d., 55.0 ± 14.0 years; 443 men and 158 women) were recruited at Nagoya City University Hospital, Ijinkai Takeda General Hospital, and Sanjukai Hospital (Table 1). Patients were excluded if they had a history of chronic urinary tract infection, renal failure, chronic diarrhea, gout, renal tubular acidosis, primary and secondary hyperparathyroidism or cancer. We excluded patients with struvite, cystine, ammonium-acid urate and uric acid stones. Patients with secondary nephrolithiasis caused by drugs, hyperparathyroidism or congenital anomalies of the urinary tract were also excluded. We also excluded anyone who had regularly taken diuretics, vitamin D or Ca supplements 1 or more times per week in the 6 months leading up to the diagnosis of Ca nephrolithiasis or the interview, because these substances may confound the results. The control group consisted of 201 Japanese subjects (mean age ± s.d., 53.2 ± 20.7 years; 166 men and 35 women) without a history of nephrolithiasis or a family history of kidney stone disease (Table 1). All control subjects underwent a medical examination and radiography (non-contrast-enhanced computed tomography) examination, along with routine blood and urine assays, for confirmation. All the patients with nephrolithiasis and the control subjects were recruited from the same racial, ethnic, geographical and environmental strata. The protocol was previously approved by the appropriate Institutional Review Boards, and informed consent was obtained from all the patients and control subjects enrolled in this study.

We assessed the effect of genetic variations on serum phosphorus, Ca, urate and creatinine levels; estimated glomerular filtration rate (eGFR); and body mass index. We analyzed a total of 230 case and control samples. In this analysis, disease status was used as a covariate. eGFR was calculated using the following formula:[10] eGFR (ml min^{-1} · 1.73 m^{-2}) = 194 × serum creatinine (mg per 100 ml)$^{-1.094}$ × age$^{-0.287}$ (× 0.739 for female subjects).

The subjects (601 cases and 201 controls) examined in this study were distinct populations, independent from those examined in the previous study (all samples were obtained from the BioBank Japan Project).[9] We performed a meta-analysis combining data from the current study with those from the GWAS study (both cases and controls). DNA samples from the 5796 patients with nephrolithiasis and 17 344 control subjects used in the prior GWAS study[9] were obtained from the BioBank Japan project, 'the Leading Project for Personalized Medicine' in the Ministry of Education, Culture, Sports, Science and Technology.[11]

Table 1 Demographic and clinical parameters of patients with nephrolithiasis and the controls

Characteristics	*Cases*	*Controls*	P-*value**
Total number	601	201	—
Sex (%)			
Male	443 (73.7)	166 (82.5)	0.88
Female	158 (26.3)	35 (17.5)	
Age			
Age (mean ± SD)	55.0 ± 14.0	53.2 ± 20.7	0.76
Primary age (%)			
⩾55 years	309 (51.4)	—	—
<55 years	282 (46.9)	—	—
Stone number (%)			
Single	247 (41.1)	—	—
Multiple	354 (58.9)	—	—
Stone frequency (%)			
Primary	278 (46.3)	—	—
Recurrence	323 (53.7)	—	—
Stone component (%)			
Ca oxalate (>50%)	539 (89.7)	—	—
Ca phosphate (>50%)	62 (10.3)	—	—

*P<0.05 deemed statistically significant.

DNA extraction

Genomic DNA was extracted from whole-blood samples by using a standard protocol. Whole-blood samples from patients and control subjects were centrifuged at 3000 r.p.m. for 10 min at 4 °C. The buffy coat was isolated from the blood samples. After lysis of the red blood cells in a lysis buffer, the samples were mixed in cell lysis buffer for several days. Protein precipitation solution was added to precipitate the contaminating cellular proteins. Finally, total genomic DNA was isolated by precipitation with 95% isopropanol and 80% ethanol.

Selection of single-nucleotide polymorphisms and genotyping

We selected a total of seven single-nucleotide polymorphisms (SNPs) at three loci at 5q35.3, 7p14.3 and 13q14.1, which had previously been reported to be significantly associated with nephrolithiasis ($P<5\times10^{-8}$).[9]

All the selected SNPs were subjected to genotyping using a multiplex PCR-Invader assay (Third Wave Technologies, Madison, WI, USA), using a detailed methodological protocol described previously.[11] Multiplexed amplification was performed with 20 ng of genomic DNA. The total reaction volume of 50 µl for each sample consisted of 50 pmol of each primer, 10 units of Ex-Taq DNA polymerase and 0.55 µg of TaqStart (Clontech Laboratories, Palo Alto, CA, USA). Samples were amplified in the GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA, USA). Allele-specific oligonucleotide pairs and invasive probes were designed and supplied by Third Wave Technologies. Fluorescence resonance energy transfer probes were labeled with carboxyfluorescein (FAM) or VIC dye (Applied Biosystems), corresponding to each allele. Signal intensity was calculated as the ratio of FAM or VIC to ROX, an internal passive reference dye. The total reaction volume of 10 µl for each sample contained 0.5 µl of signal buffer, 0.5 µl of fluorescence resonance energy transfer probes, 0.5 µl of structure-specific cleavage enzyme, 1 µl of allele-specific probe mix and 2 µl of PCR product diluted 1:10 (v/v) in signal buffer. Samples were incubated at 95 °C for 5 min, followed by incubation at 63 °C for 15 min, in an ABI7900 Real Time PCR system (Applied Biosystems).

