Karol, S., Larsen, E., Cheng, C. et al. *Genetics of ancestry-specific risk for relapse in acute lymphoblastic leukemia*. **Leukemia** 31, 1325–1332 (2017).

https://doi.org/10.1038/leu.2017.24

Genetics of ancestry-specific risk for relapse in acute lymphoblastic leukemia

Seth E. Karol, MD,
Comprehensive Cancer Center, St. Jude Children's Research Hospital, Memphis, TN, USA

Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA

Eric Larsen, MD,
Department of Pediatrics, Main Medical Center, Portland, ME, USA

Cheng Cheng, PhD,
Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, USA

Xueyuan Cao, PhD,
Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, USA

Wenjian Yang, PhD,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Laura B. Ramsey, PhD,
Department of Pharmacy Research, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

Christian A. Fernandez, PhD,
Department of Pharmaceutical Sciences, University of Pittsburgh, Pittsburgh, PA, USA

Joseph R. McCorkle, PhD,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Steven W. Paugh, PhD,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Robert J. Autry, MS,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Elixabet Lopez-Lopez, PhD,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USADepartment of Pediatrics, Main Medical Center, Portland, ME, USA

Barthelemy Diouf, PhD,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Sima Jeha, MD,
Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA

Ching-Hon Pui, MD,
Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA

Elizabeth A. Raetz, MD,
Department of Pediatrics, University of Utah, Salt Lake City, UT, USA

Naomi J. Winick, MD,
Department of Pediatrics, University of Texas, Southwestern Medical Center, Dallas, TX, USA

William L. Carroll, MD,
Perlmutter Cancer Center, Department of Pediatrics and Pathology, New York University Langone Medical Center, New York, NY, USA

Stephen P. Hunger, MD,
Department of Pediatrics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Mignon L. Loh, MD,
Department of Pediatrics, University of California School of Medicine, San Francisco, CA, USA

Meenakshi Devidas, PhD,
Department of Biostatistics, Colleges of Medicine, Public Health & Health Professions, University of Florida, Gainesville, FL, USA

William E. Evans, PharmD,
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Jun J. Yang, PhD, and
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Mary V. Relling, PharmD
Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA

Abstract

The causes of individual relapses in children with acute lymphoblastic leukemia (ALL) remain incompletely understood. We evaluated the contribution of germline genetic factors to relapse in 2,225 children treated on Children's Oncology Group trial AALL0232. We identified 302 germline single nucleotide polymorphisms (SNPs) associated with relapse after adjusting for treatment and ancestry and 715 additional SNPs associated with relapse in an ancestry-specific manner. We tested for replication of these relapse-associated SNPs in external data sets of antileukemic drug pharmacokinetics and pharmacodynamics and an independent clinical cohort. 224 SNPs were associated with rapid drug clearance or drug resistance, and 32 were replicated in the independent cohort. The adverse risk associated with black and Hispanic ancestries was attenuated by addition of the 4 SNPs most strongly associated with relapse in these populations [for blacks: model without SNPs hazard ratio (HR) =2.32, $P=2.27\times10^{-4}$, model with SNPs HR=1.07, P=0.79; for Hispanics: model without SNPs HR=1.7, $P=8.23\times10^{-5}$, model with SNPs HR=1.31, P=0.065]. Relapse SNPs associated with asparaginase resistance or allergy were overrepresented among SNPs associated with relapse in the more asparaginase intensive treatment arm (20/54 in Capizzi-methorexate arm vs. 8/54 in high-dose methotrexate arm, P=0.015). Inherited genetic variation contributes to race-specific and treatment-specific relapse risk.

Introduction

Despite significant improvements in outcome,(1–3) relapse remains the leading cause of treatment failure for children with acute lymphoblastic leukemia (ALL) and occurred in 11 to 36% of those with high-risk B-precursor ALL.(4–10) Mechanisms by which genomic variation influence relapse risk could involve somatically acquired mutations or inherited genetic variations, which could affect intrinsic resistance to chemotherapy(11–13) or host pharmacokinetics of anti-leukemic agents.(14–16)

Some studies report that black and Hispanic children with ALL have inferior outcomes to non-Hispanic white children.(17–21) Reasons for these differences are likely multifactorial, including differences in treatment adherence and access to therapy,(22–24) in the incidence of favorable and unfavorable presenting features and cytogenetics,(25–27) and in the frequency of genetic variants affecting pharmacokinetics and pharmacodynamics of antileukemic agents which segregate with ancestry.(28) It remains uncertain whether racial disparities persist with modern intensive ALL regimens.

We performed a genome wide association study (GWAS) in a large cohort of children with high-risk B-ALL to identify inherited genetic variations associated with relapse. We performed an analysis adjusting for both treatment and ancestry to identify single nucleotide polymorphisms (SNPs) which increased risk across ancestries (ancestry-agnostic SNPs). Because racial disparities in relapse persisted in this trial, we also performed analyses within each of the three largest ancestral groups (white, black, Hispanic) to identify ancestry-specific variations associated with relapse. We also interrogated relapse SNPs for associations with risk of central nervous system (CNS) relapse, relapse among patients randomized to receive either escalating-dose methotrexate and asparaginase (i.e., Capizzi regimen) or high-dose methotrexate during the first interim maintenance (IM1), and for

associations with the pharmacokinetics of antileukemic agents or the intrinsic sensitivity of leukemia cells to chemotherapy. Finally, to assess robustness of relapse SNPs across different therapies, we tested for replication in an independent cohort.

Methods

Patients and treatment

For the discovery cohort, germline DNA was obtained at remission in children and young adults with newly diagnosed B-precursor ALL enrolled on COG AALL0232 (NCT00075725, https://clinicaltrials.gov/ct2/show/NCT00075725).(8) This protocol involved a 2×2 factorial randomization for induction steroid (prednisone ×28 days vs. dexamethasone ×14 days) and interim maintenance 1 regimen (Capizzi escalating-dose methotrexate with pegylated-asparaginase vs. high-dose methotrexate). Exclusion criteria are described in Figure 1 and the Supplementary Methods. The replication cohort comprised children treated on prior generation protocols who would have met the eligibility criteria of AALL0232 (Supplementary Methods and Supplementary Table 1).

All studies were approved by the institutional review boards of participating institutions, and all patients and/or guardians provided age appropriate consent/assent in accordance with the Declaration of Helsinki.

Genotyping and genetic ancestry

Genotyping and genetic imputation was performed as described in the Supplementary Methods. Genetic ancestry was defined using STRUCTURE v2.2.3.(29) For categorization of patients into discrete ancestral groups, individuals were classified based on inferred genetic ancestry as white [Northern European (CEU) >90%], black [West African (YRI) >70%], Hispanic [Native American(30) >10% and Native American greater than West African], or Other, including Asian [East Asian (CHB/JPT) >90%].

Quality control steps for both patients and SNPs are detailed in the Supplementary Methods.

Identification of relapse associated SNPs

The approaches to perform GWASs for relapse are detailed in the Supplementary Methods. GWASs were performed to identify SNPs using an ancestry-agnostic (Supplementary Table 2 and Supplementary Figure 1) and an ancestry-specific approach (Supplementary Figures 2a–c).

Treatment arm and site specific annotation of relapse SNPs

SNPs associated with relapse were further characterized in subsets of patients based on their IM1 randomization (the Capizzi arm with escalating-dose methotrexate plus pegylated-asparaginase vs. the high-dose methotrexate arm) while adjusting for induction randomization, rapid early response, and ancestry as categorical variables. Additionally, SNPs were tested for their association with CNS relapse (isolated or combined with other sites), with isolated hematologic or other extramedullary relapse treated as competing risks. Significant association thresholds for all analyses were determined by profile information

criteria (Ip),(31) which balances false positives and negatives while addressing the effects of multiple testing.

