American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

Common variation at 1q23.3, 2p23.3, 2q33.3, and 2p21 influences risk of acute myeloid leukemia

Tracking no: BLD-2025-031266R2

Diyanath Ranasinghe (Newcastle University, United Kingdom) Wei-Yu Lin (1Translational and Clinical Research Institute, Newcastle University Centre for Cancer,) Sarah Fordham (Newcastle University, United Kingdom) Abrar Alharbi (Department of Clinical Laboratory Sciences, faculty of Applied Medical Sciences, Taibah University, Medina, Saudi Arabia) Nicola Sunter (Newcastle University, United Kingdom) Claire Elstob (Newcastle University, United Kingdom) Mohammed Nahari (Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Najran University, Najran, Saudi Arabia) Yaobo Xu (Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Najran University, Najran, Saudi Arabia) Catherine Park (Newcastle University, United Kingdom) Eric Hungate (University of Chicago, United States) Anne Quante (Institute of Human Genetics, Technical University of Munich (TUM), TUM School of Medicine and Health, Klinikum rechts der Isar, TUM University Hospital, Munich, Germany) Konstantin Strauch (Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center of the Johannes Gutenberg University Mainz, Germany) Christian Gieger (Helmholtz Zentrum Munich, Germany) Andrew Skol (4University of Chicago Comprehensive Cancer Center, United States) Thahira Rahman (Newcastle University, United Kingdom) Lara Sucheston-Campbell (Wayne State University School of Medicine, United States) Theresa Hahn (Roswell Park Comprehensive Cancer Center, United States) Alyssa Clay-Gilmour (University of South Carolina, United States) Gail Jones (Freeman Hospital, United Kingdom) Helen Marr (Freeman Hospital, United Kingdom) Graham Jackson (Department of Haematology, Freeman Hospital,) Tobias Menne (Newcastle upon Tyne Hospitals NHS Foundation Trust, United Kingdom) Matthew Collin (12Department of Haematology, Freeman Hospital, United Kingdom) Adam Ivey (Alfred Hospital, Australia) Robert Hills (University of Oxford, United Kingdom) Alan Burnett (Paul O'Gorman Leukaemia Research Centre, University of Glasgow, United Kingdom) Nigel Russell (Guy's and St Thomas' NHS Foundation Trust, United Kingdom) Jude Fitzgibbon (Queen Mary University of London, United Kingdom) Richard Larson (University of Chicago, United States) Michelle Le Beau (University of Chicago, United States) Wendy Stock (University of Chicago, United States) Olaf Heidenreich (Princess Maxima Centrum for Pediatric Oncology, Netherlands) Amir Enshaei (Translational and Clinical Research Institute, Newcastle University, United Kingdom) Dumni Gunasinghe (Newcastle University, United Kingdom) Zoë Hawking (Newcastle University, United Kingdom) Holly Heslop (Newcastle University, United Kingdom) Devi Nandana (Translational and Clinical Research Institute, Newcastle University, United Kingdom) Bingjing Di (Newcastle University, United Kingdom) Anna Plokhuta (Newcastle University, United Kingdom) Imogen Brown (Newcastle University, United Kingdom) David Allsup (Hull York Medical School, United Kingdom) Richard Houlston (Insitute of Cancer Research, United Kingdom) Andrew Collins (University of southampton, United Kingdom) Paul Milne (Newcastle University, United Kingdom) Jean Norden (Newcastle University, United Kingdom) Anne Dickinson (Newcastle University, United Kingdom) Beverley Lendrem (Newcastle University, United Kingdom) Ann Daly (Newcastle University,) Louise Palm (Birmingham Women's and Children's Hospital, United Kingdom) Kim Piechocki (Birmingham Women's Hospital, United Kingdom) Sally Jeffries (Birmingham Women's Hospital, United Kingdom) Martin Bornhäuser (School of Cancer and Pharmaceutical Research, Department of Haematology, King's College London, United Kingdom) Christoph Röllig (Technical University of Dresden, Germany) Heidi Altmann (University Hospital Carl Gustav Carus, TU Dresden, Germany) Leo Ruhnke (Universitätsklinikum Dresden, Germany) Desiree Kunadt (Department of Internal Medicine I, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany, Germany) Lisa Wagenführ (University Hospital, TU Dresden, Germany) Heather Cordell (Newcastle University, United Kingdom) Rebecca Darlay (Newcastle University, United Kingdom) Mette Andersen (Rigshospitalet, Denmark) Maria Fontana (University of Bologna, Italy) Giovanni Martinelli (Institute of Hematology, University of Bologna, Italy) Giovanni Marconi (University of Bologna, Italy) Miguel Ángel Sanz (Hospital Universitario y Politécnico La Fe, Spain) José Cervera (Hospital Universitario La Fe, Spain) Ines Gomez-Segui (Hospital Universitari i Politecnic La Fe, Spain) Thomas Cluzeau (CHU Nice, France) Chimene Moreilhon (CHU of Nice, France) Sophie Raynaud (Pasteur Hospital, CHU de Nice, France) Heinz Sill (Medical University of Graz, Austria) Maria Teresa Voso (Tor Vergata, Italy) Hervé Dombret (Université Paris Cité, France) Meyling Cheok (UMR9020 CNRS - U1277 Inserm - Université de Lille - CHU de Lille, France) Claude Preudhomme (chru of Lille, France) Rosemary Gale (University College London, United Kingdom) David Linch (University College London, United Kingdom) Julia Weisinger (Semmelweis University, Hungary) Andras Masszi (National Institute of Oncology, Hungary) Daniel Nowak (Department of Hematology and Oncology, Medical Faculty Mannheim, Heidelberg University, Germany) Wolf Hofmann (Department of Hematology and Oncology, Medical Faculty Mannheim, Heidelberg University,) Amanda Gilkes (Cardiff University, United Kingdom) Kimmo Porkka (Helsinki University Central Hospital Comprehensive Cancer Center, Finland) Jelena Milosevic Feenstra (Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria, Austria) Robert Kralovics (Medical University of Vienna, Austria) Junke Wang (Department of Systems Biology, UT MD Anderson Cancer Center, United States) Manja Meggendorfer (MLL Munich Leukemia Laboratory, Germany) Torsten Haferlach (MLL Munich Leukemia Laboratory, Germany) Szilvia Krizsán (Semmelweis University, Hungary) Csaba Bödör (Semmelweis University, Hungary) Brian Parkin (Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan, United States) Sami Malek (University of Michigan, United States) Friedrich Stölzel (Uniklinikum Schleswig-Holstein, Kiel, Germany, Germany) Kenan Onel (Roswell Park Comprehensive Cancer Center, United States) James Allan (Newcastle University, United Kingdom)

Abstract:
Acute myeloid leukemia (AML) is a complex hematological malignancy with multiple disease sub-groups defined by somatic mutations and heterogeneous outcomes. Although genome-wide association studies (GWAS) have identified a small number of common genetic variants influencing AML risk, the heritable component of this disease outside of familial susceptibility remains largely undefined. Here we perform a meta-analysis of four published GWAS plus two new GWAS, totalling 4710 AML cases and 12938 controls. We identify a new genome-wide significant risk locus for pan-AML at 2p23.3 (rs4665765; P=1.35x10-8; EFR3B, POMC, DNMT3A, DNAJC27) which also significantly associates with patient survival (P=6.09x10-3). Our analysis also identifies three new genome-wide significant risk loci for disease sub-groups, including AML with deletions of chromosome 5 and/or 7 at 1q23.3 (rs12078864; P=7.0x10-10; DUSP23) and cytogenetically complex AML at 2q33.3 (rs12988876; P=3.28x10-8; PARD3B) and 2p21 (rs79918355; P=1.60x10-9; EPCAM). We also investigated loci previously associated with risk of clonal hematopoiesis (CH) or clonal hematopoiesis of indeterminate potential (CHIP) and identified several variants associated with risk of AML. Our results further inform on AML etiology and demonstrate the existence of disease sub-group specific risk loci.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: DS and W-YL collated data, conducted data analysis and drafted the manuscript. SEF, AA, NS, CE, MHN, YX, CPa, EH, AQ, KS, CG, AS, TR, LS-C, THah, AIC-G, GLJ, HJM, GHJ, TM, MCo, PM, AI, RKH, AKB, NHR, JF, RAL, MMLB, WS, OH, AE, DJ, ZH, HH, DN, BD, AP, ITB, DJA, RSH, AC, JN, AMD, CL, AKD, LP, KPi, SJ, MB, CR, HA, LR, DK, LW, HJC, RD, MKA, MCF, GMart, GMarc, MAS, JC, IG-S, TC, CM, SR, HS, MTV, HD, MCh, CPr, REG, DL, JWe, AM, DN, W-KH, AG, KPo, JDMF, RK, JWa, MM, THaf, SK, CB, BP, SNM, FS and KO collated data and/or advised on data analysis. JMA collated data, analysed data, directed the research, obtained funding and drafted the manuscript.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Full GWAS meta summary statistics will be deposited in Newcastle University server.

Clinical trial registration information (if any):

Common variation at 1q23.3, 2p23.3, 2q33.3, and 2p21 influences risk of acute myeloid leukemia

Diyanath Ranasinghe[1*], Wei-Yu Lin[1*], Sarah E Fordham[1], Abrar Alharbi[1,2], Nicola J Sunter[1], Claire Elstob[1], Mohammed H Nahari[1,3], Yaobo Xu[1], Catherine Park[1], Eric Hungate[4], Anne Quante[5], Konstantin Strauch[6,7,8], Christian Gieger[7], Andrew Skol[4], Thahira Rahman[1], Lara Sucheston-Campbell[9], Theresa Hahn[10], Alyssa I Clay-Gilmour[11], Gail L Jones[12], Helen J Marr[12], Graham H Jackson[12], Tobias Menne[12], Mathew Collin[12], Adam Ivey[13], Robert K Hills[14], Alan K Burnett[15], Nigel H Russell[16], Jude Fitzgibbon[17], Richard A Larson[4], Michelle M Le Beau[4], Wendy Stock[4], Olaf Heidenreich[1,18,19], Amir Enshaei[1], Dumni Gunasinghe[1], Zoë L Hawking[1], Holly Heslop[1], Devi Nandana[1], Bingjing Di[1], Anna Plokhuta[1], Imogen T. Brown[1], David J Allsup[20], Richard S Houlston[21], Andrew Collins[22], Paul Milne[23], Jean Norden[23], Anne M Dickinson[23], Clare Lendrem[23], Ann K Daly[23], Louise Palm[24], Kim Piechocki[24], Sally Jeffries[24], Martin Bornhäuser[25,26,27], Christoph Röllig[27], Heidi Altmann[27], Leo Ruhnke[27], Desiree Kunadt[27], Lisa Wagenführ[27], Heather J Cordell[28], Rebecca Darlay[28], Mette K Andersen[29], Maria C Fontana[30,31], Giovanni Martinelli[30], Giovanni Marconi[31], Miguel A Sanz[32], José Cervera[32], Inés Gómez-Seguí[32], Thomas Cluzeau[33], Chimène Moreilhon[33], Sophie Raynaud[33], Heinz Sill[34], Maria Teresa Voso[35], Hervé Dombret[36], Meyling Cheok[37], Claude Preudhomme[37], Rosemary E Gale[38], David Linch[38], Julia Weisinger[39], Andras Masszi[40], Daniel Nowak[41], Wolf-Karsten Hofmann[41], Amanda Gilkes[42], Kimmo Porkka[43], Jelena D Milosevic Feenstra[44], Robert Kralovics[44], Junke Wang[45], Manja Meggendorfer[46], Torsten Haferlach[46], Szilvia Krizsán[47,48], Csaba Bödör[47], Brian Parkin[49], Sami N. Malek[49], Friedrich Stölzel[50], Kenan Onel[51], James M Allan[1]

[1]Translational and Clinical Research Institute, Newcastle University Centre for Cancer, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.