Statistical analysis

The statistical differences between the case and controls group were analyzed by the χ^2-square test. The association of SNPs with nephrolithiasis was tested by a Cochran–Armitage trend test.[12] We assumed a significance level of $P=7.14\times10^{-3}$ (0.05/7). The odds ratios (ORs) were calculated using the non-susceptible allele as a reference. The combined analysis of the results of this study and the previous GWAS was performed using the Mantel–Haenszel method. Heterogeneity across the two cohorts was examined using the Breslow-Day test.[13] We conducted association and QTL analyses using the plink-1.06 toolset (pngu.mgh.harvard.edu/~purcell/plink/).

RESULTS

Genetic polymorphisms at 5q35.3, 7p14.3 and 13p14.1

The seven SNPs at 5q35.3, 7p14.3 and 13q14.1 loci, which were significantly associated with nephrolithiasis, were genotyped using the Invader assay. Table 2 shows the genotype frequencies of polymorphism among all subjects. The genotype frequencies of seven

Table 2 Results of association analysis for nephrolithiasis in the Japanese population

					Cases				*Controls*						
Chr no.	*Chr location*	*SNP*	*Gene*	*Minor and major alleles*	*n(11)*[a]	*n(12)*[a]	*n(22)*[a]	*MAF*	*n(11)*[a]	*n(12)*[a]	*n(22)*[a]	*MAF*	P[b]	*OR*[c]	*95% CI*[c]
5	176794191	rs12654812	*RGS14*[d]	T/C	97	299	205	0.410	22	87	91	0.328	$\mathbf{3.12 \times 10^{-3}}$	1.43	(1.13–1.81)
5	176798306	rs11746443	*RGS14*[d]	T/C	43	249	303	0.282	12	75	113	0.248	1.81×10^{-1}	1.19	(0.92–1.54)
7	30915478	rs12669187	*FAM188*[e]	T/C	29	182	389	0.200	4	47	149	0.138	$\mathbf{6.40 \times 10^{-3}}$	1.57	(1.14–2.15)
7	30937178	rs1000597	No gene	G/A	40	197	356	0.234	4	66	131	0.184	4.26×10^{-2}	1.35	(1.01–1.80)
13	42690060	rs7981733	*DGKH*	A/G	57	246	298	0.300	28	95	78	0.376	$\mathbf{5.00 \times 10^{-3}}$	1.41	(1.11–1.78)
13	42702711	rs1170155	*DGKH*	A/G	66	256	277	0.324	33	93	75	0.396	9.84×10^{-3}	1.37	(1.08–1.73)
13	42754522	rs4142110	*DGKH*	A/G	115	266	220	0.413	43	96	61	0.455	1.52×10^{-1}	1.19	(0.95–1.49)

Abbreviations: Chr, chromosome; CI, confidence interval; MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism.
We analyzed 601 nephrolithiasis cases and 201 controls. Significant ($P<7.14 \times 10^{-3}$ (0.05/7) values are in bold.
[a]*n* (11), number of subjects with homozygous genotypes for the minor allele; *n* (12), number of subjects with heterozygous genotypes; *n* (22), number of subjects with homozygous genotypes for the major allele.
[b]*P*-value obtained from the Cochrane-Armitage trend test.
[c]ORs and CIs have been calculated using the non-susceptible allele as the reference.
[d]RGS14, RGS14-SLC34A1-PFN3-F12.
[e]FAM188B, INMT-FAM188B-AQP1.

SNPs among case and control subjects were distributed in accordance with the Hardy–Weinberg equilibrium.

Each of the three loci in this study possesses a significant number of SNPs, assuming a significance level of $P=7.14 \times 10^{-3}$ (0.05/7). Three SNPs at three regions—SNP rs12654812 at the 5q35.3 region containing *RGS14-SLC34A1-PFN3-F12* genes, rs12669187 at the 7p14.3 region containing *INMT-FAM188B-AQP1* and SNP rs7981733 at the 13p14.1 locus of the *DGKH* region—were significantly associated with nephrolithiasis ($P=3.12 \times 10^{-3}$, OR = 1.43; $P=6.43 \times 10^{-3}$, OR = 1.57; and $P=5.00 \times 10^{-3}$, OR = 1.41, respectively). As shown in Table 2, the T allele of rs12654812, the T allele of rs12669187 and the G allele of rs7981733 were found to increase the risk for nephrolithiasis development. Two SNPs—SNP rs1000597 at the 7p14.3 region and SNP rs1170155 at the 13p14.1 region—also indicated suggestive associations with nephrolithiasis ($P<0.05$). The risk alleles in all seven SNPs of this study are identical to those of the previous analysis.[9]

Next, we selected the most significant SNPs from each of the three genomic regions and examined the association of these SNPs with several clinical parameters. These SNPs did not associated with serum Ca, urate, creatinine, eGFR and body mass index levels, but the risk allele of rs12654812 was associated with lower serum phosphorus levels ($P=0.0353$; Table 3) although the association was not statistically significant after a multiple testing correction.

