

RESEARCH

Open Access



Exome-wide association study reveals 7 functional variants associated with ex-vivo drug response in acute myeloid leukemia patients

Anil K. Giri^{1,2,3*†}, Jake Lin^{1†}, Konstantinos Kyriakidis¹, Garima Tripathi¹ and Henrikki Almusa¹

Abstract

Acute myeloid leukemia (AML) is an aggressive blood cancer characterized by poor survival outcomes. Further, due to the extreme molecular heterogeneity of the disease, drug treatment response varies from patient to patient. The variability of drug response can cause unnecessary treatment in more than half of the patients with no or partial therapy responses leading to severe side effects, monetary as well as time loss. Understanding the genetic risk factors underlying the drug response in AML can help with improved prediction of treatment responses and identification of biomarkers in addition to mechanistic insights to monitor treatment response. Here, we report the results of the first Exome-Wide Association Study (EWAS) of ex-vivo drug response performed to date with 175 AML cases and 47 drugs. We used information from 55,423 germline exonic SNPs to perform the analysis. We identified exome-wide significant ($p < 9.02 \times 10^{-7}$) associations for rs113985677 in *CCIN* with tamoxifen response, rs115400838 in *TRMT5* with idelalisib response, rs11878277 in *HDGFL2* with entinostat, and rs2229092 in *LTA* associated with vorinostat response. Further, using multivariate genome-wide association analysis, we identified the association of rs11556165 in *ATRAID*, and rs11236938 in *TSKU* with the combined response of all 47 drugs and 29 nonchemotherapy drugs at the genome-wide significance level ($p < 5 \times 10^{-8}$). Additionally, a significant association of rs35704242 in *NIBAN1* was associated with the combined response for nonchemotherapy medicines ($p = 2.51 \times 10^{-8}$), and BI.2536, gefitinib, and belinostat were identified as the central traits. Our study represents the first EWAS to date on ex-vivo drug response in AML and reports 7 new associated loci that help to understand the anticancer drug response in AML patients.

Keywords CCIN, NIBAN1, Entinostat, AML, Drug response

[†]Anil K. Giri and Jake Lin contributed equally to this work.

*Correspondence:

Anil K. Giri

anil.kumar@helsinki.fi

¹Institute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland

²Foundation for the Finnish Cancer Institute (FCI), Tukholmankatu 8, Helsinki 00290, Finland

³iCAN Digital Precision Cancer Medicine Flagship, University of Helsinki and Helsinki University Hospital, Haartmaninkatu 8, PO Box 20, Helsinki FI-00014, Finland



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Acute myeloid leukemia (AML) is a biologically heterogeneous aggressive hematological malignancy [1, 2]. The inherent heterogeneity not only influences the clinical presentation of the disease but also affects the response and toxicity towards a therapy [3, 4]. Hence, huge inter-individual variability exists for drug response in AML patients [5–7].

The variability in the treatment response of AML patients is driven by a broad spectrum of genetic (e.g., somatic and germline variants) and nongenetic alterations (e.g., age, cancer cell maturation state, prior drug treatment) present in different patient groups [7, 8]. Recently, large scale projects like BeatAML have studied somatic mutations affecting drug response in AML [9, 10]. However, the germline variants contributing to the drug response in AML remain largely underexplored except for a few drugs (e.g. cytarabine) [11–15]. It is mostly because of a lack of well-characterized AML cohorts with both genetic and drug treatment data.

Further, the number of drugs that can be explored in a clinical cohort is limited because patients are usually treated with one or two common drugs approved in the clinics, hence neglecting the other drugs. Furthermore, the collection of clinical responses to antiAML medicine in a large patient cohort usually requires a long follow-up time as treatment response is often assessed over multiple time points, including induction remission, minimal residual disease status, relapse, and overall survival. Furthermore, such study requires large resources and a network of clinicians that could be beyond the capability of most academic laboratories [16].

The drug sensitivity and resistance testing (DSRT) platform can measure the ex-vivo response for hundreds of drugs at multiple doses in patient-derived primary cells [17–20]. The ex-vivo drug response can provide a reasonable estimate for the clinical response of the patient towards the drug [3, 9] and has been increasingly used in guiding the treatment for refractory/relapse patients with no available standardized treatment. Further, the ex-vivo drug response measured in patient-derived-primary samples has been used as a surrogate for clinical response to identify genetic variants affecting treatment response in a patient.

Here, using the germline information from the exome-sequence and ex-vivo response data from 47 drugs measured in 175 AML patients, we performed an exome-wide association study (EWAS) to identify functional variants affecting drug response in AML. The result identifies 4 novel loci near *CCIN*, *TRMT5*, *HDGFL2*, and *LTA* genes affecting the ex-vivo response of tamoxifen, idelalisib, entinostat, and vorinostat in AML at exome-wide significance levels. Further, we also used a novel and efficient statistical framework “MetaPhat” [21] to conduct and

decompose multivariate GWAS signals affecting ex-vivo drug response in AML patients. The analysis identified the association of loci in *NIBANI*/, *ATRAID*, and *TSKU* with combined drug response in AML.

Methods

AML patient’s cohort and ethical approval

One hundred seventy-five bone marrow aspirates and peripheral blood samples (leukemic cells) and skin biopsies (noncancerous cells for germline genomic information) from high-risk AML patients were obtained after informed consent with approval (No. 239/13/03/00/2010, 303/13/03/01/2011), latest amendment 7 dated June 15, 2016. Latest HUS study permit HUS/395/2018 dated February 13, 2018]. Patient characteristics are summarized in Supplementary Table 1.

The FIMM healthy donor cohort contained elderly individuals undergoing hip replacement surgery and healthy young adults volunteers. The age of healthy donors ranges between 19 and 78 years. Bone marrow aspirates were obtained with signed informed consent from these elderly healthy donors under the approval of the Tampere University Hospital Ethics Committee, Tampere, Finland (R15174) in accordance with the Declaration of Helsinki. Helsinki samples from adult patients with AML and young healthy donors were collected under the same approved ethical protocol as AML samples.

Exome-sequencing of the samples

Genomic DNA was isolated from the skin biopsies samples using the DNeasy Blood & Tissue Kit (Qiagen) and was used for exome sequencing. For exome sequencing, 3 µg DNA was fragmented and processed according to the NEBNext DNA Sample Prep Master Mix protocol. Exome capture was performed using the Nimblegen SeqCap EZ v2 capture Kit (Roche NimbleGen). Sequencing of exomes was done using HiSeq1500, 2000, or 2500 instruments (Illumina), and approximately 4×10^7 bp paired-end reads were sequenced per sample. The data was curated and analyzed as described before [9, 20].