[2]Department of Clinical Laboratory Sciences, faculty of Applied Medical Sciences, Taibah University, Medina, Saudi Arabia.

[3]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Najran University, Najran, Saudi Arabia

[4]University of Chicago Comprehensive Cancer Center, Chicago, Illinois. USA.

[5]Institute of Human Genetics, Technical University of Munich (TUM), TUM School of Medicine and Health, Klinikum rechts der Isar, TUM University Hospital, Munich, Germany.

[6]Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.

[7]Ludwig-Maximilians-Universität München, Chair of Genetic Epidemiology, IBE, Faculty of Medicine, Munich, Germany.

[8]Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center, Johannes Gutenberg University, Mainz, Germany.

[9]Karmanos Cancer Institute and Department of Oncology, Wayne State University School of Medicine, Detroit, MI, USA.

[10]Department of Cancer Prevention, Roswell Park Cancer Institute, Buffalo, New York, USA.

[11]Arnold School of Public Health, Department of Epidemiology & Biostatistics, University of South Carolina, USA.

[12]Department of Haematology, Freeman Hospital, Newcastle upon Tyne Hospitals National Health Service Foundation Trust, Newcastle upon Tyne, UK.

[13]Department of Pathology, Alfred Hospital, Melbourne, Australia.

[14]Nuffield Department of Population Health, University of Oxford, Oxford, UK.

[15]Paul O'Gorman Leukaemia Research Centre, University of Glasgow, UK.

[16]Department of Haematology, Guy's and St Thomas' Hospital, London, UK.

[17]Barts Cancer Institute, Queen Mary University of London, London, UK.

[18]Prinses Maxima Center of Pediatric Oncology, Utrecht, The Netherlands.

[19]Department of Hematology, University Medical Center Utrecht, Utrecht, The Netherlands.

[20]Centre for Biomedicine, Hull York Medical School, University of Hull, UK

[21]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.

[22]School of Human Development and Health, Faculty of Medicine, University of Southampton, Southampton SO16 6YD, UK

[23]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.

[24]West Midlands Regional Genetics Laboratory, Birmingham Women's Hospital, Birmingham, UK.

[25]School of Cancer and Pharmaceutical Research, King's College, Department of Haematology, King's College London, London, UK

[26]National Center for Tumor Diseases NCT, Partner site Dresden, Dresden, Germany.

[27]Medizinische Klinik und Poliklinik I, University Hospital Carl Gustav Carus Dresden, Technical University of Dresden, Dresden, Germany.

[28]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK.

[29]Department of Clinical Genetics, University Hospital Rigshospitalet, Copenhagen, Denmark.

[30]Institute of Hematology "L. and A. Seràgnoli", University of Bologna, Bologna, Italy.

[31]University of Bologna, Hematology Unit, Ospedale S. Maria delle Croci, Ravenna (RA), Italy.

[32]Hematology Service, Hospital Universitario y Politécnico La Fe, Valencia, Spain and CIBERONC, Instituto de Salud Carlos III, Madrid, Spain.

[33]Université Côte d'Azur, CHU Pasteur, Service d'Hématologie Biologique, Nice, France.

[34]Division of Hematology, Medical University of Graz, Austria.

[35]Università di Roma Tor Vergata, Dipartimento di Biomedicina e Prevenzione, Rome, Italy.

[36]Hôpital Saint-Louis, Institut Universitaire d'Hématologie, Université Paris Diderot, Paris, France.

[37]University of Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Jean-Pierre Aubert Neurosciences et Cancer, F-59000 Lille, France.

[38]Department of Haematology, University College London Cancer Institute, London, UK.

[39]1st Department of Internal Medicine, Semmelweis University, Budapest, Hungary.

[40]Hematology and Lymphoma Unit, National Institute of Oncology, Budapest, Hungary.

[41]Department of Hematology and Oncology, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany.

[42]Department of Haematology, University of Cardiff, UK.

[43]Helsinki University Hospital Comprehensive Cancer Center, Hematology Research Unit Helsinki, University of Helsinki, Helsinki, Finland.

[44]Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria.

[45]Department of Systems Biology, UT MD Anderson Cancer Center, Houston TX, USA.

[46]MLL Munich Leukemia Laboratory, Munich, Germany.

[47]MTA-SE Lendulet Molecular Oncohematology Research Group, Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary.

[48]Pediatric Center, Semmelweis University, Budapest, Hungary.

[49]Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan, Ann Arbor, Michigan, USA.

[50]Division of Stem Cell Transplantation and Cellular Immunotherapies, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany.

[51]Departments of Cancer Prevention and Control, Medicine, and Clinical Genomics, Roswell Park Comprehensive Cancer Center, Buffalo, New York, USA.

*These authors contributed equally.

Correspondence:

James Allan, james.allan@newcastle.ac.uk

Data availability

Full summary-level association data from meta-analyses of pan-AML, complex AML, del5/7 AML and cytogenetically normal AML are available via the GWAS catalog (study accession numbers GCST90707271, GCST90707272, GCST90707273, GCST90707274). Data availability for cases and controls recruited to GWAS1, GWAS2, GWAS3 and GWAS4 has been reported

previously[6]. Genotyping data and/or samples for GWAS5 cases and controls are available by application to the Finnish Hematology Registry and Clinical Biobank (https://www.fhrb.fi/) and The National Institute for Health and Welfare (THL) Biobank of Finland (https://thl.fi/en/research-and-development/thl-biobank). Genotyping data for GWAS6 cases are available via the NCBI Gene Expression Omnibus under accession numbers GSE21107[8] (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM527831, GSE61323[10] (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61323, GSE23452[9] (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE23452.

Genotyping data for GWAS6 controls are available via application to the database of Genotype and Phenotype (dbGAP)(10.1038/ng1007-1181) under accession number phs000021 (GAIN: Genome-Wide Association Study of Schizophrenia).

eQTL data is available from the eQTLGen consortium via http://www.eqtlgen.org/cis-eqtls.html.
URLs: Michigan Imputation Server, https://imputationserver.sph.umich.edu/index.html#!; Haplotype Reference Consortium, http://www.haplotype-reference-consortium.org/; eQTLGen Consortium, http://www.eqtlgen.org/cis-eqtls.html; 1000 Genomes Project, https://www.internationalgenome.org/; PLINK, https://www.cog-genomics.org/plink2/; SNPTEST2, https://www.well.ox.ac.uk/~gav/snptest/

Key points

- Discovered novel susceptibility loci for pan-AML and disease subtypes, including risk variants common to both clonal hematopoiesis and AML.
- Discovered a novel AML susceptibility variant on chromosome 2p23.3 (localised to DNMT3A) that also associates with patient survival.

Abstract

Acute myeloid leukemia (AML) is a complex hematological malignancy with multiple disease sub-groups defined by somatic mutations and heterogeneous outcomes. Although genome-wide association studies (GWAS) have identified a small number of common genetic variants influencing AML risk, the heritable component of this disease outside of familial susceptibility remains largely undefined. Here we perform a meta-analysis of four published GWAS plus two new GWAS, totalling 4710 AML cases and 12938 controls. We identify a new genome-wide significant risk locus for pan-AML at 2p23.3 (rs4665765; P=1.35x10^{-8}; *EFR3B, POMC, DNMT3A, DNAJC27*) which also significantly associates with patient survival (P=6.09x10^{-3}). Our analysis also identifies three new genome-wide significant risk loci for disease sub-groups, including AML with deletions of chromosome 5 and/or 7 at 1q23.3 (rs12078864; P=7.0x10^{-10}; *DUSP23*) and cytogenetically complex AML at 2q33.3 (rs12988876; P=3.28x10^{-8}; *PARD3B*) and 2p21 (rs79918355; P=1.60x10^{-9}; *EPCAM).* We also investigated loci previously associated with risk of clonal hematopoiesis (CH) or clonal hematopoiesis of indeterminate potential (CHIP) and identified several variants associated with risk of AML. Our results further inform on AML etiology and demonstrate the existence of disease sub-group specific risk loci.

Introduction

Acute myeloid leukemia (AML) is the most common acute leukemia in Europeans and comprises multiple sub-groups defined by somatic genetic/epigenetic alterations and heterogenous clinical outcomes[1]. The existence of rare constitutional genetic variants predisposing to AML with high penetrance demonstrates a role for genetics in disease susceptibility[2,3]. However, strong familial susceptibility to AML is rare, and the prevailing evidence suggests that for the majority of individuals the genetic risk for AML is determined by co-inheritance of multiple independent low penetrance genetic variants[4–6].

To identify novel AML risk loci we conducted a meta-analysis of four published genome-wide association studies (GWAS)[6] and two new GWAS, incorporating 4710 AML cases and 12938 controls of European ancestry, and report the identification of new pan-AML and AML sub-group specific risk loci. This is the largest AML GWAS to date and provides further evidence for the existence of common low-penetrance susceptibility alleles, as well as evidence for heterogeneity in genetic risk across disease sub-groups.

Methods

Study Participants

GWAS1, GWAS2, GWAS3 and GWAS4 comprised 1119, 931, 991 and 977 AML cases, respectively, and 2671, 2477, 1612 and 3728 controls, respectively, as described previously[6]. GWAS5 comprised 351 cases from the Finnish Hematology Registry and Clinical Biobank genotyped on the Illumina Omni Express Exome BeadChip. For controls, we used publicly available Illumina Omni Express Exome BeadChip data on 1055 individuals from The National Institute for Health and Welfare (THL) Biobank of Finland (Health 2000 and Health 2011 studies). GWAS6 comprised Affymetrix SNP6.0 array data on 341 AML cases of European ancestry recruited to The Cancer Genome Atlas (TCGA) project (phs000178/GRU)[7] or via hematology clinics at the University of Michigan Comprehensive Cancer Center[8–10]. For controls, we used Affymetrix SNP6.0 array data on 1395 healthy individuals of European ancestry from the GAIN: Genome-Wide Association Study of Schizophrenia project (phs000021/GRU).