Table 3 QTL analysis for serum levels of phosphorus, calcium, urate and creatinine; eGFR; and BMI

Traits	*SNP*	*MAF*	*Beta*	*s.e.*	P[a]
Serum phosphorus					
	rs12654812	0.438	−0.185	0.08688	0.0353
	rs12669187	0.186	0.045	0.12270	0.7147
	rs7981733	0.324	−0.053	0.09875	0.5937
Serum calcium					
	rs12654812	0.394	0.081	0.05545	0.1454
	rs12669187	0.166	0.018	0.07422	0.8055
	rs7981733	0.361	−0.017	0.06093	0.7815
Serum urate					
	rs12654812	0.394	0.135	0.1607	0.4015
	rs12669187	0.185	0.013	0.2060	0.9480
	rs7981733	0.332	0.158	0.1763	0.3725
Serum creatinine					
	rs12654812	0.394	−0.012	0.03887	0.7628
	rs12669187	0.162	−0.023	0.05218	0.6666
	rs7981733	0.355	−0.028	0.04197	0.5093
Estimated glomerular filtration rate					
	rs12654812	0.396	−0.922	3.238	0.7760
	rs12669187	0.163	3.217	4.346	0.4599
	rs7981733	0.352	2.604	3.516	0.4597
Body mass index					
	rs12654812	0.388	0.322	0.4622	0.4878
	rs12669187	0.150	−0.392	0.6243	0.5306
	rs7981733	0.356	0.097	0.4949	0.8446

Abbreviations: BMI, body mass index; eGFR, estimated glomerular filtration rate; MAF, minor allele frequency; QTL, quantitative trail locus; SNP, single-nucleotide polymorphism.
QTL analysis for serum levels of phosphorus ($n=104$), calcium ($n=224$), urate ($n=182$) and creatinine ($n=229$); eGFR ($n=225$); and BMI ($n=158$). eGFR is calculated by formulas using blood test results, age and gender. eGFR ($\mathrm{ml\,min^{-1} \cdot 1.73\,m^{-2}}$) = 194 × serum creatinine (mg per 100 ml)$^{-1.094}$ × age$^{-0.287}$ (× 0.739 if female).
[a]*P*-values were obtained from logistic regression analysis using disease status as a covariate.

Cumulative effect of nephrolithiasis risk alleles

Subsequently, we examined the cumulative effects of these three SNPs (rs12654812, rs12669187, and rs7981733) in the *RGS14*, *FAM188B* and *DGKH* genes on nephrolithiasis susceptibility. The risk for nephrolithiasis development increased with an increasing number of risk alleles (T, T and G on rs12654812, rs12669187 and rs7981733, respectively) for three SNPs (Figure 1). Individuals with three or more risk alleles showed >5.9-fold higher risk (5.8-fold in first-stone cases and 6.4-fold in recurrent-stone cases) of developing nephrolithiasis than those with one risk allele alone. Additionally, individuals with no risk alleles showed 0.36-fold lower risk of developing nephrolithiasis than those with one risk allele. Taken together, our findings are indicative of the additive effects of variants in three loci on nephrolithiasis susceptibility.

Meta-analysis with a previous BioBank cohort

Finally, we conducted a meta-analysis of the current and the previous study (two-stage GWAS and one replication) by using a total of 5796 patients and 17 344 controls[9] with a fixed-effects model using the Mantel–Haenszel method. All four stages exhibited similar trends of association. The Mantel–Haenszel *P*-values for independence were 7.65×10^{-15} for rs12654812 (OR = 1.17; 95% confidence