Association with orthogonal pharmacologic data

SNPs associated with relapse (ancestry-specific or ancestry-agnostic) were evaluated for association with drug resistance in HapMap cells lines (prednisone, asparaginase, mercaptopurine, methotrexate polyglutamate accumulation), primary ALL cells from newly diagnosed patients (prednisone, vincristine, mercaptopurine, asparaginase, *in vivo* leukocyte count decrease following methotrexate), or for association with increased drug clearance (asparaginase allergy, methotrexate clearance, dexamethasone clearance), as described in the Supplementary Methods. SNPs were considered supported by orthogonal data if the risk allele for relapse was associated (at $P<0.05$) with *in vitro* drug resistance, decreased methotrexate polyglutamate accumulation, smaller leukocyte decrease after methotrexate, more rapid drug clearance, or greater incidence of asparaginase allergy.

Evaluation of relapse-associated SNPs in replication cohort

Relapse-associated SNPs were evaluated in an independent replication cohort (n=719) for their association with relapse using a Cox proportional hazard regression, with patients censored at the time of competing events (i.e. remission death, second malignancy) or last follow-up and adjusting for treatment categorized into 6 groups (Supplementary Table 2).(5, 9, 32) AALL0232 ancestry-agnostic SNPs were evaluated in all patients while adjusting for treatment and ancestry. AALL0232 ancestry-specific SNPs were evaluated in the same ancestry subset of the replication cohort while adjusting for treatment and, in blacks and Hispanics, percent ancestry. The replication cohort SNPs were evaluable if they passed quality control steps as described for the discovery cohort (Supplementary Methods). Differences in genotyping platforms between the discovery and replication cohorts, as well as the smaller size of the replication cohort, resulted in 595 of the 1,017 relapse SNPs from the discovery cohort being evaluable in the replication cohort. Validated SNPs were those associated with relapse at $P<0.05$ and with identical risk alleles.

Quantitative contribution of SNPs to ancestral differences in relapse

To identify SNPs which most contributed to ancestry-associated differences in relapse risk, a classification and regression tree analysis was performed separately in blacks and Hispanics considering treatment arm and validated ancestry-agnostic and ancestry-specific SNPs as potential branches. Branches were limited to two levels with each new branch needing to contain at least 20% of the initial ancestral patient group (representing ~1% or at least 22 patients from the discovery cohort for the smallest group, those with black ancestry). The impact of these SNPs on the risk of relapse associated with black or Hispanic ancestry was then evaluated in a competing risk regression model of relapse including the SNPs, treatment, and ancestry.

Statistical analysis

Statistical and bioinformatics analyses were performed using R versions 3.2.2, including the “survival”, “cmprsk”, “rpart”, and “forestplot” packages. Association studies of orthogonal phenotypes were performed either in R or PLINK version 1.07.

Results

Patient Characteristics

Of 3,084 children and young adults enrolled on AALL0232, germline genotype and relapse data were available for 2,652, and 2,225 were included in the GWAS for relapse (Figure 1). To identify covariates to include in the GWAS, we examined the importance of treatment group and ancestry on relapse risk. Consistent with findings in the entire randomized cohort, (8) patients treated with Capizzi-methotrexate had a higher relapse risk than those treated with high-dose methotrexate (Supplementary Table 3). Because patients with slow early response did not differ by their induction steroid assignment but did differ by IM1 randomization, patients with slow early response were combined for multivariable and GWAS analyses (Supplementary Table 2). Blacks [$P=2.66\times10^{-4}$, hazard ratio (HR)=2.31] and Hispanics ($P=2.17\times10^{-5}$, HR=1.77) had an increased relapse risk compared to whites (Supplementary Table 3). The effects of ancestry and treatment groups remained significant in multivariate analyses (Supplementary Table 3, Figure 2). Blacks and Hispanics also had a higher risk of any CNS relapse than whites (P=0.016, HR=2.54 for blacks; P=0.0018, HR=2.08 for Hispanics).

Association of SNPs with relapse

Following quality control steps, 11,180,806 SNPs were evaluated for their association with relapse. A total of 302 SNPs representing 175 unique genetic loci (LD blocks) were associated with relapse in an analysis adjusting for both treatment and percent genetic ancestry (i.e. their association with relapse was “agnostic” to ancestry; Supplementary Table 4, Supplementary Figure 4). An additional 715 SNPs representing 424 unique genetic loci were associated with relapse in ancestry-specific analyses, with 280 SNPs (179 loci) associated with relapse in Hispanics, 258 SNPs (167 loci) in blacks, 173 SNPs (72 loci) in whites, 2 SNPs (2 loci) in both blacks and whites, and 2 SNPs (1 locus) in both blacks and Hispanics (Supplementary Tables 5–7, Supplementary Figures 3, 5–7).

Of the 1,017 relapse SNPs, 192 were associated with relapse in patients treated on the Capizzi arm, 186 in patients treated on the high-dose methotrexate arm, and 18 in both treatment groups; 621 SNPs were not associated with relapse in either group alone but were associated with relapse in the combined cohort (Supplementary Tables 4–7, Supplementary Figures 4–7).

Of the 302 ancestry-agnostic SNPs, 54 were also associated with an increased risk of CNS relapse (Supplementary Table 4, Supplementary Figure 4). Of these, 25 were associated with increased CNS relapse in patients treated on the Capizzi arm, 14 in patients on the high-dose methotrexate arm, and 4 in patients on both arms. Because of the association between ancestry and CNS relapse risk, we evaluated ancestry-specific SNPs for their association

with CNS relapse and identified 18 SNPs associated with increased CNS relapse risk in whites, 38 SNPs in blacks, and 52 SNPs in Hispanics (Supplementary Tables 5–7, Supplementary Figures 5–7).

Because of the importance of minimal residual disease (MRD) in defining high-risk patients, (33) we also evaluated relapse SNPs for their adverse impact in the 1,931 patients with end of induction (day 29) MRD less than 0.1%. 617 SNPs remained significant at the previously defined significance threshold, including 209 ancestry-agnostic SNPs (Supplementary Tables 4–7, Supplementary Figures 4–7).

Association of relapse SNPs with orthogonal pharmacologic data

To explore possible mechanisms underlying the 1,017 SNPs associated with relapse, we tested for their association with orthogonal phenotypes including *in vitro* resistance to chemotherapy, decreased response to methotrexate *in vivo*, increased chemotherapeutic drug clearance *in vivo*, and asparaginase allergy *in vivo.* Of the 302 ancestry-agnostic SNPs, 54 were associated with one resistance/clearance phenotype and 10 were associated with more than one such phenotype (Supplementary Table 4). Of the 715 ancestry-specific SNPs, 128 were associated with one resistance/clearance phenotype and 32 with more than one phenotype (Supplementary Tables 5–7). 36 of the 162 relapse SNPs associated with CNS relapse were associated with at least one resistance/clearance phenotype.

Of the 54 relapse SNPs associated with intrinsic leukemic asparaginase resistance (N=24 SNPs) or asparaginase allergy (N=30 SNPs), 20 were associated with relapse in the Capizzi arms, which included additional doses of asparaginase, compared to only eight associated with relapse in the high-dose methotrexate arms (Fisher's P=0.015). In contrast, relapse SNPs associated with decreased intracellular methotrexate polyglutamates (N=15 SNPs), rapid methotrexate clearance (N=19 SNPs), or decreased *in vivo* response to methotrexate (N=42 SNPs) were balanced equally in their association across IM randomization arm (19 of 76 SNPs significant in the Capizzi arm, 13 of 76 significant in the high-dose arm; Fisher's P=0.32).

Relapse SNPs were associated with both pharmacokinetic and pharmacodynamic phenotypes. For example, the relapse SNP rs10496350 was associated with asparaginase allergy (which results in decreased exposure to asparaginase), and patients carrying at least one copy of the C risk allele had a higher (P adjusted for treatment and ancestry $=2.94\times10^{-5}$) five-year cumulative incidence of relapse (CIR, 37.5%) than did patients with GG genotype (five-year CIR 13.3%) as well as double the risk (P=0.006) of allergy (23.3% for CC or CG genotype vs. 10.8% for GG genotype, Figure 3). Relapse SNPs were also associated with resistance to chemotherapeutic agents: for example, rs743535 (intronic within *CYP2E1*) was associated with both vincristine resistance (median lethal concentration for 50% of cells 0.27 μM for GG genotype vs. 2 μM for GA/AA genotypes, P=0.016) and increased five-year CIR (12% for the GG genotype vs. 20.6% for the GA or AA genotypes, $P=2.42\times10^{-4}$, Figure 4).

