



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 (Translational 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 (University of Chicago Comprehensive Cancer Center, United States) Tahira 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 (Department 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 (Institute 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-Seguí (Hospital Universitario I Politecnico 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 (CHU 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; $P=1.35 \times 10^{-8}$; EFR3B, POMC, DNMT3A, DNAJC27) which also significantly associates with patient survival ($P=6.09 \times 10^{-3}$). 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; $P=7.0 \times 10^{-10}$; DUSP23) and cytogenetically complex AML at 2q33.3 (rs12988876; $P=3.28 \times 10^{-8}$; PARD3B) and 2p21 (rs79918355; $P=1.60 \times 10^{-9}$; 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, KP_i, SJ, MB, CR, HA, LR, DK, LW, HJC, RD, MKA, MCF, GMart, GMarc, MAS, JC, IG-S, TC, CM, SR, HS, MTV, HD, MCh, CP_r, REG, DL, JWe, AM, DN, W-KH, AG, KP_o, 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):

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

3
 4 Diyanath Ranasinghe^{1*}, Wei-Yu Lin^{1*}, Sarah E Fordham¹, Abrar Alharbi^{1,2}, Nicola J Sunter¹, Claire
 5 Elstob¹, Mohammed H Nahari^{1,3}, Yaobo Xu¹, Catherine Park¹, Eric Hungate⁴, Anne Quante⁵,
 6 Konstantin Strauch^{6,7,8}, Christian Gieger⁷, Andrew Skol⁴, Thahira Rahman¹, Lara Sucheston-
 7 Campbell⁹, Theresa Hahn¹⁰, Alyssa I Clay-Gilmour¹¹, Gail L Jones¹², Helen J Marr¹², Graham H
 8 Jackson¹², Tobias Menne¹², Mathew Collin¹², Adam Ivey¹³, Robert K Hills¹⁴, Alan K Burnett¹⁵, Nigel
 9 H Russell¹⁶, Jude Fitzgibbon¹⁷, Richard A Larson⁴, Michelle M Le Beau⁴, Wendy Stock⁴, Olaf
 10 Heidenreich^{1,18,19}, Amir Enshaei¹, Dumni Gunasinghe¹, Zoë L Hawking¹, Holly Heslop¹, Devi
 11 Nandana¹, Bingjing Di¹, Anna Plokhuta¹, Imogen T. Brown¹, David J Allsup²⁰, Richard S Houlston²¹,
 12 Andrew Collins²², Paul Milne²³, Jean Norden²³, Anne M Dickinson²³, Clare Lendrem²³, Ann K Daly²³,
 13 Louise Palm²⁴, Kim Piechocki²⁴, Sally Jeffries²⁴, Martin Bornhäuser^{25,26,27}, Christoph Röllig²⁷, Heidi
 14 Altmann²⁷, Leo Ruhnke²⁷, Desiree Kunadt²⁷, Lisa Wagenführ²⁷, Heather J Cordell²⁸, Rebecca
 15 Darlay²⁸, Mette K Andersen²⁹, Maria C Fontana^{30,31}, Giovanni Martinelli³⁰, Giovanni Marconi³¹,
 16 Miguel A Sanz³², José Cervera³², Inés Gómez-Seguí³², Thomas Cluzeau³³, Chimène Moreilhon³³,
 17 Sophie Raynaud³³, Heinz Sill³⁴, Maria Teresa Voso³⁵, Hervé Dombret³⁶, Meyling Cheok³⁷, Claude
 18 Preudhomme³⁷, Rosemary E Gale³⁸, David Lynch³⁸, Julia Weisinger³⁹, Andras Masszi⁴⁰, Daniel
 19 Nowak⁴¹, Wolf-Karsten Hofmann⁴¹, Amanda Gilkes⁴², Kimmo Porkka⁴³, Jelena D Milosevic
 20 Feenstra⁴⁴, Robert Kralovics⁴⁴, Junke Wang⁴⁵, Manja Meggendorfer⁴⁶, Torsten Haferlach⁴⁶, Szilvia
 21 Krizsán^{47,48}, Csaba Bödör⁴⁷, Brian Parkin⁴⁹, Sami N. Malek⁴⁹, Friedrich Stölzel⁵⁰, Kenan Onel⁵¹, James
 22 M Allan¹

23
 24 ¹Translational and Clinical Research Institute, Newcastle University Centre for Cancer, Faculty of
 25 Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.

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

28 ³Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, Najran
 29 University, Najran, Saudi Arabia

30 ⁴University of Chicago Comprehensive Cancer Center, Chicago, Illinois. USA.

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

33 ⁶Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg,
 34 Germany.

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

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

- 39 ⁹Karmanos Cancer Institute and Department of Oncology, Wayne State University School of
40 Medicine, Detroit, MI, USA.
- 41 ¹⁰Department of Cancer Prevention, Roswell Park Cancer Institute, Buffalo, New York, USA.
- 42 ¹¹Arnold School of Public Health, Department of Epidemiology & Biostatistics, University of South
43 Carolina, USA.
- 44 ¹²Department of Haematology, Freeman Hospital, Newcastle upon Tyne Hospitals National Health
45 Service Foundation Trust, Newcastle upon Tyne, UK.
- 46 ¹³Department of Pathology, Alfred Hospital, Melbourne, Australia.
- 47 ¹⁴Nuffield Department of Population Health, University of Oxford, Oxford, UK.
- 48 ¹⁵Paul O’Gorman Leukaemia Research Centre, University of Glasgow, UK.
- 49 ¹⁶Department of Haematology, Guy’s and St Thomas’ Hospital, London, UK.
- 50 ¹⁷Barts Cancer Institute, Queen Mary University of London, London, UK.
- 51 ¹⁸Prinses Maxima Center of Pediatric Oncology, Utrecht, The Netherlands.
- 52 ¹⁹Department of Hematology, University Medical Center Utrecht, Utrecht, The Netherlands.
- 53 ²⁰Centre for Biomedicine, Hull York Medical School, University of Hull, UK
- 54 ²¹Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.
- 55 ²²School of Human Development and Health, Faculty of Medicine, University of Southampton,
56 Southampton SO16 6YD, UK
- 57 ²³Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University,
58 Newcastle upon Tyne, UK.
- 59 ²⁴West Midlands Regional Genetics Laboratory, Birmingham Women’s Hospital, Birmingham, UK.
- 60 ²⁵School of Cancer and Pharmaceutical Research, King’s College, Department of Haematology,
61 King’s College London, London, UK
- 62 ²⁶National Center for Tumor Diseases NCT, Partner site Dresden, Dresden, Germany.
- 63 ²⁷Medizinische Klinik und Poliklinik I, University Hospital Carl Gustav Carus Dresden, Technical
64 University of Dresden, Dresden, Germany.
- 65 ²⁸Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK.
- 66 ²⁹Department of Clinical Genetics, University Hospital Rigshospitalet, Copenhagen, Denmark.
- 67 ³⁰Institute of Hematology "L. and A. Seràgnoli", University of Bologna, Bologna, Italy.
- 68 ³¹University of Bologna, Hematology Unit, Ospedale S. Maria delle Croci, Ravenna (RA), Italy.
- 69 ³²Hematology Service, Hospital Universitario y Politécnico La Fe, Valencia, Spain and CIBERONC,
70 Instituto de Salud Carlos III, Madrid, Spain.
- 71 ³³Université Côte d’Azur, CHU Pasteur, Service d’Hématologie Biologique, Nice, France.
- 72 ³⁴Division of Hematology, Medical University of Graz, Austria.

- 73 ³⁵Università di Roma Tor Vergata, Dipartimento di Biomedicina e Prevenzione, Rome, Italy.
- 74 ³⁶Hôpital Saint-Louis, Institut Universitaire d'Hématologie, Université Paris Diderot, Paris, France.
- 75 ³⁷University of Lille, Inserm, CHU Lille, UMR-S 1172 - JPArc - Centre de Recherche Jean-Pierre
76 Aubert Neurosciences et Cancer, F-59000 Lille, France.
- 77 ³⁸Department of Haematology, University College London Cancer Institute, London, UK.
- 78 ³⁹1st Department of Internal Medicine, Semmelweis University, Budapest, Hungary.
- 79 ⁴⁰Hematology and Lymphoma Unit, National Institute of Oncology, Budapest, Hungary.
- 80 ⁴¹Department of Hematology and Oncology, Medical Faculty Mannheim, Heidelberg University,
81 Mannheim, Germany.
- 82 ⁴²Department of Haematology, University of Cardiff, UK.
- 83 ⁴³Helsinki University Hospital Comprehensive Cancer Center, Hematology Research Unit Helsinki,
84 University of Helsinki, Helsinki, Finland.
- 85 ⁴⁴Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria.
- 86 ⁴⁵Department of Systems Biology, UT MD Anderson Cancer Center, Houston TX, USA.
- 87 ⁴⁶MLL Munich Leukemia Laboratory, Munich, Germany.
- 88 ⁴⁷MTA-SE Lendulet Molecular Oncohematology Research Group, Department of Pathology and
89 Experimental Cancer Research, Semmelweis University, Budapest, Hungary.
- 90 ⁴⁸Pediatric Center, Semmelweis University, Budapest, Hungary.
- 91 ⁴⁹Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan,
92 Ann Arbor, Michigan, USA.
- 93
- 94 ⁵⁰Division of Stem Cell Transplantation and Cellular Immunotherapies, University Hospital
95 Schleswig-Holstein, Campus Kiel, Kiel, Germany.
- 96 ⁵¹Departments of Cancer Prevention and Control, Medicine, and Clinical Genomics, Roswell Park
97 Comprehensive Cancer Center, Buffalo, New York, USA.

98

99 *These authors contributed equally.