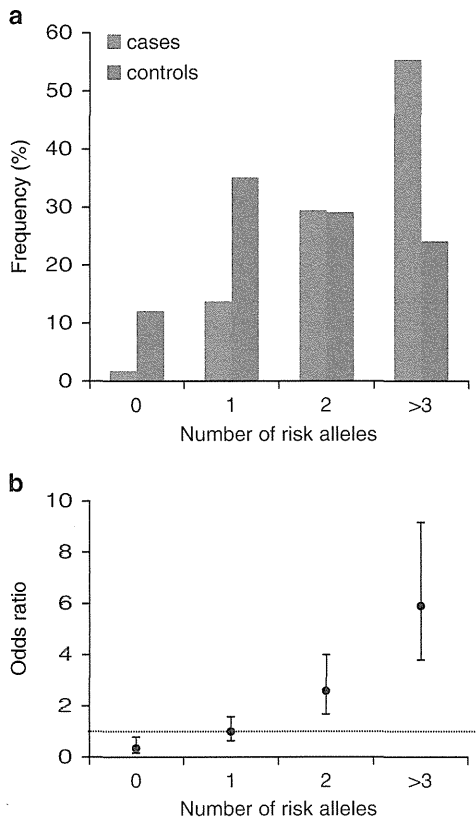

Figure 1 Cumulative effect of nephrolithiasis risk alleles. (**a**) Distribution of risk alleles in nephrolithiasis cases (red bars) and controls (blue bars). (**b**) The OR for nephrolithiasis increases with increase in the number of risk alleles. The ORs are relative to the maximum number of risk alleles (number of risk allele = 1) in the three SNPs (rs12654812, rs12669187, and rs7981733). The vertical bars correspond to 95% confidence interval . The horizontal line indicates the null value (OR = 1).

interval, 1.12–1.22), 7.86×10^{-14} for rs12669187 (OR = 1.22; 95% confidence interval, 1.16–1.29) and 1.06×10^{-9} for rs7981733 (OR = 1.15; 95% confidence interval, 1.10–1.21; Table 4).

DISCUSSION

Nephrolithiasis is one of the most common urological diseases in the developing world. Analysis of the chemical composition of urinary stones has demonstrated that Ca oxalate stones are the most prevalent stones, occurring in ~75–80% of cases of nephrolithiasis.[1] Many case–control association studies and linkage studies have been performed to identify common genetic variations associated with nephrolithiasis. However, most of the reported variations have not been successfully reproduced experimentally. This study was performed to verify the relationship of genetic variants that were previously reported to be associated with nephrolithiasis in the Japanese population.

Here, we successfully validated three nephrolithiasis loci (5q35.5, 7p14.33 and 13q14.1) in an independent cohort. These three loci were associated with nephrolithiasis, at least in the Japanese population. Furthermore, this study showed that the accumulation of risk alleles of SNPs in these three loci increases the risk of nephrolithiasis development. The distribution of SNPs in this study was confirmed by comparison with previous GWAS data. The subjects examined in this study were independent from those of the BioBank Japan Project. As 17 344 control subjects from BioBank in previous report[9] had other diseases, this complex background of control subjects may lead to false-positive or false-negative results. In contrast, nephrolithiasis was specifically excluded in our 201 control samples by medical examination (computed tomography and ultrasonography). In this replication study, the three loci showed a stronger association (OR, 1.43, 1.57 and 1.41) than those reported in the previous GWAS. Thus, these three SNPs are common nephrolithiasis loci in the Japanese population.

The results of an analysis of the protein-encoding genes near the 5q35.3, 7p14.3 and 13q14.1 loci are suggestive of several intriguing potential associations with nephrolithiasis. The *SLC34A1* gene located

Table 4 Meta-analysis of current study and the previous GWAS

SNP	*Test*	*Gene*	*Minor and major alleles*	*Case MAF*	*Control MAF*	P[a]	*OR*[b]	*95% CI*[b]	P_{het}[d]
rs12654812	GWAS			0.397	0.346	1.98×10^{-5}	1.24	(1.13–1.37)	
	Stage 2			0.381	0.354	5.57×10^{-4}	1.13	(1.05–1.20)	
	Replication	*RGS14*	T/C	0.386	0.350	5.26×10^{-5}	1.17	(1.08–1.26)	
	This study			0.410	0.328	3.12×10^{-3}	1.43	(1.13–1.81)	
	Combined[c]					7.56×10^{-15}	1.17	(1.12–1.22)	0.136
rs12669187	GWAS			0.223	0.176	1.04×10^{-6}	1.34	(1.19–1.51)	
	Stage 2			0.214	0.184	4.56×10^{-6}	1.21	(1.11–1.31)	
	Replication	*FAM188B*	T/C	0.205	0.182	1.59×10^{-3}	1.16	(1.06–1.27)	
	This study			0.200	0.138	6.40×10^{-3}	1.57	(1.14–2.15)	
	Combined[c]					7.86×10^{-14}	1.22	(1.16–1.29)	0.090
rs7981733	GWAS			0.283	0.342	4.08×10^{-7}	1.32	(1.19–1.47)	
	Stage 2			0.314	0.341	5.09×10^{-4}	1.13	(1.06–1.21)	
	Replication	*DGKH*	A/G	0.318	0.334	7.05×10^{-2}	1.07	(0.99–1.16)	
	This study			0.300	0.376	5.00×10^{-3}	1.41	(1.11–1.78)	
	Combined[c]					1.06×10^{-9}	1.15	(1.10–1.21)	0.003