Collection of patient samples and associated clinico-pathological information was undertaken with written informed consent. All studies were conducted in accordance with the Declaration of Helsinki and received local institutional review board or national research ethics approval. For GWAS5, ethical approval was granted by the Finnish Hematology Registry and Clinical Biobank (FHRB) and by the THL biobank (BB2018_63). For GWAS6, AML cases recruited via the University of Michigan Comprehensive Cancer Center was approved by the University of Michigan Institutional Review

Board (IRBMED #2004-1022). Access to the TCGA AML cases and GAIN controls was approved by the National Institute of Health (#9683). Information on ethical approvals for all other studies has been reported previously[6]

Genotyping and genome-wide quality-control procedures

Genotype calling was performed using Illumina GenomeStudio software or Affymetrix Genotyping Console software v4.2.0.26. Data handling and analysis was performed using R v4.2.1, PLINK v1.9b4.4 and SNPTEST v2.5.6. Rigorous variant and sample quality control metrics were applied to all six GWAS (Supplementary Figure 1). Specifically, we excluded variants with a call rate less than 95%, with departure from Hardy-Weinberg equilibrium (HWE; $P<10^{-3}$) or with significant differences ($P<0.05$) in missingness between cases and controls. Individual samples with a call rate of <95% or with extreme heterozygosity rates (+/- 3 standard deviation) were also excluded. Individuals were removed with estimated relatedness pihat >0.1875, both within and across GWAS. Ancestry was assessed using principal component analysis and super-populations from the 1000 genomes project as a reference, with individuals of non-European ancestry excluded based on the first two principal components (Supplementary Figures 2, 3 and 4)[6].

The majority of AML cases were genotyped using DNA extracted from cell/tissue samples (blood and bone marrow) taken during AML remission. For GWAS5, AML cases were primarily genotyped using DNA extracted from cell/tissue samples (blood and bone marrow) taken during disease presentation. We employed a stringent HWE cut-off in order to eliminate SNPs potentially affected by somatic copy number alterations. Furthermore, we also used Nexus Copy Number v10 (Bionano Genomics, California) data from 351 AML cases genotyped using samples with high somatic cell content to interrogate Log R ratio and B allele frequency at loci carrying risk variants for AML susceptibility. We found no significant evidence of somatic alterations affecting the risk loci at 11q13.2 (rs11421) or 6p21.32 (rs3916765)[6]. For the risk locus at 2p23.3 (rs4665765; *EFR3B*, *POMC*, *DNMT3A*, *DNAJC27*) we identified 4 cases with deletions and one case with gain. For the risk locus at 1q23.2 (rs12078864; *DUSP23*) we identified 3 cases with trisomy 1 and one case with a 1q gain. For the risk locus at 2p21 (rs79918355; *EPCAM*) we identified one case with a gain and one case with a deletion. For the risk variant at 2q33.3 (rs12988876; *PARD3*) we identified 1 case with a deletion. These data suggest that significant genotyping errors due to somatically acquired allelic imbalance in AML cases are unlikely.

Imputation, genome-wide association testing and meta-analysis

Genome-wide imputation was performed using the Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html) and the Haplotype Reference Consortium reference haplotype panel (http://www.haplotype-reference-consortium.org/) following pre-phasing

using ShapeIT (v2.r790)[11]. All variants with an INFO score <0.6 or a MAF <0.02 were excluded from subsequent analysis.

For each GWAS, association tests were performed for all cases (pan-AML), cytogenetically normal AML, complex karyotype AML and AML with deletions of chromosome 5 and/or 7, assuming an additive genetic model with nominally significant (P<0.05) principal components included in the analysis as covariates. Association summary statistics were combined for all six GWAS, in fixed effects models. Individual GWAS with less than 50 cases in any sub-group were excluded from subtype specific analyses. Cochran's Q statistic was used to test for heterogeneity and the I^2 statistic was used to quantify variation due to heterogeneity. Conditional analysis was conducted using the GCTA conditional and joint analysis (COJO) pipeline v1.94.4. A stepwise model selection with a GWAS P value cut-off of $5x10^5$ and collinearity cutoff of 0.9 was used to select independent associations from the summary statistics.

Fine mapping and functional annotation of causal SNPs

Sum of single effects (SuSiE) model was used in conjugation with PolyFun, to incorporate functional annotations as precomputed prior causal probabilities to improve statistical fine-mapping accuracy. Five GWAS significant loci from three AML subtypes were run through the pipeline to deduce 95% credible sets assuming a maximum of five causal variants per locus (K=5). In-sample LD was calculated from the controls (N=12938) using LDstore 2.0. The SNP2GENE tool within the FUMA pipeline (https://fuma.ctglab.nl) was used for functional annotation of fine-mapped causal variants, which included positional mapping, functional consequences on genes using ANNOVAR and chromatin interaction mapping.

Relationship between AML susceptibility variant genotype and patient survival

The relationship between AML risk variants and overall survival was evaluated in a total of 1725 AML patients (excluding acute promyelocytic leukemia) from the UK[12,13], Germany[14,15], Hungary, Finland (https://www.fhrb.fi/) and the United States[7–10]. Patients were treated with conventional intensive AML therapy including ara-C, daunorubicin and best supportive care. A subset of high-risk patients in the German cohort were treated with stem cell transplantation[15]. Overall survival was defined as the time from diagnosis to the date of last follow-up or death from any cause. Cox regression analysis was used to estimate allele specific hazard ratios and 95% confidence intervals. In order to control for index (collider) bias potentially introduced during the selection of cases for the survival analysis, we applied the corrected weighted least squares (CWLS) method[16], which uses the slope of a weighted regression of prognostic variants on risk variant associations as the bias correction factor to re-estimate the effects of variants on disease progression. We used 52408 LD pruned

(r2<=0.1, 250 kb SNP window) post-imputation variants with ≥0.98 imputation score as independent instrument variables from the GWAS summary statistics.

Interrogation of previously reported CH/CHIP variants as AML susceptibility variants

A total of 59 statistically significant ($P<5x10^{-8}$) variants were identified from published GWAS analyses of clonal haematopoiesis (CH), clonal haematopoiesis of indeterminate potential (CHIP) and their subtypes stratified by the two primary CH genes, *DNMT3A* and *TET2*. From Kar et al. (2022)[17], 15 independent lead variants that replicated in the UK Biobank cohort and that retained significance after conditioning on the lead variants were identified from analyses of pan-CH, *DNMT3A* mutation-positive CH and *TET2* mutation-positive CH. From Kessler et al (2022)[18], 52 lead variants were identified by LD thresholding and then by replication in the Geisinger MyCode Community Health Initiative (GHS) cohort from analyses of pan-CHIP, *DNMT3A* mutation-positive CHIP and *TET2* mutation-positive CHIP. We also included the independent lead variant at the *TCL1A* locus (rs2887399) from the CHIP GWAS reported by Weinstock et al (2023)[19]. Pan-AML and cytogenetically normal AML association *P* values were adjusted for multiple testing using Bonferroni correction.

We also interrogated the lead AML susceptibility variants identified in our study in the CH/CHIP datasets reported Kar *et al* (2022)[17] and Kessler *et al* (2022)[18].

Results

Meta-analysis of AML genome-wide association studies (GWAS)

We conducted 6 independent genome-wide association studies with AML cases and controls of European ancestry (GWAS1-6), four of which (GWAS1-4) have been reported previously[6]. A total of 4710 AML cases and 12938 healthy controls passed the study level quality control (Supplementary Figures 1-4) with common autosomal single nucleotide variants numbering between 250880 to 436068 genotyped across the 6 GWAS. We further improved the genomic resolution by imputing >7 million variants using the Haplotype Reference Consortium panel[20]. After excluding variants with an INFO (imputation quality) score of <0.6 and a minor allele frequency (MAF) <0.02, association tests were conducted for 5646403 autosomal variants common to all 6 GWAS. Considering the genetic and biological heterogeneity of AML, we calculated odds ratios (OR) for all AML cases (pan-AML)(N=4710) and three major AML subtypes; cytogenetically normal AML (N=1580), complex karyotype AML (N=358) and AML with deletions affecting chromosome 5 and/or chromosome 7 (del(5/7) AML)(N=319)[21]. Nominally significant ($P<0.05$) principal components in each GWAS were used as covariates to control inflation of the test statistic ($0.99>\lambda_{GC}<1.07$) for each analysis (Supplementary Figures 5–8).

Meta-analysis of six GWAS identified the previously reported signal for pan-AML at 11q13.2 (*KMT5B, CHKA, ALDH3B1, NDUFS8, TCIRG1*)[6], although with a new lead variant at this locus (rs11481 (MAF 0.34, INFO scores 0.83-0.96), $P = 3.58 \times 10^{-8}$) located 100 kb centromeric to the previous lead variant (rs4930561) (r^2=0.26) (Figures 1 and 2)[6]. Meta-analysis also identified a new signal for pan-AML surpassing genome-wide significance at 2p23.3 (rs4665765 (MAF 0.46, INFO scores 0.98-0.99); $P = 1.35 \times 10^{-8}$; *EFR3B, POMC, DNMT3A, DNAJC27*) (Figures 1 and 2). Meta-analysis of cytogenetically normal AML across six GWAS studies identified the previously reported signal at 6p21.32 (rs3916765 (MAF 0.12, INFO scores 0.90-0.99), *HLA-DQA2*)[6] (Figures 1 and 2), which has also been validated in an independent study[22]. However, our analysis did not reveal any new signals for cytogenetically normal AML surpassing the threshold for genome-wide significance. Meta-analysis of GWAS with sufficient cases in each sub-group identified genome-wide significant association signals for del(5/7) AML at 1q23.2 (rs12078864 (MAF 0.31, INFO score 0.83-0.96), $P = 7.0 \times 10^{-10}$, *DUSP23*) and for complex karyotype AML at 2p21 (rs79918355 (MAF 0.028, INFO score 0.62-0.96), $P = 1.6 \times 10^{-9}$, *EPCAM*) and 2q33.3 (rs12988876 (MAF 0.042, INFO score 0.82-0.94), $P = 3.28 \times 10^{-8}$, *PARD3B*) (Figures 1 and 2).

There was no significant evidence of heterogeneity (P<0.05) for association with AML for any of the risk variants across the GWAS included in each meta-analysis (Figure 2). Analysis conditioning on the top variant at each susceptibility locus did not identify any evidence of additional associations ($P<10^{-4}$), with the exception of the signal at 6p21.32 (*HLA-DQA2*) for cytogenetically normal AML (rs1794275, $P=1.92 \times 10^{-5}$) (Supplementary Figures 9-14).

Statistical fine-mapping of AML association signals

To determine the most credible causal variant at each association signal we conducted statistical fine-mapping using sum of single effects (SuSiE) model incorporating functional annotations of variants as prior probabilities to improve fine-mapping accuracy.

Fine-mapping indicated one 95% credible set for the pan-AML signal at 2p23.3, which captured the lead variant rs4665765 at this locus (OR 1.16, 95% CI 1.10-1.22; $P=1.35 \times 10^{-8}$) in high LD ($r^2>0.8$) with rs2164808 (OR 1.16, 95% CI 1.10-1.22; $P=1.38 \times 10^{-8}$). Based on the posterior inclusion probability (PIP), rs2164808 (PIP=0.75) was considered the most likely causal SNP over the lead SNP (PIP=0.24) at 2p23.3 (Supplementary Figure 15a, Supplementary Table 7). rs2164808 maps to exon 23 of *EFR3B* and is a nonsense variant with a CADD score of 16.5 (Supplementary Table 1). Of note, this variant is located within 200 kb of *DNMT3A* (Figure 2), which is frequently somatically mutated in AML[1,23]. Interrogation of data from the eQTLGen consortium identified rs2164808 as a significant *cis*-expression quantitative trait locus (eQTL) for *DNMT3A* ($P=2.867 \times 10^{-4}$) as well as *DNAJC27* ($P=7.27 \times 10^{-28}$), *CENPO* ($P=3.09 \times 10^{-5}$), *POMC* ($P=1.98 \times 10^{-3}$) and *ADCY3* ($P=3.66 \times 10^{-3}$)

(Supplementary Table 2), suggesting that this variant (or genetically linked variants in linkage disequilibrium) affects expression of numerous local genes.