100

101 Correspondence:

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

103

104 **Data availability**

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

109 previously⁶. Genotyping data and/or samples for GWAS5 cases and controls are available by
110 application to the Finnish Hematology Registry and Clinical Biobank (<https://www.fhrb.fi/>) and The
111 National Institute for Health and Welfare (THL) Biobank of Finland ([https://thl.fi/en/research-and-](https://thl.fi/en/research-and-development/thl-biobank)
112 [development/thl-biobank](https://thl.fi/en/research-and-development/thl-biobank)). Genotyping data for GWAS6 cases are available via the NCBI Gene
113 Expression Omnibus under accession numbers GSE21107⁸
114 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM527831>, GSE61323¹⁰
115 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61323>, GSE23452⁹
116 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE23452>.

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

120

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

126

127

128 **Key points**

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

133 **Abstract**

134 Acute myeloid leukemia (AML) is a complex hematological malignancy with multiple disease sub-
135 groups defined by somatic mutations and heterogeneous outcomes. Although genome-wide
136 association studies (GWAS) have identified a small number of common genetic variants influencing
137 AML risk, the heritable component of this disease outside of familial susceptibility remains largely
138 undefined. Here we perform a meta-analysis of four published GWAS plus two new GWAS, totalling
139 4710 AML cases and 12938 controls. We identify a new genome-wide significant risk locus for pan-
140 AML at 2p23.3 (rs4665765; $P=1.35 \times 10^{-8}$; *EFR3B*, *POMC*, *DNMT3A*, *DNAJC27*) which also
141 significantly associates with patient survival ($P=6.09 \times 10^{-3}$). Our analysis also identifies three new
142 genome-wide significant risk loci for disease sub-groups, including AML with deletions of
143 chromosome 5 and/or 7 at 1q23.3 (rs12078864; $P=7.0 \times 10^{-10}$; *DUSP23*) and cytogenetically complex
144 AML at 2q33.3 (rs12988876; $P=3.28 \times 10^{-8}$; *PARD3B*) and 2p21 (rs79918355; $P=1.60 \times 10^{-9}$; *EPCAM*).
145 We also investigated loci previously associated with risk of clonal hematopoiesis (CH) or clonal
146 hematopoiesis of indeterminate potential (CHIP) and identified several variants associated with risk of
147 AML. Our results further inform on AML etiology and demonstrate the existence of disease sub-
148 group specific risk loci.

149 **Introduction**

150 Acute myeloid leukemia (AML) is the most common acute leukemia in Europeans and comprises
151 multiple sub-groups defined by somatic genetic/epigenetic alterations and heterogenous clinical
152 outcomes¹. The existence of rare constitutional genetic variants predisposing to AML with high
153 penetrance demonstrates a role for genetics in disease susceptibility^{2,3}. However, strong familial
154 susceptibility to AML is rare, and the prevailing evidence suggests that for the majority of individuals
155 the genetic risk for AML is determined by co-inheritance of multiple independent low penetrance
156 genetic variants⁴⁻⁶.

157
158 To identify novel AML risk loci we conducted a meta-analysis of four published genome-wide
159 association studies (GWAS)⁶ and two new GWAS, incorporating 4710 AML cases and 12938
160 controls of European ancestry, and report the identification of new pan-AML and AML sub-group
161 specific risk loci. This is the largest AML GWAS to date and provides further evidence for the
162 existence of common low-penetrance susceptibility alleles, as well as evidence for heterogeneity in
163 genetic risk across disease sub-groups.

164

165 **Methods**

166 **Study Participants**

167 GWAS1, GWAS2, GWAS3 and GWAS4 comprised 1119, 931, 991 and 977 AML cases,
168 respectively, and 2671, 2477, 1612 and 3728 controls, respectively, as described previously⁶. GWAS5
169 comprised 351 cases from the Finnish Hematology Registry and Clinical Biobank genotyped on the
170 Illumina Omni Express Exome BeadChip. For controls, we used publicly available Illumina Omni
171 Express Exome BeadChip data on 1055 individuals from The National Institute for Health and
172 Welfare (THL) Biobank of Finland (Health 2000 and Health 2011 studies). GWAS6 comprised
173 Affymetrix SNP6.0 array data on 341 AML cases of European ancestry recruited to The Cancer
174 Genome Atlas (TCGA) project (phs000178/GRU)⁷ or via hematology clinics at the University of
175 Michigan Comprehensive Cancer Center⁸⁻¹⁰. For controls, we used Affymetrix SNP6.0 array data on
176 1395 healthy individuals of European ancestry from the GAIN: Genome-Wide Association Study of
177 Schizophrenia project (phs000021/GRU).

178

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

185 Board (IRBMED #2004-1022). Access to the TCGA AML cases and GAIN controls was approved by
186 the National Institute of Health (#9683). Information on ethical approvals for all other studies has
187 been reported previously⁶

188

189 **Genotyping and genome-wide quality-control procedures**

190 Genotype calling was performed using Illumina GenomeStudio software or Affymetrix Genotyping
191 Console software v4.2.0.26. Data handling and analysis was performed using R v4.2.1, PLINK
192 v1.9b4.4 and SNPTEST v2.5.6. Rigorous variant and sample quality control metrics were applied to
193 all six GWAS (Supplementary Figure 1). Specifically, we excluded variants with a call rate less than
194 95%, with departure from Hardy-Weinberg equilibrium (HWE; $P < 10^{-3}$) or with significant differences
195 ($P < 0.05$) in missingness between cases and controls. Individual samples with a call rate of $< 95\%$ or
196 with extreme heterozygosity rates (± 3 standard deviation) were also excluded. Individuals were
197 removed with estimated relatedness $\text{pihat} > 0.1875$, both within and across GWAS. Ancestry was
198 assessed using principal component analysis and super-populations from the 1000 genomes project as
199 a reference, with individuals of non-European ancestry excluded based on the first two principal
200 components (Supplementary Figures 2, 3 and 4)⁶.

201

202 The majority of AML cases were genotyped using DNA extracted from cell/tissue samples (blood and
203 bone marrow) taken during AML remission. For GWAS5, AML cases were primarily genotyped
204 using DNA extracted from cell/tissue samples (blood and bone marrow) taken during disease
205 presentation. We employed a stringent HWE cut-off in order to eliminate SNPs potentially affected by
206 somatic copy number alterations. Furthermore, we also used Nexus Copy Number v10 (Bionano
207 Genomics, California) data from 351 AML cases genotyped using samples with high somatic cell
208 content to interrogate Log R ratio and B allele frequency at loci carrying risk variants for AML
209 susceptibility. We found no significant evidence of somatic alterations affecting the risk loci at
210 11q13.2 (rs11421) or 6p21.32 (rs3916765)⁶. For the risk locus at 2p23.3 (rs4665765; *EFR3B*, *POMC*,
211 *DNMT3A*, *DNAJC27*) we identified 4 cases with deletions and one case with gain. For the risk locus
212 at 1q23.2 (rs12078864; *DUSP23*) we identified 3 cases with trisomy 1 and one case with a 1q gain.
213 For the risk locus at 2p21 (rs79918355; *EPCAM*) we identified one case with a gain and one case with
214 a deletion. For the risk variant at 2q33.3 (rs12988876; *PARD3*) we identified 1 case with a deletion.
215 These data suggest that significant genotyping errors due to somatically acquired allelic imbalance in
216 AML cases are unlikely.

217

218 **Imputation, genome-wide association testing and meta-analysis**

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

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

224

225 For each GWAS, association tests were performed for all cases (pan-AML), cytogenetically normal
226 AML, complex karyotype AML and AML with deletions of chromosome 5 and/or 7, assuming an
227 additive genetic model with nominally significant ($P<0.05$) principal components included in the
228 analysis as covariates. Association summary statistics were combined for all six GWAS, in fixed
229 effects models. Individual GWAS with less than 50 cases in any sub-group were excluded from
230 subtype specific analyses. Cochran's Q statistic was used to test for heterogeneity and the I^2 statistic
231 was used to quantify variation due to heterogeneity. Conditional analysis was conducted using the
232 GCTA conditional and joint analysis (COJO) pipeline v1.94.4. A stepwise model selection with a
233 GWAS P value cut-off of 5×10^{-5} and collinearity cutoff of 0.9 was used to select independent
234 associations from the summary statistics.

235

236 **Fine mapping and functional annotation of causal SNPs**

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

245

246 **Relationship between AML susceptibility variant genotype and patient survival**

247 The relationship between AML risk variants and overall survival was evaluated in a total of 1725
248 AML patients (excluding acute promyelocytic leukemia) from the UK^{12,13}, Germany^{14,15}, Hungary,
249 Finland (<https://www.fhrb.fi/>) and the United States⁷⁻¹⁰. Patients were treated with conventional
250 intensive AML therapy including ara-C, daunorubicin and best supportive care. A subset of high-risk
251 patients in the German cohort were treated with stem cell transplantation¹⁵. Overall survival was
252 defined as the time from diagnosis to the date of last follow-up or death from any cause. Cox
253 regression analysis was used to estimate allele specific hazard ratios and 95% confidence intervals. In
254 order to control for index (collider) bias potentially introduced during the selection of cases for the
255 survival analysis, we applied the corrected weighted least squares (CWLS) method¹⁶, which uses the
256 slope of a weighted regression of prognostic variants on risk variant associations as the bias correction
257 factor to re-estimate the effects of variants on disease progression. We used 52408 LD pruned

258 ($r^2 \leq 0.1$, 250 kb SNP window) post-imputation variants with ≥ 0.98 imputation score as independent
259 instrument variables from the GWAS summary statistics.