Abbreviations; CI, confidence interval; GWAS, genome-wide association study; MAF, minor allele frequency; OR, odds ratio; SNP, single-nucleotide polymorphism.
In this analysis, 6397 (5796 in Biobank and 601 in this study) nephrolithiasis cases and 17535 (17344 in Biobank and 201 in this study) controls were included.
[a]*P*-value obtained from the Cochran–Armitage trend test.
[b]ORs and CIs were calculated using the non-susceptible allele as the reference.
[c]Combined: ORs and *P*-values for the independence test were calculated by the Mendel–Haenszel method in the meta-analysis.
[d]The *P*-values of heterogeneities (P_{het}) across the four stages were examined using the Breslow-Day test.

in the *RGS14-SLC34A1-PFN3-F12* region encodes NPT2a, a member of the type IIa sodium-phosphate cotransporter family, which is highly expressed in the kidney.[14] The NPT2a protein family, located on the apical membrane of renal proximal tubular epithelial cells, is essential in maintaining phosphate homeostasis. Mutations in *SLC34A1* have been reported to cause hypophosphatemic nephrolithiasis and osteoporosis in humans[15] and severe renal phosphate wasting and hypercalciuria in mice.[16] In addition, previously published GWAS reports revealed the association of variations in the *SLC34A1* locus with kidney function[17] and serum phosphorus concentration.[18] In the QTL analyses in this study, rs12654812, located near *SLC34A1*, was associated with serum phosphorus levels. This result is consistent with the results of a previous study in which variants near *SLC34A1* were associated with serum phosphorus levels.[18] The SLC34A1 mutation was also shown to reduce renal phosphate resorption and recurrent nephrolithiasis.[19] As *SLC34A1* encodes the protein NPT2, a member of the type II sodium-phosphate cotransporter family, the presence of SNP rs12654812 is associated with reduced NPT2 function, resulting in subsequent reduction of serum phosphorus levels and an increased risk of nephrolithiasis development.

Targeted deletion of the *Npt2a* gene in mice resulted in increased urinary excretion of inorganic phosphate (Pi), an ~80% decrease in renal brush border membrane Na/Pi cotransport, and hypophosphatemia, which leads to increased levels of serum calcitriol (1,25(OH)2D, a biologically active vitamin D metabolite), intestinal Ca channel overexpression, increased intestinal Ca hyperabsorption, hypercalcemia and hypercalciuria. NPT2a null mice develop renal deposits of apatitic Ca phosphate,[18] a primary constituent of idiopathic Ca oxalate stones.[20] Ca oxalate stones develop while attached to Randall's plaques or Ca phosphate crystal deposits in renal tubules.[21] Randall's plaques are composed of poorly crystalline biologically active Ca phosphate, reported to begin at the basement membrane of the loops of Henle.[22,23] Ultrastructural investigation of Npt2a null mice uncovered the Ca phosphate crystals formed in the tubular lumina; these crystals were organized as microspheres. It has been suggested that variations in this region could regulate renal function and subsequently affect the risk of nephrolithiasis development.

In the previous study, rs11746443 at 5q35.3 was significantly associated with the reduction in eGFR by QTL analysis.[9] However, we found no significant association with eGFR in this study. This could be partially explained by the smaller sample size in this study ($n = 229$) compared with that of previous study ($n = 27\,323$).

The SNPs rs12669187 and rs1000597 on chromosome 7p14.3 are located on the *FAM188B* gene and between the *FAM188B* gene (5.2 kb downstream), and the *AQP1* gene (14.2 kb upstream), respectively. However, the role of FAM188B in the pathogenesis of nephrolithiasis has not yet been elucidated. Aquaporin-1 is abundantly expressed in the kidney, functioning as a water channel.[24] Moreover, *Aqp1* null mice exhibit reduced osmotic permeability in the membrane of the kidney proximal tube and become severely dehydrated after water deprivation, suggesting an important role for aquaporin-1 in the urinary concentration mechanism.[24] Therefore, the SNPs rs12669187 and rs1000597 are likely to be associated with the regulation of FAM188B and/or AQP1 expression and may affect the urine concentration process and increase the risk of nephrolithiasis development.

The SNP rs7981733 on *DGKH* is significantly associated with nephrolithiasis. DGKH, which is highly expressed in the brain, is potentially associated with psychiatric disorders such as bipolar disorder and major depressive disease,[25] but its involvement in renal function or Ca homeostasis has not been reported previously. Therefore, further functional analyses are essential to elucidate the role of this variation as a causative factor for nephrolithiasis.

During examination of the cumulative effect of three SNPs in *RGS14*, *FAM188B* and *DGKH*, 1 risk allele was found to be the most common (35%) in control subjects. The risk of nephrolithiasis development increases with increase in the frequency of risk alleles for the three SNPs (Figure 1). In addition, patients with recurrent stones exhibited higher ORs (OR = 6.4) than patients who had developed a stone for the first time (OR = 5.8). Our findings clearly indicate the existence of an additive effect of these three variants on nephrolithiasis susceptibility; therefore, SNP analysis would be an effective predictor of nephrolithiasis risk and recurrence.