Briefly, the sequenced raw exome reads were trimmed using Trimmomatic (version 0.36) to remove adapters and low-quality segments. The trimmed data were aligned to the human reference genome (GRCh build 37) using Burrows-Wheeler Aligner (bwa version 0.7.12, Wellcome Trust Sanger Institute, Cambridge, UK). After alignment, the potential PCR duplicates were removed using Picard (Broad Institute of Harvard and MIT, Cambridge, MA, USA), and BAM files were sorted, and indexed using SAMtools (Wellcome Trust Sanger Institute, Cambridge, UK). From the aligned reads, GATK (version 4.1.3.0) was used to joint call the samples for genetic variants. Variant annotation was performed using ANNOVAR and Ensembl version 6851.

We performed a stringent quality control analysis of the genetic data (Supplementary Fig. 1). We removed 1 sample with a genomic call rate of less than 95%, and 9 samples with missing information for >10 of the drug response. We also removed variants without proper annotation (without rsid), call rate (<90%), MAF<0.01, and Hardy Weinberg equilibrium p-value<0.05) were removed.

Drug resistance and sensitivity testing

We measured ex-vivo response for a library of 47 commercially available chemotherapeutic and targeted oncology compounds using the DSRT platform at the high throughput biomedicine unit at FIMM as described before [17, 22]. The library consists of 33 approved anti-cancer drugs that have the potential to be repurposed as antiAML therapy due to relevant mechanisms of action, and 14 investigation compounds (Supplementary Table 2). Out of the 33 approved drugs, 6 (cytarabine, daunorubicin, idarubicin, mitoxantrone, doxorubicin, and azacitidine) are in clinical use for the treatment of AML and others are strong candidates for repurposing against AML based on similar disease-related mechanisms. Out of 14 investigational compounds, 8 are being tested in clinical trials as antiAML drugs. The other 6 (ABT.751, BI.2536, Canertinib, Masitinib, Luminespib, Pictilisib) have been included in the study because they target cancer mechanisms relevant to AML biology.

Briefly, drugs were plated in a 10-fold dilution series of five concentrations covering a 1,000-fold concentration range on clear-bottomed 384-well plates (Corning #3712), using an Echo 550 Liquid Handler (Labcyte). The concentration ranges were selected for each compound separately to investigate their full dynamic range of the dose-response relationships. Next, 20 µl of mononuclear cells (AML cells) from freshly collected bone marrow and peripheral blood specimens (approximately 10,000) was added to a pre-drugged plate containing the mononuclear cell medium (MCM; PromoCell) supplemented with 0.5 µg/mL gentamicin and 2.5 µg/mL amphotericin B [3, 22]. We isolated mononuclear cells (MNC) from freshly collected AML patients' bone marrow/peripheral blood samples using the Ficoll-Paque centrifugation because it separates the mononuclear cells from the granulocytes and red blood cells, hence enriching leukemic cells.

The cell was maintained in the culture media at 37 °C with 5% CO₂ in a humidified incubator. After 72 h incubation, cell viability was measured using CellTiter-Glo (Promega) reagent in duplicate using a PHERAstar FS (BMG Labtech) plate reader. As positive (total killing) and negative (non-effective) controls, we used 100µM benzethonium chloride and 0.1% dimethyl sulfoxide

(DMSO), respectively, when calculating the relative inhibition percent.

The percentage inhibition was calculated by normalizing the cell viability to negative control wells containing only 0.1% dimethyl sulfoxide (DMSO) and positive control wells containing 100 µM cell-killing benzethonium chloride (BzCl). The drug responses passing the data quality assessment were included in further analysis. Drug sensitivity score (DSS), a measure of drug efficacy for each drug was calculated from a dose-response curve fitted using the cell inhibition level at 5 different drug concentrations for each patient as described before [23]. Next, selective DSS (sDSS) was calculated by normalizing drug responses against 17 healthy controls by subtracting the average DSS values for 17 healthy samples from the DSS score in the patient sample to estimate the selective efficacy of drugs against cancer cells. We used the selective drug sensitivity score as the measure for drug response of genetic association analysis.

Single SNP exome-wide association analysis

We selected 47 drugs for an analysis whose ex-vivo response was measured in all the patients. The sDSS scores for each drug were inverse normalized using the inbuilt command in R (<http://www.r-project.org/>) as a requirement for linear regression as it requires the outcome variable to follow a normal distribution. We performed linear regression-based association analysis using PLINK [24] to test the association of SNPs with inverse normalized sDSS. All other genetic parameters including minor allele frequency (MAF) and Hardy Weinberg equilibrium were estimated using the inbuilt command in PLINK. Cell culture medium for the drug assay, disease stage (relapse/diagnosis), and first five genetic principal components were used as covariates in the model. The principal components for the genotype matrix were estimated using the --pca command in PLINK. All other statistical analyses and plotting were performed using R programming. We used the Bonferroni correction for multiple tests ($n = 55423$ tests) and set the statistical significance and suggestive threshold to P-values less than 9.02×10^{-7} and 1×10^{-4} , respectively.

Multivariate exome-wide association analysis using metaphat

Several of the drugs measured in the experiment have a similar mechanism of action and act through the same pathway. As these functionally similar drugs have a correlated response in the studied samples, we used our novel method to conduct a multivariate exome wide association study of a linear combination of correlated traits using univariate results for 55,423 SNPs results as implemented in MetaPhat [21].

Further, using MetaPhat, we decomposed the multivariate associations into sets of central biomarker traits to understand the possible mechanism of action for the identified variant. Briefly, MetaPhat uses canonical correlation analysis (CCA) in its algorithm to identify traits of the highest statistical importance to each associated variant in multi-trait association studies. It identifies the subsets of the most irreplaceable traits at each associated variant by starting from all traits and removing one trait at a time until only one trait remains based on the lowest CCA score. In the process, it identifies driver traits defined as the traits that have been dropped based on the lowest CCA score at the step where the multivariate P-value was for the first time no longer genome-wide significant. Similarly, it also identifies a subset of traits, termed as the optimal subset which provides the lowest Bayesian Information Criterion (BIC score) [25] for the genetic association model. Finally, it identifies the central traits associated with the genetic variant as the union of the driver and optimal traits.