Fine mapping of the pan-AML signal at 11q13.2 indicated one credible set, with lead variant rs11481 (OR 1.17, 95% CI 1.11-1.24; P=$3.58x10^{-8}$) also identified as the most credible causal variant (PIP=0.87) (Supplementary Figure 15b, Supplementary Table 7). rs11481 maps to *RP11-802E16.3* (ENSG00000255031), a noncoding natural antisense transcript (ncNAT) to *CHKA* with a CADD score of 11.2 (Supplementary Table 1). rs11481 is a significant *cis*-eQTL for *ALDH3B1* (*P*=$4.86x10^{-44}$), *RP5-901A4.1* (*P*=$4.56x10^{-41}$), *RP11-802E16.3* (*P*=$1.84x10^{-20}$) *CHKA* (*P*=$7.14x10^{-11}$), *NDUFS8* (*P*=$4.05x10^{-10}$), *DOC2GP* (*P*=$3.01x10^{-5}$), *TBC1D10C* (*P*=$3.47x10^{-3}$) and *MRPL21* (*P*=$4.45x10^{-2}$) (Supplementary Table 3).

Two variants were included in the 95% credible set for the complex karyotype signal at 2p21. The lead variant at this locus (rs79918355, OR 2.80, 95% CI 2.0-3.91; P=$1.6x10^{-9}$) was implicated as the most credible causal variant (PIP=0.91) (Supplementary Figure 16a, Supplementary Table 7), which localises to an intronic region of the *AC073283.4* long non-coding RNA and is a significant *cis*-eQTL for *MCFD2* (*P*=$1.04x10^{-5}$) and *MSH2* (*P*=$7.23x10^{-3}$)(Supplementary Table 4).

The complex karyotype AML signal at 2q33 fine-mapped to one credible set only including the lead variant at this locus (rs12988876, OR 2.29, 95% CI 1.71-3.08; *P*=$3.28x10^{-9}$, PIP=0.99) (Supplementary Figure 16B), which maps intronic to and is a significant eQTL for *PARD3B* (*P*=0.03) (Supplementary Table 5).

One 95% credible set was deduced from the del(5/7) AML association signal on chromosome 1q23.2, capturing the top 8 SNPs at this locus. The lead variant at this locus (rs12078864, OR 1.72, 95% CI 1.45 – 2.04; P=$7x10^{-10}$) was identified as the most credible causal variant (PIP=0.46) (Supplementary Figure 17 and Table 7) and is a significant *cis*-eQTL for *DUSP23* (*P*=$7.99x10^{-84}$), *SLAMF8* (*P*=$3.56x10^{-11}$), *PEX19* (*P*=$1.82x10^{-3}$), *DARC* (*P*=$6.01x10^{-3}$), *FCRL6* (*P*=$1.49x10^{-2}$), *C1orf204* (*P*=$2.23x10^{-2}$), *FCER1A* (*P*=$2.23x10^{-2}$), *RP11-404F10.2* (*P*=$4.78x10^{-2}$) and *CD84* (*P*=$4.91x10^{-2}$)(Supplementary table 6).

Reported risk variants for clonal hematopoiesis (CH), Clonal haematopoiesis of indeterminate potential (CHIP) and their association with AML.

Clonal haematopoiesis (CH) is an age-related non-malignant condition defined by the expansion of haematopoietic stem cells (HSC) and progenitor cells in healthy individuals following the acquisition of somatic driver mutations. Clonal haematopoiesis of indeterminate potential (CHIP) is CH driven by a somatic mutation in a gene recurrently mutated in myeloid malignancy (variant allele frequency ≥0.02), with *DNMT3A*, *TET2* and *ASXL1* being the most commonly affected[24]. The presence of CH

identifies individuals with an increased risk of developing AML[25] and recent studies have reported constitutional genetic variants associated with risk of developing CH[17–19,26,27].

A total of 59 CH/CHIP risk variants reported by Kar et al. 2022[17], Kessler et al. 2022[18] and Weinstock et al. 2023[19] were annotated in our AML GWAS. Of these, 7 CH/CHIP variants were significantly associated with risk of either pan-AML or CN-AML after correction for multiple testing, including variants at the *TERT* locus (rs2736100, rs2853677, rs7705526), the *ATM* locus (rs10890839, rs11212666, rs228606) and the *MSI2* locus (rs188761458) (Figure 4). The risk of AML for all 7 variants was in the same direction as the reported risk of CH/CHIP. A further 18 reported CH/CHIP variants were nominally significantly associated with risk of either pan-AML or CN-AML ($P<0.05$), but these did not retain significance after correction for multiple testing (Supplementary Table 8). For 14 of these 18 variants, the risk of AML was in the same direction as the reported risk of CH/CHIP (Supplementary Table 8). The remaining 4 variants were all at the *TCL1A* locus and the reported risk of CH/CHIP was in opposing directions for *TET2*-mutated and *DNMT3A*-mutated CH/CHIP[17,18]. However, the risk of AML was in the same direction as the risk of *TET2*-mutant CH/CHIP for all 4 variants (Supplementary Table 8). In summary, 7 risk variants for CH/CHIP were significantly associated with AML, providing further evidence for shared genetic susceptibility between these conditions.

All 6 AML susceptibility variants reported in our study were annotated in the CH/CHIP datasets published by Kar et al (2022)[17] and/or Kessler et al (2022)[18], but none were significantly associated with risk of CH/CHIP after multiple testing correction (Supplementary Table 9). However, one variant at the 11q13.2 risk locus for pan-AML (rs11481, *CHKA*) was nominally significantly associated with risk of CH (OR 1.04, 95% CI 1.00-1.07, $P=0.024$)[17], but this did not retain significance after correction for multiple testing (Supplementary Table 9).

Risk variants for AML and their impact on patient survival

The relationship between the identified AML risk variants and survival was evaluated in 1725 AML patients from the UK, Germany, Hungary, Finland and the United States of America. Five of the six AML susceptibility variants identified via GWAS (rs11481 (11q13.2, *CHKA*), rs3916765 (6p21.32, *HLA-DQA2*), rs12078864 (1q23.2, *DUSP23*), rs79918355 (2p21, *EPCAM*), rs12988876 (2q33.3, *PARD3*)) were not significantly associated with overall survival in univariate analysis (Supplementary Figure 18) prior to or after index bias correction (Supplementary Figure 18). However, the pan-AML risk variant at 2p23.3 (rs4665765, *EFR3B, POMC, DNMT3A, DNAJC27*) was nominally significantly associated with overall survival (HR 1.13, 95% CI 1.04-1.24, $P=6.09\times10^{-3}$), with the risk allele for AML being associated with inferior outcome. This variant retained significance with an increased

effect size after index bias correction (HR 1.25, 95% CI 1.14-1.38, P=3.89x10^{-6}) and after adjustment for multiple testing (P=2.33x10^{-5}) (Supplementary Figure 18).

Discussion

By conducting a meta-analysis of six independent genome-wide association studies we report two pan-AML susceptibility loci and four AML sub-group-specific loci, two of which are reported previously[6]. We used statistical and functional fine-mapping methods to identify the most credible causal variant at each locus and were able to prioritize high confidence genes which could serve as strong candidates for functional validation experiments.

We explored three genes localised to the pan-AML susceptibility signal at 2p23.3. The most credible causal variant (rs2164808) at this locus introduces a premature stop codon in *EFR3B*. The *EFR3A/B* family are paralogous proteins that contribute to AT1 signalling regulating G-protein-coupled receptors[28]. EFR3B and EFR3A also form a complex to recruit phosphatidylinositol 4-kinase (PI4K) to the plasma membrane, with high expression of PI4K associated with inferior survival in myeloid leukemia[29]. The lead variant and most credible causal variant at this locus are both *cis*-eQTL for *DNAJC27* (*RBJ*), which encodes small GTPase that promotes development of numerous human cancers via MEK/ERK signalling[30–32]. Of note, the most credible causal variant is located within 200 Kb of *DNMT3*, encoding a DNA methyltransferase that is frequently somatically dysregulated in AML[1], where loss of function disrupts global genomic methylation in hematopoietic progenitor cells leading to leukemogenesis[23,33]. Intriguingly, the AML risk variant is also significantly associated with inferior overall survival.

The new lead variant at the 11q13.2 pan-AML association signal (rs11481) localises to the *CHKA* gene and is in modest genetic linkage with the variant (rs4930561) (linkage disequilibrium r^2=0.26) previously reported to be associated with pan-AML localised to the *KMT5B* gene[6]. Analysis conditioning on rs11481 did not identify any evidence of additional associations ($P<10^{-4}$) at this locus, suggesting that both variants are part of the same association signal. rs11481 maps to a noncoding exonic region of *CHKA*, and specifically to a noncoding natural antisense transcript (ncNAT) *RP11-802E16.3* (ENSG00000255031). Although ncNATs are regulatory RNA molecules that modulate cellular processes such as growth and differentiation, their role in cancer pathogenesis remains unknown. However, there is evidence that *RP11-802E16.3* regulates expression of *CHKA*[34]. Consistent with this model, rs11481 is a significant *cis*-eQTL for *CHKA* ($P=7.14\times10^{-11}$), where high expression drives tumour progression and metastasis of several human cancer[35–37]. Moreover, expression of *CHKA* is implicated in the pathogenesis of B-cell malignancies and T-ALL via promotion of cell survival and proliferation[38]. rs11481 is also a significant *cis*-eQTL for *ALDH3B1* ($P=4.86\times10^{-44}$) and *NDUFS8* ($P=4.05\times10^{-10}$). *ALDH3B1* encodes a member of the aldehyde dehydrogenase superfamily that protects cells from oxidative stress by catalysing the reversible oxidation of endogenous and exogenous aldehydes[39]. High expression of aldehyde dehydrogenase family members is associated with chemotherapy resistance and inferior survival in AML[40].

Furthermore, AML stem cells are acutely sensitive to small molecule inhibitors of aldehyde dehydrogenases, identifying this family of enzymes as a therapeutic vulnerability in AML[41]. *NDUFS8* encodes a subunit of the mitochondrial NADH:ubiquinone oxidoreductase complex I, responsible for NADH oxidation as part of the respiratory chain[42]. Rare constitutional variants in *NDUFS8* are associated with attenuated mitochondrial respiration in AML cells and are mutually exclusive with somatically acquired mutations in isocitrate dehydrogenase 1 (*IDH1*)[43], suggesting two alternative genetic mechanisms via which mitochondrial function is dysregulated in AML pathogenesis. Although rs11481 is not a *cis*-eQTL for *KMT5B* (*SUV420H1*), a role for this lysine methyltransferase cannot be excluded. *KMT5B* is frequently silenced via hypermethylation in numerous human cancers[44–46] and somatic mutation is reported in transformation of myelodysplastic syndrome to AML[47].