The previous GWAS data in Caucasian populations showed that CLDN14 is associated with nephrolithiasis.[26] In the Japanese population, SNP rs2835349, located 19 kb upstream of the *CLDN14* gene, exhibited a probable association with the *CLDN14* gene ($P = 6.33 \times 10^{-5}$; OR = 1.22) in the GWAS stage.[9] However this SNP was not identified in the replication cohort ($P = 0.624$; OR = 0.98). Thus, the *CLDN14* gene is unlikely to have a strong impact on nephrolithiasis development in the Japanese population. Furthermore, another genome-wide meta-analysis has shown that Ca-sensing receptor variants are associated with serum Ca level in Caucasians and Indian Asian[27] However, SNPs in the *CASR* gene was not associated with nephrolithiasis risk in the previous GWAS (data not shown).[9]

In agreement with the aforementioned findings, associations between the three novel loci (identified by GWAS and replication studies) and nephrolithiasis were observed. Our results helped elucidate the crucial roles of genetic factors related to nephrolithiasis development. Furthermore, functional analysis is necessary to completely elucidate the role of these variations in the development of nephrolithiasis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Ms Ichikawa, Ms Kawamura, Ms Noda and Ms Kasuga for their secretarial assistance. This work was partly supported by MEXT/JSPSKAKENHI (grant numbers 23249074, 23592375, 24659716 and 24689060); the BioBank Japan Project funded by the Japanese Ministry of Education, Culture, Sports, Science and Technology; the Takeda Science Foundation; and the Japanese Urological Association. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

1 Yasui, T., Iguchi, M., Suzuki, S. & Kohri, K. Prevalence and epidemiological characteristics of urolithiasis in Japan: national trends between 1965 and 2005. *Urology* **71**, 209–213 (2008).
2 Strohmaier, W. L. Course of calcium stone disease without treatment. What can we expect? *Eur. Urol.* **37**, 339–344 (2000).
3 Tasca, A. Metabolic syndrome and bariatric surgery in stone disease etiology. *Curr. Opin. Urol.* **21**, 129–133 (2011).
4 Taylor, E. N., Stampfer, M. J. & Curhan, G. C. Obesity, weight gain, and the risk of kidney stones. *J. Am. Med. Assoc.* **293**, 455–462 (2005).
5 Worcester, E. M. & Coe, F. L. Clinical practice. Calcium kidney stones. *N. Engl. J. Med.* **363**, 954–963 (2010).
6 Curhan, G. C., Willett, W. C., Rimm, E. B. & Stampfer, M. J. Family history and risk of kidney stones. *J. Am. Soc. Nephrol.* **8**, 1568–1573 (1997).
7 Goldfarb, D. S., Fischer, M. E., Keich, Y. & Goldberg, J. A twin study of genetic and dietary influences on nephrolithiasis: a report from the Vietnam EraTwin (VET) Registry. *Kidney Int.* **67**, 1053–1061 (2005).