This analysis tested each SNP separately for its simultaneous associations with all 47 phenotypes, 18 chemotherapies, and 29 nonchemotherapies drugs. A multivariate approach had the dual advantages of achieving data reduction and increasing statistical power compared with running 47 separate univariates EWAS. Since we did not have a replication cohort, we performed a statistical experiment to validate the true nature of the identified association in multivariate EWAS by estimating the chance of observing such an association in a random dataset. For this, we assigned random patient identity to each row of the phenotype file while keeping the correlation between the drug response the same and performed the univariate and multivariate association analysis with the genotypes using the same parameters.

Results

Single SNP association with the ex-vivo drug response identified association of 4 locus at exome-wide significance level

After stringent quality control, we analyzed a total of 55,423 genetic markers, including all exonic variants that passed the quality control, for their association with single-drug response in 175 AML patients using linear regression adjusting for disease stage, cell culture medium, and the first 5 genetic principal components. We detected an exome-wide significance level association for rs113985677 in *CCIN* with tamoxifen ($p = 3.68 \times 10^{-7}$), and rs115400838 in *TRMT5* with the idelalisib response ($p = 4.20 \times 10^{-7}$). ClinVar reports rs115400838 as Benign for TRMT5-related disorder. We also observed an EWAS level significance for rs11878277 in *HDGFL2* with entinostat ($p = 4.94 \times 10^{-7}$), and rs2229092 in *LTA* with vorinostat response ($p = 8.28 \times 10^{-7}$) (Table 1).

Table 1 SNPs showing association with ex-vivo drug response at exome-wide significance levels in the analysis

CHR	POS	SNP	Position	Shift/Polyphen 2 predictions	Gene	A1/A2	N	MAF	Effect size (SE)	P	Phenotype	Mechanism	FDA Approval status
9	36170290	rs113985677	Missense R (Arg) > Q/p (Gln/Pro)	Tolerated/NA	CCIN	A/G	175	0.015	4.45 (0.84)	3.68×10^{-7}	Tamoxifen	Selective estrogen receptor modulator	Approved
14	61446147	rs115400838	Missense R (Ser) > P (Cys)	Deleterious/NA	TRMT5	A/T	175	0.002	7.58 (1.44)	4.20×10^{-7}	Idelalisib	PI3K inhibitor	Approved
19	4498000	rs11878277	Synonymous S (Ser) > S (Ser)	Tolerated/NA	HDGFL2	A/G	175	0.020	10.23 (1.96)	4.94×10^{-7}	Entinostat	HDAC inhibitor	Phase 2
6	31572980	rs2229092	Missense H (His) > P (Pro)	Tolerated/Benign	LTA	C/A	173	0.060	-4.9 (1)	8.28×10^{-7}	Vorinostat	HDAC inhibitor	Approved

Chromosomal positions of SNPs are based on National Center for Biotechnology Information genome build 38. The effect size was calculated with respect to the minor alleles. The effect predictions for each SNPs were obtained from Shift (<https://sift.bii.a-star.edu.sg/>) and Polyphen 2 (<http://genetics.bwh.harvard.edu/pph2/>) websites. The association results presented were obtained from genotyped data in 175 patients. SE: standard error

The distribution of the association p-value across different chromosomes for important traits have been shown in Fig. 1. A good agreement was observed between the theoretical p-value distribution and calculated p-values for these drugs as shown in Supplementary Fig. 2. The genomic inflation factor (λ) for the fitted model ranges between 0.98 and 1.01 which indicates the homogeneity of analyzed samples.

Further, adjusting the association analysis with blast percentages and cancer cell differentiation stage as additional covariates along with disease stage, cell culture medium, and the first 5 genetic principal components in the 101 patient samples identified 3 additional variants at exome-wide significance level. We observed exome-wide significance for rs701564 in major histocompatibility complex, class II, DQ beta 1 (*HLA-DQB1*)

with luminespib ($P=1.72 \times 10^{-7}$), rs2306234 in BLK Proto-Oncogene (BLK) with sepantronium bromide ($P=3.27 \times 10^{-7}$), and rs35094336 in Charged Multivesicular Body Protein 4 C (*CHMPC4*) with fludarabine ($P=8.95 \times 10^{-7}$) as shown in Supplementary Table 3.

Suggestive association in exome-wide association analysis

Further, we observed associations of SNPs with the response of multiple drugs with a similar mechanism of action at the sub-EWAS level of significance ($P < 10^{-4}$) suggesting a true nature of the association. For example, we observed the association of SNPs in the same genes with drugs having similar mechanisms of action. For example, rs2394516 and rs2394516 in *OR2J1* (olfactory receptor 2J1) were associated with both cladribine (purine analog) and clofarabine (another purine analog)

