**Table S1.** Sample number and demographic characteristics in each ethnic groups in each cohort

| cohort | | CCRLP | | | | Kaiser GERA | | | | CCLS | | COG | WTCCC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| population | | AFR | EAS | LAT | NLW | AFR | EAS | LAT | NLW | LAT | NLW | NLW | NLW |
| sample count | **case** | 124 | 318 | 1878 | 1162 | - | - | - | - | 750 | 472 | 1504 | - |
| | **control** | 141 | 318 | 2023 | 1229 | 1926 | 4699 | 6387 | 56112 | 504 | 340 | - | 2931 |
| | **female** | 106 | 286 | 1701 | 1075 | 1207 | 2831 | 4088 | 33294 | 565 | 321 | 686 | 1269 |
| | **male** | 159 | 350 | 2200 | 1316 | 719 | 1868 | 2329 | 22818 | 689 | 491 | 818 | 1662 |
| | **B-ALL** | 68 | 195 | 1210 | 652 | - | - | - | - | 183 | 111 | 380 | - |
| | **T-ALL** | 17 | 32 | 120 | 110 | - | - | - | - | 39 | 25 | 65 | - |
| **mean age of diagnosis (stdev)** | | 5.31 (3.41) | 4.49 (3.31) | 4.85 (3.54) | 4.51 (3.30) | - | - | - | - | 5.61 (3.59) | 5.70 (3.54) | - | - |

AFR: African American; EAS: East Asian American; LAT: Latino American; NLW: non-Latino White American

**Table S2**. Summary statistics for the previously reported susceptibility variants, the top variant in the loci from our meta-analysis, and the linkage disequilibrium between the two variants in NLW and LAT.

| | Reported SNP | | | | Top SNP in this study | | | | $r^2$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Gene | Chr | Pos | rsID (reference) | P-value | Chr | Pos | rsID | P-value | NLW | LAT |
| *C5orf56* | 5 | 131765206 | rs886285 (ref (1)) | 0.63 | 5 | 131811182 | rs11741255 | $1.69 \times 10^{-4}$ | 0.35 | 0.19 |
| *BAK1* | 6 | 33546930 | rs210143 (ref (1)) | $4.49 \times 10^{-8}$ | 6 | 33546837 | rs210142 | $4.27 \times 10^{-8}$ | 1 | 1 |
| *IKZF1* | 7 | 50470604 | rs4132601 (ref (2)) | $1.13 \times 10^{-33}$ | 7 | 50477144 | rs10230978 | $3.92 \times 10^{-34}$ | 0.98 | 0.97 |
| 8q24 | 8 | 130156143 | rs4617118 (ref (3)) | $1.04 \times 10^{-12}$ | Same | | | | | |
| *CDKN2A* | 9 | 21970916 | rs3731249 (ref (4)) | $1.29 \times 10^{-18}$ | 9 | 21975319 | rs36228834 | $1.90 \times 10^{-18}$ | 0.99 | 1 |
| *TLE1* | 9 | 83747371 | rs76925697 (ref (5)) | $5.37 \times 10^{-2}$ | 9 | 83728588 | rs62579826 | $1.06 \times 10^{-2}$ | 0.81 | 0.98 |
| *GATA3* | 10 | 8104208 | rs3824662 (ref (6)) | $4.24 \times 10^{-9}$ | Same | | | | | |
| *PIP4K2A* | 10 | 22852948 | rs7088318 (ref (7)) | $6.50 \times 10^{-19}$ | 10 | 22853102 | rs7075634 | $2.42 \times 10^{-19}$ | 0.96 | 0.97 |
| *BMI1* | 10 | 22423302 | rs11591377 (ref (8)) | $8.21 \times 10^{-10}$ | 10 | 22374489 | rs1926697 | $5.24 \times 10^{-10}$ | 0.84 | 0.88 |
| *ARID5B* | 10 | 63723577 | rs10821936 (ref (9)) | $4.78 \times 10^{-67}$ | 10 | 63721176 | rs7090445 | $7.36 \times 10^{-70}$ | 0.98 | 0.99 |
| *LHPP* | 10 | 126293309 | rs35837782 (ref (10)) | $6.90 \times 10^{-4}$ | Same | | | | | |
| *ELK3* | 12 | 96612762 | rs4762284 (ref (10)) | $2.42 \times 10^{-3}$ | 12 | 96645605 | rs78405390 | $4.68 \times 10^{-5}$ | 0.13 | 0.22 |
| *CEBPE* | 14 | 23589057 | rs2239633 (ref (2)) | $3.0 \times 10^{-14}$ | 14 | 23589349 | rs2239630 | $2.12 \times 10^{-21}$ | 0.74 | 0.78 |
| *IKZF3* | 17 | 38066240 | rs2290400 (ref (3)) | $2.09\ 10^{-6}$ | 17 | 37957235 | rs17607816 | $1.42 \times 10^{-7}$ | 0.02 | 0.22 |
| *IGF2BP1* | 17 | 47092076 | rs10853104 (ref (5)) | $2.93 \times 10^{-2}$ | 17 | 47217004 | rs6504598 | $4.87 \times 10^{-4}$ | 0.02 | 0.02 |
| *ERG* | 21 | 39789606 | rs8131436 (ref (11)) | $6.97 \times 10^{-5}$ | 21 | 39784752 | rs55681902 | $9.36 \times 10^{-6}$ | 0.62 | 0.65 |

