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

Sook Wah Yee1, Joel A Mefford2, Natasha Singh3, Mary-Elizabeth Percival3, Adrian Stecula1, Kuo Yang1, John S Witte2, Atsushi Takahashi4, Michiaki Kubo4, Koichi Matsuda5, Kathleen M Giacomini1,6,7 and Charalambos Andreadis3,6,7

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 \( \geq 3 \) years and may be cured is \( \approx 5–70\% \).3,6

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...
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.16 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 phenotype19 (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 used 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 (0.05/169) 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, RRMI, 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/Ideselect.php. The search options used in this step were: SNP data set = 1000 Genomes Pilot 1 in CEU population panel, r² threshold = 0.8 and distance limit = 500 kbp.

2. Potential regulatory functions were identified by searching the following databases: Regulome database20 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

Journal of Human Genetics
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 \(>0.01\). 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

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 lymphoblastoid 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.
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: \( SLC2A3, DCK, RRM1, GSTM1, ABCC3, MSH3, GSTT1, \) 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 \( \geq 3\% \). Other SNPs in the genes of the cytarabine pathway including \( NTSC2 \) 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 \( \geq 1\% \)), 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 \( SLC2A3 \) 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 \textit{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%)                        | Female 67 (53.6%)                           |
|                                  | Male 78 (50.6%)                        | Female 58 (46.4%)                           |
| Age (years)                      | Median (s.d., range)                   | Median (s.d., range)                        |
| Age at diagnosis                 | 47.0 (13.0, 18-72)                     | 40.5 (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%)                   | 1996–2000 41 (26.6%)                        |
|                                  | 2001–2010 77 (50.0%)                   | 56 (44.8%)                                  |
| De novo/secondary AML, N (%)     | De novo 147 (95.5%)                    | Secondary 110 (88.0%)                       |
|                                  | De novo 7 (4.5%)                       | Secondary 15 (12.0%)                        |
| Risk, N (%)                      | 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.)                          | 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 \((–\log_{10} \) 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 \( \geq 1\% \) 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 \((P < 3 \times 10^{-5})\).
Pharmacogenomic studies of AML response to chemotherapy. 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 regions and the ENCODE Chip-Seq and DNase I peaks (Supplementary Table 2).

Our study was focused primarily on 154 AML patients of European ancestry, consisting of cytarabine, etoposide and busulfan, followed by ASCT. Polymorphisms in pathway genes and AML response were also included (Supplementary Table 1, Figure 1 and Figure 2). The goals of this study were to identify new SNPs previously reported to be associated with AML response could be replicated. 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.

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.

### DISCUSSION

Previous pharmacogenomic studies of AML response to chemotherapy have been limited to a small number of candidate genes. 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.

Polymorphisms in pathway genes and AML response

*Journal of Human Genetics*
Overall, the most significant SNP in our analysis was the intronic variant in \(\text{ABCC3}, \text{rs4148405} (HR = 3.1, P = 9.5 \times 10^{-6})\). \(\text{ABCC3}\) 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 \(\text{ABCC3}\) in DFS in AML patients. First, in a previous study a promoter variant in \(\text{ABCC3}, \text{rs4793665}\) 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 \(\text{Abcc2}\) \(-/-\), \(\text{Abcc3}\) \(-/-\) mice, higher etoposide glucuronide levels were observed in the liver,\(^{32}\) consistent with a potentially important role of \(\text{ABCC3}\) in etoposide pharmacokinetics. Third, cell lines transfected with \(\text{ABCC3}\) show greater resistance to etoposide.\(^{33}\) Finally, higher \(\text{ABCC3}\) expression levels in leukemia cells are associated with poor outcome in children with leukemia.\(^{34,35}\)

Data from Regulome database suggest that the \(\text{ABCC3}\) variant, \(\text{rs4148405}\), is in a functional location in the genome, as this region demonstrates direct evidence of binding through ChIP-seq studies.\(^{20,36}\) Transcription factors that have a role in hepatic gene regulation (for example, \(\text{CEBPB, USF1, FOXA1}\)) have DNA response elements within this gene region (see \text{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 \(\text{rs4148405}\) is associated with higher expression levels of \(\text{ABCC3}\) in the liver and/or leukemia cells, and thus reduced levels of etoposide in the tumor cells. Although the SNP \(\text{rs4148405}\) is found at a \(\text{MAF of} \gg 10\%\) 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 \(\text{rs4148405}\) and potential causative SNPs may have confounded the analysis.

In addition to a SNP in \(\text{ABCC3}\), we also identified 23 SNPs in seven other genes (\(\text{SLC28A3, DCK, RRM1, GSTM1, GSTT1, MSH3 and SLC22A12}\)) 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,
Thus, the SNPs in these genes, which are involved in drug metabolism, drug response and disease risk in leukemia patients. Notably, GSTM1 and GSTT1 deletions have been implicated in various phenotypes associated with leukemia, including drug response, busulfan pharmacokinetics and disease risk. 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, brain and LCL. 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, 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. 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. In these genome-wide association studies, the minor allele T has been associated with higher uric acid levels. Although speculative, our study and the previous study, 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, 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. 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). 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 IC50 of another nucleoside analog, gemcitabine, in LCL. 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). 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, RRMI 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 deoxy-
cytidine triphosphate, which is important in building blocks for
DNA replication. Studies have shown that AML blasts cells with high
levels of deoxythymidine triphosphate are resistant to cytarabine, and
that there is a significant correlation between RRM1 and RRM2 gene
expression levels and deoxythymidine triphosphate levels after cytar-
abine treatment in AML blast cells.13 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 IC50 values in LCLs.16,17 A few SNPs in
the eight genes were significantly associated with response in AML
patients,49,53 our findings suggest that the SNPs in ABCC3
associated with AML response to chemotherapy that include
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 Par Mattson for their help in
creating plots in the R-project.
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 leukemia (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 leukemia patients. *Pharmacogenetics* **14**, 759–768 (2004).

39 Xiao, Z., Yang, L., 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., Crew, 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’Alò, F., Groner, S. et al. Prognostic role of glutathione S-transferases 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 Ernst-Murugesen, 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., Damarraju, 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)