Replication cohort

Of the 1,017 relapse SNPs, 595 were evaluable in the independent replication cohort of 719 patients and 32 replicated (representing 19 loci). Of 138 evaluable ancestry-agnostic SNPs, seven were associated with increased relapse in the replication cohort. 25 of the 457 evaluable ancestry-specific SNPs were also associated with increased relapse in the replication cohort in the same ancestry as was identified in the AALL0232 cohort, including three which increased relapse risk in blacks, 18 in Hispanics, and four in whites (Table 1). Of the 32 replicated SNPs, four were associated with an increased relapse in patients treated with high-dose methotrexate, two in patients treated with Capizzi-methotrexate, and two in both cohorts. Of the seven replicated ancestry-agnostic SNPs, four were associated with at least one unfavorable pharmacological phenotype: rs41530849 in *PTPN14* with both rapid methotrexate clearance and *in vitro* asparaginase resistance, rs743535 in *CYP2E1* with *in vitro* vincristine resistance, intergenic SNP rs2463380 with rapid methotrexate clearance, and the missense SNP rs16843643 in *FARP2* with a diminished *in vivo* response to high-dose methotrexate. Additionally, four ancestry-agnostic SNPs and 12 Hispanic-specific SNPs that were associated with CNS relapse in the discovery cohort were replicated in the independent replication cohort, and 23 SNPs were significant among MRD negative patients (Table 1).

SNP contribution to excess relapse risk in black and Hispanic patients

Using classification and regression trees, we identified two SNPs in blacks (rs4710143 and rs16843643), and in Hispanics (rs9325870 and rs743535) most contributing to their excess relapse risk. In a multivariate model, these four SNPs attenuated the adverse risk associated with black (P=0.79) and Hispanic (P=0.065) ancestry group status (Figure 5a). Additionally, ancestry did not improve the ability to predict relapse if SNPs and treatment group were already known (ANOVA P=0.19 comparing a model with treatment and SNPs as covariates to a model with treatment, SNPs, and ancestry). Patients carrying at least one risk allele for any of the four SNPs had higher relapse than did patients without any risk alleles, regardless of their ancestry (Figure 5b). These variants were less prevalent in whites, with the average white patient carrying 0.21 risk alleles (of a possible eight, range in whites 0–2) compared to a mean of 1.28 in black patients (range 0–5), 0.79 in Hispanics (range 0–4), and 0.63 in patients of other ancestry (range 0–4; Mann-Whitney $P<1\times10^{-15}$).

Discussion

Relapse in high-risk B-ALL remains a significant problem, and most patients who relapse do not survive. Although evaluation of early treatment response and MRD identifies many patients at high risk for relapse, many patients who relapse do not carry these adverse features.(33, 34) Further identification of adverse biologic features is needed to allow further refinements in therapy.

In this study, we focused on three primary implications of this genetic analysis: whether host genetic variation explained ancestry-related differences in relapse, whether the importance of genetic variation differed by major treatment arms, and how genetic variations were replicated for orthogonal pharmacologic phenotypes and in an independent ALL cohort. In

this GWAS, we identified 1,017 SNPs associated with increased relapse risk in children with high-risk B-ALL. We identified both SNPs associated with relapse risk regardless of patient ancestry (ancestry-agnostic) as well as SNPs associated with relapse in an ancestry-specific fashion. Of these relapse SNPs, 7 ancestry-agnostic and 25 ancestry-specific SNPs were also associated with an increased relapse risk in an independent replication cohort (Table 1).

Importantly, we identified genetic variants associated with increased relapse risk in an ancestry-specific manner across two generations of B-ALL protocols. The identified SNPs contribute to the higher risk of relapse in blacks and Hispanics but also identify patients in each ancestral group at high risk of relapse. Using only four SNPs (rs4710143, rs16843643, rs9325870, and rs743535), we identified 73% of blacks and 57% of Hispanics at high-risk of relapse (Figure 5b). These SNPs were also associated with relapse risk in whites and patients of other ancestry. However, more than 50% of blacks and Hispanics carry at least one risk allele in these SNPs compared to 20% of whites, suggesting the increased relapse risk attributable to these SNPs is disproportionately distributed to blacks and Hispanics, simply on the basis of racial differences in allele frequency. The addition of ancestry group to a model including these SNPs and treatment group failed to improve the model (ANOVA P=0.19) suggesting these SNPs attenuate the adverse impact of ancestry on relapse. These data mirror findings in other malignant(35, 36) and non-malignant diseases(37–42) in which variants strongly associated with ancestry may be the cause of discrepant disease outcomes in different ancestral populations. Such variants offer the opportunity for therapy modification and risk stratification when their effects are stable across multiple settings, as are the replicated ancestry-specific variants identified in this study (Table 1).

One of the principles of discovery research in pharmacogenomics is that the variants identified in any study will be influenced by the therapy that has been given. Because the randomly assigned methotrexate treatment arm had a significant effect on treatment outcome in AALL0232, we had a unique opportunity to test whether some genomic variants associated with relapse were more important in those receiving one treatment arm (high-dose methotrexate) versus the other (Capizzi methotrexate plus asparaginase). Interestingly, the SNPs directly associated with methotrexate pharmacology did not differentially distribute between the two treatment arms (Fisher’s P=0.32), but relapse SNPs associated with asparaginase resistance or asparaginase allergy did cluster in the Capizzi arm (Fisher’s P=0.015). Those in the Capizzi arm received more asparaginase but less methotrexate than those in the high-dose methotrexate arm. The association with asparaginase resistance/allergy in the Capizzi arm suggests that asparaginase exposure was more critical to preventing relapse among the patients whose methotrexate exposure was low (Capizzi treatment), and that treatment with high-dose methotrexate diminishes the importance of maximizing asparaginase.

Therapeutic and patient differences may also explain differences in the SNPs associated with relapse in this cohort compared to prior analyses. In prior GWAS of ALL relapse risk and MRD,(43, 44) the majority of patients were NCI standard-risk, in contrast to the high-risk population studied here. Moreover, all patients in the discovery cohort of this study also received delayed intensification and MRD-directed therapy intensification, whereas many patients in the prior GWASs(43, 44) did not. In a review of the SNPs previously associated

with relapse or MRD,(43, 44) we identified five (rs35229355, rs7517671, rs10883699, rs7350429, and rs6773449) that associated with relapse ($P<0.05$) after adjusting for both treatment and ancestry in the current discovery cohort. However, these SNPs did not reach the Ip selected P value threshold, nor were they replicated at least 20 times during iterative resampling. This finding highlights the importance of population and therapeutic differences on the association of pharmacogenomic variants and outcome. It is encouraging that many of the SNPs identified in the current GWAS were associated with relapse among patients treated on both the high-dose methotrexate and the Capizzi escalating-methotrexate/asparaginase arms, suggesting that some of these variants may be prognostic across therapies.

The analysis of relapse SNPs' association with orthogonal pharmacologic phenotypes suggests mechanisms through which some relapse SNPs may be exerting their effects on relapse risk. Relapse SNPs were associated with both pharmacokinetic and pharmacodynamic phenotypes. For example, rs6786341 (an intronic variant in lactoferrin) was associated with more rapid methotrexate clearance, a phenotype which has previously been associated with decreased methotrexate polyglutamate accumulation and increased relapse risk.(45, 46) The rs743535 variant in *CYP2E1* was associated with resistance to vincristine (Figure 4). Variants in this gene have previously been implicated in inferior survival in non-Hodgkin's lymphoma(47) and non-small cell lung cancer,(48) potentially due to resistance to chemotherapeutic agents used in those diseases. Variants near *LZTS1*, which include promoter and enhancer marks in neural tissues,(49) were associated with CNS relapse in Hispanics. Suppression of this gene has previously been implicated in metastatic potential in multiple solid tumors,(50–52) suggesting these variants may alter leukemic trafficing into the CNS, thereby altering CNS relapse risk. Other identified CNS relapse SNPs likely contribute to CNS relapse through alterations in leukemic drug resistance or rapid drug clearance, as 36 of 162 CNS relapse SNPs were also associated with pharmakokinetic or drug resistance phenotypes.