We report two signals significantly associated with risk of complex karyotype AML, at 2p21 and 2q33. The most likely causal variant at 2p21 (rs79918355) is a *cis*-eQTL for *MCFD2*, which encodes a protein that, along with LMAN1, forms a cargo receptor complex for transport of coagulation factors[48]. Rare constitutional variants in *MCFD2* cause combined deficiency of coagulation factor V and VIII, a recessive bleeding disorder[49]. MCFD2 also has a role in the regulation of stem cell survival and pluripotency[50,51] and somatic mutations have been reported in leukemic cells from Fanconi anemia patients who developed AML[52], a disease which often presents with a complex karyotype due to inherent chromosome instability[53]. The lead variant at the 2q33 signal for complex karyotype AML (rs12988876) maps intronic to *PARD3B* and is an eQTL for this gene (P=0.03). *PARD3B* encodes a member of the PARD3 family of proteins that regulate cell polarity and centrosome localization[54]. PARD3B is a homologue of PARD3, which also functions as a scaffolding protein that interacts with numerous intracellular signalling molecules, many of which are dysregulated in cancer, including members of the PI3K/AKT and MAPK pathways such as PTEN and JNK[54]. PARD3B is implicated in prostate cancer aetiology and expression levels have been associated with survival in both colorectal and breast cancer [55,56], suggesting a potential role in numerous human cancers.

The lead variant at the 1q23.2 signal for del(5/7) AML is a significant *cis*-eQTL for *DUSP23* and *SLAMF8*. *DUSP23* encodes a dual specific phosphatase that regulates MAP kinase signalling, impacting cell proliferation, growth and survival[57]. DUSP23 also plays a role in regulating cell adhesion/migration[58] and high expression in blast cells is as an independent prognostic marker for inferior survival in AML[59]. High DUSP23 expression is also reported in $CD4^+$ T-cells from patients with systemic lupus erythematosus, where it is thought to regulate DNA methyltransferase activity, including *DNMT3A*, which is frequently dysregulated in AML via somatic mutation[1,23]. *SLAMF8* encodes a cell surface glycoprotein that is a member of the signalling lymphocytic activation molecule (SLAM) family involved in regulating the development and function of a wide range of

immune cells, such as T lymphocytes, B cells, neutrophils, dendritic cells, macrophages and eosinophils[60,61]. *SLAMF8* is upregulated in AML with *KMT2A* (*MLL*) gene partial tandem duplication (an alteration reported in AML) and knockdown significantly decreased leukemic cell growth[62], implicating *SLAMF8* as having oncogenic function in AML.

We identify genetic variants at 5p15.3, 11q23 and 17q22 previously associated with risk of CH/CHIP as also being associated with risk of AML, with functionality predicted to operate via effects on local genes *TERT*, *ATM* and *MSI2*, respectively[17,18,63].

In conclusion, we performed a genome-wide meta-analysis incorporating six AML GWAS of European ancestry and report four new susceptibility loci for pan-AML or subtype specific disease. We also identify a common variant at 2p23.3 that significantly associates with patient survival and several genetic variants that associate with both CH/CHIP and AML. Functional interrogation is warranted to decipher the molecular mechanisms by which the loci identified in this study modify AML risk and patient outcome.

Acknowledgements

This work was funded by Blood Cancer UK (to JMA; #06002 and #13044) and the Medical Research Council DiMeN Doctoral Training Programme. We are grateful to the Finnish Hematology Registry and Clinical Biobank and The National Institute for Health and Welfare (THL) Biobank of Finland for providing access to AML patient data and samples.

Author Contributions

DS and W-YL collated data, conducted data analysis and drafted the manuscript. SEF, AA, NS, CE, MHN, YX, CPa, EH, AQ, KS, CG, AS, TR, LS-C, THah, AIC-G, GLJ, HJM, GHJ, TM, MCo, PM, AI, RKH, AKB, NHR, JF, RAL, MMLB, WS, OH, AE, DJ, ZH, HH, DN, BD, AP, ITB, DJA, RSH, AC, JN, AMD, CL, AKD, LP, KPi, SJ, MB, CR, HA, LR, DK, LW, HJC, RD, MKA, MCF, GMart, GMarc, MAS, JC, IG-S, TC, CM, SR, HS, MTV, HD, MCh, CPr, REG, DL, JWe, AM, DN, W-KH, AG, KPo, JDMF, RK, JWa, MM, THaf, SK, CB, BP, SNM, FS and KO collated data and/or advised on data analysis. JMA collated data, analysed data, directed the research, obtained funding and drafted the manuscript. All authors approved the final version of the manuscript.

Conflict of Interest

The authors declare no conflicts of interest.

References

1. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016;374(23):2209-2221. doi:10.1056/NEJMoa1516192

2. Godley LA, Shimamura A. Genetic predisposition to hematologic malignancies: management and surveillance. *Blood*. 2017;130(4):424-432. doi:10.1182/blood-2017-02-735290

3. Rio-Machin A, Vulliamy T, Hug N, et al. The complex genetic landscape of familial MDS and AML reveals pathogenic germline variants. *Nat Commun*. 2020;11(1):1044. doi:10.1038/s41467-020-14829-5

4. Fletcher O, Houlston RS. Architecture of inherited susceptibility to common cancer. *Nat Rev Cancer*. 2010;10(5):353-361. doi:10.1038/nrc2840

5. Khera A V, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*. 2018;50(9):1219-1224. doi:10.1038/s41588-018-0183-z

6. Lin WY, Fordham SE, Hungate E, et al. Genome-wide association study identifies susceptibility loci for acute myeloid leukemia. *Nat Commun*. 2021;12(1):6233. doi:10.1038/s41467-021-26551-x

7. The Cancer Genome Atlas Research Network. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *New England Journal of Medicine*. 2013;368(22):2059-2074. doi:10.1056/NEJMoa1301689

8. Parkin B, Ouillette P, Wang Y, et al. NF1 inactivation in adult acute myelogenous leukemia. *Clinical Cancer Research*. 2010;16(16):4135-4147. doi:10.1158/1078-0432.CCR-09-2639

9. Parkin B, Erba H, Ouillette P, et al. Acquired genomic copy number aberrations and survival in adult acute myelogenous leukemia. *Blood*. 2010;116(23):4958-4967. doi:10.1182/blood-2010-01-266999

10. Parkin B, Ouillette P, Yildiz M, Saiya-Cork K, Shedden K, Malek SN. Integrated genomic profiling, therapy response, and survival in adult acute myelogenous leukemia. *Clinical Cancer Research*. 2015;21(9):2045-2056. doi:10.1158/1078-0432.CCR-14-0921

11. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2011;9(2):179-181. doi:10.1038/nmeth.1785

12. Goldstone AH, Burnett AK, Wheatley K, et al. Attempts to improve treatment outcomes in acute myeloid leukemia (AML) in older patients: the results of the United Kingdom Medical Research Council AML11 trial. *Blood*. 2001;98(5):1302-1311. https://www.ncbi.nlm.nih.gov/pubmed/11520775

13. Burnett AK, Hills RK, Milligan DW, et al. Attempts to optimize induction and consolidation treatment in acute myeloid leukemia: results of the MRC AML12 trial. *J Clin Oncol*. 2010;28(4):586-595. doi:10.1200/JCO.2009.22.9088

14. Schaich M, Parmentier S, Kramer M, et al. High-dose cytarabine consolidation with or without additional amsacrine and mitoxantrone in acute myeloid leukemia: Results of the prospective randomized AML2003 trial. In: *Journal of Clinical Oncology*. Vol 31. 2013. doi:10.1200/JCO.2012.46.4743

15. Röllig C, Bornhäuser M, Kramer M, et al. Allogeneic stem-cell transplantation in patients with NPM1-mutated acute myeloid leukemia: Results from a prospective donor versus no-donor analysis of patients after upfront HLA typing within the SAL-AML 2003 trial. *Journal of Clinical Oncology*. 2015;33(5). doi:10.1200/JCO.2013.54.4973

16. Cai S, Hartley A, Mahmoud O, Tilling K, Dudbridge F. Adjusting for collider bias in genetic association studies using instrumental variable methods. *Genet Epidemiol*. 2022;46(5-6):303-316. doi:10.1002/gepi.22455

17. Kar SP, Quiros PM, Gu M, et al. Genome-wide analyses of 200,453 individuals yield new insights into the causes and consequences of clonal hematopoiesis. *Nat Genet*. 2022;54(8):1155-1166. doi:10.1038/s41588-022-01121-z

18. Kessler MD, Damask A, O'Keeffe S, et al. Common and rare variant associations with clonal haematopoiesis phenotypes. *Nature*. 2022;612(7939):301-309. doi:10.1038/s41586-022-05448-9

19. Weinstock JS, Gopakumar J, Burugula BB, et al. Aberrant activation of TCL1A promotes stem cell expansion in clonal haematopoiesis. *Nature*. 2023;616(7958):755-763. doi:10.1038/s41586-023-05806-1

20. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48(10):1279-1283. doi:10.1038/ng.3643

21. Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood*. 2022;140(12):1345-1377. doi:10.1182/blood.2022016867

22. Laflamme R, Lisi V, Hébert J, et al. Replication of a GWAS signal near HLA-DQA2 with AML using a disease-only cohort and external population-based controls. *Blood Neoplasia*. Published online May 2025:100118. doi:10.1016/j.bneo.2025.100118

23. Ley TJ, Ding L, Walter MJ, et al. *DNMT3A* mutations in acute myeloid leukemia. *New England Journal of Medicine*. 2010;363(25):2424-2433. doi:10.1056/NEJMoa1005143

24. Asada S, Kitamura T. Clonal hematopoiesis and associated diseases: A review of recent findings. *Cancer Sci*. 2021;112(10):3962-3971. doi:10.1111/cas.15094

25. Young AL, Tong RS, Birmann BM, Druley TE. Clonal hematopoiesis and risk of acute myeloid leukemia. *Haematologica*. 2019;104(12):2410-2417. doi:10.3324/haematol.2018.215269

26. Zink F, Stacey SN, Norddahl GL, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood*. 2017;130(6):742-752. doi:10.1182/blood-2017-02-769869

27. Stacey SN, Zink F, Halldorsson GH, et al. Genetics and epidemiology of mutational barcode-defined clonal hematopoiesis. *Nat Genet*. 2023;55(12):2149-2159. doi:10.1038/s41588-023-01555-z

28. Bojjireddy N, Guzman-Hernandez ML, Reinhard NR, Jovic M, Balla T. EFR3s are palmitoylated plasma membrane proteins that control responsiveness to G protein-coupled receptors. *J Cell Sci*. Published online January 1, 2014. doi:10.1242/jcs.157495

29. Jiang X, Huang X, Zheng G, et al. Targeting PI4KA sensitizes refractory leukemia to chemotherapy by modulating the ERK/AMPK/OXPHOS axis. *Theranostics*. 2022;12(16):6972-6988. doi:10.7150/thno.76563

30. Wang Y, Shen X, Wang Q, et al. Non-canonical small GTPase RBJ promotes NSCLC progression through the canonical MEK/ERK signaling pathway. *Curr Pharm Des*. 2022;28(42):3446-3455. doi:10.2174/1381612829666221117124048