Reported SNP: published lead SNP; Top SNP in this study: top SNP at each locus within 1Mb of the previously reported SNP from our meta-analysis. Note that out of the 16 loci, three (8q24.21, *IKZF3*, and *BMI1*) were initially identified and five (*IKZF1, PIP4K2A, ARID5B, CDKN2A, CEBPE*) were previously shown to be replicated using a smaller but largely overlapping subset of this dataset(3,8) . For these loci, our findings here would not necessarily constitute an independent replication. Gene names (gene) are given based on the nearest gene unless the variant is in gene desert. Chromosome (Chr) and position (Pos) are given in hg19 coordinates. r2 denotes the squared correlation of the reported SNP and our top SNP in NLW from discovery cohort; NLW and LAT denote the non-Latino white and Latino cohorts, respectively.

**Table S4.** Association statistics of the three novel variants in each population-specific GWAS

| SNP | | AFR | | | EAS | | | LAT | | | EUR | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | CCRLP case | CCRLP control | GERA control | CCRLP case | CCRLP control | GERA control | CCRLP case | CCRLP control | GERA control | CCRLP case | CCRLP control | GERA control |
| | sample size | 124 | 141 | 1926 | 318 | 318 | 4699 | 1878 | 2023 | 6387 | 1162 | 1229 | 56112 |
| 6q23 rs9376090 MYB/HBS1L | MAF (C) | 0.043 | 0.052 | 0.055 | 0.115 | 0.106 | 0.129 | 0.114 | 0.122 | 0.172 | 0.170 | 0.192 | 0.236 |
| | Rsq | 0.740 | | 0.840 | 0.540 | | 0.570 | 0.820 | | 0.840 | 0.750 | | 0.940 |
| | OR (P-value) | 0.661(0.228) | | | 1.04(0.803) | | | 0.893(0.075) | | | 0.684(1.19E-11) | | |
| | HetPVal | 0.006 | | | | | | | | | | | |
| 10q21 rs9415680 NRBF2/JMJD1C | MAF (A) | 0.090 | 0.053 | 0.060 | 0.393 | 0.354 | 0.322 | 0.405 | 0.360 | 0.249 | 0.180 | 0.162 | 0.149 |
| | Rsq | 0.820 | | 0.850 | 0.940 | | 0.940 | 0.980 | | 0.970 | 0.970 | | 0.970 |
| | OR (P-value) | 1.566(0.098) | | | 1.336(0.0012) | | | 1.197(1.79E-05) | | | 1.146(0.0183) | | |
| | HetPVal | 0.3741 | | | | | | | | | | | |
| 10q21 rs10998283 TET1 | MAF (A) | 0.029 | 0.024 | 0.033 | 0.145 | 0.165 | 0.141 | 0.200 | 0.166 | 0.139 | 0.145 | 0.124 | 0.123 |
| | Rsq | 0.950 | | 0.940 | 0.880 | | 0.880 | 0.980 | | 0.980 | 0.980 | | 0.980 |
| | OR (P-value) | 0.896(0.788) | | | 1.014(0.916) | | | 1.270(3.22E-06) | | | 1.244(0.0005) | | |
| | HetPVal | 0.3567 | | | | | | | | | | | |