8 Vezzoli, G., Terranegra, A., Arcidiacono, T. & Soldati, L. Genetics and calcium nephrolithiasis. *Kidney Int.* **80,** 587–593 (2011).
9 Urabe, Y., Tanikawa, C., Takahashi, A., Okada, Y., Morizono, T., Tsunoda, T. *et al.* A genome-wide association study of nephrolithiasis in the Japanese population identifies novel susceptible Loci at 5q35.3, 7p14.3, and 13q14.1. *PLoS Genet.* **8,** e1002541 (2012).
10 Matsuo, S., Imai, E., Horio, M., Yasuda, Y., Tomita, K., Nitta, K. *et al.* Revised equations for estimated GFR from serum creatinine in Japan. *Am. J. Kidney Dis.* **53,** 982–992 (2009).
11 Nakamura, Y. The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.* **5,** 696–697 (2007).
12 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
13 Breslow, N. E. & Day, N. E. Statistical methods in cancer research. Volume II—The design and analysis of cohort studies. *IARC Sci. Publ.* **82,** 1–406 (1987).
14 Weinman, E. J., Cunningham, R., Wade, J. B. & Shenolikar, S. The role of NHERF-1 in the regulation of renal proximal tubule sodium-hydrogen exchanger 3 and sodium-dependent phosphate cotransporter 2a. *J. Physiol.* **567,** 27–32 (2005).
15 Prié, D., Huart, V., Bakouh, N., Planelles, G., Dellis, O., Gérard, B. *et al.* Nephrolithiasis and osteoporosis associated with hypophosphatemia caused by mutations in the type 2a sodium-phosphate cotransporter. *N. Engl. J. Med.* **347,** 983–991 (2002).
16 Beck, L., Karaplis, A. C., Amizuka, N., Hewson, A. S., Ozawa, H., Tenenhouse, H. S. *et al.* Targeted inactivation of Npt2 in mice leads to severe renal phosphate wasting, hypercalciuria, and skeletal abnormalities. *Proc. Natl Acad. Sci. USA* **95,** 5372–5377 (1998).
17 Köttgen, A., Pattaro, C., Böger, C. A., Fuchsberger, C., Olden, M., Glazer, N. L. *et al.* New loci associated with kidney function and chronic kidney disease. *Nat. Genet.* **42,** 376–384 (2010).
18 Kestenbaum, B., Glazer, N. L., Köttgen, A., Felix, J. F., Hwang, S. J., Liu, Y. *et al.* Common genetic variants associate with serum phosphorus concentration. *J. Am. Soc. Nephrol.* **21,** 1223–1232 (2010).
19 Lapointe, J. Y., Tessier, J., Paquette, Y., Wallendorff, B., Coady, M. J., Pichette, V. *et al.* NPT2a gene variation in calcium nephrolithiasis with renal phosphate leak. *Kidney Int.* **69,** 2261–2267 (2006).
20 Khan, S. R. Calcium phosphate/calcium oxalate crystal association in urinary stones: implications for heterogeneous nucleation of calcium oxalate. *J. Urol.* **157,** 376–383 (1997).
21 Coe, F. L., Evan, A. P., Worcester, E. M. & Lingeman, J. E. Three pathways for human kidney stone formation. *Urol. Res.* **38,** 147–160 (2010).
22 Randall, A. The origin and growth of renal calculi. *Ann. Surg.* **105,** 1009–1027 (1937).
23 Evan, A. P., Lingeman, J. E., Coe, F. L., Parks, J. H., Bledsoe, S. B., Shao, Y. *et al.* Randall's plaque of patients with nephrolithiasis begins in basement membranes of thin loops of Henle. *J. Clin. Invest.* **111,** 607–616 (2003).
24 Ma, T., Yang, B., Gillespie, A., Carlson, E. J., Epstein, C. J. & Verkman, A. S. Severely impaired urinary concentrating ability in transgenic mice lacking aquaporin-1 water channels. *J. Biol. Chem.* **273,** 4296–4299 (1998).
25 Barnett, J. H. & Smoller, J. W. The genetics of bipolar disorder. *Neuroscience* **164,** 331–343 (2009).
26 Thorleifsson, G., Holm, H., Edvardsson, V., Walters, G. B., Styrkarsdottir, U., Gudbjartsson, D. F. *et al.* Sequence variants in the CLDN14 gene associate with kidney stones and bone mineral density. *Nat. Genet.* **41,** 926–930 (2009).
27 Kapur, K., Johnson, T., Beckmann, N. D., Sehmi, J., Tanaka, T., Kutalik, Z. *et al.* Genome-wide meta-analysis for serum calcium identifies significantly associated SNPs near the calcium-sensing receptor (CASR) gene. *PLoS Genet.* **6,** e1001035 (2010).

OPEN ACCESS Freely available online

PLOS | GENETICS

Meta-Analysis of Genome-Wide Association Studies Identifies Six New Loci for Serum Calcium Concentrations

Conall M. O'Seaghdha[1,2¶], Hongsheng Wu[1,3,4], Qiong Yang[1,3], Karen Kapur[5], Idris Guessous[6,7,8¶], Annie Mercier Zuber[9], Anna Köttgen[10,11¶], Candice Stoudmann[9], Alexander Teumer[12], Zoltán Kutalik[5,13], Massimo Mangino[14], Abbas Dehghan[15], Weihua Zhang[16,17], Gudny Eiriksdottir[18], Guo Li[19], Toshiko Tanaka[20], Laura Portas[21], Lorna M. Lopez[22], Caroline Hayward[23], Kurt Lohman[24], Koichi Matsuda[25], Sandosh Padmanabhan[26], Dmitri Firsov[9], Rossella Sorice[27], Sheila Ulivi[28], A. Catharina Brockhaus[29,30], Marcus E. Kleber[31,32], Anubha Mahajan[33], Florian D. Ernst[12], Vilmundur Gudnason[18,34], Lenore J. Launer[35], Aurelien Mace[5,13], Eric Boerwinckle[36], Dan E. Arking[37], Chizu Tanikawa[25], Yusuke Nakamura[25], Morris J. Brown[38], Jean-Michel Gaspoz[39], Jean-Marc Theler[7], David S. Siscovick[19,40], Bruce M. Psaty[19,41,42], Sven Bergmann[5,13], Peter Vollenweider[43], Veronique Vitart[23], Alan F. Wright[23], Tatijana Zemunik[44], Mladen Boban[45], Ivana Kolcic[44], Pau Navarro[23], Edward M. Brown[46], Karol Estrada[47], Jingzhong Ding[46], Tamara B. Harris[35], Stefania Bandinelli[48], Dena Hernandez[49], Andrew B. Singleton[49], Giorgia Girotto[28], Daniela Ruggiero[27], Adamo Pio d'Adamo[28], Antonietta Robino[28], Thomas Meitinger[50,51], Christa Meisinger[52], Gail Davies[22], John M. Starr[22], John C. Chambers[16,17,53], Bernhard O. Boehm[54,55], Bernhard R. Winkelmann[56], Jie Huang[57], Federico Murgia[21], Sarah H. Wild[58], Harry Campbell[58], Andrew P. Morris[33], Oscar H. Franco[15], Albert Hofman[15], Andre G. Uitterlinden[47], Fernando Rivadeneira[47], Uwe Völker[12], Anke Hannemann[59], Reiner Biffar[60], Wolfgang Hoffmann[61], So-Youn Shin[62], Pierre Lescuyer[63], Hughes Henry[64], Claudia Schurmann[12], The SUNLIGHT consortium, The GEFOS consortium, Patricia B. Munroe[65], Paolo Gasparini[28], Nicola Pirastu[28], Marina Ciullo[27], Christian Gieger[29], Winfried März[32,66], Lars Lind[67], Tim D. Spector[14], Albert V. Smith[18,34], Igor Rudan[58], James F. Wilson[58], Ozren Polasek[44], Ian J. Deary[22], Mario Pirastu[21], Luigi Ferrucci[35], Yongmei Liu[68], Bryan Kestenbaum[69], Jaspal S. Kooner[16,70,71], Jacqueline C. M. Witteman[15], Matthias Nauck[59], W. H. Linda Kao[11,72], Henri Wallaschofski[59¶], Olivier Bonny[9,73¶], Caroline S. Fox[1,74¶*], Murielle Bochud[6¶*]