31. Chen T, Yang M, Yu Z, et al. Small GTPase RBJ mediates nuclear entrapment of MEK1/MEK2 in tumor progression. *Cancer Cell*. 2014;25(5):682-696. doi:10.1016/j.ccr.2014.03.009

32. Liu Q, Zhu H, Zhang C, Chen T, Cao X. Small GTPase RBJ promotes cancer progression by mobilizing MDSCs via IL-6. *Oncoimmunology*. 2017;6(1):e1245265. doi:10.1080/2162402X.2016.1245265

33. Yang L, Rau R, Goodell MA. DNMT3A in haematological malignancies. *Nat Rev Cancer*. 2015;15(3):152-165. doi:10.1038/nrc3895

34. Bellido Molias F, Sim A, Leong KW, et al. Antisense RNAs influence promoter usage of their counterpart sense genes in cancer. *Cancer Res*. 2021;81(23):5849-5861. doi:10.1158/0008-5472.CAN-21-1859

35. Zhang L, Chen P, Yang S, et al. CHKA mediates the poor prognosis of lung adenocarcinoma and acts as a prognostic indicator. *Oncol Lett*. 2016;12(3):1849-1853. doi:10.3892/ol.2016.4810

36. Asim M, Massie CE, Orafidiya F, et al. Choline kinase alpha as an androgen receptor chaperone and prostate cancer therapeutic target. *J Natl Cancer Inst*. 2016;108(5):djv371. doi:10.1093/jnci/djv371

37. Hu L, Wang RY, Cai J, et al. Overexpression of CHKA contributes to tumor progression and metastasis and predicts poor prognosis in colorectal carcinoma. *Oncotarget*. 2016;7(41):66660-66678. doi:10.18632/oncotarget.11433

38. Gokhale S, Xie P. ChoK-Full of Potential: Choline Kinase in B Cell and T Cell Malignancies. *Pharmaceutics*. 2021;13(6):911. doi:10.3390/pharmaceutics13060911

39. Marchitti SA, Brocker C, Orlicky DJ, Vasiliou V. Molecular characterization, expression analysis, and role of ALDH3B1 in the cellular protection against oxidative stress. *Free Radic Biol Med*. 2010;49(9):1432-1443. doi:10.1016/j.freeradbiomed.2010.08.004

40. Dancik GM, Varisli L, Tolan V, Vlahopoulos S. Aldehyde dehydrogenase genes as prospective actionable targets in acute myeloid leukemia. *Genes (Basel)*. 2023;14(9):1807. doi:10.3390/genes14091807

41. Venton G, Pérez-Alea M, Baier C, et al. Aldehyde dehydrogenases inhibition eradicates leukemia stem cells while sparing normal progenitors. *Blood Cancer J*. 2016;6(9):e469-e469. doi:10.1038/bcj.2016.78

42. Xu W, Fang H, Cao X, et al. NADH:ubiquinone oxidoreductase core subunit S8 expression and functional significance in non-small cell lung cancer. *Cell Death Dis*. 2025;16(1):321. doi:10.1038/s41419-025-07638-5

43. Bassal MA, Samaraweera SE, Lim K, et al. Germline mutations in mitochondrial complex I reveal genetic and targetable vulnerability in IDH1-mutant acute myeloid leukaemia. *Nat Commun*. 2022;13(1):2614. doi:10.1038/s41467-022-30223-9

44. Bröhm A, Elsawy H, Rathert P, et al. Somatic cancer mutations in the SUV420H1 protein lysine methyltransferase modulate its catalytic activity. *J Mol Biol*. 2019;431(17):3068-3080. doi:10.1016/j.jmb.2019.06.021

45. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet*. 2013;14(11):765-780. doi:10.1038/nrg3554

46. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet*. 2016;17(5):284-299. doi:10.1038/nrg.2016.13

47. Mian SA, Rouault-Pierre K, Smith AE, et al. SF3B1 mutant MDS-initiating cells may arise from the haematopoietic stem cell compartment. *Nat Commun*. 2015;6(1):10004. doi:10.1038/ncomms10004

48. Zhang B, Kaufman RJ, Ginsburg D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. *Journal of Biological Chemistry*. 2005;280(27):25881-25886. doi:10.1074/jbc.M502160200

49. Guy JE, Wigren E, Svärd M, Härd T, Lindqvist Y. New insights into multiple coagulation factor deficiency from the solution structure of human MCFD2. *J Mol Biol*. 2008;381(4):941-955. doi:10.1016/j.jmb.2008.06.042

50. Liu H, Zhao B, Chen Y, et al. Multiple coagulation factor deficiency protein 2 contains the ability to support stem cell self-renewal. *The FASEB Journal*. 2013;27(8):3298-3305. doi:10.1096/fj.13-228825

51. Toda H, Tsuji M, Nakano I, et al. Stem cell-derived neural stem/progenitor cell supporting factor is an autocrine/paracrine survival factor for adult neural stem/progenitor cells. *Journal of Biological Chemistry*. 2003;278(37):35491-35500. doi:10.1074/jbc.M305342200

52. Chang L, Cui Z, Shi D, et al. Polyclonal evolution of Fanconi anemia to MDS and AML revealed at single cell resolution. *Exp Hematol Oncol*. 2022;11(1):64. doi:10.1186/s40164-022-00319-5

53. Merfort LW, Lisboa M de O, Cavalli LR, Bonfim CMS. Cytogenetics in Fanconi Anemia: The importance of follow-up and the search for new biomarkers of genomic instability. *Int J Mol Sci*. 2022;23(22):14119. doi:10.3390/ijms232214119

54. Atashrazm F, Ellis S. The polarity protein PARD3 and cancer. *Oncogene*. 2021;40(25):4245-4262. doi:10.1038/s41388-021-01813-6

55. Catterall R, Lelarge V, McCaffrey L. Genetic alterations of epithelial polarity genes are associated with loss of polarity in invasive breast cancer. *Int J Cancer*. 2020;146(6):1578-1591. doi:10.1002/ijc.32691

56. Li T, Liu D, Lei X, Jiang Q. Par3L enhances colorectal cancer cell survival by inhibiting Lkb1/AMPK signaling pathway. *Biochem Biophys Res Commun*. 2017;482(4):1037-1041. doi:10.1016/j.bbrc.2016.11.154

57. Gao PP, Qi XW, Sun N, et al. The emerging roles of dual-specificity phosphatases and their specific characteristics in human cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*. 2021;1876(1):188562. doi:10.1016/j.bbcan.2021.188562

58. Gallegos LL, Ng MR, Sowa ME, et al. A protein interaction map for cell-cell adhesion regulators identifies DUSP23 as a novel phosphatase for β-catenin. *Sci Rep*. 2016;6(1):27114. doi:10.1038/srep27114

59. Liu X, Zhuang H, Li F, Lu Y, Pei R. Prognostic Significance of Dual-Specificity Phosphatase 23 Expression in Acute Myeloid Leukemia. *J Blood Med*. 2024;Volume 15:35-50. doi:10.2147/JBM.S437400

60. Magari M, Nishioka M, Hari T, et al. The immunoreceptor SLAMF8 promotes the differentiation of follicular dendritic cell-dependent monocytic cells with B cell-activating ability. *FEBS Lett*. 2022;596(20):2659-2667. doi:10.1002/1873-3468.14468

61. Wang G, Abadía-Molina AC, Berger SB, et al. Cutting edge: Slamf8 Is a negative regulator of Nox2 activity in macrophages. *The Journal of Immunology*. 2012;188(12):5829-5832. doi:10.4049/jimmunol.1102620

62. Bill M, Goda C, Pepe F, et al. Targeting BRD4 in acute myeloid leukemia with partial tandem duplication of the MLL gene. *Haematologica*. 2021;106(9):2527-2532. doi:10.3324/haematol.2020.271627

63. Agarwal G, Antoszewski M, Xie X, et al. Inherited resilience to clonal hematopoiesis by modifying stem cell RNA regulation. BioRxiv. doi:10.1101/2025.03.24.645017

Figure Legends

Figure 1 – Manhattan plots from acute myeloid leukemia meta-analysis of genome-wide association studies. For each GWAS, association tests were performed for all AML cases (pan-AML), cytogenetically normal AML, del(5/7) AML and complex karyotype AML assuming an additive genetic model, with nominally significant principal components included in the analysis as covariates. Association summary statistics were combined for variants common to all GWAS, in fixed effects models using PLINK (GWAS5 was excluded from the meta-analysis of del(5/7) AML and cytogenetically complex AML due to low case numbers). Manhattan plots show negative $\log_{10}$ (fixed effects meta *P* values, *Y* axis) for pan-AML (a), cytogenetically normal AML (b), del(5/7) AML (c) and cytogenetically complex AML (d) over 22 autosomal chromosomes. Risk loci are annotated with chromosome position and local genes. All statistical tests were two-sided and no adjustments were made for multiple comparisons. The horizontal red line denotes the threshold for statistical significance in a genome-wide association study ($P < 5.0 \times 10^{-8}$).

Figure 2 – Forest plots for loci associated with acute myeloid leukemia. Study cohorts, sample sizes (case and controls (con)), imputation (info) score, effect allele, effect allele frequencies (EAF) and estimated odds ratios (OR) for rs11481 (pan-AML) (a), rs4665765 (pan-AML) (b), rs3916765 (cytogenetically normal AML) (c), rs12078864 (del(5/7) AML) (d), rs79918355 (complex karyotype AML) (e) and rs12988876 (complex karyotype AML) (f). The vertical line corresponds to the null hypothesis (odds ratio (OR) = 1). The horizontal lines and square brackets indicate 95% confidence intervals (95% CI). Areas of the boxes are proportional to the weight of the study. Diamonds represent combined estimates for fixed-effect and random-effect analysis. Cochran's Q statistic was used to test for heterogeneity where $P_{HET}>0.05$ indicates non-significant heterogeneity. The heterogeneity index, I^2 (0-100) was also measured which quantifies the proportion of the total variation due to heterogeneity. All statistical tests were two-sided and no adjustments were made for multiple comparisons.

Figure 3 – Regional association and linkage disequilibrium plots for loci associated with acute myeloid leukemia. For each GWAS, association tests were performed for pan-AML cases, cytogenetically normal AML, del(5/7) AML and cytogenetically complex AML assuming an additive genetic model, with nominally significant principal components included in the analysis as covariates. Association summary statistics were combined for variants common to all 6 GWAS, in fixed effects models using PLINK. Negative $\log_{10}$-transformed *P* values (left *Y* axis) from the meta-analysis of 6 GWAS are shown for variants at 11q13 (a) and 1p31.1 (b) for pan-AML, and at 6p21.32 (c) for cytogenetically normal AML. Negative $\log_{10}$-transformed *P* values (left *Y* axis) from the meta-analysis of 5 GWAS are shown for variants at 1q23.2 (d) for del(5/7) AML, and at 2p21 (e) and 2q33.3 (f) for cytogenetically complex AML. All statistical tests were two-sided and no adjustments

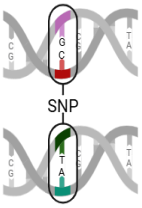
were made for multiple comparisons. The lead variant at each location is indicated by a purple diamond and the blue line shows recombination rate (right *Y* axis). All plotted variants were either directly genotyped or had an imputation score of >0.6 in all GWAS datasets. R^2 values were derived from the 1000 genomes project.