AFR: African American; EAS: East Asian American; LAT: Latino American; NLW: non-Latino White; SNP: locus position, rsID and the nearest genes for the lead SNP in each locus; Rsq: estimate of imputation quality from Michigan Imputation Server, computed in CCRLP and GERA cohort separately; MAF: Minor Allele Frequency computed in CCRLP and GERA cohort separately with imputed genotype dosage using PLINK2. The minor allele for each SNP is noted in parentheses; OR, Odds Ratio from logistic regression in each population given with respect to the minor allele; HetPVal: P-value from Cochran's Q-test for heterogeneity for meta-analysis across four populations. Note that all associated SNPs have passed post-imputation filtering to control for batch effect and systematic bias between CCRLP and GERA (Supplementary Information).

**Table S7.** Summary of conditional analysis to identify secondary associations at known loci.

| Gene | Chr | Pos | rsID | Risk allele | OR | $P_{conditional}$ | $P_{discovery}$ | $r^2$ |
|---|---|---|---|---|---|---|---|---|
| *IKZF1* | 7 | 50459043 | rs78396808 | A | 1.632 | $3.46x10^{-26}$ | $2.7x10^{-16}$ | *0.06 |
| *CDKN2A/B* | 9 | 21993964 | rs2811711 | T | 1.355 | 7.2x10-$^{10}$ | $1.85x10^{-11}$ | 0.01 |
| *CEBPE* | 14 | 23592617 | rs60820638 | A | 1.193 | $5.38x10^{-8}$ | 0.102 | 0.16 |
| *IZKF3* | 17 | 37983492 | rs12944882 | T | 1.204 | $7.71x10^{-10}$ | $2.81x10^{-7}$ | 0.02 |

For each of the four significant association after conditional analysis, we show the genomic coordinates in hg19, effect size (OR), the P-values with or without conditioning on the lead SNP from the discovery meta-analysis in the locus, and the r2 between the lead SNP and secondary association.

Chr, chromosome; Pos, Position in hg19; OR, odds ratio (effect size); $P_{conditional}$, p-value from the conditional analysis; $P_{discovery}$, p-value from meta-analysis without conditioning on any SNP; $r^2$, squared correlation of the conditioned SNP and the most significantly associated SNP from conditional analysis.

*calculated in Latino population as the variant was filtered out due to low MAF in NLW cohort.

**Table S8.** Conditional analysis at the CEBPE locus.

| rs60820638 | | rs2239635 | | rs2239630 | |
|---|---|---|---|---|---|
| **OR** | **P-value** | **OR** | **P-value** | **OR** | **P-value** |
| 1.051 | 0.106 | 1.289 | 3.06E-18 | 1.316 | 2.12E-21 |
| 1..193 | 5.38E-08 | 1.127 | 0.002017 | | * |
| | * | 1.294 | 3.34E-19 | 1.394 | 6.44E-28 |
| 1.019 | 0.5214 | | * | 1.209 | 7.15E-07 |
| 1.019 | 0.5214 | 1.027 | 0.56 | | * |
| | * | | * | 1.353 | 3.10E-10 |
| 1.177 | 2.47E-05 | | * | | * |

The conditioned SNPs are marked as *. The first row indicates the odds ratio and p-value from the marginal single-variant analysis. Subsequent rows indicate the odds ratio and p-values when one or two variants were included in a conditional multivariate analysis