260

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

262 A total of 59 statistically significant ($P < 5 \times 10^{-8}$) variants were identified from published GWAS
263 analyses of clonal haematopoiesis (CH), clonal haematopoiesis of indeterminate potential (CHIP) and
264 their subtypes stratified by the two primary CH genes, *DNMT3A* and *TET2*. From Kar et al. (2022)¹⁷,
265 15 independent lead variants that replicated in the UK Biobank cohort and that retained significance
266 after conditioning on the lead variants were identified from analyses of pan-CH, *DNMT3A* mutation-
267 positive CH and *TET2* mutation-positive CH. From Kessler et al (2022)¹⁸, 52 lead variants were
268 identified by LD thresholding and then by replication in the Geisinger MyCode Community Health
269 Initiative (GHS) cohort from analyses of pan-CHIP, *DNMT3A* mutation-positive CHIP and *TET2*
270 mutation-positive CHIP. We also included the independent lead variant at the *TCL1A* locus
271 (rs2887399) from the CHIP GWAS reported by Weinstock et al (2023)¹⁹. Pan-AML and
272 cytogenetically normal AML association *P* values were adjusted for multiple testing using Bonferroni
273 correction.

274

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

277

278

279 **Results**

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

281 We conducted 6 independent genome-wide association studies with AML cases and controls of
282 European ancestry (GWAS1-6), four of which (GWAS1-4) have been reported previously⁶. A total of
283 4710 AML cases and 12938 healthy controls passed the study level quality control (Supplementary
284 Figures 1-4) with common autosomal single nucleotide variants numbering between 250880 to
285 436068 genotyped across the 6 GWAS. We further improved the genomic resolution by imputing >7
286 million variants using the Haplotype Reference Consortium panel²⁰. After excluding variants with an
287 INFO (imputation quality) score of < 0.6 and a minor allele frequency (MAF) < 0.02 , association tests
288 were conducted for 5646403 autosomal variants common to all 6 GWAS. Considering the genetic and
289 biological heterogeneity of AML, we calculated odds ratios (OR) for all AML cases (pan-
290 AML)(N=4710) and three major AML subtypes; cytogenetically normal AML (N=1580), complex
291 karyotype AML (N=358) and AML with deletions affecting chromosome 5 and/or chromosome 7
292 (del(5/7) AML)(N=319)²¹. Nominally significant ($P < 0.05$) principal components in each GWAS were
293 used as covariates to control inflation of the test statistic ($0.99 > \lambda_{GC} < 1.07$) for each analysis
294 (Supplementary Figures 5–8).

295

296 Meta-analysis of six GWAS identified the previously reported signal for pan-AML at 11q13.2
297 (*KMT5B*, *CHKA*, *ALDH3B1*, *NDUFS8*, *TCIRG1*)⁶, although with a new lead variant at this locus
298 (rs11481 (MAF 0.34, INFO scores 0.83-0.96), $P = 3.58 \times 10^{-8}$) located 100 kb centromeric to the
299 previous lead variant (rs4930561) ($r^2=0.26$) (Figures 1 and 2)⁶. Meta-analysis also identified a new
300 signal for pan-AML surpassing genome-wide significance at 2p23.3 (rs4665765 (MAF 0.46, INFO
301 scores 0.98-0.99); $P = 1.35 \times 10^{-8}$; *EFR3B*, *POMC*, *DNMT3A*, *DNAJC27*) (Figures 1 and 2). Meta-
302 analysis of cytogenetically normal AML across six GWAS studies identified the previously reported
303 signal at 6p21.32 (rs3916765 (MAF 0.12, INFO scores 0.90-0.99), *HLA-DQA2*)⁶ (Figures 1 and 2),
304 which has also been validated in an independent study²². However, our analysis did not reveal any
305 new signals for cytogenetically normal AML surpassing the threshold for genome-wide significance.
306 Meta-analysis of GWAS with sufficient cases in each sub-group identified genome-wide significant
307 association signals for del(5/7) AML at 1q23.2 (rs12078864 (MAF 0.31, INFO score 0.83-0.96), $P =$
308 7.0×10^{-10} , *DUSP23*) and for complex karyotype AML at 2p21 (rs79918355 (MAF 0.028, INFO
309 score 0.62-0.96), $P = 1.6 \times 10^{-9}$, *EPCAM*) and 2q33.3 (rs12988876 (MAF 0.042, INFO score 0.82-
310 0.94), $P = 3.28 \times 10^{-8}$, *PARD3B*) (Figures 1 and 2).

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

316 **Statistical fine-mapping of AML association signals**

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

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

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

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

340 Two variants were included in the 95% credible set for the complex karyotype signal at 2p21. The
341 lead variant at this locus (rs79918355, OR 2.80, 95% CI 2.0-3.91; $P=1.6 \times 10^{-9}$) was implicated as the
342 most credible causal variant (PIP=0.91) (Supplementary Figure 16a, Supplementary Table 7), which
343 localises to an intronic region of the *AC073283.4* long non-coding RNA and is a significant *cis*-eQTL
344 for *MCFD2* ($P=1.04 \times 10^{-5}$) and *MSH2* ($P=7.23 \times 10^{-3}$) (Supplementary Table 4).

345 The complex karyotype AML signal at 2q33 fine-mapped to one credible set only including the lead
346 variant at this locus (rs12988876, OR 2.29, 95% CI 1.71-3.08; $P=3.28 \times 10^{-9}$, PIP=0.99)
347 (Supplementary Figure 16B), which maps intronic to and is a significant eQTL for *PARD3B* ($P=0.03$)
348 (Supplementary Table 5).

349 One 95% credible set was deduced from the del(5/7) AML association signal on chromosome 1q23.2,
350 capturing the top 8 SNPs at this locus. The lead variant at this locus (rs12078864, OR 1.72, 95% CI
351 1.45 – 2.04; $P=7 \times 10^{-10}$) was identified as the most credible causal variant (PIP=0.46) (Supplementary
352 Figure 17 and Table 7) and is a significant *cis*-eQTL for *DUSP23* ($P=7.99 \times 10^{-84}$), *SLAMF8*
353 ($P=3.56 \times 10^{-11}$), *PEX19* ($P=1.82 \times 10^{-3}$), *DARC* ($P=6.01 \times 10^{-3}$), *FCRL6* ($P=1.49 \times 10^{-2}$), *Clorf204*
354 ($P=2.23 \times 10^{-2}$), *FCERIA* ($P=2.23 \times 10^{-2}$), *RP11-404F10.2* ($P=4.78 \times 10^{-2}$) and *CD84* ($P=4.91 \times 10^{-2}$)
355 (Supplementary table 6).

356 **Reported risk variants for clonal hematopoiesis (CH), Clonal haematopoiesis of indeterminate** 357 **potential (CHIP) and their association with AML.**

358 Clonal haematopoiesis (CH) is an age-related non-malignant condition defined by the expansion of
359 haematopoietic stem cells (HSC) and progenitor cells in healthy individuals following the acquisition
360 of somatic driver mutations. Clonal haematopoiesis of indeterminate potential (CHIP) is CH driven by
361 a somatic mutation in a gene recurrently mutated in myeloid malignancy (variant allele frequency
362 ≥ 0.02), with *DNMT3A*, *TET2* and *ASXL1* being the most commonly affected²⁴. The presence of CH

363 identifies individuals with an increased risk of developing AML²⁵ and recent studies have reported
364 constitutional genetic variants associated with risk of developing CH^{17–19,26,27}.

365 A total of 59 CH/CHIP risk variants reported by Kar et al. 2022¹⁷, Kessler et al. 2022¹⁸ and Weinstock
366 et al. 2023¹⁹ were annotated in our AML GWAS. Of these, 7 CH/CHIP variants were significantly
367 associated with risk of either pan-AML or CN-AML after correction for multiple testing, including
368 variants at the *TERT* locus (rs2736100, rs2853677, rs7705526), the *ATM* locus (rs10890839,
369 rs11212666, rs228606) and the *MSI2* locus (rs188761458) (Figure 4). The risk of AML for all 7
370 variants was in the same direction as the reported risk of CH/CHIP. A further 18 reported CH/CHIP
371 variants were nominally significantly associated with risk of either pan-AML or CN-AML ($P < 0.05$),
372 but these did not retain significance after correction for multiple testing (Supplementary Table 8). For
373 14 of these 18 variants, the risk of AML was in the same direction as the reported risk of CH/CHIP
374 (Supplementary Table 8). The remaining 4 variants were all at the *TCL1A* locus and the reported risk
375 of CH/CHIP was in opposing directions for *TET2*-mutated and *DNMT3A*-mutated CH/CHIP^{17,18}.
376 However, the risk of AML was in the same direction as the risk of *TET2*-mutant CH/CHIP for all 4
377 variants (Supplementary Table 8). In summary, 7 risk variants for CH/CHIP were significantly
378 associated with AML, providing further evidence for shared genetic susceptibility between these
379 conditions.

380
381 All 6 AML susceptibility variants reported in our study were annotated in the CH/CHIP datasets
382 published by Kar et al (2022)¹⁷ and/or Kessler et al (2022)¹⁸, but none were significantly associated
383 with risk of CH/CHIP after multiple testing correction (Supplementary Table 9). However, one
384 variant at the 11q13.2 risk locus for pan-AML (rs11481, *CHKA*) was nominally significantly
385 associated with risk of CH (OR 1.04, 95% CI 1.00-1.07, $P = 0.024$)¹⁷, but this did not retain
386 significance after correction for multiple testing (Supplementary Table 9).