We identified several novel inherited risk variants for relapse in a large population of children with high-risk B-precursor ALL. Several of these are associated with the increased relapse risk specific to black and Hispanic ancestry and may contribute to the adverse outcomes attributed to "race." Many of these variants are associated with "inherited" leukemic resistance or rapid clearance of chemotherapy. These findings may allow personalized therapy to further improve outcomes for children with high-risk B-ALL.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding/Support

The work was supported by the National Institutes of Health [grant numbers GM 92666, GM 115279, CA142665, CA 21765, CA 36401, CA98543 (COG Chair's grant), CA98413 (COG Statistical Center), CA114766 (COG Specimen Banking), U01-HG04603, RC2- GM092618, R01-LM010685, 5T32-GM007569]; Leukemia Lymphoma Society (grant number 6168); and by the American Lebanese Syrian Associated Charities.

Role of funding source

The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

References

1. Hunger SP, Lu X, Devidas M, Camitta BM, Gaynon PS, Winick NJ, et al. Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2012; 30(14):1663–9. [PubMed: 22412151]
2. Schrappe M, Moricke A, Reiter A, Henze G, Welte K, Gadner H, et al. Key treatment questions in childhood acute lymphoblastic leukemia: results in 5 consecutive trials performed by the ALL-BFM study group from 1981 to 2000. Klinische Padiatrie. 2013; 225(Suppl 1):S62–72. [PubMed: 23700060]
3. Pui CH, Pei D, Sandlund JT, Ribeiro RC, Rubnitz JE, Raimondi SC, et al. Long-term results of St Jude Total Therapy Studies 11, 12, 13A, 13B, and 14 for childhood acute lymphoblastic leukemia. Leukemia. 2010; 24(2):371–82. [PubMed: 20010620]
4. Place AE, Stevenson KE, Vrooman LM, Harris MH, Hunt SK, O'Brien JE, et al. Intravenous pegylated asparaginase versus intramuscular native Escherichia coli L-asparaginase in newly diagnosed childhood acute lymphoblastic leukaemia (DFCI 05–001): a randomised, open-label phase 3 trial. The lancet oncology. 2015; 16(16):1677–90. [PubMed: 26549586]
5. Pui CH, Campana D, Pei D, Bowman WP, Sandlund JT, Kaste SC, et al. Treating childhood acute lymphoblastic leukemia without cranial irradiation. N Engl J Med. 2009; 360(26):2730–41. [PubMed: 19553647]
6. Conter V, Valsecchi MG, Parasole R, Putti MC, Locatelli F, Barisone E, et al. Childhood high-risk acute lymphoblastic leukemia in first remission: results after chemotherapy or transplant from the AIEOP ALL 2000 study. Blood. 2014; 123(10):1470–8. [PubMed: 24415536]
7. Stary J, Zimmermann M, Campbell M, Castillo L, Dibar E, Donska S, et al. Intensive chemotherapy for childhood acute lymphoblastic leukemia: results of the randomized intercontinental trial ALL IC-BFM 2002. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2014; 32(3):174–84. [PubMed: 24344215]
8. Larsen EC, Devidas M, Chen S, Salzer WL, Raetz EA, Loh ML, et al. Dexamethasone and High-Dose Methotrexate Improve Outcome for Children and Young Adults With High-Risk B-Acute Lymphoblastic Leukemia: A Report From Children's Oncology Group Study AALL0232. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2016; 34(20): 2380–8. [PubMed: 27114587]
9. Bowman WP, Larsen EL, Devidas M, Linda SB, Blach L, Carroll AJ, et al. Augmented therapy improves outcome for pediatric high risk acute lymphocytic leukemia: results of Children's Oncology Group trial P9906. Pediatric blood & cancer. 2011; 57(4):569–77. [PubMed: 21360654]
10. Marshall GM, Dalla Pozza L, Sutton R, Ng A, de Groot-Kruseman HA, van der Velden VH, et al. High-risk childhood acute lymphoblastic leukemia in first remission treated with novel intensive chemotherapy and allogeneic transplantation. Leukemia. 2013; 27(7):1497–503. [PubMed: 23407458]
11. Diouf B, Crews KR, Lew G, Pei D, Cheng C, Bao J, et al. Association of an inherited genetic variant with vincristine-related peripheral neuropathy in children with acute lymphoblastic leukemia. JAMA : the journal of the American Medical Association. 2015; 313(8):815–23. [PubMed: 25710658]
12. Gregers J, Green H, Christensen IJ, Dalhoff K, Schroeder H, Carlsen N, et al. Polymorphisms in the ABCB1 gene and effect on outcome and toxicity in childhood acute lymphoblastic leukemia. The pharmacogenomics journal. 2015; 15(4):372–9. [PubMed: 25582575]
13. Moriyama T, Nishii R, Perez-Andreu V, Yang W, Klussmann FA, Zhao X, et al. NUDT15 polymorphisms alter thiopurine metabolism and hematopoietic toxicity. Nature genetics. 2016; 48(4):367–73. [PubMed: 26878724]

14. Ramsey LB, Janke LJ, Edick MJ, Cheng C, Williams RT, Sherr CJ, et al. Host thiopurine methyltransferase status affects mercaptopurine antileukemic effectiveness in a murine model. Pharmacogenetics and genomics. 2014; 24(5):263–71. [PubMed: 24710003]
15. Kawedia JD, Liu C, Pei D, Cheng C, Fernandez CA, Howard SC, et al. Dexamethasone exposure and asparaginase antibodies affect relapse risk in acute lymphoblastic leukemia. Blood. 2012; 119(7):1658–64. [PubMed: 22117041]
16. Fernandez CA, Smith C, Yang W, Mullighan CG, Qu C, Larsen E, et al. Genome-wide analysis links NFATC2 with asparaginase hypersensitivity. Blood. 2015; 126(1):69–75. [PubMed: 25987655]
17. Abrahao R, Lichtensztajn DY, Ribeiro RC, Marina NM, Keogh RH, Marcos-Gragera R, et al. Racial/ethnic and socioeconomic disparities in survival among children with acute lymphoblastic leukemia in California, 1988–2011: A population-based observational study. Pediatric blood & cancer. 2015; 62(10):1819–25. [PubMed: 25894846]
18. Goggins WB, Lo FF. Racial and ethnic disparities in survival of US children with acute lymphoblastic leukemia: evidence from the SEER database 1988–2008. Cancer Causes Control. 2012; 23(5):737–43. [PubMed: 22450738]
19. Kadan-Lottick NS, Ness KK, Bhatia S, Gurney JG. Survival variability by race and ethnicity in childhood acute lymphoblastic leukemia. JAMA : the journal of the American Medical Association. 2003; 290(15):2008–14. [PubMed: 14559954]
20. Bhatia S, Sather HN, Heerema NA, Trigg ME, Gaynon PS, Robison LL. Racial and ethnic differences in survival of children with acute lymphoblastic leukemia. Blood. 2002; 100(6):1957–64. [PubMed: 12200352]
21. Kahn JM, Keegan TH, Tao L, Abrahao R, Bleyer A, Viny AD. Racial disparities in the survival of American children, adolescents, and young adults with acute lymphoblastic leukemia, acute myelogenous leukemia, and Hodgkin lymphoma. Cancer. 2016; 122(17):2723–30. [PubMed: 27286322]
22. Bhatia S, Landier W, Hageman L, Kim H, Chen Y, Crews KR, et al. 6MP adherence in a multiracial cohort of children with acute lymphoblastic leukemia: a Children's Oncology Group study. Blood. 2014; 124(15):2345–53. [PubMed: 24829202]
23. Pui CH, Sandlund JT, Pei D, Rivera GK, Howard SC, Ribeiro RC, et al. Results of therapy for acute lymphoblastic leukemia in black and white children. JAMA : the journal of the American Medical Association. 2003; 290(15):2001–7. [PubMed: 14559953]
24. Bhatia S, Landier W, Hageman L, Chen Y, Kim H, Sun CL, et al. Systemic Exposure to Thiopurines and Risk of Relapse in Children With Acute Lymphoblastic Leukemia: A Children's Oncology Group Study. JAMA oncology. 2015; 1(3):287–95. [PubMed: 26181173]
25. Pollock BH, DeBaun MR, Camitta BM, Shuster JJ, Ravindranath Y, Pullen DJ, et al. Racial differences in the survival of childhood B-precursor acute lymphoblastic leukemia: a Pediatric Oncology Group Study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2000; 18(4):813–23. [PubMed: 10673523]
26. Perez-Andreu V, Roberts KG, Xu H, Smith C, Zhang H, Yang W, et al. A genome-wide association study of susceptibility to acute lymphoblastic leukemia in adolescents and young adults. Blood. 2015; 125(4):680–6. [PubMed: 25468567]
27. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, et al. Novel susceptibility variants at 10p12.31–12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. Journal of the National Cancer Institute. 2013; 105(10):733–42. [PubMed: 23512250]
28. Yang JJ, Cheng C, Devidas M, Cao X, Fan Y, Campana D, et al. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. Nature genetics. 2011; 43(3):237–41. [PubMed: 21297632]
29. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155(2):945–59. [PubMed: 10835412]
30. Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, Brutsaert TD, et al. A genomewide admixture mapping panel for Hispanic/Latino populations. American journal of human genetics. 2007; 80(6):1171–8. [PubMed: 17503334]