1 National Heart, Lung, and Blood Institute's Framingham Heart Study and Center for Population Studies, Framingham, Massachusetts, United States of America, **2** Renal Division, Massachusetts General Hospital, Boston, Massachusetts, United States of America, **3** Department of Biostatistics, Boston University, Boston, Massachusetts, United States of America, **4** Department of Medical Biology, University of Split, School of Medicine, Split, Croatia, **5** Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland, **6** Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Lausanne, Switzerland, **7** Unit of Population Epidemiology, Division of Primary Care Medicine, Department of Community Medicine and Primary Care and Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland, **8** Geriatric Unit, Azienda Sanitaria Firenze (ASF), Florence, Italy, **9** Department of Pharmacology and Toxicology, University of Lausanne, Lausanne, Switzerland, **10** Renal Division, Freiburg University Hospital, Freiburg, Germany, **11** Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America, **12** Interfaculty Institute for Genetics and Functional Genomics, Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany, **13** Swiss Institute of Bioinformatics, Lausanne, Switzerland, **14** King's College London, St. Thomas' Hospital Campus, London, United Kingdom, **15** Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands, **16** Catheter Lab, Cardiology, Ealing Hospital, Southall, Middlesex, United Kingdom, **17** Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom, **18** Icelandic Heart Association Research Institute, Kopavogur, Iceland, **19** Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, United States of America, **20** Clinical Research Branch, National Institute on Aging, Baltimore, Maryland, United States of America, **21** Institute of Population Genetics, CNR-Traversa La Crucca, Reg. Baldinca Li Punti, Sassari, Italy, **22** Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, United Kingdom, **23** MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Edinburgh, United Kingdom, **24** Cardiology Group, ClinPhenomics GmbH&Co KG, Frankfurt-Sachsenhausen, Germany, **25** Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan, **26** BHF Glasgow Cardiovascular Research Centre, Division of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, Scotland, **27** Institute of Genetics and Biophysics 'Adriano-Buzzati Traverso', CNR, Napoli, Italy, **28** Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", Trieste, Italy, **29** Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany, **30** Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians University Munich, Munich, Germany, **31** Department of Internal Medicine II – Cardiology, University of Ulm Medical Centre, Ulm, Germany, **32** Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany, **33** Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, United Kingdom, **34** University of Iceland, Reykjavik, Iceland, **35** Laboratory of Epidemiology, Demography and Biometry, National Institute on Aging, Bethesda, Maryland, United States of America, **36** University of Texas Health Science Center at Houston, Houston, Texas, United States of America, **37** McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **38** Cambridge Institute of Medical Research, University of Cambridge, Cambridge, United Kingdom, **39** Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands, **40** Departments of Medicine and Epidemiology, University of Washington, Seattle, Washington, United States of America, **41** Group Health Research Institute, Group Health Cooperative, Seattle, Washington, United States of America, **42** Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, Washington, United States of America, **43** Department of Medicine, Internal Medicine, Lausanne University Hospital, Lausanne, Switzerland, **44** Faculty of Medicine, University of Split, Split, Croatia, **45** Department of Pharmacology, Faculty of Medicine, University of Split, Split, Croatia, **46** Division of Laboratory Medicine, Geneva University