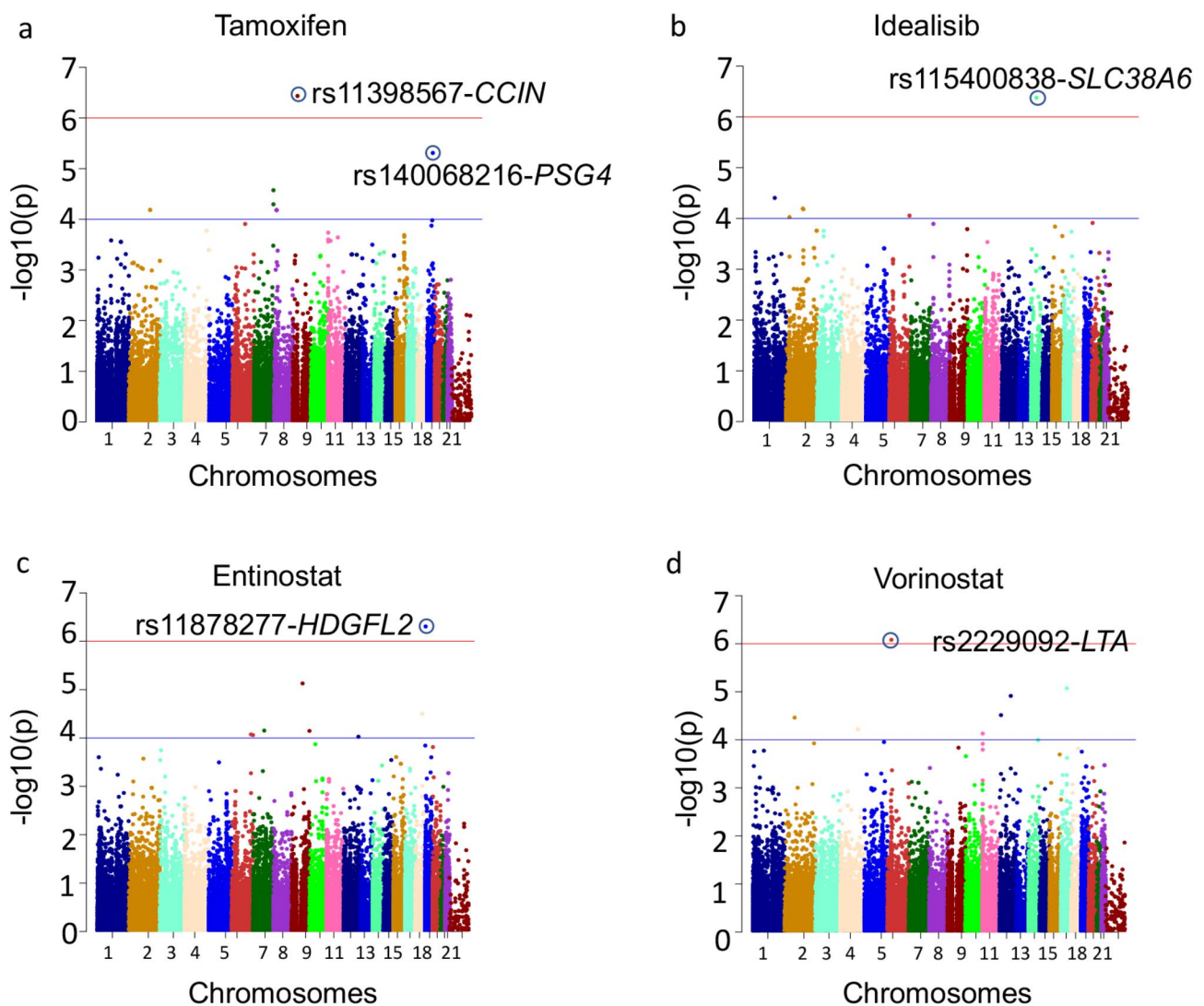


Fig. 1 Manhattan plot depicting the $-\log_{10}(p)$ values for SNPs associated with (a) tamoxifen (b) idelalisib, (c) entinostat, and (d) vorinostat across different chromosomes. The p-value has been obtained using linear regression adjusted for disease stage (diagnosis/relapse), cell culture medium and first 5 principal components from genotype data.

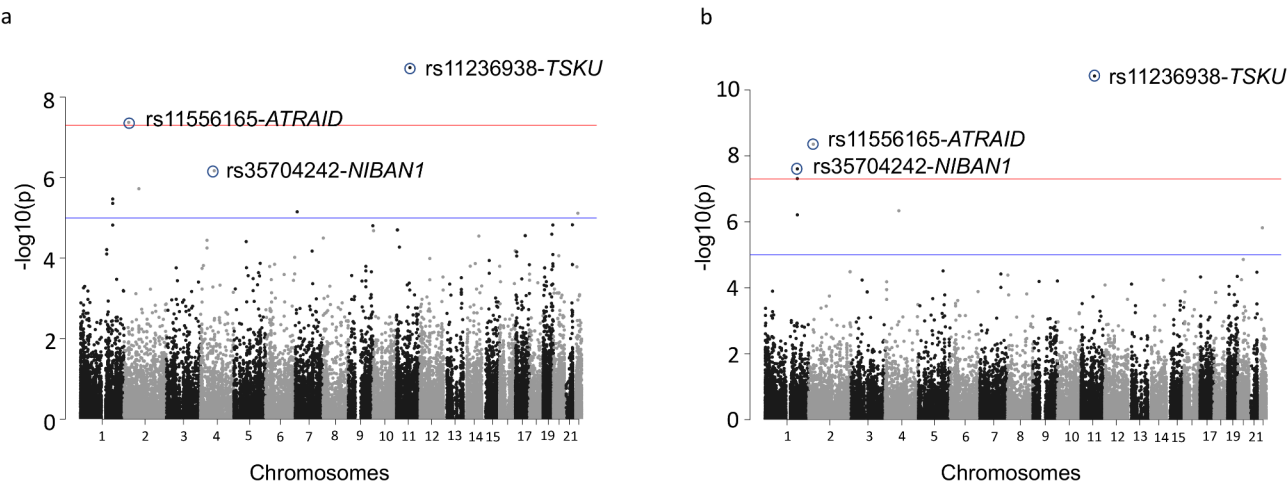


Fig. 2 Multivariate association analysis of exonic SNPs with the linear combination of drug response **(a)** Manhattan plot depicting the $-\log(10)$ p-values for SNPs associated with all 47 drugs across different chromosomes in multivariate analysis. **(b)** Manhattan plot depicting the $-\log(10)$ p-values for SNPs associated with only kinase inhibitors

Table 2 SNPs showing association with ex-vivo response of nonchemotherapy drugs as group at geome-wide significance levels in the multivariate analysis

CHR	BP	Variant (Gene)	Allele	Shift/Poly-phen 2	MAF	Pval	Central traits	Drug group
11	76507403	rs11236938-G/A (<i>TSKU</i>)	S (Ser) > N (Asn)	Tolerated/benign	0.03	3.8×10^{-11}	Tamoxifen, BI.2536, Belinostat	Non chemotherapy
2	27438615	rs11556165-G/C (<i>ATRAID</i>)	D (Asp) > H (His)	Tolerated/NA	0.04	3.98×10^{-9}	Belinostat, Fingolimod, Alvocidib	Non chemotherapy
1	184764824	rs35704242-C/T (<i>NIBAN1</i>)	D (Asp) > N (Asn)	Tolerated/NA	0.04	2.51×10^{-8}	BI.2536, Gefitinib, Belinostat	Non chemotherapy

suggesting a similar mechanism for influencing the drug response by identified variants (Supplementary Table 3).