**Table S9. Association of PRS with childhood ALL in replication cohorts**

| | | P-value | Beta | AUC |
|---|---|---|---|---|
| **COG/WTCCC** | **Known** | 2.99E-62 | 1.055 | 0.681 |
| | **All** | 6.93E-63 | 0.998 | 0.684 |
| **CCLS NLW** | **Known** | 2.22E-17 | 1.027 | 0.673 |
| | **All** | 2.03E-17 | 0.970 | 0.671 |
| **CCLS LAT** | **Known** | 4.78E-23 | 0.944 | 0.679 |
| | **All** | 5.75E-24 | 0.912 | 0.678 |

P-value, Beta: summary statistics from testing the association between PRS and case-control status in logistic regression; Known: PRS constructed with known loci; All: With known and novel loci reported in current paper.; AUC: Area Under Curve

**Table S10. Familial relative risk explained by the most significant SNPs from each risk loci associated with childhood ALL**

| | | | CCRLP/GERA | | | | | CCLS | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | OR | RAF | | FRR | | OR | RAF | | FRR | |
| Gene | SNP | Risk Allele | | LAT | NLW | LAT | NLW | | LAT | NLW | LAT | NLW |
| *C5orf56* | 5:131811182:G:A | G | 1.128 | 0.683 | 0.582 | 0.003 | 0.003 | 1.207 | 0.251 | 0.419 | 0.006 | 0.008 |
| *MYB/HBS1L** | 6:135411228:T:C* | T | 1.274 | 0.831 | 0.758 | 0.006 | 0.008 | 0.874 | 0.861 | 0.772 | 0.002 | 0.003 |
| *BAK1* | 6:33546837:T:C | C | 1.199 | 0.713 | 0.719 | 0.005 | 0.005 | 1.022 | 0.271 | 0.286 | 0.000 | 0.000 |
| *IKZF1* | 7:50459043:G:A | A | 1.425 | 0.124 | NA | 0.031 | NA | 1.415 | 0.200 | 0.010 | 0.020 | 0.001 |
| *IKZF1* | 7:50477144:G:A | A | 1.445 | 0.260 | 0.266 | 0.026 | 0.026 | 1.402 | 0.247 | 0.263 | 0.021 | 0.022 |
| 8q24 | 8:130156143:A:G | G | 1.289 | 0.153 | 0.174 | 0.008 | 0.009 | 1.325 | 0.893 | 0.839 | 0.005 | 0.008 |
| *CDKN2A* | 9:21975319:T:A | A | 1.879 | 0.023 | 0.029 | 0.014 | 0.018 | 2.223 | 0.021 | 0.024 | 0.025 | 0.027 |
| *CDKN2A* | 9:21993964:T:C | T | 1.355 | 0.883 | 0.860 | 0.006 | 0.008 | 1.342 | 0.915 | 0.854 | 0.005 | 0.007 |
| *TLE1* | 9:83728588:C:T | C | 1.249 | 0.965 | 0.961 | 0.001 | 0.001 | 1.494 | 0.027 | 0.037 | 0.005 | 0.007 |
| *LHPP* | 10:126293309:A:G | G | 1.117 | 0.558 | 0.634 | 0.003 | 0.002 | 1.136 | 0.501 | 0.371 | 0.003 | 0.003 |
| *BMI1* | 10:22374489:G:A | G | 1.250 | 0.767 | 0.759 | 0.007 | 0.007 | 1.200 | 0.206 | 0.246 | 0.005 | 0.006 |
| *PIP4K2A* | 10:22853102:T:C | C | 1.319 | 0.686 | 0.598 | 0.013 | 0.015 | 1.163 | 0.252 | 0.395 | 0.004 | 0.005 |
| *ARID5B* | 10:63721176:C:T | C | 1.642 | 0.408 | 0.336 | 0.052 | 0.052 | 1.910 | 0.581 | 0.674 | 0.071 | 0.058 |
| *JMJD1C/NRBF2 ** | 10:65020890:A:G* | A | 1.203 | 0.271 | 0.145 | 0.006 | 0.004 | 1.065 | 0.369 | 0.146 | 0.001 | 0.000 |
| *TET1** | 10:70329064:G:A* | A | 1.232 | 0.146 | 0.123 | 0.005 | 0.005 | 1.178 | 0.176 | 0.105 | 0.004 | 0.002 |
| *GATA3* | 10:8104208:C:A | A | 1.209 | 0.288 | 0.182 | 0.007 | 0.005 | 1.250 | 0.383 | 0.184 | 0.011 | 0.007 |
| *ELK3* | 12:96645605:A:C | C | 1.228 | 0.078 | 0.099 | 0.003 | 0.004 | 1.183 | 0.952 | 0.920 | 0.001 | 0.002 |
| *CEBPE* | 14:23589349:A:G | A | 1.394 | 0.527 | 0.461 | 0.023 | 0.023 | 1.292 | 0.565 | 0.468 | 0.013 | 0.014 |
| *CEBPE* | 14:23592617:A:C* | A | 1.193 | 0.702 | 0.706 | 0.005 | 0.005 | 1.210 | 0.701 | 0.676 | 0.006 | 0.006 |
| *IKZF3* | 17:37957235:T:C | C | 2.108 | 0.012 | 0.020 | 0.012 | 0.020 | 1.526 | 0.997 | 0.992 | 0.000 | 0.001 |
| *IKZF3* | 17:37983492:T:C* | T | 1.204 | 0.549 | 0.480 | 0.007 | 0.007 | 1.136 | 0.603 | 0.475 | 0.003 | 0.003 |
| *IGF2BP1* | 17:47217004:C:G | C | 1.113 | 0.639 | 0.645 | 0.002 | 0.002 | 1.071 | 0.620 | 0.633 | 0.001 | 0.001 |
| *ERG* | 21:39784752:T:C | C | 1.145 | 0.288 | 0.255 | 0.003 | 0.003 | 1.190 | 0.714 | 0.736 | 0.005 | 0.005 |
| | | all | | | | 0.250 | 0.233 | | | | 0.217 | 0.198 |
| | | novel | | | | 0.030 | 0.029 | | | | 0.016 | 0.016 |
| | | known | | | | 0.220 | 0.204 | | | | 0.202 | 0.182 |