387

388 **Risk variants for AML and their impact on patient survival**

389 The relationship between the identified AML risk variants and survival was evaluated in 1725 AML
390 patients from the UK, Germany, Hungary, Finland and the United States of America. Five of the six
391 AML susceptibility variants identified via GWAS (rs11481 (11q13.2, *CHKA*), rs3916765 (6p21.32,
392 *HLA-DQA2*), rs12078864 (1q23.2, *DUSP23*), rs79918355 (2p21, *EPCAM*), rs12988876 (2q33.3,
393 *PARD3*)) were not significantly associated with overall survival in univariate analysis (Supplementary
394 Figure 18) prior to or after index bias correction (Supplementary Figure 18). However, the pan-AML
395 risk variant at 2p23.3 (rs4665765, *EFR3B*, *POMC*, *DNMT3A*, *DNAJC27*) was nominally significantly
396 associated with overall survival (HR 1.13, 95% CI 1.04-1.24, $P = 6.09 \times 10^{-3}$), with the risk allele for
397 AML being associated with inferior outcome. This variant retained significance with an increased

398 effect size after index bias correction (HR 1.25, 95% CI 1.14-1.38, $P=3.89 \times 10^{-6}$) and after adjustment
399 for multiple testing ($P=2.33 \times 10^{-5}$) (Supplementary Figure 18).

400 Discussion

401 By conducting a meta-analysis of six independent genome-wide association studies we report two
402 pan-AML susceptibility loci and four AML sub-group-specific loci, two of which are reported
403 previously⁶. We used statistical and functional fine-mapping methods to identify the most credible
404 causal variant at each locus and were able to prioritize high confidence genes which could serve as
405 strong candidates for functional validation experiments.

406 We explored three genes localised to the pan-AML susceptibility signal at 2p23.3. The most credible
407 causal variant (rs2164808) at this locus introduces a premature stop codon in *EFR3B*. The *EFR3A/B*
408 family are paralogous proteins that contribute to AT1 signalling regulating G-protein-coupled
409 receptors²⁸. *EFR3B* and *EFR3A* also form a complex to recruit phosphatidylinositol 4-kinase (PI4K)
410 to the plasma membrane, with high expression of PI4K associated with inferior survival in myeloid
411 leukemia²⁹. The lead variant and most credible causal variant at this locus are both *cis*-eQTL for
412 *DNAJC27* (*RBJ*), which encodes small GTPase that promotes development of numerous human
413 cancers via MEK/ERK signalling³⁰⁻³². Of note, the most credible causal variant is located within 200
414 Kb of *DNMT3*, encoding a DNA methyltransferase that is frequently somatically dysregulated in
415 AML¹, where loss of function disrupts global genomic methylation in hematopoietic progenitor cells
416 leading to leukemogenesis^{23,33}. Intriguingly, the AML risk variant is also significantly associated with
417 inferior overall survival.

418 The new lead variant at the 11q13.2 pan-AML association signal (rs11481) localises to the *CHKA*
419 gene and is in modest genetic linkage with the variant (rs4930561) (linkage disequilibrium $r^2=0.26$)
420 previously reported to be associated with pan-AML localised to the *KMT5B* gene⁶. Analysis
421 conditioning on rs11481 did not identify any evidence of additional associations ($P<10^{-4}$) at this locus,
422 suggesting that both variants are part of the same association signal. rs11481 maps to a noncoding
423 exonic region of *CHKA*, and specifically to a noncoding natural antisense transcript (ncNAT) *RP11-*
424 *802E16.3* (ENSG00000255031). Although ncNATs are regulatory RNA molecules that modulate
425 cellular processes such as growth and differentiation, their role in cancer pathogenesis remains
426 unknown. However, there is evidence that *RP11-802E16.3* regulates expression of *CHKA*³⁴.
427 Consistent with this model, rs11481 is a significant *cis*-eQTL for *CHKA* ($P=7.14\times 10^{-11}$), where high
428 expression drives tumour progression and metastasis of several human cancer³⁵⁻³⁷. Moreover,
429 expression of *CHKA* is implicated in the pathogenesis of B-cell malignancies and T-ALL via
430 promotion of cell survival and proliferation³⁸. rs11481 is also a significant *cis*-eQTL for *ALDH3B1*
431 ($P=4.86\times 10^{-44}$) and *NDUFS8* ($P=4.05\times 10^{-10}$). *ALDH3B1* encodes a member of the aldehyde
432 dehydrogenase superfamily that protects cells from oxidative stress by catalysing the reversible
433 oxidation of endogenous and exogenous aldehydes³⁹. High expression of aldehyde dehydrogenase
434 family members is associated with chemotherapy resistance and inferior survival in AML⁴⁰.

435 Furthermore, AML stem cells are acutely sensitive to small molecule inhibitors of aldehyde
436 dehydrogenases, identifying this family of enzymes as a therapeutic vulnerability in AML⁴¹. *NDUFS8*
437 encodes a subunit of the mitochondrial NADH:ubiquinone oxidoreductase complex I, responsible for
438 NADH oxidation as part of the respiratory chain⁴². Rare constitutional variants in *NDUFS8* are
439 associated with attenuated mitochondrial respiration in AML cells and are mutually exclusive with
440 somatically acquired mutations in isocitrate dehydrogenase 1 (*IDHI*)⁴³, suggesting two alternative
441 genetic mechanisms via which mitochondrial function is dysregulated in AML pathogenesis.
442 Although rs11481 is not a *cis*-eQTL for *KMT5B* (*SUV420H1*), a role for this lysine methyltransferase
443 cannot be excluded. *KMT5B* is frequently silenced via hypermethylation in numerous human
444 cancers⁴⁴⁻⁴⁶ and somatic mutation is reported in transformation of myelodysplastic syndrome to
445 AML⁴⁷.

446 We report two signals significantly associated with risk of complex karyotype AML, at 2p21 and
447 2q33. The most likely causal variant at 2p21 (rs79918355) is a *cis*-eQTL for *MCFD2*, which encodes
448 a protein that, along with *LMAN1*, forms a cargo receptor complex for transport of coagulation
449 factors⁴⁸. Rare constitutional variants in *MCFD2* cause combined deficiency of coagulation factor V
450 and VIII, a recessive bleeding disorder⁴⁹. *MCFD2* also has a role in the regulation of stem cell
451 survival and pluripotency^{50,51} and somatic mutations have been reported in leukemic cells from
452 Fanconi anemia patients who developed AML⁵², a disease which often presents with a complex
453 karyotype due to inherent chromosome instability⁵³. The lead variant at the 2q33 signal for complex
454 karyotype AML (rs12988876) maps intronic to *PARD3B* and is an eQTL for this gene ($P=0.03$).
455 *PARD3B* encodes a member of the *PARD3* family of proteins that regulate cell polarity and
456 centrosome localization⁵⁴. *PARD3B* is a homologue of *PARD3*, which also functions as a scaffolding
457 protein that interacts with numerous intracellular signalling molecules, many of which are
458 dysregulated in cancer, including members of the PI3K/AKT and MAPK pathways such as PTEN and
459 JNK⁵⁴. *PARD3B* is implicated in prostate cancer aetiology and expression levels have been associated
460 with survival in both colorectal and breast cancer^{55,56}, suggesting a potential role in numerous human
461 cancers.

462 The lead variant at the 1q23.2 signal for del(5/7) AML is a significant *cis*-eQTL for *DUSP23* and
463 *SLAMF8*. *DUSP23* encodes a dual specific phosphatase that regulates MAP kinase signalling,
464 impacting cell proliferation, growth and survival⁵⁷. *DUSP23* also plays a role in regulating cell
465 adhesion/migration⁵⁸ and high expression in blast cells is as an independent prognostic marker for
466 inferior survival in AML⁵⁹. High *DUSP23* expression is also reported in CD4⁺ T-cells from patients
467 with systemic lupus erythematosus, where it is thought to regulate DNA methyltransferase activity,
468 including *DNMT3A*, which is frequently dysregulated in AML via somatic mutation^{1,23}. *SLAMF8*
469 encodes a cell surface glycoprotein that is a member of the signalling lymphocytic activation
470 molecule (SLAM) family involved in regulating the development and function of a wide range of

471 immune cells, such as T lymphocytes, B cells, neutrophils, dendritic cells, macrophages and
472 eosinophils^{60,61}. *SLAMF8* is upregulated in AML with *KMT2A (MLL)* gene partial tandem duplication
473 (an alteration reported in AML) and knockdown significantly decreased leukemic cell growth⁶²,
474 implicating *SLAMF8* as having oncogenic function in AML.