Similarly, associations of rs201274224 in *PIK3CD* were observed with daunorubicin (topoisomerase II inhibitor), idarubicin (topoisomerase II inhibitor), and fludarabine (a purine analog) response suggesting a common mechanism of action probably through DNA damage and repair pathways. In total, we observed 295 suggestive associations within 248 unique SNP associated with the 47-drug response (Supplementary Fig. 3, Supplementary Table 4).

Multi-variate association analysis identified an association of 4 additional SNPs at the exome-wide significance level

In the multivariate analysis, first, we checked the association of 55,423 SNPs with a combined response of 47 drugs (Supplementary Fig. 4). We observed an EWAS level association for rs11236938-G/A (*TSKU*) with the combined response ($p = 2.51 \times 10^{-9}$, Fig. 2a, Supplementary Table 5) of tamoxifen, BI.2536, and belinostat as the central traits. Further, the analysis also identified an association of rs11556165-*ATRAID* ($p = 4.26 \times 10^{-8}$, Supplementary Table 5) with the combined response of belinostat, fingolimod, and alvocidib as the central traits at the exome-wide significance level.

Next, we also performed multivariate analysis to check the association of SNPs with only a group of drugs with specific mechanisms of action (e.g., kinase inhibitors). The analysis of response from the 29 non-chemotherapy drugs identified an association of rs11236938-G/A in *TSKU* with combined drug response of tamoxifen, BI.2536, and belinostat ($p = 3.8 \times 10^{-11}$, Fig. 2b; Table 2) as the central trait. Further, rs11556165-*ATRAID* was also associated with EWAS significance level ($p = 3.98 \times 10^{-9}$) with the combined response of non-chemo drugs with belinostat, fingolimod, alvocidib as the major trait. Similarly, rs35704242-C/T in *NIBAN1* was associated with the combined nonchemotherapy drug response and, BI.2536, gefitinib, and belinostat as the main traits ($p = 2.51 \times 10^{-8}$). The observed inflation factor for the p-values were 1.3 for all drug analysis and 1.17 for only nonchemotherapy drug response analysis (Supplementary Fig. 5).

Validation of the identified association To check the validity of the association and significance level of the p-value, we also performed an association analysis by creating 200 sets of randomized phenotypes (200×47 phenotypes) where we kept the correlation between responses of the 47 drugs the same. The univariate association analysis

did not identify any signals at the exome-wide significance level at any of the random phenotypes.

However, only 8 out of 200 times any exome-wide significant association was observed when the responses of all 47 drugs were analyzed in the multivariate analysis. Similarly, only 5 EWAS-level significant associations were observed out of 200 times when only the response of 29 non-chemo drugs were analyzed together (Supplementary Table 6). These analyses suggest the true nature of identified associations and revealed a maximum of 4% chance of the identified signals as false positive.

Discussion

Genetic risk factors affecting the response of most cancer drugs are poorly defined [26, 27] in AML. Most of the large-scale genomics study in AML have explored the effect of somatic mutations on drug response. Although specific somatic mutations have been associated with the drug response for some of the targeted therapy (e.g. FLT3 mutation for midostaurin response), most patients lack such clear somatic drug response markers [28–30] indicating a need to also study the role of multiple germline variants affecting drug response. However, there are few studies that have explored the influence of germline variant on drug response except for a few drugs in AML [11, 12, 31]. Further, these studies exploring the effect of the germline variants on drug response have only analyzed the polymorphisms in genes encoding transporters, metabolizers, or molecular targets of chemotherapy agents neglecting most parts of the genomes [14, 31]. In AML, large exome-wide association studies to identify pharmacogenetic variants affecting drug response is lacking because it requires a large amount of time, network and resources to collect the clinical response data. For example, follow-up period for clinical studies in AML patients could last up to 107 months which is a longtime for an individual laboratory to afford [16]. Our study, which comprised 175 AML cases profiled for an ex-vivo response of 47 drugs, represents one of the important studies performed to date.

We identified 4 EWAS significant variants affecting the ex-vivo drug response in our cohort overall using the univariate association analysis. However, we could not validate these signals because of the lack of other cohorts having this kind of data. We identified the association of rs113985677, a missense variant (changes arginine to glutamine or proline) in *CCIN* (calcin) with tamoxifen ex-vivo response. Calcin is a cytoskeleton protein and has been associated with lateral sinus thrombosis, spermatogenic Failure 9, globozoospermia [32, 33] and is required for the formation of the sperm head and maintain the nuclear structure in mice [34]. Downregulation of the cytoskeleton genes, connexin 43, induces tamoxifen-resistance in breast cancer [35] cells via activation of

c-Src/PI3K/Akt signaling and Epithelial to mesenchymal transition (EMT) of cells. The identified variant in the calcin gene may alter the tamoxifen response by altering the cytoskeleton dynamics during EMT.

Further, the variant is an expression quantitative trait locus for *GLIPR2* gene in the esophagus and salivary gland ($P=6.5 \times 10^{-12}$, and 7.9×10^{-8}) where the presence of alternate allele is associated with increased expression of *GLIPR2*. *GLIPR2* is particularly expressed in leukocytes and has a role in EMT transition [36], and autophagy [37], suggesting a possible role in drug response via these mechanisms. However, a functional study will be needed to understand the exact role of genes and variants.

We also found a significant association between rs115400838, an upstream variant in *SLC38A6*, and a missense variant in *TRMT5* with an idelalisib response. *SLC38A6* is a glutamine and glutamate transporter [38] and has been associated with angiotensin-converting enzyme inhibitor-induced cough [39]. The available evidence suggests that *SLC38A6* may be a direct transporter of idelalisib itself or its substrates. We also observed a significant association of rs11878277 (*HDGFL2*) with entinostat in the univariate analysis. Entinostat is a histone deacetylase inhibitor and is under clinical trial for multiple cancers. *HDGFL2* (Hepatoma-Derived Growth Factor-Related Protein 2) is dispensable for Mixed Lineage Leukemia (MLL)-rearranged (MLL-R) leukemogenesis [40] and regulates gene expression via HRP2-DPF3a-BAF complex formation [41].