* denotes novel risk loci and variants from the current study

Gene: nearest gene unless the variant is in gene desert; RAF: risk allele frequency in controls
FRR: The proportion of familial relative risk explained as calculated per Schumacher et al.

**Table S11. Heritability estimates using common SNPs (MAF > 0.05) stratified to 4 quartiles of LD score bins with REML or PCGC regression**

| | REML | | | | PCGC | | | |
|---|---|---|---|---|---|---|---|---|
| | NLW | | LAT | | NLW | | LAT | |
| | Heritability | SE | Heritability | SE | Heritability | SE | Heritability | SE |
| **group1** | 0.105 | 0.026 | -0.011 | 0.018 | 0.104 | 0.039 | 0.042 | 0.029 |
| **group2** | 0.023 | 0.023 | 0.001 | 0.015 | 0.021 | 0.035 | 0.010 | 0.028 |
| **group3** | 0.054 | 0.018 | 0.021 | 0.012 | 0.059 | 0.026 | 0.040 | 0.021 |
| **group4** | 0.021 | 0.010 | 0.030 | 0.008 | 0.017 | 0.013 | 0.019 | 0.012 |
| **sum** | 0.203 | 0.032 | 0.041 | 0.020 | 0.202 | 0.047 | 0.111 | 0.036 |

LAT: Latino American; NLW: non-Latino White

**Table S12.** Genetic correlation between LAT and NLW

| | Variance | SE |
|---|---|---|
| **V(G)_LAT** | 0.018 | 0.003 |
| **V(G)_NLW** | 0.021 | 0.002 |
| **C(G)** | 0.014 | 0.002 |
| **V(e)_LAT** | 0.112 | 0.003 |
| **V(e)_NLW** | 0.064 | 0.002 |
| **V(G)/Vp_L_LAT** | 0.073 | 0.013 |
| **V(G)/Vp_L_NLW** | 0.223 | 0.021 |
| **rG** | 0.714 | 0.130 |
| **n** | 22710 | |