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

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

484 **Acknowledgements**

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

489

490 **Author Contributions**

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

499

500 **Conflict of Interest**

501 The authors declare no conflicts of interest.

502

503

504 **References**

- 505 1. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute
506 myeloid leukemia. *N Engl J Med*. 2016;374(23):2209-2221. doi:10.1056/NEJMoa1516192
- 507 2. Godley LA, Shimamura A. Genetic predisposition to hematologic malignancies: management
508 and surveillance. *Blood*. 2017;130(4):424-432. doi:10.1182/blood-2017-02-735290
- 509 3. Rio-Machin A, Vulliamy T, Hug N, et al. The complex genetic landscape of familial MDS and
510 AML reveals pathogenic germline variants. *Nat Commun*. 2020;11(1):1044.
511 doi:10.1038/s41467-020-14829-5
- 512 4. Fletcher O, Houlston RS. Architecture of inherited susceptibility to common cancer. *Nat Rev*
513 *Cancer*. 2010;10(5):353-361. doi:10.1038/nrc2840
- 514 5. Khera A V, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases
515 identify individuals with risk equivalent to monogenic mutations. *Nat Genet*.
516 2018;50(9):1219-1224. doi:10.1038/s41588-018-0183-z
- 517 6. Lin WY, Fordham SE, Hungate E, et al. Genome-wide association study identifies susceptibility
518 loci for acute myeloid leukemia. *Nat Commun*. 2021;12(1):6233. doi:10.1038/s41467-021-
519 26551-x
- 520 7. The Cancer Genome Atlas Research Network. Genomic and Epigenomic Landscapes of Adult
521 De Novo Acute Myeloid Leukemia. *New England Journal of Medicine*. 2013;368(22):2059-
522 2074. doi:10.1056/NEJMoa1301689
- 523 8. Parkin B, Ouilllette P, Wang Y, et al. NF1 inactivation in adult acute myelogenous leukemia.
524 *Clinical Cancer Research*. 2010;16(16):4135-4147. doi:10.1158/1078-0432.CCR-09-2639
- 525 9. Parkin B, Erba H, Ouilllette P, et al. Acquired genomic copy number aberrations and survival in
526 adult acute myelogenous leukemia. *Blood*. 2010;116(23):4958-4967. doi:10.1182/blood-
527 2010-01-266999
- 528 10. Parkin B, Ouilllette P, Yildiz M, Saiya-Cork K, Shedden K, Malek SN. Integrated genomic
529 profiling, therapy response, and survival in adult acute myelogenous leukemia. *Clinical Cancer*
530 *Research*. 2015;21(9):2045-2056. doi:10.1158/1078-0432.CCR-14-0921
- 531 11. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of
532 genomes. *Nat Methods*. 2011;9(2):179-181. doi:10.1038/nmeth.1785
- 533 12. Goldstone AH, Burnett AK, Wheatley K, et al. Attempts to improve treatment outcomes in
534 acute myeloid leukemia (AML) in older patients: the results of the United Kingdom Medical
535 Research Council AML11 trial. *Blood*. 2001;98(5):1302-1311.
536 <https://www.ncbi.nlm.nih.gov/pubmed/11520775>
- 537 13. Burnett AK, Hills RK, Milligan DW, et al. Attempts to optimize induction and consolidation
538 treatment in acute myeloid leukemia: results of the MRC AML12 trial. *J Clin Oncol*.
539 2010;28(4):586-595. doi:10.1200/JCO.2009.22.9088

- 540 14. Schaich M, Parmentier S, Kramer M, et al. High-dose cytarabine consolidation with or without
541 additional amsacrine and mitoxantrone in acute myeloid leukemia: Results of the prospective
542 randomized AML2003 trial. In: *Journal of Clinical Oncology*. Vol 31. 2013.
543 doi:10.1200/JCO.2012.46.4743
- 544 15. Röllig C, Bornhäuser M, Kramer M, et al. Allogeneic stem-cell transplantation in patients with
545 NPM1-mutated acute myeloid leukemia: Results from a prospective donor versus no-donor
546 analysis of patients after upfront HLA typing within the SAL-AML 2003 trial. *Journal of Clinical
547 Oncology*. 2015;33(5). doi:10.1200/JCO.2013.54.4973
- 548 16. Cai S, Hartley A, Mahmoud O, Tilling K, Dudbridge F. Adjusting for collider bias in genetic
549 association studies using instrumental variable methods. *Genet Epidemiol*. 2022;46(5-6):303-
550 316. doi:10.1002/gepi.22455
- 551 17. Kar SP, Quiros PM, Gu M, et al. Genome-wide analyses of 200,453 individuals yield new
552 insights into the causes and consequences of clonal hematopoiesis. *Nat Genet*.
553 2022;54(8):1155-1166. doi:10.1038/s41588-022-01121-z
- 554 18. Kessler MD, Damask A, O’Keeffe S, et al. Common and rare variant associations with clonal
555 haematopoiesis phenotypes. *Nature*. 2022;612(7939):301-309. doi:10.1038/s41586-022-
556 05448-9
- 557 19. Weinstock JS, Gopakumar J, Burugula BB, et al. Aberrant activation of TCL1A promotes stem
558 cell expansion in clonal haematopoiesis. *Nature*. 2023;616(7958):755-763.
559 doi:10.1038/s41586-023-05806-1
- 560 20. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for
561 genotype imputation. *Nat Genet*. 2016;48(10):1279-1283. doi:10.1038/ng.3643
- 562 21. Döhner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022
563 recommendations from an international expert panel on behalf of the ELN. *Blood*.
564 2022;140(12):1345-1377. doi:10.1182/blood.2022016867
- 565 22. Laflamme R, Lisi V, Hébert J, et al. Replication of a GWAS signal near HLA-DQA2 with AML
566 using a disease-only cohort and external population-based controls. *Blood Neoplasia*.
567 Published online May 2025:100118. doi:10.1016/j.bneo.2025.100118
- 568 23. Ley TJ, Ding L, Walter MJ, et al. *DNMT3A* mutations in acute myeloid leukemia. *New England
569 Journal of Medicine*. 2010;363(25):2424-2433. doi:10.1056/NEJMoa1005143
- 570 24. Asada S, Kitamura T. Clonal hematopoiesis and associated diseases: A review of recent
571 findings. *Cancer Sci*. 2021;112(10):3962-3971. doi:10.1111/cas.15094
- 572 25. Young AL, Tong RS, Birmann BM, Druley TE. Clonal hematopoiesis and risk of acute myeloid
573 leukemia. *Haematologica*. 2019;104(12):2410-2417. doi:10.3324/haematol.2018.215269
- 574 26. Zink F, Stacey SN, Norddahl GL, et al. Clonal hematopoiesis, with and without candidate
575 driver mutations, is common in the elderly. *Blood*. 2017;130(6):742-752. doi:10.1182/blood-
576 2017-02-769869

- 577 27. Stacey SN, Zink F, Halldorsson GH, et al. Genetics and epidemiology of mutational barcode-
578 defined clonal hematopoiesis. *Nat Genet.* 2023;55(12):2149-2159. doi:10.1038/s41588-023-
579 01555-z
- 580 28. Bojjireddy N, Guzman-Hernandez ML, Reinhard NR, Jovic M, Balla T. EFR3s are palmitoylated
581 plasma membrane proteins that control responsiveness to G protein-coupled receptors. *J Cell*
582 *Sci.* Published online January 1, 2014. doi:10.1242/jcs.157495
- 583 29. Jiang X, Huang X, Zheng G, et al. Targeting PI4KA sensitizes refractory leukemia to
584 chemotherapy by modulating the ERK/AMPK/OXPHOS axis. *Theranostics.* 2022;12(16):6972-
585 6988. doi:10.7150/thno.76563
- 586 30. Wang Y, Shen X, Wang Q, et al. Non-canonical small GTPase RBJ promotes NSCLC progression
587 through the canonical MEK/ERK signaling pathway. *Curr Pharm Des.* 2022;28(42):3446-3455.
588 doi:10.2174/1381612829666221117124048
- 589 31. Chen T, Yang M, Yu Z, et al. Small GTPase RBJ mediates nuclear entrapment of MEK1/MEK2 in
590 tumor progression. *Cancer Cell.* 2014;25(5):682-696. doi:10.1016/j.ccr.2014.03.009
- 591 32. Liu Q, Zhu H, Zhang C, Chen T, Cao X. Small GTPase RBJ promotes cancer progression by
592 mobilizing MDSCs via IL-6. *Oncoimmunology.* 2017;6(1):e1245265.
593 doi:10.1080/2162402X.2016.1245265
- 594 33. Yang L, Rau R, Goodell MA. DNMT3A in haematological malignancies. *Nat Rev Cancer.*
595 2015;15(3):152-165. doi:10.1038/nrc3895
- 596 34. Bellido Moliás F, Sim A, Leong KW, et al. Antisense RNAs influence promoter usage of their
597 counterpart sense genes in cancer. *Cancer Res.* 2021;81(23):5849-5861. doi:10.1158/0008-
598 5472.CAN-21-1859
- 599 35. Zhang L, Chen P, Yang S, et al. CHKA mediates the poor prognosis of lung adenocarcinoma
600 and acts as a prognostic indicator. *Oncol Lett.* 2016;12(3):1849-1853.
601 doi:10.3892/ol.2016.4810
- 602 36. Asim M, Massie CE, Orafidiya F, et al. Choline kinase alpha as an androgen receptor
603 chaperone and prostate cancer therapeutic target. *J Natl Cancer Inst.* 2016;108(5):djv371.
604 doi:10.1093/jnci/djv371
- 605 37. Hu L, Wang RY, Cai J, et al. Overexpression of CHKA contributes to tumor progression and
606 metastasis and predicts poor prognosis in colorectal carcinoma. *Oncotarget.*
607 2016;7(41):66660-66678. doi:10.18632/oncotarget.11433
- 608 38. Gokhale S, Xie P. ChoK-Full of Potential: Choline Kinase in B Cell and T Cell Malignancies.
609 *Pharmaceutics.* 2021;13(6):911. doi:10.3390/pharmaceutics13060911
- 610 39. Marchitti SA, Brocker C, Orlicky DJ, Vasiliou V. Molecular characterization, expression
611 analysis, and role of ALDH3B1 in the cellular protection against oxidative stress. *Free Radic*
612 *Biol Med.* 2010;49(9):1432-1443. doi:10.1016/j.freeradbiomed.2010.08.004