Similarly, a significant association was observed for a missense variant rs2229092 (H (His) > P (Pro) in Lymphotoxin Alpha (*LTA*) gene with vorinostat response, an HDAC inhibitor. The *LTA* SNP rs2229092 lies at 6p21.33 within HLA region, however we did not observe any linked-SNP in the same LD region in the Finnish population. *LTA*, formerly known as tumor necrosis factor-beta (TNF- β) is a toxin produced from lymphocytes (e.g., T and B) that can kill other cell types. It acts as a causative factor for autocrine and paracrine activation of canonical and noncanonical NF- κ B signaling as well as promotes JAK2/STAT6 signaling in Hodgkin lymphoma [42] resulting in prevention of apoptosis. Furthermore, *LTA* shapes the expression of lymphoid and myeloid-specific genes in blood. Given the wide role of *LTA* in immunity and tumor survival [43], *LTA* may influence vorinostat response by modulating cancer survival-related gene expression in synergy with vorinostat.

We also identified the association of rs701564 in *HLA-DQB1* with luminespib ($P=1.72 \times 10^{-7}$), rs2306234 in *BLK* with sepantronium bromide ($P=3.27 \times 10^{-7}$), and rs35094336 in *CHMPC4* with fludarabine ($P=8.95 \times 10^{-7}$). However, the decreased sample size also decreased the power of the study increasing the chances

of false association. Hence, these associations need validations in other cohorts with larger sample size.

The multivariate analysis identified rs11556165-*ATRAID* as one of the significant SNPs at the EWAS level with belinostat, fingolimod, alvocidib responses. *ATRAID* (all-trans retinoic acid-induced differentiation factor), is a genetic target of nitrogen-containing-bisphosphonates (N-BP), a commonly used drug to treat osteoporosis and other bone-related diseases [44]. Additionally, *ATRAID* binds to SLC37A3 (solute carrier family 37 member A3), a lysosomal membrane protein with 12 transmembrane segments, facilitating NBP transport [45]. We also identified the association of rs35704242 in Niban apoptosis regulator 1 (*NIBAN1*) multi-drug response with BI 2536, gefitinib, and belinostat as the central trait set. *NIBAN1* protein is highly expressed in multiple cancers such as renal, colorectal, and thyroid cancer [46, 47] and may play an important role in cell survival during cellular stress such as ultraviolet irradiation cell migration and proliferation [46], cell autophagy [48] and immune microenvironment [49].

We also observed a significant association of tsukushi (*TSKU*), a small leucine-rich proteoglycan, with overall kinase ($p = 2.51 \times 10^{-8}$) with tamoxifen, BI.2536, belinostat as the central traits. Tsukushin is an extracellular matrix protein and is also a hepatokine. *TSKU* is involved in different signaling pathways including the BMP pathway, notch pathway, transforming growth factor- β (TGF- β) pathway, mitogen-activated protein kinase (MAPK)/fibroblast growth factor (FGF) pathway, and β -catenin/Wnt pathway [50]. The absence of *TSKU* increases cleaved poly ADP-ribose polymerase (PARP) protein levels and caspase 3/7 activity thereby increasing carcinogenicity in neuroblastoma cells [51]. Further, *TSKU* expression is strongly associated with poor overall survival in lung cancer [52]. Given its role in carcinogenicity and tumor survival, *TSKU* may affect drug response in cancer. However, further functional study is needed to reveal the exact mechanism by which *TSKU* affects drug response in AML.

In summary, our exome-wide analysis identified genetic loci in *CCIN* with tamoxifen, *TRMT5* with idelalisib, *HDGFL2* with entinostat, and *LTA* with vorinostat responses. Similarly, the multi-trait association analysis identified association of variants in *TSKU* with combined response of tamoxifen, BI.2536, belinostat; *ATRAID* with combined response of belinostat, fingolimod, alvocidib, *NIBAN1* with combined response of belinostat, fingolimod, alvocidib. These loci are near genes with potential role in the various steps of drug response mechanisms such as drug transport (e.g. rs115400838 near *SLC38A6* gene, rs11556165 near *ATRAID*), drug target mechanism such as apoptosis pathways (e.g. rs2229092 in *LTA*), and drug resistance (e.g. *CCIN* with with tamoxifen)

indicating a larger role of germline variants on overall antiAML drug response by affecting underlying drug-related process. However, since we used growth inhibition in patient derived primary cells as a measure for drug response, the measured response may not directly correlate to in-vivo or patient response as the model misses some of the important drug-related process such as hepatic metabolism, and excretion present in the in-vivo model and actual patients.

Further, our study does not include all the drugs approved for AML treatment (e.g. venetoclax) and a large portion of the studied drugs are chemotherapy. Hence, it misses some of the important targeted therapy approved for AML as well as other cancers. Secondly, we also note that our study uses the skin tissue for genotyping and leukemic cells from bone marrow/peripheral blood tissue for drug sensitivity testing. We used skin DNA for genotyping because it is the most common source of germline DNA for hematological cancer genetic studies. However, skin can also acquire genetic changes with passing of age and under ultraviolet radiation exposures [53]. Hence, there is a weak possibility for spurious association of acquired genetic changes in the skin tissue with drug response.

Our results provide a basis for initial exploration of the role of genes associated with this locus. However, replication studies to confirm the observed associations are needed in larger patient group. However, the lack of another cohort for validation is a major limitation of the study. Though being a novel and significantly large cohort for a rare disease like AML, our study is still smaller than genome-wide association studies reported in other malignancies and is potentially limited in its power to detect associations. Additionally, considering the increasing evidence that AML tumors are genetically heterogeneous, genetic predisposing factors likely vary across patients. Although we attempted to address this potential weakness by performing correlation-based MetaPhat analysis, these analyses are based on drugs with different targets and mechanisms of action, resulting in the potential to miss associations due to the lack of statistical power.

Abbreviations

AML	Acute myeloid leukemia
TSKU	Tsukushi
NIBAN1	Niban Apoptosis Regulator 1

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12920-025-02130-7>.

Supplementary Material 1

Acknowledgements

We are grateful to Prof Mark J Daly and Prof Matti Pirinen for valuable discussions regarding the possible approaches to validate these findings.

Author contributions

AKG conceived the project. AKG, JL, KK, GT, HA implemented and tested the method. AKG and JL wrote the manuscript. All authors provided critical feedbacks and important contributions to the final manuscript.

Funding

Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital).