31. Cheng C, Pounds SB, Boyett JM, Pei D, Kuo ML, Roussel MF. Statistical significance threshold criteria for analysis of microarray gene expression data. Stat Appl Genet Mol Biol. 2004; 3 Article36.
32. Borowitz MJ, Devidas M, Hunger SP, Bowman WP, Carroll AJ, Carroll WL, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia and its relationship to other prognostic factors: A Children's Oncology Group study. Blood. 2008; 111(12):5477–85. [PubMed: 18388178]
33. Borowitz MJ, Wood BL, Devidas M, Loh ML, Raetz EA, Salzer WL, et al. Prognostic significance of minimal residual disease in high risk B-ALL: a report from Children's Oncology Group study AALL0232. Blood. 2015; 126(8):964–71. [PubMed: 26124497]
34. Pui CH, Pei D, Coustan-Smith E, Jeha S, Cheng C, Bowman WP, et al. Clinical utility of sequential minimal residual disease measurements in the context of risk-based therapy in childhood acute lymphoblastic leukaemia: a prospective study. The lancet oncology. 2015; 16(4):465–74. [PubMed: 25800893]
35. Rocconi RP, Lankes HA, Brady WE, Goodfellow PJ, Ramirez NC, Alvarez RD, et al. The role of racial genetic admixture with endometrial cancer outcomes: An NRG Oncology/Gynecologic Oncology Group study. Gynecologic oncology. 2016; 140(2):264–9. [PubMed: 26603970]
36. Hernandez-Suarez G, Sanabria MC, Serrano M, Herran OF, Perez J, Plata JL, et al. Genetic ancestry is associated with colorectal adenomas and adenocarcinomas in Latino populations. Eur J Hum Genet. 2014; 22(10):1208–16. [PubMed: 24518838]
37. Torgerson DG, Capurso D, Ampleford EJ, Li X, Moore WC, Gignoux CR, et al. Genome-wide ancestry association testing identifies a common European variant on 6q14.1 as a risk factor for asthma in African American subjects. The Journal of allergy and clinical immunology. 2012; 130(3):622–9 e9. [PubMed: 22607992]
38. Corvol H, De Giacomo A, Eng C, Seibold M, Ziv E, Chapela R, et al. Genetic ancestry modifies pharmacogenetic gene-gene interaction for asthma. Pharmacogenetics and genomics. 2009; 19(7): 489–96. [PubMed: 19503017]
39. Alarcon-Riquelme ME, Ziegler JT, Molineros J, Howard TD, Moreno-Estrada A, Sanchez-Rodriguez E, et al. Genome-Wide Association Study in an Amerindian Ancestry Population Reveals Novel Systemic Lupus Erythematosus Risk Loci and the Role of European Admixture. Arthritis Rheumatol. 2016; 68(4):932–43. [PubMed: 26606652]
40. Molineros JE, Maiti AK, Sun C, Looger LL, Han S, Kim-Howard X, et al. Admixture mapping in lupus identifies multiple functional variants within IFIH1 associated with apoptosis, inflammation, and autoantibody production. PLoS genetics. 2013; 9(2):e1003222. [PubMed: 23441136]
41. Niewold TB, Kelly JA, Kariuki SN, Franek BS, Kumar AA, Kaufman KM, et al. IRF5 haplotypes demonstrate diverse serological associations which predict serum interferon alpha activity and explain the majority of the genetic association with systemic lupus erythematosus. Annals of the rheumatic diseases. 2012; 71(3):463–8. [PubMed: 22088620]
42. Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, et al. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. Nature genetics. 2008; 40(10): 1175–84. [PubMed: 18794856]
43. Yang JJ, Cheng C, Devidas M, Cao X, Campana D, Yang W, et al. Genome-wide association study identifies germline polymorphisms associated with relapse of childhood acute lymphoblastic leukemia. Blood. 2012; 120(20):4197–204. [PubMed: 23007406]
44. Yang JJ, Cheng C, Yang W, Pei D, Cao X, Fan Y, et al. Genome-wide interrogation of germline genetic variation associated with treatment response in childhood acute lymphoblastic leukemia. JAMA : the journal of the American Medical Association. 2009; 301(4):393–403. [PubMed: 19176441]
45. Evans WE, Crom WR, Stewart CF, Bowman WP, Chen CH, Abromowitch M, et al. Methotrexate systemic clearance influences probability of relapse in children with standard-risk acute lymphocytic leukaemia. Lancet. 1984; 1:359–62. [PubMed: 6141424]
46. Mikkelsen TS, Sparreboom A, Cheng C, Zhou Y, Boyett JM, Raimondi SC, et al. Shortening infusion time for high-dose methotrexate alters antileukemic effects: a randomized prospective clinical trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011; 29(13):1771–8. [PubMed: 21444869]

47. Han X, Zheng T, Foss FM, Lan Q, Holford TR, Rothman N, et al. Genetic polymorphisms in the metabolic pathway and non-Hodgkin lymphoma survival. American journal of hematology. 2010; 85(1):51–6. [PubMed: 20029944]

48. Haque AK, Au W, Cajas-Salazar N, Khan S, Ginzel AW, Jones DV, et al. CYP2E1 polymorphism, cigarette smoking, p53 expression, and survival in non-small cell lung cancer: a long term follow-up study. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry. 2004; 12(4):315–22.

49. Roadmap Epigenomics C. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518(7539):317–30. [PubMed: 25693563]

50. Lin CW, Chang YL, Chang YC, Lin JC, Chen CC, Pan SH, et al. MicroRNA-135b promotes lung cancer metastasis by regulating multiple targets in the Hippo pathway and LZTS1. Nature communications. 2013; 4:1877.