LAT, Latino American; NLW, Non-Latino white; V(G), genetic variance; C(G), genetic covariance; V(e), residual variance;V(G)/Vp_L, proportion of variance explained by all common SNPs, transformed to underlying scale based on proportion of cases and disease prevalene; rG, genetic correlation; n, sample size

**Table S13.** Heritability estimates using two strata of SNP MAF (0.01-0.05, > 0.05), each stratified into 4 quartiles of LD score with REML or PCGC

| | REML | | PCGC regression | |
|---|---|---|---|---|
| | **Heritability** | **SE** | **Heritability** | **SE** |
| **group1_rare** | 0.075 | 0.035 | 0.079 | 0.052 |
| **group1_common** | 0.092 | 0.026 | 0.094 | 0.039 |
| **group2_rare** | 0.043 | 0.015 | 0.022 | 0.020 |
| **group2_common** | 0.007 | 0.023 | 0.012 | 0.035 |
| **group3_rare** | 0.007 | 0.008 | 0.004 | 0.012 |
| **group3_common** | 0.046 | 0.018 | 0.053 | 0.026 |
| **group4_rare** | 0.010 | 0.005 | 0.006 | 0.006 |
| **group4_common** | 0.017 | 0.010 | 0.014 | 0.013 |
| **sum** | 0.298 | 0.043 | 0.285 | 0.066 |

## References

1. Vijayakrishnan J, Qian M, Studd JB, Yang W, Kinnersley B, Law PJ, et al. Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk. Nat Commun. 2019 Nov 25;10(1):5348.

2. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, Sheridan E, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet. 2009 Sep;41(9):1006–10.

3. Wiemels JL, Walsh KM, de Smith AJ, Metayer C, Gonseth S, Hansen HM, et al. GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. Nat Commun. 2018 Dec;9(1):286.

4. Vijayakrishnan J, Henrion M, Moorman AV, Fiege B, Kumar R, Inacio da Silva Filho M, et al. The 9p21.3 risk of childhood acute lymphoblastic leukaemia is explained by a rare high-impact variant in CDKN2A. Sci Rep. 2015 Dec;5(1):15065.

5. Vijayakrishnan J, Qian M, Studd JB, Yang W, Kinnersley B, Law PJ, et al. Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk. Nat Commun. 2019 Dec;10(1):5348.

6. Perez-Andreu V, Roberts KG, Harvey RC, Yang W, Cheng C, Pei D, et al. Inherited GATA3 variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. Nat Genet. 2013 Dec;45(12):1494–8.

7. Xu H, Yang W, Perez-Andreu V, Devidas M, Fan Y, Cheng C, et al. Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. J Natl Cancer Inst. 2013 May 15;105(10):733–42.

8. de Smith AJ, Walsh KM, Francis SS, Zhang C, Hansen HM, Smirnov I, et al. BMI1enhancer polymorphism underlies chromosome 10p12.31 association with childhood acute lymphoblastic leukemia: *BMI* 1 enhancer polymorphism in ALL. Int J Cancer. 2018 Dec 1;143(11):2647–58.

9. Treviño LR, Yang W, French D, Hunger SP, Carroll WL, Devidas M, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet. 2009 Sep;41(9):1001–5.

10. Vijayakrishnan J, Kumar R, Henrion MYR, Moorman AV, Rachakonda PS, Hosen I, et al. A genome-wide association study identifies risk loci for childhood acute lymphoblastic leukemia at 10q26.13 and 12q23.1. Leukemia. 2017 Mar;31(3):573–9.

11. de Smith AJ, Walsh KM, Morimoto LM, Francis SS, Hansen HM, Jeon S, et al. Heritable variation at the chromosome 21 gene ERG is associated with acute lymphoblastic leukemia risk in children with and without Down syndrome. Leukemia. 2019 Nov;33(11):2746–51.