AKG has received grant support from the Academy of Finland (Grant No: 359690), Mary and Georg C. Ehrnrooth Foundation (Grant No: 202210066), Otto A. Malm Foundation (Grant No: 16-16116-11), Maud Kuistila Memorial Foundation (Grant No: 2021-0270 C), K. Albin Johanssons stiftelse sr Foundation, and Finnish Foundation for the Advancement of Laboratory Medicine. Jake Lin has received support from the academy of Finland (Grant No: 325999).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

The study has been approved by the Helsinki University Hospital Ethical Committee (No. 239/13/03/00/2010, 303/13/03/01/2011, latest Helsinki University Hospital study permit HUS/395/2018 dated February 13, 2018). A signed informed consent with protocols in accordance with the Declaration of Helsinki (study acronym HRUHLAB2) has been obtained for all the study participant before their enrollment under the study.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 10 January 2025 / Accepted: 20 March 2025

Published online: 04 April 2025

References

1. Cancer Genome Atlas, Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, et al. Genomic and epigenomic landscapes of adult de Novo acute myeloid leukemia. *N Engl J Med*. 2013;368(22):2059–74.
2. Quek L, David MD, Kennedy A, Metzner M, Amatangelo M, Shih A, et al. Clonal heterogeneity of acute myeloid leukemia treated with the IDH2 inhibitor enasidenib. *Nat Med*. 2018;24(8):1167–77.
3. lanevski A, Lahtela J, Javarappa KK, Sergeev P, Ghimire BR, Gautam P, et al. Patient-tailored design for selective co-inhibition of leukemic cell subpopulations. *Sci Adv*. 2021;7(8):eabe4038.
4. Giri AK, lanevski A, Aittokallio T. Genome-wide off-targets of drugs: risks and opportunities. *Cell Biol Toxicol*. 2019;35:485–7.
5. Daver N, Garcia-Manero G, Basu S, Boddu PC, Alfayez M, Cortes JE, et al. Efficacy, safety, and biomarkers of response to Azacitidine and nivolumab in relapsed/refractory acute myeloid leukemia: A nonrandomized, Open-Label, phase II study. *Cancer Discov*. 2019;9:370–83.
6. Giri AK, lanevski A. High-throughput screening for drug discovery targeting the cancer cell-microenvironment interactions in hematological cancers. *Expert Opin Drug Discov*. 2021;17:181–90.
7. Schroeder T, Rautenberg C, Kruger W, Platzbecker U, Bug G, Steinmann J, et al. Treatment of relapsed AML and MDS after allogeneic stem cell transplantation with decitabine and DLI-a retrospective multicenter analysis on behalf of the German cooperative transplant study group. *Ann Hematol*. 2018;97:335–42.
8. Bottomly D. Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. *Cancer Cell*. 2022;40:850–e864859. Long N, Schultz AR, Kurtz SE, Tognon CE, Johnson K et al.
9. Malani D. Implementing a functional precision medicine tumor board for acute myeloid leukemia. *Cancer Discov*. 2021;12:388–401. Kumar A, Brück O, Kontro M, Yadav B, Hellesøy M et al.
10. Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature*. 2018;562:526–31.
11. Hartford CM, Duan S, Delaney S, Mi M, Kistner S, Lamba EO et al. J. K. Population-specific genetic variants important in susceptibility to cytarabine arabinoside cytotoxicity. *Blood*. 2009;113:2145–53.
12. Gamazon ER, Lamba JK, Pounds S, Stark AL, Wheeler HE, Cao X, et al. Comprehensive genetic analysis of cytarabine sensitivity in a cell-based model identifies polymorphisms associated with outcome in AML patients. *Blood*. 2013;121:4366–76.
13. Lamba JK. Genetic factors influencing cytarabine therapy. *Pharmacogenomics*. 2009;10:1657–74.
14. Zhu KW, Chen P, Zhang DY, Yan H, Liu H, Cen LN, et al. Association of genetic polymorphisms in genes involved in Ara-C and dNTP metabolism pathway with chemosensitivity and prognosis of adult acute myeloid leukemia (AML). *J Transl Med*. 2018;16:90.
15. Williams MS, Amaral FM, Simeoni F, Somerville TC. A stress-responsive enhancer induces dynamic drug resistance in acute myeloid leukemia. *J Clin Invest*. 2020;130:1217–32.
16. Tiong IS, Wall M, Bajel A, Kalro A, Fleming S, Roberts AW, et al. How comparable are patient outcomes in the real-world with populations studied in pivotal AML trials? *Blood Cancer J*. 2024;14:54.
17. lanevski A, Giri AK, Gautam P, et al. Prediction of drug combination effects with a minimal set of experiments. *Nat Mach Intell*. 2019;1:568–77.
18. lanevski A, Giri AK, Aittokallio T. SynergyFinder 3.0: an interactive analysis and consensus interpretation of multi-drug synergies across multiple samples. *Nucleic Acids Res*. 2022;50:W739–43.
19. lanevski A, Giri AK, Aittokallio T. SynergyFinder 2.0: visual analytics of multi-drug combination synergies. *Nucleic Acids Res*. 2020;48:W488–93.
20. Wahnschaffe L, Braun T, Timonen S, Giri AK, Schrader A, Wagle P et al. JAK/STAT-Activating genomic alterations are a hallmark of T-PLL. *Cancers (Basel)*. 2019;11(12).
21. Lin J, Tabassum R, Ripatti S, Pirinen M. MetaPhat: detecting and decomposing multivariate associations from univariate Genome-Wide association statistics. *Front Genet*. 2020;11:431.
22. lanevski A, Timonen S, Kononov A, Aittokallio T, Giri AK. SynToxProfiler: an interactive analysis of drug combination synergy, toxicity and efficacy. *PLoS Comput Biol*. 2020;16(2):e1007604.
23. Yadav B, Pemovska T, Sz wajda A, Kuleskiy E, Kontro M, Karjalainen R, et al. Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies. *Sci Rep*. 2014;4:5193.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–75.
25. Schwarz G. Estimating the dimension of a model. *Ann Statist*. 1978;6(2):461–4.
26. Menden MP, Casale FP, Stephan J, Bignell GR, Iorio F, McDermott U, et al. The germline genetic component of drug sensitivity in cancer cell lines. *Nat Commun*. 2018;9(1):3385.
27. Hariprakash JM, Vellarikkal SK, Keechilat P, Verma A, Jayarajan R, Dixit V, et al. Pharmacogenetic landscape of DPYD variants in South Asian populations by integration of genome-scale data. *Pharmacogenomics*. 2018;19(3):227–41.
28. Stone RM, Mandrekar SJ, Sanford BL, Laumann K, Geyer S, Bloomfield CD, et al. Midostaurin plus chemotherapy for acute myeloid leukemia with a FLT3 mutation. *N Engl J Med*. 2017;377:454–64.
29. Little P, Lin DY, Sun W. Associating somatic mutations to clinical outcomes: a pan-cancer study of survival time. *Genome Med*. 2019;11(1):37.
30. Levatic J, Salvadores M, Fuster-Tormo F, Supek F. Mutational signatures are markers of drug sensitivity of cancer cells. *Nat Commun*. 2022;13:2926.
31. Huang RS, Duan S, Kistner EO, Bleibel WK, Delaney SM, Fackenthal DL, et al. Genetic variants contributing to daunorubicin-induced cytotoxicity. *Cancer Res*. 2008;68(9):3161–8.
32. Celse T, Cazin C, Miettinen F, Martinez G, Martinez D, Thierry-Mieg N, et al. Genetic analyses of a large cohort of infertile patients with globozoospermia, DPY19L2 still the main actor, GGN confirmed as a guest player. *Hum Genet*. 2021;140(1):43–57.
33. Li YZ, Wu RF, Zhu XS, Liu WS, Ye YY, Lu ZX, et al. Identification of a novel deletion mutation in DPY19L2 from an infertile patient with globozoospermia: a case report. *Mol Cytogenet*. 2020;13:24.
34. Zhang XZ, Wei LL, Jin HJ, Zhang XH, Chen SR. The perinuclear Theca protein Calicin helps shape the sperm head and maintain the nuclear structure in mice. *Cell Rep*. 2022;40:111049.