- 613 40. Dancik GM, Varisli L, Tolan V, Vlahopoulos S. Aldehyde dehydrogenase genes as prospective
614 actionable targets in acute myeloid leukemia. *Genes (Basel)*. 2023;14(9):1807.
615 doi:10.3390/genes14091807
- 616 41. Venton G, Pérez-Alea M, Baier C, et al. Aldehyde dehydrogenases inhibition eradicates
617 leukemia stem cells while sparing normal progenitors. *Blood Cancer J*. 2016;6(9):e469-e469.
618 doi:10.1038/bcj.2016.78
- 619 42. Xu W, Fang H, Cao X, et al. NADH:ubiquinone oxidoreductase core subunit S8 expression and
620 functional significance in non-small cell lung cancer. *Cell Death Dis*. 2025;16(1):321.
621 doi:10.1038/s41419-025-07638-5
- 622 43. Bassal MA, Samaraweera SE, Lim K, et al. Germline mutations in mitochondrial complex I
623 reveal genetic and targetable vulnerability in IDH1-mutant acute myeloid leukaemia. *Nat*
624 *Commun*. 2022;13(1):2614. doi:10.1038/s41467-022-30223-9
- 625 44. Bröhm A, Elsayy H, Rathert P, et al. Somatic cancer mutations in the SUV420H1 protein lysine
626 methyltransferase modulate its catalytic activity. *J Mol Biol*. 2019;431(17):3068-3080.
627 doi:10.1016/j.jmb.2019.06.021
- 628 45. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of
629 the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet*.
630 2013;14(11):765-780. doi:10.1038/nrg3554
- 631 46. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in
632 cancer aetiology and progression. *Nat Rev Genet*. 2016;17(5):284-299.
633 doi:10.1038/nrg.2016.13
- 634 47. Mian SA, Rouault-Pierre K, Smith AE, et al. SF3B1 mutant MDS-initiating cells may arise from
635 the haematopoietic stem cell compartment. *Nat Commun*. 2015;6(1):10004.
636 doi:10.1038/ncomms10004
- 637 48. Zhang B, Kaufman RJ, Ginsburg D. LMAN1 and MCFD2 form a cargo receptor complex and
638 interact with coagulation factor VIII in the early secretory pathway. *Journal of Biological*
639 *Chemistry*. 2005;280(27):25881-25886. doi:10.1074/jbc.M502160200
- 640 49. Guy JE, Wigren E, Svärd M, Härd T, Lindqvist Y. New insights into multiple coagulation factor
641 deficiency from the solution structure of human MCFD2. *J Mol Biol*. 2008;381(4):941-955.
642 doi:10.1016/j.jmb.2008.06.042
- 643 50. Liu H, Zhao B, Chen Y, et al. Multiple coagulation factor deficiency protein 2 contains the
644 ability to support stem cell self-renewal. *The FASEB Journal*. 2013;27(8):3298-3305.
645 doi:10.1096/fj.13-228825
- 646 51. Toda H, Tsuji M, Nakano I, et al. Stem cell-derived neural stem/progenitor cell supporting
647 factor is an autocrine/paracrine survival factor for adult neural stem/progenitor cells. *Journal*
648 *of Biological Chemistry*. 2003;278(37):35491-35500. doi:10.1074/jbc.M305342200

- 649 52. Chang L, Cui Z, Shi D, et al. Polyclonal evolution of Fanconi anemia to MDS and AML revealed
650 at single cell resolution. *Exp Hematol Oncol.* 2022;11(1):64. doi:10.1186/s40164-022-00319-5
- 651 53. Merfort LW, Lisboa M de O, Cavalli LR, Bonfim CMS. Cytogenetics in Fanconi Anemia: The
652 importance of follow-up and the search for new biomarkers of genomic instability. *Int J Mol*
653 *Sci.* 2022;23(22):14119. doi:10.3390/ijms232214119
- 654 54. Atashrazm F, Ellis S. The polarity protein PARD3 and cancer. *Oncogene.* 2021;40(25):4245-
655 4262. doi:10.1038/s41388-021-01813-6
- 656 55. Catterall R, Lelarge V, McCaffrey L. Genetic alterations of epithelial polarity genes are
657 associated with loss of polarity in invasive breast cancer. *Int J Cancer.* 2020;146(6):1578-1591.
658 doi:10.1002/ijc.32691
- 659 56. Li T, Liu D, Lei X, Jiang Q. Par3L enhances colorectal cancer cell survival by inhibiting
660 Lkb1/AMPK signaling pathway. *Biochem Biophys Res Commun.* 2017;482(4):1037-1041.
661 doi:10.1016/j.bbrc.2016.11.154
- 662 57. Gao PP, Qi XW, Sun N, et al. The emerging roles of dual-specificity phosphatases and their
663 specific characteristics in human cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on*
664 *Cancer.* 2021;1876(1):188562. doi:10.1016/j.bbcan.2021.188562
- 665 58. Gallegos LL, Ng MR, Sowa ME, et al. A protein interaction map for cell-cell adhesion
666 regulators identifies DUSP23 as a novel phosphatase for β -catenin. *Sci Rep.* 2016;6(1):27114.
667 doi:10.1038/srep27114
- 668 59. Liu X, Zhuang H, Li F, Lu Y, Pei R. Prognostic Significance of Dual-Specificity Phosphatase 23
669 Expression in Acute Myeloid Leukemia. *J Blood Med.* 2024;Volume 15:35-50.
670 doi:10.2147/JBM.S437400
- 671 60. Magari M, Nishioka M, Hari T, et al. The immunoreceptor SLAMF8 promotes the
672 differentiation of follicular dendritic cell-dependent monocytic cells with B cell-activating
673 ability. *FEBS Lett.* 2022;596(20):2659-2667. doi:10.1002/1873-3468.14468
- 674 61. Wang G, Abadía-Molina AC, Berger SB, et al. Cutting edge: Slamf8 Is a negative regulator of
675 Nox2 activity in macrophages. *The Journal of Immunology.* 2012;188(12):5829-5832.
676 doi:10.4049/jimmunol.1102620
- 677 62. Bill M, Goda C, Pepe F, et al. Targeting BRD4 in acute myeloid leukemia with partial tandem
678 duplication of the MLL gene. *Haematologica.* 2021;106(9):2527-2532.
679 doi:10.3324/haematol.2020.271627
- 680 63. Agarwal G, Antoszewski M, Xie X, et al. Inherited resilience to clonal hematopoiesis by
681 modifying stem cell RNA regulation. *BioRxiv.* doi:10.1101/2025.03.24.645017

682

683

684

685 **Figure Legends**

686 **Figure 1 – Manhattan plots from acute myeloid leukemia meta-analysis of genome-wide**
687 **association studies.** For each GWAS, association tests were performed for all AML cases (pan-
688 AML), cytogenetically normal AML, del(5/7) AML and complex karyotype AML assuming an
689 additive genetic model, with nominally significant principal components included in the analysis as
690 covariates. Association summary statistics were combined for variants common to all GWAS, in fixed
691 effects models using PLINK (GWAS5 was excluded from the meta-analysis of del(5/7) AML and
692 cytogenetically complex AML due to low case numbers). Manhattan plots show negative log₁₀ (fixed
693 effects meta *P* values, *Y* axis) for pan-AML (a), cytogenetically normal AML (b), del(5/7) AML (c)
694 and cytogenetically complex AML (d) over 22 autosomal chromosomes. Risk loci are annotated with
695 chromosome position and local genes. All statistical tests were two-sided and no adjustments were
696 made for multiple comparisons. The horizontal red line denotes the threshold for statistical
697 significance in a genome-wide association study ($P < 5.0 \times 10^{-8}$).

698
699 **Figure 2 – Forest plots for loci associated with acute myeloid leukemia.** Study cohorts, sample
700 sizes (case and controls (con)), imputation (info) score, effect allele, effect allele frequencies (EAF)
701 and estimated odds ratios (OR) for rs11481 (pan-AML) (a), rs4665765 (pan-AML) (b), rs3916765
702 (cytogenetically normal AML) (c), rs12078864 (del(5/7) AML) (d), rs79918355 (complex karyotype
703 AML) (e) and rs12988876 (complex karyotype AML) (f). The vertical line corresponds to the null
704 hypothesis (odds ratio (OR) = 1). The horizontal lines and square brackets indicate 95% confidence
705 intervals (95% CI). Areas of the boxes are proportional to the weight of the study. Diamonds
706 represent combined estimates for fixed-effect and random-effect analysis. Cochran's Q statistic was
707 used to test for heterogeneity where $P_{\text{HET}} > 0.05$ indicates non-significant heterogeneity. The
708 heterogeneity index, I^2 (0-100) was also measured which quantifies the proportion of the total
709 variation due to heterogeneity. All statistical tests were two-sided and no adjustments were made for
710 multiple comparisons.

711
712 **Figure 3 – Regional association and linkage disequilibrium plots for loci associated with acute**
713 **myeloid leukemia.** For each GWAS, association tests were performed for pan-AML cases,
714 cytogenetically normal AML, del(5/7) AML and cytogenetically complex AML assuming an additive
715 genetic model, with nominally significant principal components included in the analysis as covariates.
716 Association summary statistics were combined for variants common to all 6 GWAS, in fixed effects
717 models using PLINK. Negative log₁₀-transformed *P* values (left *Y* axis) from the meta-analysis of 6
718 GWAS are shown for variants at 11q13 (a) and 1p31.1 (b) for pan-AML, and at 6p21.32 (c) for
719 cytogenetically normal AML. Negative log₁₀-transformed *P* values (left *Y* axis) from the meta-
720 analysis of 5 GWAS are shown for variants at 1q23.2 (d) for del(5/7) AML, and at 2p21 (e) and
721 2q33.3 (f) for cytogenetically complex AML. All statistical tests were two-sided and no adjustments

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

726

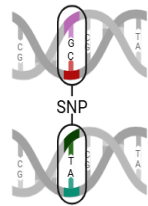
727 **Figure 4 – Forest plots for variants reported to be associated with risk of clonal hematopoiesis**
728 **(CH) and/or clonal hematopoiesis of indeterminate potential (CHIP) and their association with**
729 **risk of AML.** Forest plots show 7 variants reported by Kar *et al* (2022)¹⁶ and Kessler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