51. Onken MD, Worley LA, Harbour JW. A metastasis modifier locus on human chromosome 8p in uveal melanoma identified by integrative genomic analysis. Clinical cancer research : an official journal of the American Association for Cancer Research. 2008; 14(12):3737–45. [PubMed: 18559591]

52. Wang XX, Zhu ZM, Su D, Lei T, Wu X, Fan Y, et al. Down-regulation of leucine zipper putative tumor suppressor 1 is associated with poor prognosis, increased cell motility and invasion, and epithelial-to-mesenchymal transition characteristics in human breast carcinoma. Hum Pathol. 2011; 42(10):1410–9. [PubMed: 21419475]

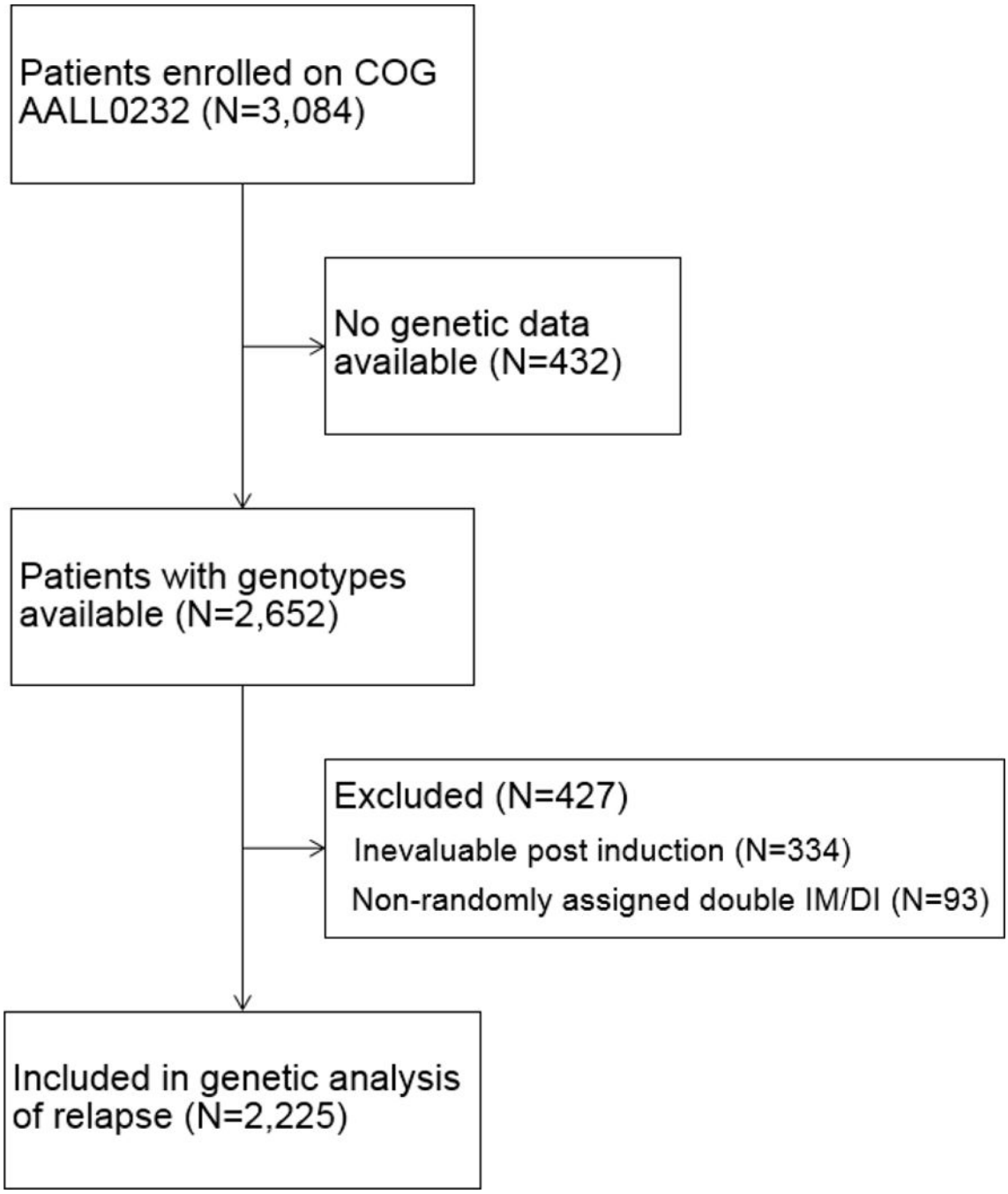


Figure 1.
Consort diagram of AALL0232 discovery cohort

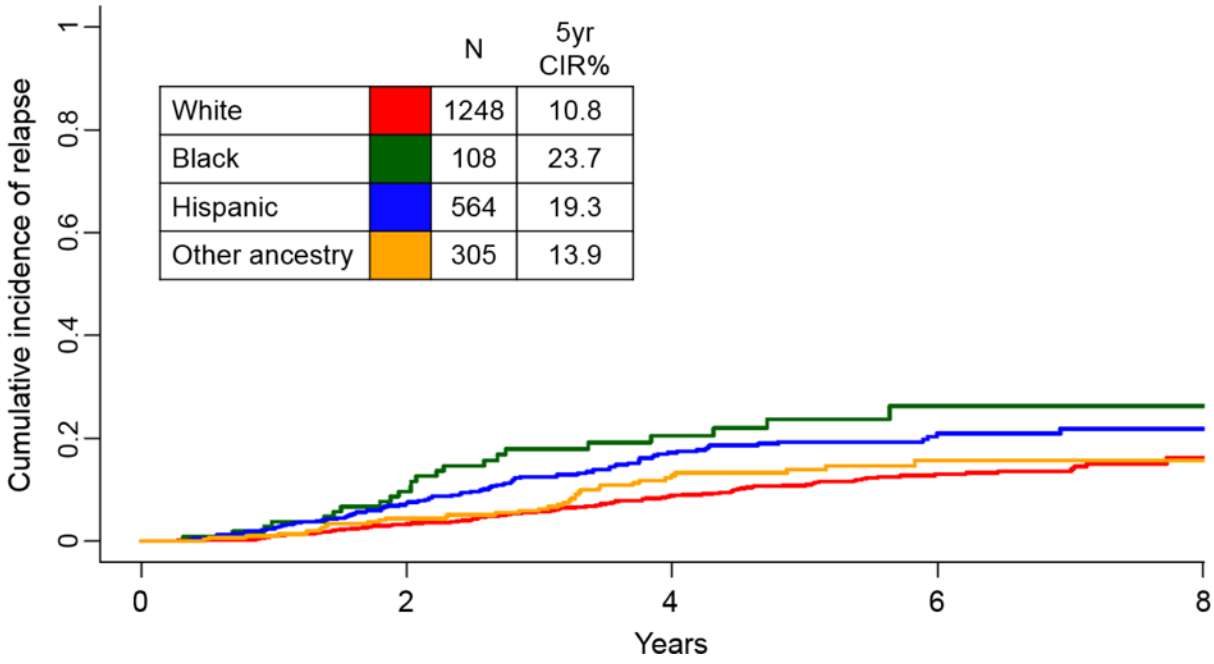


Figure 2. Association of non-white genetic ancestry with increased relapse risk

Non-whites had an increased risk of relapse in the discovery cohort. The five-year cumulative incidence of relapse was higher in blacks [23.7%, 95% confidence interval (CI) 14.7–32.7%, $P=2.27\times10^{-4}$, HR=2.32] and Hispanics (19.3%, 95% CI 15.7–22.9%, $P=8.23\times10^{-5}$, HR=1.7) than whites (10.3%, 95% CI 8.9–12.8%). P values are adjusted for treatment.