35. Wu DP, Zhou Y, Hou LX, Zhu XX, Yi W, Yang SM, et al. Cx43 deficiency confers EMT-mediated Tamoxifen resistance to breast cancer via c-Src/PI3K/Akt pathway. *Int J Biol Sci*. 2021;17:2380–98.
36. Huang S. GLIPR-2 overexpression in HK-2 cells promotes cell EMT and migration through ERK1/2 activation. *PLoS ONE*. 2013;8:e58574. Liu F, Niu Q, Li Y, Liu C, Zhang L, et al.
37. Aalberts M, van Dissel-Emiliani FM, van Adrichem NP, van Wijnen M, Wauben MH, Stout TA, et al. Identification of distinct populations of prostasomes that differentially express prostate stem cell antigen, Annexin A1, and GLIPR2 in humans. *Biol Reprod*. 2012;86:82.
38. Gandasi NR, Arapi V, Mickael ME, Belekar PA, Granlund L, Kothegala L et al. Glutamine uptake via SNAT6 and Caveolin regulates Glutamine-Glutamate cycle. *Int J Mol Sci*. 2021;22(3).
39. Hallberg P, Persson M, Axelsson T, Cavalli M, Norling P, Johansson HE, et al. Genetic variants associated with angiotensin-converting enzyme inhibitor-induced cough: a genome-wide association study in a Swedish population. *Pharmacogenomics*. 2017;18(3):201–13.
40. Van Belle S, El Ashkar S, Cermakova K, Matthijssens F, Goossens S, Canella A et al. Unlike its paralog LEDGF/p75, HRP-2 is dispensable for MLL-R leukemogenesis but important for leukemic cell survival. *Cells* 2021;10(1).
41. Zhu X, Lan B, Yi X, He C, Dang L, Zhou X, et al. HRP2-DPF3a-BAF complex coordinates histone modification and chromatin remodeling to regulate myogenic gene transcription. *Nucleic Acids Res*. 2020;48(12):6563–82.
42. von Hoff L, Kärger E, Franke V, McShane E, Schulz-Beiss KW, Patone G, et al. Autocrine LTA signaling drives NF- κ B and JAK-STAT activity and myeloid gene expression in hodgkin lymphoma. *Blood*. 2019;133(13):1489–94.
43. Fhu CW, Graham AM, Yap CT, Al-Salam S, Castella A, Chong SM, et al. Reed-Sternberg cell-derived lymphotoxin- α activates endothelial cells to enhance T-cell recruitment in classical hodgkin lymphoma. *Blood*. 2014;124:2973–82.
44. Yu Z, Surface LE, Park CY, Horlbeck MA, Wyant GA, Abu-Remaih M, et al. Identification of a transporter complex responsible for the cytosolic entry of nitrogen-containing bisphosphonates. *Elife*. 2018;10:7.
45. Surface LE, Burrow DT, Li J, Park J, Kumar S, Lyu C, et al. *ATRAID* regulates the action of nitrogen-containing bisphosphonates on bone. *Sci Transl Med*. 2020;12(544):eaav9166.
46. Feng X, et al. miR-4521-FAM129A axial regulation on CcRCC progression through TIMP-1/MMP2/MMP9 and MDM2/p53/Bcl2/Bax pathways. *Cell Death Discov*. 2019;5:89.
47. Diana P, Carvalheira GMG. NIBAN1, exploring its roles in cell survival under stress context. *Front Cell Dev Biol*. 2022;10:867003.
48. Ji H, et al. AKT-dependent phosphorylation of Niban regulates nucleophosmin- and MDM2-mediated p53 stability and cell apoptosis. *EMBO Rep*. 2012;13:554–60.
49. Zhang R, et al. Tissue Treg secretomes and transcription factors shared with stem cells contribute to a Treg niche to maintain Treg-Ness with 80% innate immune pathways, and functions of immunosuppression and tissue repair. *Front Immunol*. 2020;11:632239.
50. Deng X, Li Y, Guo C, Zhao Z, Yuan G. Novel roles of Tsukushi in signaling pathways and multiple disease processes. *BioFactors*. 2021;47:512–21.
51. Ichikawa T, Horie-Inoue K, Ikeda K, Blumberg B, Inoue S. Steroid and xenobiotic receptor SXR mediates vitamin K2-activated transcription of extracellular matrix-related genes and collagen accumulation in osteoblastic cells. *J Biol Chem*. 2006;281:16927–34.
52. Huang H, Zhang D, Fu J, Zhao L, Li D, Sun H, et al. Tsukushi is a novel prognostic biomarker and correlates with tumor-infiltrating B cells in non-small cell lung cancer. *Aging*. 2021;13(3):4428–51.
53. King C, Fowler JC, Abnizova I, Sood RK, Hall MWJ, Szeverenyi I, et al. Somatic mutations in facial skin from countries of contrasting skin cancer risk. *Nat Genet*. 2023;55:1440–7.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.