740

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

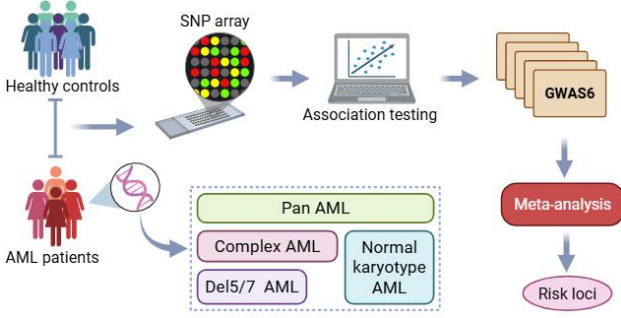
Context of Research

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



Patients and Methods

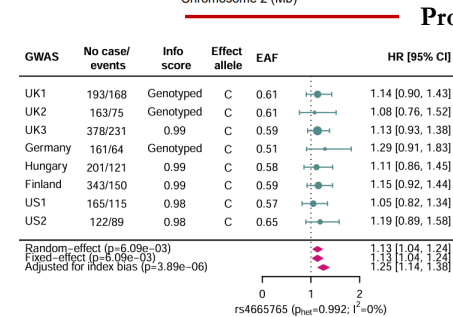
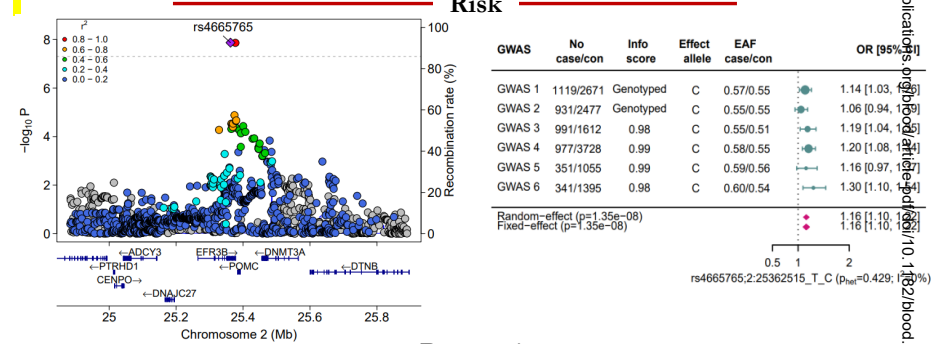
- 4710 AML patients and 12938 controls of European ancestry genotyped using SNP arrays.



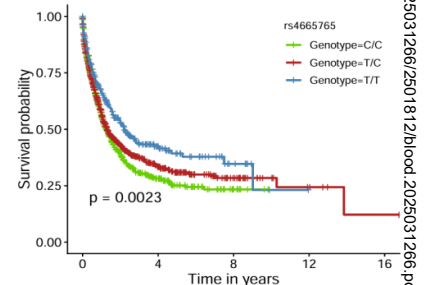
- We performed a meta-analysis of six GWAS to identify risk alleles for pan-AML and sub-groups.

Main Findings

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



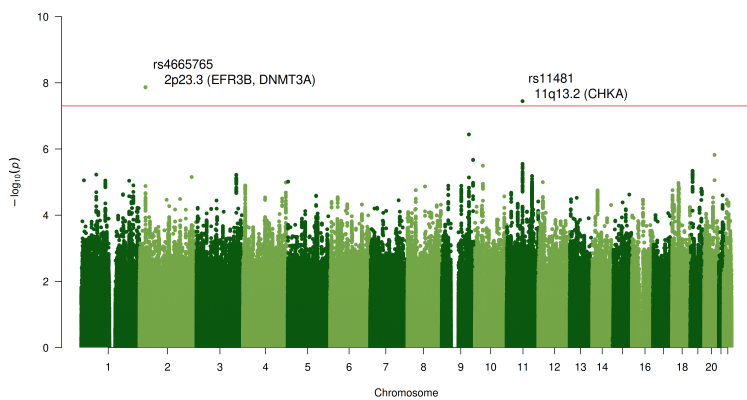
Prognosis



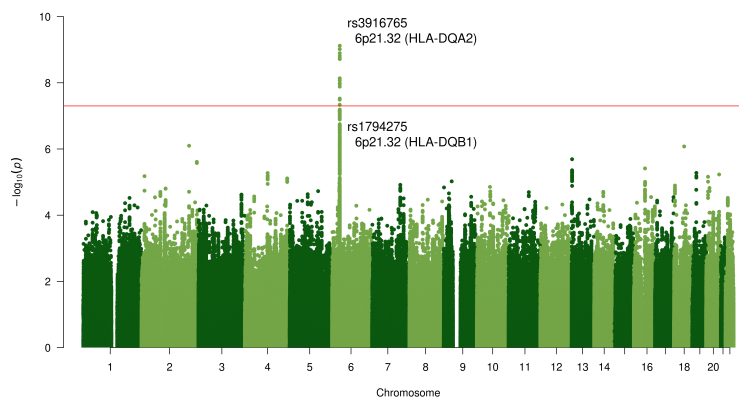
CONCLUSIONS: We identify a new genome-wide significant risk locus for pan-AML and three new risk loci for disease sub-groups, including AML with deletions of chromosome 5 and/or 7 and cytogenetically complex AML. We also identify variants previously associated with risk of clonal hematopoiesis (CH) that 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

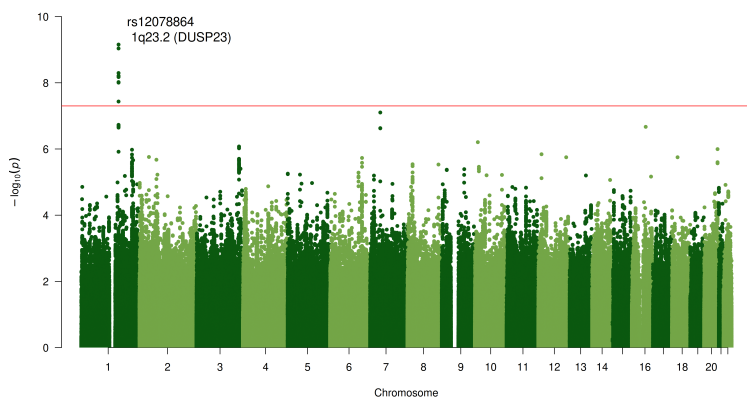
(a) Pan-AML



(b) Cytogenetically normal AML



(c) Del(5/7) AML



(d) Complex karyotype AML

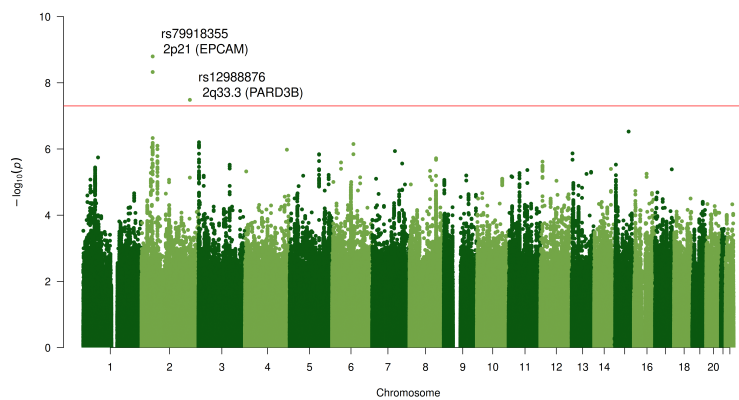
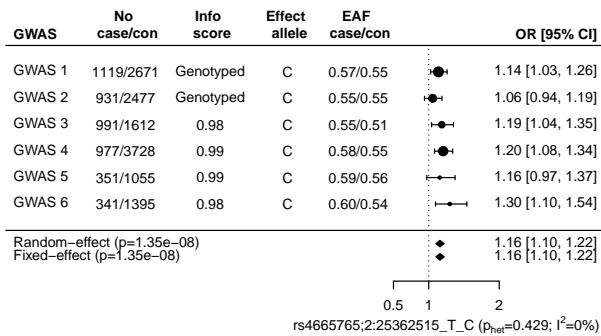
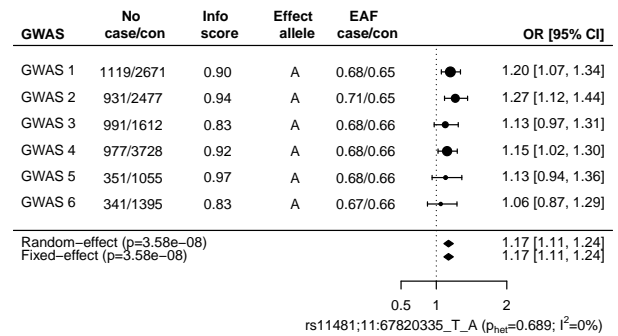