White: >90% CEU; black: >70% YRI; Hispanic: >10% Native American and Native American >YRI

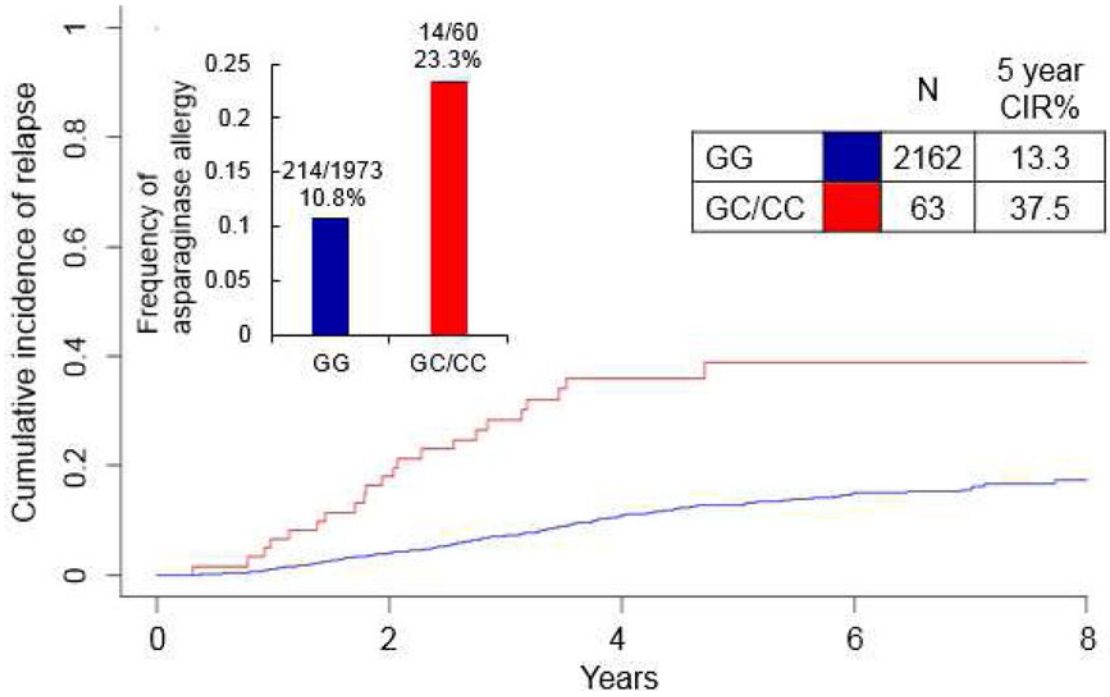


Figure 3. ***NPAS2*** **SNP rs10496350 is associated with asparaginase allergy and increased relapse risk**

Patients in the discovery cohort carrying the at least one copy of the C risk allele of rs10496350 had a higher five-year cumulative incidence of relapse (37.5%) than did those with the GG genotype (13.3%, P adjusted for treatment and ancestry $=2.94\times10^{-5}$). Patients carrying the risk allele also experienced a higher rate of allergy (23%) than did patients carrying the GG genotype (11%, P=0.006).

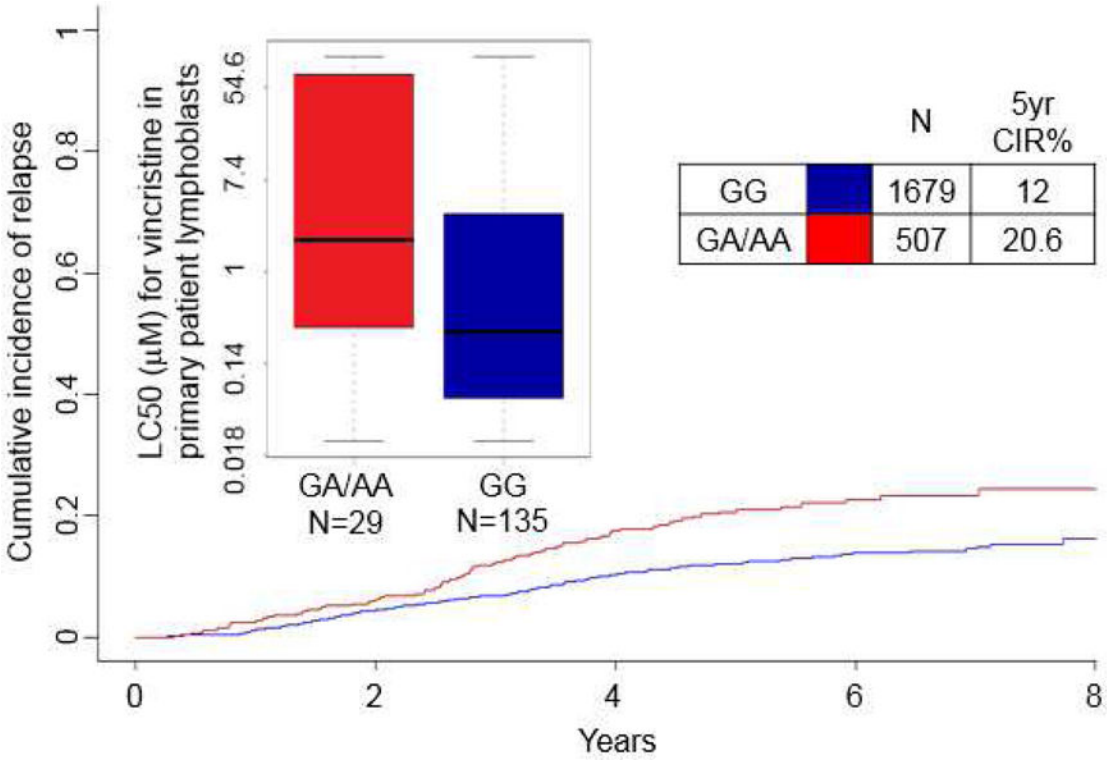
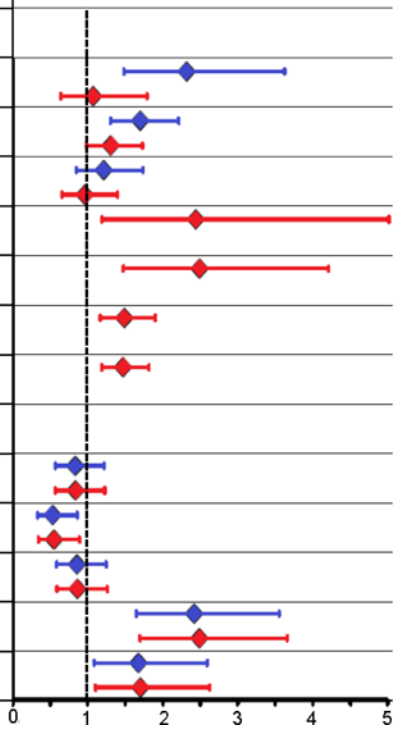


Figure 4. *CYP2E1* SNP rs743535 associated with both *in vitro* vincristine resistance and increased relapse risk

rs743535 was associated with increased relapse risk (multivariate $P=2.42\times10^{-4}$). In primary patient lymphoblasts, presence of one or more A risk alleles decreased sensitivity to vincristine (median LC50 with A allele 2 µM, median LC50 with GG genotype 0.27 µM, P=0.016).

a

	P: model without SNPs	HR (95% CI): model without SNPs	P: model with SNPs	HR (95% CI): model with SNPs
White	Reference		Reference	
• Black	$2.3x10^{-4}$	2.3 (1.5-3.6)	0.79	1.1 (0.64-1.7)
• Hispanic	$8.2x10^{-5}$	1.7 (1.3-2.2)	0.065	1.3 (0.98-1.7)
• Other ancestry	0.29	1.2 (0.85-1.7)	0.83	0.96 (0.66-1.4)
rs4710143	NA		0.015	2.4 (1.2-5)
rs16843643	NA		$6.5x10^{-4}$	2.5 (1.5-4.2)
rs9325870	NA		0.0014	1.5 (1.2-1.9)
rs743535	NA		$3.6x10^{-4}$	1.5 (1.2-1.8)
Dex/ Capizzi	Reference		Reference	
• Pred/ Capizzi	0.35	0.83 (0.57-1.2)	0.36	0.84 (0.57-1.2)
• Dex/ HD	0.0099	0.53 (0.33-0.86)	0.015	0.55 (0.34-0.89)
• Pred/ HD	0.41	0.85 (0.59-1.3)	0.44	0.86 (0.59-1.3)
• SER w/ Capizzi	$6.7x10^{-6}$	2.4 (1.7-3.6)	$3.5x10^{-6}$	2.5 (1.7-3.7)
• SER w/ HD	0.02	1.7 (1.1-2.6)	0.016	1.7 (1.1-2.6)