Figure 4 – Forest plots for variants reported to be associated with risk of clonal hematopoiesis (CH) and/or clonal hematopoiesis of indeterminate potential (CHIP) and their association with risk of AML. Forest plots show 7 variants reported by Kar *et al* (2022)[16] and Kessler *et al* (2022)[17] to be associated with risk of CH or CHIP at GWAS significance ($P < 5 \times 10^{-8}$) and their significant association with pan-AML and/or cytogenetically normal AML (CN-AML) after correction for multiple testing. SNP, Single nucleotide polymorphism; CHR, Chromosome; Gene, nearest gene; OA/EA, Non-effect allele/Effect allele; Case/Con, Number of cases/Number of controls; Trait, tested phenotype (CN-AML; CHIP, clonal hematopoiesis of indeterminate potential; CHIP-DNMT3A, DNMT3A mutation-positive CHIP; CHIP-TET2, TET2 mutation-positive CHIP; CH, clonal hematopoiesis; CH-DNMT3A, DNMT3A mutation-positive CH; CH-TET2, TET2 mutation-positive CH); Study, GWAS study; OR (95% CI), Odds ratio and 95% confidence intervals; *P* value (unadjusted), fixed effect *P* value of the association test; *P* value (adjusted), fixed effect *P* value adjusted for multiple testing using the Bonferroni method.

Role of common variation at 1q23.3, 2p23.3, 2q33.3 and 2p21 in risk of acute myeloid leukemia

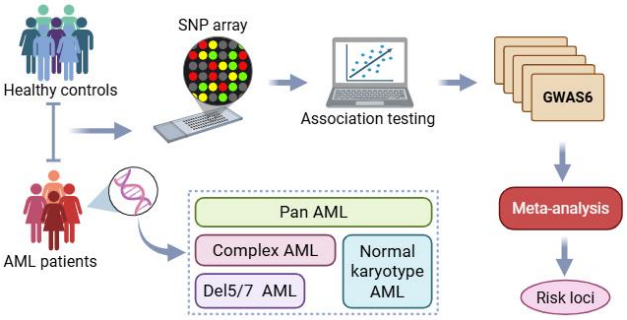
Context of Research

- Acute myeloid leukemia (AML) is a complex hematological malignancy with multiple disease sub-groups.
- The genetic risk of AML is unclear.



Patients and Methods

- 4710 AML patients and 12938 controls of European ancestry genotyped using SNP arrays.

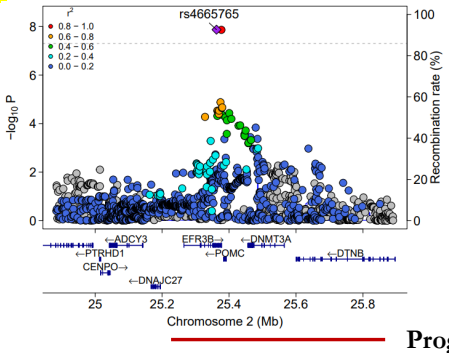


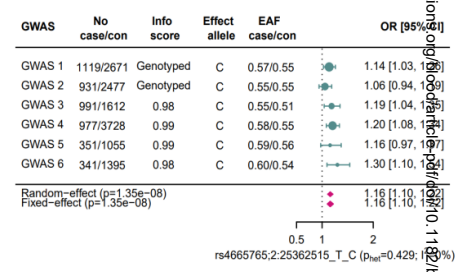
- We performed a meta-analysis of six GWAS to identify risk alleles for pan-AML and sub-groups.

Main Findings

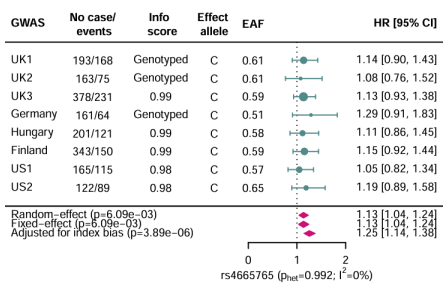
- We identify 4 new risk alleles for AML, including one close to *DNMT3A* that is also prognostic.

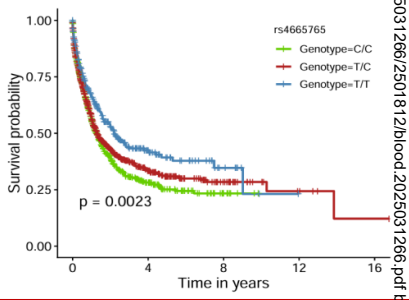
Risk





Prognosis





CONSLUSIONS: We identify a new genome-wide significant risk locus for pan-AML and three new risk loci for disease sub-groups, including AML with deletions of chromosome 5 and/or 7 and cytogenetically complex AML. We also identify variants previously associated with risk of clonal hematopoiesis (CH) that also associate with risk of AML. Our results further inform on AML etiology and demonstrate the existence of disease sub-group specific risk loci and shared genetic susceptibility with CH.

Figure 1

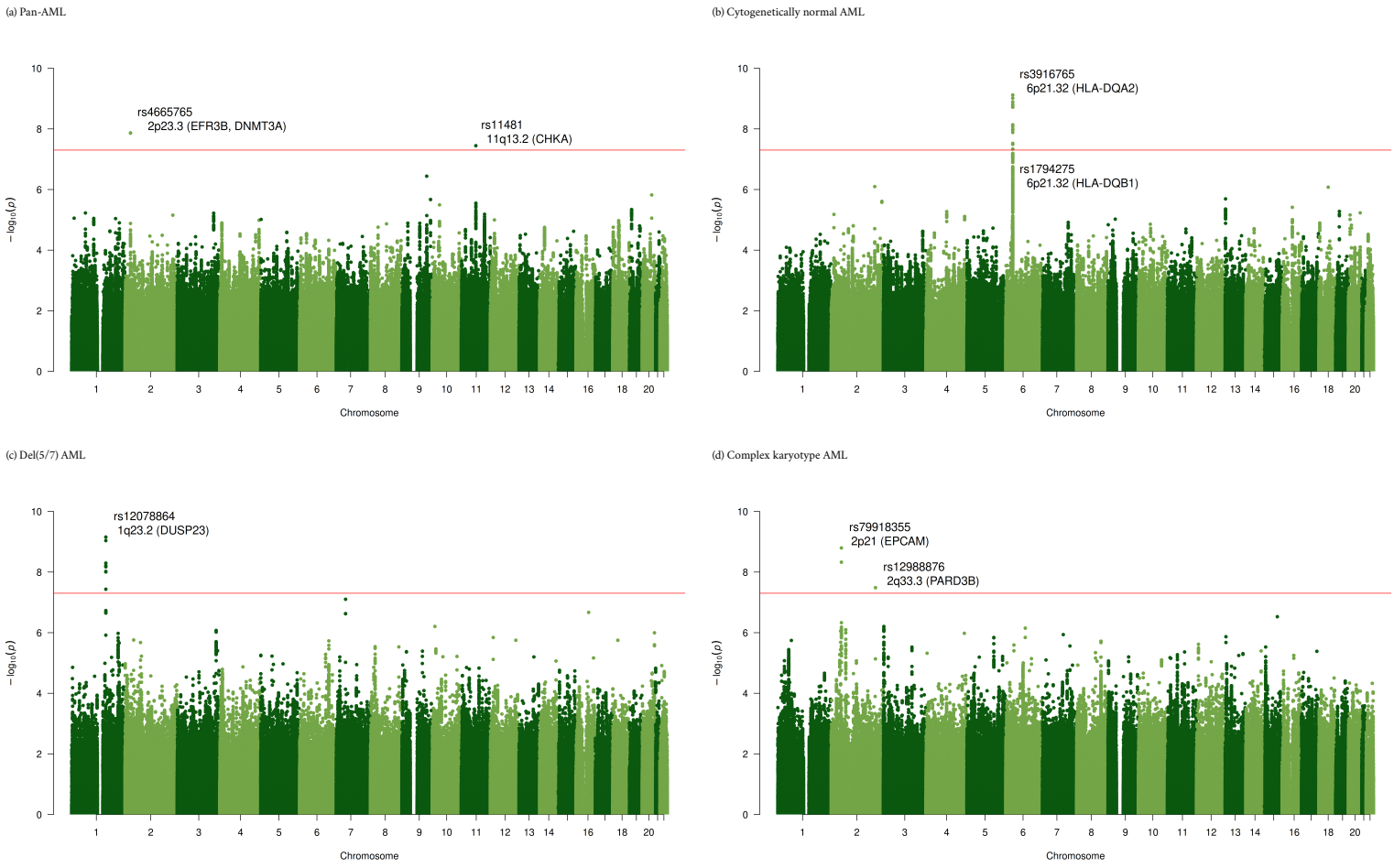
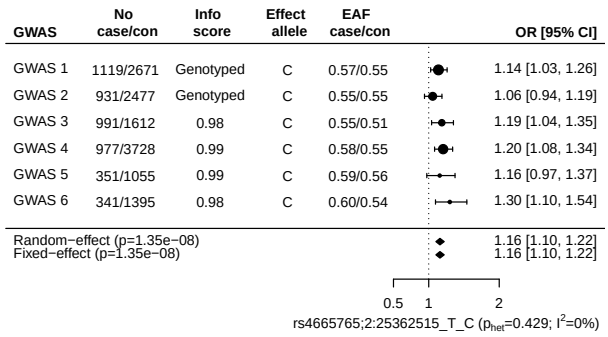
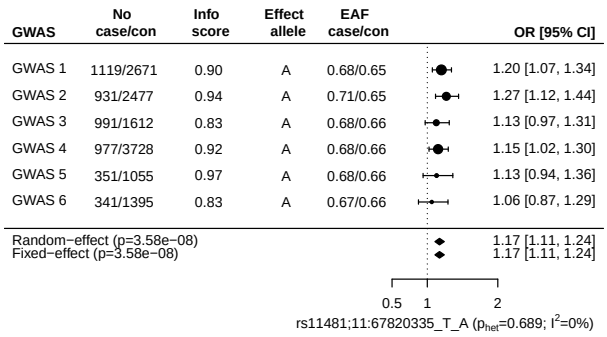