Figure 2

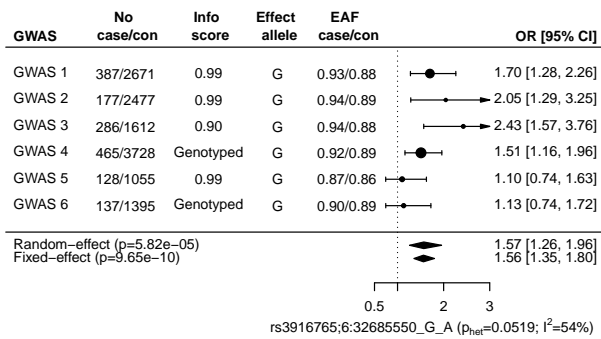
(a) Pan-AML 2p23.3



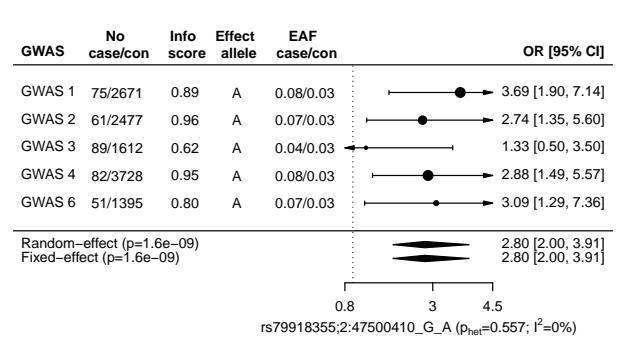
(b) Pan-AML 11q13.2



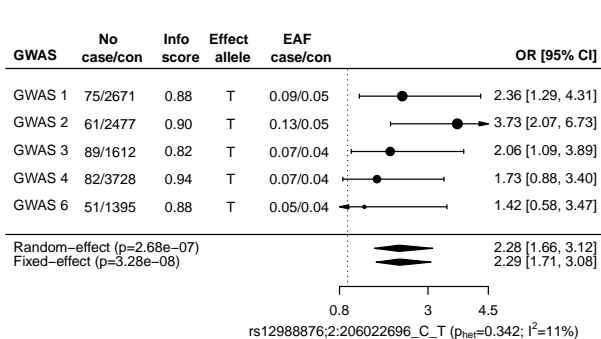
(c) Cytogenetically normal AML 6p21.32



(d) Complex karyotype AML 2p21



(e) Complex karyotype 2q33.3



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

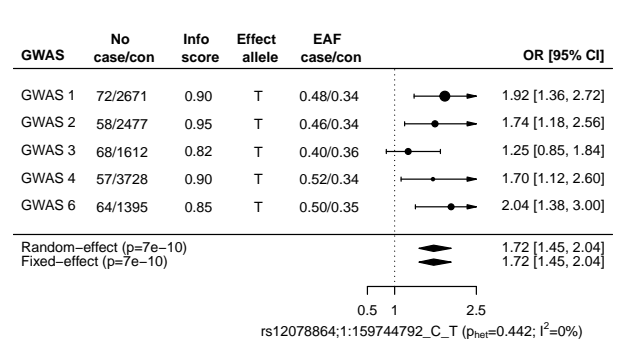
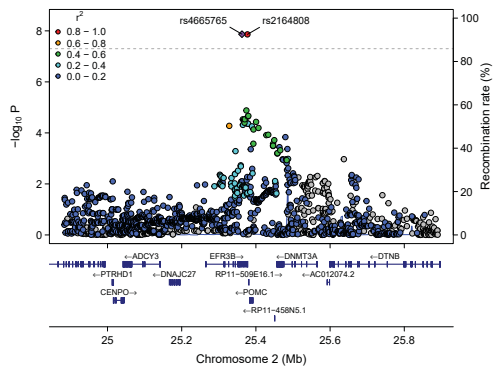
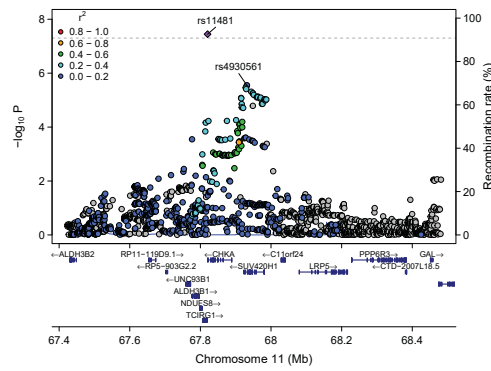


Figure 3

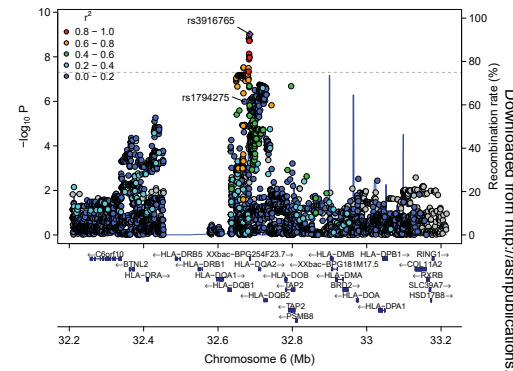
(a) Pan-AML 2p23.3



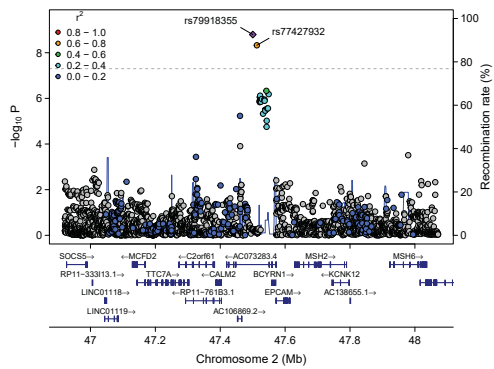
(b) Pan-AML 11q13.2



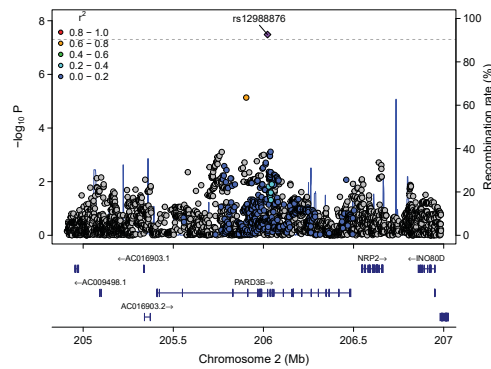
(c) Cytogenetically normal AML 6p21.32



(d) Complex karyotype AML 2p21



(e) Complex karyotype 2q33.3



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

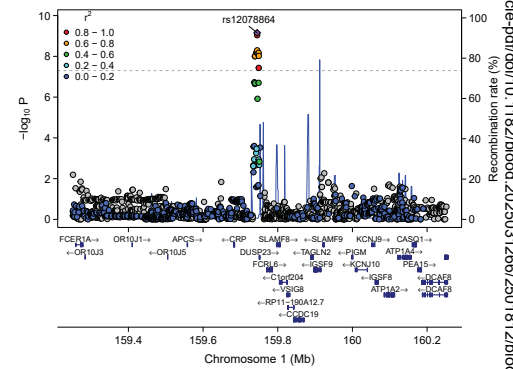
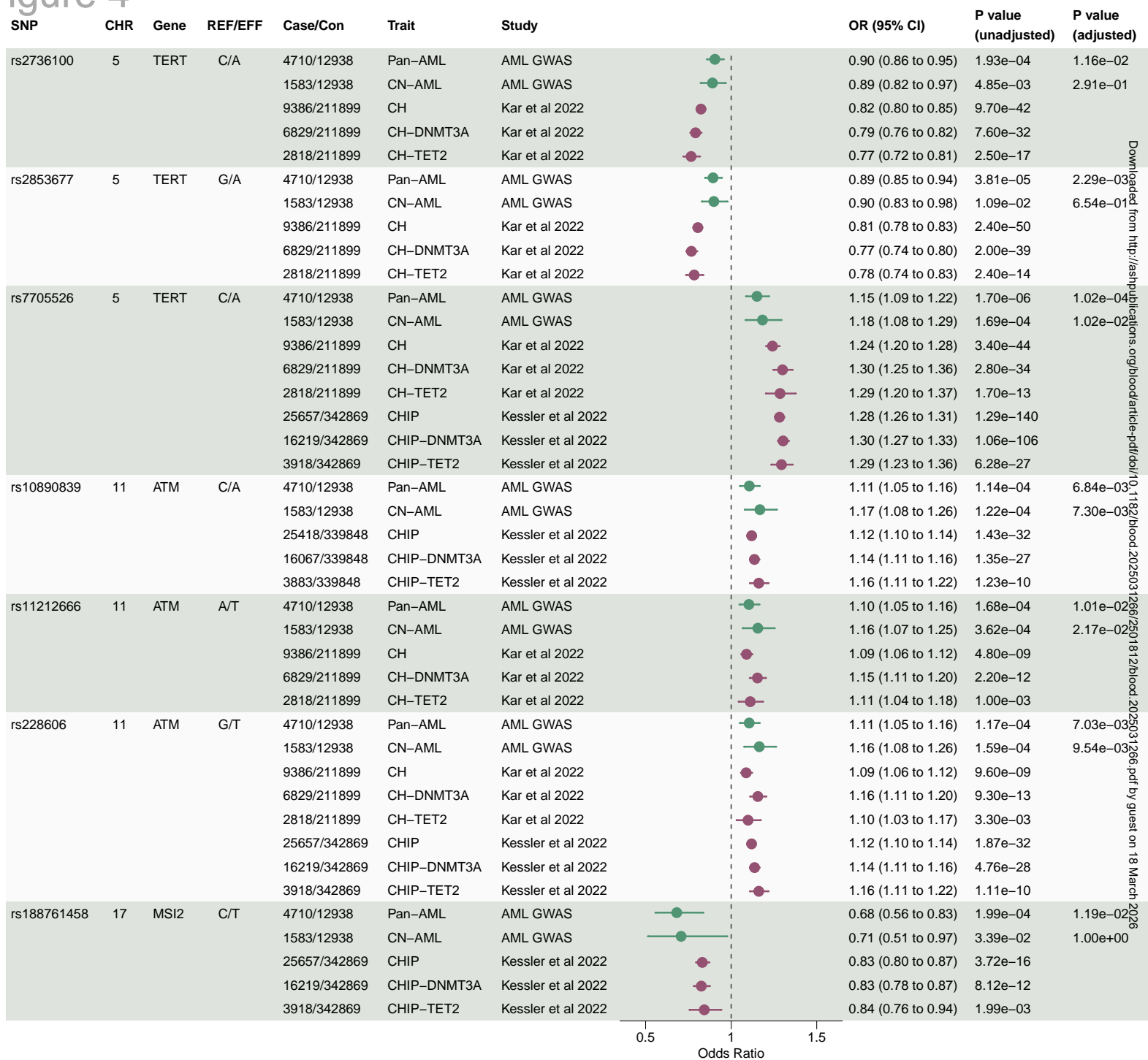


Figure 4



Downloaded from <http://ashpublications.org/blood/article-pdf/doi/10.1182/blood.2025031266/2501812/blood.2025031266.pdf> by guest on 18 March 2025