b

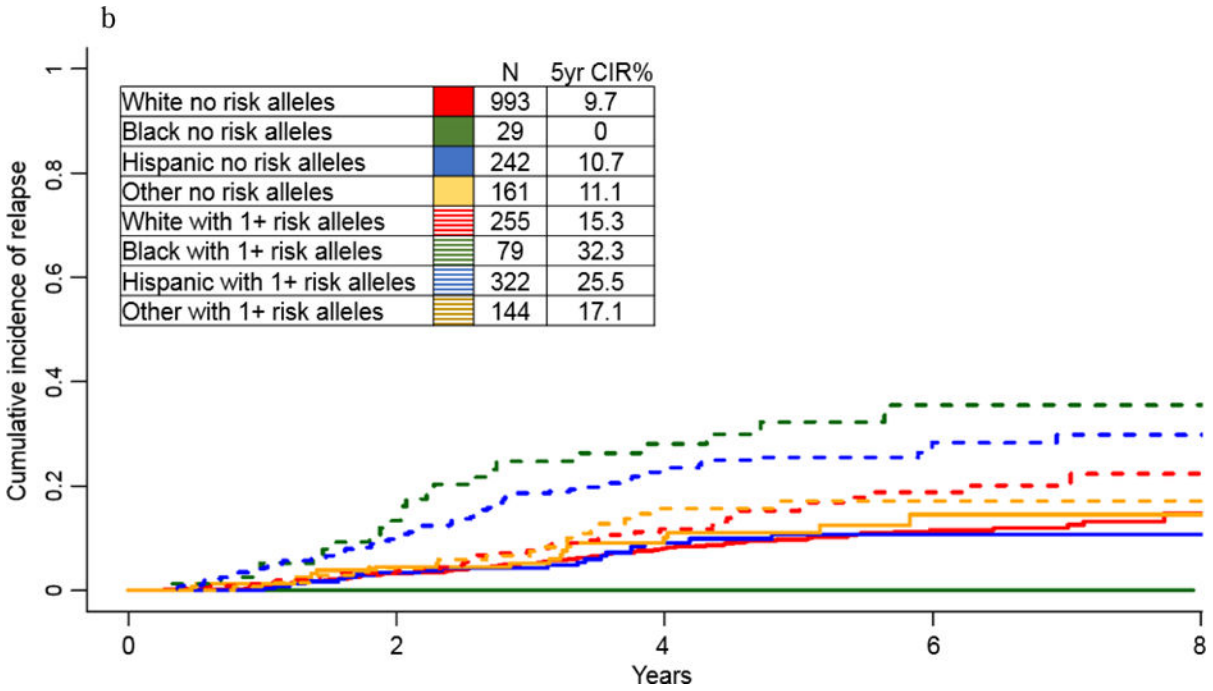


Figure 5. Relapse SNPs attenuate the adverse impact of black and Hispanic ancestry

a: **Forest plot of relapse risk comparing multivariable models with and without four relapse SNPs**

b: **Presence of a risk allele in any of the four SNPs confers high relapse risk regardless of ancestry**

Risk alleles in any of four SNPs (rs4710143, rs16843643, rs9325870, and rs743535) confer increased relapse risk regardless of ancestry. (A) In multivariate models, these SNPs largely attenuate the adverse effect of black or Hispanic ancestry, while leaving unchanged the association between treatment arm and relapse. Treatment arms are described in Supplementary Table 2: For the rapid early response patients, Dex/Capizzi, Pred/Capizzi, Dex/HD, Pred/HD refer to the induction steroid (dex = dexamethasone, pred = prednisone) and the interim maintenance (Capizzi=escalating dose methotrexate plus asparaginase, HD = high-dose methotrexate). For the slow early response patients (SER), induction steroid groups were combined. (Hazard ratio (HR) from model without SNPs shown in blue, models with SNPs shown in red).

(B) Whites, blacks, and Hispanics carrying risk alleles for any of these SNPs (dashed lines) have higher five-year relapse risks than do those without any risk alleles (solid lines) [15.3% vs. 9.7% (P=0.025) for whites, 32.3% vs. 0% ($P=1.28\times10^{-4}$) for blacks, and 25.5% vs. 10.7% ($P=3.72\times10^{-6}$) for Hispanics]. P values are adjusted for treatment.

Table 1

SNPs associated with relapse in discovery (n=2,225) and replication (n=719) cohorts

rsID	Gene	Risk allele	RAF	P: discovery cohort	Hazard Ratio (95% CI): discovery cohort	P: replication cohort	Ancestry	Additional phenotypes
rs41530849	*PTPN14*	T	0.006	4.26E-06	3.87 (2.17–6.89)	0.019	agnostic	HD arm, MTX clearance, PPL ASP
rs10205940		A	0.224	6.85E-06	1.42 (1.22–1.66)	0.044	agnostic	HD arm, Capizzi arm, CNS
chr23: 9863426	*SHROOM2*	T	0.008	1.04E-05	2.45 (1.64–3.64)	0.049	agnostic	HD arm, Capizzi arm, CNS
rs2463380		G	0.22	3.98E-05	1.52 (1.24–1.86)	0.045	agnostic	HD arm, CNS, MTX clearance
rs2710418	*NELL2*	T	0.031	4.99E-05	1.98 (1.42–2.76)	0.021	agnostic	HD arm
rs743535	*CYP2E1*	A	0.124	5.00E-05	1.54 (1.25–1.9)	0.045	agnostic	PPL Vinc
rs16843643	*FARP2*	C	0.012	0.000226	2.95 (1.66–5.25)	0.031	agnostic	Capizzi arm, CNS, MTX WBC response
rs775491	*BEST3*	A	0.304	0.000265	1.61 (1.24–2.07)	0.0086	white	
rs156008	*PCSK1*	A	0.158	0.000297	1.65 (1.26–2.16)	0.024	white	
rs4710143	*RNASET2*	G	0.074	0.000579	4.92 (1.98–12.2)	0.014	black	Capizzi arm
rs202408		C	0.144	0.000789	3.56 (1.7–7.49)	0.021	black	
rs7860525		T	0.134	0.00175	2.79 (1.47–5.31)	0.016	black	
rs9325870	*LZTS1*	C	0.205	1.84E-05	2 (1.46–2.75)	0.036	Hispanic	CNS
rs16999479	*DSCAM*	G	0.016	0.000219	4.02 (1.92–8.42)	0.046	Hispanic	CNS
rs141707566	*GRIN2A*	C	0.014	0.000289	2.76 (1.59–4.77)	0.037	Hispanic	HD arm
rs12535024	*DDC*	C	0.181	0.00103	1.76 (1.26–2.47)	0.0499	Hispanic	
rs6786341	*LTF*	T	0.012	0.00173	4.14 (1.7–10.1)	0.038	Hispanic	MTX clearance

rsID	Gene	Risk allele	RAF	P: discovery cohort	Hazard Ratio (95% CI): discovery cohort	P: replication cohort	Ancestry	Additional phenotypes
rs16945138	*DNAH9*	T	0.007	0.00186	7.5 (2.11–26.7)	0.014	Hispanic	
rs6651255	*GSDMC*	C	0.425	0.00222	1.59 (1.18–2.13)	0.0029	Hispanic	

RAF: Risk allele frequency; CI: confidence interval

Characteristics of validated SNPs are shown for the discovery cohort, with one SNP for each locus shown (SNPs removed through LD pruning are shown in Supplementary Tables 4–7). Bolded SNPs were significant at the Ip determined significance threshold when evaluated among patients who were end-induction minimal residual disease negative. SNPs are ordered by ancestry of discovery, with SNPs associated with relapse while adjusting for both treatment and ancestry (i.e. “ancestry agnostic”) labeled as agnostic and ancestry-specific SNPs labeled with their associated ancestry group. Additional phenotypes include: association with relapse among patients treated on either first interim maintenance arm [Capizzi arm, HD (high-dose methotrexate) arm], association with CNS relapse (CNS), as well as association with *in vitro* resistance among primary patient lymphoblasts to asparaginase (PPL ASP) or vincristine (PPL Vinc), more rapid methotrexate clearance (MTX clearance), or diminished white blood cell decrease after *in vivo* methotrexate treatment (MTX WBC response).