Figure 2

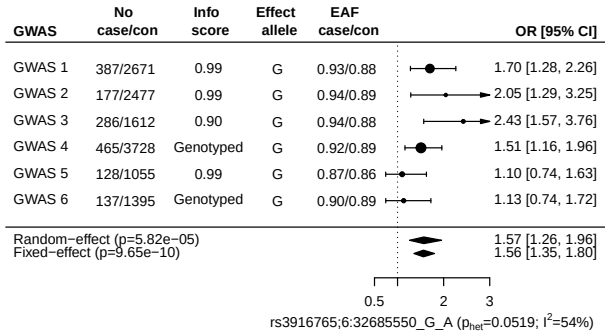
(a) Pan-AML 2p23.3



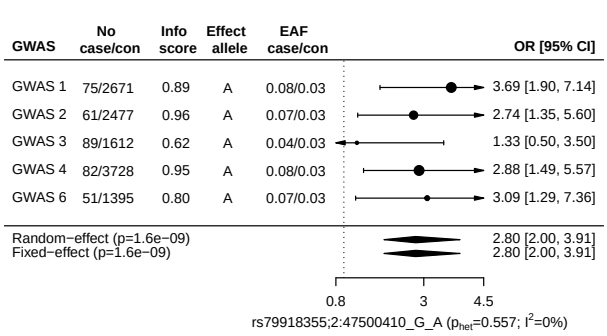
(b) Pan-AML 11q13.2



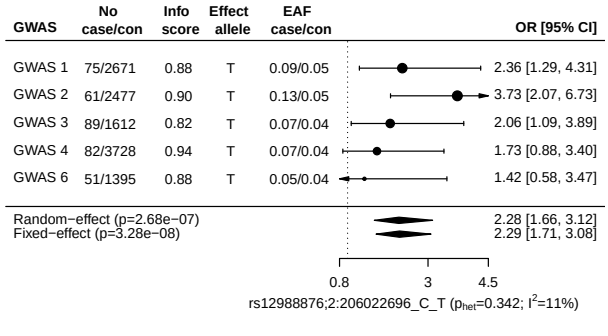
(c) Cytogenetically normal AML 6p21.32



(d) Complex karyotype AML 2p21



(e) Complex karyotype 2q33.3



(f) Del(5/7) AML 1q23.2

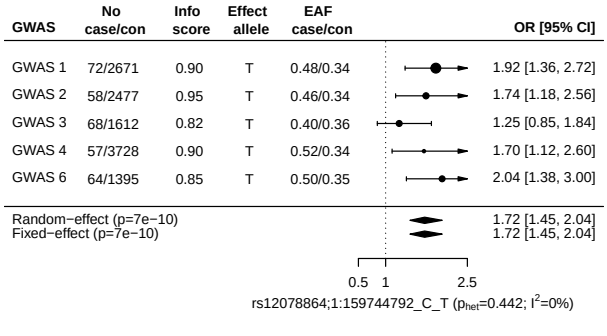
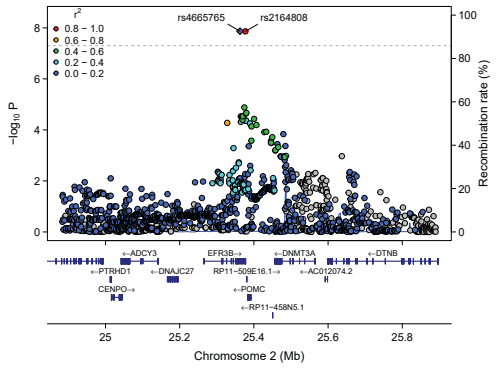
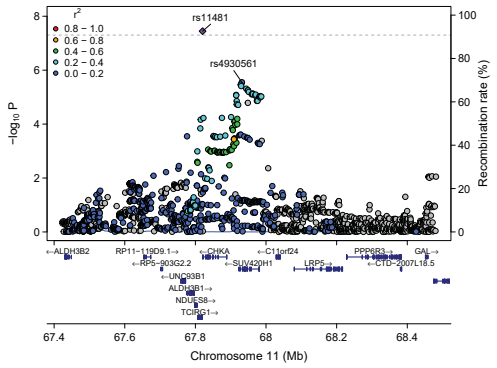


Figure 3

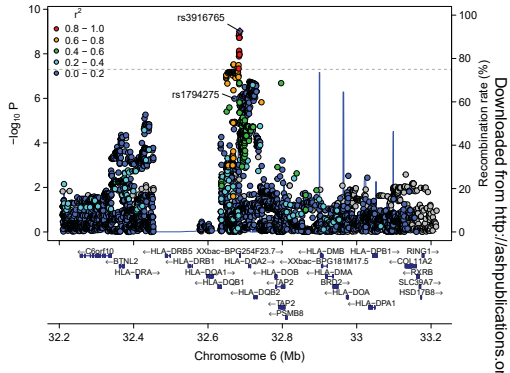
(a) Pan-AML 2p23.3



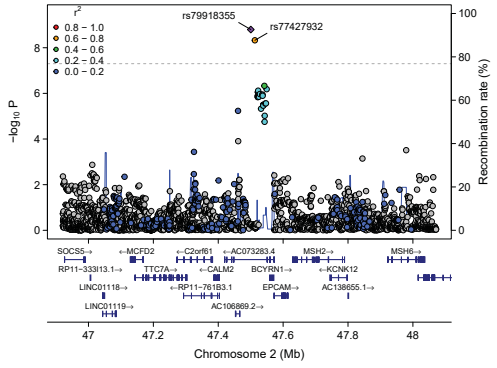
(b) Pan-AML 11q13.2



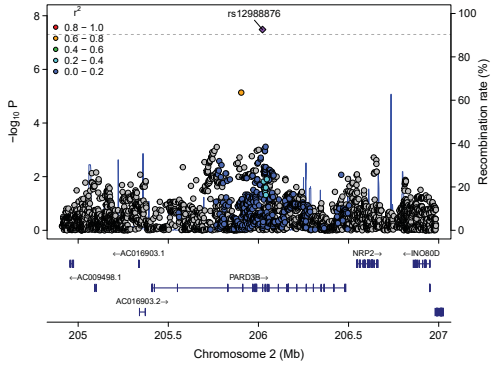
(c) Cytogenetically normal AML 6p21.32



(d) Complex karyotype AML 2p21



(e) Complex karyotype 2q33.3



(f) Del(5/7) AML 1q23.2

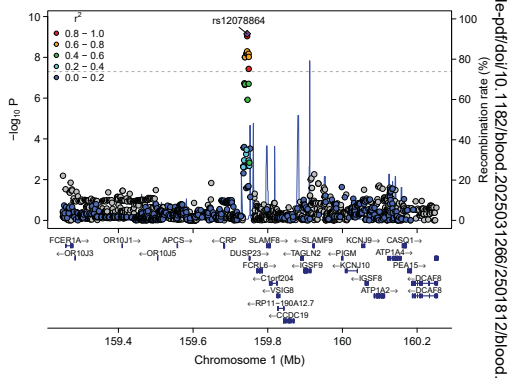


Figure 4

SNP	CHR	Gene	REF/EFF	Case/Con	Trait	Study	OR (95% CI)	P value (unadjusted)	P value (adjusted)
rs2736100	5	TERT	C/A	4710/12938	Pan−AML	AML GWAS	0.90 (0.86 to 0.95)	1.93e−04	1.16e−02
				1583/12938	CN−AML	AML GWAS	0.89 (0.82 to 0.97)	4.85e−03	2.91e−01
				9386/211899	CH	Kar et al 2022	0.82 (0.80 to 0.85)	9.70e−42	
				6829/211899	CH−DNMT3A	Kar et al 2022	0.79 (0.76 to 0.82)	7.60e−32	
				2818/211899	CH−TET2	Kar et al 2022	0.77 (0.72 to 0.81)	2.50e−17	
rs2853677	5	TERT	G/A	4710/12938	Pan−AML	AML GWAS	0.89 (0.85 to 0.94)	3.81e−05	2.29e−03
				1583/12938	CN−AML	AML GWAS	0.90 (0.83 to 0.98)	1.09e−02	6.54e−01
				9386/211899	CH	Kar et al 2022	0.81 (0.78 to 0.83)	2.40e−50	
				6829/211899	CH−DNMT3A	Kar et al 2022	0.77 (0.74 to 0.80)	2.00e−39	
				2818/211899	CH−TET2	Kar et al 2022	0.78 (0.74 to 0.83)	2.40e−14	
rs7705526	5	TERT	C/A	4710/12938	Pan−AML	AML GWAS	1.15 (1.09 to 1.22)	1.70e−06	1.02e−04
				1583/12938	CN−AML	AML GWAS	1.18 (1.08 to 1.29)	1.69e−04	1.02e−02
				9386/211899	CH	Kar et al 2022	1.24 (1.20 to 1.28)	3.40e−44	
				6829/211899	CH−DNMT3A	Kar et al 2022	1.30 (1.25 to 1.36)	2.80e−34	
				2818/211899	CH−TET2	Kar et al 2022	1.29 (1.20 to 1.37)	1.70e−13	
				25657/342869	CHIP	Kessler et al 2022	1.28 (1.26 to 1.31)	1.29e−140	
				16219/342869	CHIP−DNMT3A	Kessler et al 2022	1.30 (1.27 to 1.33)	1.06e−106	
				3918/342869	CHIP−TET2	Kessler et al 2022	1.29 (1.23 to 1.36)	6.28e−27	
rs10890839	11	ATM	C/A	4710/12938	Pan−AML	AML GWAS	1.11 (1.05 to 1.16)	1.14e−04	6.84e−03
				1583/12938	CN−AML	AML GWAS	1.17 (1.08 to 1.26)	1.22e−04	7.30e−03
				25418/339848	CHIP	Kessler et al 2022	1.12 (1.10 to 1.14)	1.43e−32	
				16067/339848	CHIP−DNMT3A	Kessler et al 2022	1.14 (1.11 to 1.16)	1.35e−27	
				3883/339848	CHIP−TET2	Kessler et al 2022	1.16 (1.11 to 1.22)	1.23e−10	
rs11212666	11	ATM	A/T	4710/12938	Pan−AML	AML GWAS	1.10 (1.05 to 1.16)	1.68e−04	1.01e−02
				1583/12938	CN−AML	AML GWAS	1.16 (1.07 to 1.25)	3.62e−04	2.17e−02
				9386/211899	CH	Kar et al 2022	1.09 (1.06 to 1.12)	4.80e−09	
				6829/211899	CH−DNMT3A	Kar et al 2022	1.15 (1.11 to 1.20)	2.20e−12	
				2818/211899	CH−TET2	Kar et al 2022	1.11 (1.04 to 1.18)	1.00e−03	
rs228606	11	ATM	G/T	4710/12938	Pan−AML	AML GWAS	1.11 (1.05 to 1.16)	1.17e−04	7.03e−03
				1583/12938	CN−AML	AML GWAS	1.16 (1.08 to 1.26)	1.59e−04	9.54e−03
				9386/211899	CH	Kar et al 2022	1.09 (1.06 to 1.12)	9.60e−09	
				6829/211899	CH−DNMT3A	Kar et al 2022	1.16 (1.11 to 1.20)	9.30e−13	
				2818/211899	CH−TET2	Kar et al 2022	1.10 (1.03 to 1.17)	3.30e−03	
				25657/342869	CHIP	Kessler et al 2022	1.12 (1.10 to 1.14)	1.87e−32	
				16219/342869	CHIP−DNMT3A	Kessler et al 2022	1.14 (1.11 to 1.16)	4.76e−28	
				3918/342869	CHIP−TET2	Kessler et al 2022	1.16 (1.11 to 1.22)	1.11e−10	
rs188761458	17	MSI2	C/T	4710/12938	Pan−AML	AML GWAS	0.68 (0.56 to 0.83)	1.99e−04	1.19e−02
				1583/12938	CN−AML	AML GWAS	0.71 (0.51 to 0.97)	3.39e−02	1.00e+00
				25657/342869	CHIP	Kessler et al 2022	0.83 (0.80 to 0.87)	3.72e−16	
				16219/342869	CHIP−DNMT3A	Kessler et al 2022	0.83 (0.78 to 0.87)	8.12e−12	
				3918/342869	CHIP−TET2	Kessler et al 2022	0.84 (0.76 to 0.94)	1.99e−03